Vol.2014-CSEC-65 No.5
Vol.2014-IOT-25 No.5
2014/5/22

権威DNSサーバのクエリログの可視化による攻撃の発見と分析

渡辺 拳竜[1]　松井 一乃[1]　池部 実[2]　吉田 和幸[3]

概要：ボットなどからの攻撃は，その多くが攻撃対象ネットワークの探索から始まる．例えば，spam 送信の際にはその送信先を決定するために MX レコードを問合せる．また，PTR レコードを用いてホストの存在を確認するホスト探索攻撃がある．これらの攻撃は攻撃対象ネットワーク内の権威 DNS サーバに対して問合せをする．本研究では，権威 DNS サーバのクエリログを用いて spam 送信やホスト探索攻撃を検知することを目的とする．特定のレコードについて分析することは攻撃発見の手がかりとなる可能性がある．そこで，本論文ではこれらの攻撃を検知するための前段階として，MX レコードと PTR レコードに着目した分析をする．権威 DNS サーバに対して MX レコードを問合わせた送信元 IP アドレスと同時期のメールサーバに対する spam の送信元 IP アドレスの上位 16 ビットについてヒルベルト曲線を用いて可視化し，比較した結果や，PTR レコードを問合わせた IP アドレスを集計し，問合せ数上位 5 件の IP アドレスについて分析した結果について報告する．

キーワード：DNS，spam，ログ解析，可視化

Visualization of query log of authoritative DNS server for attack analysis and detection

Kenryu Watanabe[1]　Kazuno Matsui[1]　Minoru Ikebe[2]　Kazuyuki Yoshida[3]

Abstract: Attackers sweep networks and look for target hosts. For example, the attacker queries MX record to authoritative DNS server for spam sending. Moreover, the attacker queries PTR record to the DNS server for host sweep. In this paper, we aim to detect some attack using queries-log of the authoritative DNS server. Therefore, we analyzed MX and PTR records of queries-log in Oita University. We use a Hilbert curve to map the first and second octets of source IP addresses that sent spam and queried MX record. And, we counted up the daily PTR records of queries-log.

Keywords: DNS, spam, Log Analysis, Visualization

1. はじめに

インターネットの発達と普及に伴い，我々の生活にはネットワークが不可欠な存在になっている． Web ページの閲覧や，メールなどのインターネット上のサービスはドメイン名を用いるため DNS(Domain Name System) を欠かすことはできない．DNS はドメイン名と IP アドレスの対応付けをする仕組みであり，現在のインターネットの根幹を支える重要なサービスのひとつである．DNS サーバには，権威 DNS サーバとキャッシュ DNS サーバの 2 つの機能がある．DNS は木構造上にドメイン名空間を分割した分散データベースとして機能する．各ドメイン名空間を

[1] 大分大学大学院工学研究科知能情報システム工学専攻
Course of Computer Science and Intelligent Systems, Graduate School of Engineering, Oita University
[2] 大分大学工学部知能情報システム工学科
Department of Computer Science and Intelligent Systems, Faculty of Engineering, Oita University
[3] 大分大学学術情報拠点情報基盤センター
Center for Academic Information and Library Services, Oita University

管理する DNS サーバが権威 DNS サーバであり，ユーザからのクエリを受信し，権威 DNS サーバへ問合せをするのがキャッシュ DNS サーバである．

警察庁の調査報告によると，DNS サーバを割り出すためのポートスキャンや標的型メール攻撃などのインターネットにおける不正通信は依然として多い状況にある [1]．攻撃者がインターネット上に存在するホストを攻撃対象としたとき，そのホストが所属するドメイン空間を管理している権威 DNS サーバには何らかの問合せがなされる可能性がある．例えば，spam 送信の際には権威 DNS サーバへ MX レコードを問合せる．普段から頻繁にメールをやりとりしている通常のメールサーバの場合では，配送先のメールサーバの MX レコードが身近なキャッシュ DNS サーバに保持されていることが多く，権威 DNS サーバを問合せることは少ないが，spam を送信するメールサーバは攻撃対象としたドメインの権威 DNS サーバへ spam 送信の直前に MX レコードを問合せることが考えられる [2]．また，攻撃対象ネットワーク内の IP アドレスを逆引きすることでホスト名を収集するホスト探索攻撃が存在する．そこで本研究では，権威 DNS サーバのクエリログを用いて spam 送信やホスト探索攻撃を検知することを目的とする．本論文では，これらの攻撃を検知するための前段階として，MX レコードを問合せた送信元 IP アドレスと spam 送信元として判断された IP アドレスを可視化し，分析する．また，PTR レコードを問合せる送信元 IP アドレスの問合わせ数を集計することで，ホスト探索攻撃を発見するための分析をする．

