Doctoral Dissertation
Doctoral Program in Electrical, Electronics and Communications Engineering
(32th cycle)

Adaptive Network Densification with Small Cell Mobile Base Stations Carried by Vehicles

Foroogh Mohammadnia

* * * * * *

Supervisor
Prof. Marco Ajmone Marsan

Doctoral Examination Committee:
Prof. Gianluca Rizzo, Referee, HES-SO
Prof. Renato Lo Cigno, Referee, University of Brescia

Politecnico di Torino
2020

I hereby declare that, the contents and organisation of this dissertation constitute my own original work and does not compromise in any way the rights of third parties, including those relating to the security of personal data.

. . . .

Foroogh Mohammadnia

Turin, 2020

Summary

Network densification is the evolutionary process that contributed the most during the last decades to allow radio access networks (RANs) cope with the exponential growth in traffic and number of users. Network densification consists in a significant increase in the number of base stations (BSs) that offer service over a given area. This implies a progressive reduction of the area of cells, leading to the concept of small cells (SCs), and possibly to the coexistence of overlapping layers of BSs with large coverage (macro BSs) and SC BSs. Expectations for the deployment of SC BSs were extremely high, but in practice deployment has been limited because of cost and practical difficulties in identifying and instrumenting SC BS sites. In addition, the intrinsic variability in time and space of the traffic demand in a RAN creates traffic peaks that move during time, thus making the utilization of SC BSs high only for possibly short periods.

To improve SC BS deployment cost and utilization, in this dissertation we advocate the introduction of mobile SC BSs carried by vehicles, exploiting the correlation between density of vehicles and peaks of data traffic, and we quantify the benefits of the proposed approach.

After discussing the issues related to the evolution of RAN architectures in the last decades (Chapter 1), we discuss in some detail the opportunities and challenges inherent in RAN densification exploiting SC BSs on vehicles (Chapter 2). Next, we exploit available real data to quantify the correlation in time and space between vehicular traffic and data traffic in cellular networks (Chapter 3). Our results indicate that correlation exists, as expected, even if the available data for vehicular traffic mostly refer to commercial vehicles, that correspond to a small fraction of vehicles on the road, and may not be fully representative of the density of vehicles on the streets of a large metropolitan area. For example, we observe high values of vehicular traffic early in the morning, when many van deliveries occur, and less traffic in the typical rush hours of employees and students. Nevertheless, correlations tend to be higher in dense urban areas and during high-demand time periods, where and when RAN densification is most needed.

After this preliminary study, we delve into the analysis of the performance of a RAN exploiting SC BSs carried by vehicles. We first study the achievable throughput and fairness in some areas of the city of Milan, Italy (Chapter 4), comparing

the maximum throughput that can be obtained with traditional fixed macro and SC BSs to the one achieved with the same fixed macro BSs complemented with mobile small cell base stations. We perform our analysis at different times of the day, using real data for both telecom and vehicular traffic. As an example, studying the main railway station area in Milan, Italy, we see that the use of mobile SC BSs achieves throughput gains up to 120% over fixed access infrastructures with only macro BSs, and equivalent throughput to the deployment of fixed SC BSs. In addition to the computation of the maximum achievable RAN throughput, we also look at the throughput that maximizes the end user proportional fairness.

The next step of our analysis looks at the issues posed by the wireless backhaul connection from mobile SC BSs to the fixed part of the RAN, which is possibly one of the most delicate issues in RAN architectures exploiting mobile SC BSs. While in the initial throughput analysis we assumed the availability of an ideal link between mobile SC BSs and the macro BS (i.e., a link with unconstrained bandwidth, and no interference with transmissions to/from end users, which could correspond to a millimeter wave connection between mobile SC BSs and macro BS), in this refined analysis we look at simpler technologies for the implementation of the wireless backhaul link. In particular, we look at two alternatives.

1. Out-band backhaul: The link between macro BSs and mobile SC BSs exploits a dedicated bandwidth (or dedicated time slots), different from the one used to connect end users to either mobile SC BSs or macro BSs, so that backhaul transmissions do not interfere with transmissions to/from end users.

2. In-band backhaul: The link between macro BSs and mobile SC BSs exploits the same bandwidth and time slots as transmissions to/from end users. This means that backhaul transmissions and transmissions to/from end users can interfere with each other.

The resulting throughput values (presented in Chapter 5) are lower than in the ideal case, as expected, but comparable to the ones obtained with fixed SC BSs, thus proving that the option offered by the exploitation of mobile SC BSs can be an interesting approach for RAN densification with high efficiency at reduced cost. Finally, we summarize our findings and discuss a number of possible directions for further steps in this research topic (Chapter 6).

Acknowledgements

The last but by no means, least section to write after these four years, however I like to keep it short.

Firstly, I will give a huge thanks to my parents, my brother and my sister for their support and patience from far through this long journey, allowing me to make all the choices to be here. I would like to generally thank to all of my friends and anyone who has been near me to support and encourage me during my education, even the authors of the books that I read and learnt from, you are all truly appreciated.

And finally, I like to openly thank my supervisor Marco Ajmone Marsan and my dear colleague Christian Vitale, who kept pushing me during my PhD career, with their passion for this work and immense patience. Thank you very much everyone for everything.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 RAN Evolution

The Radio Access Network (RAN) is the pivotal part of a telecommunication network that serves mobile users (also called a mobile network - MONET). RANs exist since the advent of mobile networks and cellular technology, and they implement the connection between end user equipment (UEs) and the core of the telecommunication network. The key RAN elements comprise UEs and base stations (BSs) with their antennas and transmission/reception equipment.

RANs have evolved dramatically over the years, with the introduction of new generations of mobile networks. In the early days, RANs were just high-power BSs covering an entire city area, with no need for handover procedures, because there was no continuity between areas covered by different BSs. With the advent of each new mobile network generation, the RAN architecture and functionality have advanced and evolved substantially. When digital mobile networks and 2G were introduced, the RAN became the key part of a mobile digital network, and this continued with successive generations and will be true with 5G as well. Each new mobile network generation adds innovative functions to the RAN, aiming to meet the new upcoming demands and to support new applications[1].

With the introduction of 2G (normally known as GSM – Global System for Mobile Communications), data transfer in the RAN went from analog to digital, and with the arrival of 3G (also known as UMTS – Universal Mobile Telecommunication System) the first packet-switched data transfers appeared, in order to support higher data rates. While in 2G the access to the available spectrum resources was based on time and frequency division, in 3G, the code division approach offered higher spectral efficiency, leading to improved data rates [**1**]. A general rendering of the evolution of wireless technologies is presented in Figure 1.1.

[1] https://www.sdxcentral.com/5g/definitions/Radio access network

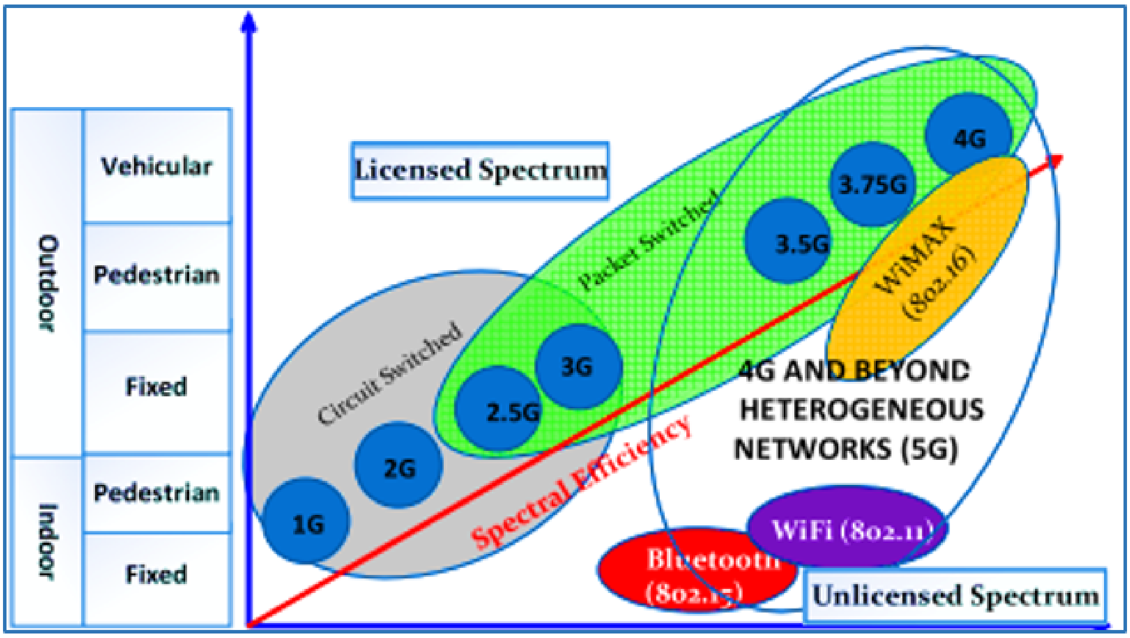


Figure 1.1: Wireless Generations Evolution [1]

Data transfer became fully IP-based in 4G, or LTE (the Long-Term Evolution of UMTS) in which transmissions on the radio interface became packet-based, and mobile networks became heterogeneous, in terms of cell sizes, cell layouts, and radio access technologies (RATs). The advent of 5G brings many new features to the RAN, from a new radio transmission interface to software-defined networking (SDN) and network function virtualization (NFV). The 5G architecture is based on a variety of cells with different coverage areas: macro, micro, pico, femto cells, possibly complemented by relays that improve the coverage quality at cell borders, and possibly exploiting different RATs.

This is coherent with the heterogeneous network concept (HetNet). A HetNet including cells of different reach and various technologies offers several benefits in comparison with a traditional homogeneous wireless network, such as higher reliability, increased spectrum efficiency and improved coverage. Better reliability is achieved because when one of the RATs in the HetNet suffers a fault, connections can still survive over another RAT. Spectrum efficiency and coverage can also improve by using a variety of RATs.

While the ever increasing number of users and the exploding growth of traffic demand are the obvious key drivers for the introduction of 5G, several other motivations exist: the need for higher capacity, the call for new services, the desire for improved data rates, the requirement of reduced latency, the need for lower cost, and the request for higher QoS levels.

1.1.1 Emerging Technologies in 5G

The main emerging technologies that will become part of a 5G RAN, include the followings [**5**].

A new radio interface, comprising novel transmission waveforms and multiple access approaches, multi-antenna and multi-node transmissions, involving different types of transmission/ reception technologies (e.g., multi-hop), as well as new radio resource management procedures.

Massive Machine Communications (MMC) for the implementation of the Internet of things (IoT).

Ultra-dense networks, which result from the network densification process that allowed the boost of the network capacity and energy efficiency and can be instrumental in the reduction of latency. We will return to the network densification aspect in the next chapter.

Massive MIMO technologies, that are instrumental in countering fading, with beneficial effects also on latency [**6**].

Interference management technologies, that allow universal reuse of frequencies, even in complex heterogeneous cell layouts. Two types of interference management will be exploited in 5G [**7**]:

1) Advanced receiver interference management technique. The advanced interference management technique at the receiver can limit the impact of interference by attempting to decode the interfering signal, and successively subtracting it from the received signal, thus producing a cleaner version of the useful signal [**7**].

2) Joint scheduling and interference management technique. The joint scheduling and interference management technique tries to reduce interference by scheduling the transmissions of possibly interfering signal sources in appropriate time slots.

Spectrum Sharing, which allows a higher efficiency in the utilization of the available spectrum. Of the two main possible spectrum sharing approaches, centralized and distributed, the distributed ones act more efficiently [**7**].

Device to Device (D2D) communications for the direct interconnection of devices in proximity of each other, without (or with marginal) involvement of the RAN infrastructure [**8**]. These are particularly useful over congested RAN areas or at the cell edge, where the BS signal is weak. Through D2D, an ad hoc network is created, and devices are able to directly communicate with each other. When this happens, the BS can control the resource allocation fully or partially, or it may not have any control. Therefore, we can classify D2D communications into four main groups. Figure 1.2 illustrates such a classification.

1) Device relaying with BS-controlled link formation; this type of communication is usually beneficial for a device that is located at the cell edge and that therefore has weak signal. In this scenario, the devices are communicating through other devices. This communication scenario helps the device to reach a better quality of service and simultaneously it increases the battery life. For control link formation the BS is in active communication with the relays [**8**].

2) Direct Device to Device communication with BS-controlled link formation; in this type of communication the devices are transferring data between each other without the involvement of a BS, but for the link formation they are supported by the macro BS [**8**].

3) Device relaying with device-controlled link formation; in this scenario, the devices are communicating through other devices and link formation is managed by devices [**8**].

4) Direct Device to Device communication with device-controlled link formation; in this sort of communication, devices communicate directly with each other and the link formation is self-controlled without participation of any BS. So, the whole spectrum that is available should be employed by the source and destination devices to create the least interference with other devices [**8**].

Ultra-Dense Networks, which are necessary to satisfy the ever-growing traffic demand. In order to reach ultra-dense layouts, HetNets or heterogeneous networks are key factors. With HetNets, RANs will be more dynamic and flexible, but the emergence of ultra-dense networks will generate new challenges of interference, backhauling and mobility. To overcome the mentioned challenges and issues a new network layer is required, and interference reduction techniques should be more dynamic and flexible and open to changes and variations [**1**].

Millimeter waves, that represent one of the most innovative aspects of 5G cellular networks, and in general a new frontier for the wireless industry [**1**],[**9**]. Millimeter waves are expected to allow operators to cope with the huge traffic growth expected in coming years, and to offer multi-giga bit per second data rates to end users.

1.1.2 5G Wireless Future

Now we consider in more details the characteristics of the mmwave technology, which is a key component of the study we conduct in chapter 4, where we assume

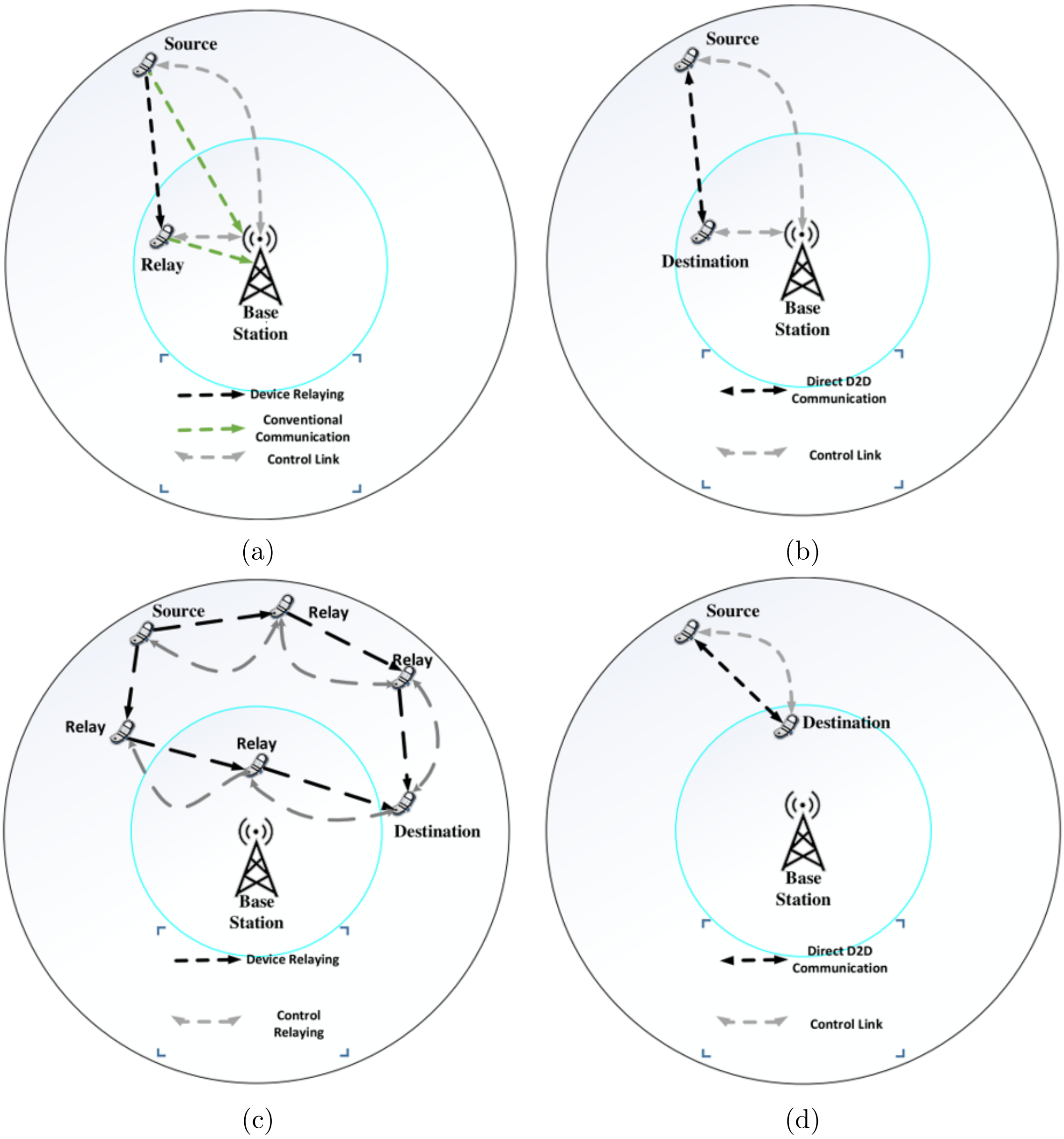


Figure 1.2: Device to Device communication description. Plot (a) refers to device relaying communication with base station controlled link formation, plot (b) refers to direct device to device communication with base station controlled link formation, plot(c) refers to device relaying communication with device controlled link formation, plot (d) refers to direct device to device communication with device controlled link formation [**8**].

to be able to exploit mmwaves for backhaul connections.

The adoption of millimeter waves in the range from 10 to 100 GHz for wireless communications requires large gain steerable antennas, massive MIMO and adaptable beamforming, as well as components that can work efficiently in mmwave frequency band [**10-12**]. The use of mmwave carrier frequencies, and the consequent wider bandwidth allotments permit very high data rates, as well as higher numbers of users and reduced latency. The bandwidth increase will be instrumental for serving densely populated areas with much better QoS with respect to the current 4G networks. One of the weak points of mmwaves is the higher attenuation in their propagation through air, and especially through materials due to obstructions along the path between transmitter and receiver. Because of this, the distance between mmwave terminals and BSs must be kept small. In this dissertation, where mmwaves are used to connect small cell BSs to macro BSs through a wireless backhaul connection, the distance between the small cell BS and the macro BS must be short. As a result, a denser deployment of BSs on the service area must be implemented.

The most important propagation concerns with regard to millimeter wave propagation for 5G cellular communications are:

1) Path Loss, the path loss in free space depends on the frequency of the carrier. It goes from 0.06 dB/km at 18 GHz to 7.5 dB/km at 73 GHz.

2) Blocking, millimeter wave signals are susceptible to blocking due to obstructions. The attenuation due to obstruction depends on material. Typical values go from 130 dB/m at 18 GHz to 420 dB/m at 73 GHz [**13**].

3) Diffraction, millimeter wave signals encounter smaller amounts of diffraction than microwave signals and show reflective propagation that makes them much more at risk in presence of obstructions.

4) Link Access, the narrow beams generated by antennas with high directionality create problems in the set up of links among users and BSs [**13**]. This problem becomes increasingly challenging in high mobility contexts.

It must be observed that signal absorptions also bring a positive aspect, since they increase the separation of cells by attenuating the interference generated by more distant BSs with respect to the one serving the user. Experiments have shown that the propagation losses of millimeter wave frequencies are manageable, but require controlling the energy of the emitted beam by means of large antenna arrays [**13**].

Signal propagation and interference are extremely relevant whenever wireless backhaul is used to link repeaters or small cell BSs to the wired part of the RAN. In these cases, the use of mmwave technologies is very interesting because of beam directionality and reduced interference, both with other mmwave links and with traditional microwave cellular signals. This aspect is very relevant to the topic of this dissertation, since backhauling of small cell BSs on vehicles is only possible through wireless connections.

1.2 Need for Densification

Network densification is the approach that allowed RANs to cope with the massive increase of traffic of the past decades, more than the introduction of new generations, new spectrum, and new transmission technologies. Experts estimate that denser cell deployments had an effect which is over an order of magnitude higher than all other improvements combined. Network densification is a key focus of this dissertation, and before getting to the details of our research questions we address three general issues: 1) what is network densification, 2) what is the importance of network densification, and 3) what is the role of the network operators in implementing this novel technology and strategy.

What is Network Densification? A simple definition for the network densification approach is to deploy additional cells in areas with high traffic, so as to grow the total available capacity thanks to the fact that each cell provides additional capacity to the RAN in the deployment area[2].

The cells that are deployed in the areas that are short of capacity can provide more capacity where it is required and even help offloading traffic from neighboring sites. Inner-city areas (also called dense urban areas) and popular venues (such as stadiums) are possible candidates for network densification due to high density of mobile users.

In order to provide densification in an area served by a set of macro BSs, which form the so-called macro cell layer, some supplementary low-power BSs are added to the network within the coverage area of the macro cell layer. This leads to what is known as heterogeneous or multilayered network architectures. The small, pico, micro cells layers (we will use the term small cell, and we will call small cell BSs the base stations generating small cells) do not necessarily cover the whole area served by the macro cell layer, and are just situated to boost capacity and improve the data rates where it is needed.

Network densification is sometimes considered to include densification in space (dense deployment of small cells) and densification in frequency (use of wider radio spectrum in different bands). Spatial densification is achieved by self-organizing networks and intercell interference management. Moreover, in order to achieve the benefits of network densification, it is necessary to pair it with adequate backhaul densification (i.e., adequate increase of backhaul link capacity)[3].

[2] https://www.electronicsforu.com/resources/learn-electronics/bdma-technology-5g-network

[3] https://www.electronicsforu.com/resources/learn-electronics/bdma-technology-5g-network

1.2.1 Network Densification Challenges

Network densification can be very effective in increasing network capacity, but it also raises several challenges that are shortly discussed here.

Interference

One of the main challenges that network densification must face is the increase of the interference in the network. This is especially true in HetNets (heterogeneous networks that comprise different cell sizes), where interference management has been one of the difficulties. The problem is critical when BSs with different coverage areas use the same licensed frequency spectrum portion.

If we consider a two-tier HetNet, we can discuss two different types of interference in the HetNet.

Co-tier Interference – Co-tier interference, or intra-tier interference, refers to the interference between BSs of the same tier. Particularly critical in the scenarios considered in this dissertation is the interference between small cell BSs, since they are located in random locations, possibly at very short distances from each other [20].

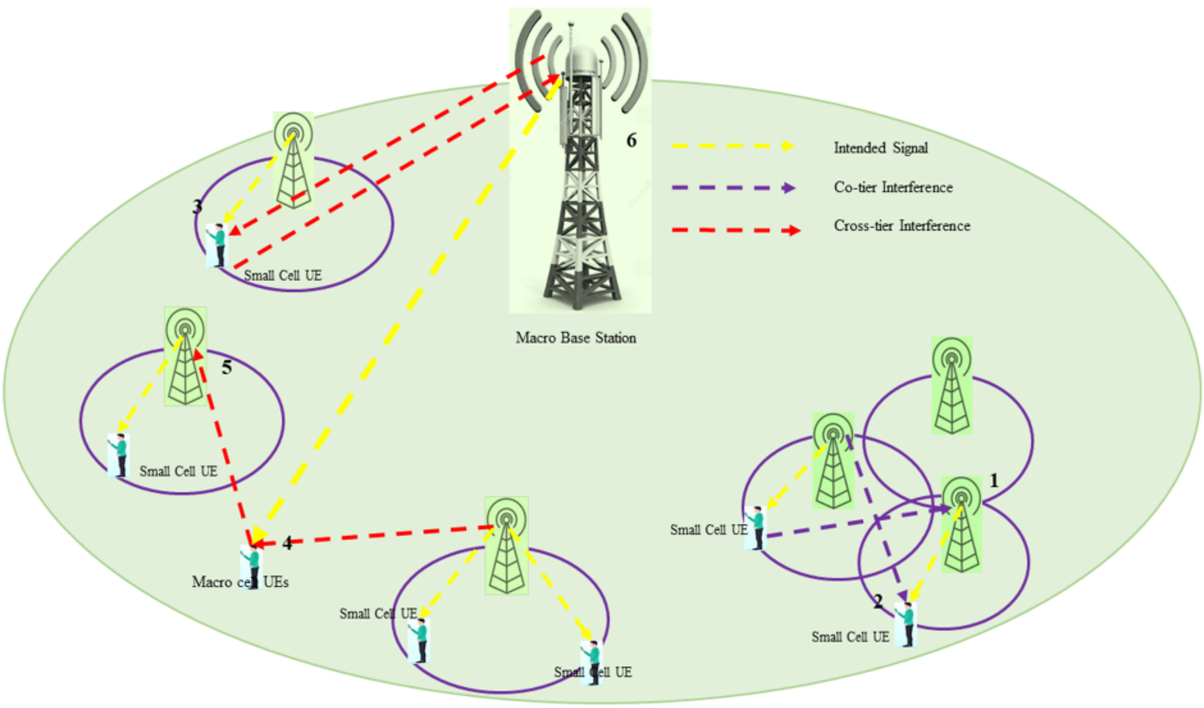


Figure 1.3: Inter tier interference and intra tier interference in uplink and downlink scenarios.

The interference scenario is depicted in Figure 1.3, Case (1): The uplink co-tier interference between the small cell BS and the small cell user equipment (UE). Case (2): The downlink co-tier interference between the small cell BS and a small cell UE.

Cross-tier Interference – Cross-tier interference is generated by the BSs of different tiers. This type of interference is caused because of two main reasons: the large difference between the transmitted power of BSs of the two tiers and the asymmetry of the covered areas by cells of each of the two tiers (and the corresponding path loss) [**21**].

Case (3) in Figure 1.3 is showing downlink cross-tier interference between one of the small cell UEs and a MBS. Case (4) depicts downlink interference between a macro cell user at the edge and small cell BS. Case (5) illustrates uplink cross-tier interference when a macro UE is at the cell edge and must transmit high power in order to make up for path loss and shadowing effect. Case (6) displays uplink cross-tier interference of a MBS and a small cell UE.

Energy Efficiency

Energy efficiency can have many possible definitions. Energy efficiency cannot just be the amount of power consumed by a network, since many aspects need to be accounted for: amount of carried traffic, quantity of served users, extension of the served area, frequency spectrum used, etc. As a result, a comprehensive definition of energy efficiency is complex. Here we just use the following definition: the network energy efficiency is the aggregate bit rate obtained on the bandwidth of 1Hz for a given power. This means that we measure energy efficiency in bits per second, per Hertz, per kilowatt. In other words, energy efficiency equals spectral efficiency divided by network power. Given the increase in the number of BSs, network densification is likely to increase power consumption unless the increase in capacity and spectral efficiency compensate the power increase.

Positioning of a huge number of BSs

The energy efficiency of network densification was studied for example in [**22**], where the authors show that the initial steps of densification allow energy efficiency to grow considerably, but as densification continues, energy efficiency experiences limitations due to interference increase. This is truer in an indoor space since in an indoor space the spectrum is not utilized well [**2**].

Positioning mechanism

The strategy that is used in small cell deployment has a remarkable effect on the amount of interference and energy efficiency of the network. There are three different strategies to deploy small cells and manage user attachment. The first user attachment strategy is an open access method in which any user around the small cell can connect to the small cell [**2**]. The second strategy is called a closed access procedure in which just the users that are subscribed to the small cell are allowed to connect to the intended small cell BS. The third scheme is the hybrid scheme in which only subscribed users can enjoy the quality of service. The effect of these accessing schemes have been discussed, and it was concluded that closed procedures are less efficient than open access methods in terms of energy.

Operation strategy

When it comes to network operation, many factors are affecting the network energy efficiency. Power control and sleep modes are among the most effective to increase energy efficiency. Currently, in LTE systems all BSs remain active 24/7, even if they are not in use by services. Controlling the activation and deactivation of BSs according to traffic conditions can greatly reduce the network power consumption, as well as reducing non-essential overheads and idle mode signaling.

Handover and Mobility Management

Handover allows cellular networks to guarantee the quality of service of active users [**23**], since the handover process helps a UE that is in a cell and is already connected to the cell BS, to transfer its connection to another cell BS while keeping the quality of service at an acceptable degree.

Managing handovers and associations of UEs to BSs have always been a challenge in RANs of different generations, and network densification, by adding many small cells to the network makes this already existing challenge even more complex to tackle. The issue is that the area that is covered by small cells is much smaller in comparison with macro cells, so that handovers become more frequent [**21**]. This calls for smart associations of UEs to BSs, possibly accounting for UEs mobility characteristics.

1.2.2 How does network densification with vehicles contribute to the state of the art?

In this dissertation we explore the possibility of an opportunistic crowdsourced densification of RANs, exploiting SC BSs carried by public or private vehicles along their daily routes. We call MoBS the mobile SC BSs, and mobile mobile network (MOMONET) the approach to densification based on MoBS.

The idea of exploiting cars as active components of a telecommunication network is not new. Many studies in vehicular networks have investigated the possibility of combining an Access Point (AP) in cars and a cellular interface. In this case the connected cars can play the role of a moving access point and can provide connection for UEs in the car proximity. For example, some studies considered mobile vehicular gateways that use Wi-Fi for vehicle-to-vehicle (V2V) and LTE for vehicle-to-infrastructure (V2I) communications.

