A Markovian Model for ATM Traffic Generation

Joaquim E. Neves*, and Mário J. Leitão**

* DEI, Universidade do Minho, Azurém, P-4800 Guimarães
* DEEC-FEUP, Universidade do Porto, Rua dos Bragas, P-4000 Porto
INESC - Instituto de Engenharia de Sistemas e Commputadores
Largo Mompilher 22, P-4000 Porto
Phone: + 351 2 2094000 Fax: + 351 2 318692
Email: eneves@inescn.pt, mleitao@inescn.pt

Abstract

Many different traffic sources are transmitted simultaneously using the Asynchronous Transfer Mode (ATM) over the Broadband Integrated Services Digital Network (B-ISDN). Because of the ATM traffic statistical characteristics, the resource allocation in B-ISDN must be related to traffic parameters and the quality of service negotiated in the establishment of each call. A model for ATM traffic characterisation and generation control of the calls and cells is proposed. Traffic parameters calculated analytically and by simulation are compared.

1 Introduction

The International Consulting Committee for Telephones and Telegraphics (CCITT) has published recently a set of recommendations where the Broadband Integrated Services Digital Networks (B-ISDN) are defined [1], [2]. The Asynchronous Transfer Mode (ATM) provides the transport and the switching of B-ISDN services in fixed size data packets called cells.

The transport of B-ISDN services, like interactive video telephony, high quality video and audio broadcast programs or data file transfer, requires a call establishment phase and the adaptation of the information flow in ATM cells [2]. A call is a concatenation of Virtual Path Connections (VPC) and Virtual Channel Connections (VCC) within which the cells are transmitted [3].

ATM cells can be transported in the existent Plesiochronous (PDH) or Synchronous (SDH) Digital Hierarchies or in the new cell-based transmission systems [4], [5].

At the call establishment, the user has to negotiate with B-ISDN control entity the traffic characteristics of the call and the quality of service requested. The network control entity can accept the request and allocate network resources for the support of the service, or propose a lower quality of service, and in the limit, reject the call, if not enough network resources are available.

This paper presents a model for ATM traffic generation suitable for network simulation and resource allocation purposes. In the next section the traffic of ATM services will be characterised at the user network interface. Section 3 describes the simulation of traffic in a ATM network node. Results of simulation are presented and compared with the traffic parameters calculated analytically. Section 4 summarises the main results of this paper.

2 ATM Traffic Model

At call establishment the B-ISDN user and the network control entity negotiate the traffic source characteristics and the required quality of service. B-ISDN has mechanisms for usage parameter control by policing the call traffic at user interface and taking appropriate actions if the usage values of the information flow parameters are exceeded in a virtual channel or virtual path. The traffic parameters that can be negotiated and policed by the network may be the peak and the average bit rate, the peak duration and the burstiness [3].

Most of ATM traffic models define the generation of ATM traffic by bursty sources [6], [7], [8], [9] where the parameters that characterise the sources are the same which are policed by the network. Those models only cope with some particular types of sources or the highly multiplexed traffic in the network node interface. The model proposed here has capabilities to characterise generally any single ATM traffic source or mixed sources.

2.1 Description of the proposed model

The proposed ATM traffic model used in the ATM traffic generation at B-ISDN user network interface is defined by the following three functional levels:

Generation level
Synchronisation level
Adaptation level

The relationship between each level of the model is identified in figure 1. In the generation level any traffic source can be defined by a Markovian state space and associated timing relations of the information flow events. The following parameters will be used within the generation level:

Number of states (N);
State transition probability matrix ($P_{ij}, 0 < i \leq N, 0 < j \leq N$);
Quantum of the duration of each state ($D_i, 0 < i \leq N$);
Probability Distribution Function (PDF) *of the time between events in each state, specifying the first* ($T_i, 0 < i \leq N$) *and other moments;*
Probability Distribution Function (PDF) *of the event duration in each state, specifying the first* ($C_i, 0 < i \leq N$) *and other moments.*

Since in B-ISDN, the information flow events can be connection requests or generation of ATM cells, each ATM traffic source will be actually defined by two Markovian state spaces. As it will be shown later, the number of expected calls and the average cell rate can be easily calculated in the generation level of the proposed model.

The synchronisation level incorporates the timing characteristics of the environment, which includes the discrete nature of the generation entities. The synchronisation level is characterised by two parameters:

The time unit (TU);
Maximum number of events per time unit (NE).

