

Performance Analysis of Telecommunication Systems based on Time-Scale Separation ^{*}

Luis Tello-Oquendo, Vicent Pla and Jorge Martinez-Bauset

ITACA Research Institute
Universitat Politècnica de València, Spain

Abstract. The increasing complexity of the telecommunication systems has made modeling more challenging. A commonly used modeling tool are Markov chains, but the existence of several different components (e.g., user types) often renders its analysis computationally intractable. When these components operate at sufficiently separated time-scales, the quasi-stationary approximation has proven to be accurate and highly efficient. However, while the computational efficiency of this method is maintained no matter what the separation of time-scales is, its precision deteriorates rapidly as the separation narrows. In this paper, we present a new approximate method that extends the outreach of the quasi-stationary approximation by trading-off computational effort in exchange of increased accuracy. The proposed method is iterative in nature and its accuracy can be improved by performing more iterations.

1 Introduction

A large variety of telecommunication systems are inevitably large and complex, mainly due to the interactions among their components as different user types or traffic categories. Rapid progress in technology has also made modeling more challenging. A commonly used modeling tool for performance and dependability analysis of such systems are continuous-time Markov chains (CTMC). One of the main advantages of using Markovian models is that it is general enough to capture the dominant factors of system uncertainty and, in the meantime, it is mathematically tractable [15].

Incorporating all the important factors into the models often results in a large state space of the underlying Markov chain, rendering the analysis problem computationally intractable. Therefore, it is very necessary to develop computationally efficient approximations techniques to compute the performance parameters of the system under study.

In general, it is not possible to divide a large system into its smallest irreducible subsystems completely separable from one another and treat each subsystem independently, we have to deal with situations in which the systems are only nearly decomposable, in the sense that, there are weak links among the irreducible subsystems, which dictate the occasional regime changes of the system. An effective way to treat such near decomposability is the time-scale

^{*} This work has been supported by the Spanish “Ministerio de Economía y Competitividad” through the project TIN2010-21378-C02-02.

separation. That is, we set up the systems as if there were two time-scales, fast vs. slow [15].

When the events of the components of a system occur at sufficiently separated time-scales, the simplest approximation, producing easily computable and accurate results, is the so called quasi-stationary (or, quasi-static) approximation (QSA) [2, 3, 5, 7, 14].

In [4] the authors introduce a new approximation method, also based in time-scale decomposition, called Generalized Quasi-Stationary Approximation (GQSA), that provides a way to trade off computational complexity and accuracy. They apply it to an integrated services system that serve short-lived non-real-time and long-lived real-time traffic. This new approximation aims to improve the accuracy at the price of higher computational cost. In [13] it was shown that, while the new GQSA improves the accuracy in some instances, it does not occur in all of them; and more importantly, it is difficult to predict in which cases accuracy can be enhanced by GQSA.

Taking into account that QSA provide accurate results only when the separation of time-scales is considerable (quasi-stationary regime), the purpose of this paper is to provide an accurate and high efficient approximation method for the performance analysis of telecommunications systems, to improve the accuracy of QSA when the system moves away from the quasi-stationary regime, i.e., enlarge the region of validity of QSA when the separation of time-scales is not enough.

We analyzed two types of telecommunication systems: a cognitive radio system (CRS) with two user types [8], and an integrated services system (ISS) with two traffic categories [11]. These systems present qualitative important differences at the model level, as is described in Sect. 2.

The approximation that we propose is based on the phase-type (PH) distribution which can be characterized as the time to absorption in a finite Markov chain. The two-dimensional Markov process of each system (CRS and ISS) can be represented with a number of transient states and an absorbing state as is detailed in Sect. 3.

We approach the problem with an iterative method. Taking QSA as an initial approximation to the solution vector, we proceed to modify this approximation in such a way that, as the number of iterations increases, the result becomes closer and closer to the exact solution. This is an important advantage with respect to the GQSA approach, because we can use the previous result as the new initial approximation. If indeed there is little change, we should expect to compute the new result in relatively few iterations. Also, an iterative process may be halted once a prespecified tolerance criterion has been satisfied, and this may be relatively lax [12].

