Proceedings of the 2012 Winter Simulation Conference
C. Laroque, J. Himmelspach, R. Pasupathy, O. Rose, and A. M. Uhrmacher, eds.

TRAFFIC MODELING WITH A COMBINATION OF PHASE-TYPE DISTRIBUTIONS AND ARMA PROCESSES

Jan Kriege
Peter Buchholz

Department of Computer Science,
TU Dortmund,
D-44221 Dortmund, GERMANY

ABSTRACT

The adequate modeling of correlated input processes is necessary to obtain realistic models in areas like computer or communication networks but is still a challenge in simulation modeling. In this paper we present a new class of stochastic processes which has been developed for describing correlated input processes and combines acyclic Phase-type distributions to model the marginal distribution with an ARMA process to capture the autocorrelation. The processes are an extension of ARTA processes, a well established input model in stochastic simulations. For the new process type we propose a fitting algorithm that allows one to approximate arbitrary sets of joint moments and autocorrelation coefficients and investigate the effect of different sets of approximated quantities on the quality of the fitted process empirically. We furthermore present an efficient way to generate random numbers and show how the processes can be easily integrated into simulation models.

1 INTRODUCTION

It is known for more than a decade that the load in computer networks has a high variance and a strong autocorrelation over several lags (Leland et al. 1994, Paxson and Floyd 1995). Consequently, the adequate modeling of traffic processes in simulation models is of outstanding importance for the accurate and meaningful analysis of networks. However, building stochastic models of correlated arrival streams is a complex problem which is not supported by available software tools for input modeling (Law and McComas 2003) and is also only briefly considered in simulation textbooks (Law and Kelton 2000).

A general and flexible approach for modeling correlated input streams is the ARTA process (Biller and Nelson 2005) which combines autoregressive processes with different distributions. This model has been extended in (Kriege and Buchholz 2011a) to combine autoregressive processes with so called Phase-type distributions, a very general and flexible class of distributions. The new class of models is denoted as *CAPP* (Correlated Acyclic Phase-type Processes). Both mentioned approaches, ARTA processes and CAPPs, try to fit the distribution function of the observed process and add the dependency structure using an autoregressive model which is parameterized to approximate the coefficient of autocorrelation. It is known that the coefficient of autocorrelation captures only part of the dependency structure, additional measures are the joint distribution or higher order joint moments.

In this paper, we present an extended approach that allows one to fit the parameters of a CAPP according to higher order joint moments. It is shown how the models are parameterized and an approach is presented to integrate CAPPs into simulation models of computer networks using the simulation tool OMNeT++ (Hornig and Varga 2008). By building models for two network traces it is shown that the integration of higher order joint moments results in more accurate simulation models. The structure of the paper is as follows: In the next section we introduce the basic ingredients of our approach and describe the background. Section 3 introduces the fitting approach to parameterize a CAPP. Afterwards, we show how CAPPs can

be used in simulation models and in Section 5 some experimental results are presented. The paper ends with the conclusions.

2 BASIC DEFINITIONS AND RELATED WORK

In this section we briefly introduce the basic ingredients of our approach.

2.1 Traces and their Characteristics

A trace $\mathscr{T} = (t_1, t_2, \ldots, t_r)$ is a sequence of interevent times t_i that describe the time distance between events i and $i+1$ ($1 \leq i < r$). Traces are measured in networks and build the base for a load description in simulation models. From the trace various characteristics and statistical properties can be computed. Common measures are the empirical moments and the variance

$$\hat{\mu}_i = \frac{1}{r}\sum_{j=1}^{r}(t_j)^i, \qquad \hat{\sigma}^2 = \frac{1}{r-1}\sum_{i=1}^{r}(t_i - \hat{\mu}_1)^2$$

which are estimates for the moments $E[X^i]$ and variance $Var(X)$ of the stochastic process X_t that is assumed to generate the load. For our work the empirical lag-h autocorrelation coefficients $\hat{\rho}_h$ and the joint moments $\hat{\nu}_{k,l,h}$ are of importance which are given by

$$\hat{\rho}_h = \frac{1}{(r-h-1)\hat{\sigma}^2}\sum_{i=1}^{r-h}(t_i - \hat{\mu}_1)(t_{i+h} - \hat{\mu}_1) \qquad \text{and} \qquad \hat{\nu}_{k,l,h} = \frac{1}{r-h}\sum_{i=1}^{r-h}(t_i)^k(t_{i+h})^l, \tag{1}$$

respectively. $\hat{\nu}_{k,l,h}$ is an estimate of the joint moment $E[X_t^k, X_{t+h}^l]$ and $\hat{\rho}_h$ is a normalized estimate of $E[X_t^1, X_{t+h}^1] = E[X_t, X_{t+h}]$. Both are measures of the process X_t.

2.2 Phase-Type Distributions

Phase-type (PH) distributions describe independent identically distributed random variables as absorption times of a finite Markov chain (Neuts 1979, Neuts 1981). A PH distribution of order n consists of a Markov chain with n transient and one absorbing state and is defined by an $n \times n$ matrix D_0 and an initial distribution vector π. Matrix D_0 is the generator of an absorbing continuous time Markov chain. Events are generated whenever an absorption occurs and the process is restarted immediately afterwards as defined by distribution π. The moments and the cumulative distribution function of a PH distribution are given by

$$\mu_i = E(X^i) = i!\pi M^i \mathbf{1} \qquad \text{and} \qquad F_X(t) = 1 - \pi exp(D_0 t)\mathbf{1} \tag{2}$$

where $M = -(D_0)^{-1}$ and $\mathbf{1}$ is a column vector of ones of length n.

The class of PH distributions contains several subclasses like Erlang, Hyper-Exponential or Hypo-Exponential that exhibit special structures for matrix D_0 and vector π. For example acyclic Phase-type distributions are characterized by an upper triangular matrix D_0.

For fitting the parameters of Phase-type distributions according to estimates resulting from a trace, various approaches exist that can be roughly divided into two classes. Algorithms from the first class try to maximize the likelihood value and work on the complete trace. Often Expectation Maximization (EM) algorithms are used for this purpose. For the second class some characteristics like moments are derived from the trace and the PH distribution is fitted to these characteristics. General PH distributions have been fitted in (Asmussen, Nerman, and Olsson 1996) using an EM algorithm, but the approach is very slow in the general case. Faster EM algorithms have been tailored to different subclasses of PH distributions. E.g., in (Khayari, Sadre, and Haverkort 2003) an EM algorithm was used to fit Hyper-Exponential distributions and in (Thümmler, Buchholz, and Telek 2006) the EM algorithm has been applied for fitting Hyper-Erlang distributions. An example approach for fitting PH distributions according to empirical moments can be found in (Buchholz and Kriege 2009).

