

Approximate Analysis of Wireless Systems Based on Time-Scale Decomposition

Luis Tello-Oquendo, Vicent Pla and Jorge Martinez-Bauset
Dept. of Communications, Universitat Politècnica de València (UPV)
ETSIT, Camí de Vera s/n, 46022 Valencia, Spain
Email: luiteloq@teleco.upv.es, {vpla,jmartinez}@dcom.upv.es

Abstract—Markov chains are a widely used modeling tool for wireless communication networks. The system size and the existence of different user types often make the analysis of the Markov chain computationally intractable. When the events of each user type occur at sufficiently separated time scales, the so-called quasi-stationary approximation (QSA) has proven to be accurate and highly efficient. Recently, a generalization of the quasi-stationary approximation (GQSA) has been introduced. The new approximation aims to improve the accuracy at the price of higher computational cost. In this paper, we carry out a comparative study of the accuracy and computational cost of both approximation methods QSA and GQSA. In particular, we explore the evolution of accuracy as the separation between time scales varies, and the trade-off between accuracy and computational cost. Our results indicate 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 the new method.

Index Terms—Wireless systems, cognitive radio systems, integrated services systems, traffic analysis, quasi-stationary approximation, time-scale decomposition.

I. INTRODUCTION

Continuous-time Markov chains (CTMC) are commonly used for modeling communication systems in order to study their performance. However, when the size of the systems is large, the computational cost to calculate their performance is greatly increased. Therefore, it is very necessary develop various approximations techniques to reduce the computational cost. One of these computationally efficient approximations, based in time-scale decomposition and often used is the quasi-stationary approximation (QSA) [1]–[3].

In [4] the authors introduce a new 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 (ISS) that serve short-lived non-real-time and long-lived real-time traffic.

Using the newly introduced method, we assess its behavior in a Cognitive Radio System (CRS), which at the model level present qualitative important differences with respect to the resources available for each type of user and the service rates in each state of the model as is described in Section II. The aim is to explore GQSA in both CRS and ISS.

The Cognitive Radio concept proposes to boost spectrum utilization by allowing cognitive users (secondary users, SU) to access the licensed wireless channel in an opportunistic manner so that interference to licensed users (primary users, PU) is kept to a minimum [5]. Since in CRS there are different types of users, the cardinality of state space increases

rapidly with the number of channels, implying the growth of the computational complexity related with the solution of the CTMC associated with the system. For this reason an approximation is typically required.

On the other hand, in an ISS, future generation broadband networks are expected to support a large variety of applications, typically grouped into two broad categories: real-time (RT) (e.g. voice and video) and non-real-time (NRT) (e.g. web-browsing, email and file-transfer) [6]. When the number of channels is large the computational complexity of solving the CTMC associated with the system becomes prohibitive. Therefore, computationally efficient approximations are required [4].

We approach the problem from the traffic perspective and develop two analytical models, one for CRS and another one for ISS, to evaluate the performance of these systems. We consider that the dynamics of the user types (in CRS) or traffic types (in ISS) operates at sufficiently separated time-scales, allowing to use approximation methods based on time-scale decomposition to simplify the computations.

The contribution of this paper is threefold. First, we evaluate the GQSA by applying it to a system (CRS) different from the one studied in [4] (ISS). Second, in both systems (CRS and ISS) we assess the behavior of the approximation when the separation of time-scales vary from the quasi-stationary regime to the fluid regime. Third, we analyze the trade-off between accuracy and computational cost of the approximation methods based on time-scale decomposition.

The rest of the paper is structured as follows. In Section II we describe the Markov models and detail the characteristics of the systems. Section III presents QSA and GQSA approximations based on time-scale decomposition to simplify the computations. Section IV detail the numeric evaluation and show the results of performance metrics of the systems to validate the accuracy and computational cost of approximations. Finally, the conclusions are presented in Section V.

II. DESCRIPTION OF MODELS AND EXACT ANALYSIS

In this section, we detail the characteristics of the systems and describe the CTMC models associated with them.

