

標的型攻撃に対する知的ネットワークフォレンジックシステム LIFTの開発 ー標的型攻撃マルウェアの解析と亜種の予測ー

渋谷健太¹ 久山真宏¹ 佐藤信¹ 三村聡志¹ 松本隆¹ 佐々木良一¹

概要: 近年, 特定の組織や個人を対象とする標的型攻撃が社会問題となっている。標的型攻撃では, 攻撃者が事前に対象となる組織や個人の情報を収集し, その情報を元にカスタマイズしたマルウェアが使用されており, シグネチャベースの対策だけでは攻撃を防ぐ事が困難となっている。そこで著者らは LIFT (Live and Intelligent Network Forensic Technologies) システムおよび LIFT システムを拡張した Super-LIFT システムの研究開発を行っている。本研究では, Super-LIFT システムで必要となる攻撃データの分析・ルール化の為, 標的型攻撃に使用されるマルウェアの内, 日本年金機構にも使用された「EMDIVI」と先行するマルウェアの亜種を収集・解析し, 解析結果を機能ごとに比較することで, 今後発生する亜種の機能の予測を行う。本研究の結果から各マルウェアの機能がどのように変化するかを明らかにすることができた。また, 多次元尺度法を用いた分析の結果マルウェア間で亜種の発生パターンは異なることが明らかになった。

Development of intellectual network forensic system LIFT against targeted attacks - Analysis and prediction of malware subspecies for targeted attacks -

KENTA SHIBUYA¹ MASAHIRO KUYAMA¹ MAKOTO SATO¹
SATOSHI MIMURA¹ TAKASHI MATSUMOTO¹ RYOICHI SASAKI¹

1. はじめに

2015年6月に起きた日本年金機構での情報流出の被害事例に挙げられるように近年, 特定の組織や個人を攻撃対象とする標的型攻撃が社会問題となっている[1]。標的型攻撃に使用されるマルウェアでは, 攻撃者が事前に入手した標的の情報を元に従来のマルウェアをカスタマイズしたマルウェアが使用されることが多く見られる。

このように従来のマルウェアをカスタマイズしたものは亜種と呼ばれ, 増加の一途を辿っている [2]。アンダーグラウンドのブラックマーケットでは開発ツールの取引が行われており[3], マルウェアに関する技術的な知識がなくても作成が容易に効率的になったことが, 爆発的に増加させている原因だと考えられる。

標的型攻撃の対策では, マルウェアの亜種の増加が急激であることに加え, 亜種ごとにカスタマイズがなされていることにより, シグネチャの作成や定義ファイルの更新が亜種の登場に追いつかず, 対策が後手に回っていることが現状の問題として挙げられる[4]。

このような状況の中で, 著者らは標的型攻撃に対処するため, 各種ログや攻撃事象の特徴から人工知能技術を用いて応急対応のガイドや自動運転の支援を行う LIFT (Live and Intelligent Network Forensic Technologies) システムおよ

び, その発展形である Super-LIFT の開発を行ってきた。これにより, 既に存在したマルウェアに対応したインシデントに対しては発生事象を正しく認識でき, 対応が可能であることを示した。

しかし, マルウェアの亜種の増加に伴い, 現状の方式で増加する亜種に対応していくには, マルウェアの亜種を解析し, その都度チューニングする必要があり, 対策が後手になってしまう。そこで本研究では, 後手に回る対策に先手を打つため, 亜種の発生パターンを分析し, 今後発生する亜種の予測を行う。

数多く出現しているマルウェアの亜種の中でも, 特に日本年金機構の情報流出の際に使用された「EMDIVI」[5]と「EMDIVI」より以前も以前に出現しており, 使用率が高い「PlugX」, 「POISONIVY」の3種それぞれの亜種に対し, 分析を行った。

マルウェアの収集には Virus Total[6]を用いて亜種を収集し, オンライン型のサンドボックスである Lastline[7]を用いて解析を行った。そして, マルウェアごとの機能の変化を比較し, 先行する「PlugX」, 「POISONIVY」の亜種の発生パターンから, 「EMDIVI」の今後発生する亜種の予測ができないか分析を行った。

亜種の分析に関してはいろいろな研究が行われているが[8], 本研究のように「PlugX」や「POISONIVY」の亜種を解析し, 「EMDIVI」の亜種の予測を行った研究は, 著者が調査した範囲では見当たらなかった。

