

Mobile Edge Computing for 5G Internet of Things

Haojun Huang, Wang Miao, Geyong Min and Chunbo Luo

November 26, 2018

Contents

1	Mobile Edge Computing for 5G Internet of Things	1
1.1	Introduction	1
1.1.1	5G Networks	1
1.1.2	Mobile Edge Computing	2
1.1.3	Mobile Edge Computing in 5G	3
1.2	Convergent Network Architecture of 5G and MEC	4
1.2.1	5G Network Architecture	4
1.2.2	Mobile Edge Computing Architecture	6
1.2.3	Integrated 5G Mobile Edge Computing Architecture	7
1.3	Technical Considerations in 5G Internet of Things	8
1.3.1	D2D Communications	9
1.3.2	Cross-layer Communications	10
1.3.3	Computation Offloading	11
1.3.4	Personalized Network Service	13

1.4	Emerging Application Scenarios	14
1.4.1	Location Service	14
1.4.2	IoT Gateway	15
1.4.3	Internet of Vehicles	16
1.5	Existing Challenges and Future Directions	17
1.5.1	Architecture Innovation	17
1.5.2	Computation Optimization	18
1.5.3	Management and Orchestration	18
1.5.4	Resource Allocation	19
1.5.5	Security and Privacy	20
1.6	Conclusion	20

Chapter 1

Mobile Edge Computing for 5G Internet of Things

1.1 Introduction

1.1.1 5G Networks

5G mobile networks, characterized by ultrahigh transmission throughput, ultrahigh connection density, and ultrahigh mobility, have been proposed as the next telecommunication standard beyond the current 4G/IMT Advanced standards. It will revolutionize future ubiquitous and pervasive networking, and provide high Quality-of-Experience (QoE) services to billions of users globally. 5G networks are designed with the aim of satisfying the performance requirements of three kinds of applications, services and business trends,

- Firstly, 5G technologies are expected to provide high-speed wireless communication services to mobile users. According to ITU MIT-2020 specification [1], 5G wireless communication networks should provide up to 20 Gbps to support various broadband-hungry applications, including high-definition videos, virtual reality, and 3D online

gaming [2].

- Secondly, 5G technologies are required to offer ultra-reliable and ultra-low latency services to support the applications in critical communication scenarios. For instance, 5G has been regarded as a key pillar technology for realizing the Industry 4.0 by German government [3] and autonomous robots and driving by Japan Government [4].
- Thirdly, 5G technology needs to support huge amount of network connections raised by Internet-of-Things (IoT)-related applications. According to Statista Report, it forecasts that the number of devices connecting to the network will reach up to 75 billion by the end of 2025 [5]. Therefore, 5G should be designed, deployed and optimized to meet the large-scale connections of IoT devices in the near future.

These three aspects of performance requirements are closely interacting with each other. For instance, the successful deployment of autonomous driving requires the 5G network to provide large amount of broadband, low transmission latency and huge amount of communication connections simultaneously. In this chapter, our focus is on the scenario of 5G networks to support IoT application and services. With the increasing deployment of IoT devices and applications, a large amount of the data will be generated during the service provisioning. For system optimization and service quality guarantee, the data generated is needed to be sent to remote cloud-based datacenter for processing, making latency-sensitive IoT application suffering from performance degradation. To address this issue, cloud resources are migrated to the edge of the network to support local data processing. With this background, this chapter presents the architecture, challenges and key technologies of complementarily integrating mobile edge computing with 5G enabled IoT services. Future work is discussed in order to guide the researchers with the similar research interests to make contributions in this area.

1.1.2 Mobile Edge Computing

Mobile Edge Computing(MEC) originated from a computing service platform, is a new network architecture to reduce the distance between the end users and computing resources,

enabling unprecedented benefits for data processing, service provisioning and resource optimization. MEC has been widely adopted by industry companies, e.g. IBM and Nokia Siemens Networks, to provide services to mobile users [9]. The basic idea is to migrate the cloud computing resources from the remote datacenters to the edge of the mobile access network to improve the utilization of computing and storage resources.

Bringing computation resources closer to mobile users is characterized by the close range, low latency and location awareness. This realizes the localization of mobile service provisioning, improves their business capabilities, reduces latency and optimizes network performance. MEC has very broad applications in today's mobile communications, e.g. Computational Offloading, Content Delivery, Mobile Big Data Analytics, Edge Video Caching, Collaborative Computing, Connected Cars, Healthcare, Smartgrids, etc.

1.1.3 Mobile Edge Computing in 5G

In the future 5G application scenario, 5G networks are required to quickly and efficiently respond to complex and variable services by leveraging cloud-based resource pools, fully meeting their performance requirements [11]. In this context, flexibility is a key feature for 5G networks to expand network resources and improve service provisioning. However, it is a challenging task to exploit the existing mobile network architecture to provide the flexibility required by 5G scenario. MEC has been considered as a promising technology for 5G network to revolutionize the manner of how resources are managed, provided and optimized [9] [10] [12]. MEC enables the mobile operators to transform their operation modes, reduce the Operational Expenses (OPEX) and enhance revenue capability. Compared with the traditional centralized data center, MEC is deployed at the edge of 5G networks. The resources in MEC could be dynamically changed and optimized based on the services and applications associated. Furthermore, the computation resources of mobile devices could also be considered in the MEC architecture to conducted the related tasks, entirely improving the system performance. The convergence of 5G network and MEC creates a new business ecosystem for network operators, service providers and end users.

