Stochastic System Identification for ATM Network Traffic Models: a Time Domain Approach[1]

Katrien De Cock and Bart De Moor[2]

Published in *Proc. of the IX European Signal Processing Conference (EUSIPCO-98)*, Rhodes, Greece, September 1998, pp. 2137-2140

[1]This report is available by anonymous ftp from *ftp.esat.kuleuven.ac.be* in the directory *pub/SISTA/decock/reports/98-22.ps.gz*

[2]K.U.Leuven, Dept. of Electrical Engineering (ESAT), Research group SISTA, Kardinaal Mercierlaan 94, 3001 Leuven, Belgium, Tel. 32/16/32 17 09, Fax 32/16/32 19 70, WWW: *http://www.esat.kuleuven.ac.be/sista*. E-mail: *katrien.decock@esat.kuleuven.ac.be, bart.demoor@esat.kuleuven.ac.be*. Bart De Moor is a senior Research Associate with the F.W.O. (Fund for Scientific Research-Flanders). Katrien De Cock is a Research Assistant with the I.W.T. (Flemish Institute for Scientific and Technological Research in Industry). Work supported by the **Flemish Government** (Administration of Science and Innovation (Concerted Research Action MIPS: *Model-based Information Processing Systems*, Bilateral International Collaboration: *Modelling and Identification of nonlinear systems*, IWT-Eureka SINOPSYS: *Model-based structural monitoring using in-operation system identification*), FWO-Vlaanderen: *Analysis and design of matrix algorithms for adaptive signal processing, system identification and control, based on concepts from continuous time system theory and differential geometry*, *Numerical algorithms for subspace system identification: Extension towards specific applications*, FWO-Onderzoeksgemeenschappen: *Identification and Control of Complex Systems*, *Advanced Numerical Methods for Mathematical Modelling*); **Belgian Federal Government**(Interuniversity Attraction Pole IUAP IV/02: *Modelling, Identification, Simulation and Control of Complex Systems*, Interuniversity Attraction Pole IUAP IV/24: IMechS: Intelligent Mechatronic Systems); **European Commission:** (Human Capital and Mobility: *SIMONET: System Identification and Modelling Network*, *SCIENCE-ERNSI: European Research Network for System Identification.*)

STOCHASTIC SYSTEM IDENTIFICATION FOR ATM NETWORK TRAFFIC MODELS: A TIME DOMAIN APPROACH

*Katrien De Cock and Bart De Moor**
K.U.Leuven, ESAT-SISTA,
Kardinaal Mercierlaan 94, 3001 Leuven, Belgium,
Tel. 32/16/32 19 25, Fax 32/16/32 19 70,
WWW: http://www.esat.kuleuven.ac.be/sista.
e-mail: katrien.decock@esat.kuleuven.ac.be, bart.demoor@esat.kuleuven.ac.be.

ABSTRACT

In our paper we discuss a new time domain approach to the traffic identification problem for ATM networks. The Markov modulated Poisson process is identified in two steps.
By applying a nonnegative least squares algorithm we obtain in a very fast way a description of the first order statistics of the data. This first order characterisation includes also an estimate of the model order. Consequently, we are able to identify a Markov modulated Poisson process without a priori knowledge of the model order.
The identification of the second order statistics is based on unconstrained optimisation algorithms.

Keywords: traffic identification, Markov modulated Poisson process, ATM

1 INTRODUCTION

Our paper is concerned with the modelling of the arrival process of cells in one node of the network. The arrival process a_k $(k = 0, 1, 2, \ldots)$ is the aggregated number of cells that arrive per time unit. This sequence is measured and forms the input of our identification algorithm.
An important innovation described in this paper is the fast characterisation of the first order statistics with automatic determination of the model order.
In Section 2 we discuss the identification problem, our identification method is explained in Section 3 and we end with some conclusions in Section 4.

