プロキシのログからの機械学習による RAT の検知方式

三村 守 †　　大坪 雄平 †‡　　田中 英彦 †

† 情報セキュリティ大学院大学
221-0835 神奈川県横浜市神奈川区鶴屋町 2-14-1
dgs104101@iisec.ac.jp

‡ 警察庁
100-8974 東京都千代田区霞が関 2-1-2

あらまし RAT(Remote Access Trojan または Remote Administration Tool) を用いた標的型攻撃による情報漏洩の被害は深刻である．標的型攻撃に用いられる RAT の通信を検知するために，これまでに多くの対策が提案されているが，その多くはパケット単位でのネットワーク監視を必要としている．しかしながら，実際に標的型攻撃を受けた組織において，パケット単位での記録を長期間保存している可能性は低い．このような場合，プロキシのログ等の限られた情報から，RAT の痕跡を検知する必要がでてくる．本稿では，プロキシのログに記録される受信量や時間等の挙動から RAT の特徴点を抽出し，機械学習により RAT の痕跡を検知する方式を提案する．さらに，実際の RAT の痕跡を含むプロキシのログを分析し，提案方式の有効性を示す．

Detecting RAT Activity in Proxy Server Logs with Machine Learning

Mamoru Mimura†　　Yuhei Otsubo†‡　　Hidehiko Tanaka†

†Institute of Information Security
2-14-1 Tsuruya-cho, Kanagawa-ku, Yokohama-city, Kanagawa 221-0835, JAPAN
dgs104101@iisec.ac.jp

‡National Police Agency
2-1-2 Kasumigaseki, Chiyoda-ku, Tokyo 100-8974, JAPAN

Abstract APT attacks that use RATs (Remote Access Trojan or Remote Administration Tool) and steal data inflict serious damages to the organization. Though there are some countermeasures to detect RAT activity in APT attacks, many previous methods require capturing packets. However, when an organization is actually attacked, there is little possibility that the organization have been storing captured packets for a long term. In such cases, we might have to detect RAT activity in limited information such as proxy server logs. In this paper, we make feature vectors from behavior such as time and sizes in proxy server logs, and propose how to detect RAT activity with machine learning. Finally, we analyze proxy server logs including actual RAT activity, and show the performance of our methods.

1　はじめに

特定の組織を狙った標的型攻撃による情報漏洩の被害は深刻である．2015 年には，多くの組

織でマルウェアへの感染や情報漏洩の可能性が公表され，大きな社会問題となっている．これらの標的型攻撃では，端末を遠隔操作するためのRAT(Remote Access Trojan またはRemote Administration Tool) が使用されている．標的型攻撃に用いられるRATは，難読化されて送り込まれることが多く，ウイルス対策ソフト等の従来の対策技術では検知することは困難である．RATを用いた標的型攻撃に対しては，これまでに動的解析を実施するサンドボックスタイプの製品や，各種のログを集約してその相関関係から感染を検知する手法等，様々な対策が提案されている．しかしながら，すべての組織で十分な対策を実施することは現実的に困難であり，実際に攻撃は対策が不十分な組織で発生している．このような組織では，攻撃者の痕跡を調査するための十分なログが記録されていないことも少なくない．したがって，プロキシのログ等の限られた情報から，RATの痕跡を検知する必要がでてくる．しかしながら，近年の標的型攻撃に用いられる主要なRATは，独自のプロトコルを用いずに，一般的なHTTPで通信を実施し，自身の通信を正規の通信に紛れこませることを意図した動作が指摘されている [1]．したがって，膨大な容量のプロキシのログからRATの痕跡を検知するのは，コマンド＆コントロール（以下Ｃ＆Ｃ）サーバのアドレスが不明な場合や，マルウェアの通信に固有の文字列が含まれない場合には困難である．

そこで本稿では，RATのＣ＆Ｃサーバのアドレスが未知であり，通信に固有の文字列が含まれない場合にも，同じタイプのRATであればその挙動が似ている可能性があるという点に着目する．既知のタイプのRATの挙動を機械学習により習得させることができれば，新たなＣ＆Ｃサーバを用いたり，通信に含まれる文字列を変化させた亜種も検知することが期待できる．本稿では，Ｃ＆Ｃサーバのアドレスが未知であり，マルウェアの通信に固有の文字列が含まれない場合にも，プロキシのログからHTTPベースのRATの通信を検知することを目標する．

