

Research Article

A Service-Oriented Approach for Radio Resource Management in Virtual RANs

Behnam Rouzbehani ¹, Luis M. Correia,¹ and Luísa Caeiro²

¹IST/INESC-ID, University of Lisbon, Lisbon, Portugal

²ESTG/INESC-ID, Setúbal Polytechnic Institute, Setúbal, Portugal

Correspondence should be addressed to Behnam Rouzbehani; behnam.rouzbehani@tecnico.ulisboa.pt

Received 19 February 2018; Revised 25 May 2018; Accepted 12 June 2018; Published 10 July 2018

Academic Editor: Anna Zakrzewska

Copyright © 2018 Behnam Rouzbehani et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Virtualisation, as a key role player of future mobile communications, promotes the idea of service-oriented architectures. This paper proposes a model of Radio Resource Management (RRM) for emerging Virtual Radio Access Networks, based on the interaction between two separated management entities: Common-RRM (CRRM) to coordinate the radio resources among the Radio Access Technologies (RATs) and a centralised virtualisation platform on top of it, called Virtual-RRM (VRRM), which is responsible for service orchestration among Virtual Network Operators, enabling the definition of various services and policies, separately from vendors and underlying RATs. The main objective of VRRM is to satisfy the Service Level Agreements associated with different service classes to the highest possible level, within the framework of proportional fairness. On the other hand, CRRM is in charge of mapping the demanded capacity of each service onto the most suitable RATs. The model is further extended to deal with extreme situations of resource shortage, resulting from high traffic loads, by introducing delay to lower priority services. The performance of the proposed model is evaluated in a practical multi-RAT scenario. Results confirm that the isolation of service classes is consistent with the introduced serving weights, while all the demanded capacities from different services are responded by the most suitable RATs. Finally, independent of the variation of traffic load, 100% of the aggregated capacity is used.

1. Introduction and Motivation

The rapid increase of demand for new services in recent years has imposed crucial requirements to network operators. However, service provision in most of the existing telecom networks is tightly coupled with costly inflexible infrastructure, which has been specifically designed to deliver a limited range of services [1]. In order to overcome this barrier, management and operation functions have to be customised, to enable support for particular service types. Virtualisation, as a key enabler of the *Network-as-a-Service* paradigm for future mobile communications, can significantly facilitate the provision of customised services, by decoupling networking functionalities from the underlying infrastructure. In this framework, a common physical infrastructure can be effectively shared, in an isolated manner, among coexisting tenants, called Virtual Network Operators (VNOs); each

VNO will be able to implement its own specific rules and regulations over the shared set of resources, without necessarily being aware of the underlying virtualisation process [2]. This is of great importance for the design of convergent mechanisms and coordination of the available resources across the emerging wireless Heterogeneous Networks (Het-Nets). In this regard, the main difference between the services that can be provided by future mobile implementations, compared to the majority of current technologies, such as LTE, lies in the *granularity* of the level of service customisation [3]. While in LTE all packets in a bearer are treated the same way, service flows in Virtual Radio Access Network (RAN) can be flexibly customised in a more granular way to sufficiently differentiate in between their requirements.

Radio Resource Management (RRM) is one of the key functionalities of cellular networks, which has a direct influence on the Quality of Service (QoS) of users, as well as on

the performance of higher layers [4]. With the introduction of new applications, which has led to the increased number of connected users, the problem of RRM has become particularly challenging, since the various applications have different and often conflicting needs [5]. Therefore, each service has to be managed independently, and the available radio resources must be allocated on a priority basis. This is the point when the context of *resource slicing* becomes interesting: each resource slice is defined to address the specific requirements of a service with a certain degree of performance isolation, to ensure that, regardless of the variation of network status, the desired performance level of independent slices is always met [6]. Such a flexible mechanism of radio resource slicing and management can be achieved through the virtualisation of radio resources. This way, an RRM algorithm should not only maximise the performance of different slices, but also the usage of the overall pool of shared resources [7].

Although there are quite extensive studies to address different challenges in traditional wireless networks, when it comes to the evolutionary technologies for future wireless communications, the existing techniques have to be modified in order to accommodate the specifications of new services and architectures. Specifically concerning the problem of RRM in Virtual RAN, there is a lack of effort to thoroughly cover the key parameters, such as *customised* service provisions, *isolation* between virtual slices, *fairness*, and the mechanisms of *interaction* between different entities. In [8], a model for RAN virtualisation is proposed, which provides a solution for the mapping of virtual network elements onto radio resources, as well as providing an algorithm of radio resource negotiation and allocation in a general term, which does not include specifications of different services into account. The problem of slice scheduling and performance isolation in Virtual RAN is addressed in [9], but, like the previous work, different needs of applications in network slicing, as well as fairness, are neglected. Although a fairness concept based on roaming price is defined in [10], and some QoS parameters are addressed in the proposed generalised virtualisation framework, the matter of isolation between services and operators is not covered.

Economic models for pricing based on optimising an objective function have been proposed with the purpose of balancing network throughput and users' fairness as two competing interests [11]. Among the available mechanisms, proportional fairness has proved to be an effective approach when the objective is to maximise the average long-term users' data rate [12, 13]. In this regard, it is suggested that, by employing a logarithmic utility function in the slice scheduler of Virtual RAN, an effective mechanism of fairness based on the concept of proportional fairness can be achieved among virtual slices [14]. A model of RRM is proposed in [15] for 5G wireless infrastructures, which aggregates users' traffic across a Wi-Fi/LTE network, by considering an α -fair mechanism in the objective function, including proportional fairness as its special case; however, there is no effort to address the specifications of different services. Another approach for cooperative RRM in 5G heterogeneous cloud RANs is proposed in [16], which considers a modified max-min technique as an alternative to proportional fairness,

although still the evaluation of the effect of different service parameters on model performance is neglected.

This paper proposes a model for *RAN-as-a-Service*, which deals with the *high-level* management of Virtual RAN, considering the concept of network slicing according to the specific customised service requirements. The contribution of this work is twofold. First, the proposed models of Virtual-RRM (VRRM) in previous works [17, 18] are extended and modified in order to realise the *separation* in the role of Infrastructure Providers (InP) and VNOs. This way, the task of mapping the demanded capacity from different virtual slices onto the underlying physical RATs is defined in Common-RRM (CRRM), in order to promote the notion of *end-to-end slicing*, since it was not covered previously. To accommodate this function, a new mechanism of *cooperation* between the two entities (i.e., VRRM and CRRM) is also proposed to define which information has to be exchanged in order to achieve an efficient interaction, while keeping a desired level of isolation.

Second, regarding the fairness problem, while in the previous paper [17], the definition of fairness is to minimise the deviation of services data rates from a nominal fixed value, the framework has been changed to a more flexible and accurate one, to cope with the concept of *proportional fairness*. This change comes with the price of a more complex objective function compared to the previous linear one. However, since VRRM deals with the high-level network management, including VNOs' policies, which does not require to be changed so often, this level of complexity is tolerable, while the mechanism of CRRM is still a low complexity linear mapping from the capacity demands onto the underlying suitable RATs.

It is worth mentioning that although some user-based metrics, such as users' average data rate or average percentage of delayed users, are defined to evaluate the model, the current work is generally not intended to address the traditional problem of RRM at the *end-user scale* or the associated metrics, such as service delay time. The reason is that the slice scheduling of service flows is performed in upper layers, and also that this scheduling does not need to further modify the existing mechanisms of RRU schedulers in lower layers [19]. Furthermore, since this work deals with a higher layer management of the network, the objectives can be addressed by an abstract vision of the underlying infrastructure. Accordingly, more detailed assumptions for the lower layers, such as power control, antenna configuration, scheduling, mobility management, channel information, and protocols, are out of the scope of this paper. Regarding the *scalability* of VRRM with a proportional fairness mechanism of scheduling, as well as service orchestration in *multitenant* environment, another work has been already developed [20], which specifically addresses these issues and proves the efficiency of the proposed VRRM model; therefore, neither of these topics are covered in this work.

