A Joint Admission Control & Resource Management Scheme for Virtualized Radio Access Networks

Behnam Rouzbehani[1], Vladimir Marbukh[2], Kamran Sayrafian[2]

[1]IST-University of Lisbon/INESC-ID
Av. Rovisco Pais 1, 1049-001
Lisbon, Portugal

[2]National Institute of Standards & Technology
100 Bureau Drive, Stop 8910
Gaithersburg, MD, USA

Abstract - **A virtualization platform is responsible for allocation and/or aggregation of radio resources from different access technologies owned by the Infrastructure Providers (InPs). It is also responsible for the distribution of the total capacity among Virtual Network Operators (VNOs). This distribution should comply with requirements specified in the Service Level Agreements (SLAs) of each VNO. The admission control process ensures adherence to the minimum service level requirements during network operation i.e. users' arrivals/departures. A joint resource management and admission control methodology is proposed in this paper that achieves optimality with respect to the aggregate system utility. The proposed approach is evaluated by simulating a sequence of scenarios with increasing number of arrivals. The results confirm maximal utilization of the available capacity in the system while all SLAs are satisfied.**

Keywords- Virtualization, Service Level Agreements (SLA), admission control, distributed resource allocation.

I. INTRODUCTION

The use of *service-oriented* architecture in the next generation of mobile network standards enables a variety of potential benefits to both users and network operators. These benefits include (a) more flexibility in sharing and utilization of network resources (b) wider range of customized services to suit users' requirements and (c) reduction in the CAPEX/OPEX costs [1]. *Virtualization* supports service-oriented architecture through decoupling of the services and functionalities from the underlying Radio Access Networks (RANs) [2]. It transforms the physical infrastructure into multiple logical networks that are shared among different *tenants,* i.e., Virtual Network Operators (VNOs), all functioning in an isolated manner. In contrast to traditional Mobile Network Operators (MNOs), VNOs do not own the infrastructure. Instead, they obtain the capacity from a centralized virtualization platform and enforce their own service requirements and policies in the process of Radio Resource Management (RRM) [3].

The diversity in users' Quality of Service (QoS) requirements has driven the emergence of resource slicing along with virtualization [4]. Resource slicing achieves a dual goal of performance isolation for different services as well as optimization. Performance isolation ensures efficiency of independent slices regardless of the variation of different parameters in the network (e.g., traffic load or channel condition) [5]. Performance optimization in virtualized Heterogeneous Networks (Het-Nets) should not only optimize the performance of various slices but also maximize the utilization of the overall shared resources [6]. To the best of the authors' knowledge, there are no comprehensive studies that thoroughly cover key issues such as differentiated service provisioning, performance isolation, and fairness for RRM in virtualized Het-Nets.

Scalability limitations of centralized RRM create a need for decentralized resource management approaches in future mobile communication networks [7], [8]. In [9], a distributed RRM model based on non-cooperative game theory has been proposed for dense 5G networks where each Base Station (BS) tries to maximize its payoff. While this model achieves energy efficiency, it does not consider the customized specifications and requirements of different services. An adaptive two-layer decentralized RRM with slow and fast timescales for adaptation of the central manager and users has been proposed in [10]. However, network virtualization and slicing concepts which are key enablers of 5G have not been considered. Another distributed RRM approach with a focus on multi-connectivity in 5G has been described in [11]. While the proposed approach aims at reducing the processing costs and signalling overhead, it lacks the notion of RAN slicing, isolation, as well as service orientation.

This paper proposes a joint admission control and RRM for virtualized RANs by extending the scheme presented in [12]. The system architecture follows the concept of network slicing proposed in the latest release of 3GPP standards for 5G service-oriented architectures [13]. The distributed resource management approach described in [12] overcomes the scalability issues of the centralized RRM discussed in [14], [15]. It maximizes the aggregate system utility using a two-stage distributed optimization with pricing adaptation on a *fast* and *slow* time scales [16], [17]. At the faster time-scale, and assuming that VNO capacities do not change, users adjust their rates based on the congestion pricing. At the slower time scale each VNO adjusts its own capacity according to its assigned congestion price. This is done subject to the total aggregate capacity of the system. This decentralization takes advantage of the dual role of congestion prices that are used for both adjustment of the users' rates and VNO capacity expansion/reduction.

The scheme proposed in [12] doesn't consider admission control as a part of the optimization process. Joint admission control and resource management with user's rate constraint using aggregate utility maximization generally requires some degree of centralization [18]. The joint optimization proposed in this paper assumes a limited form of centralization which leads to higher system utility.

