Budapest University of Technology and Economics
Department of Networked Systems and Services

Representation optimization in Markovian models

Ph.D. dissertation summary

András Mészáros

supervised by
Dr. Miklós Telek

November 12, 2018

1 Introduction

Markov chain based models are used in many fields, including queueing theory and its applications [19, 20, 6], reliability analysis [30], the analysis of chemical reactions [3], machine learning [4], and speech recognition [27] to model distributions, stochastic processes, and decision problems. Markovian models are popular for multiple reasons. Firstly, they are very flexible, that is, they can be used to approximate a wide variety of behavioural patterns. Secondly, they often have a physical interpretation that can be tied to the problem to be modelled. Thirdly, efficient analytical and numerical methods have been developed to fit Markovian models to numerical data and to calculate their main characteristics.

Markov chain based models are generally described by a set of matrices. This description is convenient from a definition point of view, however, it is usually highly redundant, as the number of elements in the matrices is significantly higher than the number of parameters needed to characterize the model. In conjunction, an infinite number of different representations can be constructed that have the exact same stochastic properties. While these representations are equivalent in some regards, in many applications some of them are preferable to others. For example, certain numerical fitting methods work better when the total number of non-zero elements in the describing matrices is minimal. Another example is the examination of systems described by matrices of infinite size. The analysis of these systems generally involves approximations. In some cases, however, the describing matrices can be compressed to finite size such that the relevant performance measures in the associated system remain unchanged.

As can be seen from the above, finding the optimal representation of a Markovian model is an important task, however, it is also often a difficult problem and may also depend on the application of the model. In my dissertation I investigated the problem of representation optimization from different perspectives for three Markovian model classes: phase type distributions (PHs), Markov arrival processes (MAPs), and Markov decision processes (MDPs). I also examined how the non-Markovian extension of PHs and MAPs can be transformed to Markovian forms. My objective was to develop numerically efficient methods to find optimal (where analytical optimization is possible) or improved representations (when only heuristic methods are applicable) for the investigated model classes.

2 Research objectives

When using Markovian models we often build on their flexibility and tractability. By changing the representation of a model appropriately we can further improve these qualities. My primary goal during my research was to understand what are the best representations, how can we obtain these representations, what are the underlying parameters that really define given Markovian models and what are their limits. During this research I naturally encountered the non-Markovian forms of the Markovian models and, consequently, the non-Markovian extension of these models. Thus my research also touched upon the relation between these two classes. Based on the topic and the applied methodology the results presented in my dissertation can be split into the following three groups:

1. Representation transformation using elementary transformation matrices
2. Canonical representation of order 2 Markov arrival processes
3. State space reduction of Markov decision processes

In the following I give a short introduction to the above topics and summarize the main objectives of my research in these areas.

2.1 Representation transformation using elementary transformation matrices

When analytical methods to find the optimal representation for a problem are not available or the available methods are infeasible because of their computational complexity, heuristic algorithms can be devised to improve the representation. These algorithms do not guarantee optimality, but the resulting representations can be satisfactory in practice.

In my research I investigated a particular approach, which is based on similarity transformations. Matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ are similar if there is an invertible $\boldsymbol{T}$ matrix, such that

$$\boldsymbol{B} = \boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}.$$

In this case the transformation between $\boldsymbol{A}$ and $\boldsymbol{B}$ is called a similarity transformation with transformation matrix $\boldsymbol{T}$. If such a transformation with an appropriate $\boldsymbol{T}$ matrix is applied to a Markovian model, it results in a new representation of the model. In theory every other representation of a model can be obtained using a similarity transformation as long as they are of the same size.

My research goal was to develop algorithms that utilize the flexibility of representation transformations but also have low computational cost. In my thesis I present two numerical methods that apply a chain of elementary transformations to optimize the representation.

2.2 Canonical representation of order 2 Markov arrival processes

The notion of canonical forms is present in many fields of mathematics. Generally a canonical form of a class of objects is required to be unique, that is, every object of the class has to have one and only one canonical form. Canonical forms not only give us a simple, standardized way of representing mathematical objects, they can also help us understand the properties and the boundaries of the corresponding class and aid the development of more efficient fitting methods. Canonical forms have been established previously for several Markovian models, including order 2 and order 3 continuous and discrete phase type distributions [15, 26, 16] and order 2 stationary continuous time Markov arrival processes [5]. My aim was to complement these results for other Markovian classes. I examined and proposed canonical forms for order 2 stationary discrete time Markov arrival classes, as well as order 2 continuous time non-stationary and transient Markov arrival processes.

