

# 時間スケールを考慮した長期依存性トラヒックの性能解析

計 宇 生<sup>†</sup> 藤 野 貴 之<sup>†</sup> 阿 部 俊 二<sup>†</sup>  
松 方 純<sup>†</sup> 浅 野 正 一 郎<sup>†</sup>

インターネットトラヒックには長期依存性ないし自己相似性が広く観察されている。このようなトラヒックを扱うために、これまでにはいくつかの解析モデルが提案されており、その代表的な例がFBMモデルである。しかし、FBMモデルはその扱いやすさの半面、1つのハーストパラメータで表せるような限られた特性を持つ自己相似性トラヒックのモデル化にのみ有効である。本研究では、待ち行列システムにおけるトラヒックの挙動と深く関わりのある時間スケールに注目し、より一般的な相関特性を持つ長期依存性トラヒックの性能の解析方法について提案する。ここで扱うトラヒックの特性は広域ネットワークやLANで観測されたトラヒックのデータから抽出したものである。また、実際のトラヒックデータを用いたシミュレーションの結果を使って、提案方式による近似解析の結果との比較、検証を行う。

## Performance Analysis of Long-range Dependent Internet Traffic According to Relevant Time Scales

YUSHENG JI,<sup>†</sup> TAKAYUKI FUJINO,<sup>†</sup> SHUNJI ABE,<sup>†</sup> JUN MATSUKATA<sup>†</sup>  
and SHOICHIRO ASANO<sup>†</sup>

In recent studies, fractional Brownian motion has been proposed to analyze nowadays internet traffic, which are found to be fractal than the classical Poisson-based models. However, due to its simplicity, fractional Brownian motion is only efficient for approximating the performance of a class of exactly self-similar traffic, whose correlation property can be described by a single Hurst parameter. In this paper, we examine the queueing behavior of long-range dependent internet traffic which has a more general correlation property. We propose an analytical method by focusing on the relationship between time scales and queueing performance. The properties of traffic data discussed in this paper are extracted from traffic traces of real networks, such as a wide area backbone network and a LAN. Results produced by simulation using real traffic data are compared with analytical results carried out by our method.

### 1. はじめに

インターネットトラヒックには長期依存性ないし自己相似性が広く観察されている。そのような長期依存性や自己相似性はLANや、広域ネットワーク、または動画像のトラヒックなどにみられ、多重化されたネットワークトラヒックに広く存在していることが報告されている<sup>1)~4)</sup>。一般に長期依存性のトラヒックでは、トラヒックの変動具合が時間スケールの広い領域において保たれるため、従来のポアソンやマルコフモデルよりも長い待ち行列の生成につながる事が知られている。必要なネットワーク帯域の設計、サービス品質の保証のためには長期依存性トラヒックの性能をでき

るだけ正確に把握することが求められている。

本論文では、実際に観測されたネットワークトラヒックを利用して、その長期依存性がトラヒックの性能に与える影響について検討する。これまでの研究では、長期依存性トラヒックを扱うために、Fractional Brownian Motion (FBM) と呼ばれる確率モデルをはじめとするいくつかの提案がなされている<sup>5)~7)</sup>。これらは厳密な自己相似性を有するようなトラヒックをモデル化するのに有効である。しかしながら、実際に観測されたネットワークトラヒックは、より複雑な自己相関特性を持っており、単一のハーストパラメータでは表せない、いわゆるマルチフラクタルなものが多い。近年ではすでにいくつかの研究によってインターネットトラヒックのマルチフラクタルな性質が報告されており<sup>8),9)</sup>、そのモデル化についてウェブサイトに基づく方法が提案されているが、解析的に性能を与え

<sup>†</sup> 国立情報学研究所

National Institute of Informatics

ることは困難である．ここでは，異なる時間スケールにおけるトラフィックの2次の統計的特性に注目し，一般的な自己相関特性を持つトラフィックの性能を近似的に解析するための方法について提案する．時間スケールと待ち行列システムにおけるトラフィックの挙動との関連性を明らかにすることにより，一般的な長期依存性トラフィックの性能特徴を把握する．また，提案方式による性能解析の結果と，実際に観測されたトラフィックデータを使ったシミュレーションの結果との比較により，提案方式は性能要求に対応した時間スケールにおけるトラフィックの性能を評価するのに有効であることを示す．

## 2. インターネットトラフィックにおける長期依存性

本章では，長期依存性の簡単な定義を与えた後，本論文で扱うトラフィックデータの特性について述べる．一般に，定常増分を有する確率過程  $X(t)$  の自己相関関数が  $r(k)$  であり，次のような条件を満足する場合， $X(t)$  が2次の自己相似性を有するという．

$$r(k) = \frac{1}{2} \{ (k+1)^{2H} - 2k^{2H} + (k-1)^{2H} \} \quad (1)$$

$H \in (1/2, 1)$  は自己相関関数の  $k$  に対する減衰の速さを特徴付ける量で，ハーストパラメータという．

また， $t$  が離散の場合，連続した  $m$  個の  $X(t)$  の平均

$$X^{(m)}(k) = \frac{1}{m} \sum_{t=(k-1)m+1}^{km} X(t), \quad k = 1, 2, \dots$$

を考えると，2次の自己相似性は次式のようにも表される．

$$\text{Var}[X^{(m)}(k)] = \sigma^2 m^{-2(1-H)} \quad (2)$$

一方， $X(t)$  の自己相関関数  $r(k)$  の減衰が非常に緩やかであることを表す性質として，長期依存性がある．確率過程  $X(t)$  が長期依存性を有する場合，次の式が成り立つ．

$$\lim_{k \rightarrow \infty} r(k) = k^{-\beta} L(k) \quad (3)$$

$\beta \in (0, 1)$  は定数， $L(t)$  は任意の  $x > 0$  に対して，

$$\lim_{t \rightarrow \infty} L(tx)/L(t) = 1$$

を満足するような， $t$  の十分大きな領域では緩やかに変化する関数である．長期依存性の特殊な場合として， $L(t) = H(2H-1)$ ， $\beta = 2(1-H)$  であれば，式(3)は2次の自己相似性となる<sup>7)</sup>．

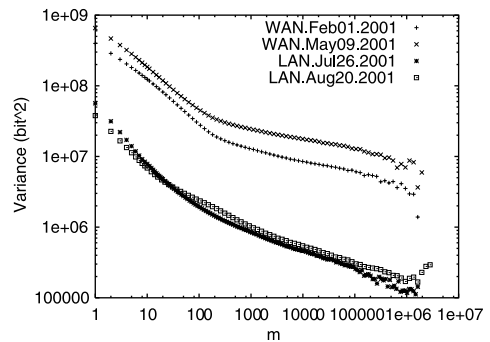


図1 観測トラフィックの Variance-Time 特性の例  
Fig.1 Examples of Variance-Time plots of observed traffic.

