분산 시뮬레이션에서의 Coverage 분석에 관한 연구

이 종 숙† · 박 형 우†† · 정 해 덕†††

요 약

본 논문에서는 분산 시뮬레이션 기법 중에 하나인 MRIP(Multiple Replications In Parallel) 시나리오에서 각종 순차적인 시뮬레이션 분석 방법들의 성능을 측정할 수 있는 포함범위(Coverage)에 대한 신뢰구간(confidence intervals) 및 속도향상(Speedup)에 대해 살펴보았다. F-분포를 기반으로 한 신뢰구간에 대한 추정기(estimator)를 단일 프로세서와 다중 프로세서 상에서 참조모델(reference model)로 $M/M/1/\infty$, $M/D/1/\infty$과 $M/H_2/1/\infty$ 큐잉 시스템을 활용하여 정상상태(steady-state)에서의 평균치를 추정하는 시뮬레이션에 적용하였다. 순차적인 포함범위 분석을 위해서는 수많은 시뮬레이션 실행(Run)들이 요구되는데, MRIP 분산 시뮬레이션 시나리오에서 다중 프로세서를 이용하여 시뮬레이션을 수행하여 최종 시뮬레이션 결과를 얻는데 걸리는 시간을 감소시켰다. 또한, LAN으로 연결된 분산 컴퓨팅 시스템에 시뮬레이션을 동시에 수행시킴으로써 쉽게 필요한 수의 시뮬레이션 실행결과(Run)를 수집할 수 있다. 이는 샘플의 수가 증가됨으로써 좀더 신뢰도가 높은 최종 신뢰구간을 시뮬레이션 수행자가 얻을 수 있게 해준다.

Quality of Coverage Analysis on Distributed Stochastic Steady-State Simulations

JongSuk R. Lee† · HyoungWoo Park†† · HaeDuck J. Jeong†††

ABSTRACT

In this paper we study the quality of sequential coverage analysis under a scenario of distributed stochastic simulation known as MRIP (Multiple Replications In Parallel) in terms of the confidence intervals of coverage and the speedup. The estimator based on the F-distribution was applied to the sequential coverage analysis of steady-state means, in simulations of the $M/M/1/\infty$, $M/D/1/\infty$, and $M/H_2/1/\infty$ queueing systems on a single processor and multiple processors. By using multiple processors under the MRIP scenario, the time for collecting many replications needed in sequential coverage analysis is reduced. One can also easily collect more replications by executing it in distributed computers or clusters linked by a local area network.

키워드 : 순차적 포함범위(Sequential Coverage Analysis), 비율(Proportions), 신뢰구간(Confidence Intervals), 분산 시뮬레이션(Distributed Simulation), 정상상태 시뮬레이션(Steady-state Simulation), 속도향상(Speedup)

1. Introduction

Statistical analysis of output data of stochastic steady-tate simulation is made difficult by the degree of serial correlation often presents in the output. Methods such as batch means, regenerative cycles, and spectral analysis are used to overcome the above problem. An important measure of the robustness of any output analysis method is the coverage of the final confidence intervals defined as the proportion of confidence intervals which contain the true value. Any good method of analysis of simulation output data should produce narrow and stable confidence intervals, and the probability of such an interval containing the true value of the estimated performance measure should be close to the assumed confidence level.

Some interesting results have been achieved in theoretical studies in terms of coverage error for confidence intervals arising in simulation output data analysis [4]. A coverage function (which is defined for all confidence levels between zero and one) to measure robustness of confidence intervals has been proposed [20], and coverage properties of confidence intervals based on the average Bayesian *posterior* probability have been studied [18]. Nevertheless, experimental analysis of coverage is still required to assess the quality of practical implementations of methods used for determining the final confidence intervals, especially in the context

† 정 회 원 : 한국과학기술정보연구원 슈퍼컴퓨팅연구실 선임 연구원
†† 정 회 원 : 한국과학기술정보연구원 슈퍼컴퓨팅센터 슈퍼컴퓨팅연구실장/책임 연구원
††† 정 회 원 : Univ. of Canterbury, NZ 연구원
논문접수 : 2002년 8월 1일, 심사완료 : 2002년 10월 10일

of stochastic steady-state simulation.

