

可変帯域リンクをもつネットワークの経路制御

山下 高生

小野 諭

yamasita@cactus.ntt.jp

ono@cactus.ntt.jp

NTT ソフトウェア研究所

あらまし これまでのコンピュータ・ネットワークは主に固定帯域リンクで構成されてきたが、ISDNの逆多重化装置やATMのVP(Virtual Path)など帯域を動的に変更可能なリンク(可変帯域リンク)の利用が可能となりつつある。可変帯域リンクは、ネットワーク設計の自由度の向上や経済的ネットワーク構成をもたらす一方で、現在用いられているルーティング・プロトコルでは、経路選択が不安定性になったり、かえって非経済的な経路が選択されるなどの問題点が生じる可能性がある。本報告では、リンク上のトラフィック量のみによって帯域変更を制御する可変帯域リンクを持つネットワークにおける最短経路選択のための経路距離を提案し、これをリング・ネットワークに適用したときの経路の安定性および経済性について検討を行なう。

A New Metric for Network with Variable Bandwidth Links

Takao Yamashita

Satoshi Ono

yamasita@cactus.ntt.jp

ono@cactus.ntt.jp

NTT Software Laboratories

Abstract Computer networks have been constructed mainly using fixed bandwidth links such as leased lines. Recently, variable bandwidth links, namely, links whose bandwidth can be changed dynamically, are also becoming available. For example, they include an Inverse Multiplexor of ISDN, and a Virtual Path (VP) of ATM networks. Although such links provide flexibilities in economical network design, they introduce following new routing problems, when conventional routing protocols are used. Firstly, it is possible that less economic routes are selected. Secondly, route oscillation may occur by bandwidth changing. In this report, we propose a new metrics of link that can be used in shortest path fast algorithms, and discuss the route stabilities under networks with variable bandwidth links.

1 序論

通信ネットワークにおいて、パケットを起点ノードから宛先ノードへ正しく導くためにはパケットが通る経路を全ての起点・宛先ノード対に対して決定しなければならない。この経路選択アルゴリズムを分類すると、集中型と分散型に分けることができる。集中型はすべてのノード対の経路を限られたネットワーク・センタで計算するのに対して、分散型は各ネットワーク・ノードが情報を互いに交換しながら協調して経路を計算する。分散型は、リンク故障やノード故障によってネットワークが分断された状態でも、経路を計算でき、耐故障性に優れているため様々なコンピュータ・ネットワークで用いられてきた。近年のコンピュータ・ネットワークで用いられている最短経路優先を用いた分散型経路選択のためのプロトコルには、RIP[2]、EGP[3]、BGP[4]などのDistance-Vector プロトコルとOSPF[1]やnew ARPANET[5]で用いられたLink- State プロトコルがある。Distance-Vector プロトコルでは分散型Bellman-Fordアルゴリズム[6]によって最短経路を計算する。Link- Stateはフラッディング[11]を用いて各ノードが自ノードに接続されたリンクの情報を他のすべてのノードに通知し、各ノードがその情報からネットワーク全体のトポロジー情報を形成して最短経路を計算する方法である。最短経路優先を用いた場合、経路の距離を決定することが必要となるが、この距離の計算方法は各リンクの帯域、信頼性、遅延時間、負荷やホップ数などのパラメータのうち、1つあるいは複数を組み合わせて行なわれる。IGRP[9]やARPANETで用いられたルーティング・プロトコルのように負荷パラメータを距離の計算に用いる適応型経路選択は、ネットワークに部分的な輻輳が存在した場合に、これを回避した経路を選択できるため、これまでのコンピュータ・ネットワークで用いられてきた。適応型経路選択では負荷という、故障やネットワークトポロジー変更と比較して頻繁に変化するパラメータを用いている。しかし、ネットワークが大規模になると、各ノードで独立に保持している経路情報がループを形成したり[12]、発振によって著しくネットワーク全体のスループットを悪化させる[7][8][10]ことを回避するためには短時間で経路を変更することは困難である。このため経路が変更されるまでの間、輻輳が続くことになる。

一方で、ISDN 逆多重化装置[13]やATM (Asynchronous Transfer Mode)のVP (Virtual Path)など動的に帯域を変化させることができるリンク(可変帯域リンク)が利用可能になりつつある。可変帯域リンクには、動的に帯域を変更可能であるという特徴の他に、動的に確保された帯域について、時間やデータ量など帯域の使用に比例したコストが必要になるという特徴がある。可変帯域リンクを用いることで、

