通信可視化システム MACIVISY (MAlware Communication Interactive VIsualization SYstem)によるマルウェア動的解析の支援

森 博志† 吉岡 克成† 松本 勉†

†横浜国立大学
240-8501 神奈川県横浜市保土ケ谷区常盤台 79-1

あらまし マルウェア動的解析では攻撃者とマルウェアとの通信やマルウェアによるリモートホスト への攻撃など，マルウェアの挙動を把握する上で重要な通信の観測が期待できる．一方でマルウェ アを長期間に渡り解析する場合や，短期間に大量の通信を行うマルウェアを解析する場合，観測される通信の量は膨大になり解析者にとって負担となる．本発表ではマルウェアによる通信を世界地 図上のアニメーションや折れ線グラフの連動により分かりやすく表現することで，マルウェア解析を支援する通信可視化システム MACIVISY (MAlware Communication Interactive VIsualization SYstem) を提案する．

Supporting Malware Sandbox Analysis with MACIVISY (MAlware Communication Interactive VIsualization SYstem)

Hiroshi Mori† Katsunari Yoshioka† Tsutomu Matsumoto†

†Yokohama National University.
79-1 Tokiwadai, Hodogaya, Yokohama 240-8501, JAPAN

Abstract Malware sandbox analysis can observe malware's important communications such as C&C and remote exploits, but it is a hard work for analysts to read log files especially when analysis is done for a long period or the analyzed malware produces large volume of traffic. In this paper we propose a communication visualization system called MACIVISY (MAlware Communication Interactive VIsualization SYstem). MACIVISY supports malware analysis by visualizing malware's communication with world map view and graph view that cooperate with each other.

1 はじめに

マルウェアを解析する手法のひとつに，マルウェアを実行しその挙動を観測することで解 析を行うマルウェア動的解析がある．マルウェア動的解析ではマルウェアを解析環境上で実 行 し，ファイルシステムやレジストリへのアクセス，通信などを観測することでマルウェアの挙動の把握を行う．動的解析手法では，攻撃者とマルウェアとの通信やリモートホストへの攻撃など，マルウェアの挙動を把握する上で重要な通信の観測が期待できる．一方でマルウェアを長期間に渡り解析する場合や，短期間に大量の通信を行うマルウェアを解析する場合，観測される通信の量は膨大になる[1]．tcpdump[2]や wireshark[3]などのツールによりキャプチャされた pcap 形式のファイルは可読形式で表示することが可能であるが，これらのツールのみで大量の通信を解析することは難しく，解析者にとって大きな負担となる．

本稿では動的解析手法により観測したマルウェアによる通信を可視化することでマルウェア解析を支援する通信可視化システム MACIVISY(MAlware Communication Interactive VIsualization SYstem)を提案する.

提案手法では動的解析により得られた通信ログを元に解析の起点となる通信データのサマリをグラフビューにより表示し, 通信の詳細を世界地図ビューにより表示する.

グラフビューでは宛先ポート毎の通信量の時間推移をグラフ表示し各ポートへの通信量の時間推移を概観することができる. 一方, 世界地図ビューでは IP アドレスから得られる地理情報に基づきマルウェアの通信先ホストを表示するこが可能であり, マルウェアがどのホストに対してどのポート番号で通信を行ったかを容易に把握可能である.

世界地図ビューではパケット単位の通信の様子をアニメーションで表示することが可能であり, マルウェアがどのホストに対してどういう内容の通信をどのタイミングで行ったのかを容易に把握することができる. 通信のアニメーション表示ではアニメーションの倍速表示やグラフビューによるアニメーションのシークが可能であり, 長期間の観測データであっても迅速にアニメーション表示したい期間を選択することができる.

これら 2 つのビューは連動しており, グラフビューに対する操作が世界地図ビューにも反映されるようになっている. 例えば, グラフビュー上で特定のポートを選択表示したりグラフ中の特定の期間をマウスでドラックすることで当該期間にフォーカスした可視化を行うことが可能であるが, 世界地図ビューではこの操作に対応して, 指定されたポートおよび期間の通信だけがビュー上に表示される.

