

Review

Network Slicing for Beyond 5G Systems: An Overview of the Smart Port Use Case

Rana Muhammad Sohaib * , Oluwakayode Onireti , Yusuf Sambo  and Muhammad Ali Imran 

James Watt School of Engineering, University of Glasgow, Glasgow G12 8QQ, UK;
Oluwakayode.Onireti@glasgow.ac.uk (O.O); Yusuf.Sambo@glasgow.ac.uk (Y.S.);
Muhammad.Imran@glasgow.ac.uk (M.A.I.)

* Correspondence: 2500800S@student.gla.ac.uk

Abstract: As the idea of a new wireless communication standard (5G) started to circulate around the world, there was much speculation regarding its performance, making it necessary to carry out further research by keeping in view the challenges presented by it. 5G is considered a multi-system support network due to its ability to provide benefits to vertical industries. Due to the wide range of devices and applications, it is essential to provide support for massively interconnected devices. Network slicing has emerged as the key technology to meet the requirements of the communications network. In this paper, we present a review of the latest achievements of 5G network slicing by comparing the architecture of The Next Generation Mobile Network Alliance's (NGMN's) and 5G-PPP, using the enabling technologies software-defined networking (SDN) and network function virtualization (NFV). We then review and discuss machine learning (ML) techniques and their integration with network slicing for beyond 5G networks and elaborate on how ML techniques can be useful for mobility prediction and resource management. Lastly, we propose the use case of network slicing based on ML techniques in a smart seaport environment, which will help to manage the resources more efficiently.

Keywords: 5G; NGMN; 5G-PPP; SDN; NFV; NS; ML; RRM



Citation: Sohaib, R.M.; Onireti, O.; Sambo, Y.; Imran, M.A. Network Slicing for Beyond 5G Systems: An Overview of the Smart Port Use Case. *Electronics* **2021**, *10*, 1090. <https://doi.org/10.3390/electronics10091090>

Academic Editor: Ana Vázquez Alejos, Manuel García Sanchez, Francisco Falcone and Leyre Azpilicueta

Received: 7 April 2021
Accepted: 3 May 2021
Published: 5 May 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Remote devices and wireless sensor networks have become a key part of our lives and the existing 4G network cannot meet the requirements of low latency and higher data rates. In recent years, the emergence of the Internet of things (IoT) has revolutionized almost every aspect of the communication field. Next-generation wireless networks make it possible for IoT devices to transform businesses by providing higher data rates and ultra-reliable low-latency communication (URLLC). Different technologies have already been developed for IoT networks, such as ZigBee, SigFox, and long range (LoRa) [1,2]. However, these technologies are not compatible with the current cellular network, and there is a gap to be filled in order to address this challenge. The architecture of previous generations of mobile networks does not meet the diversity, scalability, and performance requirements because they are not flexible. The fifth generation (5G) network was designed to overcome these challenges, and can support three sets of use cases with diverse requirements in terms of data rate, reliability, latency, and massive connectivity. Services in 5G are classified into three main categories, namely:

- Enhance Mobile Broadband (eMBB).
- Massive Machine-Type Communications (mMTC).
- Ultra-Reliable Low-Latency Communications (URLLC).

Network slicing can provide answers to fulfill the above-mentioned specifications by partitioning or slicing the physical network into multiple logical networks according to their requirements [3]. Each logical network should have its attributes, which allows for the

better utilization and allocation of the resources as compared to previous mobile networks. Logical networks provide flexibility and can be modified or operated dynamically to meet the requirements of the business segments according to service level agreement. All such logical sub-networks are treated as network slices, and a network slice consists of a set of network resources and functions. For instance, a logical sub-network can be customized to meet URLLC requirements. With network slicing, control and predictability can be achieved by reserving and isolating network resources and functions. From a security perspective, a logical sub-network can be connected to a virtual private network (VPN) by isolating it from the Internet. Due to the isolation of logical sub-networks, it will be easy to customize the customer needs according to the business model, as it was time-consuming over common infrastructure in the prior models. According to a GSMA report published in 2018 [3], NS will be the key technology of 5G opening the business opportunity worth USD 300 billion by 2025.

It will allow the enterprise community to customize the services according to their own needs [3]. The concept of dividing the network into logical networks is not new as it has been used in VPN and virtual local area networks (VLANs). In 4G network slicing, dedicated core (DECOR) permits the implementation of multiple core networks within a common infrastructure [4]. Due to the limitations of evolved packet core (EPC) architecture, it does not offer flexibility and it also considers that the radio access network (RAN) is the same for all cases [5]. Different organizations in the past have contributed to fulfilling the requirements of vertical applications. The 5G Public–Private Partnership (5G-PPP) defines the relevant network resources for management and diversity with a five layer architecture [6]:

- Service layer.
- Infrastructure layer.
- Business function layer.
- Orchestration layer.
- Network function layer.

The objective of this paper is to analyze the recent developments in network slicing and present an approach by using ML algorithms in order to manage the resources more efficiently. In this paper, we have provided an overview of network slicing in 5G networks and present some enabling technology for network slicing in beyond 5G network. Some of the enabling technologies for network slicing include software defined network (SDN), network function virtualization (NFV), machine learning, artificial intelligence, etc. SDN and NFV are envisioned as key technologies to execute the network slicing. The paper also present an overview of network slicing for the smart port 5G use case. A wide range of network infrastructure and a diverse range of services needs to be provided within the area of the seaport. This could range from offering enhanced broadband applications (e.g., virtual reality and augmented reality for passengers on cruise ship), critical services (traffic lights and water control gates at seaport), and massive connectivity (massive number of sensor) [7].

Hence, we describe the smart port scenario based on network slicing by integrating ML techniques with SDN/NFV architecture in order to manage the resources more effectively and efficiently.

2. Network Slicing in 5G

Slicing of the network should be done in an end-to-end (E2E) manner in order to provide the dedicated services to specific applications over a common physical infrastructure. There are some technical challenges for the development of network slicing over 5G network that need to be solved. These challenges include efficient management of resources, mobility prediction, security, and end-to-end orchestration. In Sections 5 and 6, we discuss how ML techniques can be utilized to manage the resources intelligently and predict the mobility. Many research projects, such as 5GEx [8], 5GNORMA [9], 5G-MEDIA [10], and 5G-XHaul [11], have provided the study of NS enabling technologies using SDN and NFV.

There is a need for optimization and new architecture to cope with new challenges. The NGMN defines the concept of NS in three layers [12]:

- Service instance layer.
- Network slice instance layer.
- Resource layer.

2.1. Service Layer

The service layer consists of specific services that are supported by the network operator or third party. It is linked to management and network orchestration (MANO) via an interface that allows it to create the dedicated slices for the specific application. The 5G network slicing architecture is shown in Figure 1. The services layer shows the supported applications and each application is depicted as a service instance.

