D2.1: Consolidated State of the Art Surveys (ESRs 3-4) and Individual Research Proposals

Project Name: Anticipatory Networking Techniques in 5G and Beyond

Acronym: ACT5G

Project no.: 643002

Start date of project: 01/05/2015 **Duration:** 48 Months

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Actions.

Document Properties

Document ID	EU-H2020-MSCA-ITN-2014-643002-ACT5G-D2.1
Document Title	D2.1: Consolidated State of the Art Surveys (ESRs 3-4) and Individual Research Proposals
Contractual date of delivery to REA	Month 10
Lead Beneficiary	Linköping University (LiU)
Editor(s)	Vangelis Angelakis, Di Yuan
Work Package No.	2
Work Package Title	Reaction Techniques
Nature	Report
Number of Pages	47
Dissemination Level	**PUBLIC**
Contributors	LiU: W. Li, V. Angelakis, D. Yuan. POLIMI: Ö. Umut Akgül, A. Capone, M. Cesana ALUD: I. Malanchini
Version	1.0

Table of Contents

1 Executive Summary **3**

2 Introduction **5**

3 State of the Art in Resource Allocation **8**

3.1 Massive MIMO 8
3.2 Millimetre Wave 10
3.3 Heterogeneous Networks 11
3.4 Non-orthogonal Multiple Access 12
3.5 Related Theory and Miscellaneous 13

4 ESR3 Research Proposal **16**

4.1 Context 16
4.2 Problem Description 17
4.2.1 Research Possibilities 17
4.2.2 Expected Results 22
4.2.3 Envisioned Methodology 22

5 State of the Art in Infrastructure Sharing **23**

5.1 Business Models 23
5.1.1 Revenue Modelling 24
5.1.2 Sharing Options 25
5.1.3 Analysis on Validity/Sustainability 27
5.2 Technical Models 30
5.2.1 Resource Sharing 30
5.2.2 Virtualization Technology 34

6 ESR-4 Research Proposal **36**

6.1 Context 36
6.2 Problem Definition & Model 36
6.2.1 Initial Assumptions 37
6.3 Expected Results 37
6.4 Long Term Impact 38
6.5 Instruments & Milestones 38

7 Conclusions **39**

Index of Figures

1 Cross-link interference in dynamic UL/DL. 19
2 The classification of literature 23
3 The technical models for sharing and the increase of different metrics with sharing options 27
4 Proposed market structure 36

1 Executive Summary

Work package (WP) 2 will be built around the objectives of network optimization and infrastructure sharing. More specifically, in WP2 investigations will be made on network resource and management optimization techniques for short-time-scale and low-load environments in dense wireless and software defined networks (SDN). On the infrastructure sharing front, WP2 will aim to derive the specifications for enabling long-time-scale and high-load virtulization and enabling sharing mobile infrastructure across providers and performance evaluation of different sharing policies.

Specific tasks for ESR3 mainly revolve around short-to-mid-term resource optimization and dense cellular network management with SDN. The envisioned goal of the short-to-mid-term resource optimization task is to provide fast algorithms for resource allocation in orthogonal frequency-division multiple access (OFDMA) based dense cellular networks with offloading capabilities. The problem in interference coordination in load-coupled systems will be addressed by examining power allocation and load balancing techniques based on multi-criteria optimization. The numerical tractability of proposed models will be verified with solution characterization and convex analysis. The task on dense cellular network management with SDN will address network management optimization with an emphasis on the aspects of user association and mobility. This task builds on coordination solutions found established the resource optimization task and derives policies for user association satisfying certain load balancing and fairness criteria. The wireless medium can be used opportunistically by exploiting time-variable interference conditions. Derived solutions can be applied through the SDN framework to alleviate potential overhead burdens.

The research on the feasibility of network sharing is performed by ESR4. The analysis is executed in two categories, i.e. business models and technologic models. In business models, the different sharing options and their applicability are investigated. It is observed that for short term sharing agreements, the economic gains of the network operators govern the decision of infrastructure sharing. On the other hand, for long term agreements, secondary metrics become dominant and shape the sharing decision. As a result of the feasibility analysis, the radio access network (RAN) sharing is concluded to be the most advantageous sharing option for long term sharing. Following the analysis on the network operator's intent on sharing, enquiries are focused on the applicability of the sharing in terms of technical necessities. First of all, during the survey on the literature, the strict dominance of spectrum sharing is noticed. This is mostly because the other type of RAN sharing approach, e.g. capacity sharing, is a type of national roaming and this action

is usually defined by service level agreements (SLAs) rather than technical analysis. As an outcome of the research, the application of a centralized controller is recognized to produce the highest spectral efficiency. Another conclusion is the fact that the usage of deviation parameter instead of using strict borders would increase the utility and the service quality. In the final part, a comparison between the SDN and network function virtualization (NFV) technologies is presented as well as an analysis of their advantages.

2 Introduction

The deployment of 4G networks proposes a more efficient and scalable network structure that can provide service to a large group of heterogeneous users with rather low cost. Thence, the stationing of small cell base stations is a novel approach. By definition, small cells are low-power wireless access points that operate in licensed spectrum. From a user's perspective, it proposes "5-bar" wireless connection, indoor and outdoor coverage and higher data rates, whereas in an operator's perspective, it proposes meeting the high demand with low cost [1]. However, ultra-densification of networks also brings many challenges such as high network complexity, energy efficiency problems, regulatory pressure, handling the heterogeneous traffic, interference, security, and network designing with plug-and-play device equipped [2]. The 4G technology also presents relatively low network speeds, e.g. up to a few gbps speed within 1 meter. Moreover, it is unable to handle the continuing exponential growth on network traffic and the increasing heterogeneity of network devices. Consequently, it is not capable of meeting the expectations of users and the network operators. The 5G technology proposes a better user experience with higher data rates and lower latency (<1ms latency). As a reaction to the constraints in 4G, one of the major expectation from 5G is providing ultra-high capacity to achieve connection of billions of things & people for low cost [3]. More generally, the offers of 5G can be listed as [2];

- Better spectral efficiency
- Better coverage
- Reduced latency (RAN latency < 1ms)
- Ultra high data rates and capacity
- Energy Savings and cost reduction

However, these promises of 5G technology also reveal bold challenges. One of the major challenge in 5G is the trade-off between cost efficiency and service utilization [2]. In an attempt to protect their customer profile, the network operators have to provide economic service provision [1]. In order to serve billions of customers, network operators are forced to buy large amounts of infrastructure resources. However, during the infrastructure purchase phase, the network operators are incapable of determining the future demands of the users. This unawareness pushes them to purchase infrastructure that would be lower or higher than the actual need. If the purchased capacity is lower than the actual demand then the network operator will be incapable of providing service to some of his customers. If the capacity is higher than the actual need then the network operator will be paying for the infrastructure that is not used.

Moreover, the user distribution of network operators complicates the enhancement of spectral and energy efficiency. Especially for suburban regions, the customer density is relatively lower than urban areas. This low density decreases the spectral utilization and energy efficiency. 5G and beyond applications demand a sizeable increment in the flexibility of the network topology as well as capability to be configured on the run. However, existing hardware dependent network models are incapable of providing such executions. Additionally, coverage area constraints are essential problem for network operators. Most of the regulations, oblige the network operators to cover regions that are not very profitable, i.e. low business potential areas [2]. These challenges decreases the business potential of the 5G networks and complicates the transition phase. Finally, the task of resource allocation is among the challenging aspects of completing an evolution from current cellular systems to future 5G. The saturation in the network resources as well as the continuing increase of network costs turn the network provision into a bad investment choice. As the economical and technological dynamics of network management change, the network operators begin bearing alternative revenue sources in conjunction with the methods to decrease their costs.

The succeeding section of the report is dedicated to resource allocation providing a survey of theoretical frameworks and specific methods some of which tackle unique challenges brought by different 5G technological factors. The study on resource allocation goes beyond the real of engineering and can become a topic of interest in social, philosophical, and political sciences. There may be different criteria for arriving at a sense of how good certain resource allocation schemes might be. Among the evaluation criteria of interest, efficiency and fairness are example factors. Works on fairness in economical, social, and other less engineering-centric dimensions may be of use in developing intuition or designing resource allocation strategies with requirements on satisfying certain constraints. A preliminary understanding experienced in the survey process has been that more room may exist in connection theoretical study and application on the topic of resource allocation.

The second part of the report focuses on the infrastructure sharing between network operators. From a network operator's perspective, sharing degrades the critically of the capacity purchasing process and lets the network operators to decrease their capital costs. Moreover, since the network operators can share their unused capacities, the operational costs will be shared between multiple operators. Besides, the network operators would be able to use the excess capacity of the other network operators in times of bursty traffic.

Another major proposal of sharing virtualized infrastructure is evolvability. Virtualized systems divide the physical infrastructure resources into virtual resource blocks and give a mapping between these resource blocks and the physical resources. While the network controller makes changes in one of the virtualized resource blocks, the network operator would be capable of continuing his service using other virtual blocks. Finally, infrastructure sharing decreases the risk of new market entry as it allows the new entrants to hire some capacity for certain amount of time. Through this time, the new operator would be able to have an exact insight of the customer responses. Moreover, by sharing the capital costs of rural coverage, the operators can reduce their risk of covering a low business potential area. In spite of these advantages, the number of existing sharing agreements are quite low. In this survey, the primary and secondary dynamics of sharing are investigated. The prior results from survey emphasize the gap between business models and the technical challenges.

The resource allocation section is organized primarily under a number of bubbles which are technological factors seen as promising candidates in 5G. With each of the subsection, if a specific technological component is concerned, a brief introduction about the technology is given outlining its potentials and challenges. Following introductory text on the 5G component, efforts are made to identify resource allocation publications that may provide methods that can be seen as responding to specific scenarios involving the different 5G components. A total of four technology components are surveyed in the resource allocation section including massive multiple-input and multiple-output (MIMO), millimetre wave, heterogeneous networks, and non-orthogonal multiple access. The last subsection deals with related theoretical studies on points including fairness, efficiency, complexity and trade-offs between different metrics. Also covered in the last subsection are papers on multi-band systems and flexible frame structure. On the analysis of infrastructure sharing, the investigations are concentrated on the viability of the infrastructure sharing. The applicability is determined under two major parts, i.e. sharing as a business model and technical aspects of sharing. In the first part, the economic and secondary motivations of different sharing options are presented. Following the viability of sharing as a business model, the technical models are investigated. In this part, the survey presents a paper based analysis on the shared entities in the literature. At the final part, the virtualization technology and its benefits are examined.

3 State of the Art in Resource Allocation

The problem of resource allocation is one studied in a wide range of subjects. Solutions to this problem have found use in a great many applications. This section of the report attempts to provide a survey of the current status with regard to resource allocation in wireless systems. As an implemented constraint on the scope and level of relavance in this survey, we focus on resource allocation strategies targeting enabling technologies of future 5G systems. The section is organized into parts that introduce the resource allocation literature under a number of technological factors that are generally envisioned to be part of 5G.

