

Towards mobile radio access infrastructures for mobile users

Marco Ajmone Marsan^{a,b,c,*}, Foroogh Mohammadnia^a, Christian Vitale^a, Marco Fiore^b,
Vincenzo Mancuso^c

^a *Electronics and Telecommunications Department, Politecnico di Torino, Italy*

^b *CNR-IEIIT, Torino, Italy*

^c *IMDEA Networks Institute, Leganes, Spain*



ARTICLE INFO

Article history:

Received 15 December 2018

Revised 13 March 2019

Accepted 15 March 2019

Available online 19 March 2019

Keywords:

Dense radio access network

Moving base station

Small cells

ABSTRACT

This paper provides a first investigation of twice-mobile networks, i.e., cellular networks where both the end users and (part of) the radio access network infrastructure are mobile. Twice-mobile networks are based on an opportunistic, dense, crowdsourced, random deployment of mobile small cell base stations carried by vehicles, and on millimetre-wave backhaul connections between the mobile small cell base stations and the fixed network elements. Thanks to the fact that vehicles carrying mobile small cell base stations roam coherently with mobile subscribers, twice-mobile networks provide adaptive broadband wireless capacity where and when users need it, thus avoiding the cost and intrinsic inefficiency of dense deployments of fixed small cell base stations. In this paper we investigate the achievable capacity under the twice-mobile network paradigm, using real-world telecom traffic and vehicle positions in two case studies in Milan, Italy. Our results show that, thanks to positive spatial correlations between mobile network demands and road traffic, mobile small cell base stations carried by vehicles ensure performance equivalent or better to that of a traditional deployment of fixed small cells, at significantly lower cost.

© 2019 Elsevier B.V. All rights reserved.

1. Introduction

Network densification is considered one of the most promising approaches to allow radio access networks (RANs) cope with the current dramatic increase in mobile data traffic, and with the emergence of environments that exhibit extreme user (or user equipment, UE) density, such as stadiums, shopping malls and conferences. RAN densification requires the deployment of large numbers of small cells (SCs) in those areas where the presence of UEs and the traffic they generate are high for some significant portion of time. The densities of base stations (BSs) that are today forecasted in the 5G vision documents are extreme: up to hundreds or even thousands per km² [1]. This entails large investments for the BSs installation and operation: the CAPEX (capital expenditures) for 5G rollout in Europe is estimated at over 50 billion euros per year [2], and it shall be largely borne by telecommunication businesses.

The number of UEs and the level of traffic they generate exhibit remarkable spatial and temporal fluctuations [3,4]. Users normally move from home to work in the morning of working days, and this makes business districts crowded during working hours.

In this period, mobile network operators (MNOs) need the capacity granted by dense small cell RAN deployments to cover business areas. However, after work, users move out of their offices, so that the RAN capacity necessary in a business district suddenly drops, and many of the installed SCs become redundant. The opposite is true for residential areas, where capacity is required in the evening, but typically not during working hours. In addition, RAN densification at different locations may be necessary in presence of traffic jams emerging during commuting times.

The explosion of mobile network traffic, coupled with the mobility of users and terminals, exacerbates such a situation. For instance, the increasingly common habit of using smartphones during crowded events, such as music concerts, sports events, or large public gatherings, gives rise to time-varying hotspots where additional network capacity is required. When considered jointly, all these situations make the traditional approach of dimensioning for peak hour traffic extremely costly: dense deployments of SCs would be needed in both business and residential districts, and extremely dense coverages should be present in, e.g., stadiums, yet all such resources would be only used for a small fraction of time. Overall, this would lead to low resource utilization, hence low return on investment, for long periods of time.

In order to improve this situation, it would be very beneficial to have a large number of SCs in business districts during the day only, and in residential districts during the evening only. One

* Corresponding author at: DET, Politecnico di Torino, Corso Duca degli Abruzzi 24, Torino 10129, Italy.

E-mail address: ajmone@polito.it (M. Ajmone Marsan).

possibility to achieve this is to deploy a dense SC coverage in all areas, switching them on and off as needed. This can bring savings in OPEX (operational expenditures), especially in relation to energy consumption [5], but does not alleviate the CAPEX entailed by cell deployment. Another possibility is to move SCs from business areas to residential areas and back, so as to have the capacity of those cells where and when needed. For instance, MNOs already use small numbers of truck-mounted BSs for the quick provision of service in areas where service is not otherwise available, or where additional capacity is temporarily needed [6], and several papers have proposed drones to support communications in disaster areas [7–9]. While these solutions may be suitable in specific use cases, they do not scale to, e.g., large metropolitan areas, where MNOs would need to run and coordinate fleets of hundreds of trucks or drones. In this paper we look at a completely new approach to implement mobile BSs that has substantial potential for scalability. Namely, we propose to use a suitably chosen set of privately-owned ground vehicles to carry (part of) the SCs that implement the dense RAN deployment and includes *mobile* small cell base station (MoBS) devices.

Our rationale is that vehicles for private ground transport have the very desirable property of moving in accordance with the mobile network subscribers. Thus, in a business district during working hours we normally have both many UEs and many vehicles. If a large enough fraction of those vehicles carries a SC, 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. In addition, situations like traffic jam areas during rush traffic hours or large events in proximity, e.g., of a stadium, will be similarly addressed in terms of network capacity provisioning.

A vehicle-based SC deployment yields multiple advantages over traditional network densification approaches that leverage fixed SCs. We summarize them as follows.

- Network densification is achieved in those areas where several connected cars are present. Given the positive correlation between the number of cars and the number of UEs in a given area, network densification moves with the peaks of data traffic demand, thus drastically increasing the efficiency of SCs. We measured such a correlation on recent real-world data, as reported in [10] and summarized later in this paper.
- Huge capacity increases are possible in dense urban areas, even with a tiny fraction of cars carrying SCs. We show that capacity gains over 100% are possible with just 1% of vehicles providing SC support.
- RAN densification is obtained very rapidly and with very limited cost for the MNO, since SCs are carried by vehicles, and are used by the MNO with a small cell as a service (SCaaS) paradigm, with no installation cost.
- Vehicles can provide, in addition to communication resources, also computing power and storage that can be exploited for the provision of extremely low latency services.
- The vehicular RAN densification solution is compatible with dedicated vehicles carrying SCs, like autonomous cars or drones that can be strategically positioned. Such vehicles would allow augmenting the deployment in a planned fashion.
- Groups of SCs can be interconnected to provide local service in areas where the cellular network infrastructure is (temporarily) not operational, e.g., following some natural disaster.
- SCs can very effectively support delay-tolerant Internet of Things (IoT) applications like smart meter reading, by downloading data while they pass by smart objects so that very little power is needed for data transfers.

Of course, the dense small cell RAN layouts resulting from vehicle-mounted SCs have quite different properties with respect

to the dense small cell RANs carefully planned by network operators, and this poses a number of challenging research issues. First of all, while planners of BS deployments suitably select BS positions so as to avoid interference and maximize efficiency, drivers travel and park their vehicles without care for network planning needs. This leads to unplanned BS deployments that call for new adaptive network management solutions. In addition, the movement of SCs, coupled with that of UEs, accelerates the overall RAN topology dynamics, requiring rapid rearrangements in the network: as an example, handovers are likely to become much more frequent, and the backhaul connectivity may rapidly change.

In this paper, we first discuss the correlation between number of vehicles and telecom traffic in a given area, by summarizing the main results in [10]. Second, we illustrate an approach for the computation of the throughput achieved with SCs on vehicles, and we show that performance gains with respect to a deployment of fixed macro BSs can be significant, comparable to the throughput gain achieved with a planned deployment of fixed SCs in equal number. We then look at the optimization of short- and long-term fairness in the allocation of throughput to end users by means of a time scheduler that alternates among configurations of active SCs. In practice, we exploit an approach whose principle is analogous to the one used for counteracting inter-cell interference by means of almost blank subframes (ABS) [11,12]. Since the computation of the maximum throughput and of the throughput that provides the optimal fairness requires the analysis of all possible configurations of active SCs, whose number is exponential in the number of SCs, we will first present an exact approach, and then a simple novel heuristic with quadratic complexity in the number of SCs. The heuristic will be shown to provide results very close to the optimum.

The rest of this paper is organized as follows. Section 2 illustrates in some detail the concept of twice-mobile networking, i.e., networks in which both the end users and (part of) the infrastructure are mobile. Section 3 overviews the related work. Section 4 discusses the correlation between number of vehicles and telecom traffic in a given area. Section 5 presents an approach for the computation of the throughput achieved with SCs on vehicles, together with our optimization for improving end-user fairness. Section 6 presents numerical results, and finally Section 7 concludes the paper.

2. Twice mobile networking

We name our approach twice-mobile networking (or TMN for short), and we call MoBS the mobile small cell base station. We argue that the impact of the TMN approach on mobile networks can be enormous. TMN may bring a revolution in wireless network design, and can redefine the possibilities of adaptive capacity provisioning. At the same time, TMN will allow a distribution of the CAPEX for the provision of wireless capacity. Indeed, similar to what happened in smart grids where solar panels belonging to end users generate electricity that is available for all, in TMN MoBSs owned by end users will provide capacity managed by the MNO for the benefit of all network end users.

