MODELING AND SOLVING THE CAPACITATED NETWORK LOADING PROBLEM

Thomas. L. Magnanti,
Prakash Mirchandani,
and Rita Vachani

OR 239-91 January 1991

MODELING AND SOLVING THE CAPACITATED NETWORK LOADING PROBLEM

Thomas L. Magnanti
Sloan School of Management, MIT, Cambridge, MA 02139

Prakash Mirchandani
Katz Graduate School of Business, University of Pittsburgh, Pittsburgh, PA 15260

Rita Vachani
GTE Laboratories, Waltham, MA 02254

January 1991

Abstract

This paper studies a topical and economically significant capacitated network design problem that arises in the telecommunications industry. In this problem, given point-to-point demand between various pairs of nodes of a network must be met by installing (loading) capacitated facilities on the arcs. The facilities are chosen from a small set of alternatives and loading a particular facility incurs an arc specific and facility dependent cost. The problem is to determine the configuration of facilities to be loaded on the arcs of the network that will satisfy the given demand at minimum cost. Since we need to install (load) facilities to carry the required traffic, we refer to the problem as the *network loading problem.*

In this paper, we develop modeling and solution approaches for the problem. We consider two approaches for solving the underlying mixed integer programming model: (i) a Lagrangian relaxation strategy, and (ii) a cutting plane approach that uses three classes of valid inequalities that we identify for the problem. In particular, we show that a linear programming formulation that includes the valid inequalities always approximates the value of the mixed integer program at least as well as the Lagrangian relaxation bound (as measured by the gaps in the objective functions). We also examine the computational effectiveness of these inequalities on a set of prototypical telecommunications data. The computational results show that the addition of these inequalities considerably improves the gap between the integer programming formulation of the problem and its linear programming relaxation: for 6 – 15 node problems from an average of 25% to an average of 8%. These results show that strong cutting planes can be an effective modeling and algorithmic tool for solving problems of the size that arise in the telecommunications industry.

In this paper, we study a problem that is becoming increasingly important in the telecommunications industry: given an organization's forecast for data and voice traffic between its various locations, what configuration of transmission facilities between the locations (nodes) will provide the necessary link capacities to carry this traffic at minimum cost? A similar problem arises in the context of transportation planning; in this setting, the traffic corresponds to freight and the transmission facilities to different types of trucks. These problems have substantial economic significance. For example, revenues to the long distance carriers from the lease of digital transmission circuits used for private communications networks are about \$1.7 billion per annum currently (*Business Communications Review*, May 1990). These revenues are generated by over 60,000 circuits and that number is expected to grow at 30 to 40% per annum by one estimate (*Telecommunications*, North American Edition, May 1990). In the transportation context, the total expenditure on trucking is estimated to reach \$276.3 billion in 1990 and expected to grow at an annual rate of 7.5% (*US Industrial Outlook*, US Department of Commerce, January 1990).

Despite the importance of these network design applications in a variety of settings, the available research on them is quite limited. The objective of this paper is to develop modeling and solution approaches for these problems. Since the models we consider are special versions of more general capacitated network design problems, we hope that this paper might also provide some useful insights for solving the notoriously difficult, general capacitated network design problem.

Before presenting a formal description of the problem that we study, we describe the telecommunications private network leasing problem that motivated this paper. Private lines are transmission facilities that customers lease from a telephone company for their exclusive use. These lines are billed on a fixed (non-usage sensitive) rate. Customers lease them for a variety of reasons. For example, from a cost perspective, rather than pay on a per usage basis, an organization might find it cheaper to lease a private line facility between any two locations that have a large amount of traffic between them. In addition, private networks offer customers greater flexibility to reconfigure the network to accommodate changes in traffic patterns, provide improved reliability, and offer higher operational control than the public, switched network. Due to rapid technological changes in the telecommunications industry, telephone companies are offering higher bandwidth (capacity) facilities to private subscribers which allow the customers to use the private networks for a variety of applications, including voice, data and video transfer. As a result, the demand for private lines has been increasing rapidly and is expected to continue to do so over the next five to ten years.

Private networks have dedicated access lines from the customer premises to the nearest telephone company switch (central office) and dedicated lines between the central offices that connect different locations of the customer's organization. The telephone industry refers to the inter central office part of the network as the backbone network. In order to send a message from location A to location B on the private network, the network must contain a path whose arcs all have the required amount of transmission capacity. Since we assume that the subscriber uses circuit switching (as opposed to packet switching) to transmit the traffic between the locations, the system uses an equal amount of capacity in both directions on all the arcs of that path. (We refer the reader to Bertsekas and Gallager, 1987 for more technical details of these concepts.)

The digital facilities that a customer leases to and between the central offices are selected from a small set of alternatives — for example, DS0 (Digital Signal Level 0), DS1 (Digital Signal Level 1), and in some cases, DS3 (Digital Signal Level 3) facilities. A DS0 facility allows the transmission of 64 kilo bits per second (kbps), the bandwidth that is required to transmit one voice call. A DS1 facility transmits at the rate of 1.54 Mega bits per second (Mbps), or offers the capacity equivalent to 24 DS0 facilities, and a DS3 facility is equivalent (in capacity terms) to 28 DS1 facilities. The tariffs for these facilities are complex; for each service type, the tariff is roughly proportional to the length of the link and the availability of several facilities of different capacities introduces strong economies of scale. Typically, a DS1 circuit, which is equivalent in capacity to 24 DS0 circuits, costs the same as only 8 to 10 DS0 circuits.

The cost of any private network corresponds to the leasing cost of the facilities installed on the arcs; the user incurs no additional routing cost. A fundamental problem that arises when designing a private network is to determine the configuration of leased facilities on the backbone links that will satisfy the projected demand at minimum cost. This problem is difficult because of the complexity of the cost structure. The optimal solution might use complicated routes for the different commodities: by aggregating traffic on some arcs, it will take advantage of the economies of scale in the tariff structure. Though researchers have successfully solved variations of the uncapacitated network design problem (for example, Balakrishnan, Magnanti and Wong, 1989), the general capacitated network design problem has proven to be considerably more difficult. The objective of this paper is to develop modeling and solution approaches for the network loading problem, which is a special version of the general capacitated network design problem. The network loading problem includes the private network leasing problem as a special case and arises in several other application contexts as well. For example, in the transportation industry, the facilities might represent trucks of fixed size and a slight variation of the model would prescribe a load plan (the assignment of trucks to routes) and the loading of freight onto trucks; see Powell and Sheffi (1983) or Leung, Magnanti and Singhal (1991).

The rest of the paper is organized as follows. Section 1 presents a formal description and formulation of the network loading problem. In Section 2, we discuss alternative solution strategies for the problem and provide motivation for our proposed solution approach. Section 3 provides a partial characterization of the mixed integer polyhedron that models the problem, and Section 4 describes our solution methodology and presents our computational results. The last section presents our conclusions, briefly discusses extensions to the model, and suggests some future research directions.

1. Network Loading Problem: Description and Formulation

The network loading problem models the design of capacitated networks for which (i) the variable flow costs are zero, and (ii) facilities of fixed capacity are available to carry flow. We can install (load) these facilities on any of the arcs of the network. The problem is to determine the number of facilities to be loaded on each of the arcs of the network to meet given point-to-point demand at minimum cost. In this paper, we assume that only two types of facilities are available. In general, we may have a choice of facilities with capacities at many different levels; in the concluding section, we indicate how to extend our results for the case of multiple facilities. In the context of the private network leasing problem, the two facilities correspond to DS0 and DS1 circuits, which are the facilities most widely available. Our model extension would permit us to consider emerging industrial practice in which telephone companies are beginning to offer DS3 facilities to private subscribers on selected segments.

We model the network loading problem with two facilities, which we refer to as the TFLP (for the **T**wo **F**acility **L**oading **P**roblem), as follows.

TFLP:

$$\text{minimize} \sum_{\{i,j\} \in A} (a_{ij}x_{ij} + b_{ij}y_{ij})$$

subject to:

$$\sum_{j \in N} f_{ji}^{k} - \sum_{j \in N} f_{ij}^{k} = \begin{cases} -d_k \; if \; i = O(k) \\ d_k \; if \; i = D(k) \\ 0 \; otherwise \end{cases} \quad for \; all \; i \in N, \; for \; all \; k \in K \tag{1}$$

$$\sum_{k \in K} \left(f_{ij}^{k} + f_{ji}^{k}\right) \le x_{ij} + Cy_{ij} \quad for \; all \; \{i,j\} \in A \tag{2}$$

$$x_{ij}, y_{ij} \ge 0 \; and \; integer \; for \; all \; \{i,j\} \in A; \; f_{ij}^{k}, f_{ji}^{k} \ge 0 \; for \; all \; \{i,j\} \in A, \; for \; all \; k \in K. \tag{3}$$

In this formulation, N denotes the set of nodes of the network, A the set of arcs, and K the set of commodities; commodity k has origin O(k), destination D(k), and demand d_k. We refer to the two types of facilities as the low capacity (LC) and the high capacity (HC) facilities; the LC facility has capacity 1 and the HC facility has capacity C. (For the telecommunications private network leasing problem with DS0 and DS1 facilities, C = 24.) The formulation contains two sets of variables: (i) design variables x_{ij} and y_{ij} that define the number of LC and HC facilities loaded on the undirected arc {i, j}, and (ii) flow variables f_{ij}^{k} that model the flow of commodity k on arc {i, j} in the direction i to j. The coefficients a_{ij} and b_{ij} represent the cost of loading a single LC and HC facility, respectively, on arc {i, j} and the objective function minimizes the total cost incurred in loading all the facilities. Constraints (1) correspond to the usual flow conservation constraints for each of the commodities at each node. The capacity constraints (2) model the requirement that the total flow (in both directions) on an arc cannot exceed the capacity loaded on that arc.

This formulation seeks to minimize the cost of the installed facilities. The model assumes that it is sufficient to install just enough capacity to meet demand and that we do not need to provide extra capacity to address reliability issues. Although this assumption is valid in situations such as the transportation of freight, it might be less so in the telecommunications setting where the reliability of the networks is of greater concern. (For an approach to telecommunications network survivability problems, see Groetschel and Monma, 1988.) However, our discussions with planners in the telecommunications industry indicate that the model we are considering is valuable as a first-cut design tool for their planning activities and that they would typically use other ancillary models to address reliability issues.

2. Modeling and Solution Approaches: Motivation

Our approach to modeling and solving the two facility loading problem is rooted in discoveries made in mathematical programming over the past two decades. Over this period, many studies of integer programming in a variety of application contexts have established that the selection of a "good" model for a problem can have a profound effect upon the performance of solution methods.

These type of modeling results assume two forms. First, in many situations, it is possible to formulate a model with different sets of decision variables and constraints. For example, it is possible to model many fixed charge network flow problems with either a small or a large number of commodities. As an example, if all the flow in a problem originates at a single node s, we can either formulate a model with a single commodity originating at node s and with demand at all other nodes, or we can formulate the problem as a multicommodity flow problem with a separate commodity defined from node s to every other node in the network. The more disaggregate formulation, if we can solve its large-scale linear programming relaxation effectively, has proved to be a much better model. Wolsey (1989) highlights the importance of this type of modeling issue for several classes of problems. Balakrishnan, Magnanti and Wong (1989) demonstrate the computational advantages of using the more disaggregate multicommodity flow formulation for the uncapacitated network design problem by showing that solving a model with approximately 2 million variables and 2 million constraints, if even approximately, is much better than solving a more aggregate model with about 45,000 variables and 2500 constraints. Wong (1984, 1980) establishes the same result in the context of the network Steiner tree and the traveling salesman problems. For related results, see Geoffrion and Graves (1974), Magnanti and Wong (1984), and Martin (1987).

The second modeling approach is embodied in the burgeoning field of polyhedral combinatorics which attempts to improve the linear programming approximation to an integer programming problem by adding (strong) valid inequalities, either a priori to the original formulation of the problem, or dynamically, via a cutting plane approach, to a series of linear programming models. The success in solving the classical nonbipartite matching problem is a landmark example illustrating the power of this modeling approach as are the strides made in recent years in solving the traveling salesman problem (see

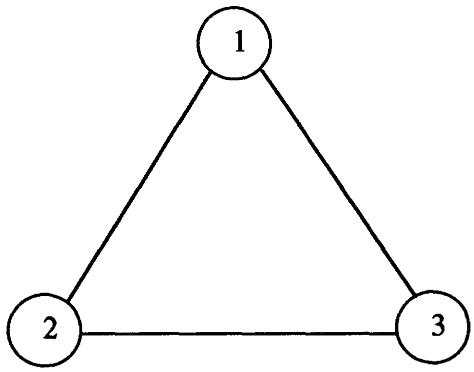
Groetschel and Padberg, 1985). The many applications of the cutting plane approach include the economic planning and linear ordering problem (Groetschel et al., 1985a, b), production planning models (Barany, Van Roy and Wolsey, 1984, Magnanti and Vachani, 1990), the fixed charge problem (Padberg, Van Roy and Wolsey, 1983), the lot sizing problem (Leung, Magnanti and Vachani, 1989, and Pochet, 1988), the spin glass problem (Barahona et al., 1988), and models for planning capacity expansion in local access telecommunications systems (Balakrishnan, Magnanti and Wong, 1990a, b). See Hoffman and Padberg (1985), Nemhauser and Wolsey (1988), and Pulleyblank (1989) for a general account of this methodology.

We use both of these modeling approaches to address our problem. We formulate the problem as a disaggregate multicommodity flow problem and we also identify a number of valid inequalities for the problem and in fact show that they are the best possible in the sense that they are facets of the underlying mixed integer polyhedron that models the problem. We also show that when applied to representative telecommunications data, the addition of these inequalities considerably improves the gap between the objective function values of the integer programming model of the problem and its linear programming relaxation: from an average of 25% on a set of small (6 – 15 nodes), but still practical, problems to an average of less than 8%.

It is easy to see that, in general, the linear programming relaxation of the TFLP provides a weak lower bound for the problem (for computational evidence, see Section 4.4 of this paper). In general, the linear programming lower bounds are weak for most capacitated network design problems and so it is much more difficult to solve these models than the uncapacitated network design problem. Our objective is to develop stronger formulations for the TFLP than its linear programming relaxation and, therefore, to develop more efficient solution techniques than a linear programming based branch and bound procedure.

An Example

To illustrate and compare different approaches for obtaining stronger lower bounds, let us consider the three-node problem shown in Figure 1. This example assumes that only HC facilities are available with $C = 24$, that the facility cost is the same on all the arcs, i.e., $b_{ij} = 20$ for all $\{i, j\}$, and that the demand between every pair of nodes (nodes 1 and 2, nodes 1 and 3, and nodes 2 and 3) is the same, i.e., $d_k = \delta$ for all k.



