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

Some Extremal Markov Chains

By J. E. MAZO

(Manuscript received March 10, 1982)

Using a Markov chain model for the motion of a particle through a V-node network, we consider the quantities n_{ij}, which are the average number of steps taken by the particle in traveling from an originating node, i, to a destination node, j. A figure-of-merit, N, for the entire network is introduced by averaging n_{ij} over i and j. We investigate which networks minimize or maximize N, either when no restriction is placed on the Markov chain, or when we restrict it so that it corresponds to random routing. By the latter we mean that at each node the particle "selects at random" lines from an undirected network graph. We show that for random routing, the complete graph has $N = (V - 1)$ and is the minimizing graph. The maximizing graph is unknown, but we establish that the worst behavior of N increases at least with the cube of the number of vertices, but no worse than the 3.5 power. Properties of the class of graphs known as barbells are useful here. The minimizing unrestricted chain corresponds to placing the nodes on a circle and proceeding unidirectionally from one node to the next. Here, $N = V/2$.

I. INTRODUCTION AND RESULTS

Our work is set against the background of the Markov chain model for the movement of a "particle" or "message" through a network of V nodes. Thus, suppose a particle originates at node i and is destined for node $j \neq i$. The particle wanders through the network toward its destination via a Markov chain, going to node n from node m with probability p_{mn}. The quantity p_{mn} is the (m, n) element of the transition matrix P of the Markov chain whose states correspond to the nodes of the network. Denote the average number of steps required by the particle to reach its destination by n_{ij}, and assume that any node is accessible from any other (the Markov chain is irreducible). Then we introduce the figure-of-merit, N, for any such chain by averaging n_{ij} over the $(V - 1)$ possible destinations and V origins of particles:

$$N \equiv \frac{1}{V(V-1)} \sum_{i=1}^{V} \sum_{\substack{j=1 \\ j\neq i}}^{V} n_{ij}. \tag{1}$$

Our problem is: Which chains minimize or maximize N, either for a random-walk Markov chain or for an unrestricted chain?

An unrestricted chain means that no restrictions are placed on the transition matrix P other than irreducibility and the requirement that $p_{ii} = 0$ for all i. The random-walk chain is a special case of interest and is defined as follows. Draw any connected undirected graph on the V nodes. If m and n are not joined by an edge of this graph set $p_{mn} = 0$, while if they are so joined set $p_{mn} = 1/l_m$, where l_m is the number of edges of the graph leaving node m. Thus, at any node, the particle chooses from the available lines "at random." This random walk on the graph has previously been used by Kleinrock in Ref. 1, where it is referred to as random routing.

Our results are:

(*i*) For random routing, the average number of steps N is minimized when the graph of the network is the complete graph. For this case $N = (V-1)$.

(*ii*) For the unrestricted chain, N is minimized when the nodes are placed consecutively on a circle and we proceed deterministically from 1 to 2 to 3, etc. Here $N = V/2$.

(*iii*) By choosing $p_{12} = p_{21} = 1 - \epsilon$, $\epsilon \to 0$, N can be arbitrarily large for the unrestricted chain (when $V > 2$).

(*iv*) We have not been able to determine the "worst" graph for random routing, but we can show for large V that $O(V^3) \leq N_{\text{worst}} \leq O(V^{3.5})$. The barbell graphs of Mitra-Weiss[2] and Landau-Odlyzko[3] are good candidates for bad graphs.

II. THE MINIMAL WALK

In this section, we demonstrate that the complete graph is the only graph which minimizes N for the random-walk problem.

The fact that the symmetry of the complete graph requires $n_{ij} = N$ for all $i \neq j$ allows a simple demonstration of the fact $N = (V-1)$ for this case. If the particle originates at node i, we go directly to our destination j with probability $1/(V-1)$, requiring only one step, or we go to another node with probability $(V-2)/(V-1)$ and then require an average of $(1 + N)$ steps to reach j from i. Thus,

$$N = \frac{1}{V-1} + \frac{V-2}{V-1}[1 + N]. \tag{2}$$

Solving (2) yields $N = (V-1)$.

To show $N > (V-1)$ for other graphs is more involved. We first give properties of those transition matrices that correspond to random routing, ending with (12) which gives the stationary probabilities for those chains. We then derive (24), which is an expression for the average of the first passage times with which we are concerned. Finally, we obtain our result by giving a lower bound to (24).

Standard results on Markov chains or positive matrices may be used without reference when needed. For the former, the reader may consult Ref. 4, while Ref. 6 is a useful source for matrices.

In this paper, we denote transposition, complex conjugation, and hermitian conjugate by the symbols T, $*$, and $\dagger$, respectively.

