SCHOOL OF OPERATIONS RESEARCH
AND INDUSTRIAL ENGINEERING
COLLEGE OF ENGINEERING
CORNELL UNIVERSITY
ITHACA, NY 14853-3801

TECHNICAL REPORT NO. 1101

August 1994

**CAPACITATED FACILITY
LOCATION: SEPARATION
ALGORITHMS AND
COMPUTATIONAL EXPERIENCE**

by

Karen Aardal¹

¹Department of Econometrics, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands. Research was supported in part by the National Science Foundation grant CCR-9307391.

Capacitated Facility Location: Separation Algorithms and Computational Experience

Karen Aardal
Department of Econometrics
Tilburg University
P.O. Box 90153
5000 LE Tilburg, The Netherlands
aardal@kub.nl

August 3, 1994

)

Abstract

We consider the polyhedral approach to solving the capacitated facility location problem. First we develop a heuristic for the separation problem based on the family of submodular inequalities. An important part of the heuristic builds on a constructive proof that two specific conditions are necessary for a large subclass of the submodular inequalities to be facet defining. The proofs show how structures that violate the two conditions can be modified to produce stronger inequalities.

One well-known subclass of the submodular inequalities is the family of flow cover inequalities. If all capacities are equal we show that for a fixed client set, the separation problem based on the flow cover inequalities can be solved in polynomial time.

The family of combinatorial inequalities were originally developed for the uncapacitated facility location, but are also valid for the capacitated problem. No computational experience with using the combinatorial inequalities has been reported so far. Here we suggest how partial output from the heuristic identifying violated submodular inequalities can be used as input to a heuristic identifying violated combinatorial inequalities.

We report on computational results from solving 60 small and medium size problems.

We study the cutting plane approach to solving the capacitated facility location (CFL) problem. The polyhedral structure of CFL has been studied by Aardal (1992) and Aardal, Pochet and Wolsey (1993), by Leung and Magnanti (1989) for the equal capacity case, and by Deng and Simchi-Levi (1993) for the case of equal capacities and unsplit demands. Cornuéjols, Sridharan and Thizy (1991) compared the strength of various Lagrangean relaxations of CFL, and relaxations obtained by completely dropping different sets of the defining constraints. They also did an extensive computational study using these relaxations and several heuristics proposed in the literature. However, very little experience in using strong valid inequalities to solving CFL has been reported on so far, except for some preliminary work by Aardal (1992).

We first introduce necessary notation, and a formulation of CFL that is non-standard as it contains redundant aggregate variables and constraints. The purpose of adding aggregate information is to provide the building blocks for two relaxations of CFL. Facet defining inequalities for these relaxations are well-known, and will be generated automatically, given our formulation, by the mixed-integer optimizer used in the computational study.

In Section 2 we describe the family of submodular inequalities for CFL as introduced by Aardal et al., and state conditions for special cases of the submodular inequalities, namely flow cover, effective capacity (EC), and single-depot inequalities, to be facet defining.

In Section 3 we first show that the separation problem based on the flow cover inequalities for a fixed client set can be solved in polynomial time if all capacities are equal. Next we suggest a general heuristic for finding appropriate subsets of depots and clients, that are then used as input to specialized heuristics for finding violated flow cover, EC and single-depot inequalities. An important part of the heuristics for identifying violated EC and single-depot inequalities builds on a proof that certain conditions are necessary for the EC inequalities to be facet defining. The proof, different from the proof in Aardal et al., is constructive and provides ways of strengthening inequalities that violate the specific necessary conditions. We also suggest how the subsets generated by the general heuristic can be used to identify violated combinatorial inequalities. These inequalities were introduced by Cho et al. (1983a) and generalize all valid inequalities developed so far for the uncapacitated facility location (UFL) problem.

The separation heuristics have been implemented as subroutines of MINTO (Savelsbergh et al. (1991)) and are used together with the system constraints of MINTO to solve CFL. In Section 4 we report on computational experience from solving 60 small and medium size test problems. The computational results indicate that the knapsack cover and flow cover inequalities, which are automatically generated by MINTO from our formulation of CFL, are important in closing substantial parts of the duality gap. However, in most cases more problem specific inequalities are needed in order to further close the gap. The heuristics we propose seem effective in generating violated EC and single-depot inequalities, and we are able to solve thirteen problems to optimality at the root node of the branch and bound tree, by generating these inequalities in addition to the MINTO system inequalities.

1 Problem Formulation

Let $M = \{1, \dots, m\}$ be the set of facilities (depots) and $N = \{1, \dots, n\}$ the set of clients. Client k has demand d_k and the capacity of depot j is m_j . The total demand of the clients in the subset $S \subseteq N$ is denoted by $d(S)$. The fixed cost of opening facility j is f_j and the per unit cost of flow between depot j and client k is c_{jk} .

Let $y_j = 1$ if depot j is open and $y_j = 0$ otherwise. The flow from depot j to client k on the arc (j, k) is denoted by v_{jk} . We want to determine which depots should be opened and how the flow should be distributed in order to satisfy all clients demand, as well as the capacity restrictions, at minimum cost. Consider the following formulation of CFL:

$$\min\left\{\sum_{j \in M} \sum_{k \in N} c_{jk} v_{jk} + \sum_{j \in M} f_j y_j : (v, y) \in X^{CFL}\right\}$$

where

$$X^{CFL} = \{(v, y) \in \mathbb{R}^{m \times n} \times \mathbb{Z}^m : \sum_{j \in M} v_{jk} = d_k \quad k \in N\} \quad (1)$$

$$v_j = \sum_{k \in N} v_{jk} \quad j \in M \quad (2)$$

$$v_j \leq m_j y_j \quad j \in M \quad (3)$$

$$0 \leq v_{jk} \leq d_k y_j \quad j \in M, k \in N \quad (4)$$

$$\sum_{j \in M} v_j = d(N) \quad (5)$$

$$y_j \leq 1 \quad j \in M\}. \quad (6)$$

By combining appropriate aggregate constraints with constraints (6), we obtain the surrogate knapsack polytope X^K , and the single-node flow polytope X^{SNF} ,

$$X^K = \{y \in \mathbb{Z}_+^m : \sum_{j \in M} m_j y_j \geq d(N), y_j \leq 1, j \in M\}, \quad (7)$$

$$X^{SNF} = \{(v, y) \in \mathbb{R}_+^m \times \mathbb{Z}_+^m : \sum_{j \in M} v_j = d(N), v_j \leq m_j y_j, y_j \leq 1, j \in M\}. \quad (8)$$

The knapsack cover inequalities and flow cover inequalities developed for X^K and X^{SNF} respectively, see Wolsey (1975) and Padberg et al. (1985), are shown to be facet defining for $\text{conv}(X^{CFL})$ by Aardal (1992). Separation heuristics based on these families of inequalities are incorporated in the system routines of the mixed-integer optimizer MINTO, and can therefore be applied directly to formulation (1)–(6). Hence, formulation (1)–(6) is useful for computational purposes, even if it contains redundant information.

2 The Submodular Inequalities and Important Special Cases

Submodular inequalities were first introduced by Wolsey (1989) for general fixed-charge network flow problems, and adapted to CFL by Aardal (1992) and Aardal et al. (1993). Let $K \subseteq N$ be a subset of clients, $J \subseteq M$ a subset of depots, and choose for each $j \in J$ a subset $K_j \subseteq K$. The *effective capacity* of depot j is defined as $\bar{m}_j = \min(m_j, d(K_j))$.

Proposition 1 (Aardal (1992).) *The function*

$$\begin{aligned}
 f(J) = \max \{ & \sum_{j \in J} \sum_{k \in K_j} v_{jk} : \\
 & \sum_{k \in K_j} v_{jk} \leq \bar{m}_j y_j \quad j \in J \\
 & \sum_{\{j \in J : K_j \ni k\}} v_{jk} \leq d_k \quad k \in K \\
 & v_{jk} \geq 0 \quad j \in J, k \in K \\
 & y_j = 1 \quad j \in J \}
 \end{aligned} \tag{9}$$

is submodular on M .

The value of $f(J)$ is the maximum flow from the depots in J to the clients in K given the arc set $\{(j, k) : j \in J, k \in K_j\}$. The difference in maximum flow when all depots are open, and when all depots except depot j are open, is called the *increment function* and is defined as $\rho_j(J \setminus \{j\}) = f(J) - f(J \setminus \{j\})$.

