DETERMINISTIC, ROUTING TO BUFFERED CHANNELS

by

Zvi Rosberg

Technical Report #318

May 1984
ABSTRACT

Consider n exponential transmission channels which transmit information with different rates. Every channel has a buffer which is capable of storing an unlimited number of messages. A new message first arrives at the controller which immediately routes it to one of the channels according to an infinite deterministic routing sequence. A cost $C^{(i)}$ per unit of staying time is charged in each of the channels and the long-run average staying cost is taken as the cost criterion. For $n = 2$ and a renewal arrival process, an $\varepsilon$-optimal routing policy is given.

For $n > 2$ and a Poisson arrival process, a lower bound to the cost criterion is found and a new routing policy, the Golden Ratio policy, is presented and its cost is found. It is also shown that for a variety of systems the Golden Ratio policy has a cost close to the lower bound.

Keywords: Communication channels, routing protocols, Golden Ratio, distributed protocols, Matrix-Geometric solutions.
1. INTRODUCTION

Message routing in a communication network was extensively studied by several authors, e.g., [Be], [EVW], [FoSa], [FGK], [Ga], [Se], [RoTo].

In most studies, static routing policies were analyzed. By static policy one means a policy which routes an outgoing message from a node with a fixed probability, independent of the past history of the entire network. Various algorithms have been proposed to find the optimal routing probabilities which minimize the long-run average delay per message (e.g., [Be], [FGK], [Ga], [Se]). However, it is often the case that non-static routing policies achieve a smaller average delay. For example, it was shown in [EVW] that for distributively routing a single stream of messages through two parallel, identical and exponential channels, the policy which sends every other arrival to one channel, minimizes the average delay. For more general networks it appears very difficult to find the minimum delay routing policies although it is clear that in most cases they are not static policies (see [EVW], [Yu], [RoTo]).

Dynamic routings were explored in [MoSe], whereby the decision on how to forward messages through the network is based on measuring the length of the queues at the network nodes. The authors gave a conceptual form of an algorithm for finding a feedback solution to the optimal control problem when the input to each node is constant in time.

In this paper we study a single node with several heterogeneous outgoing links under a subset of the distributed dynamic routing
policies and a general cost criterion. We use recent results from [Ha1], [Ha2], to bound the cost criterion and analyze a new routing policy which was suggested by us in [KoTo].

Consider \( n \) parallel transmission channels which transmit information with different rates. That is, the transmission time of a message through a channel depends on the channel.

Every channel has a buffer which is capable of storing an unlimited number of messages and can transmit at most one message at any moment of time. A message is transmitted without interruption. A new message first arrives at the controller which immediately routes it to one of the channels. In a physical environment the controller is an intelligent multiplexer or a transmission node in a network; and the channels are parallel computer components or a subset of the outgoing links respectively. For a graphical representation see Figure 1.

We assume that the controller routes the messages instantly and the messages are enqueued at the buffers of the channels. We further assume that there is no travel time between the controller and the channels.

The messages arrive at the controller as a Poisson process with rate \( \lambda \). (For \( n = 2 \) we assume only a renewal arrival process.) A message which is routed to a channel joins the end of the queue, or starts transmission immediately, if the buffer is empty.

The transmission time of a message through channel \( i \), \( 1 \leq i \leq n \), is exponentially distributed with rate \( \mu_i \). All transmission times are assumed to be independent. We also assume that the known necessary
condition for ergodicity holds, that is
\[ \lambda < \sum_{i=1}^{n} \mu_i . \]

For practical reasons we are interested in routing policies which can be distributively implemented (that is, the controller does not observe the contents of the buffers).

A general set of distributed routing policies is the following set, so called deterministic routing policies.

Suppose that a deterministic \( n \)-valued sequence \( r = (r_1, r_2, r_3, \ldots) \), \( 1 \leq r_k \leq n \), is given and the messages are numbered according to their order of arrival. For every \( k \), message \( k \) is routed to channel \( r_k \). We refer to these \( r \)'s as routing sequences.

To reflect the differences in the utility among the channels, we charge every message at channel \( i \), a cost \( C^{(i)} \), per unit of staying time in the channel (either being queued or transmitted).

Let \( V_k^{(i)}(r) \) be the total expected cost at channel \( i \), \( 1 \leq i \leq n \), until the \( k \)-th arrival at channel \( i \), using the routing sequence \( r \).

Define the total long-run average cost of using \( r \)

\[
\bar{V}(r) = \lim_{k \to \infty} \frac{1}{k} \sum_{i=1}^{n} V_k^{(i)}(r). \tag{1.1}
\]

Also, let
\[ \bar{V} = \inf_r \bar{V}(r). \]

A routing sequence \( r^* \) is optimal if
\[ \bar{V}(r^*) = \bar{V}. \]
In a previous study, [RoTo], we considered a simpler model for the routing problem. In that model the channels are assumed to be slotted, messages arrive as a binomial process, then they are "packetized" and routed to one of the channels. The number of slots which are required to transmit a message through a channel is a geometric random variable with a channel dependent parameter. The channels have no buffers and the total throughput of the channels is taken as the cost criterion. A simple deterministic routing policy (the Golden Ratio policy) was proposed there and it was shown that its throughput is within at least 98.4% of the upper bound. In this study we show that the very same policy performs very well (in most of the cases) in a more realistic model. In Section 2 we give a lower bound to $\bar{\nu}$. In Section 3 the $\epsilon$-optimal deterministic routing policy is found and in Section 4 we define and analyze the Golden Ratio routing policy. We also present performance values for various systems which show that the Golden Ratio policy is a promising one.
2. A LOWER BOUND

The lower bound to $\bar{V}$ is based on the results which have been obtained in [Hal].

