Comparative Results in GSM/GPRS Networks Modeling According to Erlang-B and Engset Traffic Models

Georgeta Budura, Cornel Balint, Marza Eugen
“Politehica” University of Timisoara
Bd. V. Parvan. No. 2
ROMANIA
georgeta.budura@etc.upt.ro, cornel.balint@etc.upt.ro

Abstract: - Resources sharing between different users and different services is a key concept of radio resources dimensioning in GSM/GPRS networks. This paper addresses the dimensioning problem of GSM/GPRS networks based on performance parameters focusing on the influence of different traffic models used. For voice traffic the Erlang-B model was considered in both cases meanwhile for data traffic we have adopted the Erlang-B model and the Engset model. Two radio resources allocation strategies: CP (Complete Partitioning) and PP (Partial Partitioning) were also considered. Based on performance parameters (blocking probability, preemption probability of voice traffic over data traffic, average total throughput and average throughput per user) we have compared the two models in order to find the adequate model to use for dimensioning purposes.

Key-Words: - GSM, GPRS, modeling, blocking and preemption probability, Erlang-B law, Engset law, average throughput

1 Introduction

The integration of General Packet Radio Service and Enhanced GPRS into GSM system world wide raises many problems. Analysis and study of data services in a cellular network require new models since traditional ones (like Erlang’s formulas) are not applicable to this kind of traffic. At the same time, the particularities of each GSM/GPRS/EDGE mainly on the resources allocation require an adaptation of the general model depending on the needs of each provider. For networks operators, equipment vendors and system integrators dimensioning rules have to be developed to plan and estimate the radio capacity that is needed for the predicted amount of data users when the radio resources are shared between circuit and packet switched services.

GSM operators have been dimensioning their networks for voice service in terms of offered voice traffic and blocking probability. The reference performance parameter for this system model is the Erlang-B formula [1]. This formula gives the proportion of calls that are blocked as a simple function of system capacity and traffic intensity. The GPRS network is designed for supporting several types of data traffic such as Wap, Web, E-Mail, etc. Therefore GPRS traffic process characterization is very demanding. Usually the data traffic is very bursty and relies to the application. A communication session may last for an extended period of time with intermittent packet transmissions.

On the other hand the GPRS service allows dynamic allocation of bandwidth resources. Wireless channels are allocated to a mobile terminal based on its traffic demands which results in a better resource utilization. This traffic behavior coupled with flexible bandwidth allocation in a GPRS network represent the starting point for constructing an appropriate model in order to evaluate the performance and to establish dimensioning rules for the system. Another major problem of GSM/GPRS networks dimensioning is the choice of strategy to partition the available cell capacity between traditional GSM and new GPRS services. The Radio Resources Manager (RRM) is in charge of optimizing the usage of radio resources, based on a specific resource sharing algorithm. Two static resources sharing schemes are used frequently:

- In the first one, called Complete Partitioning (CP), time-slots (TS) are divided into two sets and each type of traffic is allowed to use only its dedicated set.
- The second scheme, known as Partial Partitioning (PP), contains the following channel sets: one set shared between voice and data traffic and two sets each one being reserved for strict usage of its dedicated traffic: voice or data. This scheme offers many advantages: first, reserving a set of time-slots for each type of traffic allows...
guaranteeing, as in CP a minimum QoS for each type of traffic. Second, PP scheme provide a better efficiency than CP which is not suitable for maximizing radio utilization, especially in highly varying demand. Due to these advantages PP is widely implemented in a number of actually operating GSM/GPRS networks.

This paper compares the performance parameters of two models: the Erlang model and the modified Engset model, proposed to be used for dimensioning radio resources. Both models consider the voice-data interaction due to the optimization of radio resources as presented before.

Several papers have been published on traffic modeling and performance evaluation in GSM/GPRS networks. The major works in this field are based on analytical models using queuing theory and continuous-time Markov chains, and assuming an infinite number of users in the cell [1], [2] - [5]. In [7]-[9] analytical models based on discrete-time Markov chains have been proposed and a single type of traffic (data traffic) is considered. It is assumed to be generated by a finite number of users and modeled by an Erlang-like law. Other work is based on the modified Engset model [6]. In our study we first introduce some performance parameters for GPRS networks and then we propose a dimensioning method for estimating the number of TS allocated for each traffic type.

