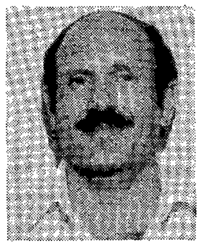
[21] D. Kazakos, "On the design of robust quantizers," presented at the 1981 Nat. Telecommun. Conf., New Orleans, LA, Dec. 1981.
[22] P. J. Huber, "A robust version of the probability ratio test," *Ann. Math. Statist.*, vol. 36, pp. 1753–1758, 1965.
[23] R. D. Martin and S. C. Schwartz, "Robust detection of a known signal in nearly Gaussian noise," *IEEE Trans. Inform. Theory*, vol. IT-17, pp. 50–55, Jan. 1971.
[24] Ky Fan, "Minimax theorems," *Proc. Nat. Acad. Sci. U.S.*, vol. 39, pp. 42–47, 1953.
[25] Y. Linde, A. Buzo, and R. M. Gray, "An algorithm for vector quantizer design," *IEEE Trans. Commun.*, vol. COM-28, pp. 84–95, Jan. 1980.
[26] J. B. Thomas, *An Introduction to Statistical Communication Theory*. New York: Wiley, 1969, Appendix E.
[27] D. G. Luenberger, *Optimization by Vector Space Methods*. New York: Wiley, 1968.
[28] E. J. McShane and T. A. Botts, *Real Analysis*. Princeton, NJ: Van Nostrand, 1959.



Dimitri Kazakos (S'69–M'73) was born in Athens, Greece, in 1944. He received the diploma in electrical, mechanical, and industrial engineering from the Hellenic National Polytechnic University, Greece, in 1967, the M.S. degree from Princeton University, Princeton, NJ, in 1970, and the Ph.D. degree from the University of Southern California, Los Angeles, in 1973, both in electrical engineering, specializing in statistical communication.

From 1973 to 1975 he held postdoctoral positions at Rice University, Houston, TX, and NASA Johnson Space Center, Houston. From September 1975 to August 1980 he was on the faculty of the Department of Electrical Engineering, State University of New York at Buffalo. In September 1980 he became an Associate Professor in the Department of Electrical Engineering, University of Virginia, Charlottesville.

On the Capacity of Multihop Slotted ALOHA Networks with Regular Structure

JOHN A. SILVESTER, MEMBER, IEEE AND LEONARD KLEINROCK, FELLOW, IEEE

Abstract—**In this paper we investigate the capacity of networks with a regular structure operating under the slotted ALOHA access protocol. We first consider circular (loop) and linear (bus) networks and then proceed to two-dimensional networks. For one-dimensional networks we find that the capacity is basically independent of the network average degree and is almost constant with respect to network size. For two-dimensional networks we find that the capacity grows in proportion to the square root of the number of nodes in the network provided that the average degree is kept small. Furthermore, we find that reducing the average degree (with certain connectivity restrictions) allows a higher throughput to be achieved. We also investigate some of the peculiarities of routing in these networks.**

I. INTRODUCTION

IN some applications it is not possible (or economically feasible) to construct conventional communication networks using wire-based transmission facilities. As a result of this many new technologies have been explored. The ALOHA system at the University of Hawaii [1], [9] is such an innovative network using a broadcast *radio* channel to interconnect the various campuses of the University which are located on different islands. Recent ongoing studies (of which this paper is part), supported by the Advanced Research Projects Agency of the Department of Defense (DARPA), have extended these ideas and developed the concept of a packet radio network (called the PRnet), wherein packets may take multiple hop paths in order to reach their destination [5], [6], [12]. Any node in the network can act as a source or destination for traffic in this network.

The simplest random access scheme is "pure ALOHA," wherein any node having a packet ready for transmission does so. Of course, collisions occur. These are resolved by the node retransmitting the packet at some (randomly chosen) later time, if no positive acknowledgment is received. It is necessary to randomize the retransmission delays in order to avoid perpetual repetition of the collision. Extensive analysis of the performance of this protocol for *centralized, fully connected* networks can be found in [1], [10], where it is determined that the maximum that the channel can be utilized is 18 percent ($1/2e$) of the channel bandwidth. A simple modification to the ALOHA scheme–slotted ALOHA, proposed in [11]–forces transmission to commence at the beginning of "slots" (time divisions of length equal to a packet transmission time). In [10], [13] we find analysis of this scheme showing that the capacity is doubled (over unslotted ALOHA), to 36 percent

Paper approved by the Editor for Computer Communication of the IEEE Communications Society for publication after presentation in part at the 12th Hawaii International Conference on System Sciences, January 1979. Manuscript received November 10, 1981; revised September 15, 1982. This work was supported by the Defense Advanced Research Projects Agency under Contract MDA 903-77-C-0272 and by the Air Force Office of Scientific Research under Contract AFOSR 80-0171.

J. A. Silvester is with the Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089.
L. Kleinrock is with the Department of Computer Science, University of California, Los Angeles, CA 90024.

0090-6778/83/0800-0974$01.00 © 1983 IEEE

$(1/e)$. These studies were concerned with centralized networks since local access networks, for which these protocols were initially designed, usually have centralized traffic requirements (the central node often being a gateway to the main network).

If the size of the network increases it may no longer be possible to have all nodes within range of one another. Repeaters must be used and we are faced with a new set of (more difficult) problems. In [4], [15] we find some capacity results for two-hop slotted ALOHA centralized nets.

