

トラヒックの長期依存性を考慮した **Passive** 計測手法の改善とその効果

小松 忠嗣[†] 川原 憲治^{††} 尾家 祐二^{††}

^{† †}九州工業大学 情報工学部 電子情報工学科

〒 820-8502 福岡県飯塚市川津 680-4

Phone:(0948)-29-7696, Fax:(0948)-29-7651

E-mail: [†]koma@nile.cse.kyutech.ac.jp, ^{††}{kawahara,oie}@cse.kyutech.ac.jp

あらまし 通信基盤としてのインターネットが普及して以来、トラヒック量は増加し、その種類も多様化している。そのため、送受信間の接続性の提供だけではなく、各トラヒックが要求するサービス品質 (Quality of Service, QoS) を満足する必要がある。そのためには、該当する送受信間のパケット損失や転送遅延、可用帯域等を計測する技術が重要になってきている。その一手法である **Passive** 計測では、測定対象経路に属するルータ/リンクの過去の情報をもとに、計測対象の現在値を推定する。代表的で簡易な推定モデルとして指数平滑移動平均 (Exponential Moving Average, EMA) があるが、ネットワークトラヒックの長期依存性の影響を反映した高精度な推定は困難だと考えられる。そこで、本研究では、トラヒックの長期依存性が指数平滑移動平均による転送トラヒック量推定に及ぼす影響をシミュレーションにより定量的に示し、推定精度能を改善するための手法を提案して、その効果を検討する。

キーワード **Passive** 計測, 長期依存性, 性能推定, 指数平滑移動平均

Improvement and Effectiveness of Passive Measurement Depending upon Long Range Dependency of Traffic

Tadashi KOMATSU[†], Kenji KAWAHARA^{††}, and Yuji OIE^{††}

^{† †}Dept. of Computer Science and Electronics, Kyushu Institute of Technology

Kawazu 680-4, Iizuka, 820-8502 Japan

Phone:(0948)-29-7696, Fax:(0948)-29-7651

E-mail: [†]koma@nile.cse.kyutech.ac.jp, ^{††}{kawahara,oie}@cse.kyutech.ac.jp

Abstract Since the Internet has been widespread as communication infrastructure, the amount of traffic becomes increasing rapidly and various types of traffic, such as voice, video, data and so on, have been transmitted. Therefore, it should satisfy Quality of Service (QoS) required by each traffic as well as its end-to-end connectivity, so that the measurement technology for inference of performance such as packet loss rate, transmission delay and available bandwidth becomes greatly important. In passive measurement, current performance can be estimated by some past and already obtained information on routers/links via the corresponding source-destination path. Exponential Moving Average (EMA) is well-known model of passive measurement and widely used because of its simplicity, however, without considering Long Range Dependency of network traffic in its parameter setting. Thus in this paper, we will quantitatively investigate the influence of LRD on EMA estimation of the amount of traffic on some router, and discuss some improvement of EMA and its effectiveness by computer simulation.

Key words Passive Measurement, Long Range Dependency, Performance Estimation, Exponential Moving Average(EMA)

1. はじめに

通信基盤としてのインターネットが普及して以来、トラヒック量は増加し、その種類も多様なものとなってきている。それにとともに、多様なアプリケーションのサービス品質 (Quality of Service, QoS) の保証が目ざされ、そのサービス品質を客観的に評価できる計測技術が重要となってきている。

計測技術の1つである Passive 計測は、計測者 (サーバ) が経路中の各ルータにおけるキュー長やリンク利用率等の過去の情報をもとに End-to-End の転送遅延などを現在の特性を推定する。このとき、ネットワークトラヒックがポアソン (Poisson) 性を有していれば、短期的な計測情報をもとに現在の性能を高精度に推定できる。しかし、現状の LAN (Local Area Network) や WAN (Wide Area Network) で観測されるトラヒックは、その変動に長期依存性 (Long Range Dependency, LRD) を示すことが報告されている [1]。その発生する要因としては、TCP (Transmission Control Protocol) のウィンドウフロー制御に起因する往復伝搬遅延時間や輻輳 [2][3]、CSMA/CD (Carrier Sense Multiple Access with Collision Detection) 方式におけるバックオフ [4] が挙げられる。トラヒック変動が LRD を示す場合、現在の性能を高精度で推定するためには、長期にわたる計測情報が必要となり、保持すべき計測情報量は LRD の強さ (Hurst パラメータ) に依存することになる。

