Analysis of a Cellular Network with User Redials and Automatic Handover Retrials

Jose Manuel Gimenez-Guzman, M[a] Jose Domenech-Benlloch,
Vicent Pla, Vicente Casares-Giner, and Jorge Martinez-Bauset

Dept. of Communications, Universitat Politècnica de València, UPV
ETSIT Camí de Vera s/n, 46022, València, Spain
{jogiguz,mdoben}@doctor.upv.es,
{vpla,vcasares,jmartinez}@dcom.upv.es

Abstract. In cellular networks, repeated attempts occur as result of user behavior but also as automatic retries of blocked requests. Both phenomena play an important role in the system performance and therefore should not be ignored in its analysis. On the other hand, an exact Markovian model analysis of such systems has proven to be infeasible and resorting to approximate techniques is mandatory. We propose an approximate methodology which substantially improves the accuracy of existing methods while keeping computation time in a reasonable value. A numerical evaluation of the model is carried out to investigate the impact on performance of the parameters related to the retry phenomena. As a result, some useful guidelines for setting up the automatic retries are provided. Finally, we also show how our model can be used to obtain a tight performance approximation in the case where reattempts have a deterministic nature.

1 Introduction

In the POTS the phenomenon of repeated attempts due to user behavior, and its analysis, have been studied, at least, since the early 70's [1]. In modern cellular networks, network driven retries of blocked handover requests (retrials) occur on top of the reattempts triggered by the user behavior during a fresh session setup (redials) [2,3].

There are important differences between redials and automatic retrials. Blocked handovers will be automatically retried until a reattempt succeeds or the user moves outside the handover area. In the former case the session will continue without the user noticing any disruption, while in the latter the session will be abruptly terminated. In contrast, persistence of redials depends on the user patience and an eventual abandonment results in session setup failure, which is less annoying than the abrupt termination of an ongoing session. Moreover, automatic retrials are rather deterministic in nature [2] while redials are affected by the randomness of human behavior. Thus, from a modeling perspective, both types of reattempts need to be considered separately giving rise to two separate orbits of retrying customers.

Y. Koucheryavy, J. Harju, and A. Sayenko (Eds.): NEW2AN 2007, LNCS 4712, pp. 210–222, 2007.

Even if only a single orbit —instead of two— is considered the resulting model is of the type of the multiserver retrial queue, for which it is known that an analytical solution is not available and only numerical approximations can be obtained (see [3,4,5] and references therein). In particular Marsan et al. [3] consider a system fairly similar to the one considered here, and propose an approximate technique for its analysis. In [6] a generalization of the approximate method in [3] was proposed for a system with only a single retrial orbit, showing a substantial improvement in the accuracy at the expense of only a marginal increase of the computational time. In this paper we extend the approximation technique of [6] to a system with two different retrial orbits (redials and retrials). The proposed method is employed to perform a numerical analysis of the system focusing on how redials and retrials impact on the system performance. As a result some guidelines for setting up the automatic retries are provided. Additionally, we propose an accurate approximation method to analyze the performance of a system with deterministic retrials, (i.e. the maximum number of retrials or the time between consecutive reattempts take fixed values). To the best of our knowledge previous performance analyses of cellular systems with retrials [3,4,6] assume that the maximum number of retrials is geometrically distributed and the time between consecutive reattempts is exponentially distributed.

The rest of the paper is structured as follows. Section 2 describes the system under study, while Section 3 discusses the system model and the analysis methodology. In Section 4 the numerical analysis of the impact of retrials/redials is carried out. Final remarks and a summary of results are provided in Section 5.

2 System Description

We consider a cellular mobile network, with a fixed channel allocation scheme and where each cell is served by a different base station, being C the number of resources in the cell. The physical meaning of a unit of resource is dependent on the specific technological implementation of the radio interface. Without loss of generality, we consider that each user occupies one resource unit. As shown in Fig. 1 there are two arrival streams: the first one represents new sessions and the second one handovers from adjacent cells. Both arrivals are considered Poisson processes with rates λ_n and λ_h respectively, being $\lambda = \lambda_n + \lambda_h$. For determining the value of λ_h we consider that the incoming handover stream is equal to the outgoing handover stream, due to the system homogeneity [7]. For the sake of mathematical tractability, the session duration and the cell residence time are exponentially distributed with rates μ_s and μ_r, respectively. Hence, the channel holding time is also exponentially distributed with rate $\mu = \mu_r + \mu_s$ and the mean number of handovers per session when the number of resources is infinite is $N_H = \mu_r/\mu_s$.

