Philippine Engineering Journal (1995) XVI (2): 1-8

ROUTING BROADCAST PACKETS ALONG A MINIMUM DIAMETER TREE

Efren F. Abaya, Ph.D.
J.J. Vergara Professor of Communications Systems
Department of Electrical & Electronics Engineering
College of Engineering
University of the Philippines
Diliman, Quezon City, Philippines

ABSTRACT

In traditional computer networks, such as X.25 and TCP/IP networks, packets generally traverse the lowest cost route going from one source to one destination. In recent years, new classes of computer and video services have emerged that transfer multiple copies of a packet from one source to many destinations (multicast or broadcast).

This paper models the broadcast routing problem in a mesh computer network as a graph theory problem with a cost function that has to be minimized. The paper proposes a new criterion for routing broadcast packets when each node in the network may be a source of broadcast packets directed to the other nodes. Constraining broadcast packets to follow a single spanning tree, it is shown that a minimum diameter spanning tree is a suitable choice for routing purposes. A heuristic for generating a minimum diameter spanning tree is presented.

INTRODUCTION

In traditional computer networks, such as X.25 and TCP/IP networks, packets generally traverse the lowest cost route going from one source to one destination. This routing strategy has proven effective for services such as batch file transfer and interactive remote login. In recent years, new classes or computer and video services have emerged that transfer multiple copies of a packet from one source to many destinations. Applications such as videoconferencing, parallel search of distributed databases, multiple address email, and email distribution lists can flood a computer network with multiple copies of a single packet that will consume the capacity or the resources of the network. This situation is similar to broadcasting or flooding algorithms.

Transmission of a packet from one source to many destinations is called multicasting [COM, TANb], or multipoint routing, or multiple destination routing (MDR) [TANa]. The extreme case, where the packet is sent to all possible destinations (except the source), is called broadcasting. The objective of routing strategies is to conserve network bandwidth or lower the cost of transmitting packets.

In networks with broadcast media such as IEEE 802.3 (Ethernet) or VSAT networks, broadcasting of packets is achieved by the normal operation of the medium. However, in networks with a mesh topology, broadcasting strategies have to be deliberately defined. Otherwise, packets may loop or return to the sender. A popular choice is to route broadcast packets along a spanning tree.

This paper models the broadcast routing problem in a mesh computer network as a graph theory problem with a cost function that has to be minimized. The purpose of this paper is to propose a new criterion for routing broadcast packets if each node in the network may be a source of broadcast packets directed to the other nodes. Constraining broadcast packets to follow a single spanning tree, it will be shown that a minimum diameter spanning tree is a suitable choice for routing purposes. A heuristic for generating a minimum diameter spanning tree is presented.

SPANNING TREES

A network is modeled by a weighted non-directed graph consisting of nodes v_i and links l_j. For the j-th link, a link weight w_j represents some general cost of transporting one packet across the link, such as physical distance between nodes, average transit time, time delay, or economic cost. Alternatively, the link weight may be made inversely proportional to link transmission capacity (bandwidth) so that a higher weight implies a less desirable link. Figure 1 shows an example of a network with link weights.

A path from a source node to a destination node is a connected sequence of nodes and links in which no node appears more than once (i.e. that has no loops). The weight or cost of the path is defined as the sum of weights w_j for the links l_i in the path.

The shortest path between two nodes v_A and v_B is that path that has the smallest path weight among all paths connecting the two nodes. The distance $d(v_A, v_B)$ between the two nodes is defined as the weight of the shortest path. The shortest path, as well as the distance, are easily obtained with standard graph theory algorithms such as Dijkstra's algorithm [DEO].

The shortest paths from a given source node v_A to all other nodes form a rooted spanning tree that we shall call a shortest path tree (sometimes called a Dijkstra tree).

Alternatively, routing algorithms can also use a minimum spanning tree (MST) or shortest spanning tree, defined as the spanning tree having the smallest weight. (The weight of a spanning tree is defined as the sum of the link weights of all the links in the tree.) The minimum spanning tree is constructed using either Prim's algorithm or Kruskal's algorithm [DEO].

Given a mesh network such as Figure 1, a spanning tree can be chosen for purposes of routing broadcast packets from one node to the other nodes. Routing along a tree is simple and attractive since from any source node there is exactly one path to every destination node. The tree ensures that every destination will be reached, and that no loops will exist for packets to circulate. Various broadcast strategies are discussed in [TAN].