2次の自己相似性や長期依存性は， $m$  に対する  $X^{(m)}(k)$  の分散値の変化 ( Variance-Time 特性 ) などで確認することができる．厳密な2次の自己相似性を有する場合，式(2)から，対数座標上では  $X^{(m)}(k)$  の分散の値が  $m$  に対して勾配が  $-2(1-H)$  の直線になることが分かる．

筆者らは広域ネットワークやLANなどのトラフィックに対して観測を行った．広域ネットワークのデータは，国立情報学研究所で運用されている学術情報ネットワーク ( SINET )<sup>10)</sup> のバックボーンノードから得られている．観測を行った2001年の時点で，SINETのバックボーンリンクはOC-3のATM回線であった．モニタリングツールOC3MON<sup>11)</sup>を使ってバックボーンノードで採取したATMセル内の情報をもとに，個々のIPパケットの到着時間とIPヘッダを含めたIPパケット長のデータを得た．また，LANトラフィックについては，tcpdump<sup>12)</sup>を使って国立情報学研究所LANが外部のネットワークと接続するイーサネットのセグメントから採取した，バックボーンデータと同様なIPパケットの到着時間とIPヘッダを含めたIPパケット長のデータである．

図1に4系列のサンプルデータの Variance-Time 特性を示す．これらのデータは2001年2月から8月にかけて，平日午後のピーク時間帯に得られたものである． $X^{(m)}(k)$  を単位時間  $\Delta t = t_{n+1} - t_n (n = 0, 1, 2, \dots)$  で時間軸を等間隔に区切った場合の時間間隔  $m\Delta t$  内の到着ビット数として，それぞれの帯状になっている点の連続は  $X^{(m)}(k)$  の分散特性を  $m$  に対して示したものである．ここでは十分な観測精度が確保できるうえ，時間スケールの十分小さな領域の特性も表現できるように， $\Delta t$  の長さを1ミリ秒とした．したがって，図1の横軸  $m$  の値はミリ秒を単位とする時間スケールとして見る事ができる．

図 1 における 2 系列のバックボーン (WAN) トラヒックデータは異なるバックボーンリンクから採取したため、ビットレートが異なり、したがって異なる分散値を示しているが、分散特性の時間スケール上での減衰の様子がよく似ている。そして 2 系列の LAN データは異なる日に同じ場所で採取しているため、2 つのビットレートは似た値を持っているが、WAN のビットレートおよび分散よりも小さい。

さらに、これらのデータの分散特性の時間スケールに対する減衰の速さは一定ではなく、 $m$  の小さい領域においては減衰が比較的速く、ポアソン特性に近い (指数分布データのハーストパラメータが 0.5 であるのに対し、ここでは 0.6 以上に相当する)。それに対して、比較的長い時間スケールにおいては、分散の減衰が緩やかになり、変動が長く持続されることを示している。図 1 から、広域ネットワークや LAN のトラヒックは長期依存性を有していることが分かる。さらに、それらの関連特性は厳密な自己相似性よりも複雑で一般的なものである。

これまでに筆者らが計測したトラヒックの分散特性のほとんどは図 1 のように、時間スケールに対しておおよそ下方に凸な形を有している。集約トラヒックは短い時間スケールにおいてポアソン性を呈する傾向があることは、古屋らの研究<sup>13)</sup>で指摘されている。これは「局所的ポアソン性」と呼ばれ、TCP プロトコルの Self-Clocking 効果によるものであると報告されている。時間スケールの大きい領域では、ネットワーク上のファイルやセッションの長さ長い裾野の分布を有していることが影響し、そのような通信トラヒックの重ね合わせが長期依存性を生成していることが証明されている<sup>14)</sup>。実際に筆者が行った観測でもファイル長の長い裾野の分布が確かめられている。このような一般的な長期依存性の生成される原因が複雑であり、その詳しい考察はここでは行わないが、以下では、時間スケールに対して下方に凸な形の分散特性を有するトラヒックに限定して議論を進める。

Norros らによって Fractional Brownian Motion (FBM) を用いて自己相似性トラヒックを近似することが提案されている<sup>5),6)</sup>。到着率が  $\lambda$ 、分散係数が  $a$  の確率過程が FBM であるとすれば、時刻  $t$  までに到着するトラヒックの量が次式で表される。

$$A(t) = \lambda t + \sqrt{\lambda a} Z_H(t) \quad (4)$$

$Z_H(t)$  は平均 0、分散  $\text{Var}[Z_H(t)] = |t|^{2H}$ 、ハーストパラメータ  $H \in [0.5, 1)$  を持つ標準 FBM 確率変数である。FBM 過程は正規性 (ガウス性) を有している

ため、 $\{\lambda, a, H\}$  の 3 つのパラメータで完全に決まる。中央極限定理によれば、たくさんの 2 次の自己相似性を有するような、同一の分布を持つ独立な確率過程 (i.i.d. 過程) の集約は FBM になる。そのため、FBM は集約トラヒックをモデル化するために有効である。

しかし、図 1 で示したように、インターネットのトラヒックは多くの場合、分散や自己相関関数のような 2 次特性が時間スケールによって異なる挙動を示している。このような特性を  $\{\lambda, a, H\}$  の 3 つのパラメータで記述することが不可能である。

### 3. 時間スケールを考慮したトラヒックの性能

文献 15) によれば、サービス率  $C$  の待ち行列の入力が  $\{\lambda, a, H\}$  で表す FBM 確率過程であるときの待ち行列長の補分布が次の式によって与えられる。

$$\Pr(Q > x) \sim e^{-\gamma x^{2(1-H)}} \text{ as } x \rightarrow \infty \quad (5)$$

ただしここでは、

$$\gamma \equiv \frac{1}{2a\lambda(1-H)} \left\{ \frac{(C-\lambda)(1-H)}{H} \right\}^{2H}$$

