An on-line Estimation Procedure for Cell-Loss Probabilities in ATM links

Costas Courcoubetis*, George Fouskas*
and Richard Weber†

Abstract

We present a methodology for the on-line estimation of the cell-loss probability of an ATM link. It is particularly suitable for estimating small probabilities of the order 10^{-6}–10^{-9}, with variance orders of magnitude smaller than traditional estimators. The method is justified by the theory of large deviations and the information it requires is based on the actual traffic flows rather than the analysis of some specific traffic models. The method is effective when there is a large degree of statistical multiplexing; in other words, when the number of input traffic sources is large. The statistical properties we require for the traffic are very general and are met by most real-time traffic source models. Implementation issues of this methodology, which demonstrate its simplicity, are also discussed.

1 Introduction

An ATM network provides support for a wide range of services which have differing bandwidth requirements, traffic pattern statistics and Quality of Service (QoS) requirements. All information flows are organized into fixed-size packets, called cells, and carried over high capacity links between switches in the network. Cells belonging to a number of different calls are statistically multiplexed in order to make efficient use of network resources. Switches are buffered, as a safeguard against those cases that cells arrive at a switch faster than they can be switched to output links. Occasionally, the buffer will not be large enough and cell loss will occur. Therefore, a measure of QoS offered by the network to a traffic stream is its cell loss probability (CLP); for most types of traffic this probability should be very small, of the order of 10^{-6} to 10^{-9}.

Accurate and timely CLP estimation is important but difficult

The network is required to attain high utilization and to provide acceptable QoS, with the help of the *Traffic and Congestion Control* mechanism [8]. The *Connection Admission Control* is a primary function of that mechanism, which must decide whether or not to accept a new call, based on the call's traffic characteristics and QoS requirements, and on the current network status. It is therefore important that the network have a mechanism for obtaining accurate and timely estimators of the CLP that occur at each link.

A generic problem that the network management system faces is that since the events of interest

*Department of Computer Science, University of Crete and Institute of Computer Science, FORTH, Greece; e-mail: {courcou,fouskas}@ics.forth.gr

†Statistical Laboratory, 16 Mill Lane, Cambridge CB2 1SB, U.K.; e-mail: rrw1@statslab.cam.ac.uk

are rare (e.g., cell-loss rates of the order of $10^{-6} - 10^{-9}$), any direct brute-force on-line estimation procedure would fail because of the large time required to make an accurate estimate. For example, assume that the CLP is equal to p, and that the cell-loss events are i.i.d. In order to estimate the CLP, with reasonable confidence interval, by simply counting the number of cells arriving and the number of cells being dropped, one would have to observe in the order of $10^3 p^{-1}$ cells passing through the buffer [15]. Now suppose that we have a 155 Mb/s link and that the load on the link is 0.80, with an anticipated CLP of the order of 10^{-6}. Then, the estimate would require approximately $3 \cdot 10^3$ seconds of network time. Such a time is *impractical* since decisions about accepting a new call in the network must be done quickly, and because the composition of the traffic can change in such a large time interval.

Indirect estimation procedures for CLP

The use of queueing theory analysis to accurately predict the cell loss rate in ATM links has been studied in a series of papers [10], [1], [16], [13]. Specifically, the GI/D/1/K model was analyzed in [10], the MMPP(2)/D/1/K in [10] and [1], the MMPP(2)/E_L/1/K in [16] and the MMPP(J)/E_L/1/K in [13]. In order to build the mathematical model for the real traffic stream, some statistics, such as the squared coefficient of variation of cell interarrival time in [10] and the index of dispersion for counts in [16], were used. The values of those statistics, known as *traffic descriptors*, are determined using real traffic data. Since the traffic descriptors are usually complex statistics, the collection and processing of that data isn't simple tasks, and often require long sampling periods to make an accurate estimate. Furthermore, the arrival process of those models is inadequate to characterize the superposition of traffic sources in realistic systems.

