# パケットマペット: パケットの特性に基づく格付け手法の提案

白畑 真[†] 南 政樹[†] 村井 純[†,‡]

[†] 慶應義塾大学大学院 政策・メディア研究科, [‡] 慶應義塾大学 環境情報学部

{true,minami,jun}@sfc.wide.ad.jp

本論文では、ワームやDoS攻撃等の不正トラフィックを識別するため、IPパケットに対する格付け手法を提案するとともに、提案手法のプロトタイプ実装の結果を報告する。提案手法では、低インタラクション型ハニーポットで収集したトラフィックと、実運用ネットワークのトラフィックではIPヘッダに異なった傾向が存在する点に着目し、IPヘッダの内容に対してベイズ推論によりスコアを算出する。そして、ネットワーク管理者の定義に基づき格付けを行う。

# Packet MUPPET: Packet Metrics and Unique Protocol Parameter Evaluation Technique

Shin Shirahata[†] Masaki Minami[†] Jun Murai[†,‡]

[†]Graduate School of Media and Governance, Keio University

[‡]Faculty of Environmental Information, Keio University

In this paper, we propose a packet-based rating technique in order to identify worms, DoS attacks and other malicious traffic, and discuss a prototype implementation of the technique.

The proposed technique involves scoring packets using a Naive Bayesian inference, by focusing on differences between traffic patterns observed in the IP packet headers in a low-interaction honeypot and a real network. The scores are rated according to definitions by the network administrator.

## 1 背景

インターネットが重要な社会インフラストラクチャとなるにつれ、その可用性に対する要求が高まっている。ネットワークの障害は、接続されている多くのホストに影響を及ぼすため、単一のホストの障害と比較して影響範囲が格段に大きい。例えば、2003年1月に大流行したSQL Slammerワームは、UDPパケットによるトラフィックによって世界中のネットワークの輻輳や、ルータへの過負荷を引き起こし、通信品質の低下やネットワークの停止などの影響を及ぼした。

## 2 目的

ネットワークをワームやDoS攻撃、DDoS攻撃に伴う不正トラフィックから保護するためには、不正なトラフィックを適切に識別し、対処する必要がある。本研究では、未知の攻撃手法に対応でき、広帯域なバックボーンネットワークに適用可能な高いスケーラビリティを備えたトラフィックの識別を目的とする。

## 3 既存手法

既存の不正トラフィック識別手法としては、ミスユース方式のNIDS(Network-based Intrusion Detection System)や、インライン型のNIDSであるIPS(Intrusion Prevension System)が用いられている。

しかし、Snort[5]などのミスユース方式のNIDSやIPSは、シグネチャ・データベースに登録されていない不正トラフィックを検知できないため、未知のワームによるアタックが発生したとしても、当該トラフィックを識別できない。

また、NIDSやIPSは、さまざまなレイヤにおいてトラフィックデータを正規化処理後にパターンマッチを行っているため、多くの計算機資源を必要とする。表1に正規化処理の例を示す。

このように、既存のNIDSやIPSはスケーラビリティに限界があるため、10Gbps Ethernetなどの広帯域化するネットワーク環境への適用は困難である。

また、PHAD[3]では、パケットのヘッダフィールドに対してPPMC法を用いてアノーマリ検知手法を提案している。PHADでは、値に対してクラスタリ

表 1　正規化処理の例

| | |
|---|---|
| ネットワーク層 | IP フラグメントの再構成 |
| トランスポート層 | TCP ストリームの再構成 (パケットロス, 重複, 破損, リオーダリングなど) |
| アプリケーション層 | アプリケーションデータの再構成<br>例: HTTP の場合:<br>/cgi-bin/dummy/../phf.cgi や<br>/%2f%63%67%69%2d%62%69%6e<br>%2f%70%68%66%2e%63%67%69 を<br>/cgi-bin/phf.cgi に正規化する必要がある。 |

ングを行っているため、特定のポート番号に対する攻撃を適切に評価できず、誤検知が多い。

## 4　要件

バックボーンネットワークにおいて不正トラフィックを識別するための要件は以下の通りである:

**網羅性**

未知のワームなどに伴うトラフィックなど、事前にデータベース等に登録が行われていない種類の不正トラフィックを識別できること。

**省資源性**

アクセスネットワークだけではなく、バックボーンネットワークなど、大量のトラフィックが発生する箇所においても適用できるよう、処理に必要な計算機資源ができるだけ少ないこと。

