Proceedings of the 1990 Winter Simulation Conference
Osman Balci, Randall P. Sadowski, Richard E. Nance (eds.)

SENSITIVITY ANALYSIS WITH REGARD TO CAPACITY EXPANSION IN NETWORK FLOW SIMULATION

Christos Alexopoulos

School of Industrial and Systems Engineering
Georgia Institute of Technology
Atlanta, Georgia 30332-0205

George S. Fishman

Department of Operations Research
University of North Carolina
Chapel Hill, North Carolina 27599

ABSTRACT

In the design of flow networks it is desirable to assess the incremental gain in network flow permitted by increasing the flow capacities of one or more components of the system. Although a well-established methodology exists for doing this for a deterministic system, probing this question for a stochastic flow network encounters many problems not present in the deterministic case. This paper provides a Monte Carlo sampling plan for investigating this issue. It allows one to conduct a sensitivity analysis for a variable upper bound on the flow capacity of a specified arc where the individual arc flow capacities are all random. The plan has two notable features. It permits estimation of the probabilities of a feasible flow for many values of the upper bound on the arc capacity from a single data set generated by the Monte Carlo method at a single value of the upper bound. Also, the resulting estimators have considerably smaller variances than crude Monte Carlo sampling would produce in the same setting. The success of the technique follows from the use of lower and upper bounds on each probability of interest where the bounds are generated from an established method of decomposing the capacity state space.

1. INTRODUCTION

Designing a flow network often calls for evaluating several alternative network configurations to decide which one best meets a specified objective. For example, given a node set $\mathcal{N}$ one may require a network to realize a flow rate of at least d from a source node s to a terminal node t $(s,t \in \mathcal{N})$ where the same set of arcs $\mathcal{A}$ holds for each alternative design but with different flow rate capacities on each alternative for a particular arc e. Suppose that arc capacities are random and, in particular, that arc e has capacity $c > 0$ with probability p_{e2} and zero capacity with probability $p_{e1} = 1-p_{e2}$. Then $g(d,c)$, the probability that a feasible flow exists when arc e has an upper capacity level c, provides one measure of performance and one may be interested in evaluating how $g(d,c)$ varies as c takes on values in $\mathcal{C} = \{\alpha, \alpha+1,..., \beta\}$, where $\alpha < \beta$ are positive integers. Conceptually, this analysis allows a designer to assess how much of an improvement in system reliability occurs as c increases. Since evaluating $g(d,c)$ belongs to the class of NP-hard problems [Alexopoulos and Fishman 1988], no polynomial time algorithm is currently available for computing $g(d,c)$ exactly. The need to evaluate this probability for a sequence of alternative capacity levels for arc e merely compounds the problem.

To overcome these limitations, this paper describes a Monte Carlo sampling plan that allows one to estimate $\{g(d,c),\ c \in \mathcal{C}\}$. The novelty of the approach comes from its ability to produce all $|\mathcal{C}|$ estimates using data generated in a single sampling experiment that uses lower and upper bounds on $g(d,c)$. The advantages of the proposed technique are twofold: First, it is a considerable improvement over the crude Monte Carlo approach and second, it extends the range of application of the method proposed in [Fishman and Shaw 1989] for estimating $g(d,c)$ for a single c. The latter approach uses known lower and upper bounds on $g(d,c)$ that hold for fixed c and, therefore, the estimation of $\{g(d,c),\ c \in \mathcal{C}\}$ requires essentially $|\mathcal{C}|$ experiments.

Let $G = (\mathcal{N},\mathcal{A},s,t)$ denote the flow network and suppose that the arcs are numbered so that $\mathcal{A} = \{1,...,a\}$. Each arc i has a random capacity B_i that takes values in the set $\{0 \le b_{i1} < ... < b_{in_i} < \infty\}$ $(n_e = 2)$ with probabilities $p_{i1},...,p_{in_i}$ respectively. If $b_{e1} > 0$, one replaces e by two arcs e' and e'' in parallel such that $B_{e'} = b_{e1}$, $B_{e''} = 0$ w.p. p_{e1} and $B_{e''} = b_{e2}-b_{e1}$ w.p. p_{e2}. Let $\Omega(c)$ denote the state space of the random vector $B = (B_1,...,B_a)$ when $b_{e2} = c$. A point b of $\Omega(c)$ can be defined as an a-tuple of values $b = (b_{1v_1},...,b_{av_a})$ where v_i is a numerical index for i going from 1 to n_i. To simplify the notation, the index v_i will also be used to denote the value b_{iv_i} itself so that a state point b will also be denoted as $v = (v_1,...,v_a)$.

