L-002

マルウェア可視化システムの実装について
An Implementation of Malware Visualization System

浦辻 和也† 松重 雄大‡ 甲斐 博† 森井 昌克‡
Kazuya Uratsuji Takahiro Matsushige Hiroshi Kai Masakatu Morii

1．まえがき

インターネットの急速な普及により多種多様なマルウェアが膨大な数で蔓延し，その被害は年々増加している．マルウェアとはウイルス，ワーム，トロイの木馬等，悪意を持って作成されたソフトウェアやコードの総称である．

日々新しいマルウェアが検出されており，アンチウィルスソフトを開発している企業などにより，その情報が公開されている．例えば，Symantec 社は，その Web サイトにおいて Security Response の Threat [1] を公開し，日々検出されるマルウェアの情報を更新している．また，ESET は Web サイト VirusRadar (BETA)[2]上でマルウェアの感染状況や症状を提供している．しかし，その内容にはマルウェアの機能が文字情報として記述されており，マルウェア間の比較をするときに一目ではその違いが分かりにくいという問題がある．

そこで，我々は，マルウェアの理解や比較を助けるため，マルウェアの可視化手法を検討し，解析結果の整理がしやすいシステムの構築をしている．例えば，[3]では，Security Response をマルウェアの解析結果として用いて，マルウェアの特徴抽出を行い，マルウェアの可視化手法を提案している．また，[4]では，ESET の VirusRadar を用いて，[3]と同様な手法でマルウェアの特徴抽出を行い，マルウェアの可視化を行っている．

本論では[3]で述べた手法を基礎として，Web 上から任意のマルウェアを選択し，マルウェアの特徴の把握や比較を行うためのシステムを提案する．

2．マルウェアの機能分類

マルウェアは，コンピュータへの進入方法や動作の特徴を解析結果から判断され，ワームやトロイの木馬等大まかな種類分けに分類される．しかし，種類分け以降の詳細な分類については明確な定義が存在しない．

詳細な分類を行うことの利点は，詳細な動作内容を把握しやすいこと，亜種間の違いが確認しやすいことが挙げられる．このためマルウェアの詳細な分類を行う多くの研究が行われている [5][6].

詳細な種類分けに関する一つの指標はベンダによる名前付けがあげられる．しかしマルウェアはセキュリティベンダごとが設けた命名規則によって名前が付けられる．例えば，Symantec 社で W32/Bugbear@mm と呼ばれるワーム型のマルウェアは，Kaspersky 社では I-Worm.Tanatos.b などと呼ばれ，その名称はセキュリティベンダによって異なる．

静的解析や動的解析を行い，その結果を用いて分類する手法も提案されているが，注目する点が異なれば分類結果も異なるという課題がある．

一般に，マルウェアは，感染，破壊活動，ネットワークへのアクセスなど特定の動作を行うものが多い．感染とは，OS，ソフトウェア，ディレクトリ構造などを調査し，データを書き換え，追加するといった行動を取る．破壊活動では，環境の調査を行い，データの書き換えを行う．ネットワークへのアクセスでも，環境の調査を行う．このようにマルウェアでは様々な情報を取得する機能や，ファイル，レジストリ，DLL に対してアクセスする機能が重要になる．

そこで，あらかじめ機能群とそれに属する機能を定義することを行う．マルウェアの解析結果からマルウェアの機能を取得し，各機能を機能群にまとめる機能分類を行えば，特定の機能の組合せによる詳細な分類が可能である．

マルウェアの機能群とは次の４つである．各機能群には複数の機能が含まれる．

[情報収集] ファイル，ソフトウェアのバージョン，パスーワードを収集するための機能をこの機能群に分類する．具体的には，「特定ファイルの調査」「ユーザ情報の収集」「特定箇所の文字列の収集」を機能として定義する．

[感染行動] マルウェアの流通，個人情報の取得，ボットネットの作成や維持を行う機能をこの機能群に分類する．具体的には，「ファイル，レジストリ，DLL の作成」「ファイル，レジストリ，DLL の書き換え」「ファイル，レジストリ，DLL の削除」を機能として定義する．

