2007－CSEC－37 (10)
2007／5／25

トラフィック解析に基づくボット検知手法

釘崎 裕司 † 笠原 義晃 †† 堀 良彰 ††† 櫻井 幸一 †††

† 九州大学大学院システム情報科学府情報工学専攻
〒 819-0395 福岡市西区元岡 744
kugisaki@itslab.csce.kyushu-u.ac.jp

†† 九州大学情報基盤センター
〒 812-8581 福岡市東区箱崎 6-10-1
kasahara@nc.kyushu-u.ac.jp

††† 九州大学大学院システム情報科学研究院
〒 819-0395 福岡市西区元岡 744
{hori, sakurai}@csce.kyushu-u.ac.jp

あらまし 近年，ボットの感染が拡大しボットネットが形成され，社会問題となっている．ボットネットに対抗するためには，脆弱性を有するコンピュータの排除が理想である．そのため，ボットに感染しているコンピュータを検知し，注意を促す情報システムが重要である．そこでボットの動作の特徴を利用する検知手法が考えられている．本稿ではIRCを利用したボットが，IRCサーバに接続する際の動作の特徴について調べた．実際に，IRCサーバでよく使われるポートのトラフィックを観測した．それにより，ボットがIRCサーバへの接続を拒否された際に，ある一定の時間間隔で再接続を試みることが確認できた．また，その時間間隔の分布を調べたところ，別のIPアドレスからの通信も類似の挙動を示していることを確認した．

Bot Detection based on Traffic Analysis

Yuji KUGISAKI † Yoshiaki KASAHARA †† Yoshiaki HORI †††
Kouichi SAKURAI †††

†The Department of Computer Science and Communication Engineering, Kyushu University, 744 Motooka, Nishi-ku, Fukuoka 819-0395, Japan,
kugisaki@itslab.csce.kyushu-u.ac.jp

††Computing and Communications Center Kyushu University 6-10-1 Hakozaki, Higashi-ku, Fukuoka, 812-8581 Japan
kasahara@nc.kyushu-u.ac.jp

†††Faculty of Information Science and Electrical Engineering, Kyushu University 744 Motooka, Nishi-ku, Fukuoka, 819-0395 Japan
{hori,sakurai}@csce.kyushu-u.ac.jp

Abstract Recently, botnet becomes a social problem due to the expansion of bot infection. Ideally, all the vulnerable computers should be fortified to counteract botnet. To do that, it is important to implement an information system which detects bot-infected computers and alerts them. Therefore, there are various studies ongoing to detect existence of bots based on characteristics of bot behavior. In this paper, we focused on bots using IRC to communicate, and examined the behavior of such bots when they connected to an IRC server. We observed the actual traffic of some ports which were often used by the IRC protocol. As a result, we confirmed that a bot tried to reconnect to an IRC server at certain intervals when the server refused the connection from the bot. Moreover, we examined the distribution of the intervals and confirmed that the communication from other IP addresses showed similar behavior.

1 はじめに

コンピュータとインターネットは，いまや我々の生活に必要不可欠なものとなっている．その便利さから，コンピュータは爆発的に普及し，低価格・大容量化によりインターネットも一般家庭に浸透した．

しかし，コンピュータのセキュリティに対する意

識が低く，また知識が不十分なままコンピュータを所持し，ネットワークに接続している利用者も増加している．現在では，ネットワーク上の脅威に対する防衛策を取らずにコンピュータをネットワークに接続すると，わずかな時間でコンピュータウイルスなど悪意あるプログラムに感染してしまう [1]．そのような状況の中で，コンピュータウイルスの一種でありながら従来のコンピュータウイルスとは区別されているボットが世界的な問題になっている．

