한국시뮬레이션학회 '99 춘계학술대회 논문집 1999.4.17 동의대학교

A New Fast Simulation Technique for Rare Event Simulation

Yun Bae KIM*, Deok Seon ROH*, Myeong Yong LEE**

* Sung Kyun Kwan University
School of Systems and Management Engineering
Suwon, Korea
82-331-290-7600(phone), 82-331-290-7610(fax)
kimyb@yurim.skku.ac.kr
** Telecommunications Network Laboratory
Korea Telecom R&D Group
Taejeon, Korea
yonglee@kt.co.kr

ABSTRACT

Importance Sampling (IS) has been applied to accelerate the occurrence of rare events. However, it has a drawback of effective biasing scheme to make the estimator from IS unbiased. Adaptive Importance Sampling (AIS) employs an estimated sampling distribution of IS to the system of interest during the course of simulation. We propose Nonparametric Adaptive Importance Sampling (NAIS) technique which is nonparametrically modified version of AIS and test it to estimate a probability of rare event in M/M/1 queueing model. Comparing with classical Monte Carlo simulation, the computational efficiency and variance reductions gained via NAIS are substantial. A possible extension of NAIS regarding with random number generation is also discussed.

1. Introduction

Simulation is a very useful tool for assessing the performance evaluation of ATM networks. The desired cell loss probability in ATM networks is in the range of 10^{-6} to 10^{-12}, according to service characteristics which make it computationally costly to use classical simulation techniques. This limitation has been reported in most simulation studies of ATM networks. The problems caused by this limitation can be classified into two ways. First, the degeneration of random number stream that might caused a repetition of random numbers. Second, excessive simulation time that may be resulted in inflating variance of estimator and makes hard to analyze the output from simulation. There are several fast simulation techniques to remedy this limitations; Importance Sampling, Parallel Simulation, Regenerative Method, and Hybrid Simulation. Smith[1997], Frost[1988], and Glynn and Iglehart[1989], provide thorough surveys on these fast simulation techniques.

To obtain a statistically valid estimate, one needs to have at least 10 independent replications if a direct method is used. Unfortunately, most simulation software or programming languages can not provide a big sequence of random numbers without degeneration (i.e., the same random numbers are repeated in a single simulation run).

IS is a powerful technique used in the areas of rare event simulation and the success of IS is reported in several papers. The fundamental idea of IS is to modify the probabilities of rare events that govern the outcomes of the simulation in a way that the original low-probability events occur more frequently. To estimate the probability of rare events, we simulate relatively high probability events with a biased sampling distribution. The sample values from a biased sampling distribution are then adjusted to make the final estimates unbiased. However, selecting an optimal sampling distribution that makes the events occur more frequently is not enough; how to make it happen is a very important issue. For example, an arbitrary selected sample distribution generates more events with an estimator of infinite variance. Hence the main step of IS is selecting an optimal sampling distribution (often system specific) which guarantees the variance reduction.

The burden of selecting an optimal sampling distribution can be eased by the system designer. The main idea of AIS is the recognition that the distribution of the samples of the error events is distributed as the optimal sampling distribution of IS. This distribution may be used to estimate the properties of optimal sampling distribution in iterative way in order to close the gaps between the optimal sampling distribution and an estimate of the optimal sampling distribution.

Most of the works in IS are focused on the selection of an estimate of sampling distribution of IS in a parametric way by Glynn and Iglehart [1989], Oh and Berger [1992, 1993], Siegmund [1976], and West [1992, 1993]. The nonparametric way studied by Givens and Raftery [1996] can provide a prominent improvement in selection of an optimal sampling distribution.

Zhang [1996] proposes a nonparametric method to estimate a sampling distribution of IS for a given system, which uses the estimated sampling distribution to generate random numbers rather than estimating the parameters of optimal sampling distribution. He extends Nonparametric Importance Sampling (NIS) to Nonparametric Adaptive Importance Sampling (NAIS), which is just iterations of NIS requiring more computation. Based on AIS, our NAIS uses the initial sampling distribution conditioned on the samples of rare enents occurred during the initial simulation run and uses Zhang's nonparametric idea to estimate the optimal sampling distribution.

The rest of this paper is organized as follows. In Section 2, we introduce the basic idea of IS method and AIS method. Section 3 is devoted to NAIS method. In Section 4, we test NAIS in an M/M/1 queueing model. Conclusions and future research areas are discussed in Section 5.

