

# コンピュータ・ネットワークにおける最適経路制御方式

小松雅治 大内康敬 中西暉 真田英彦 手塚慶一  
(大阪大学 工学部)

## §1 まえがき

コンピュータ・ネットワークをはじめとするメッセージ交換網において、メッセージの経路をいかに制御するかという問題(経路選択)は重要な問題の一つである。経路選択は、回線あるいは交換局の障害及び増設等によって時間的に変化するネットワーク・トポロジーや、時々刻々と変動するトラヒック状況に対して柔軟に対応する適応経路選択方式を用いるのが望ましい。従来、適応経路選択方式においては、ARPAネットワークの経路選択方式に代表されるように、目的局までの伝送遅延時間を基準として、各局で出力回線(経路)を決定するのが一般的である。特に、ARPAの方式では、各交換局において、網状態から予測されるメッセージ(パケット)の目的局までの伝送遅延時間から作成される経路選択表を持ち、予測遅延時間が最小となる経路を選択する。しかし、この方式では、網内のメッセージ(パケット)相互の影響を考慮に入れず、自己本意的な経路の決定がされていると考えられる。

本稿では、システム全体の平均メッセージ遅延時間を最小とするための必要十分条件を与える「最適経路配分定理」から、経路決定のための最適基準を導出し、その基準を用いた適応経路選択方式を構成し、さらに、シミュレーション結果より、本方式とARPAタイプ的方式の比較検討を行っている。

## §2 最適経路配分定理の適応経路選択方式への応用

まず、提案する適応経路選択方式の基礎となる定理について述べる。

### 「最適経路配分定理」

メッセージ交換網において、個々のチャネルに形成される平均待ち行列長 $L_i$ がメッセージの到着率 $\lambda_i$ の凸関数で与えられるとき、網全体の平均メッセージ遅延時間を最小とする経路配分を行うための必要十分条件は次式で与えられる。

ノード $s$ を発信局、ノード $d$ を目的局とするメッセージ $r_{s,d}$ を、ノード $s$ からノード $d$ への $M_{s,d}$ 本の経路 $P_1^{s,d}, P_2^{s,d}, \dots, P_{M_{s,d}}^{s,d}$ に配分率 $x_1^{s,d}, x_2^{s,d}, \dots, x_{M_{s,d}}^{s,d}$ で配分するとき、

$$\left. \begin{aligned} \sum_{i: B_i \in P_j^{s,d}} \frac{\partial L_i}{\partial \lambda_i} &= D_{s,d} && : x_j^{s,d} > 0 \\ \sum_{i: B_i \in P_j^{s,d}} \frac{\partial L_i}{\partial \lambda_i} &\geq D_{s,d} && : x_j^{s,d} = 0 \end{aligned} \right\} \quad (1)$$

が、網内の全 $(s, d)$ 対に対して成り立つことである。但し、 $B_i$ は、網を構成する各チャネルを示し、 $D_{s,d}$ は、 $(s, d)$ 対毎の定数であり、 $\sum_{j=1}^{M_{s,d}} x_j^{s,d} = 1, x_j^{s,d} \geq 0$ である。

(証明略)

上述の定理によれば、発信局から目的局までの経路に沿っての各チャネルの  $L/\lambda$  の和が最小である経路を選んでメッセージを送ることによって、平均メッセージ遅延時間を最小とすることが可能である。このことを、さらに、定量的に説明する。

メッセージ交換網において、網全体の平均メッセージ遅延時間  $T$  は

$$T = \frac{1}{\gamma} \sum \lambda_i T_i = \frac{1}{\gamma} \sum L_i \quad (2)$$