2 Basic assumptions regarding the system

As mentioned before we consider a single cell submitted to two different types of traffic: GSM voice calls and GPRS data flows. In traditional circuit-switched GSM networks, on each frequency carrier a 200 kHz bandwidth is shared between 8 voice calls. Each voice call is given a circuit, also called time-slot (TS) because it is a Time-Division multiplexing scheme (TDMA). Each voice call needs the assignment of a single time-slot for its entire duration.

GPRS data traffic uses the same radio interface as GSM voice calls hence radio resources available in the cell have to be shared among GSM and GPRS traffic. GPRS is a packet switching technology over circuit-switching based GSM system. In GPRS technology a mobile station can use several time-slots simultaneously for one application data flow to perform its transmission with a higher throughput. Each time-slot can be shared among several users by assigning different Temporary Flow Identities (TFI) to the mobiles. Each TFI identifies a GPRS physical connection called Temporary Block Flow (TBF). Up to 32 TFI’s can be allocated per TDMA frame due to the 5 bits allocated for TFI encoding at TRX level. Data flows are multiplexed by a PCU- based scheduling algorithm. In addition to time-slot partitioning, GPRS system allows time-slot aggregation: for a single mobile user the system can allocate up to d-time-slots simultaneously for downlink and up to u-time-slots simultaneously for uplink, depending on mobile station capability class (d+u). The choice of the number of TBF’s that a PDCH can have in uplink and downlink depends on the operator’s choice.

Our study is focused on the radio allocator which distributes the downlink radio channels among voice calls and GPRS data flows.

We make the following assumptions regarding the system to be modeled:
- \( t_B \) : the radio block duration, equal to 20\( \text{ms} \);  
- \( x_B \) : the number of data bytes that are transferred over one time-slot. \( \frac{x_B}{t_B} \) is the throughput offered by the RLC/MAC layer to the LLC layer. The RLC radio block size and the data rate according to the GPRS coding are indicated in TABLE 1.

<table>
<thead>
<tr>
<th>Schemes</th>
<th>CS-1</th>
<th>CS-2</th>
<th>CS-3</th>
<th>CS-4</th>
</tr>
</thead>
<tbody>
<tr>
<td>RLC block radio (bytes)</td>
<td>23</td>
<td>33</td>
<td>39</td>
<td>53</td>
</tr>
<tr>
<td>Data rate: ( \mu_{GPRS} (kbits/s) )</td>
<td>9.05</td>
<td>13.4</td>
<td>15.6</td>
<td>21.4</td>
</tr>
</tbody>
</table>

- \( TS \) : the number of time-slots of the TDMA partitioned into a contiguous set of \( TS_V \) time-slots dedicated to voice calls, \( TS_{VD} \) time-slots shared between voice and data and \( TS_D \) time-slots dedicated to GPRS; time-slots used by data \( TS_{D} + TS_{VD} \) are on a single TDMA which has a total of 8 time-slots.
- \( d \) (resp. \( u \)) : is the number of time-slots that can be used simultaneously for downlink (resp. uplink) traffic. All GPRS mobiles have the same radio capability, denoted \( d+u \).
- Voice calls have a preemptive priority over data flows on the shared part of the TDMA due to the fact that they generate the largest amount of revenue in most actual operating systems. As a consequence, if all \( TS_V \) time-slots dedicated to voice are occupied and all \( TS_{VD} \) time-slots are in use with at least one of them allocated to data, then one time-
slot assigned to GPRS traffic in the shared part of the TDMA will be reallocated to voice on the arrival of a GSM request.

