Random Forests と K-Means 法によるハイブリッド式アノマリ検知方式

高原　尚志[†1]

概要：ほとんどすべてのコンピュータがネットワークに接続されている今日，外部からのサイバー攻撃を検知することは必須である．日々新たなサイバー攻撃が生まれていることを考えると既知の攻撃は勿論，新たな攻撃にも対応した攻撃検知システムが必要である．サイバー攻撃を検知する手法として機械学習の手法が用いられている．機械学習の手法には，過去の攻撃データを学習データとして用いる教師あり学習と学習データを用いない教師なし学習がある．広く知られた教師あり学習である Random Forest を用いると高い確率で学習データにある攻撃を検知することができるが，学習データにない攻撃に対する検知率は低い．これに対して広く知られた教師なし学習である K-Means 法を用いると，学習データにあるなしの区別なく攻撃を検知することができるが，正常通信に対する攻撃通信の割合が極端に高い場合には，検知率は高くない．そこで本稿では，Random Forest で検知した攻撃を通信データから除去すること（スクリーニング）によって攻撃通信の数を減少させ，このデータに対して K-Means 法を適用することによって，K-Means 法による検知率を高め，その結果として，学習データない攻撃も含めた攻撃全体の検知率を高めることを試みた．その結果，提案手法では，既存の手法を単独で用いる場合よりも攻撃全体の検知率が向上した．

キーワード：K-Means 法，Random Forests，機械学習済攻撃検知，Detection Rate

Hybrid Anomaly Detection System with Random Forests and K-Means

Hisashi Takahara[†1]

Abstract: Today, most computers are connected to networks, making detection of cyber-attack from outside essential. Every day, new attacks are made via the Internet, so not only known attacks but also unknown attacks need to be detected. There are machine learning methods which use training data about known attacks as methods to detect unknown attacks. These have provided high detection rates in existing papers. However those detection rates include known attacks. Therefore, in this paper, we compare some representative machine learning methods by excluding known attacks from test data, and including only unknown attacks, using the well-known KDD99 dataset. The results indicated super-vised methods that were previously considered to provide good results showed low detection rates, whole the k-means method which is unsupervised maintained a high detection rate. Moreover, to verify the above, we evaluated the results with the separate Kyoto2006+ dataset. The k-means method still kept a high detection rate.

Keywords: K-Means, Random Forests, Machine Learning, Anomaly Detection, Detection Rate

1. はじめに

1.1 動機と背景

今日,インターネット上では,多くのサイバー攻撃が日々生み出されている．これに対応するため，サイバー攻撃の検知は大変重要である．サイバー攻撃の検知には，過去の攻撃パターン（シグネチャ）を参考に攻撃を検知するミスユース型検知と正常通信を定義してそれ以外の通信を攻撃とみなすアノマリ型検知がある[17]．ミスユース型検知では,過去の攻撃パターンによって攻撃か否かを決めるため，新たな攻撃を検知できないという課題がある．一方，アノマリ型検知では，正常通信以外は攻撃通信とするため，正常通信を正確に定義することが求められる．新たな攻撃にも対応した攻撃検知の方法として，機械学習を利用した方法が提案されている．機械学習の方法には，過去の攻撃データを学習データとして参考にしながら，現在の攻撃に対応する教師あり学習と学習データを用いない教師なし学習がある．教師あり学習では，過去の攻撃パターンを学習データとして，攻撃か否かを判断する．そのため，教師あり学習手法では，学習データにある攻撃（以降，学習済攻撃と称す）においては高い検知率を示すことが知られている[27][33][13]が，一方で，学習データにない攻撃（以降，非学習済攻撃と称する）に対する検知率は低い．従って，教師あり学習を用いた手法で正常通信と判定されても，その中に非学習済攻撃が混ざってしまう可能性が高い．これに対して，教師なし学習の広く知られた手法である K-Means 法では，学習，非学習にかかわらず攻撃を検知することができるが，攻撃通信の割合が正常通信に対して極端に高い場合には，検知率が低くなってしまう．これは，K-Means 法では，分類時にクラスタ内のノード数を同程度にする傾向があることに由来するものと考えられる．これを解決す

†1 新潟県立大学
University of NIIGATA PREFECTURE

るためには，攻撃通信の割合を減少させる必要があり，攻撃通信の割合を正常通信と同程度かそれよりも低くすることができれば，学習，非学習に関係なく高い確率で攻撃を検知することができる．

そこで本研究では，教師あり学習手法と K-Means 法の上記の特色を利用して，教師あり学習手法と K-Means 法を直列に組み合わせた，ハイブリッド式のアノマリ型攻撃検知手法を提案する．提案手法を用いれば，新種の攻撃が起こってからシグネチャの配布などその対策が講じられるまでの間の攻撃，いわゆるゼロデイ攻撃にも対応できるものと考えられる．

なお，既存の研究では，シグネチャや学習データにある攻撃など想定した過去ある時点までに攻撃データがある攻撃を既知の攻撃，known attack，know intrusion など，シグネチャや学習データにないなど想定した時点でデータがない攻撃を未知の攻撃，unknown attack，unknown intrusion などと称しているものもある[5] [8] [14] [29][30]が，本稿ではこれらの名称をそれぞれ学習済攻撃，非学習済攻撃に統一して使用する．

