情報処理学会論文誌 Vol.61 No.7 1249–1254 (July 2020)

テクニカルノート

LSTMによるネットワーク異常検出

浦川 侑之介[1,†1] 青木 茂樹[1,a)] 宮本 貴朗[1,b)]

受付日 2020年1月28日, 採録日 2020年4月8日

概要：本論文では，パケットのヘッダから単位時間ごとに抽出した特徴ベクトルを学習することによりネットワークの異常を検出する手法を提案する．単位時間ごとに特徴を抽出することで特徴抽出の処理にかかるコストを削減し，トラフィック量の多い大規模ネットワークに適用可能な異常検出手法の実現を目指す．特徴ベクトルの学習には，Deep Learning の一手法である LSTM（Long short-term memory）を用いる．実験では，MWS データセット，CICIDS2017 データセットおよび大阪府立大学のキャンパスネットワークのトラフィックデータに本手法を適用し，C&C サーバとの通信，DDoS 攻撃，Port scan 攻撃などの異常を検出できることを確認した．

キーワード：異常検知，Deep Learning，LSTM

Anomaly Detection of Network by LSTM

Yunosuke Urakawa[1,†1] Shigeki Aoki[1,a)] Takao Miyamoto[1,b)]

Received: January 28, 2020, Accepted: April 8, 2020

Abstract: In this research, we propose a method to detect anomalies on a network by learning feature vectors extracted from packet headers at every unit time. We reduce cost of feature extraction processing by using extracted features at every unit time and aim to realize the anomaly detection method applicable to large-scale networks with large traffic volume. We use LSTM (Long short-term memory) which is an extension of recurrent neural networks for feature vector learning. In the experiment, the method was applied to MWS dataset, CICIDS2017 dataset and traffic data of Campus Network of Osaka Prefecture University. We have detected anomalies such as C&C communications, DDoS attacks, and port scans.

Keywords: anomaly detection, Deep Learning, LSTM

1. はじめに

近年，インターネットの発達にともない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，ネットワークに対する不正な通信を検出するための侵入検知システムの研究がさかんに行われている．また，近年は画像認識や自然言語処理などの分野で Deep Learning を用いた手法が数多く提案されている．さらに，Deep Learning をネットワークの異常検知に応用している研究が注目されている．文献 [1] では，パケットのペイロードから抽出した特徴に対して CNN（Convolutional Neural Network）を適用した手法と RNN（Recurrent Neural Network）の拡張である LSTM（Long short-term memory）を適用した手法によりネットワークの異常を検出している．また，文献 [2] では RNN の拡張である GRU（Gated Recurrent Unit）と多層パーセプトロンを組み合わせた異常検出手法を提案している．文献 [2] では，TCP のセッションごとに抽出した特徴ベクトルを学習し，異常の種類を識別しているため処理に時間がかかる．そのため，大規模ネットワークへの適用が難しい．

本論文では，特徴抽出の処理にかかるコストを削減し，トラフィック量の多い大規模ネットワークに適用できるよ

[1] 大阪府立大学大学院人間社会システム科学研究科
Graduate School of Humanities and Sustainable System Sciences, Osaka Prefecture University, Sakai, Osaka 599–8531, Japan
[†1] 現在，日本電気株式会社
Presently with NEC Corpotarion
[a)] shigeki_aoki@m.ieice.org
[b)] aki@center.osakafu-u.ac.jp

うにすることを目的としたネットワークの異常検出手法を提案する．本手法では，パケットのヘッダから単位時間ごとに抽出した特徴ベクトルを LSTM で学習することによりネットワークにおける異常の有無を検出する．実験では，MWS データセット，CICIDS2017 データセットおよび大阪府立大学のキャンパスネットワークのトラフィックデータに本手法を適用し，有効性を確認した．

