July 1983 LIDS-P-1307

ASYMPTOTIC OPTIMALITY OF SHORTEST PATH ROUTING ALGORITHMS*

by

Eli M. Gafni and Dimitri P. Bertsekas**

ABSTRACT

Many communication networks use adaptive shortest path routing. By this we mean that each network link is periodically assigned a length that depends on its congestion level during the preceding period, and all traffic generated between length updates is routed along a shortest path corresponding to the latest link lengths. We show that in certain situations, typical of networks involving a large number of small users and utilizing virtual circuits, this routing method performs optimally in an asymptotic sense. In other cases shortest path routing can be far from optimal.

---

*This work was supported by DARPA under contract ONR/N00014-75-C-1183. Thanks are due to Adrian Segall who suggested the problem to the first author.

**E. M. Gafni is with UCLA, Computer Science Dept., Los Angeles, Calif. 90024. D. P. Bertsekas is with Massachusetts Institute of Technology, Laboratory for Information and Decision Systems, Cambridge, Mass.02139

I. Introduction

Most of the presently existing communication networks utilize shortest path routing as evidenced by the recent survey paper [1]. This routing method has gained popularity primarily because it is simple and handles adequately link and node failures. Relatively little is known however about the performance of shortest path routing under heavy traffic conditions since most of the practical experience reported to date relates to networks that are typically lightly loaded, e.g. the ARPANET [2].

It is customary to measure optimality of a routing scheme in terms of an objective function of the form

$$\sum_{(i,j)} D_{ij}(F_{ij}) \tag{1}$$

where F_{ij} denotes the arrival rate at the transmission queue of link (i,j). Here D_{ij} is a convex monotonically increasing function such as for example

$$D_{ij}(F_{ij}) = \frac{F_{ij}}{C_{ij}-F_{ij}}, \quad C_{ij}\text{:capacity of } (i,j) \tag{2}$$

which corresponds to the Kleinrock independence assumption [3]. There is extensive literature on the problem of minimizing (1) subject to known offered traffic for each origin-destination pair [4]-[12]. It makes sense to evaluate routing performance in terms of an objective function such as (1), (2) in circumstances where the offered traffic statistics change slowly over time and furthermore individual offered traffic sample functions do not exhibit frequently large and persistent deviations from their averages. A typical situation is a network accomodating a large number of relatively small users for each origin-destination pair in which a form of the law of large numbers approximately takes hold (see Lemma A.1). This paper considers exclusively this type of network and its conclusions do not apply at all to more dynamic situations characterized by the presence of a few large users that can by themselves overload the network over brief periods of time if left uncontrolled. For such cases an objective function such as (1) is not appropriate and different methods of analysis are called for (see e.g. [14], [15]).

The purpose of the paper is to evaluate the performance of shortest path routing in terms of the objective function (1) when the length of each link (i,j) is periodically calculated as $D'_{ij}(F_{ij})$--the first derivative of D_{ij} evaluated at the average rate F_{ij} at queue (i,j) during the preceding period. The first derivative relation between link lengths and objective function is motivated by the well known optimality condition that a routing optimizes the objective (1) if and only if it routes traffic exclusively along paths of minimum first derivative length (see e.g. [4], [13]). It is known that this type of shortest path routing is strictly suboptimal although it is believed to be close to optimal for lightly loaded networks. Furthermore for datagram networks shortest path routing is prone to oscillations which can be severe if the length functions D'_{ij} are chosen poorly [17], [18]. Indeed the original adaptive shortest path algorithm implemented in 1969 on the ARPANET exhibited violent oscillatory behavior which was restrained only after using the device of adding a bias to each link length at the expense of considerable loss of adaptivity ([16], [19], [20]).

A key feature of a datagram network is that each packet of a user pair is not required to travel on the same path as the preceding packet.

Therefore the "holding time of each communication path" (the maximum time that a user pair will continue to use the path after it is changed due to a shortest path update) is one packet long. As a result a datagram network reacts very fast to a shortest path update with all traffic switching to the new shortest paths almost instantaneously.

The situation is quite different in a virtual circuit network where every conversation is assigned a fixed communication path at the time it is first established. There the "holding time of the communication path" (as loosely described above) is often large relative to the shortest path updating period. As a result the network reaction to a shortest path update is much more gradual since old conversations continue to use their established communication paths and only new conversations are assigned to the most recently calculated shortest paths.

The main result of this paper is that the performance of shortest path routing approaches the optimal achievable by <u>any</u> other method if

$$\frac{\text{Shortest Path Updating Period}}{\text{Average Holding Time of the Communication Path}} \to 0 \tag{3}$$

and

$$n_w \to \infty, \qquad \gamma_w \to 0, \qquad n_w\gamma_w = \text{constant} \tag{4}$$

where n_w is the average number of active conversations for the generic origin-destination pair w, and γ_w is the communication rate of each conversation. Assumptions (3), (4) together with additional Poisson-like assumptions on the offered traffic statistics are formulated in the next section. The main result in Section 3 provides also bounds on the suboptimality of the shortest path method when the assumptions (3) and (4) are satisfied only approximately. Roughly speaking the theorem states that the average value of the cost (1) of the shortest path method converges to a neighborhood of the optimal cost at a natural rate which is independent of how fast the shortest paths are updated. However the size of the neighborhood is "proportional" to the extent of violation of assumptions (3) and (4).

2. Problem Formulation

Consider a network with a set of nodes N and a set of directed links L. We are given a set W of ordered node pairs referred to as origin-destination (OD) pairs. For each OD pair $w \in W$ we are given a nonempty set of directed paths P_w joining the origin node and the destination node of w. Conversations for each $w \in W$ arrive according to a Poisson process with mean rate $\frac{\bar{\lambda}_w}{\varepsilon}$ where $\bar{\lambda}_w$ is given and ε is a positive parameter the effect of which we wish to study. Each conversation for OD pair w is assigned upon arrival to a path $p \in P_w$ according to a rule to be described shortly and uses this path for the entire time of its duration assumed to be exponentially distributed with mean $\frac{1}{\mu_w}$. We assume that the Poisson arrival processes and duration times of conversations are independent, and each path can carry unlimited conversations, so the number of active conversations for each OD pair evolves as in an M/M/∞ queueing system. It follows ([21], p. 101) that if $n_w(t)$ is the number of active conversations for w at time t then its mean and variance satisfy

$$\lim_{t\to\infty} E\{n_w(t)\} = \frac{\bar{\lambda}_w}{\varepsilon\mu_w}, \quad \lim_{t\to\infty} \operatorname{var}\{n_w(t)\} = \frac{\bar{\lambda}_w}{\varepsilon\mu_w}. \tag{5}$$

Path assignment for each conversation is determined according to the following shortest path rule:

At times $t = kT$ secs, $k = 0, 1, \ldots,$ where $T > 0$ is given, the length of each link (i,j) is calculated as $d_{ij}[F_{ij}(t)]$ where $F_{ij}(t)$ is the communication rate on link (i,j) given by

$$F_{ij}(t) = \sum_{w\in W} \gamma_w \sum_{\substack{p\in P_w \\ (i,j)\in p}} n_p(t). \tag{6}$$

Here $n_p(t)$ is the number of active conversations assigned on path p at time t, $\sum_{\substack{p\varepsilon P_w \\ (i,j)\varepsilon p}} n_p(t)$ is the total number of conversations of OD pair w using (i,j) at time t, and γ_w is the communication rate per conversation of OD pair w. All conversations of OD pair w arriving at times $t\varepsilon[kT,(k+1)T)$ are assigned on a path $p\varepsilon P_w$ which is shortest relative to the link lengths $d_{ij}[F_{ij}(kT)]$. (Ties between paths are assumed resolved according to a fixed deterministic rule).

