Coláiste na Tríonóide, Baile Átha Cliath
Trinity College Dublin
Ollscoil Átha Cliath | The University of Dublin

PhD Thesis

Radio Access Network and Spectrum Sharing in Future Mobile Networks

Author:
Jacek Kibiłda

Supervisor:
Prof. Luiz A. DaSilva

14th September 2016

Declaration

I declare that this thesis has not been submitted as an exercise for a degree at this or any other university and is entirely my own work.

I agree to deposit this thesis in the University's open access institutional repository or allow the Library to do so on my behalf, subject to Irish Copyright Legislation and Trinity College Library conditions of use and acknowledgement.

Signed:

Jacek Kibiłda, 14th September 2016

Summary

The goal of this work is to resolve some of the resource management problems arising as a result of the changes to mobile network ownership and control model brought about by some of the recent trends in commercial mobile networks, such as the inter-operator network sharing, the formation of mobile virtual network operators, and the increasing role of over-the-top service providers. The problems we tackle can be stated in the form of the following research questions:

- What statistical models are representative to inter-operator resource sharing in spatially distributed mobile networks?
- What is the efficiency of inter-operator resource sharing in spatially distributed mobile networks?
- How network resource sharing affects utilities obtained by traditional mobile network operators and over-the-top service providers?

To address these questions we apply a range of analytical tools, including stochastic geometry, optimization and game theory, and, whenever possible, we populate our models with real mobile network data.

First, we address the problem of statistical representation of inter-operator mobile network deployment data. For that purpose we apply the tools from spatial statistical analysis and stochastic geometry. We find that there is a high degree of spatial dependence between the infrastructure deployment made by multiple operators. We see spatial dependence of similar strength occurring over and over again for similar areas in different countries. In consequence, we propose a stochastic geometry model that captures this dependence and is representative of the deployments made by multiple operators in different countries.

Then, we look into the fundamental trade-offs and efficiencies involved in inter-operator resource sharing over spatially distributed mobile networks, represented using stochastic geometry models. We observe that each type of sharing has its own distinctive trait. While spectrum sharing alone provides an increase in the data rate at the expense of coverage, radio access network sharing results in improved coverage. When the two sharing types are combined, we observe significant increase in the average user data rate and minor improvements to the network coverage as compared to the non-sharing scenario. Consequently, combination of the two approaches does not linearly scale the respective coverage and data rate gains. Yet, as we show, even a simple mechanism of coordination in spectrum sharing allows operators to reduce the impact of this trade-off and improve their coverage without hampering the user data rates. Additionally, we observe that the way in which spectrum is being shared makes a difference to both the coverage and data rate performance. In particular, when coverage is the main concern, it is more beneficial for operators to allow for their transmitters to dynamically select among the pooled spectrum bands, rather than simply bond the available channels.

Thirdly, we study the utilities obtained by traditional mobile network operators and over-the-top service providers from resource sharing in mobile networks. The first part of this work involves modelling the ability of over-the-top service providers to dynamically select a set of infrastructure that meets the requirements of the service. We find that considering network infrastructure of multiple mobile operators as a collective may lead to building up coverage provisioning networks which require just a fraction of all the available transmitters. We find also that low density areas benefit considerably from infrastructure sharing, which may make the case for infrastructure deployments in currently under-served rural areas. In the second part we consider new infrastructure deployments made in response to an increase in the mobile demand. We construct a model of interaction between over-the-top service providers and mobile network operators engaging in a new deployment. We show that, under some conditions, cooperation is a desirable outcome to this model. These conditions include the price of the radio access infrastructure and, to some extent, also the spatial clustering in mobile demand.

All of this builds groundwork for the vision of future wireless mobile networks, which we term Networks without Borders. This vision is based on a marketplace of virtual network operators that construct networks from a pool of shared resources, such as base stations, spectrum, backhaul, and core network servers. These resources are sourced from traditional industry players as well as from individuals. As we show, the proposed vision results in a substantially different mobile communications value-chain, and leads to a number of socio-political challenges.

Acknowledgements

One of the many things I have learnt during this PhD project is that you can improve your final draft of an article/thesis chapter by removing the often pretentious and repetitive opening paragraph. Well, after removing the opening paragraph along with the blather it contained, let me tribute the people whom I shared my PhD experience with.

First, this research would not have happened if it was not for the support of Prof. Luiz DaSilva – a mentor, a teacher, an authority, a great researcher and a friend. You are a great support and I am delighted to have had the chance to work with you as my supervisor. I owe my research personality to you. Luiz, "you are the real MVP".

I have to also recognize the role that CTVR, later transformed into CONNECT, played in keeping me inspired and providing me with an intellectually stimulating environment. Being at CTVR made the PhD experience remarkably rewarding and allowed me to grow both as a researcher and a human being. There is a number of CTVR/CONNECT people that have contributed to this, here I will name just a few (no particular order intended) but there were numerous other people during my time here which in one way or another have also contributed to the "CTVR spirit". First and foremost, I would like to give credits to Prof. Linda Doyle for organizing such a fantastic place and being an inspiring leader. I would like to credit our administrative staff Shirley, Catherine, Jennifer, Leylah, Monica, Andrew, Frank, Mark, Ultan, and Des for helping me out with whatever administrative issues I may have had during my time here. I would also like to credit all the (more) senior members of CTVR for helpful comments, coffee clubs, all the ideas we have explored together and companionship. Thanks Tim, Nicola, Marco, Paul, Irene, Seamas, Johann, Nick, Francesco, Giorgios, Maice, Andrea, Yi, Yong, Aleksandra, Avishek, Christian, Dennis (not a doctor!), Ismael, Kostas and Joao Paulo. Special thanks goes to fellow PhD students with whom I shared this troublesome journey. It has been a pleasure to share one office space with you, Danny, Paolo, Emanuele, Justin, Jonathan, Jasmina, Elma, Jernej, Boris, Giannis, Francisco, Pedro, Carlo, Arman, Parna, Marcelo, Nima, Merim, Conor, Neal, Liam, Joao, Andrei, Cezary, Quentin, and Diarmuid. I would also like to give special acknowledgements to Jessica, for exposing me to the world of expression, and to Eamonn, for being a good friend and an endless source of random facts.

Obviously, all of this would not have happened if it had not been for the support of my family. I would especially like to thank my parents Ewa and Bogdan, my brother Maciej, and my grandparents for their support and care. I know you are proud of me and I would like to dedicate this thesis to you.

Alongside my family, I would like to acknowledge my good friends who supported me during the entire duration of this PhD project: Pstrągi, Pat, Pyza, Wasyl, Tompor, Marcin, Kaśka, Plaża, Padraig, Wiola and Bartek – thank you for having been there for me when I needed it the most.

Finally, I would like to credit all the people who provided me with their invaluable comments on

various stages of writing of this thesis: Prof. Luiz DaSilva, my reviewers Prof. Gustavo de Veciana and Prof. Marco Ruffini, as well as my colleagues Jonathan van de Belt, Paolo di Francesco, Emanuele di Pascale, Carlo Galiotto, Diarmuid Collins, and Elma Avdic.

Contents

List of Figures

List of Tables

1 Introduction

Introduction

"Society has never allowed that which is necessary to existence[1] *to be controlled by private interest."*

- Theodore Vail, *"Telephone: the First Hundred Years" by John Brooks*

In many developed regions around the world commercial mobile communications systems have achieved over 100% penetration with a coverage larger than 90% of the region's area. Despite a failure of profit-based economy to deliver wireless access to rural areas of the developing world, this is an enormous success and the global statistics such as the ICT Facts & Figures for 2015 published by the International Telecommunications Union (ITU) clearly confirm that[2]. The story of the development of mobile communications systems (see [1–3]) attests that this success would not have been made possible if it were not for the support of powerful institutional agencies which oversaw equipment and protocol standardization, consistent addressing, radio spectrum allocation, inter-operation with wireline services, as well as controlled the structure of the local deployment and management of the network. This institutional support has enabled major telecommunications operators, which were at the time public switched telephony network (PSTN) companies, to incentivize development and the large-scale deployment of mobile communications systems. A PSTN served mainly as a telecommunications service provider, offering circuit-switched telephony as its core product with mobile telephony being only a value-added proposition to its wireline counterpart [2]. The earliest successful launch of a public mobile communications system, the analogue Nordic Mobile Telephony (NMT) in Finland, Denmark, Sweden and Norway, is an example of that.

After the era of PSTN conglomerates providing telecommunications services to the public, came the era of mobile network operators (MNOs), as we know them today. MNOs arrived to the scene with the birth of a digital successor to the first-generation analogue mobile communications systems – the Groupe Special Mondiale (GSM), later changed to Global System for Mobile Communications. The development and introduction of GSM was supported by the European Communities (predecessor to the European Union), which saw the digital standard of mobile communications as yet another way of creating a common European economic market. In 1987 a number of European countries signed a memorandum to implement a pan-European digital mobile communications service with international roaming capabilities. The memorandum stated that "operational networks shall be procured in each of the countries by the network operators" [4]. A network operator in this context had an exclusive license to some spectrum band and owned an extensive wireless and backbone infrastructure. This allowed it to inter-connect its subscribers within the same network

[1] In a speech from 1915 the At&T's president refers to the telephone service.

[2] The statistics were compiled and published in correspondence to the deadline for achievements of the UN Millennium Development Goals.

or with other subscribers/services within another network, for example, another mobile network or the Internet. The mobile network owned by an MNO had a very particular architecture which reinforced the model of single-ownership of the infrastructure, spectrum, and subscriber base. This architecture is principally comprised of the radio access network (RAN) and the core network. The RAN is composed of base stations and user devices, while the core network consists of a number of switches and routers, as well as computer servers that provide authentication, billing or user location tracking functions.

This model of network ownership and control fared well when new infrastructure had to be laid out and mobile penetration was still relatively low. In the early 2000s the mobile communications market was in a rapid expansion phase fostered by economic growth, with high-tech industry stocks booming and wide availability of credits from deregulations in the financial market. However, the situation started to change with the economic downturn of 2001. MNOs, previously ecstatically bidding up the license prices, now had to turn their attention to reducing their capital and operational costs. As it turned out, one of the most important, and the easiest to apply, cost reduction measures was to share ownership and control over parts of the network with other MNOs. In the main, inter-operator sharing can be applied to various elements of the network infrastructure, such as masts, antennas, base stations or backhaul, as well as radio spectrum. A natural consequence of opening up wireless network infrastructure to other MNOs was to allow new actors that do not have a license to spectrum to act as wireless service providers. This very idea lies at the core of what is a Mobile Virtual Network Operator (MVNO) as we know it today. Some MVNOs were created to utilize extra capacity available in the networks, some to reach new customer market segments, and others emerged as a result of anti-trust and maximum usage fee policies imposed by the communications regulators to boost competition and improve the availability of wireless services [5].

The emergence of network sharing and the MVNOs allowed for a break-down of ownership and control in mobile networks. We envision this break-down to continue along with the increase in popularity of mobile data communications, which is associated with a phenomenon known as the cost-demand offset [6]. In plain words, the demand for capacity grows faster than the per unit cost of capacity reduces. Indeed, upon observing fast-paced growth in mobile data demand (see, for example, [7]), we have started to set ever-more ambitious targets as to the expected capacity of mobile networks [8]. Yet revenues from mobile data services are decoupled from the traffic, hence also from operational and capital costs (as pointed out in [9]), which have been steadily growing over the years. In terms of some recently available numbers: the world-wide mobile capital costs have been predicted to grow from around $216 billion in 2014 to $224 billion by 2017, as reported during the Mobile World Congress (MWC) in 2015, while the annual fees for radio spectrum usage are also subject to a major price revision; for example, in the UK the Office for Communications (Ofcom) has put out a proposal that will result in an increase in annual spectrum fees from £64.4 million to £199.6 million for all domestic mobile operators.

One of the reasons why MNOs cannot capitalize on the increase in the mobile data traffic is the fact that end-user activity with contemporary mobile devices relies predominantly on over-the-top services. From that perspective mobile access is treated simply as a bit-pipe, and service differentiation happens only at the application layer. This situation is further complicated by the so called "net neutrality" policy [6], which, in many developed mobile markets world-wide, forbids MNOs to apply any service-level traffic engineering. The end effect is that it is the over-the-top service providers (OTTs), such as Google or Netflix, that collect revenue that would otherwise be

generated from the traditional set of mobile services, such as short-messaging service (SMS) or multimedia messaging service (MMS), offered by the MNO.

We believe that future mobile networks will be characterized by resource sharing and virtualization, in an architecture that we call Networks without Borders (NwoB) (see [10, 11]). NwoB is based on the idea that, from the perspective of OTTs, all network resources will be shared and available to the creation of virtual and dynamic instances of networks tailored to the over-the-top service that will be offered through them. Our vision is based on networks that will be constructed from a pool of shared resources such as base stations, spectrum, core network components, cloud resources, processing capabilities, etc. These resources will be sourced from traditional industry players as well as crowd-sourced from individuals, therefore changing the definition of an infrastructure provider as represented by mobile network operators. A new breed of service providers will aggregate and control slices of these resources to enable the creation of customized virtual networks instantiated to deliver specialized services to end-users. By leveraging virtualization and relying on dynamic and heterogeneous ownership of infrastructure and spectrum, this new view of mobile networks brings the sharing economy even deeper into the mobile networks.

The remainder of this chapter is organized as follows. First, we briefly go through recent trends that motivate the resource management problems tackled in this thesis and, indeed, the entire vision of future networks we propose herein. Subsequently, we discuss the scope, outline, and contributions made in this thesis, to conclude with a section on dissemination and out-reach made during this PhD project.

1.1 The path towards Network without Borders

The traditional view of a mobile network as a system that simply delivers voice and data services to a geographical area no longer reflects the reality of mobile networks. In recent years we have witnessed various trends in commercial mobile networks which pave the way towards service-driven networks and our vision of Networks without Borders. These trends include the inter-operator network sharing, the formation of MVNOs, the increasing role of OTTs and the virtualization of the wireless access. In Fig. 1.1 we illustrate how these trends evolve towards our vision. Starting from the left we have MNOs (infrastructure providers) which exclusively own and operate resources. Then, these operators share their networks (typically based on bilateral agreements), and allow other entities such as MVNOs to utilize their resources under their own brand to their own subscribers (again, based on one-to-one agreements). Nowadays we see a new breed of MVNOs which has evolved towards entities, such as Google Project Fi, which span multiple networks and multiple infrastructure providers, including also WiFi networks. In order to arrive at the NwoB vision we need one more step where infrastructure providers pool their resources and virtual operators create dynamic networks that consist of resources from various corporate or private infrastructure providers.

Our proposal of the Networks without Borders vision came three years before Google's expansion to mobile service provisioning. In Fig. 1.2 we provide a timeline of events which we deem important to the emergence of our proposal, intertwined with the evolution of mobile communications standards:

- the first commercial NMT system deployment in Scandinavia;
- the first commercial GSM deployment in Finland;

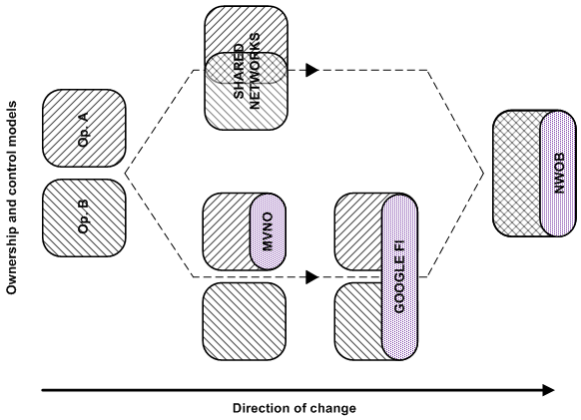


Figure 1.1: Network ownership and control model evolution towards the vision of Networks without Borders.

- the Telecommunications Act in the US;
- formation of the first MVNO in the UK;
- the first commercial Universal Mobile Telecommunications System (UMTS) deployment in Japan;
- the first passive sharing agreements in the US/Sweden;
- the first active sharing agreements in the Netherlands/US;
- the first[3] national roaming agreement in Germany;
- the first commercial Long Term Evolution (LTE) deployment in Sweden;
- launch of Facebook Zero;
- launch of Google Free Zone;
- Networks without Borders;
- launch of Google Project Fi.

This thesis is about resolving some of the resource management problems arising as a result of the changes to mobile network ownership and control model brought about by some of the highlighted trends, such as the inter-operator network sharing or the emergence of OTTs. We also see these changes converging towards a specific vision of future wireless networks that we present in the conclusion of this thesis.

1.2 Scope

As part of this PhD project we approach a number of mobile network resource management problems, using a variety of analytical and simulation-based techniques, with the main objective being to build a comprehensive and fundamental view of resource sharing in future wireless mobile networks. The results we obtain serve us as the basis for evaluation of the potential and limitations of our vision for future wireless networks that relies on the idea of shared access to virtualized radio access infrastructure and spectrum resources, which we term Networks without Borders. In general, there are three research questions that we address:

[3]It is worth noting that, in the end, the agreement was blocked by the European Commission, which postponed the introduction of national roaming in Europe for a number of years, as we will explain in Chapter 3.

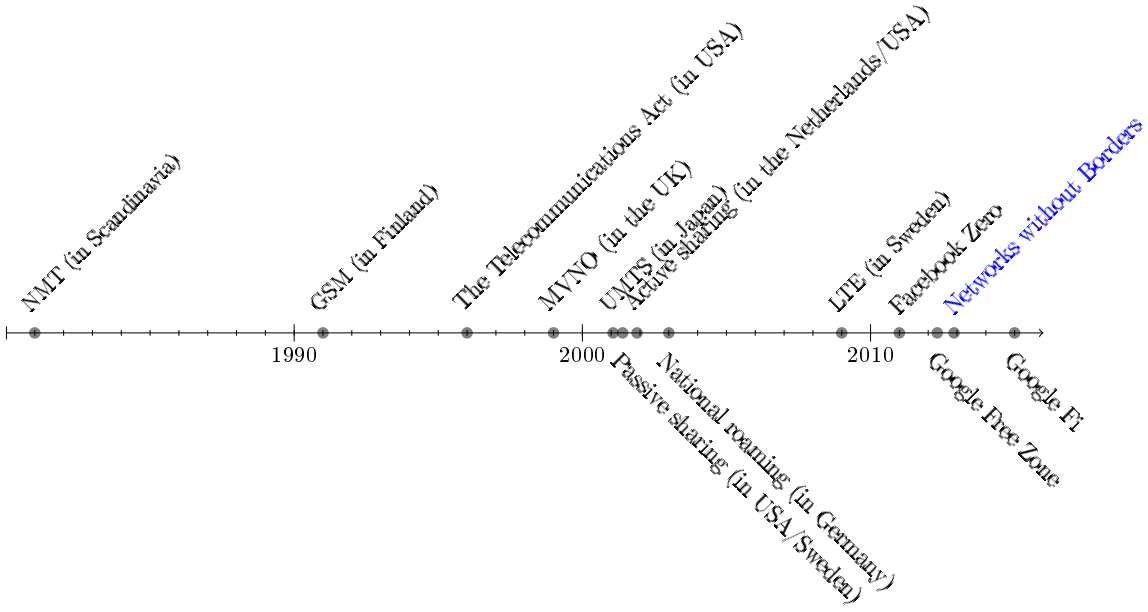


Figure 1.2: Chronology of the key trends in commercial mobile networks towards Networks without Borders.

What statistical models are representative to inter-operator resource sharing in spatially distributed mobile networks?

Assessing the performance of shared wireless mobile networks is of high importance to the cellular networks industry, as the fundamental gains and limitations of such shared networks invariably affect business decisions made by the entities involved in the mobile communications value-chain. Yet the state-of-the-art research does not provide satisfactory answers to the questions about the statistical features of multi-operator mobile networks and the kinds of models that could be used to analyze the efficiencies of mobile network resource sharing in such networks. We address this problem by applying spatial statistical analysis to various real mobile network datasets available through websites of national communications regulators. Based on the observed statistical features we construct statistical models which we use to replicate these features to the desired level of accuracy. Subsequently, we use these models to answer the next research question.

What is the efficiency of inter-operator resource sharing in spatially distributed mobile networks?

In the research to date we find evidence for mobile network sharing leading to improvements in coverage, or transmission rates, but little is known in general about the dependency between various mobile network resource sharing policies and spatial properties of shared networks. A key aspect of this analysis is recognizing the fundamental trade-offs and fungibilities involved in resource sharing, and the resulting network performance gains. Addressing this question requires that we develop models of shared mobile network resource usage over spatially distributed networks. Leveraging on the spatial statistical results obtained while addressing the previous question, we develop a stochastic geometry framework to study the efficiency of various infrastructure and spectrum sharing policies under different spatial models of radio access infrastructure deployment. Such an approach allows us to generalize the results over many possible spatial distributions and densities of the networks of the sharing operators. Moreover, in any scenario of spectrum sharing, we also consider the impact

of inter-operator interference coordination (or lack thereof) and different ways in which spectrum can be aggregated.

How network resource sharing affects utilities obtained by traditional mobile network operators and over-the-top service providers?

Over-the-top service providers, such as Google or Facebook, are asserting their role in the mobile networks landscape. OTTs have a fundamentally different business model to the incumbent MNOs, which is likely to affect the technological decisions they will make. From that perspective it is important to understand the potential of the existing networks to satisfy the requirements of these new entities, as well as the strategies that they may take up to provide mobile services to their users. One interesting aspect of the problem is the ability of OTTs to dynamically select a set of infrastructure that meets the requirements of the end-user or the service. These requirements may have a variety of forms, from the basic quality of service aspects such as coverage or capacity, to more complex ones such as indoor coverage or low latency, or a combination thereof. Another aspect is related to capacity extensions of the existing infrastructure deployments. OTTs face a dilemma of whether to invest in their own radio access infrastructure, similarly to Google's WiFi deployments made in some North American cities, or enter into a service level agreement (SLA) with an MNO, following Google's Project Fi example. In order to model these aspects and address the question, we use optimization combined with game theory. This allows us to focus our attention on the relationship between the strategies that these entities may have and the resulting efficiency of a shared network. In this case an important aspect of the efficiency is related to costs accrued by the entities from using certain types of radio access infrastructure.

1.3 Outline

This thesis can be divided into three parts. In the first part, which consists of the first three chapters of this thesis, we provide a setting and motivation for the vision of shared future wireless networks, and we detail trends in commercial mobile networks which pave the way towards this vision. Eventually, we provide background information on the analytical tools we use within the scope of this thesis. In the second part, we provide a description of our technical work. In each technical chapter we individually motivate the specific research problem, present the related work, formulate the corresponding system model, discuss any analytical and numerical results, and provide a follow up discussion. The last part of this thesis is devoted to the description of the broader vision behind our work and key challenges ahead for future wireless systems that stem from this vision. The chapters of our thesis provide the following information:

Chapter 2 – Key trends in wireless mobile networks

In Chapter 2 we discuss three trends that pave the way towards the vision of Networks without Borders. These trends include: the inter-operator network sharing, the formation of MVNOs, and the increasing role of OTTs. We present each trend comprehensively, providing definitions and numerous commercial examples of related activities. We also discuss how each of the trends sets the path for the type of future for wireless networks we envision in this thesis.

Chapter 3 – Background

In the Background chapter we discuss the main analytical tools we apply in this thesis to address the research questions posed in the previous section. In our work we foresee two types of studies: performance evaluation, using tools of stochastic geometry and spatial statistical analysis, and decision-making modelling, using optimization and game theory. In this chapter we present key definitions and some intuition behind usage of these tools, along with references to manuscripts that give a more in-depth treatment of the topics.

Chapter 4 – Multi-operator patterns in wireless mobile networks

In Chapter 4 we are concerned with spatial statistical analysis of multi-operator real network data. Specifically we apply stochastic spatial analysis to multi-operator base station location information from Ireland, Poland, and the UK. Effectively, we construct models of joint multi-operator deployments that pass goodness-of-fit tests. Subsequently, we evaluate usefulness of the selected models through wireless network-relevant analysis. We enrich our analysis with a study of downlink transmit power allocations in the UK data.

Chapter 5 – Inter-operator resource sharing

In Chapter 5 we study the dependency between various network sharing policies and spatial properties of shared networks, and the resulting efficiencies. We develop a stochastic geometry model of a shared mobile network that incorporates various infrastructure and spectrum sharing policies under different spatial models of radio access infrastructure deployment. Whenever the Poisson point process (PPP) is used, we produce closed-form analytical results. Yet one can only go so far with mathematical modelling and our study of coordination mechanisms and non-PPP geometries, which introduce a dependence that is considered in the state-of-the-art as intractable, is based on Monte Carlo (MC) simulations, with the PPP models being used to validate them. We use the results obtained to discuss the impact of sharing policies and parameters such as spatial clustering level, coherence bandwidth, or operator size on the sharing efficiency.

Chapter 6 – OTT-MNO sharing

In Chapter 6 we explore two problems: the selection of a minimum cost subset of infrastructure resources by an OTT, and the cooperation between an OTT and an MNO in serving mobile demand of the OTT. First, we model the problem of minimum cost selections as an optimization problem which we solve to optimality for an input dataset consisting of real base station locations for the country of Poland and another dataset consisting of synthetic locations generated from stochastic models of deployment. Second, we model, following Google Project Fi, an OTT that, in addition to providing content, provides wireless connectivity to its mobile subscribers. It does so either by building its own infrastructure or entering into an SLA with some MNO, to let the MNO serve its demand. We model the problem as a cooperative game between an OTT and an MNO, for which we define the corresponding utility space and the non-cooperative point related to the Nash equilibrium, and, using the Nash bargaining framework, we derive an analytical expression for the bargaining solution of the game. This allows us to study the impact the spatial distribution of mobile demand

and infrastructure cost have on the outcome of cooperation and the utilities obtained by both OTTs and MNOs.

Chapter 7 – Broader vision

In the last chapter we present our vision of future wireless mobile networks which is based on the idea that network resources are subject to virtualization, and made available to the creation of dynamic instances of networks tailored to the over-the-top service that is offered through them. Indeed, Google Project Fi can be considered an early example towards this vision whereby an over-the-top service provider becomes also an infrastructure provider, a mobile operator, and a mobile virtual operator. What we discuss in this chapter is that the proposed vision results in a substantially different mobile communications value-chain, and that it leads to a number of socio-political challenges, which are to be addressed should the change happen.

1.4 Contributions

In this section we summarize research contributions made as part of this thesis. The contributions we present range from conceptual insights to analytical results, yet it is worth noting that one general contribution of this thesis lies in recognizing a specific research direction and proposing a methodology to study it. We group our contributions according to the chapter in which they appeared.

Chapter 2

In this chapter we discuss background and related work. As part of our related work discussion we make the following contribution:

- We outline trends in the evolution of the wireless network ownership and control model; specifically, we *postulate that there exists a relationship between inter-operator network sharing, mobile virtual network operators, and service providers becoming mobile service providers.*

Chapter 4

In this chapter we focus on the spatial statistical analysis of multi-operator real network data. As a result of our analysis we find that:

- There is a *high degree of spatial dependence between the infrastructure deployment made by multiple operators*, hence mobile deployments cannot be modelled with a superposition of independent realizations of a point process describing a single-operator deployment.
- *Clustered point processes best represent base station clustering occurring in multi-operator mobile deployments*, with the log-Gaussian Cox process providing the most compelling fitness results.
- Similar *clustering behaviour occurs over and over again for similar areas* in different countries, which provides an evidence for the universality of our model.

- *Downlink macro- and microcell transmit power allocations* in a real 3rd generation (3G) network, both in a single and multi-operator case, *can be considered independent and identically distributed*, and can be modelled via random assignment from a probability distribution model.

Chapter 5

In the fifth chapter we utilize stochastic geometry to assess the trade-off between spectrum and radio access network sharing among mobile network operators. Our contributions can be summarized as follows:

- We *derive analytical expressions for the coverage probability and the average data rate for a shared network* created as a union of networks distributed according to the PPP for the scenarios of infrastructure and spectrum sharing both with channel bonding and dynamic channel selection.
- We show that while *spectrum sharing alone provides an increase in the achievable data rate* (at the expense of coverage), *infrastructure sharing results in improved coverage*. When the two sharing types are combined, we observe significant increase in the average user data rate and improvements to the network coverage as compared to the non-sharing scenario. Consequently, *combination of the two approaches does not linearly scale* the respective coverage and data rate gains.
- Yet, as we show, even *a simple mechanism of coordination in spectrum sharing allows operators to reduce the impact of this trade-off* and improve their coverage without hampering the user data rates.
- We observe that *the way in which spectrum is being shared makes a difference to both the coverage and data rate performance*. In particular, when coverage is the main concern, it is more beneficial for operators to allow for their transmitters to dynamically select among the pooled spectrum bands, rather than simply bond the available channels.
- We show that *clustering between the networks of the sharing operators has a significant impact on the efficiency* of resource sharing.
- We show that *the relative density of the networks of the sharing operators has a significant impact on how the operators perceive sharing gains*; specifically, it is the larger operator that gets the most benefits from spectrum sharing.

Chapter 6

In the sixth chapter we explore some dilemmas that OTTs face when entering the mobile communications market. In that respect we make the following contributions:

- We show that *considering a multi-operator network infrastructure as a collective* may lead to building up coverage provisioning networks which *require just a fraction of all the available base stations.*
- We find that *the lowest density areas may benefit greatly from infrastructure sharing*, which may provide a justification for joint infrastructure investments in typically under-served areas, with an effect on the digital divide.
- We *derive an analytical expression for the Nash bargaining solution (NBS)* of a cooperative deployment game between an OTT and an MNO.

- We show that *the OTT–MNO cooperation can be sustainable* only if the cost of infrastructure to the MNO is not excessive.
- We find that *the spatial clustering of the OTT demand affects the deployment price* and may lead to reductions in fees that the MNO imposes on the OTT for its traffic.

Chapter 7

In the concluding chapter we focus on laying out the vision of future wireless mobile networks, which we term Networks without Borders. Within the scope of this chapter we make two contributions:

- We *propose a vision for future wireless mobile networks*, which is based on a market-place of virtual network operators that construct networks from a pool of shared resources.
- We *postulate key socio-political challenges* associated with the vision.

1.5 Dissemination

This section lists both peer-reviewed and non-peer-reviewed dissemination undertaken during the PhD project. The work that is directly relevant to this thesis is marked with an asterisk. The remainder of the dissemination has been pursued under the scope of other projects and collaborations.

1.5.1 Peer-reviewed

Journals

1. ∗J. Kibiłda, N. J. Kaminski, and L. A. DaSilva, "Radio Access Network and Spectrum Sharing in Mobile Networks: A Stochastic Geometry Perspective", (in review).
2. P. Di Francesco, J. Kibiłda, F. Malandrino, N. J. Kaminski, and L. A. DaSilva, "Sensitivity Analysis on Service-Driven Network Planning", (in review).
3. ∗J. Kibiłda, B. Galkin, and L. A. DaSilva, "Modelling Multi-Operator Base Station Deployment Patterns in Cellular Networks", IEEE Transactions on Mobile Computing, 2016 (to appear).
4. ∗L. E. Doyle, J. Kibiłda, T. K. Forde, and L. A. DaSilva, "Spectrum without Bounds, Networks without Borders", Proceedings of the IEEE, vol. 102, no. 3, pp. 351–365, March 2014.

Conferences

1. ∗J. Kibiłda, F. Malandrino, and L. A. DaSilva, "Incentives for Infrastructure Deployment by Over-the-Top Service Providers in a Mobile Network: A Cooperative Game Theory Model", IEEE International Conference on Communications (ICC), May 2016.
2. B. Galkin, J. Kibiłda, and L. A. DaSilva, "Deployment of UAV-mounted Access Points According to Spatial User Locations in Two-tier Cellular Networks", IFIP Wireless Days, March 2016.

3. ∗J. Kibiłda, P. Di Francesco, F. Malandrino, and L. A. DaSilva, "Infrastructure and Spectrum Sharing Trade-offs in Mobile Networks", IEEE Dynamic Spectrum Access Networks (DySPAN), September 2015.
4. ∗B. Galkin, J. Kibiłda, and L. A. DaSilva, "Stochastic Modelling of Downlink Transmit Power in Wireless Cellular Networks", 20th IEEE International Workshop on Computer Aided Modelling and Design of Communication Links and Networks (CAMAD), September 2015.
5. A. Selim, J. Kibiłda, P. Sutton, and L. E. Doyle, "Real-Time Interference Reduction for OFDM-based Dynamic Spectrum Access Networks", IEEE DySPAN, April 2014.
6. ∗J. Kibiłda, and L. A. DaSilva, "Efficient Coverage through Inter-operator Infrastructure Sharing in Mobile Networks", in Proceedings of IFIP Wireless Days, November 2013.
7. ∗L. A. DaSilva, J. Kibiłda, P. Di Francesco, T. K. Forde, and L. E. Doyle, "Customized Services over Virtual Wireless Networks: The Path towards Networks without Borders", Future Network and Mobile Summit (FNMS), July 2013.
8. J. Kibiłda, J. Tallon, K. Nolan, and L. A. DaSilva, "Whitespace Networks Relying on Dynamic Control Channels", 8th International Conference on Cognitive Radio Oriented Wireless Networks (CROWNCOM), July 2013.
9. J. Tallon, J. Kibiłda, T. K. Forde, L. A. DaSilva, and L. E. Doyle, "Receiver-driven handover between independent networks", IEEE DySPAN, October 2012.

1.5.2 Non peer-reviewed

Work presented in this thesis has also been disseminated through other means than scientific articles such as scientific posters or seminar talks. The relevant dissemination includes:

1. J. Kibiłda, "Stochastic geometry model for coordinated infrastructure and spectrum sharing in wireless mobile networks", Information Theory and Applications (ITA) Workshop, San Diego (CA), 31 January – 5 Friday 2016 (*poster*).
2. J. Kibiłda, "Spectrum and Infrastructure Sharing Between Mobile Network Operators: A Stochastic Geometry View", Simons Conference on Networks and Stochastic Geometry, Austin (TX), 18–21 May 2015 (*poster*).
3. J. Kibiłda, "A thing about network sharing", seminar at Poznan University of Technology, Poznań (PL), 19 November 2014 (*talk*).
4. J. Kibiłda, "Networks without Borders: coverage study on a shared wireless infrastructure", University College Cork, Cork (IE), 20 May 2013 (*talk*).

1.5.3 Art projects

Together with Jessica Foley, a fellow PhD candidate at CTVR, we worked on a collaborative arts project entitled "Green-graphs & IRIS". The idea behind the project was to (literally) illustrate some of the models discussed in this thesis, as visual and verbal impressions of the limitations in comprehension and understanding of cross-field research. This collaborative project has led to the creation of posters illustrating (on a white background) base station deployments in Ireland, embedded with additional information derived from combinatorial interpretations of the said deployments. The images were followed by hand-drawn traces made by Jessica and a film-essay reporting the collaborative process.

The project was exhibited at:

1. The CTVR Communicating Communications Plenary, September 2013 (see Fig. 1.3).
2. As part of *Difference Engine – Accumulator III* at Limerick City Gallery of Art, October 2013 (see Fig. 1.4).
3. The Expanded Lyric conference at Queens University Belfast, April 2014[4].

Figure 1.3: One of the images from "Green-graphs & IRIS" presenting multi-operator base station deployment around the Cork city area, with cell radius denoting the local value of a specific summary characteristic and the colour denoting freuquency used by the base stations.

Figure 1.4: Hand-drawn traces of the graphs created as part of the "Green-graphs & IRIS" project, exhibited at Limerick City Gallery of Art.

[4]The video can be found at `https://vimeo.com/89157216`.

2 Key trends in wireless mobile networks

Key trends in wireless mobile networks

"The impact of new technologies is invariably misjudged because we measure the future with yardsticks from the past."

- Stephen Baker, *Can Social Media Sell Soap?, The New York Times*

In the Introduction we listed trends in the evolution of commercial mobile networks which pave the way towards the vision of Networks without Borders. These trends include the inter-operator network sharing, the formation of Mobile Virtual Network Operators (MVNOs), and the increasing role of over-the-top service providers (OTTs). In the following chapter, we detail how each of the trends has manifested itself and how it possibly sets the path for the future of wireless networks we present in this thesis. As such we emphasize the inter-operator network sharing as it comes in a variety of flavours and each of the flavours has its distinct impact on the general consensus about wireless mobile network ownership, architecture, and regulation. The following text is loosely based on our positioning paper, presented at the Future Networks and Mobile Summit in 2013 (see [10]).

2.1 Inter-operator network sharing

The origins of mobile network sharing can be traced back to the US Telecommunications Act of 1996 (see [12]). The intention of this act was to break monopolies in the delivery of shorthaul connectivity in both wireline and wireless networks by mandating telecommunications infrastructure providers to: (i) lease parts of their infrastructure to competitors; (ii) provide access at wholesale rates to any service they offer; (iii) introduce reciprocal call termination rates. According to [13], it was not until around the year 2000 when the first sharing deals were struck.

In the early 2000s the standardization and spectrum license auctioning for the 3rd generation (3G) mobile systems has resulted in operators securing spectrum licenses despite their costs exceeding estimated business case indicators [14]. However, the mobile communications market was in a rapid expansion phase fostered by economic growth, with high-tech industry stocks booming and wide availability of credits from deregulations in the financial market. The belief was that this situation will continue to last. In such a context, sharing deals were rather uncommon and the technological scenarios were limited to a handful of cases, with the main concerns of operators being the competitive advantage and possible equipment compatibility issues. The situation started to change with the economic downturn of 2001. Mobile operators, previously ecstatically bidding up the license prices, now had to turn their attention to reducing their capital and operational costs. As it turned out, one of the most important, and the easiest to apply, cost reduction measures was

network sharing. Depending on the type of sharing applied, the savings may be as high as 40% in capital costs, and 15% percent in operating costs spanned over a five-year period [15]. Due to savings occurring to both capital and operating costs, network sharing may be a viable option for both new network roll-out (either greenfield deployment or expansion) and network consolidation scenarios. The first documented examples of network sharing agreements we could trace occurred in 2001. These examples include the Groupe Special Mondiale (GSM) site sharing deal between T-Mobile and O2 in Germany, the GSM radio access network (RAN) sharing deal between T-Mobile and Cingular in the US, and Telenor and Three joint venture 3GIS to deploy a shared Universal Mobile Telecommunications System (UMTS) network (interestingly, the last one involved also spectrum sharing).

In response to the new trend in 2003 the 3rd Generation Partnership Project (3GPP) put together an initial report on the possible models for network sharing [16], which was followed by standards defining the scenarios (see [17]) and functional architectures for the shared networks (see [18]). At around the same time the European Commission started promoting network (facility) sharing for the benefit of "town planning, public health or environmental reasons" [19]. Yet the predominant policy supported by the European Commission was that of service-level competition based on a minimum degree of independence in the control of network infrastructure. As it was understood back then, some types of network sharing may lead to "operators offering similar network coverage, quality and transmission speeds", and, perhaps even, "collusion as operators exchange confidential information, increasing the predictability of commercial behaviour by operators" [20]. All these actions were seen as limiting mobile network operators (MNOs) capability to provision differentiated services, hence restricting competition. Effectively, any type of sharing that would involve sharing of operating frequencies or core network elements would not have been permitted under European law. This situation did not change until the ruling of the European Court of First Instance (CFI) in the case of T-Mobile and O2 appealing against the European Commission's decision to block their national roaming agreement. The ruling stated (as summarized in [20]) that "whether or not the service is available should not be a key competitive parameter but rather a basic requirement". In other words connectivity is a right, and MNOs shall be free to define the terms and conditions of network sharing agreements.

Indeed, in the state of the art we can find various forms of network sharing agreements. From our perspective the most relevant way to classify them is based on the level of control and ownership over the key network elements, as it allows us to discuss the architectural consequences of those sharing agreements. In [13] we have three types of sharing agreements defined:

- passive sharing,
- active sharing,
- national roaming.

In the following we will discuss each sharing type and provide some examples. Yet, due to commercially sensitive nature of the sharing agreements, except for a handful of case studies, details of the agreements are generally not available to the public.

2.1.1 Passive RAN sharing

Passive sharing, often referred to as *site sharing* or *co-location*, is the most wide-spread form of sharing, employed by mobile operators as early as 2000/2001. Passive sharing refers to sharing

agreements in which passive elements of the radio access infrastructure, such as cabinets, cooling equipment, power supply, mast and, in some cases, antennas, are jointly utilized by the sharing operators. In the main, passive sharing does not require a new network configuration or tampering with the processing units available in the active elements of the network. From a business perspective passive sharing can be realized through bilateral agreements either between mobile operators, or between mobile operators and third parties, so called tower companies, which provide passive telecommunications infrastructure to the operators [21]. Over the years, passive sharing has been encouraged by the European Commission and the national regulators as a way for mobile operators to limit the number of telecommunication sites required. The intention was to reduce the social impact of mobile network deployments[1]. There exist several examples of regulators promoting passive infrastructure sharing, ranging from the Indian Department of Telecommunications subsidizing setting up and managing mobile communications towers[2], the Canadian Radio-television and Telecommunications Commission (CRTC) mandating site and antenna sharing as part of the auctioning of 2 GHz band (see [20]), to Bahrain where the Telecommunications Regulatory Authority publishes on its website templates of mast sharing agreements (see [22]).

As mentioned previously passive sharing has been ongoing for a while now and we have various commercial examples of passive sharing world-wide. Some interesting examples include:

- Cornerstone, a joint company of O2 and Vodafone established to manage the existing masts of both operators, as well as plan and deploy new masts [23].
- Cross-european Vodafone and Telefonica passive sharing agreements, which involve countries like Germany, Ireland, the UK, Czech Republic, and Spain. The agreements supported sharing of power supplies, air ducts, and, in some cases, also antennas and backhaul (see [24, 25]).
- Vodafone and TIM in Italy [26].
- Passive sharing fully mediated by tower companies, for example, in Costa Rica and Panama (passive sharing in these countries was an imminent effect of passive infrastructure investments made by tower companies which preceded mobile network deployments made by mobile network operators) [27].
- State-controlled passive sharing in Ghana, which came as a result of the American Tower Corporation having a monopoly to build and manage passive network infrastructure for the country (see [22]).

In principle, site sharing is considered a sharing technique which does not limit operators' capability to individually set network configuration parameters. Yet that is a misleading point of view, as the sharing operators, when deciding on the details of site sharing, have to take into account, for example, load bearing capacity of towers, antenna azimuths/tilts, capacity of the cabinets, or the number of antennas that can be placed on a given site with respect to local landscape planning regulations [26]. In a way passive sharing may be considered an example of a crude form of sharing which, while bearing some socially desirable outcomes, does not allow the sharing operators to increase the flexibility of their networks, and as such is insufficient to give rise to service-driven networks.

[1] Communications masts and antennas are considered to have a negative visual impact on the landscape and wireless communications equipment may be also thought of as the source of ambient noise, electromagnetic interference or health concerns to the local communities.

[2] Interestingly, the condition to receive the subsidy was that the deployed infrastructure is used by at least three mobile operators (see [20]).

2.1.2 Active RAN sharing

Active sharing refers to scenarios in which active elements of the radio access infrastructure such as base stations or base station controllers are shared. This requires the mobile network architecture to support virtualization, multi-tenancy, and roaming. Despite the latter being one of the foundational features of the early mobile network architectures, the pre-existing architectures had provided no native support for network sharing [17]. As noted previously, this changed once network sharing has received some recognition, and the key sharing network architectures to date are tightly coupled with the 3GPP's functional network architecture (see, for example, [28]). A typical feature of each active network sharing architecture to date is a shared RAN, while the main difference lies in how the spectrum resources and the core networks are being split and utilized.

