A TWO STAGE HEURISTIC FOR DESIGNING DATA COMMUNICATION NETWORKS

By

R. Sridharan

&

N. Ravi

WP885

WP
1990
(885)

W P No. 885
August 1990

The main objective of the working paper series of the IIMA is to help faculty members to test out their research findings at the pre-publication stage.

INDIAN INSTITUTE OF MANAGEMENT
AHMEDABAD-380 056
INDIA

PURCHASED
APPROVAL
GRATIS/EXCHANGE
PRICE
ACC NO.
VIKRAM SARABHAI LIBRARY
I. I. M. AHMEDABAD.

A Two Stage Heuristic for Designing Data Communication Networks

R. Sridharan
Indian Institute of Management, Ahmedabad 380 056, India
and

N. Ravi
Telecom Commission, New Delhi, India

Abstract

In this paper, we present a two-stage heuristic which identifies the most economical way to connect the nodes of a data communication network. We formulate this network design problem as a star-star concentrator location problem. To solve this problem, a Lagrangian relaxation procedure is proposed. The first stage of the procedure, namely the solution to the relaxed problem, identifies the optimal locations of the transit nodes. Then a greedy type heuristic, as well as an optimal procedure are used to identify the linkage of these transit nodes to the other nodes. Computational results are provided and the results are also compared with another method.

Introduction

A data communication network consists of a number of nodes (points, terminals etc.) to be linked with each other so that data (of any desired form) can be transmitted from one node to another within the network. Normally, these nodes are dispersed over a wide geographical area typically consisting of many clusters. Figure 1 gives a schematic representation of the nodes and a possible network linking these nodes.

In practice the network usually consists of transit nodes, or concentrator nodes, which are linked to each other by high speed lines, or digital links. The rest of the nodes are linked to these transit nodes via low speed lines or analogue links. A node is linked to only one transit node. Normally, the transit nodes are linked in the form of a star, which forms the main network, to a central node. The ordinary nodes are then linked to the transit nodes also in the form of a star network, refer to figure 1.

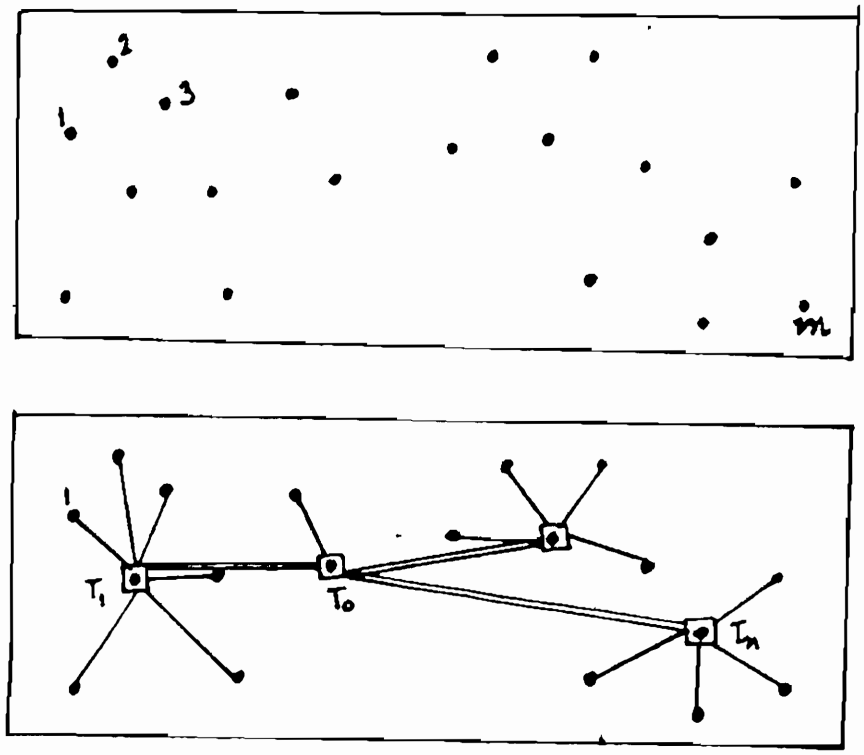

Figure 1. A typical network problem and solution.