A schematic picture of a portion of a TMN is shown in Fig. 1. The urban area is covered by standard cells, defined by the deployment of fixed macro/micro BSs, as we know today. In addition to this traditional coverage, vehicles carrying MoBSs, either parked or driving, define a temporary dense coverage of those portions of the urban area where the UE concentration is higher, through small cells that overlap with macro/micro cells, and provide the necessary additional capacity where needed, when needed. Each BS, either a fixed or vehicle-mounted one, is connected to the core network through a backhaul link, which can be either wired or wireless for fixed BSs, but must (obviously) be wireless for

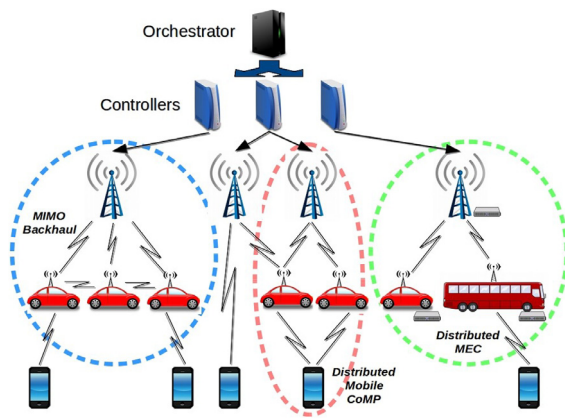


Fig. 1. The twice mobile networking concept.

MoBSs. We assume that the connection between a MoBS and the fixed network is implemented with a millimetre-wave link, so as to have sufficient capacity for the support of the traffic generated by all end users connected to MoBSs, and to avoid interference with the lower frequency channels connecting either fixed BSs or MoBSs to UEs.

As shown in the picture, a millimetre-wave link is also available between MoBSs. This is extremely important to allow the creation of a backhaul network, so that even if the direct link from a MoBS to the fixed network is not available due to obstacles, other backhaul opportunities are possible through neighbor MoBSs. In addition, the MoBS-MoBS link will permit the creation of standalone networks of MoBSs where the fixed network infrastructure is not working, like for example in disaster areas.

Finally, we observe in Fig. 1 the presence of an orchestrator and several controllers, that are instrumental to the real-time management of the extremely complex TMN environment. MoBSs must be switched on and off according to their utility in service provisioning, and their operation parameters (e.g., their output power) must be adapted so as to maximize the system performance.

In addition to communication resources, MoBSs can also provide computing power that can be exploited with fog [13] and mobile edge computing [14] paradigms to achieve the necessary computation with the minimum latency. Controllers and orchestrator are also charged with the management of these computing resources.

The number of critical research questions raised by the TMN vision is fairly large. Next, we briefly discuss the most obvious technical concerns and opportunities for innovation linked to TMN, although we acknowledge that our coverage is not exhaustive and many lesser aspects would also deserve investigation.

2.1. 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 today's 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 [15,16]. These approaches are being studied in the context of dense small cell RANs by ARTEMIS, with their pCell (personal cell) technology [17,18]. 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 [19], with the advantage of also reducing the handover rate. Interestingly, 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. Indeed, a static ABS approach is not suitable, and advanced stochastic ABS configurations can be leveraged instead, as the dynamic distributed scheme proposed in [20].

2.2. Uplink connections

The study of the downlink performance is simplified by the lack of contention in the access to wireless resources. When the uplink must be considered, it is necessary to account for the signalling between UEs and MoBSs 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 TMN approach, where MoBSs collect requests of medium size groups of users, alleviating contention on fixed BSs.

2.3. Wireless backhaul implementation

The selection of the technology for the implementation of the wireless backhaul link is possibly one of the most delicate choices for the TMN architecture, due to high capacity requirements, and to MoBSs mobility. The possibility of operating MoBSs as non-transparent layer-3 relays [21–23], and connecting them to the corresponding macro/micro BS through the X2 and S1 interfaces [24,25] is not likely to provide sufficient capacity. 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 millimetre-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 millimetre-wave links in urban environments can be problematic, due to both distance and obstacles. It is thus very important to study for what fraction of time a (possibly multihop) millimetre-wave connection from the MoBSs to the fixed network can be active. This would allow computing the effective data transfer capabilities of the link, the estimation of the latency introduced due to interruptions, and the dimensioning of the storage that must be associated to MoBSs to survive periods of absence of connectivity.

2.4. Use of SDN and NFV together with ML

The extreme complexity of the TMN system, with MoBSs (equipped with CPU and memory for data processing and storage) moving around the service area, over which also end users roam, requires a very careful real-time management of network operations. Management operations include: the (de)activation of MoBSs so as to optimize throughput and fairness, the control of UE associations to BSs or MoBSs, the management of the millimetre-wave backhaul links, the control of handovers, the dispatch of autonomous vehicles to the locations where they are needed, etc. This cannot be obtained with a human in the loop, and must be completely automated. On the one hand, artificial intelligence algorithms are necessary to infer from past observations of the system the control rules for future system behaviour [27]. On the other

hand, suitable SDN (Software Defined Networking) solutions need to be in place, so as to enable the real-time orchestration of network resources, as well as a highly flexible adaptation of network operating parameters – especially if cooperative techniques are to be implemented [26].

In addition, the time-varying nature of the RAN layout calls for the virtualization of a number of network functions, for instance those dedicated to the dynamic caching of content stored by MoBS. Similarly, the system naturally lends itself to implement services under fog and edge computing paradigms.

2.5. Impact of self-driving cars or drones

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. The same development may also happen for private electrical cars, as in the plans of Tesla Motors [28]. The possibility of moving cars as desired in the periods when they are not hired is an extremely interesting feature for the effective implementation of the TMN concept, since it allows for the reduction of unpredictability in the system, and opens the way to the optimization of the spatial distribution of MoBSs. This adds an important degree of freedom in the management of the MoBSs layout, and can yield significant performance gains. 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 centres. Similar results can also be achieved by using drones to fill spatial and/or temporal bandwidth gaps due to occasional mismatches between traffic demand and availability of MoBSs.

2.6. Mobility and handovers

A careful assessment of the TMN concept requires investigating the operation of the system under mobility of both MoBSs and UEs. In particular, mobility implies handovers of UEs from one MoBS to another, or from a MoBS to a fixed BS, or from a fixed BS to a MoBS. In addition, it is necessary to investigate the activation/deactivation of MoBSs by looking at their positions relative to other MoBSs and fixed BSs, depending on the interference that is generated and the throughput that is gained. The impact of handovers can be mitigated by the simultaneous connection of UEs to multiple BSs, for example with approaches similar to CoMP.

2.7. Exploitation for IoT

The many emerging IoT applications are expected to either exploit specific technologies, like LORA [29], or the narrowband version of 5G [30]. However, TMN could provide a very interesting alternative to these approaches, since the loose requirements on latency of some IoT applications (such as smart meter reading) combined with the requirements of extremely low power can provide a very interesting match with MoBSs passing in close proximity of meters. MoBSs can opportunistically download data from smart objects, to be forwarded to the collection centre. How much power can be saved thanks to proximity, what delay can be tolerated when waiting for a MoBS passage, how much storage is required at the smart object and at the MoBS, and how to alert a smart meter for the presence of a MoBS in close proximity are all very relevant issues to be investigated.

2.8. Exploitation in disaster areas

Up to now, we have always assumed that MoBSs operate in conjunction with fixed BSs. However, MoBSs are also capable of direct communications with each other, and this allows them to

establish a localized opportunistic cellular network providing standard cellular communication services in an area where infrastructure is not present because of digital divide issues, or because of natural or human-caused disasters (earthquake, flooding, terrorist attack, etc.).

2.9. Business model issues

The possibility of providing broadband wireless connectivity on demand by exploiting the equipment of many private and public actors has the potential of opening a new market, similar to the now common energy market, where MoBSs provide bandwidth much like solar panels on rooftops provide energy. Owners of vehicles carrying MoBSs will play a role similar to prosumers in the smart grid. Therefore, the business model issue for the TMN concept is extremely interesting. Incentives should be offered to car owners to carry MoBSs, in order to motivate them to participate in a TMN. The extra cost for the purchase of an equipped car, or for the retrofitting of an older car with a MoBS calls for a compensation, which should come from the MNO or the tower operator that exploits the capacity of the temporary dense MoBSs layout.

3. Related work

Adaptivity has traditionally been considered a very desirable feature in networking, and some classes of networks have been based on adaptivity to the environment, in particular sensor networks [31,32], ad-hoc networks [33], delay tolerant networks [34], and networking paradigms based on opportunistic communications, such as Floating Content [35,36] or Hovering Information [37].

The increasing softwarization of telecommunications, combined with the coming of the age of artificial intelligence, are creating the conditions for extreme adaptivity of wireless (and wired) networks. It is now possible to dynamically reallocate and in some cases transfer resources so as to match users requests in terms of volume and performance. The specific case that we consider in this paper concerns adaptivity in the densification of cellular networks. The idea of adaptive cellular network densification is new, but on several occasions vehicles have already been considered as elements of a telecommunication network. Some studies in the domain of vehicular networks considered the possibility of integrating a Wi-Fi Access Point (AP) in cars, and possibly also a cellular interface, so that a vehicle can provide connectivity to the surrounding cars. For example, Sadek et al. [38] proposed the use of mobile vehicular gateways that exploit Wi-Fi for vehicle-to-vehicle (V2V) communications and LTE for vehicle-to-infrastructure (V2I) communications. The study in [39] introduced the concept of Virtual APs, which allow extending the reach of roadside access points: vehicles that receive a message will store and rebroadcast it to non-covered areas. In [40–42], parked vehicles were exploited, in addition to roadside units, to improve the performance of video downloads and other services toward travelling cars. In [43], mobile nodes were used for the collection of data from sensors.

