社団法人 情報処理学会 研究報告
IPSJ SIG Technical Report

2009－ＤＰＳ－138 (27)
2009－CSEC－ 44 (27)
2009／3／6

アプリケーションレイヤ由来の情報を用いたトレースバック手法の設計と実装

井澤 志充　　大島 龍之介　　国峯 泰裕

株式会社クルウィット
{izawa, ryu, kunimine}@clwit.co.jp

あらまし　能動的な警戒手段としてのトレースバックプラットフォームの実現に向けてアプリケーションレイヤ由来情報を用いたトレースバックアルゴリズムを提案する。
まず本研究で提案するアプリケーショントレースバックの基本手法について述べ、次に対象とするアプリケーションの一例としてDNS反射攻撃を例にその手法について述べる。
最後にIPトレースバックとの連携を例示しISP環境でのトレースバックプラットフォームの構築例をあげその実用性を示す。

Design and implementation of Trace Back system using derived from application layer's information

Yukimitsu Izawa,　Ryunosuke Ohshima,　Yasuhiro Kunimine

Clwit Inc.
{izawa, ryu, kunimine}@clwit.co.jp

Abstract　We propose Technique of Trace Back system using derived from application Layer's information to build actively-actuated Trace Back platform. We summarized some scheme to Trace Back Internet traffic. Next, we describe about our application Trace Back framework and DNS reflection trace back.

1. はじめに

近年、インターネットの普及およびネットワーク技術の発展によりインターネットは社会インフラとしての役割を担うようになった。それに伴い、DoS(Denial of Serve)やDDoS(Distributed Dos)、ウイルス発信などのサイバー攻撃は、社会に与える影響を増大させている。

これらサイバー攻撃に対し攻撃元探査を行うことで攻撃元がどのホストであるか、あるいは複数あるネットワーク境界のどれから攻撃パケットが流入してきているのかを明らかにする手法であるトレースバック技術のうち、アプリケーションレイヤ由来の情報を利用する手法について述べる。

本研究では、能動的な警戒手段としてのトレースバックプラットフォームの実現に向けてアプリケーションレイヤ由来情報を用いたトレースバックアルゴリズムを提案する。

まず本研究で提案するアプリケーショントレースバックの基本手法について述べ、次に対象とアプリケーションとしてDNS反射攻撃を例にその手法について述べる。

2. 先行・関連研究調査

アプリケーショントレースバックアルゴリズムを開発するにあたって、まず、トレースバックの先行研究分野であるIPトレースバックについて文献調査を行った。また、関連分野である踏み台検出手法について文献調査を行い、その手法や特徴についてまとめる。

2.1 IP トレースバック

まず、関連研究であるIPトレースバックの概要と長所・短所について整理する。

インターネット上の各通信ノードには郵便における住所に相当するIPアドレスが割り宛てられており、通信に用いられる。しかし、IP パケットの発 IP アドレスに偽のアドレスが書かれていた場合には、発信ノードを特定することができない。このようにIPパケットに偽の発IPアドレスを書き込む行為はIP詐称(IP spoof)と呼ばれている。IP詐称を行うと、着信ノードからの返信を受けとることができないため、発信ノードから着信ノードへの単方向の通信しか行えない。しかし、発信ノードの身元を隠蔽することができるため、DoS 攻撃やアイドルスキャンに悪用されることがある。

IPトレースバックは、IP詐称されたパケットの通信経路を追跡することを目的とした技術である。数多くの先行研究がなされており、逆探知パケット方式[1][2]、マーキング方式[3]、ハッシュベース方式[4]などの手法が提案

されている。

どの手法もネットワーク上に観測点を配置して、パケットの通過を検出することによりトレースバックを可能とするものである。それぞれの手法では、あるパケットがある観測点を通過したことを記録/通知する方法に違いがある。各手法の対比を表に示す。

表 1 各トレースバック手法の対比

	逆探知パケット方式	マーキング方式	ハッシュベース方式
パケットへの書き込み	なし	あり	なし
単一パケットトレース	不可能	可能	可能
観測点との通信	なし(通信パケットを利用)	サンプリングごとに発生	トレースごとに発生

本研究で提案するアプリケーショントレースバック手法でもIPトレースバック手法と同様に観測点を配置した構成が必要になると思われる。観測点との通信方式としては、通信パケットへの変更を行う方式やサンプリングごとに通信が発生する方式はデプロイ面で問題がある。ハッシュベース方式のようにトレース時のみ通信を行う方式が導入しやすいものと思われる。