The FGC (Fractional Guard Channel) policy is characterized by only one parameter t ($0 \leq t \leq C$). New sessions are accepted with probability 1 when there are less than $L = \lfloor t \rfloor$ resources being used and with probability $f = t - L$, when there are exactly L resources in use. If there are more than L busy

Table 1. Transition rates

Transition	Condition		Rate
$(k,m,s) \to (k+1,m,s)$	$0 \le k \le L-1$	$m < Q_n$ & $s < Q_h$	λ
		$m < Q_n$ & $s = Q_h$	$\lambda + \beta_h$
		$m = Q_n$ & $s < Q_h$	$\lambda + \beta_n$
		$m = Q_n$ & $s = Q_h$	$\lambda + \beta_n + \beta_h$
	$k = L$	$m < Q_n$ & $s < Q_h$	$\lambda_h + f\lambda_n$
		$m < Q_n$ & $s = Q_h$	$\lambda_h + \beta_h + f\lambda_n$
		$m = Q_n$ & $s < Q_h$	$\lambda_h + f(\beta_n + \lambda_n)$
		$m = Q_n$ & $s = Q_h$	$\lambda_h + \beta_h + f(\beta_n + \lambda_n)$
	$L < k \le C$	$m < Q_n$ & $s < Q_h$	λ_h
		$m < Q_n$ & $s = Q_h$	$\lambda_h + \beta_h$
		$m = Q_n$ & $s < Q_h$	λ_h
		$m = Q_n$ & $s = Q_h$	$\lambda_h + \beta_h$
$(k,m,s) \to (k+1,m,s-1)$	$0 \le k \le C-1$	$1 \le s \le Q_h - 1$	$s\mu_{ret}$
		$s = Q_h$	α_h
$(k,m,s) \to (k,m,s-1)$	$k = C$	$1 \le s \le Q_h - 1$	$s\mu_{ret}P_{ih}$
		$s = Q_h$	$\alpha_h P_{ih}$
$(k,m,s) \to (k+1,m-1,s)$	$0 \le k \le L-1$	$1 \le m \le Q_n - 1$	$m\mu_{red}$
		$m = Q_n$	α_n
	$k = L$	$1 \le m \le Q_n - 1$	$m\mu_{red}f$
		$m = Q_n$	$\alpha_n f$
$(k,m,s) \to (k,m-1,s)$	$k = L$	$1 \le m \le Q_n - 1$	$m\mu_{red}(1-f)P_{in}$
		$m = Q_n$	$\alpha_n(1-f)P_{in}$
	$L < k \le C$	$1 \le m \le Q_n - 1$	$m\mu_{red}P_{in}$
		$m = Q_n$	$\alpha_n P_{in}$
$(k,m,s) \to (k-1,m,s)$	$1 \le k \le C$		$k\mu$
$(k,m,s) \to (k,m,s+1)$	$k = C$		$\lambda_h(1 - P^1_{ih})$
$(k,m,s) \to (k,m+1,s)$	$k = L$		$\lambda_n(1 - P^1_{in})(1-f)$
	$L < k \le C$		$\lambda_n(1 - P^1_{in})$

Note: $\alpha_n = M_n\mu_{red}(1-p_n)$, $\beta_n = M_n\mu_{red}p_n$
$\alpha_h = M_h\mu_{ret}(1-p_h)$, $\beta_h = M_h\mu_{ret}p_h$.

resources, new sessions are no longer accepted. Handovers are accepted while the system is not completely occupied.

When an incoming new session is blocked, according to Fig. 1, it joins the redial orbit with probability $(1 - P^1_{in})$ or leaves the system with probability P^1_{in}. If a redial is not successful, the session returns to the redial orbit with probability $(1-P_{in})$, redialing after an exponentially distributed time with rate μ_{red}. Redials are able to access to the same resources as the new sessions.

Similarly, P^1_{ih}, P_{ih} and μ_{ret} are the analogous parameters for the automatic retrials. Making $P^1_{ih} = 0$, at least one retrial will be performed. In that case, if the system were so loaded that the probability of a successful retrial could be considered negligible, the time elapsed since the first handover attempt until the system finally gives up and the session is dropped will be a sum of X iid exponential rv of mean μ^{-1}_{ret}. In our model the discrete rv X follows a geometric

