動的解析に基づく評価用マルウェアの選定方式に関する検討

渡辺　喬之†　　畑田　充弘†

†NTT コミュニケーションズ株式会社
108-8118　東京都港区芝浦 3-4-1　グランパークタワー16F
{tkyk.watanabe, m.hatada} @ntt.com

あらまし サイバー攻撃対策において重要なマルウェア対策技術が数多く提案されている．マルウェア対策技術の性能評価には，その対策技術の目的に応じて多様な検体が必要となり，様々な検体の収集・分類手法が提案されている．本稿では，アンチウイルスソフトでは検知できない未知マルウェアを対象とした動的解析の結果から，その挙動に基づく評価用マルウェアの選定方式を提案し，4900個の未知マルウェアを対象とした実験結果を示す．実験結果から，複数クラスタに分類された未知マルウェアから評価用マルウェアを選定可能であることを示す．また，D3M2010~2014のマルウェア検体のクラスタへの分類結果を考察する．

A Study on Malware Selection Methodology for Evaluation based on Dynamic Analysis

Takayuki Watanabe†　　Mitsuhiro Hatada†

†NTT Communications Corporation.
Gran Park Tower 16F, 3-4-1, Shibaura, Minato-ku, Tokyo 108-8118, JAPAN
{tkyk.watanabe, m.hatada} @ntt.com

Abstract Variety of anti-malware technologies have been proposed as one of effective measures against cyber-attacks. In terms of evaluation of anti-malware technologies, various kinds of malware samples and clustering methodologies of malware collection are necessary for corresponding to defensive objectives. This paper proposes a malware selection methodology for evaluation based on malicious activities of malware which are undetected by antivirus software. We applies the methodology to 4900 undetected malware and its results show malware samples for evaluation are determined from clustered and undetected malware. We also report that D3M2010~2014 malware samples are classified into appropriate clusters.

1　はじめに

アンチウィルスや侵入防御システムを代表とするセキュリティ技術では検知することのできない未知のマルウェアが急増している．Check Point の調査では 2012 年から 2013 年にかけて発見された未知のマルウェアは 8300 万種類，その増加率は 144%にも上ると報告されている[1].

また，同社は未知のマルウェアを検知できたアンチウィルスエンジンは全体の 10%未満と報告しており，アンチウィルスによる未知のマルウェアへの対策の限界を示している．

一方で，このような未知のマルウェアに対抗するための様々なマルウェア対策技術も存在する．ファイアウォールによる IP アドレス，ポート番号を利用したアクセス制御や，シグネチャ型の侵入検知/防御システムはマルウェアの感染防止策として用いられる．アプリケーションホワイトリストは正規のアプリケーションに関する情報を予めホワイトリストに登録しておくことで，想定外のソフトウェアの実行を制御する．また，エンドポイントセキュリティはクライアント端末において，マルウェアの検知だけでなく，不正な通信を検知するホスト型侵入検知/防御，感染端末のフォレンジックなど複合的な機能を提供する．

このように，マルウェア対策技術は数多く存在している．しかしながら，それらの技術の性能評価のために，無数に存在する検体の全てを用いて検証を行うことは極めて困難である．また，検体数が多ければ良いわけではない．

そのため，マルウェア対策技術の性能評価を効率良く行うための評価用マルウェア(以下，評価用マルウェア)の選定方法の確立は重要な課題である．企業等において，マルウェア対策技術はインターネットゲートウェイやエンドポイントに導入が進んでおり，標的型攻撃等の新たなマルウェア対策技術を組み込む場合，アンチウイルスソフトでは検知できない未知のマルウェアに対して有効であることは，新規技術導入の有用性を判断するための一つの指標となる．

そこで本稿では，アンチウイルスソフトでは検知できない未知のマルウェアを対象とした動的解析の結果を分類し，その挙動に基づいた評価用マルウェアの選定方式を提案する．また，提案手法の実現性を検証するために，4900 の未知マルウェアを対象とした実験結果示す．実験結果より，複数クラスタに分類された未知マルウェアから評価用マルウェアの選定が可能であることを確認した．また，MWS Datasets 2014[2]の D3M2010~2014 に該当するマルウェア検体のクラスタへの分類結果を報告し，クラスタによるマルウェアの挙動の特徴を考察する．

