DÉPARTEMENT DE GÉNIE INDUSTRIEL

RAPPORT TECHNIQUE : EP75 - R - 35
Classification: Library of Congress

ON THE INTEGER - VALUED VARIABLES IN THE LINEAR VERTEX PACKING PROBLEM

BY: Jean-Claude Picard, Ecole Polytechnique
Maurice Queyranne, Ecole Polytechnique

July 1975

Ecole Polytechnique de Montréal

C.P. 501
Snowdon
Montréal 248

ON THE INTEGER - VALUED VARIABLES IN THE LINEAR VERTEX PACKING PROBLEM*

By

Jean-Claude Picard
Ecole Polytechnique, Montreal

and

Maurice Queyranne
Ecole Polytechnique, Montreal

* This research was supported by National Research Council of Canada GRANT A8528 and RD 804.

ON THE INTEGER-VALUED VARIABLES IN THE LINEAR VERTEX PACKING PROBLEM

Picard - Queyranne

ERRATA

p. 1	Line 5	$(v_i,v_j) \in E$	$(v_i,v_j) \notin E$
	Line 8	adjacent to).	adjacent to it).
p. 2	6 lines from bottom	$y_j \geqslant 0,1$	$y_j = 0,1$
	5 lines from bottom	is showed by these	is shown by the
p. 3	Line 7	in the section 2; in the section	in section 2; in section
	4 lines from bottom	partionned in nine	partitioned into nine
p. 4	First line under Fig. 1	$\sum_{i \in A_{j,k}} c_i x_i$	$\sum_{i \in A_{j,k}} c_i$
	Second line under Fig. 1	Let X^3 a solution	Let X^3 be a solution
p. 5	First line	Let X^4 now a solution	Let X^4 now be a solution
	First line under (2.4)	Finally, let X^5 a solution	Finally, let X^5 be a solution
p. 6	Lines 2 & 3	and then	and thus
	Line 6	Trotter write "determine	Trotter "determine
p. 7	Line 4	two distinct optimum ... a maximal	two optimum ... a different maximal
	Line 9	may be call	may be called
	Line 14	its adjacents	its adjacent
	3 lines from bottom	in which $v' \in V$	in which $v' \in V'$
p. 9	First line under (6)	together with S	together with $\bar{S}$
	First line under THEOREM IV	in the step (2)	in step (2)
p. 10	Line (i)	(i) Let V_i	(i) Let v_i
	Second line under (i)	in the step (2)	in step (2)
	Line 10 from bottom	to have the	to having the

1. INTRODUCTION

The definitions and notations given here are from [6]. Let $G = (V,E)$ be a finite, undirected, loopless graph with weights c_i on vertices $v_i \in V$. A vertex packing $(v.p)$ is a subset $P \subseteq V$ for which $v_i, v_j \in P$ implies $(v_i, v_j) \in E$. The weight $c(P)$ of a $v.p$ is defined as $c(P) = \sum_{v_j \in P} c_j$. There is no loss of generality in assuming that $c_j > 0$ for all $v_j \in V$, and that there is no isolated vertex in G (i.e. a vertex with no edge adjacent to).

Determining a maximum weighted $v.p.$ may be formulated as the integer program:

$$(VP) \quad \begin{aligned} & Max\ cX \\ & s.t. \\ & \quad AX \leqslant 1_m \\ & \quad x_j = 0,1\ ;\ j = 1,2,\ldots,n \end{aligned}$$

in which $m = |E|$, $1_m = (1,\ldots,1)$ is an m-vector of $1's$ and A is the $m x n$ edge-vertex incidence matrix of G.

Relaxing the integrality constraints to $X \geqslant 0_n$, gives the v.p linear program (VLP).

By the transformation:

$$U = 1_m - X \qquad (1\text{-}1)$$

we obtain the integer program:

$$\begin{array}{lll} & & Min\ cU \\ (CP) & s.t. & AU \geqslant 1_m \\ & & u_i = 0,1\ ;\ i = 1,2,\ldots,n \end{array}$$

This problem is the one of finding a minimum weighted covering of edges by nodes (cf. [1]); here we simply call it the covering problem. Let (CLP) be the linear relaxation of (CP); by (1-1), we obtain (CLP) from(VLP).

