

情報ネットワークにおけるSNMP時系列のハーストパラメータ推定

OHKUBO, Tomofumi / 秋元, 仁 / SHINO, Hideaki / AKIMOTO, Hitoshi / 大久保, 智史 / YANA, Kazuo / 篠, 秀明 / 八名, 和夫

(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)

14

(開始ページ / Start Page)

135

(終了ページ / End Page)

140

(発行年 / Year)

2001-03-31

(URL)

<https://doi.org/10.15002/00024929>

情報ネットワークにおける SNMP 時系列のハーストパラメータ推定

篠 秀明 大久保 智史 秋元 仁 八名 和夫
法政大学工学部電子情報学科

あらし

近年の WWW(World Wide Web) の利用を中心とするインターネットの急速な普及によりネットワーク上には大量のトラフィックが生じており、さらにそのトラフィックは増加の傾向にある。そこで、ネットワークトラフィックの統計的性質を明らかにすることによって分散制御などをおこない負荷を分散することでネットワークの効率的に運用することが必要となる。ネットワークトラフィックには自己相似性とよばれる時間スケールの変換に対して変動性が変わらないという性質がみられるという報告がなされており、同時にこの自己相似性がパケット棄却率やファイル転送遅延時間などといったネットワークの性能の低下をもたらすという報告も存在する。そこで本論文では、SNMP ネットワークトラフィック時系列における自己相似性の有無の検証をするために R/S 解析を用いて自己相似性の強さを示す指標であるハーストパラメータを推定した。

1 まえがき

近年のインターネットの急速な普及により WWW の利用が急増しネットワーク上には大量のトラフィックが生じており、そのトラフィックはさらに増加する傾向にある。しかし、データ通信を行なう媒体の通信回線には転送容量が決まっており容量を越える状態またはそれに準じる状態ではスループットの低下、遅延等が発生する。そのためトラフィック量を軽減しスループットを向上するようなシステムが要求されており、現状ではキャッシュサーバの導入や回線容量を増すなどの措置がとられている。また、大規模ネットワークではネットワーク構造が複雑になりがちであり無駄になる回線が存在している事がある。そういった状況ではネットワークトラフィックの特性を把握しネットワークを効率的に運用することが必要となり、それはトラフィックの分散制御などをおこない負荷を分散する事によって可能となる。トラフィックの最適な分散制御を行なうためにはネットワークトラフィ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が必要となる。

近年の研究において LAN(Local Area Network) や WAN(Wide Area Network) などにおけるネットワークトラフィックに対して自己相似性とよばれる時間スケールの変換に対して変動性が変わらないという性質がみられるという報告がなされており [1]、同時にこの自己相似性がパケット棄却率やファイル転送遅延時間などといったネットワークの性能の低下をもたらすという報告も存在する [2]。

SNMP はネットワーク管理システム専用のプロトコルであり、大規模なネットワークでも比較的簡単にネットワークの管理、監視やトラフィックの監視が可能である。そこで本報告では、SNMP により得られるネットワークトラフィック時系列について自己相似性の強さを示す指標であるハーストパラメータをスケール変換解析 (R/S 解析) によって推定する事で自己相似性の強さを推定した。

vspace1.0cm

2 自己相似性

本章では、自己相似性の定義と自己相似性の度合を表すハーストパラメータについて説明する。

$X = (X_i; i = 1, 2, 3, \dots)$ は平均 μ , 分散 $\sigma^2, k \geq 0$ に対し X_i と X_{i+k} の間の自己相関係数 $\rho(k)$ を持つ定常過程とする。また、 $m = 1, 2, 3, \dots$ に対して $X^{(m)} = (X_i^{(m)}; i = 1, 2, 3, \dots)$ は自己相関係数 $\rho^{(m)}(k)$ を持つ定常過程とする。ここで $X_i^{(m)}$ は式 (1) のように表される。

$$X_i^{(m)} = \frac{1}{m} \sum_{j=(i-1)m+1}^{im} X_j \quad (1)$$

自己相似性は時間スケールの変換に対し不変に保たれる量として、有限次元分布と 2 次の特性量 (分散と自己共分散) を考えると以下の 3 つの定義が存在する。

厳密な自己相似性 (exactly self-similarity) : すべての $m \geq 1$ に対して式 (2)(3) をみたす $X^{(m,H)}$ が X と同じ確率法則に従う時、確率過程 X は自己相似性をもつという。