1.2 既存研究

文献[29]の中で Reda らは，教師あり機械学習手法である Random Forest[20]と教師なし機械学習手法である K-Means 法[2]に独自の改良を加えた手法を組み合わせたハイブリッド方式を提案している．また，検証用のデータセットには，広く知られた攻撃検知用ベンチマークデータセットである KDD99[35][36]を独自に改良したデータセットを用いている．

文献[29]では，まず Random Forest をミスユース型検知として用い，その後，K-Means 法を改良した方式をアノマリ型検知として用いている（以降 Reda 方式と称す）．この際，各特徴量に独自の方式で重み付けを行い，K-Means 法を適用するという改良を加えている（以降，weighted K-Means=wk-Means 法と称す）．また，Random Forest で攻撃と判定された通信データの中からランダムに通信データを抽出し，これを正常通信と判定された通信データに加え，このデータに対して wk-Means 法を適用している．

文献[29]では，

①ミスユース型検知では，DR はそれほど高くない（92.73%）が，一方で FAR は低く抑えることができる（0.54%）

②アノマリ型検知では，検知率（Detection Rate=DR）は非常に高い(99%)が，一方で誤検知率（False Alarm Rate=FAR）も高くなってしまう（12.6%）

というミスユース型検知とアノマリ型検知の特色を示した上で，この特色を活かした上記のハイブリッド方式を提案し，DR を高く維持しながら（98.3%），FAR を低く抑える（1.6%）ことを実現している．

ところで，文献[29]では，Random Forest で正常通信と判定されたデータに，攻撃と判定されたデータを加え，そのデータに対して wk-Means 法を適用しているが，Random Forest で攻撃と判定された通信データを完全に除去（スクリーニング）したデータに対して，wk-Means 法を適用すれば，更なる DR の向上が期待できると考えられる．しかし，wk-Means 法の詳細なプログラムは明らかにされていないため，本稿では，スクリーニングしたデータに対して通常の K-Means 法を適用し，3.1 で示す方法で簡易的に Reda 方式を実現した（以降，これを簡易 Reda 方式と称す）．

1.3 提案と貢献

前述の既存研究を踏まえて，本稿では，広く知られた教師あり学習方式であり，3 章での検証の結果，教師あり学習の中で最も高い DR を示した Random Forest(RF)を用いて（表 2）学習済攻撃をスクリーニングした後，こちらも広く知られた教師なし学習方式である K-Means 法を用いて非学習済攻撃を検知するハイブリッド式アノマリ検知方式を提案する．これにより学習済攻撃の検知に優れている RF と学習，非学習に関係なく攻撃を検知できる K-Means 法のそれぞれの特色を活かして高い確率での攻撃検知が期待できる．また提案手法では，RF と K-Means 法を用いているが，両者ともに広く知られた方式であるため，機械学習用アプリケーションにおいてもライブラリが用意されているものがあり，特に手法に改良を加えていない限り，容易に追試を行うことも可能である．また，本稿では K-Means 法におけるクラスタ数を 2 として検証を行っているが，これに加えて，適切なクラスタ数を設定できる X-Means 法での検証も合わせて行った．

1.4 関連研究との比較

1.4.1 機械学習の各手法

本稿では，機械学習の手法として，提案手法で用いられる RF，K-Means 法以外にも，比較のため DT，Naïve Bayes(NB)，Support Vector Machine(SVM)などの手法を用いる．ここでは，それぞれの手法を関連研究として述べる．

K-Means 法は，1967 年に J. MacQueen によって命名され[2]，学習データを用いない，教師なしの機械学習手法であるクラスタリングの代表的な手法のひとつで，アルゴリズムは次の通りである[24][10][7][18][38][23].

（１）予めクラスタ数 k を決め，各クラスタに対する代表値を設定する．（初期の代表値を決める方法としては，与えられたデータをランダムに k 個選んで，各クラスタの代表値とする方法などがあるが，本稿では，ランダム関数を用いて初期値を発生させる方法を用いた．）

（２）各ノードに対して各クラスタの代表値からの距離を測定して，最も短いクラスタに対象ノードを所属させる．

（３）所属したノードの平均値を計算して，改めてクラスタの代表値とする．

（４）代表値が変わらなくなるまで，（２），（３）を繰り返す．

（５）代表値が変わらなくなった時点で，クラスタリング

を終了し，各ノードの所属を決定する．

K-Means 法には，球形のクラスタを形成する傾向がある，クラスタの大きさや濃度を均等に仕様とする，外れ値の影響を受けやすいなどの課題が指摘されている[39].

提案手法で用いられる RF は複数の DT を組み合わせて用いるアンサンブル手法ではるが，そのもととなる DT は，特徴値ごとに条件を設定して，最も分類効率が大きな特徴値から順に分類を行うことにより，ノードの所属カテゴリを予測する手法である[26] [32].

RF は，2001 年に Leo Breiman によって提案された[20]が，複数の DT を用いて分類を行うアンサンブル手法[32]である．各 DT において，ランダムに抽出された特徴量の中から最も分類能力の高い特徴量を選択して，分類を行うことを繰り返す．攻撃検知に応用した場合，高い分類能力を示すことが広く知られている[27][33][13].