Proposition 2 (Aardal (1992).) *Let $J \subseteq M$, $K \subseteq N$ and $K_j \subseteq K$ for each $j \in J$. The submodular inequality*

$$\sum_{j \in J} \sum_{k \in K_j} v_{jk} + \sum_{j \in J} \rho_j(J \setminus \{j\})(1 - y_j) \leq f(J) \tag{10}$$

is valid for $\text{conv}(X^{CFL})$.

Definition 1 $J \subseteq M$ is a *cover* with respect to M and N if $\sum_{j \in J} m_j = d(N) + \lambda$ with $\lambda > 0$.

For the submodular inequalities to be valid we do not require the set J to be a cover. However, in all special cases that are studied here, we assume that J is a cover.

The first special case we shall consider is the well-known family of *flow cover* inequalities developed by Padberg et al. (1985) for X^{SNF} . Let $(x)^+ = \max(0, x)$ and let $v'_j = \sum_{k \in K} v_{jk}$, $j \in M$.

Lemma 3 *Let $K \subseteq N$. If J is a cover with respect to M and K , and if $K_j = K$ for all $j \in J$, then $f(J) = d(K)$ and $\rho_j(J \setminus \{j\}) = (m_j - \lambda)^+$, where $\lambda = \sum_{j \in J} m_j - d(K) > 0$.*

In the flow cover case, the submodular inequalities (10) take the following form:

$$\sum_{j \in J} v'_j + \sum_{j \in J} (m_j - \lambda)^+(1 - y_j) \leq d(K). \tag{11}$$

Given our formulation of CFL, MINTO will automatically recognize the single-node flow structure (8) and generate flow cover inequalities for $K = N$.

In the general case, where $K_j \subset K$ for at least one $j \in J$, there exists no closed-form expression for $\rho_j(J \setminus \{j\})$, which makes it difficult to characterize facet defining submodular inequalities. By using maximum flow arguments, it is straightforward to show that $\rho_j(J \setminus \{j\}) \geq (\bar{m}_j - \lambda)^+$ for all $j \in J$. Aardal et al. (1993) completely characterized the subclass of facet defining submodular inequalities having $\rho_j(J \setminus \{j\}) = (\bar{m}_j - \lambda)^+$, which are referred to as *effective capacity* (EC) inequalities, and two subclasses of submodular facets with $\rho_j(J \setminus \{j\}) > (\bar{m}_j - \lambda)^+$ for at least one $j \in J$, called single- and multi-depot inequalities. We begin by describing the EC inequalities.

Theorem 4 (EC inequalities, Aardal et al. (1993).) *Let $J \subset M$ be a cover with respect to M and K , and let $Q \subset J$ be the subset of depots for which $\bar{m}_q < m_q$. Assume that $\sum_{j \in M} m_j > d(N) + m_r$ for all $r \in J$. The EC inequality*

$$\sum_{j \in J} \sum_{k \in K_j} v_{jk} + (\bar{m}_j - \lambda)^+(1 - y_j) \leq d(K) \quad (12)$$

defines a facet of $\text{conv}(X^{CFL})$ if and only if

- (1) *for each pair of depots q_1, q_2 , $K_{q_1} \cap K_{q_2} = \emptyset$,*
- (2) *$K_j = K$ for all $j \in J \setminus Q$,*
- (3) *$(\cup_{q \in Q} K_q) \subset K$,*
- (4) *$\bar{m}_q > \lambda$ for all $q \in Q$,*
- (5) *if $|Q| \leq 1$, then $\exists j \in J$ with $\bar{m}_j = m_j > \lambda$.*

Remark: If $Q = \emptyset$ in Theorem 4, we obtain the facet defining *flow cover* inequalities (11).

From the point of view of separation, it is natural to ask what to do in case we generate a structure that violates one or more of conditions (1)-(5) of Theorem 4. Below we give a proof, different from the proof by Aardal et al., that conditions (1) and (4) are necessary for the EC inequalities (12) to be facet defining, by showing that an inequality that violates these conditions can be strengthened. The strengthening is constructive, and forms an important part of the separation heuristic presented in Section 3.2.

Proof of necessity of Condition (1).

We will show that if all clients belonging to any intersection of sets K_q are removed, and the simple inequality $\sum_{j \in M} \sum_{k \in K^\cap} v_{jk} \leq d(K^\cap)$ is added, where $K^\cap$ is the set of removed clients, then a stronger inequality is obtained. Let

$$\begin{aligned} K^\cap &= \bigcup_{q \in Q} (\cup_{j \in Q \setminus \{q\}} (K_j \cap K_q)), \\ Q^\cap &= \{j \in Q : K_j \cap K^\cap \neq \emptyset\}, \\ K' &= K \setminus K^\cap, \\ K'_j &= K_j \setminus K^\cap, \quad j \in J. \end{aligned}$$

The new values of the effective capacities are:

$$\begin{aligned} \bar{m}'_j &= \min(m_j, d(K')), \quad j \in J \setminus Q, \\ \bar{m}'_j &= \bar{m}_j, \quad j \in Q \setminus Q^\cap, \\ \bar{m}'_j &= \bar{m}_j - \sum_{k \in K_j \cap K^\cap} d_k, \quad j \in Q^\cap, \end{aligned}$$

giving a new value of the excess capacity, $\lambda' = \sum_{j \in J} \bar{m}'_j - d(K')$.

Assume that $\sum_{j \in J \setminus Q} m_j > d(K \setminus (\cup_{j \in Q} K_j))$, in which case $\lambda' > 0$. If $\sum_{j \in J \setminus Q} m_j \leq d(K \setminus (\cup_{j \in Q} K_j))$, no tight point with $y_j = 1$ for all $j \in \{j \in J : \bar{m}_j > \lambda\}$ exists.

We now show that the inequality

$$\sum_{j \in J} \sum_{k \in K'_j} v_{jk} + \sum_{j \in M} \sum_{k \in K^\cap} v_{jk} + \sum_{j \in J} (\bar{m}'_j - \lambda')^+ (1 - y_j) \leq d(K') + d(K^\cap) \quad (13)$$

is stronger than the EC-inequality based on the sets J , K and K_j .

Since $\lambda' > 0$, inequality (13) is valid. The right-hand sides of both inequality (13) and the EC inequality based on J , K and K_j are equal, since $d(K') + d(K^\cap) = d(K)$. All coefficients of the variables v_{jk} are equal in the two inequalities except the coefficients of v_{jk} , $j \in Q$, $k \in K^\cap \setminus K_j$ and $j \in M \setminus J$, $k \in K^\cap$, which have value one in inequality (13) and zero in the EC inequality. Hence, what remains to show is that $(\bar{m}'_j - \lambda')^+ \geq (\bar{m}_j - \lambda)^+$ for all $j \in J$.

Case 1: $j \in J \setminus Q$

$$\begin{aligned} \bar{m}'_j - \lambda' &= \min(m_j, d(K')) - \sum_{l \in J} \bar{m}'_l + d(K') = \min(m_j, d(K')) - \\ &\sum_{l \in J \setminus Q} \min(m_l, d(K')) - \sum_{l \in Q \setminus Q^\cap} \bar{m}_l - \sum_{l \in Q^\cap} (\bar{m}_l - \sum_{k \in K_l \cap K^\cap} d_k) + d(K) - d(K^\cap) = \\ &= - \sum_{l \in (J \setminus Q) \setminus \{j\}} \min(m_l, d(K')) - \sum_{l \in Q \setminus Q^\cap} \bar{m}_l - \sum_{l \in Q^\cap} (\bar{m}_l - \sum_{k \in K_l \cap K^\cap} d_k) + \sum_{l \in J} \bar{m}_l - \lambda - d(K^\cap) = \\ &= - \sum_{l \in (J \setminus Q) \setminus \{j\}} \min(m_l, d(K')) + \sum_{l \in J \setminus Q} \bar{m}_l + \sum_{l \in Q^\cap} \sum_{k \in K_l \cap K^\cap} d_k - \lambda - d(K^\cap) = \\ &\bar{m}_j - \lambda + \sum_{l \in (J \setminus Q) \setminus \{j\}} (\bar{m}_l - \min(m_l, d(K'))) + \sum_{l \in Q^\cap} \sum_{k \in K_l \cap K^\cap} d_k - d(K^\cap) \geq \bar{m}_j - \lambda, \end{aligned}$$

where the inequality follows from the second term being nonnegative, and from $\sum_{l \in Q^\cap} \sum_{k \in K_l \cap K^\cap} d_k \geq d(K^\cap)$.

