Capacity Analysis in a Multi-service Mobile Broadband System

Fernando J. Velez[1,2], Luis M. Correia[2]
[1] Department of Electromechanical Engineering
University of Beira Interior
6200 Covilhã, Portugal
[2] Instituto de Telecomunicações, Instituto Superior Técnico
Technical University of Lisbon
Av. Rovisco Pais, 1049-001 Lisboa, Portugal
fjv@ubi.pt, luis.correia@lx.it.pt

Abstract

Multi-service traffic engineering will be a key aspect in cellular planning for Mobile Broadband Systems, the feasibility of the aggregate traffic model being crucial. An MMPP model is proposed for the modelling of the superimposition of data and video IPP sources. Given the correspondence between applications and their bearer service components, an algorithm for the Bernoulli case of the Bernoulli-Poisson-Pascal model is used to compute the blocking probability. The supported number of users is higher in scenarios with a lower maximum load per user. In a cell with 384 channels of 384 kb/s, it varies from 20 to 26 users/cell in urban scenarios, while with 288 channels/cell, although the number of channels is only 25 % lower, the supported number of users decreases 50 %. High terminal mobility strongly degrades the performance in the roads scenario.

1 Introduction

Around ten years from now, Mobile Broadband Systems (MBS) [1] will play an important role in the mobile communications market, mostly in urban areas to cover hotspots in the centre of large cities, where a very high demand is foreseen. Although it is not necessary that future MBS will be based on ATM technology, the work from RACE-MBS [1] and ACTS-SAMBA [2] projects considered so, this being the approach one follows in this paper.

In ATM networks the available resources are shared in a way that allows multiplexing of different traffic sources. As far as the different sources do not take these peak values simultaneously, for a fixed number of users, the network can use less resources than would be required if resources were assigned according to the peak amounts required by each user, and a gain exists from this statistical multiplexing [3].

In order to model multi-service traffic in MBS one needs to know the model for the air interface access, as well as for the slot arrival process. Because traffic can be generated from different mixtures of voice, data or video sources, it is important to obtain performance measures for resource usage, making use of the characteristics of the frame structures [2] and of the MAC (Medium Access Control) protocols. The DSA++ MAC protocol, which is nowadays being under consideration in the wireless ATM (HIPERLAN-type2) standardisation process of the ATM Forum and the Broadband Radio Access Networks project of ETSI [4], allows considering connection-oriented communications [5]. By allocating a so-called container, formed by a number of slots, a base station (BS) defines channels like in circuit-switched connections. Thus, the methodologies for circuit-switched network analysis supporting heterogeneous traffic can be applied, while the MAC protocol guarantees that the maximum delay is kept under a threshold value that does not affect the performance of applications [4], namely real-time ones.

Therefore, the identification of relevant models for the characterisation of voice, data and video traffic sources is needed, in view of finding a unified model to evaluate the Quality of Service (QoS), which depends on the aggregate traffic. Hence, keeping in mind that these models are needed for MBS cellular planning and optimisation purposes, the implementation feasibility of the aggregate traffic model is crucial in the choice of the basic model(s) for the sources. An analytical approach is sought, instead of a simulation one.

In Section 2, one gives a brief overview of the models for bursty traffic, the choice of an approximated one for the aggregate traffic being justified. In Section 3 the Bernoulli-Poisson-Pascal model is proposed for the computation of the blocking probability. Then, in Section 4, the service components characterisation parameters are presented, and the user activation model for service components is briefly referred, the correspondence between applications and service components being described. Some important aspects on channel usage are also discussed. Finally, in

Section 5, examples are worked out for actual deployment scenarios. One determines the supported number of active users in a cell, and results are presented for the spectral efficiency. The impact of terminal mobility in the results is also briefly described. Conclusions are drawn at the end.

2 Aggregate Traffic

The difficulty in the multi-service traffic analysis comes from modelling the arrival process that results from superimposing a number of independent sources. When evaluating the performance of ATM networks, it is important that the superimposition of arrival processes is easy, or alternatively one has to find models for the superimposition.

In this work one adopted for model superimposition [3]. This approach consists in first determining models for individual processes, and then superimposing them. This method is applicable only when the superimposed model can be obtained directly from individual models such as a Markov-Modulated Poisson Process (MMPP) one. If MMPPs are used as individual process models, their superimposition can be described by an MMPP with a larger state space. An example for the superimposition of video and voice traffics is presented in [3].

MMPP is a type of Markov-modulated models, an extremely important class of traffic models where an explicit notion of state is introduced into the description of a traffic stream. An auxiliary Markov process is evolving in time and its current state controls (modulates) the probability law of the traffic process. MMPP combines the simplicity of the modulating (Markov) process with that of the modulated (Poisson) process. In this case, the modulation mechanism simply stipulates that in state k of the process, arrivals occur according to a Poisson process at rate λ_k. As the state changes, so does the rate. The introduction of MMPP processes allows to model time-varying sources, while keeping the analytical solution of the related queuing performance tractable [6].

