マルウェア対策のための研究用データセット
～MWS Datasets 2013～

神薗 雅紀†1†2　畑田 充弘†3 寺田 真敏†4　秋山 満昭†5　笠間 貴弘†1　村上 純一†6

†1 独立行政法人 情報通信研究機構 184-8795 東京都小金井市貫井北町 4-2-1
†2 株式会社セキュアブレイン 先端技術研究所 102-0083 東京都千代田区麹町 2-6-7 麹町 RK ビル 4F
†3 エヌ・ティ・ティ・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
†4 株式会社日立製作所 212-8567 神奈川県川崎市幸区鹿島田 1-1-2 新川崎三井ビル
†5 NTT セキュアプラットフォーム研究所 180-8585 東京都武蔵野市緑町 3-9-11
†6 株式会社 FFRI 150-0013 東京都渋谷区恵比寿 1-18-18 東急不動産恵比寿ビル 4 階
Email: †1masaki_kamizono@nict.go.jp, †2masaki_kamizono@securebrain.co.jp,
†3m.hatada@ntt.com, †4 masato.terada.rd@hitachi.com,
†5akiyama.mitsuaki@lab.ntt.co.jp, †1kasama@nict.go.jp, †6murakami@ffri.jp

あらまし マルウェアの脅威に対し様々な対策研究が盛んに行われているが，近年の脅威は攻撃の多様化や高度化により，研究を進める上で基礎となる「研究素材の収集と共有」が非常に困難な状況になってきている．このような課題に対し，必要となる情報を収集して客観的な評価と研究成果の共有を容易にするための研究用データセット（MWS Datasets 2013）を作成した．本稿では，MWS Datasets 2013を構成するCCC Dataset 2013, D3M 2013, FFRI Dataset 2013, PRACTICE Dataset 2013, NICTER Darknet Dataset 2013の概要を紹介する．

Datasets for Anti-Malware Research
~ MWS Datasets 2013 ~

†1†2Masaki kamizono, †3Mitsuhiro Hatada, †4Masato Terada,
†5Mitsuaki Akiyama, †1Takahiro Kasama, †6Jyunichi Murakami

†1 National Institute of Information and Communications Technology
4-2-1 Nukui-Kitamachi, Koganei, Tokyo, 184-8795 JAPAN
†2 Advanced Research Laboratory,Securebrain Corporation
Kojimachi RK Bldg,6-7 Kojimachi 2-chome,Chiyodaku, Tokyo, 102-0083 Japan
†3 NTT Communications Corporation
Gran Park Tower 16F, 3-4-1 Shibaura, Minato-ku, Tokyo 108-8118, Japan
†4 Hitachi, Ltd. 890 Kashimada, Saiwai-ku, Kawasaki, Kanagawa 212-8567
†5 NTT Secure Platform Laboratories
3-9-11 Midori-cho, Musashino-shi, Tokyo, 180-8585, JAPAN
†6 FFRI,Inc. 4F EBISU TOKYU Bld., 1-18-18 Ebisu,Shibuya-ku,Tokyo,Japan

Abstract There has been a lot of researches on countermeasures against the threats by malware. By diversification and the advancement of the recent attack, collection and sharing the data which are necessary to push forward a study become very difficult. For such a problem, anti-Malware engineering WorkShop (MWS) collected the data which were necessary to push forward a study and made data set (MWS Dataset 2013) for studies to evaluate the proposals objectively and share the research achievements. This paper presents an overview of MWS 2013 Datasets which comprised of CCC Dataset 2013,D3M 2013, FFRI Dataset 2013, PRACTICE Dataset 2013,NICTER Darknet Dataset 2013.

