

Inter-arrival Time Distribution for Channel Arrivals in Cellular Telephony

José Ignacio Sánchez, Francisco Barceló and Javier Jordán

Abstract-- *In this paper different probability density functions are fitted to the inter-arrival time in a channel of a Cellular Mobile Telephony system. The approach is entirely experimental: the data set to be fitted has been obtained on an actual system in operation. The Kolmogorov-Smirnov (K-S) goodness-of-fit test is used in order to establish a ranking of the best fitting probability density functions. From this study it can be concluded that the arrivals to a channel in a cell are according to a smooth process.*

Index terms-- *Traffic modelling, cellular telephony, mobile telephony, voice modelling.*

I. INTRODUCTION

The design and performance evaluation of Mobile Telephony Networks is usually carried out by using some basic concepts of Queuing Theory as well as assuming certain statistical distributions for the inter-arrival time and the channel occupancy time processes. A poor knowledge of these distributions contributes to an inefficient design of the network resources because the engineer must be conservative to cope with the possible error margin. On the other hand, the accurate knowledge of the distributions allows an accurate design of the resources. This shows up in a better use of the radio resources used in this kind of networks and makes it possible the development of new services offered through cellular networks such as data services.

Some cellular networks make use of the excess of capacity in their cells in order to transmit data or short data messages during the periods in which the channels are not being used for voice traffic (voice idle periods). As public cellular networks must achieve a very low Grade of Service (GoS) measured as blocking probability, the channel load should be kept low and idle periods are significant. To effectively design this kind of networks a better understanding of the duration and frequency of these idle periods is needed. Because of this, voice traffic statistics in these networks need to be better understood.

In the past, it has been widely used the negative exponential

distribution to model the call and channel holding time. However different studies showed that the call duration is better fitted by other distributions in fix telephony [1], public cellular mobile [2] and private mobile radio (PMR) [3]. The channel holding time has been also showed to fit lognormal distributions better than the exponential [4, 5, 6].

The arrival processes have usually been considered to be Poissonian. But the particularities found in Cellular Mobile Telephony Systems make this assumption very suspicious. This task has been entered upon in other studies from a theoretical point of view: in [7] it is concluded that the arrival traffic is Poissonian while in [8] the same traffic appears to be smooth traffic. In [6] a field study of the inter-arrival time to the channel is presented showing a good fit with the exponential distribution which agrees with the Poissonian assumption. In [9] an experimental approach to the inter-arrival time in PAMR systems proves that the arrival traffic is smooth, but the measured systems are not cellular and have some very specific particularities.

In our study we deal with the statistical modelling of the inter-arrival time process in Cellular Mobile Telephony systems. An experimental approach is proposed, based on the analysis of real traffic data. In Section II a review of some previous results is presented along with the theoretical model of the system. Section III describes the main details of the empirical procedure followed in our study and the statistical procedure selected to analyse the collected data. In Section IV the fitting distributions are presented along with the reasons to select some distributions while disregarding others. The fitting of the selected distributions is presented in Section V for different channel loads. In Section VI other statistical results over the inter-arrival and idle time processes are presented. The obtained results and conclusions of Section VII tend to confirm that the arrival traffic to a channel in cellular telephony is smooth in agreement with some results reached by analytical models like the one developed in [8].

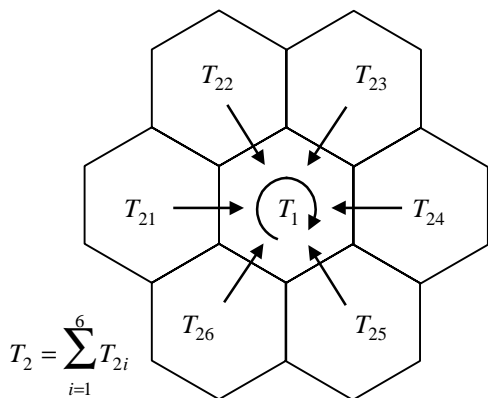
II. THEORETICAL MODEL AND PREVIOUS RESEARCH

A. A simple theoretical model

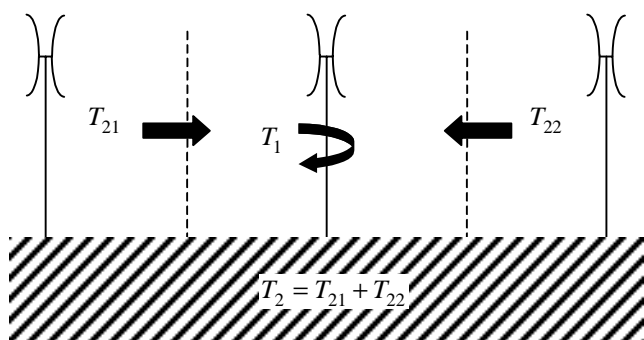
In every cell of a Cellular Mobile Telephony network the offered traffic is the result of two traffic streams of different sources:

E.T.S. de Ingeniería de Telecomunicación de Barcelona (UPC) Mod. C3 Campus Nord, c/ Gran Capitán, s/n 08034 Barcelona. E-mail: barcelo@mat.upc.es

