Making the case for a real-time market of wireless resources with dynamic network slicing and sharing

Antonio Capone¹, Matteo Cesana¹, Ilaria Malanchini², Vinay Suryaprakash²
¹Politecnico di Milano (Italy), ²Nokia Bell Labs (Germany)

Abstract—This position paper presents a new approach to infrastructure sharing for mobile networks based on the idea that network resources, both at the wireless access and the core network, can be traded dynamically on a real-time market where virtual operators or tenants compete to obtain the resources to serve their users. The negotiation procedure can be automated by allowing tenants to define their policies, objectives and constraints and designing a pricing approach that can guarantee fairness. Moreover, the whole system can be made controllable by a regulation authority that can use global constraints and pricing policies in order to guarantee fair competition, social benefits, and a level of revenues for the infrastructure providers able to guarantee capacity expansion and technology update.

I. INTRODUCTION

Mobile networking industry is fast moving towards fifth generation technology (5G) and already starting research on beyond 5G evolution. Even if enhanced version of mobile broadband (eMBB) services will continue to be an important pillar of operators offer, the new application domains of massive machine type communications (mMTC) and ultra-reliable and low latency communications (URLLC) are expected to play an important role [1].

Unfortunately, technological evolution of mobile networks and the fast growth of traffic volume require large infrastructure investments and place an enormous strain on the already decreasing profitability of mobile operators [2]. Hence, it is common belief that the introduction of 5G will mark also a change in the business modus operandi of mobile operators.

Some of the novelties being introduced in the architecture of mobile networks can be instrumental to this business change. In particular, the concepts of network function virtualization and network slicing allow operators to optimize the usage of resources and make their infrastructure more flexible to meet the needs of the very diverse set of vertical applications that are expected for next generation networks. Generally speaking, a network slice is managed by a tenant and can be dedicated to support a specific application with its own resources in the different network segments (access, transport, and core) and its chain of virtual functions and services. Specialized network slices can extend the offer of operators to the end users and also to other business entities that can act as tenants and develop their own solutions.

As a matter of fact, it is easy to see network slicing as a technical tool able to open up the mobile network to the longly predicted evolution towards infrastructure sharing [3]. For operators, sharing a common infrastructure can potentially reduce their capital and operational expenditures and allow them to offer to their customers better prices, a larger number of services, and a better quality of service. Sharing can be undertaken at various levels, as indicated in the Organization for Economic Co-operation and Development (OECD) report [3]. However, the most comprehensive method of sharing, known as active sharing, consists of multiple mobile virtual network operators (MVNOs) that lease or rent the infrastructure from an infrastructure provider (InP). Service Level Agreements (SLAs) regulate relations between the parties who intend to share the infrastructure and they usually take the form of contracts which are enforced over a long period of time.

The concept of tenant of network slices can clearly be associated to that of MVNO assuming that the same operator can manage several specialized slices to offer services to its customers or even that general purpose slices are offered by infrastructure providers and then further subdivided using a layered structure of virtual resources. MVNOs can even rent slices from different infrastructure providers according to the most convenient options in different areas and combine them together to define the service for its own users. Although this scenario appears rather different from the current one, it is reasonable to expect that both economic advantages and regulatory forces may drive the system in this direction.

This type of active sharing, which obviously implies spectrum sharing, may not be particularly appealing if based on today’s SLAs with long term contracts, since the MVNOs will not have the ability to accommodate traffic fluctuations and might often find themselves in scenarios where they risk to be unable to cater services to their customers or forced to rent more resources than needed to account for uncertainty. Another risk is that the market value of resources offered by infrastructure providers decreases and prevent them to gain the money necessary for the system to evolve in terms of technology and capacity.

We argue that the technological evolution of 5G provides the opportunity to transform mobile networks into a resource that can be flexibly partitioned and traded on short time scales, a few seconds or minutes, as already done for other network based resources such as electricity or natural gas. In order for such a system to work, viz. for it to be profitable for all the parties involved, each of them should have a good understanding of their own budgets, the implications of short-term fluctuations on them, and an accurate estimate of their traffic load and mix.

