ダウンロードホストに着目したマルウェアの活動傾向分析

石井　宏樹 †　　佐藤　和哉 †　　田端　利宏 ‡

† 岡山大学工学部
700-8530 岡山市津島中 3-1-1
{h-ishii, sato}@swlab.cs.okayama-u.ac.jp

‡ 岡山大学大学院自然科学研究科
700-8530 岡山市津島中 3-1-1
tabata@cs.okayama-u.ac.jp

あらまし　近年，マルウェアの一種であるボットによる被害が問題となっている．これに対して，ネットワークを介したマルウェアの感染動作，およびボット感染時の情報に着目した研究が行われている．一方で，ダウンロードホストを重視した調査は少ない．本稿では，ダウンロードホストとマルウェアの関係，およびマルウェアの活動の特徴に着目し，研究用データセット CCC DATAset 2008 の攻撃元データから分析したマルウェアの活動傾向について述べる．

Analysis of Malware Activities Based on Download Hosts

Hiroki Ishii†　　Kazuya Sato†　　Toshihiro Tabata‡

†Faculty of Engineering, Okayama University
3-1-1, Tsushima-naka, Okayama 700-8530, Japan

‡Graduate School of Natural Science and Technology, Okayama University
3-1-1, Tsushima-naka, Okayama 700-8530, Japan

Abstract　Recently, damage caused by bot is one of serious problem. The bot is a kind of malware. Therefore, there are studies based on infection behavior of malware through network and on the information of the bot infected state. On the other hand, there is few studies based on download hosts. In this paper, we describe an analysis of malware activities analyzed from the attack data of "CCC DATAset 2008". Specifically, we focus on the relation between download hosts and malware and the characteristics of malware activities.

1　はじめに

現在，ネットワーク上には多数のマルウェアが存在している．文献 1) によると，2007 年度の日本国内におけるマルウェアの被害報告数は，63,726 件であり，被害件数は多い．また，近年，マルウェアの一種であるボットによる被害が大きな問題となっている．ボットとは，ネットワーク経由で操作可能なマルウェアのことである．

ボットの特徴の 1 つとして，感染した計算機で「ボットネット」と呼ばれるネットワークを形成することが挙げられる．攻撃者は，形成されたボットネットを用いて，スパムメールの大量送信や特定サイトへの DDoS 攻撃といった攻撃活動を行う．このため，ボットにはウイルスやワームよりも被害の規模が大きいという特徴がある．また，ボットは，自分自身の存在を隠蔽する機能，および攻撃者によってボットの更新を可能にする機能を持つ．これらの特徴により，ボットの発見と対策が困難になっている [2)]．

以上のことから，ボットをはじめとするマルウェアの解析や対策法の検討が行われている．文献 3) では，Nepenthes を用いて収集したマルウェア情報を元に，ネットワークを介したマルウェアの感染動作に着目した調査が行われてい

る．文献 4) では，フィールド調査により，ボット感染時の動作，感染後の動作，およびボットのソースコードの解析が行われている．また，ボットネットの基本的対策として，攻撃を受けた際の対策，攻撃力を軽減するための対策，およびボットに利用される計算機等を減らすための対策がある．一方で，ダウンロードホストを重視した研究は少ない．

そこで，本論文では，研究用データセット CCC DATAset 2008 の攻撃元データ（以降，攻撃元データ）のダウンロードホストに着目し，マルウェアの特徴を分析した．本調査の目的は，マルウェアの特徴を明らかにし，マルウェアのダウンロードを抑制するための指標となる情報を示すことである．具体的には，ダウンロードホストの活動期間，IP アドレスとマルウェアの種類数の関係，シグネチャファイル配布前後でのダウンロード回数の傾向，マルウェアの 1 日の平均ダウンロード回数と活動期間との関係，および IP アドレスとマルウェア名の関係について述べる．

2　調査内容

調査項目を以下に示す．

(1) ダウンロードホストの活動期間
(2) ダウンロードホスト IP アドレスとマルウェアの種類数の関係
(3) シグネチャファイル配布前後でのダウンロード回数の傾向
(4) マルウェアの 1 日の平均ダウンロード回数と活動期間の関係
(5) ダウンロードホスト IP アドレスとマルウェア名の関係