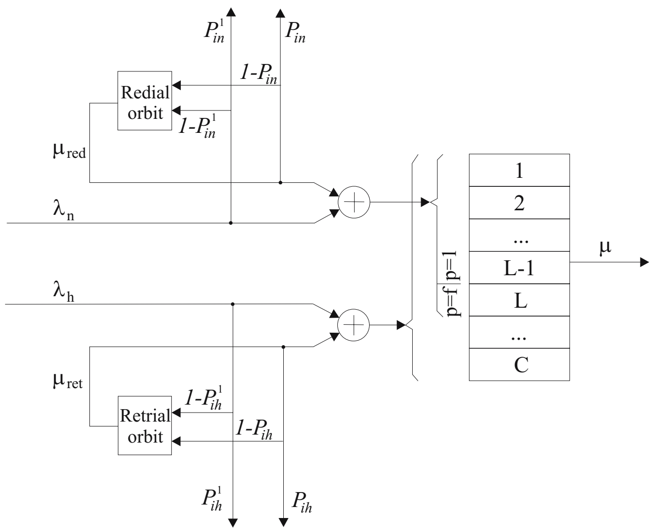


Fig. 1. System model

distribution with mean $1/P_{ih}$, hence the total time from the first attempt until abandonment is described by an exponential rv of rate $\mu_r' = \mu_{ret} P_{ih}$. In the light of the above discussion, our model represents a situation in which the blocked handover requests will keep retrying while the user remains within the handover area, being the sojourn time modeled as an exponential rv of rate μ_r'. This assumption has been shown to have a low impact on the performance measures of interest [8].

3 System Model and Performance Analysis

The model considered can be represented as a tridimensional (k, m, s) Continuous Time Markov Chain (CTMC), being one dimension (k) the number of sessions being served, the second dimension (m) the number of sessions in the redial orbit and the third dimension (s) the number of sessions in the retrial orbit. The main mathematical features of this queueing model are the fact of having two infinite dimensions (the state space of the model is $\{0, \ldots, C\} \times \mathbb{Z}_+ \times \mathbb{Z}_+$) and the space-heterogeneity along them is produced by the retrial and redial rates, which respectively depend on the number of customers on the retrial and the redial orbits.

It is known that the classical theory (see, e.g., [9]) is developed for random walks on the semi-strip $\{0, \ldots, C\} \times \mathbb{Z}_+$ with infinitesimal transitions subject to conditions of space-homogeneity. When the space-homogeneity condition do not hold the problem of calculating the equilibrium distribution has not been addressed beyond approximate methods [10], [11]. Indeed, if we focus on the simpler case of multiserver retrial queues (with only one retrial orbit) it can emphasize the absence of closed form solutions for the main performance characteristics when $C > 2$ [12].

As it is clear that in our case it is necessary to resort to approximate models and numerical methods of solution, in [6] we developed a generalization of the approximation method proposed in [3]. The new methodology is applied to both retrial and redial orbits, reducing the state space to a finite set by aggregating all states beyond a given occupancy of the orbits: Q_n (Q_h) defines the occupancy from which the states in the redial (retrial) orbit are aggregated. By increasing the values of Q_n and/or Q_h the considered state space in the approximation is enlarged and the accuracy of the solution improves at the expense of a higher computational cost. Due to that aggregation two new parameters for each orbit are introduced. The parameter M_n denotes the mean number of users in the redial orbit conditioned to those states where there are at least Q_n users in the orbit, i.e. $M_n = E(m|m \geq Q_n)$. The probability that after a successful redial the number of users in the redial orbit does not drop below Q_n is represented by p_n. For the retrial orbit the parameters M_h and p_h are defined analogously.

As a result of the aggregation the state space of the approximate model is $S = \{(k, m, s) : 0 \leq k \leq C; 0 \leq m \leq Q_n; 0 \leq s \leq Q_h\}$ where states of the form $(\cdot, Q_n, \cdot)$ represent the situation where at least Q_n users are in the redial orbit. Likewise the states of the form $(\cdot, \cdot, Q_h)$ represent the situation where at least Q_h users are in the retrial orbit. The transition rates for the approximate model are shown in Table 1.

In order to compute the steady-state probabilities of the system ($\pi(k, m, s)$) the actual values of the parameters M_n, p_n, M_h and p_h should be known. By balancing the probability fluxes across the vertical and horizontal cuts of the transition diagram, and equating the rate of *blocked first attempts* to the sum of the rates of successful and abandoning reattempts, the parameters above are expressed in terms of the steady-state probabilities

$$p_h = \frac{\sum_{m=0}^{Q_n} \pi(C, m, Q_h)}{\sum_{m=0}^{Q_n} \left[\pi(C, m, Q_h) + \pi(C, m, Q_h - 1)\right]} \tag{1}$$

$$M_h = \frac{\lambda_h (1 - P_{ih}^1) \left(\sum_{m=0}^{Q_n} \left[(\pi(C, m, Q_h) + \pi(C, m, Q_h - 1)\right] \right)}{\mu_{ret} \left(\sum_{k=0}^{C-1} \sum_{m=0}^{Q_n} \pi(k, m, Q_h) + P_{ih} \sum_{m=0}^{Q_n} \pi(C, m, Q_h) \right)} \tag{2}$$

