18

Diffusion Models to Study Nonstationary Traffic and Cell Loss in ATM Networks

T. Czachórski[a], J. M. Fourneau[b] and F. Pekergin[c]

[a]Polish Academy of Sciences, Computer Science Department, IITiS PAN, 44-100 Gliwice, ul. Bałtycka 5, Poland

[b]Laboratoire PRISM, Université de Versailles-Saint Quentin, 45, Avenue des Etats-Unis, 78035 Versailles Cedex, France

[c]Laboratoire LIPN, Université de Paris-Nord, Avenue J.B.Clement, 93430 Villetaneuse, France

Abstract

We study the effects of nonstationary traffic patterns in a network of ATM nodes. Dynamic behaviour of ATM networks is of interest due to the highly nonhomogenous nature of the load: periods of basic activities are interleaved with bursty periods of demands. The models frequently used to predict transient behaviour of these networks are based on fluid approximation. Usually they assume Poisson arrivals and consider only mean values of queues. Here, we propose a diffusion model which takes into account general input process and allows us to study the dynamics of nonstationary traffic along virtual path, to approximate transient distributions of queues and transient distributions of response times of one or several nodes. It also permits the estimation of time-varying loss rates due to limited capacity of buffers.

Keyword Codes: C.2.3; I.6.5
Keywords: Network Operations; Model Development

1. INTRODUCTION

BISDN will support various classes of multimedia traffic with different bit rates and different quality of service requirements. Traffic in the networks is expected to be very bursty and nonstationary. Because of the available bandwith, these networks will not be bandwidth-limited but latency-limited systems. Flow control and congestion control algorithms will not be based as previously on the window mechanism. Alternative techniques to prevent the congestion have been proposed and analysed [1]. Transmission delays and cell loss probabilities are the most commonly used decision criteria in admission control. If long term time-averaged values of these variables are used, there is a danger that during periods of temporal network congestion a large number of cells

D. D. Kouvatsos (ed.), *Performance Modelling and Evaluation of ATM Networks*

may be lost even when the long term time-average value of loss rate is kept small, see realistic examples in [2, 3].

Therefore, the dynamics of flows becomes an important factor which requires to be modelled. It is studied by transient solutions of corresponding models.

The direct approach lies in transient analysis of continuous-time Markov chains. Markov models are able to express various synchronization constraints related to control machanisms in a network but their application encounters several numerical problems, such as *size* of the models, which easily overpass 50 000 or 100 000 states, *ill conditionning* and *stiffnes* of the equations. A considerable effort have been already taken to overcome these problems. Explicite differential solution methods (Runge-Kutta), special stable implicit methods, an uniformization (randomization) method based on the reduction of continuous-time Markov chains to a discrete-time Markov chain subordinated to a Poisson process were proposed and tested [4]. Also some efficient approximative methods based on the use of Krylov subspaces (the original matrix of an infinitesimal generator is projected into Krylov subspace; the new matrix has the same eigenvalues but its dimention is considerably smaller) were applied [5]. Nevertheless, this approach is still a challenge to one's skills and numerical experience of a modelling person.

Another approach consists in the use of fluid approximation [6]. First order differential equations, refferred to as fluid flow equations, of the type $\frac{dX(t)}{dx} = f_{in}(t) - f_{out}(t)$ are developed; $X(t)$ is the average number of customers in the system, $f_{in}(t)$, $f_{out}(t)$ are average flows *in* and *out* of the system. The equations may include the classes of customers and the effects of non-premptive priority; they are solved numerically.

Diffusion approximation lies between the both extremities: the use of two first moments of interarrival and service time distributions improves the approximation because fluid approximation is based on the first moments only. The computation effort related to diffusion approximation is considerably smaller than in numerical analysis of Markov chains. There are already diffusion models applied to ATM networks [7], transforming diffusion process to Ornstein-Uhlenbeck process. Our approach uses a special method of solution to diffusion equation with instantaneous return process acting as boundary conditions; this method represents the desired probability density function (pdf) of the process by a spectrum of pd functions of other simpler diffusion processes [8]. That helps us to obtain analytical solution which is relatively easily computable and was already applied by us in the analysis of a single ATM node [9].

The article is organised as follows. Sections 2 reviews briefly basic concepts of diffusion approximation and recalls its main results concerning the steady-state multiple class models of a single server and of an open network of servers. Section 3 summarizes the diffusion model of the single ATM node with a space priority push-out mechanism presented and validated in [9]. It is based on $G/G/1/N$ diffusion model and on an iterative procedure reflecting the mechanism of the buffer management. The procedure can be easily replaced if a new mechanism is to be studied. Section 4 presents our transient solution of the diffusion model of a single multiclass $G/G/1$ server and of a network of $G/G/1$ servers. Time-dependent arrival and service processes are admitted as well as time-dependent routing matrices. The results are expressed in terms of Laplace transforms and numerically converted. Section 5 applies both types of results (steady-state and transient) to a network of stations in series representing a virtual path in an ATM network. The queueing network of $G/G/1$ stations is used to study the dynamics of nonstationary traffic along virtual path. Then the $G/G/1/N$ *Push-Out* models can evaluate the time-dependent cell loss at each node subjected to time-varying load. Two classes of cells with distinct loss requirements are considered. The interaction between virtual circuit traffic and local traffic is taken into account; both types of traffic have two classes of customers and are subjected to changes in time. An equivalent simulation model would be costly because of very low loss rates and consequently long simulation runs which should be performed.

The diffusion approach seems to be a natural tool to deal with the transient states and particularily well suited to model ATM networks.

2. STEADY-STATE DIFFUSION APPROXIMATION OF A SINGLE SERVER AND OF AN OPEN QUEUEING NETWORK

Let $A(x)$, $B(x)$ denote the interarrival and service time distributions in a service station. The distributions are general but not specified, the method requires only their two first moments. The means are: $E[A] = 1/\lambda$, $E[B] = 1/\mu$ and variances are $\mathrm{Var}[A] = \sigma_A^2$, $\mathrm{Var}[B] = \sigma_B^2$. Denote also squared coefficients of variation $C_A^2 = \sigma_A^2\lambda^2$, $C_B^2 = \sigma_B^2\mu^2$. $N(t)$ represents the number of customers present in the system at time t.