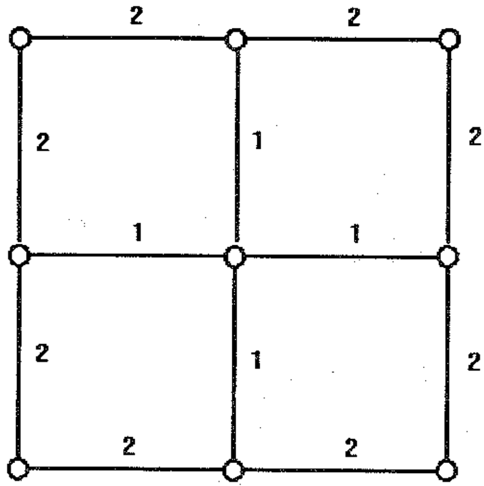


Figure 1. A weighted graph representing a computer network.

In some protocols used by LAN bridges, nodes cooperate in handling broadcast packets so that only one packet is sent along each link. When a node with two or more outgoing links receives a broadcast packet, it will forward one copy along each outgoing direction. With this strategy, the total cost of routing packets from one source to all destinations is the tree weight.

$$W = \sum_i W_i \qquad \text{(Eq. 1)}$$

where the summation is taken over all links i in the tree. Clearly, the sum is minimized by routing packets along a minimum spanning tree.

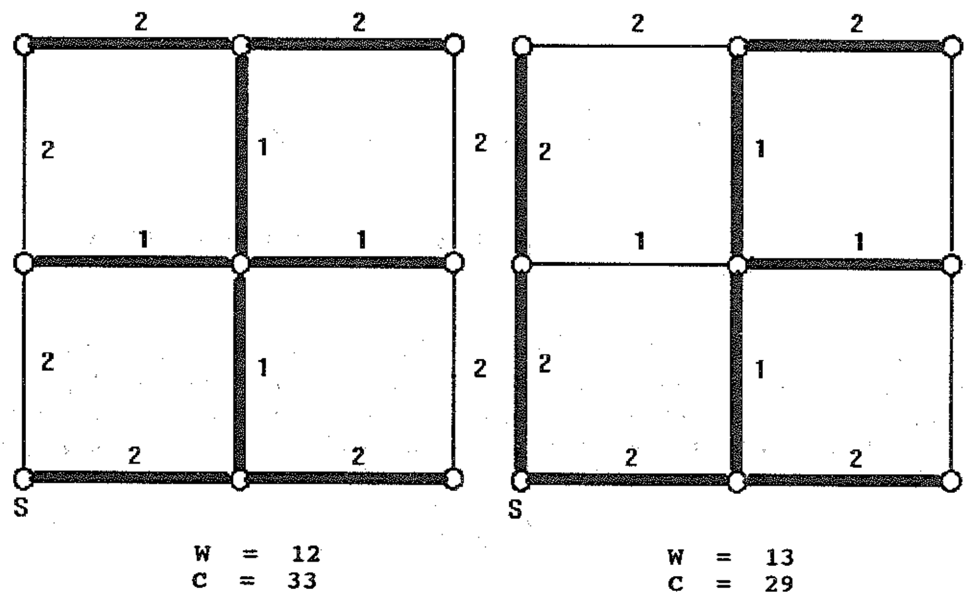
However, there are cases where a separate packet must be sent from the source to each destination. For example, blind courtesy copies (typically called bcc) of email carry a different address in the mail header for each destination, although the body of the mail will be identical for all copies. Mass mailing, or mailing lists, also typically send one separate copy of mail for each addressee.

In this case, the cost of routing packets from one source node S is given by the total weight of all paths originating from S (i.e. the total distance of all nodes from S):

$$C = \sum_i c\left(S, v_j\right) \qquad \text{(Eq. 2)}$$

where the summation is taken over all nodes j. The cost $c(S, v_j)$ is the weight of the path taken by packets going from node S to node v_j. Then, the lowest cost is obtained by routing along a shortest path tree rooted at node S so that $c(S, v_j) = d(S, v_j)$. Such a tree can be constructed using Dijkstra's algorithm to simultaneously find the shortest paths to all nodes. For example, the Open Shortest Path First (OSPF) routing algorithm is based on Dijkstra trees [PER].

We observe that in general, the shortest spanning tree and the shortest path tree are different. This is illustrated by Figure 2 for a source node S at the lower left corner of the graph. Secondly, we observe that there will be different shortest path trees from different source nodes in the graph. For example, if the tree in Figure 2(b) were used for the source at the lower right node, the total distance of all nodes would be C = 34, which is obviously not the lowest possible cost.