2.3 State space reduction of Markov decision processes

Markov decision processes (MDPs) are used in decision problems of many areas including telecommunication [11], economy [8], robot navigation [22], and reinforcement learning [18]. A Markov decision process is defined by an $(S, A, \boldsymbol{Q}(a), \boldsymbol{C}(a))$ tuple, where S is the set of states of the model, A is the set of decisions that can be made at any given state, $\boldsymbol{Q}(a)$ is the decision dependent generator matrix of the process and $\boldsymbol{C}(a)$ is a decision dependent diagonal reward rate matrix. A traditional problem in Markov decision processes is to find a policy (state-decision mapping) $\pi^*(s) \in \{\pi(s) : S \to A\}$ that maximizes the long term reward rate.

The computational time of the optimal policy is polynomial in the number of states and the number of possible decisions for the classical solution algorithms [25]. This means that, if the state space is large or infinite, the optimization using only the traditional methods can be time consuming or infeasible. My goal during my research was to make such cases more tractable using state space compression.

3 Methodology

The three different areas of my research lend themselves naturally to three different approaches. Representation transformation is an analytically hard problem especially for larger Markovian models. This motivated me to focus on developing heuristic algorithms. The proposed algorithms are based on simple results of linear algebra and were tested using numerical experiments. The research of canonical representations required a purely analytical approach in which I used only elementary mathematical tools including fixed point analysis and basic calculus. In the research of state space reduction of Markov decision processes I mainly used the matrix analytic methodology [23] which is a popular technique for calculating characteristics of Markov chain based models.

4 Results

4.1 Representation transformation using elementary transformation matrices

Thesis I.1. *[C2] I devised a heuristic algorithm for transforming the representation of phase type distributions as part of a Markov arrival process (MAP) fitting procedure that provides higher maximal lag-1 autocorrelation than previous methods.*

Fitting MAPs to statistical or traffic trace data of packet inter-arrival times is a challenging problem for which there is no generally accepted procedure. The most general method that has been proposed is the expectation maximization (EM) algorithm [9]. This algorithm, however, is computationally expensive and has numerical stability problems. To address these issues, some methods use heuristics, or fit MAPs with special structures [2, 13]. One heuristic approach is to perform the fitting in two steps [10, 14]. In the first step we fit a PH to the marginal distribution of the inter-arrival times and in the second step we extend this to a MAP. The advantage of this approach is that it divides the problem into two parts and the parameters to be fitted, thus it decreases the computational complexity of the problem. Additionally, there are multiple PH fitting methods available from which one can choose according to the characteristics of the given problem. One of the concerns with this heuristic is that the representation of the PH obtained in the first step affects the flexibility of the potential dynamic behaviour of the resulting MAP. For example some PH representations can only be extended to MAPs with uncorrelated inter-arrival times. To address this problem the representation of the PH has to be optimized using some transformation method. In [10] Buchholz and Kriege propose a method for the task. In my work I developed a representation transformation algorithm that is more flexible than the one of Buchholz and Kriege. I compared the two methods using numerical experiments which show that the new transformation method enables better fitting in general at the cost of higher computational complexity.

Thesis I.2. *[C4] I devised an improved version of an existing heuristic algorithm for transforming matrix exponential distributions and rational arrival processes to Markovian representation and demonstrated through numerical experiments that it has higher success rate than existing methods.*

PHs and MAPs are defined with vectors and matrices in which the elements have strict sign constraints. Matrix exponential distributions (MEs) and rational arrival processes (RAPs) are generalized classes of PHs and MAPs, where these sign constraints are lifted. This generalization has multiple advantages. For some distributions and processes it can lead to a more efficient description (where the describing vectors and matrices are smaller). It also enables the usage of fitting methods that do not keep the sign constraints of the Markovian classes (see e.g. [24]). The main problem with the non-Markovian extensions is that there is no general method to decide whether a non-Markovian description defines a valid distribution or process (i.e., one with non-negative (joint) probability density function). One way to confirm that a non-Markovian description describes a valid stochastic model is by transforming it to a Markovian representation. If the transformation is successful, the description corresponds to a real stochastic process. Currently there are only heuristic methods for this task. In my research I improved an existing representation transformation method. The new version uses more complex transformation steps than the original one, while keeping the computational costs low. I examined several possibilities and showed in numerical experiments that they find Markovian representations with a higher ratio than the original method, at the cost of higher computational cost.