2.2 アプリケーショントレースバック

IP トレースバックはネットワーク層レベルでの経路追跡を目的とした技術であるのに対し、アプリケーショントレースバックは、トランスポート層以上の通信を追跡するための手法をさす造語であり、IP トレースバックのみでは追跡不能なアプリケーションレベルでの通信の追跡を目的として本研究で開発する技術である。関連する先行研究としては、踏み台検出(Stepping-stone detection)がある。

2.3 踏み台検出

IP トレースバックのみでは追跡不能な「踏み台ホスト」を用いた通信の追跡を行う手法として、踏み台検出(stepping-stone detection)が研究されている。踏み台検出とは、複数の中継ホスト(踏み台ホスト)を介した通信の追跡を目的としたものであり、各ホストの入セッションと出セッションの対応付けを行うことにより、順次踏み台ホストを辿ってゆく方法で、通信の経路を追跡する手法である。

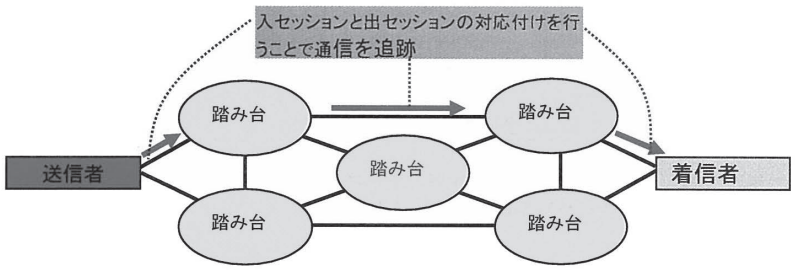

図 1 踏み台検出手法の概要

提案システムを設計するうえで考慮すべき点を次にまとめた。

ネットワークベース手法が有利

ISP のような大規模なネットワークにおいては、多種多数のホストが存在しているため、ホストベースよりもネットワークベースの方がデプロイ面で有利である。提案システムはネットワークベース手法とするのが良い。精度向上等の目的で、ホストベース的な手法を用いる場合には、末端クライアントではなく、ファイアウォール等の中継ノードや、Web、Mail などの主要サーバに限定するべきである。

統計的手法のみに頼らない

統計的手法のみを用いた手法はセッション対応付けの精度が低く実用に難がある。提案システムでは、統計的手法のみに頼ったアルゴリズムは避けるべきである。

セッション切断後のトレースができるものとする

手法によっては、トレース対象のセッションが張られた状態でのみトレースが可能なものもあった。インシデントハンドリングでの応用を考慮すると、セッションが切断された後であってもトレースが可能であるものが望ましい。踏み台検出方式の多くは、Telnet や FTP などの NVT(Network Virtual Terminal)による対話的プロトコルによる踏み台通信の検出を対象にしているものであった。

3. 対象アプリケーションの検討

日本のインターネットにおける総被害のほとんどは、ウイルスによるものである。ウイルスの被害や検出数が増加している。ウイルス感染経路は、MSBlaster が爆発的に蔓延した 2001 年 8 月を除けば、概ね 95%以上はメール経由である。それゆえにウイルスの最大の感染ルートであるメールに焦点を絞り、メールからの感染を確実に追跡できるならば、ウイルス感染の大部分を減少させる可能性をもち、さらにウイルス作成者を検挙に導くことも期待できる。本研究ではウイルスメールに対応したアプリケーショントレースバックアルゴリズムを実装し、研究室レベルながら高い確率でウイルスメールの分別と踏み台ホストの推定を行うことができた。しかしながら並行して行っていた、アプリケーショントレースバックの法的な要件検討の結果、メールのアプリケーショントレースバックを実環境へ適用するには、社会的な認識やインターネット運用と法律との成熟度の関係より時期尚早との結論に至った。そこで、近年脚光を浴びている、インターネットの基幹をなす DNS サーバへの攻撃として DNS 増幅攻撃に着目した。本研究では対象アプリケーション DNS を選択し、DNS 増幅攻撃の追跡を行う手法についてのアプリケーショントレースバックアルゴリズムを提案する。

4. フレームワーク

これまでの検討をベースに、アプリケーショントレースを実現するためのアルゴリズムとして、ハッシュベースのアプリケーショントレース手法を提案する。

本手法はハッシュベースの IP トレースバックを参考に考案したものであり､次の手順で通信経路の追跡を行う。

通信経路となるネットワークに観測点を配置する。各観測点では、トレース対象アプリケーションの通信パケットから、通信セッションを特定するための特徴情報を抽出し、そのハッシュ値を算出してデータベースに記録する。セッション観測の補助手段として、通信経路となるルータやサーバ、ホストなどからのログ情報を利用することも検討する。必要に応じてログ情報からセッションを特定するための特徴情報を抽出し、そのハッシュ値を算出してデータベースに記録する。

