マルウェア対策ユーザサポートシステムを用いた CCC DATASet 2010 の解析

川口 信隆[*1]　余田 貴幸[*1]　川口 龍之進[*1]
寺田 真敏[*1]　笠木 敏彦[*2]　星澤 裕二[*3]
衛藤 将史[*4]　井上 大介[*4]　中尾 康二[*4]

*1 株式会社日立製作所　神奈川県川崎市幸区鹿島田 890　nobutaka.kawaguchi.ue@hitachi.com
*2 KDDI 株式会社 東京都千代田区飯田橋 3-10-10
*3 株式会社セキュアブレイン 東京都千代田区麹町 2-6-7
*4 独立行政法人情報通信研究機構 東京都小金井市貫井北町 4-2-1

あらまし 今日，日々数千から数万の新種のマルウェアが発生している．これに伴い，シグニチャを用いたマルウェア検知方式では，シグニチャの更新が新種の発生頻度に追いつかなくなりつつある．そこで，我々は，シグニチャに依らず，動的解析によりマルウェアを検知して駆除する「マルウェア対策ユーザサポートシステム」の研究開発を行っている．本システムでは，ユーザからの要請に応じてユーザ PC 内からマルウェアである可能性のある擬陽性ファイルを検知し，それらをマルウェア動的解析システムに送信する．擬陽性ファイルを動的解析した結果，マルウェアと判定されたものについては，駆除ツールを自動生成し，ユーザ PC に配布・実行してマルウェアを無害化する．本稿では，本システムを用いて CCC DATASet 2010 検体を解析した結果について報告する．

Analyzing CCC DATASet 2010 using User Support System against Malware

Nobutaka Kawaguchi[*1]　Takayuki Yoda[*1]　Tatsunoshin Kawagcuhi[*1]
Masato Terada[*1]　Toshihiko Kasagi[*2]　Yuji Hoshizawa[*3]
Masashi Eto[*4]　Daisuke Inoue[*4]　Koji Nakao[*4]

*1Hitachi,Ltd., 890 Kashimada, Saiwai-Ku, Kawasaki, Kanagawa　nobutaka.kawaguchi.ue@hitachi.com
*2 KDDI Corporation, 3-10-10 Iidabashi, Chiyoda, Tokyo
*3 SecureBrain Corporation, 2-6-7 Koujimachi, Chiyoda, Tokyo
*4 National Institute of Information and Communications Technology, 4-2-1 Nukui-Kitamachi, Koganei, Tokyo

Abstract With the increasing number of new malware species, traditional malware detection approaches relying on signature files are being less effective, since it is quite difficult for anti-virus vendors to keep up with the frequent appearance of new malware species. In this situation, we are developing a system called Anti-malware User Support System, which detects malware files using dynamic analysis and remove them from user PCs. This system first finds suspicious files from a user PC by means of a client agent. Then, the suspicious files are sent to and analyzed by a malware analysis system. Finally, this system removes detected malware files in the user PC by generating, sending and executing custom-made removable tools automatically. In this paper, we analyze malware files in the CCC DATASet 2010 using the proposed system and show the results.

1. はじめに

今日，日々数千から数万規模のマルウェアの新種が出現している．これに伴い，シグニチャファイルを基にユーザ PC 内からウィルスを検知するアンチウイルスソフトは，シグニチャファイルの更新が新種マルウェアの出現頻度に間に合わず，検知率が低下している．

一方で，近年では多数のセキュリティベンダや研究機関が，マルウェア動的解析システム

[1][2][3][6]や，非マルウェアのホワイトリストデータベース[4]などの，独自のマルウェア対策機能を開発して公開している．これらの中には最新のセキュリティ技術を用いた優れたものが多数ある．特に，マルウェア動的解析システムは，マルウェアの挙動を基に検知を行うため，シグニチャファイルに依存せずに，マルウェアを発見することができる．しかし，個々の機能のみでは包括的なマルウェア対策とはならず，セキュリティに関する知識が乏しい一般ユーザが有効活用するのは難しいという問題がある．