The rest of this paper is organized as follows. Section II describes the system architecture. Section III formulates the constrained aggregate utility maximization problem for joint admission control and resource management. Sections IV and V describe the resource management scheme and the admission control strategy in details. Section VI outlines a case study with a simple user arrival/departure process along with the simulation results and relevant discussions. Finally, concluding remarks and directions for future research are expressed in section VII.

II. System Architecture

Figure 1 shows the mechanism of service-oriented RAN slicing and resource management along with interaction of different entities in the system. The Virtual-RRM (VRRM) module is a centralized virtualization platform which is responsible for configuring the RAN protocol stack and QoS metrics according to the slice requirements. Those requirements are enforced by different VNOs based on their specific policies. As an example, assume that VNOs *A* and *B* provide two types of services with different requirements. For *slice A* with high throughput requirements, radio flow *A* is configured to support multi-connectivity. Therefore, slice *A* is using the resources from 2 different radio access points. On the other hand, the network *slice B* is configured with only one connection according to the provided policy.

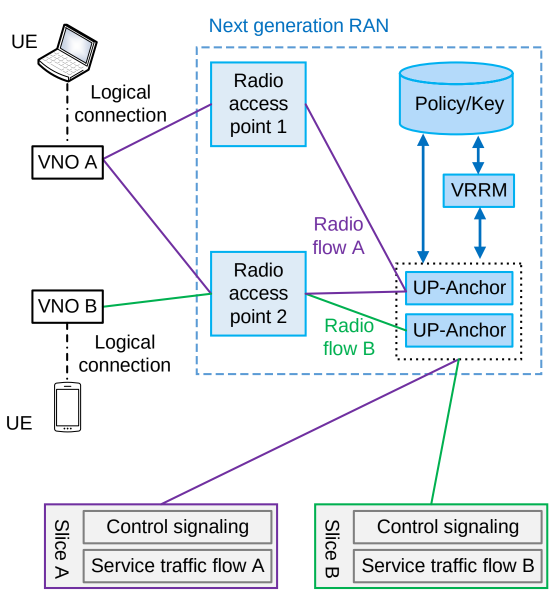

Figure 1. Service-oriented RAN slicing

The User Plane Anchor (UP-Anchor) is responsible for distributing the traffic flow in each slice. A RAN slice is composed of a control plane and a separate data plane. The required capacity allocation is subject to the SLA agreements between the VNOs and users. We consider the following three categories of SLA contracts [12]:

- Guaranteed Bitrate (GB): This is the highest priority category where a minimum threshold for data rate assignment must always be guaranteed regardless of the traffic load variation and network status. In addition, the assigned data rate cannot exceed a maximum threshold for this SLA category.
- Best effort with minimum Guaranteed (BG): This is the second highest priority category for which a minimum level of data rate is guaranteed. Higher data rates are served in a best effort manner if available.
- Best Effort (BE): This is the lowest priority category for which there is no level of service guarantees and users are served in a pure best effort manner.

III. Utility Maximization and Constraints

Let I_{sv} be set of users obtaining service from slice $s = 1,..,S$ of VNO $v = 1,..,V$, where sets I_{sv} with different (s,v) do not overlap, e.g., $I_{km} \cap I_{ln} = \emptyset$ if $(k,m) \neq (l,n)$, $k,l \in \{1,..,S\}$ and $m,n \in \{1,..,V\}$. We assume that the preference of each user $i \in I_{sv}$, for rate R can be quantified by the increasing and concave utility function $u_s(R)$. Here, we consider a logarithmic utility function

$$u_s(R) = \lambda_s \log(R), \tag{1}$$

where parameter $\lambda_s > 0$ characterizes user $i \in I_{sv}$ service weight. Following Network Utility Maximization (NUM) framework [16], we assume that the goal of system management is maximization of the aggregate utility

$$U_\Sigma(R_{svi}) = \sum_{s=1}^{S} \lambda_s \sum_{v=1}^{V} \sum_{i \in I_{sv}} \log(R_{svi}), \tag{2}$$

over vector of rates (R_{svi}) allocated to users $i \in I_{sv}$, $s = 1,..,S$; $v = 1,..,V$. This maximization is a subject to the following capacity and contractual constraints. The total capacity allocated to all users serviced by VNO v, $i \in I_{sv}$, $s = 1,..,S$ cannot exceed the VNO v capacity C_v:

$$\sum_{s=1}^{S} \sum_{i \in I_{sv}} R_{svi} \leq C_v, v = 1,..,V. \tag{3}$$

Also, the aggregate capacity allocated to all VNOs cannot exceed the total system capacity C^{VNNO}:

$$\sum_{v=1}^{V} C_v \leq C^{VNNO}. \tag{4}$$

The above constraints are due to slice s data rate guarantees to a user $i \in I_{sv}$, R_s^{min} and R_s^{max} respectively:

$$0 \leq R_s^{min} \leq R_{svi} \leq R_s^{max}, \quad s = 1,..,S, \quad v = 1,..,V \tag{5}$$

The second set of constraints is due to guarantees on the minimum capacity of each VNO v, $C_v^{min} \geq 0$:

$$C_v \geq C_v^{min}, v = 1,..,V. \tag{6}$$

In the rest of this paper we propose a distributed solution to the aggregate utility (2) maximization:

$$\max_{(C_v)} \max_{(R_{svi})} \sum_{s=1}^{S} \lambda_s \sum_{v=1}^{V} \sum_{i \in I_{sv}} \log(R_{svi}) \tag{7}$$

subject to constraints (3)-(6).

It is known that the solution to this optimization problem results in proportionally fair resource allocation [16]. Due to lower bounds in (5) and (6), optimization problem (3)-(7) may not have a feasible solution. This possibility necessitates an admission control process. Optimization problem (3)-(7) should be solved every time the set of users changes i.e. due to user arrivals/departures. Assuming that resource optimization occurs on a faster time-scale than changes in the set of users, solution to (3)-(7) is discussed in the rest of the paper.

IV. USER RATE & VNO CAPACITY ADAPTATION

This section discusses user rate and VNO capacity adaptation, given that the optimization problem (3)-(7) has a feasible solution for the set of users, i.e., system has sufficient capacity to satisfy minimum rate requirements for all users currently present in the system. Following [17] define the net utility of user $i \in I_{sv}$ to be

$$U_i(R) = \lambda_s \log R - p_v R, \tag{8}$$

where p_v is the price of a unit of bandwidth offered by the VNO v.

By solving the individual convex optimization problem $max_{R\geq 0}\, U_i\,(R)$ subject to constraint (5), each user i calculates its rate as follows:

$$R_i(p_v) = \begin{cases} \lambda_i/p_v & if \;\; R_i^{min} \leq \lambda_i/p_v \leq R_i^{max} \\ R_i^{min} & if \;\; \lambda_i/p_v < R_i^{min} \\ R_i^{max} & if \;\; \lambda_i/p_v > R_i^{max} \end{cases} \tag{9}$$

This solution, which is shown in Figure 2, is determined by condition that the slope of user utility coincides with the current congestion price within the domain of $[R_i^{min}, R_i^{max}]$.

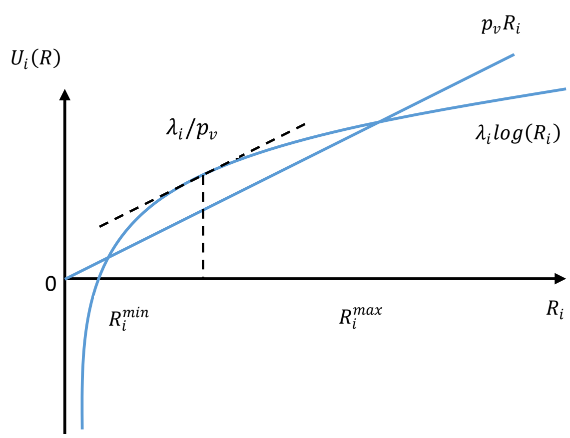

Figure 2. Individual user's rate optimization

Due to the lower bound constraints in (5), rate allocation (9) may not be feasible. In this case, VNO v capacity deficit

$$\sum_{s=1}^{S} \sum_{i \in I_{sv}} R_i^{min} - C_v > 0 \;\; , v = 1,..,V \tag{10}$$

is arbitrarily allocated to users that are currently present in this VNO.

