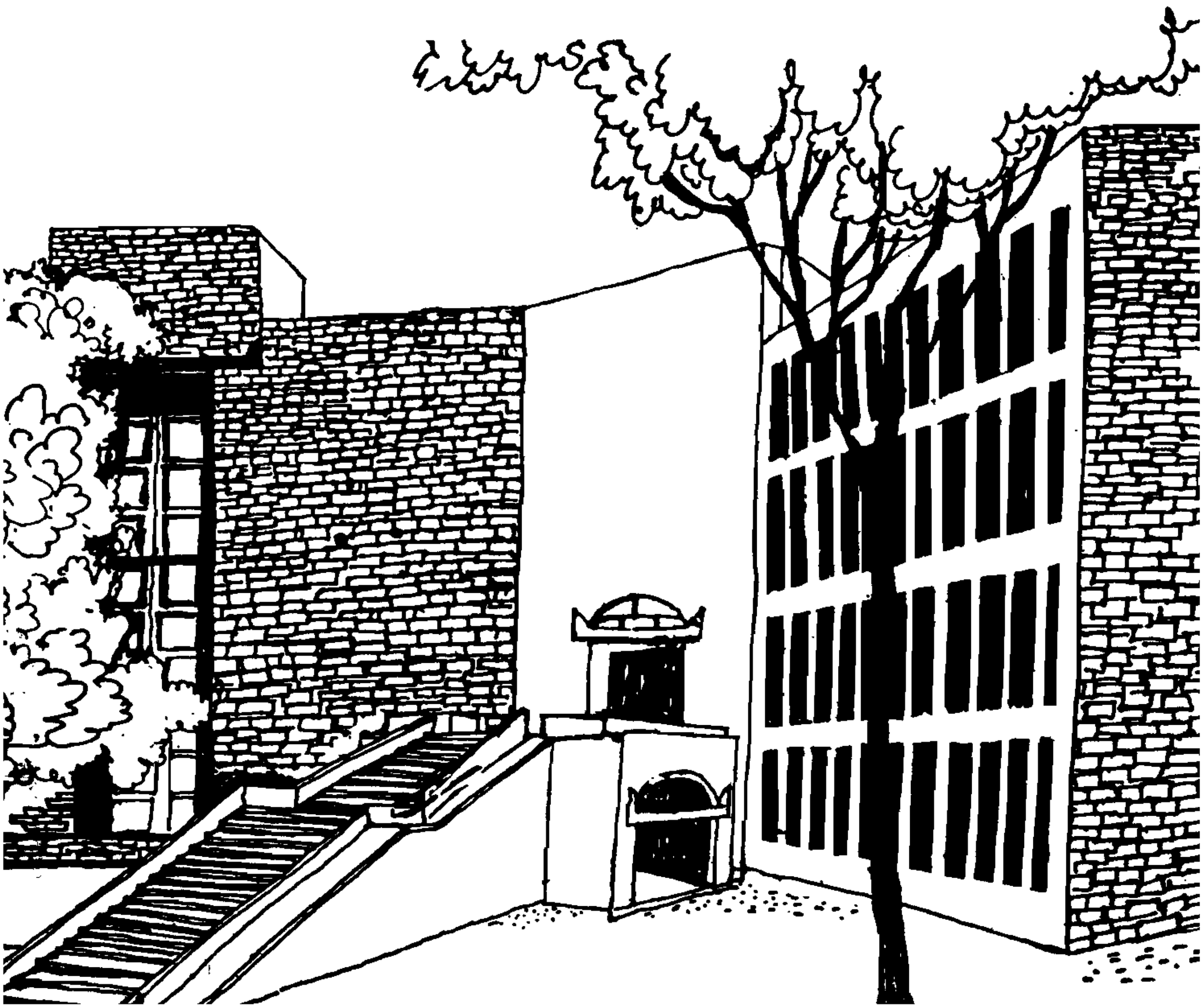
विद्याविनियोगाद्विकासः

IIM

AHMEDABAD

Working Paper



AN EXACT FORMULATION AND ALGORITHM FOR TWO COMMODITY CAPACITATED NETWORK DESIGN

By

Trilochan Sastry

W.P.No.99-10-09 /1554
October 1999

INDIAN INSTITUTE OF MANAGEMENT
AHMEDABAD-380 015
INDIA

AN EXACT FORMULATION AND ALGORITHM FOR TWO COMMODITY CAPACITATED NETWORK DESIGN

Trilochan Sastry

Indian Institute of Management, Ahmedabad

email: sastry@iimahd.ernet.in

October 1999

Abstract

We study the capacitated version of the two commodity network design problem, where capacity can be purchased in batches of C units on each arc at a cost of $w_{ij} \geq 0$, and $d_k \geq 0$ units of flow are sent from source to sink for each commodity k. We characterize optimal solutions for the problem with fixed costs and no flow costs, and show that either $\lfloor d_k/C \rfloor C$ or $(\lfloor d_k/C \rfloor - 1)C$ units of each commodity are sent on a shortest path, and the remaining flows possibly share arcs. We show that the problem can be solved in polynomial time. Next, we describe an exact linear programming formulation, i.e., one that guarantees integer optimal solutions, using $O(m)$ variables and $O(n)$ constraints. We also interpret the dual variables and constraints of the formulation as generalizations of the arc constraints and node potentials for the shortest path problem. Finally, we discuss several other variations of the single and two commodity problems.

Key words and phrases: capacitated network design, two commodity, integer optima

1 Introduction

This paper studies the capacitated two commodity network design problem, and is a sequel to Sastry [6], which studies the uncapacitated problem. The uncapacitated version of the multicommodity problem has been studied by Balakrishnan et.al. [1], who have solved large instances of the problem to optimality. Hu [3], Sakarovitch [5] and Seymour [7] have studied the two commodity maximum flow problem with no fixed or flow costs on arcs. However, the capacitated version has received relatively less attention. At the same time, the capacitated problem is important and has applications in telecommunications and transportation where large investments are required for building such networks. It is therefore important to obtain efficient solution procedures.

The capacitated two commodity network design problem **CTC** can be described as follows. Consider an undirected graph $G = (N, A)$, with node set N, arc set A, and origin destination pairs s_k, t_k, with demand of d_k units between each pair for $k = 1, 2$. Capacity can be purchased in batches of C units on each arc $(i, j) \in A$ at cost $w_{ij} \geq 0$. Flow costs are assumed to be zero. The objective is to minimise the total cost while satisfying demand between both origin destination pairs. The problem can be formulated as follows.

Problem CTC

$$\text{Min} \sum_{(i,j)\in A} w_{ij} y_{ij}$$

subject to:

$$\begin{aligned} \sum_j (x^k_{ji} - x^k_{ij}) &= \begin{cases} -d_k & \text{if } i = s_k \\ d_k & \text{if } i = t_k \\ 0 & \text{otherwise} \end{cases} \\ C y_{ij} &\geq x^1_{ij} + x^1_{ji} + x^2_{ij} + x^2_{ji} \\ x, y &\geq 0; y \text{ integer.} \end{aligned}$$

Let $m = |A|$ and $n = |N|$ denote the number of arcs and nodes respectively. The arcs are undirected and have symmetric cost, i.e., $w_{ij} = w_{ji}$. The flow variables x^k_{ij} are directed. This problem is also known as the *network loading* problem. In the next section we characterise optimal solutions of **CTC**, and in Section 3, describe a simple $O(n^2)$ algorithm to solve the problem. In

Section 4 we give an explicit reformulation for the problem in $O(m)$ variables and $O(n)$ constraints and show that it always has an integer optimal solution. In Section 5 we discuss several other variations of the one and two commodity network design problems, and in Section 6 we present the conclusions.

2 Characterizing Optimal Solutions

In this Section, we characterize optimal solutions. We describe three types of solutions, called the *independent*, *star* and *shared* solutions, and later, show that one of these is an optimal solution for **CTC**. This leads naturally to the polynomial algorithm described in the next Section. We first define a few terms. A *free arc* (i,j) has $0 < x_{ij}^1 + x_{ji}^1 + x_{ij}^2 + x_{ji}^2 < Cy_{ij}$. A *free cycle* has only free arcs in the cycle. Assume without loss of generality that $y_{ij} = \lceil (x_{ij}^1 + x_{ji}^1 + x_{ij}^2 + x_{ji}^2)/C \rceil$ for all arcs in any optimal solution, since all costs are nonnegative. Assume also that both x_{ij}^k and x_{ji}^k are not greater than zero in any optimal solution. Assume that $s_k \neq t_k$ for either commodity: otherwise there is only one commodity at most, and the problem reduces to the shortest path problem. Similarly, assume that either $s_1 \neq s_2$ or $t_1 \neq t_2$.