著者らは，これらのマルウェア対策機能を連携させることで，新種マルウェアの発見から駆除までの包括的なマルウェア対策を実現する「マルウェア対策ユーザサポートシステム」の研究開発を進めている．

本システムでは，マルウェア擬陽性ファイル検知機能を用いて，ユーザ PC 内から擬陽性ファイル（マルウェアである可能性があるファイル）を発見する．発見されたファイルは，既知マルウェア判定機能，マルウェア解析機能により解析される．ファイルがマルウェアと判断された場合，駆除ツール生成機能が駆除ツールを自動生成する．最後に，駆除ツールをユーザ PC 上で実行して，マルウェアの駆除を完了する．

個々の機能は包括的対策には不十分であっても，本システムを介して複数の機能が連携することで，一般ユーザを対象とした包括的なマルウェア対策を実現することが可能となる．

本論文では，システムの設計と実装について述べるとともに，現在開発中のシステムを用いて，CCC DATASet 2010[7]で指定されたマルウェア検体を解析した結果について報告する．

2. 関連研究

複数のマルウェア対策機能を組み合わせることで，高度なマルウェア対策を実現する手段に関する既存研究は数少ない[5][8]．CloudAV[5]は，一般的なアンチウイルスソフトや動的解析システムなどの複数のマルウェア検知機能を統合して検知を行うプラットフォームである．CloudAV では，ユーザ PC を監視し，アクセスが発生したファイルを解析対象ファイルとして，解析システムに送信する．解析システムでは，複数のマルウェア検知機能を用いてファイルを分析する．そして，分析結果を統合して，最終的な検知結果を求める．

しかし，CloudAV では，ファイルがマルウェアと判断された場合にも，駆除ツールは生成されないため，マルウェアに対する包括的対策を実現していない．また，擬陽性ファイルを発見する機能を有しておらず，PC 中の全ファイルが解析対象となり，解析システムやネットワークに大きな負荷がかかるという問題がある．

3. システム設計

3.1. 概要

マルウェア対策ユーザサポートシステムは様々なマルウェア対策機能を連携させることで，マルウェアの検知から駆除までの包括的なマルウェア対策を行う．図 1 に本システムの概要を示す．

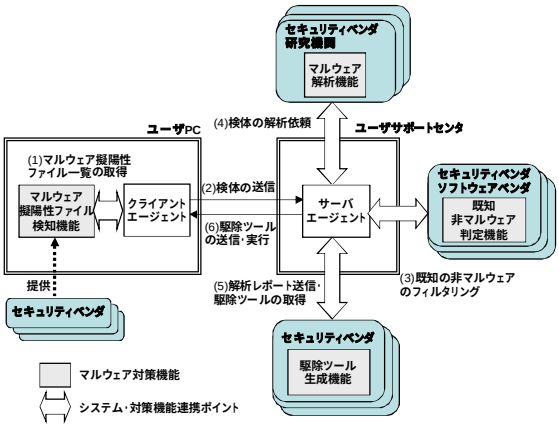

図 1　システム概要

本システムは，マルウェア対策に必要な手順を 4 種類のマルウェア対策機能（マルウェア擬陽性ファイル検知機能，既知非マルウェア判定機能，マルウェア解析機能，駆除ツール生成機能）に分割する．マルウェア擬陽性ファイル検知機能は解析対象ファイルの発見を，既知非マルウェア判定機能とマルウェア解析機能はマルウェア解析と検知を，駆除ツール生成機能はマルウェアの駆除ツールの自動生成を行う．

また，これらの機能を連携させるために，クライアントエージェント（CA），サーバエージェント（SA）という 2 つの機能を設ける．CA はユーザ PC 上で，SA はシステムを統括するユーザサポートセンタ上で動作する．

