SCIS 2018 2018 Symposium on
Cryptography and Information Security
Niigata, Japan, Jan. 23 - 26, 2018
The Institute of Electronics,
Information and Communication Engineers

プロキシログのクラスタ間遷移に着目した異常検知手法の評価

Evaluation of Anomaly Detection Algorithms on Cluster Sequence of Proxy Server Logs

石井将大 * Masahiro Ishii
猪俣漸 * Zen Inomata
オユンビレグチンゲン † Chingun Oyunbileg
京山剛大 † Takehiro Kyouyama
黒田渓介 * Keisuke Kuroda
佐藤敦 ‡ Atsushi Sato
菅谷光啓 † Mitsuyoshi Sugaya
田中圭介 * Keisuke Tanaka
千葉龍一郎 * Ryuichiro Chiba
西脇利知 * Toshikazu Nishiwaki
橋本幸典 † Yukinori Hashimoto
廣野壮志 † Soushi Hirono
松浦知史 * Satoshi Matsuura
森健人 * Kento Mori

あらまし 組織内ネットワークへの不正侵入などのサイバー攻撃による犯罪は大きな社会問題である．セキュリティ機器や，イベントログ管理システムを用いた対策が行われているが，これらの機器が出力する大規模なデータを分析するオペレータへの負担が非常に大きく，通信データそのものの管理コストも大きい．さらに，通信の良性・悪性を判断し，教師ラベルを付与することは非常に困難である．本論文では，管理コストが低い HTTP プロキシログに対してクラスタリングを行い，ホストごとの通信遷移によるクラスタシーケンスを用いてログにスコアを付け，攻撃検知と攻撃分類の可能性について評価を行う．

キーワード プロキシログ解析，異常検知，クラスタリング，特徴選択，クラスタ遷移

1 はじめに

組織内ネットワークへの不正侵入などのサイバー犯罪は大きな社会問題である．IoT 技術の普及に伴い，ネットワーク機器が増加することによって，今後もサイバー犯罪は増加の一途を辿ると考えられている．このことから，サイバー攻撃対策の一つである攻撃検知システムの性能向上が望まれている．

攻撃検知システムとしては，セキュリティ機器やイベントログ管理システムなどが用いられているが，これらが出力する大規模なデータを分析するオペレータの負担は非常に大きい．そのため，機械学習を用いてオペレータの負担を軽減することが注目されている．しかし，教師あり学習が必要とする教師データを得るためには，マルウェアなど様々な攻撃による悪性通信ログの解析が必要であり，特に大規模な通信ログデータに対して教師ラベルを付与し，教師データを用意することは非常に困難である．そのため，教師なし学習を用いた異常検知の研究が行われている．異常検知手法は様々なものがあり，正常モデルを確率的に定め，そのパラメータをデータから推定して異常を検知するものや，対象データの類似度などを用いて他のデータから離れたものを検知するものが考えられている．また，ネットワーク通信データを利用した検知システムにおいて，機械学習の入力データのサイズは一般的に非常に大きく，データそのものの管理コストが大きいことも困難な点である．

本論文では，教師データが不要な教師なし学習を用いた異常検知に着目して不審な通信を検知することを目的とする．比較的管理コストが低いデータである HTTP プロキシサーバのログデータを対象として，既存の悪性データ通信セットを利用して検知システムの評価を行った．また，機械学習のアルゴリズムにおいて，検知精度の向上と計算コストの削減のために適切な特徴を選択す

* 東京工業大学 本研究の一部は野村総合研究所，JST CREST JPMJCR14D6，JST OPERA，JSPS16H01705，17H01695 の助成を受けています．本研究では MWS 組織委員会が提供する MWS Datasets 2017 を利用しています．
† NRI セキュアテクノロジーズ
‡ 野村総合研究所

ることが重要であるため，本論文においては教師なし学習のための特徴選択アルゴリズムを採用した攻撃検知を行い，特徴選択を行わない場合との比較をした．

2 関連研究

ネットワークに対する攻撃検知に関して，通信のクラスタシーケンスを用いたものとして佐藤，山口，嶋田，高倉の研究 [11] がある．佐藤らは，セッションデータにクラスタリングを適用し，セッションデータ全体をクラスタに分割することで得られる各ホストの通信シーケンスを用いて不自然な遷移パターンを検知している．具体的には，クラスタシーケンスから特定の長さの遷移パターンを抽出し，その出現頻度からそれぞれのパターンにスコア付けを行う．