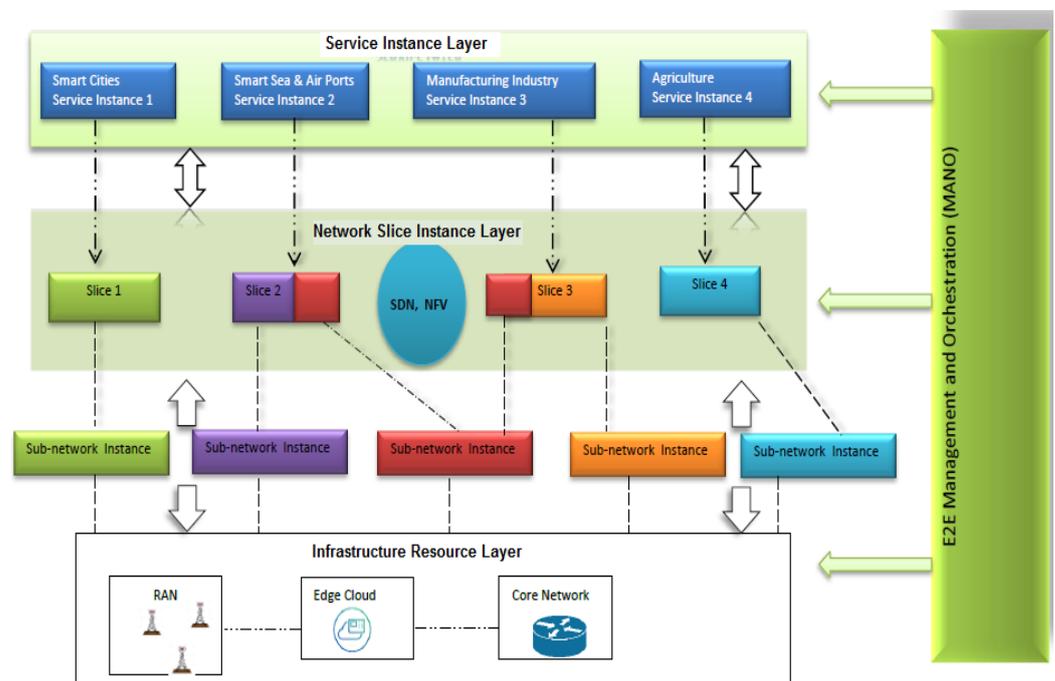


Figure 1. 5G network slicing architecture.

2.2. Network Instance Layer

In network function layer, a blueprint is used to initiate the network slice instance and it presents the network attributes which are essential for it. This can also be shared over multiple services instances as it may consist of more than one sub-instances [13]. A network slice instance can be logically separated from the rest of network slice instances. Sub-network instance consists of logical or physical resources and set of network functions, defined by sub-network blueprint [13]. A sub-network instance can split into more than one slices.

2.3. Resource or Infrastructure Layer

It comprises of actual physical and virtual resources of the RAN, networking nodes, cloud nodes, and associated links. Through virtualization, the flexibility of network resources can be achieved, which helps to create the core network slices. Spectrum resources can be shared and flexible management of these resources can be achieved by slicing the RAN. Scheduling technique can be useful to fulfill the task of radio resource sharing. It helps to dedicate the resources according to the application requirements.

In addition to the layers, NGMN also described the E2E management and orchestration concept for NS. For a specific service, it called upon the request of applicable network

function, manages the physical and virtual resources, allocates the relevant functions, and monitors each slice via application programming interface (API). It orchestrates the virtual resources and maps them to physical infrastructure resources.

3. Enabling Technologies

Virtualization can be achieved by using enabling technologies such as SDN and NFV and it helps in implementing network slicing in 5G. In this section, we discuss the key network slicing enabling technologies.

3.1. Software Defined Networking (SDN)

The purpose of SDN is to replace the existing networking system with programmable services by separating the data plane and control plane (CP) [14]. It offers the flexibility, security, and less operating cost as compared to a traditional network. Due to the centralized architecture, it is easy to manage and run the operations across the larger network. The control plane is easy to configure, as it provides useful management functions. Controllers help to forward the packets in an SDN via APIs. A comparison of the existing network and SDN is shown in Figure 2. SDN is lean to undertake the traditional network issues by offering the centralized programmable services, which are well suited for the heterogeneous environment. The controller in SDN organizes the slicing of the network instance by using the protocols. It helps to create the client and server environment along with the installation of their concerned protocols as defined in the Open Network Foundation (ONF) SDN architecture [15]. SDN also offers the ability to promptly answer to the changing network environments and enterprise community.

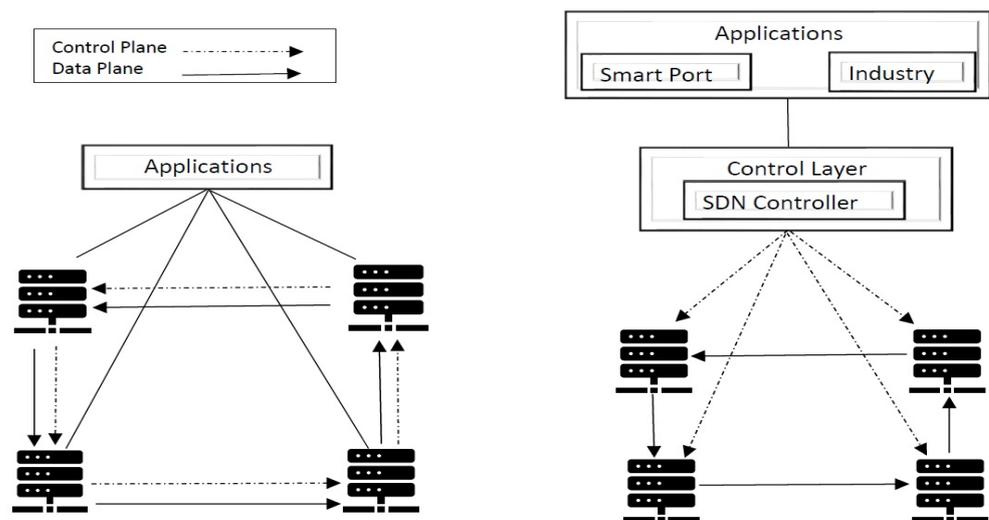


Figure 2. Existing network (left) vs. SDN (right).

3.2. ONF SDN Architecture

ONF SDN architecture is ideal for 5G network slicing because it meets the application services requirements in an effective manner, as the CP dynamically configures the data plane resources to carry out customized services to the application plane. As described in [16], controllers and resources (e.g., infrastructure resources) are envisioned as key elements. A controller is defined as a programmable centralized unit, which manages the resources in an effective manner and delivers services to the underlying data plane. As shown in Figure 3, controllers are an essential part of the architecture, as they form a link between clients and resources via client and server contexts. A controller needs all the details to mediate with a respective client. The client context consists of the resource group and support function. Moreover, the resource group provides responsive approach and enable its connection with the controller. Comparisons of different SDN controllers has been shown in Table 1. Client operations are supported by the client support function,

based on protocols that enable the clients to interact with the controller. A controller needs all the details from the server context in order to mediate with the respective resources. The controller performs the virtualization and the clustering of the underlying resources. Orchestration enables the controller to dynamically forward the respective resources to different resource groups. Controllers are connected to applications and resources via northbound interface and southbound interface, respectively. These interfaces help users and applications to connect with the network.

