

トレンド除去変動解析法によるIPトラフィック長期記憶性の自己組織化解析

A Self-Organizing Approach to Measuring Long-Range Dependence in IP Traffic using Detrended Fluctuation Analysis

馬杉 正男*

Masao MASUGI*

1. はじめに

IPネットワークトラフィックに観測される自己相似性とネットワーク品質との関係が、近年注目されている。しかしながら、トラフィック自己相似性の評価尺度(ハースト指数等)の導出については、通常、対象信号が定常性・弱定常性を有することを前提条件としており、実環境データへの適応性が課題として残されている。

そこで本報告では、非定常信号への拡張を目的として提案された Detrended Fluctuation Analysis (DFA) [1], [2] に着目した。以下、自己組織化処理[3]を用いる方法により、IPネットワークトラフィックに観測される自己相似性の時間変動パターンの可視化法について、DFAをベースに検証する。

2. 自己組織化手法を用いたトラフィックのDFA解析

(1) DFA手法の概要

上述したように、自己相似性の評価尺度であるハースト指数等の導出法は、通常、対象信号が定常性・弱定常性を有することを前提条件とする。しかしながら、解析対象とする信号には、多くの場合、非定常な変動成分が含まれるため、従来の評価尺度は適応上の制約があると考えられる。こうした課題を解決手段として、非定常信号に含まれる緩やかな変動成分を除去する Detrended Fluctuation Analysis (DFA) 手法が提案されている[1]。

まず、 $x(i) \{i=1, 2, \dots, N\}$ を時系列データとすると、全体平均 M より、新たな時系列 $y(k)$ を定義する。

$$y(k) = \sum_{i=1}^k \{x(i) - M\} \quad (1)$$

次に $y(k)$ を等間隔 n の区間で分割後、各区間内において最小2乗近似直線 $y_n(k)$ (ローカルトレンド) を導出し、

$$F(n) = \left[\frac{1}{N} \sum_{k=1}^N \{y(k) - y_n(k)\}^2 \right]^{1/2} \quad (2)$$

を定義する。全時間スケールに対して $F(n)$ を計算し、 $\log(n)$ vs. $\log(F(n))$ のプロットよりスケーリング指数 α が得られる。

ここで、対象信号は、 $\alpha \approx 0.5$ の場合に無相関性、 $\alpha > 0.5$ の場合に長期相関性(あるいは長期記憶性)を有することが知られており、非定常信号のスケーリング解析が可能となる。

(2) 自己組織化処理の概要

本報告では、文献[3]の方法を用いて入力データを2次元平面上に自己組織化マッピングする処理法を採用する。まず、多次元入力ベクトル x と2次元出力層上の参照ユニット i が参照ベクトル m_i により結合した状況を想定する。ここで、離散ステップを t として、次式による処理に基づいて学習を行う。

$$m_i(t+1) = m_i(t) + h(t) \{x(t) - m_i(t)\} \quad (3)$$

ここで、 $h(t)$ は学習率係数を考慮した近傍関数であり、離散ステップ t の増加とともに単調減少する。学習過程では、距離 $|x(t) - m_i(t)|$ を最小にするノード i を探索し、類似度に対応するように入力ベクトルが2次元出力層上にマッピングされる。

(3) トラフィック解析法

ネットワークトラフィックとネットワーク品質との関係を考えた場合、自己相似性のみに着目するのは必ずしも十分とは言えない。すなわち、ネットワーク利用帯域が限定された状態で、長期記憶効果が高くてもネットワーク品質がそれほど低下するわけではなく、各種パラメータを統合的に判断する必要がある。

そこで本報告では、自己組織化マップを用いる方法により、異なるスケーリング領域で観測される2つの指数 α と平均スループットを入力パラメータとする自己組織化処理を実施し、ネットワークトラフィックの時間変動状態を可視化する。

3. 評価例

(1) トラフィック測定とDFA解析

DFA手法の評価例として、インターネットよりNTT武蔵野R&Dセンタに流入するIPネットワークトラフィックを実測対象とした(時間分解能=10ms)。

まず、ピーク値で規格化した実測トラフィック例を図1に示す。同図より、バースト性を有する激しい振動が実測トラフィックに観測され、自己相似的な変動特性を有することがわかる。続いて、10分単位長(6万ポイント)を1データとしたスケーリング指数の導出例を図2に示す。同図より、 $n=2.2$ 付近を前後に2つのスケーリング指数 α_j ($j=1, 2$) が導出され、実測トラフィックが異なる確率的な構造を有することがわかる。

続いて、平日9:00~11:30の間、10分単位のデータを5分間隔で取得し、規格化した平均スループットと2つのスケーリング指数の関係を図3に示す。なお、同図のプロット番号は、測定開始からのデータ順序に対応し、平均スループットは最大値で規格化している。同図より、①異なる領域より導出された各指数

*: 日本電信電話株式会社 NTT環境エネルギー研究所, NTT Energy and Environment Systems Laboratories, NTT Corporation.