According to the central limit theorem, the number of customers arriving in sufficiently long time interval $[0,t]$ may be approximated by the normal distribution with mean λt and variance $\sigma_A^2\lambda^3 t$. Similarly, the number of customers served in this time is approximately normally distributed with mean μt and variance $\sigma_B^2\mu^3 t$ provided that the server is busy all the time. Consequently, the changes of $N(t)$ within interval $[0,t]$, $N(t) - N(0)$, have approximately normal distribution with mean $(\lambda - \mu)t$ and variance $(\sigma_A^2\lambda^3 + \sigma_B^2\mu^3)t$.

Diffusion approximation [10] replaces the process $N(t)$ by a continous diffusion process $X(t)$ which incremental changes $dX(t) = X(t+dt) - X(t)$ are normally distributed with the mean βdt and variance αdt, where β, α are coefficients of the diffusion equation which defines the conditional probability density function $f(x,t;x_0)$ of $X(t)$. The both processes $X(t)$ and $N(t)$ have normally distributed changes; the choice $\beta = \lambda - \mu$, $\alpha = \sigma_A^2\lambda^3 + \sigma_B^2\mu^3 = C_A^2\lambda + C_B^2\mu$ ensures the same ratio of time-growth of mean and variance of these distributions.

Limit boundaries reflecting constraints of $N(t)$ should be imposed on $X(t)$. In [11] diffusion approximation of a $G/G/1/N$ station was studied as a process $X(t)$ which is defined on the closed interval $x \in [0,N]$. Within the interval $x \in (0,N)$ $X(t)$ is a diffusion process; when it comes to $x = 0$, it remains there for a time exponentially distributed with a parameter λ_0 and then it returns to $x = 1$; when the process comes to $x = N$, it remains there for a time which is exponentially distributed with a parameter λ_N and then it starts at $x = N - 1$. The time when the process is at $x = 0$ corresponds to the idle time of the system. The sojourn time at $x = N$ corresponds to the time during which the queue is full and the arriving cells are rejected; the jump to $x = N - 1$ corresponds to the departure of the customer which has been served; a place in the queue becomes available again. Therefore, we choose $\lambda_0 = \lambda$ and $\lambda_N = \mu$.

Diffusion equations defining function $f(x,t;x_0)$, the pdf of the process, are the following [11]:

$$\begin{aligned}
\frac{\partial f(x,t;x_0)}{\partial t} &= \frac{\alpha}{2}\frac{\partial^2 f(x,t;x_0)}{\partial x^2} - \beta\frac{\partial f(x,t;x_0)}{\partial x} + \\
&\quad +\lambda_0 p_0(t)\delta(x-1) + \lambda_N p_N(t)\delta(x-N+1)\ , \\
\frac{dp_0(t)}{dt} &= \lim_{x\to 0}\left[\frac{\alpha}{2}\frac{\partial f(x,t;x_0)}{\partial x} - \beta f(x,t;x_0)\right] - \lambda_0 p_0(t)\ , \\
\frac{dp_N(t)}{dt} &= \lim_{x\to N}\left[-\frac{\alpha}{2}\frac{\partial f(x,t;x_0)}{\partial x} + \beta f(x,t;x_0)\right] - \lambda_N p_N(t)\ ,
\end{aligned} \tag{1}$$

where $\delta(x)$ is Dirac delta function.

In stationary state, when $\lim_{t\to\infty} p_0(t) = p_0$, $\lim_{t\to\infty} p_N(t) = p_N$, $\lim_{t\to\infty} f(x,t;x_0) = f(x)$, eqs.(1) become ordinary differential ones and their solution, if $\varrho = \lambda/\mu \neq 1$, may be expressed as:

$$f(x) = \begin{cases} \frac{\lambda p_0}{-\beta}(1-e^{zx}) & \text{for} \quad 0 < x \leq 1\ , \\ \frac{\lambda p_0}{-\beta}(e^{-z}-1)e^{zx} & \text{for} \quad 1 \leq x \leq N-1\ , \\ \frac{\mu p_N}{-\beta}(e^{z(x-N)}-1) & \text{for} \quad N-1 \leq x < N\ , \end{cases} \tag{2}$$

where $z = \frac{2\beta}{\alpha}$ and p_0, p_N are determined through normalization.

The steady-state solution does not depend on the distributions of the sojourn times in boundaries but only on their first moments. The boundary conditions with instantaneous returns from $x = 0$ to $x = 1$ and from $x = N$ to $x = N-1$ make the model insensitive to the system utilization: diffusion model is not a heavy-traffic approximation and gives reasonable results also for light loads.

When the input stream λ is composed of K classes of customers and $\lambda = \sum_{k=1}^{K} \lambda^{(k)}$ (all parameters concerning class k will have an upper index with brackets) then the joint service time pdf is defined as

$$b(x) = \sum_{k=1}^{K} \frac{\lambda^{(k)}}{\lambda} b^{(k)}(x)\ ,$$

hence

$$\frac{1}{\mu} = \sum_{k=1}^{K} \frac{\lambda^{(k)}}{\lambda}\frac{1}{\mu^{(k)}}\ , \quad \text{and} \quad C_B^2 = \mu^2 \sum_{k=1}^{K} \frac{\lambda_{(k)}}{\lambda}\frac{1}{\mu_{(k)}^2}(C_B^{(k)^2}+1) - 1\ . \tag{3}$$

Similarily,

$$C_A^2 = \sum_{k=1}^{K} \frac{\lambda^{(k)}}{\lambda} C_A^{(k)^2}\ . \tag{4}$$

The above parameters yield α, β of the diffusion equation; function $f(x)$ approximates the distribution $p(n)$ of customers of all classes present in the queue: $p(n) \approx f(n)$ and the probability that there are n_k customers of class k is

$$p_k(n_k) = \sum_{n=n_k}^{N} \left[p(n) \binom{n}{n_k} \left(\frac{\lambda_k}{\lambda}\right)^{n_k} \left(1 - \frac{\lambda_k}{\lambda}\right)^{n-n_k} \right], \qquad k = 1, \ldots, K\,. \tag{5}$$

In a $G/G/1$ station the queue is not limited, hence the unique barrier for the diffusion process is placed at $x = 0$; in the system of equations (1) the last (i.e. third) equation disappears as well as the last term in the first equation. In solution (2) we put $p_N \equiv 0$, $N \to \infty$.

