Input-Output Statistical Inference for Switching Processes

Alexandru Murgu
Department of Mathematical Information Technology
University of Jyvaskyla, FIN-40351 Jyvaskyla, FINLAND
E-mail: murgu@julia.math.jyu.fi

Oleg Gorokhov
Department of Mathematics and Mechanics
Ulyanovsk State University, Ulynovsk, RUSSIA
E-mail: gorokhov@math.jyu.fi

Innokenti V. Semoushin
Department of Mathematics and Mechanics
Ulyanovsk State University, Ulynovsk, RUSSIA
E-mail: isem@apt.ulsu.ru

ABSTRACT

This paper deals with the on-line change detection problem of Markov Modulated Poisson Processes (MMPPs) in the framework of switching systems. The dependency between the input and output statistical patterns in a black box viewpoint is analyzed. The abrupt change is treated as a transition from a state of the underlying Markov chain to another. We derive a generalized likelihood ratio (GLR) statistical test to detect the abrupt changes, simultaneously with the estimation of unknown distribution parameters using a likelihood scheme. The input-output mapping has a black-box representation in the class of ARMAX models.

Key words. Probabilistic Routing, Markov Modulated Poisson Process, Statistical Inference, Least Squares, Change Detection.

1. Introduction and Motivation

There are two main approaches to finding the relations between input and output of particular system. First one is a model-based approach, in which we should take into account by some reason the model of considered phenomena and describe it with a set of parameters. These parameters are to be estimated via information extraction from statistical experiment. Through this model with estimated parameters one can establish the relation between the input and output of the system. In the black box or statistical approach, we do not assume the inner system functionality be described by some model. Instead of it, models for input-output information flows are considered. The behaviour of the system is represented by all possible changes in the model parameters for output as a result of changing the model parameters for input.

The problem of statistical description of switching stochastic service systems is a topic interesting in itself. The model of input-transactions output of such system is the set of N traffic rate processes which are modelled as Markov modulated Poisson processes (MMPP). An MMPP is a doubly stochastic Poisson process whose intensity rate is modulated by an n-state underlying Markov chain. The MMPP is parameterized by a vector of intensities of the primary Poisson process $\Lambda \in R^n$ and the matrix of transition probabilities $\mathbf{Q} \in R^{n\times n}$ of Markov chain. The relation between parameters of input and output MMPP is to be found via jointly performed change detection and parameter estimation.

It is known that CUSUM test gives the infinimum of the worst mean delay for a class of statistical test, with given a priori rate of false alarms in the case of known parameters before and after the change. The application of this test to Markov modulated time series is considered in [7]. In the most realistic situation (for example network traffic) the parameters before and after the change are not known. On-line change detection algorithm is based on generalized likelihood ratio (GLR) test for MMPP model of traffic rate process.

2. Problem Statement

Let us define the topology of physical switching network represented as a graph $G_p = (N, A)$, where N is the set of nodes and A is set of links [6]. For each physical link $a \in A$, there are m_a channels with a finite capacity. The logical topology associated with the switching network is represented as a graph $G_l = (N, L)$, where L is the set of logical links. Each logical link $l \in L$ corresponds to a simple path in G_p. Let denote by W the set of origin-destination (O-D) pairs and by P, the set of all logical paths,while P_w is the set of logical paths that use $l = (i, j) \in L$.

path $p \in P$ x_p^t (packets/second), $t \in [0,T]$, where T is the time horizon. Let us denote by f_{ij}^t (packets/second) the aggregate rate of flows over the logical links connecting the nodes i and j. Each link $a_{ij} = (i,j) \in L$ has a capacity C_{ij}^t at time step $t, t \in [0,T]$.

For all commodities or traffic classes $w \in W$ we assume that input traffic can be represented as n-state MMPP characterized by the set of parameters $\left\{\Lambda_w^{in},\ Q_w^{in}\right\}_{w \in W},\ \Lambda_w^{in} \in R^n,\ Q_w^{in} \in R^{n \times n}$. The *departure flow traffic* is also modelled as n-state MMPP for each commodity with parameters $\left\{\Lambda_w^{dep}, Q_w^{dep}\right\}$ for all $w \in W$.

Taking into account capacity constraints of links in network that are reflecting in possible saturation behaviour of departure traffic for less priority commodities, it can be noted that the number of states for model of departure traffic may be reduced. In more complicated cases, some criteria of choosing the number of Markov chain states can be used.