ある計測量が長期的な視野で変動する場合の推定技術は、「テクニカル分析」として広い分野で盛んに研究されている [5]。その一推定モデルとして指数平滑移動平均 (Exponential Moving Average, EMA) があり、推定時の過去の計測情報の量や比重を単一の係数で調整する単純な方式であるため広く利用されている [6]。しかし、トラヒック変動が LRD を示すネットワーク環境においては、特性の計測/推定時間 (Aggregation Time, AT, 以降は集約時間と呼ぶ) によって長期依存性を示す局所的な Hurst パラメータの値が変動することが一般的であり、これまで、集約時間 AT が EMA の推定性能に及ぼす影響についてはあまり詳しく述べられていない。

そこで本研究では、まず、転送トラヒックが長期依存性を有するネットワーク環境をシミュレートして、ある集約時間におけるルータの転送トラヒック量を計測し、EMA によって推定する。このとき、実測値と推定値の差から得られるエラー率を推定精度を示す性能指標と定義して、集約時間の推定精度に及ぼす影響を示す。次に、Hurst パラメータが EMA の推定精度に及ぼす影響について調査し、EMA の係数を Hurst パラメータに依存して決定する推定方式の効果を示す。さらに EMA 推定方式の問題点を改善するための方式を提案し、その効果を確認する。

以下、2 章において集約時間と Hurst パラメータを定義し、3 章で各推定方式を説明する。4 章でシミュレーションモデルを説明し、5 章で推定結果を示す。そして 6 章でまとめる。

2. 長期依存性を示すパラメータの定義

トラヒックの長期依存性を考慮した Passive 計測を実現するにあたり、本章では長期依存性を示すパラメータを定義する。

2.1 集約時間 (Aggregation Time, AT)

ある単位時間をスロットと定義し、スロット t の計測トラヒック量を Y_t とする。さらに m スロットあたりの計測トラヒック量 $Y_t^{(m)}$ を次式で表す。

$$Y_t^{(m)} = \sum_{i=t-m+1}^t Y_i \quad (1)$$

ここで、 m を集約時間 (AT) と呼ぶ。

2.2 長期依存性 (Long Range Dependency, LRD) の度合

LRD は計測値に時間的な相関がある性質のことで、その相関係数は Hurst パラメータ H で表される。 H は、 $0.5 \leq H < 1$ の範囲をとり、 $H = 0.5$ では無相関を示し、 H が 1 に近づくにつれて「LRD の度合」が高くなることを示す [7]。この H を計測値から推定する簡単な手法として、Aggregate Variance 推定がある [8]。

まず、 $Y_t^{(m)}$ の標本分散 $Var(Y^{(m)})$ を以下の $\widehat{Var}(Y^{(m)})$ によって推定する。

$$\widehat{Var}(Y^{(m)}) = \frac{1}{N/m} \sum_{j=1}^{N/m} (Y_j^{(m)})^2 - \left(\frac{1}{N/m} \sum_{j=1}^{N/m} Y_j^{(m)} \right)^2 \quad (2)$$

ただし、 N は全サンプル数を指し、集約時間 m のときの全サンプル数を N/m で表す。このとき、 $Var(Y)$ と $Var(Y^{(m)})$ の関係は漸近的に $H^{(m)}$ を含む式

$$Var(Y^{(m)}) \sim \frac{Var(Y)}{m^{2(1-H^{(m)})}} \quad (3)$$

で示すことができる。ここで、 $H^{(m)}$ は集約時間 m における Hurst パラメータであることから、local Hurst 値と呼ばれている。式 (1)、(2) より、 $H^{(m)}$ を求めると

$$H^{(m)} = 1 + \frac{\log Var(Y^{(m)}) - \log Var(Y)}{2 \times \log m} \quad (4)$$

$$\triangleq 1 - \frac{\beta^{(m)}}{2}$$

となる。上式により、集約時間 m を横軸、標本分散 $Var(Y^{(m)})$ を縦軸にプロットし、両対数表示した場合の傾きで $H^{(m)}$ を求めることができる。したがってその減少率が小さければ集約時間 m によらず、LRD が強いことを示す。

3. Passive 計測における推定モデル

本章ではこれまでに計測されているトラヒック量系列 $\{Y_0^{(m)}, \dots, Y_{t-1}^{(m)}\}$ からスロット t におけるトラヒック量 $Y_t^{(m)}$ の推定を行うモデルについて示す。このとき、推定トラヒック量を $\hat{Y}_t^{(m)}$ と定義する。本章ではまず、EMA について示す。次に Hurst 値を考慮した EMA の改善手法について提案する。

3.1 指数平滑移動平均 (Exponential Moving Average, EMA)

