情報処理学会論文誌 Vol.57 No.9 1986–2002 (Sep. 2016)

推薦論文

通信プロトコルのヘッダの特徴に基づく不正通信の検知手法

小出 駿[1,†1,a)] 鈴木 将吾[1] 牧田 大佑[1,2] 村上 洸介[3] 笠間 貴弘[2] 鈴木 未央[2] 島村 隼平[4]
衛藤 将史[2] 井上 大介[2] 中尾 康二[2] 吉岡 克成[5] 松本 勉[5]

受付日 2015年12月3日, 採録日 2016年6月2日

概要：独自のネットワークを実装したマルウェアやツールが生成したパケットは，ヘッダに固有の特徴を持つことがある．本論文では，TCP ヘッダのシーケンス番号，IP ヘッダの ID 値，および DNS のヘッダの ID 値等のフィールドの値を組み合わせたシグネチャを作成し，単一のパケットから送信元のマルウェアを特定する手法を提案する．マクロ解析とミクロ解析の相関分析により提案手法の有効性を示し，不正通信の分析事例を報告する．

キーワード：不正通信検知，ネットワークスタック，ダークネット観測

Detection Method for Malicious Packets with Characteristic Network Protocol Header

Takashi Koide[1,†1,a)] Shogo Suzuki[1] Daisuke Makita[1,2] Kosuke Murakami[3]
Takahiro Kasama[2] Mio Suzuki[2] Jumpei Shimamura[4] Masashi Eto[2] Daisuke Inoue[2]
Koji Nakao[2] Katsunari Yoshioka[5] Tsutomu Matsumoto[5]

Received: December 3, 2015, Accepted: June 2, 2016

Abstract: The packets from malware and network tools that have their own implementation of network stack may have characteristic packet headers. In this paper, we propose a technique for packet detection by generating signatures using sequence number in the TCP header, ID in the IP header, ID in the DNS header, and so on. By comparing the correlation between macro- and micro-analysis, we confirm the effectiveness of our technique and report the analysis case of malicious packets.

Keywords: packet detection, network stack, darknet monitoring

[1] 横浜国立大学
Yokohama National University, Yokohama, Kanagawa 240–8501, Japan
[2] 情報通信研究機構
National Institute of Information and Communications Technology, Koganei, Tokyo 184–8795, Japan
[3] KDDI 株式会社
KDDI Corporation, Shinjuku, Tokyo 163–8003, Japan
[4] 株式会社クルウィット
clwit Inc., Shinagawa, Tokyo 141–0031, Japan
[5] 横浜国立大学大学院環境情報研究院/横浜国立大学先端科学高等研究院
Graduate School of Environment and Information Sciences/Institute of Advanced Sciences, Yokohama National University, Yokohama, Kanagawa 240–8501, Japan
[†1] 現在，NTT セキュアプラットフォーム研究所
Presently with NTT Secure Platform Laboratories, NTT Corporation
[a)] koide-takashi-mx@ynu.jp

1. はじめに

多くの OS はネットワークを介したデータの送受信を行うための API として，ソケットを提供している．ソケットを使用して送信された通信パケットは，そのヘッダ内に固有の特徴が見られることがあるため，送信元の OS を特定することが可能である．このような技術を OS Fingerprinting と呼び，その技術を実装した代表的なソフトウェアである p0f [1] は TCP パケットを受動的に観測し，シグネチャと照合することで送信元の OS を判定する．一方，DDoS 攻

本論文の内容は 2014 年 10 月のコンピュータセキュリティシンポジウム 2014 にて報告され，コンピュータセキュリティ研究会主査により情報処理学会論文誌ジャーナルへの掲載が推薦された論文である．

撃（Distributed Denial of Service attack）やネットワークスキャンを行うマルウェアは，高速かつ効率的にパケットを送信するため，OS の通信機能を使用せずに，独自に実装されたネットワークスタックを使用して通信を行うことが多い．それらのマルウェアから発生したパケットは，OS に固有の特徴を持たない代わりにマルウェアに固有の特徴を持つ可能性がある．

独自のネットワークスタックを実装したマルウェアから発生したパケットを分析する研究として，文献 [2] は，OS の動作モードが最高権限の Ring0（カーネルモード）で実装されたフルカーネルマルウェア（FKM）について，p0f で OS を判定できないパケットは FKM から発生したパケットであると推測している．さらに，p0f で生成されたシグネチャを使用して FKM の可能性が高い TCP パケットを検知している．また，文献 [3] は，OS が生成したパケットと独自のネットワークスタックを実装したマルウェアが生成したパケットを区別するために，単体のパケットのヘッダから得られる情報と連続したパケットのヘッダの変化の情報をもとに分類する手法を提案している．また，文献 [4] は IP ヘッダの TTL（Time to Live）値に着目し，異常に大きいホップ数の TTL を持つパケットは悪性である可能性が高いと推測している．しかし，これらの研究は独自のネットワークスタックを実装したマルウェアから発生したパケットのヘッダの特徴を用いて，パケットの悪性判定をするにとどまっている．

我々は，マルウェア動的・静的解析およびツールのソースコードの分析によるミクロ解析と，ダークネットやハニーポットの観測結果の分析によるマクロ解析を行い，特定のマルウェアやツールは，パケット生成の際に，特徴的なヘッダフィールドの設定方法を実装していることを確認した．その設定方法は 2 種類に大別でき，1 つは，パケットの特定のヘッダフィールドにマルウェアやツールに固有の値を設定することであり，もう 1 つは，複数のヘッダフィールドの値を用いて計算を行い，計算結果を別のフィールドに設定することである．これらの設定方法に着目することで，パケットのヘッダの特徴と，送信元のマルウェアやツールを紐付けることが可能であると考えられる．

そこで本論文では，パケットヘッダの各フィールドの値を組み合わせたシグネチャを用いて，観測したパケットに対してパターンマッチングを行い，独自のネットワークスタックを実装したマルウェアとツールを特定する手法を提案する．文献 [2], [4] が分析に使用した p0f は，ヘッダフィールドとシグネチャのフィールドを 1 対 1 に照合して OS の判定を行っている．これと同様の方法でパケットの照合を行い，さらに p0f が判定に扱っていないヘッダフィールドにまで分析対象を拡張することで，前者の設定方法によるパケットの特定を行う．また，後者の設定方法によるパケットは，既存手法の 1 対 1 の照合では特定が不可能である．そこで，ヘッダフィールドの設定に用いた計算と同様の手順による検算を行うことでパケットの特定を行う．本論文で定義するシグネチャの構成と作成方法に基づき，マクロ解析とミクロ解析による特徴抽出を行った結果，12 種のマルウェアとツールから 19 個のシグネチャを作成した．

本論文では，提案手法によって独自のネットワークスタックを実装したマルウェアやツールの特徴を抽出し，その特徴を用いてパケットから送信元を特定可能であることを示すために，ミクロ解析とマクロ解析の結果の相関分析による 2 つの検証実験を行う．マルウェア Morto とネットワークスキャンツール ZMap [5] について，ミクロ解析として Morto のマルウェア動的解析と ZMap のソースコード分析の結果を使用し，マクロ解析としてそれぞれの出現時期のダークネット観測の結果を使用して相関分析を行う．その結果，ミクロ解析で確認したパケットの特徴と，マクロ解析で得た Morto と ZMap から送信された可能性の高いそれぞれのパケットの持つ特徴が一致していることを確認した．以上の結果から，本手法が独自のネットワーク実装から送信されたパケットの分類と特定に有効であることを示した．

次に，提案手法を用いた実際のインターネット上の不正通信の分析事例を報告する．本論文で作成したシグネチャを用いて，我々が運用しているハニーポットの観測データを分析する．まず，DRDoS 攻撃（Distributed Reflection Denial of Service attack）の機能を持つマルウェアによる通信を分析した結果，本論文で分析したマルウェアは，攻撃者によって実際の DRDoS 攻撃に使用されていないと判断した．次に，組み込み Linux 機器を狙ったマルウェアによる通信を分析した結果，送信元のマルウェアは不明であるが，パケットと実際の攻撃・侵入方法を結び付けることができた．

さらに，提案手法を用いた継続的なネットワーク観測・分析システムの実装を行い，本システムで得られた観測データの分析結果について報告する．国立研究開発法人情報通信研究機構（以下，NICT）の保有するサイバー攻撃観測・分析・対策システム NICTER（Network Incident analysis Center for Tactical Emergency Response）[6] の相関分析システムに，本手法の検知アルゴリズムと本論文で作成したシグネチャを導入する．約 27 万（2015 年 11 月現在）の IP アドレスで構成されるダークネットで観測したパケットに対してパターンマッチングを行う仕組みを構築することで，マルウェアやツールによって送信された大量のネットワークスキャンパケットをリアルタイムに検知する仕組みを構築した．さらに，本システムで得られた観測結果を用いて，DRDoS 攻撃に関係するネットワークスキャンパケットを分析する．その結果，DRDoS 攻撃に悪用される踏み台のサーバ（リフレクタ）を探索するための通信は，

ZMap から発生したと思われるパケットが多くを占めていることが判明した．

2. マクロ解析とミクロ解析によるパケットヘッダの特徴抽出

本章では，独自のネットワークスタックを持つマルウェアやツールによって送信されたパケットは，ヘッダに固有の特徴を持つことを例示するために，マルウェア動的解析，静的解析（リバース・エンジニアリング）およびオープンソース・ソフトウェア（以下，オープンソース）のソースコードの分析によるミクロ解析と，ダークネットの観測結果の分析によるマクロ解析を行い，観測されたパケットのヘッダを分析する．分析に使用したダークネットの観測データは，NICTER のサイバーセキュリティ情報遠隔分析基盤 NONSTOP（NICTER Open Network Security Test-out Platform）[9] が提供している/16 ネットワーク（65536 IP）で構成されるダークネットの観測結果である．また，以降のダークネット分析も同一のダークネットによる観測データを使用する．

