一般社団法人　電子情報通信学会
THE INSTITUTE OF ELECTRONICS,
INFORMATION AND COMMUNICATION ENGINEERS

信学技報
IEICE Technical Report
NS2020-41(2020-09)

[招待講演] DNS通信分析によるボットネット通信の検知・遮断技術の研究最前線

飯田　勝吉† 　一瀬　　光†† 　金　　勇†††

† 北海道大学　情報基盤センター
〒 060-0811 札幌市北区北 11 条西 5 丁目
†† 東京工業大学　技術部
††† 東京工業大学　学術国際情報センター
E-mail: †iida@iic.hokudai.ac.jp, ††hichise@nap.gsic.titech.ac.jp, †††yongj@gsic.titech.ac.jp

あらまし　ボットネットが大きな社会問題となっている。ボットネットとはマルウェアの一種であるボットが構築する論理的なネットワークである。本稿では DNS 通信分析によるボットネット通信の検知、遮断技術を概説する。
キーワード　ボットネット, DNS, トラフィック分析, 検知・遮断.

[Invited talk] Recent Research Trends in Detection and Blocking System of Botnet Communications by DNS Traffic Analysis

Katsuyoshi IIDA†, Hikaru ICHISE††, and Yong JIN†††

† Information Initiative Center, Hokkaido University,
Kita 11, Nishi 5, Kita-ku, Sapporo-shi, 060-0811, Japan.
†† Technical Dept., Tokyo Institute of Technology, Tokyo, Japan.
††† Global Scientific Information & Computing Center, Tokyo Institute of Technology, Tokyo, Japan.
E-mail: †iida@iic.hokudai.ac.jp, ††hichise@nap.gsic.titech.ac.jp, †††yongj@gsic.titech.ac.jp

Abstract　Botnet, a logical network used by bot-type malware, is becoming a social issue. In this paper, we give an outline of the state-of-the-art technologies to detect and block botnet communication using DNS traffic analysis.
Key words　Botnet, DNS, traffic analysis, detection and blocking.

1. はじめに

ボットネットの脅威が増大している。例えば、2019 年 10 月以降に日本国内で感染事例が急増し、2020 年 2 月に 3200 以上の日本国内の組織に感染が確認されているマルウェア Emotet [1] もボットネットを利用している。また [2] によると、Spamhaus Block List (SBL) に登録されているボットネットが利用するサーバ（Command & Control (C&C)）の数は、2014 年から 2019 年にかけて約 2.4 倍に増加している。

ボットネットとは、感染した PC などの端末を C&C サーバが遠隔操作するボットと呼ばれるマルウェアにおいて、感染端末と C&C サーバ間で構成される論理的なネットワークを指す。端末がボットに感染すると、感染端末は C&C サーバの指示に基づき、機密情報の収集、DDoS 攻撃、迷惑メールの送信、マルウェアの拡散などの悪性行動をとる [3, 4]。企業や大学などの組織ネットワークの管理者にとっては、そのような意図が無かったとはいえ、ボット感染端末が外部組織等への攻撃に加担することとなり、組織防衛上、組織ネットワーク内の感染 PC の出来るだけ速やかな発見が求められる。

そのため、ボットネットの検知の研究はこれまで多数行われており、本稿ではそれらの研究動向を概説する。

2. ボットネットの通信プロトコル

ボットネットでは C&C サーバと感染端末が何らかの通信プロトコルによって通信する。これまでに利用が確認されている通信プロトコルは多数存在する（表 1）。初期に使われていた通信プロトコルは IRC であったが、ネットワーク管理者による IRC 通信の発見や遮断が容易となったため、ボットの製作者が利用する通信プロトコルは、**通信の秘匿**が可能なものに変わってきた [3, 4]。その 1 例として DNS がある。DNS はインターネットの基盤となる通信プロトコルであり、一律に遮断するとインターネットの運用に重大な支障が出るため、ボット製作者

