Submitted for European Simulation Symposium 2002

SPECTRAL ANALYSIS FOR CONFIDENCE INTERVAL ESTIMATION UNDER MULTIPLE REPLICATIONS IN PARALLEL

Gregory C. Ewing
and
Krzysztof Pawlikowski

Department of Computer Science
University of Canterbury
Christchurch, New Zealand

Donald McNickle

Department of Management
University of Canterbury
Christchurch, New Zealand

ABSTRACT

Multiple Replications in Parallel (MRIP) have been proposed for speeding up sequential stochastic discrete-event simulation. In this scenario, multiple processors concurrently produce statistically equivalent sequences of observations that are pooled by a global analyser. Such distributed production of observations allows to shorten the time spent on collecting the number of observations needed for obtaining the final results with acceptably small statistical errors.
The distributed generation of output data in MRIP simulation causes that new estimators (linear combinations of ordinary sequential estimators) have to be applied. In this paper we discuss the basic properties of an MRIP estimator of steady-state mean value that is such a linear combination of non-distributed estimators of the method of Spectral Analysis proposed by Heidelberger and Welch (SA/HW). The MRIP version of SA/HW has been assessed on the basis of its practical implementation in Akaroa2.

1 INTRODUCTION

Sequential stochastic discrete-event simulation, i.e. stochastic simulation with on-line analysis of output data, is generally accepted as the most effective way for securing representativeness of samples of observations collected during simulation (Heidelberger and Welch 1983, Law 1983, Law and Kelton 1992). In this scenario, a simulation experiment is stopped when the statistical error of estimates reaches the required (low) level.

Practical applications of sequential simulation are often hindered by extremely long times required for collecting satisfactory large numbers of observations needed for producing the results with acceptably small statistical errors. This problem can be overcome by executing stochastic simulation in Multiple Replications in Parallel (MRIP) scenario; see (Pawlikowski, Yau and McNickle 1994). In this scenario, multiple processors concurrently produce statistically equivalent sequences of observations that are pooled by a global analyser. Such distributed production of observations allows to shorten the time spent on collecting the number of observations needed for obtaining the final results with a required level of statistical errors; see (Pawlikowski and McNickle 2001) for theoretical limitations of the resulted speedup.

The concept of MRIP has been implemented in fully automated way in Akaroa2; see (Ewing, Pawlikowski and McNickle 1999). A user of Akaroa2 needs only specify a required (relative) statistical error, confidence level for each performance parameter whose mean value is sought, and the number of processors of a local computer network to be used as simulation engines, see Figure 1. During the simulation, Akaroa2's central controlling process (akmaster) repeatedly estimates the confidence interval of each mean value, at the specified confidence level. When the required statistical precision of the all results has been reached, the simulation is automatically stopped.

However, the distributed generation of output data in MRIP simulations causes that special estimators, constructed as linear combinations of ordinary sequential estimators, have to be used. In this paper we discuss the basic properties of an MRIP estimator of steady-state mean values, defined as a linear combination of non-distributed estimators of the method of Spectral Analysis proposed by Heidelberger and Welch (SA/HW) in (Heidelberger and Welch 1981). The MRIP version of SA/HW is here assessed on the basis of its practical implementation in Akaroa2.

Our focus on this method of simulation output data analysis is motivated by the fact that SA/HW is the only currently known method of sequential estimation of steady-state mean values in which designers have

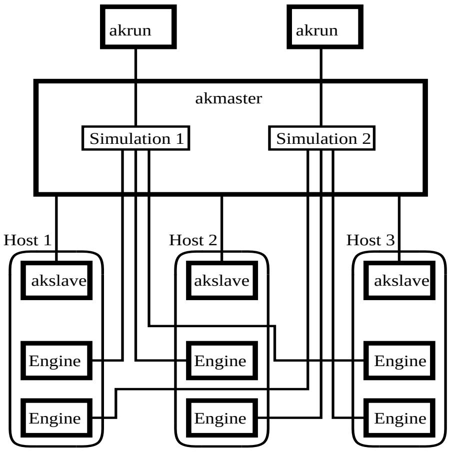


Figure 1: Block diagram of a typical Akaroa2 process structure, showing the central controlling process ("akmaster") with two simulations in progress, each using three simulation engines spread over different physical processors. The "akrun" processes provide the user interface for launching simulations, monitoring their progress and reporting the results.

large freedom of deciding about the granularity of sequential data analysis. Having selected the appropriate granularity one should be able to achieve the speedup limited only by the truncated Amdahl law formulated in (Pawlikowski and McNickle 2001).