The design of the data communication network involves the optimal way of connecting these nodes together. The objective is usually to minimize the cost in linking the nodes together. There exist a number of methods to achieve such a network optimization. Some of these techniques use exact procedures Chandy and Russel (1972), Matsui (1978), and Tabourier (1973), while others use heuristic methods to arrive at an approximate solution Bahl and Tang (1972), Karnaugh (1976), Hoang (1982), Greenhop and Campbell (1984), and Mirzaian (1985).

A heuristic procedure is proposed in this paper based on the Lagrangian relaxation technique. A formulation of the network design problem is given in the next section. In the following section the heuristic procedure is described and then the computational results are discussed and in the final section we provide some concluding remarks.

Problem Formulation

In this section a formulation for the data communication network design problem as described above will be given. This network design problem is formulated as a star-star concentrator location problem as described in Mirzaian (1985). Let

P_i refer to the node site i,
T_j refer to the transit node site j,
T_0 refer to the central node site,
c_{ij} = the cost of connecting an ordinary node P_i to a transit node T_j,
f_j = the cost of installing, operating and connecting a transit node T_j to the central site T_0,
k = the capacity of a transit node, i.e. no more than k nodes be connected to a transit,
m = the number of nodes, also the capacity of the central node T_0,
n = the number of transit node sites.

Define

$$x_{ij} = \begin{cases} 1 & \text{if } P_i \text{ is connected to } T_j \\ 0 & \text{otherwise} \end{cases}$$

and

$$y_j = \begin{cases} 1 & \text{if } T_j \text{ is a selected transit node} \\ 0 & \text{otherwise.} \end{cases}$$

$1 <= i <= m, \quad 0 <= j <= n.$

The network design problem can now be formulated as

$$Z = \min \sum_{i} \sum_{j} c_{ij} x_{ij} + \sum_{j} f_j y_j \quad \ldots \quad (P)$$

subject to

$$\sum_{j=0}^{n} x_{ij} = 1, \; i = \;, \ldots, m \quad \ldots \quad (1)$$

$$\sum_{i} x_{ij} \le k y_j, \quad = 1, \ldots, n \ldots \quad (2)$$

$$x_{ij}, y_j \in \{0,1\} \quad \ldots \quad (3).$$

The objective is to minimize the cost of installing, connecting to the central node and operating the transit nodes and the cost of connecting each node P_i to one of the chosen transit nodes. Constraint (1) along with constraint (3) ensures that each ordinary node P_i is connected to only one transit node. Constraint (2) ensures that a tranist node T_j is connected to not more than k ordinary nodes. The central node T_0 has capacity m and therefore does not get affected by the capacity constraints relating to the transit nodes.

In the next setion, a Lagrangian based heuristic procedure to find an approximate solution to this network design problem is given.

VIKRAM SARABHAI LIBRARY
INDIAN INSTITUTE OF MANAGEMENT
VASTRAPUR, AHMEDABAD-380056

Heuristic Procedure

In this section, we propose a heuristic procedure for the network design problem described above. This heuristic procedure is based on a Lagrangian relaxation of the problem. The heuristic solution involves the identification of a lower bound, namely the solution of the relaxed problem, and the identification of an upper bound, that is, a feasible solution to the original problem. The methods adopted for identifying the lower bound, the upper bound and the heuristic procedure for the network design problem will be given now.

