Technical University of Budapest, Department of Telecommunication

Stochastic models and telecommunication applications

(Theses of Ph.D dissertation)

author Levente Bodrog
scientific supervisor Prof. Dr. Miklós Telek

Budapest, 2010

 bodrog@hit.bme.hu

1 Introduction

The thesis considers the flexibility of second order Markovian arrival processes and evaluate the performance of load-balanced switches. They are a widely considered part of the stochastic modeling and the telecommunication field.

According to the subjects the thesis can be divided into two main parts. In the first part the correlation bounds of the canonical second order acyclic Markovian arrival processes are given. These bounds are used, in [J2], to prove the equivalence of arrival processes having marginal distribution with second order rational Laplace transform.

Based on the minimal canonical representation a second order Markovian arrival process (MAP(2)) fitting method is developed in which the MAP(2) bounding surface is simply decomposed into tractable subsurfaces for the numerical optimization. This decomposition helps the numerical methods to work more reliable than in case of general usage, with or without considering the exact, but complex, MAP(2) bounding surface.

We also developed a joint density function fitting algorithm utilizing the minimal property of the canonical form. The main idea behind this approach is a numerically efficient, matrix manipulation based, algorithm to compute the Euclidean distance of two joint density functions of two different MAPs. We go forward with this approach as we can determine the Euclidean distance of infinite joint density functions too.

The second part of the thesis covers the performance evaluation of load-balanced (LB) switches. In the first step [1] analyzed the cell (equal size packet) loss probability of the LB switch, proposed in [3] and [4], due to finite buffer capacity. We start with the performance evaluation of the switch with variable size packets developing the detailed model of the switch. It turned out that the detailed model leads to Markov models of large state space – with exponential complexity $\left(O\left(N^N\right)\right)$ in terms of the switch size (N). As a response to the state space explosion we come up with an ON/OFF input process based approximate model of the switch. While the state space of the ON/OFF model is smaller than the detailed one it has still exponential complexity $\left(O\left(2^N\right)\right)$. At this point a restrictive assumption is applied, namely identical input processes are considered, by which the size of the state space remains linear with the switch size.

As the main problem with the packet loss is the longer packet reassembly delays, we also proposed a packet loss minimization technique inside the LB switch.

2 Research objectives

The thesis proposes to analyze and to understand the behavior of the second order Markovian arrival processes as well as to utilize the knowledge on them.

Furthermore it also an aim to give the detailed model and to improve the performance of a new, scalable switching architecture, the load-balanced switches.

2.1 Second order Markovian arrival processes

Due to their compactness second order arrival processes play an important role in stochastic modeling and here we give their exact correlation bounds. This is allowed by the introduction of the new parametrization of the processes involving only four parameters derived from the four parameters defining the minimal canonical acyclic MAP(2) representation.

The determination of the correlation bounds completes the moment bounds of MAP(2)s in terms of the basic moment set. This recognition led us to utilize this knowledge in MAP(2) fitting.

First we introduce the decomposed numerical fitting method in which we simply decompose the optimization based fitting into, evidently given, smaller and easier to solve subproblems.

In the next step we develop a MAP reduction method which is able to reduce MAP sizes in an optimization based way. At this optimization based step we utilize again our knowledge on the minimal representation of MAP(2)s.

2.2 Load-balanced switches

As a new and promising switch architecture, we turn our attention to load-balanced switches. They seem scalable due to their deterministic and simple control and they can provide high throughput even for non uniform traffic. Contrary to the initial performance evaluation of the switching architecture, assumed to be equipped with infinite buffers, the authors of [1] point out that in the more realistic hardware setting, with finite buffers, there can be cell loss, for which an analytical model is created. In our work we go forward and do the performance evaluation of the load-balanced switches with variable size packets and also propose a packet loss minimization technique.

3 Research methodology

In the present work we do analytical considerations and Markovian modeling as well. For the determination of the correlation bounds of the second order MAPs (MAP(2)) we do mathematical analysis. While in case of the decomposed numerical MAP(2) fitting we used elementary geometry and numerical optimization. During MAP reduction we used matrix analytical methods and geometry to determine the Euclidean distance of the MAP joint density functions.

The validation of the fitting methods were performed by the comparison of the cumulative distribution functions (CDF), the correlation structures and the queueing behaviors. The CDFs and the correlation structures are determined using elementary computations while the queueing behavior is observed in an M/D/1 system using matrix analytic methods for the solution of the M/G/1 type structure.

During the performance evaluation of the load-balanced switches we assume Markovian behavior by which we can also use the elementary matrix analytical computations for the solution of these models.

The results of the analytical models of the load-balanced switches are compared to simulation results which are gained with a simulation tool developed in `c++`.

4 New results

The contribution of the present thesis can be arranged into two main groups. The first focuses on the second order Markovian arrival processes (MAP(2)s) and the other focuses on load-balanced (LB) switches.

The first group discusses the exact correlation bounds and the fitting of MAP(2)s. The second group can be divided into two further subgroups. In the first one the performance evaluation, namely the packet loss analysis, of the LB switches with heterogenous input processes is given while in the second subgroup the performance evaluation, namely the packet loss and packet delay analysis, and a packet loss minimization technique of the LB switches with homogenous input processes are proposed.

4.1 Second order Markovian Arrival Processes

4.1.1 Markovian canonical form of second order matrix exponential processes

Given an arrival process $X(t) = \{X_0, X_1, \ldots, X_n\}$ with joint density function

$$f(\mathbf{x}) = f(x_0, x_1, \ldots, x_n) = \boldsymbol{\pi} e^{\mathbf{D}_0 x_0}\mathbf{D}_1 e^{\mathbf{D}_0 x_1}\mathbf{D}_1 \cdot \cdots \cdot e^{\mathbf{D}_0 x_n}\mathbf{D}_1 \mathbb{1}, \qquad (1)$$

represented by the matrix pair $(\mathbf{D}_0, \mathbf{D}_1)$, we say that it is a second order Markovian arrival process if $f(\mathbf{x})$ is a real density function, $\mathbf{D}_0$ is a transient Markovian generator and $\mathbf{D}_1 \geq 0$ such that $-\mathbf{D}_0\mathbb{1} = \mathbf{D}_1\mathbb{1}$. $\boldsymbol{\pi}$ is the solution of the linear system of equations $\boldsymbol{\pi}(-\mathbf{D}_0)^{-1}\mathbf{D}_1 = \boldsymbol{\pi} \quad \boldsymbol{\pi}\mathbb{1} = \mathbb{1}$, i.e., it is the steady state solution of the discrete time Markov chain embedded after an arrival. $\mathbb{1}$ is the appropriate size column vector of ones.

