パケットヘッダのパターン解析による不正アクセスの検出

-提案手法による結果と評価-

松永 武　安井 浩之　松山 実

武蔵工業大学

1. はじめに

IDS(Intrusion Detection System)の不正アクセスを検出する方法には，不正検出と異常検出がある．不正検出はシグネチャと呼ばれる不正アクセスの情報をあらかじめ登録しておき，監視したデータと比較することで不正アクセスを検出する．一方，異常検出はネットワークトラフィックやシステム情報からプロファイルと呼ばれる通常状態を定義し，監視時の入力情報と比較して，差異が大きいとき異常とする．この方法は未知の不正アクセスを検出することが期待できるが，プロファイル作成に関する決定的な手法は確立しておらず，実用的なシステムにおいては不正検出が用いられている．[1]

本稿では, TCP プロトコルのパケット遷移情報に規則性があると仮定し，同一のポートの隣り合うパケットデータの相関関係をプロファイルして異常検出する方法を提案し[2], 試行した結果と評価を報告する．

2. 異常検出方法

2.1. パケットヘッダの時系列データ

プロファイルを作成するための入力データは IP ヘッダのパケット長と TCP ヘッダの Code Bit(UGR,ACK,PSH,RST,SYN,FIN) をポート別に分割したデータ採用した．

2.2. プロファイル方法

ポート別のパケット長の時系列データを $L(t)$とし，隣り合うパケット長 $L(t)$,$L(t+1)$を 1 つの組として 2 次元のマップ上にカウントしていく．

$x=L(t)$　$y=L(t+1)$　$f_{qL}(x,y) \leftarrow f_{qL}(x,y)+1$

以降，このマップを時系列相関マップ f_q と呼ぶ．Fig.1 は時系列でのパケット長のプロファイル方法を示した図である．

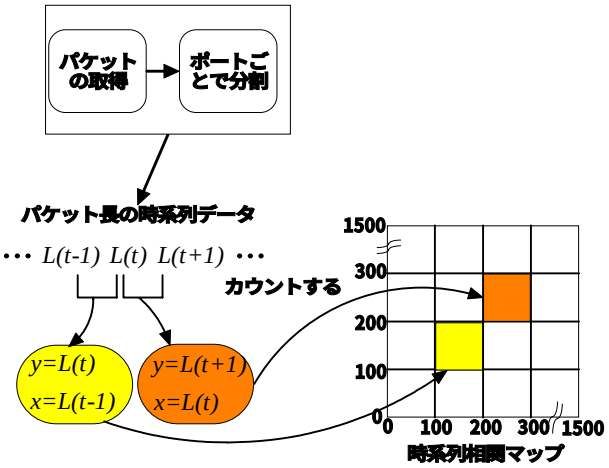

Fig.1 時系列相関マップの作成方法(パケット長)

Code Bit のデータは 6 ビットあり, 数値に変換した後，パケット長と同様の方法で時系列相関マップを作成している．

ポート別にパケット長と Code Bit の時系列相関マップを作成し，プロファイルデータを作成し，測定データとの比較により検出を行う．(Fig.2)

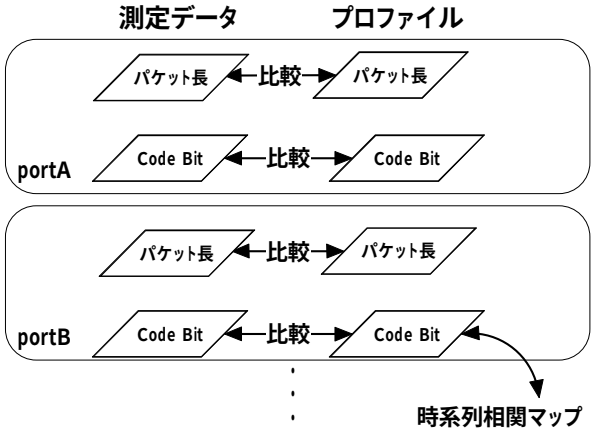

Fig.2 検出方法

Anomaly detection with pattern analysis of packet header
Takeshi Matunaga, Hiroyuki Yasui ,Minoru Matuyama
Musashi Institute of Technology