計測トラヒック量系列 $\{Y_0^{(m)}, \dots, Y_{t-1}^{(m)}\}$ から、 $\hat{Y}_t^{(m)}$ を推定する単純な方法に移動平均 (Moving Average, MA) がある。この方式では過去に得られた計測トラヒック量の加重平均により $\hat{Y}_t^{(m)}$ を求めるが、その重みを指数的に設定する方式を指数平滑移動平均 (Exponential MA, EMA) と呼び、次式で表すことができる。

$$\hat{Y}_t^{(m)} = (1 - \alpha) \hat{Y}_{t-1}^{(m)} + \alpha Y_{t-1}^{(m)} \quad (5)$$

ここで、 $\alpha (0 \leq \alpha \leq 1)$ は平滑化係数である。式 (5) より EMA では前回の推定値と計測値の 2 情報から推定が可能となる。

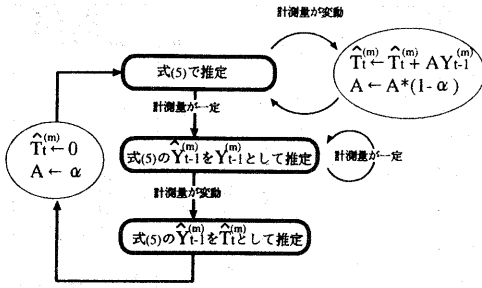


図1 EMA リセット方式のアルゴリズム

3.2 動的平滑化係数 $\alpha_t^{(m)}$ を用いた EMA(提案方式 1)

前節で示した EMA では、指数平滑係数 α は計測時刻によらず固定である。しかし、長期依存性を示すトラフィックを計測対象とする場合、その値を変更することでより高精度で推定できると考えられる。そこで、式 (5) を以下のように変更する。

$$\hat{Y}_t^{(m)} = (1 - \alpha_t^{(m)})\hat{Y}_{t-1}^{(m)} + \alpha_t^{(m)}Y_{t-1}^{(m)} \quad (6)$$

ここで、 $\alpha_t^{(m)} = f(H_t^{(m)})$ とし、計測/推定集約スロット数 m で、時刻 t の時の local Hurst パラメータ $H_t^{(m)}$ の関数で定義する (本稿で扱う詳細な関数形については後述する)。なお、 $H_t^{(m)}$ は、式 (4) の定義より、集約時間 m で時刻 t の標本分散 $\hat{V}_t^{(m)}$ を EMA を用いて式 (7) より推定し、集約時間 $m = 1$ の場合の $\hat{V}_t^{(1)}$ との関係式 (8) から導出する。

$$\hat{V}_t^{(m)} = (1 - \alpha_{t-1}^{(m)})\hat{V}_{t-1}^{(m)} + \alpha_{t-1}^{(m)}(Y_{t-1}^{(m)} - \hat{Y}_{t-1}^{(m)})^2, \quad (7)$$

$$H_t^{(m)} = 1 + \frac{\log \hat{V}_t^{(m)} - \log \hat{V}_t^{(1)}}{2 \times \log m}. \quad (8)$$

3.3 EMA リセット方式 (提案方式 2)

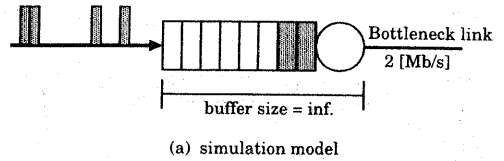
EMA、および動的平滑化係数を用いた EMA とともに、現在値の推定において「直近の過去」の計測量の比重が高い。このことは、しばらく計測量に変動がなく、ある時を境に急激に変動する時点の推定誤差を非常に大きくしてしまう可能性がある。そこで、通常は式 (5) に従い推定するが、計測量に変化しなければ、直ちに式 (5) の $\hat{Y}_{t-1}^{(m)} = Y_{t-1}^{(m)}$ として推定する。その後、計測量に変化する点 (これをトリガ点と呼ぶ) の推定には、 $\hat{Y}_{t-1}^{(m)}$ を次式で得られる $\hat{T}_t^{(m)}$ に置き換えて式 (5) を適用する方式を提案し、そのアルゴリズムを図 1 に示す。以降、EMA リセット方式と呼ぶ。

$$\hat{T}_t^{(m)} = \alpha Y_k^{(m)} + \alpha(1 - \alpha)Y_{k+1}^{(m)} + \alpha(1 - \alpha)^2 Y_{k+2}^{(m)} + \dots \quad (9)$$