である。この漸近的な値は次の下限値を時間間隔  $t$  について求めることによって得られている。

$$\begin{aligned} \Pr(Q > x) &\geq \max_{t \geq 0} \Pr(A(t) > Ct + x) \\ &= \max_{t \geq 0} \bar{\Phi} \left( \frac{(C-\lambda)t + x}{\sqrt{a\lambda t^H}} \right) \end{aligned} \quad (6)$$

ただし、 $\bar{\Phi}(\cdot)$  は標準正規分布の補分布関数であり、

$$\bar{\Phi}(z) = \frac{1}{\sqrt{2\pi}} \int_z^\infty e^{-\frac{y^2}{2}} dy$$

である。

一般的に性能評価で考慮しなければならないトラヒックの関連特性の範囲は、許容できる待ち行列の長さや遅延の大きさに対応した時間スケールにも依存する<sup>16)</sup>。たとえば、実際のシステムにおいてバッファの長さがある有限な大きさである場合、それに最も関係する時間スケールにおけるトラヒックの関連特性がバッファあふれ率を左右する。このような性能に最も関係する時間スケールは適正時間スケール (Relevant Time Scale)<sup>17)</sup> またはクリティカル時間スケール (Critical Time Scale)<sup>18)</sup> と呼ばれており、自己相似性トラヒックが正規性を有する場合の適正時間スケールが次のように導出されている。

待ち行列長 (またはバッファ長) が  $x$  のときにおける FBM 過程の適正時間スケール  $t$  は次式で表される。

$$t = \frac{x}{C - \lambda} \frac{H}{1 - H} \tag{7}$$

この適正時間スケールはすなわち式 (6) の等号右側の最大値が得られる時間間隔  $t$  の値である。式 (7) から分かるように、適正時間スケールは待ち行列長  $x$ 、剰余帯域  $C - \lambda$  および到着過程のハーストパラメータ  $H$  に依存するが、到着過程の 2 次モーメントには依存しない。式 (7) の  $t$  と  $x$  の関係によって、待ち行列長と時間スケールが直接関係づけられることになる。

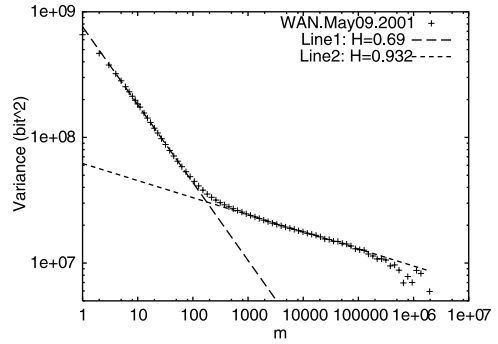
以下では、このような適正時間スケール  $t$  を式 (2) における  $m$  (厳密には  $m\Delta t$ )、すなわち Variance-Time 特性図の横軸の値と対応させることによって、異なる時間スケールにおけるトラヒックの分散特性をトラヒックの性能解析に利用する方法について述べる。

3.1 複数本の直線による分散特性の近似

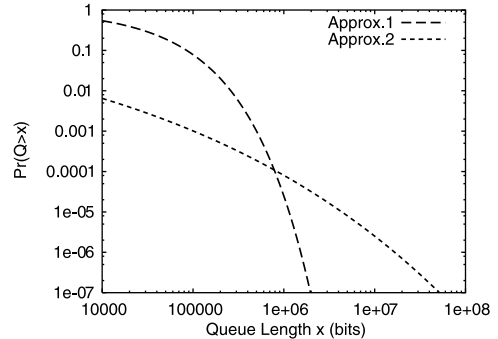
ここでは、実測されたトラヒックの分散特性の形に応じて、複数本の直線による近似の方法について述べる。トラヒックデータとして 2001 年 5 月 9 日午後 3 時から採取した 10,000 秒間の SINET トラヒックを例に使う。観測トラヒックの分散特性 (Variance-Time 特性) が横座標  $m \approx 200$  を曲がり目とする折れ線に近い形をしているため、図 2 (a) のように、それをまず 2 本の直線で近似的に表してみる。比較的急な勾配を持つ直線を「Line1」と呼び、そのハーストパラメータを  $H_1$  と記す。また、比較的緩やかな勾配を持つ直線を「Line2」と呼び、そのハーストパラメータを  $H_2$  と記す。2 本の直線の勾配が  $-2(1 - H_1) < -2(1 - H_2)$  の関係を有することから、 $H_1 < H_2$  である。実際の到着過程の分散特性は (2 つの直線の交点の近くを除けば) およそ 2 つの直線の縦軸値の大きい方を通ることになる。

Line1 および Line2 はそれぞれ 1 つの 2 次の自己相似性確率過程に対応する。さらに正規性を仮定すれば、それぞれの直線に対応したものは、1 つの FBM 過程として考えることができる。Line1 と Line2 に対応する FBM 過程にそれぞれ式 (5) を適用すると、図 2 (b) に示す Approx.1 と Approx.2 の 2 つの待ち行列長の補分布の曲線が得られる。同図では 2 つの性能曲線が交わるところの様子がよく分かるように、縦、横の両軸とも対数スケールで示した。また、待ち行列に加わる負荷 ( $\lambda/C$ ) を 75% とした。

ハーストパラメータの異なる 2 つの FBM 過程が異なる待ち行列長の分布を示すことが図 2 (b) から分かる。図 2 (a) の 2 つの直線と実際の到着過程の分散特性との関係から、実際の到着過程の待ち行列長の補分布は図 2 (b) の 2 つの曲線の中の最も性能の悪い部分を



(a) 5 月 9 日データの Variance-Time 特性と 2 つの直線による近似



(b) (a) の直線に対応した 2 つの FBM 過程の待ち行列長の補分布  
図 2 2 つの FBM モデルによるトラヒックの特性と性能の近似の例

Fig.2 Traffic approximated by two FBM processes.