The second order Markovian arrival processes, with matrix representation $(\mathbf{D}_0, \mathbf{D}_1)$, has second order phase-type distributed (PH(2)) marginals. A PH(2) distribution has the cumulative distribution function

$$F(t) = 1 - \boldsymbol{\pi} e^{\mathbf{D}_0 t}\mathbb{1} = 1 - \begin{pmatrix} p & 1-p \end{pmatrix} e^{\begin{pmatrix} -\lambda_1 & \lambda_1 \\ 0 & -\frac{\lambda_1}{\alpha} \end{pmatrix} t} \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \qquad (2)$$

The correlation structure of MAP(2) is

$$\operatorname{corr}(X_0, X_k) = \gamma^k \frac{p(p+\alpha-1)}{p(p-2)-\alpha^2}, \qquad (3)$$

where γ gives the shape of the geometrically decaying correlation structure.

An order m Markovian arrival process can be defined by the so called basic moment set containing m^2 parameters [J1].

In the second order case the four equivalent parameters used in (2) and (3) are $\{\lambda_1, p, \alpha, \gamma\}$. It can be shown that this tuple can be transformed into several others among $\{\lambda_1, p, \alpha, \gamma\} \equiv \{\lambda_1, \lambda_2, a, b\} \equiv \{\mu_1, n_2, n_3, \gamma\}$. The simple reason of transforming the tuple, in the sequel, is to gain the most suitable parameter set

- $\{\lambda_1, p, \alpha, \gamma\}$ to give the defining expressions of the MAP(2) set in (2) and (3) for the determination of the correlation bounds in Table 1,
- $\{\lambda_1, \lambda_2, a, b\}$ to express the minimal canonical form of the MAP(2) set in (4) and (5) and
- $\{\mu_1, n_2, n_3, \gamma\}$ to visualize the MAP(2) correlation bounds in a natural way in Figure 1.

Now using the definitions in (2) and (3) we can give the correlation bounds of the MAP(2) class in the following thesis.

Thesis 1.1. *I give the correlation (γ) bounds of the second order Markovian arrival processes in Table 1 in terms of the pair (p, α) defining the APH(2) marginal distribution of the arrival process. [J2]*

There is a canonical form of second order acyclic MAP(2)s (AMAP(2)s) given which at first sight appears to be even more constrained than arbitrary AMAP(2)s, due to an enforced zero element in matrix $\mathbf{D}_1$ besides the upper triangular matrix $\mathbf{D}_0$. But it is shown in [J2] that every MEP(2) can be transformed into this canonical form. This transformation lets prove the equivalence of classes AMAP(2), MAP(2) and MEP(2) in [J2].

For the determination of the general canonical form we use the parameter set $\{\lambda_1, \lambda_2, a, b\}$. Depending on the sign of the correlation parameter there are two variants of the canonical form.

Definition 1. The first canonical representation of MAP(2)s is defined as

$$\mathbf{D}_0 = \begin{pmatrix} -\lambda_1 & (1-a)\lambda_1 \\ 0 & -\lambda_2 \end{pmatrix}, \qquad \mathbf{D}_1 = \begin{pmatrix} a\lambda_1 & 0 \\ (1-b)\lambda_2 & b\lambda_2 \end{pmatrix}. \tag{4}$$

The second canonical form is given by

$$\mathbf{D}_0 = \begin{pmatrix} -\lambda_1 & (1-a)\lambda_1 \\ 0 & -\lambda_2 \end{pmatrix}, \qquad \mathbf{D}_1 = \begin{pmatrix} 0 & a\lambda_1 \\ b\lambda_2 & (1-b)\lambda_2 \end{pmatrix}, \tag{5}$$

where $0 < \lambda_1 \leq \lambda_2$, $0 \leq a \leq 1$ and $0 \leq b \leq 1$. Additionally, we require that

- $a, b \neq 1$ in the first canonical form (for recurrency) and
- $b \neq 0$ in the second canonical form (for recurrency) and
- $\lambda_1 \neq \lambda_2$, if $a = 1$ in the second canonical form.

For correlated processes, a and b must be nonzero.

Several characteristics of the canonical form depend only on parameters a and b

Corollary 1. *The correlation parameter γ of the first canonical form is given by*

$$\gamma = ab. \tag{6}$$

Table 1: Lower and upper γ bounds for the three possible combinations of the (p, α) parameters

condition	γ bound	
	lower	upper
$1-p<\alpha$	$-\frac{1-p}{p+\alpha}$	$\frac{\alpha+p(\alpha+p-1)-2\sqrt{p\alpha(-1+p+\alpha)}}{(p+\alpha)^2}$
$\alpha>1-p \wedge \frac{1-p}{p+\alpha}<1$	$-\frac{1-p}{p+\alpha}$	1
$\alpha>1-p \wedge 1 \leq \frac{1-p}{p+\alpha}$	$\frac{p}{p+\alpha-1}$	1

The correlation parameter γ of the second canonical form is given by

$$\gamma = -ab. \tag{7}$$

The phase probability vector at stationary arrival epochs in case of the first canonical form is

$$\boldsymbol{\pi} = \begin{pmatrix} \frac{1-b}{1-ab} & \frac{b-ab}{1-ab} \end{pmatrix}, \tag{8}$$

In case of the second canonical form, it is

$$\boldsymbol{\pi} = \begin{pmatrix} \frac{b}{1+ab} & 1-\frac{b}{1+ab} \end{pmatrix}. \tag{9}$$

The $\{\lambda_1, \lambda_2, a, b\}$ based canonical forms have two advantages. They are simple and their parameter bounds are straightforward, as λ_1, λ_2 are positive rates and a, b are probabilities. However it is only the new parametrization using the tuple $\{\lambda_1, p, \alpha, \gamma\}$ that enables to show the identity relationship between MEP(2)s and AMAP(2)s in [J2]. Then the correlation bounds of the second order arrival processes are given in Table 1 in terms of $\{\lambda_1, p, \alpha, \gamma\}$ while the moment bounds of the PH(2) marginal of the class is known and is given, e.g., in [2].

