Belgian Journal of Operations Research, Statistics and Computer Science Vol. 33 (1,2)

A minimax assignment problem on a linear communication network

J. Karkazis

Department of Business Administration
University of the Aegean
8, Michalon str.
82100 Chios, Greece

Abstract

A system of n communication centres C_i $(i = 1, 2, \ldots, n)$ is considered. The communication centres are to be located at a set of n prespecified positions i $(i = 1, 2, \ldots, n)$ lying in sequence on a linear network. Each communication centre C_i transmits messages to every other centre C_j at a rate of t_{ij} messages per unit of time $(t_{ii} = 0, \forall i)$. The messages that centre C_i transmits to centre C_j are sensed by all intermediate centres.

In this paper a branch-and-bound method is introduced dealing with the following minimax assignment problem :

$$\text{minimise } (\max_i t_{a(i)} : a \in A)$$

where A represents the set of all possible assignments of positions i to centres $C_{a(i)}$. The term $t_{a(i)}$ represents the traffic at i which is defined as the number of messages compiled by the centre $C_{a(i)}$ (at i) per unit of time, all inward and outward messages (at i) included.

Based on numerical experience, a comparative analysis of the minimax assignment problem and the classical minisum Quadratic Assignment Problem (QAP) is performed.

Keywords : Communication network, location problem, minimax assignment

1. INTRODUCTION

A system of n centres C_u $(u=1,2,\dots,n)$ called users, are to be placed at a set on n prespecified positions i, $i=1,2,\dots,n$ lying in sequence on a linear "network" as shown in figure 1. Each user C_u generates traffic (messages) for every other user C_v $(v=1,2,\dots,n \,/\, v \neq u)$ at a fixed rate of t_{uv} messages per unit of time $(t_{uu}=0)$. The messages that the user at position i transmits to the user at j are sensed by all intermediate users.

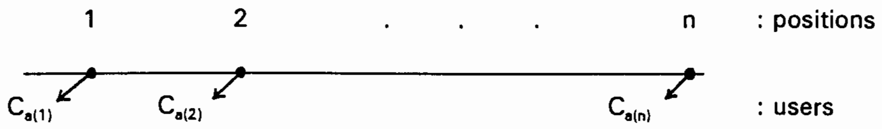


FIGURE 1. The Linear Network

We consider, next, an assignment a of position i to user a(i), $i=1,2,\dots,n$.

The traffic intensity $t_{a(i)}$ at i is defined as the number of messages sensed by the user $C_{a(i)}$ at i (per unit of time) with all inward and outward messages (at i) being included. Let $C_{a(i)}$ be the user to whom position i is assigned. Then, the traffic intensity $t_{a(i)}$ at i is given by:

$$(1) \qquad t_{a(i)} = \Sigma^{i-1}{}_{k=1}\, \Sigma^{n}{}_{l=i+1}\, (t_{a(k)a(l)} + t_{a(l)a(k)}) + \Sigma^{n}{}_{j=1} (t_{a(i)a(j)} + t_{a(j)a(i)})$$

Note at this point that the first group of summation terms represents the "passthrough" traffic whereas the second group represents the "inward-outward" traffic.

Setting, next, $T_{ij} = t_{ij} + t_{ji}$ $\forall$ i,j, it is immediate that:

$$(2) \qquad T_{ij} = T_{ji} \quad \forall\, i,j$$

Due to relation (2) relation (1) can be written:

(3) $$t_{a(i)} = \Sigma^{i-1}{}_{k=1} \Sigma^{n}{}_{l=i+1} T_{a(k)a(l)} + \Sigma^{n}{}_{j=1} T_{a(i)a(j)}$$

The key quantity in the problem is the maximum traffic position i* defined as follows:

(4) $$t_{a(i^*)} = \max_i t_{a(i)}$$

In this paper we develop a branch-bound algorithm for the minimization of the maximum traffic of the system. Formally the problem can be stated as follows:

(5) $$\text{minimize}_{a\epsilon A}\ t_a \equiv t_{a(i^*)}$$

where A is the set of all possible assignments of the users to the positions of the network. The above formulation will be called thereon "minimax" problem.