1　はじめに

標的型攻撃を始め，高度かつ複雑化したサイバー攻撃が国際的な問題となっており，国家および企業レベルでの対策が急務となっている．このような背景を踏まえ，マルウェア対策やそこから派生する様々な研究が盛んに行われているが，研究を進める上で様々な課題が浮き彫りとなってきている．そのうちの一つに「共通の研究素材がないこと」が挙げられる．共通の研究素材とは，研究技術の評価に用いるマルウェアのサンプルや，感染前後の通信データなどのことを指し，サイバー攻撃を可能な限り網羅的に，かつ攻撃の進化に合わせて適切に選択されたものが望ましい．しかし，研究素材となるこのようなデータは，主に研究者がハニーポットなどを構築して自ら作成し，各々の提案手法や対策手法の有効性や妥当性を評価してきた．このため，同じ研究テーマに取り組んだ場合でも，研究素材が異なるために，その研究を相互に比較することが困難であった．

また，もう一つの課題として「研究素材そのものを収集すること自体が難しくなってきている」ことである．例えば，Drive-by-Download 攻撃における攻撃通信や最終的にダウンロードおよび実行されるマルウェアを研究素材として収集する例を考えてみる．近年の Drive-by-Download 攻撃サイトは様々な解析および検知回避機能等を有しており，情報を収集する環境によって期待した情報が得ることができず，定性的にも定量的にも研究素材としての収集が難しくなってきている．また，BOT の C&C サーバとの通信を収集する場合においても，近年の C&C サーバは短時間で活動を停止するため，期待した通信データを収集することが困難である．なお，研究用データを収集することが困難となってきている傾向は，マルウェアを含むサイバー攻撃による脅威全般に当てはまる．

このような課題がある中，高度かつ複雑化したサイバー攻撃に対峙していくため，(1)サイバークリーンセンター運営連絡会が運用しているハニーポットで収集したデータを活用した研究用データセット:CCC Dataset 2013，(2)研究者コミュニティが収集した Web 感染型マルウェアの観測データ:D3M 2013，(3)マルウェア動的解析データ:FFRI Dataset 2013，(4)総務省「国際連携によるサイバー攻撃予知・即応に関する実証実験」プロジェクトで得られたマルウェア長期観測データ:PRACTICE Dataset 2013，(5)独立行政法人 情報通信研究機構が運用する NICTER にて観測したダークネットパケットデータ:NICTER Darknet Dataset 2013 を含む，MWS Datasets 2013 を提供し，研究成果を共有する場として「マルウェア対策研究人材育成ワークショップ(MWS2013)」を開催する．

本稿では，2 章にて関連研究として他のデータセットを紹介し，3 章以降に MWS Datasets 2013 の概要を述べる．なお，CCC Dataset の 2008～2012 年分に関しては既に畑田ら[1][2][3]が報告しているため本稿では概要を省略する．2013 年分については，サーバ型ハニーポット 12 台により 2013 年 2 月 1 日～2013 年 2 月 28 日にかけて収集したユニークな 7,028 個のマルウェア検体のハッシュ値(SHA1)となっている．最後に，7章で過去の MWS Datasets の概要と利用状況について報告する．