4.2 Canonical representation of order 2 Markov arrival processes

Thesis II.1. *[J1] I showed that order 2 discrete time stationary Markov arrival processes (DMAP(2)) are equivalent to order 2 discrete time stationary rational arrival processes, and any order 2 discrete time stationary Markov arrival processes can be given using one of the following four minimal parameter Markovian canonical forms*

- *Form 1:* $\lambda_1, \lambda_2 > 0, \gamma \geq 0$

$$\boldsymbol{D_0} = \begin{bmatrix} 1-\lambda_1 & (1-a)\lambda_1 \\ 0 & 1-\lambda_2 \end{bmatrix}, \quad \boldsymbol{D_1} = \begin{bmatrix} a\lambda_1 & 0 \\ (1-b)\lambda_2 & b\lambda_2 \end{bmatrix},$$

- *Form 2:* $\lambda_1, \lambda_2 > 0, \gamma < 0$

$$\boldsymbol{D_0} = \begin{bmatrix} 1-\lambda_1 & (1-a)\lambda_1 \\ 0 & 1-\lambda_2 \end{bmatrix}, \quad \boldsymbol{D_1} = \begin{bmatrix} 0 & a\lambda_1 \\ b\lambda_2 & (1-b)\lambda_2 \end{bmatrix},$$

- *Form 3:* $\lambda_1 < 0, \lambda_2 > 0, \gamma \geq 0$

$$\boldsymbol{D_0} = \begin{bmatrix} 1-\beta_1 & a\beta_1 \\ \frac{1}{a}\beta_2 & 0 \end{bmatrix}, \quad \boldsymbol{D_1} = \begin{bmatrix} (1-a)\beta_1 & 0 \\ (1-\frac{1}{a}\beta_2)b & (1-\frac{1}{a}\beta_2)(1-b) \end{bmatrix},$$

- *Form 4:* $\lambda_1 < 0, \lambda_2 > 0, \gamma < 0$

$$\boldsymbol{D_0} = \begin{bmatrix} 1-\beta_1 & a\beta_1 \\ \frac{1}{a}\beta_2 & 0 \end{bmatrix}, \quad \boldsymbol{D_1} = \begin{bmatrix} 0 & (1-a)\beta_1 \\ (1-\frac{1}{a}\beta_2)b & (1-\frac{1}{a}\beta_2)(1-b) \end{bmatrix},$$

where $0 < \lambda_1 \leq \lambda_2$ *are the eigenvalues of the DMAP(2),* $0 < \beta_1, \beta_2 < 1$*, and* $0 \leq a, b \leq 1$*.*

Thesis II.2. *[J1] I determined the moment and correlation parameter boundaries of order 2 discrete time stationary Markov arrival processes for the canonical forms of Thesis II.1. (The boundaries can be found in [J1].) I also showed that, for any set of* f_1, f_2, f_3 *moments and* γ *correlation parameter within these boundaries, the parameters of the canonical form from Thesis II.1 can be given using the formulas in Table 1, where*

$$
\begin{aligned}
h_1 &= 2{f_1}^2 - 2f_1 - f_2,\ h_2 = 3{f_2}^2 - 2f_1 f_3,\\
h_3 &= 3f_1 f_2 - 6(f_1 + f_2 - f_1^2) - f_3,\\
h_4 &= h_3^2 - 6h_1 h_2,\\
z &= \frac{h_2}{|h_2|},\\
p &= \frac{-z(h_3 - 6f_1 h_1) + \sqrt{h_4}}{zh_3 + \sqrt{h_4}}\\
k_1 &= (1-\gamma)(p + \beta_1 + \beta_2 - p\beta_2) - 1 + \beta_1,\\
k_2 &= 4\beta_1(k_1 - \beta_1 + \gamma - \beta_2\gamma),\\
k_3 &= (1-\gamma)(-p(1-\beta_2) - 2\beta_2) - \gamma(1-\beta_1),\\
k_4 &= k_3 + \beta_2 + \gamma - \beta_2\gamma.
\end{aligned}
$$

can. form	parameters	
Form 1	$\lambda_1 = 1 - \frac{h_3 - z\sqrt{h_4}}{h_2}$	$a = \frac{d_1 - \sqrt{d_2}}{2(1-s_1)}$ $b = \frac{d_1 + \sqrt{d_2}}{2(1-s_2)}$
Form 2	$\lambda_2 = 1 - \frac{h_3 + z\sqrt{h_4}}{h_2}$	$a = \frac{-\gamma(1-s_2)}{p(1-s_2)(1-\gamma) - \gamma(1-s_1)}$ $b = \frac{p(1-s_2)(1-\gamma) - \gamma(1-s_1)}{1-s_2}$
Form 3	$\beta_1 = \frac{12f_1^2 - 3f_2(4+f_2) - 2f_3 + 2f_1(-6+3f_2+f_3)}{(3f_2^2 - 2f_1 f_3)}$	$a = \frac{k_1 + \sqrt{k_1^2 - k_2}}{2\beta_1}$ $b = 1 - \frac{a\gamma(1-\beta_2)}{(1-a)(a-\beta_2)}$
Form 4	$\beta_2 = \frac{-3f_2(2-2f_1+f_2) + 2(-1+f_1)f_3}{12f_1^2 - 3f_2(4+f_2) - 2f_3 + 2f_1(-6+3f_2+f_3)}$	$a = \frac{k_3 + \sqrt{k_3^2 + 4\beta_2 k_4}}{2k_4}$ $b = -\frac{a\gamma(1-\beta_2)}{(1-a)(a-\beta_2)}$

