Copyright © 1978, by the author(s). All rights reserved.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission.

MODULAR DECOMPOSITION IN STOCHASTIC TRANSPORTATION NETWORKS

by

A.W. Shogan

Memorandum No. UCB/ERL M78/86

18 September 1978

ELECTRONICS RESEARCH LABORATORY

College of Engineering University of California, Berkeley 94720

MODULAR DECOMPOSITION IN STOCHASTIC TRANSPORTATION NETWORKS

Andrew W. Shogan

University of California, Berkeley

ABSTRACT

Consider a flow network having random arc capacities and having associated with each node n a "supply-demand random variable" Y whose absolute value equals the supply available at the node when Y assumes a non-negative value and the demand required by the node when Y assumes a non-positive value. A fundamental problem is the computation of the reliability R, that is, the probability that the random variables will assume values that permit a feasible flow. Upon adapting the graph theoretic concepts of "cutnode" and "block," it is possible to identify a "block-module," an independent, non-trivial subnetwork that has one and only one node (the "cutnode") connected to nodes outside the subnetwork. The reliability of the network will increase by a known factor after a "block-modular decomposition" that consists of a transformation of the cutnode's supply-demand random variable and the deletion of the remainder of the block-module. Provided the original network possesses at least one block-module, R can be determined from a sequence of blockmodular decompositions that reduce the original network to a single node whose reliability is easily computed.

-1-

Research sponsored by the Department of Energy, Electric Energy System Division, Contract EC-77-S-01-5105.

For the purposes of this paper, a <u>stochastic transportation net-</u> <u>work</u> is a flow network having random arc capacities and having associated with each node n a <u>supply-demand random variable</u> Y_n whose absolute value equals the supply available at the node when Y_n assumes a nonnegative value and the demand required by the node when Y_n assumes a non-positive value. An important application of stochastic transportation networks, and one that motivated their consideration in this paper, is their use in models of electrical power networks.

The range of a random variable is a set consisting of values the random variable assumes with non-zero probabilities. The range of a random arc capacity is assumed to be a finite set of non-negative integers, and the range of a node's supply-demand random variable is assumed to be a finite set of integers. In general, then, the range of a node's supply-demand random variable may include both positive and negative integers so that the node may supply units in some realizations of the network and demand units in others. Based on the range of its supply-demand random variable Y_n , a node n is referred to in one of four ways: source, sink, intermediate node, or random source-sink. In particular, if the range of Y_n consists solely of a single value v (i.e., Y_n is a constant), the node is a source if v > 0, a sink if v < 0, and an intermediate node if v = 0; however, if the range of Y_n consists of at least two values, node n is a random source-sink. Reference is made to two special types of random source-sinks; a random source (random sink) is a random source-sink whose range includes only non-negative (non-positive) integers.

-2-

It is assumed that the joint probability mass function of the random variables of the network is known. The exact degree to which independence must be assumed is discussed briefly later in this introduction and in greater detail in Sections 4 and 5.

Given a realization of the stochastic transportation network, the realization is <u>feasible</u> (infeasible) and the network <u>functions</u> (fails) if a (no) flow exists satisfying the following constraints: (i) the flow in each arc is no greater than the value assumed by its random arc capacity, and (ii) for each node n with Y_n assuming a value v, the flow out of node n minus the flow into node n is at most v when $v \neq 0$ and equals 0 when v = 0. The fundamental problem is the computation of the <u>reliability</u>, that is, the probability that the random variables will assume values such that the stochastic transportation network functions.

A special class of stochastic transportation networks is that in which every arc capacity is a binary random variable, one node is a source having a supply of 1, one node is a sink having a demand of 1, and all other nodes are intermediate nodes. The literature refers to such networks by a variety of names; <u>binary reliability networks</u> is used herein.

As the bibliographies in [12] and [13] illustrate, there exists an extensive literature treating binary reliability networks. However, despite the importance of the problem, the literature treating the most general stochastic transportation network is scanty. The state-of-theart is best represented by the algorithm of Doulliez and Jamoulle [6] which computes the reliability of a stochastic transportation network by efficiently partitioning the set of all possible network realizations. The algorithm as presented in [6] assumes that each node of the network

-3-

is either an intermediate node, a sink, or a random source; however, in light of the procedure of Section 4 of this paper for eliminating random source-sinks, the algorithm can be applied to the most general stochastic transportation networks. An earlier algorithm of Doulliez [5] and a later algorithm of Pang and Wood [11] are similar to [6] but not as efficient.

A technique long used in analyzing binary reliability networks is modular decomposition (cf. [2] and [3]). Intuitively, modular decomposition reduces computational effort by first identifying a complex but specially structured subnetwork whose random variables are independent of the random variables outside the subnetwork and then replacing this complex subnetwork with a simpler subnetwork. This paper demonstrates that the technique of modular decomposition is also useful in analyzing stochastic transportation networks. After Section 2 introduces some additional notation and definitions, Sections 3 and 4 develop two types of modular decompositions: series-parallel-modular decomposition and block-modular decomposition. Section 5 explains the relationship between block-modular decomposition and the graph theoretic concepts of "cutnode" and "block", and Section 6 describes an algorithm used in block-modular decomposition. Finally, Section 7 discusses both computational aspects of block-modular decompositions and areas for future research.

2. ADDITIONAL NOTATION AND DEFINITIONS

The concepts of <u>undirected arc</u>, <u>path</u>, <u>cut</u>, and <u>flow</u> have their usual meanings (cf. [7], pp. 2-10). All arcs are undirected; nodes are adjacent if an arc of the network connects them, and an arc is <u>incident</u>

-4-

to the two nodes it connects. An arbitrary indexing of the arcs of the network permits reference to the arc connecting node i and node j and having index k in one of two ways: arc (i,j) or arc k.

Given a network N, the <u>subnetwork defined by a set of nodes</u> S consists of the nodes of S and every arc of N incident to a pair of nodes both in S. The subnetwork is <u>proper</u> if it is not N itself and is <u>nontrivial</u> if it contains more than one node.