$C = 24$
$b_{ij} = 20$ for all $\{i, j\}$
$d_k = \delta$ for all k

Figure 1. Three-node example

The optimal solution to this problem depends on δ. If $\delta \bmod (24) \leq 12$ (we use the convention that $\delta \bmod (C) = C$ if δ is an integer multiple of C), then the optimal solution loads $\lfloor \delta/24 \rfloor$ facilities on any one of the arcs of the network and $\lceil \delta/24 \rceil$ facilities on the other two arcs for a total cost of $60*\lfloor \delta/24 \rfloor + 40$. If $\delta \bmod (24) > 12$, then the optimal solution loads $\lceil \delta/24 \rceil$ facilities on all three arcs with a corresponding cost of $60*\lceil \delta/24 \rceil$. In the solution of the linear programming relaxation of the formulation for this example, $y_{ij} = \delta/24$ for all $\{i, j\}$ and the corresponding optimal objective value is $60*\delta/24$. The gap between the optimal solution value for the problem and the value of its linear programming relaxation could be large depending upon the value of δ. To obtain better lower bounds, we consider two different approaches.

First, consider a *Lagrangian relaxation approach* to solving the TFLP. When using this approach, we can dualize either constraints (1) or (2). If we relax constraints (2), because the resulting Lagrangian subproblem is a network flow problem which satisfies the integrality property (i.e., its linear programming relaxation has an integer optimal solution), the Lagrangian dual problem gives the same lower bound as the linear programming relaxation of TFLP (Geoffrion, 1974). On the other hand, if we relax constraints (1) using multipliers v_i^k, then the resulting Lagrangian subproblem, with $v_{O(k)}^k = 0$, is:

$$\text{minimize} \sum_{\{i,j\} \in A} \left\{ a_{ij} x_{ij} + b_{ij} y_{ij} + \sum_{k \in K} (f_{ji}^k - f_{ij}^k)(v_i^k - v_j^k) \right\} + \sum_{k \in K} v_{D(k)}^k d_k$$

subject to: (2) and (3).

To this problem, we add the following *upper bound* constraints which do not affect the optimal objective value of the original formulation (since we can always delete the flow around cycles for any commodity), but improve the Lagrangian lower bound; we refer to the resulting Lagrangian subproblem as **P(LAG)**.

$$f_{ij}^k + f_{ji}^k \leq d_k \qquad \text{for all } \{i,j\} \in A, \text{ for all } k \in K. \tag{4}$$

Note that by dualizing the mass balance constraints (1), we have decoupled the problem into separate subproblems, one for each arc of the network. The subproblem P(LAG) does not satisfy the integrality property and, therefore, we can expect the lower bound obtained from the Lagrangian dual to be stronger than that obtained from the linear programming relaxation of TFLP. The subproblem for each arc can be solved efficiently by an incremental strategy of "loading" the "profitable" commodities (relative to the facility costs) on each arc. Vachani (1988) uses this Lagrangian relaxation strategy, with subgradient optimization to update the Lagrange multipliers and improve the Lagrangian bound, to solve the TFLP. Her results show that the lower bounds from using this approach indeed improve upon the linear programming relaxation bounds. For the three-node example of Figure 1 as well, the Lagrangian dual value improves upon the linear programming relaxation value; however, the gap between the Lagrangian lower bound and the optimal solution value varies with δ. For example, if $\delta = 12$, then the Lagrangian dual value is equal to the optimal solution value of 40 (with a choice of $v_1^1 = v_1^2 = v_2^3 = 0$ for the origin nodes, $v_2^1 = v_3^2 = v_3^3 = 10/9$ for the destination nodes, and $v_3^1 = v_2^2 = v_1^3 = 5/9$), whereas if $\delta = 13$, the Lagrangian dual value is 42.16 (with a choice of $v_1^1 = v_1^2 = v_2^3 = 0$ for the origin nodes, $v_2^1 = v_3^2 = v_3^3 = 40/37$ for the destination nodes, and $v_3^1 = v_2^2 = v_1^3 = 20/37$) which is considerably lower than the optimal value of 60.

The second, *polyhedral approach* to obtaining better lower bounds uses results about the polyhedral structure of the problem to strengthen the formulation. To illustrate this approach, again consider the example of Figure 1. Since the demand between node 1 and the other two nodes of the network (nodes 2 and 3) is 2δ units, the network must contain at least $\lceil 2\delta/24 \rceil$ HC facilities between node 1 and the other two nodes to carry this traffic. Thus, the constraint

$$y_{12} + y_{13} \geq \lceil 2\delta/24 \rceil \tag{5a}$$

is valid for the problem. Similar constraints for the other two nodes

$$y_{12} + y_{23} \geq \lceil 2\delta/24 \rceil \tag{5b}$$

$$y_{13} + y_{23} \geq \lceil 2\delta/24 \rceil \tag{5c}$$

are also valid. Note that the linear programming relaxation of the TFLP requires only that

$$y_{12} + y_{13} \geq 2\delta/24$$

and thus, the three (cutset) constraints (5a, 5b and 5c) will strengthen the linear programming relaxation if $\lceil 2\delta/24 \rceil$ is significantly larger than $2\delta/24$. For example, if δ = 12, i.e., 2δ is a multiple of 24, then constraints (5) are not effective at all since the solution to the linear programming relaxation satisfies these inequalities. On the other hand, if δ = 13, then the addition of these three constraints in the linear programming relaxation is sufficient to obtain an optimal integer solution.

The example of Figure 1 shows that the two solution approaches result in stronger lower bounds than the linear programming relaxation, though their performance depends upon the value of δ; one of them is more effective when δ = 12 whereas the other is more effective for δ = 13. However, changing the demand to 8 units shows that neither the Lagrangian approach nor including constraints (5) in the linear programming relaxation are sufficient for obtaining a good lower bound. With δ = 8, the optimal solution value for the linear programming relaxation is 20, the optimal integer solution has value 40; the Lagrangian lower bound is 30, and the enhanced linear program with constraints (5) has the same linear programming bound of 30. To obtain a linear programming formulation with an optimal solution cost of 40, we need to identify additional valid inequalities. The following (three-partition) inequality, obtained by adding constraints (5), dividing the aggregate constraint by 2 and then rounding up the righthand side to the next nearest integer since the lefthand side of the inequality is integral, serves this purpose

$$y_{12} + y_{13} + y_{23} \geq \lceil 1/2 \{ \lceil (16)/24 \rceil + \lceil (16)/24 \rceil + \lceil (16)/24 \rceil \} \rceil = 2. \tag{6}$$

Our discussion of the two different solution approaches raises the following questions: (i) can we identify situations in which one approach is likely to provide better lower bounds than the other and, more importantly, (ii) can we combine the two approaches to obtain lower bounds stronger than those that would be obtained from using either approach by itself? This paper provides a partial answer to these questions and develops one way of combining the two approaches.

As is evident from the three-node example, constraints (5) strengthen the linear programming relaxation of the TFLP when their righthand sides are not multiples of C. The linear programming relaxation of the TFLP provides sufficient capacity on all arcs to carry the flow; however, the y_{ij} variables might be fractional. Including inequalities (5) eliminates some of these fractional solutions, but not all of them. On the other hand, in the solution to the Lagrangian subproblem, the values of the y_{ij} variables will be integer, but since the Lagrangian subproblem relaxes the flow conservation constraints, it does not guarantee sufficient capacity on the arcs to carry all the demand of the original problem. These observations show that the Lagrangian dual and polyhedral solution approaches are complimentary and suggest that incorporating information from the Lagrangian subproblem into the linear programming-based approach that includes valid inequalities for the problem might prove useful in eliminating additional fractional solutions.

The strategy that we adopt to combine the two approaches is to identify facets of the Lagrangian subproblem that are valid inequalities (in fact, facets) for the TFLP. We use these inequalities to strengthen the formulation of the TFLP. These inequalities apply to individual arcs of the network (corresponding to the Lagrangian subproblem) and relate the flow of the commodities on the arc with the capacity on the arc; in contrast (the cutset) inequalities (5) apply to a set of arcs across a cutset. In fact, we identify inequalities that together with constraints (2) and (4) completely characterize the convex hull of the feasible solutions to the Lagrangian subproblem and show that our method of combining the two solution approaches for the TFLP guarantees a lower bound that is at least as strong as that obtained from either method independently. The next section provides our main technical results and the following section then discusses the solution method in more detail.

3.0 Polyhedral Results

In this section, we discuss the computational complexity of the TFLP and formally define three classes of inequalities — cutset, arc residual capacity, and three-partition — two of which we illustrated earlier for situations with only HC facilities. We prove that these inequalities are valid for the TFLP, and that they define facets of the underlying polyhedron. We then prove that if we include the arc residual capacity inequalities and the upper bound constraints in the linear programming relaxation of problem [TFLP], the resulting linear programming lower bound is at least as strong as that obtained using Lagrangian relaxation. We defer the proofs of most of these results to the Appendix.

As indicated by the next proposition, the TFLP is difficult from a computational complexity point of view. To show this result, we reduce the combinatorial 3 partition problem (which is known to be strongly NP-complete, Garey and Johnson, 1979) to the TFLP; thus, the TFLP belongs to the class of strongly NP-hard problems.

Proposition 1. *TFLP is strongly NP-hard.*

Proof. See the Appendix.

Since the TFLP is NP-hard, we do not expect to be able to provide a complete characterization of the convex hull of its feasible solutions (Groetschel, Lovasz and Schrijver, 1981; Karp and Papadimitriou, 1982). However, as our computational results in the next section show, the partial characterization that we obtain is sufficient to reduce the integrality gap significantly. For details of polyhedral terminology used in this paper, we refer the reader to the books by Schrijver (1986) and Nemhauser and Wolsey (1988).

Let Conv(TFLP) denote the convex hull of feasible solutions to TFLP. We next establish the dimension, denoted by dim(Conv(TFLP)), of Conv(TFLP). The formulation TFLP contains $2*|A| + 2*|A|*|K|$ variables and $(|N|-1)*|K|$ nonredundant equality constraints. Therefore, $\dim(\text{Conv(TFLP)}) \leq 2*|A| + 2*|A|*|K| - (|N|-1)*|K|$. Proposition 2 shows that dim(Conv(TFLP)) is exactly equal to this bound. This proof uses arguments similar to those used in Theorem 4 and we therefore omit it.

Proposition 2. *Dim(Conv(TFLP))* $= 2*|A| + 2*|A|*|K| - (|N|-1)*|K|$.

Suppose

$$\sum_{\{i,j\}\in A} \alpha_{ij}x_{ij} + \sum_{\{i,j\}\in A} \beta_{ij}y_{ij} + \sum_{k\in K}\sum_{\{i,j\}\in A} \left(\gamma_{ij}^{k}f_{ij}^{k} + \gamma_{ji}^{k}f_{ji}^{k}\right) \geq \delta \tag{7}$$

represents any valid inequality for Conv(TFLP). Furthermore, let $(\mathbf{x}^1, \mathbf{y}^1, \mathbf{f}^1)$ belong to Conv(TFLP). The fact that $(\mathbf{x}^2, \mathbf{y}^1, \mathbf{f}^1)$ also belongs to Conv(TFLP) whenever $\mathbf{x}^2 \geq \mathbf{x}^1$, implies that $\alpha_{ij} \geq 0$ for all $\{i,j\}$. Similarly, $\beta_{ij} \geq 0$ for all $\{i,j\}$. Notice that this argument does not apply for the coefficients γ_{ij}^{k} or γ_{ji}^{k} since the flow conservation constraints restrict the flows in the network: increasing the flow on one arc might require us to change the flow on some other arc. Thus, all of the valid inequalities for TFLP of the form (7) will have $\alpha_{ij} \geq 0$ and $\beta_{ij} \geq 0$ for all $\{i,j\}$. In the following discussion, we describe three classes of valid inequalities that satisfy these conditions.

3.1 The cutset inequality

In our discussion of our solution appproach for the TFLP, we introduced the cutset inequality (5) through an example for situations with only HC type facilities. We now generalize this inequality for the TFLP, i.e., for situations with both LC and HC type facilities. The *cutset inequality* for the TFLP is described by

$$X_{S,T} + rY_{S,T} \geq r\lceil D_{S,T}/C \rceil \qquad \text{for all S, T; } S \subset N,\ T = N\backslash S. \tag{8}$$

In this expression, $r = D_{S,T} \bmod(C)$. By convention, we set $r = C$ if $D_{S,T}$ is an integer multiple of C.

This facet defining inequality has several noteworthy properties. First, although it applies to a formulation in the space of the $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{f}$ variables, the inequality does not contain the flow variables. Second, suppose $X_{S,T}$ were to be always equal to 0 (i.e., the underlying cost structure were of a pure staircase form), and the problem contained $(|N|-1)$ commodities, each with unit demand and a common origin (say node 1), and with destinations at nodes $(2, 3,..., |N|)$. Now, if $C \geq |N|-1$, then the cutset inequalities reduce to one set of constraints of the cutset formulation of the minimum spanning tree, i.e., $Y_{S,T} \geq 1$. Moreover, for this special case, the optimal solution to the loading problem is a minimum spanning tree. Third, when C = 1, the problem contains an optimal solution with $Y_{S,T} = 0$ (assuming, without loss of generality, that $b_{ij} \geq a_{ij}$ for all $\{i,j\} \in A$) so that we can remove these variables from the problem formulation). Inequality (8) then reduces to $X_{S,T} \geq D_{S,T}$, which can be derived by aggregating the flow conservation and capacity constraints across the cutset.

We use the Chvátal-Gomory procedure to derive the cutset inequality (8) and thus to establish its validity. This technique consists of repeatedly taking linear combinations of already known valid inequalities, and then using integrality arguments to round up (or down) coefficients.

Proposition 3. *The cutset inequality (8) is valid for the TFLP.*

Proof. For any feasible solution to the problem, the aggregate capacity across the cutset must be no less than the demand across the cutset. Thus, the "aggregate capacity demand inequality" is

$$X_{S,T} + CY_{S,T} \geq D_{S,T} = qC + r$$

for a suitable choice of the nonnegative integer q. If r = C, then the cutset inequality is valid because it is equivalent to the aggregate capacity demand inequality. We start with the aggregate capacity demand inequality and use induction to establish the validity of the cutset inequality for r < C. Consider the inequality

$$X_{S,T} + (C - v)\, Y_{S,T} \geq q(C - v) + r. \tag{9}$$

If v = 0, then inequality (9) is the aggregate capacity demand inequality. We will show that if inequality (9) is valid for v = u, for $0 \leq u \leq C-r-1$, then the inequality is also valid for v = u + 1.

Since the aggregate design variable $X_{S,T}$ is nonnegative, $\frac{1}{[C-(u+1)]} X_{S,T} \geq 0$, Adding this inequality to inequality (9) with v = u, we obtain

$$\frac{(C-u)}{[C-(u+1)]} X_{S,T} + (C-u)\, Y_{S,T} \geq q(C-u) + r.$$

Thus,

$$\begin{aligned} X_{S,T} + [C-(u+1)]\, Y_{S,T} &\geq q[C-(u+1)] + r\,\frac{[C-(u+1)]}{C-u} \\ &= q[C-(u+1)] + r - \left(\frac{r}{C-u}\right). \end{aligned}$$

Now, we can use integrality arguments to round up the righthand side to the nearest integer because the left hand side is necessarily integer. Since $r < C - u$, we obtain inequality (9) for $v = u + 1$ and so the proof is complete. ⊗

Figure 2 pictorially depicts this derivation in the aggregate space of $X_{S,T}$ and $Y_{S,T}$ variables when $D_{S,T} > C$. In this figure, the lightly shaded region denotes the convex hull of feasible solutions to TFLP. We start with the aggregate capacity demand inequality (defined by line *KL* in Figure 2). If $r < C$, we generate a new valid inequality (defined by the line *MN*). We repeat this process of tightening the inequality until we reach the cutset inequality (line *OP*). Notice that at each stage we rotate the "current inequality" about point $Z = (q, r)$ in the anti-clockwise direction; thus, we "cut off" a part of the feasible region (the triangle, *LNZ*, at the first stage) to the linear programming relaxation at that stage.