The adaptation level acts as a finite length buffer for the events provided by the generation level. Two parameters characterise this level:

FIFO length (FL);
Minimum time between events (MT).

The adaptation level performs the low pass filter functions for the events, to warrantee at the B-ISDN user network interface the peak cell rate and consequently the burstiness negotiated at the establishment of the call.

2.2 Generation level

The discipline imposed by the state transition probability matrix determines the time evolution of staying or leaving a traffic state. The quantum duration of each state and the probability distribution function of the time between events are the specific parameters of each traffic state. Figure 2 illustrates the generation of events in the generation level of the proposed traffic model when the PDF of the time between events in each state is characterised only by the first moment (exponential distribution, for example). The probability of residence in each state of a Markovian process can be expressed by:

$$\overrightarrow{\pi} = \overrightarrow{1}.[\overrightarrow{U} + \overrightarrow{P} - \overrightarrow{I}]^{-1} \tag{1}$$

where $\overrightarrow{P}$ is the state transition probability matrix, $\overrightarrow{I}$ is the identity matrix, $\overrightarrow{1}$ is a unitary vector, and $\overrightarrow{U} = \overrightarrow{1^t}.\overrightarrow{1}$. ($\overrightarrow{1^t}$ denotes the transposed of $\overrightarrow{1}$).

The quantum of the duration of each state is a deterministic parameter, but in general the time that the traffic source stays in each state is stochastic with a geometrical distribution. In case the probability of staying in the same state is zero the duration of this state is equal to the quantum of the duration.

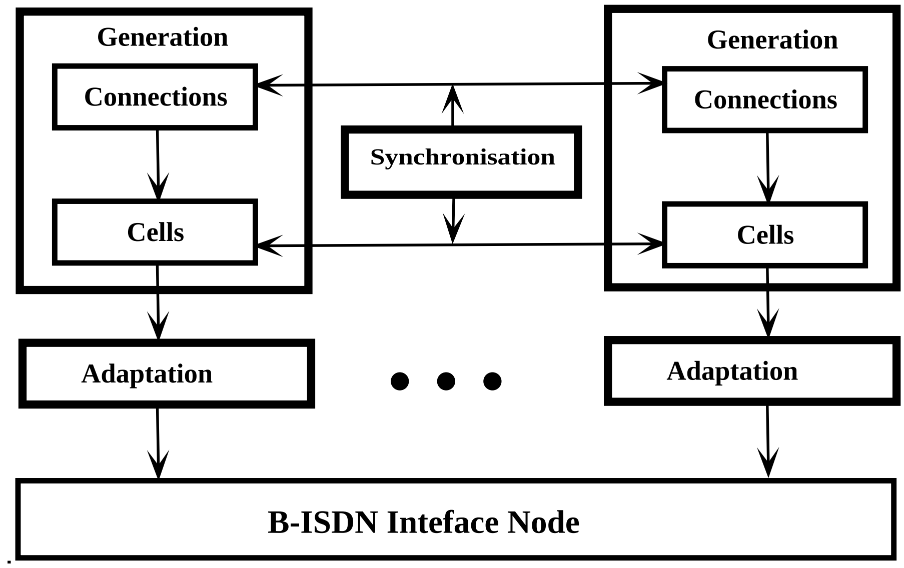

Figure 1: The traffic model

Since $(1 - P_{ii})$ is the probability of the generation process leaving a state i, the average time $\overline{D}_i$ within this state is given by

$$\overline{D}_i = \frac{D_i}{1 - P_{ii}} \tag{2}$$

The average number of events E_i generated within the quantum duration of a traffic state i, is the ratio of the quantum duration D_i and the average time between events T_i:

$$E_i = \frac{D_i}{T_i} \tag{3}$$

The number of events generated during the average time $\overline{D}_i$ within a state i of traffic activity has geometrical distribution with average $\overline{E}_i$ given by:

$$\overline{E}_i = \frac{E_i}{1 - P_{ii}} = \frac{\overline{D}_i}{T_i} = \frac{D_i}{1 - P_{ii}}.\frac{1}{T_i} \tag{4}$$

After calculating the probability of residence π_i, the mean time between event generation $\overline{T}$ can be derived from:

$$\overline{T} = \frac{\sum_{i=1}^{N} D_i.\pi_i}{\sum_{i=1}^{N} \frac{D_i.\pi_i}{T_i}} = \frac{\sum_{i=1}^{N} D_i.\pi_i}{\sum_{i=1}^{N} E_i.\pi_i} \tag{5}$$