ボットに感染したコンピュータは攻撃者からの命令を受け，その命令通りに動作する．これらボットに感染したコンピュータが何千，何万台とつながってネットワークを作り，ボットネットが構築される．これは攻撃者が自由に操ることができるコンピュータが何万台と存在しているということである．このボットネットが，以前から存在していたネットワークセキュリティ上の脅威であるDDoS攻撃やスパムメール，またフィッシングなどに利用されたりしている．さらにはボットネットをスパム業者に貸すことにより不当に利益を得るなど，攻撃者の資金の源にもなっている [2]．

ボットネットの問題は日本に限るものではない．東アジアや欧米諸国においても深刻な問題となっている．警察庁の発表によると，2005年7月から12月の時点でボットに感染しているコンピュータが多い国として，第一位がアメリカ，第二位が中国，第三位が日本，第四位がドイツとなっている [3]．アメリカやオランダなどでは，ボットネット構築により不当に利益を得たとして逮捕者も出ている [4][5]．

このように世界中でボットが広がっていることを考えると，ボットの根絶にはインターネットサービスプロバイダや国といった枠を超えて対策を練らなければならない．しかし，実際には各国の事情や法律が関わるため難しいと思われる．例えば，日本がボットを規制するような何らかの法律を施行しても，他国のボットは規制できない．またプロバイダもボットから送られるトラフィックかそうでないか，厳密な監視をするといった対策もユーザのプライバシーの観点から難しいであろう．

ボットネットが蔓延している根本的な問題は，セキュリティ対策を施さずにコンピュータをネットワークに接続する利用者にある．そのため，ボットネットを完全に根絶するのは難しいであろうが，ボットに感染しないための予防策を考案すること，またボットに感染したコンピュータを検知しボットを駆除するといったことは可能であり有用である．ボット検知の手法の一つに，ボットの行動の特徴を利用して検出するものがある．この手法はシグネチャを用いてボットを検出する手法よりも広い範囲のボットに対応できるので，注目されている [6]．しかし完全に検知することが可能なわけではなく，誤検知が多いことも問題である [7]．単一の手法を用いるだけでは誤検知や見逃しが多くなるので，複数の手法を組み合わせることが重要であると考えられる．

そのため，本論文ではボットの動作について監視・解析を行い，新たな特徴を見つけ出すことを目的とした．ボットは攻撃者から指令を受ける際に，その指令を伝えるサーバと接続しなければならない．そこでクライアントがサーバと接続する際の動作について監視したところ，ボットと疑わしいクライアントに共通する動作を観測した．そこで実際にそのような動作を行うクライアントがサーバと通信する時間間隔についても調査し，類似のパターンをとることが見られた．

本論文の構成は以下の通りである．2章で背景となっているボットネットの概要と，本研究で着目したIRCの概要を説明する．3章で既存の検知手法について述べる．4章で提案手法について述べる．5章で実際に実験した結果を記し，6章で全体のまとめをする．

2 背景

2.1 ボットネット

ボットネットとはボットなどのウイルスによって，外部の人物によってコントロールされるようになった複数のコンピュータで構成されるネットワークのことである [8]．一般的にボットは感染したコンピュータを利用することを目的としている．そのため，感染したコンピュータのデータの破壊や動作の不安定を引き起こすタイプのウイルスとは区別される．また，そのような被害をもたらす行動を取ることは少ないため，ユーザは感染に気が付きにくい．

ボットはコンピュータに感染すると，常駐し攻撃者からの指令を待つ．攻撃者からの指令を受けるとその指令を実行する．その指令は例えば，DDoS攻撃やスパムメール送信などといったものである．

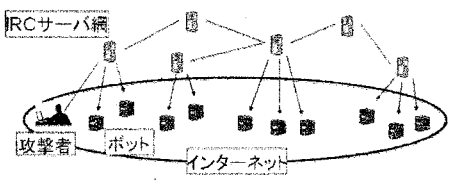

図 1: IRCを利用するボットネットの概念図

図1は一般的なボットネットの概念図である．ボッ

トに感染したコンピュータは指令サーバに接続する．指令サーバとは，攻撃者からボットに感染したコンピュータへの指令を中継するものである．この制御の中心となるサーバをC&C(Command and Control)サーバなどとも呼ぶ．また攻撃者のことをハーダー(Herder)とも呼ぶ．攻撃者の指令は指令サーバを中継してボットに感染したコンピュータに届く．攻撃者からの指令を受け取ったボットは，その指令通りに活動を行う．このように，攻撃者の指令により実際に行動をとるのはボットに感染したコンピュータである．すなわちボットに感染すると，知らないうちにDDoS攻撃やスパムメール送信などの犯罪の加害者になることがある．