Other studies introduced virtual APs that helps to increase the access range of roadside access points. In this case, vehicles receive a message, they store it, and finally re-broadcast it into the areas that are not covered. In some cases even parked cars and vehicles were exploited, in addition to roadside units, to increase the quality of video downloading and other services.

All these previous studies and approaches are different from the MOMONET concept that we investigate in this dissertation, since here we employ vehicles as a support for small cell BSs to achieve RAN densification.

MOMONET creates a logical integration of the small cell mobile BSs carried by vehicles with the traditional fixed BSs. Specifically, thanks to the correlation between vehicle density and telecommunication traffic load, the mobile small cell BSs add capacity, where and when needed, to better serve the end users of a traditional RAN. This capability lets end-user terminals freely transfer their services between a macro BS and a mobile vehicular small cell BS.

1.3 PhD Publications

Part of the work reported in this dissertation was already published in international conferences and journals.

Most of the results in Chapter 3 derive from the paper

- Foroogh Mohammadnia, Marco Fiore, Marco Ajmone Marsan, "Adaptive densification of mobile networks: Exploring correlations in vehicular and telecom traffic," 17th Annual Mediterranean Ad Hoc Networking Workshop (Med-Hoc-Net 2018), 20-22 June 2018, Capri, Italy.

Most of the results in Chapter 4 derive from the papers

- Foroogh Mohammadnia, Christian Vitale, Marco Fiore, Vincenzo Mancuso, Marco Ajmone Marsan, "Mobile Small Cells for Adaptive RAN Densification: Preliminary Throughput Results," 17th IEEE Wireless Communications and Networking Conference (IEEE WCNC 2019), 15-18 April 2019, Marrakech, Morocco.

- Marco Ajmone Marsan, Foroogh Mohammadnia, Christian Vitale, Marco Fiore, Vincenzo Mancuso, "Towards mobile radio access infrastructures for mobile users," Elsevier Ad Hoc Networks, Vol. 89, pp. 204-217.

The contents of chapter five are still not published.

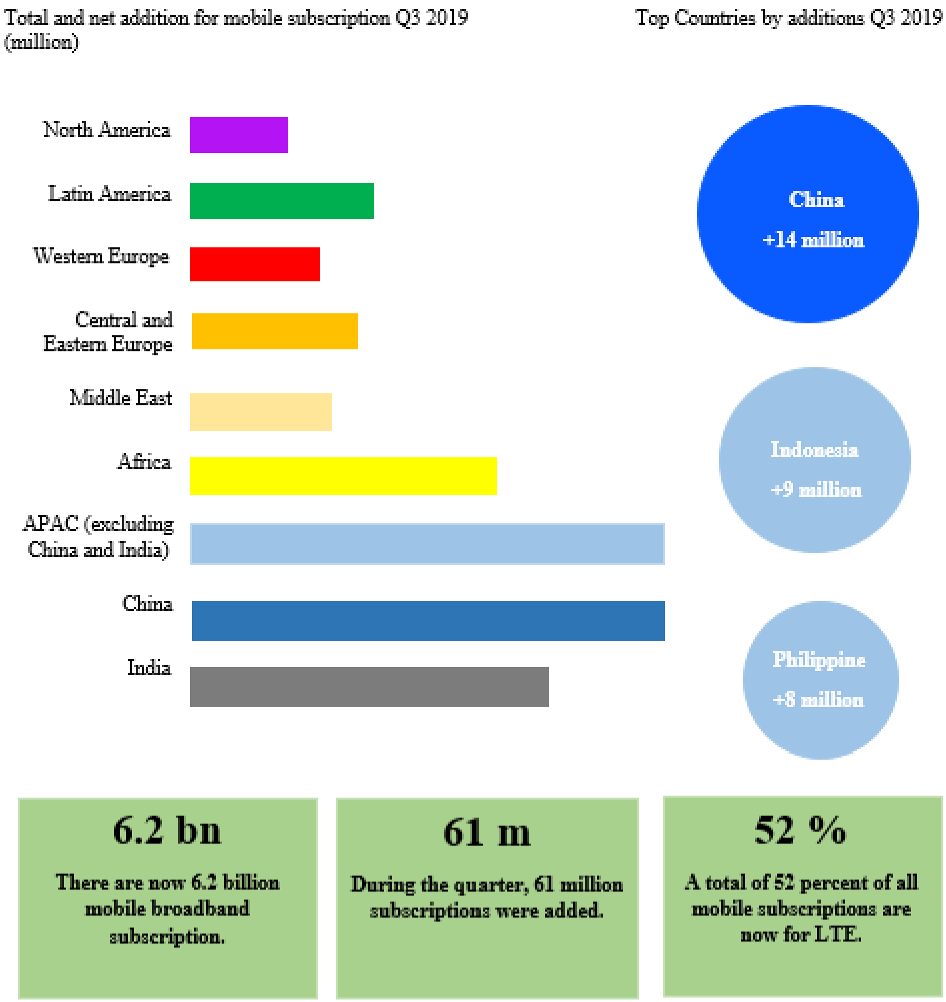


Figure 1.4: Explosive growth of mobile traffic and mobile subscriptions[a]

[a]https://www.cisco.com/

Chapter 2

Densification with Small Cells on Vehicles

2.1 MOMONET

The use of large numbers of Small Cells (SCs) for the densification of Radio Access Networks (RANs) is regarded as a most promising approach for the provision of broadband services to large numbers of mobile network end-user terminals. However, the deployment of the SC base stations (BSs) necessary to make dense RANs a reality implies huge investments, which Mobile Network Operators (MNOs) are reluctant to approve in a period of declining profits.

A major reason for the high SCs deployment cost is that peaks of bandwidth demand vary significantly in time and space [**25**], [**26**], thus making the SC utilization highly variable, and low on average. For instance, users tend to commute from home to work in the morning of working days, making business districts crowded during working hours.

In this period, MNOs need high-capacity RANs therefore, a dense presence of SCs needs to be located in business areas. However, after work, users move out of their offices, so that the demand for RAN capacity becomes much lower in business districts, and many of the installed SCs become redundant. The opposite is true for the residential areas, where capacity is mostly necessary in the evening, so that a dense RAN layout becomes useful then, not during working hours.

In addition, during commuting and in the event of traffic jams or accidents, dense RAN layouts are needed where road traffic is congested. The cost of a RAN which is dense everywhere can be extremely high, and the associated long periods of low utilization imply a poor return on investment for the MNO.

A possible approach to make high RAN capacity available where and when needed is to use mobile small cell base stations (MoBSs) carried by vehicles. This approach leads to a network where not only end users, but also part of the network infrastructure is mobile. We name such an architecture a Mobile Mobile Network or MOMONET.

The idea of mobile BSs is not new. Operators have exploited truck-mounted BSs for the quick provision of coverage in areas where service was not otherwise available or where additional capacity was needed [1]. Some truck-mounted BSs were periodically moved from beach areas in summer to skiing resorts in winter. However, in this thesis we discuss a new, much broader use for mobile BSs, looking at the possible impact of vehicle-mounted SC BSs on the implementation of dense RANs based on SCs.

The feasibility of RAN densification with vehicle-mounted SC BSs raises a number of research questions that require accurate investigation. For example: the selection of algorithms for the association of end users to BSs and for handover mitigation; the management of interference among transmissions from traditional BSs, from small cell BSs, and from end users; the selection of technologies for the implementation of small cell BSs on vehicles, and for the implementation of backhaul wireless connections; the reduction of the impact of the energy consumption of small cell BSs on vehicles power systems and the impact of self-driving cars.

Vehicles have the desirable property of moving together with end users, so that in a business district during working hours we normally have both many end users and many vehicles. If a significant fraction of those vehicles carry a SC BS, a temporary dense SC deployment is created. Quite nicely, this temporary dense SC deployment will be recreated in residential districts when drivers return home with their cars, as well as in traffic jams.

The deployment of MoBSs on public and private vehicles has the potential to significantly reduce investment cost for MNOs. Moreover, it can open an entirely new market, where wireless bandwidth generated by MoBSs plays the same role of energy produced by rooftop solar panels.

MOMONET like any other new technology has its own pros and cons and challenges. The main network challenges posed by mobile base stations are interference management, energy efficiency and mobility and handover management that were discussed briefly in chapter one, but in addition to these issues, developing efficient backhauling solutions for MoBSs is one of the most important challenges in 5G cellular systems where many small cells are deployed on vehicles. The backhaul solution that is mostly matched to the availability of existing backhaul infrastructures and service demand can consist in a heterogeneous backhauling. Indeed, network operators will face the backhauling challenge that requires that traffic should be trsansfered from fixed small cells to the core network in an efficient and cost-effective manner. Among backhaul solutions for small cell mobile base stations, there are wireless and wired solutions that have their own pros and cons. According to some research outcomes, the most effective solution for backhauling is the combination

[1] http://www.dael.com/en/telecom/cell-on-wheels

of different options. In our scenario, due to MoBS mobility we just discuss the wireless options.

Overall, this thesis tackles some of the key aspects of RANs comprising MoBSs, the correlation between telecom traffic and vehicle densities, the achievable throughput and fairness, and the impact of backhaul links from MoBSs to fixed BSs.

Key challenges to investigate will consist in understanding to which extent the mobility of base stations is effective in bringing capacity and services where traffic peaks, and to design a broadband and reliable wireless mobile backhaul to connect MoBSs to the core of the network. There are no data and no results available in the literature on these challenges.

2.2 MOMONET Concept and Architecture

2.2.1 The Novelty of MOMONET

RAN densification requires the deployment of large numbers of small cells in those areas where the number of users, both humans and machines, and the traffic they generate, is very high, at least for some significant portion of time.

However, the number of mobile users and the level of traffic they generate exhibit remarkable spatial and temporal variations [**25**], [**26**].

Users, humans in this case, normally move from home to work in the morning of working days, and this makes business districts crowded during working hours, approximately from 9 am to 6 pm, in this period, MNOs need the capacity of many small cells of their dense RAN in business areas.

However, after 6 pm, users move out of their offices, so that the RAN capacity necessary in a business district becomes much lower, and many of the installed small cells become redundant. The opposite is true for the residential areas, where capacity is necessary in the evening, so that a dense RAN layout becomes necessary then, not during working hours.

What is more, operators experience daily massive migrations of users, which can last hours in the morning and in the evening. During such migrations, there is a huge demand for data and infotainment. Therefore, network capacity needs to be provided not only in business and residential districts, but also on the move. This problem is not fully new; it was partially present also in old-style fixed telephone systems, and it is one of the key management challenges in present day RANs.

Indeed, in modern cellular networks, radio resources, active frequencies, are moved following usage patterns, so that precious bandwidth is not wasted in low traffic areas. However, moving wireless bandwidth does not reduce the investment cost for infrastructure, and BSs that in extreme cases can remain inactive for days, because they serve areas that only become crowded during weekends, increase the cost of the infrastructure, hence of service.

The mobility of users and terminals, coupled with the explosion of traffic, has made the situation much worse, and the increasingly common habit of using smartphones during crowded events, such as a rock concert or a football match makes the problem even harder.

The traditional approach of dimensioning for peak traffic is extremely costly, requiring the dense deployment of small cells in both business and residential districts as well as extremely dense coverage of stadiums, and leads to low resource utilization hence low return on investment for long periods of time.

In order to obtain the wireless bandwidth that is necessary to meet the forecasted explosive increase of mobile traffic, it will be necessary to have a large number of small cells in business districts during working hours, in residential districts during evenings, over commuting paths at the beginning and the end of working shifts, and around stadiums during special events not to mention city squares during protest gatherings, highway portions in the event of a traffic jam, etc.

One possibility is to deploy a dense small cell coverage in all areas, switching them on and off as needed[2]. This can bring substantial savings in operational expenditures (OPEX), especially those related to energy [**27**], but does not alleviate the capital expenditures (CAPEX) due to cell deployment.

Vehicle-mounted small cell BSs for the implementation of dense RANs, exploit the fact that vehicles move with people, and so does network capacity in a MOMONET. This is a disruptive concept with respect to what has been done so far.

MOMONET will use the correlation between the spatio-temporal distribution of vehicles and that of traffic demand, in order to physically move small cell base stations – mobile BSs (MoBSs) from business areas to residential areas and back, so as to have the capacity of those cells where and when needed. This will lead to a dramatic decrease in the amount of installed RAN infrastructure and will ensure high bandwidth connectivity as well as the support of a very large number of simultaneous connections.

Indeed, vehicles have the nice property of moving together with end users (here we mostly refer to humans, but vehicles can also be host to several connected devices).

If a significant fraction of those vehicles carry a small cell BS, a temporary dense small cell deployment is created. Quite nicely, this temporary dense small cell deployment will be recreated in residential districts when drivers return home with their cars, and on the road as they move.

Note that the correlation between number of vehicles and number of end users as well as between telecommunication traffic and number of potential MoBS positions, was shown in several works, starting with the "Real Time Rome" project of Carlo

[2] http://www.dael.com/en/telecom/cell-on-wheels

Ratti's MIT Senseable City Lab[3],[4], but must be carefully measured with recent real data.

In addition, parking lots are good candidate location to provide some form of infrastructure support to a MOMONET, e.g., by providing recharging facilities for electric and hybrid cars in order to enhance the autonomy of the MoBS equipment.

A similar consideration can be made for the need for sensing and computing, and the resulting traffic demand. MOMONET will provide the ideal ecosystem for an effective implementation of the MEC paradigm and, hence, the support of services requiring extremely low latency and/or low overhead. Indeed, local sensing and computing needs are also typically proportional to user density.

Many IoT services have a local sensing component. Events of some significance for an IoT based system (e.g., a road accident, a sudden medical condition on a pedestrian, or the freeing up of a parking slot) are usually more frequent when and where users concentrate.

In addition, the need for computing resources (e.g., the algorithm for inferring a given medical condition, for detecting the degree of occupancy of a parking space, or for distinguishing a road accident from a normal traffic jam) correlates positively, spatially and temporally with user mobility patterns. This is the reason why much of IoT sensing is devoted to either characterizing the human patterns directly, or to capture conditions and events (traffic jams, aggressions) whose frequency and distribution in space are tightly related to human patterns.

Similarly, infotainment service demand moves with people, and thereby computing resources used to provide those kind of services (e.g., video streaming and caching), which are now deployed at the edge of the network (following the MEC paradigm) need to move as well, to be used in an efficient way.

Of course, the dense small cell RAN layouts resulting from MoBSs have different properties with respect to the dense small cell RANs carefully planned by network operators, and this poses a number of challenges.

First of all, while planners of BS deployments carefully select the BS positions so as to avoid interference and maximize the effectiveness of BSs in serving users, drivers obviously drive and park their vehicles with no awareness of network planning issues.

This results in random BS deployments, where by random we do not mean that cars can be in any position over the area of interest, rather, that their positions are not planned and not controlled, which call for careful adaptive network management approaches. In addition, the movement of MoBSs, coupled with the movement of end users, increases the overall dynamics of the RAN, and can generate a much greater number of rearrangements in the network.

[3]MIT. Available: http://senseable.mit.edu

[4]MIT. Available: http://senseable.mit.edu/realtimerome

With regard to the type of vehicles that could be good candidates to carry MoBSs, public transport vehicles, such as buses and taxis can be a good choice, since they are moving most of the time, with higher presence in those areas where a larger number of people are present.

In addition, electrical or hybrid vehicles, either private or belonging to a car-sharing fleet, are quite interesting, since the car density correlates well with the density of end users, as we discussed, and the presence of a large energy storage unit in the car enables operation of the MoBS also while the car is parked and the engine is off.

The opposite is true for traditional combustion engine cars, which have very limited power storage available when the engine is off, but can devote extra power to the MoBS while moving. An advantage of vehicles of a car-sharing fleet is in the fact that a car-sharing operator would be in a position to sign an agreement with one or more MNOs or infrastructure/tower operators for the provision of small cells as a service. Indeed, the car sharing company can play a role similar to that of companies that manage telecom sites and towers for base stations.

2.2.2 MOMONET Architecture

The key idea of MOMONET is to provide a new paradigm for mobile cells in a mobile world. Such a paradigm will be designed base on the analysis and understanding of spatial and temporal behaviour of vehicles and telecom traffic. Specifically, MOMONET's contribution to research and innovation is twofold; the first is developing new analytical tools for the study of mobility data, and the second consists in designing a cellular network theory for MoBS-operated networks.

Developing this new research area is mandatory since there are many differences between a MOMONET and a RAN that just comprises fixed macro and micro base stations. The activity in MOMONET will specifically focus on the realization of the following key actions.

1. Gathering of empirical data obtained from the measurement of the behaviour of vehicles of the car-sharing operator participating in the consortium of data stored in open online repositories, and of data traffic traces.

2. Data analysis on mobility and traffic generation pattterns.

3. Development of analytic and modeling tools for describing and predicting data traces and correlation between mobility patterns and capacity demand patterns in space and time.

4. Interference mitigation for the optimization and coexistance of mobile radio access and mobile wireless backhaul.

5. Developing new mobile backhaul management methods that use wireless technology more than fibre technology.

6. Smart management of mobile handovers.

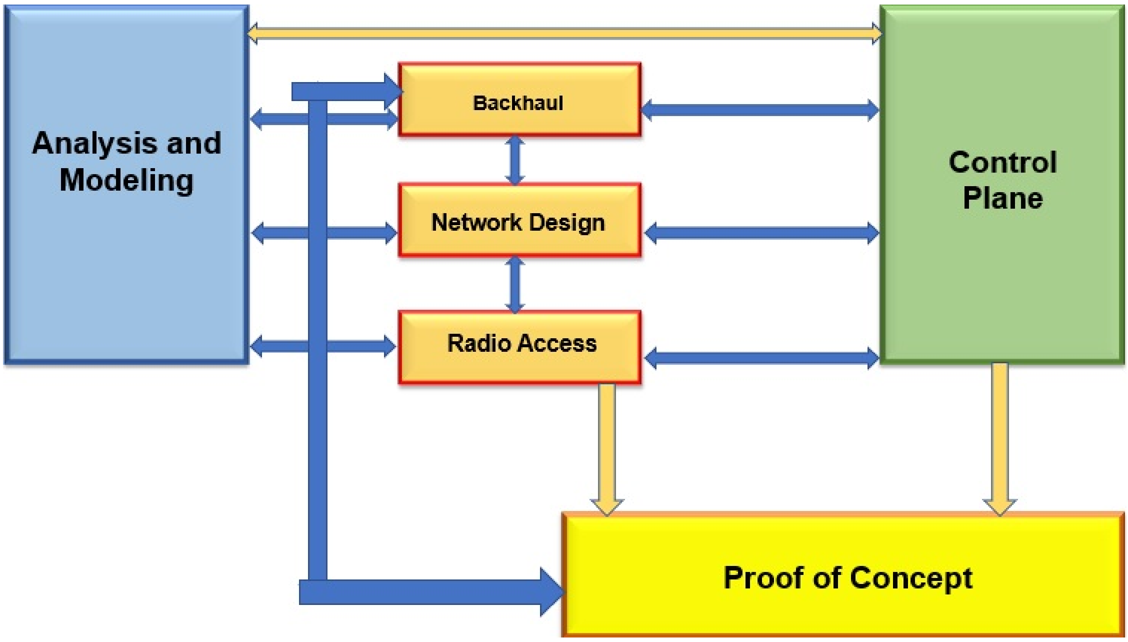


Figure 2.1: Overall methodological approach to research and innovation in MOMONET

7. Adoption of the SDN/NFV paradigms to monitor all available resources on the backhaul and also to virtualize network functions across MoBS. Figure 2.1 depicts how research and innovation on data analysis/modelling and architectural design will be coordinated.

Next, we briefly discuss the data analysis and network design in MOMONET, which constitute the main focus of this dissertation.

1. Data Analysis: This component in the figure relates to mobility/data traffic analysis and statistical models that describe user behaviours and traffic generation. The purpose of mobility and data traffic analysis is to do a descriptive statistical analysis of spatio-temporal patterns found in user behaviour, vehicle behaviour and data transmissions. Therefore a precise analysis of the available data is required to make the best architectural decisions. The data analysis will be performed on multiple scales and according to multiple data sources. Being multi dimensional does not just refer to the analysis variables, but also to multiple dimensions considered in the analysis. For example, space

and time are the key variables to indicate when and where the capacity should move.

Some descriptive and visual analysis were complemented by visualisation libraries in Python to identify the data patterns, correlations, etc. The multi dimensionality of the data, our analysis and also the large volume of the data analysis posed a great challenge to implement the visual part of our analysis and study.

Based on the statistical analysis that is performed on collected data, models are created that serve different purposes.

- Interpretable Models: This type of models describes the spatio-temporal behaviour of end-users, vehicles and data traffic. Interpretable models let researchers identify which variables are related to the problem.
- Predictive Models: The predictive models allow forecasting the system state in the near future (short-term forecasting), i.e., forecasting the value of the variables that are critical for network operation. Usually predictive models are more complex than interpretable models, but forecasting is necessary and critical for network management and operation. In fact, predictions help network operators to anticipate possible problems, and help to optimize the network resources use.

2. Network Design: This part of our work relates to small cell mobile base stations and their connection to the rest of the network and to end users.

   An overview of the basic MOMONET networking architecture elements is presented in Figure 2.2. The use, connection and control and joint optimization of such elements is the core of MOMONET operation. MOMONET basic architecture will include the following elements:

   - Mobile Users: End user terminals (UEs) that use standard or new communication technologies, like Visible Light Communication (VLC), millimeter wave communications (mm Wave), white spaces cognitive radio.
   - MoBSs: Mobile small cell BSs possibly implementing just part of a BS protocol stack. These are new components that are specific to the MOMONET concept, and they can cover a large variety of transmission technologies and frequencies.
   - Fixed cell BSs: Legacy BSs (defining macro cells, or small cells, or umbrella cells), as per the 3GPP specifications for 4G networks and beyond. In addition to legacy 4G/5G radio access, these BSs will implement the wireless backhaul for MoBSs, and therefore will support wireless backhaul technologies to be studied and identified in the project, e.g., cognitive radio over whitespaces, mmWave with CPRI-like connectivity, etc.

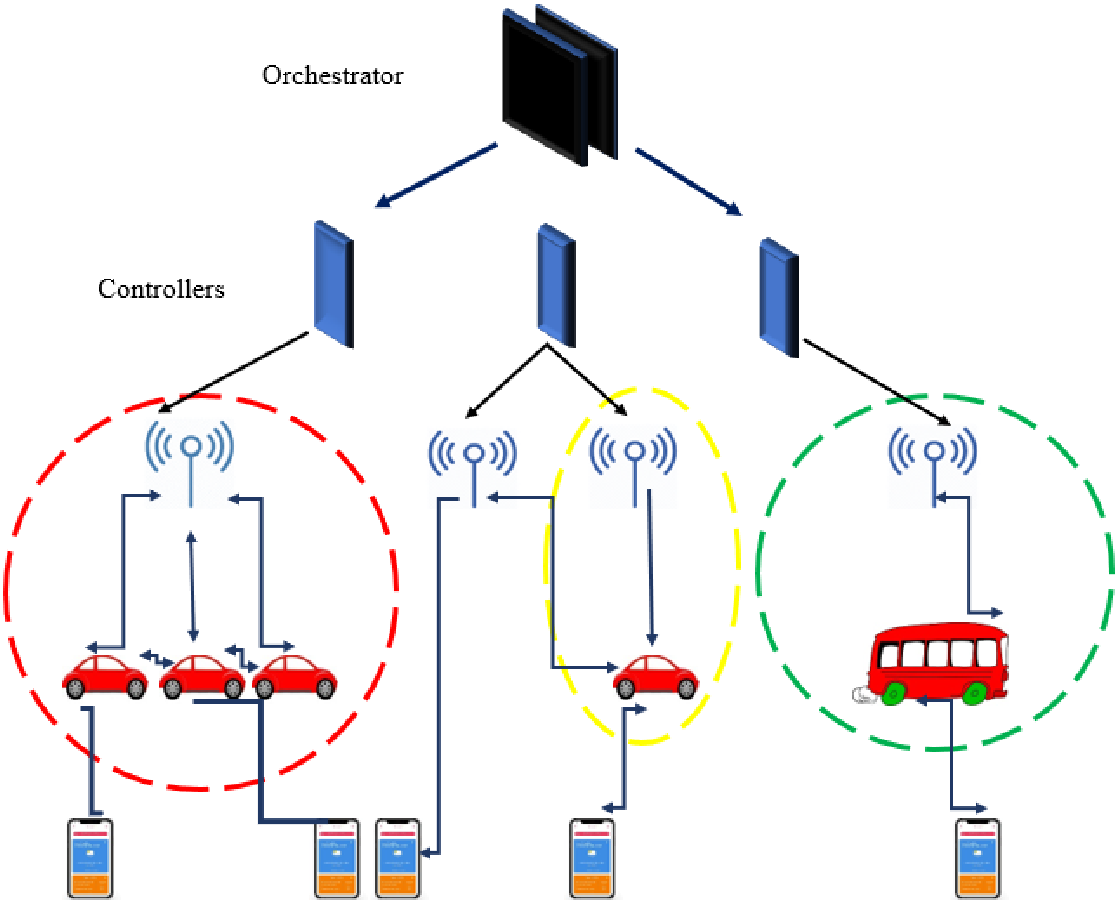


Figure 2.2: MOMONET Network Architecture

- Controllers and Orchestrators: MOMONET is based on a control plane managed by a network of SDN controllers that are orchestrated by algorithms using data analysis results and context-aware machine learning algorithms. This includes controlling and coordinating network resources.
- MEC: MEC in MOMONET can provide low-latency and low-overhead services following the users in space and time as mobile base stations in MOMONET do. MEC requires dynamically handling network slices and computational resources.
- Backhaul: In the MOMONET scenario, backhaul is a wireless connectivity between mobile base stations and fixed macro cells.
- Radio Access: Connectivity between mobile users and base stations, fixed and mobile. The radio access protocols that are adopted in MOMONET will be partially new and partially using the existing cellular technologies.

  - MoBS to UE, and vice versa: The radio access for MoBSs, on the one hand, appears as a normal 4G/5G radio access. On the other hand, it must exploit the presence of multiple MoBSs to guarantee QoS and privacy/security to the users, e.g., by means of channel-opportunistic MoBS access, MU-MIMO, Coordinated Multi Point (CoMP) obtained by means of mobile coordinated BSs. (the MoBSs), interference cancellation, physical security, etc.
  - Fixed cell to UE, and vice versa: This part of the network/protocols adopts legacy4G/5G radio access protocols.
  - MoBS to BS, and vice versa: The MOMONET mobile backhaul must provide support for dynamic reconfiguration of single and multi-hop physical paths for reliable and continuous connectivity of MoBSs to BSs.

- Coordination between MoBSs: We consider some mechanisms and protocols to intensify coordination and cooperation between mobile base stations (under the control of an SDN controller, this is not discussed in this thesis).

2.3 MOMONET Application Examples

2.3.1 Commuting in megacities

Objective: To provide seamless broadband wireless access at low cost to masses of people living in megacities and commuting every day, with limited infrastructure cost for mobile network operators.

Scenario description – The most obvious application scenario for MOMONET refers to a dense urban environment, such as a megacity of several million inhabitants, which is subject to the daily movement of large numbers of people that commute from home to work in the morning, and return home in the evening. This daily commuting involves people, vehicles, today mostly private cars and public transport vehicles; tomorrow probably mostly vehicles of car-sharing services, and telecom traffic.

Telecom traffic is generated by human end users, and by the smart objects that travel with them and their cars. What end users want, is wireless access connectivity in the place where they are, which means in their work place during working hours, along their route during their commute, and at home in other periods. Providing this capacity with traditional means implies the deployment of enough broadband wireless capacity in most places, which is extremely costly for mobile network operators.

The MOMONET approach will provide adaptive bandwidth where needed, when needed with low cost.

2.3.2 Flash crowd

Objective: To provide broadband wireless access at low cost to sporadic gatherings of end users in places with otherwise low data traffic demand.

Scenario description – This scenario refers to environments such as a portion of a highway in a rural area, where the car density is normally low, and consequently the telecom traffic demand is low. However, if an accident along the highway creates a traffic jam, the vehicle density becomes extremely high, and the telecom traffic explodes, because of both the end user density and the need to communicate.

What end users want in this case, is broadband wireless access connectivity in an area where an MNO has no incentive to deploy it. Providing the necessary wireless capacity with MoBSs carried by the vehicles trapped in the traffic jam is quite natural. This is what will become possible with a MOMONET.

2.3.3 Disaster recovery

To provide possibly broadband wireless access for emergency as well as standard applications in disaster areas.

Scenario description – The habit of being connected anytime and anywhere has driven our society into a critical reliance on telecommunication infrastructures and this poses one of the biggest threats to our society as a whole. Resiliency of telecommunication infrastructures during and after natural disasters such as earthquakes, hurricanes, floods, etc., has become one of the biggest challenges in our modern connected world.

In recent cases, such as the Katrina hurricane in 2005, and the 2009 earthquakes in Abruzzo, we have seen that communication infrastructures can be wiped out or at least be massively disrupted. This infrastructure most particularly includes cellular operations that can be disrupted locally or in wide areas due to power outages or equipment failures.

The MOMONET concept can be a possible early solution to communication needs in disaster areas before normal infrastructure can be made operational again. Vehicles equipped with a MoBS can connect in a multihop manner to bridge the communication gap during or after disasters.

Given the wide availability of vehicles both moving and parked, a temporary dynamic infrastructure can be set up and used, first of all for emergency communication services, but also to allow end users to connect to relatives.

2.3.4 Quantitative Digital Divide Reduction

Objective: Dynamically bring high quality, high bandwidth connectivity to marginal areas, both in developed and in developing countries.