The rest of the paper is organised as follows. Section 2 represents the conceptual network architecture along with the functionalities of the involved components. Section 3 describes the analytical model of VRRM, CRRM, and the mechanism of interaction in between each other. In order to

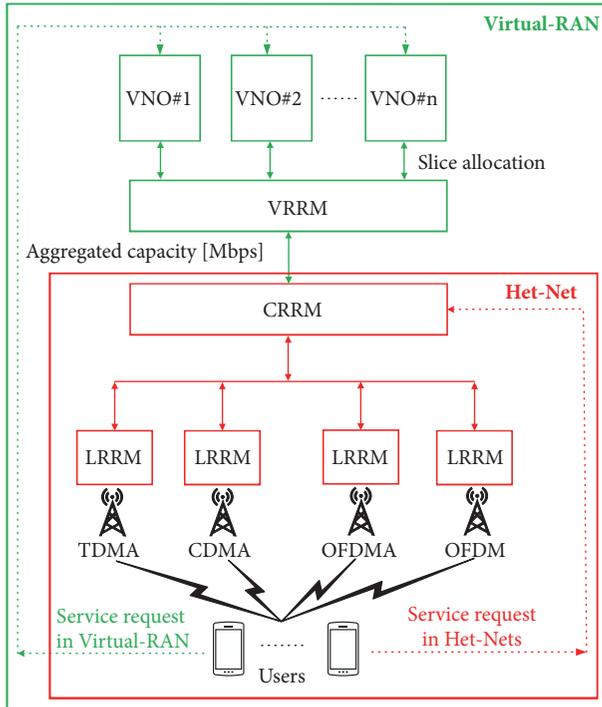


FIGURE 1: Differences between traditional Het-Net and Virtual RAN.

evaluate the proposed model, a reference scenario with some assumptions is defined in Section 4, followed by the analysis of results obtained from simulations in Section 5. Finally, the paper is concluded in Section 6, by highlighting the main achievements.

2. Network Architecture and Main Assumptions

In order to point out the main differences between traditional Het-Nets and the proposed model for Virtual RANs and to explain the functionalities of the involved parties in the RRM model, a software-based hierarchical network architecture is shown in Figure 1, which is a conceptual architecture consistent with the suggested 3GPP business model for *wholesale-only* network sharing, in which Infrastructure Providers (InPs) do not offer service to end users, rather selling capacity to businesses that do not own the infrastructure [14] (details on physical implementations as well as on different interface protocol features can be found in [21]).

The service connectivity request in typical Het-Nets is directed to CRRM, as the usual entity for network management, and processed centrally to be assigned to a suitable RAT according to a decision criterion. However, in the proposed architecture for Virtual RAN, demand for a specific service and capacity goes directly to the linked VNO as the service provider. This capacity needs to be delivered respecting the Service Level Agreements (SLAs) between InP and VNO.

In contrast to the existing Het-Nets, the role of network operators is separated from InPs; accordingly, VNOs on top of the hierarchy do not own the infrastructure, rather sharing

the radio resources from different RATs owned by InPs. As a result, from their perspective, it is not important by which technology they are being served, as long as the SLAs are satisfied. VNOs ask for *Capacity-as-a-Service* from a centralised virtualisation platform called VRRM [17], which does not exist in the current architecture of Het-Nets and is placed on the top of CRRM. VRRM is in charge of managing the total available capacity provided by CRRM, through aggregating all the Radio Resource Units (RRUs) from different RATs, which can be OFDM (related to Wi-Fi), OFDMA (related to LTE), CDMA (related to UMTS), and TDMA (related to GSM), and sharing the capacity from separated slices associated with different services of VNOs. By providing isolation and element abstraction, VRRM enables each VNO to deploy its own protocol stack over the same set of RRUs per RAT (e.g., resource-blocks in LTE, codes in UMTS, time-slots in GSM, and carriers in Wi-Fi) [22], therefore, promoting the notion of multitancy in a virtualised environment with several existing access techniques.

VRRM has to closely interact with CRRM, by translating VNOs' requirements and different SLAs into a set of management policies for the lower levels [17]. These policies contain information about the capacity demanded from each service, as well as their priorities. In return, CRRM provides VRRM with the monitoring reports and information (e.g., the available aggregated capacity) to enhance its performance. CRRM is in charge of mapping the resources of different RATs to satisfy the requests of VRRM, by demanding each RAT to provide a portion of the available capacity to be assigned to end users. The service-to-RAT association mechanism takes the load and suitability of each RAT for performing specific services into account. The performance of CRRM is being optimised, based on the information coming from Local-RRMs (LRRMs), which are in charge of managing the allocation of local RRUs from each RAT to the connected end users.

It is assumed that VNOs can provide the four service classes [23], i.e., *conversational*, *streaming*, *interactive*, and *background*, considering that the design and implementation of future mobile communications, including 5G RAN architecture, should support the 3GPP service classes [4]. As mentioned, since service customisation is one of the main aspects of RAN slicing, on top of these service classes three types of SLAs are also proposed in this work to define the level of service guarantees in terms of priority and contracted capacity, which can be modified according to the VNOs' policies. The three general categories of SLAs are considered as follows:

- (i) Guaranteed bitrate (GB): the highest priority category, for which minimum and maximum thresholds for data rate assignment have to be always guaranteed, regardless of the variation of traffic load and network status; therefore, users are always expecting a good quality in return of a relatively higher service price.
- (ii) Best effort with minimum guaranteed (BG): the second highest priority, for which just a minimum level of data rate is guaranteed, and higher data rates are served in a best effort manner in case of availability;

from the users' viewpoint, a service with acceptable quality and affordable price is expected.

- (iii) Best effort (BE): the lowest priority type, for which there is no level of service guarantees and users are served in a pure best effort manner; consequently, in extreme case of high traffic loads, BE users are the first ones who suffer.

These definitions of SLAs are in line with 5G network design assumptions, since the QoS profile of service flows should carry the information about whether a service traffic is categorised as GB or non-GB [3].

In addition, one should note that, although only data rates are being considered in here for the definition of the SLAs, in a more global perspective (which is out of the scope of this paper but is being considered for future work), one should take other parameters, like link latency, connection reliability, and connectivity capacity.

3. Radio Resource Allocation Model

3.1. VRRM Approach. The primary goal of VRRM is to maximise the usage of the aggregated capacity, which is calculated and provided by CRRM, in order to satisfy the contracted SLAs to the highest possible level, considering services' priority, while distributing capacity according to the concept of proportional fairness [24], subject to some constraints, including maximum achievable capacity, predefined SLA thresholds, and access of users to different RATs.

To realise these goals, VRRM's analytical model is formulated as a constrained concave optimisation problem. The objective function in (1) is defined with the aim of balancing between efficiency and fairness when allocating the resources in a network with heterogeneous services. The utility function, f_{VRRM} , is a measure of efficiency, mapping the portion of network bandwidth assigned to users onto a real number, quantifying the expected users' satisfaction, given the allocated resources; the logarithmic behaviour implies that users' satisfaction increases with the increasing rate as the allocated bandwidth increases. f_{VRRM} is executed in independent time intervals, representing the VRRM decision windows according to the framework of [25]. In this regard, all network parameters, such as number of users or allocated data rates, are assumed to be fixed during each time interval, representing the parameter's average value in this period.

$$\begin{aligned} & \max_{\mathbf{w}^{usr}} f_{VRRM}(\mathbf{w}^{usr}) \\ & = \max_{\mathbf{w}^{usr}} \sum_{k=1}^{N^{srv}} \lambda_k \log \left(\sum_{i=1}^{N_k^{usr}} w_{k,i}^{usr} \frac{R_{k[\text{Mbps}]}^{srv,max}}{R_{[\text{Mbps}]}^{CRRM}} \right) \end{aligned} \quad (1)$$

where

- (i) N_k^{usr} : number of users performing service k ,
(ii) N^{srv} : number of provided services,
(iii) $R_k^{srv,max}$: maximum assignable data rate for service k ,
(iv) $R_{[\text{Mbps}]}^{CRRM}$: total available capacity provided by CRRM to VRRM,