1.2 Convergent Network Architecture of 5G and MEC

Following the introduction of 5G and MEC, the focus of this subsection is to investigate the integration of 5G and MEC to support IoT applications and services. To obtain this architecture, the existing 5G and MEC architectures are comprehensively analyzed and complementarily merged. The new network architecture should satisfy the system performance requirements of IoT applications and services. Therefore, we will present the basic architecture of 5G and MEC. Then the complementary aspects of two architectures will be investigated with the aim of proposing a new architecture, named MEC-enabled 5G architecture.

1.2.1 5G Network Architecture

5G networks mainly consist of three functional planes, the access plane, the control plane, and the forwarding plane respectively.

- The access plane is mainly responsible for utilizing multi-site collaboration, multi-connection mechanism and multi-system convergence technology to construct a flexible access network.
- The control plane is built based on reconfigurable network functions, providing on-demand access, mobility and session management, refined resource management and comprehensive capability optimization.
- The forwarding plane is part of a distributed data forwarding, responsible for managing dynamic anchor settings and providing packet processing and transmission.

Within 5G architecture, a modular-based pattern is used by the control plane to design and build the dedicated logical network for specific applications, e.g. remote machine control. Various network functions are selected and chained to meet the application needs. The control plane is the core component for the 5G network architecture to design network service, schedule the underlying resources, and utilize the access plane and forwarding plane to provide the end-to-end services. 5G network architecture consists of three layers as shown

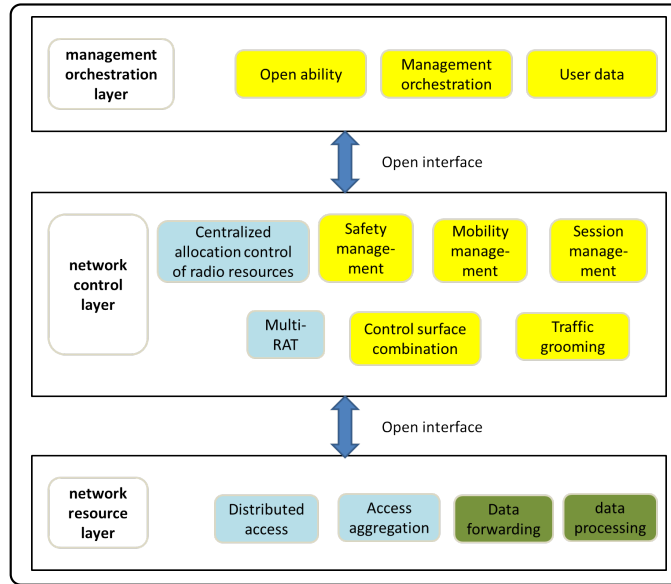


Figure 1.1: 5G network function view

in Fig. 1.1, including the management and orchestration layer, the network control layer and the network resource layer.

In order to further improve the service flexibility, 5G adopts a three level network architecture, named Distributed Unit (DU), Centralized Unit (CU) and core networks (5GC). For 5G networks, BaseBand Unit (BBU) function will be reconstructed into two functional entities, DU and CU, which together form a Next generation NodeB (gNB). Within a gNB, one CU can be connected to one or more DUs. Different function splitting schemes between the CU and the DU are used in 5G to adapt to different communication scenarios and communication requirements [8]. CU devices are responsible of managing the edge application deployment, providing high-level protocol stack functions, and supporting partial core network function, while DU devices mainly deal with physical layer functions.

5G is not a brand new network architecture, but evolves from 4G networks. Fig. 1.2 shows how to map the network functions from 4G to 5G during the network evolution.

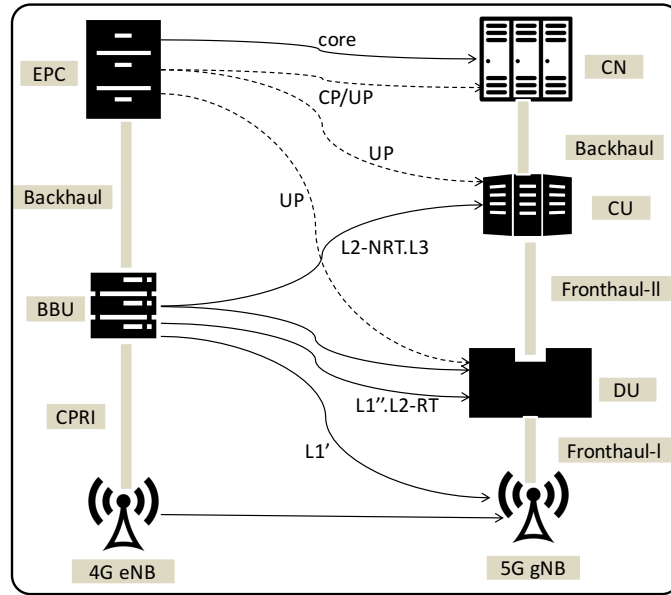


Figure 1.2: The evolution from 4G to 5G

1.2.2 Mobile Edge Computing Architecture

The architecture of the MEC is shown in the Fig. 1.3, which is a three-level network architecture, including the mobile edge system level structure, the mobile edge host level structure and MEC User Equipments (UE) respectively.

