修士学位論文要約（令和 4 年 3 月）

# 教師なし学習を用いた低レート DoS 攻撃検知に関する研究

榎場 叶耀
指導教員：菅沼 拓夫

# A Study on Low-Rate DoS Attack Detection Using Unsupervised Learning

Kiyoaki KAYABA
Supervisor: Takuo SUGANUMA

In recent years, the damage caused by DoS attacks has been increasing year by year, and low-rate DoS attacks have been observed, in which communication continues for a long time with a small number of packets, occupying the session. While supervised learning-based detection methods have been proposed as a countermeasure for low-rate DoS attacks and have shown high detection performance, collecting attack data in a real environment is a challenge for system construction. In this study, we propose a new detection method using unsupervised learning that does not require attack data. In evaluation experiments, we have shown that AutoEncoder, an unsupervised learning algorithm, is effective in detecting low-rate DoS attacks.

## 1. 序論

近年，DoS 攻撃による被害は年々増加しており，少ないパケット数で長時間にわたり通信を続け，セッションを占有する低レート DoS 攻撃が観測されている．対策として，教師あり学習を用いた検知手法 [1] が提案されており，高い検知性能と複雑な設定を必要としない検知システムの構築を達成している．一方で，学習用データセットには攻撃データが必要となり，実環境における攻撃データの収集が課題となっている．

本研究では，攻撃データを必要としない教師なし学習を用いた新たな検知手法を提案する．既存のネットワークから収集したトラフィックデータを用いて検知モデルを構築し，定期的にトラフィックデータを検知システムに入力することで，低レート DoS 攻撃の有無を判定する．実験では，データセットによる性能評価を行い，提案手法が低レート DoS 攻撃検知に対して有効であることを示した．

## 2. 関連研究と課題

本研究の対象である Slow HTTP DoS 攻撃は，低レート DoS 攻撃の一種であり，少ないパケット数で長時間にわたり HTTP 通信の送受信を続けることによって Web サーバの TCP セッションを占有し，正規のユーザがアクセスできないように妨害する攻撃である．Slow HTTP DoS 攻撃の概要を図 1 に示す．一般的な大規模 DoS 攻撃は，検知を行う IDS や検知と防御を行う IPS によって対策を行うが，Slow HTTP DoS 攻撃は IDS/IPS による検知が困難であるという課題が挙げられる．

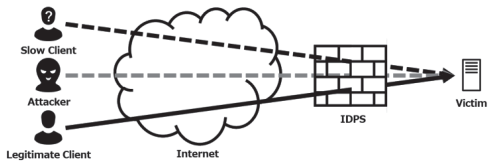


図 1: Slow HTTP DoS 攻撃の概要

低レート DoS 攻撃の検知手法として，教室あり学習を用いた検知手法 [1] が提案されている．教師あり学習を含む機械学習を用いた手法では，他のネットワークで収集したデータを使用して検知モデルを構築した場合に検知性能が低下するため，新たなデータを用意する必要がある．教師あり学習を用いた手法は，学習用データセットとして攻撃データが必要となるため，実システムに対して攻撃を行うことができない場合は，検知モデルの構築が困難であるという課題が挙げられる．

## 3. 提案

本研究では，前章で述べた課題の解決として，攻撃データを必要としない教師なし学習アルゴリズムを用いた新たな低レート DoS 攻撃検知手法を提案する．提案手法の概要を図 2 に示す．

本手法は，検知モデルを構築するためのトレーニングフェーズと Slow HTTP DoS 攻撃の有無を判定するランタイムフェーズの 2 つのフェーズに分かれる．

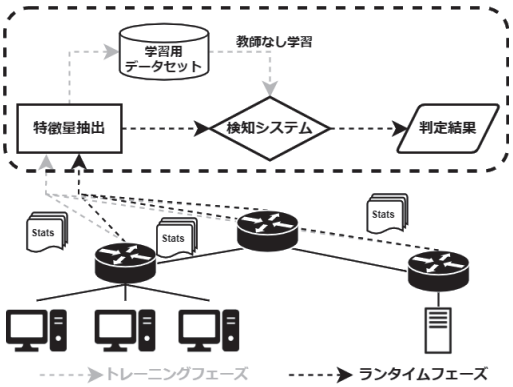


図 2: 提案手法の概要

表 1: 統計情報から抽出した特徴量

| 特徴量 | 説明 |
|---|---|
| duration | コネクション確立から終了までの秒数 |
| fwd_packets | 上りの総パケット数 |
| bwd_packets | 下りの総パケット数 |
| packets | フローの総パケット数 |
| fwd_bytes | 上りの総バイト数 |
| bwd_bytes | 下りの総バイト数 |
| bytes | フローの総バイト数 |
| fwd_pps | 上りのパケット毎秒 |
| bwd_pps | 下りのパケット毎秒 |
| pps | フローのパケット毎秒 |
| fwd_Bps | 上りのバイト毎秒 |
| bwd_Bps | 下りのバイト毎秒 |
| Bps | フローのバイト毎秒 |
| src | クライアントの IP アドレス: ポート番号 |

