{コラム}

MWS Cup 2009

2009年10月19日〜21日　MWS 2009
10月19日 10:30-12:00 MWS Cup 2009 解析の部
17:20-18:00 MWS Cup 2009 発表の部

竹森敬祐（(株)KDDI研究所）　細井琢朗（東京大学）

企画者として振り返ってみる

2009年10月に，第2回マルウェア対策研究人材育成ワークショップ（MWS 2009）が富山国際会議場で開催された．本年の特徴は，論文発表のみならず，日頃の研究・運用で培ったノウハウや開発ツールのレベルを競うMWS Cup 2009の企画である．MWS Cupは，マルウェア対策に関する人材の育成と新たな技術やツールの発掘を目的にした，実践的な競技会である．攻撃通信データの検知精度を競う解析の部と，解析に利用したノウハウや開発ツールの応用性をアピールする発表の部から構成される．

解析の部は，クリーンな環境で収集された複数の正常通信データと，本特集の「研究用データセットを用いたマルウェア対策研究人材育成ワークショップ」で述べている複数の攻撃通信データ（CCC DATAset 2009攻撃通信データ）を見分ける技術を競う．参加者は既存のパケット解析ツールや独自開発のツールを駆使して，感染の有無やマルウェアの種別を90分以内に判定した．発表の部は，解析の部で利用したノウハウや開発ツールが，教育や運用の現場で活用でき得る美しさを持った方式であることを，3分間でアピールして芸術点を競う．大学の先生，ISPの技術者，JPCERT/CCのアナリストが審査委員となって採点した．

図-1は，MWS Cup 2009解析の部の様子である．参加は，1名から10数名までのチーム制であり，学生チーム×3，社会人チーム×3，学生と社会人の混成チーム×1など，産学連携による人材育成に資する顔ぶれとなった．解析中は，全員がPCの画面に釘付けとなり，日頃の努力を競い合う集中した90分間となった．発表の部では，手作業による解析や開発ツールのアルゴリズムなど，攻撃通信データを見抜くための実践的な工夫が次々と発表された．


図-1　MWS Cup 2009 解析の部の様子

競技用データと出題

本競技会を実践的な場にするために，インターネット上で実際に行われている攻撃通信データを用いる必要がある．そこで，MWS 2009の研究論文の執筆に利用されるCCC DATAset 2009攻撃通信データに注目し，参加者に提供されていない翌日のデータを出題した．

出題内容も現場での運用を想定し，複数の通信データの中から攻撃通信データを抽出する課題，抽出した攻撃通信データの特徴からマルウェア名を言い当てる課題，そのマルウェアが数分後に行う通信パターンを予測する課題とした．

競技用データ

MWS Cup 2009を運営する上で，競技が実践的であり，かつ実社会で直面する対策推進上の課題が含まれていることが重要である．そこで企画側である我々も，マルウェア通信について探求し，数カ月前から競技用データの作成に取り組んだ．以下，正常通信（White）データとマルウェア通信（Black）データを用いて，競技用CD-ROMを作成する過程を説明する．

White と Black の通信データ

- マルウェアに感染していないPCの通信データとして，攻撃を受けても感染しないパッチ適用済PCに向けた

解析の部（テクニカルコンポーネント）　　小計：　/ 50 点

	問 1	問 2	問 3
	c1	TROJ_AGENT.ARWZ	連鎖感染，TCP（135）スキャン，IRC 通信
	c4	PE_VIRUT.AV	IRC 通信
	c5	TROJ_DLOADR.CBK	連鎖感染，C&C 接続，外部 DNS へ MX クエリ
	c8	BKDR_RBOT.ASA	何もしない
	c9	WORM_ALLAPLE.IK	ICMP スキャン
点	4/0/-2	+3 / 0 / -1	+3 / 0 / -1

発表の部（アーティスティックコンポーネント）　　小計：　/ 50 点

	審査委員	A	B	C	D	平均
育成性	教育で利用できるか？（/15）					/15
実用性	産業で利用できるか？（/15）					/15
芸術性	美しいか？（/20）					/20

