

# 修士学位論文

題目

資源予約型サービス / ベストエフォート型サービスにおけるユーザ効  
用の比較

指導教官

宮原 秀夫 教授

報告者

渡邊 啓

平成 12 年 2 月 15 日

大阪大学 大学院基礎工学研究科  
情報数理系専攻 計算機科学分野

資源予約型サービス/ベストエフォート型サービスにおけるユーザ効用の比較

渡邊 啓

内容梗概

従来、インターネットではスループットや遅延などの通信品質 (QoS) を保証しないベストエフォート型サービスのみが提供されてきた。しかし、音声や動画像などの実時間転送に対する要求が高まり、そのためユーザコネクションに対してネットワーク資源を確保することによって通信品質を保証する資源予約型サービスについても検討されるようになってきている。しかしながら、資源予約型サービスについては、パケット処理のオーバーヘッドなどの問題とともにスケーラビリティの問題も指摘され、「ベストエフォート型のままでよいのか？それとも資源予約型が必要か？」という基本的な議論も引き続いて行われている。本論文では、ユーザがネットワークアプリケーションを使用した時に得られる通信品質とそれに対して必要なコストの関係をユーザ効用と定義し、ベストエフォート型サービスと資源予約型サービスの2つのサービスモデルにおけるネットワークアプリケーションのユーザ効用を示し、それらの比較評価を行っている。その結果、回線容量が十分にあればベストエフォート型サービスでも十分な効用が得られること、ただし、そのためには適切なネットワーク設計によって十分な回線が準備できること、そうでない場合には予約型サービスが必要であることなどを定量的に示した。

主な用語

通信品質、ベストエフォート型サービス、資源予約型サービス、ユーザ効用、料金制度

## 目次

1	はじめに	5
2	サービスモデルとアプリケーション分類	9
3	ユーザ効用の導出	13
3.1	モデル化	13
3.2	非適応型アプリケーション	13
3.2.1	予約型サービスに適用した場合	13
3.2.2	ベストエフォート型サービスに適用した場合	13
3.3	適応型アプリケーション	15
3.3.1	予約型サービスに適用した場合	15
3.3.2	ベストエフォート型サービスに適用した場合	17
3.4	比較評価	17
4	多段モデルへの拡張	21
4.1	モデル化	21
4.2	非適応型アプリケーション	21
4.2.1	予約型サービスに適用した場合	21
4.2.2	ベストエフォート型サービスに適用した場合	22
4.3	適応型アプリケーション	23
4.3.1	予約型サービスに適用した場合	23
4.3.2	ベストエフォート型サービスに適用した場合	23
4.4	比較評価	24
5	料金制度を考慮した評価	27
5.1	モデル化	27
5.1.1	ベストエフォート型サービスに適用した場合	27
5.1.2	資源予約型サービスに適用した場合	31

5.2 比較評価 . . . . .	33
6 むすび	36

## 図目次

1	料金制度 . . . . .	10
2	QoS 関数 . . . . .	11
3	QoS 関数 $\alpha(\cdot)$ . . . . .	19
4	QoS 関数 $\beta(\cdot)$ . . . . .	19
5	非適応型アプリケーションの平均、最悪効用 . . . . .	20
6	適応型アプリケーションの平均、最悪効用 . . . . .	20
7	拡張モデル . . . . .	21
8	非適応型アプリケーションの平均効用 . . . . .	25
9	非適応型アプリケーションの効用差 . . . . .	25
10	適応型アプリケーションの平均効用 . . . . .	26
11	適応型アプリケーションの効用差 . . . . .	26
12	ネットワークモデル 2 . . . . .	28
13	状態遷移図 1 . . . . .	29
14	状態遷移図 1 . . . . .	32
15	実時間アプリケーションの効用 $U_{r,b}, U_{r,r}$ . . . . .	35
16	Web ユーザの効用— ユーザ数が増加 $U_{w,b}, U_{w,r}$ . . . . .	35

## 1 はじめに

従来、インターネットは、個々のコネクションに対する通信品質を保証しない、ベストエフォート型サービスのみを提供してきた。これは、もともと ARPANET にはじまるコンピュータネットワークが遠隔端末 (telnet) やファイル転送 (ftp) や電子メールなどいわゆるデータ系アプリケーションのために開発されてきたことによる。しかしながら、実時間転送を含むマルチメディアアプリケーションの重要性が認識されるようになり、ベストエフォート型サービスを前提とした実時間アプリケーションの開発も盛んに行われるようになった。しかし、ベストエフォート型サービスは、もともと実時間転送の通信品質を保証する機構を持たず、またユーザ数の増大によって回線容量が絶対的に足りなくなると、その輻輳のために品質が著しく劣化するという問題は避けられない。

そのため、インターネット上で実時間転送の通信品質を保証するアーキテクチャとして、ISPN (Integrated Service Packet Network) [1] が検討されるようになった。ISPN では、従来のデータ系通信サービスに加えて、コネクションごとに通信開始前に資源を予約することによって、実時間通信に対してその通信品質を保証する。その具体的な実現手法が、RSVP (ReSerVation Protocol) [2] などのシグナリングプロトコルと、WFQ (Weighted-Fair Queueing) [3] などのパケットスケジューラである。

しかし、ISPN に関してはスケーラビリティによる限界も指摘されている。すなわち、

1. シグナリングプロトコルが必要になるため、ネットワークが大規模になると適用が困難になる
2. パケットスケジューラのアルゴリズムの複雑性により、コネクション数の増大に対応できない

などの問題である。また、ネットワークが輻輳してくると、予約型サービスでは必然的に呼損を引き起こす。ベストエフォート型サービスにおいても、輻輳によってユーザが許容できないほどの品質劣化が発生するが、輻輳に動的に適応する輻輳適応型の処理メカニズムを導入すれば通信の接続性が維持でき、その方がより重要であるという考え方も成立する。

ベストエフォート型サービスを対象として、輻輳適応型メカニズムを実装した実時間アプ

リケーションとして文献 [4] が挙げられる。このアプリケーションでは、受信側からのフィードバック情報により、送信側はネットワークの輻輳状態を検知し、それに応じて発生レートを増減させるメカニズムである。このようなアプリケーションを使用することによって、接続性は確かに保証される。ただし、その場合、品質劣化が始まるトラフィック量の限界を輻輳適応型メカニズムを導入しない場合の限界点をより高めただけであり、トラフィック量が増大すれば機能しえなくなる。

また、輻輳適応型メカニズムはベストエフォート型サービスだけでなく、予約型サービスにも適用可能である。すなわち、予約型サービスにおける輻輳適応型メカニズムでは、シグナリングを通じた帯域再交渉により、利用可能な帯域に応じて送信レートを増減させる。これは、例えば動画像通信においては量子化定数の調整によって実現可能である [5]。この場合においても呼損は発生しうるが、いったんコネクション接続が許されれば、その後はコネクションに対して許容可能な通信品質の保証が実現できる。

以上に述べた論点に対して、予約型サービスとベストエフォート型サービス、それぞれを用いた場合に、ある与えられたトラフィック量に対してどの程度の品質が提供されるかは明らかになっていない。そこで、本研究では、実時間転送を行うアプリケーションをそれぞれのサービスに適用した場合に得られる通信品質を表す指標としてユーザ効用を導入し、そのユーザ効用を用いることによって2つのサービスの比較評価を示す。

特に、予約型サービスでは帯域保証のメカニズムによって最低通信品質が保証できるのに対して、ベストエフォート型サービスではそのような保証がないという根本的な違いがある。そのため、ベストエフォート型サービスに輻輳適応型メカニズムを適用した場合においても、品質が通信途中で許容できないほどに劣化する可能性は避けられない問題であり、その点も考慮した評価を行う (第3章)。

また、一般に回線交換型ネットワークでは、経由するノードの多い、すなわち、ホップ数の多いコネクションは、ホップ数の小さいコネクションよりも呼損率が高くなり、ホップ数が大きくなるに従って呼損率が急激に悪くなることが知られている。予約型サービスは基本的に回線交換ネットワークと同様の性質を持ち、その面からも予約型サービスは大規模なネットワークへの適用が難しくなると考えられる。そこで、本論文では、コネクションが複

数のノードを経由する場合についても評価を行っている (第 4 章)。

ユーザの満足度 (ユーザ効用) を表す指標としては、上記した通信品質が考えられるが、他に、ユーザが通信する際に支払う通信コストもユーザの満足度の指標と考えられる。上述の評価では、通信コストは、2 つのサービスともに帯域に比例するものとして扱ってきた。しかし、現在、ISP で広く採用されている料金制度は、ISP への接続時間を基準にしたものが多い。例えば、接続時間に比例した料金をユーザに課す従量制や、月単位で一定料金を課す定額制、あるいはそれらの組み合わせによるものなどである。これらの違いは、ユーザの利用形態に応じたものであり、例えば、メール中心のユーザは従量制を利用し、ヘビーユーザは定額制を利用する方が割安感が得られるであろう。メール中心のユーザであれば、必要なデータのみを転送する時間だけ ISP に接続するため、接続時間において転送時間が占める割合は高くなる。しかしながら、Web 利用を中心としたユーザであれば、接続時間の内、転送時間が含む割合は、メール中心のユーザよりも小さくなるのが十分に考えられる。従って、接続時間のみでは、ユーザが利用するネットワーク資源量を正しく表現しているとは言えない [6]。これは、ISP が提供しているサービスが、各コネクションに対してネットワーク資源の保証を行わないベストエフォート型サービスであるためである。

一方、資源予約型サービスでは、ユーザに対して利用資源量に応じた料金を課すことも可能になると考えられる。また、現在の単純な料金制度では収入予測やトラフィック管理が単純であるのに対して、資源予約型サービスでは、トラフィック管理が困難になり、また、トラフィック予測が正しく行われなければ収入予測も困難である可能性は大きい。そこで、本論文ではユーザ効用をさらに、通信品質とそれにかかるコストによって定義し、料金制度やユーザの利用形態を考慮可能な評価モデルをまず構築する。特に、現在、実時間アプリケーションとともに、インターネットのトラフィックの主流を占めている Web トラフィックをモデル化の対象とし、それぞれをベストエフォート型サービスと資源予約型サービスに適用した場合を考える (第 5 章)。

最近、ISPN モデルに対する反省として、ベストエフォートサービスと比較して「より良い」サービスを提供するため枠組みとして diff-serv (differentiated service) モデルが議論されている [7]。本論文に述べるモデル化と diff-serv モデルとの対応は以下のように捉えるこ

とができる。

1. 帯域ブローカ [8] などによるコネクション設定手続きがある場合：帯域がコネクションに対して保証されるという意味で、予約型サービスに対応する。
2. PHB (Per-Hop Behavior) [9] を用いたパケットスケジューラのみを適用する場合：従来のデータ系トラヒックに対して、実時間トラヒックをルータが優先処理するとすれば、実時間トラヒックのみを考えた場合は、実時間トラヒックに対するサービス形態はベストエフォートサービスと同じになる。