$$p_n = \frac{\zeta_1}{\zeta_2} \quad ; \quad M_n = \frac{\lambda_n (1 - P_{in}^1) \zeta_2}{\mu_{red} \zeta_3} \tag{3}$$

where

$$\zeta_1 = \sum_{k=L+1}^{C} \sum_{s=0}^{Q_h} \pi(k, Q_n, s) + (1 - f) \sum_{s=0}^{Q_h} \pi(L, Q_n, s)$$

$$\zeta_2 = \sum_{k=L+1}^{C}\sum_{s=0}^{Q_h}\big[\pi(k,Q_n-1,s)+\pi(k,Q_n,s)\big]+(1-f)\sum_{s=0}^{Q_h}\big[\pi(L,Q_n-1,s)+\pi(L,Q_n,s)\big]$$

$$\zeta_3 = \sum_{k=0}^{L-1}\sum_{s=0}^{Q_h}\pi(k,Q_n,s)+f\sum_{s=0}^{Q_h}\pi(L,Q_n,s)+(1-f)P_{in}\sum_{s=0}^{Q_h}\pi(L,Q_n,s)+P_{in}\sum_{k=L+1}^{C}\sum_{s=0}^{Q_h}\pi(k,Q_n,s)$$

The global balance equations, the normalization equation and Eqs. (1)–(3) form a system of simultaneous non-linear equations, which can be solved using — for instance— the iterative procedure sketched next: set $p_n = p_h = 0$, $M_n = Q_n$ and $M_h = Q_h$ and compute the steady-state probabilities using the algorithm defined in [13], now compute M_n, p_n, M_h, p_h using Eqs. (1)–(3) and start again. In all of our numerical experiments we repeated the iterative procedure until the relative difference between two consecutive iterations was less than 10^{-4} for all four parameters.

The most common performance parameters used in cellular systems are the blocking probabilities of both new sessions (P_b^n) and handovers (P_b^h). Additionally, it is also used the probability of having a handover failure, denoted as forced termination probability (P_{ft}), which is given in terms of the non-service probability (P_{ns}^h), i.e. the probability that a handover request and all its subsequent reattempts are blocked. Moreover, we define the mean number of redials (retrials) per user as u_n (u_h) and the mean number of users in the redial (retrial) orbit as N_{red} (N_{ret}).

$$P_b^n = \sum_{k=L+1}^{C}\sum_{m=0}^{Q_n}\sum_{s=0}^{Q_h}\pi(k,m,s) + (1-f)\sum_{m=0}^{Q_n}\sum_{s=0}^{Q_h}\pi(L,m,s)$$

$$u_n = \frac{\mu_{red}}{\lambda_n}(1-P_{in})\Big[\sum_{k=L+1}^{C}\sum_{m=0}^{Q_n-1}\sum_{s=0}^{Q_h}m\pi(k,m,s)+M_n\zeta_1+(1-f)\sum_{m=0}^{Q_n-1}\sum_{s=0}^{Q_h}m\pi(L,m,s)\Big]+$$
$$+(1-P_{in}^1)\Big[(1-f)\sum_{m=0}^{Q_n}\sum_{s=0}^{Q_h}\pi(L,m,s)+\sum_{k=L+1}^{C}\sum_{m=0}^{Q_n}\sum_{s=0}^{Q_h}\pi(k,m,s)\Big]$$

$$N_{red} = \sum_{k=0}^{C}\sum_{m=0}^{Q_n-1}\sum_{s=0}^{Q_h}m\pi(k,m,s) + M_n\sum_{k=0}^{C}\sum_{s=0}^{Q_h}\pi(k,Q_n,s)$$

$$P_b^h = \sum_{m=0}^{Q_n}\sum_{s=0}^{Q_h}\pi(C,m,s) \quad ; \quad P_{ft} = \frac{N_H P_{ns}^h}{1+N_H P_{ns}^h}$$

$$P_{ns}^h = \frac{\mu_{ret}}{\lambda_h}P_{ih}\Big[\sum_{m=0}^{Q_n}\sum_{s=0}^{Q_h-1}s\pi(C,m,s)+M_h\sum_{m=0}^{Q_n}\pi(C,m,Q_h)\Big]+P_{ih}^1\sum_{m=0}^{Q_n}\sum_{s=0}^{Q_h}\pi(C,m,s)$$

$$u_h = \frac{\mu_{ret}}{\lambda_h}(1-P_{ih})\Big[\sum_{m=0}^{Q_n}\sum_{s=0}^{Q_h-1}s\pi(C,m,s)+M_h\sum_{m=0}^{Q_n}\pi(C,m,Q_h)\Big]+(1-P_{ih}^1)\sum_{m=0}^{Q_n}\sum_{s=0}^{Q_h}\pi(C,m,s)$$

$$N_{ret} = \sum_{k=0}^{C}\sum_{m=0}^{Q_n}\sum_{s=0}^{Q_h-1}s\pi(k,m,s) + M_h\sum_{k=0}^{C}\sum_{m=0}^{Q_n}\pi(k,m,Q_h)$$