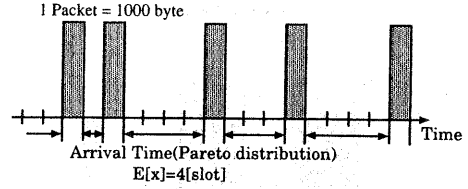
ここで、 k は前回のトリガ点を示す。すなわち、 $\hat{T}_t^{(m)}$ は、トラフィック量の変動が生じた時点からの計測量に指数的な重みを付けて加重した加重移動平均 (Weighted MA, WMA) 値を示し、この値を式 (5) の $\hat{Y}_{t-1}^{(m)}$ と置き換えることによって、過去の変動度合を反映させて推定精度を向上できることが期待される。また、変動しないときは直前の計測値をそのまま推定値とすることにより平均エラー率の改善が可能であると考えられる。このときの WMA の重み α は EMA の平滑化係数 α と同じにする。

4. シミュレーションモデル

計測対象ネットワークとしてバッファサイズが ∞ のルータ



(a) simulation model



(b) traffic model

図2 パケット到着モデル

一段を仮定し、流入トラフィックが LRD を有する場合のトラフィック量を推定する。トラフィック到着モデルとして、パケット到着モデルとフロー到着モデルの二つを考慮する。

4.1 パケット到着モデル

まず、図 2(a) に示すようなシミュレーションモデルを説明する。ルータの出力リンクの帯域は 2[Mb/s] とする。到着する IP パケットの長さは 1000[byte] とし、1 パケットの転送時間を 1 単位スロットと定義する。一回のシミュレーション時間は $N = 18 \times 10^6$ スロットとする。

ルータには、図 2(b) に示すようなパレート分布にしたがってパケットが到着するものとする。その時、パケットの到着間隔の確率分布関数 $F(x)$ は

$$F(x) = 1 - \left(\frac{k}{x}\right)^a \quad (10)$$

で表される。ここで、 $a(1 < a < 2)$ は shape パラメータ、 $k(k \leq x)$ は location パラメータと呼ばれている。また、平均到着間隔は $E[x] = ak/(a-1)$ 、Hurst パラメータは $H = (3-a)/2$ で定義されるため、shape パラメータ a を設定することで一意に $H^{(m)}$ 値を決定することができる。なお、本シミュレーションでは $E[x] = 4[\text{slot}]$ 、すなわち、平均到着トラフィック量は 0.4[Mb/s] とする。

4.2 フロー到着モデル

図 3(a) に示すように、20 の source-destination pair が一つのルータで多重されるモデルをフロー到着モデルと定義する。ルータのバッファサイズやリンク容量、パケット長についてはパケット到着モデルと同様である。

各ソースから発生するトラフィックは、図 3(b) に示すように、パケット転送区間とアイドル区間が確率的に交互に生じる on-off モデルに従うものとする。1 パケット転送区間を 1 フローと定義し、16 個のパケットが連続して送信されるものとする。フローの発生間隔は shape パラメータ $a = 1.5$ のパレート分布にしたがうものとし、ルータへの平均到着トラフィック量が 0.4[Mb/s] となるように調整する。

4.3 性能指標: エラー率

図 2(a)、図 3(a) において、集約時間 m あたりのルータの転送トラフィック量の計測値を $Y_t^{(m)}$ 、推定値を $\hat{Y}_t^{(m)}$ とした場合、推定精度を示す指標として、以下の平均エラー率 $\overline{Err}^{(m)}$ と最大エラー率 $Err_{\max}^{(m)}$ を定義する。

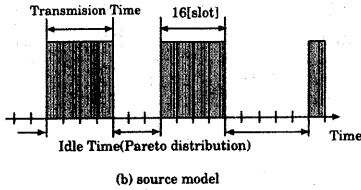
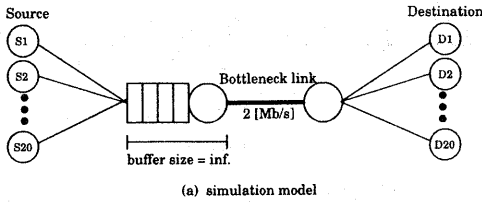


図3 フロー到着モデル

$$\overline{Err}^{(m)} = \frac{1}{N/m} \sum_{i=1}^{N/m} \frac{|Y_i^{(m)} - \hat{Y}_i^{(m)}|}{m} \quad (11)$$

$$Err_{\max}^{(m)} = \max \left[\frac{|Y_i^{(m)} - \hat{Y}_i^{(m)}|}{m} \right], \quad i = 1, \dots, N/m \quad (12)$$

