Proceedings of the 1988 Winter Simulation Conference
M. Abrams, P. Haigh, and J. Comfort (eds.)

Stochastic flow networks: How component criticality changes with component reliability

Christos Alexopoulos
Georgia Institute of Technology
and
George S. Fishman
University of North Carolina at Chapel Hill

ABSTRACT

Consider a directed flow network whose arcs have random two-state capacities. The paper describes how a single Monte Carlo experiment allows one to estimate the sensitivity of the probability that a particular cutset is critical, given that the maximum flow between a pair of nodes takes values in an interval, in response to varying arc reliabilities. The technique improves considerably on alternative potential methods.

THE PROBLEM

Flow networks model several distribution systems, electricity and water supply systems being examples. In practice, several of the parameters of a flow network are random variables and the network is called stochastic. The model under study is a directed network $\mathcal{G} = (\mathcal{N},\mathcal{A})$ with node set $\mathcal{N} = \{1,\dots,n\}$ and arc set $\mathcal{A} = \{1,\dots,a\}$. Assume that the nodes do not restrict the transmission of flow through the network and that the arcs have finite capacities which are independent *two-state* random variables B_i with states 0 and b_i for $i\in\mathcal{A}$. Let $X_i = 1$ if $B_i = b_i$ and $X_i = 0$ otherwise. Also, let s and t denote a pair of selected nodes in $\mathcal{N}$. For given capacity state-vector $x\in\mathcal{X}$, where $\mathcal{X} = \{0,1\}^a$, let $\Lambda(x)$ denote the *value of a maximum s–t flow*. Every directed minimal s–t cutset (or *s–t cutset* hereafter) $\mathcal{C}$ separates $\mathcal{N}$ into two sets X and $\bar{X}$ such that $X \cup \bar{X} = \mathcal{N}$, $X \cap \bar{X} = \emptyset$, $s\in X$ and $t\in\bar{X}$. The *capacity* of $\mathcal{C}$, denoted by $Z(\mathcal{C},x)$, is defined as $\sum_{i\in(X,\bar{X})} b_i x_i$ and satisfies $Z(\mathcal{C},x) \geq \Lambda(x)$. The cutset $\mathcal{C}$ is called *critical* when $Z(\mathcal{C},x) = \Lambda(x)$. For a review on maximum flows see Ford and Fulkerson (1962).

Fix two maximum flow values $l < u \leq \Lambda(b_1,\dots,b_a)$ and a cutset $\mathcal{C}$. This paper proposes an efficient Monte Carlo sampling plan for estimating the variation of the probability $h(\mathcal{C},q)$ that the cutset $\mathcal{C}$ is critical given that the value of a maximum s-t flow $\Lambda(X)$ lies in the interval $(l,u]$ in response to varying reliability vectors q in a set $\mathcal{S}\subset(0,1)^a$, where $q_i = \mathrm{pr}(B_i=b_i)$ for $i\in\mathcal{A}$. We call the vectors in $(0,1)^a$ *points*. The evaluation of $h(\mathcal{C},q)$ at a single point q is an intractable (*NP-hard*) problem.

Identifying s-t cutsets $\mathcal{C}^*$ with high $h(\mathcal{C}^*,q)$, for fixed q, as well as points q for which the probabilities $h(\mathcal{C}^*,q)$, for a fixed s-t cutset $\mathcal{C}^*$, become large or small are very important issues in repairing and designing stochastic networks.

For a fixed point q, let $g(q)$ denote the probability that $l < \Lambda(X) \leq u$ and let $f(\mathcal{C},q)$ denote the probability that $\mathcal{C}$ is a critical s-t cutset and $l < \Lambda(X) \leq u$. We then define

$$h(\mathcal{C},q) = \begin{cases} f(\mathcal{C},q)/g(q) & \text{if } g(q) > 0 \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

Hereafter, assume $g(q) > 0 \quad \forall\, q \in \mathcal{S}$.

