情報処理学会論文誌 Vol.54 No.6 1761–1770 (June 2013)

推薦論文

R/S Poxレッグライン特性

高橋 秋典[1,a)] 五十嵐 隆治[1] 上田 浩[2] 岩谷 幸雄[3] 木下 哲男[4]

受付日 2012年10月2日, 採録日 2013年3月1日

概要：本論文では，トラフィック時系列に現れる非定常的特徴を定量化するR/S Poxレッグライン特性，ならびにこの特性を用いた周期的時系列に対する周期推定法を提案する．本特性は，R/S解析法によるR/S Pox Diagramにおいて，時系列の周期性を表す特徴点を境に折れ曲がるプロット形状を明確に示すものである．この特徴点を定量化し周期を推定できれば，周期的特徴を持つ長期的ポートスキャンなどの低レート攻撃を検知することができると推測される．本手法の有効性は，シミュレーション時系列および実環境トラフィック時系列に対する周期推定により示す．

キーワード：R/S解析法，R/S Pox Diagram，トラフィック変化検知，周期推定

R/S Pox Leg-line Characteristics

Akinori Takahashi[1,a)] Ryuji Igarashi[1] Hiroshi Ueda[2] Yukio Iwaya[3] Tetsuo Kinoshita[4]

Received: October 2, 2012, Accepted: March 1, 2013

Abstract: This paper proposes an R/S Pox Leg-Line characteristics to quantify a non-stationary of the Internet traffic time series, and a method of period estimation using the characteristics. To sum up the major feature of the proposal characteristics, the dispersive configuration of an R/S Pox Diagram bends after a characteristic point expressing a period of a traffic time series. If the period is estimated by quantifying the characteristic point, it is thought that a low-rate attack having a cycle such as a long-term port scan is detected. The confirmation of the proposal is practiced by using a simulation and a real traffic data.

Keywords: R/S analysis, R/S Pox Diagram, Network Traffic detection, period estimation

1. はじめに

インターネットのパケットトラフィック時系列は従来モデルであるポアソン過程には従わず，時系列の自己相似性に起因する長期記憶過程であることが発見されてから [1]，その自己相似性の要因を調査する研究が行われてきた．その要因には，上位層プロトコルのTCP輻輳制御 [2], [3] や輻輳・非輻輳の臨界領域の影響 [4]，また，ネットワークアプリケーションによる影響 [5] や，DDoS攻撃のような非定常的な異常トラフィックによる影響 [6] などの報告があり，トラフィック事象変化に対して自己相似性の様相が変化するということが観測されている．

この自己相似性を表すハーストパラメータ H の導出法の1つであるR/S解析法は，Hurstがナイル川の流量変動の統計的解析に導入した後 [7]，Mandelbrotらにより数学的な基礎付けがなされた統計的解析法 [8] で，Lelandら [1] が初めて，ネットワークトラフィックの自己相似性の解析に導入したものである．この解析法は“グラフ的な方法”と呼ばれるもので，ハーストパラメータ H は観測時系列

1 秋田大学大学院工学資源学研究科
Graduate School of Engineering and Resource Science, Akita University, Akita 010–8502, Japan
2 京都大学学術情報メディアセンター
Academic Center for Computing and Media Studies, Kyoto University, Kyoto 606–8501, Japan
3 東北学院大学工学部
Faculty of Engineering, Tohoku Gakuin University, Tagajo, Miyagi 985–8537, Japan
4 東北大学大学院情報科学研究科
Graduate School of Information Sciences, Tohoku University, Sendai, Miyagi 980–8578, Japan
a) akinori@ie.akita-u.ac.jp

本論文の内容は2012年9月のFIT2012第11回情報科学技術フォーラムにて報告され，同プログラム委員長により情報処理学会論文誌ジャーナルへの掲載が推薦された論文である．

データにおける任意長区間から算出されるR/S統計量をグラフにしたR/S Pox Diagramのプロット点群の傾きより導出される．このR/S解析法に対しても様々な検討が行われており，R/S Pox DiagramのハーストパラメータH計算範囲の妥当性[9], [10]や計算高速化のためのアルゴリズム改良[11]などが提案されている．

これらの研究においては，R/S解析法から導出されるハーストパラメータHに着目しトラフィック事象に対する検討が行われているが，非定常的に突発的トラフィック量増加を示すレベルシフトが発生する場合や，間欠的に到着し，かつ到着期間内では周期列となるようなトラフィック時系列の場合，R/S Pox Diagramに特徴的なプロット形状が現れることが報告されている[12], [13], [14]．著者らの調査においては，特に周期的時系列に対しては，プロット点群が1つの傾きではなく，途中から折れ曲がり，2つの傾きを呈することが観測されている．

これらのプロット形状は，非定常性を表す特徴であると考えられるが，信頼性に欠けるハーストパラメータH推定を与えてしまうため，統計解析においては除去すべき対象として検討されてきた．Mandelbrotらがコンピュータシミュレーションを用いた研究で，ランダム過程に決定論的過程，具体的にはsin波を重畳したとき，そのR/S Pox Diagram形状に変化が現れることを指摘していた[8]．しかし，非定常性が重畳されたときのR/S Pox Diagramの形状変化を事象変化検出に積極的に利用しようとした研究は見当たらない．著者らは，その非定常性に対するプロット形状の定量化は事象変化の検知に積極的に応用できると考えている．また，R/S解析法の計算過程において，R/S統計量を求める任意長区間のサイズによって未計算部分が存在し，直近の取得データが解析に反映されないことがある．これに対して問題視する研究も見当たらない．

本研究では，まずR/S解析法の未計算区間により時間特性に対する即応性が低下する問題点を指摘し，その改良法を提案する．さらに，その改良法を用いたR/S Pox Diagramに現れる特徴的なプロット形状の要因について明らかにし，その形状を定量的に扱えるR/S Pox レッグライン特性を提案する．この名称は，周期的時系列に対して折れ曲がるようなプロット点群の形状が人間の脚部に似ていることに由来する．この特性を用いて周期的時系列に対する周期推定法を検討し，シミュレーション時系列に対する性能評価を行った．また，その際，実環境から観測された長期的ポートスキャン攻撃トラフィック[15]に対して適用を試み，そのパケット到着間隔を推定することで本手法の有効性を示した．

本研究の構成を以下に示す．2章では，R/S解析法を用いたR/S Pox Diagramの導出ならびに未計算区間に対する改良について述べる．3章では，周期性を有する時系列に対するR/S Pox Diagramの特徴について述べ，4章ではその特徴を定量的に扱うR/S Pox レッグライン特性を提案する．5章では，シミュレーションならびに実トラフィック時系列データに対する適用を試み，6章でまとめる．

