Botnetの命令サーバドメインネームを用いたBot感染検出方法に関する研究

朝長　秀誠 †　　　田中　英彦 †

† 情報セキュリティ大学院大学
〒 221-0835 神奈川県横浜市神奈川区鶴屋町 2-14-1

{mgs051103,tanaka}@iisec.ac.jp

あらまし　近年，コンピュータウィルスの一種である Bot が大きな問題となっている．その理由は，Bot に感染したコンピュータ (Bot 感染コンピュータ) が攻撃者の命令を仲介する命令サーバを中心にネットワーク (Botnet) を形成し，攻撃者からの命令を受信することである．Bot はそのプログラムがインターネット上で公開されているものが多く，それを用いることで誰でも簡単にその亜種を作成することができる．このため，現在主流となっているシグネチャマッチングタイプのアンチウィルスソフトウェアでは，パターンファイルの作成が追いつかなくなり，Bot の検出が難しくなっている．そのため，現在は Bot の挙動を用いる検出方法に対する研究が主流になっている．本研究では，Bot が DNS サーバに命令サーバの FQDN(Fully Qualified Domain Name) をクエリする挙動を用いることで Bot を検出する方法を提案する．

Research on method of detecting Bot infection using Botnet C&C server domein name

Syusei Tomonaga†　　　Hidehiko Tanaka†

†INSTITUTE of INFORMATION SECURITY
2-14-1, Turuya-chou, Turumi-ku, Yokohama-shi, Kanagawa 221-0835, Japan

{mgs051103,tanaka}@iisec.ac.jp

Abstract　テストテテストストテストテストテストテストテストテストテストテストテストテスト

1　はじめに

近年，コンピュータウィルスの一種である Bot が大きな問題となっている．その理由の一つは，Bot 感染コンピュータが攻撃者の命令を仲介する命令サーバを中心にネットワークを形成し，攻撃者からの命令を受信することである．従来のワームやコンピュータウィルスはコンピュータに感染すると予めプログラムされた同一の動作しか行わないことに対して，Bot は攻撃者からの命令を受信することで一斉に同じ攻撃動作を行ったり，その動作を変更したりすることができる．ここで Bot 感染コンピュータによって構成されるネットワークを Botnet と呼ぶ．Botnet の多くは IRC(Internet Relay Chat) プロトコルを利用して命令サーバと Bot 感染コンピュータ間の通信を行う．

また，Bot の種類がその亜種を含めて非常に多いことも一つの理由である．Bot はそのプログラムが

インターネット上で公開されているものが多く，これを用いることで誰でも簡単にその亜種を作成することができる．また，公開されている Bot プログラムの中には GUI が含まれているものもあり，プログラムの知識がない人でも Bot を作成することができる．このため，現在主流となっているシグネチャマッチングタイプのアンチウィルスソフトウェアでは，パターンファイルの作成が追いつかなくなり，Bot の検出が難しい．定点観測によって収集された Malware の 80 %が未知のものである [1] という結果からもパターンファイルによる Bot の検出は難しいことが分かる．

このため，現在は Bot の挙動を用いる検出方法に対する研究が主流になっている．本研究では，Bot が DNS サーバに命令サーバの FQDN をクエリする挙動を用いることで Bot を検出する方法について検討する．

本稿では，Botnet のメカニズムついて説明した後，その後，Bot を検出する方法，システムの構成を説明した後に，今後の研究課題を述べる．

2 Botnet とは

Botnet は Bot 感染コンピュータが攻撃者の命令を仲介する命令サーバを中心に形成するネットワークである (図 1)．Bot 感染コンピュータが攻撃者の命令を受信するために命令サーバと通信を行う際に，IRC のようなスター型と，P2P のようなランダム型が考えられるが，本研究ではスター型を対象とする．現在の主流がスター型であり，ランダム型の通信用途は，スター型の補足的な役割であるために P2P を利用しても検出する可能性がある．

