NPS55-87-007

NAVAL POSTGRADUATE SCHOOL

Monterey, California

BOOLEAN AND GRAPH THEORETIC FORMULATION OF
THE SIMPLE PLANT LOCATION PROBLEM

P. M. DEARING
P. L. HAMMER
B. SIMEONE

AUGUST 1987

Prepared for:
Naval Postgraduate School
Monterey, CA 93943-5000

NAVAL POSTGRADUATE SCHOOL
MONTEREY, CALIFORNIA

Rear Admiral R. C. Austin
Superintendent

D. A. Schrady
Provost

Reproduction of all or part of this report is authorized.

This report was prepared by:

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b RESTRICTIVE MARKINGS
2a SECURITY CLASSIFICATION AUTHORITY	3 DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution is unlimited
2b DECLASSIFICATION/DOWNGRADING SCHEDULE	
4 PERFORMING ORGANIZATION REPORT NUMBER(S) NPS55-87-007	5 MONITORING ORGANIZATION REPORT NUMBER(S)
6a NAME OF PERFORMING ORGANIZATION Naval Postgraduate School 6b OFFICE SYMBOL (If applicable) Code 55	7a NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State, and ZIP Code) Monterey, California 93943-5000	7b ADDRESS (City, State, and ZIP Code)
8a NAME OF FUNDING/SPONSORING ORGANIZATION 8b OFFICE SYMBOL (If applicable)	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State, and ZIP Code)	10 SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO / PROJECT NO / TASK NO / WORK UNIT ACCESSION NO

11 TITLE (Include Security Classification)

BOOLEAN AND GRAPH THEORETIC FORMULATIONS OF THE SIMPLE PLANT LOCATION PROBLEM

12. PERSONAL AUTHOR(S)
DEARING, P.M., Hammer, P.L. (Rutgers University), SIMEONE, B. (Rutgers University)

13a. TYPE OF REPORT	13b TIME COVERED	14 DATE OF REPORT (Year, Month, Day)	15 PAGE COUNT
Technical	FROM ____ TO ____	August 1987	18

16. SUPPLEMENTARY NOTATION

17 COSATI CODES			18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	location theory, pseudo-Boolean functions, set covering, vertex packing, graph theory, integer programming

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

The simple plant location problem is formulated as the minimization of a pseudo-Boolean function. This form of the problem is then transformed into a set covering problem and also into a weighted vertex packing problem on a graph. These formulations are compared to simil formulations in the literature and to the "standard" integer programming formulation.

20 DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	21 ABSTRACT SECURITY CLASSIFICATION Unclassified
22a NAME OF RESPONSIBLE INDIVIDUAL Prof. P.M. Dearing	22b TELEPHONE (Include Area Code) (408) 646-2594 22c OFFICE SYMBOL Code 55

DD FORM 1473, 84 MAR — 83 APR edition may be used until exhausted. All other editions are obsolete

SECURITY CLASSIFICATION OF THIS PAGE

☆ U.S. Government Printing Office: 1986 606

BOOLEAN AND GRAPH THEORETIC FORMULATIONS OF THE SIMPLE PLANT LOCATION PROBLEM

by

P. M. Dearing*
Clemson University
and Naval Postgraduate School

P. L. Hammer**
Rutgers University

and

B. Simeone
Universita' di Roma "La Sapienza"
and Rutgers University

Abstract

The simple plant location problem is formulated as the minimization of a pseudo-Boolean function. This form of the problem is then transformed into a set covering problem and also into a weighted vertex packing problem on a graph. These formulations are compared to similar formulations in the literature and to the "standard" integer programming formulation.

*Partially supported by NSF Grant number ISP-8011451 (EPSCoR) and the National Research Council

**Partially supported by NSF Grant number DMS-83-05569

INTRODUCTION

In this paper the simple plant location problem, SPLP, is formulated as the minimization of a pseudo-Boolean function. This formulation is then transformed into two other discrete optimization problems: a set covering problem and a weighted vertex packing problem on a graph. These three formulations of the SPLP are compared to similar formulations that have appeared in the literature and the differences are discussed.

The SPLP is described as follows. Let $P = \{i: i = 1, \ldots, p\}$ be the index set of potential locations for plants (or plant sites) in some space such as a network or the plane. If a plant is opened at location i, a fixed cost f_i is incurred. Let $D = \{j: j = 1, \ldots, d\}$ be the index set of customers. The unit cost of transportation between customer j and plant location i is given by c_{ij} for each $i \in P$ and $j \in D$. Each customer has a demand which must be met by the opened plants and is assumed to be one unit. If a customer's demand is different from one unit, it is scaled to one and the transport costs are scaled accordingly. Assume further that the capacity for each opened plant is sufficient to meet the demand of all customers (hence the alternative name "uncapacitated plant location problem"). The SPLP is then to choose a subset of locations from P at which plants are opened and to specify the transportation between opened plants and all customers so as to meet customer demand and to minimize the total fixed cost plus the total transportation cost.

As an example, consider the network in Figure 1 where there is a customer at each node and a plant site at each of the nodes 1, 3 and 4. Each edge number is the transportation cost accross that edge. For each plant site i and customer j, the transportation cost c_{ij} is the minimum cost over all paths between i and j. Table 1 gives the fixed cost f_i and the transportation costs c_{ij} for each $i \in P$ and $j \in D$.

