リフレクター攻撃における増幅器探索通信の解析

芳賀 夢久 †　笹生 憲 †　森 達哉 †　後藤 滋樹 †

† 早稲田大学　基幹理工学部
169-8555　東京都新宿区大久保 3-4-1
{yumehisa.haga,saso,mori}@nsl.cs.waseda.ac.jp, goto@goto.info.waseda.ac.jp

あらまし DNS や NTP 等の UDP を用いた通信サービスを濫用したリフレクター攻撃が世界的に増加傾向にある．大規模なリフレクター攻撃を実施する事前準備として，攻撃者は帯域増幅器となるホストを探索する．本研究ではそのような探索通信に着目し，帯域増幅器の探索を実施している組織やツールの弁別を狙いとする．弁別により攻撃者の動向が明瞭になることが期待できる．UDP はヘッダーの情報量が少なく得られる情報が限定的であるが，IP アドレスと通信パターンの解析に加え，リフレクター攻撃に使われるプロトコルのヘッダー情報を解析することでより多くの情報を獲得可能である．本研究ではアプリケーションヘッダ情報から増幅器探索通信に特徴的なフィンガープリントを症例対照研究のアプローチにより抽出する．IP アドレス・通信パターン分析と抽出したフィンガープリントをダークネットおよび実運用中のネットワークに適用した結果，増幅器探索通信の多くは学術・研究機関や企業等による研究・調査目的の活動であることが明らかになった．それらの通信をフィルターすることにより，攻撃者の活動トレンドがより明瞭になることが期待できる．

Analysis of amplifier probing for reflector attacks

Yumehisa Haga †　Akira Saso †　Tatsuya Mori†　Shigeki Goto

†School of Fundamental Science and Engineering, Waseda University
3-4-1 Okubo, Shinjuku-ku, Tokyo 169-8555, JAPAN
{yumehisa.haga,saso,mori}@nsl.cs.waseda.ac.jp, goto@goto.info.waseda.ac.jp

Abstract Reflector attacks, which leverage UDP-based services that reply to queries without requiring prior connection establishment, are becoming increasingly serious threats. It has been reported that adversaries who intend to employ large-scale reflector attacks perform probing *bandwidth amplifiers* as an initial setup of the attack. This work focuses on such probing activities. We aim to classify organizations or tools that perform such probing by analyzing one-way traffic from senders. Such classification enables us to extract actual attack threats because it can eliminate potential noises generated by *legitimate* probing employed by organizations such as research institutes. Because UDP headers has fewer information, compared to TCP headers, we develop a method that extracts intrinsic characteristics by analyzing protocols of services that are used for reflector attacks. We report our analysis results using traffic collected from both darknet and campus network.

1　はじめに

背景: 2014 年 2 月 11 日，DDoS 攻撃によるトラフィック流入量の世界記録 400Gbps が樹立された [1]．この流量は日本におけるブロードバンド全契約者の総ダウンロード流量の約 15%[1] に匹敵する超大規模な攻撃であり，その標的は世界中で 1 億以上のユーザが利用するノートサービスを提供する Evernote

[1] 2013 年 11 月時における我が国のブロードバンド契約者の総ダウンロードトラヒック総量は 2.6 Tbps [2].

であった．この超大規模の DDoS 攻撃により，多くの利用者がサービス利用不能になったことが報告されている [1]．上述の DDoS 攻撃には NTP 増幅攻撃が利用された．この攻撃では IP アドレスを詐称したバックスキャッタ通信を利用すること，コネクション確立が不要な UDP を用いること，クエリ通信量と比較してレスポンスサイズが格段に大きいことが特徴であり，一般にリフレクター攻撃と総称される．

リフレクター攻撃は古くから知られた攻撃方法であり，環境が整えば比較的単純な方法によって大規模化が可能である．大規模なリフレクター攻撃を行うためには，まずオープンな状態で増幅器として機能するホストの候補を探索し，大量にリストアップする必要がある．増幅器を探索する単純な方法は全 IP アドレス空間を対象として，探査パケットを生成する事である．そのような全 IP アドレス空間を対象とした探査活動の存在は古くから報告されており [3]，ダークネットにおいても広く観察されている [4]．

一方，超大規模なリフレクター攻撃の顕在化はネットワークオペレーターや研究者の関心を大いに誘引した．その結果，増幅器として動作するホストの調査や新たなリフレクター攻撃の実現可能性の検証等を目的として，増幅器の探索活動が活発に行われている．例えば [5, 6] はインターネット上のあらゆるホストに対してサービスがオープンな状態の DNS サーバおよび NTP サーバを収集・調査するプロジェクトである．また Rossow が発表した文献 [7] では DNS，NTP も含めた 14 個の異なる UDP 上のアプリケーションプロトコルを対象とし，増幅器の探索と攻撃可能性の検証分析を行っている．