For a routing sequence $r$ let $N_{j}(i)(r)$ be the number of messages in the buffer of channel $i$ (including the one which is being transmitted) just prior to the $j$-th arrival at channel $i$. Also, let

$$N_{j}(i)(r) = \lim inf_{k \to \infty} \frac{1}{k} \sum_{j=1}^{k} E[N_{j}(i)(r)]$$

where the expectation is taken with respect to the probability measure which is introduced by the arrival process, the transmission times and the routing sequence $r$.

From the definition

$$V_{k}(i)(r) = \sum_{j=1}^{k} \sum_{x} C(i)x E[I(N_{j}(i)(r) = x)],$$

where $I(A)$ is the indicator function of the set $A$ and the second sum is taken over all possible $x$'s.

Thus, under stationary conditions, the limit in (1.1) exists and

$$\bar{V}(r) = \sum_{i=1}^{n} C(i)N(i)(r).$$

(2.1)

Let $0 \leq p(i) < 1$, $1 \leq i \leq n$, and let $A_{k}(i)(r)$ be the number among the first $k$ arrivals, which are routed to channel $i$, under the routing sequence $r$. It was shown by Hajek that (see [Hal, Sec. 5]):

**Theorem 2.1 (Hajek)** Let $r$ be a routing sequence.
If \( \liminf_{k \to \infty} \frac{A_k^{(i)}(x)}{k} \geq p(i) \) and \( p(i) < \min(1, \frac{\mu_i}{\lambda}) \) then

\[
\bar{N}(i)(x) \geq \frac{x(i)}{1-x(i)},
\]

where \( x(i) \) is the unique solution of the equation

\[
\left( \frac{1}{1+(1-x(i)) \frac{\mu_i}{\lambda}} \right)^{1/p(i)} = x(i).
\]

Moreover, if \( p(i) > \frac{\mu_i}{\lambda} \) then \( \bar{N}(i)(x) = \infty \).

**Corollary 2.1:** If the conditions of Theorem 2.1 hold and if \( \frac{1}{p(i)} \) is an integer, then the lowest expected number of messages in buffer \( i \) is obtained by routing every \( \frac{1}{p(i)} \) message to channel \( i \).

**Proof:** It is well known that (see e.g. [K1, p.251]) in a GI/M/1 queueing system, where the arrival process is composed of every \( \frac{1}{p(i)} \) arrival from a Poisson process, \( \frac{x(i)}{1-x(i)} \), is the stationary expected number of messages in the buffer.

Since every message must be routed to one of the channels and \( C(i) \) are non-negative constants, a lower bound to \( \bar{V} \) is obtained by solving the following extremum problem:

\[
\min \sum_{i=1}^{n} \frac{C(i) x(i)}{1-x(i)},
\]
subject to

\[
\left( \frac{1}{1+\frac{x(i)}{\mu_i}} \right)^{1/p(i)} = x(i),
\]

\[0 \leq x(i) \leq \min(1, \frac{\lambda}{\nu_i}),\]

\[0 \leq x(i) \leq \min(1, \frac{\lambda}{\nu_i}),\]

\[\sum_{i=1}^{n} p(i) = 1.\]

Let \( \alpha_i = \frac{\mu_i}{\lambda} \) and \( g_i(x) = \frac{\ln(1+ (1-x)\alpha_i)}{-\ln x} \).

It is easy to verify that the extremum problem can be expressed as a function of the \( x(i) \)'s only:

\[
\min \sum_{i=1}^{n} \frac{c(i) \lambda(i)}{1-x(i)} \quad (2.3)
\]

subject to:

\[0 \leq x(i) \leq \min(1, \frac{1}{\alpha_i}),\]

\[0 \leq g_i(x(i)) \leq \alpha_i,\]

\[\sum_{i=1}^{n} g_i(x(i)) = 1.\]  (2.4)

Since \( g_i(x) \) increases in the interval \( (0, \min(1, \frac{1}{\alpha_i})) \), the set of feasible solutions is convex. Furthermore, the cost function is an increasing convex function. Thus, we can use the Lagrangian multipliers technique.
Define the Lagrangian

\[ F(x(1), x(2), \ldots, x(n), \eta) = \sum_{i=1}^{n} \frac{C(i)x(i)}{1-x(i)} - \eta (1 - \sum_{i=1}^{n} g_i(x(i))). \]

The \( x(i) \)'s which solve (2.3), (2.4) are the unique solution of

\[ \frac{\partial F}{\partial x(i)} = \frac{C(i) x(i)}{(1-x(i))^2} + \eta \frac{\partial g_i(x(i))}{\partial x(i)} = 0, \quad 1 \leq i \leq n \]

\[ \frac{\partial F}{\partial \eta} = \sum_{i=1}^{n} g_i(x(i)) - 1 = 0, \quad (2.5) \]

in the intervals \((0, \min(1, \frac{1}{\alpha_i}))\), \(1 \leq i \leq n\).

The solution cannot be expressed in a close form. However, by using Newton's algorithm, one can efficiently solve (2.5) and get the optimal \( x(i) \)'s.

Let \( x(i)^* \), \( 1 \leq i \leq n \), be the optimal \( x(i) \)'s and \( p^{(i)*} = g_i(x(i)^*) \) be the "desirable routing proportions".

From Corollary 2.1 it follows that, if there is a policy which routes every \( \frac{1}{p^{(i)}} \) message to channel \( i \), for every \( i \), then the lower bound will be feasible.