Lower Bound

A Lagrangian relaxation LR(u) of the network design problem is obtained by dualizing the demand constraints (1) with multipliers u_i. The relaxed problem so obtained is

$$Z(LR(u)) = \min \sum_i \sum_j (c_{ij} - u_i)\, x_{ij} + \sum_j f_j y_j + \sum_i u_i$$
subject to
$$\sum_i x_{ij} <= k y_j, \text{ for all } j,$$
$$x_{ij},\ y_j \in \{0,1\}, \text{ for all } i,\ j.$$

The solution to LR(u) for a given u is obtained as follows. The problem LR(u) separates into n independent (0,1) knapsack problems LR(u;j) on x_{ij} for each j, see Bitran et al. (1977). This can be written as

$$Z_j(LR(u;j)) = \min \sum_i (c_{ij} - u_i) x_{ij} + f_j$$
subject to
$$\sum_i x_{ij} <= k$$
$$x_{ij} = 0,\ 1 \text{ for all } i.$$

Let x_{ij}^* be the optimal solution for these knapsack problems. Then an optimal solution of LR(u) is given by the following theorem.

Theorem 1 : An optimal solution of LR(u) is

$$y_j = \begin{cases} 1 \text{ if } f_j + \sum_i (c_{ij} - u_i) x_{ij}^* < 0 \\ 0 \text{ otherwise} \end{cases} \qquad \ldots \quad (6)$$

$$x_{ij} = \begin{cases} x_{ij}^* \text{ if } y_j = 1 \\ 0 \text{ otherwise.} \end{cases} \qquad \ldots \quad (7)$$

Proof : First we note that LR(u) breaks up into n independent (0,1) knapsack problems, one for each j. We denote these knapsack problems as $Z_j(LR(u;j))$. In each $Z_j(LR(u;j))$ it is clear that $x_{ij}^* = 0$ for $i : (c_{ij} - u_i) >= 0$. Now, for all i : $(c_{ij} - u_i) < 0$, x_{ij}^* is given by the solution to the knapsack problem. Now, consider $Z(LR(u))$. This can be written as

$$Z(LR(u)) = \min_{y_j=0,1} \sum_j Z_j(LR(u;j))y_j + \sum_i u_i$$

$$= \min_{y_j=0,1} \sum_j [(c_{ij}-u_i)x_{ij}^* + f_j]y_j + \sum_i u_i$$

It can be seen that $y_j = 0$ when $(c_{ij}-u_i)x_{ij}^* + f_j >= 0$ and $y_j = 1$ otherwise. This completes the proof of the theorem.

Therefore, a lower bound for the network design problem is found by simply solving at most n knapsack problems for a given value of u_i. But, studies on problems with similar structure, as the Capacitated Plant Location Problem have shown that the lower bound obtained by adding suitable surrogate constraints are tighter, see Cornuejols, Sridharan and Thizy (1989), and Sridharan (1989). Therefore, we further strengthen this relaxation by adding a surrogate constraint

$$\sum_j ky_j >= m.$$

We will refer to the Lagrangian relaxation with the surrogate constraint as $LR_S(u)$ and the n independent knapsack problems on x_{ij} as $LR_S(u;j)$. It can immediately be seen that $LR_S(u;j)$ is the same problem as $LR(u;j)$.

This addition of the surrogate constraint will modify the y_j's defined in (6) but will leave the values of x_{ij}'s unaltered as given in (7). The optimal values of y_j's with the surrogate constraint can now be computed as given below.

Let x_{ij}^* be defined, as before, as the solution to $LR(u;j)$. It is immediate that x_{ij}^* will also be the solution to $LR_s(u;j)$. Once we have the x_{ij}^*'s, the following (0,1)-knapsack problem on the y_j variables results.

$$Z(LR_s(u)) = \min_{y_j=0,1} \sum_j Z_j(l \quad u;j))y_j + \sum_i u_i$$

$$= \min_{y_j=0,1} \sum_j [(c_{ij}-u_i)x_{ij}^* + f_j]y_j + \sum_i u_i$$

subject to

$$\sum_j ky_j >= m \qquad \ldots \qquad (8)$$

Let y_j^* be the optimal solution to the knapsack problem defined above. Then,

<u>Theorem</u> <u>2 :</u> An optimal solution to $LR_s(u)$ is given by the x_{ij}^* defined in (7) and y_j^* defined as the optimal solution to the knapsack problem defined above.

