International Journal of Computer Trends and Technology
ISSN: 2231 – 2803 / https://doi.org/10.14445/22312803/IJCTT-V68I10P109

Volume 68 Issue 10, 49-55, October 2020

Review Article

A Literature Review of Network Function Virtualization (NFV) in 5G Networks

Srinivasan Sridharan[1]

[1]*Principal Engineer – System Design, T-Mobile USA Inc.*

Received Date: 21 September 2020
Revised Date: 20 October 2020
Accepted Date: 24 October 2020

Abstract - *Network Function Virtualization (NFV), a crucial enabling technology for new-generation networks (5G), offers network users and providers an efficient platform to optimise resource utilisation. This survey paper provides a detailed outline of the scope of 5G networks and NFV in 5G. It presents the fundamentals of 5G networks. It investigates various studies conducted on NFV and explores the significant findings and recommendations provided in these works. It also presents a comparative study of existing works pertaining to this research region for analysing the technical gaps still prevailing in the domain. The paper provides the gaps in existing studies for developing better techniques and taking suitable steps to support future applications in 5G networks.*

Keywords - *5G, Cloud Computing, NFV, SDN, VNF*

I. INTRODUCTION

Recently, wireless technology has significantly emerged as the predominant trend in networking [1]. The explosively rising demands for wireless service create both opportunities and challenges in mobile networks, introducing the notion of fifth-generation (5G) wireless systems [2], [3]. The requirements and features of distinct services differ in 5G [4], [5]. The coordination and management of diverse applications, user demands, and networks necessitate a flexible and open network for ensuring resource allocation in the network with better efficiency. These demands can be fulfilled through mobile network virtualisation [6], which coordinates network resources and integrates diverse wireless networks. Numerous efforts have been put to enhance the 5G networking process and develop better technologies for fulfilling user demands. 5G centres its design goals around versatility, scalability, and efficiency to satisfy the device connectivity and capacity improvement. 5G networks should be efficient concerning cost, resource management, and energy to sustain their commercial feasibility. Connecting numerous terminals and several battery-operated components entails developing versatile and scalable network functions that deal with a broad level of service requirements, including high data rate for multimedia, low data rate for machine-type communication, and low power. The versatility, scalability, and efficiency goals of 5G network orient not only towards exploring innovative but simpler network function implementations of 5G.

The network functions of 5G encounter critical architectural and functional challenges, despite their performance excellence. For instance, coordinated multipoint (CoMP) can enhance the cell end-user experience through combined and coordinated signal transmissions from multiple terminals, antennas, sites, and cells to enhance uplink and downlink performance. Although, CoMP accomplishes this objective with increased backhauling, signalling overhead, computation, and equipment cost. Additionally, many components require a device-centric framework, specialised hardware, and ultra-densified systems/networks, which are still not well defined. Ultimately, 5G should coexist with other generations like 4G, 3G, and 2G. This demand solely increases complexity and cost. These challenges can be efficiently handled through implementing network functions of 5G by means of software components utilising the Network Function Virtualization (NFV) [7] epitome.

A rising group of standardisation bodies and industries is encouraging research and expansion of NFV epitomes to enhance flexibility and cost-efficacy. In NFV, network functions are implemented by vendors in software components known as Virtual Network Functions (VNFs), which are deployed on cloud infrastructure or massively distributed servers instead of dedicated hardware. For instance, NFV pools various signal processing resources in the cloud instead of utilising the devoted baseband processing modules at each site. This resource pooling minimises the signalling and computational overhead, improves flexibility, and optimises cost. The service provider triggers a specific resource for merely specific terminals/connections in the entire network rather than triggering the resources uselessly at every site. Typically, NFV is capable of overcoming certain 5G challenges through a) maximising the resource catering of VNFs for energy efficiency and cost, b) Scaling and mobilising VNFs from one resource to another, c) Ensuring VNFs coexistence with non-NFVs, d) Ensuring VNFs performance operations, maximum latency, and failure rate and tolerable expected packet loss.

