Input Format for Multicommodity Flow Problems

Kimberly L. Jones
Irvin J. Lustig
Program in Statistics and Operations Research
Department of Civil Engineering and Operations Research
Princeton, New Jersey 08544
EMAIL: `irv@basie.princeton.edu`

November, 1992

1 Introduction

This document describes a proposed input format for the multicommodity network flow problem. This format allows for different definitions of commodities such as products (apples and oranges), trees (flow into a destination), or origin-destionation pairs, and reduces the required storage space by eliminating redundancy. It is hoped that this input format becomes the standard among all researchers studying multicommodity network flows and encourages the accumulation of data sets for benchmark purposes.

Since it is believed that the nature of a commodity can affect the performance of a solution algorithm, problems can be specified in a number of different ways. See Jones *et al* [1] for a discussion of these effects with respect to Dantzig-Wolfe decomposition.

The first section uses some notation to define the different formulations of multicommodity flow problems. Traditionally, it was up to the person formulating the problem to label a commodity as a product, flow into a destination, or as an origin-destination pair. The format allows all of these formulations.

2 Problem Formulation

Consider the following linear program:

$$
\begin{array}{lrcll}
\textbf{(MCNF):} & \text{minimize} & & \displaystyle\sum_{\kappa=1}^{\mathcal{K}} (c^\kappa)^T x^\kappa & \\
\text{subject to} & & & & \\
& B^\kappa x^\kappa & = & d^\kappa & \kappa \in \mathcal{K}, \\
& \displaystyle\sum_{\kappa=1}^{\mathcal{K}} x_{ij}^\kappa & \leq & u_{ij}, & \forall (i,j) \in A, \\
& x_{ij}^\kappa & \geq & 0, & \forall (i,j) \in A, \kappa \in \mathcal{K}.
\end{array}
\tag{1}
$$

The linear program in (1) is a generic formulation for a MCNF problem where

$\mathcal{K}$ is the set of commodities;

B^κ is the node-arc incidence matrix over the network $G = (N, A)$ for commodity κ;

x^κ is a vector of arc flows for commodity κ;

d^κ is a vector of supply/demand requirements for commodity κ; and,

u_{ij} is the upper bound on arc (i, j) over the commodities κ.

The linear program in (1) displays the typical block-angular structure of a MCNF problem in that the first set of constraints are the network flow conservation constraints for each commodity and the second set of constraints are the mutual capacity, or *bundle* constraints over the commodities. For ease of discussion, only models without individual capacity constraints on each x_{ij}^κ are considered.

We discuss three formulations of the MCNF problem where in each a commodity is defined in a unique way. In the first formulation, we define a commodity as a product that travels between a specific origin and a specific destination. This problem is termed the *origin-destination problem* (ODP) where κ represents the triplet (k, s, t) such that k is the product, s is the specific origin, and t is the specific destination. The (ODP) is representative of a crew scheduling problem, where the identity of a crew member k must be maintained while satisfying origin and destination constraints. In the second formulation, we define a commodity as a product that travels to a specific destination from multiple origins, or vice versa, from a specific origin to multiple destinations. This problem is termed the *destination specific problem* (DSP) where the commodity κ is the pair (k, t) that identifies the product k with the specific destination t. The (DSP) is typically seen in the traffic assignment problem, where vehicles k must be routed through a network from multiple origins to a common destination t. Finally, in the third formulation, we define a commodity as a product that must travel through a network from multiple origins to multiple destinations. This problem is termed the *product specific problem* (PSP), where the commodity κ represents the singleton k as the product. The (PSP) is representative of the multi-fleet allocation problem where a specific equipment type k must handle a set of transportation services. The (PSP) is the formulation that is traditionally indicated when referring to a MCNF problem.