通る曲線で近似することが考えられる。しかし、2 つの曲線の繋ぎ目のところでは滑らかさが欠けてしまっている。それには次の 2 つの要因が考えられる。(ア) Line1 と Line2 の交点の近くでは、元の到着過程の分散特性が 2 つの直線のどちらよりも分散の値がやや大きくなっている。そのため Approx.1 と Approx.2 の交点の近くでは、2 つの直線による近似よりも実際の性能が悪いと考えられる。(イ) 2 つの直線による近似では交点において分散値の不連続な変化が生じているため、両方の直線による性能への影響を考える必要がある。(イ)については次の節でもう少し詳しく述べる。

3.2 分散特性の勾配の変化が急な場合

上述の要因 (イ) は以下の理由による。

式 (6) の等号右側の  $\Phi(\cdot)$  関数の括弧の中を  $z(t)$  と置くと、 $z$  の値が小さいほど式 (6) の左辺の確率が大きい。したがって、2 つの FBM 過程の性能を比較するには、それぞれについての  $z$  の値を比較すればよい。2 つの直線が同じ到着過程の異なる時間スケールにおける分散特性を表しているため、 $z$  関数におけるサービス率  $C$  および到着率  $\lambda$  が 2 つの FBM 過程に

ついて同一である．また，Line1 と Line2 に対応する FBM 過程の変動係数をそれぞれ  $a_1$  と  $a_2$ ，適正時間スケールを  $t_1$  と  $t_2$  で表す．2 つのハーストパラメータは  $H_1 < H_2$  であるため，同じ待ち行列長に対する適正時間スケールは  $t_1 < t_2$  である．

Line1 と Line2 の交点の横座標を  $T_m$  とする．また，ある待ち行列長  $Q_1$  に対して  $t_2 = T_m$  となるとき  $t_1$  を  $T_1$  とし，別な待ち行列長  $Q_2$  に対して  $t_1 = T_m$  となるとき  $t_2$  を  $T_2$  とすると，式 (7) より，

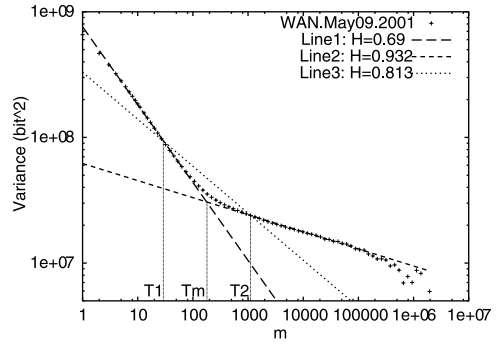
$$Q_1 = T_m(C - \lambda) \frac{1 - H_2}{H_2} = T_1(C - \lambda) \frac{1 - H_1}{H_1}$$

$$Q_2 = T_m(C - \lambda) \frac{1 - H_1}{H_1} = T_2(C - \lambda) \frac{1 - H_2}{H_2} \quad (8)$$

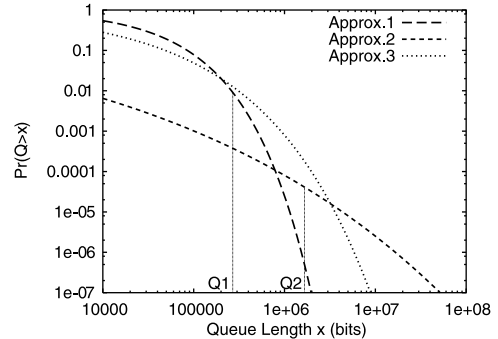
であり， $Q_1 < Q_2$ ， $T_1 < T_m < T_2$  である．この場合，待ち行列長が  $Q_1$  より小さい領域では Line1 に対応する FBM 過程，逆に待ち行列長が  $Q_2$  より大きい領域では Line2 に対応する FBM 過程のほうがより悪い性能（すなわち待ち行列長の補分布上でより大きい確率）を持つことが知られている<sup>17)</sup>．しかし，中間的な待ち行列長，すなわち  $Q_1 < x < Q_2$  の場合，どちらの性能が悪いかは一概にはいえない．むしろ，2 つの FBM 過程の性能が接近しているため，両方からの影響を同時に考慮する必要がある．

そこで，図 3(a) で示すように，横座標が  $T_1$  である Line1 上の点と横座標が  $T_2$  である Line2 上の点を通るような第 3 の直線 Line3 による近似を考える．Line3 の勾配が Line1 と Line2 の間の値を持つため，Line3 のハーストパラメータ  $H_3$  は  $H_1$  と  $H_2$  の間の値を持っている．表 1 には 2001 年 5 月 9 日のデータを例とした場合の 3 つのハーストパラメータの値とそれぞれの直線が分散の値が最も大きくなっている時間スケールの領域をまとめて示す． $T_1$  と  $T_2$  の間の領域では，Line3 の分散の値が最も大きい．また，実際のトラフィックの分散特性も Line1 と Line2 が交差するあたりでは，2 つの直線よりも分散の値がやや大きくなっているため，それを分散値が比較的大きい Line3 で近似することによって安全側の評価ができると思われる．

Line3 に式 (5) を適用すると，図 3(b) の Approx.3 で示す第 3 の性能曲線を得ることができる．3 つの性



(a) 5 月 9 日データの Variance-Time 特性と 3 本の直線による近似



(b) (a) の直線に対応した 3 つの FBM 過程の待ち行列長の補分布  
図 3 3 つの FBM モデルによるトラフィックの特性と性能の近似の例

Fig. 3 Traffic approximated by three FBM processes.

表 1 図 3(a) の 3 つの直線の関係  
Table 1 Summarization of three lines used in Fig. 3(a).

	ハーストパラメータ	分散値が最も大きい領域
Line1	$H_1 = 0.69$	$T_1$ の左側
Line2	$H_2 = 0.932$	$T_2$ の右側
Line3	$H_3 = 0.813$	$T_1$ と $T_2$ の間

能曲線の最も性能が悪い（最も右上の）部分をつなげると，図 2(b) よりも滑らかな性能曲線が得られる．2001 年 5 月 9 日のデータの例では，上述の  $T_1$ ， $T_m$ ， $T_2$  および  $Q_1$ ， $Q_2$  は表 2 に示すような数値になる．図 3(b) において， $Q_1$  は Approx.1 と Approx.3 の交点の横座標よりもやや右側であり， $Q_2$  は Approx.3 と Approx.2 の交点の横座標よりも左側である．すなわち，待ち行列長が  $Q_1$  から  $Q_2$  までの間よりもやや広い領域の性能が Line3 によって得られている．

一般に，Line1 と Line2 の勾配の差が小さければ，対数座標上での  $T_1$  と  $T_2$  の間隔が狭くなる．その間隔が十分小さければ，Line1 と Line2 だけで近似することも可能である．逆に Line1 と Line2 の勾配の差が大きいほど， $T_1$  と  $T_2$  の間隔が大きくなる．図 4

$z$  関数にある  $H$  の値に  $H_1$  と  $H_2$  を， $a$  に  $a_1$  と  $a_2$  をそれぞれ代入して 2 つの FBM 過程についての  $z_1$  と  $z_2$  を求めると， $t_1 \leq T_1$  すなわち  $t_2 \leq T_m$  のとき  $z_2(t_2) \geq z_1(t_2) \geq z_1(t_1)$  である．最初の不等号は Line1 と Line2 の位置関係により， $\sqrt{a_2} \lambda t_2^{H_2} \leq \sqrt{a_1} \lambda t_2^{H_1}$  であることから，次の不等号は  $t_1$  が  $z_1(t)$  を最小にする点であることからである．同様に， $t_2 \geq T_2$  すなわち  $t_1 \geq T_m$  のとき  $z_1(t_1) \geq z_2(t_1) \geq z_2(t_2)$  である．

表 2 図 3 の各横座標の値  
Table 2 Abscissa values in Fig. 3.