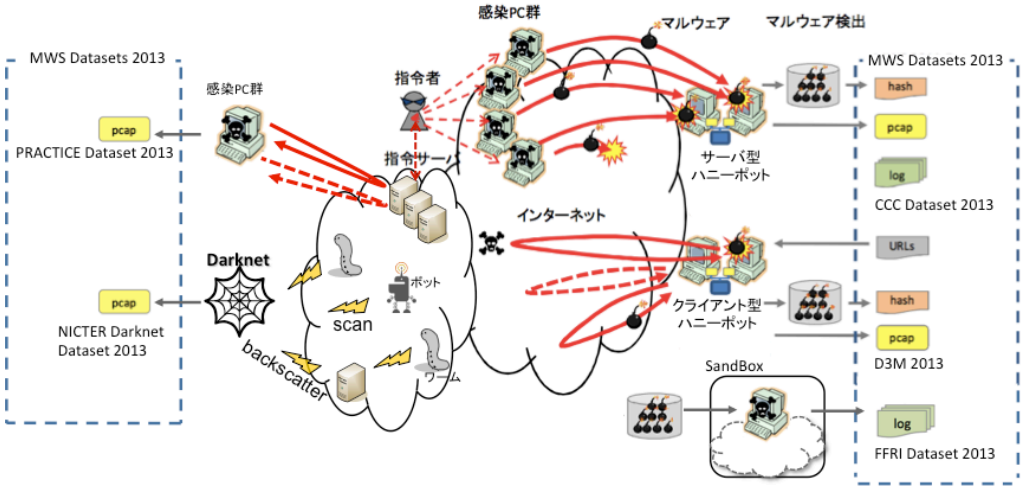

図 1 MWS Datasets 2013 概要

2　関連データセット

研究用データセットとしては，現在でも侵入検知システムの評価に用いられる DARPA Intrusion Detection Evaluation Data Sets[4]が挙げられる．これは 2000 年のデータセットが最新であり，近年のサイバー攻撃を考慮しているとは言いがたい．また，the 2009 Inter-Service Academy Cyber Defense Exercise datasets[5]は，サイバー防御演習時のデータセットでありマルウェアによる攻撃を想定したものではない．一方で，大規模セキュリティ関連データの収集と分析をもとに，より良いデータとナレッジの共有を図る BADGERS2011[6]や，コンピュータ・ネットワークの運用データをレポジトリとして蓄積し，インフラ防護と脅威評価に活用する PREDICT[7]，広域ネットワークの情報・状況を分析し，幾つかのタイプのデータセットを提供する CAIDA[8]などのプロジェクトもある．様々なデータセットが提供されているが，MWS Datasets は，図 1 に示すようにサーバ型ハニーポットやクライアント型ハニーポットにて収集した情報，マルウェア動的解析情報，ダークネットパケット情報など，マルウェア対策を含むサイバー攻撃対策を研究する上で必要となる情報を可能な限り網羅的に，かつ攻撃の進化に合わせて適切に選択したデータセットとなっている点が，他のデータセットと比較して優れていると言える．次章以降，各データセットの概要を示す．

3　D3M 2013

D3M（Drive-by-Download Data by Marionette） 2013 は，NTT セキュアプラットフォーム研究所の高対話型の Web クライアント型ハニーポット(Marionette[9])で収集した攻撃通信データ，マルウェア検体，マルウェア検体の通信データの 3 つを収録した Web 感染型マルウェアの観測データ群である．Marionette は脆弱性に対する攻撃を受けるが，ダウンロードされたマルウェアの実行を許可しない．取得したマルウェアは動的解析システム（BotnetWatcher[10]）にて解析される．

CCC Dataset はいわゆるサーバ型ハニーポットで収集したデータである．一方，D3M は近年脅威となっている Web ブラウザの脆弱性を利用して制御を奪い，マルウェアを強制的にダウンロード及びインストールさせる Drive-by-download 攻撃を捉えた研究用データセットとなっている．D3M 2013 は，感染手法の検知ならびに解析技術の研究のための「攻撃通信データ」，マルウェアの解析技術の研究のための「マルウェア検体」および「マルウェア通信データ」から構成される．データセットの取得環境を図 2に示す．以下，それぞれのデータについて概要を述べる ．

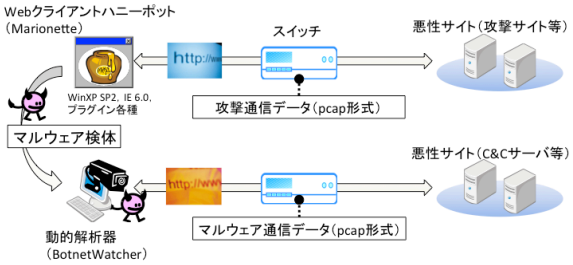

図 2 研究用データセット D3M 取得環境

3.1　攻撃通信データ

Web クライアントハニーポットの通信を tcpdump でパケットキャプチャした PCAP 形式のファイルである．ハニーポットの OS は WindowsXP SP2，ブラウザは Internet Explorer 6.0，プラグインが Adobe Reader，Flash Player，WinZip，QuickTime，Java であり，何れもセキュリティパッチは未適用である．ハニーポットはインターネット接続されており，パケットキャプチャは上流ネットワークにあるスイッチのミラーポートで行っている．データ収集日は 2012 年 8 月 2 日，2012 年 10 月 2 日，2013 年 2 月 26 日であり，日毎に 1 ファイル，計 3 ファイルである．巡回対象 URL は公開されているブラックリスト[11]に登録されている URL の中から，各データ収集日に攻撃を検知した URL をあらかじめ抽出したものを用いており，参考情報として D3M 2013 とともに提供している．各収集日においてアクセスした URL は同一とは限らず，また，入力 URL から派生する URL（リダイレクト，スクリプト読み込み，画像読み込みなど）は記載されていない．CCC Dataset と同様に D3M においても，過去のデータとの傾向を比較分析することができるよう D3M 2011 および D3M 2012 も提供している．