Bot は DoS 攻撃や SPAM メール送信をはじめとする様々な攻撃機能が組み込まれたコンピュータウイルスの一種である．ワームと同様にメールの添付ファイルから感染したり，OS やサービスの脆弱性を利用してネットワークから感染したりする．本研究では，システムの脆弱性によるネットワークから感染する Bot を対象にしてその検出方法を検討する．

命令サーバは攻撃者が Dynamic DNS を利用し，独自に用意した IRC サーバで，主に第三者のサーバが無断で利用される．Bot 感染コンピュータはこの命令サーバの FQDN を元に DNS サーバで名前解決をして得られた IP アドレスが示す命令サーバに接続する．攻撃者は命令サーバを介して複数の Bot に命令を出すことで Bot 感染コンピュータに何らかの攻撃動作を同時に実行させることができる．本研究では，この Bot が DNS サーバに名前解決をするという特徴を用いることで Bot の感染を検出する方法について検討する．

図 1: オーソドックス Botnet

3 提案方式

3.1 提案方式の特徴

本研究では Botnet の命令サーバドメイン (FQDN ブラックリスト) を Bot から自動的に抽出し，それを元にネットワークの DNS アクセスの中から Bot 感染を検出する方法を提案する．現在主流となっているシグネチャマッチングタイプのアンチウィルスソフトは専門のウイルス解析者がバイナリに含まれる特長的な文字列を元にシグネチャを作成している．よって，ウイルスについて知識のない人にはこの方法でオリジナルのシグネチャを作成しウイルス検出を行うことは難しい．本方式では，Bot に含まれる命令サーバドメインをシグネチャに使用するのでウイルス解析者でなくてもシグネチャの作成が可能である．また，命令サーバドメインという決められた値を利用するので自動的にシグネチャを作成することも可能になる．

この方法を採用する理由は，ネットワークアドレスにより流行する Malware が違うことが上げられる．現在の Bot は，Blaster を代表とするワームとは異なり，強力な感染能力を持つわけではない．攻撃者からの指示に従って，感染の範囲や規模を制御されている．そのため Bot は，ある範囲のネットワークに特化して流行する傾向にある．Bot の感染メカニズムは IP アドレスの最上位 8 ビットや，16 ビットが一致するアドレス群での感染が多い．よって，

自らのネットワークに適用するウイルス検知システムは，自ら作成することで，自分自身のネットワークで流行している Bot の感染を阻止できる可能性が高いといえる．

3.2 提案方式の構成

本方式は，3 つのシステムからなる．

- Malware 収集システム
- Malware 解析システム
- Bot 検出システム

Malware 収集システムでは，FQDN ブラックリストを作成するために Bot を収集する．Malware の収集にはハニーポットを用いる．ハニーポットとは，クラッカーの侵入手口や Malware の振る舞いを研究するために，ネットワーク上に設置された，脆弱性を持つシステムのことである．このハニーポットは 2 種類に分類することができる．

- ローインタラクション型ハニーポット
- ハイインタラクション型ハニーポット

ローインタラクション型ハニーポットとは，システムやサービスをエミュレートすることによって機能するもので，送られてきたコマンドに対して考えられるレスポンスを返す．よって，進入される危険性は少ない．それに対してハイインタラクション型ハニーポットとは，進入可能な実際のアプリケーションを備えたシステムで，オペレーティングシステム全体とアプリケーション全体を提供する．よって，システムに侵入したブラックハットの詳細な情報を得ることができる．しかし，実際に侵入されるというリスクを持っている．

Malware 解析システムでは，Malware 収集システムで収集された Malware の中から Bot のバイナリを解析し，それから得られた Botnet の命令サーバの FQDN を Bot 検出システムに転送する．Malware の解析には静的解析と動的解析を組み合わせた方法を用いる．その理由は，Bot に複数の FQDN が含まれている場合があり [1]，動的解析または，静的解析のみで FQDN を抽出しようとする際，全ての FQDN を抜き出せない可能性があるからである．

