

待ち行列網におけるフロー・コントロール

京大 工学部 高橋 豊
重田 信夫
長谷川利治

1. はじめに.

フロー・コントロールとは、コミュニケーション・システム、道路交通網等のように待ち行列網として数学的にモデル化できる多くのシステムにおいて、特定のジョブ・クラスがリソースを独占することを防ぎ、与えられたトラヒックを円滑に処理するためのアルゴリズム体系である。ここでのジョブとは、トランザクション、パケット、メッセージ等を総称している。フロー・コントロールは大別してルーティング・コントロール (Routing Control) と輻輳コントロール (Congestion Control) とからなるが、見方によればネットワーク・トポロジーも含め得るかも知れない。ルーティングコントロールとは、混雑箇所を回避し、何らかの評価基準のもとでネットワーク全体の最適運用を目指し、与えられたトラヒックの経路を制御するものである。一方輻輳コントロール

ルは、過度のトラヒックが、集中した結果、もはやルーティング・コントロールだけでは、システムの望ましい運営ができない場合に有効なものであり、通常何らかの形での入力制限が行われる。

輻輳とは、負荷と遅延との関係で見れば図1のように曲線が漸近線に近づき、急速に増加するところ、あるいは、負荷とスルー・プットの関係では、図2、図3に見るように漸近線に近づき、頭打ちになるところ、又は負荷が増大するとともにスルー・プットが低下し始めるところであり、この現象はパケット交換網、道路交通網等においてしばしば典型的に見受けられる。

以下では、輻輳コントロールに関し、現在までに実際のシステムにおいて用いられている種々の手法を述べ、これ

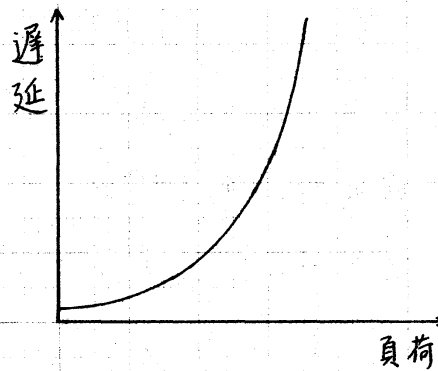


図1. 負荷-遅延

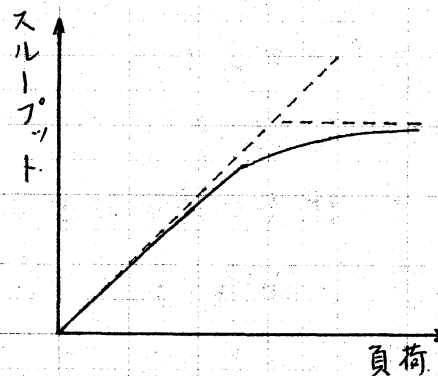


図2. 負荷-スルー・プット

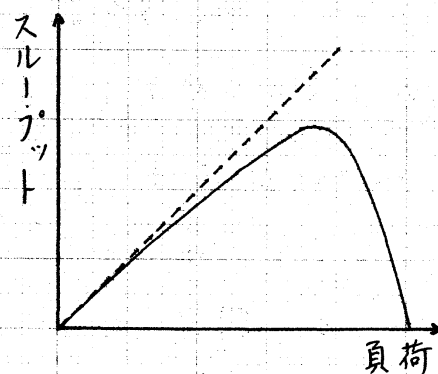


図3. 負荷-スルー・プット

りを複合化した新規の手法を提唱し、これを採用した場合の持性量を近似解析する。

2. 輻輳コントロール

輻輳が発生すると種々の弊害を生じるが、主要なものは、次のようなものである。

- a) 遅延の急激な増加。
- b) スルー・プットの低下。
- c) デッド・ロックへ陥る可能性の増加。

これらを回避するために実際に用いられている手法には、次のようなものがある。

1) local control

せいぜい隣接処理施設までのローカルな情報に基づき入力トラヒックの増減を行う。従って輻輳の原因がはるか遠くの処理施設の入力トラヒックにある場合には無力となる。典型的な例として、各処理施設が定まった時間を超えて連続的に稼働すれば入力を禁止する手法がある。

2) central control

ネットワーク全体のトラヒックを集中的に監視、制御する方法であるが、制御のための情報・指令の伝送のためにトラヒックが増大し、非効率的である。

3) end-to-end control

O-D (origin-destination) の組合せごとに、許容されるジョブ数を予め決めておき、これを超えないように制御する。この方法は、ARPA, GE Information Services Network 等において用いられているものであり、通常閉鎖型待ち行列網として数学的にモデル化し解析することができる。

4) input buffer limiting control

入力ジョブと他処理施設からの通過ジョブを分離し、入力ジョブの待ち行列長に制限を加える。

5) link buffer limiting control

通過ジョブが作る待ち行列長に制限を設け、自処理施設からの入力ジョブもある程度確保する方法である。

6) congestion control table scheme

輻輳制御表がネットワーク内を巡回し、各処理施設は必要とする情報を取り出し、ネットワーク全体の状況を把握する。また自身の状態に関し更新を行う。これらの情報に基づき各処理施設が入力ジョブを制御する方法であり、例えば Cyclades (French Nationwide Computer Network) がある。

