

Rounding algorithms for covering problems

Dimitris Bertsimas^{a,*}, Rakesh Vohra^{b,2}

^a *Massachusetts Institute of Technology, Sloan School of Management, 50 Memorial Drive, Cambridge, MA 02142-1347, USA*

^b *Department of Management Science, Ohio State University, Ohio, USA*

Received 1 February 1994; received in revised form 1 January 1996

Abstract

In the last 25 years approximation algorithms for discrete optimization problems have been in the center of research in the fields of mathematical programming and computer science. Recent results from computer science have identified barriers to the degree of approximability of discrete optimization problems unless $P = NP$. As a result, as far as negative results are concerned a unifying picture is emerging. On the other hand, as far as particular approximation algorithms for different problems are concerned, the picture is not very clear. Different algorithms work for different problems and the insights gained from a successful analysis of a particular problem rarely transfer to another.

Our goal in this paper is to present a framework for the approximation of a class of integer programming problems (covering problems) through generic heuristics all based on rounding (deterministic using primal and dual information or randomized but with nonlinear rounding functions) of the optimal solution of a linear programming (LP) relaxation. We apply these generic heuristics to obtain in a systematic way many known as well as new results for the set covering, facility location, general covering, network design and cut covering problems. © 1998 The Mathematical Programming Society, Inc. Published by Elsevier Science B.V.

1. Introduction

Given our inability to efficiently solve several discrete optimization problems (in particular NP-hard problems) exactly, it is natural to ask whether it is possible to approximate them. We will focus on minimization problems, but a parallel theory can be developed for maximization problems (see for example [1]). Algorithm A constitutes an approximation algorithm for minimization problem Π with guarantee $f(n)$, if for each instance I of size n of Π , Algorithm A runs in polynomial time in

* Corresponding author. E-mail: dbertsim@aris.mit.edu.

¹ Research partially supported by a Presidential Young Investigator Award DDM-9158118 with matching funds from Draper Laboratory.

² Research partially supported by a Deans Summer Fellowship of the College of Business of the Ohio State University.

n and returns a value $Z_A(I)$ such that $Z_A(I) \leq f(n)Z^*(I)$, where $Z^*(I)$ is the optimum solution value for instance I . The obvious question is: *Is it possible to approximate optimization problems, whose decision version is in NP within a certain guarantee $f(n)$?*

Recent progress in interactive proof systems [2,3] showed that there are classes of problems which cannot be approximated with a guarantee better than $f(n)$ unless $P = NP$.

In the last 25 years there have been several approximation algorithms proposed for NP-hard problems. Unfortunately, the proposed algorithms are problem specific. Typically, a particular approximation algorithm works for a specific problem and the insights gained from a successful analysis of a particular problem rarely transfer to another.

As a result, in our opinion, the following questions have not yet received satisfactory answers:

1. *What is the reason that certain problems can be approximated better than others?*
2. *Is there a systematic way to generate good (in particular best possible) approximation algorithms?*

In the present paper we attempt to provide some insights to these questions by focusing on general covering problems of the type:

$$\begin{array}{ll} \text{minimize} & cx \\ \text{subject to} & Ax \geq b, \\ & x \in X, \end{array} \quad (1)$$

where the entries in A, b, c are nonnegative integers and the set X is either $\{0, 1\}^n$ or Z_+^n . There is a very large collection of discrete optimization problems that have covering formulations (see Sections 2–5).

Our contributions in this paper are as follows.

1. We propose and analyze a randomized rounding heuristic with a nonlinear rounding function. With the exception of the work of Raghavan and Thompson [4], and Srivastav and Stangier [5,6] most applications of randomized rounding have been to problems in which all 0–1 solutions to an integer programming problem are feasible. The papers just mentioned used linear rounding functions to obtain approximation results for special cases of the set packing problem. We expand the power of the method by considering nonlinear rounding functions as well as applications of the method to problems, in which there is a probability that the rounded solution is infeasible. We apply the heuristic to set covering, facility location, network design and general covering problems. We show that our method matches the best possible approximation bounds for all these problems.
2. Motivated by the problem of covering cuts in a graph that encompasses several important graph problems (the minimum spanning tree, connectivity problems, the steiner tree problem, the matching problem, etc.) we identify an interesting connection between the analysis of approximation algorithms and random graph theory. In particular, we formulate a natural general conjecture that could have interesting applications in approximation algorithms.

3. We propose and analyze a generalization of a well known generic dual heuristic for general covering problems. Our analysis in this part of the paper illustrates the importance *to the design of a good approximation algorithm for a discrete optimization problem of a deep understanding of the structural properties of its LP relaxation*. Our method can be seen as an extension of the work of Goemans and Williamson [7], Williamson et al. [8] and Goemans et al. [9] in the approximability of cut covering problems to general covering problems.

The outline of the paper is as follows: In Section 2, we propose our first generic heuristic that uses randomization and apply it to the following problems: set covering, facility location and arbitrary covering problems. In Section 3, we propose our second generic heuristic that uses information from the LP dual problem and apply it to general covering problems obtaining the best known guarantee for the problem. In Section 4, we examine a special covering problem involving covering cuts in an undirected graph. We identify connections between the design of approximation algorithms for this problem and the theory of random graphs with nonuniform probabilities. As an application, we use the randomized heuristic in a particular network design problem. The last section contains some discussion of open problems.

2. Randomized rounding with nonlinear rounding functions

In this section we study discrete optimization problems for which we have an integer programming formulation:

$$\begin{aligned}
 (\text{IP}_1) \quad & \text{IZ}_1 = \text{minimize} && cx \\
 & \text{subject to} && Ax \geq b, \\
 & && x \in \{0, 1\}^n,
 \end{aligned} \tag{2}$$

where the entries in A, b, c are nonnegative integers. We consider its LP relaxation:

$$\begin{aligned}
 (\text{LP}_1) \quad & Z_1 = \text{minimize} && cx \\
 & \text{subject to} && Ax \geq b, \\
 & && x \in [0, 1]^n.
 \end{aligned}$$

Throughout the paper we denote with IZ (Z) the value of the optimal IP (LP) solution. We propose the following general randomized scheme to approximate Problem (IP_1) :

A generic randomized heuristic

1. Solve the LP relaxation (LP_1) obtaining the solution x^* .
2. Round the solution as follows: $P\{x_j = 1\} = f(x_j^*)$.
3. Use a derandomization procedure to find a deterministic approximation.

Randomized rounding heuristics were first proposed by Raghavan and Thompson [4] to approximate a VLSI design problem (a min max-type of problem). They used $f(x) = x$ to round. Yannakakis [10] interprets an algorithm of Johnson [11] for

MAXSAT as a randomized rounding heuristic with a linear rounding function ($f(x) = x$) and Goemans and Williamson [12] use nonlinear rounding functions $f(x)$ to round, matching the best known guarantee (obtained in [10] using maximum flow techniques) of $\frac{3}{4}$ for MAXSAT. The natural questions that arise in the context of the generic randomized heuristic are:

1. How does one choose the rounding function $f(x)$?
2. Is it possible to find an integer feasible solution always?
3. How does the cost of the obtained solution compare with that of the LP relaxation?
4. How does one derandomize?

Our plan is to address these issues with respect to particular examples.

2.1. The set covering problem

Given a set $N = \{1, \dots, n\}$, weights $c_j \geq 0$ for $j \in N$ and a family of m sets $S_i \subseteq N$, find a set $S \subseteq N$ such that $|S \cap S_i| \geq 1$, such that $\sum_{j \in S} c_j$ is minimized. In order to formulate the problem we let $a_{ij} = 1$ if $j \in S_i$. Then the set covering problem can be formulated as follows:

$$\begin{aligned}
 (\text{IP}_2) \quad & \text{IZ}_2 = \text{minimize} && \sum_{j=1}^n c_j x_j \\
 & \text{subject to} && \sum_{j=1}^n a_{ij} x_j \geq 1, \quad i = 1, \dots, m, \\
 & && x_j \in \{0, 1\}.
 \end{aligned}$$

Lund and Yannakakis [3] proved that the set covering problem cannot be approximated within a factor smaller than $O(\log m)$ unless $P = NP$, i.e., if there exists a polynomial time algorithm within $\frac{1}{4} \log m$ from the optimal value, then $P = NP$. Johnson [11] and Lovász [13] propose a greedy heuristic with value Z_G for the problem with $c_j = 1$, such that

$$\frac{Z_G}{\text{IZ}_2} \leq H(D) \leq \log D + 1,$$

where $H(D) = \sum_{i=1}^D 1/i$ and D is the maximum columns sum. Chvátal [14] extends the heuristic for the case of general $c_j \geq 0$ proving the same guarantee. Bronniman and Goodrich [15] describe an approximation algorithm with a bound that is a logarithmic function of the VC dimension of the constraint matrix $A = \{a_{ij}\}$.

Applying the randomized heuristic, we solve the LP relaxation and find the solution x_j^* with value Z_2 . We round as follows:

$$P\{x_j = 1\} = f(x_j^*) = 1 - (1 - x_j^*)^k, \quad k = \log D,$$

where $D = \max_j |D_j|$, with $D_j = \{i : j \in S_i\}$. The interpretation is that we flip a coin that has probability x_j^* of giving heads, k times. If, in any of these k flips the coin shows heads, we round x_j to one, otherwise we round to zero. Let x_H be the solution vector given by the heuristic (notice that the solution is not always feasible). Let Z_H be the cost of the proposed solution.