本論文では, 第2章で先行研究, 第3章で本研究をどの

¹ 東京電機大学
Tokyo Denki University

ような手順で行うか、第4章で第3章の手順を行うことで得られた結果、第5章で考察、第6章で今後の展望について述べる。

2. 先行研究

本研究に先行する研究として、著者らが所属する東京電機大学内の情報セキュリティ研究室およびサイバーセキュリティ研究所内で行なわれている LIFT システムおよび Super-LIFT システムの研究開発が挙げられる。各研究の概要は以下の通りである。

2.1 LIFT

LIFT (Live and Intelligent Network Forensic Technologies) システムとは、2013年後期より東京電機大学のサイバーセキュリティ研究所内のLIFTプロジェクト内で開発されているシステム[9]であり、高度な知識を持ったシステム管理者を複数有しない組織であっても標的型攻撃への対応を可能にすることを目的とし、各種サーバのログやアラートといった各種セキュリティ情報を元に、ルールベースシステムを用いて専門家の状況判断や攻撃推定を再現し、適切な対応の算出とガイドを行うことで、運営者が適切な対応が行えるような統合管理を行うシステムである。

LIFTシステムの概要は図1の通りである。

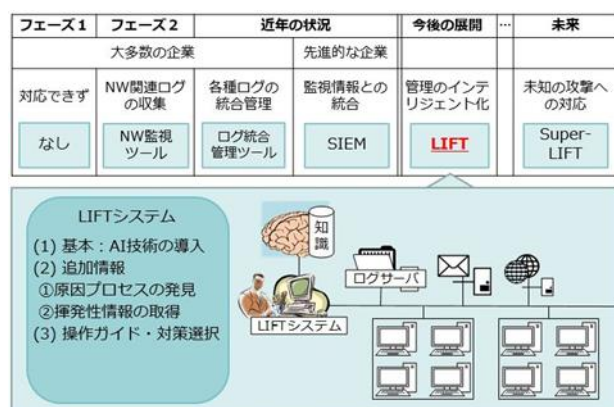


図1 LIFTシステムの概要

LIFTシステムの主な機能は、現在晒されている攻撃を各種セキュリティ情報から総合的に判断し、推定する「攻撃の推定」機能、推定した攻撃に対して適切な対応を動的に選択、管理者の指示に従ってガイドや実行を行う「応急対応」を行う機能、レポート出力や証拠保全を行うことで、事後対応を行うセキュリティ技術者を支援する「事故対応」を行う機能の3つである。

「応急対応」は「攻撃の推定」の後に行なわれ、「事故対応」は他の2つの機能と平行して行なわれている。

LIFTシステムの機能の概要図は図2、機能のフローは図3の通りである。

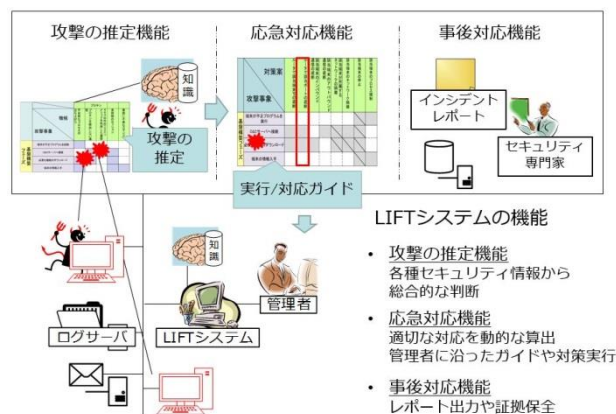


図2 LIFTシステムの機能概要

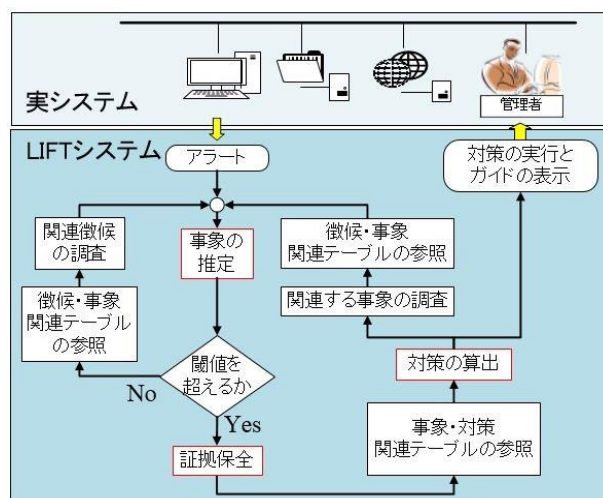


図3 LIFTシステムの機能フロー

2.2 Super-LIFT

Super-LIFTは、人工知能技術を利用することで、LIFTシステムだけでは対処が困難である未知の攻撃に対応する事を目的としている。Super-LIFTでは、LIFTシステムに対し「AI利用攻撃ケース自動生成システム」、「対応ルール自動生成システム」の2つのサブシステムを追加することで未知の攻撃に対する対策を行う [10]。