2. R/S解析法

2.1 導出手順

R/S解析法[1]によるR/S Pox Diagramの導出手順を**図1**に示す．

まず，対象とする観測データから解析対象とする時系列データ$\{X_t, t = 1, 2, \cdots, N\}$を得る．パケットトラフィックにおいては，時間軸上での点過程であるトラフィックを，計測単位時間Δtごとの到着パケットを計数して得ることになる．この時系列$\{X_t\}$に対して，図1に示すような区間長$n_m, m = 1, 2, \cdots, M$となる区間を定める．n_mは$5 \leq n_m \leq N$で，mは$mn_m \leq N$となるようにした区間数である．たとえば，$N = 3000$のとき，

- $n_m = 5$では，$m = N/n_m = 3000/5 = 600$
- $n_m = N/2 = 1500$では，$m = N/n_m = 3000/1500 = 2$
- $n_m = N = 3000$では，$m = N/n_m = 3000/3000 = 1$

などとなる．ただし，以後は記号の単純化のために，任意長区間n_mをnと記す．

いま，時系列長N内で互いに重複せず，区間長がnとなるようなm個の区間を考える．このm個の各区間（各々は区間長がnである）において，以下のような統計量を求める．まず，$1 \leq k \leq n$なるkに対し，式(1)より当該区

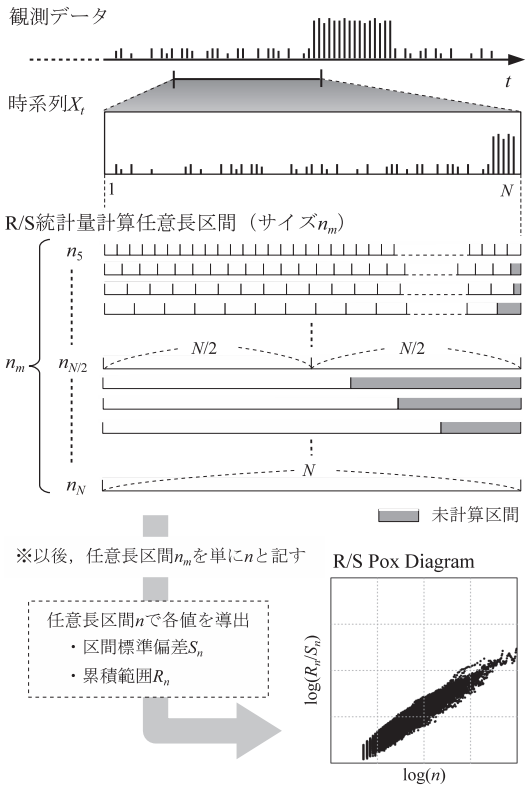


図1 R/S解析法
Fig. 1 Method of R/S analysis.

間の区間平均 $\bar{X_n}$ を求める．

$$\bar{X_n} = \sum_{k=1}^{n} X_k/n \tag{1}$$

同一区間内での累積和と線形な傾向 $k\bar{X_n}$ との差を表す値 W_k を式 (2) より求める．

$$W_k = \sum_{j=1}^{k} X_j - k\bar{X_n} \tag{2}$$

さらに，式 (3) より，この W_k から累積範囲 R_n を求める．ここで，$\max\{0, W_1, \cdots, W_k, \cdots, W_n\}$ は W_k の最大値，$\min\{0, W_1, \cdots, W_k, \cdots, W_n\}$ は W_k の最小値である．

$$\begin{aligned} R_n = &\max\{0, W_1, \cdots, W_k, \cdots, W_n\} \\ &- \min\{0, W_1, \cdots, W_k, \cdots, W_n\} \end{aligned} \tag{3}$$

この累積範囲 R_n と式 (4) から求められる当該区間標準偏差 S_n の比を用いて，R/S 統計量を導出する．

$$S_n = \sqrt{\sum_{k=1}^{n} X_k^2/n - \bar{X_n}^2} \tag{4}$$

ここから，任意長区間 n を Δn ずつ増加させながらそれぞれの R/S 統計量を導出し，式 (5) に基づいてハーストパラメータ H を推定する．

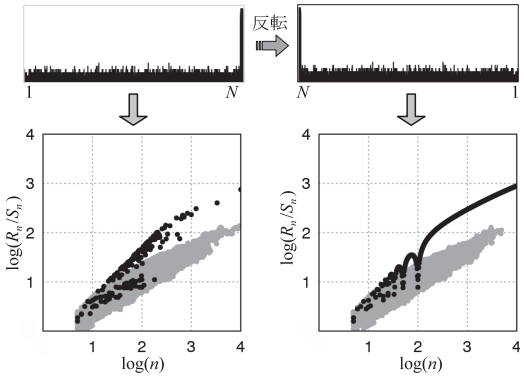
$$\log(R_n/S_n) = H\log(n) + \log c \tag{5}$$

具体的には $\log(R_n/S_n)$ を被説明変数，$\log(n)$ を説明変数，$\log c$ を定数とした回帰モデルと想定し，最小 2 乗法によりハーストパラメータ H を推定する．これをグラフ的に表すと，横軸 $\log(n)$，縦軸 $\log(R_n/S_n)$ のグラフにプロットされた点群の傾きを求めることになる．このグラフが R/S Pox Diagram である．

2.2 R/S 解析法の未計算区間

R/S 解析法は，サイズ N の時系列 X_t に対して任意長区間 n を定め R/S 統計量を導出する．しかし，図 1 に示すように，N が n で割り切れる場合，X_t 全区間データに対して計算が行われるが，割り切れない場合は X_t の時間軸上後半の部分に計算されないデータが存在する．これは n が大きくなるに従い，未計算区間も大きくなることになる．

通常，時系列解析において観測時系列の時間軸は時刻 t が小さい方向が“過去”を表し，大きい方向が“直近”を表している．つまり，未計算区間にはより直近のデータが存在することになり，解析に反映されにくくなる．観測対象によっては即応性が重視されることもあり，直近に発生したデータが解析に反映されないという事実は，特徴量の経時変化特性の検討上，問題となる場合がある．



(a) 時系列X_t　　(b) 反転時系列X_t'

図 2　反転時系列の効果

Fig. 2　R/S Pox Diagram characteristic due to reversed treatment of time series.