で与えられる。但し、 $\gamma = \sum \lambda_i$  である。そこで、図1に示すようなノード  $N_1$  からノード  $N_2$  への経路が  $P_1$  と  $P_2$  の2本があるネットワークを考えると、平均メッセージ遅延時間は

$$T = \frac{1}{\gamma} \left( \sum_{i \in P_1} L_i(\lambda_i) + \sum_{i \in P_2} L_i(\lambda_i) + \sum_{i \in P_1, P_2} L_i(\lambda_i) \right) \quad (3)$$

で与えられる。今、ノード  $N_1$  から  $N_2$  へのメッセージが微小量  $\Delta\lambda$  だけ増加したものとする。この増分メッセージ  $\Delta\lambda$  を  $P_1$  及び  $P_2$  で伝送した場合の平均メッセージ遅延時間を、それぞれ  $T_1$ ,  $T_2$  とすると

$$T_1 = \frac{1}{\gamma + \Delta\lambda} \left[ \sum_{i \in P_1} \left( L_i(\lambda_i) + \frac{\partial L_i}{\partial \lambda} \Big|_{\lambda_i} \cdot \Delta\lambda \right) + \sum_{i \in P_2} L_i(\lambda_i) + \sum_{i \in P_1, P_2} L_i(\lambda_i) \right] \quad (4)$$

$$T_2 = \frac{1}{\gamma + \Delta\lambda} \left[ \sum_{i \in P_1} L_i(\lambda_i) + \sum_{i \in P_2} \left( L_i(\lambda_i) + \frac{\partial L_i}{\partial \lambda} \Big|_{\lambda_i} \cdot \Delta\lambda \right) + \sum_{i \in P_1, P_2} L_i(\lambda_i) \right] \quad (5)$$

となる。従って、

$$\left. \begin{aligned} \sum_{i \in P_1} \frac{\partial L_i}{\partial \lambda} \Big|_{\lambda_i} < \sum_{i \in P_2} \frac{\partial L_i}{\partial \lambda} \Big|_{\lambda_i} & \quad \text{ならば} \quad T_1 < T_2 \\ \sum_{i \in P_1} \frac{\partial L_i}{\partial \lambda} \Big|_{\lambda_i} > \sum_{i \in P_2} \frac{\partial L_i}{\partial \lambda} \Big|_{\lambda_i} & \quad \text{ならば} \quad T_1 > T_2 \end{aligned} \right\} \quad (6)$$

となる。とらば、経路に沿っての各チャネルでの遅延時間の和、あるいは、 $\sum_{i \in P_1} T_i$  と  $\sum_{i \in P_2} T_i$  の大小関係によって、増分メッセージの経路を定めると網全体の平均メッセージ遅延時間を最小にできることは限らない。なぜなら、

$$\sum_{i \in P_1} T_i < \sum_{i \in P_2} T_i \quad \text{ならば} \quad \sum_{i \in P_1} \frac{\partial L_i}{\partial \lambda} \Big|_{\lambda_i} < \sum_{i \in P_2} \frac{\partial L_i}{\partial \lambda} \Big|_{\lambda_i}$$

が成り立つとは限らないからである。

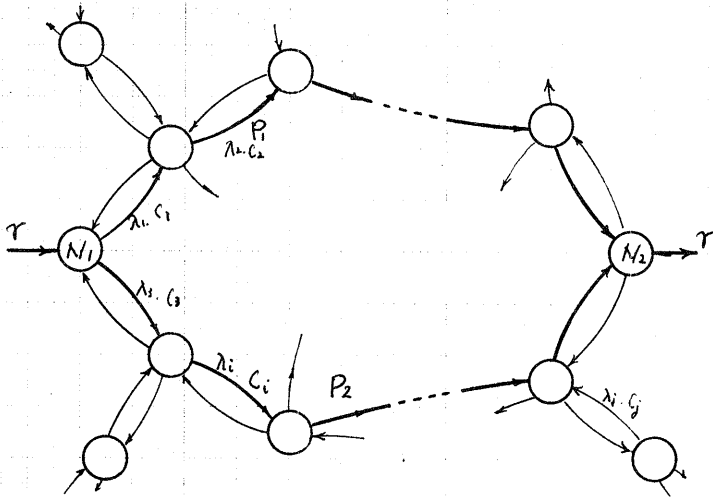


図1 2経路モデル

次に、チャネル毎の待ち行列長が  $l$  の場合の  $2L/2\lambda$  に対応する量  $Z$  について考察する。この量を  $Z(l)$  とし、待ち行列長が  $l$  である状態確率を  $P(l)$  とすると、