Table 1: Parameter matching formulas for DMAP(2)

Thesis II.3. *[C5] I showed that class of order 2 continuous time non-stationary Markov arrival processes and order 2 continuous time non-stationary rational arrival processes are equivalent and any order 2 continuous time non-stationary*

MAP($\pi_0, \boldsymbol{H_0}, \boldsymbol{H_1}$) can be represented in the

$$(\underline{\delta}, \boldsymbol{D_0}, \boldsymbol{D_1}) = (\underline{\pi_0}\boldsymbol{T}, \boldsymbol{T}^{-1}\boldsymbol{H_0}\boldsymbol{T}, \boldsymbol{T}^{-1}\boldsymbol{H_1}\boldsymbol{T})$$

canonical form, where $\boldsymbol{T}$ is the transformation matrix which transforms the MAP($\boldsymbol{H_0}, \boldsymbol{H_1}$) stationary MAP(2) to the $(\boldsymbol{T}^{-1}\boldsymbol{H_0}\boldsymbol{T}, \boldsymbol{T}^{-1}\boldsymbol{H_1}\boldsymbol{T})$ canonical form according to [5].

Thesis II.4. *[J2] I showed that the class of order 2 continuous time transient Markov arrival processes and order 2 continuous time transient rational arrival processes are equivalent and any order 2 continuous time transient Markov arrival process can be represented in one of the following five minimal parameter Markovian canonical forms*

- *Form 1: $\gamma > 0$*

$$\boldsymbol{D_0^{(1)}} = \begin{bmatrix} -\lambda_1 & \hat{c}\lambda_1 \\ 0 & -\lambda_2 \end{bmatrix}, \ \boldsymbol{D_1^{(1)}} = \begin{bmatrix} \hat{a}\lambda_1 & 0 \\ \hat{d}\lambda_2 & \hat{b}\lambda_2 \end{bmatrix}, \ \underline{\delta}^{(1)} = \begin{bmatrix} (1-\hat{a}-\hat{c})\lambda_1 \\ (1-\hat{b}-\hat{d})\lambda_2 \end{bmatrix},$$

  where $0 < \lambda_1 \le \lambda_2$, $\hat{b} > \hat{a}\frac{\lambda_1}{\lambda_2}$,

- *Form 2: $\gamma > 0$*

$$\boldsymbol{D_0^{(2)}} = \begin{bmatrix} -\bar{\lambda}_1 & \bar{c}\bar{\lambda}_1 \\ (1-\bar{b}-\bar{d})\bar{\lambda}_2 & -\bar{\lambda}_2 \end{bmatrix}, \ \boldsymbol{D_1^{(2)}} = \begin{bmatrix} \bar{a}\bar{\lambda}_1 & 0 \\ \bar{d}\bar{\lambda}_2 & \bar{b}\bar{\lambda}_2 \end{bmatrix}, \ \underline{\delta}^{(2)} = \begin{bmatrix} (1-\bar{a}-\bar{c})\bar{\lambda}_1 \\ 0 \end{bmatrix},$$

  where $0 < \bar{\lambda}_1, \bar{\lambda}_2$, $\bar{b} > \bar{a}\frac{\bar{\lambda}_1}{\bar{\lambda}_2}$,

- *Form 3: $\gamma < 0$*

$$\boldsymbol{D_0^{(3)}} = \begin{bmatrix} -\lambda_1 & \check{c}\lambda_1 \\ 0 & -\lambda_2 \end{bmatrix}, \ \boldsymbol{D_1^{(3)}} = \begin{bmatrix} 0 & \check{a}\lambda_1 \\ \check{b}\lambda_2 & \check{d}\lambda_2 \end{bmatrix}, \ \underline{\delta}^{(3)} = \begin{bmatrix} (1-\check{a}-\check{c})\lambda_1 \\ (1-\check{b}-\check{d})\lambda_2 \end{bmatrix},$$

  where $0 < \lambda_1 < \lambda_2$, $\check{d} > \check{a}\frac{\lambda_1}{\lambda_2}$,

