The Effect of Channel Quality on Virtual Radio Resource Management

Sina Khatibi, Luis M. Correia
INOV – INESC Inovação
Rua Alves Redol, N. 9, 7Dto.
1000-029 Lisboa,
Portugal
Phone: +351 21 310 04 32
Fax: + 351 21310 04 72
Email: sina.khatibi@inov.pt
      luis.correia@inov.pt
The Effect of Channel Quality on Virtual Radio Resource Management

Sina Khatibi, Luis M. Correia
INOV-INESC / Instituto Superior Técnico, University of Lisbon Portugal
Email: {sina.khatibi, luis.correia}@inov.pt

Abstract - This paper proposes a model for the management of virtual radio resources, considering different assumptions on terminals’ Signal-to-Interference-plus-Noise Ratio (SINR). The model has two main parts: estimation and allocation of the available resources. In first part, a technique for obtaining probability functions of the network throughput based on the available radio resources is introduced. In addition, three approaches for estimating the network capacity, other than a general one, are proposed, i.e., optimistic, realistic, and pessimistic. In the second part, a portion of the estimated network capacity is allocated to each of the services of Virtual Network Operators (VNOs). Meeting the Service Level Agreements (SLAs) in addition to increasing the efficiency of resource usage are the key objectives in resource allocation. The numeric results for practical scenarios show that the cellular network capacity, itself, can vary from 0.9 Gbps in the pessimistic approach up to 5.5 Gbps in the optimistic one. Moreover, the effect of this capacity variation on the allocation of the virtual radio resources to the VNOs with different SLAs is also studied. Finally, the isolation of the service classes and the VNOs by means of virtualisation of radio resources are clearly demonstrated through numeric results.

I. INTRODUCTION

While operators have scarce radio resources available, the proliferation of smart devices and traffic-hungry applications exponentially increased mobile data demand. It is expected that the monthly global data traffic will surpass 10 ExaByte in 2017 [1]. In addition, due to drastic variations of traffic, both geographically and temporally [2], the common provisioning approach used in Radio Access Networks (RANs), i.e., considering only busy hours, leads to an inefficient resource usage with relatively high Capital and Operational Expenditure (CAPEX and OPEX) costs. Hence, operators using scarce radio resources have to find a practical, flexible, and cost-efficient solution for their networks where they can also share their physical infrastructure.

In addition, traffic offloading, e.g., to Wi-Fi Access Points (APs), has recently proved to be a valuable complementary solution. Late studies [3, 4] suggest that an acceptable portion of traffic can be offloaded to APs just by allowing users to delay their delay-tolerant data for a maximum pre-specified interval, until reaching an AP. Since almost all Mobile Terminals (MTs) today have other connectivity capabilities, offloading approaches are based on using them instead of expensive cellular bands, whenever it is possible. Furthermore, [5] analysed this approach from the energy-saving perspective, the economics of traffic offloading being discussed in [6].

Recently, the sharing of network infrastructure using Network Function Virtualisation (NFV), which is meant to transform the way operators architect their networks[7], has become an active research topic. Based on this, the concept of virtualisation of the radio resources is introduced in [8]. This approach suggests to serve multiple Virtual Network Operators (VNOs) over the same infrastructure offering them isolation and flexibility in addition to network element abstraction and multi-RAT (Radio Access Technology) support, by aggregating and managing all radio resources. In this approach, VNOs do not deal with the physical resources, but they just request wireless connectivity from a set of infrastructure providers. Virtualisation of radio resources increases the flexibility of handling network traffic changes in RAN in addition to offering pay-as-you-go Connectivity-as-a-Service (CaaS) to VNOs, while enabling new business models for network operators and infrastructure providers.