We first describe briefly an efficient Monte Carlo *importance sampling* plan in Alexopoulos and Fishman (1988) for estimating $h(\mathcal{C},q)$ at a single point q and then we show how this plan can be extended for estimating this flow performance measure at all points in $\mathcal{S}$. The sampling plans use an upper bound on the flow probability $g(q)$, $q\in[0,1]^a$ to gain their advantage over a crude Monte Carlo sampling plan. Let $P(x,q) = \prod_{i=1}^{a} q_i^{x_i}(1-q_i)^{1-x_i}$ denote the p.m.f. of the random vector $X = (X_1,\dots,X_a)$. Also, let $\mathcal{P}_1,\dots,\mathcal{P}_I$ denote arc-disjoint directed s-t paths and $\mathcal{C}_1,\dots,\mathcal{C}_J$ denote arc-disjoint s-t cutsets. Define $\Lambda_1(X) = \sum_{i=1}^{I} [\min_{j\in\mathcal{P}_i} b_j] \prod_{j\in\mathcal{P}_i} X_i$, that is the maximum amount of flow that can be transmitted from s to t through the paths $\mathcal{P}_1,\dots,\mathcal{P}_P$ and $\Lambda_2(X) = \min_{j=1,\dots,J} \sum_{i\in\mathcal{C}_j} b_i X_i$. One can readily show that

$$g(q) \le g_U(q) = \mathrm{pr}[\Lambda_2(X) > l,\ \Lambda_1(X) \le u] \tag{2}$$

$$= \mathrm{pr}[\Lambda_2(X) > l] - \mathrm{pr}[\Lambda_1(X) \le u].$$

For $x \in \mathscr{X}$ let $\phi(x) = I_{(l,u]}(\Lambda(x))$, $\phi_U(x) = I_{(l,\infty)}(\Lambda_2(x))I_{(-\infty,u]}(\Lambda_1(x))$ and $\psi(\mathscr{C},x) = I_{\{0\}}(Z(\mathscr{C},x) - \Lambda(x))$ where $I_{(c,d]}(x)$ is the indicator function on the interval $(c,d]$.

To take advantage of this upper bound, one constructs a conditional distribution

$$Q(x,q) = \phi_U(x)P(x,q)/g_U(q) \tag{3}$$

(see ref. [AF] for details) and draws K independent samples $X^{(1)},\dots,X^{(K)}$ from it. An estimate for $h(\mathscr{C},q)$ is given by

$$\tilde{h}_K(\mathscr{C},q) = \tilde{f}_K(\mathscr{C},q)/\tilde{g}_K(q) \qquad \text{if } \tilde{g}_K(q) > 0$$
$$= 0 \qquad \text{otherwise} \tag{4}$$

where

$$\tilde{f}_K(\mathscr{C},q) = g_U(q)\,\frac{1}{K}\sum_{k=1}^{K}\phi(X^{(k)})\psi(\mathscr{C},X^{(k)})$$

and (5)

$$\tilde{g}_K(q) = g_U(q)\,\frac{1}{K}\sum_{k=1}^{K}\phi(X^{(k)})$$

are unbiased estimates of $f(\mathscr{C},q)$ and $g(q)$, respectively, with variances

$$\mathrm{var}\ \tilde{f}_K(\mathscr{C},q) = f(\mathscr{C},q)[g_U(q)-f(\mathscr{C},q)]/K \le g_U(q)^2/K$$

and

$$\mathrm{var}\ \tilde{g}_K(\mathscr{C},q) = g(q)[g_U(q)-g(q)]/K \le g_U(q)^2/K.$$

From Alexopoulos (1988) one has

$$\mathrm{E}\tilde{h}_K(\mathscr{C},q) = h(\mathscr{C},q)[1-(1-g(q))^K]$$

and (6)

$$\mathrm{var}\ \tilde{h}_K(\mathscr{C},q) = [g(q)/g_U(q)]h(\mathscr{C},q)[1-h(\mathscr{C},q)]/K+o(K^{-1})$$

where $o(y)$ denotes a function of y such that $o(y)\to 0$ as $y\to 0$. To order $1/K$ this variance improves on the variance of a crude estimate by a factor of $g_U(q)$ demonstrating the efficiency of the importance sampling plan based on the bound information.