4 Results and Discussion

In this section a number of numerical examples are presented with the purpose of illustrating the capabilities and versatility of our model and the analysis methodology. The numerical analysis is also aimed at assessing the impact on performance of varying the values and/or distributions of the system parameters.

For the numerical experiments a basic configuration is used and then the different parameters are varied, normally a single variation is introduced in each experiment. Thus, unless otherwise indicated, the value of the parameters will be those of the basic configuration: $C = 32$, $N_H = \mu_r/\mu_s = 2$, $\mu = \mu_r + \mu_s = 1$, $t = 31$, $P_{ih} = P_{in} = 0.2$, $\mu_{red} = 20$, $P^1_{ih} = P^1_{in} = 0$, $\mu'_r = 10\mu_r$ and then $\mu_{ret} = 100/3$.

4.1 Approximate Methodology

Here we evaluate the accuracy of the approximate analysis as a function of Q_h and Q_n. For a given performance indicator I and given values of Q_h and Q_n the relative error introduced by the approximate model is estimated by $\epsilon_I(Q_n, Q_h) = \left|\frac{I(Q_n+1,Q_h+1)}{I(Q_n,Q_h)} - 1\right|$. In Fig. 2 the relative error estimate is plotted as a function of $Q_h = Q_n$ taking as performance indicators N_{red} and N_{ret}. As it might be expected, except for a very short transient phase, the value of $\epsilon_I(Q_n, Q_h)$ decreases when the values of Q_h and Q_n increase, and also, that a higher load (given by λ_n) results in a poorer accuracy. The curves also show that a good accuracy can be achieved with relatively low values of Q_h and Q_n, having been observed in all the numerical examples we have carried out. Moreover, in all the numerical results shown hereafter the values of Q_h and Q_n have been chosen so that $\epsilon_{N_{red}}(Q_n, Q_h) < 10^{-4}$ and $\epsilon_{N_{ret}}(Q_n, Q_h) < 10^{-4}$.

4.2 Redimensioning with Redials

Due to the human behavior, users normally redial if a previous attempt has been blocked. Network operators, however, do not consider redials as such simply because they are not able to distinguish between first attempts and redials, therefore every incoming session is regarded as a first attempt. Without that distinction, a resource over-provisioning can occur because for each user requesting a session whose first attempt is blocked several new session requests are actually accounted (one per attempt).

In order to evaluate the magnitude of over-provisioning the following experiment was carried out. We start from a basic situation in which the QoS objectives ($P^n_b \le 0.05$ and $P_{ft} \le 0.005$) are fulfilled and consider several values of load growth. For each value of the load increment, the amount of resources (C) is redimensioned in order to meet the QoS objectives. The redimensioning process is done using the complete model and a simplified model where redials are considered as fresh new calls, i.e. $\lambda'_n = \lambda_n + \mu_{red}N_{red}$. Figure 3 shows a sample of results from the redimensioning process which reveal that ignoring the existence of redials can produce a significant over-provisioning.

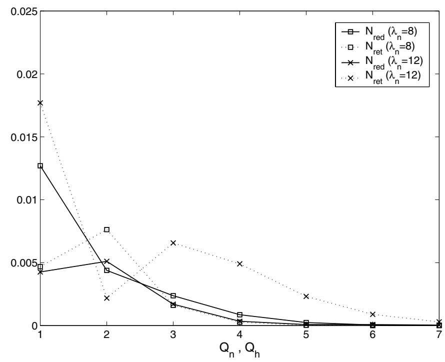
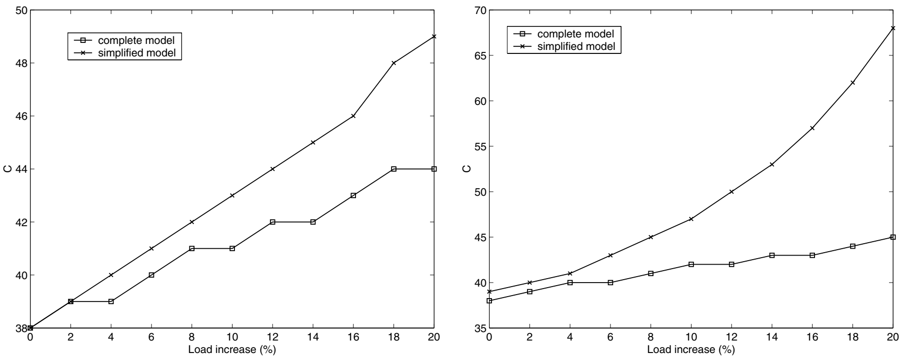


Fig. 2. Accuracy of the approximate methodology



(a) $P_{in} = 0.1$.

