UWL REPOSITORY

repository.uwl.ac.uk

A probabilistic threshold-based bandwidth sharing policy for wireless multirate loss networks

This is the Accepted Version of the final output.

UWL repository link: https://repository.uwl.ac.uk/id/eprint/2686/

Alternative formats: If you require this document in an alternative format, please contact: open.research@uwl.ac.uk

A Probabilistic Threshold-based Bandwidth Sharing Policy for Wireless Multirate Loss Networks

I. D. Moscholios, V. G. Vassilakis, M. D. Logothetis, *Senior Member, IEEE,* and A. C. Boucouvalas, *Fellow, IEEE*

Abstract—We propose a probabilistic bandwidth sharing policy, based on the threshold (TH) policy, for a single cell of fixed capacity in a homogeneous wireless cellular network. The cell accommodates random input-traffic originated from K service-classes. We distinguish call requests to new and handover and therefore the cell supports $2K$ types of arrivals. If the number of in-service calls (new or handover) of a service-class exceeds a threshold (different for new and handover calls of a service-class), a new or handover arriving call of the same service-class is not always blocked, as it happens in the TH policy, but it is accepted in the system with a predefined state-dependent probability. The cell is analysed as a multirate loss system, via a reversible continuous-time Markov chain, which leads to a product form solution (PFS) for the steady state distribution. Thanks to the PFS, the calculation of performance measures is accurate, but complex. To reduce the computational complexity, we determine performance measures via a convolution algorithm.

Index Terms—Call admission, call blocking, threshold, product form, wireless networks.

I. INTRODUCTION

PERFORMANCE modelling in wireless cellular networks requires the selection of appropriate multirate loss or queueing models which can be described via efficient convolutional or recursive formulas. The latter are essential since they reduce the computational complexity and therefore can be invoked in network planning and dimensioning procedures.

Considering call-level traffic in a single cell of a homogeneous wireless cellular network which accommodates different service-classes with different bandwidth requirements, a bandwidth sharing policy is essential since it gives access to the necessary bandwidth needed by the services and affects performance measures, such as call blocking probabilities (CBP). The simplest policy is the complete sharing (CS) policy, where a new call is accepted in the system if its bandwidth is available [1]. The CS policy leads to recursive CBP formulas but: 1) it is unfair to calls with higher bandwidth requirements since it leads to higher CBP [2] and 2) it does not provide different treatment to handover calls. These reasons motivate research on policies, such as the bandwidth reservation (BR), the multiple fractional channel reservation (MFCR), or the threshold (TH) policy (see e.g., [2]- [4]). The BR policy achieves CBP equalization among calls of different service-classes at the cost of substantially increasing the

CBP of service-classes with lower bandwidth requirements. The MFCR policy extends the BR policy by allowing the reservation of real (not integer) number of channels [2]. Both policies lead to non-reversible continuous time Markov chains and consequently to approximate CBP formulas (e.g., [2]).

We concentrate on the TH policy, because it is broadly applicable in wired, wireless and satellite networks (e.g., [3]-[7]) and it is attractive in access tree networks. Under the TH policy, the number of in-service calls of a specific service-class must not exceed a pre-defined threshold (dedicated to the service-class), after the acceptance of a new call of this service-class.

We propose a probabilistic TH policy (PrTH) for a single cell of fixed capacity that accommodates Poisson arriving calls of different service-classes. In this new policy, if bandwidth is available, call acceptance above a threshold is permitted with a predefined probability dependent on the service-class, the type of call (new or handover) and the system state. The proposed loss system is described by a reversible continuous time Markov chain, resulting in a product form solution (PFS) for the steady state distribution. Thanks to the existence of PFS, we determine accurately CBP and system's utilization via a convolution algorithm.