3.2　マルウェア検体

Web クライアントハニーポットで収集した Web

感染型マルウェアのハッシュ値をテキスト形式で記載したファイルである. 3.1節で収集した検体であり, 攻撃通信データに含まれる検体である.

3.3 マルウェア通信データ

3.2節のマルウェア検体を取得から 24 時間以内に動的解析器で実行した際の通信のフルキャプチャデータである. 動的解析器はインターネットに接続した環境でマルウェアを動作させており, ボットなどの遠隔制御されるマルウェアの動的解析が可能である. なお, 外部ホストやネットワークに対する攻撃は BotnetWacher 内の仮想インターネット環境に転送することで, 解析時の安全性を実現している.

4 FFRI Dataset 2013

FFRI Dataset 2013 は, (株)FFRI が独自に収集した計 2644 件のマルウェアを動的解析し得られたマルウェアの解析ログ群である. CCC Dataset および D3M は攻撃通信データやマルウェアの通信データ, マルウェア検体そのものをデータセットとしているが, FFRI Dataset はマルウェアの端末内での挙動に着目する. データセットの仕様について, 以下に概要を述べる.

4.1 マルウェア

マルウェアは全て PE(Portable Executable)形式かつ, Windows プラットフォーム上にて実行可能な実行形式ファイルである. 内訳は以下の通りである.

(1) 2012 年 9 月から 2013 年 3 月の期間に収集されたマルウェアをランダムサンプリングし, 抽出されたマルウェア:2461 件

(2) MITB(Man in the Browser)攻撃の機能を有した SpyEye[12]マルウェア:1 件

(3) PC 遠隔操作事件[13]で利用された遠隔操作マルウェア:1 件

(4) 韓国サイバーテロ[14]において利用された MBR 破壊マルウェア:1 件

(1)は Web クローリング等において広く世界中から収集された比較的新しいマルウェアであり, 少なくとも 1 社以上のアンチウイルス製品にてマルウェアであると判定されることを確認している. 収集された全数からランダムサンプリングを行っており, その内訳は収集時点におけるインターネット上のマルウェアの感染トレンドを反映している. 当該マルウェア検体を利用した評価により評価手法の現実的な有効性を確認することを目的として選定されている. (2)から(4)については国内外にて発生した事件において利用されたマルウェア, ないしはその亜種である. 具体的には, (2)は 2012 年 10 月末に実際に発生した不正送金事件において利用されたマルウェアの亜種であり, (3)及び(4)は, 実際の事件において利用されたマルウェアの一部である. 実際に社会的に広く認知された事件において利用されたマルウェアの一例として, 評価に利用されることを目的に選定されている. なお, データセットはこれらマルウェアの動的解析の結果であり, 当該マルウェア自体は含まない.

4.2 動的解析

前述の各マルウェアをオープンソースのマルウェア解析ツールである Cuckoo Sandbox[15]を用いて動的解析し, 解析ログを生成している. Cuckoo Sandbox は, 仮想化された Windows ゲスト内にマルウェアをコピー, 実行, 実行時挙動の記録, ゲスト環境の復元と言った一連の解析作業を自動化するソフトウェアパッケージである. マルウェアの動的解析は, ネットワーク接続を有する専用のマルウェア解析環境上に Cuckoo Sandbox による解析システムを構築し, 1 マルウェア当り 90 秒間実行したものとなっている. また, Cuckoo Sandbox は, VirusTotal[16]と連携する機能を有しており, 解析対象ファイルのハッシュ値に基づいて VirusTotal に問い合わせを行うことで, 当該時点での各アンチウイルス製品での検出状況を取得することができる. 本データセットの解析ログは, 解析を実施した 2013 年 3 月時点での当該検出状況を含んでいる. 表 1に解析ログに含まれる具体的な項目の概要をまとめる.