5. シミュレーション結果

トラヒックの長期依存性がEMAの推定性能に及ぼす影響を調査するために、前章で説明した、パケット到着モデルとフロー到着モデルにおいて、計算機シミュレーションにより推定精度を評価する。

5.1 パケット到着モデル

まず、パケット到着モデルを仮定し、EMAの推定精度について評価し、3.2, 3.3節で提案した方式の改善度合について調査する。

5.1.1 EMA

トラヒック量 $Y_t^{(m)}$ をEMAによって推定した $\hat{Y}_t^{(m)}$ の精度について評価するために、集約時間 m と local Hurst パラメータ $H^{(m)}$ の影響について示し、良好な特性を得るための平滑化係数 α について調査する。

a) 集約時間 m の影響

図4に集約時間 m に対するEMAの推定性能を示す。図中の Avg Err, Max Err はそれぞれ平均エラー率 $\overline{Err}^{(m)}$ 、最大エラー率 $Err_{\max}^{(m)}$ とする。このときの local Hurst パラメータを $H^{(m)} = 0.60$ と仮定する。まず、この図4より、集約時間 m が大きくなると平均、最大エラー率ともに小さくなることから、これは、 m スロットあたりでトラヒック量を計測/推定することで、推定値の分散が小さくなるためである。

次に、EMAの平滑化係数 α が推定誤差に及ぼす影響について考察する。図4の平均エラー率について着目すると、 α が小さいと改善されることがわかるが、その差はごくわずかである。また、最大エラー率についてみてみると、集約時間 m が10を越えると、同様に α を小さく設定した方が改善されることがわかる。図中の“Best Performance”は、最大エラー率を最も小さくする時の α に設定（その時の値は図中の () 内に記載）した場合の特性をプロットしたもので、これから、全般にわたつ

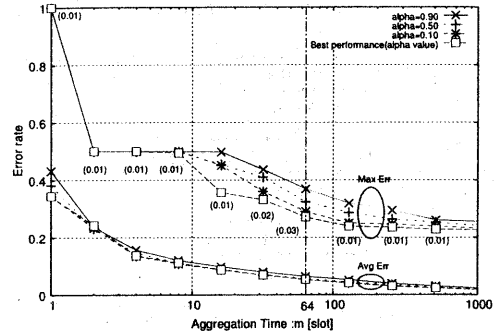


図4 集約時間 m に対する推定誤差 (平滑化係数 α による影響)

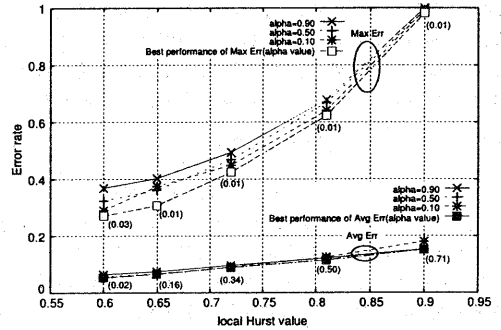


図5 $H^{(m=64)}$ に対する推定誤差 (平滑化係数 α による影響)

て $\alpha \sim 0.01$ と設定すればよいことがわかる。先に述べたように、集約数 m の増加により、推定値の分散は小さくなる傾向にあるが、トラヒックに長期依存性がある場合、その減少度合は小さい。そのような場合にEMAを用いてトラヒック量の推定を行う場合には、 α を大、すなわち、直近の過去の計測量に重きをおくよりも、 α を小、すなわち、より古い計測量の比重を高くする必要があることを示している。

b) local Hurst パラメータ $H^{(m)}$ の影響

前節では、local Hurst パラメータを $H^{(m)} = 0.60$ とした場合、平滑化係数 α をなるべく小さくすることで、平均/最大エラー率を小さくできることを示したが、本節では、 $H^{(m)}$ の値をかえて調査する。図5に $H^{(m)}$ を横軸とした場合の推定誤差について示す。ここで、集約時間 $m = 64$ とする。

図より、まず、 $H^{(m=64)}$ の増加にともない、平均、最大エラー率ともに増加することがわかる。次に、EMAの平滑化係数 α が推定誤差に及ぼす影響について考察する。図中の“Best performance of Avg Err”は、平均エラー率を最小にする α に設定（その時の値は図中の () 内に記載）した場合の平均エラー率特性、“Best performance of Max Err”は、最大エラー率を最小にする α に設定（その時の値は図中の () 内に記載）した場合の最大エラー率特性を示す。この図より、最大エラー率を改善するためには α は小さくしてよいが、平均エラー率を改善するためには、 $H^{(m=64)}$ の増加にあわせて α を大きくする必要がある。その理由は以下のように考えられる。 $H^{(m=64)}$ の増加、すなわち、長期依存性の度合いが高い場合、トラヒック到着のバースト性が非常に高くなるため、平均的に誤差を小さくするためには、 α を大きく、すなわち、直近の過去の計測量の比重を高くして