[破壊活動] データやシステムの破壊，外部からの進入経路を構築するバックドア作成等を行う機能をこの機能群に分類する．具体的には，「バックドアの作成」「ファイルシステムの破壊」を機能として定義する．

[外部への動作] 特定サイトへのアクセスやネットワーク上のほかのコンピュータへのアクセス，メールの配信，特定サーバへのログイン等を行う機能をこの機能群に分類する．具体的には，「サイトへのアクセス」「メールの送信」「サーバへのログイン」「ダウンロード」「ネットワーク共有」を機能として定義する．

3．マルウェアの機能抽出とマルウェアの可視化

本研究では，マルウェアの解析結果として，Symantec 社が公開している Security Response を用いる．Security Response の Threat にはマルウェアの最新情報が更新された日付順に掲載されており，マルウェアの名前をクリックするとその内容を見ることができる．マルウェアの内容には Summary や Technical Details などがあり，Technical Details にはどのような動作をするかが英語で自然言語的に記述されている．

マルウェアの種類，動作内容（機能），機能の対象となるものがいくつ存在するかなどがマルウェアの特性になる．その解析のため Technical Details を最初から順に読み込み，そのマルウェアにはどのような機能があるのかを抽出する．

† 愛媛大学，Ehime University
‡ 神戸大学，Kobe University

マルウェアの種類は Technical Details 上の Type に示されている文字列により分類される．Type には数種類の文字列があり，Worm, Trojan, Virus などがある．3D モデルの形として Worm は球を用い，Trojan は直方体を用いる．

Technical Details の標準的な記述形式は，まずは英語でマルウェアの感染方法やマルウェアの挙動などの機能が記述され，その記述に続いて，ファイルや DLL などの機能の対象（以下，機能対象と呼ぶ）が列挙される．

本研究では，表 1 で示した英単語を使って，機能を説明する英文が表 1 の中のどの機能と対応するかを判断する．一つの動作について何個の動作対象があるか，その数を記録し，3D モデルの特徴に用いる．

表 1 機能群の分類とモデル

機能群	機能	英単語	機能モデルの色	機能モデルの形
情報収集	特定ファイルの調査	search	赤	楕円体
	ユーザ情報の収集	collect	緑	
	特定箇所の文字列収集	password	青	
感染行動	ファイル，レジストリ，DLL の作成	create, copy,	赤	円錐
	ファイル，レジストリ，DLL の書き換え	modify, add	緑	
	ファイル，レジストリ，DLL の削除	delete	青	
破壊活動	バックドアの作成	back, door	赤	お椀型
	ファイルシステムの破壊	end	緑	
外部への動作	サイトへのアクセス	connect	赤	つの型
	メールの送信	mail	緑	
	サーバへのログイン	server	青	
	ダウンロード	download	黄色	
	ネットワーク共有	network	シアン	

例えば，W32.Ackpra.A はワームの一種であり，利用可能なすべてのネットワーク共有やリムーバブルディスクによって自身をコピーすることによって拡散し，有害なファイルをダウンロードする．

W32.Ackpra.A の Technical Details の内容に従い，マルウェアを表現する 3D モデルとして球を用い，球面上に機能モデルを機能対象の数だけ配置する．W32.Ackpra.A の場合，感染行動（ファイル，レジストリ，DLL の作成）が 134，外部への動作（ダウンロード）が 13，情報収集（特定個所の文字列収集）が 29，情報収集（特定ファイルの調査）が 13，などのデータが得られる．