なお、関連する研究として、文献 [10] において予約型サービスとベストエフォート型サービスの比較が行われている。文献 [10] では、ネットワークのコネクション数をあらかじめ定められた確率分布として与えてネットワークのモデル化を行っている。それによって、コネクション数に対するユーザ効用の平均を求め、その総和を各ネットワークサービスがユーザに提供できる性能として定義し、予約型サービスとベストエフォート型サービスを比較している。したがって、コネクション数が動的に変化する場合について考慮していないという問題がある。一方、本論文ではユーザ効用を導入し (3章参照)、より具体的なトラヒックに対するモデル化により評価を行っている。

以下、まず、2章において、本論文で対象とするサービスモデルとアプリケーションを導入する。次に、3章では、アプリケーションを輻輳適応型、非適応型に分類し、それぞれにおいて得られるユーザ効用を導出し、また、数値例を用いてユーザ効用の比較を行う。次に、4章では、3章における単一回線モデルを拡張して、複数段のノードを有するモデルにおけるユーザ効用を導出し、その数値例を示す。そして、5章では、負の効用としての料金制度を明示的に考慮したユーザ効用を対象とした評価を行い、その数値例を示す。最後に6章にまとめと今後の課題を述べる。

## 2 サービスモデルとアプリケーション分類

本論文に述べるモデル化においてはサービスモデルを以下のように考える [11]。

1. 予約型サービス： ネットワーク内の物理資源をコネクションに対してあらかじめ確保することによって、QoS 保証が可能となる。そのため、RSVP のようなシグナリングプロトコルがネットワークに必要となる。資源を確保できない場合は呼損とする。  
なお、資源予約型サービスにおいて、予約の対象となる資源については帯域やバッファなどさまざまなものが考えられ、また、通信品質の指標としてもスループットや遅延、パケット棄却率などが考えられる。しかし、本論文では、その実現性を考慮して帯域のみを予約の対象とし、その結果、スループットを通信保証の対象と考える [12]。実時間トラヒックが存在し、資源予約型サービスを利用する場合にも、データ系トラヒックは資源予約されていない帯域を使用するものとする。従って、帯域を予約するユーザは、帯域を優先利用するための料金を支払うと考える。
2. ベストエフォート型サービス： 従来のインターネットのように、ネットワーク内の資源を各コネクションに対してあらかじめ確保するようなことはしない。従って QoS 保証が実現できず、ネットワーク資源不足による呼損は発生しない。ベストエフォート型サービスの料金制度として、以下の料金制度の中で、定額制を用いる。
  - (a) 従量制  
従量制では、ユーザが ISP を利用した時間に比例した料金が課せられる。すなわち、図 1(a) のようなグラフで表すことができる。
  - (b) 定額制  
一ヶ月などを単位として、接続時間に関わらず料金が一定の課金制度である ( 図 1(b) )。
  - (c) 定額 + 従量制  
従量制と定額制の中間の料金制度として、ある一定時間までは料金が変わらず、その時間を越えると利用時間に比例した料金が必要となる定額 + 従量制度が考えられる ( 図 1(b) )。

一方、アプリケーション分類としてさまざまな分類方法が考えられるが、本論文では、

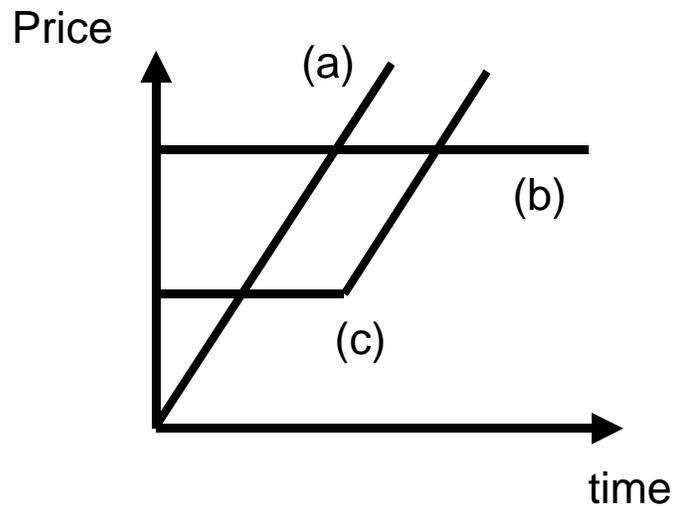


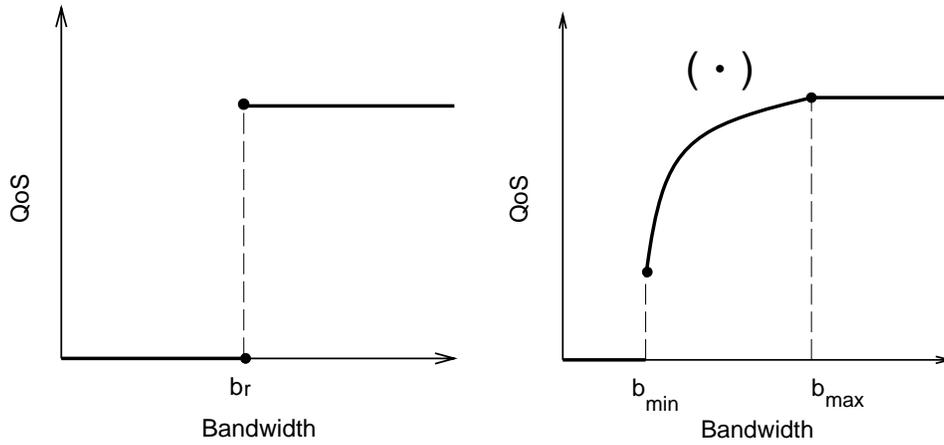
図 1: 料金制度

1. データ系アプリケーション (Web トラフィック)
2. 片方向実時間アプリケーション
3. 対話型実時間アプリケーション

に分類し、その内、対話型実時間アプリケーションを比較評価の対象とする。対話型実時間アプリケーションは遅延の制約に厳しいため、その保証を行うために、コネクションに対して帯域を保証すると考える [12]。

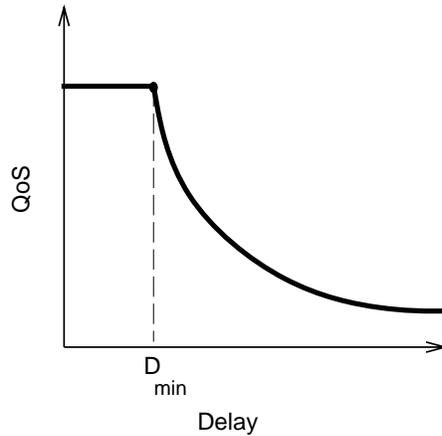
また、実時間アプリケーションをネットワークの輻輳状態に応じてその発生レートを変動させることの可能な帯域適応型と、そのようなメカニズムを持たない非適応型に分類する。例えば、音声アプリケーションの場合、PCM 音源を利用したアプリケーションは帯域非適応型であり、一方、データ圧縮した ADPCM 音源を利用したアプリケーションは、帯域適応型アプリケーションと考える。また、ネットワークの輻輳状態に適応する実現形態として遅延適応型メカニズム [13] もあるが、遅延適応型メカニズムは本来、片方向実時間アプリケーションに対して適用されるものであり、ここでは考えない。

また、本論文では、帯域適応型、非適応型の 2 つの対話型実時間アプリケーションにおいて得られる QoS を評価の対象とし、帯域に対して QoS 関数が定義されているものとする。ただし、ここで考える通信品質はコネクションが設定された後の通信品質であり、例えば、



(a) 帯域非適応型アプリケーション ( $\alpha(\cdot)$ )

(b) 帯域適応型アプリケーション ( $\beta(\cdot)$ )



(c) データ系アプリケーション ( $\gamma(\cdot)$ )

図 2: QoS 関数

SN 比や MOS によって決定されるものである。すなわち、帯域非適応型アプリケーションでは、図 2(a) に示すような QoS 関数  $\alpha(\cdot)$  を考える。すなわち、帯域がある値  $b_r$  より小さい場合、そのアプリケーションの通信品質は著しく劣化するため 0 になり、それ以上の場合は十分な通信品質が得られるものとする。一方、適応型アプリケーションにおいては、図 2(b) に示すように帯域に応じた QoS 値が定められる QoS 関数  $\beta(\cdot)$  を考える。すなわち、許容できる通信品質を得るために必要な最低限の帯域  $b_{min}$  が定義され、それ以下の帯域になると品質が著しく低下し、0 となる。一方、 $b_{min}$  を越えると、得られた帯域に応じて通信品質が向上していくものとする。ただし、最大値  $b_{max}$  を越えるとそれ以上の向上はない。

ベストエフォート型サービスを利用するアプリケーションとして、現在のトラフィックのほとんどを占める http コネクションによる通信、すなわち、Web トラフィックを想定する。Web トラフィックの通信品質としてドキュメント転送遅延時間を考える。そのために、Web トラフィックの QoS 関数として図 2(c) に示すようなドキュメント転送遅延時間に対して定義される関数を導入する。すなわち、遅延時間が  $D_{min}$  以下の場合、ユーザは十分な通信品質が得られていると感じ、遅延時間が  $D_{min}$  以上になると、遅延時間が大きいほどコネクションの通信品質が低下すると感じるというものである。

以上まとめると、本論文では、対話型実時間アプリケーションを

1. 帯域非適応型
2. 帯域適応型

の 2 つに分類し、それぞれを

- (a) 予約型サービス
- (b) ベストエフォート型サービス

を提供するネットワークに対して適用した場合のユーザ効用を比較検討する。

さらに、料金制度を考慮した評価として、ネットワークサービスを

1. ベストエフォートサービス
2. 資源予約型サービス

と考え、それぞれのサービスに対して

1. Web トラフィック
2. 帯域非適応実時間アプリケーション

を適用した場合のユーザ効用を求め、2 つのネットワークサービスを比較評価する。

### 3 ユーザ効用の導出

#### 3.1 モデル化

本論文では、まず、基本的なモデルとして回線容量  $C$  を持つ単一回線を考え、コネクションはそれを共有するものとする。コネクションは生起率  $\lambda$  のポアソン分布に従って到着し、コネクション継続時間は平均  $1/\mu$  の指数分布に従う。システムに対する負荷を  $a = \lambda/\mu$  とする。また、議論の簡単化のため、すべてのコネクションは同じ QoS 関数を持つものとする。さらに、コネクション設定時間については、コネクション接続時間に比べて十分小さいとして無視するものとする。

以下、非適応型アプリケーション、適応型アプリケーションの順に、各コネクションが得るユーザ効用を導出する。

#### 3.2 非適応型アプリケーション

##### 3.2.1 予約型サービスに適用した場合

予約型サービスでは、コネクションは通信開始前に帯域  $b_r$  をネットワークに対して要求する。ネットワーク側は、その帯域が確保できる場合はそのコネクションを受け付けるが、保証できない場合は呼損とする。そのため、予約型サービスは、コネクション数  $m = \lfloor C/b_r \rfloor$  まで同時に受け付けることが可能で、それ以上の接続要求は呼損となる M/M/m/m 待ち行列システムでモデル化できる。

通信品質は、受け付けられた場合  $\alpha(b_r)$  で与えられるが、呼損となった場合は 0 となる。そのため、ユーザ効用  $U_{r,r}$  (Utility for Rigid application in Reservation-based service) を以下のように定義する。

$$U_{r,r}(C) = (1 - L_{r,r}(C))\alpha(b_r) \quad (1)$$

ただし、 $L_{r,r}$  は同時接続可能なコネクション数を  $m$ 、負荷を  $a$  とするアラン呼損率である。

##### 3.2.2 ベストエフォート型サービスに適用した場合

ベストエフォート型サービスにおいては、原理的に呼損は発生しない。しかし、接続されているコネクション数が  $m = \lfloor C/b_r \rfloor$  を越えた時は、帯域が等しく割り当てられるとすれば、

