様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９　（共通）

科学研究費助成事業　　研究成果報告書

令和　３ 年　 ６ 月 １７ 日現在

機関番号：６２６１５
研究種目：基盤研究(B)（一般）
研究期間：2018～2020
課題番号：１８Ｈ０３２３７

研究課題名（和文）DNSバックスキャッターによるIPv6ネットワークでの大規模スキャン検出

研究課題名（英文）Detecting IPv6 network scan with DNS backscatter

研究代表者
福田　健介（Fukuda, Kensuke）

国立情報学研究所・アーキテクチャ科学研究系・准教授

研究者番号：９０４３５５０３

交付決定額（研究期間全体）：（直接経費）　　13,200,000円

研究成果の概要（和文）：IPv4インターネットではネットワークアドレスが32ビットであるため，アドレス空間全体へのネットワークスキャンは簡単に行うことができる．しかしIPv6ではアドレス空間が非常に大きいためランダムネットワークスキャンを行うこともそのようなスキャンを検出することも容易ではない．本課題では，インターネットにおける名前解決プロトコルであるDNSの問い合わせを用いたIPv6ネットワークスキャン検出手法の提案・評価を行った．DNSバックスキャッター技術を用いることで，既存の検出センサネットワークであるダークネットやバックボーンでの検出と比べて効率的にスキャンが検出可能であることが示された．

研究成果の学術的意義や社会的意義
本課題では，ネットワーク上で起こるIPv6ネットワークスキャンを中央集権的なDNSへのクエリをルールベースの識別器を用いることで検出する技術を確立した．この技術により，局所的なネットワーク監視をネットワーク中で大規模に行う必要なく，root DNSへのクエリの観測のみから検出できることから，全インターネットで起こりうる大規模ネットワークスキャンを検出することが原理的に可能となった．ネットワーク管理者・運用者は提案手法を用いることで自ネットワークでの異常検出が容易となると期待できる．

研究成果の概要（英文）：Full network scans to the IPv4 whole address space are handy and easy due to its address size. However, random scans to IPv6 address space are not feasible and also hard to detect in current passive network sensors. In this work, we design, implement, and evaluate a new framework to detect IPv6 scans, called DNS backscatter. DNS backscatter relies on DNS queries triggered by targets when scanners send scan packets to them. Our results demonstrate that the DNS backscatter can detect network-wide IPv6 scans more effectively than existing techniques (e.g., darknet, backbone traffic analysis).

研究分野：インターネット工学

キーワード：インターネット　DNS　スキャン　セキュリティ

様　式　C－１９、F－１９－１、Z－１９（共通）

１．研究開始当初の背景

インターネット上のネットワークスキャンはホストの脆弱性を探るために用いられることから，大規模攻撃の第一歩と捉えることができる．実際，新しい脆弱性が発見されると，そのインパクトを評価するためにスキャンが行われている．このスキャンはセキュリティベンダーや研究者のような善意によるものの他に，その脆弱性を使用する悪意のあるユーザによっても行われる．そのため，このようなネットワークワイドに行われるネットワークスキャンを早期に検出することは重要な問題である．

現在のインターネットで広く用いられている IPv4 プロトコルでは，ネットワーク上の機器に割り当て可能な IP アドレスは 32bit である．近年の計算機・ネットワークの高性能化によって，32bit 空間全てにスキャンを行うためには高々数時間を要するのみであるため，カジュアルにネットワークスキャンが行われる傾向にあり，ダークネットやバックボーンでのトラフィック観測によって多数のネットワークスキャンが検出されている．それに対して，次世代のネットワークプロトコルである IPv6 では事情が大きく異なっている．IPv6 は現在その利用が急速に増えており，Google 等のコンテンツ事業者も積極的にその導入を図っている．IPv6 アドレスは 128bit のアドレス空間を持っていることから，IoT デバイス等の大量の機器がインターネットに繋がれることを想定している．このような IPv6 アドレスに関するネットワークスキャンを検出することが本課題のゴールとなる．