Super-LIFTシステムの概要は図4の通りである。

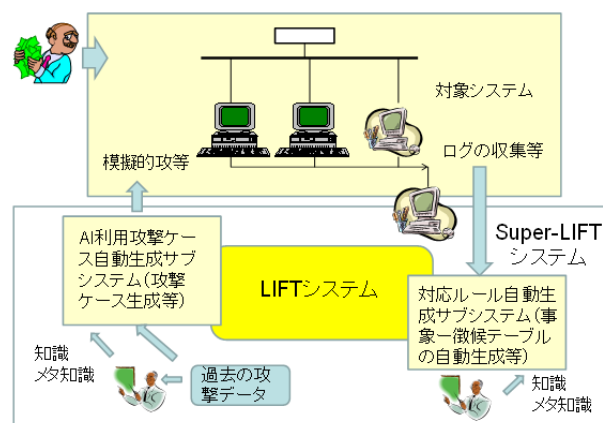


図4 Super-LIFTシステムの概要

Super-LIFTで使用される2つのサブシステムについては以下の通りである。

- AI利用攻撃ケース自動生成システム

AI利用攻撃ケース自動生成システムの概要図は図5の通りである。

このシステムでは、既存の攻撃だけでなく今後出現する未知の攻撃に対しても自動的な対応を可能とするため、収集された攻撃データを元に攻撃を生成し、実験環境内で模擬攻撃を行う。

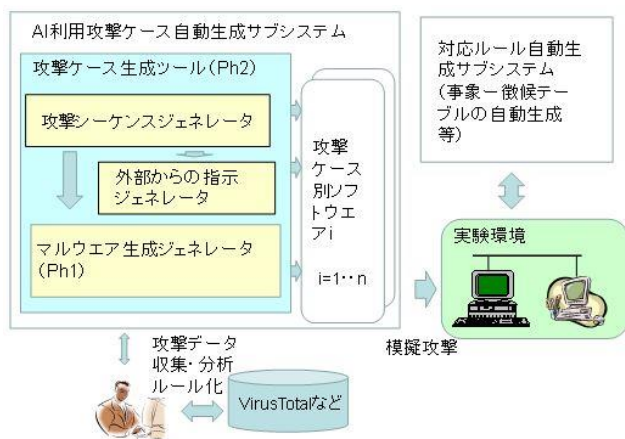


図5 AI利用攻撃ケース自動生成システム

- 対応ルール自動生成システム

対応ルール自動生成システムの概要図は図6の通りである。

このシステムでは、模擬攻撃を通して得られた結果を用いて、攻撃に対応したルールの自動生成を行う。これにより、従来人力で行っていたルール生成が自動化されている。

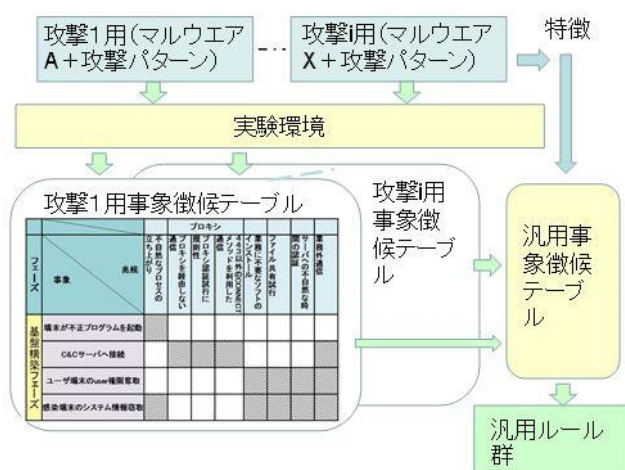


図6 対応ルール自動生成システム

以上の2つのサブシステムによりSuper-LIFTでは既存の攻撃のみでなく、未知の攻撃であっても対策を行う。本研究は、Super-LIFTシステム内のAI利用攻撃ケース自動生成システムに使用される「攻撃データの収集・分析・ルール化」部の分析・ルール化を行うものである。