2. 提案手法

本手法の異常検出の手順について述べる．本手法は学習と検出の 2 つの処理に分かれている．学習時の処理では，学習に用いるトラフィックデータを一定の単位時間で分割し，各区間に含まれるパケットのヘッダから特徴ベクトルを抽出し，正規化する．その後，抽出した特徴ベクトルを LSTM により学習する．検出時の処理では，学習に用いたデータとは別の期間に収集されたトラフィックデータを学習時と同様に一定の単位時間で分割し，特徴ベクトルを抽出し，正規化する．その後，学習済みの LSTM を用いて予測値を出力する．予測値と実際の特徴ベクトルとの誤差を算出し誤差が閾値未満であった場合，その区間は正常であるとする．一方，誤差が閾値以上の場合はその区間には学習に用いたデータには含まれない特徴が含まれると識別し，異常の発生を検出する．

2.1 特徴ベクトルの抽出および正規化

パケットのヘッダから単位時間ごとに特徴ベクトルを抽出する．パケットから特徴ベクトルを抽出する方法として，TCP のセッションごとに特徴ベクトルを抽出する方法と単位時間ごとに特徴ベクトルを抽出する方法の大きく 2 種類がある．セッションごとに抽出した特徴ベクトルを用いる場合，大規模ネットワークに適用した際に特徴ベクトルの抽出処理に時間がかかる可能性がある．一方，単位時間ごとに特徴ベクトルを抽出する場合，少ない処理時間で特徴ベクトルを抽出することができる．

そこでまず，注目しているネットワークと外部ネットワーク間のパケットを収集し，単位時間 ω で収集したトラフィックデータを分割する．単位時間で分割したものを 1 区間とし，特徴ベクトル $\boldsymbol{t}_n$ を N 区間分収集する．次に，各区間に含まれるパケットのヘッダから**表 1** に示す送信元 IP アドレス種類数やパケット中の SYN フラグ数などの 53 種類の特徴ベクトルを抽出する．その後，各特徴量に対して最大値が 1，最小値が 0 になるように正規化する．以上のように，単位時間ごとに特徴ベクトルを抽出することにより，UDP のようなセッションを用いない通信での異常も検出できる．抽出する特徴量はパケットのヘッダのみに着目しているため，DDoS や Port scan のような攻撃を異常として識別できる．また，ペイロードの情報を参照しないため，暗号化された通信にも適用できる．

表 1 抽出した特徴ベクトルの一覧
Table 1 List of Features.

パケットサイズ平均	パケット数
パケットサイズの分散	TTL 値平均
TTL 値分散	宛先 IP アドレス種類数
送信元 IP アドレス種類数	送信元ポート番号種類数
宛先ポート番号種類数	SYN パケット数
FIN パケット数	PSH パケット数
RST パケット数	URG パケット数
ACK パケット数	FIN&ACK パケット数
RST&ACK パケット数	SYN&ACK パケット数
PSH&ACK パケット数	TCP 中の RST 割合
TCP 中の SYN 割合	TCP 中の PSH 割合
TCP 中の URG 割合	TCP 中の FIN 割合
TCP 中の ACK 割合	TCP 中の RST&ACK 割合
TCP 中の PSH&ACK 割合	TCP 中の SYN&ACK 割合
TCP 中の FIN&ACK 割合	ICMP パケット数
UDP パケット数	送信元ポート番号 110 番パケット数
送信元ポート番号 22 番パケット数	送信元ポート番号 53 番パケット数
送信元ポート番号 443 番パケット数	送信元ポート番号 80 番パケット数
送信元ポート番号 25 番パケット数	送信元ポート番号 465 番パケット数
送信元ポート番号 587 番パケット数	送信元ポート番号 995 番パケット数
送信元ポート番号 993 番パケット数	送信元ポート番号 143 番パケット数
宛先ポート番号 110 番パケット数	宛先ポート番号 22 番パケット数
宛先ポート番号 53 番パケット数	宛先ポート番号 443 番パケット数
宛先ポート番号 80 番パケット数	宛先ポート番号 25 番パケット数
宛先ポート番号 465 番パケット数	宛先ポート番号 587 番パケット数
宛先ポート番号 995 番パケット数	宛先ポート番号 993 番パケット数
宛先ポート番号 143 番パケット数	