その結果得られるマルウェアの可視化モデルを図 1 に示す．

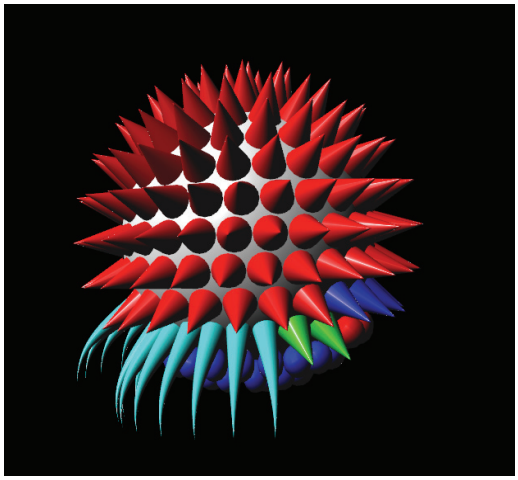

図 1：W32.Ackpra.A の可視化

本節で述べた手法は[3]で詳細に述べている．[3]では単一の Technical Details の URL を入力として可視化するシステムを提案した．

本研究では，マルウェアの分類に対してマルウェアの可視化手法の応用を検討する．

マルウェアの可視化の応用としては，

1. マルウェアの特徴や発生率を元に時間軸上で整理する
2. 特徴の把握・比較を可能にする

が考えられる．これらの目的のためには，複数のマルウェアの解析結果を処理できるシステムが必要となる．1 については ESET のデータ [2] を用いたシステム開発の必要性について [4] で議論している．本論文では，もう一方の，マルウェアの特徴を把握し比較する機能を持つシステムの提案を行う．

4．提案システムの概要

本研究で提案するマルウェア可視化システムは，可視化プログラム，マルウェアデータベース，マルウェア取得プログラムにより構成される．

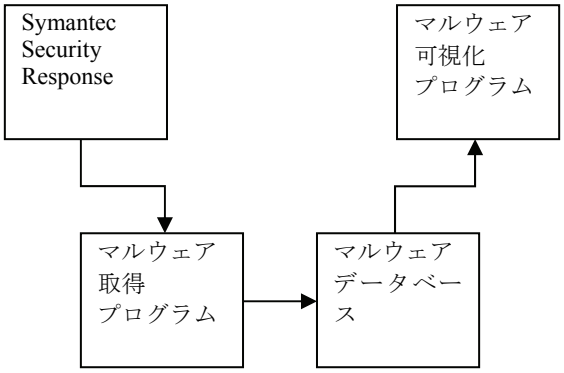

図 2：マルウェア可視化システム

図 2 に示したマルウェア取得プログラムは，シマンテックのホームページである Security Response の Threat[1]に掲載される最新のマルウェアに関する情報を取得し，マルウ

ェアデータベースの更新を行う．取得した情報は，マルウェアの機能特性を[3]で述べた手法を用いて解析する．解析結果はマルウェアデータベースへ保存する．

マルウェア可視化システムの利用者は，マルウェアデータベースから任意のマルウェアを複数選択し，マルウェア可視化プログラムで可視化することを行う．

具体的には，最新のマルウェアの取得は以下の手順で処理する．

1.Security Response の Threat の URL [1] を入力し，更新されたマルウェアがないか探す．Threat には更新日時順でマルウェアが表示されるので，前回取得済みのものは除き，新しいマルウェアのみの URL を取得する．
2.新しいマルウェアを記述した各 URL について，解析プログラムを実行し，可視化用のデータ（機能特性）を抽出する．
3.マルウェアごとに機能特性，URL などのデータをマルウェアデータベースに登録する．

上記の手順の 2 で述べた解析プログラムは，可視化用のデータとして文字列を返す．その文字列は，

- マルウェアの種類
- 機能名，機能数のリスト

を順に連結したものとして表現され，文字列はデータベースに解析結果として保存される．

5. 提案システムの詳細

データベースは MySQL を用いている．マルウェアデータベースは 1 つのテーブル malware を持ち，テーブルのエントリとして，表 2 に示す，マルウェアの id，マルウェア名（name），抽出した機能特性（extracted_data），Security Response の Threat に記載のマルウェアの URL（report）の 4 つを持つ．ただし，この report は URL 全体を記憶するのではなく，クエリに関係する部分のみを記憶する．

表 2：テーブル malware の仕様

カラム名	データ型	説明
id	int unsigned	データ登録時に自動インクリメント
name	varchar(128)	マルウェア名
extracted_data	Text	機能特性
report	varchar(32)	Security Response の Threat に記載のマルウェアの URL

図 2 に示した「マルウェア取得プログラム」は，現在，シェルスクリプト（update_malware.sh）により実装している．さらにそのシェルスクリプトの中で 3 つの perl スクリプト