However, Virtual Radio Resource Management (VRRM) is a non-trivial task, since it has to serve multiple VNOs with different requirements and Service Level Agreements (SLAs) over the same infrastructure. Furthermore, wireless links are always subject to fading and interference, hence, their performance is variable. In [9], a practical model for the management of virtual radio resources is proposed. The model, in a first step, uses a probabilistic approach to estimate the available network capacity. In the next step, the services of the VNOs are allocated by a portion of the available capacity. The novelty of this paper is the extension of the capacity estimation technique to consider three more approaches, i.e., optimistic, realistic, and pessimistic ones. In each of the approaches, the network capacity is estimated considering a pre-defined assumption. In the next step, the virtual radio resource allocation is studied, using practical scenarios when the aforementioned approach is used to estimate network capacity. Through the numeric results, the effect of channel quality on various services of the VNOs is demonstrated.

This paper is organised as follows. Section II describes the model for virtual radio resource management and the proposed extension. The scenario for evaluation of the proposed model is shown in Section III. In Section IV, the numeric results are presented and discussed. The paper is concluded in Section V.
II. RADIO RESOURCE MANAGEMENT IN V-RAN

A. Estimation of Available Resources

The management hierarchy of virtual radio resources, consists of VRRM on the top of the usual radio resource management entities of heterogeneous access networks [10], Common Radio Resource Management (CRRM) and local RMs. The role of VRRM is to translate VNOs’ requirements and SLAs into a set of polices for lower levels [8]. The estimation of the available resources, and their allocation, are the two key parts of the VRRM procedure.

The goal in the estimation of the available virtual radio resources is to obtain a relationship in the form of a Probability Density Function (PDF) between the set of available Radio Resource Units (RRUs) (e.g., time-slots in GSM, codes in UMTS, resource-blocks in LTE, and radio-channels in Wi-Fi), and the total network capacity. Based on the outcome of the estimation, the VRRM allocates a portion of this capacity to each service of each VNO.

In general, the data rate achieved by allocating an RRU to an MT depends on various parameters, such as RAT, modulation, and coding scheme, among others. In a certain configuration (e.g., certain modulation) the data rate of an RRU is a function of Signal-to-Interference-plus-Noise Ratio (SINR). Based on [8], the PDF of \( R_{b,RAT_i} \) can be given as:

\[
p_{R_b}(R_{b,RAT_i}[Mbps]) = \frac{0.5 \cdot \ln(10) \cdot \sum_{k=1}^{5} k a_k (R_{b,RAT_i})^{k-1} \cdot \frac{0.5 \cdot \ln(10) \cdot \sum_{k=1}^{5} a_k (R_{b,RAT_i})^{k}}{e^{\frac{0.5 \cdot \ln(10) \cdot \sum_{k=1}^{5} a_k (R_{b,RAT_i})^{k}}{\alpha_p}}}{e^{-\frac{0.5 \cdot \ln(10) \cdot \sum_{k=1}^{5} a_k (R_{b,RAT_i})^{k}}{\alpha_p}}} \cdot (R_{b,RAT_i}[Mbps])
\]

where:

- \( R_{b,RAT_i} \): data rate of an RRU from the \( i \)-th RAT,
- \( R_{b,RAT_i}^{\max} \): the maximum data rate of an RRU from the \( i \)-th RAT,
- \( \alpha_p \geq 2 \): the path loss exponent,
- \( a_k \): coefficients in a polynomial approximation of SINR, as a function of data rate in each RAT; for cellular networks, they are presented in [8] and [9].

Obviously, offering the same data rate to a terminal with low input SINR requires assigning comparatively more RRRUs than the one with the higher SINR. Hence, it can be claimed that a relatively higher network capacity can be achieved when all MTs have a higher SINR. Three approaches for the estimation of network capacity, in addition to the general (G) one, can be considered:

- Optimistic (OP): All RRUs are assigned to MTs with high SINR, and the data rate of each RRU satisfies:

\[
0.5 R_{b,RAT_i}^{\max} \leq R_{b}[Mbps] \leq R_{b}^{\max}[Mbps]
\]

- Realistic (RL): Half of the RRUs are assigned to MTs with high SINR, hence, data rates’ boundaries for these RRUs are as described in (2). However, the MTs to which the remaining half of RRUs are assigned have a low SINR, therefore, the data rates for these RRUs are given by:

B. Allocation of the Resources

In the next step, the services of the VNOs have to be granted with a portion of the network capacity. The allocation of the resources has to be based on services’ priority and SLAs. These SLAs can generally be categorised into three main groups, which are:

- Guaranteed Bitrate (GB), in which the VNO is guaranteed a minimum and a maximum level of data rates, regardless of the network status.
- Best effort with minimum Guaranteed (BG), where the VNO is guaranteed with a minimum level of service. The request for higher data rates than the guaranteed level is served in the best effort manner.