3 System Models

3.1 Voice traffic model

In order to model the GSM traffic, it can be assumed that the calls arrivals are a Poisson process of intensity $\lambda_v$. Each call will have a random exponential duration of mean $1/\mu_v$. The classical Markov chain model applies for voice traffic and the steady-state voice probabilities given by relation (1) are generated by the birth-death structure of this model shown in Fig. 1.

$$p_v(t) = \frac{\rho_v^t}{t!} e^{-\rho_v}, \quad \rho_v = \frac{\lambda_v}{\mu_v}, \quad t \in [0,TS_v)$$

Fig.1 The voice traffic model

The performance parameter in this case is represented by the blocking probability also known as Erlang-B formula:

$$B_{v,CP} = \frac{\rho_v^{TS_v}}{\rho_v^{TS_v} + \sum_{i=0}^{\infty} \frac{\rho_v^i}{i!}}$$

This model applies independent of the resource allocation strategies and the data model adopted. The allocation strategy influences only the total number of resources available for voice calls in equation 2 (CP: $TS_v$; PP: $TS_v + TS_{VD}$) due to the priority of voice over data calls.

3.2 Data Traffic Models

3.2.1 General features and assumptions

Data traffic is modeled assuming the following parameters:

- $N$ represents the fixed number of data mobiles in the cell. Each mobile is doing an ON/OFF traffic with an infinite number of pages
- ON periods correspond to the download of an element like a WAP, a WEB page, an email, a file, etc. Its size is characterized by a discrete random variable $X_{on}$, with an average value $E[\sigma]$.
- OFF periods correspond to the reading time of the last downloaded element, which is modeled as a random variable $T_{off}$ with an average value of $E[\tau]$ seconds.
- The maximum number of GPRS users in active transfer is given by:

$$n_{max}(TS_{D}) = \min(N,32,mTS_{D}) \quad (3)$$

$m$- is the maximum number of users that can use a single time-slot.

For constructing our model based on the ON/OFF traffic we define the average data traffic parameters as follows:

- The average rate of data arrival process:

$$\lambda_D = \frac{1}{E[\tau]} \quad (4)$$

- The average data rate per time-slot:

$$\mu_D = \frac{x_B}{E[\sigma]E[\tau]} = \frac{\mu_{GPRS}}{E[\sigma]} \quad (5)$$

Based on these two parameters we define, as shown in relation (6) a parameter $\rho_D$ that characterized data traffic, similar to $\rho_v$.

$$\rho_D = \frac{\lambda_D}{\mu_D} = \frac{t_B E[\sigma]}{x_B E[\tau]} = \frac{E[\sigma]}{E[\tau]} \frac{1}{\mu_{GPRS}} \quad (6)$$

3.2.2 Data traffic model based on Erlang law

System with complete partitioning (CP)

Based on system parameters indicated by relations (3)+(6), the data traffic model is represented by the continuous-time Markov-chain shown in Fig.2.

Fig. 2 Continuous-time Markov-chain model

As indicated in Fig. 2, the state $j$ of the Markov chain corresponds to the number of the data mobiles that are simultaneously in active transfer. The maximum bandwidth capacity they can use is $TS_{D}$.

The stationary probabilities of having $j$ data
mobile in active transfer, derived from the birth death structure of the Markov chain are:

\[
\text{for } j \in (0, j_0]:
\]

\[
p_D(j) = \frac{N!}{j! (N-j)!} \rho^j_D p_D(0)
\]

\[
\text{for } j \in (j_0, n_{\text{max}}]:
\]

\[
p_D(j) = \frac{N!}{j_0! (N-j)!} \rho^j_D p_D(0)
\]

where \(j_0\) is the maximum value of integer \(j\) satisfying the relation: \(jd < TS_D\).

Based on these distributions we have calculated the following performance parameters:
- the blocking probability, similar to the Erlang-B law [8], [10]:
\[
B^{(1)}_{CP} = \frac{N!}{j_0! (N-j_0)!} \rho_j^{n_{\text{max}}} p_D(0)
\]

- the average total throughput:
\[
X^{(1)}_{CP} = \sum_{j=1}^{n_{\text{max}}} p_D(j) j \min(d, TS_D/j) \mu_{\text{GPRS}}
\]

- the average throughput per user:
\[
X^{(1)}_{CP} = \frac{X^{(1)}_{CP}}{E[j]} = \sum_{j=1}^{n_{\text{max}}} p(j) j \min(d, TS_D/j) \mu_{\text{GPRS}}
\]

where \(E[j]\) represents the average number of data mobiles in active transfer.

