F-005

ニューラルネットワークを用いた通信データの分類手法
Neural Network for Network Traffic Classification

永澤 大智† 武田 敦志‡
Daichi Nagasawa Atsushi Takeda

1. はじめに

データセンターのシェアが行われるようになり、複数のデータセンターを仮想化して連携し、一つのデータセンターとして利用するマルチデータセンターが用いられるようになってきた。マルチデータセンターは分散されていたデータが共有されることから、より攻撃者に狙われやすくなると考えられる。

この問題に対して、Intrusion Detection System(IDS)を用いた様々なアプローチが行われている。これまで決定木アルゴリズムの実装[1]やナイーブベイズと Support Vector Machine(SVM)を組み合わせたハイブリットモデルの実装[2]等によって検知手法の改善が行われてきた。これらの手法では、DoS 攻撃、Probe 等の通信データに対しては高い検知率が達成できるが、User-to-Root(U2R)や Remote-to-Local(R2L)といった不正アクセス攻撃に関連する通信データの検知率に課題が残っている状態にある。最近では、多層ニューラルネットワークを用いたアンサンブル効果による検知率向上の研究[3]では U2R と R2L はそれぞれ 50.89%と 29.85%、SVM と Extreme Learning Machine を用いた分類器の構築[4]では U2R と R2L はそれぞれ 21.93%、31.39%を達成している。

本研究では Deep Learning の技術の一つである Residal Network(ResNet)[5]を用いて、U2R および R2L を含めた通信データの検知に適したニューラルネットワークの構築を目的とする。ResNet は Convolutional Neural Network(CNN)を発展させたもので、シンプルな仕組みでより深いニューラルネットワークの構築が可能である。そのため、従来の Machine Learning の技術による IDS に比べ、ResNet を用いることで高精度の分類器モデルの構築が期待できる。ここでは、通信データの検知に対する ResNet のネットワーク構造の比較を行い、適切なネットワーク構造を探る。

2. ニューラルネットワークを用いた攻撃検知手法

2.1 通信データの特徴を考慮した前処理

学習に使用するデータ内容に偏りがあると、過学習や未学習等が起こり、上手く学習できない可能性がある。そのため、不十分なデータを補強するデータセットの拡張処理と、正則化の強化のため入力データの CutOut 処理を実装する。

2.1.1 入力処理

通信データをそのままニューラルネットワークに入力することはできない。ニューラルネットワークへの入力のために、通信データを 0 と 1 で構成された二次元配列データに変換する。

† 東北学院大学大学院人間情報学研究科 Graduate School of Human Informatics, Tohoku Gakuin University
‡ 東北学院大学教養学部情報科学科 Department of Science, Tohoku Gakuin University

2.1.2 データセットの拡張処理

訓練データに含まれる通信データの種類に偏りがある場合、そのままでは上手く学習できないと考えられる。そこで、通信データを構成する要素に基づき、データセットの拡張を行うことを考えた。ここでは、通信データの構成要素に基づき、その出現頻度と出現の有無から、データセットの拡張を行った。

2.1.3 CutOut の実装

ニューラルネットワークの正則化を強めるため、CutOut を実装した。CutOut は入力データに対してランダムで数値 0 のマスクをすることで、ネットワークの正則化を強めることができると考えられている技術である[6]。マスクの範囲は入力データの四分の一とし、形状は固定サイズの正方形とした。また、マスクの画面外への食み出しも許容している。

2.2 攻撃検知のためのニューラルネットワーク

CNN は深層学習の分野で目覚ましい成果を挙げているが、多層ニューラルネットワークの持つ勾配消失問題も抱えたままである。これはニューラルネットワークの学習に必要な勾配が、計算の途中で非常に小さな値に収束し、消失してしまうためだと考えられる。ResNet ではこの問題に対して、ネットワークの残差を利用することで解決を図っている。層で求める出力をそのまま学習に用いるのではなく、その層への入力を参照した残差を学習に用いている。ResNet は従来の CNN を構成する Convolution 層(Conv 層)と Pooling 層を加えて、Shortcut Connection を導入することで実現できる。Shortcut Connection と Conv 層、活性化関数(ここでは ReLU 関数とする)からなるモジュール構造部分を Residual Block(ResBlock)とし、この ResBlock を複数個重ねたネットワーク構造が ResNet になる。汎化性能の高い ResNet を用いることで、U2R および R2L の検知率の向上が予想される。

