Article:

Vasilakis, Vasileios orcid.org/0000-0003-4902-8226, Moscholios, Ioannis and Logothetis, Michael (2018) Efficient radio resource allocation in SDN/NFV based mobile cellular networks under the complete sharing policy. IET Networks. pp. 103-108. ISSN: 2047-4962

https://doi.org/10.1049/iet-net.2017.0053

Efficient Radio Resource Allocation in SDN/NFV Based Mobile Cellular Networks under the Complete Sharing Policy

Vassilios G. Vassilakis, Ioannis D. Moscholios, and Michael D. Logothetis

Abstract

Novel networking paradigms, such as software-defined networking (SDN) and network function virtualization (NFV), introduce new opportunities in the design of next-generation mobile networks. The present work investigates the benefits of the emerging SDN and NFV technologies on the radio resource management (RRM) in mobile cellular networks. In particular, the aim of the RRM scheme is to enable an efficient and flexible radio resource allocation in order to assure quality-of-experience of mobile users. The authors consider the orthogonal frequency division multiple access scheme and the complete radio resource sharing policy. To enable time- and space-efficient resource allocation, the authors investigate the applicability of the well-known Kaufman-Roberts recursion in the context of new architectural and functional changes of SDN/NFV based mobile environments. Finally, they discuss the applicability of the proposed approach for more complicated resource sharing policies.

I. Introduction

Fast proliferation of smartphones has introduced the opportunity for novel services and applications in the mobile sector. At the same time, the mobile industry is moving toward more advanced networking and communication technologies. One of the factors that necessitated this move is the inability of the current long-term evolution (LTE) [1] technology to ensure sufficiently low end-to-end latency and to support large numbers of connected devices and high traffic volumes, as required in this new evolved mobile ecosystem [2]. Despite the large-scale deployments of additional base stations (BSs) and access points (APs) in the recent years, the increase in coverage and capacity of today's networks is not sufficient to assure the appropriate quality-of-experience (QoE) of mobile users (MUs) [3]. Hence, there are many attempts to push the boundaries by enabling more efficient and flexible mobility and radio resource management (RRM) and for providing native support to multiple co-existing radio access technologies (RATs) [4]. To this end, the software-defined networking (SDN) and network function virtualization (NFV) are considered as important enabling technologies [5].

SDN and NFV have recently attracted lots of research efforts and have gained a tremendous attention from both academic and industry communities. The SDN technology is the driver toward completely programmable networks, which can be achieved by decoupling the control and data planes [6], [7]. On the other hand, the NFV technology allows executing the software-based network functions on general-purpose hardware via virtualization [8], [9]. SDN and NFV, due to their complementary nature, are traditionally seen as related concepts and implemented together [10]. Some of the expected benefits of SDN/NFV include: CAPEX and OPEX reduction for network operators, by reducing the cost of hardware and automating services; flexibility in terms of deployment and operation of new infrastructure and applications; faster innovation cycles due to the creation of enhanced services/applications and new business models. Due to the aforementioned benefits, SDN/NFV will play a major role in the emerging 5th generation (5G) systems [11]. There have already been successful attempts in applying these concepts in cellular networks, both in the radio access network (RAN) [12], [13] and in the mobile core network (MCN) [14].

V. G. Vassilakis is with the Dept. of Computer Science, University of York, YO10 5GH York, UK.
I. D. Moscholios is with the Dept. of Informatics & Telecommunications, University of Peloponnese, 221 31 Tripolis, Greece.
M. D. Logothetis is with the Dept. of Electrical & Computer Engineering, University of Patras, 265 04 Patras, Greece.

This work considers an evolved ecosystem for 5G mobile cellular systems and studies the advantages of SDN and NFV on the RRM. Our considered network architecture includes: user/control plane split, as enabled by SDN; virtualization of the BS functions, as enabled by NFV; and co-existence of multiple channel access schemes. In particular, the objective of the RRM scheme is to assure the QoE of MUs in an efficient and flexible manner. To this end, we study the applicability of the well-known Kaufman-Roberts (K-R) recursion [15], [16] for the radio resource allocation (RRA) as well as the determination various system measures, such as the call blocking probabilities (CBP). The CBP, is a metric that has been traditionally used by the telecom operators to define the quality-of-service (QoS), which refers to the capability of the network to provide better service to selected network traffic. On the other hand, QoE typically refers to the acceptability of an application or service, as perceived subjectively by the end-user. QoE may be negatively influenced by high CBP and low QoS. In this work, we do not consider other factors, beyond CBP, such as subjective user perception, that may influence the QoE.

The K-R recursion was initially proposed for the traditional connection-oriented networks and since then has been extended for wired [17], [18], [19], [20], wireless [21], [22], [23], and optical networks [24], [25], [26]. In particular, the arrivals and departures of calls in the system are modelled as a continuous-time Markov chain (CTMC). Then, due to the existence of local balance (LB) between the adjacent system states, a recursive algorithm for state probabilities is derived. The algorithm is accurate and the analytical results coincide with the simulation results.

With the development of next-generation mobile technologies, as influenced by the advances in SDN/NFV, there is a need for fast, efficient, and accurate tools for RRA. To the best of our knowledge, the applicability of the K-R recursion in SDN/NFV mobile networks has not been investigated. The current work tries to fill in this gap, while considering realistic scenarios and focusing on the complete resource sharing policy. We also investigate systems of different capacities and service-classes. In all cases, as confirmed by our experimental results on general-purpose hardware, the algorithm execution time is less than a second even for systems of large capacities. This property, makes the algorithm applicable for fast RRA in a real-time manner.

This paper is organized as follows. In Section II, we present our considered architectural framework for the SDN/NFV enabled mobile systems. Section III provides the considered model for the RAN and derives the recursive algorithm for state probabilities, which can then lead to the CBP calculation. Section IV presents the experimental CBP results. In Section V, we conclude and discuss possible future directions. For the reader's convenience, in the Appendix we include the list of abbreviations used in this paper.