Case 2: $j \in Q \setminus Q^\cap$

$$\begin{aligned} \bar{m}'_j - \lambda' &= \bar{m}_j - \sum_{l \in J} \bar{m}'_l + d(K') = \\ &\bar{m}_j - \sum_{l \in J \setminus Q} \min(m_l, d(K')) - \sum_{l \in Q \setminus Q^\cap} \bar{m}_l - \sum_{l \in Q^\cap} (\bar{m}_l - \sum_{k \in K_l \cap K^\cap} d_k) + d(K) - d(K^\cap) = \\ &\bar{m}_j - \sum_{l \in J \setminus Q} \min(m_l, d(K')) - \sum_{l \in Q \setminus Q^\cap} \bar{m}_l - \sum_{l \in Q^\cap} (\bar{m}_l - \sum_{k \in K_l \cap K^\cap} d_k) + \sum_{l \in J} \bar{m}_l - \\ &\lambda - d(K^\cap) = \\ &\bar{m}_j - \lambda + \sum_{l \in J \setminus Q} (\bar{m}_l - \min(m_l, d(K'))) + \sum_{l \in Q^\cap} \sum_{k \in K_l \cap K^\cap} d_k - d(K^\cap) \geq \bar{m}_j - \lambda. \end{aligned}$$

Case 3: $j \in Q^\cap$

$$\begin{aligned} \bar{m}'_j - \lambda' &= \bar{m}_j - \sum_{k \in K_j \cap K^\cap} d_k - \sum_{l \in J} \bar{m}'_l + d(K') = \\ &\bar{m}_j - \sum_{k \in K_j \cap K^\cap} d_k - \sum_{l \in J \setminus Q} \min(m_l, d(K')) - \sum_{l \in Q \setminus Q^\cap} \bar{m}_l - \sum_{l \in Q^\cap} (\bar{m}_l - \sum_{k \in K_l \cap K^\cap} d_k) + d(K) - \\ &d(K^\cap) = \\ &\bar{m}_j - \sum_{k \in K_j \cap K^\cap} d_k - \sum_{l \in J \setminus Q} \min(m_l, d(K')) - \sum_{l \in Q \setminus Q^\cap} \bar{m}_l - \sum_{l \in Q^\cap} (\bar{m}_l - \sum_{k \in K_l \cap K^\cap} d_k) + \sum_{l \in J} \bar{m}_l - \\ &\lambda - d(K^\cap) = \\ &\bar{m}_j - \lambda + \sum_{l \in J \setminus Q} (\bar{m}_l - \min(m_l, d(K'))) + \sum_{j \in Q \setminus \{j\}} \sum_{k \in K_l \cap K^\cap} d_k - d(K^\cap) \geq \bar{m}_j - \lambda. \quad \blacksquare \end{aligned}$$

Corollary 5 *In the special case where only two client sets K_{q_1} and K_{q_2} , $q_1, q_2 \in Q$ intersect, and where Conditions (2)-(5) of Theorem 4 hold, the EC inequality based on the client set K' defines a facet.*

Proof of necessity of Condition (4).

Let $Q^- \subseteq Q$ be the set of depots in Q with $\bar{m}_j \leq \lambda$, and assume that Condition (1) of Theorem 4 holds. Here we will show that by removing all depots in Q^- and their respective client sets K_j , and by adding the inequality $\sum_{j \in M} \sum_{k \in K^-} v_{jk} \leq d(K^-)$, where K^- is the set of removed clients, we obtain an inequality that is stronger than the original EC inequality. Let

$$\begin{aligned} K^- &= \cup_{j \in Q^-} K_j, \\ K' &= K \setminus K^-, \\ K'_j &= K_j \setminus K^-, \quad j \in J, \\ J' &= J \setminus J^-. \end{aligned}$$

The new values of the effective capacities $\bar{m}_j$ are:

$$\begin{aligned} \bar{m}'_j &= \bar{m}_j, \quad j \in Q \setminus Q^-, \\ \bar{m}'_j &= \min(m_j, d(K')), \quad j \in J \setminus Q, \end{aligned}$$

giving a new excess $\lambda' = \sum_{j \in J} \bar{m}'_j - d(K')$. If $\min(m_j, d(K')) = m_j$ for all $j \in J \setminus Q$, then $\lambda' = \lambda > 0$, since by assumption $K_{q_1} \cap K_{q_2} = \emptyset$ for all $q_1, q_2 \in Q$. Let $(J \setminus Q)' \subseteq (J \setminus Q)$ be the set of depots for which $d(K') < m_j$. Then

$$\lambda' = \sum_{j \in (J \setminus Q) \setminus (J \setminus Q)'} m_j + (|(J \setminus Q)'| - 1)d(K') + \sum_{j \in Q \setminus Q^-} \bar{m}_j \geq 0.$$

We now show that the inequality

$$\sum_{j \in J'} \sum_{k \in K'_j} v_{jk} + \sum_{j \in M} \sum_{k \in K^-} v_{jk} + \sum_{j \in J'} (\bar{m}'_j - \lambda')^+ (1 - y_j) \leq d(K') + d(K^-) \quad (14)$$

is stronger than the EC inequality based on sets J , K and K_j .

Since $\lambda' \geq 0$, inequality (14) is valid. All coefficients of the flow variables v_{jk} are equal in the two inequalities except for the coefficients for $v_{jk}, j \in Q, k \in K^- \setminus K_j$ and $j \in M \setminus J, k \in K^-$, which have value one in inequality (14) and zero in the EC inequality. The right-hand sides of the equalities are equal as $d(K) = d(K') + d(K^-)$. It remains to show that $(\bar{m}'_j - \lambda')^+ \geq (\bar{m}_j - \lambda)^+$ for all $j \in J'$.

Case 1: $j \in J' \setminus Q$

$$\begin{aligned} \bar{m}'_j - \lambda' &= \min(m_j, d(K')) - \sum_{l \in J'} \bar{m}'_l + d(K') = \\ &= \min(m_j, d(K')) - \sum_{l \in Q \setminus Q^-} \bar{m}_l - \sum_{l \in J \setminus Q} \min(m_l, d(K')) + d(K) - d(K^-) = \\ &= - \sum_{l \in (J \setminus Q) \setminus \{j\}} \min(m_l, d(K')) - \sum_{l \in Q \setminus Q^-} \bar{m}_l + \sum_{l \in J} \bar{m}_l - \lambda - d(K^-) = \\ &= \bar{m}_j - \lambda + \sum_{l \in (J \setminus Q) \setminus \{j\}} (\bar{m}_l - \min(m_l, d(K'))) + \sum_{l \in Q^-} \bar{m}_l - d(K^-) = \\ &= \bar{m}_j - \lambda + \sum_{l \in (J \setminus Q) \setminus \{j\}} (\bar{m}_l - \min(m_l, d(K'))) \geq \bar{m}_j - \lambda. \end{aligned}$$

Case 2: $j \in Q \setminus Q^-$

$$\begin{aligned}
\bar{m}'_j - \lambda' &= \bar{m}_j - \sum_{l \in J'} \bar{m}'_l + d(K') = \\
&= \bar{m}_j - \sum_{l \in J \setminus Q} \min(m_l, d(K')) - \sum_{l \in Q \setminus Q^-} \bar{m}_l + d(K) - d(K^-) = \\
&= \bar{m}_j - \sum_{l \in J \setminus Q} \min(m_l, d(K')) - \sum_{l \in Q \setminus Q^-} \bar{m}_l + \sum_{j \in J} \bar{m}_l - \lambda - d(K^-) = \\
&= \bar{m}_j - \lambda + \sum_{l \in J \setminus Q} (m_l - \min(m_l, d(K'))) + \sum_{l \in Q^-} \bar{m}_l - d(K^-) = \\
&= \bar{m}_j - \lambda + \sum_{l \in J \setminus Q} (m_l - \min(m_l, d(K'))) \geq \bar{m}_j - \lambda.
\end{aligned}$$

■

If $\rho_j(J \setminus \{j\}) = (\bar{m}_j - \lambda)^+$, it means that $f(J \setminus \{j\}) = \min(\sum_{l \in J \setminus \{j\}} \bar{m}_l, d(K))$. The case where $\rho_j(J \setminus \{j\}) > (\bar{m}_j - \lambda)^+$ is difficult to analyze since, in general, we have no closed-form expression of $f(J \setminus \{j\})$. Consider a submodular inequality for which $f(J) = d(K)$ and whose support graph is connected. Aardal et al. (1993) observed that if $\rho_j(J \setminus \{j\}) > (\bar{m}_j - \lambda)^+$ for some $j \in J$, then there exists a non-trivial partition of the clients $(\bar{K}, K \setminus \bar{K})$ and the depots $(\bar{J}, J \setminus \bar{J})$ with $j \in \bar{J}$, and with the clients in $\bar{K}$ uniquely served by the depots in $\bar{J}$. This observation led to the development of two classes of facet-defining submodular inequalities having $\rho_j(J \setminus \{j\}) > (\bar{m}_j - \lambda)^+$ for at least one $j \in J$. We describe one of these classes, the single-depot inequalities, below.