The architectural framework of NFV defined by the European Telecommunication Standardization Institute (ETSI) is depicted in Figure 1. It enabled the execution and deployment of VNF on NFV infrastructure comprising a pool of resources (network, storage, and computing resources) bounded by a software layer that abstracts and rationally partitions them. The NFV infrastructure is usually a decentralised cloud infrastructure in which servers are distributed over various locations. ETSI defines network functions, including VNF. The operation, deployment, and execution of network services and VNFs in NFV infrastructure are controlled by an orchestration and management system, whose performance is steered by NFV descriptors. The procedure for exemplifying a network service is started from the operations support system of service providers.

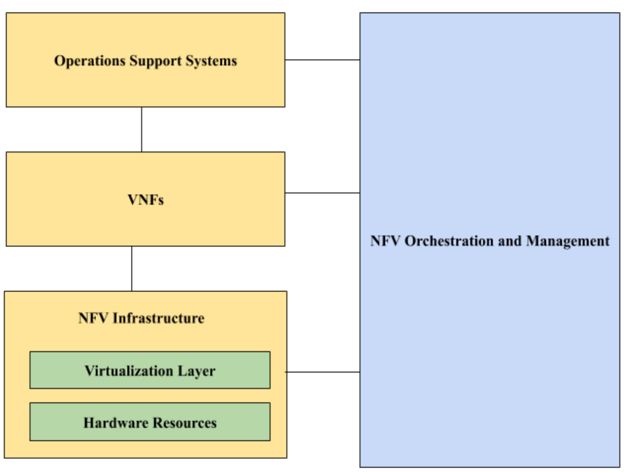

Fig. 1 Architectural framework of NFV [8]

A. Contributions to the Study

This paper explains the scope of NFV in 5G networks. The fundamentals of 5G networks are discussed, and existing studies concerning their objectives, techniques used, and future scopes are investigated. A comparative study is presented to depict the performance of related works. The major technical gaps persisting in this research are discussed. Ultimately, the significant observations and areas for improvement are presented.

II. FUNDAMENTALS OF 5G NETWORKS

The emergence of 5G technology will develop the internetwork with a substantially growing user experience range, as many of the existing limitations of traditional cellular networks will terminate to exist. The burst-type information traffic provisioning and inherent support for distinct wireless systems will be integrated into the initial deployment phases. In contrast, inefficient usage of low latency, mobile base station (MBS) processing competencies, and co-channel interference will be almost eliminated. Therefore, such an improved medium will never be responsible for low QoS, although the following characteristics need to be considered when reviewing user/customer satisfaction. Figure 2. depicts the fundamentals of 5G.