Let M be the number of stations and let us suppose at the beginning that there is one class of customers. The throughput of station i is, as usual, obtained from traffic equations

$$\lambda_i = \lambda_{0i} + \sum_{j=1}^{M} \lambda_j r_{ji}\,, \qquad i = 1, \ldots, M, \tag{6}$$

where r_{ji} is routing probability between station j and station i; λ_{0i} is external flow of customers coming from outside of network.

Second moment of interarrival time distribution is obtained from two systems of equations; the first defines C_{Di}^2 — the squared coefficient of variation of $D_i(x)$, the distribution of interdeparture times from station i — as a function of C_{Ai}^2 and C_{Bi}^2; the second defines C_{Aj}^2 as another function of $C_{D1}^2, \ldots, C_{DM}^2$:

1) The formula (7) is exact for $M/G/1$, $M/G/1/N$ stations and is approximate in the case of non-Poisson input [13]

$$d_i(t) = \varrho_i b_i(t) + (1 - \varrho_i) a_i(t) * b_i(t)\,, \quad i = 1, \ldots, M. \tag{7}$$

From (7) we get

$$C_{Di}^2 = \varrho_i^2 C_{Bi}^2 + C_{Ai}^2 (1 - \varrho_i) + \varrho_i (1 - \varrho_i)\,. \tag{8}$$

2) Customers leaving station i according to the distribution $D_i(x)$ choose station j with probability r_{ij}: intervals between customers passing this way have pdf $d_{ij}(x)$

$$d_{ij}(x) = d_i(x) r_{ij} + d_i(x) * d_i(x)(1 - r_{ij}) r_{ij} + d_i(x) * d_i(x) * d_i(x)(1 - r_{ij})^2 r_{ij} + \cdots, \tag{9}$$

Eq. (9) allows us to obtain $C_{Dij}^2 = r_{ij}(C_{Di}^2 - 1) + 1$ and

$$C_{Aj}^2 = \frac{1}{\lambda_j} \sum_{i=1}^{M} r_{ij} \lambda_i [(C_{Di}^2 - 1) r_{ij} + 1] + \frac{C_{0j}^2 \lambda_{0j}}{\lambda_j}\,. \tag{10}$$

Parameters λ_{0j}, C_{0j}^2 represent the external stream of customers.

For K classes of customers with routing probabilities $r_{ij}^{(k)}$ (let us assume for simplicity that the customers do not change their classes) we have

$$\lambda_i^{(k)} = \lambda_{0i}^{(k)} + \sum_{j=1}^{M} \lambda_j^{(k)} r_{ji}^{(k)}\,, \qquad i = 1, \ldots, M; \;\; k = 1, \ldots, K, \tag{11}$$

and

$$C_{Di}^2 = \lambda_i \sum_{k=1}^{K} \frac{\lambda_i^{(k)}}{\mu_i^{(k)2}} [C_{Bi}^{(k)^2} + 1] + 2\varrho_i(1-\varrho_i) + (C_{Ai}^2+1)(1-\varrho_i) - 1 \,, \quad k = 1,\ldots,K \quad (12)$$

$$C_{Aj}^2 = \frac{1}{\lambda_j} \sum_{l=1}^{K} \sum_{k=1}^{K} r_{ij}^{(k)} \lambda_i \left[\left(\frac{\lambda_i^{(k)}}{\lambda_i} (C_{Di}^2 - 1) \right) r_{ij}^{(k)} + 1 \right] + \sum_{k=1}^{K} \frac{C_{0j}^{(k)^2} \lambda_{0j}^{(k)}}{\lambda_j} \,. \quad (13)$$

Eqs. (8), (10) or (12), (13) form a linear system of equations and allow us to determine C_{Ai}^2 and, in consequence, parameters β_i, α_i for each station.

3. $G/G/1/N$ MODEL WITH THE PUSH-OUT POLICY OF REPLACEMENTS

The nodes of an ATM network are represented in our model by $G/G/1/N$ stations with the push-out policy of replacements. While the number n of customers in such a station is inferior to N, it acts as a conventional $G/G/1/N$ station serving two classes of customers. During saturation periods, i.e. when $n = N$, the ordinary customers in queue are being replaced by privileged ones arriving in these periods. In [9] an iterative algorithm to calculate the effective arrival rates under replacement policy is proposed. The function $f(x)$ and the values of p_0, p_N are obtained via $G/G/1/N$ model. The process enters the saturation period with probability $p(N)$. The probability ε that a class-1 customer arriving at a saturation period may replace a class-2 customer is obtained and the corrected flows $\lambda_{\text{eff}}^{(1)}$, $\lambda_{\text{eff}}^{(2)}$ are calculatted as

$$\lambda_{\text{eff}}^{(1)} = \lambda^{(1)}(1-p_N) + p_N \lambda^{(1)}\varepsilon, \qquad \lambda_{\text{eff}}^{(2)} = \lambda^{(2)}(1-p_N) - p_N \lambda^{(1)}\varepsilon. \quad (14)$$

The relative loss of class-1 and class-2 customers is

$$L^{(1)} = \frac{\lambda^{(1)} - \lambda_{\text{eff}}^{(1)}}{\lambda^{(1)}} = p_N(1-\varepsilon) \,, \qquad L^{(2)} = \frac{\lambda^{(2)} - \lambda_{\text{eff}}^{(2)}}{\lambda^{(2)}} = p_N\Big(1 + \frac{\lambda^{(1)}}{\lambda^{(2)}}\varepsilon\Big) \,. \quad (15)$$

The reader is referred to [9] for details. In the case of time-varying input the above steady-state model uses transient solution of $G/G/1/N$ station presented also in [9]. In order to correct $\lambda_{\text{eff}}^{(1)}$, $\lambda_{\text{eff}}^{(2)}$, the algorithm reflecting the push-out mechanism should be restarted every fixed time-interval chosen sufficiently small with respect to the time-scale of changes of input parameters.

In practice, the rate of loss is very small and we may neglect it in the analysis of flow dynamics. Therefore, to simplify the numerical side of the problem, we replace the network of $G/G/1/N$ stations with push-out or other mechanism by the same network of $G/G/1$ stations in order to predict the propagation of time-variable flow. Transient analysis of the $G/G/1$ queueing network is presented in the next section. Once the time-dependent input parameters for each station in the network are obtained, the stations are studied separately with the use of $G/G/1/N$*Push-Out* transient model to determine the loss probabilities as a function of time.

