進化する標的型攻撃に対抗するマルウェア自動解析技術

仲小路 博史 Nakakoji Hirofumi
鬼頭 哲郎 Kito Tetsuro
重本 倫宏 Shigemoto Tomohiro
林 直樹 Hayashi Naoki
山下 真吾 Yamashita Shingo

近年，標的型攻撃に利用されるマルウェアが高度化し，既存の入口対策では検知できないまま組織内への侵入を許してしまうケースが増えている。この場合，侵入したマルウェアの特性を解明して早急な被害拡大防止策を講じる必要がある。マルウェアの特性を解明する手法としては，マルウェアを特殊な解析環境で実行して挙動を観測する動的解析手法が行われている。一方，最近のマルウェアは実行環境を限定することで解析環境での解析を逃れるタイプも増えている。また，機密情報に寄生するマルウェアも存在し，安易に外部に解析作業を委託できない状況となっている。
それに対し，マルウェア多種環境動的解析システムは，マルウェアを多種類の解析環境で実行させることにより，環境を選ぶマルウェアも自動的に解析する。このシステムをスタンドアローンで動作させることにより，外部のサービスに頼ることなく自組織内でマルウェアの特性を解明できる。

1. はじめに

1970年代に世界で初めてコンピュータウイルスが確認されて以来，コンピュータ管理者は，半世紀近くにわたって進化し続けるコンピュータウイルスと対峙（じ）してきた。ウイルスの多様化が進んだ現在では，コンピュータウイルスに代表される悪意を持って開発されたソフトウェアを総称して「マルウェア」と呼んでいる。標的型攻撃などで世間をにぎわしているサイバー攻撃ではマルウェアが道具として利用され，金銭や機密情報の搾取，インフラシステムの破壊などを目的とした犯罪のプロ化が進んでいる。これに伴ってマルウェア自体の高度化，多様化，巧妙化もさらに進んでおり，ファイアウォールやパターン照合型のウイルス対策といった従来型の入口対策では対処が困難な状況になっている。

守る側も標的型攻撃対策の一環として従業員に向けた標的型攻撃対策訓練を実施しており，技術面だけでなく，従業員のセキュリティリテラシーの向上などの運用面の対策も進めている。このような施策が奏功し，不審なメールを受け取った従業員からの通報や不審な検体の提供を情報システム部門が受けるケースも増えており，既存のセキュリティ対策では検出できなかった，いわゆる未知のマルウェアとおぼしき検体を入手する機会が増えている。情報システム部門は，インシデント予防および対策の観点からその検体がマルウェアか否かを判断するとともに，マルウェアであった場合にそのマルウェアの有する機能を解明し，従業員が感染してしまった場合の対策などを検討して，被害の発生や拡大を防止するための内部対策や出口対策を早急に講じることが重要である。

ここでは，従来，高度な専門知識を有したマルウェア解析者が手作業で行ってきた解析作業を自動化し，効率よくマルウェアの挙動を解明するマルウェア多種環境動的解析システムについて述べる。

2. マルウェア解析の課題

検体がマルウェアか，あるいはマルウェアとしてどのような機能を有するのかを解明するには，専門家による検体の解析が必要となる。検体の解析には，リバースエンジニアリングなどの技術によって検体を実行せずに解析する静的解析手法と，検体を特殊な解析環境で実際に実行してふるまいを観測する動的解析手法がある。静的解析は，検体の機能のすべてを詳細に解明できる利点があるが，プログラムやOS（Operating System），ハードウェアの仕組みなどに関する深い知識と，コードを1行ずつ読み解くための膨大なコストが必要となる。動的解析は，難読化（コード

の暗号化など）された検体でも手を加えずに解析できるため，静的解析と比較して短時間で解析できる。静的解析だけでは分からない挙動（新たなマルウェアをインターネットからダウンロードして実行した後の挙動など）を確認できる点で有利であるが，観測中に動作しない機能はその挙動を明らかにできないという欠点もある。通常，検体を解析する際には，検体の性質，解析の目的，解析者の経験則に従って静的解析と動的解析を補完的に組み合わせて実施する場合が多い。