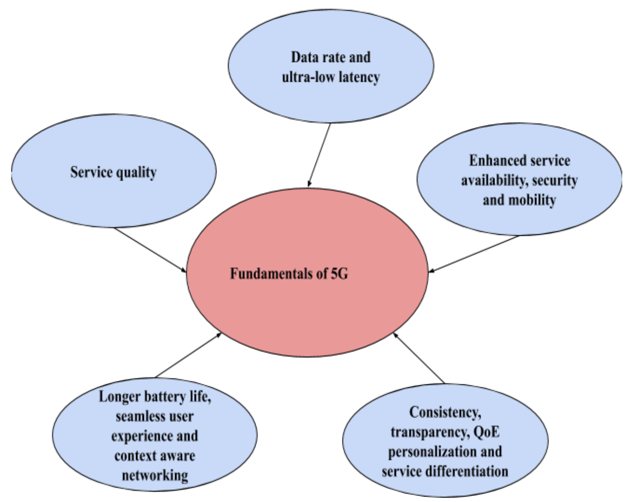

Fig. 2 Fundamentals of 5G

A. 5G Service Requirements

5G User Equipment (UE) should support a broad level of software and hardware technologies to provide uninterrupted and consistent service with exceptional quality. Although all devices must be competent enough for dealing with problems like high node density, lack of spectrum resources, and unpredictable channel conditions, limited conflicts might be present. As all wireless components frequently migrate from one cell to another, owing to their mobile nature, a significant factor that enhances the end-user experience is none other than a few handoff mechanisms to provide such transitions. Especially, some network management policies must be redesigned for better network functioning and management. Service quality and consistency influence seamless connectivity; thus, a self-healing/restoring network infrastructure is essential, irrespective of the failure cause, adjusting the desired operating channels over adjacent cells, and restoring end-users connectivity.

B. Data Rate and Ultra-Low Latency

High demands for increased data rate and ultra-low latency of users are the prime reasons for cellular generations (2G,3G,4G). The ever-increasing demands of customers have necessitated the development of 5G networks. The 5G networks are anticipated to offer users greater speed, data rate, and ultra-low latency features.

C. Enhanced Service Availability, Security, and Mobility

In a customer-centric communication system, all existing services are anticipated to be linked with a dynamically generated or pre-existing user profile once started. Specifically, when business frameworks are engaged, providing personalised content, using distinct network resources could be a requirement instead of a mere improvement. Significant deviations in customer expectations, behaviours, and priorities under enhanced quality conditions with a price were observed in [9]. Per-customer service customisation is not just the approach for achieving superior QoS metrics, as the application is considered to play a pivotal role. Not all applications influence factors or possess similar requirements despite providing an identical range of customer satisfaction. Security and mobility are also major concerns in 5G systems. The unprotected interfaces which support network programming contribute to new attacks and threats to software systems. This encourages a stable multi-level security model development for remote attestation and software integrity.

D. Consistency, Transparency, User's Qoe Personalisation, and Service Differentiation

Users always prefer consistency, service differentiation, transparency, and QoE personalisation in the network. Therefore, a network that offers an enhanced range of services with better quality simultaneously remains cryptic concerning error handling, traffic management, functionality, and resource assignment is preferable to users. Interactivity between the network and the user must not be obsolete instead of confined to the minimum. Moreover, data collection should be carried out transparently and intelligently, for example, through monitoring device-pertinent metrics and applications or employing network probes.

E. Longer Battery Life, Seamless User Experience, And Context-Aware Networking

Energy efficiency is a significant area in mobile communication that needs to undergo a redesign to satisfy the high demands of 5G networking. According to [10], a continuous rise in power usage is not feasible from a cost, battery-technology, and logistical perspective. Rapid enhancement of network density is linked directly to the raised energy demands. While the need for UE for uninterrupted connectivity and anticipated support of utilities, broad-spectrum and software enhance the total cost of computation subsides the total battery period per charge. In the wireless domain, the Internet of Things is regarded as another field that could be energy-sensitive, as wireless sensors usually possess confined operational proficiencies owing to feeble batteries. An approach like edge offloading recommends the migration of various services in the operators' domain instead of UE is considered a reasonable solution for power drain. Moreover, the expected downlink and uplink decouple, wherein the UE will employ channels from distinct MBSs or data decoupling and signalling, will enable inactive BS's power-off on-demand, possibly enhancing the QoE of the end-user indirectly through prolonged battery life. Also, introducing the QoE-conscious mechanism capable of acquiring realistic network datasets and using them in network management adds significant complexity. All network components must widen energy provisioning for controlling signalling overhead, simultaneously retaining the service standards to the greatest possible level. The quality feedback [11] enables users to utilise a QoE conscious network [12], and a seamless user experience should be implemented. In [13], an approach was presented for monitoring QoE at the user end, depending on quality deterioration. The probes were activated only in the case of quality degradation conditions.

