情報処理学会論文誌 Vol.58 No.9 1450–1463 (Sep. 2017)

NIDS評価用データセット：Kyoto 2016 Datasetの作成

多田 竜之介[1,a)] 小林 良太郎[2,b)] 嶋田 創[3,c)] 高倉 弘喜[4,d)]

受付日 2016年11月25日, 採録日 2017年6月6日

概要：サイバー攻撃に対する防御法として，ネットワークベース侵入検知システム（Network-based Intrusion Detection System：NIDS）による攻撃の検知があげられる．これまでさまざまなNIDSの高性能化に関する研究が行われ，特に機械学習によるNIDSの高性能化が議論されてきた．これまでNIDSの評価用に使われてきたデータセットには，DARPA Intrusion Detection Data SetsやKDD Cup 1999 Data，Kyoto 2006+ Datasetなどがある．しかしながら，これらのデータセットは作成から期間が経過して最新の攻撃傾向を反映できていないことや，データの収集期間が短いといった問題がある．このため我々は，Kyoto 2006+ Datasetの作成に使用され，現在も稼働し続けているハニーポットのデータを使用して新たにKyoto 2016 Datasetを作成した．本論文では，作成したKyoto 2016 Datasetについて基本的な統計情報を調査した結果を述べる．加えて，Kyoto 2016 Datasetを基本的な機械学習手法によって分類した結果を示すことで，ネットワークベース侵入検知に関する研究を行う研究者へKyoto 2016 Datasetに機械学習を適用する場合の分類精度に関する指標を与える．

キーワード：侵入検知，ネットワークベース侵入検知システム，ハニーポットデータ，機械学習

Generating Kyoto 2016 Dataset for NIDS evaluation

Ryunosuke Tada[1,a)] Ryotaro Kobayashi[2,b)] Hajime Shimada[3,c)] Hiroki Takakura[4,d)]

Received: November 25, 2016, Accepted: June 6, 2017

Abstract: Network-based Intrusion Detection System (NIDS) is widely used to counteract cyber-attacks, and thus various studies have been conducted on high-performance NIDS based on machine learning techniques. In the field of intrusion detection, we have famous datasets for evaluation like: DARPA Intrusion Detection Data Sets, KDD Cup 1999 Data, and Kyoto 2006+ Dataset. However, these datasets do not reflect recent trends of cyber-attacks because it has created in well before, and moreover they lack long-period observation. For this reason, we generated Kyoto 2016 Dataset from the traffic data obtained from the honeypot that was used to generate Kyoto 2006+ Dataset and continues operation. In this paper, we describe basic statistical information of Kyoto 2016 Dataset. Furthermore, we provide the results classified by several basic machine learning methods to give guidepost of classification precision for researchers in the field of network based intrusion detection in the case of applying these methods to Kyoto 2016 Dataset.

Keywords: intrusion detection, network-based intrusion detection system, honeypot data, machine learning

[1] 豊橋技術科学大学
Toyohashi University of Technology, Toyohashi, Aichi 441–8580, Japan
[2] 工学院大学
Kogakuin University, Shinjuku, Tokyo 163–8677, Japan
[3] 名古屋大学
Nagoya University, Nagoya, Aichi 464–8601, Japan
[4] 国立情報学研究所
National Institute of Informatics, Chiyoda, Tokyo 101–8430, Japan
[a)] tada2016@ppl.cs.tut.ac.jp
[b)] ryo.kobayashi@cc.kogakuin.ac.jp
[c)] shimada@itc.nagoya-u.ac.jp
[d)] takakura@nii.ac.jp

1. はじめに

サイバー攻撃に対する防御法として，ネットワークベース侵入検知システム（Network-based Intrusion Detection System：NIDS）による攻撃の検知があげられる．NIDSはネットワーク通信を監視し，悪性の通信を発見すると通知するものである．これまでさまざまなNIDSの高性能化に関する研究が行われ，特に機械学習によるNIDSの高性能化は近年において特に注目されている分野である [1], [2], [3], [4], [5], [6], [7], [8], [9], [10].

© 2017 Information Processing Society of Japan

Ambusaidi らは，IDS に有効な特徴選択手法を考案し，その特徴選択手法と Least Square Support Vector Machine（LSSVM）を組み合わせた IDS を提案した [2]．Om らは，k-Means と k 近傍法，ナイーブベイズを組み合わせたハイブリッド IDS を提案した [3]．Hosseini らは，蟻コロニー最適化（Ant Colony Optimization：ACO）とホタルアルゴリズム（Firefly Algorithm）を組み合わせて最適化したサポートベクタ回帰（Support Vector Regression：SVR）を用いて，Denial of Service（DoS）侵入攻撃の検知手法を提案した [4]．Eskin らは，教師なし学習手法であるクラスタベース分類法と k 近傍法，One Class SVM を使用して，アノマリ検知による侵入検知手法を提案した [5]．RT らは，Software-defined Networking（SDN）技術によって制御されるネットワークが Distributed Denial of Service（DDoS）攻撃に対して脆弱であることをあげ，Support Vector Machine（SVM）を使用した DDoS 攻撃の検知手法を提案した [6]．Mukkamala らは，ニューラルネットワークとサポートベクタマシンを用いた IDS を構築して比較を行った [7]．Masarat らは，IDS のための改良型ランダムフォレストアルゴリズムを提案した [8]．Stein らは，遺伝的アルゴリズムを用いた特徴選択と決定木を組み合わせた侵入検知手法を提案した [9]．Amor らは，IDS における決定木とナイーブベイズの分類性能の比較を行った [10]．

これらの研究の推進においては，正常通信と悪性通信の両方を含んだ通信のパケットキャプチャファイルや，通信をセッション単位でいくつかの特徴量にサマライズした通信データセットに対して提案アルゴリズムの効果を確認する形態がある．いくつかの研究組織においては，その分野で研究を実施する者の便宜をはかるため，その研究組織において採取とサマライズを行った通信データセットを提供していることがある．この侵入検知の研究領域においてこれまで NIDS の評価用に使われてきたデータセットには，DARPA Intrusion Detection Data Sets [11] や KDD Cup 1999 Data [12]，Kyoto 2006+ Dataset [13] などがある．これらは，先に示した研究 [2], [3], [4], [5], [6], [7], [8], [9], [10] においても評価のために使用された．しかしながらこれらのデータセットは，作成から期間が経過しており最新の攻撃傾向を反映できていないことや，データの収集期間が短いといった問題がある．加えて，一部のデータセットは実験用のネットワーク上で悪性通信を擬似的に作り出して作成されており，実際のネットワーク環境を反映できていない問題もある．特に評価用のデータセットとしては，最新の攻撃傾向を反映できていることやデータの収集期間が長いことが非常に重要である．現在のところ，我々の知る限りこれらを満たすデータセットは存在しない．また，多くの機関ではセキュリティポリシ [14] などにより，自組織のものであっても通信監視によって得られたデータの研究への利用を制限するようになり，研究者がサイバー攻撃のデータを入手しにくくなっている．

この現状を受け，我々は Kyoto 2006+ Dataset の作成に使用され，現在も稼働し続けているハニーポットのデータを使用して新たなデータセットを作成した．我々はこのデータセットを Kyoto 2016 Dataset と名付けた．Kyoto 2016 Dataset は以下の特徴を持つ．

- 最新の攻撃傾向を含む
- 実際のトラフィックから作られる（ハニーポットデータ）
- 長い収集期間（2006 年 11 月から 2015 年 12 月までのおよそ 10 年間）

これらの特徴を持ち合わせた Kyoto 2016 Dataset は，侵入検知の研究領域における研究活動に大きく貢献できるものであると考えている．

本論文では，作成した Kyoto 2016 Dataset について基本的な統計情報を調査した結果を述べる．加えて，Kyoto 2016 Dataset を基本的な機械学習手法によって分類した結果を示すことで，ネットワーク侵入検知に関する研究を行う研究者へ Kyoto 2016 Dataset に機械学習を適用する場合の分類精度に関する指標を与える．

以下，2 章では既存の NIDS の評価用データセットについて述べる．3 章では，Kyoto 2016 Dataset の概要について述べる．4 章では，いくつかの観点から調査した Kyoto 2016 Dataset の統計情報について述べる．5 章では，Kyoto 2016 Dataset を基本的な機械学習手法によって分類した結果について述べる．6 章では結論を述べる．

2. 関連研究

現在，NIDS の評価用データセットとしていくつかのデータセットが公開されている．ここでは，契約を交わすことなく使用でき，一般公開されているものについて述べる．

2.1 DARPA Intrusion Detection Data Sets

DARPA Intrusion Detection Data Sets は，MIT Lincoln Laboratory が作成し 1998 年から公開しているデータセットである [11]．1998 年，1999 年，2000 年の 3 年分が公開されている．このデータセットは，実験用に作成されたネットワーク環境で正常な通信の中に意図的に悪性の通信を混入させて作られたものである．含まれる攻撃の種類としては，DoS 攻撃やバッファオーバーフロー攻撃などがある．通信データは TCPDUMP 形式で公開されているため，非常に柔軟に利用できるという特徴がある．

2.2 KDD Cup 1999 Data

KDD Cup 1999 Data は，University of California Irvine, Machine Learning Repository で公開されているデータセットである [12]．このデータセットは，1998 DARPA Intrusion Detection Data Set の通信データをもとに作成されたものである．DARPA Intrusion Detection Data Sets とは

© 2017 Information Processing Society of Japan

表 1 Kyoto 2006+ Dataset の特徴量一覧
Table 1 List of Kyoto 2006+ Dataset session features.

