DES AS A REAL TIME DECISION MAKING TOOL. AN APPLICATION TO FIRE SERVICE EMERGENCY COVER

Russell Cheng

Highfield
School of Mathematics
University of Southampton
Southampton, SO17 1BJ, UK

ABSTRACT

Suppose that we have to make a decision in real time where we have to choose between a number of alternative ways of operating a system. A performance measure is available for comparing the alternatives. We also have a simulation model and can make runs of this to estimate the performance measure, but not all alternatives can be examined because there is a strict limit to the time available for simulations. A balance is needed between making a few long runs in which the performance measure is accurately estimated but only for a few alternatives, and making a large number of short runs but with the performance measure poorly estimated. We analyse how the run length should be selected to ensure that an alternative with a good performance measure will be found with high probability. We give a real example arising in the provision of fire service emergency cover.

1 INTRODUCTION

We consider the use of discrete event simulation in situations where quick decision making is needed for a combinatorially hard problem. Such decisions may need to be taken frequently, so their number can, in time, be large. But each decision is a 'one-off' reaction to a random event, and so each can be studied independently. The simulation model has to be discrete-event based to properly capture the logical complexity of the behaviour of the system. Simulation run-time is therefore likely to be long and be a serious issue.

An example of such a situation occurs in air traffic management. The Eurocontrol Central Flow Management Unit (EUROCONTROL CFMU, 2002) manages slot allocation using near real-time simulation. A possible additional application here is the delay-sharing problem discussed by Boesel (2003).

Military engagements (Wrigley and Taylor, 2003) are another area of potential application.

In this article we consider fire service emergency cover. (See also Yang et al. 2004).

In all of these examples, successful system operation requires separate rapid decisions to be taken.

We point out the effectiveness of using simulation as the basis for a decision tool that can be used in problems of this kind. However the main purpose of this article is to discuss a specific problem that such simulation based real-time decision-making gives rise to.

We consider the situation where we wish to compare the consequences of a potentially large number of different decision choices, but where the need to make a rapid decision means that n, the number of runs that can be made of the simulation model, has to be kept small.

The statistical issues have not really been fully considered from this viewpoint of simulation real-time decision-making before. In the literature, the application of ranking and selection to guarantee the quality of good solutions in stochastic optimization as discussed by Pichitlamken and Nelson (2001), Boesel et al. (2003), and Kim and Nelsom (2006), is close to the kind of situation we have in mind.

However our problem is possibly more related to the optimal computer budget allocation problem (OCBA) discussed by Chen (2002) but with the need to decide on the fraction of possible alternatives that should be examined.

In the next section we describe an example in more detail, and in Section 3 we formulate a statistical model for studying the problem. Section 4 gives examples of typical calculations using the model.

2 AN EXAMPLE

2.1 Fire Service Emergency Cover

We describe a genuine example of the problem involving the construction of a DES model of fire service emergency cover (FSEC) provided by regional fire brigades in the UK. The work involves the active participation and financial support of a UK Government department, formerly the

Office of the Deputy Prime Minister (ODPM), but recently reconstituted as the Department of Communities and Local Government (DCLG).

Regional Fire Brigades in the UK already possess a very sophisticated tool for gathering and analysing incident data in a very comprehensive way. This information is used for planning and to provide operational statistics to the UK Government. It is realised by brigade management that the data could be used to inform day to day management decisions. Indeed a probabilistic model has previously been developed and is in current use to assess the impact of possible changes in operational policy. This model runs very slowly (typically up to half an hour for one run representing a year) and is used only for planning purposes and not for real-time decision taking.

2.2 The Model

The author of this abstract was commissioned to investigate whether it would be possible to develop a simulation model that would run sufficiently fast to be used not simply for off-line planning purposes, but to evaluate risk in real time. Such a model might then be deployed as an operational tool to provide real-time advice to brigade officers in responding to actual incidents.

2.3 The 'Cover-Moves' Problem

An operational problem of particular interest is the *Cover-Moves Problem*. This occurs when a fire brigade responds to a large incident (one that needs a large, say 8 or more, number of fire appliances to attend). The incident controller then usually repositions a small number of vehicles not involved in the large incident in what are called *cover-moves*, to try to minimize risk in the remainder of the region. Here risk can be clearly defined, either as the expected fatality rate in the region, or as an overall cost that takes into account both costed expected fatalities and brigade operating costs.