It is a well-known result that any basic feasible solution to (VLP) or (CLP) is (0, 1/2, 1)-valued: this was indicated by Lorentzen [5] as a simple consequence of the work of E. Johnson [4] ; indeed the dual of (CLP) is:

$$\begin{array}{lll} & & Max\ 1_m\ Y \\ (MLP) & s.t. & Y\,A \leqslant c \\ & & y_j \geqslant 0\ ;\ j = 1,2,\ldots,m \end{array}$$

which is the linear relaxation of the c-matching problem (cf. [2])

$$\begin{array}{lll} & & Max\ 1_m\ Y \\ (MP) & s.t. & Y\,A \leqslant c \\ & & y_j \geqslant 0,1\ ;\ j = 1,2,\ldots,m \end{array}$$

The interest of studying these linear relaxations is showed by these following two results:

(*i*) (VLP) may be solved by a good algorithm: it is a result attributed by Nemhauser and Trotter [6] to Edmonds and Pulleyblank that (VLP) is equivalent to solving a maximal flow problem on a related symmetric bipartite graph, twice the size of G.

(ii) an integer-valued variable in an optimal solution to (VLP) may keep the same value in an optimal solution of (VP) (cf. [6]).

This paper shows that there exists a unique maximum set of variables that may be integer-valued in an optimal solution to VLP; this result is shown in the section 2; in the section 3, we give a labeling procedure for determining this set using a sensitivity analysis on the maximum flow in the bipartite graph of Edmonds and Pulleyblank.

2. THE MAXIMUM SET OF INTEGER-VALUED VARIABLES

For $X \in R^n$, define $I(X)$ as the set of indices i ($i = 1,2,\ldots,n$) such that x_i is integer.

Lemma I. Let X^1 and X^2 be two optimal solutions to (VLP) then there is an optimal solution X to (VLP) such that:

$$I(X) = I(X^1) \cup I(X^2)$$

Proof: Because X^1 and X^2 are (1, 0, 1/2)-valued, the indices i ($i = 1,\ldots,n$) can be partionned in nine disjoint subsets (that may be empty) $A_{j,k}$ defined by:

$$A_{j,k} = \{i \mid x_i^1 = j \text{ and } x_i^2 = k\}$$

where $j,k \in \{0,1/2,1\}$

These subsets are given in Fig. I

Values of X^2

Values of X^1	1	0	1/2
1	$A_{1,1}$	$A_{1,0}$	$A_{1,1/2}$
0	$A_{0,1}$	$A_{0,0}$	$A_{0,1/2}$
1/2	$A_{1/2,1}$	$A_{1/2,0}$	$A_{1/2,1/2}$

Fig. 1

Let $c(A_{j,k})$ denote the quantity $\sum_{i \in A_{j,k}} c_i x_i$.

Let X^3 a solution defined by:

$$x_i^3 = \begin{cases} 1 & if \;\; i \in A_{1,1} \\ 0 & if \;\; i \in A_{0,0} \\ 1/2 & otherwise \end{cases}$$

then it is obvious that X^3 is a feasible solution to (VLP); furthermore, since both X^1 and X^2 are optimal, we have:

$$cX^1 \geqslant cX^3 \quad i.e \quad c(A_{1,0}) + c(A_{1,1/2}) \geqslant c(A_{0,1}) + c(A_{0,1/2}) \tag{2-1}$$

and $cX^2 \geqslant cX^3$ i.e $c(A_{0,1}) + c(A_{1/2,1}) \geqslant c(A_{1,0}) + c(A_{1/2,0})$ (2-2)

Adding (2-1) and (2-2) gives:

$$c(A_{1,1/2}) + c(A_{1/2,1}) \geqslant c(A_{0,1/2}) + c(A_{1/2,0}) \tag{2-3}$$

Let X^4 now a solution defined by:

$$x_i^4 = \begin{cases} x_i^2 & \text{if } \quad i \in I(X^2) \\ x_i^1 & \text{otherwise (i.e. if } x_i^2 = 1/2) \end{cases}$$

then X^4 is feasible (the reader may convince himself by inspection using the grids of Fig. 2) and:

$$cX^4 = cX^2 + 1/2(c\{A_{1,1/2}\} - c(A_{0,1/2}))$$

Since X^2 is optimal, we have also:

$$c(A_{1,1/2}) \leqslant c(A_{0,1/2}) \tag{2-4}$$

Finally, let X^5 a solution defined by:

$$x_i^5 = \begin{cases} x_i^1 & \text{if } \quad i \in I(x^1) \\ x_i^2 & \text{otherwise (i.e. if } x_i^1 = 1/2) \end{cases}$$

then X^5 is feasible and:

$$cX^5 = cX^1 + 1/2(c(A_{1/2,1}) - c(A_{1/2,0}))$$

Since X^1 is optimal, we have:

$$c(A_{1/2,1}) \leqslant c(A_{1/2,0}) \tag{2-5}$$

1	1	1
0	0	0
1/2	1/2	1/2

x_1

1	0	1/2
1	0	1/2
1	0	1/2

x_2

1	1/2	1/2
1/2	0	1/2
1/2	1/2	1/2

x_3

1	0	1
1	0	0
1	0	1/2

x_4

1	1	1
0	0	0
1	0	1/2

x_5

Fig. 2

From (2-3), (2-4) and (2-5), we get:

$$c(A_{1,1/2}) = c(A_{0,1/2}) \quad \text{and then } c(X^4) = c(X^2)$$

and $$c(A_{1/2,1}) = c(A_{1/2,0}) \quad \text{and then } c(X^5) = c(X^1).$$

Hence X^4 and X^5 are two optimal solutions to (VLP) such that:

$$I(X^4) = I(X^5) = I(X^1) \cup I(X^2)$$

In their paper [6], Nemhauser and Trotter write "determine an optimum solution to (VLP) in which a maximal (but possibly not maximum) collection of variables is integer-valued". Now we can show that the parenthetic assertion is superfluous.

THEOREM II. There is a unique maximal subset of integer-valued variables yielding an optimum solution to (VLP).

Proof: this theorem easily follows from Lemma 1: let X^1 and X^2 be two distinct optimum solutions to VLP, each one having a maximal subset of integer-valued variables. Then X^4 (or X^5) defined in the proof of Lemma 1 is an optimum solution whose integer-valued collection contains both the ones of X^1 and X^2, and this is inconsistent with the hypothesis of these subsets being maximal.

Such a collection may be call the maximum subset of integer-valued variables.

3. ALGORITHM FOR DETERMINING ALL THE INTEGER-VALUED VARIABLES.

Nemhauser and Trotter propose an algorithm for determining the integer-valued variables by checking each vertex v_j as follows: set $x_j = 1$, $x_k = 0$ for all its adjacents $v_k \in N(\{v_j\})$ and solve (VLP) on the remaining subgraph induced by $V_j = V - (\{v_j\} \cup N(\{v_j\}))$. The completion of this procedure needs solving about n (VLP)-problems on subgraphs of G. In order to derive a more efficient algorithm, we recall some results about the way (VLP) may be solved.

Let V' be a copy of the vertex set V of G, in which $v' \in V$ corresponds to $v \in V$. Let $W = V \cup V' \cup \{s,t\}$, where s(resp.t) is an artificial source (resp.sink), and $H = (W, F, \bar{c})$ a network whose arcs are:

(s, v_j) with capacity c_j for all j

(v_j, v'_k) with infinite capacity for all edges $(v_j, v_k) \in E$

(v'_k, t) with capacity c_k for all k

LEMMA 3.

Let $(S;\bar{S})$ be a minimum cut in H (recall that $s \in S$, $t \in \bar{S}$)

$$\text{Set } x_j = \begin{cases} 1 & \text{if } v_j \in S \text{ and } v'_j \in \bar{S} \\ 0 & \text{if } v_j \in \bar{S} \text{ and } v'_j \in S \\ 1/2 & \text{otherwise} \end{cases} \tag{3-1}$$

then X is an optimum solution to (VLP).

Proof: this lemma is a corollary of the theorem of Edmonds and Pulleyblank cited in [6]. The equivalence between (VP) in a bipartite graph and a minimum cut is indicated by Picard and Ratliff [7]; the value of a minimum cut is then $2c(1_n - X)$.

Solving the minimum cut problem of Lemma 3 may be done by the standard maximal-flow procedure of Ford and Fulkerson [3]. Let f_{jk} be the flow on the arc (v_j, v_k) in this maximal flow. We now test each vertex v_i in the following way: replace c_i by $c_i + \varepsilon$ and check the optimality of the current solution. Taking $\varepsilon > 0$ small enough reduces the standard Ford and Fulkerson labeling routine to the following:

Labeling procedure:

(1) Discarding: discard all the vertices $v_i \in V$, such that x_i is integer-valued.

(2) Initiating: choose an unscanned vertex $v_i \in V$, label v_i and go to step 3; if all the remaining vertices v_i of V are scanned, terminate.

(3) Direct labeling: label all the vertices v_k' such that there exists an edge $(v_j, v_k') \in F$ and v_j is labeled.

(4) Test: if v_i' is labeled, then v_i is scanned and go to step (2).

(5) Reverse labeling: label all the vertices v_k such that there exists a labeled vertex v'_j and $f_{jk} > 0$.
If there is no new labeling go to step (6), otherwise go to step (3).

(6) Solution modification: the set S of all labeled vertices, together with S, defines a cut $(S, \bar{S})$ in H. Redefine X by *(3-1)* and go to step (1).