2. Importance Sampling and Adaptive Important Sampling

2.1 Importance Sampling (IS)

Let a random variable (r. v.) X be defined on the probability space (Ω, Γ, P), where Ω, Γ, and P are sample space, event space, and probability measure, respectively. The occurrence of rare event E is defined as $E \in \Gamma$. The indicator function $\phi(x)$ can be defined as follows:

$$\phi(x) = \begin{cases} 1, & \text{if } X \in E \\ 0, & \text{otherwise.} \end{cases}$$

Consider a problem of estimating the probability of rare event E :

$$E_P[\phi(x)] = \mu_{\phi(x)}, \tag{1}$$

where P is a measure with respect to the expectation is taken.

In a classical simulation, (1) can be estimated with N independent samples as follows :

$$\mu_{\phi(x)} = \frac{1}{N}\sum_{i=1}^{N}\phi(x_i).$$

By the Strong Law of Large Numbers, $\mu_{\phi(x)}$ converges to $\mu_{\phi(x)}$ as N increases. If $E_P[\phi(x)^2] < \infty$, a confidence interval of $\mu_{\phi(x)}$ can be constructed using Central Limit Theorem(CLT) as

($\mu_{\phi(x)} - Z_{\alpha/2}\sqrt{var_P[\phi(x)/N]}$, $\mu_{\phi(x)} + Z_{\alpha/2}\sqrt{var_P[\phi(x)/N]}$), where $Z_{\alpha/2}$ is the $100(1-\alpha/2)\%$ point value of standard normal distribution. Since the variance $var_P[\phi(x)]$ is not known beforehand, it should be replaced by sample variance.

Let r. v. X be defined on the probability space (Ω, Γ, P_I), where P_I is IS probability measure, and $dP_I(x)$ be absolutely continuous with respect to $dP(x)$ which has different (the rare events occur more frequently) probability measure compared to $dP(x)$. This setting implies that if the old probability measure is positive, then the new probability measure is also positive. Then

$$\mu_{\phi(x)} = E_P[\phi(x)] = \int \phi(x)dP(x)$$

$$
\begin{aligned}
&= \int \phi(x) \frac{dP(x)}{dP_I(x)} dP_I(x) \\
&= \int \phi(x) L(x) dP_I(x) \qquad (2) \\
&= E_{P_I}[\phi(x)L(x)],
\end{aligned}
$$

where $L(x) = dP(x)/dP_I(x)$ is a likelihood ratio.

Using the samples $\{(\phi(x_1), L(x_1)), (\phi(x_2), L(x_2)) \Lambda, (\phi(x_N), L(x_N))\}$ generated from P_I, an unbiased estimator μ_I is given by

$$\mu_I = \frac{1}{N}\sum_{i=1}^{N}\phi(x_i)L(x_i). \qquad (3)$$

Since the likelihood ratio $L(x) < 1$, the variance reduction is guaranteed :

$$
\begin{aligned}
E_{P_I}[\phi(x)^2 \cdot L(x)^2] &< E_{P_I}[\phi(x)^2 \cdot L(x)] \\
&= E_P[\phi(x)^2].
\end{aligned}
$$

If $E_{PI}[\phi(x)^2 L(x)^2] < \infty$, then the new confidence interval can be calculated as $(\mu_I - z_{\alpha/2}\sqrt{var_{P_I}[\phi(x)\cdot L(x)]/N}, \mu_I + z_{\alpha/2}\sqrt{\mathrm{var}_{P_I}[\phi(x)\cdot L(x)]/N})$. Shahabudin [1994] reports that there exists probability measure P_I which gives variance 0, but it requires the knowledge of the quantity of rare event. The important task in IS is to find an easily tractable measure which guarantees the variance reduction. Therefore, it is necessary to select the optimal sampling distribution that reflects the rare event *E* well. Then the theoretical optimal sampling distribution of IS can be given by

$$dP_I(x) = \phi(x)\cdot dP(x)/\mu_{\phi(x)}. \qquad (4)$$

Using (4), the original estimator can be calculated as follows :

$$\mu_{\phi(x)} = \phi(x_i)L(x_i) = \phi(x_i)\frac{dP(x)}{\phi(x_i)\cdot dP(x)/\mu_{\phi(x)}}. \qquad (5)$$

Since the optimal sampling distribution is dependent on the unknown estimator $\mu_{\phi(x)}$, the random variates x_i's are can not be generated directly from the theoretical optimal sampling distribution in (4).

Wrong selection of sampling distribution of IS in a parametric way may provide an imprecise estimator. Adaptive Important Sampling(AIS) is thus developed to overcome this problem.