Assuming a bufferless model for the real buffered-resource system, we are able to derive an upper bound of the CLP, since the buffer can save cells that should be discarded in the bufferless model. In [14], Saito and Shiomoto develop an estimation technique that relies on the peak and average rates of each traffic source. A similar in spirit technique was proposed by Murase et al. in [9]. Both techniques are based on the assumption that each call produces cells according to the worst possible traffic pattern, given the peak and average rates. In [2], Bean studies this bufferless model for the special case where the traffic sources are characterized by the exponential ON/OFF model. The work in that paper uses the Chernoff approximation, the effective bandwidth concept and the statistical technique known as bootstrap to estimate the cell loss rate when there is incomplete source information. Using those techniques we ensure that the CLP is never underestimated, but we can't ensure that it isn't too overestimated, especially when the buffer size is large.

In [7], the CLP is approximated in terms of the first and the second moment of the aggregate arrival rate and the size of the buffer. The approximation is effective only when the aggregate arrival rate is normally distributed. In [3], the CLP is given by a formula involving the mean rate of each source, its index of dispersion and the buffer size. That formula takes the form of the first two terms in an asymptotic expansion that is accurate for large buffer sizes.

In [4], the CLP is given by a function, whose form is known by a formal mathematical development based on large deviation asymptotics. The function is expressed in terms of three unknown coefficients, which depend on the system and traffic parameters, and are determined as follows. First, we select three "virtual subsystems" whose CLP is much larger than that of the actual system, and therefore can be measured with small variance in a relatively short time. Then, we accurately measure the CLP in those subsystems and solve the resulting three equations to determine the unknown coefficients. Finally, we plug in the obtained values of the coefficients into the expression of the function, and thus compute the desired CLP for the actual system. The selection of the "virtual subsystems" is a critical point in the above algorithm, but no clear selection criteria were specified in that paper.

Our approach

The key idea in our approach, developed in [5], is similar to the one found in [4], but applied on a different asymptotic for the system, which resulted in a large improvement of the variance reduction in systems of practical interest. Also, it includes "virtual subsystems" selection criteria and uses complex statistical techniques to determine the unknown coefficients, their variance and their covariance. So, we are able to estimate not only the desired CLP but also its variance.

An application of our results is towards the on-line estimation of the CLPs in actual ATM links, which serve either a single buffer in a FCFS manner, or two buffers in a complete priority manner. In the second case, our methodology estimates the cell loss rate inccured in the high-priority buffer. Furthermore, by a small variation of the method we can infer the CLP that would occur if we would accept an extra amount of traffic in the system (similarly to [4]). This variant can be used for implementing dynamic Connection Admission Control procedures in ATM networks.

The rest of the paper is organized as follows. In Section 2 we explain the model and present the CLP formula used in our procedure. In Section 3 we describe the on-line estimation methodology. Section 4 contains the implementation issues, and in Section 5 we conclude with a summary and some directions for further research.

2 The model and the CLP formula

In this section we present our model and the asymptotic CLP formula. The precision of the formula increases as a certain parameter of the systems grows. In our case this is the parameter N, defined below, which represents the "size" of the system. For details, the reader is referred to [5].

An output link of an ATM switch is modeled as a queue with a constant service rate and a buffer for B cells. We shall suppose that time is discretized into epochs, $n = 1, 2, \ldots$, and that during epoch n the cell service rate and arrival rate are both constant and equal to C and X_n cells per epoch respectively. This models the usual case where there are dedicated resources C and B for the real-time traffic on the above link (using high priority for real-time traffic ensures the first condition). The workload at the start of epoch n is denoted W_n and

$$W_{n+1} = \max\{0, \min[(W_n + X_n - C), B]\}.$$

We shall assume that $\{X_n\}$ is an ergodic process, and that $E[X_n] < C$. This condition implies that $\{W_n\}$ is ergodic.

Suppose that $\{X_n\}$ is the superposition of sources of M different types, with $N\rho_i$ sources of type i, $i = 1, \ldots, M$. Thus the traffic is defined by the vector $\rho = (\rho_1, \ldots, \rho_M)$ and the scaling parameter N. Let b and c denote respectively the amounts of buffer space and bandwidth per source, so that $B = Nb$ and $C = Nc$.