本論文の構成は以下の通りである．第 2 章で権威 DNS サーバのクエリログの分析について述べ，第 3 章で関連研究について述べる．そして第 4 章で権威 DNS サーバのクエリログの調査・分析の結果について述べ，最後に第 5 章でまとめと今後の課題について述べる．

2. 権威 DNS サーバのクエリログ分析

2.1 権威 DNS サーバのクエリログによる攻撃の発見

DNS はインターネット上の複数の権威 DNS サーバに，管理するドメイン名空間を権限移譲することで，分散データベースとして機能し，IP アドレスとドメイン名の変換やメール配送先の決定などの役割を担う．ドメイン名が必要となる通信の際には通信相手のドメイン名空間を管理する権威 DNS サーバへ問合せをする．メールを送信する際には，配送先のメールサーバを決定するため，権威 DNS サーバへ MX レコードを問合せる．この動作は spam 送信元も同様である．我々は，milter manager[3] を用いて，複数の spam 対策を組み合わせることで spam を検出するためのメールサーバを構築し，運用してきた [4]．大分大学のメールサーバに対し，spam を送信してきた IP アドレスを分析する．権威 DNS サーバに対して MX レコードを問合せた

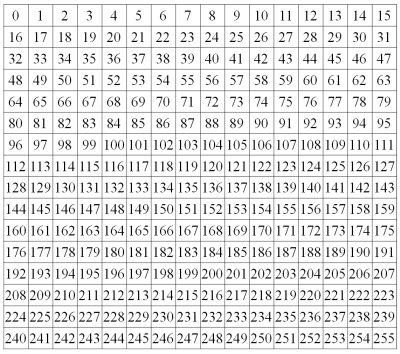

図 1 空間充填曲線を用いない 0 から 255 の分布図

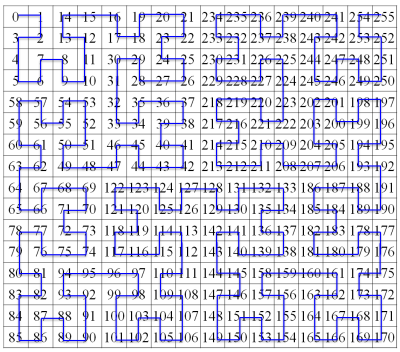

図 2 ヒルベルト曲線を用いた 0 から 255 の分布図

IP アドレスと spam 送信元と判断された IP アドレスをそれぞれヒルベルト曲線により可視化し比較することで，権威 DNS サーバのクエリログから spam 送信者を検知するための手掛かりの発見を目指す．通常 MX レコードを問合せるのは，メールサーバではなくキャッシュ DNS サーバであるため，MX レコードの問合せをする送信元 IP アドレスと spam 送信元 IP アドレスが完全に一致することは考えにくい．また，メールサーバとキャッシュ DNS サーバが分離されており，同一ネットワーク内に存在する場合であっても，サブネットマスクが 24 ビットとは限らないので，IP アドレスの上位 24 ビットを比較しても一致しない可能性がある．そのため，ヒルベルト曲線による IP アドレスの可視化は上位 16 ビットを対象とする．これにより，キャッシュ DNS サーバとメールサーバを一致した IP アドレスブロック上に捉えることが可能である．

