4Z-2

ランダムフォレストアルゴリズムを用いたネットワーク侵入検出システムの性能解析*

小池 泰輔† 梅澤 猛‡ 大澤 範高‡

千葉大学工学部情報画像学科† 千葉大学大学院融合科学研究科‡

1 はじめに

コンピュータシステムへの攻撃に対処する方法の1つとして，侵入検知システム (Intrusion Detection System:IDS) が提案されている．侵入検知システムとはネットワーク上を流れるパケットやホスト上のプロセスを監視することで不正や異常を検知するシステムである．

侵入検知手法として主流のシグネチャ検知は，事前に不正アクセスや攻撃のパターンを定義し，通信内容との比較を行うことで侵入を検知する．シグネチャ検知は定義された攻撃に対しては高い検知率を示すが，未知攻撃や亜種攻撃に対する検知率が低いという課題がある．特に亜種攻撃は既存の攻撃を部分的に変更するだけで作成可能であり，攻撃者にとって実行が容易であるため脅威となり易い．

そこで本研究では，機械学習を応用することで亜種攻撃に対応可能なシグネチャ検知型の侵入検知システムを提案する．既知攻撃の特徴を解析し，攻撃検知に有用な説明変数を抽出することで，亜種攻撃検知のためのモデル作成を図る．攻撃検知モデルの構築にあたっては，実際の侵入検知システムの動作を模した侵入検知用のベンチマークデータ解析し，ランダムフォレストによる機械学習を行った．また，得られたモデルを別のベンチマークデータに適用し，検知率の変化を測定した．

2 シグネチャ検知型 IDS

シグネチャ検知はパターンマッチングによる攻撃検知手法を指す．シグネチャとは攻撃が持つ固有の文字列やビット列を指し，パケットのペイロード部分と比較することで攻撃を検知することができる．

2.1 問題点

シグネチャ検知の主な問題点は未知・亜種攻撃への検知率の低さである．対策としては登録シグネチャを増やすことがあげられるが，通常の事象を不正と誤検知する False-Positive が増加する問題がある．誤検知が多発し警告ログが増加すると，管理者がログ解析を行う負担が増大し，実際の攻撃を警告するログを見逃す危険を生じる．

2.2 関連研究

既存研究 [2] では，予め記録されたシグネチャから類似シグネチャを作成することで未知攻撃を検出する手法を提案している．しかし既知のシグネチャに複数の類似したオプションを付加することで新たなシグネチャを作成する手法であるため，冗長なシグネチャが作成される可能性がある．それにより攻撃検知に有効ではないシグネチャが生成され，False positive が増加する危険性が生じる．そのためシグネチャの管理・調整が困難になり，侵入検知システムの性能低下を招く恐れがある．

3 提案手法

亜種攻撃は既存攻撃を部分的に変更したものでありながら，シグネチャ検知による検知が困難である問題がある．そこで本研究では機械学習を応用した侵入検知システムを提案する．教師情報として，攻撃の特徴を用いる．既知攻撃の特徴を解析し，攻撃検知に影響の低い特徴を削除することで，重要な特徴を共有する亜種攻撃への検知率向上を目指す．

機械学習アルゴリズムにはランダムフォレスト [1] を用いた．他の学習アルゴリズムでは，与えられた学習のパターンのみに特化したモデルを構築し，新たな未知のパターンに対して正しい結果を出力できない可能性がある．これを過学習と呼び，モデルの性能低下を招く恐れがある．ランダムフォレストでは高い汎化性能を得られるため過学習を防いで，精度の高い攻撃検知モデルの構築が可能となる．

提案システムの概要を図1に示す．学習データに含まれる T は目的変数を，$X_1 ... X_N$ は説明変数を意味する．説明・目的変数を持った学習データをランダムフォレストを用いて解析し，攻撃検知モデルを構築する．モデルに評価データを与えることで，対応する結果を出力する．

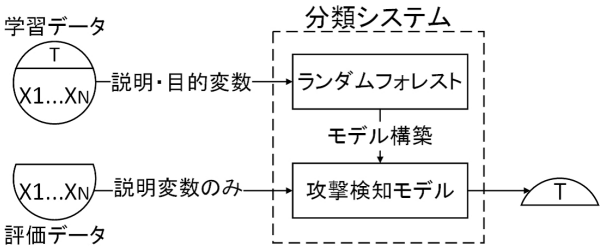

図 1: 提案システムの概要

学習フェーズでは，あらかじめ攻撃の分析が行われたトラフィックデータを学習データとし，ランダムフォレストを用いて攻撃検知モデルを構築する．

また，学習結果をもとに各説明変数の重要度を算出し，攻撃検知に必要性が低い説明変数を除いてモデルの再構築を行うことで，亜種攻撃に対する検知率向上を図る．検出フェーズでは，攻撃検知モデルによってトラフィックデータを処理することで，学習データに

*Performance Analysis of Network Intrusion Detection System using Random Forest Algorithms

†Daisuke Koike,Department of Informatics and Imaging Systems, Faculty of Engineering, Chiba University

‡Takeshi Umezawa, Noritaka Osawa, Graduate School of Advanced Integration Science, Chiba University

含まれる既知の攻撃だけでなく，既知攻撃に類似した亜種攻撃も検出することができる．

4 実験

学習および検出に用いるトラフィックデータには，KDDCup1999[3] を用いた．使用したデータセットは，8 種類の攻撃データを含み，各データには 41 次元の属性が付加されている．亜種攻撃の検知率を測定するため，学習データに含まれる 8 種類の攻撃から 1 種類を抜き出し，評価データに加えた．残る 7 種類の攻撃情報をモデル構築に用いるデータとし，作成されたモデルによって，取り除いた 1 種類の攻撃を検知できるかを確かめた．決定木 100 本でモデル構築を行い，説明変数にはデータに含まれる 41 次元の属性全てを用いた．データには 8 種類の攻撃が含まれるので，学習データと評価データの対は 8 組となる．