The problem is to analyze both the input and the output traffic patterns determined by the possible changes of the regime that is associated with the intensity of the primary Poisson process. This problem of traffic analysis is considered in the framework of statistical black-box model identification. The problem of on-line detection in traffic regime can be formulated as follows. Let θ_0 - the value of intensity parameter *be known* before change occurs. Our goal is to determine the change of regime as quick as it is possible (estimating t_a alarm time), second is to find the estimate θ_1 after change. In the continuation of experiment it is assumed that estimated parameter θ_1 on the previous segment of traffic becomes *known* parameter θ_0 at the current considered segment of traffic till the new switch of regime is determined. In this sequential scheme the error of estimating of parameter θ_1 on the previous segment will not influence on the estimating parameter θ_1 at the current segment.

2.1. Probabilistic Routing and Statistical Modeling

The statistical viewpoint enable us to determine the behaviour of the system working under routing probabilities in each node for a given commodity. At each node, the routing probabilities are computed accordingly to an ordering of forward virtual paths (VPs) at this node, for each pair source-destination [4]. Specifically, let $|N(i,k)| = n_i^k$ be the number of outgoing links from node i handling traffic for a destination of the commodity k. Thus, we have in each node i a routing table for each commodity. Note that such an ordering provides the a *priority assignment* over the forward VPs (i,j) y, $y = n_i^k$ for the highest priority. Denote the priority of a VP (i,j) for commodity k by $\pi_{i,j}^k$. Denote by $P_j^k[empty]$ the conditional probability that a forward VP (i,j) becomes *available* first, for cells of commodity k, when

$$P_j^k[empty] = \frac{1 - \varrho_{ij}}{|N(i,k)| - \sum\limits_{x=1}^{|N(i,k)|} \varrho_{ix}} \tag{1}$$

where ϱ_{ij} is the utilization of VP (i,j). It can easily be proved that taking into account that busy periods of the forward VPs are mutually independent, exponentially distributed random variables with parameters $1 - \varrho_{ix}, \forall x \in N(i,k)$ and that the probability that VPs (i,j) has the smallest busy period is

$$P_j^k[\varrho_{ij} < \varrho_{ix}] = \frac{1 - \varrho_{ij}}{\sum\limits_{x=1}^{|N(i,k)|} (1 - \varrho_{ix})}, \text{ for all } x \neq j \tag{2}$$

Given the priority assignment and the utilization values of the VPs, we establish the routing probabilities. The routing algorithm will switch a packet belonging to k to the VP (i,j) with priority y if all of the links with higher priority are busy and this VP is available. The probability P_{ij}^k that edge (i,j) is selected for commodity k with priority y is

$$P_{ij}^k = (1 - \varrho_{ij}) \prod_{x:\pi_{i,x}^k > y} \varrho_{ix} + \left(\prod_{x=1}^{n_i^k} \varrho_{ix} \right) \frac{1 - \varrho_{ij}}{\sum\limits_{x=1}^{n_i^k} (1 - \varrho_{ix})} \tag{3}$$

The first term of the equation captures the contribution of the highest priority VPs that are busy for commodity k. The second term considers the case where all the forward VPS are busy and the link (i,j) becomes the first *empty* one.

A first measure of distance which is called the *expected proximity* (EP) of the node i to node j for the commodity k, and is defined as follows

$$\mathrm{EP}_i^{k,j} = \sum_{p \in P} \frac{1}{|p|} \prod_{(u,v) \in p} (1 - \varrho_{uv}) \tag{4}$$

where P denotes all possible forward paths from i to j for commodity k on the routing table, $|p|$ is the length of a path measured in hops, and (u, v) is a link between nodes u and v. The measure $\mathrm{EP}_i^{k,j}$ has relatively large values when there are few paths p between i and j, but they are short and lightly loaded or there exists a large number of paths between i and j.

A second measure is the *maximum proximity* (MP) of node i to node j for the commodity k is which is defined as follows

$$\mathrm{MP}_i^{k,j} = \max_{p \in P} \left\{ \frac{1}{|P|} \prod_{(u,v) \in P} (1 - \varrho_{uv}) \right\} \tag{5}$$

The maximum proximity provides a finer distinction of the nodes that have better forward paths to destination.