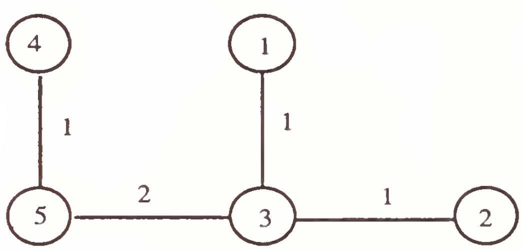

Figure 1: A tree network with customer and plant sites at nodes

		customers					
		1	2	3	4	5	f_i
	1	0	2	1	4	3	5
plant sites	3	1	1	0	3	2	6
	4	4	4	3	0	1	4

Table 1: Costs data for the SPLP of Figure 1

The approaches discussed here are not limited to a SPLP defined on a tree network, nor even to a network. The approaches apply to any SPLP defined by a matrix of nonnegative c_{ij} values and nonnegative fixed costs f_i.

In the literature there are several recent surveys of location problems that include a discussion of the SPLP. The extensive surveys by Cornuejols, Nemhauser and Wosley [3] and by Krarup and Pruzan [8] are devoted almost entirely to the SPLP. Each includes a thorough discussion of the problem's origins and several of its formulations. References [2,3] consider the "uncapacitated facility location problem" which is equivalent to the SPLP, but maximizes the total revenue from satisfying customer demands minus the cost of opening plants at the chosen sites. The formulations discussed below may be applied directly to the uncapacitated facility location problem with only minor modifications.

A BOOLEAN FORMULATION

We formulate the SPLP directly as the minimzation of a pseudo-Boolean function. This formulation follows naturally from the well known property that in some optimal solution, each customer j will receive its entire unit of demand from one open plant, namely a plant with minimum transportation cost to customer j.

For each $i \in P$, define the variable y_i to be 1 if a plant is opened at location i, and 0 otherwise. Let $\ddot{y}_i = 1 - y_i$ be the complement of y_i. For each customer j, let j(.) be a permutation of the location indices $i \in P$ so that the transportation costs from each j(i) to j are in nondecreasing order:

$$c_{j(1)j} \leq c_{j(2)j} \leq \cdots \leq c_{j(p)j}. \qquad (1)$$

Observe that in meeting the demand of customer j, the cost $c_{j(1)j}$ is incurred iff a plant is open at j(1), i.e., $y_{j(1)} = 1$; or the cost $c_{j(2)j}$ is incurred iff a plant is open at j(2) and no plant is open at j(1), i.e., $\ddot{y}_{j(1)}y_{j(2)} = 1$; or in general, the cost $c_{j(k)j}$ is incurred iff a plant is open at j(k) and no plant is open at locations with smaller cost: j(1), j(2), . . ,j(k-1), i.e., $\ddot{y}_{j(1)}\ddot{y}_{j(2)} \cdot \cdot \cdot \ddot{y}_{j(k-1)}y_{j(k)} = 1$. Finally, if no plant is open, i.e., $\ddot{y}_{j(1)}\ddot{y}_{j(2)} \cdot \cdot \cdot \ddot{y}_{j(p)} = 1$, the demand of customer j is not met. This corresponds to an infeasible solution and is avoided by including in the problem a penalty term $M_j\ddot{y}_{j(1)}\ddot{y}_{j(2)} \cdot \cdot \cdot \ddot{y}_{j(p)}$ where M_j is a large cost. Suitable values for the M_j are discussed subsequently.

To simplify the notation, let $\Pi\ddot{y}_{j(k)}$ denote the product $\ddot{y}_{j(1)}\ddot{y}_{j(2)} \cdot \cdot \cdot \ddot{y}_{j(k)}$ for each $j \in D$ and $k = 1, \ldots, p$. Furthermore, let $\Pi\ddot{y}_{j(0)} = 1$ in subsequent expressions. Then the transportation cost to customer j from the set of open plants is given by

$$T_j = \sum_{i=1}^{p} c_{j(i)j}\Pi\ddot{y}_{j(i-1)}y_{j(i)} + M_j\Pi\ddot{y}_{j(p)} . \qquad (2)$$

In the example problem, for customer 2, we have

$$T_2 = 1y_3 + 2\ddot{y}_3y_1 + 4\ddot{y}_1\ddot{y}_3y_4 + M_2\ddot{y}_1\ddot{y}_3\ddot{y}_4.$$

The total fixed cost incurred by the open plants is given by the sum of the f_iy_i. Then, a pseudo-Boolean function F (see reference [7]) of the variables $y_1, y_2, \ldots, y_p$ is defined as the total fixed cost plus the sum of the T_j's for each j in D:

$$F(y_1, \ldots, y_p) = \sum_{i=1}^{p} f_iy_i + \sum_{j=1}^{d} \left[\sum_{i=1}^{p} c_{j(i)j}\Pi\ddot{y}_{j(i-1)}y_{j(i)} + M_j\Pi\ddot{y}_{j(p)} \right] .$$

The SPLP may be written as

P1: $$\min \{ F(y_1, \ldots, y_p) : y_i = 0,1, \quad i \in P \}.$$

A formulation equivalent to P1 is reported by Hammer [6] and also included in the survey [8].

In a specific instance of the SPLP, the function F can generally be simplified by adding together those terms with identical variables. Two customers j and j' will have terms with identical variables, say $\Pi\ddot{y}_{j(k-1)}y_{j(k)}$ in T_j, and $\Pi\ddot{y}_{j'(k-1)}y_{j'(k)}$ in $T_{j'}$, whenever the sequence of plant sites, ordered by j(i) and j'(i), are identical for $i = k, k+1, \ldots, p$ for some $k < p$. This ability to aggregate like terms is a desirable feature of the pseudo-Boolean function F that is not found in most other formulations of the SPLP. In particular, since each penalty term $\Pi\ddot{y}_{j(p)}$ is identical for each customer j, all these may be added together to yield one penalty term. Let M denote the sum of the M_j over $j \in D$. Then the function F for the example problem, simplified by adding together like terms, is given as follows:

$$F(y_1,y_3,y_4) = 5y_1 + 7y_3 + 5y_4 + \ddot{y}_1y_3 + 3\ddot{y}_3y_1 + 5\ddot{y}_4y_3 + 11\ddot{y}_3\ddot{y}_1y_4 + 7\ddot{y}_4\ddot{y}_3y_1 + M\ddot{y}_1\ddot{y}_3\ddot{y}_4.$$