2.3 反転時系列による改良

この問題点を解決するために，より“直近”に近いデータを X_t の前半に，“過去”のデータを後半に反転させた時系列 X_t' を生成する．これは，式 (7) に示すように，時系列データを単純に反転させたものである．

$$X_t = \{x_1, x_2, \cdots, x_N\} \tag{6}$$

$$X_t' = \{x_N, \cdots, x_2, x_1\} \tag{7}$$

この改良法では，直近の取得データが時系列 X_t の前半 $N/2$ までに存在しているならば，どの任意長区間 n においても解析に反映されることになる．

2.4 改良法の効果

改良法の効果を示すため，突発的にトラフィックデータが増加した（レベルシフト）シミュレーション時系列を用いて，従来型の時系列 X_t と改良型の反転時系列 X_t' の 2 つから得られる R/S Pox Diagram を比較した．用いたシミュレーション時系列および導出された Pox Diagram をそれぞれ**図 2** (a)，図 2 (b) に示す．シミュレーション時系列 X_t は，分散 1.0，平均 3.0，ハーストパラメータ $H = 0.5$ と設定したサイズ $N = 10000$ の FGN（Fractional Gaussian Noize）に対し，区間 [9901-10000] に振幅 $P = 10$ のレベルシフト系列を重畳させたものである．これは，観測対象である時系列に対して，非定常的トラフィック時系列が早い段階で観測されたことを想定したものである．R/S Pox Diagram の灰色のプロット点は全区間におけるプロット点を表し，黒のプロット点はレベルシフトの変化点 $t = 9901$ が含まれる任意長区間から導出されたものを表す．

図 2 (a) から，X_t に対する R/S Pox Diagram では，レベルシフトの変化点が含まれるプロット点がまばらに観測され，特に $\log(n)$ の後半部分では未計算区間が多く存在していることが分かる．しかし，図 2 (b) より，反転時系列 X_t' ではすべての任意長区間 n に対してレベルシフトの変

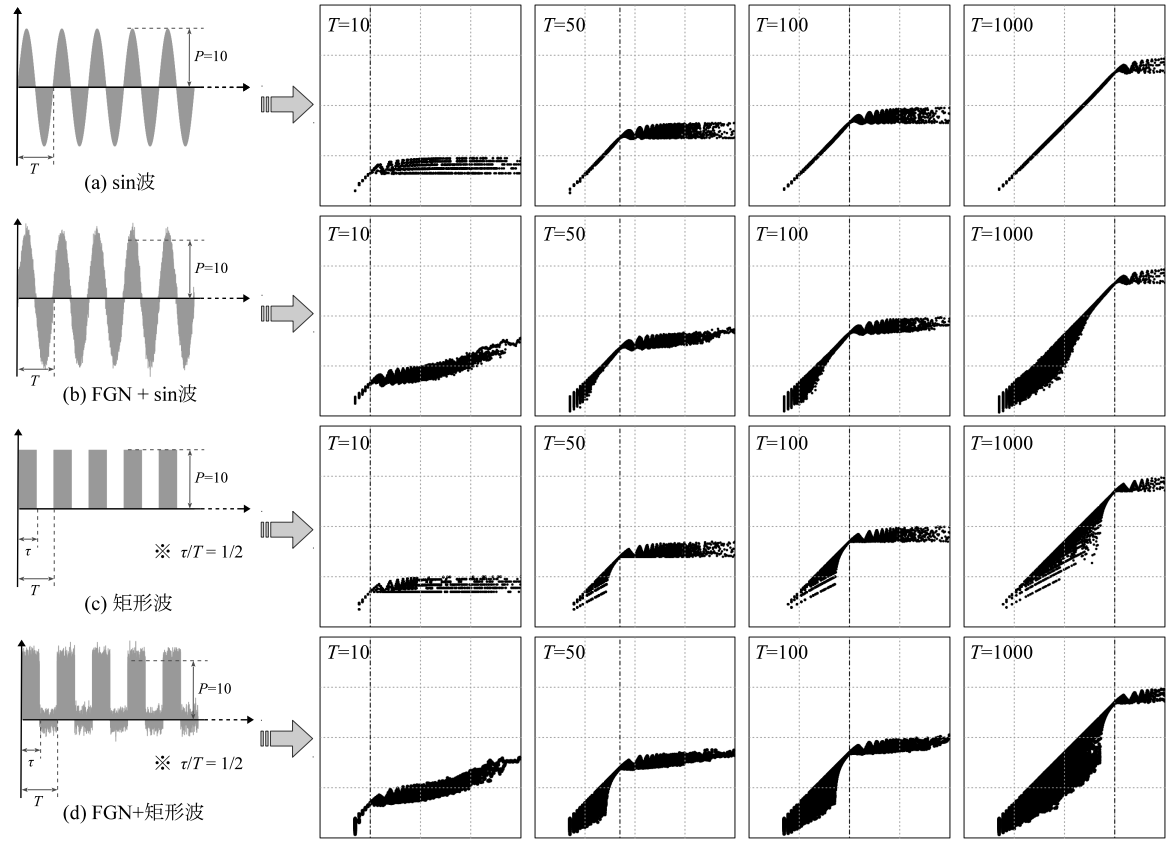


図 3 周期 T の影響
Fig. 3 Effect of period T on an R/S Pox Diagram.

化点が計算に反映されていることが分かる．これより，反転時系列を用いることで早い段階で観測される非定常的時系列に対して即応性が高まる効果が期待できる．

3. 周期性に対する R/S Pox Diagram

本研究では，間欠的に到着するパケットトラフィックのように周期性を有する時系列に着目し，その特徴を明らかにした．本解析法に関する検証は，解析に用いるパラメータが多すぎて数学的検証が困難と考えられる．そこで，本研究ではシミュレーション時系列を用いた実験的検証を行った．検証に用いたシミュレーション時系列のデータサイズ N は 10,000 とし，R/S 解析の任意長区間 n の増分 Δn は 0.01 とした．周期的時系列の波形として，周期 T の影響に対する検証には sin 波および矩形波を用いたが，振幅 P，デューティ比 D の影響に対する検証はどちらもほぼ同様な結果となったため，本論文では矩形波のみの結果を示した．

3.1 周期 T の影響

周期 T に対する R/S Pox Diagram のプロット形状変化の様子を**図 3** に示す．図 3 (a) は振幅 $P = 10$，周期 T の sin 波，図 3 (b) は分散 1.0，平均 0.0，ハーストパラメータ $H = 0.6$ とした FGN を定常状態時系列とし (a) の sin 波を重畳させたもの，図 3 (c) は振幅 $P = 10$，パルス幅 τ，周期 T とする矩形波，図 3 (d) は (b) と同様の FGN に矩形波を重畳させた時系列である．それぞれの時系列に対して周期 $T = 10, 50, 100, 1000$ と変化させたときの R/S Pox Diagram を示した．グラフの一点鎖線は $\log(n) = \log(T)$ を示している．矩形波のデューティ比 τ/T はすべて 0.5 とした．