2 関連研究

ネットワークセキュリティに関する既存の可視化手法には nicter[4]や RainStorm[5]などがある. これらの研究はネットワーク間のサイバー攻撃の検知・警告を目的として研究開発が行われている. nicter は情報通信研究機構で研究開発されているシステムでダークネットと呼ばれる未使用の IP アドレス空間に届く通信を観測することでマルウェアや攻撃者からの攻撃の観測を行う. Atlas は nicter の可視化エンジンのひとつであり, 当該可視化エンジンでは, 世界中から観測中のダークネットに対して送信される通信を可視化する. 図 1 で赤や青や黄色で表現されているオブジェクトがダークネットに到達したパケットを表し, 仰角が宛先ポートを, 色が通信の種類を表している. また nicter のスピンオフ技術である nirvana では特定ネットワークの内外の通信をそれぞれ可視化する.

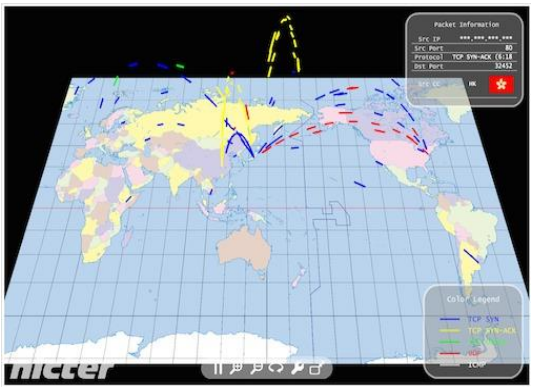

図 1 Atlas による可視化の例

RainStorm はクラス B の 2.5 倍の膨大なネットワークにおける侵入検知システム(以下 IDS と呼ぶ)の 1 日分のアラートログを可視化する. 当該システムは IDS による警告を重要度に応じて色が異なる小さなドットで時系列順にプロットして可視化を行う. 膨大なネットワークの各ホストに対するアラートを表示するため, デフォルトの可視化画面では詳細な情報を知ることは難しいがズーム機能により各ホストに対するアラートの詳細な表示を実現している.

本稿で提案する手法と上記の既存研究は, マルウェアなどによる悪性の通信を可視化することを目標としている点は類似しているが, 上記の研究は不特定のマルウェアや攻撃者から送られてくる通信データを可視化するのに対し, 本稿で提案する手法は解析環境下で解析対象となったマルウェアが送受信する通信データを可視化する点が異なる. また, 提案手法では長期間に渡る動的解析を支援するために, 全観測期間の通信の統計情報を 1 枚のグラフで示し, 更に当該グラフから解析者が指定した期間の通信に関する情報を容易にドリルダウンして可視化することが可能である.

3 動的解析環境

提案手法では動的解析環境により得られた通信を可視化の対象とする. これまで文献[6]など様々な動的解析システムが検討されているが, 本節では我々の研究拠点で運用

されている解析システム(図 2)を解析環境の例として述べる. 当該解析システムでは解析のためにマルウェアを実行させる犠牲ホストの通信をアクセスコントローラによって制御する. アクセスコントローラは解析マネージャによって予め設定されたルールに従い犠牲ホストが行う通信のうち特定のものだけをインターネットに通すことが出来る. リモートエクスプロイトなど攻撃性のある通信はインターネットには通さず, 代わりにインターネット上に存在するサーバを模したダミーサーバ群からなる擬似インターネットに転送する.

提案手法では以 上 に説明した動的解析システムなどの犠牲ホストが行う通信を観測して得られた通信挙動ログ(pcap ファイル)を可視化する.