"Random variable" and "probability mass function" are denoted by "r.v." and "p.m.f.", respectively; r.v.'s and p.m.f.'s denote their plurals. A r.v. of a stochastic transportation network is an <u>indepen-</u> <u>dent r.v.</u> if it is statistically independent of all other r.v.'s of the network. A subnetwork of the stochastic transportation network is an <u>independent subnetwork</u> if all its r.v.'s are statistically independent of the r.v.'s outside the subnetwork, even though they may have arbitrary dependence among themselves.

Given a stochastic transportation network having reliability R, another network having reliability R* is <u>c-equivalent</u>, where c is a known constant, if R = cR*. In cases where c=1, c-equivalent is shortened to <u>equivalent</u>.

Given a subset S of a set T having a finite number of elements, \overline{S} denotes its complement in T, |S| denotes its cardinality, and Pr[S] denotes its probability under some probability measure defined over T. A partition $\{S_k\}$ of S consists of disjoint subsets of S whose union equals S.

If every node i in a set S of nodes has a value v(i) associated with it, v(S) equals $\sum_{i \in S} v(i)$. For example, if the v(i) denotes the supplies (demands) at a set S of sources (sinks), then v(S) equals the

-5-

total supply (demand) within the set. Given two subsets of nodes S and T, (S,T) denotes the set of all arcs connecting a node of S to a node of T. If every arc (i,j) of the subset (S,T) has a value v(i,j) associated with it, v(S,T) equals $\Sigma_{(i,j)\in(S,T)}v(i,j)$. For example, if the v(i,j) denotes the capacities of the arcs of a cut (S,S), v(S,S) equals the capacity of the cut.

3. SERIES-PARALLEL-MODULAR DECOMPOSITION

Given a path between two nodes r and s, the <u>interior nodes</u> are all nodes of the path except r and s. A <u>series-module</u> (s-module) is a path between two nodes r and s that satisfies four conditions: (i) the path contains at least one interior node, (ii) every interior node is an intermediate node, (iii) every interior node of the path is adjacent only to other nodes of the path, and (iv) the n>2 arc capacities X_1, X_2, \ldots, X_n of the path are independent of all other r.v.'s of the network. An equivalent network results from the replacement of the arcs of the s-module by a single arc from r to s having capacity $\min[X_1, X_2, \ldots, X_n]$. Hereafter, <u>s-modular decomposition</u> refers to such a replacement.

A <u>parallel-module</u> (p-module) is a subnetwork consisting of a pair of nodes r and s joined by $n \ge 2$ arcs whose capacities X_1, X_2, \ldots, X_n are independent of all other r.v.'s of the network. An equivalent network results from the replacement of the p-module by a single arc from r to s having capacity $X_1+X_2+\ldots+X_n$. Hereafter, <u>p-modular decomposition</u> refers to such a replacement.

ş

A subnetwork is a <u>series-parallel module</u> (s-p-module) if it can be reduced to a pair of nodes joined by a single arc through a sequence of s-modular and p-modular decompositions. Hereafter, <u>s-p-modular de-</u> <u>composition</u> refers to such a reduction. The concept of s-p-modular decomposition as used here is a straightforward adaptation of a similar concept long used in the analysis of binary reliability networks and first defined by Bodin [4]. Both an s-module and a p-module are special cases of an s-p-module; Figure 1 contains a more complex s-p-module.

Hereafter, it is assumed without loss of generality that the original network under consideration contains no s-p-modules. This assumption simplifies both notation and computation.

4. BLOCK-MODULAR DECOMPOSITION

The 10-node, 15-arc transportation network of Figure 2 serves as an example throughout this section. Nodes 1, 6, and 10 are sources, each having a supply of 20 units; nodes 2, 3, 5, 7, 8, and 9 are sinks, each having a demand of 5; node 4 is the only intermediate node. The random arc capacities are independent and denoted by X_j , X'_j , and X''_j . for $1 \le j \le 5$, where each X_j takes on with equal probability one of the two values indicated in Figure 2 and where X_j , X'_j , and X''_j are identically distributed. Thus, the transportation network will be in one of 2^{15} equally likely states.

A <u>block-module</u> (b-module) is an independent, proper, and nontrivial subnetwork containing one and only one node (referred to as the <u>cut-</u> <u>node</u>) adjacent to nodes outside the subnetwork; a <u>minimal block-module</u> is a b-module containing no proper subnetworks that are also b-modules.

-7_

For example, in the network of Figure 2, the subnetworks defined by the three sets of nodes {1,2,3,4}, {4,5,6,7}, and {4,8,9,10} are three identical minimal b-modules all having node 4 as a cutnode; although the subnetwork defined by the set of nodes {1,2,3,4,5,6,7} is a b-module, it is not a minimal b-module. An explanation of the relationship of block-modules to the graph theoretic concepts of "cutnode" and "block" is delayed until Section 5.

In order to simplify the discussion, it is assumed hereafter that a b-module possesses two additional characteristics: (i) the cutnode is an intermediate node, and (ii) the b-module contains no random source-sinks. As will now be shown, these assumptions are made without loss of generality. If (i) is not true, an equivalent network results from rerouting all arcs of the b-module incident to the cutnode n into a new artificial intermediate node n' and then connecting n' and n by an artificial arc having infinite capacity; node n' then serves as the new cutnode for the b-module. Hereafter, such a procedure is referred to as inserting an artificial intermediate node n' into the subnetwork at the node n. If (ii) is not true, there exists a random source-sink) n with a supply-demand r.v. Y_n that can take on at least two values k for $-d \le k \le s$ where $d \ge 0$ and $s \ge 0$. An equivalent network with one less random source-sink results from changing node n into a sink with a constant demand of d, adding a new node n' having a constant supply of s + d, and adding a new arc (n',n) having a random capacity $X_n = Y_n + d$. When $X_n = k + d$ for $0 \le k \le s$, node n effectively becomes a source of k units since n' is able to supply the demand of d at n and still serve as a source of k units for the remainder of the network. Similarly, when $X_n = k + d$ for $-d \le k \le 0$, node n effectively becomes a sink for

-8-

٤

-k units since only d + k of its demand of d can be supplied by n'. Note that the elimination of the random source-sink does <u>not</u> require that the supply-demand r.v. Y_n be independent.