結果より，すべてのシミュレーション時系列において $n = T$ となるポイントを境にプロット点群の傾きが水平方向に折れ曲がる傾向が顕著に現れた．また，n が前述のポイントより小さい範囲では，各 n における R_n/S_n の最大値の点群による傾きが，すべてのシミュレーションにおいてほぼ 1 となった．n が前述のポイントより大きい範囲では，$T = 10$ のとき，FGN が重畳されていると傾きが若干曲線的になったが，その他はほぼ傾きが直線的に，かつ値が 0 に近くなった．この折れ曲がるポイントが明確に周期 T を示したことを用いて，2 つの傾きを表す回帰直線の交点から周期を推定する手法を検討した．

3.2 振幅 P の影響

3.1 節で用いた FGN に周期 $T = 100$ の矩形波を重畳させた時系列に対して，振幅 $P = 2$ および $P = 10$ としたときの R/S Pox Diagram を**図 4** に示す．

$P = 10$ に比べ $P = 2$ の場合，プロット点群の折れ曲がる傾向が弱くなっている．これは，定常状態に比べある程度の振幅変化がないとき，周期的特徴が顕現しにくくなると予測される．

3.3 デューティ比 D の影響

矩形波を特徴づける指標であるデューティ比に対する影響を検討する．パルス幅 τ，周期 T のとき，デューティ比 D は式 (8) で求める．

$$D = \tau / T \tag{8}$$

3.2 節と同様に FGN に周期 $T = 100$ の矩形波を重畳させた時系列において，パルス幅 τ を変化させてデューティ比 $D = 0.02, 0.1, 0.5$ としたときの R/S Pox Diagram を**図 5** に示す．

振幅の影響と同様に，D の値に応じてプロット点群の折れ曲がる特徴に影響が現れたことから，D の値によっても周期的特徴が顕現しにくくなると予測される．

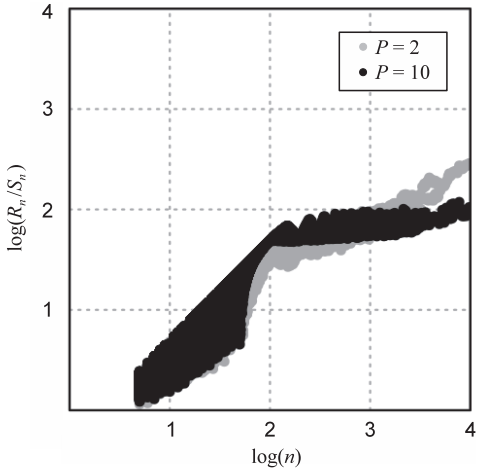


図 4 振幅 P の影響
Fig. 4 Effect of amplitude P on an R/S Pox Diagram.

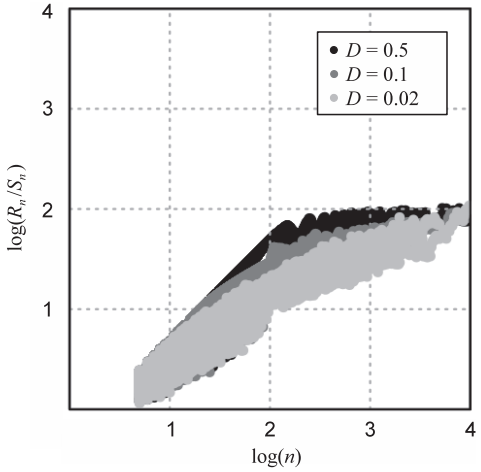


図 5 デューティ比 D の影響
Fig. 5 Effect of duty ratio D on an R/S Pox Diagram.

4. R/S Pox レッグライン特性

3 章で示した周期的時系列に対するプロット形状は，人体脚部の太腿，膝，脛と形状が非常によく似ていることから，本研究では R/S Pox レッグライン（Leg-Line）特性と呼ぶことにする．この特性から得られる特徴量を**図 6** に示す．折れ曲がるプロット形状の特徴は，脚部の太腿（Thigh）と脛（Shin）の部分に見立て，ハーストパラメータ H 導出法と同様にそれぞれの部分におけるプロット点群の傾きで表す．$n < T$ なる導出範囲 RT（Range of Thigh）における傾きを ST（Slope of Thigh），$n > T$ なる導出範囲 RS（Range of Shin）における傾きを SS（Slope of Shin）と表記する．また，$n = T$ となるポイントは，脚部の膝（Knee）と見立て，KP（Knee Point）と表記する．

本章では，この特徴量の定量化法について述べる．

4.1 各範囲における傾き（Slope Value）

R/S Pox Diagram の各範囲におけるプロット点群は，図 3 (a) のように一直線上に並ぶものや，図 3 (c)，(d) のように縦軸方向に広がるものなど様々な傾斜の形状を呈する．そこで，各範囲において 1 方向の傾きではなく，プロット点の各 n における上限点群 $\max(R_n/S_n)$，平均点群 $\mathrm{avg}(R_n/S_n)$，下限点群 $\min(R_n/S_n)$ から求まる 3 方向の傾きで表す．導出範囲 RT におけるそれぞれの傾きを $\{ST_{\mathrm{Sup}}, ST_{\mathrm{Avg}}, ST_{\mathrm{Inf}}\}$，導出範囲 RS におけるそれぞれの傾きを $\{SS_{\mathrm{Sup}}, SS_{\mathrm{Avg}}, SS_{\mathrm{Inf}}\}$ と表し，ハーストパラメータ H を求めた式 (5) と同様に最小 2 乗法を用いて推定する．

導出範囲 RT，RS の範囲長は，サンプル数とのトレードオフになる．すなわち，図 6 に示した各量の高精度導出を目指す場合，RT ないしは RS の範囲長を大とする必要があるが，この範囲長を大きくしすぎると，R/S Pox Diagram の縦軸方向でのプロット点数の減少領域をカバーしてしまい，ST や SS の変動（ゆらぎ）の増加となってしまう恐れがある．また，図 3 に示すように，周期 T の変動によって

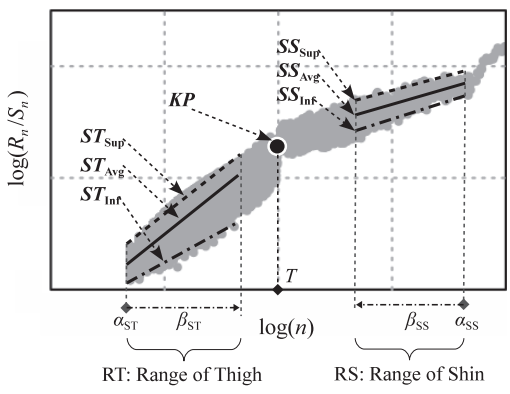


図 6 R/S Pox レッグライン特性
Fig. 6 R/S Pox Leg-line characteristics.