表 -1（上段）解析の部の正答表（下段）発表の部の採点表

	正答率	未答率	誤答率	得点率
問 1	91%	0%	9%	87%
問 2	31%	26%	43%	17%
問 3	11%	23%	66%	-10%

表 -2　解析の部の正答率／誤答率／得点率

無効な攻撃通信データ，さまざまな Web サイトから正常なアプリケーションをダウンロードしたときの通信データ，人が Internet Relay Chat（IRC）サービスを利用したときの通信データなどを（W）とする．

- MWS Cup 2009 用に CCC から提供いただいた攻撃通信データの中から，マルウェアが検出された一定時間分の通信データを（B）とした．
- （W）と（B）を混合した通信データを（B’）とした．

競技用 CD-ROM

（W）が 5 個，（B）と（B’）が 5 個の，合計 10 個の pcap 形式の通信データから競技用 CD-ROM を作成した．なお，パケットのタイムスタンプや MAC アドレスなど，攻撃情報以外の観点から，（W）と（B）や（B’）を見分けることができる情報を排除している．

こうした通信データの加工には，tcpslice[1]，tcprewrite[2]，tcpreplay[3] などの既存ツールを利用している．これらツールの活用は，研究や運用に資するノウハウとなることから，競技開催の 3 カ月前に通信データ加工の勉強会も開催した．

●解析の部の出題

解析の部は，作成された CD-ROM を競技中の 90 分間のみ貸し出して行われた．出題は，下記の 3 問である．採点は，正解（True Positive：TP）のみを評価せず，運用の現場で問題となる誤検知（False Positive：FP）を減点対象とし，合計 -20 点〜 +50 点で評価した．

（問 1）　攻撃通信データを探し出せ

競技用 CD-ROM に含まれる 10 個の通信データのうち，攻撃通信データが含まれる（B）と（B’）を 5 つ探し出せ．

（問 2）　マルウェア名を言え

上記で探し出した通信データに関連するマルウェアの名称を，10 個の選択肢の中から選べ．

（m0）BKDR_MYBOT.AH
（m1）BKDR_RBOT.ASA
・・・
（m9）WORM_DOWNAD.AD

（問 3）　今後の通信パターンを予測せよ

上記で特定したマルウェアによる通信について，今後の通信パターンを，10 個の選択肢の中から選べ．

（a0）何もしない
（a1）ICMP スキャン
・・・
（a9）連鎖感染，TCP（135）スキャン，IRC 通信

●発表の部の採点

人材育成に資する解析手法であったか，産業界で活用でき得る解析手法であったか，美しいかなど，発表を聞いて 3 つの観点から，合計 0 点〜 50 点で評価した．

- **育成性**　教育の現場で教材として活用でき得るかを採点する．
- **実用性**　産業の現場で監視ツールとして実用化でき得るかを採点する．
- **芸術性**　審査委員の知見から，総合的な美しさを採点する．

●結果発表

結果発表は，各チームの発表が終了した時点で，解析の部の技術点と発表の部の芸術点が集計され，スクリーンに映し出される．リアルタイムに順位が決まるため，フィギュアスケートの採点シーンを連想させられる．この集計の間に，審査委員がチームに向けて，学術的・産業的なアドバイスやコメントをする．

表 -1（上段）に解析の部の正答表を，（下段）に 4 名の審査委員による採点表を示す．トロイの木馬，バックドア，ワームなどさまざまなマルウェアによる通信データを出題している．また，通信パターンについても，何もしないものから複合的なものまで，さまざまである．

表 -2 に，解析の部の正答率／誤答率／得点率を示す．問 1 の攻撃通信データを探し出す課題は正答率が 91% であったが，正常と攻撃の通信データが混在する（B’）が含まれていたことで，判断に迷う場面もあった．問 2 で

は，通信セッションから容易にマルウェアを取り出すことができた TROJ_AGENT.ARWZ の正答率は高かったものの，それ以外は通信パターンからマルウェア名を推定することになり，誤答率が正答率を上回った．問 3 では，マルウェアの通信パターンがさまざまな条件によって変化するため，想定以上に難しい問いとなった．結果として，複数のアプローチを用意しておき，マルウェア通信の変化に対する柔軟な状況判断が勝敗をわけた．