*K. De Cock is a Research Assistant with the I.W.T. (Flemish Institute for Scientific and Technological Research in Industry), B. De Moor is a Senior Research Associate with the F.W.O. (Fund for Scientific Research-Flanders). This work was supported by the Flemish Government (BOF (GOA-MIPS), AWI (Bil. Int. Coll.), FWO (projects, grants, res. comm. (ICCoS)), IWT (IWT-VCST (CVT), ITA (ISIS), EUREKA (Sinopsys))), the Belgian Federal Government (IUAP IV-02, IUAP IMechS), the European Commission (HCM (Simonet), TMR (Alapedes), ACTS (Aspect), SCIENCE (ERNSI)), NATO (CRG) and industry (Electrabel).

2 THE IDENTIFICATION PROBLEM

In this section we state the identification problem. First, we show why the Markov modulated Poisson process is a good model for the arrival process of cells in an ATM network (Section 2.1). Then, in Section 2.2, we explain the structure of the Markov modulated Poisson process and we give the expressions for its first and second order statistics.

2.1 Model Choice

The task of traffic identification is to construct stochastic models, the statistics of which match those of the data. In addition, the identified model should fit into queueing analysis which is used to evaluate the network performance.
The research of Li and Hwang [6] has demonstrated that only the first order and second order statistics have a significant impact on queueing performance. Classic queueing theories have generally ignored the second order statistics entirely. By using a Markov modulated Poisson process with first order and second order statistics matching those of the measured traffic, the correlation in multimedia traffic can be captured and exploited.

2.2 The Markov Modulated Poisson Process

2.2.1 Model Structure

A Markov modulated Poisson process (MMPP) of order N consists of an N-state Markov chain in which each state i $(i = 1, 2, \ldots, N)$ represents a Poisson process with rate γ_i. In other words, a Markov modulated Poisson process is a Poisson process the rate of which is changed (modulated) according to a Markov chain. In this paper, we only consider discrete time Markov chains.
Let S_k $(k = 0, 1, 2, \ldots)$ be the stochastic process that denotes the state moving at discrete time instants between a finite number of states $1, 2, \ldots, N$. The transition probabilities depend only on the previous state, so that S_k $(k = 0, 1, 2, \ldots)$ is a Markov process. Let the matrix $P = [p_{ij}]$ be the transition matrix:

$$p_{ij} = Pr\{S_{k+1} = j \mid S_k = i\} .$$

If $\theta(k)$ is the column vector with i-th entry $Pr\{S_k = i\}$, then $\theta(k+1)^T = \theta(k)^T P$.
Throughout this paper, we will deal exclusively with the steady state case, where the state distribution θ is time independent:

$$\theta^T = \theta^T P \ .$$

Matrix P is a $N \times N$ stochastic matrix, that is $P \in \mathbb{R}^{N\times N}$, $p_{ij} \geq 0$ and $\sum_{j=1}^{N} p_{ij} = 1$. The nonnegative vector θ is the left eigenvector of the stochastic matrix P corresponding to eigenvalue 1 for which holds $\sum_{i=1}^{N} \theta_i = 1$. The Poisson parameters are arranged in a column vector γ. The ith element of γ, $\gamma_i \geq 0$ is the arrival rate of the Poisson process associated with state i.
At time step k, the MMPP (P, γ) will emit a_k arrivals generated by the Poisson process with the arrival rate γ_i if the state of the Markov chain is i.
The aim of traffic identification is to find P and γ from arrivals a_k, so that they form an MMPP with first order and second order statistics that match those of the measured traffic a_k. Afterwards, P and γ can be used in queueing analysis.