3. Markov Chain Approach to Traffic Class Migration

3.1. Markov Processes for Dynamic Modeling

Definition 3.1. (Markov Modulated Poisson Process (MMPP)). Let X_n be a n-state Markov chain with matrix of transition probabilities $\mathbf{Q}$. Doubly stochastic Poisson process is called Markov modulated Poisson process (MMPP) with vector of Poisson arrival rates $\{\lambda_i\}_{1 \le i \le n}$ if its arrival rate is λ_i when underlying Markov chain X_n is in the state i. □

Remark 3.1. MMPP with n-state is described by a pair of parameters $\Lambda = \{\lambda_i\}_{1 \le i \le n}$ and matrix of transition probabilities $\mathbf{Q}$. The probability q_{ij} denotes the probability that the chain, whenever in the state i, moves next (one unit of time later) into state j and can be referred to as a one-step transition probability. The square matrix $\mathbf{Q} = [q_{ij}]_{i \in S, j \in S}$, $S = \{1, \ldots, n\}$ is called *the one-step transition matrix.* Obviously, the following normalization constraint should be satisfied

$$\sum_{j \in S} q_{ij} = 1 \tag{6}$$

□

In order to describe the user class dynamic behaviour via MMPP models for discrete Markov chain, let us review the MMPP driven by *a continuous time Markov chain.* The differential equation formalism is an important step towards local scale characterization of events of the stochastic systems represented by the communication networks.

Definition 3.2. (Continuous Time Markov Chain (CTMC)). A stochastic process $\mathbf{X} = \{X_t\}$, $t > 0$ is continuous Markov chain with states in S if

$$q_{ij}(s) = \Pr[X(t+s) = j | X(t) = i], \ \forall i, j \in S, \ \forall s > 0 \tag{7}$$

□

Proposition 3.1. *The Chapman-Kolmogorov equation states that*

$$\mathbf{Q}(s+t) \;=\; \mathbf{Q}(s)\mathbf{Q}(t), \ \forall s > 0, \ t > 0 \tag{8}$$

where $\mathbf{Q}(0) = \mathbf{I}$. *If we set* $t = ds$, *we obtain the following equation*

$$\frac{d}{ds}\mathbf{Q}(s) \;=\; \mathbf{Q}(s)\mathbf{G} \tag{9}$$

where $\mathbf{G} = \frac{d}{ds}\mathbf{Q}(s)|_{s=0}$ *is the generator matrix of the Markov chain.* □

number of arrivals in an interval $(0,t]$ and J_t be the state of the Markov chain at the time step t. Define the matrix $\mathbf{Q}(n,t), n \in S,\ t > 0$ with elements $[q_{ij}(n,t)]_{i\in S, j\in S}$ as follows

$$q_{ij}(n,t) \quad = \quad \Pr[N_t = n,\ J_t = j | N_0 = 0,\ J_0 = i] \tag{10}$$

We call the matrix $\mathbf{Q(n,t)}$ the probability transition matrix of double process (N_t, J_t). Let Λ be the diagonal matrix associated with vector of intensities of MMPP, that is, $\Lambda = \text{diag}[\lambda_i]_{i\in S}$. The matrices $\mathbf{Q}(n,t),\ n \in S,\ t > 0$, satisfy the *Chapman-Kolmogorov* equations

$$\frac{d}{dt}\mathbf{Q}(n,t) \quad = \quad \mathbf{Q}(n,t)(\mathbf{G} - \Lambda) + \mathbf{Q}(n-1,t)\Lambda \tag{11}$$

$$\mathbf{Q}(0,0) \quad = \quad \mathbf{I},\ \forall n \in S,\ t > 0 \tag{12}$$

Multiplying (11) by z^n and summing for $n = 0, 1, \ldots,$ we obtain

$$\frac{d}{dt}Z\left\{\mathbf{Q}(z,t)\right\} \quad = \quad Z\left\{\mathbf{Q}(z,t)\right\}(\mathbf{G} - \Lambda) + zZ\left\{\mathbf{Q}(z,t)\right\}\Lambda \tag{13}$$

$$Z\left\{\mathbf{Q}(z,t)\right\}|_{t=0} \quad = \quad \mathbf{I} \tag{14}$$

where $Z\left\{\mathbf{Q}(z,t)\right\} = \sum\limits_{n=0}^{\infty} z^n \mathbf{Q}(n,t)$ is the Z-transform of matrix $\mathbf{Q}(n,t)$. Solving (13), we get the generating function of $\mathbf{Q}(n,t)$ as follows

$$Z\left\{\mathbf{Q}(n,t)\right\} \quad = \quad \exp[\mathbf{G}+(z-1)\Lambda t] \tag{15}$$