MMPP can model mixtures of voice, data and video. The performance measures, such as queuing distribution and the moments of the delay distribution can be obtained using the MMPP/G/1 queue analysis [7]. In [6], an example is given for voice that considers a two-state MMPP model, where one state is an ON one with an associated positive Poisson rate, and the other is an OFF one with associated zero-rate, such models being known as Interrupted Poisson Processes (IPP) for obvious reasons. It is also worthwhile noting that an MMPP process with P+1 states can be obtained by the superposition of P identical independent IPP sources.

More recent studies have also concluded that superimposed traffic is self-similar, but one will not follow that approach here. Instead, one is going to consider the MMPP model, whilst considering that the IPP models the various types of traffic sources.

MMPP is one of the most well known models, unlike the majority of others, which are normally exclusive for a given traffic source type and not much used. Even though, in those cases, MMPP is used as a reference for comparison purposes. As its parameters can be easily estimated from empirical data, operators can use it for cellular planning purposes. They normally have operational data for average duration of calls, and even for the characteristics of service components of each type, as well as for their generation rate. These data can be used to compute the elements of the transition matrix associated with the process, calculating the fraction of time when the process switched from a state to another.

3 BPP Model and Algorithm

The performance measure that one is interested in is the probability that an arriving customer is blocked, i.e., the connection (call) blocking probability, P_b.

Because of terminal mobility and the resulting handovers, the network planner has also to account for the handover failure probability (the probability that a user will not succeed in transferring its connection from a cell to another). The resulting connection-dropping probability is the probability of forced termination of the connection during its duration. When a single service is considered, and if one does not use guard channels for handover, the handover failure probability is equal to the blocking probability [8]. One generalises this approach, following it here for multi-service traffic.

MBS resources serve applications via different service components, i.e., the system itself serves the service components, which, in turn, serve the applications. Different applications have different durations, and different associated data rates. Each application data rate is obtained by summing the data rates of its different service components, weighted by the proportion of time they area active during the application, on average. As far as each component can be “turned” on and off several times during the application, the proportion of time it is active also depends on the number of times it is activated. Thus, one is dealing with a mixture of different service components, the correspondence between applications and service components having to be completely defined by the user model.

In the general model of a loss system with a given number of types of channels shared by J classes, a customer arrival at the resources follows a specific random process. Each customer requests a fixed number of channel units, which are granted if available. If not, the request is cleared and the customer is blocked. The classification of customers is done on the basis of their arrival process, capacity requirement, and mean holding time [9].

The service components are generated according to the Bernoulli case of a Bernoulli-Poisson-Pascal (BPP) process. The service component j capacity demand is a_j (j =1, …, J and $a_j \in$ IN). The time that these channels, once granted, will be held by the service component j is i.i.d. (independent and identically distributed), and is specified by its mean value $\overline{\tau_j}$. Its specific distribution has no influence on the calculations that one is pursuing [9], which is known as the *insensitivity* property. Thus, the capacity vector **A** is of the following type

$$\mathbf{A} = [a_j],\ j = 1,...,J \quad (1)$$

In a complete sharing policy, the set of possible states $\mathcal{N}$ is bounded as a result of the finite channel capacity

$$\mathcal{N} = \left\{ \boldsymbol{n} \in \mathrm{IN}^J : [n_1,...,n_J] \bullet \begin{bmatrix} a_1 \\ ... \\ a_J \end{bmatrix} \le c \right\}, \quad (2)$$

c being the number of available channels.

In the limit, if there are more users from an application than from another, the other can use fewer channels, Fig. 1.

When the system is in state $\boldsymbol{n}$, the time until the next arrival of a service component j's demand is exponentially distributed with parameter $\lambda_j(n_j)$. This parameter is however normalised with respect to the average component j holding time [9]; thus, a different unit of time is introduced for each customer class.

Blocking takes place if a service component j request arrives when the system is in the set

$$\mathrm{B}_j = \{\boldsymbol{n} \in \mathcal{N} : \boldsymbol{n} \cdot \mathbf{A} + a_j > c\}. \quad (3)$$

In this case, the request cannot be granted. It will be cleared and the customer blocked, which means that the system will remain in the same state.

Given the activation rate, ($-\beta_j$), and the arrival rate, α_j, in the Bernoulli case, the arrival intensity conditioned to n_j customers being in the system is [9]

$$\lambda_j(n_j) = (N_j - n_j) \cdot (-\beta_j), \quad (4)$$

with $\beta_j < 0$.

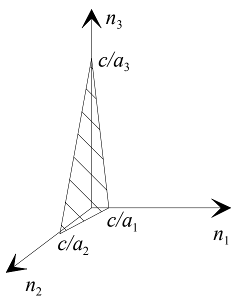

Fig. 1 Boundaries for resource usage for a complete sharing policy, $J = 3$.

Note that $N_j = -\alpha_j/\beta_j$ must be a positive integer for an equilibrium occupancy *pmf* (probability mass function) to exist [9]. Although for the particular case of exponential holding times, the BPP process can be modelled by a Markov chain [9], it is worthwhile noting that the model allows considering general distributed holding times.