以上のように，リフレクター攻撃を行う攻撃者のみならず，調査や防衛手段の考案を目的とした増幅器の探索活動が増加傾向にある．この結果，悪性通信のみが観測されると仮定されてきたダークネットでは攻撃者の探索通信に加え，上記のような調査目的の探索通信がある種の**ノイズ**として印加されるようになった．

狙いとアプローチ: 本研究の目的は増幅器の探索通信を解析し，探索を実施している組織やツールの弁別を試みることにある．攻撃者の活動とそれ以外のノイズを弁別することにより，攻撃者による攻撃準備等の動向把握がより明瞭になることが期待できる．以上の目的を達成するためのアプローチとして，2 つの方法を適用する．一つは送信元 IP アドレスの属性と通信パターンの利用，もう一つはアプリケーションヘッダ解析に基づくフィンガープリンティングである．後者はリフレクター攻撃に使われるプロトコルのヘッダー情報を解析し，探索に特徴的なパラメタ設定値等を自動的に抽出し，フィンガープリントとして出力する技術である．提案方法を評価するために /20 アドレス空間で運用されるダークネット (NICTER Dataset 2014) および /16 アドレス空間で運用されるキャンパスネットワークで計測した通信データを用いる．

貢献: 本研究の主要な貢献を下記に示す．

- 増幅器探索通信の弁別を可能とする統計的フィンガープリント抽出技術を開発した．
- ダークネットにおいて観測される増幅器探索通信ホストの約半数は調査プロジェクトもしくは学術機関による通信であり，これらは弁別可能である．

これらの発見はリフレクター攻撃に伴う探索活動の検出と攻撃予兆の発見に有益である．また，本研究で開発した 統計的フィンガープリント抽出は汎用的な技術であり，他の応用に資することが期待できる．

本論文の構成は以下の通りである．はじめに 2 章で関連研究を示した後，3 章で提案方法の概要を述べる．4 章で評価に利用したデータを説明し，5 章で分析結果を述べる．最後に 6 章にて本論文のまとめと今後の展望を述べる．

2 関連研究

本章では，リフレクター攻撃およびフィンガープリントに関連するいくつかの研究を述べる．

Rossow らは文献 [7] でリフレクター攻撃に用いられる 14 種のプロトコルについて，応答の際の増幅率を算出し，DDoS 攻撃の危険性を主張している．また，ISP の通信データを元に実際に起きているリフレクター攻撃の実態についても言及している．

中里らはダークネットで観測した DNS パケットのクエリについて分析している [8]．例えば，DNS ヘッダーの QR フィールドを参照することで，そのパケットが DNS 問い合わせなのか DNS 応答なのかがわかる．また，DNS スキャンに用いられる問い合わせ元のドメインの特徴についても報告している．

3 提案方法

本章では増幅器探索通信を分類するための提案方法として，IP アドレス属性と通信パターンを用いた方法，および統計的フィンガープリント抽出技術の概要を示す．

3.1 IP アドレス属性と通信パターン

IP アドレスの属性として，大学や研究機関等の IP アドレスを収集した Block list を利用する．この

リストは I-Blocklist [9] が世界の様々な組織の IP アドレスのリストを WEB 上で公開しているものである．本研究では，この中から 4 つのリストを選択し，ダークネットにパケットを送ってきたホストとマッチングを行った．Blocklist の詳細を表 1 に示す．

また通信パターンとしてスキャンを実施していることが明らかな探索ホストを抽出する．ここでは単一のホストがネットワーク全域を探索するパターン (Single Scanner) と，複数のホストが協調してネットワーク全域を探索するパターン (Coordinated Scanner) に分類して考える．Single Scanner の抽出は単純であり，発 IP アドレス毎のユニークな通信先 IP アドレスの数が観測しているネットワークの全アドレス数と一致したホストを抽出する．Coordinated Scanner の検出では発 IP アドレスを /24 Prefix 単位で集約し，同一 Prefix に属するホスト数が多く，かつそのホスト群が協調してスキャン通信を行っているものを抽出する．

3.2 統計的フィンガープリント抽出

統計的フィンガープリント抽出の主要なアイディアは症例対照研究 (Case-control study) のアプローチをとることにある．2 つの異なる症例に対応するグループとして増幅器探索ホスト等の通信が多いグループと，通常ホストが多いグループを考え，前者をダークネットの観測ホストとし，後者を通常のユーザが利用するネットワークの観測ホストと仮定する．通常のネットワークには探索ホストも混在しているが，通常のホスト数が圧倒的に大きいためこのような近似の仮定は妥当である．両症例間においてある特徴の出現頻度に顕著な差が認められる場合，その特徴をフィンガープリントとして抽出する．

