THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 61, No. 8, October 1982
Printed in U.S.A.

Application of Graph Theory to the Solution of a Nonlinear Optimal Assignment Problem

By M. MALEK-ZAVAREI

(Manuscript received November 5, 1981)

This paper poses an assignment problem with a nonlinear objective function. It is formulated as an integer programming problem and a graph model is used to determine its exact solution. Application of the problem in long-range homing in telephone networks is also discussed.

I. INTRODUCTION

Consider a process in which at each stage an *originating center* must be connected to exactly one of a set of *terminating centers*. The originating center has a *load* and each terminating center has a *capacity*. The originating center can be connected to a terminating center only if the terminating center has enough capacity for its load. The originating center's load and the terminating centers' capacities vary with time (i.e., from one stage to another), but they are assumed to be known at all times. The problem is to determine the optimal connection configuration at each stage of time. The costs involved are the *transmission cost* and the *rearrangement cost*. Both costs are nonlinear functions of the originating center's load. The transmission cost is the cost of connection of the originating center to a terminating center. The rearrangement cost is incurred if, in transition from one stage to the next, the connection of the originating center to a terminating center has to be changed because of insufficient capacity.

Such an assignment problem may be encountered in many applications. One important application arises in the design of hierarchical telephone networks.[1–3] The process of connecting a switching center to a center in the next level of hierarchy in the backbone route of such networks is called *homing*. The problem of determining the optimal homing configuration of a switching center over several stages during a study period can be formulated as the above assignment problem.[4]

In such a case the originating center could be an end office and the terminating centers could be toll centers (i.e., switching centers in the next level of hierarchy). The load of the originating center would then correspond to the traffic volume of the end office and the capacities of terminating centers would correspond to the switching capacities of the toll centers. The transmission cost would be the cost of trunks homing the end office on a toll center and the rearrangement cost would correspond to the cost of changing the homing configuration.

In this paper a graph model will be developed for the above nonlinear assignment problem. The solution to the problem will be converted to determining a shortest path on this graph. The algorithm developed for the solution is easily programmable on the digital computer.

II. FORMULATION OF THE PROBLEM

Consider N stages of time indicated by $t = 1, 2, \cdots N$. Let $TC(t)$ represent the set of available terminating centers at stage t. The originating center OC must be connected to a terminating center TC_k in this set at stage t. Let

$$x_k(t) = \begin{cases} 1 & \text{if } OC \text{ homes on } TC_k \in TC(t) \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

Indicate the load of the originating center at stage t by $S(t)$ and the capacity of terminating center TC_k at stage t by $C_k(t)$. The costs and the constraints of the problem can then be formulated as follows.

2.1 Costs

The transmission cost per unit distance is assumed to be a nonlinear function f of the originating center's load, as shown in Fig. 1. It includes a fixed cost that is independent of the originating center's load. (In the case of the homing problem in telephone networks, this fixed cost would represent the cost of preparing for establishing a transmission facility, e.g., laying a cable.) The fixed cost will be incurred only for a new connection. For increasing the capacity of an existing connection, only the incremental cost will be incurred.

Thus, the total transmission and rearrangement costs can be formulated as

$$\sum_{t=1}^{N} \sum_{TC_k \in TC(t)} \{f[S(t)] - f[S(t-1)]x_k(t-1)\}x_k(t)d_k \tag{2}$$

where $S(0) = 0$ and d_k is the distance between OC and TC_k. Note that $S(t)$ is known a priori for $t = 1, 2, \cdots, N$.

Equation (2) includes the rearrangement cost, i.e., the cost incurred if in transition from one stage to the next the terminating center assignment changes. However, if such a change occurs, the old con-

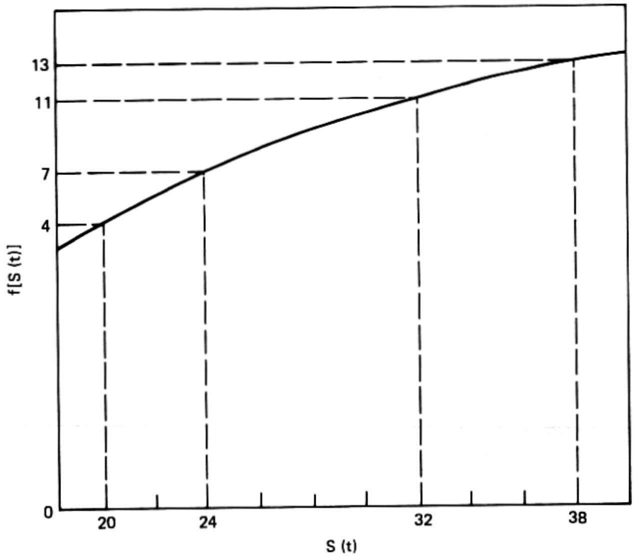

Fig. 1—Transmission cost per unit distance as a function of the originating center's load.

