3Z-2

R/S Pox レッグライン特性を用いたトラフィック状態認識法に関する研究

加賀谷享諒† 高橋秋典† 五十嵐隆治† 上田浩‡ 岩谷幸雄§ 木下哲男¶

†秋田大学工学資源学部 ‡京都大学学術情報メディアセンター
§東北学院大学大学院工学研究科 ¶東北大学大学院情報科学研究科

1 はじめに

近年ネットワークは重要な社会基盤となっており，様々なサービスが展開されている．そのため，常にネットワークから情報システムに侵入される危険性が存在する．攻撃者は侵入の前に対象ホストのポートスキャンを行い，提供されているサービスを調べる．サービスに脆弱性があればそこから攻撃や侵入が行われてしまう．ポートスキャンの検知により，そのような攻撃や侵入の抑制・発見・防止が可能となれば，ネットワーク管理者には有益な対策の一つとなる．

本研究では，トラフィック時系列の変化点をパターン認識で自動検知する手法を提案する．パターン認識に用いる特徴量として R/S Pox レッグライン特性を用いる．また識別のために，階層型ニューラルネットワークをシミュレーショントラフィックで学習させる．

2 R/S Pox レッグライン特性

R/S Pox レッグライン特性 [1] は，時系列の自己相似性の評価指標となるハーストパラメータ $H \in (0,1)$ の推定法の一つである R/S 解析 [2] から得られる 6 個の特徴量で記述できる．R/S 解析では時系列の任意長区間における R/S 統計量を計算し，これを両対数グラフにプロットした点群 (Pox Diagram) の回帰直線の傾きから H を推定する．レッグライン特性は，**図 1** のように Pox Diagram の部分点群に対して回帰直線を導出したものである．横軸前半の範囲 RT と後半の範囲 RS でそれぞれ上限点群，平均点群，下限点群の回帰直線を求め，それぞれの傾きを特徴量とする．周期

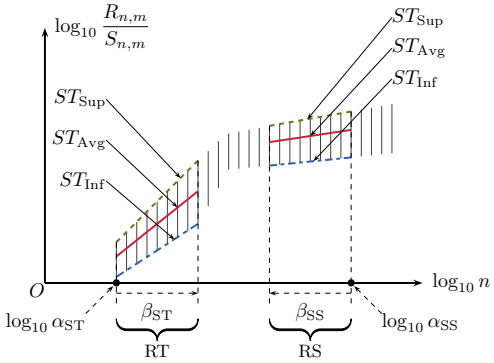

図 1: R/S Pox レッグライン特性

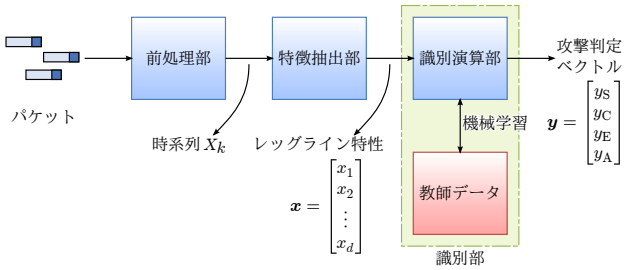

図 2: トラフィック状態認識系の構成

性を持った時系列に対して Pox Diagram は折れ曲がる性質を持つことから，レッグライン特性は時系列に含まれる周期的信号の周期推定に用いられる [1]．

3 識別システムの構成

本研究では**図 2** のような流れでパターン認識を行う．

前処理部では NIC からキャプチャしたパケットを単位時間 $\Delta t = 0.02$ 秒ごとにカウントし，時系列 $X = \langle X_0, \ldots, X_{N-1} \rangle \in \mathbb{R}^N$ を生成する．

特徴抽出部では時系列 X からレッグライン特性を計算する．**図 3** のようにサイズ $w = 3000$ の解析区間を $\Delta w = 50$ ずつずらしたものを $\nu = 3$ 個用意し，それぞれの区間でレッグライン特性を計算する．特徴ベクトルは ν 個のレッグライン特性を並べて $d = 6\nu = 18$

Study of Traffic State Recognition using R/S Pox Leg-Line Characteristics

Takaaki Kagaya† Akinori Takahashi† Ryuji Igarashi†
Hiroshi Ueda‡ Yukio Iwaya§ Tetsuo Kinoshita¶
†Faculty of Engineering and Resource Science, Akita University
‡Institute for Information Management and Communication, Kyoto University
§Engineering Graduate School of Tohoku Gakuin University
¶Graduate School of Information Sciences, Tohoku University
{s7509413, akinori, igarashi}@ie.akita-u.ac.jp