上記のような両症例におけるある特徴の出現の差が統計的に顕著であるかを判定するための手段としてオッズ比を用いる．オッズ比は 2 つの異なるグループにおける特定の事象の起こりやすさの度合いを定量化するための尺度であり，臨床疫学等の医療統計分野において広く利用されている．本研究の文脈では特にオッズ比が高い特徴は，探索ホストを特徴付けるフィンガープリントとして活用することが期待できる．

表 2 に示したノーテーションにしたがってオッズ比の計算方法を示す．表中の a, b, c, d はそれぞれのカテゴリに属するユニークなホストの数である．はじめにある増幅器探索ホストが特徴 X を有する確率 p は $p = a/(a + c)$ と計算できる．同様に通常ホストが特徴 X を有する確率 q は $q = b/(b + d)$ と

表 2: 特徴とホスト種別の分類表．

	ダークネット	通常ネットワーク
特徴 X を持つ	a	b
特徴 X を持たない	c	d

計算できる．オッズ比 OR は p, q を用いて

$$OR = \frac{p}{1-p} \Big/ \frac{q}{1-q} = \frac{ad}{bc} \tag{1}$$

と計算できる．

標本から算出したオッズ比の妥当性を定量化するために良く用いられる 95 % 信頼区間およびカイ二乗検定を用いる．本研究ではある特徴 X の出現に関してオッズ比の信頼区間の下限が 1 よりも大きく，かつカイ二乗検定による p 値が有意水準 0.05 よりも小さい場合に特徴 X は両群において有意差を持つフィンガープリントであると判定する．

本研究で用いる特徴は，例えばあるアプリケーションヘッダ中のフラグ値が特定の数値になっているか，クエリ値がある特定の文字列になっているかという情報である．本研究では特徴となる候補をれぞれのプロトコルの仕様および実装コードで定義されたヘッダの構造体を参考にして，あらゆる組み合わせで列挙した．

3.3 フィンガープリントの種類と制限

増幅器探索通信を特徴づけるフィンガープリントをホストフィンガープリントおよびクエリフィンガープリントの 2 種類にわけて考えることにする．ホストフィンガープリントはパケットを発生するホスト (OS やツール) で決まる固定長の特徴であり，主にヘッダ上のパラメタ値がそれに該当する．クエリフィンガープリントは DNS のクエリ文字列が代表例であり，パケットを発生するツールや組織によって決まる可変長の特徴であり，値の自由度が高い．

前章で述べた統計的フィンガープリント抽出はダークネットと通常ネットワークにおいて著しい差が存在することが前提となっている．DNS，NTP，SNMP (LAN 内通信を含むものとする) に関しては通常ネットワークにおける通常の利用が多いため，両者で差が生じることが期待できる．すなわち，症例対照研究のモデルに適合する．一方，chargen や NetBios 等，正常ネットワークにおいて，そのプロトコルの通常の通信がほとんど観測されないような場合は両グループによる差を期待することができないため，症例対照研究のモデルに適合しない．本研究では後者のようなケースにおいては，特に数値的に頻度が高い特徴をマニュアルで調査するアプローチをとる．

表 1: 使用した Blocklist

List name	Number of IPs	Description
Level1	787849409	Companies or organizations who are clearly involved with trying to stop filesharing.
Level2	355564796	General corporate ranges. Ranges used by labs or researchers. Proxies.
Level3	119630195	Many portal-type websites.
Edu	235111433	Contains all known Educational Institutions.

表 3: ダークネットにおけるプロトコル毎のパケット数およびホスト数.

プロトコル	ポート番号	パケット数	ホスト数
DNS	53	6,054,644	1,047
NTP	123	1,570,929	301
SNMP	161	1,571,592	154
Chargen	19	2,544,263	289
Netbios	137	1,188,119	1,129

表 4: キャンパスネットワークにおけるプロトコル毎のパケット数およびホスト数.

プロトコル	ポート番号	パケット数	ホスト数
DNS	53	33,597,847	76,871
NTP	123	2,257,449	1,352
SNMP	161	5,038,533	782
Chargen	19	5,641,617	185
Netbios	137	467,547	1,584

4 解析データとその特徴

本研究では，ダークネットおよび正常ネットワークの 2 種類の通信データを用いる．UDP によるリフレクター攻撃でよく用いられるプロトコルとして DNS, NTP, SNMP， Chargen, Netbios の 5 つを分析の対象とする．実際，これらのプロトコルは本研究で解析するダークネットにおいてもパケットの流量が最も多かったものである．

4.1 ダークネット

本研究では，ダークネットの計測データとして NICTER Darknet Dataset 2014 [10] を利用する．NICTER Darknet Dataset は NICTER の/20 アドレスのダークネットで計測されたトラヒックデータである．本研究では，2014 年 1 月 1 日から 2014 年 7 月 31 日までの 7ヶ月間のデータを分析対象とする．表 3 にダークネットにおけるプロトコル毎のパケット数およおび異なるホスト数を示す．ダークネットに届く UDP パケット数が最も多いのは DNS である．また NTP パケットは近年増加傾向にある．