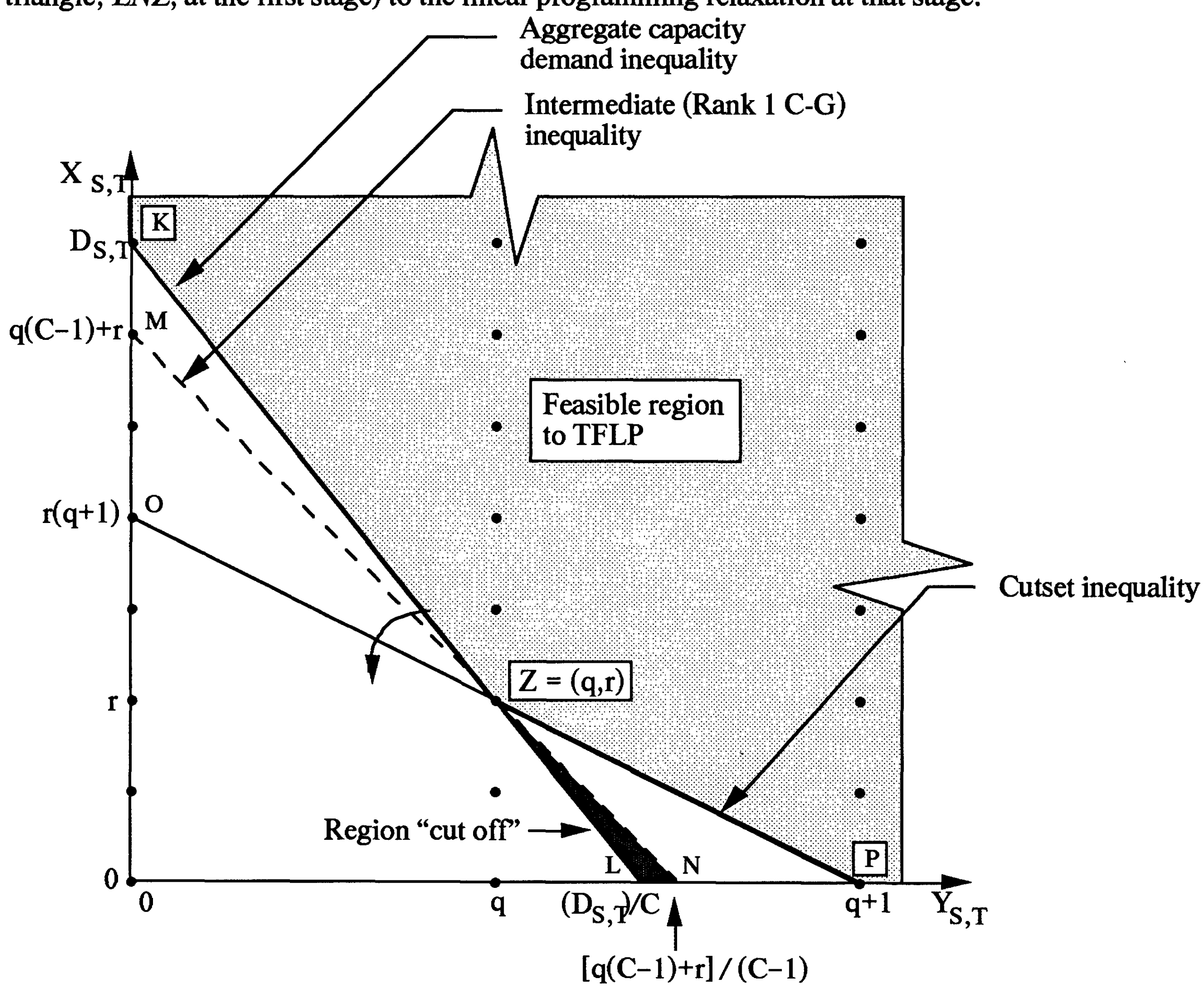


Figure 2. Pictorial interpretation of the Chvátal-Gomory procedure for the cutset inequalities.

Now, suppose S is composed of two "separated" components U and V satisfying the conditions $U \cup V = S$, $U \cap V = \phi$ and the condition that the arc set $\{U,V\} = \phi$. Then any demand from U to V must flow via T, crossing the cutset $\{S,T\}$ twice, and, therefore, the cutset inequality will not hold as an equality if $D_{U,V}$ is sufficiently large. Thus, we intuitively observe that if the subgraph induced by S or by T is not connected, then (8) could be a weak inequality. The next theorem shows that this condition on the connectivity of S and of T is necessary for the cutset inequality to be a facet.

Theorem 4. *The following conditions are necessary and sufficient for the cutset inequality (8) to be a facet of Conv(TFLP):*

1. *The subgraphs defined by S and by T are connected.*

2. $D_{S,T} > 0$.

Proof. See the Appendix.

Note that if r = C, then the cutset inequality is equivalent to the aggregate capacity demand inequality. Thus, although the cutset inequality still defines a facet under certain conditions, it does not add to the formulation of TFLP. Moreover, an immediate consequence of Theorem 4 applies to a network with existing capacities on some of the arcs. Corollary 5 shows how to modify the righthand side of inequality (8) so that we generate a facet for this situation as well.

Corollary 5. *Let $\{S,T\}$ be a partition of N and assume that the network has an existing capacity of $E_{S,T}$ installed between node sets S and T. If $D_{S,T} > E_{S,T}$, $r = (D_{S,T} - E_{S,T}) \bmod(C)$ and Condition 1 of Theorem 4 is valid, then $X_{S,T} + rY_{S,T} \geq r\lceil (D_{S,T} - E_{S,T})/C \rceil$ is a facet of Conv(TFLP).*

3.2 The arc residual capacity inequality

Magnanti, Mirchandani and Vachani (1990) have studied a core problem that arises when we use a Lagrangian approach for solving many capacitated network design models. This problem is essentially a multicommodity network design problem on a single arc with a single type of facility (HC). In their study of the convex hull of feasible solutions to this problem, they developed the arc residual capacity inequality. We show that a generalized version of this inequality defines a facet of Conv (TFLP). More importantly, if we add all the generalized arc residual capacity inequalities and the upper bound constraints (4) to formulation (TFLP), then the lower bound that we obtain from its linear programming relaxation is the same as the lower bound that we obtain if we use a Lagrangian approach to solve the TFLP.

Before introducing the generalized arc residual capacity inequality, we extend our model by adding the following logical inequality (4) to the original TFLP formulation, that is,

$$f_{ij}^k + f_{ji}^k \leq d_k \text{ for all } \{i,j\}, \text{ for all } k.$$

In addition, if $\{i,j\}$ is a bridge arc (i.e., an arc whose removal causes the network to separate into two disjoint components), we add the inequalities

$$f_{ij}^k = 0 \textit{ and } f_{ji}^k = 0 \textit{ for all } k \in K \setminus K(i,j). \tag{10}$$

In this expression, K(i,j) denotes the set of commodities whose origin and destination nodes lie on the "opposite" sides ("shores") of the arc $\{i,j\}$. We can always add the inequality (10) to the formulation for the same reason that we can add the upper bounding inequality (4) — that is, because we can assume that the solution for any commodity is cycle free.

The generalized arc residual capacity inequality, which we will henceforth refer to as simply the *arc residual capacity inequality*, is

$$\sum_{k \in L} \left(f_{ij}^k + f_{ji}^k\right) - x_{ij} - r_L y_{ij} \leq (\mu_L - 1)(C - r_L) \equiv D_L - \mu_L r_L. \tag{11}$$

In this expression, L is any subset of K, $D_L = \sum_{k \in L} d_k$, $\mu_L = \lceil D_L/C \rceil$ and $r_L = D_L \bmod(C)$.

Note that if $\sum_{k \in L} \left(f_{ij}^k + f_{ji}^k\right) = D_L$ for any subset L and $x_{ij} = 0$, then this inequality forces y_{ij} to be at least μ_L; as we have seen earlier, the linear programming relaxation without this constraint would permit the fractional solution $y_{ij} = D_L/C$. Note further that because of inequality (4), which applies to any problem with nonnegative flow costs, the arc residual capacity inequality (11) reduces to the cutset inequality (8) if $\{i,j\}$ is a bridge arc and L = K(i,j).

To verify the validity of the arc residual capacity inequality for the TFLP formulation, we rewrite the inequality as $\sum_{k \in L} \left(f_{ij}^k + f_{ji}^k\right) \le D_L - r_L(\mu_L - y_{ij}) + x_{ij}$. If $y_{ij} \ge \mu_L$, then the inequality is valid since $\sum_{k \in L} \left(f_{ij}^k + f_{ji}^k\right) \le D_L$. If $y_{ij} = \mu_L - s$ for some $s \ge 1$, then the arc residual capacity inequality reduces to $\sum_{k \in L} \left(f_{ij}^k + f_{ji}^k\right) \le D_L - r_L s + x_{ij}$ which is equivalent to or dominated by the capacity constraint $\sum_{k \in K} \left(f_{ij}^k + f_{ji}^k\right) \le C(\mu_L - s) + x_{ij}$.

The next three theorems show, in a theoretical sense, the effectiveness of the arc residual capacity inequality in tightening the linear programming relaxation of TFLP.

Theorem 6. *The arc residual capacity inequality (11) defines a facet of the extended TFLP model if and only if*

1. *If $r_L = C$, then $L = K$.*

2. *If $\{i,j\}$ is a bridge arc, then $L = K(i,j)$.*

Proof. See the Appendix.

Theorem 7. *The capacity inequality (2), the upper bound inequalities (4), the arc residual capacity inequalities (11), and the nonnegativity constraints describe the convex hull of the set of feasible solutions to P(LAG).*

Proof. Since the proof of this result is similar to the proof of a more special result given by Magnanti, Mirchandani and Vachani (1990), we do not provide the details.

Let P(LPR) denote the linear program obtained by appending all the upper bound constraints (4) and the arc residual capacity inequalities (11) to the linear programming relaxation of TFLP. Clearly, the optimal solution to P(LPR) provides a lower bound on the cost of the optimal solution to TFLP.

Theorem 8. *The lower bound provided by the optimal solution to P(LPR) is equal to the lower bound obtained from the Lagrangian relaxation approach for solving TFLP in which we relax constraint (1).*

Proof. Theorem 7 establishes that inequalities (2), (4) and (11) and the nonnegativity constraints describe the convex hull of the set of feasible solutions to P(LAG). Thus, we can replace constraints (3) of P(LAG) by (11) and the corresponding nonnegativity constraints and obtain an equivalent Lagrangian subproblem. This new (equivalent) subproblem satisfies the integrality property (Geoffrion, 1974) and, hence, it provides a Lagrangian lower bound equal to that obtained from solving P(LPR). ⊗

3.3 Three-partition inequalities

One way to view the cutset inequality is in terms of network aggregation: we aggregate the network into two "super nodes" S and T and write the inequality as a valid inequality for the resulting two node network. Building upon this idea, Magnanti, Mirchandani and Vachani (1990) have described an aggregate three-node (three-partition) inequality for the single facility case. This Chvátal-Gomory inequality, which we illustrated in Section 2, is useful for describing the convex hull of feasible solutions to the single-facility network loading problem. We describe two ways of generalizing this inequality for the two-facility case. The three-partition inequalities are motivated by the following consideration: suppose the formulation of the network loading problem consists of the flow conservation constraints, the capacity constraints, and the cutset inequalities. Then the linear programming relaxation of the loading problem on a three-node network can produce a "half-integral solution" in **y**. For example, if $C = 24$, $d_{12} = d_{13} = d_{23} = 12$, $a_{12} = a_{13} = a_{23} = b_{12} = b_{13} = b_{23}$, then $y_{12} = y_{13} = y_{23} = 1/2$ and $x_{12} = x_{13} = x_{23} = 0$ is a nonintegral optimal solution to the linear programming relaxation of the problem. (This problem is essentially our earlier example shown in Figure 1.) Notice that the arcs on which the solution **y** is half-integral form a cycle. This phenomenon occurs in larger networks for the same reason: a "half-cycle" satisfies the cutset constraints, but is cheaper than any other integral solution. The inequalities we present next are useful for cutting off such half-integral solutions. We will describe these inequalities for a network with three nodes; however, these results also apply to larger networks with three aggregate nodes.

Let 1, 2, and 3 be the three nodes of the network. Let d_{12}, d_{13} and d_{23} denote the demands between nodes 1 and 2, 1 and 3, and 2 and 3 respectively. Furthermore, if i, j and k are distinct elements of $\{1,2,3\}$, define $r_{ij} = d_{ij} \bmod (C)$ and $r_i = (d_{ij}+d_{ik}) \bmod(C)$.

Proposition 9. *Let $\underline{r} = \min(r_1, r_2, r_3)$. Then the following inequality is a valid inequality for the convex hull of feasible solutions to the two facility loading problem on a three-node, three-arc network:*

$$x_{12}+x_{13}+x_{23} + \underline{r}(y_{12}+y_{13}+y_{23}) \geq \left\lceil \frac{\underline{r}\left(\left\lceil \frac{d_{12}+d_{13}}{C} \right\rceil + \left\lceil \frac{d_{12}+d_{23}}{C} \right\rceil + \left\lceil \frac{d_{13}+d_{23}}{C} \right\rceil\right)}{2} \right\rceil. \tag{12}$$

Proof. Consider the cutset inequality with node 1 on one side of the partition and nodes 2 and 3 on the other side. This inequality is

$$x_{12}+x_{13} + r_1(y_{12}+y_{13}) \geq r_1 \left\lceil \frac{d_{12}+d_{13}}{C} \right\rceil$$

which implies

$$x_{12}+x_{13} + \underline{r}(y_{12}+y_{13}) \geq \underline{r} \left\lceil \frac{d_{12}+d_{13}}{C} \right\rceil.$$

We can similarly obtain the corresponding inequalities for nodes 2 and 3. Adding these three inequalities, dividing by 2 and using integrality arguments to round up the righthand side gives the desired result. ⊗

Proposition 10. *Consider the two facility loading problem on a three-node, three-arc network. Then all feasible solutions satisfy the inequality*

$$\begin{aligned} 2(x_{12}+x_{13}+x_{23}) + (r_{12}+r_{13}+r_{23})(y_{12}+y_{13}+y_{23}) \geq \\ (r_{12}+r_{13}+r_{23})\left(\left\lfloor \frac{d_{12}}{C} \right\rfloor + \left\lfloor \frac{d_{13}}{C} \right\rfloor + \left\lfloor \frac{d_{23}}{C} \right\rfloor + 2\right) \end{aligned} \tag{13}$$

if and only if

1. *None of the remainders r_{12}, r_{13} and r_{23} equal C.*

2. *The remainders satisfy the triangle inequality; that is,*

$$r_{12} + r_{13} \geq r_{23},\ r_{12} + r_{23} \geq r_{13} \text{ and } r_{13} + r_{23} \geq r_{12}.$$

3. *If $\max(d_{12}, d_{13}, d_{23}) > C$, then $r_{12} + r_{13} + r_{23} \leq 2C$.*

Proof.