II. SDN/NFV Based Cellular Network Architecture

In this section we briefly describe our considered SDN/NFV based cellular network architecture, shown in Fig. 1, and its main elements.

The realization of an intelligent RAN is greatly facilitated by the SDN and NFV technologies [27]. SDN enables abstraction and modularity of the network functions at the RAN level. As a consequence, a hierarchical control architecture can be implemented, in which the high control layer controls lower layers by specifying procedures and without the requirement to have access to the specific implementation details of the lower layers [28]. Such an implementation, however, requires a holistic view of the cellular network at the higher control layer to be designed by taking into account appropriate abstraction of lower layers via well-defined control interfaces. This is essential to enable programmable RRM functions, such as RRA and call admission control (CAC). On the other hand, the NFV technology allows the execution of control programs on general purpose computing/storage resources [29]. This is contrary to the traditional approach in which the BS consists of a tightly coupled software and hardware platform. Hence, an NFV-based BS may have some network functions implemented as physical network functions (PNFs), while other functions are implemented as virtual network functions (VNFs). An advantage of VNFs is that the underlying hardware can be efficiently utilized, since VNFs run on shared NFV infrastructure (NFVI).

The architecture of Fig. 1 relies on the SDN concept whose different layers are depicted in Fig. 2. In the Control Layer, the SDN controller provides a global view of the available underlying resources to

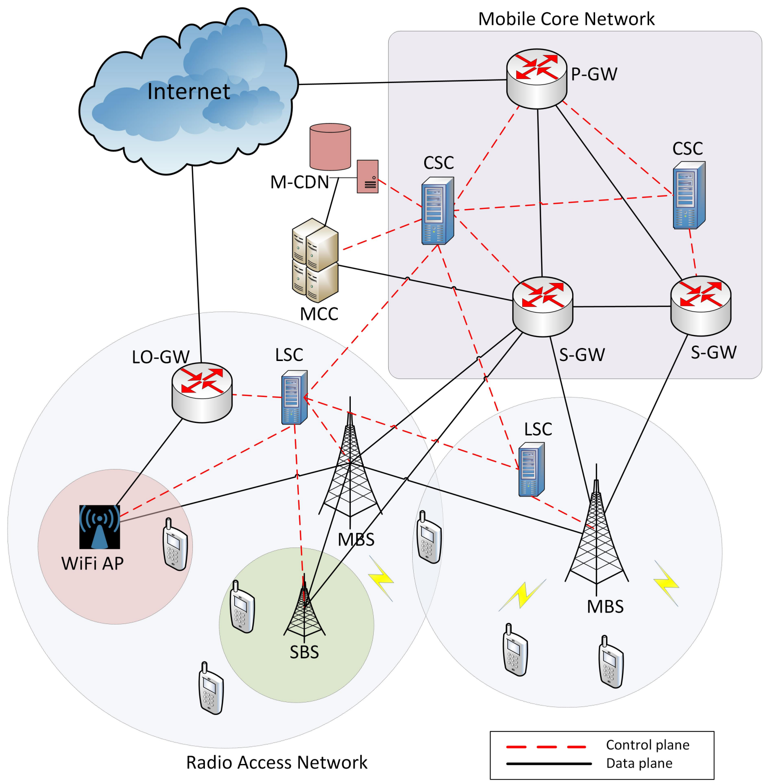

Fig. 1: SDN/NFV Based Cellular Network Architecture.

one or more network applications that are located at the Application Layer. This communication is done using the northbound open application programming interface (API). On the other hand, the southbound open API is used to configure the forwarding elements (FEs) that are located at the Infrastructure Layer. The configuration of FEs is performed by the SDN controller who sends control messages to the SDN agents located within the FEs.

The main elements at the RAN level are: small cell BSs (SBSs), macro BSs (MBSs), WiFi APs, local offload gateways (LO-GWs) [30], and MUs. These entities are controlled by the local SDN controller (LSC). The geographical area of the RAN consists of a number of clusters. Each cluster typically consists of many cells and is under the control of a single LSC. For example, in Fig. 1 the first cluster contains one MBS, one SBS, one WiFi AP, one LO-GW and four MUs, whereas the second cluster contains one MBS and three MUs. MUs can freely move between clusters or even may belong to more that one cluster at the same time.

When a network entity wishes to establish a connection, it sends the request to the corresponding LSC of the cluster. Upon receiving the request, the LSC will identify the appropriate destination address for the requested connection. In particular, the LSC will either forward the request to the appropriate in-cluster recipient (e.g, MU or MBS) or to the MCN, if the recipient is outside the cluster. To be able to perform this, the LSC maintains the knowledge of the cluster topology as well as the external connections toward the MCN and neighboring clusters. The LSC is also responsible for the multi-RAT coordination. That is, it takes the RRA decisions in geographical areas where multiple RATs are available (e.g., LTE and WiFi). In the cache-enabled mode [31], the LSC takes caching decisions within the cluster by exploiting the knowledge of content popularity and the available in-cluster resources. Another important function of the LSC is the in-cluster content routing. Upon receiving the connection request from an MU, the LSC constructs the path from the content source (if the source is within the cluster) or the border entity (if the source is outside the cluster) toward the requesting MU. The LSC then modifies the flow tables at the FEs along the content delivery path. Finally, the LCS is responsible for the MU mobility within its

