3F-6

侵入検知システムにおけるシグニチャ自動生成方法の検討

時庭康久† 鈴木清彦† 原田道明† 後沢忍†
三菱電機(株)情報技術総合研究所†

1．はじめに

未知ワームの検知対策技術として、ネットワーク通信トラヒックの観測結果から異常を検知し、侵入検知システム(IDS)のシグニチャを自動生成する方法が文献[1]等で研究されている。我々はアウトブレーク型の未知ワーム発生を想定しシグニチャを自動生成、配布、隔離を行うシステムの構築を目指し、試作評価を進めている。検討した結果を記述する。

2．研究の狙い

主なネットワーク感染型のワームは、ネットワークを介し端末を感染させ、感染パケットを送信し増殖する。感染が判明してから人手で端末を隔離したのでは被害を防げない場合があり、駆除に作業時間と余計なコストを要する。

これらの問題を解決するために以下の侵入検知システムの構築を目指した。

・イントラネット上に配置した複数の異常検知装置がリアルタイムに動作し、IDS に適用するシグニチャの候補を生成する。
・生成したシグニチャ候補の中から採用するシグニチャを決定し、各異常検知装置へ自律的に配布する。
・異常検知装置でワームを検知して防御する。

3．システム構成と動作概要

(1)システム構成

以下の装置構成で IDS 装置に適用するシグニチャの候補を自動生成する。

・異常検知装置 (IDS 装置)
ネットワーク上のトラヒックから IDS のシグニチャの候補を生成する装置。またシグニチャによりトラヒックを精査し回線上で不正な通信トラヒックを遮断する IDS 装置。
・管理装置
ワーム感染の判断を行う分析ソフトウェアが動作するサーバであり、各異常検知装置と端末からの情報を分析し精度の高いシグニチャを配布する。
・端末
端末内蔵のソフトウェアにより端末挙動情報を収集し管理装置へ提供する。

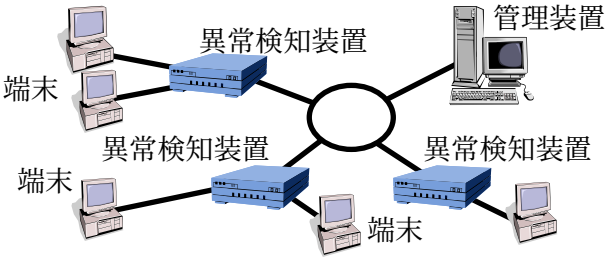

図 3-1. ネットワーク構成

(2)動作の説明

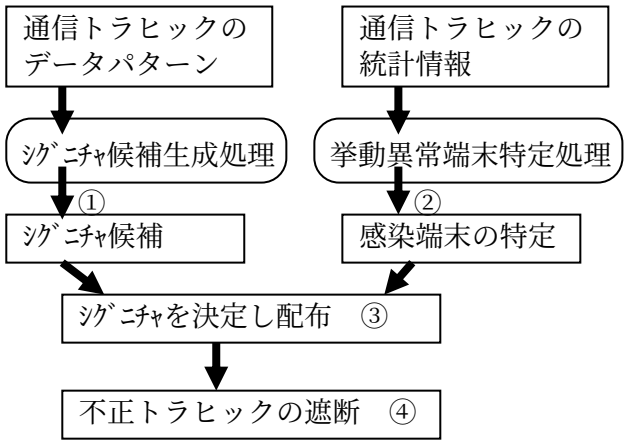

図 3-2 システムの動作の流れ

①異常検知装置は、通信トラヒックから文字列の出現頻度を計測し、ワーム固有の特徴文字列(シグニチャ候補)を抽出する。
②管理装置は、端末内蔵ソフトウェアから送受信カウンタなどの統計情報を収集し、ワーム感染挙動を示す端末を特定する。文献[2]の技術を適用し実現する。
③管理装置は、上記シグニチャ候補に対して、ワーム感染挙動を示した端末の情報等から正確なシグニチャを生成し、各異常検知装置へ配布する。
④各異常検知装置は、配布されたシグニチャを適用し、不正トラヒック(未知ワーム)を遮断する。

4．シグニチャ候補生成処理の実現方式

以下の条件を満足する文字列パターンを候補とする。

A Study of signature construction on Intrusion Detection & Prevention System
Yasuhisa TOKINIWA†, Kiyohiko SUZUKI†, Michiaki HARADA†, Shinobu USHIROZAWA†
†Information Technology R&D Center, Mitsubishi Electric Corporation 5-1-1 Ofuna, Kamakura, 247-8501 Japan