2.2 Adaptive Important Sampling (AIS)

It is obvious if we use the optimal sampling distribution then, the variance of estimate is zero as Shahabudin [1994] notes. This implies perfect estimate of $\mu_{\phi(x)}$ can be obtained in a single simulation run. The basic idea of AIS is to assume that the distribution of samples of observations collected from the rare event occurring area and the optimal sampling distribution of IS are the same. That is,

$$dP(x \mid X \in E) = \phi(x) \cdot dP(x) / \mu_{\phi(x)} = dP_I(x) . \qquad (6)$$

AIS uses the simulation results to estimate the parameters of unknown optimal sampling distribution. AIS can save the computational efforts using the probability density function (pdf) of simulation output to estimate the parameters of unknown optimal sampling distribution, and the probability of rare event simultaneously. An AIS algorithm is consist of several short simulations. For each run, $\mu_{\phi(x)}$ and $dP_I(x)$ are estimated. Then the sampling distribution of IS is modified such that its properties match the estimated properties of optimal sampling distribution to be used in the subsequent simulation runs. In this way, the sampling distribution of IS becomes more like the optimal sampling distribution and the estimate of $\mu_{\phi(x)}$ becomes more accurate as the simulation performs successively (For more detailed algorithm of AIS, see Stadler and Roy [1993]).

3. Nonparametric Adaptive Importance Sampling (NAIS)

Improperly selected family of distribution may result in variance inflation even the estimate of parameter is accurate. For example, in a linear system with long memory, the estimated sampling distribution of IS which increases variance and estimates the probability of rare event inaccurately. Thus proper selection of initial distribution is still an open problem in AIS. If a prior knowledge about the sampling distribution of IS is not available for a given system, a nonparametric approach may be more helpful.

Based on AIS, our NAIS uses the initial sampling distribution conditioned on the samples of rare enents occurred during the initial simulation run and uses Zhang's nonparametric idea to estimate the optimal sampling distribution. Silverman[1986] introduces four nonparametric methods to estimate a density function : histogram, kernel estimation, nearest neighbor, and variable kernel. We focus on the kernel estimation method.

We propose the NAIS algorithm as follows :

Step 1 : Initialize a simulation to collect the rare events samples x_i, i=1,…., p.

Step 2 : Esitmate the optimal sampling distribution $f^*_{opt}(x)$ using the kernel function estimation method :

$$f^*_{opt}(x) = \frac{1}{p}\sum_{i=1}^{p}\frac{1}{h}K(\frac{x-x_i}{h}), \quad (8)$$

where h is a smoothing parameter and $K(\ \square\ )$ is a simple rectangular kernel function such as

$$K(x) = \begin{cases} \frac{1}{2}, & if\ |x| < 1 \\ 0, & otherwise. \end{cases}$$

Step 3 : Run a simulation with $f^*_{opt}(x)$ as the optimal sampling distribution of IS, and calculate $\mu_{\phi(x)}$ as follows :

$$\mu_{\phi(x)} = \frac{1}{n}\sum_{i=1}^{n}\frac{f(x_i)}{f^*_{opt}(x_i)}\phi(x_i),$$

where n is the number of replications or the number of regeneration cycles.

5. Numerical Results

We test the proposed NAIS algorithm in an M/M/1 queueing model. Let λ and μ be the mean arrival rate and the mean service rate in an M/M/1 queueing model, respectively. Consider a problem of estimating the probability $\mu_{\phi(x)}$ that the number of customers reaches a certain queue level of A during a busy period.

Now the NAIS algorithm can be described as follows :

Step 1 : Initialize a short simulation to collect the samples of interarrival and service times when the number of customers reaches the queue level A during a busy period.

Step 2 : Use the kernel function estimation method with the samples collected from Step 1 to estimate the optimal sampling distributions of interarrival and service times to be used later simulation.

Step 3 : Proceed the simulation with estimated optimal sampling distributions from Step 2.

The sampling distributions of IS can be decided by modification of samples of interarrival and service times which occurred independently during interested busy period. Assume the sample path ω of reaching the queue level A during a busy period occurrs at time τ before the system becomes empty. If the number of departures is m, then there are $A+m-1$ arrivals during a busy period. For any busy period, if a sample path ω represents the arrivals and departures from the queue, then the likleihood function $L(\omega)$ can be defined as follows :

$$L(\omega) = \prod_{i=1}^{A+m-1} \frac{f(t_i)}{f^*(t_i)} \bullet \prod_{j=1}^{m} \frac{g(s_j)}{g^*(s_j)}, \tag{7}$$

where t_i : interarrival time for the i^{th} customer

s_j : service time for the j^{th} customer

f : pdf for interarrival times

f^* : sampling pdf of NAIS for interarrival times

g : pdf for service times

g^* : sampling pdf of NAIS for service times.