すべてのコネクションにおいて必要な帯域は得られず、通信品質は0となる。この場合、ユーザは通信を途中で終了してしまうことも考えられるが、ここではそのまま通信を継続するものとする。このようなモデル化により、対象とするシステムはサーバ数無限大の  $M/M/\infty$  待ち行列システムでモデル化でき、定常状態においてシステム内コネクション数が  $k$  である確率  $p(k)$  は

$$p(k) = \frac{e^{-a} a^k}{k!}$$

になる。従って、ユーザ効用  $U_{r,b}$  (Utility for Rigid application in Best-effort service) は

$$U_{r,b}(C) = \frac{1}{a} \sum_{k=0}^m k p(k) \alpha(b_r)$$

として定義できる。

ただし、上式において得られる効用は、コネクションを通じた平均的な値である。実際には、あるコネクションの通信途中で全体のコネクション数が増大し、通信品質が劣化する場合もある。そこで、そのような状況も考慮した評価尺度として、最悪効用を導入する。最悪効用は、コネクションの通信中における最悪の通信品質を定義する。すなわち、最悪効用  $WU_{r,b}$  (Worst Utility for Rigid application in Best-effort service) は以下のようにして導出できる。

まず、 $M/M/\infty$  待ち行列システムにおいて、あるコネクションが到着した時に他のコネクション数を  $k$  とし、そのコネクションの接続が終了するまでの間に、それ以外のコネクション数が高々  $n$  になる確率  $r_n(k)$  を考える。 $r_n(k)$  ( $k = 0, 1, \dots, n$ ) を要素とする列ベクトル  $\mathbf{r}_n$  は以下の式で与えられる。

$$\mathbf{r}_n = (\mathbf{I} - \mathbf{G}/\mu)^{-1} \mathbf{e} \quad (2)$$

ただし、 $\mathbf{e}$  はすべての要素が1の列ベクトルである。また、 $\mathbf{G}$  は  $(n+1)$  行  $(n+1)$  列の行列で、 $\exp(\mathbf{G} \cdot t)$  の  $(i, j)$  要素は、時間0で状態が  $p(i)$  にある時、時刻  $t$  までに、他のコネクション数が  $n$  以下、かつ、時刻  $t$  においてコネクション数が  $j$  である確率を与える。 $\mathbf{G}$  は、

次の式で与えられる。

$$\mathbf{G} = \begin{bmatrix} -\lambda & \lambda & & 0 \\ \mu & -(\lambda + \mu) & \lambda & \\ \ddots & \ddots & \ddots & \\ & (n-1)\mu & -(\lambda + (n-1)\mu) & \lambda \\ 0 & & n\mu & -(\lambda + n\mu) \end{bmatrix} \quad (3)$$

今、 $m$  以上のコネクションが同時にサービスされた場合、効用は 0 になると仮定しているので、あるコネクションがシステムに到着した時のコネクション数が  $m-1$  以下、かつ、そのコネクションがサービスを受けている間にコネクション数が  $m$  を越えなければ、そのコネクションは常に通信品質  $\alpha(b_r)$  を得られる。すなわち、 $WU_{r,b}$  は次の式で与えられる。

$$WU_{r,b}(C) = \sum_{k=0}^{m-1} r_{m-1}(k) p(k) \alpha(b_r)$$

なお、予約型サービスの場合は、いったんコネクションが受け付けられると通信品質は保証されるので、最悪効用は効用  $U_{r,r}$  に等しい。

### 3.3 適応型アプリケーション

#### 3.3.1 予約型サービスに適用した場合

適応型制御機能を持つアプリケーションを予約型サービスに適用する場合、以下のようなシナリオを考える。コネクションは通信開始時に、帯域  $b_{min}$ 、 $b_{max}$  をパラメータとしてネットワークに要求する。ネットワークは帯域  $b_{max}$  を保証可能な場合は、そのコネクションをそのまま受け付ける。一方、 $b_{max}$  を保証できない場合にも、他のコネクションに対して割り当てた帯域を減らせば、少なくとも最低帯域  $b_{min}$  を保証できる場合には、すでに割り当てられているコネクションに対しても割り当て帯域を減らす制御を行うことによって新しいコネクションを受け付ける。すなわち、すでに接続されているコネクション数を  $k$  とした場合、

$$b_{min} \leq C/(k+1) < b_{max}$$

であれば、割り当て帯域を  $C/(k+1)$  として受け付ける。この場合、パラメータ再交渉のメカニズムが準備されていることが前提で [5]、すでに接続されているコネクションに対しても割

当帯域を  $C/(k+1)$  とする。一方、そのような最低帯域  $b_{min}$  も保証できない場合は呼損とする。

このシステムにおいては、同時接続可能な最大コネクション数は  $m_2 = \lfloor C/b_{min} \rfloor$  となり、サーバ数が  $m_2$  である M/M/ $m_2/m_2$  待ち行列システムとしてモデル化できる。ただし、接続コネクション数が  $m_1 = \lfloor C/b_{max} \rfloor$  から  $m_2$  の間、各コネクションに対して割り当てられる帯域はコネクション数によって決定されるものとし、通信品質も割当帯域によって定められる QoS 関数  $\beta(\cdot)$  に応じて変動する。

回線容量  $C$  に対するコネクションの平均効用  $U_{a,r}$  (Utility for Adaptive application in Reservation-based service) は、以下のようにして与えられる。

$$U_{a,r}(C) = \frac{1}{a} \left\{ \sum_{k=0}^{m_1} k q(k) \beta(b_{max}) + \sum_{k=m_1+1}^{m_2} k q(k) \beta(C/k) \right\}$$

ただし、 $q(k)$  は同時接続コネクション数が  $k$  である定常状態確率であり、

$$q(k) = \frac{a^k/k!}{\sum_{j=0}^{m_2} a^j/j!}$$

である。

一方、最悪効用  $WU_{a,r}$  (Worst Utility for Adaptive application in Reservation-based service) については、割当帯域がコネクション数に応じて変動し、その結果、得られる通信品質も変動することを考慮する必要がある。そのためには、ある着目するコネクションが終了するまでの間に、他のコネクション数の最大が  $n$  になる確率  $s_n(k)$  を求めればよい。すなわち、この時、QoS は、 $\beta(C/n+1)$  となる。ただし、 $k$  はコネクション設定時の他のコネクション数を表す。 $s_n(k)$  は、前節の式 (2) に示した  $r_n(k)$  を、サーバ数有限とした  $\bar{r}_n(k)$  によって、以下の式で与えられる。

$$\begin{aligned} s_n(k) &= \bar{r}_n(k) - \bar{r}_{n-1}(k), & n > k \\ s_n(n) &= \bar{r}_n(n) \end{aligned}$$

ただし、サーバ数有限の場合、式 (3) で与えられる  $G$  の  $(n+1, n+1)$  要素が  $-n\mu$  に変わる。

$s_n(k)$  を用いて最悪効用  $WU_{a,r}$  は以下の式で与えられる。

$$WU_{a,r}(C) = \sum_{n=0}^{m_1-1} \sum_{k=0}^n s_n(k) q(k) \beta(b_{max}) + \sum_{n=m_1}^{m_2-1} \sum_{k=0}^n s_n(k) q(k) \beta(C/(n+1))$$

上式は、あるコネクションが受け付けられた時から終了するまで、他のコネクション数  $n$  が常に  $m_1 - 1$  以下であれば、その通信品質は  $\beta(b_{max})$  であり、一方、 $m_1$  以上  $m_2 - 1$  以下であれば、通信品質は  $\beta(C/(n + 1))$  であることを示している。

### 3.3.2 ベストエフォート型サービスに適用した場合

ベストエフォート型サービスに対して適応型アプリケーションを適用した場合は、非適応型アプリケーションの場合と同じく M/M/∞ 待ち行列システムでモデル化できる。ただし、ユーザ効用がコネクション数に応じて変化するところが異なる。

ユーザ効用  $U_{a,b}$  (Utility for Adaptive application in Best-effort service) は、

$$U_{a,b}(C) = \frac{1}{a} \left\{ \sum_{k=0}^{m_1} k p(k) \beta(b_{max}) + \sum_{k=m_1+1}^{m_2} k p(k) \beta(C/k) \right\}$$

となる。

また、前節と同様に、 $s_n(k)$  に当る確率を  $t_n(k)$  で表すものとする、 $t_n(k)$  を式 (2) に示した  $r_n(k)$  を用いて以下の式で与えられる。

$$t_n(k) = r_n(k) - r_{n-1}(k), \quad n > k$$

$$t_n(n) = r_n(n)$$

従って、最悪効用  $WU_{a,b}$  (Worst Utility for Adaptive application in Best-effort service) は、

$$WU_{a,b}(C) = \sum_{n=0}^{m_1-1} \sum_{k=0}^n t_n(k) p(k) \beta(b_{max}) + \sum_{n=m_1}^{m_2-1} \sum_{k=0}^n t_n(k) p(k) \beta(C/(n + 1))$$

で与えられる。

### 3.4 比較評価

以下、予約型サービスとベストエフォート型サービスの比較を行うため、2つのサービスによって得られる平均効用の差

$$D_y(C) = U_{y,r}(C) - U_{y,b}(C)$$

および、最悪効用の差

$$WD_y(C) = WU_{y,r}(C) - WU_{y,b}(C)$$

を用いて評価する。ただし、文字  $y$  は  $r$  (非適応型)、 $a$  (適応型) を示す。実装コストなどを考えない場合には、効用差が大きいほど予約型サービスが望ましいことになる。

また、非適応型、適応型アプリケーションの QoS 関数  $\alpha(\cdot)$ 、 $\beta(\cdot)$  として、音声アプリケーションを想定して図 3、4 で示すような関数を考える [14, 15]。比較に際しては、ネットワークへの負荷 ( $a = \lambda/\mu$ ) を 20 アーランに固定し、回線容量  $C$  を変更した。

まず、非適応型アプリケーションについての結果を図 5 に示す。図 5 において、平均効用  $U_{r,r}$ 、 $U_{r,b}$ 、および、その効用差  $D_r$  を見ると、回線容量  $C$  が小さい場合には、ベストエフォート型サービスのユーザ効用  $U_{r,b}$  がほぼ 0 になるため、平均効用差  $D_r$  は大きな値を示している。一方、回線容量  $C$  が大きくなると、ベストエフォート型サービスのユーザ効用  $U_{r,b}$  も大きくなるために効用差  $D_r$  が小さくなるが、その値は決して小さな値ではなく、非適応型アプリケーションに対しては予約型サービスが有効であるといえる。

また、図 5 には、最悪効用  $WU_{r,r}$  ( $= U_{r,r}$ )、 $WU_{r,b}$  とその効用差  $WD_r$  も示している。ベストエフォート型サービスにおいて、最悪効用  $WU_{r,b}$  は平均効用  $U_{r,b}$  よりかなり減少し、その結果、最悪効用差  $WD_r$  が増加していることが分かる。これは、ベストエフォート型サービスにおいては、コネクション数の変動によって通信途中の品質劣化があるためであり、予約型サービスの有効性が示されているといえる。

