NETWORK VIRTUALIZATION IN NEXT-GENERATION CELLULAR NETWORKS:

A SPECTRUM POOLING APPROACH

by

Ayman AbdelHamid

Bachelor of Science in Electrical Engineering,

Mansoura University, 2010

Master of Science in Wireless Communications and Information Technology,

Nile University, 2013

Submitted to the Graduate Faculty of

the School of Computing and Information in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2019

UNIVERSITY OF PITTSBURGH

SCHOOL OF COMPUTING AND INFORMATION

This dissertation was presented

by

Ayman AbdelHamid

It was defended on

December 19, 2019

and approved by

Dr. Prashant Krishnamurthy, Professor, Department of Informatics and Networked Systems, School of Computing and Information

Dr. David Tipper, Professor, Department of Informatics and Networked Systems, School of Computing and Information

Dr. Martin Weiss, Professor, Department of Informatics and Networked Systems, School of Computing and Information

Dr. Yi Qian, Professor, Department of Electrical and Computer Engineering, University of Nebraska

Dissertation Director: Dr. Prashant Krishnamurthy, Professor, Department of Informatics and Networked Systems, School of Computing and Information

NETWORK VIRTUALIZATION IN NEXT-GENERATION CELLULAR NETWORKS:
A SPECTRUM POOLING APPROACH

Ayman AbdelHamid, PhD

University of Pittsburgh, 2019

The hardship of expanding the cellular network market results from the tremendous high cost of mobile infrastructure, i.e. the capital expenditures (CAPEX) and the operational expenditures (OPEX). Spectrum Sharing is one of the proposed solution for the high-cost of scalability of cellular networks. However, most of the proposed spectrum pooling frameworks in the literature are mostly approached from a technical view besides there are no good cost models based on real data sets for quantifying the circumstances under which sharing the spectrum and network resources would be beneficial to mobile operators.

In this thesis, by studying different sharing scenarios in a fiber-based backhaul mobile network, we assess the incentives for service providers (SPs) to share spectrum/infrastructure in different cellular market areas/economic areas (CMA/BEAs) with different population density, allocated bandwidth (BW), spectrum bid values and considering different network topologies. Moreover, we look at the technical problem of sharing the spectrum between two SPs sharing the same base-station (BS), yet they have different traffic demand as well as different QoS constraints. We design a resource allocation scheme to provision real-time (RT), non-real-time (NRT) as well as Ultra-reliable Low Latency Communications (URLLC) traffic in a single shared BS scenario such that SPs achieve isolation, fairness and enforce their QoS constraints.

Finally, we exploit spectrum pooling to develop an approach for dynamically reconfiguring the base stations that survive a disaster and are powered by a microgrid to form a multi-hop mesh network in order to provide local cellular service.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1.0 INTRODUCTION

Within the last decade, there has been an enormous increase in the demands of mobile users due to the varieties of services that service providers (SPs) have introduced such as video streaming and massive Machine Type Communications (mMTC) [1]. In Ericsson's report (June 2017), voice over LTE (VoLTE) subscriptions were expected to exceed 540 million by the end of 2017 (more than 60 % of the voice calls on LTE smartphones) [2]. Moreover, Next-Generation Mobile Networks (NGMN) are being designed to connect a variety of industries, e.g., manufacturing and processing, intelligent transportation, smart grids, e-health, etc. Such different environments and applications will bring the challenge of speed, latency and heterogeneity. For example, Enhanced Mobile BroadBand (eMBB) service requires a very high throughput ($\geq$ 100 Mbps) to provision high data rate requests such as video gaming and tactile communications, whereas mMTC including IoT traffic encompases tremendous numbers of connections and ultra-reliable low latency crtitical communications (URLLC) flows are delay sensitive (user-plane delay ≤ 0.5 msec) and require ultra-reliable links, i.e. link reliability ≥ 0.999 [3].

Due to the diversity of applications and associated network service requirements (e.g. QoS, speed, latency, reliability and security) for each application, **the capital expenditures (CAPEX)** and **operational expenditures (OPEX)** of developing current mobile networks to cope with the requirements of NGMN are expected to be significantly high [4]. Thereby, the roll-out of mobile networks will require high sunk investments and the need to recover those by charging the user heavily for accessing mobile services [5]. In addition to CAPEX and OPEX, NGMNs need more spectrum resources to achieve the aforementioned stringent QoS requirements. In short, CAPEX, OPEX and the spectrum availability are the main issues that arise for upgrading LTE networks to 5G networks. **Wireless Network Virtualization (WNV)** has been extensively proposed in the literature as a key solution to the challenges of both high CAPEX and OPEX as well as spectrum

utilization challenges accompanied with upgrading to the NGMN [4]. Virtualizing the NGMN infrastructure saves on CAPEX and OPEX whereas virtualizing the physical spectrum resources (i.e., spectrum pooling) enhances spectrum utilization by sharing the spectrum efficiently between different mobile network operators (MNOs). In this thesis, we address some of the challenges of spectrum pooling as an approach of WNV. In the following section, we define virtualization, we discuss its challenges and we show how spectrum pooling can be exploited to solve the aforementioned challenges in NGMN. In this thesis, we will use the terms service provider (SP)/virtual entity (VE)/mobile network operator (MNO) interchangeably, without regard to the business model interpretation of each of the terms.

1.1 WIRELESS NETWORK VIRTUALIZATION (WNV)

In wired networks, virtualization has occurred for decades. Two early forms of virtualization are virtual private networks (VPNs) over wide area networks (WANs) and virtual local area networks (VLANs) in enterprise networks [6, 7]. The Xen project [8] is a primary example of virtualization, an x86 virtual machine monitor which allowed multiple commodity operating systems to share conventional hardware in a safe and resource managed fashion. Challenges of wired virtualization have been extensively studied in the literature including flexibility, heterogeneity, manageability, isolation and legacy support [9]. Nevertheless, the challenges of virtualization are of a different nature when applied to wireless networks [10]. For example, unlike wired networks where bandwidth resource abstraction and isolation can be done on hardware (e.g., port and link) basis, in wireless environment, radio resource abstraction and isolation are not straightforward, due to the inherent broadcast nature of wireless communications and stochastic fluctuation of wireless channel quality. Thus, WNV can be categorized into infrastructure virtualization (InV) and radio resource virtualization (RRV), which we interchangeably use with spectrum pooling.

Infrastructure virtualization (InV) is defined as decoupling the roles of a traditional network operator into two independent entities, infrastructure providers (InPs) who manage the physical substrate/spectrum and service providers (SPs) or Virtual Entities (VEs) who access different resources of the InP through a virtual network (VN) [11]. InV enables sharing the same infrastruc-

ture, i.e., core network, radio access network (RAN), backhaul or all of them, between different mobile network operators (MNOs), i.e., SPs, and hence, reduces the cost of establishing or expanding the backhaul/radio network. On the other side, RRV, namely spectrum pooling, is sharing radio resources, i.e., spectrum resource blocks (RBs), between different MNOs to utilize the spectrum more efficiently and also reduce the CAPEX and OPEX of expanding and upgrading current cellular networks to attain the requirements of NGMN.

Spectrum pooling as a network virtualization technique has many economical and technical challenges. First, most of the proposed spectrum pooling frameworks in the literature are solely approached from the technical side and do not explore the economic side, i.e., costs and revenues of different sharing scenarios, in order to assess the incentives to sharing entities. To the best of our knowledge, there are no good cost models based on real data sets for quantifying *the circumstances under which sharing of spectrum and network resources would be beneficial to operators*. On the technical side, pooling the spectrum between different SPs with different traffic demands and use cases is a big challenge. More precisely, allocating resource blocks (RBs) from the spectrum pool to the sharing MNOs such that both isolation, fairness and the target QoS for various traffic types is attained, is a non-trivial problem as we depict in the following chapters. Finally, it is not clear yet, how a greenfield design of a spectrum pooled network can be accomplished.

In short, this thesis is divided into three main parts: (i) understanding the economic perspective of pooling the spectrum resources between different MNOs and comparing the incentive of spectrum pooling with the incentive of infrastructure sharing, (ii) analyzing the technical side of spectrum pooling, i.e., developing an efficient spectrum sharing technique to allocate physical resources in a virtualized base station such that heterogeneous traffic demands are provisioned as well as supporting isolation and fairness between multiple SPs, and (iii) applying the spectrum pooling approach as a means to recover a cellular network from disruption during disasters, such as earthquakes, due to the disconnection between the core network and the RAN.

1.2 THESIS OBJECTIVE

The objective of this thesis is to construct an economic and efficient solution to the establishment of greenfield 5G networks by tackling the aforementioned challenges of spectrum pooling. By studying different sharing scenarios in a fiber-based backhaul network, we aim at assessing the incentives for SPs to share spectrum/infrastructure in different cellular market areas/economic areas (CMA/BEAs) and hence break the market between new roll-outs. Moreover, we look microscopically at the technical side of sharing the spectrum between two SPs sharing the same BS, yet they have different traffic demands as well as different QoS constraints. We developed two resource allocation schemes to provision real-time (RT), non-real-time (NRT) as well as Ultra-reliable Low Latency Communications (URLLC) traffic when the spectrum resource blocks (RBs) are shared between two SPs such that our schemes achieve isolation, fairness and satisfy QoS constraints. Finally, we propose exploiting spectrum pooling between different eNBs that belong to different MNOs in a disaster disrupted CMA/BEA to recover the cellular communications by forming a wireless mesh network between the survivor BSs.

1.3 THESIS STATEMENT

To get more specific, this thesis aims to tackle the following spectrum pooling scenarios and answer the accompanied design questions:

Question 1. *Given a certain Cellular Market Area/Economic Area (CMA/EA), a greenfield LTE network, a fiber-based backhaul network, two service providers (SPs) who share the spectrum, the backhaul infrastructure or both of them at the same time in this CMA/EA:* ***Which sharing scenario has more incentives for MNOs to adopt: (i) to share their own spectrum resources, (ii) to share their infrastructure resources or (iii) sharing both infrastructure and spectrum (i.e., full sharing)? Moreover, how does this incentive change with the type of the backhaul network topology, the operating frequency band, the number of BSs deployed to cover this area as well as the population density in this area?***

Question 2. *After empirically showing the conditions for economic incentives to pool the spectrum, we expose the technical part of pooling the available spectrum at a single BS between two SPs. Sharing the spectrum resource blocks (RBs) between different SPs with different traffic demands and new 5G use cases such as ultra-reliable low latency communications (URLLC) raises up the design question:* ***How can we design an efficient resource allocation scheme that allocates RBs among SPs such that the resource manager (hypervisor) would assign resources according to each virtual network's demand, i.e., traffic demand and classes of service, and the quality of channels in each virtual network, while preserving isolation between different SP/VEs, i.e., achieving high utility for one SP/VE should not affect other entities sharing the same pool of resources?***

Question 3. *To recover the cellular communications network from disruption due to natural disasters by pooling the spectrum between the survivor eNBs and establishing a wireless mesh network between the eNBs and without using any alternative technology:* ***How can we reconfigure the cellular network at the disrupted CMA with new optimal power allocations, optimal operating frequencies and optimal BW assignments per eNodeB such that the survivor eNodeB mesh network achieves the maximum average cell throughput, the target QoS constraints for each type of service and meets the transmission power limits?***

The layout of this thesis is described as follows. In Chapter 2, we present some of the recent literature review in the topics of Infrastructure Virtualization (InV) and Radio Resource Virtualization (RRV) from both the technical and economical perspective. In Chapter 3, we discuss the thesis main statement, the model assumptions and the thesis contributions as well as future work. Afterwards, in Chapter 4 we develop a cost model to assess the deployment cost of a greenfield LTE network to obtain insights into the costs and benefits of sharing the network infrastructure and radio spectrum. In Chapter 5, we propose two resource allocation schemes to enhance spectrum sharing between heterogeneous types of traffic in 5G networks in both 700 and 2100 MHz bands. Last but not least, in Chapter 6, we propose a disaster recovery algorithm based on the results concluded in Chapter 5 to recover a heterogeneous wireless network via multihop wireless mesh network using different frequency carriers for different traffic types. Eventually, Chapter 7 concludes this thesis.

2.0 BACKGROUND AND LITERATURE REVIEW

WNV has been addressed in the literature in two aspects: technically and economically. In this chapter we provide a succinct literature review that is relevant to these aspects.

2.1 THE ECONOMIC APPROACH OF WNV

Different business models for spectrum sharing were abstractly discussed with few details in [12] including sharing between equal primary devices, sharing between primary and secondary, sharing among equal secondaries and sharing among equal regional infrastructures [1]. Also, in [13] and [14] the authors discussed various scenarios of infrastructure sharing for MNOs in cellular networks, i.e. core network sharing, passive RAN sharing, active RAN sharing and etc, and abstractly specify the primal benefits and drawbacks of each, however the treatment is abstract and there are no technical simulations. Also, in [15], different sharing scenarios are investigated such as: capacity sharing, spectrum sharing and Mobile Network Operator sharing in collocated and non-collocated base stations models using a system-level 4G LTE simulation testbed. Recently, the tradeoff between the probability of coverage and the data rate was investigated in infrastructure, spectrum and full, i.e. infrastructure + spectrum, sharing scenarios [16]. In [16], J.Kibilda et al. use stochastic geometry to analyze different scenarios of sharing for independent, clustered and co-located infrastructure. Stochastic geometry is used to analyze cellular networks based on the probabilistic spatial distribution of BSs rather than on a specific network realization [17]. Abstracting each BS location to a point in Euclidean space, stochastic geometry models the probabilistic BS locations

[1] A regional infrastructure is a mobile infrastructure that covers a certain cellular market area or economic area based on the demography and topography of each region.

by a stochastic point process (PP) which describes the random spatial patterns formed by points in space. It was shown in [16] that infrastructure sharing outperforms spectrum and full sharing in terms of the coverage probability. However, spectrum sharing outperforms in terms of average user throughput. None of these research work consider the economics by looking at the actual costs of building a greenfield next-generation network or buying spectrum from the FCC.

2.2 TECHNICAL BACKGROUND OF WNV

One of the first proposed virtualized wireless network is OpenRoads [18] which was designed to replace the current WiFi network with a centralized WiFi network controlled by Network Operating System (NOX). This NOX is virtualized through different clients such that each client can control his own flows and set his own policies. As an upgrade to OpenRoads, Odin [19] is considered the first detailed WiFi virtualized architecture where the WiFi physical Access Point (AP) is virtualized between different clients using the notion of Light Virtual Access Point (LVAP). The main purpose of LVAP is to facilitate seamless mobility between APs. As an extension to Odin [19], OpenSDWN [20] is a virtualized Wifi network that takes into account the user application QoS to enable service differentiation, and allows administrators or users to specify application and flow priorities on the wired and wireless portion of the network. Later on, SoftRAN [21] proposed a centralized architecture as an alternative to the distributed control plane currently implemented in LTE networks. It abstracts out all the base stations deployed in a geographical area as a virtual "big-base station" while considering all the physical base stations as just radio elements with minimal control logic. Also, in Cloud RAN [22], base-band units (BBUs) are separated from the remote radio unit (RRU) to do signal processing and filtering. Meanwhile, BBUs are migrated into a BBU pool, which is virtualized and shared by different cell sites, and RRUs are connected to the BBU pool via the fronthaul links. Later on, [23] introduced the Hypervisor as the LTE network coordinator which is responsible for scheduling network resources between different virtual instances running on top of it, i.e. virtual networks. The Hypervisor assigns resources abstractly based on information from the individual virtual eNodeB stacks, like user channel conditions, loads, priorities, QoS requirements and information related to the contract of each of the virtual operators.

It is clearly noticeable that literature is extensively filled with research work in wireless virtualization, nevertheless most of the aforementioned work had a very abstract business analysis of both InV and RRV without delving into different possible network conditions such as: network topology, type of Backhaul infrastructure, user/population density and etc. Furthermore, inter-operator spectrum sharing has not been economically evaluated from the network operator perspective. For example, it is not clear yet if pooling the licensed spectrum will have the sufficient monetary incentive to motivate mobile network operators (MNOs) to allow partial or full spectrum sharing. Hence, in this thesis, we investigate different sharing scenarios in an LTE greenfield network under different network conditions, i.e. network topologies, backhaul types and some demographic features of different cellular market areas (CMAs)/ economic areas (BEAs).

2.3 RESOURCE ALLOCATION AND MANAGEABILITY IN VIRTUALIZED NEXT-GENERATION MOBILE NETWORKS

After analytically evaluating the economic incentives to pool spectrum resources between MNOs, we investigate a major challenge of spectrum pooling in literature which is resource management and allocation. In [24], NVS, a resource scheduling framework, is proposed to efficiently allocate resources to different virtual network slices, i.e., a group of flows that belong to a certain SP/VE, subject to its fluctuating demands. Later on, NVS was improved into a scheme called NetShare in [25] where the utility function which is maximized includes the proportional fairness (PF) demand for each SP/VE, i.e., it includes both the SP/VE's throughput in preceding time slots as well as the current user demands of an SP/VE, to schedule resources among different SP/VEs (this work assumes SP/VEs share more than one BaseStation (BS)). Also, [26] addresses resource allocation in virtualized wireless networks when using full-duplex relaying where radio spectrum, base stations and full-duplex relay stations are virtualized as virtual resources. On the same line of relaying networks, [27] accomplishes a simple form of spectrum and infrastructure sharing by exploiting multiple antennas at the relay via block diagonalization (BD). Later on, [28] used a controller for resource scheduling not only among different SPs but also taking into consideration fairness between users belonging to different SPs. The work in [29] proposes an opportunistic resource

sharing in wireless virtualized networks such that some of the physical spectrum is opportunistically accessed by users, i.e. content-based sharing.

Unlike the literature mentioned above, we address the resource allocation problem when applied to next-generation virtualized mobile networks. To the best of our knowledge, this problem has not been technically investigated in the literature. Unlike traditional cellular networks, the diversity of service classes in NGMN networks, i.e., enhanced Mobile BroadBand (eMBB), Ultra-reliable critical communications (URCC) and Machine Type Communications (MTC), makes resource allocation among different SP/VEs, sharing the same physical network and sharing spectrum, a multi-dimensional problem. The dimensions of this problem is depicted as follows.

2.3.1 Problem Dimension 1: Time-varying Resource Quality

One aspect of the problem is how the resource manager (hypervisor) should assign resources according to each virtual network's demand, i.e., traffic demand and classes of service, and the quality of channels in each virtual network, while preserving isolation between different SP/VEs, i.e. achieving high utility for one SP/VE should not affect other entities sharing the same pool of resources. In this thesis, we use LTE as the cellular system of interest, where the pooled resources are the wireless spectrum resource blocks (RBs). RBs comprise of a subset of frequency channels over a given time unit and are calculated every time slot by the RM based on the traffic demand and the type of applications used by each user. An intuitive proposition for fair resource scheduling among SP/VEs may be to simply apply the well known proportional fairness (PF) algorithm [30] that is used for scheduling mobile users of a single operator in current cellular networks. This algorithm uses the historical throughput that a mobile user has received and the channel conditions to determine whether or not resources are allocated to this user in the current transmission time interval (TTI). However, if such a PF algorithm is naively applied in the virtualized case, one possible scenario is that one SP/VE may not be able to acquire resources due to relatively bad channel conditions for its users. This clearly violates the isolation required between SP/VEs and hence PF scheduling cannot be directly applied to virtualized NGMN networks *prior to initializing the allocation process*. Thus, in our resource allocation proposed scheme, we add a prioritization step before the PF-scheduler as we will discuss in Chapter 3.

2.3.2 Problem Dimension 2: Traffic Heterogeneity in Next-generation Mobile Networks

The other aspect of the NGMN dilemma is the heterogeneity of user applications in NGMN (e.g., delay-sensitive applications, elastic applications and ultra-reliable applications). As it is well known, the characteristics of real-time communication applications differ significantly from those which are elastic [31]. For instance, in [32], it was shown that 99% of inter-arrival times for VoIP traffic are below 20 ms and the jitter delay is almost always less than 2 ms. Also, real-time traffic is transmitted in small sized packets every constant epoch of time. We can mostly model RT traffic as comprising of constant bit-rate (CBR) traffic. Furthermore, the ultra-reliable low latency communications (URLLC) is one of the emerging services in next-generation networks, i.e. 5G [4]. URLLC refers to communication services where data packets are exchanged at moderately low throughput (e.g., 50 Mbit/s) but with stringent requirements in terms of reliability (e.g., 99.999%) and latency (e.g., 1-4 ms). Examples of URLLC include reliable cloud connectivity, critical connections for industrial automation, tactile communications, and reliable wireless coordination among vehicles [33]-[34]. However, such stringent level of reliability is hard to attain in today's LTE networks (presumably the average error probability of sending a packet is no more than 10% [35]).

Based on the aforementioned resource allocation challenges, the second contribution of this thesis is developing a PF-based resource allocation scheme to handle traffic heterogenity of the RT-nonURLLC, NRT-nonURLLC, RT-URLLC and NRT-URLLC communications in virtual LTE networks. More precisely, we achieve isolation between different SPs with the Virtual Prioritized Slice (VPS) scheme. We argue by simulation results that serving the RT requests jointly for all SP/VEs using a fixed number of resource blocks (RBs) prior to scheduling positively impacts SP/VEs by reducing the user blocking probability for RT flows and in the meantime helps increase the total throughput per each individual SP/VE. While doing this, we consider the allocation of the worst resource blocks to RT traffic since they do not necessarily need the best throughput. Furthermore, we enhance the VPS scheme to provision diverse types of user traffic more efficiently where the RT/NRT-URCC requests achieve their ultra-reliability while preserving the target useful throughput. This is done through optimizing the coding rates of the transmitted packets instead of relying on the traditional CQI-coding mapping in current LTE networks which is not satisfac-

tory to establish ultra-reliable links. Moreover, we change the minimum allocation of resources Q_{min}, by taking into account the channel conditions. We propose a PF-based scheme that dynamically assigns Q_{min} to each of the SP/VEs based on their user demands as well as their historical transmission rates. We argue that such an assignment of Q_{min} improves the virtualized system performance through simulation results.

2.4 SPECTRUM POOLING IN DISASTER RECOVERY NETWORKS

2.4.1 Disaster Recovery for Cellular Networks in Literature

Due to natural disasters such as earthquakes, hurricanes and etc., base stations in cellular networks are prone to be disconnected and out of service. Unfortunately, the disruption of the cellular network in such situations put up the risk of losing more victims due to the miscommunication and the slower emergency response. Hence, in the research literature many solutions have been proposed to recover the cellular network in disastrous situations. In [36], disastrous network challenges are classified based on intent (non-malicious or malicious), scope (nodes, links and areas) and domain (wired or wireless). Wired networks such as passive optical networks (PONs) [37], which are widely used for today's backhaul networks, are sensitive to failures because multiple leaf nodes share one optical fiber. Consequently, the radio access network may get detached from the core network (the mobility management entity (MME), the packet gateway (PGW) and other core entities/functions). Hence, the RAN network becomes isolated and incapable of providing services to users. Most of the disaster recovery proposed solutions consider the replacement for the cellular communications access networks by establishing an adhoc network, using satellite communications and other alternative solutions to cellular communications. In [38] Lien at. al. proposed to use WiFi-ready notebooks to construct a MANET based group communication system to support emergency communication and information network, i.e. P2Pnet. On the same lines, Wireless Mesh Networks (WMN) were introduced as a solution during emergency and disaster recovery. The work in [39] proposed a multicast enabled dual radio wireless mesh network called Swan-Mesh where a multicast upgrade of the Adhoc On-Demand Distance Vector (AODV) protocol is

used to deliver live video and audio real time data to multiple users (multicast) during disasters. In [40] an adhoc routing protocol is developed to help communicate between BSs and user nodes where the nodes that connect directly to BSs operate in cellular mode; while nodes that do not receive signals from any BS communicate with their neighbors in adhoc mode. However, the adhoc network protocol is not compatible with cellular communication standards, e.g. 3GPP. Moreover, adhoc protocols lack efficiency in terms of handoff, synchornization and average user throughput compared to SAFECOM Public Safety Statement of Requirements Reports issued by the United States Department of Homeland Security as was shown in [41]. Satellite communications systems were introduced as a core network for emergency management and disaster recovery including satellite network topologies, service configuration, and recommended technical specifications for broadband communications during disaster or emergency times [42]. However, such type of communications requires huge infrastructure to operate and manage the network. Also, a lot of changes to the governance is needed to make the end-to-end reference architecture work as standardized by the 5G standards. The 911-network on wheels (911-NOW) solution was also introduced to enable first responders and emergency management teams to communicate mission-critical information on a secure and rapidly deployable wireless network [43]. The 911-NOW solution relies basically on deploying a basestation router (BSR) in the RAN on wheel where the radio resource and mobility management functionalities are deployed. Also, OEMAN [44] is an On-the-fly Establishment of Multihop Wireless Access Networks to quickly extend Internet connectivity from the surviving access points to form on-the-fly multihop wireless access networks that effectively help bring Internet connectivity to disaster victims.

2.4.2 Spectrum Pooling: An Enabler to Disaster Recovery Networks

Due to the challenges of integrating other technologies such as satellites and WiFi in cellular networks as aforementioned, we argue that saving some of the RANs using back-up power supplies such as micro-grids can help recover the communication network between the survivor RANs efficiently without introducing any other technology. However, in most cases of severe natural disasters, the core network is detached from the radio access network and thus the RAN capabilities become limited to the same cell unable to communicate with other cells. To solve this issue,

[45] suggests that every cluster of eNBs should be attached to an Emergency Communication Network Server (ECNS). **An ECNS is a server that encompases additional functions performed by the core network such as mobility management (MME), Authentication, Authorization and Accounting (AAA), PTP based clock synchronization and other auxiliary services**. The following part spotlights some of the literature work in micro-power grids as power supplies in cellular networks.

2.4.2.1 Powering eNBs using Micro Power Grids

The RAN can be powered using a back up power source in critical cellular communications. There are two main backup approaches used in the literature [46]: (i) either to power the RAN using a local backup power plant with batteries and a diesel generator or (ii) to use a micro-grid with varying alternatives for the local distributed generation power sources [2]. Both backup power approaches typically require some external source of energy that is delivered into the service area by one or more primary energy supply infrastructures (PESIs). Meanwhile, micro-grids show promise to achieve improved power supply resiliency to critical events over other backup approaches because well designed micro-grids, i.e., microgrids with diverse power supply from at least two distinct PESIs eliminate the single point of failure [47]. The basic requirements for micro-grids in disaster recovery networks are to operate in a stand-alone mode (i.e., the so called island mode) in which the micro-grid operates on local sources of energy (wind turbines, solar systems) with the local microgrid control system providing voltage and frequency (in the case of AC) stability for optimal power flows, and ensuring minimal load shedding and disruption during transitions from the main power grid [48].

In [49] an architecture for green cellular networks that is based on a microgrid configuration was proposed to power a cluster of seven cellular base stations mostly using renewable sources. This study shows that it is possible to feed the base stations with power obtained from renewable sources for up to 90 % of the time when using 250 kW wind turbine and at each cell site a solar panel with an area of $3\ m^2$.

[2] A local distributed generation power sources (LDPS) typically use renewable energy sources, including small hydro, biomass, biogas, solar power, wind power, and geothermal power to generate electric power. A grid-connected device for electricity storage can also be classified as an LDPS system and is often called a distributed energy storage system (DESS).

2.4.2.2 Challenges of Micro-grid Power Supplies : The cost of tens of small microgrids just to power a cellular network throughout a city during power outages is prohibitive [45]. Therefore, [45] proposes multi-user microgrids to be deployed such that the cost is shared by the smart infrastructure owners with mission critical loads (e.g., water system, cellular networks), societal important locations with mission critical loads (e.g., hospital), and the local government (police, fire, 911 service) that will utilize the communication network powered by the microgrid for public safety communications.

2.4.2.3 Spectrum Pooling Enabling Cellular-based Disaster Recovery : Building on the aforementioned proposals of micro-grid power supplies used to power up some of the RAN network and adopting the ECNS idea, we propose a multi-hop wireless mesh network between survivor BSs (SBSs) to provision mobile users without integrating any alternative technology. Regardless of the mobile network operator who owns each BS and assuming a number of BSs die after electricity blackout, we pool the available spectrum at all BSs in this area and then we reallocate the pooled resource blocks (RBs), the BS transmission power as well as the operating frequencies available at this cellular area to each BS to provision the requests of the remaining SBSs so that the inter-links (wireless links between BSs) as well as the intra-links (wireless links between a BS and corresponding users within the same cell) are maintained while achieving the maximum average throughput per cell as well as satisfying the traffic demand and QoS constraints (power and delay constraints).

In the next Chapter, we describe the thesis statement, assumptions and methodology in detail.

3.0 SUMMARY OF DISSERTATION

In this chapter, we describe the main challenges that we propose to tackle in this thesis as well as the proposed solutions, the thesis statement, the model assumptions, the thesis contributions and we conclude with suggested future work.

3.1 OVERVIEW

As discussed in Chapter 1, due to new uses cases and business models in next-generation cellular networks, the CAPEX and OPEX of building new infrastructure can be very high and impractical. Hence, the core idea of this dissertation is the establishment of greenfield next-generation cellular networks by exploring infrastructure and spectrum sharing. By studying different sharing scenarios in a fiber-based backhaul network, we aim at evaluating incentives for SPs to share spectrum/infrastructure in different CMA/BEAs and hence break the market between new roll-outs. Moreover, we look microscopically at the technical side of pooling the spectrum between two SPs sharing the same BS yet they have different traffic demands as well as different QoS constraints. In our work in Chapter 5, we developed two resource allocation schemes to provision real-time (RT), non-real-time (NRT) as well as Ultra-reliable Low Latency Communications (URLLC) traffic when the spectrum resource blocks (RBs) are shared between two SPs such that our schemes achieve isolation, fairness and guarantee the QoS constraints. We evaluate the performance of the newly proposed approaches when sharing in the LTE 700 MHz and the AWS 2100 MHz spectrum bands. Eventually, we use spectrum pooling to resolve the disruption in the cellular network due to disasters by providing the survivor eNBs with Emergency Communication Network Server (ECNS) and reconnect these eNBs together using a multi-hop wireless mesh network.

3.2 THESIS CHALLENGES AND PROPOSED SOLUTIONS

Our research work aims to address the following challenges:

3.2.1 Assessment of The Economic Incentive of Spectrum Pooling

Given a certain Cellular Market Area/Regional Economic Area (CMA/EA), a greenfield LTE-like network, a fiber-based backhaul network and two service providers (SPs), the first task of this thesis is to collect FCC data as well as backhaul infrastructure data for more than 1000 cellular market areas/economical areas (CMA/BEAs) across US in the 700 MHz as well as the 2100 MHz spectrum bands. We use the collected data as well, develop new cost models in different sharing scenarios and compare the resultant profits of these scenarios. We specifically analyze the backhaul sharing, the spectrum sharing and the backhaul+spectrum sharing scenarios. We provide answers to the following questions,

Is there an incentive for SPs to share their own spectrum resources, infrastructure resources or both of them (full sharing)? And how does this incentive change with the change of the backhaul network topology, the operating frequency band, the number of BSs deployed to cover this area as well as the population density in this area?

The proposed solutions:

1. We developed a cost model to assess the cost and revenues of establishing a greenfield network using a fiber-based backhaul and considering different network topologies as well as different CMAs with different population density, allocated BW and spectrum bid values. We show that spectrum sharing and backhaul sharing are not always beneficial. However, the sharing incentive is affected by different conditions of the network and the demography of the area of deployment. Also, from the FCC auction data, we observed that the 2100 MHz spectrum bidding value is much lower than the bidding value of the 700 MHz band. Thereby, the incentive to pool the spectrum should decrease and be less dependent on the designated population density. To confirm these inferences, we extended the proposed cost model to include the FCC data on spectrum bids for the 2100 MHz-AWS band to explore the change in the sharing

incentive for different bands.

2. Moreover, we apply our cost models to different areas, i.e., rural and urban areas [50]. Microtrenching, horizontal directional drilling (HDD) and manual trenching are well suited for urban areas, and cost approximately $50, $37 and $25 per meter respectively. However, in rural areas, more cost-effective techniques, such as direct cable burying (plouging), machine trenching (in areas not suitable for ploughing) and aerial deployment could be used instead and at a cost of approximately $7, $14 and $4 per meter, respectively. We show by numerical results that this big difference in trenching costs between urban and rural areas will affect the CAPEX cost (backhaul cost) and hence will change the sharing incentive for SPs. Accordingly, the backhaul network cost in rural areas is much less than the backhaul cost in urban areas. Thus, sharing the backhaul infrastructure in urban areas is expected to show higher incentive than in rural areas.

3.2.2 Efficient Resource Allocation in a Spectrum Pooled NGMN

Since there exists an economic incentive to share spectrum resources in CMAs with high population densities, we expose the technical part of sharing a single BS between two SPs with different traffic demands. Given two SPs who share the same RAN and agreed to share their radio spectrum to provision heterogeneous types of traffic, i.e., RT, NRT and URLLC, **How should we reallocate the pooled spectrum (RBs) so that the resource manager guarantees isolation and fairness among requests from different SPs?**

Proposed solutions:

We answer the aforementioned question by developing a resource scheduling scheme for a single shared BS scenario that comprises of the following features.

3.2.2.1 Joint Prioritization of Delay-sensitive traffic flows Unlike traditional LTE networks, in virtualized NGMN networks, it may not be the best approach to execute the prioritization step, for delay sensitive flows, for each SP/VE separately. We claim that prioritization between different traffic types should be performed for all SP/VEs *jointly* at the network scheduler prior to the

allocation of resources to SPs. This can better leverage the economies of scale effects of spectrum pooling.

To clarify this, suppose we have a total of 8 PRBs and SP1 and SP2 each have 2 RT requests at the beginning of a TTI. Also, we presume it takes 2 PRBs to serve a single RT request. As explained in Chapter 5, we assume also for simplicity that SP1's PRBs all correspond to $CQIindex = 11$ in the modulation-coding scheduling table [51] and similarly, SP2's PRBs transmission rates correspond to $CQIindex = 14$. Hence, assuming both SPs have the same transmission rates in previous time slots and assuming a proportional fairness scheduler, SP1 experience worse channel quality and consequently lower number of RBs than SP2. Accordingly, the resource manager will assign 5 RBs to SP2 and 3 RBs to SP1 (more RBs to the SP with better channel quality). As a result, this will hinder SP1 from serving all of its RT requests and it would have one dropped RT request and hence violate the isolation principle.

In contrast, let us suppose that the network scheduler prioritizes RT requests altogether first and serves them using all of the available RBs. Then, none of the VEs would encounter drops of RT requests. Using this idea, our proposed resource scheduling scheme takes into account the existence of a priority-phase *prior to allocating resources* where all RT requests from all SPs are prioritized jointly at the resource manager first in a single queue called the virtual prioritized slice (VPS), independent of the PF-criterion to ensure that RT requests are better provisioned while maximizing the throughput of NRT users.

