ネットワークトラヒックの統計的性質と確率モデル

UEDA, Takeshi / YANA, Kazuo / 八名, 和夫 / 植田, 武 / SHINO, Hideaki / 篠, 秀明

(出版者 / Publisher)
法政大学計算科学研究センター
(雑誌名 / Journal or Publication Title)
法政大学計算科学研究センター研究報告 / Bulletin of Computational Science Research Center, Hosei University
(巻 / Volume)
13
(開始ページ / Start Page)
127
(終了ページ / End Page)
136
(発行年 / Year)
2000-03-31
(URL)
https://doi.org/10.15002/00024886

ネットワークトラヒックの統計的性質と確率モデル

篠 秀明　植田 武　八名 和夫
法政大学工学部電子情報学科

あらまし
情報ネットワークの性能向上を目指すためにはネットワークトラヒックの統計的性質を明らかにすることが設計上重要である。また、その統計的性質を知ためにネットワークトラヒックをモデル化しそのパラメタを推定することが有用である。本稿では 2 状態マルコフ過程が多数重畳している重畳マルコフ過程を用いたモデル化を行ない、情報ネットワークトラヒックのパラメタを推定する方法でトラヒックの性質を明らかにする。ここでは SNMP(Simple Network Management Protocol) により得られる情報ネットワークトラヒックに対し、計算機シミュレーションによってモデルの妥当性を検討した上で実データの解析を行なった。

まえがき

近年、情報化が急速に進み情報ネットワークの重要性がますます高まっている。今やほとんどの学校や会社では、コンピュータ同士を LAN(Local Area Network) などの情報ネットワークで結び情報の利用・管理などに利用している。また、それらの大多数は WAN(Wide Area Network) にも接続されていて、各 LAN 同士で情報の交換なども行なえるようになっている。高性能な情報ネットワークの設計を行なうためには、将来、トラヒックが大量に生じるであろうと思われる場所に大容量回線を割り当てることやプロキシサーバなどによってトラヒックの流量を動的に変えることなどが考えられるが、そのためには情報ネットワークトラヒックの変動特性についてその確率的構造を解析し理解する事は設計上重要であり不可欠である。また、工学的立場から見てもその構造を理解することは大変興味深いことである。

現在では SNMP(Simple NetworkManagement Protocol) などにより情報ネットワークトラヒックが簡単にモニター可能となり多くの LAN で活用されているが、もっぱら SNMP はネットワーク負荷の監視に用いられるのみで、 SNMP から得られるトラヒック時系列データに対し解析を行ない, トラヒックの発生間隔等の特徴を抽出することや制御を行なうことなど積極的な応用がなされていないのが現状である。そこで本稿では、情報ネットワークトラヒックのモデル化を行ない、パラメタを推定することで、その確率的構造について検討する。また、モデル化はトラヒックの発生自体が通信パケットを送信するか否かの 2 状態から説明できると考え、トラヒックを発生している状態と全く発生させていない 2 状態を遷移するような要素トラヒックが基本となってそれが複数重畳して総トラヒックになっているものとするトラヒック基本モデルを考えた。また、状態の遷移には従来、通信工学等でモデルリングに用いらるマルコフ過程モデルを用いた。マルコフ過程とは状態が複数存在しているときにある任意の時刻 t の状態 $S(t)$ がわかっていれば、時刻 τ だけ後の状態に $S(t+\tau)$ 存在する値がとる確率を求めることができる性質 (マルコフ性) をもつ確率過程のことである。つまり、現在以降の状態が過去の状態とは無関係に定まることを意味していて、ランダムに状態が変化するネットワークトラヒックの遷移を表しやすい。これらの 2 つのモデルを組み合わせて 2 状態のマルコフ過程が複数重畳してネットワークトラヒックを形成する重畳マルコフ過程モデルによってトラヒックのモデリングを行なった。

1　ネットワークトラヒックのモデル化

本章では、トラヒック時系列データをモデル化する際に用いた基本的モデルとそれにマルコフ過程を応用した時に得られる重畳マルコフ過程モデルとそのモデルから得られるパラメタについて述べる。

1.1　ネットワークトラヒック基本モデル

ネットワーク上に総数 N の単一のトラヒックを発生させるネットワーク端末が存在しているモデルを考えたとき、任意の i 番目の単一トラヒック (以後要素トラヒックと呼ぶ) は、図1のようにネットワーク負荷を全く生じさせない状態 (状態 0 とする) と u[bps](単位トラヒックと呼ぶ) の負荷を生じささせる状態 (状態 1 とする) の 2 状態を遷移して、それぞれにトラヒックを発生させているとする。そして、これらの各要素トラヒックは互いに独立であるとすれば、図2で示すように、それらが複数重畳して総トラヒックを形成していると考えられる。
もし、ある時刻 t においてネットワークに接続されている端末数が $N(t)$、各端末が確率 $P(t)$ でそれぞれ u_i[bps] のトラヒックを生じさせているとき、その時刻の総トラヒック量は $T(t)$ は

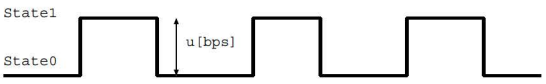

図 1. 要素トラヒック

$$T(t) = \sum_{i=1}^{N(t)} u_i(t) \tag{1}$$

と表現できる。各単位トラヒック $u(t)$[bps] が各端末で一定と考えると、 u_i は

$$u_i = \begin{cases} u(t) & : \text{確率} P(t) \\ 0 & : \text{それ以外} \end{cases} \tag{2}$$

となる。また、トラヒック量の平均 μ は $N(t),P(t),u(t)$ が定常であるなるならば

$$\mu = E[T(t)] = N(t)P(t)u(t) \tag{3}$$

と表すことができる。

1.2 マルコフ過程モデル

前節では、各要素トラヒックが複数重畳して総トラヒックを形成していると考える基本モデルについて述べたが、本報告ではその基本モデルをもとに各要素トラヒックの状態遷移が時間によらず定常である 2 状態マルコフ過程に従っているものと仮定してモデル化を試みた。そこでまず、本報告で提案する重畳マルコフ過程モデルについて説明し、その後モデルとそのパラメタについて述べる。

