

# Network Slicing for 5G: Challenges and Opportunities

Xin Li,<sup>1</sup> Mohammed Samaka,<sup>2</sup> H. Anthony Chan,<sup>3</sup> Fellow, IEEE, Deval Bhamare,<sup>2</sup> Lav Gupta,<sup>4</sup> Senior member, IEEE, Chengcheng Guo<sup>1</sup>, and Raj Jain,<sup>4</sup> Fellow, IEEE

<sup>1</sup>*School of Electronic Information, Wuhan University, Wuhan, Hubei 430072, China*

<sup>2</sup>*Department of Computer Science and Engineering, Qatar University, Doha, Qatar*

<sup>3</sup>*Huawei Technologies, Plano, TX, USA*

<sup>4</sup>*Department of Computer Science and Engineering, Washington University in St. Louis, St. Louis, MO 63130, USA*

**Abstract** — *Network slicing for 5G is receiving significant attention from the telecommunications industry as a means to provide network as a service (NaaS) for different use cases. Network slicing is a technology which allows network operators to build multiple virtual networks on a shared infrastructure. With network slicing, service providers can deploy their applications and services flexibly and quickly to accommodate specific requirements of diverse services such as augmented reality, online games, e-health and others. As an emerging technology with a number of advantages, network slicing has raised many issues for the industry and academia alike. In this article, we discuss the background and related work in network slicing and propose a framework for 5G network slicing. Finally, we discuss the challenges of network slicing and future research directions.*

**Keywords** — *5G, Core Network, Network Functions Virtualization, Network Slicing, Radio Access Network, Software Defined Networking, Vertical Services.*

## I. INTRODUCTION

Research in 5G mobile networks has gained momentum in both academia and industry recently. 5G mobile networks will carry a large plethora of mobile devices and provide faster network connection speed. In addition, mobile data traffic in future 5G networks will experience explosive growth. According to the latest Cisco forecast [1], by 2020, the number of devices connected to mobile networks will exceed the world's population reaching 11.6 billion. At that time, mobile data traffic will be 30.6 Exabytes per month, which is 8 times that in 2015. However, the most challenging aspect is that 5G networks will support a wide variety of use cases as we are entering the era of Internet of Everything (IoE). Some applications such as ultra-high definition (UHD) video and augmented reality need high-speed, high-capacity communications, while others like mission critical Internet of Things (IoT) and autonomous vehicles require ultra-low latency, ultra-reliable services.

The traditional mobile communication networks employ the one-size-fits-all approach to providing services to mobile devices, regardless of the communication requirements of vertical services. This design philosophy cannot offer differentiated services. Hence, it is necessary for the research community to explore new techniques to address the challenges associated with supporting vertical industries in 5G networks.

Software Defined Networking (SDN) and Network Functions Virtualization (NFV) have been proposed as key technologies to build software-defined, virtualized and cloudified 5G systems in recent years. SDN [2] decouples network control and data forwarding. Network control functions can run as applications independently in the logically centralized controllers. NFV [3] decouples specific network functions from dedicated and expensive hardware platforms to general-purpose commodity hardware. Network operators can implement a variety of virtual network functions over the standard commodity servers. In addition, Mobile Edge Computing (MEC) [4] as a key emerging technology in 5G is expected to serve low-latency communication that is one of the use cases in future 5G. It moves computing, storage and networking resources from remote public clouds closer to the edge of the network. Thus, mobile clients can request virtual resources within the access network and experience low end-to-end delay.

The concept of “network slicing” [5] has been proposed to address the diversified service requirements in 5G under the background of the aforementioned technologies. Network slice is an end-to-end logical network that is provisioned with a set of isolated virtual resources on the shared physical infrastructure. These logical networks are provided as different services to fulfil users’ varying communication requirements. Network slicing provides a network-as-a-service (NaaS) model, which can be very flexible to allocate and

eprint

IEEE Internet Computing, Vol. 21, Issue 5, August 2017, pp. 20-27

reallocate resources according to the dynamic demands such that it can customize network slices for diverse and complex 5G communication scenarios.