Lemma 1 *There exists an optimal solution with no free cycles.*

Proof
Suppose there is a free cycle Φ. We say that arc $(i,j) \in \Phi(c)$ if the arc is in the cycle, and node i comes before node j when we traverse the arc as we go around the cycle in a clockwise direction. Let

$$\Phi^+(k) = \{(i,j) \in \Phi(c) : x_{ij}^k \geq 0\},$$

$$\Phi^-(k) = \{(i,j) \in \Phi(c) : x_{ji}^k \geq 0\},$$

denote the set of arcs that have nonnegative quantity of commodity k flow on it in the same and opposite directions respectively. Let

$$\Delta^- = \min \{Cy_{ij} - x_{ji}^1 - x_{ji}^2 : (i,j) \in \Phi^-(1) \cap \Phi^-(2)\},$$

$$\Delta^-(1) = \min \{Cy_{ij} - x_{ji}^1 - x_{ij}^2 : (i,j) \in \Phi^-(1) \cap \Phi^+(2)\},$$

$$\Delta^-(2) = \min \{Cy_{ij} - x_{ij}^1 - x_{ji}^2 : (i,j) \in \Phi^-(2) \cap \Phi^+(1)\},$$

$$\Delta^+(k) = \min \{x_{ij}^k > 0 : (i,j) \in \Phi^+(k)\}.$$

Let $\Delta = \min \{\Delta^-, \ \Delta^-(1), \ \Delta^-(2), \ \Delta^+(1), \ \Delta^+(2)\}$. By the definition of a free cycle, $\Delta^- > 0$, $\Delta^-(1) > 0$, and $\Delta^-(2) > 0$, and hence $\Delta > 0$. If $\Delta = \Delta^-(1) = Cy_{ij} - x_{ji}^1 - x_{ij}^2$ for some arc $(i,j) \in \Phi^-(1) \cap \Phi^+(2)$, then by sending Δ units of commodity 1 in the anticlockwise direction around the cycle, we obtain a solution where arc (i,j) is no longer free. Similarly, if Δ equals $\Delta^-(2)$ or Δ^-, then by sending Δ units of commodity 2 in the anticlockwise direction around the cycle, we obtain a solution where arc (i,j) is no longer free.

If $\Delta^+(1) = \Delta = x_{ij}^1$ for some arc $(i,j) \in \Phi^+(1)$, then by sending Δ units of commodity 1 in the anticlockwise direction, we obtain a solution with $x_{ij}^1 = 0$. If $x_{ij}^2 + x_{ji}^2 = 0$, we are done. Otherwise, calculate Δ again, and repeat the procedure. Either $\Delta = \min \{\Delta^-, \ \Delta^-(1), \ \Delta^-(2)\}$, in which case total flow on some arc equals y_{ij} and we are done, or we obtain a solution where flow of some commodity on some arc reduces to zero. Notice that once flow of commodity k on arc $(i,j) \in \Phi^+(k)$ reduces to zero, the flow in the clockwise direction does not increase in subsequent iterations. By repeating this procedure at most m times, we either obtain a solution with no free cycle, or a solution in which each arc in the cycle has only one commodity flowing on it. If we now repeat the procedure once more, we are done.

Remark 1 *If there are commodity specific flow costs, there could be an optimal solution with two nodes i^* and j^* that have two arc distinct paths between them, carrying $r_k < C$ units of commodity k on each path. However, re-routing of a commodity flow from one path to the other might not be possible because flow costs on the other path are too high.*

Hereafter, we only consider optimal solutions with no free cycles. To simplify the notation, we assume that in any feasible solution (x,y) for **CTC**, we replace arc (i,j) by y_{ij} parallel arcs. Let this be the graph $G(x,y) = (N, A^*)$ associated with the feasible solution (x,y). Arbitrarily assign C units of commodity k flow to $\lfloor (x_{ij}^k + x_{ji}^k)/C \rfloor$ of these parallel arcs. If

$$0 < \sum_{k=1}^{2} (x_{ij}^k + x_{ji}^k - C\lfloor (x_{ij}^k + x_{ji}^k)/C \rfloor) \le C,$$

assign the remaining flow to another parallel arc. Otherwise, if $0 < x_{ij}^1 + x_{ji}^1 - C\lfloor (x_{ij}^1 + x_{ji}^1)/C \rfloor$, assign this quantity to one arc, and if $0 < x_{ij}^2 + x_{ji}^2 -$

$C\lfloor(x_{ij}^2 + x_{ji}^2)/C\rfloor$, assign it to another arc. If the flow on an arc satisfies the conditions

$$0 < x_{ij}^k + x_{ji}^k - C\lfloor(x_{ij}^k + x_{ji}^k)/C\rfloor \text{ for both commodities, and}$$

$$\sum_{k=1}^{2}(x_{ij}^k + x_{ji}^k - C\lfloor(x_{ij}^k + x_{ji}^k)/C\rfloor) \leq C,$$

then the arc is *shared*, otherwise it is *independent*. For any feasible solution, a *path* P between nodes i and j connects the two nodes and has capacity C on each arc in the path. However, there may be several paths of capacity C between nodes i and j.

For any commodity k, a *complete* path P_k connects s_k to t_k. Path P is *shared* if it has at least one shared arc, and a *shared forward path* (*shared reverse path*) if both commodities flow in the same direction (opposite directions) on all shared arcs. A path if *fully shared* if all arcs in it are shared. A path is *independent* if none of the arcs in the path is shared. Note that this definition implies that no cost is saved even if both commodities flow between the same pair of nodes on independent paths. Commodity k has a *complete independent path* if no complete path for the other commodity shares any arcs with it. An *in junction* node for commodity k has commodity k flow entering it from at least two arcs, and flow leaving it on exactly one arc. An *out junction* node for commodity k has commodity k flow leaving it on at least two arcs, and flow entering it on exactly one arc.

Let $d_k = \mu_k C + r_k$ where μ_k is a nonnegative integer and $0 < r_k \leq C$. Thus, if d_k is a multiple of C, we define $r_k = C$. Consider any feasible solution with only independent paths. We can re-route all independent flows of commodity k on any shortest path from s_k to t_k without increasing cost. Thus we can obtain a solution of at most equal cost by setting $y_{ij} = \mu_k + 1$ on all arcs $(i,j) \in G(N,A)$ on a shortest path from s_k to t_k: however, if some arc (i,j) belongs to the shortest path for both commodities, we set $y_{ij} = \mu_1 + \mu_2 + 2$. Such a solution is *independent*. For example, in Figure 1, if $C = 10$. $d_1 = d_2 = 6$, and commodity 1 flows on path $s_1 - k_1 - k_2 - l_1 - t_1$ and commodity 2 flows on path $s_2 - k_2 - k_1 - l_2 - t_2$, then we have an independent solution. A cycle *with one end* (*with two ends*) is a cycle with capacity C on all arcs, and has one node (two nodes) through which all flows enter or leave the cycle.

Lemma 2 *There exists an optimal solution that has no cycles with one or two ends.*

Proof

If there is a cycle with one end, we can delete it. Let P, Q be two paths between nodes i^* and j^* with capacity C on all arcs, such that all flows enter or leave the cycle through these two nodes. Without loss of generality assume that $\sum_{(ij)\in P} w_{ij} \leq \sum_{(ij)\in Q} w_{ij}$. Decrease y_{ij} by 1 on all arcs $(i,j) \in Q$, and increase y_{ij} by 1 on all arcs $(i,j) \in P$.

Corollary 1 *There exists an optimal solution (x, y) such that any two paths P^1, P^2 in $G_{(x,y)}$ have at most one fully shared connected path $P = P^1 \cap P^2$.*

Proof

Consider any optimal solution (x, y), and any flow decompostion into paths. If paths P^1 and P^2 are independent, then they do not share any arcs in $G(x.y)$. Suppose the paths are not independent. Consider all shared arcs $(i, j) \in P^1 \cap P^2$. If these arcs do not form a connected path, then there must be a cycle with two ends.