Unfortunately, an equally distance routing policy for all the channels is almost never feasible. The infeasibility springs from two sources:

(1) Usually, \( \frac{1}{p^{(i)}} \) are not integers.

(2) Even when all the \( \frac{1}{p^{(i)}} \) are integers, the implied inter-routing distances for different channels will usually clash. (See Example 4.1 in [RoTo].)
A situation in which the lower bound is clearly obtained is the homogeneous case (i.e., \( \mu_i = \mu, 1 \leq i \leq n \), for some \( \mu \)). In this case, by symmetry, \( p(i)^* = \frac{1}{n} \) and the optimal deterministic routing policy is the so called round-robin policy.

In Section 4 we present a deterministic routing policy, the Golden Ratio policy (see also [RoTo]), which approximates the optimal solution of (2.3), (2.4) by routing approximately every \( \frac{1}{p(i)^*} \) message to channel \( i \), for every \( i, 1 \leq i \leq n \).

For \( n = 2 \), we derive an \( \epsilon \)-optimal policy for every \( \epsilon > 0 \). This is done in the next section.
3. AN $\varepsilon$-OPTIMAL ROUTING POLICY FOR TWO CHANNELS

In this section we assume that $n = 2$ and the arrival process is a renewal process.

We also observe the process at a slightly different set of imbedded points. Let $r$ be a given routing sequence and $Y_k^{(i)}(r)$, $i = 1, 2$, be the number of messages in the buffer of channel $i$ (including the message which is being transmitted) just prior to the arrival of message $k$ at the system.

Let

$$\bar{Y}_k^{(i)}(r) = \lim_{k \to \infty} \frac{1}{k} \sum_{j=1}^{k} E(Y_j^{(i)}(r)),$$

$$\bar{W}(r) = C(1)\bar{Y}^{(1)}(r) + C(2)\bar{Y}^{(2)}(r),$$

and

$$\bar{W} = \inf_r \bar{W}(r).$$

A routing sequence $r(\varepsilon)$ is $\varepsilon$-optimal if

$$\bar{W}(r(\varepsilon)) \leq \bar{W} + \varepsilon.$$

For a single channel, a feasible lower bound to $\bar{Y}^{(1)}(r)$ (in the context of Theorem 2.1) is given in the outstanding paper [Ha2]. (Theorem 3.1 below.) We shall show that the same routing sequence (or a very close sequence) is also feasible for two channels.

For every $Q \leq p \leq 1$ and a constant $\alpha$, define the following routing sequence $r^*(p, \alpha) = (r_1^*(p, \alpha), r_2^*(p, \alpha), \ldots)$

$$r_k^*(p, \alpha) = \begin{cases} 1 & \text{if } \lfloor (k+1)p + \alpha \rfloor - \lfloor kp + \alpha \rfloor = 1 \\ 2 & \text{if } \lfloor (k+1)p + \alpha \rfloor - \lfloor kp + \alpha \rfloor = 0, \end{cases}$$

(3.1)

where $\lfloor x \rfloor$ is the greatest integer smaller than or equal to $x$. 

Technion - Computer Science Department - Technical Report CS0318 - 1984
We shall omit both arguments \( p, \alpha \), or one of them, whenever they will be clear from the context.

Let \( r^* \) be the complement sequence of \( r^* \) (i.e., every 1 is replaced by 2 and vice versa).

**Theorem 3.1 (Hajek)** Let \( r \) be a routing sequence.

If \( \liminf_{k \to \infty} \frac{A_k^{(1)}(r)}{k} \geq p \) and \( p < \min(1, \frac{\mu_1}{\lambda}) \), then

\[
\tilde{\gamma}^{(1)}(r^*(p)) \leq \tilde{\gamma}^{(1)}(r).
\]

For the proof see Theorem 6.2 in [Ha2].

Let \( r \) be any routing sequence. Define

\[
p(i)_r = \liminf_{k \to \infty} \frac{A_k^{(i)}(r)}{k}, \quad i = 1, 2.
\]

**Corollary 3.1**

(a) If \( \tilde{p}^{(i)} > \frac{\mu_i}{\lambda} \) for some \( i \), then \( \bar{W}(r) = \infty \)

(b) If \( \tilde{p}^{(i)} < \frac{\mu_i}{\lambda} \) for \( i = 1, 2 \), then

\[
\bar{W}(r) \geq C^{(1)}(1)\tilde{\gamma}^{(1)}(r^*(\tilde{p}^{(1)}), \alpha_1) + C^{(2)}\tilde{\gamma}^{(2)}(r^*(\tilde{p}^{(2)}), \alpha_2),
\]

where \( \alpha_1, \alpha_2 \) are any constants.

**Proof:** Part (a) follows from a well known result on the ergodicity of \( G/G/1 \) queues. Part (b) is a direct consequence of Theorem 3.1 and the assumption \( C^{(i)} > 0 \).

Note that the ergodicity assumption \( \lambda < \mu_1 + \mu_2 \) implies that there are routing sequences with finite \( \bar{W}(r) \).
Henceforth, we consider only routing sequences \( r \), for which
\[
\lim_{k \to \infty} \frac{A_k^{(1)}(r)}{k}
\]
exists. Denote by \( A \) this set of sequences.

Note that for every \( \varepsilon > 0 \), it is easy to construct a routing sequence \( r \),
with \( \lim \inf_{k \to \infty} \frac{A_k^{(i)}(r)}{k} < \varepsilon \), for \( i = 1, 2 \). Thus, without the restriction
to the set \( A \), \( \bar{W}(r) \) can approach zero as close as we please. Also from practical reasons, one would not use \( r \notin A \), since the delay of the messages would extremely fluctuate.