The cell loss rate can be expressed as $L(c, b, N) = E[(W_n + X_n - C - B, 0)^+]$, when W_n and X_n have their stationary values. A related measure is the proportion of time the buffer is full, which we denote $\Phi(c, b, N) = P(W_n = B)$. Assuming that D is the average number of cells lost in an overflowing epoch, the relationship between the CLP and the probability of buffer overflow in an epoch is $L = D\Phi$. If the epoch's duration is equal to the cell transmission time, then D is less than the number of ATM switch input links, and so we can state that Φ is a good approximation of the CLP.

The large N asymptotic

The large N asymptotic with which we are concerned takes the following form.

$$\Phi(c, b, N) = \exp(-NI(c, b) + g_1(c, b, N)), \tag{1}$$

where $\lim_{N\to\infty} g_1(c, b, N)/N = 0$. Both $I(c, b)$ and $g_1(c, b, N)$ depend on ρ but as we have done above, this is suppressed. The rate $I(c, b)$ is found as the solution to an optimization problem posed in terms of time dependent logarithmic moment generating functions.
Asymptotic (1) suggests that for large values of N (number of traffic sources) the frequency of buffer overflow $\Phi(N)$ can be approximated by

$$\Phi(N) = AN^{\xi} \exp(-NI), \tag{2}$$

where A and ξ are constants. This is because we could neglect the terms varying slower than $\log(N)$ in an expansion of $g_1(c, b, N)$ and absorb them in the constant A.

3 The on-line estimation methodology

The following example illustrates the basic idea of our methodology. Consider an ATM link that is carrying 1000 sources, with a loss rate of about 10^{-10}. As we have already mentioned, it is difficult to measure such a small cell-loss rate directly. However, if we measure the loss rate p in a "virtual subsystem", that is carrying 200 sources and has bandwidth and buffer space equal to the one-fifth of the corresponding values of the actual link, then, according to the large N asymptotic result, we should obtain a value of about 10^{-2}. Such a cell-loss rate is much larger than that of the actual link, and therefore can be measured with small variance in a relatively short time. *The interesting question consists on how to get the CLP value of the real system using the CLP estimates of the "virtual subsystems".*

Assume that $\Sigma = \{1, \ldots, N\}$ is our real system, where N is the number of calls which are served by the ATM link. A subsystem $\Sigma_i \subseteq \Sigma$ is described by the calls it serves, and let $N_i = |\Sigma_i|$ be the number of those calls. Our methodology is described by the following steps:

1st) Select K "virtual subsystems" $\Sigma_1, \ldots, \Sigma_K$.

2nd) Estimate the cell loss rate Φ_i directly, for each subsystem Σ_i, $i = 1, \ldots, K$.

3rd) Determine the unknown coefficients A, ξ and I, by solving the K linear equations

$$\log \Phi_i = \log \Phi(N_i) = \log A + \xi \log N_i - IN_i, \quad i = 1, \ldots, K, \tag{3}$$

4th) Calculate the desired quantity $\log \Phi(N) = \log A + \xi \log N - IN$

There are some issues that have to be clarified. The first concerns the choice of the K "virtual subsystems". A small subsystem, i.e. a subsystem that serves few calls, will result in estimators with small variance but which will contribute too large extrapolation errors since the approximation formula (2) becomes accurate only when there is a large degree of statistical multiplexing. On the other hand, if we select a large subsystem, then we face the problem of the large amount of time required to obtain an accurate CLP estimate directly. In order to characterize the accuracy of the CLP estimate in each subsystem, we obtain its variance. Those values are taken into consideration when determining the unknown coefficients A, ξ and I.