Theorem 1.

$$\frac{E[Z_H | x_H \text{ is feasible}]}{Z_2} \leq \frac{\log D}{(1 - 1/D)^D}.$$

In particular,

$$\frac{IZ_2}{Z_2} \leq \frac{\log D}{(1 - 1/D)^D}.$$

By derandomizing the randomized scheme we can find in polynomial time a feasible solution H such that

$$\frac{Z_H}{Z_2} \leq \frac{\log m}{(1 - 1/m)^m}.$$

Proof. Let V_i be the event that constraint i is violated by x_H . Let U_i be the event that constraint i is satisfied. Let $F = \bigcap_{i=1}^m U_i$ be the event that the solution x_H is feasible. Then

$$\begin{aligned} P\{V_i\} &= P\left\{\sum_{j=1}^n a_{ij}x_j = 0\right\} = P\left\{\sum_{j \in S_i} x_j = 0\right\} = \prod_{j \in S_i} P\{x_j = 0\} = \prod_{j \in S_i} (1 - x_j^*)^k \\ &< \prod_{j \in S_i} e^{-kx_j^*} = e^{-k\sum_{j \in S_i} x_j^*} = e^{-k\sum_{j=1}^n a_{ij}x_j^*} \leq e^{-k}. \end{aligned}$$

Notice also that

$$P\{V_i \cap V_j\} = \prod_{j \in S_i \setminus S_j} P\{x_j = 0\} \prod_{j \in S_j \setminus S_i} P\{x_j = 0\} \prod_{j \in S_i \cap S_j} P\{x_j = 0\} \geq P\{V_i\}P\{V_j\}.$$

Therefore,

$$\begin{aligned} P\{U_i \cap U_j\} &= 1 - P\{V_i\} - P\{V_j\} + P\{V_i \cap V_j\} \\ &\geq 1 - P\{V_i\} - P\{V_j\} + P\{V_i\}P\{V_j\} = P\{U_i\}P\{U_j\}, \end{aligned}$$

which is intuitively obvious, since $P\{U_i | U_j\} \geq P\{U_i\}$. In general,

$$P\left\{\bigcap_i U_i\right\} \geq \prod_i P\{U_i\}.$$

Moreover,

$$E[Z_H | F] = \sum_{j=1}^n c_j P\{x_j = 1 | F\} = \sum_{j=1}^n c_j \frac{P\{F | x_j = 1\}}{P\{F\}} P\{x_j = 1\}.$$

Let D_j^c be the complement of D_j . Then

$$\frac{P\{F | x_j = 1\}}{P\{F\}} = \frac{P\{\bigcap_{i \in D_j^c} U_i\}}{P\{\bigcap_{i=1}^m U_i\}} \leq \frac{1}{P\{\bigcap_{i \in D_j} U_i\}},$$

because $P\{\bigcap_{i=1}^m U_i\} \geq P\{\bigcap_{i \in D_j^c} U_i\}P\{\bigcap_{i \in D_j} U_i\}$. Since

$$P\left\{\bigcap_{i \in D_j} U_i\right\} \geq \prod_{i \in D_j} P\{U_i\} = \prod_{i \in D_j} (1 - P\{V_i\}) \geq (1 - e^{-k})^{|D_j|},$$

we obtain that

$$\begin{aligned} E[Z_H | F] &\leq \sum_{j=1}^n c_j (1 - e^{-k})^{-|D_j|} (1 - (1 - x_j^*)^k) \\ &\leq (1 - e^{-k})^{-D} \sum_{j=1}^n c_j (1 - (1 - kx_j^*)) \\ &= k(1 - e^{-k})^{-D} \sum_{j=1}^n c_j x_j^* = \frac{kZ_2}{(1 - e^{-k})^D}, \end{aligned}$$

where $D = \max_j |D_j|$. Choosing $k = \log D$, we obtain that

$$E[Z_H | F] \leq \frac{\log D}{(1 - \frac{1}{D})^D} Z_2 = O(\log D) Z_2.$$

This implies that there exists a feasible solution with cost at most $O(\log D) Z_2$, proving that:

$$\frac{IZ_2}{Z_2} \leq O(\log D).$$

We will now derandomize this randomized scheme to find a feasible solution with value within $O(\log m)$ from Z_2 . Notice that the guarantee is somewhat weaker as $D \leq m$. We introduce the following potential function:

$$\Phi(X_1, \dots, X_n) = \sum_{j=1}^n c_j X_j + M \sum_{i=1}^m 1_{\{A_i(\vec{X})\}},$$

where

$$1_{\{A_i(\vec{X})\}} = \begin{cases} 1, & \sum_{j=1}^n a_{ij} X_j = 0, \\ 0, & \text{otherwise} \end{cases}$$

and M is a large constant to be chosen later. Notice that if $\Phi(\vec{X}) = \sum_{j=1}^n c_j X_j \leq M$, \vec{X} is feasible, while if \vec{X} is infeasible $\Phi(\vec{X}) \geq \sum_{j=1}^n c_j X_j + M > M$. Therefore, the value of M distinguishes feasibility. If we select

$$X_j = \begin{cases} 1 & \text{with probability } 1 - (1 - x_j^*)^k, \\ 0 & \text{with probability } (1 - x_j^*)^k, \end{cases}$$

we obtain that

$$E[\Phi(\vec{X})] = \sum_{j=1}^n c_j (1 - (1 - x_j^*)^k) + M(1 - P\{F\}) \leq kZ_2 + M(1 - (1 - e^{-k})^m).$$

If we choose M so that $kZ_2 + M(1 - (1 - e^{-k})^m) \leq M$, then $E[\Phi(\vec{X})] \leq M$. We then use the method of conditional expectations (see [16, p. 223]), i.e., at the very first step we select $X_1 = 0$ if $E[\Phi(\vec{X}) | X_1 = 0] \leq E[\Phi(\vec{X}) | X_1 = 1]$; otherwise we select $X_1 = 1$. Let x_1, \dots, x_{r-1} be the values of the variables that the algorithm selected in the first $r - 1$ steps. At step r , we select $X_r = x_r$ in order to minimize over $x_r \in \{0, 1\}$ $E[\Phi(x_1, \dots, x_{r-1}, X_r, \dots, X_n) | X_r = x_r]$. In this way we obtain deterministically in polynomial time a solution x_H with $\Phi(x_H) \leq E[\Phi(\vec{X})] \leq M$, i.e., feasible. The cost of this solution is $Z_H = c'x_H = \Phi(x_H) \leq M$. Our goal is then to choose k, M in order to optimize the bounds, i.e., we solve:

$$\begin{aligned} & \text{minimize} && M \\ & \text{subject to} && kZ_2 + M(1 - (1 - e^{-k})^m) \leq M, \\ & && k, M \geq 0. \end{aligned} \tag{3}$$

Choosing $k = \log m$ and

$$M = \frac{\log m}{(1 - \frac{1}{m})^m} Z_2 = O(\log m)Z_2,$$

we can deterministically find, using the method of conditional expectations, a feasible solution x_H to the set covering problem with guarantee $Z_H/Z_2 = O(\log m)$. \square

Remarks.

1. If we choose $k = \log D + c$, we obtain that

$$\frac{IZ_2}{Z_2} \leq \frac{\log D + c}{(1 - e^{-c}/D)^D} \sim e^{e^{-c}}(\log D + c),$$

when D is sufficiently large. We can then optimize the bound by choosing the best c .

2. Another rounding function that also leads to an $O(\log m)$ approximation algorithm is the following. If $x_j^* \geq 1/\log m$, then set $x_j = 1$. If $x_j^* < 1/\log m$, then set $x_j = 1$ with probability $x_j \log m$. The proof that this leads to an $O(\log m)$ approximation proceeds along the same lines as before. We will next generalize this randomized rounding scheme to general covering problems.

2.2. *Constrained general covering problems*

We consider the problem

$$\begin{aligned} \text{(IP}_3\text{)} \quad & IZ_3 = \text{minimize} && \sum_{j=1}^n c_j x_j \\ & \text{subject to} && \sum_{j=1}^n a_{ij} x_j \geq b_i, \quad i = 1, \dots, m. \\ & && x_j \in \{0, 1\}, \end{aligned}$$

where a_{ij}, b_i, c_j are nonnegative integers. Notice that we can assume without loss of generality that $a_{ij} \leq b_i$ for all i, j , because otherwise we can replace a_{ij} with b_i and the optimal value IZ_3 will remain unchanged. Let Z_3 be the value of the LP relaxation. The best known approximation bounds for the problem are due to:

1. Hall and Hochbaum [17], generalizing earlier work of Hochbaum [18], Balas [19] and Bar-Yehuda and Even [20], who propose a dual heuristic H for the case that a_{ij} are restricted to be 0 or 1 such that $Z_H/Z_3 \leq f$, where $f = \max_i \sum_{j=1}^n a_{ij}$. We refer to this bound as *the row-sum bound*.
2. Dobson [21] and independently Fisher and Wolsey [22], generalizing earlier work of Johnson [11], Lovász [13] and Chvátal [14] for the set covering problem, who propose a greedy heuristic for the problem and prove that if $d = \max_{1 \leq j \leq n} \sum_{i=1}^m a_{ij}$, then $Z_G/Z_3 \leq H(d)$. If a_{\max} is the largest entry in matrix A , then the result is essentially $Z_G/Z_3 \leq O(\log ma_{\max})$. We refer to this bound as *the column-sum bound*.

