Technical Report No. 32-697

# Optimal Communication Nets

R. J. McEliece



JET PROPULSION LABORATORY
CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA

April 15, 1965

*Technical Report No. 32-697*

# *Optimal Communication Nets*

*R. J. McEliece*

Mahlon Easterling

Mahlon Easterling, Manager
*Communications Systems Research*

JET PROPULSION LABORATORY
CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA

April 15, 1965

Prepared Under Contract No. NAS 7-100
National Aeronautics & Space Administration

# CONTENTS



# FIGURES

## ABSTRACT

23278

The following problem is posed: If $\bar{\lambda}$ is the average separation between vertices in a finite, connected, undirected graph and $m$ is the total number of edges, which graph with $n$ vertices minimizes $m\bar{\lambda}$? The unique solution is shown to be the star graph. Variance of this is done with weighted paths and probabilistic weights. Author

## I. INTRODUCTION

In a recent book (Ref. 1), Kleinrock considers the following problem: given a set of $n$ terminals, and a pattern of the communication traffic between them (and certain other constraints), what would be the best possible configuration of communication links between the terminals? Here, "best possible" has been interpreted to mean the configuration which minimizes the mean time that a message is in the communication net. Kleinrock's discussion has suggested the following graph-theoretic investigations.

If the *distance* $\mu(a,b)$ between two vertices $a$ and $b$ of a graph is defined as the minimum number of edges in a path that joins the two vertices, then we may speak of the *average separation* $\bar{\lambda}$ in a finite, undirected, connected graph:

$$\bar{\lambda} = \frac{\sum_{(a,b)} \mu(a,b)}{\binom{n}{2}} \qquad (1)$$

The summation is taken over all unordered pairs of vertices $(a,b)$ where $a \neq b$, and $n = |G|$ is the number of vertices of $G$ (frequently called the *order* of $G$.) In this discussion, it will always be assumed that $G$ is connected and undirected. The problem of finding the minimum of this quantity with respect to all graphs on $n$ vertices is a trivial one; $\bar{\lambda} = 1$ when and only when $G$ is the complete graph $U_n$ on $n$ vertices. (The problem of maximizing $\bar{\lambda}$ is less easy, but the answer is that $\bar{\lambda}_{max} = (n+1)/2$, attained by the chain $L_n$, to be defined below.) But it seems as if $U_n$ "uses too many edges" in attaining the minimum and is therefore in some sense inefficient. This leads to the consideration of the quantity $m\bar{\lambda}$, where $m$ is the number of edges in $G$. Therefore, the problem of minimizing the quantity of $m\bar{\lambda}$ (for a fixed $n$) under various circumstances will be considered.

First note that in Eq. 1 $\bar{\lambda}$ was computed with the assumption that all paths were to be given equal weight; (that is, $2/n(n-1)$. In the first part of this discussion, the assumption will be retained; we will relax it later.

Now it is appropriate to compute the quantity $m\bar{\lambda}$ for several of the simplest graphs.

1. The *complete graph* $U_n$ on $n$ vertices:

$$m\bar{\lambda} = \binom{n}{2}\frac{\binom{n}{2}}{\binom{n}{2}} = \frac{n(n-1)}{2} \sim \frac{n^2}{2}$$

2. The *chain* $L_n$ (the graph formed by joining the vertices $V_i$ to $V_j$ if, and only if, $|i\text{-}j| = 1$):

$$m\bar{\lambda} = (n-1)\frac{\sum_{n \geqq i > j \geqq 1} (i-j)}{\binom{n}{2}} = \frac{2}{n}\cdot\frac{n}{6}(n^2-1)$$

$$= \frac{n^2-1}{3} \sim \frac{n^2}{3}$$

3. The *ring* $R_n$ (here the edges are $V_1V_2, V_2V_3, \cdots, V_{n-1}V_n, V_nV_1$):

Here it can be shown that

$$m\lambda = \begin{cases} \dfrac{n^3}{4(n-1)}, & n \text{ even} \\ \dfrac{n(n+1)}{4}, & n \text{ odd} \end{cases} \sim \frac{n^2}{4}$$