These approaches are different from TMN, which leverages vehicles as a support for small cell BSs of a dense RAN. TMN allows the seamless integration of MoBSs with the normal macro- and micro-BSs at fixed locations, so that the MoBSs provide additional capacity, when and where needed, to the end users of an otherwise traditional RAN. End-user terminals can thus freely and transparently transfer their services between a traditional fixed BS and a MoBS. TMN generalizes the concept of early deployments of SC BSs onboard public transport vehicles to serve passengers in ultra-dense urban environments [44–47], and that of using vehicle-carried SC BSs in public safety networks [48].

A concept similar to TMN has been briefly hinted at in recent works [49,50], where the authors mention the possible use of 4G or 5G small cells within moving and parked cars, or on drones, so as to serve users both indoor and outdoor. However, no results in this direction have been published so far.

4. Evidence of correlation between telecom and vehicular traffic

In a recent paper [10], we looked at the correlation between vehicular traffic and telecom traffic, and then estimated the average distance from UEs to MoBSs in an urban area. In our analysis we used a dataset that contains values of telecom and vehicular traffic during 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 rectangles, as shown in Fig. 2. The rectangles have minimum size of about 220 m by 330 m, but sides can be 2, 4, 8 times longer, as can be seen on the map. The time granularity of the dataset is 15 min (96 intervals per day). For each time interval, the dataset reports the total amount of mobile network data traffic (uplink and downlink voice and data, no signaling) recorded in every rectangle by a leading Italian MNO, as well as the GPS positions of active (i.e., not parked) vehicles tracked by InfoBlu, a large Italian infomobility services provider (about 4 million vehicles, on a national scale). The total number of time intervals is 61 times $96 = 5,856$. Considering the number of area rectangles for which data are reported separately, we have a total of 3,373,056 data points.

It is important to remark that the penetration rates of the two measurement technologies are quite different. While the telecom traffic in the dataset corresponds to more than one third of the overall MNO demand, the fraction of the monitored vehicles is less than 1%, as it mainly encompasses probe vehicles and commercial fleets monitored by InfoBlu. These percentages are consistent with the goal of our investigation: we can expect a small percentage of vehicles to act as MoBS carriers (at least in an initial phase), and would like them to cover the largest possible fraction of the total telecom traffic.

We start by showing, as an example, the average telecom and vehicular daily traffic patterns in a rectangle covering the main railway station in Milan (namely, “Milano Centrale” railway station). Fig. 3 in plots (a) and (b) shows the average telecom and

vehicular traffic patterns for all days of the week for the selected rectangle (telecom traffic in the left plot and vehicular traffic in the right plot). First of all, while in telecom traffic clear behaviors exist, vehicular traffic (at our penetration level) is highly irregular. In the considered railway station area, telecom traffic is very low at night, then sharply increases around 8 A.M. in working days, and remains close to its peak value from 9 A.M. to 9 P.M., then sharply decreases. The peak of telecom traffic in weekends is about one third of that in working days. Vehicular traffic is very bursty, peaking in the afternoon and evening. A qualitatively similar behaviour is observed in a rectangle covering a business area, as shown in Fig. 3, plots (c) and (d).

The inspection of the raw data about telecom and vehicular traffic for all rectangles and time intervals does not provide many clues to justify correlation; rather, it reveals that telecom traffic dynamics vary widely in space and time, and vehicular traffic is generally much more bursty than telecom traffic. This latter effect is mainly due to the reduced fraction of vehicles carrying MoBSs that will characterize early TMN deployments.

The results about correlation from a spatial perspective are summarized by the scatterplots in Fig. 4. The abscissa refers to the telecom traffic value, and the ordinate reports the vehicular traffic value. Each dot corresponds to one rectangle over the city of Milan (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). The figure outlines correlations emerging across the geographical area. Plots also show the least square linear regression line.

The results above lead to the following conclusions: i) there exists a general positive correlation between vehicular and telecom traffic, but ii) the correlation is low, due to high variability of telecom traffic. A sensible question is then whether some hours and geographical areas show especially strong correlations, and are thus suitable periods and locations for MNOs to take advantage of the TMN paradigm. We first investigated which periods during the day show high correlation. Fig. 5 (left plot) shows the curves of the Pearson correlation coefficient versus time for the average working day and for the average weekend day. We observed that correlation is low overnight in working days and in the early morning of weekends. On the contrary, the correlation coefficient values are higher (in the range from 0.3 to 0.5) during high-demand periods, i.e., in working periods 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, TMN is most effective exactly when densification is needed, that is, in presence of mobile network traffic demand peaks.

Concerning the geographical dispersion of correlation, Fig. 5 (right plot) shows a heatmap where the color of each rectangle corresponds to the Pearson correlation coefficient computed over an average day (dark blue means low correlation values, while red means high correlation values). 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 rectangles, hence they follow more predictable patterns. On the other hand, and more importantly, those are also the rectangles where the demand for radio access network capacity is higher, hence densification is needed the most. We conclude that, in our reference scenario, TMN proves to be mainly exploitable not only when, but also where network densification is needed the most.

The next question concerns the distance between MoBSs and the UEs that can use the MoBSs to transmit and receive mobile network traffic. Since our dataset reports exact vehicle positions within each rectangle, but only the originating area (and not the

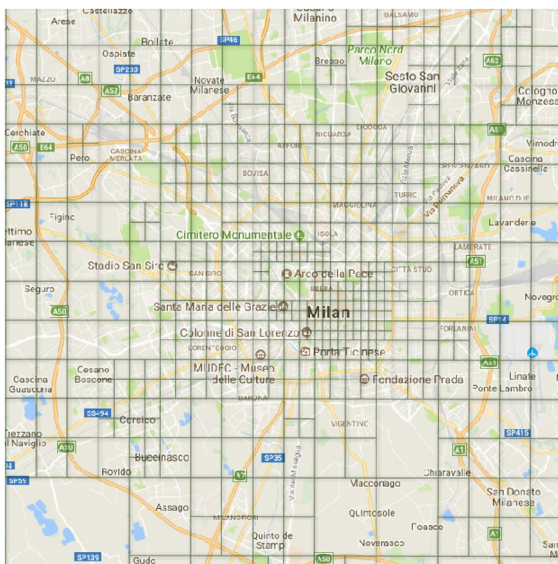


Fig. 2. The subdivision of the Milan area into 576 rectangles.

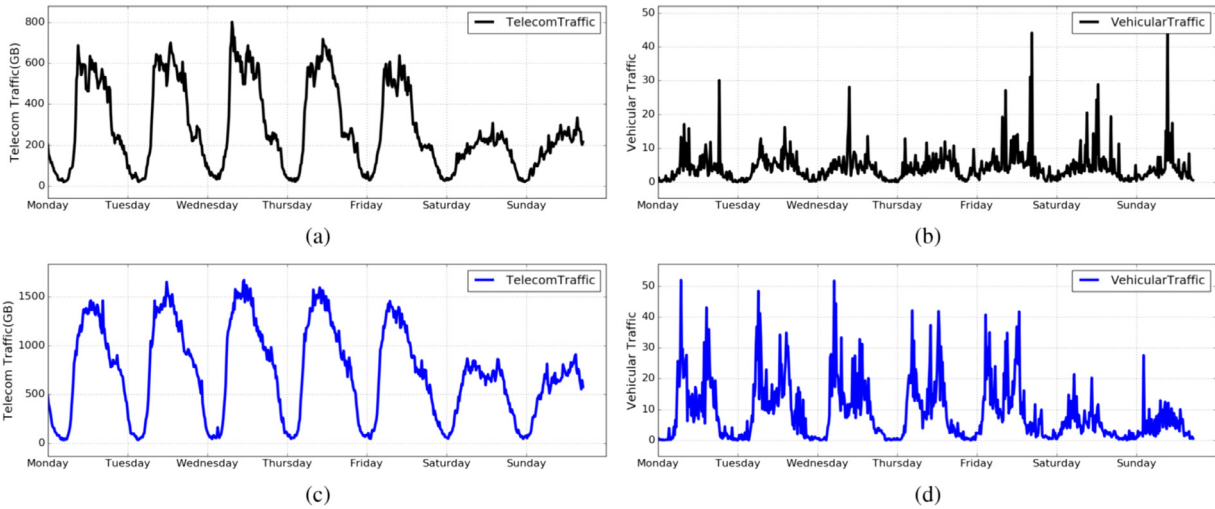


Fig. 3. Average telecom and vehicular traffic (left and right, in GB and number of cars, respectively) for all days of the week in selected rectangle areas. Plots (a) and (b) refer to a railway station, plots (c) and (d) to a business area.

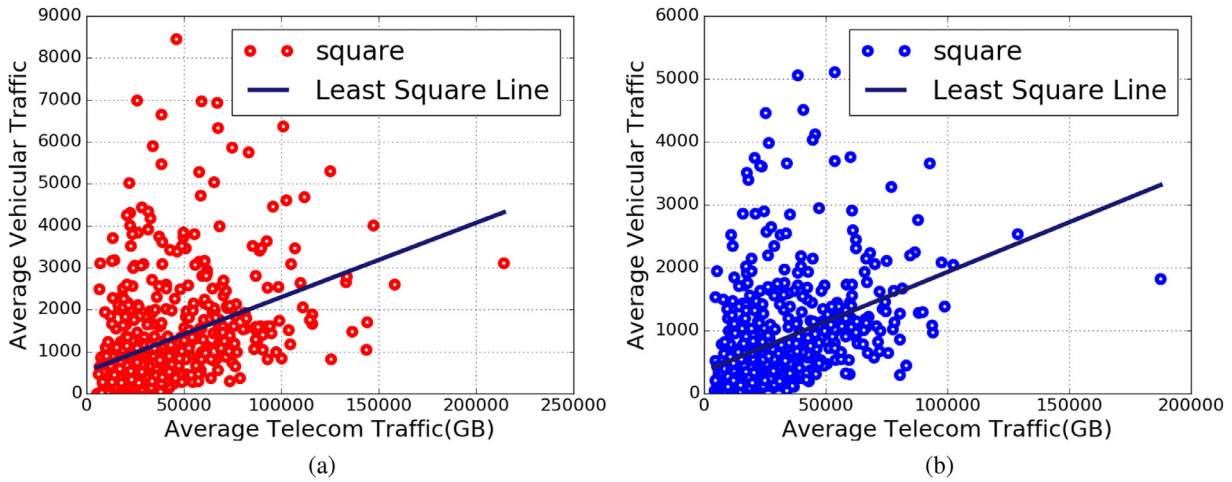


Fig. 4. Correlation between average telecom traffic (in GB) and average vehicular traffic (in number of cars). Plot (a) refers to 44 working days and plot (b) to 17 weekend days.