The mean duration of events $\overline{C}$ is given by:

$$\overline{C} = \frac{\sum_{i=1}^{N} C_i.D_i.\pi_i}{\sum_{i=1}^{N} D_i.\pi_i} \tag{6}$$

2.2.1 Average Number of Calls

Independently of the call duration distribution, when the traffic state duration is much higher than the mean duration of the calls C_i, the average number of active calls A_i at the end of state i, is given by:

$$A_i = \frac{C_i}{T_i} \tag{7}$$

On the other hand, the mean call duration is much higher than the traffic state duration, the average number of calls A_i generated in each traffic state i and still active at the end of the state is equal to the number of calls generated in that sate, which can be derived from (3):

$$A_i = E_i = \frac{D_i}{T_i} \tag{8}$$

After calculating the probability of residence π_i and the average number of active calls A_i generated in each state i, the average number of active calls $\overline{A}$, is equal to:

$$\overline{A} = \overline{C}.\frac{\sum_{i=1}^{N} A_i.\pi_i}{\sum_{i=1}^{N} D_i.\pi_i} \tag{9}$$

However, if the generation time t is much smaller than the mean call duration, $\overline{A}$ is equal to:

$$\overline{A} = t.\frac{\sum_{i=1}^{N} A_i.\pi_i}{\sum_{i=1}^{N} D_i.\pi_i}, \tag{10}$$

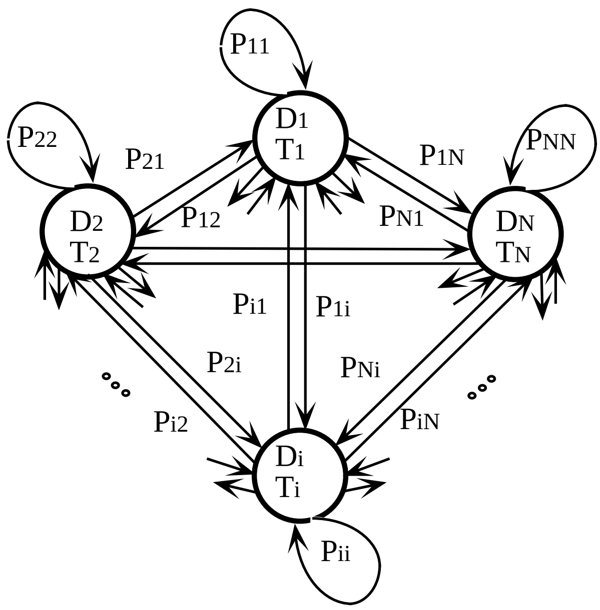

Figure 2: Traffic generation model

2.2.2 Average and Burst Cell Rate

In the cell model, the events are cells of equal length, so their duration $\overline{C}$ is fixed. ($\overline{C} = C_i,\ \forall i \leq N$).

The average cell rate λ observed during a time window t_W can be defined as the ratio of the number of cells generated $\overline{E}_W$ over the time t_w

$$\lambda = \frac{\overline{E}_W}{t_W} \tag{11}$$

When the time window is long compared within the traffic state durations, the average cell rate λ is the inverse of the mean time between event generation of the cell generation model:

$$\lambda = \frac{\overline{E}_W}{t_W} = \frac{1}{\overline{\overline{T}}} = \frac{\sum_{i=1}^{N} E_i.\pi_i}{\sum_{i=1}^{N} D_i.\pi_i} = \frac{\sum_{i=1}^{N} \frac{D_i.\pi_i}{T_i}}{\sum_{i=1}^{N} D_i.\pi_i} \tag{12}$$

If the time window is of the same order of magnitude of the traffic state quantum durations, the evaluation of the average number of cells should take into account the state transition defined by the state transition probability matrix.