トレース実行時には、まず、トレース対象通信の特徴情報を抽出し、そのハッシュ値を算出する。このハッシュ値を持つ通信の通過/非通過を各センサに問い合わせて回答を得る。通過した観測点をつなぎあわせることで、当該セッションの通信経路を特定する。

4.1 機能構成

アプリケーショントレースを実現する APTB システムは、APTB プローブ, APTB コリレータ, APTB コントーラの各機能要素で構成される。

APTB プローブは、通信経路となるネットワークを観測してトレース対象アプリケーションの通信に関する情報を取り出し、ログを生成する機能である。各 AS やサイトの境界ネットワークに設置し、入セッションと出セッションをそれぞれ監視する。

APTB コリレータは、APTB プローブや各種サーバ、ファイアウォール、ルータ,IDS 等から受信したログ情報からセッションを特定するための特徴情報を抽出する機能である。各サイトや AS ごとに配置する。どの情報源からログを収集するのかについては、トレース対象とするアプリケーションごとに検討する必要がある。抽出した特徴情報から、入セッションと出セッションの相関を取ることにより、サイトや AS 単位での踏み台検出を行う。抽出した特徴情報は、ハッシュアルゴリズムにより匿名化したうえでデータベースに格納する。このデータベースを用いて APTB コントローラからのトレース要求に対して、当該セッションの通過/非通過を返答する。

APTB コントローラトレースバック管理者や、上位トレースバックシステムからのアプリケーショントレース要求を受け付けて、各 APTB コリレータにトレース要求を発行し、APTB コリレータからの返答を集約して、トレース結果を提示する機能である。

5. DNS トレースアルゴリズム

DNS トレースアルゴリズムが対象とする攻撃を DNS 反射攻撃あるいはその一種である DNS 増幅攻撃である。DNS 反射攻撃について図 3 に示す。

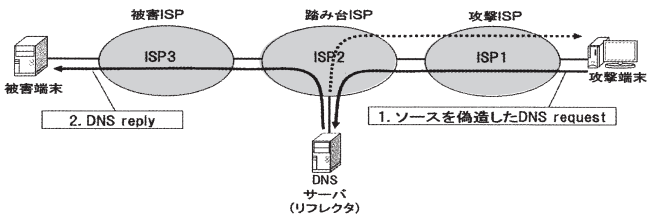

図 3　DNS 反射攻撃

DNS 反射攻撃は、攻撃者がソース IP アドレスを偽造した DNS クエリを DNS サーバ(リフレクタ)に送信することで、DNS リプライを攻撃用のパケットとして利用する攻撃である。

この攻撃の特徴は、直接の攻撃用パケットは正規の DNS サーバ(リフレクタ)によって送信されているという点である。つまり被害端末側から IP アドレスベースでのトレースバックを試みても、リフレクタまでしかたどれず、真の攻撃者(この場合はソースを偽造した DNS クエリを送信した端末)にはたどり着けない。

ここで、中継 DNS サーバから被害 DNS サーバへの DNS 通信を惹起した DNS 通信が特定できればさらに発信元端末を追跡することができる。言い換えると、DNS 反射攻撃の原因になった受信 DNS 通信(1 次 DNS 通信)と、中継 DNS サーバから送信された DNS 通信(2 次 DNS 通信)を同定できれば、中継 DNS サーバを踏み台にした DNS 反射攻撃の経路を追跡することが可能である。

そこで、本課題では、この 1 次 DNS 通信と 2 次 DNS 通信を同定することにより、中継 DNS サーバを踏み台にした DNS 反射攻撃を追跡する手法について検討した。

5.1 アルゴリズム開発方針

外部から受信した DNS 反射攻撃の原因となる DNS 通信(ここでは 1 次 DNS 通信と呼ぶ)と 1 次 DNS 通信より中継 DNS サーバから送信される DNS 通信(ここでは 2 次 DNS 通信と呼ぶ)を対応付けるアルゴリズムについて検討した。

1 次 DNS 通信と 2 次 DNS 通信は、違う内容になる。そのため、単純一致で同定することはできない。しかし、1 次 DNS 通信と 2 次 DNS 通信に共通の特徴を持っている。

本研究では、DNS 通信に含まれる特徴情報のうち、変化されない特徴や、変化が限定的な特徴を抽出し、これを用いて 1 次 DNS 通信と 2 次 DNS 通信の同定を行なうものとした。