2.3 Autoregressive Moving Average Processes

Autoregressive Moving Average Processes of order p and q are well established in time series modeling (see (Box and Jenkins 1970)). An $ARMA(p,q)$ process is defined as

$$Z_t = \alpha_1 Z_{t-1} + \alpha_2 Z_{t-2} + ... + \alpha_p Z_{t-p} + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + ... + \beta_q \varepsilon_{t-q} + \varepsilon_t \tag{3}$$

where the α_i are p autoregressive coefficients, the β_j are q moving average coefficients and the values ε_t called *innovations* have normal distributions with mean zero and variance σ_ε^2. For $q = 0$ the process becomes an *Autoregressive Process* ($AR(p)$) and for $p = 0$ an *Moving Average Process* ($MA(q)$).

Methods for fitting the parameters of ARMA processes according to measures of a trace are available and are implemented in standard statistical software tools (Box and Jenkins 1970). Observe, from Eq. 3 that the marginal distribution of these models is always given as a weighted sum of $N(0,\sigma_\varepsilon^2)$ random variables. Therefore, only models with marginal distributions that can be derived from a normal distribution can be adequately captured by ARMA processes.

2.4 Autoregressive-To-Anything Processes

ARTA processes (Biller and Nelson 2005, Cario and Nelson 1996) combine an $AR(p)$ base process as defined in Sect. 2.3 with an arbitrary marginal distribution F_Y and are defined as a sequence $Y_t = F_Y^{-1}[\Phi(Z_t)]$ ($t = 1,2,...$) where F_Y is the marginal distribution, Φ is the standard normal cumulative distribution function and $\{Z_t; t = 1,2,...\}$ is a stationary Gaussian $AR(p)$ process that is constructed such that the distribution of the $\{Z_t\}$ is $N(0,1)$ (cf. (Cario and Nelson 1996)). The probability-integral transformation $U_t = \Phi(Z_t)$ ensures that $U(t)$ has uniform distribution on $(0,1)$ (cf. (Devroye 1986)) and the transformation $Y_t = F_Y^{-1}[U_t]$ results in a time series $\{Y_t, t = 1,2,...\}$ with the desired marginal distribution F_Y. (Cario and Nelson 1996, Biller and Nelson 2005) established a relation between the autocorrelation structures of the *ARTA* process and the base process, gave an efficient numerical procedure to construct an $AR(p)$ base process such that the *ARTA* process has the desired autocorrelations that are e.g., estimated from a trace and presented two approaches for fitting ARTA processes to a trace. The first approach from (Cario and Nelson 1996) only determines the $AR(p)$ base process for a given marginal distribution and a trace, while the second fits a Johnson marginal distribution and the base process.

ARTA processes are applicable for all marginal distributions F_Y for which a closed-form expression for the inverse cdf exists or for which F_Y^{-1} can be computed efficiently using a numerical procedure. For the interesting class of Phase-type distributions, for which it is known that they can represent every distribution with a strictly positive density in $(0,\infty)$ (O'Cinneide 1990), the cdf contains a matrix exponential (cf. Eq. 2) and F_Y^{-1} cannot be computed in an efficient way. In (Kriege and Buchholz 2011a) a different approach to combine Phase-type distributions with a base process is presented that will be summarized in the following section. Afterwards this approach is extended to use arbitrary joint moments instead of only autocorrelation coefficients.

3 FITTING CORRELATED ACYCLIC PHASE-TYPE PROCESSES

A common measure for the dependence are autocorrelation coefficients which are directly related to the joint moment $E[Y_t, Y_{t+h}]$. However, more general joint moments $E[Y_t^k, Y_{t+h}^l]$ capture additional parts of the dependence structure of a trace or a stochastic process which is not visible in the first joint moments which define the autocorrelation coefficients. In (Kriege and Buchholz 2011a) Correlated Acyclic Phase-type Processes (CAPP) that combine an acyclic Phase-type distribution with an $ARMA(p,q)$ base process have been introduced and a fitting algorithm for CAPPs using the empirical autocorrelation coefficients has been presented. In the following the main results from (Kriege and Buchholz 2011a) are summarized. Afterwards we establish the theoretical framework to fit CAPPs according to arbitrary joint moments. In Sect. 5 we will investigate the effect of different sets of joint moments on the fitting quality.

3.1 Correlated Acyclic Phase-Type Processes

As already mentioned CAPPs of order (n,p,q) (denoted as $CAPP(n,p,q)$) combine an acyclic Phase-type distribution of order n with an $ARMA(p,q)$ base process. The acyclic Phase-type distribution is expressed in terms of its elementary series (Cumani 1982), where each series describes one path from an initial state to the absorbing state, and the $ARMA(p,q)$ process is used to select series, thereby introducing correlation into the sequence of Phase-type distributed random variables. More precisely, we can divide an PH distribution with initial probability vector π and transition rate matrix D_0 into m elementary series where each series describes one path from an initial state to the absorbing state. The i-th series is then defined in terms of a vector Λ_i that contains the transition rates of the states of the series and an associated probability τ_i that is computed from the the transition rates along the path and the initial probability of the first state of the path, i.e. for a series consisting of the states $i_1, i_2, \ldots, i_k$ the vector of transition rates is given by $\Lambda_i = (-D_0(i_1,i_1), -D_0(i_2,i_2), \ldots, -D_0(i_k,i_k))$ and the initial probability is given by

$$\tau_i = \pi_{i_1} \frac{D_0(i_1,i_2)}{-D_0(i_1,i_1)} \frac{D_0(i_2,i_3)}{-D_0(i_2,i_2)} \cdots \frac{t(i_k)}{-D_0(i_k,i_k)}$$

where $t(i_k)$ denotes the transition rate from state i_k to the absorbing state. An example for an acyclic Phase-type distribution and its elementary series is given in Fig. 1.