- (v) $w_{k,i}^{usr}$: assigned weight to user i , performing service k , ranging in $[0, 1]$,
(vi) λ_k : weight of service k , prioritising data rate assignment,
(vii) \mathbf{w}^{usr} : vector of users' weights, to obtain the long-term average users' data rate, according to the contracted SLAs provided by VNOs to VRRM, as well as other constraints:

$$\mathbf{w}^{usr} = \left[w_{1,1}^{usr}, w_{1,2}^{usr}, \dots, w_{1,N_1^{usr}}^{usr}, \dots, w_{N^{srv},1}^{usr}, w_{N^{srv},2}^{usr}, \dots, w_{N^{srv},N_k^{usr}}^{usr} \right]^T \quad (2)$$

This way, VRRM builds a bridge between the functionalities of MAC and higher layers, by optimising the allocation of radio resources for different applications [26]. Serving weights λ_k define services' priorities and enable the tuning of the portion of the data rate that will be assigned to each service, being assumed that they are positive integer numbers: a service with higher priority has a higher serving weight, and the lowest value is always assumed to be 1. On the other hand, user's weights, $w_{k,i}^{usr}$, are the desired parameters to be obtained by solving the optimisation problem, indicating the share of each user from the maximum achievable data rate. Therefore, the average long-term data rate of each user has to fall in the acceptable range of data rate variation defined in the SLA contracts as follows:

$$R_{k[\text{Mbps}]}^{srv,min} \leq w_{k,i}^{usr} R_{k[\text{Mbps}]}^{srv,max} \leq R_{k[\text{Mbps}]}^{srv,max} \quad (3)$$

where

- (i) $R_k^{srv,min}$: minimum assignable data rate for service k .

In addition, there is a logical constraint, which specifies that the total bandwidth allocated to all users performing different services cannot surpass the total aggregated capacity provided by CRRM. Therefore, the whole network capacity is subject to an upper bound for the total used bandwidth assigned to users:

$$\sum_{k=1}^{N^{srv}} \sum_{i=1}^{N_k^{usr}} w_{k,i}^{usr} R_{k[\text{Mbps}]}^{srv,max} \leq R_{[\text{Mbps}]}^{CRRM} \quad (4)$$

The last constraint indicates the fact that the access of users to different RATs depends on their location. Assuming that a certain number of users are enjoying full coverage of all available RATs, while the rest has access just to the cellular ones, the following condition can be used to define the limit of access of the latter group to cellular RATs:

$$\sum_{k=1}^{N^{srv}} \sum_{i=1}^{N_k^{usr,cell}} w_{k,i}^{usr} R_{k[\text{Mbps}]}^{srv,max} \leq R_{[\text{Mbps}]}^{CRRM,cell} \quad (5)$$

where

- (i) $N_k^{usr,cell}$: number of users with access only to cellular RATs, performing service k ,

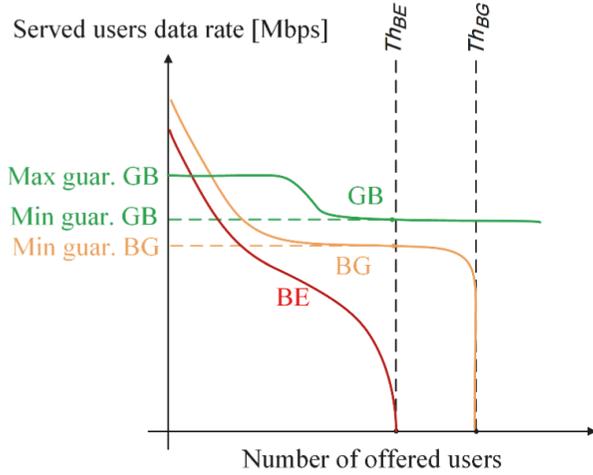


FIGURE 2: General policy of data rate allocation for different services.

- (ii) R^{CRRM}_{cell} : aggregated capacity obtained from the RRUs of cellular RATs, provided by CRRM.

The proposed problem is solved by CVX [27], which is a modelling system based on MATLAB, developed by Stanford University for disciplined convex programming. The method used to solve the problem is primal-dual interior point [28], the termination tolerance of the function being set to 10^{-5} .

The algorithms of slicing and data rate allocation are based on *prioritisation* of services and network pricing, which also include an admission control policy to guarantee the required SLAs, being a common approach for service-based RAN slicing [29, 30]. Moreover, in the extreme situation when network capacity is not enough to satisfy all users with at least the minimum guaranteed service level, delay is introduced to some of the low priority users, releasing the required capacity to serve the remaining ones with the minimum acceptable data rate. In this respect, the number of delayed users during each VRRM decision window corresponds to the average number of users whose service requests are not granted during this period. The general trend for data rate allocation is illustrated in Figure 2, representing the expected behaviour of VRRM under the specific model assumptions. For an increasing number of users, up to threshold Th_{BE} all users are served, but after that, BE users are discarded, since they have the lowest priority, in order to provide VRRM the possibility of serving all BG and GB ones with the minimum contracted data rates; after this point, capacity is not sufficient to address all contracted data rates; therefore, a reasonable decision is to start delaying just enough number of BG users to free capacity for GB ones. Threshold Th_{BG} represents the point when no BG user is left to be delayed, and the VRRM mechanism has to start delaying some of the GB ones as the last alternative.

3.2. Interaction between CRRM and VRRM. Since VRRM is a centralised entity, responsible for allocating the required capacity of each service defined by the associated VNO, based on the specific policies of that VNO regarding the

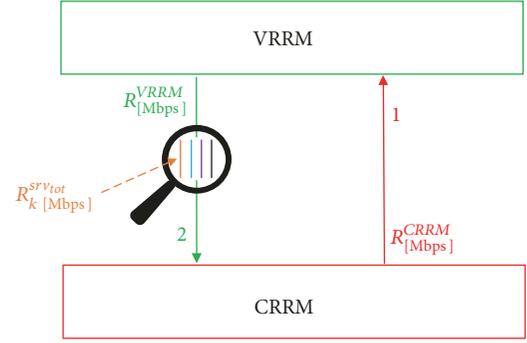


FIGURE 3: Interaction between CRRM and VRRM.

contracted SLAs and serving priorities, CRRM has to provide VRRM with the information about the maximum aggregated capacity, which is accessible from all available RATs. Then, based on this information, VRRM calculates the optimum way of distributing capacity among the different services, to satisfy their requirements. One should note that CRRM does not pass the information about the available capacity of individual RATs to VRRM, since VNOs do not own the physical infrastructure, and VRRM is not concerned about dealing with the management of the resources of the various RATs.

On the other hand, the total demanded data rate of each service, which is calculated by VRRM, should be assigned to suitable RATs, according to the specification of each access technology and requirements of different services. This way, VRRM has also influence on the decision procedure of CRRM. It is also notable that VRRM does not pass the information about the policies of individual VNOs for serving their users, such as contracted SLAs, to CRRM. Figure 3 represents the level of interaction between CRRM and VRRM.

Information about the available aggregated capacity, R^{CRRM} , takes place in the first step, provided by CRRM as a VRRM decision-making parameter. Then in the next step, VRRM calculates the total demanded capacity of each service, R^{Srv}_{k} , to be returned to CRRM for RAT selection process; R^{VRRM} represents the aggregation of all the calculated demands performed by VRRM.

3.3. Service-to-RAT Assignment Mechanism of CRRM. In order to take advantage of the specifications of different technologies in a multi RAT network and to achieve multiplexing gain, it is necessary to define a precise cooperation mechanism among RATs. The proposed technique used for mapping demand onto the available capacity of each RAT is fundamentally a *network-centric* approach, since the main objective is to maximise the *global* utilisation of radio resources. However, since the contracted SLAs are considered as a decision-making parameter in the calculation of demands by VRRM, a preestablished user's satisfaction level is also projected as a constraint to be satisfied by the CRRM model. A *service-based* policy is applied to the process, in order to ensure that mapping the various services onto the

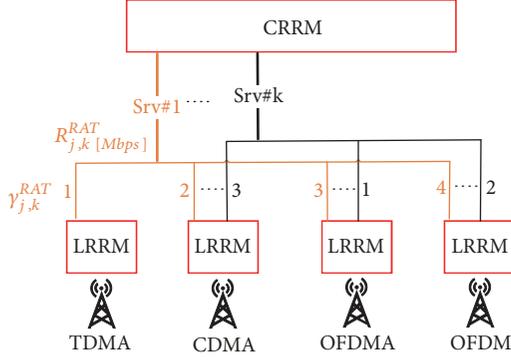


FIGURE 4: CRRM mechanism of mapping demands to the most suitable RATs.

different access technologies is based on the feasibility and suitability of both characteristics [31]; e.g., it is not feasible to serve a video streaming user with GSM, or it is preferable to use Wi-Fi for file sharing. Accordingly, for a specific service type, a prioritised list of suitable RATs is defined in the model to arrange the available common radio resources [32].