The system level structure is composed of MEC system level management and the third-party users [10]. The system-level management is mainly responsible for policy design, upper level service analysis and guiding mobile edge host to provide MEC services. The host-level structure includes host-level management software and host infrastructure. MEC host consists of MEC platform, MEC application and virtualization infrastructure. Host-level management software is in charge of the management and operation of mobile edge platform and virtualized infrastructure, including MEC platform manager and virtualization infrastructure manager. MEC platform and MEC applications share and exchange the information they stored during the whole service provisioning, such as the information of device discovery, network topology, network load and so on. MEC platform has the detailed information of the end UEs, e.g. location, bandwidth, network traffic etc., and provides the MEC application the operating environment during the service provisioning. In order to

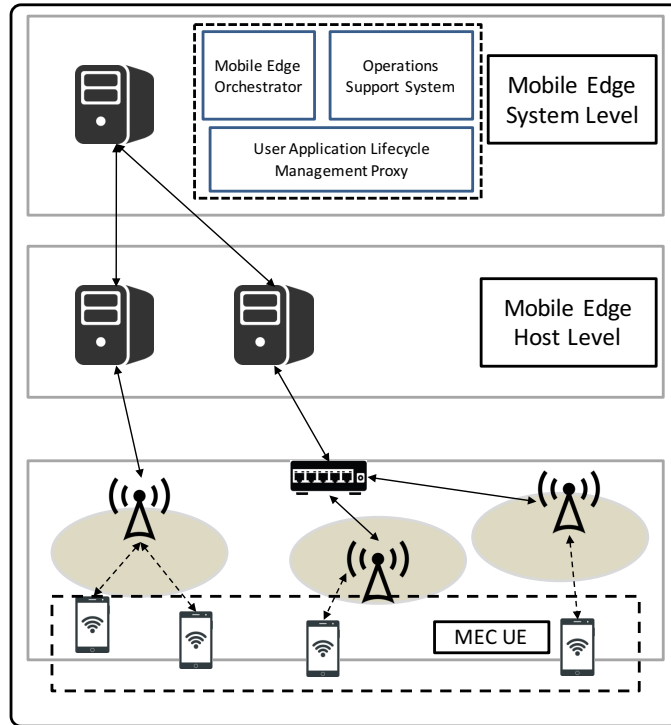


Figure 1.3: MEC Network Architecture

create an open ecosystem, the business rules of MEC services and Domain Name resolution System (DNS) proxy/server of MEC platform manager and application managers are also included in the system-level and host-level structures of the MEC architecture [11].

1.2.3 Integrated 5G Mobile Edge Computing Architecture

After introducing the network architecture, practical deployment and detailed components of 5G and MEC technologies, this subsection aims to combine these two architectures to improve the service provisioning capabilities of 5G networks for IoT applications. As we all know that the 5G networks are proposed to provide huge amount of network broadband and support large scale of device connection. In 5G networks, the wide network broadband is achieved by exploiting the millimeter communication, Multiple-Input-Multiple-Output (MIMO) and small cells technologies. For supporting the large scale of network connections, 5G uses the tremendous computation and storage resources from remote datacenter, and utilizes network function virtualization and software defined networking technologies to

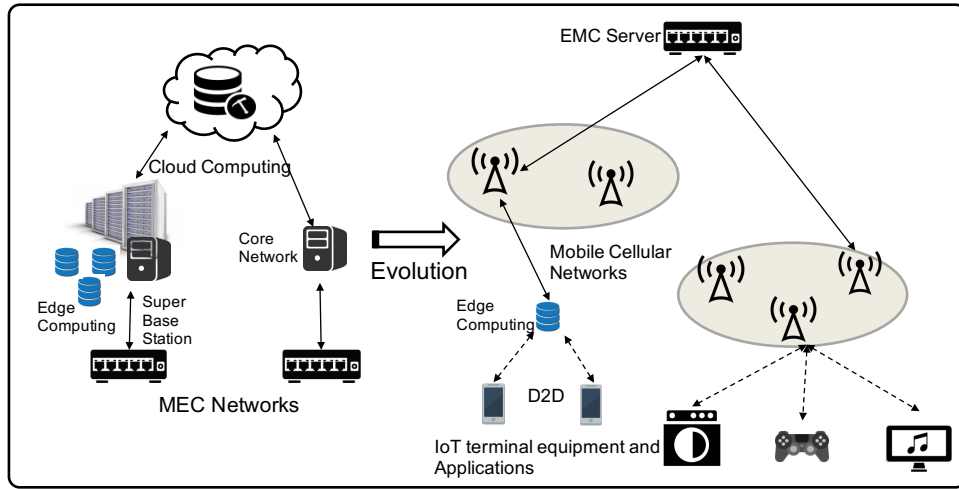


Figure 1.4: A 5G IoT infrastructure that combines MEC

virtualize the network resources for achieving an end-to-end optimized system for service provisioning. However, one issue that 5G network suffers from is the high latency, which could not meet the requirements of the emerging IoT applications [13] [14]. For solving this issue, MEC can be deployed in 5G gNB to eliminate the latency in the core network transmission, enhancing the service provisioning capability of 5G network for small-scale and ultra-low-latency services and application scenarios. As shown in Fig. 1.4, the future 5G mobile communication network will be a heterogeneous communication network that includes both the centralized Base Station (BS) and multiple distributed BSs [12]. For integrating the MEC and 5G networks, Fig. 1.4 shows a multi-level computing network that provides edge computing and cloud computing functions. Within this architecture, MEC computing resources are allocated in LTE eNB, 5G gNB, super 5G BS, the edge of core networks to provide the computing and storage resources for end-users. The new network architecture brings 5G technology more computation and storage resources, significantly improving the system performance and service provisioning capabilities.