次に、適応型アプリケーションに関する結果を図 6 に示す。全体の傾向は先の図 5 と同じであるが、ベストエフォート型サービスにおける通信品質が大幅に改善され、その結果、平均効用差  $D_a$ 、最悪効用差  $WD_a$  がかなり小さくなっていることがわかる。ただし、回線容量が小さい場合は依然ベストエフォート型サービスの効用はほとんど 0 に近く、その場合、予約型サービスの提供は必須であることもわかる。なお、回線容量が大きい場合、ベストエフォート型サービスでも十分適用可能であるように見えるが、その前提は適切な回線容量が準備されていることであり、今後、ベストエフォート型サービスにおいて必要な回線容量を把握することが必須である [16]。

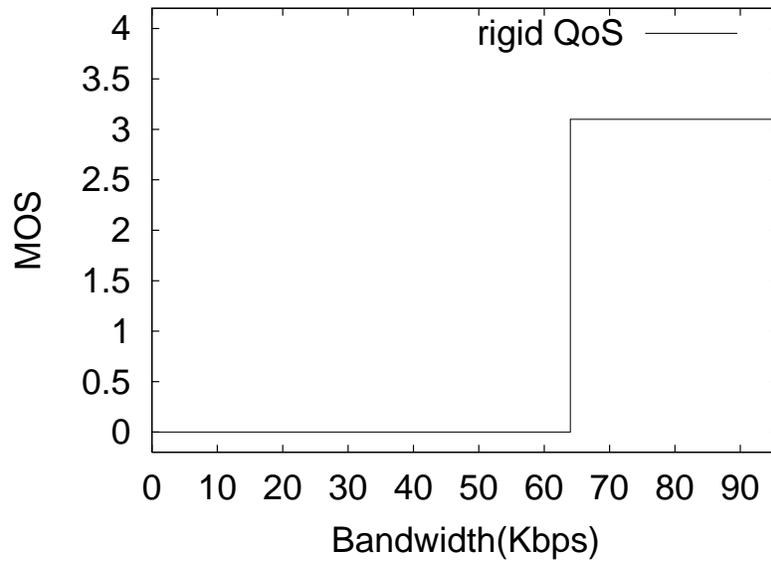


図 3: QoS 関数  $\alpha(\cdot)$

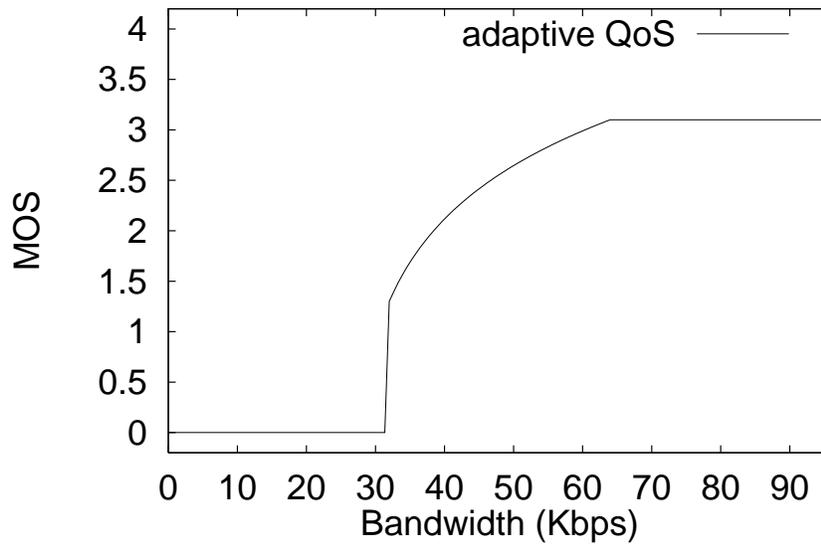


図 4: QoS 関数  $\beta(\cdot)$

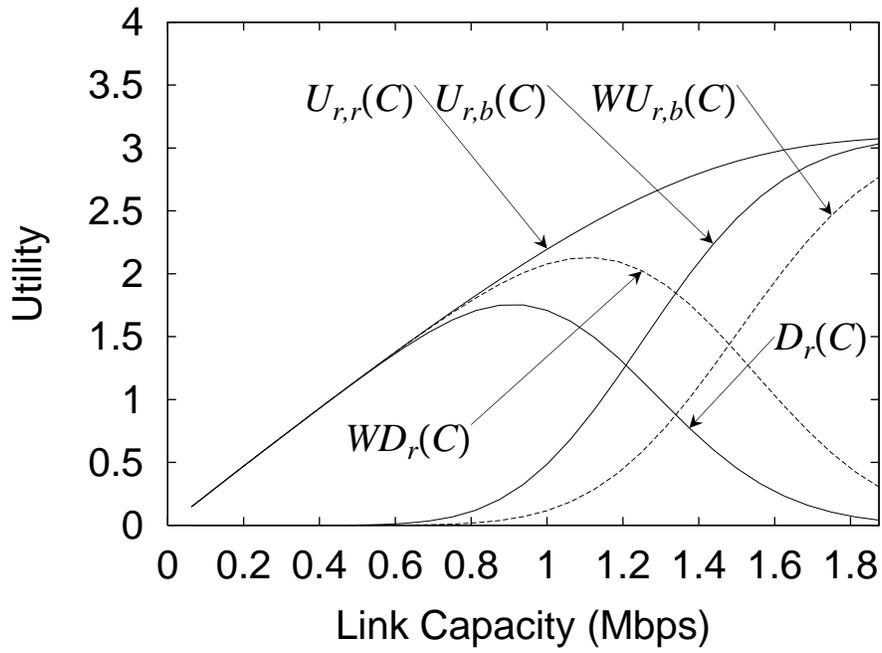


図 5: 非適応型アプリケーションの平均、最悪効用

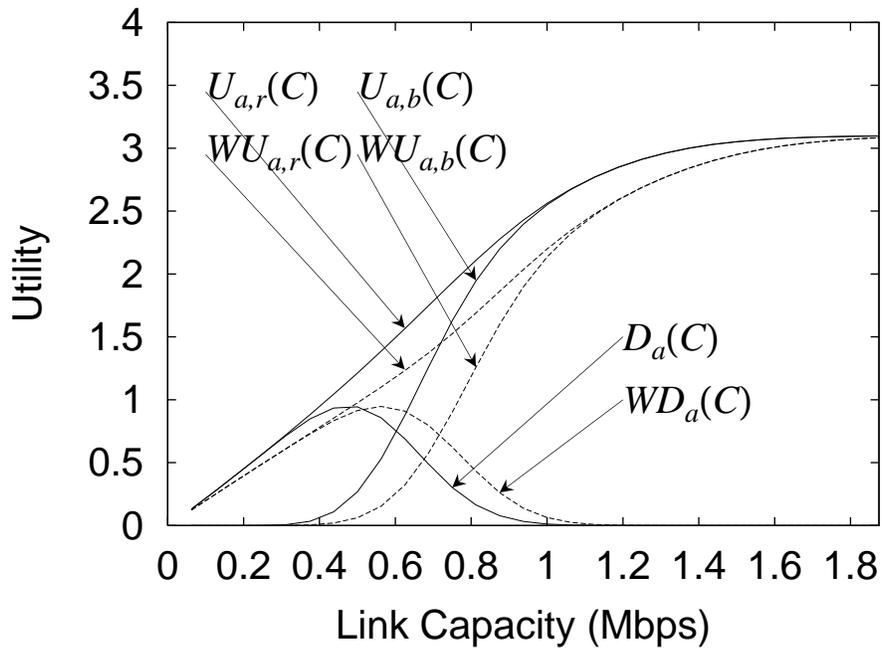


図 6: 適応型アプリケーションの平均、最悪効用

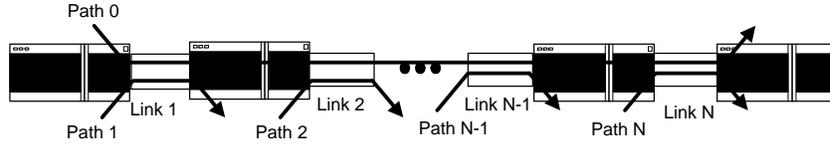


図 7: 拡張モデル

## 4 多段モデルへの拡張

### 4.1 モデル化

今、図 7 に示すように、 $N$  段のノードからなるモデルを考え、1つのパスのみ  $N$  ホップ (パス 0 とする) とし、他のパス  $i$  ( $1 \leq i \leq N$ ) はそれぞれ 1 ホップでパス 0 と回線を共有するとする。すなわち、パス数は全体で  $N + 1$  である。パス  $i$  におけるコネクションは、ポアソン到着 (平均  $\lambda_i$ ) で到着し、指数分布の継続時間 (平均  $1/\mu_i$ ) とする。また、パス  $i$  の負荷を  $a_i (= \lambda_i/\mu_i)$  とする。回線容量はすべて  $C$  とする。先と同様、コネクション設定時間については、コネクション接続時間に比べて十分小さいとして無視する。

以下では、 $N$  段のリンクを経由するパス 0 のユーザ効用のみを導出していく。

### 4.2 非適応型アプリケーション

#### 4.2.1 予約型サービスに適用した場合

各コネクションは帯域  $b_r$  を要求する。パス 0 上のコネクションに対しては、すべての回線上で帯域  $b_r$  が確保できる場合のみコネクション接続が受け付けられる。各回線ともコネクション数  $m (= \lfloor C/b_r \rfloor)$  まで同時受け付け可能であり、それ以上の接続要求は呼損となる。パス  $i$  ( $0 \leq i \leq N$ ) 上のコネクション数を  $u_i$  とすると、定常状態確率  $P_{u_0, \dots, u_N}$  は、以下の式で与えられる。

$$P_{u_0, \dots, u_N} = P_{0, \dots, 0} \prod_{i=0}^N \frac{a_i^{u_i}}{u_i!} \quad (4)$$

ただし、 $P_{0, \dots, 0}$  は以下の式で与えられる。

$$P_{0, \dots, 0} = 1 / \sum_{u_0=0}^m \left[ \frac{a_0^{u_0}}{u_0!} \times \prod_{i=1}^N \left( \sum_{u_i=0}^{m-u_0} \frac{a_i^{u_i}}{u_i!} \right) \right] \quad (5)$$

パス0上のコネクションが呼損となるのは少なくとも1つの回線におけるコネクション数が  $m$  に等しい場合である。すなわち、呼損率  $LL_{r,r}$  は、

$$LL_{r,r}(C) = \sum_{\{u_0, \dots, u_N; u_0 + u_i = m (1 \leq i \leq N)\}} P_{u_0, \dots, u_N}$$