PENALTY VALUES

It is convenient to rewrite the function F, using the common penalty term M, as:

$$F(y_1, \ldots, y_p) = \sum_{i=1}^{p} f_iy_i + \sum_{j=1}^{d}\sum_{i=1}^{p} c_{j(i)j}\Pi\ddot{y}_{j(i-1)}y_{j(i)} + M\Pi\ddot{y}_{j(p)} .$$

The following discussion shows how the value of F depends on the penalty term M. Observe that a set of variable values $(y_1, \ldots, y_p)$ is feasible to P1 if and only if $\Pi\ddot{y}_{j(p)} = 0$. Furthermore, if $(y_1, \ldots, y_p)$ is an infeasible solution, then $F(y_1, \ldots, y_p) = M$, the penalty value. Consider a minimum feasible solution $(y^*_1, \ldots, y^*_p)$, and let $z^* = F(y^*_1, \ldots, y^*_p)$ denote the optimal objective function value.

The first claim is that with penalty value $M < z^*$, each minimum solution to F is infeasible. This follows since each infeasible solution has the objective function value $M < z^* \leq F(y_1, \ldots, y_p)$ for each feasible solution $(y_1, \ldots, y_p)$. Alternatively, with a penalty cost of $M \geq z^*$ each minimum solution to F has optimal value z^*. This follows since for each feasible solution $(y_1, \ldots, y_p)$, $z^* \leq F(y_1, \ldots, y_p)$, and for each infeasible solution $(y'_1, \ldots, y'_p)$, $z^* \leq M = F(y'_1, \ldots, y'_p)$. This same argument shows that with penalty cost $M > z^*$, each minimum solution to F is feasible.

Thus for all $M \geq z^*$, the optimal objective function value of P1 is z^*, and any upper bound on z^* will suffice for a value of M. Values for the penalty coefficients M_j may be obtained from any feasible solution $(y^*_1, \ldots, y^*_p)$. Define

$$f^* = \sum_{i=1}^{p} f_iy^*_i , \quad \text{and} \quad c^*_j = \min \left\{ c_{j(i)j}\Pi\ddot{y}^*_{j(i-1)}y^*_{j(i)} : i = 1, \ldots, p \right\},$$

that is, f^* is the total fixed cost of this solution and c^*_j is the minimum transport cost to customer j for this solution. Then each penalty cost M_j may be chosen as any value greater than $c^*_j + f^*/d$.

A SET COVERING FORMULATION

Problem P1 is next transformed into a set covering problem. This formulation has fewer variables and constraints than a set covering formulation of the SPLP by Kolen [9], while it retains the combinatorial properties of his formulation. It is also related to a covering formulation for the uncapacitated facility location problem reported in reference [2] but obtained by a different approach.

For each variable $y_{j(i)}$ which appears in uncomplemented form in expression (2) substitute $y_{j(i)} = 1 - \ddot{y}_{j(i)}$ and simplify. Expresion (2) then becomes, for customer j,

$$T'_j = c_{j(1)j} + \sum_{i=1}^{p-1}\left(c_{j(i+1)j} - c_{j(i)j}\right)\Pi\ddot{y}_{j(i)} + \left(M_j - c_{j(p)j}\right)\Pi\ddot{y}_{j(p)}. \tag{3}$$

In the example problem, for customer 2, we have

$$T'_2 = 1 + \ddot{y}_3 + 2\ddot{y}_1\ddot{y}_3 + (M_2 - 4)\ddot{y}_1\ddot{y}_3\ddot{y}_4.$$

Then the function F has an equivalent form, denoted by F', which is the sum of the fixed costs plus the sum of the T'_j for each j in D:

$$F'(y_1, \ldots, y_p) = \sum_{i=1}^{p} f_i y_i + \sum_{j=1}^{d} T'_j .$$

The function F' has a natural interpretation. For each customer $j \in D$, the cost $c_{j(1)j}$ is a constant that is incurred by any solution to the SPLP. That is, the minimum transport cost to customer j, $c_{j(1)j}$, must always be incurred. This cost may be zero if a plant site and a customer location coincide. Each term $(c_{j(i+1)j} - c_{j(i)j})$ is the incremental transportation cost incurred by customer j if plants at j(1), j(2), . . ,j(i) are not opened. In this case, the transportation cost to j is at least $c_{j(i+1)j}$, and if j(i+1) is open the cost is exactly $c_{j(i+1)j}$.

As above, the function F' may be simplified by adding together like terms. For the example problem this yields:

$$F'(y_1,y_3,y_4) = 2 + 5y_1 + 6y_3 + 4y_4 + \ddot{y}_1 + 2\ddot{y}_3 + 4\ddot{y}_4 + 7\ddot{y}_1\ddot{y}_3 + 2\ddot{y}_3\ddot{y}_4 + (M - 18)\ddot{y}_1\ddot{y}_3\ddot{y}_4 .$$

A further reduction is possible for F' between pairs of complementary variables y_i and $\ddot{y}_i$ using the identity $y_i + \ddot{y}_i = 1$. Any two terms of the form $ay_i + b\ddot{y}_i$ may be rewritten as follows:

$$ay_i + b\ddot{y}_i = \{ (a-b)y_i + b, \text{ if } a \geq b; \text{ or } 0y_i + (b-a)\ddot{y}_i + a, \text{ if } a \leq b\}. \qquad (4)$$

In the second case, an uncomplemented variable y_i with zero cost may be given the value one, and all terms containing the complementay variable $\ddot{y}_i$ will have value zero and may be eliminated from F. For the example problem the reduction (4) yields:

$$F'(y_1,y_3,y_4) = 9 + 4y_1 + 4y_3 + 0y_4 + 7\ddot{y}_1\ddot{y}_3 + 2\ddot{y}_3\ddot{y}_4 + (M - 18)\ddot{y}_1\ddot{y}_3\ddot{y}_4.$$

Here y_4 is retained in the function F' in order to illustrate the next transformation.