3. 亜種の機能分析

本研究では、①標的型攻撃で使用されるマルウェアの亜種を収集し、②収集した検体についてサンドボックス環境での動的解析を行う。③動的解析により得られた結果から機能に関する情報を取得し、④収集したデータを元に、亜種の機能の予測を行う。

以降、各手順について説明する。

3.1 マルウェアの収集

3.1.1 対象マルウェア

収集の対象となるマルウェアは、Trend Micro社の国内標的型攻撃レポート[11]に基づき使用率が高く、日本年金機構での攻撃に使用され、今後出現する亜種の予測の対象となる「EMDIVI」の亜種と、先んじて出現しているマルウェアの内、出現率の高い「PlugX」、「POISONIVY」のマルウェア亜種の3種とする。

選定に用いたマルウェアの使用率は図7の通りである。

また、3種の亜種の特徴は下記のとおりである。

A) EMDIVI

2014年に登場したマルウェアで、現在最も使用率が高い。日本の企業、組織に対して多く使用され、日本年金機構における情報流出の際にも使用された。

B) PlugX

2012年後半に登場した亜種で、「EMDIVI」に次ぐ使用率の高さで、未だ増加傾向にある。

C) POISONIVY

2010年前後に登場した亜種で、今回扱う3種の中では最も古くから登場している亜種である。

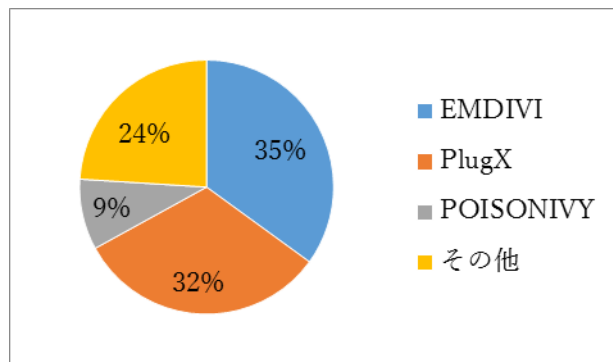


図7 標的型攻撃におけるマルウェアの使用率

3.1.2 収集方法

対象となるマルウェアの収集に際し、ファイル、URLのオンラインスキャンサービスであるVirusTotal[6]を使用した。このVirusTotalでは、40種類以上のウィルス対策製品での検知結果を閲覧、検体の入手が可能となっている。

収集方法については「EMDIVI」、「PlugX」、「POISONIVY」とマルウェアの名前を検索のキーワードとして使用することで、VirusTotal内のウィルス対策製品により該当するマルウェアとして検知された検体の検索が行えるため、該当した検体に関して収集を行った。また、出現期間に関しては、2015年10月までに出現したマルウェアを対象とした。

3.2 マルウェアの解析

収集したマルウェアの解析に際し、lastline社が提供するLastline[7]というオンライン型サンドボックスの、特にLastline Analystという機能を使用して解析を行った。このLastline Analystでは、検体を投稿することでサンドボックス環境での動的な解析を行うことが可能である。

Lastline Analyst による解析結果として、解析の概要を表示する部分に加え、動作環境ごとに詳細な情報をレポートとして得ることが出来る。

3.3 解析結果の利用

分析に使用するマルウェアの情報は、Lastline Analystにより得られた解析結果の概要部分をマルウェアの機能として、検体がVirusTotalへ初投稿された日時情報をマルウェアの出現日時としてそれぞれ使用する。

マルウェアの出現日時として、検体を解析した結果から得られたタイムスタンプ情報でなく、VirusTotalへ初投稿された日時情報を利用した。これは、検体から得られた情報であれば攻撃者によって改竄されている恐れがあるが、攻撃者の手から離れたVirusTotal内の情報であれば改ざんの可能性が低いと考えたためである。

また分析に関しては、「EMDIVI」の亜種を解析した際に得られた機能が、「EMDIVI」、「PlugX」、「POISONIVY」の亜種でそれぞれどのように出現しているか期間ごとの比較、マルウェアの種類と出現日時の関係を多次元尺度方[12]を用いての比較をそれぞれ行う。

4. 結果

第3章で示した手順に従い、「EMDIVI」の亜種を65検体、「PlugX」の亜種を114検体、「POISONIVY」の亜種を76検体収集し、サンドボックス環境で動的な解析を行った。