In this position paper, we present a new approach to infrastructure sharing based on the idea that network resources,
both at the wireless access and the core network, can be traded dynamically as in a real-time market. In particular, in Sec. II we review the current state of the art in wireless sharing. Afterwards, we highlight the technical and economical advantages of long-term active sharing and show a first feasibility assessment of the wireless market envisioned in Sec. III and Sec. IV, respectively. Sec. V concludes the paper highlighting the main open challenges.

II. STATE-OF-THE-ART IN WIRELESS SHARING

Historically, operators differentiated their offers to entice new customers based primarily on coverage and capacity guarantees. As explained in the Introduction, however, mobile network operators – faced with an explosion in the types of services expected to be catered to by wireless technologies in the near future, the tightening constraints on spectrum availability, and a sluggish growth in profits – have turned to network sharing in order to reduce expenditure while adhering to service guarantees. Driven by necessity, operators in different countries started to engage in various forms of network sharing based on their local regulations and individual service assurances [3]. Furthermore, the OECD concluded in [3] that network sharing is beneficial not only for helping operators maintain their service guarantees and reduce expenditure, but also for the overall health of the economy as a whole, since sharing ensures the continued existence of a competitive and diverse market (by making it harder for larger operators to effectively monopolize the market).

There are four main types of sharing delineated in [3], namely: passive sharing, e.g., sharing sites, masts, and antennae; active sharing, e.g., radio access network (RAN) sharing; core network sharing; and, network roaming. Since network roaming has long been a mainstay of networks to ensure users’ connectivity across geographic (mostly national) boundaries and because core network sharing – usually provided as services to network operators by third party service providers/vendors – is also widely practiced, research has focused mostly on the first two categories, i.e., passive and active sharing. As illustrated by the examples listed, passive sharing is defined as the joint use of the space or the physical infrastructure resulting in operators splitting the costs of the power supply, air conditioning, lease of the sites, etc. Studies such as [4]–[6] indicate a potential cost savings of 15%–25% depending upon the deployment scenario (i.e., dense urban, urban, sub-urban, or rural). Additionally, this type of sharing poses very few technological difficulties and allows mobile operators to retain complete autonomy over their own network.

However, the OECD report [3] – based on the analysis of data from around 35 countries – states: “In general, savings from active sharing are greater than for passive sharing as a higher proportion of costs are shared”. Active sharing entails sharing all radio access network equipment including the site, mast, antenna, and backhaul. The company managing the site typically leases the whole package to a network operator and also carries data to their core network. This method, though financially more lucrative, is technologically quite difficult to implement and – perhaps more importantly – restricts the autonomy mobile operators have over their networks.

Since resolving these issues simultaneously is quite difficult, works usually tend to handle them one at a time. Recent papers such as [7]–[10] focus on the technical aspects required to facilitate active sharing. E.g., the authors of [9] provide a generic (technology agnostic) method to efficiently utilize hardware as well as spectrum. A potential solution to the issue of operator autonomy (while using active sharing) is considered to be network slicing, [11]. The basic idea comprises the creation of a network slice, which is a customizable, dedicated, logical (and – in most cases – virtual) entity capable of functioning like an independent network though a common physical infrastructure is in use. Based on this concept, works such as [12]–[16] provide dynamic approaches to slicing in which dedicated resources can be allocated to various services depending on their requirements.

At first glance, from the description above, it appears that the salient issues faced by active sharing have been mitigated. A closer examination, however, reveals that a serious discrepancy remains, namely: although the technical solutions proposed take varying channel conditions and a fluctuating traffic mix into account, the solutions – based mostly on long-term service level agreements (SLAs), proposed for providing network operators their desired autonomy – do not. This results in scenarios where operators often find themselves facing a surfeit or paucity of network resources in their day-to-day operations and therefore, being unable to make the most of active sharing.