1.3 Technical Considerations in 5G Internet of Things

For building a high-performance IoT system, the idea of edge computing could be used to enhance the efficiency for data transmission and processing, coupled with the upper-layer

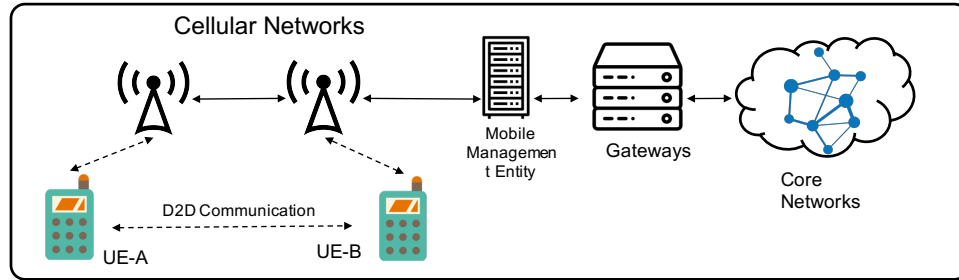


Figure 1.5: D2D Communications

network architecture. For practically supporting IoT applications, different applications have different requirements for network transmission, e.g. bandwidth, latency, jitter and so on. For instance, the automatic driving has been considered as a key application for IoT technology. It is unrealistic to exploit the remote cloud resources to process the requests from fast-moving cars. Once the remote cloud hardly responds the request in time, the problems, e.g. fatal accidents, may occur on the road, resulting in serious economic and society losses. Future researches are required to investigate new data processing technologies to address these technical problems. With this background, this part introduces D2D communications, cross-layer communications, computation offloading and customized network services in 5G IoT scenario.

1.3.1 D2D Communications

Cellular network based D2D communication, as shown in Fig. 1.5 enables that user data can be transmitted directly between terminals without being transited through the network [13]. As shown in the following figure, the working mechanism of D2D communication is significantly different from that traditional communication system.

Similar to the Machine to Machine(M2M) concept in the Internet of Things, the aim of D2D communication is to reduce the load of BSs through creating direct communications among UEs over a range of distances. Prior to the advent of D2D technology, similar communication technologies have emerged, such as Bluetooth (short-range time division duplex communication), Wi-Fi Direct (faster transmission speed and longer transmission distance) and Qualcomm's Flash LinQ technology (greatly improved the transmission distance of Wi-

Fi). Among these three technologies, Bluetooth has been widely deployed and supported by various wireless communication devices, e.g. mobile phone, laptop and IoT devices.

The main difference between D2D communication and ISM band-based short-range communication technology such as Bluetooth and wireless local area network (WLAN) is that it uses the licensed frequency band for information transmission. Therefore the interference environment is controllable, and the data transmission can be guaranteed. For Bluetooth, it requires the end users to manually establish the communications. The WLAN needs to perform user-defined settings for the access point (AP) before communications are established. D2D communication does not require the above processes, providing better Quality of Experiences (QoE). In addition, D2D could support a large number of information interactions, provides higher transmission data rate and enable better QoS guarantee for service provisioning.

1.3.2 Cross-layer Communications

At the edge of the MEC-enabled 5G IoT system, a large amount of IoT sensors are deployed and linked into the system, which are built into mobile phones, televisions, washing machines, electric lights and so on. These devices are heterogeneous, and intelligent IoT applications require these devices to communicate with each others. Generally speaking, these devices are connected to the network with different protocols including WiFi and Bluetooth. In order to realize the communications among different devices, the network gateway is used to enable the communication between two networks. A gateway is a computer system or device that acts as a conversion. The gateway is a translator between two systems that use different communication protocols, data formats, and network architectures. As the scale of the IoT applications continues to increase, the number of gateways that need to be purchased and deployed will increase dramatically, resulting in an excessive cost of IoT deployment and application.

Considering that WiFi, Bluetooth and Zigbee all work at the band of 2.4 GHz, it is reasonable to incorporate these three technologies under a unified communication frame as

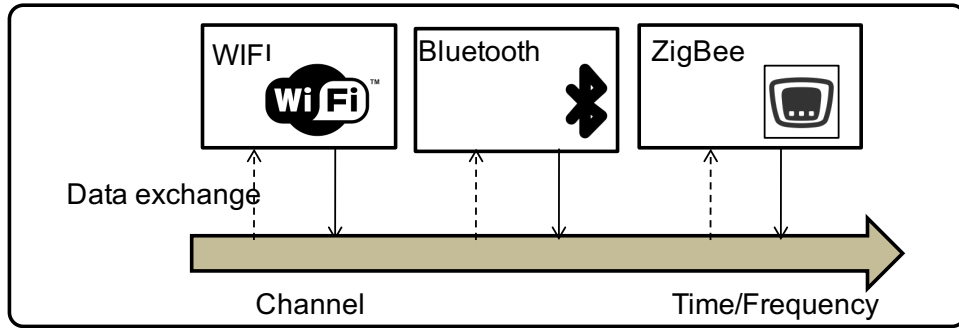


Figure 1.6: Cross-technology coordination

shown in Fig. 1.6. For instance, to realize the interconnection among different devices with different protocols, it is possible to transmit the data signals generated by these three devices into a single-sided proprietary channel. This design provides the benefits of enabling the direct communication between heterogeneous senders and receivers, allowing heterogeneous devices to receive broadcast simultaneously from a sender with overlap ping frequencies (e.g., Bluetooth to WiFi and ZigBee) and supporting a sender with a wider bandwidth (e.g., WiFi) to reach multiple narrower-band receivers (e.g., WiFi to multi-channel ZigBee) [15]. In addition, this solution does not impose any modifications on the hardware devices in use and create any traffic within the network and the whole design is transparent to the upper level wireless communication systems. As shown in Fig. 1.6, this design solves the problem of cross-technology coordination, enabling heterogeneous devices to be easily interconnected.