**System with partial partitioning (PP)**

We denote by \(TS_{\text{max}}(t)\) the number of time-slots that data mobiles can use when there are \(t\) voice calls in the system:
\[
TS_{\text{max}}(t) = TS_D + TS_{\text{VD}} - \max(0, t - TS_V)
\]

with \(t\) taking the values: \(t \leq TS_V + TS_{\text{VD}}\).

We also consider \(N_{\text{max}}(t)\), the maximum number of data mobiles that can simultaneously be in active transfer, when there are \(t\) pending voice calls. \(N_{\text{max}}(t)\) can be derived, for \(t \leq TS_V + TS_{\text{VD}}\), as follows:
\[
N_{\text{max}}(t) = \min(32, 7TS_{\text{max}}(t), mTS_{\text{max}}(t), N)
\]

The model applied to this system is a bi-dimensional Markov chain proposed in [11]. In order to reduce the bi-dimensional Markov chain complexity the probability of a generic state \((t, n)\) was approximated with the conditional product-form probability as indicated by relation (13).
\[
p(t, n) = p_v(t)p_D(n|t)
\]

As mentioned before, voice traffic can be modeled with the Erlang model. Due to the fact that there are currently \(t\) voice mobiles in communication, the data traffic is modeled by the Erlang-like model developed in [8] with a number of time-slots equal to \(N_{\text{max}}(t)\).

The stationary probabilities of having \(n\) data mobiles in active transfer conditioned by the state \(t\) of voice calls can be derived as follows:
\[
\text{for } n \in (0, n_0(t)]
\]

\[
p_D(n|t) = \frac{N!p_D^n(0|t)}{n! (N-n)!}
\]

\[
\text{for } n \in (n_0(t), N_{\text{max}}(t)]
\]

\[
p_D(n|t) = \frac{N!p_D^n(0|t)}{n_0(t)! (N-n_0(t)-1)! (N-1)!}
\]

The PP system’s performance parameters can be derived easily from the detailed stationary probabilities, \(p_D(n|t)\):
- the data blocking probability formula is as indicated in [10]:
\[
B^{(1)}_{PP} = \sum_{i=0}^{TS_V + TS_{\text{VD}}-N_{\text{max}}(t)} p(t)p_D(N_{\text{max}}(t)|t)
\]

- the preemption probability (the probability that a data transfer ends prematurely due to voice call preemption) is:
\[
B^{(1)}_{P,pp} = \sum_{i=TS_V + TS_{\text{VD}}-N_{\text{max}}(t)}^{TS_V + TS_{\text{VD}}-N_{\text{max}}(t)+1} p(t)p_D(n|t)
\]

- for the average total throughput we propose the formula:
\[
X^{(1)}_{PP} = \sum_{i=0}^{TS_V + TS_{\text{VD}}-N_{\text{max}}(t)} \sum_{n=0}^{N_{\text{max}}(t)} p(t)p_D(n|t) \min(nd,TS_{\text{max}}(t)) \mu_{\text{GPRS}}
\]

- the average throughput per user:
\[
X^{(1)}_{PP} = \frac{X^{(1)}_{PP}}{E[n]}
\]

where \(E[n]\) represents the average number of data mobiles in active transfer.

**3.2.3 Data traffic model based on the modified Engset law**

In order to construct our model we consider the same continuous-time Markov chain represented in Fig.2.

**System with complete partitioning (CP)**

According to the Engset model [6] and the data traffic parameters mentioned before, the steady-state
probability \( p_D(j) \) can be expressed as indicated by relation (19).

\[
p_D(j) = p_D(0) \prod_{i=1}^{j} \min(d(TS_D), \rho'_D) \quad (19)
\]

Based on these probabilities we have calculated the performance parameters: blocking probability \((B_D^{(2)})\), average total throughput \((X_D^{(2)})\) and the average throughput per user \((X_{u,D}^{(2)})\) and have founded similar formulas as presented before for the Erlang-model.