3.2.2.2 Provisioning URLLC traffic using URLLC coding Optimizer

The second feature of our proposed scheme is serving URLLC traffic. The LTE CQI calculations in the MCS table is designed to maximize the user throughput such that the UE power and delay constraints should be met. Also, it is presumed that the average error probability of sending a packet is no more than 10% [35] for a maximum delay of 1 - 10 msec. Figure 1 shows the reliability (%) versus the user-plane delay in traditional LTE networks, i.e. traditional CQI mapping. This figure is produced assuming AWGN channels for uplink and downlink in a cell with a radius of 500 meters and a total of 2 mobile users. We assume that one of the users is downloading a file in a size of 0.5 MBytes and the other one has a URLLC request with a packet size of 100 Bytes and an arrival rate of 2 packets per TTI. We generate the locations of the users in the cell randomly, we iterate the simulation

1000 times and we take the average over 1000 observations. Using the MCS table, we simulate an LTE cell using MATLAB and derive the average link reliability. In Figure 1, we plot the average link reliability versus different maximum delay values. Similarly, we plot the average BS antenna transmission power versus different delay constraints in Figure 2. From Figure 1, we can clearly see that achieving a user plane latency as low as 0.5 msec makes the target 5G ultra-high reliability unattainable, i.e. $\simeq 0.999$. Moreover, the transmitting power reaches the maximum allowable transmission power by a pico BS, i.e. 38 dBm according to 3GPP standard [52]. Thus, the LTE network reliability is very loose compared to the stringent requirement by URLLC applications. Therefore, we argue that the traditional LTE MCS table should be adjusted for URLLC requests to accommodate ultra-reliable users in NGMN networks, i.e., our e-VPS resource scheduler designed in Chapter 5.

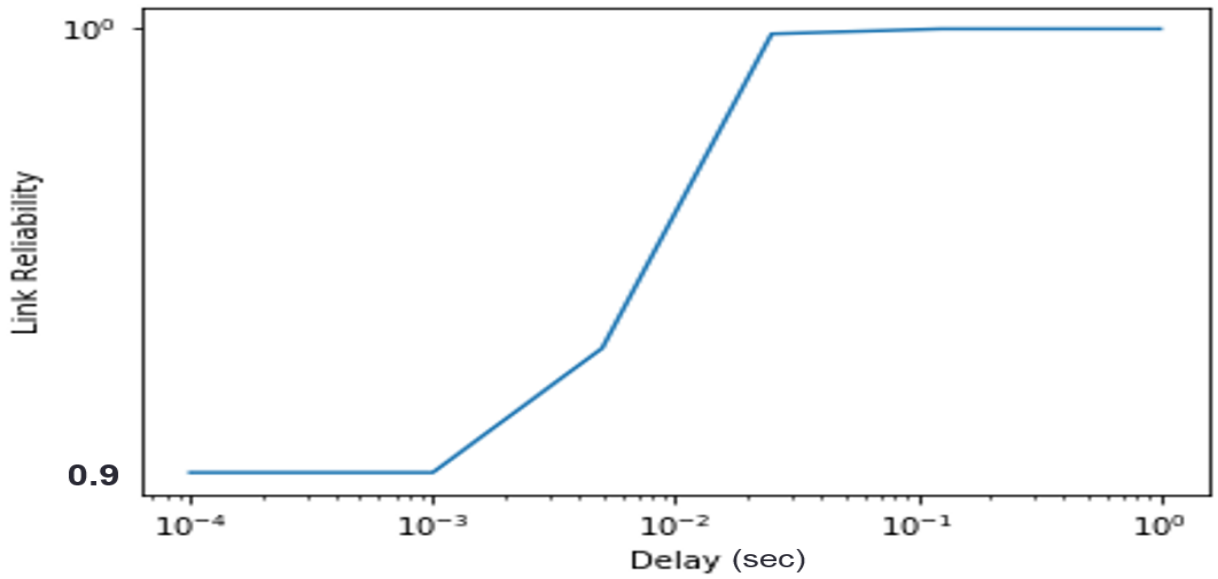


Figure 1: Reliability versus user delay constraint in traditional LTE networks

3.2.2.3 Dynamic Allocation of Q_{min} to Virtual Entities The last feature that we propose in our scheme is allocating a dynamic minimum reserved number of RBs for SP i, named Q^i_{min}, such that resources do not remain unused or no SP/VE is assigned fewer resources than possible to improve the average throughput and user blocking probability. Unlike nowadays static Q_{min} that

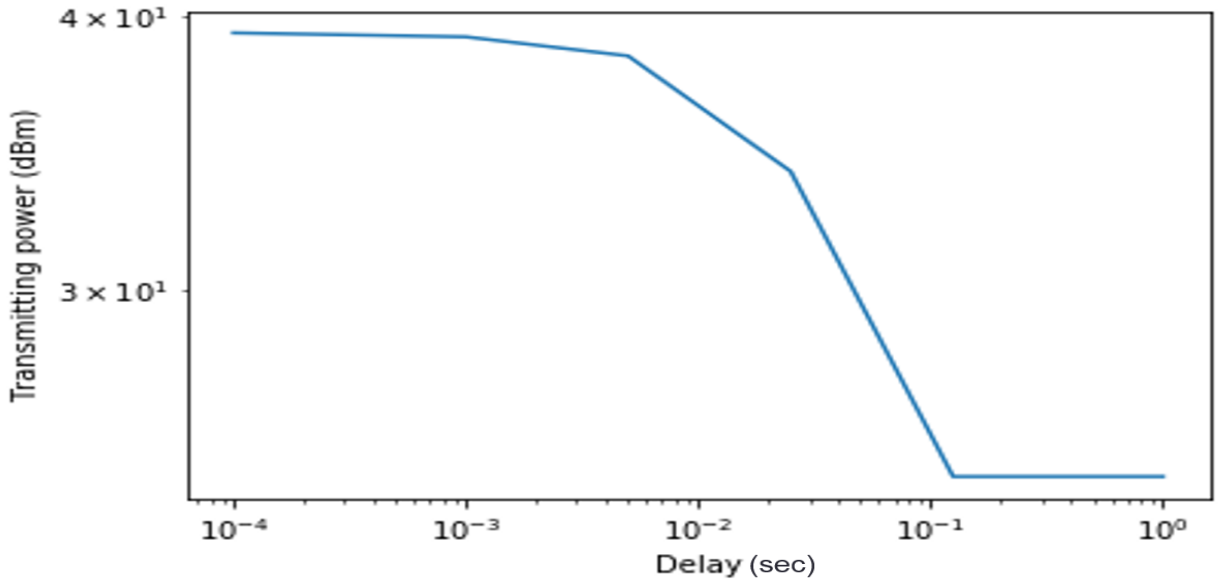


Figure 2: Transmitting Power versus user delay constraint in traditional LTE networks

are stated in SLAs, since the traffic demand changes with time and thus the number of allocated resources changes as well, the minimum number of allocated resources Q_{min} should not be fixed per an SLA. However, Q_{min} should be computed every number of TTIs based on the traffic demand provisioned by each SP and the transmission rates in former TTIs. Thus, in this proposal, we present an efficient method to derive Q^i_{min} for SPi based on both the average long term transmission rate of SPi in previous slots allocated to its users as well as the current demands of this SP/VE, albeit at an SP/VE level than a user/flow level.

We compare our scheme to a recently proposed scheme called NetShare and we show with simulation results that our newly proposed scheme outperforms NetShare in terms of the average cell throughput, the bandwidth utilization as well as link reliability in case of provisioning URLLC traffic.

3.2.3 Using Spectrum Pooling as a Disaster Recovery Technique in Cellular Networks

Given a certain CMA in a disaster situation where the backhaul network is severely affected and assuming a number of surviving BSs that belong to different SPs, we answer the following question,

Can we use spectrum pooling in next-generation networks as an enabler to recover post-disaster cellular networks without integrating any alternative technology? How can we optimally allocate different frequency bands (700, 2100 MHz) to satisfy different 5G QoS constraints?

The Proposed disaster recovery approach:

We propose a multi-hop wireless mesh network between survivor BSs (SBSs) to provision mobile users without integrating any alternative technology. Regardless the mobile network operator who owns each BS and assuming a number of BSs die after electricity blackout, we pool the available spectrum at all BSs in this area and then we reallocate the pooled physical resource blocks (PRBs), the BS transmission power as well as the operating frequencies available at this cellular area to each BS to provision the requests of the remaining SBSs so that the inter-links (wireless links between BSs) as well as the intra-links (wireless links between a BS and corresponding users within the same cell) are maintained while achieving the maximum average throughput per cell as well as satisfying the traffic demand and QoS constraints.

Unlike the standardized X2 links between BSs, the interlinks between BSs in the wireless mesh network have to operate on ultrahigh frequencies (UHF) to ensure better signal propagation in high interference and fading cellular environment. And hence, BSs are able to provision user equipment (UE) traffic and satisfy the QoS constraints for different applications (BS transmission power, user delay and throughput constraints). Yet, operating on the same frequency bands as for the intralinks will cause severe interference for UEs. More specifically, to establish line of sight (LoS) connections between BSs, using UHF frequencies will impose significant interference to the intra-cell radio links. Therefore, designing the cellular wireless mesh network is not a trivial problem. We argue that selecting the carrier frequency of the interlinks based on intra-cell transmitting powers as well as intra and inter-cell traffic demand is very crucial in designing the links between SBSs so

that the interference between inter-links and intra-links is minimized.

To address the above problem, we formulate a joint RB-power optimization problem to find the optimal transmission powers , the optimal inter-cell frequencies and the optimal intra-cell frequencies. We also, in this part of the thesis, we examine the feasibility of recovering the cellular network using a wireless mesh network between SBSs such that the UE QoS needs and power constraints are still viable. Moreover, in the second phase, we allocate the intra-cell frequencies such that the different QoS for the real-time, non-real time and URLLC services are satisfied. As we can see in Figures 3 and 4, using the 700 MHz bands achieves better reliability ($\geq$ 99.999). **We show by numerical results that applying the previously proposed e-VPS scheme at the 700 MHz band achieves higher reliability. However, the link reliability is deteriorated when sharing the spectrum at the 2100 MHz. Hence, we assign the lower intra-cell frequencies to URLLC traffic to enhance the reliability while achieving the minimum throughput and the maximum allowable delay. On the other hand, we assign the 2100 MHz-AWS frequencies to the non-real time requests that need very high data rates.**

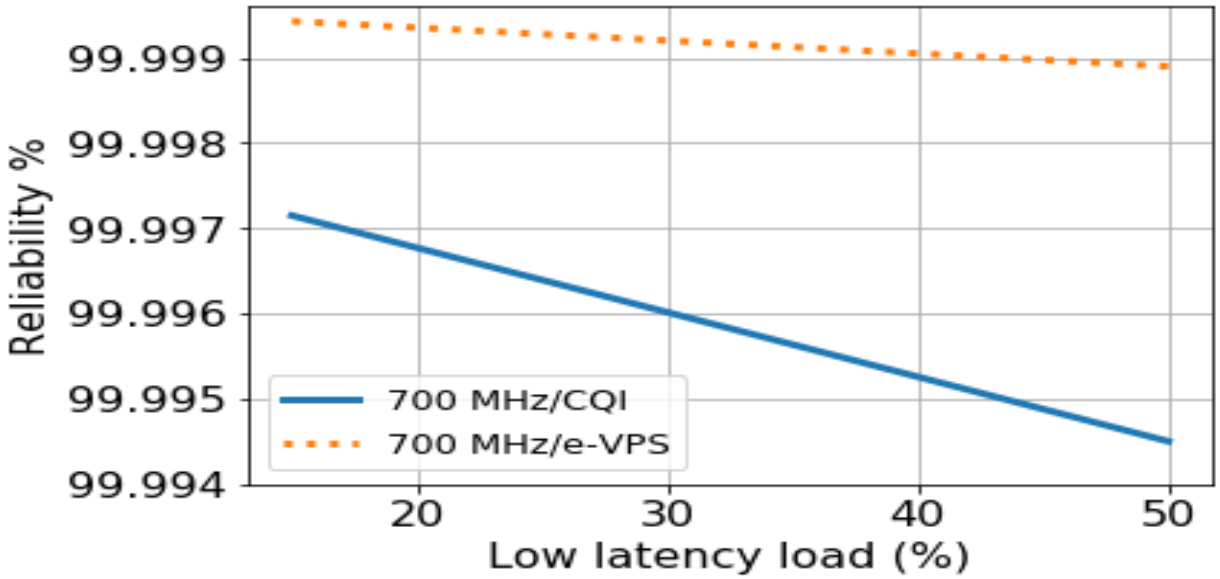


Figure 3: Reliability versus URLLC load in the 700 MHz band

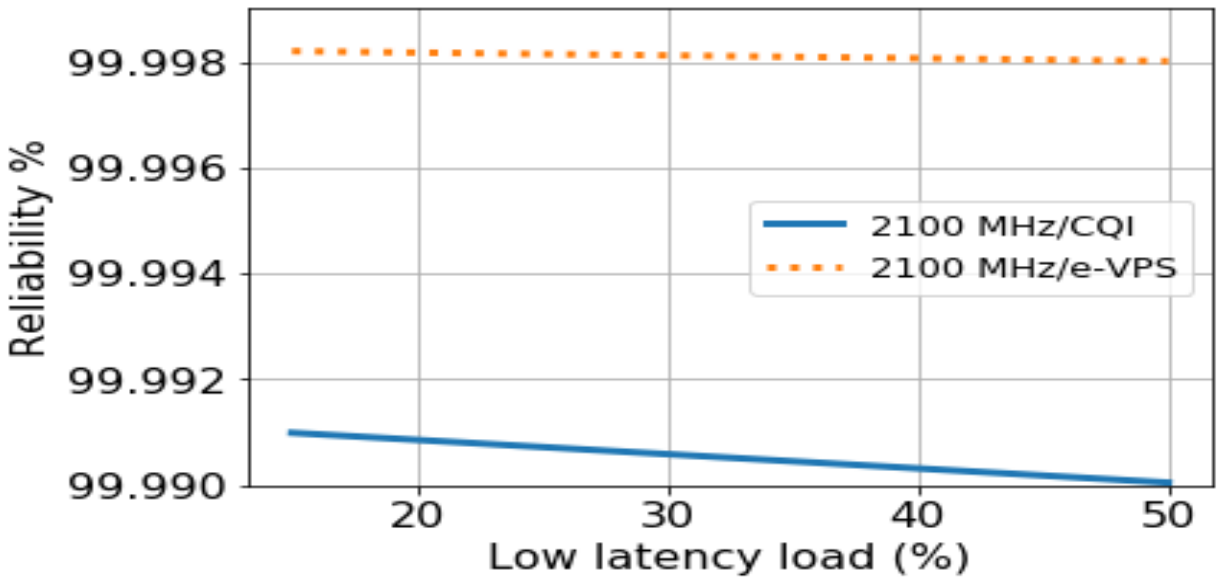


Figure 4: Transmitting Power versus URLLC load in the 2100 MHz band

3.3 SCOPE AND ASSUMPTIONS

Throughout this research, we assume the following:

1. The considered cellular network is an LTE network with fiber-based backhaul networks.
2. We assume all cells are of the same size to simplify the analytical cost model and use some approximations.
3. Cells can be micro or femto cells.
4. We assume a reuse factor of (n = 1), i.e. all frequencies are used in all cells using Fractional Frequency Reuse (FFR).
5. The backhaul network can be based on either a tree or a ring topology.
6. The user traffic is uniformly distributed across the cell.
7. The number of users per cell is directly depending on the population density at the CMA/BEA region.
8. All network deployment costs are estimated for urban areas only (disregarding the difference in costs between urban and rural areas).

Notwithstanding their importance, there are some related issues left outside the scope of this work:

1. The non-homogeneity of topography of the CMA/BEA: the assumed topography is homogeneous and regular-shaped, i.e., square or rectangle.
2. Different cell sectorization: we assume a fractional frequency reuse of 1.
3. Reducing the computational complexity of the proposed algorithms: despite the high computational cost, we assume a time granularity of 1 TTI.
4. Optimizing the time granularity of updating the RM with the channel state information, Q_{min} as well as the user demand.

3.4 METHODOLOGY

We answer the first question by developing a cost model to assess the cost and revenues of establishing a greenfield LTE network using a fiber-based backhaul and considering different network topology as well as different CMAs with different population density, allocated BW and spectrum bid values. We show that spectrum sharing and backhaul sharing is not always beneficial, however sharing incentive is affected by different conditions of the network and the demography of the area of deployment.

Afterwards, we answer the second question by proposing two efficient schemes for resource allocation in virtualized mobile networks. These schemes uses prioritization as well as a new channel coding optimizer to schedule RT requests and URLLC, respectively, and hereafter provisions the NRT requests using the proportional fairness scheduler. We compare our schemes to a recently proposed scheme called NetShare in terms of cell throughput, bandwidth utilization and reliability. We apply the proposed schemes to inter-band sharing, i.e. LTE-AWS, and intra-band sharing, i.e. using 700 MHz LTE band. We show the throughput gain achieved by inter-band sharing. However, intra-band sharing (in the 700 MHz spectrum band) achieves higher reliability for URLLC traffic. Eventually, we exploit the different gains from inter-band and intra-band sharing to best pool the spectrum between different BSs in post-disaster cellular networks. In particular, we pool the spectrum in the dead area between SBSs so that the total throughput is maximized while provisioning heterogeneous classes of traffic (real-time and non-real time). We will formulate a maximization

problem to maximize the cell throughput of the wireless mesh network as well as satisfying the power and QoS constraints. The target of the optimization problem is to find the inter-cell frequencies as well as the optimal transmitting power and bandwidth to be allocated to inter and intra-links.

3.5 FUTURE WORK

This thesis does not address the computational complexity of the proposed algorithms. Moreover, our assumption of updating the channel state information as well as Q_{min} every Transmission Time Interval (TTI) will add significantly to the complexity of the scheduling algorithms. How can this complexity be reduced is an important futuristic task to explore.

On another side, a main assumption of our model is the fixed resource assignment of RT requests. This may not be feasible for HD video streaming as an example. This is because having a high number of RBs assigned to RT requests will limit the spectrum resources to RT traffic exclusively and thus reduce the NRT traffic throughput significantly. Hence, it is worth studying such scenarios in future work and enhance the resource scheduling algorithm to take into account various RT applications.

3.6 LIST OF PUBLICATIONS

- [53] A. AbdelHamid, D. Tipper, and P. Krishnamurthy, "Recovery and optimization of post-disaster cellular networks," in *2019 15th International Conference on the Design of Reliable Communication Networks (DRCN)*. IEEE, 2019, pp. 16–20.

- [54] A. AbdelHamid, P. Krishnamurthy, and D. Tipper, "Resource allocation for heterogeneous traffic in lte virtual networks," in *Mobile Data Management (MDM), 2015 16th IEEE International Conference on*, vol. 1. IEEE, 2015, pp. 173–178.

3.7 CONTRIBUTION

In this thesis proposal, we proceed as follows.

- In [53], we proposed a QoS-aware disaster recovery algorithm in cellular networks. He formulated and numerically solved a convex optimization problem to find the optimal frequencies and bandwidth allocations for inter-links and intra-links.
- In [55], we proposed the analytical cost models for different sharing scenarios and conducted the data analysis and mining.
- Unlike the existing literature, our proposal in [55] develops an analytical model and uses real population data, spectrum bids from FCC, and network component costs (fibre, splicing, etc.) from FiberStore to obtain insights about spectrum and infrastructure sharing in US CMA/BEAs.
- In [54], we developed the virtual Prioritized Slice (VPS) scheme, a resource allocation scheme to schedule the RT and NRT traffic that belong to different VEs sharing the same eNB as well as the same spectrum. Afterwards, we validated the scheme using Matlab simulations.
- In [56], we enhance the resource allocation scheme proposed in [54] to include a new 5G use case, i.e. ultra-reliable low latency critical communications (URLLC). Also, in [56], we developed a minimum allocation criterion for the sharing SPs such that both isolation and fairness are achieved.
- To the best of our knowledge, our proposed scheme in [56] is the first work to handle resource allocation considering the provision of ultra-reliable critical communications (URLLC) in next-generation networks.

4.0 INFRASTRUCTURE VIRTUALIZATION VS RADIO RESOURCE VIRTUALIZATION: AN ECONOMIC FRAMEWORK

To the best of our knowledge, there are no good cost models and data sets for quantifying *the circumstances under which sharing of spectrum and network resources would be beneficial to operators*. The objective of this chapter of the thesis is to (i) develop cost models that can help assess such benefits (ii) use these models with publicly available data to obtain insights into the costs and benefits of sharing the network infrastructure and radio spectrum in real world data. Unlike the existing literature, we develop an analytical model and we use real census population data, spectrum bids, and network component costs (fibre, splicing, etc.) to obtain insights from the model. We rely on data from the FCC for the spectrum costs (expressed as the net bid amount) in a total of 1054 cellular market areas (CMAs)/ economic areas (BEAs) [57]. The infrastructure fiber costs are obtained from **FiberStore** [58]. We use the model to show how the backhaul and spectrum costs change with the population density, the backhaul network topology as well as the number of Base Stations (BSs) deployed in these CMAs/BEAs.

4.1 SYSTEM MODEL

In this section we describe our model assumptions and specify the backhaul construction. We assume that all cellular networks are LTE-based networks that operate in the 700 MHz and 2100 MHz bands. For simplicity, we assume the area A in our model a regular area which is structured such that the number of BSs aggregated per hop $= 3$ and $\ell \leq h$ as shown in Figure 5. We mean by h hops $\times$ ℓ levels : the number of hops needed to reach the core network $\times$ the number of eNodeBs per each hop as shown in Figure 5. We consider an area A of N_{eN} e-NodeBs (eNBs)

such that A is h hops $\times$ ℓ levels as shown by the 3 hops $\times$ 3 levels example in Figure 6. Without loss of generality, we use $h = \frac{N_{eN}}{\alpha}$, where α is a parameter to control the relationship between ℓ and h. Assuming the system model in Figure 5, the total number of BSs in area A can be given as

$$N_{eN} = \ell\big(3h + 1\big) \tag{4.1}$$

Using $h = \frac{N_{eN}}{\alpha}$ and equation (4.6), ℓ can be written as a function of h as follows.

$$\ell = \frac{\alpha h}{3h + 1} \tag{4.2}$$

It is worth noting that it is very hard to obtain an empirical value for α. Nevertheless, in Section 4.2, we prove by numerical results and using the mathematical cost model that α does not significantly affect the incentive of sharing network resources. Also, we assume a homogeneous network such that all BSs have coverage in a hexagon shape inscribed in a circle with radius R.

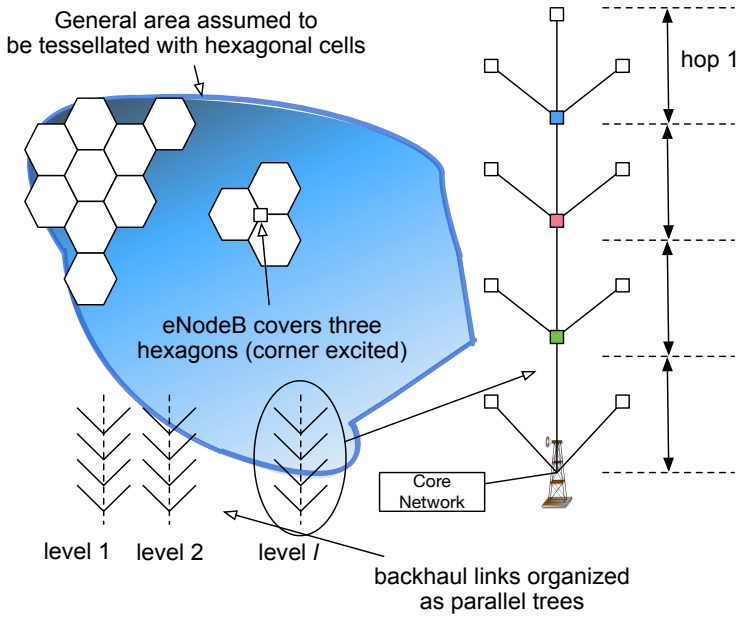


Figure 5: A Schematic Diagram for 4 hops by 3 levels backhaul topology in non-sharing scenario

These eNodeBs are connected together using a fiber based backhaul network. In the conducted analysis, we consider two different topologies: the tree and ring network topology.

4.1.1 Dimensioning the LTE Network

In order to get the different costs that contribute to establishing a greenfield LTE network, we need to calculate the number of BSs covering a certain CMA as well as the mile.sq areas covered by each BS. In this paper, we consider different approaches to get the number of BSs in a certain CMA/BEA.

4.1.1.1 Approximate Approach: In this approach, we use the data for the BS density from Deloitte 5G research study [59]. Deloitte roughly estimates the number of base stations per 10000 people in americas to be almost 5. Hence, we can get the number of base stations in every CMA according to the population density in such CMA.

4.1.1.2 Set Coverage Approach: For simplicity and assuming the area structure in Figure 5, we apply the set coverage optimization problem [60]. Given that $\mathbf{S}$ is the set of BSs in area A and $\tau_{\mathbf{i}}$ is a subset of the set $\mathbf{T}$ of test points where $\tau_{\mathbf{i}}$ is the subset of test points within the cell of BS $i \in \mathbf{S}$, the set coverage problem can be formulated as follows

$$\begin{aligned} \underset{\mathbf{x}}{\text{Minimize}} \quad & \mathbf{c}^T.\mathbf{x} \\ \text{subject to} \quad & \\ & \gamma_{i,j} \geq \gamma_{th}, \quad i \in \mathbf{S}, \\ & \mathbf{P_c^i} \geq P_{target}, \quad i \in \mathbf{S}, \\ & \mathbf{x} \in \{0,1\} \end{aligned} \tag{4.3}$$

where x is a binary vector such that $x_i = 0$ if at point i there is no need to establish a mobile BS and $x_i = 1$ when there should be a mobile BS at point i to achieve the target SINR threshold γ_{th} and the target probability of coverage $\mathbf{P_c} \geq P_{target}$. Also, $\mathbf{c}$ is the cost vector where c_i indicates the total cost of establishing a BS at point i.

After deriving the BS location vector $\mathbf{x}$, the number of eNBs (N_{eN}) needed to cover a certain CMA/BEA area can be calculated as the sum of elements of vector $\mathbf{x}$, i.e. $N_{eN} = \sum_{i \in \mathbf{S}} x_i$.

4.1.1.3 Real Data Approach

In this approach, we use real data that was gathered by OpenCellid [61] featuring all the BSs located in US. The data contains the BS location, the cell range, the location area code, and other features. We use the location area code, the longitudes and latitudes to classify BSs that belong to a certain CMA and group them together. Accordingly, we can find the number of BSs in a certain CMA based on the aforementioned classification.

4.2 A COST MODEL FOR A SINGLE MOBILE NETWORK: NO SHARING

In this section, we analytically derive the cost of establishing a single MNO in both tree and ring topology.

4.2.1 The Backhaul Cost in Tree Topology

As shown in Figure 6 we assume that the fiber backhaul is connected together in a tree topology such that there are three types of nodes, the source nodes (S-node), the kth-hop aggregating eNodeBs and the last hop eNodeBs. The S-node is an enodeB that forwards traffic to other nodes but has no aggregation capabilities. The kth-hop aggregating node is an eNodeB that aggregates the traffic from the eNodeBs in a former hop where $k = 2, 3, ...$, e.g., the 2nd hop aggregating nodes aggregate the traffic originated from the eNodeBs in the 1st hop and so forth. Also, the last hop nodes are the group of eNodeBs which are directly connected to the core network. It is worth noting also that in our work we assume that each aggregating node $k - 1$ aggregates only 3 of the kth hop nodes. In addition, according to [62] we assume two types of fibers, the feeder fiber (the bold segments) and the distribution fiber (the non-bold segments) as shown in Figure 6. The feeder fibers have higher capacity, i.e., 40 GB fibers, and hence are used in the aggregating hops. However, the distribution feeders have less capacity, i.e., 10 GB fibers, hence they carry the data from the S-nodes to the aggregating nodes.

Figure 6 shows that we have two types of fiber segments, short segments with length $2.7R$ and long segments with length $5.2R$ according to the cell geometry shown in Figure 7.

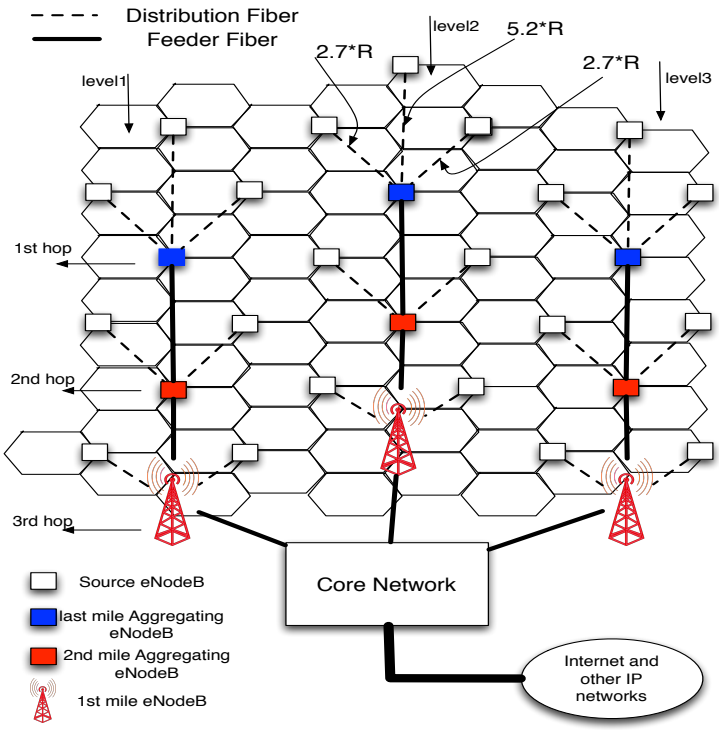


Figure 6: A Schematic Diagram for 3 hops by 3 levels backhaul tree topology in no-sharing scenario

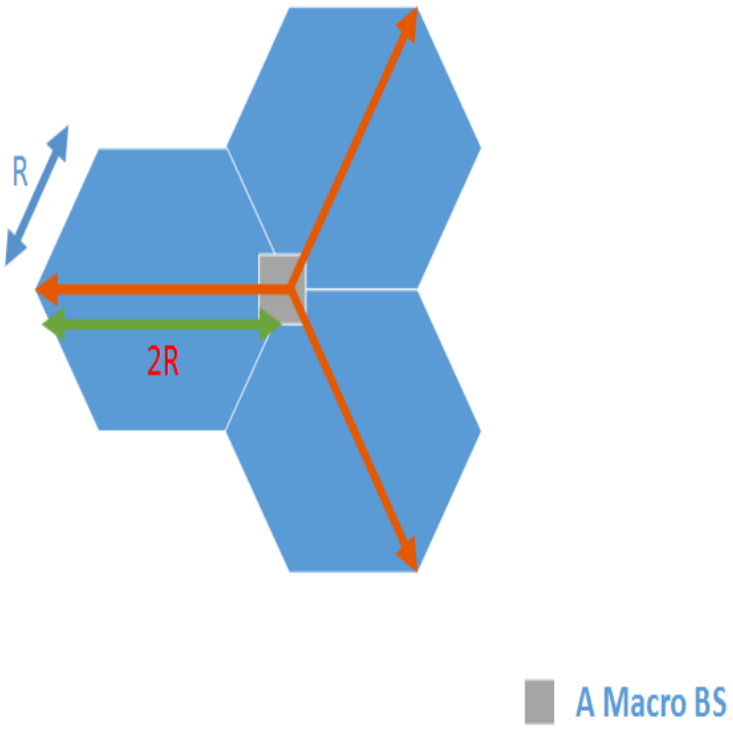


Figure 7: A Schematic Diagram for a standard Macro BS

Table 1 summarizes the infrastructure equipment and the corresponding symbols that we use in our model. From the backhaul perspective, each eNodeB should be equipped with an optical transceiver, an optical network unit(ONU), an optical line terminal (OLT) and a passive splitter-combiner (PSC) if it is an aggregating eNodeB [62]. It is common for fiber networks that fiber cables are not deployed as a single segment from source to destination. The total fiber length is chunked into a number of fixed length segments L, e.g., if the distance between source and destination is 10 miles and L = 2 miles, hence we need 5 fiber segments [63]. To connect between these segments, fusion splices are used [64].

In this section, we estimate the total backhaul cost in an area A with a total census of C_{en}. The census is the total population of a CMA/BEA. According to the LTE dimensioning in Section 4.1.1, for an area A, we can find the total expected number of eNodeBs with a radius R, where a standard macro BS covers three hexagon cells via sectorization as shown in Figure 7. Each of the blue cells in Figure 7 is a regular hexagon with a side length R. Now, we can find the total backhaul cost of

Table 1: Symbols of the prices for different fiber network components

component	symbol	cost
Distribution Fiber	d	Pr_d
Feeder Fiber	f	Pr_f
Fiber Trenching	tr	Pr_{tr}
Optical Line Terminal (OLT)	OLT	Pr_{OLT}
Optical Network Unit (ONU)	ONU	Pr_{ONU}
10 GBASE transceiver	$trans10$	$Pr_{trans10}$
40 GBASE transceiver	$trans40$	$Pr_{trans40}$
Passive Splitter-Combiner (PSC)	PSC	Pr_{psc}
Fusion splices	sp	Pr_{sp}
WDM Multiplexer	mux	Pr_{mux}

the N_{eN} eNodeBs as follows

$$\begin{aligned} C_{bk} &= Fiber\ Cost + trenching\ cost + Transceiver\ Cost + PSC\ cost + splicing\ cost \\ &= C_{fiber} + C_{trench} + C_{trans} + C_{psc} + C_{splice} \end{aligned} \tag{4.4}$$

where the five terms are given by the following equations respectively.

$$\begin{aligned} C_{fiber} &= \ell \times (2 \times 2.7R \times h \times Pr_d + 5.2R \times Pr_d + 5.2R \times (h-1) \times Pr_f) \\ C_{trench} &= Pr_{tr} \times \ell \times h \times (2 \times 2.7R + 5.2R) \\ C_{trans} &= N_{eN} \times (Pr_{OLT} + Pr_{ONU}) + (N_{eN} - \ell \times h) \times Pr_{trans10} + \ell \times h \times Pr_{trans40} \\ C_{psc} &= \ell \times h \times Pr_{psc} \\ C_{splice} &= Pr_{sp} \times (\frac{\ell \times h \times (2 \times 2.7R + 5.2R)}{2} - 1) + Pr_{splicer} \end{aligned} \tag{4.5}$$

Here the first equation represents the total length of the used fiber in the whole area A multiplied by the price per unit distance of the distribution or the feeder fiber depending on which of them is used. Then, the second equation stands for the total trenching cost and the third equation is

the total transceiver costs which are split into two types based on which type of fiber is used, i.e., 10GB or 40GB. Also, it includes the costs of the OLTs and ONUs. The next equation is the cost of splitters and combiners and eventually the last one represents the splicing cost. The number of fusion splices that are used are estimated to be: $\frac{1}{2} \times$ the number of existing fiber segments -1.
For more insights, we write the five aforementioned cost equations in terms of the number of BSs N_{eN}. Assuming the system model in Figure 5, the total number of BSs in area A can be estimated as

$$\begin{aligned} N_{eN} &= \ell\big(3h+1\big) \\ &= \ell h R\big(\frac{3}{R} + \frac{1}{Rh}\big) \\ &\simeq 3\ell h \end{aligned} \tag{4.6}$$

This is since $\frac{1}{Rh} \lll 1$ so it can be neglected. Now, we can write the cost equations in 4.5 as

$$\begin{aligned} C_{fiber} &= N_{eN} \times R \times [(1.8 + \frac{1.7}{h})Pr_d + \frac{1.7h}{h-1}Pr_f] \\ C_{trench} &= 3.5 \times N_{eN} \times R \times Pr_{tr} \\ C_{trans} &= N_{eN} \times [Pr_{OLT} + Pr_{ONU} + \frac{2}{3}Pr_{trans10} + \frac{Pr_{trans40}}{3}] \\ C_{psc} &= \frac{N_{eN}}{3}Pr_{psc} \\ C_{splice} &= Pr_{sp} \times (1.75N_{eN}R - 1) + Pr_{splicer} \end{aligned} \tag{4.7}$$

From equation (4.7) we can apparently see that the backhaul cost is proportional to the number of eNBs and the cell area (a larger R reflects a larger cell area).
The parameter α in equation (4.2) does not affect the results of the cost model. The reason for this is the dominance of the fiber cost over all other backhaul costs as shown in Figure 8. Meantime, the length of the deployed fiber relies solely on the number of eNodeBs in the CMA/BEA rather than the number of hops h or the number of levels ℓ.