Lemma 6 (Aardal et al. (1993).) *Let C^{EC} be an EC-component with client set K^{EC} , depot set J^{EC} , and arc set $\{(j, k) : j \in J^{EC}, k \in K_j \subseteq K^{EC}\}$, and such that J^{EC} , K^{EC} and $\{K_j\}_{j \in J^{EC}}$ satisfy the conditions of Theorem 4. The set $Q^{EC} \subset J^{EC}$ is the set of depots in J^{EC} having $\bar{m}_j < m_j$. Let P be a set of additional depots with client set $K_p = K^{EC} \cup \bar{K}_p$, $\bar{K}_p \neq \emptyset$, $p \in P$, where the clients in $\bar{K}_p$ are served uniquely by depot p , and such that $m_p > d(\bar{K}_p)$ for all $p \in P$. Let $J = J^{EC} \cup P$, $K = K^{EC} \cup_{p \in P} \bar{K}_p$ and $\lambda = \sum_{j \in J} \bar{m}_j - d(K)$.*

Then $\rho_j(J \setminus \{j\}) = (\bar{m}_j - \lambda)^+$ for all $j \in J^{EC}$ and $\rho_p(J \setminus \{p\}) = d(\bar{K}_p) > (\bar{m}_p - \lambda)^+$ for all $p \in P$.

In Figure 1 we show the structure of the support graph corresponding to a single depot inequality.

Figure 1. Single-depot structure.

Theorem 7 (Single-depot inequalities, Aardal et al. (1993).) *Assume that $f(J) = d(K) < \sum_{j \in J} \bar{m}_j$, and that $\sum_{j \in M} m_j > d(N) + m_r$ for all $r \in J$. Then the submodular inequality (10) based on the single-depot structure, as defined in Lemma 6, defines a facet of $\text{conv}(X^{CFL})$ if and only if*

- (1) $\rho_q(J \setminus \{q\}) > 0$ for all $q \in Q^{EC}$
- (2) if $|Q^{EC}| \leq 1$, then $\exists j \in J^{EC} \setminus Q^{EC}$ with $\rho_j(J \setminus \{j\}) > 0$.

3 Separation Heuristics

The separation problem based on a certain family $\mathcal{F}$ of valid inequalities can be formulated as follows.

Given a fractional solution (v^, y^*) , show that (v^*, y^*) satisfies all valid inequalities in $\mathcal{F}$, or find an inequality belonging to $\mathcal{F}$ that is violated by (v^*, y^*) .*

Given a client set $K \subseteq N$, the separation problem based on the set of flow cover inequalities (11), can be formulated as follows. Let $z_j = 1$ if $j \in J$ and $z_j = 0$ otherwise.

$$\begin{aligned} \max \quad & \sum_{j \in M} [v_j'^* + (m_j - \lambda)^+(1 - y_j^*)]z_j - d(K) \\ \text{s.t.} \quad & \sum_{j \in M} m_j z_j = d(K) + \lambda \\ & \lambda > 0 \\ & z_j \in \{0, 1\} \end{aligned} \tag{15}$$

Problem (15) is a knapsack problem and hence NP-hard. However, if $m_j = m$ for all $j \in M$ we have the following result.

Theorem 8 *If $m_j = m$ for all $j \in M$, the separation problem based on the flow cover inequalities can be solved in polynomial time.*

Proof. Assume that $d(K)$ is not an integer multiple of m . Let $l = \lceil d(K)/m \rceil$ and $r = d(K) - m(l - 1)$. Problem (15) can now be formulated as:

$$\begin{aligned} \max \quad & \xi = \sum_{j \in M} [v_j'^* - r y_j^*]z_j + rl - d(K) \\ \text{s.t.} \quad & \sum_{j \in M} z_j \geq l \\ & z_j \in \{0, 1\} \end{aligned} \tag{16}$$

First we show that if $\xi > 0$, then $\sum_{j \in M} z_j \geq l$, i.e. all solutions $z \in \{0, 1\}^m$ giving $\xi > 0$ are feasible in (16). Let $f = \sum_{j \in M} z_j - \sum_{j \in M} y_j^* z_j$. For given $z \in \{0, 1\}^m$ we get

$$\begin{aligned} \xi &= \sum_{j \in M} [v_j'^* - r y_j^*]z_j + rl - d(K) \leq m \sum_{j \in M} y_j^* z_j - r \sum_{j \in M} y_j^* z_j + rl - d(K) = \\ &= m \sum_{j \in M} z_j - m f - r \sum_{j \in M} z_j + r f + rl - r - m(l - 1) = (m - r)(\sum_{j \in M} z_j - (l - 1)) - f(m - r) = \\ &= (m - r)(\sum_{j \in M} z_j - (l - 1) - f). \end{aligned}$$

It now follows that if $\xi > 0$, then $\sum_{j \in M} z_j \geq l$ since $(m - r) \geq 0$ and $f \geq 0$. Solving the separation problem (16) can now be reduced to letting $z_j = 1$ if $(v_j'^* - r y_j^*) > 0$ and $z_j = 0$ otherwise, and then checking if the resulting value of ξ is positive. $\blacksquare$

A heuristic for solving (15) for a given client set K is given by Van Roy and Wolsey (1987). For fixed client set $K \subseteq N$, the coefficients $(\bar{m}_j - \lambda)^+$ and $\rho_j(J \setminus \{j\})$ of the y_j -variables in the EC and the general submodular inequalities respectively, depend on the arc set. Moreover

$\rho_j(J \setminus \{j\})$ depends on the choice of depot set J , and in order to determine the value of each coefficient $\rho_j(J \setminus \{j\})$, we need to apply a maximum flow algorithm. All these difficulties makes the separation problem based on the EC and the general submodular inequalities very hard to analyze, and it is not possible to solve any of these problems as a single integer program. We have therefore developed a heuristic for identifying violated EC inequalities, and to further extend the EC inequalities to single-depot inequalities. The heuristic consists of two parts; the first part, presented in the next subsection, aims at identifying good candidate sets J and K , and the second part, described in Section 3.2, determines an appropriate arc set and possible P -depots.

3.1 A General Heuristic for Identifying Sets J and K

For all three special cases of submodular inequalities presented in the previous section we need to identify an EC component, i.e. a client set K , a cover J and an arc set where at least one depot in J serves all clients in K . To get a single-depot structure we also need to find a set P of depots and client sets $\bar{K}_p$ for all $p \in P$, see Lemma 6. The following heuristic produces sets J and K , with J defining a cover with respect to M and K . Depending on the structure of the graph induced by J , K and the active arcs in the fractional solution, and on the characteristics of the instance, we then try to identify violated flow cover, EC, and single-depot inequalities. The heuristic is based on three observations regarding the EC inequalities

$$\sum_{j \in J} \sum_{k \in K_j} v_{jk} + \sum_{j \in J} (\bar{m}_j - \lambda)^+ (1 - y_j) \leq d(K). \quad (12)$$

Let (v^*, y^*) denote the current fractional solution.

Observation 1 *Inequality (12) can be violated only if at least one of the y_j -variables has a fractional value.*

To guarantee that J contains at least one fractional depot, we choose as initial depot set $J^1 = \{j^1\}$ where $0 < y_{j^1}^* < 1$.

Observation 2 *The value of the flow $\sum_{j \in J} \sum_{k \in K_j} v_{jk}$ should be as close as possible to $d(K)$.*

At iteration t , given the current client set K^{t-1} and the depot set $J^t = \{j^1, \dots, j^t\}$, let K^{t-1} be augmented by the set of clients served by j^t that are not already in K^{t-1} , giving set K^t . Moreover, we regard as candidates for J , all depots, apart from the depots already in J^t , that serve clients in K^t . The choice of depot to include next in J is based on the following observation.