また，権威 DNS サーバのクエリログに痕跡が現れる攻撃にはホスト探索攻撃がある．ホスト探索攻撃は，攻撃対象ネットワーク内 IP アドレスを PTR レコードにより逆引き問合せをすることで，そのネットワークに所属するホストの存在を確認する攻撃である．これを権威 DNS サーバのクエリログにより発見するため，1 日ごとの PTR レコード問合せ数上位 5 件の送信元 IP アドレスについて分析する．

2.2 ヒルベルト曲線

ヒルベルト曲線は [5][6] は空間充填曲線の一つである．

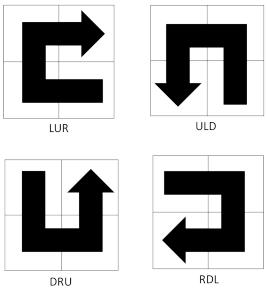

図 3 ヒルベルト曲線の 4 つの基本図形

空間充填曲線は多次元空間の情報を 1 次元空間に写像する手法として用いられる．空間充填曲線の代表例としてはヒルベルト曲線の他に，ルベーグ曲線が挙げられる．二次元空間充填曲線は平面上において近隣のノードが 1 次元でも近傍に配置されるように空間充填するため，平面上のノード近接性と変換後の一時現状での近接性に相関関係を持たせることができる [7]．空間充填曲線は，それを用いない場合 (図 1) と比較すると，連続した IP アドレスを 1 つの塊として写像できる．これに対して図 2 に示すヒルベルト曲線は，IP アドレスブロックが大きく離れることなく近接性を保持したまま描写できる．そのためヒルベルト曲線は，IP アドレスがどの IP アドレスブロックに存在しているかを調査する際に直感的に違いを把握できる利点がある．そのため，本論文ではヒルベルト曲線を用いて，MX レコードの問い合わせた送信元 IP アドレスと spam 送信元と判断された IP アドレスをそれぞれ可視化する．

ヒルベルト曲線を図 3 に示す，カタカナのコの文字の形状をした 4 つの基本図形があり，再帰的に呼び出す．基本図形を LUR(Left-Up-Right)，ULD(Up-Left-Down)，DRU(Down-Right-Up)，RDL(Right-Down-Left) とする．

この基本図形を 4 つのルールに従って再帰的に呼び出すことでヒルベルト曲線ができる．4 つのルールを以下に示す．↑，↓，→，←は，各方向への描画を示す．例としてルール 1 について説明する．まず，ULD(n) が呼び出されると，LUR(n-1) を呼び出す．LUR(n-1) の描画が終了すると，LUR(n-1) の描画が終了した場所から上方向に線を描画し，ULD(n-1) を呼び出す．以後同様に,ULD(n-1), ←,ULD(n-1), ↓,RDL(n-1) の順で呼び出し実行する．n=0 になるまで繰り返す．

- ルール 1
  ULD(n)=LUR(n-1), ↑,ULD(n-1), ←,ULD(n-1), ↓,RDL(n-1)
- ルール 2
  DRU(n)=RDL(n-1), ↓,DRU(n-1), →,DRU(n-1), ↑,LUR(n-1)
- ルール 3
  RDL(n)=DRU(n-1), →,RDL(n-1), ↓,RDL(n-1), ←,ULD(n-1)

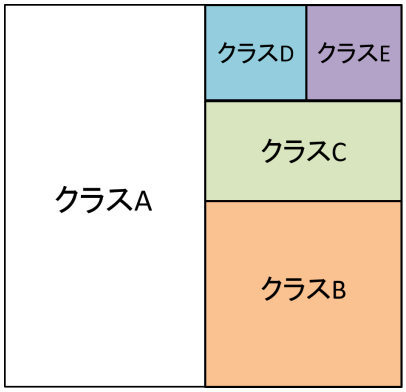

図 4 ヒルベルト曲線を用いて表現したクラス A〜E のクラスフル IP アドレス

- ルール 4
  LUR(n)=ULD(n-1), ←,LUR(n-1), ↑,LUR(n-1), →,DRU(n-1)

上記のルールに従ってヒルベルト空間曲線を描くと 2 の n 乗の正方形になる．また，ヒルベルト曲線を用いて表現したクラス A〜E のクラスフル IP アドレスの分布を図 4 に示す．