以降，本論文の構成は以下の通りである．2 章では関連研究とその課題について触れる．3 章では，マルウェア選定手法の説明を行う．また，4 章では，提案手法を用いた実験結果を示し，5 章で提案手法の有用性と課題について考察する．最後に，6 章で本稿のまとめを行う．

2 関連研究

プログラムコードの類似性に焦点をおきマルウェアを分類する方法[3,4]は，機械的に未知のマルウェアの識別を可能にしている．[3]はマルウェアのアンパッキング，逆アセンブル結果を機械語命令単位で比較することで，類似度の算出及びマルウェアの分類を行っている．[4]はマルウェアを逆アセンブルした結果得られたプログラムを関数単位で切り出し，比較することでマルウェアの機能の類似度，マルウェア同士の類似度を算出している．

マルウェアの挙動に焦点をおいた分類方法[5,6]も提案されている．[5]は長期間潜伏，常駐するマルウェアが OS の自動実行リストへ自身を登録する傾向に着目し，ビヘイビアブロッキング法に基づいたマルウェアの分類を行っている．また，[6]はサンドボックス解析で収集した大量の API コールログを用いて，マルウェアの特徴ベクトルを生成し，マルウェアのクラスタリングを行う手法を提案している．

マルウェアの挙動とアンチウイルスソフトで判定したマルウェアの名称を組み合わせて既存の

マルウェアの亜種の種類を推定する手法もある[7]. この手法は，マルウェアを解析用ネットワークで実際に実行し，その挙動をデータベースに蓄積する. また，マルウェアをアンチウイルスソフトでスキャンし，既知のマルウェアにはその名称を，未知のマルウェアには“unknown”の情報をデータベースに蓄積する. データベースに蓄積されたマルウェア間の挙動の類似性をハミング距離から算出し，マルウェアの分類を行う. この時，未知のマルウェアは最も類似した既知のマルウェアと同じ種類のマルウェアの亜種と判定する.

ここで挙げたマルウェアの分類や推定に関わる研究は，我々の目的である評価用マルウェアの選定にも有用ではあるが，分類されたものがどのような挙動のクラスタであるのかを把握するのが難しく，アンチウイルスソフトで検知できないマルウェアを対象としたい点においても適切とはえいないものがある.

3　評価用マルウェアの選定手法

図 1 に提案手法の流れを記載する. 提案手法では，まず，2 種類のハニーポットから構成される検体収集システムによりマルウェアと思われる検体を収集する. 次に，収集した検体をアンチウイルスソフトでスキャンし，既知のマルウェアをスクリーニングする. このとき，マルウェアと判定されなかった検体を評価用マルウェア候補とする. また，評価用マルウェア候補はサンドボックスにより動的解析する. 解析結果より，評価用マルウェア候補をクラスタリングし，クラスタの中から評価用マルウェアの選定を行う. 以降ではそれぞれのステップの詳細を述べる.

3.1　検体の収集

評価用のマルウェアを選定するためには，アンチウイルスソフトで検知されない未知のマルウェアを収集する必要がある. 本研究ではワーム等リモートからの攻撃を対象とする待ち受け型のハニーポットおよび Web サイトを巡回してドライブ・バイ・ダウンロード攻撃を検知するWebクライアント型のハニーポットで検体の収集を行っている.