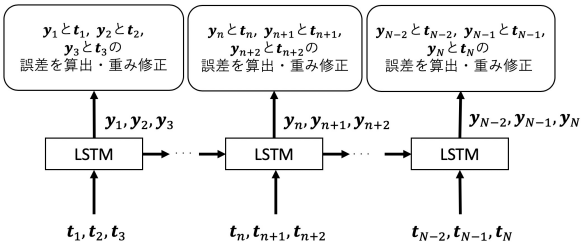

図 1 LSTM の学習の概要
Fig. 1 Overview of Learning of LSTM.

2.2 LSTM による学習

特徴ベクトルの抽出および正規化の後，RNN の拡張である LSTM で学習する．LSTM では，長期的な情報を保持して時系列データを学習できる．提案手法における LSTM の学習方法の概要を**図 1** に示す．連続する 3 区間の特徴ベクトル $\boldsymbol{t}_n$，$\boldsymbol{t}_{n+1}$，$\boldsymbol{t}_{n+2}$ を LSTM に入力し，予測値 $\boldsymbol{y}_n$，$\boldsymbol{y}_{n+1}$，$\boldsymbol{y}_{n+2}$ を出力させる．出力されたベクトル $\boldsymbol{y}_n$，$\boldsymbol{y}_{n+1}$，$\boldsymbol{y}_{n+2}$ と入力した区間の特徴ベクトル $\boldsymbol{t}_n$，$\boldsymbol{t}_{n+1}$，$\boldsymbol{t}_{n+2}$ をそれぞれ比較し，誤差を算出する．誤差の算出には平均二乗誤差（MSE）を用いた．得られた誤差を基に誤差逆伝播により LSTM の重みの更新を行い，出力が入力したベクトルに近づくように学習を進める．学習は事前に定めたエポックに達した場合に終了することとしている．

2.3 異常検出

学習した LSTM を用いて異常を検出する．学習時のトラフィックデータから抽出した特徴ベクトルと，検出に用いるトラフィックデータから抽出した特徴ベクトルが類似していた場合は，正しく入力された特徴ベクトルを予測することができる．一方，類似していない場合は入力された特徴ベクトルを正しく予測することができないと考えられる．そこで，予測値と入力された特徴ベクトルとの差分を異常度とし，異常度があらかじめ設定した閾値を上回っている区間を異常，下回っている区間を正常として識別する．異常度を算出する式を式 (1) に示す．

$$\begin{aligned} \textit{Anomaly score} = &(\boldsymbol{t}_n^{'} - \boldsymbol{y}_n^{'})^2 + (\boldsymbol{t}_{n+1}^{'} - \boldsymbol{y}_{n+1}^{'})^2 \\ &+ (\boldsymbol{t}_{n+2}^{'} - \boldsymbol{y}_{n+2}^{'})^2 \end{aligned} \tag{1}$$

ここで，$\boldsymbol{t}_n^{'}$，$\boldsymbol{t}_{n+1}^{'}$，$\boldsymbol{t}_{n+2}^{'}$ は入力された特徴ベクトル，$\boldsymbol{y}_n^{'}$，$\boldsymbol{y}_{n+1}^{'}$，$\boldsymbol{y}_{n+2}^{'}$ は学習済み LSTM に特徴ベクトル $\boldsymbol{t}_n^{'}$，$\boldsymbol{t}_{n+1}^{'}$，$\boldsymbol{t}_{n+2}^{'}$ を入力した際の予測値である．

3. 実験

MWS データセット，CICIDS2017 データセットおよび大阪府立大学のキャンパスネットワークのトラフィックデータを用いた実験を行い，本手法の有効性を確認した．