3.1 Determination of the A, ξ and I coefficients

In our methodology, the solution of the linear equations (3) is based on the regression analysis. In order to describe that solution technique we rewrite the equation (3) using a different notation

$$y(N_i) = \sum_{k=1}^{3} a_k X_k(N_i) \tag{4}$$

where $y(N_i) = \log \Phi(N_i)$, $X_1(N_i) = 1$, $X_2(N_i) = \log N_i$, $X_3(N_i) = N_i$, and the a_1, a_2, a_3 are the unknown coefficients with $a_1 = \log A$, $a_2 = \xi$ and $a_3 = -I$.

If we denote by y_i the logarithm of the direct CLP estimate for the i-th "virtual subsystem", i.e. $y_i = \log \Phi_i$, $i = 1, \ldots, K$, then we can determine the unknown coefficients by solving the following minimization problem

$$\textit{minimize over } a_1, a_2, a_3 : \qquad \chi^2 \equiv \sum_{i=1}^{K} \left[\frac{y_i - \sum_{k=1}^{3} a_k X_k(N_i)}{\sigma_i} \right]^2 \tag{5}$$

where σ_i is the standard deviation of the y_i estimator. This procedure, known as "weighted least-squares fitting", is the maximum likelihood estimation of the fitted parameters if the y_i estimation errors are independent and normally distributed. A simple estimation of σ_i is presented in subsection 3.3, so we can assume that the values of σ_i is known.

Function χ^2 obtains its minimum when the following equality holds

$$\begin{aligned}
\frac{\partial \chi^2}{\partial a_j} = 0 \Leftrightarrow \quad & \sum_{i=1}^{K} \frac{1}{{\sigma_i}^2} \left[y_i - \sum_{k=1}^{3} a_k X_k(N_i) \right] X_j(N_i) = 0 \quad \Leftrightarrow \\
\Leftrightarrow \quad & \sum_{i=1}^{K} \frac{1}{{\sigma_i}^2} \sum_{k=1}^{3} a_k X_k(N_i) X_j(N_i) = \sum_{i=1}^{K} \frac{y_i X_j(N_i)}{{\sigma_i}^2} \quad \Leftrightarrow \\
\Leftrightarrow \quad & \sum_{k=1}^{3} \sum_{i=1}^{K} \frac{X_k(N_i) X_j(N_i)}{{\sigma_i}^2} a_k = \sum_{i=1}^{K} \frac{y_i X_j(N_i)}{{\sigma_i}^2} \quad \Leftrightarrow \\
\Leftrightarrow \quad & \sum_{k=1}^{3} \beta_{jk} a_k = \gamma_j
\end{aligned} \tag{6}$$

where $\beta_{jk} = \sum_{i=1}^{K} \frac{X_k(N_i) X_j(N_i)}{{\sigma_i}^2}$ is the (j, k) element of a 3×3 matrix and $\gamma_j = \sum_{i=1}^{K} \frac{y_i X_j(N_i)}{{\sigma_i}^2}$ is the j-element of a vector of length 3.

The set of linear equations (6) can be solved using classical numerical methods, such as LU decomposition and backsubstitution or Gauss–Jordan elimination, but these methods are susceptible to roundoff error. There exist a very powerful set of techniques, known as singular value decomposition, that fixes the roundoff problem [6]. A very efficient implementation of a singular value decomposition technique is presented in [11] and is recommended for the solution of equations (6).

Let (δ_{ij}) denote the inverse matrix of (β_{ij}). We can calculate the values of the unknown coefficients using the following equations

$$a_j = \sum_{k=1}^{3} \delta_{jk}\gamma_k = \sum_{k=1}^{3} \delta_{jk} \sum_{i=1}^{K} \frac{y_i X_k(N_i)}{\sigma_i{}^2} = \sum_{i=1}^{K}\sum_{k=1}^{3} \frac{\delta_{jk} X_k(N_i)}{\sigma_i{}^2} y_i \ , \quad j = 1, 2, 3 \tag{7}$$