Assume that the capacities are independent random variables and let $P(v)$ denote the probability that the system is in state v. Then

$$P(v) = \prod_{i=1}^{a} p_{iv_i}. \tag{1}$$

Let $\Lambda(v,c)$ denote the value of a maximum s-t flow when the arc capacities are v and $b_{e2} = c$ and fix $d > 0$. For any state v and fixed $c > 0$ define the *structure function*

$$\phi(v,c) = \begin{cases} 1 & \text{if } \Lambda(v,c) \geq d \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

Then, for fixed demand d, the *network reliability* $g(c) \equiv g(d,c) = \Pr[\Lambda(B,c) \geq d]$ can be written as

$$g(c) = \sum_{v \in \Omega(c)} \phi(v,c)P(v). \quad (3)$$

The limitations on the exact computation of (3) have prompted a search for alternative methods for approximating $g(c)$, among which the Monte Carlo method is a prime candidate. For fixed c, Fishman and Shaw [1989] used available lower and upper bounds on $g(c)$ generated by a decomposition of the state space $\Omega(c)$ due to Doulliez and Jamoulle [1972] to replace the sampling distribution $\{P(v)\}$ with a modified distribution that produces an estimate of $g(c)$ with greater accuracy than one would obtain for the same cost, were one to sample directly from $\{P(v)\}$. This ability to incorporate information on bounds into a Monte Carlo sampling plan fits well with the philosophical attitude that one should use as much deterministic information as possible before turning to a sampling experiment to sharpen one's estimate of $g(c)$.

Section 2 describes the essential features of crude Monte Carlo and then describes the incorporation of the Doulliez and Jamoulle (D-J) decomposition. Section 3 extends this decomposition method to the case in which we wish to estimate $\{g(d,c),\ c \in \mathscr{C}\}$ from data generated on a single experiment. Section 4 describes the experiment and Sec. 5 illustrates how the extension works in practice.

2. MONTE CARLO METHOD

Suppose that one draws N i.i.d. samples from from $\Omega(\alpha)$ using the distribution $\{P(v)\}$. Let $V_i^{(n)}$ denote the state of arc i in the nth trial, and let $V^{(n)} = (V_1^{(n)},\ldots, V_a^{(n)})$. Also, suppose in each trial n one performs the steps:

1. For $c \in \mathscr{C}$: set $\phi(V^{(n)},c) = 0$.
2. If $\Lambda(V^{(n)},\alpha) \geq d$: For $c \in \mathscr{C}$ set $\phi(V^{(n)},c) = 1$; end replication.
3. If $V_e^{(n)} = 1$: end replication.
4. If $\Lambda(V^{(n)},\alpha) + \beta - \alpha < d$: end replication.
5. Find $k = \min\{c : \alpha < c \leq \beta \text{ and } \Lambda(V^{(n)},\alpha) + c - \alpha \geq d\}$.
6. If $\Lambda(V^{(n)},k) \geq d$: For $k \leq c \leq \beta$ set $\phi(V^{(n)},c) = 1$.

Note that step 5 is based on the fact that $\Lambda(V^{(n)},c) = \min\{\Lambda(V^{(n)},\infty),\ \Lambda(V^{(n)},\alpha) + c - \alpha\}$. Then the crude Monte Carlo estimators

$$\bar{g}_N(c) = \sum_{n=1}^{N} \phi(V^{(n)},c)/N \qquad c \in \mathscr{C} \quad (4)$$

of $g(c)$ are unbiased with

$$\text{var}\ \bar{g}_K(c) = g(c)[1-g(c)]/N. \quad (5)$$

Two sources principally contribute to the cost per replication. One is the time to sample a capacity state v from $\{P(v);\ v \in \Omega(\alpha)\}$ and the other is the time to determine a maximum flow at most twice. For sampling, the cutpoint method [Fishman and Moore 1984] takes $O(|\mathscr{A}|)$ time on the average. For general networks once the capacity state v is given, a maximum flow can be determined via a maximal flow algorithm in $O(|\mathscr{N}|^3)$ time [Papadimitriou and Steiglitz 1982].