Observation 3 *The chance of violation increases if depots having a small value of $s_j = \bar{m}_j y_j^* - \sum_{k \in K_j} v_{jk}^*$ are included in J .*

To see that this is true, replace $\sum_{k \in K_j} v_{jk}^*$ by $\bar{m}_j y_j^* - s_j$ and $(\bar{m}_j - \lambda)^+$ by $(\bar{m}_j - \lambda)$ in the EC inequality (12), and use $d(K) = \sum_{j \in J} \bar{m}_j - \lambda$. Replacing $(\bar{m}_j - \lambda)^+$ by $(\bar{m}_j - \lambda)$ makes (12) weaker. This weaker form of (12) can now be written as

$$\lambda - \lambda \sum_{j \in J} (1 - y_j^*) \leq \sum_{j \in J} s_j.$$

In practice it will often be the case, especially if the problem instance is tightly capacitated, that $s_j = 0$ for several depots considered as candidates for set J . To break ties, we choose as depot j^{t+1} , i.e. as the depot by which J^t should be augmented, the depot with largest flow contribution $\sum_{k \in K^t} v_{jk}^*$ among the candidate depots having minimum value of s_j . This is to avoid including a depot that only marginally contributes to the flow whereas the effective capacity is fully counted for.

The algorithm consists of two major loops. At every iteration of the outer loop, we choose a depot with fractional value as the first depot j^1 in the set J . This is repeated until all depots with fractional value have been considered. The inner loop is initialized with the set J^1 ($K^0 = \emptyset$). The sets K^t and J^t are then iteratively augmented according to the second and third observations. If at a certain iteration t , J^t defines a cover with respect to M and K^t , we check whether a flow cover, EC, or single-depot inequality based on (J^t, K^t) is violated. If the set of candidate depots, that can be used to augment J^t , is empty the inner loop is terminated. We also stop augmenting J^t and K^t if we at n subsequent iterations have generated a cover, but failed to generate a violated inequality. In our experiments we use $n = 2$.

3.2 Identifying Violated Flow Cover, EC and Single-Depot Inequalities

Given sets J and K we temporarily define K_j as $K_j = \{k \in K : v_{jk}^* > 0\}$, $\bar{m}_j = \min(m_j, d(K_j))$ for all $j \in J$, and $Q = \{j \in J : \bar{m}_j < m_j\}$. We distinguish between the following three cases.

Case 1: $Q = \emptyset$ and $|J| = 2$

We redefine K_j as $K_j = K$ for all $j \in J$, giving a flow cover inequality (11).

Case 2: $Q = \emptyset$ and $|J| > 2$

We try to identify P -depots as described below.

Case 3: $0 < |Q| \leq |J|$

We first check if any of the depots in Q have intersecting client sets, i.e. if $K_{q_1} \cap K_{q_2} \neq \emptyset$ for any $q_1, q_2 \in Q$. If that is the case, we remove all clients belonging to any such intersection. This rule is based on the proof that Condition (1) of Theorem 4 is necessary for (12) to be facet defining. Given the resulting sets J' , Q' , K' and K'_j for all $j \in J'$, we update the effective capacities $\bar{m}_j$ and check whether there are any depots in Q' for which $\bar{m}_j \leq \lambda$. These depots and their client sets K'_j are removed from J' and K' , which is a rule based on the proof that Condition (4) of Theorem 4 is necessary for the EC inequalities to be facet defining. We again update the effective capacities and obtain sets J'' , Q'' , K'' and K''_j for all $j \in J''$. Let K^r denote the set of removed clients and let for simplicity $J = J''$, $Q = Q''$, $K = K''$ and $K_j = K''_j$.

If $|J| > 2$ we try to identify depots that could be considered as P -depots, see Lemma 6. As candidates for P we consider depots p that at the current fractional solution uniquely serve a subset $\bar{K}_p \subset K$ of the clients and for which $d(\bar{K}_p) < m_p < d(K)$. For a single-depot structure we require the depots that constitute the EC-component to define a cover with respect to the clients in the EC-component. Therefore, after choosing appropriate depots as candidates for P , denoted $\text{cand}(P)$, we check whether (i) the resulting set $J \setminus \text{cand}(P)$ defines a cover with respect to $K \setminus (\cup_{p \in \text{cand}(P)} \bar{K}_p)$ and if (ii) $(J \setminus \text{cand}(P)) \setminus Q \neq \emptyset$. If this is the case we have generated a single-depot structure with $P = \text{cand}(P)$, otherwise we need to remove depots from $\text{cand}(P)$.

and include them in $J \setminus \text{cand}(P)$ until conditions (i) and (ii) are satisfied. Depots are removed from $\text{cand}(P)$ in order of increasing values of $d(\overline{K}_p)$. This is done in order to keep depots with high values of $\rho_j(J \setminus \{j\}) = d(\overline{K}_p)$ in the set P if possible.

When the algorithm terminates, we have the final depot and client sets J and K , which can be partitioned into the sets $(Q, P, J \setminus (Q \cup P))$ and $(\cup_{p \in P} \overline{K}_p, K \setminus (\cup_{p \in P} \overline{K}_p))$ respectively. Moreover we have the set K^r of clients that in Case 3 were removed from the initial client set. If the corresponding inequality is violated, the following variables are stored:

$$\begin{aligned} v_{jk}, \quad & j \in Q, k \in K_j, \\ & j \in P, k \in K \setminus \cup_{p \in P \setminus \{j\}} \overline{K}_p, \\ & j \in J \setminus (Q \cup P), k \in K \setminus \cup_{p \in P} \overline{K}_p, \\ & j \in M, k \in K^r, \\ y_j, \quad & j \in J. \end{aligned}$$

All variables v_{jk} have coefficient 1, and the coefficient of y_j is $-\rho_j(J \setminus \{j\}) = -(\overline{m}_j - \lambda)^+$ for $j \in J \setminus P$, and $-\rho_j(J \setminus \{j\}) = -d(\overline{K}_p)$ for $j \in P$. The right-hand side of the inequality is $d(K) + d(K^r) - \sum_{j \in J} \rho_j(J \setminus \{j\})$.

3.3 Identifying Violated Combinatorial Inequalities

Given $J \subseteq M$, $K \subseteq N$ and $K_j \subseteq K$ for all $j \in J$, let β be the minimum number of depots in J needed to reach – or cover – all clients in K . β is called the *covering number*. Let S be a zero-one $|J| \times |K|$ matrix with $s_{jk} = 1$ if $k \in K_j$ and $s_{jk} = 0$ otherwise. The matrix S is the *adjacency matrix* corresponding to the bipartite graph $G^S = (V^S, U^S, E^S)$ with $V = \{j \in J\}$, $U = \{k \in K\}$, and $E^S = \{(j, k) : j \in J, k \in K_j\}$. Assume that (i) $|J|, |K| \geq 3$, that (ii) the graph G^S is connected, and that (iii) each column of S has at least one 0 and one 1. Changing an element s_{jk} from zero to one corresponds to adding an edge (j, k) in G^S . If no such change can take place without decreasing β , S is said to be *maximal*.

Cho et al. (1983a,b) developed the class of *combinatorial inequalities*

$$\sum_{j \in J} \sum_{k \in K_j} \frac{1}{d_k} v_{jk} - \sum_{j \in J} y_j \leq |K| - c, \quad (17)$$

where $c \leq \beta$, for the uncapacitated facility location (UFL) problem. They showed that these inequalities define facets for the convex hull of feasible solutions to UFL if and only if $c = \beta$, and the adjacency matrix S corresponding to the support graph of the inequality is maximal and satisfies conditions (i)-(iii) given above. Since UFL is a relaxation of CFL, the inequalities (17) are also valid for $\text{conv}(X^{CFL})$. Aardal et al. (1993) showed that (17) define facets of $\text{conv}(X^{CFL})$ if, in addition to S being maximal and $c = \beta$, the part of the instance defined by the sets J , K and K_j is not too tightly capacitated. For details see Aardal et al., Theorem 20.

The class of combinatorial inequalities is very general, and subsumes all classes of valid inequalities developed so far for UFL. The previously known special cases, see Padberg (1973), Cornuéjols et al. (1977), Guignard (1980) and Cornuéjols and Thizy (1982), all have the property that the adjacency matrix corresponding to the bipartite support graph of the inequality is cyclic. A cyclic adjacency matrix has the same number of ones in every row; it is possible to permute

the rows and the columns such that in every row all ones appear in sequence and are shifted one position to the right for every next row. This implies that all depots serve equally many clients and that all clients are served by equally many depots. Although a fair amount of attention has been given to the various subclasses of combinatorial inequalities, very little work has been done on the separation problem based on the general class, and no computational experience with any of the subclasses is reported in the literature. Here we present a heuristic for generating violated combinatorial inequalities that uses the sets J and K , produced by the general heuristic described in Section 3.1, as input.