<u>Proof</u> <u>:</u> Omitted as the proof is very similar to the one given for theorem 1.

For a given u, a value for the lower bound is obtained for the network design problem by solving at most (n+1) knapsack problems when the proposed surrogate constraint is added. This lower bound is computed at every lagrangian iteration and the value is updated as is necessary.

<u>Upper Bound</u>

Every Lagrangian iteration to identify a given lower bound returns a set of y_j's fixed at one. The rest of the y_j's are fixed at zero. In the network design problem once the values of y_j's are fixed either at zero or one the problem reduces to a single source transportation problem as given below.

Let $J^* = \{ j : y_j$ fixed as one in the procedure$\}$

Then the single source transportation problem will be,

$$\min \sum_i \sum_{j \in J^*} c_{ij} x_{ij}$$

subject to

$$\sum_{j \in J^*} x_{ij} = 1, \quad i = 1, \ldots, m;$$

$$\sum_i x_{ij} <= k, \quad j \in J^*;$$

$$x_{ij} = (0,1).$$

The solution to this single source transportation problem provides an upper bound to the network design problem. The single source transportation problem itself is NP-hard and therefore polynomial time bounded algorithms do not exist. We use a heuristic procedure called "HRSTIC" for identifying an approximate solution for this problem. We also solved all the network design problems using an exact procedure, "TSTU", to find the solution to the single source transportation problem developed by Nagelhout and Thompson (1980). A comparison is provided in the next section.

The upper bound is not necessarily computed at every Lagrangian iteration. Upto five different combinations of y_j's fixed at one for which the upper bound has already been computed are stored. Only when the Lagrangian iteration returns a new combination of y_j's, other than the five mentioned above, the upper bound procedure is invoked. This has resulted in significant savings in computation time. The upper bound value is continuously updated and the best value is always used for computing the value of the Lagrangian multipliers in the subgradient procedure described later.

The Heuristic Procedure

The heuristic procedure for solving the network design problem is described now.

Step 1: Identify the initial upper bound with all y_j's fixed at one and solving the single source transportation problem. The solution to this transportation problem will return a set of transit locations that will be servicing the terminal nodes. Add the fixed costs of those transit locations that are used in the final solution to the objective value of the single source transportation problem to arrive at the initial upper bound Z^{UB}. Initialize the Lagrange multipliers $u_i = \min_j c_{ij}$. Go to step 2.

Step 2: Solve the (0,1) knapsack problems LR(u;j) and $LR_s(u)$ to compute the lower bound Z_{LB}. Update the lower bound if necessary. With the set of y_j's fixed at one, go to step 3.

Step 3: Update the upper bound, if necessary, using either "HRSTIC" or "TSTU". If the upper bound is equal to the lower bound, stop; an optimal solution has been found. If the number of Lagrangian iterations has exceeded 100, stop. If the lower bound converges to a particular value, stop. Else, go to step 4.

Step 4: Update the Lagrange multipliers using the subgradient approach. If all subgradients are zero, stop. Else, go to step 2.

The subgradients for u_i are computed as follows. Let x_{ij}^* and y_j^* be the optimal solutions to the Lagrangian problem. Then the subgradients NU(i) for u_i are

$$NU(i) = \sum_{j \in J^*} x_{ij}^* - 1.$$

The Lagrange multipliers are then updated as follows.

$$u_i^{k+1} = u_i^* + t_k NU(i)$$

where

$$t_k = \lambda(Z^{UB} - Z_{LB})/Norm.$$

We start with an initial λ value of 0.6 and halve the value every twelve iterations. Some computational results for this approach are provided in the next section.

Results

The procedure was coded in FORTRAN77 and run on a DEC2060 timesharing system at Carnegie-Mellon University. A code was also written in Microsoft FORTRAN to run on a Personal Computer. The PC results were obtained on a IBM PC/XT machine. The test problems were obtained from Mirzaian (1985) to make comparisons. Both the "HRSTIC" procedure and the "TSTU" procedure were used for identifying the upper bounds. The results for the test problems are provided in Table 1.