以下に，各項目の詳細を述べる．

(1) ダウンロードホストの活動期間

マルウェアの実行ファイルは，ダウンロードホストからダウンロードされる．このため，ダウンロードホストの活動期間には特徴があると考えられる．活動期間とは，調査期間内で初めてダウンロードされた日から，以降全くダウンロードされなくなる日までの期間のことである．調査期間は，2007 年 11 月 1 日から 2008 年 4 月 30 日までの 6ヶ月間である．

本調査では，ダウンロードホスト IP アドレス（258,711 個）のうち，ダウンロード回数の上位 20 個を抽出して，1 日単位の傾向を分析した．抽出した IP アドレスの数が 20 個である理由は，調査期間中の総ダウンロード回数 2,942,221 回に対し，上位 20 個の総ダウンロード回数は 1,116,347 回で全体の約 38 %を占めていることである．このため，上位 20 個を調査することにより，十分な傾向分析ができると判断した．

(2) ダウンロードホスト IP アドレスとマルウェアの種類数の関係

数種類のマルウェアがダウンロードされるダウンロードホスト IP アドレスからの通信を遮断することで，複数のマルウェアからの感染を防げると考えられる．そこで，調査期間中に総ダウンロード回数の多いダウンロードホスト IP アドレスから，ダウンロードされるマルウェアの種類数の傾向を調査した．本調査では，(1) で抽出した，20 個のダウンロードホスト IP アドレスを使用した．

(3) シグネチャファイル配布前後でのダウンロード回数の傾向

シグネチャファイルの配布により，マルウェアを駆除できるようになると，攻撃者は，解析されたマルウェアを使用しなくなる．あるいは，使用困難になると考えられる．このため，解析されていない状態（UNKNOWN）と解析後の状態（マルウェア名が付けられている）の比較から，ダウンロード回数の減少などの傾向が得られると考えられる．

そこで，マルウェアのハッシュ値ごとのダウンロード回数の変化とシグネチャファイルの配布日の関連について調査した．

(4) マルウェアの 1 日の平均ダウンロード回数と活動期間との関係

マルウェアの存在が明らかになると，解析が行われるため，攻撃者はそのマルウェアを使用しなくなる．あるいは，使用困難になると考えられる．このため，1 日の平均ダウンロード回数が多いマルウェアほど存在が明らかになりやすいため，活動期間が短いと考えられる．一方，1 日の平均ダウンロード回数が少ないマルウェアほど，活動期間が長いと考えられる．

そこで，マルウェアごとの 1 日の平均ダウンロード回数とそのマルウェアの活動期間について調査した．1 日の平均ダウンロード回数とは，

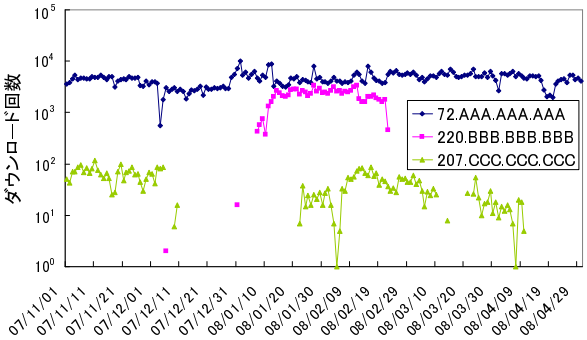

図 1　日付とダウンロード回数の関係

表 1　抽出した 20 個の活動期間の分類

特徴	個数
(1) 長期間	5 個
(2) 短期間	12 個
(3) 複数期間	3 個

マルウェアの総ダウンロード回数をマルウェアの活動期間で割ったものである．

(5) ダウンロードホスト IP アドレスとマルウェア名の関係

マルウェアの実行ファイルは，ダウンロードホストからダウンロードされる．このとき用いられる IP アドレスは，攻撃者によって指定される．このため，マルウェア名とダウンロードホスト IP アドレスには，関連性があると考えられる．また，特定のマルウェアに着目したとき，用いられるダウンロードホスト IP アドレスの活動期間には，関連があると考えられる．