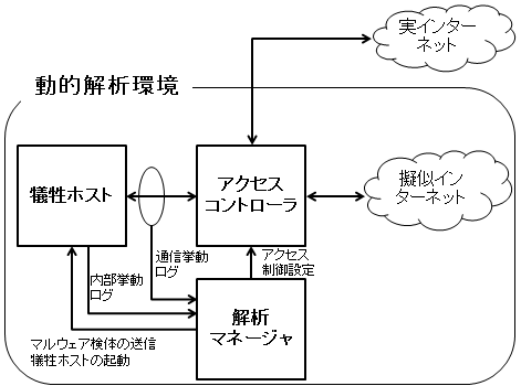

図 2 動的解析環境例

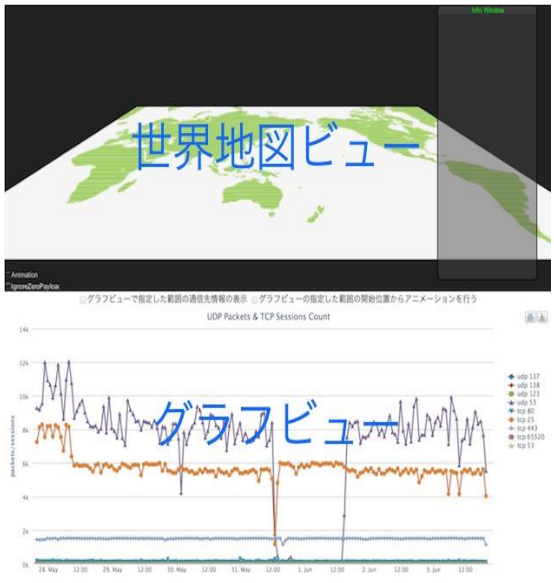

図 3 世界地図ビューとグラフビュー

4 MACIVISY における可視化手法

本章では我々が提案する可視化システム MACIVISY について述べる. MACIVISY は動的解析環境において観測した通信をグラフビューと世界地図ビューの 2 つのビューにより可視化する. グラフビューでは, 観測した通信量の時間推移を示し, 世界地図ビューでは観測した通信の宛先などの詳細な情報を可視化することで解析者を支援する(図 3).

4.1 グラフビュー

グラフビューでは解析環境から送信されたパケットの宛先ポート毎の通信量の時間推移を示す. TCP 通信に関してはセッション数で, UDP 通信に関してはパケット数でカウントし, 折れ線グラフで描画する. 当該ビューではまず解析者に観測期間全体の通信量の時間推移を示すことで, 解析者はマルウェアによる通信の概要を容易に把握できる. また当該ビューはインタラクティブに動作し, 以下の機能を持つ.

1. 指定したポートの表示・非表示切り替え

通信先ポート数が多い場合やポートごとの通信量に大きな差がある場合, 解析者が知りたい情報が見えづらくなる場合がある. そこで提案提案手法ではグラフ右側の凡例ポートをクリックすることにより注目するポートを選択することが可能である(図 4).

2. 期間指定によるグラフのドリルダウン

解析期間が長期に渡る場合, 短い期間の情報が見えづらくなる. 当該機能を利用するとグラフの内容を指定した期間にドリルダウンすることができる(図 4).

3. グラフビューへの操作の世界地図ビューへの反映

グラフビューから通信の詳細について更に可視化したい時間帯を指定することで指定した期間の通信の様子を世界地図ビューで可視化することができる.

図 4 グラフビューによる情報の絞り込み

4.2 世界地図ビュー

世界地図ビューでは解析環境で観測した通信の宛先ホストの位置やパケットの流れを示す．通信先ホストの位置情報は Maxmind 社[7]の GeoIP データベースにより IP アドレスを元に取得する．また解析対象のマルウェアが動作する環境は，解析環境が存在する地点の上空に描画し，ローカルアドレスなど GeoIP データベースに登録されていない IP アドレスは解析環境付近上空に描画する．なおデフォルトでは解析環境は日本上空に表示されるようにしている．可視化する通信を選択すると当該ビューには観測期間全体における解析環境の通信先ホストが表示される(当該機能は 2014 年 08 月 22 日時点では未実装である)．また，当該ビューにはグラフビューにより選択された期間の通信先ホストの情報を表示する機能と，パケットの流れをアニメーションで表示する機能がある．

前者の機能では指定された期間において解析環境と通信を行ったホストを通信量に比例した太さを持つ線で結んで示す．それぞれのホストを表すオブジェクト上には通信に使用したポート番号を描画する．また，HTTP や SMTP などのマルウェアに頻繁に利用されるプロトコルに対応するポートに関しては数字の代わりにプロトコルに対応するアイコンを描画することで通信内容の把握を容易にする（図 5）．