Note that the constraint matrix for all three formulations is of the form

$$\begin{pmatrix} B^1 & & & \\ & B^2 & & \\ & & \ddots & \\ & & & B^{|\mathcal{K}|} \\ I & I & \cdots & I \end{pmatrix}, \tag{2}$$

where

$$\mathcal{K} = \begin{cases} \{(k,s,t) : d^\kappa = d^k_{st} \neq 0\} = \mathcal{K}(ODP) & \text{for (ODP)}, \\ \{(k,t) : d^\kappa = d^k_t \neq 0\} = \mathcal{K}(DSP) & \text{for (DSP)}, \\ \{k : d^\kappa = d^k \neq 0\} = \mathcal{K}(PSP) & \text{for (PSP)}. \end{cases} \tag{3}$$

The differences among the formulations are the number of *primary blocks* $|\mathcal{K}|$ and the density of the right-hand side vector d^κ. The differences in $|\mathcal{K}|$ are noted in (3) where the (ODP) has potentially many more primary blocks than (DSP) and (PSP) due to the combinatorial aspects of all the products, origins, and destinations in the problem. If we assume that the underlying network is identical for all commodities and identical for each of the three formulations, then the node-arc incidence matrix $B^\kappa = B$ for all $\kappa \in \mathcal{K}$. Hence, the number of constraints in each primary block is the same for each of the three formulations, but the number of primary blocks varies according to $|\mathcal{K}|$.

The right-hand side vector d^κ varies significantly among the formulations. This is seen by investigating the number of non-zeros in each d^κ for the three formulations. For the (ODP) and a given triplet (k,s,t) corresponding to commodity κ, there are only two nonzeros in each $d^\kappa = d^k_{st}$ that correspond to the origin s and the destination t. For the (DSP) and a given pair (k,t) corresponding to commodity κ, each $d^\kappa = d^k_t$ contains multiple nonzeros for the supply constraints at the multiple origins and a single nonzero for the demand constraint at the specific destination t. Finally, for the (PSP) and a given product k corresponding to commodity κ, $d^\kappa = d^k$ contains multiple nonzeros for the supply/demand constraints at the multiple origins and multiple destinations. Hence, the (ODP) has potentially many more primary blocks than (DSP) or (PSP) in the linear programming formulation, but d^κ is significantly sparser than in the other two formulations. This is an important concept in the formulation of the MCNF problem and plays a significant role in a decomposition procedure.

3 Input Format Overview

The following input format is proposed. The data will be split into 4 sections: a node file for general information about the network, a link file, a supply/demand file, and a mutual capacity file. The naming conventions for a problem with name `xxx` are to name the 4 files `xxx.nod`, `xxx.arc`, `xxx.sup` and `xxx.mut`, respectively. Each of these files will be detailed in this report along with examples as to how it will be general enough to handle all different types of multicommodity network flow problems and how it should significantly condense the amount of redundant information.

4 Node File

The node file contains four integers specified in the following order: *number of products, number of nodes, number of links, and the number of bundled links.* Note that for typical problems of type (DSP) and (ODP), the number of products will typically be 1. In this case, the number of commodities (either number of destinations or number of origin-destination pairs) would be inferred from the supply/demand file (See Section 6). The following example states that the multicommodity network has 4 products, 48 nodes, and 150 links, of which 80 are bundled.

Node Input Format
4
48
150
80

5 Link File

The link file is specified in the following order: *fromnode, tonode, product, cost, individual capacity, origin, destination, and mutual capacity pointer.* The mutual capacity pointer is an integer specifying the instance of the bundled link. The mutual capacity upperbounds are found in the *mutual capactiy file* which is described in section 7. To indicate that the information in a specific column is to be replicated over all possible values for that column or is not applicable to the problem, the value of -1 is placed in the column. This will be further explained in the 2 examples below.