2.1 マルウェア Morto

マルウェア Morto のミクロ解析として動的解析を行い，発生したパケットを分析する．Morto は，RDP（Remote Desktop Protocol）を利用して Windows の端末やサーバに感染活動を行うワームである．Morto に対する注意喚起は Microsoft，F-Secure 等によって 2011 年 8 月後半に行われている [10], [11]，このマルウェアは，ローカルネットワークとインターネットに向けてネットワークスキャンを行い，RDP ポート（3389/TCP）を待ち受けているホストに対して，「admin/admin」といった安易なユーザ名とパスワードによる辞書攻撃を行い，侵入を試みることが知られている．Morto 検体のマルウェア動的解析を行い，通信を観測する．解析を行ったマルウェア動的解析環境は，文献 [12] と同様である．また，実際の攻撃や侵入を防ぐため，SYN パケットの送信のみ許可するアクセス制御を行っている．

- 検体ハッシュ値：0475c97ddb96252febff864fb778b460（MD5）
- 解析時間：2012 年 8 月 26 日 10:52〜16:54（6 時間）
- 実行環境：Windows XP SP2

解析の結果，多数の 3389/TCP 宛の SYN パケットを観測した．それらのパケットの宛先 IP アドレスに着目すると，実行環境の周辺のローカル IP アドレスに対するネットワークスキャンと，グローバル IP アドレスに対するネットワークスキャンに通信を分類可能であることが分かった．それぞれの宛先に対する 3389/TCP 宛 SYN パケットの 1 分あたりのパケット数を**図 1** に示す．ローカル IP アドレスに向けたパケットを分析すると，すべてのパケットが

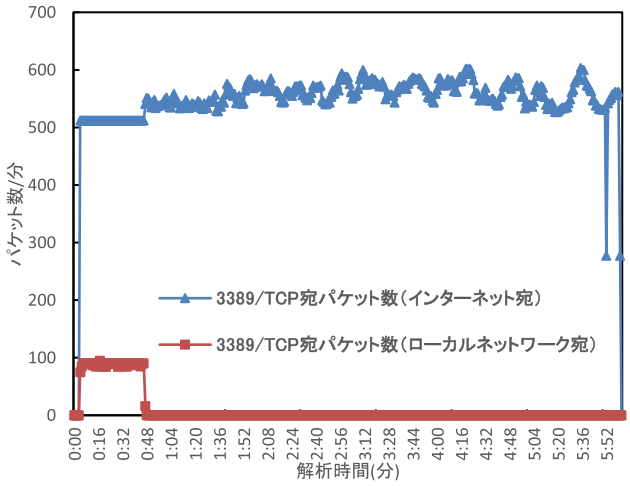

図 1 Morto から発生した 3389/TCP 宛通信
Fig. 1 Number of packets to TCP port 3389 from Morto.

Windows XP の特徴を有していた．一方，グローバル IP アドレスに向けた SYN パケットを分析すると，Windows XP の特徴を有している少数のパケットと，どの OS の特徴も有さない多数のパケットに分類することができた．後者のパケットは TCP ヘッダのシーケンス番号に 2406000322，送信元ポート番号に 4935，IP ヘッダの ID 値に 9496 がつねに設定されるといった特徴を有していた．以上の結果から，Morto は，高速にパケットを送信する必要のあるインターネットに向けたネットワークスキャンでは，独自実装のネットワークスタックによってパケット生成を行い，RDP サービスが稼働しているホストを発見すると，OS のソケットを使用してコネクションを確立し，実際の侵入を試みると推測される．

2.2 ネットワークスキャンツール ZMap

ネットワークスキャンツール ZMap のミクロ解析としてソースコードを分析する．ZMap はミシガン大学が開発し，2013 年 8 月に公開されたオープンソースの高速ネットワークスキャンツールである．開発元によると，実行環境によっては IPv4 の全アドレス空間を 5 分でスキャン可能と述べられている．このツールのソースコードを分析した結果，生成された TCP SYN・UDP・ICMP echo request パケットは，IP ヘッダの ID 値につねに 54321 が設定される等の特徴を持つことが分かった．

2.3 ネットワークスキャンツール Masscan

ネットワークスキャンツール Masscan [13] のミクロ解析としてソースコードを分析する．Masscan は 2013 年に公開されたオープンソースの高速ネットワークスキャンツールである．開発元によると，実行環境によっては IPv4 の全アドレス空間を 6 分でスキャン可能であり，毎秒 1000 万回のパケット送信が可能であると述べている．文献 [14] は，このツールは生成する SYN パケットの IP ヘッダの

ID 値に，宛先 IP アドレス，宛先ポート番号，シーケンス番号の 3 つの値の排他的論理和を設定すると述べている．さらに，我々はソースコードの UDP パケット生成部分を分析すると DNS（53/UDP），NetBIOS（139/UDP），SNMP（161/UDP）の 3 種のパケットも同様の計算を行うことを特定した．プロトコルごとに，DNS ヘッダの ID 値，NetBIOS ヘッダの ID 値，SNMP ヘッダの ID 値に TCP パケットのシーケンス番号を置き換えることで同様の計算を行う．TCP，DNS，NetBIOS，SNMP パケットの IP ヘッダの ID 値の計算方法を式 (1)，(2)，(3)，(4) に示す．

$$\text{Ip.id} = \text{Ip.dstaddr} \oplus \text{Tcp.dstport} \oplus \text{Tcp.seq} \quad (1)$$

$$\text{Ip.id} = \text{Ip.dstaddr} \oplus \text{Udp.dstport} \oplus \text{Dns.id} \quad (2)$$

$$\text{Ip.id} = \text{Ip.dstaddr} \oplus \text{Udp.dstport} \oplus \text{Ntb.id} \quad (3)$$

$$\text{Ip.id} = \text{p.dstaddr} \oplus \text{Udp.dstport} \oplus \text{Snmp.id} \quad (4)$$

2.4 DRDoS 攻撃の機能を持ったマルウェア

DRDoS 攻撃の機能を持ったマルウェア IptabLes [15]，XOR botnet [16] のミクロ解析としてマルウェア動的・静的解析による分析について述べる．IptabLes，XOR botnet は DRDoS 攻撃の機能を実装したマルウェアであり，ボットネットを構成する．これらのマルウェアは 2014 年頃からセキュリティベンダ等によって報告されている．

2.4.1 マルウェア IptabLes

マルウェア IptabLes の検体を，文献 [12] の動的解析環境のゲスト OS を Ubuntu10.04 に替えた環境でマルウェア動的解析を行う．その結果，独自に実装されたネットワークスタックによって生成されたと思われる SYN Flood パケットと DNS Flood パケットを観測した．このうち，SYN Flood パケットは，848 byte のペイロードが付加されており，シーケンス番号に 848，IP ヘッダの ID 値に 0 が設定される等の特徴を持っていた．

- 検体ハッシュ値：b826fb1253a52a3b53afa3b7543d7694（MD5）
- 解析時間：2014 年 7 月 17 日 15:30〜16:29（25 時間）
- 実行環境：Ubuntu10.04（32 bit）

一方，観測した DNS Flood パケットは，SYN Flood パケットと同様に独自実装のネットワークスタックによって送信されたと推測されるが，観測したパケットからヘッダフィールドの設定方法を類推することはできなかった．そこで，このマルウェア検体の静的解析を行い，パケットの生成方法を分析する．その結果，DNS flood パケットは，まずランダム値（Random_value）を生成して IP ヘッダの ID 値に設定し（式 (5)），その値をもとに式 (6)，(7) の計算によって DNS ヘッダの ID 値と送信元ポート番号を設定することが分かった．

$$\text{Ip.id} = \text{Random\_value} \quad (5)$$

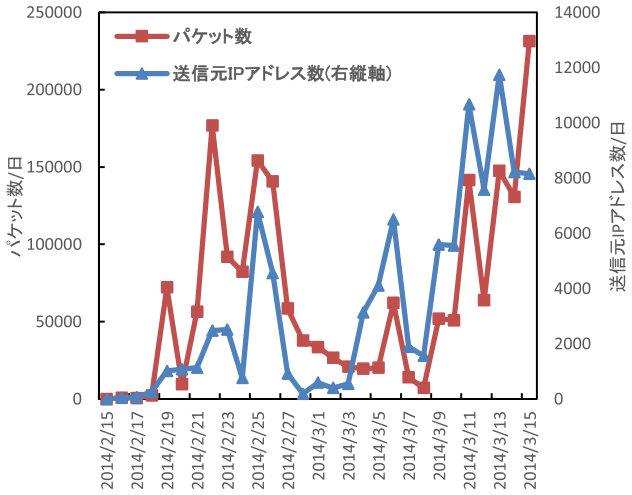

図 2 組み込み Linux 機器を狙ったマルウェアによる通信（22/TCP，23/TCP，8080/TCP，32764/TCP，58455/TCP 宛）

Fig. 2 Number of packets and source hosts to embedded Linux devices.

（ネットワークバイトオーダに変換しない）

$$\text{Dns.id} = \text{Random\_value} \ \&02345 \quad (6)$$

（ネットワークバイトオーダに変換）

$$\text{Udp.srcport} = \text{Random\_value} \ \%4996 + 1400 \quad (7)$$

（ネットワークバイトオーダに変換しない）

2.4.2 マルウェア XOR botnet

マルウェア XOR botnet の検体について，2.4.1 項と同様に静的解析を行い，パケットの生成方法を分析する．

- 検体ハッシュ値：7c903107ebc28a2e98e3247dbde71f02（MD5）

その結果，DNS Flood パケットはランダムに生成した 1 つの値（Random_value）を IP ヘッダの ID 値，送信元ポート番号，DNS ヘッダの ID 値に設定していることが分かった（式 (8)）．

$$\text{Ip.id} = \text{Udp.srcport} = \text{Dns.id} = \text{Random\_value} \quad (8)$$

2.5 組み込み Linux 機器を狙ったマルウェアによる通信

マクロ解析として，ダークネットで観測された組み込み Linux 機器を狙ったと思われるパケットを分析する．2014 年 2 月 16 日から，ダークネットで観測される 22/TCP（SSH），23/TCP（Telnet），8080/TCP，32764/TCP，58455/TCP 宛のパケットの一部に，シーケンス番号に 1112425812，IP ヘッダの ID 値に 0，ウィンドウサイズに 300 が設定されている SYN パケットを観測しており，それらのパケットは，複数の IP アドレスから大量にダークネットに向けて送信されている．2014 年 2 月 15 日から 3 月 15 日までの期間にダークネットで観測された，その特徴を持つ通信の 1 日あたりのパケット数と送信元 IP アドレス数を**図 2** に示す．