The conventional interval estimator based on the normal approximation has been widely used in experimental coverage analysis [7, 10, 14-16, 20, 21]. However, alternative, more efficient and accurate interval estimator based on the F-distribution of proportions is pointed out for analysis of proportions in sequential steady-state simulations in [13]. A comparative study of properties of three interval estimators (based on the normal distribution, the *arcsin* transformation, and the F-distribution) can be found in [13].

There are a number of related facts with coverage analysis. First, it is naturally limited to analytically tractable systems only, 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 silation studies [3]. 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.

Another fact is that coverage analysis requires execution of multiple, independent replications of simulations. Very large numbers of replications are often needed for determining coverage with satisfactory precision. Traditionally, coverage analysis was performed with a fixed number of replications, ranging between 50~500 replications [8-10, 19]. However, newer results of coverage analysis ([11, 12], and [16]) clearly show existence of high initial instability of coverage in the region of 50~500 replications for three different methods of mean value analysis : non-overlapping batch means and SA/HW (spectral analysis in its version proposed by Heidelberger and Welch [6]), and regenerative cycles. To avoid taking the final result from this region, coverage analysis has to be done over a sufficiently larger sample of data or sequentially as recommended in [6]. In any case, one needs to estimate the proportion of confidence intervals covering the theoretical value of interest.

Sequential simulation analysis, however, raises its own problems [16]. One problem is that some of the simulation experiments may stop after an abnormally short run, when the stopping criterion for the sequential simulation is temporarily satisfied. Addressing this issue, some rules for the sequential analysis of coverage have been formulated in [16]. As in the case of ordinary sequential simulation, sequential coverage analysis is continued until the final result is obtained with the required statistical error. Thus, the properties of interval estimators of proportions used for determining precision of coverage play a crucial role in the sequential coverage analysis.

The other problem is that a sequential simulation of even moderately complex simulation models in engineering and computer science is often computationally intensive and requires long runs in order to obtain the final results at a desired level of the statistical error. For instance, in sequential steady-state simulations of an $M/M/1/\infty$ queueing system with traffic intensities of $\rho = 0.99$ and $\rho = 0.999$, the estimations of the mean response time require roughly 8.3 minutes and 1.3 days on a Pentium II with 350Mhz, to achieve an estimate with the relative statistical error of at least 5% respectively [13]. For an open queueing system with traffic intensities of $\rho = 0.99$ and $\rho = 0.999$, the required times to get the steady-state mean response time are approximately 3 hours and 7.3 days respectively [13]. The obvious solution is to speedup such a sequential simulation by executing it in distributed computer systems, possibly using computers or clusters linked by a local area network.

In this paper, we discuss the interval estimator based on the F-distribution in the context of its application on sequential coverage analysis of the SA/HW method in sequential steady-state simulations of the $M/M/1/\infty$, $M/D/1/\infty$, and $M/H_2/1/\infty$ queueing systems. To see whether the interval estimator based on the F-distribution works in the case of very time consuming simulation experiments or not, we have executed all our experiments under a scenario of distributed stochastic simulation known as MRIP (Multiple Replications In Parallel) [17], in which multiple workstations within a local area network work as independent simulation engines, producing data for global output data analysers. The quality of sequential coverage analysis for the SA/HW method executing it with different numbers of workstations is also discussed in terms of the confidence intervals of coverage and the speedup.