で与えられる。

以上より、パス0のユーザ効用  $LU_{r,r}$  は、3.2.1節の導出と同様に、

$$LU_{r,r}(C) = (1 - LL_{r,r}(C))\alpha(b_r)$$

で与えられる。

#### 4.2.2 ベストエフォート型サービスに適用した場合

ベストエフォート型サービスではコネクションの接続はすべて処理されるものとする。従って、各回線上のコネクション数に関する状態を独立に考えることによって、各コネクション数が  $u_0, \dots, u_N$  である状態確率を求められる。すなわち、定常状態確率  $Q_{u_0, \dots, u_N}$  は以下の式で与えられる。

$$Q_{u_0, \dots, u_N} = \prod_{i=0}^N p_i(u_i)$$

ただし、 $p_i(u_i)$  は以下の式で与えられる。

$$p_i(u_i) = \frac{e^{-a_i} a_i^{u_i}}{u_i!}$$

パス0上のコネクションが使用できる帯域は、パス  $i$  ( $1 \leq i \leq N$ ) 上のコネクションの接続数によって決まる。すなわち、どれか1つの回線上で接続数が  $m$  より大きい場合、パス0のコネクションは帯域  $b_r$  より小さい帯域で通信が行われることになり、通信品質は0となる。従って、パス0のユーザ効用  $LU_{r,b}$  は、

$$LU_{r,b}(C) = \frac{1}{a_0} \sum_{u_0=0}^m \sum_{u_1=0}^{m-u_0} \cdots \sum_{u_N=0}^{m-u_0} u_0 Q_{u_0, \dots, u_N} \alpha(b_r)$$

### 4.3 適応型アプリケーション

#### 4.3.1 予約型サービスに適用した場合

本節では、適応型アプリケーションを予約型サービスで使用した場合を考える。この場合、各回線はコネクション数  $m_2 (= \lfloor C/b_{min} \rfloor)$  まで同時接続可能である。従って、定常状態確率  $R_{u_0, \dots, u_N}$  は、式 (4)、(5) における  $m$  を  $m_2$  に置き換えることによって求められる。

各状態  $R_{u_0, \dots, u_N}$  における各コネクションに対する帯域割当には、さまざまなものが考えられるが、ここでは、Max-Min フロー制御 [17] に基づく帯域割当がシグナリング制御によって実現されていると仮定する。すなわち、 $u_0, \dots, u_N$  の中で最大の値を  $u_{max} = \max\{u_0, \dots, u_N\}$  とすると、接続コネクション数が最大である回線において、各コネクションの割当帯域は回線容量  $C$  をコネクション数  $u_0 + u_{max}$  で等分した値になる。従って、パス 0 上の各コネクションの割当帯域は  $C/(u_0 + u_{max})$  である。パス  $i$  ( $i = 1, \dots, N$ ) 上のコネクションに対しては、残りの帯域を等分に割り当てることになる。すなわち、その割当帯域は  $(C - \frac{u_0 C}{u_0 + u_{max}})/u_i$  である。

以上より、パス 0 のユーザの通信品質は  $\beta(C/(u_0 + u_{max}))$  で与えられる。すなわち、パス 0 上のコネクションのユーザ効用  $LU_{a,r}$  は、

$$LU_{a,r}(C) = \frac{1}{a_0} \sum_{u_0=0}^{m_2} \sum_{u_1=0}^{m_2-u_0} \cdots \sum_{u_N=0}^{m_2-u_0} u_0 R_{u_0, \dots, u_N} \beta\left(\frac{C}{u_0 + u_{max}}\right)$$

#### 4.3.2 ベストエフォート型サービスに適用した場合

最後に、適応型アプリケーションをベストエフォート型サービスに適用した場合を考える。議論の簡単化のため、ここでは、適応型アプリケーションはフィードバック情報に従ってレート制御を行い、その結果、送信レートが Max-Min フロー制御によって決定される割当帯域に等しくなるようなレート制御を行う機構を持つと仮定する。定常状態確率は非適応型アプリケーションの場合と同じになり、通信品質は予約型サービスの場合と同様に  $\beta(C/(u_1 + u_{max}))$  で与えられる。従って、パス 0 のユーザ効用  $LU_{a,b}$  は、

$$LU_{a,b}(C) = \frac{1}{a_0} \times \sum_{u_0=1}^{\infty} \cdots \sum_{u_N=1}^{\infty} u_0 Q_{u_0, \dots, u_N} \beta(C/(u_N + u_{max}))$$

#### 4.4 比較評価

本節では多段モデルにおける数値例を示す。そのために以下では、パス 0 上の接続の平均効用の効用差

$$LD_y(C) = LU_{y,r}(C) - LU_{y,b}(C), \quad y = r \text{ or } a$$

を示す。以下の数値例では、パス  $i$  ( $0 \leq i \leq N$ ) の負荷 ( $a_i = \lambda_i / \mu_i$ ) はすべて等しく 10 アーランとし、回線容量  $C$  は各回線とも等しい値に設定し、段数を変化させている。

まず、図 8 に、回線容量  $C$  を 0.5、1.0、1.5、2 Mbps とした場合の非適応型アプリケーション (QoS 関数については図 3 参照) のユーザ効用  $LU_{r,r}$ 、 $LU_{r,b}$  を示す。また、その効用差  $LD_r$  を図 9 に示す。なお、図におけるリンク数 1 の場合が前章に示した結果と一致する。すなわち、図 9 からわかるように、回線容量が大きくなる (回線容量  $C$  が 0.5 Mbps から 1.5 Mbps) につれて効用差は増大するが、ある程度以上大きくなるとベストエフォート型サービスにおける効用の増大によってその差は小さくなった (図では回線容量  $C = 2$  Mbps)。

図 9 に示すように、中継リンク数が増大してもその傾向は変わらないが、効用差が最大となる容量が変化する。これは、容量が小さい場合にはリンク数の増加によって予約型サービスにおける呼損率が大きくなり、その結果、効用が減少し、ベストエフォート型サービスとの差が小さくなるためである (図 8 参照)。ただし、回線容量が十分に大きい場合 (図では  $C = 2$  Mbps)、効用差  $LD_r$  は小さな値を示しており、回線容量が負荷に対して十分あれば、回線数が増加してもどちらのサービスもその効用がほとんど変化しない、すなわち、前節の結果と同様にベストエフォート型サービスでも十分な効用が得られるといえる。

図 10 は、回線容量  $C$  を 0.25、0.5、0.75、1 Mbps とした場合の適応型アプリケーションの効用  $LU_{a,r}$ 、 $LU_{a,b}$  を示している。また、その効用差  $LD_a$  を図 11 に示す。図 10、11 より、非適応型アプリケーションの場合と同じ傾向が見られるが、その効用差は非適応型アプリケーションの場合に比べて小さな値を示している。すなわち、回線容量が十分であればベストエフォート型サービスでもよいといえるが、その前提は接続到着率 (呼量) が把握でき、それに見合った十分な回線が用意できることである。現実にはそれは困難であり、その場合は予約型サービスが必須であるといえる。