The term $\sum_{k=1}^{3} \frac{\delta_{jk} X_k(N_i)}{\sigma_i{}^2}$ is independent of y_i. If the CLP estimators for each "virtual subsystem" are mutually independent, then

$$\begin{aligned}
\mathrm{Var}(a_j) &= \sum_{i=1}^{K} \left(\sum_{k=1}^{3} \frac{\delta_{jk} X_k(N_i)}{\sigma_i{}^2} \right)^2 \mathrm{Var}(y_i) = \\
&= \sum_{i=1}^{K} \left(\frac{1}{\sigma_i{}^2} \right)^2 \left(\sum_{k=1}^{3} \delta_{jk} X_k(N_i) \right)^2 \sigma_i{}^2 = \\
&= \sum_{i=1}^{K} \frac{1}{\sigma_i{}^2} \sum_{k=1}^{3} \delta_{jk} X_k(N_i) \sum_{l=1}^{3} \delta_{jl} X_l(N_i) = \\
&= \sum_{k=1}^{3}\sum_{l=1}^{3} \delta_{jk}\delta_{jl} \left[\sum_{i=1}^{K} \frac{X_k(N_i) X_l(N_i)}{\sigma_i{}^2} \right] = \\
&= \sum_{k=1}^{3}\sum_{l=1}^{3} \delta_{jk}\delta_{jl}\beta_{lk}
\end{aligned}$$

Since (δ_{ij}) is the inverse matrix of (β_{ij}), the following equation holds

$$\mathrm{Var}(a_j) = \sum_{k=1}^{3} \delta_{jk} I_{jk} = \delta_{jj} \tag{8}$$

where (I_{ij}) is the identity matrix.

Assuming that

$$Z_{ij} = a_i + a_j = \sum_{k=1}^{3} \delta_{ik}\gamma_k + \sum_{k=1}^{3} \delta_{jk}\gamma_k = \sum_{l=1}^{K}\sum_{k=1}^{3} \frac{(\delta_{ik} + \delta_{jk}) X_k(N_l)}{\sigma_l{}^2} y_l$$

and performing a sequence of steps similar to the above, we can find that

$$\mathrm{Var}(Z_{ij}) = \delta_{ii} + \delta_{ij} + \delta_{ji} + \delta_{jj}$$

The covariance of the a_i and a_j coefficients is given by

$$\mathrm{Cov}(a_i, a_j) = \frac{1}{2}(\delta_{ij} + \delta_{ji}) = \delta_{ij}$$

because the (δ_{ij}) is a symmetric matrix.

Thus, we can calculate the logarithm of the real CLP, $y(N)$, by directly applying formula (4), and its variance by the following relation

$$\mathrm{Var}(y(N)) = \sum_{i=1}^{3} X_i(N)^2 \mathrm{Var}(a_i) + \sum_{(i,j)\,\in\,\{(i,j)\,:\,i,j\in\{1,2,3\}\,\wedge\,i\neq j\}} X_i(N) X_j(N) \mathrm{Cov}(a_i, a_j)$$

The goodness-of-fit is measured by the following quantity, which is called the *squared multiple regression coefficient*

$$R^2 = \frac{\sum \left(\widehat{y(N_i)} - \overline{y}\right)^2}{\sum \left(y_i - \overline{y}\right)^2}$$

where $\overline{y} = (1/K)\sum_{i=1}^{K} y_i$ and $\widehat{y(N_i)}$ is computed from (4). The squared multiple regression coefficient must lie between 0 and 1, and the closer it is to 1, the better the fit. If the value of R^2 is very close to 0, then either (i) the obtained values of the estimation errors σ_i may be wrong, or (ii) the estimation errors σ_i may not be normally distributed, or (iii) equation (4) may not be adequate to characterize the CLP in the current system configuration.

3.2 Selection of the "virtual subsystems"

Consider the K "virtual subsystems" $\Sigma_1, \Sigma_2, \ldots, \Sigma_K$, which are ordered such that $N_1 < N_2 < \cdots < N_K$, where $N_i = |\Sigma_i|$. In order for the procedure described in the previous section to be accurate, the following conditions must be satisfied.