解析の部

昨今のマルウェア通信

昨今のマルウェアの多くは，外部から指令を受けてコードの更新や攻撃を行うボットネットを構成する．こうした感染 PC の通信パターンの一例を図 -2 に示す．攻撃コードの有無のみで (B) もしくは (B') を抽出した場合，パッチ適用済み PC への無効な攻撃通信（W）を誤検知（FP）してしまうことになる．このため参加者は，下記に示す一連の通信パターンを見抜く必要がある．

- **感染フェーズ**：脆弱性を持つ 135/TCP サービスを突き，初期のマルウェアが注入される．
- **指令フェーズ**：IRC サービスを利用して指令を受信し，HTTP サービスによってマルウェアを更新する．
- **攻撃フェーズ**：外部の PC へ感染を広げるために多数の IP アドレスに対して 135/TCP サービスをスキャンする．また，スパムメールの送信なども行われる．

解析の様子

それでは実際に，参加チームが行った解析の様子を覗いてみよう．多くの参加チームは，図 -3 に示す Wireshark[4] に代表されるパケット解析ツールを用いて，ヘッダやペイロードに含まれる異常の有無や，マルウェアらしい通信パターンを探していた．図 -3 では，HTTP サービスで vot.exe というコードをダウンロードしている．このほか，マルウェアを外部から制御するための IRC 通信や不審な DNS 通信が見つかれば，(B) もしくは (B') と断定できる．多くのチームは，HTTP 通信の中からダウンロードされたコードを抽出し，これをアンチウイルスソフトで判定することでマルウェア名を特定していた．なお，参加チームには，研究論文用の CCC DATAset 2009 の前 2 日間の攻撃通信データがマルウェア名とともに提供されており，それらを基にして未来の通信パターンの予測を試みる．

参加チームの中には，限られた時間の中でこうした煩雑な作業をこなすために，多人数を動員して解析する人海戦術的なアプローチをとったり，マルウェアから発信される通信パケットの特徴を自動検出するツールを開発してきたり，さらにはパケット解析・マルウェア判定・通信パターンの予測までをも自動解析するツール（図 -4）を開発してくるなど，チームごとにさまざまな戦術をとっていた．図 -4 は，競技用 CD-ROM から図 -2 に示した通信パターンを抽出することで，その特徴を持つマルウェア名と予測される通信パターンを列挙するツールの様子である．このように MWS Cup 2009 に向けて，通信パケットを解析するための教育や，自動解析ツールの開発がなされており，MWS Cup から産業界への技術移転が期待される．

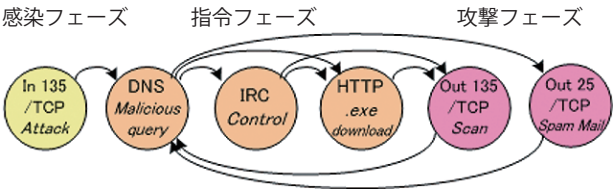

図 -2　感染 PC の通信パターンの一例

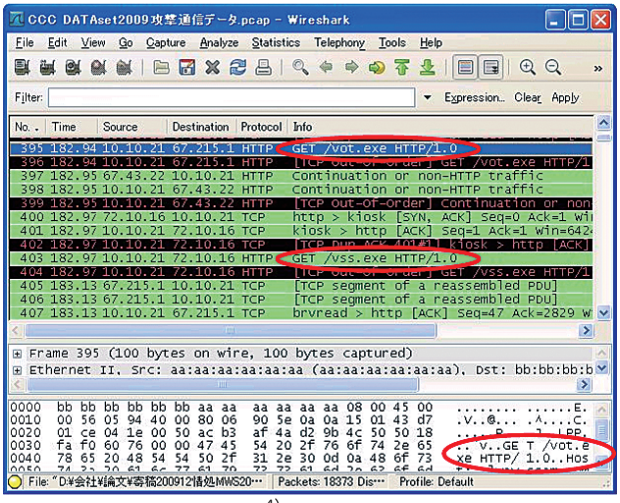

図 -3　パケット解析ツール[4]を用いた解析の一例

発表の部

参加チームは，解析の手順について，その独自性や有効性を 3 分間でアピールする．図 -5（上段）は，発表者を見守るチームメンバの様子であり，（下段）は教育・産業の現場から選ばれた 4 名の審査委員が，採点表を手に評価している様子である．