It is apparent that the classic Quadratic assignment problem (QAP) model can be applied to problems where the <u>total</u> and not the maximum traffic is to be minimized. That is QAP "minisum" model could solve problem:

(5a) $$\text{minimize}_{a\epsilon A}\ S_a \equiv \Sigma^{n}{}_{i=1} t_{a(i)}$$

The methodological and theoretical background of QAP is already rich (Finke et al (1987)). On the other hand, the optimal solutions that QAP algorithms can provide for (5a) could not guarantee that the maximum traffic is adequately low. Thus QAP can not be applied to problems where social or economic criteria are used bounding maximum "nuisance" or "load" exerted on the users of the system (Stallings (1984)). A typical application arises when a single elevator is to be replaced by two elevators, each covering continuous subsets of floors. It might be reasonable to site the connecting landing so as to minimize the traffic intensity on the busier elevator (Boffey and Karkazis (1989)).

More specifically , if {1,2,...,i} and {i,i+1,...,n} are the floors served by the first and second elevator respectively and t_{ij} represents the traffic intensity from i to j floor (e.g. it may represent the number of people expected to travel from floor i to floor j during a given period of time) then the above problem can be stated formally as follows:

(5b) $\min_i \max\{T_i^1, T_i^2\},\ T_i^1 = \Sigma^i_{k=1}\Sigma^n_{\lambda=1}(t_{k\lambda} + t_{\lambda k}),\ T_i^2 = \Sigma^n_{k=i}\Sigma^n_{\lambda=1}(t_{k\lambda} + t_{\lambda k})$

where T_i^1, T_i^2 expresses the traffic load of elevator 1, 2 respectively. The above formulation can be immediately generalized for the general case of n elevators.

The problem that will be tackled in this paper is a discrete one in the sense that centres or users occupy discrete positions in the linear network. Boffey and Karkazis (1989) proposed and analyzed a continuous version of problem (5). The problem regarded placing of bridges to alleviate congestion in a heavily loaded local-area computer network. They suggested that the generalization to that of placing (transfer) centres in a situation in which demand is arranged linearly may also be profitable. They focused their attention to models in which demand is proportional that is, per unit of time the average traffic from users in the interval (x,x+dx) to users in the interval (y,y+dy) of the linear network is p^2dxdy provided the two intervals are disjoint (p is a constant expressing the demand rate per unit of time and unit of interval length).

2. SOME THEORETICAL RESULTS

Set $t_{opt} = \min_{a \epsilon A} t_a$, $N = \{1,2,\ldots,n\}$ and $N_k = \{1,2,\ldots,k\}$, $k \leq n$ and consider a partial assignment:

(6) position i ------------------> user $C_{ak}(i)$, $i \epsilon N_k$

Denote by $A[a_k]$ the set of all assignments (of the set N onto itself) satisfying the following condition:

(6a) $\forall\ x \epsilon A[a_k]$ x -------------------> $x(i) = a_k(i)$, $i \epsilon N_k$

Given the set $A[a_k]$ we determine lowerbounds for the quantity:

$$t_{A[ak]} = \min_{x\epsilon A[ak]} t_x \tag{7}$$

Note that quantity $t_{A[ak]}$ represents the optimal value of the node (subproblem) of the branch-bound process associated with the set N_k of the preassigned positions and partial assignment a_k.

Let $\bar{N}_k = N\text{-}N_k$, $S_k = \{a_k(1), a_k(2), \dots, a_k(k)\}$ and $\bar{S}_k = N\text{-}S_k$.