In this section we show that the generic rounding heuristic with different rounding functions produces a stronger approximation than the row sum bound and a somewhat weaker version of the column sum bound. In particular, using deterministic rounding we first obtain a heuristic H with guarantee $Z_H/Z_3 \leq f$ for all covering problems (not restricted to $a_{ij} = 0$ or 1), thus generalizing all previous work regarding the row-sum bound (in Section 3 we also propose a generic rounding heuristic, but this time using dual information to achieve the same bound that has a much faster running time).

Moreover, using randomized rounding, we obtain a heuristic H with guarantee $Z_H/Z_3 \leq O(a_{\max} \log m)$, thus establishing a somewhat weaker version of the column-sum bound.

2.2.1. Deterministic rounding

We apply the generic rounding heuristic but with a deterministic rounding function as follows.

1. Solve the LP relaxation replacing $x_j \in \{0, 1\}$ with $0 \leq x_j \leq 1$. Let x^* be an optimal solution.
2. If $x_j^* \geq 1/f$, then set $x_j = 1$.
3. If $x_j^* < 1/f$, then set $x_j = 0$.

Hochbaum [18] proposes this heuristic for set covering problems. Let Z_H be the value of the above heuristic.

Theorem 2. *The rounding heuristic produces a feasible solution that satisfies*

$$\frac{Z_H}{Z_3} \leq f.$$

Proof. Let $A = \{j : x_j \geq 1/f\}$. Let $B = A^c$. Clearly $A \neq \emptyset$, because otherwise

$$\sum_j a_{ij} x_j^* < \frac{1}{f} \sum_j a_{ij} \leq 1 \leq b_i,$$

i.e., the LP solution would not be feasible. We first show that the rounding heuristic produces a feasible solution. Suppose not. Then there exists a row i such that $\sum_{j \in A} a_{ij} \leq b_i - 1$. Since the solution x^* is feasible

$$\begin{aligned}
 b_i &\leq \sum_{j \in A} a_{ij}x_j^* + \sum_{j \in B} a_{ij}x_j^* \leq b_i - 1 + \sum_{j \in B} a_{ij}x_j^* \\
 &< b_i - 1 + \frac{1}{f} \sum_{j \in B} a_{ij} \leq b_i,
 \end{aligned}$$

which is a contradiction. Moreover,

$$Z_3 \geq \sum_{j \in A} c_j x_j^* \geq \frac{1}{f} \sum_{j \in A} c_j = \frac{Z_H}{f}. \quad \square$$

2.2.2. Randomized rounding

We next apply the generic randomized heuristic with a rounding function similar to the one used in remark 2 in the end of the previous subsection as follows.

1. Solve the LP relaxation replacing $x_j \in \{0, 1\}$ with $0 \leq x_j \leq 1$. Let x^* be an optimal solution.
2. Let $k = 2a_{\max} \log m + 2$; $a_{\max} = \max_{i,j} a_{ij}$; let $R = \{j: kx_j^* \leq 1\}$.
3. If $j \notin R$, then $x_j = 1$.
 If $j \in R$, then

$$x_j = \begin{cases} 1 & \text{with probability } kx_j^*, \\ 0 & \text{with probability } 1 - kx_j^*. \end{cases}$$

In order to analyze the heuristic we need some tail estimates for sums of independent, but not identical random variables. Let $X_i, i = 1, \dots, n$ be independent random variables. X_i takes values r_{ij} with probability p_{ij} . Let $l_i = \min_j r_{ij}$, $u_i = \max_j r_{ij}$. Let $X = \sum_{i=1}^n X_i$.

Proposition 1.

$$P\{X < (1 - \delta)E[X]\} < \exp\left(-\frac{\delta^2(E[X])^2}{2 \sum_{i=1}^n E[(X_i - l_i)^2]}\right). \tag{4}$$

Proof. Let $t > 0$.

$$P\{X < (1 - \delta)E[X]\} = P\{e^{-tX} > e^{-t(1-\delta)E[X]}\} \leq \frac{E[e^{-tX}]}{e^{-t(1-\delta)E[X]}} \tag{5}$$

from Markov’s inequality. But,

$$\begin{aligned}
 E[e^{-tX}] &= \sum_j p_{ij} e^{-tr_{ij}} = e^{-tl_i} \left(1 - \sum_j p_{ij} (1 - e^{-t(r_{ij}-l_i)})\right) \\
 &\leq e^{-tl_i} \exp\left(-\sum_j p_{ij} (1 - e^{-t(r_{ij}-l_i)})\right),
 \end{aligned}$$

since $1 - x \leq e^{-x}$. Furthermore,

$$E[e^{-tX_i}] \leq e^{-tl_i} \exp\left(\sum_j p_{ij} \left(-t(r_{ij} - l_i) + \frac{t^2(r_{ij} - l_i)^2}{2}\right)\right),$$

since $e^{-x} \leq 1 - x + x^2/2$ for $x \geq 0$. Therefore,

$$E[e^{-tX_i}] \leq \exp\left(-tE[X_i] + \frac{t^2}{2}E[(X_i - l_i)^2]\right).$$

Substituting to (5) we obtain

$$P\{X < (1 - \delta)E[X]\} \leq \exp\left(-\delta tE[X] + \frac{t^2}{2}\sum_{i=1}^n E[(X_i - l_i)^2]\right).$$

Picking $t = \delta E[X] / \sum_{i=1}^n E[(X_i - l_i)^2]$ to obtain the sharpest upper bound, we prove (4). \square

Proposition 1, which appears to be new, generalizes the tail estimates for 0–1 random variables. In the case that X_i is a 0–1 random variable, $l_i = 0$, $u_i = 1$ and $E[X_i^2] = E[X_i]$, (4) reduces to the tail estimate obtained in [23]

$$P\{X < (1 - \delta)E[X]\} < e^{-\frac{\delta^2 E[X]}{2}}.$$

Similar bounds are to be found in [4].

We next analyze the generic randomized heuristic using Proposition 1.

Theorem 3. *The application of the generic randomized heuristic gives a feasible solution of value Z_H such that*

$$\frac{Z_H}{Z_3} \leq O(a_{\max} \log m).$$

Proof. We consider the set $R = \{j: kx_j^* \leq 1\}$ and its complement R^c . Let V_i be the event that constraint i is violated and $U_i = V_i^c$ the event that constraint i is satisfied. Our first goal is to find an upper bound on

$$P\{V_i\} = P\left\{\sum_{j \in R} a_{ij}x_j + \sum_{j \in R^c} a_{ij} < b_i\right\}.$$

If $\sum_{j \in R^c} a_{ij} \geq b_i$, then constraint i is satisfied. So, we assume that $\sum_{j \in R^c} a_{ij} \leq b_i - 1$. Let us first notice that

$$\sum_{j \in R} a_{ij}x_j^* \geq b_i - \sum_{j \in R^c} a_{ij} \geq 1.$$

Then,

$$P\{V_i\} = P\left\{\sum_{j \in R} a_{ij}x_j + \sum_{j \in R^c} a_{ij} < b_i\right\} = P\left\{\sum_{j \in R} a_{ij}x_j < b_i - \sum_{j \in R^c} a_{ij}\right\}$$

$$\leq P \left\{ \sum_{j \in R} a_{ij} x_j < \sum_{j \in R} a_{ij} x_j^* \right\},$$

as $b_i - \sum_{j \in R^c} a_{ij} \leq b_i - \sum_{j \in R^c} a_{ij} x_j^* \leq \sum_{j \in R} a_{ij} x_j^*$, because x^* is feasible. For $j \in R$, let $Y_j = a_{ij}$ with probability kx_j^* and 0 otherwise. Notice that $\sum_{j \in R} E[Y_j] = k \sum_{j \in R} a_{ij} x_j^*$. With $\delta = 1 - 1/k$

$$P\{V_i\} \leq P \left\{ \sum_{j \in R} Y_j < (1 - \delta) \sum_{j \in R} E[Y_j] \right\}.$$

Using (4) with $l_i = 0$ we obtain

$$P\{V_i\} < \exp \left(- \frac{(k-1)^2 (\sum_{j \in R} a_{ij} x_j^*)^2}{2k \sum_{j \in R} a_{ij}^2 x_j^*} \right) \leq e^{-\frac{(k-1)^2}{2ka_{\max}}},$$

since $\sum_{j \in R} a_{ij}^2 x_j^* \leq a_{\max} \sum_{j \in R} a_{ij} x_j^*$ and $\sum_{j \in R} a_{ij} x_j^* \geq 1$.

Let $F = \bigcap_{i=1}^m U_i$ be the event that the solution x_H is feasible. Then

$$P\{F\} = P \left\{ \bigcap_i U_i \right\} \geq \prod_i P\{U_i\},$$

since $P\{U_i | U_j\} \geq P\{U_i\}$. Thus

$$P\{F\} \geq \left(1 - e^{-\frac{(k-1)^2}{2ka_{\max}}} \right)^m.$$

Moreover,

$$E[Z_H] = k \sum_{j \in R} c_j x_j^* + \sum_{j \in R^c} c_j \leq k \left(\sum_{j \in R} c_j x_j^* + \sum_{j \in R^c} c_j x_j^* \right) = kZ_3.$$

Since $E[Z_H] \geq E[Z_H | F]P\{F\}$,

$$\frac{E[Z_H | F]}{Z_3} \leq k / \left(1 - e^{-\frac{(k-1)^2}{2ka_{\max}}} \right)^m \leq k / \left(1 - e^{-\frac{k-2}{2a_{\max}}} \right)^m.$$

For $k = 2a_{\max} \log m + 2$, we obtain

$$\frac{E[Z_H | F]}{Z_3} \leq (2a_{\max} \log m + 2) / \left(1 - \frac{1}{m} \right)^m = O(a_{\max} \log m).$$

By derandomizing as in the proof of Theorem 1 we can find deterministically a solution x_H that satisfies

$$\frac{Z_H}{Z_3} = O(a_{\max} \log m). \quad \square$$

Dobson [21] achieves a bound $O(\log \max_j \sum_i a_{ij})$ by analyzing a greedy heuristic. The bound in this case is $O(\log m + \log a_{\max})$ as opposed to our weaker bound $O(a_{\max} \log m)$. For a_{\max} constant both bounds are $O(\log m)$.