4.2 キャンパスネットワーク

ダークネットの不正な通信と比較するための正常ネットワークにおける通信として，大学で実運用中の/16 アドレスのネットワークの通信データを用いる（以降キャンパスネットワークと呼ぶ）．対象データは，大学のゲートウェイで計測されたトラフィックデータであり，2014 年 7 月 24 日から 7 月 28 日にかけて，pcap フォーマットで計測したパケットキャプチャデータである．本研究では，ダークネットで解析対象とした 5 つのプロトコルに絞ってデータを収集した．表 4 にキャンパスネットワークにおけるプロトコル毎のパケット数および異なる送信元ホスト数を示す．

5 結果

本章では 3 章で提案した手法を用いて通信データを解析した結果を示す．

5.1 IP アドレス属性と通信パターン

まず，Blocklist を用いて，組織の IP アドレスとダークネットにパケットを送信したホストをマッチングした結果を示す．図 2 に 5 つのプロトコル及び UDP パケット全体に関して Blocklist とのマッチング結果を示す．リフレクター攻撃に用いられるプ 5 つのプロトコルに関して，一定の割合でリストと一致する IP アドレスが存在することが分かる．また UDP 全体に比べて，これらのプロトコルにおけるリストの一致率が高く，特に DNS においては，リスト Edu つまり教育機関からの通信の割合が大きい．このことから，様々な組織にとってリフレクター攻撃が関心の対象となっていることが推測できる．

次に通信パターンについて解析した結果を示す．図 1 は，各送信元ホストが送った宛先 IP アドレスの数を示す．1 ホストにつき，1 IP に向けてパケットを送っている通信が多いが，同時に，ダークネット全体つまり/20 の 4096 個の IP アドレスに向かってパケットを送った通信 (フルスキャン) も多いことが分かる．

ここで 4096 個の IP に向けられたフルスキャンについて着目する．ダークネットの 53 番 (DNS) ポー

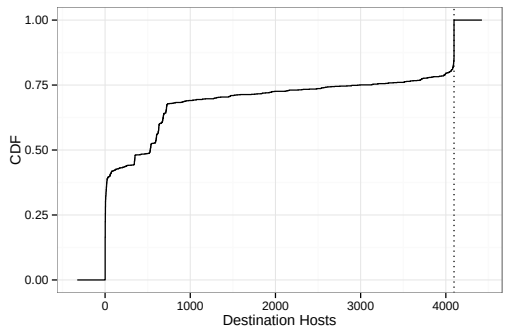

図 1: ホスト毎の送信先 IP 数.

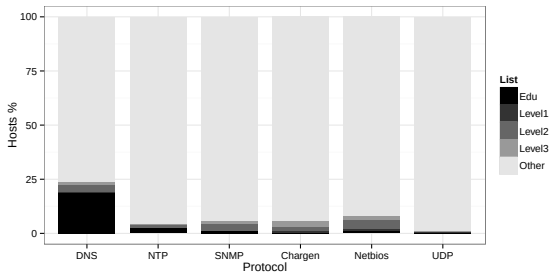

図 2: Blocklist とのマッチング結果.

トへの通信を解析した結果，一つのホストがフルスキャンを行ったもの (Single Scanner) は，53 番ポートにパケット送った全 1047 ホストの内，156 ホストあることがわかった．また，ダークネットでは，複数のホストが協調してスキャンをかけてくる通信 (Coordinated Scan) も確認した．これは，スキャン処理の負荷分散や，宛先側での IP ベースのブロックを回避するためのものと考えられる．Coordinated Scan の特徴として，同じ時期に同一の Prefix のネットワークからパケットが送られくるとこと，ホスト毎のパケットのヘッダー情報がそれぞれ酷似していることなどが挙げられる．さらにこのような特徴を持つ通信に着目したとき，いくつかの通信パターンに分けられることがわかった．Coordinated Scan の通信パターンの詳細を表 5 に示す．

また，Prefix ごとにフルスキャンを行っているものを抽出し，送信元ホスト数の多いものをいくつか挙げ，表 5 の通信パターンと照らしあわせたものを表 6 に示す．もっともホストが多かった 180.76.4.0/24 は中国の Baidu [11] の IP であり，4 月中旬から 7 月にかけてホスト毎に宛先 IP を分担し，大規模なスキャンを行っていた．通信パターン 2 のものは，ホスト毎にヘッダー情報を微妙に変化させてスキャンを行っている様子が見受けられた．また，パターン 3 のものは，パターン 2 の状態から，途中でパターン 1 の分散型のスキャンにシフトしている様子が一部確認できた．

表 5: Coordinated Scan の通信パターン分類.