5.2 開発方針

本研究で開発するアルゴリズムの開発方針を次のように定めた。

1. 高速処理可能なアルゴリズムを目指す
   - 行単位での逐次処理で抽出できる特徴を使用する。
     - DNS 通信全体を何度もスキャンするような処理は使用しない。
2. できる限り簡素なものとする
   - 簡素で拡張可能なものとする

5.3 DNS 通信特徴の検討

DNS 通信について、DNS 反射攻撃の原因となる 1 次 DNS 通信と DNS 中継サーバから送信される 2 次 DNS 通信を同定するための特徴について検討した。特徴情報としては、DNS 通信中の記述のうち、候補として次の情報を選出した。

1. DNS の問い合わせ内容

   RFC1034 で規定されている DNS の問い合わせ(Queries)の質問(Question)セクションは、1 次 DNS 通信から 2 次 DNS 通信で同一となる。これを特徴情報として使用することができると考えられる。

2. DNS の再帰問い合わせ要求

   RFC1035 で規定されている DNS のヘッダの再帰要求(Recursion Desired, RD)ビットは、1 次 DNS 通信ではセットされ、2 次 DNS 通信ではセットされていない事になる。これを DNS 通信の特徴情報として使用できると考えられる。

3. IP アドレス

   1 次 DNS 通信の送信先 IP アドレスと、2 次 DNS 通信の送信元 IP アドレスは同一となる。これを特徴情報として使用できると考えられる。

5.4 非 DNS 反射攻撃である DNS 通信との弁別に用いる情報

DNS 本課題で提案する APTB システムは DNS 反射攻撃ではない DNS 通信のトレースを目的としていない。そのため非 DNS 反射攻撃である DNS 通信と DNS 通信の高度な弁別は必要がない。不必要な処理を減らすために簡単な弁別のみを行なうものとした。非 DNS 通信との弁別には次の特徴を用いるものとする。

1. DNS 要求の有無

   DNS 反射攻撃は、DNS の問い合わせ(DNS ヘッダの QR ビットがセットされている)通信によって構成される。よって DNS の回答(DNS ヘッダの QR ビットがセットされていない)通信は DNS 反射攻撃ではない。

2. DNS の問い合わせの質問セクションの有無

DNS 反射攻撃は、DNS の問い合わせ(Queries)の質問(Question)セクションが存在する通信によって構成される。よって DNS の問い合わせの質問セクションが無い DNS 通信は DNS 反射攻撃ではない。

5.5 提案アルゴズムのまとめ

これまでの検討で導出された、DNS 通信同定のための特徴抽出アルゴリズムについて、下記にまとめた。
DNS 通信同定のための特徴抽出アルゴリズム
使用する特徴セット

1. DNS 通信同定に使用する特徴
   次の特徴を用いて DNS 通信の同定を行なう。
   1. DNS の問い合わせ内容
   2. DNS の再帰問い合わせ要求
   3. IP アドレス
2. DNS 反射攻撃ではない DNS 通信との弁別に用いる特徴
   DNS 反射攻撃ではない DNS 通信は、検索対象にならないことを想定しているため、高機能な弁別は行なわない。簡易な弁別のために次の特徴を用いる。
   1. DNS 要求の有無
   2. DNS の問い合わせの質問セクションの有無

6. IPTB と APTB 連携について

実際のネットワークにおけるインシデントのトレースバックは、本稿で提案したアプリケーショントレースバックのみでは完結せず、IP トレースバックとの相補的な組み合わせによって、より現実的なトレースバック機構を提供すると考える。

IP-TB と AP-TB が連携し、踏み台攻撃を追跡している様子が図 4 である。

図中に登場する攻撃 ISP、踏み台 ISP、被害 ISP ともにトレースバックシステムを導入していることが前提である。この TB システムは自らが所属する ISP の内部サーバの近傍に設置され、随時そのパケットのログを収集する。

攻撃が発生した場合、被害者は自らが所属する ISP の管理者に報告し、対処を依頼する。

対処を依頼された被害 ISP の管理者は、TB 管理センターに攻撃の被害報告を行い、その追跡を依頼する。

TB 管理センターでは、発生した攻撃に関してトラブルチケットベースで管理し、各 ISP 内の TB システムに問い合わせを行う。該当する攻撃を検知している TB システムはその旨を回答する。