2.3. Facility location

We are given an undirected graph $G = (V, E)$ ($|V| = n$) with costs $c_{ij} \geq 0$ for $(i, j) \in E$ and $d_i \geq 0$ for each $i \in V$. The goal is to find a set $S \subseteq V$ of facilities that minimizes $\sum_{i \in S} d_i + \sum_{i \in V} \min_{j \in S} c_{ij}$. The problem is NP-hard [1]. Hochbaum [24] presents a greedy algorithm for the problem with cost within $O(\log n)$ of the optimum. Since the problem can be formulated as a set covering problem (this is a non-standard formulation) involving $O(n^2)$ rows, this is not a surprise. Here we show that the $O(\log n)$ can be obtained directly from the standard formulation of this problem.

In this section we show that the generic randomized heuristic returns a solution with cost within $O(\log n)$ of the optimum cost, i.e., it performs as well as any heuristic can (unless $P = NP$). We start with the classical strong integer programming formulation of the problem. Let $y_i = 1$ if we locate a facility at node $i \in V$, 0 otherwise. Let $x_{ij} = 1$ if customer $i \in V$ is assigned to facility j . Then, the problem can be formulated as follows.

$$\begin{aligned}
 \text{(IP}_4) \quad \text{IZ}_4 = \text{minimize} \quad & \sum_{(i,j) \in E} c_{ij}x_{ij} + \sum_{j \in V} d_j y_j \\
 \text{subject to} \quad & x_{ij} \leq y_j, \quad (i, j) \in E, \quad j \in V \\
 & \sum_{j=1}^n x_{ij} \geq 1, \quad i = 1, \dots, n, \\
 & x_{ij}, y_j \in \{0, 1\}.
 \end{aligned}$$

Let Z_4 be the value of the LP relaxation. The randomized heuristic is as follows:

1. Solve the LP relaxation and find the solution x_{ij}^*, y_j^* .
2. Round as follows:

$$P\{y_j = 1\} = 1 - (1 - y_j^*)^k, \quad k = \log n,$$

$$P\{x_{ij} = 1 \mid y_j = 1\} = \frac{1 - (1 - x_{ij}^*)^k}{1 - (1 - y_j^*)^k},$$

$$P\{x_{ij} = 1 \mid y_j = 0\} = 0.$$

In this way the constraints $x_{ij} \leq y_j$ are always satisfied.

Let x_H be the solution generated by this algorithm and Z_H its cost. Notice that

$$P\{x_{ij} = 1\} = 1 - (1 - x_{ij}^*)^k \quad \text{and} \quad P\left\{\bigcap_j x_{ij} = 1\right\} = \prod_j (1 - (1 - x_{ij}^*)^k).$$

We bound the performance of the heuristic as follows.

Theorem 4.

$$\frac{E[Z_H \mid x_H \text{ is feasible}]}{Z_4} \leq \frac{\log n}{(1 - 1/n)^n}.$$

By derandomizing the randomized scheme we can find in polynomial time a feasible solution H such that

$$\frac{Z_H}{Z_4} \leq \frac{\log n}{(1 - 1/n)^n}.$$

Proof. Let A_i be the event that customer i is not assigned after the randomization. Let $U_i = A_i^c$. Let $F = \bigcap_{i=1}^m U_i$ be the event that the solution x_H is feasible. Then

$$\begin{aligned} P\{A_i\} &= P\left\{\sum_{j=1}^n x_{ij} = 0\right\} = P\left\{\bigcap_{j \in V} x_{ij} = 0\right\} = \prod_{j \in V} P\{x_{ij} = 0\} = \prod_{j \in V} (1 - x_{ij}^*)^k \\ &< \prod_{j \in V} e^{-kx_{ij}^*} \leq e^{-k}. \end{aligned}$$

Therefore, as before

$$P\{F\} = P\left\{\bigcap_{i=1}^n U_i\right\} \geq \prod_{i=1}^n P\{U_i\} = \prod_{i=1}^n (1 - P\{A_i\}) \geq (1 - e^{-k})^n.$$

Moreover,

$$E[Z_H] = \sum_{(i,j) \in E} c_{ij}(1 - (1 - x_{ij}^*)^k) + \sum_{j \in V} d_j(1 - (1 - y_j^*)^k) \leq kZ_4.$$

Therefore,

$$E[Z_H] \geq E[Z_H | F]P\{F\} \geq E[Z_H | F](1 - e^{-k})^n,$$

which implies that

$$E[Z_H | F] \leq \frac{kZ_4}{(1 - e^{-k})^n}.$$

Choosing $k = \log n$, we obtain that

$$\frac{E[Z_H | F]}{Z_4} \leq \frac{\log n}{(1 - \frac{1}{n})^n} = O(\log n).$$

The derandomization works exactly as in Theorem 1. \square

In order to derandomize the above procedure, we introduce the following potential function:

$$\Phi(\vec{X}, \vec{Y}) = \sum_{(i,j) \in E} c_{ij}X_{ij} + \sum_{j \in V} d_jY_j + M \sum_{i=1}^n 1\{A_i(\vec{X})\},$$

where

$$1\{A_i(\vec{X})\} = \begin{cases} 1, & \sum_{j=1}^n X_{ij} = 0, \\ 0, & \text{otherwise,} \end{cases}$$

and M is a large constant to be chosen later. Proceeding in exactly the same way as in the set covering problem we obtain that if we pick $k = \log n + \log \delta$ and

$$M = \frac{\log n + \log \delta}{1 - \frac{1}{5}} Z_{LP},$$

we can deterministically find using the method of conditional expectations a feasible solution x_H, y_H to the facility location problem with guarantee $Z_H/Z_{LP} = O(\log n)$. \square

3. Deterministic rounding through duality theory

We refer again to the general covering problem

$$\begin{aligned} (\text{IP}_5) \quad & \text{IZ}_5 = \text{minimize} \quad cx \\ & \text{subject to} \quad Ax \geq b, \\ & \quad \quad \quad x \in X, \end{aligned}$$

where A, b, c have nonnegative integer entries and X is either Z_+^n or $\{0, 1\}^n$.

We describe in this section a generic rounding heuristic that uses information from the dual of the LP relaxation of the above problem. In order to describe the dual heuristic we examine first the simpler case $X = Z_+^n$ and consider the linear programming relaxation

$$\begin{aligned} (\text{LP}_5) \quad & Z_5 = \text{minimize} \quad c'x \\ & \text{subject to} \quad Ax \geq b, \\ & \quad \quad \quad x \geq 0, \end{aligned}$$

and its dual

$$\begin{aligned} (\text{D}_5) \quad & Z_5 = \text{maximize} \quad y'b \\ & \text{subject to} \quad yA \leq c, \\ & \quad \quad \quad y \geq 0. \end{aligned}$$

The dual heuristic is as follows.

A generic dual heuristic:

1. Find an optimal solution x^*, y^* of the LP relaxation (LP_5) and its dual (D_5) , respectively.
2. If the j th constraint in the dual problem is tight, i.e., $(y^*A)_j = c_j$, then $x_j = \lceil x_j^* \rceil$. Otherwise, $x_j = 0$.

Hochbaum [18] proposed the heuristic for the case of A having 0–1 entries and b, c are arbitrary nonnegative integers. To connect this generic heuristic with the randomized rounding heuristic of the previous section, let us observe that this dual heuristic can also be interpreted as a deterministic rounding heuristic as follows: If x^*, y^* is a pair of *complementary* optimal solutions of (LP_5) and (D_5) , then the heuristic is clearly equivalent to the following deterministic rounding heuristic:

1. Find an optimal solution x^*, y^* for problems (LP_5) and (D_5) .
2. Let $x_j = \lceil x_j^* \rceil$ be the solution proposed by the heuristic.

Critical to our analysis is an understanding of the structure of an optimal dual solution. In the dual heuristic we use a particular dual optimal solution that satisfies certain structural properties. In Sections 3.1 and 3.2 we apply the dual heuristic to unconstrained ($X = Z_+^n$) and constrained ($X = \{0, 1\}^n$) covering problems respectively. In Section 3.3 we show that the dual heuristic leads to an $O(n^2)$ approximation algorithm that works without solving the LP relaxation.

3.1. Unconstrained general covering problems

We first apply the generic dual heuristic for the case $X = Z_+^n$. Let $f = \max_i \sum_{j=1}^n a_{ij} \geq 2$, otherwise (IP₅) is trivial. We apply the generic dual heuristic to (IP₅) starting with an optimal solution x^*, y^* . Let Z_H be the value of the heuristic.