$$X^{(m,H)} = (X_i^{(m,H)}; i = 1, 2, 3, \dots) \quad (2)$$

$$X_i^{(m,H)} = \frac{1}{m^H} \sum_{k=0}^{m-1} X_{i+m+k} \quad (3)$$

厳密な 2 次の自己相似性 (exactly second-order self-similarity) : すべての $m \geq 1$ に対して式 (4), (5) が成り立つとき、 X は H をハーストパラメータとする厳密な 2 次の自己相似性をもつという。

$$Var(X^{(m)}) = \sigma^2 m^{2H-2} \quad (4)$$

$$\rho^{(m)}(k) = \rho(k), k \geq 0 \quad (5)$$

これらの式 (4),(5) は $X^{(m)}$ が X と同じ分散と自己共分散をもつことを示している。

漸近的な2次の自己相似性 (asymptotically second-order self-similarity) : すべての $k \geq 0$ に対して式(6)が成り立つとき、 X は H をハーストパラメータとする漸近的な2次の自己相似性をもつという。

$$\rho^{(m)}(k) \rightarrow \rho(k) \quad \text{us} \quad m \rightarrow \infty \quad (6)$$

漸近的な2次の自己相似性とは十分大きな時間スケールの領域において、その時間スケールを変化させても2次の確率的挙動が変化しないことを意味する。

ここで H はハーストパラメータと呼ばれ、自己相似性が存在するときには $0.5 < H < 1.0$ となる。このハーストパラメータの値によって次のように元の時系列の性質が決定する。

- $0 < H < 0.5$ (短記憶性)

自己相関関数の総和が収束することから短記憶となる。

$$\sum_{k=-\infty}^{\infty} \rho(k) < \infty \quad (7)$$

- $H = 0.5$ (無相関)

$k = 0$ の点以外での自己相関関数の値が0であることから無相関である。

$$\forall k (\neq 0), \rho(k) = 0 \quad (8)$$

- $0.5 < H < 1.0$ (長記憶性)

自己相関関数の総和が発散することから長期依存性となる。

$$\sum_{k=-\infty}^{\infty} \rho(k) = \infty \quad (9)$$

- $H = 1.0$

このとき自己相関関数は k の値に依存せず一定となる。

$$\rho(k) = 1 \quad (10)$$

- $H > 1.0$

このとき自己相関関数の極限は式(11)のようになり $|\rho(k)| \leq 1$ に矛盾する。

$$\lim_{k \rightarrow \infty} \rho(k) = \infty \quad (11)$$

3 スケール変換解析 (R/S 解析)

本章では、ハーストパラメータの推定法であるスケール変換解析について説明する。

時系列データ X を m 個ずつ分けたとき、 l 番目のグループの中で i 番目までの確率変数の和がそのグループ内の平均に対して、どの程度ずれているかを表す $Z(l, m, i)$ を式(12)のように定義する。

$$Z(l, m, i) = \sum_{j=(l-1)m+1}^{(l-1)m+i} X_j - iX_l^{(m)} \quad (12)$$

この $Z(l, m, i)$ の最大値と最小値を用いて、式(13)のように、 $R(l, m)$ を定義する。

$$R(l, m) = \max_{1 \leq i \leq m} Z(l, m, i) - \min_{1 \leq i \leq m} Z(l, m, i) \quad (13)$$

さらに、標本標準偏差 $S(l, m)$ を式(14)のように定義する。

$$S(l, m) = \sqrt{\frac{1}{m} \sum_{j=(l-1)m+1}^{lm} (X_j - X_l^{(m)})^2} \quad (14)$$

$S(l, m)$ に対する $R(l, m)$ の比は R/S 統計量と呼ばれ、この R/S 統計量の対数をとることにより式(15)のように近似でき、その傾きがハーストパラメータ H となる。

$$\log E \left[\frac{R}{S} \right] \approx a + H \log m \quad (15)$$

4 シミュレーションデータの解析

本章では、実データの解析を行なう前にハーストパラメータが既知である $1/f$ 特性を持つ時系列に対し R/S 解析を用いハーストパラメータを推定し、 R/S 解析によってハーストパラメータが良好に推定されていることを確認する。始めに $1/f$ 特性について説明し、その後解析手法と解析結果を示す。