表 1 解析ログに含まれるデータ項目

項目(大見出し)	概要
info	解析の開始, 終了時刻等
yara	yara[17]の有する標準ルールセットとの照合結果
signatures	ユーザ定義シグニチャとの照合結果(未使用)
virustotal	VirusTotal に登録されている各ア

	ンチウイルス製品の検出結果
static	マルウェアファイルの静的情報(セクション構造, インポート API 等)
dropped	マルウェアが実行時に生成したファイルに関する情報
behavior	マルウェアが実行時に呼び出した API, 引数, 返り値等の情報
processstree	マルウェアが実行時に起動したプロセスの階層情報
summary	マルウェアが実行時にアクセスしたファイル, レジストリキー等の情報
target	解析対象となったマルウェアファイルの情報(ファイルサイズ, ハッシュ値等)
debug	動的解析時の Cuckoo Sandbox のデバッグログ
strings	マルウェアファイルに含まれる文字列情報
network	マルウェアが実行時に発生した通信情報

5 PRACTICE Dataset 2013

PRACTICE Dataset 2013 は総務省「国際連携によるサイバー攻撃予知・即応に関する実証実験」(略称:PRACTICE)の挙動観察システムで, 5 検体のマルウェアを長期観測(最大 1 週間程度)した際の通信トラフィック(マルウェア感染後の通信挙動)を含むデータセットである. 本データセットは, マルウェア感染後の通信挙動や通信先の分析, 検知手法の評価, 解析環境の課題提起に利用することを想定している. データセットは検体そのものを含まず, 検体情報と通信トラフィックデータで構成され, 以下に概要を述べる.

5.1 検体情報

検体情報には検体のハッシュ値(SHA1), 取得した通信トラフィックデータ(PCAP 形式)のファイル名, 解析環境の IP アドレス, 4 種類のアンチウイルス製品での検知結果(各検体収集時点で最新のパターンファイルを利用), 解析開始時刻・終了時刻, ファイルサイズを含む(表 2).

表 2 検体情報の例

項目	例
ハッシュ値(SHA1)	5b9f7***
PCAP ファイル名	practice_1.pcap
IP アドレス	10.220.0.36
検知結果	Kaspersky: 未検知 McAfee: PWS-Zbot.gen.alu Symantec: 未検知 TrendMicro: 未検知
解析時刻	2013-05-18 02:35:06 ～2013-05-25 11:59:53
ファイルサイズ	10MB

5.2 通信トラフィックデータ

通信トラフィックデータを取得している各検体の動的解析環境は BotnetWatcher[10]であり, データには, インターネットへの疎通確認や時刻同期など管理用通信も一部含んでいる. また IP アドレス, デフォルト GW, DNS サーバは DHCP で割り当てている.

TCP/UDP ともにランダムあるいは一定の high portを利用したP2P通信を行っている検体や, 繰り返し IRC 接続を試みる検体, 何らかのファイルをダウンロードしている検体など, 特徴的な通信挙動が確認できる. しかしながら, 接続先の名前解決ができないものもあり, 検体の収集直後からの解析が望ましい.

6 NICTER Darknet Dataset 2013

NICTER Darknet Dataset 2013 は, 3章で述べた D3M 2013 や5章で述べた PRACTICE Dataset 2013 と同様に攻撃通信データを提供する. 他のデータセットと大きく異なる点は, 6.1節にて述べるとおりダークネットと呼ばれるインターネット上で到達可能かつ未使用の IP アドレス空間に届いた通信データという点である.