２．研究の目的

本研究での IPv6 アドレス空間に対するネットワークスキャンを検出する既存の手法は主にパッシブなトラフィック観測によるものが多い．例えばダークネットと呼ばれるアドレスへの到着性は存在するものの実際のホストの存在しないセンサーネットワークを形成することで，そのネットワークへ到着するスキャンを検出することが可能である．ここで問題となるのは，設置したダークネットや対象とするバックボーンネットワークにどの程度ネットワークスキャンが到着するかである．前述のように IPv6 アドレス空間は 128bit あり，アドレス空間への全探査は地球の誕生から今までの時間をかけても終了することはない．そのため，ランダムなスキャンが行われたとしても，該当するネットワークにスキャンが「たまたま」到着する確率は小さいことから，これらの手法では効率良くネットワークに大規模に行われるスキャンを検出することは困難であると予想される．

本研究の目的はパッシブなトラフィック測定に基づかない新しいネットワークスキャン検出手法を提案することである．

３．研究の方法

【基本アイディア】

本研究のメインアイディアである DNS バックスキャッターについて説明する．スキャナーがネットワーク中のアドレス群(ターゲット)に対してスキャンを行うシナリオを考える(図 1)．

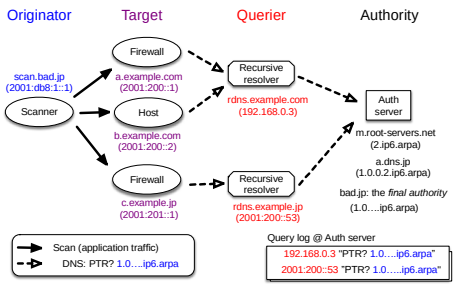

図 1: DNS バックスキャッターの原理

スキャナーがターゲットにパケットを送信する際に，そのターゲットにホスト内ファイアーウォールがあったり，ミドルボックスのファイアーウォールが存在する場合がある．想定されていないスキャナーの IP アドレスよりパケットが送られてくると，ファイアーウォールではそのパケット受信をログとして保存する．保存したログには攻撃者の IP アドレスが含まれているが，この IP アドレスのホスト名の検索が生じる．この検索には DNS が用いられることから，ターゲットからフルリゾルバー(クエリア)への DNS クエリが送られ，クエリアより権威 DNS サーバへのクエリが送られる．仮に，多数のターゲットにスキャンが送られると，短期間に攻撃者の IP アドレス-ホスト名変換を行うクエリーが各クエリアで発生する．すなわち，ネットワークスキャンを行うと，権威サーバで関連する DNS クエリが観測されることから，DNS サーバではネットワークスキャンが生じたことを知ることができる．この DNS バックスキャッターは少量の攻撃

者に対する情報を権威サーバで共有知として検出する手法と言える．今まで述べた説明は，スキャンが生じた際のバックスキャッターを示したものであるが，実際には，スキャン以外の大規模なネットワークイベントが生じた際にもバックスキャッターは生じる可能性がある．そのため，本研究では，IPv6 ネットワークにおいて，スキャンを含む大規模なネットワークイベントを観測可能であるかを検証し，その大規模なネットワークイベントからネットワークスキャンを検出することが可能であるかを示すことがメインの課題となる．

【DNS バックスキャッターの敏感性調査】
本研究トピックでは，ネットワークスキャナを準備し，その IP アドレスに関わる DNS 権威サーバを用意する．ネットワークスキャナが実際にネットワークスキャンを行った際の，ターゲットホストの挙動把握および DNS バックスキャッターの有無を確認する．ターゲットのリストは各種考えられるが，本実験では人気のあるドメインのリストである Alexa Top 1m リストの逆引きリスト，逆引き DNS のアドレスリスト，P2P ソフトウェアによって得られたアドレスリストを使用する．ネットワークスキャンに用いるプロトコル・サービスが制御パラメータとなる．

【DNS バックスキャッターの検出】
実際に稼働している権威 DNS サーバにおいて DNS 逆引きクエリを収集し，DNS バックスキャッターからネットワークスキャンを検出可能かを確認する．権威 DNS サーバとして，USC/ISI の運用する B-Root DNS サーバのログデータを使用する．先に述べたように DNS バックスキャッターは悪意のあるネットワークイベント外のイベントによっても生じる可能性がある．そこで，収集されたバックスキャッターデータからネットワークスキャンイベントを検出する必要がある．ここで問題となるのは，どのようにして他の正常なデータと異常データのラベルを得るかという点である(この正常データの収集は別の節で述べる)．正常なデータのラベルは基本的に IP アドレスのブロックより推定するが，ヒューリスティックなルールを用いて正常データの分離を行う．アイディアとしては，正常なデータを除いたデータに異常が含まれるため，その残されたデータをできるだけ小さい集合となるよう正常データのルールを生成する．また，データの収集期間とデータ量はトレードオフの関係となる．収集期間が短いと短期間でのイベント検出が可能となる一方，短い収集期間では十分な数の DNS クエリが得られない可能性があるため，このトレードオフについても調査する．

