Coverage Analysis of Sequential Regenerative Simulation *

Jong-Suk Ruth Lee
Department of Computer Science
University of Canterbury, Christchurch, New Zealand
e-mail: ruth@cosc.canterbury.ac.nz

Abstract

Regenerative simulation (RS) is a method of stochastic steady-state simulation in which output data are collected and analysed within regenerative cycles (RCs). Since data collected during consecutive RCs are independent and identically distributed, there is no problem with the initial transient period in simulated processes, which is a perennial issue of concern in all other types of steady-state simulation. In this report, we address the issue of experimental analysis of the quality of sequential regenerative simulation in the sense of the coverage of the final confidence intervals of mean values. The ultimate purpose of this study is to determine the best version of RS to be implemented in Akaroa2 [1], a fully automated controller of distributed stochastic simulation in computer network environments.

Keywords : regenerative simulation, sequential steady-state simulation, coverage analysis

1 Introduction

Sequential statistical analysis of output data in stochastic simulation, used for controlling the length of simulation, is regarded as the only practical way of securing appropriate level of credibility of the final simulation results [2]. Following this approach, simulation is progressing from one checkpoint to the next one, until a prespecified accuracy of all point estimators is obtained. Probably the most commonly used stopping criterion for sequential steady-state simulation is the *relative precision*, defined as the ratio of the half-width of the confidence interval (at a given confidence level) and the current estimate of a given estimated performance measure [3]. An experiment is stopped at the checkpoint at which the required relative precision of the final results is reached.

In non-regenerative methods of steady-state simulation output data analysis, like Spectral Analysis and Batch Means, one has to discard data collected during the initial transient periods and observe the process over sufficiently long time period later on, to obtain satisfactorily credible estimates. Determination of the length of the initial transient period is often non-trivial and likely to require sophisticated statistical techniques [3]. Therefore, regenerative method of analysis of simulation output data is very attractive alternative, because it avoids

*Technical Report TR-COSC 02/99, May 1999

this problem. In regenerative stochastic processes, regenerative cycles (RCs) produce batches of independent and identically distributed data, and the final precision of results depends on the number of RCs observed.

Standard sequential stopping rules of sequential simulation [3], like the *relative precision* can be used also in conjunction with RS (regenerative simulation). However, sequential steady-state RS can lead to inaccurate results if the simulation experiment stops too early, when the sequential stopping criterion is accidently temporarily met. Some sequential stopping rules for RS were proposed and tested by Sauer [4] and Lavenberg and Sauer [5]. Following the stopping rule proposed in [5], the simulation should be stopped when the minimum number of RCs is observed (assumed to be 10) and the required precision is reached. In [4], it was argued that the simulation run length should be associated with some minimum simulation time. As the results of our studies show, such approaches are not longer satisfactory or needed, taking into account currently available computing resources.

One of the main quality criteria used for assessing the quality of methods of simulation output data analysis in stochastic simulation is the coverage of the final confidence intervals they produce, defined as the proportion of the final confidence intervals which contain the true value of the analysed performance measure. Such experimental confidence level should be confronted with the theoretical confidence level of the final estimates. Any good method of analysis of simulation output data should produce narrow and stable confidence intervals, and the relative frequency of such an interval containing the true value of the estimated performance measure should not differ from the assumed theoretical confidence level.

In the past, coverage analyses of various sequential stopping rules for RS, including those in [4] and [5], were conducted using fixed numbers of replications (for example, 50 and 100, as [4] and [5], respectively). But, as recently argued in [6], coverage analysis should be conducted sequentially, to secure statistically accurate results. The rules of sequential coverage analysis for non-RS have been proposed in [6]. In this report, an adaptation of these rules for sequential RS is presented in Section 4. This is an enhanced version of the coverage analysis, based on F distribution, which, as shown in [7], leads to more efficient interval estimators of proportions. The numerical results of coverage analysis of the sequential RS applied for estimating steady-state means, and reported in Section 4, were obtained in our quest for the most robust method of sequential analysis of simulation output data, to be implemented in Akaroa2 [1], a fully automated controller of distributed stochastic simulation on multiple networked processors, in Multiple Replications In Parallel (MRIP) scenario [8]. The results of coverage analysis of two other methods of sequential estimation of steady-state means, namely based on Non-overlapping Batch Means and Spectral Analysis (in its version originally proposed by Heidelberger and Welch [9]) were presented in [6]. The results of coverage analysis of sequential methods of estimation of steady-state quantiles are reported in [10] and [11].

