ネットワーク型侵入検知手法を用いた KDD Cup 1999 Data と Kyoto2016 の比較

高原尚志[†1]

概要：現在，インターネット上では，日々新たな攻撃が生み出されている．そのため，未知の攻撃からコンピュータを守ることは大変重要な事柄である．攻撃からコンピュータを守る方法に，侵入検知がある．侵入検知には，ネットワーク上の通信データを用いて攻撃を検知する方法（NIDS; Network-based Intrusion Detection System）とホストコンピュータのシステムログなどを用いて検知する方法（HIDS; Host-based Intrusion Detection System）がある．検知システムを評価する方法として，評価用データセットを用いる方法が広く知られている．本稿では，NIDS に焦点を当て，その評価用データセットについて解説した後，代表的な NIDS 手法について，各データセットでの評価を行い，その結果を比較検討する．これにより，各データセットの特徴を明らかにし，読者がデータセットを選択する際に参考にしてもらえることを目指す．

キーワード：NIDS，KDD Cup 1999 Data，Kyoto2016，NSL-KDD, DARPA 1998 Data

Comparison between KDD Cup 1999 Data and Kyoto2016 Using Network-based Intrusion Detection System

HISASHI TAKAHARA[†1]

Abstract: Today, on the Internet, new cyber-attack is made. So, it is very important that PC is protected from unknown attack. There is Intrusion Detection System (IDS) as method to protect from cyber-attack. IDS have two systems. Those are Network-based Intrusion Detection System (NIDS) and Host-based Intrusion Detection System (HIDS). NIDS detects attack using traffic data and HIDS detects attack using system log data of host PC. As method to evaluate IDS, method with dataset is well-known. In this paper, focused on NIDS, those datasets is explained and evaluated with well-known methods for NIDS. Finally, the results are compared. Therefore, character of each datasets is made clear and then it is expected that this paper is referred for readers to select datasets.

Keywords: NIDS, KDD Cup 1999 Data, Kyoto2016, NSL-KDD, DARPA 1998 Data

1. はじめに

1.1 研究の背景

インターネット上では，多くのサイバー攻撃が日々行われている．これを（できれば未然に）検知すること（侵入検知）は大変重要なことである．侵入検知には，ネットワーク上の通信データを用いて攻撃を検知する手法（ネットワーク型侵入検知システム；NIDS Network-based Intrusion Detection）とホストコンピュータのシステムログデータなどを用いて攻撃を検知する手法（ホスト型侵入検知システム；HIDS Host-based Intrusion Detection System）がある[1]．侵入検知手法を評価する場合，評価用データセットを用いる方法が広く知られている[2]．著者らは，NIDS に焦点を当て，その評価用データセットの特徴について，文献[3]の中で比較し論じてきた．本稿では，文献[3]の中で比較し論じたデータセットを実際に用いて，NIDS 手法の評価を行い，データセットごとの評価結果の違いについて考察する．

1.2 研究の動機

著者らは，文献[2]の中で，NIDS の研究の手順について，次のようにまとめた．

(STEP1)　手法の理論的開発
(STEP2)　手法のソフトウエアによる実現
(STEP3)　データセットを用いた手法の評価

また，著者らは，それぞれの STEP についての課題を文献[2]の中で指摘した．更に，評価用データセットについて著者らは，文献[4]の中で，文献[2]で指摘した課題の解決方法を示した．更に，著者らは，文献[5]の中で，各データセットの用途について論じ，文献[3]の中では，各データセットの特徴について比較しまとめた．

本稿では，著者らが行ってきた上記の一連の研究の延長

†1 新潟県立大学
University of NIIGATA PREFECTURE

として，各データセットを用いて，広く知られた NIDS 手法を評価することによって，データセットごとの違いを明らかにし，考察を加える．

1.3 既存研究

広く知られた NIDS 評価用データセットに，KDD Cup 1999 Data(KDD99)[6]がある．Atilla ら[7]は，2010 年から 2015 年に発行された 65 の科学雑誌の 149 の論文をサーベイした結果，侵入検知ための NIDS 手法の評価に用いられている評価用データセットの 77.8%は KDD99 であることを明らかにしている．そこで本稿では，NIDS 評価用データセットとして，広く知られた KDD99 と日本の京都大学から提供されている評価用データセットである Kyoto Data(Kyoto)について取り上げる．ここでは，KDD99 と Kyoto2 について，その特徴を示した既存研究について述べる．なお，各データセットの詳細については，2 章で説明する．

1.3.1 KDD Cup 1999 Data

Lippmann ら[8]は，広く知られた NIDS 用データセットである KDD99 のもととなったデータセットである 1998 DARPA Intrusion Detection Evaluation Data Sets（DARPA 1998）[9]について，取得方法から統計的な分析まで詳細に説明している．

Saffaa ら[10]は，KDD99 や機械学習の手法についての今までの研究を整理し，KDD99 を NIDS 手法で評価し，TPR(True Positive Rate)[11][12]（3.1.4 章参照）や FAR(False Alarm Rate)[11][12]（3.1.4 章参照），学習データからモデルを作成するまでの時間(Training Time)などを示している．

