インターネットのトポロジ特性を考慮した確率的パケットマーキング手法の提案

國分 淳次† 金岡 晃† 岡本 栄司†

筑波大学システム情報工学研究科
305-8573 茨城県つくば市 天王台 1-1-1
kokubun@cipher.risk.tsukuba.ac.jp kanaoka@cs.tsukuba.ac.jp
okamoto@risk.tsukuba.ac.jp

あらまし 近年, インターネットの安全性を脅かすサービス拒否 (DoS: Denial of Service) 攻撃や分散サービス拒否 (DDoS: Distributed Denial of Service) 攻撃が問題となっており, 対策の 1 つに確率的パケットマーキング (PPM: Probabilistic Packet Marking) 手法がある. この対策は送信元を追跡する技術で, 大量のパケットを要することが問題である. また, 既存研究ではインターネットのトポロジ特性が考慮されていない. そこで本稿ではトポロジ特性を利用する手法を提案する. これにより, 送信元の特定を既存手法より少ないパケット数で攻撃元の特定を可能にした.

Proposal of Probabilistic Packet Marking Method using Topology Characteristic of the Internet

Junji Kokubun† Akira Kanaoka† Eiji Okamoto†

†University of Tsukuba, Graduate School of Systems and Information Engineering
Ten-nohdai 1-1-1, Tsukuba, Ibaraki 305-8573, Japan.
kokubun@cipher.risk.tsukuba.ac.jp kanaoka@cs.tsukuba.ac.jp
okamoto@risk.tsukuba.ac.jp

Abstract Denial of Service (DoS) attacks and Distributed Denial of Service (DDoS) attacks have become Internet-wide threats. Although the defence mechanisms are widely investigated, there are few researches using topology characteristic of the Internet. In this paper, the authors proposed probabilistic packet marking method using topology characteristic of the internet, and a method to reduce the number of packets needed to identify the attacker.

1 はじめに

1.1 背景と目的

近年, インターネットの普及・商用利用が進み, ネットワーク上での攻撃による被害が拡大している. その中でも, サービス提供者が意図しないにも関わらず悪意ある第三者からの攻撃によってサービスの提供ができなくなる, サービス拒否 (DoS: Denial of Service) 攻撃と分散サービス拒否 (DDoS: Distributed Denial of Service) 攻撃は大きな問題になっている.

このDDoS/DoS攻撃を追跡するための手法としてはIPトレースバックという技術があり, 様々な手法が提案されている. その 1 つに確率的パケットマーキング (PPM: Probabilistic Packet Marking) 手法がある. これは流れているパケッ

トそのものにルータ情報を確率的に書き込む (マーキングする) 方法である. 被害ノードは攻撃パケットに書きこまれた情報から攻撃元を特定することができる. これまでに, マーキング方法が異なる様々な PPM 手法が提案されている [4, 5, 6, 8, 9]. しかし, これらの手法は現実のインターネットが持つネットワークの特徴が考慮されていない.

そこで本稿では, 現実のインターネットに見られるスモールワールド性を持つネットワークに対して有効な手法を提案する. これにより, 既存手法より少ないパケット数で攻撃元を特定することを可能にした.

1.2 本論文の構成

まず, 第 2 章では関連研究として既存研究の PPM 手法とインターネットの構造について説明する. 次に第 3 章にてインターネットの特性を考慮した提案手法を説明する. そして, 第 4 章にて, 提案手法の有効性を示すためのシミュレーションについて述べる. 最後に第 5 章でまとめと今後の課題を述べる.

2 関連研究の調査

2.1 PPM 手法

ここでは PPM 手法の関連研究を簡単に紹介する. どの手法もトレースバックのために情報を IPv4 ヘッダに入れる手法で, 各手法によって書き込む情報と情報量が変わる.

Savage ら [4] が 2000 年に PPM 手法を提案し, それ以降 Savage らの手法を応用した PPM 手法が多く提案されている. Savage らの手法では各ルータは通過するパケットに隣接する 2 つのルータの情報を XOR(排他的論理和) して書き込む. このルータの情報とは IP アドレスとそのハッシュ値をインターリーブした値 (bit 単位に分解して, bit 単位ずつ交互に接続した値) にして 8 分割にしたものである. また, ルータを通過する際に距離情報の値をインクリメントすることにより攻撃経路を形成することができる.