**高速性**

トラフィックを取得後、迅速に不正トラフィックを識別できること。

## 5　前提

筆者らは 2004 年より低インタラクション型ハニーポットを運用している。運用の結果、ハニーポットで収集したトラフィックと、研究ネットワークのトラフィックを IP ヘッダに着目して分析したところ、異なった傾向が存在することが明らかになった。

**低インタラクション型ハニーポット**

本研究では、ハニーポットに低インタラクション型ハニーポットの Dumnet[4] を利用した。Dumnet は、以下の機能により、あらかじめ定義した IP アドレス空間にホストが存在するように振る舞う。

**ICMP エミュレーション機能**　指定された IP アドレス宛の ICMP ECHO Request メッセージに対して、ECHO Reply メッセージを応答する。

**TCP エミュレーション機能**　SYN フラグがセットされた TCP パケットに対して、SYN+ACK フラグをセットしたパケットを返送することで、3-way handshake を成立させ、全ての TCP ポートが開かれているかのように動作する。

**ハニーポットのトラフィック**

本研究においては、未使用であった IP アドレス空間をインターネットに広報し、当該プレフィックスすべてをハニーポット専用に割り当てた。

未使用アドレス空間に設置したハニーポットでは、ハニーポット以外のサービスを提供していない。V. Yegneswaran らの研究 [6] によれば、未使用 IP アドレス空間に対するトラフィックは、攻撃対象のプローブやスキャン、ワームがほとんどである。

このため、本研究においてはハニーポットで収集したトラフィックを全て不正トラフィックと定義する。また、研究ネットワークで収集したトラフィックを、非不正トラフィックと定義する。

**ネットワークの構成**

ハニーポットのトラフィックと、比較対照データである研究ネットワークのトラフィックを収集した。

図 1 に実験に用いたネットワーク環境の概要を示す。

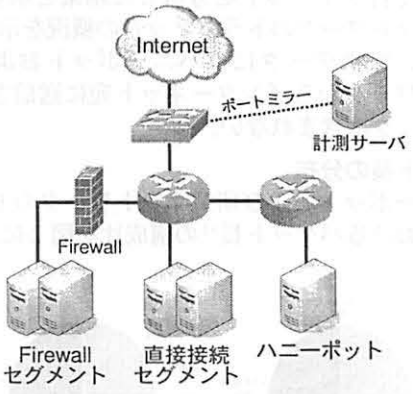


図 1　実験に用いたネットワーク環境のトポロジ

Firewall セグメントでは、Firewall 内部からインターネット方向へのコネクションの確立のみが許可されている。Firewall セグメントには、グローバル IPv4 アドレスが 320 個割り当てられている (ネットワークアドレスおよびブロードキャストアドレスを含む。以下のアドレスも同様)。

直接接続セグメントは、インターネットにそのまま接続されており、コネクションの確立に制限はない。直接接続セグメントに割り当てられている IP アドレス数は 384 個である。Firewall セグメントと直接接続セグメントをあわせて研究ネットワークと呼ぶ。

ハニーポット用セグメントは、ルータを介して直接接続セグメントに接続されている。このハニーポットに対しては、他の目的で利用されていない/16 の IP アドレス空間をルーティングし、当該プレフィックス宛の全トラフィックが転送されるよう設定した。

**トラフィックの収集**

研究ネットワークの上流に位置するネットワークの L2 スイッチにおいてポートミラーを設定し、計

表 2　収集データの概略

| | ハニーポット | 研究ネットワーク |
|---|---|---|
| アドレス数 | 44,284 | 701 |
| 受信パケット数 | 759,556 | 3,669,169 |
| 通信先アドレス数 | 7,245 | 7,936 |
| 転送トラフィック (KBytes) | 9,663 | 318,562 |
| 平均スループット (bit/s) | 21,989 | 724,909 |
| 平均スループット (pps) | 211 | 1,027 |

計測時間:2005 年 9 月 1 日 12 時から 13 時 (JST)

測サーバで実験ネットワークおよびハニーポット宛てのトラフィックを一括して収集した。なお、解析時には宛先 IP アドレスに基づきハニーポットのトラフィックと研究ネットワークのそれを分離した。

### IP ヘッダの各フィールドの傾向

低インタラクション型ハニーポット、および研究ネットワークに送信されたトラフィックの IP ヘッダについて各フィールドを分析した結果を示す。表 2 は両ネットワークのトラフィックの概況を示した。