The choice of a worthwhile combination of cover-moves is a good example of a problem in stochastic combinatorial optimization. It is not usually possible, certainly in real time, to identify the best solution. The real question is whether a worthwhile operational solution can be found.

The strict (policy driven) operating requirement for the cover-moves problem is that a solution has to be found within *one minute* of the notification of occurrence of a large incident.

We consider the kind of cover-moves solution achievable. In one example of a typical large incident, consideration was given to selecting 3 vehicles for cover-moves out of 16 available vehicles located in 11 stations. The 3 vehicles were to be sent to 3 out of the 6 stations that had supplied vehicles to attend the large incident. A simple combinatorial calculation shows that there are 25800 distinct cover-move combinations possible. In fact for operational reasons only 230 combinations needed to be considered in this situation. Figure 1 shows the EDF of the fatality counts corresponding to all 230 cover-move combinations in this case. Each run simulated an entire year's operation under the conditions of the large incident. The entire set of runs took nearly 10 minutes, so this EDF could not have been produced within the one minute stipulation.

The EDF was calculated using fairly extended run lengths and so might be regarded as providing a fairly reliable estimate of the distribution of the expected fatality rate for the 230 different possible cover-move combinations. If this EDF were available (but not knowing fatality rates corresponding to individual cover-move combinations), it would possible to give a precise assessment of the probability that the *best* of a given number of cover-move combinations randomly selected for examination by simulation, lies in the top 5% (as measured by lowest risk), say, of all the possible combinations.

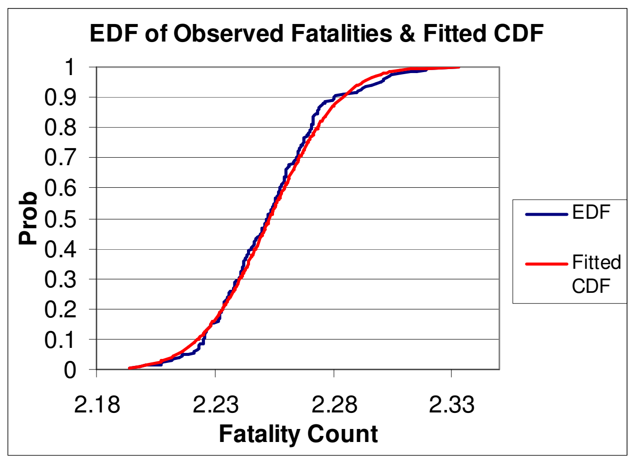

Figure 1: Distribution of Fatalities Count for 230 Cover-Move Choices.

However we are not able to construct this EDF under the real time requirement of doing this within a minute, even assuming that individual runs simulating one year each are long enough. In the next section we give a more precise formulation of the problem that takes into account the uncertainty (because of finite run length) of estimating the expected fatality count, and how this affects the choice of the best run length to use.

3 STATISTICAL MODEL OF THE PROBLEM

We can formulate the simulation run length problem as follows. Let K be the number of alternative decisions under consideration. Let the true performance measure of the system operating under the ith alternative be X_i. We consider the situation where we select alternatives at random. The performance X of a randomly selected alternative is therefore a random variable taking values in the set $\{X_i\}$. Suppose that the total time available for carrying out simulation runs is T and that $n < K$ simulation runs are to be made each of length $t = T/n$.

For the ith simulation run let the *observed* performance measure be

$$Y_i = \mu + X_i + \eta + \varepsilon_i,\ i = 1,2,...,n. \quad (1)$$

All the terms on the right: η, X_i, ε_i, $i = 1,2,...,n$ are assumed to be mutually independent random variables.

We assume that the performance measure, X_i, is a rate quantity (averaged over time) so that it is independent of the run length t. Moreover its variability stems purely from the random process of selecting alternatives to investigate, and *not* from simulation variability.