The expected number of arrivals number in the interval $(0,t]$ can be derived from the expression of the generating function, that is,

$$\mathbf{E}[N_t] \quad = \quad \pi\frac{\partial Z\left\{\mathbf{Q}(n,t)\right\}}{\partial z}|_{z=1}\mathbf{e} \tag{16}$$

where $\mathbf{e} = [1, \ldots, 1]^T$ and π is the steady state vector of the Markov chain such that $\pi\mathbf{Q} = \pi$. Finally, we obtain

$$\mathbf{E}[N_t] = \pi\Lambda\mathbf{e}t = \sum_{s\in S} \pi_i \lambda_i t \tag{17}$$

Remark 3.2. The component π_i of steady state vector π can be interpretated as asymptotically reached probability to be in the state $i \in S$ without regard of the history, or it can be expressed as

$$\lim_{t\to\infty} \Pr[X(t) = j | X(0) = i] = \pi_i,\ \forall i, j \in S \tag{18}$$

□

The type of developments presented in this section is widely used to characterize the queueing processes associated to the routing and switching device in the communication network.

3.2. Markov Chain Description of Traffic Migration

Let us consider that the traffic sources dynamics which are characterized by a MMPP process are subject to changes their operation regimes. We will call this feature, traffic class migration, and we attempt to use the available results from the theory of discrete Markov chains in order to encode the most significant feature of their behaviour.

Let us denote τ_j the first passage time to state j, defined as $\tau_j = \min\{n \geq 1, X(n) = j\}$. The probability distribution of τ_j is of great importance in what follows. Let us define

$$f(n,i,j) = \Pr_i(\tau_j = n),\ n \geq 1 \tag{19}$$

and

$$f(i,j) = \sum_{n\geq 1} f(n,i,j) = \Pr_i(\tau_j < \infty) \tag{20}$$

$$\Pr_i[\tau = \infty] = f(\infty, i, j) = 1 - f(i, j) \tag{21}$$

Therefore, $f(n, i, j),\ n = 1, 2, \ldots, \infty$, is the probability distribution of τ_j under the probability $\Pr_i$ (i.e., given the Markov chain starts in state i). The probabilities considered can be obviously expressed in a way that involves the Markov chain itself. Then, we have

$$\{\tau_j = 1\} = \{X(1) = j\} \tag{22}$$

$$\{\tau_j = n\} = \{X(m) \neq j,\ 1 \geq m \geq n-1,\ X(n) = j\}\ ,\ n \geq 2 \tag{23}$$

$$\{\tau_j = \infty\} = \{X(m) \neq j, m \geq 1\} \tag{24}$$

$$\{\tau_j < \infty\} = \{X(m) = j \text{ for at least one value of } m \geq 1\} = \bigcup_{m\geq 1} \{X(m) = j\}\ , \tag{25}$$

Therefore

$$f(1, i, j) = \Pr_j[X(1) = j] = p(i, j) \tag{26}$$

$$\begin{aligned} f(n, i, j) &= \Pr_j[X(m) \neq j,\ 1 \geq m \geq n-1,\ X(n) = j] && (27)\\ &= \sum_{i_m \neq j,\ (1\leq m\leq n-1)} p(i, i_1)p(i_1, i_2)\ldots p(i_{n-1}, j),\ n \geq 2 && (28)\end{aligned}$$

$$f(\infty, i, j) = \Pr_i[X(m) \neq j,\ m \geq 1] \tag{29}$$

$$f(i, j) = \Pr_i[X(m) = j \text{ for at least one value of } m \geq 1] = \Pr_i\Big[\bigcup_{m\geq 1} \{X(m) = j\}\Big] \tag{30}$$

The last equalities lead to

$$\Pr_i[X(n) = j] \geq f(i, j) \geq \sum_{m\geq 1} \Pr_i[X(m) = j] \tag{31}$$

for all $n \geq 1$, where

$$\sup_{n\geq 1} p(n, i, j) \geq f \geq \sum_{m\geq 1} p(m, i, j) \tag{32}$$

for all states i and j. It follows at once that $i \to j$ if and only if $f(i, j) > 0$ and that $i \leftrightarrow j$ if and only if $f(i, j)f(j, i) > 0$. Notice that by homogenity we also have

$$f(1, i, j) = \Pr[X(m+1) = j|X(m) = i] \tag{33}$$

$$f(n, i, j) = \Pr[X(m+l) \neq j,\ 1 \leq l \leq n-1,\ X(m+n) = j|X(m) = i],\ n \geq 2 \tag{34}$$

for any m for which the conditional probabilities are defined. The following results describe the qualitative properties of the Markov chain that are relevant in describing the traffic migration processes.