The blocking probability of the service component j is given by the ratio between the expectation of the number of blocked requests, per time unit, and the total number of service component requests, per time unit. It depends on ($-\beta_j$) and α_j, which on their own depend on other parameters, as defined later. The algorithm for its computation is given in [9], where an efficient auxiliary recursion is used to compute the channel occupancy *pmf*.

4 Bearer Service Components

4.1 Characterisation

The MMPP model is adequate to compute each service component blocking probability, considering the aggregate traffic. In particular, considering voice, data and video sources being modelled by the IPP model, one can use the Bernoulli case of the BPP model [9]. This is because the service component arrival process, with parameters a_j, α_j and β_j, can be interpreted as the superposition of $N_j = -\alpha_j/\beta_j$ IPP sources, where the channel requirement in the state ON is a_j [9].

The basic components are audio, data and video, a correspondence existing between types of information supporting applications (e.g., moving pictures, sound or multimedia) and service components, Fig. 2.

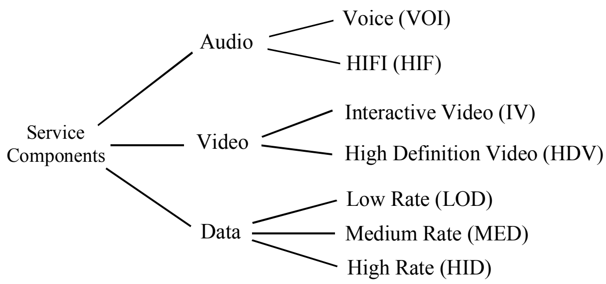

Fig. 2 Service components.

The audio component is not considered because it is not relevant for MBS traffic analysis purposes, since it is a very low data rate component. For data and video, a sub-division exists that can be mapped into the set of service components "to be used", Table 1. B_{sj} represents the bit rate associated with service component j. The way the basic service component (BAS) replaces the LOD and the IV1 components at 384 kb/s is shown in this Table, as well as how the MD3 (medium data rate data, type 3, at 1 920 kb/s) component stands both for the actual MD3 and for interactive video components (IV2) at 1 920 kb/s. Note that $B_{sj} = a_j \cdot B_{s1}$

Table 1 – Definition of actual service components.

Service component sub-division			j	a_j	B_{sj} [kb/s]
Data	Video	"to be used"			
LOD	IV1	BAS	1	1	384
MD1	-	MD1	2	3	1 152
MD2	-	MD2	3	4	1 536
MD3	IV2	MD3	4	5	1 920
-	HDV	HDV	5	21	8 064
HID	-	HID	6	83	31 872

This mapping can only be established because of two reasons

- under the hypothesis one has been assuming, pairs LOD / IV1 and IV2 / MD3 do have equal characteristics;
- these pairs are not used simultaneously by the considered mixture of applications, thus, discrimination between the components in each pair is direct for each application.

From now on, one assumes the use of a basic bit rate of 384 kb/s.

4.2 Correspondence with Applications

One considers the set of applications defined in [11], operating in the BCC (business city centre), URB (urban) and ROA (main roads) scenarios, their main parameters being presented in [10].

The usage of service component j is characterised by defining the number of times it is accessed during application k, $n_{j|k}$, and by its service rate given application k, $\mu_{j|k}$. Application k data rate is given by [12]

$$b_k = \sum_{j=\mathrm{BAS}}^{\mathrm{HID}} \frac{n_{j|k} \cdot \mu_{j|k}^{-1}}{\mu_k^{-1}} \cdot a_j \qquad (5)$$

where μ_k is the service rate associated with application k. There is a distinction between the up- and downlinks. The asymmetry ratio defines the ratio of b_k between the down- and uplinks. For example, HVT is a symmetric real time short duration application with b_{HVT} = 1920 kb/s and burstiness (the peak to mean data rates ratio) equal to 1. In contrast, ENP is a strongly asymmetric long duration bursty application.

Given the average duration and the range of foreseen values for the burstiness and for the asymmetry ratio [10], one defined the applications characteristics, its numerical definition strongly depending on the service components definition. In Table 2 one presents the application k data rates and its burstiness both for up- and downlinks, and also its holding time, $\bar{\tau}_k$.

The characteristics of service components that are not permanent can be summarised as follows

Table 2- Application parameters.

Application	Abbre-viation	$\bar{\tau}_k$ [min]	b_k [kb/s]		Burstiness	
			UP	DOWN	UP	DOWN
HD Video-telephony	HVT	3	1920	1920	1	1
ISDN-Videoconference	IVC	30	384	384	1	1
Mobile Video Surveillance	MVS	120	1920	1	1	480
HDTV Outside Broadcast	HOB	50	8068	1924	1.11	1.43
Wireless LAN Interconnect	WLI	15	146	4032	7.90	7.91
Data File Transfer (FTP)	FTP	0.33	19	384	23.85	1
Professional Images	PIM	10	384	8064	1	1
Desktop Multimedia	DMM	5	63	49	6.06	7.90
Mobile Emergency Serv.	MES	20	2731	2731	4.08	4.08
Mobile Repair Assistance	MRA	40	2328	2328	4.78	4.69
Mobile Tele-working	MTW	20	1930	1930	1.59	1.59
Freight & Fleet Managemt.	FFM	5	2736	2736	4.07	4.07
Electronic Mailbox Service for Multimedia	EMB	1	63	1536	6.06	1
E-commerce	ECO	5	16	49	24.15	7.90
Multimedia Library	MML	40	5	2328	240	1.32
Tourist Information	TIN	15	76	243	15.06	7.90
Remote Procedure Call	RPC	5	10	194	120	7.90
Urban Guidance	UGD	5	1935	1935	1.59	1.59
Assistance in Travel	ATR	20	1935	1935	1.59	1.59
TV Programme Distribut.	TVD	90	0	8064	-	1
E-newspaper	E-NP	20	1	243	480	7.90