III. RELATED WORK

In [14], an outline of sophisticated standardisation views of network slicing (NS). The crucial challenges in NS faced by mobile operators include difficulties in attaining end-to-end NS, stable migration, interoperability, and roaming. The work highlighted that academia and the mobile industry should emphasise base station virtualisation and wireless resource sharing to formulate appropriate requirements and create standardised slices. Moreover, the industry needs to provide a solution for ensuring NS's interoperability and ensuring the seamless and stable movement from physical networks to virtual ones.

In [15], 5G architecture design was presented for NS and building NFV, Software Defined Network (SDN) technologies. It emphasised schemes that provide effective substrate resource utilisation for NS. Furthermore, key concepts like service-sensible Quality of Service (QoS) control, service-customised mobility, and network orchestration were considered. Additionally, the word signified the necessity of tackling open issues such as granularity control and slice management.

In [16], an overview of enabling technologies like NFV and SDN for 5G. It explained how these enabling technologies complemented one another in 5G networks. It described the 5G slicing notion and the prominent challenges. It highlighted a few unaddressed challenges for ensuring an envisaged 5G networking system. In [17], the concept of NS as a service was presented for providing customised services. Furthermore, the service agreement and service orchestration were introduced to demonstrate the service management architecture across distinct service frameworks. However, with advancing requirements, the necessity of improving SDN/NFV technologies for developing better communication frameworks was suggested.

In [18], optimal deployment of VNF for 5G NS and assignment of resources by considering the 5G service requirements were presented. The proposed framework could completely exploit the resources in the cloud, with large or limited computational potential, through effective load distribution in-network and utilising computational resources wherever required. This enhanced the efficiency of resource usage and the number of VNF chains supported by the system in terms of static solutions. However, realistic implementation limits like configuration delays, instantiation, and orchestration were not considered. The necessity of considering the dynamic settings for avoiding service outages through artificial intelligence solutions was recommended in the proposed Study. The VNF forwarding graph (FG) and VNF placement design were presented in [19] for 5G networks. Initially, the VNF-FGs were generated as per service requests of the network. Then the VNFs were mapped to physical resources in the VNF-FGs by modifying the VNF environment by introducing many mapping nodes and physical nodes for practicality and completeness, followed by the formulation of VNF positioning optimisation issue for achieving lesser link utilisation and lesser bandwidth consumption simultaneously. The proposed approach attained better performance concerning link utilisation and bandwidth consumption. However, other issues associated with implementing NFV were not discussed in the Study.

In [20], an approach was presented for overcoming the performance deterioration issue of virtualised access points occurring due to NFV implementation. An effective packet processing method was proposed to boost mobile network virtualisation performance. Virtualisation was achieved by locating the management frame in user space and locating data and control frames in kernel space to minimise packet processing delays. In [21], a framework for mobile network virtualisation comprising three planes, namely control, cognitive, and data planes, was presented. A control method depending on cell clustering was employed with optimised resource utilisation efficiency from a network virtualisation implementation perspective.

Furthermore, two distinct use cases were analysed to illustrate the proposed method's behaviour concerning enhancing the resource efficacy and user experience. In [22], integrating fog computing (FC) and NFV was explored to enhance the handover performance of 5G networks. The prime reason for integrating NFV was its network flexibility enhancement characteristic whilst decreasing the total overhead. It proposed a direct-X2 handover method utilising the benefits of both virtual machines and edge caching. It performed better with respect to handover management than the traditional networks. However, further optimisation is required in this context.

IV. COMPARATIVE STUDY

Table 1. Comparative performance analysis of existing studies

Ref	Objectives	Technology & Findings	Future Scope
[8]	To achieve NFV-based NS for 5G.	NFV: Presented the options offered by NFV technology for enabling NS.	Further research about security orchestration, multilayer orchestration, VNF granularity is required.
[23]	To achieve NS in 5G networks.	NFV & SDN: Presented a scenario combining NFV and SDN to address NS's realisation.	Proper resource management techniques are required to facilitate resource sharing between the slices without compromising performance.
[24]	To achieve better NS in 5G networks.	NFV & RAN: Presented fundamental architectural principles required for accommodatin g NS in 5G systems and its core networks.	Better techniques should be explored for improving NS performance.