- On one hand we have “fresh” traffic due to the calls which are originated inside the limits of the cell (T_1). T_1 is entirely caused by new calls. The infinite population hypothesis could be accepted for the T_1 stream if the number of terminals which are able to start a new call inside the tacked cell is large (much larger than the number of available channels). This happens very often in public telephony environments. In this case the number of call attempts due to T_1 can be assumed to follow a Poisson distribution
- On the other hand, the “hand-off” traffic is due to the calls handed-off from surrounding cells (T_2). It is more difficult to accept that the traffic created by T_2 attempts is Poissonian. This is because T_2 is a traffic stream that has already been carried by channels in the neighbouring cells and therefore coming from a population of no more than the total number of channels of all the surrounding cells together. As T_2 is originated by a finite population it should be modelled as smooth traffic.



a) Urban network scenario



b) Highway scenario

Fig. 1 Different traffic streams in a cellular network

This can be observed in Figure 1.a, where a typical urban scenario scheme of a cellular network is shown. It can be seen how the central cell receives hand-off traffic from just six neighbouring cells if an hexagonal pattern is used. If all cells are assumed to have the same number of channels, the number of traffic sources able to originate hand-off traffic T_2 is six times the number of channels in the tacked cell. As the number of sources is not much larger than the number of channels in the Base Station (BS), this traffic should never be considered to be Poissonian. This is even more obvious in a highway scenario as the one shown in Figure 1.b. In this particular case hand-off traffic arrives from only two neighbouring cells. Therefore, even if the number of channels assigned to the BS is higher due to the higher mobility of the mobile terminals in this scenario, the maximum number of incoming hand-off calls will still be low (twice the number of channels per cell). Theoretically, the handed-off traffic in this latter case should be very smooth, specially if the number of channels per BS is not very high.

Following the former discussion, as the overall traffic is the result of adding T_1 and T_2 streams - this is adding a Poisson traffic and a smooth one - this overall traffic should be better modelled as smooth traffic. The smoothness obviously depends on factors such as cell shape and size (a larger size reduces the hand-off and the share of T_2), speed and mobility of terminals (higher speed means more hand-off and then higher T_2), traffic intensity, etc.

B. Analytical studies

In [7] a theoretical model of the cellular network is proposed in order to model the overall traffic. Traffics, both fresh and hand-off, are characterised by its first moment only (average time between attempts to size a channel). The Poisson assumption is accepted because the analytical results agree with those reached through simulations. That is to say, offered traffic is handled like if it was Poisson and as the results obtained in such a way agree with those obtained from simulations the conclusion is that hand-off traffic can be handled like Poisson traffic.

In [8] handed-off traffic is modelled by the first two moments and it is shown that this traffic is in fact a smooth traffic process (with coefficient of variation of the inter-arrival time lower than one). The two moments representation of the hand-off traffic is superior to the single moment representation. The overall traffic in this case is showed be smooth for the case of hexagonal cell shape. Peakedness as low as 0.85 are obtained, and it can be easily concluded that much lower peakedness would be obtained for the highway scenario.

C. Field studies

The main problem found in field studies is that while

analytical research can deal with the attempts of sizing a channel, field studies can only see the actual seizures. Attempts which happen during blocking periods are lost in the field research, so arrivals and attempts are more different the higher is the blocking probability. Note that public cellular systems should achieve a very low blocking probability to give acceptable GoS.

In [6] the arrivals to the set of resources in a cell are studied, and the negative exponential distribution is fitted to the inter-arrival time distribution with a significance better than 15%. This agrees with the Poisson arrival hypothesis.

III. DATA ACQUISITION, PRE-PROCESSING AND STATISTICAL TOOLS

A. Data Acquisition

The real traffic data used in this study were obtained through a scanning receiver controlled by a Personal Computer (PC). The program that controls the scanner generates .log report of activity files. Each line of these files describes the time of the beginning of the activity, modulation and strength of the signal and duration of the activity. A scheme of the working station is shown in Figure 2.

The system monitored is a public cellular system in Barcelona based on the TACS standard (very similar to AMPS). Frequency Modulation (FM) is used, so the detection of the carrier in the down-link is sufficient for the knowledge of the channel occupancy. The advantage of detecting the down-link is the higher and more stable power level which helps to reduce the annoying effects of noise and interference.

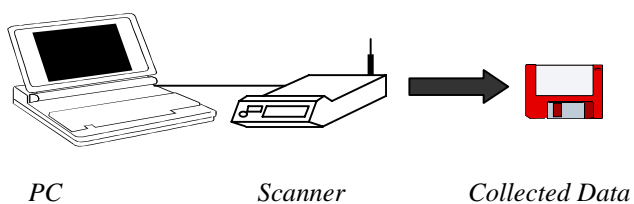


Fig. 2 Working Station

Twenty three frequencies belonging to different cells were scanned during a period of one month. Because of the co-channel and adjacent-channel interference present in cellular networks, we had to select the appropriate frequencies to be scanned. This was done by doing an aural survey of the channels in order to select frequencies free of interference that could distort the collected data.

In a first stage the frequencies were scanned during whole day periods in order to determine the busy hours. As the investigated system provides three different rate periods, the daily occupancy shows three spikes along the day (see

Figure 3); the busy hour was taken from the high rate busy period around ten hours which has the higher interest for traffic studies. Note that this first spike is more natural. Spikes at 16 and 20 hours are coincident with the beginning of the lower rate period and are in fact delayed traffic (traffic that waits to get economical advantage but which would have appeared before if the rate was kept).