Next, the SPLP is transformed by replacing each term $\Pi\ddot{y}_{j(i)}$ in F' by a variable $z_{\Pi j(i)}$, where $\Pi j(i) = \{j(1)j(2)...j(i)\}$, and by adjoining the constraints

$$z_{\Pi j(i)} \geq \Pi\ddot{y}_{j(i)} \qquad i \in P,\ j \in D. \qquad (5)$$

But each constraint of (4) is equivalent, for zero-one variables and their complements, to the constraint

$$\sum_{k=1}^{i} y_{j(k)} + z_{\Pi j(i)} \geq 1$$

These transformations, applied to the general expression for F', yield the set covering problem:

$$\text{P2: } \min \sum_{i=1}^{p} f_i y_i + \sum_{j=1}^{d} \Big[\sum_{i=1}^{p-1} (c_{j(i+1)j} - c_{j(i)j}) z_{\Pi j(i)} + (M_j - c_{j(p)j}) z_{\Pi j(p)} \Big] + \sum_{j=1}^{d} c_{j(1)j}$$

$$\text{s. t. } \sum_{k=1}^{i} y_{j(k)} + z_{\Pi j(i)} \geq 1 \qquad i \in P,\ j \in D$$

$$y_i,\ z_{\Pi j(i)} = 0,1 \qquad i \in P,\ j \in D$$

Problem P2 will have at most d(p-1) + p + 1 variables and d(p-1) + 1 constraints.

When F' is simplified by adding together like terms and by applying the reduction (4), the number of variables $z_{\Pi j(I)}$ and the number of constraints in P2 will be decreased accordingly. For the example problem, with F' simplified and M' = (M - 18), problem P2 becomes:

$$\min 9 + 4y_1 + 4y_3 + 0y_4 + 7z_{13} + 2z_{34} + M'z_{134}$$

$$\begin{array}{llllllll} \text{s. t.} & y_1 & + y_3 & & + z_{13} & & & \geq 1 \\ & & y_3 & + y_4 & & + z_{34} & & \geq 1 \\ & y_1 & + y_3 & + y_4 & & & + z_{134} & \geq 1 \end{array}$$

all variables zero - one.

The above example suggests further simplifications that may apply to specific problems. Any y_i variable with a zero cost coefficient may be set to one in any optimal solution. The corresponding column and all rows with a one in that column may be eliminated. In the example problem, setting $y_4 = 1$ and eliminating columns and rows leaves one constraint: $y_1 + y_3 + z_{13} \geq 1$, from which it is easily seen that the two alternative optimal solutions are to open plants at sites {1,3} or {1,4} with a cost of 13. Thus the set covering formulation may be interpreted as preprocessing of the problem data that may specify some plant sites to be opened a priori.

The well known set covering 'reduction rules' [10] will generally apply to problem P2. For example, since the column corresponding to a variable y_i always dominates a column corresponding to a variable $z_{\Pi j(k)}$ for $i \in \Pi j(k)$, the column of $z_{\Pi j(k)}$ can be eliminated if the cost coefficient of y_i is not greater than that of $z_{\Pi j(k)}$. In particular, since the coefficient of y_i is less than M, the variable $z_{\Pi j(p)}$ and its column can always be eliminated.

Row reductions may follow. If the variable $z_{\Pi j(k)}$ has been eliminated from a row by a column reduction, this row may now be dominated by some other row. In this case, the dominating row may be eliminated.

Problem P2 is quite similar to Kolen's set covering formulation of the SPLP [9] with the following differences. First, Kolen's formulation does not allow the simplifications we get from adding together like terms. Each such simplification eliminates at least one variable and constraint. His formulation has a penalty variable $z_{\Pi j(p)}$ and its corresponding constraint for every j in D, whereas these terms are combined into one term and constraint in problem P2. These observations imply the constraint matrix of problem P2 is a proper subset of the constraint matrix of Kolen's formulation. The example problem is taken from Kolen [9]. His set covering formulation of the example problem requires 17 constraints and 20 variables.

The set covering reduction rules also have more limited application to Kolen's formulation. In general, these rules would eliminate the columns corresponding to the variables $z_{\Pi j(p)}$ for each j in D, but since other terms are not agregated, fewer columns may be eliminated. For the example problem, only the rows and columns corresponding to the $z_{\Pi j(p)}$ variables could be eliminated. Furthermore, in Kolen's formulation, no y_i variables have zero coefficients so that no plants can be set open as in the example problem.

Kolen showed that the constraint matrix of his set covering formulation was totally balanced [9] when the transportation costs were weighted distances in a tree network and the nodes served both as customer and plant locations. Since the constraint matrix of P2 is a submatrix of the matrix of Kolen's formulation, and since a submatrix of a totally balanced matrix is totally balanced, this same property holds for the formulation P2.

Problem P2 is closely related to a formulation reported in reference [2] for the uncapacitated facility location problem and obtained by a canonical reduction of the matrix of c_{ij} values. This formulation also aggregates terms where possible and so reduces the number of variables and constraints.

A WEIGHTED DOMINATION PROBLEM

The set covering problem P2 is equivalent to a weighted domination problem on a bipartite graph. To see this equivalence, problem P2 is expanded by adding redundant variables and constraints so that the expanded matrix of constraint coefficients is the vertex-vertex incidence matrix of a bipartite graph.