(b) $P_{in} = 0$.

Fig. 3. Resource redimensioning with and without considering redials

4.3 Impact of Automatic Retrial Configuration

If the network operator enables the automatic retrial option the blocked handover attempts will be automatically retried while the user remains within the handoff area. We consider a fixed mean sojourn time in the handover area ($\mu_r' = 20/3$) and study the impact of varying the retrial rate (μ_{ret}). Note that for varying μ_{ret} while μ_r' is kept constant the value of P_{ih} is varied accordingly using their relationship, $\mu_r' = \mu_{ret} P_{ih}$.

Figure 4 shows that a higher value of μ_{ret} results in a lower forced termination probability but also a higher mean number of retrials per session. While the former is a positive effect the later is not that much as it entails an increased signaling load. In order to gain a further insight into the existing tradeoff between P_{ft} and u_h we define the overall cost function $C_T = \beta \lambda_n P_{ft} + \lambda_h u_h$. The choice of the value for β may depend on many factors and a suitable value can vary widely from one situation to another, thus we have used a wide range of values,

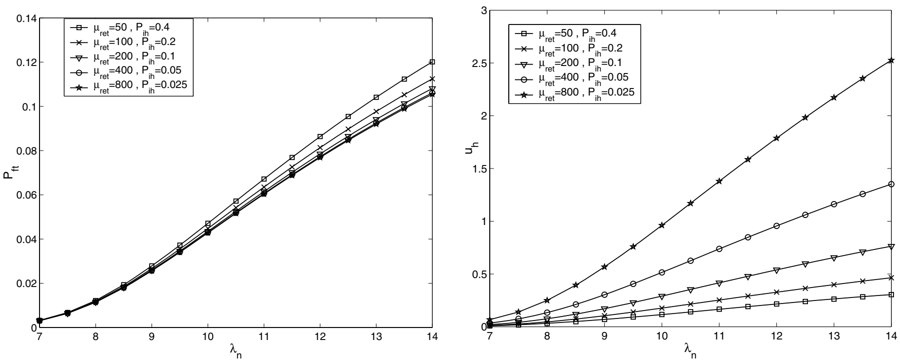
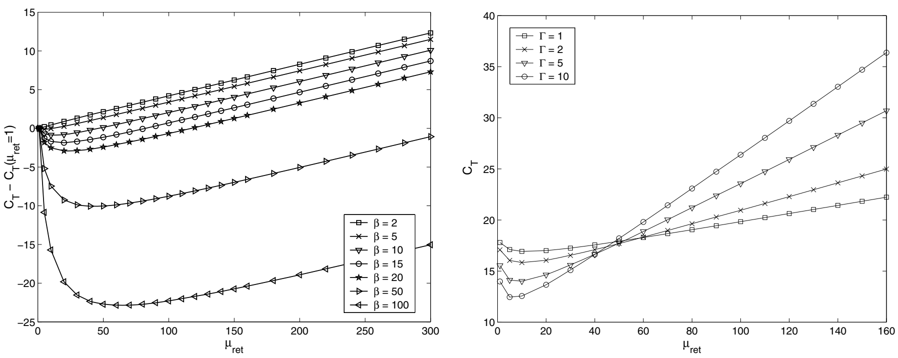


Fig. 4. Performance parameters for different retrial configurations



(a) Increment $(C_T(\mu_{ret}) - C_T(1))$ when μ_{ret} varies, $\Gamma = 1$.

(b) Absolute value, $\beta = 10$.

Fig. 5. Cost function, $\lambda_n = 12$

$\beta = \{2, 5, 10, 15, 20, 50, 100\}$. We also explored the effect of varying the mean sojourn time in the handover area $1/\mu'_r$ (actually a normalized parameter with respect to $1/(C\mu)$ has been used $\Gamma = C\mu/\mu'_r$).

The shape of cost curves in Fig. 5(a) shows the existence of an optimal configuration point. Both the relevance of the optimal configuration point and the value of the retrial rate at which it is attained increase when the weight factor β is increased. Moreover, Fig. 5(b) shows that the optimal value of μ_{ret} is rather insensitive to the mean value of the sojourn time in the handover area.