NB は，各特徴量が独立であると仮定して，事前に与えられる学習データから尤度を推定し，評価データが所属するカテゴリを求める方法である[4] [32]．また，SVM は，1963 年に V. N. Vapnik らによって線形モデルが[1]，1992 年に Bernhard E. Boser らによって非線形モデルが提案された[3]が，学習データをもとに各データを分離する超平面の中でマージンを最大化するものを求めてモデルを形成して，評価データを識別する．

1.4.2 スクリーニング

提案手法では，RF でデータをスクリーニングした後，K-Means 法で攻撃検知を行う．ここでは，スクリーニングを用いた研究について述べる．

文献[5]の中で村井らは，グラフベースの半教師あり学習を用いて評価データの中から既知の攻撃（学習済攻撃）や設定ミスによる異常通信をスクリーニングすることによって，未知の攻撃（非学習済攻撃）の検知を容易にする方式を提案している．

文献[14]の中で Song らは，学習データから正常通信のみを抽出し，クラスタリングすることによって数個のクラスタに分割した上で，正常通信のモデルを作成，これを用いて one-class SVM により，未知の攻撃（非学習済攻撃）を検知する手法を提案している．

1.4.3 ハイブリッド方式

文献[8]の中で山田らは，シグネチャを用いてミスユース型検知を行うシグネチャ型侵入検知手法（Intrusion Detection System=IDS）を用いて直近の検知結果を得た後に，これをもとにした学習データを自動作成，作成した学習データを用いて教師ありの機械学習手法のひとつである決定木(Decision Tree=DT)[26]にて検知を行う方式を提案している．この方式を用いれば，同じ環境の直近の結果を学習データとして検知を行うことができるため，シグネチャ型検知において検知された攻撃に類似した攻撃，いわゆる亜種を検知することができる．山田らの方式では，教師あり学習である決定木のみを用いているが，学習データの影響を受けない教師なし学習手法とも組み合わせることにより，更に新種の攻撃への対応も期待できる．

文献[30]の中で Kim らは，決定木（C4.5）(Decision Tree=DT)[26]と one-class SVM[12][22]を組み合わせることによって既知の攻撃（学習済攻撃）だけではなく未知の攻撃（非学習済攻撃）も検知するハイブリッド方式を提案している．文献[30]では，新たなトラフィックが来ると，まず決定木によってそのトラフィックが（既知の）攻撃（学習済攻撃）か否かを判別し，攻撃判定の場合にはブロックし，攻撃でない場合には次のステップとして，one-class SVM によってそのトラフィックが攻撃か否かを判別する．具体的な数値は明らかにされていないが，示されている ROC 曲線などから，これにより，DR を高め，FAR を低く抑えることができることが示されている．しかし，SVM は K-Means 法などと比べて処理時間が多くかかることが広く知られているため，処理時間を短縮することが課題となると考えられ，one-class SVM の代わりに K-Means 法を用いることにより処理時間の短縮が期待できる．

2. 提案手法

2.1 提案手法の流れ

提案手法では，次の 2 つの STEP で攻撃検知を行う（図 1）.

(STEP1)　RF によるスクリーニング

RF にて攻撃を検知し，検知した攻撃を元の通信データから削除する

(STEP2)　K-Means 法による攻撃検知

(STEP1)で作成されたデータ（元の通信データから RF で検知された攻撃を削除したデータ）を用いて，K-Means 法による攻撃検知を行う

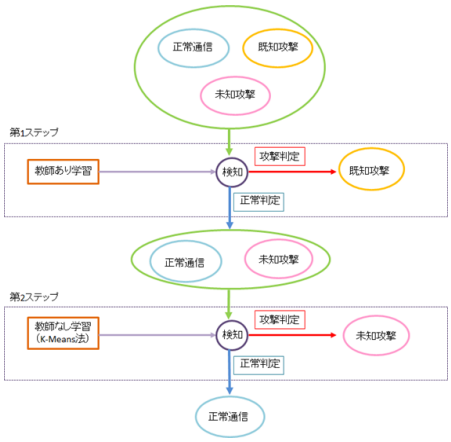

図 1 提案手法

Fiture1. Proposed Method

ここで, STEP1 と STEP2 のいずれかの時点で攻撃と判定された場合には, システム全体としての判定を攻撃とする.

3. 検証実験

3.1 目的

非学習済攻撃の検知に関して，教師あり学習手法及び教師なし学習手法を単独で使用した場合及び本稿と同様に RF と K-Means 法を用いた Reda らのハイブリッド手法 [29] と提案手法の検知率（DR）を比較して，提案手法が有効であることを示す.

ただし，文献[29]では，使用したプログラムやデータの詳細が明らかにされていないため，次のような形で検証を行う（簡易 Reda 方式）.