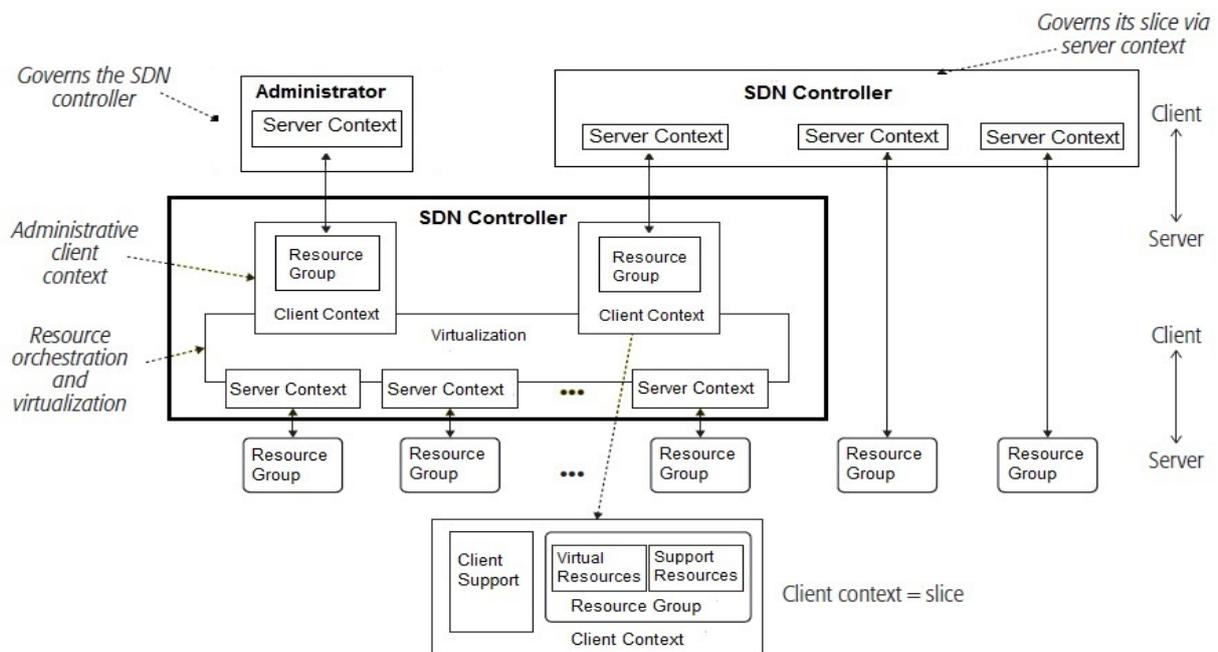


Figure 3. ONF SDN network slicing architecture [15].

Table 1. Comparison of SDN controllers.

SDN Controllers	Throughput (millions/sec)	Latency μ s	Architecture
Floodlight	0.2	25	Centralized
Onix	2.2	15	Centralized
ONOS	1.2	50	Distributed
OpenDaylight (ODL)	0.1	10	Centralized

The administrator in the architecture enables the configuration of the controller and also establishes the client and server contexts. Openflow allows the transmission between infrastructure resources and the controller, and it also helps to procure the traffic management. Openflow forwards the traffic between layer 2 and layer 3 networks and optimally updates the forwarding protocols, acting as a interface between data plane and control plane. Although, it heavily depends on the controller to establish the status of the transmitted packets [17]. Some of the 5G network slicing challenges, such as scalability and security, can be raised by Openflow due to the occurrence of a processing delay between controller and layer 2/layer 3 devices [17]. It can be overcome by advanced stateful forwarding technologies such as OpenState as it provides flexibility by increasing the network virtualization ability.

3.3. Network Function Virtualization (NFV)

We have discussed the SDN architecture and presented a detailed review of its functionalities. Due to its limitations, it lacks the ability to effectively operate the network

slices and its respective resources [18]. In contrast, the NFV architecture helps to manage the resources more effectively [19]. NFV enables the virtualization attributes in order to remodel the network functions into virtual network functions (VNF), which can be combined with each other for different network services. NFV uses end nodes for the execution of virtualization to implement the network functions (e.g., VPN, content delivery network (CDN), and firewalls), and can be allocated dynamically. As compared to SDN, by using NFV, infrastructure resources can be effectively allocated to VNFs. It will allow the Internet service providers (ISPs) to instantiate new services according to the customer demands without changing the infrastructure hardware. Further, it will also cut short the operational expenditure (OPEX) and capital expenditure (CAPEX). It will provide flexibility and becomes easy to manage, if an ISP needs to update a function then configuration only be required at ISP end, also results in decreasing the implementation time. NFV provides quality of service (QoS) and serve the low latency requirements by separating the software entities from hardware. Decoupling offers flexibility since both can execute their functions separately and dynamically which allows ISPs to install more services using the same infrastructure resources. Conventional models of virtual machines (VMs) concepts do not meet the performance and high-availability requirements of the cloud environment [20].

3.4. Management and Orchestration

NFV MANO system provides a dynamic control of network infrastructure resources, it helps to manage cloud environment and network infrastructure resources in SDN and VNF manner. The NFV MANO framework presented by ETSI envisioned as reference model is illustrated in Figure 4 [21]. The framework is grouped into the following functional blocks:

- Management and Orchestration (MANO): Manages and coordinates the virtualization in the architecture. It consists of the following three entities:
 - (1). Virtualized infrastructure manager (VIM): Helps to manage the network functions virtualization infrastructure (NFVI) resources.
 - (2). VNF manager (VNFM): Helps to provide connectivity and to configure the VNFs.
 - (3). Orchestrator: Helps to manage the life cycle of network infrastructure resources and ensures the integrity of process-to-process service.
- Network Functions Virtualization Infrastructure (NFVI): It is connected to the VNFs.
- Network Management System (NMS): It consists of element management (EM) and operation/business support system (OSS/ BSS). Collectively, it helps to provide general management by communicating with MANO functions.

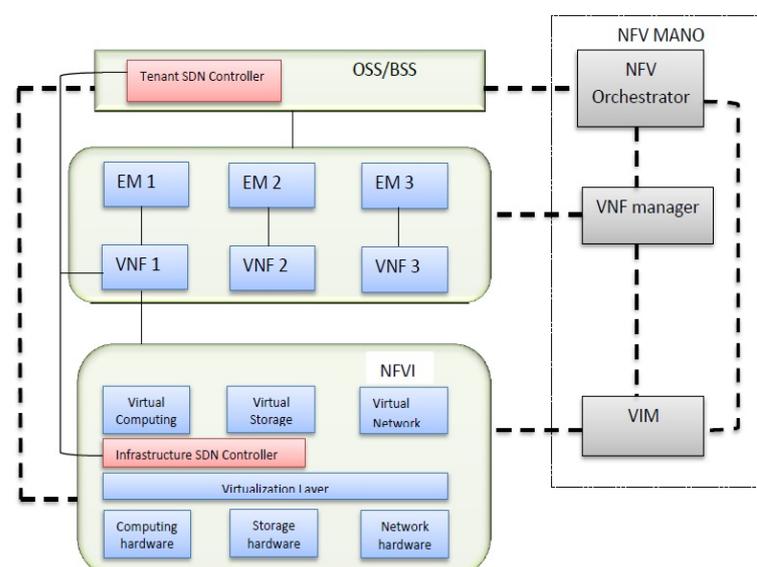


Figure 4. NFV reference architecture integrated with SDN controllers.

Two SDN controllers are mentioned in the architecture namely, the infrastructure SDN controller (ISDNC) and the tenant SDN controller (TSDNC).

- Infrastructure SDN Controller (ISDNC): Dynamically manages the infrastructure resources, which enables them to communicate with the VNFs [22].
- Tenant SDN Controller (TSDNC): Operate and dynamically manages the underlying data plane resources via interface.

These controllers need to communicate with each other to manage the network slices effectively.

4. Machine Learning and AI for Network Slicing Beyond 5G Systems

It has been observed that beyond 5G systems will be based on virtualization due to their ability of better optimization; however, these systems should support URLLC, mMTC, and eMBB. Development of beyond 5G systems not only has larger dimension but it also involves complexity due to the presence of the heterogeneous networks. In order to overcome these issues, a new intelligent architecture needs to be proposed that should offer flexibility and agility. A new intelligent architecture should have strong ability to learn, recognize, and adapt itself according to the application requirements. According to [23], artificial intelligence (AI)-enabled techniques will be applied to meet the above mentioned requirements for a 6G network. Machine learning (ML) and deep learning (DL) algorithms are envisioned as the most popular AI-enabled techniques for beyond 5G networks. It includes supervised learning, unsupervised learning, reinforcement learning, and neural networks. In the following, we provide a brief overview of supervised and unsupervised learning for network slicing.

4.1. Supervised Learning for Network Slicing

Predefined labeled data are used by supervised learning algorithms to train the system model in order to forecast possible outcomes. It can be described by following function:

$$g = h(X) \quad (1)$$

where X represents the predefined input variable and g represents the outcome. Hence for a new input data, the algorithm predicts the output by learning from the model. It can be further categorized into two techniques namely, regression and classification. Regression algorithms predicts and forecasts the real output values based on the input statistical elements by using the training data. Classification algorithms classify or distinguish the output based on discontinuous values such as categories. Supervised learning mechanism can be used in order to provide data integrity, localization, security, and tracking of objects [24,25]. Naive Bayes algorithm can be used for the localization [26]. By determining the possibilities, Bayesian algorithm can manage the missing or incomplete data sets, which is quite useful for localization.