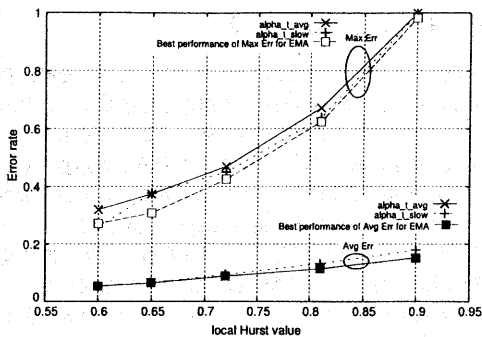


図6 動的平滑化係数 $\alpha_t^{(m)}$ の変化方式における影響

推定する必要がある。

以上のことから、ネットワークのトラフィック特性が長期依存性を有する場合、EMA 推定においては、その度合を示す Hurst パラメータの値に応じて平滑化係数を変更する必要があることがわかる。

5.1.2 動的平滑化係数 $\alpha_t^{(m)}$ を用いた EMA(提案方式 1)

これらの調査結果から、トラフィック推定に EMA を用いた場合、集約時間 m より $H^{(m)}$ による影響が大きいことがわかる。したがって、以降では $H^{(m)}$ についての議論する。

前節の結果より、EMA の平滑化係数を Hurst パラメータに応じて動的に変化する方式(提案方式 1)の有効性について調査する。ここで、集約時間は $m = 64$ とする。提案方式 1 において、式(7)、(8)より求めた集約時間を $m = 64$ とした場合のスロット t における local Hurst パラメータ $H_t^{(m=64)}$ の関数として平滑化係数 $\alpha_t^{(m)}$ を以下のように定義する。

$$\alpha_{avg}^{(m)} = 2.25 \times H_t^{(m=64)} - 1.33 \quad (13)$$

$$\alpha_{slow}^{(m)} = (H_t^{(m=64)} - 0.6)^2 \quad (14)$$

ここで、 $\alpha_t^{(m)}$ は前節において平均エラー率を最小にする α と $H_t^{(m=64)}$ の関係を一次式 $\alpha = a \times H^{(m)} + b$ で近似し、その定数 a, b を最小自乗法により求めた場合の回帰式である。また、式(14)は式(13)より $H^{(m=64)}$ の増加に対して緩やかに増加するような関数で性能比較のためにとりあげる。これらの平滑化係数の動的変化の効果を確認するために、図6に $H^{(m=64)}$ を横軸とした場合の推定誤差について示す。なお、5.1.1 b) で求めた平均/最大エラー率の最小値についても示す。図より式(13)を用いた方式では、original の EMA で得られる平均エラー率の最小値とほぼ同じ特性を得ることができるが、 $H^{(m=64)}$ が小さい時の最大エラー率特性が劣化してしまう。一方、式(14)を適用して平滑化係数を変化する場合、式(13)を用いた場合と比べて平均エラー率は同等の性能を達成しつつ、最大エラー率を若干改善することが可能であることがわかる。

以上より、local Hurst パラメータに応じた平滑化係数を動的に変化する方式により original EMA で得られる最良の特性を得ることができるが、さらなる改善には至らないと言える。

5.1.3 EMA リセット方式(提案方式 2)

前節で EMA の平滑化係数 α を時々刻々と変化する local Hurst パラメータに応じて変化する方式を提案し original の EMA との性能比較を行ったが、以下の点でエラー率の改善に限界があるものと考えられる。EMA はその定義(式(5))より、 α の値によらず、直近の過去に計測された値の比重が高い。したがって、

表1 EMA と EMA リセット方式との比較

Hurst 値	Avg Err
	EMA : EMA Reset
0.60	0.054 : 0.054 (± 0.000)
0.65	0.066 : 0.065 (-0.001)
0.72	0.089 : 0.087 (-0.002)
0.81	0.115 : 0.113 (-0.002)
0.90	0.151 : 0.149 (-0.002)