Scenario description – Global communication access has been included in fundamental human rights. Yet, connectivity and network availability are far from granted to anyone. Even when connectivity is available, its quality is extremely varied, and all areas in the planet with a low population density and possibly less affluent population remain poorly connected.

The availability of MoBS on vehicles immediately opens the possibility of bringing connectivity to the place where the vehicles go, be in a suburban area in Europe or a remote village in developing countries.

A vehicle mounted MoBS can use much more efficient antennas than portable devices, so it can naturally reach a backhaul/backbone access unreachable for handheld devices. As its main goal is not bringing connectivity in this place, but help densification in crowded places, the CAPEX is spread on different goals, so that good quality connectivity in remote areas may become affordable.

As a few MoBSs will serve many people and households, the presence of few equipped vehicles in interested areas is sufficient to achieve measurable benefits, and these few vehicles does not necessarily belong to the people that are affected by the digital divide: they can be public vehicles in some cases or the vehicles of the few affluent people in the local community in others.

2.4 MOMONET Research Issues

The research challenges raised by the MOMONET concept are many. Below we discuss some of the most obvious ones.

In order to verify the feasibility of the MOMONET approach, first of all it is necessary to study what is today the correlation between the density of end users (hence of telecom traffic), and the density of vehicles in urban areas, extrapolating the present situation to a likely evolution in dense urban areas increasingly served by vehicles of car-sharing fleets and in particular by electrical autonomous driving cars, as predicted by the recent report by Elon Musk of Tesla.[5]

Association of end users to BSs and handover mitigation

The selection of the BS to which an end user terminal should associate is very important for the effective operation of a cellular network and in particular of a MOMONET.

User association optimization has been thoroughly studied in the literature [**28**], [**29**], although the presence of mobile points of access has been so far

[5] E. Musk, "Master Plan, Part Deux". Available: https://www.tesla.com/blog/master-plan-part-deux.

evaluated only in terms of device-to-device communications [30]. Of course, for the system to provide the desired adaptive capacity, it is necessary that most terminals associate with the vehicle-mounted BSs.

MOMONET therefore needs novel algorithms, that, using a bias in the values of signal-to-noise ratio, favour the user association to MoBSs. Users associations of end users to base stations in our scenario will be discussed later in this thesis.

Management of interference among transmissions from traditional BSs, from small cell BSs, and from end users

The random, variable layout of small cell BSs carried by vehicles is very different from the carefully planned layouts of traditional cellular networks. This poses very challenging questions on the management of the interference among the transmissions between BSs (vehicular and fixed, unless they operate on separate frequency bands) and end user terminals, both in the downlink and in the uplink directions.

The approach which mimics the behaviour of today's RANs tries to reduce interference as much as possible, by carefully controlling transmission power and end user associations to BSs.

Technologies for the implementation of small cell BSs on vehicles

The implementation of a small cell BS on board of a car (MOMONET MoBS) poses a number of technological challenges, especially those related to the mechanical stress of the BS hardware, to the access to stable power sources, to the implementation of the BS antenna, to the choice of the most adequate digital transmission approach, and to the location of processing, storage and caching, for cloud, fog [31], and mobile edge computing applications [32].

Technologies for the implementation of backhaul wireless connections

One of the most delicate choices in MOMONET concerns the selection of the technology for the implementation of the wireless backhaul link. This is indeed critical in the MOMONET architecture, due to the high capacity requirements and the small cell BS mobility.

We will investigate the backhaul impact on performance according to some possible technologies for the backhaul implementation.

Energy consumption in MoBSs

The amount of energy necessary to run a MoBS on board of a vehicle must be determined as a function of the hardware efficiency, of the antenna gain, and of the required coverage. In general, it is possible to expect power consumptions of the order of a fraction of Watt, to cover distances of the order of a few tens of meters in the vehicle proximity. Power consumption may need to increase, to reach end users that are further away from the BS, or behind the obstacles.

Integration of the MOMONET approach in the Internet of Things (IoT) and in the smart grid (SG)

The fact that vehicles in a dense urban scenario travel all avenues, streets, and alleys, can be exploited for the implementation of many classes of the services that will come with the IoT and the smart grid (SG).

Indeed, any machine device and smart object is likely to be within reach of a street along which some vehicles carrying a small cell BS are passing during the day. This makes the connection between the machine end user terminal and the car possible, with minimum drain of the energy available at the machine end user terminal, for a limited amount of time, which can however be sufficient for transferring the few data necessary for the service in many cases.

Exploitation of the SDN and NFV approaches in MOMONET

The complexity of the MOMONET scenario is extreme, with MoBSs, possibly equipped also with CPU and memory for data processing and storage, moving around the dense urban area, over which also end users roam. Algorithms and protocols for a very dynamic control and adaptation of the numerous network operating parameters are necessary.

To this end, it is possible to exploit the SDN paradigm [**33**], which allows the virtualization and the real-time orchestration of network resources. In addition, the continuous variability of the RAN layout requires the virtualization of a number of network functions, as well as to implement services with approaches such as cloud, fog [**31**] and MEC [**32**].

Impact of self-driving cars

The diffusion of self-driving cars seems particularly attractive for car-sharing fleets, since they allow moving cars to the areas where they are most in demand, something that is done today for both bike and car-sharing fleets, at high cost, because of the need for human intervention. The same development may also happen for private electrical cars, as in the plans of Tesla Motors[6] and Google Self-Driving Car Project[7].

MOMONET can exploit the possibility of moving shared cars as desired in the periods when they are not hired. This allows the reduction of the unpredictability in the system, and to take some steps toward the optimization of

[6]https://www.tesla.com/blog/master-plan-part-deux

[7]Google Self-Driving Car Project. Available: www.google.com/selfdrivingcar

the spatial distribution of small cell BSs (jointly with the optimization of car positions as required by the car-sharing service).

Such a feature adds an important degree of freedom in the management of the dense small cell layout, and can provide significant performance improvements. For example, autonomous cars could be programmed to drive periodically over paths with minimal traffic, in order to collect IoT data to be transferred to processing centers.

Resilience

The complexity of MOMONET calls for significant attention to resilience [**34**], [**35**], and to the mitigation of possible vulnerabilities to malicious behaviours [**36**], [**37**], [**38**]. Note that the distributed nature of the MOMONET radio access system makes it suitable for the provision of wireless access capabilities in cases of partial failure of the traditional fixed BSs, and even for the limited connectivity of end users in the case of a total blackout of the fixed network elements [**39**], [**40**].

Indeed, in those cases (for example due to a natural disaster), the vehicle-mounted BSs allow the exchange of information among terminals within reach of the same BS, and even of multiple BSs, if an appropriate backhaul system has been implemented.

2.4.1 Tackled Research Issues

As we mentioned at the beginning of this section, in order to verify the feasibility of the MOMONET paradigm in practice, first of all we study the existence of correlation between telecom traffic and vehicular traffic. Indeed in chapter 3 of this dissertation we quantify the spatial, temporal and spatio-temporal correlation between these two traffics to visualize and verify the presence of correlation, and to identify the places and the time periods where a reasonable level of correlation exists, so that the MOMONET approach can be effective.

After studying correlation, we quantify the throughput and fairness improvements achievable with the adoption of the MOMONET approach. We do this by looking at specific areas within the city of Milano, like the main Railway Station and the Politecnico. In order to compute throughput and fairness, we account for the end user associations to MoBSs and BSs. The considered association algorithms select BSs according to the strongest received signal, hence to distance between end users and BSs (we assume simplified propagation models, not accounting for buildings layout). In order to reduce interference, we assume that some of the MoBSs within the area can be muted if they generate too much interference because they are very close to macro BSs or to other MoBSs. Muting is based on an interference threshold, but more elaborate approaches can be devised. We also investigate the possible wireless backhaul solutions for MoBSs, one of the most delicate design choices in MOMONET, and we quantify the impact of different backhaul solutions on network performance.

The rest of this thesis is structured as follows:

In chapter 3, I study the correlation between vehicular and telecommunications traffic, showing that, even considering only the very small percentage of vehicles for which real data are available, a visible correlation is present, both in time and in space.

In Chapter 4, I applied existing approaches so as to define the optimization problems that allow the computation of the maximum achievable throughput and fairness in a portion of a RAN comprising both fixed macro base stations and mobile small cell base stations. Numerical results show significant increases in both metrics. Increases are of the same order of magnitude as those achieved with regular placements of fixed small cell base stations.

In chapter 5, I discuss first the technological alternatives for the implementation of the wireless backhaul connection from mobile small cell BSs to fixed macro base stations and then the achievable throughput and fairness with three different wireless backhaul options.

Chapter 3

Exploring correlations between vehicular and telecom traffic

As discussed before, the MOMONET concept is based on the integration of traditional fixed macro/micro BSs with MoBS carried by private or public vehicles (buses, taxis, cars, vans, ...). This can provide a temporary network densification through the crowdsourced deployment of large numbers of small cells in the areas and in the periods in which traffic demand peaks, thanks to the correlation between vehicle and end user densities, hence traffic demand.

In this chapter we explore such correlation using real data for both telecom traffic and vehicle movement.

3.1 Dataset Description

For the analysis of the correlation in time and space between vehicular traffic and telecom traffic, we used three datasets.

The first dataset is the telecom traffic dataset that provides information about telecommunication events (SMSs, phone calls, Internet data transfer). The dataset is the output of data filtering over the Call Detail Records (CDRs) generated by an Italian MNO cellular network.

CDRs log the user activity for billing purposes and for network management. For the generation of this dataset, we have access to the CDRs related to the following activities:

- Sent SMS: A CDR is generated whenever a user sends an SMS.
- Incoming Call: A CDR is generated whenever a user receives a call.
- Outgoing Call: A CDR is generated whenever a user issues a call.
- Internet access: a CDR is generated whenever:

  - A user starts an internet connection.
  - A user ends an Internet connection.
  - During the same connection one of the following limits is reached:
    * 15 minutes from the last generated CDR.
    * 5 MB from the last generated CDR.

This dataset covers a large geographical area in Piedmont and Lombardy. The overall area is divided into 1114 rectangles with five different sizes.

For simplicity, they are called squares in the rest of the chapter. Within this area, we used the data for the city of Milan, which comprises 576 squares out of 1114 main squares.

Data are generated using two types of aggregation. With spatial aggregation, different activity measurements are provided for each square. With temporal aggregation, activity measurements are obtained by temporally aggregating CDRs in timeslots of 15 minutes.

The original dataset is composed of 5 files, one for each kind of traffic: sent SMSs, received SMSs, issued Calls, received Calls and Internet traffic.

The files referring to SMSs and phone calls have a very similar structure and comprise records containing:

- Time interval: The beginning of the time interval expressed as the number of milliseconds elapsed from the Unix Epoch on January 1st, 1970 at UTC. The end of the time interval can be obtained by adding 900000 milliseconds (15 minutes) to this value. Type: numeric.
- Square ID: ID of the square that is part of the grid. Type: numeric.
- Number of issued/received calls: The number of issued calls or the number of received calls. Type: Numeric.
- Country code: The phone country code of the person sending the call or the phone country code of the person receiving the call.

  Depending on the measured activity this value assumes different meanings that are explained later. Type: Numeric.
- Call Time: Total time of the calls in seconds. Type: Numeric.

The files referring to Internet access comprises records containing:

- Time interval
- Square ID

- Number of internet related CDRs: The number of CDRs generated because of internet activity. Type: Numeric.
- Country code: The phone country code of the person performing the internet connection. Type: Numeric.
- Downlink: Total data downloaded in kilobytes. Type: Numeric.
- Uplink: Total data uploaded in kilobytes. Type: Numeric.

Note that if for a given combination of the Square id, the Time interval and the Country code at least 3 events are not recorded, the record is missing from the dataset.

The second dataset provides the GPS data collected by Infoblu during March and April 2015. Infoblu is the Italian leader in the field of info-mobility. Infoblu has been a national company since 2001, and allows drivers to receive traffic reports directly on their navigator.

The info-mobility applications offered by this company provide information about the traffic conditions detected on the Italian highway network, main state roads and Metropolitan areas in Rome, Milan, Turin, Bologna, Florence, Naples, Genoa, Venice and Verona.

The data in this dataset includes geographical position data, speed, travel ID and time information from devices in vehicles.

These data are the necessary source for vehicle traffic information and for most intelligent transportation systems (ITS), so every vehicle with an active connected device acts like a sensor for the road network. Based on these data, traffic congestion can be identified, travel times can be calculated, and traffic reports can be rapidly recorded.

Each record in the file has the following structure:

- Travel ID: A unique string identifying a trip. A trip starts when the engine is switched on and lasts until the engine is switched off for at least 30 minutes. The length of the field is 32 characters.
- Timestamp: Timestamp indicating the exact time in which GPS was acquired. The time is in UTC time and its format is "yyyy-mm-dd hh:mm: ss". No time zone specification is used.
- Latitude: Latitude in WGS84 coordinates of the acquired GPS position.
- Longitude: Longitude in WGS84 coordinates of the acquired GPS position.
- Vehicle category. Kind of vehicle generating the GPS position. The value can be empty because for some vehicles the category is not available. The types of vehicles involved in the dataset are: Motor bike, car, van, truck, each with specific code, and there is also one code for the all other types of vehicles.

The third dataset is a shapefile that contains the whole area map separated in squares which are reported in UTM coordinates.

3.1.1 Data Preprocessing

The traces collected by the MNO for telecom traffic need to be preprocessed because of the existence of redundant and incomplete information in some traces and the necessity of aggregating the data related to each of the squares. The data was already aggregated temporally. The preprocessing includes three steps.

First the country code column was eliminated, then the timeslot column was eliminated since a separate data sheet was created for each timeslot, and finally the data reported in each sheet was aggregated for each square. The final obtained result of preprocessing the telecom data was the spatial and temporal distribution of cellular traffic.

For vehicular traffic, the data preprocessing consisted in ten steps, but we just summarize the most relevant ones.

The traces for different dates were separated, then the time column that was reported in real time was converted to Unix Epoch. The data aggregation in time was the next step. Since the geographical positions of trips are reported in latitude and longitude, they were converted to one geographical point.

3.2 Correlation Analysis

The analysis of correlation is necessary to evaluate the possibility of cellular network architectures where densification is obtained with MoBSs. This is an interesting solution for creation of on-demand capacity via temporary dense small cell deployments, where and when densification is needed. According to raw data analysis, vehicular traffic at penetration rates expected for small-cell-carrying vehicles is much more bursty than telecom traffic.

Note that the correlation between the number of vehicles and the number of MNO end user terminals, hence between the telecommunication traffic and vehicular traffic, was assumed in several works [**41**], [**42**], starting with the Real Time Rome project of the MIT senseable City Lab[1],[2], but has never been carefully measured in real-world settings.

Here we aim at filling this gap. We used measured data on two months of telecom and vehicular traffic in the city of Milan, Italy, and evaluated their spatio-temporal correlation.

[1]MIT. Available: http://senseable.mit.edu/,

[2]MIT. Available: http://senseable.mit.edu/realtimerome/,

3.2.1 Prepared Dataset

As mentioned before, we used a dataset that contains telecom and vehicular traffic of 61 days (44 working days and 17 weekend days) for the city of Milan, Italy, organized according to a subdivision of the area into 576 squares, as shown in Figure 3.1.

The squares (actually, rectangles) have minimum size of 220 m $\times$ 330 m, but sides can be 2, 4, 8 times bigger, as can be seen on the map. The time granularity of the dataset is 15 minutes (96 intervals per day). The total number of time intervals is $61 \times 96 = 5,856$. Considering the number of area squares for which data are reported separately, we have a total of 3,373,056 data points [**43**].

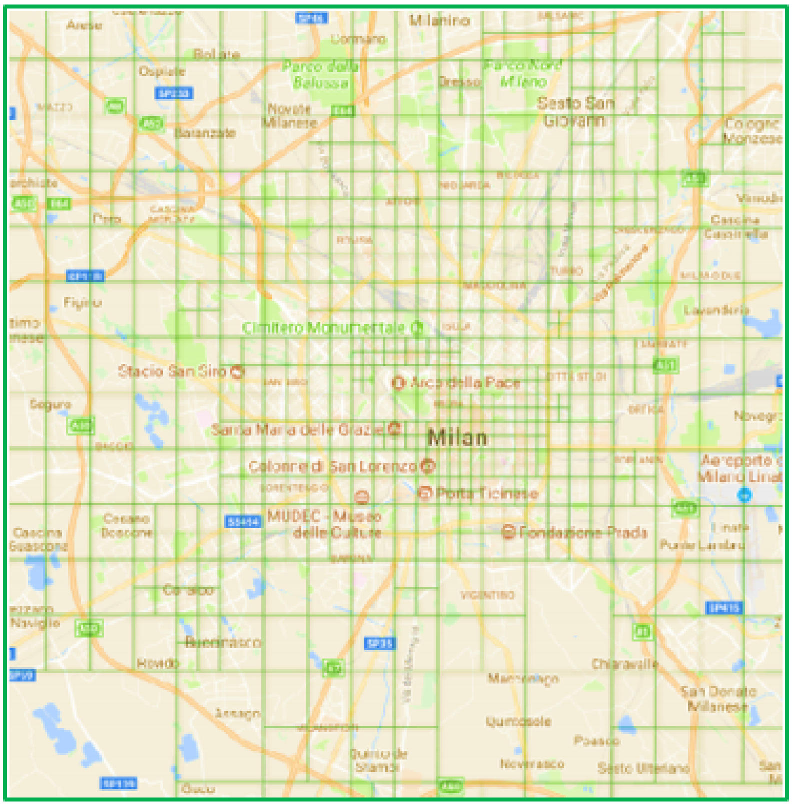


Figure 3.1: The subdivision of the Milan area into 576 squares.

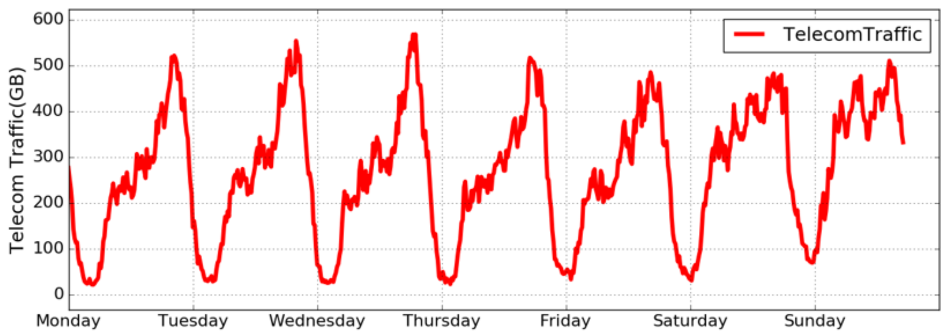
It is important to remark that the penetration rate of the two measurement technologies are different. The telecom traffic corresponds to more than one third of the overall mobile network demand, while the fraction of the monitored vehicles is less than 1%, as it mainly encompasses probe vehicles and commercial fleets monitored by InfoBlu.

We stress that these settings are consistent with the goal of our investigation: We consider a small percentage of vehicles to act as MoBS carriers, and would like them to cover the largest fraction possible of the total telecom traffic.

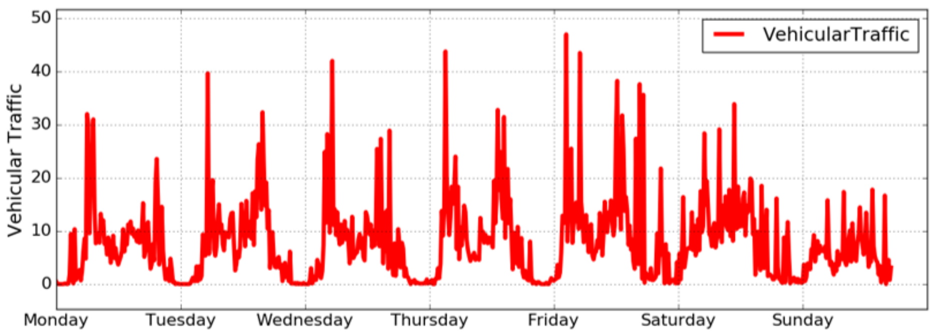
3.3 Data Visualization

Before diving into a deep analysis of mobile data traffic, we first visualize the spatio-temporal distribution of some sample squares. We consider squares of different types, based on several major land uses in these areas: a) residential, b) business, c) train station, d) airport, and e) soccer stadium.

Figure 3.2 shows the average telecom and vehicular traffic patterns for all days of the week for the residential selected square (telecom traffic in the top plot and vehicular traffic in the lower plot). We observe a number of interesting features in both telecom and vehicular daily traffic patterns.



(a)

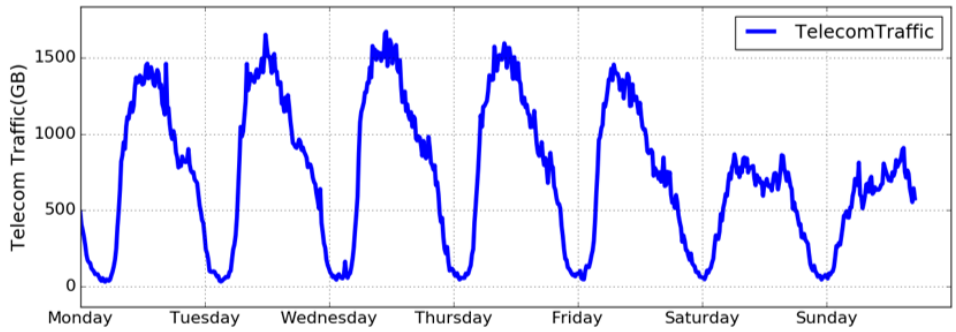


(b)

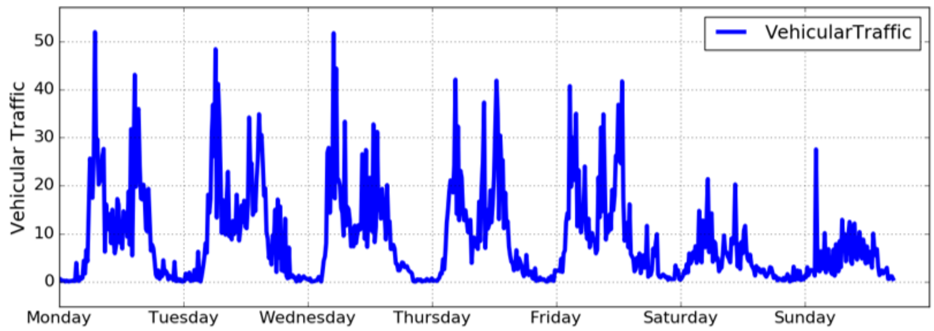
Figure 3.2: Average telecom and vehicular traffic (up and down, respectively) for all days of the week in selected square areas. Plots (a) and (b) refer to a residential area.

First of all, while in telecom traffic clear behaviors exist, the vehicular traffic is extremely irregular. In a residential area, plots (a) and (b) in Figure 3.2, we can clearly see that the telecom traffic is very low at night (roughly from 2 am to 7 am), it increases (sharply early in the morning, then more mildly) until 10 pm, and finally starts declining again.

No big differences exist among days, except for the fact that the traffic on Sundays is higher during the night and around midday. Vehicular traffic exhibits instead many ups and downs, with the highest peaks between 7 am and 10 am; also, vehicular traffic on Sundays is lower than in other days. The data embeds fundamental temporal patterns of mobile data traffic.



(a)

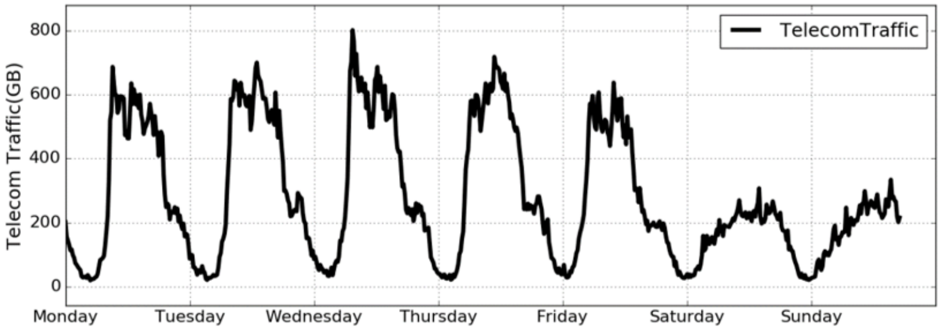


(b)

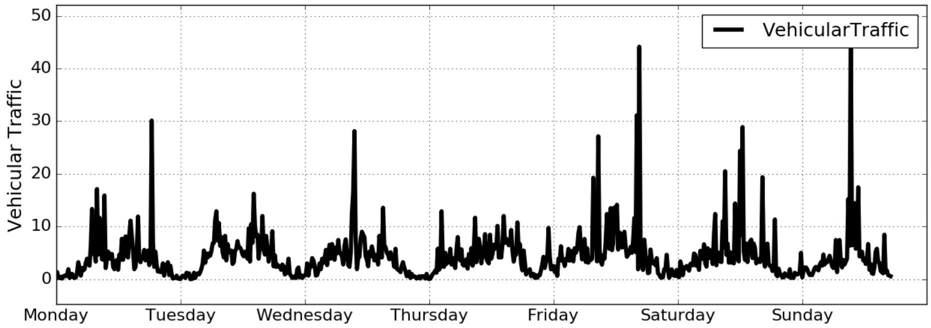
Figure 3.3: Average telecom and vehicular traffic (up and down, respectively) for all days of the week in selected square areas. Plots (a) and (b) refer to a business area.

In a business area, plots (a) and (b) in Figure 3.3, telecom traffic is very low at night (roughly from 2 am to 7 am), it increases sharply around 8 am in working days, and then gradually decreases after 2 pm. The peak of telecom traffic in weekends is about half of the one in working days. Vehicular traffic again exhibits high variability, with peaks at around 8 am and between 3 pm and 5 pm and is lower on Sundays.

In a train station area, plots (a) and (b) in Figure 3.4, telecom traffic is very low at night, then sharply increases around 8 am in working days, and remains close to its peak value from 9 am to 9 pm, then sharply decreases. The peak of telecom



(a)



(b)

Figure 3.4: Average telecom and vehicular traffic (up and down, respectively) for all days of the week in selected square areas. Plots (a) and (b) refer to a train station area.

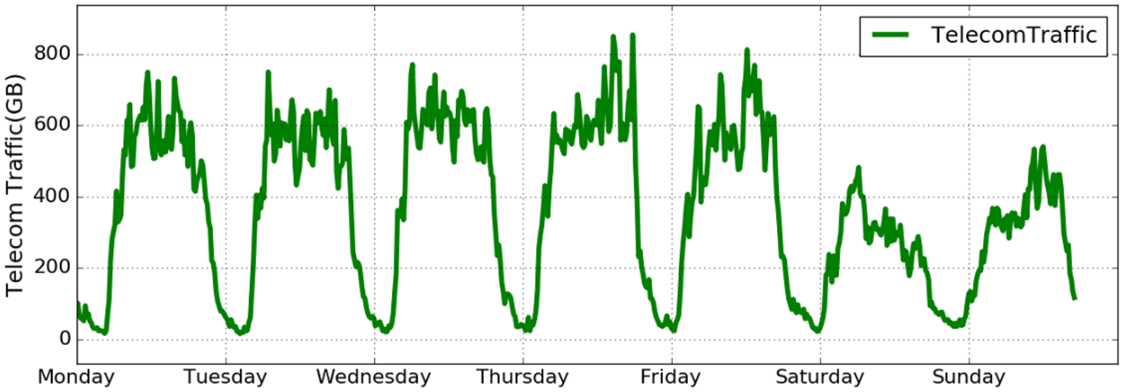
traffic in weekends is about one third of that in working days. Vehicular traffic is very bursty, as usual, peaking in the afternoon and evening.

In an airport area, plots (a) and (b) in Figure 3.5, telecom traffic is very low at night, sharply increases around 5 am, and remains close to its peak value from 8 am to 10 pm, then sharply decreases. The peak of telecom traffic in weekends is about two-thirds of the one in working days. Vehicular traffic as usual exhibits many ups and downs along most of the day.

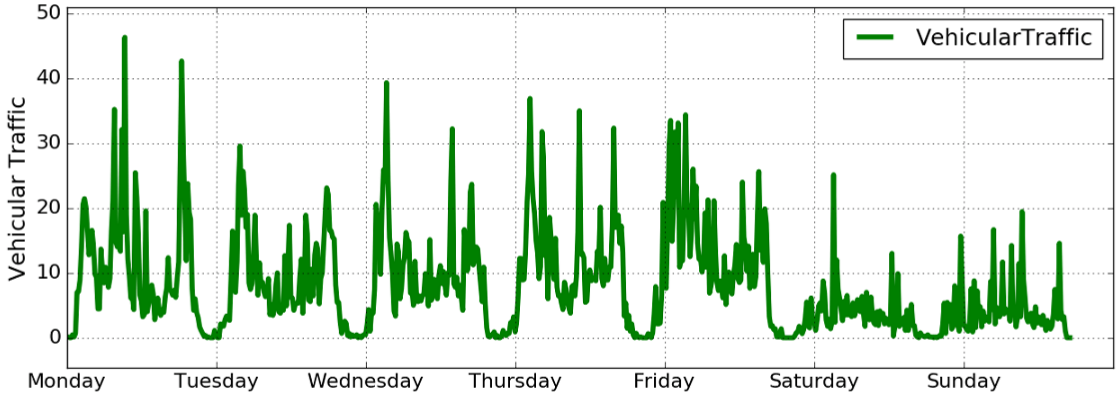
Finally, in a soccer stadium area, plots (a) and (b) in Figure 3.6, telecom traffic in plot(a) exhibits very high peaks around 9 pm on Saturdays and Sundays, when soccer matches are played (and concerts take place). Otherwise, traffic is close to zero.