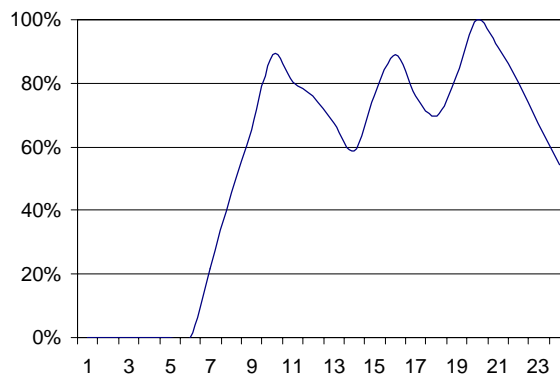


Fig. 3 Load along the time of the day (relative to the maximum load).

Among all the collected data three samples were selected for our study. We will refer to these samples as Heavy, Medium and Light Load. The Heavy Load sample corresponds to calls made in channels of average channel load around 0.6 Erlangs. In the same way, the Medium and Light Load samples correspond to calls in channels of average channel load around 0.5 Erlangs and 0.4 Erlangs respectively. To build up the three samples we have considered data in the 0.1 Erlang intervals [0.35,0.44], [0.45,0.54] and [0.55,0.64].

All these three samples were built up by the aggregation of the data collected at different scanned frequencies. This could be done so because all the scanned channels belong to the same cell and therefore present the same statistical properties.

We haven't considered heavier loads in order to avoid masking arrivals that may occur during active periods of the scanned channel. Anyway, loads heavier than 0.65 Erlangs represented less than a 1% of all the amount of data (4 out of 450 hours), so we think that the results obtained from the selected data accurately represent the nature of the inter-arrival process. Note that in order to keep a low blocking probability, loads can never be very heavy in a system for public voice service.

B. Pre-processing

Although the scanned channels are supposed to be free of interference, some short interference and cuts may still occur and distort the samples. Therefore the samples were

filtered in order to eliminate interference, short transmission cuts and other undesired effects.

This was basically done in two ways by a simple program written in C language. Short cuts that often occur due to fading were eliminated by merging registered activities separated by idle periods of less than 2 seconds. And interference was stripped by eliminating activities of less than 1 second. These bounds of 2 and 1 second respectively were checked by aural monitoring to give very good results for the desired purpose: more than 95% of the rejected data were actual cuts or interference and more than 95% of cuts and interference were rejected. The pre-processing is the same used in [4, 9] where further information can be found.

C. Statistical Tools

For our statistical analysis we have to follow these three steps:

- choose the probability density function to be fitted.
- fit the chosen function to the sample data.
- test the goodness of the fit.

In Section IV we describe the probability density functions chosen and we briefly explain the reasons for our choice.

The parameters of the different probability density functions fitted are calculated by making use of the Maximum Likelihood Estimation (MLE) [10]. We have chosen this method instead of Minimum Error or Moments method. This was motivated by the better results obtained when we compared the fitted distribution with the sample data used for the fit (goodness-of-fit test). Moreover we have seen that the MLE follows the shape of the empirical histogram better than the Moments Estimation (ME). The reason for this is that the ME uses the moments of the sample but not the sample itself to calculate the parameters of the fitted distribution, as the MLE does.

Once we have the parameters, we test the suitability of the fitting distributions by making use of the Kolmogorov-Smirnov (K-S) goodness-of-fit test. This is achieved by using Matlab programs that select, for every distribution tested, the parameters of the best fit from the K-S goodness of fit test point of view. The K-S is very simple to apply and has been widely used to fit telecommunications traffic [1]. The K-S test avoids the problems related with the bin size present in the chi-squared [6] and the Anderson-Darling [2]. This latter is more powerful than the K-S and chi-squared but the nature of the measured data makes the extra complexity not worthy.

The K-S test provides two figures of statistical interest, the modified K-S distance D and the level of significance α . The lower (higher) D (α) is, the better the fit will be. This is what will allow us to establish a ranking among those fitting distributions that pass the K-S test with a certain

level of significance (5% in our study). The level of significance α can be easily obtained from D and the size of the sample n with these two formulas [11]:

$$D_n = \frac{D}{\sqrt{n} + 0.12 + \frac{0.11}{\sqrt{n}}} \quad (1)$$

$$\alpha = 10^{(\log 2 - 2nD_n^2)} \quad (2)$$

We have seen that the results of the K-S test depend on the size of the sample. For sizes of less than 1000 the best fit was not clearly distinguishable from the other fitted distributions. As the size of the sample was closer to 1000 the best fit was easily distinguished. If the size of the sample was over 2000, the fitting began getting worse for all the distributions tested. The reason for this is that the sample data is not really originated from the fitted distribution, and as the sample grows it becomes more and more random. This is why the three samples used in our study have sizes around 1000.

It is important to stand out that we use the K-S test just to establish a ranking among the different theoretical distributions, according to the significance resulting from the K-S test. We don't try to find out the actual distribution behind the inter-arrival process. Once we have this ranking, the choice of the distribution used in any application will be a compromise between the level of significance and particular features that may be of our interest. Some of these features could be, for example, the simplicity of implementation in a computer program, the shape of the p.d.f. or that it takes into account a particular aspect of the system. The last one is the reason why we tried the Erlang-1- k distribution, as we will explain in section IV.

IV. FITTING DISTRIBUTIONS

In this Section the set of p.d. functions selected as candidates to fit the empirical distribution is introduced. Not every possible p.d.f. is an acceptable candidate to fit the underlying data. Two basic concepts have been taken into account to choose the candidates, obtaining the two following classes of distributions.