Theorem 5. *The dual heuristic produces a solution with*

$$\frac{Z_H}{Z_5} \leq f + 1.$$

Proof. Let x^* and y^* be an optimal primal–dual pair for (LP₅) and (D₅). The heuristic solution $x_j = \lceil x_j^* \rceil$ is feasible, since

$$\sum_{j=1}^n a_{ij} x_j = \sum_{j=1}^n a_{ij} \lceil x_j^* \rceil \geq \sum_{j=1}^n a_{ij} x_j^* \geq b_i.$$

Let $K = \{j : x_j^* > 0\}$. Then,

$$Z_H = \sum_{j \in K} c_j \lceil x_j^* \rceil = \sum_{j \in K} \left(\sum_{i=1}^m a_{ij} y_i^* \right) \lceil x_j^* \rceil,$$

since $c_j = \sum_{i=1}^m a_{ij} y_i^*$ for $j \in K$ by complementary slackness. Since $\lceil x_j \rceil \leq x_j^* + 1$ and $\sum_{j \in K} a_{ij} x_j^* > b_i$ implies that $y_i^* = 0$ by complementarity we obtain

$$\begin{aligned} Z_H &\leq \sum_{i=1}^m \left(b_i + \sum_{j \in K} a_{ij} \right) y_i^* = \sum_{i=1}^m b_i y_i^* + \sum_{i=1}^m \sum_{j \in K} a_{ij} y_i^* \\ &\leq Z_5 + f \sum_{i=1}^m y_i^* \leq Z_5 + f \sum_{i=1}^m b_i y_i^* = (f + 1)Z_5, \end{aligned}$$

since $b_i \geq 1$. \square

3.2. Constrained general covering problems

In this section we consider the case $X = \{0, 1\}^n$, which we call the constrained covering program. We denote by IZ_6 and by Z_6 the optimal objective function value of this integer program and its linear relaxation, respectively. We prove that if we apply the generic dual heuristic, with a particular dual optimal solution, we obtain a feasible solution with

$$Z_H \leq f Z_6,$$

where, as before, $f = \max_i \sum_{j=1}^n a_{ij}$. The analysis is considerably more complicated, but it leads to an $O(n^2)$ approximation algorithm. This algorithm can be seen as the generalization of the algorithms of Goemans and Williamson [7], Williamson et al. [8] and Goemans et al. [9] for the cut covering problem (see Section 4).

The linear programming dual of the problem is:

$$(D_6) \quad Z_6 = \text{maximize} \quad \sum_{i=1}^m b_i y_i - \sum_{j=1}^n z_j$$

$$\text{subject to} \quad \sum_{i=1}^m a_{ij} y_i - z_j \leq c_j, \quad j = 1, \dots, n,$$

$$y_i \geq 0, \quad z_j \geq 0.$$

If (y, z) is a dual feasible solution we let

$$K(y, z) = \{j : z_j > 0\},$$

$$A(y, z) = \left\{ i : \sum_{j \in K(y, z)} a_{ij} = b_i \right\},$$

$$B(y, z) = \left\{ i : \sum_{j \in K(y, z)} a_{ij} < b_i \right\},$$

$$C(y, z) = \left\{ i : \sum_{j \in K(y, z)} a_{ij} > b_i \right\}.$$

We first show that there exists an optimal dual solution with particular properties.

Proposition 2. *There exists an optimal dual (y^*, z^*) solution with the following two properties:*

1. *Property 1:*

$$\text{Either } z_j^* = 0 \text{ for all } j \text{ or } y_i^* = 0 \text{ for all } i \in A(y^*, z^*) \cup C(y^*, z^*). \quad (6)$$

2. *Property 2:*

$$\sum_{j \in K(y^*, z^*)} z_j^* \leq \sum_{i \in B(y^*, z^*)} (b_i - 1) y_i^*. \quad (7)$$

Proof. We prove (7) first. Let (y, z) be an optimal dual solution. Note that if $j \in K(y, z)$, then by complementarity, $x_j = 1$. From complementary slackness $y_i = 0$ for all $i \in C(y, z)$, since for all $i \in C(y, z)$,

$$\sum_{j=1}^n a_{ij} x_j \geq \sum_{j \in K(y, z)} a_{ij} x_j = \sum_{j \in K(y, z)} a_{ij} > b_i.$$

If $y_i = 0$ for all $i \in A(y, z)$ then (7) of the proposition holds with $y^* = y, z^* = z$. Suppose then, there exists a $p \in A(y, z)$ such that $y_p > 0$. Let

$$\theta = \min \left\{ y_p, \min_{j \in K(y,z): a_{pj} > 0} \frac{z_j}{a_{pj}} \right\}.$$

Notice that $\theta > 0$. We define a new dual solution (y', z') as follows:

$$y'_i = \begin{cases} y_i, & i \neq p, \\ y_i - \theta, & i = p, \end{cases} \quad z'_j = \begin{cases} z_j, & j \notin K(y, z), \\ z_j - a_{pj}\theta, & j \in K(y, z). \end{cases}$$

The process by which we obtain (y', z') from (y, z) we call *reduction*. We show first that the solution (y', z') is dual feasible:

By the definition of θ , $y'_i \geq 0$, $z'_j \geq 0$. Moreover, if $j \in K(y, z)$

$$\sum_{i=1}^m a_{ij}y'_i - z'_j = \sum_{i=1}^m a_{ij}y_i - a_{pj}\theta - z_j + a_{pj}\theta = \sum_{i=1}^m a_{ij}y_i - z_j \leq c_j.$$

If $j \notin K(y, z)$, then,

$$\sum_{i=1}^m a_{ij}y'_i - z'_j = \sum_{i=1}^m a_{ij}y_i - a_{pj}\theta - z_j \leq c_j.$$

The new dual feasible solution (y', z') is still optimal, because

$$\sum_{i=1}^m b_i y'_i - \sum_{j=1}^n z'_j = \sum_{i=1}^m b_i y_i - b_p \theta - \sum_{j=1}^n z_j + \sum_{j \in K(y,z)} a_{pj} \theta = \sum_{i=1}^m b_i y_i - \sum_{j=1}^n z_j,$$

since $p \in A(y, z)$.

Each time we perform a reduction operation either $y'_p = 0$ or $z'_j = 0$ for some j such that $a_{pj} > 0$. Therefore, by repeating the reduction operation either we find an optimal dual solution that has $z_j^* = 0$ for all j or $y_i^* = 0$ for all $i \in A(y^*, z^*)$. In addition, since (y^*, z^*) is optimal, it should satisfy complementary slackness, i.e., $y_i^* = 0$ for all $i \in C(y^*, z^*)$. Therefore, we have proved (6).

We now proceed to prove (7) by showing

$$\sum_{j \in K(y^*, z^*)} z_j^* \leq \sum_{i \in B(y^*, z^*)} \sum_{j \in K(y^*, z^*)} a_{ij} y_i^*.$$

Notice that the above inequality implies (7), since for each $i \in B(y^*, z^*)$, $\sum_{j \in K(y^*, z^*)} a_{ij} \leq b_i - 1$.

If $y^* = 0$, $z^* = 0$ then (7) is satisfied. Since $y^* = 0$ and $z_j^* > 0$ cannot be optimal in the dual (notice that $y = 0, z = 0$ is feasible in the dual), there exists a $y_p^* > 0$ for some $p \in B(y^*, z^*)$. We apply a reduction operation again, i.e., construct a dual solution:

$$y'_i = \begin{cases} y_i^*, & i \neq p, \\ y_i^* - \theta, & i = p, \end{cases} \quad z'_j = \begin{cases} z_j^*, & j \notin K(y^*, z^*), \\ z_j^* - a_{pj}\theta, & j \in K(y^*, z^*), \end{cases}$$

where $\theta = \min(y_p^*, \min_{j \in K(y^*, z^*): a_{pj} > 0} z_j^*/a_{pj}) > 0$. We first observe that, as before, the new solution is dual feasible (the only difference with the previous derivation, how-

ever, is that the new solution is not necessarily optimal, since $p \in B(y^*, z^*)$ rather than $p \in A(y^*, z^*)$.

Given a dual feasible solution (y, z) we define

$$L(y, z) = \sum_{j \in K(y^*, z^*)} z_j \quad \text{and} \quad R(y, z) = \sum_{i \in B(y^*, z^*)} \sum_{j \in K(y^*, z^*)} a_{ij} y_i.$$

We then want to show that $L(y^*, z^*) \leq R(y^*, z^*)$. The second important observation is that after a reduction operation this inequality is preserved, i.e.,

$$L(y^*, z^*) \leq R(y^*, z^*) \quad \text{if and only if} \quad L(y', z') \leq R(y', z').$$

The reason is that

$$L(y', z') = \sum_{j \in K(y^*, z^*)} z'_j = \sum_{j \in K(y^*, z^*)} (z_j^* - a_{pj} \theta) = \sum_{j \in K(y^*, z^*)} z_j^* - \theta \sum_{j \in K(y^*, z^*)} a_{pj}.$$

Moreover,

$$R(y', z') = \sum_{i \in B(y^*, z^*)} \sum_{j \in K(y^*, z^*)} a_{ij} y_i^* - \theta \sum_{j \in K(y^*, z^*)} a_{pj}.$$

Therefore, since both the left and the right hand side of the inequality decrease by the same amount the inequality is preserved.

Each time we perform a reduction operation either $y'_p = 0$ or $z'_j = 0$ for some j such that $a_{pj} > 0$. Therefore, by repeating the reduction operation t times, either we find a dual solution that has $z'_j = 0$ for all j or $y'_i = 0$ for all i . In the first case ($z'_j = 0$ for all j), clearly $0 = L(y', z') \leq R(y', z')$ and therefore, $L(y^*, z^*) \leq R(y^*, z^*)$.