Network slicing will be the fundamental feature of 5G networks. Slice-based 5G has the following significant advantages when compared with the traditional networks: (1) Network slicing can provide logical networks with better performance than one-size-fits-all networks. (2) A network slice can scale up or down as service requirements and the number of users change. (3) Network slices can isolate the network resources of one service from the others. The configurations among various slices don't affect each other. Therefore, the reliability and security of each slice can be enhanced. (4) Finally, a network slice is customized according to service requirements, which can optimize the allocation and use of physical network resources.

The rest of the paper is organized as follows. We introduce the related work on network slicing in Section II. Section III discusses how to implement network slicing in 5G from the perspective of network architecture. A framework for 5G network slicing is presented in Section IV. Then, we discuss some challenges and issues for future network slicing research in Section V. Finally, we conclude this paper in Section VI.

## II. RELATED WORK

Network slicing makes use of the concept known as network virtualization. Network virtualization enables flexible and dynamic network management to address the problem of network ossification by allowing multiple heterogeneous and service-specific virtual networks to share single substrate network.

In addition, the emerging technologies SDN and NFV are considered as the necessary tools to implement network slicing. The following work summarizes use of SDN and NFV to implement 5G network slicing.

A flexible 5G network architecture is proposed in [6] to support the coexistence of heterogeneous services and to quickly create services. The authors propose to use SDN and NFV into the architecture to realize the sharing of resources by different services and orchestrating resources automatically. The concept of resource slicing is similar to network slicing in the proposed architecture. Resource slices, which include virtual resources and virtual network functions, are tailored on-demand according to the service categories, i.e., Slice-as-a-Service (SlaaS). A NFV-based platform called Multi Datacenter service ChaIN Emulator

(MeDICINE) has been designed for network services in [7]. It allows management and orchestration (MANO) system to deploy virtual network resources for network services in a multi-domain infrastructure. The design and implementation of this platform show that NFV plays an important role in realizing network slicing. The authors in [8] propose a novel slice-based 5G architecture based on SDN, NFV and cloud computing. They design the elements required to implement network slicing and present a validation prototype. The concept of "network store" in this work can achieve dynamic 5G network slicing. The authors discuss building a multi-tenant and multi-service end-to-end 5G mobile network architecture using SDN, NFV, and network slicing in 5G NORMA project [9].

Mobile operators, hardware manufacturers and open source communities are also actively studying the ways to implement 5G network slicing. Ericsson and NTT DOCOMO successfully showed a dynamic 5G network slicing proof of concept (PoC) on June 9, 2016 [10]. Ericsson in [11] discussed the decomposition schemes of radio access network (RAN) functions that can support network slicing. It provides a meaningful guidance for implementing flexible and resilient RAN slices. Very recently, Huawei with other three enterprises released a network slicing white paper [12]. They discuss several key technologies such as network management system and security to achieve service-guaranteed network slicing, which will be helpful for network industries. ONAP [13], established in February 2017, is working on a cloud-centric and SDN/NFV-based network platform, which may lay the groundwork for the implementation of 5G network slicing. In addition, it is worth mentioning that the Wireless World Research Forum (WWRF) is currently launching a working group for 5G network slicing.

In the next section, we discuss some details on network slicing for 5G networks and explore why 5G needs network slicing, what is network slicing and how to implement network slicing in 5G.

## III. NETWORK SLICING IN 5G

5G networks need to integrate multiple services with various performance requirements such as high throughput, low latency, high reliability, high mobility, and high security into a single physical network infrastructure and to provide each service with a customized logical network, i.e., network slicing. The 3<sup>rd</sup> Generation Partnership Project (3GPP) has identified network slicing as one of the key technologies to

achieve the aforementioned goals in future 5G networks. Some 3GPP work items have the features of network slicing, for example, the dedicated core (DÉCOR) feature supports the operator to deploy multiple dedicated core networks by sharing a common “Public Land Mobile Network (PLMN)”. An exhaustive description about how the next generation wireless system will support network slicing is given in [14].