1.2.1 重畳マルコフ過程モデル

ここではネットワークトラヒックモデルを考慮してそれをマルコフ過程モデルに応用することを考える。総数 N の単一のトラヒックを発生させるネットワーク端末が存在しているモデルを考えたとき、各要素トラヒックは互いに独立でそれらが多数重畳して総トラヒックを形成すると考えられる。各要素トラヒックはトラヒックを全く生じさせない時を状態 0 とし、単位トラヒック u[bps] のトラヒックを生じている時を状態 1 とする。状態の遷移は時間によらず定常であり、図 3の様に表すことができる。また, このときの各状態確率を p_0,p_1 とする。また、状態 0 から 1 および状態 1 から 0 への遷移確率を $\pi_{01}(\tau)$、 $\pi_{10}(\tau)$、遷移率を λ_{01},λ_{10} とする。

Δt が非常に 0 に近いのならば次の関係式が成り立つ。

$$\begin{cases} \pi_{01}(\tau) = \lambda_{01}\Delta t \\ \pi_{10}(\tau) = \lambda_{10}\Delta t \end{cases} \tag{4}$$

各状態に存在している確率である状態確率 (State Probability) について考える。時刻 $t+\Delta t$ に状態 0 にいる確率は次のように表すことができる。

$$p_0(t+\Delta t) = p_0(t)(1-\lambda_{01}\Delta t) + p_1(t)\lambda_{10}\Delta t \tag{5}$$

この式を変形し極限をとると

$$\lim_{\Delta t\to 0}\frac{p_0(t+\Delta t)-p_0(t)}{\Delta t} = -\lambda_{01}p_0(t)+\lambda_{10}p_1(t) \tag{6}$$

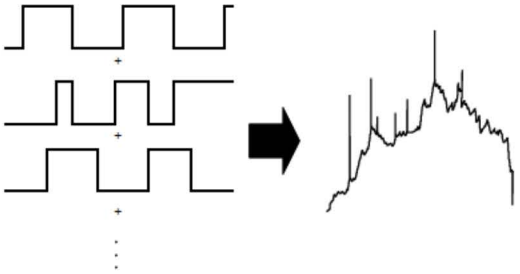

図 2. 要素トラヒックと総トラヒック

となる。 $\lim_{\Delta t\to 0}\frac{p_0(t+\Delta t)-p_0(t)}{\Delta t} = p_0'(t)$ なので, 次式のような微分方程式の形になる。

$$p_0'(t) = -\lambda_{01}p_0(t)+\lambda_{10}p_1(t) \tag{7}$$

状態遷移が定常なので、状態確率は一定である。したがって、 $p_0'(t)=0$ となり上式の微分方程式は次のようになる。

$$0 = -\lambda_{01}p_0(t)+\lambda_{10}p_1(t) \tag{8}$$

また、各状態確率の総和が 1 になるので次式のように表せる。

$$p_0+p_1 = 1 \tag{9}$$

従って、それらの式を連立すると

$$\begin{cases} p_0 = \frac{\lambda_{10}}{\lambda_{01}+\lambda_{10}} \\ p_1 = \frac{\lambda_{01}}{\lambda_{01}+\lambda_{10}} \end{cases} \tag{10}$$

となり、状態確率は遷移率 λ_{01},λ_{10} で表すことができる。

次に状態を遷移する確率である遷移確率 (Transition Probability) について考える。遷移確率の定義式はある時刻の状態が $S(t)=i$ でその τ 後の状態が $S(t+\tau)=j$ であったとき,

$$\pi_{ij}(\tau) = P[S(t+\tau)=j|S(t+\tau)=i] \tag{11}$$

となる。上式と $Chapman-Kolmogorv$ の方程式から行列の式が導かれ、その式を微分すると

$$\Pi'(\tau) = \Pi(\tau)\Lambda \tag{12}$$

という方程式が導かれる。つまり、実際には

$$\begin{bmatrix} \pi_{00}'(\tau) & \pi_{01}'(\tau) \\ \pi_{10}'(\tau) & \pi_{11}'(\tau) \end{bmatrix}$$

$$= \begin{bmatrix} \pi_{00}(\tau) & \pi_{01}(\tau) \\ \pi_{10}(\tau) & \pi_{11}(\tau) \end{bmatrix}\begin{bmatrix} -\lambda_{01} & \lambda_{01} \\ \lambda_{10} & -\lambda_{10} \end{bmatrix} \tag{13}$$

となり、この式から

$$\begin{cases} \pi_{00}'(\tau) = -\lambda_{01}\pi_{00}(\tau)+\lambda_{10}\pi_{01}(\tau) \\ \pi_{11}'(\tau) = \lambda_{01}\pi_{10}(\tau)-\lambda_{10}\pi_{11}(\tau) \end{cases} \tag{14}$$

と展開できる。これを初期条件を $\pi_{00}(0)=1$,$\pi_{11}(0)=1$ として解くと次のようになる。

$$\begin{cases} \pi_{00}(\tau) = p_0+p_1e^{-(\lambda_{01}+\lambda_{10})|\tau|} \\ \pi_{11}(\tau) = p_1+p_0e^{-(\lambda_{01}+\lambda_{10})|\tau|} \\ \pi_{01}(\tau) = 1-\pi_{00} \\ \qquad = 1-p_0+p_1e^{-(\lambda_{01}+\lambda_{10})|\tau|} \\ \pi_{10}(\tau) = 1-\pi_{11} \\ \qquad = p_1+p_0e^{-(\lambda_{01}+\lambda_{10})|\tau|} \end{cases} \tag{15}$$

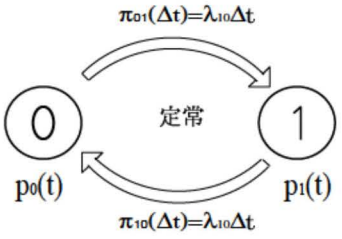

図 3. 2 状態マルコフ過程モデル状態遷移図

1.2.2　重畳マルコフ過程モデルの各パラメタ

前節で導いた確率の式を用いて本節では重畳マルコフ過程モデルを用いてネットワークトラヒックのパラメタを算出する。

まず、要素トラヒックの発生トラヒックの平均 $\overline{x}(t)$ は次のようになる。

$$\overline{x}(t) = u \cdot p_1 + 0 \cdot p_0$$

$$= \frac{\lambda_{01} u}{\lambda_{01} + \lambda_{10}} \tag{16}$$

また、総トラヒックの平均 $\overline{y}(t)$ は以下のように表すことができる。

$$\overline{y}(t) = \frac{N\lambda_{01} u}{\lambda_{01} + \lambda_{10}} \tag{17}$$