- 新規マルウェアの取得: list_newmalware.pl
- 機能特性を取得: analysis_malware.pl
- マルウェアデータベースに登録: insert_int_db.pl

を利用し，システムを実現している．

図 2 に示したマルウェア可視化システムは，現在開発中のため非公開であるが，Web ブラウザを使ってアクセスできるように実装している．

Web サーバには CentOS6.5 上の apache を利用しており，コンピュータは CPU:Core2Duo, Memory:2GB のノートパソコンを利用している．

マルウェア可視化システムは，可視化可能なマルウェアの名前を列挙する「ホーム画面」と，ホーム画面から選択したマルウェアを可視化するための「可視化画面」から構成される．

マルウェア可視化システムの Web サイトの画面遷移図を図 3 に示す．

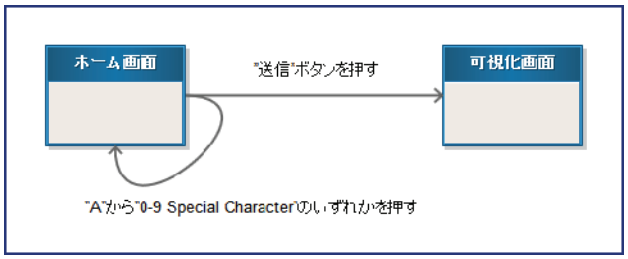

図 3：画面遷移図

5.1 ホーム画面

ホーム画面は本 web サイトにアクセスした場合，初めに表示される画面であり，図 4 に示す．

図 4：ホーム画面

またホーム画面を構成する要素を以下に示す．

1. “A”から”Z”，”0-9 Special Character”までのタブ
2. 送信ボタン
3. リセットボタン
4. マルウェア名を列挙するウィンドウ

要素 1 のいずれかのタブをクリックすることで，その文字を頭文字とするマルウェアを昇順に並べて要素 4 に表示する．

デフォルトでは A から始まるマルウェアが列挙して表示される．“0-9 Special Character”のタブはマルウェア名の頭文字が”A”から”Z”以外のものを列挙する．

要素 4 のマルウェア名の左側には選択のためのチェックボックスが設置されている．比較したいマルウェアをチェックボックスを利用して選択する．

要素 2 の送信ボタンをクリックすることで可視化画面へと移動する．

5.2 可視化画面

ホーム画面で送信ボタンを押された場合，可視化画面へと移動する，可視化画面ではホーム画面で選択されたマルウェアの可視化モデルを，選択順に左から並べ，中央揃えで描画する．図 5 はホーム画面で 2 つのマルウェアを選択した際の可視化画面である，また可視化モデルはマウスの

ドラッグによって回転させ様々な角度でマルウェアの特徴と確認することができる．同時に可視化可能な検体の数に制限はないが，現在の実装ではディスプレイの大きさで制限される．

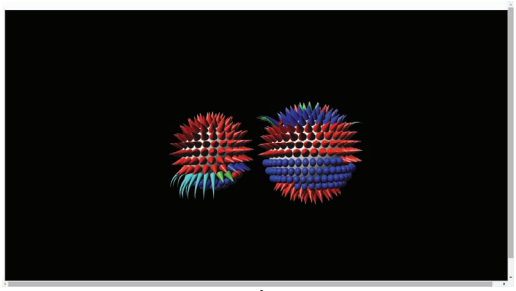

図 5：可視化画面

6. 可視化例

本節では，3 つのワーム型マルウェアと 2 つのトロイの木馬型マルウェアを可視化し，可視化モデルにより得られる効果について説明する．

図 6 は以下のワーム型のマルウェアを可視化した結果である．

- W32.Beagle.CX@mm (左)
- W32.Beagle.CY@mm (中央)
- W32.Beagle.DA@mm (右)

これらのマルウェアは，SMTP エンジンを使用して他の脅威であるトロイの木馬のコピーを送信するタイプの大量メール送信ワームである．

マルウェアに感染した際に送信されるトロイの木馬の種類が異なるため，異なるマルウェアとして分類されるが，基本的には同じ特徴を持つマルウェアの亜種になる．そのため，図 6 に示すように，似た形状の可視化モデルが得られる．