【IPv6 アドレスの収集・解析】
DNS バックスキャッターは，スキャンのような悪意のあるイベント以外にも多数の正常なイベントに関連して発生する．そのため，バックスキャッターとして検出された IPv6 アドレスが正常なものであるのかについて議論する必要がある．そこで，本トピックでは，アクティブ・パッシブな方法を用いてネットワーク上で使用される IPv6 アドレスを収集し，その傾向を把握する．実際には，メール・DNS・ウェブ・NTP・P2P 等のアプリケーションを用いて，そのアプリケーションに接続される IPv6 アドレスを対象とする．また，比較として，ネットワークバックボーン中に現れる IPv6 アドレスを収集し傾向を明らかにする．

４．研究成果

【DNS バックスキャッターの敏感性調査】
DNS バックスキャッターの敏感性を調査するために，上記の 3 リストを用いてランダムサンプリングしたアドレスに対してネットワークスキャンを行った結果を述べる(図 2)．

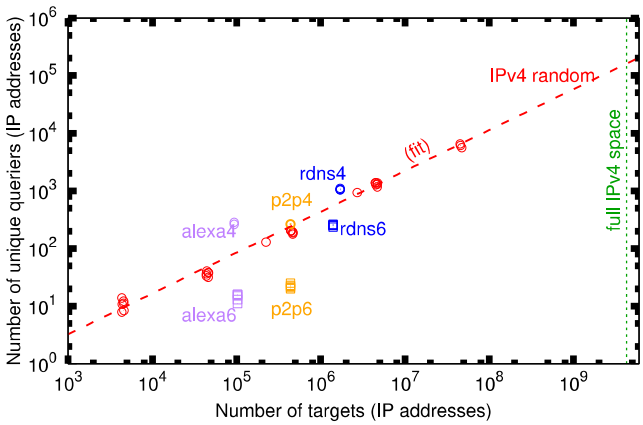

図 2: DNS バックスキャッターの敏感性

図中の横軸は送信した IP アドレス数，縦軸は観測された DNS バックスキャッターの数(クエリ

アの IP アドレス数)である．図中には，IPv4 アドレスおよび IPv6 アドレスに対して送られたスキャンの結果がプロットしてある．IPv4 アドレスでは，赤の点線のフィッティングにあるように，32 ビットの全空間をスキャンした際には，10 万のオーダのクエリアからの応答が期待できることが読み取れる．それに対して，IPv6 スキャンでは，各プロットを比較してみると，その 1-10%程度のクエリアのみが観測可能であることがわかる．つまり IPv6 アドレス空間では，IPv4 アドレス空間よりもスキャンに関する感応性が低いことを意味しており，ネットワーク上のセキュリティを考えた際には，より深刻な問題となると言える．
また，異なるトランスポートプロトコル・ポートを用いたスキャンに対する返答を調査したところ，IPv6 では，ping (62.9%)，ssh (27.8%)，http (44.8%)，dns (4.7%)，ntp (9.5%)のアプリケーション返答を得た．それに対して，IPv4 では，ping (57.8%)，ssh (30.0%)，http (35.4%)，dns (6.3%)，ntp (5.9%)との結果を得た．この両者を比較してみると，IPv4 および IPv6 のサービス別の傾向の違いは読み取れないことがわかった．これは，ファイアーウォール等の設定において，IPv4・IPv6 の違いなく，アクセスリストの設定を行っていることを示唆している．
以上，IPv6 ネットワークスキャンによる実験で DNS バックスキャッターを用いることでネットワークスキャンが原理的に検出可能であることが示された．