また，既知攻撃の特徴選択により，亜種攻撃の検知率に変化が見られるか調査した．攻撃検知に有用な説明変数の抽出を行うため，8 種類全ての攻撃を含む学習データからモデルを構築し，説明変数の重要度を算出した．算出結果より攻撃検知への重要度が 0.010 以下の説明変数を削除し，残りを有用な説明変数として抽出した．抽出した説明変数の有用性を確認するため，モデルを再構築し，各亜種攻撃への検知率を評価した

8 種類全ての攻撃を含む学習データから抽出した攻撃検知に用いられる説明変数の重要度を図 2 に示す．縦軸はモデル構築に用いた説明変数を，横軸は変数の重要度 (ジニ係数) を示しており，上から順に重要度の高い説明変数を意味している．

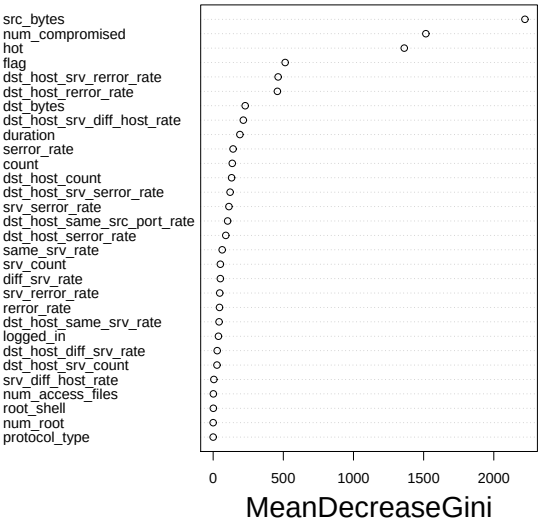

図 2: 説明変数の重要度

図 2 より攻撃検知に影響の少ない説明変数 12 個を削除し，残りの説明変数のみを用いて構築したモデルと 41 個全ての説明変数を用いて構築したモデルの精度結果を図 3 に示す．

図 3 では，TP_A,FP_A は全ての説明変数を，TP_B,FP_B は抽出された説明変数を用いたモデルの検知率を示す．TP(True-Positive) は，データセットに含まれる不正パケットのうち，検知に成功したものの割合を示す．FP(False-Positive) は，攻撃とは無関係の通常パケットを，不正パケットとして誤検知したもの

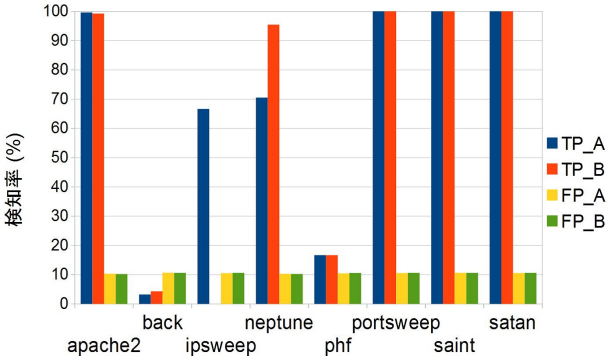

図 3: 検知精度評価

の割合を示す．説明変数選択により False-Positive の割合には大きな変化は見られなかった．攻撃 neptune に関しては検知数が 201 個から 272 個となり，検知率が 71%から 96%へ向上するなど，攻撃検知の性能には大きな変化が見られた．しかし攻撃 back に関しては検知数が 107 個から 143 個となり，検知率は 3.2%から 4.3%に向上したが，元々の検知率が不十分であり改善が必要であることがわかった．

5 考察

説明変数選択よって，back,neptune といった Dos 攻撃に対する検知率向上がみられた．しかし攻撃 back に関しては，向上後も検知率が低く，実用的な値とするためには更に対応が必要である．また ipsweep(攻撃数 3) のように攻撃数が少ない場合，攻撃検知数が 2 個から 0 個に減少した．これは説明変数を除去することによって，攻撃を抽出するための情報が減ってしまったためであると考えられる．

6 まとめ

機械学習の 1 種であるランダムフォレストを用いて既知攻撃の特徴を学習し，攻撃検知に有用な説明変数の抽出・選択を行うことで亜種攻撃への検知率向上を目指した．本研究では KDDCup1999 による攻撃情報の解析を行い，攻撃検知モデルの構築と評価を行った．その結果，説明変数選択によって検知精度に影響を与えることがわかった．今後は，攻撃の関係を分類し，お互いに亜種の関係にある攻撃のみを用いて亜種攻撃を検知可能かの評価と，実環境のトラフィックデータを用いた場合の性能解析について検討を行う予定である．

参考文献

[1] Breiman, L.: Random Forests, Machine Learning, Vol.45, No.1, pp.5-32 (2001).

[2] 川崎孝弘，服部峻，久保村千明，亀田弘之：シグネチャ型侵入検知システムにおける類似シグネチャの自動生成による未知攻撃検出手法，情報処理学会第 74 回全国大会講演論文集 Vol.2012， No.1，pp.591-593（2012）．

[3] UCI KDD Archive: KDD Cup 1999 Data, UCI (online), available from <http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html> (accessed 2014-01-14).