4.1.2 Canonical form based second order Markovian arrival process fitting

Fitting is always a complex problem even if one knows the fitting model exactly. In case of MAP the problem is more complex as the borders of the MAP class is not known in the general case, i.e., in general the fitting by MAP results in a high dimensional general optimization problem. Knowing the border one can apply optimization constrained on the border but in case of non-analytical or complex border the constrained optimization can also lead to numerically instable methods or untractable results.

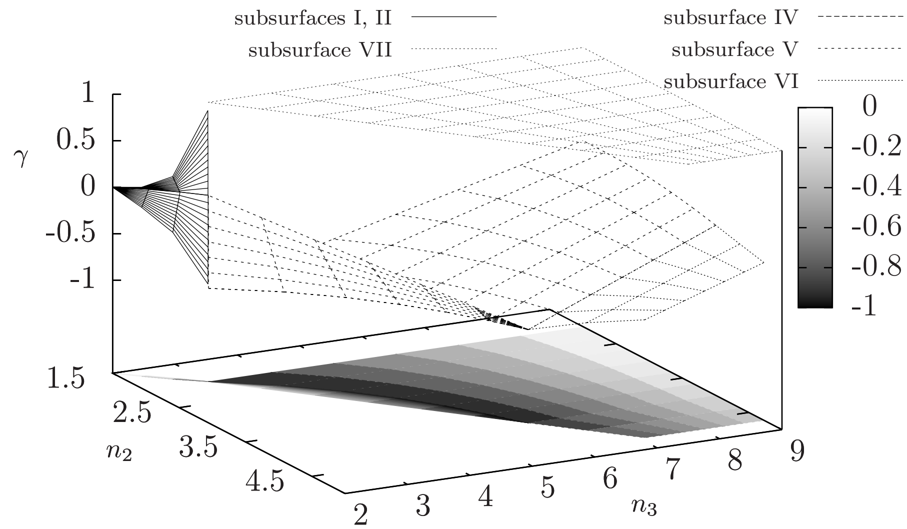

Figure 1: The bounds of the MAP(2) class in the normalized moment space

Decomposed numerical fitting method After having the exact moment bounds of the second order arrival processes we can utilize this knowledge to improve MAP(2) fitting. One evidently given utilization is to improve the general optimization based approach by constraining the optimization on the MAP(2) bounds in the moment space, or in any equivalent space to which the moment bounds can be transformed. In the followings we do the moment bound based approach and go a bit forward than constrained optimization and do piecewise the constrained optimization.

Thesis 1.2. *I improved the effectiveness of the moment (or any equivalent) distance optimization based MAP(2) fitting algorithms by the decomposition of the MAP(2) boundaries, given in Table 1, into their naturally given subsurfaces on which the constrained optimization is numerically tractable as all the subsurfaces are given analytically. [C11]*

The different γ bounds appearing in Table 1 implies the natural decomposition of the moment distance optimization based MAP(2) fitting method. Indeed, the decomposed fitting method considers these naturally got subsurfaces depicted in Figure 1.

Given an outer (non MAP(2)) point in the normalized moment space, using the knowledge on the γ bounding surface, the decomposed fitting method utilizes that the distance between an outer point and a region lies on the border of that region, i.e., an optimization constrained on the bound is performed. As in case of the MAP(2) correlation the bound is a piecewise function, the decomposed fitting method go forward and considers the subsurfaces of the bounds separately. The method goes through the bounding subsurfaces, given in Figure 1, minimizes the distance between each subsurface and the outer point and it results in the set of distances between the input and the subsurfaces. Finally it returns with the minimum of the set of the minima. This is expressed briefly in Algorithm 1.

Algorithm 1 `decomposed numerical fitting method`

INPUT: $M = (n_2, n_3, \gamma)$ the outer point
OUTPUT: $(\mathbf{D}_0, \mathbf{D}_1, d)$ the closest MAP(2) and its distance from M

$d = \infty$
while there is unchecked subsurface **do**
  find the closest point ($\tilde{M}$) on the actual surface from M
  calculate the Euclidean distance of $\tilde{M}$ and M $\tilde{d} = d(M, \tilde{M})$
  if $\tilde{d} < d$ **then**
    $d = \tilde{d}$
    $\hat{M} = \tilde{M}$
  end if
  consider the “next” subsurface
end while
$(\mathbf{D}_0, \mathbf{D}_1) \leftarrow \hat{M}$
return $(\mathbf{D}_0, \mathbf{D}_1, d)$