Due to the way that the sets J and K are generated, the bipartite graph $G = (V, U, E)$ where $V = \{j \in J\}$, $U = \{k \in K\}$ and $E = \{(j, k) : v_{jk}^* > 0, j \in J, k \in K\}$ is connected. Moreover, $\sum_{j \in J} m_j > d(K)$. Determining β , which is equivalent to determining the minimum number of rows of the adjacency matrix corresponding to G , such that there is at least one 1 in every column, is an NP-hard problem. We therefore start by approximating β by the cardinality of a minimum subset $I \subseteq J$ such that $\sum_{j \in I} d(K_j) > d(K)$, where K_j initially is defined as $K_j = \{k \in K : v_{jk}^* > 0\}$. It is clear that $|I| \leq \beta$. Assuming that the right-hand side of the inequality is $|K| - |I|$, and that all depots serve equally many clients, as in the case of cyclic adjacency matrices, it is possible to calculate the maximal outdegree t of each depot by using $|I| = \lceil |K|/t \rceil$. For each depot we sort the values v_{jk}^* , $k \in K$, in nondecreasing order and we redefine K_j to be the set of clients corresponding to the t highest values of v_{jk}^* . If the corresponding inequality is violated, we store variables v_{jk} , $j \in J$, $k \in K_j$, with coefficient $1/d_k$ and variables y_j with coefficient 1. The right-hand side of the inequality is $|K| - |I|$.

4 Computational Experience

We solve 60 capacitated facility location problems of four different sizes and three different capacity levels. Each test problem is represented by the notation $v-xyyz$, where v is the capacity level, xx the number of clients, yy the number of depots, and z the number of the instance in the series of test problems having equal values of v , xx and yy . The different capacity levels we consider are $v = 1, 2, 3$ corresponding to ratios $\sum_{j \in M} m_j/d(N) = 1.5, 2$ and 3 respectively. The number of variables, constraints and nonzeros corresponding to the different problem sizes are reported in Table 1. For more details concerning the test problems we refer to Cornuéjols et al. (1991), who developed this set of instances to provide more difficult and general test problems than the classical Kuhn-Hamburger instances.

problem type	# variables	# constraints	# nonzeros
25 × 8	216	242	832
50 × 16	832	883	3264
50 × 33	1716	1767	6732
50 × 50	2600	2651	10200

Table 1. Problem characteristics.

For the computational experiments we use MINTO 1.6a (Savelsbergh et al. (1991)) with the CPLEX 2.1 callable library, implemented on a SUN Sparc ELC. We refer to the knapsack cover

inequalities and the flow cover inequalities with $K = N$, which are automatically generated by MINTO from the substructures (7) and (8), as *MINTO inequalities*. The *user inequalities* are the EC, single depot and combinatorial inequalities. We consider the flow cover inequalities with $K \subset N$ as a special case of the EC inequalities. Three solution strategies are used, all starting from the LP-relaxation of CFL.

- I. Branch and bound.
- II. Branch and cut using knapsack cover inequalities generated by MINTO.
- III. Branch and cut using MINTO inequalities and User inequalities at the root node, and MINTO inequalities further down in the branch and bound tree.

The reason that we investigate the performance of the knapsack inequalities alone – Strategy II – is that our experience showed that they are the single most effective class of inequalities among the classes used in this study. We only generate user inequalities at the root node – Strategy III – since the main purpose is to test the effect the reduction of the duality gap in the root node has on generating fewer branch and bound nodes and thereby on reducing the total time needed to solve the problems. The results from using Strategy I–III are reported in Tables 2-4 below. The duality gap is calculated as $(z_{IP} - z_{LP})/z_{IP}$, where z_{IP} and z_{LP} denote the optimal solution to CFL and the LP-relaxation of CFL respectively. The percentage duality gap closed by the inequalities is determined by $(z_{CUT} - z_{LP})/(z_{IP} - z_{LP})$, where z_{CUT} is the optimal value of the LP-relaxation with all inequalities generated at the root node added to the formulation. In Tables 3 and 4, *% time reduction* refers to the percentage reduction in total time of Strategy II and III respectively, compared to Strategy I, whereas *average % time reduction* is the average percentage reduction over the instances in the same set. The *user time* is the time needed to generate User inequalities. Both user time and total time are given in seconds after being rounded down. All results are summarized in Table 5.

problem	duality gap (%)	# B&B nodes	time
1-25081	5.7	31	5
1-25082	6.4	70	17
1-25083	5.9	21	6
1-25084	1.3	9	4
1-25085	5.5	71	13
1-50161	2.2	61	59
1-50162	0.4	19	17
1-50163	2.3	143	129
1-50164	2.1	141	173
1-50165	0.9	43	43
1-50331	1.5	1361	1813
1-50332	1.2	797	860
1-50333	0.2	31	66
1-50334	1.0	1349	2606
1-50335	1.6	1617	2050
1-50501	0.3	143	278
1-50502	0.1	67	116
1-50503	0.4	361	681
1-50504	0.2	123	283
1-50505	0.0	1	75
2-25081	5.9	23	8
2-25082	10.3	125	34
2-25083	7.5	79	25
2-25084	2.2	9	6
2-25085	5.2	19	7
2-50161	3.7	129	145
2-50162	2.6	95	103
2-50163	0.5	19	30
2-50164	2.5	345	512
2-50165	0.0	1	15
2-50331	1.5	399	686
2-50332	1.2	691	1560
2-50333	1.5	259	556
2-50334	0.7	239	493
2-50335	1.3	685	1232
2-50501	0.2	143	296
2-50502	1.0	1625	4050
2-50503	0.1	45	164
2-50504	0.1	33	129
2-50505	0.02	3	67
3-25081	7.1	11	5
3-25082	1.5	7	5
3-25083	5.8	33	10
3-25084	4.5	45	13
3-25085	4.1	7	5
3-50161	4.1	127	160
3-50162	5.4	109	189
3-50163	1.3	17	38
3-50164	2.0	31	104
3-50165	2.1	7	23
3-50331	2.5	187	1727
3-50332	2.9	766	5188
3-50333	1.6	249	626
3-50334	0.8	57	194
3-50335	2.1	193	844
3-50501	1.4	299	1479
3-50502	1.1	443	1949
3-50503	0.9	1003	3162
3-50504	0.5	273	666
3-50505	0.7	199	686

Table 2. Results from using Strategy I.

problem	# cover inequalities	% gap closed	# B&B nodes	time	% time reduction	average % time reduction
1-25081	5	94.0	3	4	20.0	28.9
1-25082	10	78.1	9	5	70.6	
1-25083	5	69.0	5	6	0.0	
1-25084	3	100.0	1	4	0.0	
1-25085	4	82.8	7	6	53.8	
1-50161	31	68.1	17	51	13.6	22.9
1-50162	2	48.0	13	20	-17.6	
1-50163	16	82.0	13	48	62.8	
1-50164	33	33.7	35	109	37.0	
1-50165	24	52.3	15	35	18.6	
1-50331	34	81.8	41	173	90.5	57.1
1-50332	118	65.7	129	476	44.7	
1-50333	0	0.0	31	66	0.0	
1-50334	130	10.8	175	975	62.6	
1-50335	46	78.4	49	255	87.6	
1-50501	3	100.0	1	54	80.6	39.7
1-50502	9	37.2	7	93	19.8	
1-50503	10	32.4	7	107	84.3	
1-50504	43	38.9	47	356	-25.8	
1-50505	LP optimal	-	-	-	-	
2-25081	4	100.0	1	3	62.5	48.6
2-25082	10	74.3	7	7	79.4	
2-25083	6	85.5	5	8	68.0	
2-25084	1	100.0	1	4	33.3	
2-25085	5	86.6	3	7	0.0	
2-50161	16	77.6	19	58	60.0	41.5
2-50162	20	30.3	13	64	37.9	
2-50163	4	51.5	9	36	-20.0	
2-50164	12	62.0	23	68	86.7	
2-50165	LP optimal	-	-	-	-	
2-50331	13	86.0	31	125	81.8	50.3
2-50332	58	54.3	51	450	71.2	
2-50333	122	54.1	89	769	-38.3	
2-50334	42	76.6	23	213	56.8	
2-50335	25	78.3	49	248	79.9	
2-50501	8	82.5	7	76	74.3	22.7
2-50502	158	59.4	225	1551	61.7	
2-50503	6	0.0	31	170	-3.7	
2-50504	5	0.0	31	153	-18.6	
2-50505	0	0.0	3	67	0.0	
3-25081	4	100.0	1	7	-40.0	-12.8
3-25082	5	92.8	3	7	-40.0	
3-25083	6	43.3	7	11	-10.0	
3-25084	4	98.9	3	7	46.1	
3-25085	4	100.0	1	6	-20.0	
3-50161	15	52.3	13	76	52.5	9.8
3-50162	20	53.6	15	98	48.1	
3-50163	9	21.0	9	55	-44.7	
3-50164	22	82.4	15	80	23.1	
3-50165	5	39.7	5	30	-30.4	
3-50331	119	64.4	107	1076	37.7	48.2
3-50332	78	66.2	83	495	90.5	
3-50333	108	0.2	77	591	5.6	
3-50334	9	27.8	25	120	38.1	
3-50335	45	42.2	45	260	69.2	
3-50501	57	66.0	37	352	76.2	46.4
3-50502	79	17.4	125	987	49.3	
3-50503	105	6.3	125	997	68.5	
3-50504	56	0.0	79	693	-4.1	
3-50505	31	12.0	73	396	42.3	

Table 3. Results from using Strategy II.