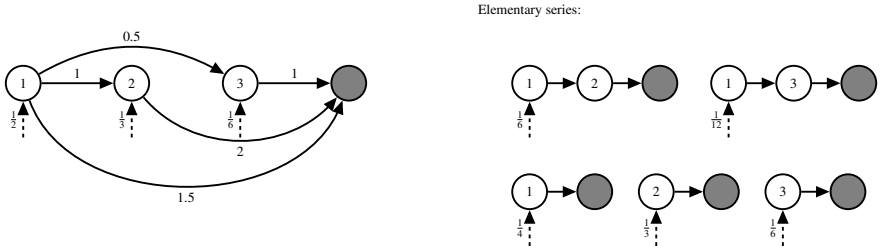


Figure 1: An APH and its elementary series

We then define disjoint subintervals $[\underline{b}_i, \bar{b}_i]$ on $[0,1]$ for each series and an indicator function $\delta(U,i)$, i.e.

$$\begin{array}{rcll} \underline{b}_1 & = & 0 & \\ \bar{b}_i & = & \underline{b}_i + \tau_i & i = 1,...,m \\ \underline{b}_i & = & \bar{b}_{i-1} & i = 2,...,m \end{array}, \qquad \delta(U,i) = \begin{cases} 1, & U \in [\underline{b}_i, \bar{b}_i) \\ 0, & \text{otherwise.} \end{cases} \tag{4}$$

Now, let $\{X_t^{(\Lambda_i)}\}$ be sequences of iid random variables with Hypo-Exponential distribution described by a vector of rates Λ_i with length S_i that contains the transition rates of the i-th series and let $\{Z_t\}$ be a sequence of correlated random variables with standard normal distribution generated by an $ARMA(p,q)$ process as defined in Eq. 3. Setting $U_t = \Phi(Z_t)$ where Φ is the standard normal cdf we obtain a sequence $\{U_t\}$ of correlated random variables with uniform distribution on $(0,1)$ (cf. (Devroye 1986)) and can construct the CAPP description

$$Y_t = \sum_{i=1}^{m} \delta(U_t, i) X_t^{(\Lambda_i)} \tag{5}$$

that uses the elementary series to describe a sequence of correlated random variables with the same acyclic Phase-type distribution that the elementary series have been computed from (Kriege and Buchholz 2011a).

The autocorrelation of the CAPP is given by

$$Corr[Y_t, Y_{t+h}] = \frac{E[Y_t, Y_{t+h}] - E[Y]^2}{Var[Y]}$$

and $E[Y]$ and $Var[Y]$ are given by the Phase-type distribution and can be computed using Eq. 2. For the remaining term $E[Y_t, Y_{t+h}]$ one obtains (Kriege and Buchholz 2011a)

$$\begin{aligned} E[Y_t, Y_{t+h}] &= E\left[\left(\sum_{i=1}^{m} \delta(U_t, i) X_t^{(\Lambda_i)}\right)\left(\sum_{j=1}^{m} \delta(U_{t+h}, j) X_{t+h}^{(\Lambda_j)}\right)\right] \\ &= \sum_{i,j}\left(\left(\sum_{s=1}^{S_i} \frac{1}{\Lambda_i(s)}\right)\left(\sum_{s=1}^{S_j} \frac{1}{\Lambda_j(s)}\right)\int_{\Phi^{-1}(\underline{b}_j)}^{\Phi^{-1}(\bar{b}_j)}\int_{\Phi^{-1}(\underline{b}_i)}^{\Phi^{-1}(\bar{b}_i)} \varphi_{\rho_h}(z_t, z_{t+h}) dz_t dz_{t+h}\right) \end{aligned} \tag{6}$$

where $\varphi_{\rho_h}(z_t, z_{t+h})$ is the bivariate standard normal density function with correlation $\rho_h = Corr[Z_t, Z_{t+h}]$. In a fitting approach, parameters of the stochastic model are set to approximate the empirical measures from a trace (i.e., $E[Y^i] \approx \hat{\mu}_i$ and $E[Y_t, Y_{t+h}] \approx \hat{\nu}_{1,1,h}$).

Because the ARMA base process is short-range dependent, CAPPs are short-range dependent as well. However, it has been shown in (Horvath and Telek 2002) that in practice short-range dependent processes can approximate long-range dependent behavior sufficiently close.

3.2 Computing arbitrary Joint Moments of a CAPP

In the following we will generalize the ideas from the previous section to compute arbitrary joint moments of a CAPP. Eq. 6 already gives the formula for the first joint moment which is needed for the computation of the autocorrelation. For the computation of arbitrary joint moments $E[Y_t^k, Y_{t+h}^l]$ we obtain

$$\begin{aligned} \nu_{k,l,h} = E[Y_t^k, Y_{t+h}^l] &= E\left[\left(\sum_{i=1}^{m} \delta(U_t, i) X_t^{(\Lambda_i)}\right)^k, \left(\sum_{j=1}^{m} \delta(U_{t+j}, i) X_{t+h}^{(\Lambda_j)}\right)^l\right] \\ &= \sum_{i,j} E\left[\delta(U_t, i) {X_t^{(\Lambda_i)}}^k, \delta(U_{t+h}, j) {X_{t+h}^{(\Lambda_j)}}^l\right] \\ &= \sum_{i,j}\left(E\left[\delta(U_t, i)\delta(U_{t+h}, j)\right] E\left[X_t^{(\Lambda_i)^k}\right] E\left[X_{t+h}^{(\Lambda_j)^l}\right]\right) \\ &= \sum_{i,j}\left(E\left[\delta(\Phi(Z_t), i)\delta(\Phi(Z_{t+h}), j)\right] \mu_k^{(hexp)}(S_i, \Lambda_i)\mu_l^{(hexp)}(S_j, \Lambda_j)\right) \\ &= \sum_{i,j}\Big(\mu_k^{(hexp)}(S_i, \Lambda_i)\mu_l^{(hexp)}(S_j, \Lambda_j) \\ &\quad \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \delta(\Phi(z_t), i)\delta(\Phi(z_{t+h}), j)\varphi_{\rho_h}(z_t, z_{t+h}) dz_t dz_{t+h}\Big) \\ &= \sum_{i,j}\Big(\mu_k^{(hexp)}(S_i, \Lambda_i)\mu_l^{(hexp)}(S_j, \Lambda_j) \\ &\quad \int_{\Phi^{-1}(\underline{b}_j)}^{\Phi^{-1}(\bar{b}_j)}\int_{\Phi^{-1}(\underline{b}_i)}^{\Phi^{-1}(\bar{b}_i)} \varphi_{\rho_h}(z_t, z_{t+h}) dz_t dz_{t+h}\Big) \end{aligned} \tag{7, 8}$$

where $\mu_k^{(hexp)}(S_i, \Lambda_i)$ is the k-th moment of a Hypo-Exponential distribution with S_i phases and rate vector Λ_i. Note, that the Eq. 7 holds because for a fixed U we have that $\delta(U, i) = 1$ for exactly one i and $\delta(U, j) = 0$ for all $i \neq j$. Eq. 8 holds because $\delta(U, i) = 1$ for $U \in [\underline{b}_i, \bar{b}_i)$ and 0 otherwise. Therefore, $\delta(\cdot)$ can be omitted in the double integral and used to determine the integration bounds. Note, that $\mu_k^{(hexp)}(S_i, \Lambda_i)$ can easily be computed using Eq. 2 and for the computation of the bivariate normal integral fast numerical procedures exist (Drezner and Wesolowsky 1990). The goal is then to choose the parameters such that $E[Y_t^k, Y_{t+h}^l] \approx \hat{\nu}_{k,l,h}$.

3.3 An Algorithm for Fitting CAPPs according to Joint Moments

Using Eq. 8 we can sketch the following fitting algorithm for CAPPs using arbitrary joint moments, which is a generalization of the algorithm from (Kriege and Buchholz 2011a).

The algorithm from Listing 1 expects an acyclic Phase-type distribution, a trace file with empirical observations from a real system, the order (p,q) of the ARMA base process and the power and lag (k,l,h) of joint moments $E[Y_t^k, Y_{t+h}^l]$ as inputs. For fitting the Phase-type distribution any of the existing approaches (cf. Sect. 2.2) can be used. For our implementation we used the approaches from (Thümmler, Buchholz, and Telek 2006) that fits an Hyper-Erlang distribution with an Expectation Maximization algorithm and the moment fitting technique from (Buchholz and Kriege 2009) that fits a general acyclic Phase-type distribution.

Listing 1: Steps for fitting CAPPs

```
Inputs:
  acyclic Phase-type distribution APH(n)
  Trace T = (t_1,t_2,...,t_r)
  p,q: number of AR and MA coefficients
  J = {(k_1,l_1,h_1),(k_2,l_2,h_2),...,(k_s,l_s,h_s)}: set of empirical joint moments used for fitting

1) determine and sort elementary series of APH(n)
2) compute empirical joint moments v̂_{k,l,h}, ∀(k,l,h) ∈ J from T according to Eq. 1
3) for each (k_i,l_i,h_i) determine ARMA autocorrelation ρ_{k_i,l_i,h_i} using Eq. 8
4) minimize Eq. 9 to find an ARMA(p,q) model for the ρ_{k,l,h}
5) set variance of ARMA(p,q) according to Eq. 10
6) return CAPP(n,p,q) model with base ARMA(p,q) process and distribution APH(n)
```

In the first step of the algorithm the elementary series of $APH(n)$ are computed. Additionally, the elementary series are sorted according to their expected durations, i.e. $i \leq j \Rightarrow E\left[X^{(\Lambda_i)}\right] \leq E\left[X^{(\Lambda_j)}\right]$. This is not a restriction because the order of the series has no effect on the distribution, but a requirement for a correct modeling of the autocorrelation (Kriege and Buchholz 2011a). In the second step the empirical joint moments of the trace $\hat{\nu}_{k,l,h}$ are computed according to Eq. 1. These are the desired joint moments that the CAPP should exhibit at the end of the fitting algorithm. Note, that if $k = l = 1$ the algorithm fits according to the joint moments $E[Y_t, Y_{t+h}]$ that correspond to the lag-h autocorrelation coefficients.

Steps 3-5 of the algorithm from Listing 1 describe the construction of the $ARMA(p,q)$ base process. The $ARMA(p,q)$ base process has to observe two requirements. First, it has to exhibit an autocorrelation structure such that the $CAPP(n,p,q)$ approximates the given set of joint moments estimated from the trace in step 2. Second, the $ARMA(p,q)$ base process from Eq. 3 has a standard normal distribution, i.e., the variance (or autocovariance at lag 0) of the $ARMA(p,q)$ process has to be 1. Hence, the construction of the ARMA base process consists of three parts. We have to determine the autocorrelation structure of the base process depending on the desired joint moments of the CAPP model (step 3 of the algorithm from Listing 1), we have to fit an ARMA model according to that autocorrelation structure (step 4 of the algorithm) and we have to modify the ARMA model to have a standard normal distribution (step 5 of the algorithm).

In step 3 of the algorithm for each empirical joint moment $\hat{\nu}_{k,l,h}$ the corresponding autocorrelation $\rho_h = \rho_{k,l,h}$ of the $ARMA(p,q)$ base process is determined. Then the parameters of the ARMA process are computed such that the CAPP has the joint moment $\nu_{k,l,h}$ according to Eq. 8 match the empirical moments $\hat{\nu}_{k,l,h}$. Observe that Eq. 8 defines how to compute $\nu_{k,l,h} = E[Y_t^k, Y_{t+h}^l]$ from a given process description, but includes no method to set the process parameters to match predefined values $\nu_{k,l,h}$. In our case $\hat{\nu}_{k,l,h}$ is given (i.e., estimated from the trace) and we have to determine the corresponding $\rho_{k,l,h}$ that determine the parameters of the ARMA process. This has to be done numerically by a simple line search algorithm (Press et al. 1993). Note, that for the CAPP fitting approach using the autocorrelations instead of joint

moments from (Kriege and Buchholz 2011a) and for the original ARTA approach from (Cario and Nelson 1996) it is also necessary to determine the base process autocorrelation numerically.

Once the base process autocorrelations $(\rho_{k_1,l_1,h_1}, \rho_{k_2,l_2,h_2}, \cdots, \rho_{k_s,l_s,h_s})$ have been determined we need to construct an $ARMA(p,q)$ that exhibits these autocorrelations, which is done in step 4 of the algorithm. This can be done by using a general purpose optimization algorithm like (Nelder and Mead 1965, Nash 1990) to minimize

$$min \sum_{i=1}^{s} \left(\frac{\bar{\rho}_{k_i,l_i,h_i}}{\rho_{k_i,l_i,h_i}} - 1 \right)^2 \tag{9}$$

where $\bar{\rho}_{k_i,l_i,h_i}$ are the autocorrelations of the $ARMA(p,q)$ model that is constructed during the minimization process and ρ_{k_i,l_i,h_i} are the autocorrelations to be achieved. See e.g., (Brockwell and Davis 1998) how to compute the autocorrelations of an $ARMA(p,q)$ process.

Step 5 of the algorithm consists in adjusting the variance such that the base process has a standard normal marginal distribution. The autocorrelation of a stationary $ARMA(p,q)$ process is given by $\rho_{1,1,h} = \rho_h = \gamma(h)/\gamma(0)$, where $\gamma(h)$ is the autocovariance at lag h. Since both $\gamma(h)$ and $\gamma(0)$ depend on the variance of the innovations σ_ε^2 (cf. (Brockwell and Davis 1998)) the fraction can be reduced by eliminating σ_ε^2 and thus, ρ_h is independent of σ_ε^2. Consequently, we can set σ_ε^2 such that the $\{Z_t; t = 1,2,...\}$ have a standard normal distribution without modifying the autocorrelation structure of the base process. For a given $ARMA(p,q)$ process with $\tilde{\sigma}_\varepsilon^2$ being the variance of the white noise and $\tilde{\gamma}(0)$ being the autocovariance at lag 0 we can set the new variance to

$$\sigma_\varepsilon^2 = \tilde{\sigma}_\varepsilon^2 / \tilde{\gamma}(0) \tag{10}$$

resulting in a new $N(0,1)$ process with the same autocorrelations as the old process.

Finally, the $ARMA(p,q)$ base process and the Phase-type distribution $APH(n)$ can be combined and the $CAPP(n,p,q)$ model is returned in step 6 of the algorithm.

4 USING CORRELATED ACYCLIC PHASE-TYPE PROCESSES IN SIMULATION MODELS

Stochastic processes have to fulfill several requirements to be used in simulation models. First, fitting the processes, i.e., setting the parameters according to observations measured from a real system, should be possible in an efficient way. For CAPPs we have outlined an efficient fitting approach in the previous section. Second, random number generation, i.e., sampling from the process, should be possible in an efficient way and the integration of the processes into simulation models has to be possible without much additional effort. In the following we will outline random number generation and the integration into simulation models for CAPPs.

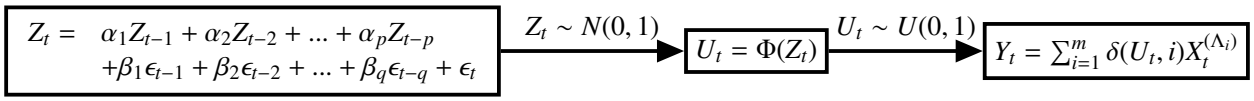


Figure 2: Steps for Generating Random Numbers from Correlated Acyclic Phase-type Processes

The basic steps for sampling from a CAPP are shown in Fig. 2. In the first step a normally distributed random number ε_t is drawn that serves as new innovation and the weighted sum according to Eq. 3 is computed to obtain the next Z_t. The Z_t is transformed into $U_t = \Phi(Z_t)$ with uniform distribution and U_t is used to determine the elementary series i for which $\delta(U_t, i) = 1$. Finally, a random number from a Hypo-Exponential distribution that corresponds to series i is drawn to determine the next element of the time series Y_t.

Since for an $ARMA(p,q)$ process it is assumed that p previous observations and q previous innovations exist, i.e., when generating Z_t the previous observations $Z_{t-1}, \cdots, Z_{t-p}$ and the innovations $\varepsilon_{t-1}, \cdots, \varepsilon_{t-q}$ are required, the $ARMA(p,q)$ should be initialized with suitable values for the previous observations and innovations when generating the first observations $Z_1, Z_2, \cdots$. For the innovations this is straightforward,

because they are iid random numbers, and consequently the $\varepsilon_{t-1},\ldots,\varepsilon_{t-q}$ can be drawn from a normal distribution with zero mean and variance σ_ε^2. For the initialization of the p previous observations we have to observe that they are correlated and hence, they have to be drawn from a multivariate normal distribution, which can be done using standard theory (Law and Kelton 2000). Let

$$\Sigma = \begin{bmatrix} r_0 & r_1 & \cdots & r_{p-1} \\ r_1 & r_0 & \cdots & r_{p-2} \\ \vdots & \vdots & \ddots & \vdots \\ r_{p-1} & r_{p-2} & \cdots & r_0 \end{bmatrix}$$

be the covariance matrix with the autocovariance values $(r_0, r_1, \cdots, r_{p-1})$ of the ARMA process and let $\mu = (\mu, \cdots, \mu)$ be a vector that contains p times the mean value μ of the ARMA process. In our case $\mu = 0$ because we required the Z_t to have standard normal distribution. Now we draw p independent random numbers $X = (X_0, X_1, \cdots, X_{p-1})^T$ from a standard normal distribution and compute the lower triangular matrix C with $\Sigma = CC^T$ by a Cholesky decomposition (Fishman 1973). Then, the $Z_{t-i}, i = 1...p$ are obtained by setting $Z = \mu + CX$.

The approaches for the simulation of CAPPs described above have been integrated into a module for OMNeT++ to allow an easy integration of stochastic processes into simulation models. OMNeT++ (Hornig and Varga 2008) is an open source simulation framework that has been developed and extensively used for the modeling of communication protocols and networks. Models for OMNeT++ consist of modules that can be arranged in a hierarchical way and that communicate via gates with message passing. So called simple modules are the basic building blocks of a model, while compound modules are a collection of simple modules and other compound modules.

In (Kriege and Buchholz 2011b) a module called OMNeT++ Arrival Process Module is presented that generates random numbers from Markovian Arrival Processes, ARTA processes and ARMA processes and has been subsequently extended to generate random numbers from CAPPs. The generation of random numbers from ARMA models and the initialization step described above are already supported by the module from (Kriege and Buchholz 2011b) and therefore it only has to be extended by a class for CAPPs that implements the transformation $U_t = \Phi(Z_t)$ and the generation of the next Y_t from U_t (cf. Fig. 2). As preparation the class computes the elementary series of the APH distribution and sorts them according to their mean values resulting in a list of the rates and the probability for each elementary series. For random number generation the class then chooses the series i for which $\delta(U_t, i) = 1$ and draws m random numbers from an exponential distribution with rates corresponding to the elementary series. The sum of this m random numbers is returned as the next sample.

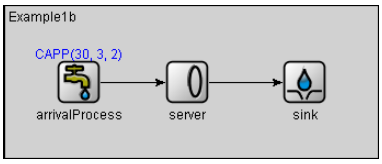


Figure 3: Simple OMNeT++ simulation model with Arrival Process module

A simple OMNeT++ example model consisting of the Arrival Process module initialized with a CAPP and a single server queue is shown in Fig. 3.

5 EXPERIMENTAL RESULTS

The algorithm from Listing 1 can be used to fit CAPPs according to an arbitrary set of joint moments $\mathscr{J}$. In the following we will investigate the effect of different sets $\mathscr{J}$ on the quality of the fitted process. We fitted CAPPs to two different benchmark traces from the Internet Traffic Archive (http://ita.ee.lbl.gov).

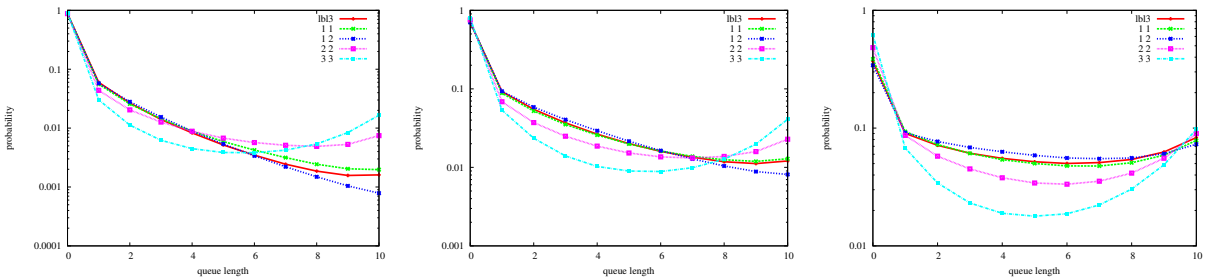
$\nu_{k,l,h}$	LBL-TCP-3	$CAPP(8,5,2)$	$CAPP(12,6,3)$	$CAPP(25,3,3)$
$\nu_{1,1,1}$	1.3013	1.30129	1.30177	1.29945
$\nu_{1,1,2}$	1.2328	1.23285	1.23363	1.23544
$\nu_{1,1,3}$	1.1957	1.19534	1.19933	1.19583
$\nu_{1,1,4}$	1.1741	1.17443	1.17007	1.17277
$\nu_{1,1,5}$	1.1575	1.15781	1.1567	1.15792
$\nu_{1,1,6}$	1.1490	1.1501	1.14836	1.14934
$\nu_{1,1,7}$	1.1434	1.14313	1.14381	1.14354
$\nu_{1,1,8}$	1.1429	1.14039	1.14106	1.14022
$\nu_{1,1,9}$	1.138	1.13727	1.13849	1.13779
$\nu_{1,1,10}$	1.13597	1.13634	1.13716	1.13639
$\nu_{2,2,1}$	17.4164	14.8483	12.9159	16.9957
$\nu_{3,3,1}$	890.542	501.286	323.586	769.042

Table 1: Joint moments of the trace LBL-TCP3 and three fitted CAPPs

The trace BC-pAug89 (Leland et al. 1994) contains a million packet arrivals observed at the Bellcore Morristown Research and Engineering facility in August 1989. The trace LBL-TCP-3 (Paxson and Floyd 1995) contains two hours of TCP traffic from the Lawrence Berkeley Laboratory and was recorded in January 1994. Since one can expect that the CAPPs provide a good approximation of the joint moments in $\mathscr{J}$ that have been used for fitting but a loose approximation of other joint moments, we will not compare the joint moments of the fitted processes and the traces as our primary measure for the fitting quality. Table 1 shows the joint moments $\nu_{k,l,h}$ of different CAPPs fitted to the trace LBL-TCP-3. For all models joint moments with $k = l = 1$ have been used for fitting and as one can see all processes provide a good approximation for these joint moments, but provide only a rough approximation for other joint moments like $\nu_{2,2,1}$. Similar observations can be made for most of the fitted processes.

Therefore, we use the results of a simple queueing model as measures to compare different input processes. The queueing model consists of a single server with a queue capacity of 10 that is fed by arrivals generated by the Arrival Process Module described in Sect. 4. The Arrival Process Module is initialized with the different CAPPs that resulted from the fitting step and we run the model with different utilization levels between $\rho = 0.3$ and $\rho = 0.9$ using the queue length distribution as result measure. To obtain reference values we use a trace driven simulation where the arrivals for the server are generated by the two traces BC-pAug89 and LBL-TCP-3.

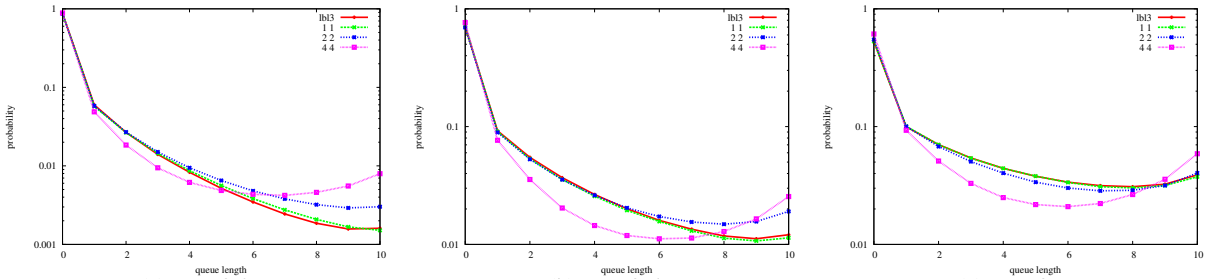
In a first series of experiments we fitted CAPPs according to empirical joint moments $\hat{\nu}_{k,l,h}, h = 1,\cdots,15$ for fixed k and l, e.g. for one CAPP the joint moments $\hat{\nu}_{1,1,1}, \hat{\nu}_{1,1,2}, \cdots, \hat{\nu}_{1,1,15}$ have been used, for another one the joint moments $\hat{\nu}_{2,2,1}, \hat{\nu}_{2,2,2}, \cdots, \hat{\nu}_{2,2,15}$ and so on. Fig. 4 shows the queue length distribution for the



(a) $\rho = 0.3$ (b) $\rho = 0.5$ (c) $\rho = 0.9$

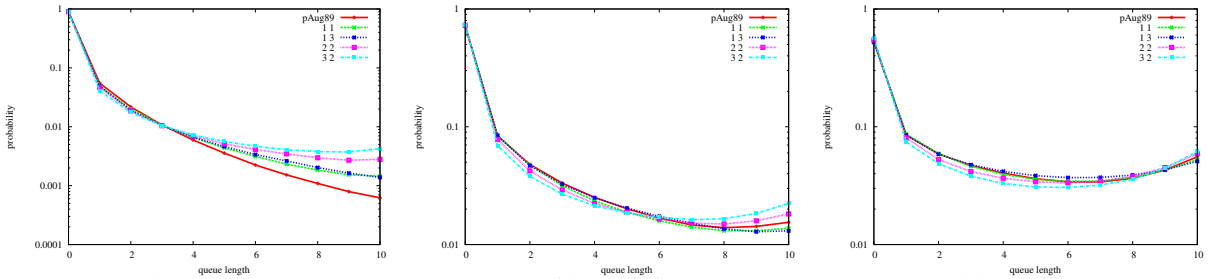
Figure 4: Queue length distribution for the trace LBL-TCP-3 and different $CAPP(12,\cdot,\cdot)$

trace LBL-TCP-3 and different CAPPs with Phase-type distributions of order 12. The labels $k\,l$ indicate which joint moments have been used for fitting, i.e they denote the joint moments $\hat{\nu}_{k,l,h}, h = 1,\cdots,15$. As



(a) $\rho = 0.3$ (b) $\rho = 0.5$ (c) $\rho = 0.7$

Figure 5: Queue length distribution for the trace LBL-TCP-3 and different $CAPP(25,\cdot,\cdot)$



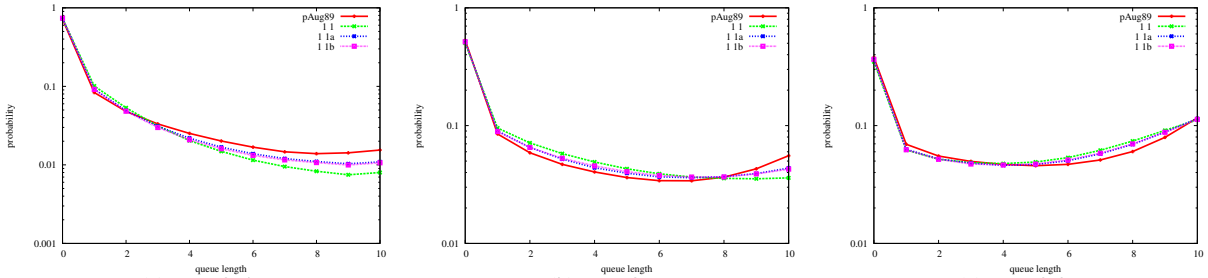
(a) $\rho = 0.3$ (b) $\rho = 0.5$ (c) $\rho = 0.7$

Figure 6: Queue length distribution for the trace BC-pAug89 and different $CAPP(25,\cdot,\cdot)$

one can see the best approximation was obtained for $k = l = 1$ followed by the CAPP with $k = 1, l = 2$. Note, that the plot for $\rho = 0.7$ was omitted in Fig. 4 because it is similar to the plot for $\rho = 0.9$. The results from Fig. 5 that show the results for CAPPs with Phase-type distributions of order 25 confirm these observations. Again we can see that the CAPPs that were fitted according to empirical joint moments $\hat{\nu}_{k,l,h}$ with smaller k and l provided the closest approximation regarding the queue length distribution.

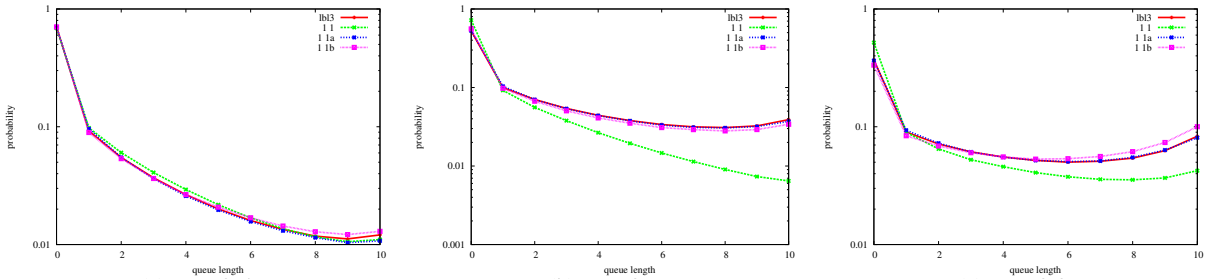
Finally, some results for the trace BC-pAug89 are shown in Fig. 6. Although in this case the approximation of the queue length distribution is better than for the previous trace for all CAPPs, the CAPPs that have been fitted according to smaller joint moments still provide the best approximation.

As we have seen from the previous examples fitting according to smaller joint moments usually yields better results than fitting according to higher joint moments only, since the empirical joint moments get very large for larger k and l values, which makes fitting more difficult. However, a combination of smaller and higher joint moments can be used to increase the quality of the approximation, which we will demonstrate in the following. Fig. 7 shows the queue length distribution for three different CAPPs with Phase-type distributions of order 6. The curve labeled `1 1` was generated by a CAPP that was fitted according to the joint moments $\hat{\nu}_{1,1,h}, h = 1,\cdots,20$. For the other two CAPPs we omitted the higher lags ($h = 16,\cdots,20$) for fitting and added higher joint moments for the lower lags instead, such that the overall number of joint moments considered for fitting remained the same. In particular, for the CAPP corresponding to the curve labeled with `1 1a` we added $\hat{\nu}_{2,2,1}, \hat{\nu}_{1,2,1}, \hat{\nu}_{2,1,1}, \hat{\nu}_{2,2,2}, \hat{\nu}_{2,2,3}$ and for the one with the curve `1 1b` we added $\hat{\nu}_{2,2,1}, \hat{\nu}_{2,2,2}, \hat{\nu}_{2,2,3}, \hat{\nu}_{2,2,4}, \hat{\nu}_{2,2,5}$. These added higher joint moments helped to capture more information of the dependance structure caused by the (usually) more important lower lags and as we can see from Fig. 7 this procedure resulted in an improvement of the approximation of the queue length distribution without increasing the computational effort for fitting. These observations are confirmed by the results from Fig. 8 that shows the results for different CAPPs that have been fitted to the trace LBL-TCP-3 in a similar setup as described above. Again, using higher joint moments for the lower lags increased the approximation of the queueing behavior, especially for the higher utilization levels.



