6Q-07　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

渡辺 英俊† 馬場 達也† 竹爪 慎治‡ 松田 栄之†
(株)NTT データ †情報科学研究所 ‡公共地域ビジネス事業本部
e-mail: †{hidetosi, baba, matu}@rd.nttdata.co.jp, ‡taketsumes@noanet.nttdata.co.jp

1. はじめに

近年、インターネットは電子商取引等のインフラとしてますます重要性が増しており、インターネットにおける不正アクセスの発信源を追跡、特定することが強く求められている。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追跡が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]。

本稿では、不正アクセス発信源追跡システムにおいて、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必要な情報およびその識別方式を提案する。

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)を置き、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、中継したパケットとその直前経路のトレーサ(これを隣接ノードと呼ぶ)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、不正アクセスを発見すると、追跡対象の不正アクセスパケットに関する情報(識別情報)をトレーサに送る。追跡対象のパケットは過去にそのトレーサが中継したものであるから、トレーサは識別情報をもとに、パケットバッファ内から追跡対象パケットの直前の隣接ノード情報を検索することができる。同様の操作を隣接ノードに対して繰り返すことで、最終的に追跡対象パケットの発信源まで、その中継経路を辿ることができる。本手順で、追跡対象パケットを識別することをパケット識別と呼ぶ。

3. パケット識別方式の要件

トレーサでの実装を考慮すると、上述したパケット識別には以下の課題が存在する。

(1) トレーサには使用可能なリソースの制約が多いため、パケットバッファの消費量が少なく、また保存処理も軽いものであること。

(2) パケットバッファからパケットを検索した時に、追跡対象パケットが検索結果に必ず含まれること。

(3) パケットバッファからパケットを検索した時に、追跡対象の送信元以外から送られたパケットが検索結果に含まれないこと。

このような要件を満たすパケット識別方式としては、以下のものが考えられる。

① トレーサはパケットバッファに中継パケットのハッシュ値を保存し、センサは不正アクセスパケットのハッシュ値を生成する。両者のハッシュ値を使ってトレーサがパケット識別を行なう。

② トレーサは中継するパケットに ID を埋め込んで中継する。パケットバッファには中継したパケットではなく、ID を保存する。センサは不正アクセスパケットに含まれている ID を取り出す。両者の ID を使ってトレーサがパケット識別を行なう。

しかし、方式①では、トレーサは絶えずハッシュ値を計算しなければならないため、トレーサの負荷が高くなるという問題がある。また、方式②では、ID を埋め込む操作は IP で規定されていないため、現行プロトコルとの適合性に問題が生ずるおそれがある。そこで本稿では、パケットが持つ情報そのものに着目してパケットを識別する方式を提案する。

4. パケットフィーチャを用いた識別方式

まず、トレーサにおいて、中継するパケットからその特徴を表わす情報を抽出し、これをパケットバッファに保存する(図 1 の(1))。この特徴情報をパケットフィーチャと呼ぶ。また、センサでも不正アクセスパケットから同様の情報を生成する。この両者を使ってトレーサがパケットを識別する(同(2))。

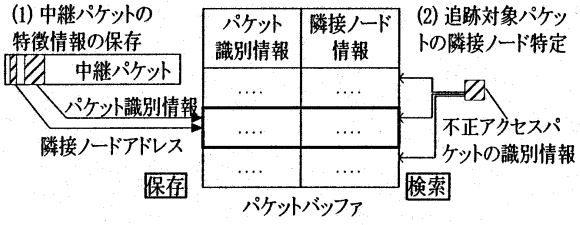

図 1 パケットフィーチャを用いた識別方式

A study of packet identifier for unauthorized access tracing system
Hidetoshi WATANABE†, Tatsuya BABA†,
Shinji TAKETSUME‡, Shigeyuki MATSUDA†,
†Laboratory for Information Technology,
‡Public Administration Community Business Sector,
NTT DATA CORPORATION

この方式の採用により、トレーサの負荷や現行プロトコルとの適合性を考慮したパケット識別が可能となる。

4.1. パケットフィーチャ

IP パケット[2]を IP ヘッダの部分とそれ以外の IP データ部に分けて、パケットフィーチャの内容を説明する。

IP ヘッダ内の Version, Header Length, Ident, Protocol, Source IP Address, Destination IP Address フィールドは、送信元から宛先まで中継される間、値が不変であり、要件(2)を満たしているのでパケットフィーチャに採用できる。特に Ident(16 ビット)は、送信元がパケットを送出する都度異なる値を付与するため、これをパケットの識別に用いれば高い識別精度が期待できる(要件(3))。