2 METHOD OF SPECTRAL ANALYSIS

The Spectral Analysis method of estimation of the variance of steady-state mean μ_X from an auto-correlated sequence of observations $x_0, x_1, ...$ was originally proposed by Heidelberger and Welch (Heidelberger and Welch 1981). The variance is obtained as the value of the periodogram $\Pi(f)$ (of the analysed sequence of abservations) at frequency $f = 0$. Because of high variability of a typical periodogram at low frequencies, in SA/HW its value at $f = 0$ is obtained through a regresion fit to the logarithm of the averaged periodogram, where fitting is done using a polynomial of degree d (typically $d \leq 2$). The fitting is done using K fixed points of the periodogram $\Pi(f)$. As was proved in (Heidelberger and Welch 1981), if $d = 2$, then the confidence interval of μ_X can be obtained using quantiles of Student t-distribution with the number of degrees of freedom *df* $= 7$ (if $K = 25$) , or *df* $= 16$ (if $K = 50$). By virtue of spectral analysis of output data, the periodogram can be calculated either over the sequence of individual observations or over the sequence of their batch means. In the latter, observations can be grouped into batches of arbitrary size m, $m \geq 1$, purely for the purpose of data aggregation.

In Akaroa2, a sequential version of SA/HW described in (Pawlikowski 1990) is used by each simulation engine participating in MRIP simulation. If P simulation engines are employed, then whenever simulation engine i, $i = 1, 2, ..., P$, reaches consecutive checkpoint of (its replication of) the simulation, it calculates the *local estimate* $\bar{X}_i(n_i)$ of an analysed mean value μ_X and the estimate of its variance $\hat{V}[\bar{X}_i(n_i)]$, using all n_i observations that it has so far generated. In the current version of Akaroa2, the latter estimates are obtained from the regression fit of paraboles, i.e. assuming $d = 2$.

The central controlling process takes the most recent local estimates produced by each of participating simulation engines whenever a simulation engine reaches its new checkpoint and combines them into a *global estimate* $\bar{X}(P)$ of mean μ_X and the estimate of variance $\hat{V}[\bar{X}(P)]$ of this pooled estimator, by using the formulas:

$$\bar{X}(P) = \frac{1}{n} \sum_{i=1}^{P} n_i \bar{X}_i(n_i)$$

$$\hat{V}[\bar{X}(P)] = \frac{1}{n^2} \sum_{i=1}^{P} n_i^2 \hat{V}[\bar{X}_i(n_i)],$$

where n_i, $n_i \geq 0$, is the number of observations from engine i used by the central analyser at a given checkpoint of its sequential analysis, and $n = \sum_{i=1}^{P} n_i$ is the total number of observations available from all engines at that checkpoint. Then, having used these estimates for calculating statistical error of results at the given checkpoint, it undertakes a decision about continuation or stopping of the simulation.

Like the original SA/HW, SA/HW in its MRIP version can be also applied to sequences of batch means, instead of to individual observations. Thus, by selecting appropriate batch size, one can reduce storage costs at each simulation engine and communication costs with the central controlling process.

Note that the pooled estimates are calculated from a set of independent sub-sequences of (correlated) observations generated by multiple simulation engines, each of which runs different, statistically independent replication of the same simulation. This results in linear increase of degrees of freedom in Student t-statistics used for obtaining the confidence interval of the pooled mean. With P participating simulation engines, such a statistic will have $df_{MRIP} = 7P$ or $16P$ degrees of freedom, depending on the assumed K and d. Thus, one could expect that the quality of the final results obtained by applying SA/HW in its MRIP version should be better than that of SA/HW in its original version (with just one simulation engine): the more degrees of freedom the more stable confidence intervals are produced. To check this supposition, let us consider the results of coverage analysis of the final results from MRIP SA/HW.

3 PERFORMANCE EVALUATION

Coverage analysis is widely used for assessing the quality of different methods used for constructing confidence intervals on the basis of simulation output data. By performing a large number of experiments we estimate the fraction of the generated confidence intervals which actually contain the true value of the parameter. If the method is accurate then when the theroretical confidence level has been set for example to 95%, this fraction should also be close to 95%.

We performed sequential analysis of coverage, using the methodology presented in (Pawlikowski, Ewing and McNickle 1998), to produce coverage of MRIP SA/HW estimates with a relative precision of 0.01 at 95% confidence level. It is worth noting that for each setting of the parameters of the reference models, getting coverage results with the statistical accuracy required meant that up to 14,000 separate experiments were needed.

Experiments were conducted for a number of reference models. Here we give only the results for an M/M/1 queueing system model with traffic intensities ranging from 0.1 to 0.9. When the degree of the fitting polynomial was fixed, the quadratic fit ($d = 2$) produced the best results when compared with $d = 1$ or $d = 3$.