Fitting high order MAPs with low order MAPs Assume that there are two MAPs given by their matrix representations $(\mathbf{D}_0, \mathbf{D}_1)$ and $(\mathbf{G}_0, \mathbf{G}_1)$ with finite joint density functions $f(\cdot)$ and $g(\cdot)$ and with stationary phase distributions $\boldsymbol{\pi}$ and $\boldsymbol{\gamma}$ respectively. Using $\mathbb{1}$, the appropriate size column vector of ones, we can express the integral of the multiplication of the finite joint density functions as

$$\begin{aligned}
L_{fg}(n) =& \int_{\mathbf{x}} f(x_1, x_2, \cdots, x_n) g(x_1, x_2, \cdots, x_n) d\mathbf{x} \\
=& \int_{\mathbf{x}} \left(\boldsymbol{\pi} e^{\mathbf{D}_0 x_1} \mathbf{D}_1 e^{\mathbf{D}_0 x_2} \mathbf{D}_1 \cdots e^{\mathbf{D}_0 x_n} \mathbf{D}_1 \mathbb{1}\right) \\
& \otimes \left(\boldsymbol{\gamma} e^{\mathbf{G}_0 x_1} \mathbf{G}_1 e^{\mathbf{G}_0 x_2} \mathbf{G}_1 \cdots e^{\mathbf{G}_0 x_n} \mathbf{G}_1 \mathbb{1}\right) d\mathbf{x} \\
=& \int_{\mathbf{x}} (\boldsymbol{\pi} \otimes \boldsymbol{\gamma}) \left(e^{\mathbf{D}_0 x_1} \otimes e^{\mathbf{G}_0 x_1}\right) (\mathbf{D}_1 \otimes \mathbf{G}_1) \times \dots \\
& \times \left(e^{\mathbf{D}_0 x_n} \otimes e^{\mathbf{G}_0 x_n}\right) (\mathbf{D}_1 \otimes \mathbf{G}_1) (\mathbb{1} \otimes \mathbb{1}) d\mathbf{x} \\
=& (\boldsymbol{\pi} \otimes \boldsymbol{\gamma}) \left(\int_{x_1} e^{\mathbf{D}_0 x_1} \otimes e^{\mathbf{G}_0 x_1} dx_1\right) (\mathbf{D}_1 \otimes \mathbf{G}_1) \times \dots \\
& \times \left(\int_{x_n} e^{\mathbf{D}_0 x_n} \otimes e^{\mathbf{G}_0 x_n} dx_n\right) (\mathbf{D}_1 \otimes \mathbf{G}_1) (\mathbb{1} \otimes \mathbb{1}) \\
=& \underbrace{(\boldsymbol{\pi} \otimes \boldsymbol{\gamma})}_{\boldsymbol{\nu}} \underbrace{\left(-(\mathbf{D}_0 \oplus \mathbf{G}_0)^{-1} (\mathbf{D}_1 \otimes \mathbf{G}_1)\right)^n}_{\mathbf{N}^n} \underbrace{(\mathbb{1} \otimes \mathbb{1})}_{\mathbb{1}} \\
=& \boldsymbol{\nu} \mathbf{N}^n \mathbb{1}.
\end{aligned} \tag{10}$$

Thesis 1.3. *Using the expression in* (10) *I give an, Euclidean distance based, MAP joint density function fitting method for both the finite and the infinite case. [C11]*

The minimization of the Euclidean distance of the joint density functions ($f(\cdot)$ and $g(\cdot)$) of the two MAPs is

$$\begin{aligned}
& \min_{\mathbf{G}_0, \mathbf{G}_1} d\big(f(\mathbf{x}), g(\mathbf{x})\big) = \\
& \quad = \min_{\mathbf{G}_0, \mathbf{G}_1} \int_{\mathbf{x}} \big(f(\mathbf{x}) - g(\mathbf{x})\big)^2 d\mathbf{x} \\
& \quad = \min_{\mathbf{G}_0, \mathbf{G}_1} \left(\int_{\mathbf{x}} f(\mathbf{x}) f(\mathbf{x}) d\mathbf{x} + \int_{\mathbf{x}} g(\mathbf{x}) g(\mathbf{x}) d\mathbf{x} - 2 \int_{\mathbf{x}} f(\mathbf{x}) g(\mathbf{x}) d\mathbf{x}\right) \\
& \quad = \min_{\mathbf{G}_0, \mathbf{G}_1} \Big(L_{ff}(n) + L_{gg}(n) - 2 L_{fg}(n)\Big).
\end{aligned} \tag{11}$$

In case of finite joint density functions we can utilize the knowledge of (10) to calculate (11). In general this is an $m(2m-1)$ dimensional optimization problem, if the size of $\mathbf{G}_0$ (and $\mathbf{G}_1$) is m. If the size of $\mathbf{D}_0$ (and $\mathbf{D}_1$) is o then the computational complexity of (10) is mainly affected by the matrix inversion of the mo size matrix $\mathbf{D}_0 \oplus \mathbf{G}_0$.

From now on we assume that the MAP, represented by $(\mathbf{G}_0, \mathbf{G}_1)$, is a MAP(2) which is a real restriction as (10) and (11) holds for any two MAP of

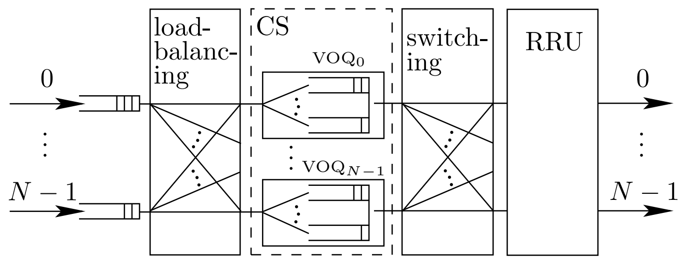

Figure 2: The overview of the load-balanced switch with N input and output ports

arbitrary size. But with this restriction we can utilize the minimal canonical forms, given in (4) and (5) and the optimization problem in (11) reduces to be a four dimensional constrained optimization, instead of the $m(2m-1) = 6$ one.

In the infinite case, if $n \to \infty$, (11) simplifies to be

$$\min_{\mathbf{G}_0, \mathbf{G}_1} \left(\lambda_f + \lambda_g - 2\lambda_{fg}\right),$$

where λ_f, λ_g and λ_{fg} are the dominant eigenvalues of the matrix (denoted as $\mathbf{N}$) in the expressions for $L_{ff}(n)$, $L_{gg}(n)$ and $L_{fg}(n)$ respectively.