最近では，動的解析を支援する解析用ソフトウェアが開発されており，OSS（Open Source Software）ではCuckoo Sandbox[1]がWebから入手可能である。このようなソフトウェアは仮想化技術を利用してサンドボックス（解析環境）内で検体を安全に実行し，さらにネットワーク通信やAPI（Application Programming Interface）コールの観測結果を詳細に取得できるため，解析を実務とする多くの専門家によって利用されている。また，ThreatExpert[2]のような検体の挙動を解析するサービスもセキュリティベンダーによって提供されており，インターネットを経由して検体を提出することにより，解析結果を入手することが可能である。

これまでに述べたように，守る側も技術やツールの進化により，以前と比較して検体の解析が容易になってきた。しかし，最近ではマルウェア開発側も自分が作成したマルウェアの検知・解析を回避するような仕掛けを組み込むようになっており，マルウェアが仮想環境やデバッグ環境，OSのバージョン，インストールアプリケーションなどのハードウェア／ソフトウェア構成を検知し，みずからが意図する攻撃の対象であるか否かを判断して動作を変える環境選択型マルウェアの存在が確認されている。また，攻撃者が用意したマルウェア配布サーバから第二のマルウェアをダウンロードさせることで，攻撃を段階的に進めるダウンローダ型マルウェアも確認されている。さらに，マルウェア配布サーバの中には，アクセス元のIP（Internet Protocol）アドレスが攻撃対象の組織である場合にだけマルウェアを配布し，それ以外の場合には正規のコンテンツを配布することでみずからの存在を隠蔽するものまである。

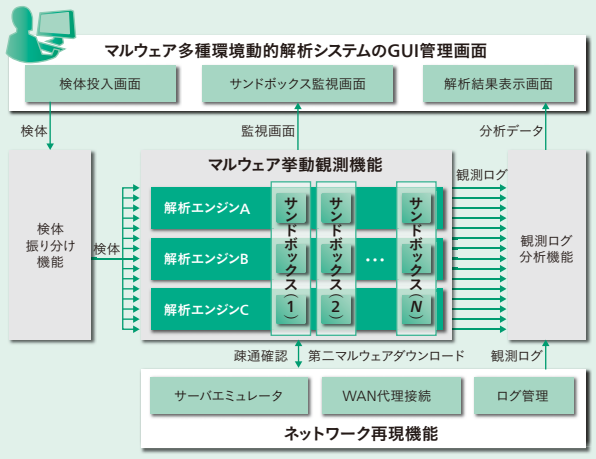
このような仕組みを持つマルウェアは，特定の環境しか用意されていない既存の動的解析ソフトウェアによる解析では，その挙動が明らかにできないケースが多い。また，標的型攻撃の対象の組織のIPアドレスを持たない外部の解析サービスに解析を委託した場合，マルウェア配布サーバが正規のサーバとして解析者に対してふるまうため，やはりその挙動を明らかにできない。さらにはPDF（Portable Document Format）ファイルなどの機密情報を含むファイルに寄生するマルウェアも確認されていることから，「マルウェア＝機密情報」であると考え，外部のサービスを利用した検体解析の依頼をためらう組織も増えてきており，自組織内でマルウェアの特性を解明したいというニーズが高まっている。

日立製作所横浜研究所は，そのような課題を解決するために，検体を複数種類の解析環境の上で自動的に動的解析し，検体のマルウェアらしさや，検体の特徴を提供する技術の研究開発を進めている。

3. マルウェア多種環境動的解析システム

マルウェア多種環境動的解析システムは，環境選択型マルウェアの解析成功率を向上させるため，複数種類の解析エンジン，複数種類のサンドボックスで解析する。このシステムのアーキテクチャを**図1**に示す。

マルウェアとおぼしき検体を解析する解析者が，このシステムの検体投入画面を通して検体を投入（アップロード）すると，検体は検体振り分け機能によってマルウェア挙動観測機能に構成されている各サンドボックスに複製されて同時に投入される。サンドボックスでは，投入された検体を自動的に実行して挙動を観測し，結果をログとして出力する。観測ログ分析機能は，マルウェア挙動観測機能から出力された大量のログ（場合によって1検体当たり数百万行，数ギガバイトに及ぶ）を収集し，各サンドボックスにおける検体の活動状況（ファイルアクセス，レジストリアクセス，ネットワークアクセスなど）を統計的に分析したり，検体による生成ファイルや，ネットワーク接続先