TB 管理センターはその情報を集約し、攻撃 ISP の管理者にその旨を連絡し個別に攻撃者への対処を依頼する。

これら TB 管理センターとの情報のやり取りはすべてハッシュ化されたパケットの情報をベースに行い、通信の秘密に抵触しない様おこなわれる。

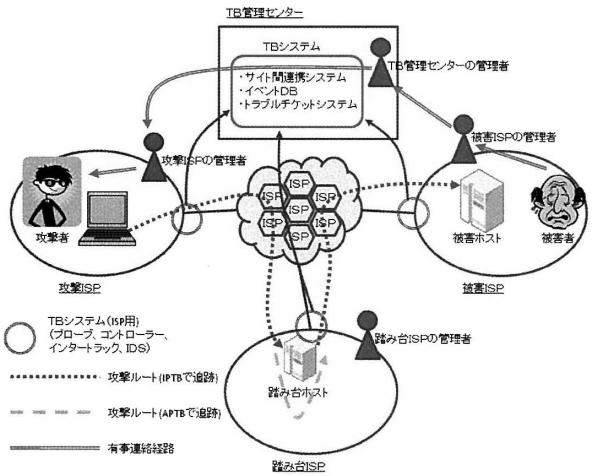

図４　APTB-IPTB 連携図

7. まとめ

アプリケーショントレースアルゴリズムを開発するにあたりまず、関連研究に関する文献調査を行い関連分野であるIPトレースバックと踏み台検出手法についてまとめた。何れも研究段階の技術であることや、そのままの形では、現実的な問題を解決する手段として応用しづらいことを把握した。

これらの調査結果をふまえて、踏み台ホストを悪用した不正行為の抑止と、インシデント対応の支援を目的としたアプリケーショントレースバックアルゴリズムについて検討し、ハッシュベースAPTB方式を提案した。提案システムでは、ネットワークベースとホストベースを組み合わせた方式を採用することで、ネットワークベース手法の特長であるデプロイ性の良さと、ホストベース手法の特長である精度の高さを持ち合わせた方式とした。提案手法の応用として、DNS反射攻撃の経路追跡を行うシステムの構成について考察した。
さらにDNS反射攻撃を例にAPTBとIPTBとの連携方法について述べた。

今後は、APTBとIPTBの連携の場を実環境に移しその有用性を確認するとともに、社会に対して有効な適用方法について検討をすすめる。

謝辞

本研究は、独立行政法人情報通信研究機構の平成17年度からの研究案件「インターネットにおけるトレースバック技術に関する研究開発」の一部である。

参考文献

[1] Steve Bellovin, Marcus Leech, Tom Taylor: ICMP Traceback Messages, Internet draft. Document: draft-IETF-iTrace-04.txt, (2003).

[2] 甲斐, 中谷, 清水, 塚本, "不正アクセスに対する高性能発信源探査方式の提案", 情報通信研究季報, Vol51, pp.41-49, (2005).

[3] S. Savage, D.Wetherall, A. Kerlin, T. Anderson, "Practical network support for IP traceback", Proc. of ACM 2000 SIGCOMM Conference, pp295-306, (2000).

[4] A. C. Snoeren, C. Partridge, L. A. Sanchez, C. E. Jones, F.Tchakountio, S. T. Kent and W. T. Strayer, "Hash-based IP Traceback", In Proceedings of SIGCOMM '01 (2001).

[5] CARRIER, B., AND SHIELDS, C. A recursive session token protocol for use in computer forensics and tcp traceback. In Proc. IEEE Infocom '02 (June 2002), (2002).

[6] JUNG, H. T., KIM, H. L., SEO, Y. M., CHOE, G., MIN, S. L., AND KIM, C. S. ,"Caller identification system in the internet environment", In Proc. USENIX Security Symposium '93 (Oct. 1993), (1993).

[7] YODA, K., AND ETOH, H., "Finding a connection chain for tracing intruders", In Proc. European Symposium on Research in Computer Security (Oct. 2000), pp. 191--205, (2000).

[8] WANG, X., REEVES, D. S., AND WU, S. F.,"Inter-packet delay based correlation for tracing encrypted connections through stepping stones", In Proc. European Symposium on Research in Computer Security (Oct. 2002), pp. 244--263, (2002).

[9] DONOHO, D. L., FLESIA, A. G., SHANKAR, U., PAXSON, V., COIT, J., AND STANIFORD, S., "Multiscale stepping-stone detection: Detecting pairs of jittered interactive streams by exploiting maximum tolerable delay", In Proc. International Symposium on Recent Advances in Intrusion Detection (Oct. 2002), pp. 17--35, (2002).

[10] STANIFORD-CHEN, S., AND HEBERLEIN, L. T.,"Holding intruders accountable on the internet", In Proc. IEEE Symposium on Security, (1995).

[11] WANG, X., REEVES, D. S., WU, S. F., AND YUILL, J., "Sleepy watermark tracing: An active network-based intrusion response framework", In Proc. International Conference on Information Security (June 2001), pp. 369--384, (2001).