We assume that $d_{ij}(\cdot)$ is a continuous strictly monotonically increasing function of F_{ij} satisfying $d_{ij}(F_{ij}) \geq 0$ for all $F_{ij} \geq 0$ and

$$|d_{ij}(\overline{F})-d_{ij}(F)| \leq L|\overline{F}-F| , \ \forall \ \overline{F},F \geq 0, \ (i,j)\varepsilon L, \tag{7}$$

where L is a given positive constant. This assumption is reasonable once the length function d_{ij} is assumed continuous. In practice the length function is sometimes taken discontinuous (e.g. the TYMNET [1]). We do not know whether and in what form our main result holds for this case. Note that the assumption (7) is not satisfied when d_{ij} is the first derivative of the function D_{ij} of (2) since this derivative increases without bound as F_{ij} approaches the capacity C_{ij}. As a practical matter this is not a problem since flow control will ordinarily not allow a link flow to get too close to capacity.

Regarding the communication rate γ_w we assume that it is of the form

$$\gamma_w = \varepsilon \overline{\gamma}_w \tag{8}$$

where $\overline{\gamma}_w$ is some constant. Thus we assume in effect that, even though the real communication rate of a conversation will be a random process, the rates γ_w used in the calculation of flows in (6) are obtained by averaging the real rates over a long period of time and over all conversations of OD pair w so that the variance of γ_w is so small that γ_w can be viewed as a deterministic quantity. Note that for each OD pair w the product

$$\text{(Mean arrival rate)} \cdot \text{(Communication rate)} = \bar{\lambda}_w \bar{\gamma}_w$$

is independent of ε. We wish to study the effect on various stochastic processes of interest of the parameters ε and T particularly as

$$\varepsilon \to 0 \quad \text{and} \quad T \to 0.$$

Taking $\varepsilon \to 0$ implies that arrival rates tend to infinity while communication rates tend to zero with the products staying constant, and approximates a situation where there are many small conversations in the network [cf.(4)]. Taking $T \to 0$ approximates a situation where updating of shortest paths is fast relative to the mean duration time of a conversation [cf.(3)].

The initial numbers $n_p(0)$ of active conversations on each path p are assumed given. These numbers together with the earlier assumptions on the arrival processes, holding times, and the routing method completely characterize the statistics of all processes of subsequent interest. Our main result can be proved in essentially the same form if $\{n_p(0)\}$ are random with given mean and variance (see Lemma A.1).

We will investigate the behavior of the processes $F(t) = \{F_{ij}(t) \mid (i,j) \varepsilon L\}$ and

$$D[F(t)] = \sum_{(i,j)\varepsilon L} D_{ij}[F_{ij}(t)]$$

where D_{ij} is some function such that

$$d_{ij}(F_{ij}) \quad = \quad D'_{ij}(F_{ij}) \overset{\Delta}{=} \text{First derivative of } D_{ij} \text{ at } F_{ij}. \qquad (9)$$

Note that, in view of our earlier assumptions, $d_{ij}(\cdot)$ uniquely defines $D_{ij}(\cdot)$ as a strictly convex, monotonically increasing function up to an additive constant.

There is a lower bound to the value of $E\{D[F(t)]\}$ achievable in the long run by <u>any</u> rule for assigning conversations to paths. This is

$$D^* \quad = \quad \min_{F \in \mathcal{F}} D(F) \qquad (10)$$

where $\mathcal{F}$ is the set of all total flows $F = \{F_{ij} | (i,j) \in L\}$ of the form

$$F_{ij} \quad = \quad \sum_{w \in W} \sum_{\substack{p \in P_w \\ (i,j) \in p}} x_p \quad , \qquad \forall\ (i,j) \in L \qquad (11)$$

where x_p are any nonnegative scalars satisfying

$$\sum_{p \in P_w} x_p \quad = \quad \frac{\overline{\lambda}_w \overline{\gamma}_w}{\mu_w} \quad , \quad \forall\ w \in W. \qquad (12)$$

In other words $\mathcal{F}$ is the set of all possible average total link rates resulting from the long term average input traffic rate $\frac{\overline{\lambda}_w \overline{\gamma}_w}{\mu_w}$ at each OD pair w (cf. (5), (8)). Note that the problem in (10) is the usual deterministic multicommodity flow problem that has been studied extensively in connection with optimal routing [4]-[13]. For any routing rule the inequality

$$D^* \quad \leq \quad \liminf_{t \to \infty} \quad E\{D[F(t)]\}$$

follows from the fact

$$D[E\{F(t)\}] \leq E\{D[F(t)]\}, \quad \forall\, t \geq 0$$

which holds by the convexity of D, Jensen's inequality, and the fact [cf. (5), (8)]

$$E\{n_w(t)\gamma_w\} \to \frac{\bar{\lambda}_w \bar{\gamma}_w}{\mu_w} \qquad \text{as } t \to \infty$$

Our main result is that as $\varepsilon \to 0$, $T \to 0$ and $t \to \infty$ the expected cost $E\{D[F(t)]\}$ corresponding to the shortest path ruie converges to the lower bound D^* while $F(t)$ converges in mean square to the unique F^* that achieves the minimum in the deterministic optimal routing problem (10).

3. Main Result

We first introduce some notation:

$x_p(t) \overset{\Delta}{=} \varepsilon\bar{\gamma}_w n_p(t)$: The communication rate on path p at time t.

$r_w(t) \overset{\Delta}{=} \sum_{p\in P_w} x_p(t)$: The total input rate of OD pair w at t.

$\bar{r}_w \overset{\Delta}{=} \frac{\bar{\lambda}_w\bar{\gamma}_w}{\mu_w}$: The long term average input rate of w.