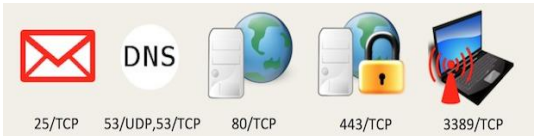

図 5 世界地図ビューで表示するアイコン一覧

後者の機能では指定された期間において観測された通信のパケットの流れをアニメーションで表示することで可視化を行う．パケットは立方体のオブジェクトとして描画し，オブジェクトの色は TCP 通信なら青色，UDP 通信なら赤色で描画する．また TCP,UDP のペイロードサイズに応じてオブジェクトの描画サイズを変える．また前者の機能と同様にマルウェアによって頻繁に利用されるプロトコルに対応するポートによる通信の場合は，描画されたパケットの上部にポートに対応するアイコンを描画する．パケットのアニメーションは一時停止の他に，再生速度を上げることが可能であり，通信頻度が低いマルウェアを解析する場合に有用である．また，解析者が任意のタイミングで 1 パケットずつ順番に描画させることができるステップ可視化機能もあり，通信頻度が低いマルウェアを解析する場合や，マルウェアによる通信の順序を詳細に解析する場合に有用である．また解析者は描画されたパケットをクリックにより選択することで当該パケットのペイロードなどの詳細な情報を得ることが可能である．

5 MACIVISY による通信解析例

本章では提案手法による 2 つの解析例を挙げる．なお，実際に提案システムを用いて解析を行った際のキャプチャ動画を研究拠点の Web ページで公開している[8]．

5.1 morto の解析例

morto は 2011 年頃に流行したマルウェアである[9][10]．morto は Windows のリモートデスクトップサービスを利用して感染を拡大する．morto に感染したホストは他のホストに対して上記サービスにより頻繁に利用されるパスワードを用いてログインを試みる．ログインに成功すると自身のコ

ピーをログイン先のホストに作成する．我々の研究拠点で当該検体を 2013 年 6 月 22 日から 13 日間動的解析して得られた通信を MACIVISY により可視化すると，図 6 のグラフが表示される．当該グラフから解析した検体は解析開始から終了まで定常的に 53/UDP 宛に 1 時間当たり 10000 パケット程度送信しており，また 3389/TCP 宛にも定常的に 1 時間あたり 1000 セッション程度の通信を行っていることが分かる．3389/TCP は Windows リモートデスクトップサービスのデフォルトポートであり，当該通信は morto が感染を拡大するために他ホストに対してログインを試みていたものと推測できる．また当該検体は DNS の TXT レコードを利用して攻撃者との C&C 通信を行うことが知られており[10]，53/UDP 宛の通信は C&C 通信であることが推測できる．

更に当該検体の通信の様子を世界地図ビューにより可視化すると 3389/TCP 宛の通信は特定の地域に偏ることなく，世界中のホストに対して行われていることが分かる（図 7）．一方で 53/UDP については，当該解析環境ではアメリカにある Google の DNS サーバを利用して名前解決を行うように設定しているにも関わらず世界中の数十か所の DNS サーバに通信している(図 8)．また本節で可視化した通信データでは確認できなかったが，2012 年 9 月の時点では，これらの DNS サーバのいくつかからは応答も得られており，ペイロードを見ると暗号化されたデータが届いていることが確認できる（図 9）．なお，これらの DNS サーバ群は DNS サービスを一般に公開しているオープンリゾルバであり，morto はこれを悪用して C&C 通信を行っている．