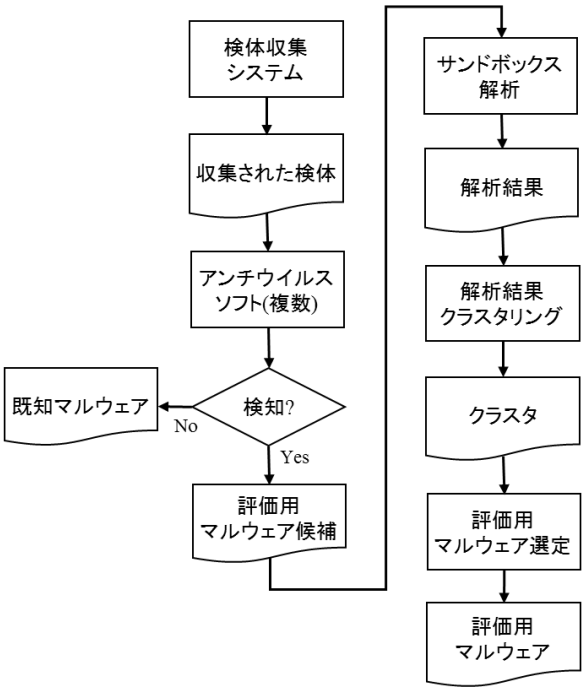

図1. 提案手法の流れ

3.2　アンチウイルスソフトによる評価用マルウェア候補の判定

3.2 で収集した検体を複数のアンチウイルスソフトでスキャンする. この時，全てのアンチウイルスソフトに検知されなかった検体を評価用マルウェア候補としてサンドボックスによる動的解析の対象とする. ただし，評価用マルウェア候補は必ずしもマルウェアであるとは限らない.

3.3　候補検体のサンドボックス解析

評価用マルウェア候補をサンドボックスで解析する. 解析結果は JSON 形式や XML 形式等の

レポートとして出力されることが多い．このレポートには，未知マルウェアの悪性度，解析対象ファイルや解析環境のメタデータ，マルウェアの挙動(アクセスしたレジストリ，ファイル，ネットワークアクティビティ)等，様々な情報が含まれる．

3.4　解析結果のクラスタリング

サンドボックス解析で出力されたレポートに基づきマルウェアのクラスタリングを行う．レポートに含まれる全ての情報を活用するのは困難なため，本研究ではマルウェアの悪性度(0-100 の整数で，値が大きいほどマルウェアである可能性が高い)とマルウェアの不正な挙動(例：Windows のプロダクト ID を読む，FTP サーバへのログイン情報を読む、等)を用いた．以下，マルウェアのクラスタリング手順を説明する．

手順 1：悪性度のラベル化

悪性度を整数から 0，1-9，10-19,…,90-99,100 とラベルに変換する．この悪性度のラベル化によりクラスタリングをした際に，悪性度が近いマルウェア同士を同一のクラスタに分類している．

手順 2：マルウェアのクラスタリング

悪性度ラベルと不正な挙動の 2 つが完全に一致するマルウェアを同一クラスタとして分類する．図 2 に本研究で用いたクラスタリングのイメージを示す．

3.5　評価用マルウェアの選定

悪性度ラベルと不正な挙動を基準として評価用のマルウェアの選定目的に応じてクラスタから試行的に検体を選定する．以下に選定方法の例を示す．

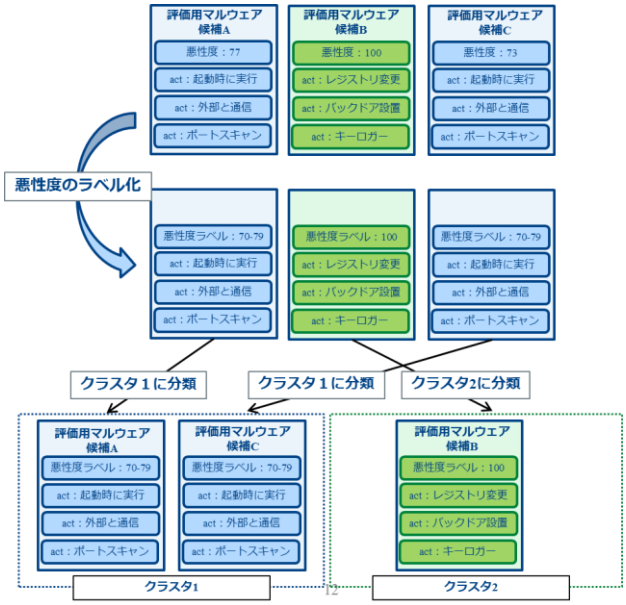

図 2.サンドボックス解析結果を利用したマルウェアのクラスタリング

① 未知の検体に対する検知性能の検証例

よりマルウェアらしいもので異なる挙動のものを選定するために、まず，悪性度ラベルが 70 以上のクラスタを全て列挙する．また，各クラスタに属する評価用マルウェア候補の中から 1 検体ずつ評価用マルウェアとして選定する．このときクラスタに複数評価用マルウェア候補が存在する場合は，以下を順に確認し該当する評価用マルウェア候補が一つに絞られたとき，評価用マルウェアとする．