4.1 $1/f$ 特性

$1/f$ 特性とはパワースペクトルが周波数 f の逆数に比例するような性質でそのことから長期依存性が認められ漸近的な自己相似性が導かれる。よってハーストパラメータ $H = 1$ となり、長記憶と非定常の境界にあるといわれている。この性質をもつ雑音は、 $1/f$ 雑音、フリッカー雑音、ピンク雑音と呼ばれ、水晶片の振動、高速道路の自動車の流れ、気温の季節的変動、音楽、心拍など、様々な現象に対して普遍的に観測されている。また、医療関係では除痛に利用する試みもなされている。しかし、その発生原因はいまだにほとんどわかっていない。

4.2 解析に使用したデータおよび解析手法

$1/f$ 特性をもつ 1000, 2000, 3000, ..., 10000 点の時系列をそれぞれ 10 例作成し、それらについて R/S 解析を用い m がデータ点数の 10 分の 1 の点までの R/S 統計量を算出し、得られた R/S 統計量に最小二乗法を用いることでハーストパラメータを推定し、各点数における 10 例のハーストパラメータの平均値を求める。さらに、こちらは 1 例であるがデータ点数がより多い事例として 10 万点の $1/f$ 特性をもつ時系列に対しても同様な手法でハーストパラメータを推定した。

4.3 解析結果

1000,2000,3000,...,10000 点の時系列の 10 例のハーストパラメータの平均値を表 (1) に示す。1/f 特性をもつ 10 万点の時系列を図 (1) にその R/S 統計量を図 (2) に示す。1/f 特性をもつ 10 万点の時系列のハーストパラメータを求めた結果、

$$H = 0.9232633$$

となった。

表 1. 各データ点数における H の平均値

データ点数	H
1000	0.72779691
2000	0.80123567
3000	0.82793679
4000	0.83419444
5000	0.85091388
6000	0.84691685
7000	0.85838569
8000	0.87163609
9000	0.87105785
10000	0.90575276

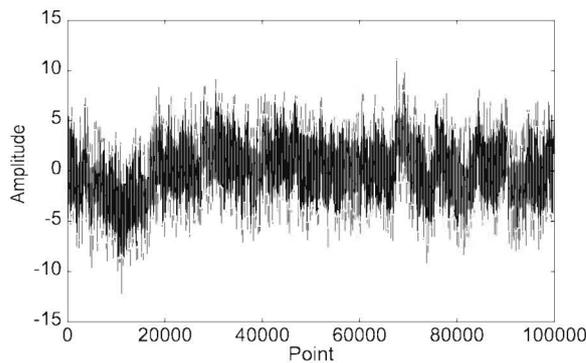


図 1. 1/f 特性をもつ 10 万点の時系列

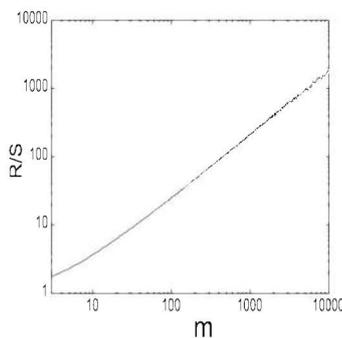


図 2. 1/f 特性をもつ 10 万点の時系列の R/S 統計量

4.4 考察

1/f 特性をもつ時系列について R/S 解析を用いハーストパラメータを推定し 10 例の平均をとった結果、1/f 特性をもつ時系列のデータ点数を大きくしていくとハーストパラメータの値が大きくなっていくことがわかった。さらにデータ点数が大きな 10 万点の時系列についてハーストパラメータを推定したところ、同様な結果が得られた。そしてそのハーストパラメータ H が $0.5 < H < 1.0$ の条件を満たすことから強い自己相似性が存在することが確認された。しかし、 $H = 1$ とならないことから 10 万点でもデータ点数が足りないと思われる。さらに、図 (2) から m が小さい範囲における R/S 統計量が全体の直線からずれていることから、この範囲における R/S 統計量が推定されるハーストパラメータの値を低くしているのではないかと推測される。