収集した検体の内、出現率の高かった2015年6月から9月の期間に関して、各マルウェアの出現数をまとめたものが図8である。

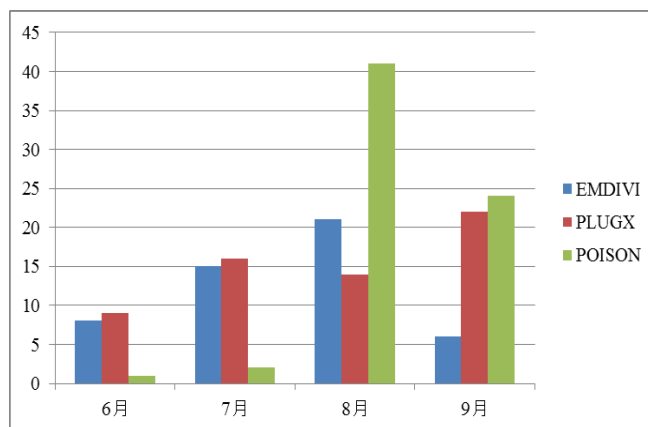


図 8 各月における亜種の出現数

また、「EMDIVI」、「PlugX」、「POISONIVY」の各マルウェアの亜種から得られた機能の内、「EMDIVI」の亜種が持つ機能について出現期間ごとに示したものが表1である。

表1より、機能の変化が顕著であった項目について以下に説明する。

● シグネチャによる検知結果

6月から9月の期間で「EMDIVI」の亜種から得られたシグネチャによる検知結果の出現率に関しては、図9のように推移している。

図9より、6月の段階では「トロイの木馬」としての検知結果が87.5%と殆どであったが、それ以降の出現率をみると減少する傾向にあり、9月に出現した検体に関しては「トロイの木馬」として検知されたものはない。これに対し「バックドア」もしくは「悪意のあるダウンロード」として検知されている検体は増加傾向にある。このことより、マルウェア自体が持つ機能が減り、外部から必要な機能を用意するといった傾向が顕著になっていることが考えられる。

また表1より、「PlugX」、「POISONIVY」の亜種についても同様に4ヶ月間のシグネチャによる検知結果を見ると、検知結果に変化は見られない。

表 1 各月における機能の出現率

		Jun-15			Jul-15			Aug-15			Sep-15		
		Em	PI	Po	Em	PI	Po	Em	PI	Po	Em	PI	Po
シグネチャ	バックドアのコード	12.5%	11.1%	100%	26.7%	0.0%	100.0%	42.9%	0.0%	100.0%	83.3%	4.5%	12.5%
	悪意のあるダウンロードのコード	0.0%	0.0%	0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	16.7%	0.0%	0.0%
	トロイの木馬のコード	87.5%	88.9%	0%	66.7%	100.0%	0.0%	57.1%	100.0%	0.0%	0.0%	95.5%	87.5%
自動起動	ログオンプロセス時に自動起動	75.0%	0.0%	0%	20.0%	0.0%	0.0%	14.3%	0.0%	0.0%	16.7%	4.5%	0.0%
	Windows起動時に自動起動の登録	0.0%	77.8%	100%	0.0%	12.5%	50.0%	0.0%	21.4%	19.5%	0.0%	22.7%	8.3%
	Windowsスタートメニューを使用した自動起動の登	0.0%	0.0%	0%	53.3%	0.0%	0.0%	38.1%	0.0%	0.0%	16.7%	4.5%	0.0%
ステルス	バックグラウンド動作時、ドキュメントを表示	25.0%	0.0%	0%	13.3%	0.0%	0.0%	14.3%	0.0%	0.0%	33.3%	0.0%	0.0%
	ファイルのダウンロード機能	0.0%	0.0%	0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	16.7%	0.0%	0.0%
ネットワーク	C&Cサーバとのトラフィックの確認	50.0%	11.1%	0%	60.0%	6.3%	0.0%	14.3%	28.6%	0.0%	0.0%	4.5%	0.0%
	到達不能なHTTPリンクへのリクエスト	62.5%	11.1%	0%	60.0%	6.3%	0.0%	28.6%	7.1%	0.0%	16.7%	0.0%	0.0%
回避	特定の画像ファイル名を検索	0.0%	0.0%	100%	0.0%	6.3%	0.0%	0.0%	0.0%	2.4%	0.0%	9.1%	0.0%
	特定のプロセスの検索	0.0%	0.0%	0%	0.0%	0.0%	0.0%	0.0%	0.0%	36.6%	66.7%	0.0%	0.0%
検索	実行中のプロセスの列挙	0.0%	11.1%	0%	0.0%	6.3%	0.0%	4.8%	78.6%	0.0%	66.7%	68.2%	0.0%
	ユーザアカウント名の取得	0.0%	88.9%	0%	0.0%	62.5%	0.0%	14.3%	35.7%	0.0%	50.0%	36.4%	0.0%
設定	デバックの権利を付与	12.5%	0.0%	0%	13.3%	0.0%	0.0%	33.3%	0.0%	0.0%	50.0%	4.5%	0.0%
例外	不要なアプリケーション/プログラム	0.0%	0.0%	0%	13.3%	0.0%	0.0%	9.5%	0.0%	0.0%	0.0%	0.0%	0.0%
メモリ	スタックオーバーフローを起こす	12.5%	11.1%	0%	0.0%	0.0%	0.0%	0.0%	14.3%	0.0%	0.0%	4.5%	0.0%