トレーニングフェーズでは，既存のネットワークから収集した統計情報を特徴量として抽出し，抽出した特徴量から保護対象となるサーバに関連するデータを学習用データセットとして蓄積する．蓄積した学習用データセットから教師なし学習アルゴリズムによって検知システムを構築する．ランタイムフェーズでは，トレーニングフェーズと同様に収集した統計情報を特徴量として抽出し，検知システムに入力することで，Slow HTTP DoS 攻撃の有無を判定する．統計情報から抽出する特徴量は Slow HTTP DoS 攻撃の特性を考慮した特徴量である．具体的な特徴量を表 1 に示す．既存のネットワークから統計情報を収集し，定常状態を学習することで，Slow HTTP DoS 攻撃の異常な通信を検知する．

## 4. 実験

実験では，データセットを用いた教師なし学習アルゴリズムの性能評価を行った．データセットには CICIDS 2017[2] を使用し，正常データのみで構成された訓練セット，正常・攻撃データで構成されたテストセットを使用した．また，教師あり学習アルゴリズムによる性能評価も行い，正常・攻撃データで構成された訓練セット，正常・攻撃データで構成されたテストセットを使用した．データの内訳は訓練セットに含まれる正常データは 50,800 件，攻撃データは 50,342 件であり，テストセットに含まれる正常データは 48,712 件，攻撃データは 50,342 件である．

表 2: データセットによる性能評価の結果

| モデル | Precision | Recall | F1 |
|---|---|---|---|
| AutoEncoder | 0.854 ± 0.064 | 0.855 ± 0.063 | 0.855 ± 0.064 |
| VAE | 0.819 ± 0.008 | 0.819 ± 0.008 | 0.819 ± 0.008 |
| Efficient GAN | 0.772 ± 0.045 | 0.775 ± 0.043 | 0.773 ± 0.044 |
| Dual-Encoder BiGAN | 0.753 ± 0.087 | 0.755 ± 0.088 | 0.754 ± 0.087 |
| SVM | 0.965 | 0.996 | 0.980 |
| RF | 0.998 ± 0.000 | 0.998 ± 0.000 | 0.998 ± 0.000 |
| CNN-LSTM | 0.964 ± 0.002 | 1.000 ± 0.000 | 0.981 ± 0.001 |

本実験で検証した教師なし学習アルゴリズムは AutoEncoder, Variational AutoEncoder (VAE), Efficient GAN, Dual-Encoder BiGAN の 4 種類である．GAN 系統は近年，注目されている異常検知アルゴリズムであり，画像データに対する異常検知では AutoEncoder 系統よりも高い検知性能を示しており，ネットワークトラフィックの異常検知への応用を検証する．教師あり学習アルゴリズムはサポートベクターマシン，ランダムフォレスト，CNN-LSTM を使用した．Precision, Recall, F1 の項目で評価し，サポートベクターマシン以外の各アルゴリズムで，実験を 10 回行い，平均値 ± 標準偏差を算出する．

実験の結果を表 2 に示す．結果より，教師なし学習アルゴリズムでは AutoEncoder が最も高い検知性能を示した．また，教師あり学習はすべてのアルゴリズムにおいて高い検知性能を示した．

## 5. 結論

本研究では，攻撃データの収集を不要とする教師なし学習を用いた新たな検知手法を提案した．実験の結果から教師なし学習が低レート DoS 攻撃の一種である Slow HTTP DoS 攻撃の検知に対して有効であることを示した．

今後の課題として，新たな異常検知アルゴリズムの応用や検知後の防御アプローチについて検討する．

## 参考文献

1) Phan, T. V. et al.: “Q-MIND: Defeating Stealthy DoS Attacks in SDN with a Machine-Learning Based Defense Framework,” Proc. of GLOBECOM (2019), pp. 1-6. 2019
2) Sharafaldin, I. et al.: “Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization,” Proc. of 4th ICISSP, pp. 1-8. 2018