□

Next. we describe the so called *shared* solution. Suppose $\mu_k C$ units of commodity k flow on a shortest path from s_k to t_k, and the remaining r_k units flow on a shared path. This is a *shared* solution, which as we show later, is possible only if $r_1 + r_2 \leq C$. Since the shared solution has at most one shared path for each commodity, Corollary 1 implies that there is one fully shared path. Hence all shared arcs are either shared forward arcs. in which case there is a *shared forward* solution, or they are shared reverse arcs. in which case there is a *shared reverse* solution. For instance. in Figure 2. if $r_1 + r_2 \leq C$, $\mu_k C$ units of commodity k flow on arc (s_k, t_k), r_1 units on path $s_1 - k - l - t_1$, and r_2 units on path $s_2 - k - l - t_2$, we have a shared forward solution. If r_2 units of commodity 2 flow on path $s_2 - l - k - t_2$, then we have a shared reverse solution.

Next. we describe the so called *star* solutions. Suppose s_1, s_2, t_1, t_2 are four distinct nodes, and suppose there is one out junction node k_1, and one in junction node l_1 for commodity 1, and one out junction node k_2, and one in junction node l_2 for commodity 2, such that k_1, k_2, l_1, l_2 are distinct nodes.

Suppose commodity 1 has two complete independent paths, P_1^1, P_2^1 that can be partitioned into arc distinct paths in $G(x, y)$ such that

$$P_1^1 = P_1(s_1, k_1) \cup P(k_1, k_2) \cup P(k_2, l_1) \cup P_1(l_1, t_1),$$

$$P_2^1 = P_2(s_1, k_1) \cup P(k_1, l_2) \cup P(l_2, l_1) \cup P_2(l_1, t_1),$$

where paths $P_1(s_1, k_1)$, $P_2(s_1, k_1)$ connect nodes s_1 and k_1, and path $P(i, j)$ connects nodes i and j. Notice that $P_1(s_1, k_1)$ and $P_2(s_1, k_1)$ are independent, and that they are arc distinct in $G(x, y)$. Commodity 2 has two complete shared paths P_1^2, P_2^2 each of which can be partitioned into arc distinct paths such that

$$P_1^2 = P(s_2, k_2) \cup P(k_2, k_1) \cup P(k_1, l_2) \cup P(l_2, t_2).$$

$$P_2^2 = P(s_2, k_2) \cup P(k_2, l_1) \cup P(l_1, l_2) \cup P(l_2, t_2),$$

and $P_1^2 \cap P_2^2 = P(s_2, k_2) \cup P(l_2, t_2)$. The four paths $P(k_1, k_2)$, $P(k_2, l_1)$, $P(l_1, l_2)$, $P(l_2, k_1)$ are all fully shared and arc distinct in $G(N, A)$, and form a *star*. Further, commodity 1 has a total of $C + r_1$ units flowing on paths P_1^1, P_2^1, and commodity 2 has a total of r_2 units flowing on P_1^2, P_2^2. For instance, $r_2/2$ units of commodity 2 flow on paths P_1^2, P_2^2, $C - r_2/2$ units of commodity 1 on path P_1^1, and $r_1 + r_2/2$ units of commodity 1 on path P_2^1. The remaining $(\mu_1 - 1)C$ units of commodity 1 and $\mu_2 C$ units of commodity 2 flow on complete independent paths. This is the *first star* solution (see Figure 1 and Example 1).

Similarly, if commodity 2 has two complete independent paths, P_1^2, P_2^2, and commodity 1 has two complete shared paths P_1^1, P_2^1, we have a *second star* solution. In this solution, commodity 2 has a total of $C + r_2$ units flowing on paths P_1^2, P_2^2, and commodity 1 has r_1 units flowing on P_1^1, P_2^1. The remaining $\mu_1 C$ units of commodity 1 and $(\mu_2 - 1)C$ units of commodity 2 are sent on complete independent paths (see Figure 1 and Example 1).

If $k_1 = k_2$, we have a star solution with a triangle having paths $P(k_1, l_1)$, $P(k_1, l_2)$, $P(l_1, l_2)$ and four paths connecting s_1 to k_1, s_2 to k_1, l_1 to t_1, and l_2 to t_2. Similarly, if $k_1 = l_2$, $l_1 = k_2$ or $l_1 = l_2$, we have a star with a triangle. If $s_1 = s_2$, $s_1 = t_2$, $s_2 = t_1$ or $t_1 = t_2$, then we have a star with a triangle, and three paths connecting each vertex of the triangle to one of the terminals. In each of these cases, there are corresponding first and second star solutions. If $k_1 = k_2$ and $l_1 = l_2$, then there is a cycle with two ends, which by Lemma 2 we need not consider. Thus a star either has three or four vertices.

INSERT FIGURE 1 HERE

The two commodity flow problem finds the maximum flow in a network if capacities are given on each arc. Let $c(s_1, t_1)$ and $c(s_2, t_2)$ denote the minimum capacity of any cut separating s_1 from t_1, and s_2 from t_2 respectively. Let $c(s, t)$ denote the minimum capacity of any cut separating s_1, s_2 from t_1, t_2 or s_1, t_2 from t_1, s_2. Let $f(k)$ denote the net value of commodity k flow leaving s_k for $k = 1, 2$. The following theorem is due to Hu [3].

Theorem 1 *(Hu [3]) Flows $f(1)$, $f(2)$ are feasible for any two commodity flow problem if and only if $f(1) \leq c(s_1, t_1)$, $f(2) \leq c(s_2, t_2)$ and $f(1) + f(2) \leq c(s, t)$.*

Given any feasible solution (x, y) to **CTC**, consider the graph $G(x, y) = (N, A^*)$. Further, let capacity of each arc be either zero or one. Let the two commodity maximum flow problem on this graph be the problem *associated* with a feasible solution for **CTC**. Hu [3] also showed that there always exists an optimal solution such that flow of a commodity on every arc is in multiples of 0.5 units. For a given commodity, a *full path* has 1 unit of flow on each arc in the path, and a *half path* has 0.5 units of flow on each arc in the path. Let $X(1, 2) \subset N$ be a subset of nodes or a cut such that

$$s_1, s_2 \in X, t_1, t_2 \in N - X, \text{ or } s_1, t_2 \in X, t_1, s_2 \in N - X.$$

and let $A(X(1, 2))$ be the set of arcs with exactly one node in $X(1, 2)$. Define

$$c(s, t) = \min \{ \sum_{(ij) \in A(X(1,2))} y_{ij} : X(1, 2) \subset N \}.$$

Cut capacities $c(s_1, t_1)$ and $c(s_2, t_2)$ are similarly defined for cuts

$$X(k) \subset N : s_k \in X(k),\ t_k \in N - X(k).$$

If $c(s, t) > \mu_1 + \mu_2 + 2$, then $c(s_k, t_k) > \mu_k + 1$ for some commodity since $c(s_1, t_1) + c(s_2, t_2) \geq c(s, t) > \mu_1 + \mu_2 + 2$. Suppose without loss of generality that $c(s_1, t_1) > \mu_1 + 1$. We can reduce y_{ij} by 1 on some complete path from s_1 to t_1 and still have a feasible flow for **CTC**. Assume therefore, that corresponding to any optimal solution for **CTC**, $c(s, t) = \mu_1 + \mu_2 + 2$ if $r_1 + r_2 > C$ and $c(s, t)$ equals $\mu_1 + \mu_2 + 2$ or $\mu_1 + \mu_2 + 1$ if $r_1 + r_2 \leq C$ in the associated problem. Define an arc in the associated problem to be *shared* if both commodities have 0.5 units flowing on it.

Theorem 2 *There exists an optimal solution OS to problem* **CTC** *such that*
(i) if $r_1 + r_2 > C$, *then OS is independent,*
(ii) if $r_1 + r_2 \leq C$ *and* $C < d_1$, $C < d_2$, *then OS is either independent, first star, second star or shared,*
(iii) if $r_1 + r_2 \leq C$, *and* $C < d_1$, $d_2 \leq C$, $(d_1 \leq C, C < d_2)$, *then OS is either independent, first star, (second star), or shared,*
(iv) if $d_1 + d_2 \leq C$, *then OS is either independent or shared.*

Proof

Consider any optimal solution (x, y) with no free cycles or cycles with one or two ends, and its associated two commodity flow problem. If any complete half path P_1 for commodity k in a maximum flow is independent. then we can increase flow on it by 0.5 units. Therefore, consider only maximum flows where each complete half path of one commodity shares arcs with one or more complete half paths of the other commodity.