The following corollary is also a direct consequence of Theorem 3.1.

\[ \text{Corollary 3.2} \quad \text{If } r \in A \text{ and } \lim_{k \to \infty} \frac{A_k^{(i)}(r)}{k} = \bar{p}^{(i)}, \quad p^{(i)} < \frac{\mu_i}{\lambda}, \quad i = 1, 2, \]
then
\[
\bar{W}(r) \geq C^{(1)} \bar{Y}^{(1)}(r^{*}(p^{(1)}, a_1)) + C^{(2)} \bar{Y}^{(2)}(r^{*}(1-p^{(1)}, a_2)),
\]
for every constants \( a_1, a_2 \).

From Corollary 3.2 it is clear that if for every \( p \), there were \( \alpha_1, \alpha_2 \) such that \( r^{*}(p, \alpha_1) = r^{*}(1-p, \alpha_2) \), then the optimal routing sequence would be \( r^{*}(p, \alpha^{*}) \) for some \( p^{*} \) and \( \alpha^{*} \).

Unfortunately, this crucial property is not true in general, however it holds for rational \( p \)'s. This, along with the continuity of \( \bar{Y}^{(1)}(r) \), are shown next.

Since \( \bar{Y}^{(i)}(r^{*}(p, \alpha)) \) is independent of \( \alpha \) we may define
\[
\bar{Y}^{*}(p) = \bar{Y}^{(1)}(r^{*}(p, \alpha))
\]
and
\[
\bar{Y}^{*}(p) = \bar{Y}^{(2)}(r^{*}(p, \alpha))
\]
Theorem 3.2 \( J_1^*(p) \) is a non-decreasing continuous and convex function in the interval \([0, \min(1, \mu_1/\lambda)]\).

Proof: From the symmetry in the definition it is sufficient to prove the theorem for \( J_1^*(p) \).

(Monotonicity). Let \( P_2 \preceq P_1 \). For every \( p \) the sum \( \sum_{i=1}^{n} r_i^*(p) \), telescopes, thus

\[
\lim_{k \to \infty} \frac{1}{k} A_{k+1}^*(r^*(P_2)) = P_2 \preceq P_1
\]

and from Theorem 3.2 \( J_1^*(P_2) \preceq J_1^*(P_1) \).

(Convexity). From Lemma 6.1 in [Ha2]

\[
E(Y_{k+1}^*(r^*)) = J_k(r_1^*, r_2^*, \ldots, r_k^*),
\]

where \( J_k(\ ) \) is a multi-modular function on \( Z^k \).

Let \( J_k(z_1, z_2, \ldots, z_k) \), \((z_1, \ldots, z_k) \in R^k\), be the lower convex envelope of \( J_k(\ ) \).

\( J_k(\ ) \) is a convex function on \( R^k \) (see [Ha2, Sec. 4]), and from [Ha2, Sec. 6]

\[
J_1^*(p) = \min_{k \to \infty} J_k^*(z),
\]

where \( z = (p, p, \ldots, p) \in R^k \).

Thus, \( J_1^*(p) \) is also convex.

(Continuity). From [Rock, Corollary 10.1, d], every finite convex function on \( k \) is continuous.
From Corollary 3.2 and Theorem 3.2

\[
\inf_{r \in \mathcal{A}} \bar{W}(r) \geq \min \{ C^*(1)(p) + C^*(2)(1-p) \}. \quad (3.2)
\]

Let \( p^* \) be the value for which the minimum in (3.2) is obtained.

We shall show below that the lower bound in (3.2) can be approached 'as close' as we please by a periodic routing sequence \( r^*(p,0) \).

Lemma 3.1 Let \( p = \ell/k \), where \( \ell \) and \( k \) are relatively prime integers, \( 0 < \ell/k < 1 \),
and let

\[
\alpha = \ell^{-1} \text{ mod } k.
\]

Then,

\[
\lfloor (n+1)p \rfloor - \lfloor np \rfloor = 1 \quad \text{if and only if} \quad \lfloor (n+1+\alpha)(1-p) \rfloor - \lfloor (n+\alpha)(1-p) \rfloor = 0.
\]

(The sequence \( r^*(1-p,0) \) shifted \( r' \) places left is the complement of the sequence \( r^*(p,0) \).)
Proof*: It is sufficient to prove only one direction, since the other direction immediately follows by counting the total number of ones.

Suppose that \( [(n+1)p] - [np] = 1 \). Then there is an integer \( a \) such that \( np < a \leq np + p \).

Let \( x = a - np \). Then \( x = i/k \), where \( i \) is an integer \( 1 \leq i \leq \ell \).

Assume by contradiction that

\[
\lfloor (n+1+a)(1-p) \rfloor - \lfloor (n+a)(1-p) \rfloor = 1.
\]

Then there is an integer \( b \) such that

\[
(n+a)(1-p) < b \leq (n+1+a)(1-p).
\]

Since \( np = a - x \) we have.

\[
-xp + x < b + a - n - a \leq \frac{i-a}{k} + 1 - p + x.
\]

Now, since \( b+a-n-a \) is an integer, by replacing \( x \) and \( p \) in \( (3.3) \) with \( i/k \) and \( \ell/k \) respectively, we obtain

\[
\frac{i-a}{k} < c \leq \frac{i-a}{k} + \frac{k-\ell}{k},
\]

where \( c \) is an integer.