7) hop priority technique

通過した処理施設数に応じて、遠くからのジョブほど高い優先順序を付けて制御する。

8) isarithmic method

ネットワーク内に一定数のパーミット (permit) があり、これらはある規則に従いネットワーク内を巡回している。ネットワーク外から到着したジョブは、順次入力待合室における待ち行列に加わり先頭から順次その処理施設において利用可能なパーミットを獲得して目的地に移動する。ジョブがネットワーク外に去ると、パーミットは解放され、回送されるが、この際の規則により固定パーミットと自由パーミットに分けられる。固定パーミット法とは、解放されたパーミットを必ず各々の決められた処理施設に回送するものであり、自由パーミット法とは、何らかの判断に基づき、適宜パーミットを回送する方法であり、例えば待ち行列長の長い処理施設へ、あるいはネットワーク通過時間が短いと予想される処理施設へ回送する方法等がある。

これらのフロー・コントロール手法は、次の様な評価基準のもとで用いられる。

- a. スループット
- b. 遅延
- c. ブロッキング確率
- d. パワー (= スループット / 遅延)

3. 複合輻輳コントロール

Davies [2] 及び Price [6] により提唱された isarithmic control 法は、ネットワーク内のジョブの数により入力を制限するという最初の試みであった。しかしこの手法単独では、輻輳防止に対し、効果を余り期待できない場合もあるために、以下に、input buffer limiting control 法と組み合わせた複合コントロール手法を述べ、解析を行う。

3.1. コントロール手法

図1で示すように、システムを構成する各ステージ（処理施設）は、2つのサブ・ステージすなわちリンク・サブステージとインプット・サブステージとから成り立っている。外部からの入力ジョブは、インプット・サブステージの待ち行列長が制限長に達していない

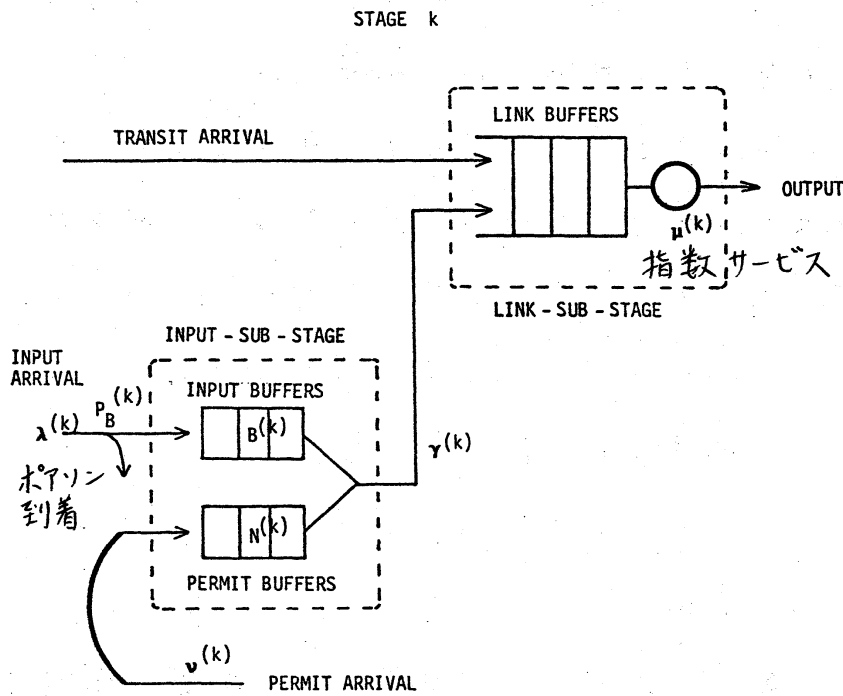


図4 各ステージの構成

いときは、待ち行列に加わることができると、既に制限長に達しているときには損失される。受け入れられたジョブは、待ち行列の先頭から順次空きパーミットを獲得し次第リンク待ち行列に加わり、リンク・サーバーにより処理を受け、これが終了すると各々決まった確率でシステムを去るか、あるいは次のステージへ移動する。ジョブが目的地へ着き、退去すると、それに付随していたパーミットは直ちに元の出発地へ回送されるものとする。またリンク待ち行列に関しては、制限のない場合及びある場合共に取り扱うが、制限のない場合にはパーミットを獲得した入力ジョブあるいは前段のステージからのジョブは、直ちにリンク待ち行列に加わることが出来る。一方制限のある場合には、リンク・バッファ一杯であるとき空き席が出来るまでインプット・バッファあるいは前段のステージで待つことになる。これをブロッキングと呼ぶ。このようにネットワーク内からのジョブがブロッキングされるとサーバーを占有しながら空き席が出来るのを待つことになり処理能力の低下を招く。