- the value $\bar{\tau}_{j|k} = \mu_{j|k}^{-1}$ = 0.055 min refers to the IPP model for Web browsing data from [4], correspon-ding to ON / OFF average durations of 3.3 s and 22.8 s, respectively (e.g., it is used to support the MD3 component in ENP at the downlink)
- other values refer to the characterisation given in [12]: a duration of 0.0083 min is used for data, e.g., for the BAS component in FTP at the uplink, while 0.5 min is used for video, e.g., for the HDV component in MES at the uplink.

Permanent service components are only activated once and have the same duration of the application it is serving. HVT is an example of an application having permanent access to a service component, the MD3 one.

The values of $n_{j|k}$ and $\mu_{j|k}$ are presented in Tables 3- 4.

Table 3- Component j given application k, uplink.

Appli-cations	$n_{j\|k}$						$\bar{\tau}_{j\|k}$ [min]					
	BAS	MD1	MD2	MD3	HDV	HID	BAS	MD1	MD2	MD3	HDV	HID
HVT	0	0	0	1	0	0	-	-	-	3	-	-
IVC	1	0	0	0	0	0	30	-	-	-	-	-
MVS	0	0	0	1	0	0	-	-	-	120	-	-
HOB	0	20	0	0	1	0	-	$8.3\cdot10^{-3}$	-	-	50	-
WLI	0	34.5	0	0	0	0	-	0.055	-	-	-	-
FTP	2	0	0	0	0	0	$8.3\cdot10^{-3}$	-	-	-	-	-
PIM	1	0	0	0	0	0	10	-	-	-	-	-
DMM	15	0	0	0	0	0	0.055	-	-	-	-	-
MES	0	10	0	1	4	0	-	$8.3\cdot10^{-3}$	-	20	0.5	-
MRA	0	20	0	1	4	0	-	$8.3\cdot10^{-3}$	-	40	0.5	-
MTW	0	20	0	1	0	0	-	$8.3\cdot10^{-3}$	-	20	-	-
FFM	0	5	0	1	1	0	-	$8.3\cdot10^{-3}$	-	5	0.5	-
EMB	3	0	0	0	0	0	0.055	-	-	-	-	-
ECO	25	0	0	0	0	0	$8.3\cdot10^{-3}$	-	-	-	-	-
MML	0	20	0	0	0	0	-	$8.3\cdot10^{-3}$	-	-	-	-
TIN	0	120	0	0	0	0	-	$8.3\cdot10^{-3}$	-	-	-	-
RPC	0	5	0	0	0	0	-	$8.3\cdot10^{-3}$	-	-	-	-
UGD	0	8	0	1	0	0	-	$8.3\cdot10^{-3}$	-	5	-	-
ATR	0	30	0	1	0	0	-	$8.3\cdot10^{-3}$	-	20	-	-
TVD	0	0	0	0	0	0	-	-	-	-	-	-
ENP	5	0	0	0	0	0	$8.3\cdot10^{-3}$	-	-	-	-	-

Table 4- Component *j* given application *k*,downlink.

Appli-cations	$n_{j\|k}$						$\bar{\tau}_{j\|k}$ [min]					
	BAS	MD1	MD2	MD3	HDV	HID	BAS	MD1	MD2	MD3	HDV	HID
HVT	0	0	0	1	0	0	-	-	-	3	-	-
IVC	1	0	0	0	0	0	30	-	-	-	-	-
MVS	30	0	0	0	0	0	$8.3\cdot10^{-3}$	-	-	-	-	-
HOB	0	20	0	1	0	0	-	$8.3\cdot10^{-3}$	-	50	-	-
WLI	0	0	0	0	0	34.5	-	-	-	-	-	0.055
FTP	**1***	0	0	0	0	0	**0.333***	-	-	-	-	-
PIM	0	0	0	0	**1***	0	-	-	-	-	**10***	-
DMM	11.5	0	0	0	0	0	0.055	-	-	-	-	-
MES	0	10	0	1	4	0	-	$8.3\cdot10^{-3}$	-	20	0.5	-
MRA	0	20	0	1	4	0	-	$8.3\cdot10^{-3}$	-	40	0.5	-
MTW	0	20	0	1	0	0	-	$8.3\cdot10^{-3}$	-	20	-	-
FFM	0	5	0	1	1	0	-	$8.3\cdot10^{-3}$	-	5	0.5	-
EMB	0	0	**1***	0	0	0	-	-	**1***	-	-	-
ECO	11.5	0	0	0	0	0	0.055	-	-	-	-	-
MML	0	20	0	1	4	0	-	$8.3\cdot10^{-3}$	-	40	0.5	-
TIN	0	0	0	34.5	0	0	-	-	-	0.055	-	-
RPC	0	0	11.5	0	0	0	-	-	0.055	-	-	-
UGD	0	8	0	1	0	0	-	$8.3\cdot10^{-3}$	-	5	-	-
ATR	0	30	0	1	0	0	-	$8.3\cdot10^{-3}$	-	20	-	-
TVD	0	0	0	0	1	0	-	-	-	-	90	-
ENP	0	0	0	46	0	0	-	-	-	0.055	-	-