There is the performance of all the numerical fitting methods evaluated according to their cumulative distribution fitting, their correlation structure fitting and their queueing behavior fitting.

Here we note again that the Euclidean distance based optimization in (11) is directly applicable to any two MAPs and its complexity can be reduced by the introduction of a minimal form of the fitting MAP, i.e., $(\mathbf{G}_0, \mathbf{G}_1)$.

4.2 Performance evaluation of the load-balanced switch with heterogenous input processes

The load-balanced switch is a two stage switching architecture which applies the recognition of [3]. That paper investigates the throughput of a one stage switch with two dimensional input queues with deterministic switching – using the round-robin (RR) pattern. It is shown that such a switch can provide for 100% throughput in case of uniform, Bernoulli, independent identically distributed input traffic. Appending a new stage in front of the one stage switch makes the non uniform arrival traffic uniform by spreading it evenly over the internal inputs (of the second stage) [5]. This is the investigated load-balanced switch architecture depicted in Figure 2.

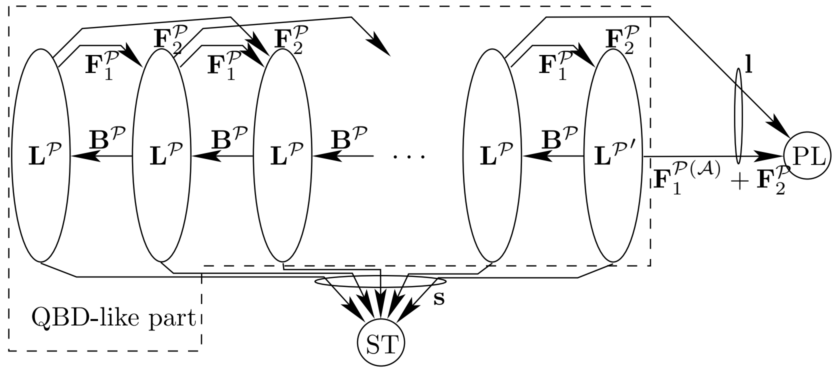

Figure 3: DTMC modeling the life cycle of the tagged packet

4.2.1 The exact model of the load-balanced switch with heterogenous input processes considered exhaustively

In the initial investigations on the load-balanced switch, [3], [5], [7], there is infinite buffers and equal size packets considered. In [1] the authors give the analytical model of the switch with finite buffers with still equal size packet assumption. In our work we go forward and give the performance evaluation of switches with variable size packets and finite buffers.

Thesis 2.1. *I give the exact analysis of the load-balanced switch with heterogenous arrival processes composed of variable size packets. [T1]*

During the analysis all the input processes are considered exhaustively, i.e., the state of all (heterogenous) inputs are considered exactly.

As the variable size *packet*s are segmented into equal size *cell*s a packet is lost if at least one of its cells is dropped at the finite central stage (CS) buffers. Using Markovian assumptions during the modeling we show that the packet loss probability is the probability of absorption in state PL of the transient discrete time Markov chain (DTMC) in Figure 3. This DTMC models the life cycle of a tagged packet ending either with successful packet transmission (absorption in state ST) or packet loss (absorption in state PL). The structure of the transient part smacks of the quasi birth-death (QBD) processes with the difference that in this case there can be more than one forward level transitions. Accordingly this part of the DTMC is denoted as "QBD-like". Its level process (horizontal dimension) represents the queue length of the CS buffer. Its phase process (vertical dimension) is the joint behavior of the heterogenous arrival processes, composed of variable size packets.

The arrival based decomposition of the phase process results in the level transition matrices $(\mathbf{B}^{\mathcal{P}}, \mathbf{L}^{\mathcal{P}}, \mathbf{F}_i^{\mathcal{P}}, \quad i \in [1, N-1])$ of the $N \times N$ switch. The level transition matrices are given with respect to cell/packet loss and successful packet transmission.

The transient DTMC, in Figure 3, is given by its initial distribution ($\boldsymbol{\pi}^{\mathcal{I}}$), the state transition probability matrix of the transient part ($\mathbf{P}^{\mathcal{P}}$) and the absorption vector to one of the absorbing states (either $\mathbf{s}$ or $\mathbf{l}$, the knowledge on other one is redundant information).

Having these properties the probability of successful packet transmission is the probability of absorbing in state ST

$$p_s = \boldsymbol{\pi}^{\mathcal{I}} \left(\mathbf{I} - \mathbf{P}^{\mathcal{P}}\right)^{-1} \mathbf{s}$$

and the packet loss probability is the probability of absorption in state PL

$$p_l = \boldsymbol{\pi}^{\mathcal{I}} \left(\mathbf{I} - \mathbf{P}^{\mathcal{P}}\right)^{-1} \mathbf{l}.$$

4.2.2 The approximate model of the load-balanced switch with heterogenous ON/OFF input processes

The problem with the exact model, in Thesis 2.1, comes from its complexity. If one considers the detailed behavior of all the inputs in the model of each input then the size of an input model scales with the number of inputs. Using such an input model to determine the phase process, i.e., the joint behavior of the inputs, of the QBD-like DTMC in Figure 3 results in a phase of size $(N+1)^N$. This is an enormous state space which can be solved only for very small switches ($N = 3$).

As response of the state space explosion we introduced the ON/OFF approximation of the input processes.

Thesis 2.2. *I give the approximating analysis of the load-balanced switch with heterogenous ON/OFF arrival process composed of variable size packets. [C8]*

In case of the ON/OFF input model the complexity of the model is reduced to be $O\left(2^N\right)$ at the expense of the exhaustive input process consideration. However the heterogenous input process consideration remained.