4. TRANSIENT SOLUTION OF $G/G/1$ DIFFUSION MODEL AND OF AN OPEN NETWORK OF $G/G/1$ STATIONS

Transient solution to $G/G/1/N$ model was presented in [8, 9]; $G/G/1$ model is its simplification.

Consider a diffusion process with the absorbing barriers at $x = 0$, started at $t = 0$ from $x = x_0 > 0$. Its probability density function $\phi(x, t; x_0)$ has the following form [14]:

$$\phi(x,t;x_0) = \frac{e^{\frac{\beta}{\alpha}(x-x_0)-\frac{\beta^2}{2\alpha}t}}{\sqrt{2\Pi\alpha t}} \left[e^{-\frac{(x-x_0)^2}{2\alpha t}} - e^{-\frac{(x+x_0)^2}{2\alpha t}} \right] . \tag{16}$$

If the initial condition is defined by a function $\psi(x)$, the pdf of the process is obtained as $\phi(x, t; \psi) = \int_0^\infty \phi(x, t; \xi)\psi(\xi)d\xi$.

The probability density function $f(x, t; \psi)$ of the diffusion process with elementary returns is composed of function $\phi(x, t; \psi)$ which represents the influence of the initial conditions and of functions $\phi(x, t - \tau; 1)$ which are pd functions of diffusion processes with absorbing barrier at $x = 0$ started earlier at time τ at points $x = 1$ with density $g_1(\tau)$:

$$f(x,t;\psi) \;\; = \;\; \phi(x,t;\psi) + \int_0^t g_1(\tau)\phi(x,t-\tau;1)d\tau \; . \tag{17}$$

Density $\gamma_0(t)$ of probability that at time t the process enters to $x = 0$ is

$$\gamma_0(t) \;\; = \;\; p_0(0)\delta(t) + [1 - p_0(0)]\gamma_{\psi,0}(t) + \int_0^t g_1(\tau)\gamma_{1,0}(t-\tau)d\tau \; , \tag{18}$$

where $\gamma_{1,0}(t)$ is the density the first passage time between $x = 1$ and $x = 0$. The function $\gamma_{\psi,0}(t)$ denote density of probability that the initial process, started at $t = 0$ at the point ξ with density $\psi(\xi)$ will end at time t. We may express $g_1(t)$ with the use of function $\gamma_0(t)$:

$$g_1(\tau) \;\; = \;\; \int_0^\tau \gamma_0(t) l_0(\tau - t)dt \; . \tag{19}$$

Laplace transforms of eqs. (18,19) give us $\bar{g}_1(s)$:

$$\bar{g}_1(s) = \Big[p_0(0) + [1 - p_0(0)]\bar{\gamma}_{\psi,0}(s)\Big] \frac{\bar{l}_0(s)}{1 - \bar{l}_0(s)\bar{\gamma}_{1,0}(s)} \tag{20}$$

and the Laplace transform of the density function $f(x, t; \psi)$ is obtained as

$$\bar{f}(x,s;\psi) = \bar{\phi}(x,s;\psi) + \bar{g}_1(s)\,\bar{\phi}(x,s;1) \, . \tag{21}$$

The inverse transforms of these functions could only be numerical but they may be easily obtained with the use of an inversion algorithm; for this purpose we have used Stehfest's algorithm [15].

The above transient solution assumes that parameters of the model are constant. However, in a network of queues the output flows of stations change contineuously, hence the input parameters of each station are also changing during transient period. We are

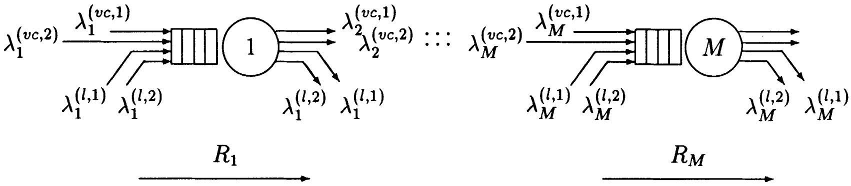

Figure 1. Queueing model of the virtual path.

obliged to discretize this changes and keep the parameters constant within relatively small time interval Δt. Transient solution at the end of each interval Δt allows us to determine $\varrho_i(t)$ and then $\lambda_i(t)$, $D_{Di}^{(k)^2}$ and D_{Ai}^2 for any station i. This solution serves also as the initial condition for the solution in the next interval: for n-th interval, $t \in [(n-1)\Delta t, n\Delta t]$ the density function of the diffusion proces of station i is $f_i(x,t;\psi_n(x))$ where $\psi_n(x) = f_i[x, t=(n-1)\Delta t; \psi_{n-1}(x)]$.

This method was already successfuly applied in [16] for another diffusion model and in analysis of a single ATM node [9].

5. DYNAMICS OF FLOW ALONG A VIRTUAL PATH IN AN ATM NETWORK

We consider a model displayed in Fig. 1: a virtual path in an ATM network is represented by a queueing network of M stations in series.
At each node two streams of customers representing virtual path cells and local traffic cells are mixed together. Each of these streams is composed of two classes of customers that correspond to priority (class 1) and ordinary (class 2) cells. Let us denote $\lambda_i^{(vc,k)}$ the throughput of class k, $k=1,2$, of customers belonging to virtual circuit in station i, $i=1,\dots,M$, and $\lambda_i^{(l,k)}$ the throughputs of local traffic. At the first station the parameters $\lambda_1^{(vc,k)}$, $C_{A1}^{(vc,k)^2}$ are given, for other stations they are obtained from equations (12-13) which for this particular topology are reduced to the form of Eqs. (25-27). The input parameters $\lambda_i^{(l,k)}$, $C_{Ai}^{(l,k)^2}$ of local traffic are given at each station. The service times are constant and equal for all customers and for all stations:

$$\frac{1}{\mu_i^{(vc,k)}} = \frac{1}{\mu_i^{(l,k)}} = \frac{1}{\mu}, \tag{22}$$

$$C_{Bi}^{(vc,k)^2} = C_{Bi}^{(l,k)^2} = 0\,, \qquad k=1,2, \quad i=1,\dots,M. \tag{23}$$

The throughput of station i is obtained as

$$\lambda_i = \sum_{k=1}^{2} \lambda_i^{(vc,k)} + \lambda_i^{(l,k)}\,, \tag{24}$$

the input parameter C_{Ai}^2 is

$$C_{Ai}^2 = \frac{1}{\lambda_i} \sum_{k=1}^{2} \left(\lambda_i^{(vc,k)} C_{Ai}^{(vc,k)^2} + \lambda_i^{(l,k)} C_{Ai}^{(l,k)^2} \right) ; \tag{25}$$

because of Eq. (23) the diffusion parameters are: $\alpha_i = \lambda_i C_{Ai}^2$, $\beta_i = \lambda_i - \mu$.