蛭田，山口，嶋田，高倉，八木，秋山の論文 [5] では，通信のクラスタシーケンスを用いてマルウェアの分類を行っている．通信プロトコルごとに異なる特徴量を用いてクラスタリングを適用し，クラスタシーケンスを得る．得られたクラスタシーケンスの類似度を用いてマルウェアを分類している．

また，一般のシンボル列に対する異常検知を広く紹介したものとしては Chandola，Banerjee，Kumar によるサーベイ論文 [3] がある．この論文では，シンボル列に対する異常検知を分類し，それぞれの分類での手法を紹介している．具体的には，シーケンス同士の間に類似度を定め，自身以外のシーケンスと類似度の低いものを異常と判定するものや，シンボル間の遷移にマルコフ性を仮定してシーケンスの生起確率を用いて異常を判定するものなどがある．

さらに，以下に HTTP プロキシサーバのログデータを利用した関連研究について述べる．HTTP プロキシログの標準的なログフォーマットに含まれる情報のうち，各項目の頻度に着目して自動的に RAT の通信を検知する手法が三村，大坪，田中 [12] によって提案されている．

Perdisci, Roberto ,Lee, Wenke ,Feamster, Nick の論文 [7] では，HTTP ベースのマルウェアを分類するために，リクエストの数，GET の数，POST の数，URL の平均の長さ，パラメータの平均数，POST で送信したデータの平均サイズ, 平均の応答のサイズを使用している．この手法ではさらにクエリの内容も分析してクラスタに分類し，マルウェアのサンプルによるクラスタと類似性を比較することでシグネチャを自動生成している．この手法では，プロキシのログから取得できる項目を用いており，その平均 に着目して挙動を抽出している．

田中，堀川，峰野，西垣 [13] はマルウェアの感染がないと想定される期間と，感染が疑われる期間のプロキシのログを比較することで，効率的にログを縮退する手法を提案した．この手法では，最終的には熟練ネットワーク管理者による判断が必要である．

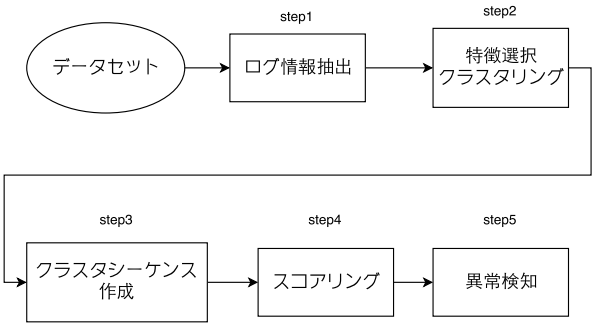

図 1: 提案手法

Ma，Saul，Savage，Voelker の論文 [6] では URL 自体から悪性サイトを検知する手法を提案している．外部情報無しに抽出できる URL に関する特徴量を複数挙げており，HTTP プロキシログから URL に関する特徴量を抽出する際の参考にした．

3 検知手法について

3.1 概要

本論文では，HTTP プロキシサーバのログデータに対してクラスタリングを行い，その中から異常なクラスタ遷移を検知する手法を検討する．HTTP プロキシサーバとしては Squid を用い，研究室内での普段の生活で利用する HTTP 通信のログを記録した．ログに必要な情報は

- リクエストに対する応答時間,
- リクエスト URL,
- クライアント IP アドレス,
- サーバ IP アドレス,
- HTTP ステータスコード,
- クライアントが受信したデータサイズ

である．本論文においては，これらの通信データを全て正常通信として扱っている．

図 1 に提案する攻撃検知手法の全体的な処理の流れを示す.

