パケットマーキングとロギングを用いた サイバー攻撃に対するトレースバック

李 鵬飛†, フォン ヤオカイ‡, 川本 淳平‡, 櫻井 幸一‡
†九州大学大学院システム情報科学府
‡九州大学大学院システム情報科学研究院
819-0395 福岡市西区元岡 744 番地
lipengfei@itslab.inf.kyushu-u.ac.jp, fengyk@ait.kyushu-u.ac.jp,
kawamoto@inf.kyushu-u.ac.jp, sakurai@inf.kyushu-u.ac.jp

あらまし インターネット技術の進歩と共にサイバー攻撃事件が多発している．サイバー攻撃への対策として，攻撃の検知・遮断技術と共に，攻撃者までのトレースバック（IP アドレスの同定）技術が不可欠である．既存のトレースバック技術には，特定の攻撃のみを対象としているおよび通信経路再構成の計算量が大きくトレースバックに時間を要するという問題がある．本研究では新しいトレースバック手法を提案する．シミュレーションにより作成したデータと既存研究で用いられている実験データを利用して，提案手法の性能を実証する．
キーワード：サイバー攻撃，追跡，トレースバック，パケットマーキング，ロギング

A Proposal for Cyber Attack Trace-back With Packet Marking and Logging

Li Pengfei†, Yaokai Feng‡, Junpei Kawamoto‡, Kouichi Sakurai‡
†Graduate School of Information Science and Electrical Engineering, Kyushu University
‡Faculty of Information Science and Electrical Engineering, Kyushu University
744 Motooka, Nishi-ku, Fukuoka 819-0395 Japan
lipengfei@itslab.inf.kyushu-u.ac.jp, fengyk@ait.kyushu-u.ac.jp,
kawamoto@inf.kyushu-u.ac.jp, sakurai@inf.kyushu-u.ac.jp

Abstract Cyber-attack incidents have become more and more frequent and serious. As a countermeasure against cyber attacks, the technology of (IP address etc.) trace back to the attackers is essential. Although many methods have been proposed for this purpose, the existing trace-back techniques suffer from several problems, including that only the specific attacks can be traced back, the tracing back is too time-consuming and traffic-path reconfiguration cannot be

guaranteed. In this study, to discover the attacker quickly and correctly, we propose a new method. Using simulation data, its performance is demonstrated.
Keyword: cyber attack, trace-back, packet marking, logging.

1. はじめに

サイバー攻撃に対する対策は, 攻撃の防衛だけでは不十分になりつつある. 攻撃者を発見し, 法的に責任を取らせることが求められている. IP トレースバックはこのために, 作られた技術であり, いくつかの手法が提案されている. パケットマーキング[1][2], ロギング[2], イングレスフィルタリング[3], エントロピー変量[4], ICMP トレースバック[5]などは様々な手法が提案されている. 中でも, パケットマーキングとロギングを用いた HIT[6]と PPIT[7]は主要なトレースバック手法である. しかし, 既存のトレースバック技術には特定の攻撃しか対応できない, およびトレースバックに長い時間を要するという問題がある. 実用的なシステムでは, 正確に攻撃者を追跡するだけでなく, 早期に攻撃者を同定する必要もある. トレースバックが遅くなれば, 被害の拡大や攻撃者に証拠隠滅の時間を与えることになる.

本論文では早期に攻撃者を発見するために, 通信経路再構成に要する時間の短い新しいトレースバック手法を提案する.

本論文の構成は次の通りである. 2 章では, 関連する既存の技術を紹介する. 3 章では, 提案手法の概要とメリットを説明する. 4 章では, 提案手法を基づいて, シミュレーションを行った結果を説明する. 5 章はまとめである. 6 章では, 結論と今後の課題を述べる.

2. 既存の関連技術

攻撃者を追跡するための代表的な方法の一つは, 通信経路の記録を利用し, 通信経路を再構成する方法である. しかし, インターネットでは匿名アクセスの可能性と国家や組織に依存しないという特性があるため, パケットの通信経路に関する記録を得ることは容易ではない. 従って, 追跡に必要な情報を記録するために, パケットの一部領域を書き換える方法とルータに記録機能を追加する方法の二種類が良く利用されている. それぞれパケットマーキングとロギングと呼ばれる. 本論文でも, このパケットマーキングとロギングが用いる.

2.1 パケットマーキング

パケットマーキングとは通信パケットに経由したルータの IP アドレスを記録する技術である. 図 1 はその概要を示す.