1) *Cognitive Radio System* : As in [7], 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

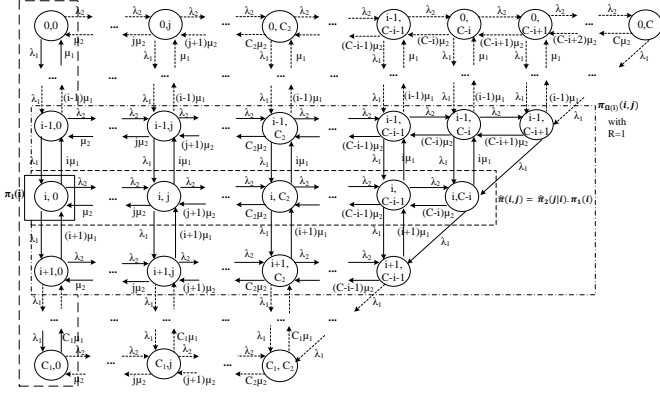


Figure 1. State-transition diagram, Cognitive Radio System.

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 [8]. This assumption is applicable to the *coexistence* deployment scenario for CRNs [9]. 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 [9].

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 $|\mathcal{S}| = \left(\frac{C_1}{2} + C_2 + 1\right) \cdot (C_1 + 1)$. 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 performance parameters are determined as follows,

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

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

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

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 $\alpha = \min(C_1, C - j)$.

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. Denote 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 $N_{rt}c$ 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 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 follow:

$$\mu_{nrt}^{(i)} = \frac{C - i \cdot c}{L}. \quad (4)$$

The system performance parameters can be developed as follows:

$$P_{nrt} = \sum_{k=0}^{N_{nrt}} \pi(k, N_{nrt}), \quad (5)$$

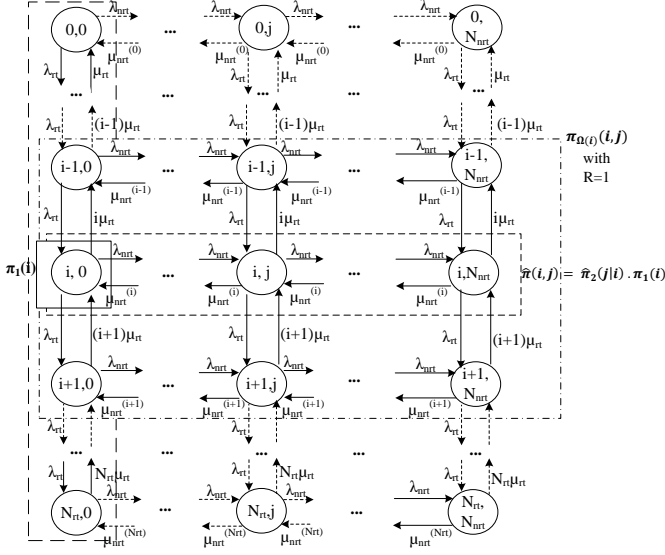


Figure 2. State-transition diagram, Integrated Services System.

$$E[X_{nrt}] = \sum_{j=1}^{N_{nrt}} \sum_{i=0}^{N_{rt}} j \cdot \pi(i, j), \quad (6)$$

$$E[D_{nrt}] = \frac{E[X_{nrt}]}{\lambda_{nrt}(1 - P_{b_{nrt}})}, \quad (7)$$

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

III. APPROXIMATIONS

In terms of the modeling approach, while Markovian models have been developed for the exact analysis of CRS or ISS, they can be numerically cumbersome [3].

This study can be problematic as the higher dimensionality may render the exact analysis computationally intractable. However, when the dynamics of different dimensions (as type of users—PU or SU—in CRS or types or traffic—RT or NRT—in ISS) of the system in analysis operates at sufficiently separated time-scales, one can resort to highly efficient approximations based on time-scale decomposition, which can greatly simplify the computations [7]. In this section, we describe a quasi-stationary approximation and a generalized quasi-stationary approximation to estimate performance parameters in both CRS and ISS.