Step1: Squid ログデータの各行に対応する HTTP リクエストとレスポンスに対して特徴量の抽出・計算を行い，特徴量ベクトルを作成する．詳細を 3.2 節で述べる．

Step2: Step1 で得られた特徴量ベクトルを元にして，クラスタリングを行う．また，解析に有用と思われ

る特徴量集合を求める．その過程で全てのデータがクラスタに分割される．詳細を 3.3 節で述べる．

Step3: Step2 でクラスタリングされた Squid ログに対し，ホスト毎に行った通信を取得し，行った通信に対応するクラスタ ID からなるクラスタシーケンスを作成する．詳細を 3.4 節で述べる．

Step4: クラスタシーケンスを用いてログに対してスコアリングを行う．詳細を 3.5 節で述べる．

Step5: Step4 でのスコアに基づいて異常な通信を検知する．ここでは閾値を適切に定め，閾値を超えるものを異常と判断する．詳細を 3.5 節で述べる．

3.2 特徴量の抽出

HTTP プロキシのログデータから特徴量抽出を行う．抽出する特徴量は複数の先行研究で使用されていたものを参考に 12 種類を選んだ．表 1 にそれらを示す．

表 1: 特徴の種類

数値特徴量 (単位)
クライアントが受信したデータサイズ (byte)
リクエストに対する応答時間（秒）
リクエスト URL の全長（文字）
リクエスト URL のパスの長さ（文字）
リクエスト URL のクエリの長さ（文字）
リクエスト URL でクエリされた値の数（個）
カテゴリ特徴量（特徴次元数）
レスポンスステータスコードの最初の桁 (5)
リクエストのファイルの有無・形式 (22)
リクエスト URL 内の IP アドレスの有無 (2)
サーバ IP アドレスの 2 進数表示 (32)

数値特徴量は，各特徴量同士の尺度を合わせるためにそれぞれ平均 0，分散 1 のデータへ正規化を行った．

カテゴリ特徴量は取りうる値同士の距離を適切に定義するのが困難なためダミー変数化を行った．レスポンスステータスコードについては，その 5 種類のクラス分けに着目したため，特徴次元は 5 となる．リクエストのファイルの有無・形式については，ファイルの送信が行われたか否かと，行われた場合は HTTP 通信で送受信される一般的なファイル形式 20 種類を選定し，それらに含まれないその他の場合を考えたため，22 次元の特徴となった．従って，数値特徴量とカテゴリ特徴量のダミー変数化により，最終的に入力データとして与える特徴ベクトルの次元は 67 である．

本論文では上記の特徴に対し，適切な特徴の組み合わせを特徴選択アルゴリズムによって探索する．

3.3 クラスタリングと特徴選択

本論文における実験では，k-means，DBSCAN，EM アルゴリズムを用いて多変量混合ガウス分布を仮定した尤度最大化によるクラスタリング (以下，EM クラスタリング) の 3 種類のクラスタリングを行う．なお，DBSCAN ではクラスタリングの際に外れ値が判定されるが，本論文では外れ値をまとめてもう一つのクラスタとして扱う．

k-means と EM クラスタリングにおいては，3.2 節で挙げた特徴量のいくつかを予め選んで実験を行うともに，ラッパー法を用いてクラスタリングと特徴選択を同時にする実験も行う．ラッパー法では，まず，探索によって特徴部分集合を抽出する．次に，その部分集合を用いてクラスタリングを行う．最後に，結果を一定の基準で評価する．以上の操作を繰り返し，その中で最も良い評価が得られたときの特徴とクラスタリング結果を採用する．以下では，k-means と EM クラスタリングにおけるラッパー法の各手順について説明する．

3.3.1 特徴探索

k-means，EM クラスタリングともに，逐次順方向探索 (SFS) によって特徴探索を行う [4]．SFS では，選択する特徴部分集合を空集合から開始し，一度につき 1 つの特徴を順次追加する．追加される特徴は，すでに選択された特徴と組み合わせたときに最大の基準値を与える特徴である．これ以上特徴を追加しても基準値が改善されない場合，探索を停止させる．

3.3.2 クラスタリングと評価基準

通常の k-means や EM アルゴリズムではクラスタ数は予め設定しておく必要がある．データに合わせて最適なクラスタ数を探索するため，大きな数 $K_{\max}$ からクラスタ数を 1 つずつ減らして，そのクラスタリング結果を評価し，最もよいものを選択する．この際，EM クラスタリングにおいては計算量を小さくするため，Bouman[2] のアルゴリズムを用いて計算を行っている．

EM クラスタリングの評価はガウス分布に対する尤度が高いほどよいものとする．k-means におけるクラスタリングと特徴選択の評価，および EM クラスタリングにおける特徴選択の評価として，散布分離性 (SS) 基準を用いている [4].