3.1 Massive MIMO

Massive MIMO technology has the potential to bridge the gap between an ever-increasing demand on throughput and connectivity and the limited spectrum resource. The central point of the massive MIMO idea is the creation of an excess in the number of service antennas. Performance improvements enabled by massive MIMO technology can be observed in a number of regards among them higher spectral efficiency, better coverage quality, and greater radiated power efficiency [4, 5, 6]. The aim of this subsection is to provide a description of the promises and challenges in massive MIMO.

A main upgrade in massive MIMO over earlier MIMO scenarios is that it allows scalability which has been a limiting factor for point-to-point MIMO and multi-user MIMO. In a point-to-point MIMO set-up, data transmission relies on links between M antennas at the base station and K antennas at the user device. Conditioned upon a number of assumptions, in the high signal to interference plus noise ratio (SINR) regime point-to-point MIMO may create a multiplexing gain characterized as $min(M, K)$ [4]. There exist several factors contributing to the infeasibility in realizing scalability beyond the scale of 8×8 for point-to-point MIMO. The requirement on a favourable propagation environment capable of supporting a high number of streams may be unrealistic. A precondition in the working mechanism of point-to-point MIMO necessitates the availability of channel knowledge at the receiver side. The acquisition of channel knowledge for the receiver depends on the transmission of pilot signals. As the number of antennas increases, the time and complexity inherent in the pilot-based channel learning process increases proportionally. Moreover, the promised multiplexing gain can not be fully exploited especially for users close to cell edges where the SINRs tend to be low. Sophisticated signal processing and electronics design for antennas also add to the trick in scaling point-to-point MIMO [4].

Similar with point-to-point MIMO, multi-user MIMO presents a difficult case for attempts at rendering scalability. A key scalability constraint stems from the need of implementing what's referred to as dirty paper coding/decoding [4]. The involved technique of dirty paper coding/decoding has a complexity with an exponential growth rate as the number of antennas increase in the system. The success of dirty paper coding/decoding is dependent on accurate channel state information and the acquisition of such information in a scaled up multi-user MIMO system translates into lengthy periods of channel estimation. Thus, the combined constraints of dirty paper coding/decoding and channel estimation paints a gloom prospect for designing a scalable multi-user MIMO system.

To break the scalability bottleneck, massive MIMO departs from the path of pursuing information theoretic capacity as is the case in multi-user MIMO and introduces a far greater number of service antennas over the number of user devices. Massive MIMO differs from point-to-point and multi-user MIMO in a few ways that reduce signal processing complexity and lessens the task of channel estimation. In a massive MIMO scenario, the availability of downlink channel state information is reduced to only the base station's side. The complexity in signal processing is lowered by opting to simple linear precoding and decoding techniques. As the number of base station antennas increase, near capacity performance can be approached by implementing linear precoding on the downlink in tandem with linear decoding on the uplink [4].

Although generally predicted to be a major enabling technology for future 5G systems, a number of challenges related to massive MIMO still require attention. As pointed out in [6], the challenges facing massive MIMO manifest mainly in the following aspects.

- Channel Reciprocity

  Operating in a time division duplex (TDD) scheme, it is believed that the propagation environment exhibits sound channel reciprocity. Mismatch to the reciprocity can arise from differences in the uplink and downlink hardware chains at the base station and user terminals. The problem of calibrating hardware chains to achieve sufficient level of reciprocity is understood to some degree and approaches exist in some cases that may help maintain the overall degree of reciprocity.

- Pilot Contamination

  The issue of pilot contamination stems from the inter-cell reuse of pilot signals. At the base station's side, pilot contamination compromises the estimation of channels between the base station and user devices in the cell. Furthermore, transmitted signals on the downlink contain inter-cellular interference as a result of pilot contamination. A study in [7] found that factors including uncorrelated noise and fast fading cease to affect performance when an unlimited number of antennas are available to the base station. The effect of pilot contamination, however, persists in the scenario with an unlimited number of base station antennas.

- Radio Propagation

  The functioning of massive MIMO requires certain conditions on the propagation environment the fulfilment of which creates what's referred to as *favourable propagation* [6].

The work in [8] represents a foray into resource allocation in a massive MIMO set-up. The main idea presented involves the utilization of massive MIMO for achieving a two-fold purpose of maximizing off-loading gain and minimizing interference. Specifically, an approach is developed in this paper that formulates the resource allocation problem as one unified network utility optimization problem. The formulated optimization problem exploits the predictability on instantaneous rates in a massive MIMO setting and can deliver improvement on the rates of cell-edge users.

3.2 Millimetre Wave

A major technological direction in 5G is the adoption of hitherto less-utilized spectrum resource in the millimetre wave (mm-wave) bands. The exploitation of millimetre wave and its integration with spectrum bands in current cellular networks offer both potentials and challenges. One underlying motivation for exploring possibilities in mm-wave bands is an urgent need to relieve the paradox between an ever-increasing demand on data and a finite amount of spectrum resource. This section of the report gives an account of the promises and research problems regarding mm-wave technology.

The envisioned mm-wave adoption chimes with another 5G technology direction which is cellular densification and the creation of a great number of small cells. The use of mm-wave in small cells provides a solution for delivering high data rate and better connectivity. Although achieving the potentials of mm-wave technology depends on sufficient capability in handling a number of challenges. The propagation characteristics of mm-wave requires careful consideration. As pointed out in [9], comparing with microwave mm-wave exhibits particularly high blockage sensitivity leading to large path loss exponents in non-line-of-sight situations. Efforts on analysing and characterising blockage sensitivity and the channel complexity are therefore called for in mm-wave research. The use of antenna arrays may be an important enabling factor for making use of mm-wave. A large antenna array may help mitigate the frequency dependence of path loss and render array gain that can remedies the effects of thermal noise bandwidth. Furthermore mm-wave presents unique challenges for designing hardware able to manage a high power consumption coming from mixed signal components including analog to digital converters and digital to analog converters.

Hybrid network architectures may provide assistance in managing the implementation difficulties due to mm-wave propagation characteristics. In a hybrid architecture scenario, mm-wave based small cells can take advantage of overlay layers operating on conventional technology. Signalling overheads on the mm-wave level can be partially assigned to overlay layers by exploiting the split between control and user planes. This functional control and user plane split may also allow mm-wave small cells to gather contex information and arrive at better informed resource allocation decisions. The process of cell discovery in mm-wave environment with multiple antennas and highly directional transmission may be a non-negligible issue. Context information based on user position can be obtained from the control plane and exploited to guide procedures in cell search [10].

3.3 Heterogeneous Networks

Heterogeneous network is seen in the literature as an evolving direction in terms of network architecture for future 5G. A description given in [11] characterizes a heterogeneous network as consisting of nodes that differ in transmission power and coverage size. Specifically, the nodes in the network can fall into the category of high power nodes (HPNs) with lager coverage area or low power nodes (LPNs) with smaller coverage area. Some of the envisioned improvements coming from the adoption of a heterogeneous network architecture include better capacity and coverage, and potentials in energy saving [11]. A motivation for this part of the report is to point out the main research issues revolving around heterogeneous networks.

The structure of a heterogeneous network exhibits a high level of flexibility which can be exploited to arrange a large number of access points with differing capabilities that utilize different parts of available spectrum resource [12]. The architectural and infrastructural idiosyncrasies of heterogeneous networks give rise to research problems targeting inter-nodal cooperation, interference cooperation, and load balancing. The significant gains in capacity and coverage of heterogeneous networks require effective means of inter-cellular and cross layer cooperation and coordination [11].

The task of spectrum allocation in a dense heterogeneous network model is considered in [13]. The paper argues that to tackle the problem of downlink spectrum allocation in heterogeneous networks novel proposals on network architecture and radio resource management are necessary. The paper outlines a network architecture model consisting of a supporting part that generates spectrum allocation patterns based on estimation and statistical analysis, and a collecting part that gathers data for feeding into the estimation and analysis process. Accompanying the proposed network architecture, a radio resource management model is developed and can function as a self-optimization entity in a self organizing network scenario. The presented radio resource management model taps into spectrum allocation patterns generated by the network architecture model and makes predictions on future spectrum allocation based on the pattern information.

Cell load optimization is the topic of interest examined in [14] with a system scenario of heterogeneous long term evolution (LTE) networks consisting of load-coupled cells. Load levels are used as indicators of resource consumption in cells. The coupling relation can be seen in the effects of a cell's load on the interference levels experienced by other cells. The objective of the paper is to formulate cell load setting strategies that maximize overall system performance measured in terms of system utility. Values for cell loads are set by solving an utility maximization problem. Furthermore, the paper described procedures for arriving at optimal demand allocation with given cell load. An argument is presented that suggests positive effects of cell load optimization on system utility.

3.4 Non-orthogonal Multiple Access

Proposed as a novel radio access concept non-orthogonal multiple access (NOMA) explores the superposition of multiple users in the power domain [15]. In its inception, the idea of NOMA is motivated by a number of evolving trends in future radio access. The transportation of multiplexing into an additional domain hitherto under-explored in conventional multiple access schemes promises significant performance gains in terms of spectral efficiency, capacity, and cell-edge user throughput [15, 16].

A core architectural component in the NOMA scheme is a downlink receiver capable of performing successive interference cancellation. The scheme of downlink NOMA with successive interference cancellation places computational requirements on user equipment (UE) as a consequence of the need to demodulate and decode signals for other user equipments in the cancellation process. The potentially hindering computational complexity requirement is however projected to be within the evolution on future UE computational and processing capabilities [15].

Gains in spectral efficiency relying singularly on NOMA fall short of the need to support future networks beyond 2020. Thus possibilities of combining NOMA with MIMO, multi-site cooperation are of interest for achieving further enhanced spectral efficiency. A proposed scheme jointly applying NOMA and multi-user MIMO is investigated in [15].

On the downlink, information is carried through multiple beams generated at the base station. Two methods of interference cancellation are applied by UEs. Successive interference cancellation handles intra-beam interference analogous to the process in a basic NOMA scenario. On the other hand, interference of an inter-beam nature are mitigated with interference rejection combing.