(a) Minimum spanning tree (b) Shortest path tree for routing packets from source S

Figure 2. Comparison of minimum spanning tree and shortest path tree rooted at node S. Cost C is computed by Eq. (2).

BROADCASTING FROM MANY SOURCES

In future applications envisioned for high speed networks, every node can potentially be a source of broadcast packets. For example, in future distributed databases, one way for a client to locate data may be to broadcast a query to all databases servers. (The alternative is to query each database sequentially.)

Assume that each destination is sent a separate copy of the broadcast packet. Suppose that we would like the broadcast packets to traverse a single tree in order to simplify the routing. Given a network graph with a mesh topology, the issue is how to pick the best spanning tree for this purpose.

In a network in which each of N nodes can be a broadcast source, the average cost per source of routing packets along a single tree is

$$\sum_{i}\sum_{j} d(v_i, v_j)/N$$

Here the distance $d(v_i,v_j)$ between source v_i and the destination v_j is taken as the weight of the unique path along the tree.

Therefore, the best spanning tree for routing broadcast packets is that which minimizes the cost criterion

$$C=\sum_{i}^{N}\sum_{j=1}^{i-1} d(V_i, V_j) \qquad \text{(Eq. 3)}$$

This cost represents the sum of distances taken between all pairs of nodes in the spanning tree.

In this cost criterion, the i-th link contributes the link weight w_i multiplied by the number of paths that traverse the link, when paths between all pairs of nodes are counted. Consider a representative tree shown in Figure 3. For link i of this tree, the number of paths is equal to the product of the number of nodes on the left side of the link (3) and the number of nodes on the right side (6). Thus, a total of 18 broadcast paths can pass through link i.

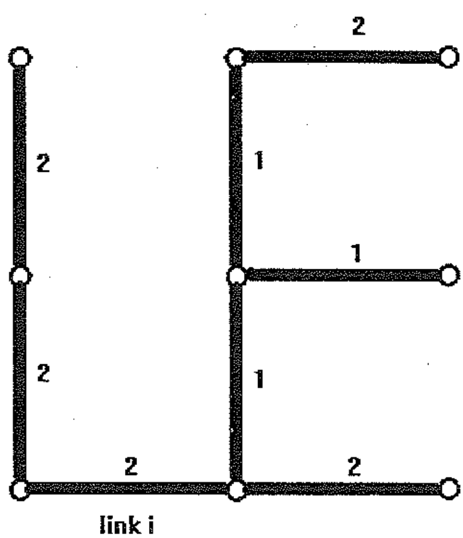


Figure 3. A typical spanning tree.

For link i of a tree, let N_{iL} be the number of nodes to the left of the link, and N_{iR} be the number of nodes to the right of the link. Then the cost (Equation 3) can be expressed alternatively as

$$C = \sum_i w_i N_{iL} N_{iR} \qquad \text{(Eq. 4)}$$

Observe that the more dissimilar the terms are, the smaller is the product N_{iL} N_{iR} (where the sum $N_{iL} + N_{iR} = N$, a constant). As the terms become closer, the larger the product becomes. Thus, the product N_{iL} N_{iR} is minimized by making one term equal to unity and the other equal to N-1, if possible.

Based on this consideration, the ideal spanning tree is a star (see Figure 4). Notice that the distance between nodes in a star network are as short as possible, so that the cost (3) is also made small. However, it is not always possible to find a spanning tree with a star topology. Then a two-level (or multi-level) star such as that shown in Figure 5 may be the best.

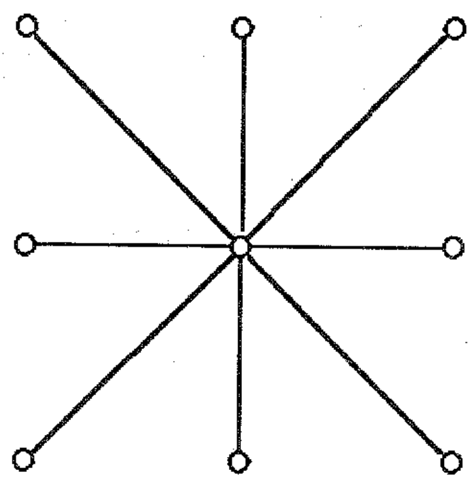


Figure 4. A spanning tree with star topology

In networks with sparse connections, we cannot always hope to find a spanning tree with a multi-level star topology. So we come back to the question of constructing a spanning tree to minimize the cost criterion Equation 3 or Equation 4.