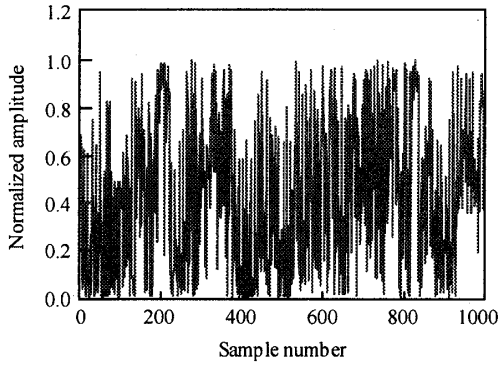


図1. 実測トラヒック例

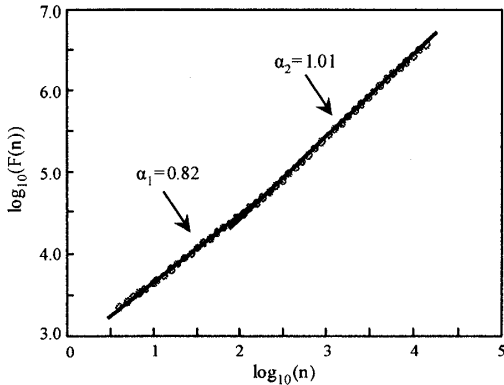
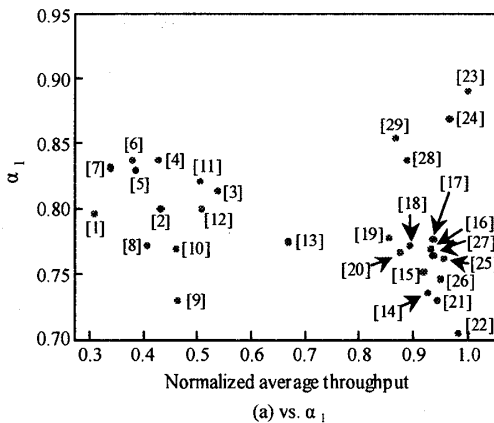
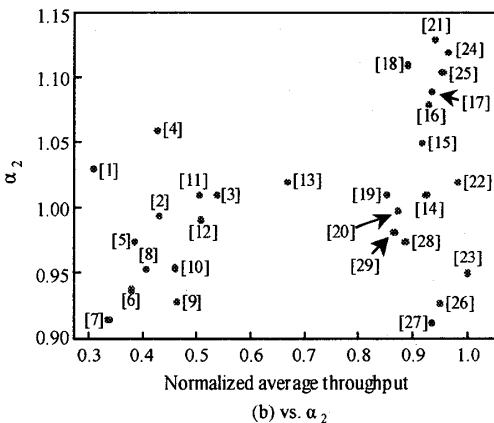


図2. スケーリング指数 α の導出例



(a) vs. α_1



(b) vs. α_2

図3. 平均スループットと指数 α の関係

が 0.70~0.89 と 0.92~1.13 に分布し、 n が大きい領域で長期記憶効果が増加する、②時間経過とともに、規格化スループットが増加傾向にある、等が確認できる。

(2) 自己組織化処理

続いて、図3のデータを元に3つの入力変数 (α_1 , α_2 , 規格化スループット) を元に自己組織化処理を実施した。10×10のマップについて、計1万回の学習を実施した際のマッピング結果を図4に示す。

同図より、① α_1 はx軸の低下とともに増加傾向にあり、左最上部が最大値をとる、② α_2 はy軸の増加とともに増加傾向にあり、最上部が最大値をとる、③規格化スループットは、x軸とy軸に沿って増加傾向にある、等が確認できる。すなわち、同図は、提案手法により、トラヒックが有する多次元的な時間変動特性が2次元平面上に効率的に可視化されることを示している。

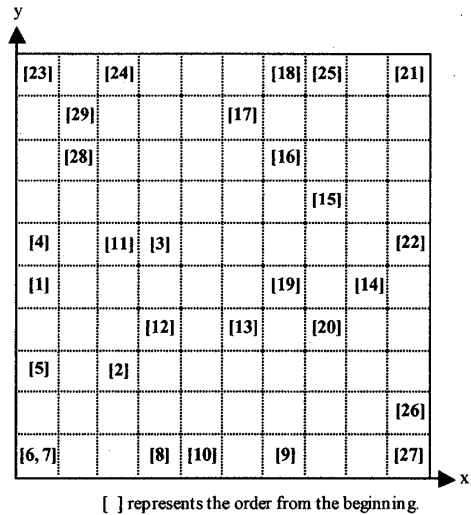


図4. マッピング例

4. まとめ

本報告では、Detrended Fluctuation Analysis (DFA) 及び自己組織化手法によりIPネットワークトラヒックの自己相似性の時間変動パターンを可視化した。今後は、ネットワーク品質とスケーリング指数の相関性評価、等を進めていく予定である。

(参考文献)

[1] C.-K. Peng, S.V. Buldrev, et al, "Mosaic organization of DNA nucleotides," Physical Review E, vol. 49, no.2, pp. 1685-1689, 1994.
 [2] M. Masugi, "Self-organizing map-based analysis of IP-network traffic in terms of time variation of self-similarity," IEICE Trans. Fundamentals, vol.E87-A, no.6, pp.1546-1554, Jun 2004.
 [3] T. Kohonen, "Self-organizing maps," Springer, Berlin, Heidelberg NY, 1997.