属性名	概要
Duration	セッションの長さ（秒）
Service	サービスの種類（http, smtp, ssh など）
Source_Bytes	送信バイト数
Destination_Bytes	受信バイト数
Count	過去 2 秒間のセッションのうち現在のセッションと宛先 IP アドレスが同じ数
Same_srv_rate	Count 特徴で該当したセッションのうち現在のセッションとサービスの種類が同じ割合
Serror_rate	Count 特徴で該当したセッションのうち “SYN” エラーが起こった割合
Srv_serror_rate	過去 2 秒間のセッションで現在のセッションとサービス種類が同じセッションのうち，“SYN” エラーが起こった割合
Dst_host_count	宛先ポートが同じ過去の 100 セッションのうち，現在のセッションと送信元 IP アドレスと宛先 IP アドレスが同じ数
Dst_host_srv_count	宛先ポートが同じ過去の 100 セッションのうち，現在のセッションと宛先 IP アドレスとサービス種類が同じ数
Dst_host_same_src_port_rate	Dst_host_count 特徴で該当したセッションのうち現在のセッションと送信元ポートが同じ割合
Dst_host_serror_rate	Dst_host_count 特徴で該当したセッションのうち “SYN” エラーが起こった割合
Dst_host_srv_serror_rate	Dst_host_srv_count 特徴で該当したセッションのうち “SYN” エラーが起こった割合
Flag	セッション終了時の接続の状態
IDS_detection	IDS の検知結果
Malware_detection	アンチウイルスソフトウェアの検知結果
Ashula_detection	シェルコードやエクスプロイトコードの検知結果
Label	セッションが正常か攻撃かを示すラベル
Source_IP_Address	月ごとにサニタイズされた送信元 IP アドレス
Source_Port_Number	送信元ポート番号
Destination_IP_Address	月ごとにサニタイズされた宛先 IP アドレス
Destination_Port_Number	宛先ポート番号
Start_Time	日の始めを始点としたセッションの開始時刻
Duration	セッションの長さ（秒）

異なり，通信データをセッション単位で扱い，各セッションについてセッションの前後関係をもとに算出した特徴量を追加して 42 次元のベクトルの形式で記録されている．セッションの前後関係から算出される特徴量には，過去 2 秒間のセッションという単位や同一ホスト間の過去 100 セッションという単位でセッションをひとまとめにし，集計して得られるものが含まれる．ペイロードはもちろん，IP アドレスやポート番号も含まれない．すべてのデータに正常か攻撃かを示すラベルが付けられ，攻撃の場合には攻撃の種類も示される．学習用と評価用にそれぞれデータが用意されており，評価用のデータには学習用に含まれない攻撃データが含まれるという特徴がある．

2.3 Kyoto 2006+ Dataset

Kyoto 2006+ Dataset は，当時 NIDS の評価用データセットとして広く用いられていた KDD Cup 1999 Data が古くなったことを受け，新たな NIDS の評価用データセットとして作成されたものである [13], [15]．通信データとして，京都大学に設置されているハニーポットのデータを使用して作成されたため，Traffic Data from Kyoto University’s Honeypots という名前で公開されている．このハニーポットは現在も稼働し続けているものの，公開されているのは 2006 年 11 月から 2009 年 8 月までの通信データから作成されたものである．データの作成方法は KDD Cup 1999 Data に準拠しており，通信データをセッション単位で扱いその前後関係から特徴量を算出して作成された．Kyoto 2006+ Dataset の特徴量は，KDD Cup 1999 Data の特徴量の一部である 14 種類と，独自に追加した 10 種類の特徴量を合わせた 24 種類である．**表 1** に，各特徴量を基本 14 特徴量，追加 10 特徴量の順に示す．

3. Kyoto 2016 Dataset の作成

Kyoto 2016 Dataset を作成するにあたり，基本的には Kyoto 2006+ Dataset の作成手順を踏襲し，データ形式も同じになるようにした[*1]．一方で，作成手順とデータ形式の一部では変更を加えた．加えた変更について以下の節で述べる．

3.1 セッション構成ツールの変更

Kyoto 2006+ Dataset では通信をセッション単位で扱う

[*1] Kyoto 2006+ Dataset の作成手順とデータ形式に関する詳細は文献 [13] で述べられている．

© 2017 Information Processing Society of Japan

ため，ハニーポットデータをセッション単位に変換する必要がある．ハニーポットデータは pcap 形式で記録されており，おおよそ 1 時間ごとに分割して保存されている．Kyoto 2006+ Dataset の作成時には，以下の手順でハニーポットデータからセッション単位への変換が行われた．

(1) mergecap を使用して 1 日超分のハニーポットデータを結合
(2) editcap を使用して正確に 1 日分を切り出すとともにパケットの重複を除去
(3) Bro IDS を使用して 1 日分の pcap ファイルからセッションを抽出

mergecap，および editcap は，ネットワーク解析ツールとして利用される Wireshark に付属するツールである [16]．mergecap はパケットキャプチャファイルの結合を行える．editcap は，パケットキャプチャファイルの重複を除去することや，ファイルサイズ単位，時間単位での分割を行える．Bro IDS はネットワーク監視用のソフトウェアであり，異常な通信の発見や，各種プロトコル，サービスに関する解析を行える [17]．

Kyoto 2016 Dataset の作成を行うにあたり，Kyoto 2006+ Dataset 作成時と同様の手順でセッションを抽出した．しかしながら，セッション構成に用いる Bro IDS について，作成当時（2016 年 10 月時点）の最新安定版であったバージョン 2.4.1 を使用したため，Kyoto 2006+ Dataset の作成時に使われた Bro IDS のバージョン 1.2.1 とは異なる[*2]．データセットの作成に使用する Bro IDS のバージョンを変更した理由は大きく分けて 2 つある．

まず，ファイルサイズが 2 GB を超えるパケットキャプチャファイルの処理に対応したことがあげられる．Kyoto 2006+ Dataset の作成時に使用された Bro IDS 1.2.1 では，2 GB を超えるパケットキャプチャファイルを扱えないという制約があったため，1 日分のパケットキャプチャファイルが 2 GB を超えてしまうとセッションを抽出できない問題があった．このため，Kyoto 2006+ Dataset では，パケットキャプチャファイルが 2 GB を超えてしまった数日分（2006 年 12 月の 11，15 日，2007 年 8 月 10 日など）のデータが欠損している．加えて，Kyoto 2006+ Dataset で扱われた期間以降（2009 年 9 月〜），ハニーポットへ到達する通信の増加にともなって 1 日分のパケットキャプチャファイルが 2 GB を超える場合が多くなった．このような背景から，Kyoto 2016 Dataset を作成する上で 2 GB 超のファイルを扱えない Bro IDS 1.2.1 では不十分であり，2 GB 超ファイルの処理に対応した Bro IDS 2.4.1 が必要だったのである．

加えて，より詳細な解析が行えるようになったことがあげられる．文献 [18] で詳しく述べられているように，Bro IDS 2.4.1 では通信のサービス種類の推定や ICMP セッションの解析において Bro IDS 1.2.1 と比較して改良が施されており，より正確なセッション抽出が行えるようになっている．

3.2 特徴量の見直しと不具合の修正

Kyoto 2016 Dataset では一部の特徴量の見直しを行った．

まず，冗長であったセッションの接続時間を示す特徴量を 1 つに統合した．Kyoto 2006+ Dataset では，1 番目の特徴量（以降，Duration と表記）と 24 番目の特徴量（以降，Duration2 と表記）がどちらもセッションの接続時間を示すものであった．実際には，Duration2 がセッションの接続時間を小数点以下 6 桁の精度で記録したものであり，Duration は Duration2 を小数点以下 2 桁まで丸めただけのものであった．よって，Kyoto 2016 Dataset では Duration を削除して Duration2 のみを残すように変更した．

次に，セッションのプロトコルを特徴量として追加した．Kyoto 2006+ Dataset では，KDD Cup 1999 Data では特徴量に含まれるセッションのプロトコルが特徴量として含まれなかった．Kyoto 2016 Dataset では，セッションのプロトコルとして “tcp”，“udp”，“icmp” のいずれかが記録されるように変更した．

さらに，セッションのサービス種類をマスクしないように変更した．Kyoto 2006+ Dataset では，セッションのサービス種類は全体を通して一意な番号に置き換えられてしまい，セッションがどのようなサービスの通信であるかは不明であった．Kyoto 2016 Dataset では，セッションのサービス種類は “smtp” や “http” というようにマスクなしで記録されるように変更した．

