Math. Program., Ser. B 92: 425–437 (2002)

Digital Object Identifier (DOI) 10.1007/s101070100284

Alper Atamtürk*

On capacitated network design cut–set polyhedra

Received: January 2, 2001 / Accepted: September 13, 2001
Published online February 14, 2002 – © Springer-Verlag 2002

Abstract. This paper provides an analysis of capacitated network design cut–set polyhedra. We give a complete linear description of the cut–set polyhedron of the single commodity – single facility capacitated network design problem. Then we extend the analysis to single commodity – multifacility and multicommodity – multifacility capacitated network design problems. Valid inequalities described here are applicable to directed network design problems with any number of facility types and any level of capacities. We report results from a computational study done for testing the effectiveness of the new inequalities.

Key words. network design – multicommodity flows – facets

1. Introduction

Given a network and a set of demands on the vertices of the network, the *capacitated network design problem* is to install integer multiples of capacities on the arcs of the network and route the flow so that the sum of capacity installation and flow routing costs is minimized. For instance, installing or leasing fiberoptic cables with certain bandwidths on a communication network in order to meet communication requirements of a number of customers, determining production and warehouse facility capacities and supply channels in a production–distribution logistics network, or equipping a set of trains on a railroad network with a number of engines with certain motive capacities can be viewed as installing capacities on the arcs of a network and routing the flow of commodities on them.

The ever increasing demand for high speed communication networks has been a major motivation for research on modeling and solving capacitated network design problems in the last decade. One modeling distinction between telecommunication network design and logistics network design problems is that in the former, capacity installed between two nodes allows traffic in both directions, thus capacity is undirected; whereas in the latter, as capacity represents availability of a transportation resource from one geographical location to another or availability of a production or storage facility from one time period until a future time period, capacity is directed. Here we consider directed capacity installation.

We say that a problem is a *single facility* network design problem if we install only a single type of facility on the arcs of the network. Routing vehicles with identical

A. Atamtürk: Department of Industrial Engineering and Operations Research, University of California, Berkeley, CA 94720-1777, USA, e-mail: atamturk@ieor.berkeley.edu

* Supported, in part, by NSF grants DMI–9908705 and DMI–0070127.

capacity in a logistics network and installing a communication network with only one cable type are examples of the single facility network design problem. In the *multifacility* problem we are allowed to install different types of facilities with varying capacities, such as fiberoptic cables with varying bandwidths, production lines or machines with different rates, or a fleet of heterogeneous vehicles with varying capacities. Depending on the application, the routed flow may be a single commodity or multiple commodities. If the flow under consideration is indistinguishable, then we say that it is a single commodity. A production–distribution problem for a single product for one time period may be viewed as a *single commodity* network design problem, whereas a multiperiod version of the same problem with release dates and due dates associated with origin and destination pairs is a *multicommodity* network design problem. The single commodity – single facility version of the problem with a single source and a single sink is known to be already $\mathcal{NP}$–hard [6].

Earlier work

The study of the polyhedral structure of capacitated network design problems has been initiated by Magnanti and Mirchandani [14] and Magnanti et al. [16]. They consider a single commodity multifacility network design problem, where the facility capacities are integer multiples of some basic capacity unit and describe the *cut–set inequalities* that have nonzero coefficients for facility capacity variables. Bienstock and Günlük [5] generalize these inequalities to flow–cut–set inequalities that have nonzero coefficients also for the flow variables for multicommodity network design problems with two types of facilities, where the capacity of the second type is an integer multiple of the capacity of the first type. Pochet and Wolsey [18] study the polyhedron of a single–arc network design problem with an arbitrary number of facilities with divisible integer capacities. Chopra et al. [6] give inequalities and extended formulations for a single commodity directed problem with two facility types. Bienstock et al. [4] consider a multicommodity single facility network design problem, whereas Günlük [10] studies a multicommodity two facility re–design problem with existing capacities.

Our contributions & outline

The objective of this paper is to study the cut–set polyhedra of single commodity – single facility, single commodity – multifacility, and multicommodity – multifacility network design problems. These polyhedra capture a significant number of the characteristics of the related network design problems and therefore it is important to understand their facial structure when developing branch–and–cut algorithms for network design problems.

In Sect. 2 we give a complete linear description of the single commodity – single facility cut–set polyhedron. Later in Sects. 3 and 4 we extend the development to single commodity – multifacility and multicommodity – multifacility network design problems, respectively. The linear description result in Sect. 2 is not only interesting by itself, but also steps of its proof are used in proving validity of the multifacility inequalities in Sect. 3.

All of the prior studies on the polyhedral structure of the multifacility capacitated network design problem assume that facility capacities satisfy a *divisibility property*, that is, the second type capacity is an integer multiple of the first type capacity, the third type capacity is an integer multiple of the second type capacity, and so on. Even though such an assumption is valid for many telecommunication applications, this is not always the case, for instance, when hub and link facilities that operate with different types of technologies are considered simultaneously. Especially in logistics network design applications, such as fleet or locomotive scheduling, production–distribution problems, vehicle capacities or machine/facility capacities do not satisfy the divisibility property, see [12,7]. Multifacility inequalities developed under divisibility assumption cannot be applied in these contexts.

One important aspect of this study is that we do not make any assumptions on either the number of facilities or the structure of capacities for the multifacility problems. Therefore, the results presented here are applicable to network design problems with an arbitrary number of facilities and arbitrary capacities. To the best of our knowledge, this is the first polyhedral study on the problem that completely drops all restrictive assumptions on capacities. We show that many classes of facet–defining cut–set inequalities introduced separately in the literature for versions of multifacility network design problems are, in fact, part of one common class, whose coefficients can be expressed in a closed form with a subadditive function. Also as we consider directed capacities, the inequalities given here have coefficients for both inflow and outflow arcs of a cut of the network. In Sect. 5 we present results from a computational study that demonstrates the effectiveness of the new inequalities in solving multicommodity – multifacility directed network design problems with arbitrary capacities.