4.2.2 The Backhaul Cost in Ring Topology

We assume the ring topology shown in Figure 9. In this topology, we assume that a ring is constructed between each 4 eNodeBs, i.e., ring cluster $= 4$. Moreover, the nodes shown in blue and

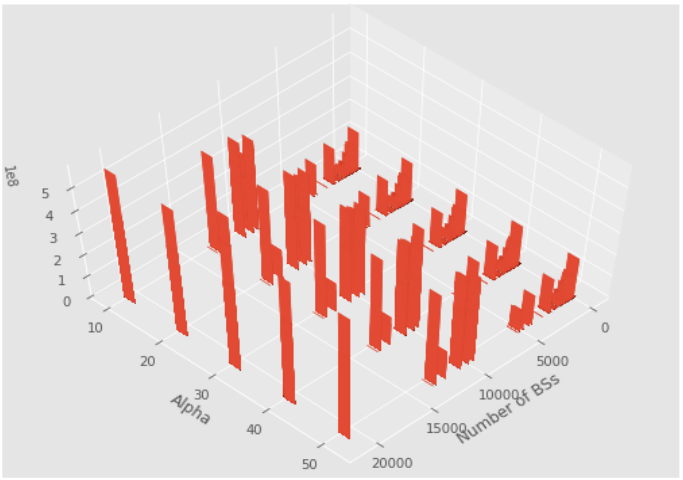


Figure 8: A 3D plot for different backhaul costs Vs. α Vs. N_{eN}

red colors are the aggregating eNBs of the network. Unlike the tree topology, the ring backhaul fiber links are all of the same length, i.e., $= 2.7R$. Moreover, the backhaul cost depends mainly on the ring size, i.e. number of eNBs per a single ring (c_r). The backhaul ring topology changes the backhaul capital fiber cost in 4.5 to be as follows

$$
\begin{aligned}
C_{fiber} &= c_r \times \ell \times 2.7R \times (Pr_d + (h-1) \times Pr_f) \\
C_{trench} &= c_r \times Pr_{tr} \times \ell \times h \times \times 2.7R \\
C_{trans} &= (c_r - 1) \times \ell \times h \times (Pr_{OLT} + Pr_{ONU}) \\
&\quad + \ell \times h \times (c_r - 2) \times Pr_{trans10} + \ell \times h \times Pr_{trans40} \\
C_{psc} &= \ell \times h \times Pr_{psc} \\
C_{splice} &= Pr_{sp} \times (\frac{c_r \times \ell \times h \times 2.7R}{2} - 1) + Pr_{splicer}
\end{aligned} \tag{4.8}
$$

In the ring topology scenario, the number of eNBs covering an $\ell \times h$ area is given as

$$N_{eN} \simeq (c_r - 1) \times \ell \times h \tag{4.9}$$

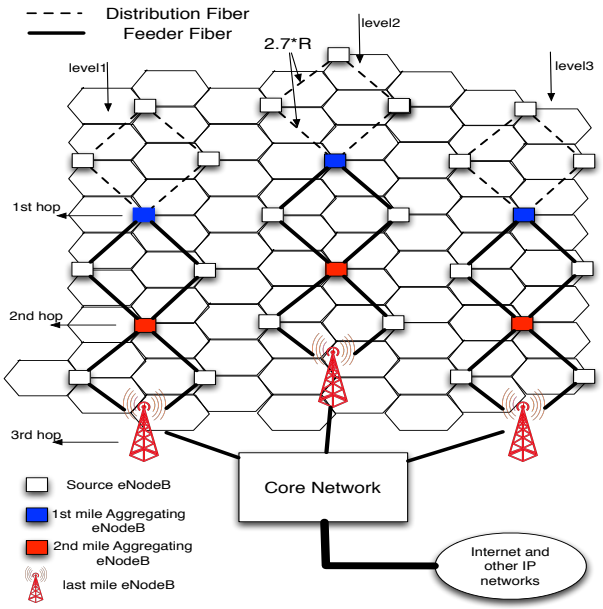


Figure 9: A Schematic Diagram for 3 hops by 3 levels backhaul ring topology in no-sharing scenario

From equation (4.9) we can see that the number of eNodeBs is mainly dependent on the ring size (c_r). Hence, c_r is the dominant factor to estimate the fiber and trenching costs. Moreover, when $c_r = 4$, the ring topology backhaul cost is the same as the tree topology cost.

4.2.3 The Operational Cost

The operational cost (OPEX) composes mainly of the man power, total power consumption, the maintenance and the rental costs. In our model, we assume that the MNO/SP/VE rents the site, the RAN and the core network. On another side, the consumed power in an eNodeB results from the transceiver (P_{Tr}), the rectifier (P_{rec}), the digital signal processor (P_{dsp}), the power amplifier (P_{pa}), the microwave dish (P_{mw}) and air cooling (P_{air}) [65]. Given that the number of sectors deployed in a BS is N_{sec} and the number of antennas is N_{ant}, hence the total power consumption

in $(watts.hour)$ Pow_T can be written as

$$\begin{aligned} Pow_T =& N_{sec} \times (N_{ant} \times (P_{Tr} + P_{pa}) + P_{rec} + P_{dsp}) \\ &+ P_{mw} + P_{air} \end{aligned} \tag{4.10}$$

Then, given the price per watt.hour, it is easy to calculate the total power consumption cost.

Furthermore, maintaining the BS periodically is mandatory to avoid abrupt failures. It is well known that MNOs hire subcontractors for maintaining infrastructure devices. Assuming the time taken to repair a certain device k is the Mean Time To Repair per month ($MTTR_k$) in hours, the total man power used in such job is M_{maint} and the salary per worker per hour is S_w, then the total maintenance cost can be written as

$$C_{maint} = \sum_k (M_{maint} S_w MTTR_k + \delta_k C_{rep,k}) \tag{4.11}$$

where $C_{rep,k}$ is the cost of the replaced item in the infrastructure and δ_k is a binary variable that takes 1 if item k is replaced and 0 otherwise.
Accordingly, the annual OPEX can be written as the sum of the costs in equations (4.10, 4.11) in addition to the man power cost and the rental cost, i.e.

$$C_{op} = 12 \times (N_{mac} \times (C_{maint} + Pow_T + C_{RAN} + C_{core}) + M_{em} S_{em}) \tag{4.12}$$

where M_{em} is the number of employees working for the MNO, S_{em} is the salary per employee per month, C_{RAN} and C_{core} are the rental cost of the RAN and core networks, respectively. Note that the total OPEX cost derived in equation (4.12) is calculated per year.

4.2.4 The Revenue of A Single MNO

Generally, the net profit is defined as the difference between the revenue and the cost for a certain entity. The main source of revenue is selling mobile services including voice, data and video to subscribers using a certain tariff per MegaByte (MB). Normally, mobile operators have multiple mobile plans for mobile users. Such plans are categorized according to the maximum allowable MB to consume per month. Fundamentally, a plan is composed of a fixed subscription fees f, a variable subscription cost V which changes according to the user plan and the excess consumption cost $\$X$ per MB. Hence, the annual revenue of an MNO in the traditional scenario can be written as

$$Rev = 12 \times N_{mac} \times N_{sub} \times (f + V.P^T + Pr_{mb}(M_{avg} - M_v)) \tag{4.13}$$

where N_{sub} is the average number of subscribers belong to the VNO at a single macro BS, V is the vector of all possible subscription categories according to the maximum consumption of MB per month M_v, P is the probability vector which includes the probability of using each of the aforementioned subscription categories (P^T is the transpose of the row vector P), M_{avg} is the average MB consumption per user per month. Also, Pr_{mb} is the per excess MB price.

Accordingly, the annual profit the MNO can be written as

$$Prof = Rev - (C_{bk} + C_{op} + C_{sp}) \tag{4.14}$$

where C_{sp} is the total cost of purchasing spectrum bands operating in this area A.

4.3 A COST MODEL FOR BACKHAUL SHARING SCENARIOS IN A GREENFIELD NETWORK

In this section, we discuss the possible changes in the cost model of the network in case that one of the backhaul sharing scenarios is applied. We assume that two MNOs/VEs/SPs (MNO 1, MNO 2) share either their backhaul networks such that both MNOs both MNOs are operating on the same frequency band, e.g., 700 MHz, in a certain area A. The cost model is essentially based on a greenfield deployment; that is, both MNOs start deploying the network infrastructure from scratch.

4.3.1 Backhaul Sharing Assumptions

In the backhaul sharing scenario, both MNOs agree to share the backhaul network in a certain area A. Since we consider a greenfield network then both MNOs build their networks from scratch, i.e., each of them has his own eNodeBs. However, in our model we assume that the sharing MNOs agree together to share the cost of the eNodeBs instead of replicating eNodeBs since they plan for the same coverage. Hence, both MNOs share the eNodeBs, the ONUs, the OLTs and the PSCs which decreases the total backhaul cost for each of the MNOs. Yet, the OPEX remains unaffected since each MNO has his own operational plan, i.e., maintenance schedule, business and marketing plan and etc.

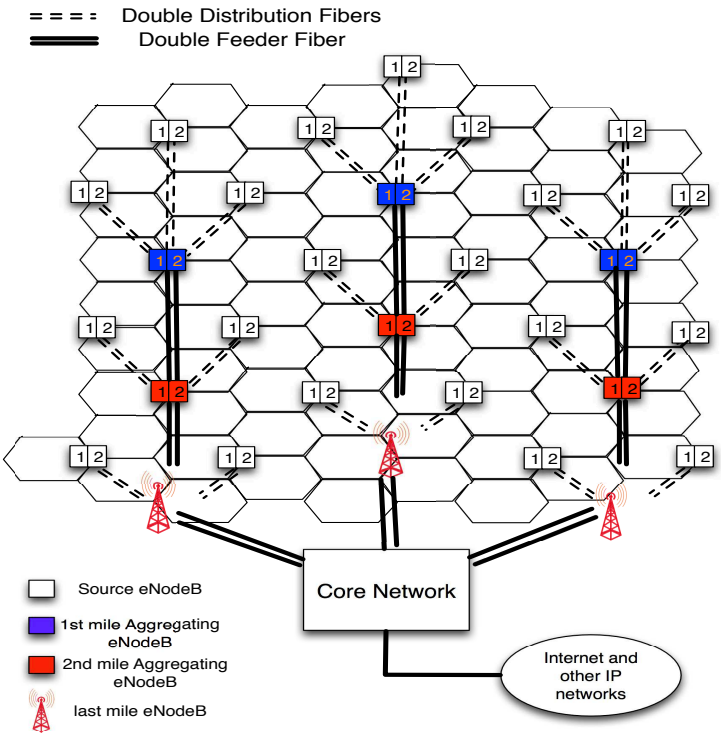


Figure 10: A Schematic Diagram for 3 hops by 3 levels LTE network in backhaul-sharing scenario with/without WDM technology

4.3.1.1 Sharing in An Intra-band Scenario

In intra-band scenario, both MNOs use the same band, i.e. 700 MHz (or 2100 MHz), and hence they have the same coverage through the area A and therefore the same number of eNodeBs N_{mac} as shown in Figure 10. We address the backhaul sharing scenario in two different cases, i.e. when no WDM (Wavelength Division Multiplexing) is applied and when WDM technology is applied. We limit the analysis here to tree topology only since the ring topology is subject to the same changes.

In NO-WDM case, each of the MNOs has to deploy his own fiber cables, yet they can share other optical devices, i.e. optical transceivers, PSCs, ... etc. On the other hand, if the WDM technology is applied, MNOs can share a single fiber cable that can carry two different wavelengths, i.e. the same as Figure 10 but with single fiber deployment instead of the double deployment. Meanwhile, in the latter case, a WDM multiplexer has to be added at each eNodeB to multiplex and demultiplex signals at their times of arrival. Accordingly, the backhaul cost in equation (4.7) is changed to be as follows. If the WDM technology is not applied, the backhaul cost can be written as

$$
\begin{aligned}
C_{fiber}^{bk-sh} &= N_{eN} \times R \times [(1.8 + \frac{1.7}{h})Pr_d + \frac{1.7h}{h-1}Pr_f] \\
C_{trench}^{bk-sh} &= 3.5 \times N_{eN} \times R \times Pr_{tr} \\
C_{trans}^{bk-sh} &= \frac{\mathbf{N_{eN}}}{\mathbf{2}} \times \\
&\qquad [Pr_{OLT,bk} + Pr_{ONU,bk} + \frac{2}{3}Pr_{trans10,bk} + \frac{Pr_{trans40,bk}}{3}] \\
C_{psc}^{bk-sh} &= \frac{\mathbf{N_{eN}}}{\mathbf{6}}Pr_{psc,bk} \\
C_{splice}^{bk-sh} &= Pr_{sp} \times (1.75N_{eN}R - 1) + Pr_{splicer}
\end{aligned}
\tag{4.15}
$$

where the bold terms are the modified terms from the single MNO equations in (4.7). In the third term, the cost of the ONUs, OLTs transceivers are changed to be $Pr_{ONU,bk}$, $Pr_{OLT,bk}$, $Pr_{trans10,bk}$ and $Pr_{trans40,bk}$. $Pr_{ONU,bk}$ and $Pr_{OLT,bk}$ are the prices of the ONU and OLT with higher number of ports. Clearly, increasing the number of ports in the ONUs and OLTs to accommodate traffic from both MNOs is less costly than duplicating the number of ONUs and OLTs in eNodeBs. Consequently, the costs of the ONUs, OLTs and transceivers are halved between both MNOs as indicated by $\frac{\mathbf{N_{eN}}}{\mathbf{2}}$. Similarly, in the fourth term, the cost of PSC is divided equally between MNOs.

Eventually, in the last term of equation (4.15), we replace the traditional PSCs by PSCs with higher number of ports.

Now, the backhaul cost in this sharing scenario can be written as

$$C_{bk}^{bk-sh} = C_{fiber}^{bk-sh} + C_{trench}^{bk-sh} + C_{trans}^{bk-sh} + C_{psc}^{bk-sh} + C_{splice}^{bk-sh} \tag{4.16}$$

Further, we can write the total profit for an MNO in the backhaul sharing scenario as

$$Prof^{bk-sh} = Rev - (C_{bk}^{bk-sh} + \frac{C_{op}}{2} + C_{sp}) \tag{4.17}$$

where the operational cost is divided equally between MNOs as shown by the third term in the equation. In the WDM case, only a single fiber backhaul is deployed since it can be shared via both MNOs using WDM technology. Yet, the cost of the added WDM multiplexers has to be considered while reevaluating the backhaul cost, i.e.

$$\begin{aligned}
C_{fiber}^{bk-wdm} &= \frac{\mathbf{N_{eN}} \times \mathbf{R}}{\mathbf{2}} \times [(1.8 + \frac{1.7}{h})Pr_d + \frac{1.7h}{h-1}Pr_f] \\
C_{trench}^{bk-wdm} &= \frac{\mathbf{3.5}}{\mathbf{2}} \times N_{eN} \times R \times Pr_{tr} \\
C_{trans}^{bk-wdm} &= \frac{\mathbf{N_{eN}}}{\mathbf{2}} \times [Pr_{OLT,bk} + Pr_{ONU,bk} + \frac{2}{3}Pr_{trans10,bk} + \frac{Pr_{trans40,bk}}{3}] \\
C_{psc}^{bk-wdm} &= \frac{\mathbf{N_{eN}}}{\mathbf{6}}Pr_{psc,bk} \\
C_{splice}^{bk-wdm} &= Pr_{sp} \times (1.75N_{eN}R - 1) + Pr_{splicer} \\
&+ \frac{N_{mac} \times Pr_{mux}}{2}
\end{aligned} \tag{4.18}$$

where Pr_{mux} is the price of a WDM multiplexer unit. Moreover, the first and the second terms are divided by 2 since the total fiber cost should be split into half between both MNOs. And now the backhaul cost in this WDM scenario can be written as

$$C_{bk}^{bk-wdm} = C_{fiber}^{bk-wdm} + C_{trench}^{bk-wdm} + C_{trans-wdm}^{bk-wdm} + C_{psc}^{bk-wdm} + C_{splice}^{bk-wdm} \tag{4.19}$$

In analog to equation (4.17), the profit of an MNO in the WDM-backhaul sharing scenario is

$$Prof^{bk-wdm} = Rev - (C_{bk}^{bk-wdm} + \frac{C_{op}}{2} + C_{sp}) \tag{4.20}$$

4.3.1.2 Sharing in An Inter-band Scenario In this scenario, we assume that the MNOs are operating in different bands, MNO 1 in 700 MHz LTE band, i.e., MNO-700, and MNO 2 in 2100 MHz AWS band, i.e. MNO-2100. Since different frequency bands leads to different coverage areas, MNOs adjust the number of eNBs to ensure a full coverage of the whole area A. Figure 11 shows a 3 hops $\times$ 3 levels example of an LTE backhaul sharing scenario. In this figure, the black cells (black hexagons) are for MNO-700 and the red cells are for the MNO-2100. The circles in black (encircling the purple eNodeBs) are the additional micro cells which the MNO-2100 has to add to have a full coverage of this area. This is because MNO-2100 has less cell coverage because higher frequencies have shorter propagation ranges.

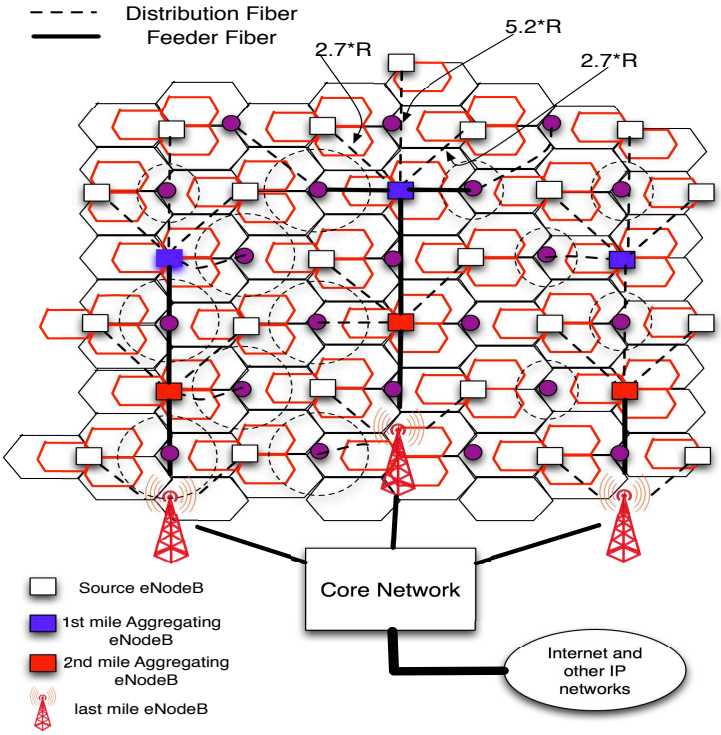


Figure 11: A Schematic Diagram for 3 hops by 3 levels LTE network in backhaul-sharing scenario with different bands (with WDM)

Again, the backhaul cost is similar to equations (4.15 and 4.18) for MNO 1 (700 MHz) in the no-WDM and WDM cases respectively. Yet, the micro cells additional cost adds to MNO2 (2100

MHZ) backhaul cost, i.e.

$$
\begin{aligned}
C_{bk,diff}^{bk-sh,2} = C_{bk}^{bk-sh} + & \underbrace{\ell \times (2h+1) \times R \times Pr_d}_{\text{add. fiber row cost}} \\
& + \underbrace{Pr_{tr} \times \ell \times (2h+1) \times R}_{\text{add. trenching cost}} \\
& + \ell \times (2h+1) \times (Pr_{OLT,bk} + Pr_{ONU,bk} \\
& + \underbrace{\ell \times (2h+1) \times Pr_{trans10}}_{\text{transceiver cost}} \\
& + \underbrace{Pr_{sp} \times (\frac{\ell \times (2h+1) \times R}{2} - 1)}_{\text{splicing cost}}
\end{aligned}
\tag{4.21}
$$

where the second term represents the row fiber cost for the additional eNodeBs. It is worth noting that we assumed that the backhaul of each of the added micro cells are deployed such that it takes the shortest path, i.e. $length = R$, and hence the lowest cost. Also, the OPEX changes for MNO-2100 to include the additional cost of micro eNodeBs, i.e.

$$
\begin{aligned}
C_{op,diff}^{bk-sh,2} = & 12 \times \Big((N_{mac} \times (C_{RAN}^{mac} + C_{core}^{mac}) \\
& + \ell \times (2h+1) \times (C_{RAN}^{mic} + C_{core}^{mic})) \times (C_{maint} + Pow_T) \\
& + M_{em} S_{em} \Big)
\end{aligned}
\tag{4.22}
$$

where C_{RAN}^{mic} and C_{core}^{mic} are the rental costs for both the RAN and the core networks for a micro eNodeB which should be less than the rental costs for a macro eNodeB, i.e., C_{RAN}^{mac} and C_{core}^{mac}. Now, the profit in this scenario can be found similar to equation (4.17).

Eventually, to reevaluate the profit in WDM based backhaul network, we replace C_{bk}^{bk-sh} in equation (4.21) by C_{bk}^{bk-wdm} from equation (4.19). Meanwhile, the backhaul cost of MNO-700 will be the same as equation (4.19).

4.3.2 Infrastructure Sharing in Ring Topology

Similar to the infrastructure sharing in tree topology, the cost of sharing the ring topology infrastructure can be estimated as

$$
\begin{aligned}
C_{fiber} &= \frac{\mathbf{1}}{\mathbf{2}} c_r \times \ell \times 2.7R \times (Pr_d + (h-1) \times Pr_f) \\
C_{trench} &= \frac{\mathbf{c_r}}{\mathbf{2}} \times Pr_{tr} \times \ell \times h \times \times 2.7R \\
C_{trans} &= \frac{(\mathbf{c_r - 1})}{\mathbf{2}} \times \ell \times h \times (Pr_{OLT} + Pr_{ONU}) \\
&\quad + \ell \times h \times (c_r - 2) \times Pr_{trans10} + \ell \times h \times Pr_{trans40} \\
C_{psc} &= \frac{\ell \times \mathbf{h} \times \mathbf{Pr_{psc}}}{\mathbf{2}} \\
C_{splice} &= \frac{\mathbf{1}}{\mathbf{2}} Pr_{sp} \times (\frac{c_r \times \ell \times h \times 2.7R}{2} - 1) + \frac{\mathbf{Pr_{splicer}}}{\mathbf{2}}
\end{aligned}
\tag{4.23}
$$

4.4 SPECTRUM SHARING SCENARIOS

In the spectrum sharing case, we assume spectrum sharing in a greenfield network. Based on the traffic demand at a certain CMA/BEA the network operator needs to plan the capacity and the coverage of his network to estimate the bandwidth to be shared. Also, we consider two cases: sharing spectrum only (no infrastructure sharing) and full sharing (spectrum and infrastructure sharing). Assuming a greenfield network, a backhaul cost C_{bk}^{ssh} and an operational cost C_{op}^{ssh}, then the profit in the spectrum sharing scenario can be written as

$$
Prof^{ssh} = Rev^{ssh} - (C_{bk}^{ssh} + C_{op}^{ssh} + C_{sp}^{ssh}) \tag{4.24}
$$

where C_{sp}^{ssh} is the spectrum cost in the spectrum sharing scenario and is calculated as

$$
C_{spi}^{ssh} = C_{spi} + N_{bs-sh} \times Pr_{MHz}^{j*} \times b_j^{sh} \tag{4.25}
$$

where C_{spi} is the basic spectrum cost for the spectrum owned by MNO i and Pr_{MHz}^{j*} is the spectrum price per MHz of a total leased bandwidth by MNO j (b_j^*) in a certain CMA/BEA area. Also,

N_{bs-sh} is the number of BSs adopting spectrum sharing in the area A. Note that Pr^{j*}_{MHz} and b^*_j are the optimal price per MHz and the optimal leased Bandwidth, respectively, such that the MNO sharing profit is higher than the single MNO profit for both MNOs, i.e.,

$$\begin{aligned} &\underset{Pr_{MHz},bi,b_j}{\text{Maximize}} && Prof_i^{ssh} + Prof_j^{ssh} \\ &\text{subject to} \\ & && Prof_i^{ssh} > \mu \times Prof_i^{single}, \quad k \in \{i,j\} \\ & && Pr^k_{MHz} > 0, \quad k \in \{i,j\} \\ & && b_k \geq \epsilon, \quad k \in \{i,j\} \end{aligned} \tag{4.26}$$

where ϵ is the minimum feasible bandwidth that can be shared, e.g. the bandwidth of two physical resource block (PRB) in LTE systems, and μ is a tuning parameter to control the spectrum sharing target incentive. Also, C^{ssh}_{bk} and C^{ssh}_{op} are the backhaul and operational costs in the spectrum sharing scenario, respectively, according to the sharing scenario as aforementioned in Sections 4.2 and 4.3. In this case, the values of C^{ssh}_{bk} and C^{ssh}_{op} in equation (4.24) are substituted by the costs of a single MNO given in equations (4.6) and (4.12), respectively. In this scenario, both spectrum and infrastructure are shared between MNOs. Thus, the costs for the CAPEX and OPEX are given by the sharing scenarios in equations (4.7 and 4.8) based on which topology is used in the backhaul (tree or ring).

4.5 NUMERICAL RESULTS

In this section, we show different results for the different sharing scenarios that we addressed in Section 4.3 and 4.4. We used **Fiber Store** [58] to get the monetary values of different fiber network components as shown in Table 2.

It is worth noting that K-PON OLT-GPON is the price of the K-ports OLT unit if used for the Gigabit Passive Optical Network(GPON) and similarly K-PON OLT-EPON for the Ethernet Passive Optical Network(EPON). Similarly, K-PON ONU-GPON is the price of K-ports ONU unit. Also, $1\times$ K PSC is the price of a $1:K$ splitter-combiner unit. Also, note that the trenching cost

Table 2: Cost values of the fiber network components

component	cost
4-PON ports OLT - GPON	$ 4275
4-PON ports OLT - EPON	$ 1352
8-PON ports OLT - GPON	$ 5063
8-PON ports OLT - EPON	$ 2292
4-PON ports ONU - GPON	$ 116
8-PON ports ONU - GPON	$ 384
$1\times$ 4 PSC	$ 30
$1\times$ 6 PSC	$ 40
$1\times$ 8 PSC	$ 55
10GBASE SFP^+ transceiver	$ 82
40GBASE QSFP transceiver	$ 750
splicer	$ 30000
fusion splice	$ 9

varies according to whether the deployment area is a rural area or an urban area [50]. Microtrenching, horizontal directional drilling (HDD) and manual trenching are well suited for urban areas, and cost approximately $50, $37 and $25 per meter respectively. The term microtrenching is used to emphasize the small size of trench in the ground where the cable is deployed, and is typically in the range of 1.5- 5cm in width and 20-40 cm in depth. HDD on the other hand, is a completely trenchless cable deployment method. Here, an inner duct is placed in a drilled path and the cable is subsequently pulled through. Both techniques are extremely effective in areas where significant disruption of surface is undesirable. However, if cost is the main concern, manual trenching could be used instead. On another note, in rural areas, more cost-effective techniques, such as direct cable burying (plouging), machine trenching (in areas not suitable for ploughing), and aerial deployment could be used instead and at a cost of approximately $7, $14 and $4 per meter, respectively. In the following numerical results, we assume the use of microtrenching in urban areas with a cost of $50 per meter. On the other side, in rural areas we assume direct cable burying with a cost of $7 per meter. Figure 12 shows the contribution of different costs in the total cost of the cellular network.

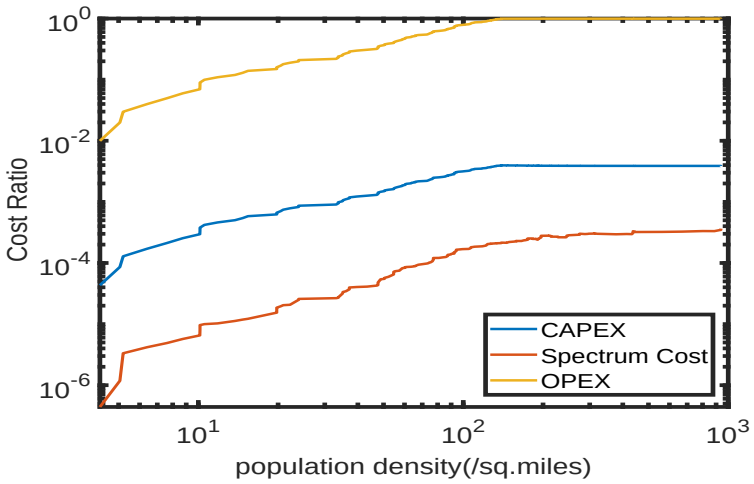


Figure 12: Cost Contribution of OPEX, CAPEX and Spectrum costs in the total cost of establishing a green-field cellular network

It is obvious that the OPEX is the dominant factor. Also, the spectrum net bidding amount is the least contributing to the total cost of the LTE network.

In this section, we first analyze the different approaches to derive the number of eNBs, i.e. set cov-

erage problem, approximation-based and real data based. Moreover, we show how the optimum price/MHz and the optimum shared BW change in the spectrum sharing scenario after applying the optimization formula in equation (4.26). Afterwards, we analyze and compare the proposed cost model in different sharing scenarios, i.e. backhaul, backhaul-WDM and spectrum sharing. Moreover, we get insights about the efficacy of the population density, the number of BSs on the net profit by using a dataset of about 1054 CMA and BEA areas in US that was derived from various FCC auction data sheets [57]. The FCC data composes of the census, the spectrum net cost, the area (in mile.sq), the used bands and bandwidths in different CMA/BEAs across US. We produce the results for tree and ring topology.

Profits are calculated as a net present value assuming a cash flow period of 15 years with an initial capital cost of $C_{bk} + C_{sp}$, i.e. the initial investments in backhaul and spectrum, and an annual interest rate of 5% [66].

4.5.1 Different Criteria to derive number of eNodeBs

Using the criteria proposed in Section 4.1.1, we derive the number of eNBs in each CMA and we compare the cumulative frequency of occurrences as shown in Figure 13. In this figure we can see that N_{eN} for the approximation approach is lower than the actual real data while the results of the set coverage problem approach high values compared to the real and approximated values. It is worth noting that we removed the outliers from the data before analyzing it to understand the effect of the different criteria clearly. In the upcoming results, we use the real values for N_{eN} unless otherwise is mentioned.

4.5.2 Optimum Spectrum Bandwidth and Price in Spectrum Sharing Scenario

Using the optimization formula in equation (4.26), we derive the optimum BW to be shared and the leasing price/Hz in spectrum sharing scenario. Figures 14a and 14b show the optimum shared BW versus the percentage of BSs that adopt sharing in a certain area A for two different areas with low and high population densities, respectively. The blue curve represents spectrum sharing only scenario wherein the red curve represents the spectrum plus backhaul sharing scenario (sharing both the backhaul and the spectrum between SPs). It is worth noting that at the cases where the optimum

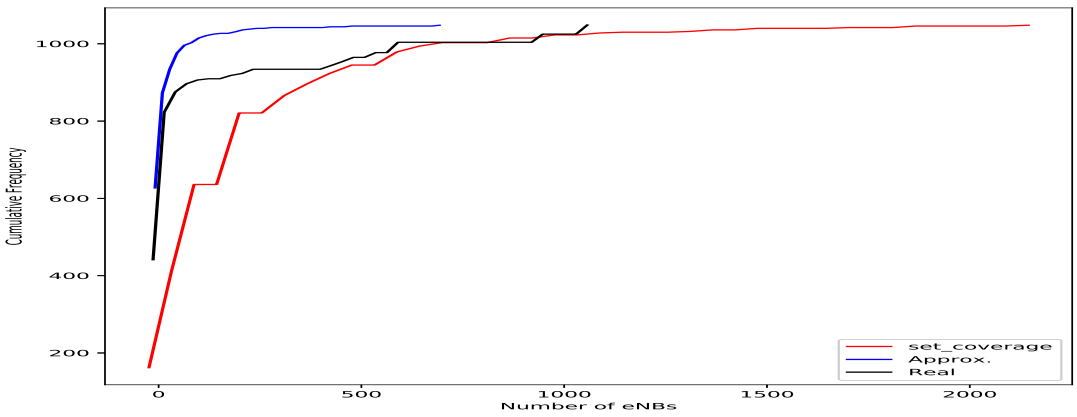
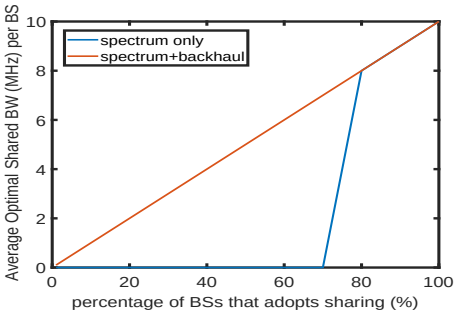


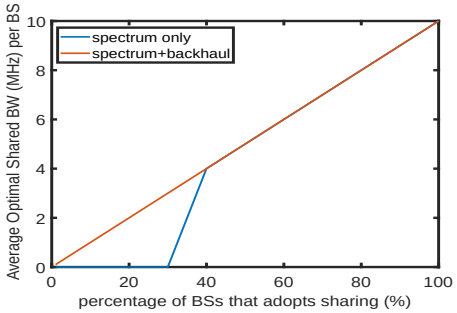
Figure 13: The cumulative frequency for $N_{eN} assuming different criteria$

BW is zero, it is an indication that the SP does not have the incentive to share the spectrum. This is due to the in-feasibility of the profit constraint in the optimization problem in equation (4.26), i.e. the profit of sharing the spectrum does not exceed the no-sharing profit. However, this sharing incentive increases with the increase of the number of BSs that adopt spectrum sharing. Moreover, the threshold at which the sharing incentive turns into a non-zero monetary value is correlated with the population density at area A. For areas with higher population densities the incentive threshold can be attained more faster than areas with lower population densities as depicted by the blue curves in Figures 14a and 14b.