In the exact DTMC model of an input process there are $N+1$ states. N of them represent the arrival to the possible destinations (outputs) and another represents the idle period of the arrival process. Considering a particular output there is only one state corresponding to arrival to that output. For example this is given in Figure 4 by state **on** for $N = 3$, input 1 and output 0. The set of states, denoted as **off** in Figure 4, represents the case when there is no arrival to the considered output. This recognition led us to the discrete Phase-type (DPH) approximation of the sojourn time of the **off** states by a

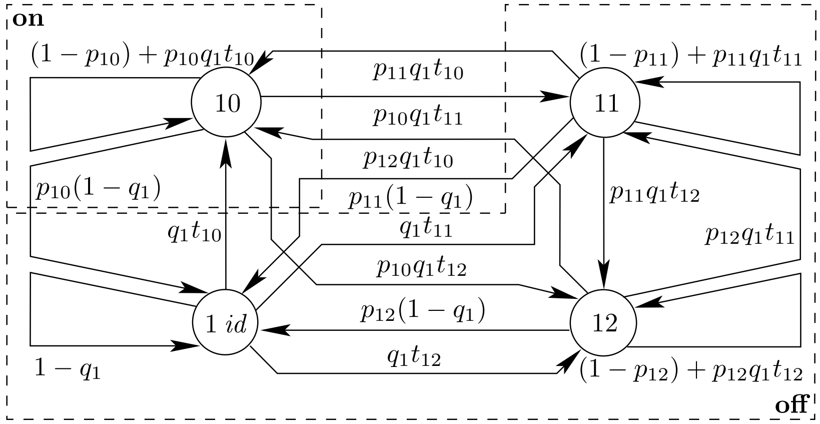

Figure 4: The exact DTMC model of the input process

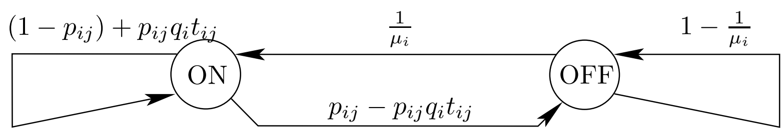

Figure 5: The graph of the ON/OFF DTMC modeling the general input-output pair

single OFF state and to introduce the two state ON/OFF model as given in Figure 5. μ_i is the time to absorption in the DPH substitute of the **off** states and accordingly $1-\frac{1}{\mu_i}$ is the sojourn probability of state OFF.

By this substitution we give the size independent model of the input processes and go on with modeling the tagged packet by the same transient, QBD-like model as in case of Thesis 2.1. As the phase process is again the joint behavior of the inputs the complexity remains exponential with the size of the switch but due to the introduction of the two state input model the base of the power is constant $O\left(2^N\right)$.

Emphasizing the difference between the two models the properties of the current transient DTMC are denoted as follows. Its initial distribution is $\boldsymbol{\varrho}^{\mathcal{I}}$, the state transition probability matrix of its transient part is $\mathbf{Q}^{\mathcal{P}}$ and its absorption vector to state ST is $\mathbf{t}$ and to state PL is $\mathbf{m}$. Using them the probability of successful packet transmission is expressed as

$$q_s = \boldsymbol{\varrho}^{\mathcal{I}}\left(\mathbf{I}-\mathbf{Q}^{\mathcal{P}}\right)^{-1}\mathbf{t}$$

and the packet loss probability is

$$q_l = \boldsymbol{\varrho}^{\mathcal{I}}\left(\mathbf{I}-\mathbf{Q}^{\mathcal{P}}\right)^{-1}\mathbf{m}.$$

4.3 Performance evaluation and packet loss minimization of the load-balanced switch with homogenous input processes

As both models, of Thesis 2.1 and 2.2, are exponential in the switch size (N) we go forward with the simplification and introduce a model with linear complexity $O(N)$. Although the model is less complex there is a restrictive assumption applied, there are identical, i.e., homogenous, input processes assumed.

As the new approach seems promising we give a numerical method to further reduce the computational needs of solving the model and a packet delay approximation too.

There is a loss minimization method also developed. Minimizing the loss is relevant as the remaining cells of a broken packet can cause more loss in the central stage (CS) and put in a claim for more storage capacity in the resequencing and reassembling unit (RRU), in Figure 2. There are several trials to solve the out of sequence problem, e.g., [4], [6], which can be applied together with the newly introduced loss minimization policy.

4.3.1 The approximate model of the load-balanced switch with homogenous input processes

Thesis 3.1. *I give an approximating model of the load-balanced switch with identical input assumption by which the model complexity is reduced to be linear in terms of the switch size (N). There is also an approximation for the packet delay as well as a fast numerical solution method given. [C7]*

The model for packet loss calculation If we assume identical input processes then their ON/OFF approximate model is also identical. Accordingly their joint behavior is described by a DTMC consists of $N+1$ states. Each state represents the number of inputs that are in ON. This DTMC can be used as the phase process of a quasi birth-like (QB-like) transient DTMC to model the tagged packet in a similar way as in both cases in Section 4.2.

The initial distribution of the transient DTMC is denoted as $\boldsymbol{\sigma}^{\mathcal{I}}$, the state transition probability matrix of the transient part is $\mathbf{S}^{\mathcal{P}}$ and the absorption vector to state ST is $\mathbf{u}$ and to state PL is $\mathbf{n}$. Using them the solution of the DTMC, i.e., the probability of successful packet transmission is

$$s_s = \boldsymbol{\sigma}^{\mathcal{I}} \left(\mathbf{I} - \mathbf{S}^{\mathcal{P}}\right)^{-1} \mathbf{u}$$

and the packet loss probability is

$$s_l = \boldsymbol{\sigma}^{\mathcal{I}} \left(\mathbf{I} - \mathbf{S}^{\mathcal{P}}\right)^{-1} \mathbf{n}.$$

Numerical solution method Basically the modeling process can divided into several parts, each of which is the calculation of a property of the transient DTMC modeling the life cycle of the tagged packet.

During the calculation of the initial distribution first we solve the cell level model of the switch which is a QBD-like DTMC. It has a state transition probability matrix with non-zero blocks in the lower subdiagonal, in the diagonal and in the first $N-1$ superdiagonals. It is repartitioned to be block tri-diagonal.