2.2.2 The first and second order statistics of arrivals

As already mentioned, only the first and second order statistics of arrivals a_k are considered. The first order statistics are described by the cumulative distribution function F, defined as $F(l) = Pr\{a_k \leq l\}$.
For an MMPP, we have

$$\begin{aligned} F_{\mathrm{M}}(l) &= \sum_{i=1}^{N} \theta_i e^{-\gamma_i} \sum_{j=0}^{l} \frac{\gamma_i^j}{j!} \\ &= \sum_{i=1}^{N} \theta_i F_{\gamma_i}(l) \end{aligned} \qquad (1)$$

where $F_{\gamma_i}(l)$ is the cumulative distribution function of the Poisson process with parameter γ_i.
The second order statistics of a_k are characterised by the autocorrelation function R, defined as $R(n) = E\{a_k \cdot a_{k+n}\}$.
The autocorrelation of an MMPP is given by:

$$\begin{cases} R_{\mathrm{M}}(0) &= \theta^T(\Gamma + \Gamma^2)e \\ R_{\mathrm{M}}(n) &= \theta^T \Gamma P^n \Gamma e, \text{ for all } n \in \mathbb{N}_0 \ , \end{cases} \qquad (2)$$

where $\Gamma = \mathtt{diag}(\gamma)$[1] and $e = [\ 1\ 1\ \ldots\ 1\ ]^T$ (see [7] for proof).

3 THE IDENTIFICATION METHOD

Our identification approach is summarised in Figure 1. It is mainly due to Yi and De Moor [8]. The main difference is that we have uncoupled the identification of the first order and the second order statistics. The first order parameters, γ and θ, and the model order N are determined by solving a nonnegative least squares problem [4] (see Section 3.1). The transition matrix P is obtained by unconstrained optimisation. This is briefly explained in Section 3.2.

3.1 Identification of the First Order Statistics

The identification of the first order parameters of an MMPP boils down to the determination of N, γ and θ so that F_{M} approximates the cumulative distribution function of the data, F_{d}, as well as possible.
The cumulative distribution function of the data sequence is a staircase function. It is computed as follows:

$$F_{\mathrm{d}}(l) = \frac{1}{m} \sum_{j=0}^{l} \sum_{k=1}^{m} \delta(a_k, j) \ ,$$

where $\delta(i,j)$ is the Kronecker delta and m is the number of data.

3.1.1 Determination of the components of a mixture distribution

The cumulative distribution function of the data, F_{d}, must be approximated by the cumulative distribution

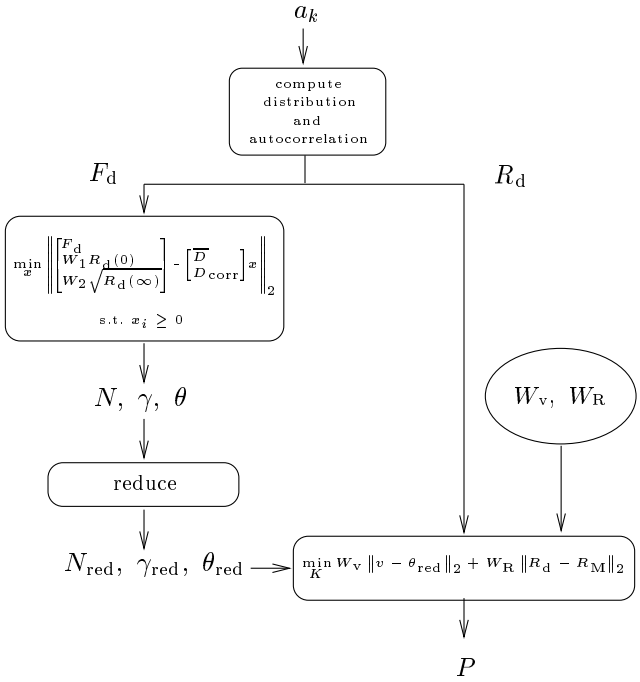


Figure 1: Identification of the MMPP where the cumulative distribution function is matched by solving the nonnegative least squares problem (5) and the autocorrelation is matched by unconstrained optimisation. The MMPP is found without a priori knowledge of the model order N. The contribution of the user is indicated by the oval.