- *Form 4: $\gamma < 0$*

$$\boldsymbol{D_0^{(4)}} = \begin{bmatrix} -\tilde{\lambda}_1 & \tilde{c}\tilde{\lambda}_1 \\ (1-\tilde{b}-\tilde{d})\lambda_2 & -\tilde{\lambda}_2 \end{bmatrix}, \ \boldsymbol{D_1^{(4)}} = \begin{bmatrix} 0 & \tilde{a}\tilde{\lambda}_1 \\ \tilde{d}\tilde{\lambda}_2 & \tilde{b}\tilde{\lambda}_2 \end{bmatrix}, \ \underline{\delta}^{(4)} = \begin{bmatrix} (1-\tilde{a}-\tilde{c})\tilde{\lambda}_1 \\ 0 \end{bmatrix},$$

  where $0 < \tilde{\lambda}_1 < \tilde{\lambda}_2$,

- *Form 5: no restriction on the sign of γ*

$$\boldsymbol{D_0^{(5)}} = \begin{bmatrix} -\lambda_1 & \dot{c}\lambda_1 \\ 0 & -\lambda_2 \end{bmatrix}, \ \boldsymbol{D_1^{(5)}} = \begin{bmatrix} \dot{a}\lambda_1 & (1-\dot{a}-\dot{c})\lambda_1 \\ \dot{d}\lambda_2 & \dot{b}\lambda_2 \end{bmatrix}, \ \underline{\delta}^{(5)} = \begin{bmatrix} 0 \\ (1-\dot{b}-\dot{d})\lambda_2 \end{bmatrix},$$

  where $0 < \lambda_1 \le \lambda_2$.

And all the parameters other than $\lambda_1, \lambda_2, \bar{\lambda}_1, \bar{\lambda}_2, \tilde{\lambda}_1, \tilde{\lambda}_2$ are between 0 *and* 1.

The standard description of stationary Markov arrival processes is composed of two square matrices. A single MAP can be described with matrix pairs of different sizes. The minimum of these is called the order of the MAP. Thus the standard description of an order n MAP has $2n^2$ elements, however, it is completely characterized by n^2 parameters, for example by $2n-1$ moments and $(n-1)^2$ joint moments of the inter-arrival times [29]. As a consequence, multiple representations describe the same process. This over-parametrization makes numerical fitting harder for the standard description. One way to overcome this problem is by finding so called canonical forms which assign a unique, minimal parameter representation to each process. Previously, canonical forms were only known for order 2 continuous time stationary Markov arrival processes [5]. In my research I devised canonical forms for order 2 discrete time stationary Markov arrival processes, order 2 continuous time non-stationary Markov arrival processes, and order 2 continuous time transient Markov arrival processes. I also proved that these process classes and their non-Markovian counterparts are equivalent. Based on the canonical representation of order 2 discrete time stationary Markov arrival processes I was also able to determine the moment and correlation boundaries of order 2 discrete time stationary Markov arrival processes.

4.3 State space reduction of Markov decision processes

Thesis III.1. *[B1] I showed that any MDP with state space S, set of decisions A, policy dependent generator matrix $Q(\pi)$, and reward rate matrix $C(\pi)$, or shortly MDP($S, A, Q(\pi), C$) has the same optimal policy as MDP($S_U, A, Q'(\pi), C'(\pi)$), where $S_U \cup S_D = S, S_U \cap S_D = \emptyset$, decisions are only made in S_U in the original MDP, and*

$$\boldsymbol{Q'}_{ij}(\pi) = \begin{cases} -\frac{1}{\tau_i(\pi)}, & \textit{if } i = j, \\ \frac{\boldsymbol{P}_{ij}(\pi)}{\tau_i(\pi)}, & \textit{otherwise,} \end{cases} \quad \textit{and} \quad \boldsymbol{C'}_{ij}(\pi) = \begin{cases} \frac{c_i(\pi)}{\tau_i(\pi)}, & \textit{if } i = j, \\ 0, & \textit{otherwise,} \end{cases}$$

with

$$\begin{aligned} \boldsymbol{P}_{ij}(\pi) &= Pr(X(\rho_{S_U \setminus i}) = j \mid X(0) = i), \\ \tau_i(\pi) &= E[\rho_{S_U \setminus i} \mid X(0) = i], \\ c_i(\pi) &= E[\int_{t=0}^{\rho_{S_U \setminus i}} C_{X(t)X(t)} dt \mid X(0) = i], \end{aligned}$$

where $\rho_{S_U \setminus i}$ is the time to visit a state in S_U different from i if the process starts in state $i \in S_U$, that is $\rho_{S_U \setminus i} = min(t | X(t) \in S_U \setminus i)$.