Suppose $r_1 + r_2 > C$. Then $c(s,t) = \mu_1 + \mu_2 + 2$, and hence, by Theorem 1, there is a maximum flow such that $f(k) = \mu_k + 1$ for each commodity in the asociated problem. Since the arc and cut capacities are all integer in the associated problem. there are an even number of complete half paths for each commodity. Arbitrarily number all half paths P_j^1 for commodity 1 from 1 through $2p$ and half paths P_j^2 for commodity 2 from 1 through $2q$, where $2p$ and $2q$ are the number of half paths for commodities 1 and 2 respectively. Let K_1 be the total cost of all shared arcs Q_1 in $\{P_1^1 \cup \ldots \cup P_p^1\} \cap \{P_1^2 \cup \ldots \cup P_q^2\}$, i.e.. arcs that have flow of 0.5 units of both commodities. Similarly. let K_2 be the total cost of all shared arcs Q_2 in $\{P_{p+1}^1 \cup \ldots \cup P_{2p}^1\} \cap \{P_{q+1}^2 \cup \ldots \cup P_{2q}^2\}$. If $K_2 \geq K_1$. we reduce the capacity of all shared arcs in Q_2 by 1. and add a parallel arc of capacity one on all shared arcs in Q_1. Notice that we can now send full flows on each of the original half paths $P_1^1, \ldots, P_p^1$, $P_1^2, \ldots, P_q^2$ since shared arcs in Q_1 now have sufficient capacity. Thus we obtain a solution for the associated problem with only full flows. Since there are at least $\mu_k + 1$ full paths for each commodity, we can send $(\mu_k + 1)C \geq d_k$ units of each commodity on independent paths in the original problem **CTC**. Hence there is an independent solution. A similar argument establishes that we can obtain an independent solution if $K_2 > K_1$.

Suppose $r_1 + r_2 \leq C$. Suppose $c(s,t) = \mu_1 + \mu_2 + 2$. Then using the same arguments as in the case $r_1 + r_1 > C$, we can obtain an independent solution. Suppose $c(s,t) = \mu_1 + \mu_2 + 1$. It follows that there must be at least one shared path in the solution to **CTC**. Further, as shown earlier, we only need to consider optimal solutions for the associated problem where each half path of one commodity shares arcs with one or more half paths of the other commodity. Consider two maximum flows in the associated problem. Solution 1 has a maximum flow of $\mu_1 + 1$ for commodity 1, and μ_2 for commodity 2. Solution 2 has a maximum flow of $\mu_2 + 1$ for commodity 2 and μ_1 for commodity 1. By Theorem 1, there always exist such maximum flows. Since maximum flows in the associated problem are in multiples of 0.5, there are two cases: (i) either there is exactly one path for each commodity on which 1 unit more of commodity k flows in solution k, or, (ii) there are two shared paths for each commodity, on which 0.5 units more of commodity k flow in solution k. Consider the following cases.

Case 1.

Let P^1 and P^2 be two complete paths for commodity 1 and 2 respectively. Suppose that in the associated maximum flow problem, P^1 has 1 unit of flow and P^2 has zero units of flow in solution 1, and P^1 has zero units of flow and P^2 has 1 unit of flow in solution 2. Further, μ_k units of commodity k flow on paths other than P^1, P^2 in $G(x,y)$. By Theorem 1, there exist such solutions. Consider all paths other than P^1 and P^2. Using arguments identical to those in the case when $r_1 + r_2 > C$, we can send these flows on independent paths. If P^1 and P^2 do not share arcs, we can increase the maximum flow. If they share both forward and reverse arcs, then there is a cycle with two ends in **CTC**. Hence, we either have a shared forward solution or a shared reverse solution, but not both.

Case 2.

Suppose there are complete paths P_1^1, P_2^1 for commodity 1 and P_1^2, P_2^2 for commodity 2 such that P_1^1, P_2^1 have 0.5 units of flow more in solution 1, and P_1^2, P_2^2 have 0.5 units of flow more in solution 2. For instance, in Figure 1, if $y_{s_1k_1} = y_{l_1t_1} = 2$, $y_{k_1l_1} = 0$, and $y_{ij} = 1$ for all other arcs (i,j), we obtain the two solutions mentioned.

Suppose flow on each of these four paths is either zero or 0.5 in the two solutions. This implies that we send at most C units of either commodity on these paths in **CTC**, and hence, we can send at least $\mu_k C$ units of each commodity on other paths. Let K_1 be the total cost of shared arcs Q_1 in

$P_1^1 \cap P_2^2$, and K_2 the cost of shared arcs Q_2 in $P_2^1 \cap P_1^2$. Assume without loss of generality that $K_2 \geq K_1$. Increase y_{ij} by 1 for all arcs (i,j) in Q_1 and decrease it by 1 for all arcs in Q_2. We can obtain a solution to **CTC** without increasing cost: send $\mu_k C$ units of both commodities on paths other than P_1^1, P_2^1, P_1^2, P_2^2. The remaining r_k units for both commodities can be sent on paths P_1^1, P_2^2.

This also establishes that if $d_1 \leq C$, then we need not consider first star optimal solutions: at most d_1 units can flow on P_1^1, P_2^1, and hence, by shifting capacity from arcs in Q_2 to arcs in Q_1, we can obtain a shared or independent solution. Similarly, if $d_2 \leq C$, then we need not consider second star optimal solutions.

Suppose flow on each of the paths P_1^1, P_2^1 is 1 unit in solution 1 and 0.5 units in solution 2, and flow on paths P_1^2, P_2^2 is zero in solution 1 and 0.5 units in solution 2. This implies that we can send $C + r_1$ units of commodity 1 and r_2 units of commodity 2 on these paths in **CTC**. Assume we can send at most $(\mu_1 - 1)C$ units of commodity 1 on paths other than P_1^1, P_2^1 in **CTC**. since otherwise, we can reduce y_{ij} by 1 on all arcs in some complete commodity 1 path. Using arguments identical to the case when $r_1 + r_2 > C$, we can send all flows other than those on $P_1^1, P_2^1, P_1^2, P_2^2$ on independent paths without increasing cost. We can still shift capacity from arcs in Q_2 to arcs in Q_1. However, some segment of path P_1^1 for commodity 1 has capacity of C in **CTC**, and we cannot send the entire $C + r_1$ units on it. Thus we cannot disturb the four paths. Suppose for one of these four paths, say P_1^2, there is some arc $(i,j) \in P_1^2$ such that $(i,j) \notin P_2^2 \cup P_1^1 \cup P_2^1$. Then we must have a free cycle or a cycle with two ends in the original problem **CTC**. Hence, there is no such arc, and we must have a first star. A similar argument shows that if commodity 2 had unit flow in solution 1 and half flow in solution 2, then we have a second star solution.