①データは本稿において他の手法で使用するものと同じものを用いる

②いずれの段階でもデータの加工は行わず, RF と K-Means 法のいずれの手法も，データ全体に対して適用する

③RF，K-Means 法のいずれか（含両方）で攻撃と判定された通信を，全体の判定においても攻撃と判定する

なお，本稿では，評価指標として DR の他に，誤検知率（False Positive Rate=FAR）と正解率(Accuracy=ACC)も合わせて示し，課題についても提示する.

3.2 環境

3.2.1 データセット

本稿では検証用のデータセットとして KDD99 を用いる.

KDD99 は，1998 DARPA Intrusion Detection Evaluation Data Set（以降，DARPA98 と称す）[37]の通信データ（TCP dump データ）をセッション単位のデータとして加工したデータセットであり，データが古い，データサイズが大きいなどの短所もあるが，現在でも多くの攻撃検知の研究で検証用データセットとして用いられている[35][31]. KDD99 には， 4,893,980 件のデータからなるフルデータ(kddcup.data)とその内の約 10%の 494,020 件のデータを抽出したデータ(kddcup.data_10_percent), そして 311,029 件のデータからなる評価データ(corrected)などが収められている．本稿では，学習データとして kddcup.data_10_percent, 評価データとして corrected を用いる．kddcup.data_10_percent の中には，97,277 件（19.69%）の正常通信と 396,743 件（80.31%）の攻撃通信が収められている. 攻撃通信は 22 種類の攻撃からなり, これが DoS, Probe, U2R，R2L の 4 つのカテゴリに分かれている．corrected の中には, 60,593 件(19.48%)の正常通信と 250,436 件(85.52%)の攻撃通信が収められている[28]. 攻撃通信は 37 種類の攻撃からなり[25]，この内，20 種類は kddcup.data_10_percent と共通の攻撃で，残りの 17 種類は新たな攻撃である．KDD99 における検証実験では, kddcup.data_10_percent を教師あり学習の学習データとして用い，corrected を評価データとして用いるが，学習データである kddcup.data_10_percent に含まれる攻撃を学習済攻撃，評価データである corrected に含まれる攻撃の内 kddcup.data_10_percent には含まれない攻撃を非学習済攻撃と称す． corrected の全攻撃通信 250,436 件の内，171,114 件（68.33%）は学習済攻撃，79,322 件（31.67%）は非学習済攻撃である.

3.2.2 ソフトウエア

本稿では，機械学習用ソフトウエアとして，オーストラリアの Waikato 大学から配布されている WEKA[40][16]を用いた．WEKA は，機械学習用のフリーソフトであり，現在，多くの機械学習の研究で用いられている[35][15].

詳細は以下の通りである.

SVM には，WEKA の SMO アルゴリズムを用い，カーネル関数として exponent 1 の Poly kernel を用いた

DT には，ID3 アルゴリズム[28][32]の拡張である C4.5 の WEKA における実装である J48 を用いた.

NB には，WEKA における実装である NaïveBayes を用いた.

RF では，選択する特徴量の数は，全特徴量数 m に対して，log2m+1 として検証実験を行った．これは，WEKA のデフォルト値である．また，今回の検証実験では，決定木数を 5 としたが, 決定木数 3 から 20 では, 各指標の標準偏差は高々0.012 であった．なお，特徴量の選択は，WEKA の random number generator に RF におけるデフォルトの seed 値 1 を与えて行った[41].

K-means 法では， WEKA の実装である SimpleKMeans にて検証を行った．初期値は，WEKA の random number generator に SimpleKMeans におけるデフォルトの seed 値 10 を与えて，クラスタの数は 2 として検証を行った.

また，本稿では，情報量基準（Bayesian Information Criterion=BIC[32][9]）をもとにして適切なクラスタ数になるまで分割を行う X-Means 法[32][6][9]による検証も行った. その際, WEKA の実装である XMeans を用い, 初期値には, WEKA の random number generator に SimpleKMeans におけるデフォルトの seed 値 10 を与えて，クラスタの数は minNumCluster（最小クラスタ数）を 2 とし, maxNumCluster（最大クラスタ数）を 8 とした.

3.2.3 本稿の STEP2(K-Means 法)におけるクラスタ数とラベリング

K-Means 法では，分類するクラスタの数と分類したクラスタにラベルを付す方法が課題となる．攻撃検知の場合，分類したクラスタに攻撃または正常通信どちらかのラベルを付す必要がある[19][11][29]．文献[29](4.3)の中で Reda らは，分類されたクラスタの内，ミスユース型攻撃検知のフェイズにおいて攻撃通信と判定された通信が多く含まれているクラスタを攻撃クラスタとするとしている. 本稿では, STEP2 の段階の前で，攻撃と判定された通信をスクリーニ

ングするため，STEP2 の段階では，各通信データに攻撃というラベルは付されていない．そのため，文献[29]のラベリング手法を用いることはできない．本稿では，K-Means 法で 2 クラスタに分類し，攻撃通信が多い方のクラスタを攻撃クラスタ,他方のクラスタを正常通信クラスタとする．しかし，実際の分類ではどの通信が攻撃通信かが分からないため，ラベリングは今後の課題である．また，3.2.2 でも述べた通り，本稿では，K-Means 法を改良した X-Means 法での検証も行った．この際，クラスタ数は 4 となったが，クラスタに含まれる攻撃が最も少ないクラスタを正常通信クラスタとし,それ以外のクラスタを攻撃クラスタとした．しかし，K-Means 法の場合同様，実際にはどの通信が攻撃通信か分からないため，ラベリングの問題が残る．

3.2.4 評価指標

以下に混同行列(Confusion Matrix)（表 2）をもとにした，各指標の定義式を示す[17][34][21][32].