One widely-applied sharing architecture, Multi-Operator RAN (MORAN), was developed by Nokia Siemens Networks. In a MORAN the radio network controller (RNC) and parts of the base station are logically partitioned between the sharing operators, while core network elements and frequency spectrum remain separated. Consequently, site-level parameters such as antenna tilts are set jointly, while cell-level parameters such as the scrambling code (required for the cell synchronization) remain specific to each sharing operator.

An alternative architecture proposed by the 3GPP is a Multi-Operator Core Network (MOCN). In MOCN dedicated core network elements from multiple operators connect to a single network controller (or a base station) located in the shared RAN. A shared RAN may operate on either pooled or dedicated frequency resources [18]. In this sharing regime both cell and site-level parameters are set jointly, which significantly limits the control over individual wireless service quality.

Another architecture proposed by the 3GPP is a Gateway Core Network (GWCN). In addition to the shared RAN, GWCN allows for sharing of some of the core network elements such as the mobility management entity (MME) (see [18]). Frequency spectrum is either pooled or dedicated to each operator. GWCN is implemented using the roaming features in the core network [26]. Since such an implementation allows for no physical or logical separation, wireless service differentiation between the sharing operators is not possible. In each architecture the user equipment behaviour is the same, and while a user equipment is required to discriminate between the sharing operators the terms of the sharing agreement are made transparent to the subscriber.

Regulators' support for active sharing (or rather lack thereof) has changed over time. In the main, regulators world-wide were reluctant to accept some types of active sharing deals, the main issue being "the entity awarded the spectrum not being the entity using it in practice" [26]. However, the cost benefits of active sharing outweighed any potential downsides and many operators world-wide decided to pursue (typically bilateral) active sharing deals; some examples include:

- MORAN:
  - T-Mobile and Three UK formed a joint venture company Mobile Broadband Network Limited to consolidate and manage their 3G infrastructure (see [20, 23]).
  - A tighter consolidation took place in Poland between local branches of Orange and T-Mobile that created a joint venture NetWorks! to consolidate existing 2^{nd} generation (2G) and 3G networks, as well as deploy a shared 4^{th} generation (4G) network. Interestingly, the consolidation has resulted in some pre-existing base stations having been decommissioned (see [29]).

  - In 2014 Eircom and Three Ireland agreed to consolidate existing sites throughout Ireland into one network and jointly deploy any new sites to meet target coverage requirements (see [30]).
  - A number of sharing agreements comes from Brazil[3] where Oi and TIM went into a sharing agreement to deploy in 12 cities a 4G network operating in 2.5 GHz band [27]. Similarly, Claro and Vivo (Telefonica) entered into a short-time sharing agreement [27].
  - In Colombia, Movistar and Tigo entered into a sharing agreement to jointly deploy Long Term Evolution (LTE) network operating in the Advanced Wireless Services (AWS) band (1.7/2.1 GHz), and consolidate and upgrade the existing infrastructure to support a multi-mode operation (see [27]).
- MOCN:
  - A combination of GWCN and MOCN that happened in Sweden between Europolitan (later Telenor and Vodafone) and Hi3G (see [31]). The two set up a joint venture 3GIS which was responsible for deployment and management of 3G infrastructure. Effectively having no subscribers 3GIS was just a provider of capacity to the two operators. Moreover, Hi3G subscribers could roam into the 2G network of Europolitan and, due to regulatory obligations which mandated at least 30% of the population to be covered by dedicated networks, both operators developed their dedicated 3G networks.
  - Sweden's Telenor and Tele2 forming a joint venture Net4Mobility dedicated to the development and maintenance of a 4G network. Interestingly, Tele2 had already formed a joint venture with another Swedish operator TeliaSonera to develop a joint 3G infrastructure, yet this cooperation was not prolonged to 4G infrastructure as TeliaSonera had decided to develop its dedicated 4G infrastructure (see [31]).
  - In Canada the MOCN architecture was utilized to provide an HSPA shared network to both Bell and TELUS (see [26]).

Wide support for active network sharing is key to our vision as it forces mobile operators to make decisions about the types and amounts of resources being used at a given time. In addition to flexibility, active network sharing shall provide means necessary for isolation between the networks of the sharing operators, which is necessary to creation of customized networks. Yet, when active network sharing was being introduced, virtualization technologies were still far from the market. Effectively, active network sharing implementations had to rely on static divisions of resources (on equipment-level) and roaming architecture (which we will discuss subsequently), which has limited the application of active sharing as well as the achievable gains [11].

2.1.3 National roaming

National roaming allows subscribers of one operator to rely on the coverage provided by a network of another operator. International roaming is a good example for this, as it allows a subscriber to use its mobile service while abroad where her home operator has no network. In the context of network sharing, roaming is used domestically as a way for competing operators to increase coverage and for greenfield operators to provide country-wide coverage prior to rolling out a country-wide network. The typical case for national roaming is that each sharing operator has its own frequency spectrum,

[3] All these sharing initiatives are fairly recent (2013-2014) and are strictly related to the infrastructural developments required by Brazilian government to support the organization of the football world cup in 2014 and the summer Olympics in 2016.

dedicated radio access network and core network, and that subscribers of the sharing operators may connect (transparently) to either of the networks. This approach requires no updates to the radio access network part, with only the core network being affected. However, as noted in [21], shared network elements may also be introduced.

National roaming (and active sharing to some extent) has been a major cause for concern to the European Commission, which has seen national roaming agreements as endangering service-level competition in the telecommunications market caused by the loss of independent control over the network parameters and therefore lack of mechanisms to provide differentiated services. As noted previously it is not until 2008 and the CFI ruling that national roaming became effectively legitimized in Europe. This included also reciprocal agreements signed to increase the pace of coverage and service quality improvement during the roll-out phase. National regulators outside of Europe were more favourable towards national roaming agreements and, for example, in Canada incumbent MNOs were required to provide roaming for a period of 5 years to new entrants to the market (see [20]). Nevertheless, competitive concerns have been the major cause for national roaming being limited to pre-specified areas only, in the case of coverage extensions, and for a limited duration, in the greenfield operator case.

Some commercial examples of national roaming deals include:

- O2 and T-Mobile national roaming deals in Germany and the UK. In Germany O2 subscribers were allowed to utilize T-Mobile's 3G network, while in the UK subscribers of both operators could use the network of the other operator if their operator did not provide coverage over a given city area[4].
- Orange and T-Mobile in the UK have provided a new opt-in service to the customers that would allow them to roam with either operator, without paying extra roaming fees [33].
- T-Mobile and H3G provided each other subscribers roaming services in Austrian rural areas [24].
- In France, Free Mobile has entered into a roaming deal with Orange to utilize Orange's existing 2G and 3G networks [24].

Since national roaming effectively removes any service-level differentiation between mobile operators, it is the most progressive form of network sharing available, and perhaps a good way to think of providing basic coverage even to the remotest areas. What differs our vision from national roaming is the fact that our vision requires some form of service-level differentiation that would enable customizable networks capable of delivering over-the-top services to subscribers at various quality points.

2.2 Mobile Virtual Network Operators

A natural consequence of opening up wireless network infrastructure to other mobile network operators was to allow new actors that do not have a license to spectrum to act as wireless service providers. This very idea lies at the core of what is an MVNO as we know it today. Originally,

[4]The ability of the two operators to roam was restricted by the European Commission, for example, depending on the area national roaming was allowed for a limited period of 3, 5, or 6 years [32]. The restriction resulted in an appeal to the CFI and a subsequent ruling which held that, on contrary to the Commission's view, the national roaming deal will lead to an increase in competition due to wider availability of 3G services [20].

MVNOs entered the European cellular market through a mix of voluntary and involuntary mechanisms [3]. Some MVNOs were created to utilize extra capacity available in the networks, some to reach new customer market segments, and others emerged as a result of anti-trust and maximum usage fee policies imposed by the communications regulators to boost competition and improve the availability of wireless services [5]. Contemporary technical solutions allowing MVNOs to provide wireless services to their subscribers are very similar to the solutions used for national roaming [5].

By definition an MVNO is a wireless communications services provider that does not own the spectrum, radio access network or backhaul [11]. The only capital and operational costs that an MVNO incurs are the ones related to the development of the policy and charging system, distribution channel, brand and marketing, and, in certain cases, also deployment of the core network. An MVNO is not responsible for wireless coverage or capacity provisioning, instead it enters into a service level agreement (SLA) with an MNO to receive bulk access to the network services at wholesale rates, which allows it to resale the network access at retail prices [34]. Such an MVNO may differentiate itself from its MNO in three main ways: traffic quality (only to a limited extent), pricing and brand. The traffic quality-based differentiation relies on switching connections between RANs of different operators, therefore it requires a specific core network architecture and multiple SLAs with multiple MNOs [5]. Obviously, even meeting all these requirements would not ensure that an MVNO can actually provide a differentiated service; as it is often the case that networks of different mobile operators in urban areas are deployed following similar spatial patterns [35].

There are three types of architectures widely applied by MVNOs [34]: a full MVNO, a light MVNO, and a Mobile Virtual Network Enabler (MVNE). A full MVNO owns a dedicated core network which inter-connects to the access networks of MNOs. A light MVNO implements the billing system and the distribution channel only, while the rest of the core and access network is provided by some MNO. An MVNE is a middle-man that provides a core network and some business functions to the MVNO and inter-connects that MVNO with radio access networks (in this case an MVNO acts as a brand and a distribution channel). The choice of the network architecture for an MVNO depends on the business model of that MVNO, for example, an MVNO that is a re-branded version of the wholesale MNO is likely to rely on the core network being delivered by the MNO, whereas an independent MVNO may be willing to control parts of the core network to withhold customer information.

It is important to note that the MVNO model necessarily requires well-developed radio access infrastructure, and may not be viable for developing markets. Also, the policies applied by communications regulators to the MVNO market differ from country to country, and, similarly to network sharing, the European Commission was in support of interventions by communications regulators if there was an unsatisfactory level of competition in the market [36]. We can differentiate three types of attitudes towards legitimacy of the existence of MVNOs in the mobile communications markets [36, 37]:

- MNOs were required to provide interface for MVNOs (Finland, Sweden, Germany, or Spain),
- MNOs were required to implement specific type of interface for MVNOs as well as apply pre-specified charges (Denmark and Ireland),
- MNOs and MVNOs cooperate voluntarily (UK, Italy, or France).

The success stories of MVNOs vary considerably from country to country, and according to the GSM Alliance (GSMA) report "Mobile Infrastructure Sharing" from 2011 there were as many as

150 MVNOs in Germany or 61 MVNOs in Netherlands, whereas only 16 or 15 of them in Italy or Sweden. An example of a successful MVNO is the Freenet Group in Germany, which has acquired 14% of the German mobile customer market, or Lycamobile, which by 2011 had 6 million subscribers across 8 European countries. An often cited example of a successful MVNO is Virgin Mobile in the UK which emerged as a 50:50 joint venture between the Virgin Group and One2One. Virgin Mobile has been buying airtime and network capacity at wholesale rates from mobile operators and reselling it to its target customers. From the perspective of a customer Virgin Mobile assumes the typical role of an MNO, which includes provisioning of voice and data services, as well as assignment of Subscriber Identity Modules (SIMs) to users (see [20, 38]). Despite these success stories a significant number of new MVNOs fail to survive on the market, even if they are a part of bigger brands such as ESPN Mobile or Dysney Mobile.

Technological limitations of the existing network architectures and equipment have limited MVNOs presence in the market to competition on pricing or branding only. However, as noted previously, even within the technological framework of the existing networks some level of service-level differentiation is possible, and that is precisely what Google was relying on when in 2015 it unveiled its plans to become a wireless service provider [39].

2.3 Rise of service providers

Over the recent years we have seen a rampant increase in mobile data usage (see [7, 8]), driven primarily by the growing popularity of smartphones and mobile applications. This increase comes at the expense of popularity of mobile voice or text messaging services, which traditionally served as a sustainable source of revenue for mobile operators, but now are giving away the spotlight to applications such as Skype or Snapchat that essentially perform the same function, but for mobile operators are all mobile data [40]. From the perspective of these applications mobile data service is simply a pipe and is expected to introduce as little costs as possible. Given growing infrastructure costs (as remarked previously), the sustainability of a model in which mobile operators are solely responsible for the augmentation of mobile capacity is challenged. This is of particular concern for service providers, such as Google or Netflix, whose very business model is based on the perception of limitless traffic (as remarked by the Federal Communications Commission (FCC) in [41]). Taking Google as our prime example, we see an increasing push among the biggest OTTs to assert their role in how mobile networks are accessed and deployed [42]. Indeed, we see examples of OTTs deploying their own networks, becoming MVNOs, or entering into some form of SLA with mobile operators or infrastructure providers. In the following, we discuss real-life examples of these practices.

2.3.1 Facebook

"Connectivity is a human right", this catchy doctrine[5] lies at the heart of Facebook's business model which relies on monetizing customer data, mostly through offering customized advertising, and anticipation (or possibly even manipulation) of the customer needs (or lack thereof). From that perspective it is understandable that Facebook is interested in bringing ever more people to stay connected. One way of achieving this is to promote mobile data usage, which requires that:

[5] Expressed by Mark Zuckerberg in August 2014 via an online report entitled "Is Connectivity A Human Right?".

(i) we[6] define some basic set of services[7] people will associate with the Internet; (ii) then, by means of cheap, or free, data plans and light-weight apps we make them affordable to the public; (iii) we deliver new (and open) technologies and platforms to share network and computing infrastructure, and devise new policies that curb the costs of licensing of key resources such as electromagnetic spectrum.

The first two have been ongoing for a while now by means of Facebook Zero, which allows Facebook users to view a text-only version of the social network website without incurring mobile data usage charges. Facebook Zero is accessible (via zero.facebook.com) to Facebook users who are also subscribers of one of the mobile operators that have entered an agreement with Facebook. Facebook Zero offers free of charge access to notifications, friend requests, or status updates, while access to the rest of the content, such as photos or games, is subject to the usual data fees. Regardless of how Facebook compensates mobile operators for data being used[8], when it is only the Facebook account (and a smartphone) that a user needs to be able to use mobile services, why would a Facebook subscriber need to even care about a subscription to the mobile carrier? Linking this with the definition of the basic set of services expressed in terms of smartphone apps, rather than traditional voice, text or data services, we make networks, at least from the perspective of a single user, over-the-top service-centric. In other words, we anticipate that a user of these new services will require a network that is crafted towards the delivery of a specific over-the-top service or a set thereof, and the wireless infrastructure will provide only the means to set it up.

The last point from Facebook's manifesto is more of a declaration of the right path of actions, rather than a well-defined project. There are two aspects to this point. The first one is related to the use of shared and open infrastructure. The second one is related to improving the cost efficiency of obtaining key network resources, with the prime target being a change in radio spectrum licensing. The belief is that more flexible modes of spectrum licensing (for example, dynamic spectrum access or spectrum pooling) will lower barriers of entry for new players and spark innovation [34].

2.3.2 Google

Google is an example of an OTT that actively and aggressively pushes its business model towards wireless networks. In 2012 Google launched in some developing countries, such as the Philippines or Nigeria, Google Free Zone service. Google Free Zone works much alike Facebook Zero as it allows subscribers of operators with whom Google has signed an agreement to access without charge light-weight versions of Google mail, Google plus or Google search engine. This includes also viewing websites that were found using Google's search engine. Google's intention is to increase usage of the mobile internet as well as encourage more users (in developing countries) to setup Google user accounts and associate their Internet experience with Google's products. In addition, over the last few years Google has been active in pursuing spectrum licenses [44], promotion of spectrum sharing in frequency bands over 24 GHz [45], as well as developing projects such as Google Loon to provide LTE wireless access to remote and rural areas [46].

However, the most prominent example of Google's activity in wireless network domain is Project

[6] The report uses "we" to refer to Facebook as an organization. Here, idealistically, we are using "we" more broadly and refer to the society at large.

[7] Unsurprisingly, this set includes Facebook's online social network.

[8] Some speculation on how this compensation may actually happen can be found in the following Quartz magazine article [43].

Fi [47], which enables users equipped in Nexus 5/6 phones with Fi SIM cards to seamlessly connect to the fastest[9] network in proximity, be it WiFi or LTE. The WiFi hotspots come from the pool of hotspots deployed by Google in several American cities such as New York or San Francisco, whereas the LTE access is provided by an MVNO that signed SLAs with Sprint and T-Mobile US. The project is a move towards more flexible network architectures in which Google is capable of offering flexible bundles which are composed of value-added services, wireless connectivity, a device (with an operating system) and guaranteed end-to-end quality.

2.3.3 Other examples

A few other Internet giants have also joined Facebook in popularizing free Internet usage. Wikimedia has launched Wikipedia Zero, a programme (partnered with some mobile operators) to let people browse Wikipedia with no mobile data charges. Twitter has signed an agreement with a Swiss mobile operator to access text-only content from Twitter site for free or at low tariffs [48]. Moreover, when discussing the increasing role of service providers the Netflix–Comcast deal is also of relevance [49]. In early 2014 Netflix, which is one of the heaviest contributors to Internet traffic, and Comcast have struck an agreement under which Netflix have agreed to pay a fee to Comcast to "receive faster and more reliable access". The deal sparked a fierce debate about its implications to the "net neutrality". The proponents of net neutrality argued that such a deal will allow some wealthy OTTs to stifle the market, in essence by creating strong ties between OTTs and infrastructure providers, while the proponents of the deal pointed out that while some service providers (particularly the ones linked to the infrastructure providers) have to share the costs of the infrastructure, others (specifically bandwidth-hungry ones) receive a free ride which hurts competition among service providers.

All these examples are intended to show a change in thinking about provisioning of over-the-top services, in particular over mobile access. A change which, if pursued further, will require a revolution in the way that mobile technologies operate, especially on the network-level. Our proposed Networks without Borders paradigm, further discussed in Chapter 7, is a reflection of this revolution and a possible path in the evolution of mobile access provisioning.

[9] Here we do not delve into the question of how the selection process is made or what does "fastest network" mean.

3 Background

Background

"The whole of life is just like watching a film. Only it's as though you always get in ten minutes after the big picture has started, and no-one will tell you the plot, so you have to work it out all yourself from the clues."

- Terry Pratchett, *Moving Pictures*

The key trends discussed in Chapter 2, and, indeed, the entire vision of Networks without Borders as exemplified in the concluding chapter of this thesis, requires various types of studies to provide answers to the research questions posed in the Introduction. In that we foresee two primary types of studies:

- *Performance evaluation* – which involves studies of stochastic models of radio access infrastructure deployment made by multiple mobile operators. Stochastic geometry tools may then be applied to these models in order to derive expressions describing coverage and achievable data rates in a shared network with a stochastic model of spatial distribution of transmitters/receivers. Whenever closed-form expressions cannot be found, numerical evaluations and computer simulations may also be applied to quantify the performance.
- *Decision making* – which involves studies of decisions that individual entities from the mobile communications market make when being involved in network resource sharing. These decisions include deciding which resources to share, which resources to utilize or how to deploy a shared network. Each entity involved can make its sharing decisions irrespective of other entities or by assuming competitive or cooperative interaction with other entities.

We perform our studies by applying three types of analytical tools:

- *Stochastic geometry* and *spatial statistical analysis.*
- *Optimization.*
- *Cooperative game theory.*

In this chapter we present key definitions and some intuition on the analytical tools applied in this thesis, as well as we provide references to manuscripts that give a more detailed view of the theories involved.

3.1 Stochastic geometry and spatial statistics

A point process can be used to model the locations of base stations (transmitters or receivers) in a wireless network (in Fig. 3.1 we have a real base station deployment from Dublin represented as

a point process). A stochastic deployment model can be applied to characterize the distribution of interference or signal-to-noise and interference ratio (SINR) which will yield performance results in the form of, for example, the cumulative distribution function of SINR which reflects the outage probability of a typical user in a wireless network. For some types of stochastic geometry models this performance can be derived in closed-form, yet for a majority of state of the art models these are out of reach, typically due to our poor understanding of dependencies involved in the model. In those cases we have either the option of applying approximations, which in some cases provide relatively tight bounds, or simply utilize the model as the basis for numerical calculations either in the form of numerical integrations (when a small number of stochastic parameters is assumed) or Monte Carlo (MC) simulations.

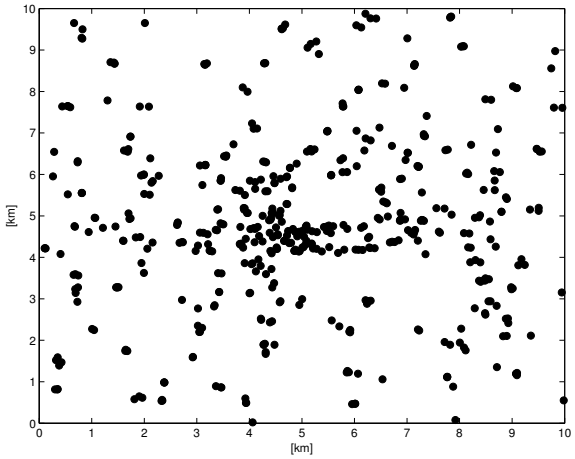


Figure 3.1: Locations of Meteor 3G base stations in the city centre of Dublin.

In order to characterize a wireless network deployment as a point process we need to define some of the underlying theoretical concepts. Let us first formally define a point process; following [50] we have:

Definition 3.1.1. A point process is a countable random collection of points that lie in some measure space. In our case this is usually the Euclidean plane, i.e., $\mathbb{R}^2$. The associated σ-algebra $\mathfrak{B}^2$ consists of the Borel sets, with the associated Lebesgue measure $v(\cdot)$.

A wireless network deployment can be represented as a point process which is a countable random set $\Phi = \{x_1, x_2, ...\} \subset \mathbb{R}^2$ with elements being random variables $x_i \in \mathbb{R}^2$. Following the random measure formalism [50], we also define $N(\mathcal{W})$ as a (random) counting measure which denotes the number of points in a set $\mathcal{W} \subset \mathbb{R}^2$. Often we will use $\mathcal{W} \subset \mathbb{R}^2$ to refer to a finite observation window and $\varphi \subset \mathcal{W}$ to refer to a collection of points in that window (i.e., φ is a sample in $\mathcal{W}$ from some point process Φ), in which case $N(\varphi)$ will refer to the number of points in the sample. Moreover, we denote $\Lambda(\mathcal{W})$ as the intensity measure of Φ, which, informally, is the mean number of points of a point process Φ in a set $\mathcal{W}$. In consequence, $\lambda(x)$ will denote the intensity function that describes the point density around location $x \in \mathbb{R}^2$. For some point processes, such as the homogeneous Poisson point process, the intensity function will be uniform, hence we will denote it as λ.

In our work we will utilize also marked point processes (see [50]). In order to obtain a *marked point process*, each point of Φ is assigned a mark m belonging to some mark space $\mathcal{M}$. This assignment results in another set $\widehat{\Phi}$ that consists of pairs of the following form (x_i, m_i), where

$x_i \in \Phi$ and $m_i \in \mathcal{M}$.

3.1.1 Point process models

There are three general classes of point process models: uniform, clustered[1] and regular. To say that a point process is uniform is equivalent to assuming that no inter-point interactions can be observed in a point pattern coming from that process. An example of a uniform point process is the homogeneous Poisson point process. Clustering means that some form of positive interaction (attraction) exists between points, leading to clustered patterns, while regularity stands for some form of negative interaction (repulsion) between points in the process. In our work we modelled a regular process as the Strauss process, while for modelling clustered patterns we used: the Gauss-Poisson process, the Matérn cluster process, the Thomas process, and the Log-Gaussian Cox process. In the following we define and briefly introduce each of the listed point processes. For a more comprehensive description we recommend [50–52].

Homogeneous Poisson point process

The homogeneous Poisson point process (PPP) with intensity λ, which describes the mean number of points to be found in a unitary area, is a point process in $\mathbb{R}^2$ that posses the following two properties:

1. The number of points $N(\mathcal{W})$ in any bounded set $\mathcal{W}$ is Poisson distributed with mean $\lambda v(\mathcal{W})$.
2. The number of points $N(\mathcal{W}_1), N(\mathcal{W}_2), \ldots, N(\mathcal{W}_k)$ in any k disjoint bounded sets is an independent random variable.

The greatest strength of the PPP is its analytical tractability, which comes from its reliance on the Poisson distribution. In spatial statistics, assuming that a pattern is a realization of a PPP is equivalent to assuming that there are no interactions between the points of that pattern.

Strauss process

The Strauss process (SP) is an example of a regular point process that comes from a family of Gibbs processes. A Gibbs process is formed by shaping the distribution of some other point process, typically a PPP, with some density function described on the space of counting measures, i.e., we reject points from the initial point process so to achieve the desired density. Realization φ of an SP with interaction radius $R \in \mathbb{R}_+$ and parameters $\beta \in \mathbb{R}_+$ and $\gamma \in [0, 1]$ can be described through the density function that takes on the following form:

$$f(\varphi) = \alpha \beta^{N(\varphi)} \gamma^{T(\varphi)}, \tag{3.1}$$

where α is a normalizing factor, and $T(\varphi)$ is the number of pairs of points (x, y) of the realization φ such that $||x - y|| < R$. Informally, R describes the maximum range at which any interaction between the points exists, β describes the intensity of the pattern, while γ is attributed to the strength of the inter-point interaction.

[1] Here we make a distinction between a clustered and cluster process. Essentially, we consider a cluster process to be a special type of a clustered processes.

Gauss-Poisson process

In the Gauss-Poisson process (GPP), clusters of points are distributed according to the PPP with intensity λ_p. These clusters consist of one or two points with probability $1-p$ and p, respectively. If a cluster has two points, one of the points is located in the center of the cluster, i.e., at the location of the parent point, and the other point is uniformly distributed on a circle with radius u centred at the parent point. Formally, the resulting pattern can be defined as follows:

$$\Phi = \bigcup_{i:x_i \in \Phi_p} (\Phi_i + x_i), \tag{3.2}$$

where $x_i \in \Phi_p$ and Φ_p is the parent PPP with intensity λ_p, and Φ_i is the untranslated daughter process, which formally can be defined as:

$$\Phi_i = \begin{cases} \{o, y_i\}, & \text{with probability } p, \\ \{o\}, & \text{with probability } 1-p. \end{cases} \tag{3.3}$$

In other words, any realization of the daughter process consists of one or two points: the first point being located at the origin $o = (0, 0)$ and the other point existing with probability p at a uniform location on a circle of radius u centred at the origin.

Matérn cluster process

The Matérn cluster process (MCP) is a doubly Poisson cluster process, i.e., parent points are generated according to a Poisson process with intensity κ, and daughter points, with the mean number of them μ, are uniformly distributed inside the circle of radius R centred at each parent point. The intensity of the daughter process takes the following form:

$$\lambda_d(r) = \frac{\mu}{v(b(o,R))} \mathbb{1}_{[b(o,R)]}(r), \tag{3.4}$$

where $\mathbb{1}_{[b(o,R)]}(r)$ is an indicator function, $b(o, R)$ denotes a circle centered at parent point o with radius R, and $v(\cdot)$ is the Lebesgue measure, which in the Euclidean space corresponds to the surface area. The intensity of the process is $\lambda = \kappa\mu$.

Thomas process

Like the MCP, the Thomas process (TP) is a doubly Poisson cluster process, where the parent points are distributed according to a Poisson process with intensity κ, while the daughter points are normally distributed, i.e., according to the intensity function:

$$\lambda_d(r) = \frac{\mu}{2\pi\sigma} \exp(-\frac{r^2}{2\sigma^2}), \tag{3.5}$$

where μ is the mean number of daughter points and σ^2 is the variance. The intensity of the process is again $\lambda = \kappa\mu$.

Log-Gaussian Cox process

In [35] we apply a subclass of Cox processes, where the logarithm of the random intensity function is a Gaussian process [53]. This subclass is called the Log-Gaussian Cox process (LGCP), with the intensity function denoted as $\lambda(x) = \exp(Y(x))$, where $Y = \{Y(x) : x \in \mathbb{R}^2\}$ is a real-valued Gaussian process with mean μ, and covariance function $\sigma^2 f(r)$, where σ^2 is the variance and $f(r)$ is some pre-defined spatial correlation function. For the LGCP to be well-defined the spatial correlation function has to be positive semi-definite [54]. This condition is met by, for example, the exponential function $f(r) = \exp(-r/\beta)$ and the stable function $f(r) = \exp(-\sqrt{r/\beta})$, where $\beta > 0$ is a scale parameter (other examples of functions that meet this condition can be found in [54]). The LGCP is fully characterized by μ, σ^2 and the parameter(s) of $f(r)$, which are easy to estimate due to the existence of closed-form expressions for the summary statistics. The process is also relatively easy to simulate based on the location dependent rejection of points generated according to the homogeneous PPP with $\lambda^* = \max(\lambda(x))$. Moreover, the intensity of the process is expressed as $\lambda = \exp(\mu + \sigma^2/2)$.

3.1.2 Summary characteristics

In spatial statistics there are numerous summary characteristics that describe and quantify local groupings of points, typically through their relative positions or inter-point distances. In the following, we briefly introduce the metrics that we will subsequently use for model fitting and goodness-of-fit tests.

Empty space function

Empty space function $F(r)$ describes the distribution of distance r from a point $o \in \mathcal{W}$ (for stationary point processes this corresponds to the origin) to the nearest point of the point process Φ:

$$F(r) = P\left(||o - \Phi|| \leq r\right) = P\left(N(b(o,r)) > 0\right), \quad r \geq 0. \tag{3.6}$$

We use the empty space function as the first-order goodness-of-fit statistic.

J-function

One way to investigate the interaction between two point patterns, each with a different type of points, observed within the same window, is to utilize the cross-type J-function [55]. The cross-type J-function is the ratio between two distributions: the distribution of distances from an arbitrary fixed point of type i to the nearest type j point, and the distribution of distances from an arbitrary fixed point of type j to the nearest type j point (i.e., the empty space distribution). Formally the cross-type J-function is given as:

$$J_{ij}(r) = \frac{1 - G_{ij}(r)}{1 - F_j(r)}, \tag{3.7}$$

where $G_{ij}(r)$ is the nearest neighbour distance function, which describes the distribution of distances from the points of type j to the points of type i, and $F_j(r)$ is the empty space function of the point pattern of type j. By this definition cross-type J-function can be used to quantify clustering or repulsion between two sets of points. In case the two sets of points are independent $J_{ij} = 1$, as is

the case with any PPP realizations. $J_{ij} < 1$ suggests that the points of different types aggregate, while $J_{ij} > 1$ suggests repulsion. However, the cross-type J-function taking value 1 for all r is not a sufficient characterization of independence between the two patterns [55]. It is also worth noting that J_{ij} is not symmetric in i and j.

L-function

The L-function is defined as $L(r) = \sqrt{K(r)/\pi}$ for $r \geq 0$, where $K(r)$ is a second-order summary characteristic which describes the mean number of points within radius r from points in point process Φ. The L-function is particularly useful in the estimation of model parameters, such as the interaction distance (maximum inter-point distance at which interaction force is at work) or interaction force (the force that describes mutual attraction or repulsion between points in a point process). For the PPP the L-function takes a convenient linear form $L(r) = r$. Hence, for some pattern φ, $L(r) < r$ will indicate regularity, while $L(r) > r$ will indicate clustering. We use the L-function to examine fitness of the estimated models with respect to second-order characteristics.

Pair-correlation function

The pair-correlation function $g(r)$ for a stationary point process Φ with intensity function λ and second order product density $\rho^{(2)}(r)$, which is the joint probability that there are two points of Φ in infinitesimally small areas surrounding two locations separated by distance r, is defined as follows [50]:

$$g(r) = \frac{\rho^{(2)}(r)}{\lambda^2}. \tag{3.8}$$

Informally, the pair-correlation function characterizes the frequency of inter-point distances in the process. Herein, we use the pair-correlation function for fitting clustered point processes to our data. The fitting method we use requires an analytical expression for the summary statistic in question. The expressions for the pair-correlation function of the analysed clustered point processes are summarized in Tab. 3.1.

Table 3.1: Analytical expressions for the pair-correlation function of the analysed clustered models, where $h(z) = 16/\pi(z\arccos(z) - z^2\sqrt{1-z^2})$ for $z \leq 1$ and $h(z) = 0$ otherwise.

Model	LGCP	MCP	TP
$g(r)$	$\exp\left(\sigma^2 f(r)\right)$	$1 + (4\pi R r \kappa)^{-1} h(\frac{r}{2R})$	$1 + (4\pi\kappa\sigma^2)^{-1} \exp(-\frac{r^2}{4\sigma^2})$

Mark-weighted K-function

The mark-weighted K-function describes the expected normalized sum of products of the mark of a point in a marked point pattern $\widehat{\varphi}$ and all of the marks belonging to points within a distance r, which can be formally expressed as follows:

$$K_{mm}(r) = \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N} \setminus \{i\}} \frac{m_i m_j \mathbb{1}_{[||x_i - x_j|| \leq r]}}{\lambda \mu^2}, \tag{3.9}$$

where $\mathcal{N} = \{1, 2, \ldots, N(\widehat{\varphi})\}$, $||.||$ denotes the Euclidean distance, $(x_i, m_i) \in \widehat{\varphi}$, λ is the density of the point pattern and μ is the mean of the mark values.

3.1.3 Spatial statistical analysis

Maximum pseudo-likelihood estimation

The method of maximum likelihood estimation relies on the likelihood function which describes the likelihood of observing a particular data sample given some probability distribution. This likelihood function is parametrized with the model parameters and the goal of the method is to find values of those parameters that maximize the sample likelihood. In practice that boils down to solving a set of partial differential equations under the assumption of convexity of the likelihood function. Now, the maximum pseudo-likelihood estimation, instead of using the likelihood function, relies on the pseudolikelihood function, which is an approximation of the joint distribution of a set of random variables [56]. As an example we can write the likelihood function of the PPP with intensity λ as:

$$L(\lambda;\varphi) = \lambda^{N(\varphi)} \exp\Big(- v(\mathcal{W})(\lambda - 1)\Big), \tag{3.10}$$

where φ lies inside the observation window $\mathcal{W}$. In this case, trivially, the maximum likelihood estimator is $N(\varphi)/v(\mathcal{W})$.

Minimum contrast estimation

In general, the minimum contrast method (MCM) seeks parameter values for the analytical expression of a summary statistic C which minimize the difference between C and its non-parametric estimate obtained from the analysed pattern(s) $\widehat{C}$, i.e.:

$$D_\theta = \int_{r_{min}}^{r_{max}} |\widehat{C}(r)^q - C_\theta(r)^q|^p dr \tag{3.11}$$

where C_θ is the theoretical expression for the summary characteristic parametrized with θ, while p and q are simulation parameters.

Mark correlation test

Testing for mark correlation in a pattern requires a two step procedure (see [57]). In the first step we test for *correlation between marks and the local density of points*, while in the second step we test for *correlation between marks of adjacent points*.

In the first step, we test the hypothesis that there exists correlation between marks and the local density of points. To perform the test we link each point of a point process to the area of the closest tile of a Voronoi tessellation of that point process (see [58]). Then, the local point density corresponds to the reciprocal of the area of that Voronoi tile. As a result, each point is assigned a pair of values: the mark and the local point density. We measure the strength of correlation between the two variables using two correlation coefficients [59]: *Pearson's coefficient*, which quantifies the strength of linear correlation, and *Spearman's coefficient*, which quantifies the monotonic relationship between the two variables (in general, this may not be linear). The *Fisher's Z transformation* is then applied to these coefficients to test the hypothesis that the two variables are correlated.

In the second step, we test the hypothesis that the mark of a point is correlated with the mark of an adjacent point. We carry out the hypothesis test using a *mark variogram test* [60, Ch. 4].

The mark variogram is a measure of variability of marks in a marked point pattern over a range of distances, defined as:

$$\gamma(r) = \frac{1}{2}\mathbb{E}\Big[(m_i - m_j)^2\Big], \tag{3.12}$$

where m_i and m_j are mark values at points a distance r apart. To statistically determine whether the variogram shows some correlation (or lack thereof) we use an *envelope test* [52]. The envelope test is an MC test in which the variogram of the real data is compared against bounds calculated from a set of marked point patterns, with mark distributions known to be uncorrelated. We create this set by randomly permuting mark allocations in the pattern obtained from real data. The marks can be considered uncorrelated if the variogram, calculated for the real pattern, is contained within the bounds obtained from minimum and maximum values of the variogram, calculated for the permuted patterns. The significance level of the test depends on the cardinality of the set of permuted patterns.

Eventually, if it turns out that the marks are independent of both local point density and other marks, then they may be considered identically and independently distributed (iid). Alternatively, if there exists correlation between: (i) marks and local point density, then the marks should be modelled as a function of the local point density, where the function is estimated using, for example, the maximum likelihood estimation [61]; (ii) adjacent marks, then geostatistical marking should be applied, wherein the spatial mark distribution is described using a random field with the value of any given mark being the value of the random field at that location.

3.1.4 Stochastic geometry in wireless communications

Application of stochastic geometry in the context of wireless communications enables the study of performance averaged over many spatial realizations of a network, with elements distributed according to some pre-specified stochastic model. As a result, general expressions (possibly even in closed-form) may be derived for various network performance parameters, such as the transmission success probability or the average user data rate, under predefined wireless channel conditions.

Signal-to-interference and noise ratio

The SINR model we use throughout the thesis is defined as follows:

$$\text{SINR} = \frac{P}{W + I} \tag{3.13}$$

where $P = h_x l(x)$ is the power of the signal received by a typical user, i.e., located in the origin $(0,0)$, from a transmitter located in $x \in \Phi$, $I = \sum_{y \in \Phi \setminus \{x\}} h_y l(y)$ is the interference power received at the origin, and W is the noise power. Any received signal is affected by two factors: pathloss $l(x)$ and power fading h_x. The pathloss function $l : \mathbb{R}^2 \to \mathbb{R}_+$ is of the form $l(x) = ||x||^{-\alpha}$, where α is the pathloss coefficient, and the power fading is Rayleigh, with mean corresponding to the transmit power $m \in \mathbb{R}_+$, i.e., $h_x \sim \exp(1/m)$. Note that we will present our results in polar coordinates which are more convenient to handle some of the required calculations for the Euclidean plane.

Coverage probability

We define the coverage probability in a cellular network as the complementary cumulative distribution of the SINR. Informally, this can also be described as the probability that a downlink user achieves a target SINR. Effectively, the coverage probability of a wireless network can be defined as:

$$p(\theta) = \mathbb{P}\Big(\text{SINR} > \theta\Big), \tag{3.14}$$

where θ is the reception threshold at the physical layer (assuming a linear receiver with interference treated as noise). Assuming exponential (e.g., Rayleigh) fading[2], the expression in Eq. (3.14) can be transformed into [62]:

$$p(\theta) = \int_{0^+}^{\infty} \exp(-sW)\mathcal{L}_I(s) f_R(r) dr, \tag{3.15}$$

where $\tilde{l}(r) \equiv l(x)$, $s = \theta/\tilde{l}(r)$, and $\mathcal{L}_I(s)$ is the Laplace transform of interference I; $f_R(r)$ is the distribution of distance to the serving base station R. For a PPP network Φ with intensity λ and the attachment to the nearest base station located at distance r apart from the typical user, the distribution of distance to the nearest base station and the Laplace transform of interference are known [62], with the latter being expressed as:

$$\mathcal{L}_I(\theta r^{\alpha}) = \exp(-\pi r^2 \lambda \mathfrak{Z}(\theta, \alpha)), \tag{3.16}$$

where $\alpha > 2$, r denotes distance to the nearest base station, and $\mathfrak{Z}(\theta, \alpha) = \theta^{\delta} \int_{\theta^{-\delta}}^{\infty} \frac{1}{1+u^{1/\delta}} du$ or, equivalently, $\mathfrak{Z}(\theta, \alpha) = \frac{\theta\delta}{1-\delta} {}_2\mathrm{F}_1(1, 1-\delta; 2-\delta; -\theta)$, where ${}_2\mathrm{F}_1(a, b; c; z)$ is the Gauss hypergeometric function, and $\delta = 2/\alpha$.

Average data rate

We define the average data rate for a typical user as the rate of a user when adaptive modulation and coding is set so that the Shannon bound[3] is achieved for the instantaneous SINR of that user. This leads to the following formal definition:

$$\tau = \mathbb{E}\left[b \log\left(1 + \frac{h_x l(x)}{W + \sum_{y \in \Phi} h_y l(y)}\right)\right], \tag{3.17}$$

where b denotes the spectrum bandwidth. Then, the average user data rate for a network Φ with intensity λ can be expressed as [62][Theorem 3]:

$$\tau = \int_{0^+}^{\infty} f_R(r) \int_{0^+}^{\infty} \exp\Big(- r^{\alpha} W(\exp(\gamma/b) - 1)\Big) \mathcal{L}_{I_r}\Big(r^{\alpha}(\exp(\gamma/b) - 1)\Big) d\gamma dr, \tag{3.18}$$

where

$$\mathcal{L}_{I_r}\Big((\exp(\gamma/b) - 1) r^{\alpha}\Big) = \exp\Big(- \pi \lambda r^2 (\exp(\gamma/b) - 1)^{2/\alpha} \int_{(\exp(\gamma/b)-1)^{-2/\alpha}}^{\infty} \frac{1}{1 + u^{\alpha/2}} du\Big). \tag{3.19}$$

[2]The strategy to obtain the coverage probability with general fading models involves using additional transforms, for example, the Plancherel-Parseval formula. More details can be found in [62, 63].

[3]Real-world mobile systems do not achieve this bound, but to account for this fact would simply require that we rescale our results.

Similarly to the coverage probability, when no noise and $\alpha = 4$ is assumed, this expression can be simplified to [62]:

$$\tau = \int_{0^+}^{\infty} \frac{1}{1 + \sqrt{\exp(\gamma/b) - 1}\arctan(\sqrt{\exp(\gamma/b) - 1})} d\gamma, \tag{3.20}$$

which can also be presented as:

$$\tau = \mathbb{E}\Big[p(\exp(\gamma/b) - 1)\Big]. \tag{3.21}$$

Applications of stochastic geometry to wireless communications

Originally stochastic geometry was used as a performance analysis tool for ad hoc networks, for example, [64–66]. However, during the last few years it has found its application in cellular networks analysis. Due to its tractability, most of the stochastic analysis for single-operator cellular networks was performed based on the assumption of uniformity and, thus, the homogeneous PPP model is being used, for example, [62, 67]. Alternatively, when repulsive inter-point interactions, related to interference management, are accounted for, then the transmission success probability (or coverage probability) was derived based on the Strauss process [68] or the Ginibre process [69]. Clustered processes, due to lack of analytically tractable models, have received much less attention. In [70] bounds on the Laplace transform of the Gauss-Poisson process were derived and subsequently applied to analyze the coverage performance of a cooperative and non-cooperative transmission in a wireless mobile network. In the context of wireless networks, stochastic geometry has been applied to a number of scenarios which include, but are not limited to, the following: heterogeneous networks [71], cognitive networks [72], device-to-device networks [73], or millimeter wave (mmWave) small cell networks [74].

An alternative, and perhaps slightly unexplored, field of application is the network planning, where spatial statistics and stochastic geometry may aid in the process of deciding base station deployment locations and network dimensioning (see [75]).