4.1.1. Support Vector Machine (SVM)

SVM is one of the supervised learning classification algorithm, it helps to classify data by separating them by finding the decision boundary, called the hyperplane. Support vectors helps to define the hyperplane. Spectrum sensing can be challenging due to heterogeneity of the network, SVM can be helpful in this regard by classifying support vectors and intelligently creating the model to able to sense the spectrum efficiently [27]. Feature selections in network slicing can also be challenging, so features can be selected according to the application by using the SVM [28].

4.1.2. k-Nearest Neighbor (KNN)

KNN is a supervised learning algorithm that is widely used in regression and classification problems. It classifies the data by finding the distance between training and test data samples. In network slicing, KNN algorithm can be used to group same application services.

4.1.3. Deep Learning

Deep learning (DL) is a supervised classification learning technique, it perceives the functioning of the human nerve system in understanding the data representation methods and establishing the communication patterns between the input and output layer. DL can be applied to classify the applications according to the requirement such as mMTC, eMBB, and URLLC [29].

4.2. Unsupervised Learning for Network Slicing

In unsupervised learning, unlabeled data are provided and the purpose of the algorithm is to find the hidden or missing information from the unlabeled data pattern. Approach of unsupervised learning categorized into classification, anomaly detection, and dimension reduction. The main purpose of clustering is to form a class of data samples into various classes according to their analogy. Dimension reduction compress the data without losing useful information. For example, principal component analysis (PCA) is one of the algorithms for dimension reduction, which helps to compress the transmitted data. By finding the linear relationship between the sensors, the least-variance elements are neglected [30].

K-means Clustering

K-means algorithm groups specific data into various sub-groups (clusters) by randomly selecting the k places, which are considered initial central points for remaining clusters. Euclidean distance is calculated from the rest of the nodes from the chosen initial points, which results in the creation of new set of clusters. Then, new central points are calculated by finding the mean value. Again, find the Euclidean distance and stop the iteration if no more changes occur in the central points. It can be represented by the following expression:

$$J = \sum_{j=1}^k \sum_{i=1}^n \|x_i^{(j)} - c_j\|^2 \quad (2)$$

Here, c_j is a central point of cluster j , k and n represents the number of clusters and number of cases, respectively, and $\|x_i^{(j)} - c_j\|$ represents the Euclidean distance. In network slicing, k-means can be used to group applications into different clusters according to their similar features. In [31], authors discussed that the number of network slices that can be created by using k-means algorithm for each mobile network operator (MNO) depending on the KPI performances. In Section 7, we have proposed an approach that how clustering can be done in network slicing in order to manage the resources intelligently.

4.3. Reinforcement Learning

In reinforcement learning (RL) algorithm, every node (sensor) continuously learn and captures information to take best decisions by interacting with the environment in order to maximize the advantage by using its own experience. Markov decision and Q-learning are widely used RL algorithms. In Q-learning, a sensor node continuously updates its reward to take the best decision by interacting with the environment. In [32], authors used RL in order to achieve the handover efficiency intelligently.

5. ML for Mobility Prediction

Beyond 5G systems will be heterogeneous, ultra dense, and with large dimension, so the mobility prediction will likely be the most formidable case of beyond 5G systems due to the continuous handovers. Mobility prediction can be achieved intelligently by applying ML approaches. Handover efficiency can be difficult in managing the URLLC, eMBB, and MTC requirements. It may also become harder to track the locations due to frequent mobility. RL can be one of the ML approaches to achieve the handover efficiency intelligently by using a centralized RL agent [32]. UE measurement reports (reference signal received power (RSRP) and reference signal received quality (RSRQ)) are handled

by centralized RL agent, as it learns through a trial-and-error learning approach, which enhances its capability to take handover decisions based on RL framework.

In order to enhance the selection of target cell during handover (HO) decision, another neural network based ML approach was proposed in [33]. Two levels of a neural network model have been proposed to determine the probability of successful downloads and average download time. In the first level as input, UE measurement reports can be used to train neural network, and collects the previous information on the quality of experience (QoE) (download completed/not completed) as output. In the second level, those UE measurement reports with completed download are used for training.

After the completion of the training, trained neural networks have been used by the handover algorithm to estimate the expected QoE. Next generation network (NGN) should meet mobility requirements of vehicular networks in heterogeneous environment. There can be frequent handovers due to mobility of high speed vehicles. In [34], the Markov-chain technique is proposed to estimate the possible location of the UE and to reduce the numbers of handover.

They proposed two handover strategies, proactive and reactive handovers, and also analyzed the user's random movement across multiple cells. Mathematically representation is as follows:

$$p_n = [p] \times [P_{n-1}] = [p_{n-1}] \times [P] \quad (3)$$

where n illustrates the number of transitions, p_n is the probability of user's location, p and P_{n-1} are the initial and current transition matrix, respectively, p_{n-1} represents the initial distribution after n , P is denoted as the initial transition matrix. Even if the user finds an RSRP with strong signal strength in its current cell, reactive handover strategy helps to reduce the unnecessary handovers. User do not need to verify its current cell RSRP and the handover is carried out at any time in the proactive handover strategy.

Signaling overhead can be increased in micro cell environment due to consistent handovers (HOs). The HO scheme in the existing RAN architecture comprises redirecting control and data channels from one base station (BS) to another, which results in increasing the signaling overhead load. In [35], authors proposed an efficient and intelligent core-network (CN) radio resource control (RRC) signaling mechanism based on control/data separation architecture (CDSA) in order to reduce the HO signaling latency. The key idea behind the CDSA approach is to split the control signals from the data signals, providing only required signaling to support coverage [36]. CDSA separation architecture has been shown in Figure 5.

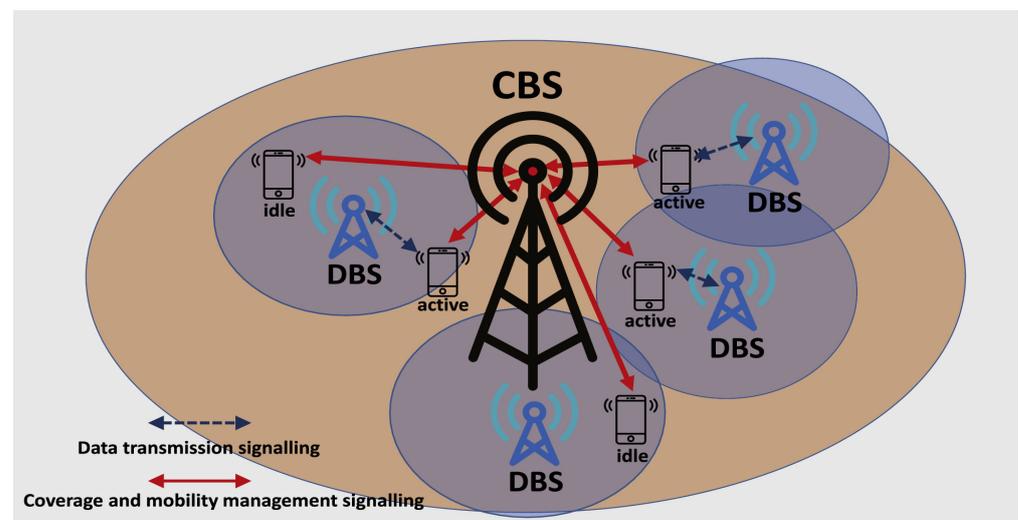


Figure 5. Control/data separation architecture.