In more recent developments such as the PRnet mentioned above, we no longer have a central node and we are thus faced with the problem of multihop networks similar to the traditional line networks such as the ARPAnet, except that a broadcast medium is used for transmissions. This greatly complicates the work of the performance analyst and relatively little work has been done in this area. In [8], [14] we considered some simple models of multihop networks with random structure and found that using a small average degree (number of nodes within range of a transmitter) is beneficial. We also noted that the effects of routing and flow control were important in these networks. In this paper we consider the simpler case of regular structures in more detail and are able to gain further insight into the routing problem.

In this paper we will be looking at multihop networks with a regular structure (i.e., in which the nodes are regularly placed on a square grid, for example), where messages are forwarded from node to node following a path defined by the routing matrix. These networks are much easier to analyze than the random networks discussed in [8] since the topology is fixed and the progress that can be made toward the destination in any hop is not dependent on any probabilistic argument.

We start by looking at one-dimensional networks. Loop networks, as discussed in Section II, are really one-dimensional line networks wrapped around a circle with the ends joined together. In Section III we look at networks generated on the line (i.e., no joining of the ends). We follow this with a discussion of two-dimensional regular networks, such as the square lattice.

The traffic matrix that we use is uniform, i.e., each node splits its traffic equally between all possible destinations. With this traffic matrix and the uniformity of the topology, we initially assume that the traffic load on all links of the network is homogeneous. This seems to be a valid assumption for any reasonable routing algorithm in loop networks, since there are no edge effects to consider. In two-dimensional networks we neglect edge effects as they are of minor importance relative to the rest of the network (the perimeter of the network will contain $O(\sqrt{n})$ nodes). We then study the effect of routing in the grid network in more detail.

II. LOOP NETWORKS

The networks considered in this section consist of n points uniformly (regularly) distributed around the circumference of a circle. Fig. 1 shows a typical loop network, consisting of eight nodes and average degree of 5 (each node can communicate with its two neighbors on either side plus itself). Each node in the network is identical in terms of traffic handled, degree, and so on. We can therefore compute the expected

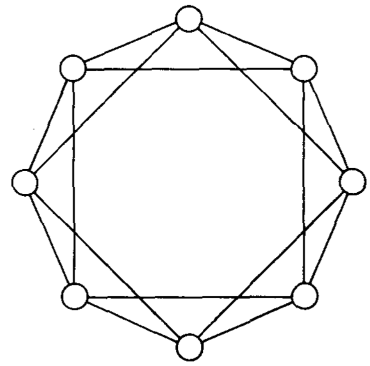


Fig. 1. Regular loop network.

number of successful transmissions per slot for the whole network (s_{net}) to be n times the probability of success for any particular node s_i. In a heavy traffic situation, i.e., in which all nodes always have a packet to be transmitted, the probability of successful transmission is identical to the local throughput that can be achieved by that node. For a homogeneous network and traffic matrix, we can divide the local throughput (summed over all nodes) by the average path length to obtain the end-to-end throughput. We consider this heavy traffic throughput to be the network capacity.

A. Network Success Rate

We start, therefore, by determining the expected number of successful transmissions per slot for an arbitrary node (i) in the network. Let N denote the degree of any node in the network (i.e., the number of neighbors plus the node itself). We will use the term "one-hop throughput" to mean the rate at which any individual node can successfully transmit packets to the next node along the path to the destination. If we let s_i denote the probability of a successful transmission in any slot by node i (one-hop throughput) then

$$\begin{aligned} s_i &= \Pr\{\text{node } i \text{ successfully transmits}\} \\ &= \Pr\{\text{node } i \text{ transmits and no node in range of the destination does}\} \\ &= p(1-p)^{N-1} \end{aligned} \tag{1}$$

where p is the probability that a node transmits in any slot. Since all nodes carry the same traffic and have the same degree, we give every node the same transmission probability. In heavy traffic (nodes are always busy), we can describe the behavior of a node in terms of this transmission probability, which corresponds to the offered traffic randomized to avoid repeated collisions. In order to find the optimum value for this transmission probability we differentiate

$$\begin{aligned} \frac{ds_i}{dp} &= (1-p)^{N-1} - (N-1)p(1-p)^{N-2} \\ &= (1-Np)(1-p)^{N-2}. \end{aligned} \tag{2}$$

Equating this to zero for optimality and excluding the case of $p = 1$ which corresponds to a minimum with zero throughput, we find that

$$p^* = \frac{1}{N}. \tag{3}$$

This is the same as found for fully connected networks in [1], as expected. We drop the asterisk and use p to represent the optimum value for the rest of this analysis. Rewriting the expression for success probability, we have

$$s_i = \frac{1}{N}\left(1 - \frac{1}{N}\right)^{N-1}. \tag{4}$$

From this we can obtain the expected number of successful transmissions per slot for the whole network s_{net}:

$$s_{\text{net}} = \frac{n}{N}\left(1 - \frac{1}{N}\right)^{N-1}. \tag{5}$$

This represents the expected number of successful packets received per slot for the whole network. It does not correspond to the throughput since a multihop path will require *many* transmissions and we are counting each hop in this expression as a contribution to the throughput. When the network is fully connected ($N = n$), we see that this reduces to the usual throughput value of $1/e$ (path lengths all being one in this case).