加えて，Kyoto 2006+ Dataset で確認された不具合の修正も行った．

まず，クラスラベルの修正を行った．Kyoto 2006+ Dataset では正常を示す “1”，既知攻撃を示す “−1”，未知攻撃を示す “−2” の 3 種類のクラスラベルが存在する．Kyoto 2006+ Dataset では悪性のセッションのうち，“IDS_detection” 特徴量と “Malware_detection” 特徴量では悪性であると検知しなかったものの，“Ashula_detection” 特徴量では検知したものを未知攻撃としている [13]．しかしながら，クラスラベルを特徴量として追加するプログラムに不具合があり，未知攻撃であるにもかかわらず既知攻撃のクラスラベルが割り振られるセッションが存在した．Kyoto 2016 Dataset では正しくクラスラベルが割り振られるように修正した．

次に，過剰に区切り文字が挿入される不具合の修正を行った．Kyoto 2006+ Dataset はタブ区切りで記録されており，タブで区切ると 24 列に分かれるはずである．しかしながら，Kyoto 2006+ Dataset の一部の行ではプログラ

[*2] セッション構成ツールの差異に関する詳細とセッション構成ツールの差異がデータセットに与える影響については文献 [18] で述べられている．

© 2017 Information Processing Society of Japan

ムの不具合によって過剰にタブが挿入されている箇所があるため，単純にタブで区切ると列数が 24 を超えてしまう場合がある．Kyoto 2016 Dataset では過剰な区切り文字が挿入されないように修正した．

4. Kyoto 2016 Dataset の統計情報

Kyoto 2016 Dataset を作成するにあたり，2006 年 11 月から 2015 年 12 月までのハニーポットデータを使用した．2006 年 11 月から 2015 年 12 月までの 3,348 日間のうち，機器のトラブルやメンテナンスによって正常に通信を記録できなかった期間が 80 日間あり，データセットに含まれるのは 3,268 日間のデータである．全 3,268 日間のセッションデータについて，いくつかの観点から統計情報の調査を行った．

4.1 観測されたセッション数の推移

全期間に観測されたセッション数と，1 日あたりの平均観測セッション数を**表 2** に示す．全期間でおよそ 8 億セッションが観測され，そのうち正常なセッションがおよそ 1.6 億セッション，悪性のセッションがおよそ 6.5 億セッションであった．平均すると，1 日あたりおよそ 25 万セッションが観測され，そのうち正常なセッションがおよそ 5 万セッション，悪性のセッションがおよそ 20 万セッションであった．

図 1 には，正常，既知攻撃，未知攻撃それぞれの月ごとの観測セッション数の推移を示す．ハニーポットへ到達した悪性通信は Kyoto 2006+ Dataset で扱われた期間以降（2009 年 9 月～）増加し始め，ピークとなった 2010 年 8 月には 1 カ月間で 14,054,993 セッションを観測した．その後悪性セッションは減少し続け，2012 年 9 月には 1 カ月間で観測した悪性セッション数が 1,747,691 まで減少した．いったんはハニーポットへ到達する悪性通信の数が沈静化したようにみえたものの，2013 年 5 月頃から再び増加し始めた．2013 年 6 月以降，2015 年 12 月まで毎月およそ 1,000 万セッション程度の悪性通信がハニーポットへ到達していた．2015 年 3 月の悪性セッション数が異常に多いのは，2015 年 3 月 27 日に 1 日で 17,812,649 セッションというバースト的な悪性セッションの増加が起こったためである．

表 2 Kyoto 2016 Dataset の観測セッション数
Table 2 Number of observed sessions of Kyoto 2016 Dataset.

	セッション数	1 日あたりの平均セッション数
正常	160,873,849	49,227
既知攻撃	640,618,555	196,028
未知攻撃	4,603,220	1,409
合計	806,095,624	246,663

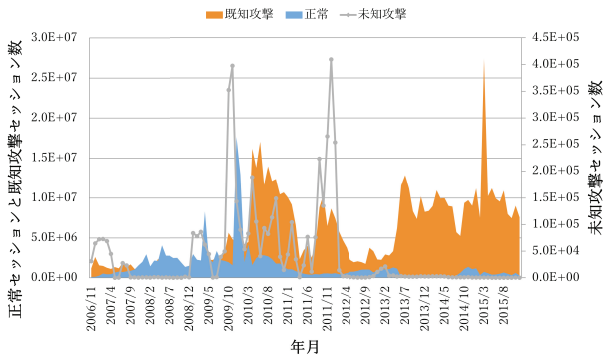

図 1 月別観測セッション数の推移
Fig. 1 Changes in number of observed sessions per month.

バースト的な悪性セッションの増加が起こった 2015 年 3 月 27 日のセッションについて調査を行った．その結果，2015 年 3 月 27 日に観測された 17,812,649 の悪性セッションのうち，およそ 99%にあたる 17,574,098 セッションが単一の IP アドレス，単一のポートから発信された TCP セッションであったことを確認した．通信は京都大学内の 271 のホスト，65,533 種類のポートをまんべんなく宛先として行われたものであった．最初の通信は午前 10 時 28 分 47 秒に開始され，最後の通信は午後 1 時 27 分 27 秒であったため，およそ 3 時間の間におよそ 1,700 万セッションの通信が観測されたことになる．MaxMind 社の GeoIP2 データベース [19] によれば，通信の発信元はセーシェルであり，オランダのホスティング会社が管理するものであった．

4.2 攻撃発信国の割合

悪性セッションの送信元 IP アドレスに基づいて悪性通信の発信国を調査した．IP アドレスの割当て国の推定には，MaxMind 社の GeoLite2 データベース [20] の 2017 年 3 月 2 日版を使用した[*3]．

図 2 に，悪性通信の発信国の割合を示す．ハニーポットへ到達した悪性通信の発信国は 240 カ国であった．悪性通信の主要な発信国としては中国，日本，台湾，アメリカ，ロシア，ブラジル，韓国などがあり，合わせて全体のおよそ 7

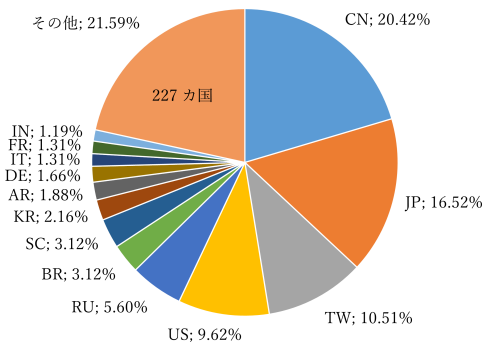

図 2 攻撃発信国の分布
Fig. 2 National distribution of attack source ip addresses.

[*3] IP アドレスブロックの移譲などにより発信国が当時と異なることがあるものの，その影響は軽微であると考えている．

© 2017 Information Processing Society of Japan

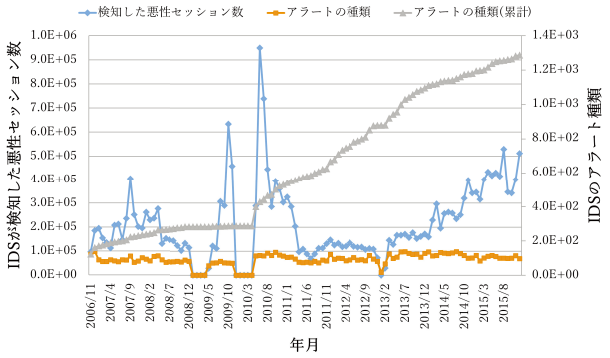

図 3 月別 IDS アラート数の推移
Fig. 3 Changes in number of observed IDS alerts per month.

割を占めることを確認した．特に中国から発信される悪性通信が非常に多く，Kyoto 2016 Dataset に含まれるおよそ 6.5 億の悪性セッションのうちおよそ 1.3 億セッションが中国から発信されたものであった．これらの国々から悪性通信が発信される割合が高い傾向は，Kyoto 2006+ Dataset 作成当時の調査と同様である [13]．また，今回の調査では，Kyoto 2006+ Dataset 作成当時の調査では大きな割合を占めなかったセーシェルからの悪性通信が 3.12%を占めることを確認した．このうち，4.1 節で述べたが，セーシェルからの悪性通信のうちおよそ 87%は DoS 攻撃によるものであった．

なお，Kyoto 2006+ Dataset 作成当時の調査でも述べられたように，Kyoto 2016 Dataset においても日本から発信されたとされる悪性セッションの大半はネットワーク機器の不適切な設定によって起こった通信であった．日本から発信された悪性とされるおよそ 1 億セッションのうち，上位 100IP アドレスについての調査でおよそ 7,000 万セッションは京都大学の IP アドレスで行われた通信であり真に悪性のものではなかった．

4.3 IDS アラート数の推移

Kyoto 2016 Dataset に含まれる IDS アラートは 2009 年までが Symantec 社 [21] の IDS によるもので，2010 年以降は Sourcefire 社（現 Cisco 社）[22] の IDS によるものである．データセットに含まれるおよそ 6.5 億の悪性セッションのうち，23,372,784 セッションで IDS がアラートを発した．平均すると，1 日あたりおよそ 7,152 セッションで IDS がアラートを発したことになる．また，全期間で観測された IDS アラートは 1,287 種類であり，1 日ごとに観測された IDS アラート種類を平均するとおよそ 37 種類であった．

図 3 に月ごとの IDS アラート数と IDS アラート種類の推移を示す．水色の線が月ごとの IDS アラートの数，橙色の線が月ごとの IDS アラートの種類，灰色の線がその月までの累計の IDS アラートの種類を示す．2009 年 1 月から 2009 年 4 月までと，2009 年 12 月から 2010 年 4 月までは IDS のトラブルかメンテナンスのために記録が停止しており，データが存在しないため IDS アラート数が 0 である．

表 3 2010 年 6 月の上位の IDS アラートの一部
Table 3 Part of the top IDS alerts observed in June 2010.