3. 実験評価

3.1 ネットワーク構造

通信データの検知率評価として、複数の ResNet のネットワーク構造の結果を比較する。ここでは、異なるネットワーク構造(図 1 に示す)と Optimizer を組み合わせた 5 つの ResNet の比較を行う。使用したネットワーク構造はモデル 1、モデル 2、モデル 3 の三種類、Optimizer は Momentum SGD と Adam の二種類である。

3.2 実装

3.2.1 データセット

通信データに対する検知精度を評価するため、KDDCup 1999Dataset(KDDCup99)をデータセットとして使用する。KDDCup99 に含まれている通信データは Normal、Probe、DoS、R2L、U2R に分類できるため、U2R および R2L の検

知率を評価できると考えられる。ここでは訓練データとして kddcup.data_10_percent、評価データとして corrected を使用した。ただし、訓練データに含まれる R2L および U2R のデータ数が少ないため、Normal、DoS、Probe、R2L、U2R がそれぞれ 2000 個になるように拡張処理を行い、これを実際の訓練データとした。

3.2.2 フレームワーク

Preferred Networks の Chainer を用いて、前処理および ResNet の構築、評価処理の実装を行った。

3.2.3 ネットワークの構成要素

ネットワークの共通パラメータとして、ResBlock=10、チャンネル数=32、バッチサイズ=64、Gaussian Noise=0.5 としている。構成要素として、Pooling 層は Average Pooling とし、Batch Normalization(Batch Norm)[7]と Dropout[8]を組み込んでいる。ResBlock は Batch Norm と活性化関数を Conv 層の前方に設置する Pre Activation 構造[9]とした。

3.3 実験結果

3.3.1 評価指標

各 ResNet に評価データを入力し、Normal、DoS、Probe、R2L、U2R の分類を行った。検知率評価のため通信データの種類に基づき、Precision、Recall、F-score を算出した。Precision と Recall はトレードオフの関係にあり、それぞれ重視する指標が異なるため、これらの評価指標だけではモデルを評価するのは難しいと考えられる。そのため、二つの調和平均をとる F-score を評価指標として扱う。F-score の定義は以下の通りである。

$$\mathrm{F-score} = \frac{2 * (precision * recall)}{(precision + reall)}$$

3.3.2 結果比較

結果をまとめたものを表 1 に示す。表 1 からデータセットの拡張処理を行わない場合、R2L と U2R の学習を上手く行えないことがわかる。モデル 2、3 を使用した場合、U2R と R2L の F-score が向上していることがわかる。モデル 2、3 では、ResBlock の間に Pooling 層を挟むことで、データ全体に Pooling 処理が行き届き、特徴を比較的正確に獲得できていると考えられる。モデル 2 では最後に Fully Connected 層に渡しているが、モデル 3 では Global Average Pooling 層に渡すことで F-score が向上している。U2R と R2L の F-score が最高値でも 31.2%、41.1%と半数を超えていない理由として、評価データに含まれる未知の通信データに適応できていない可能性が考えられる。他の通信データに比べ、U2R と R2L は訓練データと評価データの差異が大きく、学習に弊害ができているのではないかと予想される。

4. まとめ

本発表では、Deep Learning の技術の一つである ResNet を用いて、通信データの検知に適するネットワーク構造を探るためにモデルの比較を行った。適切に Pooling 層を組み込むことで、ResNet の特徴抽出の精度向上が可能であることがわかった。U2R と R2L を含む未知データの課題に関しては、二つの同一のネットワークに異なる入力をして類似点を比較する Siamese Network[10]を用いることで、未知の通信データに対する検知精度向上が見込めるのではないかと予想している。