The definition of a b-module motivates an attempt to replace the b-module with a simpler subnetwork. Given a particular realization of the random capacities of the arcs of a b-module, one of three cases may (1) the sources within the b-module not only can meet all deoccur: mands within the b-module but also can "export" units via the cutnode to the remainder of the network, (ii) the demands within the b-module can only be met if it is possible to "import" units via the cutnode from the remainder of the network, or (iii) the demands within the bmodule cannot be met, regardless of how many units can be imported via the cutnode from the remainder of the network. Intuitively, then, with respect to the remainder of the network, the b-module acts in one of three ways: (i) a source, (ii) a sink with a demand that is possible to meet, or (iii) a sink with a demand that is impossible to meet. As will now be shown, a c-equivalent network results from changing the cutnode from an intermediate node into a random source-sink and deleting the remainder of the b-module from the network.

Consider a b-module N having cutnode n. N' denotes the subnetwork of the original network defined by the cutnode n and all nodes <u>not</u> belonging to the b-module N. (Actually, N' is itself a b-module.) S (T) denotes the set of nodes in N that are sources (sinks); for any source i, a(i) denotes its supply; for any sink i, b(i) denotes its demand. Suppose N contains r arcs having indices 1,2,...,r and random capacities $X = (X_1, X_2, ..., X_r);$ let $\Omega^+ = \{x | Pr[X=x] > 0\}$. X is the <u>state vector of</u> the b-module, and Ω^+ is the <u>state space of the b-module</u>.

-9-

Given a b-module N with cutnode n, N_k for $-\infty < k < \infty$ denotes a transportation network identical to N except that n is now either a source for k units if $k \ge 0$ or a sink for -k units if k < 0. For $-\infty < k < \infty$, let $H_k = \{x \in \Omega^+ | when X = x, N_k$ functions but N_{k+1} fails}; when $0 \le k < \infty$, H_k is a set of states for which the sources within N not only can meet all the demands within N but can also "export" at most k units to N' via n; when $-\infty < k < 0$, H_k is a set of states for which the demands within N but can also "export" at most k units to N' via n; when $-\infty < k < 0$, H_k is a set of states for which the demands within N can only be met if it is possible to "import" at least -k units from N' via n. Let k_{max} equal the maximum of zero and the largest value of k for which H_k is non-empty; similarly, let k_{min} equal the minimum of 0 and the smallest value of k for which H_k is non-empty. Note that $0 \le k_{max} \le max [0,a(S)-b(T)]$ and $-b(T) \le k_{min} \le 0$. Finally, let $H_{-\infty} = \{x \in \Omega^+ | x \in H_k \text{ does not hold for } min \le k \le k_{max}\}$; that is, $H_{-\infty}$ is a set of states for which the demands within N can never be met, regardless of how many units can be "imported" from N' via n.

It is clear that the subsets H_k for $k = -\infty$ and $k_{\min} \le k \le k_{\max}$ are a partition of Ω^+ . In the remainder of this section, it will be assumed that H_k and $\Pr[H_k]$ are known for $k = -\infty$ and $k_{\max} \le k \le k_{\min}$; Section 6 contains an efficient algorithm for their computation.

For example, in the network of Figure 2, consider the b-module defined by the subset of nodes {1,2,3,4}. As will be shown in Section 6,

 $H_{10} = \{(5,15,0,0,10); (5,15,0,5,10); (5,15,5,0,10); (5,15,5,5,10)\}$ $H_{5} = \{(0,15,5,0,10); (0,15,5,5,10); (0,15,0,5,10); (5,15,5,5,0)\}$ $H_{0} = \{(0,15,5,0,0); (0,15,5,5,0); (5,15,0,0,0); (5,15,0,5,0); (5,15,5,0,0)\}$ $H_{-5} = \{(0,15,0,5,0); (5,0,0,0,10); (5,0,0,5,10); (5,0,5,0,10); (5,0,5,5,10); (5,0,5,5,0)\}$

$$H_{10} = \{(0,0,0,5,10); (0,0,5,5,10); (0,0,5,0,10)\}$$

-10-

$$H_{\infty} = \{(0,0,0,0,0); (0,0,0,5,0); (0,0,5,0,0); (0,0,5,5,0); \\(0,0,0,0,10); (0,15,0,0,0); (0,15,0,0,10); (5,0,0,0,0); \\(5,0,0,5,0); (5,0,5,0,0)\}.$$

Because each of the 32 possible state vectors is equally likely to occur, $\Pr[H_{10}] = \frac{4}{32}$, $\Pr[H_5] = \frac{4}{32}$, $\Pr[H_0] = \frac{5}{32}$, $\Pr[H_{-5}] = \frac{6}{32}$, $\Pr[H_{-10}] = \frac{3}{32}$, and $\Pr[H_{-\infty}] = \frac{10}{32}$.

Since the entire network fails if X = x where $x \in H_{\infty}$, conditioning upon whether or not $x \in H_{\infty}$ results in $R = R^* (1 - \Pr[H_{\infty}])$ where R^* is the reliability of the entire network given $x \in H_{\infty}$ does <u>not</u> occur. Consider the network obtained as follows: (i) change the cutnode from an intermediate node into a random source-sink having a supply-demand r.v. Y_n for which $Pr[H_k]$

$$\Pr[\Upsilon_{n} = k] = -\begin{bmatrix} \frac{1}{1} - \Pr[H_{\infty}] & \text{for } k_{\min} \leq k \leq k_{\max} \\ 0 & \text{otherwise} \end{bmatrix}$$

and (ii) delete the remainder of the b-module from the network. Hereafter, <u>b-modular decomposition</u> refers to such a procedure. It is clear that b-modular decomposition results in new network having reliability R^* ; that is, the new network is $(1 - Pr[H_{-\infty}])$ -equivalent to the original network.