1. 適応型経路選択の働きによって輻輳を回避するまでの間、可変帯域リンクの帯域を一時的に増加させることで輻輳を回避することが可能となり、輻輳に対して短時間で適応できるネットワークを構築できる。
2. 時間的に著しくトラフィック変動の大きいリンクに可変帯域リンクを用いることで経済的なネットワークを構築できる。

といったことが考えられる。

可変帯域リンクを含むネットワークを実現するためには、

- A 可変帯域リンクの配置方法(ネットワーク設計方法)。
- B 可変帯域リンクと固定帯域リンクという異なった性質を持つリンクを含むネットワークでの経路選択方法。

が必要であり、このうちBについて考えると、現在は可変帯域リンクの最大帯域をもとにした経路距離を用いるために、複数の経路が存在した時に可変帯域リンクにトラフィックが集まってしまう。このように、可変帯域リンクの利用は、現在の経路選択アルゴリズムではかえって非経済的な経路が選択され、可変帯域リンクの性質を十分に生かす経路選択方法が存在しない。

本報告では、帯域を常時接続されている帯域と動的に確保された帯域に区別した情報をもとにする経路距離計算方法を提案する。そして、可変帯域リンクを用いたときに問題が発生する最も簡単な場合として可変帯域および固定帯域リンクのみの経路の2つの経路が存在するようなリング・ネットワークでの安定性について検討し、従来技術である可変帯域リンクの最大帯域の情報を基にした経路距離を用いた場合と比較して、コストから決まる優先性に応じた経路選択が行なわれることをシミュレーションによって明らかにする。

第2章では可変帯域リンクに適した経路距離の計算方法について述べる。第3章では第2章で述べる経路距離の計算方法の動作特性を明らかにするために、可変帯域リンクを用いたときに問題が発生する最も簡単な場合として可変帯域リンクを含む経路と固定帯域リンクのみの経路の2つの経路が存在するようなリング・ネットワークを用いたモデルについて述べる。第4章でシミュレーションによる経路選択の有効性について述べる。第5章はまとめ、第6章で今後の課題について述べる。

2 経路距離計算方法

任意の固定帯域より可変帯域リンクについて、帯域を常時確保されている帯域と必要時に確保される帯域の和と考え、それぞれ基底帯域および付加帯域と呼ぶものとする。

可変帯域リンクと固定帯域リンクが含まれたネットワークを考えた場合、各リンクの経済的な性質から優先度の異なるリンクが存在するネットワークと考えることができる。付加帯域は、その使用に比例したコストが必要となることから考えると、使用する帯域の優先度は、まず第1に基底帯域、第2に付加帯域のうちコストが小さい帯域となる。この優先度を経路距離として実現するために、負荷が軽いときには基底帯域の情報によって距離を決定し、負荷が重いときには付加帯域および付加帯域のコストの情報を考慮して距離を決定するものとする。

ここで、それぞれの場合の距離を

軽負荷時 基底帯域を基本とした距離 d_l

重負荷時 基底帯域と付加帯域の和を表す距離 d_h

として、ノード uv 間の経路距離 $D(u, v)$ を以下のように負荷が軽い場合に支配的になる距離 $d_l(u, v)$ と、負荷が重いときに支配的になる距離 $d_h(u, v)$ の和として表現する。

$$D(u, v) = k_l d_l(u, v) + k_h d_h(u, v) \quad (1)$$

ただし、 k_l および k_h は各距離の重みを表す係数である。 $d_l(u, v)$ は経路上の基底帯域による各リンクの伝送遅延時間の和とし、

$$d_l(u, v) = \sum_{i \in P} b_i^{-1} \quad (2)$$

$$\mu_i = b_i + n_i \cdot \delta x_i \quad (3)$$

と表す。ここで、 μ_i 、 b_i および δx_i は、それぞれリンク i の帯域、基底帯域および付加帯域の変化単位を表し、 n_i は0以上の整数とし固定帯域リンクについては常に0とする。また、 P はノード uv 間の経路上のリンク集合を表す。