To address the assumptions regarding the maximisation of the total utilisation of radio resources, as well as the distribution of the traffic load among existing RATs by considering a service-based policy, the objective of CRRM can be expressed as a weighted linear function, f_{CRRM} , which maps a portion of the available capacity from each RAT onto the most suitable service:

$$\max_{\mathbf{R}^{RAT}} f_{CRRM}(\mathbf{R}^{RAT}) = \max_{\mathbf{R}^{RAT}} \sum_{j=1}^{N^{RAT}} \sum_{k=1}^{N^{srv}} \gamma_{j,k}^{RAT} \frac{R_{j,k}^{RAT} [\text{Mbps}]}{R_{CRRM}^{RAT} [\text{Mbps}]} \quad (6)$$

where

- (i) N^{RAT} : number of RATs,
- (ii) $\gamma_{j,k}^{RAT}$: assigned weight to RAT j , for performing service k , to define service-to-RAT allocation priorities,
- (iii) $R_{j,k}^{RAT}$: assigned data rate from RAT j to service k ,
- (iv) \mathbf{R}^{RAT} : vector of assigned data rates:

$$\mathbf{R}^{RAT} = \left[R_{1,1}^{RAT}, R_{1,2}^{RAT}, \dots, R_{1,N^{srv}}^{RAT}, \dots, R_{N^{RAT},1}^{RAT}, R_{N^{RAT},2}^{RAT}, \dots, R_{N^{RAT},N^{srv}}^{RAT} \right]^T \quad (7)$$

Figure 4 represents the mechanism of f_{CRRM} in distributing the traffic load associated with the different services, among the existing RATs. When a service demand is made to CRRM, the traffic load is distributed among the available RATs, with a higher portion of the load directed to the most suitable ones, in case there is enough capacity from those RATs to serve the demand. Accordingly, the services with higher priority have the possibility of using the most suitable RATs, while the remaining demand from low priority services is more likely to be directed to less suitable ones, especially under an extreme case of high traffic load.

Furthermore, there are some constraints that should be taken into consideration while optimising the objective function. An important constraint is the one that projects the demanded data rate of each service from VRRM and plays an important role in the interaction between CRRM and VRRM. From Figure 4, it is noticeable that the distribution of traffic load for each service among the different RATs should sum to the demanded data rate of that specific traffic slice, provided by VRRM; it should not be either higher, since the extra capacity will be wasted, or lower, because contracted SLAs cannot be addressed. This constraint can be expressed as follows:

$$\sum_{j=1}^{N^{RAT}} R_{j,k}^{RAT} [\text{Mbps}] = R_{k[\text{Mbps}]}^{srv,tot} \quad (8)$$

The next constraint is related to the capacity of each RAT. The aggregation of demands that are directed to each RAT cannot exceed the total available capacity of that RAT, which can be written as

$$\sum_{k=1}^{N^{srv}} R_{j,k}^{RAT} [\text{Mbps}] \leq R_{j[\text{Mbps}]}^{RAT,tot} \quad (9)$$

where

- (i) $R_j^{RAT,tot}$: total available capacity of RAT j .

The last constraint ensures that all directed loads from services to RATs are positive values:

$$R_{j,k}^{RAT} [\text{Mbps}] \geq 0 \quad (10)$$

The simplex method is adopted to find the optimal problem solution. Instead of evaluating all the possible candidates in the feasible solution space that satisfy constraints, this method examines just the *better* candidates, which are known in advance, yielding larger values of the objective function [33].

3.4. Calculating the Aggregated Capacity. The achievable data rate of each RRU depends on the SINR level of its associated user. Based on the model suggested in [34], considering an interference limited Het-Net, with the assumption that users experience channel fading with a Rayleigh Distribution (hence, the received power being described by the Exponentially Distribution), the probability of having SINR higher than an arbitrary value is derived. The authors in [17] further modified the model to find the Probability Distribution Function (PDF) of a single RRU's data rate, assuming that it varies between minimum and maximum positive values:

$$P_{R[\text{Mbps}]}(R_j^{RRU} | 0 \leq R_j^{RRU_L} \leq R_j^{RRU} \leq R_j^{RRU_H}) = \frac{e^{-(0.2/\alpha_p) \ln(10) \sum_{m=0}^5 a_m (R_j^{RRU_L} [\text{Mbps}])^m} - e^{-(0.2/\alpha_p) \ln(10) \sum_{m=0}^5 a_m (R_j^{RRU} [\text{Mbps}])^m}}{e^{-(0.2/\alpha_p) \ln(10) \sum_{m=0}^5 a_m (R_j^{RRU_L} [\text{Mbps}])^m} - e^{-(0.2/\alpha_p) \ln(10) \sum_{m=0}^5 a_m (R_j^{RRU_H} [\text{Mbps}])^m}} \quad (11)$$

where

- (i) $R_j^{RRU_L}$: lower bound for the RRU's data rate,

TABLE 1: Factor of polynomial approximation (updated from [35]).

Coefficients	TDMA (GSM)	CDMA (UMTS)	OFDMA (LTE)	OFDM (Wi-Fi)
a_0	7.95	-12.15	0.24	1.38
a_1	964.6	1.89	264.1	189.3
a_2	-293.8	-0.041	34.96	29.76
a_3	0	0	-274.4	-351.8
a_4	-871.2	0	-52.77	-38.91
a_5	109.9	0	309.8	411.5

- (ii) $R_j^{RRU_H}$: higher bound for the RRU's data rate,
- (iii) α_p : path loss exponent, where $\alpha_p \geq 2$,
- (iv) a_m : coefficients of a 5-degree polynomial, based on real network logs, Table 1.

The goal of deriving the PDF of data rate for each RRU is to find the total capacity of each RAT, as well as the aggregated network capacity, being provided from CRRM to VRRM. Therefore, the next step is to obtain the distribution functions associated with the total capacity of each existing RAT. Assuming that RAT j can assign a N_j^{RRU} number of RRUs to the connected users and that all channels are independent, the data rates of all RRUs, which are random variables, are independent as well. Therefore, the PDF of the accumulated data rate of each access technology can be expressed as the convolution of all its RRU's PDFs [36]:

$$\begin{aligned}
 P_{R[\text{Mbps}]}(R_j^{RAT_{tot}}) &= P_{R[\text{Mbps}]}(R_j^{RRU_1}) \\
 &\quad * P_{R[\text{Mbps}]}(R_j^{RRU_2}) * \dots \\
 &\quad * P_{R[\text{Mbps}]}(R_j^{RRU_{N_j^{RRU}}})
 \end{aligned} \quad (12)$$

According to the Central Limit Theorem, the convolution PDF of statistically independent random variables tends to a Normal Distribution, as the number of random variables tends to infinity, regardless of the initial PDFs [36]. Since the random variables in this work are RRUs, which are proposed to vary between a minimum and a maximum value for all the RATs, (12) is approximated by fitting a Truncated Normal Distribution for each RAT [37]. Assuming that $R_j^{RAT_{tot}}$ has a Normal Distribution, $R_j^{RAT_{tot}} \sim N(\bar{R}_j, \sigma_j)$, and that the interval for R_j^{RRU} is $R_j^{RRU} \in [R_j^{RRU_H}/2, R_j^{RRU_H}]$, then, the distribution parameters associated with each RAT within 95% confidence bounds are given in Table 2, where $R_j^{RAT_{tot}} \in [R_{b_j}^{min}, R_{b_j}^{max}]$.

