ファイアウォールログを利用したマルウェア活動の検出手法について

加藤 淳也† 門田 剛† 畑田 充弘‡ 竹内 文孝†
†NTTコミュニケーションズ株式会社 セキュリティオペレーションセンタ
‡NTTコミュニケーションズ株式会社 先端IPアーキテクチャセンタ
E-mail: {jyunya.kato, tsuyoshi.kadota, m.hatada, f.takeuchi}@ntt.com

あらまし 近年，マルウェアの感染拡大活動やDoS攻撃などが，ネットワークシステムの可用性に重大な影響を及ぼしたり，重要なサーバやクライアントのシステムダウンを引き起こすなど，企業等の事業継続性を脅かしている．本稿にて，我々はセキュリティ運用業務の実務面に重点を置き，セキュリティ対策製品において普及率の高いファイアウォールのログを利用することで，DoS攻撃・感染拡大活動などのマルウェア活動を検出させる方法を調査した．その結果，ファイアウォールのログにおける特定の条件に適合するパターンを監視することが，マルウェア活動を特定する上で有用であることが分かった．

Detection method of malware activity based on the Firewall log

Junya Kato† Tsuyoshi Kadota† Mitsuhiro Hatada‡ Fumitaka Takeuchi†
†Security Operation Center, NTT Communications Corporation
‡Innovative IP Architecture Center, NTT Communications Corporation
E-mail: {jyunya.kato, tsuyoshi.kadota, m.hatada, f.takeuchi}@ntt.com

Abstract Today, the wide-scale malware spread and DoS attacks often affect the availability of network systems and cause the system overflow, and threat the continuity of companies. In this research, we have investigated the method to detect the malware activity like DoS attacks and infection spread, by focusing on the practical aspects of the security operations and using the firewall log which has the high diffusion rate in the security products. As a result, we have found that it is helpful to monitor the patterns which comply with the specific condition of the firewall log to identify the malware activity.

1. はじめに

マルウェア対策としてIDS（侵入検知システムやIPS（侵入防止システム），WAF（Webアプリケーションファイアウォール）等のセキュリティ製品の導入が進んでいるが，これらの高価なセキュリティ製品を導入できる企業は限られている．我々はセキュリティ運用業務の実務面に重点を置き，一般的に世の中に普及しているファイアウォール(以降，FW)のログを利用してマルウェア活動を検出する方法に着目する．本稿では，まずFWのログとして分析可能な項目を挙げ，そこから得られる特徴を分析するとともに，MWS2008[1)]で得られた成果からマルウェアの検出に有効と考えられる特徴を挙げる．それぞれの特徴に対して，CCC DATAset 2009 の攻撃通信データ[2)](以降，CCC2009 攻撃通信データ)を用いた評価実験を行った結果を報告する．

2. 分析手法

2.1. FW ログ項目による分析

一般的に，感染した端末は，感染拡大，マスメーリング活動，DoS 攻撃のマルウェア活動を行うため, 効率良くマルウェアを特定する上でハニーポット発だけの通信を抽出した.

具体的には，CCC2009 攻撃通信データのパケットキャプチャデータに対して図 1に示す手順で処理を行い, ハニーポット発の通信だけを記録したFWログを生成した．生成したFWログは, 内訳としてWindows 2000 のハニーポット(以降， honey004)で 1,495,050 行，Windows XPのハニーポット(以降, honey003)で 1,337,099 行であった.