32764/TCP はメーカが意図的に作成したルータのバックドアポートであり，58455/TCP は組み込み Linux 機器を狙ったマルウェア Linux.Darlloz が作成するバックドアポートであると報告されている [17], [18]．以上のことから，これらのパケットは組み込み Linux 機器を狙ったマルウェアによるネットワークスキャンパケットであると推測される．

3. 提案手法

本章では，ヘッダフィールドを組み合わせたシグネチャを用いて，観測したパケットから送信元のマルウェアを特定する手法について説明する．

3.1 概要

p0f は pcap 形式のパケットキャプチャデータまたは，ネットワークインタフェースを通過する通信から TCP パケットを取得し，シグネチャを用いてパターンマッチングを行うことで，パケットごとに OS を判定するツールである．SYN パケットから OS を判定する際に p0f が使用するパケットの情報は，IP ヘッダの初期 TTL，TCP ヘッダの MSS（最大セグメントサイズ），ウィンドウスケール，ウィンドウサイズの 4 つのフィールドの具体的な値と，その他のヘッダフィールドのフラグビットの状態や値の有無（0 または値あり）である．IP ヘッダの ID 値，TCP ヘッダの初期シーケンス番号，送信元ポート番号等のヘッダフィールドは，OS によってパケットごとにランダム値等が動的に設定され，OS に固有の特徴は有しない．そのため p0f はそれらのフィールドの具体的な値を OS 判定に使用しない．

しかし，2 章の分析結果から，独自のネットワークスタックを実装したマルウェアやツールから発生したパケットは，OS がパケット送信ごとに動的に設定するヘッダフィールドに，つねにある固定値が設定される，またはある計算方法によって値が設定されるといった固有の特徴を持つことがあると判明した．TCP ヘッダのシーケンス番号と送信元ポート番号，IP ヘッダの ID 値，ICMP ヘッダの ID 値とシーケンス番号等がそのような特徴を持つことが多い．また，UDP のアプリケーションプロトコルヘッダの場合も，DNS ヘッダの ID 値等はマルウェアやツールに固有の特徴を持つことがある．

それらのフィールドは，サーバ・クライアント間のコネクション管理やパケットの識別のためのフィールドであり，その値は OS やアプリケーションによって動的に設定される．DDoS 攻撃やネットワークスキャン等，非対話的にパケットを送信することを目的としたマルウェアやツールは，それらのフィールドを使用する必要がない．したがって，それらのフィールドにマルウェアやツールに固有の特徴を持つ原因は，自由な値を設定しても DDoS 攻撃やネットワークスキャン等の目的を達成可能なためと考えられる．

そこで，それらのフィールドに着目し，各ヘッダフィールドの値を組み合わせたシグネチャを用いて，パケットごとにパターンマッチングを行うことで，送信元のマルウェアを特定する手法を提案する．本手法が対象とするパケットは，TCP パケット，UDP パケット，ICMP echo request パケットである．パケットとシグネチャのパターンマッチングを行う手順について説明する．本手法のパターンマッチングの基本的な仕組みは，p0f 等の OS Fingerprinting 技術を実装したソフトウェアのようにパケットのヘッダフィールドごとにシグネチャと照合する方法を採用している．まず，分析対象のパケットから本手法で使用するヘッダフィールドの値をすべて抽出する．次に，対応するシグネチャのフィールドの値を 1 つずつ照合，または計算式による検算の結果の照合を行う．すべてのフィールドがシグネチャと一致していた場合，そのパケットが陽性であると判定する．

本手法の独自性は，TCP パケットを検知する際には，p0f で使用されないシーケンス番号や送信元ポート番号等のヘッダフィールドも照合する点にある．さらに，ICMP ヘッダ，UDP ヘッダ，UDP を使用するアプリケーションプロトコルヘッダ（DNS ヘッダや NetBIOS ヘッダ等）にまで照合する対象を拡張することで，ICMP パケットと UDP パケットの検知を可能にしている．

また，Snort [7]，Bro [8] といったシグネチャ型のネットワーク型侵入検知システム（NIDS：Network Intrusion Detection System）は，主な検知方法の 1 つとして，TCP セッションを再構築しペイロードとシグネチャデータのパターンマッチングを行う．本論文で作成した SYN パケットのシグネチャで検知可能なパケットは，SYN Flood 攻撃やネットワークスキャンを目的としたパケットであり，セッション確立の要求目的のパケットではないため，ペイロードを検知する方法ではこれらのパケットを検知することはできない．しかし，提案手法はパケットのヘッダフィールドを照合するという単純な検知方法であるため，それらのシグネチャ型 NIDS に本手法を組み込むことが可能であると考えられる．本論文で作成したシグネチャによる Snort 等の NIDS 用のルールセットを公開し，各 NIDS の検知対象を拡張させることは今後の課題である．

3.2 シグネチャの構成と作成方法

提案手法のシグネチャの構成と作成方法について述べる．TCP パケット，UDP パケット，ICMP echo request パケットについて，独自のネットワークスタックを実装したマルウェアやツールに固有の特徴を持つ可能性のあるヘッダフィールドを組み合わせてシグネチャを構成する．シグネチャのフィールドは，OS がパケット送信ごとに動的に値を設定する TCP ヘッダのシーケンス番号や IP ヘッ

表 1 シグネチャのフィールド構成
Table 1 Signature format.

(a) TCP シグネチャ
(a) TCP signature

IP ヘッダ	ID 値
	TTL
TCP ヘッダ	送信元ポート番号
	宛先ポート番号
	シーケンス番号
	確認応答番号
	ウィンドウサイズ
	オプション

(b) UDP シグネチャ
(b) UDP signature

IP ヘッダ	ID 値
	TTL
UDP ヘッダ	送信元ポート番号
	宛先ポート番号
アプリケーションプロトコルヘッダ（存在する場合）	DNS，SNMP，NetBIOS ヘッダの ID 値等

(c) ICMP シグネチャ
(c) ICMP signature

IP ヘッダ	ID 値
	TTL
ICMP ヘッダ	ID 値
	シーケンス番号

ダの ID 値等と，OS が動的に値を変更しないがマルウェアやツールの特徴を持つことを 2 章の分析で確認した TCP ヘッダのウィンドウサイズやオプション等のヘッダフィールドとする．各シグネチャのフィールド構成を**表 1** に示す．たとえば，TCP パケットの場合，パケットのヘッダは IP ヘッダと TCP ヘッダで構成されるため，IP ヘッダの 2 個と，TCP ヘッダの 6 個の合計 8 個のフィールドでシグネチャは構成される．

次に，シグネチャフィールドに設定される値について説明する．フィールドごとに値の照合ができる場合は，単一の値，複数値，ワイルドカードのいずれかを設定する．また，複数のヘッダフィールドの値から算出された値を，別のヘッダフィールドの値としている場合は，計算に使用するヘッダフィールドと計算方法をシグネチャフィールドに設定する．

シグネチャの作成は，まず，マルウェア動的・静的解析結果の分析，インターネットに公開されているオープンソースのツールのソースコードの分析，ハニーポットやダークネットによる観測結果の分析から得たパケットの特徴を抽出する．表 1 に示した各フィールドについて，つねにある固定値が設定される，または，ある計算方法によって値が設定されるといった固有の特徴を基にシグネチャを構築する．このとき，検知対象とするプロトコル，攻撃やスキャン等のパケットについて，その全パケットに共通する特徴のみを適用する．つまり，対象のパケット群について検知漏れがないようにシグネチャを設定する．

3.3 適用範囲

提案手法の適用範囲は，独自のネットワークスタックを実装したソフトウェアが生成したパケットに限る．また，そのようなパケットであっても，TCP ヘッダのシーケンス番号や IP ヘッダの ID 値等の，OS が動的に設定するフィールドがランダム値等に設定されることで，ソフトウェアに固有の特徴を持たない場合は，検知に有効なシグネチャを作成することは難しいと考えられる．

2.5 節のように，ダークネットやハニーポットの観測データから特徴的なパケットを発見した場合，本手法を用いることによって，OS が送信したパケットや既知のマルウェアやツール等から送信されたパケットと，それらの未知のパケットを分類することが可能である．

また，独自のネットワークスタックを実装したマルウェアに亜種が存在したとすると，プログラム内の挙動やデータに関わる部分に変更があることが推測されるが，ネットワークスタックに変更を加えてもマルウェアとしての機能にほとんど影響はないと考えられる．したがって，マルウェアファミリは共通，または類似したネットワークスタックを実装している可能性が高いことが推測される．すべての特徴は一致しないが，特定のフィールドに設定される固有の値が一致している等，シグネチャ間の類似性が存在する場合は，同一のマルウェアファミリに属するマルウェアや，同一の攻撃者が開発したマルウェアであることを推測する手がかりになると考えられる．

4. シグネチャの作成

マクロ解析とミクロ解析によって得たパケットのヘッダを分析し，シグネチャの作成を行う．3.2 節で述べたシグネチャの作成方法に基づき，マクロ解析とミクロ解析で得たパケットから特徴抽出を行った結果，12 種のマルウェアとツールから生成される 19 種のパケットがシグネチャに適用可能な特徴を持つことが分かった．それらの特徴から作成したシグネチャを**表 A·1**，**A·2**，**A·3** に示す．

本章では，そのうち 2 章で分析を行ったマルウェア Morto，ネットワークスキャンツール ZMap・Masscan，DRDoS マルウェア IptabLes・XOR botnet，組み込み Linux 機器を狙ったマルウェアから送信されたパケットのシグネチャ作成について詳細に説明する．さらに，作成した各シグネチャの検知精度に関する考察を行う．シグネチャとのパターンマッチングの際，検知対象のマルウェアやツール以外から送信されたパケットのヘッダが偶然シグネチャと一

致することで誤検知（False-positive）となる場合がある．たとえば，IP ヘッダの ID 値は 16 bit のフィールドなので，OS はパケット生成時に 0〜65535 の値からランダムに選択する．したがって，65536 分の 1 の確率でこのフィールドがシグネチャと一致する可能性がある．そこで，OS のソケットを使用して送信されるパケットがシグネチャと偶然一致することで発生する誤検知の確率を算出する．このとき，OS やアプリケーションによって動的に設定されるフィールドに着目して分析を行う．ただし，すべてのシグネチャは，OS ごとの固有値が設定されることが多い IP ヘッダの TTL 値や TCP ヘッダのウィンドウサイズ等のフィールドを 1 つ以上照合する．それらのシグネチャフィールドが，OS の固有値（たとえば，Mac OS X や FreeBSD はウィンドウサイズを 65535 に設定する）以外の値の場合，実際の検知精度は算出した値よりも高くなる可能性がある．