One example in the literature that studies the resource allocation problem in a NOMA scenario is presented in [17]. A cellular system with multiple sub-channels and multiple users is considered. An optimization based approach is proposed to derive a channel and power allocation strategy for downlink NOMA. The objective of the formulated optimization problem maximizes the sum rate utility under constraints of a total power budget, power budgets for individual users, and a maximum number of users a sub-channel can support. Downlink successive interference cancellation in NOMA is reflected in the rate expression which treats interference from users with higher channel gains as noise while discarding interference from users with lower channel gains as a result of the cancellation process that decodes signals intended for users with weaker channel conditions. Analysis on computational complexity suggests that the problem of power and channel allocation in NOMA is NP-hard. Suboptimal solutions to the formulated optimization problem are derived with an algorithm based on Lagrangian duality and dynamic programming. A novel point in [17] is its use of an optimization-theoretic based approach in a NOMA context. Similarly based on optimization theory, the work in [18] presents downlink power allocation formulations based on measures of fairness. Two different scenarios are considered: one where instantaneous channel state information (CSI) is available at the transmitter's side and other where the transmitter only has access to average CSI. Under each of the two CSI scenarios a chosen fairness measure is used to drive the power allocation formulation. In the case with instantaneous CSI availability, a max-min fairness measure that maximizes the minimum achievable rate is selected. While the availability is limited to only average CSI, the chosen fairness measure is adjusted to min-max with the objective of minimizing maximum outage probability. The system set-up of a single channel is less complicated comparing to the multi-carrier scenario in [17]. The lack of analysis on the choices of fairness under different CSI conditions or whether other means of fairness measures could be better suited however may be a direction for future work.

3.5 Related Theory and Miscellaneous

This part of the report focuses on materials that may provide theoretical understanding for the resource allocation problem or materials that treat resource allocation without specifying specific evolving trends in 5G and thus may be considered as miscellaneous or uncategorised.

An example of works that analyse and theorise the resource allocation problem is presented in [19]. In the article, a theoretical framework made up with a number of axioms is developed. Under the axiomatic framework, a family of fairness measures that satisfy the axioms can be constructed. It is shown that individual fairness measures in the literature including Jain's index, entropy, proportional fairness, and etc., are special cases within the family of fairness measures constructed under a set of axioms. The paper further expands the axiomatic framework by making modifications on an initial set of axioms to formulate an additional two sets of axioms. The two sets of modified axioms enables the theoretical framework respectively in capturing the asymmetry in resource allocation and the trade-off between fairness and efficiency.

Theoretical insights from the axiomitization approach developed in [19] may help provide a basis for quantified understanding on the concepts of fairness in resource allocation. As pointed out in the concluding part of the paper, assumptions on the resource being of a single type and infinitely divisible, independence of the generated fairness measures from the feasible region, and the ignoring of some intricacies in the allocation process leave much room for further investigation on the topic.

A comparatively less contemporary but nonetheless potentially relevant theoretical work with implications on resource allocation can be found in [20]. Summarized in a concise and pithy manner, a main objective of this paper is to instil an understanding of the concept of layering from the perspective of optimization decomposition. It presents the argument that the network architecture characterized by protocol stacks may be holistically analysed and systematically designed from the perspective of distributively solving some global optimization problems. The functionality of the overall network architecture is seen as solving some network utility maximization problem while the individual layers handle decomposed sub-problems with the inter-layer interfaces acting as coordinators of the sub-problems. The mathematical and theoretical framework of conceptualizing networks as optimizers and layering as decomposition may aid and facilitate analysing and designing of distributed cross-layer resource allocation approaches.

The work in [21] charts the terrain of inter-metric tradeoffs in two sets of metrics namely complexity-utility-delay and complexity-stability-delay. The tradeoffs can be invoked by the application of scheduling algorithms. The study points out some observations hidden in studies considering only one or two metrics. As an example of how connections between three metrics can manifest, the paper describes that a decrease in complexity from exponential to polynomial while preserving the same stability region could generally result in exponential growth in delay. It is also shown in the paper that traoffs concerning complexity-utility-delay is different from those in the case of complexity-stability-delay. The development of a theoretical framework based on a 3D tradeoff space model (treating each metric as a single dimension) helps with arriving at a quantified understanding of performance, complexity and trdeoffs concerning the two. The mathematical theory in this paper may assist in gaining intuitive insights on the costs of achieving simplicity in scheduling algorithms.

Visions of future 5G are likely to place emphasis on the diversity and heterogeneity characteristics. For example, a 5G vision presented in [22] predicts support for a diverse range of services and devices. Thus a theoretical framework analysing resource allocation scenarios with a heterogeneous multi-resource set-up will likely better reflect the deployment situation in 5G systems. Such a study on the tradeoffs between fairness and efficiency in multi-resource allocation problem can be found in [23]. The theoretical study presented thereof develops a framework for quantitatively characterising tradeoffs between fairness and efficiency in a multi-resource scheme. Two families of fairness measures that reflect the tradeoffs differently are described, one of which codenamed fairness on dominant shares (FDS) takes consideration on both available resource and the heterogeneity in user demands on multiple types of resource, the other coined as generalised fairness on jobs (GFJ) has interest only on the available resource and disregards the heterogeneity in user demands.

In [24] the topic of resource allocation strategies satisfying certain fairness measure is investigated with a multi-band multi-cell wireless system model. The work may provide some insight for resource allocation in 5G due to its consideration of a multi-band scheme. With a chosen fairness measure based on proportions fair criteria, the work identifies four distinct components namely client scheduling, channel access, client association, and channel selection that make up the global problem of achieving weighted proportional fair. A distributed protocol is developed that jointly considers the four components to satisfy goals on weighted proportional fair. Furthermore, a greedy policy with its basis on the proposed distributed protocol is introduced that demands less implementation complexity while achieving sub-optimal results. Simulations carried out in this work suggest that it may outperform certain algorithms used as comparison cases in terms of total weighted throughput.

The resource allocation scenario in 5G will likely comprise cellular network communication and device to device (D2D) communication with the latter one generally seen as an important evolving direction. The consideration of joint resource allocation combining cellular and D2D has been explored in [25]. The paper conceptualizes a flexible

D2D mechanism for functioning in an integrated environment supporting flexible UL/DL TDD scheduling. Centralized and decentralized approaches for solving the joint resource allocation problem formulated as weighted sum maximization are described. Positive findings are made mainly in two ways. First, it is stated the inclusion of D2D in centralized resource allocation algorithms can produce performance improvement. Second, the overall system performance improves with the enabling of resource reuse across D2D and cellular communications in the uplink and downlink. The paper also argues that the adoption of flexible TDD as compared to fixed TDD will likely render better overall system performance.

A flexible frame structure has been proposed in [26] for supporting a wide range of heterogeneous service requirements in 5G. Example service cases listed in the paper include mobile broadband communication, mission critical communication and massive machine communication. The paper proposes an approach for achieving user multiplexing on share spectrum with the service requirements per link in mind. The proposed multiplexing will have flexibility and incorporates dynamic transmission time interval (TTI). A motivating factor for designing a flexible frame structure is the need to support users with diverse service requirements and radio conditions. The flexible frame structure relies on an underlying concept of in-resource physical layer control signalling which monitors user data transmission. Results presented in the paper suggest a flexible frame structure may enable a tuning of the tradeoff between control signalling and desirable low round trip time.

4 ESR3 Research Proposal

4.1 Context

An important aspect in 5G visions is scalability that enables 5G to cater to a wide range of deployment scenarios and requirements. The scalability feature is reflected in 5G's support for a high number of use cases. The following list of example use cases may help explain the different requirements on 5G and the need for scalability and flexibility [27].

- Mobile broadband access;
- High user mobility;
- Massive internet of things;
- Ultra-reliable communications;
- Extreme real-time communications.

The diversity in requirements can be observed in a number of attributes including reliability, bandwidth, latency, capacity, energy, mobility, security, etc. With the guiding principal of enhancing scalability and flexibility, the ESR3 research topic aims to investigate resource allocation under scenarios of scalable resource. More specifically, the research interest of ESR3 lies in multi-band coexistence and flexible frame structure. In the ESR3 state of art survey section, some information on flexible frame structure has been provided. Thus, before introducing the research topic specifically, the next paragraph briefly describes multi-band coexistence.

Multi-band coexistence is a 5G vision that unifies a diverse range of spectrum resource. In a multi-band scenario, licensed spectrum, shared licensed spectrum, unlicensed spectrum are all put into use to provide respectively, exclusive use, shared exclusive use, and shared use [22]. The capabilities of different frequency bands are exploited to facilitate different service classes. Legacy bands are preserved while new frequency bands are introduced. For example, frequency bands above 6 GHz which includes mm-wave can act as both access and back haul. In a more general sense of multi-band coexistence, the coexistence of TDD and FDD bands is also a consideration. The convergence of TDD and FDD in LTE has been surveyed in [28], where the potentials and challenges of a TDD/FDD coexisting network are laid out.

4.2 Problem Description

4.2.1 Research Possibilities

The research proposal for ESR3 targets a dynamic-TDD-enabled system and studies resource allocation problems in this specific network scenario. The enhanced design freedom that comes with dynamic TDD manifests mainly with two enabling technologies, namely, scalable TTI and dynamic uplink (UL) and downlink (DL).

The concept of TTI is related to resource allocation and specifies a base unit for which physical resources are allocated. In conventional networks, a fixed TTI is used which is not designed to have to flexibility for the varying latency and quality of service (QoS) requirements stemming from a wide range of services. The round-trip time (RTT) targets in 5G is envisioned to be a magnitude lower than those in LTE systems. As analyzed in [26], TTI contributes to RTT and thus flexibility in TTI size is important in satisfying RTT requirements. In a flexible frame structure framework, dynamic TTI is connected with trade-offs between spectral efficiency, latency, and link reliability.

Another enabling technology of dynamic TDD is dynamic UL/DL. The research domain of UL/DL resource allocation in TDD covers a range of interesting problems. The

ESR3 research proposal focuses on UL/DL resource allocation in dynamic TDD and to this end we outline some possible specific topics in this realm. The concept of cell load is widely used to measures the average level of resource usage in a cell [33]. The phenomenon of the coupling of cell loads is reported in many studies. Scheduling problems and load coupling problems all exhibit interesting and under-exploited aspects in the research context of dynamic UL/DL. From the perspective of scheduling, a number of new factors come into play in dynamic UL/DL that are absent in conventional scheduling problems, such as new types of interference, constraints and performance objectives, practical considerations incorporating both UL and DL. The next few paragraphs explain a scenario of dynamic UL/DL TDD systems where the interest is in scheduling and UL/DL resource allocation. After which, we look at load coupling as another example of research possibilities in dynamic UL/DL.

The trend of network densification in 5G with the vision of deploying a large number of small cells has the effect of further deepening traffic imbalance between UL and DL. Unlike the traffic profile for voice calls which is symmetric in UL and DL directions, the traffic profile in 5G can exhibit significant asymmetry for UL and DL. The ability to dynamically vary the amount of assigned UL and DL resource blocks to match the asymmetry in UL/DL traffic is highly desirable.