The general steps taken to obtain the convolution PDF are summarised in what follows. First, the relationship between SINR and the data rate of a single RRU is obtained. The information for UMTS and LTE is based on real measurements provided by Portuguese mobile operators [38], the rest being obtained according to the theoretical behaviour of SINR as a function of data rate. Considering that data rate is constant in given SINR intervals, SINR can be represented as

TABLE 2: Parameters of the truncated normal distributions.

RATs	$R_{b_j[\text{Mbps}]}^{min}$	$R_{b_j[\text{Mbps}]}^{max}$	$\bar{R}_{b_j[\text{Mbps}]}$	$\sigma_j[\text{Mbps}]$
OFDM (Wi-Fi)	3352	6704	5116	671
OFDMA (LTE)	2400	4800	3655	61.2
CDMA (UMTS)	44.1	88.2	66.2	1.63
TDMA (GSM)	0.62	1.24	0.94	0.053

a piecewise step function of data rate and further estimated by fitting a continuous fifth-degree polynomial, using the least-square technique. Then, following (11), the PDF of a single RRU is calculated, and finally the aggregation of PDFs in the convolution function of each RAT (12) yields the result.

4. Case Study Scenario

Model performance is evaluated by defining a case study scenario. It is assumed that the area under analysis is 1 km², which is uniformly covered by all cellular RATs as shown in Figure 5(a). Furthermore, a Wi-Fi Access Point (AP) based on IEEE802.11ac is placed at the centre of each LTE cell site to boost capacity. These assumptions of cell layout are just practical examples that are common among the studies that discuss the problem of resource management in 5G or future implementations from a flow level perspective [39, 40]. In this regard, the aim of considering a scenario with coexisting different RATs is to evaluate the performance of CRRM in terms of mapping service demands onto RATs according to the suitability and loading factor of these RATs. Concerning the distribution of users as presented in Figure 5(b), 25% are gathered around the *centre* (*ce*) of LTE Base Stations (BSs), while the remaining 75% are located in *off-centre* (*oc*) areas, without access to Wi-Fi, but with full access to cellular RATs. Each user performs one specific service at a time.

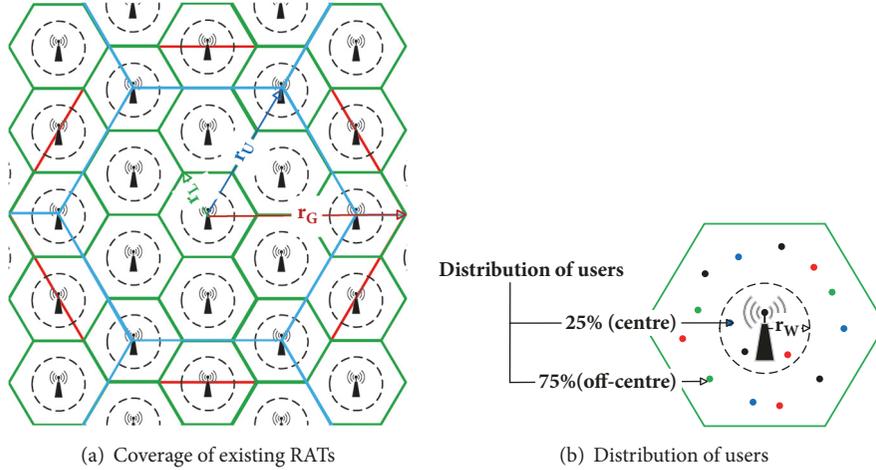
Specifications are summarised in Tables 3 and 4, where r_j^{Cell} is the coverage radius of a single BS or AP, $R_j^{RRU_{max}}$ being the maximum data rate of each RRU from different RATs. However, these values are not achievable in practice, as the quality of the physical channel is greatly influenced by SINR. Using the proposed convolution PDFs to calculate the aggregated capacity of each RAT, by choosing $\alpha_p = 3.8$ as a common value for urban outdoor environments [17], and then randomly selecting from the associated convolution PDF of RATs, $R_j^{RAT_{tot}}$ as the average value for the total available capacity of each access technology *per km²* can be obtained.

TABLE 3: RAT specifications.

RAT	# BS, AP	r_j^{Cell}	$r_{j[km]}^{Cell}$	RRU	$R_{j[Mbps]}^{RRU_{max}}$	$R_{j[Mbps]}^{RAT_{tot}}$
OFDM (Wi-Fi)	16	r_W	0.05	Carrier	0.97	504
OFDMA (LTE)	16	r_L	0.4	Res. block	0.75	350.2
CDMA (UMTS)	~ 1.7	r_U	1.2	Code	1.4	6.54
TDMA (GSM)	1	r_G	1.6	Time-slot	0.059	0.10

TABLE 4: Service parameters.

Service	Range of data rate [Mbps]		User mix [%]	λ_k	SLA
	Centre	Off-centre			
Voice (Vo) <i>Conversational</i>	[0.032, 0.064]		20	50	GB
Video streaming (Vi) <i>Streaming</i>	[2, 13]		50	30	GB
Web browsing (We) <i>Interactive</i>	[1, 862.4]	[1, 358.4]	20	6	BG
Email (Em) <i>Background</i>	[0, 862.4]	[0, 358.4]	10	1	BE

FIGURE 5: Network layout for the reference scenario ($r_G=1.6$ km, $r_U=1.2$ km, $r_L=0.4$ km, and $r_W=0.05$ km).

Summing up over all these values, one gets R^{CRRM} , as 860.8 Mbps. In order to find the maximum achievable capacity that can be provided to *oc* users, Wi-Fi should be excluded from the process, leading to 358.4 Mbps for $R^{CRRM_{cell}}$.

Concerning the assumptions for service parameters and SLAs, one service from each class is defined to be served by a single VNO. One can see in Table 4 that *Voice (Vo)* and *Video streaming (Vi)* are categorised as GB services, since they are delay sensitive and demanding almost constant data rates. Therefore, allocating capacities higher than the contracted ones will not improve their QoS [41]. On the other hand, an increase in the assigned data rate of *Web browsing (We)* and *Email (Em)* can indeed improve the users' quality of experience. The choice of these four services is according to the fact that up to 2023 more than 80% of the mobile traffic will be comprised of these services [42]. Also, concerning the assumptions for allocating the most suitable RATs to different services, the prioritised table of RAT selection for the proposed services is presented in Table 5. For each service, the RAT that is numbered as 1 is the most preferred access technology to be associated with

that service, NA representing the case that is not feasible to associate a particular service to a specific RAT.

5. Analysis of Results

The numerical trend of data rate assignment, for *ce* and *oc* users, is shown in Figure 6. It is clear that, for the same service, the data rate allocated to *ce* users is always higher than or equal to the one of *oc* users, since Wi-Fi coverage is only available in *ce* areas. When the number of users is comparatively low, all GB users are well served; as the traffic load increases, the data rates of *oc* users drop to the lowest contracted level (defined in Table 4).

Th_{BE} is the point where the algorithm stops serving the BE users in order to be able to continue serving the higher priority ones with the minimum guaranteed data rates. It is also noticeable that *Vo* is the last service to drop from 64 kbps to 32 kbps just before Th_{BE} , since it has the highest priority and the lowest demanded data rate compared to the other services.

TABLE 5: Prioritised table of service to RAT assignment.

Services	Priority of RATs			
	TDMA (GSM)	CDMA (UMTS)	OFDMA (LTE)	OFDM (Wi-Fi)
Voice	1	2	3	4
Video streaming	NA	3	1	2
Web Browsing	4	3	2	1
Email	4	3	2	1

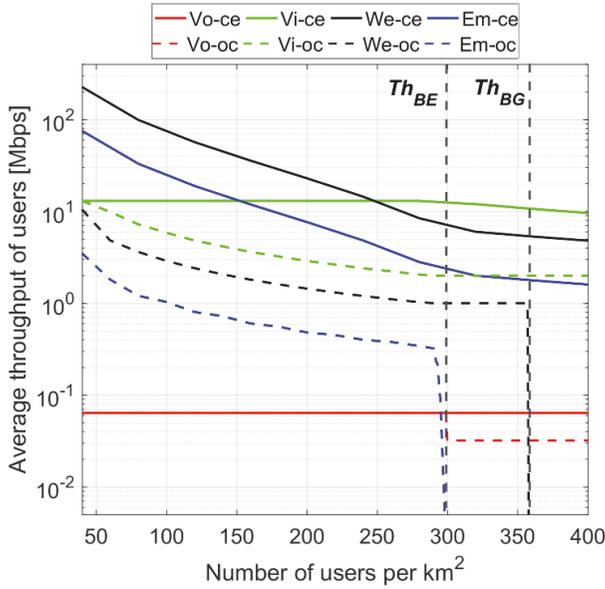


FIGURE 6: Average long-term data rate of users.