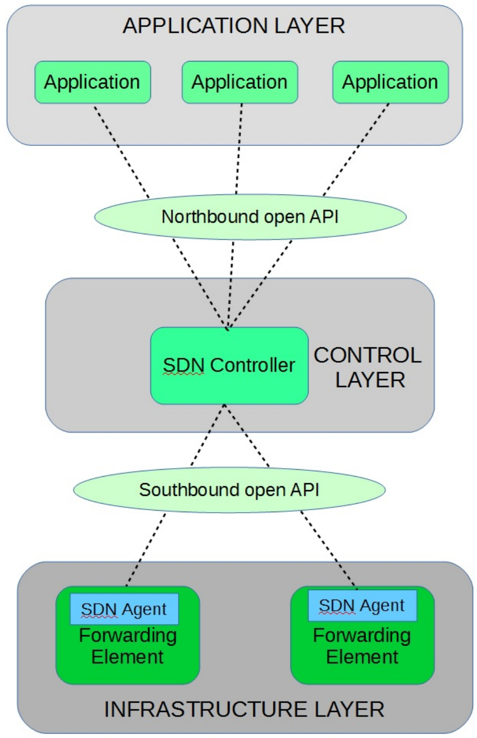

Fig. 2: Layering Concept in Software-Defined Networking.

cluster. Hence, mobility-related information does not need to be sent over to the MCN.

As mentioned in Section I, the integration of SDN with NFV is of main interest. In Fig. 3, the basic components of a virtualized RAN are shown. Multiple virtual base stations (VBSs) may run on top of the NFVI, essentially sharing the resources of the same physical infrastructure

The MCN consists of the mobile cloud computing (MCC) infrastructure, mobile content delivery network (M-CDN) servers, packet data network (PDN) gateways (PDN-GWs), and serving gateways (S-GWs). The control is performed by one or more core SDN controllers (CSCs). An CSC receives and handles the connection requests from the RAN via the corresponding LSCs. An CSC is also responsible for storage (e.g., M-CDN), compute (e.g., MCC), spectrum, and energy resources, and for providing QoS support. Finally, the PDN-GW forwards traffic to/from the Internet and other external IP networks, whereas the S-GW receives/sends traffic from/to the RAN.

III. Radio Access Network Model

Below we describe our considered RAN model in the SDN/NFV enabled cellular network. We assume a cluster of VBSs controlled by a single LSC at the RAN level (Fig. 3). The cluster has a fixed number of V of VBSs. For the purposes of the analysis it is assumed that the amount of radio resources in the RAN can be discretized and is measured in resource units (RUs). RU's definition depends on the adopted channel access scheme. When schemes based on the frequency division multiple access (FDMA) and the time division multiple access (TDMA) are used, the RU can be defined as an integer number of frequency carriers or time slots. On the other hand, when schemes based on the code division multiple access (CDMA) are used, the definition of the RU must take into account the multiple access interference [32]. To this end, the notions of the cell load and load factor has been used [33], [34]. In this work we consider the LTE orthogonal FDMA (OFDMA) scheme. We define a RU to be equal to a single OFDMA resource block (RB). For example, if the LTE channel bandwidth is 9 MHz and one subcarrier is 15 KHz, then there are in total 600 subcarriers. Since one OFDMA RB corresponds to 12 subcarriers, we have

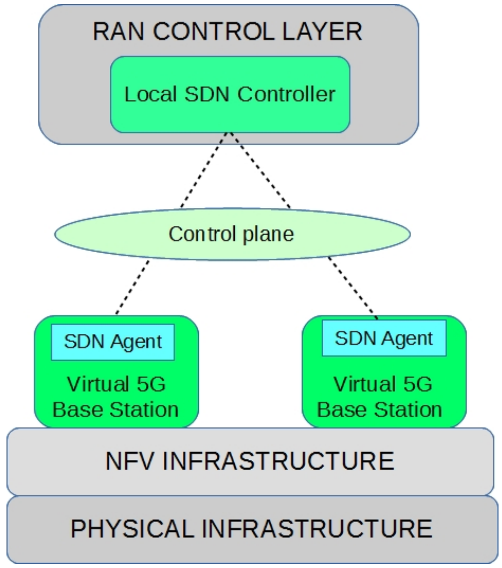

Fig. 3: SDN/NFV Based Radio Access Network.

50 RUs per channel per time slot. Similarly, a 13.5 MHz LTE channel has 75 RUs and a 18 MHz LTE channel has 100 RUs.

Let us denote by C the total number of RUs in the RAN. RUs are dynamically allocated by the LSC to the VBSs such that the VBS v $(v = 1, \ldots, V)$ receives r_v RUs. Hence, at any given moment it must hold that

$$C = \sum_{v=1}^{V} r_v \tag{1}$$

We consider K different service-classes. The service-classes are distinguished by the number of RUs requested by a single call that originates from a MU. We assume that the calls follow a Poisson distribution. The arrival rate of service-class k calls is denoted as λ_k. A service-class k call requests r_k RUs. If the requested RUs are available in the cluster, then the call is accepted in the system and stays in-service for a generally distributed holding time whose mean value is denoted as μ_k^{-1}. Such CAC policy, where all calls compete for all the available RUs, is referred to as the complete sharing policy [35]. The ratio λ_k/μ_k represents the traffic intensity.

Let us denote as n_k the number of in-service calls of service-class k. Then, the total number occupied RUs in the cluster is given by

$$j = \sum_{k=1}^{K} n_k r_k \tag{2}$$

In the following, j is considered as the system state and takes integer values in $[0, \ldots, C]$. When a new service-class k is accepted in the cluster, the system state moves from j to a higher state $j + b_k$. Similarly, when a service-class k leaves the cluster, the state moves from j to a lower state $j - b_k$. As a tutorial example consider a cluster with C RUs and $K = 2$. The calls' RU requirements are $r_1 = 1$, $r_2 = 2$. As shown in Fig. 4, the transitions toward higher states depend on the call arrival rate, λ_k, while the transitions toward lower states depend on the service rate, μ_k, (which is the inverse of the mean holding time) and the mean number, $Y_k(j)$, of service-class k calls in a given state j. We also observe that the 1st service-class calls have one blocking state: $j = C$, wheres the 2nd service-class calls have two blocking states: $j = C - 1$ and $j = C$. In general, when the RU demand is r_k, then the blocking states are: $j > C - r_k$.