For fixed c and for a specified sample size N a reduction in the variance (4) follows from using information on bounds for $g(c)$ and modifying the sampling distribution $\{P(v)\}$ [Kumamoto et al. 1977 and Fishman 1986a]. The present paper derives the bounds on the entire function $\{g(c),\ c \in \mathscr{C}\}$ from a decomposition of the capacity state spaces $\Omega(c)$ by extending the approach of Doulliez and Jamoulle. While the time to compute the bounds for all c is $O(I|\mathscr{N}|^4)$ where I is an integer ≥ 1, no special tables are needed for the sampling distribution and the cost of sampling v and determining $\Lambda(v,c)$ is no greater than in the case of crude Monte Carlo sampling. These properties make the proposed method more appealing than an alternative approach in [Fishman 1989] where the modified distribution depends on c, and the time to draw a sample from it is about twice as great as the sampling time from $\{P(v)\}$.

We describe shortly how the D-J decomposition provides useful lower and upper bounds to implement the proposed approach. Doulliez and Jamoulle describe a state space decomposition method, based on iteration, to compute $g(c)$ exactly. Although the number of iterations for exact computation can grow exponentially with the size of the problem, on each iteration the technique produces lower and upper bounds on $g(c)$ that become progressively tighter as the number of iterations increases. Any state $v \in \Omega(c)$ can be classified as either an *operating* state $(\phi(v,c) = 1)$ or a *failed* state $(\phi(v,c) = 0)$. Let $\mathscr{S}^+$ be the set of all operating states and $\mathscr{S}^-$ be the set of all failed states.

Since the state space $\Omega(c)$ has $\Pi_{i=1}^{a} n_i$ elements, it is hopeless in general to determine $\mathscr{S}^+$ (and hence $\mathscr{S}^-$) by considering each state in $\Omega(c)$ one after the other. Doulliez and Jamoulle present an iterative method for overcoming this problem. At each iteration, their algorithm determines disjoint sets of operating states, failed states and *undetermined* states. An undetermined state is one in which it is not possible at that iteration to determine $\phi(v,c)$. These undetermined states are used as input to the next iteration to determine additional operating and failed states and again any remaining undetermined states are used in the next iteration. The procedure ends with a total decomposition into operating and failed states and an exact evaluation of $g(c)$. The operating and undetermined sets that are produced in each iteration are (discrete) rectangles in the sense that: for each such set S there is a lower limiting point $\boldsymbol{l}[S]$ and an upper limiting point $\boldsymbol{u}[S]$ such that each integer $\boldsymbol{v}$ with $\boldsymbol{l}[S] \le \boldsymbol{v} \le \boldsymbol{u}[S]$ belongs to S.

For fixed c, Fishman and Shaw [1989] showed how this method can be used for deriving lower and upper bounds on $g(c)$ and presented an efficient sampling plan for estimating this quantity. However, a problem with adopting the D-J method and thereby the Fishman and Shaw method to the present setting is that an operating state for $b_{e2} = c$ is not necessarily an operating state for $b_{e2} < c$ and a failed state when $b_{e2} = c$ may not be failed when $b_{e2} > c$.

3. PROPOSED METHOD

The method proposed here also takes advantage of the D-J state-space decomposition method. In particular, it decomposes at most I subsets of $\Omega(\alpha)$ and produces operating subsets $W_1, W_2, \ldots, W_I$ of $\Omega(c)$ for all $c \in \mathscr{C}$, failed subsets of $\Omega(\alpha)$, disjoint undetermined subsets $U_1(\alpha), U_2(\alpha), \ldots, U_{M_\alpha}(\alpha)$ of $\Omega(\alpha)$, failed subsets of $\Omega(\beta)$, and disjoint undetermined subsets $U_1(\beta), U_2(\beta), \ldots, U_{M_\beta}(\beta)$ of $\Omega(\beta)$. The families $\mathscr{U}(\alpha) = \{U_1(\alpha), U_2(\alpha), \ldots, U_{M_\alpha}(\alpha)\}$ and $\mathscr{U}(\beta) = \{U_1(\beta), U_2(\beta), \ldots, U_{M_\beta}(\beta)\}$ satisfy

Property 1: $\mathscr{U}(\alpha) \subseteq \mathscr{U}(\beta)$ and the rectangles in $\mathscr{U}(\beta) - \mathscr{U}(\alpha)$ are failed subsets of $\Omega(\alpha)$.