3. 関連研究

関連研究として権威 DNS サーバを対象としたログ分析の研究について述べる．

デニスら [8] は権威 DNS サーバのログを用いて，2007 年 4 月 1 日から 2008 年 7 月 31 日の期間について，ある大学の権威 DNS サーバの DNS パケットの流量エントロピーを調査した．権威 DNS サーバが管理するドメイン空間に所属するホストがメールを送信した場合，受信側メールサーバは送信元ホストを逆引きし，ホスト名を確認する．このような特徴をもとに，権威 DNS サーバの PTR レコードのエントロピーを分析した結果，送信元 IP アドレスに関する流量エントロピーの増加に伴い，DNS クエリキーワードに関する流量エントロピーの減少する箇所では，spam ボットやホスト探索活動の挙動が観測されたことを報告した．

デニスらの研究は権威 DNS サーバを設置しているネットワーク内の spam ボットを検知するための分析である．今回，我々が検知対象とするのは外部ネットワークに存在する spam 送信元やホスト探索攻撃をする送信元である．本論文では，MX レコードを問合せる送信元 IP アドレスと spam の送信元 IP アドレスをヒルベルト曲線により可視化することで，外部ネットワークに存在する spam 送信者を検知するための分析をする．

4. 権威 DNS サーバの分析結果

4.1 MX レコードの分析

大分大学内に設置した権威 DNS サーバのクエリログを対象として分析する．分析した期間は 2014 年 2 月 1 日か

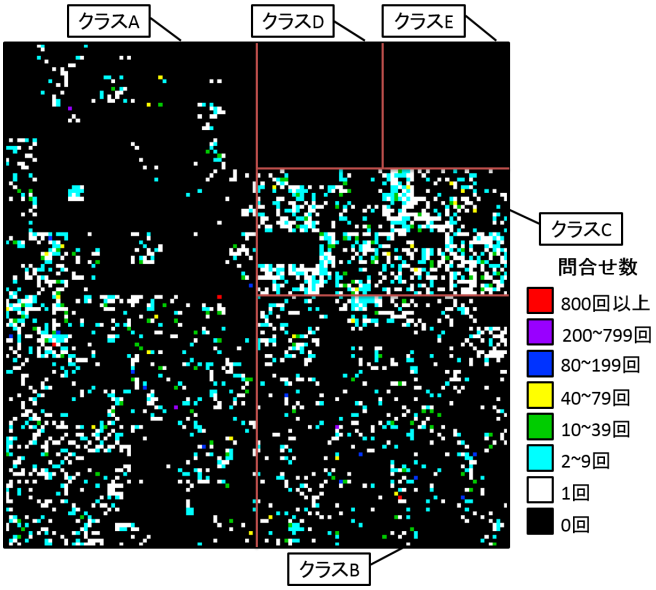

図 5　2014 年 2 月 1 日から 2 月 10 日の MX レコードを問合せをした送信元 IP アドレスの分布

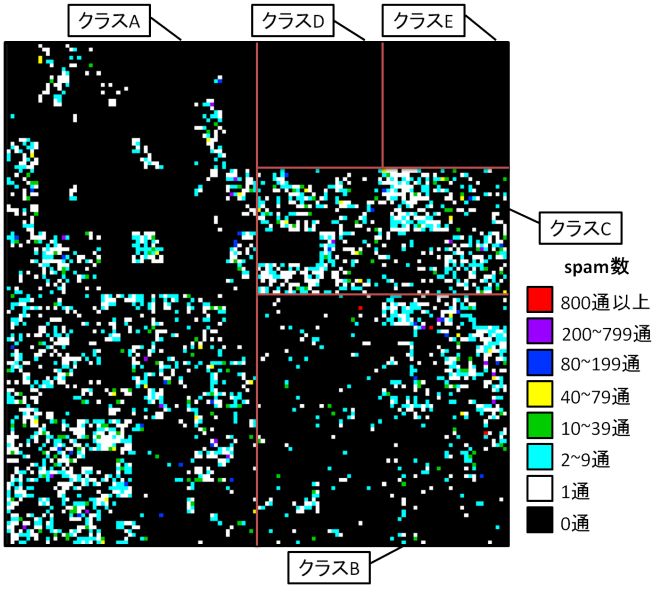

図 6　2014 年 2 月 1 日から 2 月 10 日の spam 送信元として判断された送信元 IP アドレスの分布

ら 2 月 10 日の 10 日間である．権威 DNS サーバのクエリログから MX レコードを要求した送信元 IP アドレス (図 5) とメールサーバにより spam 送信元として判断された IP アドレス (図 6) を抜粋し，それぞれの IP アドレスの第 1 オクテットと第 2 オクテットをヒルベルト曲線により可視化した．