A. Quasi-stationary Approximation

The simplest approximation, producing easily computable results is the so called quasi-stationary (or, quasi-static) approximation [10].

For instance, in CRS, the approximation can be applied by decoupling the two types of users to analyze their performances. We represent each type of CR user in one dimension of our model. In Fig. 1 we can see PUs process in the y axis and SUs process in the x axis. Since PUs utilize channel resources regardless of the existence of the SUs, the

performance analysis for PUs can be evaluated independently in an exact manner. Then the performance of SU can be approximated.

The approximation procedure consists of two stages. The first stage yields the distribution of steady-state probabilities for the PU, i.e. $\{\pi_1(i) : i = 0, \dots, C_1\}$, where $\pi_1(i)$ is the probability of finding i ongoing sessions in an $M/M/C_1/C_1$ system with only PUs.

The second stage is to calculate the j SU session distribution of steady-state probabilities by conditioning on the state distribution for the PU: $\hat{\pi}_2(j|i)$. Also, $\hat{\pi}_2(j|i)$ is the stationary probability of finding j ongoing sessions in an $M/M/C - i/C - i$ system with only SUs.

Both $\pi_1(i)$ and $\hat{\pi}_2(j|i)$ can be determined independently using simple recursions, since their corresponding CTMC are one-dimensional birth-and-death processes.

Finally, the state probability distribution of CRS can be approximated as:

$$\pi(i, j) \approx \hat{\pi}(i, j) = \pi_1(i) \cdot \hat{\pi}_2(j|i). \quad (8)$$

Sufficient accuracy can be obtained by using the QS approximation. At the QS regime, the PUs blocking probability is $P_1^{qs} = \pi_1(C_1)$; the SUs blocking probability is P_2^{qs} , the SUs throughput is Th_2^{qs} and the SUs forced termination probability is P_{ft}^{qs} and can be determined employing distribution (8) as follows [7]:

$$P_2^{qs} = \sum_{k=c_2}^C \hat{\pi}(C - k, k), \quad (9)$$

$$P_{ft}^{qs} = \frac{\lambda_1(P_2 - \hat{\pi}(C_1, C_2))}{\lambda_2(1 - P_2)}, \quad (10)$$

$$Th_2^{qs} = \sum_{j=1}^C \sum_{i=0}^{\alpha} j \mu_2 \cdot \hat{\pi}(i, j). \quad (11)$$

In the same way for ISS system, we can approximate the distribution of steady-state probabilities of the system using (8) in two stages: First stage, calculate $\pi_1(i)$ as the probability of finding i ongoing RT calls in an $M/M/N_{rt}/N_{rt}$ system with only RT traffic. The second stage is to calculate $\hat{\pi}_2(j|i)$ as the stationary probability of finding j NRT flows in an $M/M/1/N - PS$ system with only NRT traffic, i.e., using (4) for each i , the steady-state probabilities are given by:

$$\hat{\pi}_2(j|i) = \pi_n = a_2^n \pi_0, \quad (12)$$

where

$$a_2 \equiv a_2^{(i)} = \frac{\lambda_{nrt}}{\mu_{nrt}^{(i)}}; \quad \pi_0 = \frac{1 - a_2}{1 - a_2^{N_{nrt}+1}}.$$

At the QS regime, in ISS, the NRT flows blocking probability is P_{nrt}^{qs} , NRT flow average transfer delay is given by $E^{qs}[D_{nrt}]$ and can be determined using distribution (8) as follows:

$$P_{nrt}^{qs} = \sum_{k=0}^{N_{rt}} \hat{\pi}(k, N_{nrt}), \quad (13)$$

$$E^{qs} [X_{nrt}] = \sum_{j=1}^{N_{nrt}} \sum_{i=0}^{N_{rt}} j \cdot \hat{\pi}(i, j), \quad (14)$$

$$E^{qs} [D_{nrt}] = \frac{E^{qs} [X_{nrt}]}{\lambda_{nrt}(1 - P_{nrt}^{qs})}. \quad (15)$$

B. Generalized Quasi-stationary Approximation

To make this paper self-contained, in this section we describe the GQSA proposed in [4]. For the sake of conciseness, we describe it only for the CRS.