表 2 混同行列

Table2 Confusion Matrix

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

DR = TP / (TP+FN) (式 1)

FAR = FP / (FP+TN) (式 2)

ACC = (TP+TN) / (TP+FN+FP+TN) (式 3)

ここで，TP (True Positive)は Positive と正しく予測されたノード数，FP (False Positive)は，実際は Negative であるが Positive と予測されたノード数，FN (False Negative)は，実際は Positive であるが Negative と予測されたノード数，TN (True Negative)は Negative と正しく予測されたノード数である．本稿では，Positive を Attack（攻撃通信），Negative を Normal（正常通信）と読み替えるものとする．

3.3 準備（前処理）

検証実験を行う前に，データを正規化[32]した．また，用いる特徴量は，KDD99 では全 41 個の特徴量の内，テキストデータである Protocol_type，Service，Flag の 3 つの特徴量を除外した 38 個の特徴量を用いた． 前述の通り，KDD99 では，corrected（評価データ）の全攻撃 37 種類から kddcup.data_10_percent_correctd（学習データ）に含まれる攻撃 20 種類を除外した 17 種類の攻撃は学習データに含まれない攻撃であるため非学習済攻撃とし，この攻撃のみを含んだデータを作成し，非学習済攻撃のみの検知を行う際の評価データとした．

3.4 結果

それぞれの手法の攻撃全体及び非学習済攻撃のみについて DR を測定した結果が，表 3.及び図 3.，攻撃全体に対して DR, FAR と ACC を測定した結果が表 4 及び図 4.である．DR は，値が 1 に近づくほど検知率が高く，ACC も値が 1 に近づくほど正確な検知が行われていることとなる．これに対して，FAR は値が 0 に近づくほど誤検知率が小さいことになり，良好な結果となる．従って，DR 及び ACC の値が 1 に近く，FAR の値が 0 に近い状態が最も理想的な状態となる．

表 3 学習，非学習済攻撃に対する検知率（KDD99）

Table3 Detection Rate to Known and Unknown Attack (KDD99)

KDD99(DR)	ALL	NEW
K-Means	0.656	0.986
SVM	0.902	0.091
DT	0.907	0.082
NB	0.900	0.099
RF	0.910	0.100

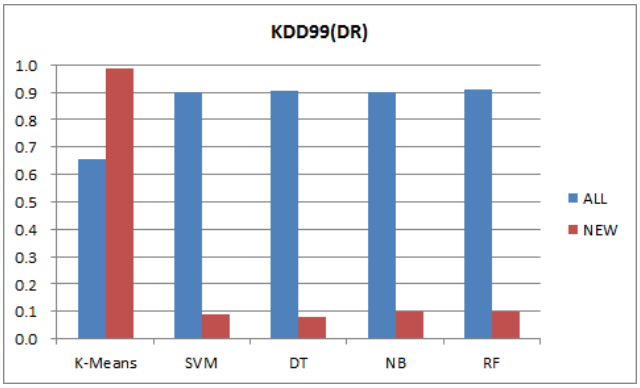

図 3 学習，非学習済攻撃に対する検知率（KDD99）

Fiture3 Detection Rate to Known and Unknown Attacks (KDD99)

表 4　各手法による評価値（KDD99）

Table2 Evaluation Value of each method(KDD99)

KDD99(ALL)	DR	FAR	ACC
K-Means	0.656	0.015	0.720
SVM	0.902	0.016	0.918
DT	0.907	0.017	0.921
NB	0.900	0.026	0.914
RF	0.910	0.005	0.926
Reda(RF+K-Means)	0.910	0.019	0.924
RF+X-Means	1.000	0.624	0.878
New(RF+K-Means)	0.995	0.554	0.888

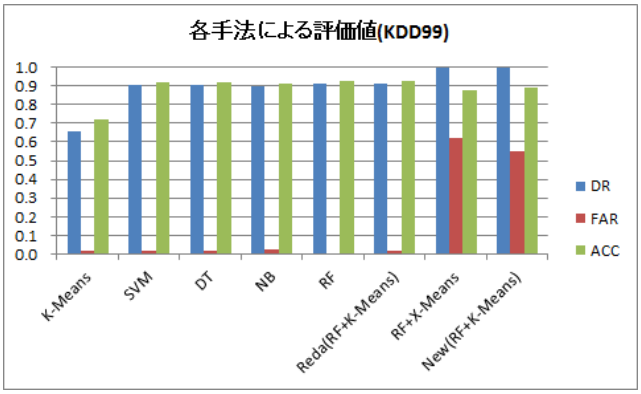

図 4 各手法による評価値（KDD99）

Fiture3 Evaluation Value of each method (KDD99)