時間スケール	待ち行列長
$T_1 = 29(\text{ms})$	$Q_1 = 269\text{Kbits}$ $Q_2 = 1,660\text{Kbits}$
$T_m = 180(\text{ms})$	
$T_2 = 1,100(\text{ms})$	

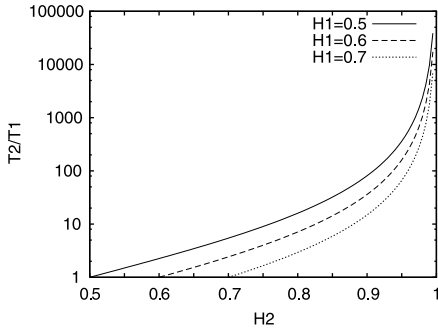


図 4 図 3(a) における  $T_2/T_1$  と  $H_2$  の関係  
Fig. 4 Relationship between  $T_2/T_1$  and  $H_2$  in Fig. 3 (a).

に  $H_1$  を固定したときの  $T_2/T_1$  と  $H_2$  の関係を示す。同図から分かるように、 $H_1$  と  $H_2$  の値が離れるほど、 $T_1$  と  $T_2$  の値で決まる Line3 の性能支配領域が広がる。特に  $H_2$  が 1 の付近になると、 $T_2/T_1$  の値が急激に大きくなる。 $H_2$  の値が大きくなって安全側の評価になるが、ある程度の精度を確保するためには、 $H_2$  が 1 から多少離れていることが条件となる。

上述の方法を分散特性が図 1 のような下方に凸な折れ線に近い形をしている場合に一般化できる。折れ曲がった点の数が  $i$  ( $i \geq 1$ ) であるとすると、 $i$  個の点で交差する  $i+1$  本の直線に対応する FBM 過程で元のトラヒックの性能を近似する。そこから得られた性能曲線がもし図 2 (b) のように交点付近で滑らかさが欠けらば、さらに図 3 (a) のように  $j$  ( $0 \leq j \leq i$ ) 本の直線を書き添えて、計  $i+j+1$  本の直線を得る。元の長期依存性トラヒックの性能はこの  $i+j+1$  本の直線に対応する FBM 過程の性能で近似することができる。この場合、それぞれの直線に対応するハーストパラメータも、性能支配領域を時間スケールの小さい方から大きい方へ並べる順に、値が大きくなっていく。

4. シミュレーション結果との比較

この章では観測されたトラヒックデータを使ったシミュレーションによってその性能を調べ、前述の近似解析方法の結果との比較を行う。

ここでは、2001 年 5 月 7 日および 5 月 9 日の午後 3 時頃から測定したそれぞれ 10,000 秒の長さを持つ 2 本のバックボートトラヒックデータトレースを使う。

表 3 2 本のデータトレースのミリ秒ごとの到着ビット数の統計値  
Table 3 Statistics of the numbers of bits arrived per mini-second for the two data traces.

測定日	平均	分散	歪度	尖度
2001 年 5 月 7 日	$6.02 \times 10^4$	$6.09 \times 10^8$	-0.20	-0.46
2001 年 5 月 9 日	$6.15 \times 10^4$	$6.56 \times 10^8$	-0.14	-0.50

10,000 秒はおよそ最長のピーク時間帯の持続時間である。表 3 にこの 2 本のデータトレースのミリ秒単位における到着ビット数の統計値を示し、観測データの中で最も正規分布に近いものである。

シミュレーションは個々のパケットの到着時間と長さを使って待ち行列における挙動を調べた。ただし、流体近似による解析の結果と合わせるため、有限バッファ時の損失はパケット単位ではなく、ビット単位でのバッファあふれと仮定して、ビット総数中のあふれたビット数の割合からバッファあふれ率の計算を行った。

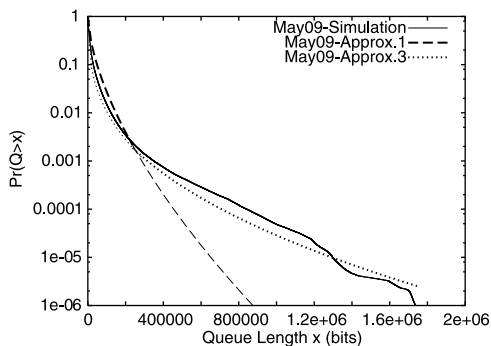
4.1 待ち行列長の補分布

まず、長期依存性トラヒックが無限バッファを持つサーバでサービスされるとき待ち行列長の分布を調べた。図 5 に 5 月 9 日のデータの待ち行列長の補分布のシミュレーション結果を示す。比較のため、前述の 3 つの FBM 過程による近似解析方法の結果も点線で示した。3 つの性能曲線の性能を支配する部分を太めの線で強調し、他の部分は細い線で示してある。図 5 の (a), (b), (c) における負荷はそれぞれ 70%, 75%, 80% である。

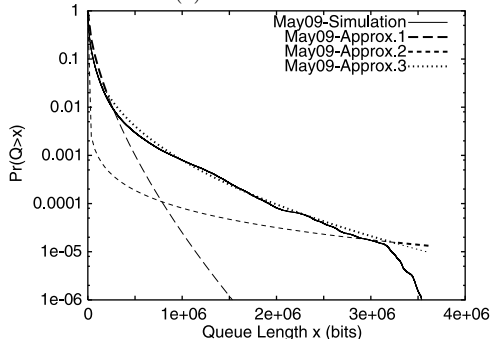
シミュレーションで使えるデータトレースの長さが限られているため、イベントの発生が希少である場合に得られる確率の値には誤差が生じやすい。そのため、待ち行列長の大きい区間では結果の揺れと確率値の急激な減少が起きている。図 5 の (a), (b), (c) ではそれぞれ確率値が  $2 \times 10^{-6}$ ,  $2 \times 10^{-5}$ ,  $10^{-3}$  より小さい部分がそれにあたる。そのような部分を除けば、3 つの FBM 性能曲線による近似解析の結果はシミュレーションの結果とよく一致している。