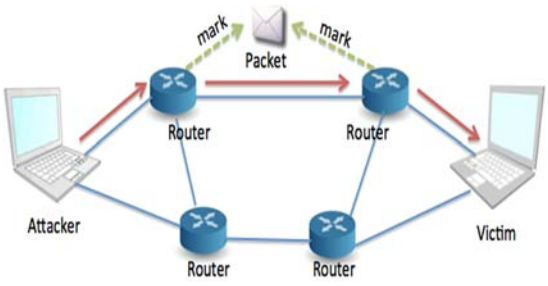

図 1　パケットマーキング

パケットマーキングでは，ルータが情報（ルータの IP アドレスなど）を書き込み，パケットを受信した端末が書き込まれた情報から攻撃ルートを推定する. パケットマーキングは二種類があり, 一定な確率でマークする確率的パケットマーキング PPM 法[8]と常にマークする確定的パケットマーキング DPM[9]法である．PPM はマーキング確率を低く設定（一般的には, 0.04 が用いられる）することにより, ルータの処理時間やパケットデータサイズの増大コストを低減することができる．一方，DPM ではパケットがエッジルータ[10]を経由する際に，エッジルータの情報をパケットに書き込む. エッジルータ経由時のみマーキングを行うことでコストを抑えている.

しかし, これらのパケットマーキング手法にはそれぞれの問題がある. 確率的パケットマーキング PPM はマーキング確率を考慮して通信経路を復元する必要があり, 再構成に要する計算量が大きくトレースバックが遅い[11]．また，確率的とは言え，パケットにルータ情報を追加するため, パケットサイズが段々増加する．一方，確定的パケットマーキング DPM は，一般的に負荷の大きいエッジルータがパケット操作を行う必要があり，高い性能を必要とする.

2.2 ロギング

ロギングとは通信パケットの ID などの情報を, そのパケットが経由したルータに記録するものである[2]．ルータに記録機能を追加し，パケットがルータを通過する際，ルータがパケットの情報を一定の形式で時系列に記録し保存する. ロギングは大量なパケットを送るフラッディング攻撃だけではなく，単一パケットを送る攻撃も追跡することができる. 本手法の弱点としては大量なパケット情報が記録され, ルータの負担が増大する. ルータが全てのパケットを記録する場合, 高性能のルータが必要である. そのためのコストは高い. さらに高速かつ大量の攻撃が行われた場合, 全てのパケットを記録し保存することは不可能である．従って，DDoS 攻撃のような大量のパケットを送る攻撃には効果的でない.

3. 本研究での提案

3.1 提案の詳細

本提案の基本的なアイデアは次のとおりである. ルータに次の二つの操作を行わせる. 一つ目は，ルータの IP アドレスをパケットに記録させることであり，二つ目は，パケットの情報を記録することである．図 2 に示したように, 各通信パケットに２つの記憶領域（マーク領域）を用意する．そして，各パケットに２つのルータのアドレスを保存させる．通信パケット（P とする）が最初のルータ（R_1 とする）を経由した際，ルータ R_1

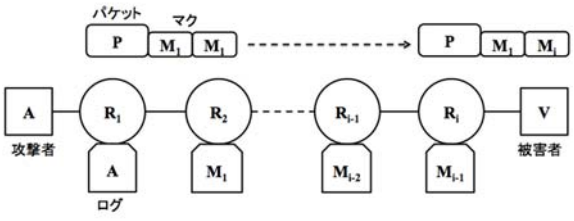

図 2　提案の基本アイデア

は，一定な確率 p で R_1 のアドレス（M_1 とする）を P にマーキングする．併せて，P の送信元 IP を R_1 に記録する．P が次のルータ（R_2 とする）を経由する際，R_2 が P のマーク領域を調べ，もし R_1 が P に記録されていなければ，これから経由したルータではパケットへのマーキングもルータのロギングも行わない．もし R_1 が P にマーキングされていれば，R_2 もマーキングを行う．即ち，P のもう１つのマーク領域に R_2 のアドレス（M_2 とする）をマークする．これで，P の２つのマーク領域は(M_1, M_2)となる．同時に P の始点 IP アドレスの情報を R_2 にログする．以降経由するルータ（R_i とする）では，P にあるマーク M_{i-1}(ルータ R_{i-1} のアドレス)を R_i に記録してから，R_{i-1} の代わりに R_i のアドレス M_i を P にマーキングする．即ち，P にあるマーク M_{i-1} を消して，R_i のアドレス M_i を入れる．これで，P の２つのマーク領域は(M_1, M_i)となる．このように，P の 2 番目のマーク領域の内容が変更し続ける．また，経由した各ルータには一歩前のルータのアドレスを記録させる．最終的に，各パケットには，最初と最後に経由したルータのアドレスが記録冴えれており，全ルータのログから経路を復元することができる．