なお、収集データにはハニーポットおよび研究ネットワークからインターネット宛に送信されるトラフィックは含まれない。

### パケット長の分布

ハニーポットおよび研究ネットワークのトラフィックにおけるパケット長*1の構成比を図 2 に示す。

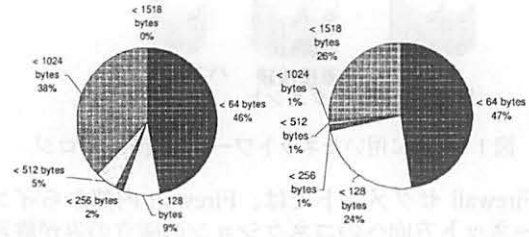


図 2　ハニーポット (左) および研究ネットワーク (右) のトラフィックにおけるパケット長の構成比率

研究ネットワークとハニーポットでは、パケット長の分布に違いが見られる。研究ネットワークでは 512〜1024 バイトのパケットが全体の 0.8% に過ぎないのに対し、ハニーポットでは 37.6% にも達する。これは、ワームのパケット長がこの範囲にある影響と考えられる。

一方、研究ネットワークでは 1024〜1518 バイトのパケットが全体の 25.7% を占めるのに対し、ハニーポットでは 0.8% にしか過ぎない。FTP やメールなどのデータの転送により、大きなパケットが多くなることが考えられる。

また、パケットごとの推定ホップ数*2の出現頻度を図 3 と図 4 に示す。これらのグラフからは、ハニーポットと通信したホストの推定ホップ数の分布と研究ネットワークの推定ホップ数の分布に違いが見られる。

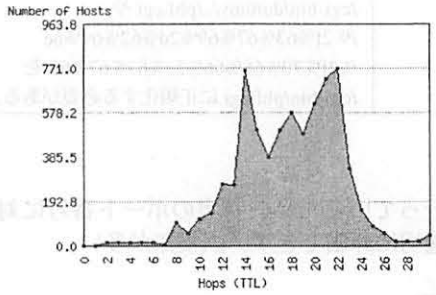


図 3　ハニーポットと通信したホストの推定ホップ数

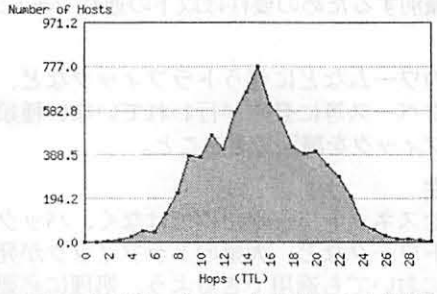


図 4　研究ネットワーク内ホストと通信したホストの推定ホップ数

この理由として、研究ネットワークは日本国内など、ホップ数の少ないホストとの通信が多いのに対し、ハニーポットでは研究ネットワークよりも国外などのホップ数の多いホストとの通信が多いことが推察される。

次に、各ネットワークにおけるトランスポート層プロトコルの構成比率を表 3 に示す。

この表からは、ハニーポットと研究ネットワークでは、トランスポート層プロトコルの構成比率が異なることがわかる。特に、ハニーポットのトラフィックには、研究ネットワークよりも UDP および ICMP パケットが含まれる可能性が高いといえる。

## 6　設計

4 節で定義した要件と 5 節で述べた実験の結果より、IP ヘッダのフィールドをパラメータとして、個々

*1 IP パケットにおけるデータグラム長

*2 受信した IP パケットの TTL と多くの OS の TCP/IP スタックのデフォルト TTL の値との差異からホップ数を推定する。TTL の値が 8 以下の場合には TTL-1、32 以下の場合には 32-TTL、64 以下の場合には 64-TTL、128 以下の場合には 128-TTL、129 以上の場合には 255-TTL の値を推定ホップ数とする。

表3　トランスポート層プロトコルの構成比率

| プロトコル種別 | ハニーポット | 研究ネットワーク |
|---|---|---|
| TCP | 82.91% | 97.94% |
| UDP | 5.90% | 1.10% |
| ICMP | 11.19% | 0.08% |
| IPIP | 0.00% | 0.42% |
| その他の IP | 0.00% | 0.45% |
| 合計 | 100.00% | 100.00% |

表4　設定したパラメータ

| ネットワーク層共通パラメータ | 推定ホップ数 (TTL) | | |
|---|---|---|---|
| | パケット長 | | |
| トランスポート層プロトコル別パラメータ | TCP | 宛先ポート番号 | TCP フラグ |
| | UDP | 宛先ポート番号 | |
| | ICMP | タイプ | コード |