[1] $\mathtt{diag}(\gamma) = \begin{bmatrix} \gamma_1 & 0 & \cdots & 0 \\ 0 & \gamma_2 & \ddots & 0 \\ \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & \gamma_N \end{bmatrix}$

function of an MMPP. This implies that F_{d} must be approximated by a nonnegative linear combination of cumulative distributions of Poisson processes (see (1)):

$$F_{\mathrm{d}} \approx D\theta \ , \tag{3}$$

where $(F_{\mathrm{d}})_l = F_{\mathrm{d}}(l-1)$ and $F_{\mathrm{d}} \in \mathbb{R}^{L+1}$ (L is the maximum number of arrivals of the given set of data). The columns of the matrix $D \in \mathbb{R}^{(L+1)\times N}$ correspond to the cumulative Poisson distribution functions F_{γ_i} $(i = 1, 2, \ldots, N)$.
If the states (i.e., the γ_i's) that form part of the MMPP which models a given data set, are given, then the matrix D is fixed and θ can be determined by solving the system of equations (3).
Without knowledge of the states or the number of states, the same approach yields N, γ and θ if in (3) D is replaced by an enlarged version $\overline{D}$. This matrix $\overline{D}$ has much more columns than D and the same number of rows. Each column $\overline{D_j}$ of $\overline{D}$ represents a *possible* state or *possible* Poisson cumulative distribution. In fact, the domain of possible γ_i's, e.g., $[0.001, L]$, is discretized with a step h in order to get a broad choice of states: $\gamma_j = \gamma_1 + (j-1)h$. The first Poisson parameter, γ_1, has to be chosen small enough in order to incorporate possible zero-valued a_k's. The computed cumulative distribution function F_{d} must be reconstructed as a nonnegative linear combination of the columns of $\overline{D}$:

$$\min_x \|F_{\mathrm{d}} - \overline{D}x\|_2, \text{ subject to } x_i \geq 0 \ . \tag{4}$$

Let x^* be the optimal value of x. The number of non-zero components of x^* gives the MMPP model order N. The indices of the non-zero components of x^* give the γ_i's. The θ_i's are given by the values of the non-zero components of x^*.

3.1.2 *Fitting the autocorrelation function at $n = 0$ and $n = \infty$*

Because the autocorrelation function $R_{\mathrm{M}}(n)$ of an MMPP at $n = 0$ and $n = \infty$ only depends on the first order parameters γ and θ:

$$\begin{cases} R_{\mathrm{M}}(0) &= \displaystyle\sum_{j=1}^{N} \theta_j(\gamma_j + \gamma_j^2) \\ R_{\mathrm{M}}(\infty) &= \displaystyle\left(\sum_{j=1}^{N} \theta_j\gamma_j\right)^2 \end{cases}$$

it is very important that these vectors are chosen in such a way that the autocorrelation function can be matched. Therefore, we add the following two rows to the matrix $\overline{D}$

$$D_{\mathrm{corr}} = \begin{bmatrix} W_1(\gamma_1+\gamma_1^2) & W_1(\gamma_2+\gamma_2^2) & \ldots & W_1(\gamma_{L+1}+\gamma_{L+1}^2) \\ W_2\gamma_1 & W_2\gamma_2 & \ldots & W_2\gamma_{L+1} \end{bmatrix}$$

and the vector F_{d} is extended with $W_1 R_{\mathrm{d}}(0)$ and $W_2\sqrt{R_{\mathrm{d}}(\infty)}$, where W_1 and W_2 are weighting factors. Solving the nonnegative least squares problem

$$\min_x \left\| \begin{bmatrix} F_{\mathrm{d}} \\ W_1 R_{\mathrm{d}}(0) \\ W_2\sqrt{R_{\mathrm{d}}(\infty)} \end{bmatrix} - \begin{bmatrix} \overline{D} \\ D_{\mathrm{corr}} \end{bmatrix} \cdot x \right\|_2 , \tag{5}$$
$$\text{subject to } x_i \geq 0$$

will then give the model order N, the Poisson parameters γ and the state distribution vector θ that approximate the cumulative distribution function of the data and the autocorrelation function at $n = 0$ and $n = \infty$.

