R/S Pox Diagram のプロット度数分布に着目した異常検知手法に関する検討

○髙橋俊彦[†]　髙橋秋典[†]　五十嵐隆治[†]　上田浩[‡]　岩谷幸雄[§]　木下哲男[¶]
秋田大学[†]　京都大学[‡]　東北学院大学[§]　東北大学[¶]

インターネットが現代社会で重要性を増すに伴い，ネットワーク攻撃も巧妙化している。その中で，攻撃検知を困難とさせるための低レートパケットトラヒックによる長期的ポートスキャンがある。我々はこの長期的ポートスキャンの周期性に着目して，R/S Pox レッグライン特性と呼ばれる特徴量を用いて異常検知を行うという研究を進めている。この手法は，パケットの到着間隔を表す周期の範囲を限定しており，範囲外となる長周期の場合，異常を判別するのは困難であった。我々はこの問題点に対して，パケット時系列におけるレベルシフト的トラヒック量変化に対する R/S Pox Diagram の特徴的プロット分布を定量化することで異常を検知する手法を提案した。検証として，長周期シミュレーション時系列に対する異常検知を行った結果，従来法では検知困難だった周期範囲も検知可能となることを示した。

A Study of Network Anomaly Detection based on Plotted Dots Frequency Distribution of R/S Pox Diagram

○Toshihiko Takahashi[†]　Akinori Takahashi[†]　Ryuji Igarashi[†]
Hiroshi Ueda[‡]　Yukio Iwaya[§]　Tetsuo Kinoshita[¶]
Akita University[†] Kyoto University[‡] Tohoku Gakuin University[§] Tohoku University[¶]

1. はじめに

インターネットは現代社会における重要な基盤の一つとなっており、多種多様な役割を全うしている一方、このネットワークが脅かされる事により重篤な被害に遭う恐れがある。巧妙化するネットワーク攻撃に対する異常検知手法として，トラヒック時系列の自己相似性の様相変化に着目して，長期的ポートスキャンや低レート DoS 攻撃のような攻撃トラヒックの周期性を検知する手法[1][2]が提案されている。これは，トラヒックの周期性に対する R/S Pox レッグライン特性の特徴的変化を定量化するものであるが，特徴量を導出する観測時系列において，周期が長い攻撃トラヒックの場合，検知不能となる問題点がある。

そこで，本研究ではこの問題点に対して，R/S Pox Diagram のプロット度数分布に着目し，長周期の周期的時系列発生時のトラヒック量変化に対する特徴を捉え，異常検知を行う手法について検討した。

2. R/S Pox レッグライン特性

R/S Pox レッグライン特性による異常検知手法[1]の概要について述べる。まず，計測単位時間Δt 毎に到着パケットを計数した観測時系列 X_t（データサイズ N）に対して，時系列内で重複せずに区間長 n となる任意長区間を定める。この各区間において，当該区間の区間平均，および同一区間内での累積和と線形的な傾向との差から求められる累積範囲 R_n と当該区間の標準偏差 S_n の比を用いて，R/S 統計量を導出する。観測時系列の任意長区間 n 全てで R/S 統計量を導出した後，任意長区間を Δn ずつ増加させ R/S 統計量を導出する。図 1 に示すように，横軸 log(n)，縦軸 log(R_n/S_n)にプロットしたグラフが R/S Pox Diagram と呼ばれ，R/S Pox レッグライン特性は，導出範囲 RT および RS におけるプロット点の各 n における上限点群 max(R_n/S_n)，平均点群 avg(R_n/S_n)，下限点群 min(R_n/S_n)のそれぞれの傾きを特徴量とする。この特徴量を用いて，任意長区間 n と周期的時系列の周期 T が等しくなる点 KP を導出することで時系列の周期性の存在を検知することができる。

しかし，従来手法による周期推定範囲は制限が設けられており[1]，範囲外の短周期，長周期に対しては検知不能となる問題点がある。そこで本研究では，突発的トラヒック量増加時の R/S Pox Diagram のプロット度数分布に着目し，その特徴を用いた異常検知手法について検討し，従来法との検知性能について比較を試みた。

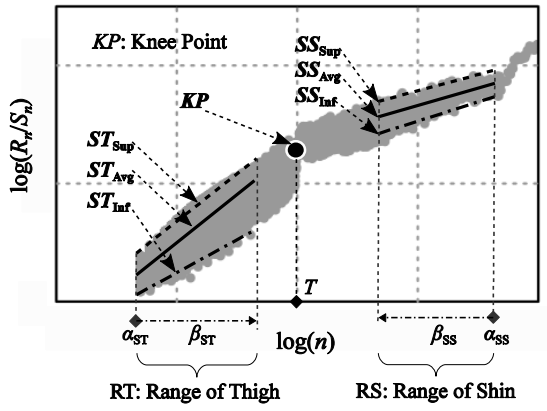

図 1　R/S Pox レッグライン特性の特徴量

3. 提案手法

R/S Pox Diagram のプロット点は，周期性を有する時系列の場合，特徴点 *KP* を境にプロット形状が折れ曲がる特徴を有している。また，観測時系列内にレベルシフト的なトラヒック量変化が発生した場合，任意長区間 n の累積範囲 R_n が増加することから，当該区間の R/S 統計量は大きくなると推測される。つまり，長周期ではあるがトラヒック量が変化する点が複数観測された場合，R/S Pox Diagram の各 n におけるプロット点がグラフ上部に集中すると考えられる。この傾向を定量化する手法を検討し，異常検知に適用する手法を提案する。