For random routing, the transition matrix P is given by

$$P = DA, \tag{3}$$

where A is the symmetric adjacency matrix of the graph and is defined by

$$a_{ij} = \begin{cases} 1 & \text{if } i \neq j \text{ and } i \text{ and } j \text{ are joined} \\ & \text{by an edge} \\ 0 & \text{otherwise,} \end{cases}$$

and D is a diagonal matrix with diagonal elements

$$d_{ii} = 1/l_i, \tag{4}$$

l_i being the ith row sum of A and also, therefore, the number of edges of the graph incident on node i. The matrix P is stochastic, that is, it has nonnegative elements and the rows sum to one. Further, the assumed connectivity of the network graph implies that P is irreducible. The facts that P is stochastic and irreducible imply that the largest positive eigenvalue of P is unity and has multiplicity one. All other eigenvalues of P have modulus less than, or equal to, one. Set $\lambda_1 = 1$ and let $\lambda_2, \cdots, \lambda_V$ be the remaining eigenvalues of P.

We now investigate the eigenvalue and eigenvector structure of P. The matrices $P = DA$ and $Q = D^{1/2}AD^{1/2}$ differ by a similarity transformation ($P = D^{1/2}QD^{-1/2}$) and so have the same eigenvalues. Since Q is real symmetric, the λ_i are real. Further Q has a complete set of orthonormal eigenvectors $\phi^{(i)}$

$$\begin{aligned} Q\phi^{(i)} &= \lambda_i\phi^{(i)} \\ \phi^{(i)\dagger}\phi^{(j)} &= \delta_{ij}. \end{aligned} \tag{5}$$

If we denote the eigenvectors of P and P^T, which correspond to λ_i by $\mathbf{U}^{(i)}$ and $\mathbf{W}^{(i)}$, respectively, then we clearly have

$$\begin{aligned} \mathbf{U}^{(i)} &= \alpha_i D^{1/2}\phi^{(i)} \\ \mathbf{W}^{(i)} &= \alpha_i^{-1} D^{-1/2}\phi^{(i)*} \\ \mathbf{W}^{(i)T}\mathbf{U}^{(j)} &= \delta_{ij}. \end{aligned} \tag{6}$$

The α_i in (6) are any convenient constants. In addition, we have the spectral representation

$$P = \sum_{i=1}^{V} \lambda_i \mathbf{U}^{(i)} \mathbf{W}^{(i)T}. \tag{7}$$

Since the rows of P sum to unity, the eigenvector $\mathbf{U}^{(1)}$ may be chosen to be

$$\mathbf{U}^{(1)} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}. \tag{8}$$

The eigenvector $\mathbf{W}^{(1)}$,

$$P^T \mathbf{W}^{(1)} = \mathbf{W}^{(1)}, \tag{9}$$

is then the stationary probability vector for the Markov chain

$$\mathbf{W}^{(1)} = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_V \end{pmatrix}, \tag{10}$$

obeying the normalization of (6).

Equation (9) is easy to solve when P corresponds to random routing. Define a vector $\mathbf{Y}$ by $(\mathbf{Y})_i = l_i$, so $D\mathbf{Y} = \mathbf{U}^{(1)}$. Then, using (3),

$$P^T \mathbf{Y} = AD\mathbf{Y} = A\mathbf{U}^{(1)} = \mathbf{Y}. \tag{11}$$

Thus, the stationary probability vector for the chain has components

$$p_i = \frac{l_i}{2L} > 0, \tag{12}$$

where L equals the number of edges in the graph. Using (12) in (6), we find

$$(\phi^{(1)})_i = \sqrt{p_i}. \tag{13}$$

We next derive the expression (24) for N that will be our point of departure. Let $f_{ij}(n)$, $n = 1, 2, \cdots$ be the first passage probability from node i to node j at time n. Then, if P^n is the nth power of the transition matrix P, we have (Feller, in Ref. 4, p. 352)

$$\begin{aligned} f_{ij}(1) &= P_{ij} \\ f_{ij}(n) &= (P^n)_{ij} - \sum_{k=1}^{n-1} (P^k)_{jj} f_{ij}(n-k), \qquad n \geq 2. \end{aligned} \tag{14}$$

In terms of the generating functions, defined for $|s| < 1$ by

$$F_{ij}(s) = \sum_{n=1}^{\infty} f_{ij}(n)s^n \tag{15}$$

$$P_{ij}(s) = \sum_{n=1}^{\infty} (P^n)_{ij}s^n, \tag{16}$$

(14) is equivalent to

$$F_{ij}(s) = \frac{P_{ij}(s)}{1 + P_{jj}(s)}, \qquad |s| < 1. \tag{17}$$

Since all eigenvalues λ_i of P satisfy $|\lambda_i| \leq 1$, (16) shows that (17) may be rewritten for $i \neq j$ as

$$F_{ij}(s) = \frac{[I - sP]_{ij}^{-1}}{[I - sP]_{jj}^{-1}}, \qquad \begin{matrix} i \neq j \\ |s| < 1. \end{matrix} \tag{18}$$