The first example lists the information for 2 links in a 4 commodity network where a commodity is specified to be a product, such as apples or oranges. Since origin-destination pairs are not applicable to this problem, there is a -1 in every instance for these 2 columns. The link $1 \rightarrow 47$ is not bundled which is indicated by a 0 in the column for mutual capacties. This link is duplicated 4 times since the cost on the link is different for all commodities. The second link, $2 \rightarrow 48$, is the first bundled link and has a cost of -300 for all 4 commodities. The second example is an origin-destination specific problem where a commodity is a specific origin-destination pair. For this example, a -1 is in every instance of the product column indicating that this information is irrevalent. However, if the OD pairs were decomposed from a different type of formulation, this information might still be relevent and thus the column would be useful. The link $1 \rightarrow 47$ is a link which is not bundled and has a cost of 98 and an individual capacity of 40 for all OD pairs. The link $3 \rightarrow 30$ is not bundled and has a cost of 20 for all OD pairs that have a destination of node 47. Lastly, the link $3 \rightarrow 31$ is the first bundled link for this network and has a cost of -340 for all OD pairs that have a destination of node 48.

Link File Input Format

Product Specific							
From	To	Prod	Cost	Ind Cap	Origin	Dest	Mut Ptr
1	47	1	98	40	-1	-1	0
1	47	2	110	50	-1	-1	0
1	47	3	60	46	-1	-1	0
1	47	4	45	49	-1	-1	0
2	48	-1	-300	-1	-1	-1	1

Origin-Destination Specific							
From	To	Prod	Cost	Ind Cap	Origin	Dest	Mut Ptr
1	47	-1	98	40	-1	-1	0
3	30	-1	20	-1	-1	47	0
3	31	-1	-340	-1	-1	48	1

6 Supply/Demand File

The supply/demand file is specified in the following order: *origin, destination, product, and flow*. This type of input format is able to handle whether the commodities are apples and oranges (products), flow into a destination or out of an origin (trees), or origin-destination pairs (paths). For the cases where commodities are products or trees, the placement of a -1 in the origin or destination column indicates whether a node is a supply or demand node. The usual way of indicating a demand is to place a negative value in the flow column. However, if we require all values in the flow column to be positive, a -1 in the origin column will indicate that it is a demand. This enables us to always expect positive elements in the final column, as would be the case if it were an origin-destination problem. A -1 in the product column has the same meaning as in the arc file: to duplicate over all possible values or the information is not applicable. In the case of the (ODP) formulation, when both the origin and destination columns have positive elements, there may be a -1 in the product column indicating that this is not relevant information. But if there is a positive element, it may mean that the OD pair is also tagged by a product or tree commodity.

The following examples below show how the supply/demand file for all three definitions of a commodity can be used in an identical format. First, we give the example of a 2 commodity (product) network with 2 supply nodes and 2 demand nodes.

Supply/Demand Input File for Product Formulation			
Origin	Destination	Product	Flow
1	-1	1	25
1	-1	2	50
2	-1	1	15
2	-1	2	30
-1	3	1	15
-1	3	2	60
-1	4	1	25
-1	4	2	20

The next example is that for the tree formulation where a commodity is flow into a destination. This example has is a 2 commodity network with 2 destinations.

Supply/Demand Input File for Tree Formulation			
Origin	Destination	Product	Flow
1	-1	1	25
2	-1	1	15
3	-1	1	10
-1	5	1	50
1	-1	2	10
2	-1	2	30
3	-1	2	10
-1	6	2	50

The final example is for an 6 commodity OD formulation, where the product column is not applicable so a -1 is present in every instance of that column.

Supply/Demand Input File for OD Formulation			
Origin	Destination	Product	Flow
1	5	-1	25
2	5	-1	15
3	5	-1	10
1	6	-1	50
2	6	-1	40
3	6	-1	10

7 Mutual Capacity File

The file for the mutual capacities contain 2 columns in the following order: *mutual capacity pointer, and upper bound.* The pointer values are listed in sequential order from 1 to the number of bundled links in the network. These bundled links are referenced from the links file, as outlined in section 5. Below is an example of a mutual capacity file that has 5 bundled links.

Mutual Capacity Input Format	
Pointer	Bound
1	200
2	159
3	249
4	100
5	218

References

[1] K. L. Jones, I. J. Lustig, J. M. Farvolden, and W. B. Powell, *Multicommodity network flows: The impact of formulation on decomposition*, Tech. Rep. SOR 91-23, Princeton University, Department of Civil Engineering and Operations Research, Princeton, NJ, 1991.