The remainder of the paper is organized as follows. Section 2 describes the Markov models, and it details the characteristics of the systems to analyze. Section 3 presents the proposed approximation method based on time-scale separation and PH distribution. Section 4 details the numeric evaluation of the systems and the results of the performance metrics computed with the proposed approximation method at different time-scales to validate its accuracy. Finally, the conclusions are presented in Section 5.

2 Description of Models

In this section, we detail the characteristics of the systems in which we applied the proposed approximation method to evaluate its performance. We describe the two-dimensional CTMC models associated with them.

2.1 Cognitive Radio System

As in [8], we model the PU and SU traffic at the session (connection) level and ignore interactions at the packet level (scheduling, buffer management, etc.). We assume an ideal MAC layer for SUs, which allows a perfect sharing of the allocated channels among the active SUs (all active SUs get the same bandwidth portion), introduce zero delay and whose control mechanisms consume zero resources. In addition, we also assume that an active SU can sense the arrival of a PU in the same channel instantaneously and reliably. In this sense, the performance parameters obtained can be considered as an upper bound.

The Cognitive Radio System has C_1 *primary channels* (PCs) that can be shared by PUs and SUs, and C_2 *secondary channels* (SCs) only for SUs. Let $C = C_1 + C_2$ be the total number of channels in the system. Note that the SCs can be obtained from e.g. unlicensed bands, as proposed in [1]. This assumption is applicable to the *coexistence* deployment scenario for CRNs [10]. Alternatively, as it might be of commercial interest for the primary and secondary networks to *cooperate*, the secondary channels may be obtained based on an agreement with the primary network [10]. A SU in the PCs might be forced to vacate its channel if a PU claims it to initiate a new session. As SUs support *spectrum handover*, a vacated SU can continue with its ongoing communication if a free channel is available. Otherwise, it is *forced to terminate*.

For the sake of mathematical tractability, Poisson arrivals and exponentially distributed service times are assumed. The arrival rate for PU (SU) sessions is λ_1 (λ_2), their service rate is μ_1 (μ_2), and requests consume 1 (1) channel when accepted. We denote by (i, j) the system state, when there are i ongoing PU sessions and j SU sessions. The set of feasible states is $\mathcal{S} := \{(i, j) : 0 \leq i \leq C_1, 0 \leq i + j \leq C\}$ and the cardinality of \mathcal{S} is $|\mathcal{S}| = (\frac{C_1}{2} + C_2 + 1) \cdot (C_1 + 1)$. The state-transition diagram of the system is depicted in Fig. 1. Given the set of feasible states and their transitions in a CTMC, we can construct the global balance equations and the normalization equation. From these, we calculate the steady-state probabilities denoted as $\pi(i, j)$.

The system's performance parameters are determined as follows:

$$P_1 = \sum_{k=0}^{C_2} \pi(C_1, k) \quad , \quad P_2 = \sum_{k=C_2}^C \pi(C - k, k), \quad (1)$$

$$P_{ft} = \frac{\lambda_1(P_2 - \pi(C_1, C_2))}{\lambda_2(1 - P_2)}, \quad (2)$$

$$Th_2 = \sum_{j=1}^C \sum_{i=0}^Z j \mu_2 \cdot \pi(i, j), \quad (3)$$

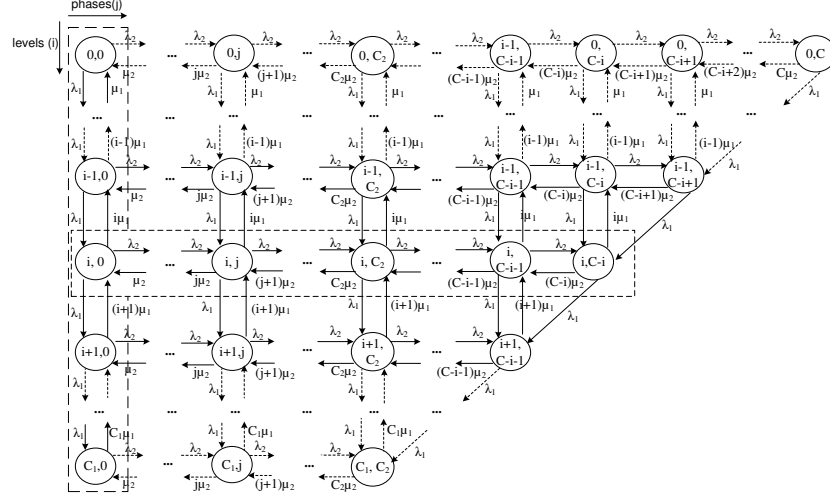


Fig. 1. State-transition diagram, Cognitive Radio System.