次に分散を求める。分散の定義式 $\sigma_x^2 = E[x^2] - E^2[x]$ より要素トラヒックの分散 σ_x^2、は以下のように表せる。

$$\sigma_x^2 = \frac{\lambda_{01} u^2}{\lambda_{01} + \lambda_{10}} - \frac{(\lambda_{01} u)^2}{(\lambda_{01} + \lambda_{10})^2} \tag{18}$$

$$= \frac{\lambda_{01}}{\lambda_{01} + \lambda_{10}} \left(1 - \frac{\lambda_{01}}{\lambda_{01} + \lambda_{10}}\right) u^2 \tag{19}$$

$$= \frac{\lambda_{01}\lambda_{10}}{(\lambda_{01} + \lambda_{10})^2} u^2 \tag{20}$$

また、総トラヒックの分散 σ_y^2 は $\sigma_y^2 = N\sigma_x^2$ なので次式で表すことができる。

$$\sigma_y^2 = \frac{N\lambda_{01}\lambda_{10} u^2}{(\lambda_{01} + \lambda_{10})^2} \tag{21}$$

次に要素トラヒックの自己相関関数 $R_x(\tau)$ を求める。自己相関関数は定義式から次のように表すことができる。

$$R_x(\tau) = \overline{x(t)x(t+\tau)} \tag{22}$$

$$= u^2 \cdot P[S(t) = 1 \text{かつ} S(t+\tau) = 1] \tag{23}$$

$$= u^2 \frac{\lambda_{01}}{(\lambda_{01} + \lambda_{10})^2} \left(p_1 + p_0 e^{-(\lambda_{01} + \lambda_{10})|\tau|}\right) \tag{24}$$

$$= \left(\frac{\lambda_{01} u}{\lambda_{01} + \lambda_{10}}\right)^2 + \frac{\lambda_{01}\lambda_{10} u^2}{(\lambda_{01} + \lambda_{10})^2} e^{-(\lambda_{01} + \lambda_{10})|\tau|} \tag{25}$$

同様にして総トラヒックの自己相関関数 $R_y(\tau)$ を求めると

$$R_y(\tau) = \left(\frac{N\lambda_{01} u}{\lambda_{01} + \lambda_{10}}\right)^2 + \frac{N\lambda_{01}\lambda_{10} u^2}{(\lambda_{01} + \lambda_{10})^2} e^{-(\lambda_{01} + \lambda_{10})|\tau|} \tag{26}$$

と表すことができる。次に求めた自己相関関数を用いて自己共分散関数も導く。

要素トラヒックの自己共分散関数 $C_x(\tau)$ は、

$$C_x(\tau) = R_x(\tau) - \overline{x(t)}^2 \tag{27}$$

$$= \frac{\lambda_{01}\lambda_{10} u^2}{(\lambda_{01} + \lambda_{10})^2} e^{-(\lambda_{01} + \lambda_{10})|\tau|} \tag{28}$$

同様に総トラヒックの自己共分散関数 $C_y(\tau)$ は次のようになる。

$$C_y(\tau) = \frac{N\lambda_{01}\lambda_{10} u^2}{(\lambda_{01} + \lambda_{10})^2} e^{-(\lambda_{01} + \lambda_{10})|\tau|} \tag{29}$$

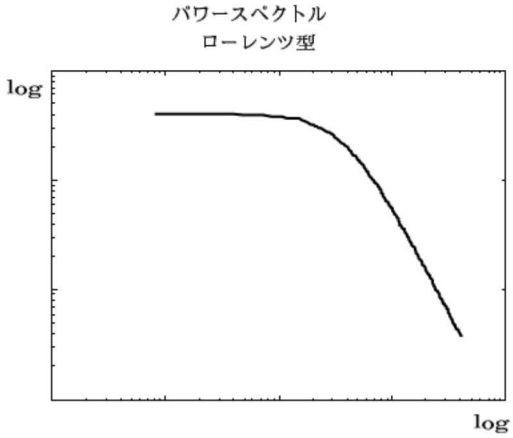

図 4. 重畳マルコフ過程モデルのパワースペクトル

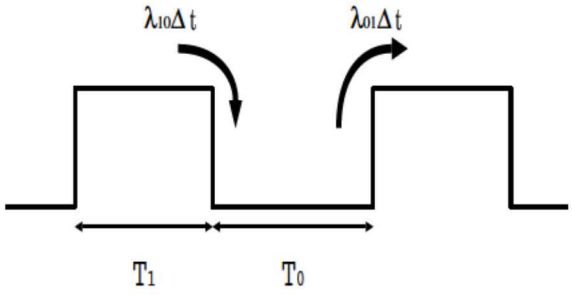

図 5. 持続時間

求めた自己共分散関数をフーリエ変換すると、ウィナー・ヒンチンの定理よりパワースペクトル $P_x(\omega), P_y(\omega)$ が求まる。まず要素トラヒックのパワースペクトル $P_x(\omega)$ から求める。

$$P_x(\omega) = \frac{2\lambda_{01}\lambda_{10} u^2 / (\lambda_{01} + \lambda_{10})}{(\lambda_{01} + \lambda_{10})^2 + \omega^2} \tag{30}$$

同様に総トラヒックのパワースペクトル $P_y(\omega)$ を求めると

$$P_y(\omega) = \frac{2N u^2 \lambda_{01}\lambda_{10} / (\lambda_{01} + \lambda_{10})}{(\lambda_{01} + \lambda_{10})^2 + \omega^2} \tag{31}$$

このパワースペクトルの式は両対数を取ると図4の様なローレンツ型を示す。

状態 0 および状態 1 の持続時間を T_0,T_1 とすると、それぞれの分布関数は次のように表すことができる。

$$\begin{cases} F_{T_0}(t) = P\{T_0 \le t\} = 1 - e^{-\lambda_{01} t} \\ F_{T_1}(t) = P\{T_1 \le t\} = 1 - e^{-\lambda_{10} t} \end{cases} \tag{32}$$

従って、それぞれの密度関数は次のようになる。

$$\begin{cases} f_{T_0}(t) = \frac{dF_{T_0}(t)}{dt} = e^{-\lambda_{01} t} \\ f_{T_1}(t) = \frac{dF_{T_1}(t)}{dt} = e^{-\lambda_{10} t} \end{cases} \tag{33}$$