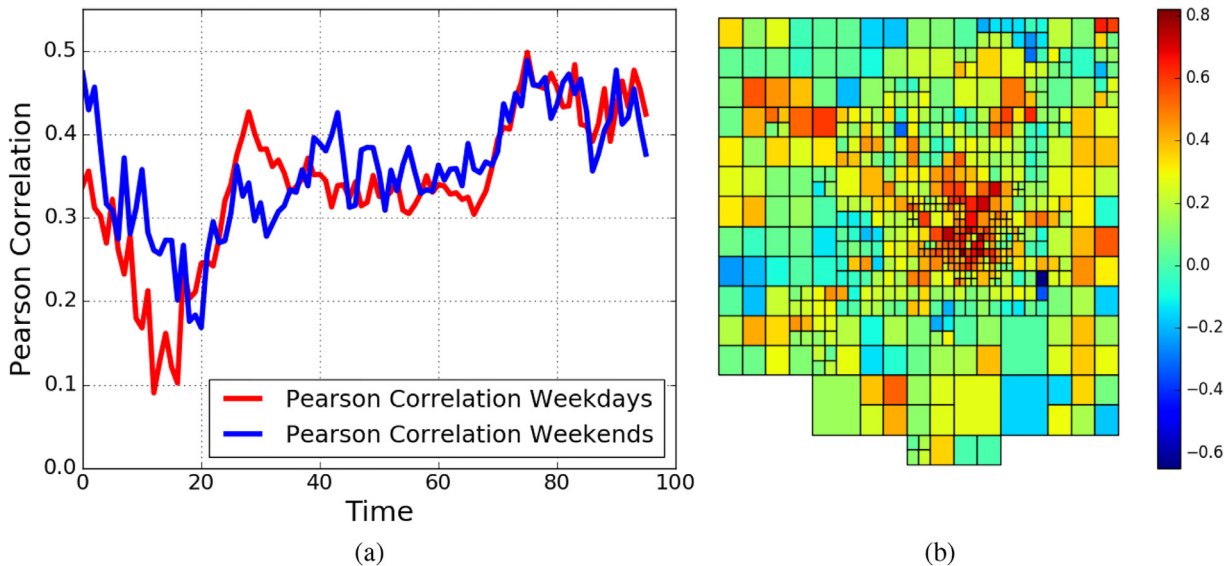


Fig. 5. Pearson correlation coefficient versus time interval between average telecom and vehicular traffic over all rectangles and over 44 working days and 17 weekend days in left plot; Pearson correlation coefficient heatmap between telecom and vehicular traffic for each Milano rectangle, averaged over 61 days in right plot.

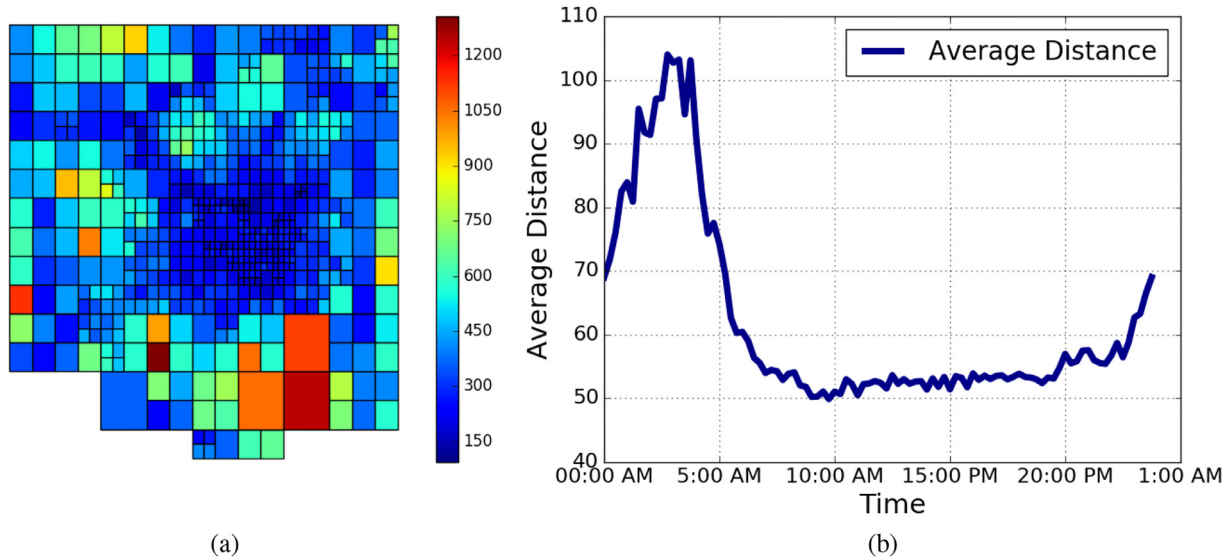


Fig. 6. Average distance in meters between vehicles and origins of telecom traffic for all rectangle areas over one day (March 2, 2015) in left plot; average distance in meters between vehicles and origins of telecom traffic versus time of day for 158 rectangles in the center of Milan averaged over 61 days in right plot.

exact end user terminal location) for telecom connections, we randomly place end user terminals within the area covered by each rectangle. 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 TMN, since it is a proxy for the connection data rate.

The heatmap in Fig. 6 (left plot) shows for all 576 rectangles that cover Milan the average distance between MoBSs and user terminals (dark blue means short distances, while red means long distances). The shortest average daily distances are found in the city center and are around 120m, while the longest average daily distances are observed in peripheral rectangles and grow up to around 1000m. MoBSs are hardly usable under such distances. However, when looking in more detail at what happens in the city center, the conclusion is different. Results in Fig. 6 (right plot) show the evolution of the average distance over time between end user terminals and the closest MoBS, in the 158 rectangles in the center of Milan averaged over all 61 days. Between 9 A.M. and 8 P.M. average distances are around 50m, so that in this portion of the city we can expect a large fraction of terminals to be within reach of a MoBS. These values nicely complement the correlation results, and confirm that TMNs have the potential to 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.

5. Throughput calculation and optimization

In this section, we summarize the results presented in [51] for the performance evaluation of systems with BSs that can be muted with ABS on a millisecond timescale and relay groups. That work is relevant to us since it studies how to orchestrate dynamic groups of users by avoiding interference. Specifically, we can see a MoBS as a relay station in the model described in [51], while UEs that connect to MoBSs are end users in that work. Therefore, we present the results in a TMN scenario and use the TMN terminology previously defined in this article rather than the one presented in [51].

Results from Vitale et al. [51] are here applied to the TMN case to evaluate analytically the downlink throughput performance achieved when MoBSs are in place. Such throughput characteriza-

tion allows us to assess the gain enabled by MoBSs, and to present an optimization that copes with the classical interference problem observed in dense scenarios. Here, we use the model to unveil how to select, at each given time, which MoBSs should be allowed to transmit (active MoBSs) and which ones should be muted instead, so to control interference. 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 (Section 5.1). Afterwards, we compute the downlink throughput obtained by a UE (Section 5.2) and, based on this result, we design an optimization that selects active MoBSs with the aim of maximizing overall throughput or user fairness (Section 5.3). Finally, given the complexity of the introduced optimization, we present for the first time an easy-to-deploy heuristic that can be used also when the number of MoBSs is large (Section 5.4).

5.1. Transmission efficiency

We consider a cellular access network with a set \mathcal{B} of interfering BSs. A BS belonging to \mathcal{B} is either a standard fixed BS or a MoBS. In the following, we consider short time slots where the location of UEs and MoBSs can be considered as fixed. Furthermore, we assume that users attach to BSs (either fixed ones or MoBSs) according to the strongest signal received. Transmission efficiency (i.e., the average number of bits per symbol transmitted), 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 Third Generation Partnership Project [52] 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:

$$\zeta_i(t) = \sum_{k \in \mathcal{M}} b_k [F_{\text{SINR}}^t(T_k^{\max}) - F_{\text{SINR}}^t(T_k^{\min})], \quad (1)$$

where \mathcal{M} is the set of MCSs, b_k is the number of bits per symbols for MCS k , (T_k^{\min}, T_k^{\max}) is the interval of the SINR for which MCS k is the best option, and F_{SINR}^t is the Cumulative Density Function (CDF) of the SINR at time t . In practice, (1) evaluates the transmission efficiency of UE i in time slot t from the value of b_k and the probability of UE i to use a specific MCS, given the experienced SINR distribution.

The CDF of the SINR depends on the radio propagation between the BSs and the UE. As pointed out in [53], 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.