Network slicing refers to partitioning of one physical network into multiple virtual networks, each architected and optimized for a specific application/service. Specifically speaking, a network slice is a virtual network that is created on top of a physical network in such a way that gives the illusion to the slice tenant of operating its own dedicated physical network. A network slice is a self-contained network with its own virtual resources, topology, traffic flow and provisioning rules. There may be various network slices to meet the specific communication needs of different users in the future mobile network systems. For example, a massive industrial IoT slice may need a light 5G core, no handover but a large number of connections. On the other hand, a mobile broadband slice may need a high capacity core, full feature mobility and low latency. Slices are logically isolated, but resources can be shared among them. Figure 1 illustrates the network slicing concept.

Network slicing needs to be implemented in an end-to-end manner to meet diverse service requirements. Each slice may have its own network architecture and protocols. 5G network slicing includes slicing 5G radio access network (RAN), 5G core network and even end-user devices. 5G RAN slicing can be implemented through logical abstraction of physical radio resources such as spectrum and physical hardware such as a base station. SDN and NFV can configure the virtual network resources flexibly, which include network bandwidth, server processing capability, network element processing capability, etc., to build the core network slices for specific service requirements. RAN slices and core network slices can either be dedicated to a class of service users or shared between multiple classes of service users. In order to provide an identical slice for a specific service, a common function "slice selecting function" is needed. It is responsible for selecting RAN and core network slices to form the end-to-end slice.

From the RAN point of view, the main task of RAN slice is to achieve the sharing and flexible management of spectrum resources. Therefore, the software-defined RAN is a promising approach to implement RAN slicing.

The sharing of radio resources between different slices can be accomplished by scheduling conducted by the controller in software-defined RAN. The controller can allocate radio resources to a slice according to service request characteristics such as traffic load. Moreover, different radio access methods for RAN slices are needed in terms of diverse service requirements. For example, besides the mobile broadband access in the traditional cellular networks, the critical communication slice in 5G needs low-latency radio access. In addition, the implementation of RAN slicing needs logical decompositions of RAN functions to figure out which functions are common functions and which functions are provided with specific hardware [11].

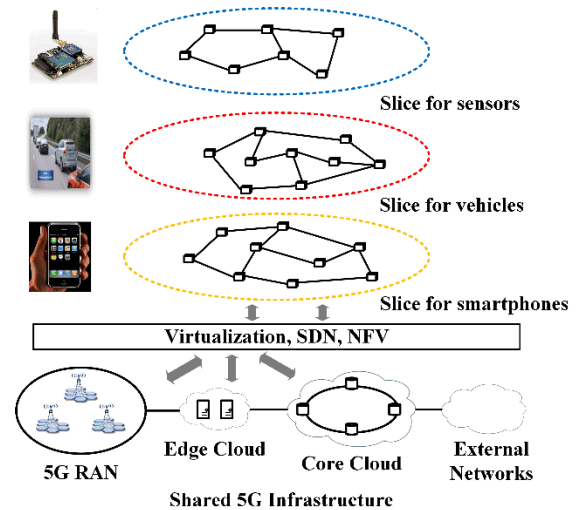


Fig. 1 Conceptual illustration of network slicing

5G core network slicing can provide tailored logical networks for different services or verticals in an agile and adaptable fashion. The key technologies of core network slicing are SDN and NFV. SDN technology can be used to separate the control plane and data plane of the core network so that the control plane and data plane can be deployed independently. The control plane can be centralized in order to facilitate management, while the data plane can be distributed. For example, the data plane in a network slice for low-latency services can be distributed on the network edge in combination with the MEC technology. NFV technology can provide the necessary virtual network functions for the data plane of the core network slices according to the service type. These virtual network functions can be scaled on-demand as the service changes dynamically.