This implies that

\[
\frac{i-a}{k} - \lfloor \frac{i-a}{k} \rfloor \leq \frac{k-\ell}{k} = \frac{\ell}{k}.
\]

But on the other hand, since \( 1 \leq i \leq \ell \),

\[
\frac{i-a}{k} - \lfloor \frac{i-a}{k} \rfloor = \frac{(i-\ell) \mod k}{k} = \frac{i-1}{k} < \frac{\ell}{k},
\]

which is a contradiction. \( \square \)

---

* The lemma was suggested by the author and proved by Gadiel Seroussi.
Theorem 3.3  For every $\epsilon > 0$, there is a periodic routing sequence $r^*(p,0)$, which is $\epsilon$-optimal in $A$.

Proof: Since the rationals are dense within the real numbers, the theorem follows directly from Lemma 3.1 and the continuity of $J^*_i(p)$. The periodicity is a consequence of $p$ being rational.

The only thing left now, is to determine the $\epsilon$-optimal value $p$. In the next section, we express $\tilde{y}^{(i)}(r)$ for every periodic sequence. Hence, $J^*_1(p)$ can be evaluated for every rational $p$. Since $J^*_1(p)$ is monotonic, an effective search can be conducted for determining the $\epsilon$-optimal $p$. 
4. THE GOLDEN RATIO ROUTING

Let \( p^{(i)}_i, 1 \leq i \leq n, \sum_{i=1}^{n} p^{(i)} = 1 \) be any desirable routing proportions (e.g., the \( p^{(i)*} \) from Section 2).

As can be seen from Section 3, \( \bar{N}^{(i)}(r) \) is minimized when the routings to channel \( i \) are distributed in almost equal distances. Equal distances for all the channels are almost never feasible (see Section 2).

We define a routing policy, the Golden Ratio policy which attempts to distribute the routings to each channel \( i \), in almost equal distances.

Let \( N \) and \( N^{(i)}_i, 1 \leq i \leq n \), be given integers such that

\[
\lfloor p^{(i)}_i N \rfloor \leq N^{(i)}_i \leq \lceil p^{(i)}_i N \rceil,
\]

and

\[
\sum_{i=1}^{n} N^{(i)}_i = N,
\]

where \( \lfloor x \rfloor \) is the smallest integer greater than or equal to \( x \).

Thus

\[
\lim_{N \to \infty} \frac{N^{(i)}_i}{N} = p^{(i)}.
\]

Let \( \text{frac}(y) = y - \lfloor y \rfloor \), \( a_j = \text{frac}(j\phi^{-1}) \) and

\[
\Omega_N = \{a_j | j = 0, 1, \ldots, N-1\},
\]

where

\[
\phi^{-1} = \frac{\sqrt{5} - 1}{2} \approx 0.6180339887
\]

is the Golden Ratio.

Consider all the arrivals in periods of \( N \) consecutive messages each. The \( t \)-th smallest point of \( \Omega_N \) is identified with the \( t \)-th message in each period.
Definition 4.1 The Golden Ratio policy, $\pi_{GR}(N)$, is the policy which assigns to channel $i$ in every period, the messages corresponding to the points

$$\{a_j \mid \sum_{m=1}^{i-1} N(m) \leq j < \sum_{m=1}^{i} N(m)\}.$$  

It will be convenient to identify the points 0 and 1, and thus the points $a_j$ are distributed over a circle $C$.

Example 4.1 Suppose $n = 3$, $p^{(1)} = 1/2 \pm \epsilon_1$, $p^{(2)} = 3/8 \pm \epsilon_2$, $p^{(3)} = 1/8 \pm \epsilon_3$, where $\epsilon_i > 0$ are arbitrarily small and $x^{(1)} + x^{(2)} + x^{(3)} = 1$.

Taking $N = 8$, $N^{(1)} = 4$, $N^{(2)} = 3$, and $N^{(3)} = 1$, $\pi_{GR}(8)$ routes to channel 1 the messages corresponding to $6\cdot \phi^{-1}$, $\frac{2}{\phi}$, and $\frac{3}{\phi}$; to channel 2 the messages corresponding to $\frac{4}{\phi}$, $\frac{5}{\phi}$, and $\frac{6}{\phi}$; and to channel 3 the messages corresponding to $\frac{7}{\phi}$. Thus, the Golden Ratio policy keeps routing every $n$ consecutive messages to the channels in the following cyclic order: "1,2,1,3,2,1,2,1".

Remark 4.1 When $\mu_i = \mu$, $1 \leq i \leq n$, the desirable routing proportions from Section 2, are $p^{(i)} = 1/n$. Taking $N = n$ and $N^{(i)} = 1$, implies that $\pi_{GR}(N)$ is the same as the Round-Robin policy, which is also optimal.

Remark 4.2 In [ItRo] it was shown that if $N$ is a Fibonacci number, the Golden Ratio policy $\pi_{GR}(N)$ distributes the routings to each channel $i$, $1 \leq i \leq n$, in at most three different distances and the distances are uniformly mixed.
To evaluate $\tilde{V}(r)$ for a golden ratio sequence $r$ (a sequence generated by a Golden Ratio policy) we need to find $\bar{N}^{(i)}(r)$.

By the method of phases (see [Ne, Ch.1]) the process 
\[ \{N_k^{(i)}(r) | k \geq 1\} \]

can be imbedded into a countable state time homogeneous Markov-chain and $\bar{N}^{(i)}(r)$ can be expressed as follows.

(Actually, the method of phases is suitable for every periodic routing sequence.)