For N busy periods (it could be obtained either N independent replications or N regenerative periods), the probability $\mu_{\phi(x)}$ can be calculated as follows :

$$\mu_{\phi(x)} = \frac{1}{N}\sum_{i=1}^{N} L(\omega_i) \bullet \phi(\omega_i), \qquad i = 1,\dots N, \tag{8}$$

where $\phi(\omega) = \begin{cases} 1, & \text{if the number of customers reaches the queue level } A \\ 0, & \text{otherwise.} \end{cases}$

We run the Monte Carlo simulation and the NAIS simulation varying the queue level of A for 10,000 busy periods for 10 times.

Table 1 and Table 2 show the results of the NAIS simulation and the Monte Carlo simulation, respectively : the estimate $\mu_{\phi(x)}$ of the probability that the number of customers reaches the queue level of A, the standard deviation (SD) of $\mu_{\phi(x)}$, and the half-width of confidence interval for $\mu_{\phi(x)}$.

Table 1. NAIS and Monte Carlo simulation result of M/M/1 queueing model (10,000 busy periods)

($\lambda = 0.3$, $\mu = 0.5$)

Queue level(A)	$\mu_{\phi(x)}$		SD		Half-width	
	Monte Carlo	NAIS	Monte Carlo	NAIS	Monte Carlo	NAIS
10	2.53E-03	2.36E-03	2.96E-04	1.02E-04	5.80E-04	2.00E-04
15	1.40E-04	1.25E-04	9.66E-05	4.88E-06	1.89E-04	9.56E-06
20	3.00E-05	7.37E-06	4.83E-05	2.25E-07	9.47E-05	4.42E-07
25		6.61E-07		3.10E-08		6.08E-08
30		6.41E-08		4.82E-09		9.45E-09
35		3.27E-09		1.46E-10		2.86E-10
40		5.38E-11		4.78E-11		9.36E-11

In Table 1, the probability $\mu_{\phi(x)}$ decreases as the queue level of A increases. As shown in Table 2, the probability below 10^{-5} can not be estimated in the Monte Carlo simulation since the number of rare events are not sufficient. The curves in Figure 1 are the NAIS and Monte Carlo (MC) simulation results of $\mu_{\phi(x)}$.

Figure 1. Probability of customer reaches the level A during busy period

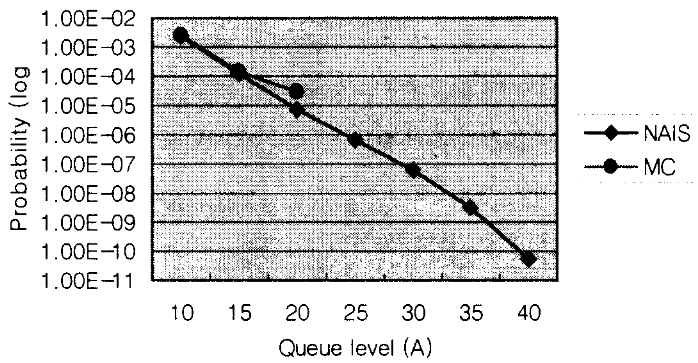
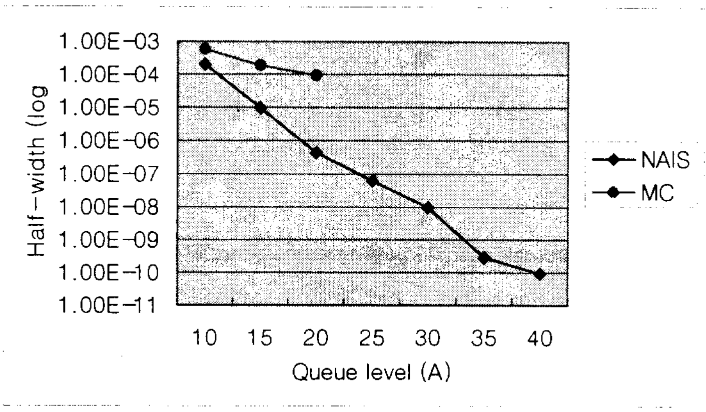


Figure 2 shows the half-widths of confidence intervals of the NAIS and MC simulations. We can see that the variance reduction is obtained by the NAIS. Since the variance is reduced, the confidence intervals become tighter which implies the estimates become more accurate.