In the second case ($y'_i = 0$ for all i) we show that the solution $(y^*, z^* - z')$ is also dual feasible contradicting the optimality of (y^*, z^*) .

Observe that the r th time we perform the reduction operation the solution is updated as follows: ($y^0 = y^*, z^0 = z^*$); $y^r = y^{r-1} - v^{r-1}$ and $z^r = z^{r-1} - w^{r-1}$, where

$$v_i^r = \begin{cases} 0, & i \neq p_r, \\ y_i^* - \theta^r, & i = p_r, \end{cases} \quad w_j^r = \begin{cases} 0, & j \notin K(y^{r-1}, z^{r-1}), \\ a_{p_r, j} \theta^r, & j \in K(y^{r-1}, z^{r-1}). \end{cases}$$

If $y^t = 0$, then $y^0 = \sum_{r=0}^{t-1} v^r$ and $z^0 - z^t = \sum_{r=0}^{t-1} w^r$.

For $j \notin K(y^t, z^t)$, $z'_j = 0$ and then

$$\sum_{i=1}^m a_{ij} y_i^0 - (z_j^0 - z_j^t) \leq c_j,$$

since (y^0, z^0) is dual feasible.

For $j \in K(y^t, z^t)$,

$$\sum_{i=1}^m a_{ij} y_i^0 - (z_j^0 - z_j^t) = \sum_{r=0}^{t-1} \left[\sum_{i=1}^m a_{ij} v_i^r - w_j^r \right] = \sum_{r=0}^{t-1} (a_{p_r, j} \theta^r - a_{p_r, j} \theta^r) = 0 \leq c_j,$$

since $j \in K(y^t, z^t)$, then $j \in K(y^r, z^r)$ for $r \leq t - 1$.

Therefore, this case cannot happen, otherwise (y^*, z^*) is not optimal. This proves (7). \square

Notice, that the proof above is constructive. Starting with an arbitrary optimal dual solution found by a polynomial LP algorithm, we apply the reduction operation at most $\min(m, n)$ times to arrive at a dual solution that satisfies the properties (6) and (7). We then apply the generic dual heuristic to this particular optimal dual solution to obtain a feasible solution x_H with value Z_H .

Theorem 6. *If we run the generic dual heuristic with an optimal dual solution (y^*, z^*) satisfying (6) and (7),*

$$Z_H \leq f Z_6.$$

Proof. We showed in Proposition 2 that there exists an optimal dual solution (y^*, z^*) satisfying (6) and (7). Moreover, we can construct such a solution in polynomial time starting with an arbitrary optimal dual solution. Under the generic dual heuristic

$$x_j = \begin{cases} 1 & \text{if } \sum_{i=1}^m a_{ij}y_i^* - z_j^* = c_j, \\ 0 & \text{if } \sum_{i=1}^m a_{ij}y_i^* - z_j^* < c_j. \end{cases}$$

If x^* is an optimal solution to the LP relaxation, which is complementary to (y^*, z^*) , then

$$\sum_{j=1}^n a_{ij}x_j \geq \sum_{j=1}^n a_{ij}x_j^* \geq b_i,$$

and therefore, the solution x_H is feasible.

Let $K = \{j : \sum_{i=1}^m a_{ij}y_i^* - z_j^* = c_j\}$. The value of the heuristic is

$$Z_H = \sum_{j \in K} c_j = \sum_{j \in K} \sum_{i=1}^m a_{ij}y_i^* - \sum_{j \in K} z_j^*.$$

If $z_j^* = 0$ for all j , then

$$Z_H = \sum_{j \in K} \sum_{i=1}^m a_{ij}y_i^* \leq f \sum_{i=1}^m y_i^* \leq f \sum_{i=1}^m b_i y_i^* = f Z_6.$$

If $y_i^* = 0$ for all $i \in A(y^*, z^*) \cup C(y^*, z^*)$ and $\sum_{j \in K(y^*, z^*)} z_j^* \leq \sum_{i \in B(y^*, z^*)} (b_i - 1)y_i^*$, then

$$\begin{aligned} Z_H &= \sum_{j \in K} \sum_{i \in B(y^*, z^*)} a_{ij}y_i^* - \sum_{j \in K(y^*, z^*)} z_j^* \leq f \sum_{i \in B(y^*, z^*)} y_i^* \\ &\leq f \left(\sum_{i=1}^m b_i y_i^* - \sum_{j \in K(y^*, z^*)} z_j^* \right) = f Z_6. \end{aligned}$$

In both cases $Z_H \leq f Z_6$. \square

3.3. *Algorithmic implications of the generic dual heuristic*

The generic dual heuristic does not provide a better guarantee than the deterministic rounding heuristic of Section 2.2. Moreover, the analysis is considerably more complicated. The importance of the dual heuristic lies not so much in the guarantee it produces but rather in its proof. It suggests how a fast approximation algorithm for the constrained covering problem with performance guarantee f can be designed. Specifically, it is sufficient to construct a dual feasible solution (y, z) with the following properties:

1. The solution obtained by letting $x_j = 1$ when $\sum_{i=1}^m a_{ij}y_i - z_j = c_j$ is primal feasible.
2. $\sum_{j=1}^n z_j \leq \sum_{i=1}^m (b_i - 1)y_i$.

The proof of Proposition 2 hints at how such a dual solution is to be constructed; by working the reduction process in reverse. We now describe an $O(n^2)$ approximation algorithm designed along these lines. Note that Hall and Hochbaum [17] propose an $O(n^2)$ approximation algorithm for the special case of the problem with $a_{ij} = 0, 1$. Our algorithm is different.

The algorithm works in phases (very much in the spirit of Williamson et al. [8]). The output of phase t is a dual feasible solution (y^t, z^t) and a vector x^t . Each phase consists of a greedy type set covering algorithm. In the first phase, a primal solution, x^1 that covers every row at least once is generated. Any column j such that $x_j^1 = 1$ is deleted as well as any row i such that $\sum_j a_{ij}x_j^1 \geq b_i$. This ensures that no column is ever selected twice and constraints that are satisfied are ignored. In the second phase we generate a primal solution x^2 using the remaining columns that covers all remaining rows at least once. At the end of this phase we delete all columns j such that $x_j^2 = 1$ and all rows i such that $\sum_j a_{ij}x_j^2 \geq b_i - \sum_j a_{ij}x_j^1$. In general, at the end of the t th phase we generate a solution x^t such that:

$$\sum_{j=1}^n a_{ij}x_j^t \geq 1 \quad \text{if and only if} \quad b_i - \sum_{p=1}^{t-1} \sum_{j=1}^n a_{ij}x_j^p \geq 1.$$

Hence, if the algorithm terminates after k phases, the vector $x = \sum_{p=1}^k x^p$ will be primal feasible.

The input to phase $t + 1$ is:

1. $I(t) = \left\{ i : b_i - \sum_{p=1}^t \sum_{j=1}^n a_{ij}x_j^p \geq 1 \right\}$ (the set of rows not yet completely covered).
2. $J(t) = \left\{ j : x_j^p = 0 \forall p \leq t \right\}$ (the set of variables not yet set to 1, i.e., the remaining columns).
3. (y^t, z^t) .

Given $I(t)$, $J(t)$ and (y^t, z^t) from phase t , then phase $t + 1$ is as follows.

PHASE $t + 1$

Step 1: For all $j \in J(t)$ (set of columns to be picked in this phase) set $S_j = \{ i : i \in I(t), a_{ij} > 0 \}$; $g(S_j) = \sum_{i \in S_j} a_{ij}$; $K^{t+1} = \emptyset$; $y_i = 0, i \in I(t)$; $w_j = c_j - \sum_{i=1}^m a_{ij}y_i^t, j \in J(t)$.

Step 2: Let $\theta = \min_{j \in J(t)} w_j / g(S_j)$; $r = \operatorname{argmin}_{j \in J(t)} w_j / g(S_j)$ (select the column with smallest reduced cost to column sum ratio).

Step 3: Set $K^{t+1} = K^{t+1} \cup \{r\}$; $y_i = \theta$, $i \in S_r$ (selection of y_i in this manner ensures that $\sum_{i \in S_r} a_{ij} y_i = w_r$).

Step 4: Set $w_j = w_j - \sum_{i \in I(t)} a_{ij} y_i$, $j \in J(t)$ (update the reduced cost of column j); $S_j = S_j \setminus S_r$ (remove rows that are covered; updating the reduced costs in this way and the choice of θ ensure that at termination, $\sum_{i \in S_j} a_{ij} y_i \leq w_j \forall j \in J(t)$).

Step 5: If $\bigcup_{j \in K^{t+1}} S_j = I(t)$, (every remaining row covered at least once) go to step 6. Otherwise go to step 2.

Step 6: If $j \in K^{t+1}$, set $x_j^{t+1} = 1$; otherwise $x_j^{t+1} = 0$; $y_i^{t+1} = y_i^t + y_i$, $i \in I(t)$; $z_j^{t+1} = \max(\sum_{i=1}^m a_{ij} y_i^{t+1} - c_j, 0)$, $j = 1, \dots, n$.