[25]	To develop optimised energy-conscious 5G NFV.	NFV: Achieved 38% energy saving.	Optimised techniques for addressing resource assignment issues should be inspected.
[26]	To assure dynamic QoS/QoE in realistic NFV-based 5G networks.	RAN & SDN: Proposed a better solution for addressing QoS/QoE requirements.	Schemes for integrating other parameters influencing the 5G performance must be investigated.
[27]	To facilitate the future generation underwater communication network development through exploiting SDN, NFV paradigms.	SDN, NFV: Presented SDN-enabled underwater solutions such as optimal throughput underwater routing, fault recovery methods, and underwater flexibility management.	Energy-proficient schemes should be developed for tackling constrained energy-resource issues.
[28]	To achieve better end-to-end NS.	NFV & RAN: Provided a good solution for end-to-end NS.	An optimal number of resource blocks assigned to distinct slices and the overall assignment of resources for orchestration must be determined.
[29]	To develop a software system for 5G.	SDN & NFV: Exploited novel concepts of network virtualisation and network cloudification for providing a resilient, flexible, and scalable network architecture.	The significant tradeoffs between latencies and reliabilities need to be taken into account.
[30]	To investigate the problems pertaining to the network utility deterioration during NFV implementation and develop a better solution.	NFV: Enhanced network performance without compromising the service quality.	More generic implementation of the proposed approach is required, and a dynamic pricing framework needs to be developed by considering the networking traffic patterns and market dynamics.

V. RESEARCH GAPS

Some of the open-ended issues are still unaddressed in the 5G networks. Managing the liveliness of software, particularly from the perspective of the necessity or trend in the disintegration of network function into sub-function, has become tedious. The distribution of network functions into distinct execution platforms has become more pervasive in the case of highly programmable hardware. Also been tedious to ensure interoperability among distinct vendors, particularly in the cloud environment of enormously degraded network functions. Furthermore, the appropriate method of mapping and translating the tenant's and customers' business demands over service providers' infrastructure has not been known. Moreover, managing and assigning massive groups of slices within the inter-administrative and intra-administrative domain has become challenging. Also, the technique or level of automating system/network management to reduce the necessity of manual efforts and intervention has become sophisticated.

The privacy and security concerns originating from slicing are a primary impediment to adopting multitenancy schemes. The network programming interfaces create new attacks and threats to software systems, necessitating a multi-level security model comprising mechanisms and policies for user authentication, remote attestation, software integrity, and dynamic threat recognition and mitigation. Orchestration and management in multitenant environments are tough. For flexible resource assignment to slices, the policy which governs the orchestrator needs to tackle situations wherein there is a frequent variation in resource demands within short timescales. This can be accomplished by proper coordination between the orchestrator and slice-explicit functional blocks, enabling both the orchestrator and slice-explicit functional blocks to perform configuration and management operations quickly. Furthermore, computationally competent resource assignment techniques should be developed.

VI. CONCLUSION

NFV being an enabling technology, offers numerous benefits to users. Although, through the inherent uncertainties and network dynamics, NFV adoption is slowed down in several budding networking applications. This has encouraged us to survey various aspects of NFV in 5G. This paper described the scope of NFV in 5G networks. It described the fundamentals of 5G networks. It investigated several types of research performed on NFV and explored the crucial observations and suggestions provided in the studies. It also presented a comparative study of existing works concerning their objectives, enabling technologies, findings, and future directions. The paper revealed that despite continuous advances in this domain, there is still massive scope for NFV development. It provided the major gaps prevailing in literary studies for taking relevant decisions and designing better schemes by addressing the technical gaps and challenges for supporting further applications in 5G systems.