2.2 IRC(インターネット・リレー・チャット)

攻撃者がボットに指令を伝える際には，その通信手段として一般的にIRCが使われる[7]．IRCとは，TCP/IPプロトコル上でサーバを介して，クライアントとクライアントがテキストデータを交換することにより会話を行うシステムのことである．IRCのマルチキャスト配送機構により，容易に多数のボットに指令を送ることが可能であるため，攻撃者がボットへの指令の送信にIRCを用いる要因となっている．

ボットに感染したコンピュータは感染するとIRCサーバに接続を試みる．IRCに接続したコンピュータは指定されたチャンネルに参加し，指令を待つボットとなる．指令はユーザとして参加している攻撃者の発言として下され，指令を受け取ったボットはその指令を実行する．IRCサーバがボットに感染して操作されている場合もある．ある一つのIRCサーバが止められてもボットネット全体が止まることは無く，攻撃者を突き止めるのは困難である．

3 既存の検知手法

ここでは，既存の検知手法について代表的なものを挙げる．

3.1 シグネチャベースの検知

ここで言うシグネチャとはパケットの特徴情報のことである．事前にボットのパケットの特徴情報を登録しておき，そのシグネチャに一致したパケットを検出する．これは単純なバイト列を比較するものであるため処理が簡単であり，また定義された不正パケットに対しては確実な検知が可能である．

しかし事前に定義ファイルを作成する必要があるため，未知のボットに対しては対応することができない．また，常に新しいシグネチャの登録が必要となるためシグネチャのデータベースの肥大化が起こり，性能に影響を及ぼしたり管理の負担が増大するといった問題もある．また新たなボットが発見されてから，それに対する定義ファイルを作成し適用されるまでには時間差があるという欠点もある．ボットはソースコードがネット上で公開されており，毎日多数の亜種が作成されている．そのため，シグネチャベースの検知手法の欠点がそのまま弱点となってしまう．

また，最近では従来のような不特定多数を狙うものではなく，特定の企業や団体などに標的を絞ったボットも増えている．このような特徴を持つボットに対しては，従来のウイルスのようなシグネチャによる検出はあまり有効とはいえない．標的となる企業や団体向けにボットが作られており，対象とする範囲が狭いためボットのデータを集めることが困難だからである．

3.2 IRCクライアントとしての振る舞いを利用した検知

現在，ボットネットはそのコントロールにIRCを利用することが多い．そこでIRCの特徴を利用してボットネットを検知する手法がある．

IRCサーバへの接続にはTCPポート番号6667を使用することが規定されているが，多くのIRCサーバでは多数のクライアントを収容するために6667番だけではなく，6666番から6669番といった隣接ポートも利用する．たとえばNIDS(Network-Based Intrusion Detection System) Snortは6666番から6669番までのポートが利用されるトラフィックをIRCトラフィックとして観測するルールを利用している．このような一般的に用いられるポートに流れるIRCトラフィックをモニタする．

IRCクライアントは，利用者がメッセージ交換を行うためのインターフェースであり，利用者が対話的に操作するものであった．一方で，ボットのようにプログラムされメッセージ交換を行うIRCクライアントも現れた．これらの2種類のクライアントにおいて，IRCサーバとの対話的なコマンドおよびメッセージ交換における振る舞いは異なっているのではないかと考える．そこで，ここではプログラムによって動作が決定されるIRCクライアントの振る舞いに着目する．