The squared coefficient of variation in the output stream is expressed as

$$C_{Di}^2 = (1 - \varrho_i) C_{Ai}^2 + \varrho_i (1 - \varrho_i) \tag{26}$$

and for the cells of the virtual circuit becomes:

$$C_{Di}^{(vc,k)^2} = \frac{\lambda_i^{(vc,k)}}{\lambda_i} (C_{Di}^2 - 1) + 1 \, . \tag{27}$$

For virtual path cells we take $C_{Ai}^{(vc,k)^2} = C_{D(i-1)}^{(vc,k)^{\;2}}$, $i = 2, \ldots, M$, — the output of station $i-1$ is directly the input of station i.

In steady state, Eqs. (22-26) may be applied directly; all parameters of these equations are constant and the virtual circuit throughput is the same for all stations:

$$\lambda_i^{(vc,k)} = \lambda_1^{(vc,k)}, \qquad i = 2, \ldots, M. \tag{28}$$

The service times are constant, hence the density functions of the number of customers n present in a station and of the time R spent in this station have the same shape: waiting time is equal to n/μ. The solution $f_i(x)$ of diffusion equation for station i gives the approximation of the number of customers in this station. It is also an approximation of the response time R_i (waiting time plus service time) spent by cells in this station: its pdf $r_i(y)$ is

$$r_i(y) = \mu f_i(\mu y - 1). \tag{29}$$

The argument of f is shifted by 1 because of service time of a considered customer which should be also taken into account. We can estimate the joint response time for m stations forming virtual circuit:

$$r_{1 \ldots m}(y) = r_1(y) * \cdots * r_m(x) \, , \qquad m = 2, \ldots, M. \tag{30}$$

In transient analysis, all parameters λ_i, C_{Ai}^2, C_{Di}^2 are changing with time. We should distinguish the flows *in* and *out* of a station. Their densities will be denoted $\lambda_{i,in}(t)$ and $\lambda_{i,out}(t)$. The output of a station is modelled as

$$\frac{1}{\lambda_{i,out}(t)} = E[d_i(t)] = \frac{1}{\lambda_{i,in}(t)} [1 - \varrho_i(t)] + \frac{1}{\mu} \, , \tag{31}$$

$$C_{Di}^2(t) = C_{Ai}^2(t)[1 - \varrho_i(t)] + \varrho_i(t)[1 - \varrho_i(t)] \, , \tag{32}$$

where $\varrho_i(t) = 1 - p_{i0}(t)$.

At the moment of the change of input parameters the composition of existing queue is determined by old parameters, i.e the probability that a customer belongs to class k is defined by old values of $\lambda_i^{(k)}$ and λ_i. This old composition, together with new ϱ_i will determine the output parameters of the queue for a certain time. If the service time is constant and equal to $1/\mu$, this time is equal x/μ where x is the queue length, hence the pdf of this time has the same form as the queue pdf at the moment $t = t_{ch}$ of the change of the input parameters:

$$q(t) = f(x, t_{ch}; x_0)\delta(t - x/\mu) = \frac{1}{\mu} f(t\mu, t_{ch}; x_0). \tag{33}$$

Our algorithm assumes that parameters change at the beginning of each time-interval Δt and remain fixed within this interval. Denote the value of any parameter, e.g. λ_i, within the j-th interval, i.e. for $t \in [j\Delta t, (j+1)\Delta t)$, by $\lambda_i(j)$. Let $\theta^{(k)}_{i,in}(j)$ be the ratio of class k customers in the input stream of station i within interval j: $\theta^{(k)}_{i,in}(j) = \frac{\lambda^{(k)}_{i,in}(j)}{\lambda_{i,in}(j)}$. At the end of the interval $(j-1)$ the queue length distribution $p_i(n_i; j-1) = f_i[n_i, t = j\Delta t, \psi(x, j-1)]$ is known, as well as the distribution $p_i(n_i; j-2)$ at the end of the previous interval. Using these two distributions we may determine the distribution $\pi_i(n_i, j-1)$ of the number of customers that came during interval $(j-1)$ and are present in the queue at $t = j\Delta t$: $p_i(n_i; j-1)$ is the convolution of $p_i(n_i; j-2)$ [probabilities are shifted by the number of customers served during period $(j-1)$] with $\pi_i(n_i, j-1)$.

We can distinguish several zones in the queue i. Each of them is characterized by its proper ratio of class-1 and class-2 customers, corresponding to the ratio in the input stream at the epoch when the customers of the zone arrived. At the head of the queue there is a zone composed of the most ancient customers. Their origin cannot be more distant than that of the period $(j-h)$, where $h = \frac{N}{\mu\Delta t}$, N is the capacity of the queue and $\mu\Delta t$ is the number of customers which may be served during Δt (suppose for simplicity that N is a multiple of $\mu\Delta t$). This zone with the ratio $\theta^{(k)}_{i,in}(j-h)$ exists with the probability

$$P[n_i > (h-1)\mu\Delta t; j-h+1] = \sum_{n_i=(h-1)\mu\Delta t+1}^{N} \pi_i(n_i; j-h+1)$$

which is equal to the probability that at the end of the period $(j-h)$ there was in the queue more than $(h-1)\mu\Delta t$ customers which arrived during the last period: only in this case at the beginning of j-th interval there are still some of them which have not been served. Otherwise this zone has already vanished and at the head of the queue are customers from a more recent period. Thery are from $(j-h+1)$ interval which composition is characterized by $\theta^{(k)}_{i,in}(j-h)$, provided that at the end of this interval there was in queue more that $(h-2)\mu\Delta t$ customers that arrived during this interval, etc. Therefore the composition of the departure stream of the queue i during period j is expressed as