以下，第２節では関連研究について説明し，本研究との違いを明確にする．第３節では提案方式とその実装について説明し，第４節では実際のRATの痕跡を含むプロキシのログを用いて実験を実施する．第５節では実験の結果を踏まえて提案方式の実用性を評価し，最後にまとめと今後の課題について示す．

2　関連研究

HTTPベースのRATの通信の検知に関連する研究としては，ネットワークの監視によってマルウェアのＣ＆Ｃサーバとの通信を検知するための研究と，プロキシのログから不正な接続先を検知するための研究が挙げられる．以下，提案方式とこれらの研究との違いについて説明する．

2.1　ネットワーク監視による手法

文献 [2] では，パケットサイズ，パケット数，到着間隔等の特徴量を用い，Ada Boostで通常の通信と不正な通信を区別することで，マルウェアへの感染を検知する手法を提案している．文献 [3] では，パケット数，データサイズ，セッション時間，アクセス回数およびアクセス時間の標準偏差を特徴ベクトルとして，Support Vector MachineによりＣ＆Ｃトラフィックを抽出する手法を提案している．文献 [4] では，セッション毎に合計パケット数，初期段階のセッションの存続時間，データサイズ，パケット数およびパケットの平均データサイズを特徴ベクトルとし，決定木とRandom Forestsにより RATによる通信か否かを判定する手法を提案している．これらの手法では，パケット単位でのトラフィックの監視が必要である．

文献 [5] では，DNSクエリの挙動を分析し，不正な未知のドメインを検知する手法を提案している．この手法では，ISP規模でのDNSクエリの監視が必要である．

提案方式では，RATによる通信か否かの判定に，Support Vector Machine(以下SVM) 及びRandom Forests(以下RF) を用いている点が従来の研究と共通している．しかしながら，ネッ

トワーク監視を必要とせず，プロキシのログのみを対象としている点が異なっている．

2.2 プロキシのログを使用する手法

文献 [6] 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ーに分類し，マルウェアのサンプルによるクラスターと類似性を比較し，シグネチャを自動生成している．この手法は，プロキシのログから取得できる情報を用いている点は共通しているが，NCSA 共通ログフォーマット[1]に含まれないパラメータを利用している点と，教師なし学習を用いている点が異なっている．したがって，プロキシでこれらのパラメータが記録されていない場合，この手法を適用することはできない．提案方式では，NCSA 共通ログフォーマットに含まれるパラメータのみを利用し，教師あり学習モデルである SVM と RF を使用している．さらに，リクエストの数やサイズのほかにも間隔や path の特徴を考慮している．

文献 [7] では，プロキシのログからクライアントのアドレス，訪問先のアドレスおよびリクエストの数を用い，クライアントと共通するサーバに着目してグループに分類し，疑わしいドメインの検出を支援する手法を提案している．この手法では，疑わしいドメインの検出をブラックリスト等の他の手法に依存している．

文献 [8] では，DNS のログ，プロキシのログ等を使用し，内部ホストの訪問履歴とその User Agent から，組織全体の希少な訪問サイト，User Agent の傾向，ドメインの類似性等を分析し，通常状態と比較することで異常なドメインを検出する手法を提案している．この手法では，プロキシ以外にも DNS のログを必要としている．また，疑わしいドメインを検出するために，ドメインの登録情報，ブラックリスト等の外部からの情報を必要としている．提案方式は，プロキシのログのみを使用し，外部からの情報を必要としない．

文献 [9] では，マルウェアの感染がないと想定する期間のプロキシのログと，マルウェアの感染を疑う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ている．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．

提案方式は，教師あり学習モデルである SVM と RF を用いて自動的に不正な通信を識別し，人手による判定を必要としない．

3 提案方式

3.1 前提条件

本稿で提案する検知方式を実現するための前提条件は，プロキシのログに以下の情報が記録されていることである．

- 時刻
- クライアントの識別子（クライアントの IP アドレスかユーザ名）
- リクエストの内容（メソッド，URL および User Agent を含む.）
- HTTP ステータスコード
- クライアントが受信したサイズ