4.1 マルウェア Morto

2.1 節の分析から，マルウェア Morto から送信されたインターネットに対するネットワークスキャンパケットに特徴を持つことが判明している．その特徴からシグネチャ Morto_scan を作成する（表 A·1）．このシグネチャのフィールドのうち，Windows や Mac OS X 等の一般的な OS の多くが動的に設定するフィールドは，32 bit のシーケンス番号，16 bit の送信元ポート番号，16 bit の IP ヘッダの ID 値である．OS がそれらのフィールドにランダム値を設定すると仮定した場合，単一の値がそれぞれのシグネチャフィールドに設定されているため，2^{64} 分の 1 の確率で誤検知が発生する可能性がある．

また，このシグネチャで検知できるパケットがルータ等を通過することで送信元ポート番号が変更される可能性を考え，Morto_scan の送信元ポート番号にワイルドカードを設定したシグネチャ Morto_scan_NAT を作成する（表 A·1）．このシグネチャの精度も同様に分析すると，誤検知が発生する確率は 2^{48} 分の 1 である．

4.2 ネットワークスキャンツール ZMap

2.2 節の分析から，ZMap によって生成された TCP SYN・UDP・ICMP echo request パケットは，IP ヘッダの ID 値に 54321 が設定される等の特徴を持つことが判明している．その特徴から，シグネチャ Zmap_tcp，Zmap_udp，Zmap_icmp を作成する（表 A·1，A·2，A·3）．

4.1 節と同様に誤検知について分析すると，OS によって動的に設定されるフィールドは，Zmap_tcp と Zmap_udp は IP ヘッダの ID 値であり，OS が生成したパケットを誤検知する確率は，2^{16} 分の 1 である．Zmap_icmp は ICMP ヘッダの 16 bit のシーケンス番号と IP ヘッダの ID 値に単一の値を持つため，誤検知の発生確率は，2^{32} 分の 1 である．

4.3 ネットワークスキャンツール Masscan

2.3 節の分析から，Masscan によって生成された TCP SYN・DNS・NetBIOS・SNMP パケットは特徴的な計算方法により IP ヘッダの ID 値を算出することが判明している．観測したパケットからそれらの 3 つのフィールドの値を抽出し，検算することで Masscan によって生成されたパケットかを判定可能なため，シグネチャ Masscan_tcp，Masscan_dns, Masscan_ntb, Masscan_snmp を作成する（表 A·1，A·2）．

これらのシグネチャの OS によって動的に設定されるフィールドは IP ヘッダの ID 値であるため，誤検知の発生確率は 2^{16} 分の 1 である．

4.4 DRDoS 攻撃の機能を持ったマルウェア

DRDoS 攻撃の機能を持ったマルウェア IptabLes，XOR botnet のシグネチャ作成について述べる．

4.4.1 マルウェア IptabLes

2.4.1 項の分析から，マルウェア IptabLes から送信される SYN Flood パケットの特徴を基にシグネチャ Iptables_tcp_1 を作成する（表 A·1）．OS が動的に設定するフィールドは，TCP ヘッダのシーケンス番号と IP ヘッダの ID 値に単一の値が設定されているため，誤検知の発生確率は 2^{48} 分の 1 である．

また，DNS Flood パケットの特徴を基に，シグネチャ Iptables_dns を作成する（表 A·2）．このシグネチャは，OS やアプリケーションが動的に設定する 16 bit の DNS ヘッダの ID 値，送信元ポート番号を式 (6)，(7) の計算手順で検算することで検知を行うため，誤検知が発生する確率は 2^{32} 分の 1 である．

4.4.2 マルウェア XOR botnet

2.4.1 項の分析から，マルウェア XOR botnet の DNS Flood パケットは IP ヘッダの ID 値，送信元ポート番号，DNS ヘッダの ID 値に同じ値がパケット送信ごとに設定されるという特徴を持つことが判明している．その特徴を基に，シグネチャ Xor_dns を作成する（表 A·2）．

このシグネチャは，OS やアプリケーションが動的に設定する IP ヘッダの ID 値，送信元ポート番号，DNS ヘッダの ID 値を用いて検知を行うため，誤検知が発生する確率は 2^{48} 分の 1 である．

4.5 組み込み Linux 機器を狙ったマルウェアによる通信

2.5 節の分析から，ダークネットで観測される 22/TCP, 23/TCP, 8080/TCP, 32764/TCP, 58455/TCP 宛のパケットの一部は共通の特徴を持つことが判明している．組み込み Linux 機器を狙ったマルウェアによる通信と思われるそのパケットの特徴を基に，シグネチャ Dark_embedded_linux_1 を作成する．このシグネチャの OS によって動的に設定されるフィールドは TCP ヘッダのシーケンス番号と IP ヘッ

ダの ID 値であるため，誤検知の発生確率は，2^{48} 分の 1 である．ただし，このシグネチャのウィンドウサイズは 300 であり，p0f（version3.07）の OS 判定シグネチャでは，このウィンドウサイズを固定値とする OS は存在しない．したがって，実際の検知精度はさらに高いと考えられる．

5. 検証実験

本章では，まず，提案手法の有効性を示すために，マルウェア Morto に関係する通信とネットワークスキャンツール ZMap から発生した通信について，ミクロ解析とマクロ解析の結果の相関分析を行う．双方の分析で得られたそれぞれのマルウェア・ツールから送信された可能性の高いパケットに対して，本手法の特徴抽出を行う．相関分析として，それぞれの特徴を比較し，それらが一致していた場合，本手法によって送信元の実装が共通であるパケットの特定と分類が可能であると判断する．次に，本手法による誤検知の発生について検証を行い，4 章で述べた各シグネチャの誤検知率が実際の不正通信検知を行う際に，十分な検知精度を持つかを考察する．

5.1 マルウェア Morto

マルウェア Morto に関係する通信について，マルウェア動的解析で得たパケットと，ダークネット観測で得た Morto 発生時期に急増しているパケットの特徴を比較する．

5.1.1 ミクロ解析

ミクロ解析として，マルウェア Morto の動的解析で得られたパケットの特徴を抽出する．2.1，4.1 節の分析結果を利用し，シグネチャ Morto_scan の特徴を，マルウェア Morto から発生したインターネットに対するネットワークスキャンパケットの特徴として相関分析に利用する．

5.1.2 マクロ解析

マクロ解析として，Morto のおおよその発生時期と思われる 2011 年 7 月 1 日から 8 月 9 日までのダークネット観測で観測した 3389/TCP 宛の SYN パケットの特徴を抽出し，その期間に増加している特徴的なパケットは Morto から送信された可能性が高いと推測し分析を行う．その結果，TCP ヘッダのシーケンス番号や IP ヘッダの ID 値にそれぞれ単一の値が設定されている 3 種類の特徴的なヘッダパターンを持つパケットを確認した．それらのパケットをパターン 3389_1，パターン 3389_2，パターン 3389_3 と分類し，日ごとの送信元 IP アドレス数を**図 3** に示す．

パターン 3389_1 の特徴を持つパケットは，2011 年 7 月 12 日から 27 日までの期間に多数のパケットを観測しているが，それ以降パケット数は減少し，現在は少数のパケットがダークネットで観測されている．パターン 3389_2 の特徴を持つパケットは，2011 年 7 月 20 日から 8 月 1 日の期間にのみ確認されたパケットであり，以降は現在までダークネットで確認されていない．パターン 3389_3 の特徴を持つパケットは，対象期間中に最も大量に確認されたパケットであり，2011 年 7 月 30 日の発生以降は継続的に複数のホストから送信されており，現在も大量のパケットをダークネットで観測することが可能である．また，3 種のパケットはそれぞれ異なるホストから送信されていることから，同一のネットワーク実装から送信されたパケットである可能性は低い．

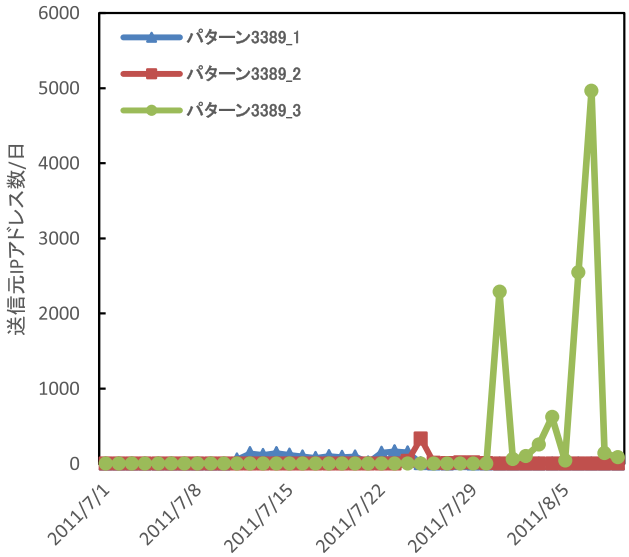

図 3 Morto 発生時期に観測された特徴的な 3389/TCP 宛通信の送信元 IP アドレス数

Fig. 3 Number of source hosts of the characteristic packets to TCP port 3389 around the appearance time of Morto.