3.1 MWS データセットにおける実験

MWS データセット中の BOS データセット [3] を用いた実験を行った．BOS データセットでは，マルウェアを実行したホストの通信をキャプチャした複数の 24 時間分のトラフィックデータとトラフィックデータに対応するマルウェアの進行度が示されており，マルウェアにより通信が発生したか，またどのように通信が行われたかが示されている．進行度ごとの説明を**表 2** に示す．本論文では，進行度 2 のトラフィックデータを異常を含まない正常なトラフィックデータ，進行度 7，8 のトラフィックデータを異常を含むトラフィックデータとして実験に用いた．ここで，特徴ベクトル抽出の際の単位時間は 30 秒とし 2880 区間に分割し実験に用いた．本研究では，C&C サーバとの通信を含む区間を異常，それ以外の区間を正常として実験した．また，進行度 2 のトラフィックデータを用いて学習した際のエポック数に対する学習誤差は，エポック数が 20 を超えた辺りから収束することを確認した．そのため，MWS データセットを用いた実験では，エポック数を 20 として実験を行った．

学習データに進行度 2 のトラフィックデータ，テストデータに進行度 8 のトラフィックデータを用いて ROC 曲線および AUC 値の算出を行った．テストデータのラベル付けとして 1 区間に C&C サーバとの通信のパケット数が P 以上の場合に異常とした．P の値を 1，10，20 と変化させたときの ROC 曲線および AUC 値を**図 2** に示す．図 2 に示すとおり，1 区間に 10 パケット以上の C&C サーバとの通信が含まれるときを異常とした際に AUC 値が 0.94 と最も高くなった．また，1 区間に 20 パケット以上の場合は AUC 値が 0.87 となり，10 パケット以上の場合より減少する結果となったが高い AUC 値となった．C&C サーバとの通信のパケットが 1 区間に 10 パケット以上含まれていた場合，正常な区間から抽出した特徴ベクトルとの差異が生まれ，高い精度で分類できることを確認できた．一方で，1 区間に C&C サーバとの通信のパケットが 1 パケット以上の場合は AUC 値が 0.72 と最も低かった．1 区間に含まれる C&C サーバとの通信のパケットが少ない場合は，抽出した特徴ベクトルが正常の特徴ベクトルと類似したため AUC 値が低い結果になったと考えられる．

次に，学習データが正常通信のみの場合と異常通信を含む場合の検出結果の比較を行った．学習データとテストデータの組合せごとの閾値，AUC，TPR（Recall），FPR，Precision，F-measure を**表 3** に示す．また，テストデータに対するラベル付けは 1 区間に C&C サーバとの通信のパケット数が 10 パケット以上含まれる場合を異常として検出実験を行った．いずれの組合せでも AUC，TPR（Recall）については高く，FPR については低い結果となった．

進行度 2 のトラフィックデータで学習を行い，進行度 7

表 2 進行度の説明
Table 2 Explanation of progress.

進行度	説明
1，2	通信発生なし
3，4，5	通信発生したが，C&C サーバとの攻撃通信不成立
6，7，8	通信発生かつ C&C サーバとの攻撃通信成立

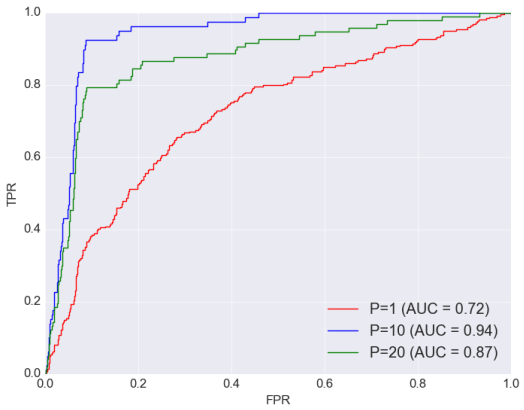

図 2 P の値ごとの ROC 曲線（BOS データセット）
Fig. 2 ROC Curve when P is changed (BOS dataset).

表 3 検出結果（BOS データセット）
Table 3 Result of detection (BOS dataset).