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

$$F_{\text{SINR}}(x) = 1 - e^{-\lambda_S N x} \prod_{j=1}^k \frac{\lambda_{I_j}}{\lambda_{I_j} + x \lambda_S}. \quad (2)$$

Proof. The proof can be easily obtained from the following expression, in which $f(\cdot)$ is an exponential pdf:

$$\begin{aligned} F_{\text{SINR}}(x) &= \Pr \left\{ \frac{S}{N + \sum_{j=1}^J I_j} \leq x \right\} \\ &= \int_0^\infty \int_0^\infty \dots \int_0^\infty \Pr \left\{ S \leq x \left(N + \sum_{j=1}^J I_j \mid I_j = y_j \right) \right\} \\ &\quad \cdot \prod_{j=1}^J f_{I_j}(y_j) dy_j \quad \square \end{aligned}$$

The average received power levels can be computed with standard distance-based path loss models [54].

5.2. User throughput

We are now able to compute the average throughput of UE i in time slot t , $\Gamma_i(t)$, by multiplying $\zeta_i(t)$ obtained with (1) times the average number of symbols per second D_i available for i at the attached BS.

$$\Gamma_i(t) = D_i \zeta_i(t). \quad (3)$$

D_i mainly depends on the scheduler used by the BS. Like in [51], we assume that the Equal Time Scheduler (ETS) is in force, so that each UE receives on average the same amount of symbols, which yields:

$$D_i = \frac{K}{N_i}, \quad (4)$$

where K is the number of symbols per second available for data transmission, and N_i is the number of UEs attached to the BS serving UE i .

5.3. Throughput optimization

We now exploit the fact that muting and reactivating transmissions at a BS is today possible at millisecond time-scale with ABS, as discussed in [52], therefore causing no handovers. Thus, by alternating subsets of BSs $B \subseteq \mathcal{B}$ to transmit, we can control the interference in the system and the SINR distributions without having to continuously deal with BS attachment procedures. Hence, selecting the right subsets of BSs 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, while all other MoBSs are muted. For instance, groups of MoBSs can be scheduled sequentially so that B 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 considering as interfering BSs in (3) only the ones included in B :

$$\Gamma_i(t) = \sum_{B \subseteq \mathcal{B}} P_B \Gamma_i^B(t). \quad (5)$$

Obviously, if the BS b to which i is attached is not included in B , then $\Gamma_i^B(t) = 0$.

In order to maximize the average user fairness, as well as the overall system throughput, it is then sufficient to optimize over P_B . Specifically, we present a convex optimization that maximizes 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, select $P_B, \forall B \in \mathcal{B}$, so to:

$$\begin{aligned} &\text{maximize} && \frac{1}{N_t} \sum_i \log \left(\sum_{B \subseteq \mathcal{B}} P_B \Gamma_i^B(t) \right); \\ &\text{subject to:} && \sum_{B \subseteq \mathcal{B}} P_B = 1, \\ &&& P_B \in [0, 1], \quad \forall B \in \mathcal{B}. \end{aligned} \quad (6)$$

ABS 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.

5.4. Heuristic solution of the PFM problem

Problem PFM can be solved with standard optimization tools (we use the MATLAB optimization toolbox), provided the number of subsets 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 problem PFM becomes rapidly problematic.

However, by examining the optimal solutions in those cases that can be handled with the MATLAB optimization toolbox, 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 very few MoBSs. In practice, our optimization generally schedules all MoBSs and BSs together, except for some few MoBS heavily impaired by interference. In order to serve also the UEs of such unlucky MoBSs, the optimization reserves a portion of resources with very low interference, e.g., with no other MoBSs active. We thus decided to implement a simple heuristic approach to optimization that only considers the subsets B comprising:

- no active MoBSs (one subset)
- just one active MoBS (n subsets)
- two active MoBSs ($\binom{n}{2}$ subsets)
- all but two active MoBSs ($\binom{n}{2}$ subsets)
- all but one active MoBS (n subsets)
- all active MoBSs (one subset)

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

5.5. Implementation

The PFM problem that we just described naturally maps onto 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 TMN environment, which implies a significant consumption of resources (that we do not consider in this paper) 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.

6. Numerical results

In this section, we describe the settings in which we computed the throughput achievable in TMN scenarios, and we present the corresponding numerical results. We use the same dataset for telecom traffic and car sharing vehicle positions as in Section 4, although here we focus on narrower areas with peculiarly distinct characteristics. Specifically, our evaluation focuses on two representative case studies. The first one is the geographical area shown in Fig. 7, which comprises the central railway station in Milan, Italy. The second is an area of smaller size, covering the Polytechnic University of Milan (POLIMI).

6.1. Use case: downtown railway station area

The area in Fig. 7 is located in Milan downtown, and it is divided into nine rectangles in our traffic dataset. For each of those, we have data about the mobile network traffic and the number of probe vehicles, at 15-min time intervals and over 61 days in April 2015. The mobile network traffic refers to data connections and voice calls of one of the largest Italian MNOs, hence amounts 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 in the area, since they are only those managed by an Italian fleet management operator. We assume that each of the tracked vehicles is equipped with a MoBS, which is coherent with the expected limited 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. on April 15, 2015, the reported vehicles positions are as shown in Fig. 8 as blue/green markers, together with the positions of fixed macro BSs as red dots. We clearly see that fixed BS positions are

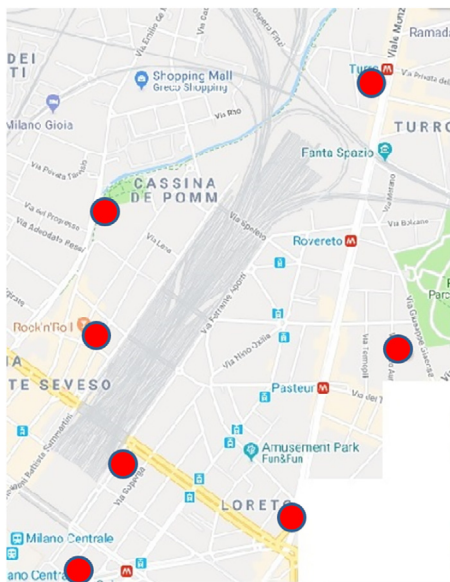


Fig. 7. The area of the central railway station in Milan with 7 fixed macro BSs (red dots). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

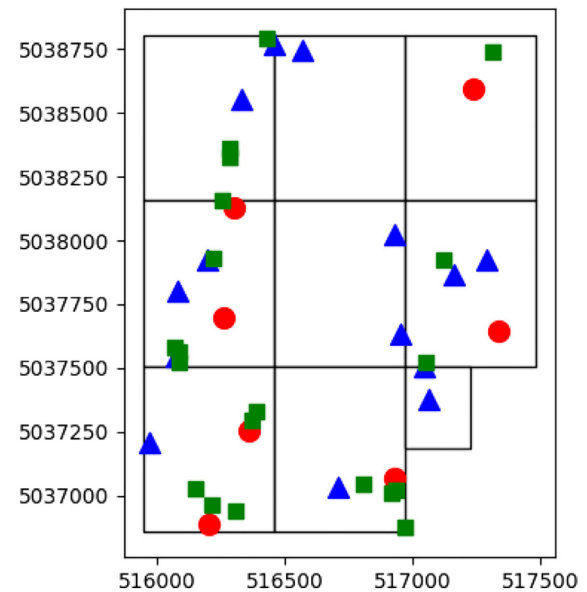


Fig. 8. The area of the central railway station in Milan with 7 macro BS (red dots) and 34 vehicle positions (blue triangles and green squares; the 20 green squares correspond to discarded MoBSs) at 8 A.M. on April 15, 2015. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

nice 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 their MoBSs. It is thus necessary to decide which MoBSs to use, so as to obtain acceptable interference (this is a responsibility of the Orchestrator in Fig. 1). In this paper, we preselect the set of useful MoBSs, that correspond to blue triangles in Fig. 8, with a simple greedy heuristic algorithm. Recalling that useful BSs are always scheduled by the PFM optimization (otherwise attached UEs experience no throughput), our algorithm aims at filtering available BSs in order to reduce the overall interference. Assuming that the transmission power of fixed macro BSs is 30 dBm, and the one of MoBSs is 20 dBm, we compute for each pair of MoBSs (or each pair comprising one MoBS and one fixed BS) the average received signal at both ends, following the path loss model in [54], and we discard the MoBS 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 discard the MoBS with the highest second value of pairwise generated interference. The process ends when no MoBS generates interference levels higher than -75 dBm. Discarded MoBSs are denoted by green squares in Fig. 8.

As regards UEs, since for each rectangle in the map of Fig. 8 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 GB of reported traffic. UE associations to fixed macro BSs or MoBSs follow a maximum received power criterion. Using the transmission powers mentioned above, the resulting UE associations are as reported in Fig. 9 in the case of fixed macro BS only, and in Fig. 10 in the case when both fixed macro BSs and MoBSs are active.

The total number of MoBSs, the number of active MoBSs, and the total data traffic for each time slot are reported in Fig. 11.

By applying the procedure outlined in Section 5, 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

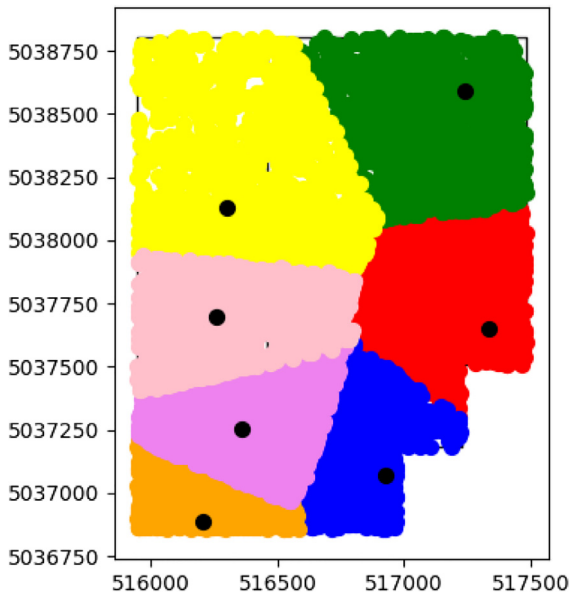


Fig. 9. UE associations to the 7 fixed macro BSs for the case of Fig. 8.