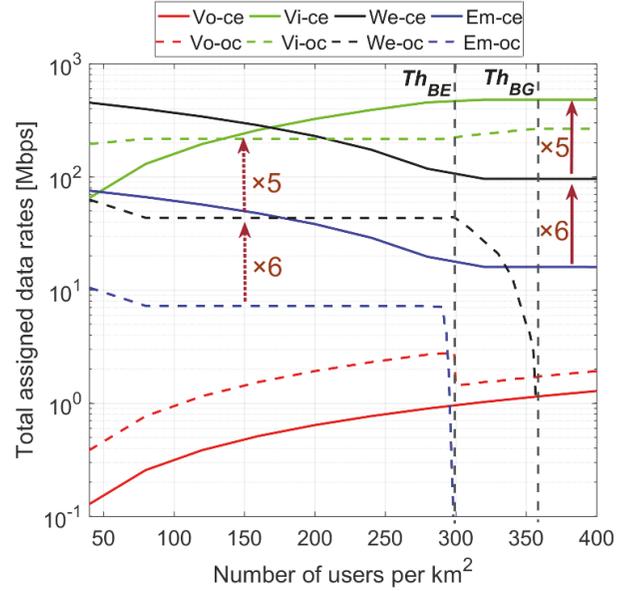


FIGURE 7: Total allocated capacity to the services.

Figure 7 shows the total data rate assigned to each service, as another evaluation metric. The results from this figure, including the different thresholds, are compatible with those of Figure 6. When there is no limitation for data rate assignment in both *ce* and *oc* areas, the ratio of the total data rate allocated to *Vi*, *We*, and *Em* is proportional to their service weights, 30, 6, and 1, respectively. The total used network capacity, one of the most important parameters, is the aggregation of data rate values at each point for both *ce* and *oc* users, being almost equal to the average total available capacity of VRRM obtained per km^2 . Hence, it is shown that, independent of the variation of traffic load and configuration of several parameters in the proposed scenario, the VRRM algorithm is capable of distributing all the available capacity to address the demanded data rates based on the proposed SLAs.

The percentage of served users is shown in Figure 8. As all users are served before Th_{BE} , the values for both *ce* and *oc* in this part remain constant, being associated with the ones defined for the user mix in Table 4. The percentage of served *ce* users in each service does not change by increasing the number of offered users up to 400, as capacity is enough to serve them all. However, at Th_{BE} , the values for *Em-oc* go to zero, meaning that all BE users located in *oc* are delayed in order to provide enough capacity to serve the rest of *oc* users, leading to a slight increase in the values of other services.

After Th_{BE} , the algorithm starts discarding *We-oc*, as BG users. At this point, since a minimum data rate is guaranteed for BG users, the algorithm delays just enough number of *We-oc* ones to release capacity for the rest of *oc* users to continue their service with the minimum contracted data rates, in order to maximise serving the higher priority users up to Th_{BG} , where no BE users left to be delayed. The same process takes place after Th_{BG} for GB services, starting from *Vi-oc*. It is also noticeable that at Th_{BG} , all available capacity at *oc* is divided between *Vi* and *Vo*, the number of *Vi* users being 2.5 times higher than *Vo* ones, which is proportional to their ratio of traffic share (50% and 20% for *Vi* and *Vo*, respectively).

In order to evaluate the performance of the CRRM mechanism of service-to-RAT assignment, as well as the interaction between the CRRM and VRRM, the procedure of allocating capacity from the available RATs, to satisfy the demanded data rate of each service, is shown in Figure 9.

Starting from *Vo* in Figure 9(a), one can see that the demanded data rate is provided by GSM and UMTS in both *ce* and *oc* areas. As GSM is more suitable to serve voice users, the CRRM algorithm gives the higher priority to this RAT. However, the demanded data rate of voice is higher than the total available capacity of GSM; therefore, the rest of *Vo* traffic load is served by UMTS, which is the second highest priority RAT. One can also notice that before Th_{BE} , since there is no

TABLE 6: RATs and services matching, before Th_{BG} .

RAT	Voice		Video		Web		Email	
	Ce	Oc	Ce	Oc	Ce	Oc	Ce	Oc
TDMA (GSM)	✓	✓	NA	NA	×	×	×	×
CDMA (UMTS)	✓	✓	×	×	×	✓	✓	✓
OFDMA (LTE)	×	×	✓	✓	×	✓	×	✓
OFDM (Wi-Fi)	×	NA	✓	NA	✓	NA	✓	NA

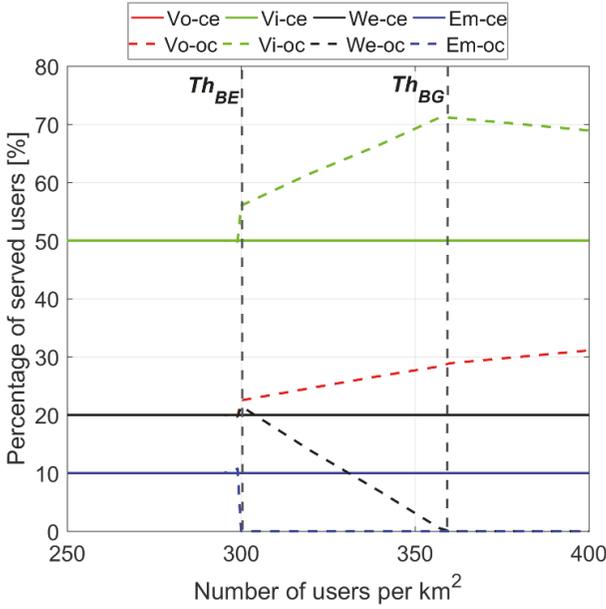


FIGURE 8: Percentage of served users.

limitation in data rate assignment, capacity share of *oc* users for both RATs is three times higher than the *ce* ones, which is in accordance with the ratio of traffic (i.e., 75% and 25% accordingly) in these two areas.

Regarding the *Vi* demand in Figure 9(b), although the number of *oc* users is three times higher than *ce* ones, the demanded data rate in centre is mostly higher, except for lower traffic rates, since this group of users have access to Wi-Fi. Accordingly, for *oc* users, the available LTE capacity as the most preferred RAT is enough to serve the demanded data rates until Th_{BE} . After that, *oc* users have to start using capacity from the second available preferred RAT, i.e., UMTS, while at each point for *ce* users the capacity of LTE and Wi-Fi is sufficient for the total demand.

For *We* users, the allocated capacity from different RATs is presented in Figure 9(c). It is noticeable that the whole traffic of *ce* users can be served just by Wi-Fi as the most preferred RAT. Regarding the *oc* users, before Th_{BE} , LTE as the most suitable RAT can cover the demand. However, by increasing the required data rate of *Vi-oc* users between Th_{BE} and Th_{BG} , there will not be enough capacity from LTE to serve the rest of *oc* traffic; therefore, CRRM will choose the second suitable RAT, UMTS, besides LTE to address all the demanded capacity of *We-oc* users.

The last service is *Em*, which is shown in Figure 9(d). One can see that the demand of *ce* users is mostly responded

by Wi-Fi, except for a tiny portion which is covered by UMTS. However, for *oc* users the whole requested data rate is allocated from LTE and UMTS, which are the first and second preferred available RATs, respectively. The absence of LTE as the second priority RAT in *ce* areas is due to the fact that *Em* is categorised as a BE service; therefore, it has the lowest priority, and the unallocated capacity of LTE is not sufficient to cover both demands from *ce* and *oc* users. Alternatively, a small part of *ce* users' demand is directed to UMTS as the third preferred choice for RAT selection, besides Wi-Fi.