Malware 解析は，動的解析から行う．WindowsOS 上で Malware を実行し，Bot とその他の Malware に分ける．そして，Bot のみを実行し，実行時の DNS サーバとの通信をキャプチャし，その中から命令サーバの FQDN を抽出する．次に静的解析を行う．Bot のバイナリを実行しメモリ上に展開されたプロセスをメモリダンプし，デバッガを用いてバイナリを逆アセンブルする．そのコードの中に含まれる B 全ての FQDN を抽出し，命令サーバの FQDN かを判断する．静的・動的解析で抽出した FQDN を FQDN ブラックリストとして Bot 検出に利用する．動的・静的解析時には Malware の耐解析機能が FQDN 抽出の妨げとなる．現在，全ての Malware の解析が可能な状態ではないが，このような Malware の耐解析機能を無効化する対策を今後組み込むことで，この方式の精度の向上を図ることができる．4.2.1 では，現在行っている耐解析機能を無効化する方法について述べる．Malware 解析システムの流れを図 2 で示す．

図 2: Malware 解析の流れ

Bot 検出システムでは，Malware 解析システムで得られた FQDN を元に FQDN ブラックリストを作成し，それを元に DNS サーバへのアクセスから Bot のアクセスを検出する．Malware 解析システムで Bot から抽出した FQDN ブラックリストは，Bot が命令サーバに接続するために DNS サーバに A レコードの正引きをする FQDN のリストである．そのため，DNS サーバへの正引きアクセスと FQDN ブラックリストを照合し，一致した場合は Bot からのアクセスと考えられる．Bot 検出ステップの流れを図 3 で示す．

図 3: Bot 検出システムの流れ

4 実装

本章では 3 章で述べた方式を実現するシステムについて説明する．

4.1 Malware 収集システム

本研究では Malware 収集にローインタラクション型ハニーポットである Nepenthes v1.70[6] を用いる．このハニーポットは Malware 収集用のハニーポットで，Malware から攻撃を受けた場合，受け側のエミュレートを行い，仕込まれるファイルを取得する．その際，知られている脆弱性を模倣して，Malware との通信を行うことで Malware のファイルをダウンロードするが実行することはない．ローインタラクション型ハニーポットを用いた場合 Malware が感染するようなリスクは低くなる．しかし，既知の脆弱性を用いる Malware しか収集することができず，実際の WindowsOS をハニーポットとして利用した場合と比較して，Malware の収集能力が低下してしまう問題点がある．

ハニーポットで収集した Malware は随時 Malware 解析環境に転送する．

4.2 Malware 解析システム

Malware 解析システムでは Malware 解析環境を構築し，その中で Malware の解析を行い命令サーバの FQDN を抽出する．マシンを 2 つ準備し，Malware 実行クライアントと仮想サーバとする．Malware 解析環境の構成を図 4 に示す．

図 4: Malware 解析環境

4.2.1 Malware 実行クライアント

Malware 実行クライアントでは Malware を実行し，実行された不正プロセスのメモリダンプを行い，解析を行う．Malware 実行クライアント上では VMware Workstation5.0 を用意．その上で WindowsXP　(SP0) をインストールし，その上で Malware を実行する．VMware Workstation はスナップショット機能があり，Malware 実行後の WindowsOS を感染以前の状態に容易に戻すことが可能である．また，コマンドラインでの実行か可能なため，自動的に仮想マシンを立ち上げて，解析を行うことができる．

しかし，現在の Malware の多くは仮想マシン上で実行されたことを検知することで，プロセスを停止させるものが存在する．これは仮想マシン上で Malware を実行させて動作解析を行うことが，Malware 解析方法の 1 つだからである．Malware が仮想マシンを検知する可能性があるポイントは以下のようなものである．

1. VMware Backdoor I/O
2. 仮想ホストの MAC アドレス
3. Video BIOS
4. SCSI デバイスのデバイス名
5. IDE デバイスのデバイス名

6. vmnat の MAC アドレス

このような仮想マシン上から得られる情報を元に VMware は検知することができる．Malware の持つ VMware を検知する耐解析機能を回避するためには，これらの情報が仮想マシン上から得られないようにする必要がある．それを可能にするために [3] では，VMware のバイナリコードを書き換えるパッチを提供している．このパッチの特徴としては，(1)〜(5) の偽装が可能となる．しかし，これだけでは vmnat の MAC アドレスは変更できないので，バイナリを直接書き換える必要がある．それ以外にも vmx ファイルにオプションとして付け加えることで，耐解析機能を回避することが可能である．