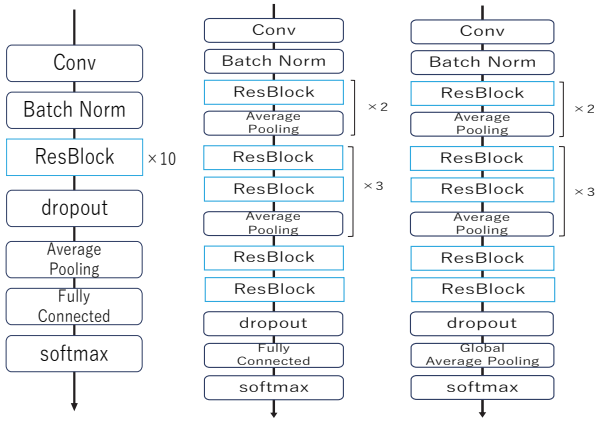

図 1　モデル 1, モデル 2, モデル 3

構造	拡張処理		Normal	Probe	DoS	U2R	R2L
モデル1 + MomentumSGD	なし	Precision	0.897	0.698	0.962	0.000	0.000
		Recall	0.988	0.702	0.893	0.000	0.000
		F-Score	0.940	0.700	**0.926**	0.000	0.000
モデル1 + MomentumSGD	あり	Precision	0.931	0.517	0.974	0.040	0.305
		Recall	0.974	0.756	0.819	0.629	0.233
		F-Score	0.952	0.614	0.890	0.076	0.264
モデル1 + Adam	あり	Precision	0.921	0.677	0.973	0.030	0.224
		Recall	0.961	0.750	0.852	0.557	0.177
		F-Score	0.941	**0.711**	0.908	0.057	0.198
モデル2 + Adam	あり	Precision	0.956	0.468	0.986	0.064	0.393
		Recall	0.975	0.956	0.827	0.614	0.253
		F-Score	**0.961**	0.628	0.900	0.116	0.308
モデル3 + Adam	あり	Precision	0.953	0.576	0.972	0.213	0.351
		Recall	0.953	0.821	0.869	0.586	0.496
		F-Score	0.953	0.677	0.918	**0.312**	**0.411**

表 1　各ネットワークの分類結果

参考文献

[1] Shailendra Sahu, B M Mehter, “Network Intrusion Detection System Using J48 Decision Tree”, ICACCI2015
[2] Amit D Sagale, Swati G Kale, “Combining Naïve Bayesian and Support Vector Machine for Intrusion Detection System”, IJCAT, Volume1(2014)
[3] Maxine Labonne, Alexis Olivereau, Baptiste Polvé, Djamal Zeghlache, “A Cascade-structured Meta-Specialists Approach for Neural Network-based Intrusion Detection”, CCNC(2016)
[4] Wathiq Laftah AI-Yaseen, Zulaiha Ali Othman, Mohd Zakree Ahmad Nazri, “Multi-level hybrid support vector machine and extreme learning machine based on modified K-means for intrusion detection system”, Expert System With Applications(2017)
[5] Kaiming He, Xiangyu Zhang, Shaoguing Ren, “Deep Residual Learning for Image Recognition”, Proceeding of the IEEE conference on computer vision and pattern recognition(2016)
[6] Terrance DeVries, Graham W Taylor, “Improved Regularization of Convolutional Neural Networks with Cutout”(2017)
[7] Sergey Loffe, Christian Szegedy, “Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift”(2015)
[8] Nitish Srivastave, Geoffery Hinton, Alex Krizhevsky Ilya Sutskever, Ruslan Salakhutdinov, “Dropout: A simple Way to Prevent Neural Networks from Overfitting”, Journal of Machine Learning Research15(2014)
[9] Kaiming He, Xiangyu Zhang, Shaoging Ren, Jian Sun, “Identity Mapping in Deep Residual Networks”, European conference on computer vision(2016)
[10] Gregory Koch, Richard Zemel, Ruslan Salakhutdinov, “Siamese Neural Networks for One-shot Image Recognition”, ICML Deep Learning Workshop(2015)