アンサンブル学習を用いたネットワーク異常検出の精度向上

2017SC064 佐古竜基
指導教員：石原靖哲

1　はじめに

近年，コンピュータやスマートフォンなどのデバイスの普及により，インターネットの規模が拡大している．それに伴い，ネットワークを通じたサイバー攻撃が増加している．サイバー攻撃への対策として，機械学習を用いたネットワークの異常を検知する技術に注目が集まっている．

文献 [1] では，DNN を用いた手法が提案されており，正解率は 75.9% がとなっている．文献 [2] では，アンサンブル学習での手法が提案されており，KNN とランダムフォレストと決定木を用いている．この手法での正解率は 83.8% となっている．文献 [3] では，DNN や決定木などを用いたアンサンブル学習での手法が提案されている．この手法ではデータの主成分分析や交差検証を行っており，正解率は 85.2% となっている．しかし，DNN は単純構造なニューラルネットワークであり，複雑構造のニューラルネットワークを用いることでさらに正解率を向上させることができると考えられる．

そこで，複雑構造を持つニューラルネットワークを用いて，アンサンブル学習により正解率を向上させることを本研究の目的とする．各攻撃に対しての分類を行い，正解率の高い攻撃に対して重みを加えることで，全体の正解率を向上させる手法を提案する．

2　機械学習手法の概要

- **多層パーセプトロン (MLP)**：複数の層から構成される最も単純なニューラルネットワークである．
- **LSTM**：中間層にループ構造を持つ再帰型ニューラルネットワーク (RNN) の拡張であり，時系列データを扱うことができる．しかし，長期的な学習を行うことで計算コストが大きくなってしまう．
- **GRU**：LSTM と比較して，計算量を減らすことで高速に学習することが可能である．計算コストは低いが，LSTM と同等の結果を得ることが出来る．
- **CNN**：CNN(畳み込みニューラルネットワーク) は，畳み込み層のフィルタを利用してデータを変換する．隠れ層は畳み込み層とプーリング層で構成される．
- **SVM**：統計的学習理論に基づくパターン認識手法であり，学習の最適解として求めた分離超平面による線形識別を行っている．
- **アンサンブル学習**： 複数の異なるモデルを独立して学習させ，各モデルの出力結果を結合することで，単一のモデルより高い精度を得ることが出来る手法である．結合方法は 2 種類存在し，ソフトなアンサンブルは，ラベルの推定確率どうしの平均を取る方法である．ハードなアンサンブルは，分類器単位でラベルを求め，その多数決を取る方法である．

3　単一手法に対する予備評価

3.1　実行環境

実装するにあたり，Google Colaboratory を用いた．Google Colaboratory はブラウザで動作する Jupyter Notebook 環境で，Google によって研究・教育目的に無償提供されている．プログラミング言語は Python を用い，機械学習手法を構築するライブラリは Keras と scikit-learn を用いる．

3.2　データセット

本研究で使用するデータセットは NSL-KDD である．NSL-KDD データセットは 4 種類のデータファイルから構成されている．データセットには訓練データ 125,973 件とテストデータ 22,544 件のデータが含まれている．それぞれのデータの中には，41 個の特徴があり，それらは数値の範囲に違いがあるため，正規化の処理を行う．学習にはミニバッチ勾配降下法を用い，最適化アルゴリズムには Adam を用いる．

3.3　評価方法

テストデータを用いてモデルの性能評価を行う．評価には，Accuracy (正解率) と Precision (適合率) , Recall (再現率) を用いる．Accuracy (正解率) は，異常/正常のデータを正しく識別した割合である．Precision (適合率) は，異常と予測したものの内，真に異常であったものの割合である．Recall (再現率) は，異常全体の内，異常と予測したものの割合である．

3.4　単一手法に対する予備評価

ラベルを異常と正常の 2 種類とし，5 手法を用いて分類を行う．評価結果を表 1 に示す．その結果，正解率はどの手法も約 78% となった．全体を通して再現率が低く，適合率が高くなっているため，見逃しが多いことがわかる．

次に各攻撃に対する分類を行う．異常データを 4 種類の攻撃 (DoS，U2R，R2L，PROBE) に分類し，それぞれの攻撃に対しての正解率を求める．各攻撃に対する評価結果を表 2 と表に示す．その結果，SVM と多層パーセプトロンは DoS 攻撃に対して 90% を超える正解率となった．

4　アンサンブル学習による提案と実装

4.1　提案手法

本研究ではソフトなアンサンブルの出力値や重み，しきい値などを調整した手法を提案する．最終的な分類は異常

表 1　各手法に対する評価結果

Algorithm	Accuracy	Precision	Recall
LSTM	78.17%	96.81%	63.76%
GRU	78.83%	96.76%	64.98%
CNN	79.76%	93.30%	69.42%
MLP	78.50%	96.79%	64.37%
SVM	78.27%	97.33%	63.57%