注：略語説明　GUI（Graphical User Interface），WAN（Wide Area Network）

図1｜マルウェア多種環境動的解析システムのアーキテクチャ
マルウェアが感染・動作しやすいように複数種類のサンドボックス（解析環境）を用意し，そこで得られた挙動観測ログをこれまでの解析ノウハウを基に自動分析して報告する仕組みにより，従来は高度な技術を持った専門家が解析した作業を自動かつ短時間・高成功率で行うことに成功した。

URL（Uniform Resource Locator）を抽出したりする。これらの処理を同時並行で自動的に実施するため，解析時間を大幅に短縮でき，さらに解析作業の夜間バッチ化も実現できる。

このシステムの特徴を以下に述べる。

3.1　マルウェア挙動観測機能

マルウェア多種環境動的解析システムは，検体を数十種類のサンドボックスで解析することにより，環境選択型マルウェアの解析成功率の向上を実現している。サンドボックス群は解析エンジンやハードウェア，ソフトウェアの種類やバージョン，設定内容などの異なる組み合わせによって構成される。サンドボックスのバリエーションが多いほど環境選択型マルウェアの解析成功率は向上するが，使用できる物理マシンのリソースやライセンスの制約により，すべての組み合わせを用意することは現実的ではない。

このシステムでは，選定要素として，(1) 解析エンジン，(2) ハードウェア，(3) アーキテクチャ，(4) OS，(5) アプリケーションの5つを定義し，サンドボックス構成を設計した（**表1**参照）。

それらの選定要素のうちマルウェア多種環境動的解析システムの解析エンジンは，前述したCuckoo Sandboxを含む合計3種類を採用している点が特徴である。解析エンジンは種類によってサポートする仮想マシンが異なっているため，解析エンジンの多種化は解析性能のほか，仮想機能検知機能を有するマルウェアの解析にも効果が期待できる。

OSやアプリケーションは，種類やバージョンなどのバリエーションが極めて多く，組み合わせ爆発の原因となる。このシステムはサンドボックスをマルウェアに感染させ，可能な限り多くの挙動を解明することを目的としていることから，マルウェア開発者の視点に立脚してマルウェアが感染および動作しやすい環境，つまり攻撃の影響を受けやすい環境を優先的に選定した。そこでOS構成は，マルウェアの感染が多く報告されているWindows XP以降の主要OSをService Packまで区別して設計した。また，アプリケーション構成は，脆（ぜい）弱性が多いアプリケーション，すなわち脆弱性情報の公開数の多いアプリケーションを優先的に選定した。脆弱性情報の公開数の調査にあたっては，脆弱性対策情報データベース「JVN iPedia[3)]」の2012年1月1日から2013年8月16日までの情報を利用した。

3.2　ネットワーク再現機能

近年のマルウェアは，ネットワーク接続機能を有し，マルウェア配布サーバに接続して第二のマルウェアをダウンロードしたり，C&C（Command and Control）サーバと接続して遠隔操作を受けたりすることが知られている。また，マルウェアの中には，解析から逃れるために，感染直後にネットワークの疎通性を確認することによって，解析環境でないことを確かめるタイプも確認されている。

マルウェア多種環境動的解析システムは，**図1**に示したようにネットワーク再現機能を有している。この機能は，Webサーバ，FTP（File Transfer Protocol）サーバ，DNS（Domain Name System）サーバなどの主要なサーバをエミュレーションしてサンドボックス内の検体からの各種サーバ向けリクエストに応答するサーバエミュレータ機能，インターネットに代理接続してマルウェア配布サーバやC&Cサーバと通信するWAN（Wide Area Network）代理接続機能（開発中）を具備する。これらにより，ダウンローダ型マルウェアが特定のWebサーバからファイルをダウンロードして実行するまでの挙動を高精度に再現し，観測することができる。