3.1.3 *Reduction of the Model Order*

The computational complexity of the identification of the second order statistics (see Section 3.2) and the complexity of the queueing analysis highly depend on the model order. For the identification of the autocorrelation $\mathcal{O}(N^3)$ operations are needed per function evaluation in the optimisation algorithm and the computational complexity of the existing queueing analysis techniques is at least proportional to N^3 [1, 3]. Therefore it is advantageous to reduce the model order.
In our approach it is possible to reduce the model order after the identification of the first order parameters, i.e. before the stochastic matrix P is determined. The number of MMPP states is reduced by replacing similar states (with Poisson parameters γ_i close to each other) by one state. This new state is characterised by

$$(\gamma_{\mathrm{red}})_i \ = \ \sum_{j=1}^{n_i} \theta_j\gamma_j$$

and its contribution to the MMPP is given by

$$(\theta_{\mathrm{red}})_i \ = \ \sum_{j=1}^{n_i} \theta_j \ ,$$

where n_i is the number of states (γ_j, θ_j) that are grouped. The value of n_i is determined by the number of states that are present (the states for which $x_i^* \neq 0$) in a certain γ-range. This range depends on the value of the γ_j's. The smaller a certain γ_j, the smaller the region in which we look for states similar to the state characterised by that γ_j, because the variance of a Poisson distribution function with parameter γ_j is equal to γ_j.
In Figure 2 we show an example of this reduction approach. The circles are the results of (5). They represent the values of x^* as a function of the proposed γ's. The model order is equal to 16, the number of non-zero components of x^*. After reduction, we have 8 states, represented by the dash-dotted lines and the stars. The cumulative distribution functions of both models are shown in the second picture of Figure 2. The full line is the cumulative distribution function of the 16th order model, the dash-dotted line represents the cumulative distribution function of the reduced model

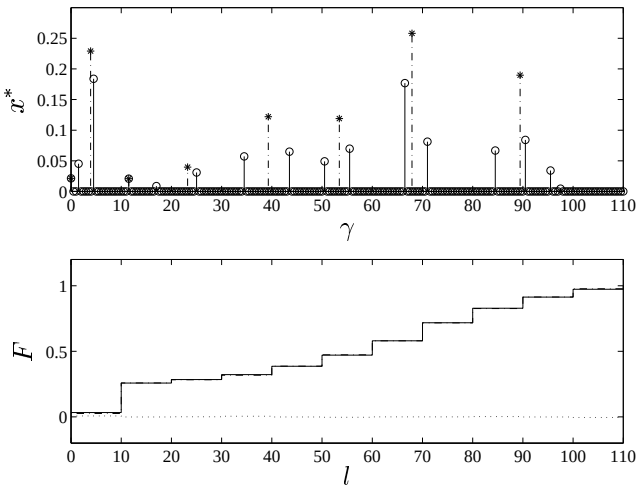


Figure 2: Reduction of a 16th order model (full line and circles) to a model of 8th order (dash-dotted line and stars). In the second picture the cumulative distribution function of both models are shown. The difference between both distribution functions is given by the dotted line.

of order 8. The difference between both distribution functions is given by the dotted line. It is clear that the reduction step does not introduce large errors in the cumulative distribution function.

3.2 Identification of the Second Order Statistics

For the identification of the second order statistics we use the following parameterisation from [8]

$$\Phi: \ \mathbb{R}^{N\times(N-1)} \to \mathbb{P}: \ K \to P = \Phi(K) \text{ with}$$

$$P = \Phi(K) = \begin{bmatrix} \frac{k_{11}^2}{1+\sum_{j=1}^{N-1} k_{1j}^2} & \cdots & \frac{k_{1(N-1)}^2}{1+\sum_{j=1}^{N-1} k_{1j}^2} & \frac{1}{1+\sum_{j=1}^{N-1} k_{1j}^2} \\ \vdots & & \vdots & \vdots \\ \frac{k_{N1}^2}{1+\sum_{j=1}^{N-1} k_{Nj}^2} & \cdots & \frac{k_{N(N-1)}^2}{1+\sum_{j=1}^{N-1} k_{Nj}^2} & \frac{1}{1+\sum_{j=1}^{N-1} k_{Nj}^2} \end{bmatrix}$$

in order to obtain an unconstrained optimisation problem. The matrix $K = [k_{ij}] \in \mathbb{R}^{N\times(N-1)}$ is the variable in the optimisation algorithm.