Given, now, some assignment $x\epsilon A[a_k]$, relation (4) can be written as follows:

$$t_{x(i^*)} = \max\{\max_{i\epsilon Nk} t_{x(i)}, \max_{i\epsilon \bar{N}k} t_{x(i)}\} \tag{8}$$

and due to (6a):

$$t_{x(i^*)} = \max\{\max_{i\epsilon Nk} t_{ak(i)}, \max_{i\epsilon \bar{N}k} t_{x(i)}\} \tag{9}$$

Set, next,

$$A_x = \max_{i\epsilon Nk} t_{ak(i)}, \quad B_x = \max_{i\epsilon \bar{N}k} t_{x(i)} \tag{9a}$$

From (3) and (9) we get:

$$A_x = \max_{i\epsilon Nk}[\Sigma^{i-1}_{\mu=1} (\Sigma^{\kappa}_{\lambda=i+1} T_{ak(\mu)ak(\lambda)} + \Sigma^{n}_{\lambda=\kappa+1} T_{ak(\mu)x(\lambda)})$$

$$+ \Sigma^{k}_{\lambda=i} T_{ak(i)ak(\lambda)} + \Sigma^{n}_{\lambda=\kappa+1} T_{ak(i)x(\lambda)}]$$

$$B_x = \max_{i\epsilon \bar{N}k}[\Sigma^{k}_{\mu=1} \Sigma^{n}_{\lambda=i+1} T_{ak(\mu)x(\lambda)} + \Sigma^{i-1}_{\mu=\kappa+1} \Sigma^{n}_{\lambda=i+1} T_{x(\mu)x(\lambda)}$$

$$+ \Sigma^{\kappa}_{\lambda=1} T_{a(\lambda)x(i)} + \Sigma^{n}_{\lambda=\kappa+1} T_{x(\lambda)x(i)}]$$

Set, now,

$$A^{*}_{x} = \max_{i \in Nk}\{A^{i}_{pt} + A^{i}_{xpt} + A^{i}_{ot} + A^{i}_{xot}\} \quad \text{where}$$

(10a) $A^{i}_{pt} = \Sigma^{i-1}_{\mu=1} \Sigma^{\kappa}_{\lambda=i+1} T_{ak(\mu)ak(\lambda)}$ (completely determined "passthrough i traffic")

(10b) $A^{i}_{xpt} = \Sigma^{i-1}_{\mu=1} \Sigma^{n}_{\lambda=k+1} T_{ak(\mu)x(\lambda)}$ (partially determined "passthrough i traffic")

(10c) $A^{i}_{ot} = \Sigma^{k}_{\lambda=i} T_{ak(i)ak(\lambda)}$ (completely determined "outward traffic")

(10d) $A^{i}_{xot} = \Sigma^{n}_{\lambda=k+1} T_{ak(i)x(\lambda)}$ (partially determined "outward traffic")

and

$$B^{*}_{x} = \max_{i \in \bar{N}k} (B^{i}_{xpt} + B^{i}_{xxpt} + B^{i}_{xot} + B^{i}_{xxot}) \quad \text{where}$$

(11a) $B^{i}_{xpt} = \Sigma^{k}_{\mu=1} \Sigma^{n}_{\lambda=i+1} T_{ak(\mu)x(\lambda)}$ (partially determined "passthrough traffic")

(11b) $B^{i}_{xxpt} = \Sigma^{i-1}_{\mu=k+1} \Sigma^{n}_{\lambda=i+1} T_{x(\mu)x(\lambda)}$ (non-determined "passthrough traffic")

(11c) $B^{i}_{xot} = \Sigma^{k}_{\lambda=1} T_{a(\lambda)x(i)}$ (partially determined "outward traffic")

(11d) $B^{i}_{xxot} = \Sigma^{n}_{\lambda=k+1} T_{x(\lambda)x(i)}$ (non-determined "outward traffic")

Given the above terminology we get the following results:

RESULT 1. A^{i}_{xpt} is totally known since

(12) $A^{i}_{xpt} = \Sigma^{i-1}_{\mu=1} \Sigma_{\lambda \in \bar{S}k} T_{ak(\mu)\lambda}$

PROOF. Immediate consequence of the inner summation of A^{i}_{xpt} over all non-assigned users (set $\bar{S}_{k}$).

RESULT 2. A^i_{xot} is totally known since

(13) $$A^i_{xot} = \Sigma_{\lambda \epsilon \bar{S}k} T_{ak(i)\lambda}$$

PROOF. As in result 1.

RESULT 3.