Their identification in bold plus an asterisk corresponds to the ABR (available bit rate) use of the respective service components. ABR applications have at least access to the amount of channels of the indicated component, using however more channels if needed and when available. Codec smoothing is assumed for video traffic.

4.3 User Model

There is a total number of *c* available channels in each cell, which are used by a number of equivalent potential users, *M*. Given a traffic mixture, one considers the model for applications activation from [12], adapted to the service components and applications of this work, Figs. 3-4.

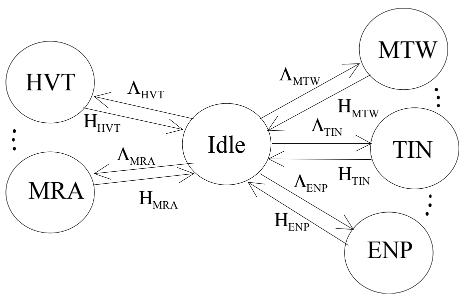

Fig. 3 Model for applications activation.

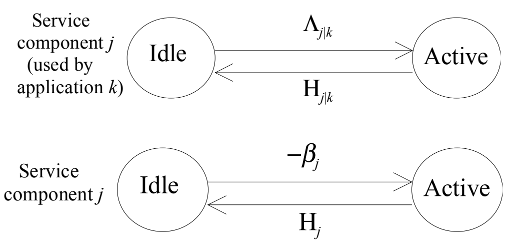

Fig. 4 Model for service components activation.

Each user can be either in an idle state or using one of the 21 applications, *k* = HVT, …, ENP, with generation rate, Λ_k, and total service rate, H_k, respectively. The total service rate is introduced in order to take into account the effect of terminal mobility in the service rate, μ_k. Once the application *k* is active, the six service components are activated with rate $\Lambda_{j|k}$ (note that $\Lambda_{j|k} = n_{j|k}\cdot H_k$) and extinguished with total service rate $H_{j|k}$, $j = 1, \ldots, J$; they can be simultaneously active, or not, and some can even not be activated for a given application.

As consequence, the activation rate of the service component *j* is given by the expectation of $\Lambda_{j|k}$, leading to

$$\left(-\beta_j\right) = \sum_{k=\mathrm{HVT}}^{\mathrm{ENP}} \Lambda_{j|k}\cdot p_k = f\cdot\sum_{k=\mathrm{HVT}}^{\mathrm{ENP}} \Lambda_{j|k}\cdot prop_k \ , \quad (6)$$

where p_k is probability of a user having an active application, *f* is the fraction of active users and $prop_k$ represents the usage of application *k* (defined as the percentage of an application use relatively to the total number of active applications), such that $p_k = f\cdot prop_k$.

If the system is stationary, the average occupancy of service component *j* is given by the following ratio

$$\left(-\beta_j^{norm}\right) = \frac{-\beta_j}{H_j} = \sum_{k=\mathrm{HVT}}^{\mathrm{ENP}} \frac{\Lambda_{j|k}}{H_{j|k}}\cdot p_k \qquad (7)$$

here called normalised activation rate, meaning that the service component *j* service rate is

$$H_j = \sum_{k=\mathrm{HVT}}^{\mathrm{ENP}} \Lambda_{j|k}\cdot\frac{\Lambda_k}{H_k} \Bigg/ \left(\sum_{k=\mathrm{HVT}}^{\mathrm{ENP}} \frac{\Lambda_{j|k}}{H_{j|k}}\cdot\frac{\Lambda_k}{H_k}\right) \qquad (8)$$

This does depend on mobility, because of the dependence of the numerator on it. If terminal mobility is considered, following an approach similar to the one for the arrival and total service rates in a single service system [8], the original value of $\Lambda_{j|k}$ has to be replaced by $\Lambda_{j|k}$ times a factor $(\mu_k+\eta_k)/\mu_k$, where μ_k and η_k are the service and the cross-over rates associated with application *k*.

This is a loss system, whose performance can be measured by the blocking probability of each service component, simplifying the analysis (because one only needs to consider the service components, and not each application).

The system average load is [12]

$$L = f\cdot b_1\cdot M \qquad (6)$$

where b_1 is the maximum load per user [12], it being given by

$$b_1 = \sum_{k=\mathrm{HVT}}^{\mathrm{ENP}} prop_k\cdot b_k \qquad (7)$$

where the values of $prop_k$ for the BCC, URB and ROA are the ones from [11]. The maximum load per user allows defining the asymmetry factor, A_f, the ratio of the maximum loads per user between the down- and uplinks, Table 5.