B-modular decomposition will consist of an additional step in those instances where the cutnode for the b-module is an artificial intermediate node n' inserted into the original b-module at a random source-sink n. In particular, an equivalent network results from changing the supply-demand r.v. of n to $Y_n + Y_n$, and then deleting the node n' and the arc (n',n) from the network.

Since a b-modular decomposition may create a new b-module in the revised network, the number of b-modular decompositions that can be performed neet not be limited to the number of b-modules in the original

-11-

network under consideration. Section 5 provides an upper bound on this quantity.

Provided the original network under consideration contains at least one b-module (and therefore two), there exists at least one sequence of k > 1 b-modular decompositions that reduces the original network to a revised network N containing no b-modules. Section 7 discusses the choice of such a sequence; for now, suppose it has made. Let n denote a node of N whose supply-demand r.v. is independent. The node that served as the cutnode for the immediately preceding b-modular decomposition is always one choice for n; for simplicity, it is assumed hereafter that n is always this node. After inserting an artificial intermediate node n' into the network at the node n, the node n' serves as the cutnode for one last b-modular decomposition. The resulting c-equivalent network consists only of the arc (n',n) and the two nodes n and n' having supply-demand r.v.'s Y and Y,, respectively; since this simple network behaves like a single random-source sink having supply-demand r.v. $Y_n + Y_n$, its reliability R^* equals $Pr[Y_n + Y_n] \ge 0$ and is easily computed as $R^* = \Sigma \Pr[Y_n, =k]\Pr[Y_n \ge -k]$ where the summation is taken over all k for which $Pr[Y_{n'} = k] > 0$. Having been reduced to a c-equivalent network consisting of a single node with reliability R^* , the original network has reliability $R = cR^*$.

For example, one possible evaluation of the reliability of the network of Figure 2 by a sequence of b-modular decomposition proceeds as follows:

(1) The subnetwork defined by the set of nodes {1,2,3,4} serves as the first b-module with node 4 serving as the cutnode. B-modular decomposition deletes all nodes and arcs of the b-module except node 4 and

-12-

results in a c-equivalent network where node 4 is now a random sourcesink having a supply-demand r.v. Y_4 . From the values given above for the $Pr[H_k]$, $c = (1 - Pr[H_\infty]) = \frac{22}{32}$ and $Pr[Y_4 = 10] = \frac{4}{22}$, $Pr[Y_4 = 5] = \frac{4}{22}$, $Pr[Y_4 = 0] = \frac{5}{22}$, $Pr[Y_4 = -5] = \frac{6}{22}$, and $Pr[Y_4 = -10] = \frac{3}{22}$.

(2) The next b-module consists of the subnetwork defined by the set of nodes {4,5,6,7} and node 4 again serves as the cutnode. This time the b-modular decomposition involves two stages since node 4 is a random source-sink. In the first stage, an artificial intermediate node 4' is inserted into the b-module at node 4 in order to serve temporarily as the cutnode for a b-module identical to the one analyzed in step (1). Deleting all nodes and arcs of the b-module except node 4' results in a $\left(\frac{22}{32}\right)^2$ -equivalent network where node 4' is now a random source-sink having an independent supply-demand r.v. Y_4 , identically distributed to Y_4 . The second stage of the b-modular decomposition combines nodes 4 and 4' into a single random source-sink having a supply-demand r.v. $W_4 = Y_4 + Y_4$, for which $\Pr[W_4 = 20] = \frac{16}{484}$, $\Pr[W_4 = 15] = \frac{32}{484}$, $\Pr[W_4 = 10] = \frac{56}{484}$, $\Pr[W_4 = 5] = \frac{88}{484}$, $\Pr[W_4 = 0] = \frac{97}{484}$, $\Pr[W_4 = -5] = \frac{84}{484}$, $\Pr[W_4 = -10] = \frac{66}{484}$, $\Pr[W_4 = -15] = \frac{36}{484}$, $\Pr[W_4 = -20] = \frac{9}{484}$.

(3) Since the revised network contains no b-modules, the final b-modular decomposition also involves two stages. In the first stage, an artificial intermediate node 4' is again inserted into the network at node 4 in order to serve temporarily as the cutnode for a b-module identical to the one analyzed in step (1). Deleting all nodes and arcs of the b-module except node 4' results in a $\left(\frac{22}{32}\right)^3$ -equivalent network where node 4' is now a random source-sink having an independent supply-demand r.v. W_4 , identically distributed to Y_4 . The second stage of the

-13-

b-modular decomposition combines nodes 4 and 4' into a single random source-sink having a supply-demand r.v. $W_4 + W_4$, and having a reliability $R^* = \Pr[W_4 + W_4, \ge 0]$; R^* is easily computed as

$$R^* = \Sigma_k = -10, -5, 0, 5, 10^{\Pr[W_4]} = k]\Pr[W_4 \ge -k] = \frac{6157}{10, 648}$$

Hence, the reliability R of the original network of Figure 2 is

$$R = cR^* = \frac{6157}{32,768}$$

As another example, consider the network of Figure 3 that arises when modeling the major electrical power transmission network serving northern California. The arcs represent 500kV transmission lines and the nodes represent generators and/or users of electricity. Whereas "s/d" within a node indicates it is a random source-sink of the most general type, "s" or "d" within a node indicate it is a random source or a random sink, respectively. The p.m.f.'s of the network's random arc capacities and random source-sinks are omitted here. The dotted lines in the figure outline subnetworks N₁, N₂, N₃, and N₄. A broad description of one possible sequence of b-modular decompositions that can be used to compute the network's reliability follows:

(1) N_1 and N_4 are both b-modules in the original network; two successive b-modular decompositions retain only the cutnodes N_1 and N_2 share with N_2 and N_3 , respectively.

(2) The revised network consists only of the newly created bmodules N_2 and N_3 having a common cutnode; two successive b-modular decompositions collapse this revised network into a single random source-sink located at the common cutnode.

It is interesting to consider the use of a sequence of b-modular decompositions to determine the feasibility of deterministic

transportation networks; that is, the special class of stochastic transportation networks in which every r.v. equals a known constant with probability one. In such a network, the state space of a b-module consists of a single state vector x. If $x \in H_{-\infty}$, the entire network must be infeasible; however, if $x \in H_k$ for $k_{\min} \leq k \leq k_{\max}$, b-modular decomposition simply adds k to the cutnode's supply-demand constant and deletes from the network all other nodes and all arcs of the b-module.