□

Example 1

Consider Figure 1, where costs are shown beside each arc, and suppose $C = 10$. If $r_1 = r_2 = 6$, then $r_1 + r_2 > C$, and we have an independent solution. For instance if $w \geq 25$, then

$$y_{s_1k_1} = y_{k_1k_2} = y_{l_1t_1} = \mu_1 + 1,\ y_{k_2l_1} = \mu_1 + \mu_2 + 2,\ y_{s_2k_2} = y_{l_2l_1} = y_{l_2t_2} = \mu_2 + 1$$

is an optimal solution. If $w = 18$ and $d_1 = 11$, $d_2 = 6$, then $r_1 + r_2 \leq C$, and we have the following first star solution of cost 50

$$y_{s_1k_1} = y_{l_1t_1} = 2, \ y_{k_1k_2} = y_{k_2l_1} = y_{k_1l_2} = y_{l_2l_1} = y_{s_2k_2} = y_{l_2t_2} = 1,$$

$$x^1_{s_1k_1} = x^1_{l_1t_1} = 11, \ x^1_{k_1k_2} = x^1_{k_2l_1} = 7, \ x^1_{k_1l_2} = x^1_{l_2l_1} = 4,$$
$$x^2_{s_2k_2} = x^2_{l_2t_2} = 6, \ x^2_{k_2k_1} = x^2_{k_1l_2} = 3, \ x^1_{k_2l_1} = x^1_{l_1l_2} = 3.$$

If $d_1 = 11$ and $d_2 = 16$, then we obtain a second star solution of cost 68. However, if $d_1 = 1$, $d_2 = 6$, then we obtain a reverse shared solution of cost 35 where commodity 1 flows on path $s_1 - k_1 - k_2 - l_1 - t_1$, and commodity 2 on path $s_2 - k_2 - k_1 - l_2 - t_2$. Similarly, if $d_1 = 11$, $d_2 = 6$ and $w = 12$, we obtain a reverse shared solution of cost 44 where the two commodities flow on paths $s_1 - k_1 - l_1 - t_1$, and $s_2 - k_2 - l_1 - k_1 - l_2 - t_2$. If $w = 12$, $w_{k_1l_2} = 100$, and $d_1 = 11$, $d_2 = 6$, then there is an independent solution of cost 49. If $w = 18$, and $w_{k_1l_2} = 100$, then there is a forward shared solution of cost 58 where 10 units flow on path $s_1 - k_1 - l_1 - t_1$, 1 unit on path $s_1 - k_1 - k_2 - l_1 - t_1$, and 6 units on path $s_2 - k_2 - l_1 - l_2 - t_2$.

3 Algorithm

Consider the two cases $r_1 + r_2 > C$ and $r_1 + r_2 \leq C$ separately. If $r_1 + r_2 > C$, then find the shortest path from s_k to t_k for both commodities. Let $a(i^*, j^*)$ denote the cost of the shortest path from node i^* to node j^* using w_{ij} as arc costs. Then the optimal cost is $(\mu_1 + 1)a(s_1, t_1) + (\mu_2 + 1)a(s_2, t_2)$. Suppose $r_1 + r_2 \leq C$. We find the least cost independent, shared and star solutions, and choose the one with minimum cost. The least cost shared solution can be found as follows. Find the shortest path from s_k to t_k for both commodities. Send $\mu_k C$ units of flow on these paths at a cost of $\mu_k a(s_k, t_k)$. The remaining r_k units can be sent on a shared path using the so called *two path* algorithm as follows. The *forward* (*reverse*) problem is to find an optimal solution if any shared path is a shared forward (reverse) path. For the forward problem we define $O_k = s_k$ and $D_k = t_k$, while for the reverse problem, we define $O_1 = s_1$, $D_1 = t_1$, $D_2 = s_2$ and $O_2 = t_2$.

Add a super source O, a super destination D, arcs (O, j) of cost $a(O_1, j) + a(O_2, j)$, and arcs (j, D) of cost $a(j, D_1) + a(j, D_2)$. The two path algorithm then uses any standard algorithm to find the shortest path between O and D. Notice that there are two passes for the algorithm, one for the forward problem, and the other for the reverse problem. Choose the shorter of the two shortest paths. If arc (O, j) belongs to the shortest path, replace it by the shortest paths from O_1 to j and from O_2 to j. Similarly, if arc (j, D) belongs to the shortest path, replace it by the shortest paths from j to D_1 and from j to D_2. Sastry [6] has shown that this gives the optimal solution, and that if $\pi_j(f)$ and $\pi_j(b)$ are the distance labels for each node $j \in N$, in the forward and reverse problem, then

$$\pi_j(f) = \min \ \{a(s_1, i) + a(s_2, i) + a(i, j) : i \in N\}, \text{ and,}$$

$$\pi_j(b) = \min \ \{a(s_1, i) + a(t_2, i) + a(i, j) : i \in N\}.$$

Algorithm two path takes $O(n^2)$ iterations if the distances $a(s_k, j)$ and $a(j, t_k)$ are known. However, these distances can be obtained by finding the simple shortest path trees rooted at nodes s_k and t_k in at most $O(n^2)$ time. The complexity of the two path algorithm is therefore $O(n^2)$. Let $OPT(s)$ be the cost of the optimal shared solution.

The first star solution can be found similarly. Let $\pi_j^1 = \min \ \{a(s_1, i) + a(t_1, i) + a(i, j) : i \in N\}$ denote the minimum cost of connecting nodes s_1, t_1 to node j. Note that these values can be found using the two path algorithm as follows. Create a new super source S, and connect arc (S, j) of cost $a(s_1, j) + a(t_1, j)$ between every node j and node S. Solve the shortest path problem between node S and all other nodes. Thus it takes $O(n^2)$ time to solve the problem. Then the optimal first star solution costs $OPT(1) = \pi_{s_2}^1 + \pi_{t_2}^1$. Similarly, let π_j^2 denote the minimum cost of connecting nodes s_2, t_2 to node j. Then the optimal second star solution costs $OPT(2) = \pi_{s_1}^2 + \pi_{t_1}^2$. The total optimal cost is therefore

$$OPT = \sum_{k=1}^{2} \mu_k a(s_k, t_k) + \min \ \{\sum_{k=1}^{2} a(s_k, t_k), OPT(s), OPT(k) - a(s_k, t_k) : k = 1, 2\}.$$

Finding $OPT(k)$ takes $O(n^2)$ for either commodity. Hence problem **CTC** can be solved in $O(n^2)$ time.

4 An Exact Formulation

We now describe an *exact* linear programming formulation for the problem and show that it always has an integer optimal solution. The formulation is suggested by the algorithm described in the previous section. When $r_1 + r_2 > C$, then the following formulation **TP** guarantees optimal integer solutions.

$$\text{Min } \nu = \sum_{(i,j)\in A} w_{ij}(f_{ij}^1 + f_{ij}^2)$$

subject to :

$$\sum_i (f_{ij}^k - f_{ji}^k) = \begin{cases} -\mu_k - 1 & j = s_k \\ \mu_k + 1 & j = t_k \\ 0 & \text{otherwise} \end{cases}$$

$$\text{All variables} \geq 0$$

Formulation **TP** is really two separate shortest path problems and hence guarantees integer solutions. Now consider the case when $r_1 + r_2 \leq C$.

Let $z_0 = 1$ if there is an independent solution, and 0 otherwise.
Let $z_s = 1$ if there is a shared solution, and 0 otherwise.
Let $z_1 = 1$ if there is a first star solution, and 0 otherwise
Let $z_2 = 1$ if there is a second star solution, and 0 otherwise.

Consider the first star solution. Let k_1, l_1 be the out and in junction nodes for commodity 1, and k_2, l_2 the out and in junction nodes for commodity 2. To find the minimum cost first star, we assume that $2z_1$ units flow out of node s_1, split into 2 at node k_1, and z_1 units enter nodes k_2 and l_2. This is called the "s_1 *flow*" which occurs *before* node k_2 or l_2. Similarly, $2z_1$ units flow out of node t_1, split into 2 at node l_1, and z_1 units enter nodes k_2 and l_2. This is the "t_1 *flow*" which occurs before node k_2 or l_2. Finally, z_1 units of commodity 2 flow from k_2 to s_2 and from l_2 to t_2. Commodity 2 flow occurs *after* node k_2 or l_2. We define the following variables.

For each arc $(i,j) \in A$ let :
$e_{ij}^1(1)$ equal the quantity of s_1 flow on the arc
$g_{ij}^1(1)$ equal the quantity of t_1 flow on the arc
$f_{ij}^2(1) = 1$ only if it has commodity 2 flow on it after k_2 or l_2
For each node $j \in N$ let $u_j^1 = 1$ if junction node j equals k_2 or l_2
Let $z_1 = 1$ if there is a first star solution.