記号	アラート種類	検出数
(a)	483-1-6, 384-1-5	521,006
(b)	2050-1-14, 4990-1-9, 2004-1-13	161,842
(c)	449-1-6	70,180
(d)	384-1-5	49,681

まず，IDS アラートの数に注目すると，2010 年 6 月の IDS アラート数が極端に多く 1 カ月で 949,682 件に上った．2010 年 6 月に IDS アラートが発されたセッションについて調査したところ，いくつかのアラートが大半を占めることを確認した．**表 3** に 2010 年 6 月の上位の IDS アラートの一部を示す．上位の IDS アラートのうち，まず記号 (a)，(d) は ICMP に関係するアラートであり，2010 年 6 月の IDS アラートの大半を占める．2010 年 6 月の IDS アラートは，全体のおよそ 76%におよぶ 724,575 件が ICMP セッションに対しての警告であった．アラートの内容を詳しくみると，(d) に含まれるアラート “384-1-5” は ICMP エコー要求を警告するものである [23]．(a) では，アラート “384-1-5” に加えて “483-1-6” も発されており，“483-1-6” は CyberKit というネットワーク管理用のソフトウェアから ICMP エコー要求が送られたことを警告するものである [24]．次に，記号 (b) は TCP に関係するアラートであり，アラート “2050-1-14” および “4990-1-9”，“2004-1-13” は Microsoft SQL Server 2000 の脆弱性を悪用しようとする試みに対しての警告である [25], [26], [27]．そして，記号 (c) に含まれるアラート “449-1-6” は traceroute の試み（最大ホップ数への到達）を警告するものである [28]．

次に，累計の IDS アラート種類に注目すると，IDS を Sourcefire 社のものに変更してからは増加し続けていることが確認できた．このことから，ハニーポットへは年々過去の悪性通信とは異なる通信が到達しており，Kyoto 2016 Dataset の悪性通信の傾向は年々変化していると考えることができる．実際に，観測された IDS のアラートには数年以内に作成されたルールによるものが含まれていた．

たとえば，アラート “31978-1-1” は 2014 年 9 月に作成されたルールによるもので，bash の脆弱性を悪用しようとする試みに対しての警告である [29]．この bash の脆弱性は “CVE-2014-7169” として知られるもので，任意のコマンドを実行できてしまう脆弱性である [30]．また，アラート “31136-1-1” は 2014 年 6 月に作成されたルールによるもので，マルウェア “Win.Trojan.ZeroAccess” に対しての警告である [31]．さらに，アラート “28556-1-1” および “28556-1-2” は，DNS Amplification Attack の試みに対しての警告である [32]．このルールは，2013 年 3 月の DNS Amplification Attack に関する US-CERT の発表に

© 2017 Information Processing Society of Japan

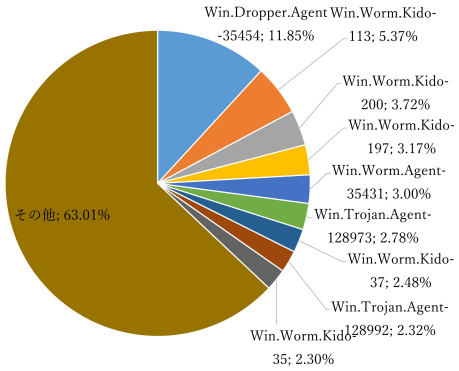

図 4 アンチウイルスソフトウェアのアラートの割合
Fig. 4 Percentage of AV alerts.

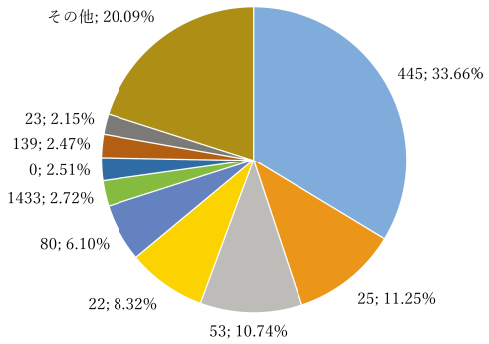

図 5 宛先ポートの割合
Fig. 5 Percentage of destination ports.

対応するものである [33]．このほか，アラート “34464-1-1” は 2015 年 5 月に作成されたルールによるもので，ASUS 社製無線ルータのファームウェアに存在する脆弱性を悪用しようとする試みに対しての警告である [34]．このファームウェアの脆弱性は “CVE-2014-9583” として知られるもので，特定の条件下で任意のコマンドを実行できてしまう脆弱性である [35]．このように，Kyoto 2016 Dataset にはさまざまな機器を対象としたさまざまな種類の悪性通信に関するデータが含まれている．

4.4 アンチウイルスソフトウェアのアラートの割合

ハニーポットの通信はアンチウイルスソフトウェアである ClamAV [36] によって監視されている．Kyoto 2016 Dataset のおよそ 6.5 億の悪性セッションのうち，464,358 セッションで ClamAV がアラートを発した．平均すると，1 日あたりおよそ 142 セッションで ClamAV がアラートを発したことになる．また，全期間で観測された ClamAV のアラートは 1,017 種類であり，1 日ごとに観測された ClamAV のアラート種類を平均するとおよそ 34 種類であった．

図 4 に全期間での ClamAV のアラートの割合を示す．観測された ClamAV のアラートは主に “Worm” や “Trojan”，“Exploit” に関するものであった．単独で最も高い割合を占めたのは，“Win.Dropper.Agent-35454” についてのアラートであった．Dropper は攻撃の初期段階で投入されるものであるために割合が高くなったと考えられる．全体的に最も高い割合を占めたのは，“Win.Worm.Kido-XXX” についてのアラートであった．Kido は Conficker として知られる，2008 年 11 月頃から爆発的な感染を引き起こしたワームである [37]．“Win.Worm.Kido-XXX” についてのアラートは 274,033 件あり，ClamAV によるアラート全体のおよそ 6 割を占めた．世界的には Conficker への感染は沈静化したものの，ハニーポットへはいまだに Conficker が到達している．

4.5 宛先ポートの分布

Kyoto 2016 Dataset のおよそ 6.5 億の悪性セッションについて，宛先ポートに関して集計した．宛先ポートの種類を平均すると，1 日あたりおよそ 2,240 種類のポートが宛先に指定されたことを確認した．

図 5 に全期間での宛先ポートの割合を示す．高い割合を占めた宛先ポートは，445 (SMB)，25 (SMTP)，53 (DNS)，22 (SSH)，80 (HTTP) などであり，これらを合わせると全体のおよそ 7 割を占めた．基本的に，高い割合を占めた宛先ポートの傾向は Kyoto 2006+ Dataset 作成時の調査と同様であった [13]．

詳しくみると，特に 445 番を宛先とするセッションが多く，観測されたセッションはおよそ 2 億にも上った．また，25 番と 53 番の割合が高いのは，ハニーポット内でメールサービスと DNS サービスが稼働しているためであると考えられる．

なお，その他の宛先ポートで通信が観測された理由の一つとして，動的に待ち受け状態を変化させるハニーポット [38] が稼働していることがあげられる．

5. Kyoto 2016 Dataset の機械学習による分類結果

Kyoto 2016 Dataset は NIDS の評価用データセットであり，特に機械学習を用いた NIDS を想定している．そこで，Kyoto 2016 Dataset を複数の基本的な機械学習手法によって分類し，分類結果から Kyoto 2016 Dataset の利用者へ，Kyoto 2016 Dataset を機械学習手法によって分類する場合の分類精度に関する指標を提供できると考える．

分類に用いる機械学習手法として，教師あり学習手法から，決定木 (Decision Tree：DT)，ランダムフォレスト (Random Forest：RF)，ナイーブベイズ (Naïve Bayes：NB)，サポートベクタマシン (Support Vector Machine：SVM) を使用した．また，教師なし学習手法から，k 近傍法 (k-Nearest Neighbor：k-NN)，One Class SVM (OCSVM) を使用した．k 近傍法による分類は，Eskin らによって示されたアルゴリズムを使用した [5]．分類したいサンプルの k 近傍との距離の平均値を外れ値度とし，しきい値を超えたかどうかで分類するというものである．各手法の実装には

© 2017 Information Processing Society of Japan

表 4 評価用データセットのサブセット
Table 4 Subset of evaluation dataset.