Vehicular traffic in plot(b) is once more irregular, peaking at around 7 am and early in the afternoon.

The investigation of the raw data about telecom and vehicular traffic that we



(a)



(b)

Figure 3.5: Average telecom and vehicular traffic (up and down, respectively) for all days of the week in selected square areas. Plots (a) and (b) refer to an Airport area.

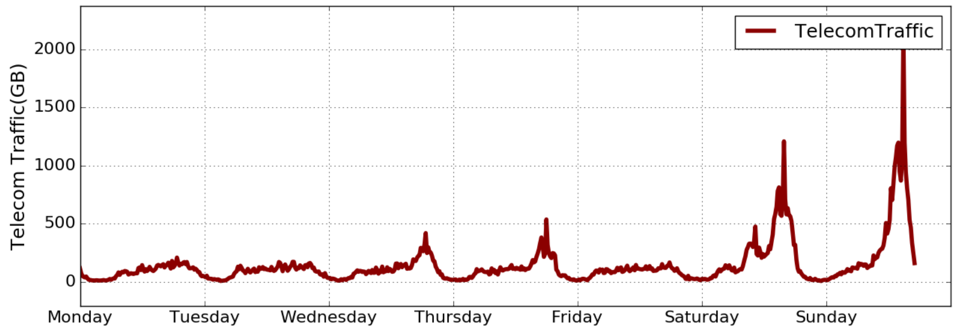
presented so far, does not provide enough clues to justify the existance of correlation; rather, it reveals that telecom traffic dynamics vary widely in space and time, and our data about vehicular traffic is much more bursty than our data about telecom traffic.

This latter effect is mainly due to the reduced fraction of vehicles carrying MoBSs that will characterize MOMONET deployments.

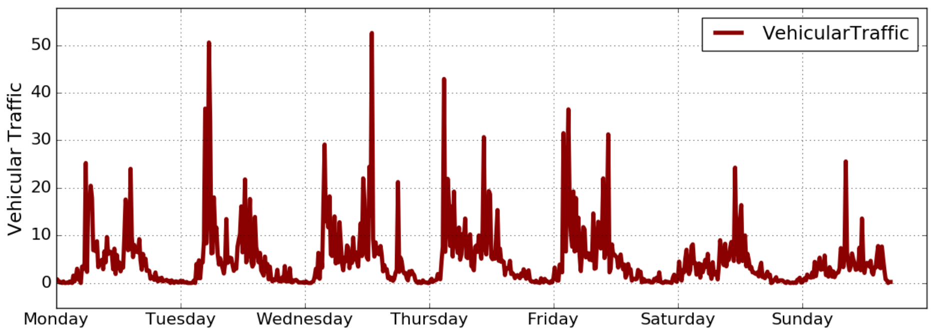
We provide a more global overview of the evolution of vehicular and telecom traffic in Figure 3.7.

There, the variation of traffic in space is visualized through heat maps, where for each square the color indicates the (telecom or vehicular) traffic level (dark blue is the lowest and dark red is the highest – the color scale is logarithmic).

The different plots portray heatmaps for one day (March 2, 2015) at five different times: 5 am, 10 am, 3 pm, 8 pm and midnight. We can see from these plots (and



(a)



(b)

Figure 3.6: Average telecom and vehicular traffic ((a) and (b), respectively) for all days of the week in selected square areas. Plots (a) and (b) refer to San Siro stadium area.

for similar ones at all time intervals in the day) that telecom traffic, in comparison to vehicular traffic, rises earlier in the morning, then it remains high later in the evening, and is more pervasive over the entire Milan area.

Even the visual presentation of the geographical distribution of vehicular and telecom traffic does not highlight evident correlations, except for the fact that traffic peaks coincide with the city center in both cases.

3.4 Different Types of Correlation

As the qualitative analysis in Section IV did not reveal evident relationships, we carry out a comprehensive study of the correlations between telecom and vehicular traffic. Specifically, we separately explore the spatial, temporal, and spatiotemporal dimensions of the problem.

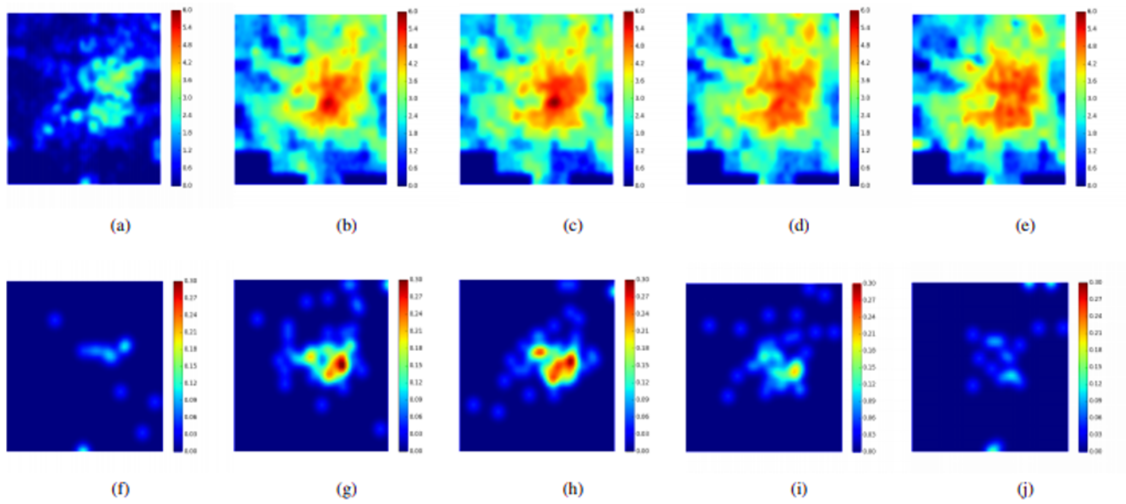


Figure 3.7: Heatmaps of Telecom (first row) and vehicular (second row) traffic on March 2, 2015. The first column shows data at 5 am; the second column at 10 am, the third at 3 pm, the fourth at 8 pm, the last at midnight. The plotted area is about 18 × 21 km.

3.4.1 Spatial Correlation

The results about correlations from a spatial perspective are summarized by the scatterplots in Figure 3.10.

There, the abscissa refers to the telecom traffic value, and the ordinate reports the vehicular traffic value for each dot corresponding to one square in the Milano area (thus 576 dots), and traffic values are averaged over all time intervals of working days (Monday through Friday; 44×96 instances - left plot) and weekends (Saturday and Sunday; 17×96 instances - right plot).

Therefore, Figure 3.10 shows correlations emerging across the geographical areas in the target metropolitan region. Plots also show the least square linear regression whose slope is 0.018 for working days and 0.016 for weekends.

The positive slope indicates that a positive correlation exists between telecom and vehicular traffic in space, despite the dispersion of points around the regression line.

3.4.2 Temporal Correlation

Correlations in the time dimension are considered in Figure 3.9, where each dot in the scatterplots is one daily time interval (for a total of 96 dots). Results are aggregated over the whole metropolitan area: the coordinates of each dot thus represent the average telecom and vehicular traffic in the time slot, averaged over 576 squares and either 44 working days in plot (a) or 17 weekend days in plot (b).

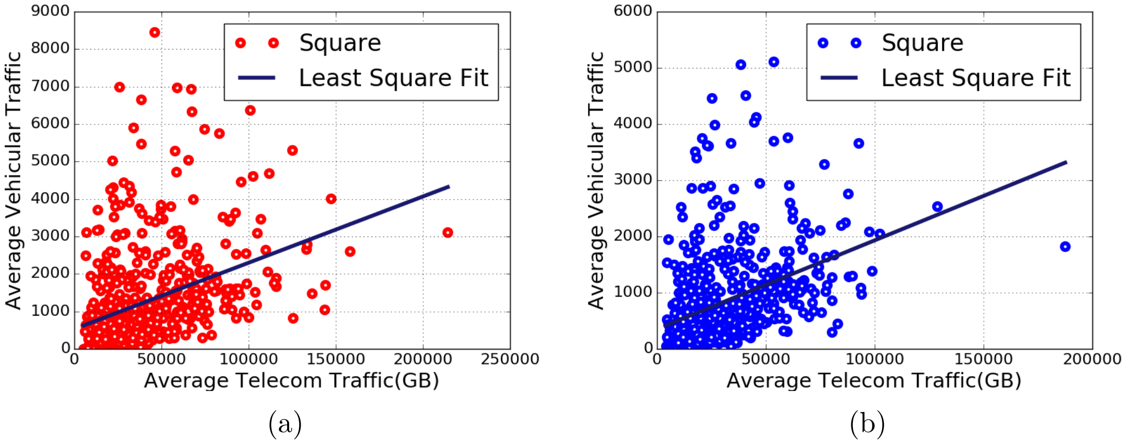


Figure 3.8: Correlation between average telecom and average vehicular traffic. Plot (a) refers to 44 working days and plot (b) to 17 weekend days.

The shape of these plots is quite interesting, if we look at them while considering the sequence of time slots. For this, plots (c) and (d) report the same data as plots (a) and (b), respectively, in a 3D plot, where the z-axis represents time of the day.

We see that a considerable amount of telecom traffic is present at 12: 15 am, while vehicular traffic is close to zero. Telecom traffic then declines, and both types of traffic are at their minimum around 3 am.

Vehicular traffic then starts to increase, while telecom traffic remains low (time runs clockwise along the dot sequences in Figure 3.9, plots (a) and (b)). Then, telecom traffic increases, and remains high until late in the evening, when vehicular traffic has become low. Finally, also telecom traffic dies out. Note that the effects are more pronounced during working days than in weekends.

Regression lines display a positive correlation also in this case, although points are typically far from a prediction by a simple linear model. The limited goodness of fit can be imputed to the temporal offset between the high activity periods of vehicular and telecom traffic explained above.

3.4.3 Spatio-temporal Correlation

We delve deeper in the analysis by disaggregating traffic over space and time. Figure 3.10 shows plots where each dot refers to one square in one 15-minute interval of the day (hence 576×96=55,296 dots in each plot). The coordinates of each dot are the average telecom and vehicular traffic recorded during either 44 working days or 17 weekend days.

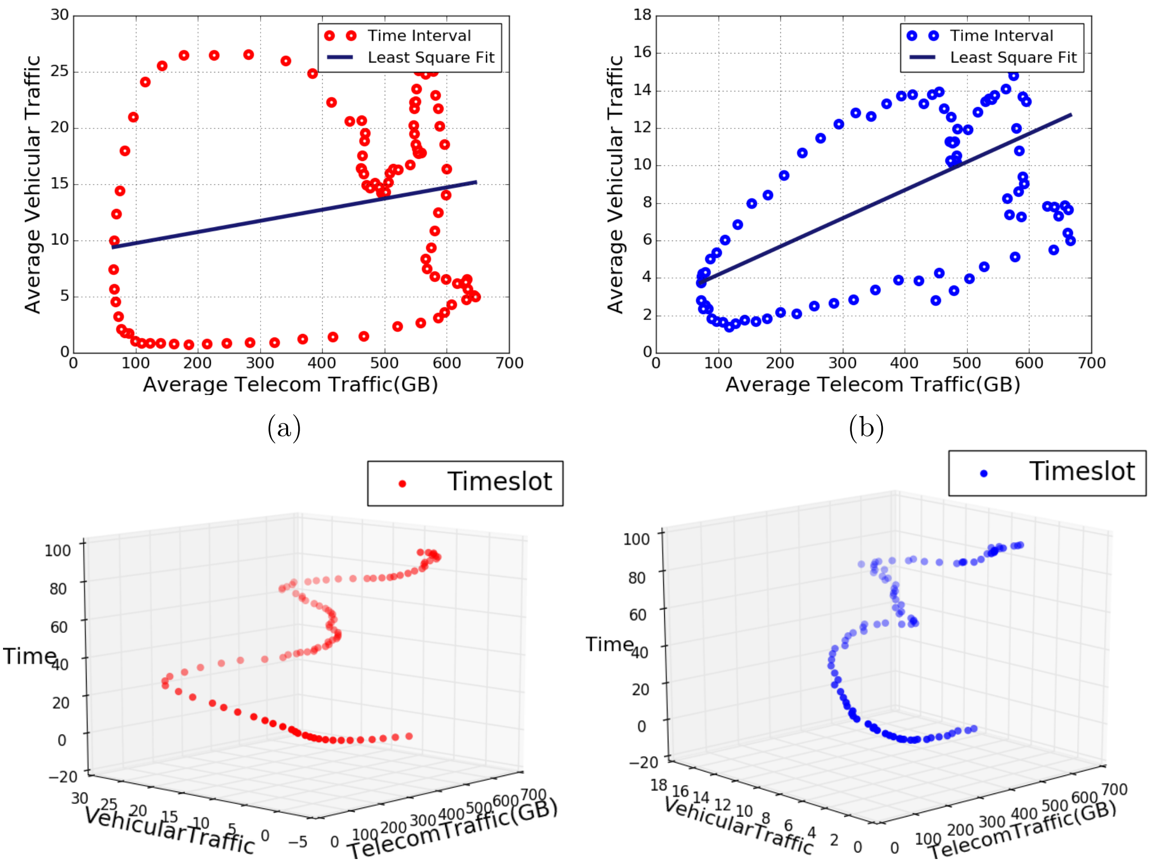


Figure 3.9: Correlation between average telecom and average vehicular traffic in each timeslot over 44 working days in plot (a) and over 17 weekend days in plot (b); plots (c) and (d) report the same data as (a) and (b), with time increasing along the z-axis.

3.4.4 When and Where is Correlation Higher?

The results above yield the following conclusions: i) there exist a general positive correlation between vehicular and telecom traffic, but ii) the correlation is low, due to a high variability of relations over time and space.

A sensible question is then whether some times and some areas show strong correlations and thus are suitable periods and locations for MNOs to take advantage of the MOMONET paradigm.

To answer this question, we leverage three different correlation coefficients:

- The Pearson correlation coefficient r, is defined as:

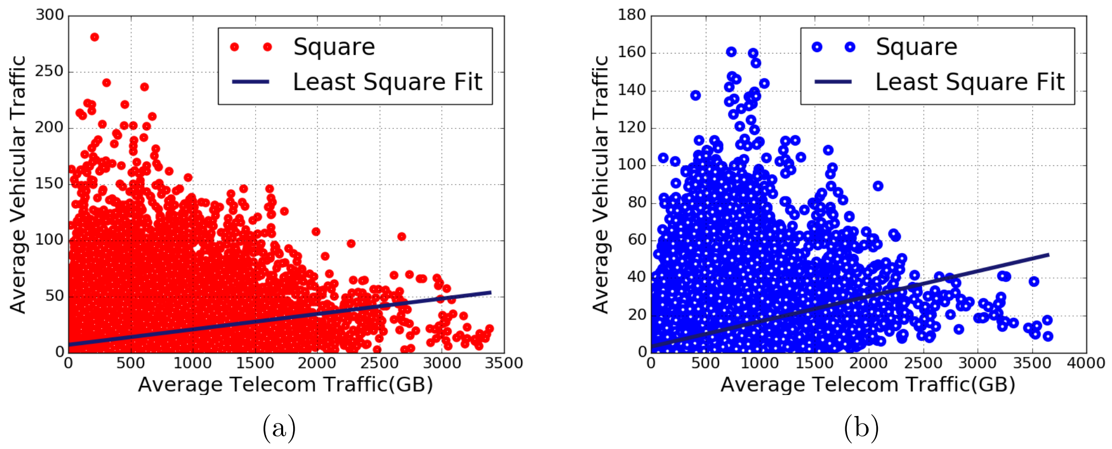


Figure 3.10: Correlation between average telecom and average vehicular traffic. Plot (a) refers to 44 working days and plot (b) to 17 weekend days.

$$r = \frac{n\sum_{i=1}^{n} x_i y_1 - \left(\sum_{i=1}^{n} x_i \sum_{j=1}^{n} y_j\right)}{\sqrt{\left[n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2\right]\left[n\sum_{i=1}^{n} y_i^2 - (\sum_{i=1}^{n} y_i)^2\right]}}, \tag{3.1}$$

where n is the number of observations for both telecom and vehicular traffic, and x_i, y_i are the telecom and vehicular traffic observations, respectively.

- The Spearman correlation coefficient s, is defined as:

$$s = 1 - \frac{6\sum_{i=1}^{n} y_i^2 d_i^2}{n(n^2-1)} \tag{3.2}$$

where d_i is the difference between the ranks (ordered positions) of corresponding telecom and vehicular traffic samples.

- The Kendall correlation coefficient q, is defined as:

$$q = \frac{n_c - n_d}{n(n-1)/2} \tag{3.3}$$

where n_c is the number of corresponding telecom and vehicular traffic samples with the same rank and n_d is the number of corresponding telecom and vehicular traffic samples with different rank. We first investigate which are the periods of the day that the correlation is higher.

Figure 3.11(a) shows the curves of the Pearson correlation coefficient for the average working day and for the average weekend day. The curves for the Spearman and Kendall correlation coefficients are not reported but are very similar.

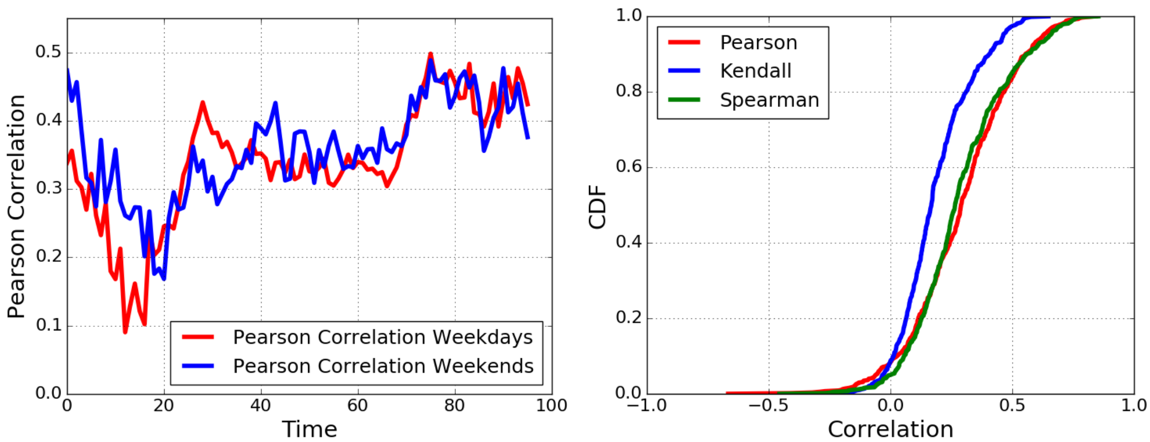
We see that the correlation is low at night in working days and in the early morning of weekends. On the contrary, the correlation coefficient values are high - in the range from 0.3 to 0.5 - in high-demand periods, i.e., in working and evening hours. These are good news, since the correlation is low in the night periods, when the demand for mobile services is typically low, and network densification is not necessary.

Instead, MOMONETs are the most effective exactly when densification is needed, that is, in presence of mobile network traffic demand peaks.

Considering the geographical distribution of correlation, Figure 3.11(b) depicts the cumulative distribution functions (CDFs) of the values of the three correlation coefficients for all spatial squares in the considered metropolitan area.

We can see that the negative correlations are few, and correlation coefficient values are positive in the (90–95) % of areas. Differences across coefficient types are also small, as the Pearson and Spearman correlation coefficients yield very similar CDFs, while the CDF of the Kendall correlation coefficient values indicates some prevalence of lower values of correlation coefficients.

However, the important observation is that the three CDFs confirm the existence of spatial correlation: depending on the considered square, the correlations vary significantly, and can reach values as high as 0.84.



(a) (b)

Figure 3.11: Pearson correlation coefficient between average telecom and vehicular traffic over all squares and over 44 working days and 17 weekend days in plot (a); cumulative distribution function of the correlation coefficients between average (over 61 days and over timeslots) telecom and vehicular traffic in each square in plot (b).

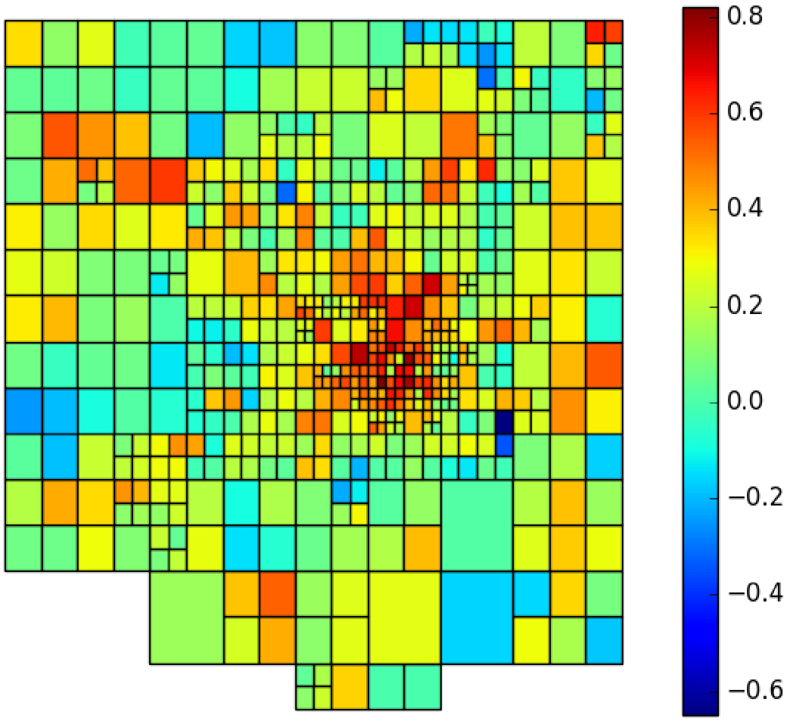


Figure 3.12: Pearson correlation coefficient heat-map between telecom and vehicular traffic for each Milano square, averaged over 61 days.

In order to understand where such high-correlation squares lie, Figure 3.12 shows a heatmap where the color of each square corresponds to the Pearson correlation coefficient computed over an average day. We can clearly see that correlation is stronger in the city center.

On the one hand, this is expected, since both vehicular and telecom traffic are higher in those squares, hence they follow more predictable patterns. On the other hand, and more importantly, those are also the squares where the demand for radio access network capacity is higher and densification is needed the most.

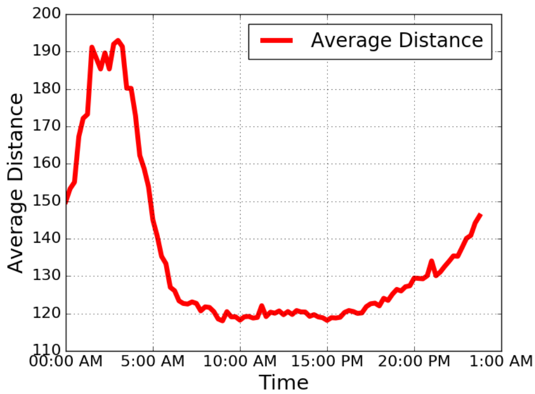
We conclude that, in our reference scenario, MOMONETs are exploitable not only when, but also where network densification is needed the most.

3.4.5 Average Distance

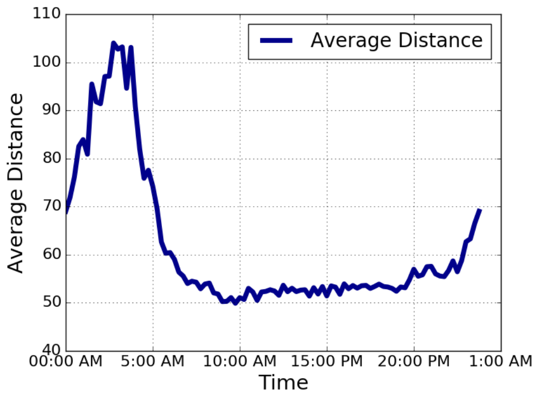
As a final complement of our analysis, we study the average distance between MoBSs in vehicles and the position of end user terminals that can use the MoBSs carried by vehicles to transmit and receive mobile network traffic.

Since our dataset reports exact vehicle positions within each square, but only the origin square (and not the exact end user terminal location) for telecom connections, we randomly position end user terminals within each square.

This is an approximation of end user positions within each area, but we do not have precise information about the exact positions of end users to drive a more realistic scenario. The average distance between the end user terminal and the closest MoBS (the one to which the end user terminal normally associates) is a metric of great interest for MOMONETs.



(a)



(b)

Figure 3.13: Average distance between vehicles and origins of telecom traffic versus time of the day averaged over 61 days and over all Milano squares in plot (a); average distance between vehicles and end users for all square areas over one day (March 2, 2015) in plot (b).

Results in Figure 3.13 (a), show the average distance over time, averaged over all spatial squares and averaged over all 61 days. The average distance is high (it can grow over 190 m) in periods of low telecom and vehicular traffic, and it stays at around 120 m during periods characterized by high network traffic, between 7 am and 6 pm roughly.

These distances risk to make connections to SCs frequently unviable, due to poor signal strength. Consider that the power emitted by a SC is typically less than 1 W, and connection depends on propagation conditions, but usually is less than 100 m.

The same observation holds when considering the geographical distribution of distances, averaged over the whole day. In Figure 3.14, we show for all 576 squares that cover Milan the average distance between the position of vehicles and the connection origin.

The shortest distances are found in the city center, around 120 m, while the longest average daily distances are observed in peripheral squares and grow up to around 200 m. Again, SCs are hardly usable under such distances.

However, when focusing on the city center, the outcome is different. In plot (b) in Figure 3.13 we plot the average distance between the position of end user

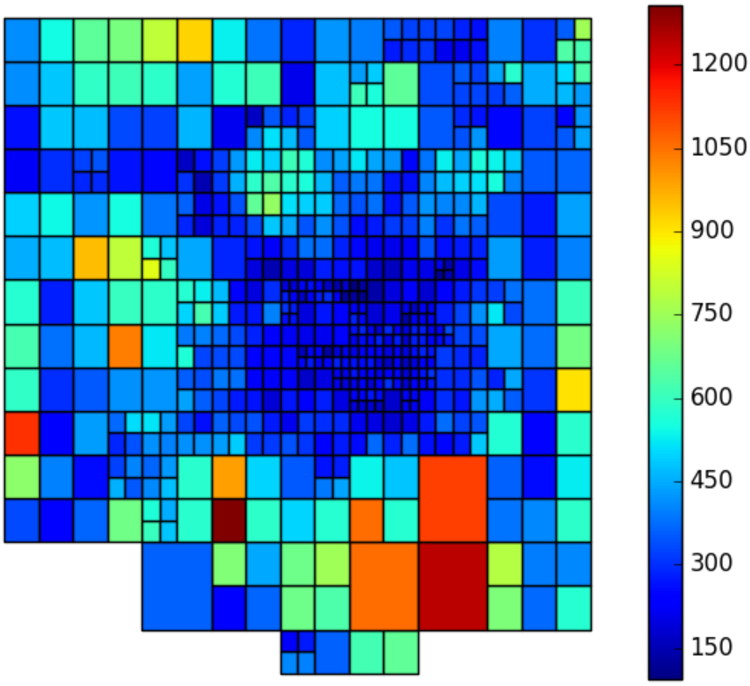


Figure 3.14: Average distance between vehicles and end users versus time of the day for 158 squares in the center of Milan averaged over 61 days.

terminals and the closest MoBS carried by a vehicle, in the 158 squares in the center of Milan averaged over all 61 days.

Between 9 am and 8 pm the distances are around 50 m, so in the city center we can expect a large fraction of terminals to be within reach of a MoBS carried by a vehicle.

These values nicely complement the results for correlation, and confirm that MOMONETs represent a suitable network paradigm to enable the adaptive densification of radio access envisioned for next-generation 5G systems, at least in the highly urbanized areas where mobile network capacity will be needed the most.

3.4.6 Conclusions

In this chapter, we used real traces of vehicular traffic and telecom traffic to study the correlation between these two types of traffic in time and in space over the city of Milano. The motivation behind this study lies in the assessment of the feasibility of the MOMONET concept, based on the densification of RANs in a metropolitan area using MoBS, i.e., small cell mobile base stations carried by vehicles.

For studying the correlation in space, we divided the area of Milano into 576 different squares and we studied the existing correlation between these two types of traffic in each square and also the total correlation in the whole metropolitan area during working days and weekends. The results show that a small positive correlation exist in the whole city, but also that correlation is stronger in the city center. Indeed, if we consider the squares within the city center (about 158 squares out of 576), we can see high values of correlation in some areas, up to 85%. This is a positive result, since densely populated areas in which both vehicular traffic and telecom traffic are high, are the best suited for the application of the MOMONET concept.

In addition, we also studied the temporal correlation during the day, and we found that correlation is higher during working hours and in the evening, when the two types of traffic are higher. Again, this is a positive result, since the adaptive densification offered by MOMONET is necessary only when traffic is high.

Finally, we characterized the average distance between the telecom traffic sources and MoBSs. For feasibility of MOMONET, we need to have shorter distances that can be covered by the low power signals emitted by MoBSs. The study of traces showed that distances between UEs and MoBSs are shorter during the times and in the places where correlation values are higher. In other words, distances are shorter in the city center during working hours, which are the periods with high correlation and high traffic values, when the MOMONET approach is useful and feasible.

Overall, the results in this chapter provide a promising basis for the viability of the MOMONET concept, and justify the investigation of the achievable throughput and fairness, as well as of the possible backhaul approaches, that we tackle in the following chapters of this dissertation.