A. Markovian stages

Statistical distributions based on Markovian stages allow to identify some memory-less properties of the source which generates the sample. Some p.d.f. of this type which are more complex than the exponential allow also to identify different streams or populations generating arrivals, and this is the case introduced in Section II.

- *Exponential*: Is the most common distribution used for modelling the inter-arrival time process in Mobile Telephony Systems due to its memory-less properties.

For this reason we use it as a reference, although we show that it doesn't fit well the data.

$$f(t) = \frac{1}{\beta} e^{-\frac{t}{\beta}} \quad (3)$$

where β (the mean time) is a scale parameter.

- *Erlang-k*: This distribution has a simple interpretation as the sum of k independent exponential random variables, i.e., it is the result of a succession of memory-less exponential stages. Its coefficient of variation (cv) is always lower than one, so it will be a good candidate for our smooth samples.

$$f(t) = \beta^{-k} \frac{t^{k-1}}{(k-1)!} e^{-\frac{t}{\beta}} \quad (4)$$

where β is a scale parameter and k a shape parameter.

- *Erlang-n,k*: Displays a great versatility, fitting random variables of coefficient of variation lower or greater than one. Note that we expect smooth traffic and hence coefficients lower than one as stated in Section II.

$$f(t) = p\beta^{-n} \frac{t^{n-1}}{(n-1)!} e^{-\frac{t}{\beta}} + (1-p)\beta^{-k} \frac{t^{k-1}}{(k-1)!} e^{-\frac{t}{\beta}} \quad (5)$$

where n and k are shape parameters, and β is a scale parameter. p and $1-p$ are the probability parameters of the *Erlang-n* and *Erlang-k* components respectively.

- *HyperErlang-k-2*: A combination of two *Erlang-k*. Is an alternative to the *Lognormal-2* (see next paragraph) with the advantage that it can be represented as a combination of memory-less stages.

$$f(t) = p\beta_1^{-k} \frac{t^{k-1}}{(k-1)!} e^{-\frac{t}{\beta_1}} + (1-p)\beta_2^{-k} \frac{t^{k-1}}{(k-1)!} e^{-\frac{t}{\beta_2}} \quad (6)$$

where the parameters have the same interpretation that for the *Erlang-n,k*.

- *Erlang-1-k*: This is a particular case of the *Erlang-n-k* with $n=1$. We have chosen this distribution because of its theoretical interest. If this distribution would be the best fit of the sample or at least be a reasonably good fit, then the incoming traffic of a cell could be described as a result of two flows, one Poissonian (the one corresponding to $n=1$) and the other *Erlang-k*. The Poissonian flow could be identified with the fresh traffic originated inside the limits of the cell, and the *Erlang-k* would be identified with the non-Poissonian (smooth) hand-off traffic. Here, the probability parameter will be interpreted as the share of the T1

traffic stream and $1-p$ as the share of the T2 traffic stream.

B. Lognormal based distributions

The skew shown by the empirical distribution does not allow the normal distribution as a candidate. In such a case the lognormal and combinations represent a good choice. In addition, these p.d. functions give the best fit in previous research on the channel holding time [1,5].

- *Lognormal*: Displays a long tail with an initial spike, and suits very well with the shape of the filtered sample histogram.

$$f(t) = \frac{1}{t\sqrt{2\pi\sigma^2}} e^{-\frac{(\log(t)-\mu)^2}{2\sigma^2}} \quad (7)$$

where σ is a shape parameter and μ a scale parameter.

- *Lognormal-k*: Provides with a great power to fit difficult empirical distributions, with big initial spikes that can't be followed by other distributions.

$$f(t) = \sum_{i=1}^k p_i [p.d.f. \log normal(\sigma_i, \mu_i)] \quad (8)$$

where σ_i and μ_i have the same interpretation that for the *Lognormal* distribution and the p_i are probability parameters of the different *Lognormal* components.

V. STUDY OF THE INTER-ARRIVAL TIME

In this Section the numerical results of our study are shown. We classify these results according to the nominal load of the sample used in every case as mentioned in Section III: heavy (0.6), medium (0.5) and light (0.4) load. We call nominal load ρ_n to the center of the considered interval. It is important to stand out that the nominal load ρ_n is not necessarily equal to the average load ρ . Note that the latter is the average load of the whole set of studied channels along the observation interval.

In all three cases, we have found that the coefficient of variation of the *inter-arrival time process* is lower than one. This fact tends to corroborate that the arrival process is smooth as conjectured in Section II.

A. Heavy Load Sample ($\rho_n = 0.6$ Erlangs)

The Heavy Load Sample consists of 1223 inter-arrival times collected among 10 different scanned channels

belonging to the same cell. All the scanned channels have a load in the range [0.55, 0.64] with average load (actual load) of 0.58 Erlangs.