2. Interval Estimator for Proportions

Binomial experiments consist of repeated trials, each with two possible outcomes, which may be labelled *success* or *failure*. The point estimator of the proportion p in a binomial experiment is simply given by the statistic

$$\hat{p} = \frac{\textit{count of } \text{successes in } \textit{sample}}{\textit{size of sample}} = \frac{X}{n} \quad (1)$$

If a binomial experiment can result in a success with probability p and a failure with probability $(1-p)$, then the probability distribution of the binomial random variable X, the number of successes in n independent experiments, is

$$b(x\,;\,n,p) = \binom{n}{x} p^x (1-p)^{n-x}, x=0,1,\cdots,n \quad (2)$$

The accuracy with which it estimates an unknown proportion p can be assessed by the width of its confidence interval at a given confidence level, i.e, by the probability

$$\Pr(\hat{p} - \triangle_1 \le p \le \hat{p} + \triangle_2) = 1-\alpha$$

where $\hat{p}$ is the estimate of the proportion p, $\triangle_1$ and $\triangle_2$ are the offset for the lower and the upper limit of the confidence interval of p, and $(1-\alpha)$ is the confidence level, $0 < \alpha < 1$. Ideally, this would mean that if the simulation experiment is repeated many times, the resulting confidence intervals would contain the parameter p in $100\times(1-\alpha)$% of cases. To determine $\triangle_1$ and $\triangle_2$ we need the exact distribution of $\hat{p}$, or at least $Var(\hat{p})$. In practice, only some approximations of these are possible.

The interval estimator of the proportion p, based on the F-distribution, is described in following ways. Confidence intervals for proportions can be formulated from the relationship between F and binomial distributions with the incomplete and the complete beta functions. The ratio of two successive terms in a binomial distribution $b(x\,;\,n,\,p)$ is

$$\frac{b(x+1\,;\,n,p)}{b(x\,;\,n,p)} = \left(\frac{n-x}{x+1}\right)\left(\frac{p}{1-p}\right),\ x=0,1,\cdots,n-1,$$

where x is the observed number of successes in the sample ; see Equation (1). Using the transformations shown, for example, in [1] and [5], the quantiles of the binomial distribution can be obtained from those of the F-distribution, as

$$\Pr\left\{F(r_1,r_2) < \left(\frac{n-n\hat{p}}{n\hat{p}+1}\right)\left(\frac{p}{1-p}\right)\right\} = \Pr\left\{\frac{(n\hat{p}+1)F(r_1,r_2)}{(n-n\hat{p})+(n\hat{p}+1)F(r_1,r_2)} < p\right\}$$

where $F(r_1,r_2)$ is a random variable with the F-distribution of $r_1 = 2\times(n\times\hat{p}+1)$ and $r_2 = 2\times(n-n\times\hat{p})$ degrees of freedom. Thus, an $100\times(1-\alpha)$% confidence interval for a proportion is given by $(\hat{p}_l, \hat{p}_u)$, where

$$\hat{p}_u = \frac{(n\hat{p}+1)f_{1-\alpha/2}(r_1,r_2)}{(n-n\hat{p})+(n\hat{p}+1)f_{1-\alpha/2}(r_1,r_2)}$$

and

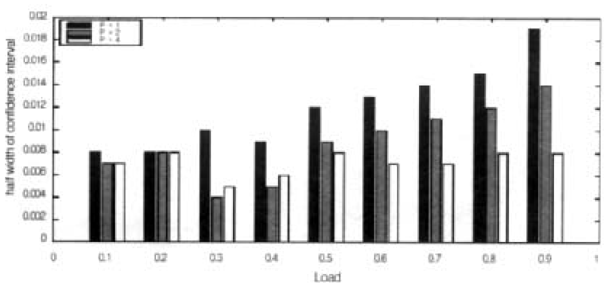
$$\hat{p}_l = \frac{n\hat{p}}{n\hat{p} + (n-n\hat{p}+1)f_{1-\alpha/2}(r_3,r_4)}$$

Here, n is the sample size, and $f_{1-\alpha/2}(r_1,r_2)$ is the $(1-\alpha/2)$ quantile of the F-distribution with (r_1,r_2) degrees of freedom, where $r_1=2\times(n\times\hat{p}+1)$ and $r_2=2\times(n-n\times\hat{p})$, while $f_{1-\alpha/2}(r_3,r_4)$ is the $(1-\alpha/2)$ quantile of the F-distribution with (r_3,r_4) degrees of freedom, where $r_3=2\times(n-n\times\hat{p}+1)$ and $r_4=2\times n\times\hat{p}$ [5].