の IP パケットの格付けを行う手法を提案する。

本手法では、IP パケットのヘッダからパラメータを抽出し、ベイズ推論によりトラフィックのプロファイルを作成する。そして、作成されたプロファイルに基づき、IP パケットの特性を判別する。

本手法では、IP ヘッダ自体を対象にスコアリングを行う。これにより、多くの計算機資源を消費するトラフィックデータの正規化が不要となる。

## 7　実装

本研究においては、Naive Bayesian の一方式である Gary Robinson-Fisher[2] 方式を用いた SPAM フィルタリングツールである bogofilter[1] 0.95.2 を元に "パケットマペット" をプロトタイプとして実装した。プロトタイプ実装では、リアルタイムでの識別ではなく、事前に保存したパケットに対して識別を行う。

パケットマペットでは、事前にハニーポットおよび研究ネットワークで取得したトラフィックを学習させる。学習は、IP パケットから表 4 のヘッダフィールドの値を抽出し、パラメータとして用いる。そしてヘッダフィールドを出現頻度データベースに登録し、ベイズ推論を用いてスコアリングを行う。

以下に bogofilter における Gary Robinson-Fisher 方式によるスコア計算式を示す。なお、$P$ はハニーポットから取得したパラメータ、$Q$ は研究ネットワークから取得したパラメータ、$S$ はスコアとする。

計算式ではパケットに対して事後確率を求めることにより、スコアを算出する。得られたスコアは、1.0 に近ければ不正トラフィックの可能性が高いことを示し、0.0 に近ければ非不正トラフィックの可能性が高いことを示す。

$$P = 1 - \sqrt[n]{(1-p_1) \times (1-p_2) \times \ldots \times (1-p_n)}$$

$$Q = 1 - \sqrt[n]{p_1 \times p_2 \times \ldots \times p_n}$$

$$P' = 1 - \chi^2(-2 \times \log P, 2n)$$

表5　収集データの概略

| | pgood | pbad | fw | U |
|---|---|---|---|---|
| プロトコル:TCP | 0.940 | 0.738 | 0.440 | − |
| TCP フラグ:ACK | 0.930 | 0.668 | 0.418 | |
| データ:40bytes | 0.539 | 0.666 | 0.553 | − |
| 19 ホップ | 0.013 | 0.027 | 0.683 | − |
| 宛先ポート:135 | 0.000 | 0.028 | 0.992 | + |
| 合計 | 0.008 | 0.992 | 0.992 | |

$$Q' = 1 - \chi^2(-2 \times \log Q, 2n)$$

$$S = \frac{(P' - Q')}{2}$$

($\chi^2$ は、カイ二乗検定を示す。)

表 5 に、ハニーポットで観測されたパケットに対するスコアリングの例を示す。なお、各項目の詳細は次の通りである。

pgood / pbad　当該パラメータを含むパケットが正常、ある異常なパケットである割合を示す。

fw　Robinson による加重インデックスの値。pgood と pbad を組み合わせ、正常であれば 0 へ、異常であれば 1 に近似した値となる。

U　"+" であればスコアを正に、"−" であれば負の方向に評価する

ネットワーク管理者は、どのスコア以上の不正トラフィックと判断するかを設定する。本機構はネットワーク管理者の設定に従い、スコアに対して格付けを行う。

## 8　評価

2005 年 9 月 1 日 12 時から 13 時までに取得したトラフィックのうち、ハニーポットのそれを不正トラフィック、研究ネットワークのそれを正常トラフィックとして学習させた。なお、日時は特に断りのない限り日本時間で表記した。

まず、学習に用いたネットワークのトラフィックを識別できているかを評価するため、学習データを取得した時間とは異なる時間を対象に、ハニーポットおよび研究ネットワークのトラフィックのスコアリングを行った。

評価対象データは、2005 年 9 月 1 日 13 時から 14 時の間に取得したハニーポットと研究ネットワークのトラフィックデータである。

図 5 では、ハニーポットのトラフィックの 78.4% は 0.9 以上のスコアを示している。一方、図 6 では、研究ネットワークのトラフィックの 96.8% は 0.1 未満のスコアを示している。