nection will represent a saving in the cost that must be considered. This saving is, in general, a nonlinear function of the rearranged load. We call this function $g[S(t)]$, which is obtained empirically for each application. Thus, eq. (2) must be modified as follows to represent the total transmission and rearrangement costs

$$\sum_{t=1}^{N} \sum_{TC_k \in TC(t)} \{f[S(t)] - f[S(t-1)]x_k(t-1)\}x_k(t)d_k$$
$$- \sum_{t=1}^{N-1} \sum_{TC_k \in TC(t)} \frac{1}{2} g[S(t)] |x_k(t+1) - x_k(t)|. \qquad (3)$$

The reason for including the coefficient ½ in the second part of eq. (3) is as follows. If in transition from one stage to the next the terminating center assignment changes from TC_k to $TC_{k'}$, the absolute value term will contribute 2 instead of 1 (1 for k and 1 for k').

2.2 Constraints

Since the originating center must be connected to exactly one terminating center at each stage, we must have

$$\sum_{TC_k \in TC(t)} x_k(t) = 1 \quad \text{for } t = 1, 2, \cdots, N. \qquad (4)$$

Also, the load of the originating center cannot exceed the capacity of the terminating center to which it may be connected; thus

$$S(t)x_k(t) \leq C_k(t) \quad \text{for } all\ TC_k \in TC(t),\ t = 1, 2, \cdots, N. \tag{5}$$

The problem can then be formulated as follows. Minimize the objective function (3) subject to constraints (4), (5) and

$$x_k(t) = 0\ or\ 1 \quad \text{for } all\ TC_k \in TC(t),\ t = 1, 2, \cdots, N. \tag{6}$$

III. A GRAPH MODEL FOR THE PROBLEM

The above problem is a nonlinear integer programming problem.[5] For most practical applications this problem will have a considerable number of variables and constraints. For example, for 10 stages and 8 terminating centers there will be $8 \times 10 = 80$ variables and $10 + 8 \times 10 = 90$ constraints in the problem. The nonlinearity of the cost function and the large number of variables and constraints render the available standard integer programming techniques (such as the branch and bound method[5]) impractical for the solution of this problem. In this section a graph model will be developed for the problem, which aids in obtaining an exact solution for it.

Define a matrix A whose rows and columns correspond to stages and terminating centers, respectively. The (t, k) element of matrix A is defined as

$$a(t, k) = \begin{cases} 1 \text{ if } OC \text{ can be connected to } TC_k \text{ at stage } t \\ 0 \text{ otherwise.} \end{cases} \tag{7}$$

Element $a(t, k)$ of matrix A can easily be obtained by determining whether $TC_k \in TC(t)$ has enough capacity for the originating center at stage t. Hence, constraint (5) will be used in the construction of matrix A. To incorporate constraint (4) and to perform the required minimization, define a directed graph[6] G whose nodes are elements 1 of matrix A. The arcs of G connect node pairs in consecutive rows of A in the direction that t increases. Then each path in G that connects a node in the first row of matrix A to a node in its last row forms a feasible solution to the problem where each node (t, k) of graph G included in such a path corresponds to $x_k(t) = 1$. Note that since exactly one node in each row of matrix A is included in such a path, constraint (4) will be satisfied.

Costs will now be assigned to the nodes and the arcs of graph G in such a way that the solution of the problem converts to solving a shortest-path problem on G. The cost associated with a vertical arc in graph G (i.e., an arc connecting a node pair in the same column of matrix A) is zero. A cross arc of graph G (i.e., an arc connecting a node pair in different columns of matrix A) from row t to row $t + 1$ has cost $-\frac{1}{2}g[S(t)]$. Each node of graph G corresponding to terminating center TC_k and stage t is assigned the cost $\{f[S(t)] - f[S(t-1)]x_k(t-1)\}d_k$. If the arc connecting the predecessor node of (t, k) to it is a cross arc, then the cost of node (t, k) is $f[S(t)]d_k$. If the arc connecting the predecessor node of (t, k) to it is a vertical arc, then the cost of node (t, k) is $\{f[S(t)] - f[S(t-1)]\}d_k$. With these assignments, the costs of the nodes and the arcs in graph G correspond, respectively, to the first and the second parts in expression (3). Hence, to solve the problem, a minimum cost path in graph G from nodes in the first row to nodes in the last row of matrix A must be found.

