Presentation

Regeneration Cycles Dependence in Confidence Estimation by Splitting

Alexandra V. Borodina

(Karelian Reseach Centre of Russian Academy of Sciences, Russia)

Introduction

A confidence estimation of rare events probabilities in $M/G/1$ system is considered. We apply special speed-up simulation technique based on a combination of splitting and regenerative approach for the queue-size process. Besides, we consider an embedded Markov process instead of the original non-Markovain one.

A comparison with the known analytical results and with the Crude Monte-Carlo (CMC) simulation shows a high efficiency of the approach. In particular, a high precision is achieved much faster than by CMC that is crucial for the rare event simulation.

Problem

Consider the queue-size process $\{\nu(t),\, t \geq 0\}$ with a state space E in a $M/G/1$ queue with a Poisson input with the independent and identically distributed (i.i.d.) interarrival times $\{\tau_n\}$ with rate λ, the i.i.d. service times $\{S_n\}$ with distribution $B(x)$ and with a finite mean ES_1.

A problem of rare event simulation is well known. We would like to estimate a *small stationary* probability

$$\gamma := P(\nu(t) \in A), \tag{1}$$

where $A \subset E$, $\gamma \approx 0$. So, following [11, 14], the event $\{\nu(t) \in A\}$ is called *rare event*, the set A is called *rare set*. Standard simulation (CMC) of γ requires enormous sample size to achieve a given accuracy, thus, the time of simulation is unacceptably huge. We seek to avoid this situation by using the technique of speed-up method Splitting [5, 7, 7].

In many cases, the estimates turn out to be unbiased [5, 7]. At the same time, it is equally important to prove also consistency and the applicability of a Central Limit Theorem (CLT) for confidence estimation.

In recent papers [2, 3], a modification of the splitting has been proposed which uses a correspondence between trajectories of a queueing Markov process obtained by the splitting and a regenerative structure. In turn, it justifies consistency of the estimate and allows us to apply regenerative simulation for confidence estimation of the (unknown) probability γ with a given precision [4, 9, 8, 15].

Regenerative interpretation

A standard spitting algorithm is typically used to estimate the (stationary) probability γ that the process hits a (rare) set A before reaching the state $\{0\}$, [5, 6, 7].

The state space E divided into $M+1$ subsets, which defined by *importance function* f:

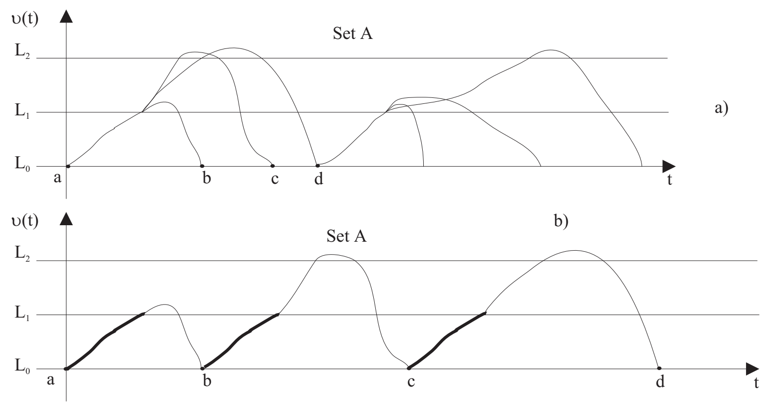
$$C_i = \{x \in E \mid f(x) \geq L_i\},\ i \in [1, M+1],$$

$A = C_{M+1}$. The trajectory of the process splits at the predefined *thresholds* L_i only. The number of splits denoted by R_i, $R_{M+1} = 1$. The trajectory of the process is terminated upon reaching A (after hitting the threshold L_{M+1}).

