L-004

ネットワークトラフィックのエントロピーに注目した異常検知

Anomaly Detection Focused on Network Traffic Entropy

柏木 宏務 *　青木 茂樹 *　宮本 貴朗 *
Hiromu Kashiwagi　Shigeki Aoki　Takao Miyamoto

1　はじめに

近年, インターネットの普及に伴ってサイバー攻撃が増加している. サイバー攻撃を防ぐために異常な通信を検知する侵入検知システム (IDS: Intrusion Detection System) の研究が盛んに行われている. IDS にはシグネチャ型 IDS とアノマリ型 IDS の 2 種類が存在する. シグネチャ型 IDS は異常な通信を予め定義して異常検知を行うため, ゼロデイ攻撃などを検知できない. そこで, 正常な通信のみを学習し, 正常な通信とは異なる通信を異常として検知することで, ゼロデイ攻撃を検知できるアノマリ型 IDS が注目されている. アノマリ型 IDS に関する手法として, ネットワークからのサイバー攻撃が発生した時のパケットの情報量に着目する手法 [1] が提案されている. 文献 [1] では, パケットのヘッダ情報から得られたエントロピーに基づいて異常を検出する EMMM 法（Entropy based Multi dimensional Mahalanobis distance Method）を提案している. この手法ではパケットを到達した順番に並べ, 一定のパケット数で分割する. そして, 分割したパケットから IP アドレスやポート番号など毎にパケット数を計測する. 次にパケットの出現確率を求め, 求めた確率からエントロピーを算出する. その後, エントロピーの変化が大きい時間を攻撃などが含まれている状態として検出している. この手法ではパケットのヘッダから得られる情報のみを利用しているため, ペイロードのみに特徴が現れる攻撃の検出が難しい.

本稿では, パケットから抽出した特徴量のエントロピーに注目したアノマリ型 IDS を提案する. まず, パケットのヘッダから特徴量を抽出する. また, ペイロードに攻撃固有の特徴が表れると仮定し, ペイロードの情報を 1Byte 単位で読み込んで 0〜255 の数値で表し, n-gram で分割して特徴量とする. 次に, 抽出した各特徴量の出現確率からエントロピーを算出して特徴ベクトルとする. その後, 生成した特徴ベクトルをクラスタリングし, 外れ値を検出することで異常を検知する. 本手法ではパケットのヘッダ情報だけでなく, ペイロードの情報を併用し, 更にエントロピーに着目することでペイロードに特徴の現れる攻撃の検出を目標としている. 実験では, CICIDS2017 データセット [2] を用いた実験により有効性を確認した.

* 大阪公立大学大学院情報学研究科 Graduate School of Informatics, Osaka Metropolitan University

表 1. 特徴量

送信元 IP アドレス	宛先 IP アドレス	送信元ポート番号
宛先ポート番号	パケットバイト数	TTL 値

2　提案手法

本手法は学習と異常検知の 2 つの処理に分かれる. 学習処理では, 学習用の正常なトラフィックデータのパケットヘッダとペイロードから特徴を抽出し, 抽出した特徴の出現確率からエントロピーを算出して特徴ベクトルとする. そして, 生成した特徴ベクトルをクラスタリングすることで, 正常な通信を複数のクラスタに分類する. 異常検知処理では, テスト用のトラフィックデータから同様の処理で特徴ベクトルを生成する. 生成したベクトルを正常な通信をクラスタリングした空間に投影し, 最も近いクラスタの中心からの距離が閾値以上であれば異常, 閾値より小さければ正常と判定する.

2.1　特徴ベクトルの抽出

ネットワークに対する攻撃を識別するために, ネットワークから収集したトラフィックデータを一定のパケット数 N で分割し, パケットのヘッダから各区間ごとに表 1 に示す 6 種類の特徴量を抽出し, その出現確率を算出する. また, ペイロードから抽出したバイナリデータを 1Byte 単位で読み込んで 0〜255 の数値で表して n-gram で分割し, 出現確率を算出する. 次に, 抽出した特徴量の出現確率を用いてエントロピーを算出する. 特徴量毎に求めるエントロピー h は, それぞれの特徴量で出現したシンボル i の出現回数から算出した出現確率 P_i により, 式 (1) で求められる. ここで m は, シンボルの種類数である.

$$h = -\sum_{i=1}^{m} P_i \log P_i \tag{1}$$

その後, 特徴量毎にエントロピーの総和を求める. ここで抽出したエントロピーの総和は, 特徴量ごとに大きく異なると考えられる. そのため各特徴量の値の平均が 0, 分散が 1 になるように標準化する. そして, 標準化した 7 次元のベクトルを特徴ベクトルとする.

なお, 区間内の全てのパケットでペイロードのサイズが 0 の場合, ペイロードのエントロピーが算出できないため, ペイロードの特徴量を 0 として特徴ベクトルを抽出している.