ゆえに、それぞれの平均 (平均持続時間) を $\overline{T_0}$,$\overline{T_1}$ とすると、

$$\begin{cases} \overline{T_0} = \int_{-\infty}^{\infty} t f_{T_0}(t) dt = \frac{1}{\lambda_{01}} \\ \overline{T_1} = \int_{-\infty}^{\infty} t f_{T_1}(t) dt = \frac{1}{\lambda_{10}} \end{cases} \tag{34}$$

と遷移率 λ_{01},λ_{10} であらわすことが出来る。

2　シミュレーション

本章では、前章で述べた重畳マルコフ過程モデルの理論的妥当性を確認するために、トラヒックが基本モデルがに矩形波の振幅が u[bps] である要素トラヒックが多重重畳し総トラヒックを形成しているものと仮定した上でその要素トラヒックの状態遷移がマルコフ過程にしたがっているものとしてシミュレーションデータを作成し、そのデータに対して解析を行なった結果を示す。

2.1　シミュレーションデータの作成方法

以下にシミュレーションデータの作成アルゴリズムを示す。

1. 状態 0 および状態 1 の平均持続時間 $\overline{T_0}$,$\overline{T_1}$, 端末数 N, 単位トラヒック u サンプリング周期 dt, 出力データ点数を与える。
2. 与えた $\overline{T_0}$,$\overline{T_1}$ から p_0,p_1 を求める。
3. 区間 (0,1) の一様乱数 x_0 を発生させて, もし x_0 が p_0 より小さいなら初期状態を 0 とし p_0 よりも大きかったのなら初期状態を 1 として N 個の要素トラヒックの初期状態をそれぞれ定める。
4. 各要素トラヒックごとに区間 (0,1) の一様乱数列を N 個作成する。
5. 各状態の持続時間は平均値が平均持続時間 $\overline{T_0}$,$\overline{T_1}$ になる指数分布に従う。一様乱数列の i 番目の一様乱数を x_i としたとき、 i 番目の持続時間は $T_i = -\overline{T}\log x_i$ で求められるので、各要素トラヒックごとに持続時間列を求める。このとき、初期状態が 0 であるなら T_{0_1},T_{1_1},T_{0_2},T_{1_2}, …逆に初期状態が 1 ならば T_{1_1},T_{0_1}, T_{1_2},T_{0_2}, …となり、交互に状態 0 および状態 1 持続時間になっている。
6. 各要素トラヒックを標本化するために、各持続時間内を何点でサンプリングできるか調べる。持続時間列の i 番目の持続時間 T_i をサンプリング周期 dt でサンプリングする時、サンプリング点数 n_i は $i-1$ 番目のサンプリングできずに残った時間を R_{i-1} とすると次の漸化式であらわすことができる。但し、 $R_0 = 0$ とし、 n_1 は始点の分が 1 点足りないので最後に n_1 に 1 を加える。 (図 6参照)

$$t_i = T_i + R_{i-1} \tag{35}$$

$$n_i = \lfloor \frac{t_i}{dt} \rfloor \tag{36}$$

$$R_i = t_i - n_i dt \tag{37}$$

7. 作り出したサンプリング点数列を使用して各要素トラヒック時系列を求める。
8. 作り出した要素トラヒック時系列データを全て足し合わせて総トラヒック時系列を作り出す。

このようにして、各要素トラヒックが、トラヒック基本モデルと重畳マルコフ過程に従うネットワークトラヒック時系列を作り出す。例としてシミュレーションを行なった結果を図 7に示す。

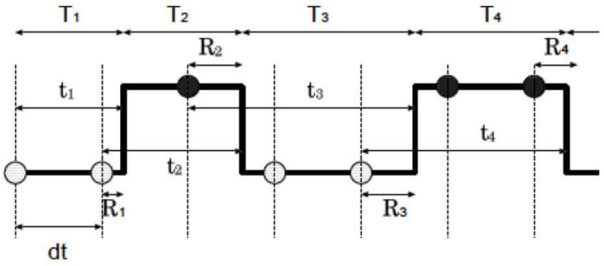

図 6. 要素トラヒックの標本化

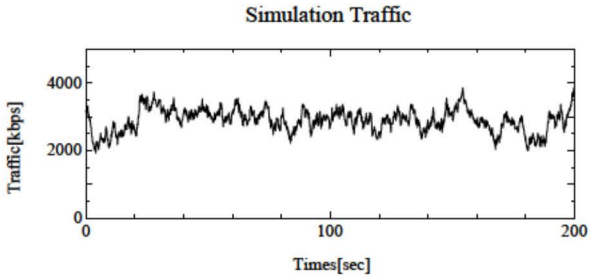

図 7. シミュレーションで作成したトラヒック時系列

2.2　シミュレーションデータの解析

2.2.1　使用したシミュレーションデータ

今回用いたシミュレーションデータは、単位トラヒック $u = 50$[Kbps], 端末数 $N = 300$ とし、各状態の平均持続時間 $\overline{T_0} = 20.0$[sec],$\overline{T_1} = 5.0$[sec] とし、それをサンプリング周期 1.0[sec] でデータ点数 2048 点でサンプリングして、求めた時系列トラヒックを解析した。

2.2.2　解析手法

重畳マルコフ過程モデルの理論的妥当性を確認するために、作り出した各シミュレーション時系列データに対して、平均値 $\overline{y(t)}$、分散 σ_y^2、自己共分散関数 C_y およびパワースペクトル $P_y(\omega)$ を推定し、モデルの理論式との比較を行なった。なお、自己共分散関数は最大ラグ $\tau_{max} = 60$[sec](12 点) で推定し、パワースペクトルは、総点 2048 点に対して、FFT 点数 256 点の Half Overlap で窓にはハミング窓を使用し、サンプリング周波数は 0.2[Hz] の高速フーリエ変換を用いて推定した。

2.3　解析結果

解析結果を以下に示す。図 8はトラヒックデータ、図 9はシミュレーションデータから得られた自己共分散関数と式 (29) から得られる重畳マルコフ過程モデルの理論自己共分散関数を図 10は FFT によって求めた片側パワースペクトルと式 (31) から求めたで理論パワースペクトル、表 1にはシミュレーション時系列データの平均、分散と重畳マルコフ過程モデルの式 (17) および式 (21) に入力条件を代入して得られた平均、分散との比較をしたものである。