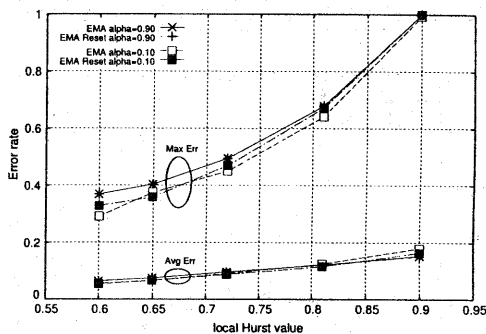


図7 EMA リセット方式と EMA の性能比較

例えば、計測量がしばらく変化せず、ある時点で急激に変化した場合、その影響を推定に反映させることができないため、その場合の誤差は大きくなる。そこで、計測量に急激な変化が生じた場合、まずその変化に敏感に反応するために、 Y_{t-1} を $\hat{Y}_t$ とする。その後計測量に変化が生じた場合、式(5)における $Y_{t-1}^{(m)}$ の値をリセットして、式(9)で表すような過去に変動が生じた区間の計測量を加重したものと置き換える EMA(EMA リセット方式)によって、誤差改善が可能か検討する。なお、本方式による推定は式(9)によってトラフィック変動の長期依存性を反映しているものと考えられる。

平滑化係数を $\alpha = 0.10, 0.90$ と設定した場合の EMA、および EMA リセット方式のエラー率 - local Hurst パラメータ特性を図7に示す。図より、平均エラー率特性は α の値によらず、両方式とも同程度の性能が得られており、最大エラー率特性においては、local Hurst パラメータ $H^{(m=64)} = 0.65$ 、 $\alpha = 0.1$ の時に EMA リセット方式が若干優れていることがわかる。図からは、平均エラー率特性の詳細な比較が困難なため、最適な α に設定した場合の EMA と EMA リセット方式のエラー率を表1に示す。この表から、平均エラー率は、local Hurst パラメータが大きい、すなわち、長期依存性の度合いが強いほど、EMA リセット方式による改善率が若干大きくなることがわかる。

5.2 フロー到着モデルにおける提案方式の効果

5.1 節では、長期依存性の度合いが original の EMA や提案方式で推定したトラフィック量の精度に及ぼす影響を調査するために、シミュレーションのトラフィックモデルとして Hurst パラメータを一意に決定できるパレート分布を仮定した。ここでは、4.2 節で定義したように現実的なトラフィックモデル(フロー到着モデル)を扱った場合の EMA と EMA リセット方式の性能を比較するために、集約時間 m に対するエラー率特性を図8に示す。ここで、各方式の平滑化係数を $\alpha = 0.10, 0.90$ と設定し、1ソー

表2 各方式における最適 α と平均エラー率

m	$H^{(m)}$	Avg Err		平滑化係数 α	
		EMA : EMA Reset		EMA : EMA Reset	
2	0.98	0.041 : 0.041(± 0.000)		0.99 : 0.99	
8	0.96	0.157 : 0.157(± 0.000)		0.99 : 0.99	
16	0.93	0.230 : 0.224(-0.006)		0.99 : 0.01	
32	0.88	0.235 : 0.221(-0.014)		0.01 : 0.01	
64	0.84	0.167 : 0.167(± 0.000)		0.01 : 0.01	
128	0.79	0.120 : 0.120(± 0.000)		0.02 : 0.01	
512	0.71	0.050 : 0.050(± 0.000)		0.08 : 0.09	
1024	0.69	0.036 : 0.036(± 0.000)		0.18 : 0.19	

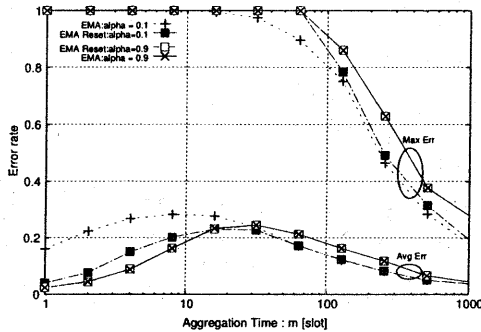


図8 EMA リセット方式と EMA の性能比較 (フロー長=16 パケット)

スから連続して転送されるパケット数 (フロー長) は 16 パケット (スロット) とする。

図8より、集約時間 m の増加につれて、最大エラー率は小さくなるが、平均エラー率は、 m がフロー長 (16) と同程度に設定された場合が最も大きくなっていることがわかる。これは、集約時間 m スロット中に全くパケットが到着しない場合と、全ソースのフロー到着が集中する場合とが存在し、推定レンジが非常に大きくなるため、平均的な誤差が大きくなるものと考えられる。そして、original の EMA の平滑化係数 α は、この m の値を境として変更する必要があることがわかる。すなわち、 $m <$ フロー長の場合 α は大きく、 $m >$ フロー長の場合 α は小さく設定しなければならない。一方、EMA リセット方式では、 $\alpha = 0.1$ とした場合、 $m <$ フロー長の範囲で、大きく平均エラー率を改善できることがわかる。また、 $m >$ フロー長の範囲でも、最大エラー率は若干劣化するが、平均エラー率は EMA 方式と同等の性能が得られる。特記すべき点は、EMA の場合と異なり、 m の値によらず、平滑化係数 α を固定 (ここでは、 $\alpha = 0.1$) としても十分に良好な特性が得られることである。