図 5 (a) における近似解析の結果には、2 つの FBM 性能曲線 (図 3 (a) における Approx.1 と Approx.3) だけが示されている。それは、最も大きいハーストパラメータを持つ Line2 の性能支配領域がシミュレーションの結果の領域からはずれているためである。また、負荷が 80% の場合 Approx.1 が待ち行列長のごく小さい領域の性能を支配しているため、Approx.1 による性能支配領域は図 5 の (c) からはほとんどみられない。

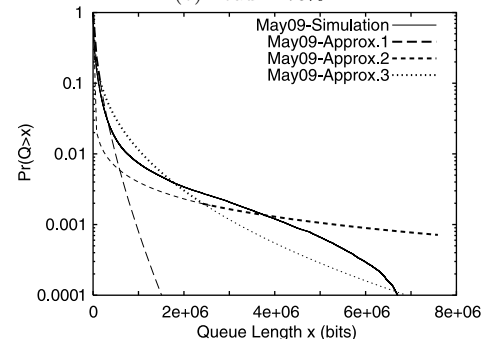
さらに、5 月 7 日のデータについても負荷 70%, 75%, 80% の下でのシミュレーションと解析の結果を



(a) Load = 70%



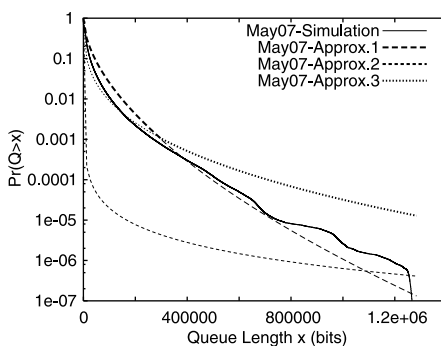
(b) Load = 75%



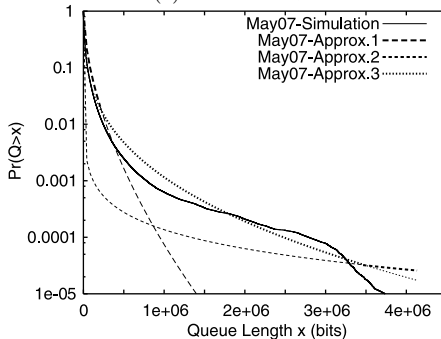
(c) Load = 80%

図5 5月9日データの待ち行列長の補分布のシミュレーション結果と解析による近似

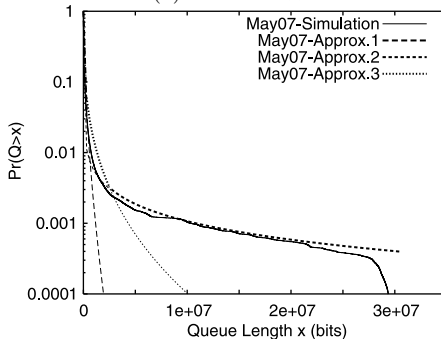
Fig. 5 Simulation results and analytical approximations of the complimentary distribution of queue length for data trace of May 9, 2001.



(a) Load = 70%



(b) Load = 75%



(c) Load = 80%

図6 5月7日データの待ち行列長の補分布のシミュレーション結果と解析による近似

Fig. 6 Simulation results and analytical approximations of the complimentary distribution of queue length for data trace on May 7, 2001.

図6に示した．図6の(a)では Approx.1, (b)では Approx.3, (c)では Approx.2が主な部分で性能を支配している様子が分かる．ただし、負荷の小さいとき近似解析の結果よりもシミュレーション値がやや小さくなっているのは、5月9日のデータと比べ、歪度が小さい(すなわち到着ビット数の周辺分布が正規分布よりも正のほうの裾が短い)ためであると推測される．

以上の結果によって示されているように、待ち行列長の短い領域の性能は短い時間スケールにおけるトラヒックの統計特性に依存し、逆に待ち行列長の長い領

域の性能は長い時間スケールにおけるトラヒックの統計特性に依存する．異なる時間スケールにおけるトラヒックの統計特性を同時に反映させることによって、比較的広い範囲の性能を近似的に得ることができる．

#### 4.2 待ち行列長と負荷の関係

無限バッファの待ち行列において、待ち行列長がある値を超える確率は負荷に関係する．待ち行列長と負荷の関係は負荷率を変化させて一定の待ち行列長を超える確率で求めることができる．図7に5月9日のデータについて待ち行列長がある値を超える確率と負

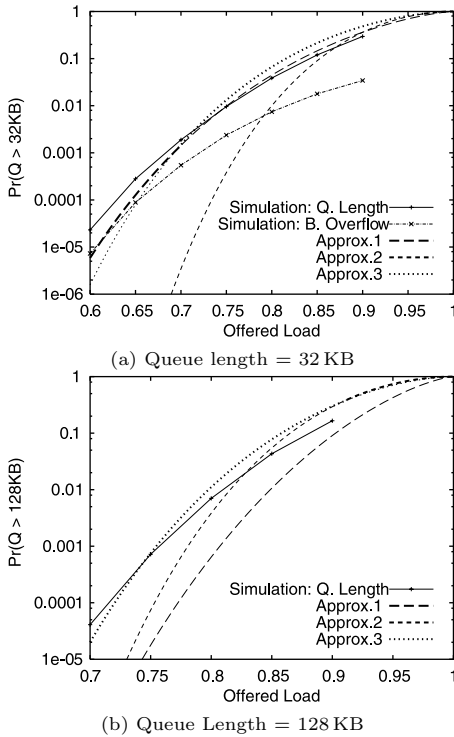


図7 2001年5月9日のデータの待ち行列長と負荷の関係のシミュレーション結果と解析による近似

Fig. 7 Simulation results and analytical approximations of the probability of queue length vs. offered load for data trace on May 9, 2001.