The optimal prices p_v^{opt} that maximize the utilization of the VNOs' available bandwidth are determined by the following distributed adaptive algorithm. The algorithm proceeds in discrete steps $k = \{1,2, \ldots\}$. At each step k, users solve the individual optimization problems resulting in (9). If constraints (3) are satisfied, i.e., the aggregate data rate of the users does not exceed the total capacity of the associated VNO, then in step $k+1$ the price $p_{v,k+1}$ is reduced in order to motivate users to buy more bandwidth. However, if the constraint (1) is not satisfied, $p_{v,k+1}$ is increased, resulting in a decrease of users' data rates. The main idea here is to maximize utilization of the available capacity in an efficient way. The price adaptation model can be expressed as follows [17]:

$$p_{v,k+1} = \left[p_{v,k} + h(\tilde{R}_{vk} - C_{vk})\right]^{+} \tag{11}$$

where

$$\tilde{R}_{vk} = max\left(C_v^{min}, \textstyle\sum_{i \in I_{sv}} R_{ik}\right) \tag{12}$$

$[x]^{+} = \max(0, x)$, and $h > 0$ is a small positive constant which regulates the tradeoff between optimality under stationary scenario and adaptability under non-stationary scenario, e.g., due to changing set of users. The main advantage of this approach is that VNOs do not have to know users' utilities which are considered private information.

In a slower time-scale each VNO adjusts its own capacity by negotiating the price with VRRM. The adaptation of capacities among the tenant VNOs (C_v) is subject to the total available capacity of VRRM is C^{VRRM} (4). The average price of a unit of bandwidth in the entire system at step $k = \{1,2, \ldots\}$ is as follows:

$$P_k^{ave} = \frac{1}{C^{VRRM}} \sum_{v=1}^{V} C_v P_{v,k} \tag{13}$$

where $P_{v,k}$ is the price of a unit of bandwidth assigned to VNO v from VRRM at step k.

We propose the following capacity adaptation algorithm for the VNOs according to [17]:

$$C_{v,k+1} = C_{v,k} + H(P_{v,k} - P_k^{ave}), \tag{14}$$

where $H > 0$ is a small constant.

Algorithm (11)-(14) increases (decreases) the capacity of a VNO if its corresponding price is higher (lower) than the average price (13). However, VNO capacity cannot fall below the lower bound in (6) due to eqaution (12).

V. ADMISSION CONTROL

This section discusses admission control strategy which allows new users if and only if the optimization problem (3)-(7) has a feasible solution with the new set of arrivals. It is shown that this admission control strategy can be implemented with a limited degree of centralization.

Consider the following admission control strategy with limited centralization. At time step k , each VNO v updates Central Controller (i.e., VRRM) on two dynamic variables: potentially accessible capacity $B_{v,k}$ and congestion price $p_{v,k}$. Potentially accessible capacity at time step k is defined as

$$B_{v,k} = C_{v,k} - \max\left(C_v^{min}, G_{v,k}\right) \tag{15}$$

where $C_{v,k}$ is the total capacity of VNO v and $G_{v,k}$ is the total minimum bandwidth guaranteed to all users already present and receiving service by VNO v according to their SLAs at step k. Central Controller evaluates two global variables i.e., the average congestion price (11) and aggregate accessible capacity $B_{\Sigma,k}$ at step k as

$$B_{\Sigma,k} = \textstyle\sum_{v=1}^{V} B_{v,k} \tag{16}$$

and informs all VNOs. VNOs use this information to adjust rates allocated to their active users and make admission decisions on new arrivals. The admission control process is shown by the flowchart in Figure 3.