3.2 Optimization and game theory

In our work we classify decisions made by entities involved in the mobile communications market as either *single-agent decisions* or *multi-agent decisions*. The explicit reference to microeconomics is intentional, as in the models we study the agents are involved in transactions that affect the supply and demand relation. Since it is not our intention to study the efficiency of the resulting market, we formulate market relations only to the extent required to motivate the agents to make technological decisions, such as *what portion of mobile demand in location p to serve through a base station in location l*. From that perspective we are primarily interested in the equilibria and strategies to reach them. For example, we wish to find a set of base stations which minimizes some cost metric, or find an allocation of spectrum to a set of users such that no other allocation yields better performance. Obviously, studying the equilibria means analyzing the characteristics of the resulting network topology, and studying the performance or efficiency measured, for example, in terms of the number of deployed base stations.

3.2.1 Single-agent decision

Single-agent decision-making involves making decisions in a scenario where one central entity collects all the relevant stimuli and processes a unified output. In our case this is a scenario where we observe the efficiency of a single entity (a mobile operator, a mobile virtual operator) operating on a pool of resources aggregated from a number of resource providers. An example of this occurs when a virtual operator signs service level agreements with $\mathcal{N}$ mobile operators (each providing independent mobile service over the same geographical region); if each service level agreement (or even a unit of infrastructure) bears some additional cost for the virtual operator, it is reasonable to assume that the virtual operator will use only a subset of the available resources to control costs.

This type of decision-making typically amounts to modelling a variation of the covering problem [76]. In general, covering models involve a number of customers and a number of facilities which provide coverage to the customers, and a decision that attaches each (or a predefined percentage) of the customers to one or a number of available facilities given certain performance-related constraints. This description can be formulated as a classical optimization problem, the set cover problem (SCP) [77]. The SCP can be formulated as follows:

$$\min_{\{x_s\}} \sum_{s \in \mathcal{S}} c_s x_s, \tag{3.22}$$

subject to

$$\sum_{s \in \mathcal{S}} a_{ps} x_s \geq 1 \quad \forall p \in \mathcal{P} \tag{3.23}$$

$$x_s \in \{0, 1\} \quad \forall s \in \mathcal{S} \tag{3.24}$$

where $\mathcal{S}$ is the set of facilities, $\mathcal{P}$ is the set of customers, c_s is the cost of a facility s, and a_{ps} is the coverage of a customer p by a facility s. The constraint (3.23) ensures that every customer is attached to at least one facility, while constraint (3.24) ensures a binary decision for every facility.

Unfortunately, as it has been shown in [78] covering-type problems belong to the class of NP-complete problems[4] and as such require special treatment, as even for relatively small input sizes finding exact solutions is typically infeasible.

Optimization has been widely used in telecommunications [79]. Specifically, coverage-type problems have been applied to model network planning decisions or base station shutdown schemes (performed to improve the energy efficiency of a network). On the network planning front a significant number of models determine the optimal placement of base stations given certain performance criteria. Herein, we refer to just a few examples, yet a lot more can be found via [79].

One of the fundamental optimization formulations for the network planning problem was proposed in [80]. The authors define a covering problem where base station placements are selected so as to provide sufficient signal quality (under pre-specified interference levels) over a target area with minimum network cost. Similarly, [81] proposes a number of integer linear programming (ILP) formulations to account for the minimization of interference or the number of blocked channels. In

[4] NP-complete problems are a special class of optimization problems, including the travelling salesman problem or the satisfiability problem, which cannot be solved by any known polynomial algorithm, and in case a polynomial algorithm is found to solve one of the problems within the class, then all other NP-complete problems are bound to have a polynomial algorithm too.

order to cope with the processing time inefficiency of solving NP-complete problems, the authors suggest usage of metaheuristic approaches and simulated annealing. The covering-type problem can also be considered from the perspective of service-oriented optimization, in which case it may be formulated as a multi-objective optimization problem, as in [82]. An additional complication to the covering problem is to allow for some level of flexibility in the transmit power allocation [83]. To account for varying interference levels related to different power allocations a number of pre-processing techniques needs to be applied and only heuristic approaches provide any hope of finding quasi-optimal solutions even for a relatively small input.

Where energy efficiency of a cellular network is concerned, it is important to make correct selections as to which base stations should be switched off or where energy-harvesting base stations should be located. In [84] the problem of selecting the smallest set of active base stations that can preserve the quality of service is studied, with the rest of deployed base stations being switched off or sent to energy saving mode. The problem is formulated as an ILP for which a heuristic based on majorization-minimization is proposed. Another way of looking at energy efficiency is through the deployment of energy harvesting units alongside base stations (see [85]). This formulation, again, requires solving a similar covering-type problem for which a greedy heuristic can be applied.

Despite the similarities in formulating the underlying optimization problems, there are still major differences between all of the above mentioned approaches and the selections made as part of Networks without Borders. When modelling those selections, an assumption has to be made as to whether the selections are made based on the existing networks or whether a clean slate approach is taken. The former assumption gives us the benefit of having a priori knowledge of the locations of wireless access points (base stations), which simplifies the problem, likely allowing us to use exact solvers. Subsets of these access points, each of which potentially belonging to a different operator, will have already been optimized (as part of the network planning process) towards meeting certain target coverage, capacity and interference levels. The latter assumption, on the other hand, requires that those locations be found, which means that for reasonable problem instances exact results will be out of reach. However, taking the latter assumption gives us the possibility of studying deployments with arbitrary network densities, and of arbitrary heterogeneity, for example, any mix of macro- and small cells. Importantly, regardless of the assumption taken all our models account for real-world data that we have from the national communications regulators.

3.2.2 Multi-agent decision

What differentiates a multi-agent decision from a single-agent decision is that a decision-making agent (entity) has to account for the existence of a number of other agents, possibly having different objectives and action spaces, making decisions which influence each others' outcomes. In this situation it is often useful to apply game theory, which was designed to study models of conflict and cooperation between rational entities. The first modelling step in game theory is to define a game by specifying three elements:

- the set of players,
- the set of actions available to each player,
- the payoff function associated with each player and each possible outcome at any time moment.

It is important to note, that it is not compulsory to apply game theory to study multi-agent systems, and sometimes[5] a multi-agent system can be translated to a single-agent system where the potential actions of other agents are considered just as a stimuli (input); such a system can then be modelled using optimization techniques.

When sharing is applied to mobile networks, one possible scenario (the one we are primarily dealing with) is the one where agents have a common interest in cooperation, for example, related to cost saving benefits of sharing. This type of multi-agent interaction is typically modelled through cooperative game theory and, when some form of negotiation process is involved, bargaining theory. In this section we focus on describing aspects of the cooperative game theory required to model sharing-type interactions between mobile operators, virtual operators and service providers.

In general, there are two main approaches that one can take towards modelling a cooperative game: the bargaining solutions and the strategic bargaining [86]. In the following we briefly introduce the first concept for the case of two players and we review some relevant state-of-the-art applications.

Let us start by assuming that we have two players, A and B, and a bargaining problem which is a pair (Ω, d), where $\Omega \subseteq \mathbb{R}^2$ and $d \in \mathbb{R}^2$. Ω is the set of possible utilities obtained as a solution to the bargaining, and $d = (d_A, d_B)$ is the disagreement point, i.e., the individual utilities obtained by players if an agreement is not reached. The key idea is that the feasible set of solutions is bounded, meaning that players may disagree, in which case the payoff received is the one related to the disagreement point. However, the preferred outcome is to agree on some other solution, which ensures an incentive to the players for taking part in the bargaining procedure. Obviously, the conflict lies in the particular agreement to be made. One way of describing an agreement is to postulate conditions defining any potential bargaining solution (without defining the strategy to reach it). Indeed, Nash has proposed a set of rational axioms that can describe any bargaining solution. For a bargaining problem $(\Omega, d) \in \Sigma$, the Nash bargaining solution has the following properties [86]:

Axiom 1: *Invariance to affine transformations*: The bargaining outcome is determined by the players' preferences rather than the particular definitions of their respective utility functions. Let f denote positive affine transformation and u^* denote the bargaining outcome of Σ, then the solution to $f(\Sigma)$ is $f(u^*)$.

Axiom 2: *Symmetry*: If the bargaining problem is symmetric, players get the same payoffs.

Axiom 3: *Independence of irrelevant alternatives*: The outcomes of two bargaining problems Σ_1 and Σ_2, where $\Sigma_2 \subset \Sigma_1$ and $u_1^* \in \Omega_2$, is identical, i.e., $u_1^* = u_2^*$.

Axiom 4: *Pareto efficiency*: Let $u^* = (u_A^*, u_B^*)$ denote the bargaining outcome of Σ, then there does not exist other bargaining outcome $u' \neq u^*$ such that $u'_i \geq u_i^*$ for all $i \in \{A, B\}$.

Obviously, this set of axioms is only a subjective way to view a bargaining problem, and a number of objections can be made as to how these axioms fail to capture some dynamics of bargaining situations (see [87]). In addition, several other axioms may be added to the list, with one of the major ones being [87]:

Axiom 5: *Monotonicity*: Each player's outcome should be proportional to its maximum gain. Let Σ_1 and Σ_2 denote two bargaining problems, such that $\Omega_2 \subseteq \Omega_1$, $u_A = u'_A$ for all $u_A \in \Omega_1$ and $u'_A \in \Omega_2$, and $u_B \geq u'_B$ for all $u_B \in \Omega_1$ and $u'_B \in \Omega_2$, then $u_1^* \geq u_2^*$.

[5] There are times when this may not be possible due to inter-dependencies in the decisions made by players; such is the case when mutual best responses are considered.

The last axiom is incompatible with the independence of irrelevant alternatives axiom, yet as we will discuss later, when it is used instead of that axiom, then some other bargaining solutions become available. Given the above axioms we can now seek bargaining solutions that satisfy them, and, indeed, such solutions do exist and some of them are well described in the literature.

Bargaining solutions

In order to find the bargaining solutions we need to make a set of additional assumptions about our bargaining problem (i.e., two-player cooperative game): (i) any contracts made as a result of bargaining are binding, (ii) lotteries are allowed (i.e., mixed correlated strategies), (iii) utilities may be transferred between players. These assumptions determine a feasible set that is closed, bounded from above and convex (by means of mixed correlated strategies). Then, a bargaining solution would simply be a point from this set that satisfies some subset of the defined axioms.

Indeed, there does exist a unique bargaining solution defined for all bargaining problems that follow the aforementioned set of assumptions and this solution will also satisfy axioms 1–4. This solution is precisely the Nash bargaining solution (NBS). Given $(\Omega, d) \in \Sigma$, with $(u_A, u_B) \in \Omega$ and $d = (d_A, d_B)$, the unique solution to the maximization problem:

$$\max_{(u_A,u_B)} (u_A - d_A)(u_B - d_B), \tag{3.25}$$

is the Nash bargaining solution [86].

It is important to understand that the NBS is one possible bargaining solution, which captures the notion of the Pareto efficiency (utility of one player cannot be improved upon without deteriorating the utility of at least one other player) and proportional fairness between players (the player with an upper hand gains more). However, in order to find these solutions one is required to re-define the set of axioms. One such re-definition is related to the, so called, Kalai-Smorodinsky bargaining solution (KSS) [88], which leads to an outcome that accounts for both maximum gains and possible losses made by both players. The KSS is the unique outcome of the bargaining problem when axioms 1,2,4 are considered and the third axiom is replaced with axiom 5. The formal definition of the KSS is made easier if we define player's preference function [89]:

$$v_A = u_A - d_A + \alpha(r_B - u_B), \tag{3.26}$$

where r_i corresponds to the maximum possible outcome for player $i \in \{A, B\}$, and $\alpha \in \{-1, 0, 1\}$ is the weighting factor. Now, different settings of α will correspond to different bargaining solutions, i.e., $\alpha = 0$ will correspond to the NBS, $\alpha = 1$ to the KSS, and $\alpha = -1$ to the utilitarian bargaining solution [90]. Then, given some α, each of the solutions can be found as a unique point maximizing the following function:

$$\max_{(v_A,v_B)} v_A v_B. \tag{3.27}$$

Therefore when defining their cooperation players need to agree on a set of axioms and the cooperative solution will be the determined through optimization of the function in Eq. (3.27) (at least for the three aforementioned solutions).

The axiomatic bargaining framework has been applied in wireless communications research to date to study rate regions of interference channels [91], dynamic spectrum access [92], resource

allocation in single-operator networks [93], resource allocation in shared networks [90], or cooperation among wireless service providers [94].

4 Multi-operator patterns in wireless mobile networks

Multi-operator patterns in wireless mobile networks

"In that Empire, the Art of Cartography attained such Perfection that the map of a single Province occupied the entirety of a City, and the map of the Empire, the entirety of a Province."

- Jorge Luis Borges, *On Exactitude in Science, Collected Fictions*

Modelling and simulating the performance of shared mobile networks is of high importance to the research community and even more so to the cellular networks industry, as the fundamental gains and limitations of such shared networks are still an open question. One key reason for this state of matters is the lack of accurate models that are representative of radio access network deployments made by multiple mobile operators. In this chapter we make a contribution to the state-of-the-art in mobile network modelling by providing a quantitative analysis of stochastic models that may be used to represent shared mobile network deployments. This chapter is largely based on our two articles [35, 57].

4.1 Introduction

A new mobile network roll-out involves three phases [95]: dimensioning, planning, and optimization. In the first phase, a rough estimate of the network layout and elements is prepared based on technical demands (such as an area to provide coverage and capacity for), forecasted subscriber base and service usage characteristics [96]. Next, network element configuration is performed, whereby base station site equipment is selected based on power budget calculations, followed by detailed topology planning, which includes coverage and capacity planning based on pathloss information and prediction models (such as the modified Hata model), and dynamic models provided by commercial network planning and optimization tools [97]. In addition, for 3^{rd} generation (3G) and 4^{th} generation (4G) systems, mobile traffic information is taken into account as it affects capacity requirements and acceptable interference levels. At this stage the cellular network is designed and ready to be deployed. During the deployment the positions of base station sites, which have been established according to the pre-specified models, become distorted by a number of non-technical aspects, such as [97]: site acquisition (availability of a given site), site legalization (license for radio emission from a given site), and site preparation (installation and availability of mast, power supply, air conditioning system, etc.). Moreover, the incumbent mobile network operators, as a natural cost minimization strategy, will tend to reuse existing sites [95]. Effectively, even initial

coverage provisioning deployment may be allocated in a way that is far from idealistic lattice-like scenarios which are typically used for mobile network performance analysis. Furthermore, as the service demand grows, and large capacities become essential, a microcellular topology is deployed, typically characterized by much lower transmit powers and rooftop level antennas, with additional pico-/small-cell systems to extend the capacity to indoor offices or places like shopping malls.

Despite these practical insights on the cellular network roll-out process, there seems to be no consensus on how to model and analyse the resulting base station deployments. In the studies to date the choice of a model is typically dependent on the type of tool applied, either stochastic geometry or computer simulation.

Stochastic geometry allows us to derive analytical expressions for various performance parameters, such as outage probability or data rate, under predefined wireless channel conditions, averaged over many realizations of a pre-specified stochastic model. For analytical tractability, much of the existing work modelling radio access network deployments (e.g., [62, 67, 98, 99]) relies on the homogeneous Poisson point process (PPP) model, which assumes no correlation in the analysed patterns. The usage of the PPP model can be analytically justified under sufficiently high log-normal shadowing, see [100]. However, spatial statistical analysis of the available mobile network deployment data to date suggests that single-operator mobile networks can be better represented using regular processes. For example, in [68] the fitted Strauss process (SP) turns out to provide the closest match to a single-operator mobile network deployment. In [101] the determinantal point process (DPP) is found to accurately represent regularities present in single-operator base station deployments. Analytical insights into these point processes are limited to some specific cases (see [69]). Yet, heuristic approaches, such as the one in [102], may be applied. Clustering in real mobile networks, as observed in large-scale single-operator deployments [75], can be modelled using cluster processes. However, analysis of these processes seems to be even harder and some level of analytical tractability has been achieved only for certain classes of cluster processes (see [70]).

Computer simulation-based approach typically relies on scenarios and models defined and published by standardization bodies, such as the 3rd Generation Partnership Project (3GPP), as a part of technical reports, for example, [103]. In those scenarios macro- and microcell base station deployments are modelled using lattices, most frequently adopting a hexagonal lattice, which creates repetitive structures to simplify inter-cell interference management [95], and "ensure optimal coverage over a given area" [97]. Even if initial deployments were performed according to a lattice-based model, the evidence to date shows that the existing networks are far from these idealistic models. Hence, to account for real-world irregularities a constant deviation may be added to a lattice-based model [97]. However, these perturbed lattices, as they are called, provide less satisfactory results than some other stochastic models (see [68, 101]). It is worth mentioning that simulations based on a hexagonal lattice model predominate in the research to date which attempts to evaluate the performance of wireless network sharing (see [104–107]).

The remainder of the chapter is structured as follows. First, we identify point processes that best describe base station deployments made by multiple mobile operators. Subsequently, we seek a way to model downlink transmit power allocation in networks composed of either single-operator deployments or multi-operator deployments. We apply spatial statistical analysis to quantify patterns available in real-world mobile networks in three European countries, and perform a series of fitness tests, which give us statistical confirmation of the correctness of the selected model(s).

4.2 Datasets

The aim of our study is to find stochastic models that characterize multi-networks aggregated for the purpose of infrastructure sharing. The input data we use consists of either real base station information or radio license information (for simplicity, we will assume that an assigned radio license is equivalent to a base station deployed in the given location), which includes spatial coordinates of the base station sites and unique cell identifiers, for Groupe Special Mondiale (GSM) and Universal Mobile Telecommunications System (UMTS) technologies in three countries: Ireland, Poland, and the United Kingdom. The data was extracted from publicly available information collected by the national telecommunications regulators. Each of the regulators collects data supplied by the local mobile network operators (MNOs), who also ensure accuracy and keep the information updated.

The Irish database contains base station site information updated by the Irish MNOs and maintained by the Irish communications regulator ComReg[1]. Based on the quarterly report from ComReg (Q4 2013), Ireland had four major MNOs (Vodafone, eircom Group Mobile, O2 and Three) together sharing 93.9% of the Irish mobile market[2]. We deem the Irish dataset as the most reliable, as ComReg requires mobile infrastructure providers to report all the masts and structures that contain "mobile base stations". Therefore, we use the Irish dataset to confirm the reliability of the other two datasets. The database for Poland contains spatially allocated radio license information from all MNOs operating in the Polish mobile market, available through the national regulator UKE[3]. By 2013 there were four major MNOs in Poland, namely Orange, Play, Plus and T-Mobile, which had a combined 98.7% of the subscriber share and 99.7% of the market revenues. The database for the UK was compiled by Office for Communications (Ofcom) based on base station site information voluntarily supplied by the local MNOs[4]. As of 2012 there were five major local MNOs: O2, Orange, T-Mobile, Three, and Vodafone, all of them together having a higher than 90% market share.

From each of the databases we have selected a number of city areas. For each city area we have focused on base station site information for two radio technologies: GSM (900 MHz) and UMTS (2100 MHz), and macro- and microcells only (if the transmit power or cell radius information was available). For each city area we have extracted this information for an observation window of size 10-by-10 km with arbitrarily selected central locations[5].

In order to build a model for downlink transmit power allocation we use the Ofcom database, which in addition to base station site location contains also information on maximum downlink transmit power, and operating frequency band. From the database we have extracted base station location and downlink maximum transmit power (in dBW) for the 3G technology for each of the four major mobile operators (O2, Orange, Three and Vodafone) in a number of city areas that represent urban scenarios: Sheffield, Edinburgh, Manchester, Liverpool and Leeds. Each of these cities has a population of $500,000 - 1,000,000$ residents with a sizeable urban topology and several hundred base stations. Similarly to the analysis of base station patterns we use an observation window of size 10-by-10 km with arbitrarily selected central locations.

[1] `http://www.askcomreg.com/mobile/mobile_sites%3b_base_stations_and_masts.36.LE.asp`
[2] Since we made our analysis, two mobile operators (O2 and Three) have consolidated their networks.
[3] `http://www.uke.gov.pl/pozwolenia-radiowe-dla-stacji-gsm-umts-lte -oraz-cdma-4145`
[4] `http://stakeholders.ofcom.org.uk/sitefinder/sitefinder-dataset/`
[5] In order to test whether our results are affected by the choice of the observation window we also looked at observation windows of size 2.5-by-2.5 km, and 5-by-5 km. In the main, the results are consistent across those three observation windows.

Table 4.1: Basic information on the extracted point patterns

Dataset	Name	Center (longitude;latitude)	Number of BSs (UMTS/GSM)	Inter-operator co-location (UMTS/GSM) [%]
Ireland	Dublin	53.3478°N;6.2597°W	447/409	5.0/4.0
Poland	Kraków	50.0614°N;19.9383°E	348/273	8.0/8.0
	Poznań	52.4°N;16.9167°E	400/362	10.0/14.0
	Warszawa	52.2333°N;21.0167°E	812/916	12.0/8.0
UK	Birmingham	52.4831°N;1.8936°W	281/144	5.0/0.0
	Leeds	53.7997°N;1.5492°W	238/115	16.0/0.0
	Liverpool	53.4°N;3°W	266/122	7.0/0.0
	London	51.5072°N;0.1275°W	1442/1513	3.0/1.0
	Manchester	53.4667°N;2.2333°W	303/155	7.0/0.0

In each case the choice of the observation window was driven by our desire to capture features of base station deployments at local scales, and to correctly assess environmental forces that underlay point distributions for multi-network deployments. In our study, we have simplified each co-location to a single base station site, which is a reasonable working assumption given that inter-operator co-location accounts on average for only 10% of all base stations. Tab. 4.1 summarizes basic information on all the selected locations.

Since the analysis covers different geographical areas as well as various statistical metrics, for clarity we have decided to focus our presentation on the study of Dublin. The results for other areas (listed in Tab. 4.1) are substantially the same and they are included in the estimated point process model, which we discuss at the end of this chapter.

4.3 Methodology

4.3.1 Base station deployment model

An elementary step in any spatial statistics analysis is the qualitative assessment, which builds rationale for further modelling decisions. Fig. 4.1 depicts four patterns representing two-operator base station deployments using: a superposition of two realizations of the hexagonal lattice model (Fig. 4.1(a)), a superposition of two realizations of the Poisson point process (Fig. 4.1(b)), a realization of a clustered point process (Fig. 4.1(c)), with operator marks assigned uniformly at random, and a real base station deployment in the Dublin city area (Fig. 4.1(d)). Clearly, we can observe a high degree of regularity with the hexagonal deployment and uniformity with the PPP deployment. The clustered deployment results in a number of smaller clusters distributed throughout the area of interest and relatively many empty spaces, in comparison to the previous two realizations. Finally, the pattern representing real deployment is clustered around the central area with blank areas forced by environmental obstacles, and somewhat uniform or repulsive behaviour outside the central area, possibly caused by frequency and coverage planning as well as low availability of shared base station sites. Qualitatively, one can conclude that the first two point patterns and the real one come from distinct families of point distributions. The relationship between the clustered pattern and the empirical pattern is less intuitive, and, as we will show later, the realizations of some of the clustered point processes do indeed convey the features of multi-operator real base station deployments.

A desirable model is one that is easy to interpret, estimate and simulate, while still preserving features of the observed base station patterns. Subsequently we will fit PPP, SP, Matérn cluster

(a) Hexagonal (b) PPP (c) LGCP (d) Real

Figure 4.1: Illustration of different types of two-operator base station deployment patterns for the city of Dublin (the Dublin city area is represented here as a superposition of the smallest unit areas from the Irish census 2011, and located within a radius of 5 km from the city's center), generated from: a) a hexagonal lattice, b) a Poisson point process, c) a clustered point process, and d) real data. The contours represent the Dublin city area, whereas the dots and triangles stand for base station site locations of two different operators. The real base station pattern is clustered in the central area, with irregular empty spaces forced by environmental obstacles and, most likely, coverage planning. Qualitatively, the clustered point process realization exhibits the closest match to the real base station pattern, with a number of smaller clusters distributed throughout the area of interest and relatively many empty spaces, while the Poisson point process pattern rather uniformly fills the Dublin area and the hexagonal pattern exhibits high regularity with little resemblance to the actual base station deployment.

process (MCP), Thomas process (TP) and Log-Gaussian Cox process (LGCP), with exponential and stable correlation functions, to the multi-network patterns under study. To fit the first two we will use the maximum pseudolikelihood method, described in Chapter 3. When the pseudolikelihood function is not easily tractable, which is typically the case with clustered point processes, we apply the minimum contrast method (MCM), described also in Chapter 3. Applying the MCM to the analysed processes is natural, as for each of those processes we can obtain closed-form expressions for the n-th order summary characteristics. Specifically, we use the pair-correlation function, for which closed-form expressions are given in Tab. 3.1. Setting $p = 2$, and $q = 1$ (for the MCP and TP), allows us to obtain closed-form expressions for the parameters of the fitted processes, which we summarize in Tab. 4.2. We derive these expressions in Appendix A.

Table 4.2: MCM expressions for model parameters

Model	Model parameters		
	I	II	III
LGCP	$\widehat{\mu} = \log(\widehat{\lambda}) - \widehat{\sigma}^2/2$	$\widehat{\sigma}^2 = \left(\frac{B_L(\beta)}{A_L(\beta)}\right)^{1/q}$	$\widehat{\beta} = \operatorname{argmax}\left(\frac{B_L(\beta)^2}{A_L(\beta)}\right)$
MCP	$\widehat{\mu} = \frac{\widehat{\lambda}}{\widehat{\kappa}}$	$\widehat{R} = \operatorname{argmax}\left(\frac{B_M(R)^2}{A_M(R)}\right)$	$\widehat{\kappa} = \frac{A_M(R)}{B_M(R)}$
TP	$\widehat{\mu} = \frac{\widehat{\lambda}}{\widehat{\kappa}}$	$\widehat{\sigma}^2 = \operatorname{argmax}\left(\frac{B_T(\sigma^2)^2}{A_T(\sigma^2)}\right)$	$\widehat{\kappa} = \frac{A_T(\sigma^2)}{B_T(\sigma^2)}$

In our analysis we look at a range of r values, where $r_{min} = \min_{x_i, x_j \in \varphi : i \neq j}\{||x_i - x_j||\}$ (as suggested in [54]) and r_{max} is set as the quarter of the side-length of the square observation window, which, according to our simulations, provides a good balance between observation of the local features and computational effectiveness. We perform our analysis using open-source statistical computing software R [108], and, in particular, we use functions from the *spatstat* library [109]. For the goodness-of-fit tests we use the envelope test, which can be interpreted as the significance test [52], with a significance level in a one-sided test equal to $1/(k+1)$, where k is the number of trials. For the envelope test we utilize both the L-function (second-order statistic) and the empty space function (first-order statistic), as suggested in [52] to minimize the bias caused by testing our model

against the data that was used to estimate model parameters.

4.3.2 Downlink transmit power allocation model

Qualitative inspection of our data allows us to conclude that the majority of mobile operators have base stations with transmit powers in the 20-30 dBW range, which implies two types of deployments: macro- and microcell. For some operators, however, we observe also small-cell deployments, where base station transmit power is below 20 dBW. These, however, occur only for Orange and Three and seemed, at the time the data was collected, to be at an early stage of deployment, as their number is significantly smaller than the number of available macro- and microcell base stations. For this reason we restrict our investigation to the macro- and microcells in the 20-30 dBW transmit power range.

Following the methodology laid out in Chapter 3, we first perform the mark spatial correlation test and, subsequently, the adjacent mark correlation test.

First, for each of the real base station deployment patterns we evaluate the correlation between local base station density and the transmit power by calculating the correlation coefficients described in Chapter 3. Using the *Fisher's Z transformation* these coefficients are converted into standard normal variables and the mark correlation hypothesis is tested for a significance level of 0.01. We report that, with the exception of the *Pearson's coefficient* for the O2 network in Manchester, all of the correlation coefficients fail the correlation hypothesis test. The results show that there is no correlation between the transmit power of a base station and the density of its neighbours, with the O2 Manchester network being an outlier. Fig. 4.2 shows an example of the typical distribution of transmit powers relative to local base station densities. The scatter plot depicts the transmit power values in the dataset and the local densities of their associated base stations. The natural logarithm of the local densities is used to more evenly distribute the values as we observe more occurrences of lower density values than high-density ones. Visual inspection of the scatter pattern confirms the result of the hypothesis test: the points show no indication of correlation to base station deployment density.

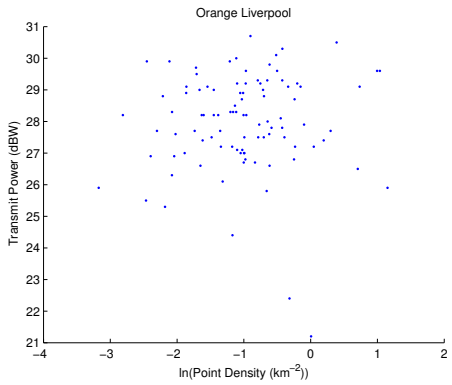


Figure 4.2: Illustrative mark spatial correlation test.

Second, we calculate the variogram for the four mobile networks in the five cities, with inter-point distance r varied between 0 and 2.5 kilometres. We use 99 Monte Carlo (MC) trials, corresponding to a significance level of 0.01, to generate uncorrelated mark distributions for the variogram hypothesis test envelope. Fig. 4.3 shows the variogram for the Vodafone network in Edinburgh. We report

that all the tested datasets fail the hypothesis test by showing a variogram value inside the envelope bounds for all values of r. As such we report that the tested networks do not display correlation in the transmit power of neighbour base stations, including the O2 Manchester network that failed the mark spatial correlation test in the previous section.

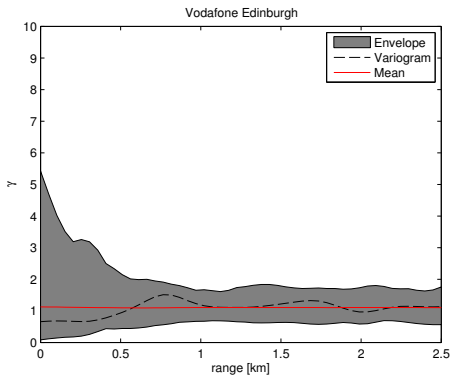


Figure 4.3: Illustrative variogram test.

In summary, our study shows low or no correlation in micro- and macrocell transmission power allocations in 3G networks. While acknowledging that lack of correlation does not imply independence, we believe it is reasonable in practice to *model* such power allocations as being identically and independently distributed (iid).

4.4 Statistical results

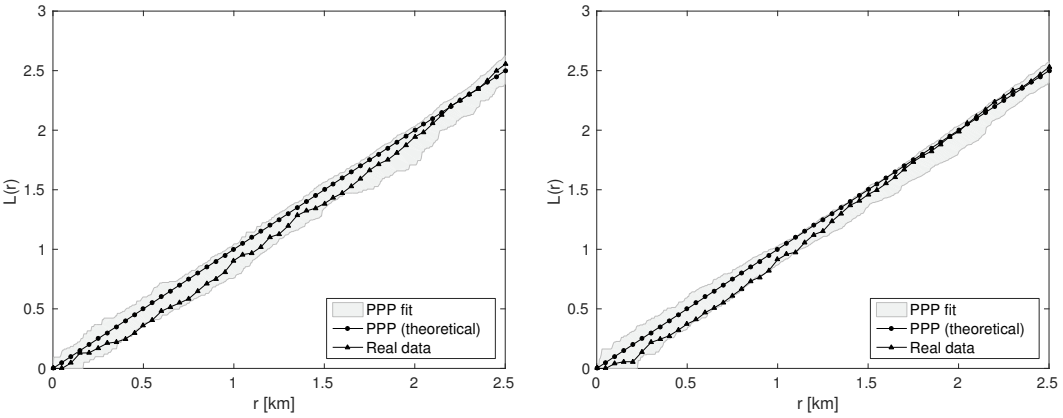
4.4.1 Base station deployment analysis

Single-operator analysis

Before we turn to the multi-network analysis, we investigate the L-function results for single-networks to confirm observations made in [68]. In Fig. 4.4 we observe the L-function results for selected GSM and UMTS deployments, along with 99 realizations of a Poisson fit (corresponding to a significance level of 0.01), and its well-known closed-form representation. The results somewhat confirm the observations made in [68], as the observed point patterns for various operators in different areas exhibit either some mild forms of inhibition or uniformity for the whole range of inter-point distances.

Cross-correlation analysis

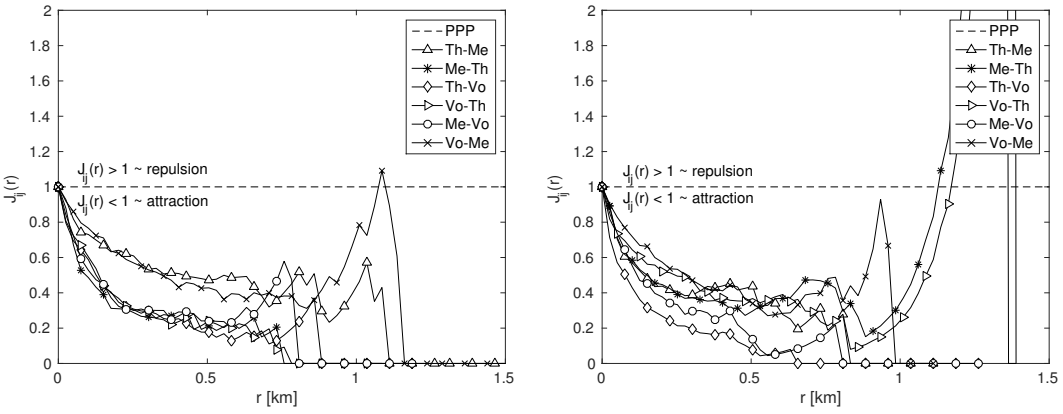
We now turn our attention to the relationship between multi-operator deployments. If single-networks were realizations of independent Poisson processes, we would expect that the observed multi-network patterns should not exhibit any forms of regularity. However, as noted previously the resulting multi-network point patterns clearly exhibit clustering. This fact and the single-operator fit would suggest that the realizations of networks of different operators are not independent from each other, which also makes sense intuitively, as one would expect that base station placements account for local phenomena, such as high availability of potential subscribers. In order to test this hypothesis, we look into the cross-type J-function results for multiple operators in Dublin. Since the



(a) GSM (b) UMTS

Figure 4.4: The L-function results for Meteor GSM and UMTS deployments in Dublin (marked with triangles) with the envelope of the fitted Poisson model at a significance level of 0.01 (grey area), and the theoretical value for the Poisson process (marked with circles).

J-function involves comparison between different patterns, we drop for a moment the assumption of simplicity and allow for base stations of different operators to co-locate. Both the results for GSM (Fig. 4.5(a)) and UMTS (Fig. 4.5(b)) confirm that there is significant clustering for the majority of inter-point distances, with some repulsive behaviour for Three-Meteor/Meteor-Three and Vodafone-Meteor pairs at larger distances, which may reflect the structure of the Irish mobile market and its existing infrastructure sharing among operators. Note that these results, as any results in empirical statistics, are merely an indication of clustering among the observed base station patterns rather than a formal proof.



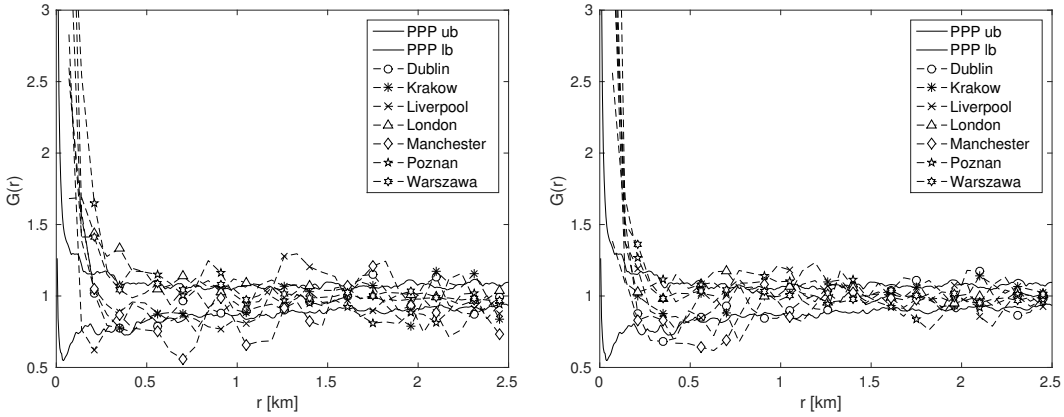
(a) GSM (b) UMTS

Figure 4.5: The cross-type J-function results for combinations of pairs of major MNOs in Dublin: Meteor (Me), Three (Th) and Vodafone (Vo). Recall that by definition J_{ij} is not symmetric and also $J_{ij} < 1$ indicates clustering (while $J_{ij} > 1$ inhibition) between the base station locations of any two MNOs.

Pair-correlation analysis

Let us now move on to describing the results of fitting a model to multi-networks. As stressed previously, we present in more details the results for Dublin, while noting that the fitting results for the remaining cities follow similar patterns, as can be seen in Fig. 4.6. In the figure we can see that

real multi-networks are characterized by significant clustering that occurs at inter-point distances below approximately 500 meters, beyond which point the patterns become somewhat uniform, with spikes of regularity or clustering occurring as a result of some unobserved local heterogeneities, most likely related to social and geographical features of the analysed areas. This result indicates that multi-networks in urban areas show similar behaviour, which gives us some confidence in the reliability of our input datasets and suggests that a single model could be potentially representative to a class of real multi-networks.



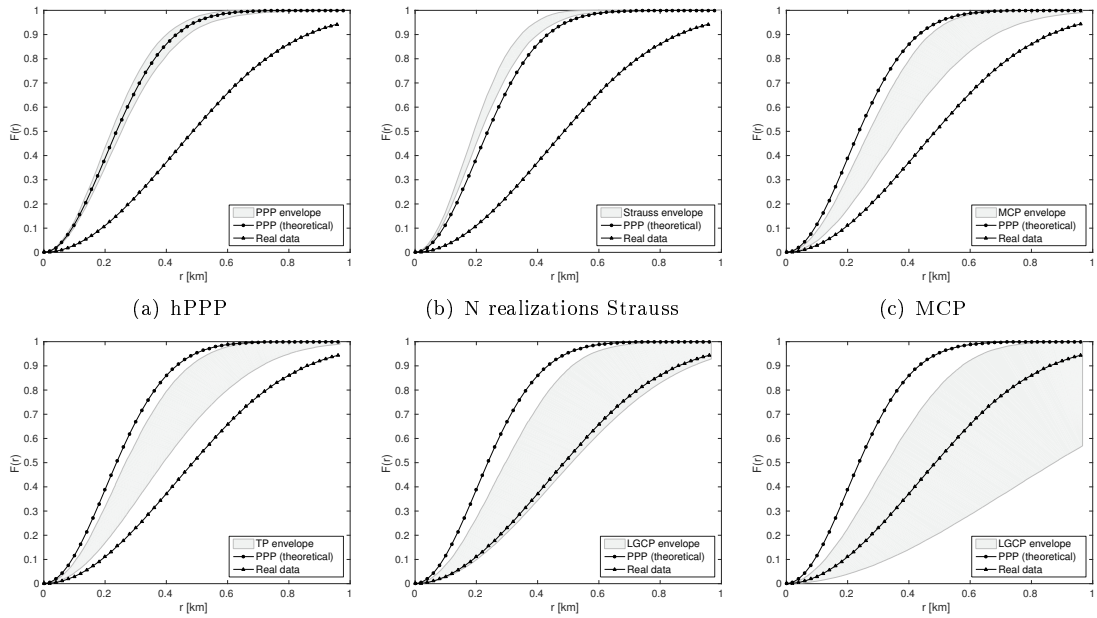
(a) GSM (b) UMTS

Figure 4.6: The pair-correlation function for real deployments in various cities; Dublin marked with triangles and the envelope of the fitted Poisson model (at a significance level of 0.01) marked with solid lines.

From Fig. 4.6 it is quite clear that any point process capable of modelling multi-networks would need to exhibit short-range clustering and be robust enough to account for some deviations in local clustering and repulsion of points at larger distances. The model that intuitively is the most promising is the log-Gaussian Cox process, which allows us to parameterize and estimate the unobserved local covariates in the form of a random field spanning the analysed area. Since our goal is to find a point process that represents features of multi-networks, we look also at other clustered point processes, namely the Matern cluster process and the Thomas process. In order to show that multi-network patterns have strong clustering features, which cannot be captured by processes describing single-networks, we also provide fitting results for other point processes: the PPP (to disprove the complete spatial randomness hypothesis [52]), and the superposition of independent SP realizations. The figures also include envelopes, which represent the maximum and minimum of 99 realizations of the fitted model, which are used to assess the goodness-of-fit of the model in the envelope test [52].

First and second order statistics test

As we can see from Fig. 4.7 and Fig. 4.8 both the PPP and the superposition of independent realizations of the SP can be rejected as candidate models, since they do not produce sufficiently large empty spaces, both in the GSM and UMTS case. Moreover, these point processes are unable to model short-range clustering as can be seen from the results obtained for the L-function in Fig. 4.9 and Fig. 4.10. Doubly Poisson clustered models (MCP and TP), seen in Fig. 4.7 and Fig. 4.8, do not pass the envelope test for the F-function, as they under-represent the occurrence of empty spaces in the patterns. Moreover, when second-order characteristics are taken into account, these point



(a) hPPP (b) N realizations Strauss (c) MCP

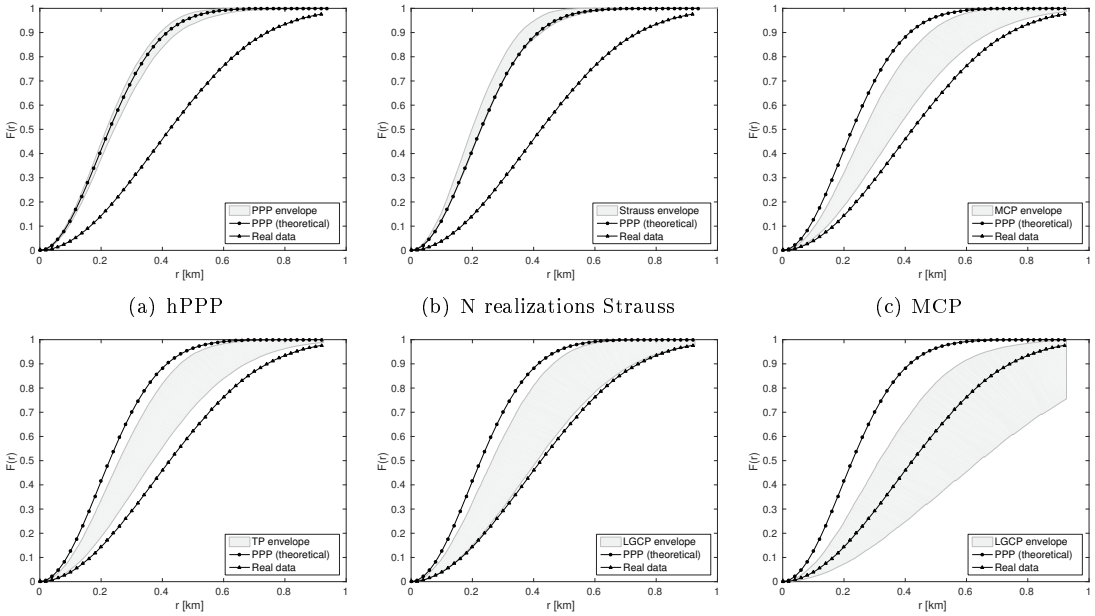
(d) TP (e) LGCP exponential (f) LGCP stable

Figure 4.7: GSM - Empty space function (F-function). The estimated F-function for the multi-network in Dublin (marked with triangles) with the envelope of the fitted point process model at a significance level of 0.1 (grey area), and the theoretical realization of the PPP (marked with circles). As we can see from the figure the "LGCP exponential" and "LGCP stable" fitted models pass the envelope test for Dublin.

processes seem to insufficiently capture clustering, while completely missing out on the long-range repulsion, see Fig. 4.9 and Fig. 4.10.