3.2. 解析

上記の手法を用いた場合には、厳密に解析することができず、従って有効な近似解法が必要となる。図5に示すMステ

一ジからなる直列型システムに対する近似解法は次のようになる。

主要な記号を次のように定義する。

$\lambda^{(k)}$: 各ステージへの入力ジョブの平均到着率

($k = 1, \dots, M$)

$\gamma^{(k)}$: 各ステージで受け入れられた入力ジョブの平均到着率

$\lambda^{(k)}$: パーミットの平均到着率

$B^{(k)}$: 入力待ち行列の制限長

$N^{(k)}$: 各ステージへ割り当てられたパーミットの数

$\mu^{(k)}$: 各リンク・サーバーの平均処理率

$P_B^{(k)}$: 入力待ち行列が制限長に達している確率

$\{P_{ij}\}$: Origin-Destination 行列

近似解法は、各パーミットが指数分布した時間間隔で戻っ

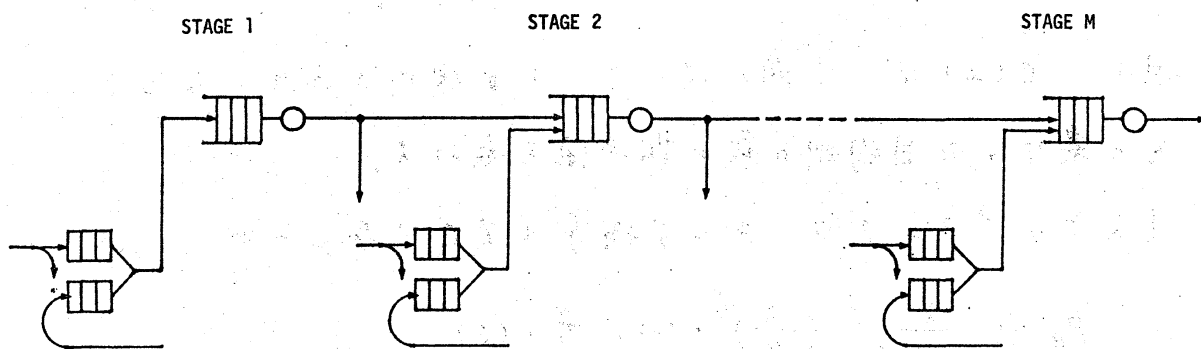


図5 Mステージ直列システム

てくるという仮定と、前段のステージからの到着がポアソン到着であるという仮定に基づいており、下記の様になる。

[ステップ1] パーミットの平均到着率を未知変数 $\nu^{(k)}$ とする。

[ステップ2] インプット・バッファとパーミット・バッファを各ステージごとに独立に取り扱う。各ステージの状態確率を $p(i)$ とし、これは次の様な確率を表わす。

$$p(i) = \text{Prob} \{ (\text{インプット・バッファの待ち行列長}) - (\text{パーミット・バッファの待ち行列長}) + N = i \}$$

従って $(N+B)$ 個の異なる状態があり、各状態確率は次式で与えられる。

$$p(i) = \left(\frac{\lambda}{\nu}\right)^i p(0) / \prod_{\ell=1}^i c(\ell) \quad (0 \leq i \leq N)$$

$$p(N+i) = \left(\frac{\lambda}{c(N)\nu}\right)^i \left(\frac{\lambda}{\nu}\right)^N p(0) / \prod_{\ell=1}^N c(\ell) \quad (0 \leq i \leq B)$$

$$p(0) = \left[\sum_{j=0}^N \left(\frac{\lambda}{\nu}\right)^j / \prod_{\ell=1}^j c(\ell) + \left(\frac{\lambda}{\nu}\right)^N / \prod_{\ell=1}^N c(\ell) \cdot \sum_{j=1}^B \left(\frac{\lambda}{c(N)\nu}\right)^j \right]^{-1}$$

但し、 $c(i)$ は、 i 個のパーミットが使用されているときにその異なる目的地の数の期待値を表わす。

[ステップ3] ブロッキング確率は次式で与えられる

$$p_B = \left(\frac{\lambda}{c(N)\nu}\right)^B \cdot \left(\frac{\lambda}{\nu}\right)^N \cdot p(0) / \prod_{\ell=1}^N c(\ell)$$

これからスルー・プロット $\gamma^{(k)}$ は、次式で与えられる。

$$\gamma^{(k)} = \lambda^{(k)} \cdot (1 - p_B^{(k)})$$

ステップ 1 ~ 3 は各ステージにおいて実行される。

[ステップ 4] 各ステージを通過するトラヒック・レート $\lambda_i^{(k)}$ (ステージ k から入り、ステージ i を通過するトラヒック・レート) を $\gamma^{(k)}$ 及び $0-D$ 行列から求める。ステージ k での平均滞在時間 $T_i^{(k)}$ は、次式で与えられる。

$$T_i^{(k)} = (\mu^{(k)} - \sum_{i=1}^M \lambda_i^{(k)})^{-1}$$

[ステップ 5] 各ステージの入カジョブがネットワークを通過するに要する平均時間 T_i が $T_i^{(k)}$ 及び $0-D$ 行列から求められる。 $\nu^{(i)}$ は、その逆数として与えられる。従って $\nu^{(i)}$ は、 $\nu^{(j)}$ ($1 \leq j \leq M$) の関数

