DNSログ解析によるDGAを用いたマルウェア検知のための予備調査

渡辺 拳竜† 池部 実‡ 吉田 和幸§

これまで，ボットネットによる攻撃行為やspam送信などのマルウェアによる悪意ある活動を検知するため，DNSサーバのクエリログを用いた検知手法を検討してきた．今回，大分大学にてDGAと呼ばれるマルウェア特有の名前解決問合せを観測した．DGAを用いて，ドメインブラックリストやURLフィルタリングを回避するマルウェアが増加している．また，DGAを用いるマルウェアが学内で発生した場合，我々がこれまでに提案した送信元IPアドレスの名前解決要求数によるマルウェア感染端末検知手法では検知できない．そこで，本論文ではDGAによって生成されたと考えられるFQDNを問合せていた送信元についてクエリログを用いて調査し，DGAを用いるマルウェアを検知するために実施した予備調査の結果を報告する．

A preliminary study for DGA-based malware detection by DNS log analysis

Kenryu WATANABE† Minoru IKEBE‡ Kazuyuki YOSHIDA§

1 はじめに

我々の生活には，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が不可欠な存在になっている．インターネットを利用するためには，DNS(Domain Name System)による名前解決を欠かすことができない．一方，攻撃者の通信，例えば，ボットネットによる攻撃行為やspam送信などの悪意ある活動においてもDNSによる名前解決は生じる．そのため，DNSサーバのクエリログには正規のユーザの問合せと悪意ある活動のための名前解決の問い合わせ要求が混在している状況にある．

我々は，これまでにDNSサーバのクエリログを分析することにより，悪意ある活動を検知することを目的として研究を進めてきた．キャッシュDNSサーバのクエリログを調査する過程で，DGA(Domain Generation Algorithm)と呼ばれるマルウェア特有の名前解決要求を観測した．そこで本研究では，DGAの問合せについての調査と今後の対策について検討した結果について報告する．

† 大分大学大学院工学研究科知能情報システム工学専攻
‡ 大分大学工学部知能情報システム工学科
§ 大分大学学術情報拠点情報基盤センター

2 マルウェアによるDGAの挙動と特徴

一部のマルウェアは，接続先ドメイン名生成の仕組みであるDGAを用いて，ランダムな英数字で構成されたFQDNを大量に生成して，MDL(Malware Domain List)[1]に代表されるドメインブラックリストやURLフィルタリングを回避するマルウェアが増加している[2].

これまでに，我々は，キャッシュDNSサーバのクエリログを用いたIPアドレスごとのFQDN問合せ数によるマルウェア検知を提案してきた[3]．しかし，DGAを用いるマルウェアはFQDNを使い捨てるため同じFQDNを繰り返して問合せることがなく，FQDN問合せ数では検知できない．また，DGAを用いるマルウェアが生成するFQDNの大半は対応する情報が存在しない．そこでクエリログに記載されているFQDNを用いて，FQDNに対応する情報が存在しない場合の回答を表すNXDOMAINを収集・分析することで，DGAを用いるマルウェアの検知に有効な特徴の発見を目指す．

3 DGAを用いるマルウェアによる名前解決問合せの調査結果

2015年1月19日から28日までの10日間の大分大学のキャッシュDNSサーバのクエリログについて調査した．1つの送信元から”uljvwkdihuyyi.tj”

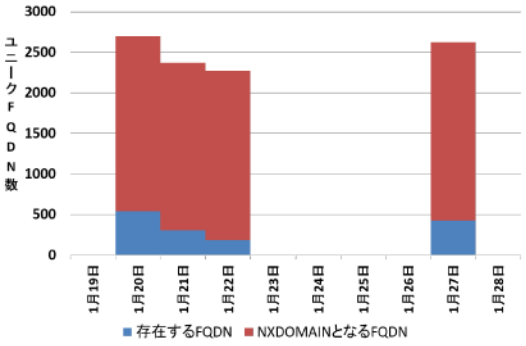

図 1　DGA ホストが問合せたユニーク FQDN 数 (2015 年 1 月 19 日から 1 月 29 日)