	# MINTO inequalities		# User inequalities								
problem	cover	flow cover	EC	single-depot	combinatorial	time user	%gap closed	# B&B nodes	total time	% time reduct.	average % time reduct.
1-25081	4	8	3	4	0	1	100.0	1	4	20.0	30.3
1-25082	7	10	5	7	0	0	81.9	5	8	52.9	
1-25083	6	9	5	4	0	1	76.1	7	6	0.0	
1-25084	1	2	0	0	0	0	100.0	1	3	25.0	
1-25085	4	12	6	3	0	0	88.5	5	6	53.8	
1-50161	15	23	9	13	0	0	86.5	9	49	16.9	47.7
1-50162	2	4	3	1	0	0	100.0	1	10	41.2	
1-50163	11	19	8	9	0	0	85.0	11	35	72.9	
1-50164	13	20	13	17	0	0	47.1	19	76	56.1	
1-50165	4	5	5	8	0	1	92.2	3	21	51.2	
1-50331	10	31	5	18	0	1	89.3	19	125	93.1	81.9
1-50332	52	65	4	11	0	0	70.1	35	208	75.8	
1-50333	0	2	1	0	0	0	100.0	1	23	65.2	
1-50334	75	97	3	14	0	1	53.8	69	488	81.3	
1-50335	22	39	1	6	0	0	81.8	25	124	94.0	
1-50501	3	5	1	35	0	1	100.0	1	56	79.9	72.4
1-50502	2	4	5	12	0	0	96.8	3	52	55.2	
1-50503	10	11	0	0	0	0	32.4	7	79	88.4	
1-50504	6	8	1	20	0	1	54.1	7	96	66.1	
1-50505	LP	opt.	—	—	—	—	—	—	—	—	
2-25081	3	5	4	3	0	0	100.0	1	4	50.0	52.4
2-25082	9	19	12	8	0	1	75.5	7	11	67.6	
2-25083	5	7	4	0	0	0	86.0	7	8	68.0	
2-25084	1	2	1	0	0	0	100.0	1	4	33.3	
2-25085	3	5	4	1	0	0	100.0	1	4	42.9	
2-50161	15	30	11	8	0	1	80.1	13	69	52.4	54.6
2-50162	11	20	4	1	0	0	30.3	13	72	30.1	
2-50163	1	2	2	5	0	0	86.0	3	15	50.0	
2-50164	12	16	7	20	0	0	66.7	15	72	85.9	
2-50165	LP	opt.	—	—	—	—	—	—	—	—	
2-50331	13	24	8	8	0	0	92.1	11	115	83.2	62.6
2-50332	33	36	4	29	0	1	67.2	27	290	81.4	
2-50333	32	43	8	23	0	1	62.1	25	177	68.1	
2-50334	17	25	7	14	0	0	81.9	11	97	80.3	
2-50335	17	23	8	10	0	1	88.0	14	124	89.9	
2-50501	5	5	4	14	0	0	85.1	5	71	76.0	52.0
2-50502	56	71	2	3	0	0	60.6	77	734	81.9	
2-50503	8	9	6	9	0	1	94.3	5	97	40.9	
2-50504	2	3	3	21	0	0	100.0	1	60	53.5	
2-50505	0	1	2	8	0	0	95.6	3	62	7.5	
3-25081	3	5	2	2	0	0	100.0	1	3	40.0	32.3
3-25082	2	3	4	2	0	1	100.0	1	4	20.0	
3-25083	4	6	6	0	3	0	46.5	7	6	40.0	
3-25084	3	5	4	5	0	0	99.1	3	5	61.5	
3-25085	3	5	2	2	0	0	100.0	1	5	0.0	
3-50161	7	14	6	2	0	1	64.2	9	58	63.8	53.6
3-50162	15	15	7	3	1	0	60.9	11	69	63.5	
3-50163	4	6	12	1	0	0	99.7	3	26	31.6	
3-50164	5	9	9	5	2	1	91.3	3	31	70.2	
3-50165	1	2	1	0	0	0	100.0	1	14	39.1	
3-50331	48	66	14	3	0	0	67.6	47	474	72.6	62.8
3-50332	56	82	20	8	0	1	73.4	49	521	90.0	
3-50333	37	47	10	2	0	0	47.4	39	394	37.1	
3-50334	9	12	7	4	1	1	45.8	11	108	44.3	
3-50335	23	34	7	4	2	1	53.3	23	254	69.9	
3-50501	10	11	11	25	1	0	85.6	9	212	85.7	45.8
3-50502	68	97	6	8	0	0	21.7	71	1317	32.4	
3-50503	55	47	7	1	0	0	29.1	67	674	78.7	
3-50504	36	46	9	9	0	0	25.5	55	800	-20.1	
3-50505	30	19	10	17	0	1	54.0	31	326	52.5	

Table 4. Results from using Strategy III.

problem	duality gap	Strategy I		Strategy II			Strategy III		
		# B& B nodes	time	% gap closed	# B& B nodes	time	% gap closed	# B& B nodes	time
1-25081	5.7	31	5	94.0	3	4	100.0	1	4
1-25082	6.4	70	17	78.1	9	5	81.9	5	8
1-25083	5.9	21	6	69.0	5	6	76.1	7	6
1-25084	1.3	9	4	100.0	1	4	100.0	1	3
1-25085	5.5	71	13	82.8	7	6	88.5	5	6
1-50161	2.2	61	59	68.1	17	51	86.5	9	49
1-50162	0.4	19	17	48.0	13	20	100.0	1	10
1-50163	2.3	143	129	82.0	13	48	85.0	11	35
1-50164	2.1	141	173	33.7	35	109	47.1	19	76
1-50165	0.9	43	43	52.3	15	35	92.2	3	21
1-50331	1.5	1361	1813	81.8	41	173	89.3	19	125
1-50332	1.2	797	860	65.7	129	476	70.1	35	208
1-50333	0.2	31	66	0.0	31	66	100.0	1	23
1-50334	1.0	1349	2606	10.8	175	975	53.8	69	488
1-50335	1.6	1617	2050	78.4	49	255	81.8	25	124
1-50501	0.3	143	278	100.0	1	54	100.0	1	56
1-50502	0.1	67	116	37.2	7	93	96.8	3	52
1-50503	0.4	361	681	32.4	7	107	32.4	7	79
1-50504	0.2	123	283	38.9	47	356	54.1	7	96
1-50505	0.0	1	75	—	—	—	—	—	—
2-25081	5.9	23	8	100.0	1	3	100.0	1	4
2-25082	10.3	125	34	74.3	7	7	75.5	7	11
2-25083	7.5	79	25	85.5	5	8	86.0	7	8
2-25084	2.2	9	6	100.0	1	4	100.0	1	4
2-25085	5.2	19	7	86.6	3	7	100.0	1	4
2-50161	3.7	129	145	77.6	19	58	80.1	13	69
2-50162	2.6	95	103	30.3	13	64	30.3	13	72
2-50163	0.5	19	30	51.5	9	36	86.0	3	15
2-50164	2.5	345	512	62.0	23	68	66.7	15	72
2-50165	0.0	1	15	—	—	—	—	—	—
2-50331	1.5	399	686	86.0	31	125	92.1	11	115
2-50332	1.2	691	1560	54.3	51	450	67.2	27	290
2-50333	1.5	259	556	54.1	89	769	62.1	25	177
2-50334	0.7	239	493	76.6	23	213	81.9	11	97
2-50335	1.3	685	1232	78.3	49	248	88.0	14	124
2-50501	0.2	143	296	82.5	7	76	85.1	5	71
2-50502	1.0	1625	4050	59.4	225	1551	60.6	77	734
2-50503	0.1	45	164	0.0	31	170	94.3	5	97
2-50504	0.1	33	129	0.0	31	153	100.0	1	60
2-50505	0.02	3	67	0.0	3	67	95.6	3	62
3-25081	7.1	11	5	100.0	1	7	100.0	1	3
3-25082	1.5	7	5	92.8	3	7	100.0	1	4
3-25083	5.8	33	10	43.3	7	11	46.5	7	6
3-25084	4.5	45	13	98.9	3	7	99.1	3	5
3-25085	4.1	7	5	100.0	1	6	100.0	1	5
3-50161	4.1	127	160	52.3	13	76	64.2	9	58
3-50162	5.4	109	189	53.6	15	98	60.9	11	69
3-50163	1.3	17	38	21.0	9	55	99.7	3	26
3-50164	2.0	31	104	82.4	15	80	91.3	3	31
3-50165	2.1	7	23	39.7	5	30	100.0	1	14
3-50331	2.5	187	1727	64.4	107	1076	67.6	47	474
3-50332	2.9	766	5188	66.2	83	495	73.4	49	521
3-50333	1.6	249	626	0.2	77	591	47.4	39	394
3-50334	0.8	57	194	27.8	25	120	45.8	11	108
3-50335	2.1	193	844	42.2	45	260	53.3	23	254
3-50501	1.4	299	1479	66.0	37	352	85.6	9	212
3-50502	1.1	443	1949	17.4	125	987	21.7	71	1317
3-50503	0.9	1003	3162	6.3	125	997	29.1	67	674
3-50504	0.5	273	666	0.0	79	693	25.5	55	800
3-50505	0.7	199	686	12.0	73	396	54.0	31	326