$$\nu^{(i)} = F^{(i)}(\nu^{(1)}, \dots, \nu^{(M)}) \quad (1 \leq i \leq M)$$

として表わされ、この連立方程は数値的に解かれ、システムの平衡状態における下記の特性能の近似値が求まる。

(a) スループット: 受け入れられたジョブのトラヒック・レートであり、ネットワーク・スループット A はこれらの和で定義する。

(b) 遅延: 各ステージに入ったジョブがネットワークを通過するに要する平均時間であり、平均ネットワーク遅延 T は、これらのトラヒックによる重みつき平均である。

(c) パワー (power) [4]: 両極端の評価基準であるスループ

ットと遅延を折衷する量であり、次式で定義される。

$$\text{パワー} = \text{スループット} / \text{遅延}$$

(d) ブロッキング確率：各ステージにおけるブロッキング確率は前述したが、ネットワーク・ブロッキング確率は次式を満たす。

$$\sum_{i=1}^M \lambda^{(i)} (1 - P_B) = \Gamma$$

以上はリンク待ち行列長に制限がない場合だったが、制限のある場合には次の様に近似解を求める。

ブロッキングにより前段のサーバーは影響を受け処理率が低下するが、この現象を考慮に入れ、実効処理率を次の様に定義する。

$$\mu_e^{(k)} = \mu^{(k)} (1 - a^{(k-1)} P_L^{(k)})$$

但し、

$$a^{(k-1)} = \frac{\text{ステージ } k-1 \text{ から } k \text{ へのトラヒック・レート}}{\text{ステージ } k \text{ のトラヒック・レート}}$$

さらに、パーミットが余っているにもかかわらず、ブロッキングのためにジョブがリンク・バッファに入れられない現象を考慮するためにパーミットの実効到着率を次式で定義する。

$$\nu_e^{(k)} = \nu^{(k)} (1 - P_L^{(k)})$$

但し $P_L^{(k)}$ はステージ k のリンク・バッファ一杯である確率。 $\mu_e^{(k)}$ 、 $\nu_e^{(k)}$ を用い、リンク待ち行列に制限がない場合と同様

にして、連立方程式を得。これを数値的に解くことにより特性量の近似値が得られる。

4. 数値結果

近似解法は、その得られた結果をシミュレーション結果と比較することにより精度が検証されるが、その前にシミュレーションの妥当性を調べるために簡単な2ステージからなるモデルを取り扱う。この場合には、平衡方程式を解き得られた厳密解をシミュレーション結果と比較し図6、

図7に示す。但しリソフ待ち行列長には制限がなく他のパラメータは次の値をとっている。 $\lambda^{(1)} = \lambda^{(2)} = 0.05, \sim 1.20$, $\mu^{(1)} = \mu^{(2)} = 1.0$, $B^{(1)} = B^{(2)} = 1$, $N^{(1)} = N^{(2)} = 2$ 。これらの結果から、シミュレーションはある程度妥当であると思われる。このモデルで与え平衡方程式は75個の

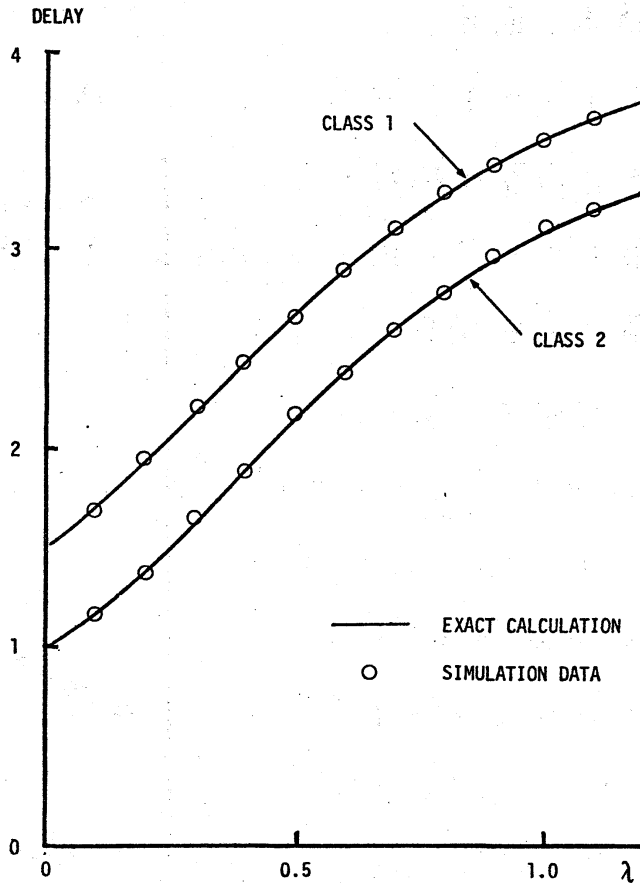


図6 到着率-遅延

未知確率を含み、これ以上複雑なモデルに対してはもはや平衡方程式を解くのは不可能になる。

次に5ステージからなるモデルで次の3通りの場合に対し、近似解法を適用した。

[ケース1] リンク待ち行列長に制限がなく各ステージに均一な到着がある場合でパラメータは次の値をとる。

$$\lambda^{(1)} = \lambda^{(2)} = \lambda^{(3)} = \lambda^{(4)} = \lambda^{(5)} = 0.05 \sim 1.20, \mu^{(1)} = \mu^{(2)} = \mu^{(3)} = \mu^{(4)} = \mu^{(5)} = 1.0$$