$\bar{r} \overset{\Delta}{=} \max_w \{\bar{r}_w\}$

$R_w \overset{\Delta}{=} |r_w(0) - \bar{r}_w|$: The initial deviation of r_w from its long term average

$R \overset{\Delta}{=} \max_w \{R_w\}$

$\mu \overset{\Delta}{=} \min_w \{\mu_w\}$

$M \overset{\Delta}{=} \max_w \{\mu_w\}$

$\bar{\gamma} = \max_w \{\bar{\gamma}_w\}$

Theorem: There exist positive constants c_1, c_2 (which depend only on the network topology, the products $\bar{\lambda}_w \bar{\gamma}_w$, and the length functions d_{ij}) such that the total link rate vector F(t) corresponding to shortest path routing satisfies for all t = kT, k = 0,1,...

$$-c_1 Re^{-\mu t} \leq E\{D[F(t)]\}-D^* \leq e^{-\mu t}[D[F(0)]-D^*] + c_2[a(\varepsilon,T) + b(\varepsilon,T)te^{-\mu t}] \tag{13}$$

where

$$a(\varepsilon,T) = \bar{r}\{\frac{(\varepsilon\bar{\gamma}\bar{r} + \sqrt{\varepsilon\bar{\gamma}(\bar{r}+R)})(e^{-\mu T}-e^{-MT})}{\bar{r}(1-e^{-\mu T})} + 2\varepsilon\bar{\gamma} + (1-e^{-\mu T})(4\bar{r}+\varepsilon\bar{\gamma})\} \tag{14}$$

$$b(\varepsilon,T) = R\{ \frac{(\varepsilon\overline{\gamma}+R+1)(e^{-\mu T}-e^{-MT}) + (1-e^{-\mu T})[\varepsilon\overline{\gamma} + (1-e^{-\mu T})(4\overline{r}+R+\varepsilon\overline{\gamma})]}{Te^{-\mu T}} \tag{15}$$

Furthermore

$$\lim_{\substack{\varepsilon\to 0\\ T\to 0}} (\limsup_{t\to\infty}\{E\{D[F(t)]\}) = D^* .$$

If in addition we assume that, for some $\ell > 0$, the length functions d_{ij} satisfy

$$\ell|\overline{F}-F| \leq |d_{ij}(\overline{F}) - d_{ij}(F)| , \quad \forall\ \overline{F},F \geq 0,\ (i,j)\varepsilon L$$

then

$$\lim_{\substack{\varepsilon\to 0\\ T\to 0\\ t\to\infty}} E\{|F_{ij}(t)-F^*_{ij}|^2\} = 0, \quad \forall\ (i,j)\varepsilon L,$$

where F^* is the unique solution of the deterministic optimal routing problem (10).

The proof of the theorem is given in the appendix. The idea of the proof is based on relations of shortest path routing with the flow deviation (or Frank-Wolfe) method [7] for solving problem (10) (see [13]). However the proof here is complicated by the fact that we are dealing with a stochastic optimization problem while the flow deviation method deals with a deterministic problem. A simpler version of the theorem that assumes that ε and T are so small that the path rates can be obtained as solutions of differential equations is given in [22].

The main implication of (13) is that, as $t \to \infty$, $E\{D[F(t)]\}$ comes within $c_2\, a(\varepsilon,T)$ of being optimal. Thus $c_2\, a(\varepsilon,T)$ may be viewed as the long-term deviation from optimality of shortest path routing. The key fact is that $a(\varepsilon,T) \to 0$ as $\varepsilon \to 0$ and $T \to 0$. The rate at which $E\{D[F(t)]\}$ approaches its long term limit depends on the largest average holding time $\frac{1}{\mu}$. There are three terms here. The first term $e^{-\mu t}[D[F(0)]-D^*]$ is proportional to the initial deviation from optimality. The other two terms are proportional to the initial deviation R of the initial OD pair rates $r_w(0)$ from their long-term averages $\overline{r}_w$.

The three transient terms in (13) characterize the rate of convergence of the algorithm. Of these terms the slowest is the one involving $t\, e^{-\mu t}$. Since for any $\delta > 0$ we have $t\, e^{-\mu t} \leq \frac{1}{\delta e}\; e^{-(\mu-\delta)t}$ we see that even this term decays "almost" as fast as $e^{-\mu t}$. Thus we can conclude that at worst, $E\{D[F(t)]\}$ converges to its long term average "almost" like $e^{-\mu t}$--a linear rate which is independent of ε and T. For specific problems the actual rate of convergence can be considerably faster and the bound $e^{-\mu t}$ is not necessarily tight. However $E\{D[F(t)]\}$ cannot converge to D* much faster than $e^{-\mu t}$ since we know that the rate of change of F(t) is constrained by the rate at which the number of old conversations on any path can decrease due to termination and this rate is precisely $e^{-\mu t}$. Thus for example if $D_{ij}(F_{ij})$ is quadratic in F_{ij} the rate of convergence of $E\{D[F(t)]\}$ cannot be faster than $e^{-2\mu t}$, while in the extreme case where $D_{ij}(F_{ij})$ is linear in F_{ij} the rate of convergence cannot be faster than $e^{-\mu t}$. Therefore there is little margin for improvement of our rate of convergence result. The conclusion is that the largest average duration $1/\mu$ of a conversation is a fundamental limiting factor in the performance of the shortest path algorithm. When $1/\mu$ is large the algorithm tends to converge slowly to a neighborhood of the optimum. This is a manifestation of the intuitively clear fact that the routing algorithm cannot perform well if poorly routed conversations last for a long time.

Appendix: Proof of the Theorem

For brevity we use the following notation in addition to the one given in the beginning of Section 3:

$$n_p^k \triangleq n_p(kT),\quad x_p^k \triangleq x_p(kT),\quad r_w^k \triangleq r_w(kT),\quad F_{ij}^k \triangleq F_{ij}(kT)$$

$$x^k \triangleq \{x_p^k | p\varepsilon P_w,\ w\varepsilon W\},\quad F^k \triangleq \{F_{ij}^k | (i,j)\varepsilon L\}.$$

We first prove some helpful lemmas. The first lemma gives some basic facts about the transient behavior of various processes of interest. In particular it shows that as $\varepsilon \to 0$ the processes $x_p(t)$ and $r_w(t)$ behave asymptotically as deterministic processes.