5. RELATIONSHIP OF BLOCK MODULES TO GRAPH THEORETIC CONCEPTS

This section may be omitted with little or no loss in continuity; it assumes an understanding of the following graph theoretic concepts as defined by Harary [9, pp. 8-37]: connectivity, component, degree of a node n (denoted by deg n), union of two or more networks, removal of a node n from a network N (denoted by N-n), tree, endnode (or leaf) of a tree, branch at a node of a tree, cutnode, nonseparable network, block, and block-cutnode tree. In this section, the last four of these concepts are adapted to apply to stochastic transportation networks and then used to describe b-modular decomposition.

Given a subnetwork N of a stochastic transportation network N* and a node n of N* not in N, N + n denotes the subnetwork N augmented by the node n and every arc of N* connecting n to a node in N. A node n of a stochastic transportation network N* is a <u>cutnode</u> if it satisfies two conditions: (i) the supply-demand r.v. Y_n is independent; (ii) N* - n consists of components N₁,N₂,...,N_p where $p \ge 2$ and the union of some proper subset of the set of subnetworks {N₁+n,N₂+n,...,N_p+n} is independent. Intuitively, a cutnode divides the network into independent

-15-

subnetworks having only the cutnode in common.

A subnetwork of a stochastic transportation network is <u>nonsepar-able</u> if it is connected, nontrivial, independent, and contains no cutnodes. A <u>block</u> is a maximal nonseparable subnetwork; that is, a block is a nonseparable subnetwork that is not a proper subnetwork of any other nonseparable network.

Let g-cutnode and g-block denote the graph theoretic definitions cutnode and block as distinguished from the definitions just given for stochastic transportation networks. Given a stochastic transportation network N*, let B_1, B_2, \ldots, B_m $(B'_1, B'_2, \ldots, B'_p)$ denote the m blocks (p g-blocks) of N*. It is well known that B'_1, B'_2, \ldots, B'_p is a division of N* into p subnetworks having two properties: (i) every arc of N* is in one and only one g-block, and (ii) every node that is not a cutnode is in one and only one g-block. It is easy to see that a similar statement is true for the cutnodes and the blocks B_1, B_2, \ldots, B_m . In general, the set of cutnodes is a subset of the set of g-cutnodes and each B_1 $(1 \le i \le m)$ is the union of one or more of B'_1, B'_2, \ldots, B'_p ; however, if every g-cutnode has an independent supply-demand r.v. and every g-block is independent, a node is a cutnode if and only if it is a gcutnode, and a subnetwork is a block if and only if a g-block.

The block-cutnode tree of stochastic transportation network has a definition identical to that in [9, p. 36]. Briefly, the block-cutnode tree contains a node corresponding to each cutnode of N* and a node corresponding to each block of N* with two nodes adjacent if and only if one node corresponds to a cutnode and the other corresponding to a block containing that cutnode.

-16-

B-modular decomposition can now be described in terms of the cutnodes, blocks, and the block-cutnode tree of the stochastic transportation network under consideration. In particular, a minimal b-module corresponds to a leaf of the block-cutnode tree and a nonminimal b-module having cutnode n corresponds to a union of at most [(deg n) - 1] branches at the node corresponding to n in the block-cutnode tree. Thus, repeated b-modular decompositions correspond to repeatedly "pruning" the blockcutnode tree. The number of blocks in the network is an upper bound on the number of b-modular decompositions that can be performed in computing the reliability. This upper bound is achieved during a reduction of the original network to a single node through a sequence of minimal bmodular decomposition; for such a decomposition, the weakest set of assumption about the r.v.'s of the network are: (i) every cutnode has an independent supply-demand r.v., and (ii) every block is independent.

An algorithm in Aho, Hopcroft, and Ullman [1] computes a graph's g-cutnodes and g-blocks (called <u>articulation points</u> and <u>biconnected com-</u> <u>ponents</u> therein) in a number of steps proportional to the number of arcs in the network. The algorithm is easily adapted to compute the cutnodes and blocks of a stochastic transportation network.

6. PARTITIONING THE STATE SPACE OF A BLOCK-MODULE

Consider a b-module N having a cutnode n, a set S of sources with supplies a(i) for $i \in S$, a set T of sinks with demands b(i) for $i \in T$, a state vector $X = (X_1, X_2, \dots, X_r)$, and state space Ω^+ . This section develops an algorithm for constructing the partition $\{H_k\}$ of Ω^+ needed to perform a block-modular decomposition. The algorithm is based on a decomposition

-17-

principle developed by Doulliez and Jamoulle in [6].

6.1 Overview of the Algorithm

Given vectors $\mathbf{m} = (\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_r)$ and $\mathbf{M} = (\mathbf{M}_1, \mathbf{M}_2, \dots, \mathbf{M}_r)$ having integer-valued components, the <u>interval having lower endpoint m and upper</u> <u>endpoint M</u> (denoted by $[\mathbf{m}, \mathbf{M}]$) is the set of vectors $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_r)$ having integer-valued components for which $\mathbf{m}_j \leq \mathbf{x}_j \leq \mathbf{M}_j$ for $1 \leq j \leq r$. Let $\Omega = [\mathbf{L}, \mathbf{U}]$ denote the smallest interval containing Ω^+ , where the size of an interval is measured by its cardinality. Such a Ω always exists since Ω^+ is a finite set of vectors all having integer-valued components. The algorithm in this section constructs a partition $\{\mathbf{G}_k\}$ of Ω , where \mathbf{G}_k is defined by replacing Ω^+ by Ω in the definition of \mathbf{H}_k . Clearly, $\mathbf{H}_k = \mathbf{G}_k \cap \Omega^+$ and $\Pr[\mathbf{H}_k] = \Pr[\mathbf{G}_k]$ for all k.