As a fast and numerically efficient algorithm, for the solution of linear equation system with block tri-diagonal coefficient matrix, we apply the folding algorithm, e.g., in [8]. During the repartitioning on the one hand we introduced the blocks of size $(N-1)(N+1)$ but on the other hand we are able to use the folding algorithm to calculate the initial distribution of the system with large buffers, i.e., state transition probability matrix containing large number of blocks.

A further important possibility, that lies in the identical input process assumption, is the change in the packet level model. The transient DTMC modeling the life of the tagged packet is now a QB-like model having upper block triangular state transition probability matrix. This recognition led us to an iterative numerical solution method for the packet loss probability (s_l) and the probability of successful packet transmission (s_s). This method reduces the inversion of the whole coefficient matrix in the solution for the DTMC to the inversion of its diagonal block.

Packet delay calculation The mean cell waiting time estimates the mean packet waiting time. This is the mean time between the entrance and the beginning of the service of a cell.

4.3.2 Packet loss minimization in load-balanced switches with homogenous input processes

During our performance evaluation it turned out that the main problem with the load-balanced switch architecture is the packet loss. It causes higher delays in the central stage and also in the resequencing and reassembly unit. Answering this problem we introduced a packet loss minimization technique which also can help to reduce system capacity and delay.

However the new technique decreases the central stage packet loss here we note that at the same time a possible packet loss appears at the inputs. When analyzing the switch with packet minimization one have to take into account both packet loss components compared to the unchanged switch architecture.

Thesis 3.2. *I give a packet loss minimization algorithm. I give also the analytical model of the switch with packet loss minimization algorithm based on the approximating model of the switch with homogenous input process assumption. [C10]*

The working mechanism of the packet loss minimization algorithm is the following. There is a controller monitoring the central stage queue lengths and if the queue length is above a predefined threshold then the *newly* arriving packets are dropped at the input.

If the buffer is almost saturated then the acceptance of the newly arrived packets would fill the buffer up and it would lead to the loss of the already and the newly accepted packets too. Instead of this, leading to further capacity waste, the arriving packet are dropped. Dropping the packets at the input drops all of its cells whilst loosing one of its cells at the central stage leaves its remaining cells in the buffers wasting the capacity and causing further losses.

During the modeling of the switch with the new algorithm mainly it have to be taken into consideration that the switch behavior is changed above the threshold.

Having a look at the loss minimization algorithm from the arrival process point of view one can say that the controller forces the input to remain idle if the central stage queue length is above the threshold. Mathematically it means the substitution of $\hat{q} = 0$ into the building blocks of the model if the queue length is above the threshold.

Here, once more, the packet loss probability at the central stage is the absorption in a transient DTMC modeling the life cycle of a tagged packet. The packet loss probability at the central stage, due to central stage buffer saturation, is the, threshold (t) dependent, probability of absorption in state PL

$$s_l^{(\mathrm{th})}(t) = \boldsymbol{\sigma}^{(\mathrm{th})^{\mathcal{I}}}(t) \left(\mathbf{I} - \mathbf{S}^{(\mathrm{th})^{\mathcal{P}}}(t)\right)^{-1} \mathbf{n}^{(\mathrm{th})}(t).$$

Here we distinguished the notations of the packet loss minimizing switch by the superscript $*^{(\mathrm{th})}$ from the values of the approximating model of Thesis 3.1. Its dependence on t is also emphasized.

Here we note again that the overall packet loss probability of the whole switch consists of another component, namely the loss at the input $s_i(t)$. Using $s_i(t)$ we can give the overall, joint input-central stage (I-CS) packet loss probability of the switch as

$$s_{\text{I-CS}}(t) = s_i(t) + \big(1 - s_i(t)\big) s_l^{(\mathrm{th})}(t),$$

Here we emphasized again that the overall loss probability depends on t – the threshold.

5 Applications

The new results on the correlation bounds of the second order MAP can be applied in several way. Two important applications are their applicability either in, moment based, arrival process fitting or in MAP reduction. These applications are the subject of Thesis 1.2 and 1.3. In practice it is useful in developing simple model of complex systems (moment based fitting) and model (MAP) size reduction, e.g., in queueing network analysis it is worth to reduce the model size eventually.

The load-balanced switch model can be used in network design. Using the predictions on the traffic volume and the aimed quality parameters one can determine the central stage buffer sizes of the applied LB switch as well as one can easily set the artificial buffer threshold based on the real (or predicted) traffic volume to decrease (or to set) the packet loss probability (with minimal buffer sizes).

References

[1] Yury Audzevich, Yoram Ofek, Miklós Telek, and Bülent Yener. Analysis of load-balanced switch with finite buffers. In *IEEE GLOBECOM*, pages 1–6, New Orleans, LA, USA, 2008.

[2] Andrea Bobbio, András Horváth, and Miklós Telek. Matching three moments with minimal acyclic phase type distributions. *Stochastic models*, pages 303–326, 2005.

[3] Cheng-Shang Chang, Duan-Shin Lee, and Yi-Shean Jou. Load-balanced Birkhoff-von Neumann switches, part I: One-stage buffering. *Computer Communications*, 25:611–622, 2002.

[4] Cheng-Shang Chang, Duan-Shin Lee, and Ching-Ming Lien. Load-balanced Birkhoff-von Neumann switches, part II: Multi-stage buffering. *Computer Communications*, 25:623–634, 2002.

[5] Isaac Keslassy, Shang-Tse Chuang, Kyoungsik Yu, David Miller, Mark Horowitz, Olav Solgaad, and Nick McKeown. Scaling Internet routers using optics. In *ACM SIGCOMM*, Karlsruhe, Germany, 2003.

[6] Isaac Keslassy and Nick McKeown. Maintaining packet order in two-stage switches. In *IEEE INFOCOM*, New York, US, June 2002.

[7] Chih-Ying Tu, Cheng-Shang Chang, Duan-Shin Lee, and Ching-Te Chiu. Design a simple and high performance switch using a two-stage architecture. In *IEEE GLOBECOM*, volume 2, pages 733–738, St. Louis, MO, USA, November 2005.