5 実データ解析

解析に用いるデータは、法政大学小金井キャンパスに設置されているルータにおいて観測される SNMP により得られるネットワークトラフィック時系列である。このネットワーク管理プロトコルである SNMP を利用するネットワーク性能計測ツールを使用することによってネットワークにおけるデータ収集、性能評価を行うことができる。本章では、解析に用いた SNMP により得られる実測ネットワークトラフィック時系列とそのデータの解析結果について述べる。

5.1 実測 SNMP 時系列

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ネットワーク管理プロトコルであり、管理対象となるネットワーク機器 (管理対象ノード) とその機器をモニタリングするプロセス (エージェント) と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)、そしてエージェントとマネージャそれぞれがもつ MIB (Management Information Base) とよばれる管理情報ベースによって構成され、マネージャは定期的エージェントがその管理対象ノードの状況を記録した MIB にアクセスし、マネージャ自身の MIB に状況を記憶するようなモデルを前提としている。トラフィック量は通信回線にどれだけデータ量が送受信されたかを表す指標であり、SNMP においてはある間隔ごとにマネージャがエージェントの MIB を参照し総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フィック量が計測できる。本学は上記のような構成になっており、トラフィック量は 5 分ごとにネットワーク管理ステーションである本研究室内のマシンから管理対象ノードである本学各ルータを参照し計測したものを解析に用いる実測 SNMP 時系列とした。

5.2 解析に使用したデータおよび解析方法

実データ解析に用いるデータは法政 TOP-SINET 間、小金井キャンパス研究室1、小金井キャンパスにおけるダイヤルアップのルータの3点で観測される1999年11月1日(月)午前0時から11月8日(月)午前0時までの1週間分のネットワークトラフィック時系列に対し、R/S解析を用い m がデータ点数の10分の1の点までのR/S統計量を算出し、得られたR/S統計量に最小乗法を用いることでハーストパラメータの推定を行った。

5.3 解析結果

法政 TOP-SINET 間、小金井キャンパス研究室1、小金井キャンパスにおけるダイヤルアップの各ルータにおける1週間分のネットワークトラフィック時系列を図(3)に、それぞれのR/S統計量を図(4)に示す。さらに、各観測ポイントにおけるハーストパラメータの値を表(2)に示す。