検証実験の結果，DR に注目した場合，KDD99 の場合，学習済攻撃も含んだ攻撃全体に対する検知率は K-Means 法が 0.656 であったのに対して，教師あり学習の各手法は 0.900 以上であった. この中でも RF の DR が最も高く 0.910 であった．これに対して，非学習済攻撃の検知率は，教師あり学習の各手法の DR が高々0.100 であるのに対して，K-Means 法の DR は 0.986 となった. この特色を利用して，まずデータ全体の検知率が最も高い RF でデータをスクリーニングした（STEP1）後, K-Means 法で検知を行う（STEP2），RF と K-Means 法のハイブリッド方式を提案した．その結果，提案手法では，DR は 0.995 となり，各手法を単独で用いた場合やデータに対してスクリーニングを行わない簡易 Reda 方式よりも高い検知率を示した（表 3，図 4）．また，提案方式では，STEP2 の K-Means 法のクラスタ数を 2 としたが，手法が適切なクラスタ数を設定する X-Means 法でも検証を行った結果，クラスタ数は 4 となり，DR は 1.000(0.99967)となった.

FAR に関しては，KDD99 においては，教師あり学習の各手法及び簡易 Reda 方式が高々0.026 であったのに対して，提案手法では，0.554 となり，RF+X-Means 法では 0.624 となった． また，ACC に関しては，教師あり学習の各手法及び簡易 Reda 方式は，KDD99 では 0.914～0.926 の範囲の値であるが，提案手法では 0.888, RF+X-Means 法では 0.888 であった.

3.5 考察

検証の結果，教師あり学習手法（SVM，DT，NB，RF）では，攻撃全体に対する DR は 0.90 以上であるが，非学習済攻撃に対する DR は 0.1 未満の値に留まるのに対して，K-Means 法では攻撃全体に対する DR が 0.656 だったのに対して非学習済攻撃の DR は 0.986 となった．これは，教師ありの学習手法が学習データに含まれる学習済攻撃の傾向の影響を受けやすいのに対して，学習データを用いない教師なし学習手法である K-Means 法では，学習済攻撃の傾向に影響を受けないため，非学習済攻撃も検知できるのだと考えられる.

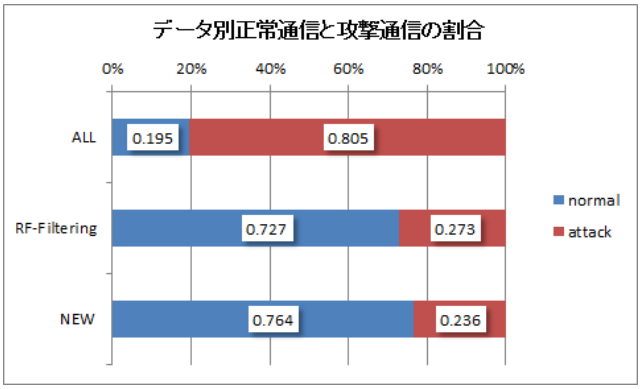

図 5　データ別正常通信と攻撃通信の割合（KDD99）

Fiture5 Ratio of Each Data Set (KDD99)

また，K-Means 法の攻撃全体に対する DR が 0.656 にとどまるのは，K-Means 法が各クラスタに属するノードの数を同程度にする傾向にあることに由来すると考えられる．つまり，通信全体の場合，正常通信と攻撃通信の割合が，0.195 対 0.805 である（図 5）ため，攻撃通信を正常通信クラスタに分類して同じ濃度にしようとしたのに対して，非学習済攻撃のみの場合は，0.764 対 0.236 と正常通信の割合が高くなる（図 5）ため，攻撃通信を正常通信クラスタに分類しようとはせず攻撃通信の DR が高くなるものと考えられる．但し，この場合，正常通信を攻撃通信クラスタに分類しようとする傾向があるため FAR は高くなると考えられる.

提案方式では，この特色を利用して，まず RF で攻撃をスクリーニングして攻撃通信の割合を減少させた後に，K-Means 法を適用するというハイブリッド方式を採用した．このため, RF 適用後の正常通信と攻撃通信の割合が，0.727 対 0.273 となり（図 5），非学習済攻撃のみを抽出した場合

と同様に正常通信の割合が攻撃通信の割合よりも高くなったため，攻撃通信を正常通信に分類せず，DR が 0.995 となったものと考えられる．また，同様の理由で，正常通信が攻撃通信に分類され，FAR も 0.554 と教師あり学習手法や簡易 Reda 方式に比べて高い値となったと考えられる．この影響で ACC も 0.888 と単独の手法や簡易 Reda 方式に比べて低い値に留まったものと考えられる．

ここで，簡易 Reda 方式は，スクリーニングを行わないため，K-Means 法の特色を十分に引き出すことができず，DR が最初のステップである RF に依存して RF と同様の 0.910 に留まったものと考えられるが， 文献[29]では，特徴量に重み付けをした wk-Means 法を用いており，これを用いた場合には，更に DR が向上する可能性がある．

本稿で用いた K-Means 法はクラスタ数を 2 としたが，X-Means 法を適用した結果，クラスタ数は 4 となり，3.2.3 の手法でラベリングを行ったところ，DR は 1.000(0.99967)となったが，FAR は 0.624 となり，K-Means 法を用いた場合の値（0.554）を上回った．