パターン	挙動
パターン 1	それぞれのホストが宛先 IP を分担してスキャンを行う.
パターン 2	すべてのホストがフルスキャン，またはそれに近いものを行う.
パターン 3	ある特定のホストのみがフルスキャンを行い，他のホストは一部の IP に対してスキャンを行う.

表 6: Coordinated Scan を行った IP Prefix (/24).

Prefix	ホスト数	パケット数	通信パターン
180.76.xx.0/24	228	143,278	パターン 1
184.105.xx.0/24	18	192,887	パターン 3
74.82.xx.0/24	15	455,732	パターン 2
89.248.xx.0/24	12	70,015	パターン 3
93.174.xx.0/24	10	65,201	パターン 3
80.82.xx.0/24	9	56,307	パターン 3
93.174.xx.0/24	8	63,079	パターン 3
94.102.xx.0/24	8	109,051	パターン 2

5.2 フィンガープリントの抽出

4 章で述べた方法でフィンガープリントを抽出した結果を示す．ダークネットとキャンパスネットワークの通信データのアプリケーションヘッダを比較し，顕著な差が見られたものを抽出する．表 7 は DNS, NTP，SNMP の 3 つのプロトコルについて抽出したホストフィンガープリントをまとめたものである．プロトコル名の下の括弧で括られている数字は，分析の対象としたフィールドの数を表している．

また，スキャンツールとして知られている Nmap, Zmap, Metasplot の 3 つのツールについて通信データを解析したところ，それぞれのアプリケーションヘッダに固有の値が見られたため，フィンガープリントの結果と照らしあわせた．その結果，ダークネットに特有な通信の中に，これら 3 つのスキャンツールを使用したと思われる痕跡を見つけることができた．スキャンツールを特定できなかったものに関しても，何らかの形で攻撃者の間でツールのコードが共有されている事も考えれる．

さらにこのフィンガープリントにより，ダークネットにおいて探索通信後のリフレクター攻撃まで意識したような通信も確認できた．例えば DNS におい

て，RD フィールドが 1 のものは，DNS の再帰検索 (recursive) を要求するものであり，オープンリゾルバの探索が可能である．また NTP において，VN フィードが 0, 1, 2 のものは NTP の古いバージョンを狙ったものであり，Mode フィールドが 6 や 7 のものは通常使われないが，問い合わせに対する応答のデータ量が大きいため，DDoS 攻撃に有効であることが知られている．

次に，DNS の問い合わせドメインに注目したクエリフィンガープリントの結果を示す．ダークネットに特有なドメインとみなす基準は前述のホストフィンガープリントと同様であるが，クエリフィンガープリントに関しては，ホスト数とパケット数それぞれに関して分析を行う．

ここで，ダークネットの DNS の問い合わせドメインを分析すると，ドメインの一部が問い合わせ毎に変化するもの (Variable Domain) 存在することが分かった．変化する部分は下位レベルのドメインで，変化の仕方は日付や IP アドレス，ランダムな数字など様々である．この Variable Domain を抽出するために，全てのドメインを”.” (ドット) で区切り，上位レベルドメインを親とする木構造を形成し，直下の子ノードが 10 個以上に分岐するパターンを Variable Domain とみなす手法をとった．これにより，Variable Domain の変化する部分を ”∗” でマスクし，同一パターンのドメインとみなすことができる．

ホスト数に注目したクエリフィンガープリントの結果を表 8 に，パケット数に注目したクエリフィンガープリントの結果を表 9 に示す．クエリフィンガープリントの結果と DNS リフレション攻撃でよく用いられる問い合わせドメインのリスト [12] (以降 Domain BL) と照らしあわせ，さらにドメイン名からスキャンツールの判別も行った．その結果，抽出されたドメインのほとんどが，Domain BL のものと一致した．だだし，その中には”dnsscan.shadowserver.org” など，明らかに何らかの組織によるサーベイ目的のスキャンであるものも含まれている．また，Domain BL に含まれなかった”∗.w.shifen.com” というドメインは，4.1 章で言及した Baidu からの問い合わせドメインとなっている．ホスト数とパケット数の両方について分析することで，一方で抽出できなかったドメインをもう一方で抽出することにも成功した．

最後に，Chargen と Netbios に関してダークネットの通信データを解析した結果，ツールを判別できたものを表 10 に示す．プロトコル名の下の括弧で括られている数字は，分析の対象としたフィールドの数を表している．

表 10: Chargen と Netbios のフィンガープリント．

Protocol	Fields	Hosts	Tool
Chargen	Data = 47:45:54:20:..	9	Zmap
(1)	Data = NULL	19	Nmap
Netbios	ID = 0x80f0	17	Nmap
(2)	flag = 0x0000	811	Metasploit