$$B^{(1)} = B^{(2)} = B^{(3)} = B^{(4)} = B^{(5)} = 1 \quad N^{(1)} = N^{(2)} = N^{(3)} = N^{(4)} = N^{(5)} = 1 \sim 4,$$

0-D行列は次である。

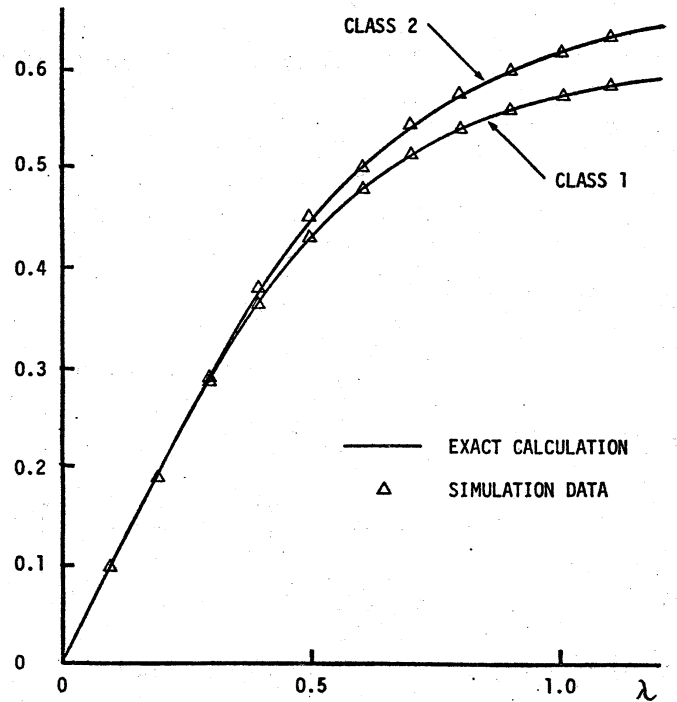


図7 到着率-スループット

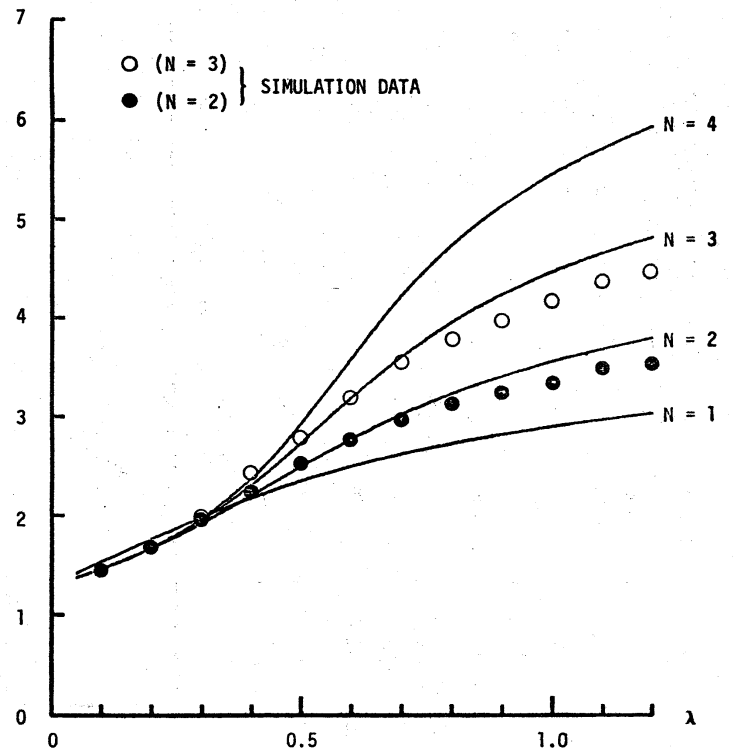


図8 到着率-遅延

$$\begin{bmatrix} 0.7 & 0.2 & 0.1 & 0 & 0 \\ 0 & 0.7 & 0.2 & 0.1 & 0 \\ 0 & 0 & 0.7 & 0.2 & 0.1 \\ 0 & 0 & 0 & 0.7 & 0.3 \\ 0 & 0 & 0 & 0 & 1.0 \end{bmatrix}$$

図8. 図9はパーミットの数が変化するとき到着率に対する平均遅延、スループットの関係もシミュレーション結果と比較して示して

おり、これらから近似解法は比較的良い近似値を与えると思われる。

また均一なシステムに対する輻輳コントロールの効果もわかる。

[ケース2] リンク待ち行列に制限がなく、各ステージに不均一な到着がある場合で、パラメータは次の値とする。 $\lambda^{(1)} = 0.05 \sim 1.20$.