3.2. 検知・駆除の手順

マルウェア検知・駆除の手順を以下に示す．

1. 先ず，ユーザはユーザサポートセンタから CA をダウンロードして，PC にインストールする．CA はマルウェア擬陽性ファイル

検知機能を定期的に実行し，PC 内のマルウェア擬陽性ファイルの一覧を取得する．

2. CA はマルウェア擬陽性ファイルを検体として，SA に送信する．ファイルは送信前に，CA・SA 間の共通鍵または SA の公開鍵により，暗号化される．また，検体と同時に，ユーザ PC 内の環境情報を送信する．環境情報には，ユーザ PC にインストールされている OS やアプリケーションの種類やバージョンなどが含まれる．
3. 検体と環境情報を受信した SA は，検体を復号して既知非マルウェア検知機能に送信する．既知非マルウェア判定機能は，検体が既知の非マルウェアであるか否かの判定結果を出力する．
   既知非マルウェア検知機能が，検体は非マルウェアと判定した場合，SA は検体のファイル名を CA に通知する．CA は，SA から通知されたファイル名をローカルホワイトリストに保存する．ローカルホワイトリストはユーザ PC 内の既知の非マルウェアの一覧を保持する．リストに含まれるファイルは以後 SA に送信されなくなる．
4. 既知非マルウェア検知機能が，検体は非マルウェアであると判定しなかった場合，SA は検体と環境情報をマルウェア解析機能に送り，検体の解析を依頼する．マルウェア解析機能は検体を解析し，検体の挙動に関する情報と検知結果を含む解析結果レポートを出力する．
   マルウェア解析機能が，検体はマルウェアでは無いと判断した場合，検体は既知非マルウェア判定機能と CA のローカルホワイトリストに登録される．
5. マルウェア解析機能が，検体はマルウェアであると判断した場合，SA は検体と解析結果レポート，環境情報を駆除ツール生成機能に送信する．駆除ツール生成機能は検体と解析結果レポートを基に駆除ツールを生成して，SA に送る．
6. SA は駆除ツールを CA に送信する．CA は駆除ツールを実行して，マルウェアを PC から駆除する．

3.3. マルウェア対策機能の要件

3.3.1. マルウェア擬陽性ファイル検知機能

マルウェア擬陽性ファイル検知機能の主要件は，ユーザ PC 内からマルウェア擬陽性ファイルを検知することである．

本機能には，極めて低い検知見逃し率が要求される．これは，本機能により検知された擬陽性ファイルがマルウェア解析機能の解析対象となるため，見逃されたファイルは解析・駆除されることなくユーザ PC 内に存在し続けることになるためである．

その一方で，検知見逃し率に比べて誤検知率はある程度許容される．これは，誤検知された検体は，後段の既知非マルウェア判定機能やマルウェア解析機能により再度分析されるため，最終的にマルウェアであると誤判定される可能性は低いためである．

3.3.2. 既知非マルウェア判定機能

既知非マルウェア判定機能の主要件は，擬陽性ファイルが既知の非マルウェアであるか否かを判定することである．

本機能はマルウェア解析機能が解析する必要があるファイル数を減らすための前処理フィルタである．このためマルウェアを非マルウェアと誤判定する確率は，極めて低い必要がある．

3.3.3. マルウェア解析機能

マルウェア解析機能の主要件は，検体の挙動を解析し，マルウェア解析結果レポートを出力することである．解析結果レポートには，解析対象ファイルがマルウェアであるか否かの情報と，ファイルやレジストリの変更・生成・削除，プロセスの実行・停止といった，マルウェアが PC に及ぼす影響に関する情報が含まれる．

3.3.4. マルウェア駆除ツール生成機能

マルウェア駆除ツール生成機能の主要件は，マルウェア解析機能から取得した解析結果レポートと検体を基に，検体を駆除して PC の状態を復旧する駆除ツールを生成することである．

駆除ツールは，マルウェアが作成・変更したファイルやレジストリの削除・修正，実行したプロセスの停止などを行う．このため，駆除ツールに不備がある場合，誤ったファイルやレジストリを削除するなどして，PC に悪影響を与える可能性がある．

また，自己改変型マルウェアに代表されるように[10][11]，マルウェアの挙動は実行環境に依存する場合があり，駆除ツール実行後もマルウェアが残存する可能性がある．実際，アンチウイルスソフトウェアなどに付属しているマルウ

ェア駆除機能では，マルウェアの完全な駆除は難しく，マルウェアが生成したファイルや追加したレジストリが残存する場合が多いと，報告されている[9].