KP も変動するため，極端な短周期および長周期に対して範囲長を事前に設定するのは困難である．そこで，本研究では観測対象を推定可能とする周期 T の範囲を限定し，その範囲において周期推定を行うことを目標とし，RT，RT の範囲長を検討した．この手法では，周期推定範囲は限られるが，観測データ生成時の時間スケールを調整すれば，実時間上で異なる周期について推定が可能となる．つまり，並列処理により広範囲または限定した周期的時系列の観測システムへの適用も検討できる．本研究では，実時間解析における計算コストを考慮したときの時系列 X_t のデータサイズ N を 10,000 以下として検討することから，有意な推定範囲として $10 \leq T \leq 1000$ を目標とした．

導出範囲 RT，RS は以下とした．すなわち図 6 において，$\log(n)$ ではなく n の値で α_{ST}，および α_{SS} を設定し，これらを各々，RT および RS の起点とする．また，$\log(n)$ 上の値で RT および RS の範囲長 β_{ST}，β_{SS} を設定する．したがって RT は α_{ST} を起点とし，範囲長 β_{ST}，RS は α_{SS} を起点とし，図示の向きの範囲長 β_{SS} となる．RT の起点 α_{ST} は 5 とし，RS の起点 α_{SS} は，区間長 n が大きくなると求められる R/S 統計量が少なくなることから，最低 3 点が求められる $n = N/3$ と設定した．また，範囲長 β_{ST}，β_{SS} は求められる傾きの精度に影響するため調整が必要となるが，本研究では経験的にそれぞれ 0.8 と設定した．

4.2 Knee Point による周期推定

周期 T の特徴となる KP は，導出範囲 RT，RS から求められた各傾きを表す回帰直線の交点の n から求める．ただし，周期的時系列が重畳されていなくても RT，RS の傾きに少しでも差異が生じた場合，KP が導出されてしまうことがある．これを回避するため，周期的時系列が重畳されると SS が小さくなることから，SS に対して重畳判定となるしきい値 γ を設定する．これを用いて，SS の値がしきい値 γ より小さくなったとき導出を行う制約条件を設ける．また，KP の値が推定目標とした $10 \leq T \leq 1000$ の範囲外となったとき，導出できなかったものとする．制約条件を考慮した周期推定手順を以下に示す．

(a) ST_{mode1} および SS_{mode2} を 1 つずつ選択する．
(b) $SS_{\mathrm{mode2}} \geq \gamma$ ならば，(a) に戻る．
(c) KP 計算
(d) $KP < 10$，または $KP > 1000$ ならば，(a) に戻る．
(e) $T = KP$

mode1，mode2 はそれぞれ，{Sup，Avg，Inf} を表しており，これらを用いたときに導出される KP の数は 9 個となる．それぞれの傾きより計算される KP を**表 1** にまとめる．例として，ST_{Sup} および SS_{Avg} から求められる交点の n を KP_{SA} と表し，以下に計算例を示す．式 (9) は，ST_{Sup}，式 (10) は SS_{Avg} の回帰直線を表し，KP_{SA} は式 (11) より求められる．ただし，C_{STS}，C_{SSA} は定数である．

表 1 導出される KP
Table 1 Knee Point KP derived by respective slopes.

	SS_{Sup}	SS_{Avg}	SS_{Inf}
ST_{Sup}	KP_{SS}	KP_{SA}	KP_{SI}
ST_{Avg}	KP_{AS}	KP_{AA}	KP_{AI}
ST_{Inf}	KP_{IS}	KP_{IA}	KP_{II}

$$\log\{\max(R_n/S_n)\} = ST_{\mathrm{Sup}} \log(n) + C_{\mathrm{STS}} \tag{9}$$

$$\log\{\mathrm{avg}(R_n/S_n)\} = SS_{\mathrm{Avg}} \log(n) + C_{\mathrm{SSA}} \tag{10}$$

$$KP_{\mathrm{SA}} = \frac{C_{\mathrm{STS}} - C_{\mathrm{SSA}}}{SS_{\mathrm{Avg}} - ST_{\mathrm{Sup}}} \tag{11}$$

より精度の高い周期推定が可能となる KP の選定は，5 章で行う．

5. シミュレーションによる性能評価

提案する R/S Pox レッグライン特性の性能評価として，3 章で示した周期，振幅，デューティ比に対する傾き ST，SS を導出し検討を行った．さらに，シミュレーション時系列および実環境から取得された長期的ポートスキャン攻撃のトラフィック時系列に対する周期推定も行った．シミュレーション時系列は図 3 (d) に基づいた FGN に矩形波による周期的時系列を重畳させたものを使用した．

5.1 周期的特徴に対する傾き

5.1.1 周期特性

周期 T に対する傾き ST, SS をそれぞれ**図 7** (a), 図 7 (b) に示す．

結果より，ST において ST_{Sup} は $20 < T \leq 1000$ のとき，安定して約 1.0 の値を保っている．ST_{Avg} および ST_{Inf} は $T = 30$ のとき最大値をとり，その後 T が増加するにつれて，値が低下している．特に ST_{Avg} は，$T = 100$ 以降，定常状態とほぼ同じ値に戻っている．ここから，ST_{Sup} はどんな周期 T に対しても安定した周期的時系列の検知ができると推測できる．また，ST_{Avg} は，$T > 100$ となる周期性には反応せず定常状態の特性を表していると推測できる．

SS では，どの傾きも $20 \leq T \leq 800$ において 0.3 以下の値となり，水平方向に傾く特徴をとらえることができた．ただし，T が 1,000 近辺になると傾きが元に戻るようになる．これは，RS の導出範囲設定に起因するものであり，β_{SS} の調整により修正が可能と考えられる．

5.1.2 振幅特性

周期 T を 50，100，500 と設定したときの振幅 P に対する傾き ST, SS を**図 8** に示す．ST においては，周期 T による値の変動はあるものの，振幅 $P > 10$ ではすべての傾きにほとんど変化が見られなくなる．つまり，定常状態に対してある程度の大きさの P であれば，観測時系列内で振幅変動があっても，安定した値を呈すると推測できる．また，周期特性と同様に ST_{Sup} は約 1.0 の値を保っているこ

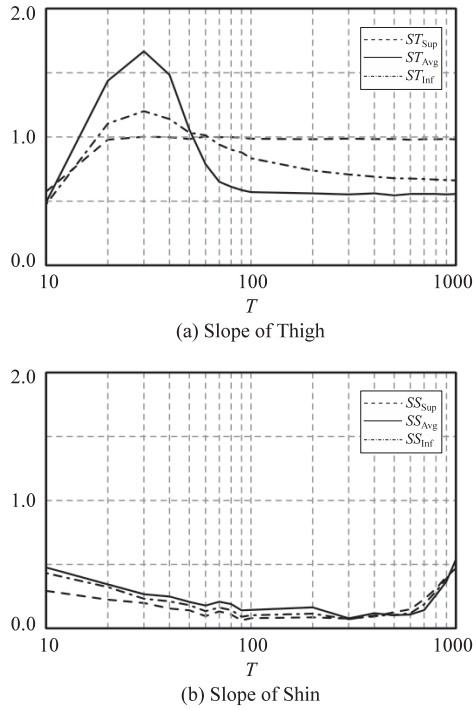


図 7 周期 T に対する傾き

Fig. 7 Slope vs. period T of simulation time series.