そこで，本調査では，マルウェア名とダウンロードホスト IP アドレス，および IP アドレスの活動期間の関連性について調査した．

3　調査結果

3.1　ダウンロードホストの活動期間

調査結果を図 1 に示す．抽出した上位 20 個のダウンロードホスト IP アドレスについて分析すると，ダウンロードホスト IP アドレスと活動期間の関連について，以下の 3 種類に分類できる．

(1) 長期間（3ヶ月以上）ダウンロードホストとして活動しているもの

例：72.AAA.AAA.AAA

(2) 短期間（3ヶ月未満）ダウンロードホストとして活動しているもの

例：220.BBB.BBB.BBB

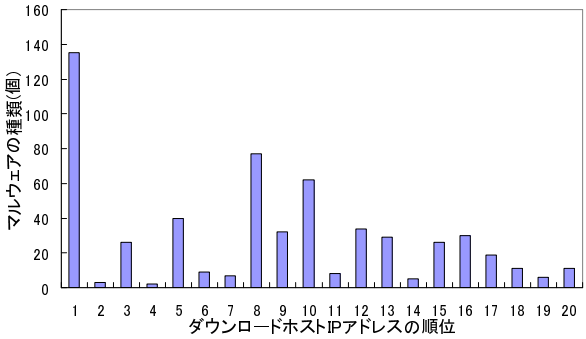

図 2　ダウンロード IP アドレス上位 20 個のマルウェアの種類数

(3) 長期間ダウンロードホストとして活動しており，活動期間の間に 1ヶ月以上活動していない期間を含むもの（複数期間）

例：207.CCC.CCC.CCC

さらに，抽出した 20 個のダウンロードホスト IP アドレスを (1)〜(3) に分類した．分類した結果を表 1 に示す．表 1 から，(2) に該当するものが多いことがわかる．

このことより，攻撃者は，ダウンロードホストが停止させられた場合にもマルウェアを頒布し続けるために，ダウンロードホストの IP アドレスを頻繁に変更していると考えられる．

3.2　ダウンロードホスト IP アドレスとマルウェアの種類数の関係

調査結果を図 2 に示す．また，各ダウンロードホスト IP アドレスにおけるマルウェア 1 種類あたりの平均ダウンロード回数，平均ダウンロード回数の標準偏差，およびマルウェアの種類数について表 2 に示す．

図 2 より，マルウェアの種類数の上位 2 つは 1 位と 8 位であり，下位 2 つは 2 位と 4 位であることがわかる．このことより，ダウンロード回数とダウンロードされたマルウェアの種類数には，関連性はみられない．

しかし，図 2，および表 2 の調査結果から，調査期間中にダウンロードされたマルウェアの種類とダウンロード回数に関して，以下の 2 種類に分類できる．

(1) ダウンロード回数の大部分を 1 から数種類の特定のマルウェアが占める

(2) ダウンロード回数の大部分を占めるマルウェアが存在しない

表 2　抽出した 20 個の平均ダウンロード回数，標準偏差，およびマルウェアの種類数

順位	平均ダウンロード回数	標準偏差	種類数
1	6125.6	9226.9	135
2	33793.7	47778.7	3
3	1296.6	1933.9	26
4	9701.0	9693.0	2
5	478.2	752.6	40
6	1523.8	1464.3	9
7	1701.7	1665.9	7
8	139.4	147.1	77
9	326.8	559.4	32
10	158.1	170.6	62
11	1178.6	1867.0	8
12	254.6	602.7	34
13	277.5	539.4	29
14	1267.0	1021.6	5
15	187.5	111.6	26
16	154.7	201.4	30
17	241.2	409.7	19
18	409.5	302.6	11
19	723.0	833.7	6
20	338.8	265.1	11

また，総ダウンロード回数の少ないダウンロードホストについても調査した結果，同様に 2 種類に分類できる．

このことより，平均ダウンロード回数，標準偏差，および種類数を基に，(1) の特徴を持つダウンロードホストを特定することで，特定のマルウェアの拡大を防止できると考えられる．また，(2) の特徴を持つダウンロードホストを特定することによって，多種のマルウェアの拡大を防止できると考えられる．