Analysis of coverage is of course limited to analytically tractable systems, since the theoretical value of the parameter of interest has to be known. Because of that, it has even been claimed that there is no justification for experimental coverage analysis, since there is no theoretical basis for extrapolating results found for simple, analytically tractable systems to more complex systems, which are subjects of practical simulation studies [12]. On the other hand, no theory of coverage for finite sample sizes exists, and in this situation, experimental coverage analysis of analytically tractable systems remains the only method available for testing validity of methods proposed for simulation output analysis. Certainly nobody is ready to accept a method of simulation output data analysis showing very poor quality in experimental studies of coverage.

2 The Properties of RS

As known, RS is based on the assumption that any regenerative process starts afresh (probabilistically) at each consecutive regenerative point. Thus, observations grouped into batches of random length, determined by successive regenerative instants of the simulated process, are statistically independent, and that includes the first RC, if the simulation starts from a regenerative state.

For instance, when simulating an M/G/1/∞ queueing system, any instant of time when this system reaches the state 0 (no customer present) represents a regenerative point at the boundary of two consecutive RCs. After any such instant of time, no event from the past influences the future evolution of the system. As a consequence of the independent and identically distributed output data within consecutive RCs, the problems related with the initial transient period and correlations between batches of data vanish [13], [14], [15], [16].

While the accuracy of the final simulation results from RS depends on the number of simulated RCs, the rate at which RCs occur depends on the simulated system. For example, in heavily loaded but stable queueing systems regenerative states can occur very rarely, making the RS very ineffective, since it becomes difficult, if possible at all, to form a reliable point estimate and its confidence interval.

As known, RS uses estimators in the form of a ratio of two variables; see for example [13]. To estimate steady-state mean EX of, for example, waiting times in a queueing system on the basis of observed waiting times $x_1, x_2, x_3, \ldots$, of consecutive customers, we collect the pairs of (secondary) output data $(a_1, y_1), (a_2, y_2), \ldots, (a_n, y_n)$ which are realisations of i.i.d. random variables A_i and Y_i, $1 \leq i \leq n$, where A_i and Y_i denote, respectively, the number of customers processed and the sum of the waiting times in ith RC. Let $\overline{y}(n)$, $\overline{a}(n)$, $s_{11}^2(n)$, $s_{22}^2(n)$, and $s_{12}^2(n)$ be the usual unbiased estimators for $E[Y]$, $E[A]$, $Var[Y]$, $Var[A]$, and $Cov[Y, A]$ for any i, respectively; that is

$$\overline{y}(n) = \frac{1}{n}\sum_{i=1}^{n} y_i, \quad \overline{a}(n) = \frac{1}{n}\sum_{i=1}^{n} a_i,$$

$$s_{11}^2(n) = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \overline{y}(n))^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} y_i^2 - (\sum_{i=1}^{n} y_i)^2/n\right),$$

$$s_{22}^2(n) = \frac{1}{n-1}\sum_{i=1}^{n}(a_i - \overline{a}(n))^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} a_i^2 - (\sum_{i=1}^{n} a_i)^2/n\right),$$

and

$$s_{12}^2(n) = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \overline{y}(n))\,(a_i - \overline{a}(n)) = \frac{1}{n-1}\left(\sum_{i=1}^{n} y_i a_i - (\sum_{i=1}^{n} y_i * \sum_{i=1}^{n} a_i)/n\right).$$