これらの情報は，すべて NCSA 共通ログフォーマットに含まれている項目であり，ほとんどのプロキシで取得可能である．提案方式では，HTTP ステータスコードが成功である通信を対象とし，クライアントの識別子と URL に含まれるホストのユニークな組み合わせを抽出する．次に，指定したログの行数毎に，時刻，リクエストの内容およびクライアントが受信したサイズから特徴ベクトルを作成する．

3.2 特徴ベクトル

提案方式において，指定する行数のログから作成する特徴ベクトルを以下に示す．

[1] Web サーバがログを記録する際の標準化されたテキストファイルの書式

① 最頻出の受信サイズ

② 最頻出の受信サイズの数

③ 最頻出のリクエストの間隔

④ 最頻出のリクエストの間隔の数

⑤ 最頻出の path の長さ

⑥ 最頻出の path の長さの数

⑦ POST メソッドの数

⑧ User Agent の長さ

特徴ベクトルは，固有の文字列を用いずに，挙動の頻度に着目して作成した．①から④は，受信サイズおよび間隔（直前のログの時刻との差分）のヒストグラムを図 1 に示すように作成し，最も頻度が高い受信サイズおよび間隔とその数とした．⑤および⑥は，図 2 に示すように RAT は同じ path に連続してアクセスするという特徴を数値化したものである．⑦は，一部の RAT は特定のメソッドを多用することに着目している．⑧は，User Agent の違いを考慮させるために選定した．

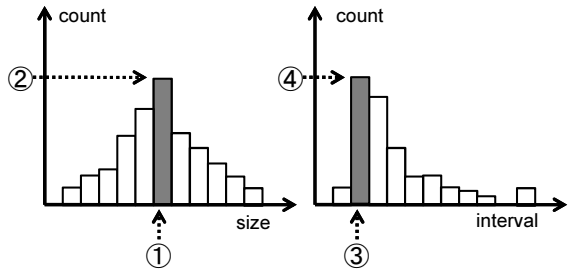

図 1: 受信サイズと間隔のヒストグラム

3.3 学習と判定

提案方式では，教師あり学習モデルである SVM と RF を用いる．そのため，教師データとして検知対象とする RAT の痕跡を含む既知のログが必要であり，かつそのログにおいて RAT の通信と通常の通信の区別がついている必要がある．提案方式の動作は，学習フェーズと判定フェーズに分類される．

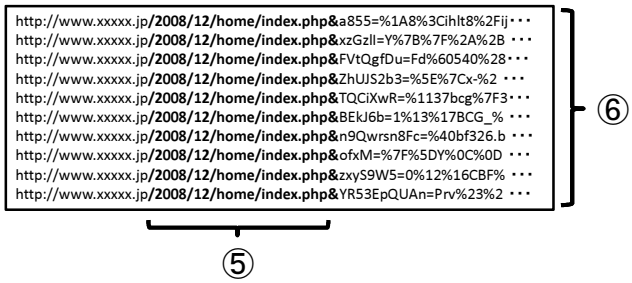

図 2: RAT のアクセスログの例

学習フェーズ

検知対象とする RAT の痕跡を含むログを読み込み，指定するログの行数毎に特徴ベクトルを作成する．次に，特徴ベクトルが RAT による通信であれば RAT の種類，それ以外であれば通常の通信のラベルを付与し，SVM または RF に学習させる．

判定フェーズ

対象とする未知のログを読み込み，学習フェーズと同様に指定するログの行数毎に特徴ベクトルを作成する．次に，その特徴ベクトルを SVM または RF に予測させ，RAT の種類か通常の通信のラベルを出力させる．

提案方式では以上の動作により，対象とするログから HTTP ベースの RAT の種類を検知する．

3.4 実装

提案方式を，Python-2.7 と機械学習のライブラリを活用して実装した．SVM については libsvm-3.2[10] を用いた．カーネル関数については，分類するデータに関する事前知識がないことから，汎用的な用途で用いられる RBF（ラジアル基底関数）カーネルを選択した．RF については scikit-learn-0.16.1[11] を用いた．SVM，RF ともに，その他のパラメータについては，デフォルトの値となっている．