送信元 IP アドレスの上位 16 ビットについて集計した結果，図 5 に存在する IP アドレスは 3475 件，図 6 に存在する IP アドレスは 4280 件であった．また，図 5 と図 6 の分布が一致していた件数は 1515 件であった．それぞれの IP アドレスの上位 16 ビットでの一致箇所を図 7 に示す．

図 7 の黄色の枠で囲まれている部分は IP アドレスブロックが密集していた箇所を示している．特に図 7 の緑の枠で囲まれている 202.0.0.0/8 の IP アドレスブロックに集中していた．この IP アドレスブロックに存在する IP アドレスについて MX レコードを問合せていた IP アドレスは 513 件，spam を送信していた IP アドレスは 195 件，両者の IP アドレスが完全一致していたものは 16 件であった．また，この 16 件の IP アドレスについて逆引きをした結果，NXDOMAIN が 5 件，企業が所有していると考えられるものが 6 件，ISP が所有していると考えられるものが 5 件存在していた．企業が所有していると考えられるものについて，whois により IP アドレスの登録者情報を調査したところ，すべてが日本の企業であり，メールマガジン配信サイトや医療系ポータルサイトが送信元であることがわかった．ISP が所有していると考えられるものについても同様に登録者情報を調査した結果，これらはインド，シンガポール，パキスタン，インドネシア，パキスタンに割り当てられている IP アドレスであった．さらにこのうち，IP アドレスを含んだ文字列と考えられるホスト名が存在していた．ホスト名の例を図 8 に示す．伏せ字の”***”は IP アドレス,”xxxxx”はプロバイダ名,”TLD”は国別トップレベルドメインを示している．図 8 のようなホスト名は，その ISP に所属するエンドユーザに割り当てられるホスト名であると考えられる．エンドユーザから直接 SMTP によりメールが送信される可能性は低く，なおかつ権威 DNS サーバへ MX レコードを問合せていたことから，これらの送信元ホストは，感染したコンピュータが存在するネットワークに依存しない，自らがキャッシュ DNS サーバの機能を備えた spam ボットであることが推測される．