3.3　観測ログ分析機能

観測ログ分析機能は，数十種類のサンドボックスから取得した大量のログからマルウェア特有の挙動を抽出する。抽出アルゴリズムの設計にあたっては，マルウェアの解析を業務として行っている専門家の高度で実績のあるマルウェア解析ノウハウ（暗黙知）に基づいて機能設計（形式知）を実施した。以下に幾つか分析機能を紹介する。

(1) デバッガ検知機能の有無判定

(2) プロセスインジェクションの有無判定

表1｜**実行環境の選定要素**

実行環境の選定要素には，解析エンジン，ハードウェア，アーキテクチャ，OS（Operating System），およびアプリケーションの5要素を定義する。

解析エンジン	ハードウェア	アーキテクチャ	OS	アプリケーション
解析エンジンA	•物理マシン •仮想マシン（VMware*1 ESXi）	•32ビット（×86） •64ビット（×64）	•Windows*3 XP（sp x） •Windows Vista*3（sp x） •Windows 7（sp x）	•Microsoft Office*3 xxxx •Adobe*4 Reader*4 xx •Internet Explorer*3 xx •Adobe Flash*4 Player xxx.x •JRE x.x •Windows Media*3 Player xx
解析エンジンB	•仮想マシン（Oracle*2 VM VirtualBox）			
解析エンジンC	•仮想マシン（VMware Workstation）			

注：略語説明ほか　JRE（Java*2 Runtime Environment）

*1 VMwareは，米国およびその他の地域におけるVMware, Inc.の登録商標または商標である。

*2 Oracle，Oracle VM VirtualBox，Javaは，Oracle Corporationおよびその子会社，関連会社の米国およびその他の国における登録商標である。

*3 Microsoft，Microsoft Office，Internet Explorer，Windows，Windows Vista，Windows Mediaは，米国Microsoft Corporationの米国およびその他の国における登録商標または商標である。

*4 Adobe，およびAdobe Reader，Flashは，米国 Adobe Systems Incorporatedの米国およびその他の国における商標または登録商標である。

feature articles

(3) 時限的発動処理の有無判定
(4) 外部ネットワーク接続判定

こうした判定対象の挙動は，マルウェアが一連の不正活動を行う中で現れる場合が多い。このため，それらの挙動の有無を判定することはマルウェアか否かを推定するのに有用となる。また，各項目の判定方法には解析ノウハウに基づいた工夫を施している。例えば，外部ネットワーク接続の判定では，マルウェア解析者が検知を逃れるために利用するマイナーなネットワーク接続手法を含む複数種類のAPIコールを監視したり，通信トラフィック，ネットワーク再現機能のログなどを分析したりすることによって多角的に判定する。

3.4　解析結果表示

マルウェア多種環境動的解析システムには，数十種類のサンドボックスでの検体の動作結果をサマリーとして表示する機能と，個々のサンドボックスの解析結果を集約して一覧表示する機能がある（**図2**参照）。

この画面を確認することにより，検体の接続先URLや，作成ファイル，生成プロセス情報のほか，16種類のウイルス対策ソフトウェアのパターンファイルと照合して得られた検知結果を確認することができる。検体がマルウェアであった場合に，ウイルス対策ソフトウェアの対応状況や，従業員がそのマルウェアに感染してしまった際に発生する可能性のあるネットワーク接続の接続先URL，従業員端末に仕掛けられたトラップ（マルウェア関連ファイル）などを容易に把握することができる。これらの情報を用いてファイアウォールやプロキシなどで接続先URLへの通信を禁止したり，ウイルス対策ソフトウェアのパターンファイルに駆除情報を追加したりすることによって，入口対策をすり抜けて従業員の端末で感染・発症した場合でも，内部対策や出口対策を活用した多層防御が可能となる。

4.　マルウェア多種環境動的解析システムの検証

あるセキュリティベンダーのウイルス対策ソフトウェアでは検知できなかった未知マルウェアとおぼしき検体数百種を，マルウェア多種環境動的解析システムのプロトタイプを用いて解析して得られた結果について報告する。