Considering wireless systems, potential applications of the PrTH policy are not only in homogeneous cellular networks, but also in heterogeneous radio access networks (RAN) based on the emerging Cloud RAN architecture. The latter is expected to provide a cost-efficient RAN with a great scalability, through the use of network function virtualization (NFV) techniques and data center processing capabilities [8]. A function that could be virtualized is the radio resource management (RRM) function, which is responsible for call admission control and radio resource allocation to mobile users (MUs). The proposed PrTH policy could be implemented at the RRM function and would enable sharing of virtualized baseband resources among MUs. In practice, this can be realised by exploiting the self organising network (SON) functionalities [9]. In particular, the centralised decision of PrTH policy parameters can be done at the network management system (NMS) level (i.e., using a centralised SON mechanism). These parameters can then be communicated to the RAN elements to configure RRM algorithms in a distributed SON fashion.

In Section II, we present the proposed policy (PrTH), show the PFS and provide a convolution algorithm for the calculation of performance measures. In Section III, we present analytical and simulation CBP results of the proposed policy and analytical CBP results of the CS, TH, BR and MFCR policies. We also conclude in Section III.

I. D. Moscholios and A. C. Boucouvalas are with the Dept. of Informatics & Telecommunications, University of Peloponnese, 221 00 Tripolis, Greece.

V. G. Vassilakis is with the Computer Laboratory, University of Cambridge, CB3 0FT Cambridge, UK.

M. D. Logothetis is with the Department of Electrical & Computer Engineering, University of Patras, 265 04 Patras, Greece.

II. THE PROPOSED MODEL

Consider a single cell of capacity C channels that accommodates calls of K service-classes under the PrTH policy. In order to distinguish between new and handover calls of a service-class, we assume that the system accommodates $2K$ service-classes. A service-class k call is new if $1 \le k \le K$ and handover if $K + 1 \le k \le 2K$. Thus, a new service-class k call has the same channel requirements with a handover service-class $K + k$ call. A service class k $(k = 1, ..., 2K)$ follows a Poisson process with arrival rate λ_k and requests b_k channels. If these channels are not available then the call is blocked and lost; otherwise:

a) If the number n_k of in-service calls of service-class k $(k = 1, ..., 2K)$ in the steady state plus the new or handover call, does not exceed a predefined threshold n_k^*, i.e., $n_k + 1 \le n_k^*$, then the new call is accepted in the system.

b) If $n_k + 1 > n_k^*$, the call is accepted with probability $p_k(n_k)$ or blocked with probability $1 - p_k(n_k)$. The set of $p_k(n_k)$ constitutes the vector:

$$\boldsymbol{p_k} = (p_k(0), p_k(1)..., p_k(n_k^*), ..., p_k(\lfloor C/b_k \rfloor)) \quad (1)$$

where $\lfloor C/b_k \rfloor$ is the maximum number of service-class k calls that the system can service.

In (1), we assume that:

1) $p_k(0) = p_k(1) = ... = p_k(n_k^*-1) = 1$, i.e., a service-class k call is accepted if the threshold n_k^* is not exceeded.

2) the probabilities $p_k(n_k^*), ..., p_k(\lfloor C/b_k \rfloor - 1)$ may be different for new or handover calls of the same service-class k. In the ordinary TH policy [3], these probabilities are all zero. In the proposed PrTH policy, they can be set either all positive, or zero after a certain number greater than n_k^*. To prioritize handover calls over new calls of a service-class k, we can choose higher values for the thresholds of handover calls, or set the corresponding probabilities of handover calls to higher values than those of new calls. E.g., assuming medium to heavy traffic-load conditions, when the arrival rate of handover calls becomes much larger than that of new calls then the probabilities of new calls should be reduced (and those of handover calls should be increased) in order to lower the CBP of handover calls. On the other hand, when the arrival rate of handover calls is quite low, then the probabilities of new calls can be increased in order to decrease the CBP of new calls.

3) $p_k(\lfloor C/b_k \rfloor) = 0$, due to lack of available bandwidth.

An accepted service-class k call has a generally distributed service time with mean μ_k^{-1}. The total offered traffic-load (in erl) of service-class k calls is $\alpha_k = \lambda_k \mu_k^{-1}$.

Let the steady state vector be $\boldsymbol{n} = (n_1, ..., n_k, ..., n_{2K})$ and $\boldsymbol{n_k^-} = (n_1, ..., n_k - 1, ..., n_{2K})$, $\boldsymbol{n_k^+} = (n_1, ..., n_k + 1, ..., n_{2K})$.