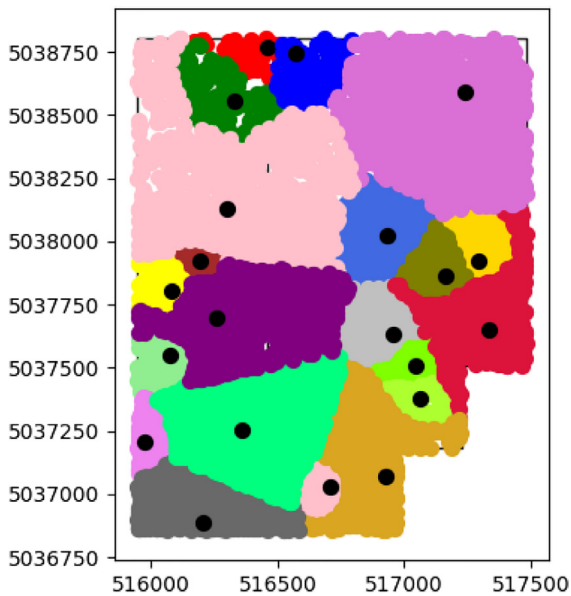


Fig. 10. UE associations to the 7 fixed macro BSs and the 14 useful MoBSs for the case of Fig. 8.

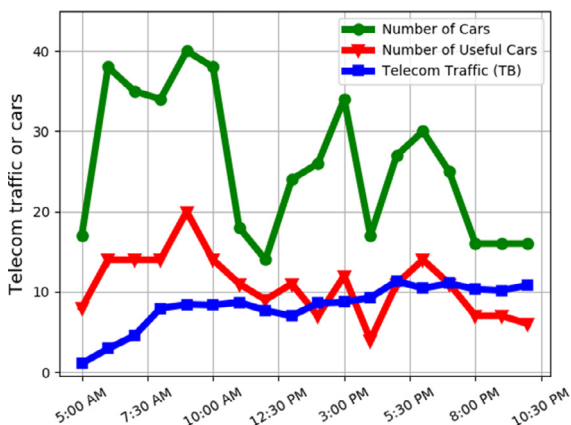


Fig. 11. Number of vehicles, number of useful MoBSs and data traffic (in TB) in the area of the central railway station in Milan on April 15, from 5 A.M. to 10 P.M.

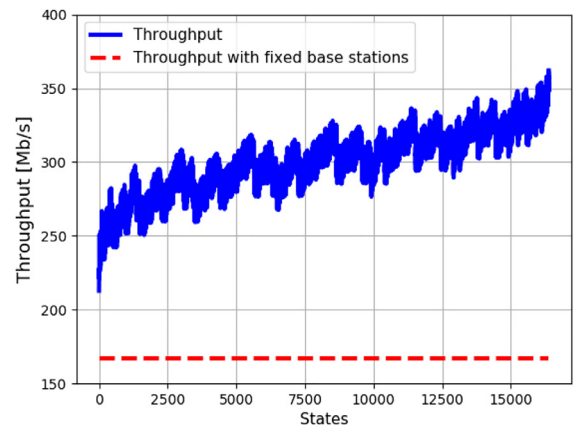


Fig. 12. Throughput of the 2^{14} MoBS configurations in the area of the central railway station in Milan on April 15, at 8 A.M.; configurations are ordered according to increasing number of active MoBSs. The red dashed horizontal line refers to the case of only fixed BS, with users attached to fixed BS only. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

procedure first computes the throughput for each user under each configuration of the system, that is, for each configuration of active/inactive (i.e., transmitting/muted) useful MoBSs.¹ 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, Fig. 12 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 BS, with users attached to fixed macro BS only.

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 the left plot of Fig. 13 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 cell coverage; (4) the throughput of the MoBS 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 MoBS configuration yielding maximum throughput exhibits the highest peaks, especially when the number of useful MoBSs is high.

¹ Note that we say that a MoBS is useful when it is not discarded by the orchestrator because it could generate excessive interference. Useful MoBSs can be active or muted in different time slots according to the chosen time schedule.

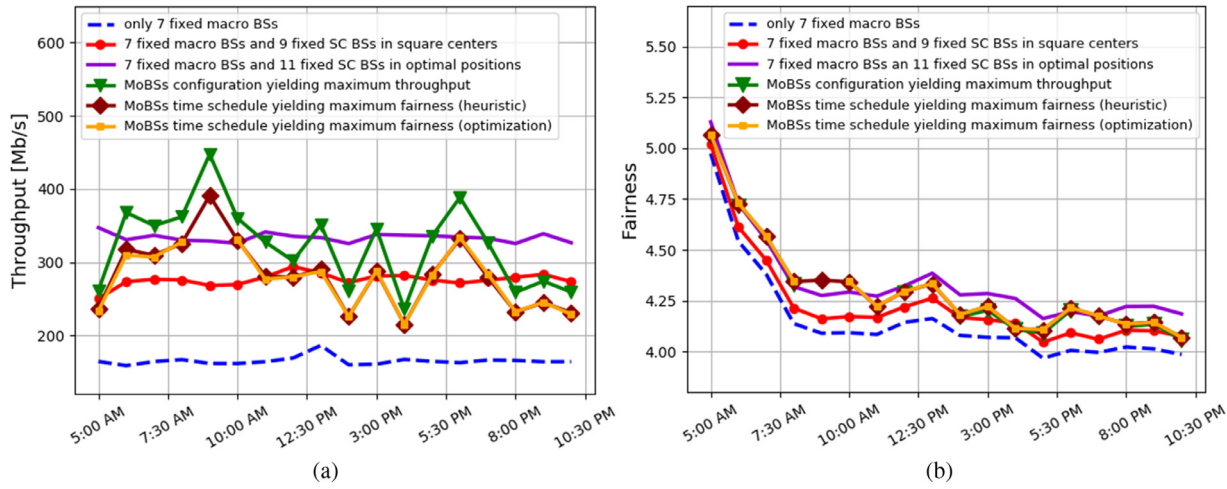


Fig. 13. Throughput in Mb/s (left) and fairness per user (right), in the area of the central railway station in Milan, on April 15 from 5 A.M. to 10 P.M., with: (i) only 7 fixed macro BSs; (ii) 7 fixed macro BSs and 9 fixed small cell BSs in rectangle centers; (iii) 7 fixed macro BSs and 11 fixed small 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).

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 MoBS 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 in the optimum fairness throughput computed using the heuristic and with the MATLAB optimization toolbox 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 Fig. 13.

It is remarkable to see that by using MoBSs we can achieve gains of up to about 150% with respect to using only macro BSs, when several vehicles are available.

The right plot of Fig. 13 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 cases of only fixed BSs, both when we consider only macro BSs, and when we consider the addition of 9 SCs in the center of rectangles.

The fairness values of the MoBS 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 $-\infty$. This proves that the use of the time scheduler is extremely important in order not to exclude some UEs from access to network resources.

6.2. 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. Fig. 14 reports, for the time interval from 8 A.M. to 8:15 A.M.

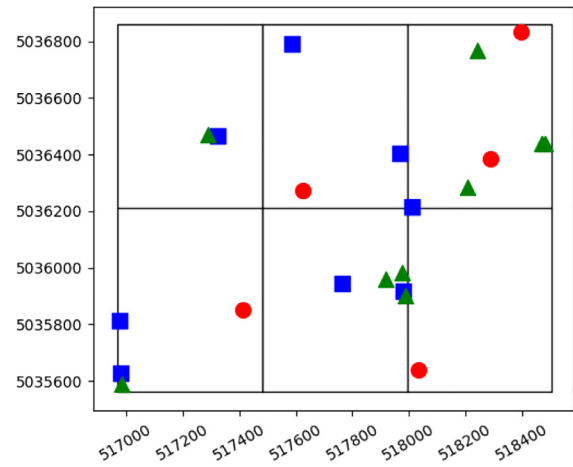


Fig. 14. The POLIMI area is covered by 5 macro BS (red circles). According to our dataset, we consider 17 vehicle positions (blue rectangles and green triangles; the 9 green triangles correspond to discarded MoBSs) at 8 A.M. on April 15, 2015. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

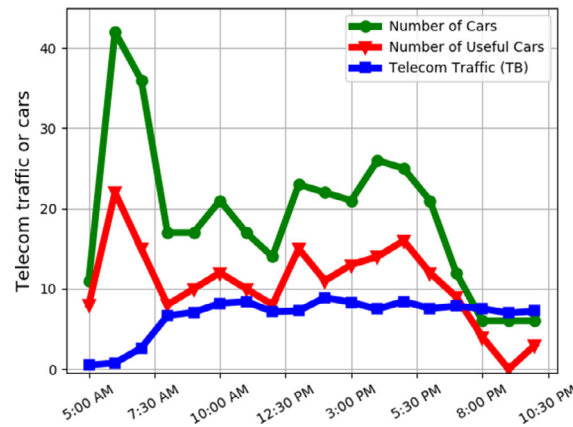


Fig. 15. Number of vehicles, number of useful MoBSs and data traffic (in TB) in the area of the Polytechnic University of Milan on April 15, from 5 A.M. to 10 P.M.