Mahbod ら[13]は，KDD99 の中で，データが冗長的であるなど KDD99 の課題を指摘し，その解決方法として，KDD99 から冗長データを削除した新たなデータセット NSL-KDD[14]を紹介し，NIDS の手法を NSL-KDD で ACC(Accuracy)[11][12]（3.1.4 章参照）をもとに評価し，その特徴を明らかにしている．

Revathi ら[15]は，NSL-KDD を用いて，NIDS の手法を攻撃カテゴリ別に ACC をもとに評価し，NSL-KDD の特徴を明らかにしている．

1.3.2 Kyoto Data

Song ら[16]は，京都大学から提供されている NIDS 評価用データセットである Kyoto2006+[17]の特徴量について説明している．また，Song ら[18]は，Kyoto2006+を作成するためのハニーポットにおけるデータの収集方法について解説をしている．更に，Song ら[19]は，Kyoto2006+について，正常通信や既知攻撃，未知攻撃などの割合やアクセスして来た国や地域の割内など，統計的な分析を行うとともに，特徴量を 14 個に絞った経緯などについて説明している．

多田ら[20]は，Kyoto2006+を更に発展させた Kyoto2016 Dataset(Kyoto2016)[17]について，収集期間や統計的な分析などを行っている．また多田らは，Kyoto2016 を用いて，NIDS 手法を評価することを通じて，Kyoto2016 の特徴を明らかにした．

1.4 解決すべき課題

現在まで，各データセットについて，その特徴を明らかにするために，NIDS の手法を用いて各データセットを個別に評価した論文は存在するが，複数のデータセットを NIDS 手法で評価し，各データセットの特徴を比較検討した論文は，著者が知る限り存在しない．NIDS 手法の評価用データセットを用いた評価がその手法の価値を決める以上，評価する場合には，各データセットの特徴を正確に把握した上で用いるデータセットを決める必要がある．さもなければ，何を評価しているかがあいまいになってしまい，その評価結果自体の価値もはっきりとしないものとなってしまう．

1.5 本研究の貢献

本研究の学術的貢献は以下の通りである．

・NIDS 手法を各 NIDS 評価用データセットで評価することによって，NIDS 評価用データセットの特徴を明らかにする

・上記で得られた評価結果をもとに，各 NIDS 評価用データセットを比較することにより，読者が NIDS 手法の評価に最適なデータセットを選択するための指針を提供する

2. データセット

本稿では，1.3 章でも指摘した通り，Atilla ら[7]が指摘した NIDS 手法の評価用に用いられているデータセットの 77.8%を占める KDD99 とその改良版である NSL-KDD，京都大学から提供されており最新の攻撃を含む Kyoto2016 について取り上げる．ここでは，その概要について述べる．

2.1 KDD Cup 1999 Data

2.1.1 作成経緯

KDD99 は，米国国防高等研究計画局（DARPA）と米国空軍研究所（AFRL）のもとでマサチューセッツ工科大学（MIT）Lincoln laboratory の The DARPA Intrusion Detection Evaluation Group が作成配布した世界初の NIDS 評価用標準データセットである DARPA98 をもとに作成された NIDS 評価用データセットで，University of California Irvine の Machine Learning Repository で公開されている．DARPA 98 がパケットキャプチャファイル形式であるのに対し，KDD99 はこれを加工し，セッションデータ形式として供給されており，現在でも多くの論文で NIDS の評価用データセットとして使用されている．しかし，日本国内の論文では，①データが古い，②冗長的である，③攻撃通信の割合が多く現実的でないなどの理由からあまり使用されていない．KDD99 の②冗長的である，③攻撃通信の割合が多く現実的でないなどの欠点を修正したデータセットとして，UNB(University of New Branswich)の CIC(Canadian Institute for Cybersecurity)から提供されている NSL-KDD がある（図

1,)．現在では，NIDS 評価用データセットとしてこちらを用いた論文も多くみられる．

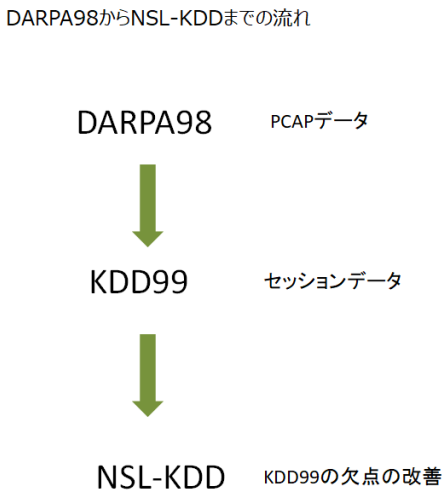


図 1. DARPA98 から NSL-KDD までの流れ
Figure1. Flow from DARPA98 to NSL-KDD

2.1.2 特徴

***KDD99**

KDD99 には，すべてのデータを収めた kddcup.data とそれから 10%を抽出した kddcup.data_10_percent，そして評価用データである corrected がある．前の 2 つ（kddcup_data と kddcup.data_10_percent）は学習用データであり，残りの 1 つ（corrected）は評価用データである．また，corrected には, kddcup.data 及び kddcup.data_10_percent に含まれない攻撃が含まれている．これにより，未知の攻撃の検知についての評価を行うことができる．上記 3 つのデータには，正常通信か攻撃通信か，攻撃通信の場合は攻撃名を記述したラベルが付されている．その為，統一した基準で，攻撃通信か正常通信かを評価することができる．なお，KDD99 では，上記のラベルが付されていないデータセットも unlabeled として配布されている．また，これらのデータは，CSV 形式で配布されている．

KDD99 は，41 個の特徴量と正常通信，攻撃通信を表すラベルからなっている．この内，41 個の特徴量の内，7 個は文字データ (Symbolic) で残りの 34 個は数値データ (Continuous)である．文字データを扱う際には，適宜数値に変換するなどの工夫が必要である. また, KDD99 の攻撃は，DOS，R2L，U2R，Probing の 4 つのカテゴリに分かれており，個別の攻撃は上記 4 つのカテゴリのいずれかに属している．更に，攻撃通信と正常通信の割合はおおよそ 8 対 2 となっている（図 2）.