期間	ラベル
2006 年 11 月～2006 年 12 月	A
2007 年 1 月～2007 年 2 月	B
2007 年 3 月～2007 年 4 月	C
2007 年 5 月～2007 年 6 月	D
2007 年 7 月～2007 年 8 月	E
2007 年 9 月～2007 年 10 月	F
2007 年 11 月～2007 年 12 月	G
2008 年 1 月～2008 年 2 月	H
2008 年 3 月～2008 年 4 月	I
2008 年 5 月～2008 年 6 月	J
2008 年 7 月～2008 年 8 月	K
2008 年 9 月～2008 年 10 月	L
2008 年 11 月～2008 年 12 月	M

表 5 学習時のサンプル数
Table 5 Number of samples in learning.

手法	攻撃サンプル	正常サンプル
DT	10,000	10,000
RF	10,000	10,000
NB	10,000	10,000
SVM	10,000	10,000
k-NN	0	2,000
OCSVM	100	10,000

Python の機械学習ライブラリである scikit-learn [39] を使用した．評価に使用したサブセットは，**表 4** に示すように，2006 年 11 月から 2008 年 12 月のデータについて，それぞれ 2 カ月ごとに 10 万件ずつ攻撃サンプルと正常サンプルに分けて無作為抽出したものである．

5.1 学習・評価時の条件

表 4 に示したデータセットのサブセットを使用して機械学習による分類を行った．分類テストは表 4 に示したラベルを使って，A で学習して B を分類，B で学習して C を分類，… というような操作を L で学習して M を分類するまで繰り返した．

サンプルの正規化として，ナイーブベイズとサポートベクタマシン，One Class SVM では scikit-learn の StandardScaler クラスによる変換を行った．k-NN では同じく scikit-learn の MinMaxScaler クラスによる変換を行った．決定木とランダムフォレストでは正規化が必要ないため行わなかった．StandardScaler クラスは，サンプルの平均を 0，分散を 1 にする標準化と呼ばれる変換を行う．MinMaxScaler クラスは，サンプルの各特徴量の値の範囲を 0 から 1 の範囲に変換する．

学習データとしては，サブセットからさらに**表 5** に示す数だけサンプルを無作為抽出したものを使用した．評価データとしては，すべての手法で攻撃サンプルと正常サンプルを 1 万件ずつ無作為抽出したものを使用した．使用した特徴は，2.3 節で示した基本 14 特徴量のうち，カテゴリ値特徴である “Service” と “Flag” を除いた 12 特徴量である．

以上の設定で 2 通りの評価を行った．学習時には各手法で分類器のパラメータを設定することができ，それによって性能が変化する．最適な分類器のパラメータは学習データと評価データに依存するため，使用するサブセットの組み合わせが変われば同時に変化してしまう．ネットワークトラフィックは時間の経過とともに性質が変化する恐れがあり，一定期間ごとにパラメータの再調整が必要だと考えられる．このため，パラメータの調整による分類への影響を評価する必要があった．まず 1 つ目として，全期間で同一のパラメータを使用する評価を行った．各手法ごとに，全期間の平均検知率が最も高かったパラメータを使用した．次に 2 つ目として，各組み合わせごとにパラメータを調整する評価を行った．各手法，各期間ごとに平均検知率が最も高かったパラメータを使用した．いずれの評価でも，全期間での平均誤検知率が 20%未満で平均検知率が最高になるように調整を行った．それぞれの評価結果から，正解率（Accuracy），適合率（Precision），検知率（True Positive Rate：TPR），誤検知率（False Positive Rate：FPR）を算出した．この際，無作為抽出による結果のばらつきを防ぐため 10 回試行時の平均値を結果として使用した．各指標の定義は以下のとおりである．

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$
$$Precision = \frac{TP}{TP + FP}$$
$$TPR = \frac{TP}{TP + FN}$$
$$FPR = \frac{FP}{FP + TN}$$

5.1.1 パラメータ固定時の評価結果

決定木，ランダムフォレスト，ナイーブベイズ，サポートベクタマシン，k 近傍法，One Class SVM のパラメータ固定時の正解率，適合率，検知率，誤検知率を，それぞれ**表 6**，**表 7**，**表 8**，**表 9**，**表 10**，**表 11** に示す．

教師あり学習の結果である表 6～表 9 の結果に注目すると，ナイーブベイズを除けば平均して検知率が 90%以上でありながら誤検知率は 10%以下であり，正解率が 90%を超えた．特にランダムフォレストでは分類精度が非常に高く，平均して検知率が 95.66%でありがなら誤検知率を 6.35%に抑えた．期間ごとにみると期間（K, L）の組合せで最も精度が高く，検知率が 98.78%と非常に高いにもかかわらず誤検知率を 1.08%に抑えた．

一方，教師なし学習の結果である表 10 と表 11 に注目すると，平均して検知率が 80%に満たないうえに誤検知率も今回の上限値である 20%に非常に近い結果となった．各

© 2017 Information Processing Society of Japan

表 6 各期間における DT の平均分類精度（固定）
Table 6 Average classification accuracy of DT under individual period (fixed parameter).

期間	正解率	適合率	検知率	誤検知率
A, B	93.74%	97.40%	89.89%	2.40%
B, C	95.85%	93.73%	98.27%	6.58%
C, D	90.50%	97.55%	83.08%	2.09%
D, E	92.18%	94.78%	89.27%	4.92%
E, F	96.55%	96.30%	96.83%	3.73%
F, G	97.39%	97.61%	97.16%	2.38%
G, H	93.68%	91.74%	96.06%	8.71%
H, I	95.77%	98.21%	93.25%	1.70%
I, J	85.45%	78.82%	97.24%	26.33%
J, K	97.58%	98.08%	97.07%	1.90%
K, L	97.55%	98.59%	96.47%	1.38%
L, M	95.71%	97.04%	94.30%	2.89%
平均	94.33%	94.99%	94.07%	5.42%

表 7 各期間における RF の平均分類精度（固定）
Table 7 Average classification accuracy of RF under individual period (fixed parameter).

期間	正解率	適合率	検知率	誤検知率
A, B	93.99%	98.08%	89.73%	1.75%
B, C	96.47%	94.16%	99.10%	6.15%
C, D	90.67%	98.23%	82.84%	1.49%
D, E	92.42%	94.71%	89.86%	5.02%
E, F	97.47%	96.35%	98.67%	3.74%
F, G	98.48%	98.11%	98.86%	1.90%
G, H	92.15%	87.89%	97.90%	13.61%
H, I	97.62%	98.46%	96.76%	1.51%
I, J	81.45%	73.65%	98.92%	36.02%
J, K	98.72%	98.55%	98.90%	1.46%
K, L	98.85%	98.92%	98.78%	1.08%
L, M	97.52%	97.47%	97.57%	2.53%
平均	94.65%	94.55%	95.66%	6.35%

表 8 各期間における NB の平均分類精度（固定）
Table 8 Average classification accuracy of NB under individual period (fixed parameter).

期間	正解率	適合率	検知率	誤検知率
A, B	81.93%	89.81%	72.30%	8.43%
B, C	90.49%	89.50%	91.75%	10.77%
C, D	89.14%	88.91%	89.43%	11.15%
D, E	88.92%	86.50%	92.25%	14.40%
E, F	84.21%	87.09%	80.32%	11.91%
F, G	85.25%	84.17%	86.83%	16.33%
G, H	75.57%	70.59%	87.66%	36.52%
H, I	85.24%	83.93%	87.18%	16.70%
I, J	67.87%	63.10%	86.06%	50.32%
J, K	81.82%	82.95%	80.37%	16.72%
K, L	92.17%	88.13%	97.45%	13.12%
L, M	85.64%	82.20%	90.96%	19.69%
平均	84.02%	83.07%	86.88%	18.84%

表 9 各期間における SVM の平均分類精度（固定）
Table 9 Average classification accuracy of SVM under individual period (fixed parameter).

期間	正解率	適合率	検知率	誤検知率
A, B	88.37%	92.29%	83.73%	7.00%
B, C	93.22%	90.05%	97.17%	10.73%
C, D	94.76%	93.35%	96.41%	6.88%
D, E	88.75%	87.38%	90.59%	13.09%
E, F	92.55%	90.58%	95.00%	9.89%
F, G	92.24%	88.92%	96.51%	12.04%
G, H	93.60%	95.84%	91.17%	3.96%
H, I	86.06%	96.94%	74.48%	2.36%
I, J	85.95%	83.53%	93.50%	21.59%
J, K	93.29%	97.14%	89.20%	2.63%
K, L	95.24%	97.34%	93.07%	2.59%
L, M	92.36%	93.31%	91.26%	6.54%
平均	91.37%	92.22%	91.01%	8.27%

表 10 各期間における k-NN の平均分類精度（固定）
Table 10 Average classification accuracy of k-NN under individual period (fixed parameter).