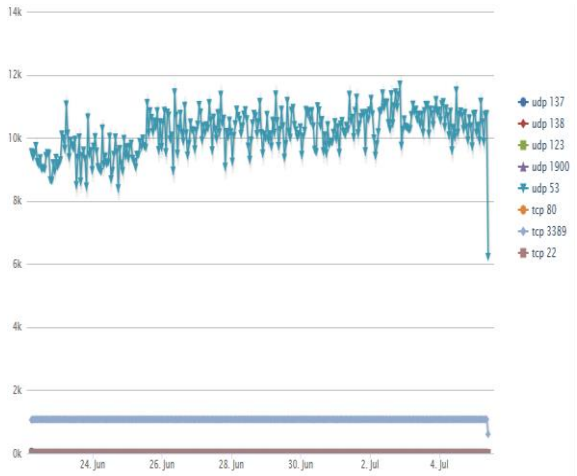

図 6　morto の通信の概要

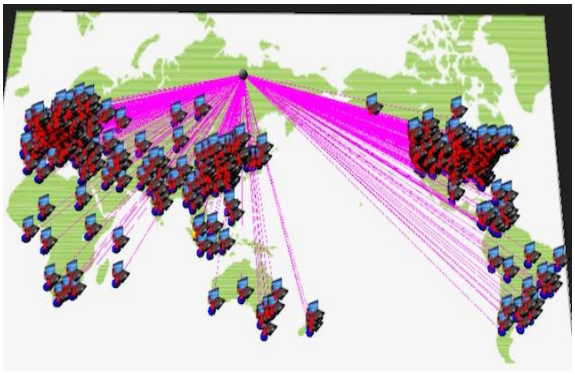

図 7　3389/TCP 宛通信量の時間推移

図 8　53/UDP 宛通信量の時間推移

```
2012-09-27 16:26:26.527450 IP 192.168.228.240.1035
 > 208.67.222.123.53: 3+ TXT? e.ppift.com. (29)
2012-09-27 16:26:26.728958 IP 208.67.222.123.53 >
192.168.228.240.1035: 3 2/0/0 CNAME e.ppiftns.in.,
 TXT "p6666U66666CD6JaeC3mDxZXekRX7pybu9WO66666666p
66664kYa7EFCQkRaD4Fc86JcZNfXlp666666p6666T66666CD6
JaeC3mDxZXekRXE=JmM9WO6666666666666666666666666666
666666666666666666666" (234)
```

図 9　morto による C&C 通信のペイロード

5.2　spybot の解析例

spybot は IRC サーバを利用して攻撃者と通信を行い DoS 攻撃などの様々な悪意ある活動を行うことが知られている[11]．当該検体を 2012 年 5 月 27 日から 7 日間観測した通信を提案システムにより可視化すると図 10 のグラフが得られる．

グラフビューから 5 月 31 日から 6 月 1 日の間に通信量の減少が激しい期間があることが分かる．さらにポートごとの通信量の変化を見ると 53/UDP，65520/TCP，443/TCP，25/TCP の 4 種のポートは上記期間において通信量が激減していたことが分かる(図 11，図 12，図 13，図 14 の赤い印の期間)．

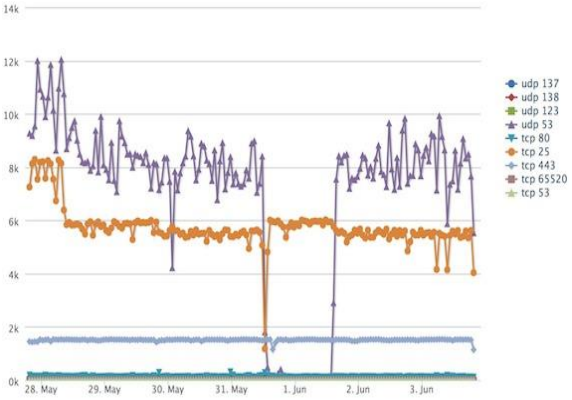

図 10　spybot の通信の概要

図 11　53/UDP 宛通信量の時間推移

図12　65520/TCP宛通信量の時間推移

図 13　443/TCP 宛通信量の時間推移

図 14　25/TCP 宛通信量の時間推移

65520/TCP宛通信は最も特徴的であり当該期間まで1 時間に1セッションの通信を行っているが，その後はセッション確立をしなくなる．当該ポートに関する通信のTCPペイロードを見るとPING,PONGという文字列が確認できる．これらの文字列はIRCプロトコルで使用されることから当該ポートではIRC通信が行われており，攻撃者とのC&C通信に使用された可能性がある．いくつかのポートにおいて通信量の変化が見られたのは当該ポートへの通信がなくなり，マルウェアの挙動が変わったものと推測できる．また世界地図ビューで当該ポート宛の通信先ホストを表示すると通信先ホストがヨーロッパにあることが分かる（図15）．