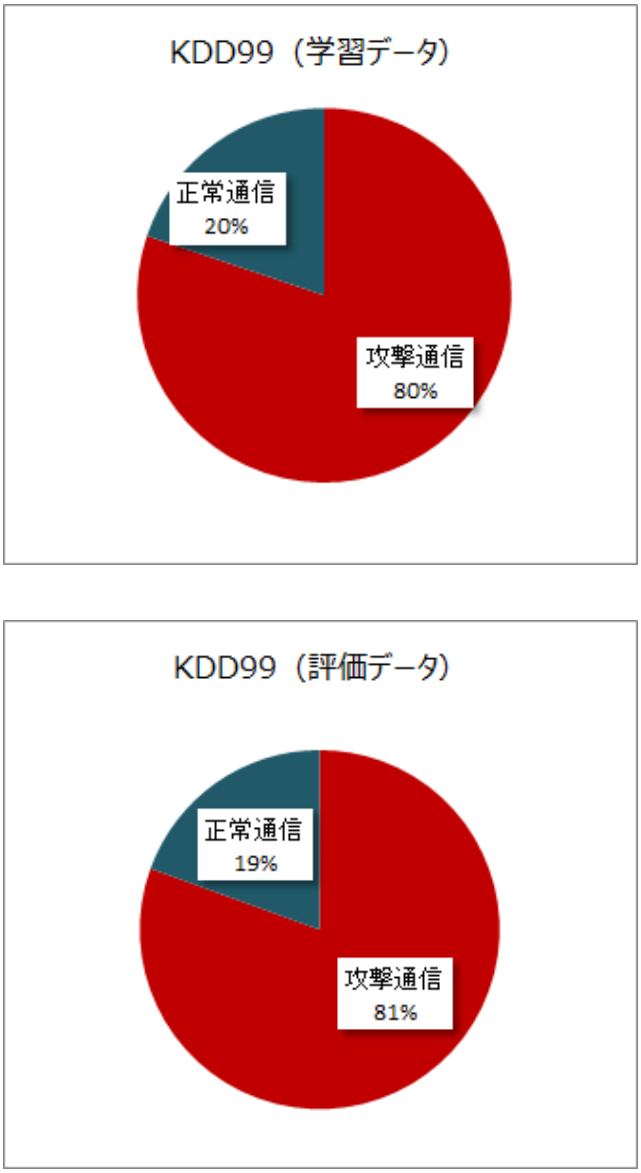


図 2. KDD99　攻撃通信と正常通信の割合
Figure2. The proportion of attack and normal

通常の攻撃通信の割合が 1%以下である[2]ことを考えれば，攻撃通信の割合が通常に比べてかなり多くなっているため，攻撃通信をカテゴリごとに抽出するなど，有効な評価を行う際には工夫が必要であると考えられる．

***NSL-KDD**

NSL-KDD には，学習用データである KDDTrain+と評価用データである KDDTest+があり，加えてそれぞれのサブセットである KDDTrain+_20Percent と KDDTest-21 がある．NSL-KDD は，一般的な TXT 形式（CSV 形式）と広く知られた機械学習用ソフトウエアである Weka（3.1.3 章参照）のデータ形式である ARFF 形式の 2 つの形式で配布されて

いる．

KDD99 に改良を加えた NSL-KDD では，KDD99 のデータが冗長的あるなどの欠点を改善し，データサイズを大幅に減少している（図 3）．

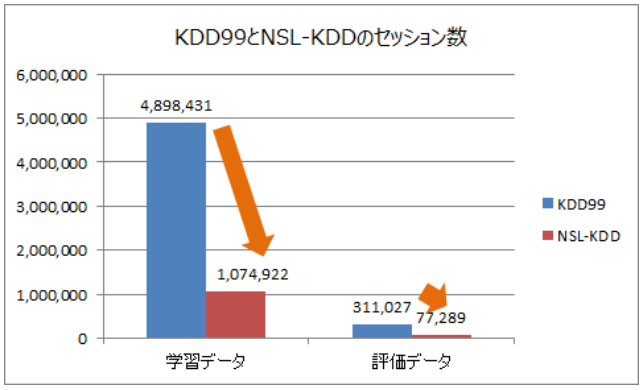


図 3. KDD99 と NSL-KDD のセッション数

Figure3. The number of Sessions of DARPA98 and NSL-KDD

また，攻撃通信の割合が多いという欠点に対しては，攻撃通信と正常通信の割合を学習データでおおよそ 3 対 1，評価データでおおよそ 4 対 6 として，正常通信の割合を大幅に増加させている（図 4）．