Consider the following formulation $\mathbf{S(1)}$.

$$\text{Min } \nu(1) = \sum_{(i,j)\in A} w_{ij}(e^1_{ij}(1) + g^1_{ij}(1) + f^2_{ij}(1))$$

subject to :

$$\begin{aligned}
\sum_i (e^1_{ij}(1) - e^1_{ji}(1)) - u^1_j &= \begin{cases} -2z_1 & j = s_1 \\ 0 & \text{otherwise} \end{cases} \\
\sum_i (g^1_{ij}(1) - g^1_{ji}(1)) - u^1_j &= \begin{cases} -2z_1 & j = t_1 \\ 0 & \text{otherwise} \end{cases} \\
\sum_i (f^2_{ij}(1) - f^2_{ji}(1)) + u^1_j &= \begin{cases} z_1 & j = s_2 \text{ or } t_2 \\ 0 & \text{otherwise} \end{cases} \\
e, g, f, u, z_1 &\geq 0
\end{aligned}$$

Consider any first star solution to **CTC**. We can obtain a corresponding solution for $\mathbf{S(1)}$ as follows. Set $z_1 = 1$, and $e^1_{ij}(1) = 2$ for all arcs (i,j) between s_1 and k_1, in the star. Similarly, set $g^1_{ij}(1) = 2$ for all arcs (i,j) between t_1 and l_1 in the star. Next, set $u^1_{k_2} = u^1_{l_2} = 1$, and $e^1_{ij}(1) = 1$ on all arcs (i,j) between nodes k_1, k_2 and between nodes k_1, l_2, and $g^1_{ij}(1) = 1$ on all arcs (i,j) between nodes l_1, k_2 and between nodes l_1, l_2 in the star. Finally, set $f^2_{ij}(1) = 1$ for all arcs (i,j) between nodes k_2, s_2 and between nodes l_2, t_2 in the star. Let $S^1(*)$ be the polyhedron defined by these constraints. The second star formulation $\mathbf{S(2)}$ can be similarly described using variables $e^2_{ij}(2)$, $g^2_{ij}(2)$, $f^1_{ij}(2)$, and u^2_j. Let $S^2(*)$ be the polyhedron defined by the constraints of $\mathbf{S(2)}$. It is easy to establish the following result.

Lemma 3 *Given a first star (second star) solution to* CTC *of cost* $W + (\mu_1 - 1)a(s_1,t_1) + \mu_2 a(s_2,t_2)$, $(W + \mu_1 a(s_1,t_1) + (\mu_2 - 1)a(s_2,t_2))$ *there exists a solution for* $\mathbf{S(1)}$ $(\mathbf{S(2)})$ *of cost* W.

Sastry [6] has described an exact linear programming formulation (one that guarantees integer optimal solutions) for the uncapacitated two commodity network design problem with fixed costs and no flow costs. This formulation is based on a characterization of optimal solutions, where each commodity flows on one path, and the two paths are either shared or independent. This suggests that we can use a similar formulation for the shared solution. For the sake of completeness, we describe the formulation here, and later show how it can be used to describe an exact formulation for **CTC**.

We define shared path P to be *maximal* if all shared arcs belong to it. Suppose there is a shared maximal path from node i^* to node j^* with commodity 1 flowing from i^* to j^*. We say that the shared path *starts* in node i^* and *ends* in node j^*. Let $\delta \in \{f, b\}$ be a parameter that denotes whether there is a forward or reverse solution respectively. If $k = 1$ or $\delta = f$, then flow of commodity k on arc (i, j) is said to occur *before* the shared path if the flow has not yet entered node i^* and is said to occur *after* the shared path if it has left node j^*. If $k = 2$ and $\delta = b$, then flow of commodity 2 on arc (i, j) is said to occur *before* the shared path if the flow has not yet entered node j^* and is said to occur *after* the shared path if it has left node i^*. We define the following 0-1 variables.

For each node $j \in N$ let:

$u_j(\delta) = 1$ if the shared path starts in node j

$v_j(\delta) = 1$ if the shared path ends in node j

For each arc $(i, j) \in A$ let :

$e^k_{ij} = 1$ only if commodity k flow occurs before the shared path

$g^k_{ij} = 1$ only if commodity k flow occurs after the shared path

$h_{ij}(\delta) = 1$ if it is a shared arc.

Consider the following formulation for the shared problem **SF**.

$$\text{Min } \nu(s) = \sum_{(i,j)\in A} w_{ij}(e^1_{ij} + e^2_{ij} + g^1_{ij} + g^2_{ij} + h_{ij}(f) + h_{ij}(b))$$

subject to :

$$\begin{aligned}
\sum_i (e^1_{ij} - e^1_{ji}) - u_j(f) - u_j(b) &= \begin{cases} -z_s & j = s_1 \\ 0 & \text{otherwise} \end{cases} \\
\sum_i (e^2_{ij} - e^2_{ji}) - u_j(f) - v_j(b) &= \begin{cases} -z_s & j = s_2 \\ 0 & \text{otherwise} \end{cases} \\
\sum_i (h_{ij}(\delta) - h_{ji}(\delta)) + u_j(\delta) - v_j(\delta) &= 0 \\
\sum_i (g^1_{ij} - g^1_{ji}) + v_j(f) + v_j(b) &= \begin{cases} z_s & j = t_1 \\ 0 & \text{otherwise} \end{cases} \\
\sum_i (g^2_{ij} - g^2_{ji}) + v_j(f) + u_j(b) &= \begin{cases} z_s & j = t_2 \\ 0 & \text{otherwise} \end{cases} \\
e, g, h, u, v, z_s &\geq 0
\end{aligned}$$

Let SC be the polyhedron defined by these constraints. It is straightforward to verify that given a complete shared path for each commodity, we can obtain a corresponding feasible solution for SC (see also Sastry [6]). We can now reformulate the capacitated two commodity network design problem **CTC** as follows. Let $\xi = 1$ if $r_1 + r2 > C$ and zero otherwise. For $k = 1, 2$ and $l = 1, 2$, define $f_{ij}^k(l)$ as zero if $k = l$. Define $e_{ij}^k(l)$, $g_{ij}^k(l)$ as zero if $k \neq l$.

Reformulation **CTC(R)**.

$$\text{Max } \nu = \sum_{(ij)\in A} w_{ij}[h_{ij}(f)+h_{ij}(b)+\sum_{k=1}^{2}(f_{ij}^k+e_{ij}^k+g_{ij}^k+\sum_{l=1}^{2}(f_{ij}^k(l)+e_{ij}^k(l)+g_{ij}^k(l)))]$$

subject to

$$\begin{aligned}
\sum_i (f_{ij}^k - f_{ji}^k) &= \begin{cases} -\mu_k - z_0 + z_k - \xi & j = s_k \\ \mu_k + z_0 - z_k + \xi & j = t_k \\ 0 & \text{otherwise} \end{cases} \\
z_0 + z_s + z_1 + z_2 + \xi &= 1 \\
e_{ij}^k, g_{ij}^k, h_{ij}(\delta), u_j(\delta), v_j(\delta) &\in SC \text{ for } k = 1, 2,\ \delta = f, b \\
e_{ij}^1(1), g_{ij}^1(1), f_{ij}^2(1), u_j^1 &\in S^1(*) \\
e_{ij}^2(2), g_{ij}^2(2), f_{ij}^1(2), u_j^2 &\in S^1(*)
\end{aligned}$$

Notice that if $r_1 + r_2 > C$ and $\xi = 1$, then all variables associated with the shared and star solutions are zero, and the problem reduces to **TP**. The reformulation has $28m + 6n + 4$ zero-one variables: 4 variables for z_0, z_s, z_1, z_2, and an additional $12m + 4n$ for the shared solution (where we count e_{ij}^k and e_{ji}^k separately), $6m + n$ for each of the star solutions, and $4m$ for variables f_{ij}^k. The original formulation for **CTC** has $5m$ variables. The reformulation has $14n$ constraints: $6n$ for the shared, and $3n$ for each of the first star and second star solutions, and $2n$ for variables f_{ij}^k. The original formulation has $O(m + n)$ constraints. Thus if $m = O(n^2)$ as in the case of complete graphs, the reformulation has fewer constraints. We show later that **CTC(R)** always has an integer optimal solution. First we motivate the result by interpreting the dual variables.