6.1 ダークネット

ダークネットとは, インターネット上で到達可能かつ未使用の IP アドレス空間のことを指す. 独

立行政法人情報通信研究機構では，インターネット上におけるセキュリティインシデントの迅速な状況把握・原因究明・対策導出を目的としたインシデント分析センタ NICTER（Network Incident analysis Center for Tactical Emergency Response）[18][19][20]の研究開発を推進しており，約 21 万 IP アドレスのダークネットに届くパケットを常時観測・分析している．今回は MWS Datasets 2013 として，そのダークネットの一部で観測されたトラフィックデータを提供する．

ダークネットに届くパケットの多くはネットワークを経由して感染を広げるタイプのマルウェアによるスキャンや，マルウェア自身がペイロードに含まれている UDP パケット，マルウェア同士が P2P ネットワークを確立するためのランデブー用のパケット，送信元 IP アドレスを詐称した DDoS 攻撃を受けているサーバからの応答（SYN-ACK）であるバックスキャッタなど，何らかの不正な活動に起因している．そのため，ダークネットに届くパケットを分析することで，インターネット上で発生している不正な活動の傾向把握が可能になる．典型的な例として 2008 年に大規模感染が発生した Conficker[21]の例を示す．図 3では，445/TCP に対するパケットのユニーク送信元 IP アドレス数（攻撃ホスト数）の推移を示しているが，図を見ると，Conficker が発生したタイミングで 2 つの観測地点において，攻撃ホスト数が同時に急増していることがわかる．これは Conficker が上記ポートにおける Server サービスの脆弱性（MS08-067[22]）を悪用して感染ホスト数を急激に拡大している動きが，ダークネット観測でも捉えられていたことを示している．

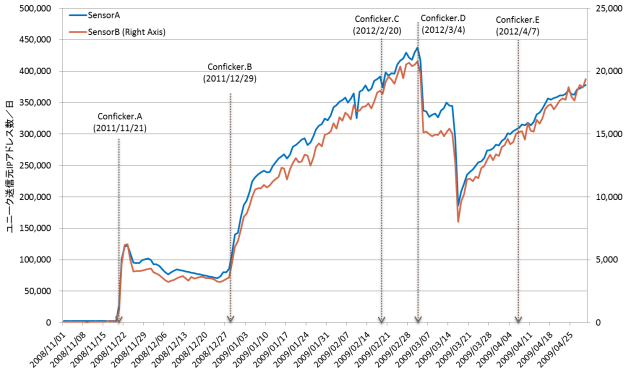

図 3 Conficker 発生時の攻撃元ホスト数（445/TCP）の推移

6.2 NONSTOP

今回の NICTER Darknet Dataset 2013 の提供には，新しい取り組みとして NICTER で開発した NONSTOP（NICTER Open Network Security Test-Out Platform）[22]を活用している．NONSTOP は各種サイバーセキュリティ情報（ダークネットトラフィック，マルウェア検体，スパムメール，マルウェア解析結果など）を外部から安全に利活用するためのプラットフォームであり，いわゆる PaaS（Platform as a Service）の形態として開発が進められている．

利用を希望するユーザは SSH クライアントとあらかじめ発行された認証用 IC カードを利用して NONSTOP へのアクセスを行い，研究内容に応じて提供される仮想マシン内で必要なサイバーセキュリティ情報にアクセスし分析を行うことになる．そのため，分析用に独自開発したツール等はローカルから仮想マシン内へファイル転送することで仮想マシン内での実行が可能となっているほか，NONSTOP 内にリポジトリを用意することで，必要な各種ライブラリ等についてもインストール可能としている．一方，仮想マシンからローカルへのファイル転送に関しては，提供したサイバーセキュリティ情報のうち外部への転送を禁止している情報の流出を防ぐ目的で，複数のフィルタ機構による検査，転送ファイルの一定期間の保存などが行われている．

6.3 NICTER Darknet Dataset 2013

NICTER Darknet Dataset 2013 では，ある固定の/20（4096 IP アドレス）のダークネットに届いたトラフィックデータ（PCAP ファイル）を提供している．本データにおいてはダークネットへ届いたパケットに対して応答は行っていないため，データセットには外部からダークネットに対するパケットしか含まれていない．また，観測地点を秘匿する目的で，データセットの宛先 IP アドレスの第 1 および第 2 オクテットは適当な値に置換している．観測期間については 2011 年 4 月 1 日～2013 年 3 月 31 日の 2 年間分を基本とし，2013 年 4 月以降のトラフィックデータに関しても順次提供を行っている．

参考までに，図 4に今回提供したダークネット

トラフィックデータの 1 日単位で観測されたパケット数とユニーク送信元 IP アドレス数(攻撃ホスト数)の前後 3 日間の移動平均を示す. 図を見ると, パケット数, 攻撃ホスト数ともにこの 2 年間で増加傾向にあることがわかる. ちなみに, 2011 年 11 月のパケット数の急激なピークは, 中国の 1 ホストからの大量のバックスキャッタ(SYN-ACK)が観測されたことが原因である.