One can compute the bound $g_U(q)$ in time polynomial in b_i, $i \in \{\bigcup_{k=1}^{I} \mathscr{P}_i\} \cup \{\bigcup_{j=1}^{J} \mathscr{C}_j\}$ and sample from the distribution $\{Q(x,q)\}$ in (3) in time $O(|\mathscr{A}|)$.

The sampling plan described above is designed to provide an estimate for $h(\mathscr{C},q)$ at a single reliability vector q. Therefore, the estimation of the function $\{h(\mathscr{C},q),\ q \in \mathscr{Q}\}$ requires, in principle, $|\mathscr{Q}|$ experiments. We now show how sampling data collected for estimating $h(\mathscr{C},q)$ at a single point can be used to provide estimates for $h(\mathscr{C},q)$ at all points $q \in \mathscr{Q}$.

Let p be a point and define the *importance functions*

$$R(x,q,p) = P(x,q)/P(x,p) \qquad x \in \mathscr{X};\ q \in \mathscr{Q}$$

$$= \prod_{i=1}^{a}\left[\frac{q_i}{p_i}\right]^{x_i}\left[\frac{1-q_i}{1-p_i}\right]^{1-x_i}.$$

Suppose one draws K independent samples $X^{(1)},\dots,X^{(K)}$ from $\{Q(x,p)\}$. Then

$$\tilde{g}_{aK}(q,p) = g_U(p)\,\frac{1}{K}\sum_{k=1}^{K}\phi(X^{(k)})R(X^{(k)},q,p)$$

and $\qquad q \in \mathscr{Q} \qquad$ (8)

$$\tilde{g}_{bK}(q,p) = g_U(q)-g_U(p)\,\frac{1}{K}\sum_{k=1}^{K}[1-\phi(X^{(k)})]R(X^{(k)},q,p)$$

are unbiased estimates of $g(q)$, $q \in \mathscr{Q}$ with

$$K\,\mathrm{var}\ \tilde{g}_{aK}(q,p) = K\,\mathrm{var}\ \tilde{g}_K(q) + \{c(q,p)g_U(p)g_U(q^*)-g_U(q)g(q)\}$$

and (9)

$$K \operatorname{var} \tilde{g}_{bK}(q,p) = K \operatorname{var} \tilde{g}_K(q)$$

$$+ \{c(q,p) g_U(p)[g_U(q^*) - g_U(q)]$$

$$- g_U(q)[g_U(q) - g(q)]\},$$

where

$$c(q,p) = \prod_{i=1}^{a} c_i(q_i,p_i)$$

$$c_i(q_i,p_i) = q_i^2/p_i + (1-q_i)^2/(1-p_i) \qquad q \in \mathscr{S} \qquad (10)$$

$$q_i^* = q_i^2/[c_i(q_i,p_i)p_i]$$

and

$$q^* = (q_1^*,\dots,q_a^*),$$

indicating that $\operatorname{var} \tilde{g}_{jK}(q,p) < \operatorname{var} \tilde{g}_K(q)$ for some $j = a,b$ provided that the corresponding expression in curly brackets is negative. One can derive unbiased estimates $\{\tilde{f}_{iK}(\mathscr{C},q,p),\ q \in \mathscr{S}\}$, $i = a,b$ of $\{f(\mathscr{C},q),\ q \in \mathscr{S}\}$ and their variances if he(she) replaces $\phi(X^{(k)})$ with $\phi(X^{(k)})\psi(\mathscr{C},X^{(k)})$ in (8) and $g(q)$, $g(q^*)$ with $f(\mathscr{C},q)$, $f(\mathscr{C},q^*)$ respectively in (9).