An undetermined set $S \subseteq \Omega(\alpha)$ with lower and upper limiting points $\boldsymbol{l}$ and $\boldsymbol{u}$ is decomposed as follows: Create a fictitious demand node T, add the arc $e' = (t, T)$ numbered with capacity d and determine a maximum s-T flow $f = (f_1 \cdots f_a, f_{e'})$ with capacities $\boldsymbol{u}$ for the arcs in $\mathscr{A}$. If the value of this flow $\Lambda(\boldsymbol{u},\alpha)$ is less than d, none of the states in S can satisfy the demand at node t and then S is a failed subset of $\Omega(\alpha)$. Suppose that $\Lambda(\boldsymbol{u},\alpha) = d$ and for each arc i define $v_i^0 = \min\{z : l_i \le z \le u_i \text{ and } z \ge f_i\}$. Obviously, all states $\boldsymbol{v}$ with $v_j^0 \le v_i \le u_i \;\forall i$ form an operating rectangle $W \subseteq \Omega(\alpha)$ and therefore of $\Omega(c)$ for $c > \alpha$. Now given the flow f, for each arc i let $\Lambda_i(\alpha)$ denote the value of maximum flow that can be transmitted from the tail of i to the head of i without using this arc. If $\Lambda_i(\alpha) < f_i$, arc i is in a minimal cut that blocks the value of maximum s-t flow below d and every state $\boldsymbol{v} \in S$ with $v_i < f_i - \Lambda_i(\alpha)$ is failed. If $\Lambda_i(\alpha) \ge f_i$, any capacity $l_i \le z \le u_i$ for arc i satisfies the demand d when all other capacities are fixed at u_k, $k \ne i$. For $i \in \mathscr{A}$ define

$$\begin{aligned} v_i^*[\alpha] &= \min\{z : l_i \le z \le u_i,\, z \ge f_i - \Lambda_i(\alpha)\} \quad \text{if } f_i > \Lambda_i(\alpha) \\ &= l_i \quad \text{if } f_i \le \Lambda_i(\alpha). \end{aligned} \tag{6}$$

Then the sets

$$F_i(\alpha) = \{\boldsymbol{v} : \boldsymbol{v} \in S \text{ and } v_i < v_i^*[\alpha]\} \quad i \in \mathscr{A}$$

are failed subsets of $\Omega(\alpha)$, the rectangles

$$\begin{aligned} L_i(\alpha) = \{\boldsymbol{v} : v_k^0 \le v_k \le u_k \text{ for } k < i,\; v_i^*[\alpha] \le v_i \le v_i^0 - 1 \\ \text{and } v_k^*[\alpha] \le v_k \le u_k \text{ for } k > i\} \quad i \in \mathscr{A} \end{aligned} \tag{7}$$

are disjoint, undetermined subsets of $\Omega(\alpha)$ and $S = W \cup \{\cup_{i=1}^a L_i(\alpha)\} \cup \{\cup_{i=1}^a F_i(\alpha)\}$.

Note that if $u_e = 2$, that is if the values $\Lambda_i(\alpha)$, $i \ne e$ were computed with capacity $b_{e2} = \alpha$, then each failed set $F_i(\alpha)$ would not be necessarily failed when $b_{e2} = c$ for $c > \alpha$. However, each failed set when $b_{e2} = \beta$ is also failed when $b_{e2} = c$ for $c < \beta$. We now determine failed and undetermined subsets of $\Omega(\beta)$ as follows: If $u_e = 1$, the rectangles $L_i(\alpha)$ are added to $\mathscr{U}(\beta)$. If $u_e = 2$, set $v_e^*[\beta] = v_e^*[\alpha]$ and, given the initial flow f and the capacities u_k for $k \ne i, e$ and $b_{e2} = \beta$, for each arc $i \ne e$ let $\Lambda_i(\beta)$ denote the value of maximum flow that can go from the tail of arc i to the head of i when this arc is deleted. For $i \in \mathscr{A}$ define

$$\begin{aligned} v_i^*[\beta] &= \min\{z : l_i \le z \le u_i,\, z \ge f_i - \Lambda_i(\beta)\} \quad \text{if } f_i > \Lambda_i(\beta) \\ &= l_i \quad \text{if } f_i \le \Lambda_i(\beta). \end{aligned} \tag{8}$$