4. The *star* $S_{n-1}$ (Fig. 1):

$$m\bar{\lambda} = (n-1)\frac{(n-1) + 2[n(n-1)/2 - (n-1)]}{n(n-1)/2}$$

$$= \frac{2(n-1)^2}{n} \sim 2n$$

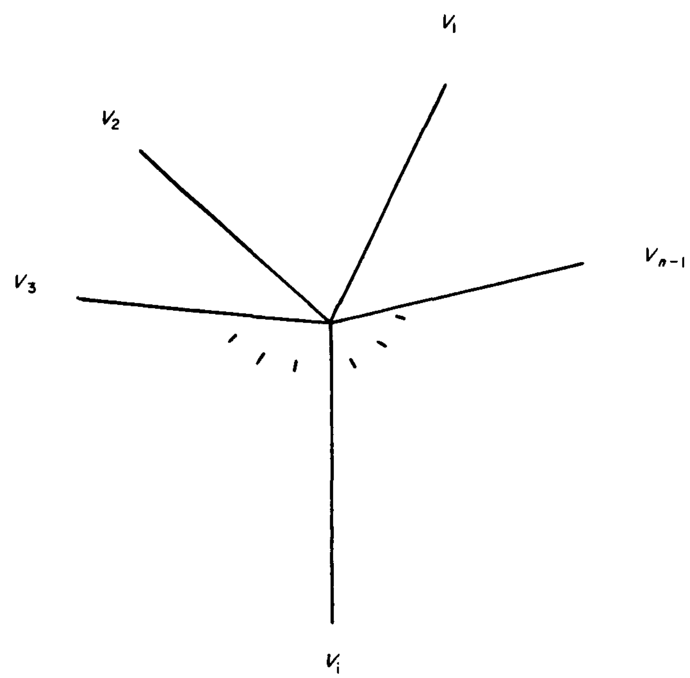


**Figure 1. The star $S_{n-1}$**

In the above examples, we see that only for the star $S_{n-1}$ does the quantity $m\bar{\lambda}$ behave linearly in $n$; for the rest it grows as $n^2$. (In fact, while the complete graph $U_n$ *minimizes* $\bar{\lambda}$, it exhibits the *maximum* $m\bar{\lambda}$ of all the examples we have given.) It can now be proved that $S_{n-1}$ does, in fact, minimize the quantity $m\bar{\lambda}$, whereas no other graph on $n$ vertices does.

## II. RESULTS IN THE CASE OF EQUAL WEIGHTS

### A. Theorem 1

If G has $n$ vertices and $m$ edges

$$m\bar{\lambda} \geqq 2m\left(1 - \frac{m}{n(n-1)}\right)$$

*Proof:* For a graph with $m$ edges, exactly $m$ of the distances between vertices will be 1. Hence the remaining $\binom{n}{2} - m$ paths must each be of length 2 or greater. Thus

$$m\bar{\lambda} \geqq m\frac{m + 2\left[\binom{n}{2} - m\right]}{\binom{n}{2}} = 2m\left(1 - \frac{m}{n(n-1)}\right)$$

This proves Theorem 1.

Now in order to obtain a *lower* bound on $m\bar{\lambda}$, which is independent of $m$ (depends only on $n$), the expression $2m[1 - m/n(n-1)]$ as a function of $m$ needs to be examined. The graph of this can be seen to be a parabola with maximum at $m = n(n-1)/2$, the maximum is also $n(n-1)/2$, and since $m \leqq n(n-1)/2$, the quantity $2m[1 - m/n(n-1)]$ will be minimized if $m$ is as small as possible. But it can be shown (Ref. 2) that $m \geqq n-1$ if the graph $G$ is connected. Consequently, we have proved Theorem 2.