Theorem 3.1. *If the Markov chain starts in state i, then the probability of returning to i at least r times equals $[f(i, i)]^r$.* □

Corollary 3.1. *Assume the Markov chain starts in state i. If $f(i, i) = 1$, then probability of returning to i infinitely often is 1. If $f(i, i) < 1$, the probability of returning to i infinitely often is 0.*

Proof. The random event

$$A = \{ \text{ the Markov chain returns to } i \text{ infinitely often } \} \tag{35}$$

is the intersection of the decreasing sequence of random events

$$\{ \text{ the Markov chain returns to } i \text{ at least } r \text{ times}\}\ ,\ r \geq 1 \tag{36}$$

Therefore, the probability of A is equal to

$$\lim_{m\to\infty} [f(i, i)] = \begin{cases} 1, & \text{if } f(i, i) = 1, \\ 0, & \text{if } f(i, i) < 1. \end{cases} \tag{37}$$

□

$$p(n,i,j) = \sum_{m=1}^{n} f(m,i,j)p(n-m,j,j) \tag{38}$$

Proof. Intuitively we may argue as follows. To be in state j at the nth step, the Markov chain should reach that state for the first time at some time m, $1 \leq m \leq n$. After that happens, it should return to j in $n-m$ steps. The rigorous proof uses the strong Markov property. We have

$$p(n,i,j) = \Pr_i[X(n) = j] = \Pr_i[\tau_j \leq n,\ X(n) = j] \tag{39}$$

since the random event $\{X(n) = j\}$ implies the random event $\{\tau_j \leq n\}$

$$p(n,i,j) = \sum_{m=1}^{n} \Pr_i[\tau_j = m,\ X(n) = j] \tag{40}$$

Because $\{\tau_j \leq n\}$ is the union of disjoint random events $\{\tau_j = m\}, 1 \leq m \leq n$

$$p(n,i,j) = \sum_{m=1}^{n} \Pr_i[\tau_j = m,\ X(\tau_j + n - m) = j] \tag{41}$$

Since $\{\tau_j = m\} = \{\tau_j + n - m = n\}$

$$p(n,i,j) = \sum_{m=1}^{n} \Pr_i[\tau_j = m] \Pr_i[X(\tau_j + n - m) = j | \tau_j = m] \tag{42}$$

$$p(n,i,j) = \sum_{m=1}^{n} \Pr_i[\tau_j = m] \Pr_i[X(\tau_j + n - m) = j | X(\tau_j) = j,\ \tau_j = m] \tag{43}$$

By Corollary 3.1 of Theorem 3.1, then we have

$$p(n,i,j) = \sum_{m=1}^{n} f(m,i,j)p(n-m,j,j) \tag{44}$$

□

Theorem 3.3. *(Doeblin's formula). For any states i and j, we have*

$$f(i,j) = \lim_{s \to \infty} \frac{\sum_{n=1}^{s} p(n,i,j)}{1 + \sum_{n=1}^{s} p(n,j,j)} \tag{45}$$

Proof. Equation (38) yields

$$\sum_{n=1}^{s} p(n,i,j) = \sum_{n=1}^{s}\sum_{m=1}^{n} f(m,i,j)p(n-m,j,j) = \sum_{m=1}^{s} \big(f(m,i,j) \sum_{n=m}^{s} p(n-m,j,j)\big), \tag{46}$$