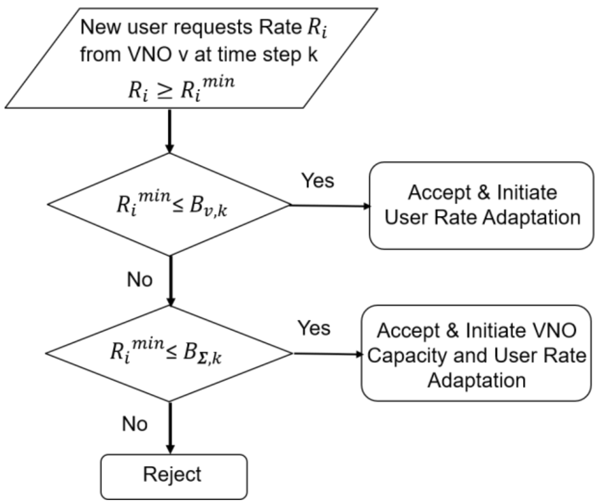

Figure 3. Admission control process

When a VNO does not have sufficient potentially accessible capacity $B_{v,k}$ at the time of a new user arrival, the admission decision will be affirmative only if the user's minimum required rate does not exceed the aggregate potentially accessible capacity $B_{\Sigma,k}$. Otherwise the decision is negative. In the case of affirmative admission decision, the arriving user is immediately admitted to the VNO v if this VNO has sufficient potentially accessible capacity $B_{v,k}$. Otherwise the arriving user is admitted conditionally (e.g. with some delay) until resource allocation algorithm provides the required capacity to this VNO. Note that two different implementations of admission control are possible here. The above implementation requires exchanging dynamic information $B_{v,k}$ and $B_{\Sigma,k}$ since decisions are made locally by each VNO. VNOs can also delegate the admission decisions to the Central Controller. In that case only $B_{v,k}$ updates are needed to be exchanged.

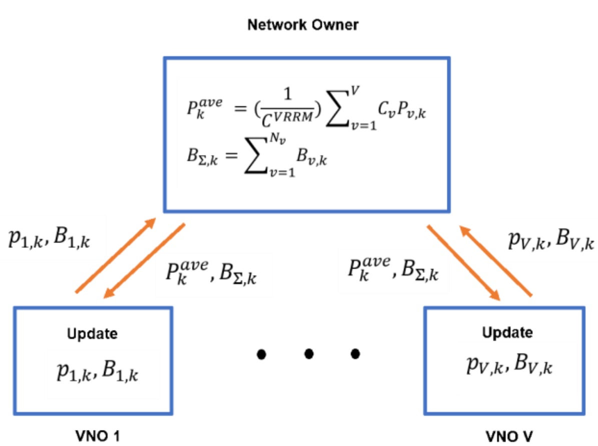

Figure 4. Two-level network management scheme

The joint admission control and RRM scheme operates on two levels i.e. VNO and system (VRRM) levels as shown in Figure 4. At the VNO level, existing users adapt their rates to the VNO bandwidth cost p_v. An arriving user is admitted if the VNO has sufficient potentially accessible capacity B_v to satisfy the user's minimum rate requirement. At the system level, each VNO v adjust its capacity based on this VNO capacity cost p_v and system-wide average bandwidth cost P^{ave}. The admission decision on a new user arrival at a VNO which does not have sufficient potentially accessible bandwidth B_v is based on the system-wide potentially accessible bandwidth B_Σ.

VI. Simulation Scenario & Results

To evaluate our proposed joint admission control and resource management strategy, the simple traffic distribution scenario in [14] with VRRM capacity of 510 Mbps has been considered in this section. Network parameters are defined in Table 1. It is assumed that 3 VNOs with different SLA types (i.e., GB, BG and BE) are providing services from 4 service classes: *Conversational* (Con), *Streaming* (Str), *Interactive* (Int.) and *Background* (Bac.) according to the class-of-service definition in UMTS. VNO GB delivers Voice (Voi), Video calling (Vic), Video streaming (Vis) and Music streaming (Mus). VNO BG serves File sharing (Fil), Web browsing (Web) and Social Networking (Soc) services, while VNO BE provides Internet of Things (IoT) and Email (Ema).

It is further assumed that at each time step k, forty new users arrive and submit their requests for service to their associated VNOs. Simultaneously, twenty users depart from the system. For simplicity, the traffic type percentages of both arrivals and departures remain the same. These values are defined as $U_{[\%]}^{srv}$ in Table 1.

Table 1 – Network Parameters

VNO	Service	Class	R_{svi} in Mbps	$U_{[\%]}^{srv}$	λ_s	C_v^{min} in Mbps
1 (GB)	Voi	Con.	[0.032, 0.064]	10	5	0.4 C^{VRRM}
	Vic		[1, 4]	10	4	
	Vis	Str.	[2, 13]	25	3	
	Mus		[0.064, 0.32]	15	1	
2 (BG)	Fil	Int.	[1, C^{VRRM}]	15	4	0.3 C^{VRRM}
	Web		[0.2, C^{VRRM}]	5	3	
	Soc		[0.4, C^{VRRM}]	10	2	
3 (BE)	Ema	Bac.	[0, C^{VRRM}]	5	4	0
	IoT		[0, 0.1]	5	4	

Figure 5 shows the capacity share of the VNOs as the number of users increases. The dashed and dotted lines represent the minimum guaranteed capacity (G_{vk}) and the minimum contracted SLAs (C_v^{min}) of each VNO respectively. Since the rate of arrivals are higher than the departures, the number of users in the system increases over time. As observed, in the very beginning, when the number of users is quite low, the minimum contracted VNO capacities (C_v^{min}) are sufficient to serve all GB users with their maximum achievable data rates.