III. TECHNICAL AND ECONOMICAL ADVANTAGES OF LONG-TERM ACTIVE SHARING

As clear from the overview on the related work in the previous section, the answer to the question “is active sharing worthy?” might not be unique, since the active sharing process encompasses both technical, technology-specific issues, as well as economical ones. The very same problem of assessing the goodness/profitability of an active sharing process is a research challenge per se, and we posit here the need of comprehensive and consistent techno-economical frameworks to do the job.

On the technical side, few elements should be defined to clearly assess the validity of a specific sharing approach: (i) the type of active sharing to be put in place (RAN sharing with/without spectrum pooling), (ii) the quality perceived by the end users in a shared network which, in turn, depends on (iii) the specific scheduling strategies in the shared network. On the economical side, the active sharing should account for: (i) the current market share of the operators participating in the sharing process, (ii) the expenses/costs to be sustained by the operators and (iii) the effect of sharing on the revenues of the MNOs, which is captured by a pricing model for the services offered on the shared network. Moreover, active sharing might be impacted by the presence/absence of a regulation authority supervising and/or incentivizing the sharing process.
A first comprehensive framework to assess RAN sharing is presented in our previous works [17,18]. The work analyzes the strategic situation in which multiple MNOs have to decide whether to invest in improving their pre-existing 3G network by deploying LTE small cells in dense urban areas and whether to share the investment (and the resulting Radio Access Network) with other MNOs.

We tackle the sharing problem in two situations: a socially-driven case in which a regulatory entity enforces sharing policies to the MNOs to optimize the quality of service perceived by the end users; in this case, we seek the sharing configuration (the MNOs that share the network infrastructure) that maximizes the end user achievable throughput, while being sustainable by the MNOs (revenues at least equalize the investment costs). Differently, in the market-driven case, we capture the competitiveness of the MNOs and their profit-maximizing nature by looking for the sharing configurations that maximize the profit for the MNOs.

The two cases share common assumptions in terms of the techno-economical characteristics. Namely

- each MNO has a given market share which is stable over time, that is, the market share is not affected by the sharing policy of the MNO;
- the cost of shared infrastructure including capex and opex is shared among all the MNOs participating in the sharing;
- the end users are characterized by a parameter capturing the willingness to pay for getting an improved service; the pricing model for the LTE service is linear in the achievable rate by the end user, that is, the revenues for the MNO grow linearly with the rate improvement in the sharing infrastructure;

The socially-driven case is formalized in a mathematical framework based on a MILP formulation to determine socially optimal configurations to optimize the quality of service perceived by all users; the market-driven is modeled by resorting to cooperative game theory to determine stable coalitions of MNOs willing to share the investment. The major results obtained by applying the proposed framework to realistic network settings can be listed as follows:

- the sharing configurations/policies are affected by the end users willingness to pay for improved services and the MNOs’ market shares;
- in the market-driven case, MNOs are better off building a unique shared RAN than creating sub-coalitions or building individual RANs due to the combined gain from spectrum aggregation and cost reduction from sharing the network infrastructure;
- in both cases (socially/market-driven), sharing is particularly appealing when users value service improvement little; regardless of the MNOs market shares, this is the case where the MNOs struggle with high investment costs.

IV. TOWARDS A REAL-TIME MARKET OF WIRELESS NETWORK RESOURCES

The proposed framework envisions the existence of a real-time wireless market that is able to dynamically negotiate and share resources (i.e., spectrum) with a very-fine granularity, both in frequency and time, while guaranteeing profitability to all parties involved. A first assessment of the feasibility of such wireless market envisioned is carried out in [19] and [20], in which we propose a first preliminary step towards more comprehensive techno-economic market models for radio access networks. The main features that we include – detailed later on in this section – are:

- real-time resource negotiation,
- possibility of service differentiation,
- capability of handling service heterogeneity,
- enabling future network expansion.

First of all, we considered two different types of stakeholders, i.e., a single infrastructure provider (InP) and multiple mobile virtual network operators (MVNOs). The former has full power in deciding when and how to allocate wireless resources as well as controlling the pricing mechanisms. The latter can make resource requests in real-time and decide whether or not to buy resources based on their own Quality-of-Service (QoS) target and budget.