### B. Theorem 2

If $G$ has $n$ vertices, then

$$m\bar{\lambda} \geqq \frac{2(n-1)^2}{n}$$

This theorem, in the light of the previous computation, shows that the star $S_{n-1}$ minimizes the quantity $m\bar{\lambda}$. Now, suppose that for some graph $G$, $m\bar{\lambda} = 2(n-1)^2/n$. Then it is clear from what has been said above that $G$ has $n-1$ edges and, furthermore, that all distances are either 1 or 2. It is shown (Ref. 2) that a graph on $n$ vertices with $n-1$ edges has at least two "pendant" vertices; a pendant vertex is one which has only one edge incident to it. Let one of these vertices be denoted by $A$, and let the other endpoint of the edge incident to $A$ be denoted by $K$. Now let $B$ be any other vertex of the graph. Then $B$ must be connected to $K$ with an edge, since otherwise the distance from $A$ to $B$ would be greater than 2. But the choice of $B$ is arbitrary. Therefore all vertices of $G$ are joined to $K$, and since $G$ has only $n-1$ edges, all edges are of the form $V_iK$. We have therefore proved Theorem 3.

### C. Theorem 3

The bound of the Theorem 2 is attained if and only if

$$G = S_{n-1}$$

In view of this result, it might be expected that a similar result would be true for the bound of Theorem 1: if $m = n - 1 + \nu$, $\nu > 0$, then the bound of Theorem 1 is attained only for graphs $G$, which have been obtained by adding $\nu$ arbitrary edges to $S_{n-1}$. But although the bound of theorem is attained for all such graphs, the bound can also be attained by others, as illustrated in Figs. 2 and 3 for the case $n = 5$, $\nu = 1$. Clearly, the graph in Fig. 3 cannot be transformed into $S_4$ by the removal of an edge.

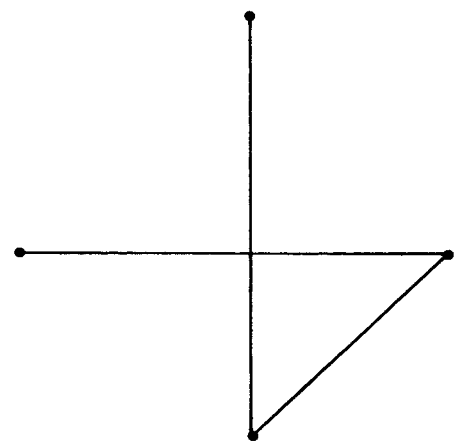

**Figure 2. A "star-like" graph attaining the bound of Theorem 1**

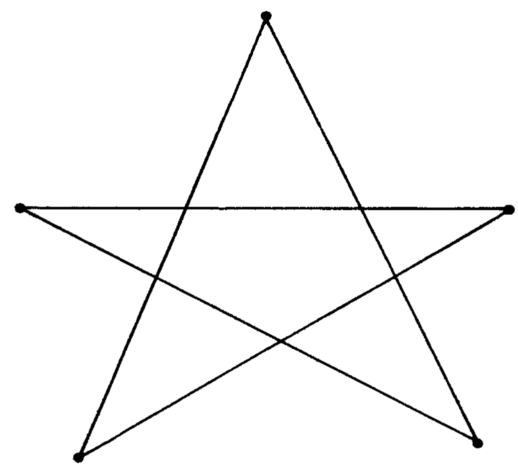

**Figure 3. A graph attaining the bound of Theorem 1 which is not "star-like"**

But there is a "partial uniqueness theorem" analogous to Theorem 3, which is presented in the next theorem.

### D. Theorem 4

If for a graph $G$ we write $\nu = m - n + 1$, and if $n > 2\nu + 3$, then the bound of Theorem 1 is attained only for graphs $G$ which have been obtained from $S_{n-1}$ by the addition of $\nu$(arbitrary) edges. (The number $\nu = \nu(G)$ is sometimes called the *cyclomatic number* of $G$.)

*Proof:* From the proof of Theorem 1, it is easy to see that the bound is attained only if all distances in the graph are either 1 or 2.

Following the notation of Berge (Ref. 2), let $|\Gamma A|$ represent the number of edges incident to the vertex $A$. Then

$$\sum_{A \varepsilon G} |\Gamma A| = 2m = 2(n - 1 + \nu)$$

since an edge incident to $A_1$ and $A_2$ is counted in $|\Gamma A_1|$ and $|\Gamma A_2|$. Consequently, the *average* number of edges incident to a vertex is $2 + (2(\nu - 1)/n)$. When $2(\nu - 1/n < 1$; i.e., when $n > 2\nu - 2$, the average is less than 3, and so there is at least one vertex $A$ such that $|\Gamma A| \leqq 2$. If $|\Gamma A| = 1$, $A$ is a pendant vertex, and we may then conclude that $G$ is of the required type by a modification of the proof given for Theorem 3.