Based on the state transition diagram (Fig. 1) of the proposed PrTH model, the global balance (GB) equation for state $\boldsymbol{n}$, expressed as *rate into state $\boldsymbol{n}$ = rate out of state $\boldsymbol{n}$*, is:

$$\sum_{k=1}^{2K} \left[\lambda_k \delta_k^-(\boldsymbol{n}) p_k(n_k-1) P(\boldsymbol{n_k^-}) + (n_k+1) \mu_k \delta_k^+(\boldsymbol{n}) P(\boldsymbol{n_k^+}) \right]$$
$$= \sum_{k=1}^{2K} \left[\lambda_k \delta_k^+(\boldsymbol{n}) p_k(n_k) P(\boldsymbol{n}) + n_k \mu_k \delta_k^-(\boldsymbol{n}) P(\boldsymbol{n}) \right] \quad (2)$$
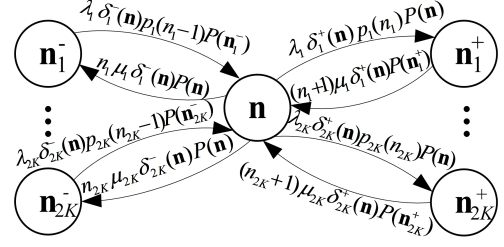

Fig. 1. State transition diagram of the proposed PrTH.

where: $\delta_k^+(\boldsymbol{n}) = \begin{cases} 1 & \text{if } \boldsymbol{n_k^+} \in \boldsymbol{\Omega} \\ 0 & \text{otherwise} \end{cases}$, $\delta_k^-(\boldsymbol{n}) = \begin{cases} 1 & \text{if } \boldsymbol{n_k^-} \in \boldsymbol{\Omega} \\ 0 & \text{otherwise} \end{cases}$,
$\boldsymbol{\Omega}$ is the state space of the system denoted by $\boldsymbol{\Omega} = \{\boldsymbol{n} : 0 \le \boldsymbol{nb} \le C, k = 1, ..., 2K\}$, $\boldsymbol{nb} = \sum_{k=1}^{2K} n_k b_k$, $\boldsymbol{b} = (b_1, ..., b_{2K})^T$ and $P(\boldsymbol{n}), P(\boldsymbol{n_k^-}), P(\boldsymbol{n_k^+})$ are the probability distributions of states $\boldsymbol{n}, \boldsymbol{n_k^-}, \boldsymbol{n_k^+}$, respectively.

The Markov chain (of Fig. 1) of the proposed PrTH model is reversible since it meets the Kolmogorov's criterion [10]: the circulation flow among four adjacent states forming a closed circle equals zero: Flow_clockwise = Flow_counter-clockwise. Because of this, local balance (LB) exists between adjacent states and the following LB equations are extracted as *(rate up = rate down)*, for $k = 1, ..., 2K$ and $\boldsymbol{n} \in \Omega$:

$$\lambda_k \delta_k^-(\boldsymbol{n}) p_k(n_k - 1) P(\boldsymbol{n_k^-}) = n_k \mu_k \delta_k^-(\boldsymbol{n}) P(\boldsymbol{n}) \quad (3)$$

$$\lambda_k \delta_k^+(\boldsymbol{n}) p_k(n_k) P(\boldsymbol{n}) = (n_k + 1) \mu_k \delta_k^+(\boldsymbol{n}) P(\boldsymbol{n_k^+}) \quad (4)$$

The system of LB equations (3)-(4) is satisfied by the PFS:

$$P(\boldsymbol{n}) = G^{-1} \prod_{k=1}^{2K} \prod_{x=n_k^*}^{n_k-1} p_k(x) \frac{\alpha_k^{n_k}}{n_k!} \quad (5)$$

where G is the normalization constant given by:

$$G \equiv G(\boldsymbol{\Omega}) = \sum_{\boldsymbol{n} \in \Omega} \left(\prod_{k=1}^{2K} \prod_{x=n_k^*}^{n_k-1} p_k(x) \frac{\alpha_k^{n_k}}{n_k!} \right)$$

