プリズムマップによる可視化を用いたマルウェアの動向解析

金子 博一† 新井 悠† 松木 隆宏†

†株式会社ラック
102-0093 東京都千代田区平河町 2 丁目 16 番 1 号 平河町森タワー
hirokazu.kaneko@lac.co.jp, y.arai@lac.co.jp, takahiro.matsuki@lac.co.jp

あらまし 近年、サイバー攻撃がネットワークインフラを脅かす存在として益々進展している。とりわけ新種のマルウェアに対抗するため、世界中で発生しているこれら攻撃の特徴を把握し、よりわかりやすい形で提示することは重要だといえる。

本稿ではMWSデータセットの各攻撃に対して、地理的可視化と統計手法による分析を行った。時間的推移を考慮し、プリズムマップを用いて地理的可視化を行うことにより、より直観的に攻撃の推移を掴むことができた。

Malware Analysis by PrismMap Visualization

Hirokazu Kaneko† Yuu Arai† Matsuki Takahiro†

†Little eArth Corporation Co., Ltd. (LAC)
Hirakawacho Mori Tower, 2-16-1 Hirakawacho, Chiyoda-ku, Tokyo
hirokazu.kaneko@lac.co.jp, y.arai@lac.co.jp, takahiro.matsuki@lac.co.jp

Abstract In recent years, cyber attacks have become one of the biggest threats for our network infrastructure. Increasing unknown malwares could be considered as one of the threatening facts. It is getting very important to understand the aspects of those attacks occurring globally, and to be able to present them in understandable ways.

This report will show the conducted research; visualization and geographic statistics of the cyber attacks retrieved from anti Malware engineering WorkShop (MWS) dataset. We have successfully identified instinctive tendencies of those attacks from chronological and geographic visualization by using prism maps. This research will also include the consideration of the similarities of known.

1 背景

近年、サイバー攻撃がネットワークインフラを脅かす存在として益々進展している。特に脆弱性が発見されてからパッチが公開されるまでの間に行われるゼロデイ攻撃が問題でとなっており、ゼロデイを利用したマルウェアも多く存在している。各種セキュリティベンダーは日々新たなマルウェアの攻撃を検知・駆除できるよう邁進しており、マルウェアの挙動や特性を掴むことは重要であるといえる。

2 目的

マルウェアの挙動や特性を掴む為、本稿では様々な視点からマルウェアの挙動を分析し、

特徴を収集する。特徴的な動きがあった場合はより深く該当の通信やマルウェアを調査し、特徴をまとめる。

3　可視化による特徴把握

マルウェアの挙動分析手法の一つとして、様々な可視化手法が提示されてきた。図 1は独立行政法人 情報通信研究機構の nicter である。

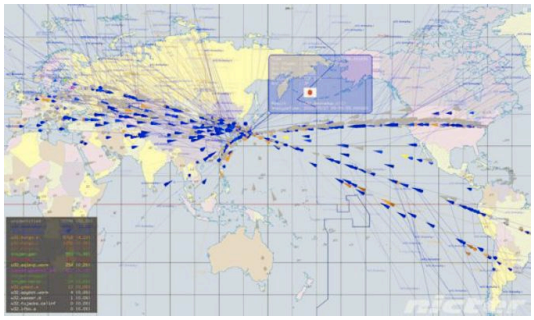

図 1 nicter 表示例

地球上の矢印は攻撃の通信を表現しており、色で攻撃の種類を表している。このシステムを使うことで、どの国からどのような攻撃が行われているか知ることができる。特に地図は人間にとって馴染みの深い可視化方法であり、より直観的にマルウェアの挙動を把握することができる。

3.1　プリズムマップ

プリズムマップは地球の国上に同形のポリゴンを生成し、攻撃の種類や量によってポリゴンの高さや色を変化させたものである。

図 2はプリズムマップの例であり、世界各国の乳児死亡率を表現している。高さで死亡率の高さを表現しており、この例では国別に表現している。また、色によって死亡率の高さのレベルを現している。

