インターネットトピック: マルウェア対策研究人材育成ワークショップ2009について

ここでは、2009年10月26日から3日間にわたって富山国際会議場で開催された、マルウェア対策研究人材育成ワークショップ2009（MWS2009）について紹介します*1。このワークショップは、情報処理学会とサイバークリーンセンター運営委員会が主催する、マルウェア解析に関する研究ワークショップです。昨年のMWS2008*2に続き今年が2回目の開催となり、研究者、学生、企業の技術者など100名以上が参加して発表や議論が活発に行われました。このワークショップは、サイバークリーンセンター*3で取得されたマルウェアの活動の情報（マルウェア検体や通信の情報）を元に、共通の解析対象としてCCC DATAset2009 を作成し、このデータセットに対して解析技術や対策技術の研究成果を共有する試みとなっています*4。

■CCC DATAset2009

発表者には、CCC DATAset2009として次の3種類のデータが提供されます。

● マルウェア検体データ

対象となる10種類のマルウェア検体のハッシュ値（検体は研究者が自ら入手する必要がある）。また、後日解析結果が提供される。

● 攻撃通信データ

ハニーポット2台に対する通信情報2日分（CCC DATAset 2009 では2009年3月13日～ 14日）。実際の通信を示すパケットダンプが提供される。

● 攻撃元データ

ハニーポット94台に対する攻撃通信の記録1年分（CCC DATAset2009では、2008年5月1日～ 2009年4月30日）。ここには、時刻、攻撃元IPアドレス、攻撃先のポート番号、感染したマルウェアに関する情報などが含まれる。

CCC DATAset2009の内容は、昨年の CCC DATAset2008 よりも拡充*5されています。発表者は、自分の研究成果をこれらのデータのいずれかに適用することで、実際の観測データを元にその有効性を検証できます。

■MWS2009の様子

昨年は一般口頭発表が22件（うち学生によるものが8件）でしたが、MWS2009では一般口頭発表30件（うち学生によるものが15件）となり、発表数が大幅に増えています。また、その内容も、マルウェア解析を効率的に行う手法から、ハニーポットの改善手法、マルウェア活動の可視化手法や活動予測の試みまで多岐にわたりました*6。発表数が示すように、学生の活躍が目立ってきたのも今回の特徴でした。IIJでは、自社で運用するMITFのハニーポットでの観測データとCCC DATAset2009の攻撃元データを比較し、観測結果の差異について大局的にまとめた結果を発表しています。同じ時期の複数の観測結果を比較することで、事象の局所性や、観測手法の精度の議論につなげられると考えています。

また、今回初の試みとして、課題として与えられたマルウェアの活動記録データを、限られた時間内に解析し、その解析技術を競う MWS Cup 2009 も開催されました。研究成果として作成したツールや、日々業務で利用する解析環境を会場に持ち込み、当日CDで配付される課題データを解析することで、その解析の早さや精度を競いました。初回にもかかわらず7チームが参加し、上位を大学のチームが占めるなど、ここでも学生の活躍が目立ちました。

マルウェア対策研究人材育成ワークショップは、共通のデータを解析することで、研究成果をフェアに比較できる場として有効なことに加えて、将来のマルウェア対策を担う人材の育成の場でもあると考えることができます。また、IIJとしては、普段あまり意見を交える機会がない学術界との交流の場としても非常に有益であり、次回以降も積極的に参加し協力していきたいと考えています。

MWS Cup 2009 会場の様子

執筆者：
齋藤 衛（さいとう まもる）
IIJサービス事業統括本部　セキュリティ情報統括部

＊1　マルウェア対策研究人材育成ワークショップ 2009（http://www.iwsec.org/mws/2009/）。情報処理学会コンピュータセキュリティ研究会主催によるコンピュータセキュリティシンポジウムCSS2009(http://www.iwsec.org/css/2009/)と同時開催。MWS2009の会場の様子は、MWS2009活動記録に掲載されている（http://www.iwsec.org/mws/2009/photo.html）。

＊2　IIJでは、昨年開催のMWS2008に続いて参加している。MWS2008（http://www.iwsec.org/mws/2008/）。また、この開催の様子は IIJ.news Vol.90 に対談として掲載している（http://www.iij.ad.jp/news/iijnews/2009/vol90.html）。

＊3　サイバークリーンセンターは総務省、経済産業省および各関連団体によるボット対策プロジェクト（https://www.ccc.go.jp/ccc/index.html）。

＊4　共通のデータを基準として研究成果を共有する試みとしては、DARPA Intrusion Detection Data Sets（1998,1999）（http://www.ll.mit.edu/mission/communications/ist/corpora/ideval/data/index.html）や Knowledge Discovery and Data Mining Tools Competition（http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html）などがある。

＊5　CCC DATAset2008 は、それぞれ、マルウェア検体データ1種類、攻撃通信データ2日分、攻撃元データ半年分のデータであった。

＊6　MWS2009の発表論文やプレゼンテーションは、著者の許諾が得られた範囲で次に公開される（http://www.iwsec.org/mws/2009/）。