Based on [11], (1) can be modified for these special case studies, where the data rate is bounded between high and low values, as follows:

\[
P_{R_b} \left( R_{b,RAT_i}[Mbps] \right) = \frac{e^{\frac{0.5 \cdot \ln(10) \cdot \sum_{k=1}^{5} a_k (R_{b,RAT_i})^{k}}{\alpha_p}}}{e^{\frac{0.5 \cdot \ln(10) \cdot \sum_{k=1}^{5} a_k (R_{b,RAT_i})^{k}}{\alpha_p}}} \cdot (R_{b,RAT_i}[Mbps])
\]

where:

- \( R_{b,low} \): low boundary for RRU data rate;
- \( R_{b,high} \): high boundary for RRU data rate, where:

\[
0 \leq R_{b,low}[Mbps] \leq R_{b,high}[Mbps] \leq R_{b}^{\max}[Mbps]
\]

The total data rate from a single RAT pool is:

\[
R_{b,tot}^{RAT_i} = \sum_{n=1}^{N_{RAT}} R_{b,n}^{RAT_i}
\]

where:

- \( N_{RAT} \): number of RRUs of the \( i \)-th RAT,
- \( R_{b,n}^{RAT_i} \): data rate from a \( i \)-th RAT pool,
- \( R_{b}^{RAT_i} \): data rate from the \( n \)-th RRU of the \( i \)-th RAT.

Assuming that channels are independent and based on [11], the PDF of a RAT’s data rate is equal to the convolution of all RRUs’ PDFs. The resource pools of different RATs can also be aggregated under the supervision of CRRM; in this situation, the total data rate from all RATs is the summation of the total data rate from each of the individual ones:

\[
R_{b,CRM}^{RAT_i} = \sum_{i=1}^{N_{RAT}} R_{b,tot}^{RAT_i}
\]

By having the number of the available resources mapped onto probability functions, the VRRM has an estimation of the total network capacity.
The key goal in the allocation procedure is to increase the total network throughput, while considering the priority of different services and the other constraints. Given this, the objective function for VRRM, $f^E_{b_j}$, is the total weighted network data rate (e.g., in Gbps), being expressed as:

$$f^E_{b_j}(R^W_{b_j}) = f^\text{cell}_{b_j}(R^\text{cell}_{b_j}) + f^\text{WLAN}_{b_j}(R^\text{WLAN}_{b_j}) - f^f_{b_j}(R^f_{b_j})$$  \hspace{1cm} (8)

where:
- $f^\text{cell}_{b_j}$: objective function for cellular RATs,
- $R^\text{cell}_{b_j}$: vector of serving data rates from cellular networks,
- $f^\text{WLAN}_{b_j}$: objective function for APs,
- $R^\text{WLAN}_{b_j}$: vector of serving data rates from WLAN,
- $f^f_{b_j}$: fairness function,
- $R^f_{b_j}$: vector of intermediate fairness variables,
- $R^W_{b_j}$: vector of serving data rates, where:

$$R^W_{b_j}[Mbps] = R^\text{cell}_{b_j}[Mbps] + R^\text{WLAN}_{b_j}[Mbps]$$  \hspace{1cm} (9)

The objective function for cellular RATs in (8) is given by:

$$f^\text{cell}_{b_j}(R^\text{cell}_{b_j}) = \sum_{i=1}^{N_{VNO}} \sum_{j=1}^{N_{srv}} W^\text{SRB}_{ji} R^\text{cell}_{b_j}$$  \hspace{1cm} (10)

where:
- $N_{VNO}$: number of served VNOs by this VRRM,
- $N_{srv}$: number of services for each VNO,
- $W^\text{SRB}_{ji}$: weight of serving unit of data rate for service $j$ of VNO $i$ by VRRM, where $W^\text{SRB}_{ji} \in [0,1]$,