The burst cell rate λ_B can be defined as the higher average number of cells $\overline{E}_B^{Max}$ generated over the burst duration t_B:

$$\lambda_B = \frac{\overline{E}_B^{Max}}{t_B} \tag{13}$$

If the burst duration is shorter than the shortest traffic state quantum duration, the burst cell rate is the inverse of the minimum mean time between events in each state, of the cell generation model.

$$\lambda_B = \frac{\overline{E}_B^{Max}}{t_B} = \frac{1}{Min(T_i)} \tag{14}$$

If the burst duration t_B is of the same order of magnitude of the traffic state durations, $\overline{E}_B^{Max}$ has to be evaluated in every state combination γ_l that have state transition probability different from zero.

$$\lambda_B = \frac{\overline{E}_B^{Max}}{t_B} = \frac{Max(\sum_{i^k \in \gamma_l} E_{i^k})}{t_B} \tag{15}$$

where

$$\gamma_l = [i^1, i^2, ..., i^K] : \sum_{k=1}^{K-1} D_{i^k} \leq t_B \wedge P_{i^k i^{k+1}} \tag{16}$$

We note that the burst cell rate is not the same as the peak cell rate defined by CCITT. The calculation of the peak is quite difficult within the generation lelvel, but it can be determined within the adaptation level.

2.3 Synchronisation level

The synchronisation level defines the time unit TU and the maximum number of events NE that can be generated in each unit. As a consequence of this discretization introduced by the synchronization level, the following singularities are highlighted:

T_i mean time between events

$T_i = 0$ *- the number of events generated every time unit UT is equal to NE, independently of the PDF;*

$T_i = \infty$ *- no cells are generated in the state i (silence state);*

D_i quantum of the duration

$D_i = 0$ *the process stays at least one time unit in the state (actually one if $P_{ii} = 0$.*

The process is deterministic if T_i, $\forall i$ is either zero or infinity and P_{ij}, $\forall ij$ is either zero or unity. This implies that the traffic model although it is generally stochastic, it can generate deterministic traffic or deterministic components of stochastic traffic. For example the model can be used to control the generation of constant bit rate traffic if the mean time between events (cells) in each state is zero (in the activity states) and infinity (in the silence states) and the probability of leaving to another specific state is unity. The state duration and the number of states can control the bit rate and the pattern of the cell stream.

2.4 Adaptation level

Cells produced by the generation level are sent to a FIFO buffer in the adaptation level and are read out with a specific minimum time MT between two consecutive cells. This allows to impose a desired peak cell rate, defined as the inverse of MT. The FIFO length FL as to be deep enough to accommodate the expected cell rate fluctuation in the cell stream provided the generation level. This process should only affect the tail of the PDF, otherwise the parameters of the generation level cease to apply.

Similar approach is used for the call generation model.

3 Traffic Simulation

The proposed ATM traffic model was used to describe and generate the traffic, in a B-ISDN access interface. The call generation associated to different services and users is a Markovian process with different state durations. Within each state, call duration and time between call births are defined by exponential distribution functions. The cells of each call are also generated by this traffic model with the appropriate parameters for each service.

3.1 Simulation of calls generation

Four call generation cases of the presented traffic model have been considered in Table 1.

Each line of Table 1 specifies the generation parameters and the average number of active calls evaluated analytically for the steady state of each case.

In the first case calls with mean duration time of 1, 20, 60 and 180 minutes have a generation rate of 1/20 call/s and generation intervals with exponential distribution.

The second case differs from the first because the call generation is interrupted during 10 minutes (silence state, note the large value of the mean time between calls) after every 10 minutes of activity state. The call generation rate in the activity state is also 1/20 call/s.

The third case has 3 states with quantum duration of 10 minutes. The duration of the most active state is deterministic ($P_{ii} = 0$) but the duration of the other states have geometrical distribution. Once the call generation reaches the most active state the call generation rate is 1/20 call/s during 10 minutes, and then leaves with equal probability to the less active or the inactive state. In the less active state the call generation rate is 60 call/s during 10 minutes, and then the call generation process can stay in the same state with probability of 30% or can leave with probability of 70% to the inactive state. In the inactive state no calls are generated and every 10 minutes the state transition can be to the same state with probability of 30% or to the most active state with probability of 70%.

The fourth case differs from the second because it has an initial state with deterministic duration of 0.1 minutes where NE calls are generated every unit time, and the transition time between states is now 60 minutes instead of 10 minutes. After leaving this initial state, the call generation alternates between one active and one inactive state like in the second case.