Figure 5 demonstrates the ability of the proposed algorithm to correct imbalances in the VNO capacities subject to the minimum capacity constraints. As the number of users in the system increases, VNO GB experiences higher congestion. The algorithm then increases the capacity of VNO GB to correct this imbalance. However, the capacity share of VNO BG will not drop below the minimum contracted SLA threshold. When the minimum requested rate of new arrivals from VNO GB gets higher than the aggregate potentially accessible capacity $B_{\Sigma,k}$, any new service requests will be rejected. At that point, and as expected, there are no BE users left in the VNO.

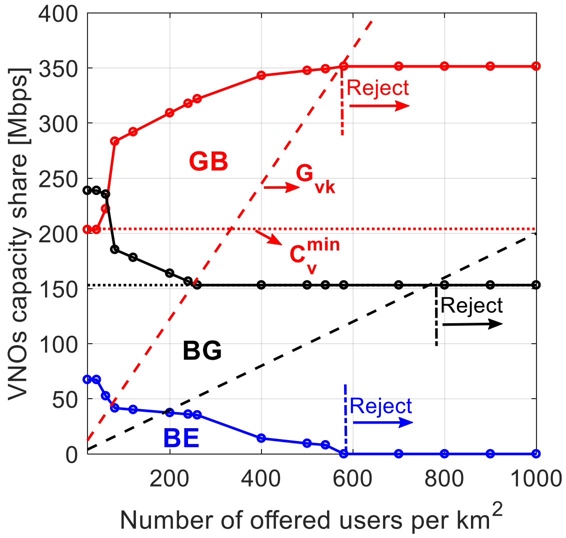

Figure 5 - Total capacity share of the VNOs

Figure 6 shows the potentially accessible capacity of each VNO as the number of users increases over time. For small number of users in the VNO GB and if the capacity is enough to meet the users predefined data rates, the potentially accessible capacity is zero for VNO GB and highest for the VNOs BG and BE. It is noticeable that when the number of users is less than 260, the minimum required capacity by users in VNO GB is less than the minimum contracted capacity for this VNO. For more number of users, the minimum required capacity becomes the dominant factor in equation 15. This explains the increase and decrease in the values of $B_{1,k}$ as the number of users increase.

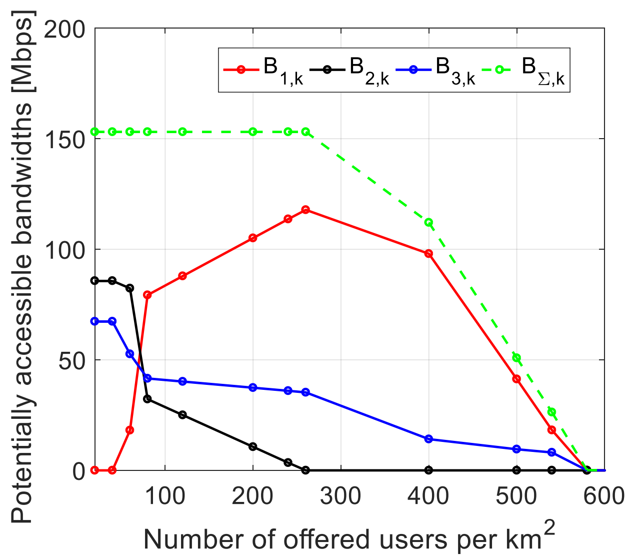

Figure 6 - Accessible bandwidth of the VNOs

The values of $B_{2,k}$ on the other hand decreases to zero, because the minimum contracted capacity for VNO BG according to its SLA remains as the dominant factor compared to the minimum required capacity for all BG users. Since there is no level of guarantees for VNO BE, all capacity share of this VNO is considered as potential available capacity to other VNOs. Therefore, the share of VNO BE capacity (in Figure 5) is the same as $B_{3,k}$. Overall, when the B_{Σ} goes to zero, the remaining capacity will not be enough to accommodate all new arrivals and at least a portion of the new users will be rejected according to the admission control process.