またIRCで用いられるコマンドにもある種の特徴が見られる．ボットであると思われるコンピュー

タは広い範囲で情報を集める傾向にある．例えば，バンド幅を計測したり，他のサーバへホストの情報を送信したりする．この情報にはオペレーティングシステム，ホストのバンド幅，ユーザ，パスワード，ファイル共有，ファイルの名前やすべてのファイルへの許可など，感染したホストの多くの細かい情報が含まれている．IRC をチャット目的で利用する場合，このような情報や情報を得るためのコマンドはまず必要とされない．よってこれらの情報からボットと判断することが可能である．

しかしこれらの検知手法にも欠点がある．まず IRC でよく使用されるポートはあるが，これは絶対ではなく自由に変えることが可能である．つまり攻撃者によってこのポート番号を変えられるとモニタリングは不可能である．また，トラフィックを暗号化されたり，ランダムノイズでフローの行動を隠されたり，異なる通信方式に切り替えられたりした場合，検知は不可能である [9].

4 提案手法

4.1 概要

そこで本研究では，IRC を利用するボットの新規の特徴を見つけることを考えた．まず IRC で一般的に使用されるポートの監視を行った．その結果，特定の IP アドレスを持つクライアントはサーバと接続する際のコマンドの流れに，その他のクライアントとは差異があることがわかった．

4.2 データ概要

調査においては，6666 番-6669 番ポートのトラフィックを用いた．本論文中で使用しているトラフィックは九州大学で 2006 年 7 月 31 日の 24 時間に流れたものである．

4.3 サーバへの接続の際の動作と特徴

IRC サーバを利用する場合には，IRC サーバと接続しなければならない．その接続には手順が規定されている．クライアントは NICK と USER のコマンドをサーバに対して送信する．この NICK とユーザのコマンドの順番はどちらが先でもよいが，両方を受け取ったのちにサーバはクライアントを登録する．この二つのコマンドが処理されてクライアントとサーバの接続が確立する．次にクライアントは，チャンネルに参加するために JOIN コマンドを送信する．チャンネルに参加することにより，クライアントが互いにメッセージのやり取りを行うことが可能である．それには PRIVMSG コマンドもしくは NOTICE コマンドを使う．

このように一般的な通信の場合は次のような流れで通信を行う．

- NICK → USER → JOIN → PRIVMSG(or NOTICE) → ...

しかし，特定の IP アドレスを持つクライアントはサーバへの接続の際に接続を拒否されるものが見られた．IRC サーバはニックネームの重複やサーバへの過負荷，また疑わしいクライアントの接続などを防ぐため，そのような事態を引き起こすようなふるまいを行うクライアントの接続を拒否する．このようなクライアントは IRC サーバに接続するために，接続が成功するまで NICK と USER を繰り返すこととなる．つまり次のような流れになることが多いと考えられる．

- NICK → (ERROR) → NICK → (ERROR) → ...
- NICK → USER → (ERROR) → NICK → USER → (ERROR) → ...

5 実験

5.1 実験結果

使用したトラフィックは 4 章でも用いた 6666 番-6669 番ポートのトラフィックである．

使用したトラフィックデータのうち，プロトコルの情報を見て IRC を利用するトラフィックに限定した．さらに上で記述した方法で疑わしいクライアントを抽出し，そのクライアントがサーバと通信を行う時間間隔について調査した．

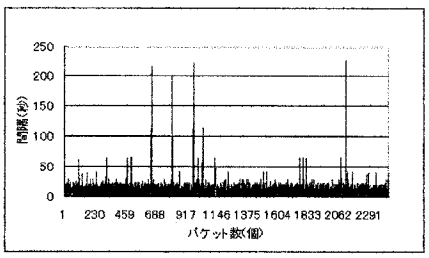

図 2: ボットと推測される IRC サーバへの通信間隔 (6667 番ポート)