【DNS バックスキャッターの検出】
本トピックでは，ルート DNS サーバにおける DNS バックスキャッター検出手法の開発を行った．対象となる DNS クエリデータは，B-root DNS サーバで得られたものを使用している．DNS バックスキャッターは正常・異常の両方のイベント検出することから，検出された IP アドレスからいベントのタイプを同定する必要がある．DNS トラフィックデータおよび口述の IPv6 アドレス収集データより，経験的に以下のタイプを定義した: CP (Google，Microsoft，Facebook，Yahoo 等のコンテンツプロバイダ)，cdn，dns，ntp，mail，web，Tor，other services，routers，tunneling，scan，spam．これらのタイプの中で scan，spam が異常を示すタイプである．B-root データで 6 ヶ月間に得られた DNS データを解析した結果，タイプを推定するには，1 週間程度の時間ビンでデータを集約する必要があることがわかった．IPv4 では 1 日単位の集約で十分であったことを考えると，IPv6 における DNS バックスキャッターの感受性が検出精度に大きく影響を及ぼしていると言える．検出されたこれらのタイプの割合は以下のとおりである．CP (70.2%)，cdn (4.2%)，well-known service (12.1%)，minor service (4%)，routers (4.28%)，spam (0.25%)，scan (0.24%)，potential abuse (1.4%)．これらの割合を見ると，70%強がコンテンツプロバイダによる正常なイベントであり，実際の異常イベントは 2%程度であることがわかる．つまり，IP アドレスの割り当ておよび使用法をリスト化することで正常なイベントを前もって取り除く手法が有効であると言える．
次に検出されたスキャンの傾向について説明する．観測期間中にダークネットおよびバックボーントラフィックの解析によって観測期間中にネットワークスキャンを検出する．このデータはある種の正解データとなる．この正解データがバックスキャッターでも検出可能であるかを検証することで，バックスキャッターと他の手法による検出の違いを評価することができる．実際には，ダークネットおよびバックボーントラフィックでは 7 つの IP アドレスが検出された．そのうち，ダークネットで検出されたアドレスは 1 つのみであり，バックボーントラフィックとダークネットではスキャンの検出精度に違いがあることがわかった．DNS バックスキャッターではこれらの 7 つの IP アドレスのうち 4 つを検出することができた．残りの 3 つの IP アドレスは未検出であるが，DNS クエリの存在は確認できており，ネットワーク規模のスキャンではなかったと考えられる．さらに，バックスキャッターでは 95 のその他の潜在的なスキャンホストを得ている．この IP アドレスには，偽陽性の可能性のあるアドレスも含まれるが，いくつかは研究目的で行われているスキャナーの IP アドレスブロックからのものであることを確認している．
以上のように，DNS バックスキャッターを用いることで，局所的なバックボーンやダークネットと言ったパッシブなセンサ以上の精度でネットワークスキャンを検出可能であることが示された．

【IPv6 アドレスの収集・解析】
本測定・解析トピックでは，7 つのアプリケーション由来の IP アドレス収集手法および，DNS の逆引きツリー探索による IP アドレス収集手法を用いて，データ収集を行った．測定期間および収集アドレス数・プリフィックス数を表 1 に示す．結果を見ると，DNS，メール，ウェブ等のサービス系のサーバによるデータ収集では多くの IP アドレスを収集することが困難であることがわかる．これは，サーバやドメインの人気度によるが，サービスに特化したアドレスを収集するには有用であるが，多数のアドレスを集める目的であればそれほど有効な手段ではないと言える．それに対して，NTP や P2P では多数のアドレスを収集できている．これは多数のクライアン

	Period	#Address	#prefix/64	#AS
DNS	425 days	12.8K	1.7K	786
Mail	354 days	4	3	1
Web	352 days	984	793	152
NTP	393 days	1.8M	1.8M	576
BitTorrent	410 days	28M	16M	2,963
Bitcoin	385 days	27K	16K	618
Traffic	420 days	2.2M	1.3M	6,695
rDNS	55 days	7.5M	118K	582
Total		38M	19M	7,369

表 1: 収集 IP アドレス数

トからの接続が期待できるアプリケーション故の結果であると言える．同様にバックボーントラフィックの解析では，多くのアドレスが収集可能であるが，これもまた，どのようなトラフィックが流れるリンクかにも依存することから，一般にこの手法が有意であるかは今後の議論が必要である．最後に DNS 逆引きによる収集手法であるが，DNS の逆引きツリーを完全に探索することは時間的な困難が伴うが，アドレス数という観点では最も効率の良い収集法であると言える．しかし，DNS 逆引きホスト登録という制約上，全てのネットワークにおいてこの設定がなされていることは期待できないため，AS 数やプリフィックスで見た場合のカバー率には問題がある．