さらに，マルウェアによっては耐駆除機能を持つため，駆除が意図した通りに行われない可能性がある.

このため，マルウェア駆除ツール生成機能は，駆除ツールが PC に悪影響を及ぼすことなく，マルウェアを確実に駆除することを検証する仕組みを持つことが望ましい.

4. マルウェア対策機能の実装

本章では，マルウェア擬陽性ファイル検知機能，既知非マルウェア判定機能，マルウェア解析機能のプロトタイプ実装について述べる．マルウェア駆除機能については，現在，設計・開発を進めている.

4.1. マルウェア擬陽性ファイル検知機能

マルウェア擬陽性ファイル検知機能として，Microsoft Authenticode を用いた検知プログラムを開発した．Authenticode は，Windows アプリケーションに電子署名を施すことで，アプリケーション発行元の検証，及び改ざんの検知を可能とする技術である．検知プログラムはユーザ PC 内の実行ファイルをスキャンし，ファイルに Authenticode 署名が施されていて，且つそのディジタル証明書が信頼できる認証局から発行されている場合に限り，非マルウェアであると判断する．現在，正規プログラムであっても Authenticode 署名が施されていないファイルも多いため，誤検知率は高い.

しかし，有効な Authenticode 署名が施されているマルウェアは著者らが知る限りでは無い．このため，検知見逃し率は低くマルウェア擬陽性ファイル検知機能に求められる性能を満たしていると言える．以降，本検知プログラムを，署名型検査プログラムと呼ぶ.

4.2. 既知非マルウェア判定機能

既知非マルウェア検知機能として，ホワイトリストを利用した判定プログラムを開発した．本プログラムは，信頼されているセキュリティ機関に非マルウェアであると判断されたファイルのハッシュ値のリストであるホワイトリストを用いる.

本プログラムは，ホワイトリストに載っている実行ファイルのみを，非マルウェアと判断する．この方法では，マルウェアを非マルウェアと誤判定する可能性が低く，既知非マルウェア判定機能に求められる性能を満たしている.

実装では約 300 万件の正規 Windows プログラムのハッシュ値を含むホワイトリストを用いた．本実装をホワイトリストサーバと呼ぶ.

4.3. マルウェア解析機能

マルウェア解析機能としては，既存のマルウェア動的解析システムの1つである，nicter ミクロ解析システム[2]（以下，nicter）を用いる.

nicter は，解析対象プログラムを解析環境内で数分間実行し，その挙動を記録する．そして，記録された挙動を基にプログラムがマルウェアであるかどうかを判定する．最後に，判定結果とマルウェアの挙動を XML 形式で示した解析結果レポートを出力する.

解析環境は，マルウェアが実行される実マシンと，マルウェアのネットワーク活動を再現するためのインターネットエミュレータ（仮想インターネット）から構成される.

マルウェアのマシン内での挙動は，API フッキングにより記録する．nicter は，仮想マシンではなく実マシン上でマルウェアを実行するため，仮想マシンを認識して挙動を変えるマルウェアに対しても対応することができる.

インターネットエミュレータ内には DNS，HTTP，SMTP，FTP，IRC などの主要なネットワークサービスを行うダミーサーバがある．マルウェアはエミュレータをインターネットと誤認識して，ネットワーク活動を行う．このため，マルウェアのパケットを実際のインターネットに送出することなく，ネットワーク活動を観測することができる.

nicter の解析では，先ず検体を実マシン上で 60 秒間実行する．検体が OS 起動時に自動実行するように，実行ファイルのパスをレジストリの RUN キーを登録を行った場合は，解析環境を再起動してさらに 60 秒間検体を実行する．実行終了後に，API ログやパケットログを基に検体がマルウェアであるかどうかを判定し，解析結果レポートに出力する.