Denoting the ith component of $\mathbf{U}^{(\mu)}$ by $U_i^{(\mu)}$, and similarly for $\mathbf{W}^{(\mu)}$, set

$$U_i^{(\mu)} W_j^{(\mu)} = b_{ij}^{(\mu)} \tag{19}$$

and note from (8) and (10) that

$$b_{ij}^{(1)} = p_j. \tag{20}$$

Then, using the spectral representation (7), (18) may be written

$$F_{ij}(s) = \frac{p_j + (1 - s) \sum_{\mu=2}^{V} \frac{1}{1 - s\lambda_\mu} b_{ij}^{(\mu)}}{p_j + (1 - s) \sum_{\mu=2}^{V} \frac{1}{1 - s\lambda_\mu} b_{jj}^{(\mu)}}. \tag{21}$$

In this form, $F_{ij}(s)$ may be analytically continued to a neighborhood of $s = 1$. Then, from (15) and a standard Abelian theorem we have

$$n_{ij} = \sum_{n=1}^{\infty} n f_{ij}(n) = \frac{dF_{ij}(s = 1)}{ds}. \tag{22}$$

Using (21) to (22), we obtain

$$n_{ij} = \frac{1}{p_j} \sum_{\mu=2}^{V} \frac{1}{1 - \lambda_\mu} (b_{jj}^{(\mu)} - b_{ij}^{(\mu)}). \tag{23}$$

Finally, using the definition (1) and relations (6) and (19), we see that the average number of steps, N, required to reach a destination with random routing for V nodes may be written

$$N = \frac{1}{V - 1} \sum_{\mu=2}^{V} \frac{1}{1 - \lambda_\mu} \left[\sum_{j=1}^{V} \frac{1}{p_j} |\phi_j^{(\mu)}|^2 - \frac{1}{V} \left| \sum_j \frac{\phi_j^{(\mu)}}{\sqrt{p_j}} \right|^2 \right]. \tag{24}$$

Again, p_j are the stationary probabilities of the Markov chain governed by transition matrix $P = DA$. The λ_μ, $\mu \geq 2$, are the nonunity eigenvalues of P and $Q = D^{1/2}AD^{1/2}$, the latter matrix having orthonormal eigenvectors $\phi^{(\mu)}$. Also $\phi^{(1)}$, the eigenvector of Q associated with the eigenvalue $\lambda_1 = 1$, is given by (13).

We now lower bound the right member of (24). A few known facts about the quantities λ_μ, $\phi^{(\mu)}$, and p_j will be exploited, but their great interdependence (they are all determined by P) will be ignored.

Set

$$a_\mu = \sum_{j=1}^{V} \frac{1}{p_j} |\phi_j^{(\mu)}|^2 - \frac{1}{V} \left| \sum_j \frac{\phi_j^{(\mu)}}{\sqrt{p_j}} \right|^2 \tag{25}$$

and rewrite (24) as

$$N = \frac{1}{V-1} \sum_{\mu=2}^{V} \frac{a_\mu}{1-\lambda_\mu}. \tag{26}$$

We assert $a_\mu > 0$, since, by Cauchy's inequality,

$$\frac{1}{V} \left| \sum_j \frac{\phi_j^{(\mu)}}{\sqrt{p_j}} \cdot 1 \right|^2 \leq \frac{1}{V} \left(\sum_j \frac{|\phi_j^{(\mu)}|^2}{p_j} \right) \left(\sum_j 1 \right) = \sum_j \frac{|\phi_j^{(\mu)}|^2}{p_j}. \tag{27}$$

However, equality can hold only if

$$\frac{\phi_j^{(\mu)}}{\sqrt{p_j}} \propto 1,$$

that is, if

$$\phi_j^{(\mu)} \propto \sqrt{p_j} = \phi_j^{(1)}.$$

Since $\phi^{(\mu)}$ is orthogonal to $\phi^{(1)}$ for $\mu \geq 2$, this cannot happen.†

With the positivity of the a_μ established, minimize (26) over the λ_μ holding the $a_\mu > 0$ fixed. During this minimization, we only impose two constraints on the $\lambda^{(\mu)}$. First, $\lambda_\mu < 1$, and second $\sum_{\mu=2}^{V} \lambda_\mu = -1$. The first has been amply discussed already while the second follows from the fact that the diagonal elements of the transition matrix P are zero and so

$$0 = \operatorname{tr} P = \sum_{\mu=1}^{V} \lambda_\mu. \tag{28}$$

Since $\lambda_1 = 1$, the second constraint follows. Introducing the second constraint via a Lagrange multiplier (call it $-\beta$), we find that the unique stationary point of

† Incidentally, if we write $n_{ij} = \sum_\mu \tau_{ij}^{(\mu)}/(1 - \lambda_\mu)$ it is not necessary that $\tau_{ij}^{(\mu)} \geq 0$. However, $\tau_{ij}^{(\mu)} + \tau_{ji}^{(\mu)} \geq 0$.