B. Path Length

In order to compute the network throughput γ, we must divide s_{net} by the expected path length in hops $\bar{l}$. In order to find the average path length, we split the network into groups such that all the members of one group are equidistant (in hops) from a given (typical) node. Thus, the first group will be those with whom a node can directly communicate, the second group will be those that are two hops away, and so on. Each group will have the same number of members, $N-1$, except possibly for the last group which will have the remainder if $(n-1)/(N-1)$ is not an integer. Let g represent the number of complete (not counting this remainder) groups; then

$$g = \left\lfloor \frac{n-1}{N-1} \right\rfloor \tag{6}$$

where $\lfloor x \rfloor$ is the largest integer less than or equal to x. There are $(n-1) - (N-1)g$ nodes in the last partial group, with path length $g+1$. This group is taken care of by the second term in the following expression for $\bar{l}$.

$$\begin{aligned}\bar{l} &= \frac{1}{n-1}\left((N-1)\sum_{i=1}^{g} i + (g+1)[(n-1) - g(N-1)]\right)\\ &= \frac{1}{n-1}\left((g+1)(n-1) - \frac{(N-1)g(g+1)}{2}\right)\\ &= (g+1) - \frac{(N-1)g(g+1)}{2(n-1)}.\end{aligned} \tag{7}$$

Recalling the example of Fig. 1, we find the number of complete groups is

$$g = \lfloor \tfrac{7}{4} \rfloor = 1 \tag{8}$$

and the average path length is

$$\bar{l} = \tfrac{10}{7}. \tag{9}$$

If there are no nodes in the special extra group, which is to say that there is no remainder in the division of $n-1$ by $N-1$, this reduces to the following "clean" expression for $\bar{l}$.

$$\bar{l} = \frac{n+N-2}{2N-2}. \tag{10}$$

This will also be a good approximation for $\bar{l}$ when the total number of nodes is large compared to the number in the last group, more precisely, when

$$\frac{(N-1)g}{2} \gg (n-1) - g(N-1). \tag{11}$$

C. Throughput

For networks where (10) applies, we compute the network throughput γ (as the network success rate divided by the average path length) to be

$$\begin{aligned}\gamma &= \frac{s_{\text{net}}}{\bar{l}}\\ &= \frac{n}{N}\left(1 - \frac{1}{N}\right)^{N-1} \frac{2N-2}{n+N-2}\\ &= \frac{2n}{n+N-2}\left(\frac{N-1}{N}\right)^{N}.\end{aligned} \tag{12}$$

Let us evaluate this expression for two interesting cases, 1) when $N = n$ (i.e., a fully connected net), and 2) when $N = 3$ (i.e., each node is connected only to his immediate neighbors). We notice that the average path length for the fully connected case is 1, and $(n+1)/4$ when each node is connected only to his neighbors. We denote the throughput for a network with $N = j$ by γ_j.

D. Fully Connected Network

For the fully connected network we have

$$\gamma_n = \left(1 - \frac{1}{n}\right)^{n-1}. \tag{13}$$

Taking the limit for large n, we find

$$\lim_{n\to\infty}\left(1 - \frac{1}{n}\right)^{n-1} = \frac{1}{e}. \tag{14}$$

We are not surprised to see that for large n we achieve a throughput which corresponds to the usual infinite ALOHA population.

E. Neighbor Communication

Considering the other extreme where each node is only connected to its two neighbors, we find (for the case where (10) is valid)

$$\gamma_3 = \frac{2n}{n+1}\left(\frac{2}{3}\right)^3 = \frac{16n}{27(n+1)}. \tag{15}$$

For a large net the throughput for the neighbor case is therefore 16/27, which is greater than $1/e$. It may be that the maximum for γ is achieved for an intermediate value. We therefore proceed to investigate the behavior of (12) for intermediate values of N in order to determine that value of n which maximizes the throughput.

F. Optimal Average Degree

Recall the throughput expression for γ_N (neglecting the last group):

$$\gamma_N = \frac{2n}{n+N-2}\left(\frac{N-1}{N}\right)^N. \tag{16}$$

Differentiating (16) with respect to N, we have

$$\frac{d\gamma}{dN} = \frac{2n}{(n+N-2)^2}\left(\frac{N-1}{N}\right)^N \cdot \left[-1 + (n+N-2)\left(\frac{1}{N-1} + \log\left(\frac{N-1}{N}\right)\right)\right]. \tag{17}$$

To find the optimal N, we must solve the equation

$$(n+N-2)\left(\frac{1}{N-1} + \log\left(\frac{N-1}{N}\right)\right) - 1 = 0. \tag{18}$$

Rewriting, we have

$$N-1 = n+N-2 + (N-1)(n+N-2)\log\left(1-\frac{1}{N}\right). \tag{19}$$

Since $N > 1$, we can expand the log to obtain

$$0 = n-1-(N-1)(n+N-2)\left(\frac{1}{N} + \frac{1}{2}\frac{1}{N^2} + o\left(\frac{1}{N^3}\right)\right). \tag{20}$$

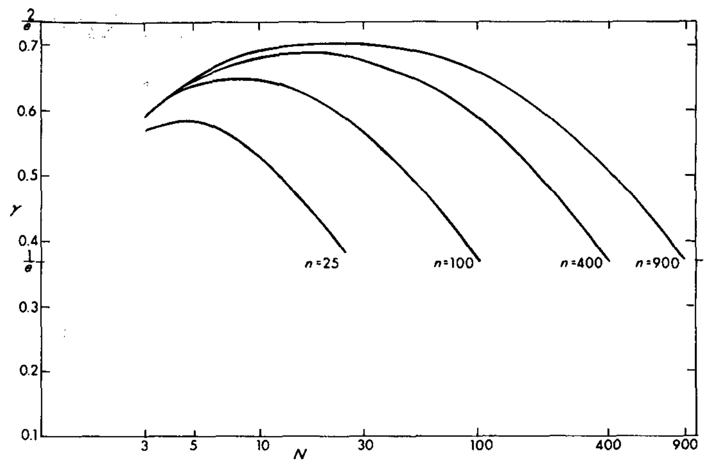


Fig. 2. Throughput versus average degree.