期間	正解率	適合率	検知率	誤検知率
A, B	79.03%	89.42%	68.14%	10.07%
B, C	79.21%	87.41%	68.14%	9.71%
C, D	81.44%	90.50%	71.85%	8.96%
D, E	84.25%	86.63%	82.80%	14.29%
E, F	78.60%	86.98%	69.27%	12.08%
F, G	74.32%	86.28%	60.22%	11.59%
G, H	65.83%	66.10%	60.03%	28.37%
H, I	85.45%	85.01%	87.72%	16.82%
I, J	61.78%	58.16%	76.67%	53.12%
J, K	78.61%	81.74%	83.63%	26.42%
K, L	74.05%	75.39%	78.35%	30.25%
L, M	77.71%	87.41%	66.40%	10.98%
平均	76.69%	81.75%	72.77%	19.39%

表 11 各期間における OCSVM の平均分類精度（固定）
Table 11 Average classification accuracy of OCSVM under individual period (fixed parameter).

期間	正解率	適合率	検知率	誤検知率
A, B	67.81%	88.32%	40.97%	5.35%
B, C	83.52%	79.57%	90.17%	23.14%
C, D	87.03%	88.47%	85.17%	11.11%
D, E	86.51%	81.41%	94.63%	21.61%
E, F	80.09%	81.82%	77.32%	17.15%
F, G	74.01%	77.35%	67.86%	19.83%
G, H	67.18%	65.95%	70.90%	36.53%
H, I	84.51%	89.10%	78.56%	9.54%
I, J	60.76%	58.20%	75.50%	53.99%
J, K	88.03%	88.29%	87.66%	11.61%
K, L	82.29%	83.81%	80.00%	15.42%
L, M	85.75%	87.08%	83.95%	12.46%
平均	78.96%	80.78%	77.73%	19.81%

© 2017 Information Processing Society of Japan

表 12 各期間における DT の平均分類精度（最適）
Table 12 Average classification accuracy of DT under individual period (optimized parameter).

期間	正解率	適合率	検知率	誤検知率
A, B	93.74%	97.40%	89.89%	2.40%
B, C	95.85%	93.73%	98.27%	6.58%
C, D	90.50%	97.55%	83.08%	2.09%
D, E	92.18%	94.78%	89.27%	4.92%
E, F	96.55%	96.30%	96.83%	3.73%
F, G	97.39%	97.61%	97.16%	2.38%
G, H	93.68%	91.74%	96.06%	8.71%
H, I	95.77%	98.21%	93.25%	1.70%
I, J	85.45%	78.82%	97.24%	26.33%
J, K	97.58%	98.08%	97.07%	1.90%
K, L	97.55%	98.59%	96.47%	1.38%
L, M	95.71%	97.04%	94.30%	2.89%
平均	94.33%	94.99%	94.07%	5.42%

期間ごとに詳しくみると，非常に高い誤検知率でありながら検知率もそれほど高くならない期間が存在する．たとえば期間（I, J）では，どちらの教師なし手法でも検知率が75%程度であるにもかかわらず誤検知率が 50%程度と非常に高くなった．教師あり手法でも同期間で同様の傾向は確認できたものの，教師なし手法での分類精度の低下は異常である．教師なし手法ではこのような期間が複数存在するため，全期間の平均誤検知率を 20%未満となるようにパラメータを調整した今回の評価では，異常な期間に影響を受けたパラメータが設定されてしまい，平均的な分類精度が低下してしまった．教師なし学習手法での極端な分類精度低下については 5.1.3 項で詳細に述べる．

5.1.2 パラメータ最適化時の評価結果

各手法ごとに加え，各期間においてもパラメータを最適化して評価を行った．なお，決定木とナイーブベイズでは，分類器の作成時にパラメータ調整を行っていないため結果はパラメータ固定時と同じである．決定木とナイーブベイズでパラメータの調整を行わなかったのは，いずれの手法でも今回調整を行ったパラメータではパラメータ調整時の分類精度の変化が非常に小さかったためである．決定木，ランダムフォレスト，ナイーブベイズ，サポートベクタマシン，k 近傍法，One Class SVM のパラメータ最適化時の分類結果を，それぞれ**表 12**，**表 13**，**表 14**，**表 15**，**表 16**，**表 17** に示す．全体の傾向としては，パラメータ固定時の結果である表 6〜表 11 とパラメータ最適化時の結果である表 12〜表 17 を比較すると，パラメータを期間ごとに調整した場合の方が分類精度が高くなる傾向が確認できた．このことから，期間ごとに最適なパラメータが変化しており，パラメータの調整が分類器の精度に大きく影響することが確認できた．

5.1.3 主成分分析によるデータセットの分析

5.1.1 項で述べたように，Kyoto 2016 Dataset では教師

表 13 各期間における RF の平均分類精度（最適）
Table 13 Average classification accuracy of RF under individual period (optimized parameter).

期間	正解率	適合率	検知率	誤検知率
A, B	93.99%	98.01%	89.80%	1.82%
B, C	96.47%	94.16%	99.10%	6.15%
C, D	90.83%	97.22%	84.05%	2.39%
D, E	92.53%	94.69%	90.10%	5.04%
E, F	97.47%	96.35%	98.67%	3.74%
F, G	98.48%	98.11%	98.86%	1.90%
G, H	92.21%	87.94%	97.94%	13.52%
H, I	97.62%	98.46%	96.76%	1.51%
I, J	81.45%	73.65%	98.92%	36.02%
J, K	98.74%	98.57%	98.91%	1.44%
K, L	98.85%	98.91%	98.80%	1.09%
L, M	97.52%	97.47%	97.57%	2.53%
平均	94.68%	94.46%	95.79%	6.43%

表 14 各期間における NB の平均分類精度（最適）
Table 14 Average classification accuracy of NB under individual period (optimized parameter).

期間	正解率	適合率	検知率	誤検知率
A, B	81.93%	89.81%	72.30%	8.43%
B, C	90.49%	89.50%	91.75%	10.77%
C, D	89.14%	88.91%	89.43%	11.15%
D, E	88.92%	86.50%	92.25%	14.40%
E, F	84.21%	87.09%	80.32%	11.91%
F, G	85.25%	84.17%	86.83%	16.33%
G, H	75.57%	70.59%	87.66%	36.52%
H, I	85.24%	83.93%	87.18%	16.70%
I, J	67.87%	63.10%	86.06%	50.32%
J, K	81.82%	82.95%	80.37%	16.72%
K, L	92.17%	88.13%	97.45%	13.12%
L, M	85.64%	82.20%	90.96%	19.69%
平均	84.02%	83.07%	86.88%	18.84%

表 15 各期間における SVM の平均分類精度（最適）
Table 15 Average classification accuracy of SVM under individual period (optimized parameter).

期間	正解率	適合率	検知率	誤検知率
A, B	88.64%	92.74%	83.84%	6.57%
B, C	93.21%	90.03%	97.18%	10.76%
C, D	94.78%	93.38%	96.42%	6.85%
D, E	88.92%	87.40%	90.95%	13.11%
E, F	92.54%	90.55%	95.00%	9.93%
F, G	93.19%	90.25%	96.86%	10.48%
G, H	93.76%	95.95%	91.37%	3.85%
H, I	86.33%	96.91%	75.05%	2.39%
I, J	87.56%	85.73%	93.09%	17.96%
J, K	93.37%	97.22%	89.30%	2.55%
K, L	95.25%	97.32%	93.11%	2.60%
L, M	92.52%	93.52%	91.38%	6.33%
平均	91.67%	92.58%	91.13%	7.78%

© 2017 Information Processing Society of Japan

表 16 各期間における k-NN の平均分類精度（最適）
Table 16 Average classification accuracy of k-NN under individual period (optimized parameter).

期間	正解率	適合率	検知率	誤検知率
A, B	88.56%	85.20%	93.34%	16.23%
B, C	88.49%	85.60%	92.77%	15.79%
C, D	90.42%	86.65%	95.64%	14.80%
D, E	87.76%	85.68%	91.55%	16.02%
E, F	85.61%	84.64%	87.53%	16.32%
F, G	76.53%	82.45%	68.66%	15.60%
G, H	66.97%	66.21%	63.66%	29.72%
H, I	86.81%	84.48%	91.35%	17.73%
I, J	64.37%	60.55%	78.66%	49.91%
J, K	91.48%	88.66%	95.25%	12.29%
K, L	82.59%	82.06%	84.64%	19.47%
L, M	84.97%	85.58%	85.26%	15.32%
平均	82.88%	81.48%	85.69%	19.93%

表 17 各期間における OCSVM の平均分類精度（最適）
Table 17 Average classification accuracy of OCSVM under individual period (optimized parameter).