$$\sum_{\mu=2}^{V} \frac{a_\mu}{1-\lambda_\mu} - \beta\Sigma\lambda_\mu, \tag{29}$$

satisfying the two constraints, occurs when

$$\lambda_\mu = 1 - \sqrt{a_\mu}\left(\frac{V}{\sum_{\nu=2}^{V} \sqrt{a_\nu}}\right). \tag{30}$$

Therefore, applying (30) to (26) gives

$$N \geq \frac{1}{V(V-1)}\left(\sum_{\mu=2}^{V} \sqrt{a_\mu}\right)^2. \tag{31}$$

Next, use the fact that the arithmetic mean exceeds the geometric mean:

$$\begin{aligned} N &\geq \frac{V-1}{V}\left(\frac{1}{V-1}\sum_{\mu=2}^{V} \sqrt{a_\mu}\right)^2 \\ &\geq \frac{V-1}{V}\left[\left(\prod_{2}^{V} \sqrt{a_\mu}\right)^{\frac{1}{V-1}}\right]^2 = \frac{V-1}{V}\left(\prod_{\mu=2}^{V} a_\mu\right)^{\frac{1}{V-1}}. \end{aligned} \tag{32}$$

Motivated by the definition of the a_μ [see (25)], introduce the real symmetric matrix M having elements

$$M_{ij} = \frac{1}{p_i}\,\delta_{ij} - \frac{1}{V}\frac{1}{\sqrt{p_i p_j}}. \tag{33}$$

Then,

$$a_\mu = \phi^{(\mu)\dagger} M \phi^{(\mu)}. \tag{34}$$

Note that

$$M\phi^{(1)} = \sum_{j=1}^{V}\left(\frac{1}{p_i}\,\delta_{ij} - \frac{1}{V}\frac{1}{\sqrt{p_i p_j}}\right)\sqrt{p_j} = 0, \tag{35}$$

while we have already shown

$$\phi^\dagger M \phi > 0 \quad \text{if} \quad (\phi, \phi^{(1)}) = 0. \tag{36}$$

Thus, M is positive semi-definite, having precisely one zero eigenvalue. Fix $p_i > 0$, $i = 1, \cdots, V$, and proceed to lower bound (32) by writing

$$\prod_{\mu=2}^{V} a_\mu \geq \min \prod_{\nu=2}^{V} \psi^{(\nu)\dagger} M \psi^{(\nu)}, \tag{37}$$

the minimum in (37) being taken over all orthonormal sets of vectors $\psi^{(\nu)}$ which satisfy

$$(\psi^{(\mu)}, \psi^{(\nu)}) = \delta_{\mu\nu}$$

$$(\psi^{(\mu)}, \phi^{(1)}) = 0, \qquad \mu, \nu = 2, \cdots, V. \tag{38}$$

An inequality of Ky Fan[5] implies that under the conditions of (38) we have

$$\min \prod_{\nu=2}^{V} \psi^{(\nu)\dagger} M \psi^{(\nu)} = \mu_1 \mu_2, \cdots, \mu_{V-1}, \tag{39}$$

where μ_i, $i = 1, \cdots, V-1$ are the nonzero eigenvalues of the matrix M, which was defined in (33). Thus,

$$N \geq \frac{V-1}{V} [\mu_1 \mu_2 \cdots \mu_{V-1}]^{\frac{1}{V-1}}. \tag{40}$$

If g_i are the components of an eigenvector of M associated with eigenvalue μ, the equation

$$\sum_{j=1}^{V} M_{ij} g_j = \mu g_i \tag{41}$$

yields, when (33) is substituted for M_{ij},

$$g_i = k \frac{1}{\sqrt{p_i}\left(\frac{1}{p_i} - \mu\right)}, \tag{42}$$

k being a normalization constant. Since for $\mu \neq 0$ we must have

$$0 = \sum_{i=1}^{V} g_i \phi_i^{(1)} = \sum_{i=1}^{V} g_i \sqrt{p_i}, \tag{43}$$

we obtain, using (42),

$$\sum_{i=1}^{V} \frac{1}{\left(\frac{1}{p_i} - \mu\right)} = 0. \tag{44}$$

Equation (44) determines the $(V-1)$ positive eigenvalues μ_i. Clearing fractions in (44) yields

$$\frac{1}{D} \sum_{i=1}^{V} \prod_{\substack{j=1 \\ j \neq i}}^{V} \left(\frac{1}{p_j} - \mu\right) = 0 \tag{45}$$

with

$$D = \prod_{i=1}^{V} \left(\frac{1}{p_i} - \mu\right). \tag{46}$$

The denominator D in (45) may be discarded and then the product of the nonzero eigenvalues of M is simply read off the remaining polynomial. Since the coefficient of μ^{V-1} in the numerator polynomial in (45) is $V(-1)^{V-1}$,

$$\mu_1 \cdots \mu_{V-1} = \frac{1}{V} \sum_{i=1}^{V} \prod_{\substack{j=1 \\ j \neq i}}^{V} \frac{1}{p_j}. \tag{47}$$