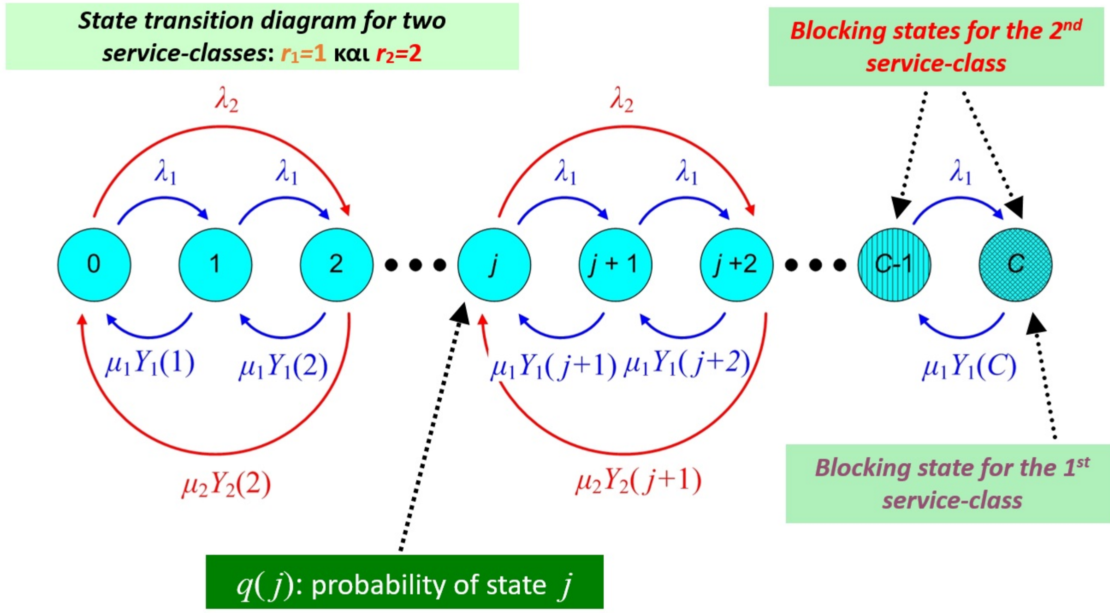

Fig. 4: State Transition Diagram.

We denote by $q(j)$ the probability of state j. Based on the above discussion, in order to calculate the CBP, B_k, of a service-class k, we need to add the probabilities, $q(j)$'s of all the blocking states

$$B_k = \sum_{j=C-r_k+1}^{C} q(j) \tag{3}$$

In order to calculate the CBP via (3), we need to determine $q(j)$'s. By exploiting the fact that the Markov chain of the model is reversible, we can write the following LB equation (rate-up = rate-down) between the adjacent states j and $j + r_k$

$$\lambda_k q(j) = \mu_k Y_k(j + r_k) q(j + r_k) \tag{4}$$

Following the analysis of [15], the unnormalized state probabilities, $\hat{q}(j)$, are accurately determined via

$$j\hat{q}(j) = \sum_{j=1}^{C} \frac{\lambda_k}{\mu_k} r_k \hat{q}(j - r_k), j = 1, \ldots, C \tag{5}$$

with $\hat{q}(0) = 1$ and $\hat{q}(j) = 0$ for $j < 0$.

Finally, $q(j)$'s are derived by dividing $\hat{q}(j)$'s with the normalization factor G

$$q(j) = \frac{\hat{q}(j)}{G}, j = 0, \ldots, C \tag{6}$$

with $G = \sum_{j=0}^{C} \hat{q}(j)$.

IV. Numerical Results

For the evaluation purposes, we simulate the RAN, as described in Section II, using the NS-3 simulator [36]. Call arrivals and departures as well as the CAC mechanism are simulated according to the model description of Section III. A total of 10 million calls has been generated in each run. When there are available RUs in the RAN cluster, the arriving call is accepted. Otherwise, the call is blocked. The simulator

TABLE I: Call arrival rate cases (1st scenario)

	1st service-class	*2nd service-class*	*3rd service-class*
case number	λ_1	λ_2	λ_3
1	4.25	2.00	1.00
2	4.50	2.10	1.02
3	4.75	2.20	1.02
4	5.00	2.30	1.04
5	5.25	2.35	1.06
6	5.50	2.40	1.06
7	5.75	2.45	1.08
8	6.00	2.50	1.10

records such blocking events to produce in the end CBP for each service-class. We also analytically calculate CBP as described in Section III. The calculation of state probabilities and CBP is based on (5) and (3), respectively. Because, the analytical model is accurate, the simulation results coincide with the analytical results.

As first experimental scenario, we simulate a cluster with 50 RUs (this corresponds to an LTE channel of 9 MHz) and $K = 3$ service-classes. The resource requirements are $r_1 = 1$ RU, $r_2 = 2$ RUs, and $r_3 = 5$ RUs. The mean call holding times are $\mu_1^{-1} = 4$ min., $\mu_2^{-1} = 2$ min., and $\mu_3^{-1} = 1$ min. Initially we study the impact of the call arrival rate on the CBP. The call arrival rates, λ_k $(k = 1, 2, 3)$, vary as shown in Table I. In the y-axis of Fig. 5, we show the resultant CBP. We observe that the CBP of the first two service-classes are relatively low and in all cases remain below the 5% threshold, which is an acceptable performance at the call level. On the other hand, due to its high resource requirement, the 3rd service-class has much higher CBP and is below the 5% only for the first three cases of the call arrival rate. Furthermore, to observe the impact of the mean holding time on the CBP, we keep fixed the call arrival rates to $\lambda_1 = 5$, $\lambda_2 = 2.3$, $\lambda_3 = 1.04$ and vary μ_1^{-1} from 1 to 6 min., as shown in the x-axis of Fig. 6. We observe that for $\mu_1^{-1} \leq 3$ min., the CBP is below the 5% threshold for all three service-classes. As μ_1^{-1} increases, the CBP, especially of the 3rd service-class, also increases reaching the unacceptable $B_3 = 25\%$ when $\mu_1^{-1} = 6$. This example indicates that particular care must be taken when planning and dimensioning the RAN to avoid high CBP.