表 1 では，4.1 のハニーポットを利用して収集した Malware を用いて，VMware Workstation5.0.0 に (1)〜(6) の特徴を無効化した状態で Malware を実行した場合と，無効化しない状態で実行した場合の比較を行った．結果として，実行可能な Malware の中で，35 個が動作するようになった．動作しなかった多くの Malware は命令サーバからの指令が受信可能かなどのネットワーク環境を検知して停止するものだったため，実際の命令サーバなどに接続できるような環境にすることで，より多くの解析が可能になると思われる．また，その他の仮想ホストを検知する特徴の回避を行う必要がある．

実行された不正プロセスのコードを得るためにはデバッガを用いる．しかし，Malware はデバッガの解析を阻止する機能を持つものが存在する．また，難読かによって FQDN 文字列を抽出できない場合も存在する．よって，この耐解析機能を回避する方法の検討が必要である．

4.2.2 仮想サーバ

仮想サーバは，各種サーバを模擬した機能と，Malware 実行クライアントに Malware を転送する機能を持つ．サーバの模擬は Malware の実行時にネットワークアクセスを把握するために，実際のネットワークに近い状況を作り出し，Malware と通信を行う．仮想サーバは DNS サーバ，IRC サーバ，SMTP サーバ，FTP サーバの 3 つから構成される．

	動作	停止	実行不能	総数
対策なし	1,169	177	308	1,654
対策あり	1,204	142	308	1,654

表 1: VMware 耐解析 Malware の調査

DNS サーバはクエリがあった場合，偽装した IP アドレスを返す．この場合，感染ホストの利用する DNS サーバ以外のアドレスにクエリしたとしても，この模擬サーバでアクセスを返す．IRC・FTP サーバは，ポートに対してアクセスがあった場合，3Way ハンドシェイクを行う．SMTP サーバはメールの送信が行われた場合，SMTP のアクセスを模擬する．

また，Malware 実行クライアントで実行した Malware の通信を制御するために iptable を用いる．52, 80 などいくつかのポートを ACCEPT し，Bot からの IRC 通信があると考えられるポートを仮想サーバの 6667 ポートにリダイレクトする．そして，DNS アクセスを含めた通信を tcpdump でキャプチャする．

4.3 Bot 検出システム

Bot の検出は DNS サーバへの A レコードの問い合わせから FQDN を抽出し，FQDN ブラックリストとマッチングをすることで行う．現在の Bot は DNS サーバへクエリする際，感染クライアントが利用するローカルの DNS サーバを利用するが，今後インターネット上の他の DNS サーバを利用する場合も考えられる．よって，ローカルの DNS サーバへのアクセス監視だけでなく，外向きの DNS サーバアクセスをネットワーク上で監視する必要がある．それ以外の方法として，外向きの DNS サーバアクセスを全てリダイレクトして，ローカルの DNS サーバへアクセスを変えてしまう方法も考えられる．本研究では，後者を採用する．

DNS サーバへのアクセスと対応付ける FQDN ブラックリストは Malware の収集数に応じて増加することが考えられる．よって，線形検索では DNS サーバのレスポンスタイムが減少してしまう．そこで本研究では，ハッシュ法などの高速検索アルゴリズムを利用する．また，FQDN ブラックリストは古くなると Botnet に利用されていない FQDN が増加し，不要な検索が増加してしまう．それを回避するために一定期間ハニーポットに感染がない Bot の FQDN は消去することで，FQDN ブラックリストの巨大化を防止する．

検出の流れは，DNS サーバへのアクセスは全てインターセプトし，A レコードの正引きだった場合に，

図 5: 検証方法

パケットに含まれる FQDN と FQDN ブラックリストを照合し，一致した場合はそのクライアントに対して 127.0.0.1 のアドレスを結果として返す．そして，アクセスのあった IP アドレスを Bot 感染端末として表示する．