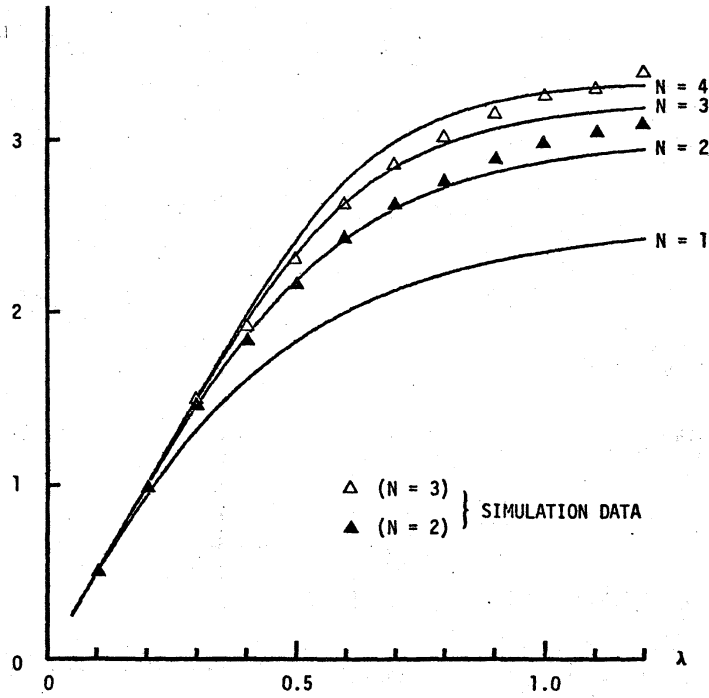


図9 到着率 - スループット

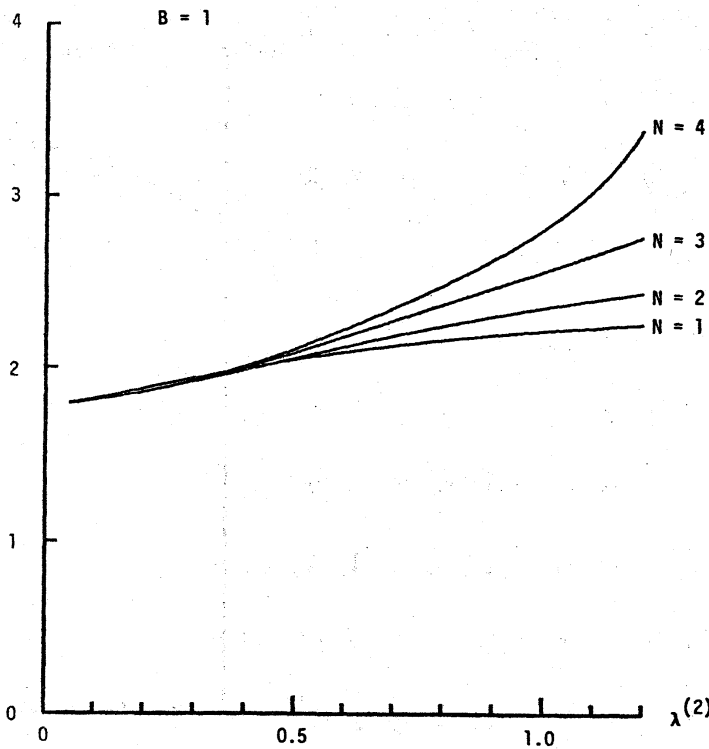


図10 到着率 - 遅延

$$\lambda^{(1)} = \lambda^{(2)} = \lambda^{(3)} = \lambda^{(4)} = 0.3$$

$$\mu^{(1)} = \mu^{(2)} = \mu^{(3)} = \mu^{(4)} = \mu^{(5)}$$

$$= 1.0, B^{(1)} = B^{(2)} = B^{(3)} = B^{(4)}$$

$$= B^{(5)} = 1, N^{(2)} = 1 \sim 4, N^{(4)}$$

$$= N^{(3)} = N^{(5)} = 2.$$

図10~12から、個々の
ステージのトラヒック
の変動に対しどのよう
なコントロールが行な
われるべきかがわかる。

例えば、 $N^{(2)} = 4$ のグラフ
が示すように、トラヒ
ックの集中しているス
テージへ多くのパーミ
ットを割り当てること
は必ずしもネットワー
ク全体から判断すると
良くはない場合もある。
[ケース3] リンク待
ち行列が有限な長さ $L^{(i)}$
に制限されている場合