4.4 Distribution of the Maximum Number and Time Between Reattemps

In real systems (e.g. GSM) the time between retrials as well as the maximum number of retrials per request take a deterministic value instead of an stochastic

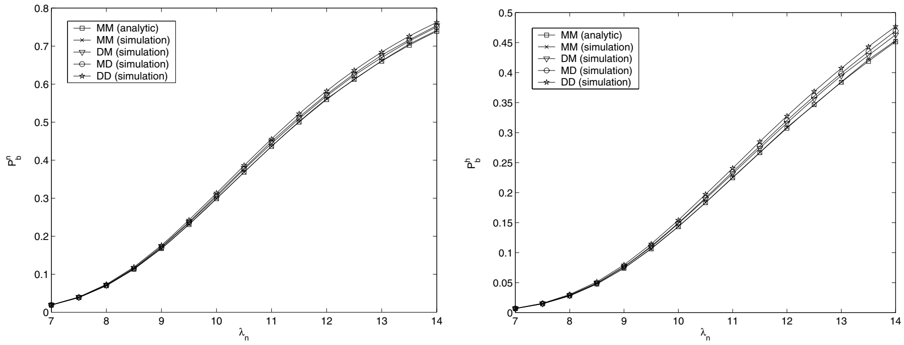


Fig. 6. Distribution of the time between reattempts: impact on P_b^n and P_b^h. Legend: XY, X (Y) $\equiv$ distribution for redials (retrials); $M \equiv$ exponential, $D \equiv$ deterministic.

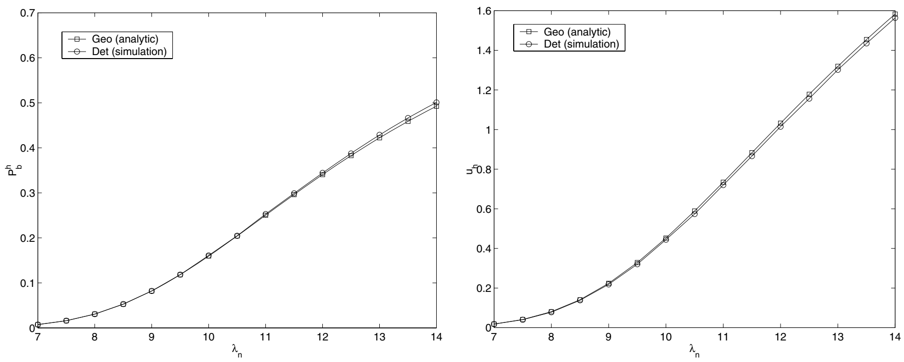


Fig. 7. Analytical approximation of a deterministic maximum number of retrials; $d = 5$, $P_{ih}^1 = 0$

one [2]. In our model, however, in order to keep the mathematical analysis tractable, we used an exponentially distributed time between retrials and a geometric distribution for the maximum number of reattempts. Here we validate these two assumptions with the help of a discrete event simulation model. In order to simplify the simulations we set $\lambda_h = 2\lambda_n$ instead of computing the equilibrium value of λ_h.

Time distribution between redials/retrials: We analyze the values of P_b^n and P_b^h when the distribution of the time between redials, retrials, or both are switched from exponential to deterministic, keeping constant its mean value. From the results in Fig. 6, and others not shown here due to the lack of space, we conclude that assuming an exponential distribution for the time between redials and/or retrials has a negligible impact in all the performance parameters of interest.

Distribution of the maximum number of reattempts: We compare a geometric distribution (after each unsuccessful attempt the user decides to abandon the system with probability P_i) with a deterministic distribution (the users leaves the system after d unsuccessful attempts). For making these two options comparable the mean number of reattempts must be the same in both cases. Note it is not the same as both distributions having the same mean as the distributions refer to the maximum number of reattempts and not to the actual number of reattempts.

While the following discussion deals only with retrials it can be easily extended to redials as well. Let q denote the blocking probability for retrials (note that in general $q \neq P_b^h$), the average number of retrials is

$$u_h^{Geo} = \sum_{n \geq 1} P_b^h (1 - P_{ih}^1)((1 - P_{ih})q)^{n-1}(1 - (1 - P_{ih})q) = \frac{(1 - P_{ih}^1)P_b^h}{1 - (1 - P_{ih})q} \quad (4)$$

$$u_h^D = (1 - q)P_b^h[1 + 2q + 3q^2 + \ldots + (d - 1)q^{d-2}] + dP_b^h q^{d-1} = P_b^h \frac{1 - q^d}{1 - q} \quad (5)$$

for the geometric and deterministic case, respectively. If we assume that both q and P_b^h take approximately the same value in both cases, by equating the right hand side of (4) and (5) we obtain

$$P_{ih} = \frac{1 - q}{q(1 - q^d)}(q^d - P_{ih}^1) \quad (6)$$

For a given value of d, by using the expressions for P_b^h and $u_h$and Eqs. (4) and (6), the value of P_{ih} that yields $u_h^{Geo} = u_h^D$ can be iteratively computed. The results shown in Fig. 7, and similar ones not shown here due to the lack of space, demonstrate that using the adjusting procedure described above, our model can provide an excellent approximation for the performance analysis of a system in which the maximum number of retrials is a fixed number.