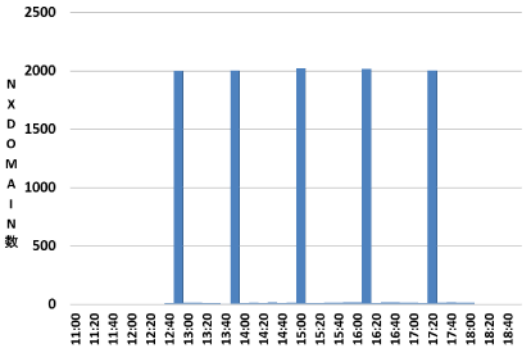

図 2　DGA ホストの NXDOMAIN 数 (2015 年 1 月 27 日)

や”bsmvtddhrdr.xxx” など DGA により生成されたと考えられる問合せを観測した (以下，この送信元を DGA ホストと呼ぶ)．10 日間における DGA ホストが問合せたユニーク FQDN 数を図 1 に示す．DGA ホストから発生する NXDOMAIN となる問合せは，すべて DGA により生成したと考えられる英数字の羅列を含んだ FQDN であった．また，NXDOMAIN となる問合せが最大であった 2015 年 1 月 27 日について，DGA ホストの NXDOMAIN となる問合せ状況を調査した結果，約 60 分間隔で DGA の FQDN を問合せる挙動が観測された (図 2)．

さらに，2015 年 1 月 27 日の全ユーザからの問合せを 1 時間ごとに集計し，問合せ数が最も多かった 12 時からの 1 時間について，IP アドレス別に NXDOMAIN となる問合せの回数を調査した．1 時間における NXDOMAIN となる問合せ数上位 5 位を示す (表 1)．IP1 は DGA ホストで，IP2 から IP5 は”wpad.iptvf.jp” やホスト名のみの問合せといった各端末の設定に起因する無効な問合せであり，DGA により生成されたと考えられる FQDN が存在しない送信元であった．IP2 から IP5 を正規ユーザと考えると，NXDOMAIN となる問合せは急激に増加することがなく，1 時間における NXDOMAIN となる問合せ数は多くとも 200 回程度であった．それに対し，DGA ホストの NXDOMAIN となる問合せ回数は正規ユーザの挙動を大きく上回っていた．

表 1　IP アドレス別の NXDOMAIN となる問合せ数 Top 5(12:00〜12:59)

IP アドレス	NXDOMAIN の回数
IP1	1,995
IP2	217
IP3	142
IP4	71
IP5	59

4　おわりに

DNS サーバのクエリログに記載された各種パラメータを用いて，悪意ある活動を発見するための手掛かりを調査した．本論文では，キャッシュ DNS サーバのクエリログを用いて，DGA によって生成した FQDN を問合せる送信元に関する調査をした．その結果，定期的なタイミングで DGA の FQDN を問合せる挙動や突発的に大量の NXDOMAIN となる問合せをする挙動が観測できた．

NXDOMAIN となる問合せは，キャッシュ DNS サーバにてネガティブキャッシュとして保持される．DGA を用いるマルウェアは活動の際に DGA の FQDN を多く問合せるため，ネガティブキャッシュが増加すると考えられる．そこで，ネガティブキャッシュの変化率を用いて DGA を用いるマルウェア検知手法を検討する．

参考文献

[1] MDL (Malware Domain List). www.malwaredomainlist.com.

[2] 2015 年 サイバー脅威と動向. http://www.verisign.com/assets/report-iDefense-trends-2015.PDF.

[3] 渡辺 拳竜, 池部 実, 吉田 和幸. DNS クエリログのクエリ数に着目した異常ホストの検出. 情報処理学会マルチメディア，分散，協調とモバイル (DICOMO2014) シンポジウム論文集, pp. 197–204, 2014 年 7 月.