確認事項 1：検証環境で実行できるファイル形式
確認事項 2：VirusTotal[10]で未登録
確認事項 3：マルウェアの Sha1 ハッシュ値(16 進数表記)が同一クラスタで一番小さい（あくまで 1 検体を選定するための条件）

② Windows のプロダクト ID を読むマルウェアに対して高い精度で検知する技術の検証例

Windows のプロダクト ID を読むという挙動を持つものでマルウェアらしさは様々なものを選ぶ．

ここで，悪性度ラベルが低いクラスタから評価用マルウェア候補を選定するのは，誤検知回避の性能検証も行うためである．具体的には，まず，Windows のプロダクト ID を読む挙動を持つクラスタを全て選出する．次に，悪性度ラベル低(0-39)・中(40-69)・高(70-100) の中から検証できる適切な数の検体を評価用マルウェアとして選定する．

上述のように検証の目的に応じて評価用マルウェアのクラスタからの選定方法を変える．本研究では，アンチウイルスソフトの検知できない評価用マルウェアを選定することが目的であるため，先の例で示した①の選定方法にて評価用マルウェアを選定する．図 3 に選定方法のイメージを示す．

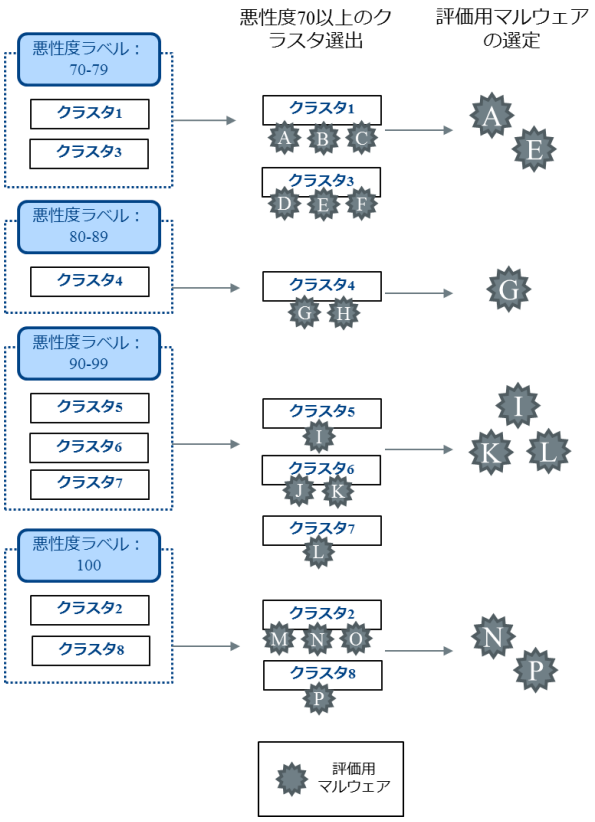

図 3.クラスタからの評価用マルウェアの選定

4 実験結果

4.1 実験結果

提案手法の実現性を検証するために，2014 年 7 月 1 日~7 日の一週間で収集した sha1 ハッシュ値でユニークな 4900 検体(700 検体/1 日)を用いた実験結果を示す．今回，4900 検体中 3 検体はサンドボックスで解析できなかったためクラスタリング非対象としている．

表 1 はサンドボックスの解析結果から明らかになった各悪性度における検体とクラスタ数の分布である．結果，4897 検体が 154 のクラスタに分類された．このクラスタを分析すると悪性度ラベルは異なるが不正な挙動が一致するクラスタが 15 組存在しており，多くは悪性度が中程度(40-69)の検体であった．