表 1　ボットネットで利用されている通信プロトコルの例

Internet Relay Chat (IRC)	Hyper Text Transfer Protocol (HTTP)
Peer-to-Peer (P2)	Domain Name System (DNS)

にとって有利であり、利用が増えている。

3. ボットネット対策研究の概況

ボットネット対策研究は多くのサーベイ論文にまとめられている [3, 5, 6]。ここではページ数の関係で各検知技術の概要のみを述べる。

- 機械学習による検知

不正な通信と正当な通信を機械学習で学習させ、それにより不正な通信を検知する方式。

- ハニーポットによる検知

仮想マシン技術やサンドボックス技術を使って、仮想の感染端末を構築し、仮想の端末に C&C サーバと通信させ、それにより不正通信のパターンを取得する方式。

- DNS 通信分析

端末がボット型マルウェアに感染すると、端末は C&C サーバとの通信を確立しようとする。もし、マルウェア内に C&C サーバの IP アドレスが直接記述 (hard-coded) されていると、当局や ISP から C&C サーバの接続停止措置を取られた際に対応が困難となるため、Domain Generation Algorithm (DGA) [7] という特殊なアルゴリズムを使って動的に変化するドメイン名を作成されるケースが存在する。そこで、DNS 通信を分析して、マルウェアに感染直後の端末が C&C サーバを探索する通信を検出する方式が研究されている [8,9]。なお、ボットネット通信のプロトコルとして DNS を利用する場合にも DNS 通信分析は有用となる。

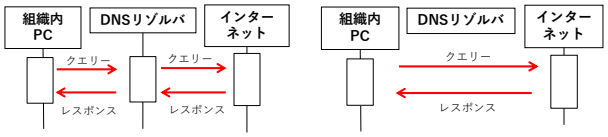
上記の方式は組み合せて利用されることがあり、たとえば機械学習によって不正な DNS 通信を判別する方式として [10, 11] がある。

4. 著者らの研究の紹介

本節では著者らが行っている DNS 通信分析によるボットネット通信の検知・遮断技術の研究を紹介する。なお、本節で紹介した一連の研究は [12] にまとめている。

4.1 DNS 型ボットネットの通信分析研究

2 節で述べた通りボットネット通信を秘匿する目的で通信プロトコルとして DNS が利用されている。DNS に複数あるレコードタイプの中でも、TXT レコードの利用が多いことが知られている。TXT レコードは EDNS 拡張 [13] を利用すると 4,000 バイト以上の任意の文字列をペイロードに格納可能なことが特徴である。ボット製作者は、この特徴を活かして暗号文のビット列を Base64 エンコードでテキスト化した文字列をドメイン名 (253 文字まで可能) に、ボットへの指示をペイロードとして TXT レコードに格納している。秘密鍵をしらない第 3 者はボットネット通信の内容の分析することが困難であり、そのためボット製作者にとっては DNS を用いた秘匿通信が可能



(a) リゾルバ経由　　(b) 外部クエリー

図 1　DNS によるボットネット通信

表 2　DNS TXT レコードの利用統計

分類	クエリ数	比率 (%)
SPF と domainkey	12,223	0.24
DNS によるサービス発見	213,978	4.30
NFSv4	3,596,481	72.14
ウィルス対策ソフト	597,901	12.00
スパム検査と DNS ブラックリスト	180,600	3.63
P2P トラッカー	446	0.01
NTP	632	0.01
その他	380,723	7.63
未確認	2,293	0.04
計	4,985,277	100.00

となる。

次に、DNS によるボットネット通信は 2 種類に大別されることが知られており、図 1 に示す。リゾルバ経由（図 1.(a)）は組織内 PC が組織内の DNS リゾルバを経由して C&C サーバと通信する。一方、外部クエリー（図 1.(b)）は DNS リゾルバを経由せずに C&C サーバと通信する。また外部クエリーはさらに 2 種類に分類でき、一切 DNS リゾルバを利用しない「直接外部クエリー」と C&C サーバの名前解決のみ DNS リゾルバを利用し、名前解決完了後はリゾルバを利用しない「間接外部クエリー」に分けられる。