Table 5 – Asymmetry factor for the various scenarios.

	BCC		URB		ROA	
	Uplink	Downlink	Uplink	Downlink	Uplink	Downlink
b_1	657	1 868	597	1 678	801	1 734
A_f	2.688		2.643		2.095	
$b_{1\text{-}UP}+b_{1\text{-}DOWN}$	2 525		2 275		2 535	

Given the supported f, the number of supported users in a cell is given by $N_{SU} = f{\cdot}M$. The spectral efficiency is defined by

$$S_{ef} = \frac{f \cdot M \cdot b_1}{c_{LINK} \cdot UBR_{slot}} \quad (8)$$

where c_{LINK} is the number of available channels for the respective link (up- or downlink) and UBR_{slot} is the user data rate directly associated with a slot, the basic channel considered here, i.e., 384 kb/s.

4.4 Channel Usage

ACTS-SAMBA proposals for a future system [2] state that, in several aspects, the foreseen characteristics of the air interface for an early stage of MBS implementation are similar to those adopted in the SAMBA Trial Platform: the same paired band, in the vicinity of 40 GHz, an OQPSK-type modulation scheme and a TDMA/FDMA scheme (Time/Frequency Division Multiple Access). It allows a wide range of service bit rates through a flexible use of the time slots concerning each available carrier frequency with an appropriate MAC protocol. Although a fixed allocation strategy is considered in our analysis, it is being based on the stochastic characterisation of service components, partly reflecting the dynamics associated with the duration of each connection. Thus, this characterisation is very useful in order to achieve some initial results for the upper bounds of blocking probability (and even for the handover failure probability).

Considering gross bit rates (*GBR*) of 40, 80 and 160 Mb/s, and the respective maximum values for cell coverage distance, R_{max}, Table 6, one obtained the user bit rate per carrier ($UBR_{carrier}$) for values of the slot duration T = 20.83, 10.42 and 5.21 μs, respectively. This is because, from the 105 bytes in each slot only 48 bytes correspond to the user (net) bit rate [2].

Table 6 – User bit rate associated to a carrier.

Case	*GBR* [Mb/s]	R_{max} [m]	T [μs]	$UBR_{carrier}$ [kb/s]	$N_{C\text{-}GHz}$
A	40	350	20.83	18 432	36
B	80	180	10.42	36 864	18
C	160	60	5.21	73 728	9

The number of carriers per GHz, $N_{C\text{-}GHz}$, was obtained, too, considering $\Delta f/R_S$ = 1.3883, where Δf is the carrier spacing and R_S is the adopted gross symbol rate (numerically equal to the *GBR*).

The user bit rate per slot, UBR_{slot} is obtained dividing $UBR_{carrier}$ by the number of slots per frame (as each slot is re-used in the next frame, i.e., only one frame later). If one considers 48, 96 and 192 slots per frame, for the respective *GBR*, one obtains UBR_{slot} = 384 kb/s in every case, it being the basic channel data rate one is considering in this paper.

In MBS one has a total bandwidth of 2 GHz available at the 40 and at the 60 GHz bands [13]. Thus, in case A (*GBR* = 40 Mb/s) one has 72 carriers in the system, each corresponding to a user bit rate of 18 432 kb/s, while in case B (*GBR* = 80 Mb/s) one only has 36 carriers available in the system, but each corresponding to a user bit rate of 36 864 kb/s.

Because the highest service component data rate is 31 872 kb/s, one needs to figure out how it is possible to achieve it. If only one carrier was available for each user then one only would have one choice, i.e., case B. However, if the system could provide more than one carrier per user, case A also could be chosen. This option is convenient, because in this case it is much easier to distribute an integer number of carriers among the cells, for different values of the reuse factor K (the number of different groups of frequency carriers). Examples are the following:

- For four operators in the system, with K = 2 there are 9 carriers per cell, while with K = 3 there are 6 carriers per cell
- For three operators in the system, there are 12, 8 and 6 carriers per cell for K = 2, 3 and 4, respectively.

Therefore, one considers case A, which implies assuming that more than one carrier can be available per user.

5 Results

The total amount of channels in each cell is a fraction of the total number of channels, depending on the number of operators and on K [14]. In this paper, one considers the cases of 384 and 288 resources per cell (the former corresponding to 3 operators and K = 3, and the latter corresponding either to 3 operators and K = 4 or 4 operators and K = 3).

Furthermore, one considers, as a simplification, that the four-leaf cell is suitable for coverage in the BCC and the URB scenarios, whereas the cigar-shaped cell is suitable for ROA scenario coverage, Figs. 5-6, where R is the coverage distance and l is the street length. One assumes the following values (corresponding to a net cell area of 8316 m^2):

- R = 100 m and l = 22 m for the BCC and the URB scenarios
- R = 150 m and l = 28 m for the ROA scenario,

The values of R are clearly in the range defined in Table 6.