以上の結果から，Morto の発生と同じ時期に急増したパターン 3389_3 の特徴を持つパケットは，Morto から送信されたパケットである可能性が高い．

5.1.3 相関分析

ミクロ解析とマクロ解析で得られた Morto から送信されたと推測されるパケットの特徴を比較する．ミクロ解析で得たシグネチャ Morto_scan の特徴とパターン 3389_3 の特徴は，TCP ヘッダのシーケンス番号と送信元ポート番号，IP ヘッダの ID 値の固有値のほか，本手法で抽出したすべてのヘッダフィールドが一致していた．マルウェア Morto の発生時期にダークネットで発生した後に急増した 3389/TCP 宛のネットワークスキャンパケットが，マルウェア動的解析によって Morto から送信されたパケットと同様の特徴を持っていたことから，提案手法によって Morto が生成したパケットの識別は可能であるといえる．さらに，Microsoft 等によって注意喚起が行われる約 1 カ月前から，Morto によるスキャンと推測される通信がすでに発生していたことが本手法によって明らかになった．このシグネチャによって検知可能なパケットは，2015 年 11 月現在も観測されることから，Morto の亜種や同様のパケット生成の実装を持ったマルウェアが現在も活動している可能性がある．また，パターン 3389_1，パターン 3389_2 の

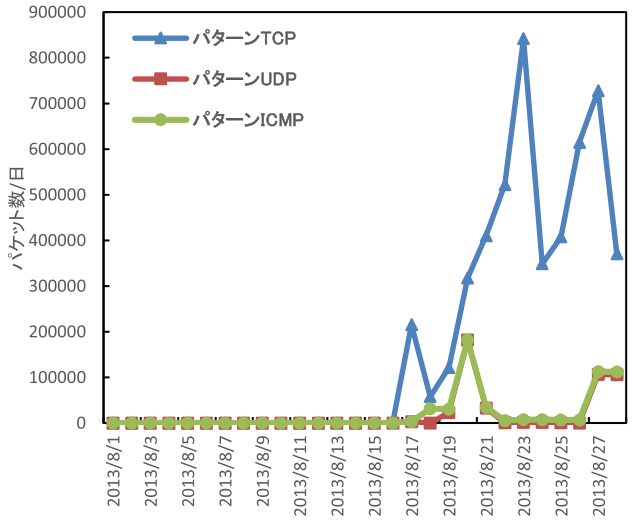

図 4 ZMap 公開時期に観測された特徴的なネットワークスキャンパケット

Fig. 4 Number of characteristic network scan packets around the time when ZMap was released.

特徴をシグネチャ Dark_dst3389_1，Dark_dst3389_2 として表 A·1 に示す．

5.2 ネットワークスキャンツール ZMap

ネットワークスキャンツール ZMap に関係する通信について，ソースコードの分析で得たパケットと，ダークネット観測で得た ZMap 公開時期に急増しているパケットの特徴を比較する．

5.2.1 ミクロ解析

ミクロ解析として，ZMap のソースコードの分析で得られたパケットの特徴を抽出する．2.2，4.2 節の分析結果を利用し，ZMap から送信された TCP，UDP，ICMP パケットのシグネチャ Zmap_tcp，Zmap_udp，Zmap_icmp の特徴を相関分析に利用する．

5.2.2 マクロ解析

マクロ解析として，ZMap 公開時期前後の 2013 年 8 月 1 日から 8 月 28 日までのダークネット観測データについて，ZMap を使用している可能性の高い組織等から送信されたパケットは ZMap の特徴を持っていると推測し，分析を行う．その結果，開発元であるミシガン大学の保有する IP アドレスや，ZMap の使用を公表している Project Sonar（Rapid7）等のプロジェクトの IP アドレスから，大量のネットワークスキャンパケットがダークネットに到達していることが判明した．それらのホストから送信されたパケットは 2013 年 8 月 15 日から発生した後に急増しており，TCP，UDP，ICMP パケットに IP ヘッダの ID 値等に固有値が設定される特徴を持っていた．それぞれのヘッダパターンの特徴をパターン TCP，パターン UDP，パターン ICMP とする．この期間にダークネットで観測された 3 種類の特徴を持つパケットの日ごとのパケット数を**図 4** に示す．これらの特徴を持つパケットは ZMap の公開時期から発生しており，ZMap を使用しているホストから大量に送信されているため，ZMap が生成したパケットである可能性が高い．また，上記の IP アドレス以外からも，これらの特徴を持つパケットが複数の IP アドレスから観測期間中に急増していることを確認している．

5.2.3 相関分析

ミクロ解析とマクロ解析で得られた ZMap から送信されたと推測されるパケットの特徴を比較する．ソースコードの分析で確認した 3 つのシグネチャのパケットの特徴と，ダークネットで観測された ZMap を使用している可能性の高いホストから送信された 3 種類のプロトコルの特徴的なパケットについて，シグネチャを構成するフィールドを比較する．その結果，IP ヘッダの ID 値の固有値等，本手法で抽出したすべてのヘッダフィールドの特徴が一致していた．ダークネットで観測された ZMap を使用して送信された可能性の高いパケットが，ミクロ解析によって抽出したパケットと同様の特徴を持っていたため，提案手法によって ZMap が生成したパケットの識別が可能であるといえる．

5.3 誤検知

提案手法のパターンマッチングにおける誤検知の発生を検証する．4 章では OS やアプリケーションが動的に設定するフィールドのデータ量に着目し，各シグネチャの精度の計算を行った．本節では，False-positive がどのくらいの頻度で発生するのかを確認するために，検証用のデータを対象に本論文で作成したシグネチャを用いて提案手法による検知を行う．ここで使用した検証用のデータは，大学研究室内（グローバル IP アドレスで構成される/24 ネットワーク（256 IP））で 2015 年 10 月 10 日から 10 月 16 日の期間に観測された通信データであり，研究室内→研究室内，研究室内→研究室外（インターネット）の TCP パケット，UDP パケット，ICMP パケットを対象とした．このネットワークでは，およそ 60 台程度の PC やプリンタ等のネットワーク機器が常時接続されている．また，観測期間に ZMap や Masscan 等のスキャンツールが使用されておらず，マルウェアに感染しているホストが存在しないことを確認している．

期間中に観測された各プロトコルの全パケット数と，検知パケット数を**表 2** に示す．検知されたパケットのすべてが 224.0.0.251（5353/UDP）宛のパケットであり，そのすべてがシグネチャ Zmap_udp によって検知された．また，それらのパケットの送信元は 3 ホストであった．このパケットは，Apple 社が開発したソフトウェア Bonjour が，サービス探索のために送信するマルチキャスト DNS パケットであると思われ，ZMap によって送信されたパケットである可能性は低い．したがって，これらのパケットは

表 2 研究室内ネットワークで検知された通信
Table 2 Number of detected packets in our laboratory's network.

日付	検知パケット数（TCP）	全パケット数（TCP）	検知パケット数（UDP）	全パケット数（UDP）	検知パケット数（ICMP）	全パケット数（ICMP）
2015/10/10	0	41422	0	2623472	0	11
2015/10/11	0	56404	11（1 ホスト）	2626603	0	4
2015/10/12	0	50847	0	2660529	0	35
2015/10/13	0	154161	0	4003154	0	404
2015/10/14	0	151818	22（2 ホスト）	4571336	0	288
2015/10/15	0	274997	0	5669666	0	272
2015/10/16	0	175040	0	2012294	0	148

偶然に Zmap_udp の特徴を持った誤検知のパケットであると判断できる．1 週間の検証期間で誤検知されたホスト数は 3 件であり，小規模なネットワークであれば管理者による対応が可能な範囲の誤検知数と考えられる．このシグネチャは本論文で作成した 19 種のシグネチャの中で誤検知の発生確率が最も高いため，他のシグネチャも同様に，十分な検知精度を持つと推察される．一方，大規模ネットワークにおいては誤検知数の増加が問題となりうる．この場合，独自のネットワークスタックを実装したマルウェアによる通信は，大量のパケットを短時間で送信することが多いことに着目し，送信元ホスト単位で検知を行うことでシグネチャとの偶然の一致を排除できる可能性がある．本論文で提案した手法の検知精度を向上させるために，ネットワークの規模や通信量に合わせてそのような分析方法を組み合わせることは，今後の課題である．

次に，本手法によるパターンマッチングの際の検知漏れ（False-negative）について考察する．まず，すべてのマルウェア等に起因する悪性パケットを対象とする場合について述べる．3.3 節で述べたように，独自実装のネットワークスタックに固有の特徴を持つパケット以外の通信は検知が不可能なため，すべて False-negative となる．次に，そのような特徴を持つパケットに限定した場合について述べる．本論文では，特定対象のマルウェアやツールが独自実装のネットワークスタックを使用して送信する，すべてのパケットが持つ同一の特徴からシグネチャを作成している．また，パケットが宛先に到達するまでに変化する可能性のある送信元ポート番号や TTL 等のヘッダフィールドは，別のシグネチャの作成や，閾値設定等によって対応している．そのため，ツールのソースコードの変更や，マルウェアのパケット生成エンジンの改ざん等，送信元の実装に変更があった場合のみ False-negative が発生する可能性がある．

6. 提案手法を用いた不正通信の分析

本章では，提案手法を用いたインターネット上の不正通信の分析事例を示す．4.4 節で作成した DRDoS 攻撃を行うマルウェアのシグネチャと，4.5 節で作成した組み込み Linux 機器を狙ったマルウェアのシグネチャを使用して，我々が運用している各ハニーポットの観測データを分析する．

6.1 DRDoS 攻撃を行うマルウェア

DRDoS 攻撃を行うマルウェアのシグネチャ Iptables_dns，Xor_dns を用いて，これらのマルウェアが実際の DRDoS 攻撃に使用されているかを確かめるために，我々が運用している DRDoS ハニーポット [19] で観測された通信を分析する．この DRDoS ハニーポットは UDP を用いた DNS や NTP 等のサーバを模擬し，攻撃者によってリフレクタ（踏み台サーバ）として悪用されることを目的としたハニーポットである．そのうち，オープンリゾルバとして動作する 7 センサの DNS ハニーポットで観測された 2015 年 1 月 1 日から 4 月 30 日までの通信データを分析対象とする．

2 つのマルウェアのシグネチャを用いたパターンマッチングの結果，偶然シグネチャに一致したと思われるごく少数のパケットは検知されたが，対象のマルウェアから発生したと思われる攻撃通信は観測できなかった．また，4.4 節のマルウェア動的解析でも，SYN Flood パケットや DNS Flood パケットを観測しているが，DRDoS 攻撃を目的として送信元を詐称している DNS パケットは観測していない．したがって，複数のセンサによる長期間の分析にもかかわらず，対象のマルウェアによる通信をいっさい観測しなかったため，攻撃者は DRDoS 攻撃を行う際に，これらのボットネットを使用しないと我々は推測している．

6.2 組み込み Linux 機器を狙ったマルウェアによる通信

ダークネットで観測された，組み込み Linux 機器を狙ったマルウェアによる通信のシグネチャ Dark_embedded_linux_1 を用いて，送信元のマルウェアの取得を試みるために，我々が運用している IoT 向けハニーポット [20] で観測された通信からパケットの分析を行う．このハニーポットは，安易なユーザ名/パスワードを設定した Telnet（23/TCP）サーバを模擬し，ログイン試行によって侵入したホストに対して，Linux ベースの仮