以上のことから、シグネチャによる検知結果の変化に関して、「EMDIVI」の亜種と「PlugX」、「POISONIVY」の亜種との間には関連が無いと考えられる。

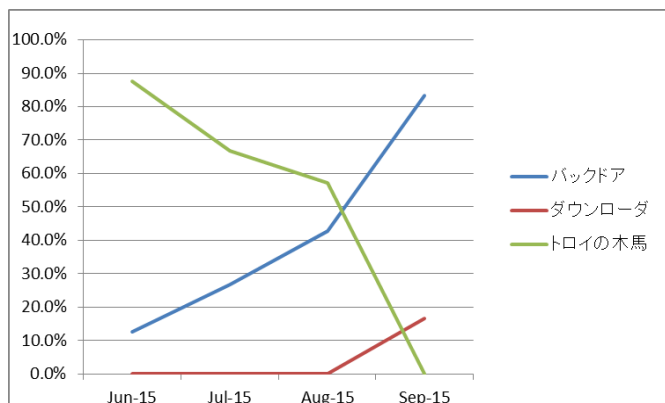


図9 シグネチャによる検知結果の推移

● ネットワークに関する機能

6月から9月の期間で「EMDIVI」の亜種から得られたネットワークに関する機能の出現率に関しては図10のように推移している。

標的型攻撃で使用するマルウェアでは通常、侵入後にC&Cサーバとの通信を行う。しかし今回の解析で得られた結果では、図10にあるように「C&Cサーバとのトラフィックの確認」、「到達不能なHTTPリンクへのリクエスト」出現率は共に減少する傾向にある。

「C&Cサーバとのトラフィックの確認」が減少している要因としては、「EMDIVI」では日本国内の正規のサイトにC&Cサーバを置くことで通信を正規のものに偽装している[13]ためだと考えられる。このことより、今後も「C&Cサーバとのトラフィックの確認」は減少すると考えられる。

「到達不能なHTTPリンクへのリクエスト」に関しても減少する傾向が見られたが、これは先に述べた通り、正規サイトにアクセスするため、「到達不能なHTTPリンクへのリクエスト」するということが今後も減少する傾向にあると考えられる。

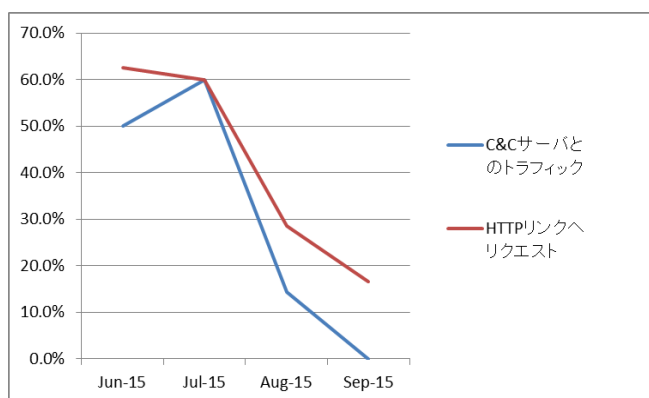


図10 ネットワークに関する機能の推移

● 検索機能

6月から9月の期間で「EMDIVI」の亜種から得られた検索機能の出現率に関しては図11のように推移している。

図11より、動作環境を検索する機能である「実行プロセスの列挙」、「ユーザアカウント名の取得」の2つの項目において、7月以降出現率の増加が確認できた。一方、「PlugX」、「POISONIVY」の亜種に関しては7月以前に確認されていた機能であるが、「EMDIVI」の亜種に関しては7月以前には確認できなかった機能である。この傾向から今後も機能「実行プロセスの列挙」、「ユーザアカウント名の取得」の2つの項目に関して多く出現する機能であると考えられる。