4 実験

4.1 実験内容

実装したプログラムとプロキシのログを用いて実験を実施する．実験環境および実験に使用するプロキシのログの概要を表 1 および表 2 に示す．このログは，2015 年に標的型攻撃を受けた組織のプロキシのログであり，NCSA 共通ログフォーマットで記録されている．このログには，２タイプの RAT による遠隔操作の痕跡が含まれていることを確認している．２タイプの RAT は，2010 年以降に出現した比較的新しい RAT であり，標的型攻撃に使用されることが多い HTTP ベースの RAT である．実験では，学習フェーズで教師データを読み込み，判定フェーズで分析データを読み込ませる．なお，今回の実験で使用する教師データは，分析データに含まれている．

表 1: 実験環境

CPU	Core i5-3450 3.1GHz
Memory	DDR3 SDRAM 8GB
HDD	Serial ATA 600
OS	Windows 7

表 2: 実験データ

	教師データ	分析データ
期間	１日	約１か月
容量	約 250MB	約 40GB
RAT の種類	２タイプ	２タイプ

4.2 実験結果

まず，ログの行数（以下ｎ）毎の特徴ベクトルの数を表 3 に示す．特徴ベクトルの総数は，教師データ，分析データともにｎを減らすたびに減少している．ｎと特徴ベクトルの総数が反比例とならないのは，指定したｎに満たないログは特徴ベクトルに反映されないためである．教師データに含まれる RAT の特徴ベクトルの数は，分析データの 10%未満となっている．

次に，実験の検知率および誤検知数を表 4 に，所要時間を表 5 示す．SVM はｎを 10 以下にすると顕著に学習時間が長くなり，検知率ではRF にやや劣るものの，誤検知数は少ない結果となった．また，ｎを 50 以上にした場合には顕著に検知率が低下した．これに対し，RF ではｎを減らしてもあまり学習時間は長くならず，全般的に高い検知率を示すが，誤検知も少し発生する結果となった．また，どちらの場合にも，ｎを５未満に減らした場合には，見逃しや誤検知数が増加した．

表 3: 特徴ベクトルの数

ログの	教師データ		分析データ	
行数ｎ	RAT	総数	RAT	総数
100	36	233	371	26116
50	78	872	818	92967
30	133	2181	1444	227798
20	211	4312	2248	443754
10	440	12251	4737	1245824
5	897	31115	9710	3135909

表 4: 検知率

ログの	検知率 (DR)		誤検知数 (FPC)	
行数ｎ	SVM	RF	SVM	RF
100	27.0%	98.7%	0	1
50	62.2%	100.0%	0	7
30	86.1%	99.4%	0	2
20	95.2%	99.3%	0	3
10	97.8%	99.5%	1	14
5	98.6%	99.2%	25	5

4.3 MWS データセットへの適用

今回の実験で使用した教師データを用い，MWS データセット [12] で追加実験を実施した．実験

表 5: 所要時間

ログの	学習時間		判定時間	
行数ｎ	SVM	RF	SVM	RF
100	31s	31s	1h18m	1h18m
50	31s	31s	1h19m	1h19m
30	32s	31s	1h20m	1h22m
20	33s	32s	1h21m	1h25m
10	42s	33s	1h26m	1h33m
5	1m08s	34s	1h40m	1h53m

に使用したデータは，BOS 2015 に含まれるすべての pcap ファイルであり，この中には RAT の痕跡が含まれている．この pcap ファイルをプロキシのログに相当する擬似ログに変換するために，まず HTTP プロトコルを抽出し，リクエストとレスポンスの対応付けを実施した．さらに，そのリクエストとレスポンスのペアから以下の情報を抽出し，擬似ログを作成した．

- 時刻
- IP アドレス
- リクエストの内容
- HTTP ステータスコード
- レスポンスのサイズ

同様の手法により，NCD in MWSCup 2014 から RAT の痕跡を含まない通常の通信の擬似ログを作成した．作成した擬似ログの概要を表 6 に示す．学習フェーズでは今回の実験に使用した教師データを読み込み，判定フェーズでは BOS 2015 および NCD in MWSCup 2014 から作成した擬似ログを読み込ませた．機械学習は RF を選択し，ｎは 30 とした．その結果，BOS 2015 から４つの不正な接続先をすべて検知し，NCD in MWSCup 2014 から誤検知は発生しなかった．