After some algebra we find

$$\frac{n}{2N^2} + o\left(\frac{1}{N}\right) + o\left(\frac{n}{N^3}\right) = 1. \tag{21}$$

In order for this equation to balance for large n, assuming negligible $o(1/N)$, etc., we have

$$N = \sqrt{\frac{n}{2}}. \tag{22}$$

The second factor of the equation for γ (i.e., $((N-1)/N)^N$) will vary from 8/27 (for $N = 3$) to $1/e$ (for large N). For large n, the first term is equal to 2, provided that N grows at a slower rate than n. We found above that at the optimum, N must be on the order of $\sqrt{n/2}$. For such a value of N, the throughput will be given by $2/e$. In fact, the exact value of N is not critical as long as it is greater than 3 and grows slower than n. We see, therefore, that the maximum throughput is $2/e$ and will be achieved for any moderate value of N. In Fig. 2 we plot the throughput as a function of the average degree, as given by (16), and see that the predicted behavior is indeed achieved.

III. LINE NETWORKS

Another one-dimensional network of interest is the line network, where nodes are regularly placed along a line (this might correspond to a network in linear countries such as Chile). In order to avoid the edge effects that occur at the ends of a line network, we either consider an infinitely long network (with locality in the traffic requirements so that the central nodes do not become overloaded) or the central section in a finite network (which will be the most heavily loaded and, thus, determines the maximum network throughput). For these two cases our homogeneity assumption that each node uses the same transmission probability is reasonable and the successful transmission rate for the network, as before, is given by

$$s_{\text{net}} = \frac{n}{N}\left(1-\frac{1}{N}\right)^{N-1}. \tag{23}$$

In order to implement locality for the infinite network, we can fix the average path length (in hops), or the distance traveled (in nodes passed over). If we fix the distance traveled to be k nodes, say, then the average path length in hops is given by

$$\bar{l} = \left\lceil \frac{k}{N/2} \right\rceil. \tag{24}$$

The network throughput is therefore

$$\gamma = \frac{n}{2k}\left(1 - \frac{1}{N}\right)^{N-1} \tag{25}$$

provided that $2k/N$ is a valid approximation for the ceiling function of (24) (i.e., $N \ll k$ or k is exactly divisible by N). We note that the network throughput is relatively insensitive to the average degree, since $(1 - 1/N)^{N-1}$ only varies from 4/9 to $1/e$.

If we consider a traffic matrix which causes the average number of hops to be a function of n (a uniform traffic matrix, for example), then the throughput is given by

$$\gamma = c\left(1 - \frac{1}{N}\right)^{N-1} \tag{26}$$

where c is some constant, depending on the specific traffic requirement. The important result here is that the throughput is independent of the average degree, provided that the average degree is smaller than the number of hops that messages take.

This result agrees with that found in [2], in which Akavia finds that in order to minimize delay, the node should use as large a range as necessary to reach the destination independent of the traffic load.

IV. TWO DIMENSIONS

In this section we proceed to consider two-dimensional networks. We first look at networks on the surface of a sphere (to avoid the edge effect problem). These correspond to loop networks in one dimension. We then consider a Manhattan (square grid) network, which is the two-dimensional equivalent of the line network considered earlier.

A. Path Length

The one-hop throughput for a two-dimensional regular network superimposed on a sphere to avoid edge effects (i.e., this corresponds to the loop network in one dimension) with fixed degree N will be the same as for the one-dimensional case.

$$s_{\text{net}} = \frac{n}{N}\left(1 - \frac{1}{N}\right)^{N-1}. \tag{27}$$

The path length for a uniform traffic matrix will now be proportional to the square root of the number of nodes, rather than proportional to the number of nodes, as we found for one-dimensional networks. (In [8], we give a detailed derivation of this relationship for random networks; a similar result applies for regular networks with a different constant of proportionality. The specific constant will depend on the exact topology; see [2] for several topologies or the next section for a square grid network. The reason that the path length is proportional to the square root of the number of nodes and the degree is that the distance is a linear measure, whereas number of nodes is a function of the area.)