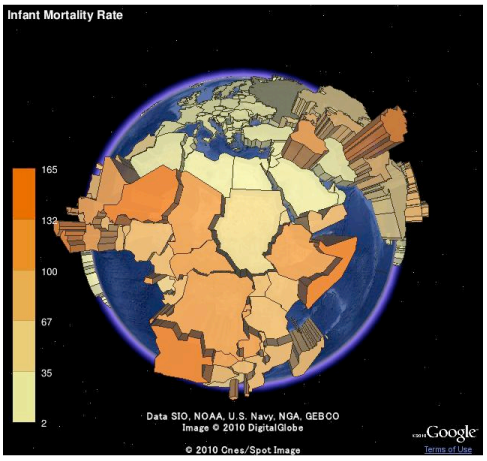

図 2 プリズムマップ - 乳児死亡率

境界が顕著に現れるため差がわかりやすく、ポリゴンが高い場合は遠くのポリゴンでも見つけやすい特徴がある。

3.2　可視化対象データ

MWS のデータセットの内、MWS2010 の 3_log を対象とした。3_log は 92 台の Honeypot の通信ログを CSV 形式で保存しており、それぞれ通信日時、通信元・先の IP・ポート番号、プロトコル名、検体ハッシュ値、ウイルス検知名、ファイル名が記述されている。また、期間は 2009 年 5 月～2010 年 4 月までの一年間である。

上記のデータセットでは Honeypot から Honeypot への通信は行われていなかった。それぞれのマルウェアのタイプによる違いを得るため、マルウェアの挙動から以下の 2 タイプに分割して可視化を行った。

- Push型マルウェア
  外部端末から管理対象のHoneypotへ攻撃を行う検体を指す
- Pull型マルウェア
  管理対象のHoneypotからHoneypot以外の端末へ接続した検体を指す。ドライブバイダウンロードなどが属する

今回は Push 型マルウェア約 4 万件、Pull 型マルウェア約 110 万件の通信を対象とした。

3.3　地理情報変換

地理的可視化を行う際、通信情報から国や緯度経度などの地理情報へ変換する必要がある。地理情報への変換はIPアドレスから緯度経度情報へ変換する。ここでは MaxMind 社の GeoIP を用いて地理情報に変換する。

GeoIP は商用目的でなければ無料で利用できるアプリケーションであり、主要諸国なら 70%～90%以上の確率で誤差 40km 以内である。本研究では国別の情報を提示しており、Web 網が一般に普及している主要諸国からの通信が多いと考えられる。そのため、GeoIP の精度で十分といえる。

今回のデータセットによる通信データは honeypot 主体の通信のため、必ず honeypot が SourceIP(送信元)か DstinationIP(送信先)に含まれている。Honeypot ではない IP アドレスを GeoIP によって地理情報に変換し、SourceIP か DestinationIP かによって Pull 型マルウェアか Push 型マルウェアかに分類しつつ変換をあらかじめ行った。

3.4　可視化方針

本研究ではマルウェアによる攻撃の通信を可視化する。攻撃の通信量が多ければ多いほどポリゴンを高く表現し、高さに応じて赤に近い色に変更しており、完全に何もない場合には薄黄色となっている。今回は欧州各国、アメリカ、中国、日本の通信量が多いことが予想されたため、遠近双方共に特徴を把握しやすいプリズムマップを採用した。

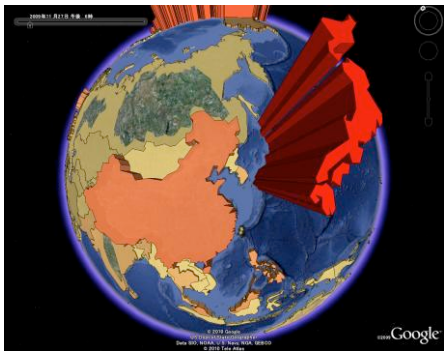

図 3 プリズムマップによる可視化例

図 3はプリズムマップを用いて可視化を行った例である。GeoIP の精度では州や県、町単位で地理情報に変換することも可能ではあるが、今回のプリズムマップでは国別にポリゴンを生成し、その高さと色で統計的情報を表現するものとした。ポリゴンは通信量に応じてプリズムの高さを高くしており、高ければ高いほど黄色から赤に変わるようになっている。また、日付毎の各国の通信量を表現しており、例えば図 3の例では、アジア圏で日本からの通信量が高く、次いで中国、インドネシアからの通信が多いことがわかる。

