4F-07　不正アクセス発信源追跡におけるパケット識別情報の有効性検証

渡辺 英俊　田中 俊介　早川 晃弘　松田 栄之
(株)NTT データ 情報科学研究所
e-mail: {hidetosi, shun, aki, matu}@rd.nttdata.co.jp

1. はじめに

インターネットにおける不正アクセスの発信源を追跡、特定するために、筆者らは、送信元 IP アドレスが偽造された場合でもその発信源を追跡可能な不正アクセス発信源追跡システムを提案している[1]。このシステムでは、追跡対象のパケットとそれ以外のパケットを識別するためにパケットフィーチャというパケット識別情報を利用している[2]。

本稿では、インターネットを流れるトラヒックを分析して、パケットフィーチャによる識別の有効性を検証する。

2. 不正アクセス発信源追跡の手順

筆者らが提案している不正アクセス発信源追跡システムでは、攻撃のターゲットとなり得るサイトに不正アクセスを検知する不正アクセスセンサ(以下、センサと呼ぶ)を設置し、不正アクセスの通過経路には不正アクセスパケットの追跡機能を搭載したルータなどの中継機器(以下、トレーサと呼ぶ)を設置する。

トレーサでは、中継したパケットの識別情報(パケットフィーチャ)と、そのパケットの直前の経路を特定する情報を関連付けて、一時的なバッファ(パケットバッファ)に保存しておく。センサは、不正アクセスを検知すると、追跡対象パケットのパケットフィーチャをトレーサに送る。追跡対象パケットは過去にそのトレーサが中継したものであるため、トレーサはセンサから取得したパケットフィーチャをもとに、パケットバッファ内から追跡対象パケットの直前の経路情報を取得できる。この操作を繰り返すことで、最終的に追跡対象パケットの発信源(以下、発信源という)まで、その中継経路を辿ることができる。

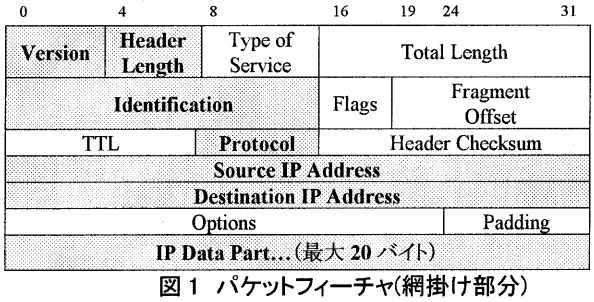

図 1　パケットフィーチャ(網掛け部分)

パケットフィーチャは図1の網掛け部分で構成されている。パケットフィーチャはパケットを一意に識別することが理想である。しかし、パケットバッファにおいて、追跡対象のパケットと同一のパケットフィーチャとなるような他のパケット(類似パケット)が存在した場合には、追跡すべき経路が複数存在することになり、発信源を一つに特定できなくなる。

そこで本稿では、トレーサ上で類似パケットが出現する可能性を検討し、パケットフィーチャによる識別の有効性を検証する。

3. パケットフィーチャの有効性

検証にあたり、発信源について以下の仮定を置くこととする。

(1) 発信源はパケットフィーチャの構造を知っている。
(2) 発信源はトレーサ内部を覗くことができず、類似パケットの送信元 IP アドレスを知ることができない。

パケットフィーチャを構成しているフィールドは、以下のように分類することができる。

(a) 固定値となるもの
Version, Header Length, Protocol, Checksum (TCP,UDP)が、これに該当する。また、Destination IP Address も、攻撃先のターゲットによって必然的に決定するので、これに該当する。
(b) 値の普遍的な傾向が不明なもの
仮定(2)より、Source IP Address がこれに該当する。
(c) 値の普遍的な傾向が分かるもの

誤追跡を引き起こす、つまり類似パケットが発生し易いようなパケットフィーチャを考えると、発信源は(c)分類のフィールドについて、最も出現しやすい値を代入したパケットを送出することで、類似パケットの出現確率を上げることができる。

4. 実データによるパケットの検証

WIDE ネットワーク上のある地点におけるトラヒック[3]を TCP,UDP それぞれについて分析し、Identification, Source Port, Destination Port のフィールドについて、

An efficacy verification of packet identifier for unauthorized access tracing system.
Hidetoshi WATANABE, Shunsuke TANAKA,
Akihiro HAYAKAWA, Shigeyuki MATSUDA,
†Laboratory for Information Technology,
NTT DATA CORPORATION