次に付加帯域間の優先性を考慮して $d_h(u, v)$ を経路上のリンクの空き帯域による伝送遅延時間を基本として、

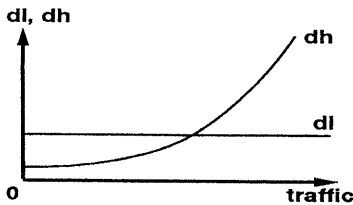


図1: 軽負荷時および重負荷時の距離

各リンクに割り当てられた優先度を考慮して次のように定める。

$$d_h(u, v) = \begin{cases} \max_{i \in P} (b_i - f_i)^{-1} & (\{i | n_i \neq 0\} = \phi) \\ \left(\sum_{i \in K} p_i \right) \max_{j \in P} ((\mu_j - f_j)^{-1}) & (\{i | n_i \neq 0\} \neq \phi) \end{cases} \quad (4)$$

ただし p_i および f_i は、リンク i の優先度とフロー（トラフィック）を表し、 $K = \{i | n_i \neq 0, i \in P\}$ とする。

d_l および d_h の負荷（フロー）による変化を図1に示す。それぞれの関数の性質から、経路の距離 $D(u, v)$ を d_l および d_h の和とすることで、負荷が軽いときには d_l が支配的になり、付加が軽い時には d_h が支配的になる。

3 特性の解析

本章では、安定性および経済性を検討するために用いるモデルについて述べる。第1節では、まず安定性および経済性を検討するために用いるネットワーク・トポロジーについて述べる。次に経路の制御方法について述べ、最後に可変帯域リンクの帯域制御方法について述べる。第2節ではシミュレーションによる評価について述べる。

3.1 解析モデル

どの経路を選ぶかという経路選択は、起点・宛先ノード対に対する経路が複数存在する場合に問題となる。経路制御の目的は、

- a 遅延時間やスループットなどのネットワーク品質とコストから決まる
- b 経路の安定性経路の最適さ

を満たすことである。bの経路の安定性とは、トラフィック自身が定常的であるときに経路が変化せず安定しているということである。適応経路制御においては経路の変更が経路上のトラフィックを変化させ、このトラフィック変化が更に経路を変化させるというフィードバック効果によって不安定になる場合がある [8][10]。

本報告では主に、経路選択、トラフィックおよび帯域の変化の相互作用について明らかにするために、複数の経路が存在する最も簡単な場合としてリングネットワークを考える。このリングネットワークにおいては

- I 2つの経路の両方とも固定帯域リンクのみ
- II 一方が可変帯域リンクを含み、他方が固定帯域のみ
- III 2つの経路の両方とも可変帯域リンクのみ

の組合せが考えられるが、このうち現在もっとも問題となると考えるIIの場合について検討する。

リング・ネットワークは図2(a)のように表されるが、これをさらに解析の容易さのために文献 [8] および文献 [10] で用いられている図2(b)のような連続ノード・モデルを利用して解析を行う。

連続ノード・モデルではリングの任意の点がノードとなり、リング・ネットワーク上の点は終点ノードからの反時計回りの距離 t で表される。 $t = 0$ かつ $t = 1$ が終点ノードであり、 $t \in (0, 1)$ は起点ノードである。トラヒッ

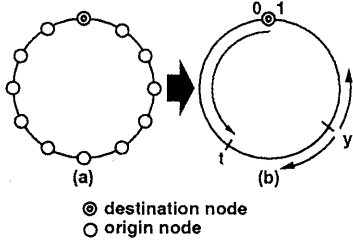


図 2: リング・ネットワーク

は起点ノードから終点ノードへながれるものとする。 t における入力トラヒックを単位長さ当たりトラヒックとして $r(t)$ 、帯域を $\mu(t)$ と表す。帯域 $\mu(t)$ は基底帯域と付加帯域の和として表される。解析に用いる $\mu(t)$ は次式のように定義し、図 4 にそのグラフに示す。

$$\mu(t) = \begin{cases} b_1 + n \cdot \delta x & (0 \leq t \leq a_1) \\ 1 & (a_1 < t < 1 - a_2) \\ b_2 & (1 - a_2 \leq t \leq 1) \end{cases} \quad (5)$$

ここで $t \in [0, a_1]$ が可変帯域リンクであり、 b_1 および b_2 が基底帯域、 $n \cdot \delta x$ が付加帯域である。

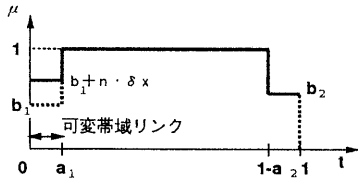


図 3: リング・ネットワークの帯域

起点ノードが終点ノードへトラヒックを流す場合、時計回りの経路および反時計回りの経路が存在する。各ノードは、その時点のルーティング y を参照し、 $t < y$ のトラヒックについては時計回り、 $t > y$ のトラヒックについては反時計回りを選択するものとする。