If, now, $|\Gamma A| = 2$, let $B$ and $C$ be the (distinct) vertices of $G$ that are joined to $A$ by an edge. If $K$ is any other vertex of the graph (which is of course not joined to $A$ by an edge), then $K$ must be joined to either $B$ or $C$ by an edge (or both), since otherwise the distance from $A$ to $K$ would be greater than 2. This shows

$$|\Gamma A| + |\Gamma B| + |\Gamma C| \geqq n + 1$$

and so the average number of vertices incident to the remaining (we assume $n > 3$) vertices is $(1 + 2\nu)/(n - 3)$. When $(1 + 2\nu)/(n - 3)$ is less than 2, that is, when $n > 2\nu + 3$, there must be at least one vertex $K$ such that $|\Gamma K| = 1$, which implies that $G$ is of the required type, thus proving Theorem 4. Note that if $\nu = 0$, Theorem 4 tells us that for $n > 3$, $S_{n-1}$ is the only graph which attains the minimum value of Theorem 2. The cases $n = 1, 2, 3$ are easily disposed of, so that Theorem 4 gives an alternate proof of Theorem 3. With $\nu = 1$, Theorem 4 also shows that no graph "larger" than that of Fig. 3 can attain the bound of Theorem 2 unless it has a star subgraph.

## III. ARBITRARY WEIGHTS

We now proceed to the more general case, where in Eq. 1 we assign a weight $w_{ij} > 0$ to each path:

$$m\bar{\lambda} = \sum w_{ij}\mu_{ij}$$

Here $\sum w_{ij} = 1$ is normalized (summations are taken over all unordered pairs $(i, j)$ where $i \neq j$), and $\mu_{ij}$ denotes the distance between the $i^{\text{th}}$ and $j^{\text{th}}$ vertices. At this point, the problem becomes a more realistic one; the $w_{ij}$ may be considered as measures of the traffic between terminals, and, in general, the traffic is not the same between each pair of terminals. In our previous discussion, of course, we set $w_{ij} = 2/n(n - 1)$ for all $i,j$. To facilitate the discussion which is to follow, let us now remember the $w_{ij}$'s (and the $\mu_{ij}$'s correspondingly) with a single subscript, so that $w_1 \geqq w_2 \geqq \cdots \geqq w_{N-1} \geqq w_N$ (here $N = n(n - 1)/2$). This renumbering may generally be accomplished in several ways.

### A. Theorem 5

If the connected graph $G$ has $n$ vertices and $m$ edges, then

$$m\bar{\lambda} \geqq m\left(1 + \sum_{k=m+1}^{N} w_k\right)$$

This bound is a minimum for $m = n - 1$. Consequently, for a given set of weights $\{w_i\}$

$$m\bar{\lambda} \geqq (n - 1)\left(1 + \sum_{k=n}^{N} w_k\right)$$

*Proof:* As in the proof of Theorem 1, notice that in $G$ there are exactly $m$ $\mu$'s equal to 1, and so the $N - \lambda$ remaining $\mu$'s must be 2 or greater. To minimize $\bar{\lambda}$, we can do no better than to have the $m$ greatest $w$'s correspond to the $\mu$'s which are 1, and the $N - m$ remaining $w$'s correspond to the $\mu$'s which are 2. Thus

$$m\bar{\lambda} \geqq m\left(\sum_{k=1}^{N} w_k + 2\sum_{k=m+1}^{N} w_k\right) = m\left(1 + \sum_{k=m+1}^{N} w_k\right)$$