[8] Jingdong Ye and San-qi Li. Folding algorithm: A computational method for finite QBD processes with level-dependent transitions. *IEEE Transactions on Communications*, 42(2/3/4):652–639, February/March/April 1994.

International articles

[J1] Levente Bodrog, András Horváth, and Miklós Telek. Moment characterization of matrix exponential and Markovian arrival processes. *Annals of Operations Research*, 160(1):51–68, 2008.

[J2] Levente Bodrog, Armin Heindl, Gábor Horváth, and Miklós Telek. A markovian canonical form of second-order matrix-exponential processes. *European Journal of Operational Research*, 190(2):459–477, 2008.

[J3] Levente Bodrog, Gábor Horváth, and Csaba Vulkán. Analytical TCP throughput model for HSDPA. *IET Software*, 3(6):480–494, December 2009.

International conference papers

[C1] Levente Bodrog, Gábor Horváth, and Miklós Telek. Comparison of simulation models for long-range dependent traffic traces. In János Levendovszky, editor, *3rd International Workshop on Rare Event Simulation and Combinatorial Optimization*, Budapest, Hungary, 7-8 September 2004.

[C2] Levente Bodrog, Gábor Horváth, Sándor Rácz, and Miklós Telek. A tool support for automatic analysis based on the tagged customer approach. In *International Conference on the Quantitative Evaluation of SysTems*, pages 323–332, Riverside, CA, USA, September 2006. IEEE Computer Society.

[C3] Levente Bodrog, András Horváth, and Miklós Telek. On the properties of moments of matrix exponential distributions and matrix exponential processes. In Dario Bini, Beatrice Meini, Vaidyanathan Ramaswami, Marie-Ange Remiche, and Peter Taylor, editors, *Numerical Methods for Structured Markov Chains*, Davis, Germany, November 2007.

[C4] Levente Bodrog, Armin Heindl, Gábor Horváth, András Horváth, and Miklós Telek. Current results and open questions on PH and MAP characterization. In Dario Bini, Beatrice Meini, Vaidyanathan Ramaswami, Marie-Ange Remiche, and Peter Taylor, editors, *Numerical Methods for Structured Markov Chains*, Davis, Germany, November 2007.

[C5] Levente Bodrog, Gábor Horváth, and Csaba Vulkán. Analytical TCP throughput model for HSDPA. In Jeremy Bradley and William Knottenbelt, editors, *24th UK Performance Engineering Workshop*, pages 89–106, London, UK, July 2008.

[C6] Levente Bodrog, Ákos Horváth, and Miklós Telek. Statistical analysis of peer-to-peer live streaming traffc. In *Proceedings of Probability and Statistics with Applications*, Debrecen, Hungary, June 2009.

[C7] Yury Audzevich, Levente Bodrog, Yoram Ofek, and Miklós Telek. Scalable model for packet loss analysis of load-balancing switches with identical input processes. In Khalid Al-Begain, Dieter Fiems, and Gábor Horváth, editors, *ASMTA*, volume 5513 of *LNCS*, pages 249–263, Madrid, Spain, June 2009. Springer.

[C8] Yury Audzevich, Levente Bodrog, Yoram Ofek, and Miklós Telek. Packet loss analysis of load-balancing switch with ON/OFF input processes. In Jeremy Bradley, editor, *Computer Performance Engineering*, volume 5652 of *LNCS*, pages 197–211, London, UK, July 2009. Springer.

[C9] Philipp Reinecke, Katinka Wolter, Levente Bodrog, and Miklós Telek. On the cost of generating PH-distributed random numbers. In *Proceedings of PMCCS*, Eger, Hungary, September 2009.

[C10] Yury Audzevich, Levente Bodrog, Yoram Ofek, and Miklós Telek. Packet loss minimization in load-balancing switch. In Khalid Al-Begain, Dieter Fiems, and William Knottenbelt, editors, *ASMTA*, LNCS, Cardiff, UK, June 2010. Springer.

[C11] Levente Bodrog, Peter Buchholz, Jan Kriege, and Miklós Telek. Canonical form based MAP(2) fitting. In *International Conference on the Quantitative Evaluation of SysTems*, Williamsburg, VA, USA, September 2010. IEEE Computer Society. accepted for publication, available at http://webspn.hit.bme.hu/~bodrog/publicat/qest10.pdf.

Hungarian journal article

[J4] Levente Bodrog, Gábor Horváth, and Csaba Vulkán. A TCP/HSDPA rendszer átvitelének analitikus modellje. *HÍRADÁSTECHNIKA*, LXIII(11):36–43, November 2008. in Hungarian.

Technical report

[T1] Yury Audzevich, Levente Bodrog, Miklós Telek, Yoram Ofek, and Bülent Yener. Variable Size Packets Analysis in Load-balanced Switch with Finite Buffers. Technical report, Technical University of Budapest, Hungary, January 2009. available at http://webspn.hit.bme.hu/~bodrog/techrep/AuBoTeOfYe09.pdf.

Referring us

[R1] András Horváth, Gábor Horváth, and Miklós Telek. A joint moments based analysis of networks of MAP/MAP/1 queues. In *Proceedings of the 2008 Fifth International Conference on Quantitative Evaluation of Systems*, pages 125–134, Washington, DC, USA, 2008. IEEE Computer Society. [J1].

[R2] András Horváth, Gábor Horváth, and Miklós Telek. A traffic based decomposition of two-class queueing networks with priority service. *Computer Networks*, 53(8):1235–1248, June 2009. [J1].

[R3] Giuliano Casale, Eddy Z. Zhang, and Evgenia Smirni. Trace data characterization and fitting for Markov modeling. *Performance Evaluation*, 67(2):61–79, 2010. [J2].

[R4] Carmen Sangüesa. On the minimal value in Maier's property concerning phase-type distributions. *Stochastic Models*, 26(1):124–140, January 2010. [J2].