A NOTE ON TREE MEDIANS

Mohan B. Sharma

IBM Corporation
11400 Burnet Road
Austin, TX 78758.

e-mail: *sharma@dss1.austin.ibm.com*

Keywords: centers, medians, communication networks.

ABSTRACT

The median invariance property of trees is described in this paper. Tree structures that are weighted (nonnegative edge weights) and unweighted (unit edge weights) are considered. It is established that the location of median(s) in trees is an invariant with respect to the weights on each tree edge and that the median of a tree T is the same if the tree T were weighted or unweighted.

I. Introduction

Network location deals with the problem of selecting one node of a network that optimizes some functions which are distance-dependent with respect to all other nodes of the network. These problems include the location of centers and medians of networks[3]. A node v of a connected graph G(V,E) is a central node if its distance to any other node in G is a minimum. The center of a graph is the set of all central nodes. The *status* of a node v in G(V,E) is the sum of distances from v to every other node in G. The median of a graph G is the set of nodes with minimum status.

With respect to facility location, the center of a network optimizes the distance to/from every other node in the network whereas the median optimizes the total distance traversed from every other node. Network location theory and problems have been discussed in great detail by Handler and Mirchandani[1]. A number of interesting results are summarized in [2,3].

In this paper, we present the median invariance property of trees. We consider median location on unweighted trees and weighted trees (trees with a nonnegative weight assigned to each tree edge) and show that the median of a tree is invariant with respect to the weights on each edge. In other words, the median of a tree T is the same if T were weighted or unweighted. The same property however, is not true for tree centers as their location in a tree is sensitive to the edge weights.

The median property is illustrated in figures 1a and 1b. The numbers marked next to each node indicate the node status. The numbers along each edge denote the edge weights in figure 1b. The median is marked by a solid circle.

The notations and definitions are described in the following section. The main result and relevent previous results are described in section III. The results and their significance are summarized in the section on conclusions.

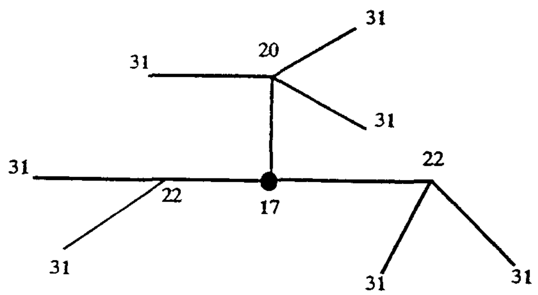

Figure 1a.

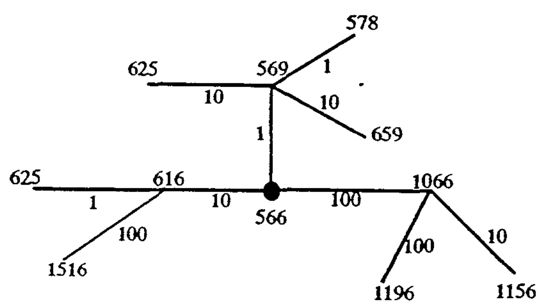

Figure 1b.

II. Notations and Definitions

Let T(V,E) denote an unweighted tree and let T_w(V,E) denote a weighted tree that is otherwise identical to T. By weighted, we imply a nonnegative number w(u,v) associated with each edge (u,v) $\in$ E(T_w). In this paper, we explicitly refer to T_w whenever necessary, otherwise we refer to the tree T in general. The distance d(u,v) denotes the length of the shortest path between nodes u and v. The status S_u of a node u in T is the sum of distances from v to every other node in T or T_w.

$$S_u = \sum_{v \in V(T)} d(u,v)$$

A node x $\in$V(T) (or $\in$ V(T_w)) is a median if S_x is the minimum over all nodes. Let (u,v) $\in$ E, then T(u,v) refers to the subtree of T rooted at v containing the node u. Let N(u) represent the set of neighbors of the node u.

III.Median Invariance

We summarize relevent results in this section and present the main result. It is known that every tree has one two medians in a tree and if there are two medians, then they must be neighbors [3].