The potential of enhanced flexibility however also gives rise to challenges in particular with regard to interference management. Asynchronous and dynamically changing UL/DL configurations in small cells can result in new types of cross-link interference, namely, UL-to-DL interference and DL-to-UL interference. Interference management is a key design concern with asynchronous UL/DL configuration across multiple cells. The detrimental effects of cross-link interference is more pronounced in networks with densely deployed small cells due to the increased risk of interference coupling experienced by neighbouring cells. Cross-link interference originating from asynchronous UL/DL configurations is illustrated in Figure 1. Cell edge users that experience poor channel conditions are especially susceptible to cross-link interference.

Interference management in dynamic TDD has been a subject of investigation in the literature. A number of dynamic UL/DL configuration strategies have been proposed with the consideration of interference mitigation. In terms of the level of synchronization in UL/DL configuration among different cells, the configuration approaches can be loosely classified into the following groups.

a) Fully synchronized UL/DL configuration.

All cells in the network apply the same UL/DL configuration.

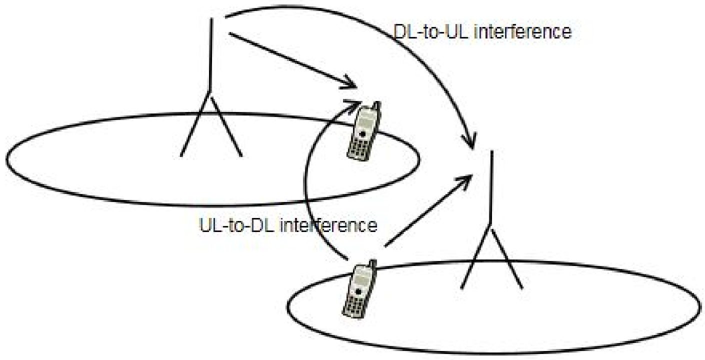

Figure 1: Cross-link interference in dynamic UL/DL.

b) Cluster-based approach.

Cells are organized into clusters. Within a cluster, the same UL/DL configuration is applied. Different clusters can adopt different UL/DL configurations.

c) Fully dynamic UL/DL configuration.

A cell adopts its UL/DL configuration based on the UL/DL traffic demand irrespective of the configuration in other cells.

The different levels of synchronization while permitting flexibility in adopting cell-specific UL/DL configurations to different extents, present varying challenges in interference management.

In the physical layer specification of LTE, TDD support includes flexibility in configuring UL/DL directions for subframes. The LTE-TDD specification supports a total of 7 static UL/DL patterns. The static UL/DL patterns provide limited configuration flexibility in terms of modulating the ratio of UL and DL subframes. However, the limited flexibility with static configuration patterns can not fully accommodate the time-varying UL/DL traffic asymmetry in 5G.

Interference is the main performance-limiting factor with the design of dynamic UL/DL configuration strategies. A number of studies have investigated the problem of dynamic UL/DL configuration in LTE-TDD. In [29], the authors surveyed potentials and challenges with dynamic UL/DL configuration. The performance evaluation therein demonstrates the advantages of dynamic TDD in delivering improved user experience and motivates further studies on interference mitigation. In [30], an analytical model for the distributions of DL SINR (signal to interference plus noise ratio) and UL SINR is developed with stochastic geometry for a network architecture consisting of phantom cells. A distributed algorithm for dynamic UL/DL configuration is proposed in [31] where

Table 1: Uplink-downlink configuration in LTE-TDD, where D represents a DL subframe, U represents an UL subframe, and S represents a special subframe.

Uplink-downlink configuration	Subframe number									
	0	1	2	3	4	5	6	7	8	9
0	D	S	U	U	U	D	S	U	U	U
1	D	S	U	U	D	D	S	U	U	D
2	D	S	U	D	D	D	S	U	D	D
3	D	S	U	U	U	D	D	D	D	D
4	D	S	U	U	D	D	D	D	D	D
5	D	S	U	D	D	D	D	D	D	D
6	D	S	U	U	U	D	S	U	U	D

base stations configure a ratio of UL/DL subframes with the objective of minimizing the overall UL/DL delay in its cell. Similar with a number of other works, the approach in [31] provides improved flexibility in UL/DL configuration however, because of limited design freedom that only permits ratio-based configuration, the full potential of dynamic UL/DL in adapting to varying traffic demands is under-exploited. The stuy in [32] is an example in the literature that jointly consideres UL/DL configuration and resource allocation. Subframes are allowed to be configured individually. Interference is mitigated with the optimization of resource allocation including channel allocation and power control. Although the formulation in [32] represents enhanced freedom in UL/DL configuration, the analysis that follows does not fully tap into the configuration flexibility as a result of its strictly suboptimal solution that essentially follows the same methodology with cluster-based approach.

In terms of scheduling in the domain of dynamic UL/DL TDD systems, research interests exist with respect to a range of problem constraints, performance metrics, the time granularity and scope of scheduling considered, etc. The ability to flexibly schedule UL/DL resources may provide the system with an edge in performance measures over conventional TDD systems. In terms of problem constraints, it is of interest to model practical constraints into the research problem. An example of potential practical constraints could be an upper limit on the number of UL/DL switches that are allowed in a frame. The inclusion of UL/DL specific practical constraints in scheduling problems is interesting since it may lead to extensions of existing research on scheduling. However, even without factoring in practical constraints, the problem of scheduling in the context of dynamic UL/DL is of theoretical interest. The perspective of the scheduling problem can be holistic, characterising network behaviour on a system level, or microscopic, looking into scheduling and resource allocation down to the level of subframes.

With respect to performance metrics, the space for consideration is also diverse. For example, we can consider energy consumption as an objective with the very practical connection to energy costs and the trend on green communications. However, implementing power control as a means of delivering energy optimization may present practical difficulties. To elaborate on this point, consider that in microscopic scheduling, power control may be specific to the level of subframes which could pose challenges in terms of implementation. The issue of defining performance metrics is important and the adoption of different performance metrics may place focus on different system aspects for the scheduling problem. Similarly, from the point of view of energy/resource consumption, the number of resource units measured in terms of subframes/frames is also interesting to be considered as an objective. A scheduling problem with the objective of minimizing total number of resource units used share similarities with the classical minimum time/length scheduling problems but in the context of dynamic UL/DL, the problem can lead to theoretical or practical extensions. Furthermore, the use of novel performance metrics in dynamic UL/DL scheduling may lead to new insights into the problem. For instance, the ratio of UL/DL resources allocated is a metric that can reflect the behaviour of scheduling algorithms and thus performance metrics that are based on UL/DL ratio formulations could potentially represent unique scheduling concerns in dynamic UL/DL TDD.

Load coupling is another example of research problems that can be considered and extended in the context of dynamic UL/DL TDD. In conventional cellular network load coupling analysis, the UL/DL scenario is not in the picture as the problem generally focuses on DL only. As a natural consequence, existing research on load coupling that only consideres DL may not fully extend to dynamic UL/DL systems. By introducing UL/DL specific considerations, the load coupling problem may represent an interesting type of research possibilities in the realm of dynamic UL/DL. The cellular network power and load coupling analysis in [33] provides a reference point for considering similar problems in dynamic TDD. The underlying factor that leads to coupling relations between cells is mutual interference. The situation of mutual interference is different in dynamic TDD as a result of new types of interference and dependencies on UL/DL resource allocations among others. The mutual interference is difficult to characterise without sufficient knowledge on the UL/DL configuration in different cells. The constraints in the load coupling problem also need to be adapted to reflect UL/DL configurations. More specifically, in conventional load coupling analysis, the load factor is assumed to be between zero and one for the system to be feasible. This feasibility constraint on load factors will need to incorporate load characterisations for both UL and DL in the case of dynamic TDD. Characterising the load coupling scenarios in dynamic TDD is thus a non-trivial research problem. Besides load coupling, incorporating considerations on power coupling in the analysis may also be considered for dynamic UL/DL systems. With modelling on load and power coupling, objectives such as energy consumption can be investigated.

4.2.2 Expected Results

As elaborated in the research possibilities section, the realm of dynamic UL/DL provides a playground for a number of interesting research problems. With the proposed research, existing work on scheduling and load coupling may be extended to dynamic UL/DL systems. For instance, in the case of scheduling, a more flexible scheduling approach exploiting dynamic UL/DL may outperform conventional TDD. An important result in treating dynamic UL/DL problems is the identification and modelling of the relations of UL/DL configurations and interference, transmit powers, and etc. Analysing and quantifying the effects of dynamic UL/DL on problems including scheduling and load coupling could provide interesting theoretical and practical results.

4.2.3 Envisioned Methodology

Several interesting problems are worth investigating in the context of dynamic UL/DL. With respect to scheduling and load coupling, the methodology will likely be based on optimization theory. The modelling and characterisation of constraints and objectives need to achieve a balance between reflecting practical considerations and theoretical insights. Special cases are of interest with optimization based methodology. By studying special cases, performance bounds may be established for general cases. Take the example of power control which in general is difficult to solve, but in some special cases in dynamic UL/DL, the problem may be solved with optimality. To summarize, modelling and optimization-based analysis are the main aspects in the envisioned methodology.

5 State of the Art in Infrastructure Sharing

Although the infrastructure sharing is not a new concept, there is surprisingly low number of papers exist in the literature. This part of the report provides a literature review on network sharing approaches and the related revenue models as well as technical approaches performed so far. As shown in Figure 2, the section is mainly divided into two categories, i.e. business models and technical models, that are investigating the literature in terms of economically appropriateness of network sharing and the enablers of sharing, respectively. Even though the introduction of 5G and beyond applications revived the studies on network sharing approaches, most of the studies in literature emphasizes the gap between the economic viability of the sharing options and the technical approaches to perform sharing. Despite the fact that no constraints were considered during the literature review, the tendency of the research trend limits the generality of the section by presenting its own constraints. More specifically, in technical models part, while the shareable resources are investigated, we observe a strict domination of the spectrum sharing in the literature. This is mostly because the capacity sharing approaches are SLA based business case solutions and there is not much analysis in a technical point of view.