Thesis III.2. *I devised a method for reducing the state space of M/G/1 type Markov decision processes.*

Thesis III.3. *I devised a method for reducing the state space of G/M/1 type Markov decision processes.*

Markov Decision processes are widely used to describe decision problems in various fields. Many decision problems are easy to formulate using MDPs, as any parameter of the examined system can be matched to a dimension in the multi-dimensional state space of an MDP. One of the drawbacks of using MDPs is that such a formulation can result in a very large or even infinite number of states. This phenomenon is known as state space explosion. To combat this problem several techniques have been developed. Some use structural properties to give an exact solution like [21] or[12], which can be used for MDPs with threshold form, or [7], which provides efficient solution methods for MDPs with factored representations. Alternatively, one can get a quasi-optimal solution by using certain approximation techniques. One possible approach is the truncation of the state space. This may happen based on the physical model (e.g. the size of the buffer is constrained), as in [28] and [17] for example. Truncating can also be done based on mathematical considerations only as discussed by [1].

In my research I developed an efficient reduction method that can be used for MDPs which are composed of a finite subset of states with decisions and a finite or infinite subset of states without decisions. More specifically the method compresses the MDP to the size of the subset with decisions. This method can be used to obtain an exact solution and does not constrain the structure of the MDP as much as solution methods relying on a factored representation or threshold form. The proposed method requires that some important parameters of the MDP can be calculated efficiently. In my thesis I show how this can be done for infinite state MDPs with QBD, M/G/1-type and G/M/1-type structures, which are the most prevalent classes of infinite MDPs in queueing problems.

5 Applications

The results of my thesis can be applied for various problems. The first thesis group can primarily be used in conjunction with the fitting of Markovian structures. The algorithm for transforming phase type distributions in Thesis I.1 was used in a heuristic MAP fitting algorithm which balances computational complexity and flexibility. Utilizing the representation transformation method in Thesis I.2 one can fit MAPs and PHs using fitting methods devised for their non-Markovian extensions, e.g. [24]. These methods can be more general as they are not restrained by the Markovian constraints. A Markovian model can be obtained by first using these more general methods and transforming the resulting non-Markovian forms to Markovian representation using the methods of Thesis I.2.

The results in the second thesis group can be used in model fitting as well. While general Markovian representations are highly redundant, canonical forms have a minimal number of parameters. Consequently fitting with canonical forms is much more efficient which results in a better fitting in general. (A demonstration of the improved performance when fitting with canonical forms can be found in [J1].) The simpler structure of canonical forms can also help in determining the boundaries of the model classes and make parameter matching possible using explicit formulas, as shown in Thesis II.2. By knowing the boundaries we can decide whether the given model class can be used to model real life data, or we need a more flexible class, while exact parameter matching provides a highly efficient way to find a Markovian model that has the required parameters.

The results in the third thesis group provide a way to efficiently solve a wide set of MDPs. These results can be utilized in any field where MDPs are used to optimize decision making. The general reduction method in Thesis III.1 can be used for MDPs with finite state space and can be applied for many types of MDPs with infinite state space. Some important types are covered by Theses III.2 and III.3. These MDP types appear in many queueing decision problems where large queues are approximated by infinite ones.

Publications

[B1] A. Mészáros and M. Telek. "Server optimization of infinite queueing systems". In: *Markov Decision Processes in Practice*. Springer, 2016.

[C1] L. Bodrog et al. "Control of queues with MAP servers: experimental results". In: *The Eighth International Conference on Matrix-Analytic Methods in Stochastic Models (MAM8)*. Calicut, Kerala, India, 2014, pp. 9–11.

[C2] A. Mészáros and M. Telek. "A two-phase MAP fitting method with APH interarrival time distribution". In: *Proceedings of the Winter Simulation Conference*. WSC '12. Berlin, Germany: Winter Simulation Conference, 2012, 425:1–425:12. URL: http://dl.acm.org/citation.cfm?id=2429759.2430310.

[C3] A. Mészáros and M. Telek. "Markov Decision Process and Linear Programming Based Control of MAP/MAP/N Queues". In: *Computer Performance Engineering, EPEW*. Vol. 8721. LNCS. Florence, Italy, Sept. 2014, pp. 179–193. DOI: 10.1007/978-3-319-10885-8_5.

[C4] A. Mészáros, G. Horváth, and M. Telek. "Representation Transformations for Finding Markovian Representations". In: *Analytical and Stochastic Modeling Techniques and Applications*. Ed. by A. Dudin and K. Turck. Vol. 7984. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2013, pp. 277–291. ISBN: 978-3-642-39407-2. DOI: 10.1007/978-3-642-39408-9_20.

[C5] A. Mészáros and M. Telek. "Canonical Form of Order-2 Non-stationary Markov Arrival Processes". English. In: *Computer Performance Engineering*. Ed. by M. Beltrán, W. Knottenbelt, and J. Bradley. Vol. 9272. Lecture Notes in Computer Science. Madrid, Spain: Springer International Publishing, 2015, pp. 163–176. ISBN: 978-3-319-23266-9. DOI: 10.1007/978-3-319-23267-6_11.