Necessity.

1. Suppose $r_{12} = C$, $r_{13} < C$ and $r_{23} < C$. Then the feasible solution

$$y_{12} = \frac{d_{12}}{C},\ y_{13} = \left\lfloor \frac{d_{13}}{C} \right\rfloor,\ y_{23} = \left\lfloor \frac{d_{23}}{C} \right\rfloor,\ x_{12} = 0,\ x_{13} = r_{13}, \text{ and } x_{23} = r_{23}$$

violates inequality (13).

2. Suppose $r_{12} + r_{13} < r_{23} < C$. Then the feasible solution

$$y_{12} = \left\lfloor \frac{d_{12}}{C} \right\rfloor,\ y_{13} = \left\lfloor \frac{d_{13}}{C} \right\rfloor,\ y_{23} = \left\lceil \frac{d_{23}}{C} \right\rceil,\ x_{12} = r_{12},\ x_{13} = r_{13}, \text{ and } x_{23} = 0$$

violates inequality (13).

3. Suppose $d_{12} > C$ and $r_{12} + r_{13} + r_{23} > 2C$. Then the feasible solution

$$y_{12} = \left\lfloor \frac{d_{12}}{C} \right\rfloor - 1,\ y_{13} = \left\lfloor \frac{d_{13}}{C} \right\rfloor,\ y_{23} = \left\lfloor \frac{d_{23}}{C} \right\rfloor,\ x_{12} = C + r_{12},\ x_{13} = r_{13} \text{ and } x_{23} = r_{23}$$

violates inequality (13).

Sufficiency.

Suppose $y_{12}+y_{13}+y_{23} \geq \left\lfloor \frac{d_{12}}{C} \right\rfloor + \left\lfloor \frac{d_{13}}{C} \right\rfloor + \left\lfloor \frac{d_{23}}{C} \right\rfloor + 2$. Then inequality (13) is clearly satisfied. So assume that $y_{12}+y_{13}+y_{23} = \left\lfloor \frac{d_{12}}{C} \right\rfloor + \left\lfloor \frac{d_{13}}{C} \right\rfloor + \left\lfloor \frac{d_{23}}{C} \right\rfloor + 2 - s$ for some integer s, $1 \leq s \leq \left\lfloor \frac{d_{12}}{C} \right\rfloor + \left\lfloor \frac{d_{13}}{C} \right\rfloor + \left\lfloor \frac{d_{23}}{C} \right\rfloor + 2$. If $s = 1$, we can assume (by symmetry) that $y_{12} \geq \left\lceil \frac{d_{12}}{C} \right\rceil$. Since $y_{13}+y_{23} \leq \left\lfloor \frac{d_{13}}{C} \right\rfloor + \left\lfloor \frac{d_{23}}{C} \right\rfloor$, a cutset argument implies that $x_{13}+x_{23} \geq r_{13}+r_{23}$. Substituting this inequality in inequality (13) and using Condition 2 proves the validity of inequality (13).

Next assume that $s \geq 2$. Then the aggregate capacity demand inequality implies that $x_{12}+x_{13}+x_{23} \geq C\,(s-2) + r_{12}+r_{13}+r_{23}$. Substituting for the lefthand side of inequality (13) and using Condition 3 proves the result. (If $\max(d_{12}, d_{13}, d_{23}) < C$, then $s = 2$, $y_{12} = y_{13} = y_{23} = 0$ and $2(x_{12}+x_{13}+x_{23})$ is at least as large as the righthand side of $2(r_{12}+r_{13}+r_{23})$; otherwise we use Condition 3.) ⊗

When implemented in our computational study along with the cutset inequality, but without the arc residual capacity inequality, these valid inequalities were modestly effective in reducing the integrality gap. With both the cutset and arc residual capacity inequalities included, the effect of adding the three-partition inequalities on the integrality gap was less pronounced.

4.0 Computational study

This section describes the results of a computational study designed to test the effectiveness of the inequalities described in Section 3. As we have shown, under suitable conditions the cutset and the arc residual capacity inequalities induce facets of the underlying polyhedron, so we know that they tighten the formulation of TFLP. Moreover, because the conditions for these inequalities to be facets are quite mild, we might be led to believe that they would be effective algorithmically in a cutting plane approach.

We have used these inequalities in an algorithmic procedure with two main phases. (For a discussion of this general approach, see Hoffman and Padberg, 1985, and Van Roy and Wolsey, 1984.) In the first phase, the algorithm uses a cutting plane approach to tighten the formulation and generate a good lower bound. If this phase terminates with a nonintegral solution, then the approach resorts to the second phase; this phase finds a good — optimal for problems up to 10 nodes — solution using branch-and-bound.

We have tested the cutting plane algorithm on a total of 126 test problems on networks with 6, 10 and 15 nodes and a variety of demand patterns (see Section 4.3). These problem sizes might appear to be small; however, the 10 and 15 node problems have approximately 45 and 65 *general* integer variables. Moreover, we have attempted to generate these problems in a way that reflects the demand and cost structures occurring in practice (that is, they are derived from real data).

4.1 Phase I

The inequalities that we developed in Section 3 could be used conceptually in two ways. For example, we could add, a priori, the cutset inequalities corresponding to all nontrivial partitions of N. However, this option would add an exponential number of constraints to the formulation. Moreover, most of these inequalities would be inactive at the optimal solution of any particular instance of the problem and are, therefore, not necessary for the solution of this problem instance.

The other option dynamically adds the inequalities in a cutting plane based algorithm. Thus, given a fractional solution for the current formulation, we identify a valid inequality that this solution violates. We adopted the second option to augment the problem formulation and used the USER subroutine of LINDO, on the VAX 6640 and 8820 computers, to automate the generation and addition of the facet inequalities.

The separation problem of the cutset inequalities for the single commodity case can be solved as a max flow problem, or more generally, as a linear program (see Mirchandani, 1989). Solving the separation problem in the multicommodity case is difficult because of the structure of the cutset inequalities: each cutset can generate a different value of the remainder, *r*. A polynomially bounded algorithm does not seem evident. (The separation problem might well be NP-hard.) In our computational study, for the size of some problems that are currently of interest to practitioners (10 to 15 nodes), we found that an exhaustive search for generating violated cuts does not consume excessive computational time (as compared with the time for re-optimizing the resulting linear program) and does reduce the integrality gap. We, therefore, adopted the following enumeration heuristic for solving the separation program associated with the cutset inequalities. This heuristic first carries out an exhaustive search of cutsets defined by sets S with small cardinality. It then uses a "growth" strategy, starting from a single node as S and sequentially building S, to identify violated inequalities.

Heuristic for identifying violated cutset inequalities

Step1: Check for violated inequalities with $|S| = 1$. Among all such violated inequalities, select the one with minimum value for $X_{S,T} + rY_{S,T} - r\left\lceil \frac{D_{S,T}}{C} \right\rceil$. If this enumeration identifies a violated inequality, return.

Step 2: Repeat Step 1, but with $|S| = 2$.

Step 3: Search sequentially through all partitions with $|S| = 3$, 4 or 5. Add the first violated inequality found. If this search does not identify a violated inequality, proceed to Step 4a; otherwise, return.

Step 4a: Initialize:

D_i := total demand originating or terminating at node i,
Z_i := total current capacity incident to node i (i.e., $\sum_{j \in N} (x_{ij} + Cy_{ij})$), and
$S := \{i^*: i^* = \underset{i \in N}{\operatorname{argmax}}\, D_i / Z_i\}$.

Step 4b: If $|S| \leq 5$, go to Step 4c. Otherwise, check if the current fractional solution violates the cutset inequality defined by S. If yes, add this inequality. Return.

Step 4c: If $|S| = |N|-6$, (print "violated inequality cannot be identified") stop. Otherwise, add node $j^* := \underset{j \in N}{\operatorname{argmin}}\, d_{i*j} \,/\, (x_{i*j} + Cy_{i*j})$ to S. Go to Step 4b.

This heuristic adds one violated cutset inequality per iteration in increasing order of $|S|$. If the heuristic cannot identify a violated cutset inequality, we first search for violations of the arc residual capacity and then violations of the three-partition inequalities. For the arc residual capacity inequality, we check for all violated inequalities with the cardinality of the commodity set (L in expression (11)) equal to 1 or 2. Since we found that the linear program solutions to large problems violate many of these inequalities, we added five such violated inequalities per iteration.

4.2 Phase II

In Phase II, we used branch-and-bound starting with the fractional solution generated by Phase I. Because the version of LINDO that we were using was not capable of solving general integer programs, we implemented this phase of the algorithm on an IBM 4381 computer using MPSX/370 version 2.0.

Prior experience has established the importance of using a good upper bound in the branch-and-bound procedure. We used the upper bound generated by a Lagrangian approach for solving the problem (see Vachani, 1988). This approach dualizes the flow conservation constraints. The relaxed problem then decomposes by arc for given values of the Lagrange multipliers; furthermore, the subproblem for each arc is a knapsack type problem that can be solved efficiently. The procedure uses subgradient optimization to tighten the lower bound. At each iteration, the method also constructs a feasible integer solution utilizing the Lagrangian solution and improves this solution heuristically.

We also used a *bootstrapping* approach for the more difficult demand topologies (see Section 4.4.1). For these problems, we obtained an upper bound after fixing some integer variables and then carrying out branch-and-bound on the remaining set of variables. (The reduced number of fractional variables accelerated the branch-and-bound phase.) Phase II of the cutting plane procedure subsequently used the better of this heuristic solution and the Lagrangian heuristic solution as an upper bound for finding the optimal solution.

4.3 Computational study design

Our test problems, although randomly generated, were based upon information provided by GTE Laboratories and are representative of cost and demand structures arising in practice.

Specifically, we tested the algorithm on 3 different network sizes: 6, 10, and 15 nodes. We generated the ordinates and abscissae of these nodes — uniformly distributed on a unit square — using a random number generator. (This random number generator satisfies Knuth's (1981) spectral test for dimensions 3, 4, 5 and 6; for all practical purposes, it has an infinite period (Press et al., 1989).) Given these points, we constructed the underlying backbone network. Recall that we are allowed to lease LC and HC facilities only on the arcs of this backbone network. For the 6 node problems, we assumed a fully connected topology. To avoid an explosive growth in the number of variables, we assumed that the 10 and 15 node networks were sparse. (The number of variables in a complete network with a commodity demand between every pair of n nodes equals $(n*(n-1) + 0.5*(n*(n-1))^2$, which for $n = 6$, 10 and 15 equals 480, 4140 and 22,260 respectively.) For sparse networks, we chose a targeted nodal degree for each node to be equal to 3 or 5 with a probability of 0.3 and 4 with a probability of 0.4. Starting from node 1, we sequentially cycled through to node $|N|$: at stage i, we determined node i's closest neighbor (in terms of Euclidean distance) with unsatisfied degree requirements. We added an arc between this pair of nodes with a probability of 0.80 and repeated the process until either (i) the topology satisfied node i's degree requirements, or (ii) we had considered all the nodes with unsatisfied degree requirements once. In case (ii), we identified node i's closest neighbor, say node j, satisfying the property that the current topology did not include arc $\{i,j\}$, and we added this arc. Consequently, this step would cause us to exceed node j's degree requirement if it had already been satisfied.

Next, we determined the LC and HC costs. Both these costs have two components: (i) a fixed cost component, and (ii) a variable cost component which is a linear function of the arc length. We determined the fixed and variable cost parameters to ensure that the generated costs are consistent with the range of tariffs offered by the long distance telephone companies at the time of this study.

We generated three different kinds of demand topologies as follows. We assumed that the probability of nonzero demand between any pair of nodes is 0.5 for the 10 node networks and 0.2 for the 15 node networks. For those pairs of nodes with nonzero demand, we chose the value of the demand in one of three different ways: (i) uniformly distributed for all pairs of nodes; (ii) uniformly distributed with a higher mean between *a* central node and all the other nodes as compared to the demand between pairs of nodes from the remaining set; and (iii) uniformly distributed with a higher mean between *two* central nodes and all the other nodes as compared to the demand between pairs of nodes from the remaining set. For each case, we generated two different levels of average demand: low and high (the variable component for calculating the high level of demand is twice the variable component for low level of demand).

Further details of the exact expressions used in the calculations of the demand and cost data and the test problems are available from the authors.

4.4 Computational results

In this subsection, we report our computational results on 126 test problems. These problems are distributed over 15 problem categories; we tested 6 to 10 problems in each problem category so that we might determine how the methodology works "on the average." Our results

- show that our methodology reduces the integrality gap from the one provided by the original linear programming formulation by 65% to 80% for the 6 and 10 node problems and approximately 55% for the 15 node problems,
- show that the average integrality gap after the completion of our cutting plane procedure is 8.13%,
- show that the approach can solve problems with up to 10 nodes (with up to 45 general integer variables) to optimality in a reasonable amount of time,
- show how to strengthen the linear programming formulation, a priori,
- identify network topologies for which the TFLP is more difficult to solve, and
- compare computationally the Lagrangian and the cutting plane based approaches.

4.4.1 Aggregate results

For the remaining part of this section, we adopt the following convention for denoting problem instances:

(i) the letter C in the first field denotes completely uniform demand between all pairs of nodes, the letter O denotes one central node, and the letter T denotes two central nodes,

(ii) the second field denotes whether the magnitude of the demand is high (H) or low (L),

(iii) the next field denotes the number of nodes in the network,

(iv) the last field contains the problem number.

We will use the following acronyms to denote the respective solutions in this discussion:

LP: Linear programming model (with the flow conservation and capacity constraints).

LPC: Cutset inequalities + three-partition inequalities + linear programming model.

LPR: Arc residual capacity inequalities + linear programming model.

LPA: All inequalities of Section 3 + linear programming model.

LLB: Lagrangian lower bound.

BES: The best integer solution obtained.

We used three performance measures for our analysis:

(i) Percentage gap := $\frac{\text{BES}-\text{LPA}}{\text{BES}}$,

(ii) Percentage improvement := $\frac{\text{LPA}-\text{LP}}{\text{LP}}$, and

(iii) Percentage gap reduction := $\frac{\text{LPA}-\text{LP}}{\text{BES}-\text{LP}}$.

The first criterion measures the final integrality gap and can be used as a performance guarantee of the heuristic used to determine BES; it is also one (rough) indicator of the time required for branch-and-bound (typically, the larger this measure, the larger the time branch-and-bound will take to solve the problem). The second criterion measures the improvement from the linear programming solution, while the third criterion combines the other two measures: it indicates the "effectiveness" of our methodology in reducing the integrality gap. (Note that for comparing the effectiveness of different approaches, we could define these performance measures using the LPC, LPR or LLB values instead of the LPA value.)

Figure 3 presents these performance measures for the problem categories that we tested.