IP データ部では、先頭 20 バイト(ただし、IP データ部が 20 バイト未満の場合はその長さまで)をパケットフィーチャとする。TCP パケットの場合、この部分はオプション部を除いた TCP ヘッダ全体となる。この中には、Sequence Number, Acknowledgement Number フィールド(各 16 ビット)がある。これは、パケットごとに異なる値を取る可能性が高く、識別精度が向上が期待できる。(要件(3))。一方 UDP, ICMP パケットの場合は、ヘッダに相当する部分に Checksum フィールド(各 16 ビット)が存在する。これも、異なるパケット間では異なる値を取る可能性が高く、要件(3)を満たすと考えられる。このようにトランスポート層のプロトコルにかかわらず、先頭 20 バイトをパケットフィーチャとすることにより、トレーサでのパケットフィーチャの生成が簡便化できる。

以上で述べたパケットフィーチャを図 2 の網掛部で示す。また、図 3 に、IP パケットに対するパケットフィーチャの大きさの比率を示す。例えば IP パケット長が 160 バイト以上の場合、この比率は 20%以下である。このように、パケットをそのまま保存する方式に比べて、パケットバッファの消費量を抑えることができる。

4.2. 識別の仕組み

上述のパケットフィーチャを用いた識別方式について説明する。

不正アクセスセンサから追跡対象のパケットに関するパケットフィーチャを受け取った時に、トレーサがパケットバッファから検索する際の比較条件を以下に示す。

a. パケットフィーチャの IP ヘッダ部分が両者で完全に一致していること。
b. パケットフィーチャの IP データ部分が両者で完全に一致していること。ただしデータ部の長さが異なる場合は、いずれか短い方までを比較対象とする。

比較条件 b は、IP パケットが中継途中で複数の IP パケットに分断されるフラグメントという現象を考慮したものである。フラグメントが起きた場合は、第 1 フラグメントの IP パケットは、IP データ部の後部が切り落とされるが前部は変化しない。そこで、センサとトレーサの双方で第 2 フラグメント以降のパケットを無視し、パケットフィーチャの IP データ部における比較では短い方まで比較することとする。

不正アクセスパケットの発信源が、送信元 IP アドレスを偽造し、かつ偽った先のホストが送信している通常のパケットと同一の Ident(IP ヘッダ)、Sequence Number(TCP ヘッダ)、Checksum(UDP,ICMP ヘッダ)を不正アクセスパケットに付与すれば、トレーサのパケット識別を意図的に誤らせることができる。しかし、不正アクセスパケットの発信源がこれらの値を予測することは非常に困難である。したがって、本章で述べた識別方式は、送信元 IP アドレスが偽造されていない場合はもちろん、偽造された場合でも高精度なパケット識別が可能である。

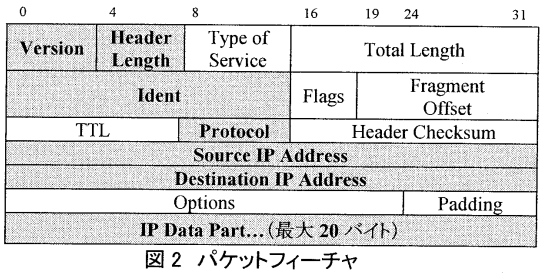

図 2 パケットフィーチャ

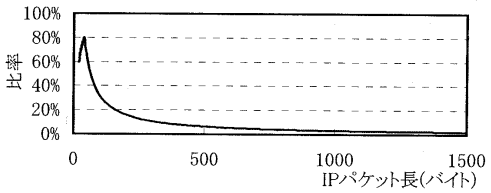

図 3 IP パケットとパケットフィーチャの比率

5. まとめ

本稿では、不正アクセス発信源追跡システムにおいて、パケットの特徴情報を利用して追跡対象パケットを識別する方式を提案した。

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-06, March 2000.
[2] Douglas E. Comer, "Internetworking with TCP/IP", Prentice Hall, 1988.