5. CCC DATASet 2010 検体の解析

マルウェア対策ユーザサポートシステムのプロトタイプ実装を用いて，CCC DATASet 2010 で指定された検体を解析した．マルウェア駆除ツール生成機能は現在設計・開発中であるた

め，実験では，ユーザ PC で検体を発見してから nicter での解析が完了するまでを評価した．

5.1. 実験環境

実験は，ユーザ PC，SA，ホワイトリストサーバ，nicter が接続されている 100 Mbps LAN 上で行った．ユーザ PC は VMWare Workstation 上の VM であり，OS は Windows XP SP3 である．SA とホワイトリストサーバは VMWare vSphere 上の VM で動作する．OS はそれぞれ CentOS5.4，Windows 2003 Server である．

検体としては，CCC DATASet 2010 で指定された検体中の 20 個を用いた．実験では，検体を1つずつユーザ PC 上にコピーした状態で，署名型検査プログラムを実行した．

5.2. 実験結果

何れの検体にも電子署名が付与されていないため，署名型検査プログラムは全ての検体を，擬陽性ファイルとして検知した．同様に，何れのファイルのハッシュ値もホワイトリストサーバには登録されていないため，非マルウェアと誤判定される検体はなく，最終的に全検体が nicter に送信され，解析された．

20 検体中，RUN キーに登録するものは 12 検体あり，多くの検体が OS 再起動後に実行を継続する仕組みを有していることがわかった．

5.2.1. 検体の処理時間

表 1 に検体の処理時間を示す．値は 20 検体における平均値である．

表 1　検体の処理時間の平均値

項目		所要時間
署名型検査プログラムの実行		0.08 秒
ローカルホワイトリストによる判定		2.17 秒
検体の CA から SA への送信		2.00 秒
ホワイトリストサーバによる判定		4.57 秒
検体の nicter への登録		2.00 秒
nicter による解析処理	RUN キー登録有り	550.05 秒
	RUN キー登録無し	239.88 秒
総時間	RUN キー登録有り	561.38 秒
	RUN キー登録無し	250.20 秒

署名型検査プログラムがマルウェアの検査を開始してから，nicter による解析が完了するまでの時間は，RUN キーに自身を登録する検体では約 9 分，登録しない検体では約 4 分かかり，その大半が nicter での解析処理時間となった．前節で述べたとおり，RUN キーに登録する検体の場合，nicter の解析環境を再起動し再度検体を実行するため，解析時間が余分にかかることになる．

5.2.2. nicter による検体解析結果

nicter による検体解析結果を表 2 に示す．

表 2 nicter による検体解析結果

番号	活動1	活動2	活動3	分類[種類/名称.亜種]
#1	IRC通信	自身のコピー (c:\windows)		BOT/A.a
#2	IRC通信	自身のコピー (システム)		BOT/A.b
#3	自身のコピー (ゴミ箱)			WORM/B.c
#4	自身のコピー (ゴミ箱)			WORM/B.d
#5	自身のコピー (ゴミ箱)			WORM/B.e
#6	IRC通信	自身のコピー (システム)		BACKDOOR/C.f
#7	IRC通信	自身のコピー (システム)		BACKDOOR/C.g
#8	IRC通信	自身のコピー (システム)		UNKNOWN
#9	プロクシ無効			DROPPER/D.h
#10	TCP リスニング			DROPPER/E.i
#11	IRC通信			DROPPER/F.j
#12	MS Phonebook の検索	プロクシ無効	自身のコピー (システム)	DROPPER/G.k
#13	IRC通信	自身のコピー (システム)		TROJAN/H.l
#14	自身のコピー (システム)			TROJAN/I.m
#15	MS Phonebook の検索	プロクシ無効		TROJAN/J.n
#16	MS Phonebook の検索	プロクシ無効		TROJAN/K.o
#17	プロクシ無効			SPAM/L.p
#18	MS Phonebook の検索	プロクシ無効		DOWNLOADER/M.q
#19	TCPリスニング			SPYWARE/N.r
#20	MS Phonebook の検索	プロクシ無効	TCP リスニング	WORM/O.s

「活動 1〜3」は，nicter が，検体をマルウェアと判断した根拠となる活動であり，解析結果レポート中に示される．リスト 1 に例を示す．この例では，マルウェアはレジストリを書き変えて，プロクシサーバの設定を無効化している．変更は，プロクシサーバ上での検知を逃れるためなどにされていると考えられる．このため，このような活動を行う実行ファイルは，マルウェアであると判断される．