SS 基準は，各クラスタに対してクラスタ内共分散行列 S_w とクラスタ間共分散行列 S_b を定義し，跡 $\mathrm{trace}(S_w^{-1}S_b)$ を計算する．$S_w^{-1}S_b$ は，クラスタ内共分散によって正規化されたクラスタ間共分散であり，この跡が大きいほど，クラスタ間の正規化された距離が大きくなり，クラスタの識別性が向上する．

SS 基準は，選択した特徴の数に評価値が依存してしまう問題があるため，そのバイアスを除去する手法も取

り入れている [4]．また，カテゴリ特徴量を含むデータを評価する際に正常に動作しないことがあるため，カテゴリ特徴量に対して微小な一様ノイズを加えている．

3.4 クラスタシーケンス作成

全セッションデータについてクラスタリングを行った後，同一ホストによるセッションに対応するクラスタ番号を時系列順に並べる．

3.5 スコアリング

ログ単位でのスコアリング手法は荒木，山口，嶋田，高倉 [9] の方法を用いた．クラスタシーケンスが得られた後にクラスタ間遷移を bi-gram によって抽出する．各遷移が全体のデータの中で出現する頻度と，あるクラスタが次のクラスタへ遷移する割合などを計算する．これによって各遷移のスコアを計算する．

シーケンス中の i 番目のログを Log_i とし，Log_i が属するクラスタを c_i とおく．クラスタ a からクラスタ b の遷移が全体の中で出現する割合を $p(a,b)$，クラスタ a からの遷移の中でクラスタ b へ移るものの割合を $q(b|a)$ とする．このとき i 番目の遷移は c_i から c_{i+1} への遷移であり，そのスコア $\mathrm{score}(i)$ は以下のように計算される．

$$\mathrm{score}(i) = p(c_i, c_{i+1})q(c_{i+1}|c_i)$$

ログ Log_i のスコア S_i は前後合わせて $2l$ 回の遷移を用いて次のように計算される．

$$S_i = \sum_{k=i-l+1}^{i+l} \mathrm{score}(k)$$

このスコアリング手法では，前後に出現割合の多い遷移が多く含まれるとスコアが増加することになる．すなわち，スコアが高いものは頻度の高い遷移が前後にあることを示しているため，正常と判定する．スコアの低いものはデータ内でのパターンが少ない遷移パターンの中に現れるものなので，攻撃のログとみなすことにする．今回の実験では，長さが l 未満のシーケンスに対しては，正常なシーケンスと判定することとする．

4 評価実験

4.1 データセット説明

実環境で観測したデータに対し，正常通信と悪性通信を正確にラベル付けすることは困難なため，本論文ではマルウェア対策のための研究用データセットの一種である D3M [1, 10] を用いて評価実験を行う．

D3M は独自に収集した悪性 URL を高対話型ハニークライアント Marionette で巡回した際の通信データ（ドライブバイダウンロード攻撃の通信など）に加え，取得したマルウェアをサンドボックスで動作させてその通信を解析したデータからなるデータセットである．通信データは pcap 形式で提供されており，宛先 URL や IP アドレスなどプロキシデータよりも多くの情報を抽出できる．

D3M の pcap データから作成した疑似プロキシデータは，ドライブバイダウンロード攻撃に関与する攻撃・異常通信として扱う．

4.2 実験データ

以下に，本実験で用いたデータの詳細を述べる．

- 正常通信データ：2017 年 10 月 2 日に Squid に蓄積されたログデータ（約 15MB，約 4 万行），
- 異常通信データ：2014 年 5 月 2 日に Marionette によって採取された pcap データから作成した疑似ログデータ（約 3800 行）．