2. Single commodity – single facility problems

In this section we consider the design problem of a directed network with a single commodity where we are allowed to install batches of a single facility type with capacity c on the arcs of the network. Let (U, V) be a nonempty partitioning of the vertices of a directed network. Let A^+ be the set of arcs directed from U to V, A^- be the arcs directed from V to U, $A = A^+ \cup A^-$, and b be the supply of U for V. For any vector v and a subset of indices S, let $v(S) = \sum_{a \in S} v_a$. Then the constraints of the network design problem with respect to the cut A can be aggregated as

$$f(A^+) - f(A^-) = b, \tag{1}$$

$$0 \leq f_a \leq c x_a, \ \forall a \in A, \tag{2}$$

where f_a denotes the flow on arc a, $a \in A$ and x_a denotes the number of facilities to install on the same arc. We assume that $b,\ c > 0$ (w.l.o.g.) and are rational numbers. Consequently, let us define the single commodity – single facility network design cut–set polyhedron as

$$\mathcal{F}_{SS} \equiv conv\{(x, f) \in \mathbb{Z}^{|A|} \times \mathbb{R}^{|A|} : \ (x, f) \text{ satisfies (1) and (2)}\}.$$

We assume that $A^+ \neq \emptyset$, because otherwise $\mathcal{F}_{SS} = \emptyset$ as $b > 0$ and $f_a \geq 0$. It is easy to check that the dimension of $\mathcal{F}_{SS}$ is $2|A| - 1$.

Optimizing a linear function over $\mathcal{F}_{SS}$ is easy: We say that flow on an arc a is fractional if $kc < f_a < (k+1)c$ for $k \in \mathbb{Z}$, $f_a - kc$ being the fraction. Let $r = b - \lfloor b/c \rfloor c$, so that $0 \le r < c$. An extreme point of $\mathcal{F}_{SS}$ has at most one arc with fractional flow f_a and the fraction is r if $a \in A^+$ and $c - r$ if $a \in A^-$. Consequently, if the minimization problem over $\mathcal{F}_{SS}$ with an objective vector $(d, e) \in \mathbb{R}^{2|A|}$ is bounded, then the optimal value is the smaller of $\lfloor b/c \rfloor \min_{a \in A^+}\{ce_a + d_a\} + \min_{a \in A^+}\{re_a + d_a\}$ and $\lceil b/c \rceil \min_{a \in A^+}\{ce_a + d_a\} + \min_{a \in A^-}\{(c - r)e_a + d_a\}$. So if there is an optimal solution, there exists one where no flow on an arc in A^- is a positive integer multiple of c. A corresponding optimal solution is clear from the description of the optimal value. In the light of polynomial equivalence of optimization and separation for polyhedra [8], separation of $\mathcal{F}_{SS}$ must also be easy.

The most general valid inequality for $\mathcal{F}_{SS}$ is given by Chopra et al. [6]. For $S^+ \subseteq A^+$, $S^- \subseteq A^-$, and $\eta = \lceil b/c \rceil$, they define the *cut–set inequality* as

$$rx(S^+) + f(A^+ \setminus S^+) + (c - r)x(S^-) - f(S^-) \ge r\eta. \tag{3}$$

Simpler forms of this inequality $x(A^+) \ge \eta$ and $rx(S^+) + f(A^+ \setminus S^+) \ge r\eta$ are given in Magnanti and Mirchandani [14] and Bienstock and Günlük [5], respectively. Since the single commodity – single facility network design problem is the simplest of the capacitated network design problems, it is important to fully understand the facial structure of $\mathcal{F}_{SS}$ before considering multicommodity – multifacility generalizations. Therefore, we are interested in knowing the conditions under which inequality (3) is strong and whether there are other classes of facets of $\mathcal{F}_{SS}$.

In the rest of this section we will answer these questions. We give the necessary and sufficient conditions for (3) to be facet–defining for $\mathcal{F}_{SS}$ and show that equality (1) and inequalities (2) and (3) are indeed sufficient to describe $\mathcal{F}_{SS}$. For completeness we begin with a simple proof of validity of the cut–set inequalities.

Proposition 1. *The cut–set inequality* (3) *is valid for* $\mathcal{F}_{SS}$.

Proof. The result holds trivially if $r = 0$. Otherwise, if $x(S^+) - x(S^-) \ge \eta$, then

$$r(x(S^+) - x(S^-) - \eta) + f(A^+ \setminus S^+) + cx(S^-) - f(S^-) \ge 0,$$

else, adding and subtracting $c(x(S^+) - \eta)$ and using $b = c(\eta - 1) + r$ (as $r > 0$)

$$(r - c)(x(S^+) - x(S^-) - \eta + 1) + cx(S^+) + f(A^+ \setminus S^+) - f(S^-) - b \ge 0.$$

□

Theorem 1. *The cut–set inequality* (3) *is facet–defining for* $\mathcal{F}_{SS}$ *if and only if* $r > 0$ *and* $S^+ \neq \emptyset$.

Proof. Necessity. If $r = 0$, then inequality (3) reduces to $f(A^+ \setminus S^+) + cx(S^-) - f(S^-) \ge 0$, which is dominated by the sum of the nonnegativity constraints of the flow variables on A^+ and the capacity constraints of the flow variables on S^-. Else if $S^+ = \emptyset$, then the inequality reduces to $f(A^+) - f(S^-) + (c - r)x(S^-) + (\eta - 1)(c - r) \ge b$, which is dominated by $f(A^+) - f(A^-) \ge b$ as $c > r$ and $\eta \ge 1$.