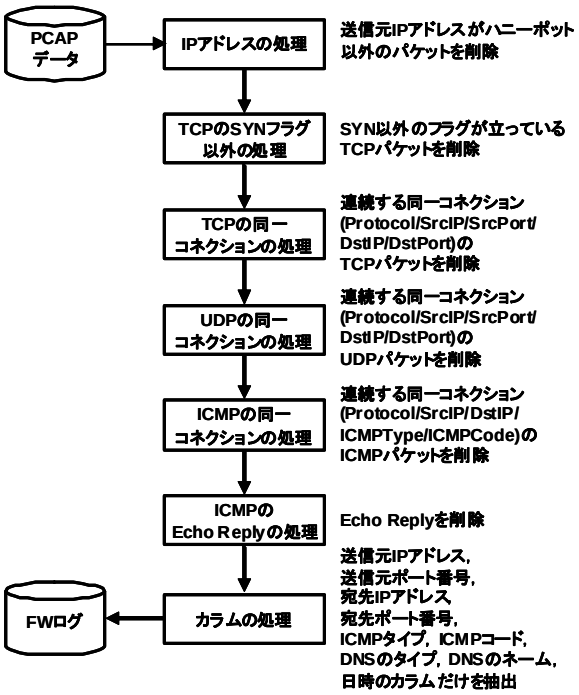

図 1　FW ログ作成ステップの流れ

厳密には, FW ログを生成する上でセッションタイムアウト内の同一コネクションを削除する必要があるが, 検知方法として利用しない通信にしか影響を及ぼさなかったので, 今回は処理を行っていない.

2.1.1. マルウェア活動の特徴

FW ログを解析した結果，マルウェア活動として，感染拡大，マスメーリング活動，DoS 攻撃，大量の DNS Query が見受けられた.

感染拡大は，135/TCP，139/TCP，445/TCP の 3 種類の宛先ポート番号と ICMP Request で多数の宛先 IP アドレスに対して行われていた．DoS 攻撃は，80/TCP の宛先ポート番号で3つの宛先IPアドレスに対して行われていた．DoS 攻撃の種類としては，コネクションの確立とコネクションの終了を繰り返し行う「TCP Connection flood」型であった．マスメーリング活動は，25/TCP の宛先ポート番号で AOL と Microsoft 社の商用メールサーバに接続を試みていた．大量の DNS Query は，53/UDP の宛先ポート番号でローカル DNS サーバ(以降, 正規の DNS サーバ)に対して行われていた.

感染拡大，大量のDNS Query，DoS攻撃のマルウェア活動におけるFWログ数と時系列の関係を一例として図 2，図 3，図 4に示す．Y軸は条件に一致するFWログ数の累積を表し，X軸は二日間の時間を表している．また，データのプロットは，秒単位で行った.

大量のDNS Queryは，感染拡大と同時に送信されていることが分かる. またそれぞれの活動に応じて特徴的なFWログの出力をしていることが分かる. これらのマルウェア活動の特徴を，表 1にまとめた．ハニーポットのOSが異なると，同じ宛先ポート番号(135/TCP)における感染拡大のFWログ数／秒が大きく異なることが分かった. 同じハッシュ値のマルウェアが異なるOSのハニーポットに感染した場合においても，同様に異なっていた.

また，感染拡大の宛先ポート番号が異なる(135/TCP と 139/TCP，135/TCP と 445/TCP)と, FW ログ数／秒が大きくことなることが分かった．これは，WORM_ALLAPLE.IK が 139/TCP と 445/TCP の感染拡大を，WORM_SWTYMLAI.CD が 135/TCP の感染拡大を行っており, 感染活動を行うマルウェアが異なっていたからである.

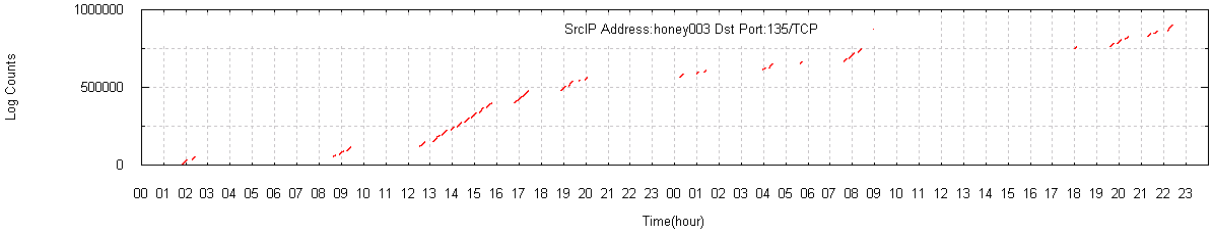

図 2　送信元 IP アドレスが honey003，宛先ポート番号が 135/TCP であるログ数(感染拡大)