上記のデータを時系列順を変えずに混合させたものを今回の実験の対象データとした．このデータから 3.2 節で述べた特徴量を抽出し，以下の 3 種類の組み合わせで各クラスタリングを行う．

(a) 全ての数値特徴量（6 次元），

(b) サーバ IP アドレスを除いた特徴量（35 次元），

(c) 全て（67 次元）．

(a) は [9] を参考にしてカテゴリデータを除いた場合として用いた．(b) は (a) との比較として，カテゴリデータを含めた場合の変化を観察することを目的とし (c) は [8] による，IP アドレスを 2 進数で表した場合の Web サイトの判別手法を参考にし，本論文においても 32 次元の特徴量としてサーバ IP アドレスを追加した．

特徴量の組み合わせと同様に，[9] で採用されていた，スコアリングの際に用いる前後のログの個数 $l = 5$ を採用した．

4.3 実験結果

検知率と実際の分類の関係を表 2 のように定めるとき，検知率と誤検知率は以下のように計算され，[9] と同様にして，本検知手法をこの検知率と誤検知率により比較評価する．

表 2: 検知結果の定義

		実際の分類	
		攻撃	正常
検知結果	攻撃	TP	FP
	正常	FN	TN

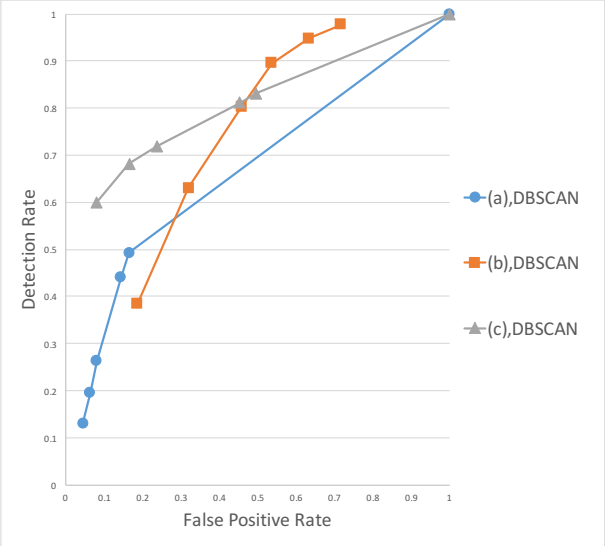

図 2: 特徴選択処理なし:DBSCAN

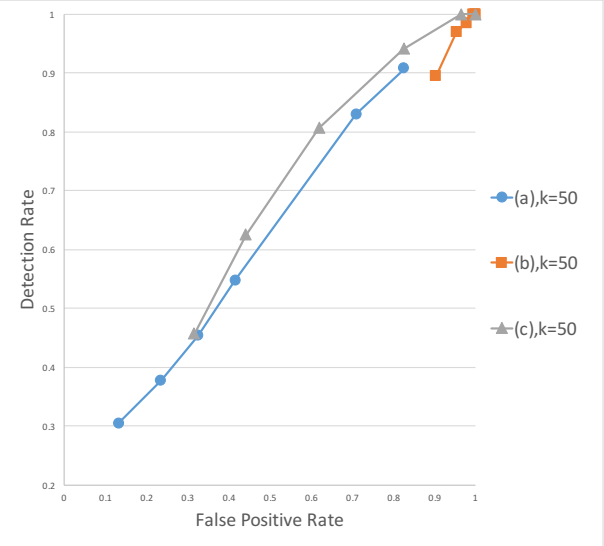

図 3: 特徴選択処理なし:k-means

$$\text{検知率 (DR)} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$\text{誤検知率 (FPR)} = \frac{\text{FP}}{\text{TN} + \text{FP}}$$

4.3.1 特徴選択処理なし

この節では特徴選択処理を行わない場合の実験結果について述べる．用いたクラスタリングアルゴリズムは k-means，DBSCAN の 2 つである．また，用いた特徴の組み合わせとして前節で紹介した 3 種類 (a), (b), (c) があり，それぞれを採用してクラスタリングを行なった結果を示す．