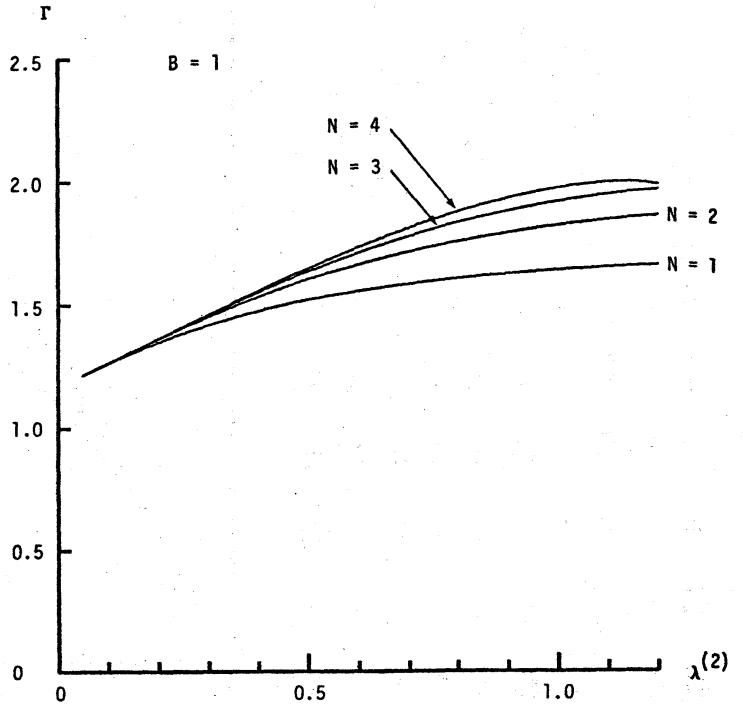


図11 到着率-スループット

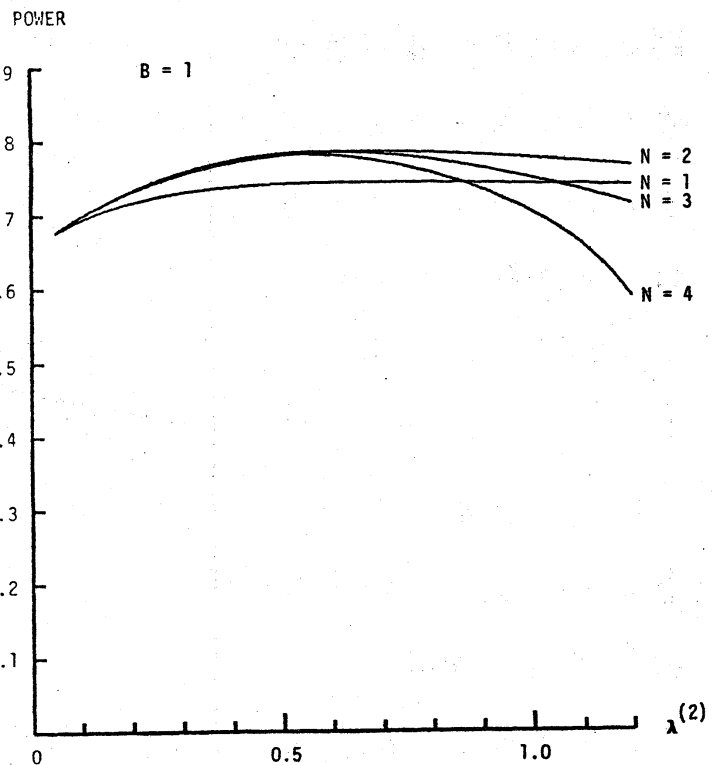


図12 到着率-パワー

い、パラメータは次の値をとる。 $L^{(1)} = L^{(2)} = L^{(3)} = L^{(4)} = L^{(5)} = 3, 4, 5,$
 ∞ 。 $B^{(1)} = B^{(2)} = B^{(3)} = B^{(4)} = B^{(5)} = 1$ 。 $N^{(1)} = N^{(2)} = N^{(3)} = N^{(4)} = N^{(5)} = 1 \sim 4$ 。
 O-D 行列は、次で与える。

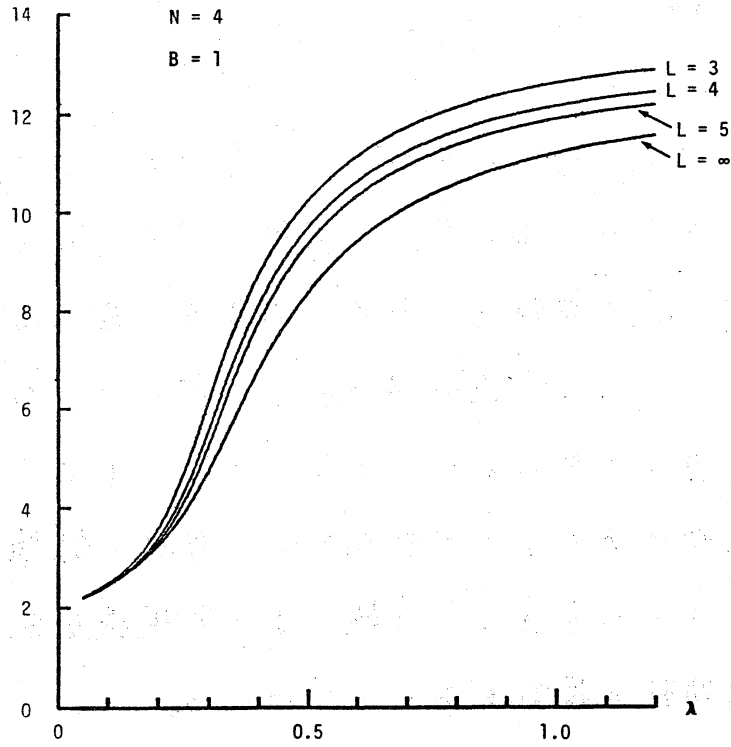
$$\begin{pmatrix}
 1/5 & 1/5 & 1/5 & 1/5 & 1/5 \\
 0 & 1/4 & 1/4 & 1/4 & 1/4 \\
 0 & 0 & 1/3 & 1/3 & 1/3 \\
 0 & 0 & 0 & 1/2 & 1/2 \\
 0 & 0 & 0 & 0 & 1
 \end{pmatrix}$$