前章のネットワーク上の距離を、このリング・ネットワークに適用する。 t における時計回り経路の終点までの距離 $D^-(t, y)$ および反時計回りの距離 $D^+(t, y)$ は次のようになる。

$$D^-(t, y) = k_l d_l^-(t) + k_h d_h^-(\mu, y) \quad (6)$$

$$D^+(t, y) = k_l d_l^+(t) + k_h d_h^+(\mu, y) \quad (7)$$

ただし、

$$d_l^-(t) = 1 - t + \frac{a_2}{b_2} - a_2 \quad (8)$$

$$d_l^+(t) = t + \frac{a_1}{b_1} - a_1 \quad (9)$$

$$d_h^-(\mu, y) = \frac{1}{\mu - F(y)} \quad (10)$$

$$d_h^+(y) = \frac{1}{b - F(1) + F(y)} \quad (11)$$

$$F(y) = \int_0^y r(s) ds \quad (12)$$

$$r(s) : \text{input source rate} \quad (13)$$

ここで、ルーティング y が $F(y) \leq F(1) - b$ または $F(y) > \mu$ となった場合にの扱いについて述べる。この場合、リング・ネットワーク上のすべての起点からのトラヒックを宛先ノードに流すことができない。通常、ある帯域のリンクにその帯域以上のトラヒックが流れこもうとした場合、輻輳制御によって起点ノードからの入力トラヒックが抑制される。このことを考慮すると式 (10) および (11) はつぎのようになる。

$$d_l^-(\mu, y) = \frac{1}{\mu - F^-(y)} \quad (14)$$

$$d_l^+(\mu, y) = \frac{1}{b_2 - F^+(y)} \quad (15)$$

$$F^-(y) = \min(F(y), s \cdot \mu) \quad (16)$$

$$F^+(y) = \min(F(1) - F(y), s \cdot b_2) \quad (17)$$

ただし、帯域 B のリンクにフロー制御を行なったトラヒックを流した場合、実際に使用可能な最大帯域の B に対する比率を s とする。この輻輳制御によるトラヒックの制限は可変帯域リンクの帯域や経路の更新の時間間隔に比べて十分に短時間でなわれ、輻輳制御によるトラヒック抑制のトランジエントな変更過程は可変帯域リンクの帯域や経路の更新に影響を与えないと仮定する。このことは現在行なわれている可変帯域リンクの帯域変更や経路の更新時間が数十秒程度以上ということと十分に現実的な仮定であると考えられる。

経路選択方法 次にルーティング y の更新方法について説明する。ある時点でのルーティングを y_{n-1} としたとき次のルーティング y_n は、

$$D^+(y_n, y_{n-1}) = D^-(y_n, y_{n-1}) \quad (18)$$

の解として求めることができる。上式に (6) から (16) 式を代入すると

$$y_n = \frac{k_l \left\{ \frac{\mu - b_2 - F^-(y_{n-1}) + F^+(y_{n-1})}{2k_c \{b_2 - F^+(y_{n-1})\}} \{ \mu - F^-(y_{n-1}) \} + \frac{a_2}{b_2} - \frac{a_1}{b_1} + a_2 - a_1 + \frac{1}{2} \right\}}{\quad} \quad (19)$$

ただし、上式によって計算した y_n が $y_n < a_1$ または $y_n > 1 - a_2$ の場合には

$$y_n \leftarrow \begin{cases} a_1 & (y_n \leq a_1) \\ 1 - a_2 & (y_n \geq 1 - a_2) \end{cases} \quad (20)$$

とする。また、ルーティング y^* が

$$D^+(y^*, y^*) = D^-(y^*, y^*) \quad (21)$$

を満たすときトラフィックおよび可変帯域リンクの帯域が変化しない間、ルーティングの更新時によってルーティングの値が変化しない平衡状態となる。

通常のネットワークでは、各ノードが非同期に経路を変更する。このことを考慮して、すべての起点ノード集合をランダムに部分集合に分割し、部分ノード集合ごとに経路を上記の方法に従って経路を変更するものとする。

可変帯域の帯域制御方法 可変帯域リンクにおける帯域 ($b + n \cdot \delta x$) の制御方法を説明する。可変帯域リンクの帯域は、可変帯域リンクを流れるフローの帯域の情報のみを用いて制御するものとする。一般的には、各 n について帯域 Bu_n および Bd_n ($Bu_n < Bd_n$) および時間 Tu_n および Td_n を定め、