図 15　65520/TCP 宛通信の宛先ホスト

25/TCP は SMTP サーバのデフォルトポートであることから，当該通信はスパムメールなどの送信を目的として行われたものだと推測できる．図 11，図 14 を見ると 53/UDP 宛通信が激減する以前の期間では 53/UDP と 25/TCP 宛通信の通信量の増減は連動していることが分かる．この時期の 53/UDP の通信内容を確認すると，実際にスパムメールの宛先メールアドレスの MX レコードの名前解決が行われており(図 16)，宛先の SMTP サーバに直接送る方法(直接スパム送信と呼ぶこととする)でスパムメール送信を行っていることがわかる．次に 53/UDP 宛通信が頻繁に行われている時期(直接スパム送信時期)と激減する時期の 25/TCP 通信の宛先を世界地図ビューにより比較すると図 17，図 18

のようになり，前者は世界中の SMTP サーバと通信しているのに対して，後者は特定の SMTP サーバとだけ通信をしていることが分かる．しかしながら，図 14 に示す通り，25/TCP 宛の通信は継続的に行われていることから spybot はこれらの特定の SMTP サーバに大量のメール送信をしていることが分かる．この事実から推測されることは，これらの SMTP サーバは他の SMTP サーバへのメールの転送が許可されたオープンリレーであり，spybot はこれらのオープンリレーを介して間接的にスパム送信を行っている(間接スパム送信)ということである．詳細は割愛するが実際に間接スパム送信を行っていることはペイロードの検証により確認済みである．

さらに長期間に渡り同様の検体を動的解析した際の 53/UDP の通信推移を示す(図 19)．53/UDP 宛通信量の増減から spybot は直接スパム送信と間接スパム送信を交互に実施していることがわかる．スパム送信方法を変更する理由として，送信元 SMTP サーバの認証やリレーサーバのブラックリストを適用している宛先 SMTP サーバに対して複数の送信方法を試す目的があると推測される．

```
2012-05-27 19:41:05.395003 IP 192.168.228.240.1114 >
 8.8.8.8.53: 28+ MX? aol.com. (25)
2012-05-27 19:41:05.408723 IP 8.8.8.8.53 > 192.168.2
28.240.1114: 28 4/0/0 MX mailin-03.mx.aol.com. 15, M
X mailin-04.mx.aol.com. 15, MX mailin-01.mx.aol.com.
 15, MX mailin-02.mx.aol.com. 15 (132)
```

図 16　spybot による MX レコードの名前解決

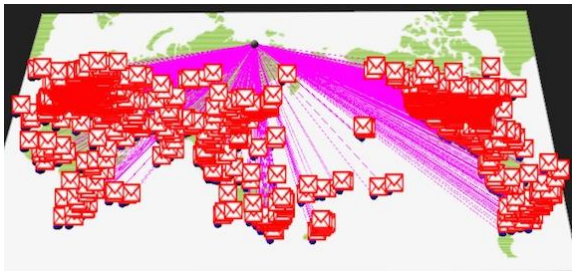

図 17　53/UDP 宛通信が激減する前の
25/TCP 宛ホスト

図 18　53/UDP 宛通信が激減した期間の 25/TCP 宛ホスト

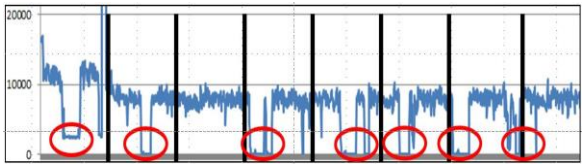

図 19　8 週間の 53/UDP 宛通信量の時間推移

80/TCP,443/TCP 宛の通信は 65520/TCP 宛通信の傾向が大きく変動した時期(図 12 参照)の前後で変動しており，この時期以前は，ロシア，日本，東南アジアに対しても通信を行っているのに対して，後半は欧米のホストが主であった．この時期を境に通信の傾向が変わった原因として，マルウェアの更新や別検体の実行が疑われるため，この時期の通信ログや内部挙動ログを精査することで，新たな知見を得られる可能性がある(図 20)．

図 20　80/TCP,443/TCP 通信の宛先ホスト

図 21　53/UDP 宛通信の宛先ホスト

```
2012-05-27 19:42:32.787853 IP 192.168.228.240.1210 >
 192.58.128.30.53: 21+ NS? com. (21)
2012-05-27 19:42:32.899903 IP 192.58.128.30.53 > 192
.168.228.240.1210: 21- 0/13/15 (509)
2012-05-27 19:42:32.900882 IP 192.168.228.240.1212 >
 192.33.4.12.53: 67+ NS? org. (21)
2012-05-27 19:42:33.029871 IP 192.33.4.12.53 > 192.1
68.228.240.1212: 67- 0/6/12 (423)
2012-05-27 19:42:33.030325 IP 192.168.228.240.1213 >
 192.36.148.17.53: 65+ NS? de. (20)
2012-05-27 19:42:33.049863 IP 192.36.148.17.53 > 192
.168.228.240.1213: 65- 0/6/10 (334)
```