We want the autocorrelation function of the model, which is dependent on P, γ and θ (see equation (2)), to be as close as possible to the autocorrelation function of the data. Because the vectors θ and γ are already known from (5), only the stochastic matrix P has to be determined. To ensure that the resulting matrix P has θ (or a vector close to θ) as its left eigenvector corresponding to eigenvalue 1, the following unconstrained optimisation problem is solved:

$$\min_{K\in\mathbb{R}^{N\times(N-1)}} W_{\mathrm{v}} \ \|\theta - v\|_2 + W_{\mathrm{R}} \ \|R_{\mathrm{d}} - R_{\mathrm{M}}\|_2 \ ,$$

where v is the left eigenvector of $P = \Phi(K)$ corresponding to eigenvalue 1 and W_{v} and W_{R} are weighting factors. In this optimisation the state parameters γ remain constant. The essence of this method is that we keep the first order parameters from (5) and impose θ to be the left eigenvector of P corresponding to eigenvalue 1. The resulting method is summarised in Figure 1.

By eliminating the optimisation of the first order statistics we obtain an algorithm that is faster than that of Yi and De Moor [8]. For more details on the computational complexity of our method in comparison with the method of Yi and De Moor, we refer to [2].

4 CONCLUSIONS

In this paper we have discussed a new method for the identification of a Markov modulated Poisson process. The first order statistics are matched by solving a non-negative least squares problem and the second order statistics are identified by unconstrained optimisation. An important advantage of this method is the automatic determination of the model order which can be reduced if necessary.

REFERENCES

[1] C. Blondia and O. Casals. "Statistical Multiplexing of VBR Sources: A Matrix-Analytic Approach", Performance Evaluation, vol. 16, 1992, pp. 5-20.

[2] K. De Cock, T. Van Gestel and B. De Moor. "Stochastic System Identification for ATM Network Traffic Models: a Time Domain Approach", Internal report ESAT-SISTA/97-90[2], 1997.

[3] S.Q. Li and C.L. Hwang. "On the Convergence of Traffic Measurement and Queueing Analysis: A Statistical-Matching and Queueing (SMAQ) Tool", IEEE/ACM Transactions on Networking, vol. 5, 1997, pp. 95-110.

[4] C.L. Lawson and R.J. Hanson. *Solving Least Squares Problems*, Philadelphia: SIAM, Classics in Applied Mathematics series, vol. 15, 1995.

[5] S.Q. Li and C.L. Hwang. "Queue Response to Input Correlation Functions: Discrete Spectral Analysis", IEEE/ACM Transactions on Networking, vol. 1, 1993, pp. 522-533.

[6] S.Q. Li and C.L. Hwang. "Queue Response to Input Correlation Functions: Continuous Spectral Analysis", IEEE/ACM Transactions on Networking, vol. 1, no. 6, December 1993, pp. 678-692.

[7] P. Van Mieghem. "The Mapping of Traffic Parameters into a MMP(N)", Technical report, Alcatel Corporate Research Center, Antwerpen, Belgium, 1995.

[8] C. Yi and B. De Moor. "Traffic Identification of ATM Networks with Optimization Algorithms", Proceedings of the 35th IEEE Conference on Decision and Control, Kobe, Japan, December 1996, pp. 277-282.

[2] This report is available by anonymous ftp from *ftp.esat.kuleuven.ac.be* in the directory *pub/SISTA/decock/reports/97-90.ps.gz*