The product of positive numbers having a fixed sum is maximized when each number is the same, so we have

$$\prod_{\substack{j=1 \\ j \neq i}}^{V} p_j \leq \left[\frac{(1 - p_i)}{V - 1} \right]^{V-1}$$

or

$$\mu_1 \cdots \mu_{V-1} \geq \frac{1}{V} \sum_{i=1}^{V} \left(\frac{V-1}{1-p_i} \right)^{V-1}. \tag{48}$$

Finally, the minimum of (48) subject to $\sum_{i=1}^{V} p_i = 1$ occurs when $p_i = \frac{1}{V}$, or,

$$\mu_1 \cdots \mu_{V-1} \geq \frac{1}{V} \sum_{i=1}^{V} \left(\frac{V-1}{1 - \frac{1}{V}} \right)^{V-1} = V^{V-1}. \tag{49}$$

Combining (49) with (40) produces our desired result:

$$N \geq \frac{V-1}{V} (V^{V-1})^{\frac{1}{V-1}} = (V-1). \tag{50}$$

It is easy to work backwards through the argument to see that the complete graph is the only one that can achieve $N = (V-1)$. Clearly, equality in the last step of the argument can only be attained if $p_i = \frac{1}{V}$. Substituting this into (24) and using the orthonormality of the $\phi^{(\mu)}$ yields

$$N = \frac{V-1}{V} \sum_{\mu=2}^{V} \frac{1}{1 - \lambda_\mu}, \tag{51}$$

which only equals the minimum when $\lambda_\mu = -\frac{1}{V-1}$, $\mu = 2, \cdots, V$. Using this, and the fact that $\lambda_1 = 1$ in (7) yields

$$P = \mathbf{U}^{(1)} \mathbf{W}^{(1)T} - \frac{1}{V-1} \sum_{i=2}^{V} \mathbf{U}^{(i)} \mathbf{W}^{(i)T}$$

$$= -\frac{1}{V-1} \sum_{i=1}^{V} \mathbf{U}^{(i)} \mathbf{W}^{(i)T} + \frac{V}{V-1} \mathbf{U}^{(1)} \mathbf{W}^{(1)T}. \tag{52}$$

The first dyad sum in the right member of (52) is the identity matrix, while the last dyad in (52) is the matrix that has all elements equal to $1/V$. Thus, P is precisely the transition matrix associated with the complete graph.

III. THE MINIMAL CHAIN

Now we turn to finding the minimum value of N for the unrestricted chain. In the introduction we stated that for this we should place the nodes on a circle and go unidirectionally from one node to the next. Clearly, in this case, $\sum_j n_{ij}$ does not depend on i, and, in fact,

$$N = \frac{1}{V-1} \sum_{j=2}^{V} n_{1j} = \frac{1}{V-1} [1 + 2 + 3 + \cdots + (V-1)] = \frac{V}{2}. \tag{53}$$

We must now show that no other setup can do as well. For this end, expressions such as (23) are not useful since the λ_μ, $\mathbf{U}^{(\mu)}$ and $\mathbf{W}^{(\mu)}$ may be complex and it would be difficult to pick out even real combinations, let alone positive ones that might be lower bounded. Rather, we retreat to an obvious generalization of the argument we used to derive (2). This generalization reads

$$n_{ij} = p_{ij} + \sum_{\substack{k \\ k \neq j}} p_{ik}(1 + n_{kj}), \qquad \begin{matrix} i, j = 1, \cdots, V \\ j \neq i \end{matrix} \tag{54}$$

or, since P is stochastic,

$$n_{ij} - \sum_{\substack{k \\ k \neq j}} p_{ik} n_{kj} = 1, \qquad \begin{matrix} i, j = 1, \cdots, V. \\ j \neq i \end{matrix} \tag{55}$$

Equation (55) is our new point of departure.