Song ら [5] の手法は Savage らの手法を改良した手法である. Savage らの手法では隣接する 2 つのルータの IP アドレスとそのハッシュ値を XOR して書き込んでいたが, Song らの手法は IP アドレスのハッシュ値だけを XOR して書き込む. この手法ではハッシュ値の情報からルータの IP アドレスを復元しなくてはならないので, 上流のルータマップの併用が必要である.

Dean ら [6] の手法はルータの IP アドレスを多項式の係数としており, 被害者は連立方程式を解くことにより IP アドレスを復元する. また, パケットに書き込む bit 数を減らすために Chinese remainder theorem(中国人剰余定理) を利用している.

岡崎ら [9] の手法も Savage らの手法を改良した手法である. 情報を書き込む IPv4 ヘッダのフィールドを多く確保し, 攻撃経路のルータが情報を書き込んだ際にその確認としてフラグを立てることにより, 効率よく攻撃経路を形成する.

Law ら [7] の手法も Savage らの手法を応用している. 各ルータのトラフィック率の変化を調べることによって, 素早く攻撃元を特定している.

Goodrich[8] の手法はチェックサムコードを利用した手法である. 攻撃者に隣接するルータの IP アドレスとチェックサムコードを接続することにより, 攻撃元だけを特定することができる. また, パケットに書き込むルータの情報を 2 回まで分割することが可能である.

2.2 インターネット・トポロジ

近年, インターネットの普及・大規模化に伴い, インターネット・トポロジの研究・調査を盛んに行われている. これらの研究・調査により, インターネットはスモールワールド性とスケールフリー性という性質を持っていることが近年明らかになっている.

スケールフリー性 (べき乗則性) とは, ノードとリンクから構成されるグラフにおいてノードに接続するリンク数で定義される次数の分布がべき乗則に従う性質のことをいう. スケールフリーネットワークでは, ある一部のノード (ハブ) は多数のリンクを持っており, 大多数のノードはごくわずかなリンクしか持っていない.

スモールワールド性とは, 任意の 2 つのノードが中間にわずかな数のノードを介するだけで接続されるという性質のことをいう. インターネット・トポロジの調査プロジェクトである Caida [1] や Dimes [2] の報告を見ると, ルータレベルで

は, あるノードから 7hop 進むと全体の 90 %以上のノードに到達しているのがわかる [3].

3 提案手法

3.1 提案手法の狙い

本稿では既存の PPM 手法では扱われていないネットワークの特徴を考慮した PPM 手法を提案する. 今までの研究ではインターネットの特徴であるスモールワールド性が考慮されておらず, ルータレベルで 31hop 先の攻撃対象者まで考慮している. そのため, パケットに書き込む限られたマーキングエリアを距離情報のために多く利用している. そこで, ルータレベルでは, あるノードから 7hop 進むと全体の 90 %以上のノードに到達できることをもとに, 7hop 以内と 8hop 以降にわけて攻撃元を特定する手法を提案する. これにより, パケットの IPv4 ヘッダに書き込む情報量が既存の PPM 手法より少なくなり, 攻撃元を特定するのに必要なパケット数の減少につながる.

3.2 マーキング法

(1) 各ルータによる準備

1. ルータは事前に自分自身の IP アドレスとそのハッシュ値を計算し, それぞれの値をインターリーブ (IP アドレスは bit 列の奇数番目, ハッシュ値は bit 列の偶数番目として, bit 単位で交互に接続) する.

2. インターリーブした値を k 分割 (以下, 分割した値をフラグメントと呼ぶ) する.

(2) 各ルータによるマーキング手順

1. 確率 p で以下の作業を行う.

   (1)0 から $k-1$ までの整数からランダムに選んだ値を j とする.

   (2)j 番目のフラグメントをパケットの IPv4 ヘッダに書き込む (以下, このフィールドをエッジフラグメントフィールドとする).

   (3) パケットの IPv4 ヘッダに番号 j を書き込む (以下, このフィールドを offset フィールドとする).