The term μ is a common mean that it enables us to take

$$E(X_i) = 0. \quad (2)$$

We write the variance of X_i as

$$Var(X_i) = \sigma^2, \quad i = 1,2,...,n. \quad (3)$$

The quantity ε_i is an 'error' term with mean zero and variance

$$Var(\varepsilon_i) = \tau^2/t = n\tau^2/T, \quad i = 1,2,...,n. \quad (4)$$

The term η is also a random variable with mean zero and variance

$$Var[\eta(t)] = \theta^2/t = n\theta^2/T. \quad (5)$$

The model thus assumes a total simulation error of $\eta + \varepsilon_i$ in estimating X, with η interpretable as the error arising from the use of common numbers.

If the simulation error is zero, so that X_i is actually known, then our best choice would be the alternative i_0 corresponding to the smallest X_i. However we do not know the true X_i but instead have the observed performance measures, Y_i. We therefore estimate the best alternative as $\hat{i}_0$, corresponding to the smallest, $Y_{(1)}$, in the expectation that when the errors ε_i are small compared to the X_i, then this will be a good choice. In what follows, for simplicity we write $X_{(1)}$ (rather than $X_{\hat{i}_0}$) for the performance measure corresponding to $Y_{(1)}$.

We write $F_X(\cdot)$, $F_Y(\cdot)$ and $F_\varepsilon(\cdot)$ for the cumulative distribution functions (cdf) of X, Y and ε; and write $f_X(\cdot)$, $f_Y(\cdot)$ and $f_\varepsilon(\cdot)$ for their probability density functions (pdf). We find that

$$\Pr(X_{(1)} < x) = F_{X_{(1)}}(x) =$$
$$1 - n\left(\int_{-\infty}^{\infty}\left\{\int_{-\infty}^{y-x} f_X(y-u) f_\varepsilon(u)[1 - F_Y(y)]^{n-1} du\right\} dy, \quad (6)$$

It should be noted that n enters this formula not only as a multiplicative factor but also through the distributional dependence of ε on n, which has been suppressed for notational simplicity.

Equation (6) allows us to find the probability, P_a, that the selected alternative, $\hat{i}_0$ (corresponding to $Y_{(1)}$), is actually one of the a alternatives with the smallest performance measure. We have

$$P_a = F_{X_{(1)}}(x_{a/K}) \quad (7)$$

where $x_{a/K}$ is the quantile

$$x_{a/K} = F_X^{-1}(a/K) \quad (= X_{(a)}) \quad (8)$$

of the distribution of the performance measure X, when this is treated as a random variable.

The best choice of n is simply that value which maximizes P_a, i.e.

$$n_{\text{opt}} = \arg\max\{F_{X_{(1)}}(x_{a/K})\}. \quad (9)$$

This can be found numerically once we know $F_X(\cdot)$ and $F_\varepsilon(\cdot)$. We consider this next.

4 DETERMINATION OF THE DISTRIBUTIONS OF X, ε AND $X_{(1)}$

The choice of an appropriate form for the distribution of X is an interesting problem in its own right. Assuming that we are interested in performance measure minimization, then we would be particularly interested in the correct form for left tail behaviour. It is possible to give examples of problems where a normal distribution would seem appropriate, and we shall only consider this case here.

The assumption of a normal distribution for that of the error components η and ε in equation (1) seems less contentious, and this will be assumed here.

We shall assume that we can make a set of simulation runs off-line in order to estimate the parameters of the normal distributions of X and ε. Let a set of n runs as given in (1) be called a *trial* and consider a set of m such trials with observations

$$y_{ij} = \mu + x_i + \eta_j + \varepsilon_{ij},\ i = 1,2,...,n,\ j = 1,2,...,m. \quad (10)$$

Simple estimates of the parameters are:

$$\hat{\mu} = (nm)^{-1} \sum_{i=1}^{n} \sum_{j=1}^{m} y_{ij} \ , \tag{11}$$

$$\hat{x}_i = m^{-1} \sum_{j=1}^{m} y_{ij} - \hat{\mu}, \ i = 1,2,...,n \ , \tag{12}$$

$$\hat{\eta}_j = n^{-1} \sum_{i=1}^{n} y_{ij} - \hat{\mu}, \ j = 1,2,...,m \ , \tag{13}$$

$$\hat{\sigma}^2 = n^{-1} \sum_{i=1}^{n} \hat{x}_i^2 \ , \tag{14}$$

$$n\hat{\tau}^2 = (nm)^{-1} \sum_{j=1}^{m} (y_{ij} - \hat{\mu} - \hat{x}_i - \hat{\eta}_j)^2 \ , \tag{15}$$

where, with no loss of generality, we have taken $T = 1$.