For the standard Splitting, an optimal distance between thresholds and an optimal number of split paths at each threshold are known [5, 7]. It is necessary to note, that if the basic process is **Markovian** (for example, queue-size process in $M/M/1$ system) then the new trajectories are i.i.d (after splitting) and we can follow the [5, 7] for γ estimation.

But, in the case of **non-Markov process** (like continuous-time queue-size process in $M/G/1$ and $GI/G/1$ systems) the new trajectories (after splitting) are not longer independent, and it may be a source of an instability of the corresponding estimates [2].

We focus on regenerative interpretation of the method. As in a standard Splitting, we divide the state space E by $M+1$ subsets C_i. But, we stop each path when the process hits zero state, and treat the obtained path as a regeneration cycle. Each trajectory started from L_0 gives the *bunch* of $D := R_1 \cdots R_{M+1}$ dependable regeneration cycles. The cycles from different bunches are independent by construction. The total number of bunches is the random variable R_0.



Splitting for an embedded Markov chain

For the queue-size process in a non-Markov system $M/G/1$ we use the **embedded Markov chain method** [1]. We consider the process at the departure instants only (there is not a guarantee to split a trajectory exactly at each level L_i).

New splitting condition: *If a trajectory starting from a region $G(i) := C_i \backslash C_{i+1}$ (at a departure instant) crosses a threshold L_{i+k}, $i + k \leq M + 1$, then it splits onto $\prod_{j=1}^{k} R_{i+j}$ paths ($R_{M+1} = 1$).*

The point estimator γ_{R_0} for γ is

$$\gamma_{R_0} = \frac{N_{M+1}}{R_0 \cdot D},$$

where N_{M+1} is the number of rare events.

Each trajectory now has a unique set of the thresholds to be split. Thus, there is a problem to obtain the optimal conditions for L_i, R_i as it was in standard Splitting.

Experimental results show the advantage of new algorithm expressed in variance reduction of point estimator for γ.

Confidence estimation

By using regenerative structure (given by splitting) and considering the dependences between the cycles in the bunch, we can now apply a regenerative simulation method [4, 8, 9, 10, 13, 15] for confidence estimation γ with the given precision.

Denote

$$Y_i := \sum_{j=(i-1)\cdot D+1}^{i \cdot D} I^{(j)}, \; i = 1, \ldots, R_0,$$

where $I^{(j)} = 1$ if the length of queue exceeds the level L_{M+1} at the regenerative cycle number j. Variables Y_i are i.i.d. The sequence $\{I^{(j)}\}_{j=1}^{D \cdot R_0}$ is regenerative with identical periods of regeneration equal to D.

From Regenerative Central Limit Theorem (RCLT) [15] under condition $EY_1^2 < \infty$ follows the expression of confidence interval for γ:

$$\left[\gamma_{R_0} \pm \frac{z_\delta \sqrt{\sum_{i=1}^{R_0} (Y_i - D \cdot \gamma_{R_0})^2}}{D \sqrt{R_0 (R_0 - 1)}} \right], \tag{2}$$

where quantile z_δ satisfies $P(N(0,1) \leq z_\delta) = 1 - \delta/2$. The point estimator γ_{R_0} defined by splitting procedure.

Results

Let the estimation precision be $x\%$ of γ (it means that the half of the interval length, given by (2), $\Delta = x\%$ of γ). Denote the half of interval without cycles dependence consideration by $\Delta_1 := \dfrac{z_\delta\sqrt{\gamma-\gamma^2}}{\sqrt{n}}$, where n the total number of regeneration cycles ($n = R_0 \cdot D$). The error that reflects the cost of dependence can be expressed as $y := \dfrac{(\Delta-\Delta_1)10^2}{\gamma}$.

Table 1.