To calculate the CBP of service-class k calls, B_k, we define the state space $\boldsymbol{\Omega}_k = \{\boldsymbol{n} : 0 \le \boldsymbol{nb} \le C - b_k, k = 1, ..., 2K\}$ which denotes the set of states for which a service-class k call will be definitely accepted or accepted with a state-dependent probability in the system. Thus:

$$B_k = 1 - G_k/G \quad (6)$$

where: $G_k = \sum_{\boldsymbol{n} \in \Omega_k} p_k(n_k) P(\boldsymbol{n})$

The computational complexity of (6) is in the order of $O(C^{2K})$ (see [11] for the CS policy). To reduce it to $O(2KC^2)$, we exploit the PFS and present a 3-step convolution algorithm for the exact calculation of CBP and system's utilization:

Let j be the occupied system's bandwidth, $j = 0, 1, ..., C$.
Step 1) Determine the occupancy distribution $q_k(j)$ of each service-class k $(k = 1, ..., 2K)$, assuming that only service-class k exists in the system:

$$q_k(j) = \begin{cases} \dfrac{q_k(0)\alpha_k^i}{i!}, & \text{for } 1 \le i \le n_k^* \text{ and } j = ib_k \\[2ex] \dfrac{q_k(0) \prod_{x=n_k^*}^{i-1} p_k(x)\alpha_k^i}{i!}, & \\[1ex] & \text{for } n_k^* < i \le \lfloor C/b_k \rfloor \text{ and } j = ib_k \end{cases} \quad (7)$$

Step 2) Determine the aggregated occupancy distribution $Q_{(-k)}$ based on the successive convolution of all service-classes apart from service-class k:

$$Q_{(-k)} = q_1 * \dots * q_{k-1} * q_{k+1} * \dots * q_{2K}.$$

By the term "successive" we mean that we initially convolve q_1 and q_2 in order to obtain q_{12}. Then we convolve q_{12} with q_3 to obtain q_{123} etc. The convolution operation between two occupancy distributions of service-class k and r is defined as:

$$q_k * q_r = \left\{ q_k(0)q_r(0), \sum_{x=0}^{1} q_k(x)q_r(1-x), \dots, \sum_{x=0}^{C} q_k(x)q_r(C-x) \right\} \quad (8)$$

Step 3) Calculate the CBP of service-class k based on the convolution operation of $Q_{(-k)}$ (step 2) and q_k as follows:

$$Q_{(-k)} * q_r = \left\{ \begin{array}{l} Q_{(-k)}(0)q_k(0), \sum_{x=0}^{1} Q_{(-k)}(x)q_k(1-x), \\ \dots, \sum_{x=0}^{C} Q_{(-k)}(x)q_k(C-x) \end{array} \right\} \quad (9)$$

Normalizing the values of (9), we obtain the occupancy distribution $q(j)$, $j = 0, 1, \dots, C$ via the formulas:

$$q(0) = Q_{(-k)}(0)q_k(0)/G, \quad (j=0)$$
$$q(j) = \left(\sum_{x=0}^{j} Q_{(-k)}(x)q_k(j-x) \right) \Big/ G, \quad j=1, \dots, C \quad (10)$$

Due to the commutative and associative properties of the convolution operation, $q(j)$'s can be determined by convolving $Q_{(-k)}$ with q_k for any k ($k = 1, \dots, 2K$), arbitrarily selected. Based on $q(j)$'s, we propose the following CBP formula for service-class k calls:

$$B_k = \sum_{j=C-b_k+1}^{C} q(j) + \sum_{x=n_k^* b_k}^{C-b_k} (1-p_k(x))q_k(x) \sum_{y=x}^{C-b_k} Q_{(-k)}(C-b_k-y) \quad (11)$$

The first term of the right hand side of (11) refers to states j where there is no bandwidth available for service-class k calls. The second term refers to states $x = n_k^* b_k, \dots, C - b_k$ where there is available bandwidth but call blocking occurs due to the PrTH policy and the threshold n_k^*. In a state x, service-class k calls occupy x channels (depicted by $q_k(x)$) while the rest $C - b_k - y$ channels (for $y = x, \dots, C - b_k$) are occupied by all other service-classes apart from k (depicted by $Q_{(-k)}$).