Signaling related to data transmission is provided on request depending on the number of active users. Control base stations (CBS) are deployed to ensure the network

connectivity and within the CBS footprint, data base stations (DBS) are deployed to provide high data rates. The CDSA provides flexibility compared to the conventional RAN architecture in implementing mobility prediction for HO management schemes. In the proposed model [35], The Markov-chain model is used to predict the future DBS HOs. Prior to the fulfillment of HO criteria, RAN signaling related to HO is executed by using the prediction outcome, which results in less signaling overhead and latency. A graphical representation is shown in Figure 6. Results show that intra-CBS HO latency is reduced by 26% as compared to conventional HO. In contrast, the inter-CBS HO can generate higher signaling load as compared to single HO. Though, it is not practical to use the Markov-chain model to predict the future DBS HOs. There are two kinds of stochastic variables (state variable and output variable) set by the hidden Markov model (HMM) [37]. These variables relate the realistic environment of observed object, and a Markov chain is formed by the state sequence. Sometimes these measured values cannot map to practical environments due to inaccuracy of observation. According to our observation, signalling reduction gain can be achieved up to 60% in inter-CBS HO. Smart cell controller can be introduced within a CBS footprint.

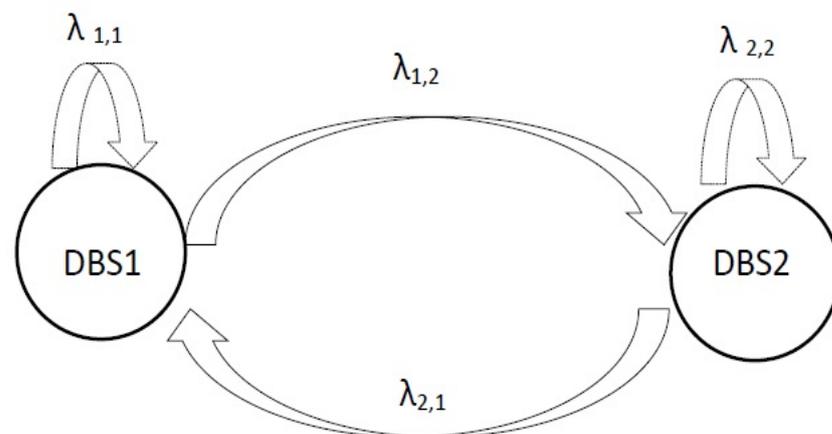


Figure 6. Discrete-time Markov Chain with 2 states (i.e., DBSs), only states 1, 2 are shown for readability.

In [38], authors proposed a scheme where a designated cell controller within a cluster handles a set of small cells and keeps the forwarding information within a localized mobility management entity (MME) for user equipments (UEs). Moreover, the data forwarding information is forwarded from the small cell controller to the serving gateway (SGW) by creating a data path tunnel. Although, this proposed scheme is not based on ML and CSDA architecture but results shows that signaling reduction gain has been achieved by up to 58.1%. Better results can be achieved by employing ML techniques and CSDA architecture as it has been discussed in [39].

6. ML for Resource Management

Network slicing will provide flexibility, heterogeneity, and scalability; on the other hand, network resources need to be highly manageable. Resource management in network slicing can be an issue and it can be addressed by applying the ML techniques. Network slicing can improve the resource efficiency by separating the infrastructure resources and network functions. In the past, network operators manages the network services by themselves but now due to network slicing they can lease them to multiple tenants. Tenants can deliver the network services to the end customers without the need of having their own infrastructure. In Figure 7, difference between existing unsliced and sliced network is shown. This new model has opened the new business opportunities, this service is called as “slice-as-a-service” (SlaaS) [40].

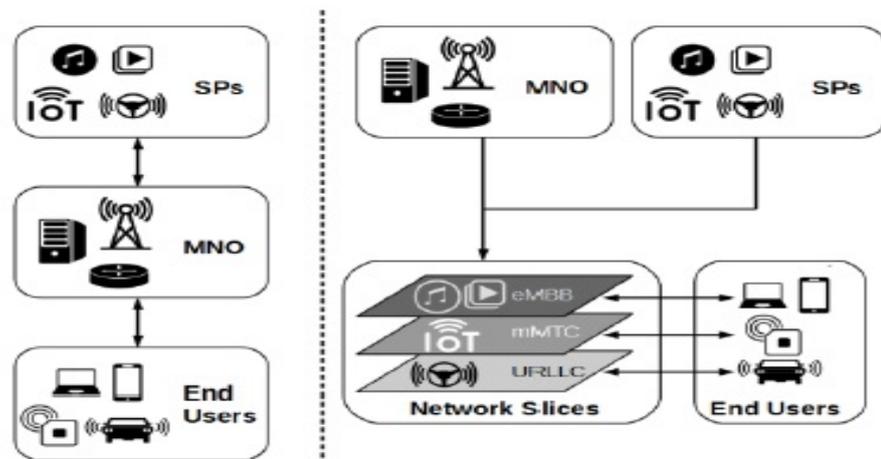


Figure 7. Unsliced network (left) vs. sliced network (Right).

SaaS can be categorized as application-as-a-service (AaaS), platform-as-a-service (PaaS) and infrastructure-as-a-service (IaaS). AaaS allows delivering the software-based services according to the end-customers demand. PaaS enables the cloud services to provide the application data. IaaS allows delivering the required network functions to manage the infrastructure resources. ML techniques can be applied to provide all these services. In [41], RL algorithm is proposed in order to enhance the network slicing policy. In [42], the authors presented the ANN-based network slicing resource management prediction model and grouped the model into three layers in order to estimate the requirements of service capacity.

7. RAN Slicing in 5G

We discussed the 5G architecture in Section 2.1, and it can be observed that it is comprised of actual physical and virtual resources of the RAN, networking nodes, cloud nodes, and associated links. Through virtualization, the flexibility of network resources can be achieved, which helps to create the core network slices. Spectrum resources can be shared, and flexible management of these resources can be achieved by slicing the RAN. Resource scheduling plays an essential role to meet the demands of RAN slicing and it can be very challenging due to different service requirements. In [43], authors present an intelligent resource scheduling scheme for 5G RAN slicing based on DL and RL. Both algorithms are used in collaborative manner in order to predict traffic pattern, which also results in improving the resources multiplexing gains. Network slicing partitions a physical network into different multiple logical networks according to their specific application requirements, and has become a key idea in addressing the heterogeneity requirements of future wireless networks.

In [44], the idea of multi-cell RAN was analyzed for the different applications that have been distributed among different tenants. It can be challenging because isolation, such as traffic isolation and the radio–electrical isolation, is required between different applications. In [45], authors have discussed the 5G RAN architecture, functions, and logical network architecture to support network slicing. Allocation of the radio resources is the main challenge in the implementation of network slicing. In RAN slicing, adaptation of conventional allocation strategies is an open research issue along with the development of new resource allocation schemes. Sharing the same radio resources can be challenging in RAN slicing for different heterogeneous applications. In [46], it has been discussed that in order to meet the heterogeneous service requirements, radio resources can be shared by tiling and puncturing in RAN slicing.

8. Performance Evaluation of Network Slicing

In network slicing, performance of the network can be measured in terms of latency, data rate, and packet loss rate; it can affect the QoS requirements needed for the different slices. In the URLLC slice, applications are considered to be delay sensitive, and ultra reliability is required to meet the QoS demands. According to [47], the committed information rate (CIR) ratios for specific data traffic depends on its characteristics. Different applications have different CIR ratios. Performance of a certain application can be affected by less CIR ratios. When there is a higher demand than the total bandwidth, a CIR ratio greater than 100% is required for the elastic and inelastic data traffic. In the case of URLLC, a CIR ratio of 120% is needed to meet the QoS requirement [47]. In the eMBB slice, a higher data rate, and in this case CIR ratio of 140%, is required to meet the 100% throughput percentage [47]; however, further study is needed to investigate how a higher number of users will impact the performance of the network slicing in terms of maximum bandwidth utilization.

