Expert Systems with Applications 37 (2010) 968–973

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

A novel algorithm to evaluate the performance of stochastic transportation systems

Yi-Kuei Lin *

Department of Industrial Management, National Taiwan University of Science and Technology, Taipei 106, Taiwan

ARTICLE INFO

Keywords:
Minimal cuts
Transportation system
Weighted capacity
Cost attribute
Multicommodity

ABSTRACT

Network analysis is a usual approach to evaluate the performance of real-life systems such as transportation systems. We construct a multicommodity stochastic flow network with weighted capacity allocation to model the transportation systems. Each arc with cost attribute has several possible capacities. The capacity weight, the consumed quantity of arc capacity by per commodity, varies with the arcs and types of commodity. Nevertheless, the system capacity is not appropriate to be treated as the maximal sum of the commodity. We define the system capacity as a demand vector **d** if the system fulfills at most **d**. The main problem of this work is to measure the quality level of a transportation system. We propose a performance index, the probability that the upper bound of the system capacity equals a demand vector **d** subject to the budget constraint. A novel algorithm based on minimal cuts is presented to generate all maximal capacity vectors meeting exactly the demand **d** under the budget B. The performance index can then be evaluated in terms of such vectors.

1. Introduction

Network analysis is a usual approach to evaluate the service quality of real-life systems such as transportation systems, logistics systems, computer systems, telecommunication systems and manufacturing systems. For a single-commodity binary-stage flow network, each arc is binary-stage with a designated capacity which will have zero level only due to any failure. The system capacity is the maximum flow from the source s to the sink t. Several authors (Aggarwal, Chopra, & Bajwa, 1982; Lee & Yum, 1999; Lee, 1980) proposed algorithms to calculate the system reliability, the probability that the system capacity is not less than the demand.

Extending to the general case that the arc is stochastic to reflect that the arc may be in failure, partial failure, maintenance, reserved by other competitors, etc., the arc capacity has several possible values and the system capacity is not a fixed number. Such a network is stochastic and called a single-commodity stochastic flow network (SSFN). Without cost attributes, several authors (Lin, Jane, & Yuan, 1995; Xue, 1985; Yeh, 1998, 2001b, 2005) presented algorithms to generate all lower boundary points for the demand d in order to evaluate the performance index, the probability that the lower bound of the system capacity equals d. In general, such a performance index is called the system reliability. A lower boundary point for d is a minimal capacity vector meeting demand d, and it can be derived in terms of minimal paths (MP) where a MP is a set of arcs whose proper subsets are no longer paths.

The max-flow min-cut theorem (Ford & Fulkerson, 1962) states that the maximum flow from s to t equals the minimum capacity among all minimal cuts (MC). If we remove some arcs from the network and thus he source disconnects the sink, then the set of such arcs is called a cut. A MC is a cut that will not be a cut after removing any arc from it. Jane, Lin, and Yuan (1993); Lin (2006b) and Yeh (2001a) used MC to generate all upper boundary points for d in order to evaluate another performance index, the probability that the upper bound of the system capacity equals d. Such a performance index is usually called the system unreliability. An upper boundary point for d is a maximal capacity vector meeting exactly the demand d, equivalently, it is a maximal system state such that the system capacity exactly equals d.

Moreover, many flow networks allow $p (p \geqslant 2)$ types of commodity to be transmitted from s to t simultaneously. Although many researchers (Fard & Lee, 1999; Ford & Fulkerson, 1974; Held, Wolfe, & Crowder, 1974; Hu, 1963; Jarvis, 1969; Rothechild & Whinston, 1966) solved the multicommodity maximum flow problem to find the maximal sum of commodity under the assumption that the arc capacity is deterministic, the system capacity is not suitable to be treated as the maximal sum of the commodity if different type of commodity consumes the arc capacity differently. For example, as shown in Table 1, the sum of the commodity in network A is larger than that in network B. This fact does not mean that network A provides the better transmission ability while commodity 2 consumes more capacity than commodity 1 does. Network B provides totally 40 units of capacity which is more than that of network A.

This paper constructs a multicommodity stochastic flow network (MSFN) to model a stochastic transportation system. We propose a novel algorithm to evaluate the performance for a MSFN.

* Fax: +886 2 27376344.
E-mail address: yklin@mail.ntust.edu.tw

0957-4174/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2009.05.080

Table 1
The total flow for two networks.