$$\begin{aligned}
\theta^{(k)}_{i,out}(j) \;=\; & P[n_i > (h-1)\mu\Delta; j-h+1]\, \theta^{(k)}_{i,in}(j-h) + \\
& P[n_i \le (h-1)\mu\Delta; j-h+1]\, P[n_i > (h-2)\mu\Delta; j-h+2]\, \theta^{(k)}_{i,in}(j-h+1) + \\
& P[n_i \le (h-1)\mu\Delta; j-h+1]\, P[n_i \le (h-2)\mu\Delta; j-h+2] \\
& \quad P[n_i > (h-3)\mu\Delta; j-h+3]\, \theta^{(k)}_{i,in}(j-h+2) + \\
& \quad \dots
\end{aligned}$$

$$
\begin{aligned}
& P[n_i \le (h-1)\mu\Delta; j-h+1]\, P[n_i \le (h-2)\mu\Delta; j-h+2] \cdots \\
& \cdots \qquad P[n_i \le (\mu\Delta; j-1]\, \theta_{i,in}^{(k)}(j-1) = \\
= & \sum_{l=2}^{h} P[n_i > (l-1)\mu\Delta; j-l+1]\, \theta_{i,in}^{(k)}(j-l) \prod_{m=l}^{h-1} P[n_i \le m\mu\Delta; m] + \\
& \theta_{i,in}^{(k)}(j-1) \prod_{m=l}^{h-1} P[n_i \le m\mu\Delta; m], \qquad (34)
\end{aligned}
$$

with $\prod_{m=a}^{b}(\cdot) \equiv 1$ if $b < a$.

The response time has time-dependent pdf $r_i(y,t;\psi) = \mu f_i(\mu y - 1, t; \psi)$. The density $r_{1\ldots m}(y,t;\psi)$ of joint response time of m stations in series may be obtained in the same way as in Eq. (30). It characterizes the random delay due to queueing in multiplexing stages, i.e. the phenomenon of jitter. Various measures of this delay, such as squared coefficient of variation or quantiles can be derived from its density.

Numerical examples

We consider a network of 4 nodes, $M = 4$; the service time is equal to the time-unit: $\frac{1}{\mu} = 1$. We suppose that the parameters of local traffic are constant and same for all stations: $\lambda_i^{(l,1)} = \lambda_i^{(l,2)} = 0.25$ and $C_{Ai}^{(l,1)^2} = C_{Ai}^{(l,2)^2} = 1$. The traffic of priority cells in virtual circuit which enters first station is a function of time: during low activity period $\lambda_1^{(vc',1)} = 0.05$, $C_{Ai}^{(vc',1)^2} = 1$; during bursts $\lambda_1^{(vc'',1)} = 0.50$, $C_{Ai}^{(vc'',1)^2} = 0.50$. Bursts have lenght of 20 units of time while interburst periods are 80 time-units long. The traffic of ordinary cells in virtual circuit has constant parameters $\lambda_1^{(vc,2)} = 0.05$, $C_{Ai}^{(vc,2)^2} = 1$.

Fig. 2 presents the flow $\lambda_{1,in}(t)$ at the entrance of first station and flows $\lambda_{i,out}$, $i = 1, \ldots, 4$ at departure of each station. The steady-state queue distribution (with low activity period parameters) was chosen as initial condition at $t = 0$, i.e. $\psi(x) = f(x)$; $\Delta t = 5$ time-units. Propagation times were not taken into account. During the first burst period some minor differences among $\lambda_{i,out}$ are visible, during next bursts the flows are practically indistinguishable. Fig. 3 presents squared coefficient of variation of interdeparture times: the influence of fixed service time and the decrease of C_{Di}^2 with the number of station is visible. Several solutions of $f_1(x,t;\psi)$ during bursty and silent periods are plotted in Figs. 4,5; the same density functions are traced in logarithmic scale in next figure in order to visualise small probabilities at the tail of the queue distribution. Fig. 7 presents the evolution of mean queue lengths $E[N_i(t)]$ corresponding to the dynamics of flows displayed in Fig. 2. Fig. 8 gives squared coefficients of variation of response time observed at stations.

In general, the first two stations ($i = 1, 2$) have different characteristics that reflect the features of the virtual circuit input stream. Then the influence of input stream is filtered and the performances of further successive stations of the virtual path ($i > 2$) become almost identical, provided that the properties of local traffic are similar.

In case of m stations in series with the same C_R^2 response time, the $C_{R_1\ldots R_m}^2$ of the total response time is $C_{R_1\ldots R_m}^2 = \frac{1}{m} C_R^2$.

Figs. 2-8 refer to the global stream of customers. In Fig. 9 the output flows $\lambda_{i,out}^{(vc,1)}(t)$, $\lambda_{i,out}^{(vc,2)}(t)$ of priority and ordinary cells of virtual circuit as well as $\lambda_{i,out}^{(l,1)}(t)$, $\lambda_{i,out}^{(l,2)}(t)$ of priority and ordinary cells of local stream are plotted and compared with $\lambda_i(t)$. Fig. 10 repeats the curve $\lambda_{i,out}^{(vc,2)}(t)$ in a more convenient scale. At the beginning of the bursty period the flow $\lambda_{i,out}^{(vc,1)}(t)$ grows proportionally to $\varrho_i(t)$ (i.e. to $\lambda_i(t)$): the composition of customers in service is determined by old, characteristic for interburst period parameters. Then the zone of higher density of class-1 virtual circuit customers comes to the head of the queue and the output flow $\lambda_{i,out}^{(vc,1)}(t)$ increases abruptly, even over its maximum input value.

Fig. 11 presents squared coefficients of variation $C_{Di}^{(vc,1)^2}(t)$, $C_{Di}^{(vc,2)^2}(t)$ for priority and ordinary cells leaving any station i in the virtual circuit. When the custemers of a given class are comparatively rare, their coefficient of variation is not far from unity, cf. Eq. (12).

Fig. 12 presents relative loss $L^{(1)}(t)$ and $L^{(2)}(t)$, as defined by Eqs. (15) of class-1 and class-2 cells at node 1 for bursty and interbursty periods as in Fig. 2. They were obtained with the use of model of Section 3; the length of the buffer $N = 15$. Similar curves were obtained for the next nodes.