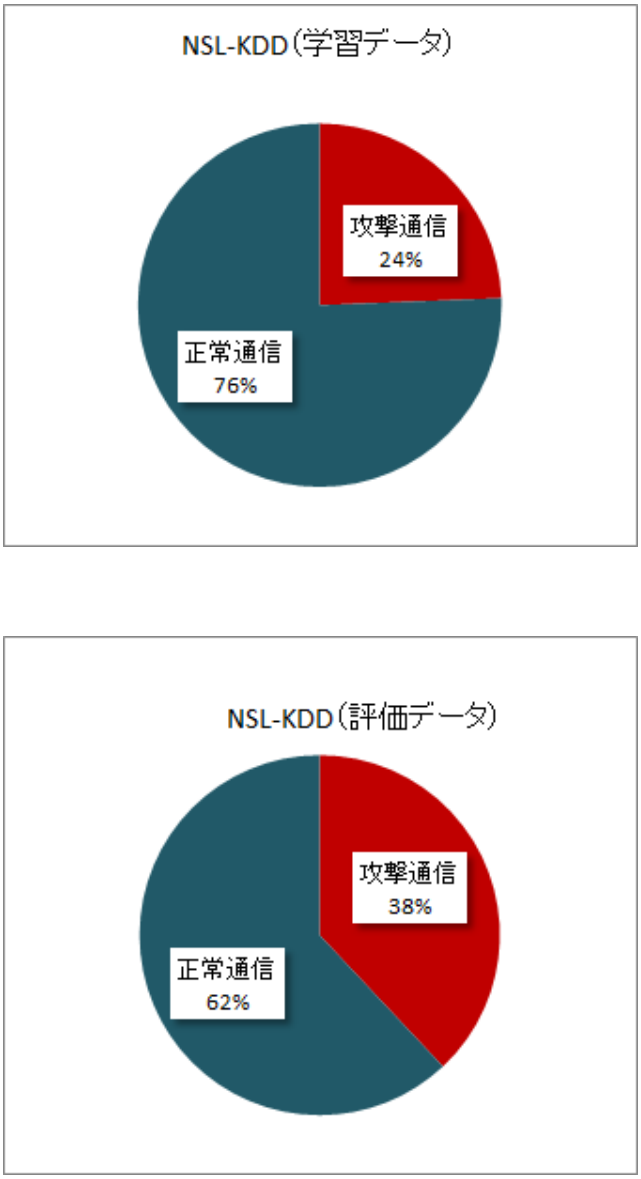


図 4. NSL-KDD　攻撃通信と正常通信の割合

Figure4. The proportion of attack and normal in NSL-KDD

2.2 Kyoto Data

2.2.1 作成経緯

Kyoto2016 は，その前身である Kyoto2006+に 2015 年 12 月までのデータを追加したものである．Kyoto2006+は，KDD99 が古くなったのを受け，京都大学に設置されているハニーポットで 2006 年 11 月から 2009 年 8 月までの期間に収集された通信データをもとに，KDD99 の特徴量の内特に影響が大きい 14 個を抽出して，セッションデータに加工し，既存の攻撃，未知の攻撃，正常通信の 3 つのラベル付けをして配布されているものである．Kyoto2006+におけるセッションデータへの加工時の課題について，改良も加えている．

2.2.2 特徴

文献[19]によれば，Kyoto2016 の観測セッション数は 806,095,624，この内攻撃通信は 645,221,775，正常通信は 160,873,849 で，攻撃通信と正常通信の割合は，おおよそ 8 対 2 であり（図 5），攻撃通信の割合が非常に多い．このため，有効な評価を行う際には，攻撃通信を一部抽出するなど工夫が必要であると考えられる．また，攻撃通信の内，既知攻撃は 640,618,555，未知攻撃は 4,603,220 で，その割合はおおよそ 99 対 1 である（図 6）．また Kyoto2016 では正常通信，未知攻撃，既知攻撃の区別はラベルで示されているが, KDD99 のように攻撃の種類まではラベルに示されていない．

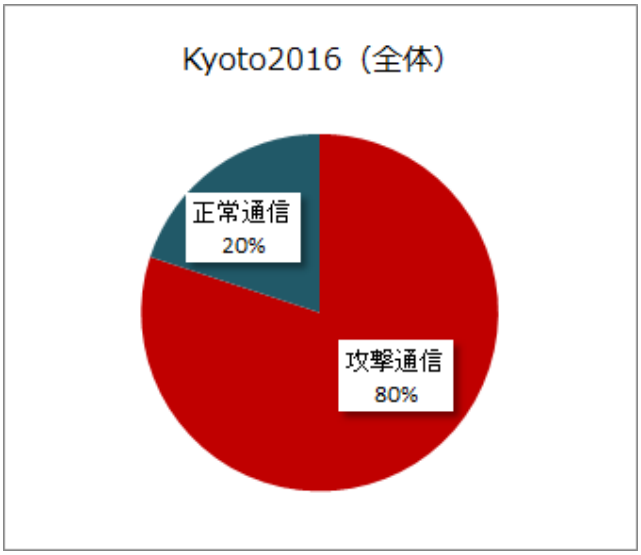


図 5. Kyoto2016　攻撃通信と正常通信の割合
Figure5. The proportion of attack and normal in Kyoto2016