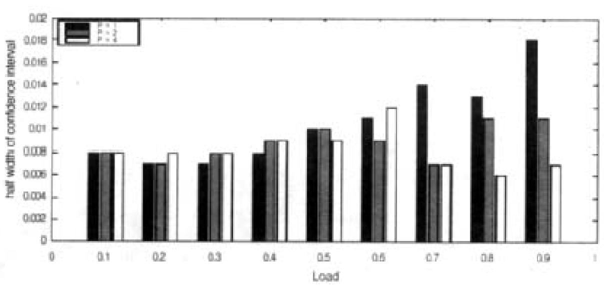
3. Numerical Results

The results of half-width of confidence interval of coverage and the speedup, which are obtained during evaluation of the SA/HW method proposed for sequential estimation of steady-state means, are reported in this section. The SA/HW method has proved to be the most satisfactory method for confidence interval estimation in sequential simulation [16]. All simulation runs were executed using of Akaroa-2, which is a fully automated simulation tool designed for running distributed stochastic simulations under the MRIP scenario, in which multiple processors operate as multiple simulation engines generating independent sequences of output data and submitting them to a global data analyser for analysis [2]. In the case of sequential steady-state simulation, Akaroa-2 automatically detects the length of initial transient period and determines the location of, and analyses the collected simulation output data at each consecutive checkpoint [2].

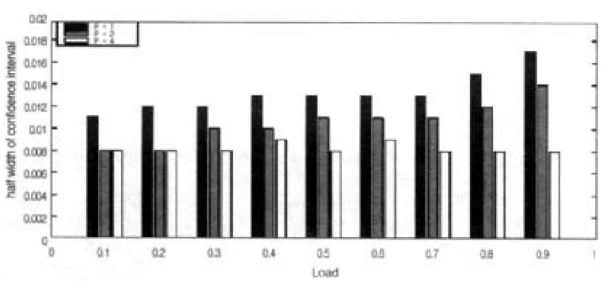
Properties of the SA/HW method were investigated both in the case of ordinary sequential steady-state simulation on a single processor and in the case of distributed simulation under the MRIP scenario. All numerical results in this paper were obtained by stopping the simulation experiments when the final steady-state results for the mean response time have reached the required relative statistical error of 5% or less, at the 0.95 confidence level. By following the proposed rules for experimental coverage analysis in [16], all coverage results were filtered of unusually short simulation runs by discarding runs which are shorter than the average of run-lengths obtained from all collected replications minus the standard deviation of run-lengths obtained from all replications. This guards against the influence of unusually short runs on experimental results. These steps are taken to ensure that the results are typical of what would be considered to be a well-managed simulation experiment. Also at least 200 confidence intervals not covering the theoretical value in sequential coverage analysis were collected. This number of observed 'bad' confidence intervals has been recommended in [16], for ensuring representativeness in the analysed data.



(a) Simulation of $M/M/1/\infty$ queueing system



(b) Simulation of $M/D/1/\infty$ queueing system



(c) Simulation of $M/H_2/1/\infty$ queueing system

(Figure 1) Half-width of confidence interval of coverage for the SA/HW method using the interval estimator based on the F-distribution (Simulation using P = 1, 2, and 4 processors under the MRIP scenario)

The influence of parallelisation of simulation in the MRIP scenario on the quality of the final results is depicted in (Figure 1). Numerical results of half-width of confidence interval of coverage are obtained from sequential coverage analysis of the SA/HW method, when estimating the mean response time using the interval estimator based on the F-distribution, in simulations of the $M/M/1/\infty$,

$M/D/1/\infty$, and $M/H_2/1/\infty$ queueing systems on a single processor (P = 1) and multiple processors (P = 2 and P = 4). For three different simulated models, as increasing the number of processors the half-width of confidence intervals at heavier traffic intensities is clearly decreasing. For example, the half-width of confidence-intervals of coverage at the traffic intensity of ρ = 0.9 has more than halved as the engaged processor increased 1 to 4. For the other interval estimators of proportion based on the normal distribution and the arcsin transformation, we also observe the very similar behaviour such as (Figure 1).