In the second scenario, we consider a cluster with 100 RUs (this corresponds to an LTE channel of 18 MHz) and $K = 5$ service-classes. The resource requirements of the service-classes are $r_k = k$ $(k = 1, 2, 3)$, $r_4 = 5$, and $r_5 = 8$ RUs. The mean call holding times are $\mu_1^{-1} = 6$ min., $\mu_2^{-1} = 3$ min., $\mu_3^{-1} = 2$ min., $\mu_4^{-1} = 2$ min., and $\mu_5^{-1} = 1$ min. Initially we study the impact of the call arrival rate on the CBP. The call arrival rates, λ_k $(k = 1, \ldots, 5)$, vary as shown in Table II. In the y-axis of Fig. 7, we show the resultant CBP. We observe that, assuming 5% as an acceptable CBP threshold, the first three service-classes are below the threshold in all the considered cases. The 4th and 5th service-classes, show considerably higher CBP. Hence, the particular RAN cluster will need to expand its capacity above the current 100 RUs to satisfy the threshold for the last two service-classes. To observe the impact of the mean holding time on the CBP, we keep fixed the call arrival rates to $\lambda_1 = 5$, $\lambda_2 = \lambda_3 = 2$, $\lambda_4 = \lambda_5 = 1$ and vary the mean holding time of the 2nd service-class as shown in the x-axis of Fig. 8. Similarly to the 1st scenario, here we also observe that the increase of μ_2^{-1} can significantly impact the CBP of all service-classes. When $\mu_2^{-1} \leq 3$ min., then the acceptable performance levels are kept for all service-classes. Finally, to observe the impact of the system capacity on the CBP, we keep fixed λ_k's and μ_k^{-1}'s, and vary the system capacity C as shown in the x-axis of Fig. 9. For $C \leq 80$, almost all CBP are above the 5% threshold. On the contrary, when $C = 100$, the performance levels of all service-classes are acceptable. Furthermore, it seems that increasing the capacity to $C = 110$ does not significantly improve the performance.

TABLE II: Call arrival rate cases (2nd scenario)

	λ_1	λ_2	λ_3	λ_4	λ_5
1	5.0	2.0	2.00	1.00	1.00
2	5.1	2.1	2.05	1.05	1.01
3	5.2	2.2	2.10	1.10	1.02
4	5.3	2.3	2.15	1.15	1.03
5	5.4	2.4	2.20	1.20	1.04
6	5.5	2.5	2.25	1.25	1.05

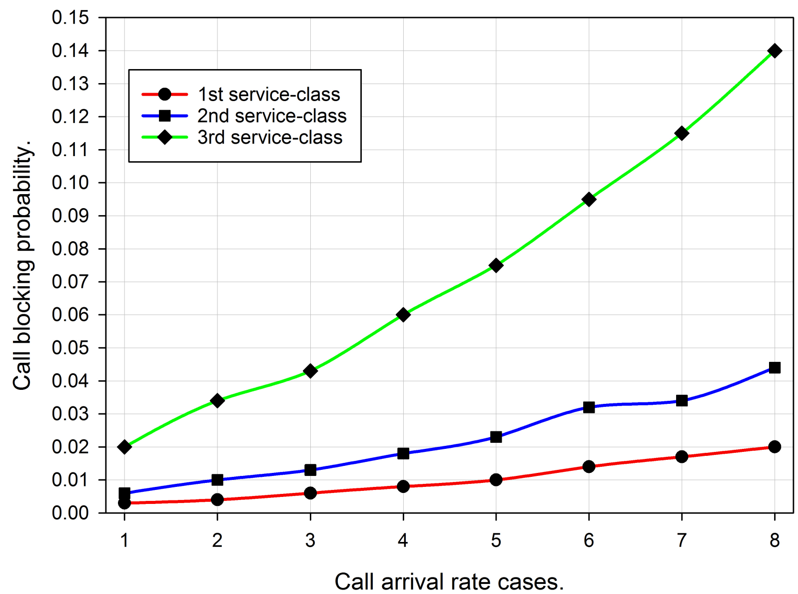

Fig. 5: Call blocking probabilities vs call arrival rate (1st scenario).

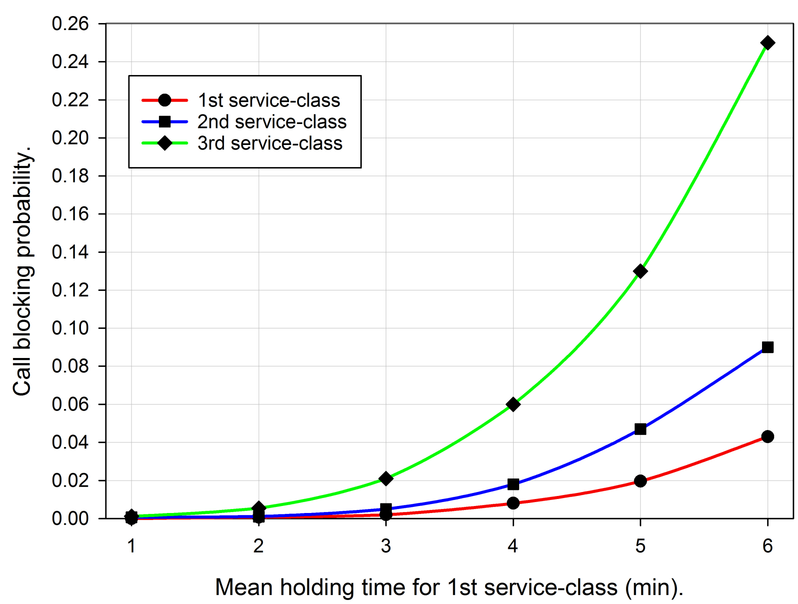

Fig. 6: Call blocking probabilities vs mean call holding time (1st scenario).