It is helpful to think of the algorithm as a branching process that produces a rooted tree. The root of the tree corresponds to Ω and every other node of the tree corresponds to an interval contained in Ω . Branching from a node corresponds to partitioning the interval into several smaller intervals. At the end of each iteration of the algorithm, the leaves of the tree correspond to a partition of Ω into intervals.

Associated with each interval is a label k; if $k \ge 0$, the cutnode n is considered to have a demand of k units, and, if $k \le 0$, n is considered to have a supply of -k units. An interval I with label k is either <u>fathomed</u> or <u>unfathomed</u>. If I has label k and is fathomed, $I \subseteq G_k$ and no further branching from I is necessary; that is, I will correspond to a leaf in the final tree produced by the algorithm. If I has label k and is unfathomed, although $I \cap G_j = \phi$ for j > k, further branching is necessary to determine if $I \cap G_k$ is non-empty.

-18-

Since $G_k = \phi$ for k > a(S) - b(T), Ω is given the label a(S) - b(T)at the algorithm's initialization; as indicated by the flow chart of Figure 4, each iteration of the algorithm consists of examining an interval I and, depending upon whether or not I $\cap G_k = \phi$, either reducing its label or partitioning it into one fathomed interval and several unfathomed intervals. Subsections 6.2-6.5 provide the details of the subroutines summarized in the flow chart, Subsection 6.6 discusses sensitivity analysis and modifications to the algorithm, and Subsection 6.7 contains an example.

6.2 The Feasibility Subroutine

Given the interval chosen is

$$I = [(m_1, m_2, ..., m_r); (M_1, M_2, ..., M_r)]$$

having label k, execute the following steps:

(a) Augment the b-module by joining each source $i \in S$ to a common fictitious source s with an arc (s,i), by joining each sink $i \in T$ to a common fictitious sink t with arc (i,t), and by joining the cutnode n to both the fictitious source s and sink t with arcs (s,n) and (n,t).

(b) Define the arc capacities $c(\cdot)$ or $c(\cdot, \cdot)$ of the augmented network as follows:

 $c(j) = M_{j} \text{ for } 1 \leq j \leq r,$ $c(s,i) = a(i) \text{ for } i \in S,$ $c(i,t) = b(i) \text{ for } i \in T,$ $c(s,n) = \max(-k,0) \text{ and } c(n,t) = \max(k,0).$

(c) Determine the max-flow from s to t in the augmented network using the Ford-Fulkerson labeling method [7, pp. 17-19] suitably modified to handle undirected arcs (cf. [7, p. 23] or [10, pp. 224-228]). For $1 \le j \le r$, denote the flow in arc j by f_j . Also, denote the min-cut by (P,\overline{P}) ; upon termination of the Ford-Fulkerson algorithm, P equals the set of labeled nodes, and \overline{P} equals the set of unlabeled nodes.

(d) It is well-known (cf. [7, pp. 38-39]) that the deterministic transportation network having arc capacities M_j for $1 \le j \le r$ and having a supply of a(i) at $i \in S$, a demand of b(i) at $i \in T$, and either a supply of -k at n if k < 0 or a demand of k at n if k > 0 is feasible if and only if the max-flow in the augmented network equals c(T,t) + c(n,t), or equivalently, $(T \cup n,t)$ is a min-cut. Hence, the interval I contains feasible points if and only if the max-flow in the max-flow in the augmented network equals c(T,t) + c(n,t).

6.3 The Partitioning Subroutine

Given the max-flow of c(T,t) + c(n,t) from s to t in the augmented network defined in the Feasibility Subroutine, execute the following steps:

(a) Let $v^{\circ} = (v_1^{\circ}, v_2^{\circ}, \dots, v_r^{\circ})$ where $v_j^{\circ} = \max(f_j, m_j)$ for $1 \le j \le r$. (b) For $1 \le j \le r$, define F_j

as the minimum possible flow on arc j given that the max-flow flow from s to t in the augmented network must equal c(T,t) + c(n,t). Then define $v^* = (v_1^*, v_2^*, \dots v_r^*)$ where $v_j^* = \max(F_j, m_j)$. Note that $v_j^* = v_j^0 = m_j$ if $f_j \leq m_j$ and that $v_j^* = v_j^0 = M_j$ if arc j is a member of min-cut obtained in part (c) of the Feasibility Subroutine. Thus, if either $f_j < m_j$ or arc j belongs to the min-cut, v_j^* can be determined immediately. However, if either of these two cases does not occur, F_j must be computed prior to determining v_j^* . The Ford-Fulkerson labeling method provides a convenient means for computing F_j . Suppose the direction of the flow in arc j is from node i to node k. Start the labeling process at node i and attempt to label node k; the only modification is that node k cannot be labeled directly from node i even if $f_j < M_j$. If a breakthrough of value ε occurs, increase by ε the flow in each forward arc of the flow augmenting path, decrease by ε the flow in each reverse arc of the flow augmenting path, and decrease by ε the flow in arc j. Thus, the value of the max-flow is unchanged but the flow in arc j has been reduced by ε . Repeat the labeling process until a breakthrough is is impossible. At termination, F_j is the current value of the flow in arc j. (Actually, since $v_j^* = \max(F_j, m_j)$, the algorithm can be terminated if the flow in arc j decreases to m_j or lower.)