Sufficiency. Let g_i and h_i denote the ith unit vectors for the continuous variables and for the integer variables, respectively. Let $\sum_{i\in A}\beta_i x_i + \sum_{i\in A}\pi_i f_i = \beta_o$ define an arbitrary hyperplane and $s \in S^+$. Since all points of $\mathcal{F}_{SS}$ satisfy $f(A^+) - f(A^-) = b$, we may add multiples of the balance equality to facet–defining inequalities without changing them; therefore without loss of generality we assume that $\pi_s = 0$. Consider the points $u^o = \eta h_s + b g_s$, and $u^i = u^o - h_s - r g_s + h_i + r g_i$, $v^i = u^i - \epsilon g_s + \epsilon g_i$ $i \in S^+$ with $0 < \epsilon < c - r$ in $\mathcal{F}_{SS}$. Suppose these points are on the hyperplane defined above. Comparing u^i and v^i, we see that $\pi_i = 0$ for all $i \in S^+$ and from u^o and u^i we have $\beta_i = \beta$ for all $i \in S^+$. From points $u^o + h_i$, we also see that $\beta_i = 0$ for all $i \in (A^+ \setminus S^+) \cup (A^- \setminus S^-)$. Similarly $u^o + \epsilon g_s + h_i + \epsilon g_i$ give $\pi_i = 0$ for all $i \in A^- \setminus S^-$. u^o and $u^o - h_s - c g_s + h_i + r g_i$ give $\beta = r\pi_i$ for all $i \in A^+ \setminus S^+$. For the rest of the coefficients, consider the points $y^i = u^o + (c-r)g_s + (c-r)g_i + h_i$ and $z^i = y^i + h_s + r g_s + r g_i$ for $i \in S^-$. Then by comparing y^i with z^i, $\beta = -r\pi_i$ for all $i \in S^-$; and from u^o and y^i we have $\beta_i = (c-r)\beta/r$ for all $i \in S^-$. Finally, plugging in these coefficients for u^o, $\beta_o = \beta\eta$. Dividing the coefficients by β/r, we arrive at (3), concluding affine independence of the described $2|A| - 1$ points on the face of $\mathcal{F}_{SS}$ induced by (3). Since $u^o + h_s \in \mathcal{F}_{SS}$ is not on the face, the face is proper.

□

Theorem 2. *$\mathcal{F}_{SS}$ is described completely by equality* (1) *and inequalities* (2) *and* (3).

Proof. Let $O(d, e)$ denote the index set of extreme optimal solutions to the optimization problem $\min\{dx + ef : (x, f) \in \mathcal{F}_{SS}\}$, where $(d, e) \in \mathbb{R}^{2|A|}$ is an arbitrary objective vector, not perpendicular to the smallest affine subspace containing $\mathcal{F}_{SS}$, i.e., $(d, e) = (d, e_{A^+}, e_{A^-}) \neq \lambda(0, 1, -1)$, $\lambda \in \mathbb{R}_+$, and therefore the set of optimal solutions is not $\mathcal{F}_{SS}$ ($\mathcal{F}_{SS} \neq \emptyset$). We will prove the theorem by exhibiting an inequality among (2) and (3) that is satisfied at equality by (x^k, f^k) for all $k \in O(d, e)$. Then, since (d, e) is arbitrary, for every facet of $\mathcal{F}_{SS}$ there is an inequality among (2) and (3) that defines it. In each step below, we either restrict our attention to a subcase without loss of generality or complete the proof.

- (i) Let $e^+_{min} = \min_{a\in A^+} e_a$. $e^+_{min} < \infty$ as $A^+ \neq \emptyset$ by assumption. We add $e^+_{min} b$ to the objective and subtract $e^+_{min} f(A^+) - e^+_{min} f(A^-)$ from it. This operation does not change the set of optimal solutions. So we may restate the assumptions on (d, e) as $e_a \geq 0$ for all $a \in A^+$, $e_a = 0$ for some $a \in A^+$, and $(d, e) \neq 0$.
- (ii) We may also assume that $d_a \geq 0$ for all $a \in A$, since otherwise the problem is unbounded, and that $e_a \leq 0$ for all $a \in A^-$, since otherwise $f^k_a = 0$ holds for all k $\in O(d, e)$. Furthermore if $d_a + c e_a > 0$ for $a \in A^-$, then $x^k_a = 0$ for all $k \in O(d, e)$. Therefore we assume that $d_a + c e_a \leq 0$ for all $a \in A^-$.
- (iii) Suppose $r = 0$. Then if $d_a > 0$ for some $a \in A$, $f^k_a = c x^k_a$ holds for all $k \in O(d, e)$; however if $d_a = 0$ for all $a \in A$, then either $e_a > 0$ for some $a \in A^+$, which implies $f_a = 0$ for $O(d, e)$, since by (i) $\exists m \in A^+$ with $e_m = 0$, or $e_a < 0$ for some $a \in A^-$, in which case the problem is unbounded for the same reason. Therefore in the sequel we assume that $r > 0$.
- (iv) If $d_a = e_a = 0$ for some $a \in A^+$, then $f^k_m = 0$ for $m \in A^+ \setminus \{a\}$ with $d_m > 0$ or $e_m > 0$ for all $k \in O(d, e)$ and we are done. If there is no such $m \in A^+ \setminus \{a\}$, then $d_m = e_m = 0$ for all $m \in A^+$. Now if $d_m + c e_m < 0$ for some $m \in A^-$,

then the problem is unbounded; otherwise from (ii) $d_m + ce_m = 0$ for all $m \in A^-$. Then by (ii) $d_m > 0$ and $e_m < 0$ for some $m \in A^-$, because otherwise $(d, e) = 0$, which contradicts (i). But since $r < c$ and $d_m + ce_m = 0$, it follows that $d_m + re_m > 0$, which implies that $f_m^k = cx_m^k$ for all $k \in O(d, e)$. So we may assume that $d_a + ce_a > 0$ for all $a \in A^+$.

(v) Let $S^+ = \text{Argmin}_{a \in A^+}\{d_a + ce_a : d_a > 0\}$. Since by (i) $\exists m \in A^+$ with $e_m = 0$ and by (iv) $d_m > 0$, it holds $S^+ \neq \emptyset$. Let $\bar{s} = d_a + ce_a$ for $a \in S^+$. Suppose $\exists a \in A^+ \setminus S^+$ such that $\bar{s} \geq ce_a$ and $d_a = 0$. Then $d_m \geq \bar{s} \geq ce_a > re_a$, which implies that $f_m^k = cx_m^k$ for $k \in O(d, e)$ and we are done. So we may assume that $\bar{s} < d_a + ce_a$ for all $a \in A^+ \setminus S^+$.