THEOREM IV. The labeling procedure is a good algorithm for finding an optimum solution to (VLP) having the maximum set of integer-valued variables.

Proof: the procedure needs at most n choices in the step (2); each labeling (steps 3 to 5) assigns at most $2n$ labels, hence it is a good algorithm.

Let X^1 be a solution to (VLP) having the maximum set of integer-valued variables.

Let X^2 be the solution given by the algorithm; as in the proof of theorem 2, let X^3 be defined by

$$x_i^3 = \begin{cases} x_i^2 & \text{if } i \in I(X^2) \\ x_i^1 & \text{otherwise} \end{cases}$$

X^3 is an optimum solution, with the maximum collection of integer-valued variables, containing all the integer-valued variables of X^2 at the same value. It will be shown that $X^2 = X^3$:

(*i*) Let V_i be a vertex such that $x_i^3 = 1$ and $x_i^2 = 1/2$ and consider the solution X obtained in the application of the procedure just before v_i was to be considered in the step (2). We have $I(X) \subseteq I(X^2)$ since the algorithm builds an increasing set of integer-valued variables, and $x_i = 1/2$. Suppose we replace c_i by $c_i + \varepsilon$ with $\varepsilon > 0$ such that $\varepsilon < Min\{f_{kl}/f_{kl} > 0\}$; the value of the current solution becomes $cX + \varepsilon/2$, though the value of X^3 becomes $cX + \varepsilon$: X is not optimal for these new weights.

The weight change leads to have the ε extra amount of flow go along the arcs (s,v_i) and (v'_i,t); the standard Ford and Fulkerson labeling routine may be used to find an augmenting path between v_i and v'_i; since $\varepsilon < Min\{f_{kl}/f_{kl} > 0\}$, it is reduced to the given labeling procedure. If v_i' is labeled, the value of the flow becomes $2c(1_n - X) + \varepsilon$ after the flow change; this is a lower bound for the minimum cut and, consequently $cX + \frac{\varepsilon}{2}$ is an upper bound for the value of any solution to (VLP); this is inconsistent with X^3 being a solution to (VLP) with $cX + \varepsilon$ value.

(*ii*) Let v_i be a vertex with $x_i^3 = 0$ and $x_i^2 = 1/2$. There is a vertex v_j, adjacent to v_i, such that $x_j^3 = 1$ (otherwise, since $c_i > 0$, we could set $x_i^3 = 1/2$); since all the integer-valued variables of X^2 keep the same value in X^3, having $x_j^2 = 0$ is impossible and then $x_j^2 = 1/2$. Having a vertex v_j such that $x_j^3 = 1$ and $x_j^2 = 1/2$ has been proved impossible in the part (*i*).

Application:

Consider the graph of Fig. 3, with weights $c = 1_{10}$.

The corresponding bipartite graph is in Fig. 4; the max-flow, shown by thick lines, corresponds with the matching (1-4, 2-3, 5-6, 7-8, 9-10). The completion of the labeling process is listed below, facing the run of the algorithm of Nemhauser and Trotter (implemented with the standard Ford and Fulkerson labeling method).

	Labeling process	Nemhauser and Trotter
Step	labeling vertices	
1	no discarding	no discarding
2	v_1	$x_1 = 1$
3	v'_2	$x_2 = 0$
	v'_3	$x_3 = 0$
	v'_4	$x_4 = 0$
	v'_5	$x_5 = 0$
5	v_3	label v_6
	v_2	" v'_7
	v_6	" v_8
2	v'_s	" v'_9
4	v_1 is scanned	" v_{10}
		" v'_8
		" v_7
		" v'_6
		flow change
		$z_1 + c_1 = 2.5 + 1 < 5$