(c) Partition the interval I into the intervals $A_{,B_{1},B_{2}},\ldots,B_{r}$, C_{1},C_{2},\ldots,C_{r} where

$$A = [(v_1^{o}, v_2^{o}, \dots, v_r^{o}); (M_1, M_2, \dots, M_r)]$$

and, for
$$1 \le j \le r$$
,

$$B_{j} = [(v_{1}^{*}, v_{2}^{*}, \dots, v_{j-1}^{*}, m_{j}, m_{j+1}, \dots, m_{r}); (M_{1}, M_{2}, \dots, M_{j-1}, v_{j}^{*} - 1, M_{j+1}, \dots, M_{r})]$$

$$C_{j} = [(v_{1}^{o}, v_{2}^{o}, \dots, v_{j-1}^{o}, v_{j}^{*}, v_{j+1}^{*}, \dots, v_{r}^{*}); (M_{1}, M_{2}, \dots, M_{j-1}, v_{j}^{o} - 1, M_{j+1}, \dots, M_{r})]$$
Note that $B_{j} = \phi$ if $v_{j}^{*} = m_{j}$ and $C_{j} = \phi$ if $v_{j}^{*} = v_{j}^{o}$. It is easy to show
that these subsets are indeed a partition of I, especially upon noting
that $\{B_{j}, 1 \le j \le r\}$ is a partition of

 $B \equiv \{x \in I \mid x_j < v_j^* \text{ for at least one } j\},\$

 $\{C_j, 1 \le j \le r\}$ is a partition of

$$C \equiv \{x \in I \mid x_k \ge v_k^* \text{ for } 1 \le k \le r \text{ and } x_j < v_j^o \text{ for at least one } j\},\$$

and {A,B,C} is a partition of I. If additional details are needed, consult [6].

(d) Given the definition of v^{o} and the fact that $x \ge v^{o}$ for all $x \in A$, it is clear that $A \subseteq G_{k}$; hence, assign the label k to interval A and consider it fathomed. Given the definition of v_{j}^{*} and the fact that $x_{j} < v_{j}^{*}$ for all $x \in B_{j}$, it is clear that $B_{j} \cap G_{k} = \phi$; hence, assign the label k-1 to B_{j} and consider it unfathomed. Because $x \ge v^{*}$ and $x_{j} < v_{j}^{o}$ for all $x \in C_{j}$, assign the label k to interval C_{j} but consider it unfathomed.

6.4 The Label Reduction Subroutine

If the Feasibility Subroutine determines that the interval I contains no feasible points, execute the appropriate step below:

(a) If $n \in P$, reset the label of I to $-\infty$ and consider I fathomed.

(b) If $n \in \overline{P}$, reset the label of I to k-D where D is defined as the amount the max-flow in the augmented network falls short of c(T,t) + c(n,t).

The justification for the above steps will now be provided.

It is well-known (cf. [7, pp. 38-39]) that \overline{P} - t is a subset of nodes in the transportation network for which the difference between the net demand within the subset and the total capacity of all arcs entering the subset is the strictly positive amount D, thus violating a

-22-

necessary condition for feasibility in the theorem due to Gale [8]. In particular, one of the following must hold:

(1) If $n \in P$, then

$$b(T \cap \overline{P}) - a(S \cap \overline{P}) - c(P-s,\overline{P}-t) = D > 0.$$

In this case, step (a) is appropriate since the excess net demand in \overline{P} - t can never be satisfied, regardless of the degree to which the boundary node n can serve as a supply.

(2) If $n \in \overline{P}$, then

$$b(T \cap \overline{P}) + k - a(S \cap \overline{P}) - c(P-s,\overline{P}-t) = D > 0.$$

In this case, step (b) is appropriate since if k were reduced by at least D, the interval I might then contain some feasible points. If k > 0 (k<0), reduction of k is equivalent to decreasing the demand (increasing the supply) at the cutnode.

6.5 The Termination Subroutine

When all existing intervals have been fathomed, execute the following steps:

(a) For each interval I, compute Pr{I}. If the arc capacities are independent,

$$\Pr\{I\} = \prod_{j=1}^{r} \left(\sum_{\substack{x_j = m \\ j}}^{M_j} \Pr\{X_j = x_j\} \right);$$

if not, use the joint p.m.f. of X to compute

$$\Pr\{I\} = \sum_{x_1=m_1}^{M_1} \sum_{x_2=m_2}^{M_2} \cdots \sum_{x_r=m_r}^{M_r} \Pr\{X = (x_1, x_2, \dots, x_r)\}.$$

(b) For $k = -\infty$ and $-b(T) \le k \le a(S) - b(T)$, G_k is simply the union of the non-overlapping intervals having label k. Hence, compute $Pr\{G_k\}$ by

$$\Pr\{G_k\} = \sum_{\{I \mid I \text{ has label } k\}} \Pr\{I\}.$$

6.6 Sensitivity Analysis and Modifications to the Algorithm

As long as L and U remain unchanged, sensitivity analysis requiring only changes in the joint p.m.f. of X is easy to perform since it will not affect the partition $\{G_k\}$ produced by the algorithm, even though $Pr{G_{L}}$ may change for some values of k. However, this property may be obtained at the expense of additional and perhaps significant computation time if many intervals I result for which $Pr{I} = 0$. Hence, in some cases (e.g., when $|\Omega^+|$ is significantly less than $|\Omega|$), it will probably be more beneficial to reduce computational effort by insuring inductively that $Pr{I} > 0$ for each interval I. More specifically, if $Pr{X_i = v_i^0}$ = 0 ($\Pr\{X_i = v_i^*\}$ = 0) after defining $v_i^o(v_i^*)$ in the Partitioning Subroutine, redefine v_i^o (v_i^*) by increasing it until a value v_i is reached for which $\Pr\{X_{j} = v_{j}\} > 0$. Then, given v^o and v^{*}, if $\Pr\{X_{j} = v_{j}^{o} - 1\} = 0$, $(\Pr{X_i = v_i^* - 1} = 0)$, instead of using $v_i^o - 1$ $(v_i^* - 1)$ in the Partitioning Subroutine as the j-th component of the upper limiting state space defining the interval C_j (B_j), use the largest value v_j less than $v_j^o - 1$ (v_i^*-1) for which $\Pr\{X_i = v_i\} > 0$. Hereafter, this modification of the algorithm will be referred to as Modification A. Although the modified algorithm will no longer produce a partition of Ω , the subsets will still be non-overlapping and their union will include each $x \in \Omega^+$. Note also that the subsets produced by the algorithm with Modification A are insensitive to changes in the joint p.m.f. of X provided Ω^+ remains unchanged.