5 Conclusions

In cellular networks, repeated attempts occur due to user redials when their session establishments are blocked and also due to automatic retries when a handover fails. The impact of both phenomena plays an important role in the system performance and it should not be ignored. However, the main feature of the Markovian model describing such a complex system is the space-heterogeneity along two infinite dimensions. Due to this fact, we develop an approximate methodology that aggregates users in the retrial/redial orbit beyond a given occupancy. Our proposal achieves a higher accuracy than other techniques while keeping computation time negligible from a human point of view.

A numerical evaluation of the system has been performed in order to evaluate the impact of the reattempt phenomena in the system performance. We have studied the effect of automatic retrials for handovers while the user remains into the handover area, giving some guidelines to the network operators in order to configure this behaviour optimally. Finally, we have shown how our model can be used to obtain a tight performance approximation when the time between reattempts and maximum number of reattempts are deterministic. Results of this approximate method are compared against those obtained by simulation, concluding that the proposed method is very accurate.

Acknowledgements

This work was supported by the Spanish Government (30% PGE) and the European Commission (70% FEDER) through projects TSI2005-07520-C03-03 and TEC2004-06437-C05-01 and by *Cátedra Telefónica de Internet y Banda Ancha* (e-BA) from the Universidad Politécnica de Valencia. Besides M. Jose Domenech-Benlloch was supported by the Spanish Ministry of Education and Science under contract AP-2004-3332.

References

1. Jonin, G., Sedol, J.: Telephone systems with repeated calls. In: Proceedings of the 6th International Teletraffic Congress ITC'6, pp. 435.1–435.5 (1970)
2. Onur, E., Deliç, H., Ersoy, C., Çaglayan, M.U.: Measurement-based replanning of cell capacities in GSM networks. Computer Networks 39, 749–767 (2002)
3. Marsan, M.A., De Carolis, G.M., Leonardi, E., Lo Cigno, R., Meo, M.: Efficient estimation of call blocking probabilities in cellular mobile telephony networks with customer retrials. IEEE Journal on Selected Areas in Communications 19(2), 332–346 (2001)
4. Tran-Gia, P., Mandjes, M.: Modeling of customer retrial phenomenon. IEEE Journal on Selected Areas in Communications 15(8), 1406–1414 (1997)
5. Chakravarthy, S.R., Krishnamoorthy, A., Joshua, V.: Analysis of a multi-server retrial queue with search of customers from the orbit. Performance Evaluation 63(8), 776–798 (2006)
6. Doménech-Benlloch, M.J., Giménez-Guzmán, J.M., Martínez-Bauset, J., Casares-Giner, V.: Efficient and accurate methodology for solving multiserver retrial systems. IEE Electronic Letters 41(17), 967–969 (2005)
7. Marsan, M.A., Carolis, G.D., Leonardi, E., Cigno, R.L., Meo, M.: How many cells should be considered to accurately predict the performance of cellular networks? In: Proceedings European Wireless (1999)
8. Pla, V., Casares, V.: Effect of the handoff area sojourn time distribution on the performance of cellular networks. In: Proceedings of IEEE MWCN, pp. 401–405 (2002)
9. Neuts, M.: Matrix-geometric Solutions in Stochastic Models: An Algorithmic Approach. The Johns Hopkins University Press, Baltimore (1981)
10. L. Bright, P.G.T.: Calculating the equilibrium distribution of level dependent quasi-birth-and-death processes. Communications in Statistics-Stochastic Models 11(3), 497–525 (1995)

11. Latouche, G., Ramaswami, V.: Introduction to Matrix Analytic Methods in Stochastic Modeling. ASA-SIAM (1999)
12. Artalejo, J.R., Pozo, M.: Numerical calculation of the stationary distribution of the main multiserver retrial queue. Annals of Operations Research 116(1-4), 41–56 (2002)
13. Servi, L.D.: Algorithmic solutions to two-dimensional birth-death processes with application to capacity planning. Telecommunication Systems 21(2-4), 205–212 (2002)