Notice that the algorithm will terminate after at most n phases, i.e., every primal variable is set to 1. Let the algorithm terminate after phase p . From the last part of step 6, (y^p, z^p) will be dual feasible. We show next that the vector $x = \sum_{t=1}^p x^t$ satisfies the complementary slackness condition, $x_k = 1$ if $\sum_{i=1}^m a_{ik} y_i^p - z_k^p = c_k$. Now $x_k = 1$ implies that at the end of some phase, $t + 1$, say, $x_k^{t+1} = 1$. Observe that at end of phase $t + 1$:

$$\sum_{i \in S_j} a_{ij} y_i \leq w_j, \quad j \in J(t), \quad \sum_{i \in S_j} a_{ij} y_i = w_j, \quad j \in K^{t+1},$$

by steps 2 and 4. So, $x_j^{t+1} = 1$ only if $\sum_{i \in S_k} a_{ij} y_i = w_k$. By the definition of w_j and y^{t+1} , this implies that $x_j^{t+1} = 1$ only if $\sum_{i=1}^m a_{ij} y_i^{t+1} = c_j$. Hence, if $x_j^{t+1} = 1$, constraint k of the dual is tight with respect to (y^{t+1}, z^{t+1}) . From Step 6, once a constraint becomes tight, it remains tight, i.e., $\sum_{i=1}^m a_{ik} y_i^r - z_k^r = c_k$, $r \geq t + 1$.

Finally we use induction to show that

$$\sum_{j=1}^n z_j^p \leq \sum_{i=1}^m (b_i - 1) y_i^p.$$

At the end of phase 1, $z_j^1 = 0$. So, $\sum_{j=1}^n z_j^1 \leq \sum_{i=1}^m (b_i - 1) y_i^1$.

Suppose that $\sum_{j=1}^n z_j^t \leq \sum_{i=1}^m (b_i - 1) y_i^t$. Notice that if $z_j^{t+1} > 0$, then

$$0 < z_j^{t+1} = \sum_{i=1}^m a_{ij} y_i^{t+1} - c_j = \sum_{i=1}^m a_{ij} y_i^{t+1} - c_j = \sum_{i \in I(t)} a_{ij} y_i - \left(c_j - \sum_{i=1}^m a_{ij} y_i^t \right).$$

Hence from step 6, $j \in \bigcup_{r \leq t} K^r$, (i.e., $x_j^r = 1$ for some $r \leq t$); because if $j \in J(t)$, then $\sum_{i \in I(t)} a_{ij} y_i - (c_j - \sum_{i=1}^m a_{ij} y_i^t) = \sum_{i \in I(t)} a_{ij} y_i - w_j \leq 0$ contradicting $z_j^{t+1} > 0$. Hence,

$$\begin{aligned} \sum_{j=1}^n z_j^{t+1} &= \sum_{j \in \bigcup_{r \leq t} K^r} z_j^{t+1} \leq \sum_{j \in \bigcup_{r \leq t} K^r} \sum_{i \in I(t)} a_{ij} y_i + \sum_{j \in \bigcup_{r \leq t} K^r} z_j^t \\ &\leq \sum_{j \in \bigcup_{r \leq t} K^r} \sum_{i \in I(t)} a_{ij} y_i + \sum_{i=1}^m (b_i - 1) y_i^t \end{aligned}$$

$$\leq \sum_{i \in I(t)} (b_i - 1)y_i + \sum_{i=1}^m (b_i - 1)y_i^t = \sum_{i=1}^m (b_i - 1)y_i^{t+1}.$$

Remark. The primal rounding heuristic of Section 2.2 produces a much shorter proof of the bound. The heuristic, however, depends on the solution of the LP relaxation. In contrast the dual heuristic leads to an $O(n^2)$ non LP-based algorithm.

4. Cut covering problems

Let $G = (V, E)$, $|V| = n$ be an undirected graph. Let $f: 2^V \rightarrow Z_+$ be a given set function. Let $\delta(S) = \{e = (i, j) \in E \mid i \in S, j \in V \setminus S\}$. In this section we consider the following family of discrete optimization problems, which belongs to the class of covering problems:

$$\begin{aligned} \text{(IP}_f\text{)} \quad \text{IZ}_f = & \text{minimize} \quad \sum_{e \in E} c_e x_e \\ & \text{subject to} \quad \sum_{e \in \delta(S)} x_e \geq f(S) \quad \forall S, \\ & \quad \quad \quad x_e \in X, \end{aligned}$$

where X is either $\{0, 1\}$ or Z_+ . In Table 1 we list classical combinatorial problems formulated using the cutset formulation (IP_f) (see [7]).

Let Z_f be the value of the LP relaxation. We first note that if the set function f is arbitrary, then the guarantees of the previous cannot be further improved. Even if we restrict our attention to set functions f taking values in $\{0, 1\}$, $\text{IZ}_f/Z_f = \Omega(n)$ as can be seen from the following example: $f(S) = 1$ for all S such that $|S| = n/2$ and

Table 1
Some classical problems encompassed using a cutset formulation

Problem	$f(S)$
Spanning tree	1 for all $S \neq \emptyset, V$
Shortest path	$1, S \cap \{s, t\} = 1$
Steiner tree	1 if $S \cap T \neq \emptyset, T$
Generalized Steiner tree	1 if $S \cap T_i \neq \emptyset, T_i, i = 1, \dots, k$
Nonbipartite matching	1 if $ S = 2k + 1$
T -join	1 if $ S \cap T = 2k + 1$
b -matching problem ($\sum_{i \in V} b(i) = 2k$)	$\begin{cases} 1 & S \geq 2, \quad \sum_{i \in S} b(i) = 2k + 1, \\ b(i) & S = \{i\} \cup V \setminus \{i\}. \end{cases}$
Network survivability	$\max_{e \in \delta(S)} r_e, r_e \geq 0$
k -connected graph	k for all $s \neq \emptyset, V$
Tree partitioning	1, if $ S \leq k, S \geq n - k$
Capacitated tree	$\begin{cases} 2 \frac{\sum_{i \in T} d_i}{Q} & S \subset V, \\ 2 \frac{\sum_{i \in T} d_i}{Q} & 0 \in S \end{cases}$

$c_e = 1$. Then since $x_e = 4/n^2$ is a feasible solution to the LP relaxation $Z_f \leq (4/n^2) (n(n-1)/2) \leq 2$. However, every integral solution should have at least $n/2$ edges implying that $IZ_f \geq n/2$, and therefore the ratio grows linearly with n . Notice that a bound of $O(n)$ follows immediately from the column sum bound, ($O(\log 2^n)$).

If we impose, however, some conditions on f we can approximate (IP_f) significantly better. In particular if f satisfies $f(\emptyset) = 0$, $f(S) = f(V - S)$ and $f(A \cup B) \leq \max\{f(A), f(B)\}$, when $A \cap B = \emptyset$ (proper functions) Goemans and Williamson [7] propose a dual heuristic with value Z_H such that

$$\frac{Z_H}{Z_f} \leq 2 \left(1 - \frac{1}{|A|} \right),$$

where $A = \{i \in V : f(\{i\}) = 1\}$. These results are further generalized in [8,9] for more general set functions. For f being proper, taking arbitrary integer values and $X = \{0, 1\}$, the bound is extended in [9] to

$$\frac{Z_H}{Z_f} \leq 2 \sum_{i=1}^{f_{\max}} \frac{1}{i} = O(\log f_{\max}).$$

The technique used in these papers is the construction of a dual feasible solution that in addition to the properties in Proposition 2 has further special properties that follow from the particular properties of f and then the use of the idea of the dual heuristic of the previous section (when the dual constraint is tight the corresponding primal variable is set to 1). In this way these results can be seen as an application of the dual heuristic.

A natural idea is to apply the randomized heuristic of Section 2. Our goal is to show that the application of the randomized heuristic gives rise: (a) to an interesting, we believe, extension of classical random graph theory and (b) to a connection of random graphs to approximability.

Table 2
Thresholds properties for the uniform random graph model

Graph property	critical p
Connectivity	$\frac{\log n + c_n}{n}$
Graph has a matching	$\frac{\log n + c_n}{2n}$
Hamiltonicity	$\frac{\log n + \log \log n + 2c_n}{n}$
k -connectivity	$\frac{\log n + (k - 1) \log \log n + c_n}{n}$

4.1. Potential applications of random graphs in approximation algorithms

The theory of random graphs has developed independently from the study of approximation algorithms in trying to understand the properties of randomly generated graphs. The most widely used model (see for example [25]) involves a graph $G_{n,p}$ on n nodes, where each edge is present independently with the same probability p . The typical theorem proved under this model (for example [25]) is of the threshold type:

Let A be a graph property; let $p = (g(n) + c_n)/n$.

$$\lim_{n \rightarrow \infty} P\{G_{n,p} \text{ has property } A\} = \begin{cases} 1, & c_n \rightarrow \infty, \\ e^{-e^{-c}}, & c_n \rightarrow c, \\ 0, & c_n \rightarrow -\infty. \end{cases}$$

As examples, we list some important properties together with the corresponding critical probabilities p .

Let us attempt to apply the generic randomized heuristic for the class of problems (IP_f). We consider the LP relaxation and find the optimal solution x^* . Let

$$p_e = P\{x_e = 1\} = 1 - (1 - x_e)^{k_f},$$

where k_f will be chosen later. In this way we obtain a random graph, in which each edge e has a probability p_e of being present. The properties of the random graph under this *nonuniform* random model have not been extensively studied in the theory of random graphs. We conjecture an intimate connection between properties of random graphs and approximation algorithms.