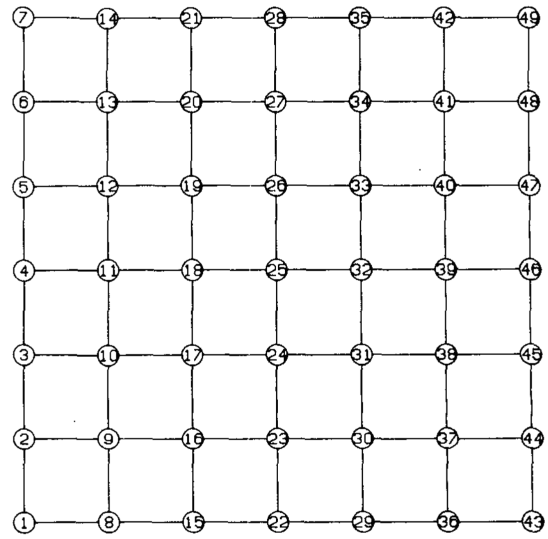


Fig. 3. A square grid network.

$$\bar{l} = c\frac{\sqrt{n}}{\sqrt{N}} \tag{28}$$

where c is some constant depending on the network topology. The throughput will be

$$\gamma = d\sqrt{\frac{n}{N}}\left(1 - \frac{1}{N}\right)^{N-1} \tag{29}$$

where d is another constant depending on the topology. The implication of this is that we should let N become as small as possible, since both $\sqrt{n/N}$ and $(1 - (1/N))^{N-1}$ increase as N decreases. The minimum value that N can take is 4 for a hexagonal tesselation (a three-connected net). For small degrees ($N = 4$ or 5), we must evaluate the proportionality constant (d). In [2], Akavia makes this comparison and finds that the optimum network is the hexagonal tesselation mentioned above.

B. Manhattan Nets

We now consider networks embedded in the plane. In particular, we consider the Manhattan (square grid) network in more detail. Fig. 3 shows a sample Manhattan network for $N = 5$ and $n = 49$. The distance metric for this network is the sum of the differences in x and y coordinates (i.e., we can only move parallel to the x or y axes).

Consider two points (x_1, y_1) and (x_2, y_2) randomly located in the unit square. The Manhattan distance between these points is $|x_1 - x_2| + |y_1 - y_2|$. Since $|x_1 - x_2|$ and

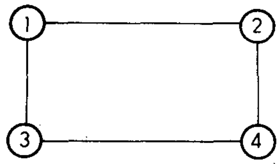


Fig. 4. A 2 × 2 grid network.

$|y_1 - y_2|$ are identically distributed independent random variables, the expected value of the distance is $d = 2E[x_1 - x_2]$, where x_1 and x_2 are uniformly distributed over [0, 1]. Thus

$$d = 2 \int_0^1 \left\{ \int_0^{x_1} (x_1 - x_2)\, dx_2 + \int_{x_1}^1 (x_2 - x_1)\, dx_2 \right\} dx_1 = \tfrac{2}{3}. \tag{30}$$

In the Appendix we derive a similar result for a discrete network and show that the average path length in an $m \times m$ square network is

$$\bar{l} = \tfrac{2}{3} m. \tag{31}$$

Assuming that edge effects can be neglected and that traffic flows are homogeneous, i.e., that each node carries the same traffic load (whether this is achievable is discussed below), we have as before (replacing n by m^2)

$$\gamma = \frac{m^2}{5} (1 - \tfrac{1}{5})^4 \frac{3}{2m} = 0.123m. \tag{32}$$

Our first assumption, that edge effects can be neglected, is similar to that for line networks, i.e., we either consider an infinite network with some locality of traffic or the center of a large network. The second assumption is harder to justify, since it is not at all clear how to route traffic in order to achieve a balanced load. Let us consider a simple routing procedure. For a uniform traffic matrix we need to specify routes for all pairs of points. It is straightforward to determine the routes for a 2 × 2 network, For a traffic pattern requiring one unit of traffic between each pair of nodes, the resulting flows for the network of Fig. 4 would be

$$F = \begin{pmatrix} 0 & 2 & 2 & 0 \\ 2 & 0 & 0 & 2 \\ 2 & 0 & 0 & 2 \\ 0 & 2 & 2 & 0 \end{pmatrix} \tag{33}$$

by using the following routing ($r_{ij} = k$ means that traffic destined to j arriving at i is forwarded to k):

$$R = \begin{pmatrix} - & 2 & 3 & 2 \\ 1 & - & 1 & 4 \\ 1 & 4 & - & 4 \\ 3 & 2 & 3 & - \end{pmatrix}. \tag{34}$$

We can now decompose a larger $m \times m$ network into 2 × 2 squares to find the routes. We define a subsquare to be a set of four nodes on the corners of a rectangle. Thus, in Fig. 5, (1, 3, 7, 9) constitute a subsquare, whereas (1, 2, 8, 3) do not.

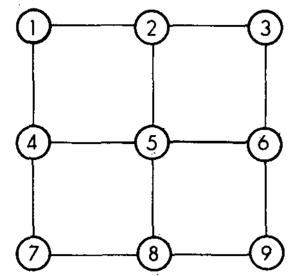


Fig. 5. A 3 × 3 grid network.

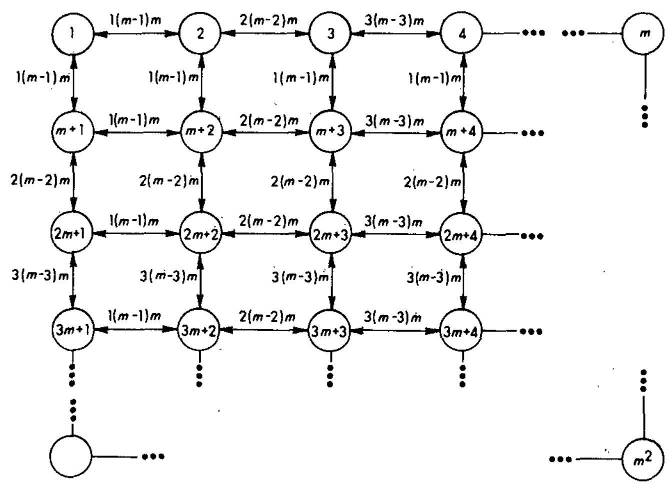


Fig. 6. Flows for an $m \times m$ network.