The system's utilization U (in channels) is given by:

$$U = \sum_{j=1}^{C} j q(j) \quad (12)$$

where the values of $q(j)$ are given by (10).

The mean number of service-class k calls in the system, $E(n_k)$, can be determined by the formula:

$$E(n_k) = \sum_{j=1}^{C} y_k(j)q(j) \quad (13)$$

where $y_k(j)$ is the average number of service-class k calls in state j and can be calculated by:

$$y_k(j) = \alpha_k \left[\sum_{x=0}^{j-b_k} p_k(x)q_k(x)Q_{(-k)}(j-b_k-x) \right] \Big/ q(j) \quad (14)$$

The rationale behind (14) is similar to (A5) of [11] and is based on the fact that LB does exist between the adjacent states $j - b_k$ and j in the PrTH model.

As a general rule, the selection of $n_k^* >> E(n_k)$ for all service-classes decreases the effect of the PrTH policy on arriving calls and therefore leads to CBP that are close to those obtained by applying the CS policy.

As far as the probability vectors are concerned, in practice, they can be defined based on the relative frequency of handover and new calls, or on traffic forecasts. Having determined $E(n_k)$ via (13), we define a fairness index for service-class k calls, FI_k, which can be monitored on a regular basis:

$$FI_k = E(n_k) \Big/ \sum_{k=1}^{2K} E(n_k) = \alpha_k(1-B_k) \Big/ \sum_{k=1}^{2K} \alpha_k(1-B_k) \quad (15)$$

A similar index, the network fairness index (NFI_k), could be initially specified by the network administrator for service-class k calls:

$$NFI_k = \alpha_k \Big/ \sum_{k=1}^{2K} \alpha_k \quad (16)$$

If the ratio FI_k/NFI_k is near unity, then there is an agreement between the initial network specifications index for service-class k and the actual monitored index. Since the probability vector for each service-class k affects CBP and $E(n_k)$, these ratios for all service-classes could provide a guide for determining the probability vectors.

III. EVALUATION AND CONCLUSION

As an application example, we consider a cell of capacity $C = 150$ channels that accommodates two service-classes, with the traffic characteristics shown in Table I.

We provide analytical CBP results for the proposed PrTH considering three scenarios (a)-(c): (a) New calls of the 1st service-class behave as in the ordinary TH policy [3], i.e., if the number of in-service calls plus a new call exceeds $n_1^* = 35$, then call blocking occurs. So, $p_1(35) = \dots = p_1(75) = 0$. New 2nd service-class calls behave as in the PrTH policy and are accepted in the system with probability $p_2(10) = \dots = p_2(20) = 0.5$, and $p_2(21) = 0$, (b) New 1st service-class calls behave as in the PrTH policy and are accepted with probability $p_1(35) = \dots = p_1(74) = 0.7$, and $p_1(75) = 0$, while 2nd service-class calls are accepted as in scenario (a) and (c) All calls (new or handover) of both service-classes behave as in the ordinary TH policy. For scenarios (a) and (b) and both service-classes let $p_3() = p_4() = 0.95$, for all states equal or above the corresponding thresholds. These CBP results are compared with the CBP results assuming the CS, the BR and the MFCR policies [1], [2]. In the BR policy, the BR parameters are $t_1 = t_3 = 5$ channels and $t_2 = t_4 = 0$ so as to achieve CBP equalization among calls (new or handover) of both service-classes. The BR parameter of a service-class k denotes the channels reserved to benefit calls of all service-classes, apart from k. In the MFCR policy, real number of channels, $t_{r,k}$, are reserved against service-class k calls. The reservation of real number of channels is achieved because $\lfloor t_{r,k} \rfloor + 1$ channels are reserved with probability $t_{r,k} - \lfloor t_{r,k} \rfloor$ while $\lfloor t_{r,k} \rfloor$ channels are reserved with probability $1 - (t_{r,k} - \lfloor t_{r,k} \rfloor)$. The MFCR parameters are $t_{r,1} = t_{r,3} = 4.7$ channels and $t_{r,2} = t_{r,4} = 0$.