Lemma 1: For all $t \geq 0$ and $w\varepsilon W$

$$E\{r_w(t)\} = \bar{r}_w + e^{-\mu_w t}[r_w(0)-\bar{r}_w] \tag{A1}$$

$$\operatorname{var}\{r_w(t)\} = \varepsilon\bar{\gamma}_w(1-e^{-\mu_w t})[\bar{r}_w + e^{-\mu_w t} r_w(0)]. \tag{A2}$$

Furthermore, for each $w\varepsilon W$, if $p_k\varepsilon P_w$ is the shortest path used for routing in the interval $[kT, (k+1)T)$ we have for all $t\varepsilon[kT, (k+1)T]$

$$E\{x_p(t)|x_p^k\} = \begin{cases} e^{-\mu_w(t-kT)} x_p^k & \text{if } p \neq p_k \\ \bar{r}_w + e^{-\mu_w(t-kT)}(x_p^k-\bar{r}_w) & \text{if } p = p_k \end{cases} \tag{A3}$$

$$\mathrm{var}\{x_p(t)\,|\,x_p^k\} = \begin{cases} \varepsilon\bar{\gamma}_w[1-e^{-\mu_w(t-kT)}]e^{-\mu_w(t-kT)}x_p^k & \text{if } p \neq p_k \\ \varepsilon\bar{\gamma}_w[1-e^{-\mu_w(t-kT)}][\bar{r}_w + e^{-\mu_w(t-kT)}x_p^k] & \text{if } p = p_k. \end{cases} \tag{A4}$$

<u>Proof</u>: Consider an $M/M/\infty$ queueing system with arrival rate Λ and service rate $\frac{1}{M}$. The probabilities $P_k(t)$ of k customers in the system at time t satisfy the differential equations ([21], p. 59, 101)

$$\dot{P}_0 = -\Lambda P_0 + MP_1$$

$$\dot{P}_k = -(\Lambda + kM)P_k + \Lambda P_{k-1} + (k+1)MP_{k+1}, \quad k = 1,2,\ldots \tag{A5}$$

Let $N(t) = \sum_{k=1}^{\infty} kP_k(t)$ and $\sigma(t) = \sum_{k=0}^{\infty} [k-N(t)]^2 P_k(t)$ be the expected value and variance of the number in the system. Multiplying (A5) by k and adding we obtain by straightforward calculation the differential equation

$$\dot{N} = -MN + \Lambda. \tag{A6}$$

Also by multiplying (A5) by $(k-N)^2$, adding, and taking into account the fact $\dot{\sigma} = \sum_{k=0}^{\infty} (k-N)^2 \dot{P}_k$ we obtain the equation

$$\dot{\sigma} = -2M\sigma + MN + \Lambda. \tag{A7}$$

The solutions of the linear differential equations (A6), (A7) can be calculated by the variations of constants formula. They are

$$N(t) = \frac{\Lambda}{M} + e^{-Mt}[N(0) - \frac{\Lambda}{M}] \tag{A8}$$

$$\sigma(t) = e^{-2Mt}\,\sigma(0) + (1-e^{-Mt})[\frac{\Lambda}{M} + e^{-Mt}\,N(0)]. \tag{A9}$$

Applying (A8) for $M = \mu_w$, $\Lambda = \frac{\bar{\lambda}_w}{\varepsilon}$, and multiplying by $\varepsilon\bar{\gamma}_w$ yields (A1). Applying (A9) for $M = \mu_w$, $\Lambda = \frac{\bar{\lambda}_w}{\varepsilon}$, $\sigma(0) = 0$, and multiplying by $\varepsilon^2\bar{\gamma}_w^2$ yields (A2). A similar application of (A8) and (A9) yields (A3) and (A4). Q.E.D.

Note that from (A1), (A2) we obtain the useful relations

$$|E\{r_w(t)\} - \bar{r}_w| \leq e^{-\mu_w t} R_w \leq e^{-\mu t} R \tag{A10}$$

$$\begin{aligned} \mathrm{var}\{r_w(t)\} &\leq \varepsilon\bar{\gamma}(1-e^{-\mu_w t})[(1+e^{-\mu_w t})\bar{r}_w + e^{-\mu_w t}[r_w(0) - \bar{r}_w]] \\ &\leq \varepsilon\bar{\gamma}(\bar{r} + e^{-\mu t}R). \end{aligned} \tag{A11}$$

The proof of Theorem 1 would be considerably simplified if the average holding time of a conversation is independent of the OD pair, i.e. $\mu_w = \mu = M$ for all $w\varepsilon W$. In fact the reader may wish to go first through the proof assuming this. To cope with the case where $\mu \neq M$ we will need to introduce the following "normalized" processes

$$\tilde{x}_p(t) = \frac{x_p(t)\bar{r}_w}{r_w(t)}, \quad \forall\, w\varepsilon W,\ p\varepsilon P_w, \tag{A12a}$$

$$\tilde{F}_{ij}(t) = \sum_{w\varepsilon W} \sum_{\substack{p\varepsilon P_w \\ (i,j)\varepsilon p}} \tilde{x}_p(t), \quad \forall\,(i,j)\varepsilon L. \tag{A12b}$$

We denote

$$\tilde{x}_p^k \triangleq \tilde{x}_p(kT), \quad \tilde{F}_{ij}^k \triangleq \tilde{F}_{ij}(kT). \tag{A12c}$$

Using the fact $\tilde{x}_p(t) \leq \bar{r}_w$, and (A1), (A11) we have

$$E\{|\tilde{x}_p(t) - x_p(t)|^2\} = E\{|\tilde{x}_p(t)[1 - \frac{r_w(t)}{\bar{r}_w}]|^2\}$$

$$\leq E\{|\bar{r}_w - r_w(t)|^2\}$$

$$\leq E\{|E\{r_w(t)\} - e^{-\mu_w t}[r_w(0) - \bar{r}_w] - r_w(t)|^2\}$$

$$\leq \mathrm{var}\{r_w(t)\} + e^{-2\mu_w t}R_w^2$$

$$\leq \varepsilon\bar{\gamma}(\bar{r} + e^{-\mu t} R) + e^{-2\mu t} R^2.$$

Since $\tilde{F}_{ij}$ and F_{ij} are sums of $\tilde{x}_p$ and x_p respectively we obtain for some constant α_{ij}

$$E\{|\tilde{F}_{ij}(t) - F_{ij}(t)|^2\} \leq \alpha_{ij}[\varepsilon\bar{\gamma}(\bar{r} + e^{-\mu t}R) + e^{-2\mu t} R^2]. \tag{A13}$$

The next lemma provides a basic estimate:

<u>Lemma 2</u>: For every vector $\tilde{F}\varepsilon F$ and every other total link rate vector F (not necessarily in F) there holds

$$D(\tilde{F}) \leq D(F) + B \sum_{(i,j)} |\tilde{F}_{ij} - F_{ij}|, \tag{A14}$$

where B is an upperbound for $d_{ij}(\tilde{F}_{ij})$ over $(i,j)\varepsilon L$ and $\tilde{F}\varepsilon F$.