As in QSA, the performance of SU can be approximated using GQSA but considering a set of i neighboring states of the PU process (y axis in Fig. 1) rather than only the state i . In this way, the joint PU and SU process is more likely to reach statistical equilibrium since the time duration that the PU events continuously remains in a set of neighboring i states is generally longer than the time duration that it continuously remains in one i state.

In GQSA a new parameter called radius denoted by R , $R \in \{0, 1, 2, \dots, \lceil \frac{C_1}{2} \rceil\}$ is introduced, to indicate the size of the set of i neighboring states that we are considering in the model to analyze. The number of rows (state i and adjacent states to i) in the state-transition diagram, Fig. 1, with a defined radius to compute the approximation is $2R + 1$. We define $\Omega(i)$ as the set of states composed of the row of i state and its $2R$ closest rows of states as follow:

$$\Omega(i) = \begin{cases} \{(i', j) \in \mathcal{S} : 0 \leq i' \leq 2R\}, & 0 \leq i < R, \\ \{(i', j) \in \mathcal{S} : i - R \leq i' \leq i + R\}, & R \leq i \leq C_1 - R, \\ \{(i', j) \in \mathcal{S} : C_1 - 2R \leq i' \leq C_1\}, & C_1 - R < i \leq C_1. \end{cases}$$

As the set of states of PU process can comprise a single PU state, QSA is considered a special case of GQSA when $R=0$.

In Fig. 1, we show the elements, probabilities and stationary distributions involved in GQSA for a CRS:

- $\pi_1(i)$ is the probability of finding i ongoing sessions in an $M/M/C_1/C_1$ system with only PUs.
- $\hat{\pi}_2(j|i)$ is the stationary probability of finding j ongoing sessions in an $M/M/C - i/C - i$ system with only SUs.
- $\hat{\pi}(i, j) = \pi_1(i) \cdot \hat{\pi}_2(j|i)$ is the approximated steady-state probability of the CRS using QSA.
- $\pi_{\Omega(i)}(i, j)$ is the stationary distribution of the new set of states defined by R for analyze the system.

Finally, the approximated (i, j) steady-state probability using GQSA is defined as follow:

$$\pi(i, j) \approx \bar{\pi}(i, j) = \pi_1(i) \cdot \frac{\pi_{\Omega(i)}(i, j)}{\sum_j \pi_{\Omega(i)}(i, j)}. \quad (16)$$

The approximate values of performance parameters for CRS are calculated using (1), (2) and (3) with the steady-state probabilities defined in (16). Also in ISS, the performance parameters are calculated using (5), (6) and (7) with the steady-state probabilities defined in (16).

IV. NUMERICAL EVALUATION AND RESULTS

In this section, we study the behavior of GQSA when the separation of time-scales vary from QS regime to fluid regime. We provide and discuss numerical results of the accuracy of the approximation in the calculation of performance parameters of CRS and ISS.

As a baseline for our study, we implemented the exact solution of the CTMC system (see Fig. 1 for CRS, and Fig. 2 for ISS) to calculate the exact values of their performance parameters. Then we implemented the GQSA, i.e. we apply the distribution defined in (16) to calculate the performance parameters of each system. We focus on evaluating the relative error (e_r) of each parameter. For instance, the relative error of the blocking probability for SUs in a CRS, $e_r(P_2)$, is computed as

$$e_r(P_2) = \frac{|P_2^E - P_2^G|}{P_2^E} \quad (17)$$

where P_2^E is the exact value of SUs blocking probability and P_2^G is the approximate value of SUs blocking probability calculated using (16).