想端末を提供するハニーポットである．マルウェアをダウンロードするためのコマンドを受信した場合，実際のマルウェアを取得することが可能である．また，これまでの調査により，シグネチャ Dark_embedded_linux_1 の特徴を持つパケットを送信するマルウェアに関する報告は存在せず，送信元のソフトウェアは判明してない．そこで，このハニーポットを用いた送信元の特定を試みる．

2015 年 2 月 22 日から 10 月 19 日の期間に観測された通信のうち，シグネチャ Dark_embedded_linux_1 で検知可能なパケットを送信した IP アドレスを対象とし，検知した日と，その次の日に対象アドレスから送信されたパケットを分析する．その結果，多数のパケットを検知し，それらの送信元の一部は，別の TCP セッションにてハニーポットと通信を行うことを確認した．3WAY ハンドシェイクを確立し，FIN パケットによってコネクションを終了するまでを 1 セッションとし，送信したホスト数と，それらのホストから観測されたセッション数を**図 5** に示す．

観測されたセッションの通信内容を分析すると，その多くが 3WAY ハンドシェイクの後，「root/admin」等の安易なユーザ名とパスワードを試行し，端末への侵入を試みることを確認した．しかし，その後コマンド入力等の通信はすべてのセッションで確認できず，侵入から 1 分後に FIN パケットを送信しコネクションを終了するという共通の特徴が見られた．この通信の送信元の OS を調べるために，p0f（version3.07）を用いて SYN パケットの OS 判定を行うと，すべての通信が「Linux 2.4.x」や「Linux 2.2.x-3.x」といった Linux 系の OS と判定された．また，送信元の IP アドレスに対して Web ブラウザ（80/TCP）でアクセスを行うと，一部のアドレスは，ルータやネットワークカメラ等の認証画面が表示されることを確認した．以上の結果から，シグネチャ Dark_embedded_linux_1 により検知可能なパケットは，感染拡大の手段の 1 つとして Telnet 認証への辞書攻撃を行う，組み込み Linux 機器を攻撃対象としたマルウェアが生成したパケットであると推測される．ダークネットで観測される 22/TCP（SSH），23/TCP（Telnet），8080/TCP，32764/TCP，58455/TCP 宛のパケットの提案手法を用いた分析によって，同様の送信元の実装と思われるパケットを抽出することができ，それらの攻撃対象や攻撃の傾向が明らかになった．

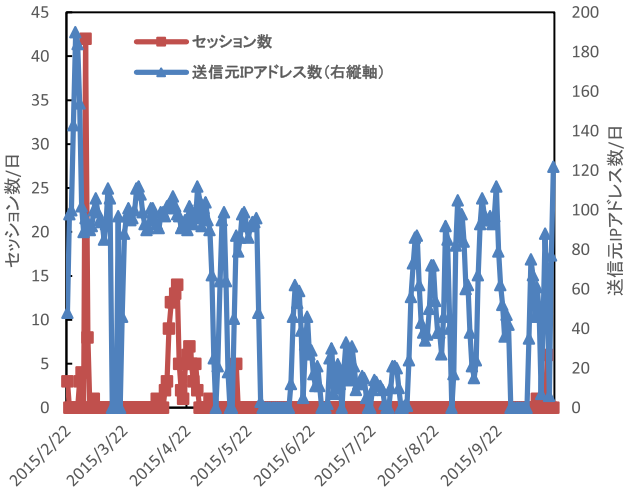

図 5 シグネチャ Dark_embedded_linux_1 によって検知された通信
Fig. 5 Number of packets and source hosts detected by Dark_embedded_linux_1.

7. 提案手法の NICTER への導入

本章では，提案手法を用いた不正通信の分析の応用例として，本手法を実装したネットワーク観測・分析システムについて説明する．NICT が研究開発を進めているインシデント分析センタ NICTER に，本手法の検知方法を組み込むことによって，ダークネットで観測された通信からマルウェアやネットワークスキャンツールによるパケットをリアルタイムに検知するシステムを構築する．

NICTER は主なコンポーネントとして国内外の大規模ダークネットを観測するマクロ解析システム，マルウェアの動的・静的解析を行うミクロ解析システムと，双方の結果の相関関係を分析しインシデントの判定を行う相関分析システムによって構成される．NICTER は約 27 万の IP アドレスを，パケットの送信元に対して応答を行わないブラックホールセンサのダークネットとして観測している．そこで観測されるパケットを入力とし，本手法を用いて観測したパケットの送信元のマルウェアやツールをリアルタイムに特定するシステムを開発する．また，本システムによって得られた分析データが，インターネット上の不正通信の実態把握に有益であることを示すため，DRDoS 攻撃に悪用されることが多いポートに対するネットワークスキャンについて分析する．

7.1 システム構成

本システムは，NICTER のダークネット観測データ収集サーバからリアルタイムにパケットデータを受け取り，提案手法によるパターンマッチングを行い，その検知結果をデータベースに保存する．本システムの処理は，大きく 3 つのフェーズに分かれており，その処理の流れを**図 6** に示す．パケット分析フェーズでは，収集サーバから受け取ったパケットデータから分析対象のプロトコルのパケットを取り出す．ヘッダ抽出フェーズではパターンマッチン

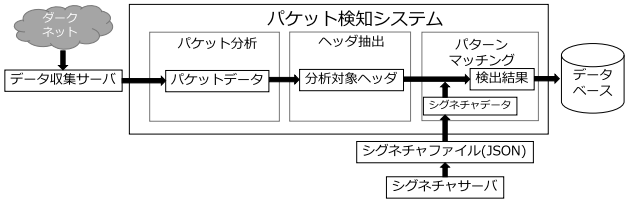

図 6 NICTER に導入した提案手法によるパケット検知システム
Fig. 6 Packet analysis system in NICTER.

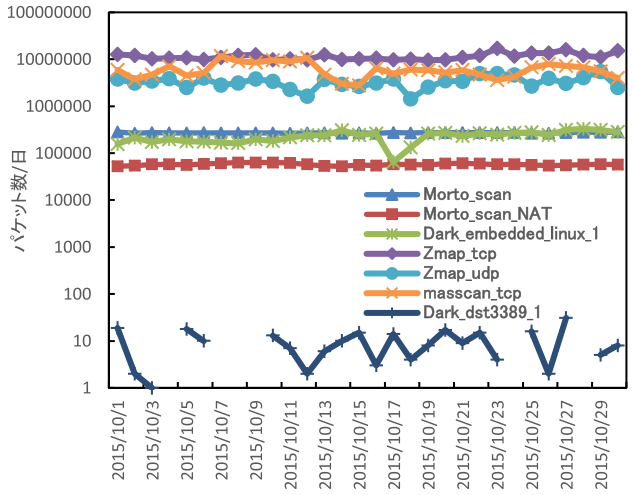

図 7 27 万 IP アドレスのダークネットで検知された通信
Fig. 7 Number of detected packets in the darknet (270 thousand IP addresses).

グに使用する各プロトコルのヘッダを抽出する．パターンマッチングフェーズでは，提案手法の検知方法によってパケットとシグネチャを照合する．その検知結果を日ごとのテーブルに 1 パケットにつき 1 レコードとして，MySQL で構築したデータベースに保存する．ここで使用するシグネチャは，日ごとにシグネチャサーバにアクセスし，最新のシグネチャファイルを取得してシステム内のシグネチャデータを更新する．最新のシグネチャファイルは Web サイト [21] で公開している．

データベースの各レコードは送信元と宛先の IP アドレスとポート番号，検知したシグネチャ名，検知時のシグネチャバージョン，ダークネットセンサの ID，パケット識別のために NICTER で番号付けされるパケット ID の 8 種のフィールドを持つ．なお，今回の分析では DNS ヘッダ等のアプリケーションプロトコルヘッダの情報は省いたため，Masscan_dns 等 UDP のシグネチャの一部は検知されない．

本システムによる検知結果の例として，2015 年 10 月 1 日から 10 月 30 日の期間に約 27 万の IP アドレスで観測された結果を**図 7** に示す．対象期間では 7 種のネットワークスキャンに関係するシグネチャのパケットが検知された．ZMap や Masscan 等のネットワークスキャンツールによるパケットが最も多く，マルウェア Morto に関係するパケットや組み込み Linux 機器を狙ったマルウェアによるパケット等，マルウェアから発生したと思われるネットワークスキャンも多数検知している．

7.2 検知結果の分析

本システムの有効性を示すため，DRDoS 攻撃に悪用されることの多いプロトコルに対するネットワークスキャンについて，我々のこれまでの分析の結果をもとに，本システムで得られた検知結果を分析する．

我々は文献 [19] で，観測しているダークネットの全アドレスのうち，/16 ネットワークに対して送信されたパケットを対象とし，DRDoS 攻撃に悪用されることの多い QOTD (17/UDP)，CharGen (19/UDP)，DNS (53/UDP)，NTP (123/UDP)，NetBIOS (137/UDP)，SNMP (161/UDP)，SSDP (1900/UDP)，27015/UDP，27960/UDP の 9 種の宛先ポート番号のパケットを対象に分析している．ここで，ダークネットの 64 アドレス以上にパケットを送信しているホストをスキャナと定義する．DNS の逆引き等の方法を用いて，それらのポートにパケットを送信するスキャナの組織情報を調べると，Shodan，Shadowserver 等のスキャンプロジェクトや，大学やセキュリティ関連企業等，調査目的と思われるホストが多数含まれることが判明した．期間中に観測したスキャナは 12,653 ホストであり，そのうち調査目的のホストは 1,296 ホスト (10.1%) である．したがって，DRDoS 攻撃に悪用するためのリフレクタを探索する通信は，攻撃者やマルウェアだけではなく，調査目的のホストによっても一定量送信されていることが分かった [19]．

そこで本論文では，ZMap のシグネチャにより検知可能なパケットと調査目的のホストの関係をプロトコルごとに分析する．/16 ネットワークのダークネットで観測された，9 種のポート番号宛の通信をシグネチャ Zmap_udp によって検知されたパケットと，Zmap_udp によって検知されないパケットに分類し，それぞれのパケット数を積み上げ棒グラフとして**図 8** に示す．