- The CLP estimators for each subsystem must be mutually independent. Thus, the subsystems must not serve common calls, i.e., $\Sigma_i \cap \Sigma_j = \{\emptyset\}$ for $i \neq j$, and
- The sizes N_i of the subsystems must be evenly spaced in the range $[N_1, N_K]$.

Since a large degree of statistical multiplexing in each subsystem is highly desired, we specify a lower threshold of the number of calls that each subsystem should serve, which we denote by n_{min}. Also, we like to use the maximum possible number of data points as inputs to the regression analysis, thus to select as many subsystems as possible.

The proposed "virtual subsystems" selection method satisfies the aforementioned conditions and is based on a simple optimization problem. If N is the number of calls which are served in the real system, then we partition the set of that calls into the $K+1$ subsets $\Sigma_1, \ldots, \Sigma_K, \Sigma^+$, such that $N_i = |\Sigma_i| = in^*$, $i = 1, \ldots, K$, and $\sum_{i=1}^{K} N_i + |\Sigma^+| = N$. The value of n^* is determined by the following minimization problem

$$\begin{aligned}
\textit{minimize} \quad & f(n) \equiv |\Sigma^+| = N - \textstyle\sum_{i=1}^{K} in = N - \frac{K(K+1)}{2}n, \\
\textit{over} \quad & n, \\
\textit{such that} \quad & f(n) \geq 0, \\
& n \text{ integer}, \quad n \geq n_{min}
\end{aligned}$$

for

$$K = \left\lfloor \frac{\sqrt{1 + 8N/n_{min}} - 1}{2} \right\rfloor$$

where $\lfloor x \rfloor$ denotes the integer part of x. That is, we minimize the number of calls that are not served by any subsystem which is used as input to the regression analysis procedure, and, at the same time, we maximize the statistical multiplexing degree in any such subsystem. A numerical solution of the above minimization problem is very simple, since $f(n)$ is a decreasing function.

A final comment is needed about the choice of the constant n_{min}. We do not have any answer yet for the problem concerning the optimal choice of the above constant. In general, for a specific ATM switch, the value of n_{min} depends on the traffic characteristics of the multiplexed calls. So, different values of n_{min} should be used for different compositions of the incoming traffic. Since the choice of the constant n_{min} should be done empirically, it is possible to select subsystems where the large degree of statistical multiplexing assumption does not hold. The data points that correspond to those subsystems are called *outliers*, and can easily turn the results of least-squares fit into nonsense. The use of *regression diagnostics* [12], which are certain quantities computed from the data points with the purpose of pinpointing the outliers, is of current study.

3.3 Estimation of the CLP measurement errors in subsystems

As we have already mention in subsection 3.1, we need to know the variance of the $\log \Phi_i$ estimator, where Φ_i is the cell loss rate measured for subsystem Σ_i, $i = 1, \ldots, K$. Assuming that L_i is the number of cells being dropped and D_i is the number of cells arriving at the ATM link, then an unbiased and consistent estimator of $\log \Phi_i$ is the $\log(L_i/D_i)$.

Let $U_j = [\Sigma_{1,j}, \Sigma_{2,j}, \ldots, \Sigma_{K,j}, \Sigma_j^+]$, $j = 1, \ldots, M$, be M different partitions of the set of calls that are served by the ATM link, which satisfy the criteria described in subsection 3.2. The notation $\Sigma_{i,j}$ denotes the i-th subsystem that belongs to the j-th partition, $i = 1, \ldots, K$, $j = 1, \ldots, M$. Of course it is true that $|\Sigma_{i,j}| = |\Sigma_{i,k}|\ \forall i, j, k$ and $\Sigma_{i,j} \neq \Sigma_{i,k}$ for $j \neq k$. Thus, we have M subsystems of the same size, each serving a different set of calls. Let $y_{i,j}$ be the estimate of the logarithm of the CLP for subsystem $\Sigma_{i,j}$. A crude estimate of the variance of the $\log \Phi_i$ estimator is given by

$$\sigma_i^2 = \frac{1}{M-1}\left\{\sum_{j=1}^{M}(y_{i,j}-\overline{y_i})^2 - \frac{1}{M}\left[\sum_{j=1}^{M}(y_{i,j}-\overline{y_i})\right]^2\right\} \tag{9}$$

where $\overline{y_i} = (1/M)\sum_{j=1}^{M} y_{i,j}$. The second sum does a good job of correcting the roundoff error in the first term.