1.3.3 Computation Offloading

The terminals of the IoT applications are deployed in the form of sensors. These devices can be homogeneous or heterogeneous. For fulfilling various task of IoT applications, devices needs to conduct various computation processing and access the network to upload the data and download the messages. However, with the explosive growth of the mobile devices and IoT applications, the computation, storages and network resource are approaching their limits, and hardly meet the performance requirements for IoT applications running on these devices [16].

It is worth noting that the limitations of mobile terminal resources are related to the mo-

bility feature. Compared with the static devices, e.g. laptop, the mobile devices are designed with the lower computation, the fewer storage and narrower broadband to meet the requirement of portability and mobility. In the design of the mobile terminals, in order to satisfy the requirements of mobility and portability, the processing capability of the terminal device, the network connection and the like are transferred [17]. In this case, computing offloading is an effective method to extend the resources for mobile devices. During computation offloading, the intensive computing tasks of IoT application are sent to remote high-performance datacenter or nearby mobile devices to be executed, and the execution results are downloaded back to the mobile devices for further processing, significantly enhancing the computation capability of mobile devices.

Nowadays, after 20 years of computation evolution, the computation offloading technologies could be classified three stages as shown in Table 1.1. In the evolution of computing migration, there are some computing migration systems, such as MAUI, Cloudlet, and CloneCloud, emerged to implement the idea of computation offloading. Designed with different research objectives, these systems have their own advantages and disadvantages and were developed based on the architectures of surrogate, cloudlet-based and cloud-based respectively. The authors in [18] demonstrated the detailed implementations for these three systems for computation offloading.

Although there have been some interesting research results reported in the literature, there are still several major problems in computing migration listed as follows. Firstly, there is a lack of a unified resource organization and management strategy to realize computation offloading [19] [20]. Secondly, there is a lack of incentives and reputation systems necessary for mobile devices to cooperate with each other. And finally, the research on the stability, security, data consistency and privacy of computation offloading have not been fully investigated. In order to realize a successful deployment of computation offloading in the practical network environment, more research efforts are expected to address the aforementioned problems.

Table 1.1: The evolution of computation offloading

Stages	Time	Research Mobile Terminal	Research Application	Research Content	Implementation
Distributed Comput- ing	1995-2000	Laptop perfor- mance/Energy consumption	Computationally intensive application	Feasibility	Manual division
Pervasive Comput- ing	2001-2008	Performance of laptops and Smartphone/ Energy consumption	Resource intensive application	System ar- chitecture and imple- mentation	Base on surrogate
Cloud Comput- ing	2009	Mobile device enhancement	Rich mobile app, augmented reality	Cloud- based computing migration system	Base on cloudnet/cloud/ Mobile devices

1.3.4 Personalized Network Service

Different IoT applications have different requirements that the network should satisfy. Therefore, the network should be managed and configured to cater the upper level applications, making personalized network services an important aspect in network evolution. For instance, 4G networks were mainly designed and deployed to provide the services for mobile phones and the network architecture and protocols are optimized for mobile devices and services. However, in the 5G era, the devices connected to the mobile networks are significantly increased and more importantly, the types of the devices, the performance requirements of applications and patterns of service provisioning are also changed a lot. This makes it important for the network to conduct data processing, understand the quality of service provisioning and optimize the network resource to satisfy the transmission requirements.

In the MEC-enabled 5G IoT scenario, the sensors connected to the edge of the Internet of Things generate data days and nights. These data is transmitted from the end device, through the network links and devices and reach the intermediate nodes. The intermediate nodes determine whether the data should be sent to the remote cloud node for processing or can be directly processed in the edge computing nodes. Therefore, for MEC-enabled 5G IoT system, the function of filtering the data should be designed for saving bandwidth and

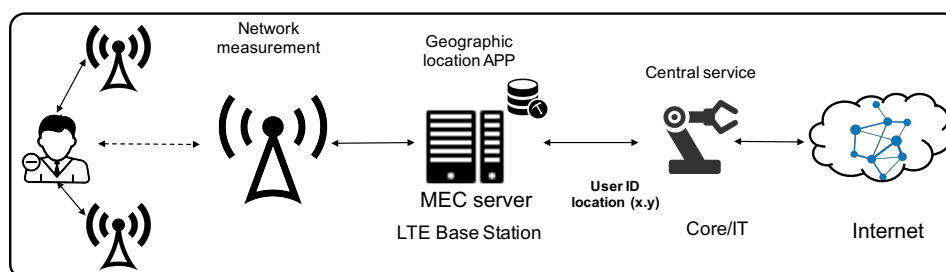


Figure 1.7: Location service

reducing the processing latency. If the filtering function is realized in the mobile devices, status of the available resources, processing time and network bandwidth should be transmitted from edge MEC to mobile devices, introducing extra transmission latency. Therefore, the filtering function should be located in MEC nodes to achieve the global optimization for IoT devices.