We use the routing found for the 2 × 2 square for each of the subsquares, which are (1, 2, 4, 5); (2, 3, 5, 6); (4, 5, 7, 8); (5, 6, 8, 9); (1, 2, 7, 8); (2, 3, 8, 9); (1, 4, 3, 6); (4, 6, 7, 9); and (1, 3, 7, 9).

Routing for the first subsquare, (1, 2, 4, 5), produces two units of flow *in each direction* on the links (1, 2), (1, 4), (2, 5), and (4, 5). Similarly, the second subsquare, (2, 3, 5, 6), would produce two units of flow on the links (2, 3), (2, 5), (5, 6), and (3, 6). Since we have already routed the traffic between 2 and 5, however, only one unit of flow is generated on the link (2, 5). The resulting flow is then two units on (1, 2), (1, 4), (4, 5), (2, 3), (3, 6), (5, 6) and three units on (2, 5). (*Note:* we have described this process for the link (1, 2), etc., but the same comments apply to the link (2, 1).) This procedure is repeated for the rest of the subsquares listed above. The resulting flow matrix is

$$F = \begin{pmatrix} 0 & 6 & 0 & 6 & 0 & 0 & 0 & 0 & 0 \\ 6 & 0 & 6 & 0 & 6 & 0 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 & 0 & 6 & 0 & 0 & 0 \\ 6 & 0 & 0 & 0 & 6 & 0 & 6 & 0 & 0 \\ 0 & 6 & 0 & 6 & 0 & 6 & 0 & 6 & 0 \\ 0 & 0 & 6 & 0 & 6 & 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 6 & 0 & 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 0 & 6 & 0 & 6 & 0 & 6 \\ 0 & 0 & 0 & 0 & 0 & 6 & 0 & 6 & 0 \end{pmatrix}$$

Using this procedure we find that the flows for an $m \times m$ network are as shown in Fig. 6, where the flows shown are in each direction on the associated link. Note that for our examples (four and nine nodes), the flows on all links were equal, whereas in general this is not true. To check that this procedure generates shortest path routes, we sum the total flow in this network. We call this total flow λ. Note that the number of links carrying a flow $i(m-i)m$ is $4m$, each direction on the m links in the ith column and ith row.

$$\lambda = 4m\left\{\sum_{i=1}^{m-1} i(m-i)m\right\}$$

$$= 4m^2\left\{m\sum_{i=1}^{m-1} i - \sum_{i=1}^{m-1} i^2\right\}$$

$$= 4m^2\left\{\frac{m(m-1)m}{2} - \frac{(m-1)m(2m-1)}{6}\right\}$$

$$= \tfrac{2}{3}m^3(m^2-1). \tag{35}$$

The total traffic carried, γ, is

$$\gamma = m^2(m^2-1). \tag{36}$$

Using a well-known result for the average path length [7], we find

$$\bar{l} = \frac{\lambda}{\gamma} = \tfrac{2}{3}m. \tag{37}$$

This checks with our previous result and, thus, the paths generated by our procedure are indeed shortest paths.

It should be clear that the flows so produced are as well balanced as possible, i.e., any other flow pattern will have a higher maximum link flow. The most heavily congested node is the one in the center of the network, which carries a flow of $m^3/4$ on each of its outgoing links. For each of these links we have that the success rate must exceed the required flow rate. For a total network traffic of γ, the flows on these links are

$$f = \frac{m^3}{4}\,\frac{\gamma}{m^2(m^2-1)}$$

$$= \frac{m\gamma}{4(m^2-1)}. \tag{38}$$

Since the nodes surrounding the center node carry approximately the same load, it is reasonable to assume that they all use the same transmission probability p and that $p = 1/5$ will give the best performance. The probability of success for this node is

$$s = p(1-p)^4. \tag{39}$$

Setting $p = 1/5$, we have that the rate at which this node succeeds to a particular neighbor, $s/4$, is

$$\frac{s}{4} = \tfrac{1}{4}\,\tfrac{1}{5}\,(1-\tfrac{1}{5})^4$$

$$= \frac{0.08192}{4}. \tag{40}$$

Therefore, since the success rate must exceed the flow requirement, we have

$$\frac{m\gamma}{4(m^2-1)} \leqslant \frac{0.08192}{4}. \tag{41}$$

Calling the maximum throughput or network capacity γ^*, and noting that the number of nodes in the network, n, is equal to m^2, we have

$$\gamma^* = 0.08192\,\frac{n-1}{\sqrt{n}}$$

$$\approx 0.08192\sqrt{n} \qquad \text{(for large } n\text{)}. \tag{42}$$

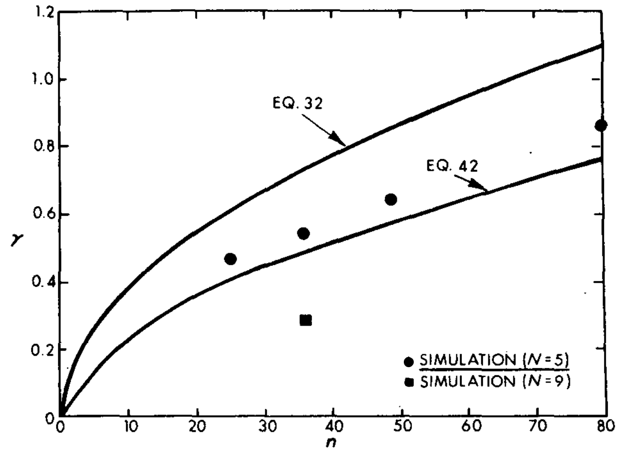


Fig. 7. Throughput for two-dimensional regular networks.