Figure 2. Half-widths of confidence intervals



Conclusion

This paper proposed a modified fast simulation technique NAIS, and demonstrated in the estimation of rare event probabilities in an M/M/1 queuing model. The experiments with NAIS in our work show the substantial gains of computational efficiency and the variance reduction in both models compared to classical Monte Carlo simulation. The difficulty of proper choice in optimal sampling distribution of IS and AIS can be eased by NAIS, since we can directly use the data collected during the simulation to modified the optimal sampling distribution regardless of the characteristics of the system. To improve the efficiency in modifying the estimate of optimal sample distribution, a more complicated kernel function, such as Gaussian density is worth of trial. We noted that the time for random number generation could be saved if we use more efficient random variate generation technique. Other than Acceptance-Rejection technique should be tested since too many random variates are discarded. It is also desired to develop a method to guarantee an optimal sampling distribution, which is invertible in regardless of density estimation methods. Easy random variate generation technique, such as inverse transformation method is much faster than the acceptance-rejection method. NAIS can be used for simulation of highly reliable systems whose general characteristics are not known beforehand.

References

1. Frost, V., Larue, W., and Shanmugan, K., "Efficient Techniques for the Simulation of Computer Communications Networks," *IEEE J. on Selected Areas in Communications*, Vol. 6, No. 1, pp.146-157,

1998.

2. Givens, G. H. and Raftery, A. E., "Local Adaptive Importance Sampling for Multivariate Densities with Strong Nonlinear Relationships," *J. of the American Statistical Association*, Vol. 91, pp.132-141, 1996.

3. Glynn, P., and Iglehart, D., "Importance Sampling for Stochastic Simulations, " *Management Science* Vol. 35, pp.1367-1392, 1989.

4. Hansen, V. G. "Detection Performance of Some Non-parametric Rank Tests and an Application to Radar, " *IEEE Trans. Inform. Theory*, Vol. 16, No.3, pp.309-318, 1970.

6. Mike West, "Modeling with Mixtures, " in *Bayesian Statistics 4*, eds. J. M. Bernardo et al., Oxford, U.K. Oxford University Press, pp.503-524, 1992.

7. Mike West, "Approximating Posterior Distributions by Mixtures," *J. of the Royal Statistical Society, Ser. B.* vol. 55, pp.409-422, 1993.

8. Oh, M. S., and Berger, J., "Adaptive Importance Sampling in Monte Carlo Integration," *J. of Statistical Computation and Simulation*, Vol. 41, pp.143-168, 1992.

9. Oh, M. S., and Berger, J., "Integration of Multimodal Functions by Monte Carlo Importance Sampling," *J. of American Statistical Association*, Vol. 88, pp.450-456, 1993.

10. Orsuk, G. and Azaharg, B. "On the Theory of Importance Sampling Applied to the Analysis of Detection Systems, " *IEEE Trans. Comm.*, Vol. 37, No. 4, pp.332-339, 1989.

11. Shahabuddin, P., "Importance Sampling for the Simulation of Highly Reliable Markovian Systems," *Management Science*, Vol. 40, No.3, pp. 333-352, 1994.

12. Shangmugan, K. S. and Balaban, P., "A Modified Monte Carlo Simulation Technique for the Evaluation of Error Rate in Digital Communication Systems," *IEEE Trans Comm.*, Vol. 28, No. 11, pp.1916-1924, 1980.

13. Siegmurd, D., "Importance Sampling in Monte Carlo Study of Sequential Test," *The Annals of Statistics*, Vol. 4, pp.673-684, 1976.

14. Silverman, B., Density Estimation, London: Chapman and Hall, 1986.

15. Smith, P., Shati, M., and Gao, H., "Quick Simulation: A Review of Importance Sampling Techniques in Communication Systems," *IEEE J. on Selected Areas in Communications*, Vol. 15, No. 4, pp.597-613, 1997.

16. Stadler, J. S. and Roy, S., "Adaptive Importance Sampling," *IEEE J. on Selected Areas in Communications*, Vol. 11, No. 3, pp.309-316, 1993.

17. Strickland, S., "Optimal Importance Sampling for Quick Simulation of Highly Reliable Markovian Systems," *Proc. '93 Winter Simulation Conference*, Los Angeles, U.S.A., pp.437-444, Dec. 1993.

18. Zhang, P. "Non-parametric Importance Sampling ," *J. of the American Statistical Association*, Vol. 91, No. 435, pp.1245-1253, 1996.