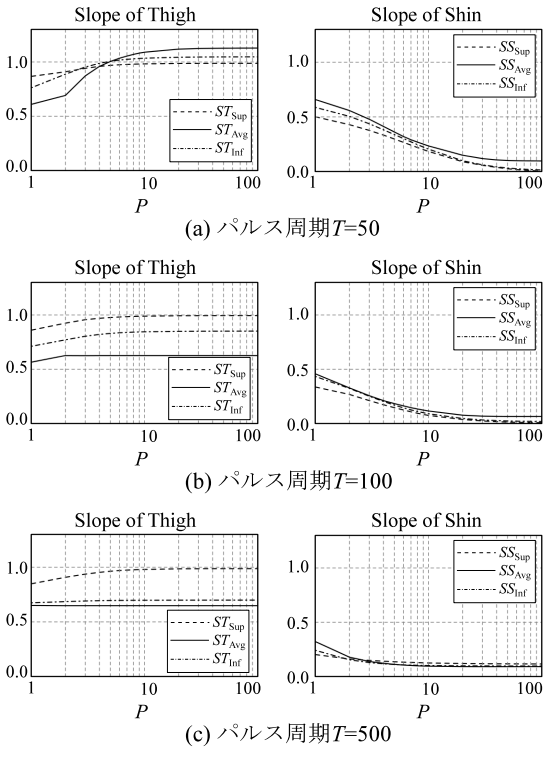


図 8 振幅 P に対する傾き

Fig. 8 Slope vs. amplitude P of simulation time series.

とから，ここでも検知指標としての有効性が推測される．

SS においては，各傾きとも P の増加にともない値が小さくなり，徐々に水平方向に傾く傾向が観測された．ただし，振幅が $1 \leq P \leq 10$ と比較的小さい場合，その傾きは周期 T の影響が大きいことも観測できた．

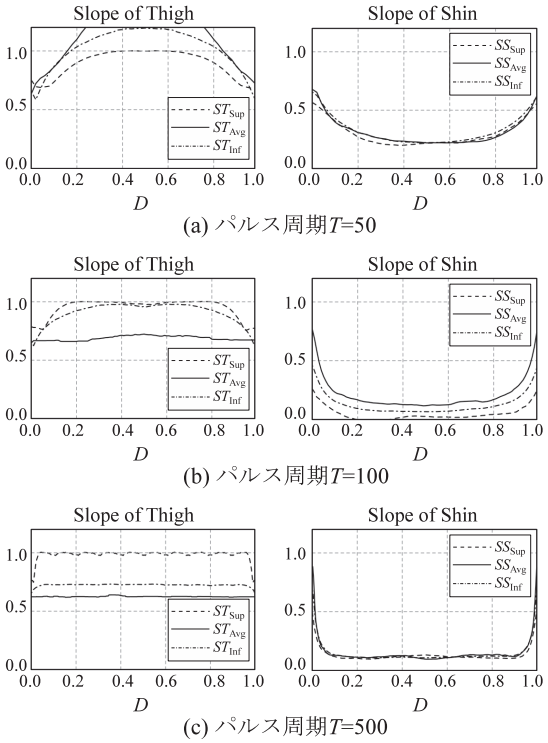


図 9 デューティ比 D に対する傾き

Fig. 9 Slope vs. duty ratio D of simulation time series.

5.1.3 デューティ比特性

周期 T を 50，100，500 と設定したときのデューティ比 D に対する傾き ST，SS を**図 9** に示す．

ST および SS のどちらの場合においても，$D = 0.5$ においてそれぞれ最大値および最小値をとることが観測された．このときの値が周期的時系列の特徴を最もとらえた値と推測される．この値が安定して観測できる D の範囲は，周期 T の値に応じて変化することが分かる．ST_{Sup} に着目してみると，$T = 50$ ではおおよそ $0.4 < D < 0.6$，$T = 100$ では $0.2 < D < 0.8$，$T = 500$ では $0.04 < D < 0.96$ で安定している．それぞれの T における安定状態開始点の D からパルス幅 τ 計算してみる．

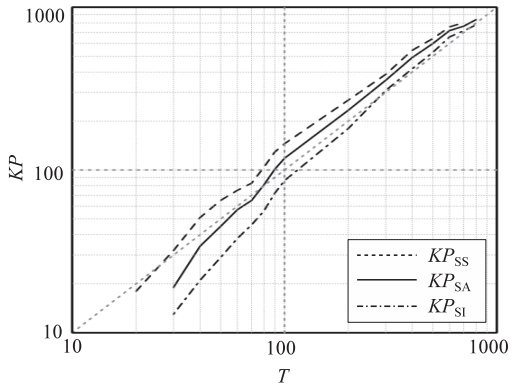
$$\begin{aligned}\tau &= D \times T \\ &= 0.4 \times 50 = 0.2 \times 100 = 0.04 \times 500 \\ &= 20 \end{aligned} \tag{12}$$

すべての T において，パルス幅 $\tau = 20$ が得られた．つまり，今回のシミュレーションにおいて，ST_{Sup} はパルス幅 $20 \leq \tau \leq T - 20$ の周期的時系列に対して安定して約 1.0 の値を示すことが分かった．

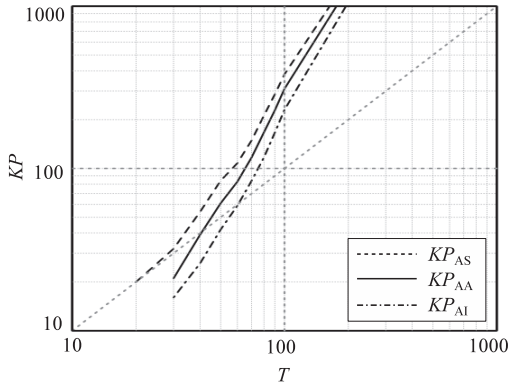
5.2 周期推定

5.2.1 シミュレーション時系列

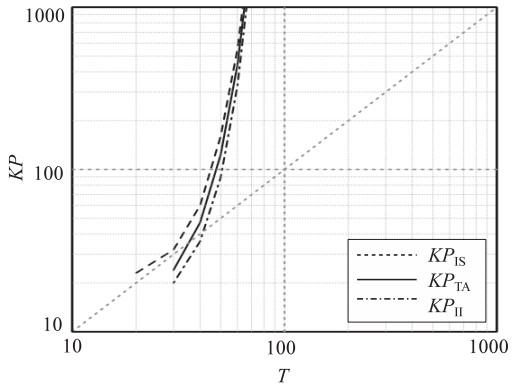
シミュレーション時系列（FGN+矩形波）を用いて周期を $10 \leq T \leq 1000$ と変化させたときの周期推定法から導出された KP を**図 10** に示す．SS のしきい値 γ は 0.3 とし



(a) ST_{Sup}利用時



(b) ST_{Avg}利用時



(c) ST_{Inf}利用時

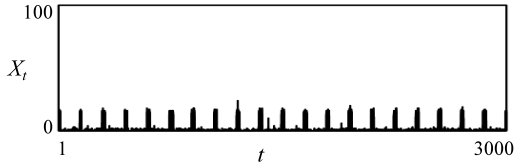
図 10 周期 T に対する KP
Fig. 10 Knee Point KP vs. period T of simulation time series.