Combining the estimates $\tilde{f}_{iK}(\mathscr{C},q,p)$, $i = a,b$ and $\tilde{g}_{jK}(q,p)$, $j = a,b$ one has four potential estimates for $L(\mathscr{C},q)$ for $q \in \mathscr{S}$, namely

$$\tilde{h}_{ijK}(\mathscr{C},q,p) = \tilde{f}_{iK}(\mathscr{C},q,p)/\tilde{g}_{jK}(q,p) \quad \text{if } \tilde{g}_{jK}(q,p)>0 \qquad i,j=a,b \qquad (11)$$

$$= 0 \qquad \text{otherwise.}$$

Fix $q \in \mathscr{S}$, let

$$A_K^{(j)} = \{|\tilde{g}_{jK}(q,p) - g(q)| < g(q)\} \qquad j=a,b$$

and observe that $\operatorname{pr}[A_K^{(j)}] \geq 1 - [1-g(p)/g_U(p)]^K$ and then $\lim_{K\to\infty} \operatorname{pr}[A_K^{(j)}] = 1$ for $j = a,b$. From Fishman (1978) (p. 55-59) one has

$$\lim_{K\to\infty} K\, \mathrm{E}[\tilde{h}_{ijK}(\mathscr{C},q,p) - h(\mathscr{C},q) \,|\, A_K^{(j)}] = h(\mathscr{C},q)$$

$$\times K \left\{ \frac{\operatorname{var} \tilde{g}_{jK}(q,p)}{g(q)^2} - \frac{\operatorname{cov}[\tilde{f}_{iK}(\mathscr{C},q,p), \tilde{g}_{jK}(q,p)]}{f(\mathscr{C},q) g(q)} \right\}$$

and $\qquad i,j=a,b \qquad (12)$

$$\lim_{K\to\infty} K \operatorname{var}[\tilde{h}_{ijK}(\mathscr{C},q,p) \,|\, A_K^{(j)}] = h(\mathscr{C},q)^2$$

$$\times K \left\{ \frac{\operatorname{var} \tilde{g}_{jK}(q,p)}{g(q)^2} - 2\frac{\operatorname{cov}[\tilde{f}_{iK}(\mathscr{C},q,p), \tilde{g}_{jK}(q,p)]}{f(\mathscr{C},q) g(q)} + \frac{\operatorname{var} \tilde{f}_{iK}(\mathscr{C},q,p)}{f(\mathscr{C},q)^2} \right\}$$

indicating the possibility that $\operatorname{var} \tilde{h}_{ijK}(\mathscr{C},q,p) < \operatorname{var} \tilde{h}_K(\mathscr{C},q)$ for some i and some j.

To anticipate the efficiency of the proposed experiment, we compare it with the $|\mathscr{S}|$ single experiments which produce the estimates $\tilde{h}_K(\mathscr{C},q)$ in (4). Namely, let $T(\mathscr{S})$ denote the mean time per replication for the proposed experiment and $T(q)$ denote the mean time per replication for the single experiment that is designed for estimating $h(\mathscr{C},q)$ for $q \in \mathscr{S}$. Then

$$\alpha(p) = \sum_{q \in \mathscr{S}} \frac{T(q)}{T(\mathscr{S})} \times \frac{\operatorname{var} \tilde{h}_K(\mathscr{C},q)}{\min\limits_{i,j=a,b} \operatorname{var} \tilde{h}_{ijK}(\mathscr{C},q,p)}$$

denotes the number of single experiments required to run to produce estimates of $\{h(\mathscr{C},q),\ q \in \mathscr{S}\}$ as accurate overall as the estimates the proposed experiment produces when all $|\mathscr{S}| + 1$ experiments run for the same time.

Note that the cutset $\mathscr{C}$ in the definitions of $f(\mathscr{C},q)$ and $h(\mathscr{C},q)$ can be replaced with a set Γ of s-t cutsets and then $h(\Gamma,q)$, that is the probability that Γ contains a critical cutset given that the value of a maximum s-t flow lies in $(l,u]$, can be estimated similarly to $h(\mathscr{C},q)$ at an incidental additional cost.