図 22 ヨーロッパのホストに対する 53/UDP 宛通信の内容

さらに，53/udp の宛先に着目すると，解析環境では DNS サーバとして Google の DNS サーバを利用するように設定しているにもかかわらず，ヨーロッパのホストに対しても通

信をしていたことが分かる(図 21). この通信内容を確認すると NS レコードのクエリであり, 通信の目的は不明である(図 22).

以上は解析の一例ではあるが提案システムによりマルウェアによる通信の内容把握の効率化が出来ると考える.

6 まとめと今後の課題

本稿ではマルウェア動的解析によって観測した膨大な通信を効率よく把握するための可視化システム MACIVISY を提案した. また, MACIVISY によるマルウェアの解析例を示した. マルウェアは観測期間中, 常に同じ動作をするとは限らず, 起動時間や攻撃者からの命令などを条件に様々な挙動を示す場合がある. そのため時間に注目して解析を進めることはマルウェアの挙動を把握する上で有効である. 提案手法は通信量の時間推移を示すグラフビューと当該ビューと連動する世界地図ビューにより, 直感的な操作で任意の期間の通信の様子を可視化することが可能であり, 時間に注目した解析を支援する.

マルウェア動的解析ではマルウェアによる通信以外にもマルウェアによるファイルの書き換えやプロセスの生成などの情報も観測することが可能であるため, これらの情報も同時に可視化できる手法の検討が今後の課題である.

謝辞 本研究の一部は, 総務省情報通信分野における研究開発委託/国際連携によるサイバー攻撃の予知技術の研究開発/サイバー攻撃情報とマルウェア実体の突合分析技術/類似判定に関する研究開発により行われた.

参考文献

[1]森博志, 吉岡克成, 松本勉, “長期間のマルウェア動的解析を支援する通信可視化手法とユーザインタフェースの提案, ” 第 58 回 CSEC・第 4 回 SPT 合同研究発表会, 2012.

[2]TCPDUMP, “TCPDUMP & LibPCAP,” http://www.tcpdump.org/ (最終閲覧日:2014/08/02)

[3]WIRESHARK, “WIRESHARK,” http://www.wireshark.org/ (最終閲覧日:2014/08/02)

[4]中尾康二, 松本文子, 井上大介, 馬場俊輔, 鈴木和也, 衛藤将史, 吉岡克成, 力武健次, 堀良彰, “インシデント分析センタ nicter の 可視化技術, ”ISEC, Vol.106, No.176, pp.83- 89, 2006.

[5]Abdullah, Kulsoom, et al. "IDS RainStorm: Visualizing IDS Alarms," VizSEC2005, セッション 1 – 1, 2005.

[6]Katsunari Yoshioka, Tsutomu Matsumoto, “Multi-Pass Malware Sandbox Analysis with Controlled Internet Connection,” IEICE Trans, E93A, No.1, pp.210-218, 2010.

[7]MaxMind, “MaxMind - IP Geolocation and Online Fraud Prevention,” https://www.maxmind.com/ja/home (最終閲覧日:2014/08/22)

[8]横浜国立大学 情報・物理セキュリティ研究拠点, “通信可視化システム MACIVISY,” http://ipsr.ynu.ac.jp/macivisy/index.html (最終閲覧日:2014/08/25)

[9]IPA, “Morto 一般向け情報, “ http://www.ipa.go.jp/files/ 000002624.pdf (最終閲覧日:2014/08/10)

[10]Symantec, “DNS レコードを利用する Morto ワーム, ” http://www.symantec.com/connect/blogs/dns-morto (最終閲覧日:2014/08/10)

[11]Symantec, “W32.Spybot.Worm,” http://www.symantec.com/ja/jp/security_response/writeup.jsp?docid=2003-053013-5943-99&tabid=2 (最終閲覧日:2014/08/13)