図 2 から図 6 まで IRC サーバを利用するボットであると疑わしいクライアントの通信について見た

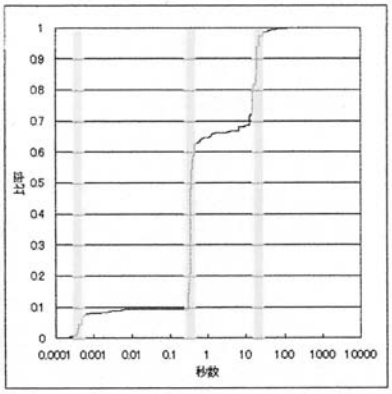

図 3: ボットと推測される IRC サーバへの通信間隔の比率 (6667 番ポート)

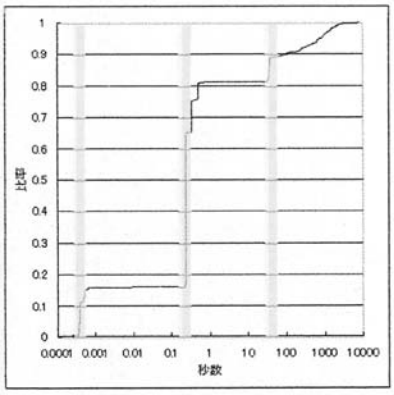

図 4: ボットと推測される IRC サーバへの通信間隔の比率 (6666 番ポート)

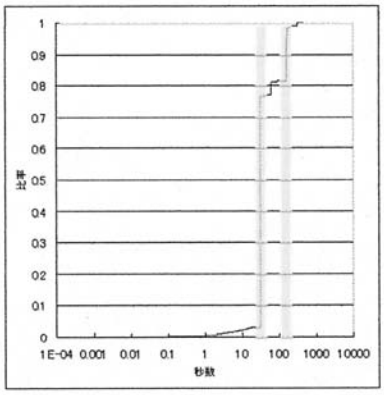

図 5: ボットと推測される IRC サーバへの通信間隔の比率 (6668 番ポート)

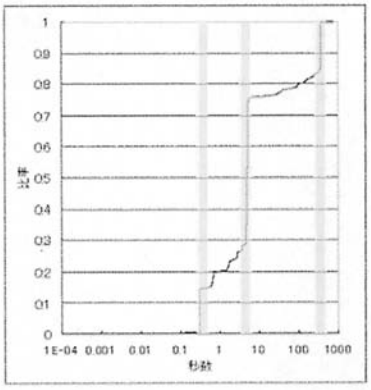

図 6: ボットと推測される IRC サーバへの通信間隔の比率 (6669 番ポート)

ものである．図 2 は IRC サーバへ接続しようとするボットが，IRC サーバと通信を行う間隔を調べたものである．いくつかの例外はあるが，ほぼ決まった間隔で通信を行っていることがわかる．このような一定の時間間隔での動作は，人によって対話的な操作が行われている IRC クライアントでは見られないものである．

図 3 から図 6 は通信の間隔をその分布で表わしたものである．x 軸は通信の間隔の秒数を対数表示したものである．y 軸は比率である．それぞれ分布にいくつかの偏りがあるという共通点があることがわかる．逆に分布の偏りにもその秒数や偏りの個数の違いのような差異があるということもわかる．例を挙げると，図 3 ではおよそ 0.2 秒から 0.3 秒の付近で全体の 5 割を占めていることがわかる．また図 4 では，0.2 秒から 0.3 秒の間隔で通信を行うことが全体のおよそ 6 割程度となっている．

詳細を述べると図 2，3 は 6667 番ポート，図 4 は 6666 番ポート，図 5 は 6668 番ポート，図 6 は 6669 番ポートを利用する通信であった．

図 7 はボットでないと推測されるクライアントについて，その IRC サーバへの通信間隔を調べたものである．図 2 と比較すると，その差がはっきりしていることがわかる．図 7 は図 2 に比べて，偏りが少なくばらつきが大きい．

また図 8 はその通信の間隔をその分布で表したものである．ボットと推測されるクライアントの分布を表した図 3 から図 6 とは異なり，その分布に偏りが少ないことがわかる．