By far the most satisfying results we have observed come from the LGCP models, which in the UMTS case passed the F-function envelope test for both models, see Fig. 4.8, and for the GSM case passed the F-function test for the model with the stable correlation function, see Fig. 4.7. Apparently, the trade-off for achieving better fitness is related to increased variance in the occurrence of empty spaces prevalent in the modelled patterns. When it comes to second order statistics, we can see in Fig. 4.9 and Fig. 4.10 that both LGCP models are able to account for the short-range clustering, while also being able to capture long-range repulsion in the patterns. In our simulations we have also tried other than the exponential and stable spatial correlation functions (for example, Gauss or spherical); however, fitting against a model based on either of the two functions gave us the most satisfying results.

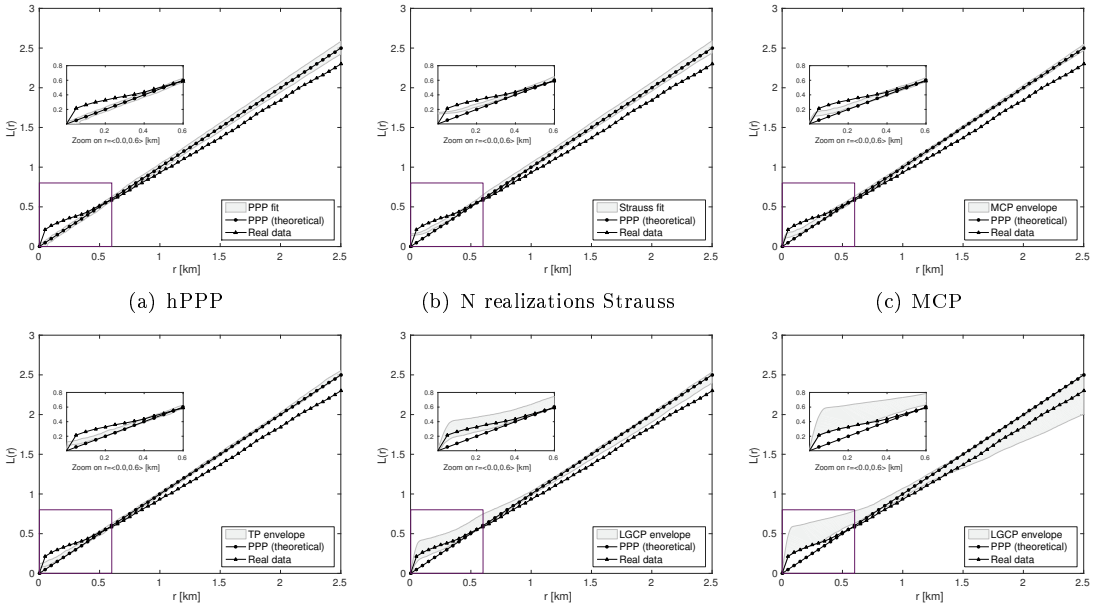
When fitting the models to other patterns from our dataset, we have observed similar results, i.e., we could model the short-range clustering with any of the clustered point process, while the long-range repulsion was best represented with the LGCP model. For some areas, however, the LGCP was insufficient to capture all the inter-point interactions; in those cases, in order to achieve fits one could adjust the intensity function of the process with some additional environmental covariates or inject an inter-point interaction force into the density function of the LGCP model. Both of these, however, destroy any potential analytical tractability of the model, with the risk of over-fitting, which goes against our goal of model universality. In addition to a number of urban cases, we have also looked at some areas with lower density of base stations, such as Athlone, Ireland. In order to obtain meaningful results for rural areas we had to increase the observation window size to 25-by-25 km. However, even increasing the observation window left us with relatively small patterns



(a) hPPP (b) N realizations Strauss (c) MCP

(d) TP (e) LGCP exponential (f) LGCP stable

Figure 4.8: UMTS - Empty space function (F-function). The estimated F-function for the multi-network in Dublin (marked with triangles) with the envelope of the fitted point process model at a significance level of 0.1 (grey area), and the theoretical realization of the PPP (marked with circles). As we can see from the figure "LGCP stable" passes the envelope test for Dublin.

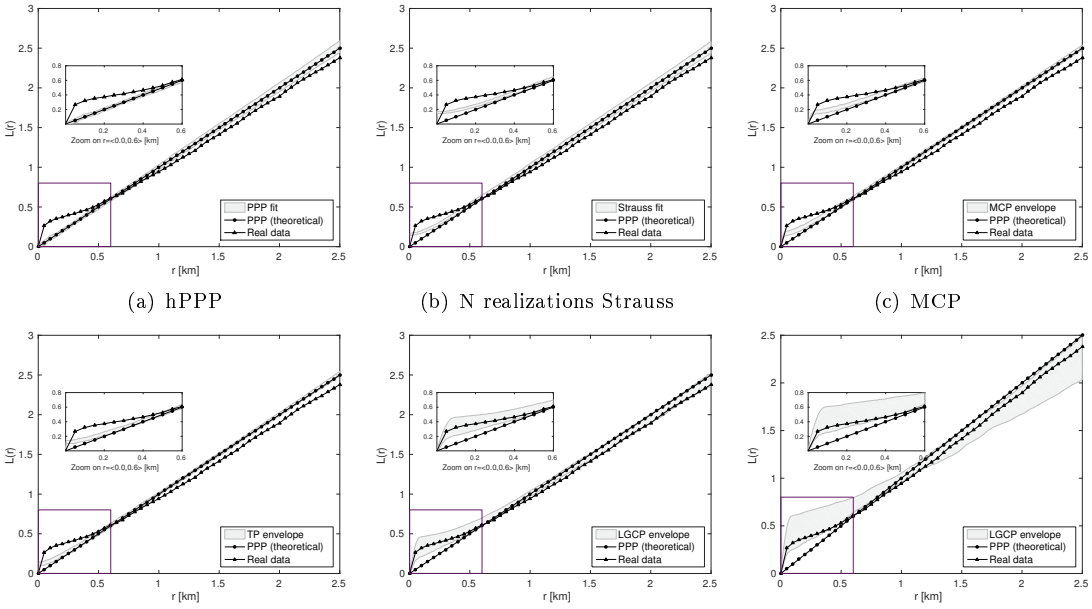


(a) hPPP (b) N realizations Strauss (c) MCP

(d) TP (e) LGCP exponential (f) LGCP stable

Figure 4.9: GSM – the L-function. The estimated L-function for the multi-network in Dublin (marked with triangles) with the envelope of the fitted point process model corresponding to a significance level of 0.01 (grey area), and the theoretical realization of the PPP (marked with circles).

that contained significant amount of empty spaces. In those patterns we could still observe some mild forms of clustering at smaller distances, which in most cases was enough to reject the PPP hypothesis. The considered clustered point processes were able to capture this clustering as well as



(a) hPPP (b) N realizations Strauss (c) MCP

(d) TP (e) LGCP exponential (f) LGCP stable

Figure 4.10: UMTS – the L-function. The estimated L-function for the multi-network in Dublin (marked with triangles) with the envelope of the fitted point process model at a significance level of 0.01 (grey area), and the theoretical realization of the PPP (marked with circles).

the following uniformity in the pattern, although the fitted models had significant variance. And since these results were not entirely consistent across the other rural areas we have analysed, we have decided to focus on modelling urban locations.

Summary

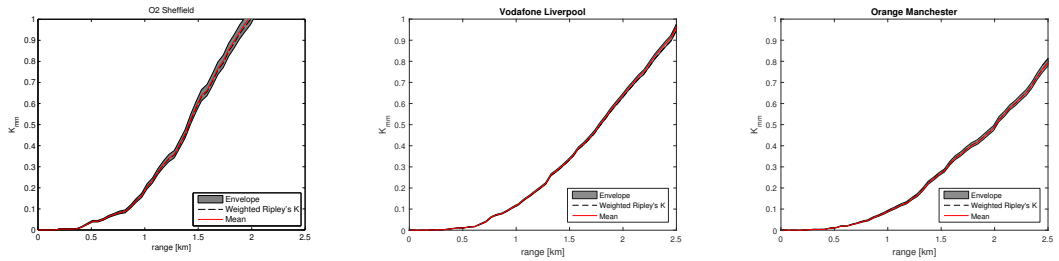
Based on the fitness results we have concluded that the most promising point process to model multi-networks in urban areas is the LGCP. Furthermore, we have observed that we can almost unambiguously classify the parameters of the models according to the type of deployment, i.e., urban, and dense urban. Tab. 4.3 summarizes the two classes of models we have found. It is interesting that for the urban scenario the standard deviation is relatively low (around 10%) for both shape parameters, i.e., σ^2 and β, which suggests similar shape of the intensity function $\lambda(x)$ for each of the geographic areas considered, with only scale parameter μ (precisely $\exp(\mu)$) having higher deviation to account for the different number of base stations in each area. For the two dense urban areas (London and Warsaw) the estimated model parameters differ significantly, which is presumably the result of much higher number of base stations and much richer structure of the area (higher concentration of base stations, larger number of neighbourhoods with concentration of base stations).

Table 4.3: Fitted model parameters

Location type	Model	Model parameters		
		σ^2	β	μ
Urban	LGCP	3.9040 ± 0.3699	0.03 ± 0.0031	-0.5634 ± 0.2956
Dense Urban (London and Warszawa)	LGCP	$2.0561, 2.7228$	$0.054, 0.0288$	$1.2665, 1.9477$

4.4.2 Downlink transmit power allocation analysis

Having demonstrated that the real-world networks have uncorrelated maximum downlink transmit powers we now carry out a goodness-of-fit test to determine if a probability distribution function (pdf) fitted to the mark distribution is feasible as a transmit power modelling tool. The goodness-of-fit test is based on an envelope test using the mark-weighted Ripley's K summary characteristic, as described in Chapter 3. The pdf of choice is a truncated normal distribution with parameters obtained via maximum likelihood fitting of the dataset mark space $\mathbb{R}_+$ using the MASS library in R [110]. Fig. 4.11 shows the envelope of the goodness-of-fit test for some selected areas.



(a) O2 Sheffield (b) Vodafone Liverpool (c) Orange Manchester

Figure 4.11: Three examples of the goodness-of-fit test for the power allocation model based on random assignment of downlink transmit powers from the truncated normal distribution.

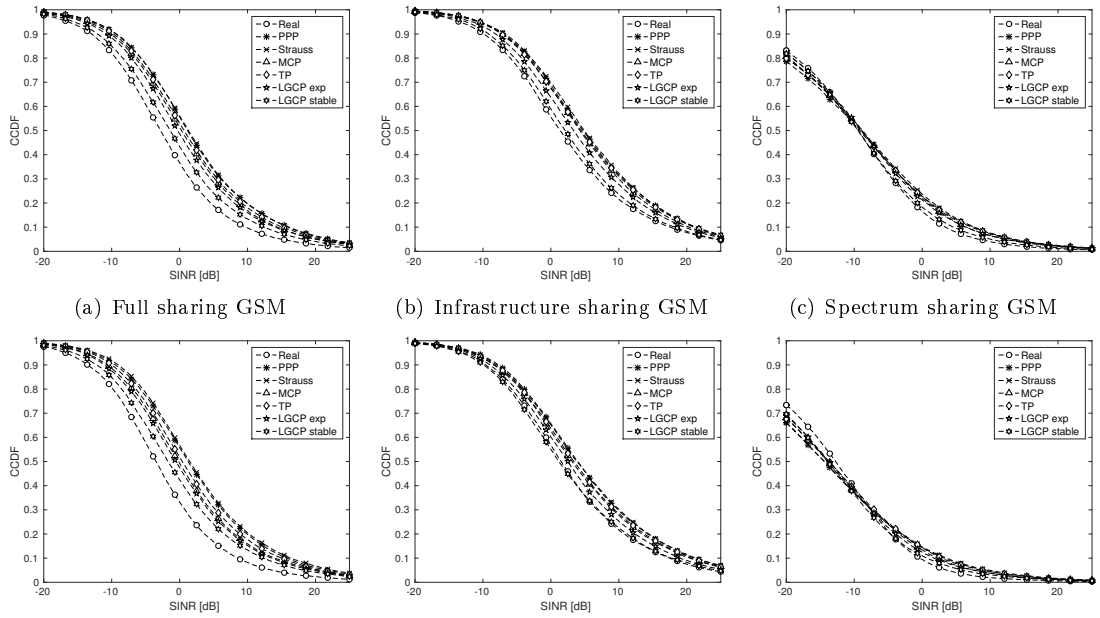
We report that a truncated normal distribution passes the goodness-of-fit envelope test for all operators across the five cities. Furthermore, the transmit power deployments for the O2, Orange and Vodafone operators can be successfully replicated using a single fitted model. Our results suggest that a truncated normal distribution with a mean μ of 28.03 dBW and a standard deviation σ of 1.41 dBW produces transmit power deployments which pass the envelope test across all cities for the three operators.

4.5 Wireless network relevant analysis

Having evaluated the statistical match between our proposed point process models and real multi-network deployments, we now turn our attention to wireless network-relevant evaluation based on the radio access network sharing scenarios of: infrastructure, spectrum and full network sharing, as studied in [99]. As a metric for our evaluation we use the *coverage probability* and the signal-to-noise and interference ratio (SINR) model as described in Chapter 3. We obtain our numerical results using MC simulations, and making the following assumptions: (i) worst-case interference is considered, i.e., all transmitters transmit simultaneously; (ii) we assign individual base stations to operators uniformly at random; (iii) noise is unitary for all the receivers; (iv) $\alpha = 4$; (v) we generate 2000 user locations, over which we average our results. We also apply edge-effect correction based on an enlarged window, i.e., we assume that the user locations lie inside a square window centred at the analysed area with the side equal to half the side of the window containing the original area.

4.5.1 Coverage probability - unitary transmit power

Fig. 4.12 presents the coverage performance of our models for each of our sharing scenarios and technologies when base station transmit power is unitary. In each case the LGCP model with the stable correlation function gets the closest to the theoretical coverage provided by a real multi-operator network. In the case of full network sharing our models significantly over-estimate the amount of coverage in the network (approx. 2 dB vertical shift for the "LGCP stable" model). However, in the case of infrastructure sharing we observe that the "LGCP stable" model provides a very tight match to a shared network performance. The spectrum sharing scenario provides significantly worse coverage, which should come as no surprise as we assumed worst-case interference. Interestingly, in the case of spectrum sharing we observe a cross-over point, which we have reported also in [99], which is related to the amount of clustering in the evaluated patterns. Clearly, a real multi-network provides higher clustering than any of our models can produce. It is also worth noting that, while statistical tests have shown that a superposition of independent realizations of the Strauss process provide a reasonably good match to the real data, according to our coverage analysis this way of modelling multi-operator patterns fails to reproduce coverage properties of multi-networks in wireless network sharing scenarios.



(a) Full sharing GSM (b) Infrastructure sharing GSM (c) Spectrum sharing GSM

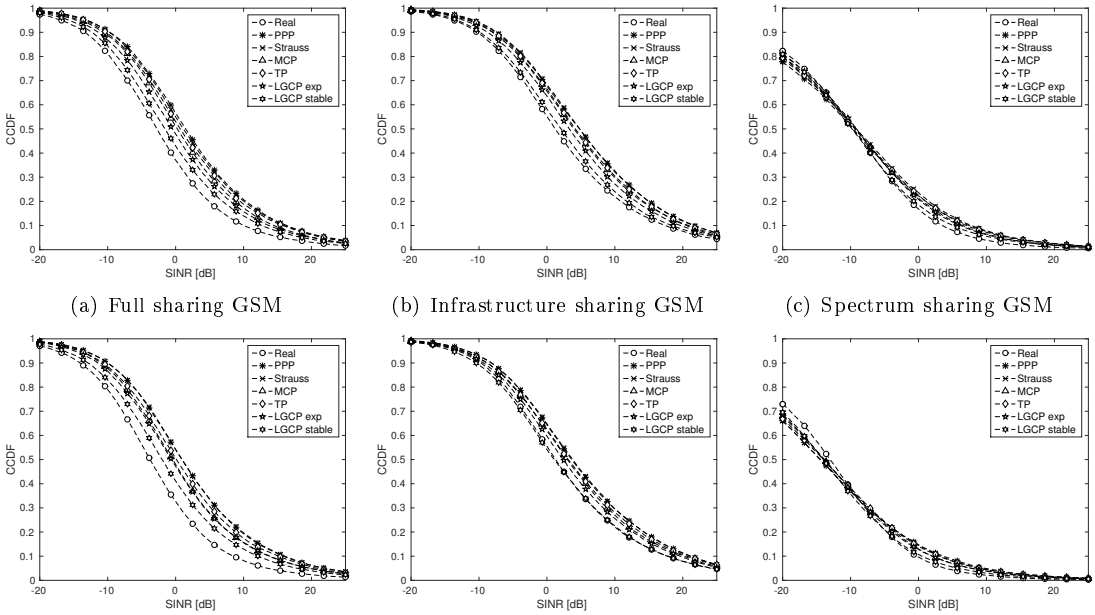
(d) Full sharing UMTS (e) Infrastructure sharing UMTS (f) Spectrum sharing UMTS

Figure 4.12: The coverage probability comparison for real deployment in Dublin (circle) and various fitted point process models, for each radio access network sharing scenario and radio technology when transmit power is unitary for all base station transmitters. Clearly, "LGCP stable" (hexagonal star) provides the closest fit to real network data.

4.5.2 Coverage probability - transmit power assigned according to the proposed model

The first thing we observe in Fig. 4.13 is base station transmit power assigned from a model based on real data has a negligible impact on the coverage of the analysed network model (the coverage

probability is degraded by roughly 1% for each SINR point), which is consistent with our prior observations [57]. This is true also when the noise level of -130 to -90 dBW is included at the receiver. This result is good news as it allows us to use, without loss of generality, a simplified model for multi-network that includes unitary transmit power.



(a) Full sharing GSM (b) Infrastructure sharing GSM (c) Spectrum sharing GSM

(d) Full sharing UMTS (e) Infrastructure sharing UMTS (f) Spectrum sharing UMTS

Figure 4.13: The coverage probability comparison for real deployment in Dublin (circle) and various fitted point process models, for each radio access network sharing scenario and radio technology when transmit power is iid according to a truncated Gaussian distribution. Again, "LGCP stable" (hexagonal star) provides the closest fit to real network data.

What we can conclude from the above results is that the chosen point processes provide a reasonable representation of real base station deployments when sharing scenarios are considered, even when exact information about operator-base station assignments is not available. Furthermore, similar results hold even when transmit powers are assigned from a model based on real data [57], which also suggests that this type of analysis may be robust to certain operator settings.

4.6 Conclusion

What we have shown here is that networks deployed by independent operators tend to cluster at shorter distances, which may correspond to areas of high demand, and repulse at long-distances, which would presumably correspond to coverage provisioning and low availability of shared sites. This result implies that if one wants to model (or simulate) the performance of a generalized shared infrastructure between N mobile operators, one cannot simply model the entire network as a superposition of N independent realizations representing single-operator networks. Instead, the approach we propose is to model such a network as a single clustered point process and indeed, as we have shown in our results, clustered point processes that represent well such networks can be found. What is even more appealing is that similar models may be used to represent networks of multiple operators in different countries (herein, we examined Ireland, Poland and the UK), which suggests that using such a model may come useful when deriving fundamental results about wireless network

sharing, or evaluating the performance of protocols and management strategies designed for shared wireless infrastructure. Moreover, the analysis of downlink transmit powers leads us to conclude that downlink transmit power allocations in a real wireless network, both in the single and the multi-operator case, may be considered iid, and a single statistical model may be used to replicate these allocations.

5 Inter-operator resource sharing

Inter-operator sharing

"When I talk about language (words, sentences, etc.) I must speak the language of every day. Is this language somehow too coarse and material for what we want to say? Then how is another one to be constructed? -And how strange that we should be able to do anything at all with the one we have!"

- Ludwig Wittgenstein, *Philosophical Investigations*

Capital and operating costs of running wireless mobile networks have been increasing steadily, while the growth in popularity of over-the-top services which utilize a flat-fee data service does not translate into a proportionate stream of revenues for mobile operators. That means mobile network operators have to seek ways for reducing their costs. It is thus inevitable that next generation mobile networks will have to rely ever more heavily on network sharing, which is considered as one of the most important cost reduction measures for mobile network operators. In the research to date we find evidence for mobile network sharing leading to improvements in coverage, or transmission rates, but little is known in general about the dependency between various network sharing policies and spatial properties of shared networks. Herein, we develop a model for a shared mobile network that allows us to study the efficiency of various infrastructure and spectrum sharing policies under different spatial models of radio access infrastructure deployment. Whenever the Poisson point process (PPP) is used, we back up our analysis with analytical results. Yet one can only go so far with analytical work and our study of coordination mechanisms and non-PPP spatial geometries, which introduce a dependence that is considered in the state-of-the-art as intractable, is based on Monte Carlo (MC) simulations, with the PPP models being used to validate them. The content of this chapter is based on [99], as well as our ongoing work efforts, which are currently under revision for publication.

5.1 Introduction

One of the prevailing pressures in the design of communications networks is known as the cost-demand offset [6]. In plain words, the demand for capacity grows faster than the per unit cost of capacity reduces. Indeed, the fast-paced growth in mobile data demand (see, for example, [7]) has led to ever-more ambitious targets for the expected capacity of mobile networks [8]. This has resulted in steadily growing costs: world-wide mobile communications capital costs have been predicted to grow from around $216 billion in 2014 to $224 billion by 2017, as reported during the Mobile World Congress (MWC'2015), while the annual fees for radio spectrum usage are also subject to a major price revision; for example, in the UK the Office for Communications (Ofcom) has put out a

proposal that will result in an increase in annual spectrum fees from £64.4 million to £199.6 million for all domestic mobile operators. Yet, the capital as well as operational cost increases have not been followed by a corresponding stream of revenues from mobile data services, which are the main services being used by mobile subscribers today [111]. The reason for this, as pointed out in [9], is that mobile data revenues are decoupled from the traffic.

One approach for mobile operators to ease the burden of these large costs is to share ownership and control over some components of the network, for example, the radio access network or the backhaul, that generate the most costs. This may mean establishing network sharing agreements with other operators, and/or subletting elements of the radio access infrastructure from other mobile operators, in which case we talk about *infrastructure sharing*. Another possibility is to re-use spectrum between operators wherever and whenever capacity expansion is needed, which we refer to as *spectrum sharing*. Indeed, we are starting to gather evidence (see [112]) that the spatial correlation in demand experienced by mobile operators is low enough to encourage this type of sharing. Both types of sharing can also be applied in combination, which we refer to as *full sharing*, to result in even larger costs reduction to the mobile operators.

During the recent decade wireless mobile network sharing has grown in popularity and it is considered an important cost reduction measure for mobile network operators [21]. Indeed, both industry and research communities have already recognized the importance of network sharing in the evolution of mobile networks. We have the 3rd Generation Partnership Project (3GPP)-defined standards for network sharing [18]: the Multi-Operator Core Network (MOCN), and the Gateway Core Network (GWCN). We have the 2012 report from the European Commission entitled "Promoting the shared use of radio spectrum resources in the internal market". We have communications regulators testing the feasibility of various spectrum sharing models, as is the case with, for example, the recent ruling of the Federal Communications Commission (FCC) to open military frequencies in 3550-3700 MHz band to mobile broadband services[1]. And most importantly we have some recent commercial examples of mobile network sharing some of which we have listed in Chapter 2.

Joint infrastructure and spectrum sharing, based on simulations of a grid network deployment in three scenarios of capacity, spectrum and virtualized resource sharing, has been studied in [106]. In [113] trade-off between the sharing of remote radio heads (RRHs) and spectrum was studied on a simulation scenario with a single macrocell RRH and a number of uniformly distributed small-cell RRHs. In [114] the interchangeability between massive MIMO antennas and shared spectrum was studied as an inter-operator auction. In [115] infrastructure sharing in a real-world multi-operator mobile network was shown to significantly reduce the number of base stations required to provide mobile service and improve coverage in under-served areas. Similarly, in [116] an optimization-based model of new infrastructural deployments by multiple mobile operators was applied to study the effects of a shared greenfield deployment and its relationship to various techno-economic parameters. Gains in spectral efficiency from inter-operator spectrum sharing were studied in [107] using computer simulations over a grid-based deployment for two spectrum sharing regimes: orthogonal spectrum sharing, where operators exclusively assign spectral resources coming from a shared pool, and non-orthogonal spectrum sharing, where the operators aggressively re-use spectrum by allowing more than a single operator to use a specific spectral resource in any given area at a particular time.

It is only recently that stochastic geometry has been applied to the studies of wireless network

[1]The ruling can be accessed here: `http://transition.fcc.gov/Daily_Releases/Daily_Business/2015/db0421/FCC-15-47A1.pdf`

sharing. In [117] stochastic geometry was applied to study the trade-off between spectrum and base station density, given variable level of user activity. In [118] a stochastic geometric model was developed to study orthogonal and non-orthogonal spectrum sharing between device-to-device networks and macrocell users. And only recently, stochastic geometry was applied to make the case for spectrum sharing in millimetre-wave (mmWave) spectrum bands (see [119]), whereby the excessive interference is mitigated with the help of directional antennas.

In our work we assess fundamental trade-offs between spectrum and infrastructure sharing, and whether (or to what extent) one can be substituted for the other. Moreover, we consider the two in combination and quantify the resulting network performance gains. To achieve our goal we define three network sharing scenarios: *infrastructure sharing*, *spectrum sharing*, and *full sharing*. For each of the scenarios we analyze the resulting network performance in terms of coverage and user data rate. Fig. 5.1 summarizes our main results and trade-offs observed.

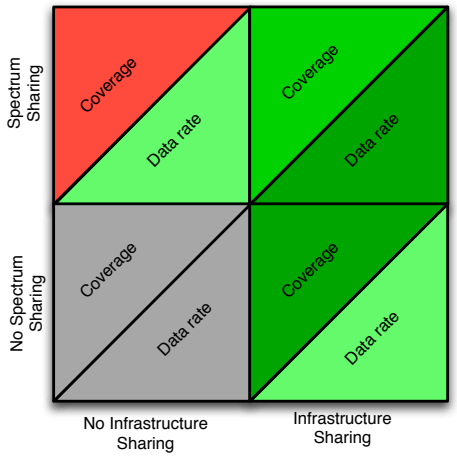


Figure 5.1: Conceptual depiction of the effect of radio access infrastructure and spectrum sharing on network coverage and data rate. Each square corresponds to the network performance of a different combination of spectrum and infrastructure sharing. Each square is split evenly between network coverage and user data rate. Saturation of the green-coloured cells reflects the gain with respect to the no sharing case (grey cells), while red-coloured cells represent negative gain. Infrastructure sharing in isolation greatly improves network coverage, but provides minor gain in terms of data rate; spectrum sharing (when worst case interference scenario is considered) provides minor gains in data rate, but degrades coverage. When applied in combination, both coverage and data rate are improved over the no sharing case, with network coverage being slightly worse than in the pure infrastructure sharing case, as a result of increased interference.

For the cases when spectrum is shared, i.e., spectrum and full sharing, we consider two methods for utilization of the shared spectral resources: channel bonding (operators simply treat the shared spectrum as one band), and best channel selection (operators dynamically select between the shared spectrum bands to obtain the diversity gain from mitigating detrimental fading effects). In addition, we investigate two methods for managing inter-operator interference, one where operators do not coordinate the use of spectrum, and another one where the use of spectrum is coordinated by means of spatial spectrum sharing policies. We summarize the resulting scenarios in Fig. 5.2.

We generalize the resulting scenarios by considering various spatial distributions of base stations. Specifically, we consider two mobile operators whose networks have been deployed either *independently*, or with *clustering*, i.e., there is strong spatial correlation between the locations of base stations of one operator and the locations of base stations of the other operator. We also consider an asymptotic case of clustering, i.e., *co-location*, where the nearest base stations of two

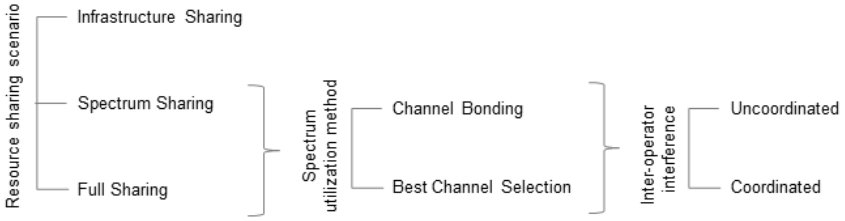


Figure 5.2: Taxonomy of the studied scenarios.

operators are located an arbitrarily small distance apart. The clustering case is of particular interest as it is representative of real radio access network deployments, which tend to be deployed to follow social and geographical features, resulting in clustered deployment patterns (see [35, 75]).

It is worth noting that the decisions on whether to share (and how) are not purely technical and involve consideration of various costs, legal/regulatory framework, operators' position in the market, corporate strategy, etc. [97]. Yet, consideration of these aspects is out of scope of this chapter and some analysis which includes listed factors may be found in Chapter 6 and the works of Markendahl et al. (see, for example, [120]).

The remainder of the chapter is structured as follows. First, we outline the system model of a shared mobile network and define the reference scenarios. Subsequently, we present analytical expressions of network performance under each of the considered scenarios for the PPP, which we then expand onto a numerical analysis of our scenarios in various spatial geometries for both uncoordinated and coordinated spectrum sharing. Eventually, we conclude with a discussion of the obtained results.

5.2 System model and reference scenarios

5.2.1 System model

We model base station deployment by a single operator as a point process defined as a random countable set[2] $\Phi_n \subset \mathbb{R}^2$, with elements being random variables $x_i \in \mathbb{R}^2$ and $n \in \mathcal{N}$, with $\mathcal{N}$ denoting the set of mobile operators. We also denote $\Phi = \bigcup_{n\in\mathcal{N}} \Phi_n$ as the set of all transmitter locations of all operators. We assume that the power of the signal received by the reference user $y_n = (0,0)$ of operator $n \in \mathcal{N}$ is affected by two factors: pathloss $l(x)$ and power fading h_x, where x denotes the location of the serving transmitter. The pathloss function $l : \mathbb{R}^2 \to \mathbb{R}_+$ is of the form $l(x) = ||x||^{-\alpha}$, where α is the pathloss exponent, while the power fading between the transmitter in x and the typical user is spatially independent and exponentially distributed with unit mean, i.e., we assume Rayleigh fading with unitary transmit power. We assume that the reference subscriber will associate with the closest transmitter of operator n, i.e., $\text{argmin}_{x\in\Phi_n} ||x - y_n||$. Note that, when infrastructure sharing is in place, the reference subscriber of operator n may associate with any network of any infrastructure sharing operator, i.e., the association policy becomes $\text{argmin}_{x\in\Phi} ||x - y_n||$. Moreover, our model applies to the downlink only and presupposes a full buffer transmission, i.e., each transmitter continuously operates in selected bands.

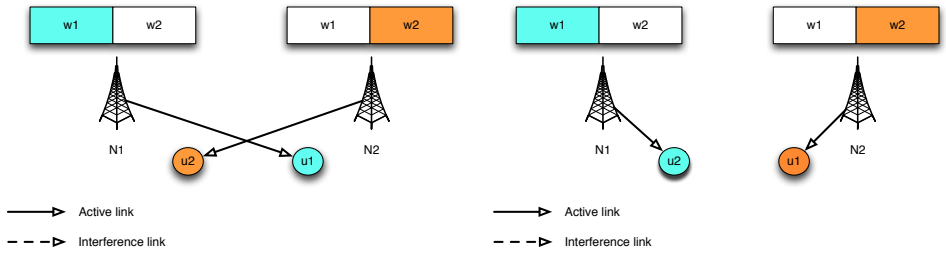
[2]Using the random set formalism we implicitly assume that base stations from more than one operator do not occupy the same location. Yet, this does not preclude base stations to be located arbitrarily close to each other, which we consider as co-location.

5.2.2 Scenarios

Herein, we are interested in modelling resource sharing in wireless mobile networks for two types of resources: spectrum and infrastructure. We assume that each mobile operator has a license to use one spectrum band b_n, and each spectrum band is of equal size and does not overlap with any other spectrum band, hence $b_n = b$ for each $n \in \mathcal{N}$. In a system where operators decide to share spectrum, exclusive spectrum usage becomes a special case. Having this in mind a spectrum sharing policy for location $x \in \Phi_m$ can be defined as a vector $\mathbf{V}_x$ of size $|\mathcal{N}|$ (equal to the number of available spectrum bands) consisting of binary variables v_x^n denoting whether a transmitter in x may utilize spectrum band belonging to operator n. In addition, the operators may share infrastructure in which case their subscribers are able to freely roam between transmitters of any of the infrastructure sharing operators. Therefore the joint network, as seen from the perspective of a reference user of operator n, will be the sum of $|\mathcal{N}|$ networks of the sharing operators $\Phi = \bigcup_{n\in\mathcal{N}} \Phi_n$. Based on this setup, in the following we describe four scenarios: *exclusive use (baseline)*, *infrastructure sharing*, *spectrum sharing* and *full sharing*.

Baseline scenario

The conventional scenario of a commercial mobile market consists of competing operators, each of which owns the network infrastructure and holds a license to spectrum, which also implies that subscribers of one operator exclusively[3] utilize the infrastructure and spectrum of their operator as depicted in Fig. 5.3(a). This serves as our baseline for comparison, and the coverage probability and the average user rate for this scenario we have given in Eq. (3.15) and Eq. (3.21), respectively.



(a) Exclusive use (b) Infrastructure sharing

Figure 5.3: Two operators (1 and 2) with subscribers (u1, u2) exclusively using their spectrum (w1, w2), and either prohibiting or allowing for access to each other's infrastructure (N1, N2).

Infrastructure sharing scenario

When infrastructure sharing is applied, subscribers of one operator are able to connect with transmitters of another operator. To represent this scenario we assume that operators pool their radio access infrastructure without pooling their operational frequencies (see Fig. 5.3(b)). Effectively, interference to the desired signal is identical to the single-operator case, i.e., it comes from Φ_n where n is the operator to which the serving transmitter belongs. In this scenario users connect to the closest transmitter belonging to the network of any of the $|\mathcal{N}|$ operators. Henceforth, the signal-to-noise

[3] Leaving international roaming out of the picture.

and interference ratio (SINR) for this scenario takes the following form:

$$\text{SINR} = \frac{h_x l(x)}{W + \sum_{y \in \Phi_n \setminus \{x\}} h_y l(y)}, \tag{5.1}$$

where W is the noise power.

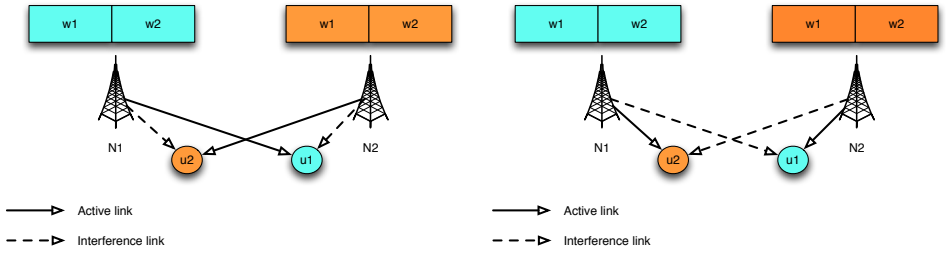
Spectrum sharing scenario

In the spectrum sharing scenario we assume that operators allow access to their spectrum bands from outside of their network, by means of the spectrum sharing policy, without sharing their infrastructure (see Fig. 5.4(a)). Effectively, interference to the desired signal of a user of operator n in spectrum band $k \in \mathcal{N}$ comes from all transmitters of operator k and the transmitters of all other operators that may access spectrum band k. The SINR for this scenario takes the following form:

$$\text{SINR}_k = \frac{v_x^k h_x l(x)}{W + \sum_{y \in \Phi \setminus \{x\}} v_y^k h_y l(y)}, \tag{5.2}$$

where $x \in \Phi_n$ is a tagged transmitter of operator n.

Let us note that when inter-operator interference is uncoordinated $v_y^k = 1$ for all $k \in \mathcal{N}$ and $y \in \Phi$. In the case when inter-operator interference coordination is in place, $v_y^k = 1$ for all $y \in \Phi_k$, while for the remainder of transmitters v_y^k is set according to the specifics of the spectrum sharing policy applied.



(a) Spectrum sharing (b) Network sharing

Figure 5.4: Two operators (1 and 2) with subscribers (u1, u2) sharing their spectrum (w1, w2), and either prohibiting or allowing for access to their infrastructure (N1, N2).

Full sharing scenario

In the full sharing scenario operators both pool their radio access infrastructure and allow for shared use of spectrum. Spectrum pooling results in increased interference as more transmitters may now operate in the same frequency band and users may connect to any of the transmitters of the sharing operators, see Fig. 5.4(b). The SINR model is essentially the same as in the spectrum sharing scenario, with the only difference being that the tagged transmitter comes from the pool of all infrastructure sharing operators, i.e., $x \in \Phi$.

5.3 Analytical results

In this section, we apply stochastic geometry to derive analytical expressions for the coverage probability and the average user data rate in the cases of spectrum, infrastructure, and full sharing between multiple mobile operators with radio access infrastructure distributed according to the homogeneous PPP with intensity λ_n each. We assume no inter-operator interference coordination, and two modes of shared spectrum usage: channel bonding and best channel selection. We validate the obtained closed-forms using MC simulations.

5.3.1 Main results

Infrastructure sharing scenario

Operators pool their radio access infrastructure without pooling their operational frequencies. Effectively, subscribers of one operator are able to attach to transmitters of another operator. When only infrastructure is shared, we get the following result.

Proposition 1. *The coverage probability of a reference user of operator n in the infrastructure sharing scenario can be expressed as:*

$$p_n(\theta) = 2\pi \sum_{i\in\mathcal{N}} \lambda_i \int_0^\infty \exp(-\theta r^\alpha W) \exp\Big(- \pi r^2 \Big(\sum_{j\in\mathcal{N}} \lambda_j + \lambda_i \mathfrak{Z}(\theta,\alpha)\Big)\Big) r dr, \tag{5.3}$$

where $\mathfrak{Z}(\theta,\alpha) = \theta^{2/\alpha} \int_{\theta^{-2/\alpha}}^\infty \frac{1}{1+u^{\alpha/2}} du$. *The average data rate of a reference user of operator n can be found as:*

$$\tau_n = \mathbb{E}_\gamma \Big[p_n \Big(\exp(\gamma/b) - 1 \Big) \Big]. \tag{5.4}$$

Proof:

The proof is provided in Appendix B.1. ■

Corollary 1. *When no noise and $\alpha = 4$ is assumed, the coverage probability of a user of operator n is given as:*

$$p_n(\theta) = \sum_{i\in\mathcal{N}} \frac{\lambda_i}{\sum_{j\in\mathcal{N}} \lambda_j + \lambda_i \sqrt{\theta} \arctan(\sqrt{\theta})}. \tag{5.5}$$

Proof: Assuming no noise and $\alpha = 4$, we can calculate the integral in Eq. (5.3). In addition, $\mathfrak{Z}(\theta, 4)$ simplifies to $\sqrt{\theta}\arctan(\sqrt{\theta})$. ■

Spectrum sharing scenario

In the spectrum sharing scenario operators bond their spectrum bands and allow for uncoordinated access to the joint band by all the transmitters of all the operators. Yet subscribers may still only connect to the transmitters of their operator. When spectrum is shared, we get the following result.

Proposition 2. *The coverage probability of a user belonging to operator n in the spectrum sharing scenario with channel bonding can be expressed as:*

$$p_n(\theta) = 2\pi\lambda_n \int_0^\infty \exp(-\theta r^\alpha W) \exp\Big(- \pi r^2 \Big((1 + \mathfrak{Z}(\theta,\alpha))\lambda_n + \mathfrak{Z}_0(\theta,\alpha) \sum_{j \in \mathcal{N}\setminus\{n\}} \lambda_j \Big) \Big) r dr. \quad (5.6)$$

where $\mathfrak{Z}_0(\theta,\alpha) = \theta^{2/\alpha}\Gamma(1+2/\alpha)\Gamma(1-2/\alpha)$. *The data rate can be expressed as:*

$$\tau_n = |\mathcal{N}| \cdot \mathbb{E}_\gamma \Big[p_n \Big(\exp(\gamma/b) - 1 \Big) \Big]. \quad (5.7)$$

Proof:

The proof is provided in Appendix B.2.

■

Corollary 2. *When no noise and $\alpha = 4$ is assumed, the coverage probability of a user of operator n is given as:*

$$p_n(\theta) = \frac{\lambda_n}{\lambda_n + \lambda_n \sqrt{\theta} \arctan(\sqrt{\theta}) + \frac{\pi}{2}\sqrt{\theta} \sum_{j \in \mathcal{N}\setminus\{i\}} \lambda_j}. \quad (5.8)$$

Proof: Similarly to Corollary 2, when no noise and $\alpha = 4$ is assumed, Eq. (5.11) can be simplified to Eq. (5.8). ■

An alternative to channel bonding is to access shared spectral resources by means of best channel selection to exploit the potential diversity gain present when the fading between the shared bands is not correlated (coherence bandwidth is large enough). Assuming best channel selection, and specializing to a two-operator case, we get the following result.

Proposition 3. *The coverage probability of a user belonging to operator n from the shared spectrum when best channel selection is applied can be expressed as:*

$$\bar{p}_{nm}(\theta) = 2p_n(\theta) - p_{nm}(\theta), \quad (5.9)$$

where $p_n(\theta)$ is the single-operator coverage probability obtained in Eq. (5.5), and $p_{nm}(\theta)$ is the joint probability of coverage for a user of operator n in two spectrum bands n and m, which can be expressed as:

$$p_{nm}(\theta) = \frac{\lambda_n}{\lambda_n + \lambda_n \frac{3}{2}\sqrt{\theta} \arctan\sqrt{\theta} + \lambda_n \frac{\theta}{2(1+\theta)} + \lambda_m \frac{3}{4}\pi\sqrt{\theta}}. \quad (5.10)$$

Proof:

The proof is provided in Appendix B.3.

■

Full sharing scenario

In the full sharing scenario operators both pool their radio access infrastructure and allow for shared use of spectrum by means of channel bonding. In consequence we get the following result.