表 2. 各観測ポイントにおける H の値

観測ポイント	H
HoseiTOP-SINET	0.929439
res-1	0.842822
DialUP.k	0.986832

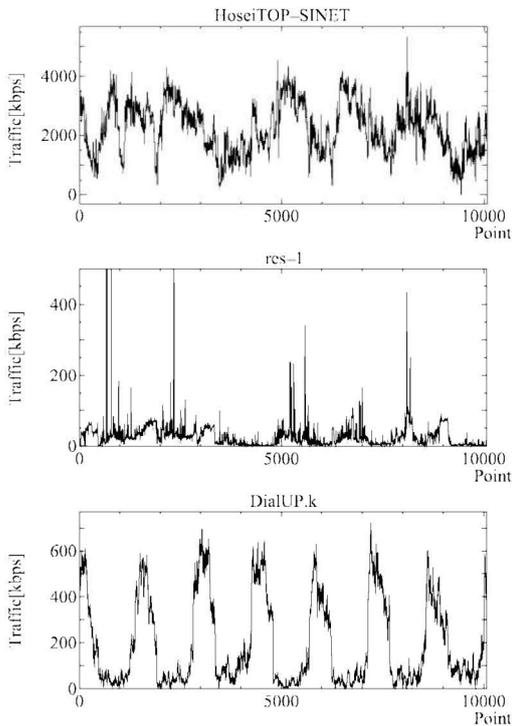


図 3. 各観測ポイントにおけるネットワークトラフィック

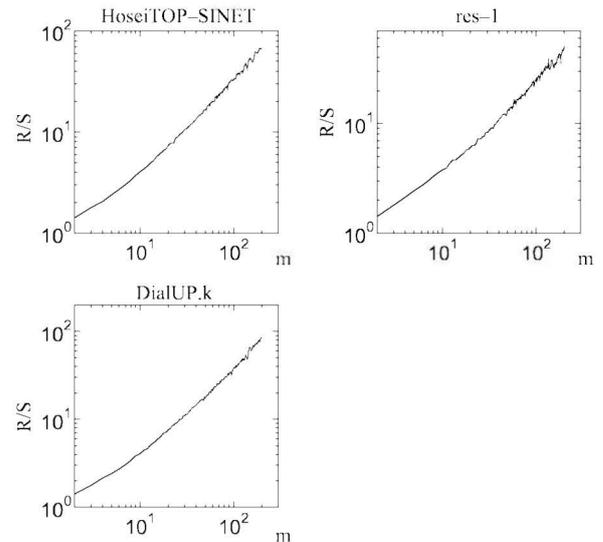


図 4. 各観測ポイントにおけるネットワークトラフィックのR/S統計量

5.4 考察

法政大学小金井キャンパスに設置されている各ルータにおいて観測される1週間分のSNMPにより得られるネットワークトラフィック時系列からハーストパラメータを推定した結果、ハーストパラメータ H が自己相似性をもつときの条件 $0.5 < H < 1.0$ を3点のルータすべてにおいて満たしていることから今回解析に使用した3点のルータで観測されるSNMPにより得られるネットワークトラフィックすべてにおいて自己相似性をもつことが示された。

6 おわりに

本報告では、ネットワークの効率的運用のために必要となるネットワークトラフィックの統計的性質を理解することを目的として、SNMPにより得られるネットワークトラフィックにおける自己相似性の存在の有無の検証を、R/S解析によって自己相似性の強さを示す指標であるハーストパラメータを推定することで行った。実際に観測されるSNMPにより得られるネットワークトラフィック時系列の解析を行う前に、ハーストパラメータが既知である $1/f$ 特性をもつ時系列についてシミュレーションを行い、ハーストパラメータが良く推定されていることを確認した。そこで、法政大学小金井キャンパスにおけるルータ3点で観測される1週間分のSNMPにより得られるネットワークトラフィック時系列について自己相似性の有無の検証を行った結果、いずれの点においてもハーストパラメータ H が $0.5 < H < 1.0$ の条件をみたしていることから観測されるネットワークトラフィック時系列が自己相似性を有していることを示した。

参考文献

- [1] W.E. Leland, W. Willinger, M.S. Taqqu, and D.V. Wilson, "On the self-similar nature of Ethernet traffic," SIGCOMM '93, pp.183-193, 1993.
- [2] 住田 義明, 大崎 博之, 村田 正幸, 宮原 秀夫, "上位層プロトコルがネットワークトラヒックの自己相似性に与える影響" 電子情報通信学会論文誌 B Vol.J82-B No.6, pp.1126-1137, June. 1999
- [3] J・フェダー著, 松下 貢, 早川 美徳, 佐藤 信一訳, "フラクタル", 啓学出版, 1991.
- [4] 小沢 利久, "トラヒックはいつ自己相似となるのか", 電子情報通信学会技術研究報告 IN2000-33, pp.1-8, 2000.
- [5] 小沢 利久, 町原 文明, 石橋 圭介, "マルチメディアトラヒック理論の最新動向", 信学誌, Vol.81, no.5, pp.506-515, May. 1998
- [6] 工藤 誠也, 高木 英明, 濱田 元, 久保田 文人, "ATM 網における圧縮動画トラヒックの自己相似性", 電子情報通信学会論文誌 B-I Vol.J81-B-I No.9, pp.549-556, September. 1998
- [7] 鈴木 良次, 佐藤 俊輔, 池田 研二, 吉川 昭, "生体信号", コロナ社, 1989.
- [8] 星宮 望, 石井 直宏, 塚田 稔, 井出 英人共著, "生体情報工学", 森北出版, 1986.

キーワード.

ネットワークトラフィック解析、SNMP 時系列、自己相似性、ハーストパラメータ、R/S 解析

Summary.

On Self-similarity of the SNMP Time Series of an Information Network

Hideaki Shino Tomofumi Ohkubo Hitoshi Akimoto and Kazuo Yana
Department of Electronic Informatics, College of Engineering, Hosei University

Network traffic is currently exploding worldwide due to the rapid development of broadband network technology and communication demand such as WWW browsing. To clarify the statistical properties of the network traffic is important for designing an efficient network system. This paper confirmed the presence of self-similarity in a university LAN traffic i.e. SNMP time series by the R/S analysis. The evidence will be useful in the prediction of network traffic and evaluation of network devices.

Keywords.

Network Traffic Analysis , SNMP Time Series , Self-similarity , Hurst Parameter , R/S Analysis