As a consequence of the *strong law of large numbers* [13], the point estimator of the mean

$$\hat{r}(n) = \frac{\overline{y}(n)}{\overline{a}(n)}$$

is strongly consistent estimator of steady-state mean EX; that is, $\hat{r}(n) \to EX$ as $n \to \infty$. Moreover, the estimator for variance

$$s^2(n) = \{s_{11}^2(n) - 2\hat{r}(n)s_{12}^2(n) + \hat{r}^2(n)s_{22}^2(n)\}$$

is also strongly consistent; that is, $s^2(n) \to Var(X)$ as $n \to \infty$.

A $100(1-\alpha)\%$ confidence interval for the steady-state mean obtained by applying RS is given by

$$\hat{r}(n) \pm \frac{s(n)z_{1-\alpha/2}}{\overline{a}(n)\sqrt{n}}, \tag{1}$$

where $z_{1-\alpha/2}$ is the $(1-\alpha/2)$ quantile of the standard normal distribution [13], [14], [15], [16].

3 Sequential Procedures for RS

This section presents in detail sequential procedures for stopping RS experiment when the required relative precision of confidence intervals is achieved. Among the few possible criteria for stopping RS, we adapt a stopping criteria which is based on the relative half-width of the confidence interval at a given confidence level $(1-\alpha)$, defined as the ratio

$$\epsilon(n) = \frac{\Delta_x(n)}{\hat{r}(n)}, \tag{2}$$

where $\Delta_x(n) = (s(n)z_{1-\alpha/2})/(\overline{a}(n)\sqrt{n})$ and $\epsilon(n), 0 < \epsilon(n) < 1$, is the relative precision of the confidence interval obtained on the basis of n RCs.

The simulation experiment is stopped at the first checkpoint for which $\epsilon(n) \leq \epsilon_{max}$, where ϵ_{max} is the required limit relative precision of the results at the $100(1-\alpha)\%$ confidence level, $0 < \epsilon_{max} < 1$.

Sequential RS is described below by a pseudocode procedure that uses the following parameters:

```
(1 - \alpha) : The assumed confidence level of the final results
               (0 < \alpha < 1)
Maximum Relative Precision (\epsilon_{max}) : The maximum acceptable
               value of the relative precision of confidence intervals
               (0 < \epsilon_{max} < 1)

PROCEDURE RegenerativeAnalysis;
{Uses the regenerative method for one ratio estimator}

  PROCEDURE GetNextRC;
```

```
    * Get a RC by collecting obs. until a regenerative point is detected.

    * Collect information of sum and length of a RC.
       - RCSum
       - RCLength

    * Collect the following statistics for estimating variance s^2(n)
         with RCSum and RCLength of RCs
       - MeanRCLength = SUM(RCLength) / NRCs;
       - MeanRCSums = SUM(RCSum) / NRCs;
       - SumofSqRCSums = SUM(RCSum*RCSum);
       - SumofSqRCLengths = SUM(RCLength*RCLength);
       - SumofRCSumbyRCLength = SUM(RCSum*RCLength);

  END GetNextRC;

  PROCEDURE UpdateStatistics;
  {Update the overall variance and the mean using their classical estimators.
   The sums are updated dynamically offering a quicker method for determining
   the overall variance, than the looping mechanism is used by the jackknife
   estimator.}

    * Update the following statistics using formulae s^2_{11}(n),
           s^2_{22}(n), and s^2_{12}(n).
      - VarTourSums = s^2_{11}(n);
      - VarTourLengths = s^2_{22}(n);
      - covariance = s^2_{12}(n);

    * Calculate the overall mean and overall variance using a simple
           ratio estimator.
      - OverallMean = MeanRCSums / MeanRCLength;
      - OverallVariance = s^2(n);

  END UpdateStatistics;

BEGIN {main procedure}

  {initialise parameters for calculating statistics from observations in RC}
  NRCs = 1;                {Number of RCs collected}
  RCSum = 0;               {Sum of the observations within a RC}
  RCLength = 0;            {Length of a single RC}
  MeanRCSums = 0.0;        {Overall mean of RC observations}
  MeanRCLength = 0.0;      {Overall mean of RC lengths}
                               {For estimating variance s^2(n)}
  SumofSqRCSums = 0.0;         {Sum of squares of sum of obs. in a RC}
```

```
  SumofSqRCLengths = 0.0;     {Sum of squares of RC lengths}
  SumofRCSumbyRCLength = 0.0; {Sum of RC lengths by RC sums}

  StopSimulation = false;
        {a condition of stopping the simulation has not been met yet}

  Call GetNextRC;

  while (not StopSimulation) {do}

    * Call GetNextRC;

    * Call UpdateStatistics;
      {To ensure the formulae s^2_{11}(n), s^2_{22}(n), and s^2_{12}(n)
       are not divided by zero, call UpdateStatistics after minimum 2 RCs
       collected.}

    * Update the value of the relative precision using Equations (1) and (2).
      if (relative precision <= Maximum Relative Precision)
           StopSimulation = true;
      else StopSimulation = false;

  enddo;

END RegenerativeAnalysis;
```