Chapter 4

MOMONET: Preliminary Throughput Results

In the previous chapter, using real data referring to the city of Milano, we have proved the existence of correlation between the number of vehicles in a given area and the volume of wireless data transmitted in the same area, also showing that correlation is stronger in the most dense urban areas, where network densification is most necessary. In addition, we have also computed the expected distance between vehicles (considering only the tracked ones, hence reaching a pessimistic result) and end user terminals, showing that in city centers they are compatible with the transmission range of a small cell. These results are the prerequisites for the viability of the MOMONET concept.

Once those prerequisites have been verified, it is necessary to investigate the effectiveness of MoBSs in increasing the RAN capacity. With this objective, in this chapter we study the capacity, i.e., the maximum achievable throughput of RANs that exploit MoBSs for adaptive densification in urban areas.

While traditional approaches for radio access network densification with fixed small cell base stations are proving ineffective and extremely costly, MoBSs can provide adaptive densification while achieving similar efficiency and lower cost. Indeed, the existence of correlations between the number of mobile network subscribers and the number of vehicles in a given area that we observed in the previous chapter allows for the spontaneous creation of temporary dense small cell deployments where and when needed.

Ultimately, this Radio Access Network densification method increases efficiency and reduces costs for the operator.

In this chapter, we first present an approach for the computation of the maximum throughput obtained by fixed base stations and mobile small cell base stations. Then, we provide initial estimates for the throughput improvements with respect to traditional deployments that only rely on fixed base stations.

Evaluations in the realistic case study of the main railway station area in Milan, Italy, reveal that the use of mobile base stations improves throughout gains up to 120 % over fixed infrastructures, while granting higher fairness among subscribers.

The computation of throughput comprises a few steps. To tackle the major problem of interference, we first present an approach for the computation of the maximum downlink throughput achievable in an area served by traditional fixed BSs and MoBSs. Then, we develop an easy-to-deploy optimization that selects active MoBSs to maximize the overall throughput or user fairness. Finally, we provide initial estimates for the throughput and fairness improvements provided by MoBSs.

Our results show that, thanks to the positive spatial correlations between mobile network demands and road traffic, mobile small cell base stations carried by vehicles guarantee equivalent or better performance in comparison with a traditional deployment of fixed small cells with significantly lower cost.

The BSs, whether fixed or vehicle-mounted, are connected to the core network through a front/backhaul link, that is either wired or wireless for fixed BSs but must obviously be wireless for MoBSs. We assume that the connection between a MoBS and the fixed network is implemented with a millimeter wave link, so as to have enough capacity for the support of the traffic generated by all end users connected to the MoBS.

Furthermore, we assume that such broadband links avoid interference with the lower frequency channels connecting either fixed BSs or MoBSs to UEs.

4.1 Interference Issues

The random, variable layout of MoBSs is very different from the carefully planned layouts of traditional cellular networks; this poses very challenging questions on the management of the interference among BSs (vehicular and fixed, unless they operate on separate frequency bands) and UEs, both in the downlink and in the uplink directions.

The approach used by todays RANs tries to reduce interference as much as possible, by carefully controlling transmission power and end user associations to BSs. More innovative approaches are possible, such as the exploitation of the many MoBSs for cooperative transmission approaches derived from MIMO or CoMP[1], [**44**].

These approaches are for example being studied in the context of dense small cell RANs by ARTEMIS, with their pCell (personal cell) technology [**45**], [**46**].

A different innovative approach can envision the geo-coordination of MoBSs, so that they collectively “look like” a single BS for terminals, possibly exploiting CoMP [**47**], with the advantage of also reducing the handover rate. Interestingly,

[1] J. S. Seybold. Introduction to RF propagation. J. Wiley Sons, 2005

also D2D (device-to-device) communications, or MoBS-to-MoBS communications could play a crucial role in reducing the access network interference.

Since MoBSs can move, they introduce inter-cell interference levels that change over time. Therefore, they exacerbate the complexity of inter-cell interference coordination mechanisms like ABS (almost blank subframes). Indeed, a static ABS approach is not suitable, and advanced stochastic ABS configurations can be leveraged instead, as the dynamic distributed scheme proposed in [**48**].

4.2 Uplink connections

The downlink performance is simplified by the lack of contention accessing the wireless resources. When the uplink is considered, the signaling between UEs and MoBSs is considered to acquire access to radio resources.

Part of the standard procedures incorporate random access algorithms, which can become critical with very large user populations. This can be a further advantage of the MOMONET approach, where MoBSs collect requests of medium size groups of users, alleviating contention on fixed BSs.

4.3 Throughput Calculation and Optimization

In this section, we summarize the results presented in[2] to evaluate the performance of the systems including BSs that can be muted on a millisecond timescale or relay groups. These results are applied to the MOMONET case to evaluate analytically the maximum downlink throughput performance achieved when MoBSs are in place.

This capacity characterization allows the assessment of the gain enabled by MoBSs, and the formalization of an optimization problem that copes with the classical issue of interference observed in dense scenarios. The result unveils how to select, at each given time, which MoBS should be allowed to transmit and which ones should be muted instead, so to control interference.

Our proposed approach just mutes MoBSs, not fixed BSs, but it can also handle the muting of fixed BS.

We start by computing the average number of bits per symbol transmitted to a specific UE by the serving BS, i.e., the so-called transmission efficiency. Afterwards, we compute the downlink throughput obtained by a UE and, based on this result, we propose an easy-to-deploy optimization that selects active MoBSs to maximize overall throughput or user fairness.

[2] C. Vitale, V. Sciancalepore, V. Mancuso, "Fair Stochastic Interference Orchestration with Cellular Throughput Boosted via Outband Sidelinks," arXiv: http://arxiv.org/abs/1809.09524

4.3.1 Transmission Efficiency

We consider a scenario with $\mathcal{B}$ set of interfering base stations. A base station that belongs to $\mathcal{B}$ is either a standard fixed base station or a MoBS. In the following, we consider short time slots in which the location of UEs and MoBSs is assumed as fixed. Furthermore, we suppose that UEs attach to base stations either fixed or MoBSs based on the strongest received signal.

Transmission efficiency, though, not only depends on the location of BSs and UEs, but also on the mapping between Signal to Interference and Noise Ratio SINR, and Modulation Coding Schemes MCSs (we refer the reader to[3] for further details on MCS mapping examples). Considering time slot t, where UEs and MoBSs locations can be considered as fixed, the transmission efficiency $\zeta_i(t)$ of UE i can be then computed as:

$$\zeta_i(t) = \sum_{k \in \mathcal{M}} b_k \left[F^t_{\mathrm{SINR}}(T^{\max}_k) - F^t_{\mathrm{SINR}}(T^{\min}_k) \right], \tag{4.1}$$

where $\mathcal{M}$ is the set of MCSs, b_k is the number of bits per symbol for MCS k, $(T^{\min}_k, T^{\min}_k]$ is the interval of the SINR for MCS k, and F^t_{SINR} is the Cumulative Density Function (CDF) of the SINR at time t. In practice, Eq. 4.1 evaluates the probability of UE i to use a specific MCS in time slot t, given the experienced SINR distribution.

The SINR CDF depends on the radio propagation between the BSs and the UEs. As pointed out in [**49**], in urban environments, UEs are most likely to experience Rayleigh fading.

For this reason, we assume that the power received by a UE both from the attached and the interfering BSs follows a negative exponential distribution, i.e., the received instantaneous signals by the UE follow a Rayleigh distribution. As a result, we obtain the following proposition.

Proposition 1. The CDF $F_{\mathrm{SINR}}(x)$, resulting from an exponential useful signal with average power $1/\lambda_S$, J independently exponentially distributed interfering signals with average power $1/\lambda_{I_j}$ and constant noise power N, is, $\forall x \geq 0$:

$\mathrm{F}_{\mathrm{SINR}}(x) = 1 - e^{-\lambda_S N x} \prod_{j=1}^{k} \frac{\lambda_{I_j}}{\lambda_{I_j} + x\lambda_S}$.(4.2)

[3]Third Generation Partnership Project (3GPP), 3GPP TS 36.423 v. 14.0.0: Evolved Universal Terrestrial Radio Access Network (E-UTRAN); X2 application protocol (X2AP)

Proof. The proof can be easily obtained from the following expression:

$$F_{\text{SINR}}(x) = \Pr\left\{\frac{S}{N + \sum_{j=1}^{k} I_j} \le x\right\} =$$
$$= \int_0^\infty \int_0^\infty \ldots \int_0^\infty \Pr\left\{S \le x \left(N + \sum_{j=1}^{J} I_j \middle| I_j = a_j\right)\right\}$$
$$\cdot \prod_{j=1}^{k} f_{I_j}(a_j)\; da_1 da_2 \ldots da_J$$

□

The average received power levels can be computed with standard distance-based path loss models.[4]

4.3.2 User Throughput

In the previous section, we computed the number of bits transmitted on average by i at time slot t, i.e., $\zeta_i(t)$. Now we can compute the average user throughput $\Gamma_i(t)$ multiplying $\zeta_i(t)$ by the average symbols per second D_i available for i at the attached base station.

$$\Gamma_i(t) = D_i \zeta_i(t). \tag{4.3}$$

D_i mainly depends on the base station scheduler. We assume that the Equal Time Scheduler (ETS) is applied so each UE receives on average the same amount of symbols, which yields:

$$\Gamma_i(t) = \frac{K}{N_b | b \text{ covers } i} \zeta_i(t), \tag{4.4}$$

where K is the number of symbols per second available for data transmission, and N_b is the number of UEs attached to the base station serving i.

4.3.3 Throughput Optimization

Muting and reactivating transmissions at a BS is currently possible in millisecond timescale (see the Almost Blank Sub-frame (ABS) tool, as an example[5]) without any handover for end users. So by changing subsets of $B \subseteq \mathcal{B}$ to transmit,

[4] J. S. Seybold. Introduction to RF propagation. J. Wiley Sons, 2005.

[5] Third Generation Partnership Project (3GPP), 3GPP TS 36.423 v.14.0.0:Evolved Universal Terrestrial Radio Access Network (E-UTRAN); X2 application protocol (X2AP).

we can control the interference in the system and the SINR distributions without having to continuously deal with the BS attachment algorithms.

Therefore, selecting the right subsets of B of $\mathcal{B}$ allowed to transmit, and the frequency at which subsets are muted, we can optimize the average user fairness or throughput in a way that is transparent to UEs. We assign to each subset B of $\mathcal{B}$ a portion P_B of transmission resources, where only BSs and MoBSs in B are allowed to transmit and other MoBSs are muted.

For instance, group of MoBSs can be scheduled sequentially, so that B can be active for a fraction P_B of the system time. The throughput of each user i when B is active, i.e., $\Gamma_i^B(t)$, can be easily obtained by considering the base stations in B as the only interfering base stations:

$$\Gamma_i(t) = \sum_{B \subseteq \mathcal{B}} P_B \Gamma_i^B(t). \tag{4.5}$$

Obviously, if the MoBS b to which i is attached is not included in B, then $\Gamma_i^B(t) = 0$. To maximize the average user fairness, as well as the overall system throughput, it is then sufficient to optimize P_B. Specifically, we propose a convex optimization to maximize proportional fairness, namely, the proportional fair muting `PFM` problem. Analogous optimization problems can be easily obtained for other fairness metrics.

Problem `PFM`:
At time t, with N_t UEs in the area, we have $P_B, \forall\ B \in \mathcal{B}$, so to:

$$\begin{aligned} \text{maximize} \quad & \sum_i \log\left(\sum_{B \subseteq \mathcal{B}} P_B \Gamma_i^B(t)\right); \\ \text{subject to:} \quad & \sum_{B \subseteq \mathcal{B}} P_B = 1, \\ & P_B \in [0,1], \quad \forall\ B \in \mathcal{B}; \end{aligned}$$

Muting patterns are generally fixed, and can be updated every second (roughly). Problem `PFM` can be therefore computed on the same time scale, so as to update $\Gamma_i^B(t)$ according to the positions of MoBSs and users.

4.4 Numerical Results

In this section, we describe the settings of our preliminary analysis of the throughput attained in MOMONET systems, and then we present the corresponding numerical results.

4.4.1 Case Study

Our initial throughput evaluation focuses on one representative case study, i.e., the geographical area shown in Figure 4.1 including Milan central railway station. In Figure 4.2, the map area covers nine squares of the main map, the whole area covered by the nine squares is distiguished by the red border. For each of those squares, we have data about the mobile traffic and the number of probe vehicles, at 15-minute time intervals and over several days in April 2015.

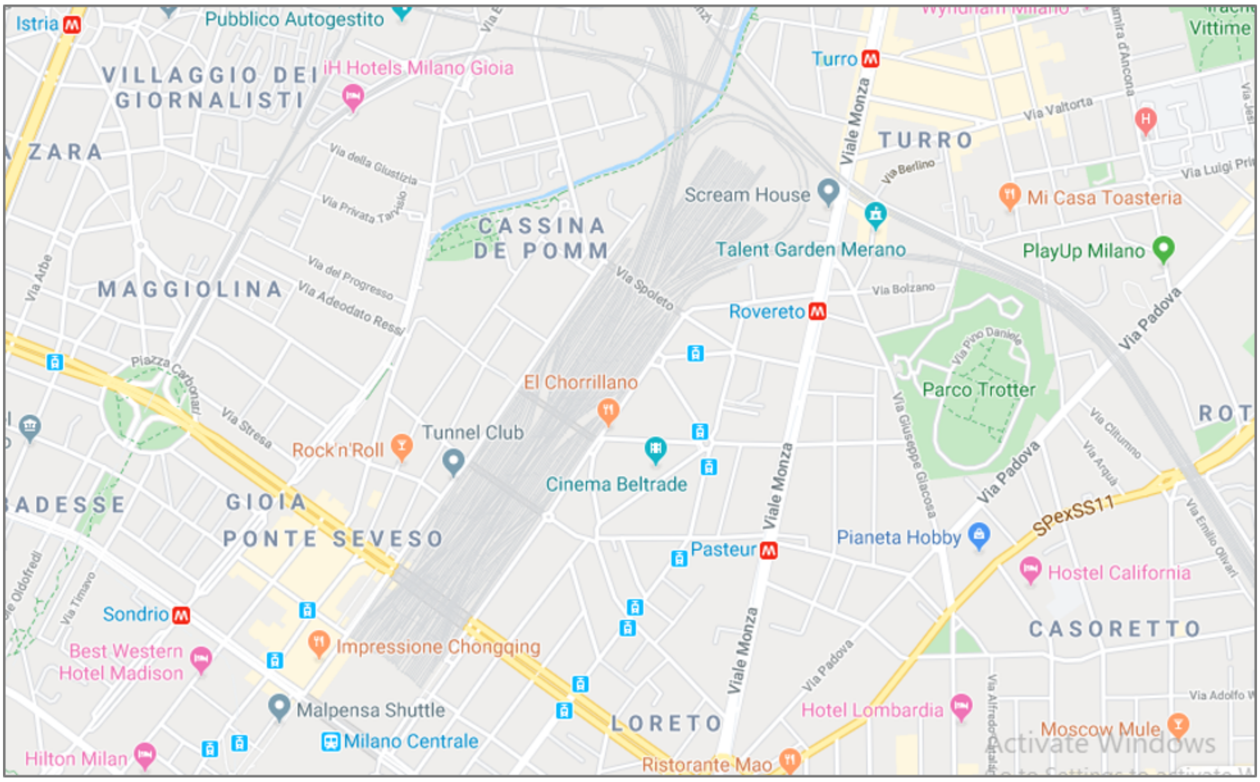
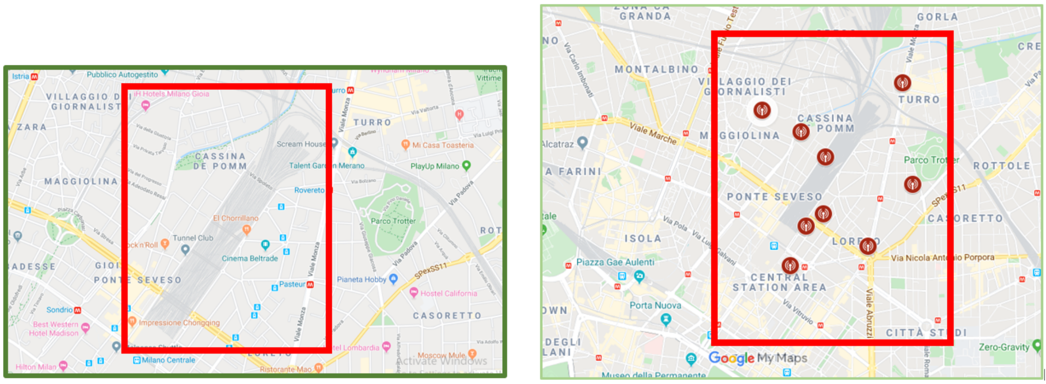


Figure 4.1: The area of the central railway station in Milan.

The mobile network traffic refers to Internet connections and voice calls of one of the largest Italian MNOs and corresponds to a large fraction of the total data traffic. For the same MNO, we have the positions of fixed BSs.

On the contrary, tracked vehicles are less than 1 % of those vehicles in the area. We assume that each of the tracked vehicles is equipped with a MoBS, which is coherent with the assumed penetration rate of mobile small cells in the vehicle population. For example, in the time interval from 8 A.M. to 8:15 A.M., April 15, 2015, the reported vehicle positions are as shown in plot(a) in Figure 4.4 as blue/green markers, together with the positions of the area of the central railway station in Milan of fixed BSs as red circles. We clearly see that fixed BSs are nicely spaced, while vehicles cluster along the main roads surrounding the railway station, and in some cases are very close to fixed BSs, risking to generate excessive interference with MoBSs.



(a) (b)

Figure 4.2: Milano train station area. Plot (a) refers to Milano rail way station including the 9 squares of Milano identified overall by the red square. Plot(b) refers to Milano railway station and fixed macro base stations.

By using the right MoBSs we can achieve reasonable interference (this is a responsibility of the Orchestrator in Figure 2.2). We choose the set of useful MoBSs that correspond to blue triangles in plot(a) in Figure 4.4, with a useful heuristic algorithm.

Recalling that useful BSs are always scheduled by the `PFM` optimization (otherwise, attached users experience no throughput), our algorithm eliminates some available MoBSs in order to reduce the overall interference.

Assuming that the transmission power of fixed BSs is 30 dBm, and the one of MoBSs is 20 dBm, for each pair of MoBSs (or each pair comprising one MoBS and one fixed BS) we compute the power of the average received signal at both ends following the path loss model in[6], and we eliminate the MoBSs with the highest value of generated interference.

Since the channel model is symmetric, when the highest value of generated interference is due to a pair of MoBSs, so that both generate the highest interference value, we eliminate the MoBS with the highest second value of pairwise generated interference. Removed MoBSs are denoted by green squares in plot(a) of Figure 4.4.

Considering UEs, since for each rectangle in the map of plot(a) in Figure 4.3 we know the total data traffic volume in the considered time slot, but we have no information about UE positions, we randomly place UEs in the rectangles, assuming that one UE is present for each 1GB of reported traffic.

[6]J.S. Seybold. Introduction to RF propagation. J. Wiley Sons, 2005.

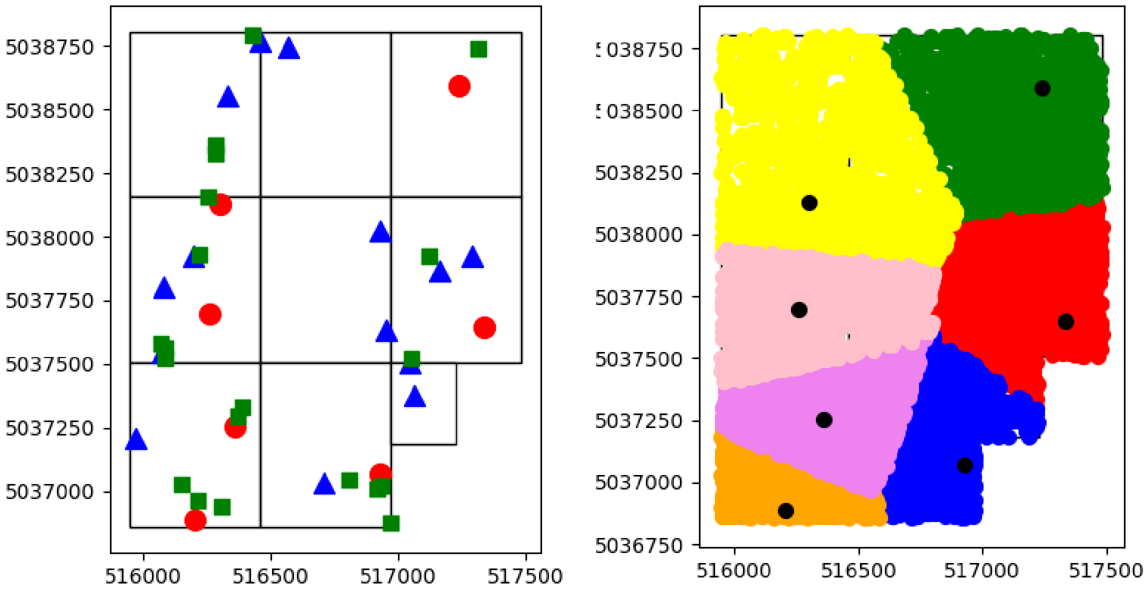


Figure 4.3: The area of the central railway station in Milan. Plot(a) shows the area with fixed BSs and vehicle positions at 8 A.M. on April 15, 2015. Plot(b) depicts UE associations to fixed BSs (i.e., without MOMONET).

UE associations to fixed BSs or MoBSs follow a maximum received power criteria. Using the transmission powers mentioned above, the resulting UE associations are as reported in plot(b) of Figure 4.3 in the case of fixed BSs only, and in Figure 4.4 is the case when both fixed BSs and MoBSs are active.

We compute the maximum achievable throughput for the interval between 5 A.M. and 10 P.M. on April 15, 2015, at one-hour spacing. The total number of cars, the number of useful (i.e., preselected) MoBSs in the area and the total data traffic for each time slot are reported in Figure 4.5.

4.5 Results

By applying the procedure outlined in the throughput calculation section, it is possible to compute for each time slot t the maximum downlink throughput as well as the throughput corresponding to maximum fairness.

The procedure first computes user throughput under each configuration of the system for each combination of active/inactive (i.e., transmitting/silent) MoBSs. If the number of useful MoBSs in the area is equal to n, the number of configurations is 2^n.

Then, by summing up the throughput of all users in a given configuration, the

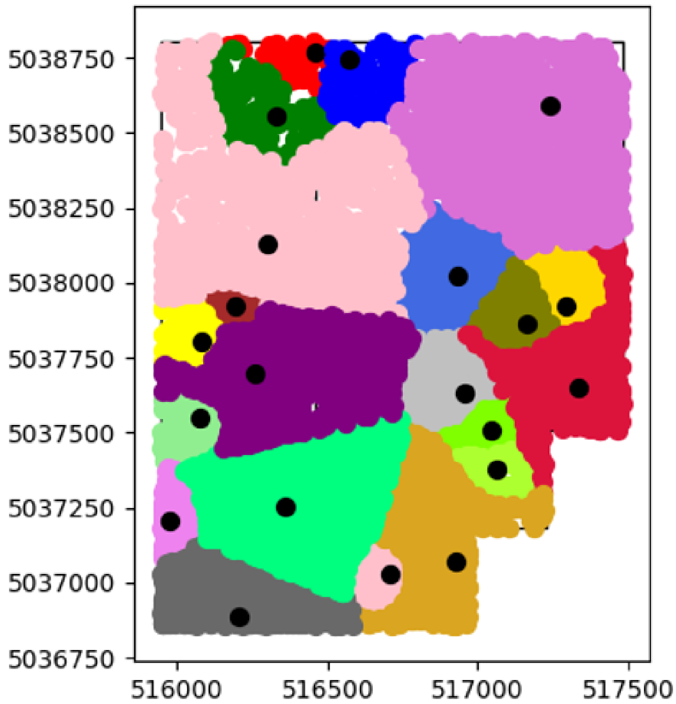


Figure 4.4: The area of the central railway station in Milan. The plot depicts the UE associations to fixed BSs and MoBSs.

total throughput of each configuration is obtained; the maximum throughput value among the throughput of all configurations is the maximum achievable throughput in time slot t.

As an example, Figure 4.6 reports the total throughput of all configurations in the considered area at 8 A.M. on April 15, ordering configurations so that all cases with one active MoBS appear first, then all cases with two active MoBSs, and so on, until the configuration with all active MoBSs is reached.

Different colors identify different numbers of active MoBSs. Configurations are ordered according to their binary representation (0 means inactive and 1 means active; the all 0 configuration mutes all MoBSs, while the all 1 configuration has all MoBSs active).

Since the number of possibly active MoBSs is 16, the total number of configurations is 65536. The red dashed horizontal line refers to the case of only fixed BSs, with users attached to fixed BSs only. In this case, the maximum throughput is reached when all 16 MoBSs are active, i.e., at the rightmost point in the graph.

However, this is not always the case: In some time slots, maximum throughput configuration does not include all MoBS. For instance, at 7 A.M. and 9 A.M. and at 5 and 7 P.M. one MoBS is excluded, and at 10 A.M. and 11 A.M. and at 1 P.M. two MoBSs are excluded.

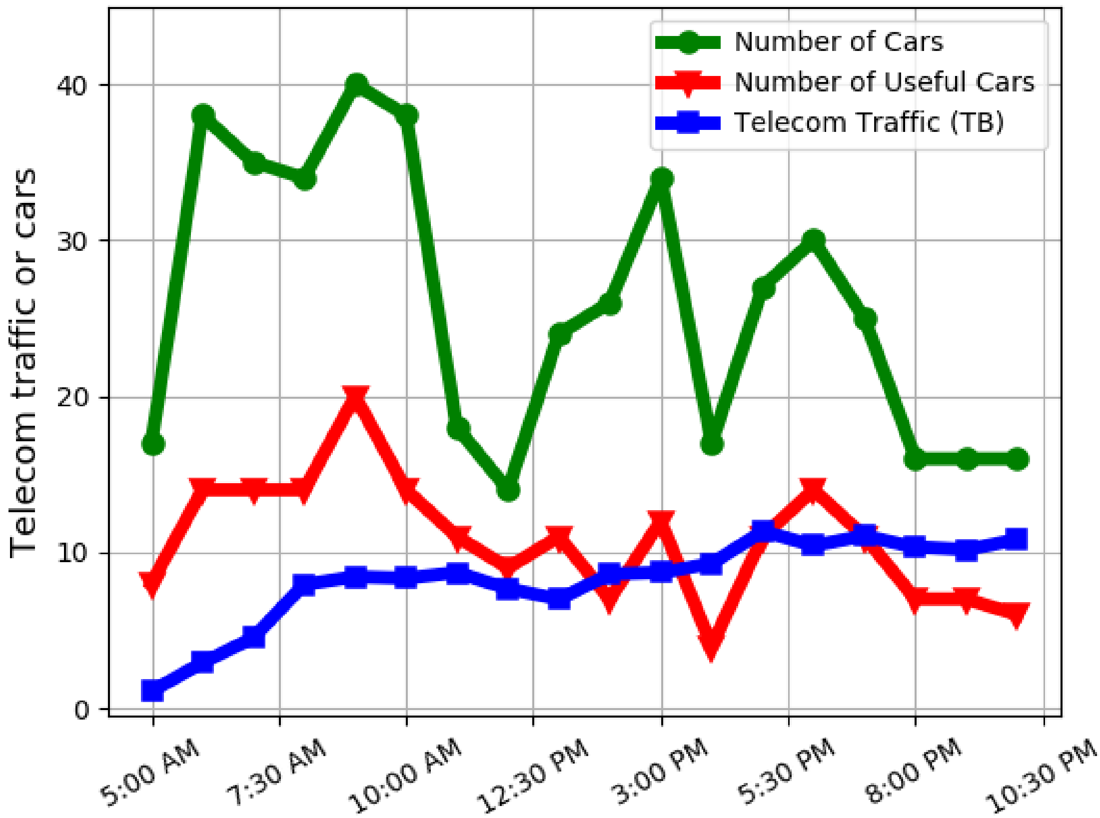


Figure 4.5: Number of vehicles, number of MoBSs and data traffic (in TB) in the area of the central railway station in Milan on April 15 from 5 A.M. to 10 P.M.

Finally, by applying a time schedule that alternates over a set of configurations, the optimized throughput is obtained by exploiting the approach of the `PFM` problem described above.

In Figure 4.7, we report for each time slot of April, 15, three throughput values: the throughput with only fixed BSs (and all UEs associated to just fixed BSs), the throughput of the MoBSs configuration yielding maximum throughput, and the throughput of the time schedule that optimizes fairness.

Of course, the throughput of the MoBSs configuration yielding maximum throughput is the highest, and the throughput with only fixed BSs is the lowest. It is remarkable to see that by using MoBSs we can achieve gains of about 150% if fairness is not an issue, and gains of almost 120% when fairness is optimized.

Figure 4.8 reports the fairness values in the same three cases. We can see that the fairness achieved when MoBSs are present is always higher than in the case of only fixed BSs.