1.4 Emerging Application Scenarios

1.4.1 Location Service

At present, the location service is mainly provided by GPS satellite system, which provides the accurate navigation information for outdoor navigation and has been widely used in daily life, Unmanned Aerial Vehicle (UAV), Google maps, autonomous driving and so on. However, as the population increases and human living habits change, the demand for indoor positioning is increasing. Its accuracy could not be satisfied by the GPS system. The commonly used indoor navigation technologies are generally based on Wi-Fi and Bluetooth. While these technologies have not been practically deployed due to the issue of accuracy and reliability. In this area, MEC system has been proposed a promising technology to estimate the location of end devices.

The advantage of MEC system is that its geographical location is close to the BSs and end devices. The whole process is shown in Fig. 1.7. The network measurement data of BSs and user data can be accessed and read with relatively small cost. Based on the

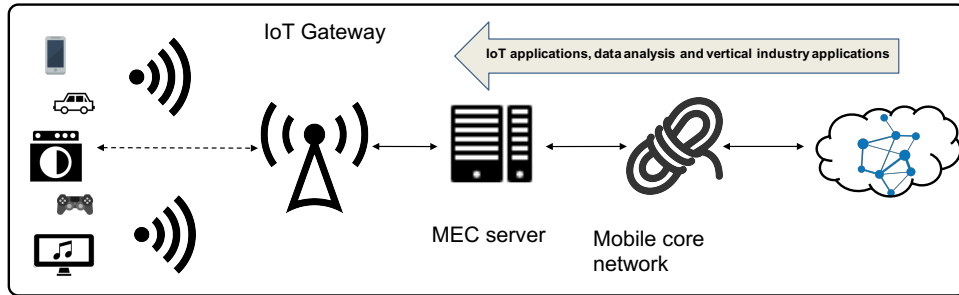


Figure 1.8: IoT Gateway

data received, the algorithms in MEC stations could conduct comprehensive analysis of measurement data to accurately estimate the position of end devices, providing location services for IoT applications.

1.4.2 IoT Gateway

Constraint to the practical environment, IoT devices located at the end of the network are allocated with limited resources in computing and storage and can usually run simple communication protocols. To support resource-hungry IoT applications, IoT gateways are deployed to perform network protocol conversion. IoT Gateway is a core equipment for the wireless sensor network to deploy IoT applications. It is responsible for building the links between a connection-aware network and a traditional communication network.

With the increasing popularity of IoT devices, various IoT applications, especially those that are sensitive to latency, are beginning to emerge. The MEC system provides conditions for supporting such IOT applications in complex network environment as shown in Fig. 1.8. By deploying the IoT gateway in the MEC-enabled 5G system, the processing power of the edge MEC servers could be used by IoT gateway to support various IoT protocols with the lower latency and the higher agility, significantly enhancing the capability of service provisioning in IoT networks.

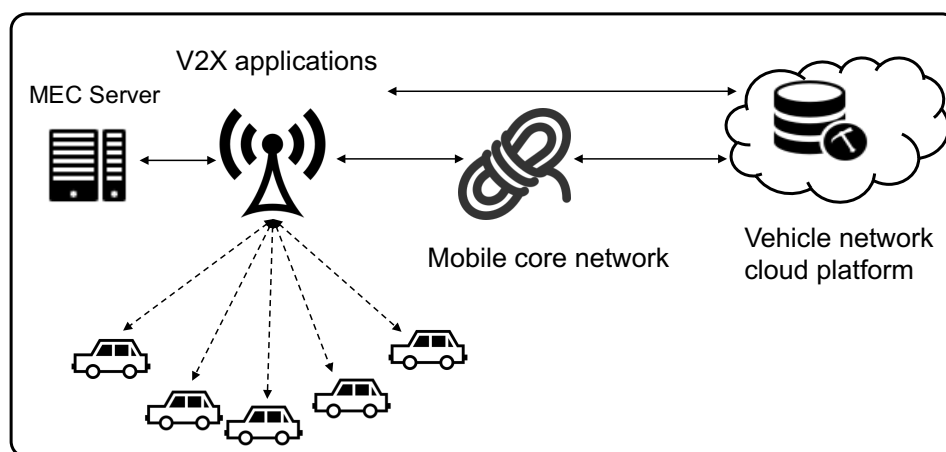


Figure 1.9: Internet of Vehicles based in MEC

1.4.3 Internet of Vehicles

The increasing number of cars connected has promoted the development of the Internet of Vehicles, e.g. various vehicle-related applications are appearing, automatic driving, smart road construction, intelligent transportation, and so on. These emerging applications in internet of vehicles pose challenging issues in network design, e.g. requiring low latency transmission, large scale of data processing, huge amount of IoT devices connected. By migrating the remote cloud resources to nearby and distributed mobile BSs, MEC brings network operators a promising solution to deal with the strict performance requirements of the vehicles-related applications. The data could be sorted and processed in the nearby MEC server and key decisions could be made based on the analysis results, which enables higher processing capability and lower transmission latency as shown in Fig. 1.9.

Other vehicle networking applications that the MEC system can support include road fault notification, traffic congestion prompts, optimal path selection, and location tracking. These applications will greatly improve the safety, efficiency and convenience of the transportation system. Therefore, the combination of Internet of Vehicles and MEC is a complementary solution and will have significant impact for smart vehicle applications.