- Tu_n の間の平均トラフィックが Bu_n を越えた場合
 $n \leftarrow \min(n + 1, n_{max})$
- Td_n の間の平均トラフィックが Bu_n を下まわった場合
 $n \leftarrow \max(n - 1, 0)$

とする。ただし、 n_{max} はその可変帯域リンクが最大帯域となるときの n の値である。

3.2 シミュレーションによる評価

連続ノードモデルを用いて、本報告で提案した経路距離を用いた場合の経路選択について評価を行なう。評価方法は、各起点ノードからリング・ネットワークに入力されたトラフィック量 (全トラフィック量) の、可変帯域リンクを含む経路と含まない経路へのトラフィック割り当てを評価する。

本報告で提案した経路距離計算方法の評価を行なうための基準について検討する。経済的に理想的な状態を考えると、可変帯域を含まない経路によって全トラフィック量を終点ノードに送ることができるときは、可変帯域リンクを含まない経路の使用率が100%となることが理想的であると考えられる。全トラフィック量を可変帯域リンクを含まない経路によって送れるときには、遅延時間の公平性を考えて可変帯域リンクを含む経路と含まない経路の最大使用率が同一になるべきであると考えられる。

シミュレーションを行なうためのパラメータについて説明する。可変帯域リンクの帯域制御を行なうためのパラメータ Bu_n および Bd_n は、

$$Bu_n = \begin{cases} 0.9b_1 & (n = 0) \\ b_1 + 0.9n \cdot \delta x & (n \geq 1) \end{cases} \quad (22)$$

$$Bd_n = \begin{cases} 0.5b_1 & (n = 1) \\ b_1 + 0.9(n - 1) \cdot \delta x & (n \geq 2) \end{cases} \quad (23)$$

とし、可変帯域リンクの帯域更新時間間隔は経路更新時間間隔の1/50とした。また、帯域については次式のように定めた。

$$\mu = \begin{cases} 1 + 0.1 \cdot n & (0 \leq t \leq 0.05) \\ 3 & (0.05 < t < 0.95) \\ 1 & (0.95 \leq t \leq 1) \end{cases} \quad (24)$$

まず、全トラフィック量が3.2、可変帯域リンクの帯域幅の初期値が1.2、ルーティング y の初期値を0.9としたと

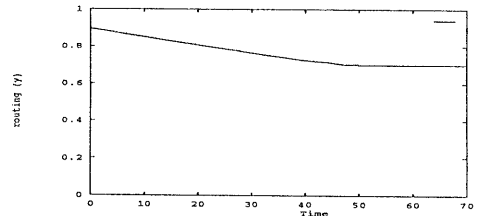


図4: ルーティング y の時間変化

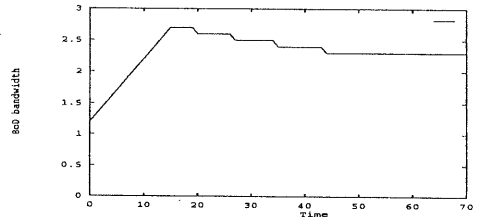


図5: 可変帯域リンクの帯域幅の時間変化

きのルーティング y および可変帯域リンクの帯域幅の時間変化を図4および図5に示す。ここに示すように、ルーティングおよび可変帯域リンク帯域幅は一定時間後に一定の値に収束する。以後、このような収束した時の状態についてのシミュレーション結果を示す。

図6にリング・ネットワークの2つの経路のうち固定帯域リンクの使用率の最大値 (縦軸) ($t=1$ での利用率) と全トラフィック量 (横軸) の関係を示す。ここで、MAXで示されるグラフは従来技術による結果であり、IDEALが上述の理想的な特性である。従来技術による結果とは、IGRPにおいて可変帯域リンクが確保できる最大帯域をもとにした経路距離によるものである。NEWが本報告で提案したリンク距離を用いた場合のグラフである。比較に用いた従来技術は、可変帯域リンクを、そのリンクで利用可能な最大帯域を持つ固定帯域リンクであるとして経路制御を行なう方法である。この図から従来技術と比較して本報告で提案した方法を用いた場合、理想的な使用率に近くなることがわかる。