9. Smart Port Use Case

Most of the seaports around the globe are equipped with fixed machinery and connected to the network via cables. Mostly, trade around the world is carried out via seaports, and most of the operations are heavily dependent on humans. Recently, COVID-19 has dealt a major blow globally, resulting in shipping costs reaching an all-time high. A reliable network is required to trace and locate the containers, and to check their custom clearance reports and temperature. The existing 4G network cannot meet these requirements due to its limitations. We present the smart port use case, where network slicing is deployed to ensure safe and efficient processes. NS based 5G trial test-bed has been implemented in port of Hamburg [7]. Results showed that in order to manage the slices, all network entities need to be involved. A major challenge in NS is to manage the resources more efficiently [7]. We propose the network slicing model based on ML techniques in order to manage the resources more efficiently. 5G network slicing can be deployed in the following smart port scenarios:

- Intelligent automation (IA).
- Virtual reality (VR).
- Smart guided vehicle (SGV).

9.1. Intelligent Automation (IA) (mMTC)

Intelligent automation (IA) concerns the intelligent controlling and monitoring of the end devices with the combination of physical entities and application services. The efficiency of the port can be achieved by improving the automation, which will also help to communicate between different parties present on the port. Smart ships are equipped with sensors in order to monitor the position of the ships and environmental data in real time. Containers can also be equipped with sensors to monitor the temperature and to trace their location. These sensors need to be connected to 5G base station via a 5G customer premise equipment (5G-CPE) wireless terminal. A data analytics mechanism is required for clustering and data aggregation due to the higher volume of data generated by sensor networks placed in the sea port. ML algorithms can be used in this regard.

9.2. Virtual Reality (VR) (eMBB)

It allows the field operator to work efficiently by allowing them to monitor the port operations in real time via video surveillance, and also enables them to visit the port in a VR environment and to receive remote assistance. It requires higher throughput and a reliable connection, which can be achieved by deploying 5G network slicing.

9.3. Smart Guided Vehicles (SGV) (URLLC)

Conventional trucks must be replaced with smart guided vehicles (SGV) in order to improve the safety and port efficiency due to their ability to consume less resources. SGV needs to be equipped with sensors and high resolution cameras. Existing 4G network

cannot meet the low latency requirements of SGV, but it can be achieved by implementing the 5G network.

9.4. Design Overview

Integration of ML techniques with SDN/NFV architecture will help to manage the resources more efficiently. Handover behavior can also be predicted by using ML algorithms. In our proposed model, we use ML techniques in order to slice the network with our above mentioned requirements. Identification of application in earlier stages is considered to be an important step to distinguish between mobile traffic, this approach can be useful to slice the network. In this regard, the SVM algorithm can be used to extract the data, which helps to classify the mobile traffic. Then K-means algorithm is used to group applications into different clusters according to similar traffic patterns. We categorize the above-mentioned requirements into three different services: application-as-a-service (AaaS), platform-as-a-service (PaaS), and infrastructure-as-a-service (IaaS). Every slice will be grouped into these three services, which will help to manage the traffic load evenly. For example, if a field operator wants to visit the port in a VR environment, they can achieve that by using the AaaS platform. In contrast, the IaaS platform will provide administration privileges to the port authority of this particular slice (eMBB). Figure 8 describes the architectural design of our model.

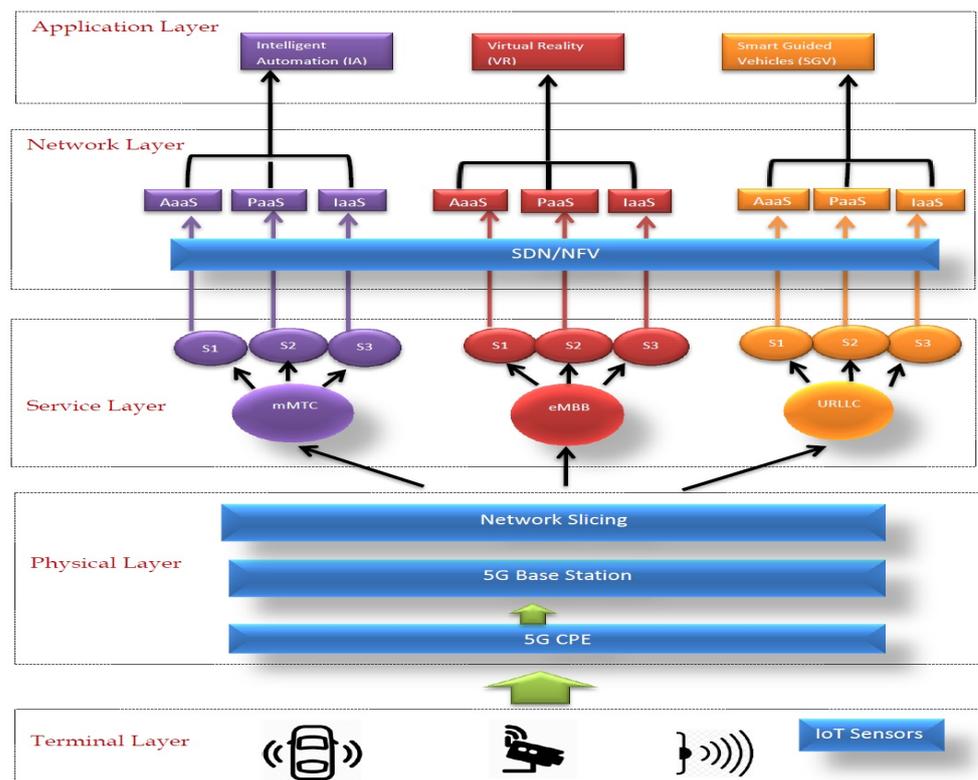


Figure 8. Proposed smart port network slicing architecture.

We grouped the model into five layers: terminal layer, physical layer, service layer, network layer, and application layer. The terminal layer is used to collect the data from the sensors and then these data are passed onto the physical layer via 5G-CPE wireless terminal. IoT devices need to be 5G-enabled in order to connect these IoT devices to a 5G base station. Thus, due to the resource limitations of IoT devices we can use 5G-CPE [48].

Data are analyzed and processed at the physical layer. The physical network is grouped into three slices (i.e., mMTC, eMBB, and URLLC). At the service layer, ML techniques are used and every sliced network is grouped into three more slices according to the services: AaaS, PaaS, and IaaS. The network layer, also called the core network,

offers various services, such as providing the control and data plane functions and storage services. the client–server environment is provided at the application layer.

9.5. Methodology

The approach consists of following steps: data are processed; features are extracted; the clustering algorithm is used to group applications into different clusters according to their similar application services.

9.5.1. Data Processing

Data mining needs to be done at the physical layer due to the higher volume of data generated by the sensors. Data mining is the process of extracting useful data patterns from a real-time continuous dynamic data stream. It is hard to store all the generated data from the sensors so it needs to be processed earlier. Useful data can be extracted by using ML algorithm. There are various data mining techniques based on ML algorithms and can be applied according to the application requirements. In [49], an ML-based data mining framework of wireless sensor network is discussed.

9.5.2. Feature Extraction

Data can be extracted based on the size of packet, TCP/UDP, or by their IP. Then, a classification algorithm, such as SVM or DL, can be applied to classify the applications such as mMTC, eMBB, and URLLC. New incoming traffic can be grouped according to the application requirements.

9.5.3. Clustering

Clustering algorithm K-means is used to group applications into different clusters according to their similar features, such as AaaS, PaaS, and IaaS. As it is an unsupervised algorithm, it identifies the patterns by using unlabeled data and group them into k number of clusters according to their similar traffic pattern. Let us suppose that if we use the size of the packet as input feature, after applying K-means algorithm, every cluster will be grouped according to bandwidth.