Then, the sets

$$F_i(\beta) = \{\boldsymbol{v} : \boldsymbol{v} \in S \text{ and } v_i < v_i^*[\beta]\} \quad i \in \mathscr{A}$$

are failed subsets of $\Omega(\beta)$ and the rectangles

$$\begin{aligned} L_i(\beta) = \{\boldsymbol{v} : v_k^0 \le v_k \le u_k \text{ for } k < i,\; v_i^*[\beta] \le v_i \le v_i^0 - 1 \\ \text{and } v_k^*[\beta] \le v_k \le u_k \text{ for } k > i\} \quad i \in \mathscr{A} \end{aligned} \tag{9}$$

are disjoint, undetermined subsets of $\Omega(\beta)$.

Observe that for each i, $\Lambda_i(\alpha) \le \Lambda_i(\beta)$ and then $v_i^*[\alpha] \ge v_i^*[\beta]$. This implies that each rectangle $L_j(\alpha)$ is a subset of a rectangle $L_m(\beta)$. In fact, the latter set is unique and has the same upper limiting point with $L_j(\alpha)$. The decomposition of $L_m(\beta) - L_j(\alpha)$ produces the disjoint undetermined subsets of $\Omega(\beta)$

$$R_i(\beta) = \{v : l_k[L_j(\alpha)] \le v_k \le u_k \text{ for } k < i, \; l_i[L_m(\beta)] \le v_i \le l_i[L_j(\alpha)] - 1 \text{ and } l_k[L_m(\beta)] \le v_k \le u_k \text{ for } k > i\} \quad i \in \mathscr{A}, \tag{10}$$

where u is the common limiting point of $L_j(\alpha)$ and $L_m(\beta)$. The rectangles $L_j(\alpha)$ and $R_i(\beta)$ are then added to $\mathscr{U}(\beta)$ and the rectangle S is removed from both $\mathscr{U}(\alpha)$ and $\mathscr{U}(\beta)$.

Once at most I subsets of $\Omega(\alpha)$ are decomposed, the remaining undetermined subsets of $\Omega(\alpha)$ and $\Omega(\beta)$ form the sets $\mathscr{U}(\alpha) = \{U_1(\alpha),\ldots,U_{M_\alpha}(\alpha)\}$ and $\mathscr{U}(\beta) = \{U_1(\beta),\ldots, U_{M_\beta}(\beta)\}$ respectively. One can easily check that $\mathscr{U}(\alpha)$ and $\mathscr{U}(\beta)$ satisfy Property 1.

Let $M = M_\beta$ and define

$$g_L = \sum_{k=1}^{I} \sum_{v \in W_k} P(v) = \sum_{k=1}^{I} \prod_{i=1}^{a} \left[\sum_{z=l_i[W_k]}^{u_i[W_k]} p_{iz} \right]$$

and

$$g_U = g_L + \sum_{m=1}^{M} \sum_{v \in U_m(\beta)} P(v) \tag{11}$$

$$= g_L + \sum_{m=1}^{M} \prod_{i=1}^{a} \left[\sum_{z=l_1[U_m(\beta)]}^{u_i[U_m(\beta)]} p_{iz} \right].$$

As a result for $\alpha \le c \le \beta$

$$g_L \le g(c) = g_L + \sum_{m=1}^{M} \sum_{v \in U_m(\beta)} \phi(v,c)P(v) \le g_U. \tag{12}$$

4. SAMPLING

For $\alpha \le c \le \beta$ let

$$H_{im} = \sum_{z=l_i[U_m(\beta)]}^{u_i[U_m(\beta)]} p_{iz} \quad i \in \mathscr{A}$$

and

$$\pi_m = \Pr[U_m(\beta)] = \sum_{v \in U_m(\beta)} P(v) = \prod_{i=1}^{a} H_{im}. \tag{13}$$