Proposition 4. *The coverage probability of a user of operator n in the full sharing scenario can be expressed as:*

$$p_n(\theta) = 2\pi\Big(\sum_{i\in\mathcal{N}}\lambda_i\Big)\int_0^\infty \exp(-\theta r^\alpha W)\exp\Big(-\pi r^2\Big(1+\mathfrak{Z}(\theta,\alpha)\Big)\sum_{i\in\mathcal{N}}\lambda_i\Big)rdr. \tag{5.11}$$

The data rate can be expressed as:

$$\tau_n = \sum_{i\in\mathcal{N}}\mathbb{E}_\gamma\Big[p_i\Big(\exp(\gamma/b)-1\Big)\Big]. \tag{5.12}$$

Proof: Let us first recall the following result: the superposition of PPPs is also a PPP, with intensity equal to the sum of the intensities of the component processes [121][Proposition 1.3.3]. In the case of the full sharing scenario, the users connect to the closest transmitter of any of the networks of the sharing operators and suffer interference from all the other transmitters. Formally a fully shared network is a PPP $\Phi = \bigcup_{i\in\mathcal{N}}\Phi_i$ with intensity $\lambda = \sum_{i\in\mathcal{N}}\lambda_i$. Therefore its coverage probability corresponds to that of the baseline case (single-operator and exclusive use of resources), yet with intensity λ. In order to calculate the average data rate we need to simply take an expectation over the sum of data rates obtained from using each spectrum band, which can be written as the sum of expected data rates obtained for each spectrum band. ■

Corollary 3. *When no noise and $\alpha = 4$ is assumed, the coverage probability of a user of operator n simplifies to:*

$$p_n(\theta) = \frac{1}{1+\sqrt{\theta}\arctan(\sqrt{\theta})}. \tag{5.13}$$

The result, which also happens to describe a single-operator PPP network [62], is obtained via straight-forward calculations.

two effects "cancel" each other out.

Similarly to the spectrum sharing scenario, instead of simply utilizing the entire spectrum band we may exploit spectrum sharing for diversity gain (apply best channel selection, instead of channel bonding), in which case we get the following result (derived for a two-operator case).

Proposition 5. *The coverage probability of a user belonging to operator n from the full network sharing when best channel selection is applied can be expressed as:*

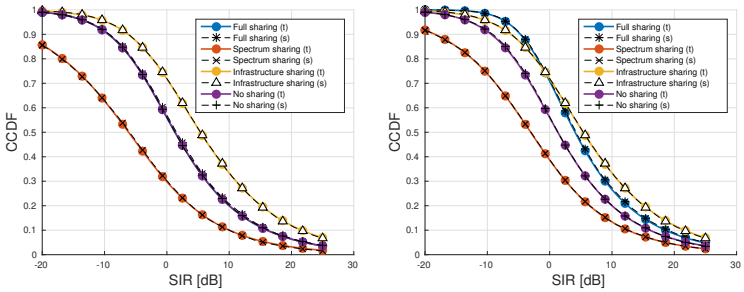
$$\bar{p}_{nm}(\theta) = 2p_n(\theta) - p_{nm}(\theta), \tag{5.14}$$

where $p_n(\theta)$ is expressed in Eq. (5.13), and $p_{nm}(\theta)$ is the joint probability of coverage for a user of operator n in two spectrum bands n and m, which can be expressed as:

$$p_{nm}(\theta) = \frac{1}{1+\frac{3}{2}\sqrt{\theta}\arctan(\sqrt{\theta})+\frac{\theta}{2(1+\theta)}}. \tag{5.15}$$

Proof:

The derivation of this result mirrors that provided in Appendix B.3, with a modification to the set of interferers in Eq. (7.27), which simplifies some of the calculations and allows us to obtain the



(a) Coverage probability (channel bonding)

(b) Coverage probability (best channel selection)

Figure 5.5: Cross-validation of the simulations and the derived closed-form expressions for the coverage probability, for a two-operator case where networks of the sharing operators are evenly-sized; in the legend, "t" - denotes an analytical result, and "s" - denotes a simulation result.

joint probability of coverage for a user of operator n in two spectrum bands n and m as:

$$p_{nm}(\theta) = 2\pi(\lambda_n + \lambda_m) \int_0^\infty \exp(-\theta r^\alpha W) \exp\Big(- \pi r^2 \big(1 + \mathfrak{Z}'(\theta, \alpha)\big)\big(\lambda_n + \lambda_m\big)\Big) r dr, \tag{5.16}$$

where $\mathfrak{Z}'(\theta, \alpha) = \theta^{2/\alpha} \int_{\theta^{-2/\alpha}}^\infty 1 - \big(1 + u^{\alpha/2}\big)^{-2} du.$ ■

5.3.2 Validation

Before we move on to discuss the results, we cross-validate the correctness of the derived closed-form expressions and our simulations. For that purpose we compare our closed-form expressions to the MC simulation results for the PPP. We perform this comparison for two and more operators (at least for the channel bonding case), with evenly and non-evenly-sized networks. However, to illustrate the correctness of our analytical expressions we present the two operator and evenly-sized networks case only. In Fig. 5.5(a) and Fig. 5.5(b) we see that our closed-form expressions tightly match the results from our simulations for both modes of utilizing the shared spectrum.

5.4 Numerical results

In our work we analyze two types of relative radio access network deployments between mobile operators: independent and clustered. For the sake of clarity of our argument, from now on we assume $|\mathcal{N}| = 2$. Now, we model the independent deployment as two independent realizations of the PPP, while the clustered deployment as either of the following two point processes: the Gauss-Poisson process (GPP), which allows us to vary the level of clustering between the two networks, and the Log-Gaussian Cox process (LGCP), which, as we have shown in Chapter 4, provides a good match with the real multi-operator mobile network deployments (both processes are defined in Chapter 3). We derive analytical results for the former case only, while the latter two, being analytically intractable[4], we evaluate using MC simulations. In both the PPP and GPP case the underlying process of distributing transmitters for a single-operator deployment is the PPP, which according to spatial statistical studies to date (see [35, 68]) may be considered a reasonable approximation of a real-world single-operator single-technology deployment.

[4]Bounds on the Laplace transform of the GPP can be found in [70].

Whenever spectrum is shared, we differentiate between two methods of utilizing shared spectral resources: (i) channel bonding (operators simply treat the shared spectrum as one band), and (ii) best channel selection (operators dynamically select between the shared spectrum bands to obtain the diversity gain from mitigating detrimental fading effects). We also distinguish between two methods for managing inter-operator interference: uncoordinated (uncoordinated inter-operator interference) and coordinated by means of spatial spectrum sharing policies (coordinated inter-operator interference). In each case we look into the coverage probability and the average user data rate for the scenarios of infrastructure, spectrum and full network sharing, for each deployment type.

When presenting the results, we assume no noise, i.e., we use the signal-to-interference ratio (SIR), and the pathloss exponent $\alpha = 4$. The analysis based on these assumptions is in line with similar investigations of network coverage and data rate made in [62, 68, 119].

5.4.1 Spectrum sharing with uncoordinated inter-operator interference

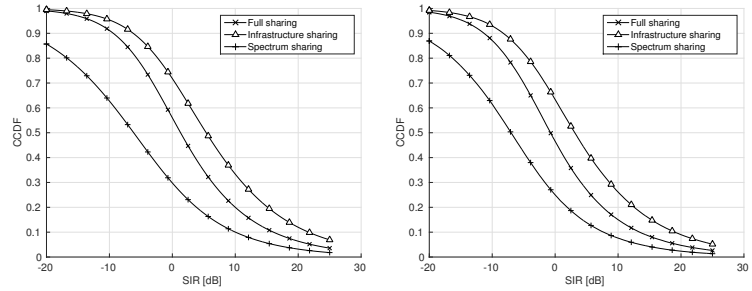
Coverage probability

The first thing we analyze is the comparative coverage results for various sharing scenarios, when networks of the sharing operators are evenly-sized. In Fig. 5.6 we immediately see that infrastructure sharing provides superior coverage as compared to spectrum and full sharing scenarios with channel bonding. This is related to the fact that with the increase in radio access infrastructure sharing, we also shorten the distance to the potential serving transmitter, i.e., we increase the strength of the signal provided by the serving transmitter. This effect does not occur when both operators co-locate, as any additional shared transmitters, will be co-located with the private ones. Let us note that, if a comparison between different scenarios is made, the LGCP distributed network (see Fig. 5.6(b)) provides worse coverage than the PPP (see Fig. 5.6(a)) or GPP (see Fig. 5.6(c)) networks. This is related to higher level of spatial clustering in our LGCP model. We can see from Fig. 5.6(c) that some level of clustering may still yield coverage performance that is almost identical to the performance of an independently distributed infrastructure (see Fig. 5.6(a)).

When, instead of channel bonding, spectrum is shared by means of best channel selection (see Fig. 5.7), we see the coverage performance in each spectrum sharing scenario being significantly improved, with the full network sharing faring very closely to the infrastructure sharing. Interestingly, full network sharing provides better coverage at low SIR values, which may be especially important for improving the performance for cell-edge users). Let us note also that in the co-located case (see Fig. 5.7(c)) the application of spectrum sharing allows an operator to switch between the performance of a fully shared network and a non-shared network (in the co-located case it provides the same performance as infrastructure sharing).

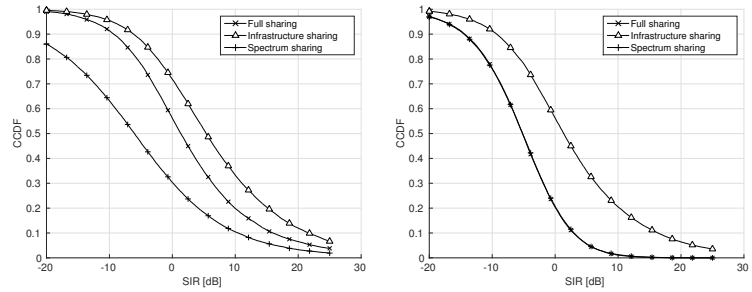
In Fig. 5.6(c) we observed the coverage performance for a fixed level of clustering, corresponding to the normalized cluster radius of 0.1 (normalized to the size of the simulation window). Now, we change our perspective and observe the impact that clustering of infrastructure has on different sharing scenarios (we assume spectrum sharing based on channel bonding). Figures 5.8(a) to 5.8(c) present the coverage performance for each of our sharing scenarios when the infrastructure is: independently distributed (ppp), co-located ($u \to 0$), and clustered with a variable cluster radius.

As expected, in Figures 5.8(a) and 5.8(b) we observe that the increase in clustering (decrease in cluster radius) deteriorates coverage (the direction of change marked with a red arrow) for both



(a) Independent networks (PPP)

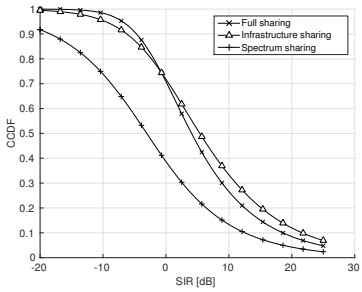
(b) Clustered networks (LGCP)



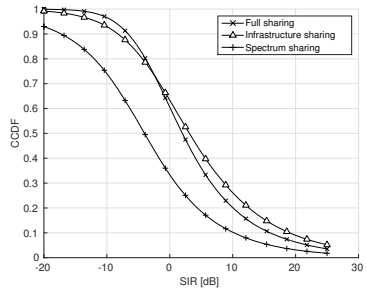
(c) Clustered networks (GPP, $u = 0.1$)

(d) Co-located networks (GPP, $u \to 0$)

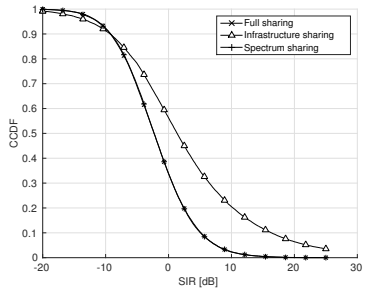
Figure 5.6: The coverage probability for different sharing scenarios with channel bonding.



(a) Independent networks (PPP)



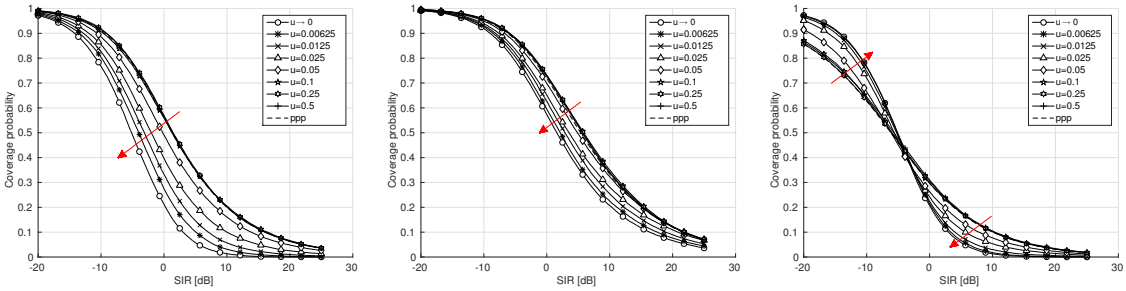
(b) Clustered networks (LGCP)



(c) Co-located networks (GPP, $u \to 0$)

Figure 5.7: The coverage probability for different sharing scenarios with best channel selection.

full and infrastructure sharing scenarios. In the case of spectrum sharing (Fig. 5.8(c)) the impact of clustering on the sharing performance is more complex. At low SIR values we observe that the co-located infrastructure provides superior coverage while the independently distributed infrastructure fares the worst. After the cross-over point (slightly below 0 dB) the situation reverses. This is related to an interplay between intra- and inter-cluster interference. Clearly, decreasing the cluster radius increases intra-cluster interference. However, as the cluster radius is decreased so is the inter-cluster interference, as the distribution of distances to the interferers outside of the serving cluster changes (pairs of interferers are getting closer to each other). Effectively, while we observe that clustering reduces the probability of obtaining SIR-values above the cross-over point (due to increased inter-cluster interference), we also observe an increase in the probability of obtaining SIR-values below the cross-over point (due to reduced intra-cluster interference), as compared to the independently distributed infrastructure.

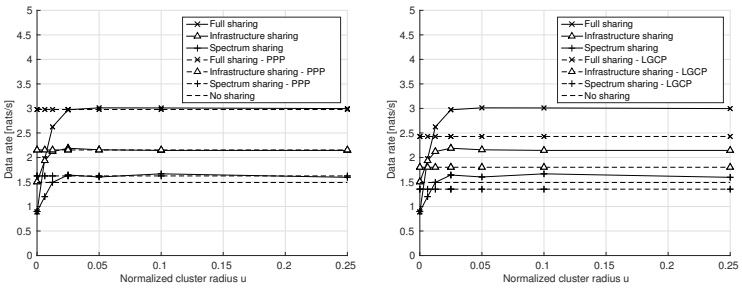


(a) Full sharing (b) Infrastructure sharing (c) Spectrum sharing

Figure 5.8: The coverage probability versus the relative cluster size in a shared clustered network (the performance of independently distributed network is marked with a dashed horizontal line).

Average data rate

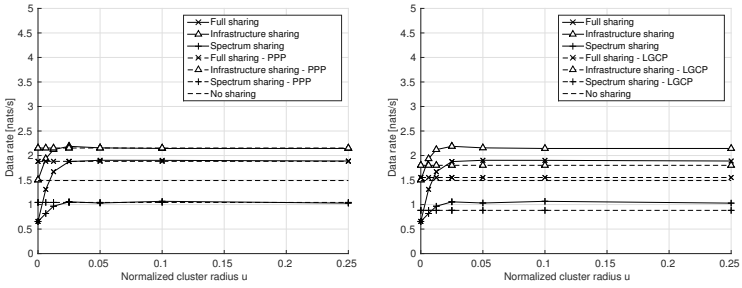
When the average user data rate is considered (see Fig. 5.9), with the cluster radius being large enough, we can observe that the rate performance with network sharing is always improved over the non-sharing scenario, which, as derived in [62], is approximately 1.49 nats/sHz. Moreover, full network sharing outperforms all other sharing scenarios for majority of the cluster radii. It is only at small cluster radii when the data rate significantly drops, and, in fact, when the infrastructure co-locates full sharing achieves the same performance as the spectrum sharing scenario. From the figure we can also see that significant clustering and co-location of operator networks have a dramatic effect on the data rate performance of the spectrum sharing scenario (due to strong interference present in the shared network), which applies both to independent (see Fig. 5.9(a)) and LGCP-distributed networks (see Fig. 5.9(b)). This result shows the need to apply smart spectrum resource allocation techniques whenever a high level of clustering is present in a multi-operator network, which is typically the case with real mobile networks [35].



(a) GPP vs PPP (b) GPP vs LGCP

Figure 5.9: The average user data rate performance versus the relative cluster size in a shared clustered network with channel bonding.

Unsurprisingly, in Fig. 5.10 we see that best channel selection results in lower average data rates than in the case of channel bonding. Yet, the diversity gain obtained from the application of best channel selection allows operators to mitigate some of the detrimental effects of increased interference, and the more clustered the network the more efficient this mitigation becomes.

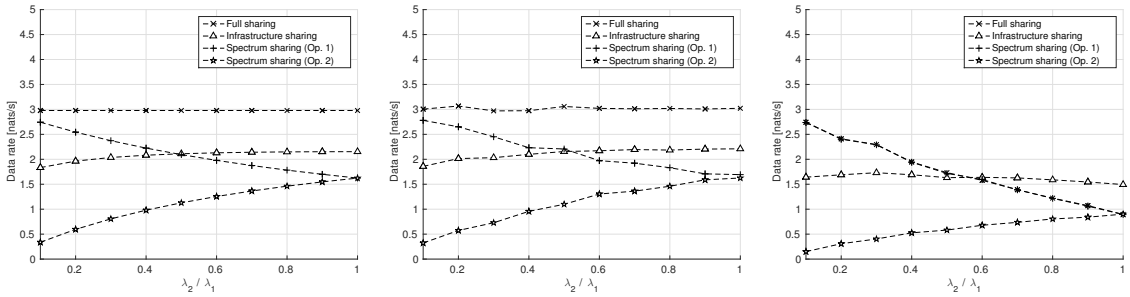


(a) GPP vs PPP (b) GPP vs LGCP

Figure 5.10: The average user data rate performance versus the relative cluster size in a shared clustered network with best channel selection.

Network density imbalance between sharing operators

We now show the performance of network sharing depending on the imbalance in the density of the networks of two operators. We specifically look at the average user data rate. We define the imbalance as the ratio between the respective intensities of their networks λ_2/λ_1, with 0.1 meaning highly unbalanced networks and 1 standing for evenly-sized networks. We can immediately note from Fig. 5.11(a) that the imbalance has no impact on the performance of independently distributed networks for the cases when infrastructure is shared, which is precisely the consequence of the result stated in Corollary 3, whereby the density of the infrastructure has no impact on the SIR performance. We can also observe that, in the case of spectrum sharing, the operator that has a denser network benefits the most from sharing, which is consistent with the fact that there is less interference created by the network of the smaller operator. Interestingly, at a specific imbalance level, for the smaller operator it does not make sense to share the spectrum anymore, at least without sharing the infrastructure. In the case of clustered deployment in Fig. 5.11(b), we observe that the infrastructure sharing gains provide additional performance benefits as soon as the network of the smaller operator is enlarged. Eventually, spectrum sharing for the larger operator outperforms infrastructure sharing when the imbalance is significant, which is related to the minuscule level of interference suffered from the smaller operator. When we look at the imbalance of a co-located network (see Fig. 5.11(c)), we can see fast degradation of performance for the spectrum sharing cases, which is again expected as increased co-location produces stronger interferers in the proximity of the desired transmitters.



(a) Independent (PPP) (b) Clustered (GPP, $u = 0.1$) (c) Co-located (GPP, $u \to 0$)

Figure 5.11: The average user data rate performance versus the imbalance in the size of the networks between the operators.

5.4.2 Spectrum sharing with coordinated inter-operator interference

In the following we perform a numerical study of the impact that a spectrum sharing policy has on inter-operator interference, and, effectively, the coverage and data rate performance. The resource sharing policy we use is based on the concept of exclusion zones, i.e., discs of arbitrary size centred at the locations of transmitters, in which a given resource cannot be shared. As illustrated in Fig. 5.12, given a resource belonging to some operator n, transmitters of operators other than n which are at a distance smaller than or equal to the radius of an exclusion zone from any transmitter of operator n cannot utilize that resource. By adjusting the exclusion zone radius operators may switch between more permissive and more restrictive attitude towards resource sharing (at an extreme resorting to exclusive use). We apply the exclusion zones to coordinate sharing of radio spectrum.

Formally, a spectrum sharing exclusion zone for a spectrum band that belongs to operator n is a disc of radius $R_n \in \mathbb{R}_+ \cup \{0\}$ centred at the location of a transmitter $y \in \Phi_n$ which prohibits transmitters of any operator other than n located within that disc from utilizing spectrum band n. For the sake of clarity of our argument we assume that R_n is the same for all operators, yet a more general case requires simply an additional index. Now, the spectrum sharing policy for some transmitter x for spectrum band $n \in \mathcal{N}$ takes the following value:

$$v_x^n = \begin{cases} 1, & \text{if } x \in \Phi_n \\ 1, & \text{if } x \in \Phi \setminus \Phi_n \vee ||x-y|| > R_n \forall y \in \Phi_n, \\ 0, & \text{otherwise.} \end{cases} \tag{5.17}$$

Loosely speaking, the spectrum sharing policy implies that only two types of transmitters may operate in spectrum band n: all the transmitters of operator n, i.e., Φ_n, and transmitters of operators other than n which are at least an R_n distance apart from any transmitter of operator n, i.e., $\widehat{\Phi}_j$ with $j \in \mathcal{N} \setminus \{n\}$. Note that $\widehat{\Phi}_j$ is a Poisson hole process (PHP) [50] which is known to be analytically intractable, with typical approximations involving thinning a PPP with a function that corresponds to the radius of the exclusion zone (see [122]). In our work instead of approximations we rely on direct results obtained by means of MC simulations.

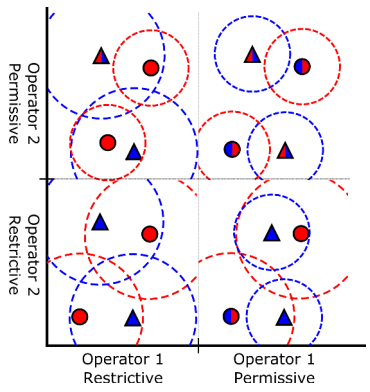
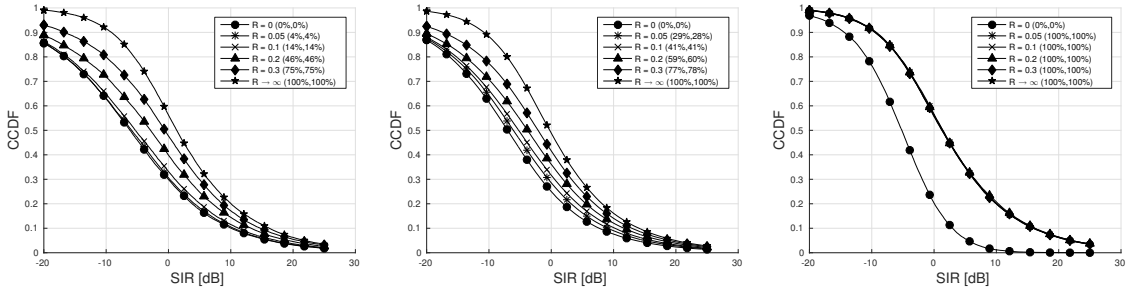


Figure 5.12: Exclusion zone-based coordination in resource sharing. When operators are restrictive in sharing of their resources, they simply increase their exclusion zones resulting in low sharing levels (bottom-left corner). When operators become more permissive in the sharing of their resources, they decrease the exclusion zones resulting in more resources being shared (top-right corner).

Coverage probability

As noted previously, when spectrum is shared without inter-operator interference coordination, the coverage probability gains from spectrum sharing become significantly deteriorated due to the presence of an increased number of interferers, some of which being located in close proximity to the reference user. When exclusion zones are applied evenly to each of the operators, we can, indeed, observe an improvement in the coverage performance (see Fig. 5.13). In Fig. 5.13, in addition to the normalized exclusion zone radius, we provide the percentages of transmitters from each operator that are not permitted to use the shared spectrum (they can only utilize the spectrum band belonging to their operator). As we can observe in Fig. 5.13(a), the elimination of the nearest interferers ($R = 0.1$) does not yield a significant improvement to the coverage performance when the networks are independently distributed. A significant improvement to the coverage performance may only be observed if at least 50% of the transmitters of the other operator are removed ($R > 0.2$). When clustering is applied (see Fig. 5.13(b)) the situation changes and improvements to the coverage performance can be observed even for smaller exclusion zone radii ($R \leq 0.1$). Interestingly, when large exclusion zone radii are applied ($R \leq 0.3$), the coverage performance of independently distributed network is slightly better than the LGCP-distributed network. This is related to the fact that the LGCP-distributed network models both clustering and repulsion. So while the clustering interferers will be removed at a relatively small exclusion zones radius, the interferers being repulsed will require larger exclusion zone radii and will likely be distributed closer to the reference user than the PPP-distributed interferers. Obviously, for co-located networks (see Fig. 5.13(c)) exclusion zones prohibit any form of spectrum sharing.



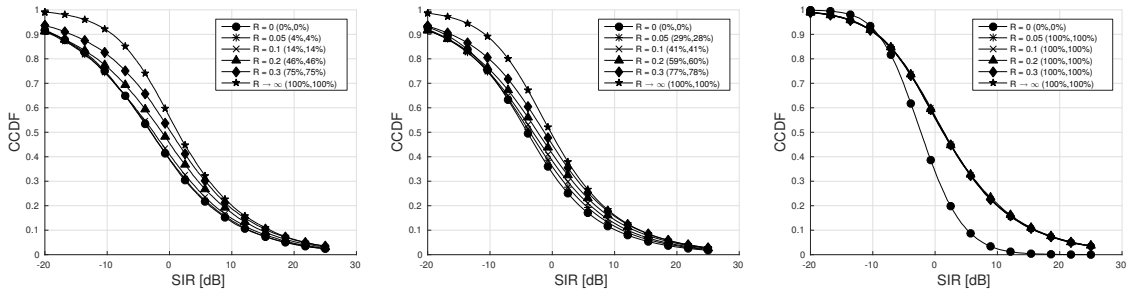
(a) Independent networks (PPP) (b) Clustered networks (LGCP) (c) Co-located networks (GPP, $u \to 0$)

Figure 5.13: The coverage probability in the coordinated spectrum sharing scenario with channel bonding, where R denotes the normalized exclusion zone radius and the percentages in brackets represent the number of "shutdown" transmitters from each operator when operating in a spectrum band of another operator.

When instead of channel bonding we apply best channel selection to shared spectrum usage (see Fig. 5.14), we can observe that exclusion zones do not impact the performance as significantly as in the channel bonding case, and it is only for high SIR–values that the coverage probability is improved. We also see that exclusion zones fare better when some-level of spatial clustering between transmitters is involved. In general, we can say that spatial coordination in spectrum sharing provides more improvement to coverage of the networks that exhibit some forms of clustering.

Average data rate

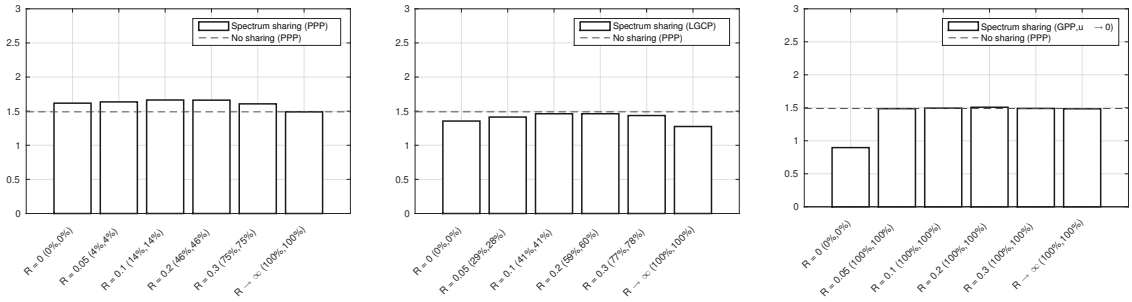
We have already shown that applying exclusion zones may lead to significant improvements to network coverage. Quite remarkably this improvement does not deteriorate the average data rate!



(a) Independent networks (PPP) (b) Clustered networks (LGCP) (c) Co-located networks (GPP, $u \to 0$)

Figure 5.14: The coverage probability in the coordinated spectrum sharing scenario with best channel selection, where R denotes the normalized exclusion zone radius and the percentages in brackets represent the number of "shutdown" transmitters from each operator when operating in a spectrum band of another operator.

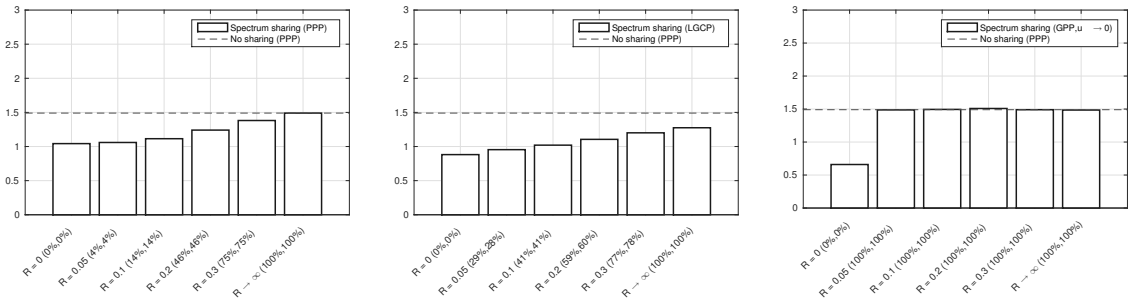
As we can see in Fig. 5.15 imposing an exclusion zone that eliminates as much as 75% of the transmitters has almost no effect on the average data rate. At first it may sound counter-intuitive, yet the exclusion zones affect not only the set of interferers, but also the probability that our tagged transmitter will be able to use shared spectrum. With spatially applied coordination these two effects seem to cancel each other out, and we receive almost constant average data rate for a range of exclusion zone radii. Contrary to coverage, more clustering has a negative impact on the average data rate as we can see in Fig. 5.15(b) and Fig. 5.15(c).



(a) Independent networks (PPP) (b) Clustered networks (LGCP) (c) Co-located networks (GPP, $u \to 0$)

Figure 5.15: The average user data rate in a coordinated spectrum sharing scenario with channel bonding, where R denotes the normalized exclusion zone radius and the percentages in brackets represent the number of shut down transmitters from each operator when operating in a spectrum band of another operator.

When best channel selection is applied to coordinated spectrum sharing (see Fig. 5.16), we can observe that the average data rate is significantly reduced (even in respect to the no sharing case) due to increase in interference. Application of the exclusion zones does not readily translate into an increase in the data rate, and large exclusion are required to result in any increase in the data rate. Interestingly, for the LGCP distributed network (see Fig. 5.16(b)) we can observe that the data rate offered by a network of a single operator is reduced with respect to a single-operator PPP-distributed network. That is related to the spatial clustering inherent to LGCP-distributed transmitters.



(a) Independent networks (PPP) (b) Clustered networks (LGCP) (c) Co-located networks (GPP, $u \to 0$)

Figure 5.16: The average user data rate in a coordinated spectrum sharing scenario with best channel selection, where R denotes the normalized exclusion zone radius and the percentages in brackets represent the number of shut down transmitters from each operator when operating in a spectrum band of another operator.

5.5 Conclusion

In this chapter, we have assessed the fundamental trade-offs between spectrum and infrastructure sharing in wireless mobile networks at various degrees of spatial correlation between the networks of sharing operators. What we have observed is that each type of sharing has its own distinctive trait, which we summarize in Fig. 5.1. While spectrum sharing alone provides an increase in the achievable data rate (at the expense of coverage), infrastructure sharing gives additional degrees of freedom when choosing the point of attachment to the network, resulting in improved coverage. When the two sharing types are combined, we observe significant increase in the average user data rate and improvements to the network coverage as compared to the non-sharing scenario (yet, lower coverage than in the pure infrastructure sharing scenario). Consequently, infrastructure and spectrum sharing cannot be simply substituted for each other, as there exists a trade-off in the coverage and data rate performance between the two. Moreover, combination of the two approaches does not linearly scale the respective coverage and data rate gains; yet it provides respectable improvement to both over the non-sharing case, which together with the potential for cost reductions may make it a commercially attractive solution.

Additionally, we observed that the way that spectrum is shared makes a difference to both the coverage and data rate performance. In particular, when coverage is the main concern, it is more beneficial for operators to pool their spectrum bands and allow for their transmitters to dynamically select the spectrum channel, rather than simply bond the available channels. Also, the relative spatial distribution, and the relative density of the networks of the sharing operators, have significant impact on the performance, and need to be taken into account when evaluating the potential benefits of network sharing. In particular, when the networks of two mobile operators are closely spatially correlated, then any gains obtained from aggregating spectrum are significantly reduced due to increase in interference. This effect may be combated by the application of coordinated spectrum sharing, which especially for networks with more spatial clustering may significantly improve the perceived coverage, without hampering the data rate performance. Eventually, the imbalance in the density of the networks of the sharing operators has a significant impact on how the two operators perceive sharing gains. Counter-intuitively, it is the larger operator that gets the most benefits from spectrum sharing, in terms of user data rate. The improvements to network coverage through

infrastructure sharing are the highest when the networks of the two operators are similarly-sized.

6 OTT-MNO sharing

OTT-MNO sharing

"We're still behaving like the rebel alliance, but now we're the Empire."

- Pete Warden, *Pete Warden's blog*

Over-the-top service providers, such as Google or Facebook, are asserting their role as the new driving force in the developments in architectures of wireless mobile networks. The recently announced Google Project Fi is, in fact, an example of a new type of a player in the mobile networks market, namely one that is simultaneously an over-the-top service provider, an infrastructure provider and a mobile virtual operator. From that perspective it is important to understand the potential of the existing networks to satisfy the requirements of these new players, as well as the strategies that they may take up to provide mobile service to their users. In this chapter we make a contribution to the state-of-the-art in modelling: the efficiency of minimum cost selections made by over-the-top service providers over shared wireless mobile networks, and the incentives for radio access infrastructure deployments by over-the-top service providers. The lion's share of this chapter is based on our two articles [42, 115].

6.1 Introduction

The new technological challenge ahead for mobile network operators (MNOs) is to meet a 1000–fold increase in demand for mobile capacity [8]. However, meeting this challenge requires that MNOs face the hard facts of ever-increasing infrastructure costs [123] and popularity of over-the-top services which "extract revenue that would otherwise be generated by the MNO from traditional voice and data services" [11]. Ensuring that mobile networks provide adequate capacity for over-the-top services is of particular concern for over-the-top service providers (OTTs), such as Google, Skype or Netflix, which rely on mobile networks to offer their services and content to their mobile users. Taking Google as our prime example, we see an increasing push among the OTTs to assert their role in shaping how mobile networks are accessed and deployed. The recently announced Google Project Fi brings about a new type of a player to the mobile networks market, namely one that is an OTT, an infrastructure provider and an Mobile Virtual Network Operator (MVNO) at the same time. Subscription with Fi allows owners of Nexus 6 smartphones to connect with the Internet using unlicensed WiFi infrastructure, contracted or deployed by Google, or, if WiFi is unavailable, the cellular network of T-Mobile or Sprint.

It is precisely this kind of behaviour that we were envisioning as part of the concept of Networks without Borders (NwoB) [10, 11]. One of the fundamental aspects underlying Networks without

Borders is the ability of service providers to dynamically select a set of wireless access points from the pool of all available resources[1] that meets the requirements of the end-user or the service, while, perhaps, bearing the lowest possible cost for an operator. These requirements may have a variety of forms, from the basic quality of service aspects such as coverage or capacity, to more complex ones such as indoor coverage or low latency, or a combination thereof. Herein, we focus solely on studying a first tractable formulation of the process, which we term coverage shaping. While in the case of Google Project Fi these resource selection decisions are performed in a distributed manner[2], the coverage shaping is a centralized mechanism, which provides us with upper bounds on the cost efficiency. Another aspect of the NwoB ecosystem, which we study in this chapter, is related to capacity extensions of the existing infrastructure deployments. Assuming that these extensions happen via network densification[3], i.e., small cell deployments, OTTs face a dilemma of whether to invest in *their own radio access infrastructure*, similarly to Google's WiFi deployments made in some North American cities, or enter into an service level agreement (SLA) with an MNO, following the Google Project Fi example. In the former case the OTT bears the full costs of the required unlicensed radio access infrastructure, while in the latter case it pays recurring charges to the MNO. The MNO, in turn, may invest this additional revenue in the densification of its network, for example, by deploying small cells. The two topics are slowly getting recognition and research work that addresses these and similar concerns is finally finding its outlet in research papers.

Let us start by noting that, in its most general form, the process of coverage shaping resembles the well-studied process of network planning, and more specifically coverage planning. Network planning involves selecting locations to install base stations, setting configuration parameters, and assigning carrier frequencies to meet the coverage and capacity demand of a given geographical area. The objective of coverage planning is to select base station locations so as to provide sufficient signal quality (under pre-specified interference levels) over a target area with minimum network cost [80]. A number of models and approaches to the coverage planning problem can be found in [79]. The more up-to-date version of the problem is how to allocate resources in a shared virtualized wireless network. In [125] the authors consider the problem of determining the optimal set of resources to be leased from multiple infrastructure providers, which can then be sliced among different service providers to satisfy their demand. The authors model the problem as a two-stage sequential stochastic allocation. In [126] a matching (through hierarchical auction) is found between spatially distributed users and frequency resources available at the nearest base station belonging to some infrastructure provider. This matching is intermediated by virtual operators that interact with a number of infrastructure providers on behalf of their subscribers. In [127] the physical infrastructure, represented as virtualized frequency-time resource blocks, is shared among multiple virtual operators. Then, allocation of such virtualized infrastructure to the virtual operators is made by means of dynamic greedy network embeddings. Similarly, in [128] a network embedding approach is considered to dynamically split virtualized resource blocks between multiple public safety operators such that resource occupancy is maximized. Physical resource blocks are also allocated among service providers in [129]. This time, however, multiple service providers cooperate in a bankruptcy game to allocate virtualized resources provided by multiple infrastructure providers such that the number of resources allocated is proportional to the traffic and fairness is ensured by means of the Shapley value [87].

[1] These resources may include assets of the current mobile networks, remote radio heads, small cells, commercial WiFi networks or household access points.

[2] The end-user device prioritizes WiFi connectivity over cellular.

[3] Network densification is considered as the most effective way to date to increase over-the-air throughput (see [124]).

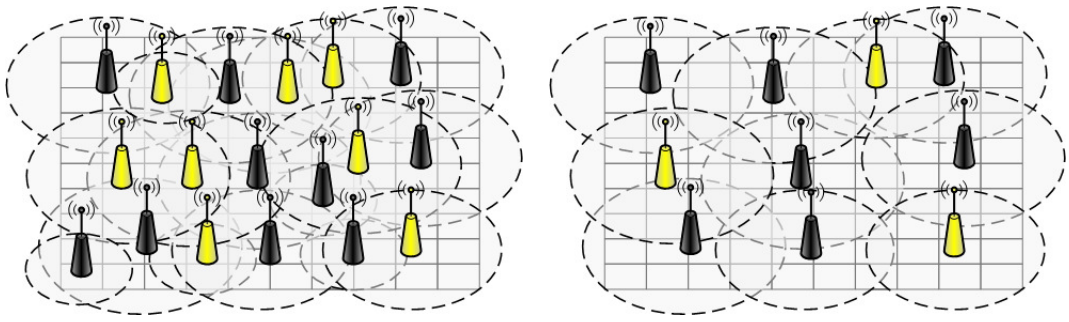
Similar type of questions are also considered within the scope of multi-tenant heterogeneous networks [130], where virtualized resources are leased between spatially adjacent infrastructure to deal with spatio-temporal variations in demand intensity.

The majority of models relevant to the dilemma that OTTs face, i.e., whether to invest in their own radio infrastructure or to enter into an SLA with an MNO, have been proposed and studied in connection with the "net neutrality" discussion. In [6] we find a summary of various possible frameworks for the interaction between OTTs and MNOs. These interactions can be classified based on the pricing regime applied, i.e., *one-sided*, where no additional fees are collected from service providers, or *two-sided*, where service providers are charged according to the terms of bilateral SLAs with infrastructure providers. Among one-sided pricing models we can find the *status quo model*, where best effort traffic of various OTTs is subject to traffic engineering techniques which limit certain types of services, for example, by blocking peer-to-peer traffic, or the *strict net neutrality model*, where network operators apply non-discriminatory traffic management, yet they cap the available capacity on a per user basis, or the *user-tiering model*, where the network operator charges users for different Quality of Service (QoS) levels delivered. Two-sided pricing involves OTTs paying either a guaranteed traffic fee corresponding to the guaranteed QoS level (*service provider-tiering model*) or a termination fee which depends on the volume of OTT traffic that terminates in the MNO network (*termination fee model*). Leaving aside the political (and business) implications, each of these approaches has its strengths and weaknesses, and can possibly be combined into more sophisticated pricing schemes, for example, the channel quality-based pricing [131]. We construct our model based on two-sided pricing and service provider-tiering. This and similar pricing schemes have been applied widely to study the impact of regulation on the wired networks and the Internet market (see [132–134]). For example, [133] studies a model with paid prioritization of content (as well as various forms of market regulations) under a system that maximizes end-user utility. A two-sided pricing model with side payments is studied in [134] where both content providers and Internet service providers charge users a subscription fee, with content providers generating additional advertisement-related revenue. Another model with side payments, with the presence of multiple competitive Internet service providers and a monopolistic content provider is studied in [135]. In [132] a similar model is studied, along with the dynamics of the relationship between Internet service providers and content providers. A somewhat related problem is that of shared deployments made by multiple infrastructure providers: in [116] multiple operators decide whether and how to share new deployments, with the problem being modelled as a mixed integer program (MIP). In addition, the OTT–MNO interaction may also be projected onto a more conventional relationship between an entrant MNO and an incumbent MNO, in which case the former may enter into a national roaming agreement with the latter [136]. In our work we model the OTT–MNO interaction using the axiomatic bargaining framework (see Chapter 3).

The remainder of the chapter is split into two parts. In the first part, we study a stochastic optimization model to deal with the problem of minimum cost resource selection in pooled wireless networks made by service providers. In the second part, we analyze the OTT–MNO interaction modelled as a cooperative game over a spatially distributed demand.

6.2 Efficiency of infrastructure pooling under minimum service requirements

Wireless systems provide radio access by means of a set of wireless access points (in cellular networks these correspond to base stations), which transmit and receive radio signals to/from end-user mobile terminals within a specified service area. We define a *service area* as a set of geographical points in which the received signal strength from the wireless access point in consideration is higher than the reception threshold, and the received interference is kept below a pre-specified level [80]. The size of the service area depends on wireless access point configurations, such as transmit power, carrier frequency or antenna tilt, as well as on radio channel effects, such as pathloss, multipath fading and shadowing. In the context of NwoB, wireless access points and their corresponding service areas are virtualized to enable "on the fly" (the time-scale of this process should reflect the daily changes in radio resource usage) reconfiguration of wireless networks. These wireless networks can be reconfigured in line with a range of objective functions. One such objective is to meet the desired coverage at a minimum base station cost, which we term *coverage shaping*. In general, coverage shaping can be stated as follows: given a set of geographical points in an area, called a universe, and a set of wireless access points (with their corresponding configurations), each of which covering a subset of elements of the universe (service areas), find a set of wireless access points with a minimum associated cost such that the union of their service areas will contain all the elements of the universe. Figure 6.1(a) depicts a multi-operator wireless network prior to coverage shaping, while Figure 6.1(b) shows a minimum cost instance of the same network. An alternative definition of coverage shaping may be stated as follows: find a minimum number of wireless access points such that the union of their service areas will contain a specified fraction of the elements of the universe to meet a pre-specified service reliability. We will refer to the former as *full coverage shaping*, while to the latter as *partial coverage shaping*.