5　本手法の有効性

ここでは，本方法の有効性の評価を行う．提案方法の場合，監視するネットワークにアクセスがある Bot を，いくつ集めることが出来るかが Bot 検出の確率を上昇させる．よって，監視ネットワークで多くの IP アドレスを利用して，ネットワーク感染する Malware を収集することで提案方法の精度が上昇する．

そこで，複数の IP アドレスを用いて構築したハニーポットで 1 週間収集した Malware が，近接 IP アドレスで 1 日に感染してくる Malware の一致率を検証する．これにより，いくつの IP アドレスを用いてハニーポットを構築することで，Bot 検出に十分な FQDN ブラックリストが作成できるかが分かる．検証に利用する計算式を図 5 に示す．

4 つの連続する IP アドレス 2 組 (60.32.xxx.xxx, 222.151.xxx.xxx) 利用して 1 週間収集した Malware を元に，それぞれの近接 IP アドレスで 1 日にアクセスがあった Malware の一致率を測定した (表 2).

Malware 収集期間:8 月 20 日〜26 日 測定日:8 月 26 日

Malware 収集地点	A B	B C	A C	A B C
Malware 一致率	0.94	0.78	0.84	1.00

Malware 収集期間:9 月 7 日〜13 日 測定日:9 月 13 日

Malware 収集地点	X Y	Y Z	X Z	X Y Z
Malware 一致率	0.94	0.94	0.94	0.94

表 2: Malware 一致率

それぞれの連続する IP アドレス群で 3 つの (A,B,C と X,Y,Z) ハニーポットと，1 つの感染ハニーポットを用意する．そして，それぞれのハニーポットで 1 週間収集した Malware を A のみ，A と B，A と B と C など組み合わせによって 1 週間に検出したユニークな Malware とした．そして，感染ハニーポットで 1 日に検出したユニークな Malware を照合して一致した Malware の数を出す．それにより，Malware の一致率を測定する．

この結果からハニーポットを近接 IP アドレスに 3 つ以上設置することで本方式の有効性が上昇すると考えられる．

6　今後の課題

今後は以下の課題を改善するような方法を検討する．

1. **耐解析機能の回避**
   現在の Malware 解析システムでは，仮想ホストやネットワーク環境を検知され全ての Malware を解析することはできない．仮想ホストについては他の仮想化技術を用いることで，解析率を向上することができる可能性がある．ネットワーク環境検知については，直接命令サーバや Web サーバにアクセスできるような環境に変更することで回避できる．その際，ポートスキャンなどの攻撃動作を外部へ出さない環境の構築が必要になる．

2. **FQDN 抽出**
   本方式では実行された不正プロセスのコードから FQDN を抽出する．Bot には命令サーバの FQDN 以外に複数の FQDN が存在する．その中から命令サーバの FQDN を抽出する方法としては，動的解析から得られた命令サーバ FQDN とのメモリ上の近接関係から判断している．しかし，この方法が全ての Bot に対応できるかは未確認である．

3. **Bot 検出の限界**
   本方式では全ての Bot を検出することはできない．Bot の特徴の 1 つとして，プログラムのアップデートを行うことができるということがある．これにより，1 つの Bot 検出の手法が開発されたとしても，その検出を回避す

るアップデートが行われた場合，その検出方法は無効となってしまう．よって，既存の Bot 感染拡散を検出する手法では全ての Bot を検出できないといえる．そのため，Bot 感染検出のためには 1 つではなく，様々な Bot の特徴を捉える方法を融合することで検出精度の向上をはかることができると考える．本方式は，その 1 つをになう方法であるといえる

参考文献

[1] 高橋正和, 村上純一, 須藤年章, 平原伸昭, 佐々木良一, ”フィールド調査によるボットネットの挙動解析”, 情報処理学会論文誌，vol.47 No.8 p2512-p2523，Aug.2006.

[2] Nepenthes - finest collection - `http://nepenthes.mwcollect.org/`

[3] French Honeynet Project-VMware fingerprinting counter measures `http://honeynet.rstack.org/tools/vmpatch.c`