[C6] A. Mészáros and M. Telek. "Canonical Representation of Discrete Order 2 MAP and RAP". In: *Computer Performance Engineering*. Ed. by S. Balsamo, W. Knottenbelt, and A. Marin. Vol. 8168. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2013, pp. 89–103. ISBN: 978-3-642-40724-6. DOI: 10.1007/978-3-642-40725-3_8.

[J1] A. Mészáros, J. Papp, and M. Telek. "Fitting traffic traces with discrete canonical phase type distributions and Markov arrival processes". In: *Int. J. Appl. Math. Comput. Sci.* 24.3 (2014), pp. 453–470. DOI: 10.2478/amcs-2014-0034.

[J2] A. Mészáros and M. Telek. "Canonical representation of order 2 transient Markov and rational arrival processes". In: *Performance Evaluation* 114 (2017), pp. 78–96.

Independent citations

[C2i1] Y. M. Ko and J. Pender. "Strong approximations for time-varying infinite-server queues with non-renewal arrival and service processes". In: *Stochastic Models* 34.2 (2018), pp. 186–206.

[C2i2] P. Buchholz, I. Felko, and J. Kriege. "Transformation of acyclic phase type distributions for correlation fitting". In: *International Conference on Analytical and Stochastic Modeling Techniques and Applications*. Springer. 2013, pp. 96–111.

[C2i3] P. Buchholz, J. Kriege, and I. Felko. "Parameter Fitting of MAPs". In: *Input Modeling with Phase-Type Distributions and Markov Models*. Springer, 2014, pp. 75–93.

[C3i1] A. Asanjarani. "QBD Modelling of a finite state controller for queueing systems with unobservable Markovian environments". In: *Proceedings of the 11th International Conference on Queueing Theory and Network Applications*. ACM. 2016, p. 20.

[C5i1] J. Rodriguez, R. E. Lillo, and P. Ramirez-Cobo. "Analytical issues regarding the lack of identifiability of the non-stationary MAP2". In: *Performance Evaluation* 102 (2016), pp. 1–20.

[C5i2] M. Bražėnas and E. Valakevičius. “On Structured Initial Solution Generation for Phase Type Fitting with EM Method”. In: *Information Technology And Control* 47.2 (2018), pp. 197–208.

[J1i1] A. Dudin, M. H. Lee, and S. Dudin. “Optimization of the service strategy in a queueing system with energy harvesting and customers’ impatience”. In: *International Journal of Applied Mathematics and Computer Science* 26.2 (2016), pp. 367–378.

[J1i2] A. Brugno et al. “Analysis of an MAP/PH/1 queue with flexible group service”. In: *International Journal of Applied Mathematics and Computer Science* 27.1 (2017), pp. 119–131.

[J1i3] S. Balsamo et al. “Applying reversibility theory for the performance evaluation of reversible computations”. In: *International Conference on Analytical and Stochastic Modeling Techniques and Applications*. Springer. 2016, pp. 45–59.

[J1i4] V. M. Vishnevskii and A. N. Dudin. “Queueing systems with correlated arrival flows and their applications to modeling telecommunication networks”. In: *Automation and Remote Control* 78.8 (2017), pp. 1361–1403.

[J1i5] Y. Xue et al. “Abnormal Prediction of Dense Crowd Videos by a Purpose–Driven Lattice Boltzmann Model”. In: *International Journal of Applied Mathematics and Computer Science* 27.1 (2017), pp. 181–194.

References

[1] E. Altman. *Constrained Markov decision processes*. Vol. 7. CRC Press, 1999.

[2] A. T. Andersen and B. F. Nielsen. “A Markovian approach for modeling packet traffic with long-range dependence”. In: *IEEE journal on Selected Areas in Communications* 16.5 (1998), pp. 719–732.

[3] D. F. Anderson and T. G. Kurtz. “Continuous time Markov chain models for chemical reaction networks”. In: *Design and analysis of biomolecular circuits*. Springer, 2011, pp. 3–42.

[4] C. Andrieu et al. “An introduction to MCMC for machine learning”. In: *Machine learning* 50.1-2 (2003), pp. 5–43.

[6] G. Bolch et al. *Queueing networks and Markov chains: modeling and performance evaluation with computer science applications*. John Wiley & Sons, 2006.

[7] C. Boutilier, R. Dearden, and M. Goldszmidt. “Stochastic dynamic programming with factored representations”. In: *Artificial Intelligence* 121.1 (2000), pp. 49–107.

[8] A. Briggs, M. Sculpher, et al. “An introduction to Markov modelling for economic evaluation”. In: *Pharmacoeconomics* 13.4 (1998), pp. 397–410.