次に、既存手法との比較を行うため、IDS の検知結果と本手法によるスコアリング結果の比較を行った。

2005 年 9 月 8 日 1 時 59 分からの 10 時 59 分に研究ネットワーク宛に到達したトラフィックから、

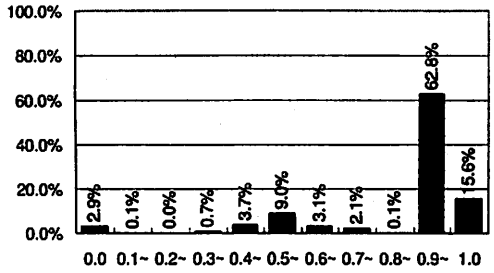


図5 ハニーポットのトラフィックのスコア分布

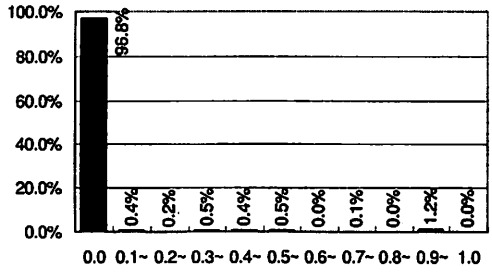


図6 研究ネットワークのトラフィックのスコア分布

NIDS がアタックとして判断したトラフィックを分離し、これに対して、プロトタイプ実装を用いてスコアリングを行った。

なお、NIDS にはミスユース型の IDS である Snort 2.3.3 を利用した。また Snort のシグネチャには、2005 年 9 月 8 日 0 時に Subscriber ライセンスで取得した Sourcefire VRT certified Rules を用いた。また、誤検知率の高い検出結果を除外するため、宛先ポート番号が 80 番および 161 番宛のトラフィックをスコアリング対象から除外した。

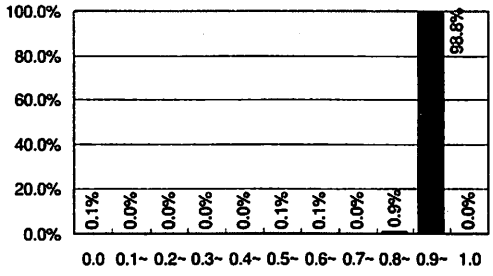


図7 IDS の検知したトラフィックのスコア分布

図 7 に IDS で検出されたトラフィックに対するスコア分布を示す。0.9 以上のスコアが 98.8% となっており、IDS の検知結果と類似したトラフィック識別結果を得ることができた。

## 9 今後の課題

プロトタイプ実装では、パラメータをトランスポート層プロトコル、パケット長、宛先ポート番号、推定ホップ数などに限定した。今後、スコアに対して各パラメータの寄与度を評価するとともに、スコアリングの精度を向上させるために、新たなパラメータの導入が課題となる。また、パケットのスコアリングに Gary Robinson-Fisher 方式を用いたが、不正トラフィックの識別に適した手法およびパラメータの検討が必要である。

さらに、本手法は IP パケットの格付けのみを行っているため、不正なトラフィックの拡散を防止できない。そのため、今後は本手法に適したトラフィックの制御手法の研究が必要である。

本手法は、IP ヘッダの各フィールドの構成比率が、ハニーポットと研究ネットワークのトラフィックが異なることに依存している。従って、トラフィックの構成に差異が少ない場合や、ヘッダの値が改ざんされた場合には、スコアリングの精度は低下する可能性が想定されるため、さらなる研究が求められる。

## 10 まとめ

本論文では、低インタラクション型ハニーポットのトラフィックと、研究ネットワークのトラフィックの IP ヘッダの傾向が異なることを明らかにした。

さらに、トランスポート層プロトコル、パケット長、宛先ポート番号、推定ホップ数などの IP ヘッダの内容を元にベイズ推論を行うことで、IP パケットに対してスコアを算出し、ネットワーク管理者の定義に基づき格付けを行う手法を提案した。また、評価の結果、本手法は既存の NIDS で検知した不正トラフィックを識別できることを明らかにした。

今後、格付けの結果をトラフィック制御の基準として活用するなどの応用が検討される。

## 参考文献

[1] Bogofilter. http://bogofilter.sourceforge.net/.

[2] Paul Graham. Better bayesian filtering, Jan 2003. http://www.paulgraham.com/better.html.

[3] Matthew V. Mahoney and Philip K. Chan. PHAD: Packet header anomaly. detection for identifying hostile network traffic, 2001. Florida. Tech. technical report 2001-04.

[4] Junichi Murakami. Dumnet. http://tf.rootkit.jp/work/dumnet/.

[5] Martin Roesch. Snort: Lightweight intrusion detection for networks. In *13th Systems Administration Conference (LISA'99)*, pp. 229–238. USENIX Associations, 1999.

[6] V. Yegneswaran, P. Barford, and D. Plonka. On the design and use of internet sinks for network abuse monitoring. In *Proceedings of Recent Advances in Intrusion Detection*, 2004.