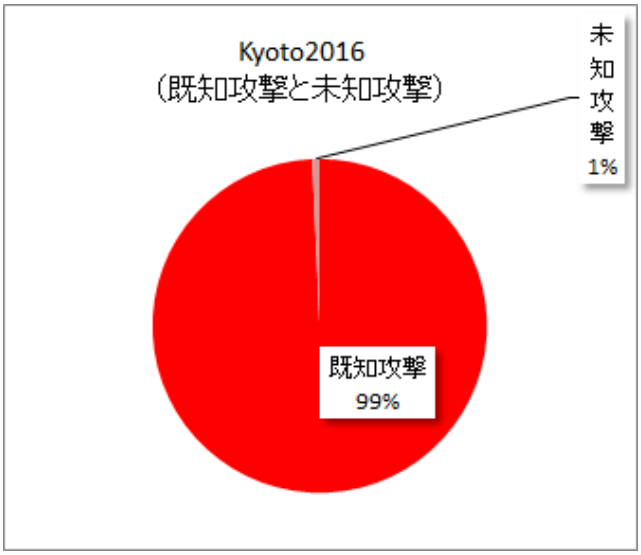


図 6. Kyoto2016　既知攻撃と未知攻撃の割合
Figure6. The proportion of known and unknown attack in Kyoto2016

2.3 データセットの取得

KDD99 はサイト[6]にデータが掲載されており，自由にダウンロードすることができる．これに対して，NSL-KDD の場合は，サイト[14]の中にはダウンロードのページはない．ダウンロードする際には，原則，サイト[14]に記載されているメールアドレスにメールを送り，ダウンロードサイト及びパスワードを取得する必要がある．取得したパスワードには有効期限（2018 年 8 月時点で 72 時間）があるため，その期限内にダウンロードする必要がある．また Kyoto2016 は, サイト[17]にデータセットが年月別で掲載されており，自由にダウンロードできる．

3. データセットの比較実験

3.1 準備

3.1.1 データセット

本稿で比較に用いるデータセットは，利用数 1 位と 2 位を占め[7]，広く知られた NIDS 評価用データセットである KDD99 と KDD99 に改良を加えた NSL-KDD を使用する．また，最新の攻撃データということを考慮して，2015 年 12 月までのデータが提供され，京都大学から配布されている Kyoto2016 も使用する．上記のデータセットの内，KDD99 については，10%サブセット（kddcup.data_10_percent）を学習用データとして使用し，評価用データとして corrected を，両者からセッションを抽出することなく使用する．また，NSL-KDD については，学習用データセットとして KDDTrain+, 評価用データセットとして KDDTest+といったいずれもフルセットを，セッションの抽出を行わない形で使用する．Kyoto2016 については，最新の攻撃への対応を評価することを考慮して，学習用データ，評価用データともに，2015 年 12 月のデータからランダムに日付を数日選択して評価を行い，その平均値を求める．この際，次の点に注意した．

- 学習用データの日付が評価用データの日付を超えない
- 学習用データと評価用データが同一の日付にならない
- 学習用データについては既存攻撃と正常通信のみを抽出し，評価用データについてはセッションの抽出を行わずに（既存攻撃，未知攻撃，正常通信を含んだもので）評価を行う

また, 特徴量については, Kyoto2016 で採用された KDD99 の 14 個の特徴量の内，テキスト形式の特徴量（Service と Flag）を除外した 12 個の特徴量で評価を行う．

3.1.2 ネットワーク型侵入検知手法

近年，NIDS の手法には，機械学習とよばれる手法が広く用いられている．本研究では，文献[7]で示されている利用数が 1 位，2 位，3 位の機械学習手法で，広く知られた機械学習の手法である Support Vector Machine（SVM）[21][22], Naïve Bayes（NB）[23]及び Decision Tree(J48)[24]を用いる．上記 3 つの機械学習手法を KDD99，NSL-KDD 及び

Kyoto2016 に適用することによって，それぞれのデータセットの特徴を明らかにする．

3.1.3 ソフトウエア

本実験では，読者による再現性を考え，独自プログラムは使用せず，広く知られた機械学習用ソフトウエアである Weka を使用する．Weka[25][26]はニュージーランドの Waikato 大学で開発された機械学習用プログラムで，一般に公開されている．だれでも参加できる Weka 専用のメーリングリストなども設置され，バグの改善などを素早く行える体制を整えている．文献[7]によると，Weka は，用いたソフトウエアを明らかにしている文献の中では，最も多く利用されている機械学習用ソフトウエアで，約 37%の論文が利用している．

3.1.4 評価指標

本研究では，文献[7]で示されている利用数が 1 位と 2 位の評価指標である TPR(DR)，FAR 及び総合指標である AUR(Area Under ROC curve, AUC とも言う) [11]を使用する．TPR は全攻撃通信の内，警告をどのくらいの割合で発することができたかを表し（式 1），FAR は全正常通信の内，誤って警告を発せられた割合を表している（式 2）．混同行列を以下のようにすると（表 1），各指標の式は以下の通りである．