次に、両方式における最適平滑化係数の値と、その時の平均エラー率を表2に示し、性能を比較する。この表より EMA リセット方式においても、EMA の場合と同様に、Hurst パラメータが大きくなる (集約時間 m が小さい) 程、平均エラー率を最小にする平滑化係数の値は大きく設定する必要がある。さらに、EMA リセット方式は、集約数 m がフロー長程度に設定された場合に、EMA の性能を改善することができるがわかる。

6. まとめと今後の課題

現状のインターネットにおいては、トラヒック変動に長期依存性 (Long Range Dependency, LRD) が存在することが示されている。そのため、トラヒック制御で重要となるネットワーク性能計測/推定においては、そのことを考慮する必要がある。本研究では、Passive 計測の一手法で広く利用されている指数平滑移動平均 (Exponential Moving Average, EMA) によりトラヒック量を推定する場合を仮定し、長期依存性が推定精度に及ぼす影響を詳細に調査し、精度改善方法について検討した。

通常の EMA により推定を行う場合、そのパラメータである平滑化係数 α の値を長期依存性の度合を示す Hurst パラメータに応じて変更する必要があることを示した。そこで、推定トラヒック量の分散を EMA で求め、それから得られる local Hurst パラメータに応じて、動的に平滑化係数 α を変更する方式を提案し、その精度を調査したが、通常の EMA の推定精度を改善するには至らなかった。そこで、その改善のために、過去のトラヒック変動度合を EMA の推定に反映する EMA リセット方式を提案し、その精度を確認したところ、Hurst パラメータが大きい場合や計測時間がフロー長と同程度の場合に、平均エラー率が改善できることが示された。

以上のように、提案した EMA リセット方式の有効性は確認できたが、その改善率はわずかなもので、ネットワーク環境が異なる場合、例えば、計測対象ルータに TCP トラヒックが到着する場合や、その数等が精度に及ぼす影響について調査する必要がある。また、EMA リセット方式において、過去のトラヒック変動度合を Hurst パラメータに応じて調整する方法で更なる改善が見込まれるが、これらの調査については今後の課題である。

謝 辞

本研究の一部は、総務省における「戦略的情報通信研究開発推進制度」、および、通信・放送機構 (TAO) における「JGN プロジェクト」(JGN-R14101) の支援を受けている。ここに記して謝意を表す。

文 献

- [1] W. Leland, M. Taqqu, W. Willinger, and D. Wilson, "On the Self-Similar Nature of Ethernet Traffic," Trans. on Networking, Vol.2, No.1, pp.1-15, Feb, 1994
- [2] H. Furuya, T. Hill, and H. Nakamura, "Influence of transmission delay on self-similar scaling behavior of aggregated TCP/IP traffic," IEICE Technical Report, SSE2000-9, Sept., 2000.
- [3] H. Furuya, T. Hill, and H. Nakamura, "Influence of network congestion on self-similar scaling behavior of aggregated TCP/IP traffic," IEICE Technical Report, IN2000-36, June, 2000.
- [4] 竹島 由晃, 福田 健介, 高安 秀樹, 高安 美佐子, "コンピュータシミュレーションによる CSMA/CD アクセス方式ネットワークトラヒックの統計物理学的解析," 電子情報通信学会論文誌, VOL.J84-B, No.5, May, 2001.
- [5] 山本 拓, "時系列分析とその経済分析への応用" 大蔵省財政金融研究所「ファイナンス・レビュー」, 第23号, June, 1992.
- [6] G. Bianchi, V. Mancuso, and G. Neglia, "Is Admission-Controlled Traffic Self-Similar?," NETWORKING, pp.327-339, 2002.
- [7] K. Park, and W. Willinger, Self-Similar Network Traffic And Performance Evaluation, Wiley Inter-Science, pp.17-27, 2000.
- [8] M.S.Taqqu, V.Teverovsky, and W.Willinger, "Estimators for long-range dependence: an empirical study," Fractals, Vol 3, No. 4, pp.785-788, 1995.