期間	正解率	適合率	検知率	誤検知率
A, B	87.10%	87.86%	86.12%	11.91%
B, C	74.49%	80.71%	64.30%	15.32%
C, D	88.14%	87.53%	88.97%	12.69%
D, E	85.85%	83.47%	89.39%	17.69%
E, F	78.94%	84.19%	71.23%	13.34%
F, G	82.56%	78.69%	89.30%	24.19%
G, H	67.14%	66.35%	69.15%	34.87%
H, I	86.87%	86.30%	87.62%	13.88%
I, J	64.85%	60.65%	84.20%	54.50%
J, K	91.82%	87.47%	97.64%	13.99%
K, L	82.10%	85.15%	77.76%	13.56%
L, M	87.36%	86.90%	88.00%	13.28%
平均	81.44%	81.27%	82.81%	19.94%

なし学習手法を使用した分類時に極端な精度低下が起こった．特に，表 16，表 17 から分かるように，期間（G, H），(I, J）での分類精度低下が大きい．この原因を調査するために，主成分分析による Kyoto 2016 Dataset の可視化を行った．主成分分析により，一部の情報は失われるものの 4 次元以上のデータを視覚で理解可能な次元に写像することができる．

主成分分析を行うにあたり，表 4 に示した A から M までの全 260 万サンプルを使用した．使用した特徴量は，5.1.1 項，および 5.1.2 項の評価と同じ 12 種類の特徴量（12 次元）である．これら 12 次元，260 万サンプルのデータに対し，scikit-learn の StandardScaler クラスによる正規化を行い，同じく scikit-learn の PCA クラスによる主成分分析で 2 次元の空間へ写像した．主成分分析の結果のうち，期間 I（2008 年 3 月から 4 月）の攻撃サンプルのみを取り出したものを**図 6**，正常サンプルのみを取り出したものを

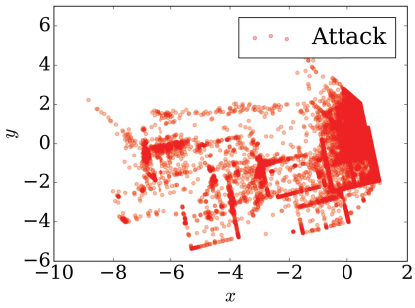

図 6 期間 I での攻撃サンプルの主成分分析結果
Fig. 6 Principal component analysis result of attack traffic samples in period I.

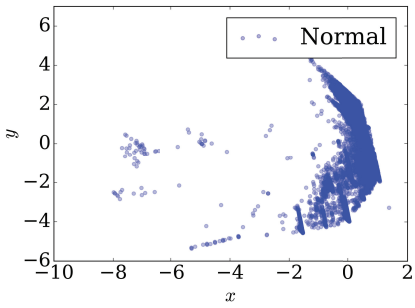

図 7 期間 I での正常サンプルの主成分分析結果
Fig. 7 Principal component analysis result of normal traffic samples in period I.

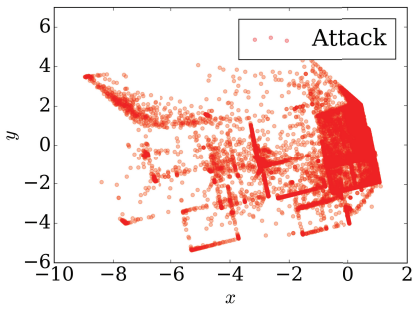

図 8 期間 J での攻撃サンプルの主成分分析結果
Fig. 8 Principal component analysis result of attack traffic samples in period J.

図 7 に示す．また，期間 J（2008 年 5 月から 6 月）の攻撃サンプルのみを取り出したものを**図 8**，正常サンプルのみを取り出したものを**図 9** に示す．図 7 をみると，期間 I の正常サンプルは中央より右側に多く分布することが分かる．一方，図 9 をみると，期間 J の正常サンプルは x 軸方向に広く分布することが分かる．期間 (G, H) についても同様に分布することを確認した．この結果から，期間（G, H），(I, J）では，評価データに学習データには含まれない正常サンプルが含まれるという仮説を立てた．

仮説が正しいことを確認するために，教師あり学習手法の結果に注目した．表 12〜表 15 をみると，期間（G, H），(I, J）では平均検知率は他期間と同程度であるものの，全体的に平均誤検知率が上昇している．平均検知率が他期間

© 2017 Information Processing Society of Japan

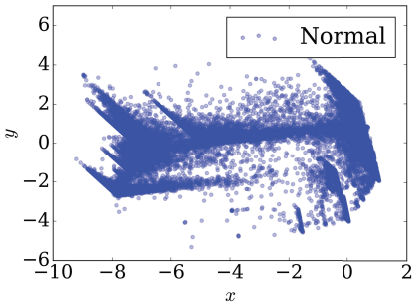

図 9 期間 J での正常サンプルの主成分分析結果

Fig. 9 Principal component analysis result of normal traffic samples in period J.

と同程度であったのは，図 6，図 8 から分かるように，学習データと評価データで攻撃サンプルの分布がほぼ変化していないためだと考えられる．平均誤検知率が上昇した原因は，仮説を用いると説明することができる．学習データに含まれる評価データの正常サンプルは，当然正しく正常サンプルであると分類できる．しかしながら，学習データに含まれない評価データの正常サンプルは分類器の作成時には未知の存在であり，境界を決定する際に考慮されない．このため，分類器は学習データに含まれない評価データの正常サンプルを正しく分類することができず，誤検知が増加したと考えられる．以上から，教師なし学習手法における一部での極端な分類精度低下の主な要因は，評価データに学習データには含まれない正常サンプルが含まれるためだと結論付けた．

6. まとめ

本論文において，我々は Kyoto 2016 Dataset を作成し，基本的な統計情報の調査と機械学習による分類を行った．

Kyoto 2016 Dataset 作成には，ハニーポットによって収集された 3,268 日間のトラフィックデータを使用した．3,268 日間のトラフィックデータは，Bro IDS 2.4.1 のセッション解析によっておよそ 8 億のセッションへと変換され，Kyoto 2016 Dataset は 8 億サンプルからなる巨大なデータセットとなった．

また，統計情報の調査によって，ハニーポットへは比較的新しい攻撃を含むさまざなま機器を対象としたさまざまな悪性通信が到達しており，傾向も年々変化していることを確認した．このことから，Kyoto 2016 Dataset は日々刻々と変化する悪性通信の実情をより正確にとらえているものだと考えることができる．

さらに，Kyoto 2016 Dataset を複数の基本的な機械学習手法によって分類した結果を示した．この分類結果は，ネットワーク侵入検知に関する研究を行う研究者へ，Kyoto 2016 Dataset へ機械学習手法を適用する場合の分類精度に関する指標を与えられるものだと考えている．

謝辞 本研究の一部は，JSPS 科研費 16K00071 の支援により行った．

参考文献

[1] Chandola, V., Banerjee, A. and Kumar, V.: Anomaly Detection: A Survey, *ACM Comput. Surv.*, Vol.41, No.3, pp.15:1–15:58 (online), DOI: 10.1145/1541880.1541882 (2009).

[2] Ambusaidi, M.A., He, X., Nanda, P. and Tan, Z.: Building an Intrusion Detection System Using a Filter-Based Feature Selection Algorithm, *IEEE Trans. Computers*, Vol.65, No.10, pp.2986–2998 (online), DOI: 10.1109/TC.2016.2519914 (2016).

[3] Om, H. and Kundu, A.: A hybrid system for reducing the false alarm rate of anomaly intrusion detection system, *2012 1st International Conference on Recent Advances in Information Technology* (*RAIT*), pp.131–136 (online), DOI: 10.1109/RAIT.2012.6194493 (2012).

[4] Hosseini, Z.S., Chabok, S.J.S.M. and Kamel, S.R.: DOS intrusion attack detection by using of improved SVR, *2015 International Congress on Technology, Communication and Knowledge* (*ICTCK*), pp.159–164 (online), DOI: 10.1109/ICTCK.2015.7582663 (2015).

[5] Eskin, E., Arnold, A., Prerau, M., Portnoy, L. and Stolfo, S.: *A Geometric Framework for Unsupervised Anomaly Detection*, pp.77–101 (online), DOI: 10.1007/978-1-4615-0953-0_4, Springer US (2002).

[6] RT, K., Selvi, S.T. and Govindarajan, K.: DDoS detection and analysis in SDN-based environment using support vector machine classifier, *2014 6th International Conference on Advanced Computing* (*ICoAC*), pp.205–210 (online), DOI: 10.1109/ICoAC.2014.7229711 (2014).

[7] Mukkamala, S., Janoski, G. and Sung, A.: Intrusion detection using neural networks and support vector machines, *Proc. 2002 International Joint Conference on Neural Networks, IJCNN '02.*, Vol.2, pp.1702–1707 (online), DOI: 10.1109/IJCNN.2002.1007774 (2002).

[8] Masarat, S., Sharifian, S. and Taheri, H.: Modified Parallel Random Forest for Intrusion Detection Systems, *J. Supercomput.*, Vol.72, No.6, pp.2235–2258 (online), DOI: 10.1007/s11227-016-1727-6 (2016).