4 Implementation issues

In this section we discuss some implementation issues of our methodology, wishing to demonstrate its simplicity.

As we can see in figure (1), the on-line estimation methodology can be also implemented in an external device, i.e. outside of the VLSI chip which performs the cell switching. When a cell arrives at the output link, we duplicate the VPI/VCI field of its header and transmit it to the estimation device. Then, using a lookup table, we find the set of subsystems $\Sigma_{i,j}$ which serve that VPI/VCI connection. Three registers, denoted $Z_{i,j}$, $D_{i,j}$ and $L_{i,j}$, are used for each "virtual subsystem". Register $Z_{i,j}$ simulates the output queue behaviour in subsystem $\Sigma_{i,j}$, and $D_{i,j}$, $L_{i,j}$ count the number of cells arriving and being dropped respectively in subsystem $\Sigma_{i,j}$. All three registers are set equal to zero at the start of an estimation period, and, at each cell arrival, are updated in the following manner

```
D_{i,j} ++
Z_{i,j} ++
if ( Z_{i,j} > B_i ) {
      L_{i,j} ++
      Z_{i,j} = B_i
}
```

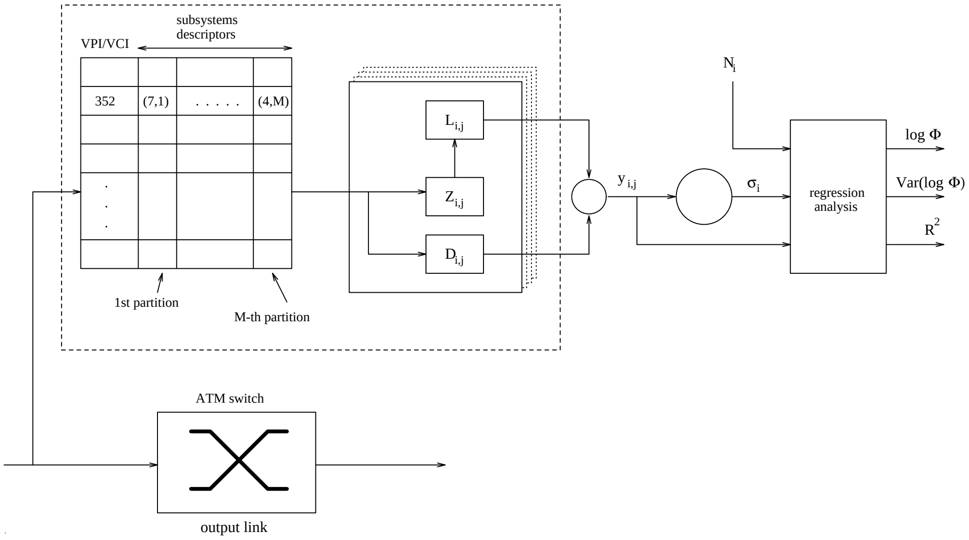

Figure 1: An implementation of the proposed on-line estimation procedure

where $B_i = N_i b$. Also, the value of $Z_{i,j}$ decreases by one every $1/(N_i c)$ seconds, and, at the same time, it is checked so that $Z_{i,j} \geq 0$.

When we want an estimate of the current CLP incurred in the ATM link, we first calculate the $y_{i,j}$ for each subsystem $\Sigma_{i,j}$, using the formula $y_{i,j} = \log(L_{i,j}/D_{i,j})$. Next, we compute σ_i^2 using (9), and insert the values, together with $y_{i,j}$ and N_i, as inputs to the regression analysis procedure described in subsection 3.1. The results of that procedure are the estimates of $\log \Phi(N)$ and $\mathrm{Var}(\log \Phi(N))$ and the squared multiple regression coefficient R^2.