図13. 到着率 - 遅延

図13, 図14からリンク・バッファの大きさ及びパケットの数がシステム全体に及ぼす影響がわかる。

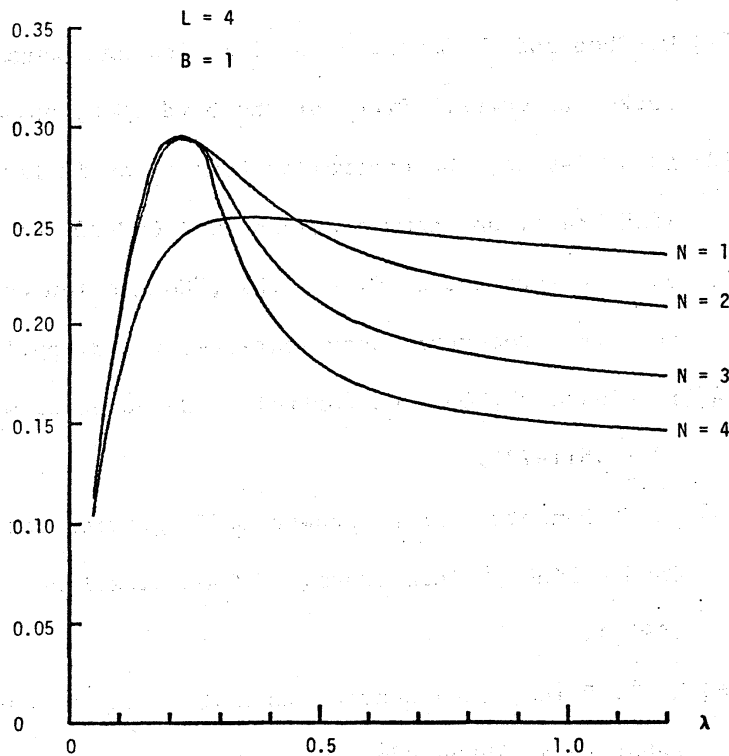


図14 到着率 - パワー

5. 終わりに.

以上、待ち行列網におけるフロー・コントロール特に輻輳コントロールに関し、実際に用いられている手法の概略を述べ、複合輻輳コントロール手法を導入し、その近似解法を示した。この近似解法は実用的には十分な精度で近似値を与えることがわかり、またこのコントロール手法も輻輳に関し有効であることが判明した。今後の課題としては、他のコントロール手法との比較、より大規模なネットワークに対する有効性の検証があると思われる。

[参考文献]

- [1] W. Chou and M. Gerla, "A Unified Flow and Congestion Control Model for Packet Networks," Proc. of the 3-rd ICC, pp.475-482(1976).
- [2] D. W. Davies, "The Control of Congestion in Packet-Switching Networks," IEEE Trans. on Communications, vol.COM-20, pp.546-550(1972).
- [3] J. L. Grange and J. C. Majithia, "Congestion Control for a Packet-Switched Network," Computer Communications, vol.3, pp.106-116(1980).
- [4] L. Kleinrock, "On Flow Control in Computer Networks," ICC'78, pp.27.2.1-27.2.5(1978).
- [5] M. C. Pennotti and M. Schwartz, "Congestion Control in Store and Forward Tandem Links," IEEE Trans. on Communications, vol.COM-23, pp.1434-1443 (1975).
- [6] W. L. Price, "Data Network Simulation Experiments at the National Physical Laboratory, 1968-1976," Computer Networks, vol.1, pp.199-200(1977).

- [7] M. Schwartz and S. Saad, "Analysis of Congestion Control Techniques in Computer Communication Networks," Flow Control in Computer Networks, North-Holland Publishing Company, pp.113-129(1979).
- [8] A. Chatterjee, N. D. Georganas and P. K. Verma, "Analysis of a Packet-Switched Network with End-to-End Congestion Control and Random Routing," Proc. of the 3-rd ICC, pp.488-494(1976).