4.1 Interpretation of dual variables and constraints

In this section, we consider the dual **DCTC(R)** of **CTC(R)** and interpret the dual variables at each node as an upper or lower bound on the distance from one of the terminal nodes, or as a bound on the cost of sending one unit from a source sink pair, s_k, t_k to one of the other terminals. The dual constraints corresponding to the first star solution, except the constraint for z_1 are

$$\begin{aligned}
\alpha_j^1(1) - \alpha_i^1(1) &\leq w_{ij} \\
\beta_j^1(1) - \beta_i^1(1) &\leq w_{ij} \\
\eta_j^2(1) - \eta_i^2(1) &\leq w_{ij} \\
\eta_j^2(1) - \alpha_j^1(1) - \beta_j^1(1) &\leq 0.
\end{aligned}$$

Let $DS^1(*)$ denote the polyhedron described by these constraints. Notice that the first three constraints of the dual are identical to those of the simple (one commodity) shortest path problem and correspond to the primal variables $e_{ij}^1(1)$, $g_{ij}^1(1)$ and $f_{ij}^2(1)$. In the simple shortest path problem, with origin s, destination t, and dual variables π_j, we can find at least two sets of optimal dual node potentials: (i) $\pi_j = a(s,j)$, or (ii) $\pi_j = a(s,t) - a(j,t)$. Either of these sets of values gives a dual optimal solution. The first set of dual values can be interpreted as an upper bound on the node distances or how 'far' nodes can be from s, and the second set of values as a lower bound on node distances, or how 'close' nodes can be to s given that the shortest distance from s to t is $a(s,t)$.

As we show later, the optimal value of dual variables $\alpha_j^1(1)$ is $a(s_1,j)$, the upper bound on distances from s_1, and $\beta_j^1(1)$ is $a(t_1,j)$, the upper bound on distances from t_1. The variables $\eta_j^2(1)$ equal π_j^1 which is the least cost of connecting nodes s_1, t_1, j in the first star solution.. The last constraint, corresponding to primal variable u_j^1, states that the minimum cost π_j^1 of connecting nodes s_1, t_1 to j is at most equal to $a(s_1,j) + a(t_1,j)$.

The constraints for the second star solution, except the constraint for z_2 are

$$\begin{aligned}
\alpha_j^2(2) - \alpha_i^2(2) &\leq w_{ij} \\
\beta_j^2(2) - \beta_i^2(2) &\leq w_{ij} \\
\eta_j^1(2) - \eta_i^1(2) &\leq w_{ij} \\
\eta_j^1(2) - \alpha_j^2(2) - \beta_j^2(2) &\leq 0.
\end{aligned}$$

Let $DS^2(*)$ denote the polyhedron described by these constraints. The interpretation of the dual is identical to that of the first star solution with the roles of the two commodities reversed. The dual constraints corresponding to the shared solution, except the constraint for z_s are

$$\begin{aligned}
\alpha_j^k - \alpha_i^k &\leq w_{ij} \\
\beta_j^k - \beta_i^k &\leq w_{ij} \\
\eta_j(\delta) - \eta_i(\delta) &\leq w_{ij} \\
\eta_j(f) - \alpha_j^1 - \alpha_j^2 &\leq 0 \\
\beta_j^1 + \beta_j^2 - \eta_j(f) &\leq 0 \\
\eta_j(b) + \beta_j^2 - \alpha_j^1 &\leq 0 \\
\beta_j^1 - \alpha_j^2 - \eta_j(b) &\leq 0.
\end{aligned}$$

Let DSC denote the polyhedron described by these constraints. As we show later, the optimal node potential α_j^k, which corresponds to the flow conservation equation for e_{ij}^k, equals $a(s_k, j)$, the upper bound on the distance from s_k to node j. Similarly, $\eta_j(f)$ corresponds to the flow conservation equation for $h_{ij}(f)$, and equals $\pi_j(f)$, the minimum cost of reaching j from nodes s_1, s_2. The node potentials β_j^k correspond to the flow conservation equations for g_{ij}^k, and equal $OPT(s)/2 - a(j, t_k)$. If we interpret $OPT(s)/2$ as the *refracted distance* between s_k and t_k, then these values are a lower bound on distances, and measure how 'close' these nodes can be to s_k given the shortest distance $OPT(s)/2$.

Again, as we show later, the dual constraints with $\eta_j(f)$ and α_j^k, correspond to the primal node variable $u_j(f)$, and ensure that the minimum cost $\eta_j(f)$ of reaching node j from s_1 and s_2, is at most equal to $a(s_1, j) + a(s_2, j)$. The dual constraints with $\eta_j(f)$ and β_j^k, correspond to the primal node variable $v_j(f)$, and ensure that the minimum cost $\eta_j(f)$ is at least equal to $OPT(s) - a(j, t_1) - a(j, t_2)$, the sum of the lower bounds on the node potentials.

Finally, the dual $\mathbf{DCTC(R)}$ of reformulation $\mathbf{CTC(R)}$, including objective function and the constraints for the variables z_0, z_s, z_1, z_2 is

$$\text{Max } \nu = \sum_{k=1}^{2} \mu_k(\gamma_{t_k}^k - \gamma_{s_k}^k) + \Theta$$

subject to

$$\begin{array}{rcl}
\gamma_j^k - \gamma_i^k & \leq & w_{ij} \\
\Theta + \gamma_{s_1}^1 + \gamma_{s_2}^2 - \gamma_{t_1}^1 - \gamma_{t_2}^2 & \leq & 0 \\
\Theta + 2\alpha_{s_1}^1(1) + 2\beta_{t_1}^1(1) - \eta_{s_2}^2(1) - \eta_{t_2}^2(1) + \gamma_{t_1}^1 - \gamma_{s_1}^1 & \leq & 0 \\
\Theta + 2\alpha_{s_2}^2(2) + 2\beta_{t_2}^2(2) - \eta_{s_1}^1(2) - \eta_{t_1}^1(2) + \gamma_{t_2}^2 - \gamma_{s_2}^2 & \leq & 0 \\
\Theta + \alpha_{s_1}^1 + \alpha_{s_2}^2 - \beta_{t_1}^1 - \beta_{t_2}^2 & \leq & 0 \\
\alpha_j^1(1), \beta_j^1(1), \eta_j^2(1) & \in & DS^1(*) \\
\alpha_j^2(2), \beta_j^2(2), \eta_j^1(2) & \in & DS^2(*) \\
\alpha_j^k, \beta_j^k, \eta_j(f), \eta_j(b) & \in & DSC \\
\text{All variables} & \geq & 0.
\end{array}$$

The first constraint is the shortest path type constraint corresponding to primal variable f_{ij}^k which send μ_k units from s_k to t_k. The dual variable γ_j^k equals $a(s_k, j)$. Thus, the dual objective function equals $\mu_1 a(s_1, t_1) + \mu_2 a(s_2, t_2) + \Theta$. If we recall that $OPT(1) = \pi_{s_2}^1 + \pi_{t_2}^1$ and $OPT(2) = \pi_{s_1}^2 + \pi_{t_1}^2$, the next constraints ensure that

$$\begin{array}{rcl}
\Theta & \leq & a(s_1, t_1) + a(s_2, t_2) \\
\Theta & \leq & OPT(1) - a(s_1, t_1) \\
\Theta & \leq & OPT(2) - a(s_2, t_2) \text{ and} \\
\Theta & \leq & OPT(s).
\end{array}$$

Clearly, this ensures that the extra cost Θ is at most equal to the cost of two independent paths, the cost of a shared solution, or the extra cost of a star solution.

4.2 Integer Optimal Solutions

We now come to the main result of this Section, where we show that the linear program **CTC(R)** is an exact formulation, i.e., it has an integer optimal solution.

Theorem 3 *The linear programming formulation* **CTC(R)** *always has an integer optimal solution.*

Proof
We prove the result by describing a dual feasible solution for **DCTC(R)** that has the same objective function value OPT as the primal solution. Let $\gamma_j^k = a(s_k, j)$.

Using the $\pi_j(\delta)$ values from algorithm two path in the shared problem, let
$\alpha_j^k = a(s_k, j)$,
$\beta_j^k = OPT(s)/2 - a(j, t_k)$,
$\eta_j(f) = \pi_j(f)$, and
$\eta_j(b) = \pi_j(b) - OPT(s)/2$.

Using the π_j^1 values from algorithm two path in the first star solution, let
$\alpha_j^1(1) = a(s_1, j)$, $\beta_j^1(1) = a(j, t_1)$, $\eta_j^2(1) = \pi_j^1$.
Using the π_j^2 values from algorithm two path in the second star solution, let
$\alpha_j^2(2) = a(s_2, j)$, $\beta_j^2(2) = a(j, t_2)$, $\eta_j^1(2) = \pi_j^2$.
Finally, let

$$\Theta = \min \{a(s_1, t_1)+a(s_2, t_2), OPT(s), OPT(1)-a(s_1, t_1), OPT(2)-a(s_2, t_2)\}.$$

Hence the dual objective function equals the minimum of the costs of the independent, shared, first star or second star solution. Therefore, if the dual values are feasible, we are done.