(vi) $\bar{s} + d_a + ce_a \geq 0$ for all $a \in A^-$ since otherwise the problem is unbounded. Let $S^- = \{a \in A^- : \bar{s} + d_a + ce_a = 0 \text{ and } d_a > 0\}$. Suppose $\bar{s} + ce_a = 0$ for some $a \in A^- \setminus S^-$ with $d_a = 0$. Then, since $d_m > 0$ for $m \in S^+$, $e_m + e_a < 0$ must hold, which implies that $f_m^k = cx_m^k$ for all $k \in O(d, e)$. Therefore, we may assume that $\bar{s} + d_a + ce_a > 0$ for all $a \in A^- \setminus S^-$.

(vii) Steps (v) and (vi) imply that in an extreme optimal solution arcs in $A^+ \setminus S^+$ and $A^- \setminus S^-$ cannot have flow that is a positive integer multiple of c, since otherwise we can improve the objective by moving the flow on such arcs to either S^+ or S^-. Then in an extreme optimal solution $f(A^+ \setminus S^+)$ equals either 0 or r; similarly in such a solution $f(A^- \setminus S^-)$ equals either 0 or $c - r$. Since the balance equality must be satisfied, this implies $(\eta - 1)c \leq f^k(S^+) - f^k(S^-) \leq \eta c$ for all $k \in O(d, e)$. Furthermore because $d_a > 0$, we have $x_a^k = \lceil f_a^k / c \rceil$ for all $a \in S^+ \cup S^-$ and all $k \in O(d, e)$. Then $\eta - 1 \leq x^k(S^+) - x^k(S^-) \leq \eta$ is satisfied for all $k \in O(d, e)$. Under this condition inequality (3) holds at equality. Proof is complete.

□

Corollary 1. *If* $r = 0$, *then equality* (1) *and inequalities* (2) *describe* $\mathcal{F}_{SS}$.

Separation

For a fixed cut A, solving the separation problem of the cut–set inequalities is trivial: given $(\bar{x}, \bar{f})$, if $r\bar{x}_a < \bar{f}_a$ for $a \in A^+$, then we include a in S^+, if $(c - r)\bar{x}_a < \bar{f}_a$ for $a \in A^-$, then we include a in S^-. However, finding the best cut set A is not as easy. For the single source single sink case, the problem of finding the best cut set can be posed as an s–t maxcut problem.

3. Single commodity – multifacility problems

Next we consider network design problems where we are allowed to install facilities of multiple types with different capacities on the arcs of the network in batches. Let c^t be the capacity of facility of type t, $t \in T$. No assumption is made on either the number of facility types or the structure of capacities (other than $c^t > 0$ and rational). In order to formulate the problem, we replace the capacity constraint (2) with

$$0 \leq f_a \leq \sum_{t \in T} c^t x_a^t, \ \forall a \in A. \tag{4}$$

Consequently, the related single commodity – multifacility cut–set polyhedron is

$$\mathcal{F}_{SM} \equiv conv\{(x, f) \in \mathbb{Z}^{|A||T|} \times \mathbb{R}^{|A|} : (x, f) \text{ satisfies (1) and (4)}\}.$$

Optimization over $\mathcal{F}_{SM}$ is $\mathcal{NP}$–hard, it becomes a knapsack problem in the special case of $|A^+| = 1$ and $A^- = \emptyset$. Therefore, due to the negative result of Karp and Papadimitriou [11], we may not expect to give an explicit linear description of $\mathcal{F}_{SM}$ as we did for $\mathcal{F}_{SS}$ unless $\mathcal{NP}$= *co–*$\mathcal{NP}$. Nevertheless, it is possible to identify nontrivial classes of facet–defining inequalities for $\mathcal{F}_{SM}$ that may be useful in branch–and–cut computations.

Before proceeding any further, we note that since we do not assume any structure on the capacities, if arc a has an existing capacity c_a, it can be included in T as a pseudo–facility. Once a valid inequality is found for this relaxation, the pseudo–facility variables can be projected to 1, to handle the capacity expansion version of the network design problem studied in [5]. Therefore, we do not address this extension here explicitly.

The following proposition specifies a class of facets of $\mathcal{F}_{SM}$ for which the coefficients of the facility variables with the same capacity are equal. This result allows us to treat such variables as a single variable and simplifies the proof of Theorem 3.

Proposition 2. *Let $\sum_{t \in T} \beta^t x^t + \pi f \geq \pi_o$ be a facet–defining inequality for $\mathcal{F}_{SM}$, different from the nonnegativity constraints. If $\pi_a = \pi_b$ for $a, b \in A^+$ or $a, b \in A^-$, then $\beta^t_a = \beta^t_b$ for all $t \in T$.*

Proof. Suppose that the claim is not true and without loss of generality $\beta^t_a > \beta^t_b$ for some $t \in T$. Since $\sum_{s \in T} \beta^s x^s + \pi f \geq \pi_o$ is facet–defining and is not $x^t_a \geq 0$, there exists $(x, f) \in \mathcal{F}_{SM}$ such that $\sum_{s \in T} \beta^s x^s + \pi f = \pi_o$ and $x^t_a > 0$. Let $(\bar{x}, \bar{f})$ be defined as $\bar{f}_a = 0, \bar{x}^t_a = 0, \bar{f}_b = f_b + f_a, \bar{x}^t_b = x^t_b + x^t_a$, and $\bar{f}_i = f_i$ for all $i \in A \setminus \{a, b\}$, $\bar{x}^s_i = x^s_i$ for all $i \in A \setminus \{a, b\}, s \in T$, $\bar{x}^s_a = x^s_a$, $\bar{x}^s_b = x^s_b$ for all $s \in T \setminus \{t\}$. Since $a, b \in A^+$ or $a, b \in A^-$, $(\bar{x}, \bar{y}) \in \mathcal{F}_{SM}$. However, $\sum_{s \in T} \beta^s \bar{x}^s + \pi \bar{f} < \pi_o$, which contradicts validity of the inequality.