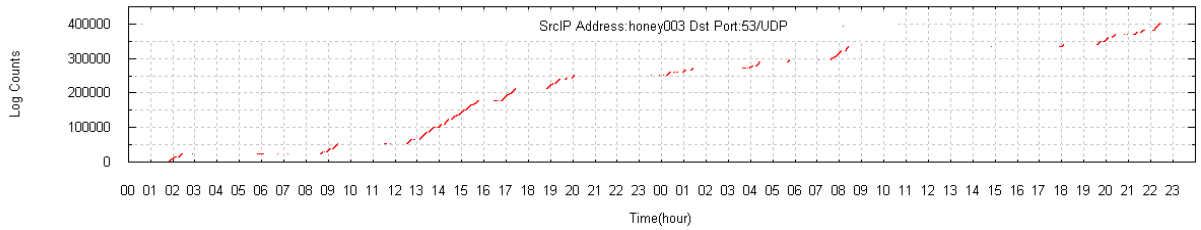

図 3　送信元 IP アドレスが honey003, 宛先ポート番号が 53/UDP であるログ数(大量の DNS Query)

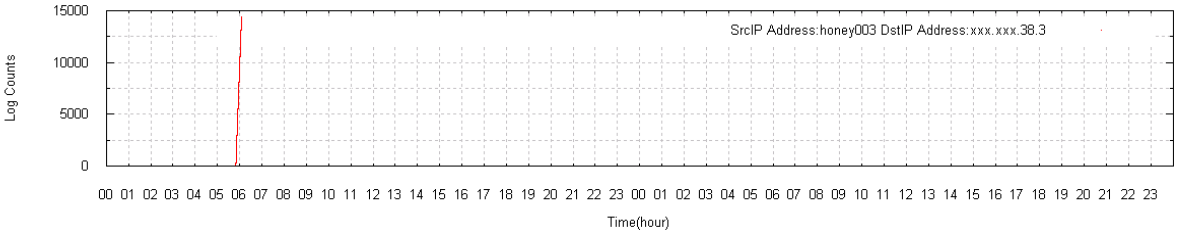

図 4　送信元 IP アドレスが honey003，宛先 IP アドレスが xxx.xxx.38.3 であるログ数(DoS 攻撃)

表 1　マルウェア活動の特徴

ハニーポット	活動内容	宛先ポート/ICMP タイプ	活動回数	(FW ログ数/秒)			標準偏差
				最小	最大	平均	
honey004 Windows 2000	感染拡大	135/TCP	27	117.13	133.04	124.57	2.31
		139/TCP	1	8.21	8.21	8.21	0.00
		445/TCP	1	8.18	8.18	8.18	0.00
		ICMP Echo Request	1	16.58	16.58	16.58	0.00
	マスメーリング活動	25/TCP	4	1.33	2.00	1.50	0.29
	DoS 攻撃	–	0	–	–	–	–
	大量の DNS Query	53/UDP	0	–	–	–	–
honey003 Windows XP	感染拡大	135/TCP	36	24.11	44.48	38.93	3.92
		139/TCP	0	–	–	–	–
		445/TCP	0	–	–	–	–
		ICMP Echo Request	0	–	–	–	–
	マスメーリング活動	25/TCP	4	1.33	2.00	1.50	0.29
	DoS 攻撃	80/TCP	3	7.29	15.52	10.66	3.52
	大量の DNS Query	53/UDP	36	9.19	21.44	17.28	2.46