(14) $$B^i_{xpt} \geqslant \Sigma^{n-i}_{w=1} \hat{K}_w(a_k)$$

where $\hat{K}_w(a_k)$: $w \epsilon \bar{S}_k$ are the elements of the set $\{K_w[a_k]: w \epsilon \bar{S}_k\}$ ordered in increasing magnitude and $K_w[a_k] = \Sigma^k_{\mu=1} T_{ak(\mu)w} : w \epsilon \bar{S}_k$.

PROOF. Since we have uncertainty over the user assigned to position i we select the one (among the unassigned) associated with the least passthrough traffic. The passthrough traffic will comprise the traffic sent by assigned users (positions 1,2,...,k) to unassigned users at positions $\{i+1,...,n\}$ that is, it is the traffic directed to n-i users. To each one unassigned user at j ($j \epsilon \bar{N}_k$) traffic T_j from assigned users is calculated and the smallest n-i T_j are summed up to give a lowerbound.

RESULT 4.

$$B^i_{xxpt} \geqslant \Sigma^{ntrans}_{w=1} \hat{K}_w$$

where $ntrans = (n-i)\phi$

and $$\phi = \begin{cases} i-k & \text{if } i > k+1 \\ 0 & \text{if } i \leqslant k+1 \end{cases}$$

Note that ntrans is the number of message transactions among non-assigned users that pass through position i and $\hat{K}_{\lambda\mu}$ $\lambda,\mu \epsilon \bar{S}_k$ ($\lambda \neq \mu$) are the elements of the set $K_{\lambda\mu} = T_{\lambda\mu}$ $\lambda,\mu \epsilon \bar{S}_k$ ($\lambda \neq \mu$) ordered in increasing magnitude.

PROOF. It is immediate from relation (11b) that B^{i}_{xxpt} consists of $(n\text{-}i)\phi$ terms. Since we have uncertainty both over the user assigned to position i and over the source and destination of the "passthrough traffic" (case of non-determined "passthrough traffic") we sum up the $(n\text{-}i)\phi$ smallest elements of the set $K_{\lambda\mu} = T_{\lambda\mu}\ \ \lambda,\mu \epsilon \bar{S}_k\ (\lambda \neq \mu)$ in order to get a lowerbound for quantity B^{i}_{xxpt}.

RESULT 5.

$$B^{i}_{xot} \geqslant \min_{w\epsilon\bar{S}k} \Sigma^{k}_{\lambda=1}\ T_{a(\lambda)w}$$

PROOF. Immediate.

RESULT 6.

$$B^{i}_{xxot} \geqslant \min_{w\epsilon\bar{S}k} \Sigma_{\mu\epsilon\bar{S}k}\ T_{\mu w}$$

PROOF. Immediate.

As an immediate consequence of the above results we get the following:

$$(15) \qquad A_x = \max_{i\epsilon Nk} \underbrace{[A^{i}_{pt} + \Sigma^{i-1}_{\mu=1}\ \Sigma_{\lambda\epsilon\bar{S}k}\ T_{ak(\mu)\lambda} + A^{i}_{ot} + \Sigma_{\lambda\epsilon\bar{S}k}\ T_{ak(i)\lambda}]}_{\text{completely determined (see results 1 \& 2)}}$$

$$(16) \qquad B_x \geqslant B^{*}_{x} \equiv \Sigma^{n-ic}_{w=1}\ \hat{K}_{w} + \Sigma^{ntrans}_{w=1}\ \hat{K}_{w} + \min_{w\epsilon\bar{S}k} \Sigma^{k}_{\lambda=1}\ T_{a(\lambda)w} + \min_{w\epsilon\bar{S}k} \Sigma_{\mu\epsilon\bar{S}k}\ T_{\mu w}$$

RESULT 7. Given an assignment $x\epsilon A[a_k]$ (node of the b-b tree) then the maximum traffic $t_{x(i^*)}$ for this node satisfies the following inequality:

$$(16a) \qquad t_{x(i^*)} \geqslant L_{ak} = A_x + B^{*}_{x}$$

with A_x, B^{*}_{x} defined as in (15) and (16).

PROOF. Immediate from (15) and (16).