□

Let $\eta^s = \lceil b/c^s \rceil$ and $r^s = b - \lfloor b/c^s \rfloor c^s$. Also for $k = \lfloor c/c^s \rfloor$, let $\phi^+_s(c) = \min\{c - k(c^s - r^s),\ (k+1)r^s\}$ and $\phi^-_s(c) = \min\{c - kr^s,\ (k+1)(c^s - r^s)\}$. Then for $s \in T$, we define the *multifacility cut–set inequality* as

$$\sum_{t \in T} \phi^+_s(c^t) x^t(S^+) + f(A^+ \setminus S^+) + \sum_{t \in T} \phi^-_s(c^t) x^t(S^-) - f(S^-) \geq r^s \eta^s. \qquad (5)$$

Theorem 3. *The multifacility cut–set inequality* (5) *is valid for $\mathcal{F}_{SM}$. It is facet–defining for $\mathcal{F}_{SM}$ if $r^s > 0$, $S^+ \neq \emptyset$, $A^+ \setminus S^+ \neq \emptyset$, and $S^- \neq \emptyset$.*

Proof. We use a cut–set inequality for the single facility restriction of $\mathcal{F}_{SM}$ for facility $s \in T$ as the basis for constructing (5). Inequality

$$r^s x^s(S^+) + f(A^+ \setminus S^+) + (c^s - r^s) x^s(S^-) - f(S^-) \geq r^s \eta^s$$

is facet–defining for $\{(x, f) \in \mathcal{F}_{SM} : x^t(A) = 0, t \in T \setminus \{s\}\}$ since $r^s > 0$ and $S^+ \neq \emptyset$. We lift this inequality with the variables that are fixed to zero to obtain a facet–defining inequality for $\mathcal{F}_{SM}$. Notice that due to Proposition 2 we can treat $\{x^t_a\}_{a\in S}$ as a single variable for $S = S^+$, $S = A^+ \setminus S^+$, $S = S^-$, and $S = A^- \setminus S^-$. Then the exact lifting coefficient ϕ^+_t of $x^t(S)$ for some $t \in T \setminus \{s\}$ is the optimal value of the following nonlinear mixed–integer optimization problem

$$\phi^+_t = \max\left\{\frac{r^s(\eta^s - x^s(S^+)) - f(A^+ \setminus S^+) - (c^s - r^s)x^s(S^-) + f(S^-)}{x^t(S)}\right\}$$
$$\text{s.t.: } (x, f) \in \mathcal{F}_{SM},\ x^v(A) = 0, v \in T \setminus \{s, t\},\ x^t(A \setminus S) = 0,\ x^t(S) \geq 1,$$

which follows from Wolsey [19]. The lifting coefficients for $x^t(A^+\setminus S^+)$ and $x^t(A^-\setminus S^-)$, $t \in T$ are zero since increasing capacities of these arcs has no effect on the lifting problem. So we first let $S = S^+$ and evaluate $z^+(c^t x^t(S^+))$, the maximum value of the numerator as a function of $x^t(S^+)$. Let us initially evaluate $z^+(c^t)$, thus $x^t(S^+) = 1$. Observe that the numerator can be increased by either decreasing $x^s(S^+)$ and $f(A^+\setminus S^+)$ to their lower bounds or increasing $x^s(S^-)$ and $f(S^-)$ as needed subject to the balance constraint and the variable upper bounds. Since both of these have the same effect on the value of the numerator we assume the latter. To be precise, without loss of generality, assume that $x^s(S^+) = \eta^s$, $f(S^+) = \eta^s c^s$ and $f(A^+ \setminus S^+) = 0$. Then the numerator is maximized by maximizing $f(S^-) - (c^s - r^s)x^s(S^-)$ subject to $f(S^-) \leq c^s - r^s + c^t$ and $f(S^-) \leq c^t x^t(S^-)$. Let $kc + \bar{r} = c^s - r^s + c^t$ with $k \in \{0, 1, 2, \dots\}$ and $0 \leq \bar{r} \leq c^t$. An optimal solution is then given by $f(S^-) = c^s - r^s + c^t$ and $x^t(S^-) = k + 1$ if $c^s - r^s < \bar{r} \leq c^s$, and by $f(S^-) = kc^s$ and $x^t(S^-) = k$ if $0 < \bar{r} \leq c^s - r^s$. Therefore, $z^+(c^t) = \min\{c^t - k(c^s - r^s),\ (k + 1)r^s\}$, where $k = \lfloor c^t/c^s \rfloor$. Notice that z^+ is a subadditive function of c^t, i.e., $z^+(a^1) + z^+(a^2) \geq z^+(a^1 + a^2)$. Consequently, $\phi^+_t = \max_{i\geq 1} z^+(ic^t)/i = z^+(c^t)$. Moreover, the lifting is sequence independent, i.e., ϕ^+_t does not change with the order in which $x^t(S^+)$ is introduced to the inequality [20, 9, 1]. So by lifting the basic cut–set inequality with $x^v(S^+),\ v \in T$ in any order, we obtain

$$\sum_{v\in T} \phi^+(c^v)x^v(S^+) + f(A^+ \setminus S^+) + (c^s - r^s)x^s(S^-) - f(S^-) \geq r^s\eta^s. \tag{6}$$

Next we lift inequality (6) with $x^t(S^-)$, $t \in T \setminus \{s\}$. Then, the coefficient of $x^t(S^-)$ is the optimal value of the following optimization problem

$$\phi^-_t = \max\left\{\frac{r^s\eta^s - \sum_{v\in T} \phi^+(c^v)x^v(S^+) - f(A^+\setminus S^+) - (c^s - r^s)x^s(S^-) + f(S^-)}{x^t(S^-)}\right\}$$
$$\text{s.t.: } (x, f) \in \mathcal{F}_{SM},\ x^v(S^-) = 0, v \in T \setminus \{s, t\},\ \text{and } x^t(S^-) \geq 1.$$

We claim that this problem has an optimal solution such that $x^v(S^+) = 0$ for all $v \in T \setminus \{s\}$. Suppose that the claim is not true and the minimum value $x^v(S^+)$ takes in any optimal solution is a positive integer p. If $kc^s + r^s \leq c^v < (k + 1)c^s$ for some $k \in \mathbb{Z}_+$, then $\phi^+(c^v) = (k + 1)r^s$ and therefore we obtain a solution with objective at least as high by setting $x^v(S^+) = 0$ and increasing $x^s(S^+)$ by $(k + 1)p$. Otherwise, $kc^s \leq c^v < kc^s + r^s$ and $\phi^+(c^v) = c^v - k(c^s - r^s)$. But in this case we obtain a solution with the same objective by setting $x^v(S^+) = 0$ and increasing $x^s(S^+)$ by kp, $f(A^+ \setminus S^+)$ by $p(c^v - kc^s)$. Contradiction.