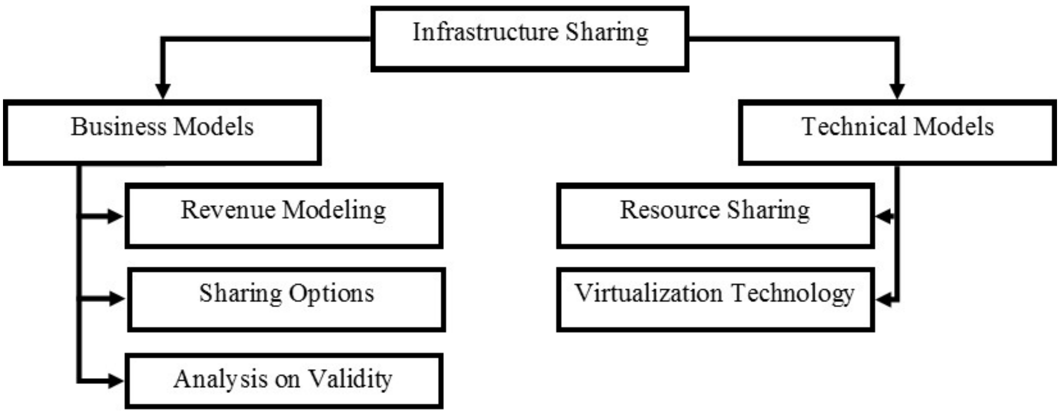

Figure 2: The classification of literature

5.1 Business Models

In this first part of the section, a general overview on the network operators' revenue model and the sharing options are presented. In order to determine the profitability of sharing approaches, a generic cost model is examined. Following the economic structure of the network operator, we investigated the mostly agreed sharing options and their connections with the network costs. Through the correlation between these two analyses, we observed that the maximum economic benefits can be achieved with the deepest sharing option. On the other hand, since the real life sharing agreements are only presenting low level superficial sharing agreements, we continued on searching the secondary dynamics that affects the network operator's decision. Based on the cross validation of network operator's concerns and the recent studies in the literature, we concluded on;

- In short term sharing, the applicability of the sharing option can be determined by the motivation of the sharing (e.g. cost function, capacity increase, providing coverage, customer profiling etc.)
- In long term sharing, secondary constraints become the governing limiters on sharing depth
- RAN sharing and mobile roaming are the most valid options for long term inter-operator sharing

5.1.1 Revenue Modelling

As the wireless resources become highly saturated, the option of increasing the revenue by providing service to more user is losing its validity. Since the revenue increase is not a valid option, the only way to increase the profit is decreasing the costs. In order to emphasize the governing dynamics of economical motivation, a precise cost function is necessary. With this objective in mind, the overall cost parameters are investigated in this part.

[35] presents a strong cost modelling argument as well as an analysis on the changes of cost functions with sharing options. The overall cost metrics are modelled under three main categories entitled as equipment cost, operation cost and transaction cost [35]. In the considered model, the equipment cost parameters are land lease cost, tower construction cost, access equipment cost, core equipment cost and spectrum license fee. The operation cost covers all the direct and indirect costs such as maintenance, planning, research and design (R&D) etc. Finally, transaction cost is the regulation and connection based costs. The assumption of network operators equally splitting the costs demands a more generic and more complex assumption of each network operator has equal power in the sharing and they all have similar sharing ratios.

An analysis over the governing cost functions are presented in [36]. The cost function of the wireless networks are divided into four main categories, i.e. operational phase costs, deployment phase costs, planning phase costs and migration phase costs. The specific sub-costs related to each cost segment are presented in Table 2. The capital

Table 2: The general structure of the cost function for wireless networks [36]

Wireless Network Cost Function			
Operational Phase	Deployment Phase	Planning Phase	Migration Phase
Licensing	Antenna Sites	R&D	Administrative
Energy	Core Equipment		Subscription
Housing			
Cooling			
Maintenance and Repair			
Operational Planning			
Marketing			

expenditure (CapEx) and operational expenditure (OpEx) distribution of each wireless networks varies with the urbanity of the network. In [36], it also measured that in urban scenarios 37% of the costs are OpEx related, whereas, 63% of the overall costs are CapEx related. However, for rural scenario, these parameters are 18% for OpEx and 82% for CapEx costs [36]. Finally, apart from the network operator's generic cost functions, the variation of cost parameters and their changes with the sharing option is presented in [37]. [37] models and investigates the CapEx and OpEx values of the resource pooling, more specifically, virtualizing resources from different network operators. In a classical wireless network topology the per user cost does not change with the number of network operators. However, in a virtualized system, the network resources may be distributed between different operators so the visible cost per operator decreases with the increase in number of network operators [37]. The analysis are conducted in two folds, i.e. the virtualization of a single base station and virtualized resources in multiple base stations. Although, it is observed that the overall cost function would decrease with sharing, they only focused on the power costs in OpEx parameter. Such an approach is valid yet power consumption covers a small fraction of OpEx. A key result of the study is emphasizing the fact that the sharing is only possible if the demand is low and the price for per unit resource is higher than the OpEx. More specifically, if the user demands are high, the network operator will need the whole resources and the sharing will not be possible.

5.1.2 Sharing Options

In this part, the different sharing options are going to be examined. The network sharing options can be studied under three major models [38] [39]

- *Business Model :* describing the parties involved in the sharing agreement, e.g.

network operators or infrastructure providers, and the agreements between these parties

- *Geographic Model :* identifies the physical footprint of individual network operators
- *Technology Model :* characterizes the technical solutions to enable sharing

The business models can be summarized as network operator- network operator sharing case and network operator infrastructure provider case. Secondly, the geographic sharing options can be investigated under four major cases, i.e. full split, unilateral shared region, common shared region and full sharing [38]. The full split sharing option is the case in which the network operators are covering the complementary areas and want to involve in a sharing to increase their coverage ratios. When one of the network operator does not have the necessary infrastructure to perform full coverage, the network operators can involve in a unilateral sharing agreement in which the operator can use others infrastructure to provide full coverage to his customers. In the common shared region model, the network operators that have their own infrastructures make an agreement to install new infrastructure to the region that is not covered by both of them. The key points in this sharing model are that the network operators have their own infrastructure in the region but not in some specific areas. They start building joint infrastructure to provide full coverage to the customers. The final geographic sharing model is the full sharing in which the network operators share all the base stations in the coverage area. In [38] and [39], the technology model of sharing is divided into passive and active sharing. Whereas, a more general and relatively more suitable approach is presented in [40]. [40] presents a classification over the shared entity of the infrastructure, more specifically they are dividing the technology model into site sharing, mast sharing, RAN sharing, network roaming, core network sharing. As presented in Figure 3, the network operators can achieve higher economic benefits with deeper sharing approaches while in long terms this approach can lead loss of business opportunities.

In the determination of sharing option, a critical definition is the maturity level of the market. In mature markets, the maintaining coverage phase is completed and the network operators mainly focus on the network capacity to differentiate their services, whereas, in developing markets the network operators try to cover as much different area as possible [39]. As a generic classification the mature wireless markets are named as capacity driven markets while the developing markets are entitled as coverage driven markets [39] [40].

Network operators have to decide on their cooperation level considering the possible gains and risks. As emphasized by [38], sharing network resources presents a trade-off between cost and business potential. More specifically, the maximum cost efficiency can

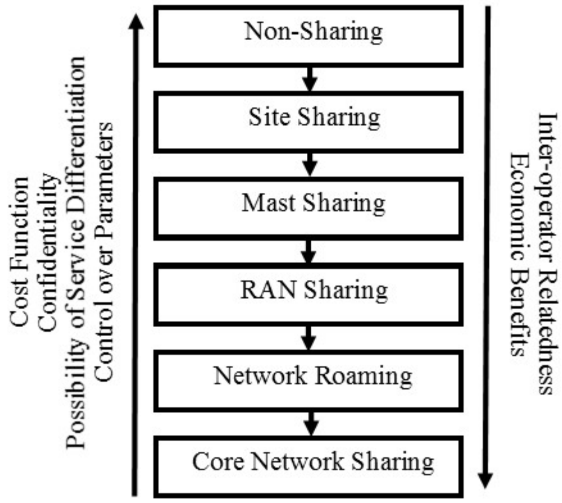

Figure 3: The technical models for sharing and the increase of different metrics with sharing options

be achieved when the network operators involve in full sharing agreements, however, this will limit their control over their infrastructure and eventually lead to massive losses in business potential [38]. [40] argues that the maturity of the market and the validity of differentiating the service by capacity or coverage is also a major bullet to determine. A quite general approach of network operators is to involve in stronger partnership and cooperation agreements for rural areas that usually have least business attraction but needs to be covered due to regulations [38][39]. Moreover, in a shared network, providing security of confidential business information or customer specific traffic data also becomes a challenge. [39] argues that outsourcing can be a better option than out tasking as a neutral provider would increase the confidentiality. It is generally approved that core network sharing may not be a feasible solution as any service or function that one operator implements can be replicated by the other as they have the same coverage areas and QoS [40]. [39] argues that the greatest economic benefits of sharing can be observed in coverage driven domains rather than capacity driven domains.

5.1.3 Analysis on Validity/Sustainability

In the previous part, the different sharing approaches in business model are defined. Despite the cost reduction objective proposes, most of the recent sharing agreements are not considering a full shared network topology. In this part, the secondary parameters that effect the network operator's decision are examined. Although, the cost reduction is the key parameter in the short time viability of sharing agreement, in long term agreements the degree of freedom in differentiability, long term expectations of network providers and the revenue split between operators also become visible as decision parameters.

The capability of differentiation is one of the most critical aspects in market structure. Each network operator has to differentiate his service in order to attract a certain group of customers. However, the biggest concern towards network sharing is the possibility that it would eventually lead to loss of competitive power as the network operators will have the same level of coverage ratio and capacity. Such a structure is infeasible both for capacity driven mature markets and coverage driven developing markets. In this respect, an evaluation over sharing viability is highly related to the possibility of differentiation. More specifically, the possibility of differentiation needs to be argued to analyse the sharing. In [41], a framework that allows the network operators to apply their own policies while sharing network resources are presented. The selected policies are designed by the combination of two different parameters, i.e. drop/admission probability and OpEx of small cells. The generic assumption is the network operator's capability to represent their policies in a standard mathematical form. The studied sharing option is the full sharing of small cells in the network (RAN sharing). In the considered framework, each network operator may use the small cell pool according to a sharing ratio constraint which is determined by the SLA. Controller tries to fulfil the sharing ratios of each network operator with the minimum number of small cell base stations. The key contribution of the framework is the fact that it allows the network operators to differentiate their services within certain limits.

Another aspect of network sharing is the long term expectations of the network operators such as migration to the individual networks [42]. The operators may need to migrate the shared network into individual networks due to the regulations or loss of economic gain. Therefore, the capability of migrating into an individual network with full re-use of invested infrastructure is critical [42]. In a generic sense, the issues to be considered for network migration are the motivation of cooperation between network operators, the coalition strategy and finally the business case evolution over time. A precisely defined SLA that respect those three matters could enable migration challenge.

Unlike most of the other studies, [43] focuses on the validity of outsourcing approaches by arguing that outsourcing non-differentiating services would increase the revenue of the network operators. As most of the critiques on network sharing mainly focus on the risks of sharing, neutrality of the infrastructure provider may be helpful for possible problems in sharing [43]. The involvement of infrastructure provider would be especially effective as the regulatory approval could be obtained more easily. Nevertheless, the current level of maturity of the service provider in the country is a decision metric for the sharing depth.

More specifically, the management success of the provided services, the coverage ratio, the technical expertise of the company determine the maturity level of the infrastructure provider [43].