where P_1 is the PUs blocking probability, which clearly coincides with the one obtained in an Erlang-B loss model with C_1 servers; P_2 is the SUs blocking probability, i.e. the fraction of SU sessions rejected upon arrival as they find the system full; P_{ft} is the forced termination probability of the SUs, i.e. the rate of SU sessions forced to terminate divided by the rate of accepted SU sessions; Th_2 is the SUs throughput, i.e the rate of SU sessions successfully completed and $Z = \min(C_1, C - j)$.

2.2 Integrated Services System

We use the same model defined in [4] for an Integrated Services System that serve real-time (RT) and non-real-time (NRT) traffic. We consider a link whose limited resources (C Mbps in total) are shared amongst RT and NRT requests. The RT traffic is given strict priority over the NRT traffic. We initially assume that all RT calls are of the same class each requiring one channel of rate c b/s during its entire service duration to meet its required QoS. We denote by N_{rt} the maximum number of channels for RT calls. When an RT call arrives, it occupies 1 channel if available; otherwise, it is blocked. We set N_{rt} , such that $N_{rt} \cdot c$ is sufficiently smaller than C to avoid starvation of the NRT traffic. Let $n_{rt}(t)$ be the number of RT calls in the system at time t , $t \geq 0$, so $\{n_{rt}(t), t \geq 0\}$ is the RT process. NRT flows are served evenly by the leftover capacity from the RT traffic according to the processor sharing (PS) discipline. Let $n_{nrt}(t)$ be the number of NRT flows in the system at time t , $t \geq 0$. Then, $\{(n_{rt}(t), n_{nrt}(t)), t \geq 0\}$ is the joint RT and NRT process. The capacity available for all the NRT traffic at time t is given by $C_{nrt}(t) = C - n_{rt}(t) \cdot c$. The bit-rate of each admitted NRT flow at time t is $c_{nrt}(t) = C_{nrt}(t)/n_{nrt}(t)$, which is updated with RT or NRT admitted arrivals or departures. To satisfy the QoS of admitted NRT flows, the maximum

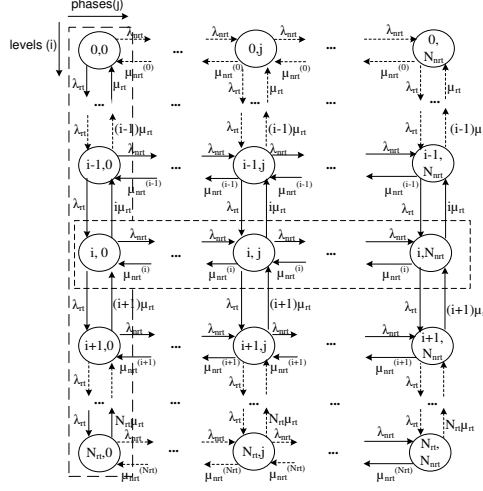


Fig. 2. State-transition diagram, Integrated Services System.

number of concurrent NRT flows is limited to N_{nrt} . Accordingly, an NRT flow arriving at time t is blocked if $n_{nrt}(t) = N_{nrt}$. We assume Poisson arrivals for RT and NRT requests with rates λ_{rt} and λ_{nrt} respectively. The service time of each admitted RT request is exponentially distributed, its service rate is μ_{rt} . On the other hand, as data sessions generate NRT traffic, their sojourn time will depend on the available resources. The size of the flows generated by the data sessions are exponentially distributed with mean L (bits).

We denote by (i, j) the system state, when there are i ongoing RT calls and j NRT flows. Let \mathcal{S} be the set of feasible states as $\mathcal{S} := \{(i, j) : 0 \leq i \leq N_{rt}, 0 \leq i+j \leq N_{rt}+N_{nrt}\}$ and the cardinality of \mathcal{S} is $|\mathcal{S}| = (N_{rt}+1)(N_{nrt}+1)$. The state-transition diagram of the system is depicted in Fig. 2. Given the set of feasible states and their transitions in a CTMC, we can construct the global balance equations and the normalization equation. From these, we calculate the steady-state probabilities denoted as $\pi(i, j)$. We must consider that the service rate of NRT flows varies according to the n_{rt} RT calls in the system as $\mu_{nrt}^{(i)} = \frac{C-i \cdot c}{L}$.