表 11: 増幅器探索ホストの分類．

分類	説明
Legitimate	宛先の IP 数が 10 未満のもの．
Survey	Legitimate ではなく，かつ Blocklist の IP と一致したかクエリフィンガープリントによりサーベイ系と判断したもの．
Scan	Legitimate でもなく，Survey でもないもの．

5.3 増幅器探索通信の内訳

これまでの手法を用いて，ダークネットとキャンパスネットワークへ通信を行ったホストを，サーベイ活動によるスキャン (Survey) を行ったホスト，サーベイ活動ではない悪性なスキャン (Scan) を行ったホスト，スキャンではない通信 (Legitimate) を行ったホストの 3 つに分類した．分類方法の詳細を表 11 に示す．

紙面の都合上，53 番 (DNS) ポートにパケットを送信したホストのみ解析の対象とする．解析の期間は，ダークネットは 2014 年 1 月 1 日から 2014 年 7 月 31 日の 7ヶ月間，キャンパスネットワークは 7 月 25 日金曜日の 8 時から 20 時までの間とした．ダークネットとキャンパスネットワークの通信の解析結果を図 3， 4 に示す．なお，ダークネットにおいて，Legitmate に相当するホストは ”Other” という表現を用いた．

図 3 より，ダークネットにおいて，様々な組織によるサーベイ目的のホストの割合が高く，およそ半数から時期によっては 9 割以上がサーベイ目的のホストであった．またこのグラフにも，4 月中旬から 7 月にかけて急増した Baidu からの通信も反映されている．図 4 より，キャンパスネットワークでは Legitimate な通信を行うホストが大部分を占めていることがわかる．また，ダークネットの通信に比べ，Survey に対する Scan の割合が大きくなっている．これは，実運用中であるキャンパスネットワークのほうが，攻撃の対象になりやすいためではないかと推察できる．

表 7: ホストフィンガープリント.

Protocol	Fields	Darknet	Campus	Odds ratio	95% CI	P-value	Note	Tool
DNS (9)	Flag = NULL	38	4	723.72	257.82 - 2031.54	<0.001		
	QR = 1	13	16	60.39	28.98 - 125.87	<0.001	Massage is not response	
	Opcode = 10	28	2	1056.10	251.25 - 4439.14	<0.001		Zmap
	AA = 1	34	27	95.52	57.42 - 158.92	<0.001	Server is an authority for domain	Zmap
	RD = 1	958	4765	162.89	130.83 - 202.8	<0.001	Do query recursively	Nmap, Metasploit
	RA = 1	11	12	68.01	29.94 - 154.47	<0.001	Server can do recursive queries	
	Z = 0	959	45827	7.38	5.93 - 9.19	<0.001		Nmap, Metasploit
	Z = 2	43	8	411.49	192.98 - 877.45	<0.001		Zmap
	Rcode = 7	11	1	816.19	105.28 - 6327.61	<0.001	A resource record set exists that should not	
	ID = 0x4567	76	115	52.24	38.83 - 70.28	<0.001		Zmap
	ID = 0x0000	49	164	22.96	16.59 - 31.79	<0.001		Nmap
NTP (10)	Flag = NULL	14	24	2.70	1.38 - 5.28	0.03		
	VN = 0	27	1	133.13	18.01 - 983.84	<0.001		Zmap
	VN = 1	28	6	23.01	9.44 - 56.1	<0.001		
	VN = 2	187	372	4.32	3.33 - 5.61	<0.001		
	Mode = 6	23	17	6.50	3.43 - 12.32	<0.001	NTP control message	
	Mode = 7	188	351	4.74	3.65 - 6.17	<0.001	Reserved for privateuse	Zmap
	Stratum = NULL	231	404	7.74	5.78 - 10.37	<0.001		Zmap
	Poll Intervall = NULL	231	404	7.74	5.78 - 10.37	<0.001		Zmap
	Reference ID = NULL	231	404	7.74	5.78 - 10.37	<0.001		Zmap
	Reference time = NULL	231	406	7.69	5.74 - 10.29	<0.001		Zmap
	Originate time = NULL	232	406	7.83	5.85 - 10.5	<0.001		Zmap
	Receive time = NULL	232	406	7.83	5.85 - 10.5	<0.001		Zmap
	Transmit time = NULL	232	406	7.83	5.85 - 10.5	<0.001		Zmap
	Transmit time = Nov 24, 2004 15:12:11.4441 UTC	31	0	n/a	n/a	<0.001		Nmap, Metasploit
SNMP (6)	Version = NULL	21	20	6.02	3.17 - 11.4	<0.001		Zmap
	Version = 1	77	17	45.00	25.32 - 79.98	<0.001	SNMPv2c	Metasploit
	Community = NULL	21	20	6.02	3.17 - 11.4	<0.001		
	Community = "Swou9riu"	3	0	n/a	n/a	<0.001		
	Community = "private"	125	0	n/a	n/a	<0.001		
	Data = NULL	23	19	7.05	3.74 - 13.31	<0.001		Zmap
	Data = 1	7	6	6.16	2.04 - 18.59	<0.001	GetNextRequest	Metasploit
	Data = 5	35	3	76.37	23.12 - 252.24	<0.001	GetBulkRequest	
	Request ID = NULL	23	19	7.05	3.74 - 13.31	<0.001		Zmap
	Request ID = 20039	25	1	151.36	20.33 - 1126.73	<0.001		
	Request ID = 1722903134	17	10	9.58	4.3 - 21.36	<0.001		
	Request ID = 925904563	16	1	90.55	11.91 - 688.34	<0.001		
	Error Status = NULL	58	22	20.87	12.23 - 35.63	<0.001		Zmap
	Variable Bindings = NULL	23	19	7.05	3.74 - 13.31	<0.001		Zmap
	Variable Bindings = 1	92	63	16.93	11.21 - 25.58	<0.001		Metasploit
	Variable Bindings = 2	9	1	48.48	6.1 - 385.54	<0.001		
	Variable Bindings = 4	16	5	18.02	6.49 - 49.98	<0.001		
	Variable Bindngs = 7	12	3	21.94	6.12 - 78.74	<0.001		