Now we will characterize $z^-(c^t x^t(S^-))$, the maximum value of the numerator as a function of $x^t(S^-)$. First suppose that $x^t(S^-) = 1$. Without loss of generality, we may assume that $x^s(A^+ \setminus S^+) = x^s(A^-) = 0$. After substituting $b + f(S^-)$ for $f(S^+)$, the numerator is maximized by maximizing $f(S^-) - r^s x^s(S^+)$ subject to $b + f(S^-) \leq c^s x^s(S^+)$ and $f(S^-) \leq c^t$. Let $kc^s + \bar{r} = b + c^t$ with $k \in \mathbb{Z}_+$ and $0 \leq \bar{r} < c^s$. This maximization problem has an optimal solution given by $x^t(S^+) = k+1$ and $f(S^-) = c^t$ if $r \leq \bar{r} < c^s - r^s$, and by $x^t(S^+) = k$ and $f(S^-) = kc - b$ if $0 \leq \bar{r} < r$. Therefore, $z^-(c^t) = \min\{c^t - kr^s,\ (k+1)(c^s - r^s)\}$, where $k = \lfloor c^t/c^s \rfloor$. Since z^- is subadditive as well, it follows that $\phi_t^- = z^-(c^t)$. Hence, the inequality is as desired.

□

Remark 1. No inflow. We remark that if the multifacility cut–set inequality (5) has coefficients only for the outflow arcs, i.e., when $S^- = \emptyset$, then its coefficients can be strengthened as $\tilde{\phi}_s^+(c) = \min\{\phi_s^+(c),\ \eta^s r^s\}$. Notice that $\tilde{\phi}_s^+(c) < \phi_s^+(c)$ for $c > \eta^s c^s$. This improvement in the coefficients of facilities with capacity larger than $\eta^s c^s$ is justified by observing that the numerator of the objective function in the first part of the proof of Theorem 3 is upper bounded by the lower bounds on $x^s(S^+)$ and $f(A^+ \setminus S^+)$ when $S^- = \emptyset$.

Remark 2. Varying capacities. Multifacility cut–set inequalities (5) are valid also for network design models with arc–varying capacities

$$f(A^+) - f(A^-) = b, \tag{7}$$

$$f_a \leq c_a\, x_a \text{ for } a \in A, \tag{8}$$

which can be seen by simply aggregating (8) over subsets of A to obtain multifacility relaxations of the form $f(S) \leq \sum_{a \in S} c_a\ x_a$ for $S = S^+, S^-, A^+ \setminus S^+, A^- \setminus S^-$. If necessary, by introducing fictitious facility variables, we get a multifacility network design relaxation with four arcs for (7)–(8). Klabjan and Nemhauser [13] study a special case of this model with $A^- = \emptyset$ in detail.

Separation

Given a cut A for each facility $s \in T$, the multifacility cut–set inequalities (5) is an exponential class by the choice of the subsets of arcs S^+ and S^-. However finding a subset that gives a most violated inequality for a point $(\bar{x}, \bar{f})$ is straightforward. If $\sum_{t \in T} \phi_s^+(c^t)\bar{x}_a^t < \bar{f}_a$ for $a \in A^+$, then we include a in S^+, and if $\sum_{t \in T} \phi_s^-(c^t)\bar{x}_a^t < \bar{f}_a$ for $a \in A^-$, then we include a in S^-. Since $\phi_s^+(c)$ or $\phi_s^-(c)$ can be calculated in constant time, for a fixed cut A a violated multifacility cut–set inequality is found in $O(|A||T|)$ if there exists any.

4. Multicommodity – multifacility problems

Finally we extend the inequalities presented in Sect. 3 for multicommodity – multifacility network design problems. Consider a nonempty partitioning (U, V) of the vertices of the network. Let b^k denote the supply of commodity k in U for V and f_a^k be flow of commodity k on arc a for $k \in K$. Recall that A^+ is the set of arcs directed from U to V, A^- is the set of arcs directed from V to U, and $A = A^+ \cup A^-$. Then the constraints of the multicommodity – multifacility problem across cut A can be aggregated as

$$f^k(A^+) - f^k(A^-) = b^k, \qquad \forall k \in K, \tag{9}$$

$$0 \leq \sum_{k \in K} f_a^k \leq \sum_{t \in T} c^t x_a^t, \ \forall a \in A. \tag{10}$$

So the corresponding multicommodity – multifacility cut–set polyhedron is

$$\mathcal{F}_{MM} \equiv conv\{(x, f) \in \mathbb{Z}^{|A||T|} \times \mathbb{R}^{|A||K|} : (x, f) \text{ satisfies (9) and (10)}\}.$$

The inequalities presented in Sect. 3 are extended for $\mathcal{F}_{MM}$ by considering single commodity relaxations of $\mathcal{F}_{MM}$ obtained by aggregating flow variables and balance equations (9) over subsets of K. For $Q \subseteq K$ let $f_Q(S) = \sum_{k \in Q} f^k(S)$ and $b_Q = \sum_{k \in Q} b^k$. From Theorem 3 it is seen that *multicommodity - multifacility cut–set inequality*

$$\sum_{t \in T} \phi^+_{s,Q}(c^t) x^t(S^+) + f_Q(A^+ \setminus S^+) + \sum_{t \in T} \phi^-_{s,Q}(c^t) x^t(S^-) - f_Q(S^-) \geq r^s_Q \eta^s_Q \tag{11}$$

is valid $\mathcal{F}_{MM}$ with $r^s_Q = b_Q - \lfloor b_Q / c^s \rfloor c^s$, $\eta^s_Q = \lceil b_Q / c^s \rceil$, and $\phi^+_{s,Q}$ and $\phi^-_{s,Q}$ defined in the same way as ϕ^+_s and ϕ^-_s using r^s_Q and η^s_Q instead of r^s and η^s.