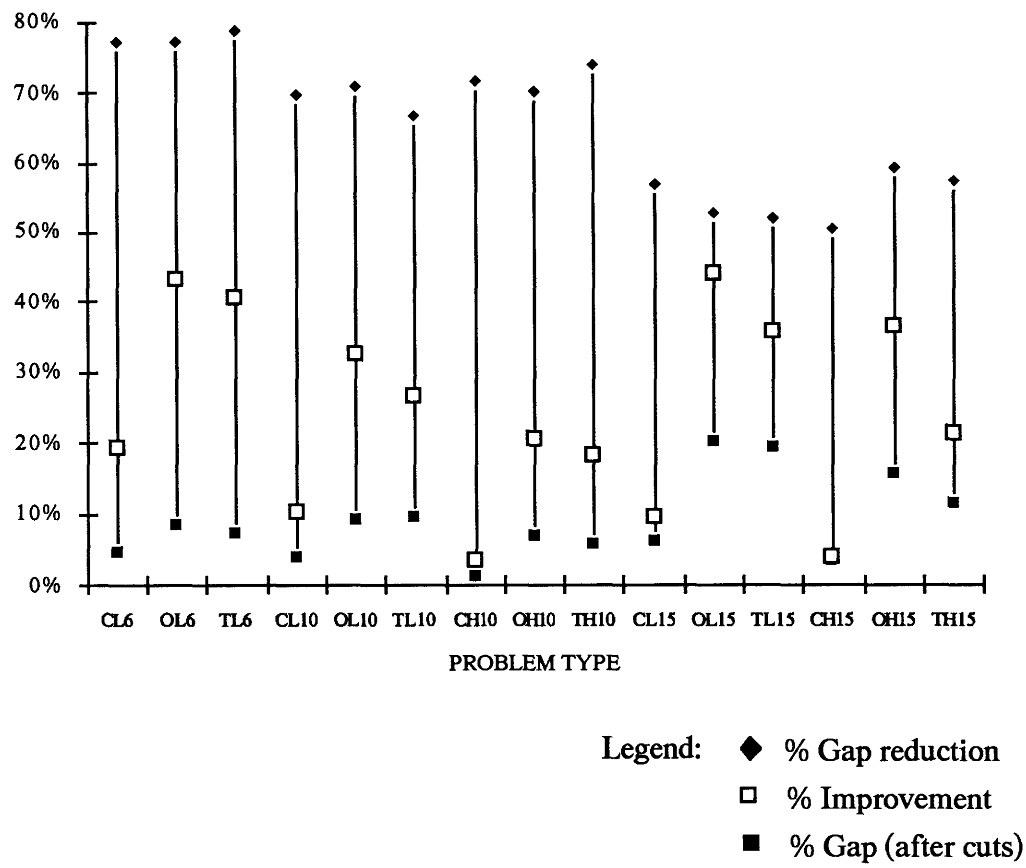


Figure 3. Average performance measures

These results indicate that the inequalities under investigation are effective in reducing the integrality gap, especially for 6 and 10 node problems. The average percentage gap is high for the O 15 and T 15 problem categories. We suspect that these larger gaps are attributable to a bad upper bound. However, because we did not run the branch-and-bound algorithm on these problem categories, we cannot substantiate this statement. Furthermore, we observe that the completely uniform demand (C) topologies have the smallest percentage gaps (see Figure 3) and seem to be the easiest to solve, and that the one and two central node (O and T) topologies are more difficult. Moreover, as the demand level increases, the percentage gap becomes smaller. Thus, for example, the percentage gap for OH 10 problems is smaller than the corresponding gap for OL 10 problems.

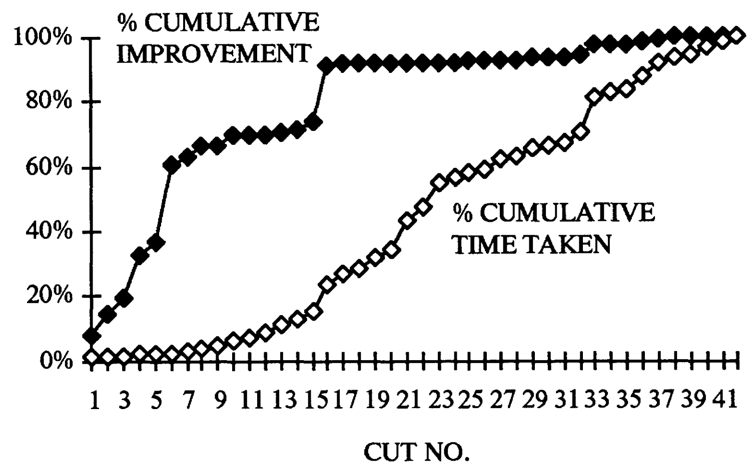
The percentage improvement, on the other hand, is the lowest for the complete demand topologies (see Figure 3 again). We can explain this apparent anomaly by observing that the linear programming relaxation of the original formulation generates a smaller integrality gap for these problem categories. We also observe that the average reduction in the integrality gap does not seem to depend on the demand pattern; this figure is between 65% and 80% for the nine problem categories with up to 10 nodes (problems for which we found the optimal solution) and approximately 55% for the 15 node problems. (Notice that although we haven't reported this data directly, our results show that the gap between the optimal objective values of the linear programming relaxation and the integer programming version of the original problem formulation TFLP is as high as 43%.)

In addition to testing the effectiveness of the polyhedral approach for solving the TFLP, this computational study was designed to identify possible ways to improve the formulation of the TFLP a priori. We collected information on the reduction of the gap after the addition of each cut. Figure 4 presents the cumulative improvement after the addition of each cut and the cumulative time taken up to that stage for two typical problems. From these figures, we observe that the "cumulative percentage improvement" for the cuts exhibit a "tailing effect." However, the improvements do jump on occasion: often when the method identifies a new class of inequality. (Recall that our separation problem heuristic searches for a violated inequality in a pre-specified order of the class of inequality.) Notice further that the cumulative time grows slowly in the beginning stages of the algorithm, but the slope tends to increase as the algorithm proceeds and the linear programs become larger. Observe that we achieve about 90% of the improvement in the integrality gap in about 50% of the total solution time.

These observations lead us to conclude that (i) we might try the effect of randomizing the order in which we select the class of inequality to be considered, and (ii) we might terminate the cutting plane procedure after we have added a predetermined number of violated inequalities.

We observed from the timing information that the method spends most of its time solving the linear program: thus, the time for facet-based optimization could be reduced by adding, say, 3 to 5 cutset inequalities simultaneously. This implementation would, however, defeat our objective of checking the progress of the algorithm at each step. This study's developmental nature prompted us to focus less attention on the algorithm's timing and to concentrate more on testing the method's effectiveness in reducing the integrality gap.

PROGRESS OF CUTTING PLANE PROCEDURE
PROBLEM CL61



PROGRESS OF CUTTING PLANE METHOD
PROBLEM CL101

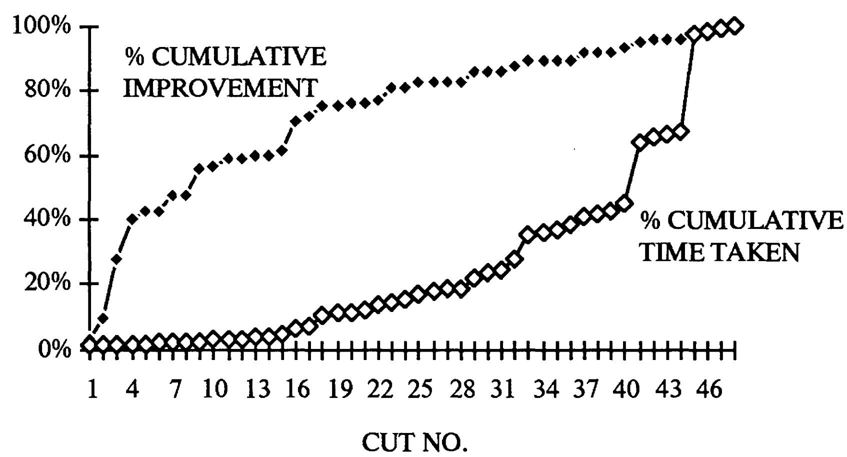


Figure 4. Progress of cutting plane procedure

Ideally, we would like to analyze the cumulative improvement and the cumulative timing information aggregated by problem category as well. We could aggregate the percentage improvement by cut number; however, this approach has two disadvantages. First, because the jumps do not occur at the same cut number across different problems, we would lose detail in the aggregation and the jumps seen in these graphs would become "averaged out." Second, the total number of cuts added varies across problems in the same category, and, therefore, aggregating might be misleading. Instead, we aggregated the percentage improvement by cutset inequality class when |S| = 1 and |S| = 2 for the three network sizes. Table I presents this information.

We observe that adding, a priori, all cutset inequalities (a polynomial number) for |S| equal to 1 or 2 can be quite effective in strengthening the formulation. As the size of the problem grows, the impact of adding only these inequalities, though still considerable, seems to be less pronounced.

Table I. Average gap reduction by inequality type

Problem category (no. of nodes)	Cutset inequalities %		
	|S|=1	|S|=2	|S|=1 & |S|=2
6	51.9%	21.4%	73.3%
10	26.8%	20.0%	46.8%
15	30.9%	15.6%	46.5%

We might note that the *total* decrease in the integrality gap due to the addition of these inequalities occurs in two stages: before and after the addition of the arc residual capacity and the three-partition inequalities. Therefore, the actual improvement in the integrality gap, if we were to include the cutset inequalities of cardinality 1 and 2 in advance, would be slightly lower than that suggested by the last column.

4.4.2 Computational Comparison of the Polyhedral and Lagrangian Methods

This section compares polyhedral methodology with the Lagrangian based approach. Admittedly, such a comparison would depend on the problem class that we are investigating, the inequalities identified and implemented for the polyhedral approach, and the implementation of the Lagrangian approach. Nonetheless, this comparison could be useful in an algorithm development process; we believe that this is the first study of its kind in this respect.

In Theorem 8, we proved that the duality gaps for the following two problems are equal: (i) the Lagrangian problem that dualizes the flow conservation constraints, and (ii) problem TFLP extended by adding all the arc residual capacity inequalities. However, in practice it is difficult to obtain the optimal solutions to both these problems. First, although the literature suggests a number of strategies for implementing the Lagrangian approach, the most commonly used strategy, subgradient optimization, does not guarantee theoretical convergence under practically feasible conditions. Our implementation uses subgradient optimization for updating the Lagrange multipliers; for this method, we could use any of a number of empirically tested alternatives for adjusting the step size from one iteration to the next. We tested several of these possibilities and in our computational evaluation we have used the best solution value found.

On the other hand, adding the arc residual capacity inequalities, a priori, would increase the size of the linear program substantially by approximately 2^{20} constraints for the 10 node problems. (For the demand patterns we considered, the 10 node problems contained approximately 20 commodities.) Instead, we added only a small subset of these possible inequalities: all those violated inequalities with the cardinality of the commodity set equal to 1 or 2. Thus, for both the Lagrangian approach and the polyhedral approach (with only the arc residual capacity inequalities), we obtained lower bounds to the actual solution values.

Figure 5 compares the average integrality gaps that we obtained using these two approaches. In this figure, LAG refers to the integrality gap that we obtained using the Lagrangian approach. ARC, CUT and ALL refer to the integrality gaps obtained using only the (1 and 2 commodity set) arc residual capacity inequalities, the cutset and the three-partition inequalities, and all the inequalities of Section 3 in the cutting plane procedure. The LAG and the ARC gaps are fairly close to each other (although the LAG gaps are slightly higher for the 15 node problems), suggesting that as the underlying network becomes larger, the polyhedral approach seems to provide better lower bounds.

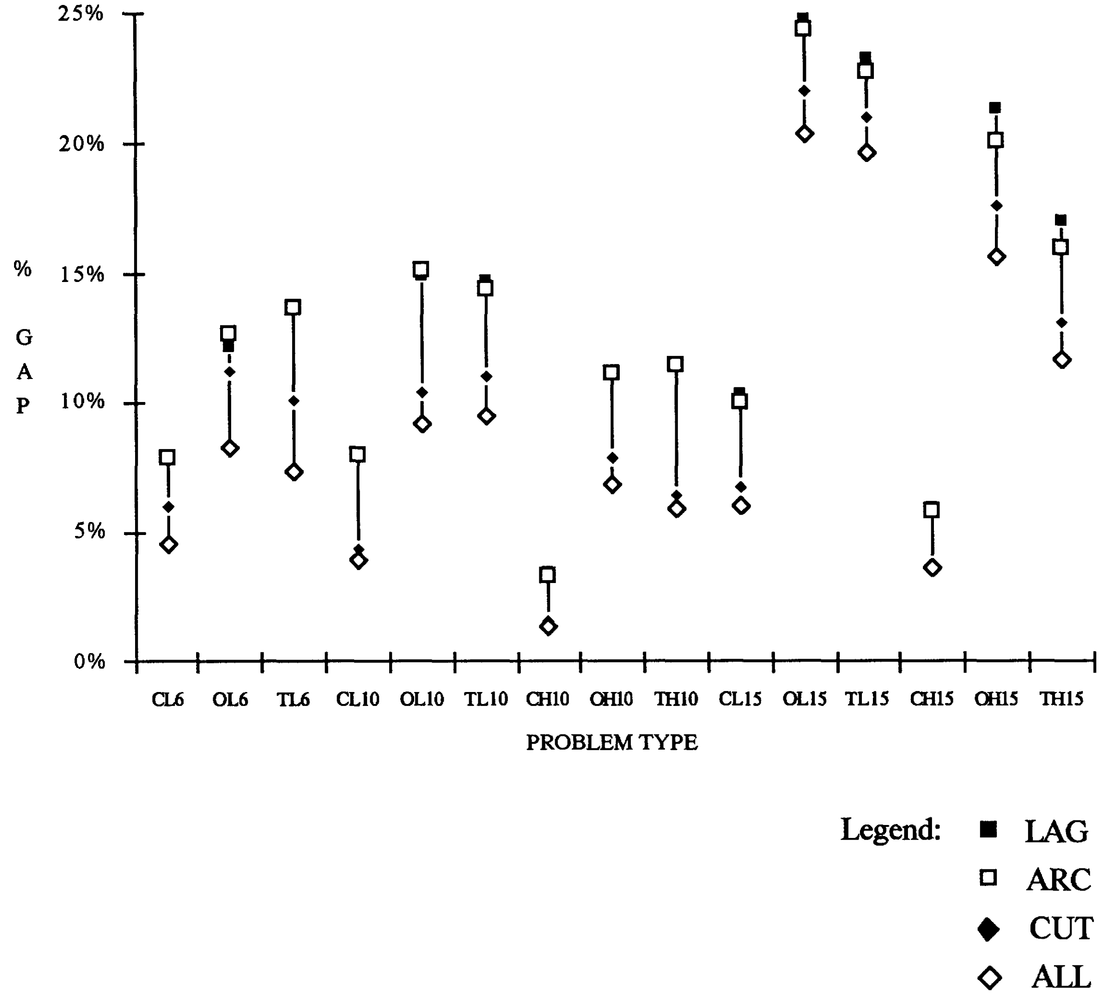


Figure 5. Percentage gap comparison for different approaches. If the symbols corresponding to LAG and ARC, or, CUT and ALL values overlap, we show only one symbol in this figure.