3.3　シグネチャファイル配布前後でのダウンロード回数の傾向

特徴が見られた代表的なものについて，調査した結果を図 3〜図 5 に示す．

図 3 のようにシグネチャファイル配布日前後でダウンロード回数が減少するマルウェアも存在する．しかし，実際には，図 4 のようにダウンロード回数がほぼ一定のマルウェア，および図 5 のように活動開始日付近のみダウンロード回数が多く，その後は急激に減少するマルウェアがほとんどである．このため，図 3 は，シグネチャファイルの配布日が図 5 の急激に減少する時期に偶然重なっただけと考えられる．

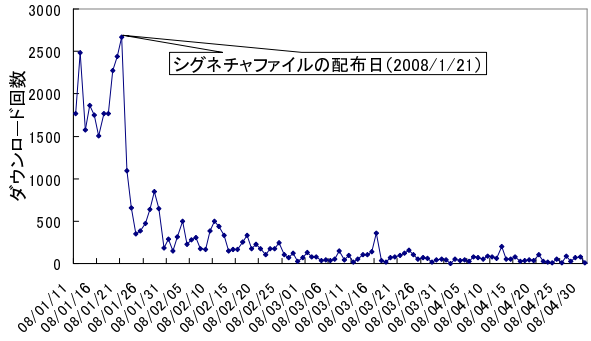

図 3　シグネチャファイル配布後にダウンロード回数が減少するマルウェア

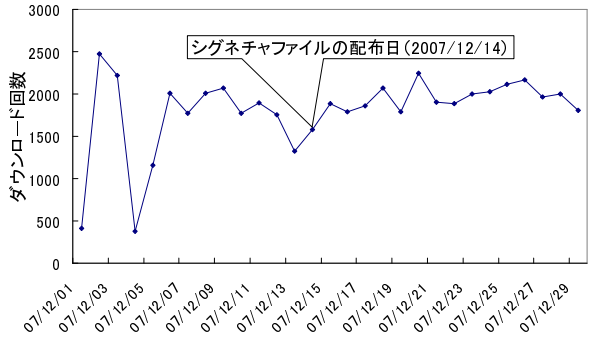

図 4　ダウンロード回数の変化がほぼ一定のマルウェア

図 3 を図 5 の一種とした場合，シグネチャファイル配布前後でのダウンロードの回数に明確な変化は見られなかったことになる．しかし，ダウンロード回数の変化は，以下の 2 種類に分類できる．

(1) 図 4 のように，マルウェアの活動中はダウンロード回数の変化がほぼ一定

(2) 図 5 のように，マルウェアの活動開始日からおよそ 10〜20 日以内の間はダウンロード回数が多く，その後はダウンロード回数が急激に減少

また，マルウェアの亜種をひとまとめにしたマルウェアの種類別に，ダウンロード回数の変化に同様の傾向が見られないか調査した．しかし，調査した結果，同じ種類のマルウェアの中でも上記の (1)，(2) の 2 種類がどちらも存在していた．このため，マルウェアの種類別ではダウンロード回数の変化に明確な特徴を持たない．

さらに，同じマルウェア名でもハッシュ値ごとに特徴があるか調査した．調査結果の一部として，BKDR_VANBOT.AD についての調査結果を表 3 に示す．ハッシュ値が eb52cab976･･･であ