Conjecture 1. *If a graph property A can be modeled using a cutset formulation (IP_f) for some set function f and there is a critical probability $p = (g(n) + c_n)/n$ for which the random graph $G_{n,p}$ (under the uniform model) has property A , then the application of the generic randomized heuristic with $P\{x_e = 1\} = 1 - (1 - x_e)^{g(n)}$ gives a solution x_H such that*

$$\frac{Z_H}{Z_f} \leq O(g(n)).$$

Alon [26] proved the conjecture for $f(S) = 1$, i.e., for the minimum spanning tree. We use his result to propose an $O(\log n)$ algorithm for a network design problem.

For example, suppose we solve the LP relaxation of the k -connected problem ($f(S) = k$). The corresponding critical probability is $p = (\log n + (k - 1) \log \log n + c_n)/n$. Then, the conjecture would imply that there exists a solution Z_H such that

$$\frac{Z_H}{Z_f} \leq O(\log n).$$

We next use the result of Alon [26] result to propose a $\log n$ approximation algorithm for a special network design problem.

4.2. A randomized approximation for a network design problem

In this section we consider the application of the generic randomized heuristic to a generic network design problem, that can be viewed as a variation of the shared fixed cost problem first considered by Rhys [27]. We are given a graph $G = (V, E)$ and we want to select edges at cost d_e , so that we can construct m trees each with costs c_e^k , where k denotes the k tree ($k = 1, \dots, m$). The problem can be formulated as follows.

$$\begin{aligned} \text{IZ}_7 = \text{minimize} \quad & \sum_{e \in E} c_e^k x_e^k + \sum_{e \in E} d_e y_e \\ \text{subject to} \quad & \sum_{e \in \delta(S)} x_e^k \geq 1, \quad k = 1, \dots, m, \quad S \subseteq V, \\ & x_e^k \leq y_e, \quad e \in E, \quad k = 1, \dots, m, \quad x_e, y_e \in \{0, 1\}. \end{aligned}$$

We apply the randomized heuristic as follows:

1. Solve the LP relaxation. Let x^*, y^* be the optimal solution.
2. Let $k = \log n$.

$$\begin{aligned} P\{y_e = 1\} &= 1 - (1 - y_e^*)^k; \\ P\{x_e^k = 1 \mid y_e = 1\} &= \frac{1 - (1 - x_e^*)^k}{1 - (1 - y_e^*)^k}; \\ P\{x_e^k = 1 \mid y_e = 0\} &= 0. \end{aligned}$$

Alon [26] proved Conjecture 1, for the case of $f(S) = 1$, i.e., he showed that if x^* is a feasible solution to P_f with $f(S) = 1$ and $P\{x_e = 1\} = 1 - (1 - x_e^*)^{b \log n}$, then with probability at least $1 - 1/n^b$ the graph is connected. Using this result, it is immediate that the value Z_H returned by the randomized heuristic satisfies:

$$\begin{aligned} \frac{E[Z_H \mid F]}{Z_7} &= O(\log n), \\ P\{F\} &\geq 1 - \frac{k}{n^b}. \end{aligned}$$

5. Concluding remarks

We presented two methods to construct approximation algorithms for covering problems:

- Randomized rounding with nonlinear rounding functions and
- Deterministic rounding using dual information.

We saw that these two approximation methods match the best known bounds for several covering problems.

Related to the question that motivated our research, whether there exists a systematic way to construct approximation algorithms, we believe that these two algo-

rithmic ideas provide a unification of methods for approximating discrete optimization problems.

We close the paper with a summary of research directions-conjectures that, we believe, will enhance our understanding of approximation methods:

1. Related to the question of which problems can be approximated better than others, we believe that as far as min–sum (as opposed to min–max) problems, we believe that only covering problems can be approximated within a $O(\log n)$ factor. Although this is not a formal mathematical statement we do not know of any example of a min–sum problem with a logarithmic or a sublogarithmic guarantee that cannot be formulated as a covering problem.
2. The randomized rounding heuristic uses the nonlinear rounding function $f(x)$. While we specified several rounding functions, we did not propose a systematic method to construct the rounding function. The problem of finding the best rounding function $f(x)$ in order to minimize $E[Z_H(f)]$ reduces to a calculus of variations problem, which seems to be difficult to solve at its full generality.
3. Investigation of Conjecture 1 seems interesting as an extension of random graph theory; a first step in this direction was taken by Alon [26]. Applications to approximability could also be interesting to explore.

References

- [1] M. Garey, D. Johnson, *Computers and Intractability: A Guide to the Theory of NP-completeness*, 1979, Freeman, San Francisco, CA.
- [2] S. Arora, C. Lund, R. Motwani, M. Sudan, M. Szegedy, Proof verification and intractability of approximation problems, 33rd IEEE Symposium on Foundations of Computer Science, 1992, pp. 14–23.
- [3] C. Lund, M. Yannakakis, On the hardness of approximating minimization problems, 33rd IEEE Symposium on Foundations of Computer Science, 1992, pp. 960–981.
- [4] P. Raghavan, C. Thompson, Randomized rounding: A technique for provably good algorithms and algorithmic proofs, *Combinatorica* 7 (1987) 365–374.
- [5] A. Srivastav, P. Stangier, Integer multicommodity flows with reduced demands, *Proceedings of the First European Symposium on Algorithms*, Lecture Notes in Computer Science, no. 726, Springer, Berlin, 1993, pp. 360–372.
- [6] A. Srivastav, P. Stangier, Tight approximations of resource constrained scheduling problems, *Proceedings of the Second European Symposium on Algorithms*, Lecture Notes in Computer Science, no. 855, Springer, Berlin, 1994, pp. 307–318.
- [7] M. Goemans, D. Williamson, (1992). A general approximation technique for constrained forest problems, *Proceedings of the Third ACM-SIAM Symposium on Discrete Algorithms*, 1992, pp. 307–316.
- [8] D. Williamson, M. Goemans, M. Mihail, V. Vazirani, A primal-dual approximation algorithm for generalized steiner network problems, *Proceedings of the 25th ACM Symposium on the Theory of Computing*, 1993, pp. 708–717.
- [9] M.X. Goemans, A.V. Goldberg, S. Plotkin, D.B. Shmoys, E. Tardos, D.P. Williamson, Improved approximation algorithms for network design problems, *Proceedings of the fifth ACM-SIAM Symposium of Discrete Algorithms*, (1994) pp. 223–232.
- [10] M. Yannakakis, On the Approximation of Maximum Satisfiability, *Proceedings of the Third ACM-SIAM Symposium on Discrete Algorithms*, 1992, pp. 1–9.

- [11] D. Johnson, Approximation algorithms for combinatorial problems, *Journal of computer and System Science* 9 (1974) 256–298.
- [12] M. Goemans, D. Williamson, A new 3/4-approximation algorithm for MAXSAT, *Third Conference on Integer Programming and Combinatorial Optimization*, 1993, pp. 313–321.
- [13] L. Lovász, On the ratio of optimal integral and fractional covers, *Discrete Mathematics* 13 (1975) 383–390.
- [14] V. Chvátal, A greedy heuristic for the set covering, *Mathematics of Operations Research* 4 (1979) 233–235.
- [15] H. Bronnimann, M.T. Goodrich, Almost optimal set covers in finite VC-dimension, *Proceedings of the 10th ACM Symposium on Computational Geometry*, 1994, pp. 293–302.
- [16] N. Alon, J. Spencer, *The Probabilistic Method.*, Wiley, New York, 1992.
- [17] N. Hall, D. Hochbaum, A fast approximation algorithm for the multicovering problem, *Discrete Applied Mathematics* 15 (1989) 35–40.
- [18] D. Hochbaum, Approximation algorithms for set covering and vertex cover problems, *SIAM Journal on Computing* 11 (1982) 555–556.
- [19] E. Balas, A sharp bound on the relation between optimal integer and fractional covers, *Mathematics of Operations Research* 9 (1984) 1–7.
- [20] R. Bar Yehuda, S. Even, A linear time approximation algorithm for the weighted vertex cover problem, *Journal of Algorithms* (1981) 198–206.
- [21] G. Dobson, Worst-case analysis of greedy heuristics for integer programming with nonnegative data, *Mathematics of Operations Research* 7 (1982) 515–531.
- [22] M. Fisher, L. Wolsey, On the greedy Heuristic for covering and packing problems, *SIAM Journal on Algebraic and Discrete Methods* 3 (1982) 584–591.
- [23] D. Angluin, L.G. Valiant, Fast probabilistic algorithms for Hamiltonian circuits and matchings, *Journal of Computer System Science* 18 (1979) 155–193.
- [24] D. Hochbaum, Heuristics for the fixed cost median problem, *Mathematical Programming* 22 (1982) 148–162.
- [25] B. Bollobas, *Random Graphs*, Academic Press, London, 1985.
- [26] N. Alon, On network reliability, 1993 (Preprint).
- [27] H. Gabow, M. Goemans, D. Williamson, (1993). An efficient approximation algorithm for the survivable network Design Problem, *Third Conference on Integer Programming and Combinatorial Optimization*, 1993, pp. 57–74.
- [28] J. Rhys, A selection problem of shared fixed costs and network flows, *Management Science* 17 (1970) 200–207.
- [29] C. Papadimitriou, M. Yannakakis, Optimization, approximation and complexity classes, *Journal of Computer and System Science* 43 (1991) 425–440.
- [30] S. Rajagopalan, V.V. Vazirani, Primal-dual RNC approximation algorithms for (multi)-set (multi)-cover and covering integer programs, *Proceedings of the 33rd IEEE Conference on the Foundations of Computer Science*, 1993, pp. 293–302.