Table 1: Parameters of call generation control

Call Generation						
Case	**Number of States**	**State Transition Probability**	**State Duration (Minutes)**	**Mean Time Between Calls (Seconds)**	**Mean Call Duration (Minutes)**	**Average Number of Calls**
1	1	1.0	60.0	20.0	1.0 20.0 60.0 180.0	3.0 60.0 180.0 360.0
2	2	0.0 1.0 1.0 0.0	10.0 10.0	20.0 1000000.0	1.0 20.0 60.0 180.0	1.5 30.0 90.0 180.0
3	3	0.0 0.5 0.5 0.0 0.3 0.7 0.7 0.0 0.3	10.0 10.0 10.0	20.0 60.0 1000000.0	1.0 20.0 60.0 180.0	1.2 23.6 70.9 141.8
4	3	0.0 1.0 0.0 0.0 0.0 1.0 0.0 1.0 0.0	0.1 60.0 60.0	0.0 20.0 1000000.0	1.0 20.0 60.0 180.0	1.5 30.0 90.0 180.0

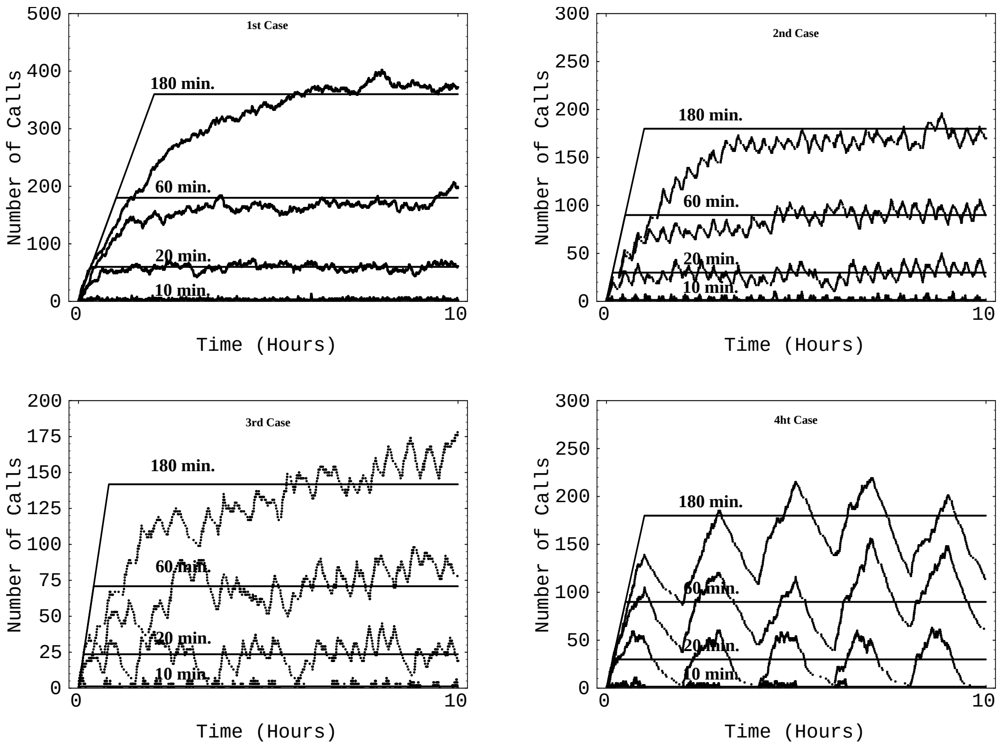

Figure 3: Number of calls during 10 hours of simulation for the cases characterized in table 1 (mean call duration as a parameter).

Table 2: Parameters of cell generation control from different services

Cell Generation					
Service	Number of States	State Transition Probability	State Duration (Seconds)	Mean Time Between Cells (Seconds)	Average Cell Rate (kcell/s)
1	1	1.0	0.10000	0.00003	33.00
2	2	0.2 0.8 0.8 0.2	0.00100 0.00200	0.00003 600.00000	11.00
3	3	0.0 0.0 1.0 0.5 0.0 0.5 0.0 0.8 0.2	0.00010 0.00010 0.00010	0.00003 0.00050 600.00000	6.90
4	4	0.0 1.0 0.0 0.0 0.0 0.0 1.0 0.0 0.0 0.0 0.2 0.8 0.8 0.0 0.0 0.2	0.00000 0.00010 0.00000 0.00010	0.00000 600.00000 0.00000 600.00000	9.70

Call generation has been simulated over 10 hours, with a simulation time unit $TU = 0.1$ s and $MT = NE = 1$. The number of calls during the simulation time is plotted in figure 3. The straight lines represent the asymptotes of the expected average number of calls calculated by equations (10) and (9).

Good agreement is observed between the simulated and calculated values.