Then

$$\pi = \Pr[\mathscr{U}(\beta)] = \sum_{m=1}^{M} \pi_m = g_U - g_L.$$

Define the binary functions

$$I_m(v,\beta) = \begin{cases} 1 & \text{if } v \in U_m(\beta) \\ 0 & \text{otherwise} \end{cases} \quad m=1,\ldots,M \tag{14}$$

and the distribution

$$Q(v,\beta) = \frac{P(v)}{\pi} \sum_{m=1}^{M} I_m(v,\beta). \tag{15}$$

Suppose one draws N independent samples $V^{(1)},\ldots, V^{(N)}$ from $\{Q(v,\beta)\}$. It is easy to show that

$$\tilde{g}_N(c) = g_L + \pi \sum_{n=1}^{N} \phi(V^{(n)},c)/N \quad c \in \mathscr{C} \tag{16}$$

are unbiased estimates of $g(c)$ with variances

$$\begin{aligned} \operatorname{var} \tilde{g}_N(c) &= [g_U - g(c)][g(c) - g_L]/N \\ &\le [g_U - g_L]^2/4N. \end{aligned} \tag{17}$$

When compared to crude Monte Carlo sampling, this sampling plan leads to a worst-case variance ratio

$$\operatorname{var} \bar{g}_N(c)/\operatorname{var} \tilde{g}_N(c) \ge R^* = 1/\left[\sqrt{g_U(1-g_L)} - \sqrt{g_L(1-g_U)}\right] \ge 1 \tag{18}$$

which one can compute before sampling begins to determine the minimal reduction in variance to be expected. Observe that the form of the sampling distribution $\{Q(v,\beta)\}$ allows one to draw a sample $V = (V_1,\ldots,V_a)$ in $O(|\mathscr{A}|)$ mean time by using the cutpoint method in Fishman and Moore (1984) as follows:

a. Sample index m from $\{1,\ldots,M\}$ with probabilities $\{\pi_m/\pi, m = 1,\ldots,M\}$.
b. For $i = 1,\ldots,a$: Sample V_i from $\{p_{iz}/H_{im}, l_i[U_m(\beta)] \le z \le u_i[U_m(\beta)]\}$.

Procedure RUN describes a Monte Carlo sampling experiment that generates unbiased estimates of $\tilde{g}_N(c)$ for $c \in \mathscr{C}$ and of their variances. Steps II.2 through II.8 utilize Property 1. Note that this procedure requires only the rectangles in $\mathscr{U}(\beta)$ and the labels $\text{LABEL}[U_m(\beta)] = 1$ if $U_m(\beta) \in \mathscr{U}(\alpha)$, $\text{LABEL}[U_m(\beta)] = \alpha + 1$ if $U_m(\beta) \notin \mathscr{U}(\alpha)$ for $m = 1,\ldots,M$.

Procedure RUN

I. Initialization.
For $\alpha \leq j \leq \beta$: Set $S_j = 0$.

II. For N independent replications:
1. Sample $V = (V_1,...,V_a)$ from $Q(v,\beta)$.
2. Set $k = \text{LABEL}\ [U_m(\beta)]$.
3. Compute the maximum flow value $\Lambda(V,k)$ with capacities V and $b_{e2} = k$.
4. If $\Lambda(V,k) \geq d$: For $k \leq j \leq \beta$ set $S_j = S_j + 1$; end replication.
5. If $V_e = 1$ or $\Lambda(V,k) + \beta - k < d$: end replication.
6. Set $c = \min\{j:\ k < j \leq \beta \text{ and } \Lambda(V,k) + j - k \geq d\}$.
7. Compute the maximum flow value $\Lambda(V,c)$ with capacities V and $b_{e2} = c$.
8. If $\Lambda(V,c) \geq d$: For $c \leq j \leq \beta$ set $S_j = S_j + 1$.

III. Compute summary statistics.
For $\alpha \leq j \leq \beta$:

$$\tilde{g}_N(j) = g_L + \pi S_j/N.$$

$$VAR[\tilde{g}_N(j)] = [g_U - \tilde{g}_N(j)][\tilde{g}_N(j) - g_L]/(N-1).$$

The representation of $[g(c) - g_L]/\pi$ as a convex combination

$$[g(c) - g_L]/\pi = \sum_{m=1}^{M} (\pi_m/\pi)\mu_m(c) \tag{19a}$$

where

$$\mu_m(c) = \sum_{v \in \Omega(\beta)} \phi(v,c)Q_m(v,\beta) \tag{19b}$$

and

$$Q_m(v,\beta) = \frac{I_m(v,\beta)}{\pi_m} P(v) \quad v \in \Omega(\beta) \tag{19c}$$