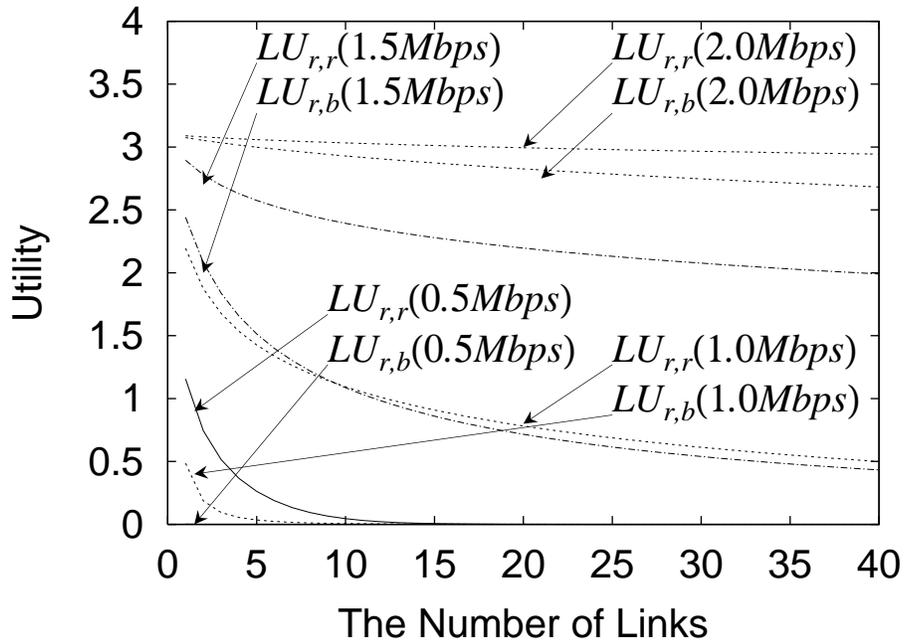


図 8: 非適応型アプリケーションの平均効用

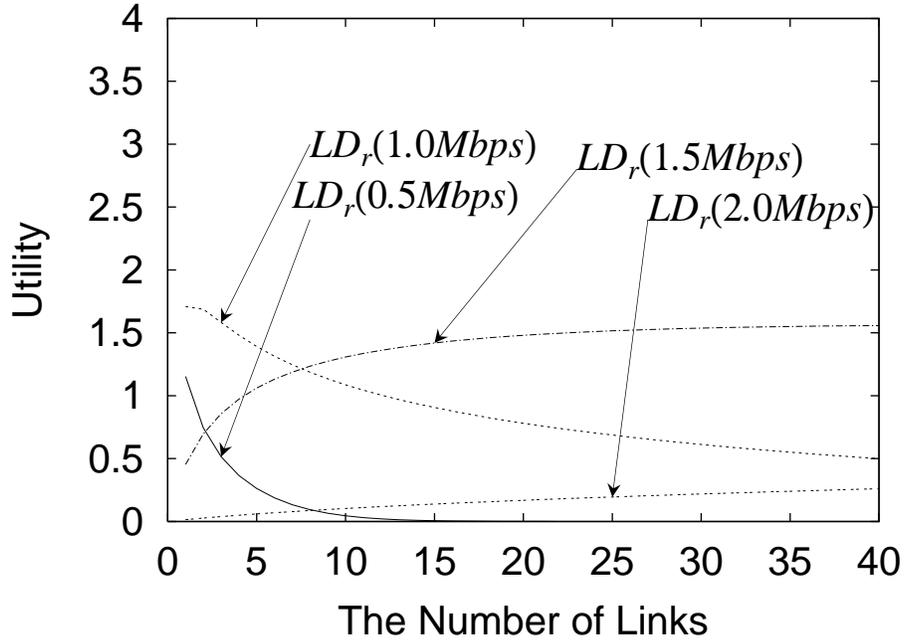


図 9: 非適応型アプリケーションの効用差

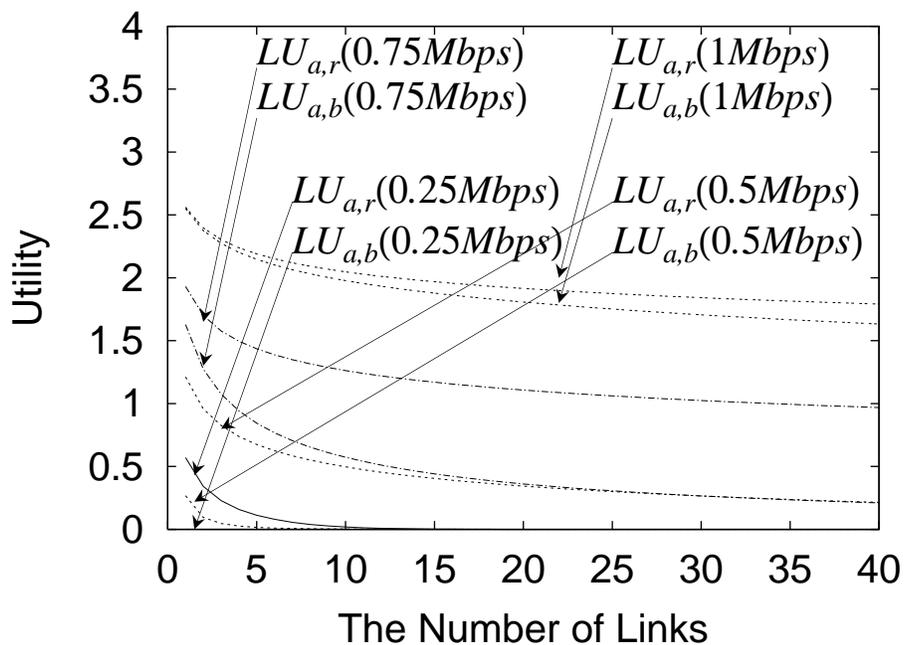


図 10: 適応型アプリケーションの平均効用

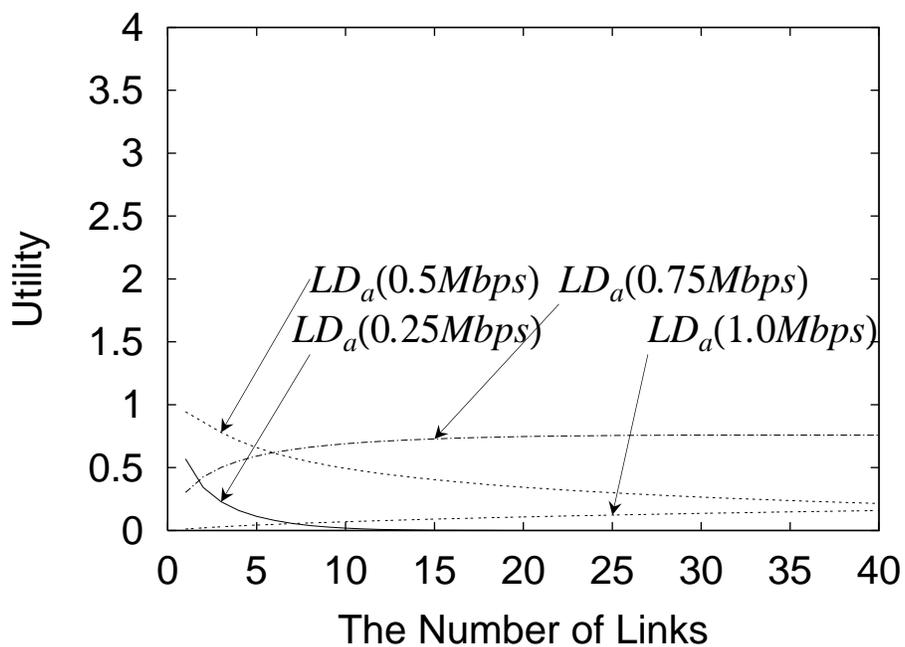


図 11: 適応型アプリケーションの効用差

## 5 料金制度を考慮した評価

本章では、実時間トラヒックと Web トラヒックが混在した時に、それらをベストエフォート型サービスに適用した場合、帯域予約型サービスに適用された場合のそれぞれについてユーザ効用を求め、その比較を行う。

まず、ネットワークモデルとして図 12 に示すモデルを考える。すなわち、ISP は、容量  $C$  を持つアクセス回線によりバックボーンネットワークに接続し、ユーザに対して電話回線を通じて  $N$  回線まで接続可能なアクセスポイントを持っているとする。すなわち、最大  $N$  人までのユーザの同時接続を可能としている。

### 5.1 モデル化

#### 5.1.1 ベストエフォート型サービスに適用した場合

ユーザには、2 種類のユーザ集合があると考え。すなわち、ユーザ集合 A を実時間アプリケーションを使用するユーザの集合とし、ユーザ集合 B をデータ系アプリケーション (Web) を使用するユーザの集合と考える (図 12)。ここでは議論の簡単化のため、集合 A のユーザは、実時間アプリケーションのみの使用し、集合 B のユーザは、Web アプリケーションのみを使用することを前提とする。まず、ユーザ集合 A は、ユーザ数  $n_A$  からなり、各ユーザからの接続要求は平均  $\lambda_A$  のポアソン分布で発生し、その接続時間は平均  $1/\mu_A$  の指数分布で与えられるものとする。実時間通信を行う時間はプロバイダとの接続時間に等しいと仮定する。ユーザ集合 B はユーザ数  $n_B$  からなり、その接続要求発生率、接続時間は、 $\lambda_B$  のポアソン分布、平均  $1/\mu_B$  の指数分布で与えられるものとする。

今、ユーザ集合 A の接続数を  $i$ 、ユーザ集合 B の接続数を  $j$  とした場合の状態を  $s_{i,j}$  とあらわす。 $s_{i,j}$  の定常状態確率を  $P_{i,j}$  とすると、 $P_{i,j}$  は、

$$P_{i,j} = P_{0,0} \times \frac{a_A^i a_B^j}{i! j!}$$
$$P_{0,0} = 1 / \sum_{i=0}^N \left[ \frac{a_A^i}{i!} \times \sum_{j=0}^{N-i} \frac{a_B^j}{j!} \right]$$

となる。ここで、 $a_A = n_A \times a_{one,A} (= \lambda_A / \mu_A)$ 、 $a_B = n_B \times a_{one,B} (= \lambda_B / \mu_B)$  である。状態  $s_{i,j}$  における実時間アプリケーション、Web トラヒックの QoS を考える。ベストエフォート

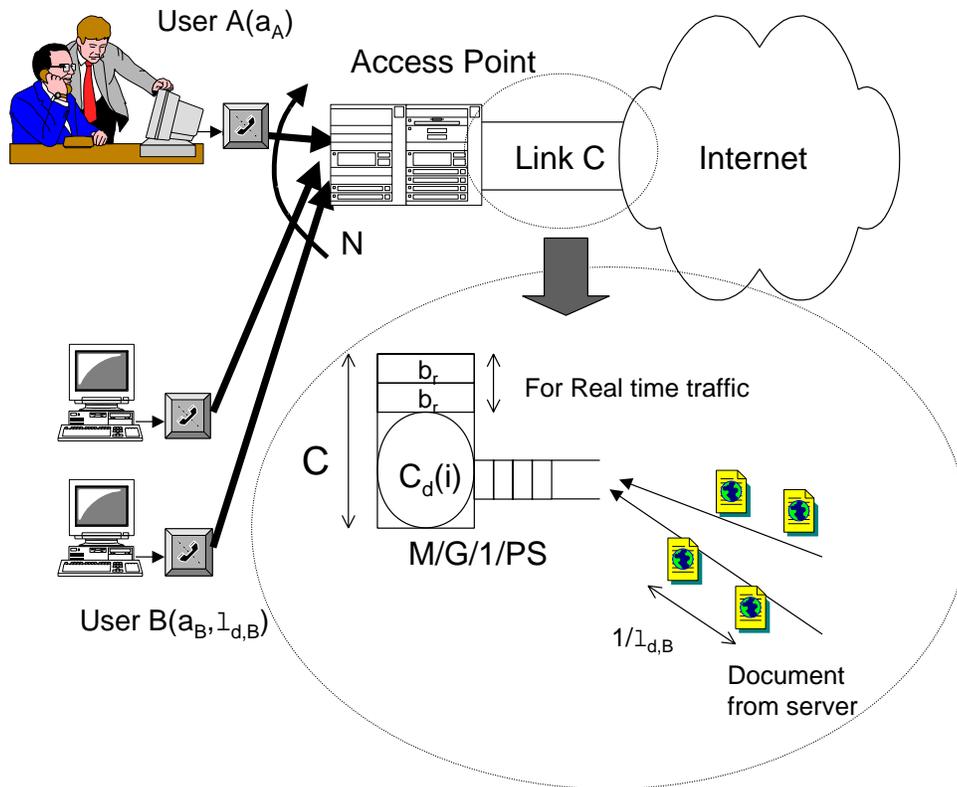


図 12: ネットワークモデル 2

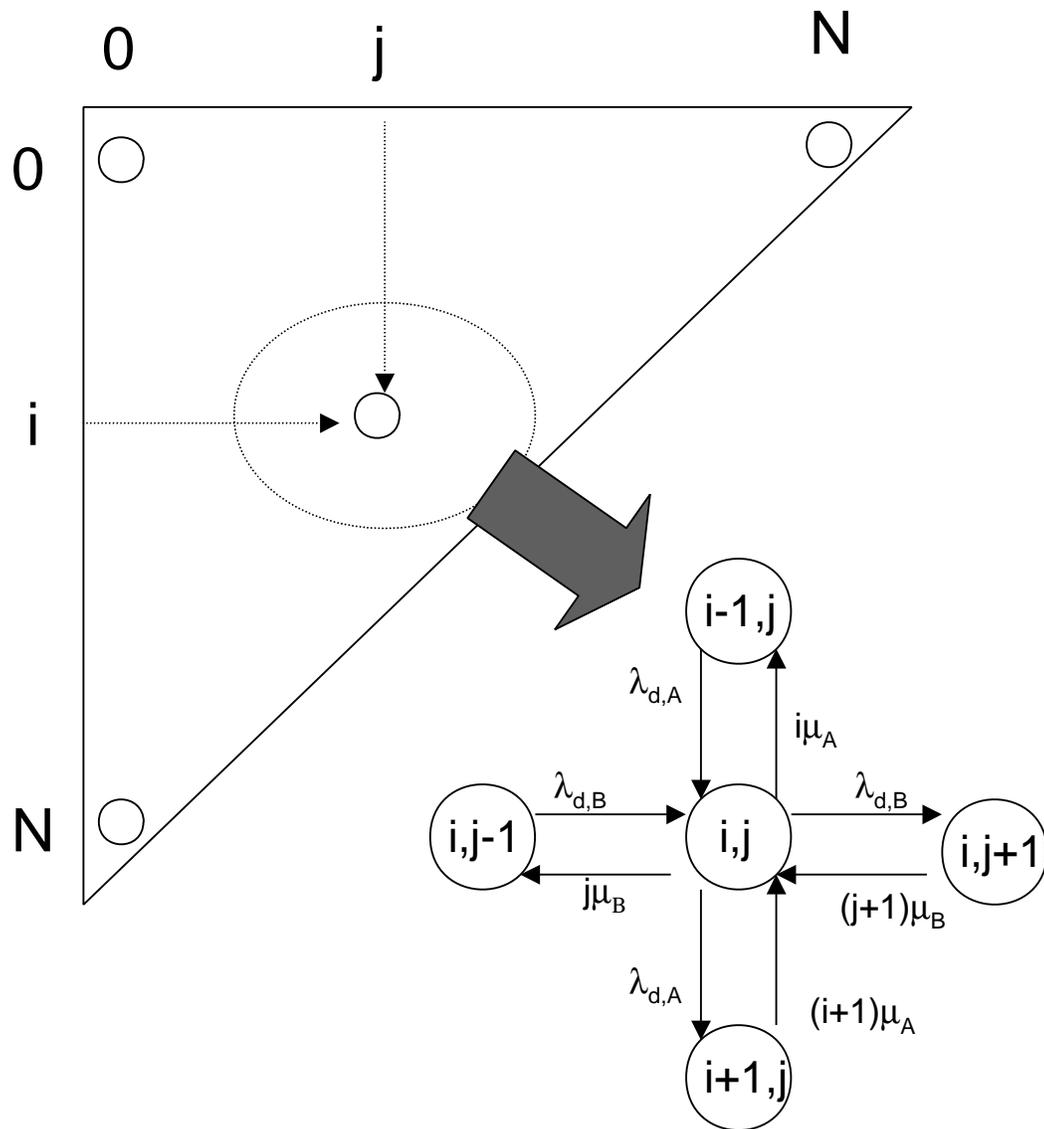


图 13: 状态迁移图 1

ト型サービスでは、実時間トラヒックはUDPにより、またWebトラヒックはTCPにより通信していると考え。UDPが存在する場合、TCPはそのウインドウフロー制御によって送信レートが減少するため [18]、ここでは実時間トラヒックは通信品質を満足できる帯域  $b_r$  で送信を行い、その未使用の帯域をWebトラヒックが使用するものとする [19]。ただし、実時間トラヒックの総帯域がアクセス回線容量を越えた場合、そのQoSは0になるものとする。すべての実時間アプリケーションが同じ帯域  $b_r$  を使用すると仮定すると、実時間アプリケーションを使用するユーザ数が  $i$  の時に、Webトラヒックが使用できる帯域  $C_d(i)$  は以下の式で与えられる。