R_1 で一定な確率 p で記録を行う理由は，攻撃者たち（特に DDoS 攻撃などの場合）が多くのパケットを送ると想定し，その中１つだけでもマーキングできれば十分に追跡できるため，各ルータの負担を軽減するためである．想定した攻撃者の発送パケット数より確率 p を調整する．既存の確率的パケットマーキングと同じように，想定した攻撃者の発送パケット数が多いほど，確率 p が減らすことができる．想定した攻撃者からパケット数が少ない場合は，$p = 1.0$ とする．即ち，R_1 は必ずマークキングされる．その場合，何かのトラブルで R_1 はマーキングされなくても，次のルータはマーキングされるので，トレースバックする際に，無事にマーキングされた最初ルータまでトレースバックできる．

3.2 提案手法のメリット

本提案手法は 4 つのメリットがある，

1) 送信元 IP の同定が速い
   パケットの 1 番目に記録されている情報は経路に関係なく，パケットが最初に経由するルータを指している．そのルータのログにはパケット送信元の IP アドレスが記録されている．そのため，容疑者を見つけることができる．
2) ルート再構成の計算量が少ない
   マークする際，複雑な計算しないので，ルート再構成する際，複雑な計算は必要がない．
3) パケットの負担が軽い
   パケットに二つマークだけを付けるので，他の確率的パケットマーキングと比べ，マークが少ない．
4) 個々のルータに依存しない
   ルータの役割を分担し，リスクを減らす．ルータの役割は大体同じである．そして，一番重要なマークはルータ R1 のマークであるが，不可欠ではない．他のマークも不可欠ではない．

4. シミュレーションによる評価

パケットをトレーシングするシミュレーションを設計して，ルータを経由する際のパケットの特徴を抽出する．このように作成したシミュレーションデータと既存研究の実験データを利用して，提案手法の性能を実証する．

本手法では，ロギングの確率はマーキングの確率と同じ，通信パケットがマークされるなら，ルータにログする．パケットにログしないなら，ルータにログしない．更に，１回の攻撃にたとえ一つの通信パケットだけが正確的にマークされた場合でも，攻撃元を追跡できる．即ち，１回の攻撃に対して，トレースバックの成功率は正確的にマークされた通信パケットが存在するかどうかにより決められる．

シミュレーションにより，通信パケットが少ない際，確率を 100%に設定し，パケット数を増加すると，確率を下がるということを想定する．通信パケット数を 100 から 10000 まで増やせ，通信パケット数が 100，300，500，1000，2000，3000，4000，5000，10000 になる時のマーク数を監視し，各確率で 10000 回のシミュレーションを実行した．図 2 と図 3 に 10000 回のシミュレーションの実行結果を示した．確率は 1/10 の場合，通信パケット数が 100 から 10000 までふやせば，10000 回のシミュレーションでは，追跡できる回数は 10000 回である．確率は 1/20 の場合，通信パケット数が 100 の時，トレースバックの成功率が下がって，99.37%になった．通信パケット数が 300 に増加した以後，成功率が 100%に上がった．確率は 1/50 の場合，通信パケット数が 1000 になった以後は，成功率は 100%になれる．更に，確率は 1/500 と 1/1000 の場合では，成功率が 100%になれるのは，通信パケット数が 10000 になる時である．