Thus,

$$\Big(1 + \sum_{n=1}^{s} p(n,j,j)\Big) \sum_{m=1}^{s} f(m,i,j) \geq \sum_{n=1}^{s} p(n,i,j) \geq \Big(1 + \sum_{n=1}^{s-s'} p(n,j,j)\Big) \sum_{m=1}^{s'} f(m,i,j), \tag{47}$$

for all $s' < s$. Since $1 + \sum_{n=1}^{s} p(n,j,j)$ dominates

$$\sum_{n=m}^{s} p(n-m,j,j) = 1 + \sum_{n=1}^{s-m} p(n,j,j) \tag{48}$$

and

$$\sum_{m=1}^{s} \big(f(m,i,j) \sum_{n=m}^{s} p(n-m,j,j)\big) \tag{49}$$

$$\sum_{m=1}^{s'}\left((f(m,i,j)\sum_{n=m}^{s-s'+m} p(n-m,j,j))\right) = \left(1+\sum_{n=1}^{s-s'} p(n,j,j)\right)\sum_{m=1}^{s'} f(m,i,j). \tag{50}$$

dividing by $1+\sum_{n=1}^{s} p(n,j,j)$ and letting first $s\to\infty$, then $s'\to\infty$, yields the equation (45). □

Remark 3.3. The above results are useful towards estimating theoretically the transition rates of the discrete Markov chain which is modulating the basic Poisson process. An extension of theses results inthe spirit of properties presented in Section 3.1 would be of major interest. □

4. Input-Output Representation of Statistical Behaviour

Let us consider an extended experiment that can be formulated as follows. Assume that for each commodity $w \in W$, we have the set of possible values for vector $\Lambda_w \in \mathcal{R}^n$ denoted by $\mathcal{L}_{in} \subset \mathcal{R}^n$. The estimated set of parameters is $\hat{\mathcal{L}}_{in} \subset \mathcal{R}^n$. Similarly, $\hat{\mathcal{L}}_{dep} \subset \mathcal{R}^n$ is the set of intensity parameters for the output traffic. The goal is to find the dependency relation between the sets $\hat{\mathcal{L}}_{in}$ and $\hat{\mathcal{L}}_{dep}$. Let $\hat{\Lambda}_{in} \in \hat{\mathcal{L}}_{in}$ and $\hat{\Lambda}_{dep} \in \hat{\mathcal{L}}_{dep}$.

$$\hat{\Lambda}_{dep} = \mathbf{f}(\hat{\Lambda}_{in}) \tag{51}$$

We concentrate on constructing linear relations as follows

$$\hat{\Lambda}_{dep} = \mathbf{G}_{LS}\hat{\Lambda}_{in} \tag{52}$$

where the equality holds in the least squares sense [5], that is,

$$\mathbf{G}_{LS} = \arg\min_{\mathbf{G}\in\mathcal{R}^{n\times n}} \sum_{\hat{\Lambda}_{in}\in\hat{\mathcal{L}}_{in},\hat{\Lambda}_{dep}\in\hat{\mathcal{L}}_{dep}} \|\hat{\Lambda}_{dep} - \mathbf{G}\hat{\Lambda}_{in}\|^2 \tag{53}$$

Finding the matrix $\mathbf{G}_{LS}$ will encode the dependencies between the statistical representations of the input and output traffic. Instead of proposing the model for the switching system and then estimating the model parameters, we consider models for input and the output traffic and in the extended experiment, the changes in the output parameters will reflect changes in the input parameters that are subject of statistical estimation [2].

The matrix $\mathbf{G}_{LS}$ can be used further in some other ways. First, the explicit usage is in the prediction of the output parameters for given vectors of input parameters which in fact may not belong to the set $\hat{\mathcal{L}}_{in}$ of extended experiment. In this case, it can be seen as an extrapolation, or as interpolation if a given vector is included into $\hat{\mathcal{L}}_{in}$. In both cases, this extended experiment may be treated as a learning process during which the matrix $\mathbf{G}_{LS}$ is being constructed. In the next stage, $\mathbf{G}_{LS}$ is used without modifications as a given model for dependencies between the input and the output. The matrix $\mathbf{G}_{LS}$ can be estimated as off-line linear regression over the sets $\hat{\mathcal{L}}_{in}$ and $\hat{\mathcal{L}}_{dep}$.

The second possible usage of $\mathbf{G}_{LS}$ is in solving the inverse problem of finding the values of the input parameters corresponding to given output parameters. This problem can be solved in the context of adaptive control (or regulation) of the switching system in order to achieve a desired level of performance with respect to given regulation criteria, which in the most situations can be expressed as a desired set of output traffic parameters.