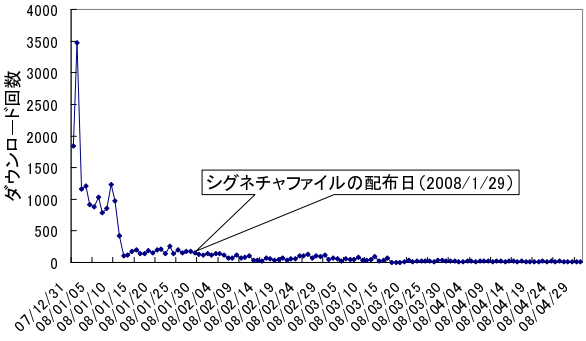

図 5　活動開始日付近のみダウンロード回数が多いマルウェア

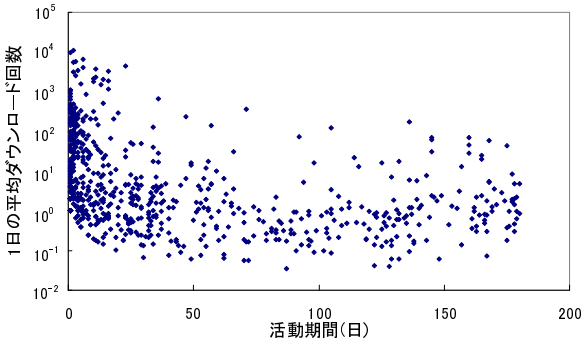

図 6　1 日の平均ダウンロード回数と活動期間の関係

表 3　各ハッシュ値の 1 日のダウンロード回数

マルウェア名	マルウェアのハッシュ値	取得日	ダウンロード回数
UNKNOWN	eb52cab976 …	2008/1/10	20
UNKNOWN	eb52cab976 …	2008/1/11	9
BKDR_VANBOT.AD	eb52cab976 …	2008/1/16	2
BKDR_VANBOT.AD	eb52cab976 …	2008/1/18	4
BKDR_VANBOT.AD	eb52cab976 …	2008/1/19	7
BKDR_VANBOT.AD	eb52cab976 …	2008/1/21	2
UNKNOWN	9d6c66bf09 …	2008/1/22	967
BKDR_VANBOT.AD	eb52cab976 …	2008/1/22	2
UNKNOWN	9d6c66bf09 …	2008/1/23	1985
BKDR_VANBOT.AD	eb52cab976 …	2008/1/23	2
BKDR_VANBOT.AD	9d6c66bf09 …	2008/1/24	2043

る BKDR_VANBOT.AD は，2008/1/11 以降，マルウェア名が UNKNOWN から BKDR_VANBOT.AD に変化している．このことは，2008/1/11 から 2008/1/15 の間にシグネチャファイルの配布が行われたことを示している．また，2008/1/22 には，ハッシュ値が 9d6c66bf09 … である BKDR_VANBOT.AD が新たにダウンロードされている．このことから，シグネチャファイルが配布された数日後に，マルウェアの更新が行われたと考えられる．2008/1/22 におけるマルウェアの更新前後の 1 日のダウンロード回数は，更新前は 2 回，更新後は 967 回である．このため，マルウェアの更新前後ではダウンロード回数に明確な変化がある．

このことから，シグネチャファイル配布後も，マルウェアの更新に対する注意が必要であることがわかる．

3.4　マルウェアの 1 日の平均ダウンロード回数と活動期間との関係

調査期間をまたいで活動しているマルウェアは，マルウェアの 1 日の平均ダウンロード回数と活動期間の正確な関係が取れない．このため，本調査では，調査開始日である 2007 年 11 月 1 日，および調査終了日である 2008 年 4 月 30 日にダウンロードされたマルウェアを調査対象から除外した．調査対象となるマルウェアは，780 個である．各マルウェアの 1 日の平均ダウンロード回数と活動期間との関係を図 6 に示す．

図 6 より，マルウェアの活動期間は，20 日以下に全体の約 56 %が集中している．また，活動期間が長くなると，1 日の平均ダウンロード回数の少ないものが比較的多くなる．さらに，1 日の平均ダウンロード回数が 10^3 以上のマルウェアの活動期間は，最長で 23 日であり，活動期間が長いものは存在しない．

これらのことより，1 日の平均ダウンロード回数が多いマルウェアほど，早期にマルウェアの解析を行い，シグネチャファイルの配布などの対策が必要となると考えられる．また，1 日の平均ダウンロード回数が少ないマルウェアでも，早期にダウンロードホストを停止させるなどのマルウェア拡散への対策が有効になると考えられる．

3.5　ダウンロードホスト IP アドレスとマルウェア名の関係

ダウンロードホスト IP アドレスとマルウェア名の関係を，図 7 に示す．図 7 は，各ダウンロードホストからのマルウェアのダウンロード回数が 500 回以上のものを示している．

図 7 より，ダウンロードホストは，多種のマルウェアがダウンロードされるダウンロードホスト，およびダウンロードされたマルウェアが 1，2 種類しかないダウンロードホストの 2 種類