4.1.1 リゾルバ経由通信の分析

最初に著者らは、キャンパスネットワーク上で TXT レコードでどんな通信が行われているか、特にリゾルバ経由通信について調査した [14, 15]。TXT レコードは自由に利用できるレコードタイプであるが、標準化された利用方法 (電子メールの迷惑メール対策用の SPF レコード、分散ファイルシステム用の NFSv4 レコードなど) が存在する。さらに、標準化はされていないが、ベンダーが独自利用しているもので、利用方法が公開されているものが存在する（ウィルス対策ソフトのセキュリティアップデートなど)。表 2 は、キャンパスネットワークで約 3 か月間取得した TXT レコードのクエリを分類したものである。事前に作成した分類に入らず、なんの目的で利用されているか判別できなかったものを「未確認」分類としている。「未確認」分類のドメイン名を “virustotal.com” で確認したところ、約 44%のドメイン名が過去に不正通信に関与したことがあることが明らかになった。つまり、TXT レコードタイプの通信を機械的に分類し、「未確認」分類だけを抽出し、詳細を調査すること効果的に不正通信の検出が可能であることを示した。

4.1.2 外部クエリー通信の研究

次に外部クエリー通信について調査した [16]。前節の研究ではキャンパスネットワークのリゾルバのログを調査したが、

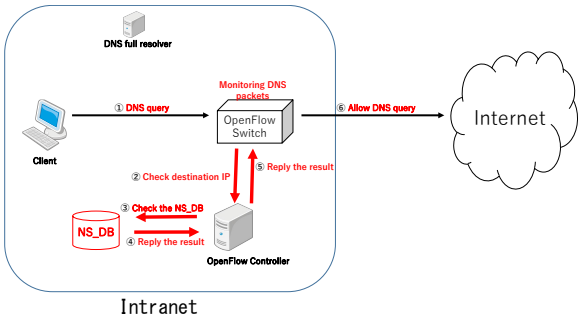


図 2　DNS を用いたボットネット通信の検知・遮断機構

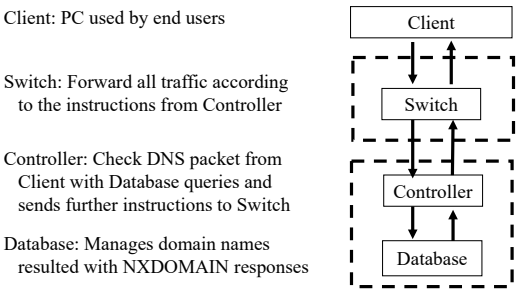


図 3　NXDOMAIN 応答に着目したボット検知機構

外部クエリーの調査をするためにはキャンパスネットワークのゲートウェイでのパケットキャプチャ環境が必要でありこれを構築し、約 3 か月間のログを取得した。ログ分析により抽出した外部 DNS サーバの IP アドレスを “virustotal.com” で調査したところ、直接外部クエリーに関しては約 8%、間接外部クエリーに関しては約 22%の IP アドレスが不正通信に関わっていたことが明らかとなった。

そもそも通常の DNS 通信はリゾルバを経由するため、外部クエリー通信（特に直接外部クエリー通信）はそれだけで不正通信のリスクが高いと考えられていた。しかし、昨今では Google や Cloudflare などがパブリック DNS リゾルバサービスが広く利用されており、外部クエリー通信というだけで通信を遮断するのは適切ではなくなってきている。そこで、リスクの高い外部クエリー通信を識別し、遮断する遮断する機構が重要となる。