3. THE BRANCH-BOUND PROCESS

3.1 Initial Assignment

In order to get an initial incumbent value, we adopt the following heuristic. Consider quantities:

(17) $T_i = \Sigma^n_{j=1} T_{ij}$

Order T_i in increasing magnitude to obtain the sequence:

$T^{max}_{wi} \quad i = 1,\ldots,n$

To location i: $(i = 1,2,\ldots,[n/2])$ assign centre $a(i) = w_{2i-1}$ whereas to location i: $(i = [n/2] + 1,\ldots,n)$ assign centre $a(i) = w_{2(n-i+1)}$.

The above process assigns centres that transmit large quantities of messages to central positions:

locations:	1	2	3	• • •	k	k+1	• • •	n-1	n
centres :	w_1	w_3	w_5		w_{2k-1}	w_{2k}		w_4	w_2

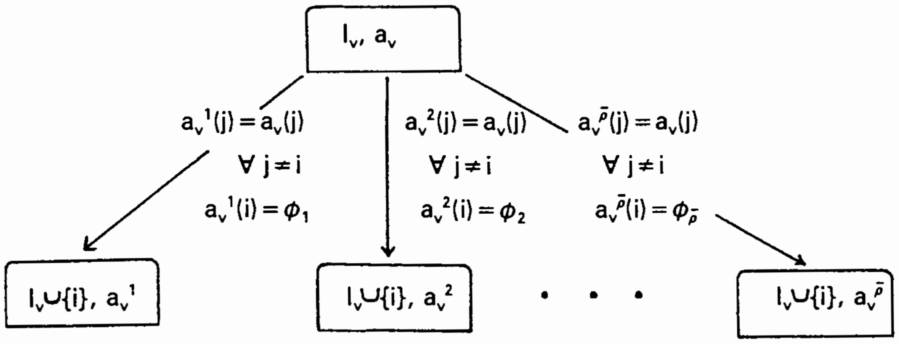
where $k = [n/2]$.

3.2 Fathoming Process

Each node v of the tree is associated with a pair (I_v, a_v) where I_v is the set of locations having centres assigned to them and a_v is the corresponding (current) assignment.

Calculate lower, L_{av}, bound for T_{av} (see relation (16a)). Let T_{opt} be the value of the current best solution up to date.

If $T_{opt} < L_{av}$ then we fathom all branches emanating from this node otherwise we consider an index $i \notin I_v$ and develop the following branches:



where $\{\phi_1, \phi_2, \ldots, \phi_{\bar{\rho}}\}$ is the set of centres not yet assigned.

3.3 Branching Policy and Priority Rules

Numerical experience with this type of problem suggests that the number of nodes (and consequently memory requirements) increases dramatically with the value of n, especially for $n \geqslant 10$. A depth first search aiming at reaching quickly terminal nodes associated with good solutions could alleviate to some extent this problem.

At each node v we evaluate a solution for the problem (that is a complete assignment of all locations) by applying the technique used in the initial assignment to the set of non-assigned locations at node v (set $\bar{I}_v = N - I_v$). The value, U_{av}, of this solution represents an upper bound for the value, U^*_{av}, of the optimal solution of the problem associated with this node.

In the context of the mentioned search strategy we experimented with 3 different priority (branching) rules:

(i) branching from the "least lowerbound" node: $\min_{v \in V} L_{av}$, where V is the set of end nodes.
(ii) branching from the "least upperbound" node: $\min_{v \in V} U_{av}$
(iii) branching from the node associated with the minimum value of quantity

$$M_v = (L_{av} + U_{av})/2$$

Note that quantity M_v, as numerical experience suggests, represents a good approximation to the optimal value U^*_{av} associated with node v.

From the above rules, rule (iii) gave by far the best results (results of table 1).

4. NUMERICAL RESULTS AND CONCLUSIONS

Traffic data between centres is represented by random numbers in the range [0,5000]. The previously described algorithm "minimax" was tested for n= 4, 5, 6, 7, 8, 9, 10, 12. Notice at this point that optimally solving QAP becomes extremely difficult for $n \geqslant 12$. Solving a n=10 problem may take up to few thousand branching nodes whereas the optimal solution for n=12 requires a number of nodes exceeding 10.000 (Finke et al (1987)).