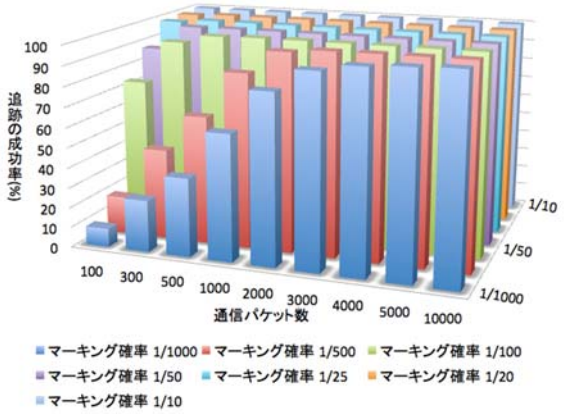

図 2　シミュレーションの結果（図示）

		マーキング確率						
		1/10	1/20	1/25	1/50	1/100	1/500	1/1000
通信パケット数	100	10000	9937	9854	8733	7356	1867	918
	300	10000	10000	10000	9976	9544	4518	2579
	500	10000	10000	10000	9999	9955	6381	3907
	1000	10000	10000	10000	10000	10000	8673	6264
	2000	10000	10000	10000	10000	10000	9821	8398
	3000	10000	10000	10000	10000	10000	9981	9486
	4000	10000	10000	10000	10000	10000	9997	9817
	5000	10000	10000	10000	10000	10000	9998	9934
	10000	10000	10000	10000	10000	10000	10000	10000

図 3　シミュレーションの結果（数値）

図 2 と図 3 のシミュレーション結果から分かるように，各確率で通信パケット数が少ない方は成功率が低いである．

図 4 に各確率 1/10，1/20，1/25，1/50 で通信パケット数を 100 から 500 までを増加させ，シミュレーションをした結果を示した．図

4 の結果から分かるように，成功率を 100%に維持するために，通信パケットが 100 の時に，マーキング確率が 1/10 以上にすべきである．通信パケット数が 300 の時，マーキング確率は 1/25 なら，まだ追跡できる．通信パケット数は 500 の時，マーキング確率は 1/50 は下限になる．

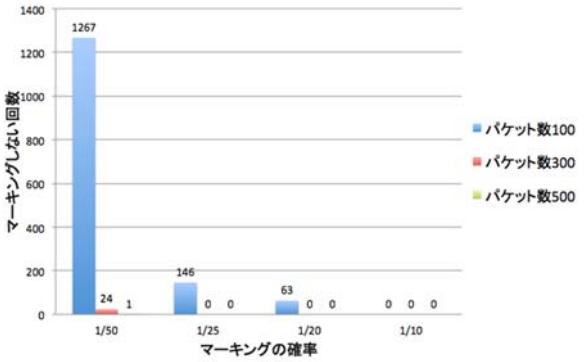

図 4　各確率で追跡失敗した回数

各確率でマークされたパケット数のデータを収集し，ルータにおける負担の大きさを分析するために，各確率が 1/10，1/20，1/25，1/50 の場合で一回のシミュレーションをした．図 5 に通信パケット数が 100 から 10000 まで増やし，各確率でパケットをマークしたシミュレーションの結果を示した．シミュレーションの結果から見ると，通信パケット数の増加と共に，ルータの負担が大きくなる．更に，マーキング確率は大きくほど大きければ，ルータにおける負担が著しくになる．一方，シミュレーションにより，パケットのマーク数は二つ種類がある．図 6 にシミュレーションで一つ通信パケットのマーク数を示した．予想と同じで，マークされないなら，通信パケットのマーク数がいつも 0 である．マークされたなら，通信パケットのマーク数

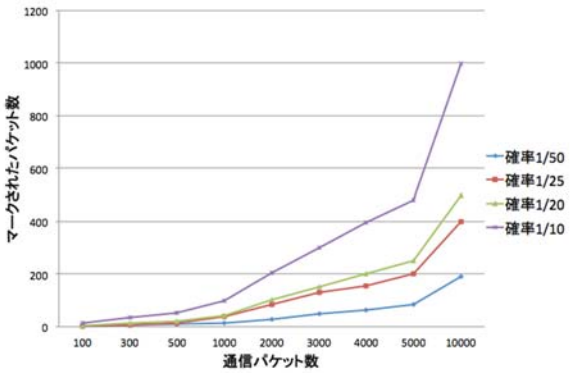

図 5　各確率でマークされたパケット数

がいつも 2 である．PPM と PPIT など既存の手法と比べると，一つパケットにマークは最大二つしかないので，オーバーロード防止する対策は必要がない．

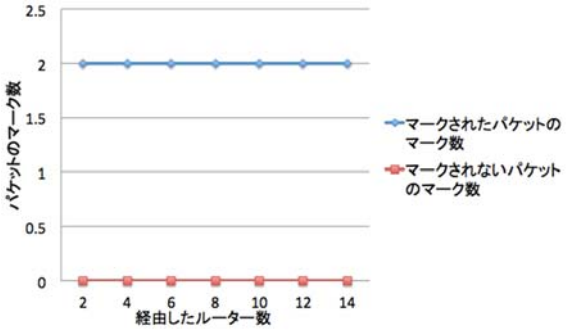

図 6　一つ通信パケットのマーク数

以上のシミュレーションにより，トレースバックの成功率が通信パケット数とマーキング確率との二つパラメーターに決められる．通信パケット数が少ないなら，マーキング確率は上がる必要がある，逆に通信パケット数が一定程度増加したら，ルータにおける負担を減少するために，マーキング確率は下がるべきである．一方，一つ通信パケットのマーク数はルータにおける負担を増加しない．