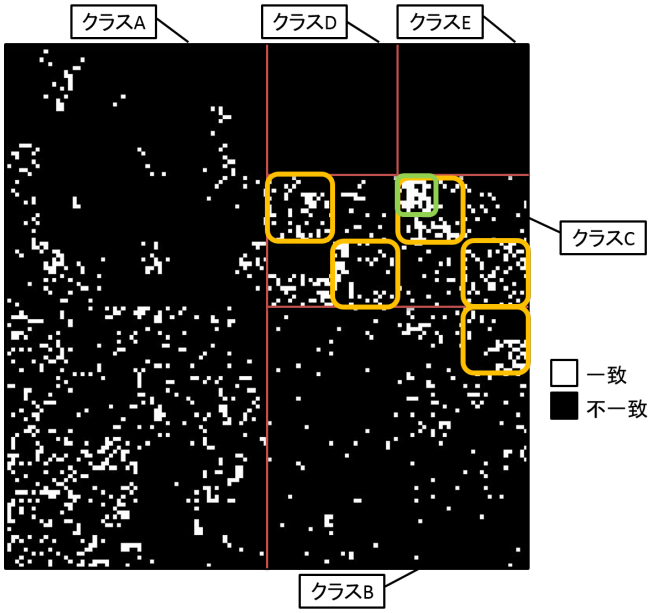

図 7　図 5 と図 6 の一致箇所

そこで，全 IP アドレスブロックの IP アドレスを対象に，MX レコードを問合せた IP アドレスと spam 送信元

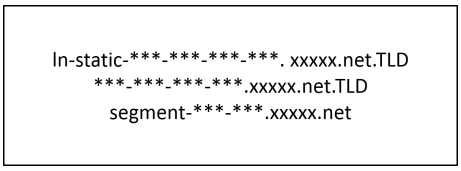

図 8 IP アドレスを含んだ文字列と考えられるホスト名

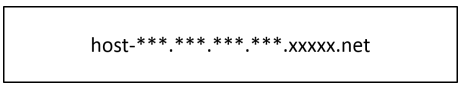

図 9 エジプトの ISP に所属する送信元 IP アドレスのホスト名

と判断された IP アドレスが完全一致するものについて調査した．MX レコードを問合せた IP アドレスは 8753 件，spam の送信元 IP アドレスは 9676 件存在していた．両者の IP アドレスが完全一致するものは 970 件であり，spam の送信元 IP アドレスの約 10 %を占めていた．完全一致していた IP アドレスのうち，エジプトの ISP に所属する送信元 IP アドレスからは権威 DNS サーバへ MX レコードを問合せたのちに spam を送信するという挙動が観測された．MX レコードの問合せ回数は 18 回，spam 送信回数は 97 回であった．この送信元 IP アドレスを逆引きして得られたホスト名は図 9 のような形であり，エンドユーザコンピュータと推測される．完全一致していた IP アドレスすべてについて逆引きをしたところ，SERVFAIL が 20 件，NXDOMAIN が 298 件，メールサーバや企業のホスト名と思われるものが 232 件，IP アドレスと思われる数字列を含んだホスト名が 490 件であった．このことから，MX レコードを問合せた IP アドレスと spam 送信元と判断された IP アドレスが完全一致するものは，spam ボットに感染したエンドユーザコンピュータである可能性が高いと推測される．

4.2 PTR レコードの分析

2014 年 2 月 1 日から 2 月 10 日の 10 日間における日ごとの PTR レコード問合せ数上位 5 件の IP アドレスについて分析した．それらの IP アドレスのクエリ数を図 10 に示す．IP1，IP2 はグルジアの ISP に所属する二つのキャッシュ DNS サーバの IP アドレスであり，2 月 1 日から 2 月 6 日は PTR レコードの問合せ数が多く，2 月 7 日以降は問合せ数は 0 回であった．IP1，IP2 のその他のレコードについては権威 DNS サーバの A および AAAA レコードの問合せのみであった．IP3 は 2 月 8 日に 636 回，2 月 9 日に 43377 回と非常に極端な増加が観測された．IP3 の問合せ内容を調査すると，2 月 8 日，2 月 9 日ともに権威 DNS サーバを A，AAAA レコードで問合せたのちに，大分大学内の IP アドレスを一つずつ順に逆引きするホスト探索活動の挙動が見られた．また，IP3 は日本のある ISP に所属しており，ホスト名は IP アドレスの数字列を含むものであったため，エンドユーザコンピュータからの問合せであると考えられる．通常エンドユーザコンピュータは ISP のキャッシュ DNS サーバを経由して問合せをするが，IP3 は通常とは異なる挙動をしていた．IP4 は ISC(Internet Systems Consortium) が所有する IP アドレスであった．IP4 からは PTR レコード以外の問合せは存在しなかった．IP4 はポート番号を固定にして，大分大学内の IP アドレスを第 3 オクテットを一つずつ順に逆引きする挙動が観測された．また，PTR レコードの 1 日の最大問合せ数 1746 回で数日に分けてホストを探索していた．IP4 は ISC がインターネット上のホスト名を調査する目的でホスト探索をしている可能性がある IP アドレスであったため，通常とは異なる挙動ではあるが異常な問い合わせとは断定できない．IP5，IP6，IP7，IP8 は IP アドレスの上位 24 ビットが一致している送信元であった．PTR レコードの問合せ内容について調査すると，大分大学内の特定セグメントの IP アドレスをランダムに問合せる傾向が観測された．これらの IP アドレスは中国の ISP が所有する IP アドレスであり，すべて逆引きが設定されていなかった．また，IP5，IP6，IP7，IP8 のその他のレコードについては権威 DNS サーバや大分大学内で DHCP 割り当てされている IP アドレスに対する A レコード，AAAA レコードの問合せが日ごとに数件存在していた．