(a) Prior (b) After

Figure 6.1: Coverage shaping on a pooled wireless infrastructure.

In the following section we formulate the full and partial coverage shaping problems as extensions to a classical combinatorial optimization formulation (the set cover problem (SCP), see [77]) which we solve using state-of-the-art MIP solvers. Eventually, we study the efficiency of minimum cost selections over pooled real cellular infrastructure in a number of locations in Poland, as well as over cellular infrastructure distributed according to some stochastic processes.

6.2.1 Network model and problem formulation

Let us define an extreme radio access network infrastructure sharing scenario in which a number of wireless access points, deployed independently by a set of infrastructure providers $\mathcal{M} \neq \emptyset$, are added to resource pool $\mathcal{S} = \bigcup_{m \in \mathcal{M}} \mathcal{S}_m$, where $\mathcal{S}_m \neq \emptyset$ is a set of wireless access points deployed and owned by infrastructure provider $m \in \mathcal{M}$. Each wireless access point $j \in \mathcal{S}$ provides wireless radio coverage corresponding to configured center frequency $f \in \mathcal{F}_m$ and transmit power $t \in \mathcal{T}$, where $\mathcal{F}_m \neq \emptyset$ is a set of non-interfering center frequencies used by infrastructure provider m, and $\mathcal{T} \neq \emptyset$ is a set of available transmit power classes. Let us also denote c_j as the cost of including wireless access point j in the instantiation of a wireless network.

Let us define an arbitrary geographical area which the operator wishes to cover. Acknowledging the advantages of systematic sampling [80], we split the area into a rectangular grid, which consists of a set of pixels $\mathcal{P} \neq \emptyset$, such that each pixel is of size r-by-r km where $r \in \mathbb{R}_+$ is the resolution of the grid. For simplification, we assume that the considered area lies in two-dimensional Euclidean space, and each pixel is represented by its centroid.

Full coverage shaping

In a cellular system, a communication link between a base station and a mobile terminal needs to satisfy some minimum QoS. Starting from analogue systems, this QoS level was strictly related to the signal-to-interference ratio (SIR) threshold, which accounted for inter-cell interference [97]. In 4^{th} generation (4G) systems this threshold relies on the signal-to-noise and interference ratio (SINR) measured by a mobile terminal, which is used to optimize system capacity and coverage for a given transmission power. If the measured SINR for a particular location is lower than the threshold, no reliable transmission link to the base station may be established [137]. Following this, we assume a mobile terminal located in pixel $i \in \mathcal{P}$ may establish a communication link with wireless access point $j \in \mathcal{S}$ only if the long-term average SINR for the access point's received signal satisfies the following condition:

$$\gamma_i = h_{ij} t_j (\eta_i + \sum_{k \in \mathcal{S}, k \neq j} \mathbb{1}_{[f_j = f_k]} h_{ik} t_k)^{-1} \geq \gamma^*, \tag{6.1}$$

where t_j denotes the downlink transmit power of wireless access point j, and h_{ij} denotes the channel gain between pixel i and wireless access point j, which consists of fading term w_{ij} and pathloss term l_{ij}. η_i denotes the additive noise power received by a mobile terminal located at pixel i. We assume this additive noise term to be normalized, i.e. $\eta_i = 1, \forall i \in \mathcal{P}$. Moreover, the indicator function $\mathbb{1}_{[f_j = f_k]}$ ensures that we account only for interference from wireless access points operating in the same frequency.

The QoS constraint (6.1) is a non-linear constraint, and, hence, might be difficult to handle. However, assuming that the pooled resources come strictly from the existing cellular networks we can note two key facts: 1) networks of different MNOs use exclusively licensed and, hence, non-interfering frequencies, 2) each of the operators has already optimized its network, and if the current frequency and power configurations are not altered the interference thresholds shall not be violated. Based on these assumptions we can re-state constraint (6.1) as the receiver sensitivity constraint:

$$r_i = h_{ij} t_j \geq r^*. \tag{6.2}$$

Receiver sensitivity level r^* and transmit power t_j are predefined network parameters, while channel state information h_{ij} can be inferred from the existing network measurements. By pre-computing the available values, and using indicator function $\mathbb{1}_{[r_i/r^* \geq 1]}$ we obtain the binary incidence matrix $\mathbf{A}$, where a single entry a_{ij} specifies whether or not wireless access point j covers pixel i. Having obtained the incidence matrix $\mathbf{A}$ we can formulate the full coverage shaping problem as follows:

$$\min_{\{x_j\}} \sum_{j \in \mathcal{S}} c_j x_j \tag{6.3}$$

subject to:

$$\sum_{j \in \mathcal{S}} a_{ij} x_j \geq 1 \quad \forall i \in \mathcal{P}, \tag{6.4}$$

$$x_j \in \{0, 1\} \quad \forall j \in \mathcal{S}, \tag{6.5}$$

which is the classical SCP formulation [77], where constraint (6.4) ensures that every point of a geographical area belongs to at least one service area, while constraint (6.5) ensures that the minimum cost choice is a binary decision made for each wireless access point.

Partial coverage shaping

Full coverage provisioning may be perceived as a special case of the partial coverage provisioning problem. In the partial coverage provisioning problem, constraint (6.4) is relaxed to allow fractional coverage service requirements. This can be achieved by replacing the right-hand side vector of ones in (6.4) with binary random vector ξ, where component ξ_i denotes a service request stemming from point i at a given point in time:

$$\xi_i = \begin{cases} 1, & \text{one or more service requests originate from point } i, \\ 0, & \text{otherwise.} \end{cases} \tag{6.6}$$

In partial coverage provisioning, the goal is to meet the new coverage constraint with probability at least $\alpha \in (0, 1]$, which represents some pre-specified service reliability level. Hence, (6.3)-(6.5) can be re-formulated as follows:

$$\min_{\{x_j\}} \sum_{j \in \mathcal{S}} c_j x_j \tag{6.7}$$

subject to:

$$\mathbb{P}\left(\sum_{j \in \mathcal{S}} a_{ij} x_j \geq \xi_i \right) \geq \alpha \quad \forall i \in \mathcal{P}, \tag{6.8}$$

$$x_j \in \{0, 1\} \quad \forall j \in \mathcal{S}, \tag{6.9}$$

where $\mathbb{P}(\cdot)$ denotes probability, and (6.8) is a joint probabilistic constraint, as used in [138]. The resulting formulation is a stochastic SCP, which in general might be hard to solve. However, following the remark from [138], if we assume independence of the components of ξ, then the deterministic equivalent integer program can be found as follows. Note that (6.8) is equivalent to

the existence of a vector $z \in \{0,1\}^{|\mathcal{P}|}$ such that $\mathbf{A}x \geq z$ and:

$$\mathbb{P}\big(z \geq \xi\big) \geq \alpha, \tag{6.10}$$

where $\mathbb{P}(\cdot)$ is a joint probability of $|\mathcal{P}|$ independent events. Taking the logarithm of both sides of (6.10) we arrive at the following inequality:

$$\sum_{i \in P} \ln\Big(\mathbb{P}\big(z_i \geq \xi_i\big)\Big) \geq \ln(\alpha). \tag{6.11}$$

Observing that $\mathbb{P}\big(1 \geq \xi_i\big) = 1$ and $\mathbb{P}\big(0 \geq \xi_i\big) = \mathbb{P}\big(\xi_i = 0\big)$, and assuming that $\mathbb{P}\big(\xi_i = 0\big)$ (the probability that no service request will originate from pixel i) is known, we arrive at the following combinatorial reformulation of (6.7)-(6.9):

$$\min_{\{x_j, z_i\}} \sum_{j \in \mathcal{S}} c_j x_j \tag{6.12}$$

subject to:

$$\sum_{j \in \mathcal{S}} a_{ij} x_j \geq z_i \quad \forall i \in \mathcal{P}, \tag{6.13}$$

$$\sum_{i \in \mathcal{P}} (1 - z_i) \ln\Big(\mathbb{P}\big(\xi_i = 0\big)\Big) \geq \ln(\alpha), \tag{6.14}$$

$$x_j \in \{0,1\} \quad \forall j \in \mathcal{S}, \tag{6.15}$$

$$z_i \in \{0,1\} \quad \forall i \in \mathcal{P}. \tag{6.16}$$

Similarly to (6.3)-(6.5), the resulting formulation is a combinatorial optimization problem for which we can apply the same set of solvers.

6.2.2 Numerical results

In this section we present exact coverage shaping results obtained with the CPLEX ILOG tool for two datasets consisting of real and random network deployments.

Optimization tools

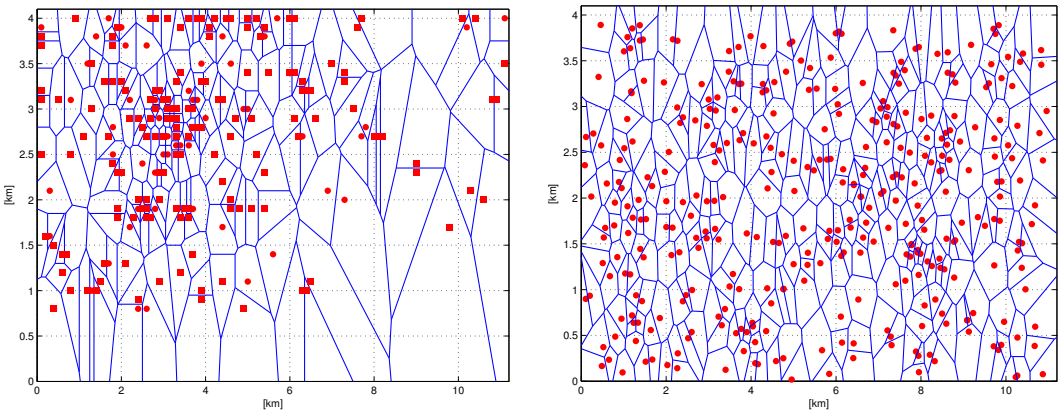
There are several exact and heuristic approaches to solving set covering problems [77]. Since both of our problems fall under the category of MIP models, we decided to rely on commercial software (IBM ILOG CPLEX 12.5 solver) that offers customizable solvers for general MIP problems and is available for use under academic licensing. In order to find an exact solution to a MIP, the CPLEX solver applies an optimized version of a branch-and-cut algorithm. This approach works well for both full and partial coverage problems, and exact solutions to reasonably large problem instances (approximately 1000 variables and 3000 linear constraints) can be found in acceptable time-scales (in the range of an hour). We ran the solver on a computer with 8 Gbytes of RAM, and an Intel Core i7-3740QM 2.70 GHz CPU with 64-bit instruction set, 8 logical cores and a cache of 6 Mbytes. The version of Linux used was 3.5.0-31-generic x86_64 GNU/Linux. The solver was set to run

deterministically the maximum number of available threads with the relative MIP gap tolerance of 3% for the full coverage problem and 5% for the partial coverage problem, and the absolute upper limit on the size of the branch-and-cut tree of 100×1024 Mbytes. The other parameters were used according to their default algorithmic settings. In the full coverage case, before a problem instance is fed to the solver we preprocess it by removing redundancies and dominations inside incidence matrix $\mathbf{A}$ [77].

Results

To evaluate the efficiency of the inter-operator infrastructure sharing we test the proposed coverage shaping models against real network deployment data from Poland, as well as against a randomly generated dataset. The real data we have contains base station deployment information from the four major MNOs operating in the Polish mobile market, namely Orange, Play, Plus and T-Mobile. For each of the MNOs we have extracted snapshots of base station geo-location data (from Groupe Special Mondiale (GSM)/Universal Mobile Telecommunications System (UMTS) radio licenses listed on-line at [139]) for four different cities that differ in deployment density (the average number of base stations deployed per unit area) and in the mix of commercial and residential areas: Warszawa, Wrocław, Olsztyn, and Świdnica. Each snapshot is a 10.9-by-4.2 km grid (which corresponds to zoom 14 in Google Maps) with a resolution of 100 meters, centered at the city center, that contains base station locations for three radio technologies: GSM in 900 MHz (GSM900), GSM in 1800 MHz (GSM1800), and UMTS in 2100 MHz (UMTS2100).

In addition, we also present results for a random deployment of base stations, following a Poisson point process (for this random geographical distribution model, we used the same intensity of base stations per technology and operator as we observed from the real deployment data). The real and random deployment for Wrocław are depicted in Figures 6.2(a) and 6.2(b), respectively; qualitatively, the deployment density in the two cases exhibits different properties, i.e. real deployment tends to have a much larger peak density centred around the central area, while (not surprisingly!) random deployment is more uniform.



(a) Empirical (urban area) (b) Random

Figure 6.2: Network deployment for Wrocław with Voronoi tessellation (dots denote single base stations, while squares denote co-located ones). Random network deployment with Voronoi tessellation (using the same intensity of base stations as in Figure 6.2(a)).

For both datasets, we model the propagation according to the modified Hata model [140], with

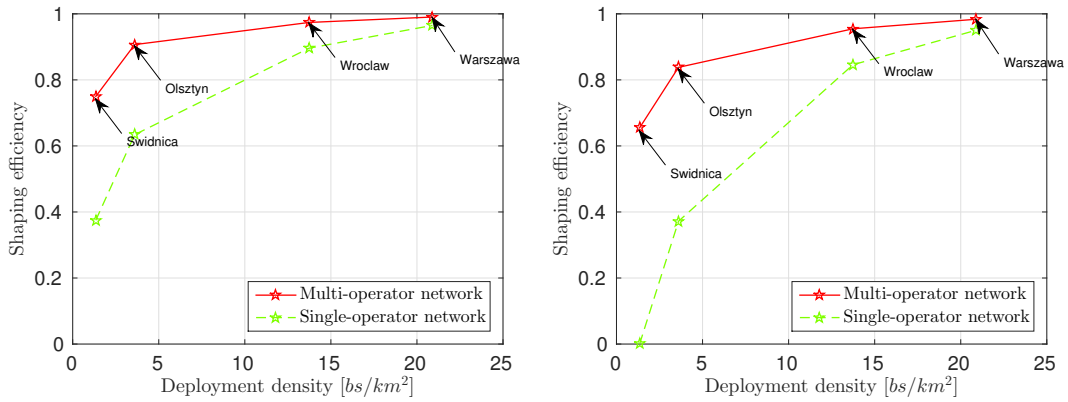
base station height of 10 meters, and mobile terminal height of 1.5 meter. In our tests, we assume two types of base station downlink transmit power allocation schemes: a) homogeneous, where base stations are assigned omnidirectional macrocell coverage with an equivalent isotropic radiated power (EIRP) of 45 dBm, b) heterogeneous, where base stations are assigned omnidirectional macro- or microcell coverage (decided based on the inter-base station distance) with EIRP of 45 and 38 dBm respectively. To ensure fair comparison for the homogeneous case we assume that all base stations have a fixed unitary cost of usage, while in the heterogeneous case we assume the cost is directly proportional to transmit power.

To quantify the gains of coverage shaping, we look at *coverage shaping efficiency*, which we define as unity minus the cost of the shaped network divided by the cost of the baseline network. For the formulation in question, the cost of the baseline network is simply the cost of all network elements belonging to the MNOs, while the cost of the shaped network represents the total cost of the minimum cost subset of the network elements that provides the desired coverage. The following analysis is intended to describe limitations and provide arguments for the feasibility of base station selections made under the scope of inter-operator infrastructure sharing and, more generally, NwoB.

Figures 6.3(a) and 6.3(b) show the full coverage shaping efficiency for the two power allocation schemes and the two datasets. We immediately observe that the highest efficiency improvements can be achieved for the areas with higher deployment density. In the densest real case (Warszawa) the gains of coverage shaping were as high as 98%. This means that a single operator, in order to provide full macrocell coverage, requires only 2% of all the existing base stations in the area (i.e., 16 out of 934). This gain is also very high in comparison to the cost of a single-operator network for both homogeneous and heterogeneous power allocations, which shows how much redundancy is introduced to the network in order to secure the capacity and possibly ensure indoor macro-, microcell coverage. For the lowest density case (Świdnica), the benefits of sharing arise as a result of the collective resource usage, which may justify potential benefits of infrastructure sharing in less urbanized areas. In this case a single operator may not need to roll-out all the base stations required to fully cover the considered area; instead, infrastructure sharing among several operators may make it economically feasible to provide service to such typically under-served areas.

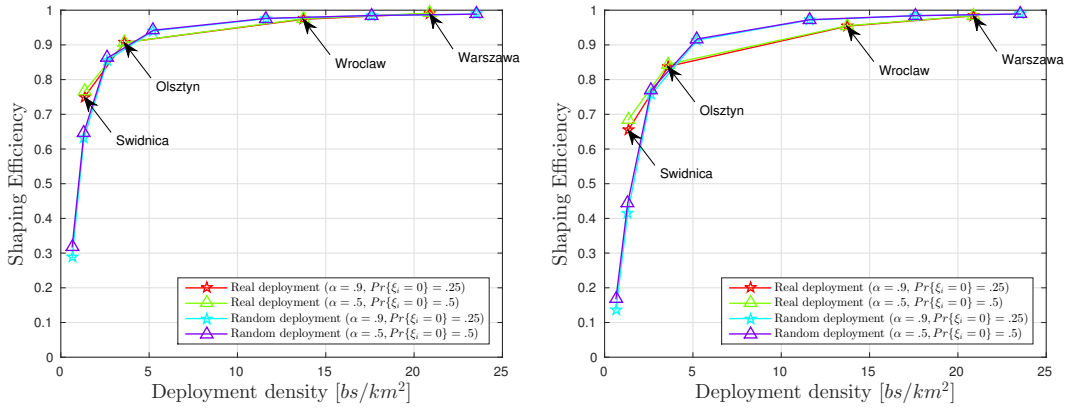
Analyzing the differences in results between homogeneous and heterogeneous power allocations, we can observe that the efficiency is reduced with the power allocated to the system. This reduction is inversely proportional to the deployment density. In addition, if we compare the results for real and random deployments, we can note that for low deployment densities (below approximately 5 bs/km^2) the gains for random and real deployments diverge. This may be due to the fact that real multi-operator deployments were planned to provide coverage for the given area following more or less the same coverage pattern for every operator (hence, resulting in redundancies), while the random deployments are more spread, resulting in lower redundancy.

Figures 6.4(a) and 6.4(b) show the partial coverage shaping efficiency. Our first observation is that relaxing the full coverage assumption brings only moderate improvements to the efficiency and, hence, the total cost of the system. This happens due to the fact that the full covering provides very few overlaps (especially for areas of high deployment density). Effectively, every base station covers large sets of pixels independently, which highly increases the probability that at least one pixel will require service. Hence, no "fringe areas" policy can be enforced even if we relax the service reliability level. One way to enforce moderately higher savings is to increase the pixel resolution maintaining the same probability of a service request or by pre-specifying the fringe areas.



(a) Homogeneous power allocation (b) Heterogeneous power allocation

Figure 6.3: Full coverage shaping efficiency for real and random deployments with homogeneous and heterogeneous power allocations.



(a) Homogeneous power allocation (b) Heterogeneous power allocation

Figure 6.4: Shaping efficiency of partial covering for real and random deployments with homogeneous and heterogeneous power allocations, with varying service reliability α and probability of no service request $\mathbb{P}(\xi_i = 0)$.

6.3 OTT-MNO interaction and service-driven deployment

As outlined in the introduction, there are two entities in our systems an OTT and an MNO. Following the example of Google Project Fi, we assume that OTTs, in addition to providing content, will provide wireless connectivity to their mobile subscribers. In order to do so the OTTs will have to decide whether to build their own infrastructure or enter into an SLA with some (or a group of) MNO, to let the MNO serve their demand. The decision is not binary, as the OTT may decide to serve a fraction of its demand using its own infrastructure, while the remaining demand will be served by the MNO and can be subject to a per unit of served traffic charge decided by the MNO. Effectively, the MNO may need to expand its network (for example, by deploying new small cells) to host the subscriber base of the OTT with the required quality, and cover the resulting costs. The decisions made by the OTT and MNO are coupled, i.e., the OTT will decide to serve a certain portion of its demand through the MNO network only if the price per traffic unit is acceptable, while the MNO will set the price to a certain level in response to the volume of the offloaded traffic, and considering factors, such as the costs of deployment and operation of its own infrastructure.

The decisions made by both OTTs and MNOs are not independent, as whatever the decision one party makes, it affects the outcome observed by the other party. A natural tool to model this type of interaction is game theory, and cooperative game theory, in particular. In cooperative games players typically have an incentive to enter into a cooperative agreement, yet they may have conflicting views on the specifics of such an agreement. In our case the OTT requires radio access infrastructure, which the MNO already has, while the MNO seeks additional revenues to cope with the increasing infrastructure costs. At the same time the OTT would like to utilize the infrastructure as cheaply as possible, while the MNO collects revenues from the usage of its infrastructure and charges the OTT (or its subscribers) accordingly. In the remainder of this section we propose the aforementioned model of cooperation between an OTT and an MNO, for which we define the corresponding utility space and the non-cooperative point related to the Nash equilibrium. Then, using the Nash bargaining framework [87], we derive an analytical expression for the Nash bargaining solution (NBS) of the corresponding game. What we show is that the spatial distribution of mobile demand has a significant impact on the outcome of cooperation, and that it unevenly affects the utilities achieved by OTTs and MNOs. Moreover, the cost of MNO's infrastructure, if excessively high, may render cooperation between an OTT and an MNO ineffective.

6.3.1 OTT-MNO interaction model

Game model

We model the interaction between an OTT and an MNO in the following way. The OTT has to provide service to a set of spatially distributed mobile demand points $\mathcal{S} \subset \mathbb{R}^2$. Each demand point is associated with some volume of demand $r \in \mathbb{N}$, and the marked set of demand points is denoted as $\widehat{\mathcal{S}} = \{(s, r_s)\}$, with the total demand (per OTT) of $R = \sum_{r \in \widehat{\mathcal{S}}} r_s$. In order to provide service to its demand points the OTT may either deploy its own radio access infrastructure (at a unit cost c_o) or enter into an SLA with the MNO. In the latter case the OTT has to pay the per traffic unit charge decided by the MNO for the demand it decides to serve with the MNO network. The MNO serves the OTT's demand using the existing Long Term Evolution (LTE) macrocell infrastructure, which bears no additional deployment costs, or, if that is not possible (due to lack of coverage or capacity), using newly deployed small cell infrastructure at a unit cost c_m. Additionally, we denote Q as the revenue received by the OTT from its mobile subscribers. Since we are interested in studying the OTT-MNO interaction we assume that Q is fixed and no further assumption about the OTT subscribers needs to be made.

From the system's description we formulate the game model as follows: $\mathcal{N} = \{\text{OTT}, \text{MNO}\}$ is the set of players; $[0,1] \times [0,1]$ is the action space, where $\theta \in [0,1]$ is the decision of the OTT, i.e., the fraction of the demand to be served by its network, in case it decides to deploy its own infrastructure (while the rest gets served by the MNO network), and $\beta \in [0,1]$ is the decision of the MNO, i.e., the per traffic unit charge (normalized as described in the following paragraph); and $\Omega = \big(u_o(\beta,\theta), u_m(\beta,\theta)\big)$ is the utility space, with $u_o(\beta,\theta)$ denoting the utility obtained by the OTT and $u_m(\beta,\theta)$ denoting the utility obtained by the MNO.

Up to now we have defined traffic-related measure r_s and various cost-related parameters, yet we have not discussed their representation. First, we consider all of our economic parameters to be per unit of traffic served during a certain time frame. This means, for example, amortizing the

deployment cost of a base station over an appropriate period. This approach allows us to construct utility functions that combine revenues and costs related to the user traffic with the costs of base station deployments. Second, all our economic parameters are normalized against c_o, which allows us to simplify the utility functions used and make the model independent of particular choices of monetary values. An additional implication of this assumption is that the traffic charge β, when normalized against c_o, can be limited to $[0, 1]$. The rationale is as follows: when the cost of offloading a unit of traffic is higher than the cost of deploying a base station per unit of traffic, then the OTT is more likely to deploy a base station than to offload its traffic to the MNO. Ultimately, the traffic can be represented as bytes, packets, flows, etc., with the selection of the traffic unit being independent from our model.

Utility space

We define $u_o(\beta, \theta)$ as follows:

$$u_o(\beta, \theta) = -m_o(\theta, \widehat{\mathcal{S}}) - \beta(1-\theta)R + Q, \tag{6.17}$$

where $m_o(\theta, \widehat{\mathcal{S}})$ denotes the number of base stations to be built by the OTT, given θ and demand $\widehat{\mathcal{S}}$. The rationale for this expression of the utility is straight–forward – the OTT has to pay for any new infrastructure it decides to build and/or pay the traffic fee for the demand served by the MNO network. In return the OTT receives revenue from its subscribers either directly, as part of a subscription fee, or indirectly, from customized advertising.

The MNO obtains the following utility:

$$u_m(\beta, \theta) = -c_m m_m(\widehat{\mathcal{S}}_{\texttt{res}}) + \beta(1-\theta)R, \tag{6.18}$$

where $m_m(\widehat{\mathcal{S}}_{\texttt{res}})$ denotes the number of base stations to be built by the MNO to serve the demand $\widehat{\mathcal{S}}_{\texttt{res}}$, where $\widehat{\mathcal{S}}_{\texttt{res}} = f(\theta, \widehat{\mathcal{S}})$, with $f : [0, 1] \times (\mathbb{R}^2, \mathbb{N}) \to (\mathbb{R}^2, \mathbb{N})$, is the residue of the OTT's demand not served by the OTT's infrastructure, which we discuss in the following paragraph. The MNO pays all the deployment costs associated with the deployment of additional base stations (we assume that the existing infrastructure bears no deployment costs) required to satisfy the residual demand, while at the same time receiving revenue proportional to the served traffic.

The utilities of both the OTT and MNO depend on the number of new base stations required to satisfy the demand, i.e., $m_o(\theta, \widehat{\mathcal{S}})$ and $m_m(\widehat{\mathcal{S}}_{\texttt{res}})$. In its generality the problem of deploying wireless access infrastructure is a complicated one consisting of various planning stages (see [95]) and requiring various smart ways to find solutions to NP-hard problems (see [79]). Since our goal is to study the relationship between an OTT and MNO, instead of directly modelling the deployment decisions (i.e., whether a base stations should be built in some given location), we are interested in aggregate decisions, such as "How many base stations are required given certain volume of spatially distributed demand?" In other words, we seek a mapping between the demand and a counting measure of the number of base stations deployed to serve the demand, i.e., $m_o(\theta, \widehat{\mathcal{S}})$ and $m_m(\widehat{\mathcal{S}}_{\texttt{res}})$, which hereafter we will refer to as the *deployment function*. We determine the deployment function for both the OTT and MNO empirically by applying the following procedure (detailed description of the procedure is provided in Appendix C).

First, we generate the set of demand points $\mathcal{S}$ as the realizations of the Gauss-Poisson process

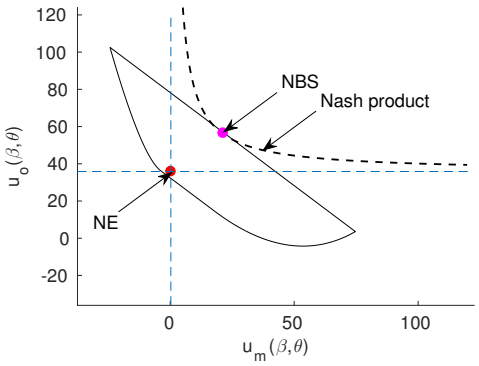


Figure 6.5: The utility space for our model obtained from drawing a convex hull over utility pairs calculated for the action space of the MNO and OTT varied over $[0,1] \times [0,1]$, for a fixed clustering level and infrastructure price; marked points denote special operating points (NE - Nash Equilibrium, NBS - Nash Bargaining Solution); while the dashed curve corresponds to the Nash product curve.

[50], with the parameter cluster probability that allows us to vary the level of clustering in $\mathcal{S}$. Subsequently, we extrapolate real-world call-detail record information from an Irish mobile operator into a continuous demand field, which we sample using points from $\mathcal{S}$ to obtain $\widehat{\mathcal{S}}$. Effectively, we construct a demand model with real demand volume information and varying level of clustering, which we control through the cluster probability. We use $\widehat{\mathcal{S}}$ to feed an optimization model which minimizes the deployment cost, given that θ fraction of demand $\widehat{\mathcal{S}}$ needs to be served using the OTT's infrastructure. The outcome of the model is a set of WiFi access points which the OTT has to deploy, and the residual demand $\widehat{\mathcal{S}}_{\mathtt{res}}$, which will be offloaded to the MNO. To find the set of new infrastructure required by the MNO to provide coverage to $\widehat{\mathcal{S}}_{\mathtt{res}}$ we use an optimization model which minimizes the deployment cost, subject to covering all of $\widehat{\mathcal{S}}_{\mathtt{res}}$ with either new small cell or existing macrocell infrastructure (in both cases we assume single-antenna LTE system). This optimization model also considers the requirement that the MNO must serve its own subscribers' demand. Now, for each value of cluster probability, we obtain a curve that represents the number of base stations required as a function of θ, to which we can fit some well-known functions. We obtained the most satisfactory fits with the quadratic function of the form $a_0 + a_1\theta + a_2\theta^2$, where the coefficients vary with the cluster probability (the obtained coefficients are given in Tab. 6.1). We will use $a_{\{0,1,2\}}$ to denote the coefficients of $m_o(\theta, \widehat{\mathcal{S}})$ and $b_{\{0,1,2\}}$ to denote the coefficients of $m_m(\widehat{\mathcal{S}}_{\mathtt{res}})$. Since our demand model $\mathcal{S}$ is stochastic, the deployment function should be interpreted as the average number of base stations to be built for a given clustering level of spatially distributed demand.

The point of this whole procedure was to determine a closed-form representation for the deployment functions, such that some level of analytical tractability is achieved. Yet, our utility space Ω, resulting from the postulated utility functions, is not convex. However, if correlated mixed strategies are permitted, then the considered utility space becomes convex [87]. An example utility space for our model is illustrated in Fig. 6.5. It was obtained from drawing a convex hull over utility pairs calculated for the action space varied over $[0,1] \times [0,1]$. The point marked as NE denotes the Nash equilibrium, which corresponds to a non-cooperative solution to our game model, while the NBS corresponds to the cooperative solution, which we find analytically in Sec. 6.3.2.

Disagreement point

In any cooperative game we need to define the disagreement point, which is the point of minimal utility (payoff) expected by each of the players if bargaining (negotiations) breaks down. Obviously, this implies that if the cooperation provides lower payoff(s) to any one of the players, she will not be interested in cooperating, and will settle for the minimal utility. Often times it makes sense to choose as the disagreement point, the point related to the Nash equilibrium (NE) of the non-cooperative game. The NE corresponds to a set of strategies selected by players such that no individual player can increase her utility by unilaterally deviating.

Proposition 6. *The Nash equilibrium strategy* $(\beta_{NE}, \theta_{NE})$ *corresponds to the following pair of actions:*

$$\beta_{NE} = 1.0,$$
$$\theta_{NE} = \begin{cases} \frac{R-a_1}{2a_2}, & \text{if } a_2 > 0, \\ \text{argmax}(u_o(\beta_{NE}, 0.0), u_o(\beta_{NE}, 1.0)), & \text{otherwise.} \end{cases}$$

Proof: Let us start by observing that, given $\beta \in [0,1]$, the MNO has a dominant strategy $\beta_{\text{NE}} = 1.0$ as $u_m(\beta_{\text{NE}}, \theta) \geq u_m(\beta', \theta)$ for all $\beta' \in [0,1]$ and $\theta \in [0,1]$. Given the dominant strategy of the MNO, we find the equilibrium strategy of the OTT by maximizing its utility, which leads to strategy $\theta_{\text{NE}} = \frac{R-a_1}{2a_2}$, if $a_2 > 0$, or $\text{argmax}(u_o(\beta_{\text{NE}}, 0.0), u_o(\beta_{\text{NE}}, 1.0))$ otherwise, which concludes the proof of the proposition. ■

The pair of outcomes in the disagreement point corresponds to the element $\big(u_o(\beta_{\text{NE}}, \theta_{\text{NE}}), u_m(\beta_{\text{NE}}, \theta_{\text{NE}})\big)$. This disagreement point should be interpreted in the following way: a rational MNO has no incentive to set any other fee than β_{NE}, therefore forcing an OTT to deploy either a full infrastructure or a significant portion of it.

6.3.2 Bargaining solution

In the following we derive a formula that describes the Nash bargaining solution to our game. An NBS is the unique solution to the following optimization problem:

$$\text{argmax} \prod_{i \in \mathcal{N}} (u_i(\beta, \theta) - d_i), \tag{6.19}$$

subject to:

$$u_i(\beta, \theta) \geq d_i, \quad \forall i \in \mathcal{N}, \tag{6.20}$$
$$\beta, \theta \in [0,1], \tag{6.21}$$

where $d_i = u_i(\beta_{\text{NE}}, \theta_{\text{NE}})$ for each $i \in \mathcal{N}$. In this formulation constraints Eq. (6.20) ensure that our solution is no worse than the disagreement point, while constraint Eq. (6.21) ensures that the strategies available to the players are feasible.

Theorem 1. *Given, the problem in Eq. (6.19), the NBS is:*

$$(\beta_{NBS}, \theta_{NBS}) = \left(\frac{Q - a_0 + c_m b_0 + d_m - d_o}{2R}, 0.0 \right). \tag{6.22}$$

Proof: Given Ω, which is non-empty, compact and convex, we can find the maximum of Eq. (6.19) by first taking the logarithm of the product in Eq. (6.19) and then applying the Lagrangian multiplier method:

$$L(\beta,\theta,\mu,\eta,\gamma)=\sum_{i\in\mathcal{N}}\log(u_i(\beta,\theta)-d_i)-\sum_{i\in\mathcal{N}}\mu_i(d_i-u_i(\beta,\theta))+\eta_o\theta+\eta_m\beta-\gamma_o(\theta-1)-\gamma_m(\beta-1), \tag{6.23}$$

where $\mu_i \geq 0$, $\eta_i \geq 0$, $\gamma_i \geq 0$. Then, the necessary and sufficient Karush-Kuhn-Tucker (KKT) conditions for optimality are:

$$\nabla_\beta L(\beta,\theta,\mu,\eta,\gamma)=0, \tag{6.24}$$

$$\nabla_\theta L(\beta,\theta,\mu,\eta,\gamma)=0, \tag{6.25}$$

$$\mu_i(d_i-u_i(\beta,\theta))=0, \quad \forall i\in\mathcal{N}, \tag{6.26}$$

$$-\eta_o\theta\leq 0, -\eta_m\beta\leq 0,$$

$$\gamma_o(\theta-1)\leq 0, \gamma_m(\beta-1)\leq 0, \tag{6.27}$$

where Eq. (6.24) and Eq. (6.25) are the stationary conditions, and Eq. (6.26) and Eq. (6.27) are the complementary slackness conditions. After making substitutions and calculating the derivatives, the stationary conditions yield the following forms:

$$\frac{-R(\theta-1)}{d_o-u_o(\beta,\theta)}+\frac{R(\theta-1)}{d_m-u_m(\beta,\theta)}-\gamma_m+\eta_m=0, \tag{6.28}$$

$$\frac{a_1-R\beta+2a_2\theta}{d_o-u_o(\beta,\theta)}+\frac{R\beta+c_m(b_1+2b_2\theta)}{d_m-u_m(\beta,\theta)}-\gamma_m+\eta_m=0. \tag{6.29}$$

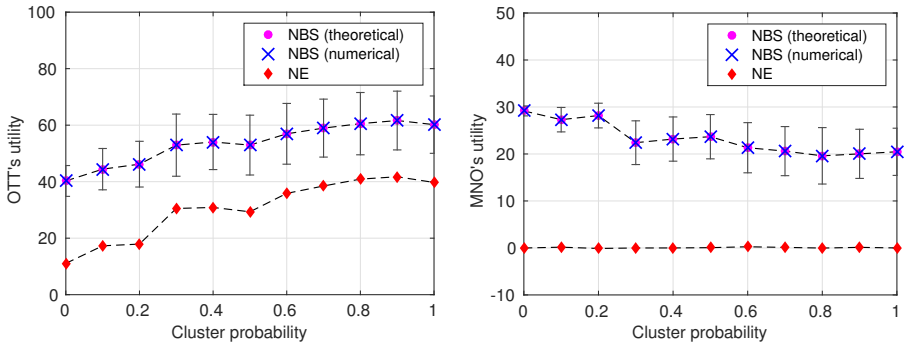
When $\mu_i = 0$, $\gamma_i = 0$, $\eta_m = 0$, and $\theta = 0.0$, we get that $\beta = (Q - a_0 + d_m - d_o + c_m b_0)/(2R)$ and $\eta_o = (2(a_1 + b_1 c_m))/(a_0 - Q + d_m + d_o + b_0 c_m)$. This solution is not only feasible (provided that $Q - a_0 + d_m - d_o + c_m b_0 < 2R$ and $Q - a_0 + d_m - d_o + c_m b_0 \geq 0$), but it happens to lie on the Pareto frontier of our utility space, precisely at the intersection of the Nash product curve and the Pareto frontier of Ω, as presented in Fig. 6.5. Any other feasible solutions will not lie on the Pareto frontier of Ω, which concludes the proof. ■

Given Theorem 1, if the OTT and MNO cooperate, the OTT serves all of its demand through the infrastructure provided by the MNO and pays the traffic fee β_{NBS}. In the following we will analyze the impact that the spatial distribution of demand and the infrastructure price have on this cooperation.

6.3.3 Numerical results and discussion

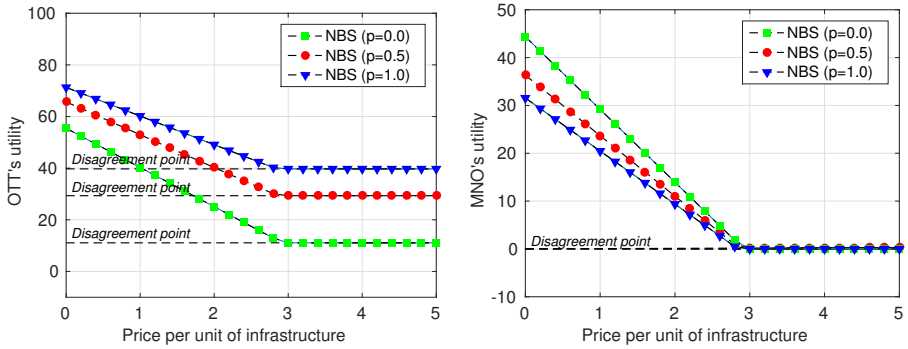
We validate our analytical solutions in Theorem 1 by applying to problem Eq. (6.19) iterative optimization methods, such as the sequential least square method. Indeed, in Fig. 6.6 and Fig. 6.7 we can see that the numerically obtained bargaining solutions match closely the ones expressed in closed-form, confirming the correctness of our derivations.

In Tab. 6.1 we present estimates of the deployment function for various demand clustering levels. Let us recall Property 6: given these estimates, we can see that $a_2 < 0$, therefore for each clustering level the NE corresponds to the pair $(\beta_{\text{NE}}, \theta_{\text{NE}}) = (1.0, 1.0)$, which means that the OTT fully deploys



(a) OTT (b) MNO

Figure 6.6: Utility obtained in the non-cooperative (NE) and cooperative (NBS) cases for different values of spatial clustering in the OTT's demand, with error bars accounting for the deviation in the number of base stations required for a given cluster probability. When the cluster probability is close to 0.0, the demand is almost uniformly distributed in space, whereas when the cluster probability approaches 1.0, the demand is generated from a number of highly concentrated clusters. The cost c_m is fixed at 1.0.



(a) OTT (b) MNO

Figure 6.7: Utility obtained from the cooperation for different price points of the MNO's infrastructure, and three values of the cluster probability p. When $c_m < 1.0$, the MNO's infrastructure is cheaper than the infrastructure deployed by the OTT. When $c_m > 1.0$, the MNO pays more for the infrastructure than the OTT does. We see that both utilities are inversely proportional to the price of the MNO's infrastructure; yet, there is a price point at which the cooperation collapses to the disagreement point.

(or contracts the deployment of) its infrastructure to serve its demand. In our numerical analysis we also assume that the revenue obtained by the OTT, i.e., Q, is fixed and equal to the total volume of the demand R.

In Fig. 6.6 we show how the utility, obtained by the OTT (Fig. 6.6(a)) and MNO (Fig. 6.6(b)), changes with the clustering of the OTT's demand. What we can immediately note from Fig. 6.6(a) is that the clustering of demand has a stronger impact on the utility obtained by the OTT. This has to do with the fact that clustering means less infrastructure is required to cover the same volume of demand, as we can observe in the NE case where the OTT deploys all the infrastructure on its own. Clearly, having to deploy less infrastructure per unit of demand makes deploying its own infrastructure a more attractive option for the OTT. To compensate for this fact an MNO has to decrease the price it charges for carrying the OTT's traffic, therefore giving up on some of its profits (even though the MNO will also have to deploy less infrastructure due to clustering of the demand).

The infrastructure cost c_m has a significant impact on the cooperation between the MNO and the OTT. In Fig. 6.7(a) and Fig. 6.7(b) we observe that the utilities of the OTT and the MNO

Table 6.1: Estimates of the deployment function for the OTT and MNO for varying levels of clustering in demand.

	OTT	MNO
Cluster probability	Parameters $\{a_0, a_1, a_2\}$	Parameters $\{b_0, b_1, b_2\}$
0.0	$\{0.1, 108.0, -12.9\}$	$\{30.8, 47.8, -17.4\}$
0.1	$\{-1.1, 134.7, -53.6\}$	$\{30.7, 47.9, -17.3\}$
0.2	$\{-3.0, 152.2, -84.4\}$	$\{26.6, 39.9, -13.1\}$
0.3	$\{-2.0, 150.0, -80.9\}$	$\{27.9, 42.3, -14.2\}$
0.4	$\{-4.3, 163.7, -100.2\}$	$\{26.7, 37.1, -10.4\}$
0.5	$\{-1.3, 151.8, -88.5\}$	$\{25.4, 35.5, -10.2\}$
0.6	$\{-3.7, 149.8, -88.1\}$	$\{23.4, 34.0, -10.5\}$
0.7	$\{-4.6, 166.4, -109.2\}$	$\{22.8, 30.1, -6.8\}$
0.8	$\{-2.4, 149.7, -81.4\}$	$\{24.3, 33.4, -9.1\}$
0.9	$\{-3.5, 162.9, -99.6\}$	$\{23.7, 29.9, -5.9\}$
1.0	$\{-5.0, 176.5, -129.3\}$	$\{20.2, 27.9, -8.0\}$

are inversely proportional to the infrastructure price, i.e., the higher the price, the lower the utility. The MNO compensates for higher infrastructure price by imposing a higher traffic fee on the OTT. When the price of MNO infrastructure becomes excessive, the cooperation is no longer feasible and the only reasonable solution for the players is to fall back to the disagreement point. The level of spatial clustering in the demand has little to no impact on the point where cooperation collapses.