Consider the function F', which has been simplified by adding together like terms and reduced by expression (4). Let S be the subset of indices i in P corresponding to those terms in F' that contain only $\ddot{y}_i$. For each i in S, $\ddot{y}_i$ represents the aggregation of one or more variables $\ddot{y}_{j(1)}$ over those indices j such that j(1) = i. To obtain problem P2 each variable $\ddot{y}_i$ was replaced by the variable z_i and the constraint $y_i + z_i \geq 1$ is included.

A new problem, called P2', is obtained by expanding problem P2 as follows. For each i in P-S, add the variable z_i to problem P2 with a zero cost in the objective function and add the constraint $y_i + z_i \geq 1$. Problem P2' now has a variable z_i and a constraint $y_i + z_i \geq 1$ for each i in

P. This version of P2' is equivalent to P2 since any feasible solution to P2 may be amended by $z_i = 1$ for i in P - S to obtain a feasible solution to P2' with no change in the objective function. Conversely, deleting $z_i = 1$ from any feasible solution to P2' gives a feasible solution to P2 with no change in the objective function.

Let m be the number of variables $z_{\Pi j(i)}$ in P2', including the z_i added above. Note that m is also the number of constraints in P2'. The coefficient matrix of P2' may be written as (I_m, A), where I_m is an mxm identity matrix with a column corresponding to each variable $z_{\Pi j(i)}$, and A is an mxp matrix with a column corresponding to each variable y_i. Observe that A contains the pxp identity matrix I_p as a submatrix since for each i in P, there is a constraint $z_i + y_i \geq 1$.

Next, the set of constraints with coeficient matrix (A^t, I_p), where A^t is the transpose of A, is added to P2'. Then the coefficient matrix of P2' becomes:

$$C = \begin{pmatrix} I_m & A \\ A^t & I_p \end{pmatrix}$$

For each i in P, the i^{th} constraint of (A^t, I_p) contains z_i and y_i (since A contains I_p) and is therefore satisfied by any solution that satisfies the constraint $z_i + y_i \geq 1$ in (I_m, A). Thus the constraints of (A^t, I_p) are redundant and problem P2' is equivalent to P2.

Since the matrix C is symmetric with ones on the diagonal, it is a vertex-vertex adjacency matrix of some graph G. The graph G is seen to be bipartite with one set of vertices corresponding to variables y_i and the other to the variables $z_{\Pi j(i)}$. Each vertex $z_{\Pi j(i)}$ is adjacent to those vertices y_i where i is in the set $\Pi j(i)$. Each vertex is weighted by the cost coefficient of the respective variable. Therefore, the set covering problem P2' is a weighted domination problem on the bipartite graph G. Figure 2 shows the bipartite graph of the example problem.

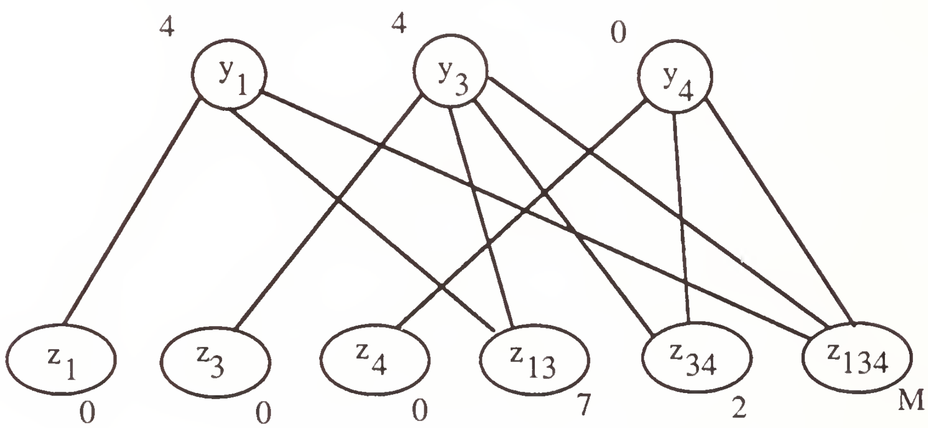

Figure 2: Bipartite Graph of the Example Problem

A VERTEX PACKING FORMULATION

The third formulation obtained for the SPLP is that of a weighted vertex packing problem on a graph. For each customer j and its permuation j(i) for i = 1, . . . , p, we have the following Boolean identity:

$$\sum_{i=1}^{p} \Pi \ddot{y}_{j(i-1)} y_{j(i)} + \Pi \ddot{y}_{j(p)} = 1. \tag{6}$$

Solving for $\Pi \ddot{y}_{j(p)}$, substituting into expression (2) and simplifying yields

$$T''_j = \sum_{i=1}^{p} (c_{j(i)j} - M_j) \Pi \ddot{y}_{j(i-1)} y_{j(i)} + M_j \tag{7}$$

for each j in D. For the example problem and customer 2,

$$T''_2 = (1-M_2)y_3 + (2-M_2)\ddot{y}_3 y_1 + (4-M_2)\ddot{y}_1\ddot{y}_3 y_4 + M_2.$$

Next substitute $y_i = 1 - \ddot{y}_i$ in each term $f_i y_i$ of F. Making these substitutions in the function F gives an equivalent form:

$$F(y_1, \ldots, y_p) = - \sum_{i=1}^{p} f_i \ddot{y}_i - \sum_{j=1}^{d}\sum_{i=1}^{p} (M_j - c_{j(i)j}) \Pi \ddot{y}_{j(i-1)} y_{j(i)} + M + \sum_{i=1}^{p} f_i .$$