学習データ	テストデータ	閾値	AUC	TPR(Recall)	FPR	Precision	F-measure
進行度 2	進行度 7	3.2	1	1	0	1	1
進行度 2	進行度 8	0.056	0.94	0.92	0.09	0.23	0.36
進行度 7	進行度 8	0.071	0.96	0.9	0.07	0.26	0.41
進行度 8	進行度 7	0.18	1	0.99	0.002	1	0.99

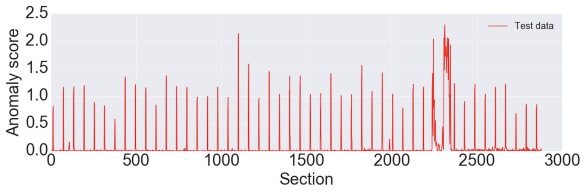

図 3 異常度のグラフ（BOS データセット）
Fig. 3 Graph of Anomaly Score (BOS dataset).

のトラフィックデータでテストを行った際，Precision が 1，F-measure が 1 と高い結果となった．これは学習データが全て正常な通信であったからであると考えられる．そのため，正常な通信の特徴をうまく学習することができ，テストデータに現れた異常を正しく識別することができた．

進行度 2 のトラフィックデータで学習を行い，進行度 8 のトラフィックデータでテストを行った際，Precision が 0.23，F-measure が 0.36 と低い結果となった．**図 3** に異常度のグラフを示す．図では横軸に区間（時刻），縦軸に異常度を示している．図 3 より，実際に異常ラベルが付与されている区間は 2244 区間から 2341 区間のあたりであるが，異常でない区間にもかかわらず定期的に異常度が高くなっていることが分かる．この区間の特徴ベクトルを確認してみると，143 番ポートに関する特徴量の値が増加していることを確認できた．学習データでは，定期的に 143 番ポートに関する特徴量が増加している区間がなかったため，検出の際に異常度が高くなったと考えられる．そのため，Precision，F-measure の値が低くなったが，学習データに存在しない通信パターンを検出できることを確認できた．

進行度 7 のトラフィックデータで学習を行い，進行度 8 のトラフィックデータでテストを行った際，Precision が 0.26，F-measure が 0.41 と低い結果になった．これは学習に用いた進行度 7 のトラフィックデータのなかに多くの異常な通信が含まれていたからであると考えられる．進行度 7 のトラフィックデータは 2880 区間の内，1650 区間に正常のラベル，1230 区間に異常のラベルが付与されている．そのため，正常の特徴と異常の特徴の両方を学習してしまい，うまく検出することができなかった．

進行度 8 のトラフィックデータで学習を行い，進行度 7 のトラフィックデータでテストを行った際，Precision が 1，F-measure が 0.99 と高い結果になった．これは，正常の特徴をうまく学習できていたためであると考えられる．学習に用いた進行度 8 のトラフィックデータの 2880 区間の内，2801 区間に正常のラベルが付与されており，ほとんどが正常な区間であった．そのため，正常な特徴が学習されてうまく検出することができた．

以上の結果より，学習に用いるトラフィックデータに異常が含まれていない，もしくは異常が少数であればテストデータの異常を高精度に検出可能であることを確認できた．

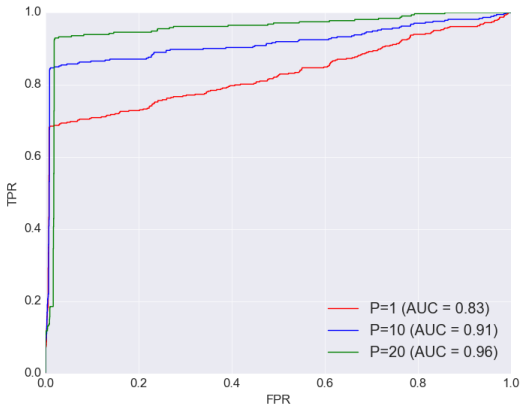

図 4 P の値ごとの ROC 曲線（CICIDS2017 データセット）
Fig. 4 ROC Curve when P is changed (CICIDS2017 dataset).