Reflecting back on the achieved results, we can note two things. In real relationships between OTTs and MNOs, other considerations, besides cost, may influence the OTT's decision of whether to deploy its own infrastructure. If the expected monetary impact of these strategic considerations could be quantified, it would be possible to incorporate them into our model. On the MNO side, we do see that excessive cost of infrastructure, which in practice includes both the deployment and the spectrum licensing fees, may cause MNOs to be very protective over who can use their network and how. This may be comparable to the case of mobile virtual operators which were able to enter the European market only after favourable regulation was in place. However, falling revenues may be a motivator for MNOs to actually enter into deals with OTTs, as we can currently observe with Google Project Fi.

6.4 Conclusion

Sharing of wireless communications infrastructure is happening more and more. We envision this trend will continue towards networks constructed on-the-fly by over-the-top service providers, which will necessitate more sophisticated understanding of the efficiencies and tradeoffs involved in decision-making over aggregate wireless network resources. In this chapter we quantified the efficiencies that result, under different coverage models, in small and large cities. Using the real data we showed that considering the multi-operator network infrastructure as a collective may lead to building up coverage provisioning networks which require just a fraction of all the available base stations, even over single-operator networks. This may be especially important for the lowest density areas, which benefit highly from infrastructure sharing, in which case a single operator may not need to roll-out all the base stations required to fully cover the considered area; instead, infrastructure sharing among several operators may make it economically feasible to provide service to such typically under-served areas, which may have its effect on the digital divide.

Subsequently, we analyzed a model of cooperation between an OTT (or a group of OTTs) and an MNO working together to deploy a wireless network crafted to the needs of a service offered by the OTT. This cooperation is sustainable only if the cost of infrastructure to the MNO is not excessive. In case it significantly exceeds the price paid by the OTT for the infrastructure, cooperation is no longer feasible (the MNO is no longer able to increase the charge on the OTT, without the OTT retracting to its own deployment). In the non-cooperative case, the OTT solely deploys an unlicensed spectrum infrastructure such as WiFi. Moreover, spatial clustering of the OTT demand means less radio access infrastructure is required to cover the same volume of demand. This means deploying its own infrastructure becomes cheaper for the OTT, and therefore more attractive. Effectively, it is the OTT that benefits the most from having spatially clustered demand.

7 Broader vision

Broader vision

"A revolution is certainly the most authoritarian thing there is; it is the act whereby one part of the population imposes its will upon the other part by means of rifles, bayonets and cannon"

- Frederick Engels, *On Authority*

New trends in commercial mobile networks have emerged. These trends include the inter-operator network sharing, the formation of Mobile Virtual Network Operators (MVNOs), and the increasing role of over-the-top service providers (OTTs). Each of the trends comes at a range of flavours each leaving a distinct trait on the general consensus about wireless mobile network ownership, architecture, and regulation. The key purpose of this thesis was to quantify the impact these trends have on modelling and performance of future mobile networks.

What we have shown in this thesis is that modelling of pooled radio access networks requires new statistical models of deployment and, indeed, such models can be found, and will be representative to many similar areas in different regions. Furthermore, allowing for resource sharing, among mobile operators only, results in new, and to some extent non-intuitive, performance trade-offs, which are further exacerbated by the spatial distribution of transmitter locations. Different resource sharing schemes, such as spectrum and radio access infrastructure, when combined together, do not produce linearly scaling performance gains. Yet a simple mechanism of inter-operator interference coordination allows operators to partially compensate for this effect. Moreover, depending on the preferred spectrum sharing method, different performance gains and trade-offs can be obtained. In particular, when coverage is the main concern, it is more beneficial for operators to allow for their transmitters to dynamically select among the pooled spectrum bands, rather than simply bond the available channels.

Radio access network and spectrum sharing are key to opening up mobile networks to a new type of OTT which offers seamless mobile service that spans networks of multiple mobile operators. When network infrastructure of multiple mobile operators is considered as a collective, a coverage provisioning network may be constructed such that just a fraction of all the available transmitters is required. This network will be considerably less infrastructure-intensive, which may make the case for shared infrastructure deployments in currently under-served rural areas. Furthermore, when growth in mobile demand necessitates new network deployments, it is reasonable for this new type of an OTT to pursue these deployments in cooperation with a mobile network operator (MNO). This cooperation would rely primarily on the MNO playing to its strength and expanding the capacity of the existing network, while the OTT provides it with a new stream of customers and revenues.

This cooperation would also be conditioned on the price of the radio access infrastructure and, to some extent, also on the spatial clustering in mobile demand.

All these studies are intended to build groundwork for the change in the way that mobile networks are being owned, controlled and, ultimately, accessed. This change, if pursued further, will require a revolution in the way that mobile networks operate. Indeed, Google Project Fi can be considered an example for this type of a change whereby an over-the-top service provider becomes also an infrastructure provider, a mobile operator, and a mobile virtual operator. Networks without Borders is a reflection of this revolution and a possible path in the evolution of mobile access provisioning. As we show, the proposed vision results in a substantially different mobile communications value-chain, which results also in a number of socio-political challenges. The following text is based on our paper, published in the Proceeding of the IEEE (see [11]).

7.1 Vision

We believe that the traditional view of wireless networks as systems which deliver voice and data services to some pre-defined geographical areas no longer reflects the reality of wireless networks. We have argued for a paradigm shift by outlining some key trends that have emerged in wireless networks ownership and control model over the last few years. In this chapter we propose a vision of wireless networks which in our view emerges from these trends. There are two reasons why we believe such vision is needed.

First, a vision is a driver for innovation. As such a particular view and understanding of the role of wireless networks and the development forces at play will direct the innovation process, especially at an industrial level which often times requires society-wide consensus about the system requirements in order to proceed with the development and implementation. An example of how the current view of wireless networks limits innovation is the 3rd Generation Partnership Project (3GPP)-based standardization of integrated mobile and WiFi networks, or IP Multimedia Subsystem (IMS), which presupposes an authoritative role of mobile network operators in setting up networks and delivering reliable wireless services to users[1], as well as reinforces the split between traditional mobile services (video or data) and over-the-top services.

Second, we want to design policies and regulations that are effective in implementing our political agenda. In our view, this is no longer possible with the status quo model of coverage or capacity provisioning, as it does not reflect all the relevant actors, changes in the business models or subscriber behaviour. The spectrum licensing process is an example of a failure of policy and regulations to recognize the growing discrepancy between the owners of exclusive spectrum licenses, i.e., the parties that bear the brunt of costs of delivering mobile services, and the parties that actually extract revenues from providing of wireless services to end users.

In the following we delineate our vision, discuss its implications to the mobile communications value-chain, as well as enumerate some key challenges.

[1] Arguably, leading to a failure in widespread adoption of these technologies.

7.1.1 Networks without Borders

The concept of Networks without Borders revolves around an over-the-top service and a network that responds to change. With Networks without Borders the physical infrastructure provides physical connections between the service provider(s) and the user(s), and as such is merely a source of heterogeneous resources subject to virtualization. Virtualization enables a competent actor to dynamically partition and aggregate these resources. In consequence sets of such resources may be combined to form intermittent networks instantiated only to deliver a particular service, possibly even to a particular user. Many such networks with differing properties may coexist at any given time, as virtualization provides isolation necessary for concurrent and uninterrupted use of common resources. Networks without Borders is about re-conceiving networks from a perspective of significantly more fluid access to spectrum and all of the resources that comprise a wireless network.

In Fig. 7.1 we illustrate the key idea behind our vision. Depending on whether we read the figure from bottom-to-top or top-to-bottom we arrive at different conclusions on how future networks may be conceived. Reading it bottom-to-top, we start with infrastructure providers that are in possession of some specific resource(s), this can be a license to a spectrum band, a set of wireless access points deployed in a given area, a data center, or an authentication service. These resources, coming from individual infrastructure providers, when subject to virtualization, may be aggregated to form resource pools. A set of such resource pools constitutes all necessary and possible elements that form a wireless network. In practice, however, only a handful of resources will be selected so as to form a network, and many such networks (yet with differing properties) may be created. Eventually, a decision is made to deliver a service (or a bundle of services) via such a network, or a set of such networks, which also happens to exhibit some desired properties. This could be thought of as a virtual operator-centric design. Changing our perspective to top-to-bottom allows us to recognize that a service provider may want to design a network that suits the needs of its service (or a bundle of services). Then, the decision is to select resources such that this network (or a network of comparable properties) can be instantiated. Obviously, any selected virtual resource has to map to some realizable physical connection. Therefore a competent actor has to either re-use existing infrastructure, whenever possible, and/or deploy additional infrastructure, if the former cannot be accomplished. This can be viewed as a service-centric design. These views are not mutually exclusive and to some extent they represent a possible framework in which future service-driven networks could operate.

In essence there are five points that make our vision substantially different to the contemporary way of looking at wireless networks.

First, we see properties of future networks to be tightly coupled with specific over-the-top services and their recipients. This approach mirrors that of the Internet, in the sense that the network infrastructure merely provides physical connectivity for whatever service is transmitted. Our view of a service is more nuanced than the traditional classification involving voice, text and data in that and we think about services in a fine-grained manner. So, for example, a Netflix-like service specifically focused on video-streaming will require a specific kind of connectivity, while a bursty machine-to-machine (M2M) service will require a very different kind of connectivity. The service provider (or a party operating on its behalf) will be able to select from a menu of virtual network operators, each offering networks of different per user throughput or latency. The idea of an over-the-top service no longer exists, as the service and virtual network over which it is delivered are interconnected,

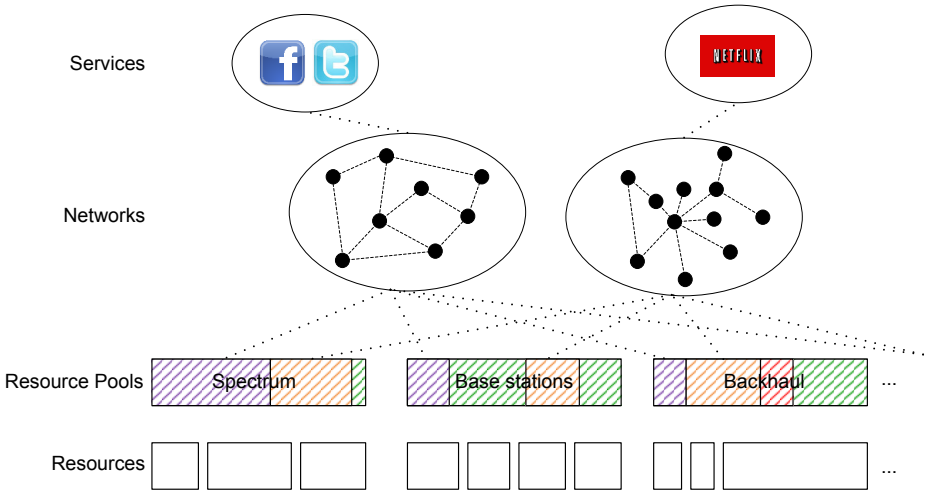


Figure 7.1: Networks without Borders explained – a vertical view.

which allows for the service provider to extract value from the network while compensating virtual operators for their services. Of course, such an approach entertains the possibility of grouping services into bundles which may be delivered jointly.

This leads to the second point, that virtualization is key. It hides from the users the notion of resource ownership, as resource pools can be partitioned and aggregated in new ways, giving the level of flexibility required to create customized networks. On a business-technological level whether this is done through bilateral agreements, as with the MNO and the MVNO relationship, or an auction mechanism becomes transparent to the user of a virtual network. Moreover, virtualization is the catalyst of sharing as it brings about isolation required to enable a concurrent use of shared resources. Virtualization may operate on a substrate of exclusive or shared resources, which may give rise to virtual networks that also operate at different price points.

Thirdly, and perhaps most significantly, the resources which make up the network are considered to come from a pool of all and every resource. This means that the borders between licensed and license-exempt spectrum or commercial and user-deployed infrastructure are removed. Spectrum is spectrum and infrastructure is infrastructure. Whatever types of resources are available or, more importantly, whatever types of resources that are appropriate for different kinds of services are used. This notion of a pool of resources leaves the door open for a wider range of contributors to the pool, and a broader call for investment in infrastructure. Given the vast need for small cells to deliver the capacities for future data consumption, it makes great sense that small cells, or indeed other resources, can be crowd-sourced from the public, for example. The "bring your own" concept can be extended. Bring your own device (BYOD) refers to the policy of permitting employees to bring personally owned mobile devices (laptops, tablets, and smart phones) to the workplace. These devices can then be used to access privileged company information and applications. We can extend this to be not just about the devices that hang on the end of the network but to include the devices/resources that constitute the network. Bring your own spectrum (BYOS), bring your own base station (BYOBS), bring your own processing power (BYOPP) are all examples of how this can be extended.

Following on from the idea of a pool of resources, opens up a way for making progress on how spectrum is accessed. More than a decade of research work has taken place which has explored a

myriad of ways in which spectrum can be accessed in a dynamic manner. There is technical work on how radios can sense white spaces, use databases for identifying spectrum, use dynamic auctions of all kinds in spot-markets or over longer time-scales, trade spectrum and much more. There is also much work on the licensed or license-exempt constructs in which these more dynamic approaches can reside. However, this has not yet translated into concrete regulatory action or change. The preferred exclusive licensing model is to be blamed for perceivable spectral scarcity, while, given our research findings to date and with the help of virtualization, we can create at least an illusion of infiniteness of the spectral resources [3].

Fifthly, our view of things allows new entities to arise which may operate on various sets of virtual resources and on various decision levels. In Fig. 7.2 we present an example realization of the simplified telecommunications market value-chain. Here we use the notion of a content provider rather than a service provider simply to invoke a broader understanding of an exchangeable (tradeable) good/service as anything ranging from provisioning of wireless access to architecting a wireless network. Also, the flow of currency may be expressed in a variety of forms dependent on the business model of entities involved in the relationship, for example, for some relationships an out-right monetary transfer can be made, yet for others, user (meta-)data may be the currency. There are a number of new roles which appear in the mobile communications value-chain[2]. First, the virtual operator is an entity that controls virtual network instances and couples a service to a particular network instance, or a set of instances. These network instances are brought to existence by network architects that select and aggregate subsets of virtualized resources which make up the physical substrate on which the virtual operators reside. Each resource may remain as part of an aggregated subset for an arbitrary period of time, and may be returned to the pool when not needed. Finally, resource aggregators manage these pools of resources. Hence, they enable partitioning and aggregation of heterogeneous resources coming from various resource (infrastructure) providers. Each role in the figure may be played by one or more entities, and a single entity may function with more than one role.

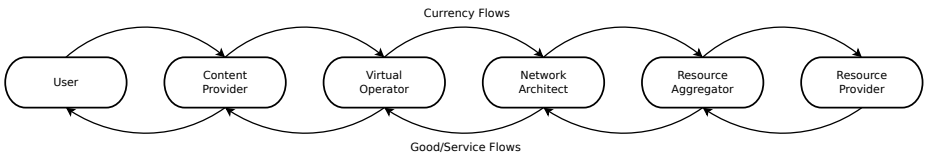


Figure 7.2: Networks without Borders – a simplified realization of the mobile communications value-chain.

All of these aspects are a mere declaration of our vision, yet what we offer here implies deep changes to how wireless services are provisioned and sold, and how virtual and physical networks are designed, deployed, and managed. These changes pose a number of challenges, which we discuss in the following.

7.2 Outlook

Data is the new currency of the information age. Companies whose very business model relies on capitalizing on user data, like Google or Facebook, not only dominate their respective markets, but also expand way beyond them, for example, towards the mobile communications market. In this

[2]More details on the conventional value-chain can be found in [3].

thesis we gave various examples, and, indeed, an entire vision, of service providers asserting their role in the mobile networks landscape. The future has caught up with our vision rather quickly and it is only last year that we have witnessed an emergence of a new type of a player in the mobile market, namely one that is simultaneously an over-the-top service provider, an infrastructure provider, a mobile operator, and a mobile virtual operator. Interestingly, this new entity arrived to the mobile market despite falling revenues of the entities already in the market, as we exemplified earlier. The new entity relies on a completely different business model that is flexible enough to allow it to alter between two sources of revenues: direct payments from subscribers and online advertising systems which capitalize subscriber data.

Mobile networks, due to their conventional ownership and control model, are a great and, to some extent, unexplored source of information about subscribers, revealing user location, mobility patterns, and even health, as mobile devices add sensors to monitor heart rate and other indicators, in addition to any user provided content. But the data is not the only reason for service providers to enter the mobile market. The other reason is control over ads distribution channels. Currently, the majority of the content available to the users of the Internet is free of charge, but it is only free of charge because, along with the content, users download ads. The point of issue is that those ads create a non-negligible overhead that needs to be transported over communications networks. For this reason some mobile operators (see [141]) are considering implementation of traffic engineering techniques (essentially ad-blocking), which would force service providers to share their ad-related revenues. However, service providers have at least two ways to avoid this scenario: implement end-to-end encryption and/or route their user traffic through preferred access networks. Both of which will be effective only if service providers secure some control over the access part of communications networks.

One promising technology that may enable service providers to exercise the required level of control over the radio access networks is network virtualization. Network virtualization allows a competent actor to dynamically partition and aggregate network resources to create virtualized networks instantiated only to deliver a particular service, possibly even to a particular user. Many such networks with differing properties may coexist at any given time, as virtualization provides isolation necessary for concurrent and uninterrupted use of common resources. Given that some service providers, such as Google, exert control over mobile devices, their operating systems, web browsers, mobile and backbone connectivity, end-to-end encryption, end-user service and even end-server hardware, we could ask, paraphrasing [142], what is the form of power and authority embodied in such virtualized networks?

Given the current knowledge, the answer to that question may only be – authoritarian monopolies/oligopolies, as there is no reason to believe that the dominance of the service providers is likely to be endangered in any immediate future [143]. The problem is that as much as we like to think of ourselves as customers to the service providers often times this is not the case, and we are typically just their products (or rather our data is). The real customers of service providers are companies that buy customized advertising channels. Therefore service providers grow stronger the more data they have about their users, and that is why Amazon is assessing the introduction of face recognition to authenticate its customer payments, that is why Facebook is experimenting with influencing its users emotional states and that is why we have Google Project Fi. And, as argued in [143], the expansion of control over user data is not just about profit maximization, it is about reinforcing the monopoly that service providers have. All of this just means we cannot simply vote

with our feet unless a political consensus on protecting our data is made.

In other words, to make sure that "connectivity is a human right", and not just a buzzword intended to promote a specific business model, action is needed and communications regulators and policy-makers have to finally recognize that there are some important aspects to be considered in respect to the evolution of mobile communication systems. First, they need to determine whether some public intervention is required to enable full network virtualization, and perhaps new infrastructure investments. If these are coupled with new policies mandating some form of spectrum sharing in the exclusively licensed spectrum bands, or opening up of new shared spectrum bands, we may invariably see the rise of new types of virtual operators as predicted in our vision. The pressure from some service providers, who will require fast expansion of the wireless network capabilities to enable a new generation of services (think, for instance, of virtual reality and self-driving cars), may yield a situation in which these virtual operators act unregulated in a market whose goal is to eventually provide connectivity to the society at large. Therefore appropriate regulations have to be placed to ensure that new actors not only harvest the profits from the market, but also provide connectivity as an egalitarian good. The Internet conglomerates, through their search engines and recommendation systems, have the means to quickly monopolize areas in which they operate, as we have already discussed, which brings us to the point that policy-makers have to re-evaluate the type of communications market they intend to promote. Eventually, all these are invariably related to the question of the expected public benefit, as compared to the status quo, and the necessary amount of regulation, and public investment, required to achieve this benefit. Similar public policy questions have been debated for the past decade in the context of dynamic access to spectrum, with concrete steps taken by regulators such as the Federal Communications Commission (FCC), in the US, and Office for Communications (Ofcom), in the UK. We advocate a broader discussion that encompasses infrastructure associated with multiple wireless access technologies, in addition to spectrum.

Moreover, we as a society have to recognize that the data coming from our mobile phone usage holds a great value to companies that capitalize on customer profiling and service customization. In a way that is good news for the low-income part of our society, as it may expect to receive cheaper, or perhaps, free of charge mobile service. But on the other hand, it is also apparent to us that due to lack of protection of users privacy we are currently witnessing the emergence of the surveillance society [144], with all its grim consequences [145]. If left unregulated, the service provider in the ecosystem of Networks without Borders gets an unprecedented access to metadata revealing user location, mobility patterns, and even health, as mobile devices add sensors to monitor heart rate and other indicators, in addition to any user provided content. Regardless of the intent of the service providers, it is in our best interest to ensure that any future discussion of privacy entertains the possibility that the same service provider may hold data on both usage patterns (as is currently true of Google, Facebook, etc.) and rich information about user routines, health and instantaneous location. Effectively, a dividing line may need to be drawn between sensitive and non-sensitive data, specification of minimum security for data storage and monitoring may need to be set in place, as well as rules defining terms of resale, processing and usage of the data. At the same time, new business models should be discussed that incentivize private entities to utilize data in a manner that produces social benefit.

All of this discussion goes way beyond the technical questions addressed in this thesis, and is, in fact, a reflection of our belief that applied science should be held accountable for its research. It is not enough to insist that (sub-)optimal solutions are "best guarantees of democracy, freedom, and

social justice" [142], while in reality the designs we make benefit only a handful of people and lead to a world in which the Internet corporations have even tighter control over our lives. We can and should do better than that. Even if we will, things may still go the wrong way, but then we will seek solace in an affordable, or perhaps even free, mobile access.

Appendix

A Minimum contrast estimation

In the following we derive expressions for parameters of the point processes analyzed in Chapter 4, applying the method of minimum contrast to the theoretical expression for pair-correlation function $g(r)$ and its non-parametric estimate $\widehat{g}(r)$.

A.1 log-Gaussian Cox process

For an Log-Gaussian Cox process (LGCP) we find the variance σ^2 for the stationary Gaussian process and the scale parameter β of the correlation function $f_\beta(r)$ by minimizing the following contrast function:

$$\int_{r_{min}}^{r_{max}} \left(\log(\widehat{g}(r))^q - (\sigma^2 f_\beta(r))^q\right)^2 dr. \tag{7.1}$$

Assuming fixed $\beta > 0$ and taking the square of the expression inside the integral, we achieve the following expression:

$$\int_{r_{min}}^{r_{max}} \log(\widehat{g}(r))^{2q} dr - 2(\sigma^2)^q \int_{r_{min}}^{r_{max}} \log(\widehat{g}(r))^q \left(f_\beta(r)\right)^q dr + (\sigma^2)^{2q} \int_{r_{min}}^{r_{max}} \left(f_\beta(r)\right)^{2q} dr. \tag{7.2}$$

Since Eq. (7.1) is a quadratic function of $(\sigma^2)^q$, the extremum of this function can be found through the first derivative test:

$$\int_{r_{min}}^{r_{max}} \log(\widehat{g}(r))^q \left(f_\beta(r)\right)^q dr - (\sigma^2)^q \int_{r_{min}}^{r_{max}} \left(f_\beta(r)\right)^{2q} dr = 0. \tag{7.3}$$

Assuming that $B_L(\beta) = \int_{r_{min}}^{r_{max}} \log(\widehat{g}(r)^q \left(f_\beta(r)\right)^q dr$ and $A_L(\beta) = \int_{r_{min}}^{r_{max}} \left(f_\beta(r)\right)^{2q} dr$, the above expression attains its minimum for:

$$\widehat{\sigma}^2 = \left(\frac{B_L(\beta)}{A_L(\beta)}\right)^{1/q}. \tag{7.4}$$

Given the second derivative test, $B_L(\beta)$ has to be larger than zero for the minimum of Eq. (7.1) to exist.

In order to derive β one needs to plug in Eq. (7.4) into Eq. (7.2), in which case one obtains:

$$\widehat{\beta} = \operatorname{argmax}\left(\frac{B_L(\beta)^2}{A_L(\beta)}\right). \tag{7.5}$$

Given $\widehat{\sigma}^2$, an estimate of the mean can be calculated as $\widehat{\mu} = \log(\widehat{\lambda}) - \widehat{\sigma}^2/2$, where $\widehat{\lambda}$ denotes a non-parametric estimate of the intensity function of the point process.

A.2 Matérn cluster process

In order to find a representative model for the Matérn cluster process (MCP) we need to find the parameters R and κ that minimize the following contrast function:

$$\int_{r_{min}}^{r_{max}} \left(\widehat{g}(r)^q - \left(1 + \frac{1}{\kappa}H(R)\right)^q\right)^p dr, \tag{7.6}$$

where $H(R) = (4\pi Rr)^{-1}h(\frac{r}{2R})$. Fixing $p = 2$ and $q = 1$, assuming fixed R and following similar logic to the LGCP case allows us to express Eq. (7.6) as:

$$\int_{r_{min}}^{r_{max}} \left(\widehat{g}(r)^2 - 2\widehat{g}(r) + 1\right) dr - \frac{2}{\kappa}\int_{r_{min}}^{r_{max}} (\widehat{g}(r) - 1)H(R)dr + \frac{1}{\kappa^2}\int_{r_{min}}^{r_{max}} H(R)^2 dr. \tag{7.7}$$

Similarly to the above, Eq. (7.6) is a quadratic function of κ^{-1}, thus, the extremum of this function can be found through the first derivative test:

$$\int_{r_{min}}^{r_{max}} (\widehat{g}(r) - 1)H(R)dr + \frac{1}{\kappa}\int_{r_{min}}^{r_{max}} H(R)^2 dr = 0. \tag{7.8}$$

Assuming that $B_M(R) = \int_{r_{min}}^{r_{max}}(1 - \widehat{g}(r))H(R)dr$ and $A_M(R) = \int_{r_{min}}^{r_{max}} H(R)^2 dr$, we can find parameter $\widehat{\kappa}$ that minimizes the above expression:

$$\widehat{\kappa} = \frac{A_M(R)}{B_M(R)}. \tag{7.9}$$

Clearly, $B_M(R)$ has to be larger than zero for the minimum to exist. Then, given $\widehat{\kappa}$, $\widehat{R}$ can be found as:

$$\widehat{R} = \text{argmax}\left(\frac{B_M(R)^2}{A_M(R)}\right). \tag{7.10}$$

Finally, the mean number of points per cluster $\widehat{\mu}$ is obtained as:

$$\widehat{\mu} = \widehat{\kappa}\widehat{\lambda}, \tag{7.11}$$

where $\widehat{\lambda}$ is the estimate of the intensity of the observed point pattern.

A.3 Thomas process

Derivations for the parameters of the Thomas process, i.e., σ^2 and κ, follow the same logic as in the MCP case and yield the following results:

$$\widehat{\kappa} = \frac{A_T(\sigma^2)}{B_T(\sigma^2)}, \tag{7.12}$$

where $B_T(\sigma^2) = \int_{r_{min}}^{r_{max}} (1-\widehat{g}(r))N(\sigma^2)dr$, $A_T(\sigma^2) = \int_{r_{min}}^{r_{max}} N(\sigma^2)^2 dr$, $N(\sigma^2) = (4\pi\sigma^2)^{-1}\exp(-\frac{r^2}{4\sigma^2})$, and:

$$\widehat{\sigma}^2 = \text{argmax}\left(\frac{B_T(\sigma^2)^2}{A_T(\sigma^2)}\right). \tag{7.13}$$

Finally, $\widehat{\mu} = \widehat{\kappa}\widehat{\lambda}$, with $\widehat{\lambda}$ being the estimate of the intensity of the observed point pattern.

B Coverage probability derivation

In the following we derive expressions for the coverage probability and the average data rate for the resource sharing scenarios presented in Chapter 5.

B.1 Infrastructure sharing

Let us recall that the superposition of Poisson point processes (PPPs) is also a PPP, with the intensity equal to the sum of the intensities of the component processes [121][Proposition 1.3.3]. Effectively, in the infrastructure sharing scenario the point process that describes the serving base station is Φ with intensity $\lambda = \sum_{i\in\mathcal{N}} \lambda_i$, and the transmitter association policy, as stated in the system model (see Chapter 5), based on the nearest distance. The probability distribution function (pdf) of the distance to the closest transmitter in the Euclidean plane (obtained from [146][Theorem 1]) can be expressed as follows:

$$f_R(r) = 2\pi\lambda r \exp(-\lambda\pi r^2). \tag{7.14}$$

Now, in the infrastructure sharing scenario the interference to which the user is exposed to depends on the network to which the user is connected (since each network operates in an exclusive spectrum band). To find the coverage probability over all networks of all the operators we need to find the expectation over the coverage probability of each individual network, which requires defining the probability of associating with the network of operator $n \in \mathcal{N}$. This association probability for a user with the network of operator n in the infrastructure sharing scenario can be found as follows:

$$\begin{aligned}
\mathfrak{A}_n &= \mathbb{E}\Big[\min_{j\in\mathcal{N}\setminus\{n\}} R_j > R_n \Big| R_n\Big], \\
&= \mathbb{E}\Big[\prod_{j\in\mathcal{N}\setminus\{n\}} \mathbb{P}\left(r < R_j\right)\Big| R_n = r\Big], \\
&= \int_0^\infty \prod_{j\in\mathcal{N}\setminus\{n\}} \mathbb{P}\left(r < R_j\right) f_{R_n}(r)dr,
\end{aligned} \tag{7.15}$$

where $f_{R_n}(r)$ is the distribution of distance from the reference user to the nearest transmitter of n, and $\mathbb{P}\left(r < R_j\right)$ may be interpreted as the probability that there is no transmitter of j closer than a distance r from the user; this can be also referred to as the null probability for network j, which is $\exp(-\pi\lambda_j r^2)$. After plugging the null probability for network j and Eq. (7.14) into Eq. (7.15), we

get that:

$$\begin{aligned}\mathfrak{A}_n &= 2\pi\lambda_n \int_0^\infty r \exp\left(-\pi r^2 \sum_{j\in\mathcal{N}} \lambda_j\right) dr, \\ &= \frac{\lambda_n}{\sum_{j\in\mathcal{N}} \lambda_j}. \end{aligned} \tag{7.16}$$

Now, the coverage probability for the infrastructure sharing scenario can be expressed as (using the expression in Eq. (3.15)):

$$p(\theta) = \sum_{i\in\mathcal{N}} \mathfrak{A}_i p(\theta|i), \tag{7.17}$$

where $p(\theta|i)$ is the coverage probability conditioned on connecting to the network of operator i, which can be expressed as:

$$p(\theta|i) = \int_{0^+}^\infty \exp(-\theta r^\alpha W)\mathcal{L}_{I_i}(\theta r^\alpha) f'_{R_i}(r) dr, \tag{7.18}$$

where the Laplace transform $\mathcal{L}_{I_i}(\theta r^\alpha)$ of the interference in network i is described using the form in Eq. (3.16), and $f'_{R_i}(r)$ is the pdf of the distance from a user to the nearest transmitter of network i, given that it is also its serving transmitter among all networks. The latter can be derived as follows:

$$\begin{aligned} f'_{R_i}(r) &= \frac{dF'_{R_i}(r)}{dr}, \\ &= \frac{d}{dr}\left(1 - \mathbb{P}\Big(R_i > r \Big| \min_{j\in\mathcal{N}\setminus\{i\}} R_j > R_i\Big)\right), \\ &\overset{(a)}{=} -\frac{d}{dr}\frac{\mathbb{P}\Big(R_i > r, \min_{j\in\mathcal{N}\setminus\{i\}} R_j > R_i\Big)}{\mathbb{P}\Big(\min_{j\in\mathcal{N}\setminus\{i\}} R_j > R_i\Big)}, \\ &\overset{(b)}{=} \frac{d}{dr}\frac{1}{\mathfrak{A}_i}\int_{-\infty}^{r} \prod_{j\in\mathcal{N}\setminus\{i\}} \mathbb{P}\left(R_j > r\right) f_{R_i}(r) dr, \\ &= \frac{2\pi\lambda_i}{\mathfrak{A}_i} \exp\Big(-\pi r^2 \sum_{j\in\mathcal{N}} \lambda_j\Big) r, \end{aligned} \tag{7.19}$$

where $F'_{R_i}(r)$ denotes the cumulative distribution function; (a) the denominator was derived in Eq. (7.16); (b) the inner probability is the null probability of network j, i.e., $\exp(-\pi\lambda_j r^2)$. Finally, after making the necessary substitutions, we get the following expression:

$$p(\theta) = 2\pi \sum_{i\in\mathcal{N}} \lambda_i \int_{0^+}^\infty \exp\Big(-\theta r^\alpha W\Big) \exp\Big(-2\pi r^2 \Big(\sum_{j\in\mathcal{N}} \lambda_j + \lambda_i \mathfrak{Z}(\theta,\alpha)\Big)\Big) r dr. \tag{7.20}$$

The derivation of the average user data rate requires plugging the obtained formula into Eq. (3.21), which concludes the proof. Note that the special case of this result for two operators was derived in [147].

B.2 Spectrum sharing with channel bonding

Let us start by observing that the point process describing the closest transmitter to a reference user of operator n is Φ_n with intensity λ_n. Therefore, following the same logic as previously, the probability of finding the closest base station at distance r is:

$$f_{R_n}(r) = 2\pi\lambda_n r \exp(-\lambda_n \pi r^2). \tag{7.21}$$

In the spectrum sharing scenario the interference to a user of operator n comes from: (i) all the transmitters of operator n that are at a distance r and further apart, which we denote as $I_r(\Phi_n)$, and (ii) all the transmitters of all operators other than n, which may be at any arbitrary distance to the reference user, which we denote as $I_0(\Phi_j)$ and $j \neq n$. Now, the total interference is $I = I_r(\Phi_n) + \sum_{j\neq n} I_0(\Phi_j)$, and its Laplace transform can be derived as follows:

$$\begin{aligned}
\mathcal{L}_I(s) &= \mathbb{E}_I\left[\exp(-sI)\right], \\
&= \mathbb{E}_I\Big[\exp(-sI_r(\Phi_n))\exp(-s\sum_{j\neq n} I_0(\Phi_j))\Big], \\
&\overset{(a)}{=} \mathbb{E}_{I_r(\Phi_n)}\Big[\exp(-sI_r(\Phi_n))\Big]\prod_{j\neq n}\mathbb{E}_{I_0(\Phi_j)}\Big[\exp(-s\sum_{j\neq n} I_0(\Phi_j))\Big], \\
&= \mathcal{L}_{I_r(\Phi_n)}(s)\prod_{j\neq n}\mathcal{L}_{I_0(\Phi_j)}(s),
\end{aligned} \tag{7.22}$$

where $s = \theta r^\alpha$, and (a) follows from the worst case scenario interference assumption, which allows us to treat the interference coming from the the networks of different operators as independent, and therefore replace the expectation of the product with the product of individual expectations.

Now, $\mathcal{L}_{I_r(\Phi_n)}(s)$ is given in Eq. (3.16), while $\mathcal{L}_{I_0(\Phi_j)}(s)$ can be obtained as follows:

$$\begin{aligned}
\mathcal{L}_{I_0(\Phi_j)}(\theta r^\alpha) &= \exp\Big(-\pi r^2\lambda_j\theta^{2/\alpha}\int_{0^+}^{\infty}\frac{1}{1+u^{\alpha/2}}du\Big), \\
&= \exp\Big(-\pi r^2\lambda_j \mathfrak{Z}_0(\theta,\alpha)\Big),
\end{aligned} \tag{7.23}$$

where $\mathfrak{Z}_0(\theta,\alpha) = \theta^{2/\alpha}\Gamma(1+2/\alpha)\Gamma(1-2/\alpha)$. Finally, when we plug Eq. (3.16) and Eq. (7.23) into Eq. (7.22), and, subsequently, to Eq. (3.15), we get the following expression for the coverage probability of operator n:

$$p_n(\theta) = \pi\lambda_n\int_{0^+}^{\infty}\exp(-\theta v^{\alpha/2}W)\exp(-\lambda_n\pi v)\exp\Big(-\pi v\Big(\lambda_n\mathfrak{Z}(\theta,\alpha)+\mathfrak{Z}_0(\theta,\alpha)\sum_{j\neq n}\lambda_j\Big)\Big)dv. \tag{7.24}$$

Calculation of the average user rate across the shared spectrum requires observing that the expectation of a sum of random variables is a sum of expectations of each individual random variable. Hence, given the expression in Eq. (3.21), we get that the average user data rate in the spectrum sharing scenario is:

$$\tau_n = |\mathcal{N}|\cdot\mathbb{E}_\gamma\Big[p_n\Big(\exp(\gamma/b)-1\Big)\Big]. \tag{7.25}$$

where $p_n(\cdot)$ is the coverage probability from Eq. (7.24), which concludes the proof.

B.3 Spectrum sharing with best channel selection

When instead of channel bonding operators select between possible bands such that a diversity gain can be obtained, the coverage probability of a reference user of operator n in the spectrum sharing scenario with two shared bands n and m can be obtained as follows:

$$\begin{aligned}
\bar{p}_{nm}(\theta) &\overset{(a)}{=} 1-\mathbb{P}\Big(\bigcup_{i\in\{n,m\}} \mathrm{SINR}_i \leq \theta_i\Big),\\
&= 1-\mathbb{P}\Big(\bigcup_{i\in\{n,m\}} h_x^i \leq \theta_i/l(x)\,(W_i+I_i)\Big),\\
&\overset{(b)}{=} 1-\mathbb{E}_{I_n}\Big[1-\exp\Big(-\theta_n/l(x)\big(W_n+I_n\big)\Big)\Big]\mathbb{E}_{I_m}\Big[1-\exp\Big(-\theta_m/l(x)\big(W_m+I_m\big)\Big)\Big]\\
&\overset{(c)}{=} \mathbb{E}_{I_n}\Big[\exp\Big(-\theta_n/l(x)\big(W_n+I_n\big)\Big)\Big]+\mathbb{E}_{I_m}\Big[\exp\Big(-\theta_m/l(x)\big(W_m+I_m\big)\Big)\Big]-\\
&\quad \mathbb{E}_{I_n}\Big[\exp\Big(-\theta_n/l(x)\big(W_n+I_n\big)\Big)\Big]\mathbb{E}_{I_m}\Big[\exp\Big(-\theta_m/l(x)\big(W_m+I_m\big)\Big)\Big], \qquad (7.26)
\end{aligned}$$

where SINR_n is the signal-to-noise and interference ratio (SINR) in spectrum band n; (a) conditioned on the distance r to the nearest transmitter in x, the probability of coverage from at least one spectrum band may be expressed in terms of the probability of outage in each shared spectrum band; (b) assuming small enough coherence bandwidth[3] and exponential fading, signal fading in different bands from the tagged transmitter is independent; (c) the first two terms are precisely the expressions for the coverage probability in the spectrum sharing scenario for operator n in each individual spectrum band n and m, while the last term, which is the joint probability of coverage in each of the spectrum bands (denoted as $p_{nm}(\theta)$), can be further simplified as follows:

$$\begin{aligned}
p_{nm}(\theta) &\overset{(a)}{=} \exp(-\theta/l(x)W)\mathbb{E}_{\Phi}\Big[\prod_{y\in\Phi}\mathbb{E}_{h^n}\Big[\exp\Big(-\theta l(y)/l(x)h_y^n\Big)\Big]\mathbb{E}_{h^m}\Big[\exp\Big(-\theta l(y)/l(x)h_y^m\Big)\Big]\Big],\\
&\overset{(b)}{=} \exp(-\theta/l(x)W)\mathbb{E}_{\Phi}\Big[\prod_{y\in\Phi}\frac{1}{\big(1+\theta l(y)/l(x)\big)^2}\Big],\\
&\overset{(c)}{=} \exp(-\theta/l(x)W)\prod_{i\in\{n,m\}}\mathbb{E}_{\Phi_i}\Big[\prod_{y\in\Phi_i}\frac{1}{\big(1+\theta l(y)/l(x)\big)^2}\Big],\\
&\overset{(d)}{=} \exp(-\theta/l(x)W)\exp\Big(-\lambda_n\int_{\mathbb{R}^2\setminus B(o,r)}\Big(1-\frac{1}{\big(1+\theta l(y)/l(x)\big)^2}\Big)dy\Big)\cdot\\
&\quad \exp\Big(-\lambda_m\int_{\mathbb{R}^2}\Big(1-\frac{1}{\big(1+\theta l(y)/l(x)\big)^2}\Big)dy\Big), \qquad (7.27)
\end{aligned}$$

where $W=W_k+W_j$ and $B(o,r)$ denotes a ball of radius r centred at the location of the reference user; (a) the set of interferers in each spectrum band is the same while the fading experienced by each interfering signal, due to our previous assumption about the coherence bandwidth, is independent; (b) given the assumption that fading is exponentially distributed; (c) we can divide the set of interferers into two independent sets: a set of transmitters of operator n and a set of transmitters of operator m; (d) using the expression for the probability generating functional (pgfl) of the PPP [50][Theorem 4.9]. Now, by switching to polar coordinates and making the necessary substitutions

[3] When coherence bandwidth is small enough, fading across different spectrum bands may be considered independent.

(and de-conditioning on r), we get that:

$$p_{nm}(\theta) = \pi\lambda_n \int_{0^+}^{\infty} \exp(-\theta r^\alpha W) \exp\Big(- \lambda_n \pi r^2 \big(1 + \mathfrak{Z}'(\theta,\alpha)\big)\Big) \exp\Big(- \lambda_m \pi r^2 \mathfrak{Z}'_0(\theta,\alpha)\Big) r dr, \quad (7.28)$$

where $\mathfrak{Z}'(\theta,\alpha) = \theta^{2/\alpha} \int_{\theta^{-2/\alpha}}^{\infty} 1 - \big(1+u^{\alpha/2}\big)^{-2} du$, and $\mathfrak{Z}'_0(\theta,\alpha) = \theta^{2/\alpha} \int_{0^+}^{\infty} 1 - \big(1+u^{\alpha/2}\big)^{-2} du$.

When no noise and $\alpha = 4$ is assumed, the joint probability of coverage probability from two spectrum bands for a user of operator n is given as:

$$p_{nm}(\theta) = \frac{\lambda_n}{\lambda_n + \lambda_n \frac{3}{2}\sqrt{\theta}\arctan\sqrt{\theta} + \lambda_n \frac{\theta}{2(1+\theta)} + \lambda_m \frac{3}{4}\pi\sqrt{\theta}}. \quad (7.29)$$

Effectively, the coverage probability for shared spectrum with best channel selection can be expressed as:

$$\bar{p}_{nm}(\theta) = 2p_n(\theta) - p_{nm}(\theta), \quad (7.30)$$

where $p_n(\theta)$ is given in Eq. (5.8), and $p_{nm}(\theta)$ in Eq. (7.29).

Applying the above formula to Eq. (3.21), allows us to obtain the average user rate, which concludes the proof.