From the above, it is clear that the estimation device functions are simple and require a small amount of memory and processing power, hence the implementation of such a device is feasible with a small cost.

5 Summary and Directions for Further Research

An important issue in ATM networks is the accurate and timely estimation of the cell-loss probabilities that occur at various switches in the network, so that correct decisions can be made about whether or not to accept more traffic or how it should be routed. In this paper, we present a methodology for estimating the cell-loss rate of an ATM link in a reasonably short time. The method uses information that is obtained from the actual traffic flows and it is effective when there is a large degree of statistical multiplexing. The method can be used by Connection Admission Control and Routing procedures in ATM networks. Also, it can be used for speeding up the simulation time for large systems.

An exhaustive simulation study could be conducted in order to study the effectiveness of the proposed on-line estimation procedure, in terms of accuracy and timeliness, using realistic traffic sources, long sequences of real video data and heterogeneous input traffic.

References

[1] A. Baiocchi, N. B. Melazzi, M. Listanti, A. Roveri, and R. Winkler. Loss Performance Analysis

of an ATM Multiplexer Loaded with High-Speed ON-OFF Sources. *IEEE JSAC*, 9(3):388, April 1991.

[2] N. Bean. *Statistical Multiplexing in Broadband Communication Networks.* PhD thesis, Jesus College, University of Cambridge, 1993.

[3] C. Courcoubetis, G. Fouskas, and R. Weber. On the Performance of an Effective Bandwidths Formula. In *Proceedings of the ITC 14*, 1994.

[4] C. Courcoubetis, G. Kesidis, A. Ridder, J. Walrand, and R.R. Weber. Admission Control and Routing in ATM Networks using Inferences from Measured Buffer Occupancy. ORSA/TIMS special interest meeting, Monterey, January, 1991. (to also appear in *IEEE Trans. Communications*).

[5] C. Courcoubetis and R. Weber. Buffer Overflow Asymptotics for a Switch Handling Many Traffic Sources. to appear in *Journal of Applied Probability, 1995.*

[6] G. Forsythe, M. Malcolm, and C. Moler. *Computer Methods for Mathematical Computations.* Prentice Hall, Inc., 1977.

[7] R. Guérin, H. Ahmadi, and M. Naghshineh. Equivalent Capacity and Its Application to Bandwidth Allocation in High-Speed Networks. *IEEE JSAC*, 9(7):968, September 1991.

[8] CCITT Recommendation I.371. Traffic Control and Congestion Control in B-ISDN. Geneva, 1993.

[9] T. Murase, H. Suzuki, S. Sato, and T. Takeuchi. A Call Admission Control Scheme for ATM Networks Using a Simple Quality Estimate. *IEEE JSAC*, 9(9):1461, December 1991.

[10] R. Nagarajan, J. Kurose, and D. Towsley. Approximation Techniques for Computing Packet Loss in Finite-Buffered Voice Multiplexers. *IEEE JSAC*, 9(3):368, April 1991.

[11] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery. *Numerical Recipes in C.* Cambridge University Press, 1994.

[12] P. Rousseeuw and A. Leroy. *Robust Regression & Outlier Detection.* John Wiley & Sons, Inc., 1987.

[13] H. Saito, M. Kawarasaki, and H. Yamada. An Analysis of Statistical Multiplexing in an ATM Transport Network. *IEEE JSAC*, 9(3):359, April 1991.

[14] H. Saito and K. Shiomoto. Dynamic Call Admission Control in ATM Networks. *IEEE JSAC*, 9(7):982, September 1991.

[15] J. Walrand. *An Introduction to Queueing Networks.* Prentice Hall, Inc., 1988.

[16] H. Yamada and S. Sumita. A Traffic Measurement Method and its Application for Cell Loss Probability Estimation in ATM Networks. *IEEE JSAC*, 9(3):315, April 1991.