5. まとめ

既存のトレースバック技術の, 特定の攻撃しか対応できないおよび通信経路再構成の計算量が大きいためにトレースバックに時間を要するという問題を解決するために, 新しい手法を提案した. 模擬実験で提案手法の有効性を確認した.

本手法ではトレースバックの時間が短い. 更に, ルート再構成の計算量が少ない. ルータの役割を分担する為, ルータにおける負担を減少し，リスクも少なくなった．そして, マークが二つだけがあるので, オーバーロード防止する対策は必要がない. 最後，結果としては，IP アドレス偽装攻撃の場合は，追跡の時間が多少影響されるが，基本的には，IP アドレス偽装攻撃にトレースバック能力が強い.

6. 今後の課題

マーキングする際, マークの数だけではなく，マークの大きさも考慮すべきである．よって, マークの内容と領域を選ぶことは重要である．既存の PPM がパケットのヘッドの識別子領域を利用する．更に，ルータ ID を作って，IP アドレスの代わりに，ルータ ID でパケットをマークするルータ ID 領域[12]という研究がある．今後の課題として，より良いマーク内容と領域を決める.

もう一つ課題は IP アドレス偽装攻撃の場合では, 本手法のトレースバック効率が下がると考えられる．そのために，本手法では，IP アドレス偽装攻撃の対策を追加することが必要である.

謝辞

この研究の一部は，科学研究費（基盤研究(C) No. 25330131)の支援を受けている．ここに記して謝意を表す.

参考文献

[1] H. Burchand and B. Cheswick, “Tracing anonymous packets to their approximate source”, in Proceedings of the 14th USENIX Conference on System Administration, 2000.

[2] R. Jain and A. Meshram, “A Survey on Packet Marking and Logging”, (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 4 (3), 2013.

[3] P. Ferguson and D. Senie, “Network Ingress Filtering: Defeating Denial of Service Attacks Which Employ IP Source Address Spoofing”, RFC 2267, Jan. 1998.

[4] S. Yu, W. L. Zhou, and W. J. Jia, “Traceback of DDoS attacks using entropy variations”, IEEE Transactions on Parallel and Distributed Systems, Vol. 22, 2011.

[5] S. M. Bellovin, “ICMP traceback messages”, Network Working Group Internet Draft, March 2000.

[6] C. Gong and K. Sarac, “A more practical approach for single-packet IP traceback using packet logging and marking”, IEEE Transactions on Parallel and Distributed Systems, Vol. 19, 2008.

[7] Y. Dong, Y. Wang, S. Su, and F. Yang, “A

Precise and Practical IP Traceback Technique Based on Packet Marking and Logging", Journal of Information Science and Engineering March 2012.

[8] S. Savage, D. Wetherall, A. Karlin and T. Anderson, "Network Support for IP Traceback", ACM/IEEE Trans. Networking, vol. 9, no. 3, pp. 226-237, 2001.

[9] A. Belenky and N. Ansari, "IP Traceback With Deterministic Packet Marking", IEEE Communication Letters, Vol.7, No.4, Apr.2003.

[10] F. Xue and S. J. Ben Yoo, "Self-similar traffic shaping at the edge router in optical packet-switched networks", Communications, 2002. ICC 2002. IEEE International Conference on. Vol. 4. IEEE, 2002.

[11] X. Yang, W. Zhou and M. Guo, "Flexible Deterministic Packet Marking: An IP Traceback System to Find the Real Source of Attacks", IEEE Trans. on Parallel and Distributed Systems. Vol. 20, No. 4, Apr. 2009.

[12] M. Muthuprasanna, G. Manimaran, M. Manzor and V. Kumar, "Coloring the inter-net: IP Traceback", in Proceedings of the 12th International Conference on Parallel and Distributed Systems, 2006.