5G network slicing brings flexibility while increasing the complexity of network management. Thus, there is a need to design automated management schemes to address such complexity. The lifecycle management is responsible for slice creation, reconfiguration, deletion, etc. It involves orchestrating and allocating infrastructure resources. Therefore, it is crucial to improve resource utilization of 5G infrastructure.

In the next section, we present our 5G network slicing framework based on our previous work [15].

#### IV. 5G NETWORK SLICING FRAMEWORK

In this section, we present our three-layer network slicing framework for 5G networks as shown in Figure 2. The proposed framework supports flexible deployment and management of diverse network applications over one common infrastructure. The framework contains 5G software-defined infrastructure (5G-SDI) layer, virtual resource layer, application and service layer and slicing MANO functional component. We use three different colors in the figure to indicate three network slices for connected vehicles, virtual reality and mobile broadband (MBB).

The bottom most layer in Figure 2 is the 5G infrastructure layer, which is composed of multi-domain 5G software-defined infrastructures with software-based control and management. The use of software-defined approach brings the following benefits to the proposed framework: (1) It provides the abstraction of resources and allows dynamic allocation of resources. (2) Virtualization of infrastructure resources can flexibly form service-specific slices. (3) Both the network provider and the slice user can manage virtual resources. The design of the SDN controller in the framework uses the hierarchical method. Each domain has a local controller and these local controllers eventually converge to the central global controller. Some examples of the 5G-SDI controllers include OpenADN [15], OpenDayLight, FloodLight and OpenStack. Cloud-based technologies such as Cloud RAN and MEC can be introduced to enable a cloudified 5G infrastructure. Virtual network functions (VNFs) in low-latency slice like the slice for connected vehicles can be placed in edge clouds while VNFs in MBB slice can be provisioned on commodity servers in the core cloud.

The virtual resource layer provides all of the virtual resources required for network slices such as radio, computing, storage, network bandwidth and others. Virtual resources and VNFs run as virtual machine (VM)

instances. These resources are in fact an abstraction of the underlying physical resources managed at 5G-SDI and can be shared across slices, for example, the VM instance shared between virtual reality slice and MBB slice in the figure. Moreover, these resources can be used on-demand by network slices, which enhances the network resource utilization.

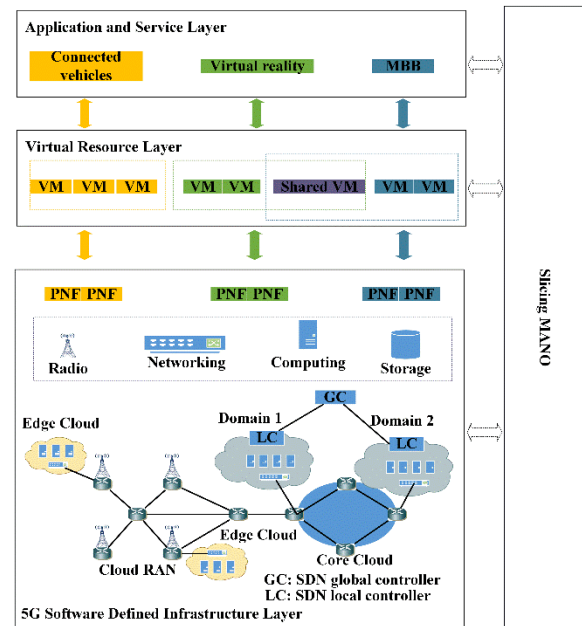


Fig. 2 5G network slicing framework

Heterogeneous service instances such as connected vehicles, virtual reality and MBB constitute the service layer. These service instances belong to multiple tenants. However, one tenant can still have several service types. The service instance informs slicing MANO of its service requirements, which is then mapped to the corresponding network slice. The dynamic nature of services mandates the network slices to be scalable at run-time. Services that have similar demands may share one network slice, and services with different service requirements can still share the slices partially.