「分類」は，市販のアンチウイルスソフトによるものである．#1 と#2，#3〜#5，及び#6 と#7 のように，種類･名称が同じマルウェアは類似した活動を行うことがわかる．

多くの検体が，ゴミ箱（C:¥RECYCLER）やシステムファイル，WINDOWS ファイルに自身のコピーを作成している．これは，ゴミ箱に隠れたり，システムファイルを装うことで，駆除を逃れようとするためと，判断できる．

また，#12，#15，#16，#18，#20 では，MS

Phonebook の検索を行っている. これは, ダイアルアップ等のリモート接続の設定ファイル中の接続先ネットワークを, 攻撃者のネットワークに変更しようとする意図があると, 判断できる.

リスト 1 解析結果レポートの例

```
<addReg>
<classification type="WORM" />
  <content>
    <registrykey>
      <![CDATA[ HKEY_CURRENT_CONFIG¥Software
      ¥Microsoft¥windows¥CurrentVersion¥Internet
      Settings]]>
    </registrykey>
    <registryname>
      <![CDATA[ ProxyEnable]]>
    </registryname>
    <registryvalue>
      <![CDATA[ 0]]>
    </registryvalue>
  </content>
</addReg>
```

6. おわりに

本稿では, 様々な組織が開発したマルウェア対策機能を連携させることで, 包括的なマルウェア対策を行う, マルウェア対策ユーザサポートシステムの設計について述べた.

本システムは, 擬陽性ファイルの発見からマルウェア駆除までを行うことで, 一般ユーザを対象とした包括的なマルウェア対策を実現することが可能になる.

現在, 我々は, システムのプロトタイプ実装を進めている. 本論文では, CCC DATASet 2010 で示された検体を用い, 擬陽性ファイルの発見から nicter 上での解析までを 4 分～9 分で完了することを示した.

今後は, 駆除ツール生成機能の設計, プロトタイプ実装を進める. また, 署名型検査プログラムを改良し, 誤検出率がより低いマルウェア擬陽性ファイル検知機能を開発する. 具体的には, マルウェアが頻繁に利用する API の組み合わせや, 固有なファイル構造などに着目することを検討している.

最終的に, 擬陽性ファイルの発見からマルウェア駆除までを 10 分以内で完了するシステムを実現することを目標としている.

謝辞

本研究は独立行政法人情報通信研究機構から委託を受けた「マルウェア対策ユーザサポートシステムの研究開発」の成果の一部を含みます. 本研究を進めるにあたって有益な助言と協力を頂いた関係者各位に深く感謝いたします.

参考文献

[1] C. Willems, et al.,“Toward Automated Dynamic Malware Analysis Using CWSandbox,” IEEE Security and Privacy Magazine, Vol.5, Issue 2, 2007.

[2] D. Inoue, et al., “Automated Malware Analysis System and its Sandbox for Revealing Malware's Internal and External Activities,” IEICE Trans. Information and Systems, Vol.E92-D, No.5, 2009.

[3] Anubis: Analyzing Unknown Binaries, http://anubis.iseclab.org/, accessed at 08/09/2010.

[4] National Software Reference Library, http://nsrl.nist.gov/, accessed at 08/09/2010.

[5] J. Oberheide, et al., “CloudAV: N-Version Antivirus in the Network Cloud, ” In Proc. of the 17th Usenix Security Symposium, July, 2008.

[6] Norman Solutions. Normand sandbox whitepaper, http://download.norman.no/whitepapers/whitepaper_Norman_SandBox.pdf, accessed at 08/09/2010.

[7] 畑田 他, “マルウェア対策のための研究用データセット～MWS 2010 Datasets～,” MWS2010 予稿集, 2010.

[8] Virustotal, http://www.virustotal.com, accessed at 08/09/2010.

[9] E. Passeriniy, et al., “How good are malware detectors at remediating infected systems?,” In Proc. of DIMVA’09, 2009.

[10] C. Kruegel, et al., “Polymorphic worm detection using structural information of executables,” In Proc. of RAID’05, 2005.

[11] G.Wicherski, “peHash: A Novel Approach to Fast Malware Clustering,” In Proc. of USENIX LEET’09 2009.