To evaluate the goodness of GQSA, in terms of computational complexity, execution time and accuracy, we studied the systems with different sizes (number of channels) and different load conditions.

To set the load conditions, we proceed as follows: setting the service rates to 1, we adjusted arrivals rates to obtain two load conditions, low (L) and high (H), which correspond to blocking probabilities $1 \cdot 10^{-3}$ and $5 \cdot 10^{-2}$, respectively.

In total we consider four load configurations for each system:

- | | |
|----|--|
| LL | low load condition for PUs (RT traffic), and low load condition for SUs (NRT traffic). |
| LH | low load condition for PUs (RT traffic), and high load condition for SUs (NRT traffic). |
| HL | high load condition for PUs (RT traffic), and Low load condition for SUs (NRT traffic). |
| HH | high load condition for PUs (RT traffic), and high load condition for SUs (NRT traffic). |

With the arrival rates adjusted to the specified load (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, in order to analyze the approximation methods from the QS regime to the fluid regime, for each value of f , the PU arrival and service rates are obtained as $\lambda_1(f) = f \cdot \lambda_1$ and $\mu_1(f) = f \cdot \mu_1$.

In CRS we analyze blocking probability, forced termination probability and throughput, considering the following values for the number of primary channels: $C_1 = \{30, 40, 50, 60, 70, 80, 90\}$ and for each of them the following values for the number of secondary channels are considering: $C_2 = \{C_1, (C_1/2), (C_1/5), (C_1/10)\}$.

In ISS we analyze blocking probability and average transfer delay for NRT traffic; keeping c and L constant with values for total link capacity of $C = \{1.92, 3.84, 7.68\}$ Mbps.

We varied the accelerating factor f to analyze the behavior of the approximations as a function of the separation of time-

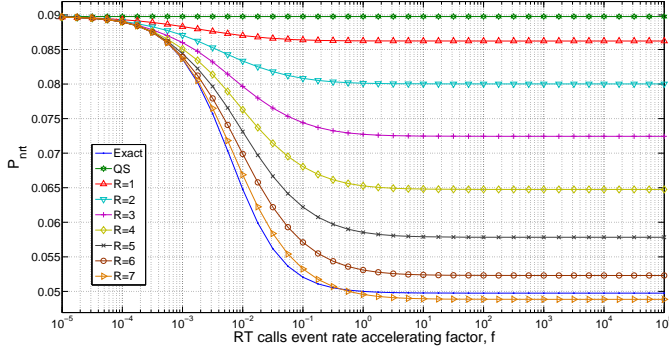


Figure 3. ISS, NRT Blocking Probability, LH load condition; $\lambda_{rt} = 10.812$, $\mu_{rt} = 1$, $N_{rt} = 22$; $\lambda_{nrt} = 0.317$, $N_{nrt} = 30$; $C = 1.92$ Mbps, $c = 64$ kbps, $L = 4$ Mb.

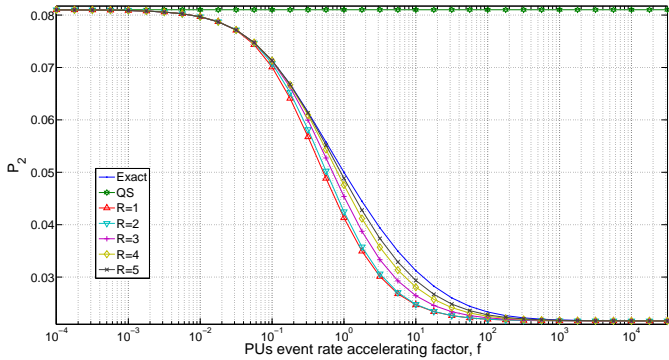


Figure 4. CRS, SUs Blocking Probability, HH load condition; $\lambda_1 = 34.596$, $\mu_1 = 1$, $C_1 = 40$; $\lambda_2 = 5.455$, $\mu_2 = 1$, $C_2 = 4$.

scales. The results are shown in Figs. 3–6; from them we can make the following observations:

- 1) The values of the performance parameter obtained by the GQSA attain the exact value when R is increased to $C_1/2$ in CRS or $N_{rt}/2$ for ISS.
- 2) As expected, when the accelerating factor f decreases the curves tend to the QS regime for all performance parameters.
- 3) In ISS, for all values of accelerating factor f and for all performance parameters, when R increases from 0 to $N_{rt}/2$ the approximations approach gradually the exact value. Figure 3 shows this behavior for P_{nrt} with load condition LH.