Typical estimates from a pilot study of the cover-moves problem gave:

$$\hat{\sigma} = 0.0322 \text{ and } \hat{\tau} = 0.000563 .$$

The distribution of $X_{(1)}$ and the probability P_a can now be calculated. The value of P_a for selected n is given in Table 1, for the cover-moves example with a = 10 (and K = 230). In this case we see that we could actually have made runs sufficiently short to have examined all the 230 alternatives. Note that even when we examine all alternatives, we cannot guarantee selection of one of the best a alternatives because we cannot completely eliminate all simulation experimental error.

Table 1: Value of P_a, calculated from Equation (7), for selected n, when a = 10, K = 230.

n	100	200	500	1000	2000
P_a	0.969	0.991	0.991	0.978	0.932

5 CONCLUSIONS

We have given a preliminary analysis of how simulation runs should be set up when there is a limit to the time available for simulation runs. This problem can arise when simulation is used in real time decision making. We have shown how the model might be used in practice in Section 4, and illustrated it with an application to a real problem arising in the provision of fire service emergency cover.

An issue of especial interest which we have not discussed, is the form that the distribution of the performance measure might take when it is randomly sampled. We have only considered the normal model for this, but there are theoretical reasons why power law distributions may be more appropriate in certain situations. It is hoped to discuss this issue elsewhere.

REFERENCES

Boesel J. (2003) Simulating Aircraft Delay Absorption. Proceedings of the 2003 Winter Simulation Conference. S. Chick, P.J. Sanchez, D. Ferris, and D.J. Morrice, eds, IEEE Piscataway. 1663 – 1669.

Boesel, J, Nelson, BL, and Kim, S-H. (2003). Using ranking and selection to “clean up” after simulation optimization. *Operations Research,* 51, 814-825.

Chen, C-H. (2002). Very Efficient Simulation for Engineering Design Problems with Uncertainty. In Modeling and Simulation-Based Life Cycle Engineering, K. Chnog, S. Saigal and S Thynell, eds) pp 291-302, London: Spon Press.

EUROCONTROL-CFMU (2002). Flow and Capacity Management Unit. http://www.cfmu.eurocontrol.int/cfmu/opsd/public/standard_page/operational_services_atfcm_management.html

Kim, S-H and Nelson, BL (2006). Selecting the Best System. Chap 17. In Simulation, Handbooks in Operations research and Management Science, SG Henderson and BL Nelson, eds. Amsterdam: North-Holland

Pichitlamken, J. and Nelson, B.L. (2001). Selection-of-the-Best Procedures for Optimization via Simulation. In *Proceedings of the 2001 Winter Simulation Conference.* IEEE Piscataway, NJ, 401-407.

Wrigley, D and Taylor, B (2003). SIMBRIG – Simple Brigade Model, Meeting the Need for Fast operational analysis support at the formation level. www.dcmt.cranfield.ac.uk/ismor/ISMOR/2003/wrigley.ppt

Yang, B., Viswanathan, K., Lertoworawanich, P. and Kumar, S. (2004). Fire station districting using simulations: Case study in Centre Region, Pennsylvania. *J. of Urban Planning and Development*, 130, 3, 117-12

AUTHOR BIOGRAPHY

RUSSELL C. H. CHENG is Professor of Operational Research at the University of Southampton. He has an M.A. and the Diploma in Mathematical Statistics from Cambridge University, England. and a Ph.D. from Bath University. He is a former Chairman of the U.K. Simulation Society, a Fellow of the Royal Statistical Society and the British Computer Society, Member of the Operational Research Society. His research interests include: variance reduction methods and parametric estimation methods. He was a Joint Editor of the *IMA Journal of Management Mathematics*. His email and web addresses are <R.C.H.Cheng@maths.soton.ac.uk> and www.maths.soton.ac.uk/staff/Cheng