The slicing management and orchestration (MANO) functional component is a key part of the network slicing framework. It manages and orchestrates the physical resources, virtual resources and network slices in a way the NFV-MANO will be built in. Slicing MANO can have sub-functions such as infrastructure manager, VNF manager, and slices orchestrator. The tasks of slicing MANO include: (1) creating and

managing VM instances by using the infrastructure resources. (2) mapping network functions to virtual resources and connecting network functions to create service chains. (3) managing the life cycle of network slices by interacting with the application and service layer, i.e., automated creation of service-oriented slices and dynamic maintenance by monitoring service requirements and virtual resources. In fact, the management of network slices may be performed at two levels, which are inter-slice management and intra-slice management. Inter-slice management is performed by the slice orchestrator while intra-slice management is implemented by running a virtual manager function as a part of the slice.

In the next section, we discuss some important research challenges and directions for network slicing in 5G networks. They include resource sharing, dynamic creation and management, isolation among network slices, mobility management, security and wireless resource virtualization in network slicing, and algorithmic aspects of resource allocation.

## V. CHALLENGES AND FUTURE RESEARCH AREAS

### A. Resource sharing

The Next Generation Mobile Networks (NGMN) has figured out that resource sharing among slice tenants is a key issue [5]. Resource sharing can be implemented by static partition sharing or elastically dynamic sharing. Due to the dynamic characteristics of network load, the dynamic resource sharing among slice tenants makes the utilization of network resources more efficient. There are some specific issues to be solved in resource sharing. For example, radio resources can be shared among RAN slices. It requires a good radio scheduling mechanism to allocate radio resources among these slices. In addition, computational resource sharing and other resource sharing need to be considered. Although resource sharing brings benefits to infrastructure providers, it also brings more challenges to other problems such as slice isolation.

### B. Dynamic slice creation and management

One of the concerns for future 5G network operators is how to allocate network resources to maximize their benefits. In this context, life-cycle management of the network slices is a critical problem to be solved. In order to accommodate the maximum number of diversified service requests, 5G network operators need to deploy virtual network functions and allocate network resources

rapidly to build network slices. In addition, they should be able to scale slices dynamically according to the varying service load. On the other hand, although a network operator has the maximum control over his network slices, one slice may still need to exercise some control over itself to improve the service quality. Accordingly, the network slicing technique should consider how to open partial permissions to each slice to configure and manage it without raising security issues. In addition, management of network slice needs to be implemented in an automated fashion to avoid manual efforts and errors.

### C. Isolation among network slices

Different services in 5G networks have unique requirements. As a consequence, dedicated virtual network resources are needed to guarantee the service quality at each slice. This requires slices to be highly isolated from each other. The isolation among network slices can be achieved via data plane isolation and control plane isolation. In general, the slice control function can be shared among slices, while for some services such as mission-critical communications the slice needs its own control function. Further, the effective isolation of network slices can ensure that the failure or security attack on one slice does not affect the operation of the other slices. Hence, the slice isolation mechanism is a key challenge of implementing network slicing.

### D. Mobility management in network slicing

Mobility aspects such as seamless handover and interference management bring new challenges to network slicing. Fast mobility handover is crucial for real-time services, which has a direct influence on the quality of service. The mobility support for network slicing is optional for some specific network slices. For example, network slices serving industrial control do not need mobility management functions due to fixed position of devices. However, network slices that do need mobility management differ in their mobility requirements. For instance, the mobility requirement of the mobile broadband slice is different from the slice for automated driving services. Thus, the design of a slice-oriented mobility management protocol is imperative to tackle mobility challenges in network slicing.

### E. Security in network slicing

Security in network slicing is a critical problem to be addressed because of resource sharing among slices. Network slices serving different types of services may have different levels of security policy requirements. Therefore, while designing network slicing security protocols, it is necessary to consider the impact on other slices and the entire network systems. In addition, security issues become more complex when network slicing is implemented on the multi-domain infrastructure. Security policy coordination mechanisms among different domain infrastructure need to be designed.