In contrast, as can be observed in Fig. 4, in CRS for all values of f , the curves corresponding to $R = 1, \dots, C_1/2$ are not between the curve for QSA ($R = 0$) and the curve for the exact values. In other words, the QSA ($R = 0$) overestimate the exact value of P_2 whereas the GQSA ($R > 0$) underestimate it. However, this behavior is not maintained for all different configurations (see Fig. 5 with $R = 1$).

- 4) Figure 5 and 6 show that to achieve a high accuracy in some system configurations, the radius needs only to be increased slightly ($R = 1$ in Fig. 5, and $R = 3$ in Fig. 6).

Note that increasing the radius not always ensures a

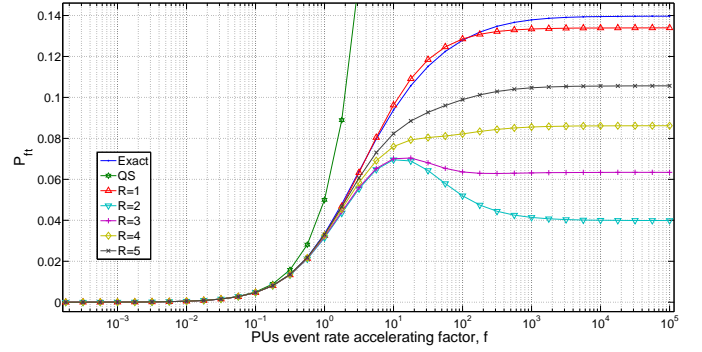


Figure 5. CRS, SUs Forced Termination Probability, HH load condition; $\lambda_1 = 34.596$, $\mu_1 = 1$, $C_1 = 40$; $\lambda_2 = 42.182$, $\mu_2 = 1$, $C_2 = 40$.

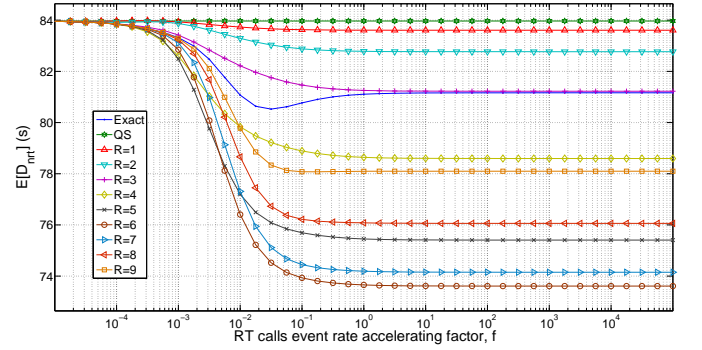


Figure 6. ISS, NRT Flow Average Transfer Delay, HH load condition; $\lambda_{rt} = 17.132$, $\mu_{rt} = 1$, $N_{rt} = 22$; $\lambda_{nrt} = 0.227$, $N_{nrt} = 30$; $C = 1.92$ Mbps, $c = 64$ kbps, $L = 4$ Mb.

gradual and monotonous convergence to the exact value. As can be seen in Fig 6, up to $R = 3$ increasing the radius improved the accuracy of the GQSA. However, increasing further the radius from 4 to 6 the accuracy of the GQSA deteriorates. Finally, as the radius increases beyond 6, the accuracy of the GQSA gradually improves again. Clearly, the trade-off between the accuracy and computational cost will discourage the use of a radius larger than $R = 3$.