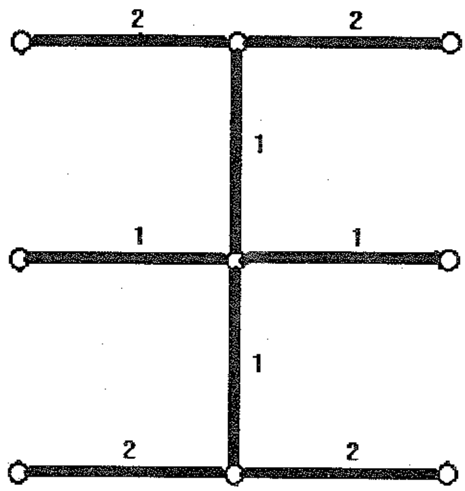


Figure 5. A spanning tree with multi-level star topology.

MINIMUM DIAMETER SPANNING TREE

Given a network graph with a mesh topology, we would like to construct a spanning tree that minimizes the cost criterion Equation 3 or Equation 4. Note that this cost is the sum of distances taken between all pairs of nodes in the spanning tree. Thus it make sense to select a tree in which the longest path is as small as possible.

We define the diameter of a graph as the largest distance between two nodes [DEO]. It follows that the diameter of a spanning tree is the longest path along the tree.

We conjecture that a minimum diameter spanning tree minimizes, or approaches the minimum of, the cost (3). Therefore, we seek to construct a tree that achieves

$$\min \max_{i,j} d\left(v_i, v_j\right) \qquad \text{(Eq. 5)}$$

where the minimum is taken over all spanning trees of the graph. Observe that in a graph with mesh topology, the minimum tree diameter D_t is lower bounded by the diameter of the graph D_g. For suppose this were not true and the diameter of the spanning tree D_t is less than the diameter of the graph Dg. Then the largest distance between two nodes is D_t and this contradicts the assumption that the largest distance between two nodes in the graph is D_g. On the other hand, it is easy to produce examples for which the minimum tree diameter is larger than the graph diameter. A common example is a graph with ring topology ("ring" or "loop" network).

To illustrate the above ideas, the graph in Figure 1 has a diameter of 6. The minimum diameter tree is shown in Figure 2a, with a cost of 116 computed from either Equation (3) or Equation (4). The shortest path tree in Figure 2b is inferior because its diameter is 10 and the cost by Equation 3 is 154.

Since there are not many algorithms for finding the minimum diameter tree in the literature we use the following heuristic:

1. Find the shortest path between the two nodes that are furthest apart in the graph, and call this the starting subtree. That is, find the path that determines the diameter of the graph. (Intuitively, this seems to be a logical choice. Suppose the two nodes that are farthest apart are denoted by v_1 and v_N; they are connected by a path having weight D_g. If this path is not included in the minimum diameter tree, then for sure D_t is strictly greater than D_g.)

2. For each node not in the starting subtree, connect the node to the starting subtree in such a way that the increase in tree diameter is minimized. If there are several connecting paths that satisfy this condition, use the one with the shortest path to the starting subtree.

This algorithm produced the spanning tree in Figure 2a.

CONCLUSION

This paper has presented a new criterion for routing broadcast packets in a computer network if each node in the network may be a source of broadcast packets directed to the other nodes. Constraining broadcast packets to follow a single spanning tree, it is shown that a minimum diameter spanning tree is a suitable choice for routing purposes. A heuristic for generating a minimum diameter spanning tree is presented.

The discussion should motivate a search for good algorithms to generate minimum diameter spanning trees. The heuristic algorithm presented can also be further investigated to determine how close its solutions are to the desired minimum diameter tree.

REFERENCES

[COM] Comer, Douglas E. (1991). Internetworking with TCP/IP vol. 1, 2nd ed. Prentice-Hall, N.J.

[DEO] Deo, Narsingh (1974). Graph Theory with Applications to Engineering and Computer Science. Prentice Hall, N.J..

[PER} Perlman, Radia (1992). Interconnections: Bridges and Routers. Addison-Wesley Publishing Co., Mass.

[TANa] Tanaka, Yoshiaki and Paul C. Huang (1993). Multiple Destination Routing Algorithms. IEICE Trans. Commun., vol. E76-B, no.5 (May 1993), pp. 544-552.

[TANb] Tanenbaum, Andrew S. (1989). Computer Networks, 2nd ed. Prentice Hall, Inc., N.J.