2.1.2. FW ログ数による検知方法

これらのマルウェア活動の特徴から最小の FW ログ数／秒の値と以下の特定の条件に適合するパターンを利用して, マルウェア活動を検知する．感染拡大に関しては，同一の送信元 IP アドレスから特定の宛先ポート番号(135/TCP，139/TCP，445/TCP)及び ICMP タイプの FW ログ数を監視する．DoS 攻撃に関しては，同一の送信元 IP アドレスから同一の宛先 IP アドレスの FW ログ数を監視する．大量の DNS Query に関しては, 同一の送信元 IP アドレスから特定の宛先ポート番号(53/UDP)の FW ログ数を監視する．なお，最小の値を利用するのは, マルウェア活動にもかかわらず, その活動を見逃してしまう可能性をなくすためである．マスメーリング活動に関しては，FW ログ数／秒の値が小さいため正常なトラフィックと区別がつきにくく, FW ログ数による検知方法はあまり有効でないと考える．

2.2. MWS2008 成果からの分析

東角ら[3)]は，マルウェアがリゾルバの逸脱や異常なDNS Queryを送信することを，竹森ら[4)]は, マルウェアがhostsファイルの一部をローカル・ループバックアドレス(127.0.0.1)に改竄することを，また畑田ら[5)]は，マルウェアの通信がブラックリストに引っかかることを報告している．これらの成果から，以下の方法による検知方法が考えられる．

2.2.1. リゾルバの逸脱による検知方法

正規の DNS サーバ以外に対する DNS Query があるかどうかでマルウェア活動を検知する．

2.2.2. DNS Query のリソースレコードのタイプとネームによる検知方法

DNS QueryのAレコードのネームに表 2のブラックリストのFQDN/IPアドレスの値が含まれているかどうかでマルウェア活動を検知する. ブラックリストはすべて 2009/6/20 版を利用した．

表 2　ブラックリスト一覧

ブラックリスト名	ユニークな FQDN	ユニークな IP アドレス
config-hosts[3)]	266,789	0
MalwareBlockList[5)]	793	573
PhishTank[5)]	860	561

DNS Query の A レコードのネームに ISP の動的 IP アドレスに割り当てられると思われる FQDN が含まれているか, あるいは IP アドレスが含まれているかどうかでマルウェア活動を検知する．

DNS Query の MX レコード，PTR レコードがあるかどうかでマルウェア活動を検知する．

3. 評価実験

3.1. FW ログ数による検知方法

マルウェア活動の特徴に基づいた FW ログ数による検知方法の有効性を確認するために評価システムを実装した．本システムは S 秒間隔で FW ログを読み込み，予め設定したプロトコル，IP アドレス，ポート番号，ICMP タイプに一致した FW ログが N 行以上ある場合にアラートを上げるシステムである. 運用面でのシステム負荷を考慮し，60 秒間隔で処理を行うようにした．(S=60[秒]，N=FW ログ[数／秒]×60[秒])

実験結果を，表 3に示す．

表 3　FW ログ数による検知方法の実験結果

ハニーポット	活動内容	宛先ポート /ICMP タイプ	活動回数	正しい検知回数	検知見逃し回数	誤検知回数
honey004 Windows 2000	感染拡大	135/TCP	27	26	1	0
		139/TCP	1	1	0	0
		445/TCP	1	1	0	0
		ICMP Echo Request	1	1	0	0
honey003 Windows XP	感染拡大	135/TCP	36	36	0	0
	DoS 攻撃	80/TCP	3	3	0	0
	大量の DNS Query	53/UDP	36	36	0	0

135/TCP の感染拡大において検知見逃しが 1 回生じたが, その他のマルウェア活動は誤検知もなく正しく検知することができた. 検知見逃しが生じた理由としては, 感染拡大の活動時間が 25 秒と短かったため，一致した FW ログ数が少なく, N 行以上にならなかったためである. 処理間隔と正しい検知の間にはトレードオフが生じるため, マルウェアの活動時間とシステム負荷に応じた適切な処理間隔 S 値を決定する必要があるであろう．