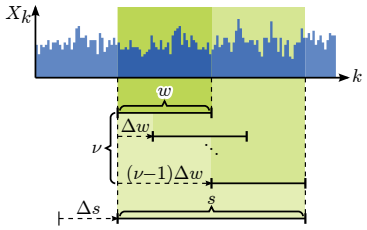

図 3: R/S 解析を行う区間とその移動

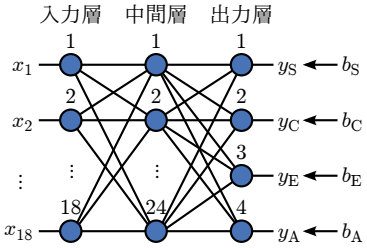

図 4: 3 層ニューラルネットワーク

次元とする．

識別部では**図 4** に示すニューラルネットワーク [3] に特徴ベクトルを入力し，その出力からトラフィック状態を推定する．ニューラルネットは 3 階層で，各層のニューロン数をそれぞれ入力層 $d = 18$ 個，中間層 24 個，出力層 4 個とする．ニューラルネットの学習には誤差逆伝播法を用い，学習終了条件はすべての教師データの教師信号 $\boldsymbol{b}$ とニューラルネットの出力 $\boldsymbol{y}$ の各成分の誤差が 0.1 以下となることとする．

識別器の学習のための教師データは，**図 5** に示す攻撃を含んだシミュレーショントラフィック時系列 X_k と攻撃フラグ A_k から生成される．A_k は X_k が攻撃であれば 1，そうでなければ 0 とする．サイズ $s = w + (\nu - 1)\Delta w = 3100$ の注目区間を Δs ずつ進めながら，X_k から特徴ベクトル $\boldsymbol{x}$ を，A_k から攻撃判定ベクトル $\boldsymbol{b}$ を生成する．$\boldsymbol{b} = [b_S, b_C, b_E, b_A]^T \in \{0,1\}^4$ は次の攻撃判定基準に従って生成される．

- b_S (攻撃開始): フラグの立上り
- b_C (攻撃継続): フラグが 1 の状態で注目区間終了
- b_E (攻撃終了): フラグの立下り
- b_A (攻撃存在): フラグに 1 が存在

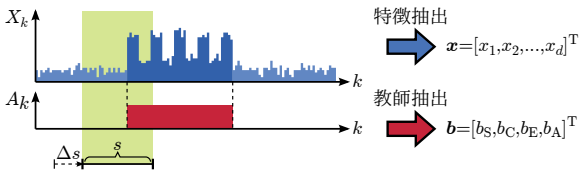

図 5: 攻撃フラグからの教師データ生成

表 1: 識別結果と理想判定の一致率

データ	開始	継続	終了	存在
2008/09/06 06 時	0.674	0.803	0.808	0.908
2008/09/06 19 時	0.757	0.615	0.577	0.916

4 実験

学習用のシミュレーショントラフィックを生成し，識別器を機械学習させた．シミュレーショントラフィックは，定常時系列を Fractional Gaussian Noise[4]（ハースト指数 H: 0.5, 0.7; 分散 $V = 1$）で表し，攻撃をパルス（波長 20, 50, 100; 振幅 15; デューティ比 0.5）で表現した時系列とした．注目区間の移動を $\Delta s = 500$ ステップとすると，生成された教師データは 260 組となった．

学習させた判別系を用いて実験を行った．実トラフィックから生成した時系列 X_k に対してパターン認識を行った．トラフィックは秋田大学ネットワークの対外接続ポートでキャプチャしたものである．特徴抽出の注目区間の移動を $\Delta s = 50$ ステップとした．**表 1** は目視により設定した攻撃フラグ A_k による判定 b_j とパターン認識結果 y_j の一致率を示す．一致率は注目区間の個数 $n_j = \lfloor (N - s)/\Delta s \rfloor + 1$ に対する各判定 j の一致数 m_j の割合 m_j/n_j である．ニューラルネットの出力 y_j は $y_j \in (0,1)$ であり，$y_j \notin \{0,1\}$ であるが，$y_j < 0.5$ を 0，$y_j \geq 0.5$ を 1 として扱った．

5 まとめ

レッグライン特性を用いたパターン認識は攻撃の存在をある程度判別可能であることが明らかとなった．一方，攻撃開始・継続・終了などといった詳しい状態の判別は現在不得手である．今後，検出性能を高めるための手法を逐次考案する．

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. 第 11 回情報科学技術フォーラム講演論文集, No. 4, pp. 9–16, Sep. 2012.

[2] W. E. LELAND. On the Self-similar nature of ethernet traffic. *Proc. ACM SIGCOMM 1993, NY, USA*, 1993.

[3] C. M. Bishop, et al. *Neural Networks for Pattern Recognition*. Clarendon Press Oxford, 1995.

[4] B. MANDELBROT. The Fractional Brownian motions, fractional noises and applications. *SIAM Rev*, Vol. 10, pp. 422–436, 1968.