#### *F. Wireless resource virtualization*

Virtualization is the most important technology to implement network slicing. Virtualization technology has evolved over the past 20 years and been applied mainly in wired networks. A lot of work has focused on core network virtualization. However, wireless communication is more complex as wireless links are time varying and vulnerable to interference. Some of the virtualization methods in wired network cannot be directly used for their wireless counterparts. Therefore, designing new virtualization mechanisms to achieve the sharing of wireless spectrum resources and virtualization of base stations plays a vital role in implementing RAN slicing.

#### *G. Algorithmic aspects of resource allocation*

How to efficiently allocate slice resources is a challenging problem to be addressed. Slice resource allocation problem is similar to the virtual network embedding problem. The formulation of this problem can use mathematical methods in operation research such as integer linear programming (ILP). The algorithms to solve the problem may adopt different strategies according to the size of a slice, i.e., the number of virtual links and virtual nodes. The resource allocation for small-scale slices can be solved by exact algorithms, while for large-scale slices heuristic or meta-heuristic strategies are more efficient. It is worth noting that, due to the dynamic characteristic of 5G network, we need to design practical resource allocation algorithms with the ability to reconfigure and migrate slice resources.

## VI. CONCLUSIONS

Network operators have begun to seek new ways to support heterogeneous services. Network slicing is emerging as a possible methodology for allowing

network operators to create service instances quickly and flexibly enabling different services to have their own logical slice instances on a shared infrastructure. Network slicing together with SDN, NFV, MEC and other technologies will become a key in the service-oriented 5G network architecture. In this paper, we first introduce the relevant background for network slicing technology. Implementation of end-to-end network slices over 5G networks has then been discussed. We have proposed a network slicing framework consisting of a hierarchical virtual resource management system. Finally, we summarize the major challenges in network slicing and give directions for future research, especially from the perspective of the network slicing implementation over 5G networks. Network slicing has just started to attract the attention of academia and industry. It is necessary to take up the related specification and standardization to promote the deployment and application of network slicing technology. We believe that network slicing will be an indispensable technology in future 5G systems.

## ACKNOWLEDGMENT

This work has been supported under the grant ID NPRP 6-901-2-370 for the project entitled "Middleware Architecture for Cloud Based Services Using Software Defined Networking (SDN)", which is funded by the Qatar National Research Fund (QNRF), by Huawei Technologies, and by China Scholarship Council (No. 201506270075). The statements made herein are solely the responsibility of the authors.

## REFERENCES

- [1] Cisco, "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2015-2020," White Paper, 2016.
- [2] McKeown N, Anderson T, Balakrishnan H, et al. OpenFlow: enabling innovation in campus networks [J]. ACM SIGCOMM Computer Communication Review, 2008, 38(2): 69-74.
- [3] ETSI, "Network Functions Virtualisation: An Introduction, Benefits, Enablers, Challenges & Call for Action", October 2012, [https://portal.etsi.org/NFV/NFV\\_White\\_Paper.pdf](https://portal.etsi.org/NFV/NFV_White_Paper.pdf)
- [4] European Telecommunications Standards Institute (ETSI), "Executive Briefing – Mobile Edge Computing (MEC) Initiative," Sophia Antipolis, 2014.
- [5] NGMN Alliance, "Description of Network Slicing Concept," January 2016
- [6] Simon C, Maliosz M, B f6 J, et al. 5G exchange for inter-domain resource sharing[C]/Local and Metropolitan Area Networks (LANMAN), 2016 IEEE International Symposium on. IEEE, 2016: 1-6.
- [7] Peuster M, Karl H, Van Rossem S. MEDICINE: Rapid Prototyping of Production-Ready Network Services in Multi-PoP Environments [J]. arXiv preprint arXiv:1606.05995, 2016.