Similarly, Figures 15a and 15b show the profit per SP of sharing the spectrum between two identical SPs in the 700 MHz A band. It is obvious that the spectrum plus backhaul sharing outperforms the spectrum sharing only scenario at all values of population densities and all percentages of shared BSs. It is important to notice that a zero profit in the figures reflects a zeros incentive but not necessarily a zero monetary value, i.e. we nullified the non-zeros profits that does not exceed the target no-sharing profit.

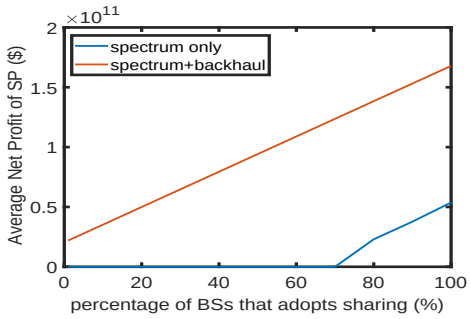


(a) Low population density = 1000 per sq.miles

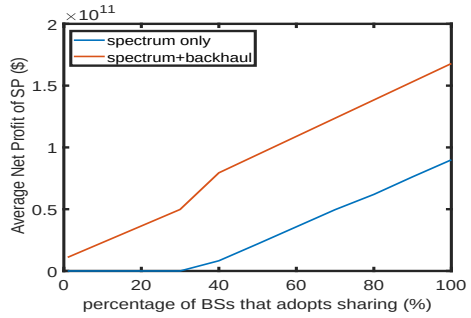


(b) High population density = 6500 per sq.miles

Figure 14: Optimum Shared BW versus the percentage of BSs deploying spectrum sharing (N_{sh-bs}) in the 700 MHz A band



(a) Low population density = 1000 per sq.miles



(b) High population density = 6500 per sq.miles

Figure 15: SP profit versus the percentage of BSs deploying spectrum sharing in the 700 MHz A band

4.5.3 Comparing The profits of Different Sharing Scenarios

4.5.3.1 The Net Profit Versus The number of BSs per CMA/BEA Using the FCC datasheets, we are able to derive the net profit for 1054 CMA/BEAs across US in different sharing scenarios by applying the cost models proposed in Section 4.2 of this paper. Figures 16a, 16b and 16c show the net profit versus the number of BSs per a CMA area A in different sharing scenarios that deploy a tree-based backhaul topology with $N_{sh-bs} = 10\%$, 50% and 90% respectively. Since it was clear and affirmative that all the sharing scenarios outperform the no-sharing scenario in terms of net profit, we do not show the no sharing curve for better presentation of the results. Interestingly in these figures we can clearly see that backhaul sharing profit outperforms spectrum sharing profit when $N_{sh-bs} = 10\%$ and the total number of eNBs ($N_e N$) is lower than 750. As the number of BSs per CMA/BEA increases, spectrum sharing outperforms backhaul sharing. Moreover, for higher N_{sh-bs}, i.e. $N_{sh-bs} = 50\%$ and 95%, spectrum sharing achieves higher profit than backhaul sharing unconditionally as shown in Figures 16b and 16c. The reason for that is that as the number of BSs increases the revenue from sharing the spectrum increases according to equation (4.25) and hence the spectrum sharing profit increases accordingly. Meanwhile, the backhaul cost decreases in the scenario of backhaul sharing. However, the increase in the revenue due to spectrum sharing outweigh the decrease in the backhaul cost as the number of eNBs per CMA as well as the number of BSs adopting spectrum sharing increases, i.e. in areas with high number of BSs.

Similarly, Tables 3 and 4 shows the ten first and last CMAs arranged with respect to the incentive of backhaul sharing, respectively. Moreover, Tables 5 and 6 shows the first and last 10 CMAs arranged with respect to the incentive of spectrum sharing. Figures 17a, 17b and 17c show the net profit versus the number of BSs in different sharing scenarios for the 2100 MHz (AWS) band with $N_{sh-bs} = 10\%$, 50% and 90% respectively. Unlike the 700 MHz band, we can see that at $N_{sh-bs} = 10\%$ spectrum sharing outperforms backhaul sharing when the number of eNBs exceeds 1400. However, at $N_{sh-bs} = 50\%$ spectrum sharing outperforms backhaul sharing when the number of eNBs exceeds 90. In conclusion, the 2100 MHz band requires higher number of sharing eNBs to achieve higher incentive than the backhaul sharing. Similar to the tree-based results, Figures 18a, 18b and 18c show the net profit versus the number of BSs per a CMA area A in different sharing scenarios that deploy a ring-based backhaul topology. We can see two main

Table 3: Highest 10 CMA/BEA profits of sharing backhaul

CMA/BEA	Increase in profit %
Syracuse, NY	8.3
Carlisle, PA	4.9
State college, PA	3.5
Buffalo, NY	3.3
Niagara Falls, NY	3.3
Augusta, GA	2.6
Aiken, SC	2.6
Harrisburgh, PA	2.4
Lebanon, PA	2.4

Table 4: Lowest 10 CMA/BEA profits of sharing backhaul

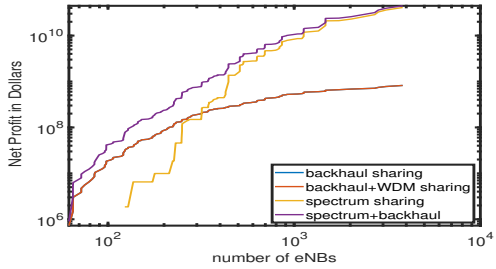
CMA/BEA	Increase in profit %
Great Falls, MT	0.73
Duluth, MN	0.76
Superior, WI	0.76
Bismarck, ND	0.77
Odessa, Tx	0.85
Midland, Tx	0.85
San Angelo, Tx	0.89
Rapid city, SD	0.93
Beaumont, Tx	1

Table 5: Highest 10 CMA/BEA profits of sharing spectrum

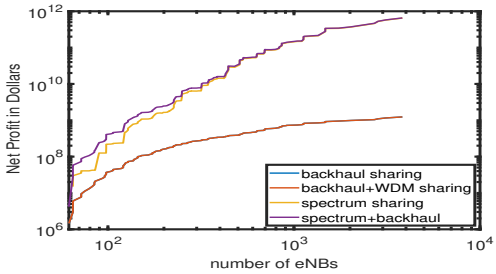
CMA/BEA	Increase in profit %
Boston, MA	100
Philadelphia-Atlantic city	95
Dallas, Tx	75.9
San Francisco, CA	45.6
Oakland, CA	45.6
San Jose, CA	45.6
Tampa, FL	45.0
San Diego, CA	36.9
Seattle, WA	30
Houston, Tx	23.3

Table 6: Lowest 10 CMA/BEA profits of sharing spectrum

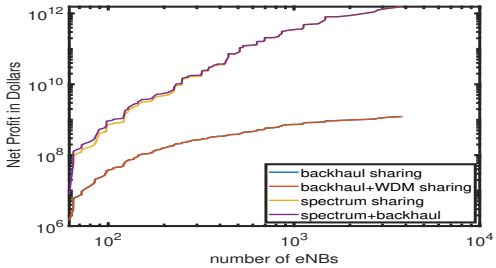
CMA/BEA	Increase in profit %
Odessa, Tx	1.13
Davenport, Iowa	1.68
Charleston, VA	2
Rochester, NY	2.23
Lincoln, NE	2.58
Evansville, IN	3.13
Fortwayne, IN	3.15
Topeka, KS	3.23
Portland, Maine	4.32
Columbia, SC	4.75



(a) $N_{sh-bs} = 10\%$

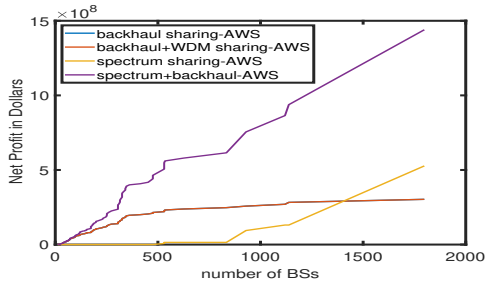


(b) $N_{sh-bs} = 50\%$

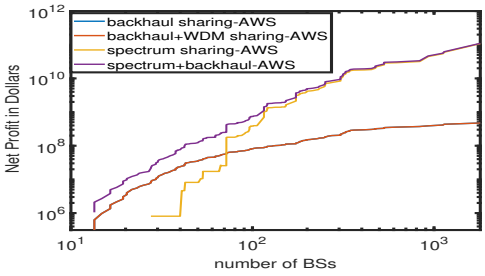


(c) $N_{sh-bs} = 95\%$

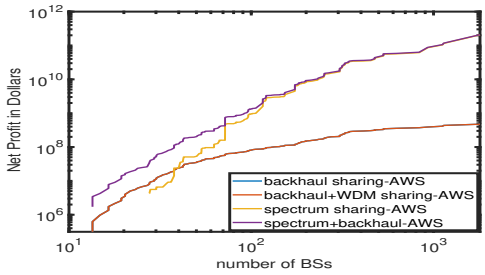
Figure 16: Net Profit versus number of BSs (N_{eN}) in different sharing scenarios for the 700 MHz band



(a) $N_{sh-bs} = 10\%$



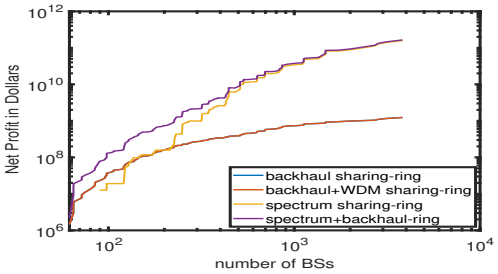
(b) $N_{sh-bs} = 50\%$



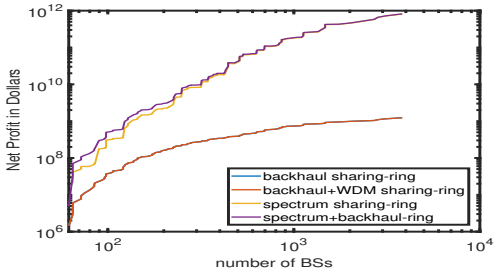
(c) $N_{sh-bs} = 95\%$

Figure 17: Net Profit versus number of BSs (N_{eN}) in different sharing scenarios for the 2100 MHz band

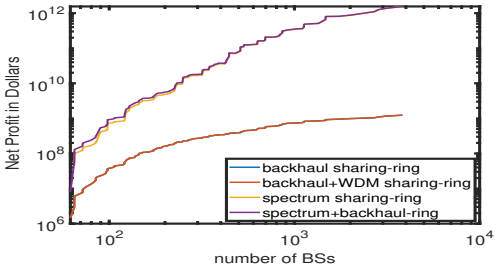
differences from the tree topology. The ring topology profits are almost 10 times greater than the the tree topology profits. Moreover, the threshold where the spectrum sharing achieves better incentive than the backhaul sharing is at $N_{eN} \simeq 200$ instead of 700 for the tree topology.



(a) $N_{sh-bs} = 10\%$



(b) $N_{sh-bs} = 50\%$



(c) $N_{sh-bs} = 95\%$

Figure 18: Net Profit versus number of BSs (N_{eN}) in different sharing scenarios

4.5.4 Mapping The Incentive of Different Sharing Scenarios Across US CMAs/BEAs

In this subsection, we project the results from the cost model in Section 4.2 on the US state map. All the numerical values are pre-processed by standardization and normalization to eliminate outliers. Since the gathered data is normally distributed, we normalized the data by subtracting the mean and dividing by the standard deviation. Figure 19 shows the CMAs where MNOs have incentive to share the network resources (either spectrum or backhaul). The numbers in the legend of the map indicate the percentage of increase in the MNO profit thanks to sharing.

Figures 20 and 21 show the spectrum sharing and backhaul sharing incentive, respectively, in different CMAs in the case of using the real data approach to derive N_{eN}. It is obvious that spectrum sharing is superior to backhaul sharing in the areas where the population density is fairly high such as California, Florida and a big part of the east-coast. However, interestingly, there are some exceptions to this consensus when some of the CMAs/BEAs with high population density experience very high cost per MHz per population such as in New York, Washington D.C., Pennsylvania and West Virginia. This ultra-expensive cost per MHz per population is due to that the spectrum net bidding cost is very high and hence sharing the spectrum adds a high spectrum lease cost to the OPEX. Recall that the spectrum lease rate is proportional to the spectrum cost per MHz per population unit per CMA/BEA. This lease value increases the total LTE network cost significantly since the OPEX cost is a dominant factor in the total cost of deploying a greenfield network as pinpointed before in Figure 12. Therefore, Sharing the backhaul network in such CMAs makes more sense than sharing the allocated spectrum.

Similarly, Figures 22 and 23 show the spectrum and backhaul incentive maps in case of using the set coverage problem to derive N_eN. Note that the numbers in the legend of the set coverage maps are ratios not percentages, i.e. $0.25 = 25\%$ increase in the profit.

4.5.5 Sensitivity Analysis

In this part, we conduct a sensitivity analysis for the profit of sharing in different scenarios. In Table 7, we increase the cost per MHz per population by multiples of 2 and we deduce the threshold of number of BSs and population density at which spectrum sharing incentive exceeds backhaul sharing incentive. We can clealry see that increasing the spectrum cost per MHz still matches the

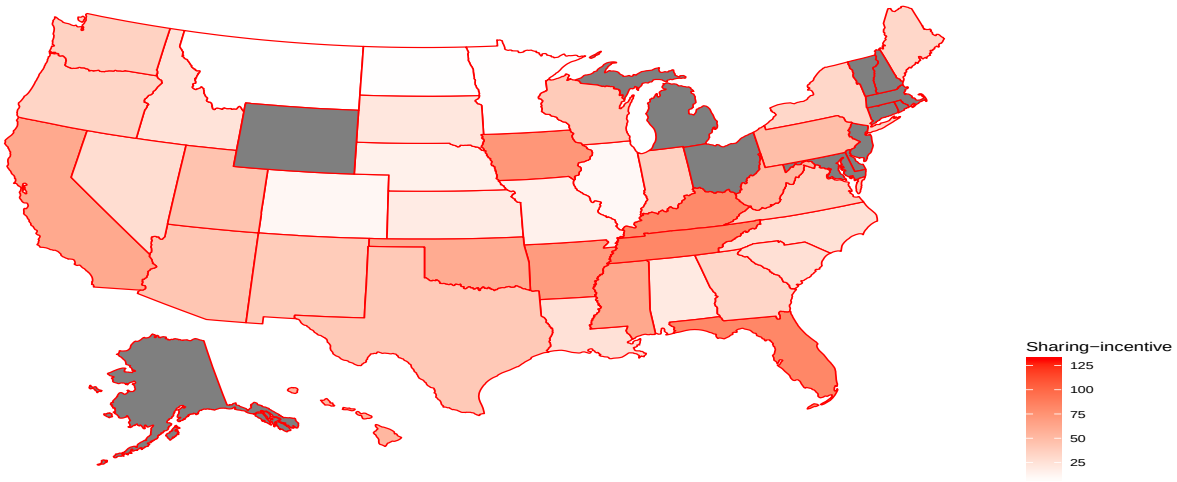


Figure 19: A Map for Sharing Incentive Across US states

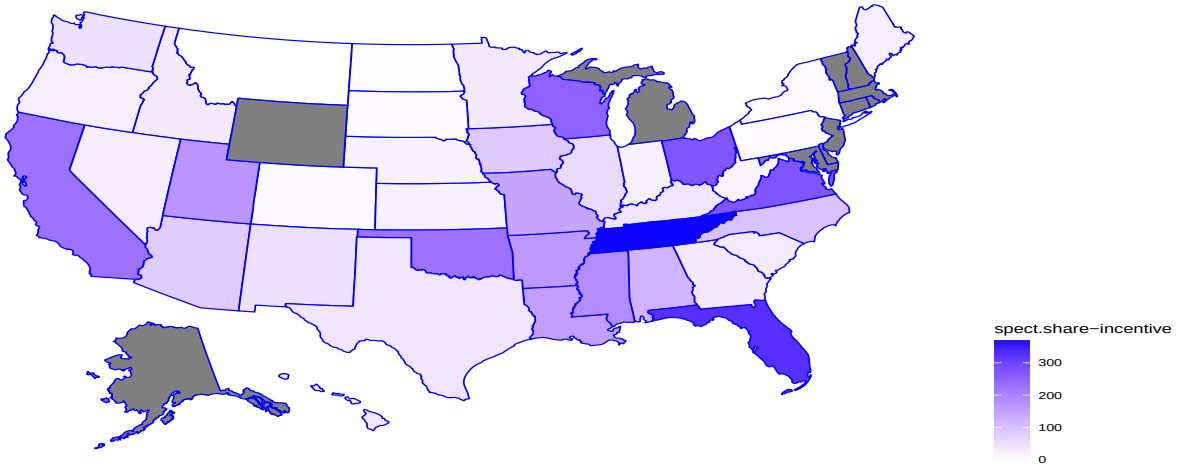


Figure 20: A Map for Spectrum Sharing Incentive Across US states

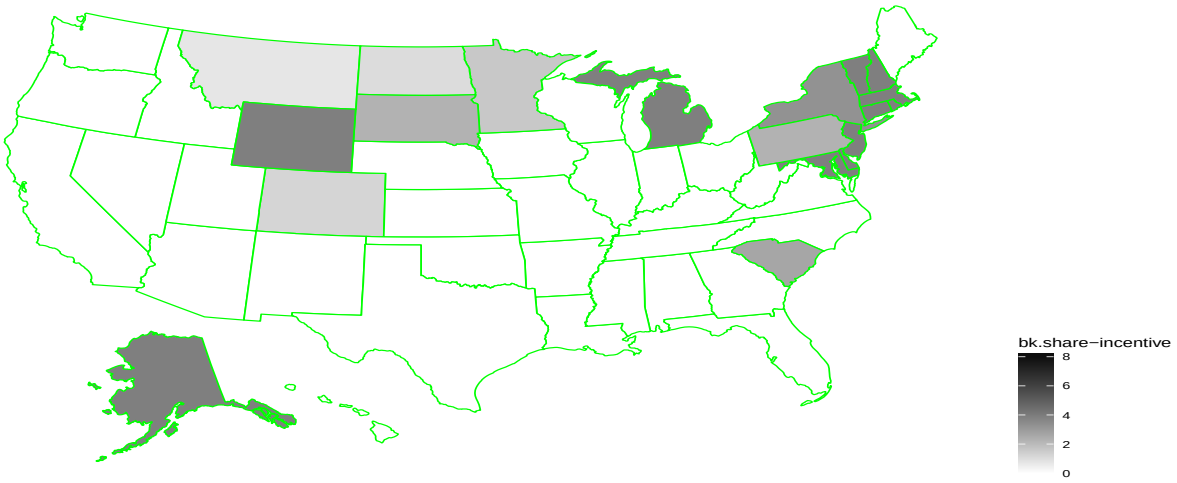


Figure 21: A Map for Backhaul Sharing Incentive Across US states

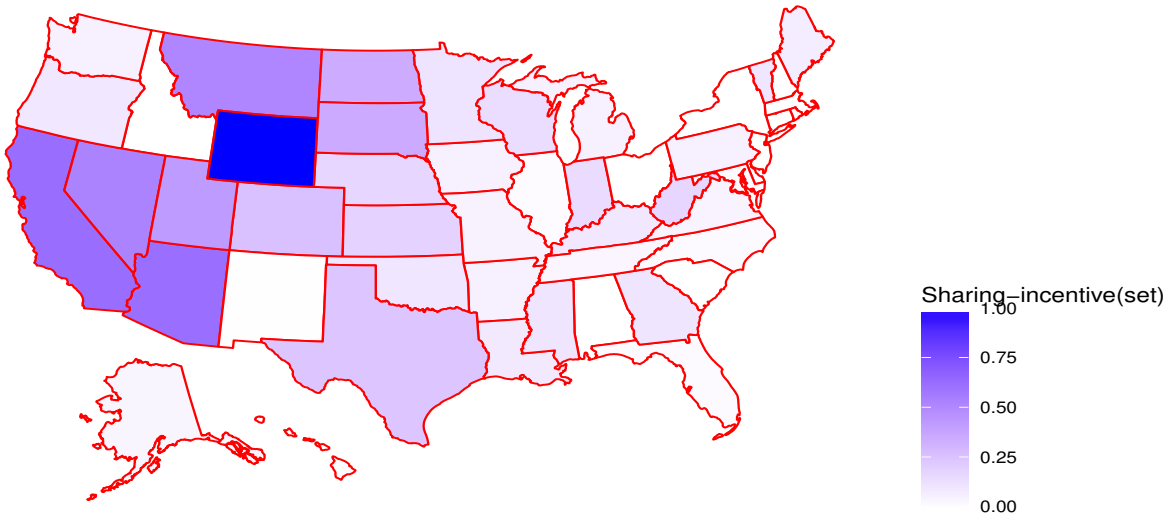


Figure 22: A Map for Spectrum Sharing Incentive Across US states-Set Coverage Approach

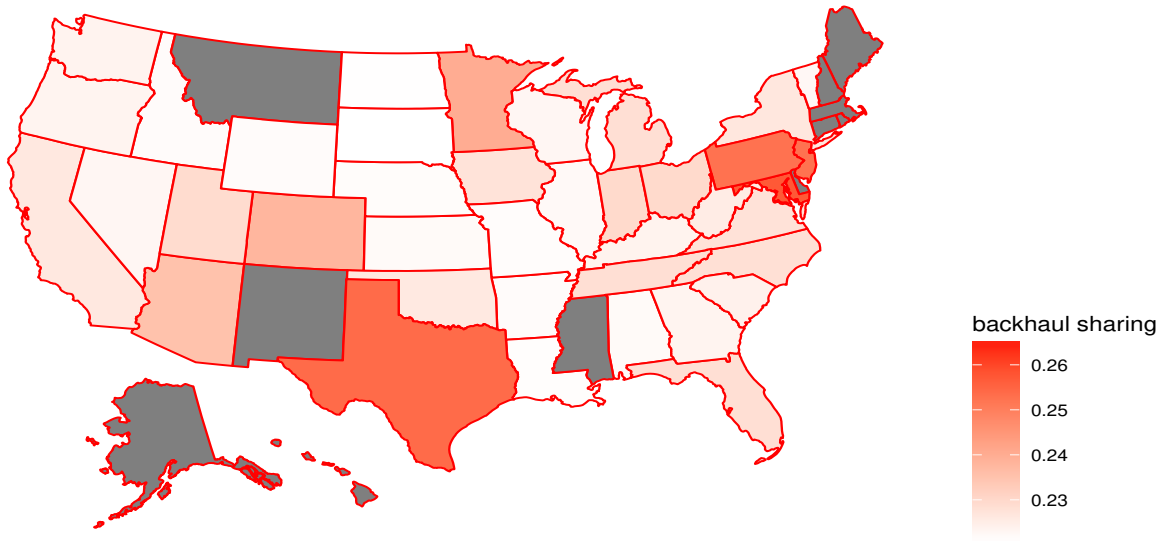


Figure 23: A Map for Backhaul Sharing Incentive Across US states-Set Coverage Approach

concluded consensus about spectrum sharing and backhaul sharing. Yet, the number of sharing eNodeBs decreases for higher spectrum coast per MHz per population, i.e., MNO/VE can achieve spectrum sharing incentive with less number of sharing eNodeBs.

Table 7: Sensitivity Analysis of sharing

spectrum cost per MHz per population	number of eNodeBs threshold	population density threshold
$\times$ 1	1400	1900
$\times 2$	800	1200
$\times 4$	300	1000
$\times 6$	100	700
$\times 10$	0	350

5.0 RESOURCE ALLOCATION IN HETEROGENEOUS VIRTUALIZED NEXT-GENERATION MOBILE NETWORKS

After exploring the economic incentive of sharing the infrastructure and spectrum resources in cellular networks, in this chapter, we propose an efficient scheme for resource allocation in virtualized mobile networks. The scheme handles two main issues: isolation between different MNOs sharing the same resources and the heterogeneity of next-generation mobile networks including new service types such as enhanced mobile broadband (eMBB), ultra-reliable critical communications (URCC) and massive machine type communications (mMTC).

5.1 SYSTEM MODEL AND MOTIVATION

In this section, we provide an overview of the system model and the motivation of the proposed VPS algorithm. Also, we spotlight the lack of ultra-reliability in current LTE cellular networks and hence examine an enhanced version of the VPS algorithm (e-VPS) to handle URLLC traffic in an efficient manner.

5.1.1 System Model and Background

In this chapter, we consider a single cell of size d_{max} served by one BS (or evolved node B – eNB in LTE) which is physically maintained by a certain InP. We assume two SP/VEs, VE1 and VE2 operating on the same eNB. We assume both VEs have mobile users in the cell with the same coverage. Also, we assume both SP/VEs share the entire spectrum (the 700 MHz-LTE bands are aggregated to form a spectrum pool of bandwidth $B = B_1 + B_2$ where B_1 and B_2 are the

bandwidths that are otherwise allocated to SP/VE1 and SP/VE2, respectively).

Since we assume SP/VEs are operating on an LTE-based physical substrate, physical resource blocks are allocated to users based on the channel quality index (CQI). The CQI is used by the eNB to infer the channel conditions for a given mobile in a given set of RBs during a TTI. In this thesis, we compute CQI according to the table in [67]. The corresponding transmission rate per Hz(η) is derived from Table 8 as below.

Table 8: CQI and MCS Table

CQI index	modulation(M-ary)	code rate x 1024	efficiency(η)
0	out of range		
1	QPSK	78	0.1523
2	QPSK	120	0.2344
3	QPSK	193	0.3770
4	QPSK	308	0.6016
5	QPSK	449	0.8770
6	QPSK	602	1.1758
7	16QAM	378	1.4766
8	16QAM	490	1.9141
9	16QAM	616	2.4063
10	64QAM	466	2.7305
11	64QAM	567	3.3223
12	64QAM	666	3.9023
13	64QAM	772	4.5234
14	64QAM	873	5.1152
15	64QAM	948	5.5547

In LTE, RT VoIP traffic is assumed to consist of 40 byte packets that are generated every 20 ms., i.e., the inter-arrival time is 20ms [68]. A fixed number of PRBs are sufficient for RT traffic packets. According to table 8, the lowest CQI corresponds to $CQI = 1$ with a code rate efficiency of 0.1523 bits/Hz. If we denote the RB capacity, i.e. the maximum number of transmitted bits per RB per TTI, by B_{rt}, we can see that B_{rt} can be computed roughly as:

$$\begin{aligned} B_{rt} &= 2 \times 12 \times 15\ KHz \times 0.1523 \text{ bits/Hz} \\ &= 54.83 \text{ Kb} \end{aligned} \tag{5.1}$$

Hence, one RB may be sufficient to serve a RT request (we do not include the overhead and control signaling here). However the LTE standard [67] specifies 2 RBs to be the least number of RBs to be acquired by a single request in a TTI.

We also note here that since the lowest CQI can still serve RT requests, it will be beneficial to allocate the worst channels to RT traffic and save the best channels, which have higher data rates for non real-time traffic. We explore this later.

5.1.2 Motivation

In this part, we explore the motives for designing the VPS and e-VPS resource allocation schemes. We start with reasoning for the joint scheduling of requests among different VEs. Then, we show that high reliability is hard to attain in current LTE systems.

5.1.2.1 Problems with Separate Scheduling for SP/VEs We examine why separate scheduling of RT traffic flows *after* allocation of RBs may result in more blocked RT requests through a simple example. In the framework shown in Figure 24, the Network Scheduler is in charge of estimating the number of spectrum resources that should be assigned to each SP/VE based on the proportional fairness allocation scheme. This scheme operates as follows:

$$\begin{aligned} \underset{c_g}{\text{maximize}} \quad & \sum_{g} \upsilon(\Gamma_g) \\ \text{subject to} \quad & \Gamma_g \geq Q^g_{min} \forall g \end{aligned} \tag{5.2}$$

where $\upsilon(\Gamma_g)$ is the utility function which is maximized to attain the best resource allocation in a TTI, $\Gamma(g)$ is the total transmission rate in a TTI for VE g and g is the slice number, (for example, slice 1 belongs to VE1, slice 2 belongs to SP/VE2, etc.). Also, Q^g_{min} is the minimum guaranteed number of resource blocks needed to satisfy the SLA for VE-g. According to [69], the proportional fairness utility function can be written in terms of a concave function as follows

$$\upsilon(\Gamma_g) = \frac{\Gamma_g}{T_g} \log(C_g) \tag{5.3}$$

where C_g is the number of RBs acquired by VE-g at the beginning of a TTI and T_g is the aggregate transmission rate obtained by VE-g in all former TTIs starting from $t = 1$ where t is the slot number. Note, that the higher T_g is, the lower the utility, whereas the higher C_g, the higher the utility. While a log utility function is compatible with NRT traffic, the RT traffic with constant

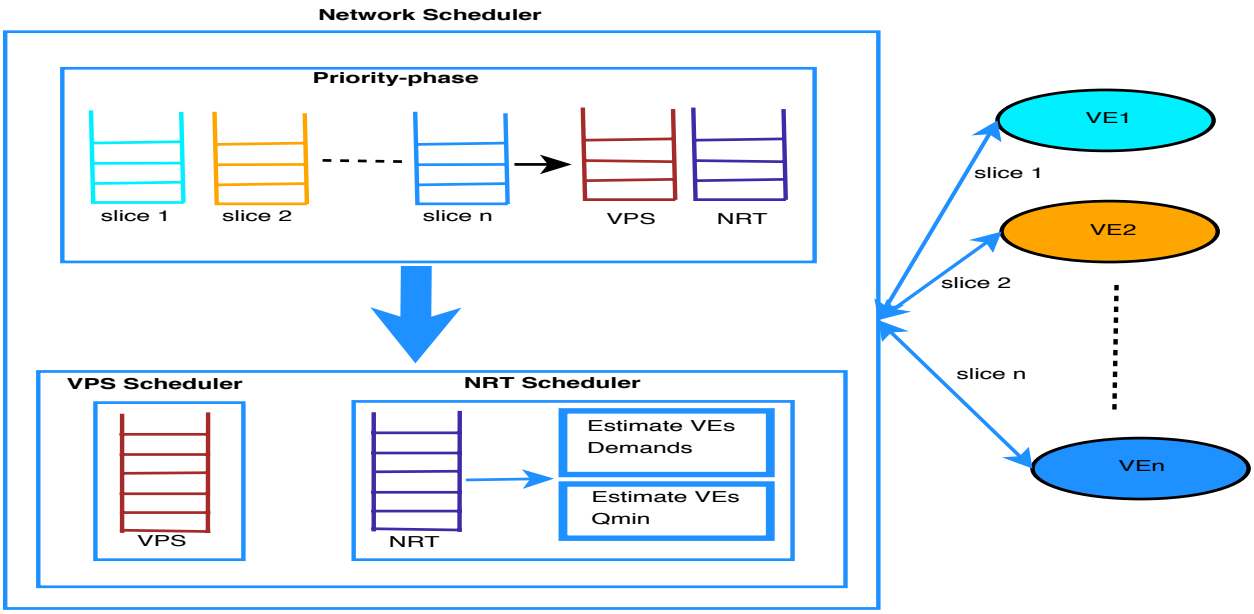


Figure 24: The new virtualization framework

inter-arrival times does not observe any improvement in utility with larger bandwidths[1]. In contrast, real-time traffic is delay sensitive and should not encounter a delay beyond a maximum acceptable delay. If the delay is larger, RT packets are dropped (or flows are blocked). Thus, one of the common methods to handle RT traffic is the notion of priority queues [70]. In LTE, users requests are initially prioritized before being served according to the criticality of the user application [68], e.g. VoIP(2), Video call(4), ... and so forth. The highest priority flow request (the lowest priority index) is served first. Unlike traditional LTE networks, in virtualized NGMN networks, it may not be the best approach to execute the prioritization step for each SP/VE separately. We claim that prioritization between different traffic types should be performed for all SP/VEs *jointly* at the network scheduler prior to the allocation of resources. This can better leverage the economy of scale effects of spectrum pooling. Furthermore, it is possible that an SP/VE would have insufficient resources to serve all real-time requests if the Network Scheduler assigns resources on a fair-throughput basis regardless of the traffic heterogeneity.

[1]This may change with different qualities of RT traffic, such as HD voice and video, which is outside the scope of this thesis.

To clarify this, we use a small example. Suppose we have a total of 8 PRBs and VE1 and VE2 each has 2 RT requests at the beginning of a TTI. We assume also for simplicity that VE1's PRBs all (with transmission rates per Hz) correspond to $CQIindex = 11$ in Table 8 and similarly, VE2's PRBs transmission rates correspond to $CQIindex = 14$. Hence, assuming $T_1 = T_2$ for simplicity and solving the optimization problem in Equation (5.2) for two VEs (VE1 and VE2) where the total utility function per Hz υ_{total} is given as,

$$\begin{aligned} &\upsilon_{total} \\ &\quad = \underset{C_1,C_2}{\text{maximize}} \quad v(\Gamma_1) + v(\Gamma_2) \\ &\quad = \underset{C_1,C_2}{\text{maximize}} \quad \Gamma_1 \log C_1 + \Gamma_2 \log C_2 \end{aligned} \tag{5.4}$$

we can derive C_1 and C_2 by solving equation (5.4). Note that Γ_1 and Γ_2 are the total transmission rates per unit time for VE 1 and VE 2, respectively. Figure 25 shows the different feasible solutions for C_1 and C_2. It is obvious that the proportionally fair allocation policy in this scenario is ($C_1 = 3, C_2 = 5$), i.e., these are the numbers of allocated RBs for VE1 and VE2 respectively which achieve the maximum total utility. We now examine the impact of this solution on the blocking rate of RT requests. First consider when prioritization is carried out individually by each SP/VE after allocation of RBs by the Network Scheduler, i.e., post-allocation priority. As we previously assumed, an RT request is served using 2 PRBs. Hence, VE1 would not be able to serve all of its RT requests and would have one dropped RT request. In the second scenario, let us suppose that the network scheduler prioritizes requests first and serves them using all of the available RBs. Then, none of the VEs would encounter drops of RT requests. This simple example is just a motivation to argue that the existence of a priority-phase *prior to allocating resources* and then serving RT requests first, independent of the PF-criterion is useful to ensure that that RT requests are better provisioned.