4 Coverage Analysis for Sequential RS

In sequential RS with a stopping rule based on the *relative precision*, inaccurate estimates can be obtained if the stopping criterion is accidently temporarily satisfied, having recorded an insufficient number of RCs. As a consequence of this, sensible practise is to ensure that estimates do not come from simulation experiments with too few RCs. Recognising the significance of this factor, we have adjusted stopping rules for sequential RS by ensuring that minimum of 30 RCs in a single simulation of $M/M/1/\infty$ queueing system have to be observed before it is stopped [17].

This minimum of 30 RCs as the shortest acceptable length of sequential RS was found experimentally and can be supported by such results as those reported in Table 1, obtained during RS of $M/M/1/\infty$ queueing system. One can see that such very short simulation runs do have very poor coverage, below 10%, for the assumed theoretical coverage of 95%. The elimination of too short simulation runs significantly improves the quality of sequential RS, as documented by the results of coverage analysis in Figures 1 and 2. These figures show the results of sequential coverage analysis of $M/M/1/\infty$ queueing system loaded at 0.5, with and without the restriction on the minimum of 30 recorded RCs as the length of simulation. The figures also show high initial instability of coverage. This phenomenon, similar to that reported in [6], has been the main motivation behind the proposal of sequential analysis of coverage. It is clear that the coverage analysis has to be done over sufficiently large sample of data (in this case: after sequential simulation is repeated sufficiently many times).

Ideally, the confidence interval of coverage for a method of simulation output data analysis should cover the confidence level assumed for the final results [4]. In practice, this criterion is hardly met by any method of simulation output data analysis, so, making this requirement weaker, we accept the method for practical applications if the confidence interval of its coverage is sufficiently close to the confidence level assumed. However, Figures 1 and 2 show that the final coverage was far away from the required level of 0.95.

As argued in [6], this could be caused by the fact that an insufficient number of bad final confidence intervals was recorded. (As in [6], a bad confidence interval means a confidence interval that does not cover the theoretical value of the estimated parameter). Following [6], we assumed that representativeness of data for coverage analysis requires that minimum 200 bad confidence intervals have to be recorded before sequential analysis of coverage can commence. Typical convergence of coverage to its final accurate level, if too short simulation runs are discarded when minimum number of 200 bad confidence intervals are recorded, is shown in Figure 3. Again one can see that the statistical "noise" introduced by too short simulation runs should be removed before correct conclusions regarding the quality of a given method of simulation output analysis (in this case: the RS) are drawn. As shown in Figure 3, this resulted in a jump of coverage from 0.9 to 0.95. Thus, the results of coverage of sequential RS reported in this report were obtained sequentially, until at least 200 bad confidence intervals have been recorded and having results coming from sequential RS not shorter than 30 RCs. These results will be additionally confronted with the results obtained following previously used method of coverage analysis, based on the fixed-sample size approach.