On the VAX 8820, the Lagrangian approach required approximately 20 to 40 seconds to solve 6 node problems, 2 to 4 minutes to solve 10 node problems, and 3 to 6 minutes to solve 15 node problems; this time also includes the time for determining the heuristic feasible solution. We implemented the polyhedral methodology on the VAX 6440 and the VAX 8820 machines. On these machines, Phase I of the procedure required 2 to 4 seconds to solve 6 node problems, 4 to 50 seconds to solve 10 node problems, and 14 to 350 seconds to solve 15 node problems when only arc residual capacity inequalities were used. When we included all the inequalities in the cutting plane procedure, the polyhedral procedure required significantly more time.

The Lagrangian and polyhedral approaches differ in three other respects:

(1) In order to obtain better lower bounds using the Lagrangian relaxation approach, we might have to add new constraints to the original problem formulation. However, doing so can make the relaxed problem much more difficult (and "inefficient") to solve. Therefore, reducing the integrality gap becomes increasingly more difficult using the Lagrangian approach. On the other hand, the polyhedral approach offers an opportunity for continuous improvement through the identification and implementation of new facets and valid inequalities.

(2) Unlike the Lagrangian approach, the polyhedral approach generates monotonically increasing lower bounds at every iteration.

(3) As the problem size becomes larger, the size of the linear program to be solved for the polyhedral approach increases rapidly (especially for fully-connected networks) and this approach might become difficult to use in practice. On the other hand, the computational burden of the Lagrangian approach does not increase as rapidly with problem size.

Figure 5 also shows that the cutset inequalities are more effective in reducing the integrality gap than are the arc residual capacity inequality across all problem categories. When both these inequalities are used together, the arc residual capacity inequalities seem to be more useful for the more difficult (i.e., the O and the T problem) problem categories.

To conclude this section, we note that the percentage gaps are still high for some problem categories, perhaps because the upper bounds are loose. Nevertheless, a further study of these network topologies might permit us to identify new valid inequalities and to improve the performance of cutting plane methods for these problems.

5.0 Conclusions

In this paper, we have modeled and developed solution approaches for a capacitated network design problem that arises in the telecommunications industry. Our model assumes that we can install a combination of two types of facilities to satisfy given point-to-point demand between various pairs of nodes of the network. We study two solution approaches to the problem: (i) a Lagrangian approach, and (ii) a cutting plane approach.

One of the objectives of this research is to compare the two approaches theoretically and computationally.

We have identified a set of arc residual capacity inequalities that when appended to the original linear programming formulation guarantee a lower bound equal to the Lagrangian lower bound. However, generating these bounds is difficult in practice because (i) the Lagrangian lower bound is difficult to achieve under practically feasible conditions, and (ii) the number of arc residual capacity inequalities grows exponentially in the number of commodities in the network. In our computational study, we have used only a polynomial subset of the arc residual inequalities and obtained a bound close to (and, in most cases, higher than) the Lagrangian lower bound.

In addition to the arc residual inequalities, we also identified two other classes of valid inequalities (the cutset and the three-partition inequalities) for the underlying polyhedron. Adding these inequalities ensures that we obtain a lower bound using the cutting plane approach that is at least as strong as the Lagrangian lower bound. Indeed, our computational results have shown that these inequalities are quite effective in reducing the integrality gap. Using the results of the computational study, we have also identified inequalities that might be added to the formulation, a priori, to reduce the integrality gap significantly without an enormous increase in the size of the linear program.

As we noted in Section 1, for telecommunications applications, subscribers might have a choice of a third facility, DS3, with capacity equal to 28 DS1 facilities. In general, consider m facilities denoted by HC(1), HC(2), ..., HC(m). Let the capacities of these facilities be λ^1C, λ^2C, λ^3C,..., λ^mC for some set of multipliers $\lambda^i \in Z^1_+$ and $\lambda^1 = 1$; the facilities are indexed so that $\lambda^j > \lambda^i$ if $j > i$. Let y^p_{ij} denote the number of facilities of type p installed on arc {i,j}. If x_{ij} denotes the number of LC facilities (with capacity 1) installed on arc {i,j} and we define aggregate variables across an {S,T} cutset as before, then it is possible to show that

$$X_{S,T} + r\sum_{p=1}^{m} \lambda^p Y^p_{S,T} \geq r\left\lceil \frac{D_{S,T}}{C} \right\rceil$$

(where $r = D_{S,T} \bmod (C)$ as earlier) is a valid inequality for the underlying multiple facility polyhedron. In fact, this inequality is facet defining under conditions similar to the conditions of Theorem 4. Thus, while we have discussed our results for the two facility loading problem, they are applicable in more general settings.

In conclusion, we would like to pose some research questions related to this research. First, under what conditions would the proposed inequalities describe the convex hull of the feasible solutions to the capacitated network loading problem? Second, can we identify additional classes of facet inequalities for the problem that might help us in reducing the integrality gap further? Finally, can we extend the formulation for other problem classes so that we obtain a bound that theoretically competes with the bound obtained using Lagrangian relaxation approaches? The answers to these questions might help us in further understanding the polyhedral structure of the capacitated network design model and other integer programming problems.

Appendix

Proposition 1. *TFLP is strongly NP-hard.*

Proof.

The three partition problem can be stated as follows:

Given 3n + 1 integers α_1, α_2, ..., α_{3n} and L satisfying $\sum_i \alpha_i = nL$, does there exist a partition of α_1, α_2, ..., α_{3n}, consisting of n sets S_1, S_2, ..., S_n, each of cardinality 3, that satisfies the property that $\Sigma\ \{\alpha_j : j \in S_i\} = L$ for all $1 \leq i \leq n$?

To transform the three partition problem into the TFLP, define a fully connected network with 3n+1 nodes. Call one of these nodes (say node 0) the central node and let $d_{0i} = \alpha_i + M$ for i=1, 2,..., 3n and $d_{ij} = 0$ otherwise, with M chosen to be a sufficiently large constant. Further, assume that the cost of installing a LC or HC facility between the central node and any of the other nodes is 1, and the cost of installing either facility between any other pair of nodes is ε (ε is strictly greater than 0 and sufficiently small). Let the capacity of a HC facility be L + 3M.

Note the following properties of any optimal solution to this TFLP.

(i) We can assume an optimal design does not use any LC facilities, since we can increase our capacity on any arc by installing a HC facility instead of a LC facility without increasing the cost.

(ii) Any feasible design must place at least n HC facilities on arcs adjacent to node 0. This result is true because the total demand is nL + 3nM and the capacity of each HC facility is L + 3M.

(iii) The cost of an optimal solution must be at least $n + 2n\varepsilon$, and any solution with this cost places 2n HC facilities on arcs {i,j} with $i \neq 0$, $j \neq 0$ and n HC facilities on arcs incident to node 0. Moreover, this design does not place multiple HC facilities on any arc {0,i} for $1 \leq i \leq n$. To establish this fact, we argue as follows. Since costs are positive, the optimal design must be a tree and contain 3n arcs. By property (ii), if the cost is $n + 2n\varepsilon$, then exactly n of these arcs are adjacent to node 0. If an optimal design places 2 or more HC facilities on any arc $\{0,j\}$, then more than 2n arcs $\{i,j\}$ with $i \neq 0$, $j \neq 0$ must contain a HC facility and thus the total cost exceeds $n + 2n\varepsilon$.

(iv) In any optimal solution with cost $n + 2n\varepsilon$, a node will act as a transshipment node for at most 2 other nodes. This result is a consequence of properties (ii) and (iii) and the fact that M is large.

We claim that we would have a Yes instance of 3PP if and only if the optimal solution to the TFLP has cost $n + 2n\varepsilon$. One direction of this claim is easy to prove. For, if we have a Yes instance to the 3PP, and the partitions are given by $S_i=\{i, 2i, 3i\}$ for $1 \le i \le n$, say, then a Yes instance of the TFLP can be obtained by installing HC facilities on arcs $\{0,i\}$, $\{i,2i\}$, and $\{i,3i\}$ for $1 \le i \le n$.

Now, assume that we have a solution to the TFLP with cost $n + 2n\varepsilon$. Then we have used exactly n HC facilities (each with capacity L + 3M) between the central node and other nodes, and exactly 2n HC facilities on the other arcs. Thus, the design is satisfying demand for n nodes directly, and the demand for the balance 2n nodes through some transshipment node. But then properties (ii) - (iv) imply that we have a 3 partition of nodes and, since the total demand is nL + 3nM, the total demand for each partition is exactly L + 3M. Consequently, we have a Yes instance to the 3PP. ⊗

Theorem 4. *The following conditions are necessary and sufficient for the cutset inequality (8) to be a facet of Conv (TFLP).*

1. *The subgraphs defined by S and by T are connected.*
2. $D_{S,T} > 0$.

Proof.

Necessity.

1. Assume that S is not connected, and let U and V be two "separated" components as defined in the discussion preceding the statement of the theorem. Let $r_U \equiv r_{U,\{V\cup T\}}$, $r_V \equiv r_{V,\{U\cup T\}}$, and $r_S \equiv r \equiv r_{\{U\cup V\},T}$ to simplify the notation in the following proof. Define single indexed aggregate design and demand variables similarly. Note that the definition of separated components implies that $X_S = X_U + X_V$ and $Y_S = Y_U + Y_V$. Note that we can

assume that $r < C$. Moreover, if $D_U = 0$ or $D_V = 0$, then we can tighten inequality (8); indeed, if $D_U = 0$ then the inequality corresponding to the cutset $\{V, U \cup T\}$ is tighter than the inequality corresponding to $\{S,T\}$. So, we assume that $D_U > 0$ and $D_V > 0$.

Now, either $r = r_U + r_V$ or $r = (r_U + r_V) \bmod (C)$. If $r = r_U + r_V$, then the Chvátal-Gomory procedure for deriving inequality (8) shows that the inequalities

$$X_U + rY_U \geq r\left\lfloor \frac{D_U}{C} \right\rfloor + r_U \text{ and } X_V + rY_V \geq r\left\lfloor \frac{D_V}{C} \right\rfloor + r_V$$

are valid. Adding these inequalities, we obtain inequality (8); thus, inequality (8) cannot be a facet.

On the other hand, if $r = (r_U + r_V) \bmod (C)$, then $r \leq \min(r_U, r_V)$. Therefore,

$$X_U + rY_U \geq r\left\lceil \frac{D_U}{C} \right\rceil \text{ and } X_V + rY_V \geq r\left\lceil \frac{D_V}{C} \right\rceil.$$

Adding these inequalities and noting that $\left\lceil \frac{D_U}{C} \right\rceil + \left\lceil \frac{D_V}{C} \right\rceil \geq \left\lceil \frac{D_U+D_V}{C} \right\rceil \geq \left\lceil \frac{D_S}{C} \right\rceil$, we obtain inequality (8); thus, it cannot be facet.

A similar argument shows that T must be connected for (8) to be a facet.

2. If $D_{S,T} = 0$, then (8) is a linear combination of the nonnegativity constraints.

Sufficiency.

To prove that the cutset inequality defines a facet, we will use an interchange argument. This argument works as follows. We define the face

$$\mathfrak{F} = \{ (x,y,f) \in Conv(TFLP) : (x,y,f) \text{ satisfies (8) as an equality} \}$$

and prove that $\dim \mathfrak{F} = \dim(Conv(TFLP)) - 1$ by showing that any other valid inequality that is satisfied as an equality by all points in $\mathfrak{F}$ is a linear combination of (8) and the equality constraints.

Let

$$\sum_{\{i,j\} \in A} \alpha_{ij} x_{ij} + \sum_{\{i,j\} \in A} \beta_{ij} y_{ij} + \sum_{k \in K} \sum_{\{i,j\} \in A} \left(\gamma_{ij}^k f_{ij}^k + \gamma_{ji}^k f_{ji}^k \right) \geq \delta \qquad \text{(I.1)}$$

represent an arbitrary inequality that is satisfied as an equality by all $(\mathbf{x},\mathbf{y},\mathbf{f}) \in \mathfrak{L}$. In this expression, each coefficient α_{ij}, β_{ij}, γ_{ij}^k and δ is a real number. The interchange argument permits us to develop the desired relationship between these coefficients. Suppose the vectors $(\mathbf{x}^1,\mathbf{y}^1,\mathbf{f}^1)$ and $(\mathbf{x}^2,\mathbf{y}^2,\mathbf{f}^2)$ belong to $\mathfrak{L}$, and every component of $(\mathbf{x}^1,\mathbf{y}^1,\mathbf{f}^1)$ equals the corresponding component of $(\mathbf{x}^2,\mathbf{y}^2,\mathbf{f}^2)$ except for components x_{pq}^1 and x_{pq}^2. Substituting these two solutions in (I.1) and subtracting the resulting equations, we obtain $\alpha_{pq} = 0$. On the other hand, if all components of $(\mathbf{x}^1,\mathbf{y}^1,\mathbf{f}^1)$ and $(\mathbf{x}^2,\mathbf{y}^2,\mathbf{f}^2)$ are equal except that $x_{rs}^1 = x_{pq}^2 = 0$ and x_{pq}^1 and $x_{rs}^2 > 0$ (i.e., we have interchanged x_{pq}^1 LC facilities in $(\mathbf{x}^1,\mathbf{y}^1,\mathbf{f}^1)$ with x_{rs}^2 LC facilities in $(\mathbf{x}^2,\mathbf{y}^2,\mathbf{f}^2)$), then a similar substitution of both solutions $(\mathbf{x}^1,\mathbf{y}^1,\mathbf{f}^1)$ and $(\mathbf{x}^2,\mathbf{y}^2,\mathbf{f}^2)$ in (I.1) shows that $\alpha_{pq} / \alpha_{rs} = x_{rs}^2 / x_{pq}^1$.

Construct a feasible solution $(\mathbf{x}^0,\mathbf{y}^0,\mathbf{f}^0)$ satisfying (8) as an equality as follows.

For all commodities $k \in \{S,S\}$ (or $k \in \{T,T\}$) connect O(k) and D(k) by $\lceil d_k / C \rceil$ HC facilities along a path fully contained in S (in T). This choice is possible because of Condition 1 of the Theorem. Send a flow of d_k along this path from O(k) to D(k).

Choose a node $u \in S$ and a node $v \in T$ for which $\{u,v\} \in A$. For all commodities $k \in \{S,T\}$ with $O(k) \in S$, connect O(k) to u by $\lceil d_k / C \rceil$ HC facilities installed on a path $\{O(k),..., u\}$ fully contained in S. Similarly, connect v to D(k) by $\lceil d_k / C \rceil$ HC facilities installed on a path $\{v,..., D(k)\}$ fully contained in T. Send a flow of d_k along these paths. Next for all commodities $k \in \{S,T\}$ with $O(k) \in T$, send a flow of d_k along some paths $\{O(k),..., v\}$ fully contained in T and $\{u,..., D(k)\}$ fully contained in S on suitably installed HC facilites. Install $\lceil D_{S,T} / C \rceil$ HC facilities on arc $\{u,v\}$. Let $(f_{uv}^k)^0 = d_k$ for all $k \in \{S,T\}$ with $O(k) \in S$ and $(f_{vu}^k)^0 = d_k$ for all $k \in \{S,T\}$ with $O(k) \in T$. Thus, we obtain a feasible solution for which $X_{S,T} = 0$ and $Y_{S,T} = \lceil D_{S,T} / C \rceil$. This solution satisfies (8) as an equality.