3.2 Simulation of cells generation

Four ATM services are characterised by the traffic model presented in Table 2. Each line of Table 2 specifies the cell generation parameters of each service and the average cell rate, which was evaluated analytically.

The first service has a cell rate of 1/0.00003 cell/s and generation instants with exponential distribution.

The second service has two states with duration geometrically distributed. When the service leaves the activity state, the cell generation is interrupted ($600 \gg 0.002$) during a time multiple of 0.002 seconds (silence state). The cell rate in the activity state is also 1/0.00003 cell/s and the duration is a multiple of 0.002 seconds. The probability of staying or leaving each state is 20% and 80% respectively.

The third service has 3 states with quantum duration of 0.0001 seconds. The duration of the most active state is deterministic and the duration of the less active and the inactive states have a geometrical distribution. Once the service reaches the most active state the cell rate is 1/0.00003 cell/s during 0.0001 seconds, and then leaves to the inactive state. In the less active state the cell rate is 1/0.0005 cell/s during 0.0001 seconds, and then the service can leave with equal probability to the most active or the inactive state. In the inactive state no cells are generated and every 0.0001 seconds the service can stay in the inactive state with probability of 20% or can leave with probability of 80% to the less active state.

The fourth service has two states of activity where a cell is generated every time unit the service stays in those states and two silence states where no cells are generated.

The quantum duration for the activity states is zero and for the silence states is 0.0001 seconds. The service stays in the first activity state only a unit of simulation time, it generates NE cells and leaves to the first silence state. After 0.0001 seconds without generating any cell, the service leaves this silence state to the second active state. In that activity state the service generates NE cells every simulation time unit.