5.1.2.2 Reliability in Wireless LTE Systems Generally, link reliability is the ability of a radio link to transmit and receive a certain amount of data successfully within a predefined deadline [71]. In this thesis, we focus on the reliability of the link between a user in the network and the network itself. To clarify more, assume a user A communicates with the network through a single link. If user A sends a packet of M_p bits (payload) and the control bits added to the payload is M_t, then

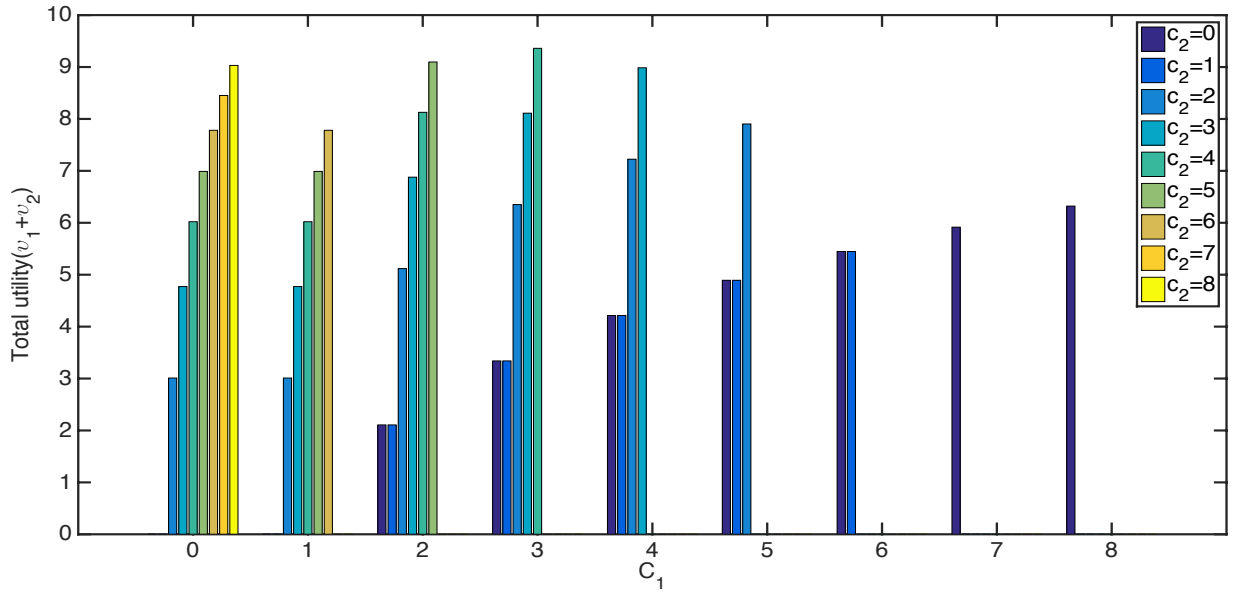


Figure 25: Total Utility Function υ versus the number of PRBs assigned to VE1 C_1 for different values of C_2

the total packet size is $K = M_p + M_t$. This data is channel coded with M_c bits for error detection and correction [72]. Hence, the total number of symbols received at the receiver is $N = K + M_c$. The coding rate R of the communication system is defined as the fraction K/N of information bits to the number of transmitted symbols. The task of the decoder at the receiver is to determine the information bits $B_1; ....; B_K$ from the N channel outputs $Y_1; ....; Y_N$. We denote the packet error probability as P_e. The packet error probability is the percentage of packets not successfully received at the receiver within a certain period of time. Hence, the number of coding bits M_c is chosen such that the error probability per packet per unit time is minimized and meanwhile the user still achieves the target throughput. As M_c increases, the total number of transmitted symbols N increases and hence the throughput deteriorates but P_e decreases and vice versa.

According to Polyanskiy et al. [73], the coding rate $\hat{R}$ is upper-bounded by the following formula

$$\hat{R} = \Phi - \sqrt{\frac{\Theta}{N}} Q^{-1}(\epsilon) + \mathcal{O}(\frac{logN}{N}) \tag{5.5}$$

where ϵ is the maximum allowable total transmission error probability and $\mathcal{O}(\frac{\log n}{n})$ comprises remainder terms of order $\frac{\log n}{n}$. Also, Q^{-1} denotes the inverse of the Gaussian Q function. Θ is the channel dispersion and Φ is the channel normalized capacity [73]. For the AWGN channel, Θ and Φ can be written as

$$\Phi(\gamma) = \log(1+\gamma) \tag{5.6}$$

$$\Theta(\gamma) = \gamma\frac{(2+\gamma)}{(1+\gamma)^2}(\log e)^2 \tag{5.7}$$

where γ is the received signal-to-noise ratio measured at the receiver. Also, a good approximation for $\mathcal{O}(\frac{logn}{n})$ in equation (5.5) is proved to be $(logn)/2n$ [74]. The formula in (5.5) gives the maximum coding rate needed for a maximum allowable total transmission error probability of ϵ and a total number of symbols N. A transmission error probability (ϵ) maps to a reliability rate of $Rel(\epsilon, N)$ where $Rel(\epsilon, N)$ for a certain link can be written as

$$Rel(\epsilon, N) = (1-\epsilon_u)(1-\epsilon_d) \tag{5.8}$$

i.e. a link that carries N symbols with a maximal uplink error probability of ϵ_u and a downlink error probability of ϵ_d is guaranteed to be reliable for a $Rel(\epsilon, N)\%$ of the whole time this link is active. We use the approach in equation (5.8) as an approximation for quantifying reliability in this thesis. Unlike other reliability approaches used in the literature, this approach is tractable.

The LTE CQI calculations in Table 8 is designed to maximize the user throughput such that the UE power and delay constraints should be met. Also, it is presumed that the average error probability of sending a packet ϵ is no more than 10% [35]. Figure 26 shows the reliability (%) versus delay in seconds in traditional LTE networks, i.e. traditional CQI mapping. This figure is produced assuming AWGN channels for uplink and downlink in a cell with a radius of 500 meters and a total of 2 mobile users. We assume that one of the users is downloading a file in a size of 0.5 MBytes and the other one has a URLLc request with a packet size of 100 Bytes and an arrival rate of 2 packets per TTI. We generate the locations of the users in the cell randomly, we iterate the simulation 1000 times and we take the average over 1000 observations. Using the modulation and coding schedule (MCS) in Table 8, we simulate an LTE cell using MATLAB and derive the average link reliability using equation (5.8). Then, we plot the average link reliability versus different maximum delay

values in Figure 26. We also plot the average antenna transmission power versus different delay constraints in Figure 27. From Figure 26, we can clearly see that achieving a user plane latency as low as 0.5 msec makes the target 5G ultra-high reliability unattainable, i.e. $\simeq 0.999$. Moreover, the transmitting power reaches the maximum allowable transmission power by a pico BS, i.e. 38 dBm according to 3GPP standard. Thus, the LTE network reliability is very loose compared to the stringent requirement by URLLC applications. Therefore, the traditional LTE resource allocation should be adjusted for URLLC requests to accommodate ultra-reliable users in NGMN networks, i.e. our e-VPS new resource scheduler.

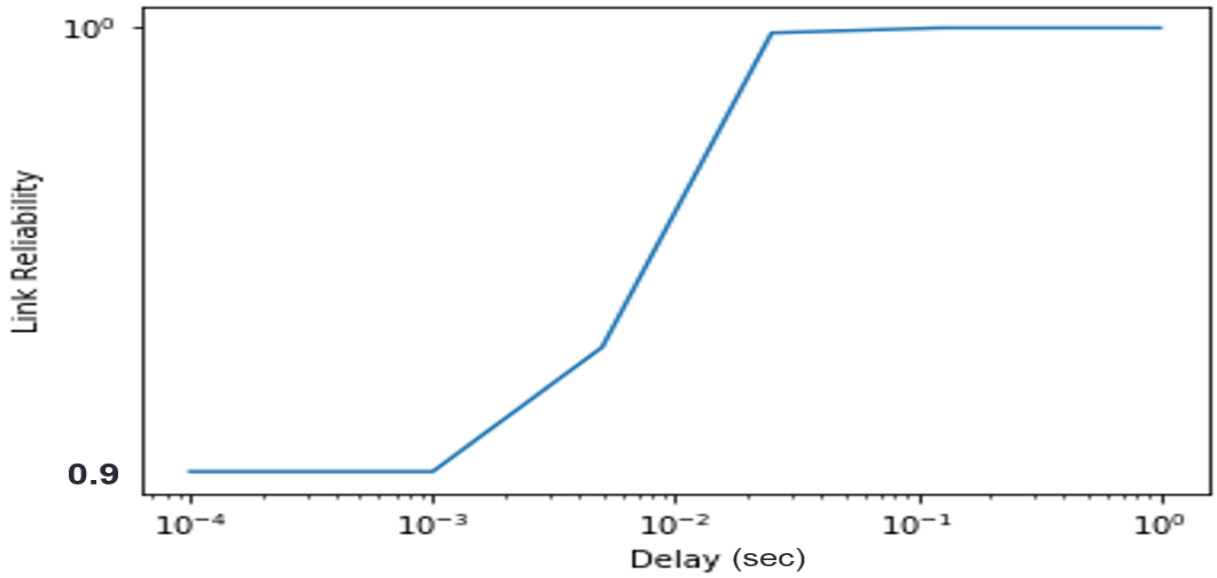


Figure 26: Reliability versus user delay constraint in traditional LTE networks

5.1.3 Problems with Static Allocation of Q_{min} RBs to Virtual Entities

The third problem that our proposed resource allocation tackle is determining the appropriate choice of the minimum reserved number of RBs for SP/VEs such that resources do not remain unused or no SP/VE is assigned fewer resources than possible to improve the average throughput and user blocking probability. In this subsection, we show the impact of changing Q_{min} on system performance through discussing various scenarios.

Intuitively, since the traffic demand changes with time and thus the number of allocated resources

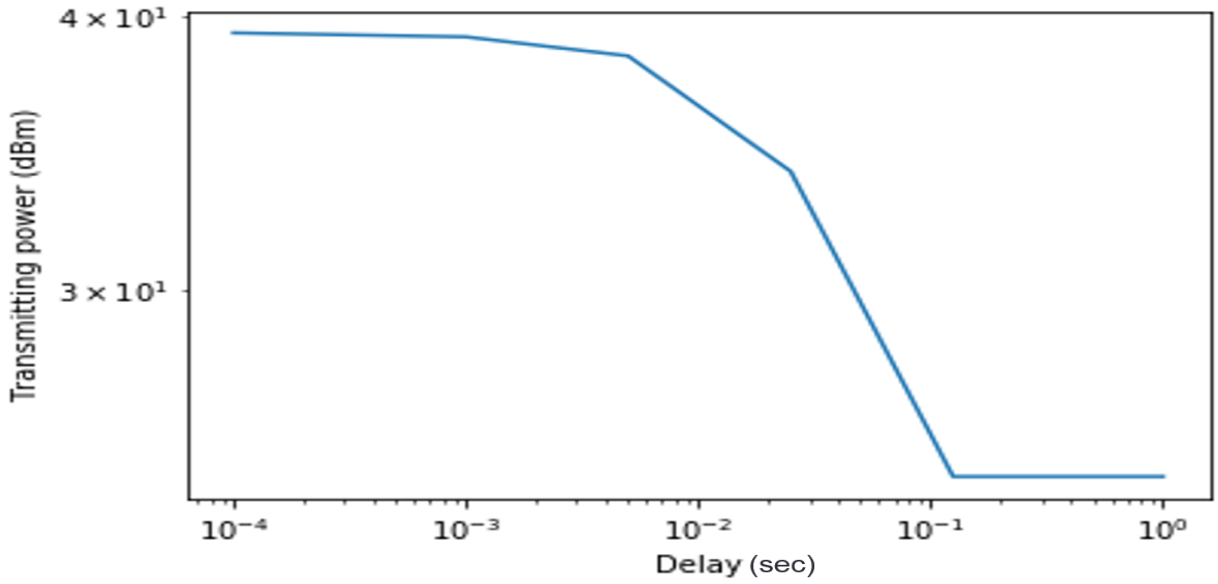


Figure 27: Transmitting Power versus user delay constraint in traditional LTE networks

changes as well, the minimum number of allocated resources Q_{min} should not be fixed. In NetShare, the minimum QoS for VE i is just proportional to the total number of resources which is agreed to be statistically reserved to such SP/VE in the SLA denoted by L_i, i.e. $Q^i_{min} = \zeta L_i$. However, we argue in this thesis that provisioning Q_{min} dynamically based on the actual number of allocated resource per SP/VE can improve the total throughput for the SP/VE. For better clarification, we simulate an LTE cell with Rayleigh fading channels and a path loss exponent of $\alpha = 2$. SP1 and SP2 have 10 and 5 users respectively and users are uniformly distributed through the cell with a radius of 500 meters. Using the NetShare scheduling, i.e. $Q^i_{min} = \zeta L_i$, we define the loss in resource utilization as the difference between the actual used RBs in transmission for VE-i denoted by C_i and the total reserved RBs L_i, i.e.,

$$loss(\%) = \frac{L_i - C_i}{L_i} \times 100\% \tag{5.9}$$

Figure 28 shows the variation of the spectrum utilization by VE1 and VE2 demands with Q_{min} assuming the algorithm used by NetShare. It is apparent that as ζ increases, the expected loss in resource utilization increases as well. In other words, the suggested Q_{min} deviates more from

the actual number of resources needed to serve the SP/VE. Next, we discuss another scenario that

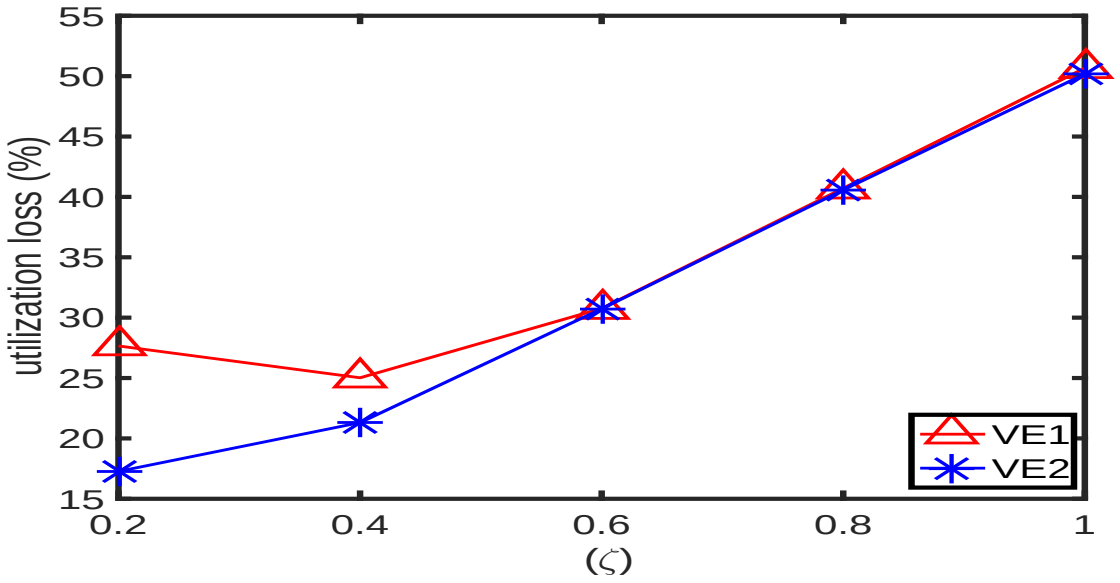


Figure 28: The utilization loss versus Q_{min} for SP/VE 1 and 2

further shows the shortcoming of the static selection of Q_{min}. Again, we conduct simulations using the aforementioned NetShare assumptions. However, in this experiment we artificially assume quite poor quality of RBs for mobile users of VE2, by changing the path-loss exponent ($\alpha_1 = 4$). We assume that VE1 has a path-loss exponent of $\alpha_2 = 2$. Since the received signal strength falls with distance d as d^{α}, mobile users of VE2 see poorer signal strength and thus CQI. We explore two scenarios, Scenario I where $Q^1_{min} = \frac{1}{2}\frac{C_1}{C_1+C_2}C_T = Q_a$ and $Q^2_{min} = \frac{1}{2}\frac{C_2}{C_1+C_2}C_T = Q_b$, i.e. Q_{min} for VE i is proportional to the number of mobile users subscribed to that entity and $\zeta = 0.5$ (we use ζ). The other scenario is Scenario II where $Q^1_{min} > Q_a, Q^2_{min} > Q_b$. Figure 29 shows the average flow throughput versus various percentages of RT requests for both scenarios. Clearly, we can see that Scenario II has a better throughput than Scenario I for VE1, but the throughput is roughly the same for VE2 in both cases. The reason for that is the very low quality RBs for VE2, hence the throughput is not improved by increasing its Q_{min}. However, VE1's throughput is clearly enhanced by increasing Q_{min} from that in Scenario I to that in Scenario II. Generally speaking, this comparison of scenarios implies that keeping Q_{min} static is probably not a good idea. Dynamically changing Q_{min} according to the quality of RBs, as well as the traffic demand

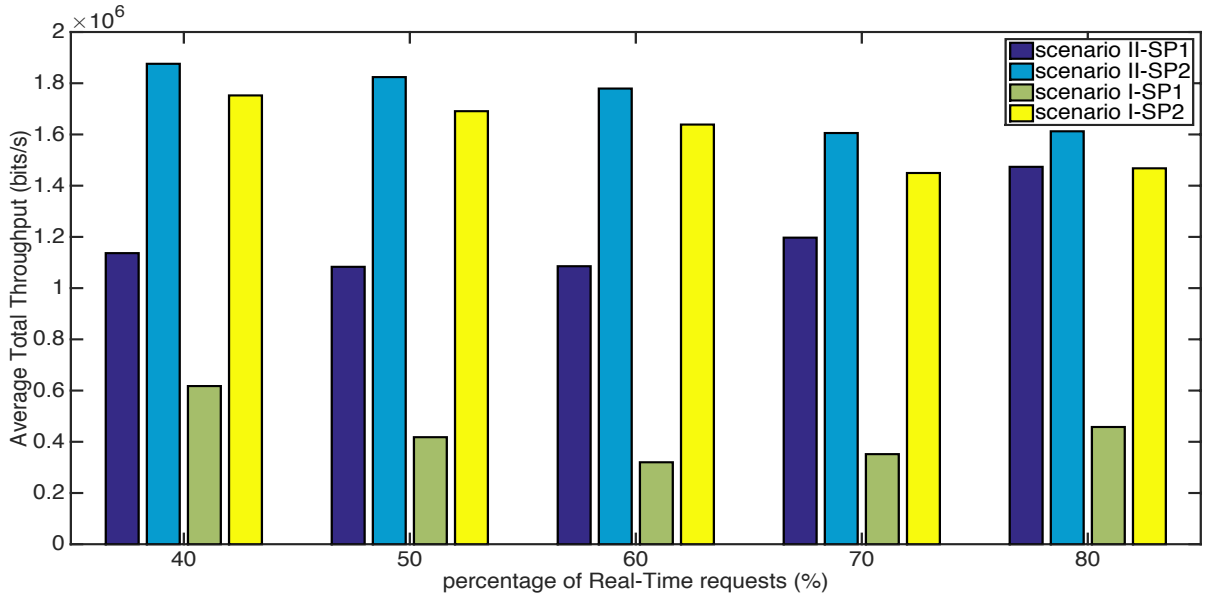


Figure 29: Average Total Throughput versus the percentage of real-time requests in both scenarios I and II

may positively affect the virtualized system's overall performance. Accordingly, in next section, we present a new method to derive Q^i_{min} for VEi based on both the average long term transmission rate of VEi in previous slots allocated to its users as well as the current demands of this SP/VE, i.e., it resembles a PF-based selection, albeit at an SP/VE level than a user/flow level.

5.2 E-VPS SYSTEM

In this section, we describe the model of our proposed approach. Unlike the virtualization framework proposed for NVS [24] and NetShare [25], we adopt the VPS framework that was proposed in previous work [54] as shown in Figure 24. In this framework, the Network Scheduler performs resource scheduling in two main phases, a priority-phase and a virtual prioritized slice (VPS) scheduling-phase as depicted in Figure 24. The primitive framework in [54] does not address URLLC requests nor does it consider dynamic changes to $Qmin$. Hence, we extend this frame-

work as shown in Figure 30 to an enhanced virtual prioritized slice, i.e. e-VPS, that includes URLLC traffic. The e-VPS framework comprises of four phases as follows.

5.2.1 Priority-Phase

In the priority phase, instead of performing prioritization between different classes of applications for each SP/VE separately after allocation of RBs (post-allocation priority), the Network Scheduler collects the flow requests from both entities VE1 and VE2, and jointly prioritizes the flows according to a priority index that takes into account delay and reliability. This priority index considers the type of application before allocation of RBs (pre-allocation priority). Flow requests from all SP/VEs are queued as shown in Figure 24 into two queues – a virtual prioritized slice (VPS) which contains all RT requests from both SP/VEs and the NRT slice containing all NRT requests from both SP/VEs. The VPS is forwarded to the scheduler to be assigned RBs regardless of the Proportional Fairness scheduling scheme in equation (5.2). This is because, as we discussed previously, each RT request needs a fixed assignment of RBs (two in this thesis). On the other hand, the NRT slice is provisioned using the PF-scheduler in equation (5.2) according to the number of requests per SP/VE as well as the PRBs quality for each SP/VE. The next step is scheduling URLLC requests within both RT and NRT traffic flows.

5.2.2 Scheduling URLLC Requests

Assuming that URLLC requests are one of the service classes in the NGMN network, we redesign our proposed VPS resource allocation scheme to fit the new scenario. We propose the enhanced VPS algorithm (e-VPS) for allocating resources to RT, NRT and URLLC requests. Basically, the URLLC requests can be RT or NRT. Examples of RT-URLLC include tactile communications, e-health and Vehicle-to-Vehicle emergency communications where a real-time reliability is inevitable to guarantee the full functionality of such type of communications. On the other hand, communicating ultra-reliably to a public cloud over short periods of time to analyze confidential data could be an example of NRT-URLLC. Hence, in the e-VPS scheduler we have four classes of traffic: RT-URLLC, RT only, NRT-URLLC and NRT only, arranged respectively according to priority. After splitting the traffic into RT (VPS) and NRT queues as depicted in Figure 30, the

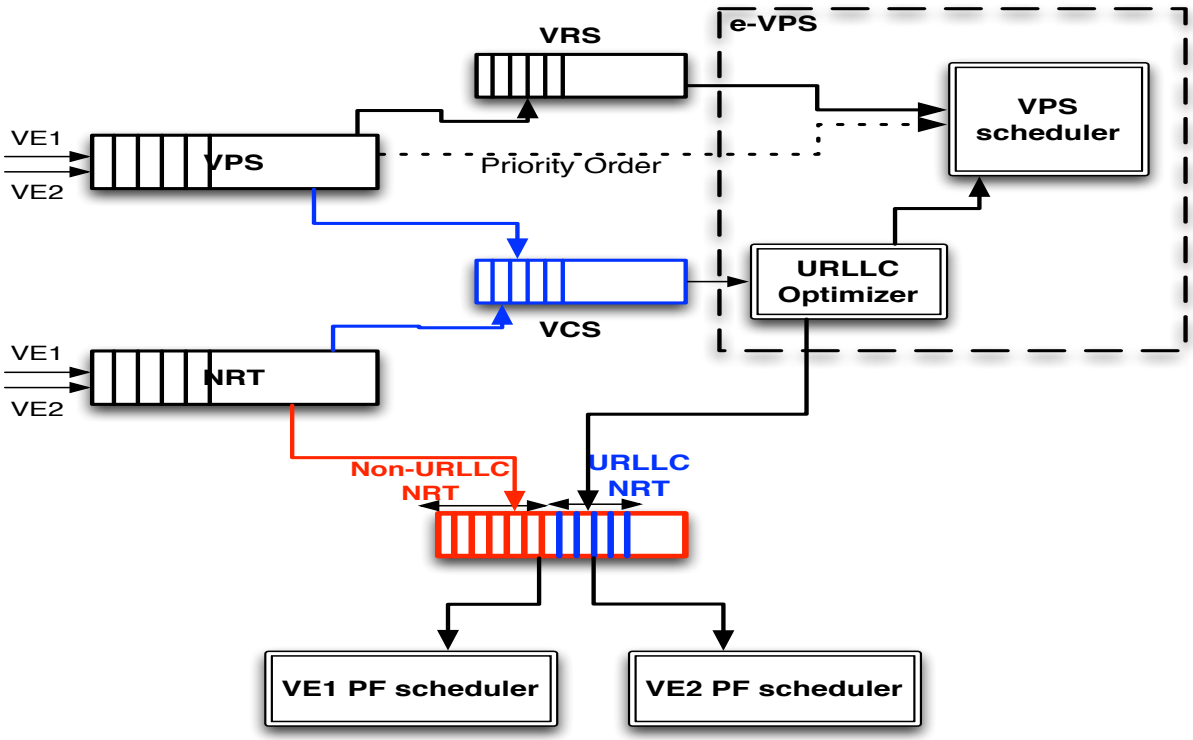


Figure 30: Scheduling Packets from different VEs with e-VPS

NRT-URLLC and RT-URLLC requests are queued into a VCS (Virtual Critical Slice), i.e. the queue in blue, as shown in Figure 30. The resources for VCS requests are scheduled through the URLLC optimizer shown in Figure 30. Here, we optimize the coding rates over different URLLC links through the URLLC optimizer as described next.

5.2.3 URLLC Coding Optimizer for URLLC Traffic

We argue that for the URLLC service class a new scheduling scheme is to be constructed based on the required reliability level ($Rel\%$), the SNRs of the available channels and the maximal allowable error rate (ϵ) given in equations (5.5-5.8) as well as traditional LTE CQI metrics. The new optimal scheduler criterion can be written as

$$\begin{aligned} \underset{N_j}{\text{maximize}} \quad & \upsilon(N_j) \\ \text{subject to} \quad & Rel\%(N_j, \epsilon_{u,j}, \epsilon_{d,j}) \geq r_j, \quad j \in \varkappa \\ & D_j^{ur-rt} \leq D_{max}^{rt}, \quad j \in \varkappa \\ & D_j^{ur-nrt} \leq D_{max}^{nrt}, \quad j \in \varkappa \\ & P_j \leq P_{max}, \quad j \in \varkappa \\ & \rho_u \geq \rho_{th}, \quad j \in \varkappa \\ & t_{up} \leq t_{co}. \end{aligned} \tag{5.10}$$

where $\varkappa$ is the set of links between mobile users and the network. P_j is the transmitting power across link j, ρ_u is the user throughput and ρ_{th} is the minimum allowable user throughput. D_j^{ur-rt} and D_j^{ur-nrt} denote the RT and NRT URLLC user-plane delay on link j which have to be lower than or equal to the maximum allowable delays D_{max}^{rt} and D_{max}^{nrt}, respectively. The optimization problem in equation(5.10) finds the optimal number of symbols N_j to be transmitted on link j (N_j) to maximize the throughput $\upsilon(N_j)$ over this link j such that the reliability is maintained at a certain level $Rel\%$, the uplink and downlink error probabilities are upper-bounded by $\epsilon_{u,j}$ and $\epsilon_{d,j}$, respectively, and the application end-to-end delay requirement can be fulfilled.

On a system level, the VCS requests are routed to the URLLC optimizer shown in Figure 30 to apply the URLLC optimization criterion in equation (5.10) to obtain the optimal coding rates ($R = K/N$) for every individual request. Also, it is worth noting that the time granularity to run

such optimization can not be less than the coherence time t_{co}, i.e, $t_{update} \leq t_{co}$. The reason for that is to make sure that the scheduler copes with the fading duration [75]. However, we do not address the issue of time granularity in this thesis. After optimally selecting the coding rates per each link in $\varkappa$, the number of resources per a request is determined as illustrated in the last phase, i.e., optimal VPS scheduler.

5.2.4 Optimal VPS Scheduler

The last phase for the virtual framework consists of the VPS scheduler and the NRT scheduler. The VPS scheduler allocates a fixed number of resources to RT requests, i.e. RBs each. More specifically, RT requests are first allocated RBs. We examine two approaches later - assigning the RBs with the lowest CQIs and assigning RBs with the highest CQIs. As noted previously, assigning RBs with the highest CQI does not materially improve the performance for RT requests. RT requests are also arranged in order of their delay budgets as explained later. The ones with the smallest delay budgets are allocated RBs earlier. Next, the NRT requests are scheduled according to Equation (5.2) using the residual RBs. Moreover, in our framework, since we propose a new dynamic Q_{min} allocation, the NRT scheduler will not only estimate the resources allocated to each SP/VE but also the appropriate assignment of Q_{min} based on a PF-criterion as we illustrate next.

5.2.5 PF-based Assignment of Q_{min}

We propose determining Q_{min} such that Q^i_{min} is proportional to $\frac{\Gamma_i}{T_i}$, i.e.

$$Q^i_{min} = \beta \frac{\Gamma_i}{T_i} \tag{5.11}$$

where Γ_i is the total estimated transmission rate for users belong to VE-i during the current slot t, T_i is the aggregate cumulative transmission rate in the preceding $t-1$ slots and β is a parameter that enforces a necessary condition which is:

$$0 < Q^i_{min} < \frac{1}{\eta} C_T \tag{5.12}$$

where $\eta \geq 2$ and C_T is the total number of available resource blocks. The β interval enforces an intuitive constraint on Q_{min} which is the minimum guarantee of any SP/VE is strictly less than

half of the available resources in the system to assure isolation between SP/VEs (assuming they are identical in demand, channel quality and etc.). Note that we assume that both SP/VEs have the same rights to access the pooled spectrum without priorities. To achieve these conditions, we define Q_{min} with the following criterion:

$$Q_{min} = \frac{\Gamma_i}{\eta\Gamma_i + \theta T_i} N_T \tag{5.13}$$

where θ is a parameter that controls the weight of T_i in selecting Q^i_{min}, and $\theta \geq 1$.

5.3 SIMULATION RESULTS

In this section, we present some results which show how the schemes proposed in this chapter outperform those that have been previously proposed in the literature. In particular, we evaluate the performance of the two new features proposed in our model, pre-allocation priority and the dynamic selection of Q_{min}. We show that these two new features are beneficial in comparison to NetShare [25]. It is worth noting that, in [25], it was shown by simulations that NetShare outperforms NVS [24], i.e., it achieves more efficient resource scheduling. Hence, we limit our comparison to NetShare only to provide more clarity to the plots. Moreover, we show by simulations the superiority of handling URLLC requests compared to current traditional LTE networks. We conducted simulations using MATLAB. The default parameters used in this simulation are shown in Table 9 unless otherwise indicated. Also, we assume that all RBs face Rayleigh fading with the parameters shown in Table 9. Our results are calculated based on an average of 5000 iterations per 20 TTIs. Also, the error bars shown in the figures correspond to the 95% confidence intervals. Mobile users of both SP/VEs are distributed uniformly in the cell.

In estimating the received signal-to-interference and noise ratio (SINR) γ, we assume that the serviced requests always undergo a given constant interference from surrounding cells during their service, for simplicity. Generally, the received SINR γ_k at user k is given as:

$$\gamma_k = \frac{P_{r,k}}{P_T + n} \tag{5.14}$$

Table 9: Simulation Parameters

operating frequency (f)	700 MHz
Transmitted BS power	$\leq$ 40 dbm
Mobile User power	20 dbm
Cell Size(d_{max})	600 m
Number of active users-SP/VE1(u_1)	20
Number of active users-SP/VE2(u_2)	5
flow length(L)	3×10^5bits
SP/VE1-Bandwidth	1.4 MHz
SP/VE2-Bandwidth	1.4 MHz
shared BW(full sharing)	2.8 MHz
total number of RBs	15
subcarrier spacing(Δf)	15 KHz
Number of subcarriers/RB	12
subframe(TTI) time	1 msec
Number of symbols/TTI	14
pathloss exponent (α)	2
RT maximum delay (D_{max}^{ur-rt})	15 msec
NRT maximum delay (D_{max}^{ur-nrt})	100 msec
minimum NRT user throughput (ρ_{th})	10 Mbps
minimum RT user throughput (ρ_{th})	1 Mbps
minimum URLLC user throughput (ρ_{th})	1 Mbps
minimum Target Reliability (r_j)	≥ 0.999

where $P_{r,k}$ is the received power at user k, P_T is the total received power from surrounding cells and n is the zero-mean additive white gaussian noise(AWGN) variance. In this thesis, we assume the free space path-loss model and accordingly $P_{r,k}$ is given by [76]

$$P_{r,k}\ \text{dB} = P_{t,k}\ \text{dB} + (\frac{\lambda^2 d_o^{(\alpha-2)}}{16\pi^2 d_k^{\alpha}})\ \text{dB} + |H_{k,m}|^2\text{dB} \tag{5.15}$$

where $P_{t,k}$dB is the power of the transmitted signal by the BS (downlink) in dB, λ is the wavelength of the transmitted signal, d_o is a reference signal for the antenna far field, usually set to 10-100 m outdoors, d_k is the distance from BS for user k and $H_{k,m}$ is the multipath fading coefficient for RB m used by user k. We assume that the received power at the user receiver $P_{r,k}$ undergoes Rayleigh fading with fading coefficients $H_{k,m}$. Then, the received SINR γ per TTI is a random variable with an exponential probability density function (pdf) [76]

$$P_\gamma(\gamma) = \frac{1}{\overline{\gamma}} exp(-\frac{\gamma}{\overline{\gamma}}) \tag{5.16}$$

where $\overline{\gamma}$ is the average received SINR, calculated from the path-loss. The CQI, for determining the throughput, is calculated using the following conversion equation based on [35]

$$j = \frac{\gamma_k(dB) - b}{a} \tag{5.17}$$

where j is the CQI value, a and b are constants which are derived based on the linear relation between the CQI value and the γ_k(dB) obtained in [77]. In [77], a and b are given as $\frac{13}{7}$ and $\frac{-55}{7}$, respectively.

We assume RT requests with exponentially distributed maximum request delay budgets. We recall that the maximum-request-delay budget is the maximum delay a request can tolerate before it is discarded and dropped. Also, we assume that the inter-arrival time for RT requests is one TTI, i.e., 1 ms. We recall that in our proposed VPS approach, each RT request is serviced by 2 PRBs as we presumed antecedently. We assume the number of RT requests u_{rt} per SP1 and SP2 is uniformly distributed within the interval $[0, u1]$ and $[0, u2]$, respectively. Also, the number of URLLC requests per SP is uniformly distributed within the interval $[0, u_{rt}]$.

5.3.1 Effectiveness of pre-allocation prioritization at the Network Scheduler in different frequency bands

In this section, we discuss the results of sharing in the 700 MHz, the 2100 MHz band and the inter-band sharing between the 700 and 2100 MHz bands.