表 6: 擬似ログの概要

	BOS 2015	NCD
期間	12 日	１日
容量	約１ MB	約８ MB
RAT の種類	１タイプ	―

5 評価

5.1 検知率

SVM についてはｎが 20 以下において，検知率は 95%以上に達した．誤検知については，ｎが 10 以上においてはほとんど発生しなかった．したがって，SVM の場合の最適なｎは，10〜20 程度であると言える．

RF についてはｎが 50 以下において，検知率は 99%以上となった．誤検知については，ｎの値にかかわらず概ね 10 件程度であった．したがって，RF の場合の最適なｎは，５〜50 程度であると言える．

誤検知や見逃しの原因は，HTTP ベースの RAT に特徴ベクトルが類似しているサイトであった．たとえば，動画のストリーミング再生や，何らかの API を提供するサイトに関しては，同一の受信サイズ，あるいは同一に近い受信サイズの通信が定期的に繰り返される傾向が認められた．また，検知率には反映されていないが，設定したｎに満たない場合にはそもそも特徴ベクトルが作成できないため，見逃しの可能性がある点にも注意する必要がある．

5.2 所要時間

SVM については，ｎを 10 以下にすると，顕著に学習時間が長くなる傾向が認められた．この原因は，特徴ベクトルが増加したためであると考えられる．判定時間については，特徴ベクトルの数に応じて緩やかに長くなる傾向が認められた．

これに対し，RF についてはｎを減らしても，学習時間はあまり長くならず，概ね一定となる傾向が認められた．判定時間については，SVM

と同様に特徴ベクトルの数に応じて緩やかに長くなる傾向が認められた．

この結果から，教師データが少なく，特徴ベクトルが少ない場合にはSVMの方が高速であることが確認できた．しかしながら，SVMでは特徴ベクトルが多くなると顕著に学習時間が長くなることから，教師データが多い場合にはRFの方が高速になるものと考えられる．これは，学習が高速なRFの一般的な特性によるものであると考えられる．

5.3 実用性

提案方式は，標的型攻撃を受けた組織のプロキシのログを詳細に分析する用途と，ネットワークをリアルタイムで監視する用途を想定している．どちらの用途においても，提案方式ではNCSA共通ログフォーマットに含まれている項目のみを用いるため，様々な組織の機器や様々な状況に適応可能であると考えられる．これに対し，既存の研究ではネットワーク監視，NCSA共通ログフォーマットに含まれない項目，外部からの情報の取得等を前提としているため，現実的には適応できる状況は限られていると言える．

詳細に分析する用途に着目すると，提案方式は特に標的型攻撃の対策が不十分であり，攻撃の痕跡がプロキシのログのみが残されている状況で特に有用であると考えられる．このような詳細に分析する用途では，検知率に優れるRFを用いることで，ほとんど見逃すことなくHTTPベースのRATを検知することが可能である．本稿における実験データは，教師データが分析データに含まれている．これは，デジタルフォレンジック等で復元した一部のマルウェアを分析し，その一部のC＆CサーバのURLが判明した場合，他にまだ発見されていない未知のC＆CサーバのURLを検知する状況に合致している．

提案方式は，SVM，RFともに約1か月分で約40GBのログを2時間以内に処理しており，リアルタイムで監視する用途での運用も可能であると考えられる．リアルタイムで運用する場合には，発生する誤検知の数をオペレータが処理できる数に抑える必要がある．10〜20程度のnを設定してSVMを選択すれば，誤検知をなるべく抑えつつ，95%以上の検知率を実現することが可能である．さらに検知率を高めるためには，RFを選択してnを5〜50程度に設定すれば，99%の検知率を実現することができる．RFの場合にはやや誤検知が発生するが，その数がオペレータが処理できる量であれば運用に支障はないものと考えられる．