	Network A	Network B	Capacity weight
Flow of commodity 1	10	5	2
Flow of commodity 2	6	10	3
Sum of the commodity	16	15	
Total capacity consumed	38 $(10\times 2+6\times 3)$	40 $(5\times 2+10\times 3)$	

The remainder of this work is organized as follows. In Section 3, the system capacity is firstly define and then a new performance index, the probability that the upper bound of the system capacity equals a given vector $\mathbf{d}$ subject to the budget B, is proposed. Such a performance index is designed to measure the service quality of a MSFN. The MC are utilized to study the relationship among flow assignments and arcs capacities. The novel algorithm in terms of MC is presented in Section 4 to generate all lower boundary points for $(\mathbf{d}, B)$, the maximal capacity vector meeting exactly demand $\mathbf{d}$ under the budget B. In Section 5, an international transportation example between manufacturing factories is shown to illustrate the proposed algorithm and how the performance index be calculated. Time complexity and storage complexity of the proposed algorithm are both analyzed in Section 6.

2. Assumptions and vector operations

Let $G \equiv (A, C, M, W)$ be a MSFN with weighted capacity allocation where $A \equiv \{a_i | 1 \leqslant i \leqslant n\}$ the set of arcs, $C \equiv (c_1, c_2, \ldots, c_n)$ with c_i the cost per unit of capacity on a_i, $M \equiv (M_1, M_2, \ldots, M_n)$ with M_i the maximal capacity of a_i, and $W = \{w_i^k | 1 \leqslant i \leqslant n, 1 \leqslant k \leqslant p\}$ with w_i^k the consumed quantity of capacity on arc a_i by per commodity k. Hence, each arc a_i has two attributes: M_i and c_i. The current capacity (or residual capacity) of arc a_i is denoted by x_i, and $X \equiv (x_1, x_2, \ldots, x_n)$ denotes the capacity vector. Apparently, the capacity weight varies with arcs and types of commodity. Without loss of generality for w_i^k, we assume that $w_i^p \geqslant w_i^{p-1} \geqslant \cdots \geqslant w_i^1 \geqslant 1$ for all arcs. G is required to further satisfy the following assumptions:

1. The flows in G must observe the flow-conservation law (Ford & Fulkerson, 1962).
2. The capacities of different arcs are statistically independent.
3. x_i is an integer-valued random variable which takes the integer value from $\{0, 1, 2, \ldots, M_i\}$ according to a given probability distribution.
4. Each node is perfectly reliable.
5. The capacity weight is a positive integer.
6. All p types of commodity are transmitted from s to t.

Vector operations are defined as follows.

$Y \geqslant X (y_1, y_2, \ldots, y_n) \geqslant (x_1, x_2, \ldots, x_n)$: $y_i \geqslant x_i$ for each $i = 1, 2, \ldots, n$;
$Y > X (y_1, y_2, \ldots, y_n) > (x_1, x_2, \ldots, x_n)$: $Y \geqslant X$ and $y_i > x_i$ for at least one i;
$\mathbf{d}' \geqslant \mathbf{d} (d'_1, d'_2, \ldots, d'_p) \geqslant (d_1, d_2, \ldots, d_p) : d'_k \geqslant d_k$ for each $k = 1, 2, \ldots, p$;
$\mathbf{d}' > \mathbf{d} (d'_1, d'_2, \ldots, d'_p) > (d_1, d_2, \ldots, d_p) : (d'_1, d'_2, \ldots, d'_p) \geqslant (d_1, d_2, \ldots, d_p)$ and $d'_k \geqslant d_k$ for at least one k;
$\mathbf{d}' + \mathbf{d} (d'_1, d'_2, \ldots, d'_p) + (d_1, d_2, \ldots, d_p) : (d'_1 + d_1, d'_2 + d_2, \ldots, d'_p + d_p)$.

3. A MSFN model

Suppose $Z_1, Z_2, \ldots$, and Z_m are m MC. With respect to each MC $Z_r = \{a_{r1}, a_{r2}, \ldots, a_{rn_r}\}$ where n_r is the number of arcs in Z_r, the vector $F_r = \left(F_r^1, F_r^2, \ldots, F_r^p\right)$ is called a flow assignment where $F_r^k = (f_{r1}^k, f_{r2}^k, \ldots, f_{rn_r}^k)$ with f_{rj}^k denoting the flow of commodity k through a_{rj}, $j = 1, 2, \ldots, n_r$, $k = 1, 2, \ldots, p$. It is feasible under the capacity vector $X = (x_1, x_2, \ldots, x_n)$ if

$$\sum_{k=1}^{p} w_{rj}^k f_{rj}^k \leqslant x_{rj} \quad \text{for } j = 1, 2, \ldots, n_r. \tag{1}$$

This inequality says that the capacity of a_{rj} consumed by all commodities cannot exceed the current capacity x_{rj}.