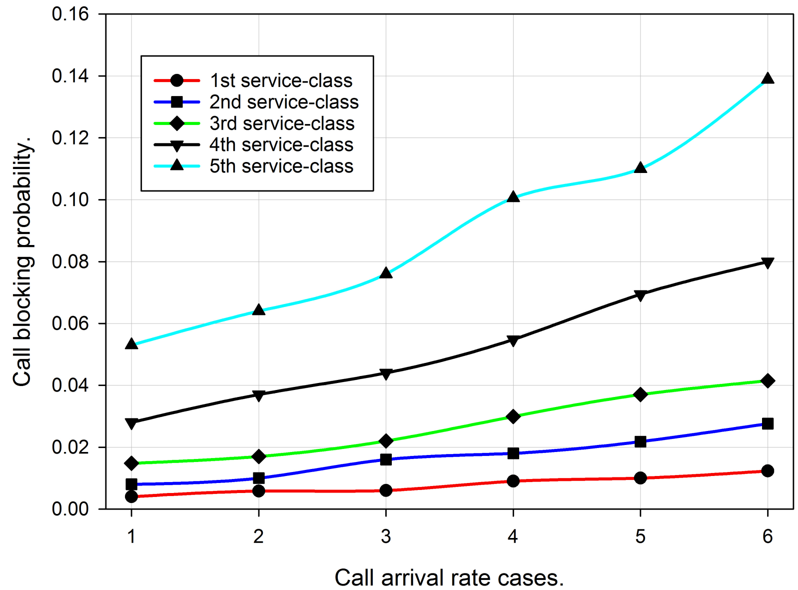

Fig. 7: Call blocking probabilities vs call arrival rate (2nd scenario).

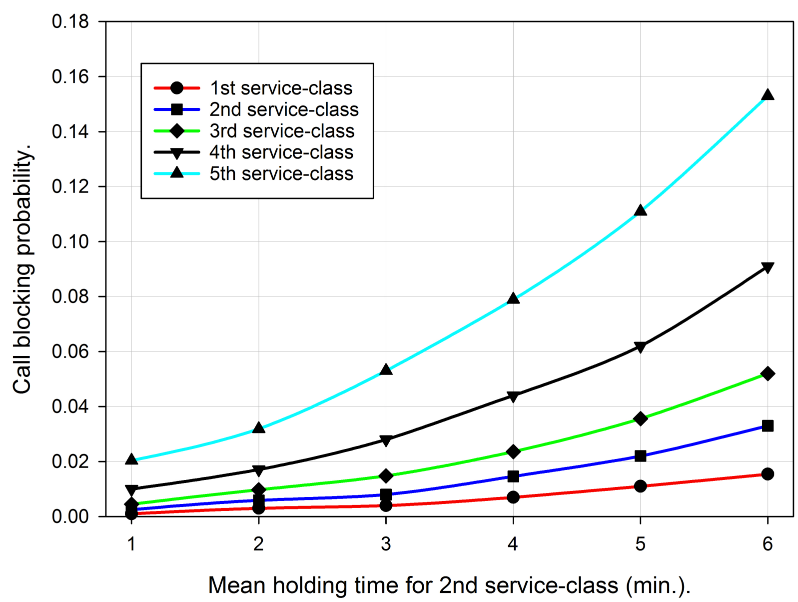

Fig. 8: Call blocking probabilities vs mean call holding time (2nd scenario).

V. Conclusion and Future Work

This work considers the SDN/NFV based cellular network architecture, the OFDMA multiple-access scheme, and the complete resource sharing policy. We study the problem of fast and efficient radio resource management, focusing on the radio resource allocation and call admission control. In particular, we investigate the applicability of the Kaufman-Roberts algorithm for determining the system state probabilities and CBP. The derived formulae can be used for call admission control and for flexible radio resource allocation to virtualized network elements, thus assuring appropriate QoS of mobile users. Since the Markov chain of the proposed model is reversible and due to the local balance between adjacent

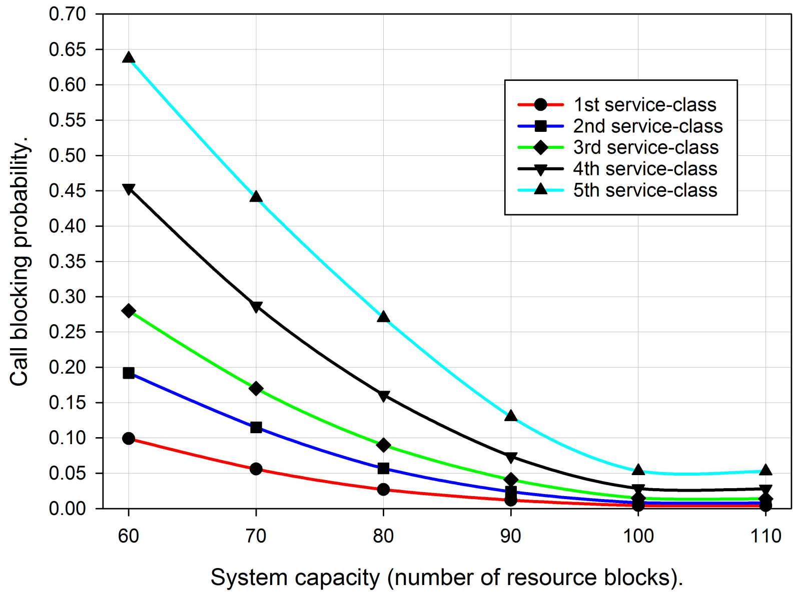

Fig. 9: Call blocking probabilities vs system capacity (2nd scenario).