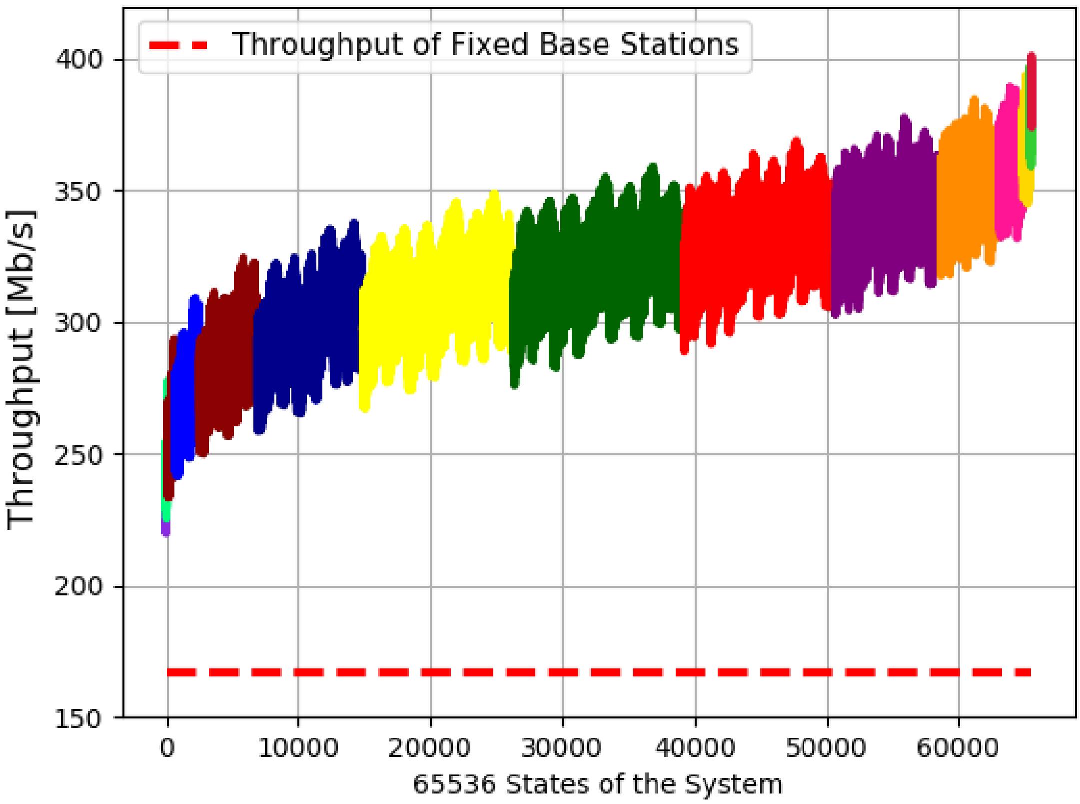


Figure 4.6: Throughput of the 2^{16} MoBS configurations in the area of the central railway station in Milan on April 15 at 8 A.M.; configurations are in order of number of active MoBSs (colors identify the number of active MoBSs; for equal number of active MoBSs the order follows the binary representation). The red dashed horizontal line refers to the case of only fixed BSs, with users attached to fixed BSs only.

The fairness values of the MoBSs configurations providing maximum throughput in some cases are not reported, because when the max throughput state (which must be used with a resource share factor $P_B = 1$ to achieve the maximum throughput configuration) is such that not all useful MoBSs are active, some users receive zero throughput all the time, so that fairness takes the value of $-\infty$.

This proves that the use of the time scheduler is extremely important in order not to exclude some UEs to access to network resources.

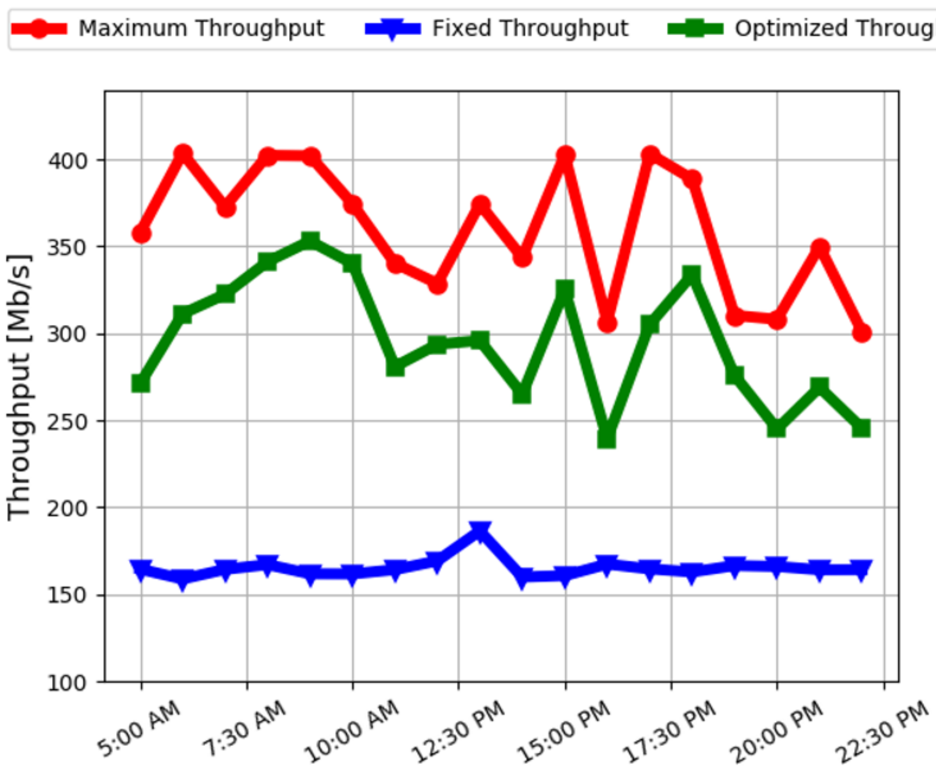


Figure 4.7: Throughput with only fixed BSs, throughput of the MoBSs configuration yielding maximum throughput and throughput of the time schedule that optimizes fairness (in Mb/s) in the area of the central railway station in Milan on April 15 from 5 A.M. to 10 P.M.

4.6 Heuristic Solution of the `PFM` Problem

Problem `PFM` can be solved with standard optimization tools, provided the number of configurations B of $\mathcal{B}$ is not too large. When the number of MoBSs to be scheduled is equal to n, the cardinality of the number of subsets B is equal to 2^n, so that the solution of the problem `PFM` becomes rapidly problematic.

For this reason, we resorted to a heuristic that only uses a subset of configurations B. However the choice of the configurations to be considered in the heuristic requires some understanding of the relevance of different types of configurations in the optimization procedure.

After implementing the optimization for the train station area in each of the 18 time slots from 5 AM to 10 PM and obtaining the optimal probability vector we tried to characterize the configurations yielding the largest probabilities.

By examining the optimal solutions, we observed that in most cases the largest values of probabilities P_B are associated with subsets of MoBSs B comprising either a large number or a few MoBSs. In practice, our optimization generally schedules

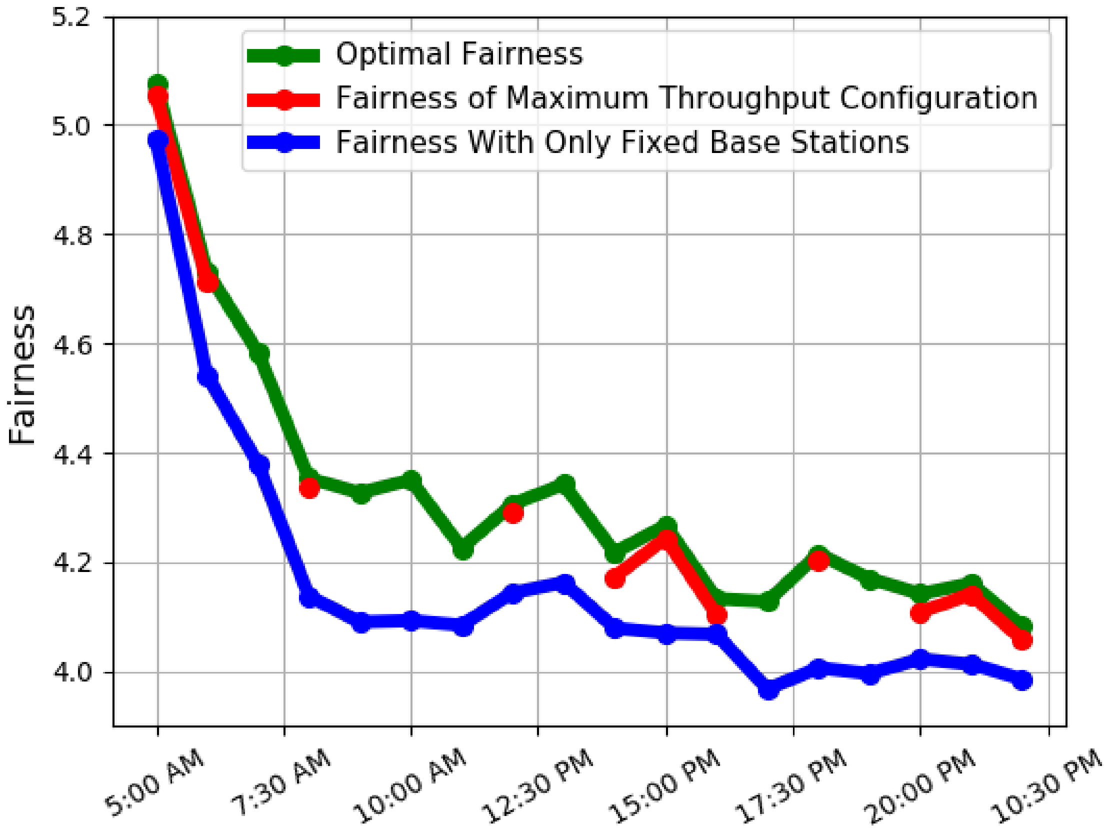


Figure 4.8: Fairness per user in the cases: i) only fixed BSs, ii) MoBSs configuration yielding the maximum throughput (points not reported corresponding to maximum throughput configurations where some users receive zero throughput), and iii) optimal fairness time scheduling, in the area of the central railway station in Milan on April 15 from 5 A.M. to 10 P.M.

all MoBSs and BSs together, except for some few MoBSs heavily impaired by interference.

Our analysis was based on a visualization of the MoBSs positions in each configuration, trying to identify common patterns. We investigated all the 18 time slots in the day, but here we just discuss the first two time slots.

The first time slot is at 5 A.M. After implementing the optimization, we extracted all probabilities that are larger than 10^{-3}. They are presented in Table 4.1 together with the corresponding binary representation of muted/unmuted MoBS configurations (0 means muted and 1 means unmuted). The number of cars in this time slot equals to 11, so that the overall number of states is 2048. Among them only 11 states have probability higher than 10^{-3}.

Table 4.1: MoBSs configurations, corresponding probabilities and binary representation at 5 A.M.

Configuration	Probability	Binary representation
2031	1.964010e-02	11111101111
1126	2.228112e-02	10001100110
258	2.991714e-02	00100000010
879	4.202345e-02	01101101111
1135	5.163970e-02	10001101111
1391	5.304009e-02	10101101111
1024	5.485979e-02	10000000000
1094	6.469376e-02	10001000110
2015	1.984901e-01	11111011111
256	2.159965e-01	00100000000
1134	2.474139e-01	10001101110

In Table 4.2, we present the MoBS configurations with probability of larger than 10^{-3} at 6 A.M. in the train station area, when 15 MoBSs are present.

Table 4.2: MoBSs configurations, corresponding probabilities and binary representation at 6 A.M.

Configuration	Probability	Binary representation
25402	2.052239e-02	110001100111010
4105	2.052792e-02	001000000001001
28671	4.220389e-02	110111111111111
16383	6.236284e-02	011111111111111
4634	6.691856e-02	001001000011010
4233	6.762712e-02	001000010001001
25376	8.552587e-02	110001100010111
25398	9.845642e-02	110001100110110
4617	1.143879e-01	001001000001001
4887	1.374714e-01	001001100010111
26558	2.137421e-01	110011110111110

From these two sets of MoBSs configurations and their binary representations we clearly see that most of the high probability states contain either most MoBSs in the unmuted state or very few MoBSs in the unmuted state. Similar patterns are observed also for all other time slots in the day.

Our understanding of this observation is that the optimizer unmutes most MoBSs, except the ones that cause the highest interference. However, in order to provide transmission opportunities also for the end users connected to those MoBSs that are excluded from the schedules where most are unmuted, the scheduler allocates time to them, either alone or with other MoBSs with which interference is low.

According to these considerations, we decided to run our heuristic on small sets of MoBSs configurations, comprising only:

- no unmuted MoBSs (one configuration)
- just one unmuted MoBS (n configurations)
- two unmuted MoBSs ($\binom{n}{2}$ configurations)
- all but two unmuted MoBSs ($\binom{n}{2}$ configurations)
- all but one unmuted MoBS (n configurations)
- all unmuted MoBSs (one configuration)

By doing so, the heuristic does not consider the entire set of 2^n elements of $B \in \mathcal{B}$. Rather, it considers $2+2n+2\binom{n}{2} = n^2+n+2$ configurations, and can thus be solved for large values of n. Numerical results in the next section show that the difference between the optimization results and the heuristic results is very small.

4.6.1 Heuristic Implementation

The `PFM` problem that we described in Section 4.3.3 is naturally a centralized implementation, which must first collect data from the MoBSs and UEs, then compute the optimal muting patterns, and redistribute those to BSs and MoBSs.

The frequency of this operation must be compatible with the extreme dynamicity of the MOMONET environment, which implies a significant consumption of resources (that we do not consider here) for the collection of data and the distribution of results. This makes the design of distributed, possibly suboptimal versions of the `PFM` problem quite a challenging and interesting topic for future work.

4.6.2 Numerical Results

The settings in which we computed the throughput achievable in MOMONET scenarios using the heuristic approach described above exploit the same dataset for telecom traffic and vehicle positions as in previous sections.

Specifically, our evaluation focuses on two representative case studies. The first is the one described in Section 4.4.1, while the second is an area of smaller size, covering the Polytechnic University of Milan (POLIMI).

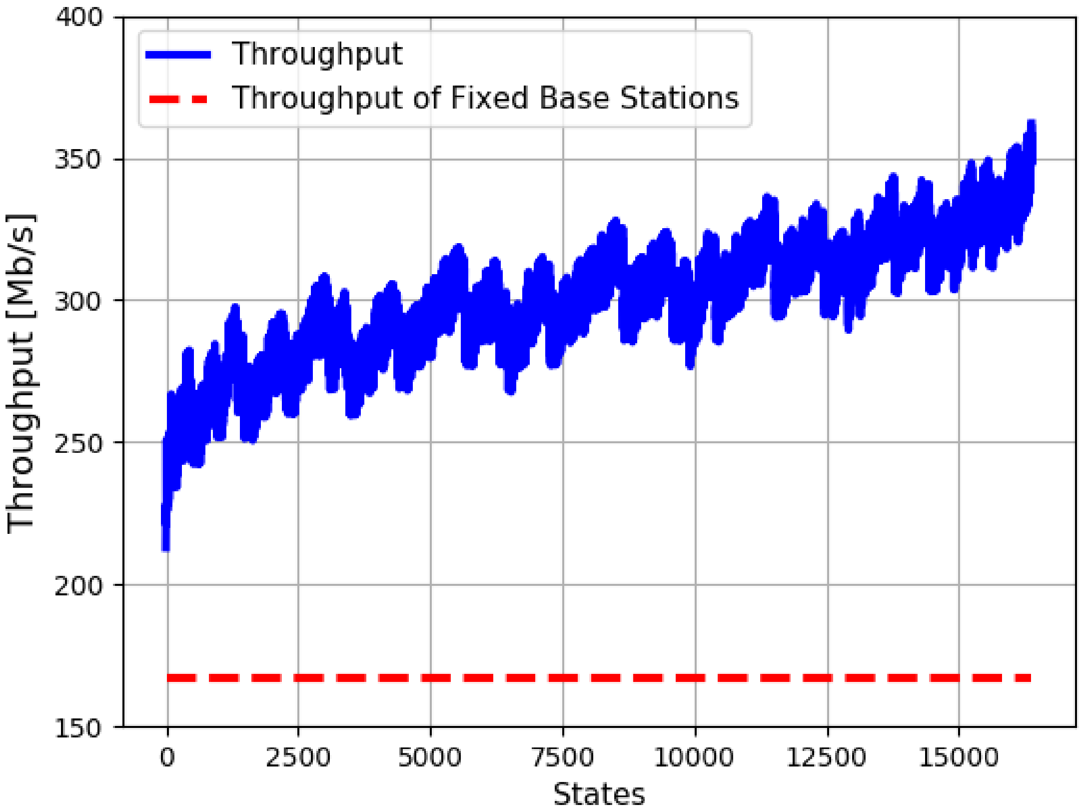


Figure 4.9: Throughput of the 2^{14} MoBSs configurations in the area of the central railway station in Milan on April 15, at 8 A.M.; configurations are ordered according to increasing number of active MoBSs. The red dashed horizontal line refers to the case of only fixed BSs, with users attached to fixed BSs only.

Use Case: Downtown Railway Station Area

By applying the procedure outlined in Section 4.6, it is possible to compute, for each time slot t, the maximum downlink throughput achievable as well as the throughput that corresponds to maximum fairness. We compute the maximum achievable throughput and the throughput at maximum fairness for the interval between 5 A.M. and 10 P.M. on April 15, 2015, at one-hour spacing.

The procedure first computes the throughput for each user in each system configuration, in fact for each configuration of active/inactive (i.e., transmitting/muted) useful MoBSs. Then, by summing over all users in a given configuration, the total throughput of each configuration is obtained.

The maximum over all configurations is the maximum achievable throughput in time slot t. As an example, Figure 4.9 reports the total throughput of all configurations in the considered area at 8 A.M. on April 15, ordering configurations so that all cases with one active MoBS appear first, then all cases with two active MoBSs, and so on, until the configuration with all active MoBSs is reached.

For equal number of active MoBSs, configurations are ordered according to their binary representation, (0 means muted and 1 means active; the all 0 configuration thus mutes all MoBSs, while the all 1 configuration has all MoBSs active). Since the number of possibly active MoBSs in the considered scenario is 14, the total number of configurations is $2^{14} = 16{,}384$. The red dashed horizontal line refers to the case of only fixed macro BSs, with users attached to fixed macro BSs only.

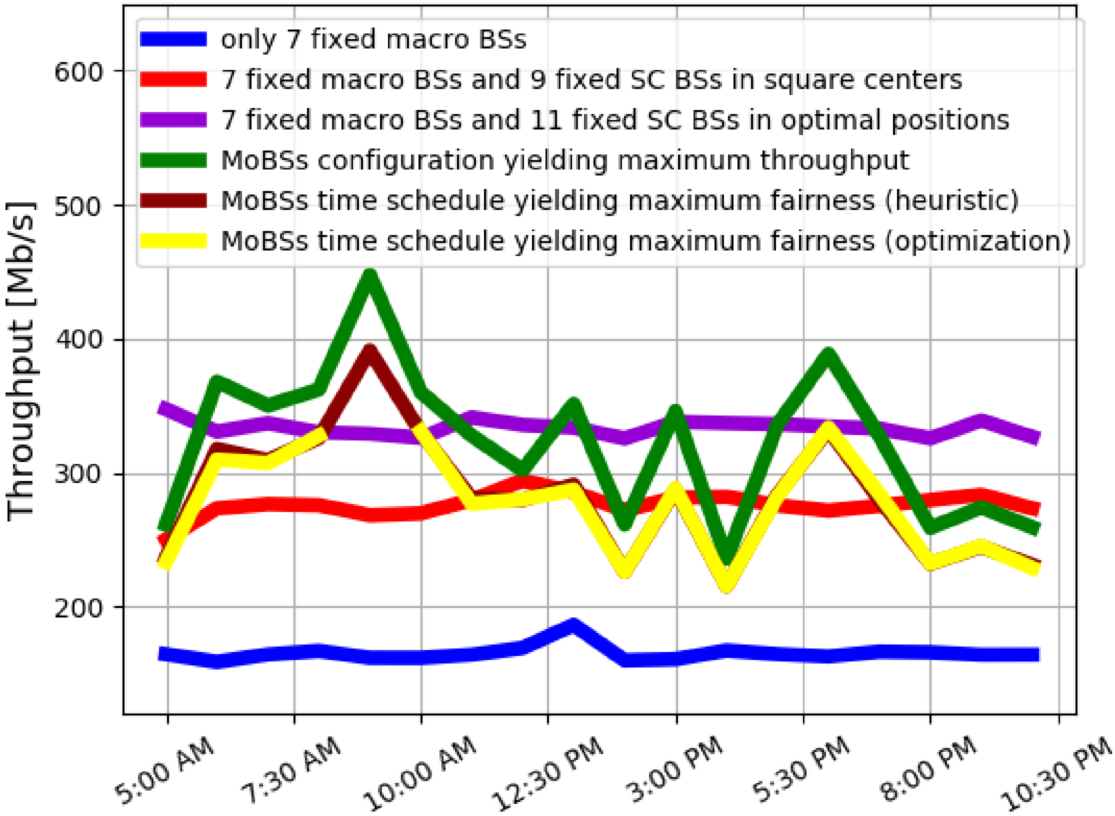


Figure 4.10: Throughput in Mb/s per user , in the area of the central railway station in Milan, on April 15 from 5 A.M. to 10 P.M., with: (i) only 7 fixed macro BSs; (ii) 7 fixed macro BSs and 9 fixed small cell BSs in rectangle centers; (iii) 7 fixed macro BSs and 11 fixed small cell BSs in optimal positions; (iv) the MoBSs configuration that yields maximum throughput; (v, vi) the MoBSs time schedule that yields maximum fairness (computed with the optimization and the heuristic).

In this case, the maximum throughput is reached when all 14 MoBSs are active, i.e., at the rightmost point in the graph. However, this is not always the case: at multiple times, maximum throughput configurations mute some MoBSs. Finally, by applying a time schedule that alternates over a set of configurations chosen according to PFM, a throughput is obtained, which optimizes fairness.

In Figure 4.10 we report, for each time slot of April 15, six throughput values: (1) the throughput with only the 7 fixed macro BSs (and all UEs associated to just fixed macro BSs); (2) the throughput with the 7 fixed macro BSs and 9 fixed SCs located in the centers of the 9 rectangles; (3) the throughput with the 7 fixed macro BSs and 11 fixed SCs strategically positioned at the edge of the macro BSs coverages; (4) the throughput of the MoBSs configuration yielding maximum throughput; (5), (6) the throughput of the time schedule over MoBSs configurations that optimizes fairness (computed using both the heuristic and the MATLAB optimization toolbox).

Of course, the throughput of the MoBSs configuration yielding maximum throughput shows the highest peaks, especially when the number of useful MoBSs is high.

As expected, the throughput with only 7 fixed macro BSs is the lowest. The throughput achieved with 9 fixed SCs in the rectangle centers is lower than the optimum fairness throughput achieved with MoBSs, except for most cases when the number of useful MoBSs is less than 9 (which happens at 5 A.M., 2 P.M., 4 P.M., 8 P.M., 9 P.M., 10 P.M.).

The throughput achieved with 11 fixed SCs in strategic positions is obviously higher than with 9 fixed SCs in the rectangle centers, and is also higher than the optimum fairness throughput achieved with MoBSs, except at 9 A.M. (when the number of useful MoBSs is 20) and 6 P.M. (when the number of useful MoBSs is 14).

This is due to the fact that the MoBSs positions are not optimal, but when a large number of MoBSs is present, the additional capacity makes up for the loss in efficiency due to suboptimal positioning.

The difference between the optimum fairness throughput computed through heuristic and the optimization algorithm is negligible. This is very good news, since the optimal solution of the PFM problem becomes problematic already with 20 MoBSs.

This is the reason why we do not show the point of the optimum fairness throughput at 9 A.M. in Figure 4.10. By using MoBSs we can meet gains of up to about 150% with respect to using only fixed BSs, when several vehicles are available.

Figure 4.11 reports the fairness values in the same six cases.

We can see that the fairness achieved when MoBSs are present is always higher than in the case of only fixed BSs, both when we consider only macro BSs, and when we consider the addition of 9 fixed SCs in the center of squares.

The fairness values of the MoBSs configurations providing maximum throughput

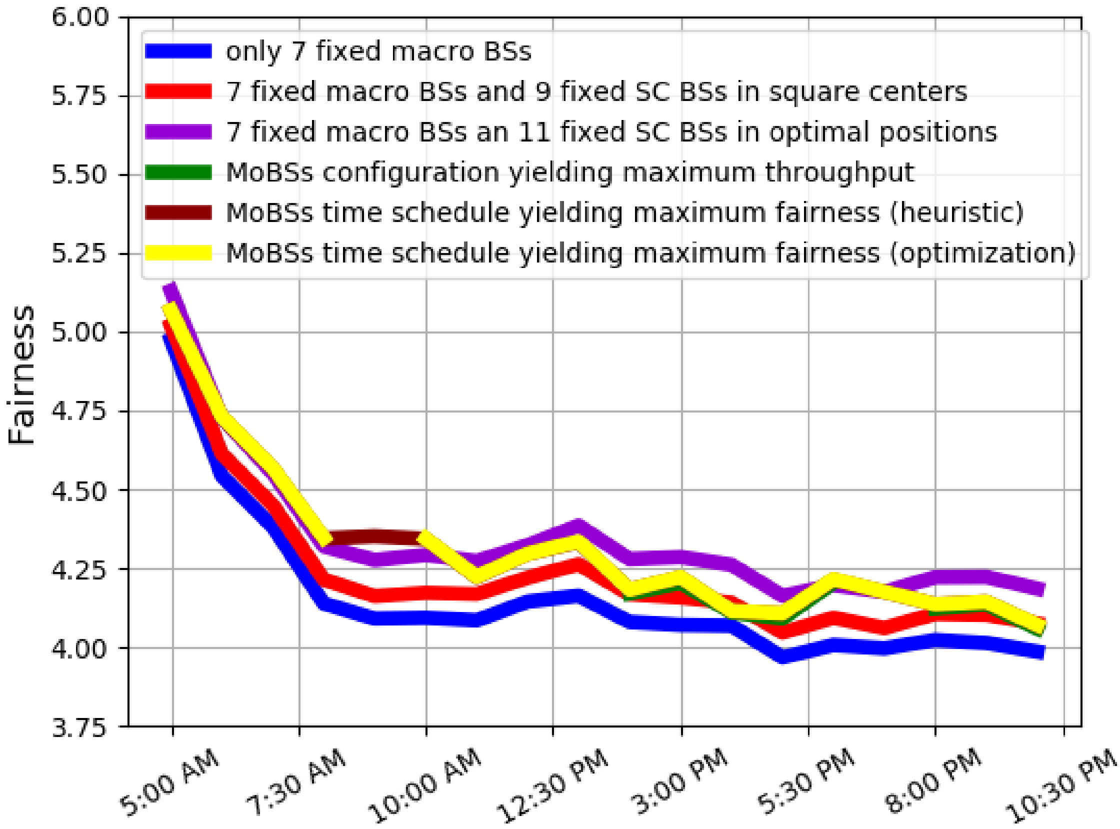


Figure 4.11: Fairness per user, in the area of the central railway station in Milan, on April 15 from 5 A.M. to 10 P.M., with: (i) only 7 fixed macro BSs; (ii) 7 fixed macro BSs and 9 fixed small cell BSs in rectangle centers; (iii) 7 fixed macro BSs and 11 fixed small cell BSs in optimal positions; (iv) the MoBSs configuration that yields maximum throughput; (v, vi) the MoBSs time schedule that yields maximum fairness (computed with the optimization and the heuristic).

in some cases are not reported, because when the max throughput state (which must be used with a resource share factor $P_B = 1$ to achieve the maximum throughput configuration) is such that not all useful MoBSs are active, some users receive zero throughput all the time, so that fairness takes the value of -∞.

This proves that the use of the time scheduler is extremely important in order not to exclude some UEs to access to network resources.

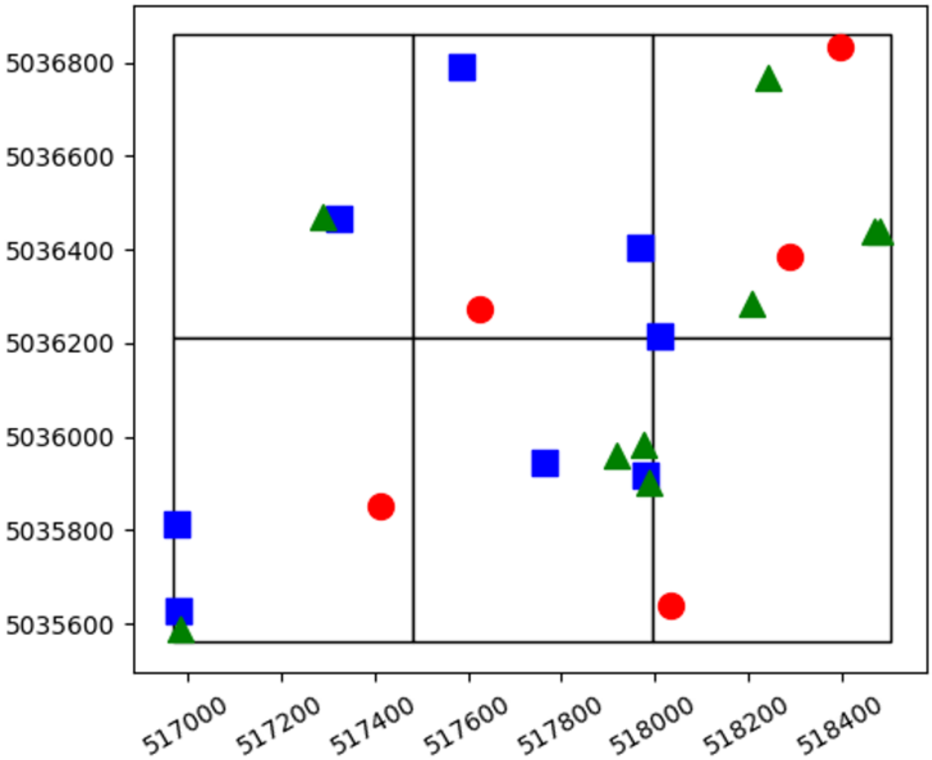


Figure 4.12: The POLIMI area is covered by 5 macro BS (red circles). According to our dataset, we consider 17 vehicle positions (blue rectangles and green triangles; the 9 green triangles correspond to discarded MoBSs) at 8 A.M. on April 15, 2015.