Let $N_i, N^{(i)}_i, 1 \leq i \leq n$, be given and let $d_j^{(i)}, 0 \leq j \leq N^{(i)}_i - 1$ be the number of arrivals between two successive routings to channel $i$ in a given period (including the message which is routed next to the channel). Clearly, the sets $\{d_j^{(i)} | 0 \leq j \leq N^{(i)}_i - 1\}, 1 \leq i \leq n$, are uniquely determined by $\pi_{GR(N)}$ and they are the same for every period.

The states of channel $i$ between successive routings will be referred as phases and be denoted by $j, 0 \leq j \leq N^{(i)}_i - 1$.

Let $a^{(i)}_{j,[j+1]}(k)$ be the probability that $k$ messages are transmitted by channel $i$ during phase $j$, given that the channel is busy during the entire phase. ($[j+1] = (j+1) \text{ mod } N^{(i)}_i$.)

Since the time required to observe $k$ arrivals from a Poisson process has an Erlang distribution

\[
a^{(i)}_{j,[j+1]}(k) = \sum_{x=0}^{k} \frac{\mu_i x (\mu_i x)^k}{k!} \frac{d_j^{(i)} d_{j-1}^{(i)}}{(d_{j-1}^{(i)})!} e^{-\lambda x} dx = \left( \frac{k + d_j^{(i)} - 1}{\lambda} \right)^k \left( \frac{\mu_i}{\lambda} \right)^k e^{-k \lambda} \frac{d_j^{(i)}}{(d_j^{(i)} - 1)!}
\]

for $0 \leq j \leq N^{(i)}_i - 1, k \geq 0$.
Also, let

\[ b_{i,j}^{(i)}[j+1]^{(k)} = 1 - \sum_{l=0}^{k} a_{i,j}^{(i)}[j+1]^{(k)} \]

\[ A_i(k) = \begin{bmatrix} 0 & a_{1,2}^{(i)}(k) & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & a_{n-1,n}^{(i)}(k) & 0 & 0 \\ a_{n,1}^{(i)}(k) & 0 & \cdots & \cdots & 0 \end{bmatrix} \]

\[ B_i(k) = \begin{bmatrix} 0 & b_{1,2}^{(i)}(k) & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & b_{n-1,n}^{(i)}(k) & 0 & 0 \\ b_{n,1}^{(i)}(k) & 0 & \cdots & \cdots & 0 \end{bmatrix} \]

Let \( R_i \) be the unique matrix which solves

\[ R_i = \sum_{k=0}^{\infty} R_i^k A_i(k) \quad (4.1) \]

and let

\[ B(R_i) = \sum_{k=0}^{\infty} R_i^k B_i(k) \]

(See [Ne, Ch. 1] for uniqueness.)
From [Ne, Sec. 1.8]

\[
\hat{R}^{(i)}_t = \xi_i R_1 (I - R_1)^{-2} e ,
\]

(4.2)

where \( e \) is the column vector with all its components equal to one, and \( \xi_i \) is the positive left-invariant eigenvector of \( B_1(R) \), that is,

\[
\xi_i = \xi_i B_1(R),
\]

normalized by

\[
\xi_i (I - R_1)^{-1} e = 1 .
\]

Remark 4.3: The matrix \( R_1 \) can be efficiently computed by successive substitutions. For further computational procedures, see [Ne, Sec. 1.9].

For a given system \( \{\lambda, \mu_1, \mathcal{C}^{(i)}_1|1 \leq i \leq n\} \) let \( LB \) be the lower bound to \( \hat{V} \) as found in Section 2, and \( \hat{V}(GR) \) be the cost of using the Golden Ratio policy. To evaluate the quality of \( \pi_{GR} \), we define the performance measure

\[
f = \frac{V(GR) - LB}{LB} .
\]

We tabulate below (Tables 1-4) the values \( f \) for various systems with \( n = 2, 5, 15, 30 \) and \( \mathcal{C}^{(i)}_1 = 1 \). Although there is a large variety of systems, we can restrict ourselves to a small descriptive set.

First, we may assume without loss of generality that \( \lambda = 1 \), since \( \hat{N}^{(i)}(r), 1 \leq i \leq n \), are not changed by transforming the original system into the system \( \{1, \frac{\mu_1}{\lambda}, \mathcal{C}^{(i)}|1 \leq i \leq n\} \). Further, from a large number of examples which we considered,
it was found that for a given $n$, there are two main parameters which affect $f$:

$$
\mu_1(1) = \frac{1}{n} \sum_{i=1}^{n} \mu_i
$$
and

$$
\mu_2(2) = \frac{1}{n} \sum_{i=1}^{n} \mu_i^2
$$

$\mu_1(1)$ determines the total load at the channels and $\mu_2(2)$ determines the variation among the rates of the channels.

In the tables below, the vectors $(\mu_1, \ldots, \mu_n)$ are not given explicitly, since there is no significant difference between vectors with the same $\mu_1(1)$ and $\mu_2(2)$.