た．図 10 (a) は，導出範囲 RT において傾き ST_{Sup} を選択したとき，RS の傾き $\{SS_{\mathrm{Sup}}, SS_{\mathrm{Avg}}, SS_{\mathrm{Inf}}\}$ から導出される $\{KP_{\mathrm{SS}}, KP_{\mathrm{SA}}, KP_{\mathrm{SI}}\}$，図 10 (b) は，傾き ST_{Avg} を選択したときに RS の傾き $\{SS_{\mathrm{Sup}}, SS_{\mathrm{Avg}}, SS_{\mathrm{Inf}}\}$ から導出される $\{KP_{\mathrm{AS}}, KP_{\mathrm{AA}}, KP_{\mathrm{AI}}\}$，図 10 (c) は，傾き ST_{Inf} を選択したときに RS の傾き $\{SS_{\mathrm{Sup}}, SS_{\mathrm{Avg}}, SS_{\mathrm{Inf}}\}$ から導出される $\{KP_{\mathrm{IS}}, KP_{\mathrm{IA}}, KP_{\mathrm{II}}\}$ を示している．グラフの対角線にある灰色の破線は，理想直線を表す．プロット点の存在しない T は，提案周期推定手法の制約条件によって導出されなかったものである．

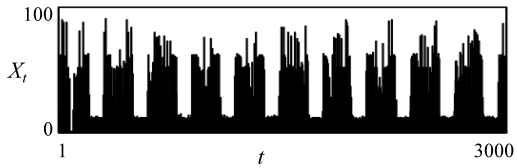
結果より，導出範囲 RT において傾き ST_{Sup} を選択した場合，他に比べ，RS のどの傾きにおいても理想直線に近似した形となり良好な周期推定が行われている．ただし，T が小さい場合は KP_{SS} が良好な結果を示し，$T = 100$ 前後では KP_{SA}，T が大きい場合は KP_{SI} とそれぞれ有効な範囲があることが示された．$T < 20$ において KP が導出されなかった原因は，5.1.1 項で示したように RT，RS の導出範囲によるものと推測される．

表 2 ポートスキャン攻撃の詳細
Table 2 Summary of the port scan attack.

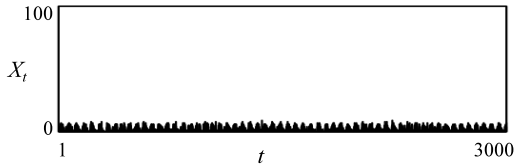
	Scan 1	Scan 2	Scan 3
計測日	2008/08/26	2008/08/27	2008/08/30
宛先ポート	4899	80	23
周期 T	約 150	約 300	約 50
パルス幅 τ	約 10～20	約 200	約 20
振幅 P	約 15	約 30～90	約 5



(a) Scan 1 (2008/8/26/09:02:16)



(b) Scan 2 (2008/8/27 03:12:15)



(c) Scan 3 (2008/8/30 11:00:00)

図 11 ポートスキャン攻撃トラフィック時系列
Fig. 11 Time series of the port scan attack.

5.2.2 実環境からのトラフィック時系列

実環境から計測されたトラフィック時系列に対する適用として，長期的ポートスキャン攻撃が含まれたトラフィック時系列を用いた周期推定を試みた．長期的ポートスキャン攻撃とは，ポートスキャンに使用する調査パケットを短時間で大量に送信せずに，時間間隔をおいて間欠的に少量の調査パケットを送信することで検知しにくくする攻撃である．間欠的に送信されるため周期的時系列の特徴を有しており，本手法によって攻撃検知，ならびに周期推定が可能な観測対象と考えられる．

周期推定に用いたポートスキャン攻撃の詳細を**表 2** に，そのときの時系列 X_t のグラフを**図 11** に示す．トラフィックデータは秋田大学キャンパスネットワークの対外接続ポートから tcpdump を用いて取得されたパケットデータで，学外から到着した TCP SYN パケットのみをカウント

表 3 周期推定結果
Table 3 Result of period estimation.

	Scan 1	Scan 2	Scan 3
KP_{SS}	129	299	64
KP_{SA}	108	194	46
KP_{SI}	86	114	32

したものである．この時系列 X_t には，攻撃トラフィック以外にも正常な通信で発生した TCP SYN パケットも含まれる．計測単位時間 Δt は 0.02 s，時系列 X_t のサイズ N は 3,000 点，つまり解析に用いたデータは 60 s 間に計測されたものである．表 2 の周期 T は，それぞれの攻撃における周期パルス間の間隔時間の平均値から求めた実測周期である．

提案手法による周期推定結果を**表 3** に示す．導出範囲 RT の傾きは，前項で示したように安定した値をとる ST_{Sup} を利用して KP を 3 点導出した．結果より，すべてのスキャン攻撃に対して，KP_{SS}，KP_{SA}，KP_{SI} のうち表 2 に示した実測周期の値により近い値を示したのは KP_{SS} であった．また，Scan 1 のようにパルス幅 τ の変動，Scan 2 のように振幅 P の変動がランダムに発生している場合でも，良好な周期推定が行われた．ここから，提案する周期推定法はパルス幅，振幅の変動に対しロバストな手法であると推測できる．

6. まとめ

本研究では，まず，ハーストパラメータ H の推定に用いられていた R/S 解析法の未計算部分の存在を指摘し，反転時系列を適用することで，直近で発生した事象が解析に反映させることを示した．これより，より高い即応性，実時間解析への適用が期待できる．