DBSCAN を適用する際に，用いる特徴によってデータのばらつきが異なるため近傍半径 eps などのパラメータを個別に設定している．それぞれの結果を図 2 に示す．それぞれのパラメータは (a): （eps = 0.001, クラスタ数：75）(b): （eps = 1.4, クラスタ数：103），(c): (eps = 3.2, クラスタ数：237）である．

k-means は実行する際にクラスタ数を指定する必要がある．今回の実験では，指定するクラスタ数を 50 と 60 とした．図 4， **??**に k-means を用いた攻撃の検知結果を示す．

4.3.2 特徴選択処理あり

この節では特徴選択処理を行った場合の実験結果について述べる．特徴選択処理を行う際に，(a), (b), (c) のそれぞれから選択される特徴部分集合は異なると考えられる．EM クラスタリングの評価関数として SS を用いた場合に選択された特徴の組をそれぞれ (d), (e), (f) とする．k-means の評価関数として SS を用いた特徴選択では (a), (b), (c) から同じ特徴部分集合が選択された．その特徴の組を (g) とする．図 5 では，EM クラスタリングの結果を，図 6 では同じ特徴量を用いた DBSCAN の結果を示す．それぞれの DBSCAN のパラメータは (d): (eps = 0.1, クラスタ数：3），(e): (eps = 0.2, クラスタ数：3），(f): (eps = 0.2, クラスタ数：3）である．

k-means を用いて特徴選択処理を行なった実験では，クラスタ数が 14 となった．DBSCAN でこれと同じ特徴を採用して行った異常検知の結果を図 7 に示す．

4.4 考察

4.3 では比較的良かった値のみ示したが、それでも DR と FPR がほぼおなじ値をとるものが多く，今回行った実験では有用な異常検知は行えていないことがわかる．原因としては

(i) 数値特徴量のみを入力とした異常検知手法の入力にカテゴリ特徴量を用いた．

(ii) 特徴量選択アルゴリズムも同様に数値特徴量を入力としたものだった．

(iii) プロキシデータから抽出した特徴量が不十分．

(iv) クラスタリング時のパラメータの吟味が不十分．

などがあげられる．(i), (ii) については数値特徴量のみに絞るか，カテゴリ特徴量を上手く加工する方法が必要である．(iii) についてはプロキシデータから得られる情報の再吟味が必要である．(iv) については，特徴量選択なしでの DBSCAN を用いた検知精度がクラスタ数が多い場合に比較的良性であることから，クラスタ数に注目してパラメータの調整を続けることで多少改善されると考えられる．