Let us now attempt to find the minimum of the expression

$$F(m) = m\left(1 + \sum_{k=1}^{N} w_k\right)$$

with respect to $m$. We have

$$F(m-1) = (m-1)\left(1 + \sum_{k=m}^{N} w_k\right)$$

$$= m + m\sum_{k=m}^{N} w_k - 1 - \sum_{k=m}^{N} w_k$$

$$F(m-1) = m\left(1 + \sum_{k=m+1}^{N} w_k\right) + mw_m - 1 - \sum_{k=m}^{N} w_k$$

$$= F(m) + \left(mw_m - 1 - \sum_{k=m}^{N} w_k\right)$$

But since $w_1 + w_2 + \cdots + w_m + \cdots + w_N = 1$, $w_1 \geqq w_2 \geqq \cdots \geqq w_N \geqq 0$, we have $mw_m \leqq 1$, and so

$$mw_m = 1 - \sum_{k=m}^{N} w_k \leqq 0$$

Actually, this is a strict inequality, since if $mw_m = 1$, $w_m > 0$ and so

$$\sum_{k=m}^{N} w_k > 0$$

Consequently, $F(m)$ is a *decreasing* function of $m$, thus minimized when $m$ is as small as possible. But we have seen that if $G$ is connected, $m \geqq n-1$. Theorem 5 is proved.

Now define $e_s(G)$ to be the minimum of $\overline{m\lambda}$ taken over *all* connected graphs $G$ with respect to a given set $S = \{w_{ij}\}$ of weights. Although for $w_{ij} = 2/n(n-1)$ we have seen that the bound of Theorem 5 is always attained, this is not the case generally. In fact, it is easy to see that the second ($m$-independent) bound of Theorem 5 is only attained *by a star* if the $n-1$ pairs $(i_1, j_1), \cdots, (i_{n-1}, j_{n-1})$ corresponding to the $n-1$ largest weights all share a common coordinate. Here, of course, if this common point is made the center of a star $S_{n-1}$, the bound is attained.

For example, (Fig. 4) if we assign the weight $AB = CD = 0.4$, $AC = BD = 0.1$, $AD = BC = 0$, the bound of Theorem 3 gives $e_s(G) \geqq 3.3$, but it is relatively easy to see that the best possible configuration gives $e_s(G) = 4.5$ (choose the star formed by $AB$, $AC$, $AD$).

Next, it might be tempting to conjecture that although the bound of Theorem 5 is not always attained, the best possible graph is always a star; but this is not true: for let

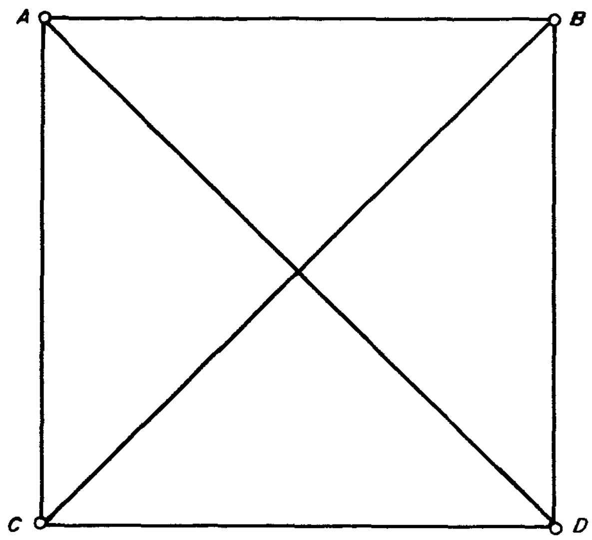


**Figure 4. Graph of order 4**

$AB = AC = BD = CD = 0.1$, $AD = BC = 0.3$ in Fig. 4. Here the bound of Theorem 5 is 3.9; all stars have the same value of $\overline{m\lambda} = 4.5$, but the graph formed by the edges $AD$, $AB$, and $BC$ has $\overline{m\lambda} = 4.2$.

Finally, we might hope that the best configuration is always attained by a graph with $n-1$ edges, but even this is not the case. Let $AB = BC = CD = AD = 0.23$, $AC = BD = 0.04$ in Fig. 4. Here the graph formed by $AB$, $BC$, $DC$, $AD$ has $\overline{m\lambda} = 4.32$, but the best graph with 3 edges has $\overline{m\lambda} = 4.5$ (any star has $\overline{m\lambda} = 4.5$).