表 2. 実験データ

	正常	異常	総数
学習用	9752	0	9752
テスト用	5556	435	5991

表 3. 異常データ

Brute Force	XSS	Portscan	DDoS
197	65	5	168

2.2 特徴ベクトルのクラスタリング

抽出した特徴ベクトルをクラスタリングし, 学習用データのセッションを分類する. まず, 抽出したベクトルの次元を主成分分析法により圧縮する. ここでは, 累積寄与率 80 %以上となる最小の次元数を用いる. 次に, 次元圧縮したベクトルを Mean-Shift 法でクラスタリングする. Mean-Shift 法は, 事前にクラスタ数を決定する必要がないクラスタリング手法である. そのため, 事前に分類対象の種類数が判明していない場合に適した手法である. 本研究では, 正常通信にも複数の種類が存在すると考え, このクラスタリング手法を用いている.

2.3 異常検知

クラスタリングを行った学習用のトラフィックデータとは別のトラフィックデータから 2.1 節と同様の処理で特徴ベクトルを抽出する. その後, 2.2 節で作成した正常な通信をクラスタリングした空間に, 次元を圧縮して投影する. 次に, 各クラスタの中心と投影したベクトルとの距離を求め, 最も距離が近いクラスタを選択する. そして, 最も近いクラスタ中心との距離が閾値未満であれば正常, 閾値以上であれば異常と判定する. 閾値はクラスタ内の最も中心から離れたベクトルとの距離とする.

3 実験

3.1 実験概要

本手法の有効性を確認するために, CICIDS2017 データセット [2] を使用して実験を行った. 実験では正常な通信として Monday-WorkingHours, 異常な通信として Brute Force, XSS を含む Thursday-WorkingHours, DDoS, Portscan を含む Friday-WorkingHours を使用した. トラフィックデータを分割するパケット数 N は 500 パケットとした. そして, データセットで攻撃元と指定されている IP アドレスからのパケットが 50 パケット以上含まれている区間を, 異常区間とした. n-gram の n の値は 4 とし, 出現回数が 1 回の文字列は除外した. また, ペイロード特徴の有効性を確認するために, ペイロード特徴を使用した場合と使用しない場合での実験結果を比較した. 実験データの正常数・異常数・総数を表 2, 異常データの攻撃ごとの内訳を表 3 に示す.

3.2 実験結果

実験結果を表 4 に示す. 実験の結果, ペイロード特徴を使用しなかった場合と比較して, 正常通信では約 9.5 %, Brute Force では約 31.0 %, XSS では約 26.2 %の検知精度向上を達成し, ペイロード特徴を使用することの有効性を確認できた. しかし, 実験全体の精度としては正常通信を正常と検知した割合が 70.0 %, 異常通信を異常と検知した割合は 66.2 %と低い結果となった.

表 4. 実験結果

	正常	Brute Force	XSS	Portscan	DDoS
ペイロードあり	70.0%(3891)	78.7%(155)	55.4%(36)	100.0%(5)	54.8%(92)
ペイロードなし	60.5%(3359)	47.7%(94)	29.2%(19)	80.0%(4)	73.2%(123)

3.3 考察

Brute Force の検知精度が実験データが少ない Portscan を除いて一番高い 78.7 %となり, ペイロードの特徴量を含まない実験結果と比較して約 31.0 %高い結果が得られた. これは, Brute Force 攻撃では, ペイロードに類似するデータが多く含まれ, その特徴を上手くとらえることができたためであると考えられる.

XSS においても, ペイロード特徴を含まない実験と比べて, 約 26.2 %の検知精度向上がみられた. これは脆弱性のあるページにリダイレクトさせる XSS 攻撃の特徴が, 一部のパケットのペイロードに現れ, それらをとらえることができたためであると考えられる.

Portscan は検知精度 100 %であるが, ポートスキャンのパケットを 50 以上含む区間が少なく, 実験データが少ない状況での結果であるためペイロードの特徴が検知精度に影響を与えたかは判断できない. 他のデータセットなどを用いてポートスキャンのデータを増やして実験することが課題として挙げられる.

DDoS においては, 検知精度が 18.4 %下がる結果となった. 実験データを精査すると, ペイロードのサイズが 0 のパケットが多く含まれ, ペイロードのエントロピーを 0 とした特徴ベクトルが多い傾向があった. このようなデータが, 検知精度を低下させる要因になったと考えられる.

また, 正常通信を含む全体での検知精度が低いことについては, 異常区間の設定方法の検討や, 学習データを更に増加させた時の実験結果を確認し, 更に有効な特徴量を選択することで改善できると考えられる.

4 まとめ

本稿では, パケットのヘッダとペイロードから抽出した特徴量の出現確率から算出したエントロピーの特徴ベクトルをクラスタリングし, 外れ値を検出することで異常を検知する手法を提案した. ペイロード特徴を使用しなかった実験と比較し, 本手法の有効性を確認することができた. 一方, 本手法では全体の検知精度が低いため, 今後の課題としては異常区間の設定方法の検討や, 学習データを増やした実験の実施などが挙げられる.

参考文献

[1] 小島俊輔, 中嶋卓雄, 末吉敏則:エントロピーベースのマハラノビス距離による高速な異常検知手法, 情処学論, Vol52, No.2, pp.656-668(2011)

[2] I. Sharafaldin, et al.:Toward generating a new intrusion detection dataset and intrusion traffic characterization in ICISSP, pp.108–116(2018)