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 Fig. 15. We can see that

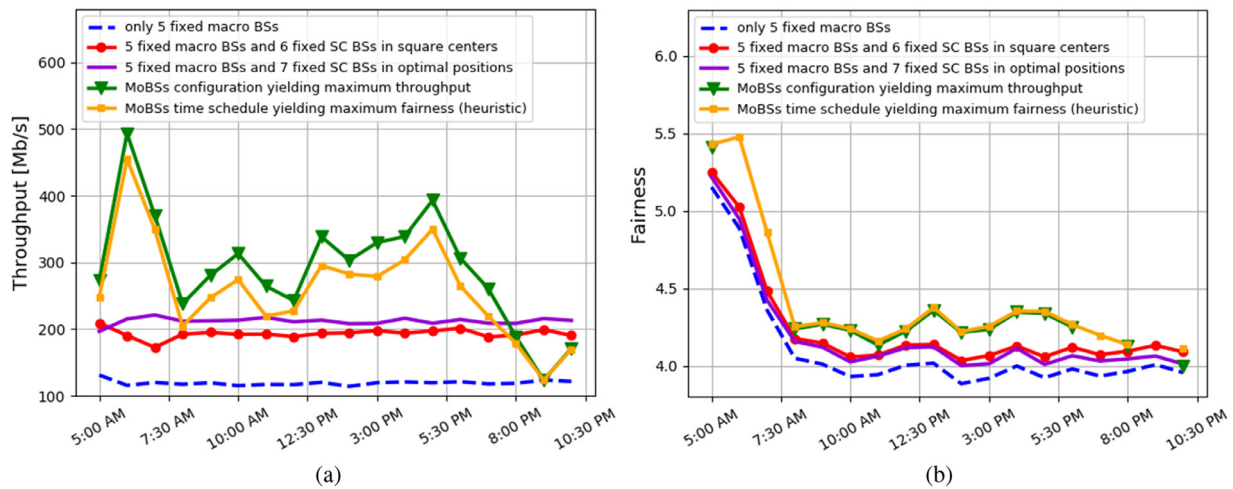


Fig. 16. Throughput in Mb/s (left) and fairness per user (right), in the area of the Polytechnic University of Milan, on April 15, from 5 A.M. to 10 P.M., with: (i) only 5 fixed macro BSs; (ii) 5 fixed macro BSs and 6 fixed small cell BSs in rectangle centers; (iii) 5 fixed macro BSs and 7 fixed small 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).

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 the left plot of Fig. 16 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) the throughput of the MoBS configuration yielding maximum throughput; (5) the throughput of the time schedule over MoBS configurations that optimizes fairness (computed using the heuristic).

Also in this case, obviously, the throughput of the MoBS configuration yielding maximum throughput exhibits the highest peaks, especially when the number of useful MoBSs is high, and the throughput with only 5 fixed macro BSs is 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 numbers 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, except for evening hours, when the number of useful MoBSs is very low.

In general we see that, even if the MoBS 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 the right plot of Fig. 16. 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.

7. Conclusions

In this paper we looked at cellular radio access network architectures where adaptive densification is achieved with small-cell mobile base stations carried by vehicles. Considering the ge-

ographical 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, we computed the throughput achievable with and without mobile small cell base stations, 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.

Our results prove that adaptive densification of radio access networks with small-cell mobile base stations carried by vehicles can be a high-performance and low-cost solution for the rollout of 5G and beyond.

Acknowledgments

This work was supported by the Siebel Energy Institute through the Seed Grant Award 59/2016 and by the European Commission through the H2020 projects 5G-TRANSFORMER (Project ID 761536). The work of V. Mancuso was supported by a Ramon y Cajal grant (ref: RYC-2014-16285) from the Spanish Ministry of Economy and Competitiveness.

References

- [1] W. Mohr, 5g empowering vertical industries, Tech. Rep., 5G PPP, 2016. [Online]. Available: <https://ec.europa.eu/digital-single-market/en/blog/5g-empowering-vertical-industries-0>.
- [2] European Commission, Identification and Quantification of Key Socio-Economic Data to Support Strategic Planning for the Introduction of 5G in Europe, Publications Office of the European Union, Luxembourg, 10.2759/56657.
- [3] A. Furno, D. Naboulsi, R. Stanica, M. Fiore, Mobile demand profiling for cellular cognitive networking, *IEEE Trans. Mob. Comput.* (99) (2016). 1–1.
- [4] A. Furno, R. Stanica, 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*, 2015. Paris.
- [5] L. Budzisz, Dynamic resource provisioning for energy efficiency in wireless access networks: a survey and an outlook, *IEEE Commun. Surv. Tutorials* 16 (4) (2014) 2259–2285.
- [6] DAEL, <http://www.dael.com/en/telecom/cell-on-wheels>, [Online].
- [7] M. Deruyck, J. Wyckmans, L. Martens, W. Joseph, Emergency ad-hoc networks by using drone mounted base stations for a disaster scenario, in: *2016 IEEE 12th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, 2016, pp. 1–7. New York, NY.
- [8] M. Erdelj, E. Natalizio, K.R. Chowdhury, I.F. Akyildiz, Help from the sky: Leveraging UAVs for disaster management, in: *IEEE Pervasive Computing*, volume 16, 2017, pp. 24–32.
- [9] A.M. Hayajneh, S.A.R. Zaidi, D.C. McLernon, M.D. Renzo, M. Ghogho, Performance Analysis of UAV Enabled Disaster Recovery Networks: A Stochastic Geometric Framework Based on Cluster Processes, in: *IEEE Access*, volume 6, 2018, pp. 26215–26230.