Instead of $\mu_1(1)$ we use a more common parameter,

$$
\rho = \frac{\lambda}{n \mu_1(1)} = \frac{1}{n \mu_1(1)}.
$$

Finally, for every $n$ we evaluate $\pi_{GR}$ for $\rho = 0.35, 0.50, 0.75$ and $\mu_2(2)$ which varies from the smallest to the largest possible value.
<table>
<thead>
<tr>
<th>$\rho$</th>
<th>$\mu^{(2)}$</th>
<th>$L_B$</th>
<th>$\bar{V}(GR)$</th>
<th>$f$</th>
<th>Period length-N</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.040</td>
<td>.582</td>
<td>.582</td>
<td>0</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>2.122</td>
<td>.580</td>
<td>.626</td>
<td>.078</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>2.367</td>
<td>.577</td>
<td>.620</td>
<td>.075</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>2.775</td>
<td>.570</td>
<td>.614</td>
<td>.076</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>3.347</td>
<td>.560</td>
<td>.587</td>
<td>.047</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>1.000</td>
<td>1.236</td>
<td>1.236</td>
<td>0</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>1.04</td>
<td>1.229</td>
<td>1.275</td>
<td>.037</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>.50</td>
<td>1.16</td>
<td>1.252</td>
<td>.036</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>1.36</td>
<td>1.168</td>
<td>1.214</td>
<td>.039</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>1.64</td>
<td>1.102</td>
<td>1.133</td>
<td>.027</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>.444</td>
<td>4.194</td>
<td>4.194</td>
<td>0</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>.462</td>
<td>4.171</td>
<td>4.234</td>
<td>.015</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>.515</td>
<td>4.095</td>
<td>4.145</td>
<td>.012</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>.604</td>
<td>3.945</td>
<td>4.004</td>
<td>.015</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>.729</td>
<td>3.655</td>
<td>3.693</td>
<td>.010</td>
<td>14</td>
<td></td>
</tr>
</tbody>
</table>

Table 1* - Performance values for $n = 2$

* All numbers are rounded up to the third decimal position.
<table>
<thead>
<tr>
<th>$p$</th>
<th>$\mu(2)$</th>
<th>LB</th>
<th>$\bar{V}$(GR)</th>
<th>$f$</th>
<th>$N$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.326</td>
<td>0.773</td>
<td>0.773</td>
<td>0</td>
<td>5</td>
<td>15</td>
</tr>
<tr>
<td>0.351</td>
<td>0.761</td>
<td>0.831</td>
<td>0.091</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.417</td>
<td>0.723</td>
<td>0.813</td>
<td>0.125</td>
<td>28</td>
<td>64</td>
</tr>
<tr>
<td>0.50</td>
<td>0.581</td>
<td>0.713</td>
<td>0.097</td>
<td>6</td>
<td>15</td>
</tr>
<tr>
<td>0.584</td>
<td>0.628</td>
<td>0.686</td>
<td>0.060</td>
<td>64</td>
<td></td>
</tr>
<tr>
<td>1.106</td>
<td>0.591</td>
<td>0.667</td>
<td>0.128</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>1.474</td>
<td>0.558</td>
<td>0.583</td>
<td>0.045</td>
<td>64</td>
<td></td>
</tr>
<tr>
<td>0.160</td>
<td>1.981</td>
<td>1.981</td>
<td>0</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>0.172</td>
<td>1.940</td>
<td>2.017</td>
<td>0.039</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>0.205</td>
<td>1.795</td>
<td>1.913</td>
<td>0.066</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>0.50</td>
<td>0.285</td>
<td>1.508</td>
<td>1.609</td>
<td>0.067</td>
<td>10</td>
</tr>
<tr>
<td>0.335</td>
<td>1.507</td>
<td>1.546</td>
<td>0.099</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>0.542</td>
<td>1.231</td>
<td>1.347</td>
<td>0.094</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>0.722</td>
<td>1.070</td>
<td>1.111</td>
<td>0.039</td>
<td>64</td>
<td></td>
</tr>
<tr>
<td>0.071</td>
<td>7.790</td>
<td>7.790</td>
<td>0</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>0.076</td>
<td>7.662</td>
<td>7.947</td>
<td>0.037</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>0.091</td>
<td>7.133</td>
<td>7.420</td>
<td>0.040</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>0.75</td>
<td>0.126</td>
<td>5.714</td>
<td>5.995</td>
<td>0.049</td>
<td>10</td>
</tr>
<tr>
<td>0.149</td>
<td>5.393</td>
<td>5.711</td>
<td>0.059</td>
<td>56</td>
<td></td>
</tr>
<tr>
<td>0.241</td>
<td>4.552</td>
<td>4.847</td>
<td>0.065</td>
<td>38</td>
<td></td>
</tr>
<tr>
<td>0.321</td>
<td>3.518</td>
<td>3.750</td>
<td>0.066</td>
<td>64</td>
<td></td>
</tr>
</tbody>
</table>