[9] P. Buchholz. “An EM-algorithm for MAP fitting from real traffic data”. In: *International Conference on Modelling Techniques and Tools for Computer Performance Evaluation.* Springer. 2003, pp. 218–236.

[10] P. Buchholz and J. Kriege. “A heuristic approach for fitting MAPs to moments and joint moments”. In: *Quantitative Evaluation of Systems, 2009. QEST'09. Sixth International Conference on the.* IEEE. 2009, pp. 53–62.

[11] D. V. Djonin and V. Krishnamurthy. “MIMO transmission control in fading channels—a constrained Markov decision process formulation with monotone randomized policies”. In: *IEEE Transactions on Signal Processing* 55.10 (2007), pp. 5069–5083.

[12] D. Efrosinin. “Controlled Queueing Systems with Heterogeneous Servers”. PhD thesis. University of Trier, 2004.

[13] A. Horváth and M. Telek. “A Markovian point process exhibiting multifractal behaviour and its application to traffic modeling”. In: *Proceedings of 4th International Conference on Matrix-Analytic Methods in Stochastic models.* 2002, pp. 183–208.

[14] G. Horváth, P. Buchholz, and M. Telek. “A MAP fitting approach with independent approximation of the inter-arrival time distribution and the lag correlation”. In: *Quantitative Evaluation of Systems, 2005. Second International Conference on the.* IEEE. 2005, pp. 124–133.

[15] G. Horváth and M. Telek. “On the canonical representation of phase type distributions”. In: *Performance Evaluation* 66.8 (2009), pp. 396–409.

[16] I. Horváth, J. Papp, and M. Telek. “On the Canonical Representation of Order 3 Discrete Phase Type Distributions”. In: *Electronic Notes in Theoretical Computer Science* 318 (2015), pp. 143–158.

[17] K. Jagannathan et al. “A state action frequency approach to throughput maximization over uncertain wireless channels”. In: *Internet Mathematics* 9.2-3 (2013), pp. 136–160.

[18] L. P. Kaelbling, M. L. Littman, and A. W. Moore. “Reinforcement learning: A survey”. In: *Journal of artificial intelligence research* 4 (1996), pp. 237–285.

[19] L. Kleinrock. *Theory, Volume 1, Queueing Systems.* New York, NY, USA: Wiley-Interscience, 1975. ISBN: 0471491101.

[20] L. Kleinrock. *Queueing systems, volume 2: Computer applications.* Vol. 66. wiley New York, 1976.

[21] Y. L. Kocaga and A. R. Ward. “Admission control for a multi-server queue with abandonment”. English. In: *Queueing Systems* 65.3 (2010), pp. 275–323. ISSN: 0257-0130. DOI: `10.1007/s11134-010-9176-z`. URL: `http://dx.doi.org/10.1007/s11134-010-9176-z`.

[22] S. Koenig and R. Simmons. “Xavier: A robot navigation architecture based on partially observable markov decision process models”. In: *Artificial Intelligence Based Mobile Robotics: Case Studies of Successful Robot Systems* (1998), pp. 91–122.

[23] G. Latouche and V. Ramaswami. *Introduction to matrix analytic methods in stochastic modeling*. SIAM, 1999.

[24] A. van de Liefvoort. *The moment problem for continuous distributions*. Tech. rep. WP-CM-1990-02. University of Missouri – Kansas City, USA: School of Computing and Engineering, 1990.

[25] M. L. Littman, T. L. Dean, and L. P. Kaelbling. “On the Complexity of Solving Markov Decision Problems”. In: *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence*. UAI’95. Montreal, Quebec;, Canada: Morgan Kaufmann Publishers Inc., 1995, pp. 394–402. ISBN: 1-55860-385-9. URL: `http://dl.acm.org/citation.cfm?id=2074158.2074203`.

[26] J. Papp and M. Telek. “Canonical representation of discrete phase type distributions of order 2 and 3”. In: *Proc. of UK Performance Evaluation Workshop, UKPEW*. Vol. 2013. 2013.

[27] L. R. Rabiner. “A tutorial on hidden Markov models and selected applications in speech recognition”. In: *Proceedings of the IEEE* 77.2 (1989), pp. 257–286.

[28] J. Slegers, I. Mitrani, and N. Thomas. “Optimal dynamic server allocation in systems with on/off sources”. In: *Formal Methods and Stochastic Models for Performance Evaluation*. Springer, 2007, pp. 186–199.

[29] M. Telek and G. Horváth. “A minimal representation of Markov arrival processes and a moments matching method”. In: *Performance Evaluation* 64.9 (2007), pp. 1153–1168.

[30] J. A. Whittaker and M. G. Thomason. “A Markov chain model for statistical software testing”. In: *IEEE Transactions on Software engineering* 20.10 (1994), pp. 812–824.