$$C_d(i) = \max\{C - i \times b_r, 0\}$$

ユーザBはISPに接続中にドキュメント転送要求を行うと考える。httpによるドキュメント転送要求は、M/G(log-normal)/1/PSでモデル化する [20] (図12参照)。ただし、ここではアクセス回線のみを評価の対象としている。ユーザ集合Bのドキュメント転送要求は平均  $\lambda_{d,B}$  のポアソン分布に従って発生する。また、要求される転送ドキュメントの長さは、その平均値を  $X_m$  とする対数正規分布で表せるものとする。

これより、 $s_{i,j}$  におけるドキュメント転送遅延時間  $D_{i,j}$  は、

$$D_{i,j} = \frac{X_m}{C_d(i) - j\lambda_{d,B}X_m}$$

となる。従って、Webトラヒックの平均  $QoS_{w,b}$  は以下の式で与えられる。

$$QoS_{w,b} = \sum_{i=0}^N \sum_{j=0}^{N-i} P_{i,j} \gamma(D_{i,j})$$

次に実時間アプリケーションユーザのQoSを求める。実時間アプリケーションは、帯域  $b_r$  を使用できれば、すなわち  $i \leq m = \lfloor C/b_r \rfloor$  であれば、 $\alpha(b_r)$  を得ることができ、一方、 $i > m$  であれば、十分な帯域が確保できないためにQoSは0とする。従って、平均  $QoS_{r,b}$  は以下の式で与えられる。

$$QoS_{r,b} = \frac{1}{a_A} \sum_{i=0}^m \sum_{j=0}^{N-i} i P_{i,j} \alpha(b_r)$$

実時間トラヒック、Webトラヒックのユーザ効用  $U_{r,b}$ 、 $U_{w,b}$  は、そのコストを考慮することによって、以下の式で与えられる。

$$U_{r,b} = h a_A Q_0 S_{r,b} - Cost_A(h a_A)$$

$$U_{w,b} = h a_B \lambda_{d,B} Q_0 S_{w,b} - Cost_B(h a_B)$$

### 5.1.2 資源予約型サービスに適用した場合

ベストエフォート型サービスと同様に、集合 A を実時間アプリケーションのユーザ集合、集合 B を Web のユーザ集合とする。資源予約型サービスにおいては、帯域予約を行うユーザは、通信開始時にあらかじめ帯域予約を要求する。すなわち、接続要求時に帯域  $b_r$  を申告し、ネットワークが帯域  $b_r$  を確保できるならば、その接続を受け付ける。ここでは、その帯域予約の対象は ISP のバックボーンネットワークへのアクセス回線のみを考える。ただし、たとえ帯域  $b_r$  を確保できる状態であっても、Webトラヒックのユーザを含めて最大接続数  $N$  によって抑えられている可能性はある。ここで、Web ユーザは、帯域予約されていない帯域を利用するとしている。ただし、実時間アプリケーションがその帯域予約によってすべての帯域を使用してしまわないように、Webトラヒックの保護帯域として最低帯域  $C_{dr}$  ( $\leq C$ ) を確保すると考える。よって、同時接続可能な実時間アプリケーションユーザ数は、最大  $m_2 = \lfloor (C - C_{dr})/b_r \rfloor$  となる。

まず、実時間アプリケーションのユーザ数が  $i$ 、Web のユーザ数が  $j$  である状態を  $u_{i,j}$  とすると、その状態遷移図は図 14 のようになる。 $u_{i,j}$  の定常状態確率  $Q_{i,j}$  は、

$$Q_{i,j} = Q_{0,0} \times \frac{a_A^i a_B^j}{i! j!}$$

$$Q_{0,0} = 1 / \sum_{i=0}^{m_2} \left[ \frac{a_A^i}{i!} \times \sum_{j=0}^{N-i} \frac{a_B^j}{j!} \right]$$

となる。また、状態  $u_{i,j}$  において Webトラヒックが使用できる帯域  $C_d(i)$  は、

$$C_d(i) = \max\{C - i \times b_r, C_{dr}\}$$

となる。状態  $u_{i,j}$  におけるドキュメント転送遅延時間  $D_{i,j}$  は、

$$D_{i,j} = \frac{X_m}{C_d(i) - j \lambda_{d,B} X_m}$$

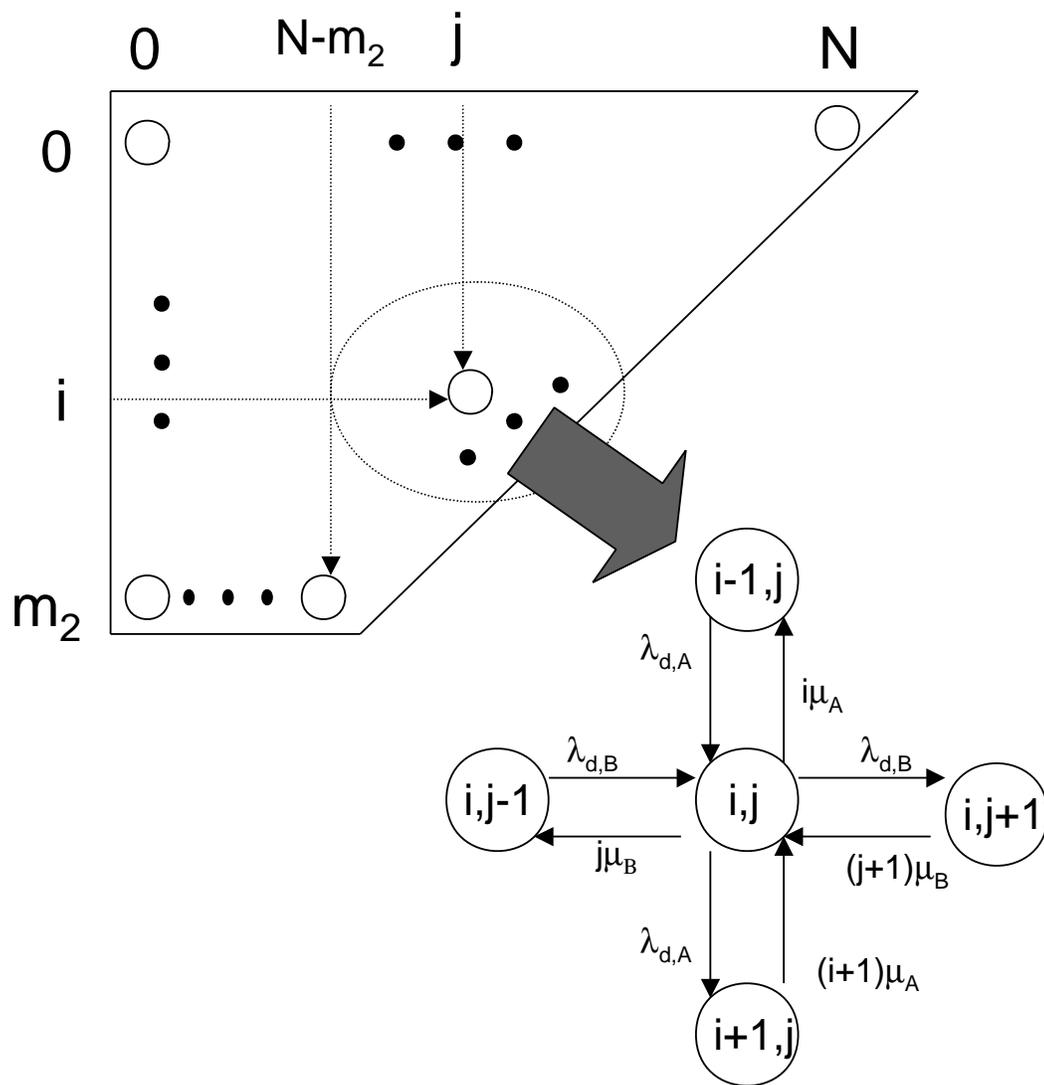


图 14: 状态迁移图 1

となり、Webトラヒックの平均  $QoS_{w,r}$  は以下の式で与えられる。

$$QoS_{w,r} = \sum_{i=0}^{m_2} \sum_{j=0}^{N-i} Q_{i,j} \gamma(D_{i,j})$$

次に実時間アプリケーションの  $QoS$  を求める。実時間アプリケーションは、接続されれば帯域  $b_r$  を使用できるので、 $QoS$  値として  $\alpha(b_r)$  を得ることができる。従って、実時間アプリケーションの呼損率を  $Pb_r$  とすると、平均  $QoS_{r,r}$  は以下の式となる。

$$QoS_{r,r} = (1 - Pb_r) \alpha(b_r)$$

ただし、 $Pb_r$  は状態確率  $u_{i,j}$  ( $i = m_2$  or  $i + j = N$ ) の総和である。すなわち、以下の式で与えられる。

$$Pb_r = \sum_{j=0}^{N-m_2} Q_{m_2,j} + \sum_{i=0}^{m_2-1} Q_{i,N-i}$$

帯域使用料金は、ユーザが使用した時間、および、帯域量に比例すると考える。すなわち、比例係数を  $p_{bw}$  とすると、各ユーザは  $p_{bw} \times a_A h \times b_r$  のコストが必要となる。従って、ユーザ効用は以下の式で与えられる。

$$U_{r,r} = h a_A QoS_{r,r} - (1 - Pb_r) h a_A p_{bw} b_r$$

$$U_{w,r} = h a_B \lambda_{d,B} QoS_{w,r} - Cost(h a_B)$$

## 5.2 比較評価

本節では、前節で求めた解析結果に基づいた評価例を示す。ここでは、実時間トラヒック量を変化させることによって、ベストエフォート型サービスならびに資源予約型サービスにおけるユーザ効用の変化を調べる。そのために、 $N = 200$ 、 $n_B = 200$  に固定し、 $n_A$  を 100 から 300 まで増加させている。集合 A のユーザについては平均接続時間を 40 時間とし、B については平均 60 時間とする。また、実時間アプリケーションの要求帯域  $b_r = 128Kbps$  とし、そのとき  $QoS_r(b_r) = 5$  とする。ドキュメント転送要求発生率  $\lambda_{d,B} = 0.2$  とした。また、Webトラヒックの保護帯域  $Cdr$  は、回線容量  $C$  の 20% とした。Webトラヒックのコスト関