(a) $\rho = 0.5$ (b) $\rho = 0.7$ (c) $\rho = 0.9$

Figure 7: Queue length distribution for the trace BC-pAug89 and different $CAPP(6,\cdot,\cdot)$



(a) $\rho = 0.5$ (b) $\rho = 0.7$ (c) $\rho = 0.9$

Figure 8: Queue length distribution for the trace LBL-TCP-3 and different $CAPP(10,\cdot,\cdot)$

6 CONCLUSIONS

The paper presents a new approach for modeling correlated input processes in computer and communication networks. The approach combines ARMA processes and Phase-type distributions and generates a process which is denoted as a CAPP. Phase-type distributions are used for an accurate fitting of the observed distribution using maximum likelihood or moment based fitting methods. Parameters of the ARMA are computed to capture higher order joint moments which is an extension of other available methods that only consider the coefficient of autocorrelation which corresponds to the first joint moment for describing the autocorrelation. An algorithm for generating the process description is presented and it is shown how the resulting CAPPs can be used in simulation models and especially in network simulation models of the software tool OMNeT++ where CAPPs are available as standard modules.

REFERENCES

Asmussen, S., O. Nerman, and M. Olsson. 1996. "Fitting Phase-Type Distributions via the EM Algorithm". *Scandinavian Journal of Statistics* 23 (4): 419–441.

Biller, B., and B. L. Nelson. 2005. "Fitting Time-Series Input Processes for Simulation". *Oper. Res.* 53 (3): 549–559.

Box, G., and G. Jenkins. 1970. *Time series analysis: Forecasting and control*. San Francisco: Holden-Day.

Brockwell, P. J., and R. A. Davis. 1998. *Time Series: Theory and Methods*. 2nd ed. Springer.

Buchholz, P., and J. Kriege. 2009. "A Heuristic Approach for Fitting MAPs to Moments and Joint Moments". In *Proceedings of the 6th International Conference on Quantitative Evaluation of SysTems (QEST 2009)*, 53–62. Los Alamitos, CA, USA: IEEE Computer Society.

Cario, M. C., and B. L. Nelson. 1996. "Autoregressive to anything: Time-series input processes for simulation". *Operations Research Letters* 19 (2): 51–58.

Cumani, A. 1982. "On the Canonical Representation of Homogeneous Markov Processes Modeling Failure-Time Distributions". *Micorelectronics and Reliability* 22 (3): 583–602.

Devroye, L. 1986. *Non-Uniform Random Variate Generation*. New York: Springer.

Drezner, Z., and G. O. Wesolowsky. 1990. "On the computation of the bivariate normal integral". *Journal of Statistical Computation and Simulation* 35:101 – 107.

Fishman, G. S. 1973. *Concepts and Methods in Discrete Event Digital Simulation*. New York: John Wiley.

Hornig, R., and A. Varga. 2008. "An Overview of the OMNeT++ Simulation Environment". In *Proc. of 1st International Conference on Simulation Tools and Techniques for Communications, Networks and Systems (SIMUTools)*.

Horvath, A., and M. Telek. 2002. "Markovian Modeling of Real Data Traffic: Heuristic Phase Type and MAP Fitting of Heavy Tailed and Fractal Like Samples". In *Performance 2002*, 405–434: Springer.

Khayari, R. E. A., R. Sadre, and B. Haverkort. 2003. "Fitting World-Wide Web Request Traces with the EM-Algorithm". *Performance Evaluation* 52:175–191.

Kriege, J., and P. Buchholz. 2011a. "Correlated phase-type distributed random numbers as input models for simulations". *Performance Evaluation* 68 (11): 1247–1260.

Kriege, J., and P. Buchholz. 2011b. "Simulating Stochastic Processes with OMNeT++". In *Proceedings of the 4th International OMNeT++ Workshop (OMNeT++ 2011)*: ICST.

Law, A. M., and W. D. Kelton. 2000. *Simulation Modeling and Analysis*. 3rd ed. Boston: McGraw-Hill.

Law, A. M., and M. G. McComas. 2003. "ExpertFit distribution-fitting software: how the ExpertFit distribution-fitting software can make your simulation models more valid". In *Winter Simulation Conference*, 169–174.

Leland, W., M. Taqqu, W. Willinger, and D. Wilson. 1994. "On the Self-Similar Nature of Ethernet Traffic". *IEEE/ACM Transactions on Networking* 2 (1): 1–15.

Nash, J. 1990. *Compact Numerical Methods for Computers: Linear Algebra and Function Minimisation*. 2. ed. Adam Hilger.

Nelder, J., and R. Mead. 1965. "A simplex method for function minimization". *Computer Journal* 7:308–313.

Neuts, M. F. 1979. "A versatile Markovian point process". *Journal of Applied Probability* 16:764–779.

Neuts, M. F. 1981. *Matrix-geometric solutions in stochastic models*. Johns Hopkins University Press.

O'Cinneide, C. A. 1990. "Characterization of Phase-Type Distributions". *Stochastic Models* 6:1–57.

Paxson, V., and S. Floyd. 1995. "Wide-area traffic: The failure of Poisson modeling". *IEEE/ACM Transactions in Networking* 3:226–244.

Press, W., S. Teukolsky, W. Vetterling, and B. Flannery. 1993. *Numerical Recipes in C - The Art of Scientific Computing*. 2nd ed. Cambridge University Press.

Thümmler, A., P. Buchholz, and M. Telek. 2006. "A Novel Approach for Phase-Type Fitting with the EM Algorithm". *IEEE Transactions on Dependable and Secure Computing* 3 (3): 245–258.

AUTHOR BIOGRAPHIES

JAN KRIEGE received the Diploma degree (2006) and the Doctoral degree (2012) from the TU Dortmund where he is currently a postdoc in the modeling and simulation group. His research interests include the modeling and analysis of logistics networks and computer and communication systems. His email address is jan.kriege@tu-dortmund.de.

PETER BUCHHOLZ received the Diploma degree (1987), the Doctoral degree (1991) and the Habilitation degree (1996) all from the TU Dortmund, where he is currently a professor for modeling and simulation. His current research interests are efficient techniques for the analysis of stochastic models, formal methods for the analysis of discrete event systems, the development of modeling tools, as well as performance and dependability analysis of computer and communication systems. His email address is peter.buchholz@tu-dortmund.de.