C Constructing a deployment model for Dublin

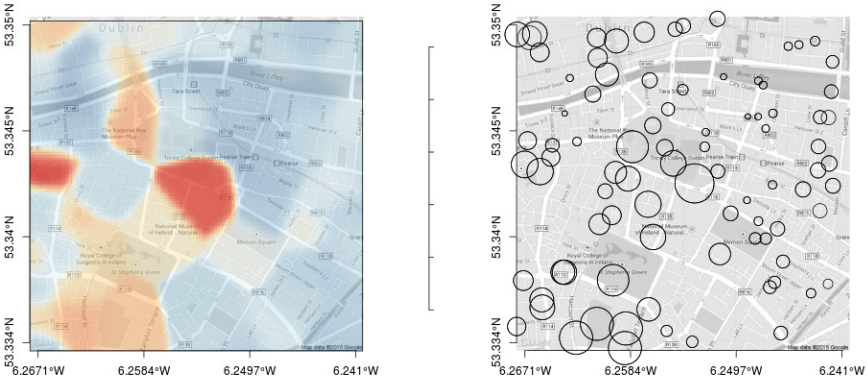
In the following we provide details on the procedure of determination of the deployment function for both the over-the-top service provider (OTT) and the mobile network operator (MNO), used in Chapter 6. The procedure consists of the following steps:

- generate the set of demand points $\mathcal{S}$;
- extrapolate real-world call-detail record information from an Irish mobile operator into a continuous demand field, and sample it using points from $\mathcal{S}$ to obtain $\widehat{\mathcal{S}}$;
- generate new base station locations for the OTT;
- generate new base station locations for the MNO;
- fit analytical functions to the curves that represent the number of base stations required as a function of θ.

The point of this procedure is to find a closed-form representation for the deployment function, such that some level of analytical tractability is achieved.

C.1 Modelling mobile demand

Our scenario is an urban scenario (the city centre of Dublin) that is rooted in present reality. Specifically, we exploit two real-world traces, provided by an Irish mobile network operator. Said traces contain call-detail records for voice calls and data sessions, collected over the entire territory of the Republic of Ireland for a period of two weeks. In order to construct our demand model we represent the location-specific demand data as a continuous demand data field (we apply the Gaussian kernel smoothing to convert our discrete information into a data field), which can be interpreted as a spatial function of the demand (see Fig. 7.3(a)). Then, for our given area, we



(a) Demand field (b) Demand process $\widehat{\mathcal{S}}$

Figure 7.3: Constructing a demand model with varying spatial clustering of demand: a) from our real data we create a continuous demand field spanning the analyzed area (red colour represents a higher volume of demand); b) we sample the demand field based on realizations of the Gauss-Poisson process for a fixed cluster probability value to obtain $\widehat{\mathcal{S}}$ (radii of the discs are directly proportional to the volume of demand originating in a given point).

generate a number of realizations of the Gauss-Poisson process (GPP), which we have already defined in Chapter 3. Shortly, the GPP is a point process on $\mathbb{R}^2$, with uniformly and independently distributed clusters, where each cluster consists of either one point or two points depending on the cluster probability. By generating realizations of the GPP with varying cluster probability, we obtain patterns of arbitrary level of clustering between points. Finally, we use such generated point patterns to sample the demand field created in the previous step to obtain $\widehat{\mathcal{S}}$ (see Fig. 7.3(b)).

C.2 Modelling base station placement

The real data we have also contains positions and coverage information for real Long Term Evolution (LTE) base stations. Additional base stations (small cells) required to cover the demand are taken from a regular grid laid over the analyzed area, where vertices of the grid form the set of candidate locations; for simplicity, we assume that any kind of radio access technology can be placed at any such location (and, in fact, multiple technologies may co-locate at any single location). We formulate the problems of determining the base station locations to serve the mobile demand as two mixed integer optimization problems.

OTT

We formulate the problem that the OTT is facing as:

$$\min_{\{x_l, y_{ls}\}} \sum_{l \in \mathcal{L}} c_l x_l, \tag{7.31}$$

subject to:

$$\sum_{s\in\mathcal{S}} \max\left(R_s - \sum_{l\in\mathcal{L}} t_{ls}y_{ls}, 0\right) \geq \theta R, \tag{7.32}$$

$$\sum_{s\in\mathcal{S}} y_{ls} \leq x_l p_l, \quad \forall l \in \mathcal{L}, \tag{7.33}$$

$$x_l \in \{0,1\}, \quad \forall l \in \mathcal{L}, \tag{7.34}$$

$$y_{ls} \in \mathbb{Z}_+, \quad \forall l \in \mathcal{L}. \tag{7.35}$$

where $\mathcal{L}$ denotes the set of base station locations, c_l is the price of a base station in location l, t_{ls} states whether a base station in location l is able to provide coverage to demand point in s, p_l is the capacity of a base station located in l. To simplify our notation we omitted the impact a technology has on the coverage and capacity. Each operator minimizes the number of deployed base stations, so as to serve a pre-specified fraction of the demand Eq. (7.32), without violating the capacity constraint Eq. (7.33), and by making binary decisions on whether to deploy a base station in location Eq. (7.34).

MNO

Now, the MNO has to serve the OTT's residual demand using the existing infrastructure, which bears no additional deployment costs, or, if that is not possible (due to lack of coverage or capacity), using newly deployed small cell infrastructure at a unit cost c_m. It is safe to assume that its goal is to maximize the revenue; in the deployment phase this is equivalent to minimizing the deployment cost, given some level of demand that needs to be served:

$$\min_{\{x_l, y_{ls}\}} \sum_{l\in\mathcal{L}} x_l, \tag{7.36}$$

subject to:

$$\sum_{l\in\mathcal{L}} y_{ls}t_{ls} \geq r_s, \forall (s, r_s) \in \widehat{\mathcal{S}}_{\texttt{res}}, \tag{7.37}$$

$$\sum_{s\in\mathcal{S}} y_{ls} \leq x_l p_l, \forall l \in \mathcal{L}, \tag{7.38}$$

$$x_l \in \{0,1\}, \quad \forall l \in \mathcal{L}, \tag{7.39}$$

$$y_{ls} \in \mathbb{Z}_+, \quad \forall l \in \mathcal{L}. \tag{7.40}$$

The MNO minimizes the number of deployed base stations to cover the residual demand $\widehat{\mathcal{S}}_{\texttt{res}}$ (demand constraint Eq. (7.37)). Again, the decisions being made are subject to the capacity constraint (Eq. (7.38)) and are binary (Eq. (7.39)), and concern whether or not to deploy a base station in location $l \in \mathcal{L}$. Note that we can model any demand the MNO has from its subscribers by adjusting available capacity p_l accordingly.

Solver

The deployment problems of the OTT and the MNO, formulated as Eq. (7.31) and Eq. (7.36), have the same structure. They are linear problems with binary and integer variables; for reasonably sized scenarios such as ours they can be solved to optimality with branch-and-cut algorithms implemented in commercial solvers. Here we utilize a commercial solver software Gurobi [148]. Let us note also that exact solvers to combinatorial optimization problems yield notorious scalability issues. Hence, if larger scenarios are considered, heuristic solutions need to be applied, such as the greedy algorithm described in [149].

Regression analysis

After we apply optimal solvers to the problems in Eq. (7.31) and Eq. (7.36), we obtain a set of curves that represents the deployment function with respect to the spatial clustering of the demand data. Having an empirical function representing the deployment may just not be enough to arrive at more general conclusions about the relationship between the MNO and the OTT. Therefore we exploit the fact that our deployment function curves have an apparent parabola-like shape, and we perform a linear (least-squares) regression on the empirical data samples. Using the least-squares estimates method implemented in R software [108], we fit our data to three different models:

- a linear function of the form $a_0 + a_1\theta$;
- a piecewise function which is a hockey-stick type of a curve that is composed of two rays meeting in a unique knot point (we perform the estimation using the SiZer package [150]);
- a quadratic function of the form $a_0 + a_1\theta + a_2\theta^2$, which we fit using the Levenberg-Marquardt algorithm [151], implemented in the minpack.lm package [152],

where $a_0, a_1, a_2 \in \mathbb{R}$.

In Fig. 7.4(a) and Fig. 7.4(b) we present examples of the fitting results. We can qualitatively assess that quadratic function provides the most satisfying fits for the deployment function of the MNO. Both functions turn out to be good estimators of the deployment sizes produced by our model. The benefit of having a quadratic function representation is that it allows us to apply the Lagrangian multipliers method, which requires convexity of the objective function being optimized (for this to hold, we also need to make the assumption that $a_2 \geq 0$, which is in line with our regression results).

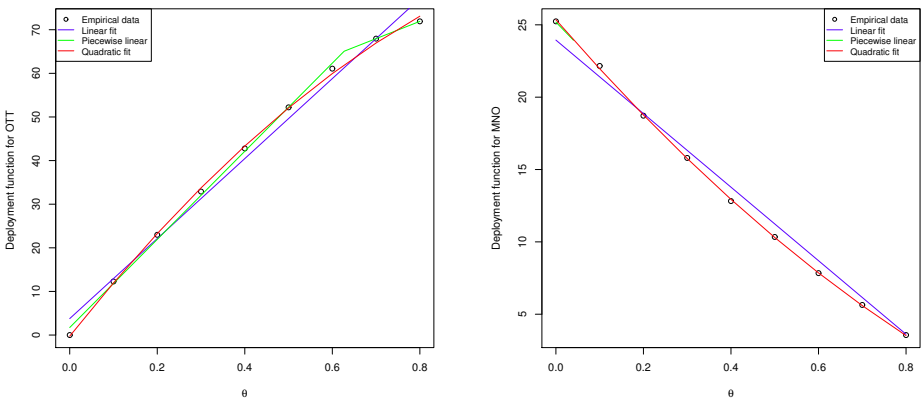
We can now substitute the fitted deployment function to Eq. (6.17), in the OTT case:

$$u_{\mathrm{OTT}} = -\left(a_0 + a_1\theta + a_2\theta^2\right) - \widehat{\beta}(1-\theta)R + Q, \tag{7.41}$$

where $a_{\{0,1,2\}}$ are the coefficients of the deployment function, and to Eq. (6.18) in the MNO case:

$$u_{\mathrm{MNO}} = -p_{\mathrm{MNO}}\left(b_0 + b_1\theta + b_2\theta^2\right) + \widehat{\beta}(1-\theta)R, \tag{7.42}$$

where $b_{\{0,1,2\}}$ are the coefficients of the deployment function.



(a) OTT deployment function (b) MNO deployment function

Figure 7.4: An example of the empirical deployment function (circles) and the fitted functions: linear, piecewise linear, and quadratic.

Acronyms

2G	2^{nd} generation
3G	3^{rd} generation
3GPP	3^{rd} Generation Partnership Project
4G	4^{th} generation
AWS	Advanced Wireless Services
CFI	European Court of First Instance
CRTC	Canadian Radio-television and Telecommunications Commission
DPP	determinantal point process
EIRP	equivalent isotropic radiated power
FCC	Federal Communications Commission
GPP	Gauss-Poisson process
GSM	Groupe Special Mondiale
GSM1800	Groupe Special Mondiale (GSM) in 1800 MHz
GSM900	GSM in 900 MHz
GSMA	GSM Alliance
GWCN	Gateway Core Network
iid	identically and independently distributed
ILP	integer linear programming
IMS	IP Multimedia Subsystem
ITU	International Telecommunications Union
KSS	Kalai-Smorodinsky bargaining solution
LGCP	Log-Gaussian Cox process
LTE	Long Term Evolution

M2M machine-to-machine

MC Monte Carlo

MCM minimum contrast method

MCP Matérn cluster process

MIP mixed integer program

MME mobility management entity

MMS multimedia messaging service

mmWave millimetre-wave

MNO mobile network operator

MOCN Multi-Operator Core Network

MORAN Multi-Operator RAN

MVNE Mobile Virtual Network Enabler

MVNO Mobile Virtual Network Operator

MWC Mobile World Congress

NBS Nash bargaining solution

NE Nash equilibrium

NMT Nordic Mobile Telephony

NwoB Networks without Borders

Ofcom Office for Communications

OTT over-the-top service provider

pdf probability distribution function

pgfl probability generating functional

PHP Poisson hole process

PPP Poisson point process

PSTN public switched telephony network

QoS Quality of Service

RAN radio access network

RNC radio network controller

RRH remote radio head

SCP set cover problem

SIM Subscriber Identity Module

SINR signal-to-noise and interference ratio

SIR signal-to-interference ratio

SLA service level agreement

SMS short-messaging service

SP Strauss process

TP Thomas process

UMTS Universal Mobile Telecommunications System

UMTS2100 Universal Mobile Telecommunications System (UMTS) in 2100 MHz

Bibliography

[1] J. West, *Institutional constraints in the initial deployment of cellular telephone service on three continents.* Idea Group Publishing, Hershey, USA/London, UK, 1999, ch. Information Technology Standards and Standardization: A Global Perspective, pp. 198–221.

[2] J. L. King and J. West, "Ma Bell's orphan: US cellular telephony, 1947–1996," *Telecommunications Policy*, vol. 26, no. 3–4, pp. 189–203, April–May 2002.

[3] T. Forde and L. Doyle, "Cellular clouds," *Telecommunications Policy*, vol. 37, no. 2–3, pp. 194–207, March–April 2013.

[4] GSM Group, "Memorandum of Understanding on the Implementation of a Pan European 900 MHz Digital Cellular Mobile Telecommunications by 1991," September 1987.

[5] M. Balon and B. Liau, "Mobile virtual network operator," in *XVth International Telecommunications Network Strategy and Planning Symposium (NETWORKS)*, October 2012, pp. 1–6.

[6] J. Krämer, L. Wiewiorra, and C. Weinhardt, "Net neutrality: A progress report," *Telecommunications Policy*, vol. 37, no. 9, pp. 794–813, October 2013.

[7] Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2011-2016. [Online]. Available: http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-520862.pdf

[8] N. Bhushan, J. Li, D. Malladi, R. Gilmore, D. Brenner, A. Damnjanovic, R. Sukhavasi, C. Patel, and S. Geirhofer, "Network densification: the dominant theme for wireless evolution into 5G," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 82–89, February 2014.

[9] B. Molleryd, J. Markendahl, J. Werding, and O. Makitalo, "Decoupling of revenues and traffic – Is there a revenue gap for mobile broadband?" in *9th Conference on Telecommunications Internet and Media Techno Economics (CTTE)*, June 2010, pp. 1–7.

[10] L. A. DaSilva, J. Kibiłda, P. Di Francesco, T. K. Forde, and L. E. Doyle, "Customized Services over Virtual Wireless Networks: The Path towards Networks without Borders," in *Proceedings of Future Network and Mobile Summit (FNMS)*, July 2013, pp. 1–10.

[11] L. Doyle, J. Kibiłda, T. Forde, and L. A. DaSilva, "Spectrum without Bounds, Networks without Borders," *Proceedings of the IEEE*, vol. 102, no. 3, pp. 351–365, March 2014.

[12] N. Economides, "The Telecommunications Act of 1996 and its impact," *Japan and the World Economy*, vol. 11, no. 4, pp. 455–483, December 1999.

[13] T. Frisanco, P. Tafertshofer, P. Lurin, and R. Ang, "Infrastructure Sharing and Shared Operations for Mobile Network Operators: From a Deployment and Operations View," in *Proceedings of IEEE International Conference on Communications (ICC)*, May 2008, pp. 2193–2200.

[14] A. Bartlett and N. Jackson, "Network planning considerations for network sharing in UMTS," in *Third International Conference on 3G Mobile Communication Technologies*, May 2002, pp. 17–21.

[15] A. Khan, W. Kellerer, K. Kozu, and M. Yabusaki, "Network sharing in the next mobile network: TCO reduction, management flexibility, and operational independence," *IEEE Communications Magazine*, vol. 49, no. 10, pp. 134–142, October 2011.

[16] "TR 22.951 V6.1.0; Universal Mobile Telecommunications System (UMTS); Service aspects and requirements for network sharing," 3GPP, Tech. Rep., 2003.

[17] "TR 22.951 V12.0.0; Universal Mobile Telecommunications System (UMTS); Service aspects and requirements for network sharing," 3GPP, Tech. Rep., October 2014.

[18] "TS 23.251 V12.1.0.; Universal Mobile Telecommunications System (UMTS); LTE; Network sharing; Architecture and functional description," 3GPP, Tech. Rep., October 2014.

[19] Directive 2002/21/EC of the European Parliament and of the Council on a common regulatory framework for electronic communications networks and services, 2002 OJ L 108/33.

[20] C. B. Lefevre, "Mobile sharing," in *Proceedings of 8th International ITU Global Symposium of Regulators*, 2008.

[21] 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, May 2011.

[22] KPMG. (2011) Passive Infrastructure Sharing in Telecommunications. online. [Online]. Available: http://www.kpmg.com/BE/en/IssuesAndInsights/ArticlesPublications/Documents/Passive-Infrastructure-Sharing-in-Telecommunications.pdf

[23] Mobile Operators Association (MOA). Network Sharing and Consolidation. online. [Online]. Available: http://www.mobilemastinfo.com/network-sharing-and-consolidation/

[24] K. Kyllesbech Larsen. (2012, July) Network Sharing Fundamentals. online. [Online]. Available: http://www.slideshare.net/KimKyllesbechLarsen/fundamentals-of-mobile-network-sharing

[25] B. Ray. (2012, June) Vodafone and O2 to merge mobile networks. The Register. [Online]. Available: http://www.theregister.co.uk/2012/06/07/vodafonica/

[26] Belgian Institute For Postal Services And Telecommunication (BIPT). (2012, January) Communication of the BIPT of 17 January 2012 containing guidelines for infrastructure sharing. online. [Online]. Available: http://www.bipt.be/public/files/en/680/3666_en_02_tech_infra_sharing_eng_final.pdf

[27] D. Leza, "Mobile Infrastructure Sharing: Trends in Latin America," ITU Regional Economic and Financial Forum of Telecommunications/ICTs for Latin America and the Caribbean, March 2014.

[28] G. Punz, *Evolution of 3G Networks: The Concept, Architecture and Realization of Mobile Networks Beyond UMTS*. Springer Science & Business Media, 2009.

[29] M. Ney. (2013, June) Deployment of RAN Sharing in Poland. online. [Online]. Available: http://www.slideshare.net/zahidtg/deployment-of-ran-sharing-in-poland

[30] M. Carroll. (2014, August) Eircom agrees long-term network sharing deal with Three Ireland. online. FierceWirelessEurope. [Online]. Available: http://www.fiercewireless.com/europe/story/eircom-agrees-long-term-network-sharing-deal-3-ireland/2014-08-29

[31] J. Markendahl and B. G. Mölleryd, "On co-opetition between mobile network operators: Why and how competitors cooperate," ser. 19th ITS Biennial Conference 2012, Bangkok, Thailand, 18 - 21 November 2012: Moving Forward with Future Technologies: Opening a Platform for All. Bangkok: ITS, 2012. [Online]. Available: http://hdl.handle.net/10419/72491

[32] Commission Decision No. 2004/207/CE (Case COMP/38.369: T-Mobile Deutschland/O2 Germany: Network Sharing Rahmenvertrag), 2003 O.J. L 75/32.

[33] T. Espiner. (2010, September) Orange, T-Mobile to let customers roam across networks. onliny. ZDNet. [Online]. Available: http://www.zdnet.com/article/orange-t-mobile-to-let-customers-roam-across-networks/

[34] T. K. Forde, I. Macaluso, and L. E. Doyle, "Exclusive sharing & virtualization of the cellular network," in *Proceedings of IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN)*, May 2011, pp. 337–348.

[35] J. Kibiłda, B. Galkin, and L. A. DaSilva, "Modelling Multi-operator Base Station Deployment Patterns in Cellular Networks," *IEEE Transactions on Mobile Computing*, 2015, (to appear).

[36] N. Corrocher and L. Lasio, "Diversification strategies in network-based services: The case of mobile virtual network operators," *Telecommunications Policy*, vol. 37, no. 11, pp. 1110–1123, December 2013.

[37] L. Cricelli, M. Grimaldi, and N. L. Ghiron, "The competition among mobile network operators in the telecommunication supply chain," *International Journal of Production Economics*, vol. 131, no. 1, pp. 22–29, May 2011, 2008.

[38] P. Xavier. (2001, September) Licensing of third generation (3G) mobile: Briefing paper. [Online]. Available: http://www.itu.int/osg/spu/ni/3G/workshop/Briefing_paper.PDF

[39] P. Suciu. (2015, May) Google's MVNO Approach May Show Up Carriers. online. TechNewsWorld. [Online]. Available: http://www.technewsworld.com/story/ect-mobile/81766.html

[40] "The Communications Market Report UK," Ofcom, Tech. Rep., August 2015. [Online]. Available: http://stakeholders.ofcom.org.uk/binaries/research/cmr/cmr15/CMR_UK_2015.pdf

[41] "Protecting and Promoting the Open Internet, GN Docket No. 14-28; FCC 14-61," Federal Communications Commission (FCC), 15 May 2014.

[42] J. Kibiłda, F. Malandrino, and L. A. DaSilva, "Incentives for Infrastructure Deployment by Over-the-Top Service Providers in a Mobile Network: A Cooperative Game Theory Model," *IEEE International Conference on Communications (ICC)*, May 2016.

[43] C. Mims. (2012, September) Facebook's plan to find its next billion users: convince them the internet and Facebook are the same. [Online]. Available: http://qz.com/5180/facebooks-plan-to-find-its-next-billion-users-convince-them-the-internet-and-facebook-are-the-same/

[44] [Online]. Available: http://www.rcrwireless.com/20080318/policy/its-over-700-mhz-auction-ends-after-38-days-261-rounds

[45] [Online]. Available: http://apps.fcc.gov/ecfs/document/view;ECFSSESSION=Jj4ZJ9dGFkqMLDTsNgBvJ32pGwhJS2DgG5tvvdbC1JcF4tnhjPDL!1175060748!1957906226?id=60001013832

[46] [Online]. Available: https://www.google.com/loon/

[47] [Online]. Available: https://fi.google.com/about/

[48] S. Datoo, "Twitter's latest deal points to ambitions in emerging markets," *The Guardian*, 2013.

[49] E. Wyatt and N. Cohen. (2014, February) Comcast and Netflix Reach Deal on Service. The New York Times. [Online]. Available: http://www.nytimes.com/2014/02/24/business/media/comcast-and-netflix-reach-a-streaming-agreement.html?_r=1

[50] M. Haenggi, *Stochastic Geometry for Wireless Networks*. Cambridge University Press, 2013.

[51] J. Moeller and R. P. Waagepetersen, *Statistical Inference and Simulation for Spatial Point Processes.* Chapman & Hall/CRC, 2004.

[52] J. Illian, A. Penttinen, H. Stoyan, and D. Stoyan, *Statistical Analysis and Modelling of Spatial Point Patterns (Statistics in Practice)*, 1st ed. Wiley-Interscience, 2008.

[53] R. G. Gallager, *Stochastic Processes: Theory for Applications.* Cambridge University Press, 2013.

[54] J. Moeller, A. R. Syversveen, and R. P. Waagepetrsen, "Log Gaussian Cox Processes," *Scandinavian Journal of Statistics*, vol. 25, no. 3, pp. 451–482, September 1998.

[55] M. Van Lieshout and A. J. Baddeley, "Indices of Dependence Between Types in Multivariate Point Patterns," *Scandinavian Journal of Statistics*, vol. 26, no. 4, pp. 511–532, December 1999.

[56] A. Baddeley and R. Turner, "Practical maximum pseudolikelihood for spatial point patterns," *Australian and New Zealand J. Statistics*, vol. 42, no. 3, pp. 283–322, September 2000.

[57] B. Galkin, J. Kibiłda, and L. A. DaSilva, "Stochastic Modelling of Downlink Transmit Power in Wireless Cellular Networks," in *20th IEEE International Workshop on Computer Aided Modelling and Design of Communication Links and Networks (CAMAD)*, September 2015, pp. 33–37.

[58] M. De Berg, M. Van Kreveld, M. Overmars, and O. C. Schwarzkopf, *Computational geometry.* Springer, 2000.

[59] H. Frank and S. C. Althoen, *Statistics Concepts and Applications.* Cambridge University Press, 1994.

[60] E. H. Isaaks and R. M. Srivastava, *An Introduction to Applied Geostatistics.* Oxford University Press, 1989.

[61] Y. Ogata and K. Katsura, "Likelihood analysis of spatial inhomogeneity for marked point patterns," *Annals of the Institute of Statistical Mathematics*, vol. 40, no. 1, pp. 29–39, March 1988.

[62] J. G. Andrews, F. Baccelli, and R. K. Ganti, "A tractable approach to coverage and rate in cellular networks," *IEEE Transactions on Communications*, vol. 59, no. 11, pp. 3122–3134, November 2011.

[63] M. Di Renzo, A. Guidotti, and G. Corazza, "Average Rate of Downlink Heterogeneous Cellular Networks over Generalized Fading Channels: A Stochastic Geometry Approach," *IEEE Transactions on Communications*, vol. 61, no. 7, pp. 3050–3071, July 2013.

[64] R. K. Ganti and M. Haenggi, "Interference and outage in clustered wireless ad hoc networks," *IEEE Transactions on Information Theory*, vol. 55, no. 9, pp. 4067–4086, September 2009.

[65] G. Alfano, M. Garetto, and E. Leonardi, "Capacity scaling of wireless networks with inhomogeneous node density: upper bounds," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 7, pp. 1147–1157, September 2009.

[66] M. Garetto, A. Nordio, C. Chiasserini, and E. Leonardi, "Information-theoretic capacity of clustered random networks," *IEEE Transactions on Information Theory*, vol. 57, no. 11, pp. 7578–7596, November 2011.

[67] J. Wang, X. Zhou, and M. Reed, "Analytical Evaluation of Coverage-Oriented Femtocell Network Deployment," in *Proceedings of IEEE International Conference on Communications (ICC)*, June 2013, pp. 4567–4572.

[68] A. Guo and M. Haenggi, "Spatial Stochastic Models and Metrics for the Structure of Base Stations in Cellular Networks," *IEEE Transactions on Wireless Communications*, vol. 12, no. 11, pp. 5800–5812, November 2013.

[69] N. Deng, W. Zhou, and M. Haenggi, "The Ginibre Point Process as a Model for Wireless Networks With Repulsion," *IEEE Transactions on Wireless Communications*, vol. 14, no. 1, pp. 107–121, January 2015.

[70] A. Guo, Y. Zhong, M. Haenggi, and W. Zhang, "Success probabilities in gauss-poisson networks with and without cooperation," in *Proceedings of IEEE International Symposium on Information Theory (ISIT)*, June 2014, pp. 1752–1756.

[71] S. Singh, H. Dhillon, and J. Andrews, "Offloading in heterogeneous networks: Modeling, analysis, and design insights," *Wireless Communications, IEEE Transactions on*, vol. 12, no. 5, pp. 2484–2497, May 2013.

[72] H. Elsawy, E. Hossain, and M. Haenggi, "Stochastic Geometry for Modeling, Analysis, and Design of Multi-Tier and Cognitive Cellular Wireless Networks: A Survey," *IEEE Communications Surveys and Tutorials*, vol. 15, no. 3, pp. 996–1019, 2013.

[73] M. Afshang, H. S. Dhillon, and P. H. Joo Chong, "Coverage and Area Spectral Efficiency of Clustered Device-to-Device Networks," *IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6, December 2015.

[74] M. Kulkarni, S. Singh, and J. Andrews, "Coverage and rate trends in dense urban mmWave cellular networks," in *IEEE Global Communications Conference (GLOBECOM)*, December 2014, pp. 3809–3814.

[75] M. Michalopoulou, J. Riihijarvi, and P. Mahonen, "Studying the Relationships between Spatial Structures of Wireless Networks and Population Densities," in *IEEE Global Telecommunications Conference (GLOBECOM)*, December 2010, pp. 1–6.

[76] R. Z. Farahani, N. Asgari, N. Heidari, M. Hosseininia, and M. Goh, "Covering problems in facility location: A review," *Computers & Industrial Engineering*, vol. 62, no. 1, pp. 368–407, February 2012.

[77] A. Caprara and P. Toth, "Algorithms for the Set Covering Problem," *Annals of Operations Research*, vol. 98, no. 1–4, pp. 353–371, December 2000.

[78] C. H. Papadimitriou and K. Steiglitz, *Combinatorial Optimization: Algorithms and Complexity*. Courier Dover Publications, 1998.

[79] E. Amaldi, A. Capone, F. Malucelli, and C. Mannino, "Optimization problems and models for planning cellular networks," *Handbook of Optimization in Telecommunication*, pp. 879–901, 2006.

[80] A. Ligeti and J. Zander, "Minimal Cost Coverage Planning for Single Frequency Networks," *IEEE Transactions on Broadcasting*, vol. 45, no. 1, pp. 78–87, March 1999.

[81] R. Mathar and T. Niessen, "Optimum positioning of base stations for cellular radio networks," *Wireless Networks*, vol. 6, no. 6, pp. 421–428, December 2000.

[82] F. Gordejuela-Sanchez and J. Zhang, "Lte access network planning and optimization: A service-oriented and technology-specific perspective," in *Proceedings of IEEE Global Telecommunications Conference (GLOBECOM)*, November 2009, pp. 1–5.

[83] L. Chen and D. Yuan, "Solving a minimum-power covering problem with overlap constraint for cellular network design," *European Journal of Operational Research*, vol. 203, no. 3, pp. 714–723, June 2010.

[84] E. Pollakis, R. Cavalcante, and S. Stańczak, "Base station selection for energy efficient network operation with the majorization-minimization algorithm," in *Proceedings of IEEE 13th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, June 2012, pp. 219–223.

[85] M. Zheng, P. Pawełczak, S. Stańczak, and H. Yu, "Planning of cellular networks enhanced by energy harvesting," *IEEE Communications Letters*, vol. 17, no. 6, pp. 1092–1095, June 2013.

[86] A. Muthoo, *Bargaining Theory with Applications.* Cambridge University Press, 1999.

[87] G. Owen, *Game Theory.* Academic Press, 1995.

[88] E. Kalai and M. Smorodinsky, "Other Solutions to Nash's Bargaining Problem," *Econometrica*, vol. 43, no. 3, pp. 513–518, 1975.

[89] X. Cao, "Preference functions and bargaining solutions," in *21st IEEE Conference on Decision and Control*, December 1982, pp. 164–171.

[90] S.-L. Hew and L. B. White, "Cooperative Resource Allocation Games in Shared Networks: Symmetric and Asymmetric Fair Bargaining Models," *IEEE Transactions on Wireless Communications*, vol. 7, no. 11, pp. 4166–4175, November 2008.

[91] E. Larsson and E. Jorswieck, "Competition Versus Cooperation on the MISO Interference Channel," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 7, pp. 1059–1069, September 2008.

[92] J. E. Suris, L. A. DaSilva, Z. Han, A. B. MacKenzie, and R. S. Komali, "Asymptotic Optimality for Distributed Spectrum Sharing Using Bargaining Solutions," *IEEE Transactions on Wireless Communications*, vol. 8, no. 10, pp. 5225–5237, October 2009.

[93] H. Boche, M. Schubert, N. Vucic, and S. Naik, "Non-symmetric nash bargaining solution for resource allocation in wireless networks and connection to interference calculus," in *Proc. 15th European Signal Processing Conference*, September 2007, pp. 1317–1321.

[94] P. Lin, J. Jia, Q. Zhang, and M. Hamdi, "Cooperation among wireless service providers: opportunity, challenge, and solution," *IEEE Wireless Communications*, vol. 17, no. 4, pp. 55–61, August 2010.

[95] J. Lempiainen and M. Manninen, *UMTS Radio Network Planning, Optimization and Qos Management: For Practical Engineering Tasks.* Kluwer Academic Publishers, 2004.

[96] L. Z. Ribeiro and L. A. DaSilva, "A Framework for the Dimensioning of Broadband Mobile Networks Supporting Wireless Internet Services," *IEEE Wireless Communications*, vol. 9, no. 3, pp. 6–13, June 2002.

[97] M. Nawrocki, H. Aghvami, and M. Dohler, *Understanding UMTS Radio Network Modelling, Planning and Automated Optimisation: Theory and Practice.* John Wiley & Sons, 2006.

[98] H. Dhillon, R. K. Ganti, F. Baccelli, and J. G. Andrews, "Modeling and analysis of k-tier downlink heterogeneous cellular networks," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 3, pp. 550–560, April 2012.

[99] J. Kibiłda, P. Di Francesco, F. Malandrino, and L. A. DaSilva, "Infrastructure and Spectrum Sharing Trade-offs in Mobile Networks," in *IEEE Dynamic Spectrum Access Networks (DySPAN)*, September 2015, pp. 348–357.

[100] B. Błaszczyszyn, M. K. Karray, and H. P. Keeler, "Using Poisson processes to model lattice cellular networks," in *Proc. of IEEE INFOCOM*, April 2013, pp. 773–781.

[101] Y. Li, F. Baccelli, H. Dhillon, and J. Andrews, "Fitting determinantal point processes to macro base station deployments," in *IEEE Global Communications Conference (GLOBECOM)*, December 2014, pp. 3641–3646.

[102] M. Haenggi, "The Mean Interference-to-Signal Ratio and Its Key Role in Cellular and Amorphous Networks," *IEEE Wireless Communications Letters*, vol. 3, no. 6, pp. 597–600, December 2014.

[103] "Further Advancements for E-UTRAN Physical Layer Aspects," 3GPP, Tech. Rep. TR 36.814 V9.0.0, March 2010.

[104] M. Bennis and J. Lilleberg, "Inter base station resource sharing and improving the overall efficiency of B3G systems," in *Vehicular Technology Conference (VTC)*, September 2007, pp. 1494–1498.

[105] L. Anchora, L. Badia, H. Zhang, T. Fahldieck, J. Zhang, M. Szydelko, M. Schubert, E. Karipidis, and M. Haardt, "Resource allocation and management in multi-operator cellular networks with shared physical resources," in *IEEE International Symposium on Wireless Communication Systems (ISWCS)*, August 2012, pp. 296–300.

[106] J. S. Panchal, R. D. Yates, and M. M. Buddhikot, "Mobile Network Resource Sharing Options : Performance Comparisons," *IEEE Transactions on Wireless Communications*, vol. 12, no. 9, pp. 4470–4482, September 2013.

[107] E. A. Jorswieck, L. Badia, T. Fahldieck, E. Karipidis, and J. Luo, "Spectrum Sharing Improves the Network Efficiency for Cellular Operators," *IEEE Communications Magazine*, vol. 52, no. 3, pp. 129–136, March 2014.

[108] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2015. [Online]. Available: https://www.R-project.org

[109] A. Baddeley and R. Turner, "Spatstat: an R package for analyzing spatial point patterns," *J. Statist. Software*, vol. 12, no. 6, pp. 1–42, February 2005.

[110] W. N. Venables and B. D. Ripley, *Modern Applied Statistics with S*, 4th ed. New York: Springer, 2002.

[111] (2015, June) Ericsson mobility report. on the pulse of the networked society. Ericsson AB. [Online]. Available: http://www.ericsson.com/res/docs/2015/ericsson-mobility-report-june-2015.pdf

[112] P. Di Francesco, F. Malandrino, and L. A. DaSilva, "Mobile Network Sharing Between Operators: A Demand Trace-Driven Study," in *Proceedings of the 2014 ACM SIGCOMM Workshop on Capacity Sharing*, August 2014, pp. 39–44.

[113] M. Marotta, N. Kaminski, I. Gomez-Miguelez, L. Zambenedetti Granville, J. Rochol, L. DaSilva, and C. Both, "Resource sharing in heterogeneous cloud radio access networks," *IEEE Wireless Communications*, vol. 22, no. 3, pp. 74–82, June 2015.

[114] I. Macaluso, H. Ahmadi, I. Gomez-Miguelez, L. E. Doyle, and L. A. DaSilva, "Substitutability of Spectrum and Cloud-based Antennas in Virtualised Wireless Networks," *IEEE Wireless Communications Magazine*, 2016, (to appear).

[115] J. Kibiłda and L. A. DaSilva, "Efficient Coverage Through Inter-operator Infrastructure Sharing in Mobile Networks," in *Proceedings of IFIP Wireless Days*, November 2013, pp. 1–6.

[116] L. Cano, A. Capone, G. Carello, and M. Cesana, "Evaluating the performance of infrastructure sharing in mobile radio networks," in *IEEE International Conference on Communications (ICC)*, June 2015, pp. 3222–3227.

[117] Y. Yang and K. W. Sung, "Technical rate of substitution of spectrum in future mobile broadband provisioning," in *IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN)*, September 2015, pp. 297–300.

[118] X. Lin, J. Andrews, and A. Ghosh, "Spectrum Sharing for Device-to-Device Communication in Cellular Networks," *IEEE Transactions on Wireless Communications*, vol. 13, no. 12, pp. 6727–6740, December 2014.

[119] A. K. Gupta, J. G. Andrews, and R. W. Heath, "On the Feasibility of Sharing Spectrum Licenses in mmWave Cellular Systems," *Arxiv*, 2016. [Online]. Available: http://arxiv.org/abs/1512.01290

[120] A. A. W. Ahmed, Y. Yang, K. W. Sung, and J. Markendahl, "On the engineering value of spectrum in dense mobile network deployment scenarios," in *IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN)*, September 2015, pp. 293–296.

[121] F. Baccelli and B. Błaszczyszyn, *Stochastic geometry and wireless networks: Volume 1: Theory*. Now Publishers Inc, 2009, vol. 1.

[122] Z. Yazdanshenasan, H. S. Dhillon, M. Afshang, and P. H. J. Chong, "Poisson Hole Process: Theory and Applications to Wireless Networks," *arXiv*, 2016. [Online]. Available: http://arxiv.org/abs/1601.01090

[123] Credit Suisse. (2011) U.S. wireless networks running at 80% of capacity. [Online]. Available: http://benton.org/node/81874

[124] M. Dohler, R. Heath, A. Lozano, C. Papadias, and R. Valenzuela, "Is the PHY layer dead?" *IEEE Communications Magazine*, vol. 49, no. 4, pp. 159–165, April 2011.

[125] M. J. Abdel-Rahman, K. V. Cardoso, A. B. MacKenzie, and L. A. DaSilva, "Dimensioning Virtualized Wireless Access Networks from a Common Pool of Resources," *13th IEEE Consumer Communications and Networking Conference*, January 2016.

[126] K. Zhu and E. Hossain, "Virtualization of 5G Cellular Networks as a Hierarchical Combinatorial Auction," *IEEE Transactions on Mobile Computing*, 2015, (to appear).

[127] J. van de Belt, H. Ahmadi, and L. Doyle, "A Dynamic Embedding Algorithm for Wireless Network Virtualization," in *IEEE 80th Vehicular Technology Conference (VTC Fall)*, September 2014, pp. 1–6.

[128] J. van de Belt, H. Ahmadi, L. Doyle, and O. Sallent, "A Prioritised Traffic Embedding Mechanism Enabling a Public Safety Virtual Operator," in *IEEE 82nd Vehicular Technology Conference (VTC Fall)*, September 2015, pp. 1–5.

[129] B. Liu and H. Tian, "A Bankruptcy Game-Based Resource Allocation Approach among Virtual Mobile Operators," *IEEE Communications Letters*, vol. 17, no. 7, pp. 1420–1423, July 2013.

[130] G. Tseliou, F. Adelantado, and C. Verikoukis, "Scalable RAN Virtualization in Multi-Tenant LTE-A Heterogeneous Networks," *IEEE Transactions on Vehicular Technology*, 2015, (to appear).

[131] F. Fund, S. Hosseini, and S. Panwar, "More bars, more bang for the buck: Channel-dependent pricing for video delivery to mobile users," in *IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, April 2014, pp. 565–570.

[132] E. Altman, M. K. Hanuwa, and R. Sundaresan, "Nonneutral network and the role of bargaining power in side payments," in *Workshop on Network Control and Optimization (NET-COOP)*, November 2010.

[133] R. Ma and V. Misra, "The public option: a nonregulatory alternative to network neutrality," *IEEE/ACM Transactions on Networking*, vol. 21, no. 6, pp. 1866–1879, December 2013.

[134] G. Kesidis, "A simple two-sided market model with side-payments and ISP service classes," in *IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, April 2014, pp. 595–597.

[135] F. Boussion, P. Maillé, and B. Tuffin, "Net neutrality debate: Impact of competition among ISPs," in *International Conference on Communication Systems and Networks (COMSNETS)*, January 2012, pp. 1–8.

[136] J. Markendahl, "Mobile network operators and cooperation," Ph.D. dissertation, Royal Institute of Technology (KTH), 2011.

[137] S. Sesia, I. Toufik, and M. Baker, *LTE - The UMTS Long Term Evolution: From Theory to Practice*, 2nd ed. John Wiley & Sons, Sep. 2011.

[138] P. Beraldi and A. Ruszczyński, "The probabilistic set-covering problem," *Operations Research*, vol. 50, no. 6, pp. 956–967, 2002.

[139] K. Błaszczyk. (2013, April) Urząd Komunikacji Elektronicznej (UKE). [Online]. Available: http://www.uke.gov.pl/pozwolenia-radiowe-dla-stacji-gsm-umts-lte-oraz-cdma-4145

[140] "Report ITU-R SM.2028-1;Monte Carlo simulation methodology for the use in sharing and compatibility studies between different radio services or systems," ITU-R, Tech. Rep., 2001-2002.

[141] L. O'Reilly. (2015, May) This ad blocking company has the potential to tear a hole right through the mobile web – and it has the support of carriers. Business Insider UK. [Online]. Available: http://uk.businessinsider.com/israeli-ad-blocker-shine-could-threaten-mobile-advertising-2015-5

[142] L. Winner, "Do artifacts have politics?" *Daedalus*, vol. 109, no. 1, pp. 121–136, 1980.

[143] N. Newman, "Search, Antitrust, and the Economics of the Control of User Data," *Yale Journal of Regulation*, vol. 31, no. 2, p. 401, 2014.

[144] L. Mearian, "Big data to drive a surveillance society," *Computerworld*, March 2011. [Online]. Available: http://www.computerworld.com/article/2507168/business-intelligence/big-data-to-drive-a-surveillance-society.html

[145] C. Doctorow, "Personal data is as hot as nuclear waste," *The Guardian*, January 2008. [Online]. Available: http://www.theguardian.com/technology/2008/jan/15/data.security

[146] M. Haenggi, "On distances in uniformly random networks," *IEEE Transactions on Information Theory*, vol. 51, no. 10, pp. 3584–3586, October 2005.

[147] S. Hua, P. Liu, and S. S. Panwar, "The urge to merge: When cellular service providers pool capacity," in *IEEE International Conference on Communications (ICC)*, June 2012, pp. 5020–5025.

[148] Gurobi Optimization. Gurobi Optimizer Reference Manual. [Online]. Available: http://www.gurobi.com

[149] A. Caprara, P. Toth, and M. Fischetti, "Algorithms for the set covering problem," *Annals of Operations Research*, vol. 98, no. 1, pp. 353–371, December 2000.

[150] D. Sonderegger, *SiZer: Significant Zero Crossings*, 2012, R package version 0.1–4. [Online]. Available: http://CRAN.R-project.org/package=SiZer

[151] J. Moré, "The Levenberg-Marquardt algorithm: Implementation and theory," in *Numerical Analysis*, ser. Lecture Notes in Mathematics, G. Watson, Ed. Springer Berlin Heidelberg, 1978, vol. 630, pp. 105–116.

[152] T. V. Elzhov, K. M. Mullen, A.-N. Spiess, and B. Bolker, *minpack.lm: R interface to the Levenberg-Marquardt nonlinear least-squares algorithm found in MINPACK, plus support for bounds*, 2013, R package version 1.1–8. [Online]. Available: http://CRAN.R-project.org/package=minpack.lm