9.5.4. Core Network

The core network acts as a bridge between data transmission and applications, providing storage, control plane, and data plane functions. After the logical slicing according to the services (AaaS, PaaS, and IaaS), information is provided to applications such as intelligent automation services (mMTC), virtual reality (VR) services (eMBB), and SGV services (URLLC). The performance of our proposed model will be evaluated to manage the resources efficiently.

10. Conclusions and Future Work

In this paper, we presented a comprehensive review of network slicing using enabling technologies, SDN and NFV, with the main focus on its architecture design and ML approach. We then discussed how ML techniques can be utilized to intelligently predict mobility and balance the load in the network. We also presented the smart port use case scenario, where ML approach was proposed to manage the resources more effectively and efficiently. There are several challenges, such as distribution of network functions, RAN Virtualization, security, and privacy issues. These issues and challenges need to be addressed in order to provide a good model for beyond 5G systems. In the future, blockchain technology can be integrated with network slicing to meet the QoS requirements of future networks. Blockchain offers better storage and management options with low latency compared to conventional database platforms.

Author Contributions: R.M.S. proposed the concept and carried out the theoretical analysis and review. O.O., Y.S. and M.A.I. gave valuable suggestions on the structure of the manuscript and also assisted proofreading. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the U.K. Engineering and Physical Sciences Research Council under Grant EP/S02476X/1.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Tomtsis, D.; Kokkonis, G.; Kontogiannis, S. Evaluating existing wireless technologies for IoT data transferring. In Proceedings of the 2017 South Eastern European Design Automation, Computer Engineering, Computer Networks and Social Media Conference (SEEDA-CECNSM), Kastoria, Greece, 23–25 September 2017; pp. 1–4. [CrossRef]
2. Alreshidi, A.; Ahmad, A. Architecting. Software for the Internet of Thing Based Systems. *Future Internet* **2019**, *10*, 153. [CrossRef]
3. Network Slicing Use Case Requirements. Available online: <https://www.gsma.com/futurenetworks/wp-content/uploads/2018/07/Network-Slicing-Use-Case-Requirements-fixed.pdf> (accessed on 12 December 2020).
4. 3GPP, TR 23.711. Enhancements of Dedicated Core Networks Selection Mechanism (Release 14). 2016. Available online: <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=2962> (accessed on 26 September 2016).
5. Kaloxylou, A.; Alexandros. A survey and an analysis of network slicing in 5G networks. *IEEE Commun. Stand. Mag.* **2018**, *2*, 60–65. [CrossRef]
6. View on 5G Architecture—5G PPP Architecture Working Group. Available online: <https://www.trust-it-services.com/sites/default/files/View%20on%205G%20Architecture%20-%205G%20PPP%20Architecture%20Working%20Group.pdf> (accessed on 19 June 2019).
7. Rost, P.; Breitbach, M.; Roreger, H.; Erman, B.; Mannweiler, C.; Miller, R.; Viering, I. Customized Industrial Networks: Network Slicing Trial at Hamburg Seaport. *IEEE Wirel. Commun.* **2018**, *25*, 48–55. [CrossRef]
8. Sgambelluri, A.; Tusa, F.; Gharbaoui, M.; Maini, E.; Toka, L.; Perez, J.M.; Paolucci, F.; Martini, B.; Poe, W.Y.; Hernandez, J.M.; et al. Orchestration of network services across multiple operators: The 5G exchange prototype. In Proceedings of the 2017 European Conference on Networks and Communications (EuCNC), Oulu, Finland, 12–15 June 2017; pp. 1–5.
9. Gramaglia, M.; Digon, I.; Friderikos, V.; von Hugo, D.; Mannweiler, C.; Puente, M.A.; Samdanis, K.; Sayadi, B. Flexible connectivity and QoE/QoS management for 5G Networks: The 5G NORMA view. In Proceedings of the 2016 IEEE International Conference on Communications Workshops (ICC), Kuala Lumpur, Malaysia, 23–27 May 2016; pp. 373–379.
10. Rizou, S.; Athanasoulis, P.; Andriani, P.; Iadanza, F.; Carrozzo, G.; Breitgand, D.; Weit, A.; Griffin, D.; Jimenez, D.; Acar, U.; et al. A service platform architecture enabling programmable edge-to-cloud virtualization for the 5G Media industry. In Proceedings of the 2018 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB), Valencia, Spain, 6–8 June 2018; pp. 1–6.
11. Tzanakaki, A.; Anastasopoulos, M.; Simeonidou, D.; Berberana, I.; Syrivelis, D.; Korakis, T.; Flegkas, P.; Mur, D.C.; Demirkol, I.; Gutiérrez, J.; et al. 5G infrastructures supporting end-user and operational services: The 5G-XHaul architectural perspective. In Proceedings of the 2016 IEEE International Conference on Communications Workshops (ICC), Kuala Lumpur, Malaysia, 23–27 May 2016; pp. 57–62.
12. NGMN Alliance. Description of Network Slicing Concept. White Paper. 2016. Available online: <https://www.ngmn.org/publications/description-of-network-slicing-concept.html> (accessed on 2 January 2021).
13. NGMN 5G Project Requirements and Architecture—Work Stream E2E Architecture (Version 1.0). Available online: https://www.ngmn.org/wp-content/uploads/160113_NGMN_Network_Slicing_v1_0.pdf (accessed on 13 January 2016).
14. McKeown, N.; Balakrishnan, H.; Peterson, L.L.; Shenker, S.; Turner, J.S.; Anderson, T.; Parulkar, G.; Rexford, J. OpenFlow: enabling innovation in campus networks. *ACM SIGCOMM Comput. Commun. Rev.* **2008**, *38*, 69–74. [CrossRef]
15. ONF TR-526, Applying SDN Architecture to 5G Slicing. April 2016. Available online: https://opennetworking.org/wp-content/uploads/2014/10/Applying_SDN_Architecture_to_5G_Slicing_TR-526.pdf (accessed on 2 January 2021).
16. ONF TR-521, SDN Architecture. February 2016. Available online: https://opennetworking.org/wp-content/uploads/2014/10/TR-521_SDN_Architecture_issue_1.1.pdf (accessed on 2 January 2021).
17. Zhu, S.; Bi, J.; Sun, C.; Wu, C.; Hu, H. Sdpa: Enhancing stateful forwarding for software-defined networking. In Proceedings of the 2015 IEEE 23rd International Conference on Network Protocols (ICNP), San Francisco, CA, USA, 10–13 November 2015; pp. 323–333.
18. Ordonez-Lucena, J.; Ameigeiras, P.; Lopez, D.; Ramos-Munoz, J.; Lorca, J.; Figueira, J. Network slicing for 5G with SDN/NFV: Concepts, architectures, and challenges. *IEEE Commun. Mag.* **2017**, *55*, 80–87. [CrossRef]
19. ETSI, Network Functions Virtualisation (NFV); Management and Orchestration 1 (2014): V1. Available online: https://www.etsi.org/deliver/etsi_gs/nfv-man/001_099/001/01.01.01_60/gs_nfv-man001v010101p.pdf (accessed on 5 January 2021).
20. Taylor, M. The Myth of the Carrier-Grade Cloud. 2015. Available online: <http://www.metaswitch.com/the-switch/the-myth-of-the-carrier-grade-cloud> (accessed on 5 January 2021).
21. ETSI GS NFV-EVE 005. Network Functions Virtualisation (NFV); Ecosystem; Report on SDN Usage in NFV Architectural Framework, v. 1.1.1, December 2015. Available online: https://www.etsi.org/deliver/etsi_gs/NFV-EVE/001_099/005/01.01.01_60/gs_nfv-eve005v010101p.pdf (accessed on 7 January 2021).