Theorem 4. *The multicommodity – multifacility cut–set inequality* (11) *is valid for* $\mathcal{F}_{MM}$*. It is facet–defining for* $\mathcal{F}_{MM}$ *if* $(S^+, A^+ \setminus S^+)$ *and* $(S^-, A^- \setminus S^-)$ *are nonempty partitions,* $r^s_Q > 0$*, and* $b^k > 0$ *for all* $k \in Q$*.*

If $S^- = \emptyset$, then the coefficients of inequality (11) can also be strengthened as described in Remark 1.

Separation

For a fixed cut of the network, the complexity of separating multicommodity cut–set inequalities is an open question even for a single facility problem. Optimization of a linear function over $\mathcal{F}_{MS}$ is $\mathcal{NP}$–hard as the facility location problem is a special case of it. For a multicommodity single facility network design problem of a single arc, cut–set inequalities (11) reduce to the residual capacity inequalities [15], for which an exact linear–time separation method is given in Atamtürk and Rajan [2]. From here it follows that, for a single facility problem, if S^+ and S^- are fixed, then we can find a subset of commodities $Q \subseteq K$ that gives a most violated inequality in linear time. Alternatively, if Q is fixed, since the model reduces to a single commodity, we can find the subsets $S^+ \subseteq A^+$ and $S^- \subseteq A^-$ that give a most violated inequality in linear time as well. However, the complexity of determining Q, S^+ and S^- simultaneously is an open question.

Example 1. We specialize inequality (11) for the network design problem with two types of facilities considered in Bienstock and Günlük [5]. Let the vectors x^1 and x^2 denote the facilities with capacities $c^1 = 1$ and $c^2 = \lambda > 1$ with $\lambda \in \mathbb{Z}$, respectively. Let Q be a nonempty subset of the commodities. Then by letting $s = 1$, we have $r_Q^1 = b_Q - \lfloor b_Q \rfloor$ and inequality (11) becomes

$$r_Q^1 x^1(S^+) + \big(r_Q^1 \lfloor\lambda\rfloor + \min\{\lambda - \lfloor\lambda\rfloor, r_Q^1\}\big)x^2(S^+) + f_Q(A^+ \setminus S^+) + \big(1-r_Q^1\big)x^1\big(S^-\big)+\big(\big(1-r_Q^1\big)\lfloor\lambda\rfloor+\min\{\lambda-\lfloor\lambda\rfloor, 1-r_Q^1\}\big)x^2(S^-)-f_Q(S^-) \geq r_Q^1\lceil b_Q\rceil,$$

which, when λ is integer, reduces to

$$r_Q^1 x^1(S^+) + \lambda r_Q^1 x^2(S^+) + f_Q(A^+ \setminus S^+) + \big(1 - r_Q^1\big)x^1(S^-) + \lambda\big(1 - r_Q^1\big)x^2(S^-) - f_Q(S^-) \geq r_Q^1\lceil b_Q\rceil. \quad (12)$$

Notice that inequality (12) is not valid for $\mathcal{F}_{MM}$ unless $\lambda \in \mathbb{Z}$. Also by letting $s = 2$, we have $r_Q^2 = b_Q - \lfloor b_Q/\lambda\rfloor\lambda$. So the corresponding multicommodity two facility cut–set inequality is

$$\min\{1, r_Q^2\}x^1(S^+) + r_Q^2 x^2(S^+) + f_Q(A^+ \setminus S^+) + \min\{1, \lambda - r_Q^2\}x^1(S^-) + \big(\lambda - r_Q^2\big)x^2(S^-) - f_Q(S^-) \geq r_Q^2\lceil b_Q/\lambda\rceil.$$

5. Computational results

In this section we present results from a computational study with the cut–set inequalities introduced in previous sections. In order to test the effectiveness of the inequalities, we implemented a branch–and–cut algorithm to solve multxicommodity – multifacility network design problems with general capacities. The cut generation module is added to CPLEX[1] branch–and–bound algorithm using callback functions of its callable library. In order to find feasible solutions to the problems early in the tree, we also incorporated a simple heuristic, which rounds up the fractional capacity variables in the LP relaxation solutions. All computations presented here were performed on a SUN Ultra 10 workstation with a time limit of 10 hours and a memory limit of 256 MB.

For generating cut–set inequalities, an important implementation question is how to choose the cut (A^+, A^-) of the network from which inequalities are generated. Since finding a most violated cut–set inequality over all cuts of the network is $\mathcal{NP}$–hard even for the special case $S^+ = A^+$ and $S^- = \emptyset$ for the single facility case [3], for simplicity of implementation, we limited the choice of cut sets to those of one and two–node subsets of the network. Thus, the number of cut sets considered for generating cut–set inequalities is at most quadratic in the number of nodes of the network. Given a cut of the network, since we do not have an efficient way of choosing a commodity subset $Q \subseteq K$ that gives a most violated inequality, we considered only subsets that are singleton commodities and the set of all commodities whose supply and demand nodes are separated by the current cut.

[1] CPLEX is a trademark of ILOG, Inc.

For the experiments we used the data set of Parker and Ryan [17], generated in conjunction with US West to test bandwidth packing algorithms for video conferencing networks (publicly available at `ftp://eng.auburn.edu/pub/pvance/data/mcf`). This set consists of 14 problems ranging in size from 19 to 29 nodes and 23 to 93 commodities. Demand for commodities range between 5 and 60. In order to convert them to multifacility network design problems, for each arc we introduced three capacity variables with capacities equal to $2^4 - 1$, $2^6 - 1$, and $2^8 - 1$, so that capacities do not satisfy the divisibility property. We solved these instances first with default CPLEX and then with generating multicommodity – multifacility cut–set inequalities (11) at each node of the search tree. The rounding heuristic is run in both cases.

A summary of the computational results is presented in Table 1. In the second column of this table we report the number of cut–set inequalities (`cuts`) added to the formulations. In the next 3 columns of Table 1 we report the objective values of LP relaxation at the root node before branching (`zroot`), the objective values of the best integer solution found (`zub`), the number of nodes explored in the search tree (`nodes`) without cut–set inequalities and with cut–set inequalities under headings (1) and (2), respectively. In the last column of Table 1, we report the CPU time elapsed in seconds if the problem is solved within time and memory limits. Otherwise, we report the percentage gap (`endgap`) between `zub` and the best lower bound in the search tree at termination in parenthesis. The letter next to the endgap indicates whether the algorithm is terminated due to memory limit (M) or time limit (T).