5. Derivation of GLR Test

In this section we derive GLR algorithm that is effectively used under the Poisson distribution hypothesis for the basic stochastic process of traffic generation [1]. We consider a parametric family of Poisson distributions $\mathcal{P} = \{\Pr_\lambda\}_{\lambda\in\Theta}, \Theta \subset R^n$. Let as usual, $\mathcal{Y}_1^N$ be a data sample of size N. The goal is to distinguish the distance between two hypothesis

$$\begin{aligned}\mathbf{H}_0 &= \{L = \Pr_{\lambda_0}\}\\ \mathbf{H}_1 &= \{L = \Pr_{\lambda_1}\}\end{aligned} \tag{54}$$

Since the Poisson process can be represented as the sum of i.i.d. exponentially distributed random variables, its density is given by

$$p_{n,\lambda}(t) = \lambda e^{-\lambda t}\frac{(\lambda t)^{n-1}}{(n-1)!},\ t \geq 0 \tag{55}$$

construction of GLR method and it is defined as follows

$$S_j^N(\lambda_0,\lambda_1) \quad = \quad \ln\frac{p_{N,j,\lambda_1}(\mathcal{Y}_j^N)}{p_{N,j,\lambda_0}(\mathcal{Y}_j^N)} \tag{56}$$

where $p_{N,j,\lambda_i}(\mathcal{Y}_j^N), i=1,2$ is the joint p.d.f. for part of the data sample for time steps $j, j+1, \ldots, N$.

Remark 5.1. In the notation of joint p.d.f., there are subscripts N and j that explicitly denote the dependency p.d.f. of time interval of considered part of N-size sample $\mathcal{Y}_1^N$. □

Under the independence assumption of observations, equation (56) becomes

$$S_j^N(\lambda_0,\lambda_1) \quad = \quad \sum_{t=j}^{N}\ln\frac{p_{n,\lambda_1}(y_t)}{p_{n,\lambda_0}(y_t)} \tag{57}$$

Let us derive the expression for joint p.d.f of $\mathcal{Y}_1^N$. We have

$$\begin{aligned} p_{N,j,\lambda}(\mathcal{Y}_j^N) &= \prod_{i=j}^{N}\lambda e^{-\lambda y_i}\frac{(\lambda y_i)^{(i-1)}}{(i-1)!} \\ &= \lambda^{N-j+1}e^{-\lambda(\sum_{i=j}^{N} y_i)}\frac{\lambda^{(j-1)+j+\cdots+N}}{(j-1)!j!\ldots N!}\prod_{i=j}^{N}y_i^{(i-1)} \\ &= \pi\lambda^{\alpha}e^{-\lambda\beta} \end{aligned} \tag{58}$$

where in the last relation, the following notations are used

$$\pi \quad = \quad \prod_{i=j}^{N} y_i^{(i-1)} \tag{59}$$

$$\alpha \quad = \quad N-j+1+\frac{1}{2}(N(N-1)-(j-1)(j-2)) \tag{60}$$

$$\beta \quad = \quad \sum_{i=j}^{N} y_i \tag{61}$$

The maximum of $p_{N,j,\lambda}(\mathcal{Y}_j^N)$ with respect to the parameter λ can be easily obtained after solving the equation

$$\frac{\partial p_{N,j,\lambda}(\mathcal{Y}_j^N)}{\partial\lambda} = 0 \tag{62}$$

Let us explicitly solve the equation (62) which can be rewritten as

$$\frac{\partial p_{N,j,\lambda}(\mathcal{Y}_j^N)}{\partial\lambda} = \pi[\alpha\lambda^{(\alpha-1)}e^{-\lambda\beta} - \lambda^{\alpha}\beta e^{-\lambda\beta}] = 0 \tag{63}$$

Consequently,

$$\hat{\lambda}(N,j,\mathcal{Y}_1^N) = \arg\sup_{\lambda}[p_{N,j,\lambda}(\mathcal{Y}_j^N)] = \frac{\alpha-1}{\beta} \tag{64}$$