In order to decrease the collision rates, it is desirable that the services with the higher serving weights receive higher data rates than the ones with the lower serving weights. Hence, the equivalent function for WLANs is:

$$f^\text{WLAN}_{b_j}(R^\text{WLAN}_{b_j}) = \sum_{i=1}^{N_{VNO}} \sum_{j=1}^{N_{srv}} \left( W^\text{SRB}_{ji} R^\text{WLAN} + W^\text{SRB}_{ji} R^\text{cell}_{b_j} \right)$$  \hspace{1cm} (11)

where:
- $W^\text{WLAN}_{ji}$: weight for session average data rate, where $W^\text{WLAN}_{ji} \in [0,1]$,
- $R^\text{max}_{b_j}$: maximum average data rate among all network services,
- $R^f_{b_j}$: average data rate for service $j$.

The weights in (10) and (11) are used to prioritise the allocation of data rates to services, being based on the SLAs between VNOs and VRRMs; it is a common practice to have the summation of all of them equal to unit.

Fairness is the other objective in the allocation procedure, given by:

$$f^f_{b_j}(R^f_{b_j}) = \sum_{i=1}^{N_{VNO}} \sum_{j=1}^{N_{srv}} \left( R^\text{max}_{b} R^f_{b_j} \right)$$  \hspace{1cm} (12)

where:
- $R^\text{max}_{b}$: minimum average data rate among all network services,
- $R^f_{b_j}$: the boundary for deviation data rate from the normalised average for service $j$ of VNO $i$, defined as follows:

$$\frac{R^\text{SRV}_{b_j}[Mbps]}{W^\text{SRB}_{ji}} - \sum_{i=1}^{N_{VNO}} \sum_{j=1}^{N_{srv}} \frac{R^\text{SRV}_{b_j}[Mbps]}{N_{VNO} N_{srv} W^\text{SRB}_{ji}} \leq R^f_{b_j}[Mbps]$$  \hspace{1cm} (13)

In addition, there are more constraints for VRRM to allocate data rates to various services, which should not be violated. The very fundamental constraint is the total network capacity estimated in last section. The summation of the entire assigned data rates to all services should always be smaller than the total estimated capacity of the network:

$$\sum_{i=1}^{N_{VNO}} R^\text{SRV}_{b_j}[Mbps] \leq R^\text{CRBM}_{b_j}[Mbps]$$  \hspace{1cm} (14)

The offered data rate to GB and BG services imposes the next constraints. The data rate allocated to these services has to be higher than a minimum guaranteed level (for both GB and BG) and lower than a given maximum (for GB only):

$$R^\text{Min}_{b_j}[Mbps] \leq R^\text{SRV}_{b_j}[Mbps] \leq R^\text{Max}_{b_j}[Mbps]$$  \hspace{1cm} (15)

where:
- $R^\text{Min}_{b_j}$: minimum guaranteed data rate for service $j$ of VNO $i$,
- $R^\text{Max}_{b_j}$: maximum data rate for service $j$ of VNO $i$.

In this paper, the optimisation is done based on the interior-point method using the linprog function of MATLAB [12].

### III. SCENARIOS

The numerical study is conducted by using a realistic scenario, in which coverage is provided through a set of Radio Remote Heads (RRHs) [14]. The RRHs are capable of supporting multiple RATs, i.e., OFDMA (based on LTE-Advance), CDMA (based on UMTS), and FDMA/TDMA (based on GSM). The considered layout has TDMA cells with the radius of 1.6 km, CDMA cells with 1.2 km, and OFDMA cells with 0.4 km. It is assumed that the coverage area is divided into serving areas, over which a VRRM is operating. In these scenarios, the serving area for each VRRM is considered to be as big as a TDMA cell, hence, each serving area is covered by 1 TDMA cell, ~1.7 CDMA cells, and 16 OFDMA ones.