3.1 プロット度数分布の定量化

R/S Pox Diagram のプロット度数分布を定量化する手法について述べる。まず，図2に示すように，R/S Pox Diagram を横軸Δx，縦軸Δy ごとに分割する領域を設定する。この分割領域におけるプロット度数 *his*[*i*][*j*]をカウントする。ここで，変数 *i*，*j* は当該領域の行番号，列番号を表す。また，Δx，Δy は経験的に0.1と定めた。

また，異常検知に用いるプロット度数分布データは，区間長 n を $1.5 \leq \log(n) \leq 2.5$ の範囲に制限した。これは，n が小さい範囲では感度が低く，また大きい範囲ではプロット点数が少ないため，検知性能が低下すると推測されるためである。この対象範囲を関心領域と呼ぶことにする。

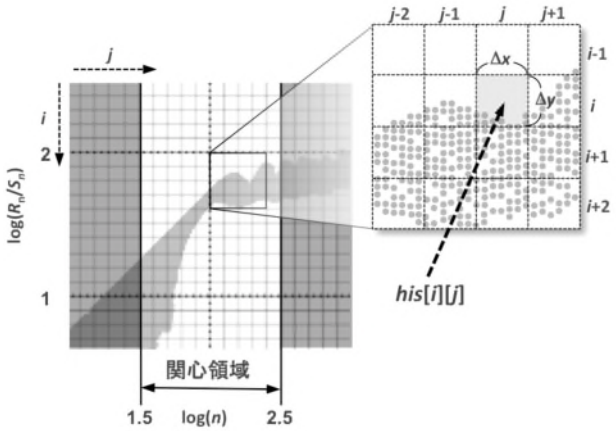

図2 プロット度数分布

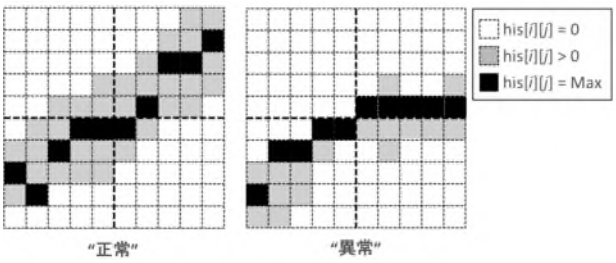

図3 検知例

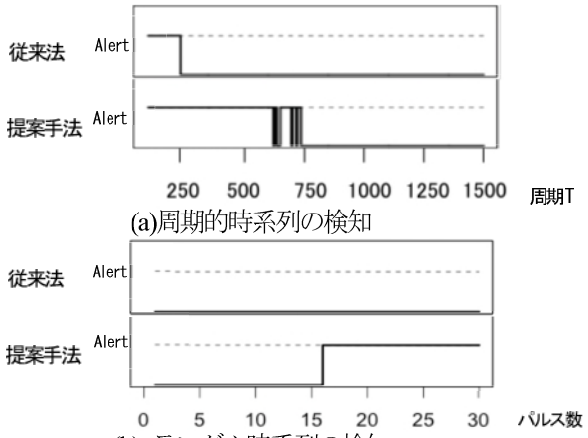

図4　シミュレーション結果

3.2 異常検知法

3.1 のように定量化されたプロット度数分布データに基づいた異常検知手法について述べる。まず、関心領域における対象列 *j* の上部より走査を行い，最初に度数 *his*[*i*][*j*]>0 となった値をその対象列の基準値 *SL*(*j*)= *his*[*i*][*j*]とする。次に，対象列 *j* の残りの度数 *his*[*i*][*j*]と *SL*(*j*)と比較して，*SL*(*j*)が対象列 *j* の最大値であった場合，異常と判別し，検知走査を終了する。*SL*(*j*)が対象列 *j* の最大値でなかった場合，対象列を *j*+1 とし，同様の走査を行う。関心領域のすべてに対して走査を行っても，異常と判別されない場合は，正常と判別して走査を終了する。検知例を図3に示す。

4. シミュレーションによる検証

データサイズ N=3000 の FGN 時系列に対して，振幅10パルス長50とし周期 T を変化させた周期的時系列，および到着時間間隔をランダムに設定し，パルス数を変化させたランダム時系列を重畳させたシミュレーション時系列を用いて性能評価を行った。周期的時系列重畳時の結果を図4(a)，ランダム時系列重畳時の結果を図4(b)に示す。

周期的時系列に対する従来法による検知では，周期 T>250 において完全に検知不能となったが，提案手法では周期 T=600 程度まで安定して異常を検知し，周期 T>750 程度より検知不能となった。また，ランダム時系列に対する検知では，従来法ではすべてのパルス数において検知不能だったが，提案手法では，パルス数16程度より検知可能となった。この結果より，提案手法は，従来法と比較して，長周期トラヒック検知およびランダムトラヒック検知の面で有効であることが示された。

5 まとめ　本手法の検知性能の向上，および時系列内のレベルシフトの個数を推定する手法について検討する。

謝辞　本研究の一部は東北大学電気通信研究所共同プロジェクト研究H27/A24の助成を受けたものである。

参考文献

[1] 高橋秋典, 五十嵐隆治, 上田浩, 岩谷幸雄, 木下哲男, R/S Pox レッグライン特性, 情報処理, Vol.54, No.6, pp.1761-1770, 2013.
[2] 加賀谷, 他, R/S Pox レッグライン特性を用いたトラフィック異常検知に関する研究, 平成26年度第1回情報処理学会東北支部研究会, p.3, Dec. 2014.