suggests an alternative sampling plan that guarantees additional variance reduction. Let $N_m = N\pi_m$, for $m = 1,...M$. If one samples $V^{(1)},..., V^{(N_m)}$ from $\{Q_m(v,\beta)\}$, then

$$\hat{\mu}_m(c) = \sum_{n=1}^{N_m} \phi(V^{(n)},c)/N_m \quad c \in \mathscr{C} \tag{20}$$

are unbiased estimates of $\mu_m(c)$ and

$$\hat{g}_N(c) = g_L + \sum_{m=1}^{M} \pi_m\hat{\mu}_m(c) \quad c \in \mathscr{C} \tag{21}$$

are unbiased estimates of $g(c)$ with

$$\text{var}\ \hat{g}_N(c) = \text{var}\ \tilde{g}_N(c) - \sum_{m=1}^{M} \pi_m[\mu_m(c) - \mu]/N$$
$$\leq \text{var}\ \tilde{g}_N(c) \tag{22}$$

where

$$\mu = \sum_{m=1}^{M} \pi_m\mu_m(c)/\pi$$

and $\text{var}\ \hat{g}_N(c) = \text{var}\ \tilde{g}_N(c)$ only if $\mu_m(c)$, $m = 1,...,M$ are all equal. The estimates $\hat{g}_N(c)$ of $g(c)$, in Eq. (21), have an additional advantage over $\tilde{g}_N(c)$, in Eq. (16): Since no need exists to sample the stratum m on each replication, the time to sample V is slightly less.

Confidence intervals based on the sample means $\tilde{g}_N(c)$ and $\hat{g}_N(c)$ that are valid for every sample size N can be computed using the procedure in [Fishman 1986b] based on a result in [Hoeffding 1963].

5. EXAMPLE

An example with $n_i = 2$ (two-state arc capacity levels) for all $i \in \mathscr{A}$ illustrates how the proposed method works in practice. Figure 1 shows a network of 10 nodes, 25 arcs, $s = 1$, $t = 10$, and $e = 3$. Except for arc 3, the notation i/q on an arc denotes arc i with capacity level $b_{i2} = q$. We assume that $b_{i1} = 0$, and take $p_{i2} = \Pr[B_i = b_{i2}] = 0.9$ and $1-p_{i2} = \Pr[B_i = 0] = 0.1$ for all $i = 1,...,25$. The objective is to estimate $g(c)$, the probability that demand $d = 56$ at node t is satisfied when b_{32} takes each of the 31 values $c = 20 + k$, $k = 0,...,30$. The maximum flow value with all capacities at their upper levels and $b_{32} = 20$ is 67.

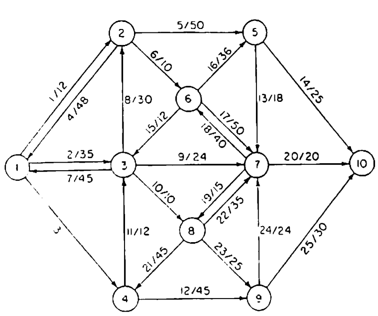

Figure 1. Network

For this special case of two-state capacities, we can simplify the presentation by defining the set of "on" arcs as

$$U_m^+(c) = \{i \in A : l_i[U_m(c)] = 2\}$$

and the set of "off" arcs as

$$U_m^-(c) = \{i \in A : u_i[U_m(c)] = 1\}$$

for each undetermined rectangle $U_m(c)$. Conversely, if $U_m^+(c)$ and $U_m^-(c)$ are given, the lower and upper boundary elements of $U_m(c)$ can be determined as

$$l_i[U_m(c)] = \begin{cases} 2 & \text{if} \quad i \in U_m^+(c) \\ 1 & \text{otherwise} \end{cases}$$

and

$$u_i[U_m(c)] = \begin{cases} 1 & \text{if} \quad i \in U_m^-(c) \\ 2 & \text{otherwise} \end{cases} .$$

Note that only the capacities of those arcs i not explicitly in these sets need to be sampled, thus reducing the computation time.