First consider the shared solution. From algorithm two path notice that $\pi_j(f) \leq a(s_1, j) + a(s_2, j)$ and that $\pi_j(b) \leq a(s_1, j) + a(j, t_2)$. Since $\pi_j(f)$ is the minimum cost of sending one unit of flow from each of the nodes s_1 and s_2 to node j, and since $a(j, t_1) + a(j, t_2)$ is an upper bound on the cost of sending one unit from node j to each of nodes t_1 and t_2, it follows that

$$\pi_j(f) + a(j, t_1) + a(j, t_2) \geq OPT(s).$$

Since $\pi_j(b)$ is the minimum cost of sending one unit of flow from node s_1 to node j, and one unit from j to t_2, and since $a(j, t_1) + a(s_2, j)$ is an upper bound on the cost of sending one unit from node j to node t_1 and from s_2 to j, it follows that

$$\pi_j(b) + a(j, t_1) + a(s_2, j) \geq OPT(s).$$

It is now easy to verify that these values of the dual variables satisfy dual feasibility for the shared problem. Similar arguments establish that dual constraints are satisfied for the two star problems.

5 Related Problems

Consder one and two commodity problems, capacitated and uncapacitated, with or without fixed and flow costs. Fixed costs can be shared between two commodities. whereas flow costs cannot. It is important to distinguish between the network loading type of capacitated problems considered here where capacity can be purchased in batches, and fixed capacity problems where at most some fixed quantity C of capacity can be added. Unless specified otherwise, we consider problems on undirected graphs.

Consider single commodity problems. The uncapacitated version of this problem with or without fixed costs, with or without flow costs is easy to solve. The shortest path from source to sink provides the optimal solution. The network loading problem with fixed costs and no flow costs, or with no fixed costs and only flow costs can be solved using a shortest path algorithm. However, if there are both fixed and flow costs, Chopra et.al.[2] showed that the problem is NP-hard. They also showed that if there are two types of *facilities* where capacity can either be added in units of C or 1, the problem is NP-hard even without flow costs.

Consider uncapacitated two commodity problems. If there are no fixed costs. but only flow costs, the problem can be solved by finding the shortest path from each source to the corresponding sink. If there are fixed costs, but no flow costs, Sastry [6] showed that the problem has an optimal solution with at most one shared path, and can be solved in $O(n^2)$ time. and described an extended linear programming formulation that guarantees integer optimal solutions. If there are both fixed and flow costs. it is not known whether the problem is in the class P or not.

Some related problems are the so called *two path* problems, where we wish to find the shortest paths from two sources to two sinks, and paths are allowed to share arcs. If the problem is defined on an undirected graph, it reduces to the uncapacitated two commodity network design problem with no flow costs. If it is on a directed graph, then the best known combinatorial algorithm solves the problem in $O(n^4)$ time (M. Natu and S.Fang [4]). This problem is equivalent to the uncapacitated two commodity network design problem on a directed graph.

Consider the network loading type of capacitated two commodity problems.

If there are no fixed costs, the problem can be solved by finding the shortest path from each source to the corresponding sink. If there are fixed costs, but no flow costs, then we have shown in this paper that the problem can be solved in $O(n^2)$ time, and have described an exact linear programming formulation that guarantees integer optimal solutions. If there are both fixed and flow costs, then clearly, the problem is NP-hard, since the single commodity version is known to be so.

The following table summarises the results known so far. Dashes in the table indicate that the particular parameter is irrelevant. The table shows the complexity of the best best known combinatorial algorithm if the problem can be solved in polynomial time. All costs are assumed to be non negative.

Problem	Capacitated	Fixed Cost	Flow Cost	Complexity
Single Commodity	No	-	-	$O(n^2)$
	Yes	Yes	No	$O(n^2)$
	Yes	No	Yes	$O(n^2)$
	Yes	Yes	Yes	NP-hard
	Yes (two facility)	Yes	No	NP-hard
Two Commodity	No	Yes	No	$O(n^2)$ Digraph: $O(n^4)$
	No	No	Yes	$O(n^2)$
	No	Yes	Yes	Not known
	Yes	Yes	No	$O(n^2)$
	Yes	No	Yes	$O(n^2)$.
	Yes	Yes	Yes	NP-hard

6 Conclusions

We have characterized optimal solutions and shown that the capacitated two commodity network design problem with fixed costs and no flow cost can be solved in polynomial time. Either all flows are on independent paths, or $\mu_k C$ units of each commodity are sent on shortest paths with the remainders sharing paths (which we call the shared solution), or $(\mu_k - 1)C$ units of one commodity and $\mu_l C$ of the other commodity flow on shortest paths, with the remaining flows forming the kth star solution. The algorithm repeatedly uses

Djikstra's algorithm on different auxiliary graphs derived from the original one. We also describe an exact linear programming formulation that guarantees integer optimal solutions, using $O(m)$ variables and $O(n)$ constraints. We also discuss several other variations of the single and two commodity problems, where the presence of both fixed and flow costs makes capacitated problems hard.

Acknowledgements

The author is indebted to Yogesh Agarwal and Mechthild Opperud (nee Stoer) for several useful discussions.

7 References

[1] A.Balakrishnan. T.L.Magnanti and R.T.Wong. A Dual Ascent Procedure for Large-Scale Uncapacitated Network Design. *Operations Research* **37** (1989) 716-740.

[2] S.Chopra, I.Gilboa and S.T.Sastry. Source Sink Flows with Capacity Installation in Batches. *Discrete App. Math.* **85** (1998) 165-192.

[3] T.C.Hu. Multi-commodity Network Flows *Operations Research* **11** (1963) 344-360.

[4] M.Natu and S.Fang. On the point to point connection problem. *Info.Proc. Letters* **53** (1995) 333-336.

[5] M.Sakarovitch. Two Commodity Network Flows and Linear Programming. *Math.Prog.* **4** (1973) 1-20.

[6] S.T.Sastry. A Characterization of the Two Commodity Network Design Problem. To appear in *Networks.*

[7] P.D.Seymour. Four-Terminus Flows. *Networks* **10** (1980) 79-86.

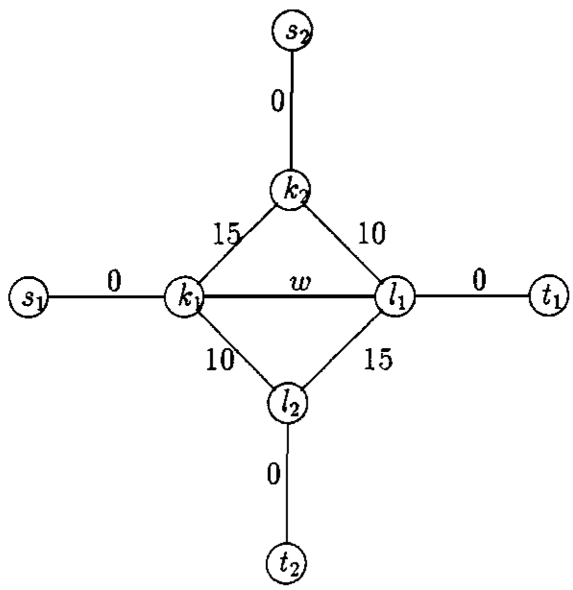
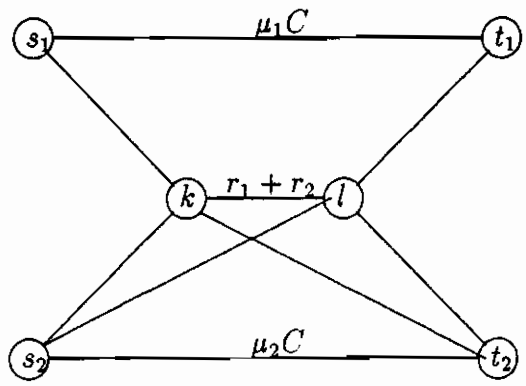


FIGURE 1



SHARED SOLUTION

FIGURE 2