1.5 Exiting Challenges and Future Directions

The MEC-enabled 5G network is a multi-level computing and communication architecture. Compared with the traditional mobile communication networks, which purely focus on data transmission, MEC-enabled 5G networks cooperatively combine the data transmission, data storage and data processing in a single network architecture, greatly reducing the end-to-end transmission latency, which is a key challenge that 5G communication network struggles to handle. Although MEC-enabled 5G network brings various benefits for network design and has attracted tremendous research interests, there are still some challenges that need further research investigation in order to realize the ambitions that MEC-enabled 5G networks promise. This subsection will briefly introduce the main challenges of MEC-enabled 5G networks, including the architecture innovation, the practical deployment and security and privacy-related issues.

1.5.1 Architecture Innovation

As discussed above, MEC-enabled 5G network will be an intelligent information system integrating computing, storage and networking resources. However, the theory for practical deploying this network architecture is not fully investigated. For example, how to globally organize, manage and optimize the network, storage and computing resources is a key issue to address. A unified management architecture is needed to allocate the system resources according to the upper level application requirements. In addition, how to define and model network utilization capacity, how to convert network utility capacity and computing power, and how to use virtualization and network technology to achieve effective resource allocation, etc., also need more scientific work. These issues all place great demands to deepen and refine the design of the MEC-enabled 5G architecture. For instance, how to improve its infrastructure so that it can balance the communication and computation is also an important research issue for MEC-enabled 5G architecture.

1.5.2 Computation Optimization

For the MEC-enabled 5G networks, the collaborative optimization of communication, storage and computing resources is one of the key issues. The IoT applications running on the UE devices have specific QoS requirements e.g. end-to-end latency. On the other hand, fulfilling the function of IoT applications requires the storage and computation resources that UE devices can not meet due to the limitation of the size, cost, and energy-consumptions. The remote cloud computing is an option for UE devices to support IoT applications. While the long distance transmission results in large transmission latency, which can't meet the QoS requirements. It is necessary to shorten the distance of the cloud resource and UE devices, or leverage the other affordable computing nodes nearby to conduct the calculation tasks. By integrating MEC to mobile communication networks, 5G networks have multi-level computing, storage and networking capabilities. However, there are still some research challenges for addressing the limited computing resources in UE nodes. For instance, how to assign the computing tasks to multiple nodes with different computing powers. This is a well-known problem, computing task unloading, and needs to be solved. Compared with the traditional offloading technologies of computing task, task offloading in MEC-enabled 5G network should take not only the computation resources, but also the network transmission and storage capacity into account. For instance, radio resources such as bandwidth and transmission power are strictly limited in wireless communication. Task offloading should consider the quality of channel transmission and storage capability in distributed servers and conduct the task division and allocation among multiple nodes for optimizing the overall system performance.

1.5.3 Management and Orchestration

For a reliable system, a systematic service management and orchestrator is responsible for managing the network topology, optimizing the network resources and providing end-to-end network services [22]. In the MEC-enabled 5G architecture, the diversity of the edge resources, the IoT devices and upper level applications highlights the importance of the service orchestration. Under the unified management of service orchestration, the components

could communicate with each other and send the tasks and data to edge nodes. In addition, the network operators can also deploy fault detection and access control and introduce new services in the networks with the guide of service orchestration.

It is worth mentioning that the diversity of data also increases the difficulty for service management. For IoT applications, a variety of IoT devices will be deployed and connected to MEC-enabled 5G networks. These devices continuously generates heterogenous data, which pour into the network domain. For the service orchestrator, how to exploit the intelligent algorithms and network resources to handle the huge of heterogenous data, analyze valuable information and optimize the network resources is an urgent problem to be solved in the future.

1.5.4 Resource Allocation

In the MEC-enabled 5G IoT system, MEC is applied to the edge of the network to improve the performance of the system. However, MEC resources are distributively deployed and controlled by different management entities [23]. The examples of these management entities include the UE terminal equipments, BS control softwares, core network routers, and the edge servers. The resource management approaches in traditional cloud computing are mainly based on a logically centralized node, which makes it difficult to directly implement the resource management of cloud computing in MEC 5G networks. Therefore, new resource managements solutions are required to develop a global resource management system for MEC 5G system. Virtualization technologies have been considered a good technology in future management system to virtualize the underlying network infrastructure and improve the resource utilization. However, the network virtualization brings new challenges to management system, such as complexity and performance guarantee, which needs more research efforts in MEC-enabled 5G scenarios.

1.5.5 Security and Privacy

In MEC-enabled 5G networks, a large amount of IoT devices will be connected to the network infrastructure. For the security of the entire network and IoT applications running this infrastructure, the network management system should authenticate these devices before providing the network service [21]. However, many resource-constrained terminals do not have enough storage and computation resources to perform the encryption operations required by the authentication protocol. Although traditional public key infrastructure, e.g. PKI-based authentication can solve the problem of secure communication, it is not originally designed for IoT application and faces serious challenges security issues posed by the unique features of IoT system, for instance, large-scale of device connections. In addition, malicious computing nodes may pretend to be legitimate nodes to steal data generated by other IoT devices, and may also abuse user data or provide malicious data to neighboring nodes to disrupt their behavior. Due to the complexity of trust management, dealing with this issue in the IoT environment can become more complicated.

In this context, novel permission authentications are needed to enhance the security and privacy of IoT system and application. For IoT applications, different devices have different permissions for data transmission and sharing. New methods are expected to detect and analyze the data received from IoT devices. The management system could be enhanced to conduct the task of authenticating and authorizing the requests from the normal nodes and rejecting the packets from malicious nodes.