The system's performance parameters can be developed as follows:

$$P_{nrt} = \sum_{k=0}^{N_{rt}} \pi(k, N_{nrt}), \quad (4)$$

$$E[X_{nrt}] = \sum_{j=1}^{N_{nrt}} \sum_{i=0}^{N_{rt}} j \cdot \pi(i, j), \quad (5)$$

$$E[D_{nrt}] = \frac{E[X_{nrt}]}{\lambda_{nrt}(1 - P_{nrt})}, \quad (6)$$

where P_{nrt} is the blocking probability of NRT flows, $E[X_{nrt}]$ is the mean number of NRT flows in the system and $E[D_{nrt}]$ is the NRT flow average transfer delay.

3 Approximation Method

3.1 PH Distributions

Consider a CTMC on a finite state space $\mathcal{S} = \{0, 1, 2, \dots, m\}$. A PH distribution is the distribution of the time until absorption in a finite Markov chain of dimension $m + 1$, where 1 state is absorbing and the remaining m states are transients. A PH distribution is uniquely given by an m dimensional row vector α and an $m \times m$ matrix T . The vector α can be interpreted as the initial probability vector among the m transient states (with $\sum_{i=0}^m \alpha_i = 1$), while the matrix T can be interpreted as the infinitesimal generator matrix among the transient states in the continuous case. The random variable that is defined as the time to absorption, is said to have a (continuous) PH distribution [9].

The infinitesimal generator for the CTMC can be written in block-matrix form as $Q = \begin{bmatrix} t & T \\ 0 & 0 \end{bmatrix}$. Here, 0 is a $1 \times m$ vector of zeros. The vector $t = (t_{10}, t_{20}, \dots, t_{m0})'$ (the prime denoting transpose) where, for $i = 1, 2 \dots m$, $t_{i0} \geq 0$, with at least one of the t_{i0} s positive, is the absorption rate from state i . The $m \times m$ matrix $T = [t_{ij}]$ is such that, for $i = 1, 2 \dots m$, with $i \neq j$,

$$t_{ij} \geq 0 \quad \text{and} \quad t_{ii} = -\sum_{j=0, j \neq i}^m t_{ij}, \quad (7)$$

that is, $t = -Te$ where e is a column vector of ones of appropriate dimension. We call the pair $(\alpha; T)$ a representation for the PH distribution. The matrix T is referred to as a *PH generator*.

To ensure absorption in a finite time with probability one, we require that every non-absorbing state is transient. This statement is equivalent to T being invertible, therefore $-(T^{-1})_{ij}$ is the expected total time spent in phase j during the time until absorption, given that the initial phase is i [6, Theorem 2.4.3].

For extensive bibliographies and comprehensive theoretical treatment of PH distributions, see [9, Chap.2]. Also [6, Chap.2], is a very readable introduction to the topic.

3.2 Absorbing Markov Chains Method (AbMC)

In the quasi-stationary approximation it is assumed that, when the process enters a level, it takes an infinitely long time to leave this level. Thus, the probability that the process is in a phase j of the level i , given that the process is in the level i , is simply given by the stationary distribution of the level i considered as an isolated CTMC.

In our method, we assume that although the sojourn time in a level will be typically large (consistently with the large separation between time-scales) it is finite. We treat each level as an absorbing Markov chain in which the transitions out the level are transitions to an absorbing state. We now obtain the probability that the process is in a phase j of the level i , given that the process is in the level i , as the fraction of time that the process spends in the phase j before absorption. Note that if we knew the initial probabilities of the phases (i.e., upon entering the level) then the conditional probabilities obtained

by this method would be the exact ones. However, unless the original CTMC has some special structure (for instance, if each level can only be entered by exactly one of its phases), these initial probabilities cannot be obtained without having the stationary distribution of the whole CTMC.

