2003年日本オペレーションズ・リサーチ学会
秋季研究発表会

ミスユース型侵入検知システムのログを利用したアノマリ検出

(株) NTTデータ　山中　啓之　YAMANAKA Hiroyuki
(株) NTTデータ　小堀　誠　KOBORI Makoto
01404540　(株) NTTデータ　*中川慶一郎　NAKAGAWA Keiichiro

1　はじめに

インターネットの利用拡大に伴い，コンピュータ・ウイルスの感染，外部ネットワークから内部ネットワークへの不正アクセスといった脅威が増えており，特に国の重要インフラの基幹をなす情報システムに対する攻撃，いわゆるサイバーテロが危惧されている．

現在，このようなシステムでは，ネットワーク・セキュリティを守るために，外部と内部の通信をコントロールするファイアウオール，ネットワーク経由の攻撃を検知する侵入検知システム (IDS: Intrusion Detection System) が導入されている．

IDSの機能は，ミスユース検出 (misuse detection) とアノマリ検出 (anomaly detection) に分類される [1]. ミスユース検出とは，予め用意されたシグネチャとパケットの中身をリアルタイムに照合し，不正を検出するもので，既知の攻撃の発見を目的としている．一方，アノマリ検出とはネットワーク・トラヒック，ログイン時刻，使用コマンドなどを判断材料として通常とは異なる振る舞いを検出するもので，未知の攻撃の発見を目的としている．

本稿では，ミスユース型IDSの出力ログを利用して，攻撃の予兆発見といったアノマリ検出を行う方法を考える．次にシステムへの実装に向けて，ユーザやシステムの振る舞いをモデル化し，実際のIDSのアラートに対して適用することによって，いくつかのデータマイニング手法の適用可能性を検討する。

2　従来研究

ネットワーク・トラヒックなどを用いたアノマリ検出は，現在のところ研究の段階であり，商用IDSのほとんどはミスユース検出機能のみを提供している．一方，ミスユース検出型IDSでは，攻撃を確実に検知するために敏感に設定されている．そのため，

(1) 監視対象が広域となり，管理するIDSの数が増えた場合，出力されるアラートの種類と数が増加し，人間による処理が困難になる．
(2) 実害がないとされるアラートはフィルタリングすることにより対応しているが，これらのアラートにも重要な情報（兆候など）が含まれている可能性がある．しかし，それを検出する方法が確立していない．
(3) 通知されたアラートに場当たり的に対応しており，「予兆があって発生したのか」あるいは「突発的に発生したのか」といった判断ができていない．

といった理由から，ミスユース検出型IDSの出力ログをデータ・マイニング手法によって解析し，

(1) 予め攻撃者やシステムの振る舞いを把握する．
(2) 攻撃の動向，兆候を把握する．
(3) 攻撃による障害が発生してから即座にその原因を探索する．

といったアノマリ検出を効率的に実施することが求められている．しかし，データ・マイニング手法の活用は試行的な段階に留まっており，通常の振る舞い (定常状態) を表現したり，データ・マイニング手法の適用可能性に関する体系的な検討がなされているわけではない．

3　アノマリ検出の方法

本研究では，以下の手順によってアノマリ検出を行う．

(1) IDSのアラート発生状態，攻撃者の行動をモデル化する．
(2) IDSの出力ログをモデルに当てはめ定常状態を定義する．
(3) 定常状態からの外れた状態と判断する基準を設定する．
(4) 基準から外れた状態と判断された場合，異常な攻撃を受けているものとしてワーニングを出す．

なお，このとき用いるIDSの出力ログは何等かの不正な操作や接続を試みたユーザの記録であるので，本研究ではこれらのユーザを攻撃者と呼ぶこととする．また，インターネットに接続しているシステムは，定常状態においても一定量の攻撃に晒されていることを前提とする．

3.1 IDSのアラート発生状態モデル

時間帯別アラート発生回数よりIDSの定常的なアラート発生状態を説明するモデルを考える．

いま，日付を $d = 1, 2, \cdots$，時間帯を $t = 0, 1, \cdots, 23$，時間帯別アラート・ログ発生回数を $\mathbf{x} = (x_{d0}, x_{d1}, \cdots, x_{dt}, \cdots, x_{d,23})$ とする．このとき，時間帯を変量として主成分分析を行い，時間帯に関する集約指標を作成する．次に各日付に対応する曜日を目的変数，時間帯を説明変数として正準判別分析を行う．