表彰

総合点，技術点，芸術点の各 1 位のチームへは，MWS 2009 実行委員長から，表彰状と大会ロゴがデザインされた T シャツが贈呈された（図 -6）．この 3 チームは，マルウェアの通信パターンの事前調査や利用するツールの習得など，競技会の開催に向けた準備を整え，

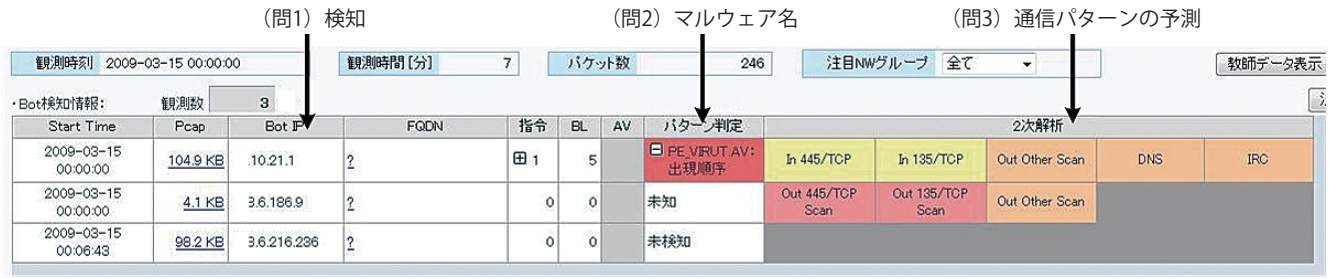

図 -4　通信パターンの抽出による感染 PC 検知の一例

図 -5　（上段）発表を控えたチームの様子（下段）審査委員よる採点の様子

技術レベルを向上させたチームであった．なお，芸術賞を受けたチームは，独自の解析ツールを開発してきており，競技中は想定外の不具合で利用できなかったものの，開発プロセスの中で得たノウハウをアピールしたことで受賞に至っている．

MWS Cup の発展に向けて

MWS Cup 2009 は，研究論文のみならず日頃鍛えた技術を実践する場として，学生から社会人まで幅広く参加できるマルウェア対策の競技会である．そして，大学などの研究者と産業界の運用者が集うことで，新たなノウハウを獲得できる人材育成の場としての狙いがある．プロ・アマが同じ土俵で競う，実力が試される競技会であり，当日のみならず日頃からの創意工夫が求められる．このため，発想が豊かな学生チームが社会人チームを上

図 -6　表彰式の様子

回るという場面もあった．今後は，学生チームには，経験と自信を付ける場として積極的に参加してほしい．社会人チームには，販売予定のツールの評価や仕上がり具合を把握する場として参加を望む．

次回は，攻撃通信データのみならず，多数のコードの中からマルウェア検体を抽出するコード解析の部も予定しており，幅広い研究者や運用者の参加を期待する．

最後に，MWS Cup は，教育と現場を結びつけるマルウェア対策五輪として，ロボットコンテストや数学五輪のような，権威ある競技会へと発展することを願っている．そして学生の皆様には，ここでの経験を履歴書の 1 ページに残していただけると幸いである．

参考文献

1）tcpslice, ftp://ftp.ee.lbl.gov/tcpslice.tar.gz
2）tcprewrite,
http://tcprelay.synfin.net/trac/wiki/tcprewrite
3）tcpreplay,
http://rcpreplay.synfin.net/trac/wiki/tcpreplay#tcpreplay
4）Wireshark, http://www.wireshark.org

（平成 22 年 1 月 5 日受付）

竹森敬祐（正会員）
takemori@kddilabs.jp
（株）KDDI 研究所ネットワークセキュリティグループ，攻撃トラフィック解析・防御技術の研究に従事．

細井琢朗（正会員）
hosoi@iis.u-tokyo.ac.jp
東京大学生産技術研究所技術職員，攻撃ホストのトレースバック，マルウェア通信解析の研究に従事．