Proof: We have by the convexity of D

$$D(F) \geq D(\tilde{F}) + \sum_{(i,j)} d_{ij}(\tilde{F}_{ij})(F_{ij}-\tilde{F}_{ij})$$

$$\geq D(\tilde{F}) - B \sum_{(i,j)} |F_{ij}-\tilde{F}_{ij}|. \qquad \text{Q.E.D.}$$

Proof of Theorem 1:

We first show the left side of (13). Let $\{x_p^*(t)\}$ be a set of path rates that solve the deterministic multicommodity flow problem

$$\begin{aligned} &\text{minimize } D(F) \\ &\text{subject to } F_{ij} = \sum_{w \in W} \sum_{\substack{p \in P_w \\ (i,j) \in p}} x_p \\ &\qquad \sum_{p \in P_w} x_p = E\{r_w(t)\}, \quad \forall\, w \in W \\ &\qquad x_p \geq 0, \quad \forall\, p \in P_w,\ w \in W. \end{aligned} \tag{A15}$$

Let $F^*(t)$ be the vector of corresponding total link rates, i.e.

$$F^*_{ij}(t) = \sum_{w\varepsilon W} \sum_{\substack{p\varepsilon P_w \\ (i,j\varepsilon p}} x^*_p(t)$$

Define the "normalized" rates

$$\hat{x}_p(t) = \frac{\bar{r}_w}{E\{r_w(t)\}} x^*_p(t) \tag{A16}$$

$$\hat{F}_{ij}(t) = \sum_{w\varepsilon W} \sum_{\substack{p\varepsilon P_w \\ (i,j)\varepsilon p}} \hat{x}_p(t).$$

Since $\hat{F}(t) = \{\hat{F}_{ij}(t)\}\varepsilon F$ we have using (A14)

$$D^* \leq D[\hat{F}(t)] \leq D[F^*(t)] + B \sum_{(i,j)} |\hat{F}_{ij}(t) - F^*_{ij}(t)| \tag{A17}$$

$$\leq D[E\{F(t)\}] + B \sum_{(i,j)} |\hat{F}_{ij}(t) - F^*_{ij}(t)|$$

$$\leq E\{D[F(t)]\} + B \sum_{(i,j)} |\hat{F}_{ij}(t) - F^*_{ij}(t)|,$$

where the last step follows using Jensen's inequality.

From (A16) we have using the fact $\hat{x}_p(t) \leq \bar{r}_w$ and (A10)

$$|\hat{x}_p(t) - x^*_p(t)| = \left| \frac{\hat{x}_p(t)}{\bar{r}_w} [\bar{r}_w - E\{r_w(t)\}] \right| \leq R\, e^{-\mu t}.$$

Since $\hat{F}_{ij}(t)$ and $F^*_{ij}(t)$ consist of sums of $\hat{x}_p(t)$ and $x^*_p(t)$ respectively we have for some constants β_{ij}

$$|\hat{F}_{ij}(t) - F^*_{ij}(t)| \leq \beta_{ij} R\, e^{-\mu t} \tag{A18}$$

Taking $c_1 = B \sum_{(i,j)} \beta_{ij}$ we obtain from (A17) and (A18)

$$D^* \leq E\{D[F(t)]\} + c_i R e^{-\mu t}$$

and the left side of (13) is proved.

To prove the right side of (13) we first fix k and consider times $t\varepsilon[kT, (k+1)T]$. We have using (7) and Taylor's Theorem

$$\begin{aligned} D_{ij}[F_{ij}(t)] &= D_{ij}(F_{ij}^k) + d_{ij}(F_{ij}^k)[F_{ij}(t) - F_{ij}^k] \\ &\quad + \int_0^1 \{d_{ij}[F_{ij}^k + \alpha(F_{ij}(t) - F_{ij}^k)] - d_{ij}(F_{ij}^k)\}[F_{ij}(t)-F_{ij}^k]d\alpha \\ &\leq D_{ij}(F_{ij}^k) + d_{ij}(F_{ij}^k)[F_{ij}(t) - F_{ij}^k] + \frac{L}{2}|F_{ij}(t) - F_{ij}^k|^2 \end{aligned}$$

By summing over all links (i,j) we obtain

$$D[F(t)] \leq D(F^k) + \sum_{(i,j)} d_{ij}(F_{ij}^k)[F_{ij}(t) - F_{ij}^k] + \frac{L}{2} \sum_{(i,j)} |F_{ij}(t) - F_{ik}^k|^2. \tag{A19}$$

We derive an upper bound for the expected value of each of the last two terms above.

Denote by d_p^k the length of path p corresponding to the link flows F_{ij}^k. We have

$$d_p^k = \sum_{(i,j)\varepsilon p} d_{ij}(F_{ij}^k) \quad , \quad \forall\ w\varepsilon W,\ p\varepsilon P_w$$

and it follows that

$$\sum_{(i,j)} d_{ij}(F_{ij}^k)[F_{ij}(t) - F_{ij}^k] = \sum_{(i,j)} d_{ij}(F_{ij}^k) \sum_{w\varepsilon W} \sum_{\substack{p\varepsilon P_w \\ (i,j)\varepsilon p}} [x_p(t)-x_p^k]$$

$$= \sum_{w\varepsilon W} \sum_{p\varepsilon P_w} d_p^k[x_p(t)-x_p^k] \tag{A20}$$

Let $p_k \varepsilon P_w$ be the shortest path used for routing in $[kT,(k+1)T)$ and define

$$\bar{x}_p^k = \begin{cases} 0 & \text{if } p \neq p_k \\ \bar{r}_w & \text{if } p = p_k \end{cases} \tag{A21}$$

Taking conditional expectation in (A20) and using (A3)

$$E\{\sum_{(i,j)} d_{ij}(F_{ij}^k)[F_{ij}(t)-F_{ij}^k]|x^k\} = \sum_{w\varepsilon W} \sum_{p\varepsilon P_w} d_p^k[E\{x_p(t)|x_p^k\} - x_p^k] \tag{A22}$$

$$= \sum_{w\varepsilon W} [1-e^{-\mu_w(t-kT)}] \sum_{p\varepsilon P_w} d_p^k(\bar{x}_p^k-x_p^k)$$

$$= \sum_{w\varepsilon W} [1-e^{-\mu_w(t-kT)}][\sum_{p\varepsilon P_w} d_p^k(\bar{x}_p^k-\tilde{x}_p^k) +$$

$$+ \sum_{p\varepsilon P_w} d_p^k(\tilde{x}_p^k-x_p^k)]$$

where $\tilde{x}_p^k$ is given by (A12). Since $\sum_{p\varepsilon P_w} \bar{x}_p^k = \sum_{p\varepsilon P_w} \tilde{x}_p^k = \bar{r}_w$ and, for each w, p_k is the shortest path we obtain using (A21)

$$\sum_{p\varepsilon P_w} d_p^k \bar{x}_p^k \leq \sum_{p\varepsilon P_w} d_p^k \tilde{x}_p^k$$

so (A22) can be strengthened to yield

$$E\{\sum_{(i,j)} d_{ij}(F_{ij}^k)[F_{ij}(t)-F_{ij}^k]|x^k\} \leq [1-e^{-\mu(t-kT)}] \sum_{w\varepsilon W} \sum_{p\varepsilon P_w} d_p^k(\bar{x}_p^k-\tilde{x}_p^k)$$

$$+ \sum_{w\varepsilon W} [1-e^{-\mu_w(t-kT)}] \sum_{p\varepsilon P_w} d_p^k(\tilde{x}_p^k-x_p^k)$$

$$= [1-e^{-\mu(t-kT)}] \sum_{w\varepsilon W} \sum_{p\varepsilon P_w} d_p^k(\bar{x}_p^k-x_p^k)$$

$$+ \sum_{w\varepsilon W} [e^{-\mu(t-kT)} - e^{-\mu_w(t-kT)}] \sum_{p\varepsilon P_w} d_p^k(\tilde{x}_p^k-x_p^k) . \tag{A23}$$

We proceed to bound each of the two terms in the right side above.