Another important aspect to determine the network operator's intent to start or continue the network sharing is the price of the unit infrastructure that is used by other network operators. Extending the previous cost analysis, it is visible that the lower value of the price is constituted by the summation of overall OpEx expenses and a fraction of the CapEx of the used infrastructure. More specifically, the guest network operator has to compensate the expenses of the host network operator as well as a fraction of the infrastructure cost. If the price of infrastructure is higher than the summation of operational and capital costs, then the guest network operator's intent to hire infrastructure will be quite low, whereas if the price is too low it may not be worth to share the network [44].

Using game theoretic approach, [44] analyses the investments of network operators on infrastructure and the sharing decisions under different traffic rates for each network operator. The calculations show that the maximum incentive to invest is reached when the traffic distribution is counter-monotone. More importantly, it is observed that the validity of sharing possibility is determined according to how the revenue is split between providers. The total revenue obtained from serving the user is considered as a division unit. If the price is equal to the half of this revenue and assuming that the network operators have similar traffic patterns, then the network operator with the lower cost function will buy the necessary infrastructure while the second network operator stays out [44].

Regardless of the improvements in business model, there also exit many open points for sharing [45]. In active sharing case, the network operator has to have full control over the QoS metrics such as coverage or the handover parameters. However, the necessity of coalition agreement can result in technical difficulties and lower QoS. Moreover, even the simplest decision metrics such as location or the size of the mast need a common choice [45]. Furthermore, the handover procedure of shared resources is not trivial. If another network operator is already using some of the shareable RAN elements, the adjustments between different operator's constraints are not simple [45]. The network sharing approach would not be applicable if the shared network becomes over-crowded, so the growing speed of the network is another decision parameter

5.2 Technical Models

In the previous part, viability of the infrastructure sharing in terms of business model is investigated. In the economical modelling part, the network operator's cost variables are reviewed and after a general analysis, using a case study, some cost functions that are used in the literature are showed. After a brief definition of network sharing options, the key parameters to assess the validity of sharing are explored. In this section, the key assumption is the operator's agreement on the network sharing. More specifically, it is assumed that the network operators have an intent to share. Although, prevalent agreements cover both passive sharing approaches and RAN sharing, we are more concerned with RAN sharing options since the passive sharing agreements are more business case rather than technical models. In the first part, resource sharing algorithms in the literature are analysed, while in the second part of the section, the necessary virtualization techniques are presented. The general results obtained from this part can be summarized as;

- A centralized controller that optimizes the network resources will achieve higher efficiency in terms of both spectral efficiency and acceptance rate
- Instead of using fixed or full sharing, application of a deviation parameter that would provide an optimization between spectral efficiency and SLA agreement is the most general and most appropriate model for sharing rate
- Application of SDN has a critical value for sharing

5.2.1 Resource Sharing

In RAN sharing, with a most general approach the shareable resources can be separated into two subgroups, i.e. spectrum and the capacity. The capacity sharing approach is a typical kind of roaming agreement in which the network operators agree to provide services to each other's users. Although, this sharing option is the simplest and the mostly applied method, it is more likely a business model as the network operators directly provide services to a rate of second network operator's users. This rate is determined in SLA and the sharing procedure is performed in the access control rather resource sharing. However, in this part the main focus is on technical methods to enable resource sharing. Due to this technicality constraint, this study is mostly covering spectrum sharing options. However, it is necessary to emphasize that this constraint is a result of existing research trends rather than a limitation to our research.

In terms of spectrum sharing, two subcategories are considered; the studies mainly focused on scheduling the spectrum resources and the papers targeting the spectrum distribution between network operators. In the following studies, different approaches and controlling actions for spectrum division and sharing are covered. As a general abstract, the spectrum sharing process can be performed with a local management or a global controller. The local management framework makes sharing decisions with low level (local) knowledge, whereas, the application of global controller enables making spectrum splitting and sharing decisions based on a global view over the network state. As a secondary division parameter, the sharing rate is considered. The two extreme cases of spectrum sharing are full sharing and fixed sharing. In fixed sharing, the network operator's sharing rate is strictly constraint by the SLA agreement. In full sharing case, the network controller can optimize the spectrum utility without strict borders.

In [46], a scheduling algorithm for N operators that share the radio resource with RAN sharing is presented. In the proposed model, there are two types of schedulers; inter-operator scheduler and intra-operator scheduler. The inter-operator scheduler decides the flows of each operator and the intra-operator scheduler distributes these resources to its customers. By using this double scheduling algorithm, they are providing the capability of differentiation in sharing. Each network operator schedules its own traffic according to quality of service (QoS) and bit error rate (BER). The analysis are performed under two parts; fixed sharing and flexible sharing. A novel argument in [46] is the consideration of new entrants on the run time. In [47], single base station is shared amongst m network operators. In order to provide a mid-level solution between fixed sharing and full sharing, a deviation parameter from the agreed sharing ratio is proposed. The scheduler may change the sharing ratios with a delta value to increase the spectral efficiency. Proposed framework covers these extreme options as special cases. Finally, [48] propose a power allocation framework between different network operators that use the same base station but different antennas. In this study, tolerance value is applied to increase the flexibility of the optimization and solve the problem of rate fairness with a heuristic approach.

[49] presents a spectrum management framework to enable spectrum sharing between operators. More precisely, this study focuses on how the interference can be avoided while network operators have high spectral efficiency. The considered spectrum manager coordinates the communications between network operators. The authors emphasize that this spectrum manager does not necessarily have to be a separate entity. The spectrum manager handles the partitioning of the spectrum and distributing this fractions to the network operators. [49] considers two stage scheduling. At the first part the spectrum is assigned to different network operators, whereas, in the second part each network operator

assign spectrum fractions to its users. Each network operator sends its traffic demand information and its preferable sub-carrier information to the spectrum manager. The spectrum manager proposes an optimal schedule for the network operators according to the demands.

In [50], an analysis on heterogeneous network (hetnet) structure that contains long-term evolution (LTE), worldwide interoperability for microwave access (WiMAX) and wireless local area network (WLAN) is presented. The applied utility function represents the profit by using a secondary network to provide the service instead of the primary network. Using neural networks, the expected average transmission rate of each service is calculated. Every base station periodically calculates the unused resources with Neural Network. The proposed framework consist of three main parts. At first part, each base station calculates the expected traffic in the next time interval. Based on this parameter, the controller decides the sharing rates. When the real traffic parameters are revealed, the controller recalculates and evaluates the validity of the performed decisions. Finally, the controller sends feedback to each base station to renew their neural network structure. In [50] using neural networks to determine the traffic rate in the next time step and based on the real data, the base stations are updating the neural network.

Most generally, the spectrum sharing option can be investigated under three major categories, i.e. spectrum sharing in unlicensed band, one licensed band and multiple licensed band [51]. In the first case, there is no owner of the spectrum and the wireless devices try to optimize their communication efficiencies individually. In the second case, which is similar to the cognitive radio, the spectrum is licensed to one user and the rest of the users try to optimize their spectral efficiency without disturbing the primary user. The third and final case is similar to the inter-operator spectrum sharing case, where the operators may use the same spectrum without disturbing others.

In [51], the third sharing option is considered. In the considered structure, each base station is modelled with M/G/c/c queue with the arrival rate λ, service rate μ_i and c_i servers. In the considered framework, base stations are transmitting their states periodically. The framework assumes that each user can connect to every base station even though the base station belongs to another operator. The scheduler tries to connect each user to the base station that can produce the highest pay off function for the user. The pay-off functions are calculated using the server rate that user can reach from a base station and the achievable data rate of user by connecting to base station. The achievable data rate depends on the channel fading, the bandwidth of the resource blocks and transmission power. Using the game theory the framework decides which base station should be assigned to which user.

The [52] proposes a spectrum access framework for evaluation of spectrum sharing algorithms. It is argued that the differential load levels of two different network operator can provide an improved system capacity and hence higher spectral efficiency gain. Users' incapability of connecting on the initial network would cause longer setup times. The low level spectrum sharing is controlled by the resource requests from the operators The inter operator signalling is minimized by the application of a queue. The privilege is given to the users that are trying to use their main network operator and in case of no free channel, then the requests are queued for a short period of time (2-3 seconds) [52]. In the sharing model two different structures are investigated. If there exist a queue, the requests are queued for a short period of time and if the request cannot be processed during this time, it is sent to another network (a foreign network). If there is not a queue in the structure, depending on the SLAs, the request can be sent to another network or it can be dropped. The key result is that under heavy traffic conditions the queuing would not affect the satisfaction ratio as both network operators are only able to serve their own users.

Although, there are an absolute dominance of spectrum sharing algorithms in the literature, other types of shared entities and their feasibility needs to be investigated as well. In [53], the authors presents a comparative guideline on various sharing options in terms of their performance. More specifically, capacity sharing, spectrum sharing, virtual spectrum sharing and virtual physical resource block sharing are the considered as inter-operator sharing approaches. The considered network topology assumes an independent sharing entity that coordinates and manages the sharing process. In this study, the capacity sharing option which is a type of roaming sharing is observed to be the most efficient sharing approach in terms of load distribution and technological complexity [53]. Through the analysis, [53] presents three important argument about the sharing success;

- For sharing to be effective, the availability of surplus resources should be in close proximity to the demand. Considering this aspect, the inter-operator sharing options were effective only when there were large differences in resource utilization between the two networks
- Division of resources usually results in more overloaded base station sectors
- The non-collocated network topology has the highest complexity and the lowest sharing success rate

[54] is investigating the viability and the possible strategies of RAN sharing and the possible gains of network operators from sharing. The feasible sharing strategies of the network operators highly depends on the maturity of the wireless market. In a mature

market, as the coverage maintaining phase is completed, national roaming agreements or RAN sharing based mobile virtual network operator concept may decrease OpEx and increase the user satisfaction [54]. However, for such an agreement to be successful, availability of unused capacity is mandatory. In case of non-excess capacity or coverage driven network structure, passive or active network sharing approaches will create a decrease in OpEx and CapEx [54]. It is a well-known fact that as the network sharing depth increases the possible business risks emerge and a stronger motivation becomes a necessity. Although both the technical and the business aspects are challenging, the active sharing approaches proposes a 40% of cost reduction [54]. Nevertheless, this outcome highly depends on the sharing policies that are applied during the accessing the resources such as giving priorities to certain network operators or allowing access based on first come first served approach [54]. In order to achieve the highest possible utilization, a generic approach is to apply a centralized controller that may monitor the resources and receive and optimize the resource request. Depending to the considered network dynamics and the network operators, the optimization of resources may be conducted by the controller over the whole network or by base station locally [54]. During the numerical analysis, it is observed that global optimization is more robust and converges the optimal solution with lower levels of sharing rates. Through analysis, they show that the network sharing allows the single network operators to serve their users with higher capacity. However, the authors also emphasize the fact that regardless of the sharing approach that is applied, the possible usable capacity has limited with the physical resources [54]. The numerical results showed that if the network operators has Poisson traffic distribution, with a sharing ratio of 0.4, the network operators will be able to increase their acceptance ratio up to 99% [54].