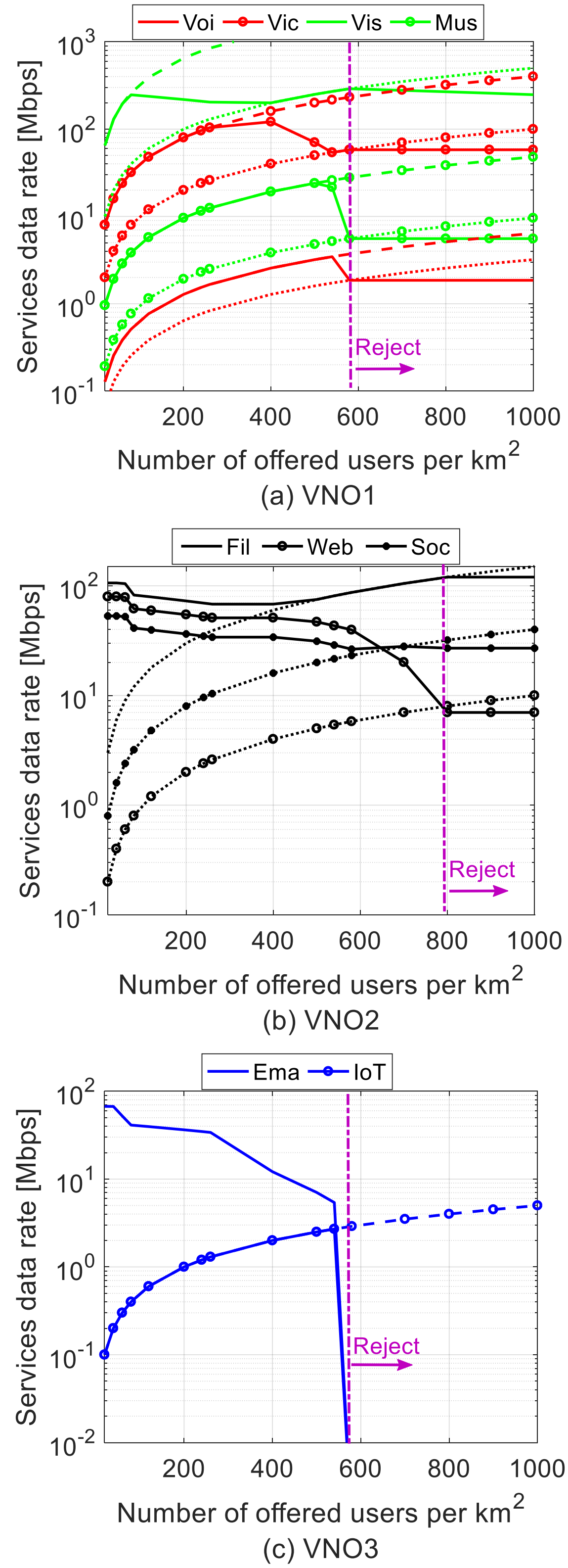

Figure 7. Share of the available capacity by services and VNOs

The capacity shares of each VNO among its connected service slices are presented in Figure 7. The dashed/dotted lines represent maximum/minimum of traffic demands for each service slice. Looking at the VNO GB, when there is enough capacity, the service data rates always vary between the minimum and maximum thresholds predefined in Table 1. However, when the number of users becomes more than 600, all new service requests to VNO GB will be rejected. As a result, there will not be any increase in the capacity share of the service slices in this VNO. This was also shown in Figure 6 when no more potential capacity was available for use by the new arrivals. Also, at that time, all BE users (either existing or new arrivals) have to be delayed or rejected since there is no guarantees for such users. For VNO BG and up to 800 number of users, since the minimum contracted capacity is higher than the minimum required data rates for its users, all new service requests are accepted. Once the total number of users in this VNO reaches 800, all further service requests will be rejected. It is also notable that when there is no capacity constraint, the share of data rates among the services is exactly proportional to their serving weights λ_i as shown in Table 1.

VII. Conclusion and Future Research

A joint admission control and RRM scheme for virtualized RANs has been proposed in this paper. This is based on the service-oriented architecture that are being considered for 5G standards. The Simulation results confirm maximal utilization of the available capacity in the system while all SLAs are satisfied. The fundamental assumption of time scale separation will be addressed in future research. Specific issues that require further studies include viability of this assumption in practical situations, performance loss in situations of comparable time scales in rate/capacity adaptation and users' arrivals/departures process, and mechanisms to mitigate this inefficiency. Another important direction of research is accounting for users with delay-sensitive services. These users are typically characterized by a S-type utility function which turns our formulization into a non-convex optimization problem. In that case, centralized solutions would be computationally intractable, and development of approximate decentralized solutions is a challenging open problem. Finally, due to highly dynamic nature of the emerging services, employing AI techniques as part of network management could be a major focus of future research.