[9] Stein, G., Chen, B., Wu, A.S. and Hua, K.A.: Decision Tree Classifier for Network Intrusion Detection with GA-based Feature Selection, *Proc. 43rd Annual Southeast Regional Conference - Volume 2*, ACM-SE 43, pp.136–141, ACM (online), DOI: 10.1145/1167253.1167288 (2005).

[10] Amor, N.B., Benferhat, S. and Elouedi, Z.: Naive Bayes vs Decision Trees in Intrusion Detection Systems, *Proc. 2004 ACM Symposium on Applied Computing, SAC '04*, pp.420–424, ACM (online), DOI: 10.1145/967900.967989 (2004).

[11] MIT Lincoln Laboratory: DARPA Intrusion Detection Data Sets, MIT Lincoln Laboratory (online), available from ⟨https://www.ll.mit.edu/ideval/data/⟩ (accessed 2016-10-17).

[12] UCI KDD Archive: KDD Cup 1999 Data, UCI KDD Archive (online), available from ⟨http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html⟩ (accessed 2016-10-17).

[13] Song, J., Takakura, H., Okabe, Y., Eto, M., Inoue, D. and Nakao, K.: Statistical Analysis of Honeypot Data and Building of Kyoto 2006+ Dataset for NIDS Evaluation, *Proc. 1st Workshop on Building Analysis Datasets and Gathering Experience Returns for Secu-*

© 2017 Information Processing Society of Japan

rity, *BADGERS '11*, pp.29–36, ACM (online), DOI: 10.1145/1978672.1978676 (2011).
[14] 高等教育機関における情報セキュリティポリシー推進部会：高等教育機関の情報セキュリティ対策のためのサンプル規程集（2015 年版補訂），国立情報学研究所（オンライン），入手先 ⟨http://www.nii.ac.jp/csi/sp/⟩（参照 2017-04-04）.
[15] Takakura, H.: Traffic Data from Kyoto University's Honeypots, Takakura, H. (online), available from ⟨http://www.takakura.com/Kyoto_data/⟩ (accessed 2016-10-17).
[16] Wireshark developers: Wireshark, Wireshark developers (online), available from ⟨https://www.wireshark.org/⟩ (accessed 2016-10-17).
[17] The Bro Project: The Bro Network Security Monitor, The Bro Project (online), available from ⟨https://www.bro.org/index.html⟩ (accessed 2016-10-17).
[18] 多田竜之介，小林良太郎，嶋田　創，高倉弘喜：新 Kyoto 2006+データセットの作成に関する検討と評価，電子情報通信学会技術研究報告，ICSS, 情報通信システムセキュリティ，Vol.116, No.328, pp.21–26 (2016).
[19] MaxMind: GeoIP2 Database Demo, MaxMind (online), available from ⟨https://www.maxmind.com/ja/geoip-demo⟩ (accessed 2017-03-20).
[20] MaxMind: GeoLite2, MaxMind (online), available from ⟨http://dev.maxmind.com/ja/geolite2/⟩ (accessed 2017-03-08).
[21] Symantec：シマンテック, Symantec（オンライン）, 入手先 ⟨https://www.symantec.com/ja/jp⟩（参照 2017-03-19）.
[22] Cisco: Sourcefire - Cisco, Cisco (online), available from ⟨http://www.cisco.com/c/en/us/services/acquisitions/sourcefire.html⟩ (accessed 2017-03-19).
[23] Cisco: Sid 1-384, Cisco (online), available from ⟨https://www.snort.org/rule_docs/1-384⟩ (accessed 2017-03-20).
[24] Cisco: Sid 1-483, Cisco (online), available from ⟨https://www.snort.org/rule_docs/1-483⟩ (accessed 2017-03-20).
[25] Cisco: Sid 1-2050, Cisco (online), available from ⟨https://www.snort.org/rule_docs/1-2050⟩ (accessed 2017-03-20).
[26] Cisco: Sid 1-4990, Cisco (online), available from ⟨https://www.snort.org/rule_docs/1-4990⟩ (accessed 2017-03-20).
[27] Cisco: Sid 1-2004, Cisco (online), available from ⟨https://www.snort.org/rule_docs/1-2004⟩ (accessed 2017-03-21).
[28] Cisco: Sid 1-449, Cisco (online), available from ⟨https://www.snort.org/rule_docs/1-449⟩ (accessed 2017-03-20).
[29] Cisco: Sid 1-31978, Cisco (online), available from ⟨https://www.snort.org/rule_docs/1-31978⟩ (accessed 2017-03-21).
[30] National Institute of Standards and Technology: CVE-2014-7169 - NVD - Detail - NIST, National Institute of Standards and Technology (online), available from ⟨https://web.nvd.nist.gov/view/vuln/detail?vulnId=CVE-2014-7169⟩ (accessed 2017-03-21).
[31] Cisco: Sid 1-31136, Cisco (online), available from ⟨https://www.snort.org/rule_docs/1-31136⟩ (accessed 2017-03-21).
[32] Cisco: Sid 1-28556, Cisco (online), available from ⟨https://www.snort.org/rule_docs/1-28556⟩ (accessed 2017-03-21).
[33] United States Computer Emergency Readiness Team: DNS Amplification Attacks — US-CERT, United States Computer Emergency Readiness Team (online), available from ⟨https://www.us-cert.gov/ncas/alerts/TA13-088A⟩ (accessed 2017-03-21).
[34] Cisco: Sid 1-34464, Cisco (online), available from ⟨https://www.snort.org/rule_docs/1-34464⟩ (accessed 2017-03-21).
[35] National Institute of Standards and Technology: CVE-2014-9583 - NVD - Detail - NIST, National Institute of Standards and Technology (online), available from ⟨https://web.nvd.nist.gov/view/vuln/detail?vulnId=CVE-2014-9583⟩ (accessed 2017-03-21).
[36] ClamAV Team: GeoLite2, ClamAV Team (online), available from ⟨https://www.clamav.net/⟩ (accessed 2017-03-19).
[37] Symantec：W32.Downadup — シマンテック日本, Symantec（オンライン）, 入手先 ⟨https://www.symantec.com/ja/jp/security_response/writeup.jsp?docid=2008-112203-2408-99⟩（参照 2017-03-19）.
[38] 大平健司，宋　中錫，高倉弘喜，岡部寿男：様々なアプリケーションへの攻撃活動を察知する汎用性の高いハニーポットシステムの構築と運用, 電子情報通信学会論文誌, D, 情報・システム = The IEICE transactions on information and systems (Japanese edition)，Vol.93, No.7, pp.1125–1134（オンライン）(2010)，入手先 ⟨http://ci.nii.ac.jp/naid/110007642046/⟩.
[39] scikit-learn developers: scikit-learn: machine learning in Python, scikit-learn developers (online), available from ⟨http://scikit-learn.org/⟩ (accessed 2016-10-17).

多田 竜之介

2017 年豊橋技術科学大学工学部情報・知能工学課程卒業．同年同大学大学院工学研究科情報・知能工学専攻博士前期課程入学．ネットワークセキュリティの研究に従事．

小林 良太郎　（正会員）

1995 年名古屋大学工学部電子情報学科卒業．1997 年同大学大学院工学研究科電子情報学専攻博士課程前期課程修了．2000 年同大学院工学研究科電子情報学専攻博士課程後期課程満了．工学博士．2000 年名古屋大学大学院工学研究科電子情報学専攻助手．2008 年豊橋技術科学大学工学部講師．2016 年同大学大学院工学研究科准教授．2017 年工学院大学情報学部コンピュータ科学科准教授．1999 年本会山下記念研究賞受賞．2002 年本会論文賞受賞．計算機アーキテクチャ，ネットワークセキュリティの研究に従事．

© 2017 Information Processing Society of Japan

嶋田 創 （正会員）

1976 年生．1998 年名古屋大学工学部情報工学科卒業．2000 年同大学大学院工学研究科情報工学専攻博士前期課程修了．2004 年同大学工学博士．2004 年名古屋大学工学部電気系 COE 研究員．2005 年京都大学大学院情報学研究科特任助手．2006 年同大学院情報学研究科助手．2009 年奈良先端大学院大学情報科学研究科准教授．2013 年名古屋大学情報基盤センター情報基盤ネットワーク研究部門准教授．低消費電力と高信頼性を兼ね備えた計算機アーキテクチャとネットワーク関連の研究に従事．電子情報通信学会，IEEE 各会員．

高倉 弘喜 （正会員）

1990 年九州大学工学部卒業．1992 年同大学大学院修士課程修了．1995 年京都大学大学院博士後期課程修了．博士（工学）．米国イリノイ州立大学訪問研究員，奈良先端科学技術大学院大学助手，京都大学講師，助教授（准教授），名古屋大学教授を経て，2015 年から国立情報学研究所教授．サイバーセキュリティ，高機能ネットワークの研究に従事．電子情報通信学会，システム制御情報学会，地理情報システム学会，ACM 各会員．

© 2017 Information Processing Society of Japan