以上，8 つの IPv6 アドレス収集手法を実際に稼働させデータ収集を行い，それぞれのデータ手法およびその結果の分析結果を得ることができた．

５．主な発表論文等

〔雑誌論文〕　計7件（うち査読付論文　7件／うち国際共著　4件／うちオープンアクセス　2件）

１．著者名 K.Fukuda, Y.Yoneya, T.Mitamura	４．巻 2020
２．論文標題 Towards detecting DNSSEC validation failure with passive measurements	５．発行年 2020年
３．雑誌名 Proceedings of IEEE/IFIP ANNET2020	６．最初と最後の頁 1-6
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/NOMS47738.2020.9110466	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 －

１．著者名 P.Mulinka, P.Casas, K.Fukuda, L.Kencl	４．巻 2020
２．論文標題 HUMAN - Hierarchical Clustering for Unsupervised Anomaly Detection and Interpretation	５．発行年 2020年
３．雑誌名 Proceedings of NoF 2020	６．最初と最後の頁 132-140
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/NoF50125.2020.9249194	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

１．著者名 P.Mulinka, K.Fukuda, P.Casas, L.Kencl	４．巻 2020
２．論文標題 WhatsThat? On the Usage of Hierarchical Clustering for Unsupervised Detection & Interpretation of Network Attacks	５．発行年 2020年
３．雑誌名 Proceedings of IEEE European Symposium on Security and Privacy Workshops	６．最初と最後の頁 574-583
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/EuroSPW51379.2020.00084	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

１．著者名 Guannan Hu and Kensuke Fukuda	４．巻 0
２．論文標題 Toward Detecting IoT Device Traffic in Transit Networks	５．発行年 2020年
３．雑誌名 Proceedings of ICAIIC2020	６．最初と最後の頁 525-530
掲載論文のDOI（デジタルオブジェクト識別子） 10.1007/978-1-7281-4985-1/20	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 －

1．著者名 K.Fukuda, J.Heidemann	4．巻 -
2．論文標題 Who Knocks at the IPv6 Door?: Detecting IPv6 Scanning,	5．発行年 2018年
3．雑誌名 Proceedings of ACM IMC 2018	6．最初と最後の頁 231-237
掲載論文のDOI（デジタルオブジェクト識別子） 10.1145/3278532.3278553	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 該当する

1．著者名 S.Mongkollusksamee, V.Visoottiviseth, K.Fukuda	4．巻 -
2．論文標題 Robust Peer to Peer Mobile Botnet Detection by Using Communication Patterns	5．発行年 2018年
3．雑誌名 Proceedings of AINTEC 2018	6．最初と最後の頁 38-45
掲載論文のDOI（デジタルオブジェクト識別子） 10.1145/3289166.3289172	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 該当する

1．著者名 新津雄大，小林諭，福田健介，江崎浩	4．巻 J103-B
2．論文標題 大規模IPv6アドレスの収集・分析	5．発行年 2020年
3．雑誌名 電子情報通信学会和文論文誌(B)	6．最初と最後の頁 223-233
掲載論文のDOI（デジタルオブジェクト識別子） 10.14923/transcomj.2019JBT0002	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 －

〔学会発表〕　計3件（うち招待講演　1件／うち国際学会　1件）

1．発表者名 小林日向，小林諭，福田健介，江崎浩
2．発表標題 IPv6エイリアス空間検出を考慮したハニーネットの検討
3．学会等名 電子情報通信学会インターネットアーキテクチャ研究会
4．発表年 2019年

１．発表者名 Kensuke Fukuda
２．発表標題 Detecting large-scale network scanners in IPv4/IPv6 networks
３．学会等名 Proceedings of FDSE/ACOMP 2019（招待講演）（国際学会）
４．発表年 2019年

１．発表者名 Y.Aratsu, S.Kobayashi, K.Fukuda, H.Esaki
２．発表標題 Collecting a large number of IPv6 addresses
３．学会等名 Internet Conference 2018
４．発表年 2018年

〔図書〕　計0件

〔産業財産権〕

〔その他〕

―

６．研究組織

	氏名 （ローマ字氏名） （研究者番号）	所属研究機関・部局・職 （機関番号）	備考

７．科研費を使用して開催した国際研究集会

〔国際研究集会〕　計0件

８．本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関			
米国	南カリフォルニア大学			