Let $\mathbf{N}^{(j)}$, $j = 1, \cdots, V$, be the $(V-1)$ dimensional vector whose components are n_{ij}, with j fixed and $i \neq j$. Also let $\tilde{P}(j)$, $j = 1, \cdots, V$, be the $(V-1) \times (V-1)$ matrix obtained by crossing out the jth row and the jth column of P. The $\tilde{P}(j)$ still have positive elements, but the irreducibility of P implies that not all rows of $\tilde{P}(j)$ can sum to one, and so the largest eigenvalue of $\tilde{P}(j)$ is strictly less than one. The equations represented by (55) may now be written

$$[I - \tilde{P}(j)]\mathbf{N}^{(j)} = \mathbf{u}, \qquad j = 1, \cdots, V, \tag{56}$$

where $\mathbf{u}$ is the $(V-1)$ dimensional vector having all components unity. From (56),

$$\mathbf{N}^{(j)} = [I - \tilde{P}(j)]^{-1}\mathbf{u} = [I + \tilde{P}(j) + \tilde{P}^2(j) + \cdots]\mathbf{u}. \tag{57}$$

If $\tilde{P}^k_{rs}(j)$ denotes the (r, s) element of the kth power of $\tilde{P}(j)$, then (1) and (57) yield

$$N = \frac{1}{V(V-1)} \sum_{j=1}^{V} \sum_{r,s=1}^{V-1} \sum_{k=0}^{\infty} \tilde{P}_{rs}^{k}(j), \tag{58}$$

where by $\tilde{P}^0(j)$ we mean the identity matrix for $(V-1)$ dimensions. Equation (58) succumbs to the application of the following:

Lemma: Let P be any stochastic $V \times V$ matrix and let $\tilde{P}(j)$, $j = 1, \cdots, V$ be the $(V-1) \times (V-1)$ matrix obtained by crossing out the jth row and jth column of P. Then

$$\sum_{j=1}^{V} \sum_{r,s=1}^{V-1} \tilde{P}_{rs}^{k}(j) \geq \begin{cases} 0 & \textit{for} \quad k \geq 0 \\ V(V-k-1) & \textit{for} \quad 0 \leq k \leq V-2. \end{cases} \tag{59}$$

Proof: The positivity of the sum for all k is trivial since $\tilde{P}(j)$ has nonnegative elements. If $k = 0$, $P^0(j)$ is the identity for all j, so the sum of all its elements if $(V-1)$, and we obtain $V(V-1)$ when we sum over j. Next consider $k = 1$ and make use of the stochastic nature of P, that is $\sum_s P_{rs} = 1$.

$$\begin{aligned} \sum_{j} \sum_{r,s} \tilde{P}_{rs}(j) &= \sum_{j} \sum_{\substack{r \neq j \\ s \neq j}} P_{rs} = \sum_{j} \sum_{r \neq j} (1 - P_{rj}) \\ &= V(V-1) - \sum_{j} \sum_{r} P_{rj} + \sum_{r} P_{rr} \geq V(V-2). \end{aligned}$$

Now proceed by induction assuming that the lemma is true for k and show it true for $(k+1)$

$$\begin{aligned} \sum_{j=1}^{V} \sum_{r,s=1}^{V-1} \tilde{P}_{rs}^{k+1}(j) &= \sum_{j} \sum_{\substack{n_1 \neq j \\ \vdots \\ n_{k+2} \neq j}} P_{n_1 n_2} P_{n_2 n_3} \cdots P_{n_{k+1} n_{k+2}} \\ &= \sum_{j} \sum_{\substack{n_1 \neq j \\ \vdots \\ n_{k+1} \neq j}} P_{n_1 n_2} P_{n_2 n_3} \cdots P_{n_k n_{k+1}} \\ &\quad - \sum_{j} \sum_{\substack{n_1 \neq j \\ \vdots \\ n_{k+1} \neq j}} P_{n_1 n_2} \cdots P_{n_{k+1} j} \\ &\geq V(V-k-1) - \sum_{j} \sum_{n_1, n_2 \cdots} P_{n_1 n_2} \cdots P_{n_{k+1}, j}. \end{aligned} \tag{60}$$

The inequality follows from the induction assumption (true for k) and is further strengthened by extending the range of summation on the negative term. In the negative term, the sums over j, n_{k+1}, $\cdots$, n_2 yield unity, and, finally, the sum over n_1 gives V. The conclusions of the lemma follow.

Now apply the lemma to (59).

$$N \geq \frac{1}{V(V-1)} \sum_{k=0}^{V-2} \left[\sum_{j} \sum_{r,s} \tilde{P}_{rs}^{k}(j) \right]$$

$$\geq \frac{1}{V(V-1)} \sum_{k=0}^{V-2} V(V-k-1) = \frac{V}{2}, \tag{61}$$

as desired.

The condition for equality in (61) obviously requires that

$$\sum_{j} \sum_{\substack{n_1 \neq j \\ \vdots \\ n_k \neq j}} P_{n_1 n_2} \cdots P_{n_k j} = \sum_{j} \sum_{n_1, \cdots, n_k} P_{n_1 n_2} \cdots P_{n_k j}, \qquad k = 1, \cdots, V-1. \tag{62}$$

Since all terms on the right member of (62) are nonnegative, any term present there but not appearing on the left must be zero. In particular, we must have for any j

$$P_{jj} = 0 \qquad \text{if} \quad k = 1 \tag{63}$$

$$\sum_{n_2 \cdots n_k} P_{jn_2} \cdots P_{n_r j} = 0 \qquad k = 2, \cdots, V-1. \tag{64}$$

These equations state that it is impossible to return to any initial state j in less than V steps. This, plus irreducibility, implies the unidirectional movement or a circle.