2.3. 検出方法

検出方法はプロファイルと測定したデータから作成した時系列マップより特徴量を抽出して行う．具体的には，各マップの分布確率 $p(x,y)$を求め，しきい値 αより分布確率が大きい場合の座標を特徴量として比較し，座標位置がマッチした数によって判別する．異常と判断するしきい値 β はプロファイル時にマッチした座標数より値を決めた．

3. 結果

時系列マップで検出した結果を以下に示す．測定期間は 30 日とし，事前に 10 日分のデータをプロファイルしている．パケット長のマップの大きさを 10×10，Code Bit を 64×64 にした．測定データ 1000 パケットごとにプロファイルデータと比較している．

Fig.3 は時系列相関マップで検出した HTTP ポートの結果であり，プロファイルと測定したデータのマッチした座標数の変化を示している．

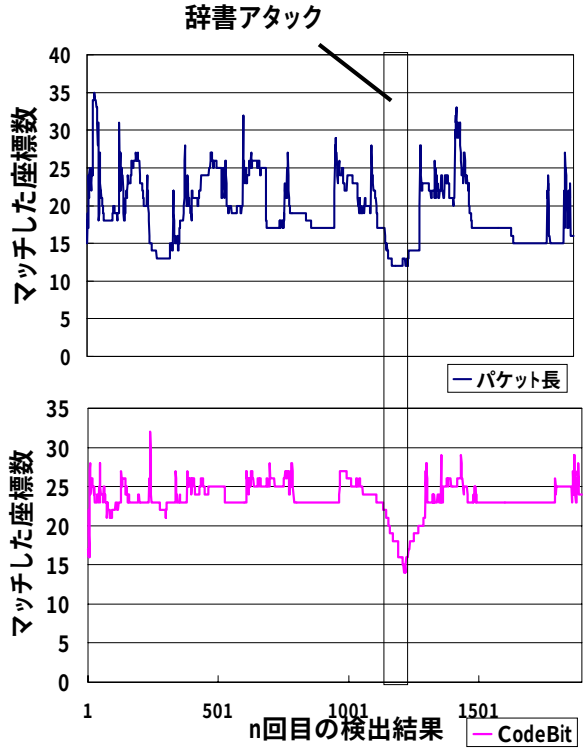

Fig.3 結果(ポート HTTP)

測定データの中には辞書アタックを行ったときのデータを含めておき，時系列マップで検出できるか測定した．Fig.3 に辞書アタックを行ったときの状態を示している．

辞書アタックによる検証すると，Fig.3 に示すように，辞書アタック時はマッチした座標数が減少していることから検出は可能である．連続してユーザ名とパスワードの認証処理を行うので，同じ大きさのパケット長が連続して大量に発生するためマッチする座標数が減少する．Code Bit についても，同じフラグの繰り返しになるためマッチする座標数が減少する．

Fig.3 の結果では誤検出率はパケット長で 5%，Code Bit で 3%になった．

4. まとめ及び今後の課題

本稿では時系列マップよる異常検出について検証し，辞書アタックツールを検出することができることが確認できた．また，Syn,Fin 攻撃の検出も期待できる．しかし，システム上ポート別に測定しているのでポートスキャンなどは検出できない．今のところ誤検出率が高く，実用的ではないが，検出手法の一つとして，有効であるといえる．

今後の課題として，誤検出率が高いことから，パラメータ（マップの大きさやしきい値など）を変更して実験を行い，有効なパラメータを特定することが必要である．

参考文献

[1]武田圭史:侵入検知システムに関する研究の現状，情報処理 Vol.42 No.12, pp.1169-1174

[2]松永武,安井浩之,松山実:パケットヘッダのパターン解析による未知の不正アクセスの検出，第 64 回情報処理学会全国大会講演論文集(3) pp.381-382