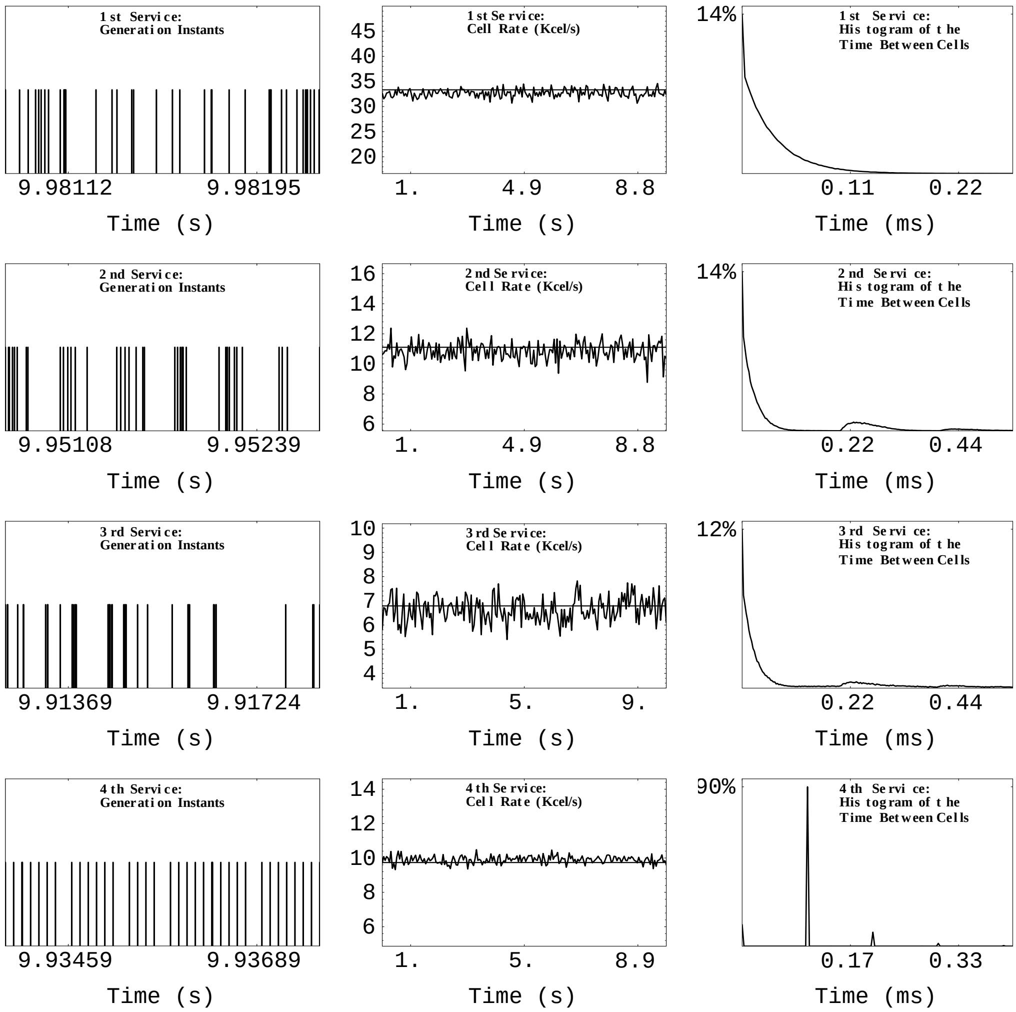

Figure 4: Cell generation from the services characterized in Table 2, during 10 seconds of simulation.

The service can stay in the same state with probability of 20% or leave to the second inactive state with probability of 80%. In the second inactive state every 0.0001 second, the service can stay in the same state with probability of 20% or leave to the first active state with probability of 80%.

The generation instants, the cell rate and the histogram of the time between cells, provided by one call of each ATM service characterized in Table 2, are plotted in figure 4.

The cell rate is calculated within a 50 ms window, during 10 seconds of simulation time, with a simulation unit $TU = 2.7\mu$s and $MT = NE = 1$.

The straight lines represent the cell rate expected analytically, by equation (12). Again good agreement is observed between the simulated and calculated values. The figures also exhibits the characteristics of the states and changes between them previously highlighted: the 1st services is a pure exponentially distributed source; in the 2nd and 3rd services the exponentially distributed generation process is modulated by a Markovian activity/silence process; the 4th service is a constante bit rate source with the cell generation intervals lightly modulated by a geometrical distribution.

4 Summary

The proposed ATM traffic model can be used in traffic generation at any B-ISDN interface and is defined by three functional levels. In the generation level, a traffic source is specified by a Markovian state space and associated timing relations of the information flow events. The synchronization level incorporates the timing characteristics of the environment. The adaptation level performs the low pass filter functions for the cell stream, to warrantee at the B-ISDN user network interface the peak cell rate. The generation level defines a set of timing relations of the information flow variations for any traffic source. The discipline imposed by the state transition probability matrix determines the time evolution of staying or leaving a traffic state. The quantum duration of each state and the probability distribution functions of the event duration and the time between events are the specific parameters of each traffic state. The quantum of the duration of each state is a deterministic parameter, but in general the time that the traffic source stays in each state is stochastic with a geometrical distribution. It is shown that, independently of the probability distribution function of each traffic state, the number of expected call connections and the average cell rate can be easily calculated in the generation level of the proposed model. Simulations support the adequacy of the model to simulate broadband ATM traffic.

References

[1] ., "Broadband aspects of ISDN," in *Draft Recommendation I.121*, Geneva CCITT, May 1990.

[2] .,"B-ISDN service aspects," in *Draft Recommendation I.211*, Geneva: CCITT, May 1990.

[3] ., "B-ISDN general network aspects," in *Draft Recommendation I.311*, Geneva: CCITT, May 1990.

[4] ., "DETR transport of ATM cells over various transmission systems including cell based systems," Report on STC meeting held in Göreme 9-13, ETSI/STC-NA5, Sept. 1991.

[5] ., "Draft TR evolution toward B-ISDN," Report on stc meeting held in Lisbon 23-27, ETSI/STC-NA5, Mar. 1992.

[6] H. Yamada e S. Sumita, "Trafic Measurement and its Application for Cell Loss Probability Estimation in ATM Networks", *A IEEE Journal on Selected Areas in Communications*, vol. 9, number 3, April 91.

[7] E. P. Rathgeb, "Modeling and Performance Comparison of policing Mechanisms for ATM Networks", *IEEE Journal on Selected Areas in Communications*, vol. 9, number 3, April 91.

[8] R. Grunenfelder, J. P. Cosmas, S. Manthorpe, e A. Odinma-Okafor, "Characterization of Video Codes as Autoregressive Moving Average Processes and Related Queueing System Performance", *IEEE Journal on Selected Areas in Communications*, vol. 9, number 3, April 91.

[9] R. M. Rodrigues-Dagnimo, M. R. Khansari, e Leon-Garcia,"Prediction of Bit Rate Sequences of Encoded Video Signals", *IEEE Journal on Selected Areas in Communications*, vol. 9, number 3, April 91.