In the x-axis of Figs. 2 and 3 the offered traffic load of new and handover calls increases in steps of 1.0, 0.2, 0.5 and

TABLE I
Traffic characteristics

Service-class	Traffic-load (erl)	Bandwidth (channels)	Threshold
1st (new)	$\alpha_1 = 20.0$	$b_1 = 2$	$n_1^* = 35$
2nd (new)	$\alpha_2 = 5.0$	$b_2 = 7$	$n_2^* = 10$
1st (handover)	$\alpha_3 = 6.0$	$b_3 = 2$	$n_3^* = 70$
2nd (handover)	$\alpha_4 = 1.0$	$b_4 = 7$	$n_4^* = 20$

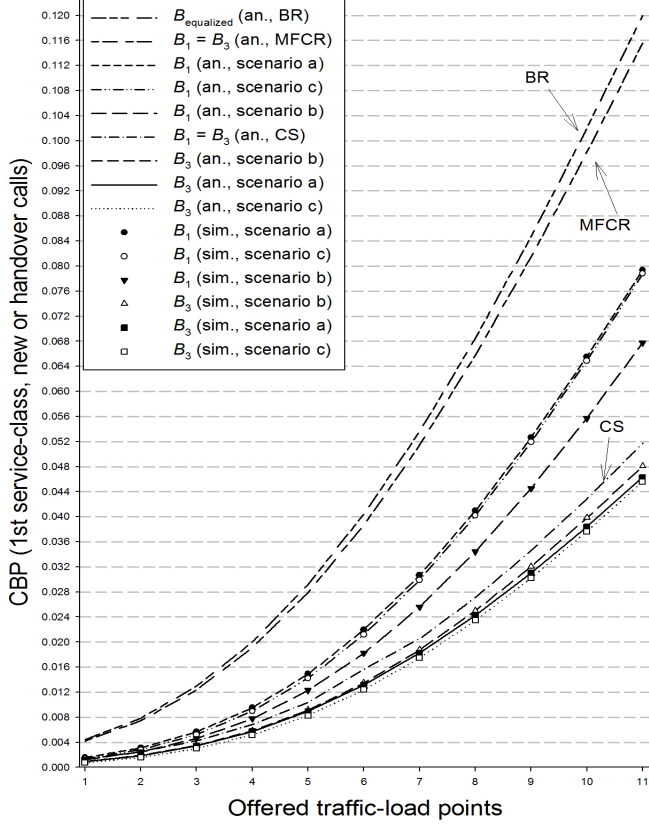Fig. 2. CBP - 1st service-class (new or handover calls).

Fig. 3. CBP - 2nd service-class (new or handover calls).

0.1 erl, respectively. So, point 1 refers to: $(\alpha_1, \alpha_2, \alpha_3, \alpha_4) = (20.0, 5.0, 6.0, 1.0)$, while point 11 to: $(30.0, 7.0, 11.0, 2.0)$. Simulation results, presented for scenarios (a)-(c), are mean values of 7 runs and are almost identical to the corresponding analytical results. Each run is based on the generation of two million calls. To account for a warm-up period, the blocking events of the first 5% of these generated calls are not considered in the results. The simulation tool used is Simscript III [12].