$$\frac{\partial L}{\partial \lambda} = \sum_{l=0}^{\infty} Z(l) P(l) \quad (7)$$

が成り立つものとする。これによって求められた  $Z(l)$  を、経路選択の一つの基準として用いることが可能である。たとえば、図1で、ノード  $N_1$  から  $N_2$  への経路  $P_1$  と  $P_2$  において中継ノードが無い場合、Smallest  $Z(l)$  方式 (Shortest queue 方式に対応する) に従って、メッセージを伝送すると、定常状態で、2つのチャネルにおける  $2L/2\lambda$  が等しくなることが期待される。

特に、メッセージ長が指数分布に従う場合の  $Z(l)$  を導出する。(シミュレーションでは、メッセージ長の指数性を仮定している) この時、各チャネルは、 $M/M/1$  待ち行列システムで記述でき、平均待ち行列長  $L$  及び状態確率  $P(l)$  は次式で与えられる。

$$L = \frac{\rho}{1-\rho} \quad (8)$$

$$P(l) = (1-\rho) \rho^l \quad (9)$$

但し、

$$\rho = \frac{\lambda}{\mu C}$$

である。従って、式(7)は、

$$\frac{1}{\mu C (1-\rho)^2} = \sum_{l=0}^{\infty} Z(l) (1-\rho) \rho^l \quad (10)$$

となり、式(10)より

$$Z(l) = \frac{1}{2\mu C} (l+1)(l+2) \quad (11)$$

が得られる。

次に、経路選択の基準として  $Z(l)$  及び  $2L/2\lambda$  を用いた適応経路選択方式と、ARPAタイプの予測遅延時間を用いる方式の構成・機能上の比較を行う。

	定理に基づく方式	ARPAタイプ的方式
ルーチンリ基準	$Z(l), 2L/2\lambda$	遅延時間
各交換局 $i$ で保持するテーブル	(1) $R(i, J, K)$ 交換局 $i$ で、出力回線 $J$ を選択したとき、 $J$ の着局 $L(i, J)$ から目的局 $K$ までの $2L/2\lambda$ の予測値を与える。	(1) $T(i, J, K)$ 交換局 $i$ で、出力回線 $J$ を選択したとき、 $J$ の着局 $L(i, J)$ から目的局 $K$ までの予測遅延時間を与える。
	(2) $R_{node}(i, J)$ 交換局 $i$ で、出力回線 $J$ を選択したとき、 $J$ での $Z(l)$ の値を与える。	(2) $T_{node}(i, J)$ 交換局 $i$ で、出力回線 $J$ を選択したとき、 $J$ での遅延時間を与える。
	(3) $R_{min}(i, K)$ 交換局 $i$ から目的局 $K$ までの最小予測 $2L/2\lambda$ を与える。	(3) $T_{min}(i, K)$ 交換局 $i$ から目的局 $K$ までの最小予測遅延時間を与える。
出力回線 $J_0(i, K)$ の決定	$J_0(i, K) = [J   \text{Min}_J \{ R(i, J, K) + R_{node}(i, J) \}]$	$J_0(i, K) = [J   \text{Min}_J \{ T(i, J, K) + T_{node}(i, J) \}]$
テーブル(1)の更新	出力回線 $J=1, N$ の着局である隣接局 $L(i, J)$ からテーブル(3)を受信すると $R(i, J, K) = R_{min}(L(i, J), K)$	$T(i, J, K) = T_{min}(L(i, J), K)$