In addition to cellular networks, it is assumed that an IEEE802.11ac AP is placed in the centre of each OFDMA cell. These OFDM AP are configured to work over a 80 MHz channel bandwidth (from European Union regulations, there are only five available channels for 80 MHz APs [13]). In contrast to cellular RATs, Wi-Fi uses the same set of links for up- and downlinks. In order to achieve coherence among all RATs, the total Wi-Fi throughput is equally divided between
up- and downlinks. Hence, the number of OFDM RRUs in each link is half of the total number of available channels, TABLE I.

<table>
<thead>
<tr>
<th>RAT</th>
<th>Number of Cells</th>
<th>Cell Radius [km]</th>
<th>System</th>
<th>(n_{\text{RRU}}^{\text{RAT}})</th>
</tr>
</thead>
<tbody>
<tr>
<td>OFDM</td>
<td>16</td>
<td>0.08</td>
<td>Wi-Fi</td>
<td>40</td>
</tr>
<tr>
<td>OFDMA</td>
<td>16</td>
<td>0.4</td>
<td>LTE</td>
<td>8 000</td>
</tr>
<tr>
<td>CDMA</td>
<td>1.7</td>
<td>1.2</td>
<td>UMTS</td>
<td>80</td>
</tr>
<tr>
<td>TDMA</td>
<td>1</td>
<td>1.6</td>
<td>GSM</td>
<td>75</td>
</tr>
</tbody>
</table>

Moreover, three VNOs are assumed to operate in this area, with the same number of subscribers. Subscribers are using smart phones and the average data rate of 6.375 Mbps is required for each one [14]. VNOs also offer the same set of services as listed in TABLE II.

The serving weights are based on general service classes: 0.4 for Conversational, 0.3 for Streaming, 0.2 for Interactive, and \(W^{\text{SRB}}\) is heuristically chosen to 0.05. Each VNO has a different SLA, i.e., GB, BG, and BE. On the ground of each service data rate, the SLAs of these VNOs are defined as follows:

- VNO GB: the data rates allocated to services are guaranteed to be in the range of 50% to 100% of the corresponding service data rate.
- VNO BG: it has best effort with a minimum 25% of the service data rate guaranteed by the SLA.
- VNO BE: it has best effort served in the best effort approach, without any guarantee.

<table>
<thead>
<tr>
<th>Service</th>
<th>Volume [%]</th>
<th>(W^{\text{SRB}})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Video Call</td>
<td>59.7</td>
<td>0.3</td>
</tr>
<tr>
<td>Video Streaming</td>
<td>4.6</td>
<td>0.4</td>
</tr>
<tr>
<td>File Sharing</td>
<td>3.5</td>
<td>0.2</td>
</tr>
<tr>
<td>Web Browsing</td>
<td>11.9</td>
<td>0.2</td>
</tr>
<tr>
<td>Social Networking</td>
<td>14.4</td>
<td>0.2</td>
</tr>
<tr>
<td>M2M</td>
<td>5.9</td>
<td>0.2</td>
</tr>
<tr>
<td>e-Health</td>
<td>25</td>
<td>0.05</td>
</tr>
<tr>
<td>Int. Transp. Serv.</td>
<td>25</td>
<td>0.2</td>
</tr>
<tr>
<td>Surveillance</td>
<td>25</td>
<td>0.05</td>
</tr>
<tr>
<td>Email</td>
<td>1</td>
<td>0.05</td>
</tr>
<tr>
<td>Music</td>
<td>3</td>
<td>0.3</td>
</tr>
<tr>
<td>VoIP</td>
<td>1</td>
<td>0.4</td>
</tr>
</tbody>
</table>

IV. RESULTS

The PDFs of the cellular networks for any of these approaches using (1) and (4) were calculated, and the relative CDFs of the cellular networks capacity are plotted in Fig. 1. The PE approach leads to the lowest and OP the highest network capacity. Between PE and OP approaches, the G and RL approaches are placed. It can be observed that as the path-loss exponent increases, the network capacity also increases due to the higher attenuation of interferences. In the next step, the number of VNOs’ subscribers swept between 300 and 1400.