表 8: ホスト数に注目したクエリフィンガープリント.

Domain	Darknet	Campus	Odds ratio	95% CI	P-value	Note
com	186	28	592.87	396.24 - 887.05	<0.001	Domain BL
census.gov	20	7	213.84	90.22 - 506.81	<0.001	Domain BL
NULL	101	67	122.39	89.28 - 167.77	<0.001	Zmap
doc.gov	14	4	260.44	85.58 - 792.56	<0.001	Domain BL
ietf.org	15	5	223.45	81.06 - 615.95	<0.001	Domain BL
isc.org	42	29	110.73	68.71 - 178.47	<0.001	Domain BL, Metasploit
dnsscan.shadowserver.org	35	33	80.53	49.85 - 130.09	<0.001	Survey, Domain BL
wradish.com	15	13	85.93	40.79 - 181.05	<0.001	Domain BL
1x1.cz	42	0	n/a	n/a	<0.001	Domain BL
∗.w.shifen.com	228	0	n/a	n/a	<0.001	Survey, Baidu

表 9: パケット数に注目したクエリフィンガープリント.

Domain	Darknet	Campus	Odds ratio	95% CI	P-value	Note
com	337,314	170	11,660.08	10,032.28 - 13,551.99	<0.001	Domain BL
doc.gov	38,328	22	9,729.13	6,405.38 - 14,777.58	<0.001	Domain BL
www.iana.org	75,813	112	3,803.81	3,160.3 - 4,578.37	<0.001	Domain BL
dnsscan.shadowserver.org	827,288	124,977	42.39	42.13 - 42.64	<0.001	Survey, Domain BL
www.google.it	335,757	112,146	17.53	17.41 - 17.65	<0.001	Domain BL
∗.openresolverproject.org	120,863	42,844	15.95	15.78 - 16.13	<0.001	Survey, Domain BL
VERSION.BIND	342,429	183,207	10.93	10.87 - 11.00	<0.001	Domain BL, Metasploit
wradish.com	108,932	61,566	9.98	9.88 - 10.08	<0.001	Domain BL
NULL	82,669	61,511	7.55	7.47 - 7.63	<0.001	Zmap
∗.openresolvertest.net	122,745	0	n/a	n/a	<0.001	Survey, Domain BL