5.3.1.1 Sharing in the 700 MHz band : We first show the effect of sharing in a 700 MHz band. Figure 31 shows the average total SP/VE throughput for VE1 versus the percentage of RT requests (compared to the total network load) for the proposed approach VPS (pre-allocation) for three cases – the best (highest) CQIs with URLLC (blue), the best (highest) CQIs case (without URLLC)(red) and the worst (lowest) CQIs case (without URLLC)(yellow). Also, the result for NetShare with URLLC (post-allocation) is shown by (purple) bars. It's obvious that the VPS approach with worst-CQI outperforms the best-CQI VPS until the network load of RT requests exceeds 60%. When the RT load surpasses 60%, the throughput of the worst-CQI case deteriorates due to the fact that our approach that assigns a fixed number of RBs, i.e. two RBs, to the RT traffic helps NRT requests to obtain RBs with higher quality, i.e. higher SNRs and hence higher throughput. This is why, as shown in Figures 35 and 36, the throughput for NRT traffic improves drastically. Similarly, we can see the same performance for VE2 in Figure 32 but the change is slight since VE2 has a much lower number of users than VE1.
Nevertheless, as the percentage of RT requests increases in the network, i.e., the network becomes more loaded, the best-CQI VPS approach outperforms the worst-CQI approach since the NRT throughput decreases while the dominant RT throughput is low due to using low-SINR channels to provision RT requests as shown in Figures 33 and 34. Also, in Figure 33, NetShare achieves better real-time throughput than our proposed e-VPS algorithm since NetShare assign RBs to VEs independently subject to the net demand of each VE irrespective of the individual traffic loads of different service types (RT, NRT and URLLC) such that the VE throughput is maximized. However, the VPS scheme allocates only 2 RBs per each RT request to better utilize the spectrum resources to provision both URLLC and NRT traffic. Moreover, when applying the proposed VPS scheme to a network with URLLC type of service, Figures 31 and 32 show that the total throughput of the VE decreases below the no-URLLC case, i.e. the blue bars compared to the yellow and the red bars. The reason for that is the URLLC optimizer assigns higher channel coding to URLLC requests to provide the target ultra-high reliability and hence the throughput decreases such that the minimum user throughput is still achievable. We can also observe that guaranteeing PRBs for RT requests does not significantly affect NRT requests since we limit the number of RBs per an RT request to 2 RBs only. To support this statement, Figure 35 shows the average NRT traffic throughput for VE1 versus the percentage of RT requests to the total number of requests. It is apparent that the VPS approach has higher NRT throughput than NetShare. This is due to that most of RBs are occupied by RT traffic flows leaving fewer capacity for NRT traffic, since RT traffic is prioritized first. Similarly, Figures 32 and 36 show the total throughput as well as the NRT throughput for SP/VE2 in different scenarios. The same conclusions apply to SP/VE2 but with a negligible change with the increase of RT traffic percentage. This is due to the low number of users served by SP/VE2, i.e. $U_2 = 5$, and hence the change in the total, RT and NRT throughput is slight.

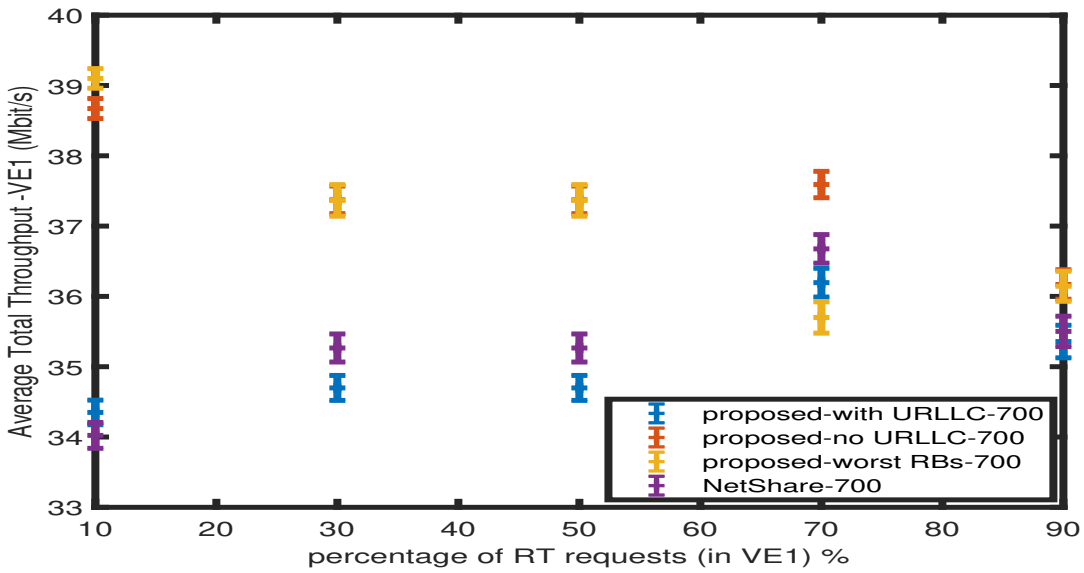


Figure 31: The Average Total throughput for SP/VE 1 for Pre-allocation and post-allocation priority cases

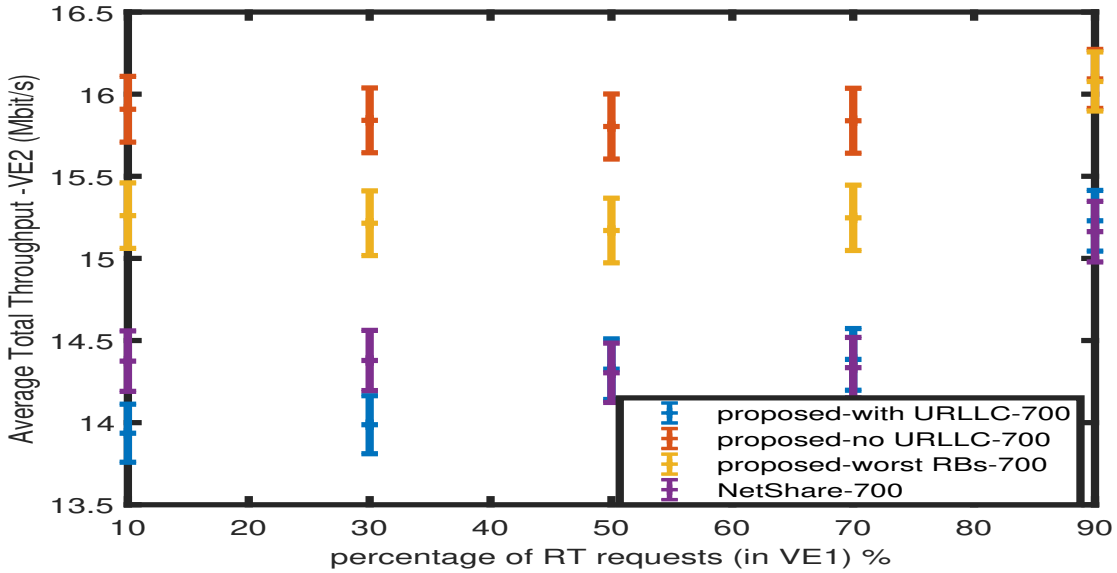


Figure 32: The Average Total throughput for SP/VE 2 for Pre-allocation and post-allocation priority cases

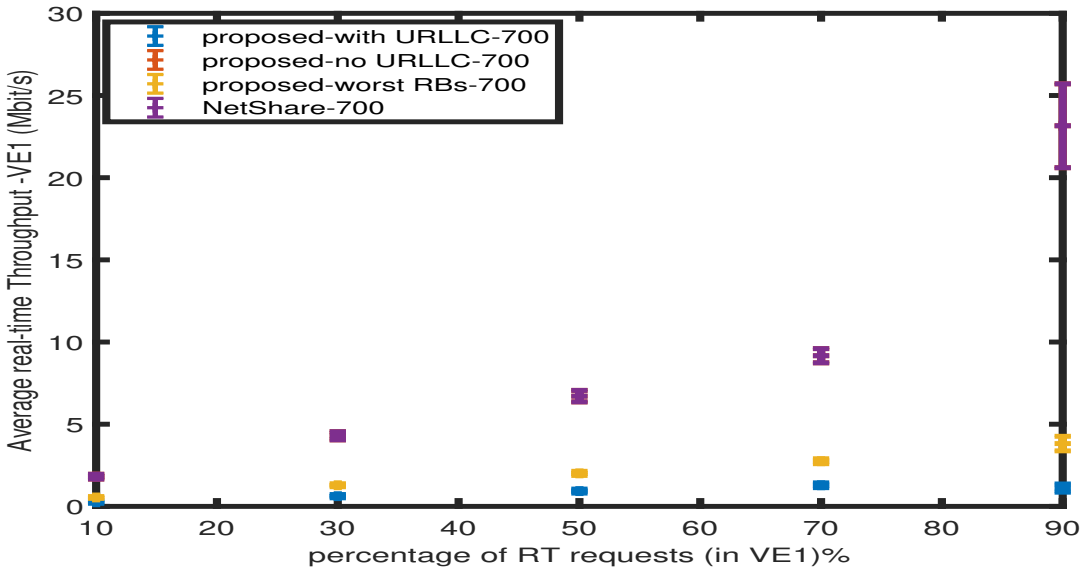


Figure 33: The real time traffic throughput versus the percentage of Real-Time requests for SP/VE1

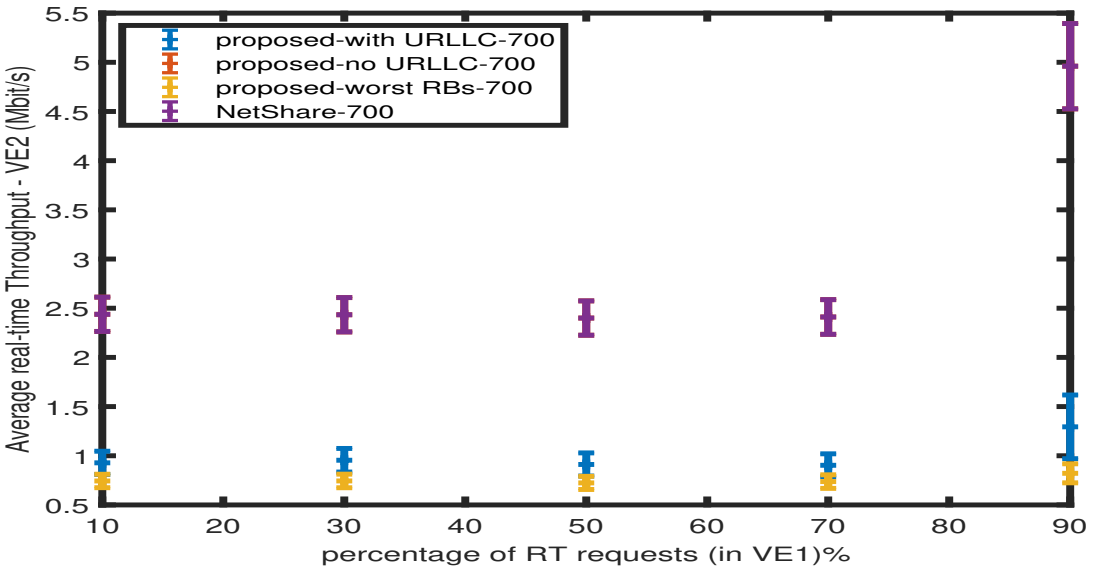


Figure 34: The real time traffic throughput versus the percentage of Real-Time requests for SP/VE2

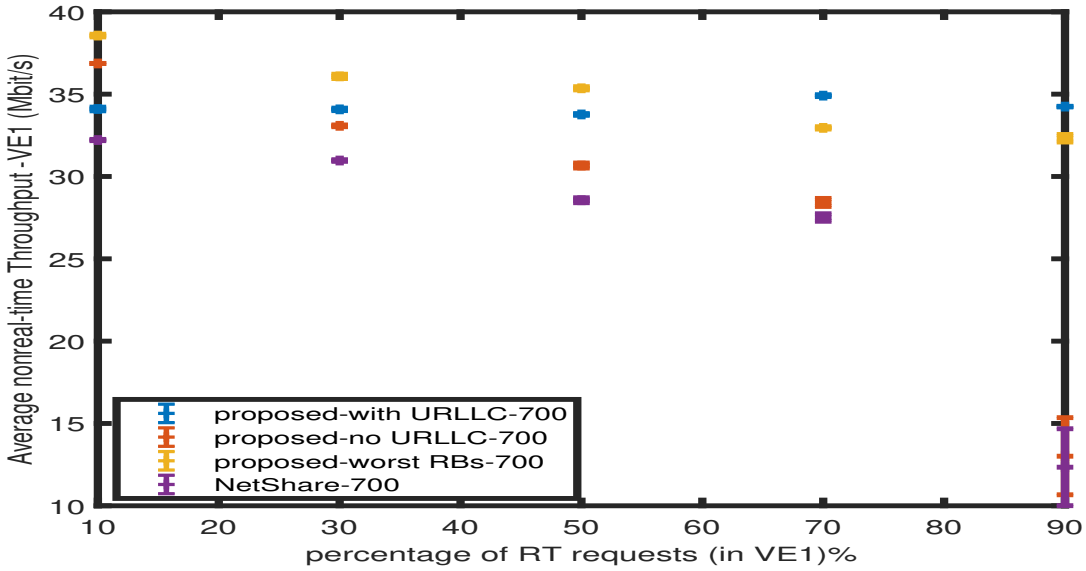


Figure 35: The Non-real time traffic throughput versus the percentage of Real-Time requests for SP/VE1

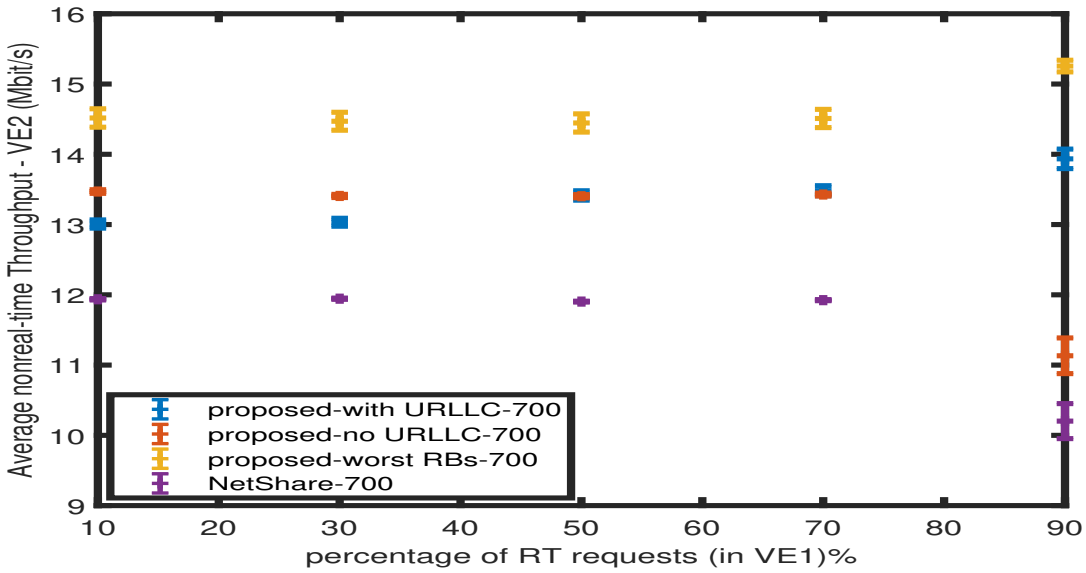


Figure 36: The Non-real time traffic throughput versus the percentage of Real-Time requests for SP/VE2

5.3.1.2 Sharing 700-2100 MHz band : In this part, we assume inter-band sharing where SP1 uses the 700 MHz band and SP2 uses the 2100 MHz band. Similarly, Figures 37, 38, 39, 40, 41 and 42 show the total throughput, the real time throughput and non-real time throughput for SP1 and SP2 respectively. From the total throuhput figure, it is clear that inter-band sharing decreases the total throughput due to higher attenuation in 2100 MHz bands. However, this deterioration is relatively small compared to using 2100 MHz bands only. Hence, in cases where 2100 MHz bands are used for higher bandwidth efficiency, 700 MHz bands can be used as supplementary frequency bands to provide better throughput relative to using the 2100 MHz band only, especially for URLLC use cases.

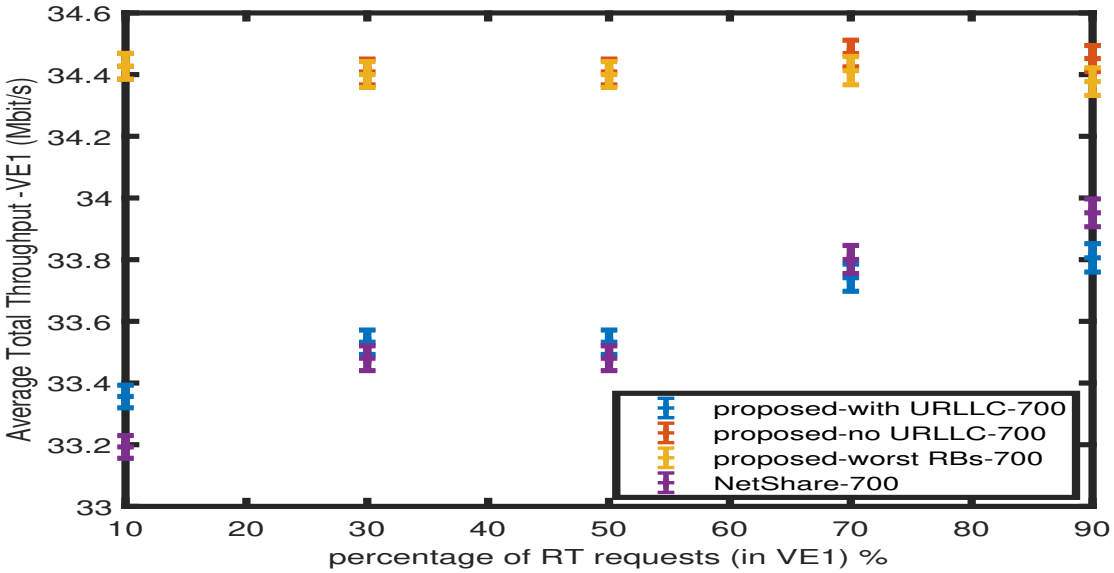


Figure 37: The Average Total throughput for SP/VE 1 for Pre-allocation and post-allocation priority cases, inter-band sharing

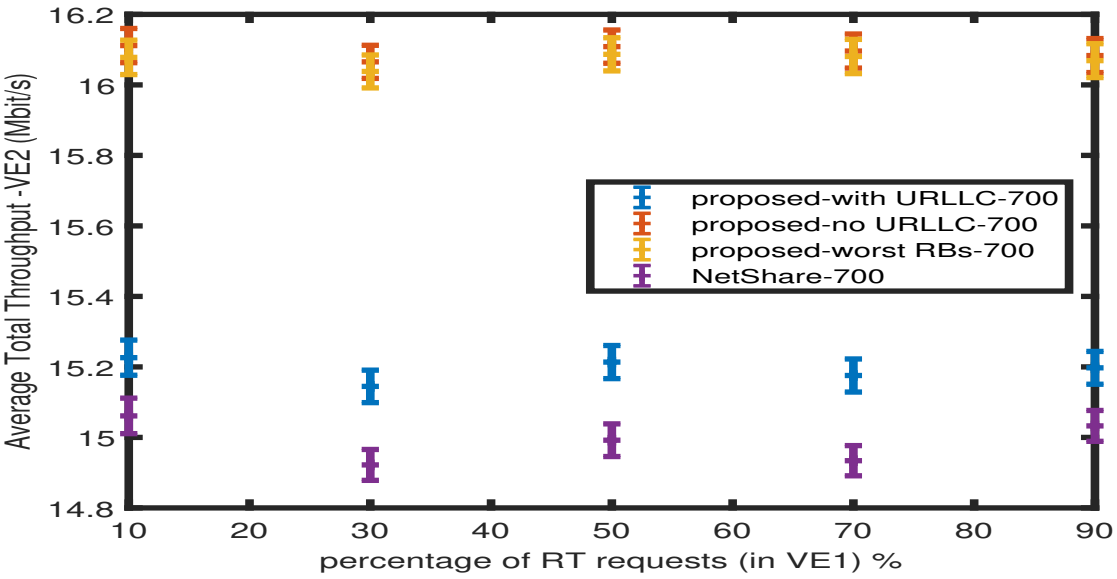


Figure 38: The Average Total throughput for SP/VE 2 for Pre-allocation and post-allocation priority cases, inter-band sharing

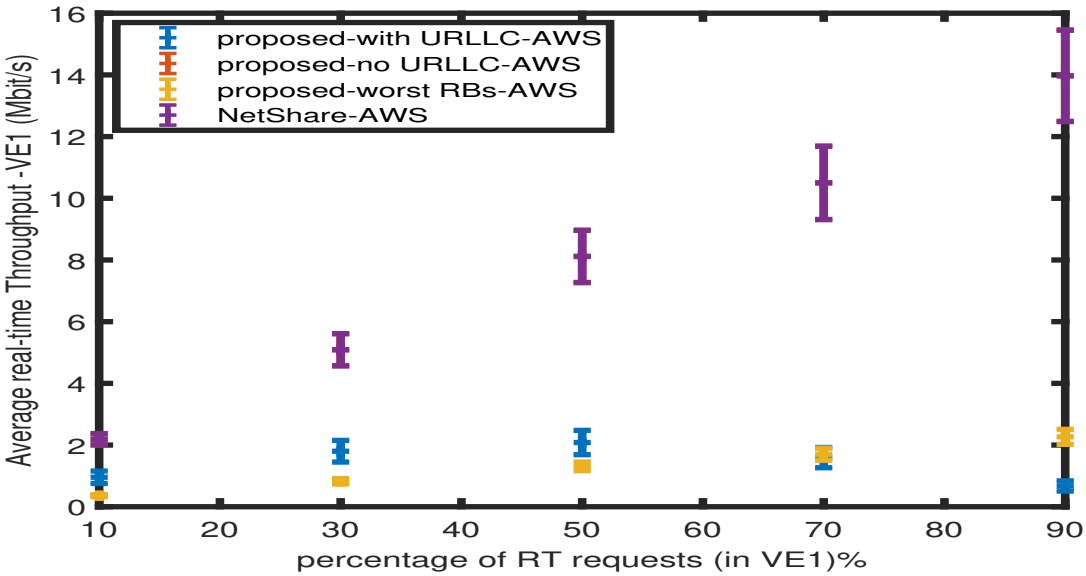


Figure 39: The real time traffic throughput versus the percentage of Real-Time requests for SP/VE1, inter-band sharing

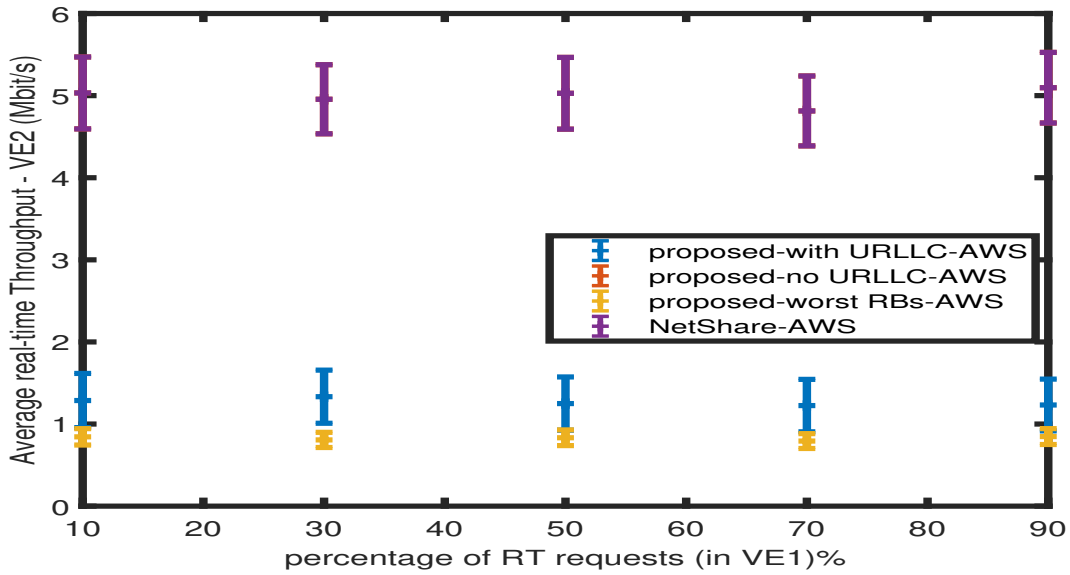


Figure 40: The real time traffic throughput versus the percentage of Real-Time requests for SP/VE2, inter-band sharing

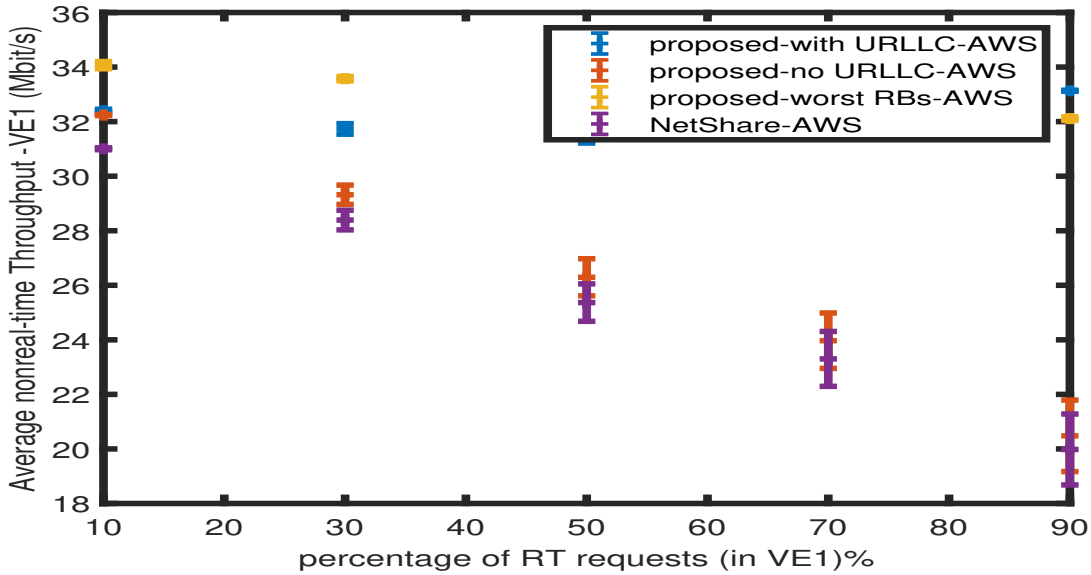


Figure 41: The Non-real time traffic throughput versus the percentage of Real-Time requests for SP/VE1, inter-band sharing

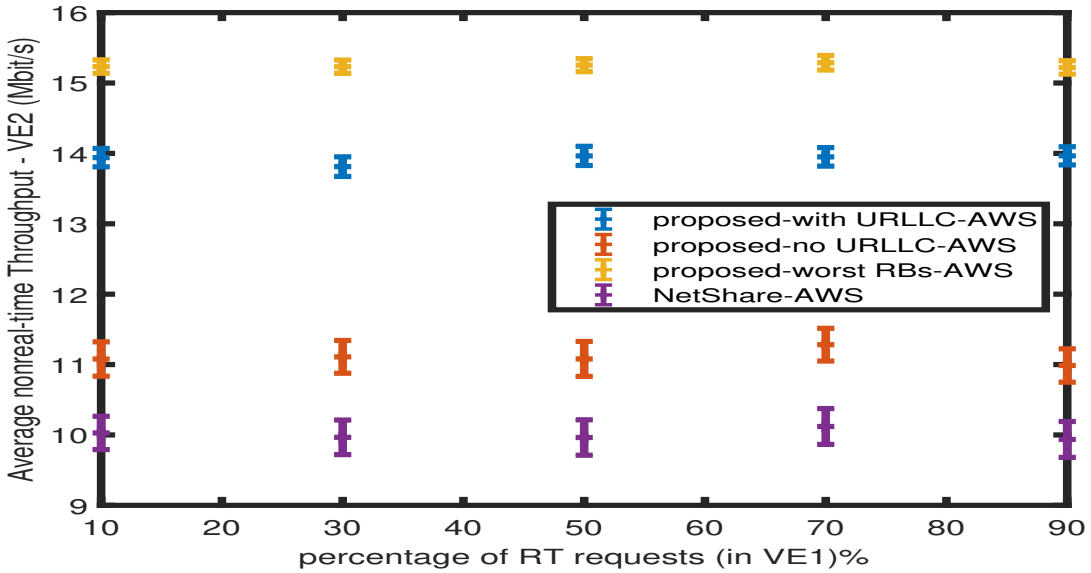


Figure 42: The Non-real time traffic throughput versus the percentage of Real-Time requests for SP/VE2, inter-band sharing

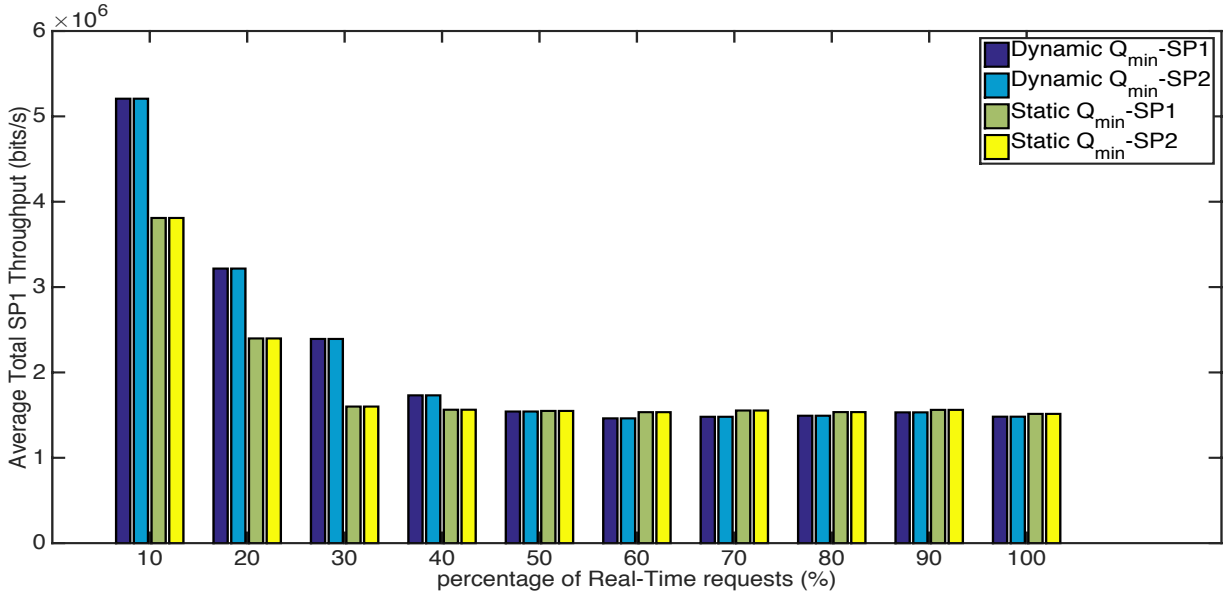


Figure 43: The Average Flow Throughput Vs. Flow Arrival Rate for both SP/VEs in dynamic Q_{min} and NetShare scenarios

5.3.2 Effectiveness of PF-based Q_{min} selection

In this set of numerical results, we study the use of PF-based dynamic Q_{min} with the proposed VPS. The values of the additional parameters used in this study are listed in Table 10. We assume these parameters hold for both SP/VEs. Figure 43 depicts the average total throughput for SP/VE1

Table 10: parameters for PF-based dynamic Q_{min} and NetShare

(a) PF-based

η	3
θ	4

(b) NetShare

L_i	$0.5N_T$
ζ	0.25

and SP/VE2 versus the percentage of RT requests. We can clearly note that the PF-based dynamic Q_{min} selection outperforms the static one, i.e. NetShare, for both SP/VEs when the percentage of RT requests is less than 50% of the total requests. This is due to the flexibility in choosing Q_{min} for both SP/VEs according to their channel conditions as well as the number of users they have resulting in improving the efficiency of resource scheduling.

5.3.3 Evaluation of The e-VPS Algorithm

Last but not least, in this part, we evaluate the e-VPS approach we proposed in Section 5.2. We assume the same network parameters in Table 9. We assume that the number of users who have URLLC requests is a uniform random variable with a mean of $0.3 \times U_x$ where $x = 1, 2$ and U_x is the number of users belongs to SPs 1 and 2, respectively. Also, we use a URLLC packet size that is exponentially distributed with an average size of $L = 100$ Bytes. The target level of reliability for the produced figures is 0.99. After running our simulations for 1000 iterations, we have 4 users of URLLC requests at SP 1 and 3 users of URLLC requests at SP 2.

Figures 44 and 45 show the average link reliability versus the percentage of URLLC traffic in 700 MHz and 2100 MHz spectrum sharing, respectively. As shown in the figures, the reliability achieved via the e-VPS scheme outperforms the traditional LTE CQI-mapping in 700 MHz band. This is because the proposed e-VPS approach uses the end-to-end link reliability as the target

optimization objective function unlike the traditional LTE MCS table which is fundamentally established to maximize the user throughput with a relatively loose constraint on the channel error rate [77]. However, in the 2100 MHz band, neither of the schemes achieve the 99.999 % reliability. More clearly, the e-VPS scheme adjusts the channel coding rates for different users based on the modulation scheme, the received SINR and the transmission power levels to attain the target URLLC reliability, i.e. ≥ 0.999, but limited to low frequencies where channels have enough quality for ulta-reliable connections.

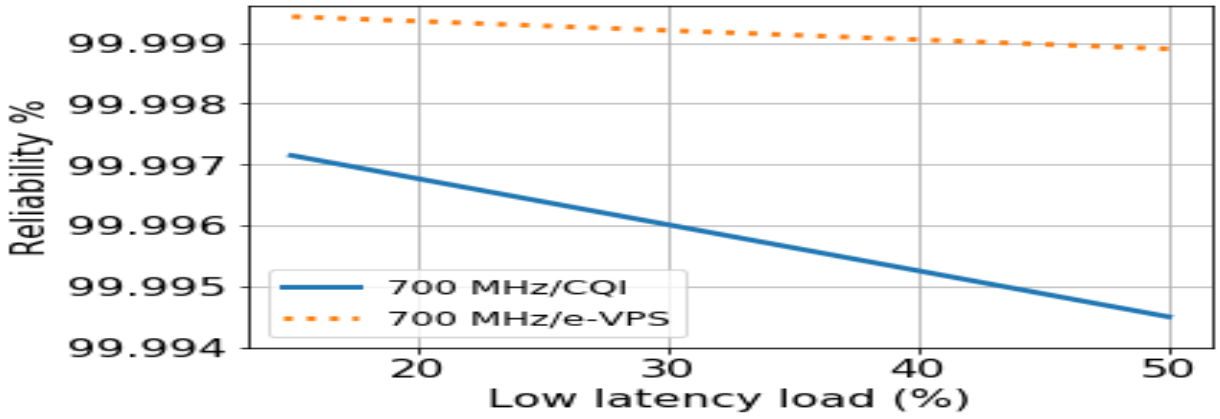


Figure 44: Reliability Vs URLLC traffic load in LTE virtual networks-700 MHz

Figures 46, 47 and 48 show the total throughput of our proposed algorithm versus NetShare in 700, 2100 and 3500 MHz bands, respectively. It is clear that higher frequency bands have worse channel quality due to lower average received signal strength, especially for users further away from the BS. And hence the enhancement gap between our proposed e-VPS (with worst channels for RT) and NetShare gets lower when using higher frequency bands.