上表の各種記号の関係図を図2に示す。

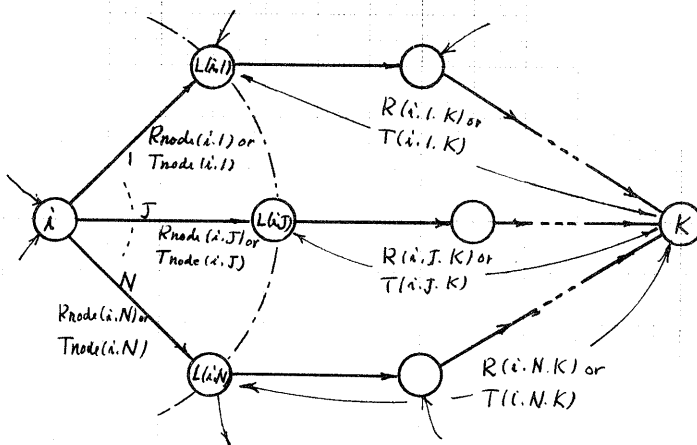


図2. ネットワークとテーブル

上の表から明らかのように、丙方式はルーティング情報の内容が異なるだけである。そのため、経路選択方式に要求されるネットワークトポロジーやトラヒック状況の変化に対する適応性、方式の簡潔性及びコストなどの問題に関して、丙方式は同等であると考えてない。そこで、丙方式の比較は、経路選択方式のパフォーマンスである平均メッセージ遅延時間に関して行えばよい。

### 5.3 シミュレーション結果とその検討

#### 3.1 シミュレーションモデル

(1) ネットワーク形態；6局14回線梯子状ネットワーク（図3）

(2) トラヒック；網外から発信局へのメッセージの到着はポアソン分布に従い、メッセージ長は平均 $1/\mu = 1000 \text{ bits}$ の指数分布に従うものとする。トラヒックマトリクス $[Y_{ij}]$ を式(12)で与える。

$$[Y_{ij}] = \begin{matrix} \begin{pmatrix} 0 & 1.2 & 0.4 & 0.8 & 0.1 & 0.3 \\ 1.2 & 0 & 1.2 & 9.6 & 0.3 & 3.6 \\ 0.4 & 1.2 & 0 & 3.2 & 0.4 & 1.2 \\ 0.8 & 9.6 & 3.2 & 0 & 0.8 & 9.6 \\ 0.1 & 0.3 & 0.4 & 0.8 & 0 & 1.2 \\ 0.3 & 3.6 & 1.2 & 9.6 & 1.2 & 0 \end{pmatrix} & (12) \end{matrix}$$

但し、 $Y_{ij}$ は、発信局をi、一ドーン目的局をjとするメッセージの網外から網(i一ドーン)への平均到着率[個/sec]を示す。

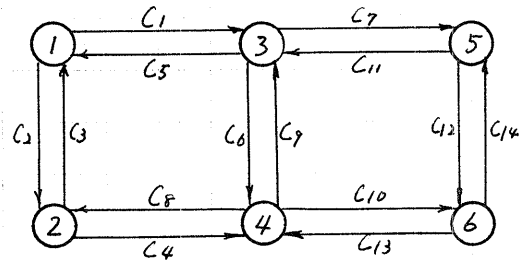


図3. 梯子状ネットワーク

(3) テーブル(2)の決定とテーブル(3)の予測。

これらの決定、予測法によって経路選択方式を次の3方式に分類する。

#### 方式(A)

テーブル(2)は、メッセージ到着時の $Z(l)$ (式(11))によって与え、テーブル(3)の予測は、あるルーティング情報更新間隔内の $Z(l)$ の平均を $Z$ とすることで行う。