In Fig. 7, we plot this expression (42) and the one given by (32), which assumes that the flow on all links in the network is equal. We also plot some simulation data in which the path between a pair of nodes was selected by randomly choosing among the set of possible shortest paths. This approach results in less balanced flows. Since the flows so produced are unbalanced, the transmission probabilities are no longer simply 1/5, but determined by the ratio of the total flow of this node over the flow in the destination environment.

We see that the simulation data fall between the two models but, indeed, grow proportional to the square root of the network size. This leads us to the interesting conclusion that balancing the flows on links in broadcast networks is not the best thing to do (as it was in traditional networks). This derives from two factors. First, we should attempt to balance total *nodal* traffic rates rather than link flows. If we were able to do so (and not affect the average path length), the per-

formance would be as predicted by (32). We were unable to devise an algorithm that accomplished this, however. Second, even balanced nodal flows are not really appropriate since when we have a large user in an ALOHA network the maximum throughput is greater than if all users are equally loaded. This fact makes determination of optimal routes for these broadcast networks much more difficult than in traditional networks where the link capacity is not a function of the load.

We also plot a single data point for a grid network with a higher degree $N = 9$. We note that the capacity has been significantly reduced for this case.

V. CONCLUSIONS

In this paper we were concerned with the capacity of broadcast slotted ALOHA networks with regular structure. We were interested in determining whether we could benefit from *spatial reuse*. We investigated the effect of the average degree on network performance and also determined optimal transmission policies. In particular we found the following.

1) For one-dimensional networks the throughput is (almost) independent of the degree and a capacity of c/e can be achieved, with the constant c depending on the form of the traffic matrix.

2) For loop networks the capacity is $2/e$ for average degrees of the order of the square root of the number of nodes in the network.

3) For two-dimensional networks we find that spatial reuse allows a capacity proportional to the square root of the number of nodes in the network to be achieved for small average degrees.

4) The intuitive routing of balancing link flows was shown to be nonoptimal for a broadcast grid network.

This last finding has an impact on the optimal routes for broadcast networks in a more general sense. We can no longer expect the traditional optimal routing algorithms such as the flow deviation algorithm [3] to apply, since link capacity is now a function of the carried load (and of the load carried on neighboring links).

The most important result of this paper is that spatial reuse is indeed beneficial for these regularly structured broadcast networks and that low connectivity is a desirable quality for increasing capacity. This result is similar to that predicted for random networks in [8]. In the light of the importance of the nonuniformity of traffic flows as discussed here, a more detailed model of random networks is warranted.

APPENDIX

AVERAGE PATH LENGTH IN A SQUARE GRID NETWORK

Consider two points (i_1, j_1) and (i_2, j_2) in an $m \times m$ square network with Manhattan norm. Then the distance between these points is given by

$$d = |i_1 - i_2| + |j_1 - j_2| \quad 1 \leqslant i_1, i_2, j_1, j_2 \leqslant m. \tag{A1}$$

The mean distance is given by

$$\begin{aligned} E[d] &= \frac{1}{m^2}\sum_{i_1=1}^{m}\sum_{j_1=1}^{m}\frac{1}{m^2-1}\sum_{i_2=1}^{m}\sum_{j_2=1}^{m}|i_1-i_2|+|j_1-j_2| \\ &= \frac{1}{m^2-1}\left\{\sum_{i_1=1}^{m}\sum_{i_2=1}^{m}|i_1-i_2| + \sum_{j_1=1}^{m}\sum_{j_2=1}^{m}|j_1-j_2|\right\} \\ &= \frac{2}{m^2-1}\sum_{i_1=1}^{m}\sum_{i_2=1}^{m}|i_1-i_2| \\ &= \frac{2}{m^2-1}\sum_{i_1=1}^{m}\left\{\sum_{i_2=1}^{i_1}(i_1-i_2) + \sum_{i_2=i_1+1}^{m}(i_2-i_1)\right\} \\ &= \frac{2}{m^2-1}\sum_{i_1=1}^{m}\left\{ i_1{}^2 - \frac{i_1(i_1+1)}{2} + \frac{m(m+1)}{2} - \frac{i_1(i_1+1)}{2} - (m-i_1)i_1\right\} \\ &= \frac{1}{m^2-1}\left\{m^2(m+1) - 2(m+1)\sum_{i=1}^{m} i + 2\sum_{i=1}^{m} i^2\right\} \end{aligned} \tag{A2}$$