Let $\{x_p^*|w\varepsilon W,\ p\varepsilon P_w\}$ be any set of path flows minimizing $D(F)$ over F i.e., any $x_p^* \geq 0$ such that

$$F_{ij}^* = \sum_{w\varepsilon W} \sum_{\substack{p\varepsilon P_w \\ (i,j)\varepsilon p}} x_p^*, \quad \forall\ (i,j)\varepsilon L.$$

Since for each w the shortest path is p_k and $\sum_{p\varepsilon P_w} x_p^* = \sum_{p\varepsilon P_w} \bar{x}_p^k = \bar{r}_w$

we have

$$\sum_{p\varepsilon P_w} d_p^k(\bar{x}_p^k-x_p^k) \leq \sum_{p\varepsilon P_w} d_p^k(x_p^*-x_p^k) \tag{A24}$$

while similarly as earlier [cf. (A20)] we have

$$\sum_{w\varepsilon W} \sum_{p\varepsilon P_w} d_p^k(x_p^*-x_p^k) = \sum_{(i,j)} d_{ij}(F_{ij}^k)(F_{ij}^*-F_{ij}^k). \tag{A25}$$

Since D is convex we obtain

$$\sum_{(i,j)} d_{ij}(F_{ij}^k)(F_{ij}^*-F_{ij}^k) \leq D(F^*) - D(F^k) = D^* - D(F^k). \tag{A26}$$

Combining (A24)-(A26) we see that

$$\sum_{w\varepsilon W} \sum_{p\varepsilon P_w} d_p^k(\bar{x}_p^k - x_p^k) \leq D^* - D(F^k). \tag{A27}$$

which provides a bound for the first term on the right in (A23).

To obtain a bound for the second term on the right of (A23) we write

$$\sum_{p\varepsilon P_w} d_p^k(\tilde{x}_p^k - x_p^k) = \sum_{p\varepsilon P_w} (d_p^k - \tilde{d}_p^k)(\tilde{x}_p^k - x_p^k) + \sum_{p\varepsilon P_w} \tilde{d}_p^k(\tilde{x}_p^k - x_p^k) \tag{A28}$$

where $\tilde{d}_p^k$ is the length of path p if each flow x_p^k is replaced by $\tilde{x}_p^k$, i.e.

$$\tilde{d}_p^k = \sum_{(i,j\varepsilon p} d_{ij}(\tilde{F}_{ij}^k).$$

Using (7) and (A13) it is easily seen that for some constant $\xi > 0$

$$E\{\sum_{p\ P_w} (d_p^k - \tilde{d}_p^k)(\tilde{x}_p^k - x_p^k)\} \leq L\, E\{\sum_{p\varepsilon P_w} \sum_{(i,j)\varepsilon p} |\tilde{F}_{ij}^k - F_{ij}|^2\}$$

$$\leq \xi[\varepsilon\bar{\gamma}(\bar{r}+e^{-\mu kT}R) + e^{-2\mu kT}R^2].$$

Using (A12) we have

$$\sum_{p\varepsilon P_w} \tilde{d}_p^k(\tilde{x}_p^k - x_p^k) = \frac{\bar{r}_w - r_w^k}{r_w^k} \sum_{p\varepsilon P_w} \tilde{d}_p^k\, x_p^k$$

$$\leq \frac{|\bar{r}_w - r_w^k|B}{r_w^k} \sum_{p\varepsilon P_w} x_p^k = B|\bar{r}_w - r_w^k|$$

where B is the constant defined in Lemma 2.

We have

$$E\{|\bar{r}_w - r_w^k|\} \le E|\{\bar{r}_w - E\{r_w^k\}|\} + E\{|E\{r_w^k\} - r_w^k|\}$$

$$\le E\{|\bar{r}_w - E\{r_w^k\}|\} + \sqrt{\operatorname{var}\{r_w^k\}}$$

where the last step follows using Jensen's inequality. Therefore using (A10) and (A11) we obtain

$$E\{|\bar{r}_w - r_w^k|\} \le e^{-\mu t} R + \sqrt{\varepsilon\bar{\gamma}(\bar{r} + e^{-\mu t} R)}$$

$$\le e^{-\mu t} R + \sqrt{\varepsilon\bar{\gamma}(\bar{r} + R)} ,$$

and

$$E\{\sum_{p\varepsilon P_w} \tilde{d}_p^k(\tilde{x}_p^k - x_p^k)\} \le B[e^{-\mu kT} R + \sqrt{\varepsilon\bar{\gamma}(\bar{r} + R)}\ ] .$$

Taking expectation over x^k in (A28) and using the inequalities above we obtain for some constant $\zeta > 0$

$$E\{\sum_{w\varepsilon W} [e^{-\mu(t-kT)} - e^{-\mu_w(t-kT)}] \sum_{p\varepsilon P_w} d_p^k(\tilde{x}_p^k - x_p^k)\}$$
$$\le \zeta[e^{-\mu(t-kT)} - e^{-M(t-kT)}][\varepsilon\bar{\gamma}(\bar{r} + e^{-\mu kT}) + e^{-2\mu kT} R^2 + e^{-\mu kT} R + \sqrt{\varepsilon\bar{\gamma}(\bar{r} + R)}\ ] \tag{A28}$$

Combining (A23), (A27), (A28), and taking expectation over x^k we obtain for some constant β_1

$$E\{\sum_{(i,j)} d_{ij}(F_{ij}^k)[F_{ij}(t) - F_{ij}^k]\} \leq [1-e^{-\mu(t-kT)}][D^* - E\{D(F^k)\}] \tag{A29}$$

$$+ \beta_1[e^{-\mu(t-kT)} - e^{-M(t-kT)}][\varepsilon\bar{\gamma}(\bar{r} + e^{-\mu kT} R) + e^{-2\mu kT} R^2 + e^{-\mu kT} R + \sqrt{\varepsilon\bar{\gamma}(\bar{r}+R)}\ ],$$

which provides the desired bound on the expected value of the next to last term in (A19).