Using the interchange argument with one of the solutions as $(\mathbf{x}^0, \mathbf{y}^0, \mathbf{f}^0)$, we can show that

(1) $\alpha_{ij} = \beta_{ij} = 0$ for all $\{i,j\} \in \{S,S\}$ or $\{T,T\}$,

(2) $r\alpha_{uv} = \beta_{uv}$, and since the choice of arc $\{u,v\}$ is arbitrary, $r\alpha_{ij} = \beta_{ij}$ for all $\{i,j\} \in \{S,T\}$, and

(3) $\gamma_{ij}^k = -\gamma_{ji}^k$ for all $\{i,j\} \in \{S,S\}$ or $\{T,T\}$, for all $k \in K$.

Now, consider arc $\{u,v\}$; since $r < C$ and $D_{S,T} > 0$, after we have installed $\lceil D_{S,T}/C \rceil$ HC facilities between nodes u and v, this link has a residual capacity of at least 1 unit. So define, for some $k_1 \in K$ and $0<\varepsilon\leq 1/2$,

$$\mathbf{y}^1 = \mathbf{y}^0$$

$$\mathbf{x}^1 = \mathbf{x}^0$$

$$\left(f_{uv}^{k_1}\right)^1 = \left(f_{uv}^{k_1}\right)^0 + \varepsilon$$

$$\left(f_{vu}^{k_1}\right)^1 = \left(f_{vu}^{k_1}\right)^0 + \varepsilon$$

$$\left(f_{ij}^{k}\right)^1 = \left(f_{ij}^{k}\right)^0 \text{otherwise.}$$

It is easy to verify that $(\mathbf{x}^1, \mathbf{y}^1, \mathbf{f}^1)$ is in $\mathfrak{L}$. Using the interchange argument again, we see that $\gamma_{uv}^k = \gamma_{vu}^k$ for $k = k_1$. But since we chose k_1 and $\{u,v\}$ arbitrarily, we conclude that $\gamma_{ij}^k = -\gamma_{ji}^k$ for all $\{i,j\} \in \{S,T\}$ and for all $k \in K$.

Using this result, we will first show that the sum of the γ coefficients corresponding to any cycle in the network equals zero. This result implies that

$$\sum_{k\in K} \sum_{\{i,j\}\in A} \left(\gamma_{ij}^k f_{ij}^k + \gamma_{ji}^k f_{ji}^k\right) \text{ is a constant.}$$

Consider any node r belonging to N. Let Δ_r denote the set of (directed) cycles originating and ending at node r from the arc set A. Note that since the arc set is undirected we may traverse a particular arc in both the directions and therefore Δ_r is nonempty. We assume that each arc is traversed at most once in each direction for all the cycles belonging to Δ_r.

Consider a particular cycle ς belonging to Δ_r. Call ς an *s-intersection cycle* if it contains exactly s $\{S,T\}$ cutset arcs. Note that two directed arcs (p,q) and (q,p) of ς may use the same undirected cutset arc $\{p,q\}$; in this case, these arcs add 1 to the intersection count, s.

Define the following feasible solution if ς is a 0 or 1-intersection cycle for some $k_1 \in K$ (we assume that arc $\{u,v\}$ is the common arc belonging to both $\{S,T\}$ and ς if ς is a 1-intersection cycle):

$$\mathbf{y}^2 = \mathbf{y}^0$$

$$x^2_{ij} = x^0_{ij} + 1 \text{ if } \{i,j\} \in \zeta\backslash\{S,T\},\ x^2_{ij} = x^0_{ij} \quad \text{otherwise}$$

$$\left(f^k_{ij}\right)^2 = \left(f^k_{ij}\right)^0 + \varepsilon,\ 0<\varepsilon\leq 1/2, \text{ if } (i,j)\in\zeta \text{ and } k = k_1$$

$$\left(f^k_{ij}\right)^2 = \left(f^k_{ij}\right)^0 \quad \text{otherwise.}$$

This solution maintains feasibility and satisfies (8) as an equality (and thus belongs to $\mathfrak{L}$). (Note that the upper bound on ε is necessary to account for the case when the residual capacity on arc $\{u,v\}$ — on which we have installed $\lceil D_{S,T} / C\rceil$ HC facilities — might be 1.) Comparing the coefficients of $(\mathbf{x^0}, \mathbf{y^0}, \mathbf{f^0})$ and $(\mathbf{x^2}, \mathbf{y^2}, \mathbf{f^2})$, we find that

$$\sum_{(i,j)\in\zeta} \gamma^k_{ij} = 0 \qquad \begin{cases} \text{for all 0 or 1-intersection cycles } \zeta\in\Delta_r, \\ \text{for all } r\in N, \text{ for all } k\in K. \end{cases}$$

Now, suppose ς is a 2-intersection cycle. Assume $\{u,v\}$ and $\{p,q\}$ are the cutset arcs belonging to cycle ς. Construct a solution $(\mathbf{x^3}, \mathbf{y^3}, \mathbf{f^3})$ as follows: send the flow of commodities belonging to $\{S,S\}$ (or $\{T,T\}$) on paths fully contained in S (or T) as we did for solution $(\mathbf{x^0}, \mathbf{y^0}, \mathbf{f^0})$. Let $y_{uv} = \left\lceil \frac{D_{S,T}}{C} \right\rceil - \frac{1}{2}$ and $y_{pq} = \frac{1}{2}$. By installing additional facilities on arcs in $\{S,S\}$ and $\{T,T\}$, route the commodities belonging to $\{S,T\}$ so that each of the arcs $\{u,v\}$ and $\{p,q\}$ contains at least $\frac{1}{2}$ units of residual capacity. We can determine a feasible flow that meets this condition because $r < C$. Note that

(i) this solution belongs to $\mathfrak{L}$, and

(ii) given this solution, we can send an additional flow of $\frac{1}{2}$ units along ς without increasing the capacity on the cutset arcs.

Now define

$$\mathbf{y}^4 = \mathbf{y}^3$$

$$x_{ij}^4 = x_{ij}^3 + 1 \text{ if } \{i,j\} \in \zeta \backslash \{S,T\},\ x_{ij}^4 = x_{ij}^3 \quad \text{otherwise}$$

$$\left(f_{ij}^k\right)^4 = \left(f_{ij}^k\right)^3 + \varepsilon,\ 0 < \varepsilon \leq 1/2, \text{ if } (i,j) \in \zeta \text{ and } k = k_1$$

$$\left(f_{ij}^k\right)^4 = \left(f_{ij}^k\right)^3 \quad \text{otherwise.}$$

Comparing the coefficients of $(\mathbf{x}^3, \mathbf{y}^3, \mathbf{f}^3)$ and $(\mathbf{x}^4, \mathbf{y}^4, \mathbf{f}^4)$, we conclude that

$$\sum_{(i,j) \in \zeta} \gamma_{ij}^k = 0 \qquad \begin{array}{l} \text{for all 2-intersection cycles } \zeta \in \Delta_r, \\ \text{for all } r \in N, \text{ for all } k \in K. \end{array}$$

Now consider an arbitrary s-intersection cycle ς. We will show that the sum of γ coefficients corresponding to the arcs of this cycle also equals 0. Let

$$\gamma_\zeta = \sum_{(i,j) \in \zeta} \gamma_{ij}^k$$

that is, γ_ζ is the sum of the γ's corresponding to the arcs of ζ. We will show that some 0-intersection cycle, say ψ, satisfies $\gamma_\zeta = \gamma_\psi$. Since we have already shown that $\gamma_\psi = 0$, this result would complete our argument.

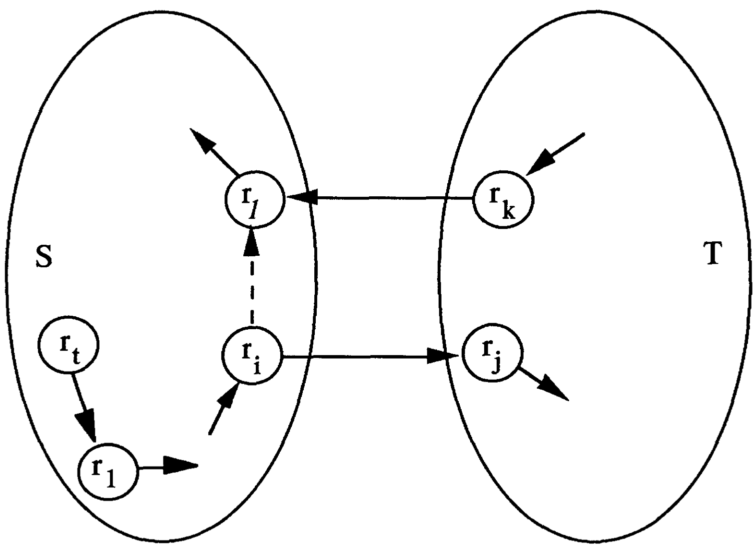


Figure I.1. Solid lines denote arcs of cycle ζ. The dashed arc (r_i,r_l) belongs to ψ.

Let ζ be defined by $\{(r_1,r_2), (r_2,r_3), \ldots, (r_t,r_1)\}$ as shown in Figure I.1. Suppose $r_1 \in S$. Let (r_i, r_j) be the first arc of the cycle that crosses the $\{S,T\}$ cutset and let (r_k, r_l) be the first subsequent arc that re-enters the set S. Notice that node r_i can equal r_l and/or r_k can equal r_j. In any case, $\{(r_i, r_j), \ldots, (r_k, r_l), (r_l, r_i)\}$ is a either a 1 or 2-intersection cycle. (Arc $\{r_l, r_i\}$ need not exist in the underlying network, but this condition does not change the essence of the following argument.) The sum of the γ's on this subcycle must equal 0, thus $\gamma_{r_i r_l}$ equals the sum of the γ's on the path $(r_i, r_j), \ldots, (r_k, r_l)$. Thus, we can replace the path $(r_i, r_j), \ldots, (r_k, r_l)$ by the arc (r_i, r_l). Repeating this argument, if necessary, we can construct a 0-intersection cycle ψ that satisfies $\gamma_\zeta = \gamma_\psi$.

We have now shown that

$$\sum_{(i,j)\in\zeta} \gamma_{ij}^k = 0 \qquad \text{for all } \zeta \in \Delta_r, \text{ for all } r \in N, \text{ for all } k \in K.$$

The above argument also implies

$$\sum_{(i,j)\,\in\,\text{path } \pi} \gamma_{ij}^k = \text{constant, say, } \gamma_{pq}^k \quad \text{for any (directed) path } \pi \text{ connecting nodes p and q.}$$

In particular, suppose we chose $p = O(k)$ and q arbitrarily in this argument, then the sum of the γ's for all $\{i,j\}$ (with proper signs) belonging to any path connecting $O(k)$ and q is the same. Let $\gamma_{O(k)q}^k$ denote this quantity. Thus by setting $v_{O(k)}^k = 0$, we can find *unique* multipliers v_i^k satisfying the condition $v_j^k - v_i^k = \gamma_{ij}^k = -\gamma_{ji}^k$. Now, using these multipliers for the flow conservation constraints, we obtain

$$\sum_{\{i,j\}\in A} \left\{\left(v_j^k - v_i^k\right)f_{ij}^k + \left(v_i^k - v_j^k\right)f_{ji}^k\right\} = \sum_{\{i,j\}\in A} \left(\gamma_{ij}^k f_{ij}^k + \gamma_{ji}^k f_{ji}^k\right)$$

$$= \left(\gamma_{O(k)D(k)}^k\right)d_k$$

$$= v_{D(k)}^k d_k.$$

This equality implies

$$\sum_k \sum_{\{i,j\} \in A} \left(\gamma_{ij}^k f_{ij}^k + \gamma_{ji}^k f_{ji}^k\right) = \sum_k \gamma_{O(k)D(k)}^k d_k$$

$$= \sum_k v_{D(k)}^k d_k$$

$$= \text{constant.}$$

We can now show that $\alpha_{ij} = \alpha$ and $\beta_{ij} = \beta$ for all $\{i,j\} \in \{S,T\}$. Choose $\{p,q\} \in \{S,T\}$, so that $\{p,q\} \neq \{u,v\}$, and $k_1 \in K$. Let P(u,p) be a path from node u to node p fully contained in S, and P(q,v) be a path from node q to node v fully contained in T. Define

$$y_{uv}^5 = y_{uv}^0 - 1$$

$$y_{ij}^5 = y_{ij}^0 \text{ for all } \{i,j\} \neq \{u,v\}$$

$$x_{uv}^5 = r - 1$$

$$x_{pq}^5 = 1$$

$$x_{ij}^5 = 1 \text{ if } \{i,j\} \in P(u,p) \text{ or } P(q,v)$$

$$x_{ij}^5 = 0 \text{ otherwise}$$

$$\left(f_{ij}^k\right)^5 = \left(f_{ij}^k\right)^0 + 1 \text{ if } (i,j) \in P(u,p) \text{ or } P(q,v) \text{ and } k = k_1$$

$$\left(f_{uv}^k\right)^5 = \left(f_{uv}^k\right)^0 - 1,\ \left(f_{pq}^k\right)^5 = 1 \text{ for } k = k_1$$

$$\left(f_{ij}^k\right)^5 = \left(f_{ij}^k\right)^0 \text{ otherwise.}$$

Using the interchange argument on the solutions $(\mathbf{x^0}, \mathbf{y^0}, \mathbf{f^0})$ and $(\mathbf{x^5}, \mathbf{y^5}, \mathbf{f^5})$, we see that

$$\alpha_{uv} + \gamma_{uv}^k = \alpha_{pq} + \sum_{(i,j) \in P(u,p)} \gamma_{ij}^k + \gamma_{pq}^k + \sum_{(i,j) \in P(q,v)} \gamma_{ij}^k$$

for $k = k_1$. Since the sum of the γ's corresponding to any cycle equals zero, we see that $\alpha_{uv} = \alpha_{pq}$. Furthermore, since arc $\{p,q\}$ was chosen arbitrarily, we obtain $\alpha_{ij} = \alpha$ and, thus, $\beta_{ij} = \beta = r\alpha$, for all $\{i,j\} \in \{S,T\}$.

Thus, (I.1) is equivalent to $\alpha X_{S,T} + r\alpha Y_{S,T} + \text{constant} = \alpha^*$, which implies $\alpha X_{S,T} + r\alpha Y_{S,T} = \alpha_0$. Since (I.1) is nonvacuous, $\alpha \neq 0$. Consequently, $X_{S,T} + rY_{S,T} = \alpha_0/\alpha = r\lceil D_{S,T}/C\rceil$ since (8) holds as an equality for all points in $\mathfrak{L}$. ⊗

Theorem 6. *The arc residual capacity inequality (11) defines a facet of the extended TFLP model if and only if*

1. *If $r_L = C$, then $L = K$.*

2. *If $\{i,j\}$ is a bridge arc, then $L = K(i,j)$.*

Proof.