To accommodate the need for fast re-negotiations of resources, the wireless market should be able to work in real-time and to make decisions on a very short time scale (i.e., seconds, minutes). As proposed in [19], the service level agreements (SLAs), e.g., in terms of sharing parameters, are then directly mapped into the resource allocation algorithms in such a way that they are able to incorporate both long term MVNOs’ expectations and real time traffic fluctuations. In this regard, the SLAs between the different parties should reflect the willingness of the MVNOs to buy resources and their flexibility in sharing them. For this reason, we envision a long term sharing ratio, able to reflect the amount of resources needed (hence requested and possibly used) on the long run and a short term deviation, which allows both the MVNOs to trade unused resources and the InP to fully exploit the flexibility and improve the scheduling efficiency. To this extent, [19] proposes an optimization framework which is able to maximize the quality achieved by the MVNOs while preserving fairness and satisfying budget requirements. Renegotiation of the sharing parameters is also included to adapt to traffic fluctuations and changing in service requirements.

Furthermore, the willingness of MVNOs to share should not prevent them from the possibility of differentiating. For instance, MVNOs should still be able to provide different services, or the same service but with different quality levels. As suggested in [20], MVNOs can compete and dynamically select the quality target to deliver to their customers, while simultaneously seeking to maximize their profits. Namely, we introduce a quality parameter that can be selected by each MVNO autonomously and determines the level that the MVNO decides to provide to its users for the required service.
This allows MVNOs to apply quality diversification in the services delivered to their customers. In [20], competition is modeled by using a non-cooperative game, wherein the Nash Equilibria show the propensity of operators to meet their users’ requirements.

Moreover, the introduced ecosystem and new functionalities will also be able to support service heterogeneity. Namely, network slicing and the use of dedicated resources for different services is beneficial for achieving the service guarantees required by the large variety of applications foreseen in future networks. To incorporate slicing functionalities within our framework, the wireless market should be able to determine the size of the network resource slices required for various active services, but it should also adapt prices accordingly.

Finally, as mentioned above, pricing mechanisms should automatically adapt the cost of resources and the price paid by each MVNO in such a way that they are proportional to what is requested and what is actually used. The cost of resources should be market driven, and prices and budgets must be directly incorporated into the short term trading operated by the wireless market. Furthermore, the InP should also be able to continuously collect revenue for network expansion. For this reason, we introduced in [19,20] the notion of “pressure cost”. Namely, the pressure cost ensures that the costs of network resources as well as future expansion scale according to the MVNOs’ demand.

V. TECHNICAL AND SCIENTIFIC CHALLENGES

In conclusion, we have discussed in this paper a new approach to infrastructure sharing based on the idea of a real-time market of wireless resources with dynamic network slicing and sharing. Based on our previous work, we presented a first feasibility assessment which shows the main functionalities envisioned and how those can be modeled and realized.

At the same time, several open points still need to be addressed to make the proposed wireless market a reality. In our opinion, the most interesting technical challenges that point to possible research directions are the following:

- A clear strategy definition should be available to virtual operators (or tenants) to allow them to play in their own best interests while also taking into account prediction information of the upcoming traffic mix as well as users’ channel conditions – elements which are critical to, e.g., adjust the slice assignments and sharing agreements.
- Pricing models play a critical role in the market, as well as the fairness guarantees that must be ensured by the infrastructure provider. Different behaviors can be induced changing pricing or resource allocation strategies, for instance considering different time scales in the market (similarly to the case of energy) and privileging operators able to accurately predict traffic and buy resources in advance.
- The critical role played by the regulation authority requires the definition of its responsibilities and interactions with the wireless market. These can take the form of constraints on the strategies of operators and infrastructure providers and on the price composition that should for instance ensure that sufficient economic resources are injected in the system to guarantee its evolution and ability to support traffic and applications.

An overview of the scenario envisioned is presented in Figure 1.
ACKNOWLEDGEMENT

The work has been partially supported by the European Union Horizon 2020 Marie Sklodowska-Curie ITN/EID project ACT5G (Grant Agreement No. 643002).

REFERENCES