The following theorem relates the status of neighboring nodes to the number of nodes in the subtrees that includes each of them. We present the theorem and proof for the general case of weighted trees which is an extension of a similar result described in [4,5].

Theorem 1.:

Let i,j be two distinct, adjacent nodes in T_w. Then, $S_i = S_j$ + w(i,j) [|V(T(j,i))| - |V(T(i,j))|].

Proof:

By definition of S_i, we have,

$$S_i = \sum_{k \in V(T)} d(i,k) = \sum_{k \in V(T(i,j))-\{j\}} d(i,k) + \sum_{k \in V(T(j,i))-\{i\}} d(i,k)$$

However, V(T) = V(T(i,j)) $\cup$ V(T(j,i)) - {i,j}.

$$d(i,k) = \begin{cases} d(j,k) + w(i,j) & \text{for } k \in V(T(j,i)) - \{i\} \\ d(j,k) - w(i,j) & \text{for } k \in V(T(i,j)) - \{j\} \end{cases}$$

$$\text{Hence, } S_i = \sum_{k \in V(T(i,j))-\{j\}} d(j,k) + w(i,j)\,[|V(T(j,i))| - 1] + \sum_{k \in V(T(j,i))-\{i\}} d(j,k) - w(i,j)\,[|V(T(i,j))| - 1]$$

$$= S_j + w(i,j)\,[|V(T(j,i))| - |V(T(i,j))|].$$

Hence, the lemma holds. □

Theorem 2:

Median location in a tree is independent of edge weights.

Proof:

Suppose v_x is a median in T and v_y in T_w, and that $v_x \neq v_y$. If $v_x \in N(v_y)$, then there exists a path P = $v_y, v_{i1}, v_{i2}, .. , v_{ik}, v_x$, k > 0, between v_y and v_x.

Since v_x and v_y are medians in T and T_w respectively, it is clear that

$$S_{v_x} \leq S_{v_{ik}} < .. < S_{v_y} \text{ in T and,}$$

$$S_{v_y} \leq S_{v_{i1}} < S_{v_{i1}} < .. < S_{v_x}, \text{in } T_w.$$

This implies that $S_{v_x} > S_{v_{ik}}$ in T, and $S_{v_x} \leq S_{v_{ik}}$ in T_w; a contradictiion from theorem 1 since $S_{v_x} = S_{v_{ik}} + w(v_x, v_{ik})[|V(T(v_{ik}, v_x))| - |V(T(v_x, v_{ik}))|]$ and $w(v_x, v_{ik}) > 0$.

If v_x and v_y are neighbors, then both must be the medians of T and T_w or else the above contradiction holds. □.

Conclusion:

The property of median invariance is of significance in communication networks , besides others. In a communication network, the cost of communication between a pair of nodes depends upon several factors such as the distance between them, link reliability, and bandwidth. As such, they are generally represented as edge weighted graphs. The location of centers and medians is important in communication networks for optimizing network activities. There are several sequential and distributed algorithms for determining centers and medians for both weighted and unweighted graphs representing the network. The median invariance property of trees suggests that a median finding algorithm in such a network can ignore the edge weights and simply treat the tree as unweighted.

Acknowledgement: The author would like to thank Dr.Jianhua Chen and Dr.S.S.Iyengar for their suggestions and technical inputs during the preparation of this paper.

References:

1. Handler, G.Y., and Mirchandani, P., "Location on networks: Theory and Applications", *MIT Press*, Cambridge, Mass., 1979.

2. Hansen, P., Labbe, M., Peeters, D., and Thisse, J.F., "Single facility location on networks", *Annals of discrete mathematics*, 31, 1987, pp.113-146.

3. Buckley, F., and Harary, F., "Distance in graphs", *Addison-Wesley*, 1990.

4. Sharma, M.B., Chen, J., and Iyengar, S.S., "Distributed algorithms for locating centers and medians in communication networks", *Proc. ACM Symp. on applied computing*, Vol.II., March 1-3, 1992, pp.808-817.

5. Korach, E., Rotem, D., and Santoro, N., "Distributed algorithms for finding centers and medians in networks", *ACM trans. prog. lang. and systems*, Vol.6., No. 3, July 1984, pp.380-401.