荷の関係の解析とシミュレーションの結果を示す。待ち行列長は、図7の(a)と(b)でそれぞれ32と128キロバイトとした。また、図7(a)ではバッファ長が有限なときのバッファあふれ率のシミュレーション結果も合わせて示した。縦軸は対数スケールである。図5と同じように、解析結果の性能支配部分は太めの線で強調した。

前述したように、解析結果は3つのFBM過程の性能より得ている。しかし、比較的短い待ち行列長の図7(a)では、Approx.1とApprox.3だけが結果に寄与しているように見え、Approx.2は負荷が100%に近いところの性能を支配している。一方、比較的長い待ち行列長の図7(b)では、Approx.2とApprox.3だけが結果に寄与し、Approx.1は負荷のかなり小さい領域の性能を支配している(同図には示されていないが、実際のApprox.1とApprox.3との交点は横軸の値が約0.35、縦軸の値が約 $10^{-50}$ にある)。また、図7(b)で負荷が70%以下の部分が示されていないのはシミュレーションの結果が得られなかったためである。

近似解析とシミュレーションの結果は全体的に比較

的一致しているが、図7(a)において、バッファ長が $B$ の場合のバッファあふれ率が無限バッファの場合の待ち行列長が $B$ を超える確率よりも小さくなっている。特に負荷が大きいときにはその差が顕著である。それはバッファあふれ率を考えた場合、バッファからあふれたものがサーバのサービスの対象から除かれるため、待ち行列長の補分布に比べ、待ち行列が見かけ上短くなるためである。バッファあふれ率が大きくなるような、負荷の大きいときやバッファ長が短いときには大きな差を生じている。したがって、待ち行列長の補分布を使ってバッファあふれ率を近似するにはバッファあふれ率の大きい領域では大きな誤差が生じやすい。

以上の結果から分かるように、異なる時間スケールにおけるトラフィックの統計特性は異なる負荷の下での性能に影響している。負荷が小さいときには短い時間スケールの統計特性が性能を支配し、逆に負荷が大きいときには長い時間スケールにおける統計特性が性能を支配する。

## 5. おわりに

インターネットトラフィックの長期依存性が明らかになって以来、様々な方法で長期依存性トラフィックのモデル化や性能解析が試みられている。よく知られているものとして、FBM過程や、Fractional ARIMA過程<sup>19)</sup>、 $M/G/\infty$ 待ち行列<sup>20)</sup>を用いたモデル化の手法が提案されているが、その中で特にFBMモデルは待ち行列システムの入力としたときの性能が容易に導出されるため性能解析によく用いられている。FBM過程は従来の正規過程に長期依存性を表すパラメータ $H$ の導入によって、長期依存性トラフィックを記述できるようになったことは意味が大きい。特定のFBM過程における $H$ が一定であるため、実際には2次の自己相似性トラフィックの記述に限定される。

一方、近年では長期依存性モデルのみでトラフィックの性能を近似することには限界があり、応用範囲を限定すれば従来のトラフィック理論の適用も可能であると指摘されている<sup>21)~23)</sup>。このような結論が得られる理由の1つは、バッファ長が短い場合または負荷があまり重くない場合には実際の性能が従来のポアソンモデルによる解析の結果とさほど違わないことである。また、本研究でも明らかになったように、インターネットトラフィックの多くは、厳密な自己相似性を有するのではなく、より一般的な長期依存性を持つものが多い。すなわち、トラフィックの2次モーメントの時間スケールに対する減衰の度合いが必ずしも一定ではない。このようなトラフィックは1つのハーストパラメータで表



せないため、厳密な 2 次の自己相似性を仮定した解析手法は必ずしも有効ではないのである。

本研究では、トラヒックの性能を支配する時間スケールに着目し、適正時間スケールという、時間スケールと待ち行列長の関係を使って一般的な長期依存性を持つトラヒックの性能を近似する方法について提案した。それは分散の減衰特性が異なる複数の時間スケールのそれぞれに対して、別々な FBM 過程をあてはめ、それぞれの FBM 過程による解析の結果を合成して元のトラヒックの性能として近似する方法である。ハーストパラメータが 0.5 の場合も FBM 過程の特殊なケースとして含まれるため、提案方式は時間スケールの短い領域でのポアソン性と時間スケールの長い領域での強い相関性を同時に表現することができる。提案方式では FBM 過程をベースに使用しているため、トラヒックの 2 次モーメントの時間スケール上の特性のみに基づいて性能の解析ができ、適用が簡単であるが、一定の近似精度を得るためにはトラヒックの正規性が要求される。したがって、集約度の比較の高いトラヒックにより適合しており、このことは、実測されたバックボーントラヒックを用いたシミュレーションの結果との比較によっても実証された。

また、文献 24) では、バッファの平均稼働時間と緩和時間を時間区分領域  $m$  と対応させ、負荷領域へマッピングする方法により、負荷ごとに異なる FBM 過程を用いて待ち行列長を近似する方法が提案されている。本論文の提案方法では待ち行列長と時間スケールを直接関連づける適正時間スケールの利用によって、バッファにおける平均稼働時間や緩和時間の算出が不要となる（実際にそれらの時間を長期依存性トラヒックについて算出する有効な方法がなく、シミュレーションに頼ることになる）。さらに、分散特性の勾配の急な変化による性能への影響を考慮した複数の FBM 過程の合成によって、負荷や待ち行列長のより広い領域における性能の推定が可能となる。

技術の進歩によってネットワークの帯域やアーキテクチャもつねに変化している。数ないし数十ギガビットオーダの帯域を持つ光ネットワークのトラヒック観測と性能評価についても興味深い話題である。ネットワークの帯域の増大の結果として、受動的なトラヒックモニタリングのみでトラヒックを正確に把握することは困難になりつつあり、能動的なトラヒックモニタリングや、部分的なモニタリングによる性能の推測も含めてより有効な性能評価の方法を検討していくことが今後の課題である。