出現する値の傾向を調べた。曜日による偏りを考慮して、2000年6月1日、7日、13日、19日の4日間におけるトラヒックを用いた。これらのトラヒックを合算し、各フィールドについて最も多く出現した値(最頻値)と、2番目に多く出現した値(第2最頻値)、それらの出現率を表1、2に集計した。このトラヒックは合計2時間47分間のパケットであり、TCPは6,330,958パケット、UDPは1,958,431パケットあった。

表1、2より、UDPのIdentificationは、最頻値が特殊な値ではなく、かつ最頻値と第2最頻値の間に著しい度数の相違もないため、(b)に分類できる。それ以外のフィールドの最頻値はいずれも特殊な値で、かつ第2最頻値は特に意味のある値ではない。したがって、(c)に分類できる。

以上の結果から、類似パケットが出現しやすいパケットは表3のようになる。TCPパケットでは、Source PortとDestination Portが共に80となるパケットはないため、ここではDestination Portを80とし、Source Portは不定とした(*部分)。

5. 類似パケットの出現確率

表3の各フィールドについて、単独項目の値の一致確率を以下のように定義する。

表1 パケットフィーチャのフィールドの最頻値(TCP)

	最頻値	出現率	第2最頻値	出現率
Ident.	0	1.7e-4	35594	8.0e-5
Src.Port	80	7.1e-2	6346	3.6e-2
Dst.Port	80	6.3e-1	6346	8.6e-2

表2 パケットフィーチャのフィールドの最頻値(UDP)

	最頻値	出現率	第2最頻値	出現率
Ident.	46115	3.9e-4	46118	3.9e-4
Src.Port	53	6.3e-1	41831	9.0e-2
Dst.Port	53	6.4e-1	36434	3.2e-2

表3 類似パケットが多いパケットフィーチャ

フィールド名	TCP	分類	UDP	分類
Version	4	(a)	4	(a)
Header Length	5	(a)	5	(a)
Identification	0	(c)	不定	(b)
Src. IP Address	不定	(b)	不定	(b)
Dst. IP Address	固定	(a)	固定	(a)
Src. Port	*不定	(b)	53	(c)
Dst. Port	*80	(c)	53	(c)

- (a)分類のフィールド

Destination IP Addressは、通過パケットの半分が攻撃ターゲットのホスト宛と仮定して確率は0.5とし、それ以外のフィールドでは確率は全て1とする。

- (b)分類のフィールド

Identification(UDP)の確率は$1/2^{16}$、Source IP Addressの確率は有効なアドレスを全体の1割と仮定して$1/(2^{32}\times 0.1)$、Source Port(TCP)の確率はハイポート(1024～65535)のみを考えて$1/(2^{16}-1024)$とする。

- (c)分類のフィールド

3章の集計で得られた各最頻値の出現率をその確率とする。

以上の検討結果より、あるパケットフィーチャがあって、パケットバッファから無作為に1パケット取り出した時に、それが類似パケットである確率rは上記の各確率の積となり、これはTCPで1.9e-19、UDPでは7.2e-15となる。

また、トレーサのパケットバッファの中に過去10分間のパケットフィーチャが格納されていると仮定し、そのパケット数をnとする。n個のパケットの中に少なくとも1つの類似パケットが出現する確率(誤認識率)は、$0<r\ll 1$でnが非常に大きい場合、

$$1-(1-r)^n \fallingdotseq n\times r$$

と近似できる。したがって、本分析に使用したトラヒック上で10分間のパケット(パケット数はTCPが平均379,288、UDPが平均117,330)に対する誤認識率は、TCPでは7.4e-13、UDPでは8.4e-10といずれも非常に小さい値となる。

6. まとめ

パケットフィーチャによる追跡は類似パケットがほとんど出現せず、追跡の精度が非常に高いことが示せた。今後は、プロトタイプシステムを開発して追跡の精度等の評価を行なう予定である。

謝辞

本研究は、通信・放送機構(TAO)の委託研究テーマ「不正アクセス発信発信源追跡技術に関する研究開発」の一環として行われているものである。

参考文献

[1] 竹爪他,"不正アクセス発信源追跡アーキテクチャの一検討",情処60全大, 6Q-6, March 2000.

[2] 渡辺他,"不正アクセス発信源追跡のためのパケット識別情報の検討", 情処60全大, 6Q-7, March 2000.

[3] "WIDE バックボーントラフィックの収集と解析", WIDE プロジェクト MAWI ワーキンググループ, http://www.wide.ad.jp/wg/mawi/