The test problems using the "HRSTIC" procedure had been solved on an IBM PC/XT machine. As we can see from the results this procedure is comparable to that of Mirzaian (1985) in terms of the solution accuracy. While the "TSTU" procedure is computationally more time consuming it definitely generates superior upper bounds. The "HRSTIC" procedure provides competing results in many of the test problems. Also, since the Lagrangian heuristic provides an upper and a lower bound the decision maker is aware of the extent to which his/her solution is close to the best that can be achieved.

Conclusion

We conclude this paper by noting that the Lagrangian based heuristic is very efficient is solving the network design problem. It is true that the formulation of the network design problem ignores some of the real life constraints and conditions. For example, the cost of a transit node is function of the number of terminal nodes to which it is connected. But, if the formulation were to consider such realities then the resulting problem may become computationally very hard to handle. In spite of such assumptions and approximations, we feel that the decision maker will benefit immensely by using the procedure we have presented here as a tool for decision making.

Table 1*

Problem Size	MIRZAIAN			LAGRANGIAN HEURISTIC TSTU			HRSTIC		
	Itns.	LB	UB	Itns.	LB	UB	Itns.	LB	UB
50X20X3	100	369	371	100	369	374	100	369	391
50X20X5	100	298	305	100	298	307	100	298	307
50X20X7	100	276	278	100	277	278	100	277	290
40X20X3	100	322	331	100	322	326	100	322	334
40X20X5	100	248	254	100	248	254	100	248	254
40X20X7	100	232	234	100	232	234	100	232	238
23X5X3	1	264	264	33	264	264	100	264	280
23X5X5	25	223	223	39	223	223	100	223	228
23X5X7	15	218	218	22	218	218	22	218	218

REFERENCES

Bahl, L.R. and Tang, D.T. (1972) "Optimization of concentrator locations in teleprocessing networks", Symp. on Computer Communication Networks and Teletraffic, Polytechnic Institute of Brooklyn, New York, USA.

Bitran, G.R., Chandru, V., Sempionski, D.E., and Shapiro, J.E., (1977) "Inverse Optimization; an application to the capacitated plant location problem", Management Science, 23, 789-810.

Chandy, K.M., and Russel, R.A. (1972) "The design of multipoint linkages in a teleprocessing tree network", IEEE Trans. on Computers, vol. C-21, No. 10, 1062-1066.

Cornuejols, G., Thizy, J.M. and Sridharan, R. (1989) "A comparison of heuristics and relaxations for the capacitated plant location problem", under revision with European Journal of Operational Research.

Greenhop, D. and Campbell, R. (1984) "Cost-201 - a procedure for the optimization of telecommunication networks", British Telecom Eng., vol. 3, 47-58.

Hoang, H.H. (1982) "Topological optimization of networks: a nonlinear mixed integer model employing generalized Bender's decomposition", IEEE Trans. on Auto. Control, Vol. AC-24, 164-169.

Karnaugh, M. (1976) "A new class of algorithms for multi-point network optimization", IEEE Trans. on Commun., Vol. COM-24, 500-505.

Matsui, S.A. (1978) "Network optimization by mixed integer programming, Nippon Univac, Tokyo, Japan.

Mirzaian, A. (1985) "Lagrangian relaxation for the star-star concentrator location problem: approximation algorithm and bounds", Networks, Vol. 15, 1-20.

Nagelhout, R.V. and Thompson, G.L. (1980) "A single source transportation algorithm", Journal of Computers and Operations Research, Vol. 7, 185-199.

Sridharan, R. (1989) "A Lagrangian heuristic for the capacitated plant location problem with single source constraints", submitted to European Journal of Operational Research.

Tabourier, Y. (1973) "All shortest distances in a graph: an improvement to Dantzig's inductive algorithm", Discrete Mathematics, Vol. 4, 83-87.