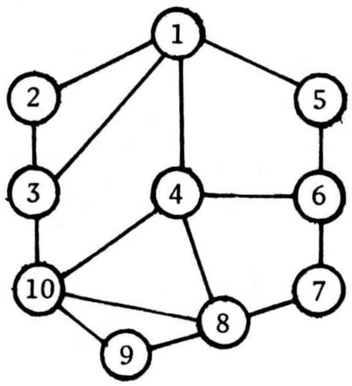

Fig. 3

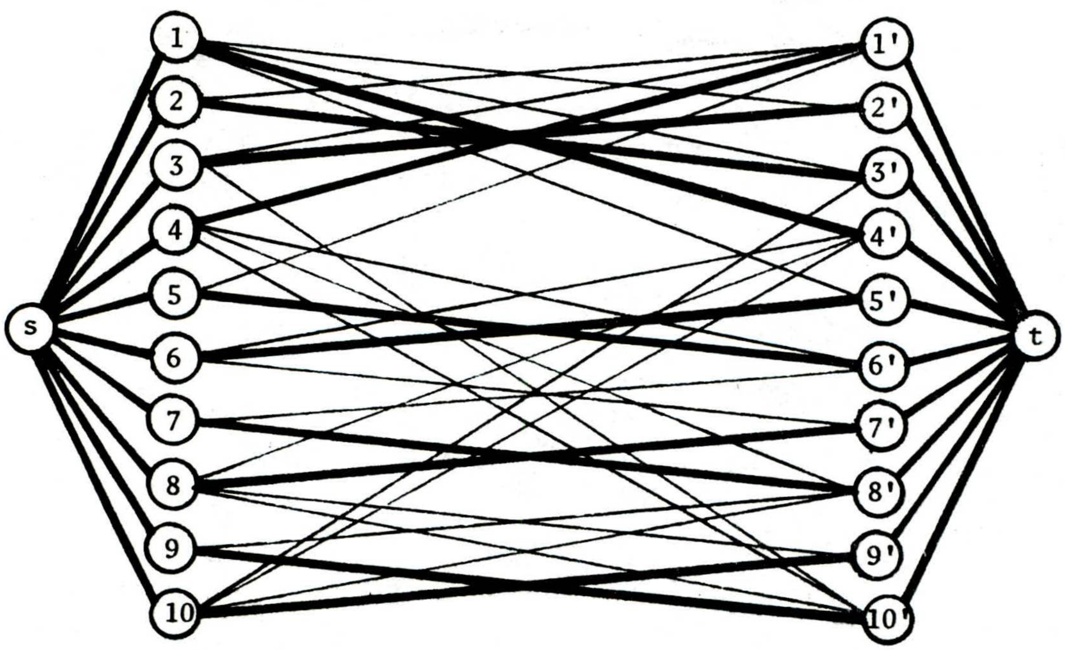

Fig. 4

	Labeling process	Nemhauser and Trotter
Step	labeling vertices	
2	v_2	$x_2 = 1$
3	v'_1	$x_1 = 0$
	v'_3	$x_3 = 0$
4	v_4	label v_4
3	v'_6	" v'_6
	v'_8	" v'_8
	v'_{10}	" v'_{10}
4	v_5	" v_5
	v_7	" v_7
	v_9	" v_9
5	Solution $x_2 = x_4 = x_5 = x_7 = x_9 = 1$	$z_2 + c_2 = 4 + 1 = 5$
	$x_1 = x_3 = x_6 = x_8 = x_{10} = 0$	Solution $x_2 = x_4 = x_5 = x_7 = x_9 = 1$
		$x_1 = x_3 = x_6 = x_8 = x_{10} = 0$

On this example, it appears that some significant simplifications are obvious (for instance, as soon as a neighbour v_j of v_i is labeled, v'_i may be labeled at the next step); others are less easy to set, and suppose more material about "alternating chains". They lead to an improved version of the labeling procedure, which

(*i*) works directly on the original graph *G*,

(*ii*) assigns labels + or - to the vertices, the main part of this being done along trees,

(*iii*) uses an immediate identification of a subset of vertices that might not be integer-valued, and then are discarded,

(*iv*) assigns at most two labels to the remaining vertices.

For seek of simplicity, we have not included here this improved algorithm; the interested readers may write to the authors to get it.

REFERENCES

[1] BALINSKI M.L. (1970) "On Maximum Matching, Minimum Covering and their Connections" in Proceedings of the Princeton Symposium on Mathematical Programming, H. Kuhn ed., Princeton University Press.

[2] EDMONDS J. and JOHNSON E.L. (1970) "Matching: A Well-Solved Class of Integer Linear Programs" in Proc. of Calgary Int. Conf. on Comb. Structures and their Appl., Gordon and Breach, New York.

[3] FORD L.R. and FULKERSON D.R. (1962) Flows in Networks, Princeton University Press.

[4] JOHNSON E.L. (1965) "Programming in Networks and Graphs" ORC 65.1, University of California, Berkeley.

[5] LORENTZEN L.C. (1966) "Notes on Covering of Arcs by Nodes in an Undirected Network" ORC 66.1, University of California, Berkeley.

[6] NEMHAUSER G.L. and TROTTER L.E. (1974) "Vertex Packings: Structural Properties and Algorithms" Mathematical Programming, Vol 8, No 2 (April 1975) 232-248.

[7] PICARD J.C. and RATLIFF H.D. "Minimum Cuts and Related Problems" (to appear in Networks)