表 1.　混同行列

Table1. Confusion matrix

	Positive（予測）	Negative（予測）
Positive（正解）	True Positive(TP)	False Negative(FN)
Negative（正解）	False Positive(FP)	True Negative(TN)

$$\mathrm{TPR} = \frac{TP}{TP+FN} \quad \cdots（式 1）$$

$$\mathrm{FAR} = \frac{FP}{FP+TN} \quad \cdots（式 2）$$

TPR が高く，FAR が低い手法が有効な手法と考えることができる．これを総合的に表したものが ROC 曲線[11]を利用した AUR である．AUR の値が 1 に近づくほど，TPR，FAR の両者を加味した上で，総合的に優れているという判断ができる．

3.2 実験結果

KDD99，NSL-KDD，Kyoto2016 の各データセットについて，NB，J48，SVM の TPR，FAR，AUR を測定した結果を表 2.に，それをグラフ化したものを図 7 に示す．

表 2.各データセットの測定結果

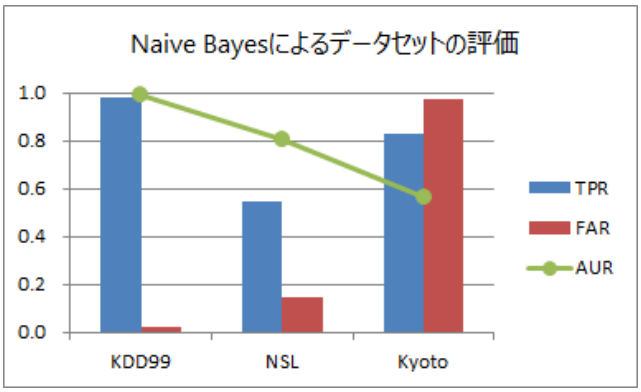
Table2. Results of measurement on each datasets

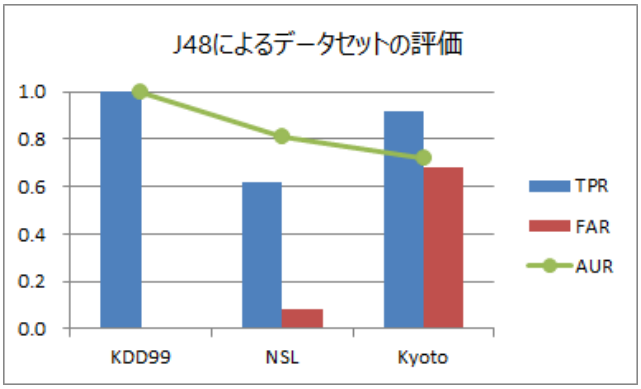
	Naïve Bayes			J48			SVM		
	KDD	NSL	Ky	KDD	NSL	Ky	KDD	NSL	Ky
TPR	0.983	0.545	0.830	0.998	0.618	0.920	0.982	0.560	1.000
FAR	0.020	0.142	0.971	0.006	0.084	0.681	0.010	0.079	1.000
AUR	0.989	0.806	0.561	1.000	0.812	0.718	0.986	0.740	0.500

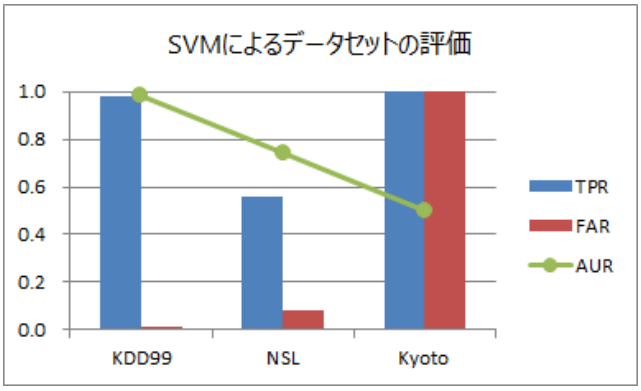
※表中，KDD は KDD99，NSL は NSL-KDD，Ky は Kyoto

図 7. 各データの測定結果

Figure7. Results of measurement on each datasets







＊実験結果の分析と考察

上記の 3 つのデータセットを整理すると以下のようになる．

・KDD99；最も広く利用されている NIDS 用セッション型データセットで，1999 年に配布されたため，古いという欠点や意図的に攻撃通信を混ぜており，冗長的であるという欠点を有する

・NSL-KDD；KDD99 の冗長的であるなどの欠点を修復したセッション型データセット

・Kyoto；KDD99 の特徴量の一部を踏襲して，2015 年 12 月までのデータを提供している NIDS 評価用セッション型データセット

以上を踏まえて，上記の実験結果を考察する．NB，J48，SVM ともに総合指標である AUR の値が KDD99, NSL-KDD, Kyoto の順で下降している．NSL-KDD では，FAR が KDD99 と比べて高々0.15 程度の増加に留まっているが，TPR は KDD99 と比べて 0.4 程度低くなっている．また，Kyoto では，TPR は KDD99 と比べてほとんど変動はないが，FAR については，KDD99 よりも値が大きい NSL と比べても 0.6 以上の増加がみられる．このようなことから，総合指標である AUR が下降しているのは NSL の場合は TPR が低くなったからであり，Kyoto の場合は FAR が高くなったからであると言える．このようなことを総合すると，KDD99 に対して，セッションの冗長性などを修正した NSL-KDD においては，攻撃の検知率(TPR)が下降し，最新の攻撃を収めた Kyoto においては誤検知（FAR）が増加する傾向にある．その原因としては，NSL-KDD の TPR の低下に関しては，学習データにおいて，攻撃通信の情報量が減ったため，攻撃通信の検知率が低下した可能性がある．また，Kyoto に関しての FAR の増加については，特に NB と SVM で顕著やように，FAR がほぼ 1.0 ということは，ほぼすべての正常通信を攻撃とみなしていると考えられ，どの NIDS 手法ともほとんど判別機能が働いていない状態であると考えられる．つまり，最新の攻撃に対応できない可能性がある．