A summary of matching RATs and services is presented in Table 6 as a comparison with the assumptions of Table 5. One can see from Table 6 that at most two RATs are assigned to a specific service, for both *ce* and *oc* areas, which are the highest priority ones, except for a small share of *Em-ce* that uses UMTS as the third priority. It is also notable that for *We-oc* and *Em-oc* users, since the first priority RAT is Wi-Fi, which is not accessible for *oc* users, they are alternatively connected to LTE and UMTS, as the second and third most suitable RATs, respectively.

6. Conclusions

This paper proposes a centralised cooperative mechanism of RRM for Virtual RANs, based on aggregation and virtualisation of all the available radio resources from different RATs. In this context, the roles of VNOs and InPs are separated, while they are interacting to maximise the overall performance efficiency. VNOs do not own the underlying infrastructure and accordingly are not dealing with the management of radio resources. Alternatively, they demand a centralised virtualisation platform, called VRRM, for *Capacity-as-a-Service* in order to satisfy their users in respect to the contracted SLAs. VRRM is in charge of managing the aggregated capacity, which is provided by InPs through the CRRM entity. On the other hand, CRRM is responsible for distributing the heterogeneous traffic among the available RATs, based on some decision criteria for defining the most suitable service-to-RAT assignment. The model is further extended to handle the extreme situations resulting from high traffic loads, when the capacity is not sufficient to provide all users with at least the minimum contracted data rate. In this case, the algorithm introduces delay to the lowest priority users, releasing enough capacity so that the rest of higher priority ones continue their service.

The main goal from the VRRM perspective is to maximise the utilisation of total capacity, which is provided by aggregating the RRU from different RATs in order to satisfy the SLAs to the highest achievable level, while maintaining a level

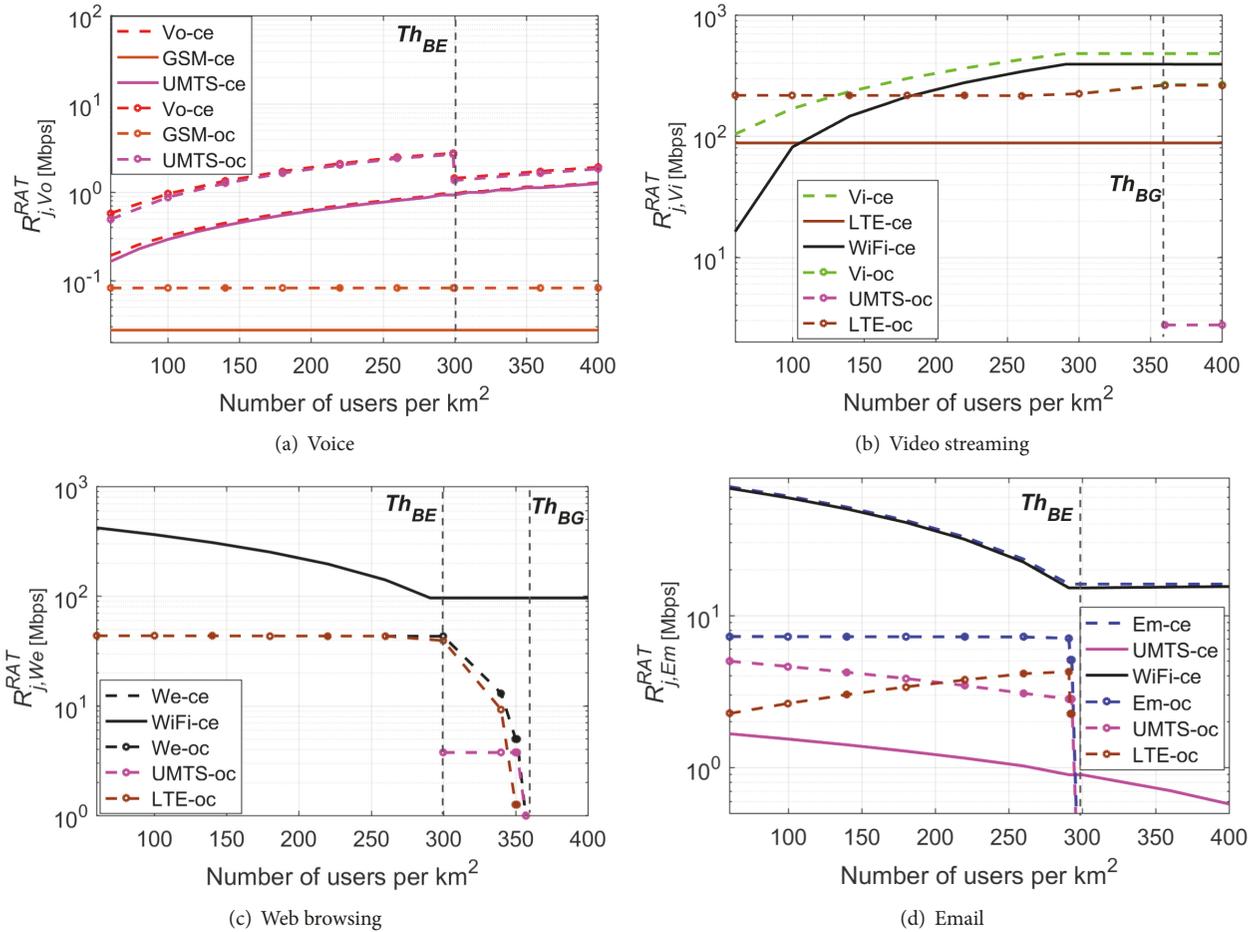


FIGURE 9: Allocation of capacity from the available RATs to each service.

of fairness among users considering the priority of services and to realise an acceptable level of isolation. To achieve this goal, VRRM has to closely interact with CRRM in order to calculate the demanded data rates of each service based on the information about the available capacity provided by CRRM, and then returning the demanded capacity to CRRM, to be provided by the available RATs. The key objective of CRRM is to associate these demands to the most suitable RATs, in order to meet the QoS requirements of each service.

The performance of the proposed model is analysed by realising a practical scenario. While all users have access to the cellular RATs, only 25% of them can use Wi-Fi. The results from the VRRM evaluation confirm that when there is no shortage of resources, the model can capture the demanded capacity of *Vo*, *Vi*, *We*, and *Em* users, according to the concept of proportional fairness, which is consistent with the predefined serving weights of 50, 30, 6, and 1, respectively. Moreover, all SLAs are satisfied independent of the variation of traffic load, which yields a desired level of isolation. Under extreme traffic loads in *oc* areas, the algorithm delays the users based on their service priorities, trying to serve the maximum number of high priority ones. In any case, 100% of the aggregated capacity is used. From a CRRM viewpoint, all the

demanded capacities from different services are addressed by maximum two available RATs, which are the highest priority ones according to the predefined table of service-RAT matching, the only exception being a small portion of BE users' demand, since they have the lowest priority.

Data Availability

The software used for solving the optimisation problems, CVX, is publicly available online at this address: <http://cvxr.com/cvx/>. The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work is developed within the framework of the COST Action CA15104, the Inclusive Radio Communication Networks for 5G and beyond (IRACON).