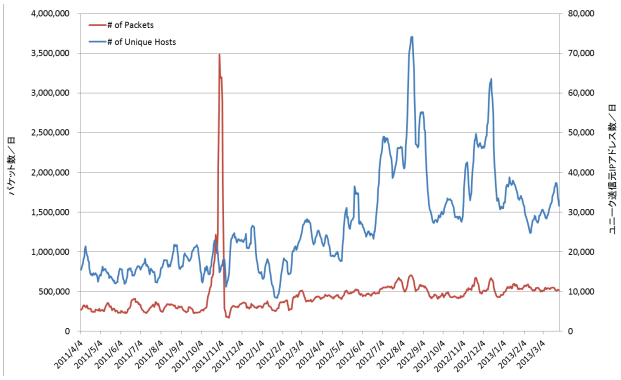

図 4 NICTER Darknet Dataset における観測パケット数・攻撃元ホスト数の推移

今回の MWS Datasets 2013 への提供では, 全体のダークネット観測網のうち/20 という限られたデータのみの提供となっているが, その中には NICTER が把握しているものだけでも, 2011 年 8 月の Morto[24]発生時の 3389/TCP に対するパケット数および攻撃元ホスト数の増加や, Carna ボットネットによるスキャン[25]など, 研究対象として非常に興味深いデータが含まれており, 本データセットがマルウェア対策研究の促進に役立つことを期待している.

7 MWS Datasets 利用状況

MWS Datasets を利用し, 研究成果を共有する場として「マルウェア対策研究人材育成ワークショップ」を 2008 年から開催している. 過去の MWS Datasets とその利用内訳を表 3 に示す. データセットにより発表件数に偏りはあるものの, 多くの方に MWS Datasets が利用されていることが分かる. また, 表 3 は「マルウェア対策研究人材育成ワークショップ」のみの発表件数を纏めており, それ意外にも MWS Datasets を利用した多数の論文が投稿されている. 2013 年度においては 25 件の発表が予定されており、内 10 件が学生による発表となっている. なお, 利用したデータセットの内訳は未定としている.

表 3 MWS Datasets 利用状況

MWS Datasets		2008	2009	2010	2011	2012	2013
CCC Dataset	マルウェア検体	5	7	6	5	7	未定
	攻撃通信データ	9	14	5	6	2	
	攻撃元データ	8	6	5	4		
MARS Dataset				1	1		
D3M				4	3	3	未定
IIJ MITF Dataset						1	
FFRI Dataset							未定
PRACTICE Dataset							未定
NICTER Darknet Dataset							未定
全部(データセット説明)			1[1]	1[2]	1[3]		1(本稿)
合計 ()内は学生発表の件数		22 (8)	28 (15)	22 (10)	20 (9)	13 (9)	25 (10)

8 おわりに

本稿では, 最新のサイバー攻撃に対峙するためのマルウェア対策や, そこから派生する様々な研究を進めるための客観的な評価基準となる研究用データセット MWS Datasets 2013 について述べた.

研究用データセット自身が研究者間での共通言語としての役割を担うことや,研究用データセットを用いて研究開発した技術等の共有により, 人材育成を含む本研究分野の発展に寄与することが期待できる.今後は, 最新の脅威を捉えた研究用データセットの作成ならびに利用環境の構築・提供など, 包括的なフレームワークを検討するとともに, 評価用として利用可能なより良い研究用標準データの作成に向け検討していきたい.

謝辞

本研究にあたって, 有益な助言とデータセット作成の協力を頂いた CCC 運営連絡会の関係者各位, 研究者コミュニティならびに総務省実証実験プロジェクトの関係者各位に深く感謝致します.