All results for sequential RS were obtained assuming the required precision of the final result 5% or less, at the confidence level of 0.95. The same stopping criterion applied in our sequential coverage analysis. Additionally, only simulation runs of minimum 30 RCs were taken into account, and the interval estimator of coverage was based on F distribution to ensure that the sequential analysis of coverage does not last excessively long [7].

The results of coverage reported in this section were obtained on the basis of simulation of $M/M/1/\infty$ queueing systems. The results of coverage of the sequential RS obtained from non-sequential analysis are presented in Figure 4, while Figure 5 show the same results obtained sequentially. One can clearly see that the sequential coverage analysis, with filtering off too short simulation runs and requiring recording of at least 200 bad confidence intervals, produces better (more reliable, as we have argued) results.

Generally, our results show that the sequential RS used for analysis of steady-state means can be considered as a good candidate for being implemented in such simulation packages as Akaroa2, where whole process of simulation output data is conducted automatically during simulation. Before the final recommendation is done, one should conduct full study of coverage of this method of simulation output analysis by including wider spectrum of its applications, over a range of standard stochastic systems and processes.

5 Conclusions

Sequential run length control of stochastic simulation is the only efficient way for securing precision of the final simulation results. In this report we have been applied the rules of sequential coverage analysis for methods of output analysis used in sequential RS used for estimation of steady-state means.

Our initial results, obtained when using M/M/1/∞ queueing systems used as the reference model, indicate the RS in its sequential version is an attractive solution for practitioners if special care is taken for avoiding too short simulation runs. Our coverage analysis of this RS is continued by studying its applications over a broader spectrum of simulation reference models. On the other hand, additional problems have to be solved before this method can be offered in fully automated simulation tools as Akaroa2. These include rules for determination of (approximate) regenerative points.

References

[1] G. C. Ewing, K. Pawlikowski, and D. McNickle. Akaroa2: Exploiting Network Computing by Distributing Stochastic Simulation. In *13th European Simulation Multiconference*, 1999.

[2] A. M. Law and W. D. Kelton. Confidence Intervals for Steady-State Simulations, II: A Survey of Sequential Procedures. *Management Science*, 28(5):550–562, 1982.

[3] K. Pawlikowski. Steady-State Simulation of Queueing Processes: A Survey of Problems and Solutions. *ACM Computing Surveys*, 22(2):122–170, 1990.

[4] C. H. Sauer. Confidence Intervals for Queueing Simulations of Computer Systems. *ACM Performance Evaluation Review*, 8(1-2):46–55, 1979.

[5] S. S. Lavenberg and C. H. Sauer. Sequential Stopping Rules for the Regenerative Method of Simulation. *IBM J. Research Development*, pages 545–558, 1977.

[6] K. Pawlikowski, D. C. McNickle, and G. Ewing. Coverage of Confidence Intervals in Sequential Steady-State Simulation. *Simulation Practice and Theory*, 6:255–267, 1998.

[7] J. R. Lee, D. McNickle, and K. Pawlikowski. Confidence Interval Estimators for Coverage Analysis in Sequential Steady-State Simulation. In *Proceedings of the Twenty Second Australasian Computer Science Conference, Auckland*, pages 87–98, 1999.

[8] K. Pawlikowski, V. Yau, and D. C. McNickle. Distributed and Stochastic Discrete-event Simulation in Parallel Time Streams. *Proceedings of the 1994 Winter Simulation Conference*, pages 723–730, 1994.

[9] P. Heidelberger and P. D. Welch. A Spectral Method for Confidence Interval Generation and Run Length Control in Simulations. *Communications of the ACM*, 25:233–245, 1981.

[10] J. R. Lee, D. McNickle, and K. Pawlikowski. Sequential Estimation of Quantiles. Technical Report TR-COSC 05/98, Department of Computer Science, University of Canterbury, Christchurch, New Zealand, 1998.