つぎに、図7にルーティング y （縦軸）と全トラフィック量（横軸）の関係を示す。この図から、リング・ネットワークという限定されたものではあるが、従来技術と比較して、より理想的なルーティング y に近くなり、改善されていることがわかる。

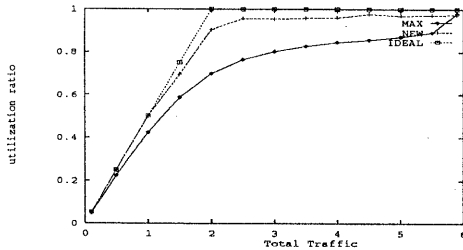


図 6: 固定帯域リンクの利用率

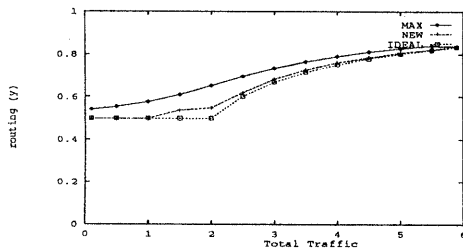


図 7: ルーティングと総トラフィック量の関係

4 まとめ

現在、可変帯域リンクを用いると最大帯域をもとにしたリンク距離を用いるために、複数の経路が存在した時に可変帯域リンクにトラフィックが集まってしまう。このように、可変帯域リンクの利用は、現在の経路選択アルゴリズムではかえって非経済的な経路が選択され、可変帯域リンクの性質を十分に生かす経路選択方法が存在しない。

本報告では、帯域を常時接続されている帯域と動的に確保された帯域に区別した情報をもとにする経路距離計算方法を提案した。そして、可変帯域リンクを用いたときに問題が発生する最も簡単な場合として可変帯域リンクを含む経路と固定帯域リンクのみの経路の2つの経路が存在するようなリング・ネットワークでの安定性について検討を行ない、可変帯域リンクの最大帯域の情報に基づいた経路距離を用いた場合と比較して、優先性に応じた経路選択が行なわれることをシミュレーションによって明らかにした。

5 今後の課題

- リング・ネットワークで2つの経路が両方も可変帯域リンクの場合の特性についての検討。
- ネットワーク・シミュレータなどを用いた一般的なネットワークでの動作のシミュレーション。
- 一般的ネットワークでの、本報告で提案したリンク距離の計算方法の妥当性の評価基準の検討。

参考文献

- [1] R.Coltun, OSPF: An internet routing protocol, *ConneXions*, vol. 3, no. 8, pp. 19-25, Aug. 1989.
- [2] C.Hedrick, Routing information protocol, RFC 1058, *Netw. Inform. Cent., SRI Int., Menlo Park, CA*, June 1988.
- [3] D.Mills, Exterior gateway protocol formal specification, RFC 904, *Netw. Inform. Cent., SRI Int., Menlo Park, CA*, Dec. 1983.
- [4] K.Lougheed and Y.Rekhter, Border gateway protocol 3 (BGP-3), RFC 1267, *SRI Int., Menlo Park, CA*, Oct. 1991.
- [5] J.McQuillan, Adaptive routing algorithms for distributed computer networks, BBN Rep. 2831, Bolt Beranek and Newman Inc., Cambridge, MA, May 1974.
- [6] L.R.Ford and D.R.Fulkerson, *Flows in Networks*, Princeton, NJ: Princeton Univ Press, 1962.
- [7] D.Bertsekas and R.Gallager, *Data Networks*, Prentice-Hall Inc., 1987
- [8] D.P.Bertsekas, Dynamic behavior of shortest path routing algorithms for communication networks, *IEEE Trans. on Auto. Control*, pp.60-74, Feb., 1982.
- [9] An introduction to IGRP. preprint, Laboratory for Computer Science Research, Rutgers University, New Jersey, August, 1991.
- [10] S.Low and P.Varaiya, Stability of a Class of Dynamic Routing Protocols (IGRP), *Infocom*, pp.610-616, Vol.2, 1993.
- [11] E.C.Rosen, The updateing protocol of ARPANET's new routing algorithm, *Comput. Networks*, vol. 4, no.1, Feb. 1980.
- [12] J.M.Jaffe and F.M.Moss, "A responsive routing algorithm for computer networks", *IEEE Trans. Commun.* vol. COM-30, No. 7, pp.1758-1762, July 1982.
- [13] 山下、小野 「ISDN 逆多重化装置の特性評価と効率の利用法」 情報処理学会研究報告 (93-DPS-61), Vol.93, No.58.