日本オペレーションズ・リサーチ学会
2004年春季研究発表会

データマイニングを用いたIDSログ情報のネットワーク監視業務への活用

㈱ＮＴＴデータ＊山中　啓之　YAMANAKA Hiroyuki
㈱ＮＴＴデータ　小堀　誠　KOBORI Makoto
01404540　㈱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や情報漏洩など、情報システムの脆弱性を狙った攻撃、いわゆるサイバー攻撃が危惧されている。その対策の一つとして、不正アクセスを検知する目的で侵入検知システム（IDS）の導入が進んでいる。

IDSからは常に大量のログが出力されるため、管理・監視の現場では重要度の高いログのみを通知するようフィルタリングしている。しかし、通知されなかったログの中にも、ネットワーク攻撃の兆候が潜んでいる可能性があるため、その有効活用が求められている。

ネットワーク攻撃の兆候や未知の攻撃の対策として、異常検出型IDSの研究が行われているが、未だ試行的な段階であり、決定的なツール・システムは現れていない[1][2]。

本稿では、不正検出型IDSログを利用した異常検出を行い、その結果を元に監視対象システムの「ネットワークコンディション」を定義し、監視業務に活用するフレームワークを提案する。

2. 提案する方式

IDSの方式は、不正検出(misuse detection)型と異常検出(anomaly detection)型に分類される。不正検出型とは、予め用意されたシグネチャと照合し攻撃を検知するもので、既知の攻撃を高速に検知するものである。一方、異常検出型のIDSは、ネットワークトラヒックやログイン時刻などから、通常と異なる振る舞いを検出するもので、未知の攻撃を検出できる可能性がある。現場に広く普及しているのは、誤検出率の問題や速度の点から、不正検出型IDSである。

本稿では、不正検出型IDSのログに対して、統計手法・データマイニング手法を適用することで、管理者のシステム状態の把握に活用するフレームワークを提案する。

以下に処理概要を示す。

①ネットワーク上に配備された複数のIDSの情報を一元集約して管理する。

②データマイニングを用いて蓄積されたIDSログから、アラートの発生状態、攻撃者の行動をモデル化し、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リアルタイムで取得された攻撃者の振る舞いをモデルに当てはめ、基準から外れた場合、異常な攻撃を受けているものと判断する。

⑤ネットワーク全体、各サイトにて検出された外れ値の数により「ネットワークコンディション」を定める。

⑥「ネットワークコンディション」にしたがって、システム全体の状態を把握し、監視レベルを変更する。

3. システム概要

3.1 全体構成

図1にシステム全体構成を示す。

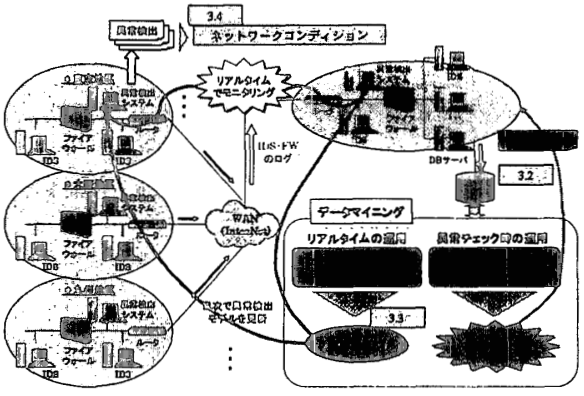

図1　システム構成

3.2 ログの収集・一元管理

監視対象のIDSで生成されたログを一元管理する。SNMP経由でファイアウォール、ルータなどの機器情報も管理可能である。

一元管理されたログは、「ネットワークコンディション」の定義以外にも、

・予め攻撃者の振る舞いを把握し、注意すべき攻撃者のリスト（ブラックリスト）を作成しておく。

・攻撃による障害発生時にその原因追求を行う。ことに利用できる。

3.3 予兆検出モデル

情報システムは、常に外部・内部より不正な操作や接続といった一定量の攻撃に晒されていることを前提とする。このような定常状態を統計手法・データマイニング手法によってモデル化する[3]。

本研究で、適用した手法の概要を以下に示す。

(1) アラート・ログ発生状態モデル

時間帯別ログ発生回数より、定常的なアラート発生状態を説明するモデルを作成する。

主成分分析・正準判別分析を行うことで時間帯に関する総合指標を作成する。

(2) 攻撃者の行動モデル

ある組み合わせや順序のアラートを発生させる長期的、短期的な攻撃があったとき、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。

分析の対象期間内に同一攻撃者から発生した一連のログの背後にある行動を長期的行動、一定の期間内に同一攻撃者から発生した一連のログの背後にある行動を短期的行動と考える。順序なしあるいは、順序付きアソシエーション・ルールとして表現でき、一定の確信度を閾値としてルールを作成する。

3.4 ネットワークコンディション

実際にアラート・ログが発生した場合に、作成されたモデルと照合し、外れ値かどうか判断する。

監視対象のサイト $i = 1,2,3, \ldots, n$ ，モデル種類（アラート・ログ発生モデル、攻撃者の行動モデル･･･など）を $j = 1,2,3, \ldots, m$ としたとき、一定時間内におけるサイト別外れ値検出回数を $x_i = (x_{i1}, x_{i2}, \ldots, x_{im})$とする。このとき、各サイトの「ネットワークコンディション」をベクトルの大きさ $|x_i|$ で定義する。また、ネットワーク全体の状態は、$\Sigma |x_i| / n$で定義される。

このベクトルの方向は、監視対象のネットワークへの攻撃の傾向を示し、ベクトルの大きさは、攻撃の規模を表している。監視者は、各サイト、ネットワーク全体の「ネットワークコンディション」を知ることで、情報システム全体に対する攻撃（図２）であるのか、特定のサイトに対する攻撃（図３）であるのかを判断することができる。

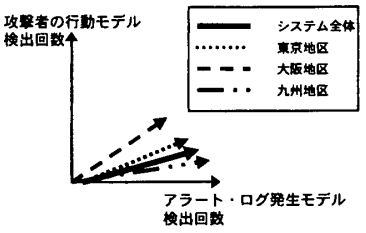

図２　システム全体への攻撃イメージ

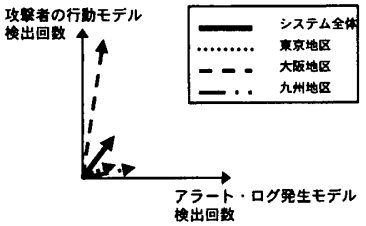

図３　特定のサイトへの攻撃イメージ

3.5 対策の実施

「ネットワークコンディション」が悪化した場合 IDS のシグネチャやファイアウォールの設定変更を行う。また、特定のサイトに集中している場合、そのサイトの監視体制を強化する運用を行う。

4. おわりに

本稿では、情報システムに配備された複数の IDS のログを一元管理し分析することにより、ネットワークの監視業務を支援するフレームワークを提案した。

問題点としては、本稿で提案したモデルが異なるシステム・異なる時期での攻撃状態を十分表現できるかということや、アソシエーション・ルールでのモデル作成条件をどのように設定すべきか明確な指標がないことがある。

また、IDS のみでなくファイアウォール，MIB 情報，sys ログなどを統合したモデルを構築し、ネットワークコンディションを定義することも今後の課題として挙げられる。

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，（2000）

[2]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 情報処理学会， pp.1169-1174

[3]山中　啓之，小堀　誠，中川　慶一郎 “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”，2003 年度日本 OR 学会秋季研究発表会