**System with partial partitioning (PP)**

The basic idea in constructing the model in this case relies on 2 assumptions: the voice calls are independent of GPRS connections and the voice and data traffic evolves at different time scales [6].

The data traffic model is constructed with respect to the Engset model and the particularities imposed by the resources sharing capability as presented above.

The data blocking probability is deduced similarly as in [6]:

\[
B_{DP}^{(2)} = \sum_{s=0}^{TS_D} p_V(s)B_D^{(2)}(\min(TS - TS_V, TS - s)) \quad (20)
\]

For the total average throughput we have used the formula:

\[
X_{DP}^{(2)} = \sum_{s=0}^{TS_D} p_V(s)X_D^{(2)}(\min(TS - TS_V, TS - s)) \quad (21)
\]

The average throughput per user is:

\[
X_{u,DP}^{(2)} = \sum_{s=0}^{TS_D} p_V(s)X_{u,D}^{(2)}(\min(TS - TS_V, TS - s)) \quad (22)
\]

Finally the preemption probability due to voice calls priority over data users is:

\[
B_{E,DP}^{(2)} = \sum_{s=TS_D}^{TS_D-1} p_V(s) \sum_{k=\max(s+1,TS_D)}^{TS_D} B_D^{(2)}(\min(TS - TS_V, TS - k)) \quad (23)
\]

**4 Comparative results**

Both models have been implemented by Matlab programs and the performance parameters mentioned before have been determined and compared. The same scenario has been experimented for the presented models: Erlang and Engset. For the CP scheme we have considered a cell equipped with a single TRX \((TS=8; TS_D=7; TS_{D}=1)\). For the PP allocation strategy the cell was also equipped with a single TRX \((TS=8; TS=3; TS_{D}=1; TS_{D}=1)\). The following system parameters have been considered in our experiments:

- voice traffic parameters: \( \rho_V = 2.94 \) Erlang; 2% lost traffic according to Erlang –B formula.
- data traffic parameters: \( E[\sigma] = 5 KB \), \( E[\tau] = 12 s \), GPRS mobile class: 4+1 and CS2 coding scheme \( (\mu_{GPRS} = 13, 4kbits/s) \).

In Fig. 3 we have represented the blocking probabilities according to the presented formulas for \( B_D^{(1)} \), \( B_{DP}^{(1)} \), \( B_D^{(2)} \) and \( B_{DP}^{(2)} \). In the figure can be observed that differences between the two models are not significant (maximum 5% for \( B_D^{(1)} \) and 2% for \( B_{DP}^{(1)} \)).

![Fig. 3 Blocking probabilities for Erlang and Engset models](image)

Figure 4 shows a perfect matching regarding the preemption probabilities according to Erlang and Engset models.

![Fig. 4 Preemption probabilities for Erlang and Engset models](image)

For dimensioning purposes average throughput per user represents an important parameter. Fig. 5 represents \( X_{u,D}^{(1)} \), \( X_{u,D}^{(1)} \), \( X_{u,D}^{(2)} \) and \( X_{u,D}^{(2)} \). The curves show significant differences (20%) between the Erlang and the Engset models only for a relative...
small number of mobiles in the cell. For the region of interest in dimensioning process, corresponding to a large number of mobiles, the two models provide identical results.

For each model (Erlang and Engset) we have proposed and calculated performance parameters: blocking probability, preemption probability, average total throughput and average throughput per user. These parameters have been compared using two allocation strategies: CP and PP.

The comparative studies based on performance parameters mentioned before show that both models can be used for dimensioning purposes. The represented results indicate very small differences between Erlang and Engset models. Regarding the computational complexity, the Engset model is much more simple than the bidimensional Markov chain used by the Erlang model.

5 Conclusion
In this paper we have implemented two traffic models dedicated to GSM/GPRS networks performances analysis. Both models consider the voice-data interaction in order to optimize radio resources and both allow performance parameters computation based on voice and data traffic loads.

References: