Simulation Modelling of Overflow Probabilities in Tandem Queuing System

D. Radev, E. Rashkova and D. Stankovski

***Abstract* – An approach for simulation modelling of overflow probability based on the rare event simulation and limited relative error is suggested. The developed algorithm for tandem queuing system help to speed-up the simulation for estimation of overflow probability, which in tandem queuing system is a rare event. Numerical and simulation results are shown.**

***Keywords* –Simulation Modelling, Markov Chains, Queuing Theory, Overflow Probability, Rare Event.**

I. Introduction

Queuing systems are convenient reference models for research and analysis of communication systems performance, where very important is to evaluate the overflow probability. In real communication systems it is impossible to receive the values of these probabilities with direct methods, that's why is necessary to involve simulation modeling [1].

The model of tandem queuing system is under study in order to be able to compare received simulation measures with known analytical or numerical results. For example, in tandem queuing system, the probability that the content of the second buffer exceeds some high level L before it becomes empty starting from a given state is a rare event, because it probability most often is less than 10^{-6}.

The estimation of rare event probability using Monte Carlo simulation requires a very long computer processing time, and cannot be implemented [4]. Lately two basic methods of the rare event simulation were developed: Importance splitting of the sample path after reach definite intermediate level between the starting level and the rare event [6]; and the Importance Sampling (IS) generation, where the probability density function is used as the rare event evaluation measure, which can be compared and changed based on the likelihood ratio of the less rare event probability density function [2].

RESTART (Repetitive Simulation Trials After Reaching Thresholds) approach is an accelerated simulation method, which belongs to the so-called importance splitting methods, used to the speed-up the rare event simulation [5,7].

The Limited Relative Error (LRE) measures the complementary distribution function of the queue occupancy and performs the Run Time Control (RTC) of the simulation. The limited relative error performs the run time control with two conditions: first the Large Sample Conditions and second the Relative Error Condition. The first condition assures that the queuing system has reached the steady state. The second condition represents a measure to estimate the relative error at the current state of the simulation.

For simulation modeling of overflow probabilities in tandem queuing system in this paper consider RESTART approach combined with limited relative error.

[1]D. Radev, Department of Communication Technique and Technologies, University of Rousse, Phone: +359 82 888 841, E-mail: dradev@abv.bg

[2] E.P.Rashkova, Department of Communication Technique and Technologies, University of Rousse; E-mail: elena_pl@gbg.bg

[3] D. Stankovski, Department of Communication Technique and Technologies, University of Rousse; E-mail: draganstankovski@yahoo.com

II. Overflow probability of tendem queues

Many models used for evaluation of queuing systems performance can be presented as discrete-time Markov chains (DTMCs) [3]. This is the case if all inter-arrival times and service times have an exponential distribution, and the quantity of interest is the overflow probability. Note that some other performance measures, like delays, cannot be obtained from the discrete-time Markov chain description. A DTMC for a Markov queuing model has a highly regular structure. First of all, the states typically can be arranged conveniently on a grid with as many dimensions as the number of queues, where each of coordinates represents the number of customers in one of the queues. Secondly, every transition in the DTMC corresponds to an elementary event in the queuing model: an arrival or a service completion at one of the queues. These events are known as "transition events" and they are defined independently of the state; i.e., there is only one transition event for a service completion at a given queue, and this single transition event corresponds to a transition out of every state in the DTMC in which this particular queue is non-empty. From the other side not all transition events are "enabled" in every state: e.g., in a state where a particular queue is empty, the service completion event of that particular queue is not possible, i.e., not enabled.

Consider the overflow probability of the total population in a queuing network consisting of two queues in tandem presented on Figure.1, where customers arrive at the first queue according to a Poisson-process with arrival rate λ. Both servers have exponentially-distributed service times with rates μ_1 and μ_2.

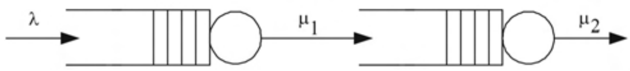

Figure 1. Two-queue tandem network.

The state of the system at any time is given by the two integer values n_1 and n_2, which are the number of customers in the first and second queues, respectively.

The difficulty of applying standard simulation techniques arises when the first queue is the bottleneck and the rare set definition is related to the value of n_2. In this case it would be better to apply one of splitting methods for rare event simulation with careful choice of the importance function [3]. On Figure 2 is presented state space of two queues in tandem network.

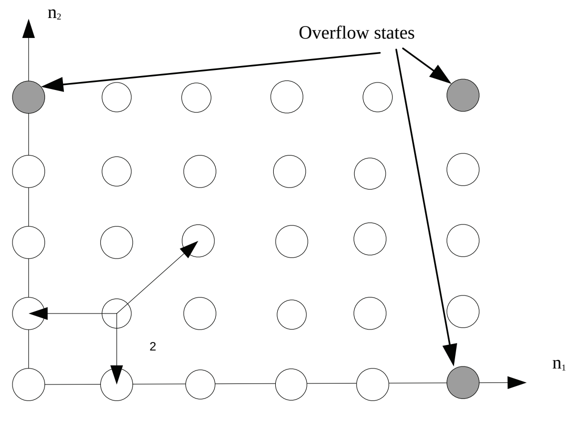

Figure 2. State space of two queues in tandem network

For the considered two-queue tandem network the rare event set A include the following three cases, where L is the level of overflow of the queues (1).

$$\begin{aligned} & n_1 + n_2 \geq L, \\ & n_2 \geq L, \\ & \mathrm{Min}(n_1, n_2) \geq L \end{aligned} \tag{1}$$

There exist three possible cases for rare event:

A. Rare event defined as $n_1 + n_2 \geq L$

For this definition of the rare event set, the possible states for rare event are $(0, n_1+n_2)$ or $(0,(n_1+n_2-1))$. The importance of these states is different. The higher value of λ_1 (for n_1+n_2), given that customer at n_1 has to be served by both servers before leaving the system, while μ_1 a customer at n_2 has to be served only at the second one. Here the bottleneck is the first queue, so the states with high value of n_1 and low value of n_2 have the highest probability.

μ

B. Rare event defined as $n_1 \geq L$

For this definition of the rare event set $n_1 \geq L$, to obtain states with closer importance, it seems that the weight given to n_2, have to be given to n_1.

C. Rare event defined as $Min(n_1,n_2) \geq L$

For $n_1 \leq L$ and $n_2 \leq L$, the rare event set is in the queue with fewer customers.

III. The LRE approach

The limited relative error approach tries to construct independent samples from a given time series by building batches that tend to have lesser autocorrelation than the original series. It leads to reduction of autocorrelation and results in confidence intervals for the mean of the simulation results. The limited relative error approach is implemented with counters to measure the quantities: n - number of trials; h_i - state frequency; l_i - number of times the queue was in state i when a cell arrived; a_i - transition frequency for decreasing number of transitions from j to i between two arrivals, with $j>I$. From these counters the quantities l_i and d_i are (2) and (3).

$$l_i = n - d_i \tag{2}$$

$$d_i = \sum_{j=1}^{k} h_j, with \quad d_0 = n \tag{3}$$

If the large sample conditions and relative error condition has to be fulfilled and the results of it are summarized in the end of the algorithm. The large sample conditions are presented as (4).

$$\begin{aligned} & n \geq 10^3, \\ & (l_i, d_i) \geq 10^2, \\ & (a_i, d_i - c_i, l_i - a_i) \geq 1 \end{aligned} \tag{4}$$

The relative error condition is presented as (5).

$$\mathrm{RE}_{\max i} = \frac{1 - d_i / n}{d_i} \cdot \mathrm{Cor}_i \tag{5}$$

The LRE approach starts with index i=1, tests the conditions (4) and (5), and if the conditions are fulfilled the LRE advances to the next index i. The simulation is terminated for the last index $i=L_m$, where L_m is the overflow conditional probability.

IV. Algorithm with RESTART

The RESTART method is a simple simulation method for the estimation of small probabilities. Applying the RESTART means to split the possible range of values of λ into regions with different importance. The given set of thresholds is L_i, $i=0,1,2,...,m$. The system reaches the state L_0, while evaluating the first interval $[0, L_0]$ the system states are saved. Once the values for the first interval are established at the time t_0, the simulation restarts from one of the previously saved states at threshold L_0. Each time when the threshold L_0 is crossed a restart is performed with one of the saved states, which is chosen randomly and the system is reload with this state.

The RESTART method is used to the speed-up the rare event simulation for estimation of overflow probability defined as (6).

$$\begin{aligned} & G_l = P\{\lambda \geq L\} = \\ & P\{\lambda \geq L | \lambda \geq L_{m-1}\} P\{\lambda \geq L_{m-1} | \lambda \geq L_{m-2}\} \\ & \dots P\{\lambda \geq L_1 | \lambda \geq L_0\} P\{\lambda \geq L_0\} \end{aligned} \tag{6}$$

The relative error $RE_{\max_i}$ is a function of the overall relative error RE_{max} and the number of thresholds m is defined as (7).

$$RE_{\max_i} = (1 + RE_{\max})^{\frac{1}{m+1}} - 1 \tag{7}$$

The proper choice of the thresholds L_i leads to a maximum speed-up according to (8).

$$\frac{G_i}{G_{i-1}} = e^{-2} \tag{8}$$

Since the complementary cumulative density function is monotonically decreasing, and the thresholds are properly chosen, the RESTART method will speed-up the simulation.

The following algorithm with RESTART/LRE for tandem queuing systems was developed:

Algorithm RESTART/LRE

Step 1 Initialization

Specification of the rare event L, number of levels m, the values of the thresholds $L_0,\dots,L_{m-1}$, and the maximal relative error $RE_{\max}$;

Definition of generated random variables β and the upper and lower boundary for generated variables $\beta_{\max}$ and $\beta_{\min}$;

Definition of parameters for arrival and service process;

Calculation maximal relative error $RE_{\max i}$ from $RE_{\max}$;

Generate model objects (traffic sources, queues, network nodes);

Initialization of counters, parameters of the system and the matrix of the state space (Fig. 2)

Definition of the rare event set and kind of the model:

Case 1: Rare event set defined as $n_1 + n_2 \geq L$

Case 2: Rare event set defined as $n_2 \geq L$

Case 3: Rare event set defined as $\mathrm{Min}(n_1 + n_2) \geq L$

Initialization of number of runs n:=0 and the simulation time s:=1;

Set the scaling factor U:=1;

Start simulation;

Step 2 Outer loop

If ($n \geq 10^3$ AND $(l_i, d_i) \geq 10^2$ AND $(a_i, l_i - a_i, d_i - a) \geq 10$)

Step 3 For i=0

While not (error< $RE_{\max}$ for L_0-1 to L_0)

Simulation continues with generating a new value of random variables β with Poisson distribution.

Estimation of generated random variable β: whether β is in the left side of L_0-1 or in the right side of L_0; Result of estimation: the new value of β is accepted if it is in the right side and the state is stored for RESTART;

Evaluation Counter=Counter+1,

End;

Step 4 Inner loop

For i=1 to m

Set the thresholds $L_0,\dots,L_m$ for the rare event L;

While not (error< $RE_{\max}$ for L_{i-1} to L_I)

Restore one random state from L_{i-1} ;

Calculation Relative Error for Complementary Cumulative Distribution Function for $G_{L_{i-1}|\beta \geq L_{i-1}}$ to $G_{L|\beta \geq L_{i-1}}$. If the $RE > RE_{\max}$ the simulation stop and the new random state is restored;

Simulation continues with generating data: new value of random variables β with Poisson distribution end estimation whether β is in the left side of L_i-1 or in the right side as L_i. The new value of β is accepted if it is in the right side and the state is stored for RESTART.

Evaluation Counter=Counter+1;

n:=n+1

Set scaling factor U:= $G_{L|\beta \geq L_{i-1}}$;

Multiplication of G-values with U for $L_{i-1},\dots, L_i$;

Continue simulation till the last value of n;

End;

Step 5 Output results

Calculation of final results:

$$\hat{G}(x) = \hat{G}_i = d_i / n$$

$$\hat{\beta} = \frac{1}{n} \sum_{i=1}^{k} d_i$$

$$\hat{\rho}(x) = \hat{\rho}_i = 1 - \frac{c_i / d_i}{1 - d_i / n}$$

$$\mathrm{Cor}[x] = \mathrm{Cor}_i = (1 + \hat{\rho}_i)(1 - \hat{\rho}_i)$$

$$\mathrm{RE}[x]^2 = \mathrm{RE}_i = \frac{1 - d_i / n}{d_i} \cdot \mathrm{Cor}_i$$

END.

V. Numerical Example

The simulation results for RESTART/LRE algorithm are received with simulation program on Visual C++ with implemented of Ptolemy as a powerful, object-oriented simulator and Microsoft Visual Studio 3.0. The simulator is available with source code for operating systems Windows 9x, NT, XP. All the simulation examples were executed on Personal computer with properties: Pentium, 1,4 GHz, 312Mb RAM, operating system Windows XP Professional. One example of a part of developed simulation software is shown on Figure 3.

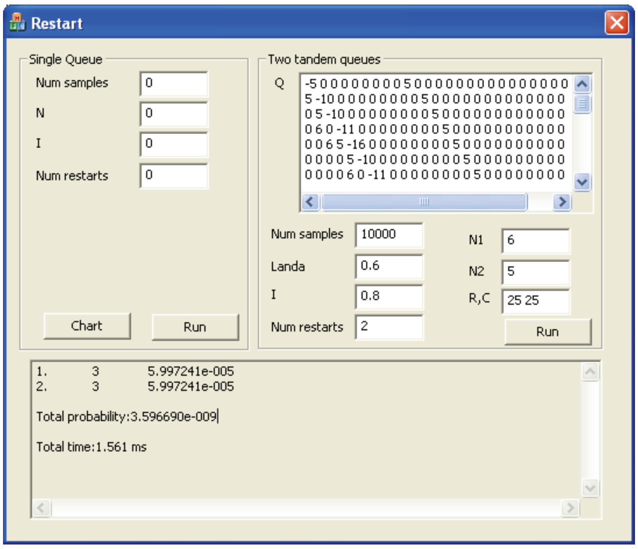

Figure 3. Program realization

Customers with Poisson arrival distribution enter the first queue with buffer size n_1=6 for service completion and enter the second one with buffer size n_2=5 for service completion as is shown on Figure 2. The mean arrival rate at the first queue is λ and the service time is exponentially distributed in each queue with mean service rates μ_1 and μ_2, respectively. The load at each queue is $\rho_1 = \lambda/\mu_1$ and $\rho_2 = \lambda/\mu_2$. The generated random variables with Poisson arrival distribution are shown on Figure 4.

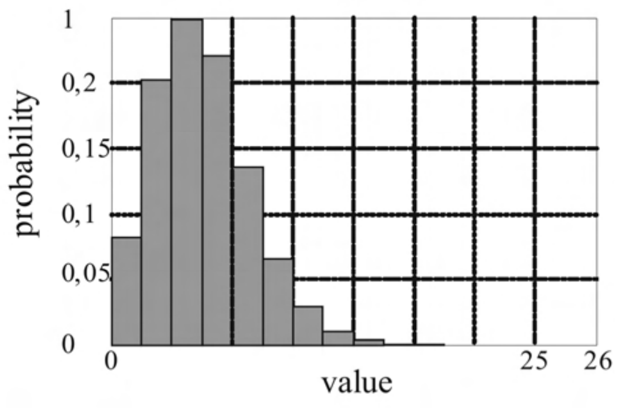

Figure 4. Generation of variables with Poisson distribution

The buffer size at each queue is assumed to be finite. Consider example with arrival rates λ=0,6 and service completion with rates μ_1=0,7 and μ_2=0,8. The number of simulation samples is n=10 000.

In all the runs, the simulation length was adjusted to have a relative half width of the 95% confidence interval (relative error) equal to 10%. The interval width was evaluated using the batch means method [3].

The investigated states defined as rare event L are in the end of the first queue (5,0), the minimum (5,4), and in the end of the second queue (0,4). The importance of each of these states is different. The thresholds were chosen as L=0,8 and L=0,9. The number of restarts in the numerical example is 2. The number of samples is n=10 000.

In Table 1 are shown the received simulation results for L=0,8; relative error RE=5%, λ=0,3, μ_1=0,4 and μ_2=0,9. For comparison are given theoretical results and simulation time.

Table 1 Results for tandem queue network with n=10 000 simulation samples,

Definition of Rare state	G_L theoretical	G_L	Time, *ms*
First case: (5,0)	0,0318598651	9,292000e-006	2,549
Second case: (0,4)	0,0020832578	6,912656e-006	2,346
Third case: (5,4)	0,0000417815	7,849558e-006	2,836

In Table 2 are shown the received simulation results for L=0,9; relative error RE=10%, λ=0,4, μ_1=0,7 and μ_2=0,9. For comparison are given theoretical results and simulation time.

Table 2 Results for tandem queues network with n=10 000 simulation samples,

Definition of Rare state	G_L theoretical	G_L	Time, *ms*
First case: (5,0)	0,0134789964	9,582591e-006	2,836
Second case: (0,4)	0,0078628157	5,983916e-006	2,513
Third case: (5,4)	0,0003996907	3,129537e-006	3,349

VI. Conclusion

The simulation modeling of overflow probability based on the rare event simulation and limited relative error was under consideration. The developed algorithm for tandem queuing system help to the speed-up the simulation for estimation of overflow probability, which in tandem queuing system is a rare event. The received simulation results for the basic reference model, two server queuing system – FIFO with a finite buffer size N, shows that suggested approach and algorithm are suitable for rare event simulation with large number of customers in complex communication networks.

Acknowledgments

This work is a part of the research project ВУ-ТН-105/2005 of Bulgarian Science Fund at Ministry of Education and Science.

References

[1] Boer. P. Analysis and Efficient Simulation of Queuing models of Telecommunication Systems. PhD thesis, Univ. of Twente, 2000.
[2] Georg, C., Schreiber, F. The RESTART/LRE method for rare event simulation. Winter Simulation Conference, Coronado, CA, USA, 1996, 390-397.
[3] Radev, D. Modeling of Rare events in broadband digital networks, Sofia, 2006.
[4] Radev, D., Rashkova, E. Algorithms for traffic flows modelling with Monte Carlo simulation TELEKOM'2006, Varna, 2006.
[5] Radev, D., Iliev, M., Arabadjieva, I. RESTART Simulation in ATM networks with tandem queue. - Proceedings of the International Conference Automatics and Informatics'2004, Sofia, Bulgaria, 37-40.
[6] Radev, D. Splitting method in rare event simulation for high - speed communication networks. - Proceedings of the International Conference Automatics and Informatics'2003, Vol.1, Sofia, 103-107, 2003.
[7] Villen Altamirano, M. Villen Altamirano, J. RESTART: a method accelerating rare event simulations, in 13th International Teletraffic Congress, Copenhagen, North- Holland, 1991, pp. 71-76.