Use Case: POLIMI

The area covering the Polytechnic University is less central in the city of Milan and comprises 6 rectangles and 5 fixed macro BSs. Figure 4.12 reports, for the time interval from 8 A.M. to 8:15 A.M. on April 15, 2015, the vehicles positions as blue/green dots, together with the positions of fixed macro BSs as red dots. The total number of MoBSs, the number of useful MoBSs, and the total data traffic for each time slot are reported in Figure 4.13.

We can see that in this case the number of tracked cars and the number of useful MoBSs exhibit higher variability with respect to the railway station area. In particular, the number of useful MoBSs varies between 22 (at 6 A.M.) and 0 (at 9 P.M.).

In Figure 4.14 we report, for each time slot of April 15, five throughput values: (1) the throughput with only the 5 fixed macro BSs (and all UEs associated to just fixed macro BSs); (2) the throughput with the 5 fixed macro BSs and 6 fixed SCs located in the centers of the 6 rectangles; (3) the throughput with the 5 fixed macro BSs and 7 fixed SCs optimally positioned at the edge of the macro cell coverage; (4)

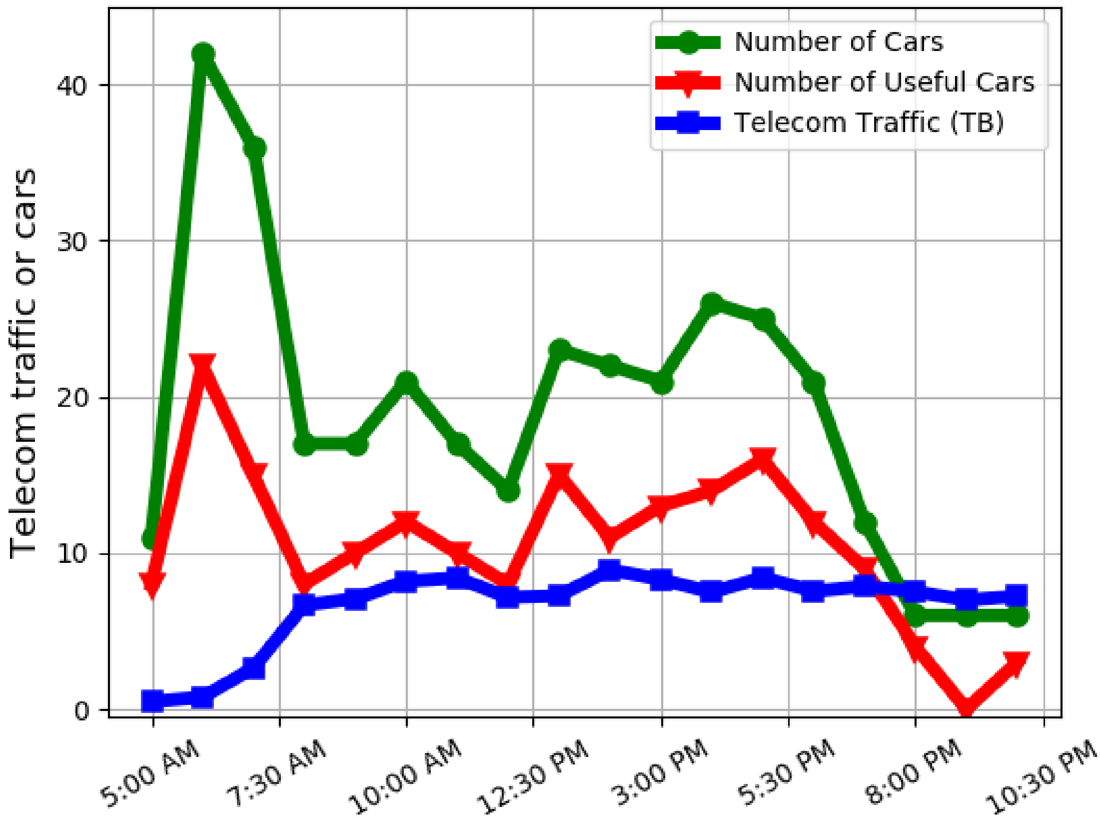


Figure 4.13: Number of vehicles, number of useful MoBSs and data traffic (in TB) in the area of the Polytechnic University of Milan on April 15, from 5 A.M. to 10 P.M.

the throughput of the MoBSs configuration yielding maximum throughput; (5) the throughput of the time schedule over MoBSs configurations that optimizes fairness (computed using the heuristic).

Also, in this case, the throughput of the MoBSs configuration yielding maximum throughput shows the highest peaks, especially when the number of useful MoBSs is high, and the throughput with only 5 fixed macro BSs is the lowest.

The throughput achieved with 6 fixed SCs in the rectangle centers is almost invariably lower than the optimum fairness throughput achieved with MoBSs, except for few cases with small number of useful MoBSs (in the evening at 8, 9 and 10 P.M., when the number of useful MoBSs are 4, 0, and 3, respectively).

The throughput achieved with 7 fixed SCs in strategic positions is slightly higher than with 6 fixed SCs in the rectangle centers (except at 5 A.M.), and in most cases significantly lower than the optimum fairness throughput achieved with MoBSs,

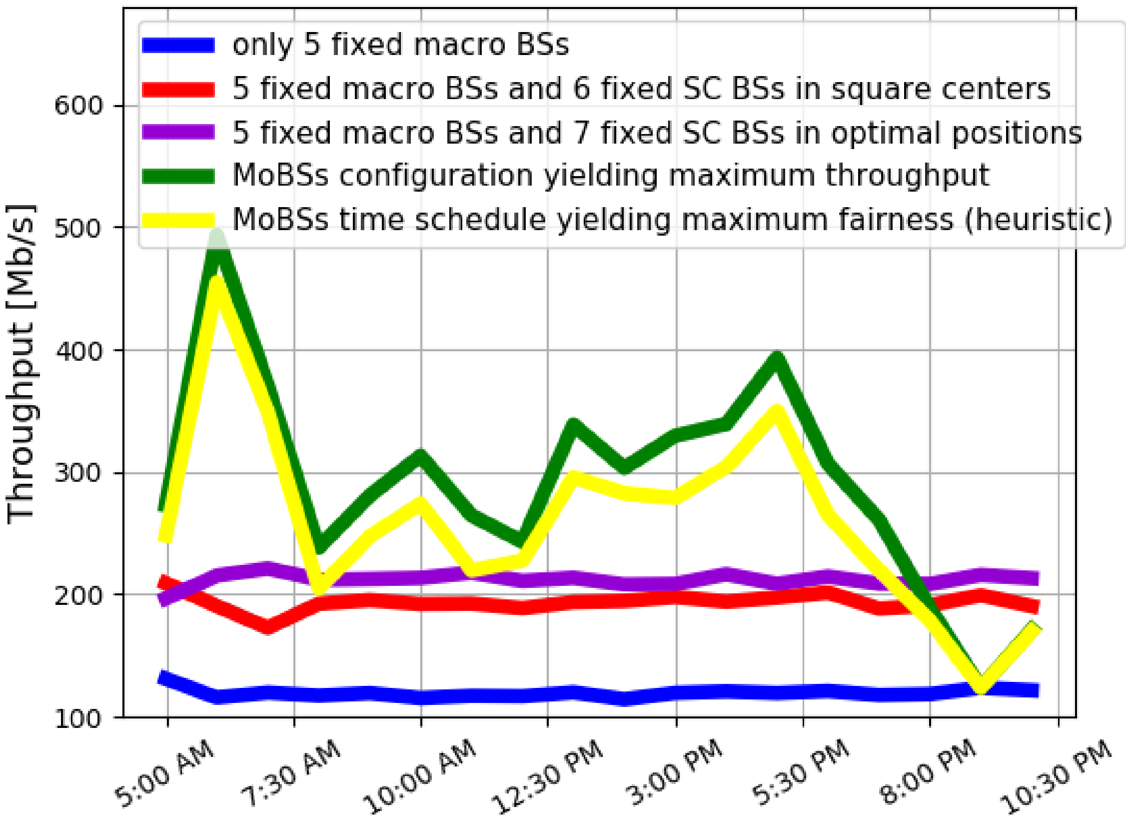


Figure 4.14: Throughput in Mb/s per user, in the area of the Polytechnic University of Milan, on April 15, from 5 A.M. to 10 P.M., with: (i) only 5 fixed macro BSs; (ii) 5 fixed macro BSs and 6 fixed small cell BSs in rectangle centers; (iii) 5 fixed macro BSs and 7 fixed small cell BSs in optimal positions; (iv) the MoBSs configuration that yields maximum throughput; (v) the MoBSs time schedule that yields maximum fairness (computed with the heuristic).

except for evening hours, when the number of useful MoBSs is very low.

In general, we see that, even if the MoBSs positions are not optimal, the maximum throughput achievable in the area can be drastically improved by the presence of MoBSs, especially in those time slots when a large number of MoBSs is present. For example, at 6 A.M., with 22 useful MoBSs, the throughput increases by about 300% with respect to the case of fixed macro BSs only.

The corresponding values of fairness are shown in Figure 4.15. We can see that the fairness achieved with the schedule resulting from the heuristic is significantly better than with only fixed macro BSs, or with fixed macro and SC BSs.

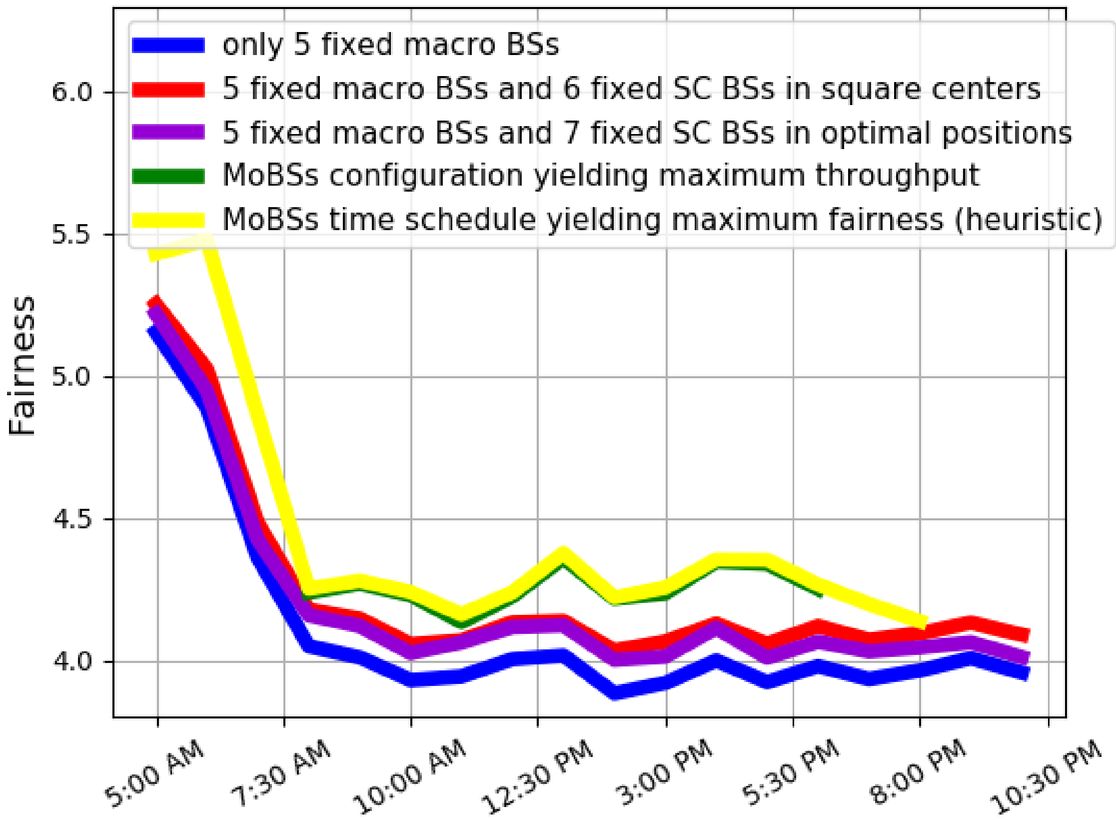


Figure 4.15: Fairness per user, in the area of the Polytechnic University of Milan, on April 15, from 5 A.M. to 10 P.M., with: (i) only 5 fixed macro BSs; (ii) 5 fixed macro BSs and 6 fixed small cell BSs in rectangle centers; (iii) 5 fixed macro BSs and 7 fixed small cell BSs in optimal positions; (iv) the MoBSs configuration that yields maximum throughput; (v) the MoBSs time schedule that yields maximum fairness (computed with the heuristic).

4.7 Conclusions

In this chapter we computed the throughput achievable with and without MoBSs, showing possible capacity increases of up to a factor 4 with respect to the case of macro base stations only, and capacity comparable to the optimal positioning of fixed small cells in similar number, considering the geographical areas around the central railway station and around the Polytechnic University in Milan, Italy, and using real data about data traffic and number of vehicles.

Our results prove that adaptive densification of radio access networks with MoBSs can be a high-performance and low-cost solution for the rollout of 5G and beyond.

Chapter 5

Wireless Backhaul between MoBSs and Macro BSs

In the previous chapter we have shown very significant capacity improvements resulting from the crowdsourced network densification with MoBSs. However, the results we presented rely on the optimistic assumption of a zero-interference infinite bandwidth backhaul connection between MoBSs and macro BSs. While such a link can in principle be obtained with a mmwave link, we know that such links have limited reach and are subject to outage due to obstructions. For this reason, it is important to investigate what is the performance achievable with the MOMONET concept when using more traditional technologies for backhaul links. This is the subject of this chapter.

In a mobile network, the backhaul network includes the intermediate links connecting the core network with the subnetworks at the periphery of the network.

Mobile end user terminals attached to one base station form a local subnetwork at the network periphery; the connection between the base station and the rest of the world starts with a backhaul link to the core of the mobile network operator network.

A backhaul link can be implemented with copper, fiber optic or wireless elements. Wireless solutions require careful design and implementation to guarantee the problem-free coexistence with transmissions between the base station and mobile terminals.

The choice of the technology for the backhaul implementation must consider many different parameters, such as capacity, reach, reliability, as well as the necessary resources, the cost, and the legal implications.

In selecting a backhaul technology, the first design choice is whether to implement a wired backhaul (using leased lines or copper/fiber cables deployed on purpose) or a wireless backhaul (either point-to-point, or point-to-multipoint; either single-hop or multi-hop). Wired backhaul solutions are often costly, and obviously not applicable to the MOMONET scenario. For this reason, they will not be considered further in this chapter.

Multi-hop wireless backhaul architectures can offer the advantage of low cost wide coverage over areas served by a number of adjacent base stations comparable to a wired backhaul connectivity[1].

A further advantage of multi-hop wireless backhaul solutions consists of easy deployment and flexibility, in the sense that they can be reconfigured to accommodate newly deployed base stations.

However, it must be considered that when wireless is chosen as the backhaul technology, some issues must be tackled, since wireless connections provide lower data rates and consume precious spectrum resources [**105**].

5.1 Wireless Backhaul in MOMONET

The technology selection for the implementation of the wireless backhaul link is possibly one of the most delicate choices for the MOMONET architecture, due to high capacity requirements, and to MoBSs mobility.

Other options, such as exploiting WiFi as a backhaul for cars parked sufficiently close to an open access point, or even a wired backhaul for electric vehicles connected to a recharge station, can be considered. However, the most promising solution seems to be a millimeter-wave multi-hop backhaul connection.

This approach does not generate interference with the much lower frequencies used in the wireless connections from MoBSs to end user terminals, and is expected to provide capacities which are largely in excess of the sum of the data rates from MoBSs to UEs. However, it must be considered that millimeter-wave links in urban environments can be problematic, due to both distance and obstacles [**50**].

5.2 5G Backhaul Perspective

As discussed in previous sections, 5G faces many challenges concerning different aspects. One of the most important bottlenecks emerging in 5G networks is the backhaul implementation. The new 5G backhaul architecture encompasses different technologies in order to accommodate extreme traffic and new upcoming demands. Indeed, the most relevant 5G property that the backhaul should consider is the diversity of service requirements.

The present situation of mobile networks is such that backhaul links are mainly implemented with microwave or fiber/copper. Proportions of the technologies differ from country to country and operator to operator [**51**]. Backhaul provision with optical fiber is not pervasive, and microwave replacement of optical fiber may not

[1] https://en.wikipedia.org/wiki/Backhaul$_{(}$*telecommunications*)

be able to cope with the traffic growth predictions. So, it is absolutely evident that providing backhaul needs innovation in 5G. Available backhaul technologies can be employed through different topologies[**51**][2].

5.2.1 Advances in mm Wave

The most interesting characteristic of mmWaves that makes them attractive for 5G is that they offer large quantities of untapped bandwidth, especially in the 60-100 GHz region of spectrum[3]

However, while this bandwidth to a large extent is not utilized, due to propagation conditions only short distances can be covered.

This creates both a problem and an opportunity. A problem because the link reach is short, but an opportunity because frequency can be reused much more frequently in space. In addition, the link that uses the mmwave frequency band surely does not generate interference with links using traditional cellular bandwidths (those that are now used to reach end user terminals).

A further characteristic of mmwave links is that they usually require Line of Sight (LoS) conditions, since reflections imply significant losses, and obstacles are likely to generate link outages. While LoS may be simple to guarantee (at least with high probability) in the case of fixed small cell backhauling, this is obviously much more problematic in the case of MoBSs.

In[4], mm Wave connection is proposed to implement ultra dense 5G networks causing data rates of nearly 10 Gbps be achieved. Also self-backhauling with mmWave-based technology and interference-aware routing are recommended to overcome the difficulty and the cost of wired fronthaul/backhaul connections.

Another backhauling technology is in-band backhauling, that is based on the reuse of the radio frequencies used to reach end user terminals to also support wireless backhaul links. Such in-band backhauling solution offers simple implementation and lower cost, thanks to hardware and spectrum reuse, but it provides a much more limited bandwidth in the backhaul connection.

A further option consists in separating the frequency bands used to reach end user terminals and to implement backhaul connections. This option will be called out-band. In this case the bandwidth of the backhaul link can be controlled, and no interference is generated between backhaul and end user connections. However,

[2] F. Barros, K. Ahl, and V. Chaillou. (2015). Creating a Connected Continent: Press Conference Warsaw. Available: http://www.ftthcouncil.eu/

[3] A. Goldsmith, "Do We Have a Shortage of Bandwidth or Imagination?,". Available: http://technical.ly/brooklyn/2014/04/30/andrea-goldsmith-bandwidth/.

[4] 3rd Generation Partnership Project, TS36.413 V9.2.1, Evolved Universal Terrestrial Radio Access (E-UTRA); S1 Application Protocol (S1 AP), 3GPP, 2010.

the efficiency of the system can be limited because of bandwidth requirements and limited efficiency due to resource separation.

A number of other proposal for the implementation of backhaul connections were put forward in different works, proposing major and minor differences and changes. Recently, for example, an article in [**52**], investigates the challenges of combining mmWave technologies and massive MIMO in 5G.

Several authors have discussed the relative advantages of point to point and point to multipoint microwave backhaul. While the point to multipoint option has the advantage of a higher traffic aggregation, hence higher efficiency in the link utilization, thanks to the aggregation of the very bursty traffic originated by several small cells, it is at risk of suffering capacity shortage in case of multiple simultaneous traffic peaks.

5.2.2 SDN in the Backhaul

SDN is currently one of the key players in facilitating 5G networks. In fact, SDN enables multiple operators to share the same physical network, and so it is an efficient tool for resource reuse and reduction of CAPEX.

For example, a research group in [**49**], proposed a new SDN tool that is called a backhaul resource manager that provides a flexible hybrid mmwave and optical backhaul with high capacity. This resource manager in the backhaul implements dynamic automated resource providing, as well as capacity-aware path computation. As a result it will improve fairness, network usage and QoE of end-to-end user.

In another recent research, the strength of SDN is investigated in optimizing the mobile backhaul network performance by searching the optimal backhaul path (according to the existing capacity and latency) and allocating the necessary wavelength. The strength and benefits of developing a backhaul based on SDN are numerous.

An SDN architecture in the backhaul creates the possibility of addition, extension and dynamic reallocation of resources in the backhaul network. This means that the adoption of SDN-based technology helps the backhaul network to share multiple technologies and to maximize the resource utilization efficiency.

5.2.3 Conclusions and Outlook

The challenges on 5G backhaul are numerous: More than 10 Gbps of capacity, less than 1 millisecond of latency, high security, synchronization of time and frequency and flexibility, low level of energy consumption and low cost.

Currently, none of these backhaul scenarios can individually meet all of these requirements, but fiber optic-based scenarios are best positioned, and HetNet backhaul solutions are emerging in parallel. If wireless backhaul is necessary, then point-to-multipoint mmwave solution seem to be the best solution.

Considering the already enormous and growing size of networks and their related parameters, together with the growing relevance of the backhaul in 5G, the adoption of the SON (self organizing network) paradigm is becoming mandatory. The principal benefit of SON is its adaptive nature that can cope with the network dynamics.

In this context, technologies like SDN play a critical role, since they will facilitate the management of the backhaul network in presence of SON operations.

5.3 Backhaul in MOMONET

The selection of the technology for the implementation of the wireless backhaul link is possibly one of the most delicate choices for the MOMONET architecture, due to high capacity requirements, and to MoBSs mobility.

In the preliminary analysis of Chapter 4, we did not consider the impact of the backhaul link since we assume the backhaul link always has the necessary capacity, and does not generate interference with transmissions from either BSs or MoBSs to UEs. This choice was justified by declaring that the backhaul implementation is based on the millimeter wave technology, so that a mmwave connection between the MoBS and the macro base station is available.

The advantage of a mm Wave (single or multi-hop) backhaul connection is that it does not generate interference with much lower frequencies used in wireless connections from MoBSs to end user terminals and is expected to provide capacities that are largely in excess of the sum of the data rates from MoBSs to UEs.

However, it must be considered that millimeter-wave links in urban environments can cause prblem due to both distance and obstacles. It is thus very important to study the possibility of using other approaches, that can at least support the connection from MoBSs to macro BSs for those periods when the mmwave connection from the MoBSs to the fixed network cannot be active.

We take into consideration two approaches as alternatives of the usage of mmwaves.

1. Out-band backhaul: The link between BSs and MoBSs exploits a dedicated bandwidth (e.g., on 2G/3G bandwidth) but, contrarily to what we assumed in the mmwave scenario, macro base stations backhaul transmissions towards MoBSs interfere with one another and backhaul link capacities are not infinite.

   We assume, like in the mmwave case, that macro BSs are always active, while the ABS approach is used to mute/unmute MoBSs both in the bandwidth for backhauling and in the bandwidth for downlink transmissions towards UEs. In addition, we consider that the traffic received by UEs attached to MoBSs is constrained by the capacity of the corresponding backhauling link.

2. In-band backhaul: The link between macro base stations and MoBSs exploits the same bandwidth and the same time intervals as the communications between BSs/MoBSs and UEs. This means that MoBSs and UEs are simultaneously served by macro BSs transmissions that can interfere with each other.

   As in previous cases, we assume BSs to be always active, while the ABS approach is used to mute/unmute MoBSs so as to reduce interference.

The notation that we use to formulate the `PFM` optimization problem in the two new cases of outband and inband backahul is the following.

- A: the set of all macro BSs;
- B: a possible subset of MoBSs;
- $\mathcal{B}$: the set of possible subsets of MoBSs;
- P_B: the probability that subset B is scheduled;
- N: users in the system;
- i: index of UEs;
- c: index of MoBSs;
- N_x: the set of users associated to x (x may be either a MoBS c or a macro BS);
- Γ_x^y: Throughput of x (either a MoBS c or a UE i), when the subset y of transmitters is active.

Please note that in the following we do not consider user attachment. Nevertheless, a strongest power approach is not enough. Indeed, it is of fundamental importance that users attach to a MoBS only if the link BS/MoBS present a better channel quality than the BS/UE link.

5.3.1 Problem `PFM-outband backhaul`

At time t, with N UEs in the area, select P_B; $\forall B \in \mathcal{B}$, so to:

$$\begin{aligned}
\text{maximize } \frac{1}{N} \quad & \sum_i \log\left(\sum_{B\subseteq\mathcal{B}} P_B\Gamma_i^B(t)\right); \\
\text{subject to:} \quad & \Gamma_c^A(t) = \left(\sum_{B\subseteq\mathcal{B}} \sum_{i\subseteq\mathcal{N}_\rfloor} P_B\Gamma_i^B(t)\right), \forall\ c \\
& \sum_{B\subseteq\mathcal{B}} P_B = 1, \\
& P_B \in [0{,}1], \quad \forall\ B\in\mathcal{B};
\end{aligned} \tag{5.1}$$

This optimization is a convex optimization problem and can be solved with off-the-shelf optimizers. Nevertheless, the problem is exponential in the number of MoBSs, and in the cases of large numbers of MoBSs it is necessary to exploit to heuristics, like the original `PFM` problem.

5.3.2 Problem `PFM-inband backhaul`

At time t, with N UEs in the area, select P_B; $\forall B\in\mathcal{B}$, so to:

$$\begin{aligned}
\text{maximize } \frac{1}{N} \quad & \sum_i \log\left(\sum_{B\subseteq\mathcal{B}} P_B\Gamma_i^B(t)\right); \\
\text{subject to:} \quad & \left(\sum_{B\subseteq\mathcal{B}} P_B\Gamma_c^B(t)\right) = \left(\sum_{B\subseteq\mathcal{B}} \sum_{i\subseteq\mathcal{N}_\rfloor} P_B\Gamma_i^B(t)\right), \forall\ c \\
& \sum_{B\subseteq\mathcal{B}} P_B = 1, \\
& P_B \in [0{,}1], \quad \forall\ B\in\mathcal{B};
\end{aligned} \tag{5.2}$$

The optimization is convex also in this case and can be solved with off-the-shelf optimizers.

5.3.3 Fairness and Throughput Maximization

The optimization problems presented above provide the values of MoBSs configurations probabilities that yield the maximum overall proportional fairness. However, a small variation of the same problem can also be used to maximize the system throughput. Indeed, by removing the log in the objective functions, the overall throughput is maximized, obviously with no consideration of fairness.

Optimizing the throughput leads to linear optimizations, which can be solved in linear time. Nevertheless, it should be noted that, with throughput optimization, the problem complexity resides in computing the throughputs Γ_x^y which again scale exponentially with the number of MoBSs.

In the next subsection we report results for throughput and fairness resulting from the fairness optimization; i.e., we report the maximal fairness, and the throughput achieved in the maximal fairness condition.

5.3.4 Numerical Results

For the derivation of numerical results with the new types of wireless backhaul connections, we again consider the scenarios of Sections 4.4 and 4.6.

In Fig.5.1 we report the throughput in the area of the main railway station of Milano at 1h intervals on April 15,2015, considering six different scenarios:

- 7 macro BSs only
- 7 macro BSs and 9 fixed SC BS in centers of squares
- 7 macro BSs and 11 fixed SC BS in optimal positions
- 7 macro BSs and MoBS with mmwave backhaul
- 7 macro BSs and MoBS with outband backhaul
- 7 macro BSs and MoBS with inband backhaul

For fixed SC BSs we assume that those are equipped with a wired backhaul. In the case of optimized fixed SC BSs positions, we assume they are placed at cell edges, at the mid point of the segment connecting two adjacent macro BSs. In the case of MoBSs, their number is as reported in Fig4.5.

As expected, the highest throughput values are obtained with 7 macro BSs and 11 fixed SC BSs in optimal positions, except for 9 A.M., when the number of MoBSs is almost twice the number of fixed SC BSs.

The throughput obtained with 7 macro BSs and 9 fixed SC BS in centers of squares is comparable in average to the one achieved with MoBSs equipped with a mmwave backhaul, whose variability follows the hourly variation in the number of MoBSs.

The hierarchy among the backhaul technologies is as expected, since the wireless mmwave backhaul case provides more throughput than the outband backhaul case, which in turn provides more throughput than the inband backhaul cases. Interestingly, even the inband wireless backhaul approach provides in some cases about 50% throughput increase with respect to the use of macro BSs only.

The corresponding results for fairness, shown in Fig.5.2, exhibit more limited differences, but the main message in this case is that all considered options offer better fairness than the standard case in which only macro BSs are used, and that, as expected, the highest fairness is achieved with SC BSs in optimal positions, as expected, because in this case SC BSs increase the throughput of least priviledged UEs.

The results for the scenario of Milan Polytechnic are presented in Figs. 5.3 and 5.4.