Under the capacity vector X, the MC Z_r is called to support the demand $\mathbf{d} = (d_1, d_2, \ldots, d_p)$ if there exists an F_r such that

$$\sum_{j=1}^{n_r} f_{rj}^k = d_k \quad \text{for } k = 1, 2, \ldots, p, \tag{2}$$

where d_k is the demand of commodity k at sink, $k = 1, 2, \ldots, p$. Under X, Z_r is said to support at most $\mathbf{d}$ if Z_r supports $\mathbf{d}$ but Z_r cannot support $(\mathbf{d} + e_1)$ (i.e., no F_r under X supports $(\mathbf{d} + e_1)$) where e_1 is a p-tuple vector that has 1 at position 1 and 0 at others. If Z_r cannot support $(\mathbf{d} + e_1)$, then Z_r cannot support $(\mathbf{d} + e_i)$ for any i since commodity 1 consumes the least quantity of capacity. The capacity vector X is defined to support $\mathbf{d}$ if under X, all MC support $\mathbf{d}$. Furthermore, X is defined to support at most $\mathbf{d}$ (i.e., X supports $\mathbf{d}$ but cannot support any $\mathbf{d}'$ with $\mathbf{d}' > \mathbf{d}$) if under X, all MC support $\mathbf{d}$ and at least one MC supports at most $\mathbf{d}$.

Definition 1. For a MSFN, the system capacity denoted by $L(X)$ is defined to be *d if X supports at most d*.

Two performance indexes $\Pr\{L(X) \geqslant \mathbf{d}\}$ and $\Pr\{L(X) \leqslant \mathbf{d}\}$ can be adopted to measure the quality level of a MSFN; the former is the probability that the lower bound of the system capacity equals $\mathbf{d}$, and the latter is the probability that the upper bound of the system capacity equals $\mathbf{d}$. In terms of MP, Lin (2001) proposed an algorithm to evaluate $\Pr\{L(X) \geqslant \mathbf{d}\}$ without budget constraints. This paper utilizes MC to evaluate $\Pr\{L(X) \leqslant \mathbf{d}\}$ but subject to budget B. That is, we study how to evaluate a new performance index $\Phi_{d,B} \equiv \Pr\{X | L(X) \leqslant \mathbf{d}$ and $C(X) \leqslant B\}$ where $C(X) \equiv \sum_{i=1}^{n} c_i x_i$ is the total cost under X.

3.1. Lower boundary points for $(\mathbf{d}, B)$ and performance index $\Phi_{d,B}$

Definition 2. A lower boundary point for (d, B) is the maximal vector in the set $\{X | L(X) = d$ and $C(X) \leqslant B\}$. Equivalently, X is a lower boundary point for (d, B) if (i) $L(X) = d$, (ii) $C(X) \leqslant B$, and (iii) $L(Y) > d$ *or* $C(Y) > B$ for any capacity vector Y with $Y > X$.

The set of lower boundary points for $(\mathbf{d}, B)$ is also the set of maximal vectors in $\{X | L(X) \leqslant \mathbf{d}$ and $C(X) \leqslant B\}$. If $X_1, X_2, \ldots, X_v$ are lower boundary points for $(\mathbf{d}, B)$, let $B_i = \{X | X \leqslant X_i\}$ for $i = 1, 2, \ldots, v$. Thus, the performance index $\Phi_{d,B}$ can be computed in terms of all lower boundary points for $(\mathbf{d}, B)$ by the following equation,

$$\begin{aligned} \Phi_{d,B} &= Pr\{X | L(X) \leqslant \mathbf{d} \text{ and } C(X) \leqslant B\} = \Pr\{X | X \\ &\leqslant X_i \quad \text{for a lower boundary point } X_i \text{ for } (\mathbf{d}, B)\} \\ &= \Pr\{B_1 \cup B_2 \cup \cdots \cup B_v\}. \end{aligned} \tag{3}$$

It can be calculated by applying methods such as the inclusion–exclusion (Griffith, 1980; Lin, 2001, 2004, 2006a, 2006b, 2007), disjoint subsets (Lin, 2004, 2006b; Soh & Rai, 2005; Xue, 1985), or state-space decomposition (Jane et al., 1993; Lin, 1998; Lin et al., 1995). Note that $\Pr\{Y \leqslant X\} = \Pr\{y_1 \leqslant x_1\} \times \Pr\{y_2 \leqslant x_2\} \times \cdots \times \Pr\{y_n \leqslant x_n\}$ by assumption 2. In particular, the inclusion–exclusion shows that $\Pr\{B_1 \cup B_2 \cup \cdots \cup B_v\} = \sum_i \Pr\{B_i\} - (-1)^2 \sum_{i<j} \Pr\{B_i \cap B_j\} - (-1)^3 \sum_{i<j<k} \Pr\{B_i \cap B_j \cap B_k\} - \cdots - (-1)^v \Pr\{B_1 \cap B_2 \cap \cdots \cap B_v\}$. An efficient way to calculate $\Phi_{d,B}$ is to generate all lower boundary points for $(\mathbf{d}, B)$ firstly.