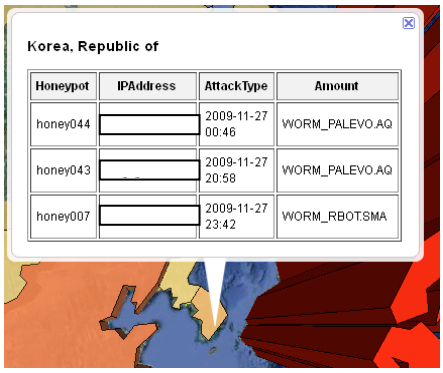

図 4 詳細情報表示例

図 4はプリズムをクリックすることによって表示させた詳細情報であり、攻撃を受けた Honeypot の番号や接続先 IP アドレス、日時、攻撃の検知名などを知ることができる。この図は日付が若い順から並んでおり、時系列に行われた通信の順番がわかりやすいようになっている。

3.5　可視化結果

今回は Push 型マルウェア、Pull 型マルウェアと種類別に分割し、可視化を行った。それによって得られた知見を記載する。

3.5.1　Push 型マルウェア

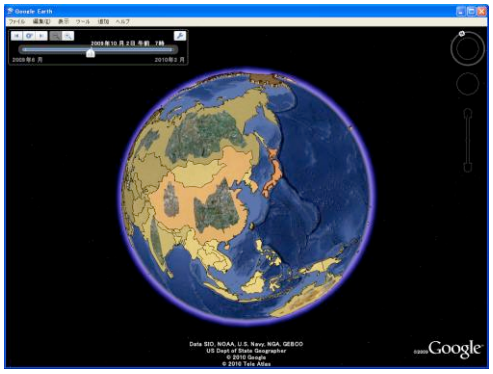

図 5 Push 型マルウェア可視化例 – 2009/10/2

2009 年 5 月～2009 年 12 月は、図 5のように日本やアメリカ、欧州各国から少量の通信がある程度であった。

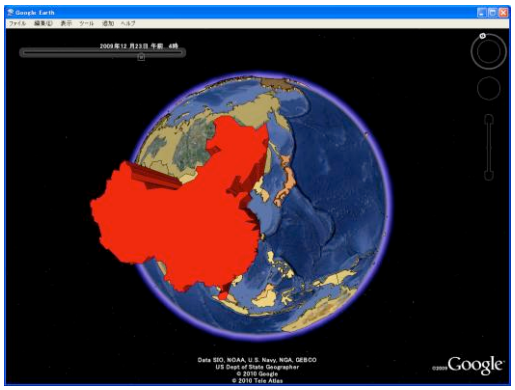

図 6 Push 型マルウェア可視化例 – 2009/12/23

しかし図 6のように、2009/12/11 を境に極端に中国から通信量が多くなることが確認できた。2009/12/11～2010/1/4 まで中国の通信量が群を抜いて多かった。

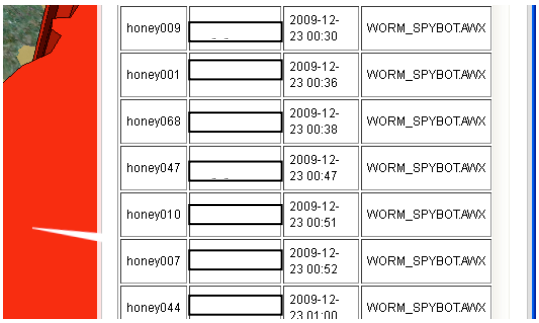

図 7 中国の詳細情報の確認 - 2009/12/23

図 7は中国の詳細情報を表示した例である。同じ IP アドレスから多くの Honeypot に対して通信が行われていることがわかる。詳細に調査すると、いままで主流だった WORM_ALLAPLE とは違い、二種類の WORM_SPYBOT による攻撃が主体であった。途中 12 月 13 日～12 月 16 日までは未知検体が検出されていたが、その直後に検出された SPYBOT と全く同じハッシュ値を持つため、SPYBOT による通信だといえる。その後、2010/1/5 から 2010/4/30 までは図 5と同じように、特別に多い通信などは確認できない状態に戻った。