References

- [1] S. Glisic, *Advanced Wireless Networks: Technology and Business Models*, John Wiley and Sons, Chichester, UK, 2016.
- [2] H. Wen, P. K. Tiwary, and T. Le-Ngoc, *Wireless Virtualization*, Springer, Montreal, Canada, 2013.
- [3] P. Marsch, O. Bulakci, O. Queseth, and M. Boldi, *5G System Design: Architectural and Functional Considerations and Long Term Research*, John Wiley and Sons, Hoboken, NJ, USA, 2018.
- [4] W. Xiang, K. Zheng, and X. Shen, *5G Mobile Communications*, Springer International Publishing, Zürich, Switzerland, 2017.
- [5] C. Liang and F. Yu, "Enabling 5G mobile wireless technologies," *EURASIP Journal on Wireless Communications and Networking*, vol. 2015, no. 218, 2015.
- [6] J. Lucena, P. Ameigeiras, D. Lopez, J. J. Ramos-Munoz, J. Lorca, and J. Folgueira, "Network slicing for 5G with SDN/NFV: concepts, architectures, and challenges," *IEEE Communications Magazine*, vol. 55, no. 5, pp. 80–87, 2017.
- [7] A. Aijaz, "Towards 5G-enabled Tactile Internet: Radio resource allocation for haptic communications," in *Proceedings of the 2016 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1–6, Doha, Qatar, April 2016.
- [8] G. Tseliou, F. Adelantado, and C. Verikoukis, "Scalable RAN virtualization in multitenant LTE-A heterogeneous networks," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 8, pp. 6651–6664, 2016.
- [9] R. Riggio, A. Bradai, D. Harutyunyan, T. Rasheed, and T. Ahmed, "Scheduling wireless virtual networks functions," *IEEE Transactions on Network and Service Management*, vol. 13, no. 2, pp. 240–252, 2016.
- [10] M. Farooq, H. Ghazzai, E. Yaacoub, A. Kadri, and M. Alouini, "Green virtualization for multiple collaborative cellular operators," *IEEE Transactions on Cognitive Communications and Networking*, vol. 3, no. 3, pp. 420–434, 2017.
- [11] M. Salem, A. Adinoyi, M. Rahman, H. Yanikomeroğlu, D. Falconer, and Y.-D. Kim, "Fairness-aware radio resource management in downlink OFDMA cellular relay networks," *IEEE Transactions on Wireless Communications*, vol. 9, no. 5, pp. 1628–1639, 2010.
- [12] X. Yu and H. Zhu, "Optimal resource management with delay differentiated traffic and proportional rate constraint in heterogeneous networks," *Journal of Communications*, vol. 9, no. 9, pp. 714–722, 2014.
- [13] C. Y. Wong, R. S. Cheng, K. B. Letaief, and R. D. Murch, "Multiuser OFDM with adaptive subcarrier, bit, and power allocation," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 10, pp. 1747–1758, 1999.
- [14] X. Costa-Perez, J. Swetina, T. Guo, R. Mahindra, and S. Rangarajan, "Radio access network virtualization for future mobile carrier networks," *IEEE Communications Magazine*, vol. 51, no. 7, pp. 27–35, 2013.
- [15] S. Singh, S. Yeh, N. Himayat, and S. Talwar, "Optimal traffic aggregation in multi-RAT heterogeneous wireless networks," in *Proceedings of the IEEE International Conference on Communications Workshops (ICC '16)*, pp. 626–631, Kuala Lumpur, Malaysia, May 2016.
- [16] M. Gerasimenko, D. Moltchanov, R. Florea et al., "Cooperative radio resource management in heterogeneous cloud radio access networks," *IEEE Access*, vol. 3, pp. 397–406, 2015.
- [17] S. Khatibi and L. M. Correia, "Modelling virtual radio resource management in full heterogeneous networks," *EURASIP Journal on Wireless Communications and Networking*, vol. 2017, no. 73, 2017.
- [18] L. Caeiro, B. Rouzbehani, and L. M. Correia, "A fair mechanism of virtual radio resource management in multi-RAT wireless het-nets," in *Proceedings of the 2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pp. 1–5, Montreal, QC, Canada, October 2017.
- [19] M. Richart, J. Baliosian, J. Serrat, and J.-L. Gorricho, "Resource slicing in virtual wireless networks: a survey," *IEEE Transactions on Network and Service Management*, vol. 13, no. 3, pp. 462–476, 2016.
- [20] B. Rouzbehani, L. M. Correia, and L. Caeiro, "Radio resource and service orchestration for virtualised multi-tenant mobile het-nets," in *Proceedings of the 2018 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1–5, Barcelona, Spain, April 2018.
- [21] R. Ferrus, O. Sallent, J. Perez-Romero, and R. Agusti, "On 5G radio access network slicing: radio interface protocol features and configuration," *IEEE Communications Magazine*, vol. 56, no. 5, pp. 2–10, 2018.
- [22] C. Liang and F. Yu, "Wireless network virtualization: a survey, some research issues and challenges," *IEEE Communications Surveys and Tutorials*, vol. 17, no. 1, pp. 358–380, 2015.
- [23] W. Song and W. Zhuang, *Interworking of Wireless LANs and Cellular Networks*, Springer, New York, NY, USA, 2012.
- [24] Q. Ye, B. Rong, Y. Chen, M. Al-Shalash, C. Caramanis, and J. G. Andrews, "User association for load balancing in heterogeneous cellular networks," *IEEE Transactions on Wireless Communications*, vol. 12, no. 6, pp. 2706–2716, 2013.
- [25] S. Khatibi and L. M. Correia, "A model for virtual radio resource management in virtual RANs," *EURASIP Journal on Wireless Communications and Networking*, vol. 2015, no. 1, pp. 1–12, 2015.
- [26] G. Miao, G. Song, and G. Li, *Energy and Spectrum Efficient Wireless Network Design*, Cambridge University Press, Cambridge, UK, 2015.
- [27] "CVX – Software for Disciplined Convex Programming," <http://cvxr.com>, Feb. 2017.
- [28] S. Mehrotra, "On the implementation of a primal-dual interior point method," *SIAM Journal on Optimization*, vol. 2, no. 4, pp. 575–601, 1992.
- [29] D. Marabissi and R. Fantacci, "Heterogeneous public safety network architecture based on RAN slicing," *IEEE Access*, vol. 5, pp. 24668–24677, 2017.
- [30] M. Jiang, M. Condoluci, and T. Mahmoodi, "Network Slicing Management and Prioritization in 5G Mobile Systems," in *Proceedings of the European Wireless 2016 – 22th European Wireless Conference*, Oulu, Finland, 2016.
- [31] E. Hossain, *Heterogeneous Wireless Access Networks: Architectures and Protocols*, Springer, Boston, MA, USA, 2009.
- [32] G. Piao, *Radio Resource Management for Integrated Services in Multi-radio Access Networks*, Springer, Kassel, Germany, 2007.
- [33] H. Chen, L. Huang, S. Kumar, and C. J. Kuo, *Radio Resource Management for Multimedia QoS Support in Wireless Networks*, Springer, Boston, MA, USA, 2004.
- [34] H. S. Dhillon, R. K. Ganti, F. Baccelli, and J. G. Andrews, "Modeling and analysis of K-tier downlink heterogeneous cellular networks," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 3, pp. 550–560, 2012.

- [35] S. Khatibi and L. M. Correia, "Modelling of virtual radio resource management for cellular heterogeneous access networks," in *Proceedings of the 2014 IEEE 25th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pp. 1152–1156, Washington DC, USA, September 2014.
- [36] A. Papoulis and S. U. Pillai, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, New York, NY, USA, 4th edition, 2002.
- [37] N. Mattew, O. Sadiku, and A. Warsame, *Signals and Systems: A Primer with MATLAB*, CRC Press, Boca Raton, FL, USA, 2015.
- [38] P. A. Carreira, *Data Rate Performance Gains in UMTS Evolution to LTE at the Cellular Level [Msc Thesis]*, Instituto Superior Técnico, Technical University of Lisbon, Lisbon, Portugal, 2011.
- [39] Mobile Cloud Networking (MCN), "Future Communication Architecture for Mobile Cloud Services," Tech. Rep., 2014, Report No. D4.3, Version 1, https://cordis.europa.eu/project/rcn/105938_en.html.
- [40] T. Shuminoski and T. Janevski, "Radio network aggregation for 5G mobile terminals in heterogeneous wireless and mobile networks," *Wireless Personal Communications*, vol. 78, no. 2, pp. 1211–1229, 2014.
- [41] N. Ferdosian, M. Othman, B. M. Ali, and K. Y. Lun, "Greedy-knapsack algorithm for optimal downlink resource allocation in LTE networks," *Wireless Networks*, vol. 22, no. 5, pp. 1427–1440, 2016.
- [42] Ericsson., *Ericsson Mobility Report*, Stockholm, Sweden, 2017, <https://www.ericsson.com/assets/local/mobility-report/documents/2017/ericsson-mobility-report-november-2017central-and-eastern-europe.pdf>.



Hindawi

Submit your manuscripts at
www.hindawi.com

