4X-05

機械学習を利用した通信異常検知システムの提案

石田 時大† 喜多 義弘† 朴 美娘† 岡崎 直宣‡
神奈川工科大学† 宮崎大学‡

1. はじめに

近年，一般ユーザがマルウェアに感染し，踏み台として利用され，ユーザが気が付かないうちに攻撃を行ってしまい，意図せず攻撃者となってしまう事案が多く発生している．

このようなマルウェアに対して，一般ユーザが行える対策としてはウイルス対策ソフトやファイヤーウォールの導入がある．それに対し，サーバを保有している企業や組織などはそれらに加え，侵入検知システム（IDS：Intrusion Detection System）や侵入防止システム（IPS：Intrusion Prevention System）を導入している．一般ユーザは異常な通信に対して対処することが難しい現状がある．

このような攻撃への対策として，機械学習を用いた不正アクセス予測に関する研究が行われている[1]．しかし，検知率が基準を達していないことや，個人ユーザを対象としていない問題がある．

本研究では，一般ユーザの保護を目的として，機械学習にユーザの通信を予め学習させ，その分類から異常を検知するシステムの構築と，一般ユーザであっても判断がしやすいインターフェイスを提案する．

2. 従来研究

シグネチャ型 IDS である Snort[2]のルールファイルを自己組織化マップで学習させる．ここでは，既存の攻撃に類似している不正アクセスを予測している[1]．この手法では現存するアクセスの亜種や類似とみなせるアクセスを 80～90%程度で予測することが可能である．しかし，Snort のルールファイルに登録されているシグネチャを利用するため，学習から大きく離れるような未知な攻撃の場合に攻撃を検知することは難しいと問題がある．

A Proposal of Communication Anomaly Detection System using Machine Learning
†Tokihiro Ishida, Yoshihiro Kita, Mirang Park, and ‡Naonobu Okazaki , †Kanagawa Institute of Technology
‡University of Miyazaki

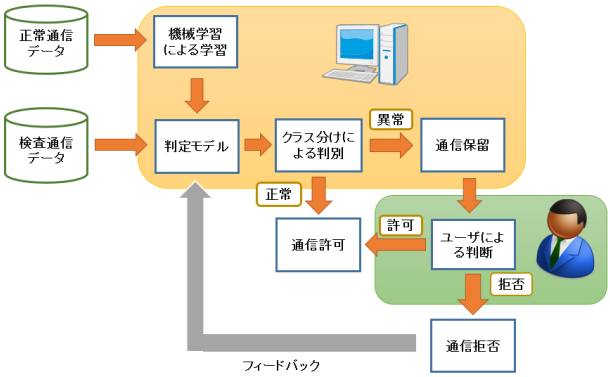

図 1 機械学習を利用した異常検知システム

3. SVM を利用した異常検知システム

本研究では，従来研究で課題とされていた未知の攻撃への対処を目的とする．そこで，異常検知システムの問題点である検知精度の向上のためサポートベクターマシン（SVM：Support Vector Machine）[3]を利用する．

また，一般ユーザの考慮し，検知支援を行うためのユーザビリティの向上を目指す．そこで，異常と判断された際にどのような通信であるかをわかりやすく告知し，通信を一時的に遮断し使用に問題が無いかをユーザに確認を行う．一定時間通信を保留することにより通常作業に支障が出ない場合，異常通信である可能性があるので通信を遮断する．

3.1 提案手法の流れ

提案手法は図 1 のような学習フェーズと判定フェーズによって構成される．

学習フェーズ

通常の通信を SVM により学習し，判定モデルを生成する．

(1) 通常通信のキャプチャを行う．
(2) 通信データのベクトル化を行う．通信データの特徴量として，通信発生時間，送信元 IP アドレス，送信先 IP アドレス，送信元ポー

ト，送信先ポート，プロトコル，通信量を利用する．
(3) ベクトル化した通信データを機械学習に入力することで，学習を行う．学習を繰り返し行うことにより判定モデルの最適化を行う．
(4) 通常通信が学習された判定モデルを生成する．

SVM を利用する際、通信データのパラメータとして属性ベクトル化をする必要がある．ベクトル化とは，データを数値化することである．数値化することで機械学習に入力することが出来る．

判定フェーズ

異常通信判定を学習フェーズで生成された判定モデルを利用して行う．

(1) 検査通信データのキャプチャを行い，検査通信データのベクトル化を行い判定モデルに入力する．
(2) 判定モデルのクラス分けにより，異常であるか正常を判別する．
(3) クラス判別を行う．
　a) 正常と判断された場合，要求されたパケットの通信を許可する．
　b) 異常と判断された通信を一定時間保留する．設定された一定時間の間に作業に支障が出るか出ないかをユーザによる判断をする．
(4) 保留時間が経過した後，ユーザに通信の通信許可を判断させる．
　a) 作業に支障が出た場合，正常通信であると判断し，通信を許可する．
　b) 作業に支障が出なかった場合，ユーザが必要としている通信でないと判断出来る．通信を求めているソフトウェアのパケットを破棄する．
(5) ユーザの判別を判定モデルにフィードバックする．

3.2 通信情報の可視化

本システムでは，一般ユーザが通信の許可や，拒否を下せる必要な情報を提示するインターフェイスを開発することを目的としている．必要な情報のみを表示することで，ユーザの判断をしやすくする．

図 2 のような，通信を行おうとしているソフトフェア名を表示し，詳細情報は詳細ボタンによって表示する．また，どのようなソフトウェアであるか判断出来ない場合に，ソフトウェア名をインターネットから検索することで，ソフトウェアの情報を取得することが出来る．

図 2　インターフェイス例

インターネット上に公開されている他のユーザによる情報を利用する事で通信を求めているソフトウェアの詳細を知ることが出来る．

4. 評価

本研究では，アノマリ検知を利用することでシグネチャを利用するミスユース検知で検知することが難しいとされた未知への攻撃を防ぐことが出来ると考えられる．かつ，異常検知方法に SVM を利用することで，従来よりも高い判別性能を保有出来ていると考えられる．

また，本研究で通常の通信から異常を検知し，利用対象である一般ユーザの保護を目的として，必要な情報のみを提示することに加え，分かりやすいインターフェイスを提案している．これにより高度な知識を持たない一般ユーザが攻撃の標的にされることを防ぐことが出来ると考えられる．

5. おわりに

本研究では，一般ユーザの保護を目的として，機械学習にユーザの通信を予め学習させ，その分類から異常を検知するシステムの構築と，一般ユーザであっても判断がしやすいインターフェイスを提案した．

今後の課題として，本提案手法の有効性を確認するために，SVMを利用した異常検知システムを実装し，既存手法との比較によって，異常検知の性能について検証する．

参考文献

[1]中山亮介，納富一宏，斎藤恵一，“自己組織化マップを用いた不正アクセス検知”，第8回情報科学技術フォーラム，pp.131-132，2009.
[2]“Snort”，Sourcefire，http://www.snort.org/
[3]Nello Cristianini，John Shawe-Taylor，“サポートベクターマシン入門”，共立出版，2005.