Define the function $G(y_1, \ldots, y_p) = - F(y_1, \ldots, y_p)$, which is a pseudo-Boolean function in posiform. Then problem P1 may be written in the equivalent form as:

$$\text{P3:} \quad \max \{G(y_1, \ldots, y_p): y_i = 0,1, \ i \in P\}.$$

The function G for the example problem, with $M_j = 5$ and without simplifications, is:

$$G(y_1, y_3, y_4) = 5\ddot{y}_1 + 6\ddot{y}_3 + 4\ddot{y}_4 + 5y_1 + 4\ddot{y}_1 y_3 + \ddot{y}_1\ddot{y}_3 y_4 + 4y_3 + 3\ddot{y}_3 y_1 + \ddot{y}_1\ddot{y}_3 y_4 + 5y_3 + 4\ddot{y}_3 y_1 + 2\ddot{y}_1\ddot{y}_3 y_4 + 5y_4 + 2\ddot{y}_4 y_3 + \ddot{y}_3\ddot{y}_4 y_1 + 4y_4 + 3\ddot{y}_4 y_3 + 2\ddot{y}_3\ddot{y}_4 y_1 - 40.$$

Hammer and Rudeanu [7], and Ebenegger, Hammer, and de Werra [5] have shown that the problem of maximizing a pseudo--Boolean function given in a posiform may be transformed into a maximum weighted vertex packing problem on a graph. In reference [5], the graph corresponding to the posiform pseudo-Boolean function is called a "conflict graph". The conflict graph has a vertex for each term in the function, and an edge connecting two vertices if the corresponding terms contain at least one complementary pair of variables. That is, if y_i is in some term of G and its complement $\ddot{y}_i$ is in some other term, there is an edge between the two vertices. The vertex weights of the graph are the coefficients of the terms in G. We call the conflict graph of the function G derived from a SPLP a "plant location graph" (PLG). The plant location graph for the expample problem, without any simplifications is shown in figure 3.

It is shown in reference [5] that the maximum value of G is equal to the value of a maximum weighted vertex packing of the conflict graph. We denote the problem of determining a maximum weighted vertex packing in the plant location graph of the function G as problem P4.

The SPLP has been formulated as a weighted vertex packing problem by Cho, Johnson Padberg, and Rao [1] and by Cornuejols and Thizzy [4]. Our formulation is closely related theirs but with some interesting differences. An obvious similarity is the highly structured form of the plant location graph. For each j in D, the set of vertices corresponding to the terms in T'_j form a clique ('rows' in Figure 3). For each i, a star is determined with the vertex corresponding to $\ddot{y}_i$ at the center and an adjacent vertex corresponding to the term containing y_i in each expression T''_j ('columns' in Figure 3).

Several differences are observed. First, the plant location graph of P4 contains all the edges of the graph constructed in references [1,4] but has the additional edges (between cliques) corresponding to complementary variables. For example, there is an edge between $\ddot{y}_1 y_3$ and $y_1 \ddot{y}_3$ in the example problem. These additional edges strengthen the vertex packing formulation.

Second, if no simplifications are made in G by adding together like terms, the plant location graph has the same number (dp+p) of vertices as does the graph constructed in [1,4]. However, each simplification of G by adding like terms reduces the number of vertices in the plant location graph. This reduction in vertices is not possible for the graph constructed in references [1,4].

Third, the reduction given by expression (4) may be applied to the function G. For each i consider the terms $(M_j - c_{j(i)j})y_{j(i)} + f_i\ddot{y}_i$, such that $j(i) = i$. For the function G, the penalty cost M_j may be chosen sufficiently large so that $M_j - c_{j(i)j} > f_i$. Then these two terms reduce to $(M_j - c_{j(i)j} - f_i)y_{j(i)} + f_i$ which eliminates the vertiex corresponding to to $\ddot{y}_i$ from the plant location graph. With these reductions, the plant location graph has at most (dp) vertices. The conflict graph resulting from the function G, after all simplifications and reductions, is called the reduced plant location graph.

For the example problem without simplifications in the function G, the plant location graph contains 18 vertices and 73 edges. However applying all possible simplifications and reductions, the function G becomes:

$$G(y_1, y_3, y_4) = 3y_3 + 5y_4 + 4\bar{y}_1 y_3 + 7y_1\bar{y}_3 + 5y_3\bar{y}_4 + 4\bar{y}_1\bar{y}_3 y_4 + 3y_1\bar{y}_3\bar{y}_4 - 25.$$

Figure 4 shows the reduced plant location graph for the function G as simplified above.

Finally, for all SPLP with p plant sites and d customers, the vertex and edge sets of the graphs constructed in references [1,4] are identical. Only the vertex weights change. However, a plant location graph may differ from one SPLP to another with the same sets P and D since the vertex set and the edge set depend directly on the costs c_{ij} and f_i. In particular, the plant location graph for a SPLP arising from a tree network may differ from a SPLP arising from a general network of the same size. This raises the following question corresponding to Kolen's result for the SPLP on a tree network: Does the plant location graph arising from a SPLP on a tree network have some special struture?