- 5) The behavior of GQSA is not monotonous in terms of accuracy in both ISS and CRS. GQSA accuracy starts being good in QS regime; as the accelerating factor moves away from the QS regime (see Fig. 4 for $10^{-1} \leq f \leq 10^0$, and Fig. 3 for $10^{-4} \leq f \leq 10^{-2}$), we observe that the curves of GQSA using small radius begin to distance the curve of the exact values, i.e. GQSA ceases to be accurate. Surprisingly, for certain values of R , as f keeps on growing and we approach the fluid regime ($f > 10^4$), the accuracy of the GQSA improves and the values obtained with it (for any value of $R > 0$) almost overlap the exact ones. This behavior is clearly observed for the blocking probability in CRS with any system size and load condition (see Fig. 4). Also we note this behavior in ISS, for the NRT flow average transfer delay with $R = 3$ and the specifications detailed in Fig. 6.

The behavior in observations 4 and 5, might be due to the

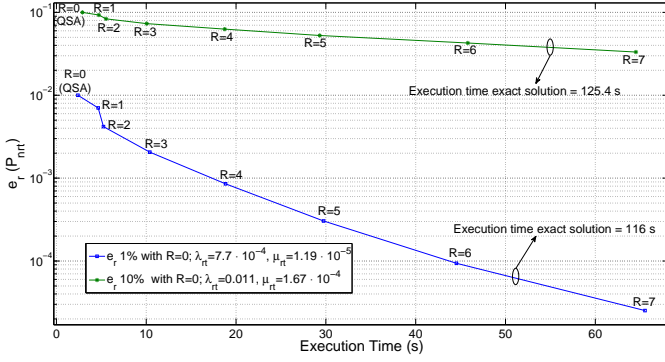


Figure 7. ISS, e_r in Blocking Probability, LL load condition; $C = 7.68$ Mbps, $c = 64$ kbps, $L = 4$ Mb; $N_{rt} = 88$, $N_{nrt} = 120$

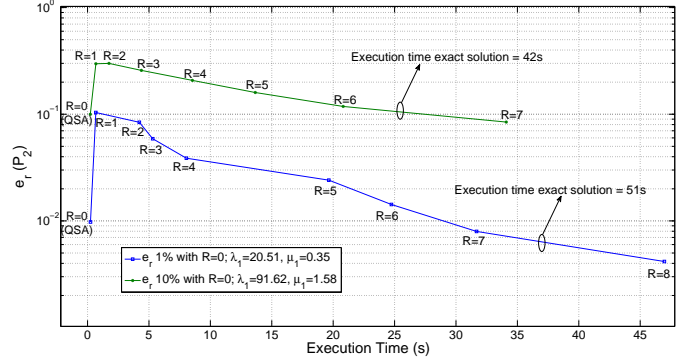


Figure 8. CRS, e_r in Blocking Probability, LL load condition; $C_1 = 80$, $C_2 = 80$; $\lambda_2 = 71.62$, $\mu_2 = 1$

way in which the subset of states $\Omega(i)$ is chosen. Note the asymmetry with respect to row i , when $0 \leq i < R$ and $C_1 - R < i \leq C_1$. This aspect requires further investigation.

In order to measure the trade-off between accuracy and computational cost, in Figs. 7 and 8 we represent the relative error e_r against the execution time for different values of the radius R . Note that $R = 0$ corresponds to the QSA case. In this curves, we have set the value of f so that in QS regime, the e_r is normalized to 1% and 10% for each performance parameter. Figures 7 and 8 show only the results for those values of R in which the execution time that do not exceed the time necessary to obtain the exact solution.

On the behavior of GQSA we note the following:

In ISS, the relative error decreases when the radius of GQSA increases. Although they are not represented here due to the space limitations, we have observed the same behavior in all performance parameters, for all load conditions and system sizes.

A rather different behavior is observed, for instance, in CRS with LL and HL load conditions. Analyzing blocking and forced termination probabilities with LL load condition and large system sizes, is better to use QSA than to use GQSA with some values for the radius.