In this case we see that the throughput at maximum fairness obtained by using MoBSs with a mmwave wireless backhaul significantly outperforms the case of 7

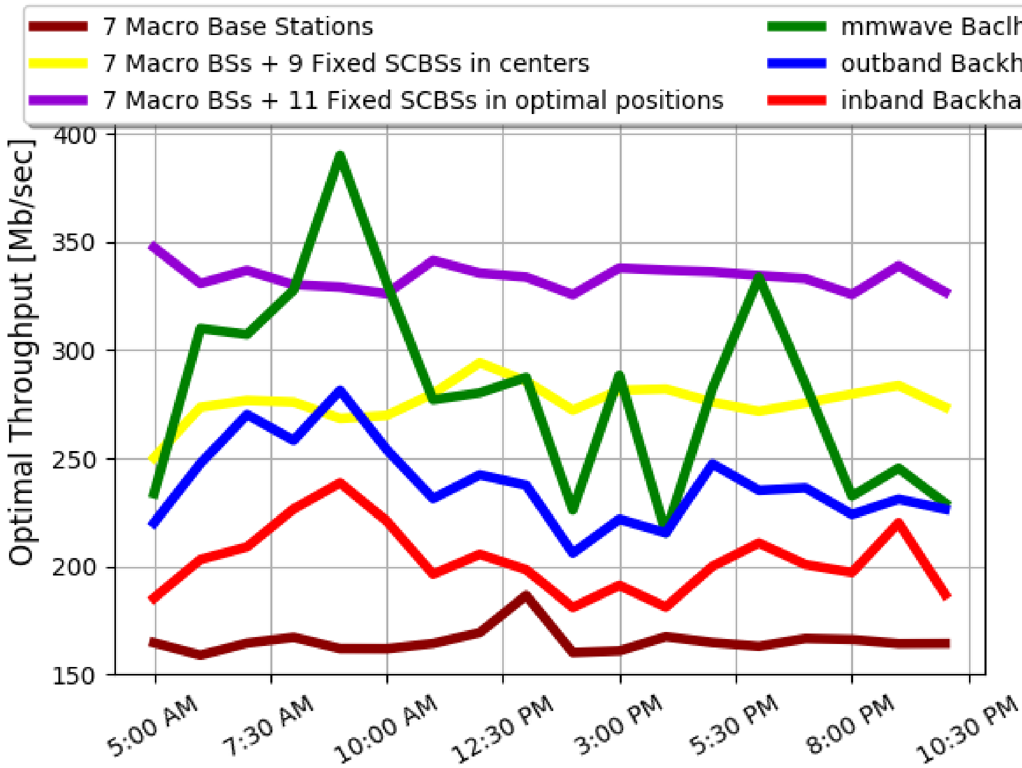


Figure 5.1: Throughput at maximum fairness in the case of the railway station scenario versus time of the day, for macro BSs only, two cases of fixed SC BSs, and different wireless backhaul technologies.

fixed SC BSs in optimal positions. This is due to the fact that the number of MoBSs at most times is much larger than 7 (see Fig. 4.13).

The same does not happen for the cases of outband and inband wireless backhaul, that however still significantly outperform the case of macro BSs only.

As regards optimal fairness (Fig. 5.4), the same hierarchy as for the railway station scenario is observed. Lowest fairness is provided by the case with macro BSs only, and highest fairness is provided by SC BSs in optimal positions. All other cases provide intermediate values, and differences are limited.

5.4 Conclusions

The comparison of the performances achievable in the cases of mmwave, outband, and inband backahul reveals that, as expected, the mmwave option leads to best performance, but even in the cases of outband and inband backhaul the

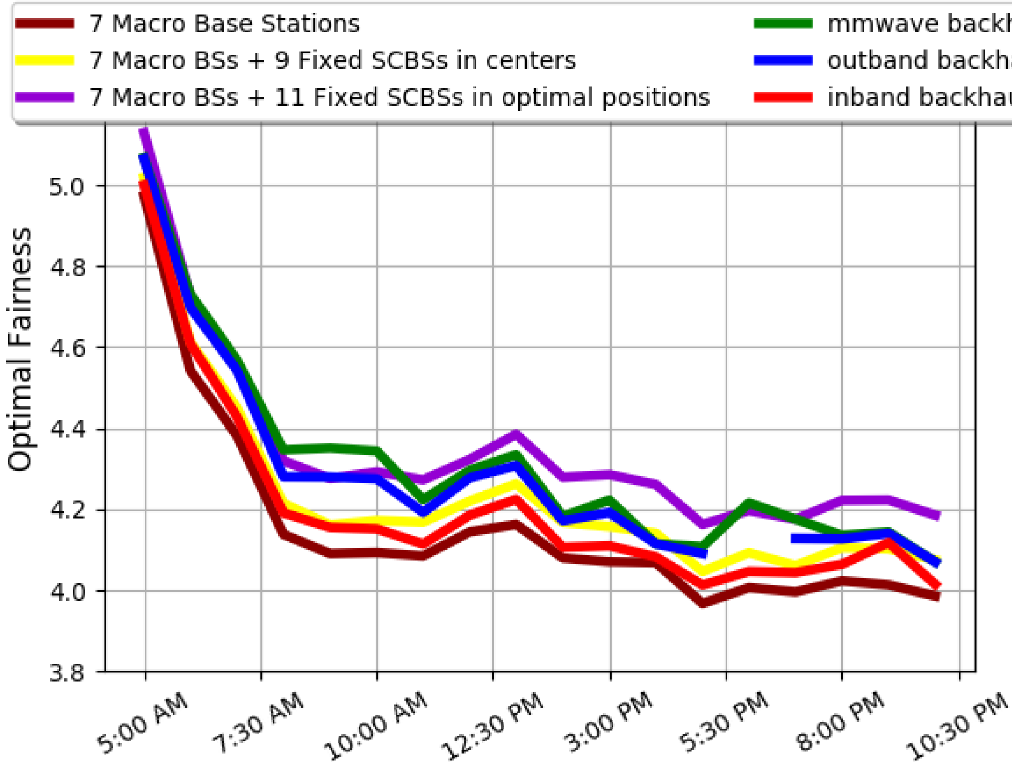


Figure 5.2: Maximum fairness in the case of the railway station scenario versus time of the day, for macro BSs only, two cases of fixed SC BSs, and different wireless backhaul technologies.

MOMONET approach can be effective. This is quite relevant, since the mmwave approach is known to be prone to link outages due to obstructions, and alternative approaches should be used whenever the mmwave link is disrupted.

A deeper analysis of the backhaul options in MOMONET should consider the possibility of exploiting multi-hop mmwave connections, or parallel mmwave and inband (or outband) connections, so that in case of outage of the direct mmwave link between the MoBS and a fixed BS, other options are available.

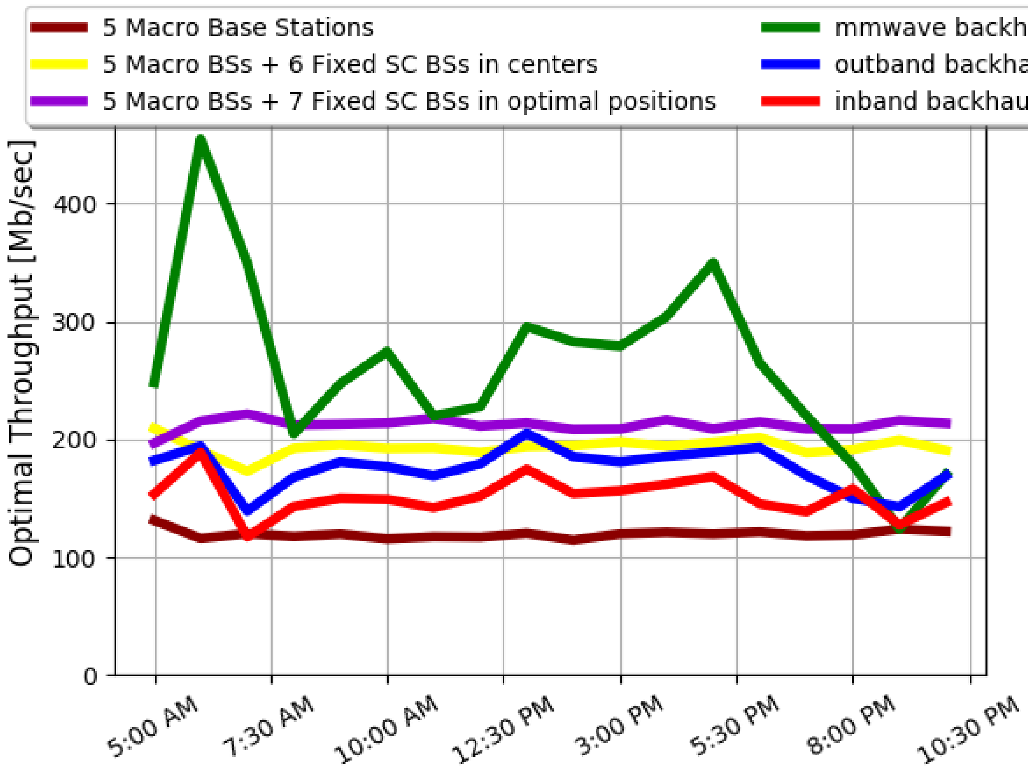


Figure 5.3: Throughput at maximum fairness in the case of the Milan Politechnic scenario versus time of the day, for macro BSs only, two cases of fixed SC BSs, and different wireless backhaul technologies.

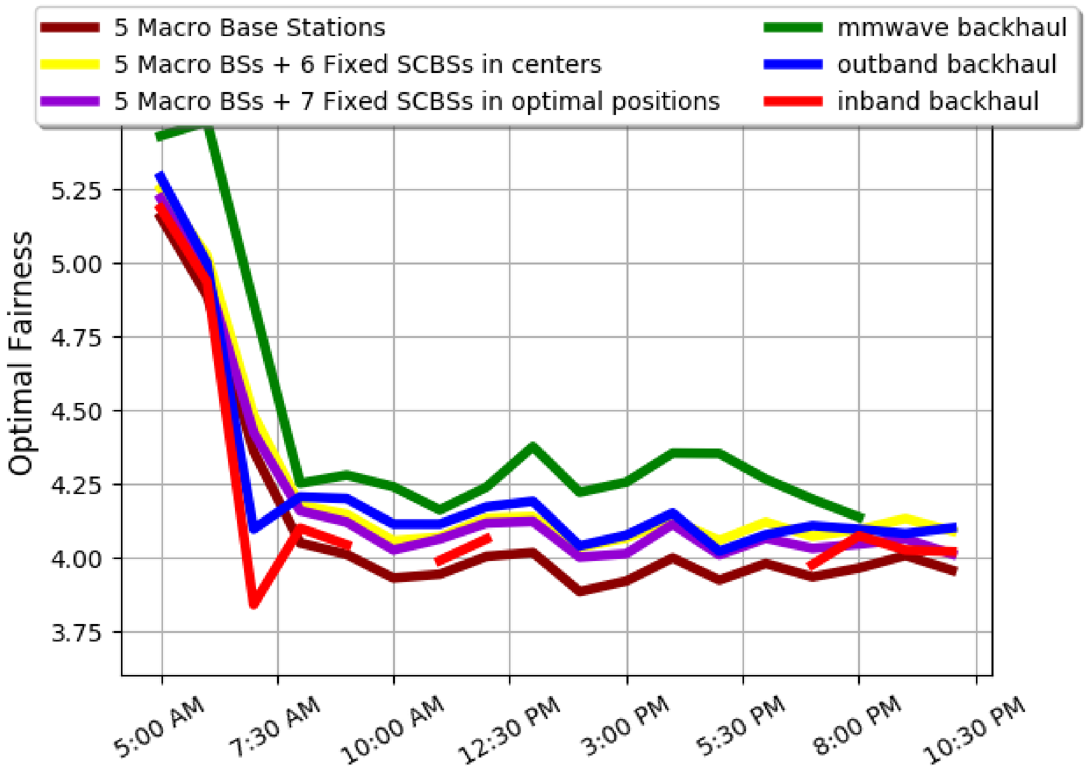


Figure 5.4: Maximum fairness in the case of Milan Polytechnic scenario versus time of the day, for macro BSs only, two cases of fixed SC BSs, and different wireless backhaul technologies.

Chapter 6

Conclusions and Discussion

This dissertation is devoted to the study of the benefits of radio access network densification through the use of small cells carried on board vehicles, aiming in particular at the assessment of the achievable capacity and fairness improvements.

My studies demonstrate how network densification with the MOMONET approach can be a viable alternative to an extensive deloyment of fixed small cell base stations.

Specifically, in chapter 3 I studied the spatial, temporal and the spatio-temporal correlation between telecom and vehicular traffic, as well as the expected distance between mobile small cell base stations and user terminals, which is a critical parameter to determine the applicability of the MOMONET concept, and in particular for the end user access to mobile small cell base stations. For these analyses I used two large datasets generated by TIM and Infoblu, that include real data about both types of traffic. After the analysis in this chapter, I concluded that there exists a reasonable amount of correlation in time and in space between the two different types of traffic and MOMONET can be applicable where and when it is needed, specially in the city center where the average distance between end users and mobile small cell base stations is much shorter.

In chapter 4, by again using the same datasets, I studied and analyzed the throughput and fairness obtainable in a MOMONET scenario, without considering the impact of the wireless backhaul link from mobile small cell base stations to macro base stations. This amounts to assuming that a millimeter wave wireless backhaul link is always available, and that it provides the sufficient capacity at any time. Under this assumption, I was able to compute huge improvements in the user throughput and fairness with respect to the case in which only macro base stations are present. In some cases the capacity over the considered portion of the radio access network was shown to grow up to 150%. I also computed under the same conditions the throughput increase when a few fixed small cell base stations are used for radio access network densification, and found that the use of mobile small cell base stations leads to similar throughput values.

The solution of the optimization problems to compute the network capacity and fairness turned out to be problematic because of the exponential increase of the problem size with the number of small cell base stations used for densification. I thus introduced simple heuristic algorithms that drastically reduce the problem complexity, and that were shown to compute results very close to the actual optimization.

The investigation of the effect of choosing different technologies for the implementation of the wireless backhaul connection from mobile small cell base stations to fixed macro base stations was the last part of my study. In this case I could reuse most of the approaches adopted in previous chapters to investigate two additional technological options.

In my analysis in Chapter 4 we assumed the backhaul connection to have infinite capacity and generate no interference with the transmissions towards the end user terminals, the first backhaul alternative assumes the availability of two separate portions of the frequency spectrum, one for the backhaul connection and one for transmissions towards end user equipment. The two spectrum portions have equal bandwidth. The second backhaul alternative assumes that only the bandwith used to reach end user equipment is available, and that it must support also the connection of mobile small cell base stations to fixed macro base stations. Numerical results show that a mmwave wireless backhaul produces, as expected, higher total throughput and fairness with respect to the two alternatives we considered, but differences remain limited.

It is worth observing that the results in this dissertation were derived using several simplifying assumptions, whose elimination would produce more realistic results. In particular, the impact of the following assumptions should be assessed in future works.

- I used a signal propagation model which is solely based on distance between transmitter and receiver, while it is well known that, especially in urban environments, propagation is strongly influenced by the built environment. This has a significant effect on the evaluation of the level of interference generated by fixed BSs and MoBSs, as well as on the capacity of channels between end users and BSs or MoBSs, and of backhaul connections between MoBSs and fixed BSs.

- I assumed the end user population to be uniformly distributed over each square of the grid covering the city of Milano, while typically end users cluster in specific areas. Clustering of end users can have a non-marginal impact on performance, with either positive or negative effects, depending on the relative positions of user clusters and BSs/MoBSs.

- I did not differentiate between indoor and outdoor users, while it is clear that MoBSs move outdoors, and can much better serve outdoor users, because of both closer proximity and better propagation conditions.

- While I based my optimization on proportional fairness, the robustness of the system performance to different optimization objectives (e.g. max-min fairness) should be explored.
- I applied ad-hoc heuristic approaches based on the inspection of the optimum solution of smaller cases, to the solution of optimization problems in cases with large number of states. However, for increased generality, it could be worth to investigate the applicability of meta-heuristic optimization approaches, such as extremal or genetic algorithms, to the optimization of cases with large numbers of components.

Despite the rather large set of topics covered in this dissertation, many more different aspects in the MOMONET architecture deserve a careful investigation. In particular, while I studied a sequence of stationary snapshots, the dynamic analysis of the behavior of a MOMONET is surely necessary, to verify the validity of the results obtained with the snapshot approach, and to observe the number and the effect of handovers of end user terminals to/from fixed macro base stations from/to mobile small cell base stations.

In addition, the problem of managing user associations to base stations of either type is also quite interesting and may reveal advantages and limitations of the MOMONET approach.

Also, while I considered that only mobile small cell base stations actively operate so as to reduce interference, the case in which also fixed macro base stations are periodically muted to reduce interference is an interesting option to investigate.

Finally, in this dissertation I concentrated on the performance of MOMONET, never considering energy consumption aspects. However, MOMONET can offer interesting advantages also along the energy dimension, since MoBSs can obtain the energy necessary for their operation directly from the vehicles that carry them. This makes the analysis particularly relevant and interesting.

I expect and hope that the studies developed in this dissertation can be a good and reasonable starting point to develop further studies by other researchers of this field.

The non-technical merits of the adoption of small cell base stations on vehicles were briefly touched in Chapter 2. However, a careful assessment of the economical aspects of the MOMONET option deserve a careful investigation.

I firmly believe that the MOMONET architecture is a viable approach for network densification given the continuous growth of telecom traffic and the need to reduce network operator capital and operational expenditures.

Bibliography

[1] A. Gupta, R. K. Jha, "A Survey of 5G Network: Architecture and Emerging Technologies," IEEE Access, vol. 3, pp. 1206-1232, 2015.

[2] B. Romanous, N. Bitar, A. Imran, H. Refai, "Network densification: Challenges and opportunities in enabling 5G," 20th IEEE International Workshop on Computer Aided Modelling and Design of Communication Links and Networks (CAMAD), IEEE, 2015.

[3] V. Chandrasekhar, J. G. Andrews, and A. Gatherer, "Femtocell networks: A survey," IEEE Commun. Mag., vol. 46, no. 9, pp. 59-67, 2008.

[4] F. Rusek et al., "Scaling up MIMO: Opportunities and challenges with very large arrays," IEEE Signal Processing Magazine, vol. 30, no. 1, pp. 40-60, 2016.

[5] A. Osseiran et al., "Scenarios for 5G mobile and wireless communications: the vision of the METIS project," IEEE Commun. Mag., vol. 52, no. 5, pp. 26-35, 2014.

[6] E. G. Larsson, F. Tufvesson, O. Edfors, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," IEEE Commun. Mag., vol. 52, no. 2, pp. 186-195, 2014.

[7] W. Nam, D. Bai, J. Lee, and I. Kang, "Advanced interference management for 5G cellular networks," IEEE Commun. Mag., vol. 52, no. 5, pp. 52-60, 2014.

[8] M. N. Tehrani, M. Uysal, and H. Yanikomeroglu, "Device-to-device communication in 5G cellular networks: Challenges, solutions, and future directions," IEEE Commun. Mag., vol. 52, no. 5, pp. 86-92, 2014.

[9] T. S. Rappaport et al., "Millimeter wave mobile communications for 5G cellular: It will work!" IEEE Access, vol. 1, pp. 335-349, 2013.

[10] F. Gutierrez, S. Agarwal, K. Parrish, and T. S. Rappaport, "On-chip integrated antenna structures in CMOS for 60 GHz WPAN systems," IEEE J. Sel. Areas Commun., vol. 27, no. 8, pp. 1367-1378, 2009.

[11] T. S. Rappaport, E. Ben-Dor, J. N. Murdock, and Y. Qiao, "38 GHz and 60 GHz angle-dependent propagation for cellular peer-to-peer wireless communications," in Proc. IEEE Int. Conf. Commun., pp. 4568-4573, 2012.

[12] T. S. Rappaport, J. N. Murdock, and F. Gutierrez, "State of the art in 60-GHz

integrated circuits and systems for wireless communications," Proc. IEEE, vol. 99, no. 8, pp. 1390-1436, 2011.

[13] J. G. Andrews et al., "What will 5G be?" IEEE J. Sel. Areas Commun., vol. 32, no. 6, pp. 1065-1082, 2014.

[14] A. R. Khan, M. Othman, S. A. Madani, and S. U. Khan, "A survey of mobile cloud computing application models," IEEE Commun. Surveys Tuts., vol.16, no. 1, pp. 393-413, 2014.

[15] M. Peng, Y. Li, Z. Zhao, and C. Wang, "System architecture and key technologies for 5G heterogeneous cloud radio access networks," IEEE Netw., vol. 29, no. 2, pp. 6-14, 2015.

[16] M. Peng, X. Xie, Q. Hu, J. Zhang, and H. V. Poor, "Contract-based interference coordination in heterogeneous cloud radio access networks," IEEE J. Sel. Areas. Commun., vol. 33, no. 6, pp. 1140-1153, 2015.

[17] M. Peng, C. Wang, V. Lau, and H. V. Poor, "Fronthaul-constrained cloud radio access networks: Insights and challenges," IEEE Wireless Commun., vol. 22, no. 2, pp. 152–160, 2015.

[18] M. Peng, C. I, C. Tan, and C. Huang, "Recent advances in cloud radio access networks," IEEE Access, vol. 2, pp. 1683-1685, 2014.

[19] M. Peng, Y. Li, J. Jiang, J. Li, and C. Wang, "Heterogeneous cloud radio access networks: A new perspective for enhancing spectral and energy efficiencies," IEEE Wirless Commun. vol. 21, no. 6, pp. 126-135, 2014.

[20] N. Saquib, E. Hossain, L. B. Le, and D. I. Kim, "Interference management in ofdma femtocell networks: issues and approaches," IEEE Wireless Communications, vol. 19, no. 3, pp. 86-95, 2012.

[21] I. Hwang, B. Song, and S. Soliman, "A holistic view on hyperdense heterogeneous and small cell networks," IEEE Commun. Mag., vol. 51, no. 6, pp. 20-27, 2013.

[22] S. Yunas, T. Isotalo, J. Niemela, and M. Val ama, "Impact of macrocellular network densification on the capacity, energy and cost efficiency in dense urban environment," International Journal of Wireless Mobile Networks (IJWMN), vol. 5, no. 5, pp. 99-118, 2013.

[23] D. Lopez-Perez, I. Guvenc, and X. Chu, "Mobility enhancements for heterogeneous networks through interference coordination," Wireless Communications and Networking Conference Workshops (WCNCW), IEEE, 2012.

[24] D. Lopez-Perez, I. Guvenc, and X. C., "Mobility management challenges in 3gpp heterogeneous networks," IEEE Commun. Mag., vol. 50, no. 12, pp. 70-78, 2012.

[25] A. Furno, D. Naboulsi, R. Stanica, and M. Fiore, "Mobile Demand Profiling for Cellular Cognitive Networking," IEEE Transactions on Mobile Computing, vol. 16, no. 3, pp. 772-786, 2016.

[26] A. Furno, R. Stanica and M. Fiore, "A comparative evaluation of urban fabric detection techniques based on mobile traffic data," in IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, IEEE, 2015.
[27] Ł. Budzisz et al., "Dynamic Resource Provisioning for Energy Efficiency in Wireless Access Networks: A Survey and an Outlook," IEEE Communications Surveys Tutorials, vol. 16, no. 4, pp. 2259-2285, 2014.
[28] W. Song, H. Jiang, and W. Zhuang, "Performance analysis of the WLAN-first scheme in cellular/WLAN interworking," IEEE Transactions on Wireless Communications, vol. 6, no. 5, pp. 1932-1952, 2007.
[29] S. Lee, K. Sriram, K. Kim, Y. H. Kim, and N. Golmie, "Vertical handoff decision algorithms for providing optimized performance in heterogeneous wireless networks," IEEE T. Vehicular Technology, vol. 58, no. 2, pp. 865-881, 2009.
[30] C. Vitale and V. Mancuso, "Energy Efficiency in Mixed Access Networks," in Proceedings of the 19th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems, 2016.
[31] F. Bonomi, "Cloud and Fog Computing: Trade-offs and applications," in Intl. Symp. Comp. Architecture (ISCA), EON Workshop, 2011.
[32] H. Li, G. Shou, Y. Hu, and Z. Guo, "Mobile Edge Computing: Progress and Challenges," in 4th IEEE International Conference on Mobile Cloud Computing, Services, and Engineering (Mobile Cloud), IEEE, 2016.
[33] D. Kreutz, F. M. V. Ramos, P. E. Veríssimo, C. E. Rothenberg, S. Azodolmolky, and S. Uhlig, "SoftwareDefined Networking: A Comprehensive Survey," in Proceedings of the IEEE, vol. 103, no. 1, pp. 14-76, 2015.
[34] A. Chorti, S.M. Perlaza, Z. Han, and H.V. Poor, "On the resilience of wireless multiuser networks to passive and active eavesdroppers," in IEEE Journal on Selected Areas in Communications, vol. 31, no. 9, pp. 1850-1863, 2013.
[35] Y. Mo, T.H.J. Kim, K. Brancik, D. Dickinson, H. Lee, A. Perrig, and B. Sinopoli, "Cyber–physical security of a smart grid infrastructure," in proceedings of the IEEE, vol. 100, no. 1, pp. 195-209, 2012.
[36] K. Govindan, and P. Mohapatra, "Trust computations and trust dynamics in mobile adhoc networks: a survey," in IEEE Communications Surveys and Tutorials, vol. 14, no. 2, pp. 279-298, 2012.
[37] R.P. Jover, "Security attacks against the availability of LTE mobility networks: Overview and research directions," 16th International Symposium on Wireless Personal Multimedia Communications (WPMC), IEEE, 2013.
[38] M. Chen, Y. Qian, S. Mao, W. Tang, and X. Yang, "Software-defined mobile networks security," in Mobile Networks and Applications, pp. 729-743, 2016.
[39] O. Altintas, K. Seki, H. Kremo, M. Matsumoto, R. Onishi, H. Tanaka, "Vehicles as Information Hubs During Disasters: Glueing Wi-Fi to TV White Space to Cellular Networks," in IEEE Intelligent Transportation Systems Magazine,

vol. 6, no. 1, pp. 68-71, 2014.

[40] O. Altintas, F. Dressler, F. Hagenauer, M. Matsumoto, M. Sepulcre, and C. Sommer, "Making Cars a Main ICT Resource in Smart Cities," 2015 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), IEEE, 2015.

[41] A. Janecek, D. Valeri o, K. A. Hummel, F. Ricciato, and H. Hlavacs, "The Cellular Network as a Sensor: From Mobile Phone Data to Real-time Road Traffic Monitoring," IEEE Transactions on Intelligent Transportation Systems, vol. 16, no. 5, pp. 2551-2572, 2015.

[42] N. Caceres, L. M. Romero, F. G. Benitez, and J. M. del Castillo, "Traffic flow estimation models using cellular phone data," IEEE Transactions on Intelligent Transportation Systems, vol. 13, no. 3, pp. 1430-1441, 2012.

[43] F. Mohammadnia, M. Fiore, M. Ajmone Marsan, "Adaptive densification of mobile networks: Exploring correlations in vehicular and telecom traffic," 17th Annual Mediterranean Ad Hoc Networking Workshop (Med-Hoc-Net), IEEE, 2018.

[44] N. M. Sadek, H. H. Halawa, R. M. Daoud, and H. H. Amer, "A Robust Multi-RAT VANET/LTE for Mixed Control Entertainment Traffic," Journal of Transportation Technologies, vol. 5, no. 2, pp. 113-121, 2015.

[45] N. Frangiadakis, D. Cmara, F. Filali, A. A. F. Loureiro, and N. Roussopoulos, "Virtual Access Points for Vehicular Networks," 1st International Conference on mobile Wireless middleware, Operating Systems, and Applications, 2008.

[46] F. Dressler, P. Handlex, and C. Sommer, "Towards a Vehicular Cloud Using Parked Vehicles as a Temporary Network and Storage Infrastructure," in WiMobCity14, 2014.

[47] F. Malandrino, C. E. Casetti, C.F. Chiasserini, C. Sommer, and F. Dressler, "Content Downloading in Vehicular Networks: Bringing Parked Cars Into the Picture," in 23rd PIMRC 2012, IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, IEEE, 2012.

[48] F. Malandrino, C. Casetti, C. F. Chiasserini, C. Sommer, and F. Dressler, "The Role of Parked Cars in Content Downloading for Vehicular Networks," IEEE Transactions on Vehicular Technology, vol. 63, no. 9, pp. 4606-4617, 2014.

[49] V. Sciancalepore, V. Mancuso, A. Banchs, S. Zaks, and A. Capone, "Enhanced content update dissemination through D2D in 5G cellularnetworks," IEEE Transactions on Wireless Communications, vol. 15, no. 11, pp. 7517-7530, 2016.

[50] M. Ajmone Marsan, F. Mohammadnia, C. Vitale, M. Fiore, V. Mancuso, "Towards mobile radio access infrastructures for mobile users," Ad Hoc Networks, vol. 89, pp. 204-217, 2019.

[51] J. Allen, and F. Chevalier, "Report for Vodafone—Mobile backhaul market: Phase 1 report," Analysis Mason, Cambridge, 2014.

[52] T. E. Bogale, and L. B. Le, "Massive MIMO and mmWave for 5G wireless HetNet: Potential benefits and challenges," in Proc. IEEE Veh. Technol. Mag., vol. 11, no. 1, pp. 64-75, 2016.
[53] D. Bojic et al., "Advanced wireless and optical technologies for small-cell mobile backhaul with dynamic software-defined management," IEEE Commun. Mag., vol. 51, no. 9, pp. 86-93, 2013.

This Ph.D. thesis has been typeset by means of the TEX-system facilities. The typesetting engine was pdfLATEX. The document class was `toptesi`, by Claudio Beccari, with option `tipotesi=scudo`. This class is available in every up-to-date and complete TEX-system installation.