We now bound the expected value of the last term in (A19). Since F_{ij}^k and $F_{ij}(t)$ are sums of path flows x_p^k and $x_p(t)$ respectively we have that there exists a constant θ such that

$$\sum_{(i,j)} |F_{ij}(t) - F_{ij}^k|^2 \leq \theta \sum_{w\varepsilon W} \sum_{p\varepsilon P_w} |x_p(t) - x_p^k|^2. \tag{A30}$$

We have

$$E\{|x_p(t) - x_p^k|^2|x_p^k\} = \mathrm{var}\{x_p(t)|x_p^k\} + [x_p^k - E\{x_p(t)|x_p^k\}]^2$$

and using Lemma 1 we obtain

$$\sum_{p\varepsilon P_w} E\{|x_p(t) - x_p^k|^2|x_p^k\} = \varepsilon\bar{\gamma}_w[1-e^{-\mu_w(t-kT)}][\bar{r}_w + e^{-\mu_w(t-kT)} r_w^k]$$

$$+ [1-e^{-\mu(t-kT)}]^2[(\bar{r}_w - x_{p_k}^k)^2 + \sum_{\substack{p\varepsilon P_w \\ p\neq p_k}} (x_p^k)^2]$$

$$\leq [1-e^{-\mu(t-kT)}][\varepsilon\bar{\gamma}_w(\bar{r}_w + r_w^k) + (1-e^{-\mu T})(\bar{r}_w + r_w^k)^2].$$

Taking expectation over x^k and using (A10), (A11) we obtain

$$\sum_{p\varepsilon P_w} E\{|x_p(t) - x_p^k|^2\} \leq [1-e^{-\mu(t-kT)}]\{\varepsilon\bar{\gamma}(\bar{r}_w+E\{r_w^k\}) \tag{A31}$$

$$+ (1-e^{-\mu T})[\bar{r}_w^2 + 2\bar{r}_w E\{r_w^k\} + (E\{r_w^k\})^2 + \mathrm{var}\{r_w^k\}]$$

$$\leq [1-e^{-\mu(t-kT)}]\{\varepsilon\bar{\gamma}(2\bar{r} + e^{-\mu kT} R)$$

$$+ (1-e^{-\mu T})[(2\bar{r} + e^{-\mu kT} R)^2 + \varepsilon\bar{\gamma}(\bar{r} + e^{-\mu kT} R)]\}$$

We now combine (A19), (A29)-(A31) to obtain for all $t\varepsilon[kT, (k+1)T]$ and some positive constant β_2

$$E\{D[F(t)]\} - D^* \leq e^{-\mu(t-kT)}[E\{D(F^k)\} - D^*] \tag{A31}$$

$$+ \beta_1[e^{-\mu(t-kT)} - e^{-M(t-kT)}][\varepsilon\bar{\gamma}(\bar{r} + e^{-\mu kT}R) + e^{-2\mu kT}R^2 + e^{-\mu kT}R + \sqrt{\varepsilon\bar{\gamma}(\bar{r}+R)}\ ]$$

$$+ \beta_2[1-e^{-\mu(t-kT)}]\{\varepsilon\bar{\gamma}(2\bar{r} + e^{-\mu kT}R) + (1-e^{-\mu T})[(2\bar{r} + e^{-\mu kT}R)^2 + \varepsilon\bar{\gamma}(\bar{r} + e^{-\mu kT}R)]\}.$$

By applying this inequality for $t = (k+1)T$, setting $c_2 = \max\{\beta_1,\beta_2\}$ and collecting terms we obtain

$$E\{D(F^{k+1})\} - D^* \leq e^{-\mu T}[E\{D(F^k)\} - D^*] \tag{A32}$$

$$+ c_2[\bar{a}(\varepsilon,T) + \bar{b}(\varepsilon,T)e^{-\mu kT}]$$

where

$$\bar{a}(\varepsilon,T) = \bar{r}\{(e^{-\mu T} - e^{-MT})(\varepsilon\bar{\gamma} + \frac{\sqrt{\varepsilon\bar{\gamma}(\bar{r}+R)}}{\bar{r}}) + (1-e^{-\mu T})[2\varepsilon\bar{\gamma} + (1-e^{-\mu T})(4\bar{r}+\varepsilon\bar{\gamma})]\} \tag{A33}$$

$$\bar{b}(\varepsilon,T) = R\{(e^{-\mu T} - e^{-MT})(\varepsilon\bar{\gamma} + R + 1) \tag{34}$$

$$+ (1-e^{-\mu T}) + [\varepsilon\,\bar{\gamma} + (1-e^{-\mu T})(4\bar{r} + R + \varepsilon\bar{\gamma})]\}.$$

Applying (A32) repeatedly for k equal to zero up to (k-1) we obtain

$$E\{D(F^k)\} - D^* \leq e^{-\mu kT}[D(F^o) - D^*]$$

$$+ c_2 \left[\frac{\bar{a}(\varepsilon,T)}{1-e^{-\mu T}} + \frac{\bar{b}(\varepsilon,T)}{Te^{-\mu T}} kTe^{-\mu kT} \right]$$

which is the desired right side of relation (13) [compare (14), (15) with (A33), (A34)].

Since

$$\lim_{\substack{\varepsilon\to 0 \\ T\to 0}} \frac{\bar{a}(\varepsilon,T)}{1-e^{-\mu T}} = 0 \text{ and } \lim_{\substack{\varepsilon\to 0 \\ T\to 0}} \frac{\bar{b}(\varepsilon,T)}{Te^{-\mu T}} < \infty$$

we see that $E\{D[F(kT)]\} \to D^*$ as $\varepsilon \to 0$, $T \to 0$ and $kT \to \infty$. It follows from (A31) that $E\{D[F(t)]\} \to D^*$ as $\varepsilon \to 0$, $T \to 0$, and $t \to \infty$.

To show the last part of the theorem we use Taylor's theorem and the hypothesis $\ell|\bar{F}-F| \leq |d_{ij}(\bar{F}) - d_{ij}(F)|$ to write for any vector $F \varepsilon \mathcal{F}$

$$D(F) = D(F^*) + \sum_{(i,j)} d_{ij}(F^*_{ij})(F_{ij}-F^*_{ij})$$

$$+ \sum_{(i,j)} \int_0^1 \{d_{ij}[F^*_{ij} + \alpha(F_{ij}-F^*_{ij})] - d_{ij}(F^*_{ij})\}(F_{ij}-F^*_{ij})d\alpha$$

$$\geq D(F^*) + \sum_{(i,j)} d_{ij}(F^*_{ij})(F_{ij}-F^*_{ij}) + \frac{\ell}{2} \sum_{(i,j)} |F_{ij}-F^*_{ij}|^2 .$$

Since F^* minimizes D over $\mathcal{F}$ we have the optimality condition $\sum_{(i,j)} d_{ij}(F^*_{ij})(F_{ij}-F^*_{ij}) \geq 0$ and it follows that

$$D(F) \geq D^* + \frac{\ell}{2} \sum_{(i,j)} |F_{ij}-F^*_{ij}|^2 , \quad \forall F \varepsilon \mathcal{F} .$$

Therefore using also Lemma 2 we have

$$D^* + \frac{\ell}{2} \sum_{(i,j)} E\{|\tilde{F}_{ij}(t) - F^*_{ij}|^2\} \leq E\{D[\tilde{F}(t)]\}$$

$$\leq E\{D[F(t)]\} + B \sum_{(i,j)} E\{|\tilde{F}_{ij}(t) - F_{ij}(t)|\}$$

Since $E\{|\tilde{F}_{ij}(t) - F_{ij}(t)|\} \to 0$ [cf. (A13)] and $E\{D[F(t)]\} \to D^*$ as $\varepsilon \to 0$, $T \to 0$ and $t \to \infty$ we obtain that $\tilde{F}_{ij}(t)$ converges in mean square to F^*_{ij}. Since $\{\tilde{F}_{ij}(t) - F_{ij}(t)\}$ also converges to zero in mean square [cf. (A13)] we obtain that $F(t)$ converges to F^* in mean square. Q.E.D.