We propose here to use the QSA to estimate the initial probabilities for the phases in each level. Then, we can obtain an estimation of the conditional probabilities for the phases as described above, i.e., as the fraction of time spent in each of them before absorption. Now, from the estimation of the conditional probabilities for the phases in each level, and the probability distribution for the levels, a new approximation for the stationary distribution of the CTMC is obtained. This way, we have obtained a refinement of the initial approximation of the stationary distribution given by the QSA. Moreover, the same process can be repeated iteratively to further improve the approximation.

Based on the basic properties of PH distributions introduced in Sect. 3.1, the iterative procedure described above can be implemented using the following equations:

$$\tilde{\pi}_i^{(k+1)} = \left[\alpha_i^{(k)} (-T_i^{-1}) \mathbf{e} \right]^{-1} \alpha_i^{(k)} (-T_i^{-1}) \quad (8)$$

$$\mathbf{v}_i^{(k)} = \pi_{i-1} \tilde{\pi}_{i-1}^{(k)} \mathbf{u}_{i-1} + \pi_{i+1} \tilde{\pi}_{i+1}^{(k)} \mathbf{d}_{i+1} \quad (9)$$

$$\alpha_i^{(k)} = \frac{1}{\mathbf{v}_i^{(k)} \mathbf{e}} \mathbf{v}_i^{(k)} \quad (10)$$

- Being k the iteration number, $\tilde{\pi}_i^{(k)}$ is the distribution of probabilities obtained by the PH distribution. The initial value is given by

$$\tilde{\pi}_i^{(0)} = \hat{\pi}(j|i), \quad (11)$$

where $\hat{\pi}(j|i)$ is the distribution of probabilities of:

- finding j ongoing sessions in an $M/M/(C-i)/(C-i)$ system with only SUs in CRS.
- finding j NRT flows in an $M/M/1/N-PS$ system with only NRT traffic in ISS.
- π is the distribution of probabilities of the elements with strict priority in the system (PUs in CRS, RT traffic in ISS). It is computed using simple recursions since their corresponding CTMC are one-dimensional birth-and-death processes.
- $\mathbf{v}_i^{(k)}$ is a vector that represent the input rates to each state of the level i . The initial value for the iterations is given by:

$$\mathbf{v}_i^{(0)} = \pi_{i-1} \tilde{\pi}_{i-1}^{(0)} \mathbf{u}_{i-1} + \pi_{i+1} \tilde{\pi}_{i+1}^{(0)} \mathbf{d}_{i+1}, \quad (12)$$

where \mathbf{u}_{i-1} is a vector with the transition rates from level $i-1$ to level i and \mathbf{d}_{i+1} is a vector with the transition rates from level $i+1$ to level i .

- $\alpha_i^{(k)}$ is the initial probability vector among the j transient states of the level i .

$$\alpha_i^{(0)} = \frac{1}{\mathbf{v}_i^{(0)} \mathbf{e}} \mathbf{v}_i^{(0)}. \quad (13)$$

Finally, the state probability distribution of the systems can be approximated as $\pi(i, j) \approx \tilde{\pi}(i, j) = \pi_i \cdot \tilde{\pi}_i^{(k)}(j)$. With this distribution of probabilities, we compute the approximate values of the performance parameters, using (1 – 3) for CRS, and (4 – 6) for ISS. Note that, if we were able to know the initial probability vector of the PH distribution (α), just an iteration would be enough to obtain the exact distribution of probabilities $\pi(i, j)$.

4 Numerical Evaluation and results

As a baseline for our study, we implemented the exact solution of the CTMC associated with each system, to calculate the exact values of their performance parameters. For the sake of comparison, we have used system sizes that allowed the exact solution to be computed within a reasonable time. In addition, we implemented both approximation methods: QSA and GQSA to validate and compare the performance of AbMC. We focus on evaluating the relative error (e_r) of each performance parameter. For instance, the relative error of the blocking probability for SUs in a CRS is computed as $e_r(P_2) = \frac{|P_2^E - P_2^{Ab}|}{P_2^E}$, where P_2^E is the exact value of SUs blocking probability and P_2^{Ab} is the approximate value of SUs blocking probability computed by AbMC method.