これら 4897 検体のユニークな不正な挙動総数は 120 通りであり，悪性度が低いラベル(1-39)を持つクラスタの不正な挙動には，http クライアントの起動，アプリケーションディレクトリ内での exe ファイルの変更などが含まれている．また，悪性度が中程度のラベル（40-69）を持つクラスタの不正な挙動は，ブラウザのクッキーの読み込み，アプリケーションの強制終了，SMTP クライアントの起動などが含まれる．悪性度の高いラベル（70-100）を持つクラスタの不正な挙動には，不正に JavaScript を Web ページに仕込む，解析対策のために一時的にスリープ状態になる，Windows が起動した時に自動起動する設定を行うなどが確認された．

マルウェアの選定プロセスにおいては，よりマルウェアらしいものを評価用マルウェアに選定するために悪性度が 70 以上のクラスタを選定対象にしたところ，評価用マルウェアを 24 検体選定することができた．これは，4897 の検体の割合の 0.49%にあたる．一方，4897 検体の分布は，悪

精度 50 未満が 3889 検体で，79.4%を占めている．収集してきた検体の悪性度が小さく，http クライアントの起動やアプリケーションディレクトリ内でのファイル変更，アドウェアの実行など一般的なアプリケーションが実施する不正な挙動を有することを鑑みると，マルウェアでないものが含まれている可能性は高い．従って，ハニーポットで収集した検体から評価用のマルウェアを効率良く選定するにあたり，提案手法は有用であると考えられる．

表 1. 各悪性度における検体とクラスタ数の分布

Malicious score ranges	# of Malware Samples	# of Clusters
0	1812	1
1-9	34	8
10-19	2	2
20-29	6	6
30-39	2035	24
40-49	719	54
50-59	91	14
60-69	148	21
70-79	35	12
80-89	5	5
90-99	3	3
100	7	4
	Total:4897	Total:154

4.2 D3M 検体のクラスタリング

提案手法を用いて，D3M 2010～2014 の 66 検体から評価用マルウェアのクラスタリングを行う．また，クラスタリングの結果から検体収集システムで収集した 4900 の検体のクラスタとの比較を行う．

表 2,3 にハニーポットで収集した検体，D3M 検体，両検体全体のクラスタリング結果を示す．表 2,3 より，D3M 検体は 17 のクラスタに分類されたことがわかる．この内，3 クラスタは 4.1 で算出したクラスタに分類され，14 クラスタ(複数手段によるパスワード奪取をするマルウェア，長期の潜伏を行うワーム，不正ファイルの生成とセキュリティ機構の無効化を行うマルウェアなどのクラスタ)は D3M 検体に固有であった．また，D3M 検体の不正な挙動は全部で 34 種類あったが，内 19 種類が表 4 に示される D3M 検体のみに見られた不正な挙動(パスワードブルートフォース攻撃，ファイル拡張子の偽装，Windows ファイアウォールの無効化など)である．この D3M データセットのみに見られた不正な挙動を有するクラスタは 10 通り存在しており，いずれも高い悪性度をもつことが確認された．この結果から D3M 検体は悪性度が高い既知の不正な挙動を有しており，アンチウィルスで検知されるマルウェア検体であると推測できる．

表 2. D3M 検体の各悪性度におけるクラスタの数の分布

Malicious score ranges	# of Malware Samples	# of Clusters
0	0	0
1-9	0	0
10-19	0	0
20-29	0	0
30-39	7	2
40-49	2	2
50-59	0	0
60-69	0	0
70-79	1	1
80-89	0	0
90-99	0	0
100	56	12
	Total:66	Total:17

表 3. データセットごとの検体数及びクラスタリングの結果

	Evaluation malwares	D3M	Evaluation malwares + D3M
# of Samples	4897	66	4963
# of Clusters	154	17	168