As shown by (9) and (12), the sampling point p affects the accuracy of the estimates in (8) and (11), therefore, the selection of it using only *a-priori* information is an important issue. Note that the quantity $c(q,p)g_U(p)g_U(q^*)$ is an upper bound on $\operatorname{var} \tilde{g}_{jK}(q,p)$, $j = a,b$ for $q \in \mathscr{S}$. One approach then chooses a sampling point p which minimizes the worst case bound $\max_{q \in \mathscr{S}} c(q,p) g_U(p) g_U(q^*)$. Other approaches, related directly to the accuracy of the estimates $\tilde{h}_{ijK}(\mathscr{C},q)$, $i,j = a,b$ in (11), are currently under investigation.

Confidence intervals (individual and simultaneous) for $h(\mathscr{C},q)$, $q \in \mathscr{S}$ are necessary for evaluating the accuracy of the estimates the proposed method produces. Since convergence to normality is not uniform in general, confidence intervals which are derived with exclusive use of statistical inequalities and are, therefore, valid for each finite sample size K deserve special attention. Such non normal confidence intervals are currently under development.

Finally, the general case of multiple capacity levels and set $\mathscr{S}$ consisting of joint p.m.fs. with independent marginals and common support is also under development by the authors of this paper.

ACKNOWLEDGEMENT

This research was supported by the Air Force Office of Scientific Research under grant AFOSR-84-0140. Reproduction in whole or part is permitted for any purpose of the United States Government.

REFERENCES

Alexopoulos, C. and G.S. Fishman (1988). Characterizing stochastic flow networks using the Monte Carlo method, Technical Report No. UNC/OR/TR-88/4, Department of Operations Research, University of North Carolina at Chapel Hill.

Alexopoulos, C. (1988). Maximum flows and critical cutsets in stochastic networks with discrete arc capacities, Ph.D. Thesis, Department of Operations Research, University of North Carolina at Chapel Hill.

Cochran, W.G. (1977). *Sampling Techniques*, John Wiley and Sons.

Fishman, G.S. (1978). *Principles of Discrete Event Simulation*, John Wiley and Sons.

Fishman, G.S. (1986). Monte Carlo estimation of the maximum flow distribution in a network with discrete stochastic edge capacity levels, Technical Report No. UNC/OR/TR-86/18, Department of Operations Research, University of North Carolina at Chapel Hill.

Fishman, G.S. (1986). Confidence intervals for a mean and a proportion in the bounded case, Technical Report No. UNC/OR/TR-86/19, Department of Operations Research, University of North Carolina at Chapel Hill.

Fishman, G.S. (1987). Maximum flow and critical cutset as descriptors of multi-state systems with randomly capacitated components, *Computers and Operations Research*, 14, 507-521.

Ford, L.R. and D.R. Fulkerson (1962). *Flows in Networks*, Princeton University Press.

AUTHORS' BIOGRAPHIES

CHRISTOS ALEXOPOULOS is an assistant professor in the School of Industrial and Systems Engineering at the Georgia Institute of Technology. He has a B.S. degree in Mathematics from the Aristotle University of Thessaloniki in Greece and a Ph.D. degree in Operations Research from the University of North Carolina at Chapel Hill. At present, he is working on methods for computing quantities characterizing the performance of stochastic networks. He is a member of the Operations Research Society of America.

Christos Alexopoulos
School of Industrial and Systems Engineering
Georgia Institute of Technology
Atlanta, Georgia 30332

GEORGE FISHMAN is professor and chairman of the Department of Operations Research at the University of North Carolina at Chapel Hill. His principal interest is the development of statistical methodology applicable to the analysis of output from discrete event digital simulation models. He is the author of *Concepts and Methods in Discrete Event Digital Simulation* published by Wiley in 1973 and of *Principles of Discrete Event Simulation* published by Wiley in 1978. He is a frequent contributor to the operations research and statistical literature on this topic. At present, he is working on variance reducing methods for network reliability estimation and on the influence of concurrent processing on simulation program structure. Professor Fishman has been simulation departmental editor for *Management Science* and is a member of the Operations Research Society of America, the Institute of Management Science and the American Statistical Association.

George S. Fishman
Department of Operations Research
University of North Carolina
Chapel Hill, North Carolina 27599