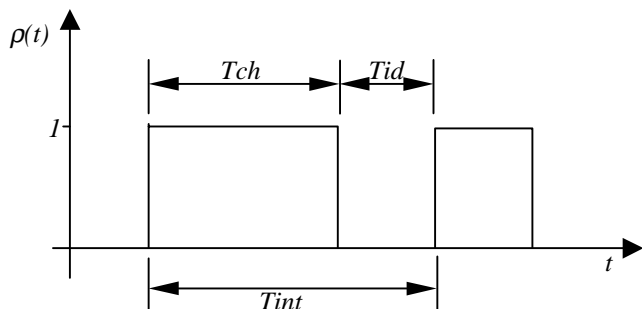


Fig. 4 Instantaneous load over a channel

In Figure 4 it can be seen that the inter-arrival time is the sum of the channel holding time and the idle time. Therefore, as the average load of the channel is the average amount of time that the channel is being used in relation to the total time, it can be found by calculating the ratio between the average channel holding time and the average inter-arrival time:

$$\rho = \frac{\bar{T}_{ch}}{\bar{T}_{ch} + \bar{T}_{id}} = \frac{\bar{T}_{ch}}{\bar{T}_{int}} \quad (9)$$

where \bar{T}_{ch} is the average channel holding time, \bar{T}_{id} the average idle time and \bar{T}_{int} the average inter-arrival time.

Obviously Equation (9) can also be directly applied to the whole set of studied channels by adding all the collected data for the channel holding time and inter-arrival or idle time from all the considered channels.

In Table I it can be seen that the lowest value of the statistical distance D (and then the best significance and fit) is obtained for the *Erlang-3,8* distribution. It is followed by a *Lognormal-2*, an *Hyper-Erlang-3-2* and a *Lognormal*. It can also be seen that the *Exponential* distribution is far from fitting the sample data.

Table I: Moments and fitting for inter-arrival time.

Fitting	Moments: $m_1 = 64.77$, $cv^2 = 0.49$ $\rho_n = 0.6$			
Exponential	$D: 6.02$	$\beta: 64.77$		
Erlang-3,8	$D: 0.59$	$\beta: 17.84$	$p: 0.875$	
Lognormal	$D: 1.13$	$\mu: 3.94$	$\sigma: 0.69$	
Lognormal-2	$D: 0.61$	$\mu_1: 4.02$	$\sigma_1: 0.63$	$p_1: 0.9$
		$\mu_2: 3.28$	$\sigma_2: 0.8$	
H-Erlang-3-2	$D: 0.67$	$\beta_1: 16.3$	$\beta_2: 33$	$p_1: 0.7$
Erlang-2	$D: 1.31$	$\beta: 32.4$		
Erlang-1,3	$D: 1.94$	$\beta: 20.27$	$1-p: 0.99$	

In Figure 5 we can see the histogram of the filtered sample and the probability density function (p.d.f.) of the best fit. Although it is not a good fit, in Figure 6 we show the histogram of the filtered sample along with its exponential fit.

The value of the modified distance (D) for the best fit is 0.59. For this value of D , the level of significance (α) is 40.6%. This is a very good result, considering that in telecommunications it is normal to work with levels of significance from 5 [1] to 15% [6].

It is interesting to see that for this case we find the smallest inter-arrival time average and biggest coefficient of variation among the three studied samples. The smallest average of the inter-arrival time can be easily explained by the heavier load over the scanned channel. Note that the average channel holding time is always the same under all loads as the samples are all taken at the same time of the day and cells are very similar.

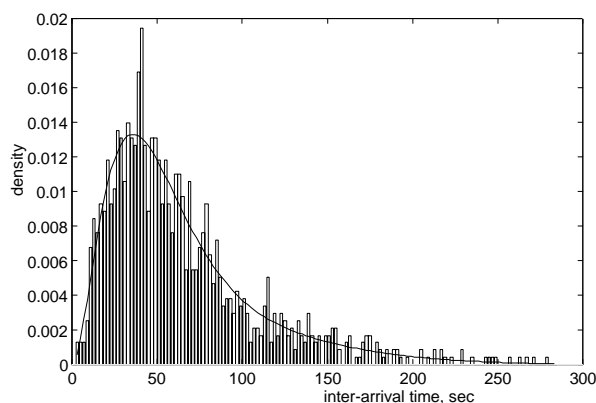


Fig. 5 Histogram of the Heavy Load sample along with the best fit pdf.

The explanation for the high coefficient of variation is not so simple. The inter-arrival time can be split in two components, the channel holding time and the idle time (see Figure 4):

- As the load increases, the idle periods are shorter and the inter-arrival takes statistical properties from the channel holding time. In [4] it is shown that the channel holding time in cellular systems has a coefficient of variation bigger than one. So as the load grows, the inter-arrival time becomes more similar to the channel holding time and therefore its coefficient of variation is bigger.
- When the load decreases the correlation between the inter-arrival and idle times is higher. The coefficient of variation of the idle time measured in this work is lower for lighter loads as shown in Section VI.

Then both effects tend to increase the coefficient of variation when the load grows and decrease it for lighter loads.

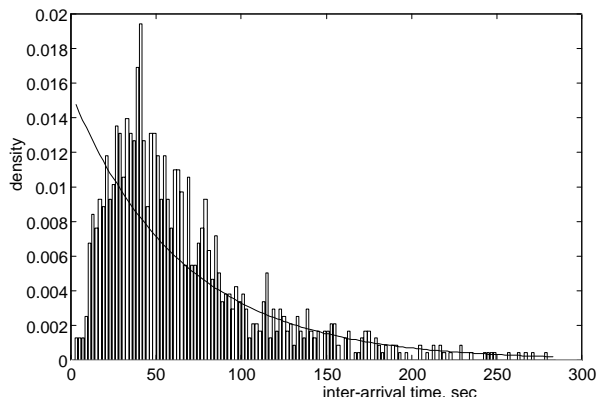


Fig. 6 Histogram of the Heavy Load sample and its exponential fit.