今回の検証で使用したサンドボックスの数は約80種類であり，1つの検体にかかる解析時間（80種類のサンドボックスすべてで解析が完了するまでの時間）は15分程度であった。それらの検体のうち，不正なサイトに関連す

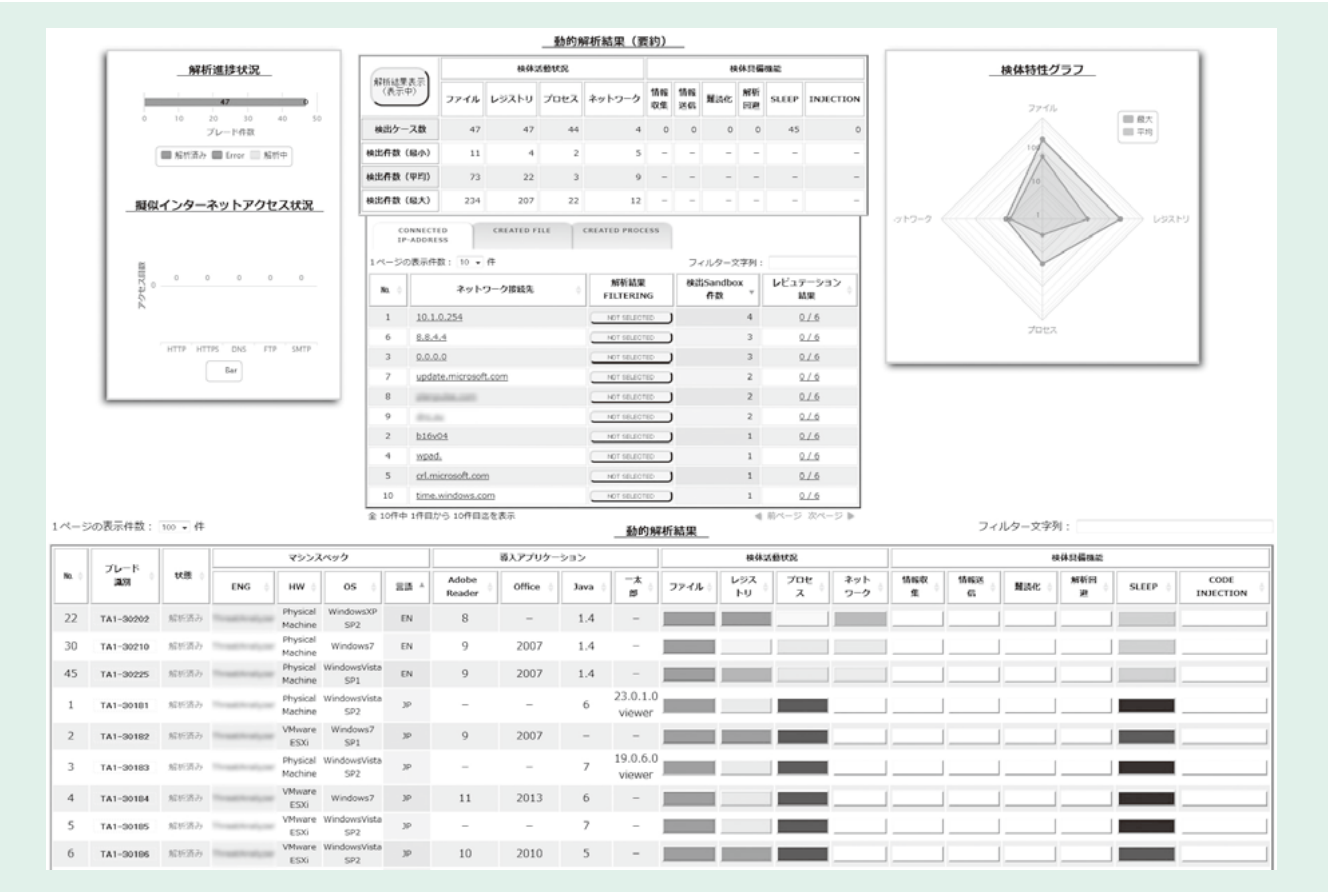


図2｜解析結果表示画面の例
数十種類のサンドボックスでのマルウェアの動作結果をサマリーとして表示する画面（上）と，個々のサンドボックスの解析結果を集約して一覧表示する画面（下）を示す。