References

[1] Z. Chang, Z. Zhou, S. Zhou, T. Chen and T. Ristaniemi, "Towards Service-Oriented 5G: Virtualizing the Networks for Everything-as-a-Service", *IEEE Access*, Vol. 6, Dec. 2017, pp. 1480–1489.

[2] Z. Feng, L. Ji, Q. Zhang and W. Li, "A Supply-Demand Approach for Traffic-Oriented Wireless Resource Virtualization with Testbed Analysis", *IEEE Transactions on Wireless Communications*, Vol. 16, No. 9, Jun. 2017, pp. 6077–6090.

[3] M. Elkhodr, Q.F. Hassan and S. Shahrestani, *Networks of the Future: Architectures, Technologies, and Implementations*, CRC Press, Boca Raton, FL, USA, 2018.

[4] C. Liang and F. Yu, "Enabling 5G mobile wireless technologies", *EURASIP Journal on Wireless Communications and Networking*, Vol. 2015, No. 218, Sep. 2015.

[5] J. Lucena, P. Ameigeiras and D. Lopez, "Network Slicing for 5G with SDN/NFV: Concepts, Architectures, and Challenges", *IEEE Communications Magazine*, Vol. 55, No. 5, May 2017, pp. 80–87.

[6] A. Aijaz, "Towards 5G-enabled Tactile Internet: Radio Resource Allocation for Haptic Communications", in *Proc. of WCNC'16 - 17th IEEE Wireless Communications and Networking Conference*, Doha, Qatar, Apr. 2016.

[7] S. Singh, S. Yeh, N. Himayat, S. Talwar, "Optimal Traffic Aggregation in Multi-RAT Heterogeneous Wireless Networks", in *Proc. of ICC'16 –52th IEEE International Conference on Communications*, Kuala Lumpur, Malaysia, May 2016.

[8] M. Gerasimenko, D. Moltchanov and R. Florea, "Cooperative Radio Resource Management in Heterogeneous Cloud Radio Access Networks", *IEEE Access*, Vol. 3, Apr. 2015, pp. 397–406.

[9] P. Sroka and A. Kliks, "Playing Radio Resource Management Games in Dense Wireless 5G Networks", *Hindawi Journal of Mobile Information Systems*, Vol. 2016, Nov. 2016, pp. 1 – 18.

[10] F. Teng and D. Guo, "Resource Management in 5G: A Tale of Two Timescales", in *Proc. of ACSSC'15 - 49th Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, USA.

[11] V. Monteiro, D. Sousa and T. Maciel, "Distributed RRM for 5G Multi-RAT Multi-connectivity Networks", *IEEE Systems Journal* (Early Access), Jun. 2018, pp. 1 – 13.

[12] B. Rouzbehani, V. Marbukh, K. Sayrafian, and L.M. Correia, "Towards Cross-Layer Optimization of Virtualized Radio Access Networks," EuCNC 2019.

[13] 3GPP TS 29.500, "5G System; Technical Realization of Service Based Architecture", Jul. 2018.

[14] B. Rouzbehani, L.M. Correia and L. Caeiro, "Radio Resource and Service Orchestration for Virtualised Multi-Tenant Mobile Het-Nets", in *Proc. of WCNC'18 – 19th IEEE Wireless Communications and Networking Conference*, Barcelona, Spain, Apr. 2018.

[15] B. Rouzbehani, L.M. Correia and L. Caeiro, "A Fair Mechanism of Virtual Radio Resource Management in Multi-RAT Wireless Het-Nets", in *Proc. of PIMRC'17 – 28th IEEE Symposium on Personal, Indoor and Mobile Radio Communications*, Montreal, QC, Canada, Oct. 2017.

[16] J. K. MacKie-Mason and H.R. Varian, "Pricing congestible network resources," *IEEE Journal of Selected Areas in Communications*, Vol. 13, No. 7, Sep. 1995, pp. 1141 – 1149.

[17] X. Lin, N.B. Shroff, and R. Srikant, "A Tutorial on Cross-Layer Optimization in Wireless Networks", *IEEE Journal on Selected Areas in Communications*, Vol. 24, No. 8, Aug. 2006, pp 1452 – 1463.

[18] M. Fazel and M. Chiang, "Network utility maximization with nonconcave utilities using sum-of-squares method," in Proc. IEEE Conf. Dec. Contr., Dec. 2005, pp. 1867–1874.