states, the analytical results coincide with the simulation results. We present numerical CBP results for two scenarios in a RAN cluster while varying the call arrival rate, the mean call holding time, and the system capacity. These observations can be used for appropriate network planning and dimensioning purposes to avoid performance degradation of some service-classes, especially of those with high resource demands. As a future research direction, we intend to extend the derived algorithm for more complicated resource sharing policies, such as the bandwidth reservation (BR) [37] and threshold (TH) [38] policies. The BR policy introduces a higher priority in order, for example, to benefit certain service-classes or to achieve CBP equalization among calls of different service-classes. Similar goals can be achieved by the TH policy that uses pre-defined thresholds for the number of in-service calls of different service-classes.

VI. Appendix: List of acronyms/abbreviations

AP: Access Point
API: Application Programming Interface
BR: Bandwidth Reservation
BS: Base Station
CAC: Call Admission Control
CBP: Call Blocking Probabilities
CDMA: Code Division Multiple Access
CSC: Core SDN Controller
CTMC: Continuous Time Markov Chan
FDMA: Frequency Division Multiple Access
FE: Forwarding Element
IP: Internet Protocol
K-R: Kaufman-Roberts
LB: Local Balance
LO-GW: Local Offload Gateway
LSC: Local SDN Controler

LTE: Long Term Evolution
M-CDN: Mobile Content Delivery Network
MBS: Macro Base Station
MCC: Mobile Cloud Computing
MCN: Mobile Core Network
MU: Mobile User
NFV: Network Function Virtualization
NFVI: Network Function Virtualization Infrastructure
OFDMA: Orthogonal Frequency Division Multiple Access
PDN: Packet Data Network
PDN-GW: Packet Data Network Gateway
PNF: Physical Network Function
QoE: Quality of Experience
QoS: Quality of Service
RAN: Radio Access Network
RAT: Radio Access Technology
RB: Resource Block
RRA: Radio Resource Allocation
RRM: Radio Resource Management
RU: Resource Unit
S-GW: Serving Gateway
SBS: Small Cell Base Station
SDN: Software-Defined Networking
TDMA: Time Division Multiple Access
TH: Threshold
VBS: Virtual Base Station
VM: Virtual Machine
VNF: Virtual Network Function

REFERENCES

[1] Sesia S., Baker M., Toufik I.: ‘LTE - The UMTS long term evolution: From theory to practice’ (John Wiley & Sons, 2011)

[2] Demestichas P., Georgakopoulos A., Karvounas D., Tsagkaris K., Stavroulaki V., Lu J. Xiong C,. Yao J.: ‘5G on the horizon: Key challenges for the radio-access network’, *IEEE Vehicular Technology Magazine*, 2013, **8** (3), pp. 47-53

[3] Pierre S.: ‘Quality of experience aware multimedia communications over heterogeneous wireless networks’, *IET Networks*, 2016, **5** (2), pp. 37-43

[4] Andrews J. G.: ‘Seven ways that HetNets are a cellular paradigm shift’, *IEEE Communications Magazine*, 2013, **51** (3), pp. 136-144

[5] López L., Caraguay A., Villalba L., López D.: ‘Trends on virtualisation with software defined networking and network function virtualisation’, *IET Networks*, 2015, **4** (5), pp. 255-263

[6] Nunes B., Mendonca M., Nguyen X., Obraczka K., Turletti T.: ‘A survey of software-defined networking: Past, present, and future of programmable networks’, *IEEE Communications Surveys & Tutorials*, 2014, **16** (3), pp. 1-18

[7] Xia W., Wen Y., Foh C. H., Niyato D., Xie H.: ‘A survey on software-defined networking’, *IEEE Communications Surveys & Tutorials*, 2015, **17** (1), pp. 27-51

[8] Mijumbi R., Serrat J., Gorricho J.-L., Bouten N., De Turck F., Boutaba R.: ‘Network function virtualization: State-of-the-art and research challenges’, *IEEE Communications Surveys & Tutorials*, 2016, **18** (1), pp. 236-262

[9] Han B., Gopalakrishnan V., Ji L., Lee S.: ‘Network function virtualization: Challenges and opportunities for innovations’, *IEEE Communications Magazine*, 2015, **53** (2), pp. 90-97

[10] Duan Q., Ansari N., Toy M.: ‘Software-defined network virtualization: An architectural framework for integrating SDN and NFV for service provisioning in future networks’, *IEEE Network*, 2016, **30** (5), pp. 10-16

[11] Agyapong P. K., Iwamura M., Staehle D., Kiess W., Benjebbour A.: ‘Design considerations for a 5G network architecture’, *IEEE Commun. Mag.*, 2014, **52** (11), pp. 65-75

[12] Gudipati A., Perry D., Li L. E., Katti S.: ‘SoftRAN: Software defined radio access network’. *Proc. 2nd ACM SIGCOMM workshop on Hot Topics in Software Defined Networking*, Aug. 2013, Hong Kong, China, pp. 25-30

[13] Li L. E., Mao Z. M., Rexford J.: ‘Toward software-defined cellular networks’. *Proc. IEEE European Workshop on Software Defined Networking (EWSDN)*, Oct. 2012, Darmstadt, Germany, pp. 7-12

[14] Jin X., Li L. E., Vanbever L., Rexford J.: Softcell: Scalable and flexible cellular core network architecture’. *Proc. 9th ACM Conf. on Emerging Networking Experiments and Technologies (CoNEXT)*, Dec. 2013, Santa Barbara, CA, USA, pp. 163-174