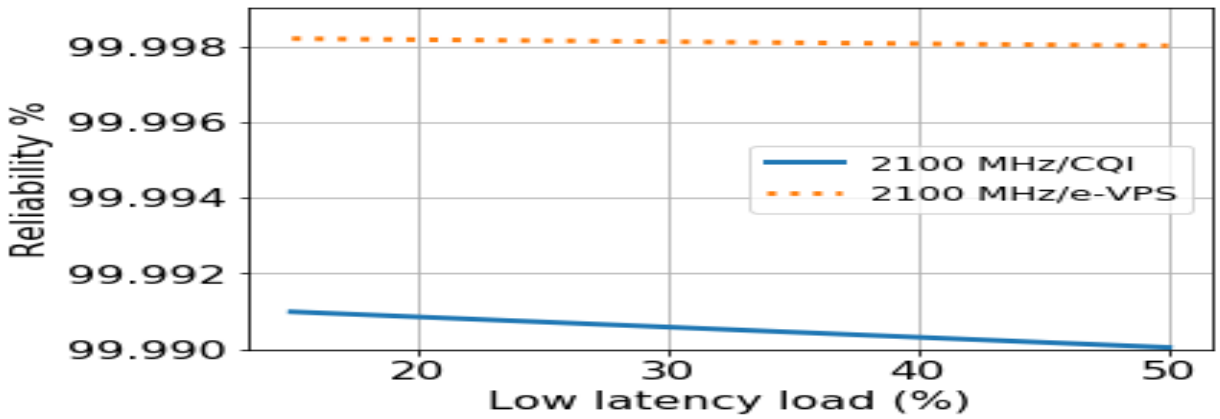


Figure 45: Reliability Vs URLLC traffic load in LTE virtual networks-2100 MHz

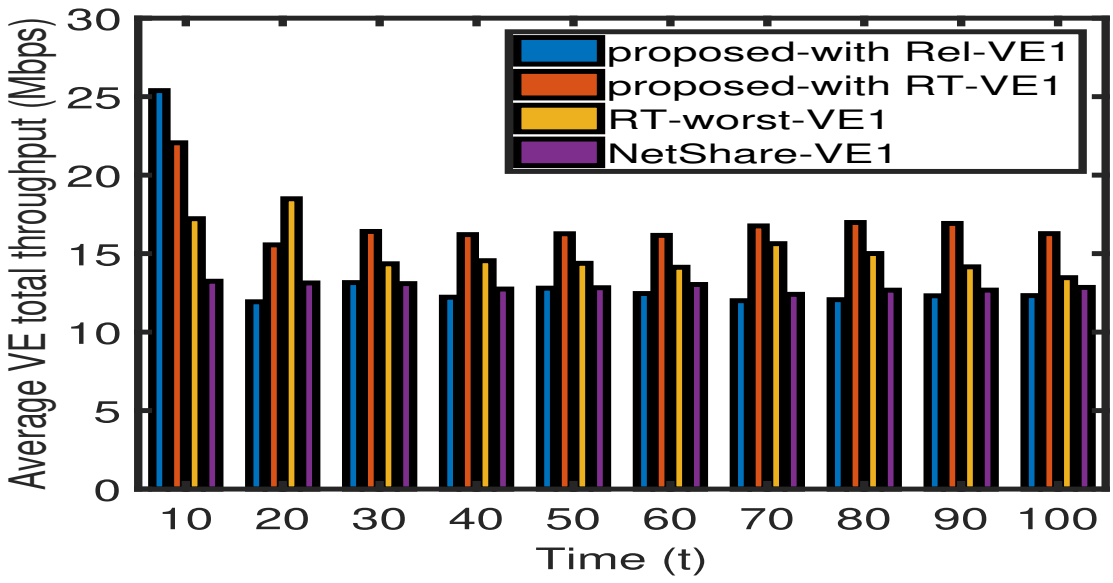


Figure 46: Total throughput for VE1 Vs. time in LTE virtual networks-700 MHz

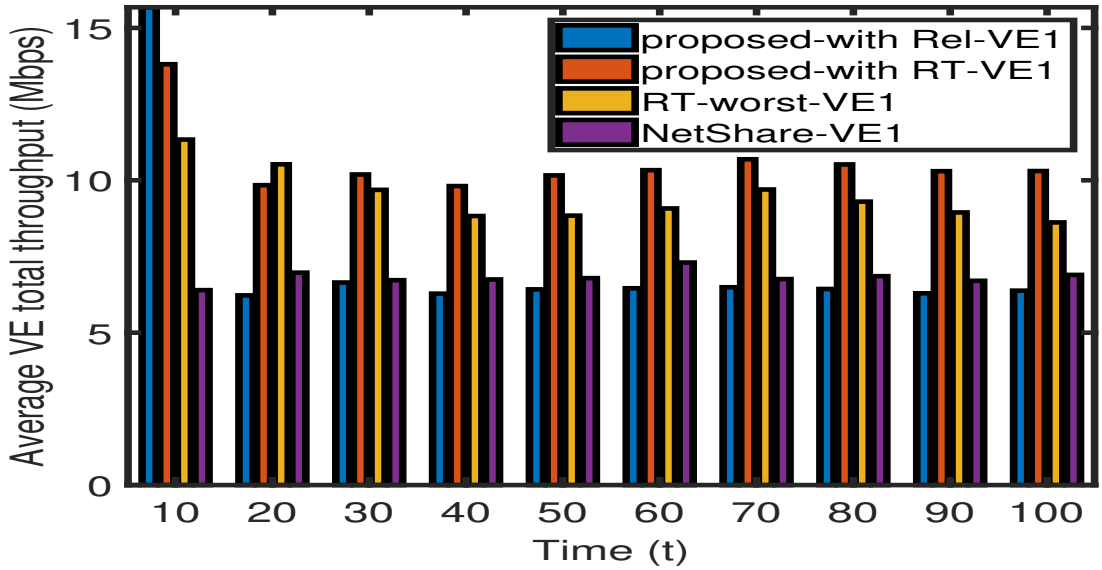


Figure 47: Total throughput for VE1 Vs. time in LTE virtual networks-2100 MHz

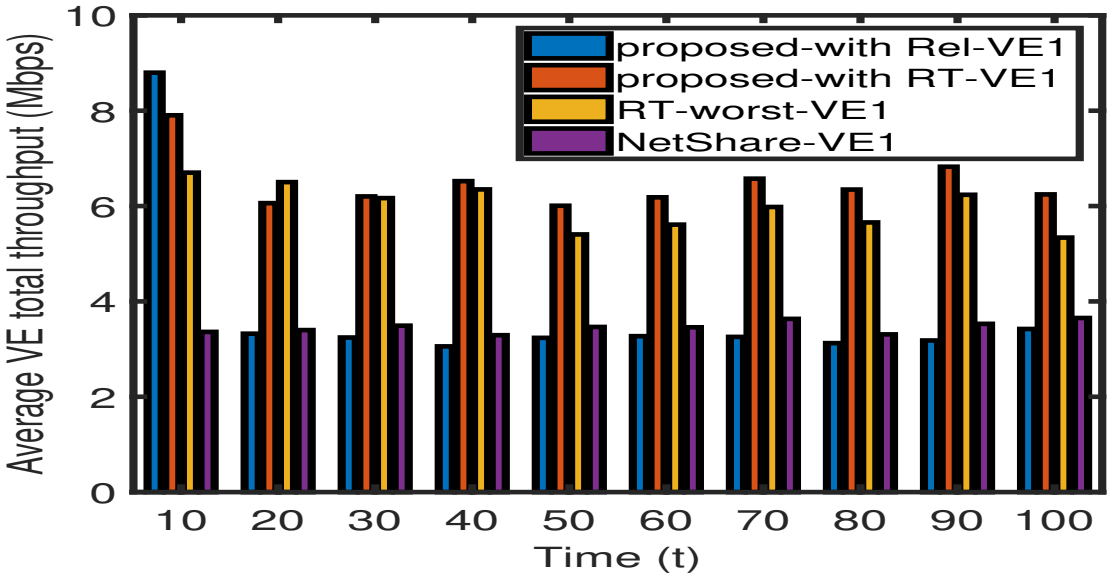


Figure 48: Total throughput for VE1 Vs. time in LTE virtual networks-3500 MHz

Also,, Figure 49 shows a degradation in the average user throughput using e-VPS due to higher coding rates when the SNR decreases with cell radius increases. Nonetheless, the e-VPS through-

put still satisfies the URLLC applications' target.

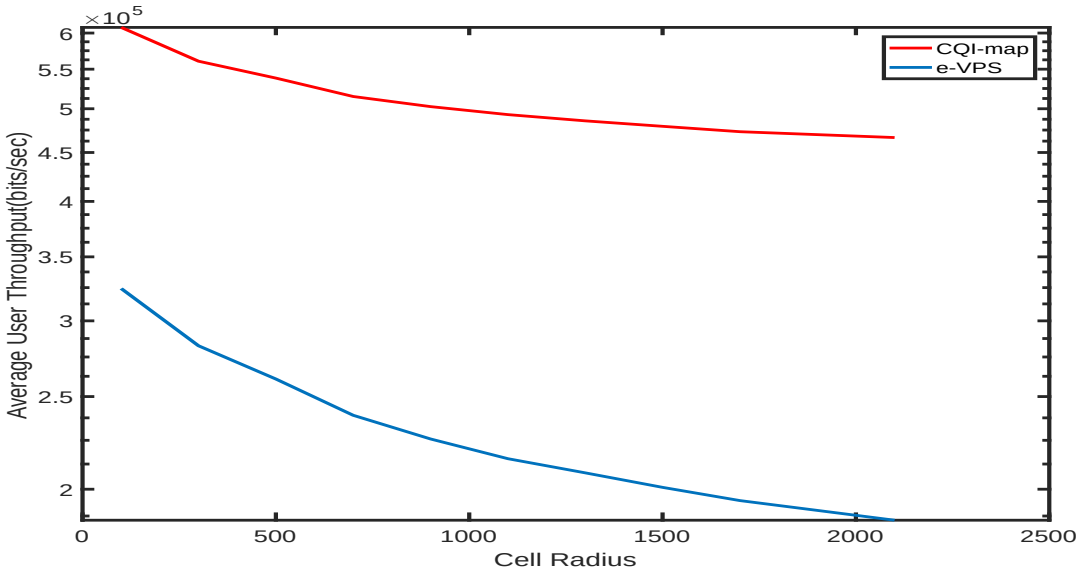


Figure 49: Average user throughput Vs Cell Radius in e-VPS LTE virtual networks

6.0 RECOVERING POST-DISASTER CELLULAR NETWORKS USING QOS-AWARE SPECTRUM POOLING

In this chapter, we examine the feasibility of recovering the disconnected cellular network using a wireless mesh network between SBSs such that the UE QoS and power constraints are still viable. In particular, we maximize the sum of the throughput across all cells such that the inter-links (wireless links between BSs) as well as the intra-links (wireless links between a BS and corresponding users) are maintained while maintaining the minimum average throughput per link as well as satisfying the traffic demand and QoS constraints. We also allocate different bands, i.e., 700, 2100 and 3500, to SBSs based on the traffic demand to enhance certain QoS such as ultra-reliability

6.1 SYSTEM MODEL

In this section, we state the assumptions of the cellular network model used in this thesis. Moreover, we briefly recall from the literature the definitions and the expressions of user throughput taking into consideration both real-time (RT) and non-real time (NRT) traffic requests.

6.1.1 System Model and Assumptions

We assume an LTE network of $|\Upsilon|$ BSs in a certain area A, where Υ is the set of all BS. Υ_c denotes the set of actively connected SBSs in area A. These SBSs belong to different mobile network operators (MNOs) and cover various cell ranges, i.e., macro cell of radius (R_{mac}), micro cell of radius (R_{mic}) and femto cell of radius (R_{fe}). Also, we assume the total BW available

at BS i is denoted by B_i. We use **a** as the subscript/superscript of the intra-links and **e** as the subscript/superscript of the inter-links. We adopt the Okumura-Hata model [78] as the path-loss propagation model over intra-links with a path loss in dB of

$$G_{ok}(dB) = U + Z\log(r) - E \tag{6.1}$$

where U is given by

$$U = 69.55 + 26.16\log(F_i^a) - 13.82\log(h_b), \tag{6.2}$$

Z is given by

$$Z = 44.9 - 6.55\log(h_b) \tag{6.3}$$

and E can be written as

$$E = 3.2\log(h_m^2) - 4.97 \tag{6.4}$$

where F_i^a is the operating frequency at cell i for intra-links, h_b is the BS antenna height, h_m is the UE antenna height and r is the distance between the BS and the UE.

For the inter BS links forming the mesh network, we assume that Line of Sight (LoS) links between SBSs exist and are subject to free space path loss which is given as

$$H_{ij}(dB) = 20\log(\frac{4\pi F_{ij}^e}{\nu}) + 20\log(d_{ij}) \tag{6.5}$$

where F_{ij}^e denotes the operating frequency over the inter-cell link between SBSs i and j, ν is the speed of light (3×10^8 meters/sec) and R_{ij} is the distance between SBS i and j.

For intra-cell links, the probability of coverage p_c at a certain cell is defined as the probability that a randomly chosen user can achieve a target signal-to-interference plus noise ratio $SINR_i^a$, i.e.,

$$p_c \triangleq \mathbb{P}\left[SINR_i^a > T\right] \tag{6.6}$$

In other words, a user is in coverage when the received SINR from the nearest BS is larger than some threshold T. The downlink SINR at the UE at a distance r from its associated BS can be expressed in dB as

$$SINR_i^a = P_i^a - G_{ok} - I_i^a - \sum_{x \neq i} I_x^e - \sigma^2 \tag{6.7}$$

where P_i^a is base station transmit power on intra-link frequency F_i^a the G_{ok} is the path-loss given in equation (6.1), I_i^a is the interference power at the UE receiver at cell i (i.e., the sum of the received

powers from all other base stations other than the home base station), I_x^e is the interference power from inter-cell links and σ^2 is the variance of the AWGN.
Similarly, the received SINR over an inter-cell link ij can be written as

$$SINR_{ij}^e = P_{ij}^e - H_{ij} - I_i - I_j - \sigma^2 \tag{6.8}$$

where P_{ij}^e is the transmitting power over link ij on frequency F_{ij}^e, I_i is the total interference power originating from cell i, I_j is the total interference power originating from cell j and H_{ij} is the free space pathloss from equation (6.5). It is worth noting that we assume that the interference over inter-cell links occurs only due to intra-cell connections. This is due to our assumption in this paper that all inter-cell links are established as a number of point-to-point single hop connections and hence they do not cause interference to each other.

According to [78], the capacity of a fading channel with receiver Channel Side Information (CSI) is given by

$$C = \int_T^\infty B \times \log_2(1 + SINR) \times Prob(SINR) \quad dSINR \tag{6.9}$$

where B is the bandwidth of the flat fading channel, SINR is the received SINR at the receiver (at the UE in case of intra-links and at the SBS in case of inter-links), T is the SINR threshold and Prob(SINR) is the probability distribution function (PDF) of the SINR.

6.1.2 Average Link Throughput in an LTE Network

In LTE networks, there are two major types of traffic: elastic traffic and real time traffic. Elastic traffic, such as web browsing and FTP, is generated by non-real time applications and carried over TCP. On the other hand, real time traffic, such as streaming, conferencing and VoIP, is very sensitive to delay and requires specific requirements to be transmitted. In this paper, we assume that the UE is able to use only $1/Q$ portion of the cell's PRBs even when there are no other active UEs, where Q is a parameter. The throughput of NRT and RT links can be calculated as follows.

6.1.2.1 RT Link Throughput RT applications such as video streaming and voice are known to have constant bit rate. More clearly, there should be a guaranteed number of PRBs every transmission time interval (TTI) dedicated to the RT flow. Thus, assuming that the RT flow arrival is Markovian, we can model the RT link as an M/D/1 queue with a service time of $1/\mu$. Assuming that every user sends a single flow per unit time and an RT flow size of L bits, the average RT link throughput in bits per unit time can be given as

$$\Phi_{u,rt} = \frac{L}{\mathbb{E}(T_{rt})} \tag{6.10}$$

where L is the RT flow size and $\mathbb{E}(T_{rt})$ is the average delay over the RT link, i.e. the total time taken to serve an RT flow of size L (service time + waiting time), and is given as [79]

$$\mathbb{E}(T_{rt}) = \frac{2 - \rho_{rt,i}}{2\mu(1 - \rho_{rt,i})} \tag{6.11}$$

where $1/\mu$ is the mean service time for an RT flow and $\rho_{rt,i}$ is the PRB utilization for RT traffic in cell i. According to 3GPP standardization, we recall that the PRB bandwidth in LTE systems is 180 KHz (Orthogonal Frequency Division Multiplexing), the average cell spectral efficiency is η_i bits/sec/Hz and the total number of PRBs available at cell i is Rb_i, the PRB utilization over the RT links per cell can be written as i.e.

$$\begin{aligned} \rho_{rt,i} &= \frac{L\lambda^i_{rt}}{C^i_{rt}} \\ &= \frac{L\lambda^i_{rt}}{180 \times \eta \times Rb_{rt,i}} \\ &= \frac{D}{\eta \times Rb_{rt,i}} \end{aligned} \tag{6.12}$$

where λ^i_{rt} is the average RT flow arrival rate (per unit time) and C^i_{rt} is the cell capacity in bits per unit time. Note that $D = L\lambda^i_{rt}/180$ for simplicity.

η is a function of the operating frequency (F^a_i and F^e_{ij}) and the transmitting power (P^a_i and P^e_{ij}). Further η can be related to the channel quality indicator (CQI) which is a function of the SINR according to [35] as

$$\eta = f(CQI) = f\left(\frac{SINR(dB) + 55/7}{13/7}\right) \tag{6.13}$$

where SINR(dB) is given by equations (6.7) and (6.8). The function $f()$ is given by the modulation coding schedule (MCS) given in the 3GPP standard [80]. Accordingly, we can put the average RT link throughput in cell i, $\Phi_{u,rt}$ in equation (6.10) as a function of the number of PRBs allocated to BS i, namely Rb_i, and the intra-link transmitting power P_i as well as the inter-links transmitting power from BS i to BS j ($P_{i,j}$) as follows

$$\Phi_{u,rt} = \begin{cases} \mu \frac{\eta(P_i^a) \times Rb_i^a - D}{2\eta(P_i^a) \times Rb_i^a - D} & \textit{for intra-links} \\ \\ \mu \frac{\eta(P_{ij}^e) \times Rb_i^e - D}{2\eta(P_{i,j}^e) \times Rb_i^e - D} & \textit{for inter-links} \end{cases} \tag{6.14}$$

The LTE cell capacity is calculated using the Shannon bound in equation (6.9). Also, in this chapter, we limit our analysis to a single class of the RT applications for simplicity.

6.1.2.2 NRT link Throughput For elastic traffic and assuming a proportional fairness (PF) scheduler, the BS can be modelled as an $M/G/Q_M$ processor sharing ($M/G/Q_M - PS$) queue [81]. Here $Q_M = Rb_i/Q$ defines the fraction of PRBs a UE is allowed to utilize (a maximum bound to the number of PRBs can be utilized by a single user in a cell). Then, the PRB utilization can be calculated as

$$\rho_{nrt,i} = \frac{S\lambda_{nrt}^i}{C_{nrt}^i} \tag{6.15}$$

where S is the average NRT flow size (bits), λ_{nrt}^i is the flow arrival rate (per unit time) and C_{nrt}^i is the radio link capacity for NRT (in bits per unit time). The radio link capacity available for NRT is a function of the RT traffic since it is computed after provisioning the higher priority traffic classes (real-time traffic in our model). Therefore, the NRT link capacity can be given as $C_{nrt}^i = C^i - \sum_{\ell} \Phi_{rt,i}$, where C^i is the total capacity of the radio link at a single BS i, Φ_{rt} is the RT link throughput (estimated from equation (6.10)) and $\sum_{\ell} \Phi_{rt,i}$ is the total throughput in bits per unit time consumed at BSi for all RT links $\in \ell$.

Based on the $M/G/Q_M - PS$ queue modelling the NRT traffic provision, the mean link delay in the NRT case is given as [79]

$$E[T_{nrt}] = \frac{S}{C_{nrt}^i(1 - \rho_{nrt,i})} \tag{6.16}$$

From queueing theory [82, 79], the average link throughput in a $M/G/Q_M - PS$ queue can be calculated by the following formula

$$\Phi_{u,nrt} = \frac{C^i_{nrt}(1-\rho_{nrt,i})}{\rho_{nrt,i}} \ln \frac{1}{1-\rho_{nrt,i}} \tag{6.17}$$

where $\Phi_{u,nrt}$ is the average throughput for a single NRT link initiated at cell i.
Similar to the RT scenario, we can put the average NRT link throughput in cell i $\Phi_{u,nrt}$ in equation (6.17) as a function of Rb_i, P^a_i and P^e_{ij}. Given that the radio link utilization for cell i in equation (6.15) can be written as $\frac{K}{\eta(P_i)\times Rb_i}$, where $K = S\lambda_{nrt}$ and the average cell spectral efficiency η is a function of the transmitting power P_i, the average NRT link throughput can be written as

$$\Phi_{u,nrt} = \begin{cases} \eta Rb^a_i \ (\eta \times Rb^a_i - K) \ \ln(\frac{\eta \times Rb^a_i}{\eta \times Rb^a_i - K}) & \textit{intra-links} \\ \\ \eta Rb^e_i \ (\eta \times Rb^e_i - K) \ \ln(\frac{\eta \times Rb^e_i}{\eta \times Rb^e_i - K}) & \textit{inter-links} \end{cases} \tag{6.18}$$

In Section 6.3, we use the RT and NRT average link throughput from equations (6.10) and (6.18), respectively, in the objective function of our joint power and PRB optimization problem.

6.2 RECONFIGURING THE CELLULAR NETWORK AFTER DISASTERS

Prior to any disaster, a cellular network operates in the default mode (D-mode) where the Radio Access Network (RAN) is connected to the core network as shown in Figure 50. Due to power outage and/or failures BSs that do not have a backup source of power are disconnected from the remaining network (inactive BSs in black, Figure 51). Moreover, the backhaul links for both SPs A and B shown in Figure 50 are affected and the connection between the RAN and the core network is cut off. Hence, we deploy an Emergency Communication Network Server (ECNS) where some critical functions of the core network and other auxiliary services are executed locally at the RAN to help resume network activity.

For example, the ECNS would perform authentication/security management, timing services, mobility management, PSAP (911) services and Domain Name System (DNS) service. Meanwhile, the SBSs operate on power supplied locally via the shared microgrid as shown in Figure

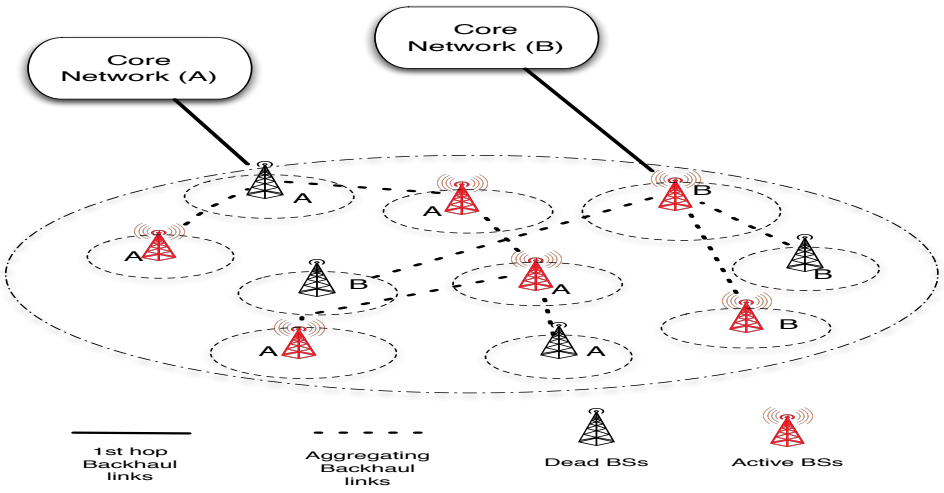


Figure 50: A Cellular network at no disaster - Normal Mode

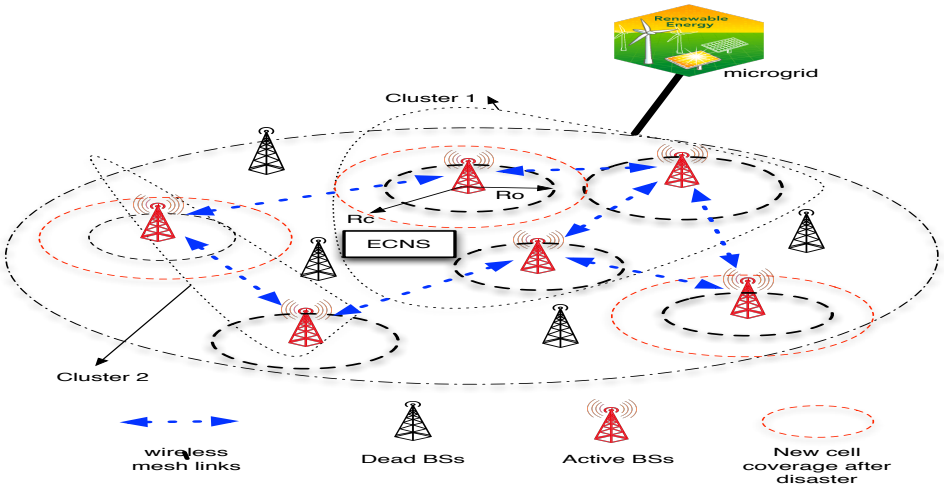


Figure 51: A Cellular network operating on power grid for network recovery - Emergency Mode

51. On the communication side, to enable mobile users to establish sessions and communicate without a core network, we establish a multi-hop wireless mesh network between SBSs as shown in Figure 51. Note,the SBSs must also communicate with the ECNS to enable the communications to the ECNS and between UEs from different cells, inter-cell links are established between BSs. Moreover, the resource manager at the ECNS server pools the total bandwidth available in the disaster area taking into account the frequency reuse factor. For example, assume the cellular network shown by the diagram in Figure 52. Since many frequency bands b_i are reused in more than one cell, then the set of the pooled frequency bands can be computed as the Union set of all non-overlapping sets of frequency bands that are used in different cells, i.e. $\eth = \bigcup_i b_i$. Moreover, the ECNS server is responsible for finding the best routes for traffic between SBSs, i.e. inter-cell links. It is worth noting that we do not consider a fully connected mesh network, however we assume the traffic between different BSs are processed following a shortest-path routing protocol such as the Destination-Sequenced Distance-Vector Routing (DSDV) or the Ad Hoc On-Demand Distance Vector Routing (AODV) [83]. Selecting the optimal wireless routing protocol is out of the scope of this paper. In Section 6.3, we propose our Mobile Network Recovery Algorithm (MNRA) algorithm to redefine the transmitting power levels, the allocated resources as well as the operating frequencies for both the intra-links (F_i^a) and inter-links (F_{ij}^e) such that the sum of the throughput of all cells in the area of interest is maximized.

6.3 MOBILE NETWORK RECOVERY ALGORITHM (MNRA)

In this section, we formulate a joint power and PRB allocation optimization problem to determine the optimal power as well as the allocation of the number of PRBs to different SBSs based on the traffic demand, QoS and power constraints.

The first step in our MNRA algorithm is to pool the available spectrum in the disaster area into one group and reassign it between all SBSs. We denote the total available PRBs by $Rb_t = \sum_{i \in \eth} Rb_o^i$, where Rb_o^i is the number of PRBs that corresponds to frequency band b_i before the disaster happens and $\eth$ is the set of frequency bands available in area A as computed in Section 6.2. The second step is to determine the set of connected SBSs based on the flow matrix $\mathbb{M}$. $\mathbb{M}$ is a $|\Upsilon| \times |\Upsilon|$ matrix

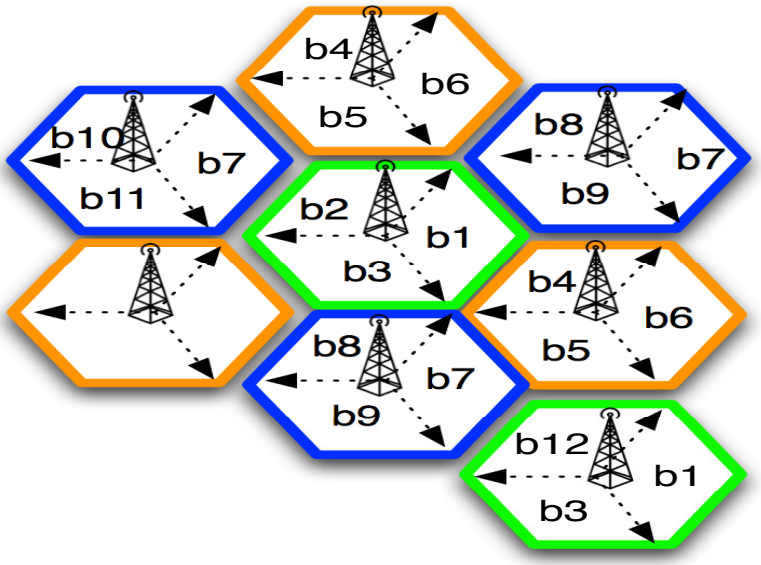


Figure 52: Network operating on multiple frequency bands with a frequency reuse factor of 3 and 120^o sectorization

where element $m(i,j)$ is the number of flows arriving over link ij, i.e. from BS i at BS j, per unit time and $|\Upsilon|$ is the cardinality of set Υ. Hence, the elements of the main diagonal of $\mathbb{M}$ represent the intra-cell flow arrivals (arrivals from UE to BS and vice versa). Now, we determine Υ_c, i.e. the set of actively connected SBSs, where $\Upsilon_c = \{i \in \Upsilon | m(i,i) \neq 0\}$. The next step is to select the inter-cell frequency $F_{ij}^e \in \Pi$, where Π is the set of operating frequencies adopted in the cellular system. Afterwards, for each $F_{ij}^e \in \Pi$ we solve the throughput maximization problem in (6.19). We denote optimal solution in every iteration y as (Δ_y), such that $\Delta \equiv (Rb_i^{a*}, Rb_i^{e*}, P_i^{a*}, P_i^{e*})$. Finally, we

derive the global optimal solution per each BS i as $\Delta^* = \underset{y}{\mathbf{Max}}\ (\Delta_y)$, for all $y = 1, 2, 3, ...|\Pi|$.

$$
\begin{aligned}
\underset{\Delta_i^*}{\text{maximize}} \quad & \sum_{i \in \Upsilon_c} \Phi^i_{u,UR} + \Phi^i_{u,rt} + \Phi^i_{u,nrt} \\
\text{subject to} \quad & \overline{SINR^a}_i \geq \gamma_{th}, \quad \forall i \in \Upsilon_c, \\
& \overline{SINR^e}_{ij} \geq \gamma_{th}, \quad \forall i, j \in \Upsilon_c, \\
& P^a_i < P_{max}, \ \forall i \in \Upsilon_c, \\
& P^e_{ij} < P_{max}, \ \forall i, j \in \Upsilon_c, \ i \neq j, \\
& E[T_{nrt}] \leq \tau, \\
& E[T_{rt}] \leq \zeta, \\
& E[T_{rt}] \leq \eta, \\
& \Phi^i_{u,rt} \geq \varphi_{rt}, \quad \forall i \in \Upsilon_c, \\
& \Phi^i_{u,nrt} \geq \varphi_{nrt}, \quad \forall i \in \Upsilon_c, \\
& F^e_{ij} \in \Pi, \quad \forall i, j \in \Upsilon_c, \ i \neq j, \\
& \sum_{i \in \Upsilon_c} Rb^a_i + Rb^e_i \leq Rb_t, \quad \forall i \in \Upsilon_c, \\
& Rb^a_i, Rb^e_i \geq 0, \quad \forall i \in \Upsilon_c, \\
& \sum_{j \in \Upsilon_c} W_j \leq P^g_v, \quad \forall j \in \Upsilon_c, \forall v \in \Omega, \\
& P^a_i, P^e_{ij}, P^g_v > 0, \quad \forall i, j \in \Upsilon_c, \forall v \in \Omega
\end{aligned}
\tag{6.19}
$$

Note that $\Delta_i = \left(\frac{Rb_t}{2||}, \frac{Rb_t}{2N_b}, \frac{P_{th}}{2}, P_{th}\right)$ is an initialization vector for Rb^{a*}_i, Rb^{e*}_i, P^{a*}_i and P^{e*}_i, respectively, to restrain the feasible region and attain the optimal solution in a lower number of iterations. Also, P_{max} is the maximum allowable transmitting power per single antenna according to FCC [52].

Also, $\Phi^i_{u,rt}$ and $\Phi^i_{u,nrt}$ are the average link throughput for RT and NRT traffic given in equations (6.14) and (6.18), respectively. The first two constraints are set to achieve the minimum average SINR per cell (γ_{th}) across both intra and inter links, $\overline{SINR}^a_i$ and $\overline{SINR}^e_{ij}$, respectively. The third and forth constraints are on the intra-link and inter-link transmitting powers at every BS i, respectively, to be no more than the maximum allowable transmitting power per single antenna by FCC (P_{max}). Recall that P^a_i is BS i transmitting power for intra-cell links. Also, $P^e_{i,j}$ is the transmitting

Algorithm 1 MNRA Algorithm

Input S,L,F_i^a,F_i^e, R_{mic},R_{mac},R_{fe},R_o^i,P_{th},γ_{th}

Output $\Delta^* = (Rb_i^{a*}, Rb_i^{e*}, P_i^{a*}, P_i^{a*})$

Calculate $Rb_t = \sum_{i \in \eth} R_o^i$

Find Υ_c

Initialize $\Delta_i = \left(\frac{Rb_t}{2|\Upsilon|}, \frac{Rb_t}{2|\Upsilon|}, \frac{P_{max}}{2}, P_{max}\right)$

for $y = 1 : |\Pi|$ **do**

Set $F_i^e = F(y) \in \Pi$

Find Δ_y = $\left(Rb_y^a, Rb_y^e, P_y^a, P_y^a\right)$ from (19)

$\Delta^* = \mathbf{Max}_y \ (\Delta_y)$

power on the inter-link from BS i to BS j. The fifth and sixth constraints are delay constraints for NRT and RT traffic requests, respectively. The expected value of the NRT delay $\mathbf{E}[T_{nrt}]$ and RT delay $\mathbf{E}[T_{rt}]$ are derived in equations (6.16) and (6.11) in Section 6.1. Next, the seventh and eighth constraints are to maintain a minimum target link (flow) throughput for both RT and NRT flows respectively. We recall that the average link throughput for RT and NRT requests are given by equations (6.10) and (6.18). The second to last constraint is a power consumption constraint on the sum of BS power consumption per cluster limited by the micro-grid capacity, where W_j is the total power consumption at BS j powered by power grid v that has a power capacity of P_v^g. However, according to [45], the LTE BS power consumption is considered negligible compared to other critical infrastructure consumption such as hospitals. Hence, we ignore this power consumption constraint in our optimization problem.