#### 方式(B)

テーブル(2)の決定は方式(A)と同じであるが、テーブル(3)については、ル

ルーティング情報更新間隔内の出力回線待ち行列長の平均から、回線利用率を推定し  $\rho/\lambda = 1/\mu C(1-P)^2$  によって与える。

### 方式(C)

テーブル(2)の決定は、メッセージ到着時の出力回線待ち行列長から、その回線での遅延時間を求めることにより行う。テーブル(3)は、ルーティング情報更新間隔内の平均待ち行列長から遅延時間を推定するのを基本として作成される。本方式は、ARPAタイプの経路選択方式である。

なお、これらの3方式において、テーブル(3)の作成にあたり、現在のルーティング情報更新間隔内の推定値と、前回更新間隔内での推定値と、それとそれ、(1- $\beta$ )の重み付け和をもって決定するものとする。(0 <  $\beta$  < 1)

(4) ルーティング情報更新間隔  $T_{ud}$  を一定とした周期更新とする。

## 3.2 シミュレーション結果及び検討

本節の結果は、待ち合せ網シミュレータ R-SSQ<sup>(5)</sup> (Revised-System Simulator for Queueing Network) を用い、シミュレーション時間が 800 sec になつたところで得たものである。なお、各種統計量は、シミュレーション時間 600 sec 程度のところで充分安定した結果を示した\*。

また、3.1で示すような分散制御形の適応経路選択方式では、ループ現象が不可避である。そのため、本シミュレーションでは、ループ現象の中で最も単純なポンポン現象に対する防止策<sup>\*\*</sup>を施している。

### (1) 一様回線容量の場合

14本の回線すべてに、一様に 10 Kbit/sec の回線容量を与える。図4に、3方式の平均メッセージ遅延時間特性を示す。なお、定線は、式(1)を満足するトラヒック配分を行う場合の遅延時間の理論値である。同図より、 $\gamma=0.5$  程度までは、3方式ともすべて同じ特性を示し、ほぼ理論値と一致する。さらに負荷が増加するに従って方式(A)(B)は(C)よりも劣化する。この原因は、表1の伝送段数の統計値からある程度推測できる。伝送段数の平均及び分散は、方式(A)(B)ともに(C)より大きく、その傾向は、負荷が大きくなる程大きくなる。また、最大伝送段数の値は、ループ現象の存在を裏付けている。仮に、発信局と目的局間の最大伝送段数は、ノード1と6、あるいは2と5の間の5段であるからである。いずれの方式を用いても、ループ現象は不可避であり、その対策を施すことが必要であるが、方式(A)(B)は、ARPAタイプの方式(C)に比べてループ現象を起しやすいことがわかる。それにもかかわらず、適当な負荷条件において、(A)(B)の平均メッセージ遅延時間が(C)と同等であるということは、混雑地域を迂回かうまく行われていることを示している。しかし、負荷

\* 仮に、平均メッセージ遅延時間の平均をシミュレーション時間毎にすると、シミュレーション時間の小さいところでは、時間の経過とともに変動するが 600 sec 位になるとほぼ一定値をとる。  
\*\* ポンポン現象は、隣接ノード間でメッセージが行き来する伝送段数2段のループ現象であり、防止策として、メッセージにフラッグを立てることによって一段を元のノードに戻すのを防ぐようにしている。

がさらに増加するに従って、その効果もループ現象による特性劣化を改善するに至らない。なお方式(A)と(B)とでは、(B)の方が優っていることが分る。また、ループ現象対策を施すことによつて、経路の安定性を増せば、両方式ともよりよく改善され、方式(C)と対等あるいは、より優れた特性を示すことが示唆された。

