アンサンブル学習を用いた IoT ネットワークおける攻撃検知

Effective Attack Detection in IoT Networks using Ensemble Learning

2241049　栗原　満洋
Mitsuhiro KURIHARA
指導教員　秋葉　知昭

This study proposes an IoT attack detection method based on ensemble learning. Random Forest and XGBoost were combined to improve detection performance, and temporal features were incorporated into the dataset. Experimental results show that the ensemble model outperformed individual models in terms of recall and F1-score, demonstrating the effectiveness of ensemble learning for IoT attack detection.

１．緒　言

近年,IoT 機器の普及に伴い,IoT ネットワークを標的としたセイバー攻撃が増加している.本研究では,IoT ネットワークにおける攻撃通信検知を目的として,時間帯・曜日別の脅威分析,SOM による可視化,Random Forest および XGBoost を用いたアンサンブル学習を適用する.さらに,モンテカルロシミュレーションによる評価を通じて,検知性能および安定性の向上を検討する[3].

２．提案手法の概要

本研究では,脅威分析,SOM による可視化,アンサンブル学習,モンテカルロシミュレーションによる評価からなる攻撃検知手法を提案する.

2.1 脅威分析

時間帯別および曜日別に攻撃発生状況を分析することで,攻撃種別ごとの時間的傾向を把握する.これらの分析から得られた時間帯情報および曜日情報を特徴量として機械学習モデルに導入し,攻撃検知性能の向上を図る.

2.2 自己組織化マップ（Self-Organizing Map：SOM）

SOM を用いることで,複数の特徴量からなる攻撃通信と正常通信の分布構造を可視化し,両者の分離性や攻撃パターンの特徴を把握する.これにより,機械学習モデル設計における特徴量選択およびモデル構成の妥当性を検討する.

2.3 アンサンブル学習

Random Forest と XGBoost により算出された攻撃クラスの予測確率を単純平均するアンサンブル学習を用いる.これにより,各モデルの極端な予測を抑制し,特定のモデルに依存しない安定した攻撃検知を実現する[2].

2.4 モンテカルロシミュレーション

攻撃発生率の変動を考慮した性能評価を行うために,モンテカルロシミュレーションを用いる.攻撃発生率を事前確率としてランダムに変化させ,ベイズの定理に基づき予測確率を補正した上で評価指標を算出することで,モデル性能の平均的傾向および安定性を検討する.

３．関連研究

既存研究では,IoT 環境を想定した攻撃通信と正常通信を生成するデータ基盤の提案や,通信量やパケット数などの特徴量を用いた機械学習による攻撃手法が報告されている.また, Precision, Recall, F1 スコアなど複数の評価指標を用いた性能評価や,SOM による通信パターンの可視化が有効であることが示されている.一方で,攻撃発生率の変動や実環境を想定した性能安定性については十分に検討されていない[1].

４．結果および考察

4.1 脅威分析の結果

時間帯, 曜日別の攻撃発生件数を図１, 図２に示す.