1.6 Conclusion

This chapter mainly introduced a new network architecture to support IoT applications, named MEC-enabled 5G networks. Firstly, we discussed the basic architecture, the main components, and the network functions of 5G and MEC. Based on the analysis of complementary aspects of 5G and MEC architectures, a novel network architecture was presented in this chapter to integrate MEC into 5G architecture to improve the service provisioning abil-

ity for IoT applications. We presented the advantages and benefits of the proposed network architecture for IoT applications. In addition, the technology considerations of providing IoT service in the MEC-enabled 5G networks were introduced, including D2D communications, cross-layer communications, computation offloading and customized network services. Following the system design considerations, we investigated three use cases of how to use MEC-enabled 5G network to support IoT applications. Finally, the challenges and future research directions were discussed with the aim of guiding and inspiring other researchers to make progressive contributions in this area.

References

- [1] M. Carugi, “Key features and requirements of 5G/IMT-2020 networks,” *Report*, 2018.
- [2] L. Tian, C. Yan, W. Li, Z. Yuan, and W. Cao, “On Uplink Non-Orthogonal Multiple Access for 5G: Opportunities and Challenges,” *China Communications*, 2017, 14(12): 142-152.
- [3] Y. Nakamura, “Japan’s 5G Policy Perspectives,” *Report*, 2016.
- [4] S. Ludwig, M. Karrenbauer and A. Fellan, “A 5G Architecture for The Factory of the Future,” *Networking and Internet Architecture*, 2018: 1-7.
- [5] Statista, “Internet of Things (IoT) connected devices installed base worldwide from 2015 to 2025 (in billions),” [Online]. Available: <https://www.statista.com/statistics/471264/iot-number-of-connected-devices-worldwide/>.
- [6] E. Bastug, M. Bennis, and M. Debbah, “Living on the Edge: The Role of Proactive Caching in 5G Wireless Networks,” *IEEE Communications Magazine*, 2014, 52(8): 82-89.
- [7] S. Li, L. Xu, and S. Zhao, “5G Internet of Things: A Survey,” *Journal of Industrial Information Integration*, 2018, 10: 1-9.
- [8] M. Palattella, M. Dohler, A. Grieco, G. Rizzo, and J. Torsner, “Internet of Things in the 5G Era: Enablers, Architecture, and Business Models,” *IEEE Journal on Selected Areas in Communications*, 2016, 34(3): 510-527.
- [9] W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu, “Edge Computing: Vision and Challenges,” *IEEE Internet of Things Journal*, 2016, 3(5): 637-646.

- [10] M. Satyanarayanan ,“ The Emergence of Edge Computing ,” *Computer* , 2017, 50(1): 30-39.
- [11] H. Pang, and K. Tan, “Authenticating Query Results in Edge Computing, ” *International Conference on Data Engineering*, 2004, 20: 560-571.
- [12] Y. Mao, C. You, J. Zhang, K. Huang, and K. Letaief, “A Survey on Mobile Edge Computing: The Communication Perspective, ” *IEEE Consumer Electronics Magazine*, 2016, 5(4): 732322 - 2358.
- [13] M. Du, K. Wang, Y. Chen, X. Wang, and Y. Sun, “Big Data Privacy Preserving in Multi-Access Edge Computing for Heterogeneous Internet of Things, ” *IEEE Communications Magazine*, 2018, 56(8): 62-67.
- [14] D. He, S. Chan, and M. Guizani “Security in the Internet of Things Supported by Mobile Edge Computing, ” *IEEE Communications Magazine*, 2018, 56(8), 56-61.
- [15] M. Song and H. Tian, “FreeBee: Cross-technology Communication via Free Side-channel, ” *International Conference on Mobile Computing and Networking*, 2015: 317-330.
- [16] D. Duolikun, T. Enokidoy, and M. Takizawa, “A Process Migration Approach to Energy-Efficient Computation in a Cluster of Servers, ” *International Conference on Broadband and Wireless Computing*, 2016: 191-198.
- [17] R. Zhang, D. Zhang, and Z. Kan, “Using Mobile Agent Techniques for Distributed Manufacturing Network Management, ” *International Conferences on Info-tech and Info-net*, 2001.
- [18] A. Milanes, N. Rodriguez, and B. Schulze ,“A Classification for the Implementations of Heterogeneous Strong Migration of Computations,” *IEEE International Symposium on Cluster Computing and the Grid*, 2007: 858-868.
- [19] C. Gao and V. C. S. Lee, “Energy Efficient Mobile Computation Offloading through Workload Migration, ” *IEEE International Conference on Smart City*, 2015: 1147-1150.
- [20] Milans, N. Rodriguez, and R. Leruslimschy, “Reflection-based Heterogeneous Migration of Computations, ” *Brazilian Symposium on Computer Networks and Distributed Systems*, 2014: 223-230.
- [21] W. Zhou, Y. Jia, A. Peng, Y. Zhang, and P. Liu, “The Effect of IoT New Features on Security and Privacy: New Threats, Existing Solutions, and Challenges Yet to Be

- Solved, ” *IEEE Internet of Things Journal*, 2018, Early Access.
- [22] P. Shantharama, A. Thyagaturu, N. Karakoc, L. Ferrari, M. Reisslein and A. Scaglione, “LayBack: SDN Management of Multi-Access Edge Computing (MEC) for Network Access Services and Radio Resource Sharing, ” *IEEE Access*, 2018, 6: 57545-57561.
- [23] A. Kiani, and N. AnsariA, “Edge Computing Aware NOMA for 5G Networks, ” *IEEE Internet of Things Journal*, 2018, 5(2): 1299-1306.