THE APPLICATION OF CONTROL VARIABLES TO THE SIMULATION OF CLOSED QUEUEING NETWORKS

S.S Lavenberg, T.L. Moeller and P.D. Welch

IBM Thomas J. Watson Research Center
Yorktown Heights, New York 10598

ABSTRACT

Control variables are presented which reduce the variance of estimators of steady state response characteristics in the simulation of a broad class of closed queueing networks. Statistical aspects of applying control variables in this context are discussed, and results of extensive empirical studies, conducted to evaluate the control variables, are summarized.

INTRODUCTION

Control variables are random variables whose expectations are known and which are correlated with an estimator of interest. They can be combined with the estimator to reduce its variance while not changing its expectation. Books containing a discussion of control variables include [1] and [2]. Here we consider the application of control variables to the simulation of a class of closed queueing networks. Statistical aspects of applying control variables in this context are discussed.

CLASS OF NETWORKS

The networks consist of a finite number of interconnected service centers and are closed, serving a fixed finite number of customers. Each service center is a single server or multiple server queue. The service times are statistically independent random variables, and the route followed by a customer is a finite state Markov chain. Multiple types of customers (different types of customers can have different service time distributions, different Markov routing chains and different priorities), blocking due to capacity constraints and priority queueing disciplines are allowed. Networks of this type are commonly used to model the contention for resources which occurs in interactive multiprogrammed computer systems [5].

CONTROL VARIABLES FOR THE NETWORKS

The control variables are, for each service center and each type of customer, the sum of the completed service times for that customer type at that service center divided by the number of service completions in the network for that customer type. These control variables, called work variables, were introduced in [3] for networks with only one type of customer. Their asymptotic expectations as the number of service completions for each customer type becomes large are known. Work variables reflect the effects of both the service time and routing random variables which drive the simulation. Thus, we expect them to be correlated with estimators of steady state response characteristics of the network. Note that if there are S service centers and T types of customers which visit every service center, the number of work variables is Q =ST.

STATISTICAL ASPECTS

Let μ be an unknown deterministic quantity to be estimated and let Y be an unbiased estimator of μ. Let $\underset{\sim}{C} = (C_1, \dots C_Q)$ be a column vector of control variables, so that $E[\underset{\sim}{C}]$ is known, and let $\underset{\sim}{a} = (a_1, \dots, a_Q)$ be a column vector of constants. Then the controlled estimator

$$Y(\underset{\sim}{a}) = Y - \underset{\sim}{a}'(\underset{\sim}{C} - E[\underset{\sim}{C}]),$$

where $\underset{\sim}{a}'$ denotes the transpose of $\underset{\sim}{a}$, is also unbiased, and the value of $\underset{\sim}{a}$, call it $\underset{\sim}{a}^*$, which minimizes $\mathrm{Var}[Y(\underset{\sim}{a})]$ can be expressed in terms of the covariance matrix of $\underset{\sim}{C}$ and the covariances between Y and $\underset{\sim}{C}$. In general these covariances are not known, so that $\underset{\sim}{a}^*$ is not known. The ratio of the resulting minimum variance to the variance of Y is given by $1 - [R(Y,\underset{\sim}{C})]^2$, where $R(Y,\underset{\sim}{C})$ is the multiple correlation coefficient between Y and $\underset{\sim}{C}$ (e.g., see [3]). We have empirically investigated several moderate sized queueing networks ($Q \le 12$) and estimated the minimum variance ratio for all work control variables. Estimators of such steady state response characteristics as mean waiting times and service completion rates were considered. Minimum variance ratios were typically below .5.

The minimum variance ratio is achieved when $\underset{\sim}{a}^*$ is known. However, $\underset{\sim}{a}^*$ is unknown and in practice it must be estimated from the data. Suppose we obtain N independent observations of the random vector $(Y,\underset{\sim}{C})$. These observations can be obtained from N independent replications of the simulation, or from a single long simulation using the method of batch means. An estimate of $\underset{\sim}{a}^*$, call it $\hat{\underset{\sim}{a}}^*$, can be constructed from these observations by replacing the covariances in the expression for $\underset{\sim}{a}^*$ by the corresponding sample covariances based on the N observations. Let $\bar{Y}$ denote the sample mean of the Y's, let $\bar{\underset{\sim}{C}}$ denote the sample mean of the $\underset{\sim}{C}$'s and let

$$\bar{Y}(\underset{\sim}{a}) = \bar{Y} - \underset{\sim}{a}'(\bar{\underset{\sim}{C}} - E[\underset{\sim}{C}]).$$

The controlled estimator we use is $\bar{Y}(\hat{\underset{\sim}{a}}^*)$. We have shown [4] that if $(Y,\underset{\sim}{C})$ is non-singular multivariate normal, then

$$\mathrm{Var}[\bar{Y}(\hat{\underset{\sim}{a}}^*)]/\mathrm{Var}[\bar{Y}(\underset{\sim}{a}^*)] = (N-2)/(N-Q-2).$$

Thus, much of the potential variance reduction is lost due to estimating $\underset{\sim}{a}^*$ unless the number of control variables Q is much less than the number of independent observations N. This can be a problem in queueing networks where the number of work variables is large.

In addition to constructing a reduced variance estimator for μ using control variables, it is important in practice to be able to estimate the variance of this estimator and to construct a valid confidence interval for μ which reflects the reduced variance. We have obtained an estimator for $\mathrm{Var}[\bar{Y}(\hat{\underset{\sim}{a}}^*)]$ which we have shown to be unbiased if $(Y,\underset{\sim}{C})$ is non-singular multivariate normal. We also have obtained a confidence interval for μ whose width is proportional to the square root of the variance estimator. We have shown that the confidence interval is valid under the multivariate normal assumption. (See [4] for proofs of these results.)

We have conducted extensive empirical investigations of the application of work control variables to the simulation of closed queueing networks using the estimator $\bar{Y}(\hat{\underset{\sim}{a}}^*)$ and the above variance and confidence interval estimators. We summarize results of these investigations on the variance reduction achieved and on the confidence interval coverage.

BIBLIOGRAPHY

1. Gaver, D.P., and Thompson, G.L., Programming and Probability Models in Operations Research, Brooks/Cole Publishing Co., Monterey, California, 1973.
2. Kleijnen, J.P.C., Statistical Techniques in Simulation Part I, Marcel Dekker, Inc., New York, 1974.
3. Lavenberg, S.S., Moeller, T.L., and Welch, P.D., "Control Variables Applied to the Simulation of Queueing Models of Computer Systems," Proceedings of the International Symposium on Computer Performance Modeling, Measurement and Evaluation, Edited by K.M. Chandy and M. Reiser, North Holland Publishing Co., New York, 1977.
4. Lavenberg, S.S., Moeller, T.L., and Welch, P.D., "Control Variables Applied to the Simulation of Closed Queueing Networks," to appear as an IBM Research Report, 1977.
5. Muntz, R.R., "Analytic Modeling of Interactive Systems," Proc. IEEE 63, 946-953 (1975).