3.2. リゾルバの逸脱による検知方法

ハニーポットからDNSサーバに送られたDNS QueryのFWログ数を宛先DNSサーバ別に表 4に示す．

表 4　宛先 DNS サーバ別の FW ログ数

ハニーポット	宛先 DNS サーバ	FW ログ数
honey004 Windows2000	正規の DNS サーバ	438
	その他の DNS サーバ	0
honey003 WindowsXP	正規の DNS サーバ	401,228
	その他の DNS サーバ	0

ハニーポットからDNSサーバに送られたDNS Queryは全て正規のDNSサーバに対してであった．CCC2009 攻撃通信データにおいてリゾルバの逸脱は見受けられなかった. しかし, CCC2008 攻撃通信データにおける東角ら[3)]の研究では, リゾルバを逸脱したマルウェアを確認していることから，リゾルバを逸脱したDNS Queryを監視することはマルウェア活動を特定する上で有用だと考える.

3.3. DNS Query のリソースレコードのタイプとネームによる検知方法

ハニーポットからDNSサーバに送られたDNS QueryのFWログ数をタイプ別に表 5に示す.

表 5　タイプ別の FW ログ数

ハニーポット	DNS Query のタイプ	FW ログ数
honey004 Windows2000	A レコード	438
	PTR レコード	0
	MX レコード	0
	その他のレコード	0
honey003 WindowsXP	A レコード	401,225
	PTR レコード	3
	MX レコード	0
	その他のレコード	0

大抵の DNS Query は A レコードであった. honey003 では, DNS Query の PTR レコードのネームに，ローカル・ループバック・アドレスのドメイン名(1.0.0.127.in-addr.arpa)を格納して送信していた．この DNS Query は, hosts ファイルの改竄による可能性が考えられる．PTR レコードの利用方法としては異常な利用方法であるため，DNS Query の PTR レコードのネームにローカル・ループバック・アドレスのドメイン名が格納されているかを監視することは, 検知方法として有用だと考える.

CCC2009 攻撃通信データにおいてDNS QueryのMXレコードは存在しなかった．しかし, 武藏ら[6)]の研究では, MXレコードのDNS Queryを送るマルウェアを確認していることから，正規のメールサーバ以外の端末がMXレコードのDNS Queryを送信しているかを監視することは同様に有用であろう.

次に，Aレコードのネームを種類別に表 6に示す.

表 6　種類別の FW ログ数

ハニーポット	種類	FW ログ数
honey004 Windows2000	config-hosts	219
	MalwareBlockList	30
	PhishTank	0
	ISP の FQDN	28
	IP アドレス	0
honey003 WindowsXP	config-hosts	250
	MalwareBlockList	35
	PhishTank	96
	ISP の FQDN	37
	IP アドレス	400,922

honey003 に着目すると，99%以上の A レコードのネームは IP アドレスであった．A レコードの利用方法としては異常な利用方法であるため, DNS Query の A レコードのネームに IP アドレスが格納されているかを監視することは同様に有用である.