表4. D3M検体のみに見られた不正な挙動一覧

Activity
ウイルスコードの検知
システムファイルの作成
メモリの書き換え
パスワード奪取プログラムの実施
Windows の重要なプロセスが走るメモリの書き換え
ルートディレクトリに自動起動に関連するファイルを作成
他のファイル拡張子を装ったファイルの作成
インターネットのセキュリティレベルの変更
隠れファイルに関する表示設定の変更
タスクモニタなどの管理者ツールを実行不能にする
Java のインストールを利用したなりすましファイルの作成
Security Center notifications 機能の無効化
Windows ファイアウォールの無効化
ユーザアカウント管理に関する通知の無効化
FTP クライアントのパスワードの奪取
メールサーバのパスワードの奪取
パスワードブルートフォース攻撃の実施
ブラウザに記憶されたパスワードの読み込み
データディレクトリ内の怪しい場所にある実行ファイルの変更

5 考察

提案手法により 4897 の検体を 154 のクラスタに分類することができた. また, 154 のクラスタから評価用マルウェア 24 検体を選定することができた. 検体の多くが悪性度の低い, 非マルウェアである可能性が高いことを踏まえると, 提案手法により効率良く評価用マルウェアの選定ができていると考えられる.

しかしながら, 検体の選定プロセスには改善の余地がある. 悪性度ラベルは異なるが, 不正な挙動が完全に一致する検体が存在している点を考えると, サンドボックスのレポートから他の情報の抽出も考えなくてはならない.

また, 今回, 悪性度の高い検体が少なかったため, 最終的に選定した評価用マルウェアが適度な数になった. 一方で, 収集した検体のマルウェアが占める割合が高くなる場合, 提案手法では評価用マルウェアを絞りきることができない. より汎用性の高い選定方式を確立するために, マルウェアが有する不正な挙動の類似度や連続性を考慮したクラスタリングに関しては今後の課題としたい.

6 まとめ

本稿では, アンチウイルスソフトでは検知できない未知のマルウェアを対象とした動的解析の結果をクラスタリングし, その挙動に基づいた評価用マルウェアの選定方式を提案した. また, 4900 の検体を用いて提案手法の妥当性を検証した. 実験結果より, 悪性度と不正な挙動を基準にすることで, 複数クラスタから評価用マルウェアを効率良く選定できることが確認された. 一方, D3M2010~2014 検体をクラスタリングしたところ, 多くはアンチウイルスソフトで検知されるマルウェアであることが示された.

参考文献

[1] Check Point (2014), 2014 CHECK POINT ANNUAL SECURITY REPORT, pp.11 - 20.

[2] 秋山満昭, 神薗雅紀, 松木隆宏, 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 2014～, 情報処理学会 研究報告コンピュータセキュリティ（CSEC） Vol. 2014-CSEC-66, No. 19, pp. 1 - 7, 2014.

[3] 岩村誠, 伊藤光恭, 村岡洋一: 機械語命令列の類似性に基づく自動マルウェア分類システム, 情報処理学会論文誌, Vol.51, No.9, pp.1622 – 1632 (2010).

[4] 東結香, 中津留勇, 猪俣敦夫, 砂原秀樹, 藤川和利: マルウェアのコードの類似度を用いた分類手法に関する一考察, コンピュータセキュリティシンポジウム 2011 論文集, pp.107 - 112 (2011).

[5] 名坂康平, 酒井崇裕, 山本匠, 竹森敬祐, 西

垣正勝：自動実行登録に基づくマルウェアの分類に関する検討と評価，コンピュータセキュリティシンポジウム 2010 論文集，pp.459 - 464(2010).
[6] 藤野朗稚，森達哉：自動化されたマルウェア動的解析システムで収集した大量 API コールログの分析，コンピュータセキュリティシンポジウム 2013 論文集，pp.618 - 625(2013).
[7] 堀合啓一，今泉隆文，田中英彦：マルウェア亜種の動的挙動を利用した自動分類手法の提案と実装，情報処理学会論文誌，Vol50，No.4，pp.1321 - 1333(2009).
[8] 堀合啓一，今泉隆文，田中英彦：定点観測によるボットネットの観測と Malware の動的挙動解析システムの提案，情報処理学会論文誌，Vol49，No.4， pp.1680 - 1691(2008).
[9] 村上純一，鵜飼裕司：類似度に基づいた評価データの選別によるマルウェア検知精度の向上，コンピュータセキュリティシンポジウム 2013 論文集，pp.870 - 876(2013).
[10] “VirusTotal” https://www.virustotal.com.