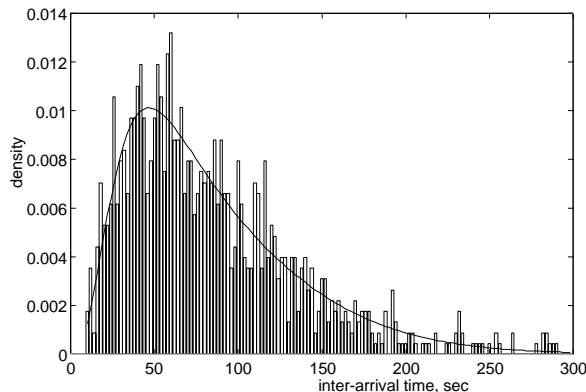


Fig. 7 Histogram of the Medium Load sample along with the best fit p.d.f.

B. Medium Load Sample ($\rho_n = 0.5$ Erlangs)

This sample consists of 1185 inter-arrival times collected among 10 different scanned channels of the same cell. All the scanned channels have a load in the range [0.45,0.54] with average (actual) load of 0.48 Erlangs obtained according to Equation (9).

Table II: Moments and fitting for inter-arrival time.

Fitting	Moments: $m_1 = 84.25$, $cv^2 = 0.38$ $\rho_n = 0.5$			
Exponential	$D: 6.52$	$\beta: 84.25$		
Erlang-3,6	$D: 0.6$	$\beta: 21.97$	$p: 0.722$	
Lognormal	$D: 1.25$	$\mu: 4.24$	$\sigma: 0.64$	
Lognormal-2	$D: 0.64$	$\mu_1: 4.61$	$\sigma_1: 0.462$	$p_1: 0.5$
		$\mu_2: 3.88$	$\sigma_2: 0.586$	
H-Erlang-4-2	$D: 0.57$	$\beta_1: 12.8$	$\beta_2: 26.6$	$p_1: 0.4$
Erlang-3	$D: 1.41$	$\beta: 20$		
Erlang-1,3	$D: 1.17$	$\beta: 29.52$	$1-p: 0.927$	

In Table II we see how that the best fit is an *Hyper-Erlang-4-2* distribution with a level of significance of 45.6%. It is followed by an *Erlang-3,6* with a level of significance of 37%. It can also be seen that functions such as the *Lognormal-2* or the *Erlang-1,3* fit the inter-arrival time sample quite well.

Here we can see how the average is bigger and the coefficient of variation smaller than in the previous case. The explanation is the same given for the heavy load case.

In Figure 7 the empirical histogram of the Medium Load sample along with the p.d.f. of the best fit - the *Hyper-Erlang-2-4* - are represented.

C. Light Load Sample ($\rho_n = 0.4$ Erlangs)

This sample consists of 1027 inter-arrival times collected among 10 different scanned channels of the same cell. All the scanned channels have a load in the range [0.45,0.44] with average 0.397 Erlangs obtained from Equation (9).

In Table III we see how the *Lognormal-2* has become the best fit, with a level of significance of 28%. In Figure 8 we can see the histogram of the Light Load sample depicted along with the pdf of the best fit.

In this case the square coefficient of variation is lower than in the first two cases. It is 0.33 and both the *Erlang-3,5* and the *Erlang-1,3* are good fits. In fact, if we look at the fitting of the *Erlang-n,k* along the three cases we see that *n* is always a 3. And when we force $n=1$ (*Erlang-1,k*) then *k* for the best fitting happens to be 3. Moreover the *Erlang-3* parts of this distributions are much better weighted than the other so it can be said that they are almost *Erlang-3* functions that have $cv^2 = 0.33$.

Table III: Moments and fitting for inter-arrival time.

Fitting	Moments: $m_1 = 90.59$, $cv^2 = 0.33$ $\rho_n = 0.4$			
Exponential	$D: 6.24$	$\beta: 90.59$		
Erlang-3,5	$D: 0.68$	$\beta: 27.99$	$p: 0.882$	
Lognormal	$D: 1.86$	$\mu: 4.32$	$\sigma: 0.63$	
Lognormal-2	$D: 0.65$	$\mu_1: 4.6$	$\sigma_1: 0.44$	$p_1: 0.5$
		$\mu_2: 3.94$	$\sigma_2: 0.63$	
H-Erlang-3-2	$D: 0.69$	$\beta_1: 30.9$	$\beta_2: 14$	$p_1: 0.9$
Erlang-3	$D: 0.8$	$\beta: 30.19$		
Erlang-1,3	$D: 0.79$	$\beta: 30.75$	$1-p: 0.973$	

As we explained in Section III it could be interesting to watch the sample data as the result of the addition of two

traffic streams. A Poissonian *fresh traffic* flow and a non-Poissonian *smooth hand-off traffic* flow.

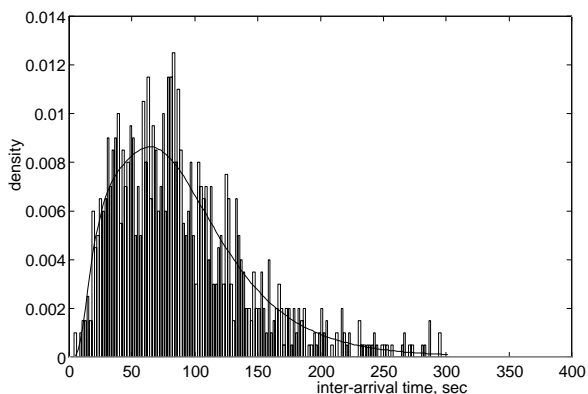


Fig. 8 Histogram of the Light Load sample along with the best fit p.d.f.