4.1.1 節、4.1.2 節の研究は、TXT レコードによる DNS ボットネット通信の一連の研究として、[17] にまとめている。

4.1.3　検知・遮断システムの研究

次に、[15–17] の分析に基づき、自動的に検知・遮断するシステムを実装・評価した [18, 19]。提案機構 (図 2) の基本的なアイデアは、外部の DNS サーバと通信する前に DNS サーバの名前解決（間接外部クエリーの一部）を行うはずであり、それをしていないと不正通信の可能性が高いことである。そのため、過去の DNS サーバの名前解決（グルー A と呼ばれるレコードが用いられる）の履歴を取得し、データベースに登録してホワイトリストとすることとし，さらにデータベースに登録されていない外部クエリの通信は不正通信の可能性が高いための遮断することとした。(注1)。そのための機構を SDN を用いて、設計実装し、当該システムを評価した [18]。評価の結果、False positive 率が 10%未満になることなどを明らかにした。

4.1.4　今後の課題

一連の研究により、DNS の特に TXT レコードを用いたボットネット通信の検知・遮断機構の基本設計とその有効性が明らかになっている。しかし、SDN を前提とした検知機構であるため、より汎用性の高い検知機構が求められる。また、一部の性能指標に関してさらなる性能向上が必要である。これらの目的を達成するため、BIND の拡張機能である DNS Response Policy Zones (RPZ) [20] を用いた実証実験を行っている。

4.2　NXDOMAIN 応答に着目した DGA を用いるボット検知の研究

3 節で示したとおり、ボットに感染した端末は C&C サーバへの接続を試みる。その際、DGA というアルゴリズムを用いて C&C サーバの探索を秘匿されることが多い。DGA には異なる多くのアルゴリズムがあり、リバースエンジニアリングや計測によりそれらの特徴や分類が明らかになっている [7, 21]。

我々は多くのボットの共通する特徴に着目し、新たな方式を提案している [22, 23]。すなわち、DGA で生成される一部のドメイン名のみが登録されるため、NXDOMAIN 応答とよばれるドメインの未登録エラーが生成されることが報告されている。また、同一のマルウェアに感染した複数の PC が同時に NXDOMAIN 応答を生成するため、複数の PC が同一種類の NXDOMAIN 応答が同時に生成する状況を検知し、それに基づき疑わしい通信を自動的に遮断する機構を設計、実装した（図 3）。

提案機構では、スイッチが NXDOMAIN 応答を受信すると SDN コントローラにそれを送信し、SDN コントローラは NXDOMAIN 応答の情報をデータベースに記録する。もし、複数の端末が同一のドメイン名を検索し、いずれも NXDOMAIN 応答を返した場合、不正な通信の可能性が高いため、以降の同様の DNS クエリーを遮断する。提案機構においてコントローラの処理手順を図 4 に示す。テストベッド環境において既知の 23 種類の DGA [24] について評価したところ、提案機構が期待通りに検知・遮断できることを明らかにした。

今後は現実のキャンパスネットワークのデータを用いて未知の DGA の検知・遮断の検証を行う予定である。

5.　おわりに

本稿では、不正な通信であるボットネット通信を発見し、遮断するために DNS 通信を分析する研究を概説した。ボットネット通信の検知・遮断技術は、機械学習やハニーポットを使われる方法が多用されてきたが、ボットネットが DGA というアルゴリズムを利用して、高度化してきたために、それを検知するための DNS 通信分析の研究が幅広く行われている。その中でも著者らが行っている 2 つの研究を紹介した。1 つ目は DNS 型ボットネット通信に対処するために、DNS TXT レコードの