3.2 攻撃者の行動モデル

本研究では，攻撃者の行動をモデル化するにあたり，分析の対象期間内に同一IPアドレスから発生した一連のアラートの背後にある行動を長期的攻撃，一定期間内に同一IPアドレスから発生したアラートの背後にある行動を短期的攻撃と考える．

回数・構成比・組み合せの有無によるモデル　長期的攻撃を攻撃元IPアドレス単位のアラート別発生回数，アラート構成比，アラートの有無 (組み合せ) のベクトルで表現し，いくつかのタイプに分類する．

いま，攻撃元IPアドレスを $i = 1, 2, \cdots, n$，アラートを $j = 1, 2, \cdots, m$，アラート別発生回数を $\mathbf{x}_i = (x_{i1}, x_{i2}, \cdots, x_{ij}, \cdots, x_{im})$ とする．このとき，$\mathbf{x}_i$ を ℓ 個 $(k = 1, 2, \cdots, \ell)$ のグループに分類し，その平均値 $\bar{\mathbf{x}}^{(k)}$ を導き出す．アラート構成比 $\mathbf{y}_i$，発生したアラートを要素とする集合 (組み合せ) Z_i の特性関数 $\phi(Z_i) = (z_{i1}, z_{i2}, \cdots, z_{ij}, \cdots, z_{im})$ が与えられたときも同様の分類を行い，$\bar{\mathbf{y}}^{(k)}$, $\overline{\phi(Z)}^{(k)}$ を求める．ただし，特性関数 $\phi(Z)$ はアラート j が発生したとき，第 j 番目の要素が1となり，発生しなかったときは0となる．

この分類を用いて，予めユーザーやシステムの振る舞いを把握したり，攻撃による障害が発生してから即座にその原因を探索することになる．なお，分類にあたっては自己組織化マップ，K-means法を用いる．

発生パターンによるモデル　ある組み合せや順序のアラートを発生させる長期的，短期的攻撃があったとき，次にどのようなアラートを発生させるかというアラート・パターンをモデル化する．

いま，発生したアラートの組み合せ $\phi(Z)$，アラートの順序 $J = (j_1, j_2, \cdots, j_s, \cdots)$ が与えられたとき，Z あるいは J のもとでアラート j が発生するというモデルは順序なしあるいは順序付きのアソシエーション・ルールとして表現できる．したがって，このときの確信度 (confidence) は条件付き確率 $\Pr(j|Z)$, $\Pr(j|J)$ として表わされる．

適用にあたってはIDSのデータから確信度が一定の閾値を越えたアソシエーション・ルールを事前に抽出し，実際にアラートが発生したとき，アソシエーション・ルールのデータベースに照合をかけることにより，攻撃の予兆をリアル・タイムに察知する．

4 適用例

4.1 適用データ

本研究では試験的に収集したIDSの出力ログを用いる．なお，このデータの詳細については以下の通りである．

期間：　2001/08/14～2002/01/25
アラート件数：　249,553件 (84種類)
攻撃元IP アドレス：4,033
データ項目：時間/攻撃元IP/攻撃先IP/攻撃元ポート/攻撃先ポート/攻撃元mac/攻撃先mac/アラート名/クラス/プライオリティ

4.2 データ・マイニング手法の適用可能性

適用結果及び本研究で用いたデータ・マイニング手法の適用可能性の検討については，当日報告する．

5 おわりに

本研究では，アノマリ検出を目的として，IDSのアラート発生状態，攻撃者の行動を表現するモデルを提案した．また，実際のミスユース型IDSの出力ログに適用することにより，IDSの異アノマリ検出機能におけるデータ・マイニング手法の適用可能性を検討した．今後の研究課題としては提案モデルのアノマリ検出能力やシステムへの実装可能性の検討，複数のファーイアウォールやIDSのアラート・ログを統合したモデルの構築などが挙げられる．

参考文献

[1] 武田 圭史, 磯崎 宏,「ネットワーク侵入検知」, ソフトバンクパブリッシング, (2002).