IP1 から IP8 までの送信元 IP アドレスについて分析した結果，PTR レコードの問合せ数が多い外国 ISP 内のキャッシュ DNS サーバや，上位 24 ビットが一致する IP アドレス群からの問合せ，ホスト探索する送信元 IP アドレスを観測したが，日ごとの PTR レコード問合せ数に着目した今回の分析だけでは，ホスト探索攻撃を検知するための明確な検知条件の発見には至らなかった．

5. おわりに

5.1 まとめ

本論文では，権威 DNS サーバのクエリログによって spam 送信やホスト探索活動といった攻撃を発見するための前段階として，MX レコードと PTR レコードに着目した分析をした．MX レコードに関しては，2014 年 2 月 1 日から 2 月 10 日の 10 日間を分析期間とし，権威 DNS サーバの MX レコードを問合せた送信元 IP アドレスと大分大学で運用しているメールサーバにより spam 送信元として判断された IP アドレスについて，それぞれの IP アドレスをヒルベルト曲線により可視化し，比較した．その結果，可視化した上位 16 ビットにおいて両者の IP アドレスが一致していた件数は 1515 件観測され，密集した IP アドレスブロックが存在していた．さらに密集した IP アドレスブロックについて調査したところ，IP アドレスが完全一致するものが 16 件存在し，ISP 内のエンドユーザに割り当てられている可能性のあるホスト名を 3 件発見した．そこ

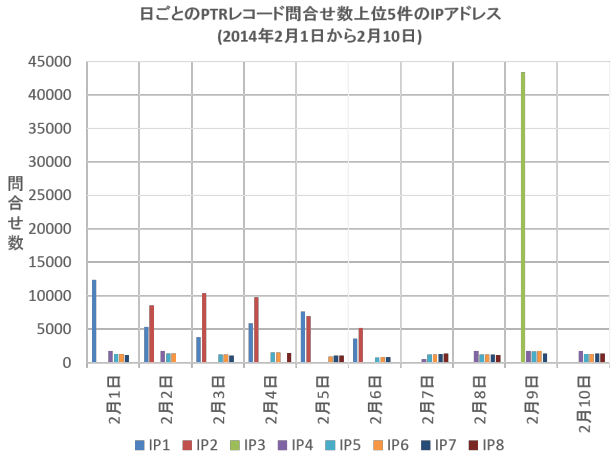

図 10 日ごとの PTR レコード問合せ数上位 5 件に存在した送信元 IP アドレス

で，全 IP アドレスブロックを対象として，IP アドレスが完全一致するものについて調査した結果，970 件の完全一致する IP アドレスを観測した．これらの IP アドレスは逆引きが設定されているものの多くが ISP 内のエンドユーザコンピュータと推測されるホスト名であった．

PTR レコードに関する分析では，2014 年 2 月 1 日から 2 月 10 日における日ごとの PTR レコード問合せ数上位 5 件の送信元 IP アドレスを分析対象とした．それらについて分析した結果，一つの送信元 IP アドレスから 43377 回の PTR レコード問合せがあった．この送信元 IP アドレスの問合せ内容を調査するとホスト探索攻撃と考えられる挙動が観測された．別の送信元 IP アドレスもまた，ホストを探索をしている挙動がみられたが，こちらは数日に分けて探索をしており，1 日あたりの問合せ数が極端に増加する日はなかった．ホスト探索攻撃は，PTR レコードを監視することで検知可能であると考えられるが，問合せ数のみをパラメータとした今回の分析では明確な攻撃検知条件を発見することはできなかった．