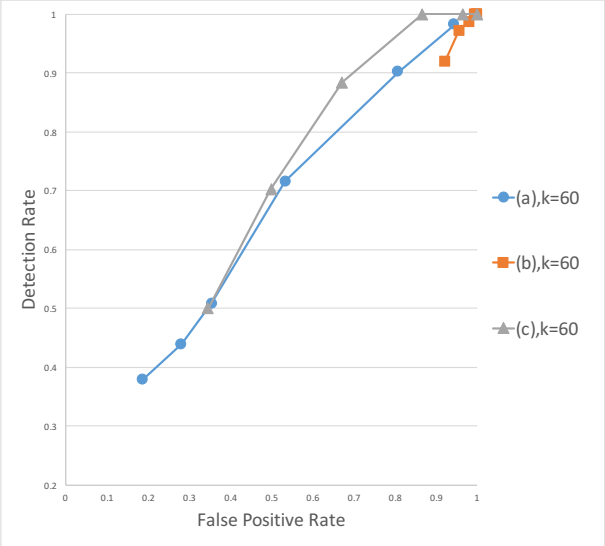

図 4: 特徴選択処理なし:k-means2

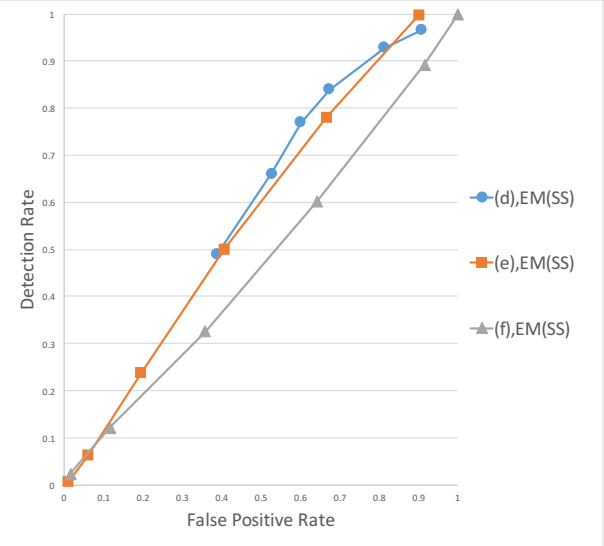

図 5: 特徴選択処理あり:EM(SS)

5 今後の課題

本論文ではプロキシデータから抽出した特徴量は全て，単一のセッションに当たるログから得られるものである．[9] でも参考にされている Kyoto2006+[1] が含む特徴量に，一定時間内での同一ホストの通信回数などがある．これを参考にして前後のセッションのログも用いることで特徴量を増やすことができると考えられる．また，今回の実験では数値特徴量が少ないことが検知精度が低いことの一因であると考えているので，数値特徴量の種類を増やし，評価実験を行う必要がある．

今回用いたクラスタリング手法の k-means と DBSCAN ではそれぞれ指定しなければいけないパラメータがあり，それらを変えながら実験を行い，良い結果が得られるパラメータを選んだ．この適切なパラメータの選択方法については調査が必要であると考えている．他にも DBSCAN で異常検知の精度が良くなるパラメータが見つかったときに，そこで得られたクラスタ数を k-means のパラメータに用いるなどの実験も行っていきたい．

異常検知の手法として，クラスタシーケンスを比較することによって不自然な遷移パターンを検知することが考えられるシーケンス検知の手法は様々あり，[3] によればどのような異常を検知したいかによって問題設定が大きく 3 つに分類される．各設定での手法を利用することによって，様々な状況に対する異常検知手法を提案することが可能と考えられる．また，本論文で行なった検知手法では検知するに至るまでに計算時間がかかるので，リアルタイムで検知することが可能な検知手法を考えることが今後の課題としてあげられる．

[1] http://www.takakura.com/kyoto_data/

参考文献

[1] Mitsuaki Akiyama, Makoto Iwamura, Yuhei Kawakoya, Kazufumi Aoki, and Mitsutaka Itoh. Design and implementation of high interaction client honeypot for drive-by-download attacks. *IEICE Trans. Commun.*, Vol. 93, No. 5, pp. 1131–1139, may 2010.

[2] Charles A. Bouman. *CLUSTER: An Unsupervised Algorithm for Modeling Gaussian Mixtures*, 1995.

[3] Varun Chandola, Arindam Banerjee, and Vipin Kumar. Anomaly detection for discrete sequences: A survey. *IEEE Trans. Knowl. Data Eng.*, Vol. 24, No. 5, pp. 823–839, 2012.

[4] Jennifer G. Dy and Carla E. Brodley. Feature selection for unsupervised learning. *Journal of Machine Learning Research*, Vol. 5, pp. 845–889, 2004.

[5] Shohei Hiruta, Yukiko Yamaguchi, Hajime Shimada, Hiroki Takakura, Takeshi Yagi, and Mitsuaki Akiyama. Evaluation on malware classification by session sequence of common protocols. In Sara Foresti and Giuseppe Persiano, editors, *Cryptology and Network Security: 15th International Conference, CANS 2016, Milan, Italy, November 14-16, 2016, Proceedings*, pp. 521–531, Cham, 2016. Springer International Publishing.