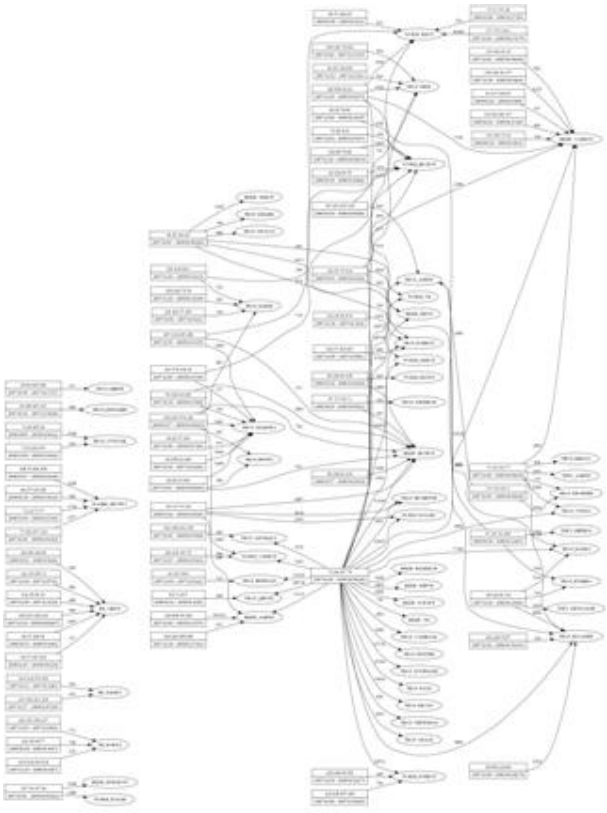

図 7　ダウンロードホスト IP アドレスとマルウェア名の関連

に分類できる．

多種のマルウェアがダウンロードされるダウンロードホストについては，活動期間が比較的長い．特に，72.AAA.AAA.AAA（活動期間：182 日）からダウンロードされたマルウェアは種類，ダウンロード回数ともに多い．一方，ダウンロードされたマルウェアが 1，2 種類しかないダウンロードホストについては，活動期間が比較的短い．例として，220.DDD.DDD.DDD（活動期間：57 日），125.EEE.EEE.EEE（活動期間：4 日）が挙げられる．

また，マルウェアについては，多数のダウンロードホストからダウンロードされるマルウェア，および 1，2 個のダウンロードホストからしかダウンロードされないマルウェアに分類できる．

多数のダウンロードホストからダウンロードされるマルウェアが利用するホストの活動期間に着目すると，長期間（3ヶ月以上）活動しているダウンロードホストが含まれている．例として，PE_VIRUT の 218.FFF.FFF.FFF（活動期間：107 日），および WORM_IRCBOT の 124.GGG.GGG.GGG（活動期間：133 日）が挙げられる．

これらのことより，活動期間の長いマルウェアのダウンロードホストを特定することで，多種のマルウェア感染の拡大を防止できると考えられる．また，特定のマルウェアが長期間ダウンロードされ，感染が拡大することも防止できると考えられる．

4　おわりに

攻撃元データを基にダウンロードホストに着目してマルウェアの特徴を分析した結果を報告した．

調査結果から，ダウンロードホスト，およびマルウェアには以下の特徴があるといえる．ダウンロードホストの活動期間は，長期間，短期間，および複数期間の 3 種類に分類可能である．ダウンロードホストからのダウンロード回数とダウンロードされるマルウェアの種類数の関連性は薄い．シグネチャファイル配布前後でダウンロード回数の変化は小さい．しかし，マルウェアの更新前後ではダウンロード回数が変化する．1 日の平均ダウンロード回数が多いマルウェアの活動期間は短い傾向にある．ダウンロードホストの活動期間が長いものほど，ダウンロードされるマルウェアの種類数が多い傾向にある．

謝辞　本研究の一部は，中島記念国際交流財団 日本人若手研究者研究助成の支援を受けて行った．

参考文献

[1] トレンドマイクロ ウイルス感染被害レポート - 2007 年度（最終版），http://jp.trendmicro.com/jp/threat/security_news/monthlyreport/article/20080108011916.html

[2] サイバークリーンセンター (CCC) — ボット ウイルスとは，https://www.ccc.go.jp/bot/index.html

[3] 藤原将志，寺田真敏，安部哲哉，菊池浩明，“ マルウェアの感染動作に基づく分類に関する検討，”情報処理学会研究報告，Vol.2008，No.21，2008-CSEC-40，pp.177–182，March 2008.

[4] 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，pp.2512–2523，August 2006.