In the above examples ($n = 4$), it was shown that the minimum of $\overline{m\lambda}$ for the *best star* was always 4.5, even though the weights assigned to the various pairs of vertices were different in the three cases considered. Further, 4.5 is the bound of *Theorem 2* for $n = 4$. This behavior can be explained by the following result.

## B. Theorem 6

For a given set of weights $S = \{w_{ij}\}$ on a set of $n$ vertices, let $e_s^*(G)$ represent the minimum value of $\overline{m\lambda}$ attained by any star $S_{n-1}$. Then it is always the case that $e^*(G) \leqq 2(n-1)^2/n$. *Remark:* Compare this result to that of Theorem 2; it states that any deviation from a flat distribution of weights can only result in a *decrease* in the minimum value of $\overline{m\lambda}$.

*Proof:* Denote the $n$ vertices by $V_1, V_2, \cdots, V_n$, and with each vertex $V_i$ associate a positive number $u_i$ as follows:

$$u_i = \sum_j w_{ij}$$

If we form the star with $V_i$ as center, it has

$$m\bar{\lambda} = (n-1)\,[u_i + 2(1-u_i)] = (n-1)(2-u_i)$$

and so the best star corresponds to the vertex $V_i$ for which $u_i$ is a minimum.

But

$$\sum_{i=1}^{n} u_i = 2$$

since in this sum each $w_{ij}$ occurs exactly twice; that is, once in $u_i$ and once in $u_j$, and $\Sigma w_{ij} = 1$. Hence

$$u_{max} = \max_i (u_i) \geqq \frac{2}{n}$$

and

$$e_s^*(G) \leqq (n-1)\left(2 - \frac{2}{n}\right) = \frac{2(n-1)}{n}$$

Comparing the results of Theorems 5 and 6, we see that the best $S_{n-1}$ has $m\bar{\lambda} < 2(n-1)$, while the best possible $m\bar{\lambda} \geqq n-1$. Therefore, although a star may not always be optimal, the best star is never worse than a factor of 2 from optimal. An extreme case of the relationship of $e_s^*(G)$ and $e_s(G)$ is given by the following example.

Let the vertices of $G$ be denoted by $V_1, V_2, \cdots, V_n$, and let $w_{ij} = 1/(n-1)$ if $j = i+1$, and 0 otherwise. Here the bound of Theorem 3 is attained by the chain $L_n$ from $V_1$ to $V_n$, while the best star has

$$m\bar{\lambda} = (n-1)\left(2 - \frac{2}{n-1}\right)$$

here the ratio $e^*(G)/e(G) = 2 - [2/(n-1)] \rightarrow 2$ as $n \rightarrow \infty$, and it was seen that 2 is the largest ratio possible.

But in a statistical sense, this example is pessimistic, as the following Section will help to show.

## IV. PROBABILISTIC WEIGHTS

Consider an elementary lemma on limiting distributions:

*Lemma:* Let $F$ be a distribution function with $F(x) = 0$ for $x \leqq 0$. Suppose $F$ has (finite) mean $\mu$ and (finite) variance $\sigma^2$. Let $X_1, X_2, \cdots, X_n$ be $n$ independent random variables with identical distribution functions $F$, and let $X_{(n)}$ represent the largest of $X_1, X_2, \cdots, X_n$. Let $Y_n$ be the random variable defined by

$$Y_n = \frac{X_{(n)}}{X_1 + X_2 + \cdots + X_n}$$

Then for every $\varepsilon > 0$,

$$Pr\{n^{1/2} Y_n > \varepsilon\} \rightarrow 0 \text{ as } n \rightarrow \infty$$

*Proof:* To find the distribution function of $X_{(n)}$, note that

$$Pr\{X_{(n)} < x\} = \prod_{i=1}^{n} Pr\{X_i < x\} = F^n(x)$$

since the random variables $X_i$ are independent.