ダークネットに到達した通信は，分析対象期間ですべてのプロトコル宛のパケットの数が増加傾向にあり，ZMap によって送信されたと思われる通信も，2014 年 3 月頃から増加していることを確認した．また，DNS や 27960/UDP 宛の通信に見られるように，プロトコルによっては大部分を ZMap から送信されたパケットが占めている時期が存在することを確認した．

次に，期間中に観測したスキャナと ZMap をスキャンに使用しているホストを比較する．スキャンに ZMap を使用していると思われるホストは 727 ホスト (5.7%) 存在し，ZMap を使用し，かつ調査目的のホストは 507 ホスト (4.0%) 存在することが分かった．今回確認できたのは DNS 逆引き情報等から確証が持てる調査目的のホストであるため，実際にはさらに多くの調査目的のホストが，ZMap を使用してネットワークスキャンを行っていると推測される．

7.3 感染ホスト情報の提供

本システムの検知結果から，マルウェア感染の疑いのあるホストの情報を用いて注意喚起をすることについて考察する．4.1，4.2 節の分析から，ダークネットで観測される

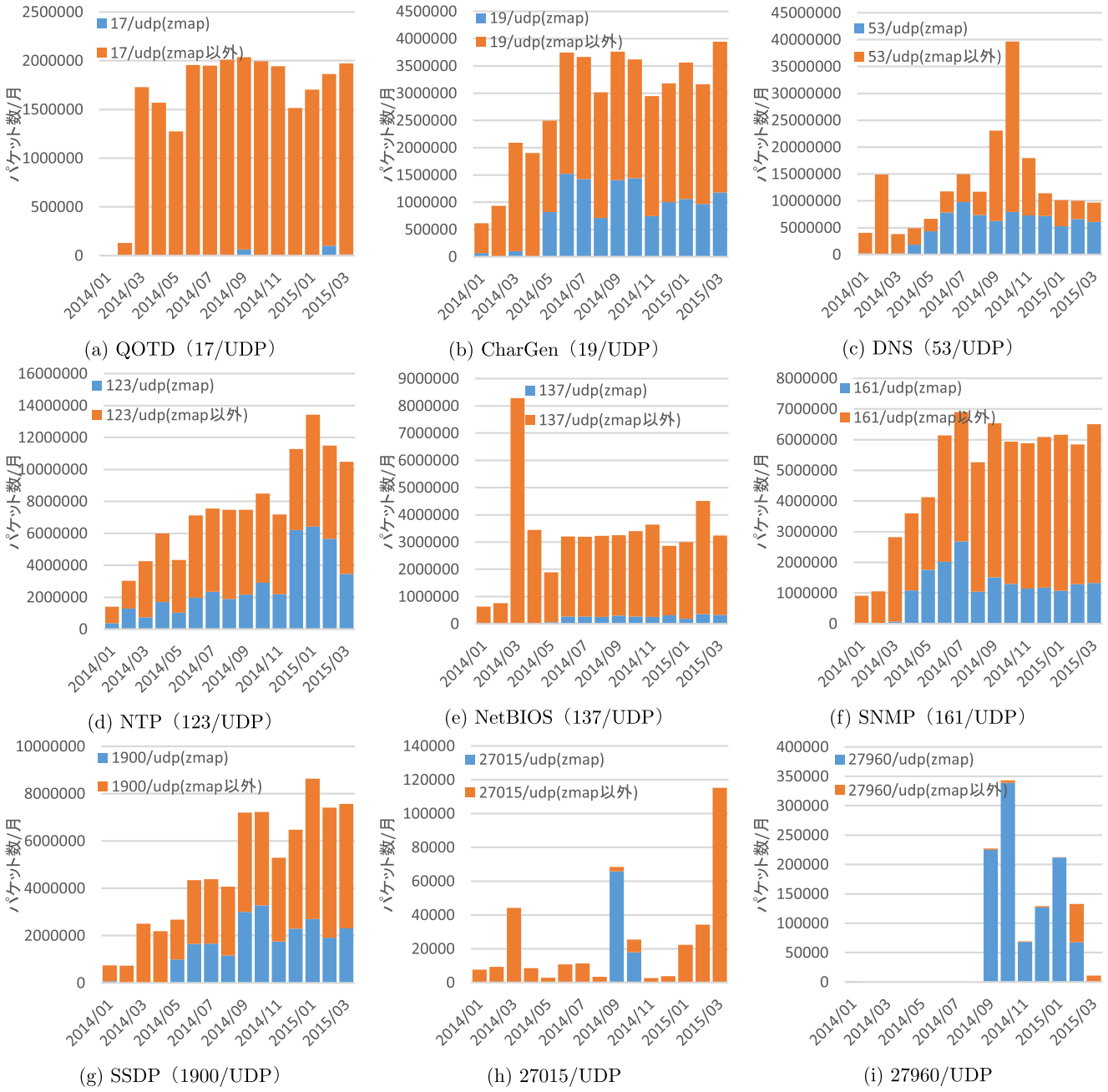

図 8 ダークネットで観測された DRDoS 攻撃に悪用されるプロトコル宛の通信の ZMap によるパケットの割合

Fig. 8 Number of packets to UDP ports abused for DRDoS attack in the darknet.

Morto に関係する通信と，組み込み Linux 機器を狙ったマルウェアによる通信は，マルウェアに感染しているホストから送信された可能性が高い．また，それらの通信の送信元を調べると，日本国内に存在する複数の IP アドレスが該当する場合がある．そこで，本システムで得られたマルウェア感染ホストの情報をもとに，国内の各 ISP と連携することで，感染の疑いのあるホストに対して注意喚起を行うことが可能である．マルウェア感染の対処を行ううえで，感染しているマルウェアとその攻撃の傾向の情報は重要な価値があると考えられる．本システムの分析結果を用いた注意喚起を行い，それらの付加情報を提供することは今後の課題である．

8. まとめと今後の課題

本論文は，観測したパケットから独自のネットワークスタックを実装したマルウェアやツールを特定する手法を提案し，Morto に関連する通信と ZMap に関連する通信の，ミクロ解析とマクロ解析の結果の相関分析によってその有効性を示した．また，本手法を用いたインターネット上の不正通信の分析事例として，DRDoS 攻撃を行うマルウェアと組み込み Linux 機器を狙ったマルウェアによる通信の分析結果を報告した．さらに，NICTER に本手法の検知

方法を適用することで，大規模ダークネットからリアルタイムに不正通信の検知を行うことが可能になった．今後の課題として，まず，本論文で作成したシグネチャを NIDS に適用させるためのルールセットを公開することがあげられる．また，独自のネットワークスタックを実装したマルウェアやツールは，大量のパケットを短時間で送信することが多いことに着目し，送信元ホスト単位で検知を行う方法を提案手法に組み合わせることで，検知精度の向上を図ることが可能である．さらに，NICTER に導入した本手法による分析システムから得られる情報をもとに，日本国内の ISP と連携することによって，感染ホストに対して注意喚起を行う仕組みを構築することは今後の課題である．

謝辞 本研究の一部は，総務省情報通信分野における研究開発委託/国際連携によるサイバー攻撃の予知技術の研究開発/サイバー攻撃情報とマルウェア実体の突合分析技術/類似判定に関する研究開発により行われた．また，本研究の一部は，文部科学省国立大学改革強化推進事業の支援を受けて行われた．

参考文献

[1] Zalewski, M.: p0f v3 (online), available from ⟨http://lcamtuf.coredump.cx/p0f3⟩ (accessed 2015-11-23).
[2] 木佐森幸太，下田晃弘，森 達哉，後藤滋樹：TCP フィンガープリントによる悪意のある通信の分析，情報処理学会論文誌，Vol.52, No.6, pp.2009–2018 (2011).
[3] 中里純二, 島村隼平, 衛藤将史, 井上大介, 中尾康二：nicter によるネットワーク観測および分析レポート—ネットワークインシデントの前兆，信学技報，Vol.113, No.95, ICSS2013-14, pp.79–84 (2013).
[4] Yamada, R. and Goto, S.: Using abnormal TTL values to detect malicious IP packets, *Proc. Asia-Pacific Advanced Network* (*APAN*), Vol.34, pp.27–34 (2013).
[5] Durumeric, Z.: Zmap: The Internet Scanner (online), available from ⟨https://zmap.io/⟩ (accessed 2015-11-23).
[6] 情報通信研究機構：NICTER（オンライン），入手先 ⟨www.nicter.jp⟩（参照 2015-11-23）.
[7] Snort (online), available from ⟨https://www.snort.org/⟩ (accessed 2015-11-23).
[8] The Bro Project: The Bro Network Security Monitor (online), available from ⟨https://www.bro.org/⟩ (accessed 2015-11-23).
[9] 竹久達也, 井上大介, 衛藤将史, 吉岡克成, 笠間貴弘, 中里純二，中尾康二：サイバーセキュリティ情報遠隔分析基盤 NONSTOP，電子情報通信学会技術研究報告，Vol.113, No.95, ICSS2013-15, pp.85–90 (2013).
[10] Microsoft: Encyclopedia entry: Worm: Win32/Morto.A, Malware Protection Center (online), available from ⟨http://www.microsoft.com/security/portal/threat/encyclopedia/entry.aspx?Name=Worm:Win32/Morto.A⟩ (accessed 2015-11-23).
[11] F-Secure Labs: Worm: W32/Morto.A Description (online), available from ⟨https://www.f-secure.com/v-descs/worm_w32_morto_a.html⟩ (accessed 2015-11-23).
[12] 吉岡克成，村上洸介，松本 勉：マルウェア感染ホスト検出のためのネットワークスキャン手法と検出用シグネチャの自動生成，情報処理学会論文誌，Vol.51, No.9, pp.1633–1644 (2010).
[13] Robert Graham: MASSCAN: Mass IP port scanner, GitHub (online), available from ⟨https://github.com/robertdavidgraham/masscan⟩ (accessed 2015-11-23).
[14] Durumeric, Z., Bailey, M. and Halderman, J.: An Internet-Wide View of Internet-Wide Scanning, *Proc. 23rd USENIX Security Symposium* (2014).
[15] [state of the internet]: IPTABLES AND IPTABLEX DDOS BOTS THREAT ADVISORY (online), available from ⟨https://www.stateoftheinternet.com/resources-web-security-threat-advisories-2014-iptables-iptablex-linux-bots-botnet.html⟩ (accessed 2015-11-23).
[16] [state of the internet]: XOR DDOS [HIGH RISK] (online), available from ⟨https://www.stateoftheinternet.com/resources-web-security-threat-advisories-2015-xor-ddos-attacks-linux-botnet-malware-removal-ddos-mitigation-yara-snort.html⟩ (accessed 2015-11-23).
[17] Gallagher, S.: Backdoor in wireless DSL routers lets attacker reset router, get admin, arstechnica (online), available from ⟨http://arstechnica.com/security/2014/01/backdoor-in-wireless-dsl-routers-lets-attacker-reset-router-get-admin/⟩ (accessed 2015-11-23).
[18] Blinka, H.: Linux.Aidra vs Linux.Darlloz: War of the Worms, AVG. Now (online), available from ⟨http://now.avg.com/war-of-the-worms/⟩ (accessed 2015-11-23).
[19] Krämer, L., Krupp, J., Makita, D., Nishizoe, T., Koide, T., Yoshioka, K. and Rossow, C.: AmpPot: Monitoring and Defending Am-plification DDoS Attacks, *Proc. 18th Interna-tional Symposium on Research in Attacks, Intrusions and Defenses* (*RAID 2015*) (2015).
[20] Pa, Y.M.P., Suzuki, S., Yoshioka, K., Matsumoto, T., Kasama, T. and Rossow, C.: IoTPOT: Analysing the Rise of IoT Compromises, *Proc. 9th USENIX Workshop on Offensive Technologies* (*WOOT'15*) (2015).
[21] 小出 駿：特徴的な TCP/IP ヘッダによるパケット検知ツール tkiwa，横浜国立大学情報・物理セキュリティ研究拠点（オンライン），入手先 ⟨http://ipsr.ynu.ac.jp/tkiwa/index.html⟩（参照 2015-11-23）.
[22] 小出 駿, 鈴木将吾, 牧田大佑, 村上洸介, 笠間貴弘, 島村隼平, 衛藤将史, 井上大介, 吉岡克成, 松本 勉：通信プロトコルのヘッダの特徴に基づく不正通信の検知・分類手法，情報処理学会コンピュータセキュリティシンポジウム 2014 論文集，pp.48–55 (2014).