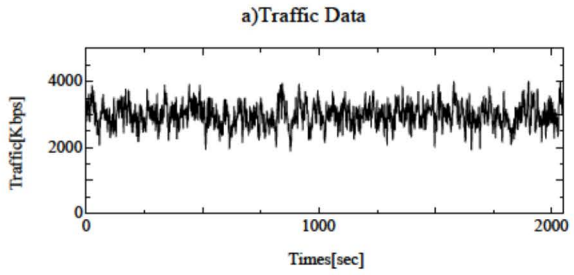

図 8. トラヒックデータ

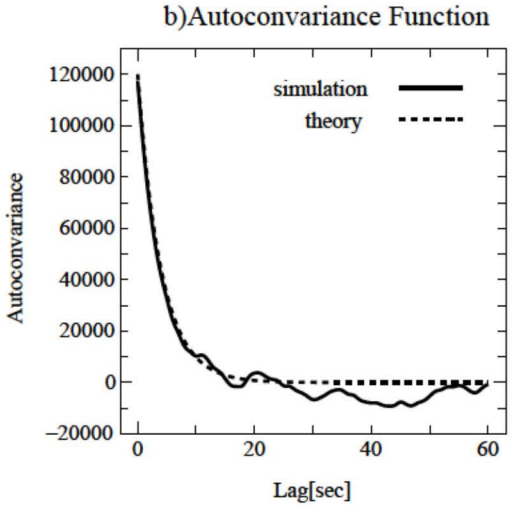

図 9. 自己共分散関数

2.4 考察

表 1からシミュレーション時系列データの平均と分散は与えられたパラメータを式 (17), 式 (21) に代入して得られる値とほぼ一致している。また、自己共分散関数も式 (29) から導かれる指数関数とラグが小さい所ではほぼ完璧に一致しており、ラグが大きくなっても大体一致しているように見受けられる。パワースペクトルも低い周波数の部分で少しずれている程度でほぼ、式 (31) から導かれる片側パワースペクトルと一致している。したがって、重畳マルコフ過程モデルで作成したシミュレーションデータは理論通りの統計的性質を有していることがわかり、定常なトラヒックへの重畳マルコフ過程によるモデリングは理論的に妥当であると考えられる。

3 ネットワークトラヒックの計測

前章では、重畳マルコフ過程モデルに従いシミュレーションデータを作成して解析を行ない重畳マルコフ過程モデルの理論的妥当性を確認した。そこで本章では、実データ解析を行なう前にネットワークトラヒックの説明をした後実データ計測に使用した SNMP と実際にどの様な方法で実データを計測し, そのデータをどの様に処理を行なって解析に用いたかを述べる。

3.1 ネットワークとトラヒック

ネットワークとは互いに関係をもちあるシステムを構築しているものであり、ここでは 0,1 の離散的電気信号をやり取りし情報を交換する情報ネットワークのことを指す。この情報の単位時間当たりの大きさをトラヒックと呼び、通信の速度、情報の密度として考えられる。

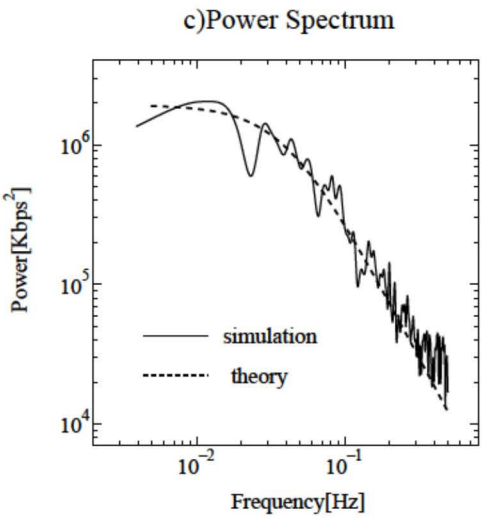

図 10. パワースペクトル

表 1. 平均と分散

	理論値	シミュレーション
平均	3000.0	120000.00
分散	2984.77	117331.39

3.2 SNMP

現在 LAN などでのネットワーク管理のツールとして、最も一般的なのが SNMP(Simple Network Management Protocal) という規約である。 SNMP は分散的にネットワークを、集中管理するために考えられた仕組みであり、管理対象となるネットワーク機器 (エージェント)、エージェントをコントロールするワークステーション (マネージャ) と MIB とよばれる管理情報ベースによって構成され、エージェントがもつ MIB にマネージャからアクセスし、ネットワークの管理対象ノードに定期的な要求を出しエージェントの管理やネットワークの状態をモニタリングする。このシステムを利用して送受信したオクテット数 (byte) などネットワークの情報を計測することができる。

3.3 ネットワークトラヒック計測

トラヒックは、通信回線にどれだけのデータ量が送受信されたかを表す指標となる。管理対象ノードにあるインターフェースを介して送受信されたデータ量は、インターフェースの送信、受信ごとに最大 $2^{32}-1$ のカウンタに MIB の累積情報として記憶される。なお、カウンタに記憶された値は最大値を越えるとゼロに戻る。この累積オクテット数を計測することでその間隔のインターフェイスにおけるトラヒック量を監視できる。すなわち、このデータ系列は累積値として与えられる為、そのサンプリング間隔におけるトラヒック量を求める場合、 1 サンプルの前の値と差分をとることで得ることができる。この得た差分を列にしたものがトラヒック時系列データである。

```
#REPORTLOG-1
D "" 133.25.253.144 100122 133.25.253.144 942031413 2382339 944017245 944017245 0 0 0
(1)(2)    (3)          (4)          (5)          (6)          (7)       (8)       (9)     (10)(11)(12)
K-25140-rt1-1.k.hosei.ac.jp ifInput "" netmgt_schema S snmp netmgt_table_key S 1 ifInOctets
          (13)              (14) (15)    (16)     (17)(18)      (19)          (20)(21)    (22)
C 501104115 netmgt_endfrow netmgt_table_key S 2 ifInOctets C 0 netmgt_endfrow netmgt_table_key
    (24)    (25)           (26)          (27)(28)  (29)    (30)(31)  (32)          (33)
S 3 ifInOctets C 1317190596 netmgt_endfrow netmgt_table_key S 4 ifInOctets C 185923476
(35)(36)    (37)(38)        (39)          (40)          (41)(42)  (43)    (44)(45)
netmgt_endfrow
    (46)

D "" 133.25.253.144 100122 133.25.253.144 942031413 2382339 944017250 944017250 0 0 0
K-25140-rt1-1.k.hosei.ac.jp ifInput "" netmgt_schema S snmp netmgt_table_key S 1 ifInOctets
C 521304216 netmgt_endfrow netmgt_table_key S 2 ifInOctets C 0 netmgt_endfrow netmgt_table_key
S 3 ifInOctets C 1418100906 netmgt_endfrow netmgt_table_key S 4 ifInOctets C 187625576
netmgt_endfrow
```