Figure 2 shows the results obtained for $d = 2$ with $K = 25$ and 50, and a single simulation engine. It can be seen that the coverage obtained agrees well with the required coverage at low to medium traffic intensities, falling off slightly at high intensities. There appears to be little to choose between $K = 25$ and $K = 50$; the latter perhaps giving a small improvement in coverage at high traffic intensities.

Figure 3 compares the results obtained from a single simulation engine to those from $P = 2$ or 4 simulation engines. It can be clearly seen that, apart from reduction of simulation time, use of multiple simulation engines leads to better quality of simulation results as measured by the coverage of the final confidence intervals.

4 CONCLUSIONS

The method of SA/HW, in its MRIP version implemented in Akaroa2, has been found experimentally to produce coverage values which agree very well with those expected. Recent work by the authors suggests that further improvements in coverage of MRIP SA/HW can be obtained by dynamically selecting the value of d (the degree of the polynomial for regression fit) at run time. Another important issue is to find an implementation of MRIP SA/HW that could offer not only a good coverage of the final results but also speedup close to the value theoretically achievable according to the Truncated Amdahl Law. For that purpose, one needs to look at the granularity of sequential data analysis at individual simulation engines. At the same time, locations of the first checkpoints should be carefully selected, to insure that a quickly finished simulation still allows simulation engines to produce valid local estimates.

In addition, the authors continue their investigations of other methods of simulation output analysis for increasing functionality of Akaroa2.

REFERENCES

Ewing, G., Pawlikowski, K. and D. McNickle. 1999. “Akaroa2: Exploiting Network Computing by Distributing Stochastic Simulation”. *Proc. European Simulation Multiconference ESM’99,* Warsaw. International Society for Computer Simulation. pp. 175-181

Heidelberger, P. and P. D. Welch. 1981. “A Spectral Method for Confidence Interval Generation and Run Length Control in Simulations.” *Communications of the ACM,* vol. 24, no. 4 (April), pp. 233-245

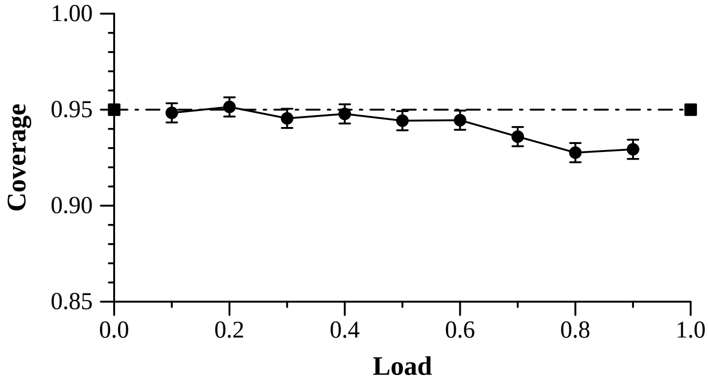
Heidelberger, P. and P. D. Welch. 1983. “ Simulation Run Length Control in the Presence of an Initial Transient”. *Operations Res.,* vol. 31, pp. 1109-1144

Law, A. M. 1983. “Statistical Analysis of Simulation Output Data”. *Operations Res.,* vol. 31, no. 6, pp. 983-1029

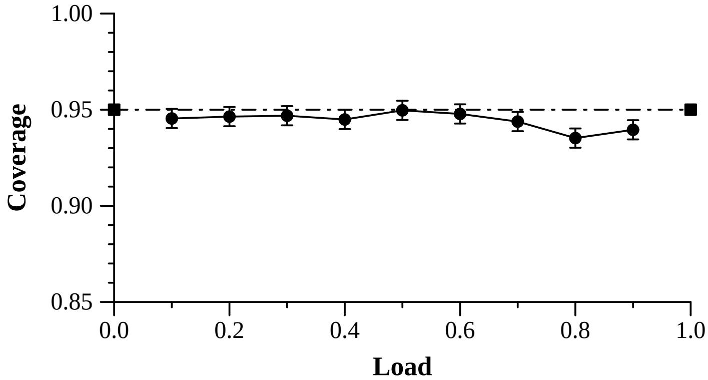
Law, A. M. and W. D. Kelton. 1992. *Simulation Modeling and Analysis.* McGraw-Hill, NY

K. Pawlikowski. 1990. “Steady State Simulation of Queueing Processes: A Survey of Problems and Solutions”. *ACM Computing Surveys*, vol.22, no. 2, June 1990, 123-170

K. Pawlikowski, G. Ewing and D. McNickle. 1998. “Coverage of Confidence Intervals in Sequential Steady-State Simulation”. *J. Simulation Practise and Theory*, vol. 6, no. 3, 1998, 255-267



(a) $K = 25$



(b) $K = 50$

Figure 2: Coverage obtained from M/M/1 queueing model running on a single simulation engine.