which, after some algebraic manipulation, gives

$$E[d] = \tfrac{2}{3}\, m. \tag{A3}$$

REFERENCES

[1] N. Abramson, "The ALOHA system—Another alternative for computer communications," in *AFIPS Conf. Proc., Fall Joint Comput. Conf.*, 1970, vol. 37, pp. 281–285.
[2] G. Y. Akavia, "Hierarchical organization of distributed packet-switching communication systems," Ph.D. dissertation, Dep. Comput. Sci., Univ. California, Los Angeles, 1978.
[3] M. Gerla, "The design of store and forward networks for computer communications," Ph.D. dissertation, School Eng. Appl. Sci., Univ. California, Los Angeles, Jan. 1973 (NTIS Doc. AD758-704).
[4] I. Gitman, "On the capacity of slotted ALOHA networks and some design problems," *IEEE Trans. Commun.*, vol. COM-23, pp. 305–317, Mar. 1975.
[5] R. E. Kahn, "The organization of computer resources into a packet radio network," *IEEE Trans. Commun.*, vol. COM-25, pp. 169–178, Jan. 1977.
[6] R. E. Kahn *et al.*, "Advances in packet radio technology," *Proc. IEEE*, vol. 66, Nov. 1978.
[7] L. Kleinrock, *Queueing Systems, Vol. 2: Computer Applications*. New York: Wiley, 1976.
[8] L. Kleinrock and J. A. Silvester, "Optimum transmission radii in packet radio networks or why six is a magic number," in *Proc. Nat. Telecommun. Conf.*, Birmingham, AL, Dec. 1978.
[9] F. F. Kuo and N. Abramson, "Some advances in radio communications for computers," in *Dig. Papers, 7th Annu. IEEE Comput. Soc. Int. Conf.*, San Francisco, CA, Feb. 1973.
[10] S. Lam, "Packet switching in a multi-access broadcast channel with applications to satellite communication in a computer network," Comput. Syst. Modeling Anal. Group, School Eng. Appl. Sci., Univ. California, Los Angeles, UCLA-ENG-7429, Apr. 1974 (also, Ph.D. dissertation, Dep. Comput. Sci., U.C.L.A.).
[11] L. G. Roberts, "ALOHA packet system with and without slots and capture," ARPA Network Inform. Cen., Stanford Res. Inst., Menlo Park, CA, ASS Note 8 (NIC 112390), June 1972.

[12] ——, "Extensions of packet communication technology to a hand held personal terminal," in *AFIPS Conf. Proc., Spring Joint Comput. Conf.*, 1972, vol. 40, pp. 295–298.
[13] ——, "ALOHA packets system with and without slots and capture," *Comput. Commun. Rev.*, vol. 5, pp. 28–42, Apr. 1975.
[14] J. A. Silvester, "On the spatial capacity of packet radio networks," Dep. Comput. Sci., Univ. California, Los Angeles, U.C.L.A. Modeling Anal. Rep. UCLA-ENG-8021, May 1980.
[15] F. A. Tobagi, "Analysis of a two-hop centralized packet radio network—Part I: Slotted ALOHA," *IEEE Trans. Commun.*, vol. COM-28, pp. 196–207, Feb. 1980.

★

John A. Silvester (M'79) was born in Kent, England, in 1950. He received the B.A. and M.A. degrees in mathematics and operations research from Cambridge University, Cambridge, England, in 1971 and 1975, respectively, the M.S. degree in statistics and computer science from West Virginia University, Morgantown, in 1973, and the Ph.D. degree in computer science from the University of California at Los Angeles in 1980. His Ph.D. dissertation, supervised by Prof. L. Kleinrock, was related to performance studies of packet radio networks.

In January 1979 he joined the University of Southern California as an Assistant Professor of Electrical Engineering, where he teaches courses in computer architecture, queueing theory, and computer networks and is actively engaged in research in broadcast communication networks. He also consults for several companies and government agencies on related topics.

Leonard Kleinrock (S'55–M'64–SM'71–F'73) received the B.S. degree in electrical engineering from the City College of New York, New York, NY, in 1957, and the M.S.E.E. and Ph.D.E.E. degrees from the Massachusetts Institute of Technology, Cambridge, in 1959 and 1963, respectively.

While at M.I.T., he worked at the Research Laboratory for Electronics as well as with the Computer Research Group of Lincoln Laboratory, Lexington, MA, in advanced technology. He joined the faculty at the University of California at Los Angeles in 1963. His research interests focus on computer networks, packet radio systems, and local area networks. He has had over 120 papers published and is the author of three books, *Communication Nets: Stochastic Message Flow and Delay* (New York: McGraw-Hill, 1964), *Queueing Systems, Vol. I: Theory* (New York: Wiley, 1975), and *Queueing Systems, Vol. II: Computer Applications* (Wiley, 1976), as well as the *Solutions Manual for Queueing Systems, Vol. I* (Wiley, 1982). He served as the Head of the U.C.L.A. Department of Computer Science Research Laboratory and is a well-known lecturer in the computer industry. He is Principal Investigator for the ARPA Advanced Teleprocessing Systems contract at U.C.L.A. and Co-Principal Investigator for the NSF Advanced Network Environment for Distributed Systems Research Project.

Prof. Kleinrock was recently elected to the National Academy of Engineering, is a Guggenheim Fellow, and serves on the Boards of Governors of various advisory councils in the computer field. He is a member of the Science Advisory Committee for IBM. He has received numerous best paper and teaching awards, including the ICC '78 Prize Winning Paper Award, the 1976 Lanchester Prize for outstanding work in Operations Research, and the IEEE Communications Society 1975 Leonard G. Abraham Prize Paper Award. In 1982, as well as having been selected to receive the C.C.N.Y. Townsend Harris Medal, he was co-winner of the L. M. Ericsson Prize, presented by His Majesty, King Carl Gustaf of Sweden, for his outstanding contributions in packet switching technology.