The total number of collected replications in sequential coverage analysis of the SA/HW method on a single processor (P = 1) and multiple processors (P = 2 and P = 4), when estimating the mean response time with the required relative statistical error of 5% or less, at the 0.95 confidence level, is presented in <Table 1>. <Table 1> shows that sequential coverage analysis executed using more processors collects more numbers of replications.

<Table 1> The number of replications collected in sequential coverage analysis of the SA/HW method on a single processor (P = 1) and multiple processors (P = 2 and P = 4) (When estimating the mean response time with the required relative statistical error of 5% or less, at the 0.95 confidence level)

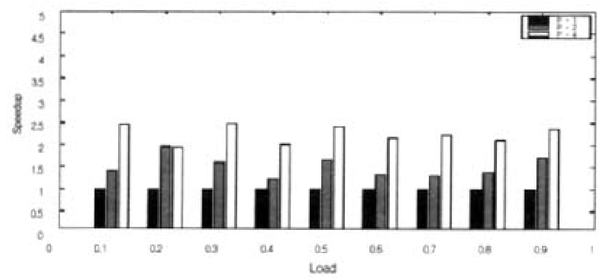
Load	$M/M/1/\infty$			$M/D/1/\infty$			$M/H_2/1/\infty$		
	P=1	P=2	P=4	P=1	P=2	P=4	P=1	P=2	P=4
0.1	3076	4381	5002	3857	4231	3936	2888	3725	4179
0.2	3439	3501	7052	4106	3718	3737	2560	3896	4131
0.3	3463	4317	5586	3450	3334	3373	2479	3056	3860
0.4	2398	3923	4775	2807	3124	2368	2431	3266	3684
0.5	2718	3248	4533	3321	5240	5615	2311	2822	4087
0.6	2403	3599	4440	3022	3539	4811	2302	2905	3601
0.7	2306	3538	4154	2385	3110	5085	2319	2806	4010
0.8	2135	3098	4051	2294	3670	4211	1961	2550	4070
0.9	2191	2545	3732	1893	2604	3638	1802	2149	3671

This suggests that conducting stochastic simulation in parallel on multiple processors under the MRIP scenario usually leads to a reduction of half-width of final confidence intervals of three considered queueing systems by increasing the sample size (say, the number of collected replications) ; see (Figure 1).

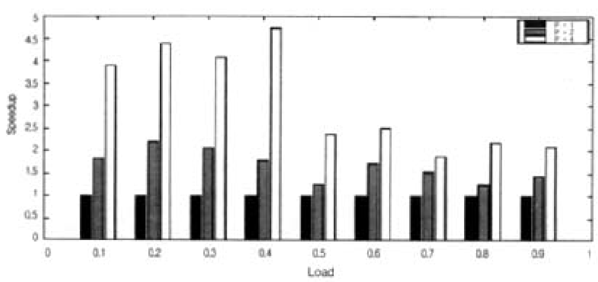
The usual measurement of effectiveness of parallel computations is the speedup. The speedups obtained when estimating the confidence interval of coverage of the SA/HW method using the interval estimator based on the F-distribution and using P = 1, 2, and 4 processors under the MRIP scenario are presented in (Figure 2). These speedups are calculated from the data set of <Table 1> by

$$Speedup = \frac{\text{the number of replication collected using a single processor}}{\text{the number of replications collected using } P \text{ processor } / \ P}$$