We evaluate the performance of the systems with different sizes (number of channels available for each type of user or flow) and different load conditions. In CRS, we analyze the blocking probability, forced termination probability and throughput of the SUs. We consider the following values for the number of primary channels: $C_1 = \{70, 80, 90, 100, 120, 140\}$, and for each of them, the following values for the number of secondary channels are considered: $C_2 = \{10, 20, 40\}$.

In ISS, we determine the blocking probability and average transfer delay of the NRT traffic. Keeping c and L constant, we consider the following values for the total link capacity of the system: $C = \{1.92, 7.68\}$ Mbps, that are similar to the ones used in [4].

To establish the load conditions, we set the service rates to 1, and then we adjust the arrival rates to obtain two load conditions: low (L) and high (H), which correspond to blocking probabilities of $1 \cdot 10^{-3}$ and $5 \cdot 10^{-2}$, respectively for each user type or traffic category, given as result four load configurations: *LL*, *LH*, *HL* and *HH*.

To assess the goodness of AbMC, we evaluate the evolution of the relative error as the separation between the time-scales varies. To accomplish this analysis, with the arrival rates adjusted to the specified load condition (*LL*, *LH*, *HL* or *HH*), we use an accelerating factor f , $10^{-5} \leq f \leq 10^5$, to accelerate or decelerate the arrival and departure events of the PUs, in the case of CRS or of the RT traffic, in the case of ISS, while keeping the offered traffic constant. For instance in a CRS, for each value of f , the PU arrival and service rates are obtained as $\lambda_1(f) = f \cdot \lambda_1$ and $\mu_1(f) = f \cdot \mu_1$. Note that the offered traffic $\frac{\lambda_1}{\mu_1} = \frac{\lambda_1(f)}{\mu_1(f)}$ is independent of f .

We varied the accelerating factor f to analyze the behavior of the approximation methods as a function of the separation of time-scales. To obtain the results with the iterative methods, we measured the time to compute *GQSA*

with the respective Radius R and iterated AbMC until that time, e.g., for $AbMC_1$ we iterated the method until the time measured to compute $GQSA_{R=1}$. In the same way for the other representations. The results are shown in Figs. 3 and 4; from them we can make the following observations:

- As expected, when the accelerating factor f decreases ($f \rightarrow 0$), the approximate values of all performance parameters evaluated tend to the QSA and the QSA results are very accurate.
- Note that increasing the radius in GQSA not always ensures a gradual reduction of the relative error [13]. As can be seen in Fig 3, $GQSA_{R=1}$ has a better accuracy in comparison with $GQSA_{R=2}$ and $GQSA_{R=3}$ for $f > 10^{-1}$. Clearly, the trade-off between the accuracy and computational cost will discourage the use of a radius larger than $R = 1$ in GQSA.
- AbMC have a predictable behavior: increasing the number of iterations to find the solution, ensures an improvement of the accuracy. Although they are not represented here due to space limitations, we have observed the same behavior in all performance parameters, for all load conditions and system sizes.

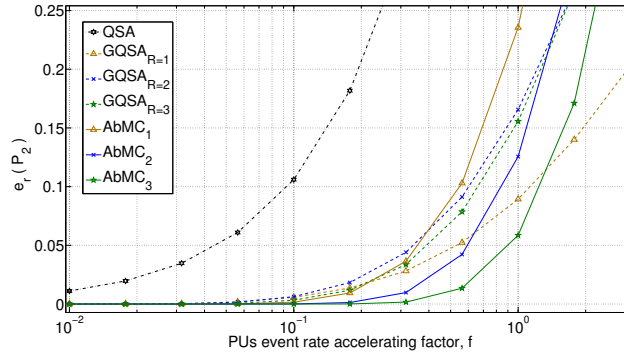


Fig. 3. CRF, Relative Error for the Blocking Probability SUs, LH load condition; $\lambda_1 = 49.239$, $\mu_1 = 1$, $C_1 = 70$; $\lambda_2 = 25.601$, $\mu_2 = 1$, $C_2 = 10$.