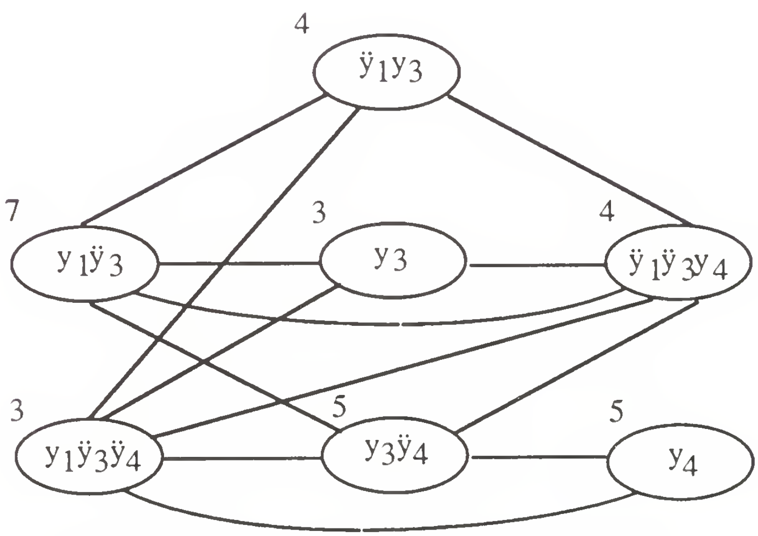

Figure 4: Reduced Plant Location Graph for G

RELATIONSHIP TO STANDARD INTEGER FORMULATIONS OF THE SPLP

Problems P1 and P3 are equivalent to the well known integer programming formulation of the SPLP, often called the "strong formulation" [3] which is given below. The variables y_i are the same as defined above. Define x_{ij} to be a binary variable indicating whether customer j is served by a plant at i ($x_{ij} = 1$) or not ($x_{ij} = 0$). Then the strong formulation may be written as follows:

$$\text{P5:} \qquad \min \sum_{i=1}^{p} \sum_{j=1}^{d} c_{ij} x_{ij} + \sum_{i=1}^{p} f_i y_i \tag{8}$$

$$\text{s. t.} \quad \sum_{i=1}^{p} x_{ij} = 1 \qquad j \in D \tag{9}$$

$$x_{ij} \leq y_i \qquad j \in D, \ i \in P \tag{10}$$

$$x_{ij} = 0,1 \qquad y_i = 0,1 \tag{11}$$

Given next is a relationship between the variables x_{ij} and the Boolean expressions in the y_i variables that were defined earlier. For any set of values for the y_i variables, $x_{ij} = 1$ if and only if a plant is open at site i and the transportation cost c_{ij} between customer j and site i is minimum over all other open plant sites. That is, for each $i \in P$ and $j \in D$ with the j(.) defined by expression (1) and $j(k) = i$,

$$x_{ij} = \ddot{y}_{j(1)}\ddot{y}_{j(2)} \cdots \ddot{y}_{j(k-1)}y_{j(k)} = \Pi\ddot{y}_{j(k-1)j}y_{j(k)}. \tag{12}$$

Expression (12) is substituted for each x_{ij} in problem P5. Thus each constraint of (10) becomes

$$\Pi\ddot{y}_{j(k-1)j}y_{j(k)} \leq y_i$$

which is always satisfied for y_i equal to either 0 or 1. With these substitutions, the constraints (10) can be discarded since they are always satisfied.

Substituting expression (12) for each x_{ij} into constraint (9) yields

$$\sum_{k=1}^{d} \Pi\ddot{y}_{j(k-1)}y_{j(k)} = 1 \qquad j \in D \tag{13}$$

Comparing (13) to the Boolean identity (6) implies that each constraint (13) is equivalent to requiring that $\Pi\ddot{y}_{j(p)} = 0$ for each j in D. Thus the constraints (9) may be eliminated from P5 if each term $\Pi\ddot{y}_{j(p)}$ is forced to be zero in an optimal solution. This is accomplished by adding to the objective function the terms $M_j \Pi \ddot{y}_{j(p)}$ where M_j is a large penalty cost, for each j in D.

Finally, the c_{ij} may be reindexed by the permutations j(.) to yield $c_{j(i)j}$, and expression (12) is substituted for each x_{ij} in the objective function. With these changes, problem P5 is written as

$$\min \sum_{j=1}^{d} \Big[\sum_{i=1}^{p} c_{j(i)j} \Pi \ddot{y}_{j(i-1)} y_{j(i)} + M \Pi \ddot{y}_{j(p)} \Big] + \sum_{i=1}^{p} f_i y_i$$

$$\text{s. t.} \quad y_i = 0,1$$

which is precisely problem P1.

Since problems P1 and P3 are equivalent, it follows that P5 is equivalent to P3. However, it is interesting to note that P3 may be interpreted as a Lagrangian relaxation of P5. Expression (12) is again substituted for x_{ij} in the objective function (8) and in each constraint (9) and (10) of problem P4. The constraints (10) are deleted as above since they are always satisfied. For the resulting problem, consider a Lagrangian relaxation with respect to the constraints (9) and with multipliers M_j. This yields the following problem.

$$\min \sum_{i=1}^{p} \sum_{j=1}^{d} c_{j(i)j} \Pi \ddot{y}_{j(i)j} y_{j(k)} + \sum_{i=1}^{p} f_i y_i + \sum_{j=1}^{d} M_i \Big(1 - \sum_{k=1}^{p} \Pi \ddot{y}_{j(k-1)} y_{j(k)}\Big)$$

$$y_i = 0,1.$$

Problem P3 is obtained by substituting $y_i = 1 - \ddot{y}_i$ in each term $f_i y_i$, collecting terms, and changing to a maximization problem by multiplying through by -1.