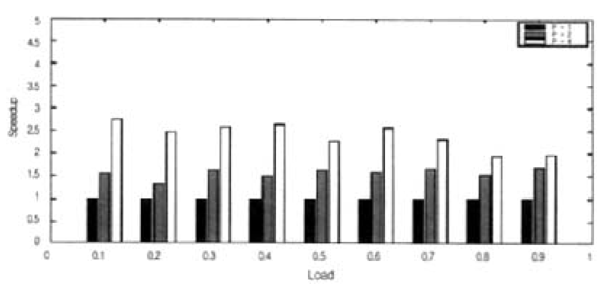
where P is the number of processors used. This result shows that the time for collecting many replications in sequential coverage analysis under the MRIP scenario is clearly reduced. The average speedup was increased about 2.5 times as the engaged processor increased 1 to 4. Using multiple processors under the MRIP scenario especially in sequential coverage analysis, one can easily collect more replications which can help to produce the better final confidence intervals.



(a) Simulation of $M/M/1/\infty$ queueing system



(b) Simulation of $M/D/1/\infty$ queueing system



(c) Simulation of $M/H_2/1/\infty$ queueing system

(Figure 2) Speedup obtained when estimating the confidence interval of coverage of the SA/HW method using the interval estimator based on the F-distribution (Simulation using P = 1, 2, and 4 processors under the MRIP scenario)

4. Conclusions

While some interesting results have been achieved in theoretical studies of coverage analysis of sequential interval estimators used in simulation output data analysis, experimental analysis of coverage is still required for assessing the quality of the practical implementations of methods used for determining confidence intervals in sequential stochastic simulation. In this paper we study the interval estimator of proportions, in the context of their applications in sequential coverage analysis under a scenario of parallel and distributed stochastic simulation known as MRIP (Multiple Replications In Parallel). The estimator based on the F-distribution was applied to sequential coverage analysis of the SA/HW method to estimate steady-state means in simulations of the $M/M/1/\infty$, $M/D/1/\infty$, and $M/H_2/1/\infty$ queueing systems on a single processor (P = 1) and multiple processors (P = 2 and P = 4).

For those three different simulated models, as increasing the number of processors the half-width of confidence intervals at heavier traffic intensities is clearly decreasing (more than 50%). This is caused by the fact that conducting stochastic simulation in parallel on multiple processors under the MRIP scenario leads to increase the sample size. In the case of sequential coverage analysis, we easily collected more replications as using multiple processors under the MRIP scenario. The average speedup was increased about 2.5 times as the engaged processor increased 1 to 4. Therefore, the time for collecting many replications needed in sequential coverage analysis is clearly reduced.

References

[1] M. Abramowitz and I. A. Stegun, "Handbook of Mathematical Functions," Government Printing Office. New York : Dover Publications, Inc., 1964.

[2] G. C. Ewing, K. Pawlikowski and D. McNickle, "Akaroa-2 : Exploiting Network Computing by Distributing Stochastic Simulation," 13th European Simulation Multiconference, Warsaw, Poland, pp.175-181, 1999.

[3] B. Fox, "Estimation and Simulation," Management Science. Vol.24, pp.860-861, 1978.

[4] P. W. Glynn, "Coverage Error for Confidence Intervals Arising in Simulation Output Analysis," Proceedings of the

1982 Winter Simulation Conference, pp.369-375, 1982.

[5] A. Hald, "Statistical Theory with Engineering Applications," John Wiley and Sons, Inc., 1952.

[6] P. Heidelberger and P. D. Welch, "A Spectral Method for Confidence Interval Generation and RunLength Control in Simulations," Communications of the ACM. Vol.25, pp.233-245, 1981.

[7] J. Judge, H. W. P. Beadle and J. Chicharo, "Correlation of HTTP Response Packet Size and Estimating Confidence intervals for Mean Packet Size and WWW Traffic Volume," Proceedings of APCC 1997, Sydney, Australia, pp.382-386, 1997.

[8] S. S. Lavenberg and C. H. Sauer, "Sequential Stopping Rules for the Regenerative Method of Simulation," IBM Journal of Research Development, pp.545-558, 1977.