・通信トラヒックで出現頻度が高い文字列であること。
・文字列を含むデータを送受信している端末の IP アドレスの分散度(送受信端末数／単位時間)が多いこと。

さらにシグニチャ候補生成の精度向上と生成処理の負荷の軽減のため入出力フィルタリング処理を採用した。

４．１ シグニチャ候補生成処理の流れ

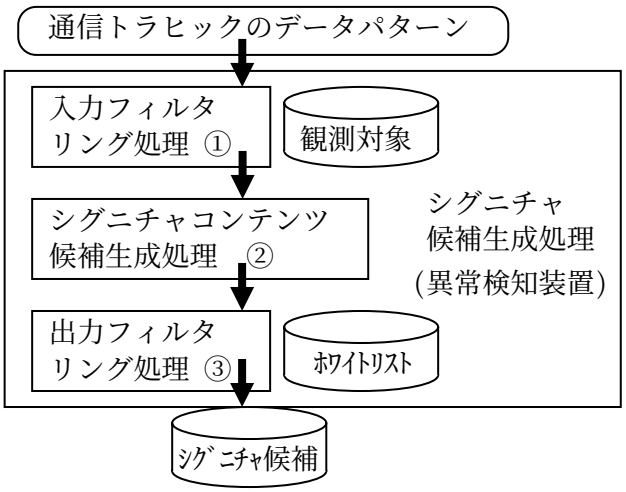

図 4-1 シグニチャ候補生成の処理の流れ

①入力フィルタリング処理はシグニチャ自動生成で適用する通信トラヒックを絞り込む。
②シグニチャコンテンツ候補生成処理を実行する。
③シグニチャの出力候補に対して出力フィルタリング処理を施し、精度を向上させる。

４．２ 入力フィルタリング処理

プロトコルの状態チェック(SPI：Statefull Packet Inspection)と組み合わせることにより、データ(文字列)や TCP/UDP のポート番号以外に例えば、「HTTP ヘッダ内」や「リクエスト URI 内」などの検知範囲に絞り込むことによって検知精度を向上させる。

後述の hash 値の計算では、プロトコルのヘッダ部とペイロード部に分けてそれぞれ行う。プロトコルのペイロード部の hash 値を間引いた値にプロトコルのヘッダ部(プロトコル番号、宛先ポート番号ッダ部(プロトコル番号、宛先ポート番号、通信方向)の hash 値を XOR した値を採用し、精度を上げるようにした。

４．３ シグニチャコンテンツ候補生成処理

文献[1]の多項式をベースとした hash 関数（Rabin fingerprint）を採用し実現した。(図 4-2 参照)

(1)ユーザ指定の固定長ウィンドウをスライドさせながら文字列の hash 値を計算する。
(2)処理量・必要メモリ量を効率化するためにユーザ指定のマスク値に一致する hash 値のみに間引くことにより、出現頻度を絞り込む。
(3)上記の頻出パターンに関して、{送信元／宛先 IP アドレス、文字列(コンテンツ) }エントリが何度出現したかを、送信元 IP アドレスと宛先 IP アドレスのそれぞれを hash 値により近似計測し(アドレス空間を小さくし)カウントし、しきい値を超えた場合、当該文字列(コンテンツ)をシグニチャ候補と判定した

４．４ 出力フィルタリング処理

誤った候補を除外するためにホワイトリストを導入し、IP アドレス、プロトコル、ポート番号、(コンテンツに含まれる)文字列の条件に一致する候補を出力候補から除外する。以前に誤検知となったシグニチャの文字列を除外する。

５．まとめ

未知ワームに対するシグニチャ自動生成方法について検討した結果を述べた。現在、本方式の効果を実証すべく評価を実施中である。

参考文献

[1] Sumeet Singh,Cristian Estan, George Varghese and Stefan Savage,“Automated Worm Fingerprinting”,6TH SYMPOSIUM ON OPERATING SYSTEMS DESIGN & IMPLEMENTATION (OSDI'04)
[2]北澤繁樹、樋口毅、原田道明、藤井誠司、“マハラノビス距離を用いた判別分析による未知ワーム感染挙動特定方式”、情報処理学会マルチメディア，分散，協調とモバイル（DICOMO 2006）シンポジウム

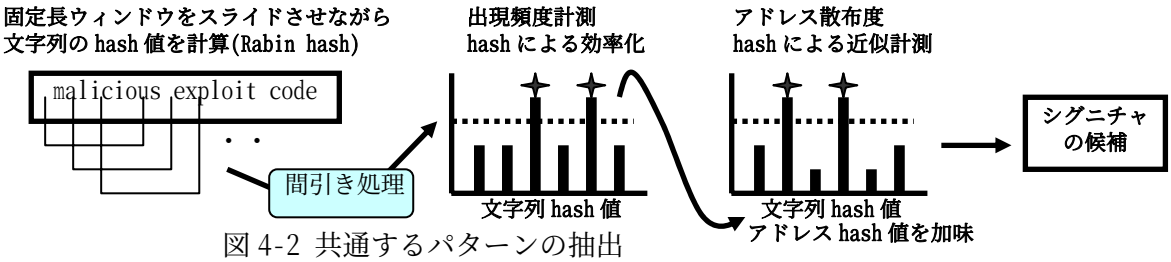

図 4-2 共通するパターンの抽出