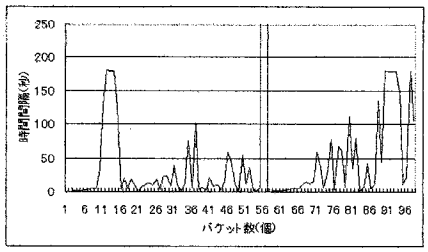

図 7: ボットでないと推測されるクライアントのIRCサーバへの通信間隔

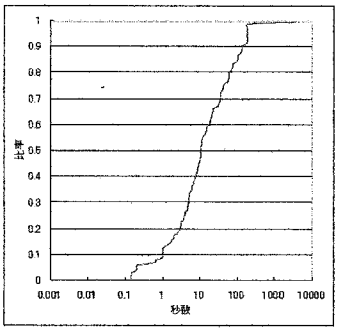

図 8: ボットでないと推測されるクライアントのIRCサーバへの通信間隔の比率

5.2 まとめと考察

本章では，利用者によって対話的に操作が行われるIRCクライアントと，ボットとの挙動が異なると考えて，IRCサーバで一般的に使用されるポートのトラフィックを調べた．そしてクライアントとIRCサーバとの通信間隔がどのくらいの時間であるかをグラフに示した．すると，ボットであると推測されるクライアントのトラフィックには共通した特徴が見られた．どのトラフィックにおいても，サーバへの通信の時間間隔には偏りがあることが確認された．反対に，共通した特徴の中にもその分布の偏りや秒数などには差異があることがわかる．

6 終わりに

本論文では，ボットの振る舞いに関する特徴をもとにボットを検知する手法について，新たな特徴について調査することを目的とした．

まず，IRCを用いたボットについてその動作の特徴について考察した．次に，そのボットの動作の特徴がトラフィックではどのようなパターンで見られるのか，IRCで一般的に用いられるポートのトラフィックをキャプチャし調べた．ボットと推測されるクライアントは，ボットでないと推測されるクライアントと比較して，異なる通信パターンを取ることが観測された．

このようにトラフィックを可視化することで，様々な利点が生まれると考えられる．

- 従来のウイルスよりも多いといわれるボットの種別の特定
- 機械学習によるボット検知の自動化

今後の課題としては，

- より多くの対象の調査 (検証や一般性の問題)
- 具体的な可視化による検知の考案
- IRCを利用しないボットへの対策

などがある．

参考文献

[1] Time to live on the network. Avantgarde Marketing & Design, http://www.avantgarde.com/xxxxttln.pdf. 2004.

[2] David Geer. Malicious Bots Threaten Network Security, Industry trends, January 2005, 18-20, 2005.

[3] 警察庁, @ police. 平成 17 年下半期 (7 月～12 月) における botnet 観測システム観測結果, *http://www.cyberpolice.go.jp/detect/pdf/20060316_botnet.pdf*.

[4] WORLD WIDE WEB, From content types, http://pc.nikkeibp.co.jp/article/NPC/20070413/268240/

[5] WORLD WIDE WEB, From content types, http://japan.cnet.com/news/sec/story/0,2000056024,20088605,00.htm

[6] 河本貴則, 秋山満昭, 横山輝明, 門林雄基, 山口英. ボットネットの協調動作に注目した検出手法の一検討, SCIS2007, 2007.

[7] Evan Cooke, Farnam Jahanian, Danny McPherson. The Zombie Roundup Understanding, Detecting, and Disrupting Botnets, Steps to Reducing Unwanted Traffic on the Internet Workshop, 39-44, 2005.

[8] Ramneek Puri. Bots and Botnet: An Overview, SANS Institute 2003, GSEC Practical Assignement Version 1.4b, 2003.

[9] James R., Binkley, Suresh Singh. An Algorithm for Anomaly-based Botnet Detection, SRUTI '06 2nd Workshop on Steps to Reducing Unwanted Traffic on the Internet, 43-48, 2006.