表 4 検出結果（CICIDS2017 データセット）
Table 4 Result of Detection (CICIDS2017 dataset).

AUC	Precision	Recall	F-measure
0.96	0.86	0.93	0.89

また，検出の際に学習データに存在しなかった新たな通信パターンが含まれていた場合についても検出できることを確認できた．一方で，学習に用いるトラフィックデータに異常が多数含まれている場合はうまく検出できなかった．

3.2 CICIDS2017 データセットにおける実験

CICIDS2017 データセット [4] を使用して実験を行った．CICIDS2017 データセットは侵入検知精度の評価用データセットであり，複数のトラフィックデータとトラフィックデータに対応する異常のラベルが付与されている．学習データには正常な通信のみを含むトラフィックデータを用い，テストデータには DDoS および Port scan を異常として含むトラフィックデータを用いた．ここで，特徴ベクトル抽出の際の単位時間は 5 秒とし，学習データは 5829 区間，テストデータは 2800 区間に分割し実験に用いた．また，正常通信のみを含むトラフィックデータを用いて学習した際のエポック数に対する学習誤差は，エポック数が 30 を超えた辺りから収束することを確認した．そのため，CICIDS2017 データセットを用いた実験では，エポック数を 30 として実験を行った．

テストデータのラベル付けとして 1 区間に異常な通信のパケット数が P 以上の場合に異常とした．P の値を 1，10，20 と変化させたときの ROC 曲線および AUC 値を**図 4** に示す．図 4 より，3.1 節の実験結果と同様に異常な通信のパケットが 1 区間にある程度以上含まれていた場合，正常な区間から抽出した特徴ベクトルとの差異が生まれ，高い精度で検出できることを確認した．

また，異常検出を行った際の AUC，Precision，Recall，F-measure を**表 4** に示す．テストデータに対するラベル付けは 1 区間に異常な通信のパケット数が 20 パケット以上

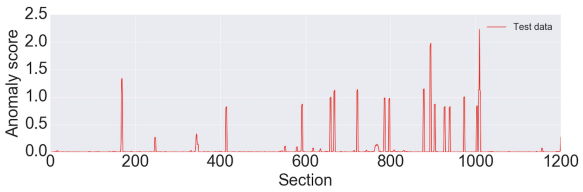

図 5 異常度のグラフ（大阪府立大学のトラフィックデータ）
Fig. 5 Graph of Anomaly Score (Traffic Data of Osaka Prefecture University).

含まれる場合を異常として検出実験を行った．異常検出の際の閾値は図 4 の P=20 のグラフにおいて TPR が 0.93，FPR が 0.02 の際の閾値である 0.044 を用いた．実験の結果，どの評価値も高い結果となった．これは，学習データには異常が含まれていなかったため，テストデータに含まれる異常を正しく識別できたためであると考えられる．また，テストデータに含まれる攻撃ごとの検出率を確認したところ，DDoS は 97.5%，Port scan は 77.3%であった．DDoS の検出率が 97.5%と高くなったのは，DDoS を含む全ての区間が連続しており，ほとんどの区間で DDoS のパケットの割合が高かったためであると考えられる．一方で，Port scan の検出率が 77.3%と DDoS に比べ低くなったのは，Port scan を含む区間の一部は連続しておらず，さらに 1 区間に含まれる Port scan のパケットの割合が少ない区間が存在したためであると考えられる．以上の結果より，1 区間に含まれる異常のパケットの割合が多ければ高い精度で検出できることを確認した．