In order to formulate GLR method in the case of *unknown* parameter λ_1 after change, the equation (56) of the log-likelihood ratio should be rewritten as

$$S_j^N(\lambda_0) \quad = \quad \ln\sup_{\lambda_1}\left[\frac{p_{N,j,\lambda_1}(\mathcal{Y}_j^N)}{p_{N,j,\lambda_0}(\mathcal{Y}_j^N)}\right] \tag{65}$$

which using (64) can be written as follows

$$S_j^N(\lambda_0) \quad = \quad \ln\left[\frac{p_{N,j,\hat{\lambda}(N,j,\mathcal{Y}_1^N)}(\mathcal{Y}_j^N)}{p_{N,j,\lambda_0}(\mathcal{Y}_j^N)}\right] \tag{66}$$

$$S_j^N(\lambda_0) = \ln\Big[\Big(\prod_{i=j}^{N}\hat{\lambda}e^{\hat{\lambda}y_i}\frac{(\hat{\lambda}y_i)^{i-1}}{(i-1)!}\Big)\Big(\prod_{i=j}^{N}\lambda_0 e^{\lambda_0 y_i}\frac{(\lambda_0 y_i)^{i-1}}{(i-1)!}\Big)^{-1}\Big]$$
$$= \ln\Big[\Big(\frac{\hat{\lambda}}{\lambda_0}\Big)^{N-j}\frac{e^{(N-j)\hat{\lambda}}}{e^{(N-j)\hat{\lambda_0}}}\frac{\hat{\lambda}^{(j-1)+j+\cdots+N}}{\lambda_0^{(j-1)+j+\cdots+N}}\Big]$$
$$= (N-j+\sum_{i=j-1}^{N})\ln\Big(\frac{\hat{\lambda}}{\lambda_0}\Big)+(N-j)(\hat{\lambda}-\lambda_0) \quad (67)$$

where $\hat{\lambda}=\hat{\lambda}(N,j,\mathcal{Y}_1^N)$ and is given by (64). Finally, the GLR statistics are given by

$$g_k(S_1^k) = \max_{j\in[1,k]} S_j^k(\lambda_0), \;\; k>0 \quad (68)$$

The stopping time of GLR test can be written as

$$t_a^{GLR} = \min_{k>0}\{g_k(S_1^k)\geq h_{GLR}\} \quad (69)$$

The threshold h_{GLR} is an adjustable parameter of the proposed test. The choice of this threshold is based on the investigation of asymptotic behaviour of the likelihood ratio. Let α be a size of test $g_k(S_1^k)$ when the asymptotic distribution of $-2\ln g_k(S_1^k)$ under hypothesis $\mathbf{H}_0$ is $\chi^2(1)$. The following relation determines the *in-control state* of $\mathcal{Y}_1^N$

$$-2\ln g_k(S_1^k)\leq \chi^2_{1-\alpha}(1) \quad (70)$$

where $\chi^2_{1-\alpha}(1)$ is the $(1-\alpha)$-th quantile of the χ^2-distribution with 1 degree of freedom.

Remark 5.2. The same asymptotic behaviour of $g_k(S_1^k)$ can be observed in the vector case of parameter. In this case

$$L\Big[-2\ln g_k(S_1^k)\Big]\Rightarrow \chi^2(l), \text{ when } L(S_1^k)=\mathbf{P}_{\lambda_0} \quad (71)$$

where l is the dimension of parameter space. □

6. Concluding Remarks

In this paper we have shown the main concepts from the telecommunication networks traffic theory and system modelling and their interplay that are useful in order to describe in an unified manner the input-output transfer processes occurring in the switching systems. Some experimental studies are to be considered in the future towards validating the theoretical framework that has been developed.

References

[1] M. Basseville, I.V. Nikiforov, *Detection of Abrupt Changes: Theory and Application*, Prentice-Hall, Englewoods Cliffs, NJ, 1993.

[2] Y. Bengio, P. Frasconi, "Input/Output HMMs for Sequence Processing", *Research Report*, AT&T Bell Laboratories, Holmdel, NJ, September 4, 1995.

[3] M. Iosifescu, *Finite Markov Processes and Their Applications*, John Wiley and Sons, New York, 1980.

[4] D. Medhi, "Multi-Hour, Multi-Traffic Class Network Design for Virtual Path-Based Dynamically Reconfigurable Wide-Area ATM Networks", *IEEE/ACM Transactions on Networking*, vol.3, no. 6, pp. 809-818, 1995.

[5] E. Mosca, *Optimal, Predictive and Adaptive Control*, Prentice-Hall, Englewoods-Cliffs, NJ, 1995.

[6] A. Murgu, *Optimization of Telecommunication Networks. Lecture Notes*, University of Jyväskylä, 1999.

[7] J. Rittwik, D. Subhrakanti, "Change Detection in Teletraffic Models", in *Proceedings of the 36th IEEE CDC*, San Diego, pp. 3984-3989, 1997.