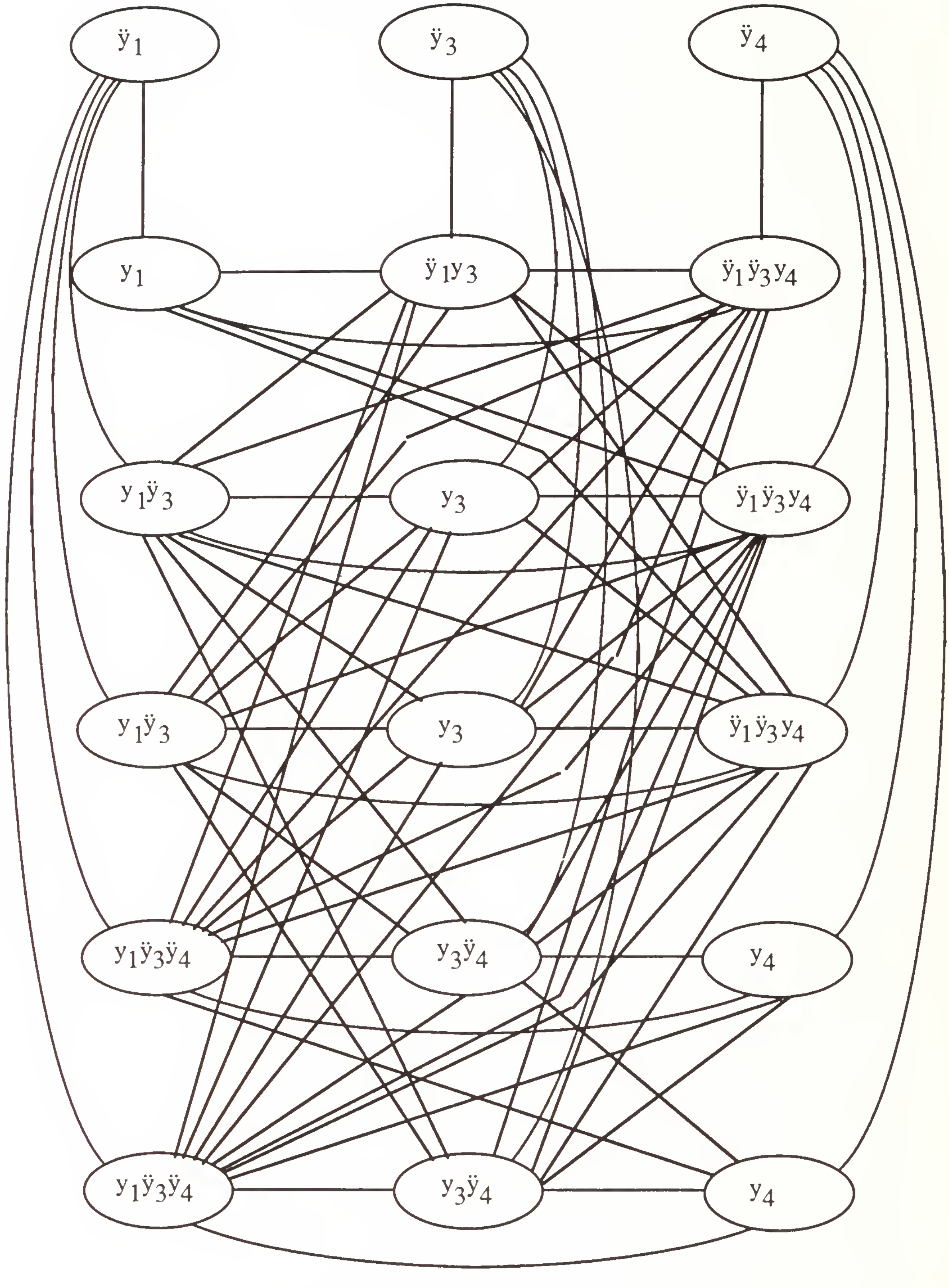

Figure 3: Conflict graph of G without simplifications

REFERENCES

[1] D. C. Cho, E. L. Johnson, M. Padberg, and M. R. Rao, "On the uncapacitated plant location problem I: Valid inequalities and facets," Math. of Opns. Res. 8, 579-589 (1983).

[2] G. Cornuejols, G. L. Nemhauser, and L. A. Wolsey, "A canonical representation of simple plant location problems and its applications," SIAM J. Alg. Disc. Math. 1, 261-272 (1980).

[3] G. Cornuejols, G. L. Nemhauser, and L. A. Wosley, "The uncapacitated facility location problem," Tech Report 605, School of OR & IE, Cornell Univ., (1984). Also, Ch. 3, Discrete Location Theory (Francis and Mirchandini, eds.) Interscience, (forthcoming).

[4] G. Cornuejols, and J. M. Thizy, "Some facets of the simple plant location polytope," Math. Prog. 23, 50-74 (1982).

[5] C. Ebenegger, P. L. Hammer, and D. de Werra, "Pseudo-Boolean functions and stability of graphs," Algebraic and Combinatorial Methods in Operations Research, R. E. Burkard, et. al. (eds.), North Holland, 83-97 (1984).

[6] P. L. Hammer, "Plant Location- a pseudo-Boolean approach," Israel J. Tech. 6, 330-332 (1968).

[7] P. L. Hammer and S. Rudeanu, Boolean Methods in Operations Research, Springer-Verlag, New York, 1968.

[8] J. Krarup and P. M. Pruzan, "The simple plant location problem: survey and synthesis," European Journal of Operations Research 12, 36-81 (1983).

[9] A. Kolen, "Solving covering problems and the uncpacitated plant location problem on trees," European Journal of Operations Research 12, 266-278 (1983).

[10] R. S. Garfinkel and G. L. Nemhauser, Integer Programming, John Wiley, New York, 1972.

DISTRIBUTION LIST

	NO. OF COPIES
Library (Code 0142) Naval Postgraduate School Monterey, CA 93943-5000	2
Defense Technical Information Center Cameron Station Alexandria, VA 22314	2
Office of Research Administration (Code 012) Naval Postgraduate School Monterey, CA 93943-5000	1
Center for Naval Analyses 2000 Beauregard Street Alexandria, Va 22311	1
Library (Code 55) Naval Postgraduate School Monterey, CA 93943-5000	1
Operations Research Center, Rm E40-164 Massachusetts Institute of Technology Attn: R. C. Larson and J. F. Shapiro Cambridge, MA 02139	1
Koh Peng Kong Ministry of Defence Blk A, Stockport Road SINGAPORE 0511	1