3.3 大阪府立大学のトラフィックデータにおける実験

大阪府立大学のトラフィックデータを使用して実験を行った．学習データとして 10 分間のトラフィックデータを用いた．テストデータとしては学習データとは別の期間に収集した 10 分間のトラフィックデータを用いた．また．テストデータ内において UTM（Unified Threat Management）で DoS 攻撃の一種である TCP flood 攻撃が観測されている．学習データ，テストデータ共に特徴ベクトルの抽出の際の単位時間は 0.5 秒とし 1200 区間に分割して実験に用いた．UTM において DoS 攻撃を観測した区間はテストデータの 810 区間目である．また，エポック数は 100 で学習した．学習したモデルにテストデータを入力した際の異常度のグラフを**図 5** に示す．図 5 より，800 区間の直前から 800 区間にかけて異常度の値が大きくなっていることが分かる．これは，UTM が観測した DoS 攻撃の影響であると考えられる．UTM がアラートを発生したのは 810 区間目であるが，UTM における検出はある一定の閾値を超えた場合のみに行われる．そのため，実際に DoS 攻撃が発生した時刻はアラート発生よりも前の時刻であるため，800 区間の直前から 800 区間にかけて異常度の値が高くなったと考えられる．また，そのほかにも異常度が高い区間がいくつか存在した．これは，本手法により UTM では検知できなかった異常を検出できた可能性があると考えられる．

4. まとめ

本論文では，パケットのヘッダから抽出した特徴ベクトルに対して LSTM で学習を行うことで異常を検出する手法を提案した．実験では，MWS データセット，CICIDS 2017 データセットおよび大阪府立大学のトラフィックデータに本手法を適用し，C&C サーバとの通信，DDoS 攻撃，Port scan 攻撃などの異常を検出できることを確認した．実験結果より，学習に用いるトラフィックデータに異常が含まれていない，もしくは異常が少数であれば異常を検出可能であることを確認できた．また，学習データに含まれていない通信パターンについても検出できることを確認できた．

今後の課題としては学習の際のパラメータや閾値の調整や実運用中のトラフィックに適用した際の実験結果の精査などがあげられる．

参考文献

[1] Liu, H., Lang, B., Liu, M. and Yan, H.: CNN and RNN based payload classification methods for attack detection, *Knowledge-Based Systems*, Vol.163, pp.332–341 (2019).
[2] Xu, C., Shen, J., Du, X. and Zhang, F.: An intrusion detection system using a deep neural network with gated recurrent units, *IEEE Access*, Vol.6, pp.48697–48707 (2018).
[3] 高田雄太，寺田真敏，松木隆宏，笠間貴弘，荒木粧子，畑田充弘：マルウェア対策のための研究用データセット—MWS 2018 Datasets，情報処理学会研究報告，Vol.2018-CSEC-82, No.38, pp.1–8 (2018).
[4] Sharafaldin, I., Lashkari, A.H. and Ghorbani, A.A.: Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization, *4th International Conference on Information Systems Security and Privacy* (*ICISSP*), pp.108–116 (2018).

浦川 侑之介

1995 年生．2018 年大阪府立大学現代システム科学域知識情報システム学類卒業．2020 年同大学大学院人間社会システム科学研究科博士前期課程修了．現在，日本電気株式会社．在学中，情報セキュリティに関する研究に従事．

青木 茂樹 （正会員）

1975 年生．1998 年大阪府立大学総合科学部卒業．2004 年同大学大学院工学研究科博士後期課程修了．同年熊本電波工業高等専門学校電子制御工学科助手．2006 年大阪府立大学総合教育研究機構講師，学術情報センター講師兼務．現在，同大学大学院人間社会システム科学研究科准教授，情報基盤センター准教授兼務．情報システム，情報セキュリティ，パターン認識に関する研究に従事．博士（工学）．電子情報通信学会，日本ロボット学会各会員．

宮本 貴朗 （正会員）

1962 年生．1985 年大阪府立大学総合科学部卒業．1987 年同大学大学院総合科学研究科修士課程修了．1988 年同大学大学院工学研究科博士後期課程退学．同年同大学計算センター助手．現在，同大学大学院人間社会システム科学研究科教授，情報基盤センター長．情報システム，情報ネットワーク，情報セキュリティに関する研究に従事．博士（工学）．電子情報通信学会，システム制御情報学会，IEEE，ACM 各会員．