(注1)：なお、幅広く普及しているパブリック DNS リゾルバの IP アドレスはホワイトリストに記載し、遮断しないようにしている。

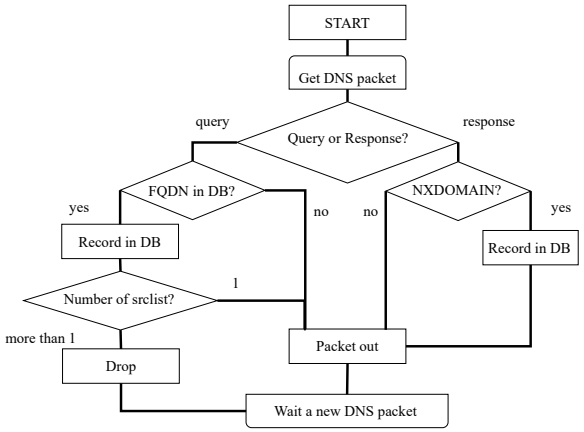


図 4　コントローラの処理手順

分析とそれに基づく検知機構の研究で、２つ目は多くの DGA に共通する特徴を利用して識別する方式として NXDOMAIN 応答に着目した検知機構の研究であった。

DNS は、名前解決というインターネットの重要なサービスの基盤として長年使われてきた。そして、その重要性や TXT レコードの性質などをつかって不正通信の媒体としても使われてきた。また、ボットに感染した端末は、C&C サーバと秘密裏に接続する必要がり、そのため DGA が発達してきた。一方、IoT 機器やサービスの急速な普及が見込まれており、IoT に特化した DNS 不正通信の研究が必要とされている [25]。さらに、DNS over TLS や DNS over HTTPS 等の暗号化を伴う名前解決手法 [26] の普及が見込まれており、それに合わせて不正通信の手法も変化すると考えられ、新たな研究開発が必要と言える。

文　献

[1] 佐條 研, “マルウエア Emotet への対応 FAQ,” `https://blogs.jpcert.or.jp/ja/2019/12/emotetfaq.html`, Dec. 2019.

[2] Spamhaus, “Botnet threat report 2019,” `https://www.spamhaustech.com/custom-content/uploads/2020/04/2019-Botnet-Threat-Report-2019-LR.pdf`, Apr. 2020.

[3] S. Khattak, et al., “A taxonomy of botnet behavior, detection, and defense,” *IEEE Commun. Surveys & Tutorials*, vol. 12, no. 2, pp. 898–924, Oct. 2013.

[4] B. Al-Duwairi, and M. Jarrah, “Chapter 1: Botnet architectures: A state-of-the-art review,” *Botnets: Architectures, countermeasures, and challenges,* ed. G. Kambourakis, M. Anagnostopoulos, W. Meng, and P. Zhou, pp. 1–31, CRC Press, Boca Raton, FL, USA, Sept. 2019.

[5] M. Feily, et al., “A survey of botnet and botnet detection,” *Proc. IEEE Int’l Conf. Emerging Security Information, Systems and Technologies*, Athens, Glyfada, pp. 268–273, June 2009.

[6] H. Binsalleeh, “Botnets: Analysis, detection, and mitigation,” *Network Security Technologies: Design and Applications*, ed. A. Amine, O.A. Mohamed, and B. Benatallah, pp. 204–223, IGI Global, Hershey, PA, USA, Nov. 2013.

[7] A. K.Sood, and S. Zeadally, “A taxonomy of domain-generation algorithms,” *IEEE Security & Privacy*, vol. 14, no. 4, pp. 46–53, July-Aug. 2016.

[8] S. Al-Mashhadi, M. Anbar, S. Karuppayah, and A.K. Al-Ani, “Review of botnet detection approaches based on DNS traffic analysis,” *Lecture Notes in Networks and Systems*, vol. 67, pp. 305–321, Springer, Singapore, May 2019.

[9] K. Alieyan, et al., “A survey of botnet detection based on DNS,” *Neural Comput. & Applications*, vol. 28, pp. 1541–1558, July 2017.

[10] A. Satoh, Y. Nakamura, D. Nobayashi, and T. Ikenaga, “Estimating the randomness of domain names for DGA bot callbacks,” *IEEE Commun. Letters*, vol. 22, no. 7, pp. 1378–1381, July 2018.

[11] A. Satoh, Y. Nakamura, Y. Fukuda, K. Sasai, and G. Kitagata, “A cause-based classification approach for malicious DNS queries detected through blacklists,” *IEEE Access*, vol. 7, pp. 142991–143001, Sept. 2019.