REFERENCES

[1] Akpakwu, G. A., Silva, B. J., Hancke, G. P., & Abu-Mahfouz, A. M. A survey on 5G networks for the Internet of Things: Communication technologies and challenges IEEE Access, 6 (2017) 3619-3647.

[2] Agiwal, M., Roy, A., & Saxena, N. Next-generation 5G wireless networks: A comprehensive surve. IEEE Communications Surveys & Tutorials, 18(3) (2016)1617-1655.

[3] Osseiran, A., Boccardi, F., Braun, V., Kusume, K., Marsch, P., Maternia, M., ... & Tullberg, H. Scenarios for 5G mobile and wireless communications: the vision of the METIS project". IEEE communications magazine, 52(5) (2014) 26-35.

[4] Ateya, A. A., Muthanna, A., Makolkina, M., & Koucheryavy, A. Study of 5G services standardisation: specifications and requirements. In 2018 10th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT) (2018)1-6).

[5] Yassein, M. B., Aljawarneh, S., & Al-Sadi, A. Challenges and features of IoT communications in 5G networks. In International Conference on Electrical and Computing Technologies and Applications (ICECTA) (2017)(1-5).

[6] Liang, C., & Yu, F. R. Wireless virtualisation for next-generation mobile-cellular networks. IEEE wireless communications, 22(1) (2015). 61-69.

[7] Abdelwahab, S., Hamdaoui, B., Guizani, M., & Znati, T. Network function virtualisation in 5G. IEEE Communications Magazine, 54(4) (2016)84-91.

[8] Chatras, B., Kwong, U. S. T., & Bihannic, N. NFV enabling network slicing for 5G. In 20th Conference on Innovations in Clouds, Internet and Networks (ICIN) (2017) 219-225.

[9] Sackl, A., Zwickl, P., & Reichl, P. The trouble with a choice: An empirical study to investigate the influence of charging strategies and content selection on QoE. Proceedings of the 9th International Conference on Network and Service Management (2013) 298-303.

[10] Andrews, J. G., Buzzi, S., Choi, W., Hanly, S. V., Lozano, A., Soong, A. C., & Zhang, J. C. What will 5G be?. IEEE Journal on selected areas in communications, 32(6) (2014) 1065-1082.

[11] Ahmad, A., Atzori, L., & Martini, M. G. Qualia: A multilayer solution for QoE passive monitoring at the user terminal. In IEEE International Conference on Communications(2017) (1-6).

[12] [Tselios, C., & Tsolis, G. On QoE-awareness through virtualised probes in 5G networks. In the IEEE 21st International Workshop on Computer-Aided Modelling and Design of Communication Links and Networks (2016) 159-164.

[13] [Ahmad, A., Floris, A., & Atzori, L. Towards QoE monitoring at user terminal: A monitoring approach based on quality degradation. The IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB) (2017). 1-6.

[14] Kim, D., & Kim, S. Network slicing as enablers for 5G services: state of the art and challenges for the mobile industry. Telecommunication Systems, 71(3) (2019) 517-527.

[15] Yousaf, F. Z., Gramaglia, M., Friderikos, V., Gajic, B., von Hugo, D., Sayadi, B., ... & Crippa, M. R. Network slicing with flexible mobility and QoS/QoE support for 5G Networks. In IEEE International Conference on Communications Workshops (ICC Workshops) (2017) 1195-1201.

[16] Yousaf, F. Z., Bredel, M., Schaller, S., & Schneider, F. NFV and SDN—Key technology enablers for 5G networks. IEEE Journal on Selected Areas in Communications, 35(11) (2017) 2468-2478.

[17] Zhou, X., Li, R., Chen, T., & Zhang, H. Network slicing as a service: enabling enterprises' own software-defined cellular networks. IEEE Communications Magazine, 54(7) (2016) 146-153.