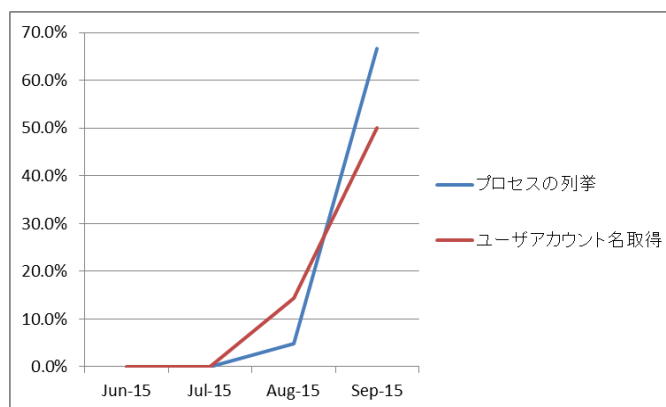


図11 検索機能の推移

また、収集した各マルウェアの亜種を多次元尺度法により、亜種ごと、期間ごとの関係を表したものが図12、図13である。

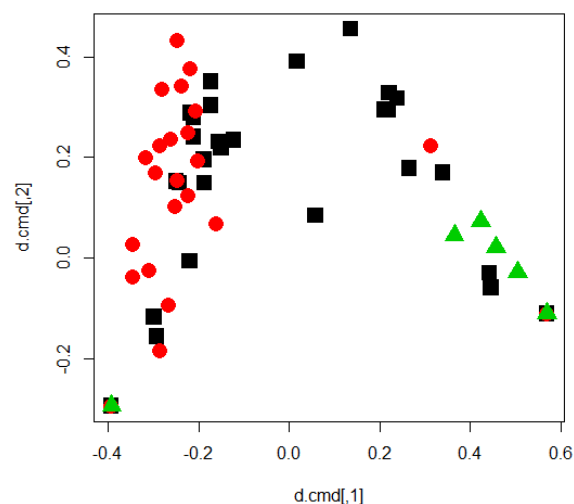


図12 マルウェアごとの機能比較

図12に関して、それぞれ「EMDIVI」の亜種を「■」、「PlugX」の亜種を「●」、「POISONIVY」の亜種を「▲」で示している。「EMDIVI」の亜種については全体に広がりが見えるが、「PlugX」の亜種は右側に、「POISONIVY」の亜種は左右の下側はそれぞれ偏りが見られた。

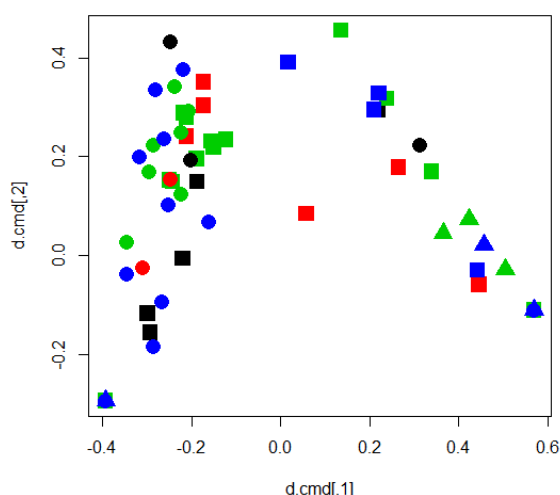


図 13 出現期間ごとの機能比較

図13に関しては、図12に対し、それぞれ6月を「黒」、7月を「赤」、8月を「緑」、9月を「青」と月ごとに分けてみた。そうしたところ、各月で広がりが見られ、関連性は見られなかった。

以上のクラスタリングの結果より、期間ごとの関連性は低く、マルウェアごとの関連性が高いと言える。

5. 考察

第4章より、各マルウェアの機能がどのように変化するかを明らかにすることで、次のような傾向がみられることが明らかになった。

- 1) 「EMDIVI」は「バックドア」として検知される比率が増大の傾向にあり、今後も続くものと予想される。
- 2) 「EMDIVI」は「C&Cサーバとのトラフィック」、「到達不能なHTTPリンクへリクエスト」に関して、共に減少する傾向がみられ、今後も減少する傾向だと考えられる。
- 3) 動作環境を検索する機能である「実行プロセスの列挙」、「ユーザアカウント名の取得」の2つの項目においては、先行する2種のマルウェアに追従する形で現れており、「EMDIVI」の予測に使用できる可能性がある。