Write

$$S_n = X_1 + X_2 + \cdots + X_n$$

and

$$G_n(y) = Pr\{n^{1/2} X_{(n)} \leqq y S_n\}$$

Then

$$1 \geqq G_n(y) \geqq Pr\left\{X_{(n)} \leqq \frac{y}{n^{1/2}} S_n \,\middle|\, \left|\frac{S_n}{n} - \mu\right| < \varepsilon\right\} \times Pr\left\{\left|\frac{S_n}{n} - \mu\right| < \varepsilon\right\}$$

By the Chebyschev inequality

$$1 \geqq G_n(y) \geqq Pr\left\{X_{(n)} \leqq \frac{y}{n^{1/2}}(n\mu - n\varepsilon)\right\}\left(1 - \frac{\sigma^2}{n\varepsilon^2}\right)$$

and by substitution

$$1 \geqq G_n(y) \geqq F^n[y n^{1/2}(\mu - \varepsilon)]\left(1 - \frac{\sigma^2}{n\varepsilon^2}\right) \tag{2}$$

for all $\varepsilon > 0$.

We now compute

$$\lim_{n\to\infty} F^n [yn^{1/2}(\mu - \varepsilon)]$$

$F$ has finite variance, so

$$\int_0^\infty x^2 f(x)\,dx = \mu_2 < \infty$$

where $F' = f$. Consequently,

$$\lim_{y\to\infty} \int_y^\infty x^2 f(x)\,dx = 0$$

which means that

$$\lim_{y\to\infty} y^2 \int_y^\infty f(x)\,dx = 0$$

But now

$$\begin{aligned} F^n [yn^{1/2}(\mu - \varepsilon)] &= \left(\int_0^{yn^{1/2}(\mu-\varepsilon)} f(x)\,dx\right)^n \\ &= \left(1 - \int_{yn^{1/2}(\mu-\varepsilon)}^\infty f(x)\,dx\right)^n \\ &= \left(1 - \frac{A(n)}{n}\right)^n \end{aligned}$$

where

$$A(n) = n \int_{yn^{1/2}(\mu-\varepsilon)}^\infty f(x)\,dx$$

But from the fact that

$$\lim_{y\to\infty} y^2 \int_y^\infty f(x)\,dx = 0$$

we see that

$$\lim_{n\to\infty} y^2 n(\mu - \varepsilon)^2 \int_{yn^{1/2}(\mu-\varepsilon)}^\infty f(x)\,dx = 0$$

and so

$$\lim_{n\to\infty} A(n) = 0$$

as well. Hence

$$\lim_{n\to\infty} F^n [yn^{1/2}(\mu - \varepsilon)] = \left(1 - \frac{A(n)}{n}\right)^n = 1$$

from elementary limit considerations. So from Eq. 2, we see that

$$\lim_{n\to\infty} G_n(y) = 1$$

But

$$Pr\{n^{1/2} Y_n > \varepsilon\} = 1 - G_n(\varepsilon)$$

and so

$$\lim_{n\to\infty} Pr\{n^{1/2} Y_n > \varepsilon\} = 0$$

This proves the lemma.

Now in Theorem 5, the bound may be rewritten as follows:

$$e_s(G) \geqq (n-1)\left(2 - \sum_{k=1}^{n-1} w_k\right)$$

Hence

$$e_s(G) \geqq (n-1)[2 - (n-1)w_1]$$

and so

$$\frac{e_s^*(G)}{e_s(G)} \leqq \frac{2(n-1)}{n[2-(n-1)w_1]} \leqq \frac{2}{2-(n-1)w_1}$$

If now, for example, the weights are considered to be $n(n-1)/2$ random samples from a distribution function with finite mean and variance (normalized so that their sum is 1), then an easy conclusion of the lemma shows that

$$Pr\{(n-1)w_1 > \varepsilon\} \to 0$$

Consequently,

$$Pr\{e^*(G)/e(G) > 1 + \varepsilon\} \to 0 \text{ as } n \to \infty$$

for every $\varepsilon > 0$. In a meaningful sense, therefore, the star "asymptotically minimizes" $m\bar{\lambda}$.

*Remark:* The lemma can be modified to show the stated result when the traffic between two terminals is assumed to be proportional to the sum or product of the "sizes" of the vertices, where now the "sizes" are assumed to be distributed according to some distribution with finite mean and variance.

## REFERENCES

1. Kleinrock, L., "Communication Nets," *Stochastic Message Flow and Relay,* McGraw-Hill Book Co., New York, 1964.

2. Berge, C., *The Theory of Graphs,* Methuen & Co., London, 1962.