In Figure 9 we show how the *Erlang-1,3* function follows quite well the shape of the sample histogram. Although its level of significance of 11.6% is not the best one, it is quite good and worth to be considered due to the theoretical aspects involved. Even though, we leave this result pendent for further studies where the share of the traffic flows T2 and T1 might be studied. The reason for this is that, as we have explained in Section IV.A , the share of the fresh traffic T1 would be equal to p (Table III) that in this case is much smaller (0.027) than the share of the handed-off traffic T2 (0.973). This result doesn't seem to be realistic.

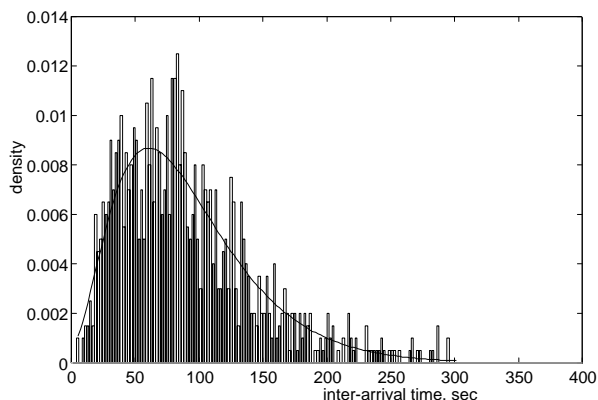


Fig. 9 Histogram of the Light Load sample along with the Erlang-1,3 fit pdf.

VI. OTHER STATISTICAL RESULTS

In this Section some additional statistical results are presented in order to reinforce the thesis of the non-exponential nature of the inter-arrival time. We will classify these results in two groups: Inter-arrival time statistics and Idle time statistics.

A. Inter-arrival time statistics

In Figure 10 the best fits for the different average channel loads are described, according to the level of significance resulting from the K-S test. We include the *Erlang-1,k* because of its theoretical interest explained in Section IV.

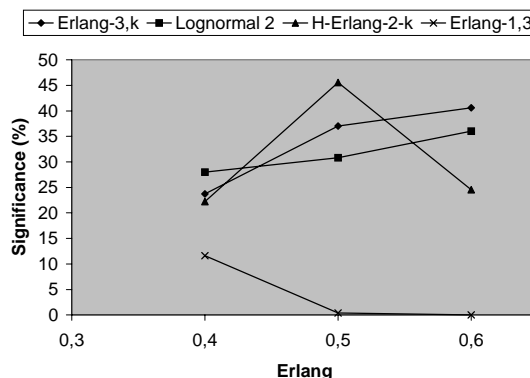


Fig. 10 Ranking of the different fitted distribution depending on the average channel load.

It can be seen that the *Erlang-3,k* is the most stable fit for the three loads studied in this paper. Moreover it is a very good fit with levels of significance among 25 and 40%.

Finally, in Figure 11 we show the average remaining time against the elapsed time. With this we try to probe in another way, that the inter-arrival time is not exponential.

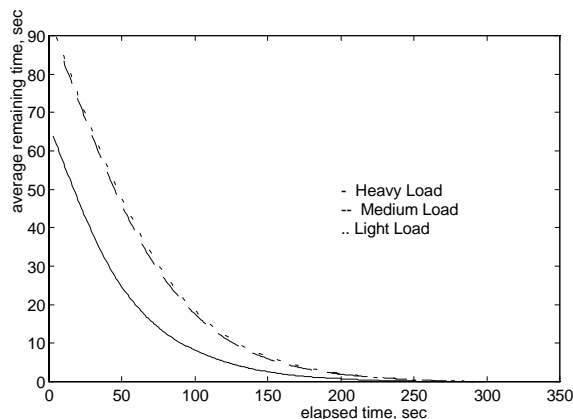


Fig. 11 Average remaining time vs elapsed time

We can see that the average remaining time clearly depends on the elapsed time. It has a decreasing dependence on it as opposite as it occurs for the call holding time [1] and the channel holding time [4, 5]. This means that as the time without new calls increases the average remaining time for a call to happen decreases.

If the inter-arrival time would be exponential, then this average remaining time would be constant and equal to its

β parameter due to the memory-less property of this distribution. As we show that this is far from being true we have one more reason, apart from the ones based on statistical fitting, to think that inter-arrival time in Mobile Telephony Systems is not exponentially distributed.

This curve was obtained as it is done in [1] for fixed Telephony circuit holding times. Each point of the curve is calculated by subtracting the elapsed time to all the values of the sample and then doing the average of the non-negative resulting values. This curves result much less scattered than in [4] for the channel holding time, the reason for this is that as it is shown in the mentioned paper, the channel holding time in cellular systems is hyper-exponential and therefore much more random than the hypo-exponential inter-arrival time.

B. Idle time statistics

As we have seen in Section V.A, the inter-arrival time is the result of the sum of the channel holding time and the idle time. Therefore, if we show that this idle time is not exponential then the inter-arrival time could never be considered exponential independently of the nature of the channel holding time.

We will try to show this in three different ways: by studying the second order statistics of the idle time of the studied samples, by studying the best fittings of the idle time and finally by studying the average remaining time as a function of the elapsed time.