Necessity.

If $r_L = C$, and $L \subset K$, then the arc residual capacity inequality is dominated by the capacity constraint for arc $\{i,j\}$. Now, suppose, that $\{i,j\}$ is a bridge arc and let $G = L \cap K(i,j)$ and $H = L \setminus G$. Also, for simplicity of notation, let $r_{ij} = r_{K(i,j)}$, $D_{ij} = D_{K(i,j)}$ and $\mu_{ij} = \mu_{K(i,j)}$. Since $\sum_{k \in G} \left(f_{ij}^k + f_{ji}^k\right) = D_G$ and $\sum_{k \in H} \left(f_{ij}^k + f_{ji}^k\right) = 0$, the arc residual capacity inequality is equivalent to $x_{ij} + r_L y_{ij} \geq \mu_L\, r_L - D_H$. If $L = K(i,j)$, then this inequality becomes $x_{ij} + r_{ij} y_{ij} \geq \mu_{ij}\, r_{ij}$.

We first show that $\mu_G\, r_L \geq \mu_L\, r_L - D_H$. Since $\mu_G = (D_G + C - r_G)/C$ and $\mu_L = (D_G + D_H + C - r_L)/C$, we can write $\mu_G\, r_L - \mu_L\, r_L + D_H$ as $r_L(r_L - r_G)/C + D_H(1 - r_L/C)$ which is nonnegative if $r_L \geq r_G$. If $r_L < r_G$, then $r_L < \min(r_G, r_H) \leq D_H$ and, therefore, $r_L(r_L - r_G)/C + D_H(1 - r_L/C) \geq D_H(r_L - r_G)/C + D_H(1 - r_L/C) \geq 0$. Thus $\mu_G\, r_L \geq \mu_L\, r_L - D_H$.

Case (i). $r_L \leq r_{ij}$.

In this case, we show that the arc residual capacity inequality $x_{ij} + r_{ij}\, y_{ij} \geq \mu_{ij}\, r_{ij}$ for $L = K(i,j)$ dominates the arc residual capacity inequality for the given choice of L.

Since $\mu_{ij} \geq \mu_G$, the arc residual capacity inequality for L = K(i,j) is stronger than $x_{ij} + r_{ij}\, y_{ij} \geq \mu_G\, r_{ij}$. The last inequality dominates $x_{ij} + r_L\, y_{ij} \geq \mu_G\, r_L$ if $r_L \leq r_{ij}$ and $y_{ij} \leq \mu_G$ (if $y_{ij} > \mu_G$, the inequality $x_{ij} \geq 0$ implies that $x_{ij} + r_L\, y_{ij} \geq \mu_G\, r_L$). Since $\mu_G\, r_L \geq \mu_L\, r_L - D_H$, the necessity of Condition 2 follows if $r_L \leq r_{ij}$.

Case (ii). $r_L > r_{ij}$.

If r_L is greater than r_{ij}, then consider the following linear combination of $x_{ij} + r_{ij}\, y_{ij} \geq \mu_{ij}\, r_{ij}$ (the arc residual capacity inequality for L = K(i,j)) and $x_{ij} + C\, y_{ij} \geq D_{ij}$ (the aggregate capacity demand inequality across arc {i,j}):

$$\left(\frac{C-r_L}{C-r_{ij}}\right)(x_{ij} + r_{ij}y_{ij}) + \left(\frac{r_L-r_{ij}}{C-r_{ij}}\right)(x_{ij} + Cy_{ij}) \geq \left(\frac{C-r_L}{C-r_{ij}}\right)\mu_{ij}r_{ij} + \left(\frac{r_L-r_{ij}}{C-r_{ij}}\right)D_{ij}.$$

Simplifying this inequality, we obtain $x_{ij} + r_L\, y_{ij} \geq r_L\, \mu_{ij} + r_{ij} - r_L$. The righthand side of this inequality is greater than $r_L\, \mu_G$ if $\mu_G < \mu_{ij}$ and so the residual capacity inequality is no stronger than a weighted combination of the other two constraints. So assume that $\mu_G = \mu_{ij}$. This assumption implies that $\mu_L \geq \mu_{ij}$, thus

$$\begin{aligned} x_{ij} + r_L y_{ij} &\geq r_L\mu_{ij} + r_L(\mu_L-\mu_{ij}) - C(\mu_L-\mu_{ij}) + r_{ij} - r_L \\ &= r_L\mu_L - C(\mu_L-\mu_{ij}) + r_{ij} - r_L \\ &= r_L\mu_L - D_L + D_{ij} \\ &\geq r_L\mu_L - D_L + D_G \\ &= r_L\mu_L - D_H. \end{aligned}$$

Therefore, since the residual capacity constraint for a bridge arc is implied by a weighted combination of two valid inequalities, it cannot be a facet.

Sufficiency.

We will use an interchange argument, similar to the one used for Theorem 4, to prove the sufficiency part of the theorem. As earlier, define $\mathfrak{L}$ to be the set of points that belong to Conv(TFLP) and satisfy (11) as an equality. Let (I.1) be an arbitrary inequality that is satisfied as an equality by all points belonging to $\mathfrak{L}$. First, construct a feasible solution $(\mathbf{x^0},\mathbf{y^0},\mathbf{f^0})$ that belongs to $\mathfrak{L}$.

Consider a (nonbridge) arc {u,v}. For each $k \in K\backslash L$, install $\lceil d_k/C \rceil$ HC facilities on a path connecting O(k) and D(k) that does not contain arc {u,v} and set $f^k_{ij} = d_k$ for all arcs lying on this path.

For each $k \in L$, consider a path connecting O(k) and D(k) that contains arc $\{u,v\}$ and install $\lceil \sum_{k \in L} d_k/C \rceil$ HC facilities on all arcs of this path except arc $\{u,v\}$. Now send d_k units by of flow for all $k \in L$ by installing $\lceil \sum_{k \in L} d_k/C \rceil$ HC facilities on arc $\{u,v\}$.

Arguments similar to those used to prove Theorem 4 permit us to show that

(1) $\alpha_{ij} = \beta_{ij} = 0$ for all $\{i,j\} \neq \{u,v\}$,

(2) $r_L \alpha_{uv} = \beta_{uv}$,

(3) $$\sum_{(i,j) \in \zeta} \gamma_{ij}^k = 0 \begin{cases} \text{for all cycles } \zeta, \text{ for all } k \in K \backslash L \\ \text{for all cycles } \zeta \text{ for which neither } (u,v) \text{ nor } (v,u) \in \zeta, \text{ for all } k \in L, \end{cases}$$

(4) $\alpha_{uv} + \sum_{(i,j) \in \zeta} \gamma_{ij}^k = 0$ for all cycles ζ for which (u,v) or $(v,u) \in \zeta$, for all $k \in L$.

Set $v_{O(k)}^k = 0$ for all k and define:

$$\theta_{ij}^k = \begin{cases} \gamma_{ij}^k & \text{for all } k \in K \backslash L, \text{ for all } (i,j) \\ \gamma_{ij}^k & \text{for all } k \in L, \text{ for all } (i,j) \neq (u,v) \text{ or } (v,u) \\ \gamma_{ij}^k + \alpha_{uv} & \text{for all } k \in L \text{ and for } (i,j) = (u,v) \text{ or } (v,u). \end{cases}$$

We can now find unique multipliers, using θ_{ij}^k as arc lengths, so that $v_j^k - v_i^k = \theta_{ij}^k$ for all (i,j) and for all k. Multiplying the flow conservation constraint for node i, commodity k by multiplier v_i^k, and adding, we obtain

$$\sum_{k \in K} \sum_{\{i,j\} \in K} \left\{ \left(v_j^k - v_i^k\right) f_{ij}^k + \left(v_i^k - v_j^k\right) f_{ji}^k \right\} = \sum_{k \in K} v_{D(k)}^k d_k$$

or

$$\sum_{k \in K} \sum_{\{i,j\} \in K} \left\{ \gamma_{ij}^k f_{ij}^k + \gamma_{ij}^k f_{ji}^k \right\} + \sum_{k \in L} \alpha_{uv} \left(f_{uv}^k + f_{vu}^k \right) = \textit{some constant, say } \Theta.$$

Thus, inequality (I.1) is equivalent to

$$\Theta - \sum_{k \in L} \alpha_{uv}\left(f_{uv}^{k} + f_{vu}^{k}\right) + \alpha_{uv}x_{uv} + r_L\alpha_{uv}y_{uv} \geq \delta$$

which proves the theorem. $\otimes$

References

BALAKRISHNAN, A., T. L. MAGNANTI, AND R. T. WONG. 1989. A Dual-Ascent Procedure for Large Scale Uncapacitated Network Design. *Operations Research,* **37**, 716–740.

BALAKRISHNAN, A., T. L. MAGNANTI, AND R. T. WONG. 1990a. Local Access Telecommunication Network Expansion: Modeling and Polyhedral Characterization, to appear.

BALAKRISHNAN, A., T. L. MAGNANTI, AND R. T. WONG. 1990b. A Decomposition Algorithm for Local Access Telecommunications Network Expansion Planning, to appear.

BARAHONA, F., M. GROETSCHEL, M. JUENGER, AND G. REINELT. 1988. An Application of Combinatorial Optimization to Statistical Physics and Circuit Layout Design. *Operations Research,* **36**, 493–513.

BARANY, I., T. J. VAN ROY, AND L. A. WOLSEY. 1984. Uncapacitated Lot-Sizing: The Convex Hull of Solutions. *Mathematical Programming Study,* **22**, 32–43.

BERTSEKAS, D., AND R. GALLAGER. 1987. *Data Networks.* Prentice-Hall, New Jersey.

GAREY, M. R., AND D. S. JOHNSON. 1979. *Computers and Intractibility: A Guide to the Theory of NP-Completeness.* W. H. Freeman & Co., San Fransisco, CA.

GEOFFRION, A. M. 1974. Lagrangian Relaxation and its uses in Integer Programming. *Mathematical Programming Study,* **2**, 82–114.

GEOFFRION, A. M., AND G. GRAVES. 1974. Multicommodity Distribution System Design by Benders Decomposition. *Management Science,* **5**, 822–844.

GROETSCHEL, M., AND O. HOLLAND. 1988. Solution of Large-Scale Symmetric Traveling Salesman Problems. Report No. 73, Institut fuer Mathematik, Universitaet Augsburg, Augsburg.

GROETSCHEL, M., M. JUENGER, AND G. REINELT. 1985a. A Cutting Plane Algorithm for the Linear Ordering Problem. *Operations Research,* **32**, 1195–1220.

GROETSCHEL, M., M. JUENGER, AND G. REINELT. 1985b. Facets for the Linear Ordering Problem. *Mathematical Programming,* **33**, 43–60.

GROETSCHEL, M., L. LOVASZ, AND A. SCHRIJVER. 1981. The Ellipsoid Method and its Consequences in Combinatorial Optimization. *Combinatorica,* **1**, 169–197.

GROETSCHEL, M., AND C. L. MONMA. 1988. Integer Polyhedra Arising From Certain Network Design Problems with Connectivity Constraints. Report No. 104, Institut fuer Mathematik, Universitaet Augsburg, Augsburg.

GROETSCHEL, M., AND M. W. PADBERG. 1985. Polyhedral Theory, in *The Traveling Salesman Problem*, eds. E. Lawler et al., John Wiley and Sons, New York.

HOFFMAN, K., AND M. W. PADBERG. 1985. LP-Based Combinatorial Problem Solving. *Annals of Operations Research,* **4**, 145–195.

KARP, R. M., AND C. H. PAPADIMITRIOU. 1982. On the Linear Characterization of Combinatorial Optimization. *SIAM Journal of Computing,* **11**, 299–311.

KNUTH, D. E. 1981. Seminumerical Algorithms, 2nd ed., Vol. 2 of *The Art of Computer Programming,* Addison Wesley, Reading, MA.

LEUNG, J., T. L. MAGNANTI, AND V. SINGHAL. 1991. Routing in Point-to-Point Delivery Systems, Transportation Science, to appear.

LEUNG, J., T. L. MAGNANTI, AND R. VACHANI. 1989. Facets and Algorithms for Capacitated Lot Sizing. *Mathematical Programming,* **45**, 331–359.

MAGNANTI, T. L., P. MIRCHANDANI, AND R. VACHANI. 1990. The Convex Hull of Two Core Capacitated Network Design Problems, Operations Research Centre Working Paper No. OR 217-90, MIT, Cambridge, MA.

MAGNANTI, T. L., AND R. VACHANI. 1990. A Strong Cutting Plane Algorithm for Production Scheduling with Changeover Costs. *Operations Research,* **38**, 456–473.

MAGNANTI, T. L., AND R. T. WONG. 1984. Network Design and Transportation Planning. *Transportation Science,* **18**, 1–55.

MARTIN, R. K. 1987. Generating Alternative Mixed Integer Programming Models using Variable Redefinition. *Operations Research,* **35**, 820–831.

MIRCHANDANI, P. 1989. *Polyhedral Structure of a Capacitated Network Design Problem with an Application to the Telecommunication Industry.* Ph. D. Dissertation, MIT Sloan School of Management, Cambridge, MA.

NEMHAUSER, G. L., AND L. A. WOLSEY. 1988. *Integer and Combinatorial Optimization.* Wiley-Interscience, New York.

PADBERG, M. W., T. J. VAN ROY, AND L. A. WOLSEY. 1983. Valid Inequalities for Fixed Charge Problems. *Operations Research,* **33**, 842–861.

POCHET, Y. 1988. Valid Inequalities and Separation for Capacitated Economic Lot Sizing. *Operations Research Letters,* **7**, 105–116..

POWELL, W. B., AND Y. SHEFFI. 1983. The Load Planning Problem of LTL Motor Carriers: Problem Description and a Proposed Solution Approach. *Transportation Research,* **17A**, 471–480.

PRESS, W. H., B. P. FLANNERY, S. A. TEUKOLSKY, AND W. T. VETTERLING. 1988. *Numerical Recipes: The Art of Scientific Computing,* Cambridge University Press, Cambridge, MA.

PULLEYBLANK, W. 1989. Polyhedral Combinatorics, in *Optimization,* eds G. L. Nemhauser et al., North Holland, Amsterdam.

SCHRIJVER, A. 1986. *Theory of Linear and Integer Programming.* Wiley-Interscience, New York.

VACHANI, R. 1988. Design of Private Telecommunication Networks, TM 0027-07-88-446, GTE Laboratories Incorporated, Waltham, MA 02254.

VAN ROY, T. J. AND L. A. WOLSEY. 1984. Solving Mixed Integer Programs by Automatic Reformulation, *Operations Research,* **35**, 45–57.

WOLSEY, L. A. 1989. Strong Formulations for Mixed-Integer Programming: A Survey. *Mathematical Programming,* **41**, 173–191.

WONG, R. T. 1980. Integer Programming Formulations of the Traveling Salesman Problem. *Proceedings of 1980 IEEE International Conference on Circuits and Computers,* 149–152.

WONG, R. T. 1984. A Dual Based Procedure for Solving Steiner Tree Problems on a Directed Graph. *Mathematical Programming,* **28**, 271–287.