Concluding, both figures show that: a) The PrTH policy clearly affects the CBP of both service-classes, compared to the CBP obtained by the CS, BR and MFCR policies; thus, it enables a fine CBP control (via the selection of probability vectors and/or the thresholds) aiming at guaranteeing certain QoS to each service-class (particularly for handover calls). b) The CBP results obtained via the CS, BR and MFCR policies fail to approximate the CBP results obtained from the proposed PrTH policy; this fact reveals the necessity of the new policy.
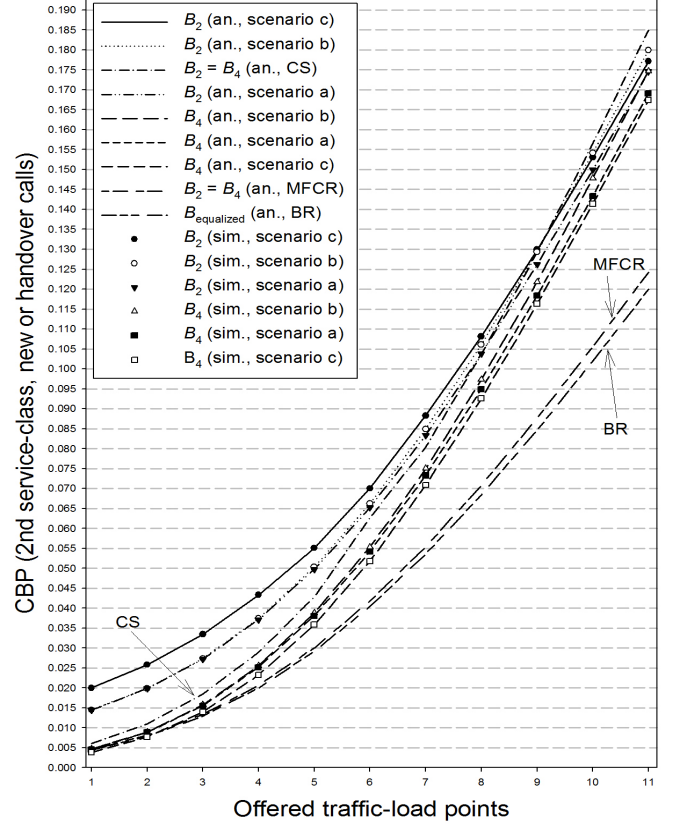
References

[1] M. Stasiak, M. Glabowski, A.Wisniewski, and P. Zwierzykowski, Modeling and Dimensioning of Mobile Networks, Wiley, 2011.
[2] F. Cruz-Pérez, J. Vázquez-Ávila, and L. Ortigoza-Guerrero, "Recurrent formulas for the multiple fractional channel reservation strategy in multiservice mobile cellular networks," IEEE Commun. Letters, vol. 8, no. 10, pp. 629-631, Oct. 2004.
[3] J. Ni et al., "Optimal and structured call admission control policies for resource-sharing systems," IEEE Trans. Commun., vol. 55, issue 1, pp. 158-170, Jan. 2007.
[4] I. Moscholios et al., "Performance metrics of a multirate resource sharing teletraffic model with finite sources under both the threshold and bandwidth reservation policies," IET Networks, vol. 4, issue 3, pp. 195-208, May 2015.
[5] A. Al Daoud, M. Alanyali, and D. Starobinski, "Pricing strategies for spectrum lease in secondary markets," IEEE/ACM Trans. Netw., vol. 18, no. 2, pp. 462-475, April 2010.
[6] X. Y. Yu and H. B. Zhu, "An efficient method for loss performance modeling of hierarchical heterogeneous wireless networks," Int. Journal of Commun. Systems, vol. 27, issue 6, pp. 956-968, June 2014.
[7] Z. Wang, P. Mathiopoulos, and R. Schober, "Channel partitioning policies for multi-class traffic in LEO-MSS," IEEE Trans. Aerosp. Electron. Syst., vol. 45, no. 4, pp. 1320-1334, Oct. 2009.
[8] A. Checko et al., "Cloud RAN for mobile networks - a technology overview," IEEE Commun. Surveys & Tutorials, vol. 17, no. 1, pp. 405-26, Sept. 2014.
[9] 3GPP 32.500 v12.1.0, "Self-Organizing Networks (SON); Concepts and requirements (Release 12)", Sept. 2014.
[10] V. Iversen, Teletraffic Engineering and Network Planning, Technical University of Denmark, May 2011.
[11] J. Kaufman, "Blocking in a shared resource environment," IEEE Trans. Commun. vol. 29, no. 10, pp. 1474-1481, Oct. 1981.
[12] Simscript III, http://www.simscript.com (last accessed: March 2016)