図 11. 実データのログ

3.4 実データの収集方法とログ形式

今回用いたデータは、法政大学小金井校舎の *LocalServer* セグメントにある計算機によって計測されたものを用いた。この計算機には SNMP ベースのネットワーク管理マネージャである *SunNetManager*(*SunSoft* 社) が搭載されており、これによって各ルータにおける一定時間の送受信のバイト数が計測されている。データを計測している場所は図 12の○がついているルータであり, それぞれそこの 10 時から 20 時までの送受信の累積オクテット数 (byte 数) を計測し、計測したデータをサンプリング周期 5[sec] でログと呼ぶ形式で出力する。例として計測されたデータのログを図 11に示す。

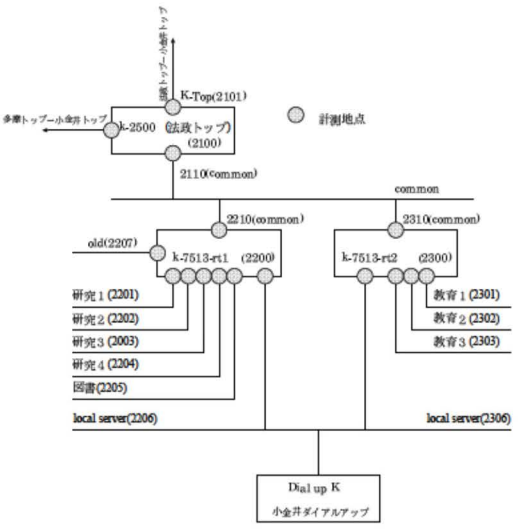

図 12. データの計測場所

このログは法政大学小金井校舎トップルータ (k-2514-rt1) で計測されたものでここでは小金井 *COMMON*、法政トップ、多摩トップと k-2514-rt1 ルータ間とのトラヒック量を計測していて、それぞれ Table Key Number として 1、3、4 が与えれている。

ログの中身は、第 1 行目の #REPORTLOG-1 は、ログを最初に書き始めるときに出力されるものであり実際の情報としてはなんら意味がない。次の D が情報を書き始める部分でここからが実際の情報が出力されている部分であり D から次の D の前の部分までが 1 行であり、同時刻の情報が記載される。実際にこのデータのログに記載されている内容は、 D を第 1 フレームとすると 1 フレーム (以後 1 フレーム= (1) のようにする) が前にも述べた様に行の先頭を表し、(2)133.25.253.144 から (8)944017245 までが収集している計算機の IP アドレス等の情報を表し、次の (9)944017245 が情報を書き込んだ時間 (UNIX TIME) で、 (13)K-2514-rt1-1.k.hosei.ac.jp が計測場所のリモートホストアドレス、(18)snmp は情報を SNMP で収集していることを表す。また、 (19)(26)(33)(40)netmgt_table_key は 2 フレーム後 (それぞれ (21)(28)(35)(42)) が計測している場所の Table Key Number で (22)(29)(36)(43)ifInOctets の 2 フレーム後の (24)(31)(38)(45) が受信してる累積オクテット数 (バイト数) となり (25)(32)(39)(46)netmgt_endofrow がその Ta-

ble Key Number での情報の終りを示している。そして次に次の何秒か後の情報がさらに D の後に記載されていきログが出力される。

この様に実データのログには解析の必要のない情報が記載されているので、必要な情報である計測時間、オクテット数のみを抜き出し、解析が行ないやすい様に加工しなければならない。そこで、次にログからデータにするまでの加工アルゴリズムを示す。

実データ加工アルゴリズム

[1] ログから計測時間, および累積オクテット数を抜きし、各テーブルキーナンバーごとに取り出す。ただし、オクテット数が $2^{32}-1$ をこえる場合にはオクテット数はゼロに戻る (図13参照) のでそこを考慮して累積オクテットを抜き出す。

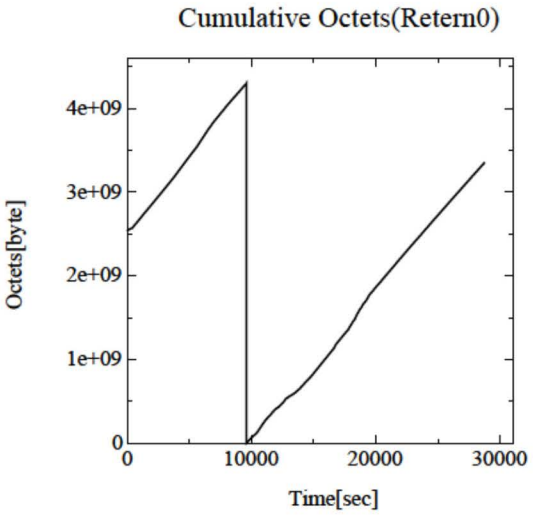

図 13. ゼロに戻った累積オクテット数

[2] 取り出したデータは、一定間隔でサンプリングされていないことがあるので、一定間隔のデータにするために線形補間を行ない等間隔のデータにする。

[3] オクテット数は図14な累積値なので前の値との差分をとる。同様に計測時間も前の値との差分をとる。

[4] 差分オクテット数は差分時間内に計測されたオクテット数なので差分時間で割り平均オクテット数 (トラヒック量 [byte/sec]) にする。

[5] データの単位を [Kbps] に直すため、 8 倍をかけ 1024 で割る。

[6] 1. から 5. で求めたトラヒック量と計測開始時間との差 (時間 [sec] となる) を出力しトラヒック時系列データを作り出す。

この様にして求めた時系列データは図 15のようになる。この 5 秒間隔のデータを使用して実データ解析を行なった。

4 実データの解析

前々章の結果から定常なトラヒック時系列データでの重畳マルコフ過程モデルの理論的妥当性が確認できた。そこで、本章では前章で説明した SNMP により計測し取得した 5 秒間隔サンプリングのトラヒック時系列データ対して解析を行なう。