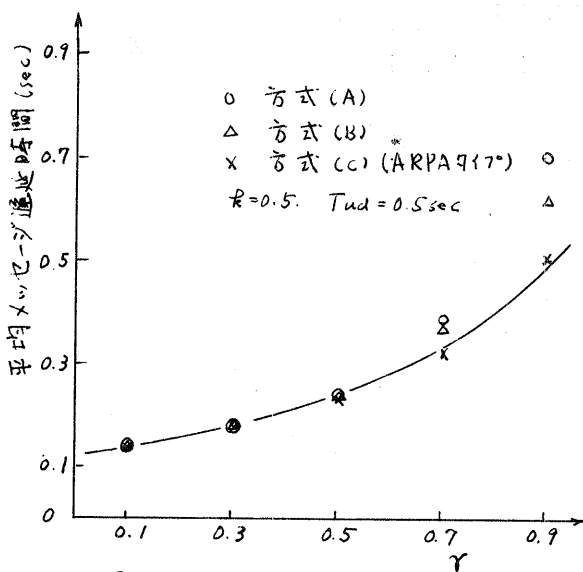


図4. 平均メッセージ遅延時間  
(一様回線容量の場合)

表1 伝送段数の統計

方式 \ $\gamma$		0.1	0.3	0.5	0.7	0.9
(A)	平均	1.280	1.356	1.522	1.763	1.918
	分散	0.270	0.451	1.027	1.375	1.634
	最大	6	9	65	17	14
(B)	平均	1.285	1.411	1.586	1.727	1.875
	分散	0.309	0.607	5.260	1.337	1.612
	最大	13	16	306	32	21
(C)	平均	1.266	1.316	1.414	1.542	1.661
	分散	0.237	0.341	0.573	0.826	1.058
	最大	4	7	23	11	12

(2) 非一様回線容量の場合

(1)の一様回線容量の場合、適当な負荷条件( $\gamma=0.5$ 程度)で、回線4, 8, 10, 13の回線利用率が、他の回線の利用率の約3倍に及び、 $R$ ので、これらの回線容量を他の回線の3倍、すなわち30 Kbit/secとして、シミュレートした結果を示す。図5に、平均メッセージ遅延時間の特性を示す。なお、実線は、式(1)を満足するトラヒック配分を行った場合の遅延時間の理論値を示す。同図より、方式(A)(B)の平均メッセージ遅延時間特性は、(C)に比べて極めて良いことが分る。さらに、(B)は(A)に比べて極めてわずかながら良い。これは、非一様回線容量の場合、 $2L/\lambda$ の基準が、 $T$ なる基準にくらべ効果的は厳格なものである。すなわち、 $C \leq C'$  のとき、 $Z(L) > Z'(L)$  あるいは  $2L/\lambda > 2L'/\lambda$  であっても、必ずしも  $T > T'$  とはならないからである。さらに、表2に、伝送段数の統計量を示す。非一様回線容量の場合にも、一様回線容量の場合と同様、伝送段数の平均、分散は  $\gamma=0.5$  位までは、方式(A)(B)の方が大きく、それ以上になると小さくなる傾向にある。ループ現象については、やはり(A)(B)の方が起しやすくなるように見える。

なお、 $\lambda$  の変化\*に対して  
 いずれの方式もほとんど  
 変化がなかったことも明  
 らかとなった。また、ル  
 ーティング情報更新間隔と  
 0.1 sec から 2.0 sec の  
 間に変化させても、遅延  
 特性の変化はほとんど無  
 かった。

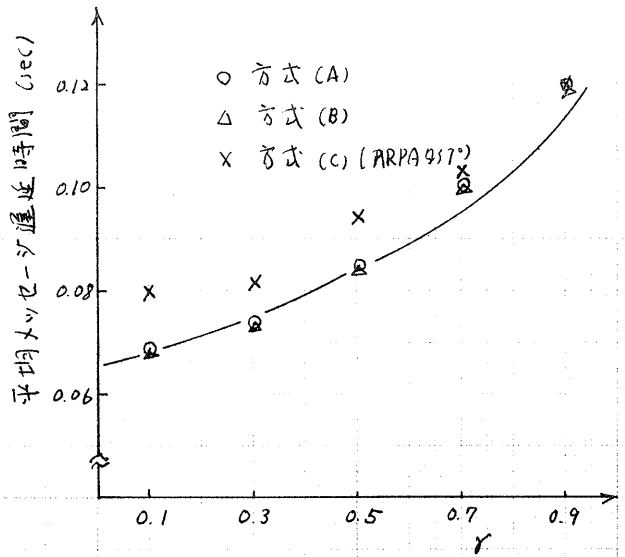


図5 平均メッセージ遅延時間  
 (非一様回線容量の場合)