Pawlikowski, K. and D. McNickle. 2001. "Speeding Up Stochastic Discrete-Event Simulation". *Proc. 12th European Simulation Symposium, ESS'01*, Marseilles, France. International Society for Computer Simulation

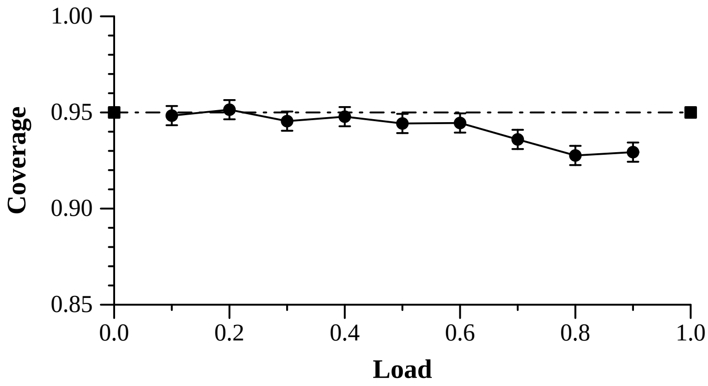
Pawlikowski, K., Yau, V. and D. McNickle. 1994. "Distributed Stochastic Discrete-Event Simulation in Parallel Times Streams". *Proc. Winter Simulation Conference WSC'94*. IEEE Press. pp. 723-730

AUTHOR BIOGRAPHIES

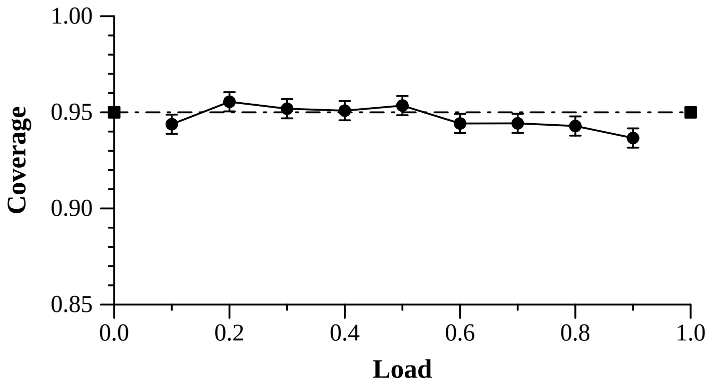
GREG EWING is a research associate in the Department of Computer Science at Canterbury, where received a Ph.D. His research interests include simulation, distributed systems, programming languages, 3D graphics and graphical user interfaces. He has made contributions to the Python programming language, and has recently been nominated for membership of the Python Software Foundation. His email address is <greg@cosc.canterbury.ac.nz> and his web page is <http://www.cosc.canterbury.ac.nz/~greg/>.

DONALD MCNICKLE is a Senior Lecturer in Management Science in the Deparmtent of Management at the University of Canterbury. His research interests include queueing theory, networks of queues and statistical aspects of stochastic simulation. He is a member of INFORMS. His email address is <d.mcnickle@mang.canterbury.ac.nz>.

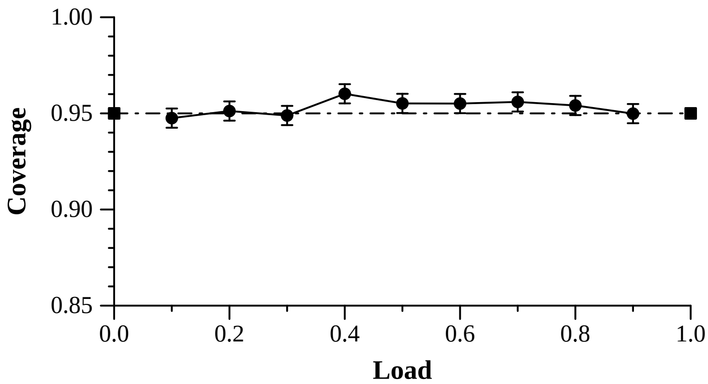
KRZYSZTOF PAWLIKOWSKI is an Associate Professor of Computer Science at the University of Canterbury. His research interests include quantitative stochastic simulation, and performance modelling of telecommunication networks. He received a PhD in Computer Engineering from the Technical University of Gdansk, Poland in 1975 He is a Senior Member of IEEE. His email address is <krys@cosc.canterbury.ac.nz> and his web page is <http://www.cosc.canterbury.ac.nz/~krys/>.



(a) 1 engine



(b) 2 engines



(c) 4 engines

Figure 3: Comparison of coverage obtained using $K = 25$ and varying numbers of simulation engines.