数については、両サービスともに定額制 3000 円を用いるとする。帯域  $b_r$  を 128Kbps とし、その時のコスト  $p_{bw}b_r$  は、現在の電話網の料金を参考にして 10 円 / 3 分とする。この場合、200 円 / 1 時間となり、ここで、負の効用として、実際の円の 100 分の 1 をコストとする。つまり、 $p_{bw}b_r \times 1 = 2$  となるように設定した。上記の条件である平均 40 時間の場合、呼損しなければコストは 80 (ベストエフォート型サービスでは 30) となる。

アクセス回線容量  $C$  を 10Mbps、15Mbps、20Mbps とした場合の、ベストエフォート型サービス、資源予約型サービスにおけるユーザ効用  $U_{r,b}$ 、 $U_{w,b}$ 、 $U_{r,r}$ 、 $U_{w,r}$  を図 15、16 に示す。図 15 より、 $C = 10$ Mbps の場合、実時間トラヒックが少なければ、ベストエフォート型サービスと資源予約型サービスのユーザ効用の差は、ほとんどない。しかしながら、実時間トラヒックが 150 まで増加すると、ベストエフォート型サービスのユーザ効用  $U_{w,b}$  が小さくなり、その結果  $U_{r,b}$  との差が大きくなる。資源予約型サービスのユーザ効用  $U_{w,r}$  も小さくなるが、帯域使用料が高いために  $U_{r,r}$  との差はベストエフォート型サービスほど大きくない。さらに実時間トラヒックが増加すると QoS が非常に小さくなってしまい、ベストエフォート型サービスのユーザ効用はどちらも 0 よりも小さくなってしまう。しかしながら、回線容量が 15Mbps に増加すると、ベストエフォート型サービスと資源予約型サービスのユーザ効用の差は小さくなり、20Mbps になると、逆にベストエフォート型サービスの方が料金が安いので、ベストエフォート型サービスのユーザ効用の方が大きくなっている。すなわち、実時間トラヒックが増加した場合、十分な回線容量を得ることができなければ、ベストエフォート型サービスでは十分なユーザ効用が得ることができず、その場合は資源予約型サービスが必要であると言える。

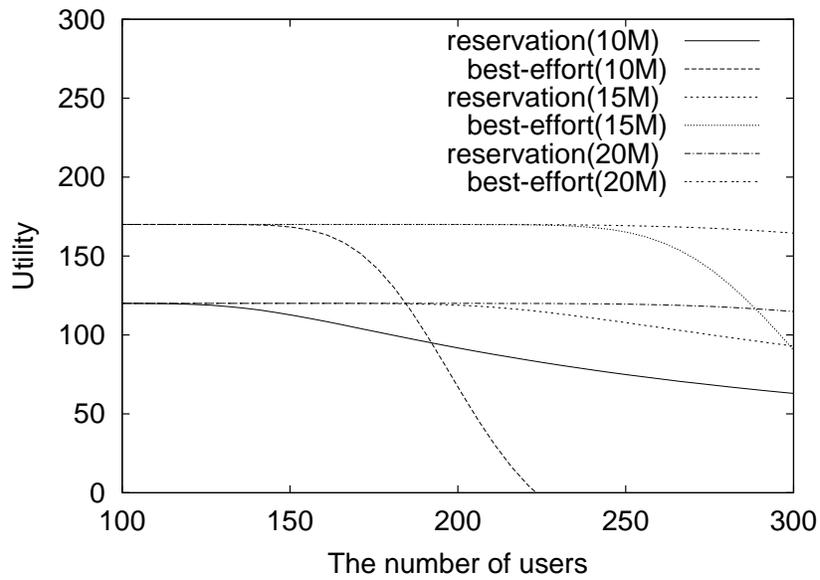


図 15: 実時間アプリケーションの効用  $U_{r,b}$ ,  $U_{r,r}$

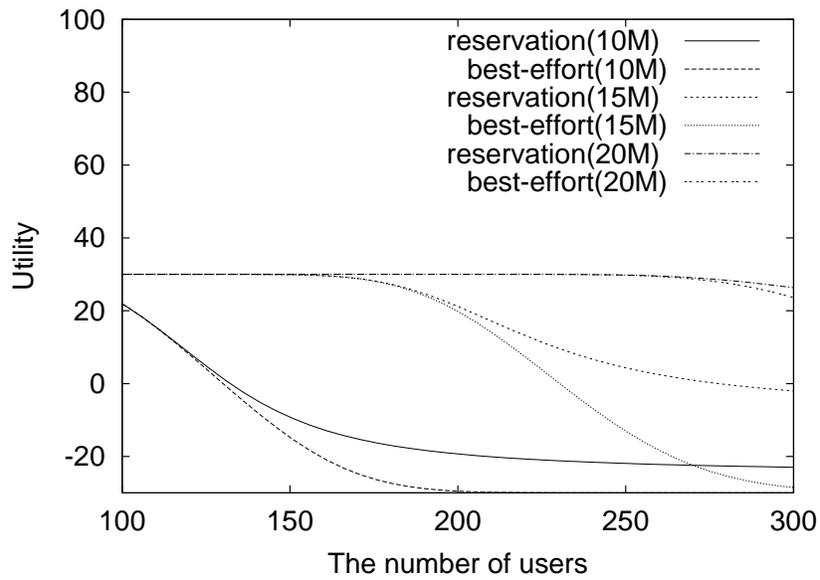


図 16: Web ユーザの効用— ユーザ数が増加  $U_{w,b}$ ,  $U_{w,r}$

## 6 むすび

本論文では、ユーザ効用を尺度として、対話型実時間アプリケーションを予約型、ベストエフォート型サービスに適用した場合の比較を行った。さらに、通信途中の品質劣化を調べるために最悪効用を求めてサービス間の比較を行った。さらに、経路リンク数がユーザ効用に与える影響を調べた。最後に、料金制度を考慮に入れたユーザ効用を調べた。その結果、回線容量が十分にあればベストエフォート型サービスでも十分な効用が得られること、ただし、そのためには適切なネットワーク設計によって十分な回線が準備できることが必要であることを明らかにした。そうでない場合には予約型サービスが必須である。なお、コネクションが必要とする帯域と回線容量は相対的なものであるため、本論文に示した結果は動画像などのアプリケーションについても十分適用できるものである。

本論文では、対話型実時間アプリケーションのみを対象とした評価を行った。今後は、片方向実時間アプリケーションにおけるユーザ効用の比較を行っていく必要がある。また、料金制度に関しては、回線容量に対するコストとユーザが負担する料金の関係を考慮してない。そこで、回線容量とそのコストの関係がネットワークサービスとそのユーザ効用に与える影響を調べていく必要がある。

## 謝辞

本論文を終えるにあたり、御指導、御教授を頂いた宮原秀夫教授に深く感謝致します。本論文において終始御指導頂いた村田正幸教授に深く感謝致します。日頃から適切な助言を頂いた大型計算機センターの下條真司教授、若宮直紀講師、情報処理教育センター大崎博之助手、経済学部部長谷川剛助手に深く感謝致します。最後に、御協力を頂いた宮原研究室、村田研究室の皆様にな心からお礼を申し上げます。

## 参考文献

- [1] R. Braden, D. Clark, and S. Shenker, “Integrated services in the internet architecture,” *Internet RFC 1633*, July 1994.
- [2] P. White, “RSVP and integrated services in the Internet: A tutorial,” *IEEE Commun. Mag.*, vol. 34, pp. 100–106, May 1997.
- [3] A. Demers, S. Keshav, and S. Shenker, “Analysis and simulation of a fair queueing algorithm,” *Proceedings of ACM SIGCOMM*, pp. 1–12, September 1989.
- [4] J-C.Bolot and T.Turletti, “Experience with control mechanisms for packet video in the Internet,” *CCR ACM Sigcom*, vol. 28, January 1998.
- [5] K. Fukuda, N. Wakamiya, M. Murata, and H. Miyahara, “MPEG-2 rate control algorithm for ATM networks with bandwidth re-negotiation,” in *Proceedings of Fifth IFIP IWQoS*, pp. 291–302, May 1997.
- [6] 渡辺啓, 村田正幸, 宮原秀夫, “ユーザ効用によるネットワークサービスの比較評価,” 電子情報通信学会技術報告 *SSE00* 掲載予定, March 2000.
- [7] “An architecture for differentiated services,” *Internet RFC 2475*, December 1998.
- [8] K. Nichols, V. Jacobson, and L. Zhang, *A Two-bit Differentiated Services Architecture for the Internet*. November 1997.
- [9] K. V.Jacobson and K.Poduri, “An expedited forwarding PHB,” *Internet RFC 2598*, June 1999.
- [10] L. Brealau and S. Shenker, “Best-effort versus reservations: A simple comparative analysis,” *CCR ACM Sigcomm*, vol. 28, October 1998.
- [11] 村田正幸, “マルチメディアコンピュータネットワークの通信品質保証,” 電子情報通信学会誌, vol. 81, pp. 362–370, April 1998.

- [12] 村田 正幸, “マルチメディアネットワークにおける通信品質保証の実現と課題,” 電子情報通信学会論文誌 (B-I), vol. J80-B-1, pp. 296–304, June 1997.
- [13] D. T. R.Ramjee, J. Kurose and H. Schulzrinne, “Adaptive playout mechanisms for packetized audio applications in wide-area networks,” *Proc. IEEE Infocom*, vol. 2, pp. 680–688, 1994.
- [14] 浅谷 耕一, 通信ネットワークの品質設計. 電子情報通信学会, February 1993.
- [15] 渡辺啓, 馬場健一, 村田正幸, 宮原秀夫, “ユーザ QoS を考慮した動的帯域管理方式の性能評価,” 電子情報通信学会論文誌 (B-I), vol. J82-B-1, pp. 549–559, April 1999.
- [16] 村田 正幸, “インターネットにおける輻輳制御とトラヒック特性,” システム / 制御 / 情報学会誌, vol. 43, pp. 141–148, March 1999.
- [17] D. Bertsekas and R. Callager, *DATA NETWORKS*, ch. 9, pp. 524–529. Prentice-Hall International Edition.
- [18] S. F. Jamshid Mahdavi, “TCP-friendly unicast rate-based flow control,” *Technical note sent to the end2end-interest mailing list*, January 1997.
- [19] Akira Watanabe, Masayuki Murata and Hideo Miyahara, “Which provides better utility to users, best-effort service or reservation-based service?,” in *Proceedings of ITCS-CSCC*, pp. 1190–1193, July 1999.
- [20] 名部正彦, 馬場健一, 村田正幸, 宮原秀夫, “インターネット・アクセスネットワーク設計のための WWW トラヒックの分析とモデル化,” 電子情報通信学会論文誌 (B-I), vol. J80-B-1, pp. 428–437, June 1997.