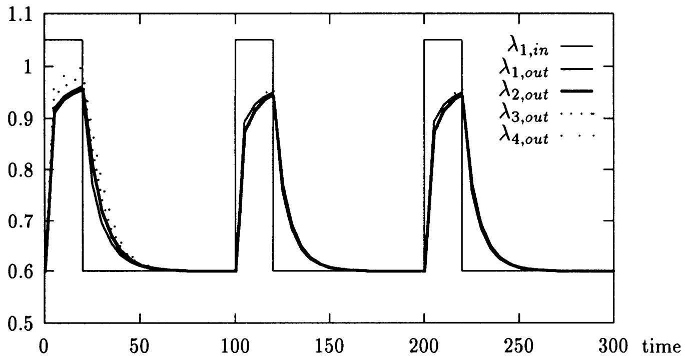

Figure 2. The density of input flow at first station and densities of output flows at successive nodes.

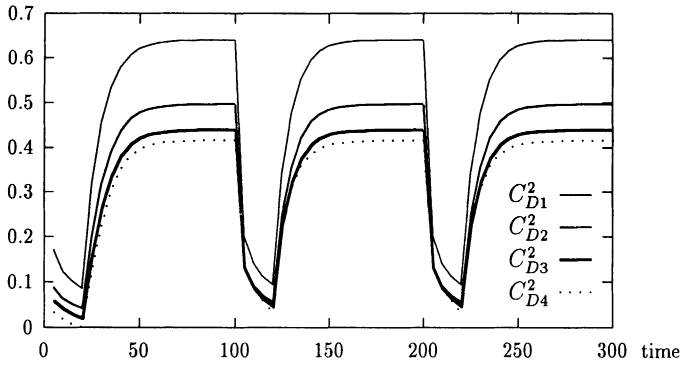

Figure 3. Squared coefficient $C^2_{Di}(t)$, $i = 1, \ldots, 4$, of interdeparture time distributions.

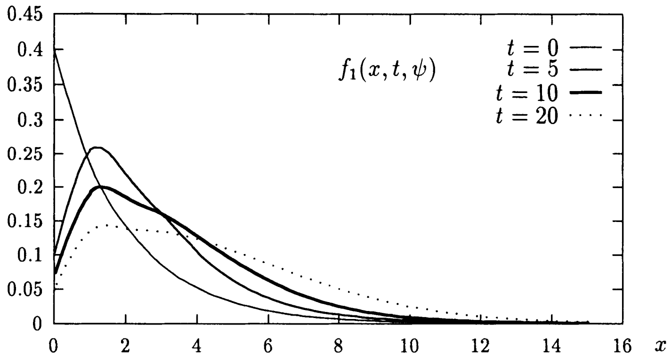

Figure 4. Transient queue length distribution $f_1(x, t; \psi)$ during bursty period.

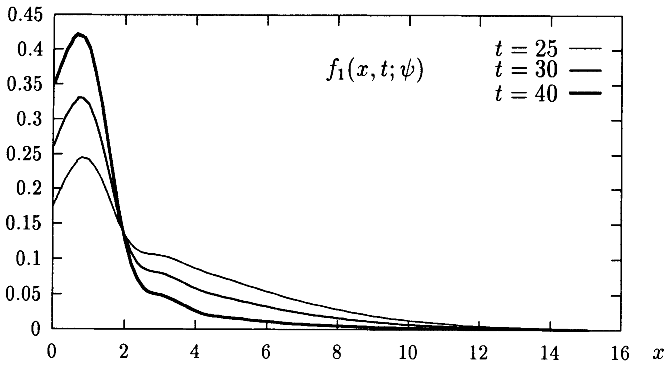

Figure 5. Transient queue length distribution $f_1(x,t;\psi)$ during silent period.

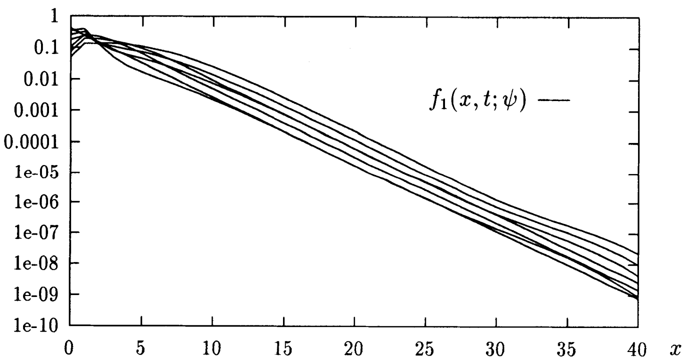

Figure 6. Queue length distributions $f_1(x,t;\psi)$, the same as in Figs. 4-5, represented in logarithmic scale.

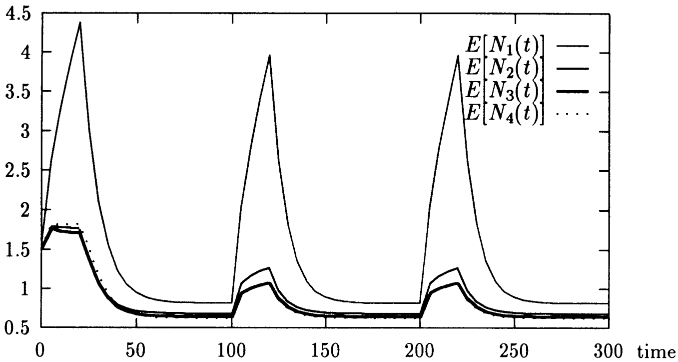

Figure 7. Time-dependent mean queue lengths $E[N_i(t)]$ at nodes i.

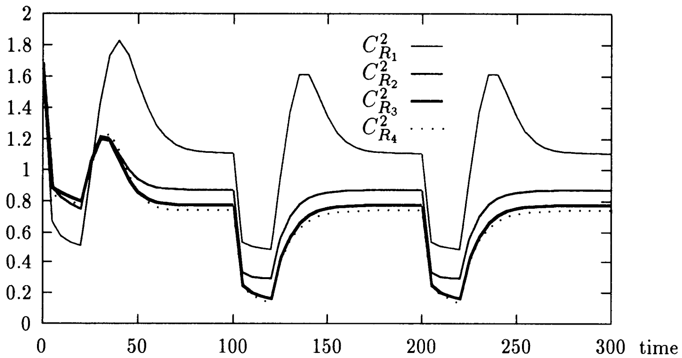

Figure 8. Squared coefficient of variation of response time at nodes i.