[12] H. Ichise, “A study on detection and blocking of DNS-based botnet communication,” Ph.D. thesis, Hokkaido University, Feb. 2020. DOI: 10.14943/doctoral.k14122

[13] J. Damas, and P. Vixie, “Extension mechanisms for DNS (EDNS0),” *IETF RFC6891*, Apr. 2019.

[14] H. Ichise, Y. Jin, and K. Iida, “Analysis of via-resolver DNS TXT queries and detection possibility of botnet communications,” *Proc. IEEE Pacific Rim Conference on Communications, Computers and Signal Processing (PACRIM2015)*, pp. 216–221, Aug. 2015.

[15] H. Ichise, Y. Jin, and K. Iida, “Analysis of via-resolver DNS TXT queries and detection possibility of botnet communications,” *IEICE Commun. Express*, vol. 5, no. 3, pp. 74-78, Mar. 2016. DOI: 10.1587/comex.2015XBL0186

[16] Y. Jin, H. Ichise, and K. Iida, “Design of detecting botnet communication by monitoring direct outbound DNS queries,” *Proc. IEEE Int’l Confrerence on Cyber Security and Cloud Computing (CSCloud2015)*, pp. 37–41, New York, NY, USA, Nov. 2015.

[17] H. Ichise, Y. Jin, and K. Iida, “Analysis of DNS TXT record usage and consideration of botnet communication detection,” *IEICE Trans. Commun.*, vol. E101-B, no. 1, pp. 70–79, Jan. 2018. DOI: 10.1587/transcom.2017ITP0009

[18] H. Ichise, Y. Jin, K. Iida, and Y. Takai, “NS record history based abnormal DNS traffic detection considering adaptive botnet communication blocking,” *IPSJ J. Information Processing*, vol. 28, pp. 112–122, Feb. 2020. DOI: 10.2197/ipsjjip.28.112

[19] H. Ichise, Y. Jin, K. Iida, and Y. Takai, “Detection and blocking of anomaly DNS traffic by analyzing achieved NS record history,” *Proc. Asia-Pasific Signal and Information Processing Association, Annual Summit and Conference 2018 (APSIPA-ASC2018)*, pp. 1586–1590, Honolulu, HI, USA, Nov. 2018.

[20] P. Vixie, and V. Schryver, “DNS response policy zones (RPZ),” *IETF internet-draft*, draft-ietf-dnsop-dns-rpz-00, Mar. 2017.

[21] D. Plohmann, et al., “A comprehensive measurement study of domain generating malware,” *Proc. USENIX Security Symp.*, Austin, USA, pp. 263–278, Aug. 2016.

[22] 井内裕貴, 金　勇, 一瀬　光, 飯田勝吉, 高井昌彰, “NXDOMAIN 応答に着目した DGA を用いるボットの検知・遮断システムの実装と評価,” 電子情報通信学会・技術研究報告, vol. 119, no. 435, IA2019-67, pp. 7–12, 2020 年 3 月.

[23] Y. Iuchi, Y. Jin, H. Ichise, K. Iida, and Y. Takai, “Detection and blocking of DGA-based bot infected computers by monitoring NXDOMAIN responses,” To appear in *Proc. IEEE CSCloud2020*, 6 pages, Aug. 2020.

[24] “Domain generation algorithm,” `https://github.com/baderj/domain_generation_algorithms`, Accessed at Jan. 2020.

[25] N. Koroniotis, N. Moustafa, and E. Sitnikova, “Forensics and deep learning mechanisms for botnets in internet of things: A survey of challenges and solutions,” *IEEE Access*, vol. 7, pp. 61764–61785, May 2019.

[26] F. Rashid, “The fight over encrypted DNS: Explained,” *IEEE Spectrum: Technology, Engineering, and Science News,* `https://spectrum.ieee.org/tech-talk/telecom/security/the-fight-over-encrypted-dns-boils-over`, Nov. 2019.