## 参考文献

- 1) Leland, W., Taqqu, M., Willinger, W. and Wilson, D.: On the self-similar nature of Ethernet traffic (extended version), *IEEE/ACM Trans. Networking*, Vol.2, No.1, pp.1–15 (1994).
- 2) Paxson, V. and Floyd, S.: Wide area traffic: the failure of Poisson modeling, *IEEE/ACM Trans. Networking*, Vol.3, No.3, pp.226–244 (1995).
- 3) Beran, J., Sherman, R., Taqqu, M. and Willinger, W.: Long-range dependence in variable-bit-rate video traffic, *IEEE Trans. Commun.*, Vol.43, No.2/3/4, pp.1566–1579 (1995).
- 4) Kato, T., Ji, Y. and Asano, S.: A study on the self-similarity of MPEG2 video traffic and the effect for transmission quality, *Proc. 10th Int. Packet Video Workshop* (2000).
- 5) Norros, I.: On the use of fractional Brownian motion in the theory of connectionless network, *IEEE JSAC*, Vol.13, No.6, pp.953–962 (1995).
- 6) Erramilli, A., Narayan, O. and Willinger, W.: Experimental queueing analysis with long-range dependent packet traffic, *IEEE/ACM Trans. Networking*, Vol.4, No.2, pp.209–223 (1996).
- 7) 小沢：長期依存性と自己相似性トラヒックのモデル化，システム，制御と情報，Vol.43, No.3, pp.117–122 (1999).
- 8) Riedi, R., Crouse, M., Ribeiro, V. and Baraniuk, R.: A multifractal wavelet model with application to network traffic, *IEEE Trans. Inf. Theory*, Vol.45, No.3, pp.992–1019 (1999).
- 9) Feldmann, A., Gilbert, A., Huang, P. and Willinger, W.: Dynamics of IP Traffic: A study of the role of variability and the impact of control, *Proc. ACM SIGCOMM*, pp.301–313 (1999).
- 10) <http://www.sinet.ad.jp/>
- 11) <http://www.nlanr.net/NA/Oc3mon/>
- 12) <http://www.tcpdump.org/>
- 13) 古屋，中村，野本：IP バックボーントラヒックの局所ポアソン性に関する時間スケールとネットワーク設計への応用，信学技報，IN2001–108, pp.9–16 (2001).
- 14) Taqqu, M., Willinger, W. and Sherman, R.: Proof of a Fundamental Result in Self-Similar Traffic Modeling, *ACM Computer Communication Review*, Vol.27, No.2, pp.5–23 (1997).
- 15) Norros, I.: A storage model with self-similar input, *Queueing Systems*, Vol.16, pp.387–396 (1994).
- 16) Montgomery, M. and De Veciana, G.: On the

relevance of time scales in performance oriented traffic characterizations, *IEEE/ACM Trans. Networking*, Vol.7, No.5, pp.629–640 (1999).

- 17) Neidhardt, A. and Wang, J.: The concept of relevant time scales and its application to queuing analysis of self-similar traffic (or is Hurst naughty or nice?), *Proc. ACM SIGMETRICS*, pp.222–232 (1998).
- 18) Ryu, B. and Elwalid, A.: The Importance of Long-Range Dependence of VBR Video Traffic in ATM Traffic Engineering: Myths and Realities, *ACM Computer Communication Review*, Vol.26, No.4, pp.3–14 (1996).
- 19) Lopez-Ardao, J., Lopez-Garcia, C., Suarez-Gonzalez, A., Fernandex-Veiga, M. and Rodriguez-Rubio, R.: On the Use of Self-Similar Processes in Network Simulation, *ACM Trans. Modeling and Computer Simulation*, Vol.10, No.2, pp.125–151 (2000).
- 20) Makowski, A. and Parulekar, M.: Buffer Asymptotics for M/G/ $\infty$  Input Processes, *Self-Similar Network Traffic and Performance Evaluation*, Park, K. and Willinger, W. (Eds.), Willy Interscience (2000).
- 21) Heyman, D. and Lakshman, T.: Long-Range Dependence and Queueing Effects for VBR Video, *Self-Similar Network Traffic and Performance Evaluation*, Park, K. and Willinger, W. (Eds.), Willy Interscience (2000).
- 22) Cao, J. and Ramanan, K.: A Poisson Limit for Buffer Overflow Probabilities, *Proc. IEEE INFOCOM*, pp.994–1003 (2002).
- 23) 古屋, 中村, 野本, 滝根, 福嶋: 実測 IP トラヒックの局所ポアソン性に関する一考察, 信学技報, IN2002–187, IA2002–43, pp.25–30 (2003).
- 24) 阿部, 藤野, 計, 松方, 浅野: 学術情報ネットワークにおける長期依存性トラヒック解析とキュー長テール分布近似, 電子情報通信学会論文誌, Vol.J86–B, No.12, pp.2487–2500 (2003).

(平成 15 年 7 月 11 日受付)

(平成 16 年 3 月 5 日採録)



計 宇生 (正会員)

昭和 59 年東京大学工学部電子工学科卒業。平成元年同大学院電気工学専攻博士課程修了。工学博士。平成 2 年学術情報センター助手。平成 7 年同助教授を経て、現在国立情報学研究所助教授。超高速ネットワークの品質保証, トラヒック制御, 性能評価等の研究に従事。信学会, IEEE 各会員。



藤野 貴之 (正会員)

平成元年明治大学工学部電子通信工学科卒業。平成 6 年同大学院電気工学専攻博士後期課程修了。工学博士。平成 6 年東京理科大学情報処理センター。平成 8 年 10 月学術情報センター助手。平成 12 年国立情報学研究所助手, 現在に至る。超高速ネットワークアーキテクチャ, ネットワーク制御方式の研究に従事。信学会会員。



阿部 俊二

昭和 55 年豊橋技術科学大学工学部情報工学課程卒業。昭和 57 年同大学院修士課程修了。同年(株)富士通研究所入社。平成 7 年学術情報センターを経て、現在国立情報学研究所助教授。工学博士(平成 8 年, 東京大学)。通信ネットワークの性能評価, トラヒック制御方式, 超高速ネットワークアーキテクチャ等の研究に従事。信学会, IEEE 各会員。



松方 純 (正会員)

昭和 54 年東京大学教養学部基礎科学科卒業。昭和 59 年同大学院工学系研究科情報工学専攻博士課程単位取得退学。同年東京大学大型計算機センター。国文学研究資料館, 宇宙科学研究所, 学術情報センターを経て、現在, 国立情報学研究所助教授。コンピュータネットワークの研究に従事。信学会会員。



浅野 正一郎 (正会員)

昭和 45 年東京大学工学部電子工学科卒業。昭和 50 年同大学院電子工学専攻博士課程修了。工学博士。東京大学大型計算機センター, 同宇宙航空研究所, 同工学部を経て, 昭和 61 年から学術情報センター。現在国立情報学研究所研究主幹・教授。超高速ネットワークの研究開発に従事。信学会フェロー。