1) Best fittings

Among all the candidate distributions, the exponential is far from being the best fit. In Figure 12 we show a ranking of the best fittings along the three different channel loads.

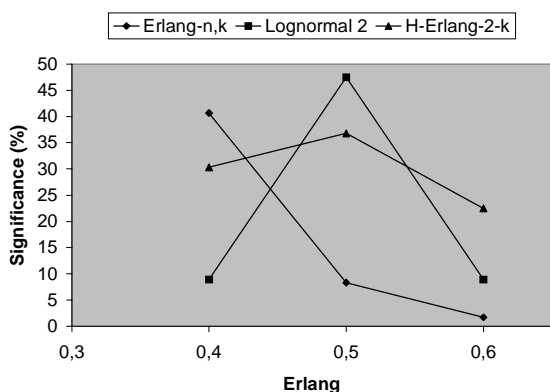


Fig. 12 Ranking of the best idle time fittings as a function of the channel load.

2) Average remaining time

As it was done above for the inter-arrival time, we try to show that the idle time is not exponential by showing that the average remaining time as a function of the elapsed time is not constant. In Figure 13 we can see that the average remaining time decreases as the elapsed time increases.

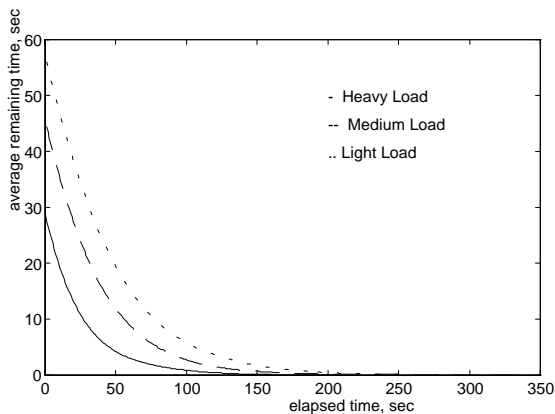


Fig. 13 Average remaining time vs. Elapsed time for the idle time process.

3) Second order statistics

For the three studied samples the coefficient of variation is lower than one. Actually, the values of this parameter are 0.8, 0.63 and 0.58 for the heavy, medium and light load samples respectively. It can be seen that the idle time is non-exponential.

VII. CONCLUSION

The coefficient of variation is $cv < 1$, i.e., the offered traffic is *smooth*. This tends to confirm of the results shown in [8] from an analytical point of view. Although the above mentioned study deals only with the hand-off traffic (T2), the fact that this traffic is smooth implies that the total offered traffic (T2+T1) cannot be, in any case, exponential. On the measured data we conclude that the inter-arrival time process is not exponential.

Although this would imply that the design and performance evaluation of these networks based on the assumption of an exponential inter-arrival time process overestimates the blocking probability, further studies over greater samples would be needed to confirm this in an experimental way.

The results presented in this paper differ from that presented in [6] where the exponential distribution agrees with the measured data. The difference must not cause surprise to the reader, because many factors influence on the inter-arrival time. Among them we can mention the mobility pattern, cell shape and size, mobile speed and others. The main conclusion of this paper is that all these

matters must be taken into account and measures are needed to effectively design the number of channels and save radio spectrum.

ACKNOWLEDGEMENT

This work was funded by Spanish CICYT Project TIC-97-0942.

REFERENCES

- [1] V. Bolotin, "Telephone circuit holding time distributions", *Proc. 14th International Teletraffic Congress*, 1994, Antibes, pp. 125-134.
- [2] Edward Chlebus, "Empirical validation of call holding time distribution in cellular communications systems", In *Telletraffic Contributions for the Information Age*, pp. 1179-1189, Elsevier, 1997.
- [3] Javier Jordán, Francisco Barceló, "Statistical of Channel Occupancy in Trunked PAMR Systems", In *Telletraffic Contributions for the Information Age (ITC 15)*, pp. 1169-1178, Elsevier, June 1997.
- [4] Francisco Barceló, Javier Jordán, "Channel holding time distribution in cellular telephony", *Electronics Letters*, Vol.34 No.2, pp. 146-147, 1998.
- [5] Francisco Barceló, Javier Jordán, "Channel holding time distribution in cellular telephony", *Proc. 9th International Conference On Wireless Communications*, 1997, Calgari, pp. 125-134.
- [6] C. Jedrzycki, V. C. M. Leung, "Probability Distribution of Channel Holding Time in Cellular Telephone Systems", *Proc of the IEEE Vehicular Technology Conf. VTC'96*, pp. 247-251, 1996.
- [7] Edward Chlebus, Wieslaw Ludwin, "Is Handoff Traffic Really Poissonian ?", *Proc of the IEEE ICUPC 95*, pp. 348-353, Nov. 1995.
- [8] M. Rajaratnam, F. Takawira, "Hand-off Traffic Modelling In Cellular Networks", *Proc of the IEEE GLOBECOM 97*, pp. 131-137, Nov. 1998.
- [9] Francisco Barceló, Sergio Bueno, "Idle and Inter-arrival Time Statistics in Public Access Mobile Radio (PAMR) Systems", *Proc of the IEEE GLOBECOM 97*, pp. 126-130, Nov. 98.
- [10] A.O. Allen, "Probability, Statistics, and Queuing Theory with Computer Science Applications", Academic Press 1990.
- [11] A.M. Law, W.D. Kelton, "Simulation Modelling and Analysis", McGraw-Hill 1991.