References

[1] M. Schwartz and T. E. Stern, "Routing Techniques Used in Computer Communication Networks", IEEE Trans. on Communications, Vol. COM-28, 1980, pp. 539-552.

[2] J. M. McQuillan, I. Richer and E. C. Rosen, "The New Routing Algorithm for the ARPANET", IEEE Trans. on Communications, Vol. COM-28, 1980, pp. 711-719.

[3] L. Kleinrock, Communication Nets: Stochastic Message Flow and Delay, McGraw-Hill, N.Y., 1964.

[4] R. G. Gallager, "A Minimum Delay Routing Algorithm Using Distributed Computation", IEEE Trans. on Communications, Vol. COM-25, 1977, pp. 73-85.

[5] A. Segall, "Optimal Routing for Virtual Line-Switched Data Networks", IEEE Trans. on Communications, Vol. COM-26, 1979.

[6] T. E. Stern, "A Class of Decentralized Routing Algorithms Using Relaxation", IEEE Trans. on Communication, Vol. COM-25, 1977, pp. 1092-1102.

[7] L. Fratta, M. Gerla, and L. Kleinrock, "The Flow Deviation Method: An Approach to Store-and-Forward Communication Network Design", Networks, Vol. 3, 1973, pp. 97-133.

[8] D. G. Cantor and M. Gerla, "Optimal Routing in a Packet Switched Computer Network", IEEE Trans. on Computation, Vol. C-23, 1974, pp. 1062-1069.

[9] D. P. Bertsekas and R. G. Gallager, Data Networks, Prentice Hall, Englewood Cliffs, N.J., 1986.

[10] D. P. Bertsekas and E. M. Gafni, "Projection Methods for Variational Inequalities with Application to the Traffic Assignment Problem", Math. Programming Study 17, 1982, pp. 139-159.

[11] D. P. Bertsekas and E. M. Gafni, "Projected Newton Methods and Optimization of Multicommodity Flows", LIDS Report P-1140, M.I.T., Aug. 1981, IEEE Trans. on Aut. Control, Vol. AC-28, 1983, pp. 1090-1096.

[12] D. P. Bertsekas, E. M. Gafni, and R. G. Gallager, "Second Derivative Algorithms for Minimum Delay Distributed Routing in Networks", IEEE Trans. on Communications, Vol. COM-32, 1984, pp. 911-919.

[13] D. P. Bertsekas, "Optimal Routing and Flow Control Methods for Communication Networks", in Analysis and Optimization of Systems, by A. Bensoussan and J. L. Lions (eds.), Springer-Verlag, Berlin, & N.Y., 1982, pp. 615-643.

[14] F. H. Moss and A. Segall, "An Optimal Control Approach to Dynamic Routing in Communication Networks", IEEE Trans. on Aut. Control, Vol. 27, 1982, pp. 329-339.

[15] B. Hajek and R. G. Ogier, "Optimal Dynamic Routing in Communication Networks with Continuous Traffic", Networks, Vol. 14, 1984, pp. 457-487.

[16] H. Schwartz, Computer Communication Network Design and Analysis, Prentice Hall, Englewood Cliffs, N.J., 1977.

[17] D. P. Bertsekas, "Dynamic Behavior of Shortest Path Routing Algorithm for Communication Networks", IEEE Trans. on Aut. Control, Vol. AC-27, 1982, pp. 60-74.

[18] D. P. Bertsekas, "Dynamic Models of Shortest Path Routing Algorithms for Communication Networks with Multiple Destinations", Proc. of 1979 IEEE Conf. on Decision and Control, Ft. Lauderdale, Fla., 1979, pp. 127-133.

[19] J. M. McQuillan, G. Falk, and I. Richer, "A Review of the Development and Performance of the ARPANET Routing Algorithm", IEEE Trans. on Communications, Dec. 1978.

[20] L. Kleinrock, Queueing Systems, Vol. II, J. Wiley, N.Y., 1976.

[21] L, Kleinrock, Queueing Systems, Vol. I, J. Wiley, N.Y., 1975.

[22] E. M. Gafni and D. P. Bertsekas, "Path Assignment for Virtual Circuit Routing" SIGCOMM 83 Symposium on Communications Architectures and Protocols, Austin, Texas, March 1983, pp. 21-25.

Distribution List

Defense Documentation Center Cameron Station Alexandria, Virginia 22314	12 Copies
Assistant Chief for Technology Office of Naval Research, Code 200 Arlington, Virginia 22217	1 Copy
Office of Naval Research Information Systems Program Code 437 Arlington, Virginia 22217	2 Copies
Office of Naval Research Branch Office, Boston 495 Summer Street Boston, Massachusetts 02210	1 Copy
Office of Naval Research Branch Office, Chicago 536 South Clark Street Chicago, Illinois 60605	1 Copy
Office of Naval Research Branch Office, Pasadena 1030 East Greet Street Pasadena, California 91106	1 Copy
Naval Research Laboratory Technical Information Division, Code 2627 Washington, D.C. 20375	6 Copies
Dr. A. L. Slafkosky Scientific Advisor Commandant of the Marine Corps (Code RD-1) Washington, D.C. 20380	1 Copy

Office of Naval Research
Code 455
Arlington, Virginia 22217

1 Copy

Office of Naval Research
Code 458
Arlington, Virginia 22217

1 Copy

Naval Electronics Laboratory Center
Advanced Software Technology Division
Code 5200
San Diego, California 92152

1 Copy

Mr. E. H. Gleissner
Naval Ship Research & Development Center
Computation and Mathematics Department
Bethesda, Maryland 20084

1 Copy

Captain Grace M. Hopper
Naval Data Automation Command
Code OOH
Washington Navy Yard
Washington, DC 20374

1 Copy

Advanced Research Projects Agency
Information Processing Techniques
1400 Wilson Boulevard
Arlington, Virginia 22209

1 Copy

Dr. Stuart L. Brodsky
Office of Naval Research
Code 432
Arlington, Virginia 22217

1 Copy

Prof. Fouad A. Tobagi
Computer Systems Laboratory
Stanford Electronics Laboratories
Department of Electrical Engineering
Stanford University
Stanford, CA 94305