5.4 制約

提案方式が機能するためには，前提条件で示したとおり，NCSA共通ログフォーマットに含まれている項目がプロキシのログに記録されている必要がある．従来の機械学習によりマルウェアの挙動を検知する手法のほとんどは，パケット単位でのネットワーク監視を前提条件としている．しかしながら，標的型攻撃の被害を受けた組織において，パケット単位での記録を長期間保存している可能性は低いものと考えられる．したがって，提案方式の前提条件はより実用的であると言える．

提案方式では，教師データとして検知対象とするRATの痕跡を含む既知のログが必要であり，かつそのログにおいてRATの通信と通常の通信の区別がついている必要がある．通常は，RATの痕跡を含むログは標的型攻撃を受けた組織が保有しているため，入手は容易ではない．検体等が入手できればハニーポットを運用することで作成することは可能である．しかしながら，教師データの作成が難しい点は，提案方式の制約である．

運用においては，設定したnが多い場合には見逃しの可能性があることに注意する必要がある．この制約に関しては，RFを選択してnを少なめに設定し，ホワイトリスト等を併用して誤検知を除外することで，ある程度は緩和することが可能である．

6 おわりに

本稿では，プロキシのログから特徴ベクトルを作成し，機械学習によりHTTPベースのRAT

を検知する方式を提案した．また，RAT による標的型攻撃を受けた組織のプロキシのログを分析し，提案方式の有効性を示した．さらに，実験結果を考察し，提案方式の実用性を評価した．

今後の課題としては，他のタイプの RAT や，他の標的型攻撃の受けた組織のプロキシのログへの適用が挙げられる．本稿では主要な 2 つのタイプの RAT に対する有効性を示したが，他のタイプの RAT に対する効果は明確ではない．他の新たなタイプの RAT が出現した場合には，特徴ベクトルを再検討する必要がでてくる可能性がある．また，リアルタイム検知システムへの応用も今後の課題である．

参考文献

[1] 標的型サイバー攻撃分析レポート 2015 年版 ～「気付けない攻撃」の高度化が進む～ (online) http://www.go-tm.jp/apt2015/ (2015-07-24).

[2] 市野 将嗣，市田 達也，畑田 充弘，小松 尚久：トラヒックの時系列データを考慮した AdaBoost に基づくマルウェア感染検知手法，情報処理学会論文誌，Vol.53, No.9, pp.2062–2074 (2012).

[3] 山内 一将，川本 淳平，堀 良彰，櫻井 幸一：機械学習を用いたセッション分類によるＣ＆Ｃトラフィック抽出，2014 年暗号と情報セキュリティシンポジウム (2014).

[4] 蒋 丹，面 和成：初期段階における Remote Access Trojan の検知手法，コンピュータセキュリティシンポジウム 2014 (2014).

[5] Babak Rahbarinia, Roberto Perdisci, Manos Antonakakis：Segugio: Efficient Behavior-Based Tracking of New Malware-Control Domains in Large ISP Networks, *Proc. 2015 IEEE/IFIP International Conference on Dependable Systems and Networks* (2015).

[6] Roberto Perdisci, Wenke Lee, Nick Feamster：Behavioral Clustering of HTTP-based Malware and Signature Generation using Malicious Network Traces, *Proc. 2010 USENIX Symposium on Networked Systems Design and Implementation* (2010).

[7] Manh Cong Tran and Yasuhiro Nakamura：A Supplementary Method for Malicious Detection, Journal of Communications, vol.9, no.12, pp.923-929 (2014).

[8] Alina Oprea, Zhou Li, Ting-Fang Yen, Sang Chin and Sumayah A. Alrwais：Detection of Early-Stage Enterprise Infection by Mining Large-Scale Log Data, CoRR, vol.abs/1411.5005 (2014).

[9] 田中 功一，堀川 博史，蜂野 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 (2014).

[10] libsvm (online) https://www.csie.ntu.edu.tw/~cjlin/libsvm/ (2015-07-24).

[11] scikit-learn (online) http://scikit-learn.org/ (2015-07-24).

[12] 神薗 雅紀，秋山 満昭，笠間 貴弘，村上 純一，畑田 充弘，寺田 真敏：マルウェア対策のための研究用データセット ～MWS Datasets 2015～，第 70 回コンピュータセキュリティ研究発表会 (2015).