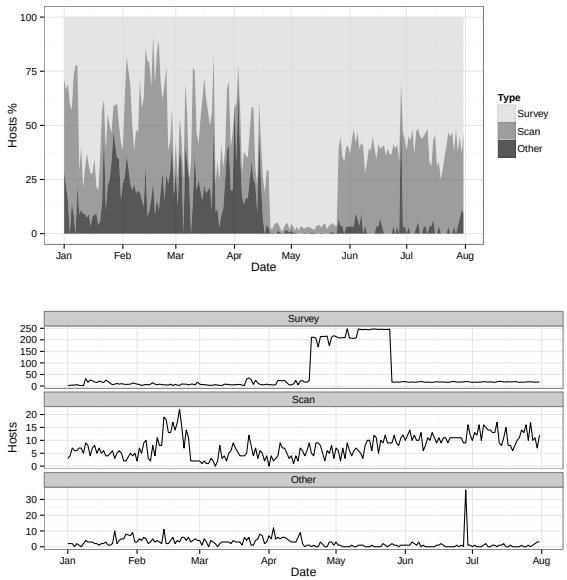

図 3: ダークネットにおける増幅器探索通信の内訳．100%積み上げグラフ (上) および折れ線グラフ (下)．

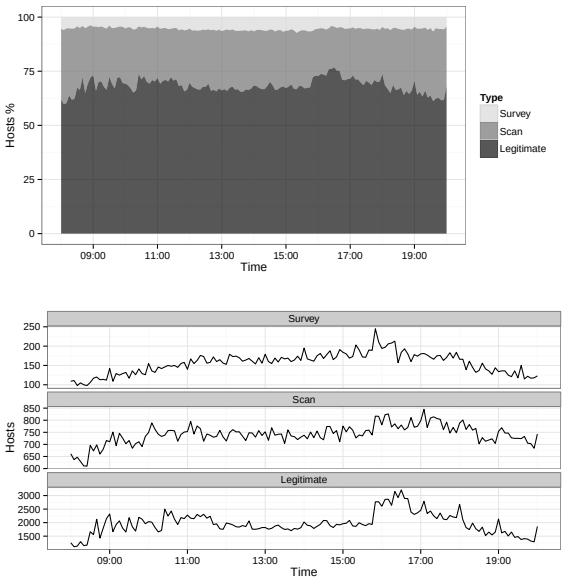

図 4: キャンパスネットワークにおける増幅器探索通信の内訳．100%積み上げグラフ (上) および折れ線グラフ (下)．

6 まとめ

リフレクター攻撃における増幅器探索通信をダークネットおよび正常ネットワークを用いて解析した．特にダークネットにおける観測では調査・研究目的の通信ホストが少なくとも半数以上を占めることを明らかにした．また，IP アドレスベースの分析の結果，複数ホストを用いた協調型のスキャンの存在を明らかにした．本研究で開発した症例対照研究を用いたフィンガープリント抽出技術はダークネットと正常ネットワークのデータ間で統計的に有意差がある特徴を統計的に検出することが主要なアイディアである．これにより増幅器探索通信ホストの特徴や使用したスキャンツールを割り出すことができた．

本研究で示した IP アドレスベースの解析とフィンガープリントを活用することにより，増幅器探索通信から調査・研究目的の通信を**ノイズ**として除外することが出来る．すなわちノイズを除去することにより，攻撃者によるスキャン通信の検出や攻撃予兆の精度を高めることが出来る．また，統計的フィンガープリント抽出手法はあらゆる通信プロトコルに対して汎用的であり，リフレクター攻撃に限らず，様々なサイバー攻撃の検知，動向把握に応用が期待できる．本研究ではフィンガープリントとして個々の特徴を独立に抽出しているが，複数の特徴が共起する現象が確認できている．これらの特徴の組み合わせを捉え，フィンガープリントの精度を向上させることは今後の課題である．

参考文献

[1] Angela Moscaritolo, “Record-Breaking DDoS Attack Nears 400 Gbps.” http://www.pcmag.com/article2/0,2817,2453157,00.asp.

[2] 総務省　総合通信基盤局, “我が国のインターネットにおけるトラヒック総量の把握.” http://www.soumu.go.jp/main_content/000279409.pdf.

[3] M. Allman, V. Paxson, and J. Terrell, “A brief history of scanning,” in *Proceedings of the 7th ACM SIGCOMM Conference on Internet Measurement*, IMC ’07, (New York, NY, USA), pp. 77–82, ACM, 2007.

[4] M. Eto, D. Inoue, J. Song, J. Nakazato, K. Ohtaka, and K. Nakao, “Nicter: A large-scale network incident analysis system: Case studies for understanding threat landscape,” in *Proceedings of the First Workshop on Building Analysis Datasets and Gathering Experience Returns for Security*, BADGERS ’11, (New York, NY, USA), pp. 37–45, ACM, 2011.

[5] Open Resolver Project. http://openresolverproject.org/.

[6] Open NTP Version (Mode 6) Scanning Project. https://ntpscan.shadowserver.org/.

[7] C. Rossow, “Amplification Hell: Revisiting Network Protocols for DDoS Abuse,” in *Proceedings of the 2014 Network and Distributed System Security (NDSS) Symposium*, February 2014.

[8] 中里純二, 島村隼平, 衛藤将史, 井上大介, and 中尾康二, “ダークネットモニタリングによる DNS トラフィック分析,” in *MWS2013*, October 2013.

[9] “I-Blocklist.” https://www.iblocklist.com/.

[10] 秋山満昭, 神薗雅紀, 松木隆宏, and 畑田充弘, “マルウェア対策のための研究用データセット ～mws datasets 2014～,” in **情報処理学会研究報告コンピュータセキュリティ（*CSEC*）**, vol. 2014-CSEC-66, pp. 1–7, 2014.

[11] “Baidu.” http://www.baidu.jp/.

[12] “No Think!.” http://www.nothink.org/honeypot_dns.php.