The decomposition procedure in Sec.3 produced 13 undetermined subsets of $\Omega(50)$ with two rounds of iterations, where in each round all the existing undetermined rectangles in $\mathscr{U}(20)$ were decomposed. The resulting bounds $g_L = 0.4651$ and $g_U = 0.6226$ and the minimal variance ratio $R^* = 40.01$ show a clear advantage over crude Monte Carlo sampling.

Table 1 presents results for estimating $g(c)$ for sample size $N = 65{,}536$. Note that the variance ratio $\text{var}\,\bar{g}_N/\text{var}\,\hat{g}_N$ measures the overall benefit that derives from combining both importance sampling and stratified sampling in the same estimator $\hat{g}_N$ when compared to crude Monte Carlo sampling as in $\bar{g}_N$, whereas $\text{var}\,\tilde{g}_N/\text{var}\,\hat{g}_N$ measures the effect of adding stratification to the importance sampling scheme. For example, for capacity $b_{32} = c = 20$, crude Monte Carlo sampling would have required 1335 observations for every observation using importance and stratified sampling together, whereas it would have required 14 observations when compared to importance sampling alone. While notably an improvement by itself, augmenting importance sampling by stratified sampling offers a guaranteed advantage, especially since the cost of using these two sampling techniques in the present context is no greater than the cost of stratified sampling alone.

Table 1. Monte Carlo Results for Demand $d = 56$ and Capacities c for Arc $3^\dagger$ $(N = 65{,}536)$

c	$\hat{g}_N$	$10^8 \times \text{var}\,\hat{g}_N$	$\dfrac{\text{var}\,\bar{g}_N}{\text{var}\,\hat{g}_N}$	$\dfrac{\text{var}\,\tilde{g}_N}{\text{var}\,\hat{g}_N}$	$\dfrac{\text{var}\,\bar{g}_N}{\text{var}\,\tilde{g}_N}$
20	.48442	.28554	93.60	14.26	1334.71
21	.53432	1.25676	40.74	3.63	147.87
22-23	.54495	1.21108	40.02	4.48	179.27
24-25	.54512	1.20697	40.00	4.57	182.82
26-30	.54565	1.97368	40.01	4.79	191.67
31	.54572	1.95757	40.01	4.83	193.25
32-33	.54654	1.82205	40.07	5.18	207.56
34-35	.55000	1.98425	40.15	4.74	190.33
36-43	.55335	2.08923	40.48	4.46	180.52
44-45	.59677	4.82953	71.06	1.07	76.03
46-50	.60023	4.40047	79.25	1.05	83.21

† Entries for $\text{var}\,\bar{g}_K$, $\text{var}\,\tilde{g}_K$ and $\text{var}\,\hat{g}_K$ are based on (5), (17) and (22), respectively, with estimates substituted for the true parameters.

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, revised 1989.

Doulliez, P. and E. Jamoulle (1972), "Transportation networks with random arc capacities," *R.A.I.R.O. 3*, 455-599.

Fishman, G.S. and L.R. Moore (1984), "Sampling from a discrete distribution while preserving monotonicity," *The Amer. Statist. 38*, 219-223.

Fishman, G.S. (1986a), "A Monte Carlo sampling plan for estimating network reliability," *Oper. Res. 34*, 581-594.

Fishman G.S. (1986b), "Confidence intervals for the mean in the bounded case," Technical Report No. UNC/OR/TR-86/19, Department of Operations Research, University of North Carolina at Chapel Hill, revised 1989.

Fishman, G.S. (1989), "Monte Carlo estimation of the maximal flow distribution with discrete stochastic arc capacity levels," *Naval Logistics Research Quarterly 36*, 829-849.

Fishman, G.S., and T.Y Shaw (1989),"Evaluating reliability of stochastic flow networks," *Prob. in the Engineering and Informational Sciences 3*, 493-509.

Hoeffding, W. (1963), "Probability inequalities for sums of bounded random variables," *J. Amer. Statist. Assoc. 58*, 13-29.

Kumamoto, H., K. Tanaka and K. Inoue (1977), "Efficient evaluation of system reliability by Monte Carlo method," *IEEE Trans. on Reliability, R-26*, 311-315.

Papadimitriou, C.H. and K. Steiglitz (1982), *Combinatorial Optimization,* Prentice Hall, Englewood Cliffs, New Jersey.

Walker, A.J. (1977) "An efficient method for generating discrete random variables with general distributions," *ACM Trans. on Math. Software 3*, 253-256.