[11] J. R. Lee, D. McNickle, and K. Pawlikowski. Quantile Estimation in Sequential Steady-State Simulation. In *13th European Simulation Multiconference*, 1999.

[12] B. Fox. Estimation and Simulation. *Management Science*, 24:860–861, 1978.

[13] G. S. Shedler. *Regenerative Stochastic Simulation.* Academic Press, Inc., 1993.

[14] D. L. Iglehart. Simulating Stable Stochastic System, V: Comparison of Ratio Estimators. *Naval Res. Logist. Quart.*, 22:553–565, 1975.

[15] M. A. Crane and A. J. Lemoine. *An Introduction to the Regenerative Method for Simulation Analysis in Lecture Notes in Control and Information Sciences.* Springer Verlag, 1977.

[16] M. A. Crane and D. L. Iglehart. Simulating Stable Stochastic Systems:III. Regenerative Processes and Discrete Event Simulations. *Operations Research*, 23(1):33–45, 1975.

[17] J. R. Lee, D. McNickle, and K. Pawlikowski. Quality of Sequential Regenerative Simulation. In *13th European Simulation Multiconference*, 1999.

Table 1: The number of too short simulation runs (less than 30 RCs) in 3000 simulation replications and their coverage ($M/M/1/\infty$, theoretical confidence level = 0.95).

Load	*Number of too short runs*	*Coverage*
0.1	158	6.3%
0.2	167	5.4%
0.3	159	4.4%
0.4	156	5.8%
0.5	166	3.6%
0.6	159	3.1%
0.7	191	4.7%
0.8	281	3.6%
0.9	450	6.0%

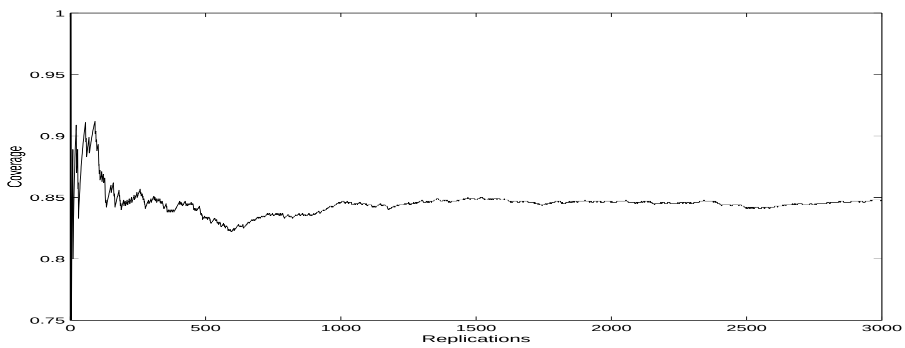

Figure 1: Convergence of coverage analysis for sequential RS with no restriction on the minimum run length (M/M/1/∞, load = 0.5).

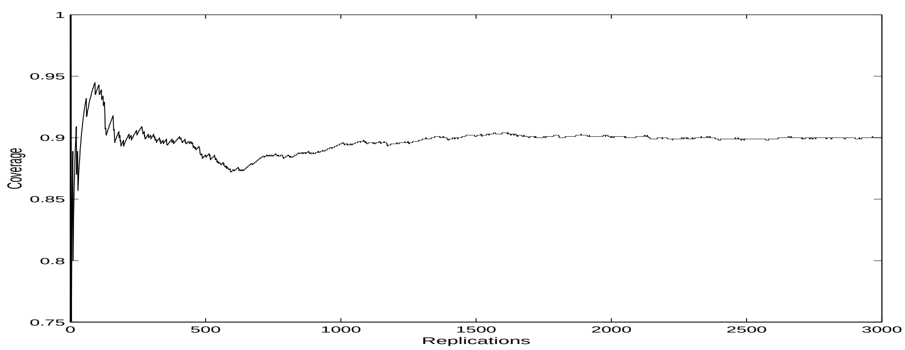

Figure 2: Convergence of coverage analysis for sequential RS with the minimum length of 30 RCs before stopping (M/M/1/∞, load = 0.5).