The following table gives the basic numerical features of the proposed algorithm performance:

n	initial assignment	initial value	optimal assignment	optimal value	initial assign.	nodes
4	3,4,1,2	8.851	3,4,1,2	8.851	0.0 %	17
5	3,2,5,1,4	14.705	5,4,2,1,3	14.362	2.4 %	32
6	2,6,4,5,3,1	28.097	6,3,5,1,4,2	20.535	36.8 %	19
7	2,5,7,4,6,3,1	43.423	7,6,5,1,2,3,4	33.629	29.1 %	138
8	1,7,6,5,8,3,4,2	55.166	8,3,7,2,1,6,4,5	41.427	35.6 %	293
9	1,8,4,7,3,9,6,5,2	69.536	9,5,2,1,7,4,3,8,6	50.641	20.5 %	241
10	1,8,10,2,5,3,4,6, 9,7	66.598	10,5,3,7,8,9,2,1, 6,4	59.261	12.4 %	1323
12	1,6,3,7,8,4,2,10, 9,11,12,5	103.675	1,6,3,7,8,4,2,10, 9,11,12,5	81.558	27.0 %	21407

TABLE 1. Performance of the "Minimax" Algorithm

Although we could not directly compare the "minimax" algorithm with those proposed for the "minisum" QAP we could conclude that:

a. The algorithm "minimax" solves optimally all problems tested.

b. The initial assignment heuristic although very trivial gives a reasonably good starting solution. Note that, the relative error of the Gilmore-Lawler QAP bounds (Finke et al (1987)) for n=6 is 4.6 %, for n=8 is 13.1 %, for n=10 is 13.9 % and for n=12 is 14.7%.

c. The number of branching nodes is relatively low and compares reasonably with the existing QAP branch-bound algorithms.

Finally, in order to numerically compare the "minimax" and QAP problems we have run the above set of test problems using as initial solution the optimal solution of the corresponding QAP. The results were surprising (see table 2). The relative error of the optimal QAP assignment used as initial solution for "minimax" problem didn't exceed 7.2% whereas in 2 cases out of 8 tested gave the optimal solution of "minimax" problem, and in all cases the number of branching nodes was drastically reduced.

n	initial assignment (optimal assignment of QAP)	initial value	optimal assignment	optimal value	% error of initial assignment	branch. nodes
4	3,4,1,2	8.851	3,4,1,2	8.851	0.0 %	17
5	4,5,1,2,3	14.377	5,4,2,1,3	14.362	0.1 %	32
6	6,3,5,4,1,2	20.535	6,3,5,1,4,2	20.535	0.0 %	9
7	5,6,7,4,3,1,2	35.002	7,6,5,1,2,3,4	33.629	4.1 %	138
8	4,6,8,5,7,2,1	44.474	8,3,7,2,1,6,4,5	41.427	7.2 %	181
9	8,6,4,3,7,9,5,2,1	51.697	9,5,2,1,7,4,3,8,6,3	50.641	1.9 %	104
10	5,1,6,9,10,8,4,7, 2,3	60.598	10,5,3,7,8,9,2,1, 6,4	59.261	2.3 %	374

TABLE 2. Performance of the "Minimax" Algorithm Using as Initial Assignment the Optimal QAP Assignment

Reversing the argument, we could conclude that "minimax" algorithm solution can provide a good initial solution for QAP.

Although the comparative numerical experience with the min-max and min-sum versions of the assignment problem could not suggest that the use of min-max problem as a preprocessing step for the solution of the min-sum problem (and vice-versa) can make it easier, when both problems are to be solved at the same time then the use of the optimal solution of one of the problems as initial solution for the other can significantly reduce the solution burden (number of branching nodes and hence solution time).

REFERENCES

Finke G., Burkard R. and F.Rendl, "Quadratic assignment problems", Annals of Discrete Mathematics 31 (1987) 61-82

Stallings W., "Local Networks: An Introduction". MacMillan, New York (1984)

Boffey B. and J.Karkazis, "Location of transfer centres of a communication network with proportional traffic", J.Opl.Res.Soc. vol.40 (1989) 729-734