さらに，R/S Pox Diagram に現れる周期的時系列の特徴に着目し，これを定量化する R/S Pox レッグライン特性を提案した．ここで示した導出範囲 RT，RS の傾きは，周期的時系列の重畳に対して顕著な変化を示すことから，事象変化の検知指標として期待できる．また，プロット形状の折れ曲がるポイントが明確に周期を示すことを利用した周期推定法も提案し，パケットトラフィックを想定したシミュレーション時系列ならびに実環境から実測された長期的ポートスキャン攻撃トラフィック時系列に対して検証を行った．この結果より，Thigh 部分の上限点群スロープ ST_{Sup} と Shin 部分の上限点群スロープ SS_{Sup} から求められる Knee Point KP_{SS} を用いた場合，良好な周期推定が行えることを示した．

今後は，導出範囲の調整による周期推定の精度向上の実現，ならびに事象の経時変化に対する特徴量の時間特性の明確化により，本特性を利用した長期的スキャン攻撃の検知手法について研究を進める．

謝辞 本研究の一部は科研費（23500077）および東北大学電気通信研究所における共同プロジェクト研究 H22/A14 の助成を受けたものである．

参考文献

[1] Leland, W.E., Taqqu, M.S., Willinger, W. and Wilson, D.V.: On the Self -Similar Nature of Ethernet Traffic, *Computer Communications Review*, Vol.23, No.4, 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 (1999).
[3] 土井博生，松田崇弘，山本 幹：TCP ふくそう制御がトラヒックのマルチフラクタル性に与える影響，電子情報通信学会論文誌 B，Vol.J88-B, No.6, pp.1029–1037 (2005).
[4] Fukuda, K., Takayasu, M. and Takayasu, H.: A cause of self-similarity in TCP traffic, *International Journal of Communication Systems*, Vol.18, No.6, pp.603–617 (2005).
[5] 上田 浩，奈須野裕，岩谷幸雄，木下哲男：確率過程による LAN トラヒックのモデル化における一考察，情報処理学会論文誌：数理モデルと応用，Vol.48, No.SIG 2 (TOM 16), pp.167–174 (2007).
[6] Li, M.: Change trend of averaged Hurst parameter of traffic under DDOS flood Attacks, *Computers & Security*, Vol.25, No.3, pp.213–220 (2006).
[7] Hurst, H.E.: A suggested statistical model of some time series which occur in nature, *Nature*, Vol.180, No.494 (1957).
[8] Mandelbrot, B.B. and Wallis, J.R.: Robustness of the Rescaled Range R/S in the Measurement of Noncyclic Long-Run Statistical Dependence, *Water Res.*, Vol.5, No.5, pp.967–988 (1969).
[9] Beran, J., Sherman, R., Taqqu, M.S. and Willinger, W.: Long-Range Dependence in Variable-Bit Rate Video Traffic, *IEEE Trans. Comm.*, Vol.43, No.2/3/4, pp.1566–1579 (1995).
[10] Taqqu, M.S., Teverovsky, V. and Willinger, W.: Estimators for long-range dependence an empirical study, *Fractals*, Vol.3, No.4, pp.785–798 (1995).
[11] Igarashi, R., Ono, S., Takahashi, A., Iwaya, Y. and Sakata, M.: Some Features of Network Traffic Depending on Protocols, *Proc. ICMR 2005*, pp.426–431 (2005).
[12] Dang, T. and Molnar, S.: On the Effects of Non-Stationarity in Long-Range Dependence Tests, *Periodica Polytechnica, Ser. El.*, Vol.43, No.4, pp.227–250 (1999).
[13] 松葉育雄：非線形時系列解析，pp.83–91, 朝倉書店，東京 (2000).
[14] Takahashi, A., Igarasjo, R., Ueda, H., Iwaya, Y. and Kinoshita, T.: Network Anomaly Detection Based on R/S Pox Diagram, *International Journal of the Society of Materials Engineering for Resources*, Vol.17, No.2, pp.186–192 (2010).
[15] 三鑰達真，吉田和幸：長期的スキャンニングを対象としたスキャン攻撃検知システム，電子情報通信学会技術研究報告，Vol.107, No.449, pp.39–44 (2008).

推薦文

研究会にて関連研究の発表があり興味深く，また手法の有効性は広く論じられるべきであると考えられる．ページ

数が 6 ページと不足しているが，体裁を整えたうえで投稿されるならば本研究会として推薦を行うものとする．

（FIT2012 第 11 回情報科学技術フォーラム
プログラム委員長　橋田浩一）

高橋 秋典　（正会員）

1991 年秋田大学鉱山学部電子工学科卒業．同年秋田大学鉱山学部助手．1996 年東京大学工学部文部省内地研究員．2007 年秋田大学工学資源学部助教．ネットワークトラフィック特性解析，トラフィック可視化等の研究に従事．電子情報通信学会，電気学会，日本素材物性学会各会員．

五十嵐 隆治　（正会員）

1974 年秋田大学鉱山学部電気工学科卒業．1974 年アキタ電子株式会社．1975 年秋田大学鉱山学部助手．1986 年工学博士（北海道大学）．1987 年 LBL：アメリカ合衆国ローレンス・バークレー国立研究所客員研究員．1987 年秋田大学講師．1995 年秋田大学助教授．2011 年秋田大学工学資源学部教授．ネットワークトラフィックの実測と確率過程による解析等の研究に従事．電気学会，電子情報通信学会，応用物理学会，計測自動制御学会，IEEE，放電学会，応用物理学放射線分科会，日本素材物性学会各会員．

上田 浩　（正会員）

2004 年豊橋技術科学大学大学院博士後期課程修了．同年東北大学電気通信研究所博士研究員．2006 年群馬大学総合情報メディアセンター助教授．2011 年京都大学学術情報メディアセンター准教授．博士（工学）．確率過程モデル，HIV と免疫系の相互作用モデル，生態系の数理モデル等の研究に従事．電子情報通信学会，日本数理生物学会各会員．

岩谷 幸雄　（正会員）

1993 年東北大学大学院修士課程修了．同年秋田大学鉱山学部助手．2000 年秋田大学工学資源学部講師．2002 年東北大学電気通信研究所助教授．2012 年東北学院大学工学部教授．博士（情報科学）．ネットワークの知的管理手法，音空間知覚過程の解明とその応用等に従事．FIT2003 論文賞，FIT2005 船井ベストペーパー賞等受賞．電子情報通信学会，日本音響学会，日本 VR 学会，米音響学会各会員．

木下 哲男　（フェロー）

1979 年東北大学大学院修士課程修了．同年沖電気工業（株）入社．1996 年東北大学電気通信研究所助教授，2001 年同大学情報シナジーセンター教授，現在，同大学電気通信研究所教授．知識工学，エージェント工学，エージェント応用システム等の研究開発に従事．情報処理学会平成元年度研究賞，同平成 8 年度論文賞，電子情報通信学会平成 13 年度業績賞等受賞．工学博士．電子情報通信学会，人工知能学会，IEEE，ACM，AAAI 各会員．