付　　録

表 A·1 TCP シグネチャ
Table A·1 TCP signatures.

シグネチャ名	IP ヘッダ		TCP ヘッダ					
	ID 値	初期 TTL	シーケンス番号	確認応答番号	送信元ポート番号	宛先ポート番号	ウィンドウサイズ	オプション
Morto_scan	9496	0〜64	2406000322	0	4935	3389	65535	
Morto_scan_NAT	9496	0〜64	2406000322	0	*	3389	65535	
Dark_dst3389_1	256	65〜128	1210253312	0	*	3389	16384	
Dark_dst3389_2	256	65〜128	2284205602	0	*	3389	512	
Dark_embedded_linux_1	0	0〜64	1112425812	0	22, 23, 8080, 32764, 58455	300	*	
Zmap_tcp	54321	129〜255	*	0	*	*	65535	
Dark_ipid0_1	0	0〜64	*	0	*	0	8192	
Iptables_tcp_1	0	129〜229	848	0	*	*	16000〜18999	848 byte payload
Srizbi_1	*	65〜128	6509	0	4099	24000	0x1F219A	
Linuxddos1_tcp_1	0	129〜255	0	0	*	*	6000	960〜980 byte payload
Masscan_tcp	式 1	129〜255	*	0	*	*	1024	
Dark_scan_1	256	65〜128	*	0	12200	*	8192	

表 A·2 UDP シグネチャ
Table A·2 UDP signatures.

シグネチャ名	IP ヘッダ		UDP ヘッダ		アプリケーションプロトコルヘッダ
	ID 値	初期 TTL	送信元ポート番号	宛先ポート番号	
Zmap_udp	54321	129〜255	*	*	
Masscan_dns	式 (2)	129〜255	*	53	
Masscan_ntb	式 (3)	129〜255	*	139	
Masscan_snmp	式 (4)	129〜255	*	161	
Iptables_dns	式 (5)	129〜255	式 (7)	53	式 (6)（DNS ヘッダ ID）
Xor_dns	式 (8)	0〜178	式 (8)	53	式 (8)（DNS ヘッダ ID）

表 A·3 ICMP シグネチャ
Table A·3 ICMP signatures.

シグネチャ名	IP ヘッダ		ICMP ヘッダ	
	ID 値	初期 TTL	ID 値	シーケンス番号
Zmap_icmp	54321	129〜255	*	0

推薦文

通信プロトコルのヘッダの特徴に基づき，マルウェアやスキャンツールの通信を識別するための手法を提案している．検体の静的・動的解析とダークネットの観測を併用することによって提案手法を評価し，その有効性を確認している点を評価し，推薦論文として推薦したい．

（コンピュータセキュリティ研究会主査　西垣正勝）

小出 駿

2016 年 3 月横浜国立大学大学院環境情報学府情報メディア環境学専攻博士課程前期修了，修士（工学）．同年 4 月日本電信電話株式会社に入社．ネットワークセキュリティの研究に従事．

鈴木 将吾

2016 年 3 月横浜国立大学大学院環境情報学府情報メディア環境学専攻博士課程前期修了．修士（工学）．同年 4 月 PwC サイバーサービス合同会社に入社．在学中，ネットワーク攻撃観測等のネットワークセキュリティの研究に従事．

牧田 大佑　（学生会員）

2014 年 3 月横浜国立大学大学院環境情報学府情報メディア環境学専攻博士課程前期修了，修士（情報学）．同年 4 月横浜国立大学大学院環境情報学府情報メディア環境学専攻博士課程後期に進学．同年 4 月より独立行政法人情報通信研究機構で研究員として勤務．ネットワーク攻撃観測等のネットワークセキュリティの研究に従事．

村上 洸介

2012 年 3 月横浜国立大学大学院環境情報学府情報メディア環境学専攻博士課程前期修了．修士（工学）．同年 4 月より KDDI 株式会社に入社し，セキュリティ運用業務およびネットワーク攻撃通信解析の研究開発業務に従事．

笠間 貴弘　（正会員）

2014 年 3 月横浜国立大学大学院環境情報学府情報メディア環境学専攻博士課程後期修了．博士（工学）．2011 年 4 月より情報通信研究機構に研究員として入所．マルウェア解析やネットワーク攻撃観測・分析等サイバーセキュリティの研究開発に従事．2011 年情報処理学会山下記念研究賞受賞．

鈴木 未央　（正会員）

2008 年 3 月奈良先端科学技術大学院大学博士後期課程単位取得退学．博士（工学）．同年 4 月より情報通信研究機構に入所．ネットワーク攻撃観測・分析等ネットワークセキュリティの研究開発に従事．

島村 隼平

2003 年 3 月芝浦工業大学電気工学科卒業．2011 年 4 月より株式会社クルウィット勤務．2005 年からダークネットをはじめとするインターネット上の不正通信の観測・分析作業に従事．

衛藤 将史

国立研究開発法人情報通信研究機構セキュリティ人材育成研究センター研究マネージャー．2005 年情報通信研究機構に入所．以降，同機構サイバーセキュリティ研究室研究員．2013 年より同機構サイバー攻撃対策総合研究センターサイバー防御研究室主任研究員（兼務）．2016 年より現職．ネットワーク運用管理技術，アプリケーショントレースバック技術，NICTER プロジェクト，IPv6 セキュリティ，ITS セキュリティ等サイバーセキュリティ関連技術の研究開発に従事．また，セキュリティ人材育成の業務および研究活動に取り組む．2007 年暗号と情報セキュリティシンポジウム（SCIS）論文賞，2009 年科学技術分野の文部科学大臣表彰（科学技術賞）等を受賞．博士（工学）．

井上 大介

2003 年横浜国立大学大学院工学研究科博士課程後期修了．2003 年通信総合研究所（現，情報通信研究機構）に入所．2006 年よりインシデント分析センター NICTER の研究開発に従事．現在，情報通信研究機構サイバーセキュリティ研究所サイバーセキュリティ研究室室長．2002 年暗号と情報セキュリティシンポジウム論文賞，2009 年科学技術分野の文部科学大臣表彰（科学技術賞），2013 年グッドデザイン賞，2014 年 Asia-Pacific Information Security Leadership Achievements 等を受賞．博士（工学）．

中尾 康二 （正会員）

1979 年早稲田大学卒業後，国際電信電話（株）に入社．KDD 研究所を経て，現在 KDDI（株）顧問，および国立研究開発法人情報通信研究機構（NICT）サイバーセキュリティ研究所主管研究員兼務．ネットワークおよびシステムを中心とした情報セキュリティ技術の研究開発に従事．電子情報通信学会等の会員．経済産業省大臣表彰賞，KPMG 情報セキュリティアウォーズ，文部科学省大臣表彰賞，情報セキュリティ文化賞，総務大臣表彰等を受賞．

吉岡 克成 （正会員）

2005 年 3 月横浜国立大学大学院環境情報学府情報メディア環境学専攻博士課程後期修了，博士（工学）．同年 4 月独立行政法人情報通信研究機構研究員．2007 年 12 月より横浜国立大学学際プロジェクト研究センター特任教員（助教）．2011 年 4 月より横浜国立大学大学院環境情報研究院准教授．マルウェア解析やネットワーク攻撃観測・検知等のネットワークセキュリティの研究に従事．2009 年文部科学大臣表彰・科学技術賞（研究部門）受賞．

松本 勉

1986 年 3 月東京大学大学院工学系研究科電子工学専攻博士課程修了，工学博士．同年 4 月横浜国立大学講師．2001 年 4 月より同大学院環境情報研究院教授．2014 年 12 月より同大学先端科学高等研究院主任研究者を兼務．ネットワーク・ソフトウェア・ハードウェアセキュリティ，暗号，耐タンパー技術，生体認証，人工物メトリクス等の「情報・物理セキュリティ」の研究教育に 1981 年より従事．1982 年にオープンな学術的暗号研究を目指した「明るい暗号研究会」を 4 名で創設．2005 年～2010 年国際暗号学会 IACR 理事．1994 年第 32 回電子情報通信学会業績賞，2006 年第 5 回ドコモ・モバイル・サイエンス賞，2008 年第 4 回情報セキュリティ文化賞，2010 年文部科学大臣表彰・科学技術賞（研究部門）受賞．