22. ETSI GS NFV-INF 001. Network Functions Virtualisation (NFV); Infrastructure Overview, v. 1.1.1, January 2015. Available Online: https://www.etsi.org/deliver/etsi_gs/NFV-INF/001_099/001/01.01.01_60/gs_NFV-INF001v010101p.pdf (accessed on 7 January 2021).
23. Letaief, K.B.; Chen, W.; Shi, Y.; Zhang, J.; Zhang, Y.J.A. The roadmap to 6G: AI empowered wireless networks. *IEEE Commun. Mag.* **2019**, *57*, 84–90. [[CrossRef](#)]
24. Janakiram, D.; Reddy, V.; Kumar, P.A. Outlier detection in wireless sensor networks using Bayesian belief networks. In Proceedings of the International Conference on Communication System Software and Middleware, New Delhi, India, 8–12 January 2006; pp. 1–6.
25. Moustapha, A.; Selmic, R. Wireless sensor network modeling using modified recurrent neural networks: Application to fault detection. *IEEE Trans. Instrum. Meas.* **2008**, *57*, 981–988. [[CrossRef](#)]
26. Morelande, M.; Moran, B.; Brazil, M. Bayesian node localisation in wireless sensor networks. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Las Vegas, NV, USA, 31 March–4 April 2008; pp. 2545–2548.
27. Azmat, F.; Chen, Y.; Stocks, N. Analysis of Spectrum Occupancy Using Machine Learning Algorithms. *IEEE Trans. Veh. Technol.* **2016**, *65*, 6853–6860. [[CrossRef](#)]
28. Singh, S.K.; Salim, M.M.; Cha, J.; Pan, Y.; Park, J.H. Machine Learning-Based Network Sub-Slicing Framework in a Sustainable 5G Environment. *Sustainability* **2020**, *12*, 6250. [[CrossRef](#)]
29. Rezwani, S.; Choi, W. Priority-Based Joint Resource Allocation With Deep Q-Learning for Heterogeneous NOMA Systems. *IEEE Access* **2021**, *9*, 41468–41481. [[CrossRef](#)]
30. Alsheikh, M.A.; Lin, S.; Niyato, D.; Tan, H. Machine Learning in Wireless Sensor Networks: Algorithms, Strategies, and Applications. *IEEE Commun. Surv. Tutor.* **2014**, *16*, 1996–2018. [[CrossRef](#)]
31. Kalyoncu, F.; Zeydan, E.; Yigit, I.O. A data analysis methodology for obtaining network slices towards 5g cellular networks. In Proceedings of the 2018 IEEE 87th Vehicular Technology Conference (VTC Spring), Porto, Portugal, 3–6 June 2018; pp. 1–5.
32. Yajnanarayana; Vijaya; Rydén, H.; Héviz, L. 5G handover using reinforcement learning. In Proceedings of the 2020 IEEE 3rd 5G World Forum (5GWF), Bangalore, India, 10–12 September 2020.
33. Ali, Z.; Baldo, N.; Mangués-Bafalluy, J.; Giupponi, L. Machine learning based handover management for improved QoE in LTE. In Proceedings of the NOMS 2016–2016 IEEE/IFIP Network Operations and Management Symposium, Istanbul, Turkey, 25–29 April 2016; pp. 794–798.
34. Ulvan, A.; Bestak, R.; Ulvan, M. Handover procedure and decision strategy in LTE-based femtocell network. *Telecommun. Syst.* **2013**, *52*, 2733–2748. [[CrossRef](#)]
35. Mohamed, A.; Onireti, O.; Imran, M.A.; Imran, A.; Tafazolli, R. Predictive and Core-Network Efficient RRC Signalling for Active State Handover in RANs With Control/Data Separation. *IEEE Trans. Wirel. Commun.* **2017**, *16*, 1423–1436. [[CrossRef](#)]
36. Capone, A.; Dos Santos, A.F.; Filippini, I.; Gloss, B. Looking beyond green cellular networks. In Proceedings of the 2012 9th Annual Conference on Wireless On-Demand Network Systems and Services (WONS), Courmayeur, Italy, 9–11 January 2012; pp. 127–130.
37. Akoush, S.; Sameh, A. Mobile user movement prediction using bayesian learning for neural networks. In Proceedings of the 2007 International Conference on Wireless Communications and Mobile Computing, Honolulu, HI, USA, 12–16 August 2007; pp. 191–196.
38. Thainesh, J.S.; Wang, N.; Tafazolli, R. Reduction of core network signalling overhead in cluster based LTE small cell networks. In Proceedings of the 2015 IEEE 20th International Workshop on Computer Aided Modelling and Design of Communication Links and Networks (CAMAD), Guildford, UK, 7–9 September 2015; pp. 226–230. [[CrossRef](#)]
39. Mohamed, A.; Onireti, O.; Hoseinitabatabaei, S.A.; Imran, M.; Imran, A.; Tafazolli, R. Mobility prediction for handover management in cellular networks with control/data separation. In Proceedings of the 2015 IEEE International Conference on Communications (ICC), London, UK, 8–12 June 2015; pp. 3939–3944. [[CrossRef](#)]
40. Sciancalepore; Vincenzo; Cirillo, F.; Xavier C.P. Slice as a service (SlaaS) optimal IoT slice resources orchestration. In Proceedings of the GLOBECOM 2017–2017 IEEE Global Communications Conference, Singapore, 4–8 December 2017.
41. Bega, D.; Gramaglia, M.; Banchs, A.; Sciancalepore, V.; Samdanis, K.; Costa-Perez, X. Optimising 5G infrastructure markets: The business of network slicing. In Proceedings of the IEEE INFOCOM 2017—IEEE Conference on Computer Communications, Atlanta, GA, USA, 1–4 May 2017; pp. 1–9.
42. Bega, D.; Gramaglia, M.; Fiore, M.; Banchs, A.; Costa-Perez, X. DeepCog: Cognitive network management in sliced 5G networks with deep learning. In Proceedings of the IEEE INFOCOM 2019—IEEE Conference on Computer Communications, Paris, France, 29 April–2 May 2019; pp. 280–288.
43. Yan, M.; Feng, G.; Zhou, J.; Sun, Y.; Liang, Y.C. Intelligent resource scheduling for 5G radio access network slicing. *IEEE Trans. Veh. Technol.* **2019**, *68*, 7691–7703. [[CrossRef](#)]
44. Sallent, O.; Perez-Romero, J.; Ferrus, R.; Agusti, R. On radio access network slicing from a radio resource management perspective. *IEEE Wirel. Commun.* **2017**, *24*, 166–174. [[CrossRef](#)]
45. Marsch, P.; Da Silva, I.; Bulakci, O.; Tesanovic, M.; El Ayoubi, S.E.; Rosowski, T.; Kalokylos, A.; Boldi, M. 5G radio access network architecture: Design guidelines and key considerations. *IEEE Commun. Mag.* **2016**, *54*, 24–32. [[CrossRef](#)]
46. Elayoubi, S.E.; Jemaa, S.B.; Altman, Z.; Galindo-Serrano, A. 5G RAN slicing for verticals: Enablers and challenges. *IEEE Commun. Mag.* **2019**, *57*, 28–34. [[CrossRef](#)]

-
47. Lin, Y.B.; Tseng, C.C.; Wang, M.H. Effects of Transport Network Slicing on 5G Applications. *Future Internet* **2021**, *13*, 69. [[CrossRef](#)]
 48. Yang, J.; Feng, L.; Jin, T.; Liang, H.; Xu, F.; Tian, L. Enabling Sector Scheduling for 5G-CPE Dense Networks. *Secur. Commun. Netw.* **2020**, *2020*, 6684681. [[CrossRef](#)]
 49. Azhar, M.; Ke, S.; Shaheen, K.; Mi, X. Data mining techniques for wireless sensor networks: A survey. *Int. J. Distrib. Sens. Netw.* **2013**, *9*, 1–24.