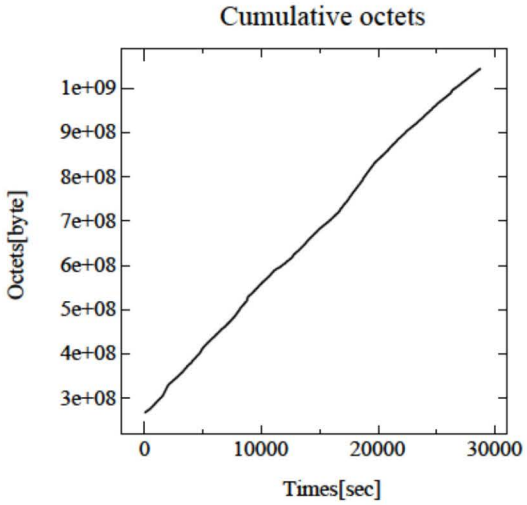

図 14. 累積オクテット数

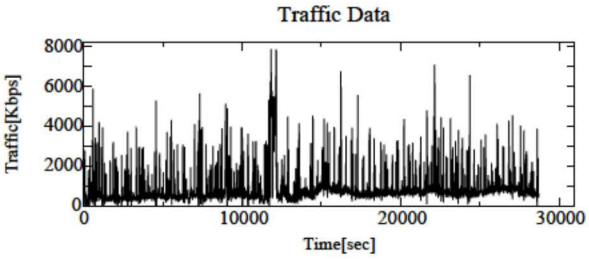

図 15. ログから取り出したトラヒック時系列データ

4.1 解析に使用したデータおよび解析方法

1999 年 11 月 9 日に得られた法政大学小金井キャンパス 7513 ルータ 1- 研究 2 の出力側データ (データ 1) に対し、そのほぼ定常な部分抜きだしそれの自己共分散関数とパワースペクトルを推定した。なお、自己共分散関数は最大 lag が 50 秒 (10 点) で求め、スペクトル推定にはデータ総点 2048 点に対し、ポイント数 256 点、サンプリング周波数 0.2[Hz]、ハミング窓、 Half Overap の高速フーリエ変換を用いた。また、 1999 年 11 月 29 日に得られた同大学同キャンパス 7513 ルータ 1- 研究 2 の出力側データについても (データ 2) に対し、同様にの自己共分散関数とパワースペクトルを推定した。なお、自己共分散関数は最大 lag が 100 秒 (20 点) で求め、スペクトル推定にはデータ総点 2748 点に対し、ポイント数 256 点、サンプリング周波数 0.2[Hz]、ハミング窓、 Half Overap の高速フーリエ変換を用いた。

4.2 解析結果

解析結果を以下に示す。まず、図 16は解析したデータ 1 のトラヒック時系列データ図 17にはトラヒック時系列データ (データ 1) から求めた自己共分散関数と指数関数 y =3.6 96044 $\times 10^5 \exp[-0.118\tau]$ を図 21は片側パワースペクトルと指数関数の時定数および $\tau = 0$ の時の値 (分散) から式

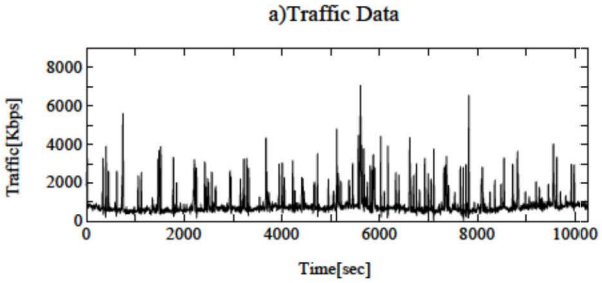

図 16. トラヒックデータ (データ 1)

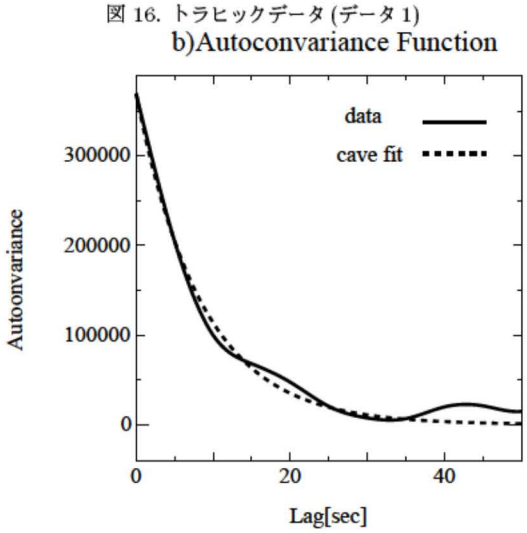

図 17. 自己共分散関数 (データ 1)

(31) に代入して求めた片側パワースペクトルをそれぞれ示し、同様にデータ 2 についても図 19にトラヒック時系列データ、図 20データ 2 の自己共分散関数と指数関数 y =5.019066 × 10^5exp[-0.071τ]、図 18にはデータ 2 のパワースペクトルおよび指数関数から算出したパワースペクトルを示す。

4.3 考察

まずデータ 1 の時系列データの解析結果において図 17においては自己共分散関数は指数関数 y =3.696044 × 10^5exp[-0.118τ] とほぼ一致している。また、図 21のパワースペクトルにおいても指数関数の $\tau = 0$ の値と時定数を理論式に代入して得られた理論式のローレンツ型パワースペクトルとほぼ同じように分布していることが見受けられ、同様にデータ 2 においても図 20の自己共分散関数は指数関数 y =5.019066 × 10^5exp[-0.071τ] とほぼ一致しており、パワースペクトルも図 18からローレンツ型のパワースペクトルを示していることが分かる。このことから、定常に近いトラヒック時系列データにおいては、トラヒック基本モデルとマルコフ過程を合わせた重畳マルコフ過程モデルによってモデル化を行なうことができると考えらる。

あとがき

本稿では、まずネットワークトラヒックは通信パケットが送信されているか否かの 2 状態から説明できると考えて、各自独立な各要素トラヒックがある一定の大きさのトラヒックを発生させている状態と全くネットワーク負荷を生じさ