[15] Kaufman J.: 'Blocking in a shared resource environment', *IEEE Trans. Commun.*, 1981, **29** (10), pp. 1474-1481

[16] Roberts J. W.: 'A service system with heterogeneous user requirements', in Pujolle G. (Ed.): 'Performance of Data Communications Systems and Their Applications' (North Holland, Amsterdam, 1981), pp. 423-431

[17] Brewka L., Iversen V., Kardaras G.: 'Integrated service resource reservation using queueing networks theory,' *IET Networks*, 2014, **3** (1), pp. 16-21

[18] Huang Y., Ko K. T., Zukerman M.: 'Performance approximation and bounds for non-real-time traffic in an integrated service system', *IET Networks*, 2014, **3** (1), pp. 22-29

[19] Moscholios I. D., Logothetis M. D., Vardakas J. S., Boucouvalas A. C.: 'Performance metrics of a multirate resource sharing teletraffic model with finite sources under the threshold and bandwidth reservation policies', *IET Networks*, 2014, **4** (3), pp. 195-208

[20] Moscholios I. D., Logothetis M. D., Vardakas J. S., Boucouvalas A. C.: 'Congestion probabilities of elastic and adaptive calls in Erlang-Engset multirate loss models under the threshold and bandwidth reservation policies', *Computer Networks*, 2015, **92**, pp. 1-23

[21] Stasiak M., Wisniewski A., Zwierzykowski P., Glabowski M.: 'Blocking probability calculation for cellular systems with WCDMA radio interface servicing PCT1 and PCT2 multirate traffic', *IEICE Trans. on Commun.*, 2009, **E92-B** (4), pp. 1156-1165

[22] Vassilakis V. G., Moscholios I. D., Logothetis M. D.: 'Quality of service differentiation in heterogeneous CDMA networks: A mathematical modelling approach', *Wireless Networks*, 2017 (in press), doi:10.1007/s11276-016-1411-z

[23] Vassilakis V. G., Moscholios I. D., Logothetis M. D.: 'Uplink blocking probabilities in priority-based cellular CDMA networks with finite source population', *IEICE Transactions on Communications*, 2016, **99** (6), pp. 1302-1309

[24] Beyranvand H., Maier M., Salehi J. A.: 'An analytical framework for the performance evaluation of node-and network-wise operation scenarios in elastic optical networks', *IEEE Transactions on Communications*, 2014, **62** (5), pp. 1621-1633

[25] Vardakas J. S., Moscholios I. D., Logothetis M. D., Stylianakis V. G.: 'Performance analysis of OCDMA PON configuration supporting multi-rate bursty traffic with retrials and QoS differentiation', *Optical Switching and Networking*, 2014, **13**, pp. 112-123

[26] Vardakas J. S., Vassilakis V. G., Logothetis M. D.: 'Loss models in traffic-groomed WDM all-optical networks'. *Proc. 6th International Symposium on Communication Systems, Networks and Digital Signal Processing (CNSDSP)*, July 2008, Graz, Austria, pp. 217-221

[27] Vassilakis V. G., Moscholios I. D., Alzahrani B. A., Logothetis M. D.: 'A software-defined architecture for next-generation cellular networks'. *Proc. IEEE International Conference on Communications (ICC)*, May 2016, Kuala Lumpur, Malaysia, pp. 1-6

[28] Chen T., Matinmikko M., Chen X., Zhou X., Ahokangas P.: 'Software defined mobile networks: Concept, survey, and research directions', *IEEE Commun. Mag.*, 2015, **53** (11), pp. 126-133

[29] ETSI GS NFV-MAN 001 (V1.1.1), 'Network Function Virtualisation (NFV); Management and Orchestration', Dec. 2014

[30] Samdanis K., Taleb T., Schmid S.: 'Traffic offload enhancements for eUTRAN', *IEEE Communications Surveys & Tutorials*, 2012, **14** (3), pp. 884-896

[31] Poularakis K., Iosifidis G., Tassiulas L.: 'Approximation algorithms for mobile data caching in small cell networks', *IEEE Transactions on Communications*, 2014, **62** (10), pp. 3665-3677

[32] Shen Y., Wang Y., Peng Z., Wu S.: 'Multiple-access interference mitigation for acquisition of code-division multiple-access continuous-wave signals', *IEEE Communications Letters*, 2017, **21** (1), pp. 192-195

[33] Staehle D., Mäder A.: 'An analytic approximation of the uplink capacity in a UMTS network with heterogeneous traffic', *Teletraffic Science and Engineering*, 2003, **5**, pp. 81-90

[34] Vassilakis V. G., Moscholios I. D., Bontozoglou A., Logothetis M. D.: 'Mobility-aware QoS assurance in software-defined radio access networks: An analytical study'. *Proc. IEEE Conference on Network Softwarization (NetSoft)*, April 2015, London, U.K., pp. 1-6

[35] Moscholios I. D., Logothetis M. D., Kokkinakis G. K.: 'Call-burst blocking of ON-OFF traffic sources with retrials under the complete sharing policy', *Performance Evaluation*, 2005, **59** (4), pp. 279-312

[36] 'Network Simulator NS-3', http://www.nsnam.com, accessed October 2017

[37] Glabowski M., Kaliszan A., Stasiak M.: 'Asymmetric convolution algorithm for blocking probability calculation in full-availability group with bandwidth reservation', *IET Circuits, Devices & Systems*, 2008, **2** (1), pp. 87-94

[38] Moscholios I. D., Vassilakis V. G., Logothetis M. D, Boucouvalas A. C.: 'A probabilistic threshold-based bandwidth sharing policy for wireless multirate loss networks', *IEEE Wireless Communications Letters*, 2016, **5** (3), pp. 304-307