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

[10] A. M. Law and W. D. Kelton, "Simulation Modeling and Analysis," McGraw-Hill, Inc., 2nd Ed., 1991.

[11] J. R. Lee and D. McNickle and K. Pawlikowski, "Quality of Sequential Regenerative Simulation," 13th European Simulation Multiconference, Warsaw, Poland, pp.161-167, 1999.

[12] J. R. Lee, K. Pawlikowski and D. McNickle, "Do Not Trust Too Short Sequential Simulation," Proceedings of the 1999 Summer Computer Simulation Conference, Chicago, Illinois, pp.97-102, 1999.

[13] Jong-Suk Ruth Lee, "On Automated Sequential Steady-State Simulation," Ph. D. Dissertation, Department of Computer Science, University of Canterbury, New Zealand, 2001.

[14] D. C. McNickle, K. Pawlikowski and G. Ewing, "Experimental Evaluation of Confidence Interval Procedures in Sequential Steady-State Simulation," Proceedings of the 1996 Winter Simulation Conference, pp.382-389, 1996.

[15] D. S. Moore and G. P. McCabe, "Introduction to the Practice of Statistics," W. H. Freeman and Company, New York, 2nd Ed., 1996.

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

[17] 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, Lake Buena Vista, Florida, pp.723-730, 1994.

[18] D. B. Rubin and N. Schenker, "Efficiently Simulating the Coverage Properties of Interval Estimates," Applied Statistics. Vol.35 (2), pp.159-167, 1986.

[19] C. H. Sauer, "Confidence Intervals for Queueing Simulations of Computer Systems," ACM Performance Evaluation Review. Vol.8 (1-2), pp.46-55, 1979.

[20] L. W. Schruben, "A Coverage Function for Interval Estimators of Simulation Response," Management Science. Vol.26, pp.18-27, 1980.

[21] R. E. Walpole and R. H. Myers, "Probability and Statistics for Engineers and Scientists," Macmillan Publishing Co., Inc., New York, 2nd Ed., 1978.

이 종 숙

e-mail : jsruthlee@hpcnet.ne.kr

1992년 충남대학교대학원 전산학과(이학석사)

2001년 Dept. of Computer Science, University of Canterbury, NZ (전산학 박사)

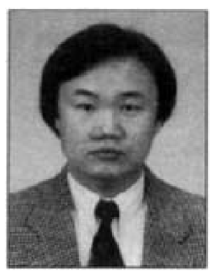
1992년~1993년 한국전자통신연구원(연구원)

1999년~2002년 Univ. of Canterbury, NZ(연구원)

2002년~현재 한국과학기술정보연구원 슈퍼컴퓨팅연구실(선임연구원)

현 수행과제 : 그리드 미들웨어 연구사업과 국가 그리드기발 구축사업

관심분야 : 그리드 컴퓨팅, 그리드 미들웨어, 병렬/분산 컴퓨팅, 분산/병렬 시뮬레이션



박 형 우

e-mail : hwpark@hpcnet.ne.kr

1985년 서울시립대 전자공학과(공학사)

1996년 성균관대 산업과학대학원 전자공학과(공학석사)

2000년 성균관대 대학원 전기전자컴퓨터공학과(공학박사)

2002년~현재 KISTI 슈퍼컴퓨팅센터 슈퍼컴퓨팅연구실장/책임연구원

현 수행과제 : 국가 그리드기반 구축사업과 그리드 미들웨어 연구사업

관심분야 : Grid 컴퓨팅, 차세대 TCP, 고성능 인터넷 관리 및 보안

정 해 덕

e-mail : joshua@cosc.canterbury.ac.nz

2002년 Dept. of Computer Science, University of Canterbury, NZ(전산학 박사)

2002년~현재 Univ. of Canterbury, NZ(연구원)

관심분야 : teletraffic modeling in telecommunication, networks and stochastic simulation