Table 2 - Performance values for $n = 5$
<table>
<thead>
<tr>
<th>$p$</th>
<th>$\mu$</th>
<th>$LB$</th>
<th>$\bar{V}(GR)$</th>
<th>$f$</th>
<th>$N$</th>
</tr>
</thead>
<tbody>
<tr>
<td>.036</td>
<td>.039</td>
<td>.040</td>
<td>.059</td>
<td>.076</td>
<td>.123</td>
</tr>
<tr>
<td>1.505</td>
<td>1.452</td>
<td>1.305</td>
<td>1.059</td>
<td>.967</td>
<td>.796</td>
</tr>
<tr>
<td>1.505</td>
<td>1.649</td>
<td>1.565</td>
<td>1.228</td>
<td>1.072</td>
<td>.960</td>
</tr>
<tr>
<td>0</td>
<td>.135</td>
<td>.198</td>
<td>.159</td>
<td>.108</td>
<td>.207</td>
</tr>
<tr>
<td>15</td>
<td>45</td>
<td>33</td>
<td>18</td>
<td>216</td>
<td>42</td>
</tr>
<tr>
<td>.018</td>
<td>.019</td>
<td>.023</td>
<td>.032</td>
<td>.037</td>
<td>.060</td>
</tr>
<tr>
<td>4.518</td>
<td>4.363</td>
<td>3.874</td>
<td>3.003</td>
<td>2.662</td>
<td>2.047</td>
</tr>
<tr>
<td>4.518</td>
<td>4.602</td>
<td>4.323</td>
<td>3.287</td>
<td>2.922</td>
<td>2.281</td>
</tr>
<tr>
<td>0</td>
<td>.055</td>
<td>.116</td>
<td>.095</td>
<td>.098</td>
<td>.114</td>
</tr>
<tr>
<td>15</td>
<td>45</td>
<td>96</td>
<td>18</td>
<td>6</td>
<td>42</td>
</tr>
<tr>
<td>.0079</td>
<td>.0085</td>
<td>.0101</td>
<td>.0140</td>
<td>.0165</td>
<td>.0268</td>
</tr>
<tr>
<td>0</td>
<td>.056</td>
<td>.042</td>
<td>.057</td>
<td>.030</td>
<td>.081</td>
</tr>
<tr>
<td>15</td>
<td>60</td>
<td>81</td>
<td>30</td>
<td>216</td>
<td>123</td>
</tr>
</tbody>
</table>

Table 3 - Performance values for $n = 15$
<table>
<thead>
<tr>
<th>$\rho$</th>
<th>$\mu(2)$</th>
<th>$\bar{v}_B$</th>
<th>$\bar{v}(GR)$</th>
<th>$f$</th>
<th>$N$</th>
</tr>
</thead>
<tbody>
<tr>
<td>.0091</td>
<td>2.639</td>
<td>2.639</td>
<td>0</td>
<td></td>
<td>30</td>
</tr>
<tr>
<td>.0097</td>
<td>2.512</td>
<td>2.890</td>
<td>.150</td>
<td></td>
<td>36</td>
</tr>
<tr>
<td>.0161</td>
<td>1.705</td>
<td>1.977</td>
<td>.160</td>
<td></td>
<td>36</td>
</tr>
<tr>
<td>.0190</td>
<td>1.513</td>
<td>1.692</td>
<td>.118</td>
<td></td>
<td>12</td>
</tr>
<tr>
<td>.0307</td>
<td>1.136</td>
<td>1.416</td>
<td>.247</td>
<td></td>
<td>78</td>
</tr>
<tr>
<td>.0614</td>
<td>.764</td>
<td>.798</td>
<td>.045</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>.0672</td>
<td>.751</td>
<td>.881</td>
<td>.174</td>
<td></td>
<td>15</td>
</tr>
<tr>
<td>.1293</td>
<td>.600</td>
<td>.611</td>
<td>.018</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>.0444</td>
<td>8.339</td>
<td>8.339</td>
<td>0</td>
<td></td>
<td>30</td>
</tr>
<tr>
<td>.0048</td>
<td>8.009</td>
<td>8.530</td>
<td>.065</td>
<td></td>
<td>90</td>
</tr>
<tr>
<td>.0057</td>
<td>7.003</td>
<td>7.835</td>
<td>.119</td>
<td></td>
<td>192</td>
</tr>
<tr>
<td>.0079</td>
<td>5.257</td>
<td>5.761</td>
<td>.096</td>
<td></td>
<td>36</td>
</tr>
<tr>
<td>.0093</td>
<td>4.563</td>
<td>5.009</td>
<td>.098</td>
<td></td>
<td>12</td>
</tr>
<tr>
<td>.0199</td>
<td>2.700</td>
<td>3.094</td>
<td>.145</td>
<td></td>
<td>210</td>
</tr>
<tr>
<td>.0302</td>
<td>1.920</td>
<td>1.989</td>
<td>.035</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>.0634</td>
<td>1.294</td>
<td>1.312</td>
<td>.014</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>.0020</td>
<td>37.803</td>
<td>37.803</td>
<td>0</td>
<td></td>
<td>30</td>
</tr>
<tr>
<td>.0021</td>
<td>36.862</td>
<td>38.046</td>
<td>.048</td>
<td></td>
<td>90</td>
</tr>
<tr>
<td>.0025</td>
<td>33.281</td>
<td>34.685</td>
<td>.042</td>
<td></td>
<td>162</td>
</tr>
<tr>
<td>.0035</td>
<td>24.614</td>
<td>26.117</td>
<td>.061</td>
<td></td>
<td>60</td>
</tr>
<tr>
<td>.0041</td>
<td>22.078</td>
<td>23.079</td>
<td>.045</td>
<td></td>
<td>444</td>
</tr>
<tr>
<td>.0067</td>
<td>16.388</td>
<td>17.473</td>
<td>.066</td>
<td></td>
<td>312</td>
</tr>
<tr>
<td>.0134</td>
<td>8.019</td>
<td>8.449</td>
<td>.054</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>.0267</td>
<td>5.073</td>
<td>5.314</td>
<td>.047</td>
<td></td>
<td>2</td>
</tr>
</tbody>
</table>

**Table 4** - Performance values for $n = 30$
In several cases when some of the channels have a very low rate $\mu_i$, it is worthwhile not to route any messages to them. Thus, it can be seen in the tables that for large $\mu^{(2)}$, the period length $-N$ is sometimes smaller than the number of the channels $-n$.

As a general conclusion from the examples above, one can say that for $p \geq 0.50$, the relative difference between the Golden Ratio policy and the lower bound mainly varies in the range of 4-10%. Moreover, it becomes smaller as $p$ becomes larger.
REFERENCES


REFERENCES (cont'd)