Table 5. Summary of the results.

From the results in Table 3, we note that the knapsack cover inequalities are quite effective in reducing the total time required to solve the problems. Given that the surrogate knapsack polytope (7) is a drastic relaxation of CFL, this is quite surprising. For the problems with capacity level $v = 3$, i.e. the loosely capacitated problems, the cover inequalities seem useful for the larger instances. For the smaller instances, a relatively large percentage of the duality gap is closed by the cover inequalities, but the gain in terms of fewer branch and bound nodes is consumed by the time spent at separation. Forty-eight of the sixty instances are solved at least as fast using Strategy II as by Strategy I. Seven of the twelve instances requiring more time in Strategy II have $v = 3$. It is worth noticing here that the loosely capacitated problems do not seem easier to solve, something that is often claimed to be the case.

If we consider Strategy III, all instances but one are solved at least as fast as by Strategy I. Thirteen problems were solved to optimality at the root node and for sixteen of the problems the percentage time reduction was more than 75 %. For the instances of size 50×33 and 50×50 with capacity level $v = 3$, the percentage duality gap closed by the inequalities added at the root node is smaller than for the other instances, and the time required to solve them remains higher than for the more tightly capacitated instances of the same size. This is not unexpected since all inequalities used in this study, except the combinatorial, make use of the capacities. For a relatively uncapacitated instance, a typical fractional solution after adding only knapsack and flow cover inequalities, has very few depots with positive y_j -values and only one with a fractional value. This fractional value is usually very small. The fractional depot serves a small portion of the flow to a set of clients almost completely served by one other depot that is just slightly short of capacity. The structure of such a solution leaves little freedom to generate inequalities of the types we have discussed. If the inequalities generated at that point are not able to cut off the remaining duality gap, the flow tends to spread drastically in the sense that every depot with a positive y_j -value will serve almost every client, and almost all y_j -variables get fractional values, which makes it very hard to identify an isolated substructure on which an inequality that is substantially violated can be based. The consequence is then that quite a large number of inequalities with a small violation are generated and the solution value hardly increases.

Very few violated combinatorial inequalities were generated, and they were all generated for the loosely capacitated problems as could be expected. However, the heuristic we applied for these inequalities uses the sets J and K from the general heuristic that was developed for the submodular inequalities. Given a set K , the set J that we generate is likely to be a good choice of depot set, but given the different character of the combinatorial inequalities compared to the submodular inequalities, we may use the wrong client set. Moreover, given the active arcs, we do not compute the exact covering number. In order to solve loosely capacitated problems it will be important to develop a good heuristic based on the class of combinatorial problems, and also to develop new classes of inequalities that use the capacities in a less restrictive way.

5 Concluding Remarks

We have studied the cutting plane approach to solving CFL. The computational experience shows the importance of, in particular, the knapsack cover inequalities, but also the flow cover inequalities. By adding aggregate variables and constraints to the standard formulation of CFL, general purpose mixed integer optimizers such as MINTO and MPSARX (Van Roy and Wolsey (1987)) will recognize the surrogate knapsack and single-node flow polytopes (7) and (8), and generate knapsack cover and flow cover inequalities as part of the system procedures. There-

fore, even if we have used redundant variables and constraints, we have in most cases gained in reduction of the duality gap, and thereby in the total time needed to solve the problem. This supports the choice of formulation (1)–(6). We also conclude that in most cases we need more problem specific inequalities to tighten the duality gap sufficiently. The heuristic suggested in Section 3 for identifying violated submodular inequalities appears effective, in particular for the more tightly capacitated instances. The results were less satisfactory for the larger, relatively uncapacitated instances. Hence, it seems useful to improve the heuristic based on the class of combinatorial inequalities, as well as to develop new classes of strong inequalities. The new classes of inequalities should, ideally, rely less on capacities than the EC or single depot inequalities, and have a support graph that to a greater extent corresponds to the structure of a typical fractional solution than the support graph corresponding to the combinatorial inequalities.

Acknowledgements

I am most grateful for the financial support given by David Shmoys, Cornell University, through NSF grant CCR-9307391. I also want to thank the Department of Mathematics and Computing Science at Eindhoven University of Technology for letting me use their computing facilities, and Martin Savelsbergh, Georgia Institute of Technology, for his advice regarding the use of MINTO.

References

- K. Aardal, “On the solution of one and two-level capacitated facility location problems by the cutting plane approach”, Ph.D. Thesis, Université Catholique de Louvain (Louvain-la-Neuve, 1992).
- K. Aardal, Y. Pochet and L.A. Wolsey, “Capacitated facility location: valid inequalities and facets”, CORE Discussion Paper 9323 (Louvain-la-Neuve, 1993) to appear in *Mathematics of Operations Research*.
- D.C. Cho, E.L. Johnson, M.W. Padberg and M.R. Rao, “On the uncapacitated plant location problem. I: valid inequalities and facets”, *Mathematics of Operations Research* (1983a) 579–589.
- D.C. Cho, M.W. Padberg and M.R. Rao, “On the uncapacitated plant location problem. II: facets and lifting theorems”, *Mathematics of Operations Research* (1983b) 590–612.
- G. Cornuéjols and J.-M. Thizy, “Some facets of the simple plant location problem”, *Mathematical Programming* 23 (1982) 50–74.
- G. Cornuéjols, M.L. Fisher and G.L. Nemhauser, “On the uncapacitated location problem”, *Annals of Discrete Mathematics* 1 (1977) 163–177.
- G. Cornuéjols, R. Sridharan and J.-M. Thizy, “A comparison of heuristics and relaxations for the capacitated plant location problem”, *European Journal of Operational Research* 50 (1991) 280–297.

- Q. Deng and D. Simchi-Levi, “Valid inequalities, facets and computational experience for the capacitated concentrator location problem”, Research Report, Department of Industrial Engineering and Operations Research, Columbia University (New York, 1993).
- M. Guignard, “Fractional vertices, cuts and facets for the simple plant location problem”, *Mathematical Programming* 12 (1980) 150–162.
- J.M.Y. Leung and T.L. Magnanti, “Valid inequalities and facets of the capacitated plant location problem”, *Mathematical Programming* 44 (1989) 271–291.
- M.W. Padberg, “On the facial structure of set packing polyhedra”, *Mathematical Programming* 5 (1973) 199–215.
- M.W. Padberg, T.J. Van Roy and L.A. Wolsey, “Valid inequalities for fixed charge problems”, *Operations Research* 33 (1985) 842–861.
- M.W.P. Savelsbergh, G.C. Sigismondi and G.L. Nemhauser, “Functional description of MINTO, a Mixed INteger Optimizer”, Memorandum COSOR 91-17, Eindhoven University of Technology (Eindhoven, 1991) to appear in *Operations Research Letters*.
- T.J. Van Roy and L.A. Wolsey, “Solving mixed-integer programming problems using automatic reformulation”, *Operations Research* 35 (1987) 45–57.
- L.A. Wolsey, “Faces for a linear inequality in 0-1 variables”, *Mathematical Programming* 8 (1975) 165–178.
- L.A. Wolsey “Submodularity and valid inequalities in capacitated fixed charge networks”, *Operations Research Letters* 8 (1989) 119–124.