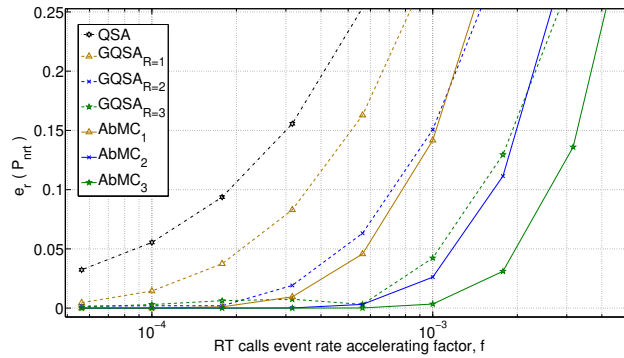


Fig. 4. ISS, Relative Error for the Blocking Probability NRT, HL load condition; $\lambda_{rt} = 38.557$, $\mu_{rt} = 1$, $N_{rt} = 44$; $\lambda_{nrt} = 1.243$, $N_{nrt} = 60$; $C = 7.68$ Mbps, $c = 64$ kbps, $L = 4$ Mb.

5 Conclusions

We provide a new approximate method called Absorbing Markov Chains (AbMC) for the performance analysis of telecommunication systems. We have validated AbMC in comparison with the quasi-stationary approximation and its recently proposed generalization. We have explored the evolution of the accuracy at different time-scales in two telecommunication systems: a cognitive radio system and an integrated services system which, at the model level, present qualitative important differences. The numerical results demonstrate that AbMC extends the outreach of the quasi-stationary approximation by trading-off computational effort in exchange of increased accuracy. AbMC has an iterative approach and as a part of future work, we will study the trade-off between accuracy and computational cost.

References

1. Al-Mahdi, H., Kalil, M.A., Liers, F., Mitschele-Thiel, A.: Increasing spectrum capacity for ad hoc networks using cognitive radios: an analytical model. *IEEE Communications Letters* 13(9), 676–678 (Oct 2009)
2. Alexiades, V., Solomon, A.D.: *Mathematical modeling of melting and freezing processes*. Taylor & Francis (1993)
3. Boxma, O.J., Gabor, A.F., Núñez-Queija, R., Tan, H.P.: Performance analysis of admission control for integrated services with minimum rate guarantees. In: *Proceedings of NGI'06*. pp. 41–47 (2006)
4. Huang, Y., Ko, K., Zukerman, M.: A generalized quasi-stationary approximation for analysis of an integrated service system. *IEEE Communications Letters* 16(11), 1884–1887 (Nov 2012)
5. Hubner, F., Tran-Gia, P.: Quasi-stationary analysis of a finite capacity asynchronous multiplexer with modulated deterministic input. *ITC-13*, Copenhagen (1991)
6. Latouche, G., Ramaswami, V.: *Introduction to Matrix Analytic Methods in Stochastic Modeling*. ASA-SIAM (1999)
7. Liu, S., Virtamo, J.: Performance analysis of wireless data systems with a finite population of mobile users. In: *Proceedings of the 19th International Teletraffic Congress ITC 19*. pp. 1295–1304 (2005)
8. Martinez-Bauset, J., Pla, V., Vidal, J., Guijarro, L.: Approximate analysis of cognitive radio systems using time-scale separation and its accuracy. *IEEE Communications Letters* 17(1), 35–38 (Jan 2013)
9. Neuts, M.: *Matrix-geometric Solutions in Stochastic Models: An Algorithmic Approach*. The Johns Hopkins University Press (1981)
10. Peha, J.: Sharing spectrum through spectrum policy reform and cognitive radio. *Proceedings of the IEEE* 97(4), 708–719 (2009)
11. Roberts, J.W.: Internet traffic, qos, and pricing. *Proceedings of the IEEE* 92(9), 1389–1399 (2004)
12. Stewart, W.J.: *Probability, Markov chains, queues, and simulation: the mathematical basis of performance modeling*. Princeton University Press (2009)
13. Tello-Oquendo, L., Pla, V., Martinez-Bauset, J.: Approximate analysis of wireless systems based on time-scale decomposition. In: *Wireless Days (WD), 2013 IFIP*. pp. 1–6 (Nov 2013)
14. Wong, E., Foh, C.: Analysis of cognitive radio spectrum access with finite user population. *IEEE Communications Letters* 13(5), 294–296 (2009)
15. Yin, G.G., Zhang, Q.: *Discrete-time Markov chains: two-time-scale methods and applications*, vol. 55. Springer (2006)