＊結論

以上をまとめると，開発した NIDS 手法を評価する際には，KDD99 では TPR が高く，FAR を低く抑えられていても, KDD99 の冗長性などの欠点を改良してより現実のネットワークデータに近づけた NSL-KDD での評価では TPR が下がってしまう可能性があることが分かった．また，最新の攻撃を含む Kyoto2016 を用いた評価では, KDD99 の評価が良好な手法でも，判別能力を失う可能性があることも合わせて分かった．

以上の結果から，NIDS 手法の評価を行う場合にはひとつのデータセットのみで行うのではなく，次の観点から複数のデータセットで行う必要があると考える．

①過去の論文の手法との比較を行うために，過去の論文で評価用に使用されてきた KDD99 を使用する

②KDD99 の欠点を改良して現実のネットワークに近づけたデータでの評価を行うために，NSL-KDD での評価を行う

③最新の攻撃への対応能力を評価するために Kyoto2016 を用いた評価を行う

特に，③の最新の攻撃への対応能力を評価することについては，本稿の実験結果より，KDD99 や NSL-KDD ではある程度の性能を発揮しても，Kyoto2016 では判別能力がほとんど発揮されない結果となっているため，ユーザが実際のネットワークに適用する際の参考資料として，重要な評価要素となると考えられる．

このように目的別に複数のデータセットによる評価を行うことにより，開発した NIDS 手法の性質をより正確に読者に伝えることができるものと考える．

なお，ハニーポットのような実験用のネットワークではなく，実際に運用されているネットワーク（実ネットワーク）のデータを用いた評価が必要であるという意見もある．確かに実ネットワークのデータによる評価ができれば NIDS 手法の評価に対して大変有益であることは間違いない．しかし，読者による検証結果の再現などの観点から，実ネットワークのデータを用いる場合には次の要件を満たす必要があるということを，著者らは文献[2]の中で指摘した．

（要件 1）取得した実ネットワークのデータが公開されていること

（要件 2）環境が異なるネットワークのデータを混ぜないこと

セキュリティポリシーが厳しい現在では，自組織のネットワークデータであっても取得が制限される場合が多く[20]，上記の要件を満たしながら実ネットワークのデータを用いて評価するのは実質的には大変困難であると考えられる．そこで，本稿では公開されている実験用データセットを用いて NIDS 手法を評価する方法を検討した．

4. まとめ

本研究を通じて，KDD99，NSL-KDD，Kyoto の 3 つの NIDS 評価用セッション型データセットの特徴を明らかにし，広く知られた機械学習手法を上記 3 つのデータセットで評価することにより，手法が有効に機能するか否かについても確認した．その結果，KDD99 で TPR も FAR も有効な値を示していても，NSL-KDD では TPR が下降し，Kyoto では FAR が場合によって 1.0 付近にまで上昇することを確認した．その原因として，NSL-KDD の TPR の低下に関しては，KDD99 の攻撃通信を整理したことにより，学習データにおける攻撃通信のモデル化に影響を与え，Kyoto における FAR の上昇に関しては，各手法が最新の攻撃の判別に対応していない可能性について指摘した．

以上の結果，過去の手法との比較には KDD99，最新の攻

撃への対応には Kyoto2016 というように目的別に適したデータセットによって NIDS 手法を評価することが必要であることを示した．特に今回の実験では，Kyoto2016 を用いた最新の攻撃への対応について，KDD99 や NSL-KDD では有効である手法も，その機能を十分に発揮できない可能性があるという結果になった．従って，KDD99，NSL-KDD，Kyoto2016 など複数のデータセットを用いて評価を行うことにより，NIDS 手法を様々な観点から評価することが重要であることを示した．このように目的別に複数のデータセットを用いて評価することにより，ユーザは，それぞれの目的やネットワーク環境に適した NIDS 手法を選択するための知見を得ることができるものと考えられる．

今後，KDD99 や NSL-KDD，Kyoto に関して，一部攻撃を削除して評価するなど様々な加工を行い NIDS 手法の評価を行うことによって，上記の原因を明らかにする予定である．また，本稿では NIDS 手法を Weka のデフォルトのままパラメータチューニングを行わずに評価を行ったが，今後，パラメータチューニングを行い，最適条件での評価を行い，更なる考察を加える予定である．