6.4 NUMERICAL RESULTS

In this section, we show some numerical results for the optimal solution of the MNRA optimization problem. Table 11 lists the parameters that are used throughout our simulations. We assume 10

BSs of different ranges as shown in Table 11. We assume 2 clusters each of 5 BSs, i.e. frequency reuse factor $= 5$. Each cluster has at least 1 Macro cell, 1 Micro Cell and 1 Pico cell. Further we assume that minimum hop routing is used. Optimizing the multi-hop paths between BSs is out of the scope of this paper. We denote the shortest path flow matrix as $\mathbb{M}$, in which element $m(i,j)$ is the number of flows per link (ij) transmitted from BS i to BS j and $m(i,i)$ represents the number of flows within the same cell. A sample of the input flow matrix is given in Table 11 (some are cut off due to space limitation). We model the path-loss over intra-links using the Okumura-Hata model [78] and inter-links as free space path loss as discussed in Section 6.1. We also assume that UEs are uniformly distributed across the cells.

The NRT traffic is modeled by an FTP application downloading a file of a size of 5 Mbytes with a minimum average user delay of 10 msec. The RT traffic is modeled by VoIP calls with a voice code rate of $\delta = 8$ Kbps, an average speech period of $\ell = 3$ seconds, average silence period of 3 seconds and a constant call duration of $d = 90$ seconds. Accordingly, the RT flow size L can be estimated as $L = 0.5 \times \delta \times \frac{d}{\ell}$. We assume 5 frequencies for inter-cell frequencies (850, 2100, 2900, 3500 and 5000 MHz). We also assume that intra-cell frequencies for the pico(femto), micro and macro cells are 2900, 850 and 700 MHz, respectively. We solve the optimization problem in (6.19) using the nonlinear constrained optimization tool in Python [84], (Sequential Least Squares Programming). The optimization is executed for 10 iterations and the results averaged over the different iterations. We also run the optimization for different values of SINR threshold (γ_{th}). We pick our initialization vector Δ_i such that the initial number of PRBs allocated to BSs is 10, the initial value for inter-cell transmission power is 14 dBm and the initial value for intra-cell transmission power is 30 dBm.

Figures 53 and 54 show the optimal number of PRBs allocated to intra-cell links and inter-cell links, respectively, versus different values of SINR threshold (γ_{th}). Also, we can clearly see that the optimal transmitting powers change with the change of the SINR threshold as shown in Figures 55 and 56. Interestingly, we can see also that the power allocations for inter-cell links are lower than the intra-cell power allocations. The reason is that the LoS free space path loss is much less than the Okumura-Hata path loss. Moreover, it is clear that femto cells use higher transmitting powers than micro and macro BSs to compensate for the high attenuation due to higher operating frequency.

Table 11: Values of Generic Network Parameters

Macro Cell radius	1500
Micro Cell radius	500
Pico Cell radius	100
BS antenna height	35 meters
UE antenna height	1.5 meters
Max power (FCC) (P_{max})	37 dBm
Total Number of PRBs (Rb_t)	100
Noise Variance (σ^2)	-90 dBm
F^a (femto)	2900 MHz
F^a (micro)	1800 MHz
F^a (macro)	700 MHz
FTP traffic model	
File Size	5 Mbyte
Maximum NRT flow delay $E(T_{nrt})$	50 msec
Minimum NRT user throughput (ϕ_{nrt})	10 Mbps
VOIP traffic model	
voice codec δ	8 kbps coding rate
call activity (ℓ)	speech:exp.dist.(3 sec)
	silence:exp.dist.(3 sec)
call duration (d)	90 seconds
Maximum RT flow delay $E(T_{rt})$	1 msec
Minimum RT user throughput (ϕ_{rt})	0.5 Mbps

Table 12: Shortest Path Flow Matrix (number of flows between BSs)

BS	1	2	3	...	10
1	2	3	2	...	3
2	7	4	8	...	4
3	8	5	2	...	1
.	.	.	.	...	.
.	.	.	.	...	.
10	4	2	3	...	4

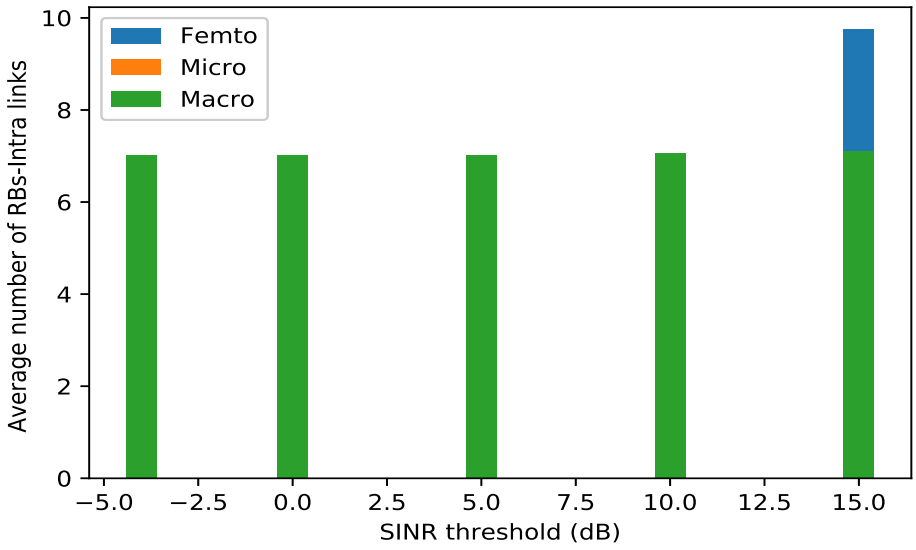


Figure 53: PRBs allocated to intra-links vs γ_{th}

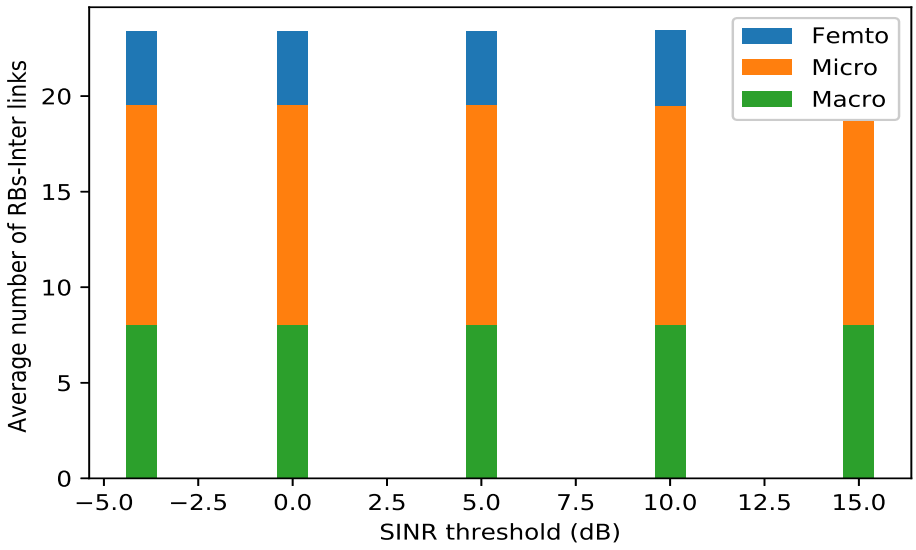


Figure 54: PRBs allocated to inter-links vs γ_{th}

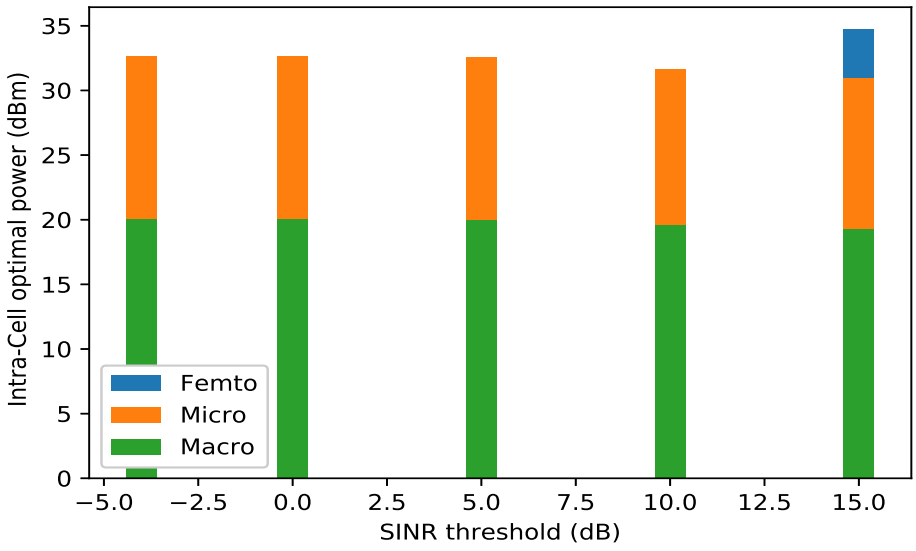


Figure 55: Average intra-cell transmit power vs γ_{th}

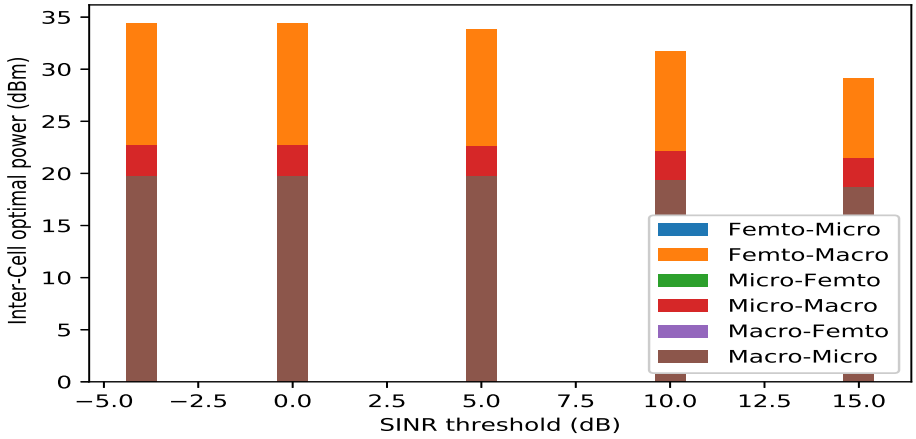


Figure 56: Average inter-cell transmit power vs γ_{th}

Figure 57 shows the optimal inter-cell frequency per cell for the pico, micro and macro cells versus the SINR threshold. It is obvious that the inter-cell links acquire for the highest available frequencies, i.e. 3500 and 5000 MHz frequency bands. This is to keep the signal attenuation and inter-link interference to minimum and achieve the target QoS per link. Figure 58 shows the average total throughput (for all cells) in bits per seconds versus the SINR threshold (γ_{th}). We can see that $throughput_{femto} > throughput_{micro} > throughput_{macro}$. This is due to that small cell networks are much more dense than micro and macro cells.

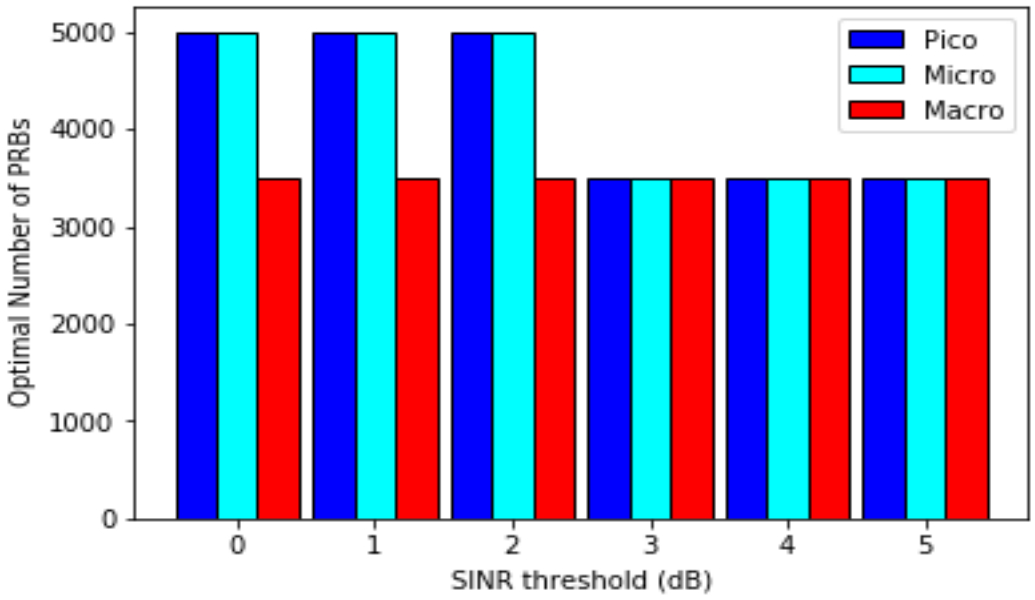


Figure 57: Inter-cell frequency (MHz) vs γ_{th}

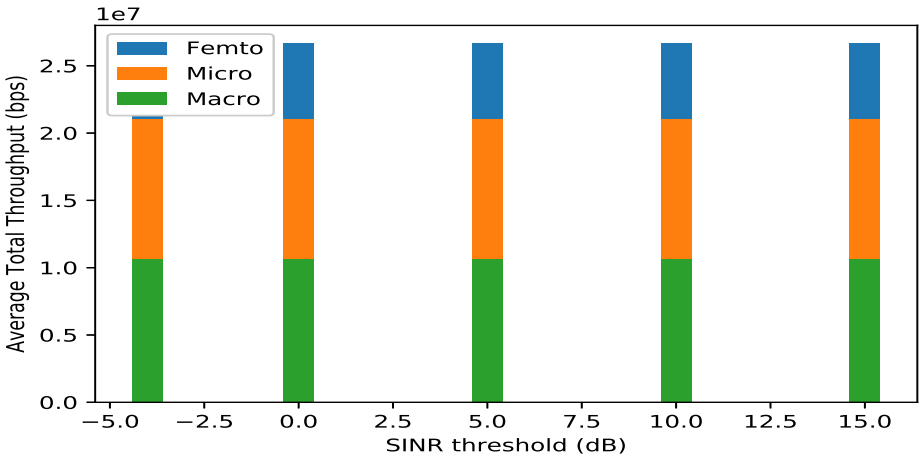


Figure 58: Average Total Throughput per BS vs γ_{th}

7.0 CONCLUSIONS

In this thesis, we considered three problems related to spectrum pooling for virtualized future wireless networks including the economic and technical aspects. We studied infrastructure and spectrum sharing to assess the economic incentives of different sharing scenarios in greenfield LTE networks. We showed that spectrum sharing between SPs has more incentive in urban CMAs with high population densities due to the ultra-high spectrum bidding value especially when many BSs share spectrum. Moreover, sharing incentive is influenced by the number of BSs that adopt sharing, the allocated BW and the frequency band (700 MHz or 2100). Also, we developed two resource allocation schemes to allocate PRBs in 5G virtualized networks such that fairness, isolation and heterogeneity between SPs are taken into consideration. We assume the provision of RT, NRT and URLLC traffic. We also designed a PF-based minimum value of the PRBs allocated to SPs to achieve high throughput and fairness. We show using numerical results that our proposed schemes outperforms NetShare, a recently proposed resource allocation scheme. We also show that we achieve the ultra-reliability constraint of URLLC while maintaining the target throughput with a new optimal channel coding algorithm. Eventually, we use different gains from inter-band sharing and intra-band sharing to enable disaster recovery in cellular networks. We share the spectrum between SBSs to form a wireless mesh network between RANs. We optimally design the wireless mesh network by finding the optimal allocated bandwidth, the optimal transmission power and the optimal frequency bands allocated to different SBSs to mitigate interference as well as to achieve different QoS constraints for URLLC, real-time and non-real time traffic.

BIBLIOGRAPHY

[1] S.-Y. Lien, K.-C. Chen, and Y. Lin, "Toward ubiquitous massive accesses in 3gpp machine-to-machine communications," *IEEE Communications Magazine*, vol. 49, no. 4, 2011.

[2] "Ericsson mobility report-june 2017." https://www.ericsson.com/assets/local/mobility-report/documents/2017/ericsson-mobility-report-june-2017.pdf, accessed:2017-06-30.

[3] "5G requirements (SMARTER, NexGen)," http://www.3gpp.org/release-14, accessed:2018-06-01.

[4] N. Alliance, "5g white paper," *Next Generation Mobile Networks, White paper*, 2015.

[5] A. Ericsson, "Shared networks–white paper: An operator alternative to reduce initial investments, increase coverage and reduce time to market for wcdma by sharing infrastructure."

[6] D. Benhaddou and W. Alanqar, "Layer 1 virtual private networks in multidomain next-generation networks," *IEEE Communications Magazine*, vol. 45, no. 4, 2007.

[7] T. Takeda, "Framework and requirements for layer 1 virtual private networks," 2007.

[8] P. Barham, B. Dragovic, K. Fraser, S. Hand, T. Harris, A. Ho, R. Neugebauer, I. Pratt, and A. Warfield, "Xen and the art of virtualization," *ACM SIGOPS Operating Systems Review*, vol. 37, no. 5, pp. 164–177, 2003.

[9] A. Haider, R. Potter, and A. Nakao, "Challenges in resource allocation in network virtualization," in *20th ITC Specialist Seminar*, vol. 18, no. 2009. ITC, Hoi An Vietnam, 2009.

[10] N. M. K. Chowdhury and R. Boutaba, "A survey of network virtualization," *Computer Networks*, vol. 54, no. 5, pp. 862–876, 2010.

[11] X. Wang, P. Krishnamurthy, and D. Tipper, "Wireless network virtualization," in *Computing, Networking and Communications (ICNC), 2013 International Conference on*. IEEE, 2013, pp. 818–822.

[12] J. M. Peha, "Sharing spectrum through spectrum policy reform and cognitive radio," *Proceedings of the IEEE*, vol. 97, no. 4, pp. 708–719, 2009.

[13] T. Frisanco, P. Tafertshofer, P. Lurin, and R. Ang, "Infrastructure sharing and shared operations for mobile network operators from a deployment and operations view," in *Network Operations and Management Symposium, 2008. NOMS 2008. IEEE*. IEEE, 2008, pp. 129–136.

[14] A. Khan, W. Kellerer, K. Kozu, and M. Yabusaki, "Network sharing in the next mobile network: Tco reduction, management flexibility, and operational independence," *IEEE Communications Magazine*, vol. 49, no. 10, 2011.

[15] J. S. Panchal, R. D. Yates, and M. M. Buddhikot, "Mobile network resource sharing options: Performance comparisons," *IEEE Transactions on Wireless Communications*, vol. 12, no. 9, pp. 4470–4482, 2013.

[16] J. Kibiłda, P. Di Francesco, F. Malandrino, and L. A. DaSilva, "Infrastructure and spectrum sharing trade-offs in mobile networks," in *Dynamic Spectrum Access Networks (DySPAN), 2015 IEEE International Symposium on*. IEEE, 2015, pp. 348–357.

[17] M. Haenggi, J. G. Andrews, F. Baccelli, O. Dousse, and M. Franceschetti, "Stochastic geometry and random graphs for the analysis and design of wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 7, 2009.

[18] M. Bansal, J. Mehlman, S. Katti, and P. Levis, "Openradio: a programmable wireless dataplane," in *Proceedings of the first workshop on Hot topics in software defined networks*. ACM, 2012, pp. 109–114.

[19] L. Suresh, J. Schulz-Zander, R. Merz, A. Feldmann, and T. Vazao, "Towards programmable enterprise wlans with odin," in *Proceedings of the first workshop on Hot topics in software defined networks*. ACM, 2012, pp. 115–120.

[20] J. Schulz-Zander, C. Mayer, B. Ciobotaru, S. Schmid, and A. Feldmann, "Opensdwn: programmatic control over home and enterprise wifi," in *Proceedings of the 1st ACM SIGCOMM Symposium on Software Defined Networking Research*. ACM, 2015, p. 16.

[21] A. Gudipati, D. Perry, L. E. Li, and S. Katti, "Softran: Software defined radio access network," in *Proceedings of the second ACM SIGCOMM workshop on Hot topics in software defined networking*. ACM, 2013, pp. 25–30.

[22] A. Checko, H. L. Christiansen, Y. Yan, L. Scolari, G. Kardaras, M. S. Berger, and L. Dittmann, "Cloud ran for mobile networks—a technology overview," *IEEE Communications surveys & tutorials*, vol. 17, no. 1, pp. 405–426, 2015.

[23] Y. Zaki, L. Zhao, C. Goerg, and A. Timm-Giel, "A novel lte wireless virtualization framework," in *International Conference on Mobile Networks and Management*. Springer, 2010, pp. 245–257.

[24] R. Kokku, R. Mahindra, H. Zhang, and S. Rangarajan, "Nvs: a virtualization substrate for wimax networks," in *Proceedings of the sixteenth annual international conference on Mobile computing and networking*. ACM, 2010, pp. 233–244.

[25] R. Mahindra, M. A. Khojastepour, H. Zhang, and S. Rangarajan, "Nec laboratories america inc., princeton, usa," in *Network Protocols (ICNP), 2013 21st IEEE International Conference on*. IEEE, 2013, pp. 1–10.

[26] G. Liu, F. R. Yu, H. Ji, and V. C. Leung, "Distributed resource allocation in full-duplex relaying networks with wireless virtualization," in *Global Communications Conference (GLOBECOM), 2014 IEEE*. IEEE, 2014, pp. 4959–4964.

[27] F. Römer, J. Zhang, M. Haardt, and E. Jorswieck, "Spectrum and infrastructure sharing in wireless networks: A case study with relay-assisted communications," *Proc. Future Network and Mobile Summit 2010*, 2010.

[28] M. I. Kamel, L. B. Le, and A. Girard, "Lte wireless network virtualization: Dynamic slicing via flexible scheduling," in *Vehicular Technology Conference (VTC Fall), 2014 IEEE 80th*. IEEE, 2014, pp. 1–5.

[29] M. Yang, Y. Li, J. Liu, D. Jin, J. Yuan, and L. Zeng, "Opportunistic spectrum sharing for wireless virtualization," in *Wireless Communications and Networking Conference (WCNC), 2014 IEEE*. IEEE, 2014, pp. 1803–1808.

[30] P. Bender, P. Black, M. Grob, R. Padovani, N. Sindhushyana, and A. Viterbi, "Cdma/hdr: a bandwidth efficient high speed wireless data service for nomadic users," *Communications Magazine, IEEE*, vol. 38, no. 7, pp. 70–77, 2000.

[31] S. Shenker, "Fundamental design issues for the future internet," *Selected Areas in Communications, IEEE Journal on*, vol. 13, no. 7, pp. 1176–1188, 1995.

[32] V. Pejovic and H. Zheng, "A real-time traffic packet scheduler for a novel tdma mac protocol."

[33] P. Popovski, "Ultra-reliable communication in 5g wireless systems," in *5G for Ubiquitous Connectivity (5GU), 2014 1st International Conference on*. IEEE, 2014, pp. 146–151.

[34] O. N. Yilmaz, Y.-P. E. Wang, N. A. Johansson, N. Brahmi, S. A. Ashraf, and J. Sachs, "Analysis of ultra-reliable and low-latency 5g communication for a factory automation use case," in *2015 IEEE International Conference on Communication Workshop (ICCW)*. IEEE, 2015, pp. 1190–1195.

[35] O. Østerbø, "Scheduling and capacity estimation in lte," in *Proceedings of the 23rd International Teletraffic Congress*. International Teletraffic Congress, 2011, pp. 63–70.

[36] E. K. Çetinkaya, D. Broyles, A. Dandekar, S. Srinivasan, and J. P. Sterbenz, "Modelling communication network challenges for future internet resilience, survivability, and disrup-

tion tolerance: A simulation-based approach," *Telecommunication Systems*, vol. 52, no. 2, pp. 751–766, 2013.

[37] K. Christensen, P. Reviriego, B. Nordman, M. Bennett, M. Mostowfi, and J. A. Maestro, "Ieee 802.3 az: the road to energy efficient ethernet," *IEEE Communications Magazine*, vol. 48, no. 11, 2010.

[38] Y.-N. Lien, L.-C. Chi, and Y.-S. Shaw, "A walkie-talkie-like emergency communication system for catastrophic natural disasters," in *Pervasive Systems, Algorithms, and Networks (ISPAN), 2009 10th International Symposium on*. IEEE, 2009, pp. 309–314.

[39] M. Iqbal, X. Wang, D. Wertheim, and X. Zhou, "Swanmesh: A multicast enabled dual-radio wireless mesh network for emergency and disaster recovery services." *JCM*, vol. 4, no. 5, pp. 298–306, 2009.

[40] G. Chen, A. Hu, and T. Sato, "A scheme for disaster recovery in wireless networks with dynamic ad-hoc routing," in *Kaleidoscope: Beyond the Internet?-Innovations for Future Networks and Services, 2010 ITU-T*. IEEE, 2010, pp. 1–6.

[41] R. Wishart, M. Portmann, and J. Indulska, "Evaluation of wireless mesh network handoff approaches for public safety and disaster recovery networks," in *Telecommunication Networks and Applications Conference, 2008. ATNAC 2008. Australasian*. IEEE, 2008, pp. 294–299.

[42] Y.-M. Lee, B.-J. Ku, and D.-S. Ahn, "A satellite core network system for emergency management and disaster recovery," in *Information and Communication Technology Convergence (ICTC), 2010 International Conference on*. IEEE, 2010, pp. 549–552.

[43] D. Abusch-Magder, P. Bosch, T. E. Klein, P. A. Polakos, L. G. Samuel, and H. Viswanathan, "911-now: A network on wheels for emergency response and disaster recovery operations," *Bell Labs Technical Journal*, vol. 11, no. 4, pp. 113–133, 2007.

[44] Q. T. Minh, K. Nguyen, C. Borcea, and S. Yamada, "On-the-fly establishment of multihop wireless access networks for disaster recovery," *IEEE Communications Magazine*, vol. 52, no. 10, pp. 60–66, 2014.

[45] A. Alqahtani, R. Abhishek, D. Tipper, and D. Medhi, "Disaster recovery power and communications for smart critical infrastructures," in *IEEE Internationl Conference On Communications, 2018. ICC 2018.* IEEE, 2018.

[46] A. Kwasinski and P. T. Krein, "Telecom power planning for natural and man-made disasters," in *Telecommunications Energy Conference, 2007. INTELEC 2007. 29th International*. IEEE, 2007, pp. 216–222.

[47] A. Kwasinski, W. W. Weaver, P. L. Chapman, and P. T. Krein, "Telecommunications power plant damage assessment for hurricane katrina–site survey and follow-up results," *IEEE Systems Journal*, vol. 3, no. 3, pp. 277–287, 2009.

[48] X. Wang, A. V. Vasilakos, M. Chen, Y. Liu, and T. T. Kwon, "A survey of green mobile networks: Opportunities and challenges," *Mobile Networks and Applications*, vol. 17, no. 1, pp. 4–20, 2012.

[49] A. Kwasinski and A. Kwasinski, "Architecture for green mobile network powered from renewable energy in microgrid configuration," in *Wireless Communications and Networking Conference (WCNC), 2013 IEEE*. IEEE, 2013, pp. 1273–1278.

[50] R. Nadiv and T. Naveh, "Wireless backhaul topologies: Analyzing backhaul topology strategies," *Ceragon White Paper*, pp. 1–15, 2010.

[51] P. Mogensen, W. Na, I. Z. Kovács, F. Frederiksen, A. Pokhariyal, K. I. Pedersen, T. Kolding, K. Hugl, and M. Kuusela, "Lte capacity compared to the shannon bound," in *2007 IEEE 65th Vehicular Technology Conference-VTC2007-Spring*. IEEE, 2007, pp. 1234–1238.

[52] "Base station (bs) radio transmission and reception (fdd) (3gpp ts 25.104 version 11.3.0 release 11)," http://www.etsi.org/deliver/etsi_ts/125100_125199/125104/11.03.00_60/ts_125104v110300p.pdf, accessed:2017-09-01.

[53] A. AbdelHamid, D. Tipper, and P. Krishnamurthy, "Recovery and optimization of post-disaster cellular networks," in *2019 15th International Conference on the Design of Reliable Communication Networks (DRCN)*. IEEE, 2019, pp. 16–20.

[54] A. AbdelHamid, P. Krishnamurthy, and D. Tipper, "Resource allocation for heterogeneous traffic in lte virtual networks," in *Mobile Data Management (MDM), 2015 16th IEEE International Conference on*, vol. 1. IEEE, 2015, pp. 173–178.

[55] A. Abdelhamid, P. Krishnamurthy, D. Tipper, and M. Weiss, "A techno-economic study of spectrum and infrastructure sharing in next-generation cellular networks," 2018, to be Submitted to Telecommunications policy , Elsevier.

[56] A. Abdelhamid, P. Krishnamurthy, and D. Tipper, "Resource scheduling for heterogeneous traffic innext-generation virtual mobile networks," 2018, submitted to European Alliance for Innovation in Mobile Communications and Applications.

[57] "Fcc auctions data," http://wireless.fcc.gov/auctions/, accessed:2015-01-25.

[58] "Fiber Store," http://www.fiberstore.com/, accessed:2015-06-28.

[59] Deloitte, "5G: The chance to lead for a decade," https://www2.deloitte.com/content/dam/Deloitte/us/Documents/technology-media-telecommunications/us-tmt-5g-deployment-imperative.pdf, 2018, accessed:2018-12-28.

[60] E. Amaldi, A. Capone, F. Malucelli, and C. Mannino, "Optimization problems and models for planning cellular networks," in *Handbook of optimization in telecommunications*. Springer, 2006, pp. 917–939.

[61] OpenCellid, "OpenCellid:The world's largest Open Database of Cell Towers," https://opencellid.org/, 2019, accessed:2019-01-30.

[62] I. Rec, "G. 984.1,"," *Gigabit-capable passive optical networks (GPON): General characteristics*, 2003.

[63] M. Hajduczenia, B. Lakic, H. J. da Silva, and P. P. Monteiro, "Optimized passive optical network deployment," *Journal of Optical Networking*, vol. 6, no. 9, pp. 1079–1104, 2007.

[64] A. D. Yablon, *Optical fiber fusion splicing*. Springer Science & Business Media, 2005, vol. 103.

[65] M. Deruyck, E. Tanghe, W. Joseph, and L. Martens, "Modelling and optimization of power consumption in wireless access networks," *Computer Communications*, vol. 34, no. 17, pp. 2036–2046, 2011.

[66] R. G. Beaves, "Net present value and rate of return: implicit and explicit reinvestment assumptions," *The Engineering Economist*, vol. 33, no. 4, pp. 275–302, 1988.

[67] E. U. T. R. Access, "Physical layer procedures (release 9), 3gpp," *Technical Specification Group Radio Access Network TS36*, vol. 213, 2010.

[68] E. Dahlman, S. Parkvall, and J. Skold, *4G: LTE/LTE-advanced for mobile broadband*. Academic Press, 2013.

[69] F. Kelly, "Charging and rate control for elastic traffic," *European transactions on Telecommunications*, vol. 8, no. 1, pp. 33–37, 1997.

[70] C. M. Aras, J. F. Kurose, D. S. Reeves, and H. Schulzrinne, "Real-time communication in packet-switched networks," *Proceedings of the IEEE*, vol. 82, no. 1, pp. 122–139, 1994.

[71] A. P. Snow, U. Varshney, and A. D. Malloy, "Reliability and survivability of wireless and mobile networks," *Computer*, vol. 33, no. 7, pp. 49–55, 2000.

[72] C. E. Shannon, "A mathematical theory of communication," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 5, no. 1, pp. 3–55, 2001.

[73] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime," *IEEE Transactions on Information Theory*, vol. 56, no. 5, pp. 2307–2359, 2010.

[74] V. Y. F. Tan and M. Tomamichel, "The third-order term in the normal approximation for the awgn channel," *IEEE Transactions on Information Theory*, vol. 61, no. 5, pp. 2430–2438, 2015.

[75] T. S. Rappaport *et al.*, *Wireless communications: principles and practice*. Prentice Hall PTR New Jersey, 1996, vol. 2.

[76] J. G. Proakis, "Digital communications, 1995."

[77] C. Mehlführer, M. Wrulich, J. C. Ikuno, D. Bosanska, and M. Rupp, "Simulating the long term evolution physical layer," in *Proc. of the 17th European Signal Processing Conference (EUSIPCO 2009), Glasgow, Scotland*, vol. 27, 2009, p. 124.

[78] A. Goldsmith, *Wireless communications*. Cambridge university press, 2005.

[79] L. Kleinrock, *Queueing systems, volume 2: Computer applications*. wiley New York, 1976, vol. 66.

[80] "User equipment (ue) radio transmission and reception (3gpp ts 36.101 version 9.1.0 release 9)," http://www.etsi.org/deliver/etsi_ts/136100_136199/136101/09.01.00_60/ts_136101v090100p.pdf, accessed:2017-09-01.

[81] K. Lindberger, "Balancing quality of service, pricing and utilisation in multiservice networks with stream and elastic traffic," *Teletraffic science and engineering*, pp. 1127–1136, 1999.

[82] A. A. Kherani and A. Kumar, "Stochastic models for throughput analysis of randomly arriving elastic flows in the internet," in *IEEE INFOCOM*, vol. 2. INSTITUTE OF ELECTRICAL ENGINEERS INC (IEEE), 2002, pp. 1014–1023.

[83] E. M. Royer and C.-K. Toh, "A review of current routing protocols for ad hoc mobile wireless networks," *IEEE personal communications*, vol. 6, no. 2, pp. 46–55, 1999.

[84] "Optimization and root finding (scipy.optimize)," https://docs.scipy.org/doc/scipy/reference/optimize.html#module-scipy.optimize, accessed:2017-09-01.