ると思われる外部サーバに接続した検体が全体の約73%を占めており，近年のマルウェアがネットワーク接続を行う性質があることを改めて確認した。また，マルウェア多種環境動的解析システムの特徴である多種類の解析環境を再現したサンドボックスでマルウェアを解析することにより，以下の条件で顕現化する環境選択型マルウェアの存在を確認した。

(1) Microsoft Office 2007/2010がインストールされた環境でのみ活動する検体

(2) Windows XP上でのみ活動する検体

(3) 物理環境でのみ活動する検体

(4) 物理環境かつWindows 7（ただしService Pack 1を除く）上でのみ活動する検体

(5) VMware ESXiおよびVMware Workstationでは活動しないがOracle VM VirtualBoxでは活動する検体

つまり，これらの性質を有する検体は，動作条件が合致しない環境での動的解析が困難であることを示している。

今回の検証により，マルウェアの解析の自動化，未知マルウェアの不正な挙動の解明，および環境選択型マルウェアの実行ならびに挙動の解明ができることを確認した。また，ネットワーク接続などのマルウェアの挙動が顕現したサンドボックスの数や，顕現したサンドボックスの環境構成の共通点を抽出することにより，顕現の容易性や環境選択型マルウェアの実行可能な環境条件を導出することができる。これらの情報は，さらなる詳細な解析のための解析環境構築の手がかりにすることができる。

5. おわりに

ここでは，標的型攻撃に利用される不審なファイルの挙動を動的解析によって自動で解明して活動内容を報告するマルウェア多種環境動的解析システムについて述べた。

このシステムは，ハーフラック型にオールインワンシステム化し，スタンドアローンで稼働できるように設計している。このシステム単体でのマルウェアの挙動解析が可能なため，第二のマルウェアを特定のIPアドレスからしかダウンロードできないダウンローダ型マルウェアや，機密情報扱いとなる検体も自組織内で閉じて解析できるようになる。このシステムを組織の情報システム部門やセキュリティオペレーションセンターに導入することにより，マルウェア解析業務を行っている専門家の作業コストを大幅に削減できるほか，専門家不在の組織でも容易にマルウェアの脅威を明らかにできるようになる。これにより，被害状況の把握や，標的型攻撃などの最新のサイバー攻撃への多層防御対策に効果を発揮できるものと期待する。

今後は，解析時間のさらなる短縮，観測ログ分析機能の拡充，そしてこのシステムによって得られたマルウェアの特性情報に基づく自動対処の研究により，安全・安心なIT（Information Technology）環境を実現できるものと考える。

参考文献など

1) Claudio “nex” Guarnieri & Cuckoo Sandbox Developers：Automated Malware Analysis – Cuckoo Sandbox, http://www.cuckoosandbox.org/
2) ThreatExpert Ltd.：ThreatExpert - Automated Threat Analysis, http://www.threatexpert.com/
3) JPCERT/CC and IPA：JVN iPedia — 脆弱性対策情報データベース, http://jvndb.jvn.jp/

執筆者紹介

仲小路 博史
日立製作所 横浜研究所 情報サービス研究センタ エンタープライズシステム研究部 所属
現在，サイバー攻撃対策技術の研究開発に従事
情報処理学会会員

鬼頭 哲郎
日立製作所 横浜研究所 情報サービス研究センタ エンタープライズシステム研究部 所属
現在，サイバー攻撃対策技術の研究開発に従事
情報処理学会会員

重本 倫宏
日立製作所 横浜研究所 情報サービス研究センタ エンタープライズシステム研究部 所属
現在，サイバー攻撃対策技術の研究開発に従事
情報処理学会会員

林 直樹
日立製作所 横浜研究所 情報サービス研究センタ エンタープライズシステム研究部 所属
現在，サイバー攻撃対策技術の研究開発に従事

山下 真吾
株式会社日立アドバンストシステムズ エンジニアリング統括本部 事業推進本部 防衛情報システム部 所属
現在，情報セキュリティ製品の開発に従事