4. まとめ

既存の研究では，KDD99 において，教師ありの機械学習手法を攻撃検知に適用した場合，0.9 以上の検知率（DR）を示していたが，非学習済攻撃のみを抽出して検証を行ったところ，KDD99 においては，教師ありの機械学習手法では，高々0.1 の DR しか得ることができないことを示した．また，教師なしの機械学習手法である K-Means 法を適用したところ，0.986 の DR を得ることができることも合わせて示した．但し，K-Means 法を通信全体に適用した場合の DR は 0.656 と他の手法に比べて低い値となったが，これは，正常通信に比べて攻撃通信の割合が極端に高いためと考えられる．このため，まず RF で攻撃通信をスクリーニングした後，K-Means 法にて分類を行う方式を提案し，その有効性を検証した．その結果，提案方式は DR において 0.995 という他の手法に比べて高い値を示すことを確認し，提案手法の有効性を示すことができた．更に，RF とクラスタ数を適切に設定する X-Means 法とを組み合わせることにより，DR を 1.000(0.99967)まで高めることができることも同時に示した．しかし，K-Means 法，X-Means 法いずれの場合も，分類したクラスタへのラベリングの問題が残ったため，今後，その方法について検討して行く予定である．

また，FAR では，他の手法が KDD99 で高々0.026 であったのに対して，提案手法（RF+K-Means 法）では 0.554，RF+X-Means 法では 0.624 となり，他の手法に比べて高い値となった．今後， DT や one-class SVM など，教師あり学習と教師なし学習の各手法の特色を活かした様々な組合せを試行し，DR を更に向上させ FAR を低く抑えることができる方法を検討して行く予定である．

謝辞 本稿の作成にあたって，九州大学の桜井幸一先生と YAOKAI FENG 先生にご指導頂きました．ここに感謝の意を表します．

参考文献

[1] V. N. Vapnik and A. Ya. Lerner: Pattern Recognition Using Generalized Portrait Method, *Automation and Remote Control.*, vol.24, no.6, p.774-780 (1963).

[2] J. MacQueen: Some methods for classification and analysis of multivariate observations. *Proc. Fifth Berkeley Symp. on Math. Statist. and Prob.*, vol. 1, p. 281-297 (1967).

[3] 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). , p.144-152 (1992).

[4] George H. John, Pat Langley: Estimating continuous distributions in Bayesian classifiers. *Proc. the Eleventh conference on Uncertainty in artificial intelligence* (UAI'95), p.338-345 (1995).

[5] 村井光，正代隆義：効果的なネットワークインシデント検知のための半教師ありデータスクリーニング， *火の国情報シンポジウム 201 論文集 5*, 2B-3, pp.1-8, 佐賀大学 (2015).

[6] Dan Pelleg, Andrew Moore: X-Means: Extending K-means with Efficient Estimation of the Number of Cluseters, *Proc. ICML2000*, pp.727-734, Stanford University (2000).

[7] 松田一孝，筧一彦，胡振江，武市正人：データマイニングのアルゴリズム記述を容易にする拡張行列演算の提案．*情報処理学会論文誌 プログラミング.*, vol.46, no.SIG 11 (PRO 26), pp.1-15 (2005).

[8] 山田明，三宅優，竹森敬祐，田中俊昭：学習データを自動生成する未知攻撃検知システム，*情報処理学会論文誌*, vol.46, no.8., p.1947-1958 (2005).

[9] 石岡恒憲: x-means 法改良の一提案ーk-means 法の逐次繰り返しとクラスターの再併合ー，*計算機統計学*, 第 18 巻，第 1 号，pp.3-13 (2006).

[10] [10]上田達也，安倍広多，石橋勇人，松浦敏雄: P2P 手法によるインターネットノードの階層的クラスタリング，*情報処理学会論文誌*, vol.47, no.4, p.1067-1076 (2006).

[11] Jungsuk Song, Hiroki Takakura, Yasuo Okabe, Yongjin Kwon: A Robust Feature Normalization Scheme and an Optimized Clustering Method for Anomaly-Based Intrusion Detection System, *Proc. Advances in Databases: Concepts, Systems and Applications*, pp.140–151 (2007).

[12] 赤名武志: "One Class Support Vector Machine を用いたミスラベルデータ検出手法の提案，*筑波大学大学院博士課程システム情報工学研究科社会システム工学専攻修士論文* (2008).

[13] Jiong Zhang, Mohammad Zulkernine, and Anwar Haque: Random-Forests-Based Network Intrusion Detection Systems, *IEEE Transactions on Systems, Man, and Cybernetics-Part C: Applications and Reviews*, Vol.38, No.5, pp.649-659 (2008).

[14] Jungsuk SONG, Hiroki TAKAKURA, Yasuo OKABE, and Yongjin KWON: Unsupervised Anomaly Detection Based on Clustering and Multiple One-Class SVM, *IEICE TRANS. COMMUN.*, Vol.E92-B, No.6, pp.1981-1990, (2009).