謝辞

本研究は JSPS 科研費 JP17K00187 の助成を受けたものです．この場を借りて，感謝の意を表します．

参考文献

[1] Glass-Vanderlan, Tarrah R., Iannacone, Michael D.. “A Survey of Intrusion Detection Systems Leveraging Host Data ”. https://arxiv.org/pdf/1805.06070.pdf, (参照 2018-09-08).
[2] 高原尚志．ネットワーク型侵入検知における研究結果の検証に関する一検討．ソフトウェアエンジニアリングシンポジウム 2018(SES2018)論文集. 2018, pp.170-173 .
[3] 高原尚志．ネットワーク型侵入検知システム評価用セッション型データセットに関する一考察．コンピュータセキュリティシンポジウム 2018(CSS2018), 2018（印刷中）.
[4] 高原尚志．ネットワーク型侵入検知の評価用データセットに関する提案．情報処理学会研究報告. 2018, Vol.2018-SE-199, No.16, p.1-6 .
[5] 高原尚志．マルウェア検知評価用データセットに関する一考察．信学技報. 2018, IN2018-33(2018-09), p.65-70.
[6] “KDD Cup 1999 Data”. http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html, (参照 2018-09-08).
[7] Atilla Ozgur, Hamit Erdem. “A review of KDD99 dataset usage in intrusion detection and machine learning between 2010 and 2015. https://peerj.com/preprints/1954.pdf”. (参照 2018-09-08)
[8] Richard P. Lippmann, David J. Fried, Issac Graf, Joshua W. Haines, Kristoper R. Kendall, David McClung, Dan Weber, Seth E. Webster, Dan Wyschogrod, Robert K. Cunningham, and Marc A. Zissman. Evaluating Intrusion Detection Systems: The 1998 DARPA Off-line Intrusion Detection Evaluation. Proceedings DARPA Information Survivability Conference and Exposition(DISCEX'00). 2000l, Volume2, p.12-26.
[9] “MIT Lincoln Laboratory DARPA Intrusion Detection Evaluation”. https://www.ll.mit.edu/r-d/datasets/1998-darpa-intrusion-detection-evaluation-data-set, (参照 2018-09-08).
[10] Saffa O. Al-mamory, Firas S. Jassim. Evaluation of Different Data Mining Algorithms with KDD CUP 99 Data Set. Journal of Babylon University, Pure and Applied Sciences. 2013, vol.21, no.8, p.2663-2681.
[11] 荒木雅弘．フリーソフトではじめる機械学習入門．森北出版株式会社, 2014, 262p.
[12] Chordia Anita S, Sunil Gupta. An Effective Model for anomaly IDS to Improve the Efficiency. Proc. 2015 International Conference on Green Computing and Internet of Things (ICGCIoT). 2015, p.190-194 .
[13] Mahbod Tavallaee, Ebrahim Bagheri, Wei Lu, and Ali A. Ghorbani. A Detailed Analysis of the KDD CUP 99 Data Set. Proc. 2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications (CISA 2009). 2009, p.1-6 .
[14] “NSL-KDD Datasets Research Canadian Institute for Cybersecurity UNB (online)”. http://www.unb.ca/cic/datasets/nsl.html, (参照 2018-09-08).
[15] S. Revathi, A. Malathi. A Detailed Analysis on NSL-KDD Dataset Using Various Machine Learning Techniques for Intrusion Detection. International Journal of Engineering Research & Technology (IJERT). 2013, vol.2, Issue 12, p.1848-1853.
[16] Jungsuk SONG, Hiroki Takakura, and Yasuo Okabe, “Description of Kyoto University Benchmark Data”. http://www.takakura.com/Kyoto_data/BenchmarkData-Description-v5.pdf, (参照 2018-09-08).
[17] “Traffic Data from Kyoto University's Honeypots ”. http://www.takakura.com/Kyoto_data/, (参照 2018-08-09).
[18] Jungsuk SONG, Hiroki TAKAKURA, Yasuo OKABE. Cooperation of Intelligent Honypots to Detect Unknown Malicios Code. In Proceedings of Workshop on Information Security Threats Data Collection and Sharing (WOMABT). 2008, p.31-39.
[19] Jungsuk Song, Hiroki Takakura, Yasuo Okabe, Masashi Eto, Daisuke Inoue, Koji Nkao. Statistical Analysis of Honeypot Data and Building of Kyoto 2006+ Dataset for NIDS Evaluation. In Proceedings of The 1st International Workshop on Building Analysis Datasets and Gathering Experience Returns for Security (BADGERS’11). 2011, p.29-36, (2011).
[20] 多田竜之介，小林良太郎，嶋田創，高倉弘喜. NIDS 評価用データセット：Kyoto 2016 Dataset の作成．情報処理学会論文誌. 2017, Volume 58, No.9, p.1450-1463.
[21] V. N. Vapnik and A. Ya. Lerner. Pattern Recognition Using Generalized Portrait Method. Automation and Remote Control. 1963, vol.24, no.6, p.774-780.
[22] Bernhard E. Boser, Isabelle M. Guyon, Vladimir N. Vapnik. A Training Algorithm for Optimal Margin Classifiers. Proc. the fifth annual workshop on Computational learning theory (COLT ’92). 1992, p.144-152.
[23] George H. John, Pat Langley. Estimating continuous distributions in Bayesian classifiers. Proc. the Eleventh conference on Uncertainty in artificial intelligence (UAI'95). 1995, p.338-345.
[24] Kotsiantis, S B.. Decision trees: a recent overview. Artificial Intelligence Review. 2013, vol. 39, no. 4, p.261-283.
[25] The University of Waikato. “Weka 3 - Data Mining with Open Source Machine Learning Software in Java”. http://www.cs.waikato.ac.nz/ml/weka/, (参照 2018-09-08).
[26] Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, Ian Witten. The WEKA data mining software: an update. *ACM SIGKDD Explorations Newsletter*. 2009, vol.11, no.1, p.10-18.