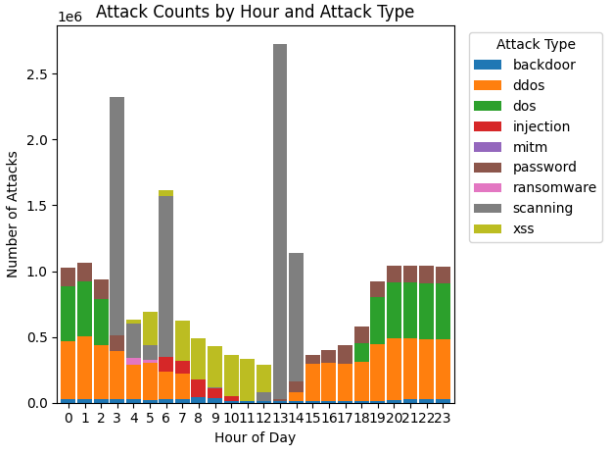

図１時間帯別分析

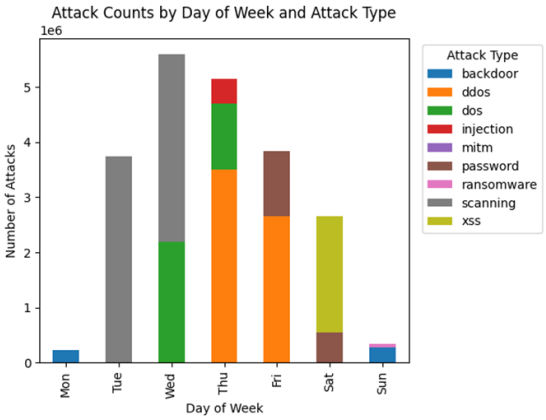

図２　曜日別分析

攻撃通信の発生状況を時間帯および曜日別に分析した結果,攻撃種別ごとに明確な偏りが確認された.このことから,時間帯および曜日情報は攻撃検知に有効な特徴量であると判断できる.本研究では,これらの特徴量を機械学習モデルに導入し,検知性能の向上を図る.

4.2 SOM による可視化結果

SOM による可視化の結果,攻撃通信は一様に分布せず,特定の領域に集中して配置されることが確認された.これは攻撃通信が共通した特徴量を持つ複数のパターンに分類可能であることを示す.

また,一部には境界的な分布も見られ,多様な攻撃特性が存在することが示唆された.これらの結果から,攻撃通信には非線形な特徴構造が内在していることが判断できる.

4.3 アンサンブル学習の性能評価

アンサンブル学習モデル性能評価の値を表１に示す.

表１アンサンブル学習モデルの性能評価結果

	Precision	Recall	F1-score
正常(0)	0.4765	0.5109	0.4931
攻撃(1)	0.9872	0.9853	0.9862
Accuracy			0.9732

攻撃通信に対する再現率および適合率はいずれも高く,高精度な攻撃検知が可能であることが確認された.一方で正常通信の誤検知は一部残るものの,単体モデルと比較して F1 スコアは改善しており,誤検知の抑制に一定の効果が見られた.これらの結果から,アンサンブル学習は検知性能と安定性の両立に有効であると言える.

4.4 モンテカルロシミュレーションの評価結果

モンテカルロシミュレーションの結果の値を表２に示す.

表２モンテカルロシミュレーション結果

	Precision	Recall	F1
Mean	0.999361	0.880400	0.932814
Std	0.000148	0.101689	0.061536
Min	0.999197	0.624052	0.768443
Max	0.999764	0.963005	0.980767

適合率は 0.9994 と非常に高く,誤検知の少ない判定が維持された.再現率および F1 スコアも高い平均値を示し,標準偏差が小さいことから,事前確率の変動に対しても安定した検知性能を有することが確認された.

以上より,提案するアンサンブル学習は実環境を想定した条件下においても有効であると言える.

５. 今後の課題

本研究の課題として,正常通信を攻撃通信として誤検知する傾向が残っている点が挙げられる.今後は,判定閾値の動的調整や通信の文脈情報を考慮した特徴量の導入により,誤検知低減を図る必要がある.また,本研究で用いたデータセットは分類が比較的容易であった可能性があるため,今後は実ネットワーク環境に近いデータを用いて提案手法の汎化性能を検証することが課題である.

６. 結　言

本研究では,IoT 環境におけるネットワーク攻撃検知を目的として,脅威分析と機械学習を組み合わせた手法を検討した.時間帯・曜日別分析および SOM により,攻撃通信の特徴的パターンを確認し,Random Forest と XGBoost を用いた,アンサンブル学習によって,誤検知の抑制と性能の安定化を実現した.さらに,モンテカルロシミュレーションにより,実環境を想定した条件下でもアンサンブル学習の有効性が示された.以上より,本研究で提案した手法は実運用を想定した IoT 環境における攻撃検知に有用であると考えられる.

文　献

[1]A. Mehrban and S. Karimi Geransayeh, “Ransomware threat mitigation through network traffic analysis,” *arXiv*, 2024. https://arxiv.org/abs/2401.15285

[2] IBM, 「アンサンブル学習とは」, IBM Japan, (2025), https://www.ibm.com/jp- ja/topics/ensemble-learning

[3] IPA, 情報セキュリティ 10 大脅威 2024. https://www.ipa.go.jp/security/10threats/10threats2024.html