[18] De Domenico, A., Liu, Y. F., & Yu, W. Optimal Virtual Network Function Deployment for 5G Network Slicing in a Hybrid Cloud Infrastructure. IEEE Transactions on Wireless Communications. (2020).

[19] Cao, J., Zhang, Y., An, W., Chen, X., Sun, J., & Han, Y. VNF-FG design and VNF placement for 5G mobile networks. Science China Information Sciences, 60(4) (2017) 040302.

[20] Wang, X., Xu, C., Zhao, G., & Yu, S. Tuna: An efficient and practical scheme for wireless access points in 5G networks virtualisation. IEEE Communications Letters, 22(4) (2017)748-751.

[21] Feng, Z., Qiu, C., Feng, Z., Wei, Z., Li, W., & Zhang, P. An effective approach to 5G: Wireless network virtualisation. IEEE Communications Magazine, 53(12) (2015) 53-59.

[22] Qiu, Y., Zhang, H., Long, K., Sun, H., Li, X., & Leung, V. C. Improving the handover of 5G networks by network function virtualisation and fog computing. In IEEE/CIC International Conference on Communications in China (ICCC) (2017) 1-5.

[23] Ordonez-Lucena, J., Ameigeiras, P., Lopez, D., Ramos-Munoz, J. J., Lorca, J., & Folgueira, J. Network slicing for 5G with SDN/NFV: Concepts, architectures, and challenges. IEEE Communications Magazine, 55(5) (2017) 80-87.

[24] Rost, P., Mannweiler, C., Michalopoulos, D. S., Sartori, C., Sciancalepore, V., Sastry, N., . & Aziz, D. Network slicing to enable scalability and flexibility in 5G mobile networks. IEEE Communications Magazine, 55(5) (2017) 72-79.

[25] Al-Quzweeni, A. N., Lawey, A. Q. Elgorashi, T. E., & Elmirghani, J. M. Optimised energy-aware 5G network function virtualisation. IEEE Access, 7(2019) 44939-44958.

[26] Pedreno-Manresa, J. J., Khodashenas, P. S., Siddiqui, M. S., & Pavon-Marino, P. Dynamic QoS/QoE assurance in realistic NFV-enabled 5G access networks. At the (19th international conference on transparent optical networks (ICTON) (2017) (1-4).

[27] Akyildiz, I. F., Wang, P., & Lin, S. C. SoftWater: Software-defined networking for next-generation underwater communication systems. Ad Hoc Networks, 46 (2016) 1-11.

[28] Afolabi, I., Bagaa, M., Taleb, T., & Flinck, H. End-to-end network slicing is enabled through network function virtualisation. In IEEE Conference on Standards for Communications and Networking

(CSCN). (2017) 30-35
[29] Akyildiz, I. F., Wang, P., & Lin, S. C. software: A software-defined networking architecture for 5G wireless systems. Computer Networks, 85 (2015) 1-18.
[30] Cheng, X., Wu, Y., Min, G., & Zomaya, A. Y. Network function virtualisation in dynamic networks: A stochastic perspective. IEEE Journal on Selected Areas in Communications, 36(10) (2018) 2218-2232.
[31] Sridharan, S. Systems and methods for authorisation and billing of users for wireless charging. US Patent No.9,972,037. Washington, DC: US Patent and Trademark Office. (2018).
[32] Sridharan, S. Ue initiated the evolved packet core (epc) and ip multimedia subsystem (ims) network usage optimisation algorithm for lte capable smartphones connected to wireless lan (wi-fi) network. US Patent No.20160088677A1. Washington, DC US Patent and Trademark Office. (2016).
[33] Sridharan, S. Network initiated an evolved packet core (epc) and ip multimedia subsystem (ims) network usage optimisation algorithm for lte capable smartphones connected to wireless lan (wi-fi) network . US Patent No.20160088678A1. Washington, DC: US Patent and Trademark Office. (2016).