The reason why the lemma is exact in this case is that each $\tilde{P}(j)$ is, then—except for a reordering of nodes—a canonical Jordan block with all zeros in the matrix except for $(V-2)$ ones on the appropriate off-diagonal.

IV. THE WORST WALK

We have been unable to describe the worst setup for random routing. For $V = 3$, 4, and 5, numerical work shows that arranging the vertices on a straight line gives the worst cases. In fact, for V nodes on a straight line it is possible to show that

$$N = \frac{V^2 - 1}{3}, \tag{65}$$

which is significantly worse than $(V-1)$, the best attainable with random routing. However, for large V one can do worse than (65). Our result for this situation is

$$O(V^3) \leq N_{\text{worst}} \leq O(V^{3.5}). \tag{66}$$

In (66) N_{worst} denotes the largest value of N obtained from any of the connected network graphs on the V vertices.

For the lower bound, assume $V = 6m - 1$ and consider the bar-

Fig. 1—An 11-node barbell.

bell graphs described by Landau-Odlyzko.[3] This class of graphs has $(2m-1)$ nodes connected in a straight line with complete graphs of $2m$ nodes attached to each end of the line by a single edge of the graph. If $m = 1$, this only describes a straight line, but for $m = 2$ or greater, the barbell nature is evident. The case for $V = 11$ is shown in Fig. 1.

In Ref. 3 the authors show that for such graphs λ_2, the second largest eigenvalue of $P = DA$, satisfies

$$\lambda_2 \geq 1 - \frac{\gamma}{V^3} \tag{67}$$

with $\gamma = 54 + O\left(\frac{1}{V}\right)$. Then from (26), (67), and (34),

$$\begin{aligned} N &\geq \frac{1}{V-1}\frac{a_2}{1-\lambda_2} \geq O(V^2)\phi^{(2)\dagger}M\phi^{(2)} \\ &\geq O(V^2)\mu_{V-1}, \end{aligned} \tag{68}$$

μ_{V-1} being the smallest nonzero eigenvalue of M. Equation (44) shows

$$\mu_{V-1} > \frac{1}{\max p_i}. \tag{69}$$

Since the nodes of the complete graphs in the barbell each have $O(V)$ incident edges and there are $O(V^2)$ edges in the barbell, (12) implies that $\max p_i = O\left(\frac{1}{V}\right)$. Equations (68) and (69) then give

$$N \geq O(V^2)O(V) = O(V^3). \tag{70}$$

Although we will not give the details here, it is possible to show that, in fact, $N = O(V^3)$ for barbell graphs.

Our upper bound will be based on (57), but first we need the result that the largest eigenvalue of $\tilde{P}(j)$ satisifes

$$\lambda_1(j) \equiv \lambda_1[\tilde{P}(j)] \leq 1 - O\left(\frac{1}{V^3}\right). \tag{71}$$

Just as the matrix $\tilde{P}(j)$ was defined in Section III, now introduce matrices $\tilde{D}(j)$ and $\tilde{A}(j)$ by eliminating the jth row and column of D and A, respectively. Then,

$$\tilde{P}(j) = \tilde{D}(j)\tilde{A}(j). \tag{72}$$

Also set

$$\tilde{Q}(j) = \tilde{D}^{1/2}(j)\tilde{A}(j)\tilde{D}^{1/2}(j) \tag{73}$$

so that $\tilde{Q}(j)$ is symmetric and has the same eigenvalues as $\tilde{P}(j)$. In fact, we have

$$\lambda_1[\tilde{P}(j)] = \lambda_1[\tilde{Q}(j)] = \max_{\sum y_i^2=1} \mathbf{y}^\dagger \tilde{Q}(j)\mathbf{y}$$

$$= \max_{\sum l_i z_i^2=1} \mathbf{z}^\dagger A(j)\mathbf{z} = \max_{\substack{x_j=0 \\ \sum l_i x_i^2=1}} \mathbf{x}^\dagger A\mathbf{x}. \tag{74}$$

Note that in (74) we have returned to the $V \times V$ adjacency matrix A by introducing the constraint $x_j = 0$. If S is the set of those (ordered) pairs of indices which corresponds to vertices connected by an edge of the graph, we have[†]

$$\mathbf{x}^\dagger A\mathbf{x} = \sum_{(m,n)\in S} x_m x_n = \frac{1}{2}\sum_S [x_m^2 + x_n^2 - (x_m - x_n)^2]$$

$$= \sum_m l_m x_m^2 - \frac{1}{2}\sum_S (x_m - x_n)^2. \tag{75}$$

Using (75) in (74) we easily obtain

$$\lambda_1(j) = 1 - \frac{1}{2} \min_{\substack{x_j=0 \\ \sum l_i x_i^2=1}} \sum_S (x_m - x_n)^2. \tag{76}$$

Let x_k be the component of x having the largest square. Then

$$1 = \sum l_i x_i^2 \leq (V-1) l_{\max} x_k^2, \tag{77}$$

or

$$x_k^2 \geq \frac{1}{l_{\max}(V-1)}, \tag{78}$$

where $l_{\max}$ (or $l_{\min}$) is the maximum (or minimum) of the l_i, $i = 1, \cdots, V$.