表2 伝送段数の統計

方式 \ $\rho$		伝送段数				
		0.1	0.3	0.5	0.7	0.9
(A)	平均	1.308	1.303	1.325	1.359	1.363
	分散	0.275	0.348	0.459	2.657	0.939
	最大	5	6	21	273	110
(B)	平均	1.267	1.285	1.301	1.347	1.379
	分散	0.245	0.323	0.395	0.685	0.575
	最大	6	21	27	93	41
(C)	平均	1.253	1.272	1.301	1.371	1.428
	分散	0.233	0.255	0.447	0.848	0.792
	最大	3	5	49	97	27

#### 5.4 まとめ

本稿では、平均メッセージ遅延時間を最小とするための必要十分条件を与える「最道経路配分定理」から、経路選択に用いる最適基準  $Z(\lambda)$  あるいは  $Z(\lambda)/\lambda$  \* 特に、 $\lambda = 0$  の場合、方式 (C) は、よく知られた Shortest Queue + bias 方式となる。バイアス項は、現在局から目的局までの無負荷時の遅延時間 ( $1/\mu_i$  の和) を与える。また、方式 (A) (B) は、Shortest Queue + bias 方式に対応して、Smallest  $Z(\lambda) + bias$  方式ということになる。この方式はバイアス項もまた、Shortest Queue + bias 方式と同じバイアス項をとることになる。すなわち

$$Z(0) = \frac{1}{\mu_C} \quad \text{あるいは} \quad \frac{\partial Z}{\partial \lambda} \Big|_{\lambda=0} = \frac{1}{\mu_C}$$

となるからである。



を求め、これらを用いた分散制御形の適応経路選択方式を構成し、シミュレーション結果より、ARPAタイプの遅延時間を基準に用いた従来の方式との比較検討を行った。その結果、本方式は、網内の中継局工でのメッセージ相互間の影響を考慮に入れた経路選択が可能であることが明らかとなった。一方、本方式はルーティング現象を起し易い欠点をもつ。しかし、この欠点は分散制御形の方式に共通する欠点であり、従来の方式も例外ではない。本方式は、ルーティング対策を施したり、ルーティング情報更新間隔を非同期にしたり、あるいは、自局のルーティング情報  $R_{node}$  を平均値で与えて経路の安定性を増すことにより、(2)あるいは、(3)の効果が十分に発揮され、従来の方式(遅延時間を基準とする)より、さらに改善されることか期待できる。

#### [謝辞]

御討論下さった大阪大学手塚研の諸氏に感謝する。

#### [参考文献]

- (1) J. M. McQuillan, "Adaptive Routing Algorithms for Distributed Computer Networks," NTIS, U.S. Department of Commerce, 1974
- (2) Harry Rudin, "On Routing and Delta Routing; A Taxonomy and Performance Comparison of Techniques for Packed-Switched Networks," IEEE TRANS. ON COM., Vol. COM-24, No. 1, Jan., 1976.
- (3) 小松、中西、真田、手塚, "一般サービス分布をもつデータ通信網の最適経路配分," 信学会交授研報, SE 76-38, 1976
- (4) H. Sanada & Y. Tezuka "Route Assignment on Computer Network," Pacific Area Computer Communication Network System Symposium, Aug. 1975.
- (5) 大内, "計算機網におけるシミュレーション技法に関する研究" 大阪大学卒業論文, 1976