表 2　各攻撃に対する Accuracy の評価結果

Algorithm	DoS	U2R	R2L	PROBE	正常
LSTM	88.61%	99.11%	87.78%	91.81%	74.86%
GRU	88.76%	99.11%	87.78%	95.68%	75.62%
CNN	82.70%	99.11%	87.78%	89.20%	75.91%
MLP	93.40%	99.11%	87.78%	94.38%	76.93%
SVM	92.68%	99.13%	87.77%	94.88%	76.40%

と正常の 2 種類であり，攻撃の分類は行わないものとする．提案手法の構造を図 1 に示す．

4.1.1　重みの追加

各攻撃に対する評価結果より，SVM と多層パーセプトロンは DoS 攻撃に有用であることが示された．それにより，LSTM，GRU，CNN の 3 手法で平均を取り，SVM と多層パーセプトロンは DoS の重みを加えるだけとする．

4.1.2　出力値の調整

3 手法で出力値を算出し，それを平均化する前に調整する．各手法の出力値は 0 から 1 の連続変数であり，正解と異なる結果を出力した値の中に，0.4 から 0.6 までの出力値が多数存在する．これを防ぐために，平均化前の出力値が 0.3 未満を 0，0.6 より大きい場合に 1 として変換し、平均化を行う．

4.1.3　しきい値の調整

平均化後，0 と 1 のラベルに変換するためのしきい値は，標準で 0.5 が使われている．しかし，0.1 から 0.5 の間にも異常データが多数含まれており，正常データは少ししか含まれていないため，しきい値を 0.08 として変換を行う．これは，試行錯誤の結果の最適な数値である．

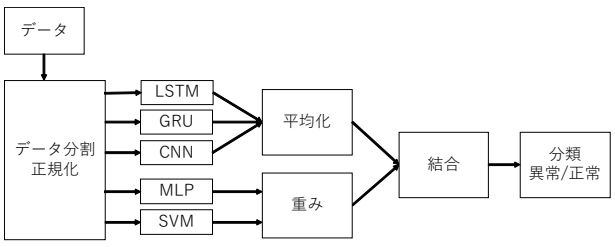

図 1　アンサンブル学習の構造

表 3　全ての評価結果

Algorithm	Accuracy	Precision	Recall
LSTM	78.17%	96.81%	63.76%
GRU	78.83%	96.76%	64.98%
CNN	79.76%	93.30%	69.42%
MLP	78.50%	96.79%	64.37%
SVM	78.27%	97.33%	63.57%
ソフトなアンサンブル	78.95%	96.77%	65.19%
ハードなアンサンブル	78.19%	96.87%	63.75%
Gao らのアンサンブル [3]	85.2%	86.5%	85.2%
提案手法	**86.83%**	**87.89%**	**86.76%**

4.2　比較と考察

全ての評価結果を表 3 に示す．単一手法では，どの手法も同じような結果となっている．アンサンブル後も単一手法とあまり変わらない結果となった．しかし，ソフトなアンサンブルをベースに調整した結果，単一で最も高い CNN の 79.76% から 85.66% まで向上させることができた．また，再現率も向上させることができ，これは見逃しが減ったことになる．Gao らの提案 [3] と比較すると，正解率は僅かながら高い結果となった．

5　まとめ

本研究では機械学習アルゴリズムである LSTM，CNN，GRU，多層パーセプトロン，SVM の 5 手法を用いてネットワークトラフィックの異常を検知する手法を実装し，それらをアンサンブル学習によって組み合わせる手法を提案した．さらに，提案手法の有効性を確かめるために，NSL-KDD データセットを用いて実験を行った．その結果，Gao らの提案 [3] より向上させることができた．

今後の課題としては，テストデータに対して未知の攻撃を完全には検知できていないため，主成分分析や交差検証などで対策する必要がある．また，他のデータセットでの実装も行う必要がある．

参考文献

[1] S. Rawat, A. Srinivasan, and R. Vinayakumar. Intrusion detection systems using classical machine learning techniques versus integrated unsupervised feature learning and deep neural network. *ArXiv*, Vol. abs/1910.01114, , 2019.

[2] P. Illy, G. Kaddoum, C. M. Moreira, K. Kaur, and S. Garg. Securing fog-to-things environment using intrusion detection system based on ensemble learning. In *2019 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1–7, 2019.

[3] X. Gao, C. Shan, C. Hu, Z. Niu, and Z. Liu. An adaptive ensemble machine learning model for intrusion detection. In *IEEE Access*, Vol. 9, pp. 82512–82521, 2019.