IV. AN ALGORITHM FOR THE SOLUTION

Since graph G includes no cycles, a labeling method similar to that used in the "shortest-path algorithm"[6] may be employed to determine paths of minimum cost. Label each node in the first row of A by its corresponding cost. The labels of nodes in the other rows of matrix A will be determined according to the following rule:

$$\begin{matrix}\text{Label of}\\ \text{node } (t, k)\end{matrix} = \min_{\text{all predecessor nodes}} \left\{ \begin{matrix}\text{Node label of its}\\ \text{predecessor node}\\ +\text{ cost of the arc}\\ \text{connecting them}\\ +\text{ proper node cost}\end{matrix} \right\} \tag{8}$$

where

$$\begin{matrix}\text{proper node cost}\\ \text{of node } (t, k)\end{matrix} = \begin{cases} f[S(t)]d_k \text{ if the arc} \\ \text{arriving at node } (t, k) \text{ is} \\ \text{a cross arc,} \\ \{f[S(t)] - f[S(t-1)]\}d_k \text{ if} \\ \text{the arc arriving at node} \\ (t, k) \text{ is a vertical arc.} \end{cases} \tag{9}$$

With the above labeling method, the label of each node will be the minimum cost of path(s) in graph G connecting nodes in the first row of matrix A to it. An optimal solution to the problem is obtained by finding the node(s) with minimum label in the last row of matrix A and determining its predecessor(s) from which the label was produced and continuing to the first row of matrix A. Each node thus obtained will then identify the proper terminating center assignment at the corresponding stage.

The above method is essentially a forward dynamic programming technique.[7] It can also handle the case where a terminating center has been initially assigned to the originating center. In such a case, simply adjoin a row on top of the first row of matrix A such that all its elements are zero except the one corresponding to the initial terminating center assigned to the originating center, which is 1. In that case all paths in graph G will originate from this new node (and its

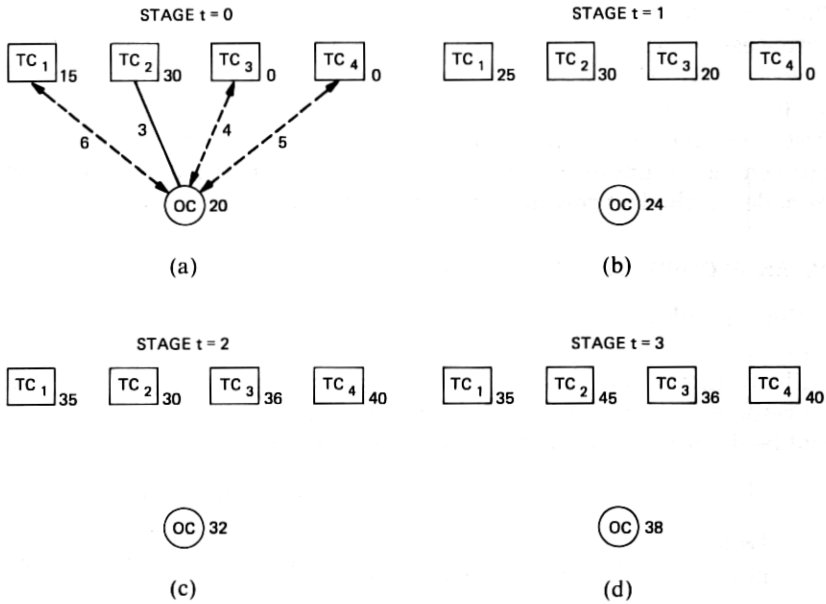

Fig. 2—Load, capacities, and distances in the example for stages $t = 0$ through $t = 3$.

label can be assumed to be zero). Note that for cost calculation in such a case, $S(0)$ will have its actual value, not zero.

V. EXAMPLE

We will illustrate the above method by applying it to a problem involving three stages, $t = 1, 2, 3$, and four terminating centers TC_k, $k = 1, 2, 3, 4$. Assume that initially (i.e., at $t = 0$) terminating center TC_2 has been assigned to the originating center. This is illustrated in Fig. 2a where the load of the originating center and the capacities of the terminating centers (in proper units) are indicated adjacent to them. Also, the distance between the originating center and each terminating center is indicated in Fig. 2a (i.e., $d_1 = 6$, $d_2 = 3$, $d_3 = 4$, $d_4 = 5$ in proper units). Figures 2b, 2c and 2d show the originating center's load and the terminating centers' capacities in stages 1, 2 and 3, respectively. Note that

$$S(0) = 20, S(1) = 24, S(2) = 32, S(3) = 38.$$

This configuration results in the matrix A shown in Fig. 3a. The corresponding graph G in which nodes are distinguished by circles and arcs are shown by directed links between the nodes is shown in Fig. 3b. Assume, for convenience, that the cost characteristic shown in Fig. 1 does not change with time and that