また、多次元尺度法を用いて得られた結果からは、マルウェア間で亜種が発生するパターンの方が、時期による変化よりも大きい傾向にあり、「EMDIVI」の亜種の動向を予測できるという結論には至らなかった。

しかし、今回の機能比較に用いたデータが2015年6月から9月までとデータ量として満足とはいえない。そこで、データ量の増加を図る必要がある。

また、本研究で収集および動的な解析を行った検体の中にはシグネチャによる検知結果のみの検体が存在した。この解析結果が、本当に検体に機能が存在しないことによるものか、マルウェアが持つ耐解析機能により Lastline 上で動作を行わなかったためであるか、更なる調査の必要があると考える。

6. おわりに

本研究では、日本年金機構の情報流出の際に使用された「EMDIVI」と、先行する「PlugX」、「POISONIVY」、それぞれの亜種で同時期に確認した機能を比較した。その結果、各マルウェアの機能がどのように変化するかを明らかにすることができた。また、多次元尺度法を用いた分析の結果、マルウェア間で亜種の発生パターンは異なることが明らかになった。現状では、先行するマルウェアの亜種の動向から「EMDIVI」の亜種の動向を予測することは困難かもしれないと考えているが、さらに調査期間を延ばして分析を続けたいと考えている。

そのため、今後はデータ量増加に向け、2015年10月以降に出現した対象マルウェアの亜種を収集していく共に、VirusTotalに投稿されていないような検体を収集する手段の検討を行う。そのほか、変化が起きた背景にどのような社会的な事象があるか調査を行うことで更なる機能の予測および本研究で得られた結果の検証を行う。

また、本研究では検体の収集・解析・分析全てを手動で行っている。この部分を、各サービスを連携させることで自動化させ、より多くのデータの収集を行う。

参考文献

- [1] MOTEX:日本年金機構の情報漏えい、明日は我が身！標的型攻撃メールの脅威、MOTEX(オンライン), 入手先<<http://www.motex.co.jp/nomore/casestudy/2879/>>.
- [2] G DATA:G DATA SECURITYLABS MALWARE REPORT, G DATA (online), available from<https://public.gdatasoftware.com/Presse/Publikationen/Malware_Reports/GData_PCMWR_H2_2014_EN_v1.pdf>.
- [3] Symantec:アンダーグラウンドのブラックマーケット:盗難データ、マルウェア、攻撃サービスの取引が盛況、Symantec(オンライン), 入手先<<http://www.symantec.com/connect/ja/blogs-367>>.
- [4] Canon:ヒューリスティック「基礎編」未知のウイルスも検出して駆除する ESET のヒューリスティック機能、Canon(オンライン), 入手先<http://canon-its.jp/eset/malware_info/technology/140626/>.
- [5] ZDNet Japan: 年金機構事件で発覚-感染に気付かれず潜伏する「Emdivi」の恐ろしさ、ZDNet Japan(オンライン), 入手先<<http://japan.zdnet.com/article/35065996/>>.
- [6] VirusTotal: <https://www.virustotal.com/>
- [7] Lastline Analyst: <https://www.lastline.com/platform/analyst>
- [8] 大久保 諒, 森井 昌克, 伊沢 亮一ほか: マルウェアの部分コードによる類似度判定および機能推定, コンピュータセキュリティシンポジウム 2012 論文集, Vol.2012, No.3, pp. 9-14 (2012).
- [9] 比留間裕幸, 橋本一紀, 柿崎淑郎ほか: 標的型攻撃に対する知的ネットワークフォレンジックシステム LIFT の開発 (その 1) - 予兆検知と対策方法の提案 -, DICOMO 2015, pp. 29-37 (2015).
- [10] 佐々木良一, 八槇博史: 標的型攻撃に対する知的ネットワークフォレンジックシステム LIFT の開発 (その 3) - 今後の構想 -, DICOMO2015, pp.44-50 (2015).
- [11] TREND MICRO:国内標的型サイバー攻撃分析レポート 2015 年版, TREND MICRO(オンライン), 入手先<https://app.trendmicro.co.jp/doc_dl/select.asp?type=1&cid=161>.
- [12] MACROMILL:MDS(多次元尺度構成法)とは, MACROMILL(オンライン), 入手先<<http://www.macromill.com/landing/words/b016.html>>
- [13] TREND MICRO:Flash Player のゼロデイ脆弱性「CVE-2015-5199」による標的型攻撃を国内で確認, TREND MICRO(オンライン), 入手先<<http://blog.trendmicro.co.jp/archives/11944>>.