Now, from the connectivity of the graph, vertex j is attached to another vertex t. So, using (76),

$$\lambda_1(j) \leq 1 - x_t^2, \tag{79}$$

since $x_j = 0$. If $t = k$, (78) and (79) yield

$$\lambda_1(j) \leq 1 - \frac{1}{(V-1)l_{\max}} \qquad (t = k). \tag{80}$$

If $t \neq k$, there exists a chain of r distinct edges joing vertices t and k.

[†] The remainder of this demonstration is entirely inspired by Ref. 3.

Clearly, $r \leq d$, where d is the diameter of the graph. Then, using the basic trick of Ref. 3,

$$x_k - x_t = (x_k - x_{k_1}) + (x_{k_1} - x_{k_2}) + \cdots + (x_{k_{r-1}} - x_t) \tag{81}$$

and so, by Cauchy's inequality,

$$(x_k - x_t)^2 \leq \frac{r}{2} \sum_S (x_m - x_n)^2 \leq \frac{d}{2} \sum_S (x_m - x_n)^2. \tag{82}$$

Equations (76) and (82) thus yield

$$\lambda_1(j) \leq 1 - \frac{(x_k - x_t)^2}{d}. \tag{83}$$

Combine (79) and (83) by averaging and then minimize over the numerical value of x_t to obtain, with (78), for $t \neq k$,

$$\lambda_1(j) \leq 1 - \frac{1}{2}\left[x_t^2 + \frac{(x_k - x_t)^2}{d}\right] \leq 1 - \frac{1}{2(1 + d)(V - 1)l_{\max}}. \tag{84}$$

Both cases (80) and (84) are included in

$$\lambda_1(j) \leq 1 - \frac{1}{2(1 + d)(V - 1)l_{\max}} \leq 1 - \frac{1}{2V^3}, \tag{85}$$

since $d \leq (V - 1)$ and $l_{\max} \leq (V - 1)$.

To complete the upper bound on N, start with (58) and write (suppressing the j-dependence of the matrices in the notation)

$$\begin{aligned} N &= \frac{1}{V(V-1)} \sum_{j=1}^{V} \mathbf{u}^\dagger [I - \tilde{P}(j)]^{-1} \mathbf{u} \\ &\leq \frac{\|\mathbf{u}\|^2}{V(V-1)} \sum_j \|[I - \tilde{P}(j)]^{-1}\| \\ &= \frac{1}{V} \sum_j \|\tilde{D}^{1/2}[I - \tilde{Q}]^{-1}\tilde{D}^{1/2}\| \\ &\leq \frac{1}{V} \sum_j \|\tilde{D}^{1/2}\| \, \|\tilde{D}^{-1/2}\| \, \|[I - \tilde{Q}]^{-1}\| \\ &\leq \sqrt{\frac{l_{\max}}{l_{\min}}} \frac{1}{V} \sum_{j=1}^{V} \frac{1}{1 - \lambda_1(j)}. \end{aligned} \tag{86}$$

From $l_{\min} \geq 1$, $l_{\max} \leq (V - 1)$, (85) and (86) we thus have

$$N \leq 2V^3\sqrt{V - 1}, \tag{87}$$

as desired.

Is it possible that the complete graph is optimal for random routing because adding any edge to a graph decreases N? No. Many edges must be added to the straight-line graph to form the barbell, but the former has $N = O(V^2)$ while for the latter, $N = O(V^3)$.

REFERENCES

1. L. Kleinrock, *Communication Nets*, New York: McGraw-Hill, 1964, Chap. 6.
2. D. Mitra and A. Weiss, "Analysis of Delay-Differential Equations Arising in Communication Network Synchronization," in Proc. IEEE Int. Symp. Circuits and Systems, Houston, Texas, 1980, pp. 839–43.
3. H. J. Landau and A. M. Odlyzko, "Bounds for Eigenvalues of Certain Stochastic Matrices," Linear Algebra and Applications, *38* (June 1981), pp. 5–15.
4. W. Feller, *An Introduction to Probability Theory and Its Applications*, New York: John Wiley, 1957, 2nd ed.
5. E. F. Beckenbach and R. Bellman, *Inequalities*, New York: Springer-Verlag, 1965, p. 76, Theorem 24.
6. Peter Lancaster, *Theory of Matrices*, New York: Academic Press, 1969.