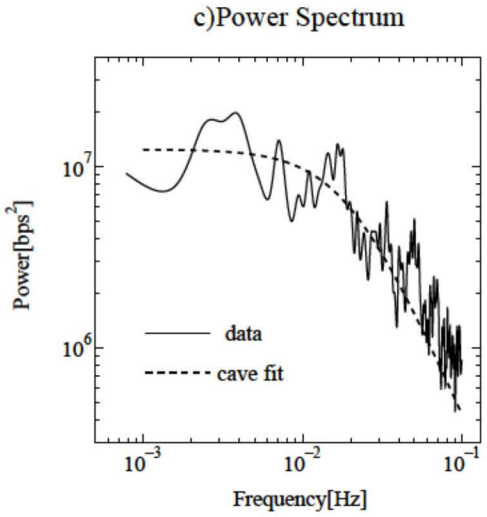

図 18. パワースペクトル (データ 1)

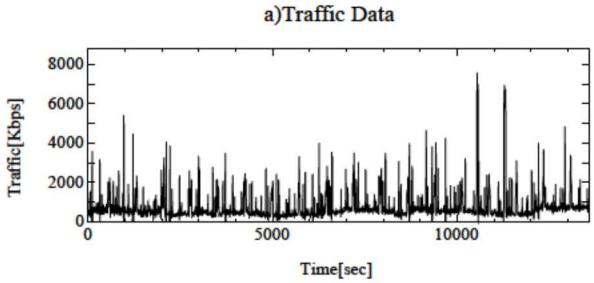

図 19. トラヒックデータ (データ 2)

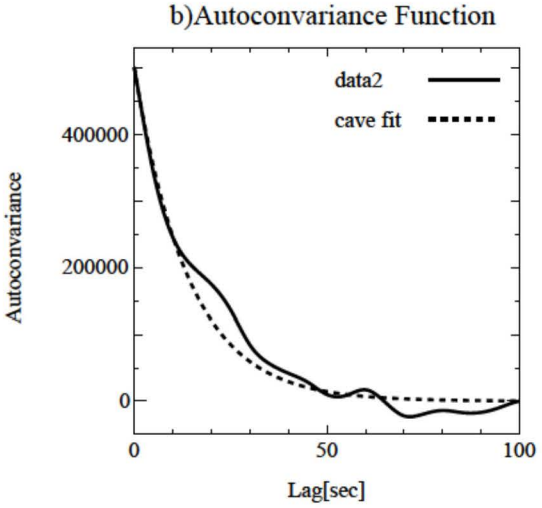

図 20. 自己共分散関数 (データ 2)

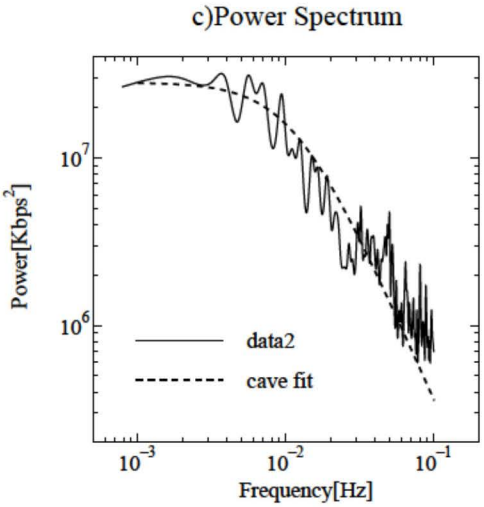

図 21. パワースペクトル (データ 2)

せていない状態の 2 状態を遷移しているとし、それらが複数重畳して総トラヒックとなるネットワークトラヒック基本モデルを考え、それに状態遷移がマルコフ過程にしたがってるとした重畳マルコフ過程モデルを提案し、そのモデルから得られるトラヒックの平均、分散等のパラメタの算出を行なった。そして、その理論的妥当性を確認するため重畳マルコフ過程モデルに従うシミューレション方法を説明し、実際にシミュレーションによってトラヒック時系列データを作成して解析を行なった。その結果、理論式から得られるパラメタとシミュレーションによるトラヒック時系列データから得られるパラメタを比較することによってネットワークトラヒックに対する重畳マルコフ過程モデルの妥当性を確認した。その後、本稿で実データ解析を行なう際に用いたデータについてその収集場所やデータの処理の仕方などの説明を行なった上で実データ解析を行ない重畳マルコフ過程モデルがネットワークトラヒックをモデル化できることを確認した。このことにより、ネットワークトラヒックの各統計量が簡単なモデル化を行なうことにより導き出せることや逆に統計量から、他のパラメタを推定できる可能性を見出すことができた。

参考文献

[1] ロバート・ B・クーバー, 秋丸 春夫, “通信トラヒック工学,” オーム社.

[2] Mark A. Miller 著, トップスタジオ訳, 宇野俊夫監修, “SNMP インターネットワーク管理,” 翔泳社, 1998.

[3] 名部雅彦, 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” 信学論 (B-I),Vol.J80-B-I,No.5,pp.428-437,1997.

[4] H.Michiel, and K.Laevens, ” Teletraffic engineering in a Boad-Band era,” Proc.IEEE,Vol.85,No.12,DEC 1997.

[5] D.Lucantoni, ” New results on the single server queue with a batch Markovian arrival process,” Stochastic Models, vol.7,pp.1-46,1991.

[6] C.Blondia,”A discrete-time batch Markovian arrival process as B-ISDN traffic model,” Belgian J. Oper.Res.Star.Comput.Sci.,vil.32,no.3,pp.3-23,1992.

キーワード.

ネットワークトラヒック、マルコフ過程、SNMP

Summary.

Statistical Properties of the Information Network Traffic and its Modeling.

Hideaki Shino Takeshi Ueda Kazuo Yana
Department of Electrical Infomatics , Colleage of Engineering , Hosei University

Statistical modeling of the information network traffic is important in network design which enables a large scale computer simulation before an actual system construction. The modeling may also be used for statistical estimation of useful parameters such as mean traffic duration or mean bandwidth of a single traffic. We have examined the statistical properties of actual network traffic of our campus and found that the autocorrelation function fits well with a single exponential curve. A Markov model is compatible with this oversrvation. A useful application of the Markovian model for estimating traffic parameters based on observed traffic time series will be described.

Keywords.

Network Traffic , Markov ModelingProcess , SNMP