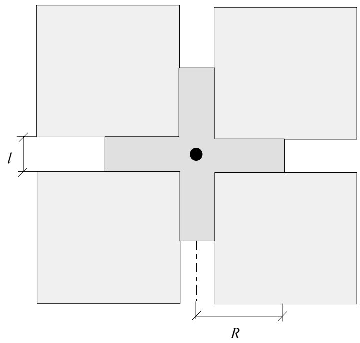

Fig. 5 Four-leaf cell.

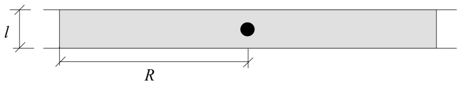

Fig. 6 Cigar-shaped cell.

One considered, as an approximation, that the number of potential users for the considered environment is M_{ENV} = 250 user per cell in the BCC scenario and M_{ENV} = 100 user per cell in the URB and the ROA scenarios. Taking the density factors from [11] into account, one obtained the number of resources in each link, c_{LINK}, Table 7. When the resources are all occupied, the number of users is obtained directly from the load by $M_{max\_load} = L_{max}/b_1$, Table 7.

Table 7 –Assumptions for M_{max_load} and c_{LINK}.

		M_{max_load}		c_{LINK}	
No. channel/cell		384	288	384	288
BCC	UP	55	41	88	66
	DOWN	55	41	296	222
URB	UP	60	45	89	67
	DOWN	61	46	295	221
ROA	UP	55	41	99	70
	DOWN	60	45	285	218

Results for $P_b(f)$ were obtained in the various scenarios for the up- and downlinks for $M = M_{ENV}$ and $M = M_{max\_load}$, Fig. 7 (example for the URB scenario, considering 384 channels/cell). In order to balance the results for the supported f between links, the resources were distributed between them with slight changes relatively to the ones defined by the asymmetry factor (for P_b = 2%), Table 7.

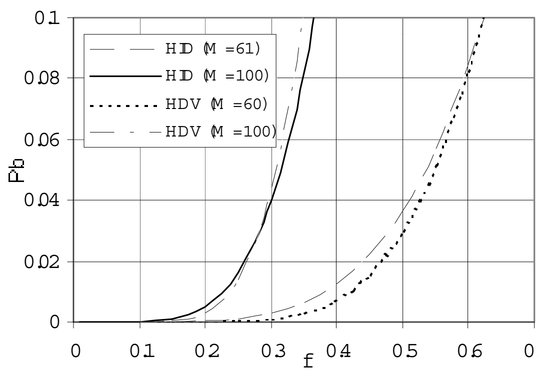

Fig. 7 $P_b(f)$ for the URB scenario, 384 channel/cell.

As the considered mixtures of applications do not use the HID service component at the uplink, it only limits the system at the downlink, whereas the HDV service component limits the uplink. The relevant differences in these curves come from different values of M. Between them, the values of the product $f{\cdot}M$ are similar, they being however slightly higher for lower values of M. For P_b = 2%, the supported f in this scenario is only 26 %, it being however approximately 44 % of the maximum load, as it can be seen from the curves for $M = M_{max\_load}$ (the values from M_{max_load} for the URB scenario were extracted from Table 7).

Although the results for the fraction of active users are quite interesting, the comparison among the three different scenarios is difficult, as the number of potential users in a cell is different (although the cells have the same net area).

So, alternatively one presents results for the number of supported users in a cell and for the spectral efficiency, Tables 8 and 9. Worst-case values of the supported f between the up- and downlinks are considered to compute the latter.

Table 8 – Number of supported users in a cell.

N_{SU}		$M = M_{ENV}$		$M = M_{max\_load}$	
No. channels/cell		384	288	384	288
BCC	UP	21.2	11.7	23.8	13.0
	DOWN	20.5	10.2	21.8	10.6
URB	UP	26.4	15.0	28.2	16.0
	DOWN	26.1	13.4	26.9	13.9
ROA	UP	22.6	11.2	24.4	12.1
	DOWN	21.1	10.6	21.7	10.9

Table 9 – Spectral efficiency in the worst cases.

Worst case S_{ef}		$M = M_{ENV}$		$M = M_{max\_load}$	
No. channels/cell		384	288	384	288
BCC	UP	0.44	0.29	0.46	0.30
	DOWN	0.35	0.23	0.37	0.24
URB	UP	0.50	0.34	0.52	0.36
	DOWN	0.40	0.28	0.41	0.28
ROA	UP	0.47	0.34	0.49	0.35
	DOWN	0.35	0.23	0.36	0.23

These type of results can be fed into the process of the optimisation of MBS cellular planning. As these parameters vary with the coverage distance, it is important to refer the differences, which basically come from a different value of the potential number of active users per cell. This change in M is however very easy to obtain because it varies linearly with R. In Fig. 8 one presents an example for 288 resource/cell in the BCC scenario, from which one concludes that the balance between links is maintained as R varies.

Although one does not present detailed results for the influence of mobility in the results, one explores the differences, considering that applications mobility is defined according to their deployment scenarios [11].