5.2.2 Virtualization Technology

Next generation wireless systems require adaptive control both in terms of monitoring and decision making, infrastructure exchange on demand and achieving maximum spectral efficiency with an efficient energy management. However, the substantial network structure is not sufficient for 5G applications. More specifically, fixed hardware based architectures of existing cellular systems encounter with scalability and vendor specific device configuration problems [55]. In order to provide evolvable and efficient network management architecture, logically centralized network controllers are considered for the next generation wireless networks. The benefits of logically centralized network control can be listed as [55];

- Energy efficient and green infrastructure
- Better spectrum utilization
- Adaptability to the heterogeneous needs and traffic
- Smart control and load balancing between base stations and network operators

SDN and NFV are two candidates for 5G and beyond network applications. NFV proposes abstraction to network functions, e.g. routing decisions, and implement them in a centralized controller that is detached from the local devices [55]. The separated network functions can be implemented in software either locally or on cloud. The momentous advantage of applying NFV is to perform flexible and evolvable network management with functional updates that is not strictly constraint by the underlying physical infrastructure [55].

On the other hand, SDN proposes splitting the network controller and the forwarding plane equipments completely, i.e. controller and data plane respectively. By separating fixed management structure and implementing logically centralized structure, SDN offers real time control without loss of scalability. While the controller structure provides network policies, data plane elements, e.g. switches, performs as local control agents for simple actions [55]. Separating the control functions and infrastructure resources especially important to perform sharing actions. In virtualized systems, physical infrastructure resources can be abstracted and sliced into virtual resources holding certain corresponding functionalities, and shared by multiple parties through isolating each other [56].

The virtualization procedure done by mobile virtual network operator can be considered as the mapping process between virtual resources and physical resources [56]. With an efficient sharing approach, it is estimated that up to 40 percent of $60 billion used for OpEx and CapEx can be saved by operators over a five-year period. Moreover, the dedicated physical resources are incapable of respond to the changes in the technology. In a virtualized system the system upgrade is simpler as any upgrading and maintenance in one slice will not affect other running services.

6 ESR-4 Research Proposal

The research area of ESR4 covers the performance and enablers of infrastructure and resource sharing. On the tendency of this goal, as the first research problem, a preliminary market model for wireless resource sharing in a multi-operator network is proposed.

6.1 Context

Through the study of literature, it is observed that all the current works are mainly focusing on two major points, i.e. technical models to enable infrastructure sharing in multi-tenant networks and economic models for different sharing types and entities. However, a techno-economical model that would connect these two aspects and eventually lead to a dynamic context aware market structure is missing. Moreover, the dynamic nature of resource sharing requires the negotiations to be rapidly concluded. Therefore, in the first step of the research activities for ESR 4, a preliminary definition and analysis of the context aware resource market will be performed. Based upon the work in [57], a novel sharing model that would enable dynamic resource allocation and pricing framework will be proposed. The main novelty of the work is a techno-economic approach to infrastructure sharing in multi-operator networks. Unlike existing studies, the framework proposes an integrated negotiation and pricing model that is performed dynamically in short time scale yet governs the decisions in long time scale.

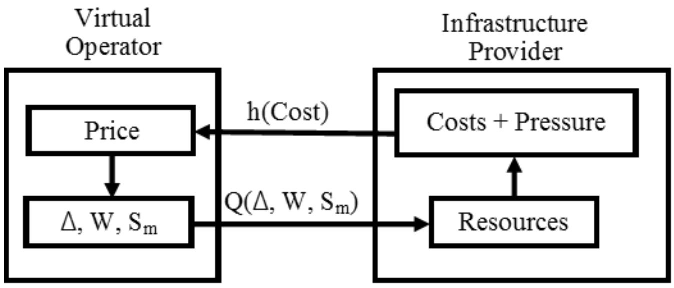

Figure 4: Proposed market structure

6.2 Problem Definition & Model

In the proposed framework, the defined resource market model contains two players, i.e. network operators and infrastructure provider. As a market structure, the infrastructure provider sets the price of the resources according to the costs and demands of network operators and the network operators reacts to the price by varying their demands (Figure

4). The pricing concept will be divided into two major parts that are the fixed costs such as OpEx and CapEx and an extra cost, called "pressure cost", that stimulates the effective usage of resources and also provides additional revenue for future capacity expansions. For the network operators whose QoS utility is equal to their expectations, the pressure cost will be zero. On the other hand, for the operators who have lower QoS than their requests, this price will be greater than zero and, over relatively long term, the collected revenue will be reinvested for capacity expansion. With this first problem, an analysis and validation of optimal resource allocations and price values will be performed.

6.2.1 Initial Assumptions

- All the resources can be dynamically allocated/negotiated which is the first simple model of a radio resource market.
- The prices are dynamically determined according to the agreed service parameters and the operator's gap between the target and the actual quality of service levels.
- All the resources are identical and have the same importance for the network operators.
- Infrastructure provider uses the collected revenue for capacity expansion.

6.3 Expected Results

In this first research problem a dynamic pricing and negotiation framework is defined. The main objective of this problem is to capture the behaviour of the market model with different cost parameters and different system metrics. Consequently the expected results of the first research problem can be listed as below.

- Analysis and validation of optimal resource allocations and price values
- Effects of pricing on the network operators' demand on owned resources and their flexibility
- Analysis of operators intent to share resources for different agreement durations, different values of pressure cost, and different number of operators with heterogeneous market shares.

6.4 Long Term Impact

In the long term, the outcomes of this study are expected to lead to the definition of a context-aware resource market for short term infrastructure sharing with trading and pricing framework. Furthermore, this study will be a first step towards analysis of the network capacity evolution over time under different policies for price negotiations and investments allocations.

6.5 Instruments & Milestones

In the beginning phase of the study, mathematical programming tool is going to applied to analyse the validity of the model. Furthermore, an optimization model will be defined and previously mentioned analysis will be performed on this model. The considered milestones for this research problem are as follows.

- Milestone 1: The first milestone is the design of the optimization problem. At the first simplified model will be design to maximize service satisfaction with cost parameters and price variables. Following the design and analysis of this simplified model, advanced models with different objective functions and constraints will be designed.
- Milestone 2: Implementation of the optimization problem is the second milestone of the research problem. The optimization problem will be modelled using a modelling language and a mathematical programming solver (MATLAB, GUROBI).
- Milestone 3: Following the implementation of the optimization problem, numerical analysis on synthetic instances will be performed. A set of instances of the considered problem will be defined as well as numerical analysis of the optimal solutions obtained with the solver.
- Milestone 4: In case of encountering any infeasibility or long computational time, an heuristic based algorithm will be designed and the analysis will be performed through that algorithm. In this context, the concept of "long time" is highly depends on the time windows of the network operators. At each time window the algorithm has to present an optimal solution. If the computational time is greater than the minimum time window, than the applied problem is considered to be infeasible and the heuristic algorithm will be used.

7 Conclusions

This report provides a technical survey on topics related to WP2 (reaction techniques). The report comprises introduction on the evolving trends in 5G, literature survey on resource allocation and review on infrastructure sharing.

The section dedicated to resource allocation gives information on a number of 5G technological components and surveys some existing literature that either develops theoretical understanding or proposes resource allocation strategies responding to different challenges stemming from 5G components. Some of the surveyed methods work to satisfy certain requirements such as fairness measures. The methods may be centralized or distributive. The main resources of concern usually are spectrum and power. Some novel proposals on multi-band and flexible frame structure are also given. The whole of this section works to demonstrate existing and ongoing research forming part of a knowledge base for tacking resource allocation challenges in 5G systems.

In the second part of the report, an analyse on the infrastructure sharing and the viability of sharing options are presented. In business model part, we analyse the economic and secondary metrics and present an overview of the cost analysis as well as sharing options. Over the study, it is observed that the network operator's costs can be categorized under operational phase costs, migration phase costs, deployment phase costs and finally planning phase costs (Table 2). Through the analysis, decrescent character of the costs with respect to the increasing sharing depth is appeared. In a more general sense, the necessity of investigating the network sharing actions under short term and long term agreements is noticed. In short term sharing, the cost function can be the decision metric of the sharing depth. On the other hand, for long term agreements, the secondary parameters become dominant on network operator's decision of cooperation. It is also concluded that, considering both cost function and secondary parameters, the RAN sharing or mobile roaming are the most valid options for long term sharing.

Following the business analysis of network operators, a review over the current sharing options is presented. In the first part of the analysis, the main focus was the resource sharing options. Through the enquiry in the literature, it is noticed that most of the literature is dominated by spectrum sharing options. This is mostly because of the fact that the capacity sharing approaches are a type of roaming agreement and the improvements can be applied in strictly defined SLAs rather than technological aspects. Based on the understanding from survey, it is valid to claim that the most agreeable option is to use a centralized controller to share the network resources. Moreover, in terms of sharing rate, the application of a deviation rate is the best feasible solution as it covers fixed sharing and full sharing cases as special cases. The investigation is concluded with the analysis of virtualization technology and its importance in sharing model. With this objective, software defined networks and network function virtualization are investigated. Through our analysis, the importance of SDN as an enabler technology is underlined.

References

[1] S. Patil, V. Patil, and P. Bhat, “A review on 5G technology,” *International Journal of Engineering and Innovative Technology (IJEIT)*, vol. 1, no. 1, pp. 26–30, 2012.

[2] P. Demestichas, A. Georgakopoulos, D. Karvounas, K. Tsagkaris, V. Stavroulaki, J. Lu, C. Xiong, and J. Yao, “5g on the horizon: Key challenges for the radio-access network,” *Vehicular Technology Magazine, IEEE*, vol. 8, no. 3, pp. 47–53, Sept 2013.

[3] Q. C. Li, H. Niu, A. T. Papathanassiou, and G. Wu, “5G network capacity: key elements and technologies,” *Vehicular Technology Magazine, IEEE*, vol. 9, no. 1, pp. 71–78, 2014.

[4] T. L. Marzetta, “Massive MIMO: An introduction,” *Bell Labs Technical Journal*, vol. 20, pp. 11–22, 2015.

[5] V. Jungnickel, K. Manolakis, W. Zirwas, B. Panzner, V. Braun, M. Lossow, M. Sternad, R. Apelfrojd, and T. Svensson, “The role of small cells, coordinated multipoint, and massive MIMO in 5G,” *IEEE Communications Magazine*, vol. 52, no. 5, pp. 44–51, May 2014.

[6] E. Larsson, O. Edfors, F. Tufvesson, and T. Marzetta, “Massive MIMO for next generation wireless systems,” *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186–195, February 2014.

[7] T. L. Marzetta, “Noncooperative cellular wireless with unlimited numbers of base station antennas,” *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590–3600, 2010.