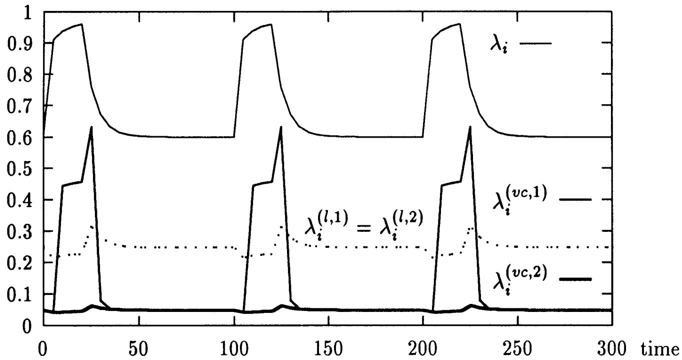

Figure 9. The densities $\lambda_{i,out}^{(vc,1)}(t)$ and $\lambda_{i,out}^{(vc,2)}(t)$ of priority and ordinary cells in virtual circuit compared with density of the global stream.

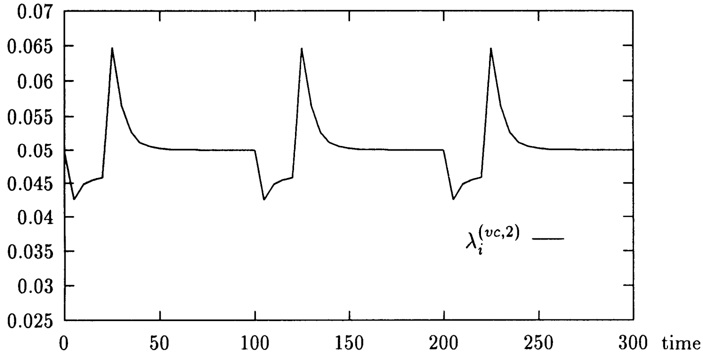

Figure 10. The density $\lambda_{i,out}^{(vc,2)}(t)$ of ordinary cells in virtual circuit stream.

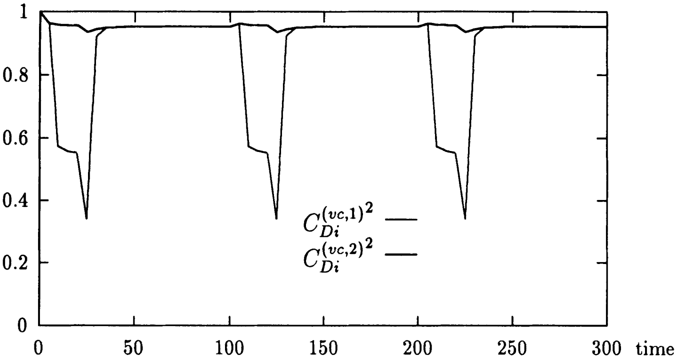

Figure 11. Squared coefficients of variation $C_{Di}^{(vc,1)^2}(t)$, $C_{Di}^{(vc,2)^2}(t)$ of interdeparture times from any station i in virtual circuit.

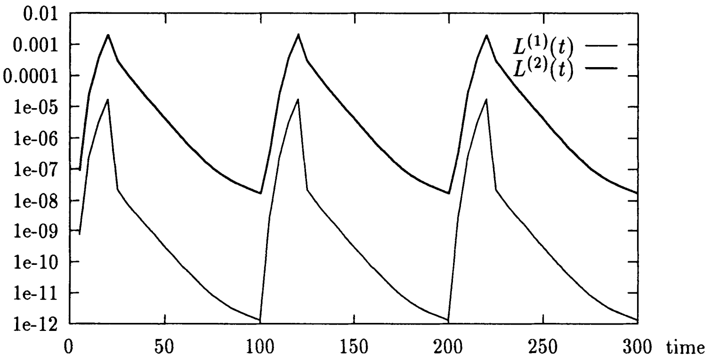

Figure 12. Relative loss $L^{(1)}(t)$ and $L^{(2)}(t)$ of class-1 and class-2 cells at node 1 for bursty and interbursty periods.

6. CONCLUSIONS

Diffusion approximation seems to be particularily well suited to model time-dependent flows in ATM networks because they are composed of the flows of large number of small cells. It allows us to take into account the variances of incoming flows, the priority of cells and gives the estimations of time-dependent queue lengths at each node and time-dependent response times for one node or a series of nodes. Hence, the time-dependent cell loss as well as jitter may be predicted. The numerical effort lies rather in carefull programming to ensure satisfactorily small computation errors then in consumed CPU time; typical examples demand few minutes of a workstation time.

REFERENCES

1. J. W. Roberts (ed.), COST92: Performance Evaluation and design of multiservice networks, Final Report, Office for Official Publications of the European Communities, Luxemburg, 1992.
2. T. Kamitake and T. Suda, Proc. IEEE GLOBECOM '89 (1989) 49.4.1.
3. S-Q. Li, IEEE Trans. Comm., 37 (1989) 1192.
4. A. Reibman and K. Trivedi, Comput. Opns. Res., 15 (1988) 19.
5. B. Philippe and R. B. Sidje, Transient Solutions of Markov Processes by Krylov Subspaces, IRISA Publication interne No. 736 (1993).
6. S. Sharma and D. Tipper, Proc. of IEEE International Conf. on Communications ICC '93, (1993) 352.
7. H. Kobayashi and Q. Ren, Proc. of IEEE International Conf. on Communications ICC '93, (1993) 1047.
8. T. Czachórski, Bulletin of Polish Academy of Sciences (Technical Sciences), No. 4 (1993).
9. T. Czachórski, J. M. Fourneau and F. Pekergin, Proc. IEEE Conf. on Computer Communications INFOCOM '92 (1992).
10. G. F. Newell, Applications of Queueing Theory, Chapman and Hall, London, 1971.
11. E. Gelenbe, J. ACM, No. 2 (1975).
12. E. Gelenbe and G. Pujolle, Acta Informatica, Fasc. 7 (1976).
13. P. J. Burke, Operations Research, No. 6 (1956) 699.
14. D. R. Cox and H. D. Miller, The Theory of Stochastic Processes, Methuen, London, 1965.
15. H. Stehfest, Comm. of ACM, No. 1 (1970) 47.
16. A. Duda, IEEE J. on Selected Areas in Communications, No. 6 (1986).