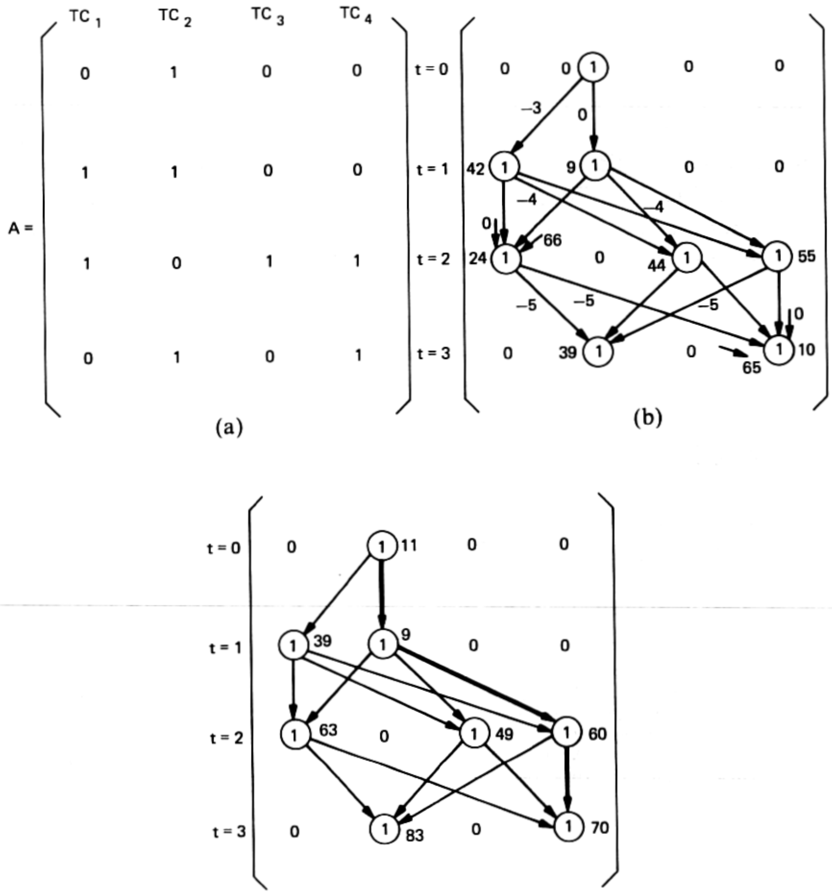

Fig. 3—Construction of the optimal solution in the example. (a) The matrix. (b) The graph with costs of arcs and nodes. (c) Node labels and the optimal path (in heavy lines).

$$f(20) = 4,\ f(24) = 7,\ f(32) = 11,\ f(38) = 13,$$

as indicated in Fig. 1. Also assume that

$$g(20) = 6,\ g(24) = 8,\ g(32) = 10.$$

From the above information the proper node costs can be calculated. For example, for node (1,1) (corresponding to stage $t = 1$ and terminating center TC_1) the proper node cost is

$$f[S(1)]d_1 = (7)(6) = 42,$$

and for node (1,2) (corresponding to stage $t = 1$ and terminating center TC_2) it is

$$\{f[S(1)] - f[S(0)]\}d_2 = (7 - 4)(3) = 9.$$

The costs of arcs and the proper node costs in graph G are written adjacent to them in Fig. 3b. Note that all the vertical arcs have zero cost and all the cross arcs from any stage to the next have the same (negative) cost. The node labels obtained by the above procedure are shown in Fig. 3c. The path distinguished by heavy lines in Fig. 3c is the minimum-cost path corresponding to the optimal solution. Thus, the optimal assignments are as follows:

stage	0	1	2	3
terminating center	TC_2	TC_2	TC_4	TC_4

REFERENCES

1. M. Malek-Zavarei, "Optimal Scheduling of Switching Centers for Telephone Systems," Proc. World Telecommunications Forum, Geneva, Switzerland, Oct. 6–8, 1975.
2. J. J. Pilliod, "Fundamental Plans for Toll Telephone Plants," B.S.T.J., *31* (September 1952), pp. 832–50.
3. R. Syski, *Introduction to Congestion Theory in Telephone Systems*, London: Oliver & Boyd, 1960.
4. M. Malek-Zavarei, "Computer-Aided Design of Telephone Systems," Computers & Electrical Engineering, *4,* No. 1 (January 1977), pp. 37–51.
5. R. S. Garfield and G. L. Nemhauser, *Integer Programming*, New York: Wiley, 1972.
6. L. R. Ford, Jr. and D. R. Fulkerson, *Flows in Networks*, Princeton, N.J.: Princeton University Press, 1962.
7. S. E. Dreyfus and A. M. Law, *The Art and Theory of Dynamic Programming*, New York: Academic Press, 1977.