Fig. 2 presents the allocated data rate to VNO GB in conjunction with minimum and maximum guaranteed data rate. As long as one has the data rates within the acceptable region (i.e., shown by the solid colour), there is no violation to the SLA and guaranteed data rate. It can be seen that in OP approach up to 600 subscribers, the maximum guaranteed data rate is offered to this VNO. Considering the other approaches, as the number of the subscribers increases the allocated data rate moves toward the minimum level of the guaranteed data rate. Finally, in the PE approach, as the number of subscribers passes 1100, violations to minimum guaranteed data rate can be observed: this means that the network capacity is too low so that the model faced resource shortage and it has to violate the minimum guaranteed data rate of this VNO.

FIG. 1. CDF of the cellular networks data rate for different approaches.

FIG. 2. The allocated data rate to VNO GB in different approaches.

The allocate data rates to VNO BG and BE are also plotted in Fig. 3. Comparing the allocated data rate to all three VNOs in OP and RL approaches, it can be seen that not only the VNO GB is served, acceptable data rates are also allocated to VNO BG and BE. However, in G and PE approaches, the VNO BG and BE are assigned with comparatively lower data rates than VNO GB. They also affect more in resources shortage until the point that VNO BE is stopped being served as the 1100 subscribers under PE approach.

FIG. 3. The allocated data rate to VNO BG and VNO BE in different approaches.
Furthermore, Fig. 4 illustrates the variation of the capacity assigned to conversational and background classes of VNO GB in different approaches. It can be seen that the assigned data rates to the conversational class (i.e., the class with the highest service weight) for the OP and RL approaches is the maximum guaranteed data rate. As the capacity reduces in G and PE approaches, the model manages to keep the allocated data rate in the acceptable range without any violations. In contrast, background class services face violation of minimum guaranteed data rate in the PE approach. As the number of subscribers passes 1100 subscribers, the violations become so severe that the services do not receive any data rate.

Fig. 4. The allocated data rate to service classes of VNO GB.

V. CONCLUSIONS:

The two key parts of the proposed model for management of virtual radio resources are the estimation of available resources and their allocating to different services of VNOs. In this paper, an extension for the technique used in the first part is suggested that enables the model to consider different assumptions on the terminals input SINR and their effects on the total network throughput. In addition, three cases representing three approaches to estimating the network capacity are introduced: optimistic, realistic, and pessimistic. The goal in the allocation procedure is to allocate the available resources among different services of the VNOs in order to increase the resource usage efficiency, satisfy the guaranteed data rates, and meet the fairness criteria.

The results show that the network capacity increases from 0.9 Gbps in PE to 5.5 Gbps OP approach. The effect of these capacity changes on the allocated data rates to different VNOs and their services classes are presented through series of plots. It is shown that when there is enough capacity, not only the guaranteed VNO is satisfied but also the best effort VNOs are also well served. However, as the network capacity decreases due to channel quality (G and PE approaches), the best effort VNOs are affected more than the guaranteed one. The same situation is shown in the service class level too. The conversational and the streaming classes are the classes with the highest serving priority (i.e., serving weights). These service classes are allocated with the higher data rates generally than the other two service classes. When there are resources shortage, i.e., in G and PE approaches, the violations start relatively by background and interactive classes.

Finally, this paper not only discusses the effect of channel quality on VRRM procedure but it also shows the isolation offered by virtualisation of radio resources. It shows through practical scenarios, that the model is able to isolate the offer service quality to the guaranteed VNO from the network situation. As evidence to this claim, the VNO GB and particularly its conversational class can be considered where the requested service quality regardless of network situation is offered. In addition, the flexibility of the applied model made the equilibrium among different services or different VNOs possible.

ACKNOWLEDGMENT

The research leading to these results was partially funded by the European Union’s Seventh Framework Programme Mobile Cloud Networking project (FP7-ICT-318109).

REFERENCES