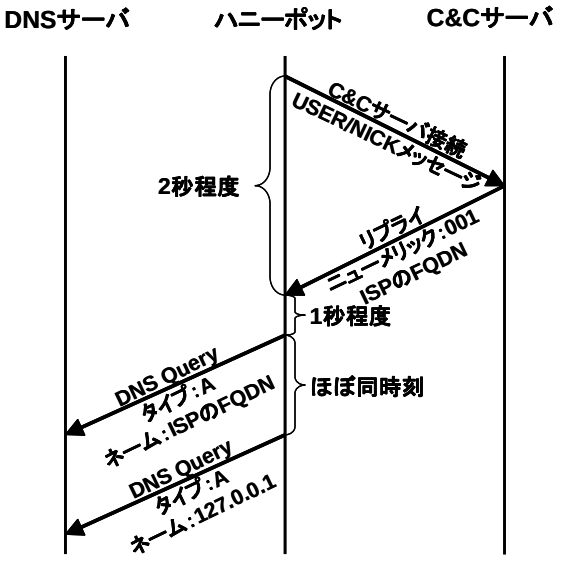

図 5　DNS Query(ISP FQDN) 発生メカニズム

また，AレコードのネームにISPのFQDNが含まれていた. この発生メカニズムを図 5に示す．これはハニーポットがC&Cサーバに IRC接続した際に，C&Cサーバがボットのグローバルアドレスの逆引き結果(ISPのFQDN)をボットに返していた．その後にボットがDNS QueryのAレコードのネームにそのISPのFQDNを格納して送信していた．このDNS Queryと同時刻(1 秒以内)に，ハニーポットがDNS QueryのAレコードのネームにローカル・ループバック・アドレス(127.0.0.1)を格納して送信していた(WindowsXPのみ).

よって，DNS Query の A レコードのネームに自身のグローバル IP アドレスの FQDN が格納されているか，ローカル・ループバック・アドレスが格納されているかを監視することは同様に有用であろう. ブラックリストにおいては，config-hosts が最も検知率が高いが，他のブラックリストと比較してリストに登録している数が多いため, 実環境におけるシステム負荷は高いであろう.

4. 今後の課題

本手法の検知精度を向上させるためには，以下の課題がある.

OS・マルウェアの種類の網羅性

実環境では様々なOSのクライアント･サーバがFW配下に存在するため，今回利用したハニーポットのOSとは別のOSやサービスパックでの調査が必要である．特に，WindowsXP SP2 以降ではセキュリティ機能の強化によりTCPハーフコネクション数に制限がかけられている．よって，表 1に示したマルウェア活動の特徴が異なることが予測される．また，ハニーポットに感染したマルウェアの種類が少なかったことから, 様々なマルウェアに感染させて同様に調査する必要がある.

誤検知の評価

CCC2009 攻撃通信データにはユーザが端末を操作して業務を行う等の正常なトラフィックが流れていない. よって，業務等の正常なトラフィックが流れる実環境においてこれらの検知手法を適用し誤検知が発生しないか評価する必要がある.

5. おわりに

FW ログからマルウェア活動を検出する方法を調査した. マルウェア活動の特徴と特定の条件に適合するパターンを利用して, FW ログから感染拡大，DoS 攻撃，大量の DNS Query を行うマルウェアを検知できることが分かった．また，リゾルバの逸脱や DNS Query のリソースレコードのタイプとネームによる検知もマルウェア活動を検知する上で有効であることが分かった. 今後は，本手法の検知精度を向上させるための課題に取り組む予定である.

謝辞

本研究の遂行にあたって，有益な研究用データセットを提供して頂いたサイバークリーンセンターに感謝する.

参考文献

1) マルウェア対策研究人材育成ワークショップ, http://www.iwsec.org/mws/2008/

2) 畑田充弘，中津留勇，寺田真敏,篠田陽一：マルウェア対策のための研究用データセットとワークショップを通じた研究成果の共有, MWS2009

3) 東角芳樹，鳥居悟：DNS 通信の挙動からみたボット感染検知方式の検討，情報処理学会・CCC 運営委員会, MWS2008, pp.13-18, 2008

4) 竹森敬祐，磯原隆将，三宅優，西垣正勝：ボットネットおよびボットコードセットの耐性解析，情報処理学会・CCC 運営委員会, MWS2008, pp.49-54, 2008

5) 畑田充弘，寺田真敏：複数観測データを用いたボットネットの活動分析に関する一考察，情報処理学会・CCC 運営委員会, MWS2008, pp.87-92, 2008

6) 武藏泰雄，松葉龍一，杉谷賢一：プロトコル異常検知による A レコード型 DNS パケット分散サービス妨害攻撃の阻止，情報処理学会研究報告, IPSJ SIG Notes 2005-DSM-38-(5)