Table 1. Computational results with cut–set inequalities

prob	cuts	zroot		zub		nodes		time/(endgap)	
		(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
1	363	5976.94	6904.49	7235	7235	861	97	3	3
2	386	11330.32	12455.66	12640	12640	210	12	2	3
3	3706	8047.65	14052.10	21710	19442	78275	1249	(53.3)M	(26.2)T
4	2675	4876.68	8057.67	10600	9055	187485	9010	(32.2)M	(3.1)T
5	1551	3744.33	6911.48	9355	8105	240596	15857	(26.5)M	6182
6	1417	9049.53	13848.16	17822	15240	114702	100296	(35.1)M	(4.3)T
7	966	12462.94	14994.64	15426	15426	437520	966	12651	487
8	793	6862.23	10307.06	13890	12535	245389	33474	(20.3)M	1248
9	3979	8638.45	13007.36	21453	16326	96670	1688	(51.4)M	(17.3)T
10	151	7864.63	8454.83	8985	8985	3928	1821	7	7
11	598	2680.54	3963.21	4356	4356	4367	295	11	16
12	1306	14641.09	18452.35	23526	20590	128042	176473	(26.2)M	(1.0)T
13	776	10169.40	11823.91	13385	13385	1403873	94870	15434	2801
14	756	6129.42	8435.79	9566	9566	385848	28975	3022	707

(1) without cut–set inequalities, (2) with cut–set inequalities.

Both methods solved 4 out of the 14 instances very quickly. Runs without cut–set inequalities reached the memory limit for 7 of the remaining instances, whereas runs with the cut–set inequalities reached the time limit for 5 of these remaining instances. A comparison of the columns of Table 1 for the three harder instances that could be solved with or without the cuts, shows that number of nodes and solution times decreased dramatically when cut–set inequalities were added in the search tree. The cut–set inequalities consistently improved the LP relaxation bounds. For the instances that could not be solved to optimality, the values of the feasible solutions found and the endgaps are significantly better for the runs with the cut–set inequalities.

We should also mention that the runs without cut–set inequalities resulted very large search trees with little bound improvement on the optimal values. When cut–set inequalities were added, more time was spent in improving the lower bounds at each node of the search tree, but as a result, smaller number of nodes was explored and overall much less memory was required.

It is encouraging to see that adding cut–set inequalities even from a restricted choice of cut sets improved the computations significantly. These experiments clearly show the usefulness of the cut–set inequalities for multicommodity – multifacility network design problems with arbitrary capacities.

Acknowledgement. I am grateful to Oktay Günlük for a careful review of an earlier version of this paper.

References

1. Atamtürk, A. (2001): Sequence independent lifting for mixed–integer programming. Research Report BCOL.01.02, University of California, Berkeley, May 2001. Available at `http://ieor.berkeley.edu/~atamturk`
2. Atamtürk, A., Rajan, D. (2000): On splittable and unsplittable flow capacitated network design arc–set polyhedra. Research Report BCOL.00.02, University of California, Berkeley, July 2000. Available at `http://ieor.berkeley.edu/~atamturk`
3. Bienstock, D. (2001): Personal communication
4. Bienstock, D., Chopra, S., Günlük, O., Tsai, C.-Y. (1998): Minimum cost capacity installation for multicommodity networks. Math. Program. **81**, 177–199
5. Bienstock, D., Günlük, O. (1996): Capacitated network design – Polyhedral structure and computation. INFORMS J. Comput. **8**, 243–259
6. Chopra, S., Gilboa, I., Sastry, S.T. (1998): Source sink flows with capacity installation in batches. Discrete Appl. Math. **85**, 165–192
7. Florian, M., Bushell, G., Ferland, J., Guérin, G., Nastansky, L. (1976): The engine scheduling problems in a railway network. INFOR **14**, 121–138
8. Grötschel, M., Lovász, L., Schrijver, A. (1981): The ellipsoid method and its consequences in combinatorial optimization. Combinatorica **1**, 169–197
9. Gu, Z., Nemhauser, G.L., Savelsbergh, M.W.P. (2000): Sequence independent lifting in mixed integer programming. J. Comb. Optim. **4**, 109–129
10. Günlük, O. (1999): A branch-and-cut algorithm for capacitated network design problems. Math. Program. **86**, 17–39
11. Karp, R.M., Papadimitriou, C.H. (1980): On the linear characterizations of combinatorial optimization problems. In: Proceedings of the 21st Annual Symposium on Foundations of Computer Science, pp. 1–9
12. Kim, D., Barnhart, C., Ware, K., Reinhardt, G. (1999): Multimodal express package delivery: A service network design application. Transport. Sci. **33**, 391–407
13. Klabjan, D., Nemhauser, G.L. (2001): A polyhedral study of integer variable upper bounds. Draft
14. Magnanti, T.L., Mirchandani, P. (1993): Shortest paths, single origin-destination network design, and associated polyhedra. Networks **23**, 103–121
15. Magnanti, T.L., Mirchandani, P., Vachani, R. (1993): The convex hull of two core capacitated network design problems. Math. Program. **60**, 233–250
16. Magnanti, T.L., Mirchandani, P., Vachani, R. (1995): Modeling and solving the two-facility capacitated network loading problem. Oper. Res. **43**, 142–157
17. Parker, M., Ryan, J.A. (1994): A column generation algorithm for bandwitdh packing. Telecom. Systems **2**, 185–195
18. Pochet, Y., Wolsey, L.A. (1995): Integer knapsack and flow covers with divisible coefficients: Polyhedra, optimization, and separation. Discrete Appl. Math. **59**, 57–74
19. Wolsey, L.A. (1976): Facets and strong valid inequalities for integer programs. Oper. Res. **24**, 367–372
20. Wolsey, L.A. (1977): Valid inequalities and superadditivity for 0/1 integer programs. Math. Oper. Res. **2**, 66–77