#	γ	D	n	x	y	time (sec.)
1	2,92E-07	1	5,48E+07	49	0	21006,5
2	2,80E-07	10	2,86E+08	66,6472	44,7	10104,8
3	4,87E-07	100	1,55E+09	68,8478	61,7	5657,25
4	2,52E-09	5,03E+08	7,20E+13	47,7	47,2	77,265
5	3,80E-09	1,16E+09	7,65E+13	34,0	33,7	69,281
6	2,14E-09	2,37E+09	2,21E+14	21,9	21,7	152,593
7	1,57E-09	5,90E+09	5,89E+14	9,6	9,4	309,781
8	5,46E-09	3,28E+09	1,65E+14	6,1	5,9	352,89
9	2,46E-09	5,90E+09	4,69E+14	5,6	5,5	1023,39
10	2,34E-08	4,44E+09	1,66E+13	3,7	3,4	2550

The numerical results have confirmed that the dependence between regeneration cycles in the bunches (caused by dependence of trajectories having a common pre-history) essentially influences on interval length.

It is easy to see from Table 1 that $x \approx y$ (excepting the case of small D). Thus, it is impossible to consider the regeneration cycles from one bunch as independent. Otherwise the error coincides with accuracy.

The case $D = 1$ corresponds the CMC (in this case all cycles are independent), which gives error $y = 0$. Nevertheless, the CMC method is strongly loses on time in comparison with offered speed-up approach for confidence estimation with the given accuracy.

References

[1] S. Asmussen. Applied Probability and Queues, 2nd ed., Springer, NY, 2003.

[2] A. V. Borodina, E. V. Morozov. Confidence estimation of buffer overfull probability by using speed-up regenerative simulation of $M/M/1$.// Proceedings IAMR Karelian Research Center RAS, vol. 7, 2006, 125-135. (in Russian)

[3] A. V. Borodina, E. V. Morozov. Speed-up regenerative simulation of a high load probability for a single-server queue. Journal of Applied and Industrial Mathematics (in Russian, accepted for publication in 2007).

[4] M. Crane, D. L. Iglehart. Simulating stable stochastic systems, III: Regenerative processes and discrete-event simulations. Operations Research 23: 33-45.

[5] M. Garvels. PhD Thesis. "The splitting method in rare event simulation", The University of Twente, The Netherlands May, 2000.

[6] P. Glasserman, P. Heidelberger, P. Shahabuddin, and T. Zajic. A look at multilevel splitting. In H. Niederreiter, editor, Monte Carlo and Quasi Monte Carlo Methods 1996, Lecture Notes in Statistics, vol. 127, 1996, 99-108, Springer Verlag.

[7] P. Glasserman, P. Heidelberger, P. Shahabuddin, and T. Zajic. Splitting for rare event simulation: analysis of simple cases. Proceedings of the 1996 Winter Simulation Conference,1996, 302-308.

[8] P. W. Glynn. Some topics in regenerative steady state simulation. Acta Applic. Math. 34, 1994, 225-236.

[9] P. W. Glynn, D. L. Iglehart. Conditions for the applicability of the regenerative method. Management Science 39, 1993, 1108-1111.

[10] P. W. Glynn, D. L. Iglehart. A joint central limit theorem for the sample mean and regenerative variance estimator. Annals of Operations Research 8, 1987, 41-55.

[11] P. E. Heegaard. A survey of Speedup simulation techniques. Workshop tutorial on Rare Event Simulation, Aachen, Germany, 1997.

[12] P. Heidelberger. Fast simulation of rare events in queuieng and relaibility models, Performance Evaluation of Computers and Communications Systems Springer-Verlag, LN in Computer Sci., v. 729, 1993, 165-202.

[13] E. Morozov and I. Aminova. Steady-state simulation of some weak regenerative networks, European Transactions on Telecommunications ETT, Vol. 13, No. 4, July/August, 2002, 409-418.

[14] P. Shahabuddin. Rare event simulation in stochastic models. Proceedings of the WSC 1995, IEEE Press., 1995, 178-185.

[15] K. Sigman, R. Wolff. A review of regenerative processes. SIAM Review, Vol. 35, No. 2, pp.269-288. 1993.