-24-

Changes in the joint p.m.f. of X affecting L and U can frequently be analyzed without applying the algorithm to the entire revised state space. For example, suppose that after considering the state space $\Omega = [L,U]$, sensitivity analysis leads to a revised state space $\Omega^* = [L,U+\Delta]$ where $\Delta = (\Delta_1, \Delta_2, \dots, \Delta_r)$ and $\Delta \ge 0$. Since Ω^* can be partitioned into the non-overlapping intervals $\{\Omega, I_1 I_2, \dots, I_r\}$ where, for $1 \le j \le r$, $I_j = [(L_1, L_2, \dots, L_{j-1}, U_j + 1, L_{j+1}, \dots, L_r); (U_1, U_2, \dots, U_{j-1}, U_j + \Delta_j, \dots, U_r + \Delta_r)]$, revision of $\Pr\{H_k\}$ is possible upon application of the algorithm to each I_i for $1 \le j \le r$. Note that $I_i = \phi$ if $\Delta_i = 0$.

Changes resulting in a revised state space of the form $\Omega^* = [L, U-\Delta]$, $\Omega^* = [L+\Delta, U]$, or $\Omega^* = [L-\Delta, U]$ can be treated in a similar manner. However, changes resulting in simultaneous increases and decreases in U and/ or L are more difficult to handle; rather than reapplying the algorithm to Ω^* , it may be possible to treat such changes more efficiently by viewing them as a sequence of changes of the four types just discussed.

Let N' denote the subnetwork of the original network defined by the cutnode and all nodes <u>not</u> belonging to N. Assume without loss of generality that N' contains no random source-sinks. Let S' (T') denote the set of nodes of N' that are sources (sinks); for any source 1, a(1) denotes its supply; for any sink i, b(1) denotes its demand. The partition of Ω produced by the above algorithm is insensitive to all changes made within N'; however, this peoperty is obtained at the expense of computation time that would be unnecessary if it were known with certainty that a(S') and b(T') would remain constant regardless of the sensitivity analysis conducted in N' and that b(T') < a(S) - b(T) or a(S') - b(T') < b(T). In particular, the following modifications result

-25-

in a reduction in computational effort: (i) if b(T') < a(S) - b(T), assign the label b(T') to Ω at the initialization step and (ii) if a(S') - b(T') < b(T), fathom an interval when its label falls below b(T') - a(S') in part (d) of the Partitioning Subroutine or part (b) of the Label Reduction Subroutine.

6.7 An Example

Figure 5 contains the tree produced by the algorithm when it is applied using Modification A to the subnetwork of Figure 2 defined by the set of nodes $\{1,2,3,4\}$. Within a node of the tree, the second and third lines contain the interval's upper endpoint M and lower endpoint m, respectively; the first line of node contains the interval's initial label and any subsequent changes (denoted by \rightarrow). The remainder of Figure 5 is self-explanatory. Examination of the intervals corresponding to leaves of the tree confirms the partition $\{H_k\}$ given in the example of Section 4.

7. CONCLUSION

Provided the stochastic transportation network under consideration contains at least one block-module, block-modular decomposition is an alternative to existing methods for computing the network's reliability. Instead of computing the reliability of one large network, it is possible to analyze a sequence of smaller subnetworks. Block-modular decomposition is particularly useful in the analysis of large electrical power networks.

A computer code implementing block-modular decomposition is under

development. Upon its completion, experiments will provide empirical answers to the following questions:

(1) In comparison to existing methods, when does a sequence of block-modular decompositions result in a significant reduction in computation time?

(2) In performing a sequence of block-module decompositions that reduce the original network to a single node, are there heuristic rules for selecting the next block-module if the objective is to minimize the total computation time? (For example, one possibility is to always select the minimal block-module whose state space has the smallest cardinality.)

REFERENCES

- [1] Aho, A. V., Hopcroft, J. E., and Ullman, J. D., <u>The Design and</u> <u>Analysis of Computer Algorithms</u>, Addison-Wesley, Reading, Mass., 1974, pp. 179-187.
- [2] Birnbaum, Z. W. and Esary, J. D., "Modules of Coherent Binary Systems," <u>SIAM Journal of Applied Mathematics</u>, Vol. 13, No. 2 (June 1965), pp. 444-462.
- Bodin, L. D., "Approximations to System Reliability Using a Modular Decomposition," <u>Technometrics</u>, Vol. 12 (May 1970), pp. 335-344.
- [4] Bodin, L. D., "The Catalogue Ordering Problem," Technical Report ORC 67-39, Operations Research Center, University of California, Berkeley, Ca. (1967).
- [5] Doulliez, P., "Probability Distribution Function for the Capacity of a Multiterminal Network," <u>Revue Francaise d'Automatique, Infor-</u> matique et Recherche Operationnelle, Vol. 1 (1971), pp. 39-49.
- [6] _____ and Jamoulle, E., "Transportation Networks with Random Arc Capacities," <u>Revue Francaise d'Automatique, Informatique et</u> <u>Recherche Operationelle</u>, Vol. 3 (November 1972), pp. 45-59.
- [7] Ford, L. R. and Fulkerson, D. R., <u>Flows in Networks</u>, Princeton University Press, Princeton, New Jersey, 1962.
- [8] Gale, D., "A Theorem on Flows in Networks," <u>Pacific Journal of</u> <u>Mathematics</u>, Vol. 7 (1957), pp. 1073-1082.
- [9] Harary, F., Graph Theory, Addison-Wesley, Reading, Mass., 1969.
- [10] Hillier, F. S. and Lieberman, G. J., <u>Operations Research</u>, Holden-Day, Inc., San Francisco, Ca., 1974.

-28-

- [11] Pang, C. K. and Wood, A. J., "Multi-area Generation System Reliability Calculations," <u>IEEE Transactions on Power Apparatus and</u> <u>Systems</u>, Vol. PAS-94, No. 2 (March/April 1975), pp. 508-517.
- [12] Shogan, A. W., "Sequential Bounding of the Reliability of a Stochastic Network," <u>Operations Research</u>, Vol. 34, No. 6 (November/ December 1976), pp. 1027-1044.
- [13] Shogan, A. W., "A Decomposition Algorithm for Network Reliability Analysis," <u>Networks</u>, Vol. 8, No. 3 (1978), pp. 231-251.







Figure 2



Figure 3