参考文献

[1] 畑田充弘, 他: マルウェア対策のための研究用データセットとワークショップを通じた研究成果の共有, CSS2009(MWS2009) (2009.10)

[2] 畑田充弘, 他: マルウェア対策のための研究 用データセット ～MWS 2010 DATA SETS～, CSS2010(MWS2010) (2010.10)

[3] 畑田充弘, 他: マルウェア対策のための研究 用データセット ～MWS 2011 DATA SETS～, CSS2011(MWS2011) (2011.10)
[4] MIT Lincoln Laboratory, DARPA Intrusion D etection Evaluation Data Sets, http://www.ll.mit. edu/mission/communications/ist/corpora/ideval/i ndex.html
[5] B. Sangster, et al.: Toward Instrumenting N etwork Warfare Competitions to Generate Label ed Datasets, 18th USENIX Security Symposium CSET'09 (2009.08)
[6] BADGERS2011: Building Analysis Datasets and Gathering Experience Returns for Security, http://iseclab.org/badgers2011/ (2011.04)
[7] PREDICT: the Protected Repository for the Defense of Infrastructure Against Cyber Threats, https://www.predict.org/
[8] CAIDA: The Cooperative Association for Internet Data Analysis http://www.caida.org/home/
[9] Mitsuaki Akiyama, et al: Design and Implementation of High Interaction Client Honeypot for Drive-by-download Attacks, IEICE Transactions on Communication, Vol.E93-B No.5 pp.1131-1139 (2010.05)
[10]Kazufumi Aoki, et al: Controlling Malware HTTP Communications in Dynamic Analysis System using Search Engine, IEEE CSS2011
[11]MALWARE DOMAIN LIST http://www.malwaredomainlist.com/
[12]SPYEYE http://about-threats.trendmicro.com/malware.aspx?language=jp&name=SPYEYE
[13]IPA「濡れ衣を着せられないよう自己防衛を！」～踏み台として悪用されないために～ http://www.ipa.go.jp/security/txt/2012/11outline.html
[14]FFRI BLOG 2013-03-27 緊急レポート: 韓国サイバー攻撃マルウェア詳細解析結果 http://www.ffri.jp/blog/2013/03/2013-03-27.htm
[15]Cuckoo Sandbox: Automated Malware Analysis http://www.cuckoosandbox.org/
[16]VirusTotal - Free Online Virus, Malware and URL Scanner https://www.virustotal.com/ja/
[17]yara-project - A malware identification and classification tool https://code.google.com/p/yara-project/
[18]K. Nakao, K. Yoshioka, D. Inoue, and M. Eto, "A Novel Concept of Network Incident Analysis based on Multi-layer Observations of Malware Activities," The 2nd Joint Workshop on Information Security (JWIS07), pp. 267-279, 2007.
[19]D. Inoue, M. Eto, K. Yoshioka, S. Baba, K. Suzuki, J. Nakazato, K. Ohtaka, and K. Nakao, "nicter: An Incident Analysis System Toward Binding Network Monitoring with Malware Analysis," WOMBAT Workshop on Information Security Threats Data Collection and Sharing (WISTDCS 2008), pp. 58-66, 2008.
[20]K. Nakao, D. Inoue, M. Eto, and K. Yoshioka, "Practical Correlation Analysis between Scan and Malware Profiles against Zero-Day Attacks based on Darknet Monitoring," IEICE Trans. Information and Systems, Vol. E92-D, No.5, pp. 787-798, 2009.
[21]"Windows を Conficker ワームから守る," http://technet.microsoft.com/ja-jp/security/dd452420.aspx
[22]マイクロソフト セキュリティ情報 MS08-067 - 緊急 http://technet.microsoft.com/ja-jp/security/bulletin/ms08-067
[23]竹久, 井上, 衛藤, 吉岡, 笠間, 中里, 中尾 "サイバーセキュリティ情報遠隔分析基盤 NONSTOP," 電子情報通信学会 情報通信システムセキュリティ研究会(ICSS), pp. 85-90, Jun 2013.
[24]"W32.Morto," http://www.symantec.com/ja/jp/security_response/writeup.jsp?docid=2011-082908-4116-99
[25] "Carna botnet scans confirmed," http://blog.caida.org/best_available_data/2013/05/13/carna-botnet-scans/