3.5.2　Pull 型マルウェア

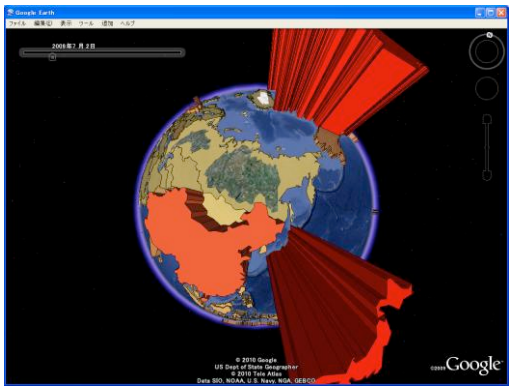

図 8 Pull 型マルウェア可視化例 – 2009/7/2

Push 型データに比べて Pull 型データはデータ量が多いため、ポリゴンも顕著である。図 8は 2009 年 5 月～2009 年 12 月までのプリズムマップの一例を表示しており、カナダや日本、中国が多いことがわかる。

honey069		2009-12-01 00:04	TROJ_BUZUS.BEZ
honey069		2009-12-01 00:04	UNKNOWN
honey069		2009-12-01 00:05	TROJ_BUZUS.BMH
honey019		2009-12-01 00:05	UNKNOWN
honey019		2009-12-01 00:05	TROJ_BUZUS.AQP
honey019		2009-12-01 00:05	UNKNOWN
honey019		2009-12-01 00:06	TROJ_BUZUS.BMH

図 9 カナダの詳細情報 - 2009/12/1

このとき、図 9のようにカナダへの通信に着目すると、おおよそ一つの Honeypot から同じ時間に 3～5 件の通信が行われている。この例では、honeypot69に対して2件の未知検体とトロイの木馬による通信が発生しており、honeypot19 に対して未知検体 1 件とトロイの木馬 2 件の通信が発生していた。このように、おおよそ 1 件～2 件はトロイの木馬に属するものであり、残り 1～2 件未知検体となっていた。また、単一の IP アドレスに対して通信を行うものであり、多くの Honeypot から接続要求が行われている。

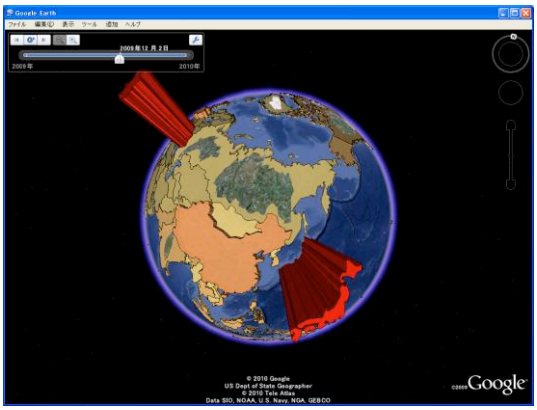

図 10 Pull 型マルウェア可視化例 - 2009/12/2

しかし、図 10のように 12 月初頭からカナダの攻撃量が激減し、その代わりにウクライナの攻撃量が激増している。

honey080		2009-12-03 12:50	UNKNOWN
honey080		2009-12-03 12:50	TROJ_BUZUS.BMH
honey080		2009-12-03 12:50	UNKNOWN
honey053		2009-12-03 12:51	TROJ_BUZUS.BMH
honey053		2009-12-03 12:51	TROJ_BUZUS.BMH
honey053		2009-12-03 12:51	UNKNOWN
honey053		2009-12-03 12:51	UNKNOWN
honey053		2009-12-03 12:51	TROJ_BUZUS.BEZ
honey053		2009-12-03 12:51	UNKNOWN

図 11 ウクライナの詳細情報 - 2009/12/3

図 11のようにウクライナへの攻撃の詳細情報を確認すると、おおよそ同じ Honeypot から同じ時刻に 3～5 件の通信が発生しており、一つはトロイの木馬に属し、残りは未知検体に属するという、12 月前のカナダの通信と酷似していた。また、対象の IP アドレスも一つだけ指定していることが判明した。