In Fig. 7 for the curve with initial $e_r = 1\%$, we observe that from $R = 0$ to $R = 1$, e_r decreases in 30% while the execution time increases by 94%. When the initial approximation is poorer ($e_r = 10\%$ with $R=0$), accuracy improves more slowly as we increase R .

Fig. 8 shows that e_r increases abruptly from $R = 0$ to $R = 1$ and then decline gradually. The same behavior is observed no matter what the initial e_r is (1% or 10%).

V. CONCLUSIONS

In this paper we have studied two approximation methods based on time-scale decomposition for the analysis of cognitive radio systems and integrated services systems which, at the model level, present qualitative important differences. We have modeled them as continuous time Markov chains. We assessed the behavior of the approximations when the separation of time-scales vary from the QS regime to the fluid regime. We have measured the trade-off between accuracy and computational cost. During the study, we illustrate how GQSA display a behavior in terms of accuracy, not previously

encountered in the analysis of other systems. The numerical results demonstrate, contrary to what one may expect, that the relative error of the performance parameters using GQSA does not always decrease when R is increased, i.e. increasing the radius not always improves the accuracy, in some cases it deteriorates; therefore, the computational cost necessary to gain in accuracy can be very high in comparison to use QSA in order to evaluate the performance of the systems. Knowing when the relative error decreases, and when it does not, depends in a complex way on several factors. Some of them (type of system, load conditions, system sizes) were discussed in the paper while other require further investigation, due to it is difficult to predict in which cases accuracy can be enhanced by the new method. An unexpected finding is that in some specific cases, GQSA is a good approximation not only for QS regime but also in the fluid regime, where the difference in the separation of time-scales is negligible.

REFERENCES

- [1] F. Hubner and P. Tran-Gia, "Quasi-stationary analysis of a finite capacity asynchronous multiplexer with modulated deterministic input," *ITC-13, Copenhagen*, 1991.
- [2] S. Liu and J. Virtamo, "Performance analysis of wireless data systems with a finite population of mobile users," in *Proceedings of the 19th International Teletraffic Congress ITC 19*, 2005, pp. 1295–1304.
- [3] O. J. Boxma, A. F. Gabor, R. Núñez-Queija, and H.-P. Tan, "Performance analysis of admission control for integrated services with minimum rate guarantees," in *Proceedings of NGI'06*, 2006, pp. 41–47.
- [4] Y. Huang, K. Ko, and M. Zukerman, "A generalized quasi-stationary approximation for analysis of an integrated service system," *IEEE Communications Letters*, vol. 16, no. 11, pp. 1884–1887, Nov. 2012.
- [5] I. F. Akyildiz, W. Lee, M. Vuran, and S. Mohanty, "A survey on spectrum management in cognitive radio networks," *IEEE Communications Magazine*, vol. 46, no. 4, pp. 40–48, 2008.
- [6] J. W. Roberts, "Internet traffic, qos, and pricing," *Proceedings of the IEEE*, vol. 92, no. 9, pp. 1389–1399, 2004.
- [7] J. Martinez-Bauset, V. Pla, J. Vidal, and L. Guijarro, "Approximate analysis of cognitive radio systems using time-scale separation and its accuracy," *IEEE Communications Letters*, vol. 17, no. 1, pp. 35–38, Jan. 2013.
- [8] H. Al-Mahdi, M. A. Kalil, F. Liers, and A. Mitschele-Thiel, "Increasing spectrum capacity for ad hoc networks using cognitive radios: an analytical model," *IEEE Communications Letters*, vol. 13, no. 9, pp. 676–678, Oct. 2009.
- [9] J. Peha, "Sharing spectrum through spectrum policy reform and cognitive radio," *Proceedings of the IEEE*, vol. 97, no. 4, pp. 708–719, 2009.
- [10] V. Alexiades and A. D. Solomon, *Mathematical modeling of melting and freezing processes*. Taylor & Francis, 1993.