[8] Q. Ye, O. Y. Bursalioglu, H. C. Papadopoulos, C. Caramanis, and J. G. Andrews, “User association and interference management in massive MIMO HetNets,” *CoRR*, vol. abs/1509.07594, 2015. [Online]. Available: http://arxiv.org/abs/1509.07594

[9] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *Communications Magazine, IEEE*, vol. 52, no. 2, pp. 74–80, 2014.

[10] A. Capone, I. Filippini, and V. Sciancalepore, "Context-based cell search in millimeter wave 5G networks," *arXiv preprint arXiv:1501.02223*, 2015.

[11] R. Hu and Y. Qian, "An energy efficient and spectrum efficient wireless heterogeneous network framework for 5G systems," *IEEE Communications Magazine*, vol. 52, no. 5, pp. 94–101, 2014.

[12] P. Demestichas, A. Georgakopoulos, D. Karvounas, K. Tsagkaris, V. Stavroulaki, J. Lu, C. Xiong, and J. Yao, "5G on the horizon: Key challenges for the radio-access network," *IEEE Vehicular Technology Magazine*, vol. 8, no. 3, pp. 47–53, 2013.

[13] T. R. Omar, A. E. Kamal, and J. M. Chang, "Downlink spectrum allocation in 5G HetNets," in *2014 International Wireless Communications and Mobile Computing Conference (IWCMC)*. IEEE, 2014, pp. 12–17.

[14] I. Siomina and D. Yuan, "On optimal load setting of load-coupled cells in heterogeneous LTE networks," in *2014 IEEE International Conference on Communications (ICC)*. IEEE, 2014, pp. 1254–1259.

[15] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi, "Non-orthogonal multiple access (NOMA) for cellular future radio access," in *2013 IEEE 77th Vehicular Technology Conference (VTC Spring)*. IEEE, 2013, pp. 1–5.

[16] Z. Ding, Y. Liu, J. Choi, Q. Sun, M. Elkashlan, C. I, and H. V. Poor, "Application of non-orthogonal multiple access in LTE and 5G networks," *CoRR*, vol. abs/1511.08610, 2015. [Online]. Available: http://arxiv.org/abs/1511.08610

[17] C. K. H. L. Lei, D. Yuan and S. Sun, "Joint optimization of power and channel allocation with non-orthogonal multiple access for 5G cellular systems," *Proceedings of IEEE GLOBECOM*, 2015.

[18] S. Timotheou and I. Krikidis, "Fairness for non-orthogonal multiple access in 5G systems," *CoRR*, vol. abs/1504.02300, 2015. [Online]. Available: http://arxiv.org/abs/1504.02300

[19] T. Lan, D. Kao, M. Chiang, and A. Sabharwal, "An axiomatic theory of fairness in network resource allocation," Tech. Rep., 2011. [Online]. Available: http://www.princeton.edu/~chiangm/fairness.pdf

[20] M. Chiang, S. H. Low, J. C. Doyle *et al.*, "Layering as optimization decomposition: A mathematical theory of network architectures," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 255–312, 2007.

[21] Y. Yi, A. Proutière, and M. Chiang, "Complexity in wireless scheduling: Impact and tradeoffs," in *Proceedings of the 9th ACM international symposium on Mobile ad hoc networking and computing*. ACM, 2008, pp. 33–42.

[22] P. John E. Smee, "5G design across services," Tech. Rep., 2015. [Online]. Available: http://johannesbergsummit.com/wp-content/uploads/sites/6/2014/11/Tuesday_3_John-Smee.pdf

[23] C. Joe-Wong, S. Sen, T. Lan, and M. Chiang, "Multiresource allocation: Fairness–efficiency tradeoffs in a unifying framework," *IEEE/ACM Transactions on Networking*, vol. 21, no. 6, pp. 1785–1798, 2013.

[24] I. Hou, P. Gupta *et al.*, "Distributed resource allocation for proportional fairness in multi-band wireless systems," in *2011 IEEE International Symposium on Information Theory Proceedings (ISIT)*. IEEE, 2011, pp. 1975–1979.

[25] V. Venkatasubramanian, F. Sanchez Moya, and K. Pawlak, "Centralized and decentralized multi-cell D2D resource allocation using flexible UL/DL TDD," in *2015 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, March 2015, pp. 305–310.

[26] K. Pedersen, F. Frederiksen, G. Berardinelli, and P. Mogensen, *A flexible frame structure for 5G wide area*, 9 2015.

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

[28] GTI, "TDD/FDD convergence white paper v1.1," 2015.

[29] Z. Shen, A. Khoryaev, E. Eriksson, and X. Pan, "Dynamic uplink-downlink configuration and interference management in TD-LTE," *IEEE, Communications Magazine*, vol. 50, no. 11, pp. 51–59, 2012.

[30] B. Yu, S. Mukherjee, H. Ishii, and L. Yang, "Dynamic TDD support in the LTE-B enhanced local area architecture," in *2012 IEEE, Globecom Workshops (GC Wkshps)*. IEEE, 2012, pp. 585–591.

[31] M. S. ElBamby, M. Bennis, W. Saad, and M. Latva-aho, "Dynamic uplink-downlink optimization in TDD-based small cell networks," in *2014 11th International Symposium on Wireless Communications Systems (ISWCS)*. IEEE, 2014, pp. 939–944.

[32] A. M. El-Hajj and Z. Dawy, "Joint dynamic switching point configuration and resource allocation in TDD-OFDMA networks: optimal formulation and suboptimal solution," *Transactions on Emerging Telecommunications Technologies*, 2014.

[33] C. K. Ho, D. Yuan, L. Lei, and S. Sun, "Power and load coupling in cellular networks for energy optimization," *IEEE Transactions on Wireless Communications*, vol. 14, no. 1, pp. 509–519, 2015.

[34] M. Chiang, "Nonconvex optimization of communication systems," *Advances in Mechanics and Mathematics*, vol. 3, pp. 137–196, 2008.

[35] Y.-K. Song, H. Zo, and S. Lee, "Analyzing the economic effect of mobile network sharing in Korea," *ETRI Journal*, vol. 34, no. 3, pp. 308–318, 2012.

[36] B. Lannoo, M. Kantor, L. Wosinska, K. Casier, J. Van Ooteghem, S. Verbrugge, J. Chen, K. Wajda, and M. Pickavet, "Economic analysis of future access network deployment and operation," in *11th International conference on Transparent Optical Networks (ICTON 2009)*. IEEE, 2009, pp. 70–73.

[37] M. M. Rahman, C. Despins, and S. Affes, "Analysis of CAPEX and OPEX benefits of wireless access virtualization," in *Communications Workshops (ICC), 2013 IEEE International Conference on*. IEEE, 2013, pp. 436–440.

[38] Investelecom, "Network infrastructure sharing: Operations and technology manager."

[39] T. Frisanco, P. Tafertshofer, P. Lurin, and R. Ang, "Infrastructure sharing for mobile network operators; from a deployment and operations view," in *Information Networking, 2008. ICOIN 2008. International Conference on*. IEEE, 2008, pp. 1–5.

[40] GSMA, "Mobile infrastructure sharing," 2012.

[41] I. Malanchini and M. Gruber, "How operators can differentiate through policies when sharing small cells," in *Vehicular Technology Conference (VTC Spring), 2015 IEEE 81st*. IEEE, 2015, pp. 1–5.

[42] A. Ericsson, "White paper-shared networks."

[43] L. L. Lurin, P., "Re-visiting the nancial benets of managed services and infrastructure sharing for Telcos."

[44] R. Berry, M. Honig, T. Nguyen, V. Subramanian, H. Zhou, and R. Vohra, "On the nature of revenue-sharing contracts to incentivize spectrum-sharing," in *INFOCOM, 2013 Proceedings IEEE*. IEEE, 2013, pp. 845–853.

[45] D.-E. Meddour, T. Rasheed, and Y. Gourhant, "On the role of infrastructure sharing for mobile network operators in emerging markets," *Computer Networks*, vol. 55, no. 7, pp. 1576–1591, 2011.

[46] S. A. AlQahtani, "Adaptive rate scheduling for 3G networks with shared resources using the generalized processor sharing performance model," *Computer Communications*, vol. 31, no. 1, pp. 103 – 111, 2008. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0140366407004173

[47] I. Malanchini, S. Valentin, and O. Aydin, "An analysis of generalized resource sharing for multiple operators in cellular networks," in *Personal, Indoor, and Mobile Radio Communication (PIMRC), 2014 IEEE 25th Annual International Symposium on*, Sept 2014, pp. 1157–1162.

[48] O. Aydin, D. Aziz, and E. Jorswieck, "Radio resource sharing among operators through MIMO based spatial multiplexing in 5G systems," in *Globecom Workshops (GC Wkshps), 2014*, Dec 2014, pp. 1063–1068.

[49] J. Luo, J. Eichinger, Z. Zhao, and E. Schulz, "Multi-carrier waveform based flexible inter-operator spectrum sharing for 5G systems," in *Dynamic Spectrum Access Networks (DYSPAN), 2014 IEEE International Symposium on*, April 2014, pp. 449–457.

[50] J. Mei, P. han Ho, H. Ji, Y. Li, and X. Y. Wang, "Service-aware optimal spectrum sharing algorithm in heterogeneous wireless networks," in *Communications (ICC), 2010 IEEE International Conference on*, May 2010, pp. 1–5.

[51] Y.-T. Lin, H. Tembine, and K.-C. Chen, "Inter-operator spectrum sharing in future cellular systems," in *Global Communications Conference (GLOBECOM), 2012 IEEE*, Dec 2012, pp. 2597–2602.

[52] G. Salami and R. Tafazolli, "Interoperator dynamic spectrum sharing (analysis, costs and implications)," *International Journal of Computer Networks (IJCN)*, vol. 2, pp. 47–61, 2010.

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

[54] A. P. Avramova and V. B. Iversen, “Radio access sharing strategies for multiple operators in cellular networks,” in *Communication Workshop (ICCW), 2015 IEEE International Conference on*. IEEE, 2015, pp. 1113–1118.

[55] I. F. Akyildiz, S.-C. Lin, and P. Wang, “Wireless software-defined networks (W-SDNs) and network function virtualization (NFV) for 5G cellular systems: An overview and qualitative evaluation,” *Computer Networks*, vol. 93, pp. 66–79, 2015.

[56] C. Liang, F. Yu, and X. Zhang, “Information-centric network function virtualization over 5G mobile wireless networks,” *Network, IEEE*, vol. 29, no. 3, pp. 68–74, May 2015.

[57] I. Malanchini, S. Valentin, and O. Aydin, “Wireless resource sharing for multiple operators: Generalization, fairness, and the value of prediction,” *Computer Networks*, vol. 100, pp. 110 – 123, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1389128616300366