これまでは未知検体による通信が全体の通信のうち 2/3 を占めていたが、12 月 13 日以降は未知検体がほとんど検出されなくなり、代わりに WORM_KOLAB.EA の検体が検出されるようになった。ハッシュ値を確認したところ、12 月 12 日以前のほとんどの未知検体と一致したため、特定のワームとトロイの木馬に狙われていたことがわかる。

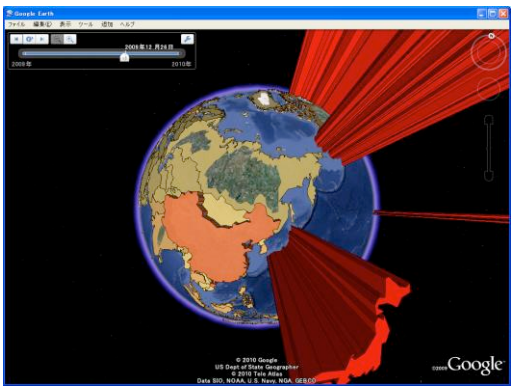

図 12 Pull 型マルウェア可視化例 - 2009/12/26

更に 2009/12/25 を境に、今度はウクライナの通信量が激減し、その代わりにアメリカの通信量が増大するようになる。図 12 は

2009/12/25 以降の代表例として、2009/12/26 のデータを表示している。この通信も先ほどのウクライナの例と同じような挙動を示しており、感染活動を行う検体名であった。更にハッシュ値を照合したところ、ウクライナで通信していた検体と一致することがわかった。

3.5.3 考察

Push 型マルウェア、Pull 型マルウェアのそれぞれにおいて、可視化による特徴を把握した。

Push 型マルウェアによる通信は、2009 年 12 月初頭から一時的に中国から SPYBOT による攻撃が行われていた。バンクーバーオリンピック開催時期や、Google 社が中国からの撤退を考慮する時期が重なるため、時事的な原因である可能性はあるが、詳細は不明である。

Pull 型マルウェアはカナダ、ウクライナ、アメリカの該当の攻撃期間を切り出して調査したところ、およそ特定のトロイの木馬やワームによるマルウェアが主体となっていることがわかった。また、時期が連続しており、一度につき 3~5 回の通信が行われる、検体のハッシュ値が一致する、IP アドレスは一つだけを対象とするなどの行動が似ていることから、ボットネットのハーダーが拠点を変更しているといった可能性が挙げられる。

また、双方共に未知検体を含んだ攻撃を行っていることが特徴的であるといえる。ベンダーによる早急な対策が功を奏しているが、攻撃者もマルウェアの流行に素早く転化していることがよくみてとれるだろう。特に同じような目的の未知検体を扱っていたことは、攻撃者の攻撃に関する関心の高さを伺わせる。

4 まとめ

MWS データセットの 3_log に対して地理的可視化手法を用いて分析を行った。Push 型マルウェア、Pull 型マルウェアに分割して可視化を行った結果、特に Pull 型マルウェアではハーダーが使用するボット群を変更するような動作が見られたため、特徴の抽出に成功したといえる。

また、全体として、攻撃者は未知検体に素早く切り替えていることがわかる。今回はベンダーの素早い対応により検知することができていた。検知されるようになってからどの程度で攻撃者が検体を変更するかといった行動分析なども考えられるため、今後はこのような視野を持った研究に結び付けたい。

参考文献

[1] nicter
http://www2.nict.go.jp/pub/whatsnew/press/h22/100601/100601.html

[2] GeoIP City
http://www.maxmind.com/app/city_accuracy

[3] 金子博一, 小池英樹, Google Map と GeoIP を用いた分散 Honeypot のログ解析と視覚化, 情報処理学会 CSS(Computer Security Symposium) 2007.

[4] 向坂真一, 小池英樹, 内部ネットワーク監視を目的とした時間・論理・地理情報の統合的視覚化システム, 情報処理学会論文誌, pp.503-512, Vol.49, No.1, 2008.

[5] 秀島裕介, 小池英樹, 複数拠点におけるサイバー攻撃監視のための IPMatrix, 情報処理学会 CSS(Computer Security Symposium) 2006.

[6] Prism Map
http://thematicmapping.org/api/prism.php