5.2 今後の課題

本論文の分析結果より，spam ボットに感染した疑いのあるエンドユーザからの MX レコードの問合せおよび spam 送信を観測した．これらのエンドユーザはホスト名に IP アドレスと思われる数字列を含んでいることが多かった．このことから，メールサーバの spam 対策に用いられる S25R(Selective SMTP Rejection)[9] のようなルールセットを作成することで，エンドユーザから送信される spam を権威 DNS サーバのクエリログから検知することができ，MX レコードの問合せ段階で spam を遮断することが可能であると考えられる．また，今回の分析では，MX レコードの問合せと spam 送信者の IP アドレスを 10 日間という期間で可視化し分析をしたが，より長期間のデータを対象とした場合について分析する必要がある．また，今回は IP アドレスの第 1 オクテット，第 2 オクテットを対象とした上位 16 ビットについて可視化を試みたが，上位 16 ビットの可視化では関係のないネットワークに存在する IP アドレス同士を一致した IP アドレスブロックとして捉えてしまう場合がある．そのため，今後は可視化するビット長を変更することで，新たに得られる特徴量を分析していく．また，今回用いた spam 送信者のデータは，メールサーバ側での様々な spam 対策の過程で検出された送信元 IP アドレスをまとめた状態で使用しているため，どの spam 対策の段階で検出されたものかまでは考慮していない．そのため，各 spam 対策の段階で検出された IP アドレスと，権威 DNS サーバへの MX レコードの問合せもしくはその他のレコードの問合せに違いが現れるか調査する．

PTR レコードの分析では極端なホスト探索攻撃をする送信元 IP アドレスが観測されたが，ホスト探索を検知するための特徴量としてはまだ不十分である．今後は問合せ数だけではなく，別のパラメータを考慮した分析をしていくことで検知基準の発見を目指す．

本論文では MX レコードと PTR レコードを対象としたが，その他のレコードを分析することで検知可能な攻撃が考えられるため今後分析していく．

参考文献

[1] 情報技術解析平成 25 年報～平成 25 年中のインターネット観測結果等～，入手先 (https://www.npa.go.jp/cyberpolice/detect/pdf/H25_nenpo.pdf)(参照 2014 年 4 月 9 日)
[2] 山井成良，岡山聖彦，宮下卓也，繁田展史，丸山伸，中村素典，”発信元詐称 spam メールに起因するバウンスメール集中への対策方法”，情報処理学会論文誌，Vol.47，No.4，pp.1010-1020，2006 年 4 月
[3] milter を用いた効果的な迷惑メール対策，入手先 (http://milter-manager.sourceforge.net/)(参照 2014 年 4 月 5 日)
[4] 松井一乃，金高一，池部実，吉田和幸，” milter manager を用いたメールサーバの運用における導入の効果”，マルチメディア、分散協調とモバイルシンポジウム 2013 論文集 ，pp.772-778，2013 年 7 月
[5] B.Irwin，N.Pilkington，“High Level Internet Scale Traffic Visualization Using Hilbert Curve Mapping”， proceedings of the VizSEC 2007，Mathematics and Visualization，Springer Berlin Heidelberg，pp147-158，2007 年 10 月
[6] R.Munroe，Map of the Internet，入手先 (http://www.xkcd.com/195/)(参照 2014 年 4 月 3 日)
[7] C.Muelder，KL.Ma，” Rapd Graph Layout Using Space Filling Curves”，Proceedings of the IEEE Transactions Visualization and Computer Graphics 2008，Vol.14，No.6，pp.1301-1308，Nov.2008
[8] デニス・アルトゥロ・ルデニャ・ロマニャ，杉谷賢一，久保田真一郎，武藏泰雄，”DNS によるスパムボットとホスト探索活動の検知”，情報処理学会研究報告，2008-IOT-3，pp.1-6，2008 年 9 月
[9] 阻止率 99 % のスパム対策方式の研究報告，入手先 (http://www.gabacho-net.jp/anti-spam/anti-spam-system.html)(参照 2014 年 4 月 5 日)