The Application of Phase Type Distributions for Modelling Queuing Systems

Eimutis VALAKEVICIUS
Department of Mathematical Research in Systems
Kaunas University of Technology
Kaunas, LT - 51368, Lithuania

ABSTRACT

Queuing models are important tools for studying the performance of complex systems, but despite the substantial queuing theory literature, it is often necessary to use approximations in the case the system is non-markovian. Phase type distribution is by now indispensable tool in creation of queuing system models. The purpose of this paper is to suggest a method and software for evaluating queuing approximations. A numerical queuing model with priorities is used to explore the behaviour of exponential phase-type approximation of service-time distribution. The performance of queuing systems described in the event language is used for generating the set of states and transition matrix between them. Two examples of numerical models are presented – a queuing system model with priorities and a queuing system model with quality control.

Keywords: Queuing Approximation, Distribution Fitting, Phase Type distributions, Markov Chain, Numerical Model

1. INTRODUCTION

The use of queuing models is a basic tool for studying systems involving contention for resources. Major application areas include computing systems, telecommunication systems, and manufacturing systems. Simple models involving restrictive distributional assumptions often provide essential insights into the behaviour of a queuing system. However, to obtain more precise and detailed information about system behaviour, in general one must employ a model that allows specification of greater detail about the actual input distributions (e.g. the service –time and inerarrival-time distributions) of the system.

Use of phase-type (PH) distributions is a common means of obtaining tractable queuing models [1-4]. The approach of this research is to begin with service-time distribution to be approximated. Simple three-moment approximation, along with more refined approximation taking into account distribution shape, is presented for original input distribution. Then both the original and approximating distributions are used in modelling the queue with simple priorities.

It is known that creation of analytical models requires large efforts. Use of numerical methods permits to create models for a wider class of systems. The process of creating numerical models for systems described by Markov chains consists of the following stages: 1) definition of the state of a system; 2) creating equations describing Markov chain; 3) computation of steady state probabilities of Markov chain; 4) computation measures of the system performance. The most difficult stages are obtaining the set of all the possible states of a system and transition matrix between them. In the paper is used a method for automatic construction of numerical models for systems described by Markov chains with a countable space of states and continuous time.

2. APPROXIMATION OF SERVICE TIME DISTRIBUTION

Let us consider the *M*/*G*/c queue. In this multi-server model with *c* servers the arrival process of customers is a Poisson process with rate λ and the service time *T* of a customer has a general probability distribution function *G*(*t*). It is assumed that $\rho = \lambda E(S)/c$ is smaller than 1. The *M*/*G*/*c* queue with general service times permits no simple analytical solution, not even for the average waiting time. Useful approximation can be obtained by the mixture and convolutions of exponential (phase-type) distributions. Then a Markov chain with a countable space of states and continuous time can represent the evolution of the system. Suppose we let $m_k, k = \overline{1,3}$ denote the *k*th non - central moment (i.e. $E[T^k]$, where T is a random variable of service time).

Define a random variable *X* in this way:

$$X = \begin{cases} X_1 & \text{with prob.} \quad p_2; \\ X_1 + X_2 & \text{with prob.} \quad p_1, \end{cases} \tag{1}$$

where X_1 and X_2 are independent random variables having exponential distribution with parameters μ_1 and $p_2\mu_2$ respectively; $p_1 + p_2 = 1$. It is easy to verify that the density function of *X* is given by

$$f(x) = \mu_1 e^{-\mu_1 x} + \frac{p_1\mu_1}{p_1\mu_2 - \mu_1}\left(\mu_1 e^{-\mu_1 x} - p_2\mu_2 e^{-p_2\mu_2 x}\right) \tag{2}$$

Moment matching is a common method for approximating distributions, especially in the area of queuing approximations. Though two-moment queuing approximations are common, they may lead to serious error when the coefficient of variation, ν (the standard deviation divided by the mean), is high [1, 5]. The first three moments of any no degenerate distribution with support on [0,∞) can be matched by the distribution (2).

To obtain the values of the parameters μ_1, μ_2, p_1 and p_2 of approximation, a complex system of non-linear equations needs to be solved:

$$\begin{cases} \dfrac{1!\cdot p_2\mu_1}{\mu_2 p_2-\mu_1}\left(\dfrac{\mu_2-\mu_1}{\mu_1^2}-\dfrac{\mu_2 p_1}{\mu_2^2 p_2^2}\right)=m_1; \\ \dfrac{2!\cdot p_2\mu_1}{\mu_2 p_2-\mu_1}\left(\dfrac{\mu_2-\mu_1}{\mu_1^3}-\dfrac{\mu_2 p_1}{\mu_2^3 p_2^3}\right)=m_2; \\ \dfrac{3!\cdot p_2\mu_1}{\mu_2 p_2-\mu_1}\left(\dfrac{\mu_2-\mu_1}{\mu_1^4}-\dfrac{\mu_2 p_1}{\mu_2^4 p_2^4}\right)=m_3; \\ p_1+p_2=1. \end{cases} \qquad (3)$$

The solution of the system is the following [6]:

$$\mu_2=\frac{g_2-g_1^2}{g_1^3-2g_1g_2+g_3},\ g_k=\frac{m_k}{k!},\ k=\overline{1,3};$$

$$\mu_1=\frac{1+\mu_2 g_1\pm\sqrt{(1-\mu_2 g_1)^2+4\mu_2{}^2\left(g_2-g_1^2\right)}}{2g_1-2\mu_2\left(g_2-g_1^2\right)};$$

$$p_1=\frac{\mu_2(\mu_1 g_1-1)}{\mu_2(\mu_1 g_1-1)+\mu_1};$$

$$p_2=\frac{\mu_1}{\mu_2(\mu_1 g_1-1)+\mu_1}. \qquad (4)$$

3. NUMERICAL MODEL OF QUEUING SYSTEM WITH SIMPLE PRIORITY

First of all the created software is tested with a simple system that has analytical formulas to calculate system measures.

Suppose that there are two classes of customers in a queuing system. Their service times follow a lognormal distribution. The arrival process is Poisson with parameters λ_1 ir λ_2 respectively. We shall suppose that class 1 has higher simple priority than class 2. This queuing model has l_1 and l_2 waiting positions for each class of customers to await service respectively. Let us calculate the mean number of customers in the queue and the mean waiting time of a customer in each class.

Let us assume that service-time distribution is approximated by expression (2). The scheme of considering a system is represented in Fig. 1.

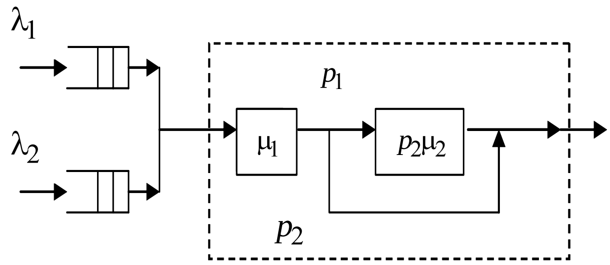

Fig. 1. *Queuing system with simple priority*

A new customer can not be accepted for the servicing while a previous one has not passed throughout all the phases of service.

A Markov chain with the countable space of states and continuous time can describe the functioning of such a system. To construct a numerical model of the system the approach proposed in [7] is applied.

The set of events in the system:

$$E=\{e_1, e_2, e_3, e_4, e_5\},$$

where

e_1 –a customer arrived from class 1;

e_2 –a customer arrived from class 2;

e_3 – completed service in the first phase with probability p_2;

e_4 –completed service in the first phase with probability p_1;

e_5 –completed service in the second phase.

The set of all feasible states of the system is:

$$N=\{(n_1, n_2, n_3, n_4)\},\ n_1=\overline{0,l_1};\ n_2=\overline{0,l_2}$$

where

n_1 – number of customers from class 1 present in the system;

n_2 – number of customers from class 2 present in the system;

$$n_3=\begin{cases} 0, \text{if the system is empty}; \\ 1, \text{ if a customer from class 1 is being served}; \\ 2, \text{ if a customer from class 2 is being served}; \end{cases}$$

$$n_4=\begin{cases} 0, \text{if the system is empty}; \\ 1, \text{if a customer is being served in the first phase}; \\ 2, \text{if a customer is being served in the second phase}. \end{cases}$$

The mean number of customers $L^{(1)}$ and $L^{(2)}$ in the queue and the mean waiting time $W^{(1)}$ and $W^{(2)}$ of a customer in each class are given by the following formulas

$$L^{(j)} = \sum_{n_j=1}^{l_j} \sum_{n_1,n_2,n_3,n_4} n_1 \pi(n_1, n_2, n_3, n_4)$$

$$W^{(j)} = \frac{L^{(j)}}{\lambda_j}, \quad j = 1,2 \,,$$

where $\pi(n_1, n_2, n_3, n_4)$ is the steady state probability of the system state. As an example, describe the event e_3 in the event language.

e_3 : **IF** $n_4 = 1$
  if $n_3 = 1$ **then** $n_1 \leftarrow 0$
    else $n_2 \leftarrow 0$
  end if
  if $n_1 > 0$ **then** $n_3 \leftarrow 1$
    else if $n_2 > 0$
    then $n_3 \leftarrow 2$
    else
    $n_3 \leftarrow 0 \quad n_4 \leftarrow 0$
    end if
  end if
 end IF
Return Intens $\leftarrow \mu_1 p_2$
END e_3

The created software, using the description of events, generates the set of feasible states of the system, the matrix of transition rates between them and the stationary probabilities of the states. Applying the obtained probabilities, it is possible to compute the desired characteristics of the system performance.

3.1 Results

If the number of waiting positions for service in each class of customers is unlimited, e.g. $l_1 = \infty$ and $l_2 = \infty$, then values $W^{(i)}$ and $L^{(i)}$ can be calculated by the analytical formulas :

$$L_q^{(1)} = \lambda_1 \cdot W_q^{(1)} = \lambda_1 \cdot \frac{(\lambda_1 + \lambda_2) E^2(X)(1 + \nu_X^2)}{2(1 - \lambda_1 E(X))};$$

$$L_q^{(2)} = \lambda_2 \cdot W_q^{(2)} =$$

$$\lambda_2 \cdot \frac{(\lambda_1 + \lambda_2) E^2(X)(1 + \nu_X^2)}{2(1 - \lambda_1 E(X))(1 - (\lambda_1 + \lambda_2) E(X))};$$

$$\nu_X = \sigma(X)/E(X),$$

where $E(X)$ and $\sigma(X)$ are the mean and the standard deviation of the service time.

Suppose that service-time is distributed according to the lognormal distribution with probability density

$$g(x) = \frac{1}{\alpha x \sqrt{2\pi}} \exp\left[-(\ln x - \lambda)^2 / 2\alpha^2\right], \quad x > 0$$

with parameters $\alpha = 0.9,\ \lambda = -0.05$. The first three non-central moments of the distribution are the following: $m_1 = 1.42618$, $m_2 = 4.57225$ and $m_3 = 28.3606$.

The results of the analytical model with parameters

$$\lambda_1 = 0.2, \lambda_2 = 0.09, \quad l_1 = \infty, l_2 = \infty$$

are the following:

$$L_q^{(1)} = 0.1855\,,\ L_q^{(2)} = 0.1423\,.$$

The results of the numerical model with the following values of parameters:

$$\lambda_1 = 0.2, \lambda_2 = 0.09,$$

$$\mu_1 = 0.74437, \mu_2 = 0.227796$$

$$p_1 = 0.018503, l_1 = 25, l_2 = 25$$

are:

$$L_q^{(1)} = 0.1855\,, L_q^{(2)} = 0.1424\cdot$$

As it is seen from the results, software calibration is successful and we can move forward with analyzing queuing systems that do not have analytical formulas to calculate various system measures

4. NUMERICAL MODEL OF QUEUING SYSTEM WITH QUALITY CONTROL

Suppose that there is one flow of customers and two queuing systems. Their service time follow a lognormal distribution. The arrival time process is Poisson with parameter λ. This queuing model has l_1 and l_2 waiting positions before each queuing system to await service. This system holds quality control that redirects with probability *p* already processed customers to go through both queuing systems again. Let us calculate the mean number of customers in both queues and the mean waiting time of a customer in each queue.

A new customer can not be accepted for the servicing while a previous one has not passed throughout both phases of queuing system.

The set of events in the system:

$$E = \{e_1, e_2, e_3, e_4, e_5, e_6, e_7, e_8, e_9\},$$

where

e_1 – a customer arrived to the first queuing system;

e_2 – completed service of the first queuing system in the first phase with probability p_1;

e_3 – completed service of the first queuing system in the first phase with probability p_2;

e_4 – completed service of the first queuing system in the second phase;

e_5 – completed service of the second queuing system in the first phase with probability p_1;

e_6 – completed service of the second queuing system in the first phase with probability p_2, and a customer has passed quality control;

e_7 – completed service of the second queuing system in the first phase with probability p_2, and a customer has failed quality control;

e_8 – completed service of the second queuing system in the second phase, and a customer has passed quality control;

e_9 – completed service of the second queuing system in the second phase, and a customer has failed quality control;The set of all feasible states of the system is:

$$N = \{(n_1, n_2, n_3, n_4, n_5, n_6)\}, n_1 = \overline{0, l_1}; n_4 = \overline{0, l_2},$$

where

n_1 – number of customers in the first queue;

$$n_2 = \begin{cases} 0, & \text{if the first phase of the first queuing system is empty;} \\ 1, & \text{if a customer is being served in the first phase of the first queuing system.} \end{cases}$$

$$n_3 = \begin{cases} 0, & \text{if the second phase of the first queuing system is empty;} \\ 1, & \text{if a customer is being served in the second phase of the first queuing system.} \end{cases}$$

n_4 – number of customers in a second queue;

$$n_5 = \begin{cases} 0, & \text{if the first phase of the second queuing system is empty;} \\ 1, & \text{if a customer is being served in the first phase of the second queuing system.} \end{cases}$$

$$n_6 = \begin{cases} 0, & \text{if the second phase of the second queuing system is empty;} \\ 1, & \text{if a customer is being served in the second phase of the second queuing system.} \end{cases}$$

Let us assume that service-time distribution is approximated by expression (2). The scheme of examining system is represented in Fig. 2.

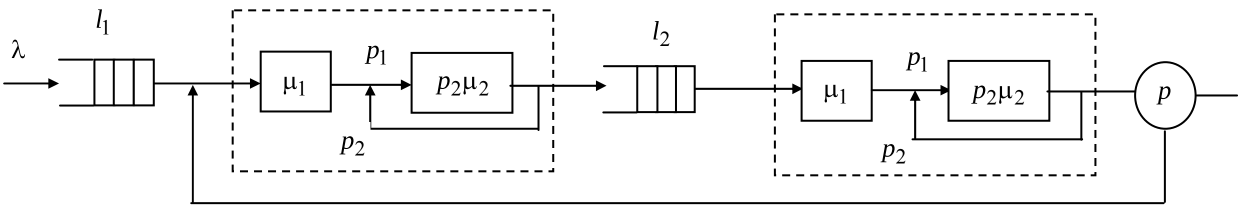

Fig.2. *Queuing system with quality control*

The mean number of customers $L^{(1)}$ and the mean waiting time $W^{(1)}$ of a customer in the first queue are given by the following formulas

$$L^{(j)} = \sum_{m_1=1}^{l_1} \sum_{n_1,n_2,n_3,n_4,n_5,n_6} m_1 \pi(n_1, n_2, n_3, n_4, n_5, n_6)$$

$$W^{(j)} = L^{(j)} / \lambda_1, \quad j = 1,2,$$

where $\pi(n_1, n_2, n_3, n_4, n_5, n_6)$ is the steady state probability of the system state.

As an example, describe the event e_3 in the event language.

```
e3 :   IF   n3 > 0
          if  n5 < 1  and  n6 < 1
               then n5 ← n5 + 1
               else n4 ← n4 + 1
          end if
          if   n1 > 0     then
     n1 ← n1 − 1;  n2 ← n2 + 1;  n3 ← n3 − 1;
               else n3 ← n3 − 1
          end if
       End IF
Return  Intens ← μ2 p2
END e3
```

The created software in C++, using the description of events, generates the set of feasible states of the system, the matrix of transition rates between them and the stationary probabilities of the states. Applying the obtained probabilities, it is possible to compute the desired measures of the system performance.

4.1 Results

The results of the numerical model with the following values of parameters:

$$\lambda = 1, \quad \mu_1 = 0.74437, \quad \mu_2 = 0.227796$$

$$p_1 = 0.018503, \quad p = 0.9, \quad l_1 + l_2 = 20$$

are :

$$L_q^{(1)} = 8.9849, \; L_q^{(2)} = 7.6322$$

$$W_q^{(1)} = 12.0706, \; W_q^{(1)} = 33.4796.$$

5. Conclusions

A method and created software for automatic construction of numerical models for systems described by Markov chains together with queuing approximation allows:

- to analyse complex non-markovian queuing systems applying Markov chains theory,
- queuing systems with the infinite (countable) space of states can be approximated by the finite space of states.

5. REFERENCES

[1] W. Whitt, "On approximations for queues, III: Mixtures of exponential distributions", **AT&T Bell Labs Tech Journal,** Vol. 63, No. 1, 1984, pp. 163-175.

[2] M.A. Johnson, "An empirical study of queuing approximations based on phase-type distributions", **Commun. Statist.-Stochastic Models**, Vol. 9, No. 4, 1993, pp. 531-561.

[3] J.A. Luhman, M.J. Johnson, "Behaviour of queuing approximations based on sample moments", **Applied Stochastic Models and Data analysis**, Vol. 10, 1994, pp. 233-246.

[4] M. J. Faddy, "Examples of fitting structured phase-type distribution", **Applied Stochastic Models and Data Analysis**, Vol. 10, 1994, pp. 247-255.

[5] M. A. Jonhson, M. R. Taaffe, "An investigation of phase distribution moment matching algorithms for use in queuing models", **Queuing systems**, Vol. 8, No. 2, 1991, pp. 129-147.

[6] E. Valakevičius, "On Approximation of Non-Markovian Queuing Models", In: **The Second International conference "Simulation, gaming, training and business process reengineering in operations"**, Riga, Latvia, 2000, pp. 56-59.