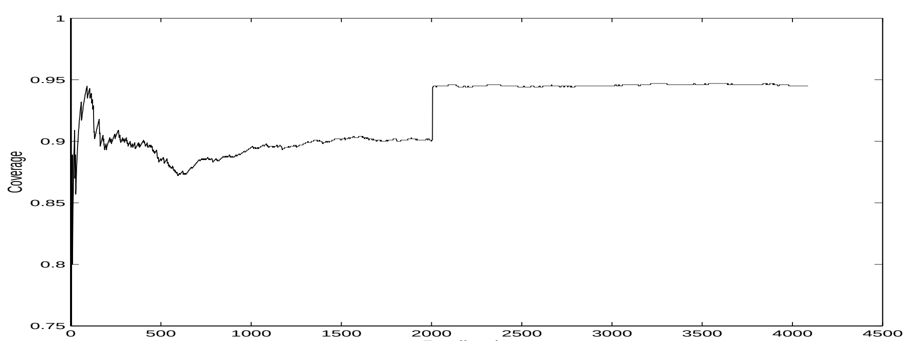

Figure 3: Convergence of coverage analysis for sequential RS with the minimum length of 30 RCs, and 200 bad confidence intervals (M/M/1/∞, load = 0.5).

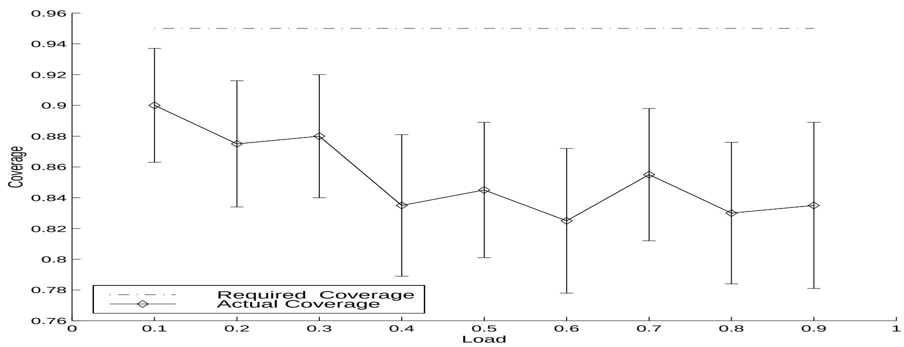

Figure 4: Non-sequential coverage analysis of sequential RS (200 replications; M/M/1/∞).

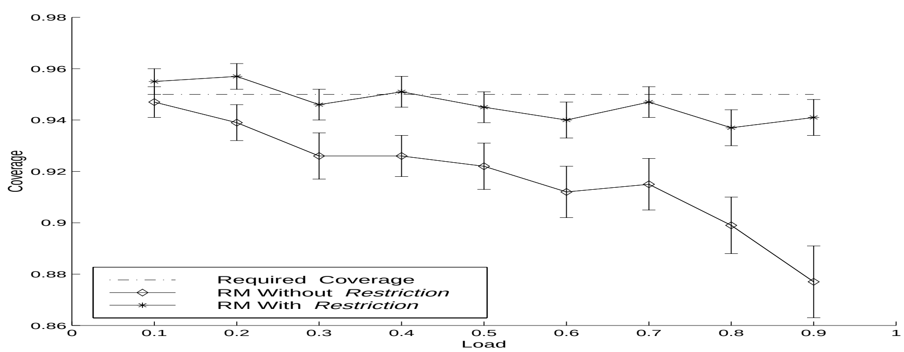

Figure 5: Sequential coverage analysis of sequential RS without and with the restriction on the minimum run length and the number of bad confidence intervals (M/M/1/∞).