- [10] F. Mohammadnia, M. Fiore, M.A. Marsan, Adaptive densification of mobile networks: Exploring correlations in vehicular and telecom traffic, in: The 17th Annual Mediterranean Ad Hoc Networking Workshop (Med-Hoc-Net 2018), 2018, pp. 20–22. Capri, Italy.
- [11] M. Cierny, H. Wang, R. Wichman, Z. Ding, C. Wijting, On number of almost blank subframes in heterogeneous cellular networks, in: IEEE Transactions on Wireless Communications, volume 12, 2013, pp. 5061–5073.
- [12] 3GPP, Evolved universal terrestrial radio access (e-UTRA) and evolved universal terrestrial radio access network (e-UTRAN); overall description; stage 2, tech. spec. TS 36.300, Ver. 10.8.0, 3GPP, 2012.
- [13] F. Bonomi, Cloud and fog computing: trade-offs and applications, Intl. Symp. Comp. Architecture (ISCA), EON Workshop, 2011.
- [14] H. Li, G. Shou, Y. Hu, Z. Guo, Mobile edge computing: progress and challenges, in: 4th IEEE International Conference on Mobile Cloud Computing, Services, and Engineering (MobileCloud), 2016, Oxford.
- [15] K. Zheng, L. Zhao, J. Mei, B. Shao, W. Xiang, L. Hanzo, Survey of large-scale MIMO systems, IEEE Commun. Surv. Tutorials 17 (3) (2015) 1738–1760.
- [16] Q. Li, R.Q. Hu, Y. Qian, G. Wu, Cooperative communications for wireless networks: techniques and applications in LTE-advanced systems, IEEE Wireless Commun. 19 (2) (2012).
- [17] Artemis Networks, <http://www.artemis.com/>, [Online].
- [18] S. Perlman, A. Forenza, An introduction to pcell, 2015.
- [19] S. Bassoy, H. Farooq, M.A. Imran, A. Imran, Coordinated multi-point clustering schemes: asurvey, in: IEEE Communications Surveys & Tutorials, volume 19, 2017, pp. 743–764. Secondquarter.
- [20] V. Sciancalepore, I. Filippini, V. Mancuso, A. Capone, A. Banchs, A multi-traffic inter-cell interference coordination scheme in dense cellular networks, in: IEEE/ACM Transactions on Networking, volume 26, 2018, pp. 2361–2375.
- [21] C. Hoymann, W. Chen, J. Montojo, A. Golitschek, C. Koutsimanis, X. Shen, Relaying operation in 3GPP LTE: challenges and solutions, IEEE Commun. Mag. 50 (2) (2012) 156–162.
- [22] M. Iwamura, H. Takahashi, S. Nagata, Relay technology in LTE-advanced, NTT DOCOMO Tech. J. 12 (2) (2010) 29–36.
- [23] K. Loa, C.C. Wu, S.T. Sheu, Y. Yuan, M. Chion, D. Huo, L. Xu, IMT-advanced relay standards, IEEE Commun. Mag. 48 (8) (2010) 40–48.
- [24] 3rd Generation Partnership Project, TS 36.420 V9.0.0 (2009-12), X2 general aspects and principles (Release 9), 3GPP, 2009.
- [25] 3rd Generation Partnership Project, TS36.413 V9.2.1, Evolved Universal Terrestrial Radio Access (E-UTRA): S1 Application Protocol (S1 AP), 3GPP, 2010.
- [26] D. Kreutz, F.M.V. Ramos, P.E. Verssimo, C.E. Rothenberg, S. Azodolmolky, S. Uhlig, Software-defined networking: a comprehensive survey, in: Proceedings of the IEEE, volume 103, 2015, pp. 14–76.
- [27] C. Zhang, P. Patras, H. Haddadi, Deep learning in mobile and wireless networking: a survey, 2018, arXiv:1803.04311 [cs.NI].
- [28] E. Musk, Master plan, part deux, [Online]. Available: <https://www.tesla.com/blog/master-plan-part-deux>.
- [29] U. Raza, P. Kulkarni, M. Sooriyabandara, Low power wide area networks: an overview, IEEE Commun. Surv. Tutorials 19 (2) (2017) 855–873. Secondquarter, doi: 10.1109/COMST.2017.2652320.
- [30] G.A. Akpakwu, B.J. Silva, G.P. Hancke, A.M. Abu-Mahfouz, A survey on 5g networks for the internet of things: communication technologies and challenges, IEEE Access 6 (2018) 3619–3647.
- [31] M.Z. Hasan, H. Al-Rizzo, F. Al-Turjman, A survey on multipath routing protocols for qos assurances in real-time wireless multimedia sensor networks, IEEE Commun. Surv. Tutorials 19 (3) (2017) 1424–1456. Thirdquarter.
- [32] M. Asif, S. Khan, R. Ahmad, M. Sohail, D. Singh, Quality of service of routing protocols in wireless sensor networks: a review, in: IEEE Access, volume 5, 2017, pp. 1846–1871.
- [33] F. Cadger, K. Curran, J. Santos, S. Moffett, A survey of geographical routing in wireless ad-hoc networks, in: IEEE Communications Surveys & Tutorials, volume 15, 2013, pp. 621–653. Second Quarter.
- [34] Y. Cao, Z. Sun, Routing in delay/disruption tolerant networks: a taxonomy, survey and challenges, in: IEEE Communications Surveys & Tutorials, volume 15, 2013, pp. 654–677. Second Quarter.
- [35] E. Hyytia, J. Virtamo, P. Lassila, J. Kangasharju, J. Ott, When Does Content Float? Characterizing Availability of Anchored Information in Opportunistic Content Sharing, INFOCOM, Shanghai, China, 2011, p. 31233131.
- [36] G. Manzo, M.A. Marsan, G. Rizzo, Performance modeling of vehicular floating content in urban settings, in: 2017 29th International Teletraffic Congress (ITC 29), 2017, pp. 99–107. Genoa.
- [37] A.A.V. Castro, G.D.M. Serugendo, D. Konstantas, Hovering informationselforganizing information that finds its own storage, in: Autonomic Communication, Springer, 2009, p. 111145.
- [38] N.M. Sadek, H.H. Halawa, R.M. Daoud, H.H. Amer, A robust multi-RAT VANET/LTE for mixed control & entertainment traffic, J. Transp. Technol. (5) (2015) 113–121.
- [39] N. Frangiadakis, D. Cmara, F. Filali, A.A.F. Loureiro, N. Roussopoulos, Virtual access points for vehicular networks, in: MOBILWARE 2008, 1st International Conference on MOBILE Wireless MiddleWARE, Operating Systems, and Applications, 2008. Innsbruck, Austria.
- [40] F. Dressler, P. Handlex, C. Sommer, Towards a vehicular cloud using parked vehicles as a temporary network and storage infrastructure, WiMobCity 14, 2014. Philadelphia, PA, USA.
- [41] F. Malandrino, C.E. Casetti, C.F. Chiasserini, C. Sommer, F. Dressler, Content downloading in vehicular networks: Bringing parked cars into the picture, 23rd PIMRC 2012, IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, 2012. Sydney, Australia.
- [42] F. Malandrino, C. Casetti, C.F. Chiasserini, C. Sommer, F. Dressler, The role of parked cars in content downloading for vehicular networks, in: IEEE Transactions on Vehicular Technology, volume 63, 2014, pp. 4606–4617.
- [43] D. Borsetti, C. Casetti, C.F. Chiasserini, M. Fiore, J.M. Barcel-Ordinas, Virtual data mules for data collection in road-side sensor networks, Second International Workshop on Mobile Opportunistic Networking (MobiOpp '10), 2010. New York, NY, USA.
- [44] Y. Sui, I. Guvenc, T. Svensson, On the deployment of moving networks in ultra-dense urban scenarios, in: 1st International Conference on 5G for Ubiquitous Connectivity, 2014. Akaslompolo.
- [45] A. Panno, D. Mastro Simone, New challenge: moving network based on mmwave technology for 5g era, in: 2015 International Conference on Computer, Information and Telecommunication Systems (CITS), 2015. Gijon.
- [46] S. Jangsher, V.O.K. Li, Resource allocation in moving small cell network, IEEE Trans. Wireless Commun. 15 (7) (2016) 4559–4570.
- [47] H. Yasuda, A. Kishida, J. Shen, Y. Morihiro, Y. Morioka, S. Suyama, A. Yamada, Y. Okumura, T. Asai, A study on moving cell in 5g cellular system, in: 2015 IEEE 82nd Vehicular Technology Conference (VTC2015-Fall), 2015. Boston.
- [48] M. Shin, S.T. Shah, M.Y. Chung, S.F. Hasan, B.C. Seet, P.H.J. Chong, Moving small cells in public safety networks, in: 2017 International Conference on Information Networking (ICOIN), 2017. Da Nang.
- [49] S. Andreev, V. Petrov, M. Dohler, H. Yanikomeroğlu, Future of ultra-dense networks beyond 5g: harnessing heterogeneous moving cells, 2017, arXiv:1706.05197.
- [50] M. Pous-Fenollar, P. Ferti, Mobile crowdcell, MWC, 2016. Barcelona.
- [51] C. Vitale, V. Sciancalepore, V. Mancuso, Fair stochastic interference orchestration with cellular throughput boosted via outband sidelinks arXiv:1809.09524.
- [52] 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).
- [53] V. Sciancalepore, V. Mancuso, A. Banchs, S. Zaks, A. Capone, Enhanced content update dissemination through d2d in 5g cellular networks, IEEE Trans. Wireless Commun. 15 (2016) 75177530.
- [54] J.S. Seybold, Introduction to RF Propagation, J. Wiley & Sons (2005).



Marco Ajmone Marsan is a full professor at the Electronics and Telecommunications Department of the Politecnico di Torino in Italy, and a part-time research professor at IMDEA Networks Institute in Leganes, Spain. Marco Ajmone Marsan obtained degrees in EE from the Politecnico di Torino in 1974 and the University of California, Los Angeles (UCLA) in 1978. He received a honorary doctoral degree in Telecommunication Networks from the Budapest University of Technology and Economics in 2002. Since 1974 he has been at Politecnico di Torino, in the different roles of an academic career, with an interruption from 1987 to 1990, when he was a full professor at the Computer Science Department of the University of Milan.

Marco Ajmone Marsan has been doing research in the fields of digital transmission, distributed systems and networking. He has published over 350 papers in the leading conferences and journals of his research area. He is also coauthor of two books: Performance Models of Multiprocessor Systems (MIT Press, 1987) and Modelling with Generalized Stochastic Petri Nets (John Wiley, 1995). Marco Ajmone Marsan has been a member of the editorial board and of the steering committee of the “ACM/IEEE Transactions on Networking”. He is a member of the editorial boards of the journals “Computer Networks” and “Performance Evaluation” of Elsevier, and of the “ACM Transactions on Modeling and Performance Evaluation of Computer Systems”. He served in the organizing committee of several leading networking conferences, and he was general chair of INFOCOM 2013. Marco Ajmone Marsan is a Fellow of the IEEE, a member of the Academy of Sciences of Torino, and a member of Academia Europaea. Marco Ajmone Marsan was the Vice-Rector for Research, Innovation and Technology Transfer at the Politecnico di Torino from 2005 to 2009. From 2002 to 2009 he was the Director of the Istituto di Elettronica e Ingegneria dell’Informazione e delle Telecomunicazioni of the Italian National Research Council. He was the Italian delegate in the ICT and IDEAS committees of FP7.



Foroogh Mohammadnia received a Master degree in electrical engineering from Shiraz University of Technology, Shiraz, Iran in August 2016. She is currently pursuing the Ph.D. degree in electrical, electronics and communication engineering at Politecnico di Torino, Torino, Italy. Her research interests are Mobile Mobile Networks and Big Data analysis.



Christian Vitale is currently a PostDoc at KIOS Center of Excellence in Cyprus. He received his PhD from Universidad Carlos III de Madrid, while working at IMDEA Networks Institute as a researcher. In his career, he worked at Politecnico di Torino, as a PostDoc, at Telecom ParisTech and NEC Europe as visiting researcher. His work focuses on the analytical modeling of wireless network technologies, such as 4G/5G or IEEE 802.11, and on the design of mechanisms improving network efficiency.



Marco Fiore is a researcher at CNR-IEIT, Italy, a Royal Society visiting research fellow, and a Marie Curie fellow. He received a PhD degree from Politecnico di Torino, Italy, and a HDR degree from Universite de Lyon, France. He was associate professor at INSA Lyon, France, associate researcher at Inria, France, visiting researcher at Rice University, TX, USA and Universitat Politecnica de Catalunya, Spain, and visiting research fellow at UCL, UK. His current research interests are in the fields of mobile networks, network traffic analytics and privacy.



Vincenzo Mancuso is Research Associate Professor at IMDEA Networks Institute, Madrid, Spain, and recipient of a Ramon y Cajal research grant of the Spanish Ministry of Science and Innovation. Previously, he was with INRIA Sophia Antipolis (France), Rice University (Houston, TX, USA) and University of Palermo (Italy), from where he obtained his MSc and his Ph.D. in Electronics, Computer Science and Telecommunications. He has authored more than 110 peer-reviewed publications focusing on Internet QoS and on the analysis, design, and experimental evaluation of opportunistic and adaptive protocols and architectures for wireless networks. He is currently working on analysis and optimization of wireless access networks and on the measurements and assessment of mobile broadband networks.