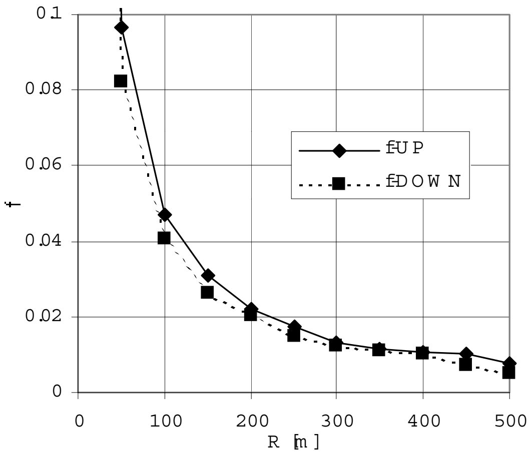

Fig. 8 Supported *f* as a function of *R* for the BCC scenario, 288 channel/cell.

While in the BCC scenario mobility has no influence, it has important consequences in the URB and ROA scenarios. Only 81-82 % of the traffic is supported in the URB scenario relatively to the case of absence of mobility. In the ROA scenario the influence is more drastic: only 28 % of the traffic is supported in the 384 channels/cell case relatively to the case of absence of mobility, and 21 % in the 288 resource/ cell case.

6 Conclusions

An MMPP model was proposed for the modelling of the superimposition of MBS data and video IPP sources. Given the correspondence between applications and their bearer service components, an algorithm for the Bernoulli case of the Bernoulli-Poisson-Pascal model was used to compute the blocking probability. One concludes that for the scenario where the maximum load per user is lower, i.e., the URB one, the number of supported users is higher. Comparing the cases of 288 and 384 channel/cell one also concludes that, although the number of channels is only 25 % lower, the supported number of users decreases 50 %. This problem could be solved by decreasing the data rate of the basic channel, i.e., the slot, leading to a higher number of slots in a cell, thus, to a different channel granularity. High terminal mobility strongly degrades system performance.

7 References

[1] L. Fernandes, "Developing a System Concept and Technologies for Mobile Broadband Communications," *IEEE Personal Communications Magazine*, Vol.2, No.1, Feb. 1995, pp. 54-59.

[2] M. Prögler and S. Svaet (eds.), *MBS Performance Evaluation*, ACTS-SAMBA Deliverable A0204/TN/PK/DS/P/014/b1, ACTS Central Office, Brussels, Belgium, 1999.

[3] H. Saito, *Teletraffic Technologies in ATM Networks*, Artech House, Boston, MA, USA, 1994.

[4] G. Anastasi, L. Lenzini, E. Mingozzi, A. Hettich and A. Krämling, "MAC Protocols for Wideband Wireless Local Access: Evolution Towards Wireless ATM," *IEEE Personal Communications Magazine*, Vol. 5, No. 5, Oct. 1998, pp. 53-64.

[5] A. Krämling, M. Scheibenbogen and T. Lohmar, "Dynamic Channel Allocation in Wireless ATM Networks," in *Proc. of ICT'98 – International Conference on Telecommunications*, Porto Carras, Greece, Jun. 1998.

[6] V.S. Frost and B. Melamed, "Traffic Modelling for Telecommunications Networks," *IEEE Communications Magazine*, Vol. 32, No. 3, Mar. 1994, pp. 82-89.

[7] J.P. Rebelo, *Approximation of ON-OFF Processes via MMPP Processes for Average Delay Analysis in an ATM Queue* (in Portuguese), MSc Thesis, Instituto Superior Técnico, Technical University of Lisbon, Lisbon, Portugal, Dec. 1996.

[8] B. Jabbari, "Teletraffic Aspects of Evolving and Next-generation Wireless Communication Networks," *IEEE Personal Communications Magazine*, Vol. 3, No. 6, Dec. 1996, pp. 4-9.

[9] G.A. Awater and H.A. van de Vlag, "Exact Computation of Time and Call Blocking Probabilities in Large, Multi-traffic, Multi-resource Loss Systems," *Performance Evaluation,* Vol. 25, No. 1, Mar. 1996, pp. 41-58.

[10] F.J. Velez and L.M. Correia, "Classification and Characterisation of Mobile Broadband Services," in *Proc. of VTC'2000 Fall-IEEE Vehicular Technology Conference*, Boston, MA, USA, Sep. 2000.

[11] F.J. Velez and L.M. Correia, "Deployment Scenarios for Mobile Broadband Communications," in *Proc. of PIMRC' 2000-11th IEEE International Symposium on Personal, Indoor and Mobile Radio Communication*, London, UK, Sep. 2000.

[12] F.J. Velez, "Multi-service Traffic Analysis in Mobile Broadband Systems," in *Proc. of 4th ACTS Mobile Communications Summit*, Sorrento, Italy, Jun. 1999.

[13] F. Velez, L. M. Correia, "Capacity Trade-offs in Mobile Broadband Systems using Guard Channels for High Mobility Handover," in *Proc. of PIMRC' 98 - 9th IEEE International Symposium on Personal Indoor, and Mobile Communications*, Boston, Massachusetts, USA, Sep. 1998

[14] F.J. Velez and L.M. Correia, "Optimization Criteria for Cellular Planning of Mobile Broadband Systems in Linear and Urban Coverages," in *Proc. of 2nd ACTS Mobile Communications Summit*, Aalborg, Denmark, Oct. 1997.