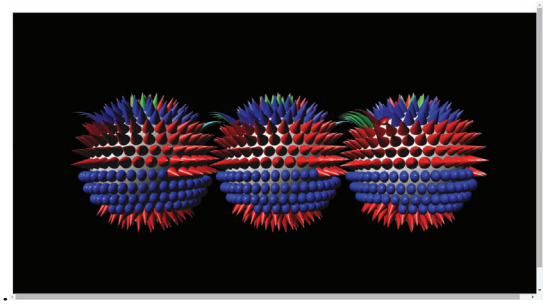

図 6：ワーム型マルウェアの可視化結果

図 7 は以下のトロイの木馬型のマルウェアを可視化した結果である．

- Trojan.Ransomcrypt.L （左）
- Trojan.Ransomcrypt.M （右）

これらのマルウェアは，侵入先のコンピュータのファイルを暗号化した後，ファイルを復号するために代金を支払うようユーザーに要求する．

基本的な動作が同じであるため亜種であると分類されているが，作成するファイルの種類や数が異なり，暗号化の対象となるファイルの拡張子の種類が異なる．亜種であるが特徴が大きく異なることが可視化モデルにより直感的にわかる例になっている．

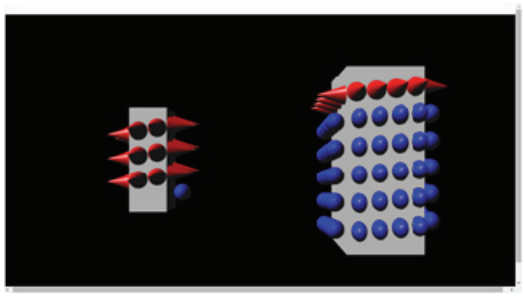

図 7：トロイの木馬型マルウェアの可視化結果

いずれの例も亜種を比較したものであるが，マルウェアの特性を視覚的に理解することに成功している．

7. まとめ

本研究では，マルウェアの解析結果に対して機能分類を行い，3D モデルで可視化する手法を用いて，マルウェアの機能を比較するシステムの開発を行った．われわれが提案した可視化モデルを用いることで，マルウェアの機能の類似性を直感的に把握することが可能になった．

6 節で可視化例を与えているが，図 6 がアンチマルウェアベンダーの分類でも名前から亜種と判定でき，可視化した結果も亜種であろうことが納得できる．

しかし，図 7 ではアンチマルウェアベンダーでの分類で亜種とされていながら，機能の対象が大きく異なるという結果になっている．これは本手法で提案する可視化の大きな優位性になる．もちろん解析レポートを注意深く吟味すれば理解できるが，簡単に判別できるという点で大きな利点になる．

しかし，そのような利点を持つ一方で，可視化モデルでは機能の詳細が隠れてしまい，各マルウェアの詳細について検討を進めるというような専門的な用途には利用できない．今後，ユーザが選んだ機能について機能の詳細を表示するよう変更し，類似性を直感的に確認できる特徴を持ちつつ，差異を深く理解できるよう改善する必要がある．

参考文献

[1] Symantec Security Report, Threats, http://www.symantec.com/security_response/landing/threats.jsp
[2] ESET VirusRadar(BETA), http://www.virusradar.com/en/home/world
[3] 浦辻和也，松重雄大，甲斐博，森井昌克，"Malware visualization based on the behavior and its classification," 第 13 回情報科学技術フォーラム(FIT2014), 2014.
[4] 松重雄大，浦辻和也，甲斐博，森井昌克，”マルウェアの可視化とその応用に関する研究”, コンピュータセキュリティシンポジウム 2014 論文集, 2014(2), pp.1142-1147, 2014.
[5] 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 50(4)，pp.1321-1333，2009.
[6] 藤原将志, 今田真敏, 安倍哲哉, 菊池浩明，“マルウェアの感染動作に基づく分類に関する検討,” 情報処理学会研究報告, CSEC, コンピュータセキュリティ, 2008(21), pp.177-182, 2008.