[15] Tavallaee, Mahbod, Bagheri, Ebrahim, Lu, Wei, Ghorbani, Ali-A: A Detailed Analysis of the KDD CUP 99 Data Set. *Proc. the Second IEEE Symposium on Computational Intelligence for Security and Defence Applications 2009*, pp.1-6 (2009).

[16] [16]Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, Ian Witten: The WEKA data mining software: an update, *ACM SIGKDD Explorations Newsletter*, vol.11, no.1, p.10-18 (2009).

[17] 山田明: ネットワーク侵入検知システムの高度化に関する研

究, *東北大学博士学位論文* (2009).
[18] [18]Jain, Anil K.. Data clustering: 50 years beyond K-means, *Pattern Recognition Letters*, Vol. 31, No. 8, pp.651-666 (2010).
[19] Moriteru Ishida, Hiroki Takakura, Yasuo Okabe: High-Performance Intrusion Detection Using OptiGrid Clustering and Grid-based Labelling, *Proc. 2011 IEEE/IPSJ International Symposium on Application and the Internet (SAINT 2011)*, pp.11-19 (2011).
[20] LEO BREIMAN: Random Forests, *Machine Learning*, vol.45, no.1. p.5-32 (2011).
[21] Natesan, P, Balasubramanie, P, Gowrison, G.: Improving the attack detection rate in network intrusion detection using adaboost algorithm, *Journal of Computer Science*, Vol. 8, No. 7, pp.1041-1048 (2012).
[22] Mennatallah Amer, Markus Goldstein, Slim Abdennadher: Enhanceing One-class Support Vector Machines for Unsupervised Anomaly Detection, *Proc. the ACM SIGKDD Workshop on Outlier Detection and Description (ODD '13)*, pp.8-15 (2013).
[23] [23]Supreet Kaur, Usvir Kaur: A Survey on Various Clustering Techniques with K-means Clustering Algorithm in Detail, *International Journal of Computer Science and Mobile Computing*, vol.2, no.4, p.155-159 (2013).
[24] [24]Mukesh Kumar Choudhar, Mandeep Singh Saini: Palvee. Classification by K-Means Clustering, *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)*, vol.2, no.5, p.1684-1688 (2013).
[25] 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)*, vol.2, Issue 12, p.1848-1853 (2013).
[26] Kotsiantis, S B.: Decision trees: a recent overview, *Artificial Intelligence Review*, vol. 39, no. 4, p.261-283 (2013).
[27] Saint Murat GIRAY, Aydin Goze POLAT: Evaluation and Comparison of Classification Techniques for Network Intrusion Detection, *Proc. 13th International Conference on Data Mining Workshops (ICDMW 2013)*, p.335-342 (2013).
[28] 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*, vol.21, no.8, p.2663-2681 (2013).
[29] Reda M. Elbasiony, Elsayed A. Sallam, Tarek E. Eltobely, Mahmoud M. Fahmy: A hybrid network intrusion detection framework based on random forests and weighted k-means, *Ain Shams Engineering Journal*, Volume 4, Issue 4, pp.753-762, (2013).
[30] Gisug Kim, Seungmin Lee, Sehun Kim: A novel hybrid intrusion detection method integrating anomaly detection with misuse detection, *Expert Systems with Applications*, Volume 41, Issue 4, pp.1690-1700 (2014).
[31] Martina Troesch and Ian Walsh: Machine Learing for Network Intrusion Detection, *Final Report for CS 229*, Stanford University (2014).
[32] 荒木雅弘: フリーソフトで始める機械学習入門, *森北出版株式会* (2014).
[33] Sundus Juma, Zaiton Muda, M.A. Mohamed, Warusia Yassin: Machine Learning Techniques for Intrusion Detection System: A Review, *Journal of Theoretical and Applied Information Technology (JATIT)*, vol.72, no.3, p.422-429 (2015).
[34] 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)*, p.190-194 (2015).
[35] Atilla Ozgur, Hamit Erdem: A review of KDD99 dataset usage in intrusion detection and machine learning between 2010 and 2015 (online), available from <https://peerj.com/preprints/1954/> (accessed 2016-05-21).
[36] Irvine: KDD Cup 1999 Data (online), available from <http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html> (accessed 2016-05-21).
[37] Lincoln Laboratory, Massachusetts Institute of Technology (online), available from <https://www.ll.mit.edu/ideval/data/1998data.html> (accessed 2016-05-21).
[38] Johannes Blömer, Christiane Lammersen, Melanie Schmidt, Christian Sohler: Theoretical Analysis of the k -Means Algorithm - A Survey (online), available from <https://arxiv.org/pdf/1602.08254v1> (accessed 2016-05-21).
[39] Tan,Steinbach, Kumar: K-Means Cluster Analysis (online), available from <https://www.yumpu.com/en/document/view/26521444/k-means-cluster-analysis-chapter-3-3-ppdm-cl-ass> (accessed 2016-05-21).
[40] [40]The University of Waikato: Weka 3 - Data Mining with Open Source Machine Learning Software in Java (online), available from <http://www.cs.waikato.ac.nz/ml/weka/> (accessed 2016-05-21).
[41] RandomForest (online), available from <http://weka.sourceforge.net/doc.dev/weka/classifiers/trees/RandomForest.html> (accesed 2016-05-21).