[6] Justin Ma, Lawrence K. Saul, Stefan Savage, and Geoffrey M. Voelker. Learning to detect malicious

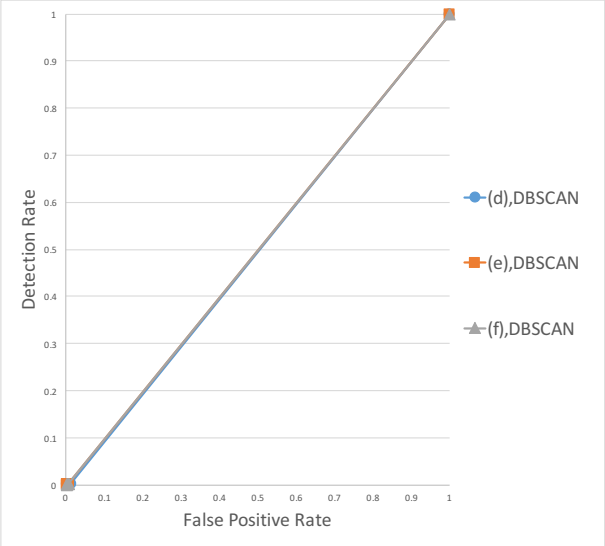

図 6: 特徴選択処理あり:DBSCAN

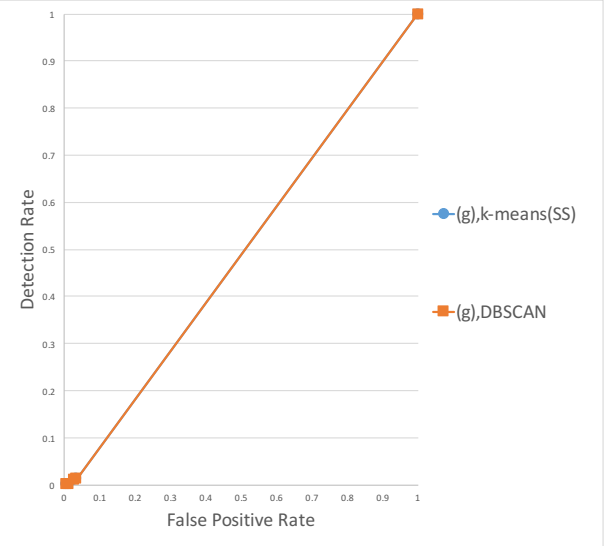

図 7: 特徴選択処理あり:k-means

URLs. *ACM TIST*, Vol. 2, No. 3, pp. 30:1–30:24, 2011.

[7] Roberto Perdisci, Wenke Lee, and Nick Feamster. Behavioral clustering of http-based malware and signature generation using malicious network traces. In *Proceedings of the 7th USENIX Conference on Networked Systems Design and Implementation*, NSDI'10, pp. 26–26, Berkeley, CA, USA, 2010. USENIX Association.

[8] 金澤しほり, 中村嘉隆, 稲村浩, 高橋修. 悪性 IP アドレスの分布特徴に基づく未知の web サイトの判別手法. コンピュータセキュリティシンポジウム 2017 論文集, 第 2017 巻, pp. 1076–1084, oct 2017.

[9] 荒木翔平, 山口由紀子, 嶋田創, 高倉弘喜. 通信のクラスタ間遷移に基づくサイバー攻撃検知手法. コンピュータセキュリティシンポジウム 2015 論文集, 第 2015 巻, pp. 1066–1072, oct 2015.

[10] 高田雄太, 寺田真敏, 村上純一, 笠間貴弘, 吉岡克成, 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 datasets 2016〜. 情報処理学会研究報告コンピュータセキュリティ（CSEC）, 第 2016-CSEC-74 巻, pp. 1–8, jul 2016.

[11] 佐藤正明, 山口由紀子, 嶋田創, 高倉弘喜. セッションデータのシーケンスに注目した異常な通信パターンの検出. 第 31 回 暗号と情報セキュリティシンポジウム (SCIS2014), 2014.

[12] 三村守, 大坪雄平, 田中英彦. HTTP ベースの通信挙動に基づく RAT 検知システムの試作. コンピュータセキュリティシンポジウム 2016 論文集, 第 2016 巻, pp. 488–495, oct 2016.

[13] 田中功一, 堀川博史, 峰野博史, 西垣正勝. ログ解析によるマルウェア侵入検知手法の提案. マルチメディア、分散協調とモバイルシンポジウム 2014 論文集, 第 2014 巻, pp. 522–529, jul 2014.