- [8] Nikaein N, Schiller E, Favraud R, et al. Network store: Exploring slicing in future 5g networks[C]//Proceedings of the 10th International Workshop on Mobility in the Evolving Internet Architecture. ACM, 2015: 8-13.
- [9] Rost P, Banchs A, Berberana I, et al. Mobile network architecture evolution toward 5G [J]. IEEE Communications Magazine, 2016, 54(5): 84-91.
- [10] [https://www.nttdocomo.co.jp/english/info/media\\_center/pr/2016/0613\\_00.html](https://www.nttdocomo.co.jp/english/info/media_center/pr/2016/0613_00.html)
- [11] Westerberg Erik, 4G/5G RAN architecture: how a split can make the difference, <https://www.ericsson.com/assets/local/publications/ericssontechnologyreview/docs/2016/etranarchitecture.pdf>
- [12] White Paper, 5G Service Guaranteed Network Slicing White Paper, March, 2017
- [13] ONAP, <https://www.onap.org/>
- [14] 3GPP TR 23.799 V14.0.0, "Study on Architecture for Next Generation System (Release 14)", December, 2016.
- [15] Paul S, Jain R. Openadn: Mobile apps on global clouds using openflow and software defined networking[C]//2012 IEEE Globecom Workshops. IEEE, 2012: 719-723.

**Raj Jain** is the Barbara J. and Jerome R. Cox, Jr., Professor of Computer Science and Engineering at Washington University in St. Louis. Previously, he was one of the Cofounders of Nayna Networks, Inc. He was a Senior Consulting Engineer at Digital Equipment Corporation. Contact him at [jain@wustl.edu](mailto:jain@wustl.edu).

**Xin Li** is a Ph.D. student in the School of Electronic Information at Wuhan University, China. He is currently a visiting Ph.D. student in Professor Raj Jain's group at Washington University in St. Louis, USA. Contact him at [lixin\\_itlab@whu.edu.cn](mailto:lixin_itlab@whu.edu.cn).

**Mohammed Samaka** is an associate professor of Computer Science in the Department of Computer Science and Engineering (CSE), College of Engineering at Qatar University. He obtained his PhD and Master from Loughborough University in England and a Post Graduate Diploma in Computing from Dundee University in Scotland. He obtained his Bachelor from Baghdad University. Contact him at [samaka.m@qu.edu.qa](mailto:samaka.m@qu.edu.qa)

**H. Anthony Chan** is currently with Huawei Technologies, USA and honorary professor at The University of Hong Kong, with current research in Wireless Internet. He is an IEEE Fellow, Distinguished Lecturer of IEEE Communication Society and IEEE Reliability Society. He has authored/co-authored 240 papers, a handbook, and 30 US patents. Contact him at [h.anthony.chan@huawei.com](mailto:h.anthony.chan@huawei.com).

**Deval Bhamare** is currently working as a post-doctorate at Qatar University. He has earned his dual PhD from IITB-Monash Research Academy, a joint venture between Indian Institute of Technology, Bombay, India and Monash University, Melbourne, Australia. His areas of research include Network Optimization, Middleware Architecture for Cloud Based Services, Cloud Security and Software Defined Networks. Contact him at [devalb@qu.edu.qa](mailto:devalb@qu.edu.qa).

**Lav Gupta** is a senior member of IEEE. He is currently a doctoral candidate in CSE at WUSTL. He has worked for 15 years in the area of telecommunications planning and deployment. He has also worked for many years as faculty in Access Network Planning in India and the UAE. Contact him at [lavgupta@wustl.edu](mailto:lavgupta@wustl.edu).

**Chengcheng Guo** is currently a professor and a Ph.D. supervisor in the School of Electronic and Information, Wuhan University, China. His research interests are internet and communication technology, wireless mesh networks, industry control networks and real-time and reliability communications. Contact him at [netccg@whu.edu.cn](mailto:netccg@whu.edu.cn).