   (4) パケットの IPv4 ヘッダに距離情報として 0 を書き込む (以下, このフィールドを distance フィールドとする).

2. 確率 $(1-p)$ で以下の作業を行う.

   (1) オフセットフィールドの値が同じ番号 (j 番目) のフラグメントとエッジフラグメントフィールドの値と XOR を行う. 算出した値をエッジ ID と呼ぶ.

   (2) エッジフラグメントフィールドに算出したエッジ ID の値を書き込む.

   (3)distance フィールドの値をインクリメントする.

3. 下流のルータにパケットを転送する.

このように流れているパケットの追跡に必要な情報がマーキングされて被害者に届く. PPM 手法ではこの追跡に必要な情報を IP ヘッダ部に書き込んでおり, それぞれの提案手法によって書き込む場所と情報量は違う. 図 1 に IPv4 のパケット構造を示す.

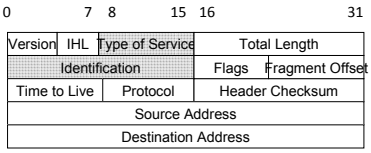

図 1: IPv4 のパケット構造

Savage ら [4], Song ら [5], Dean ら [6] の手法は IPv4 ヘッダにある identification フィールドの 16bit を利用している. また, 岡崎ら [9] と Goodrich[8] は IPv4 ヘッダにある identification フィールドの 16bit と ToS(Type of Service) フィールドの 8bit を利用している. これは PPM 手法のフラグメント化されたパケットがインターネットを流れるトラヒック全体から見てもごくわずかな量であるとの前提のもとに, 彼らは idenidentification フィールドと ToS フィールドを利用している. そこで我々も identification フィールドの 16bit を使う手法と identification フィールドの 16bit と ToS フィールドの 8bit を使った手法を提案する.

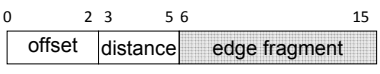

図 2: 16bit 版のマーキング情報

まず, identification フィールドの 16bit を使うマーキング情報の構成を図 2 に示す. 16bit のうち, フラグメントの順番を書き込む offset フィールドを 3bit, 距離の情報を書き込む distance フィールドを 3bit, ルータの情報を書き込むエッジフラグメントフィールドを 10bit とする. そうすると, 1 パケットあたり 10bit のデータを書き込むことができる. Savage らの手法ではフラグメントのフィールドが 8bit しかなかったため、分割数 k が 8 必要だったが, 提案手法では 7 にすることが可能である. さらにハッシュ値のサイズを 32bit から 28bit にすると, インターリーブした値は 60bit になるので, 分割数 k を 6 にすることが可能となり, 受信パケット数を減らすことが可能である.

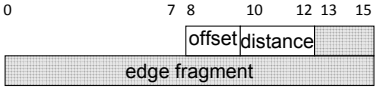

図 3: 24bit 版のマーキング情報

次に identification フィールドの 16bit と TOS フィールドの 24bit を使うマーキング情報の構成を図 3 に示す. 24bit のうち offset フィールドを 2bit, ディスタンスフィールドを 3bit, エッジフラグメントフィールドを 19bit とする. これにより, 1 パケットあたり 19bit のデータを書き込むことができるので, 分割数 k を 4 にすることが可能である. さらにハッシュ値を 32bit から 25bit にすると, インターリーブした値は 57bit になるので, 分割数 k を 3 にすることが可能となり, 受信パケット数を減らすことが可能である.

3.3 攻撃元の特定

ここでは被害者が攻撃元を特定するまでの手順を示す.

1. 被害者は受信したマーキングパケットの中から同じ distance フィールドの offset フィールドを調べ, k 分割されたフラグメントを結合し, エッジ ID を集める.
2. 集めたエッジ ID と隣接するルータのエッジ ID と XOR を行い, インターリーブした値を算出する.
3. 算出したインターリーブした値から IP アドレスとハッシュ値を戻す.
4. IP アドレスをハッシュ化してハッシュ値と比較をする.
5. 値が一致すれば正当なインターリーブした値とし, 一致しない場合はフラグメントの違う組み合わせを試す.
6. 攻撃元にあるルータとその下流にあるルータのマーキングパケットから, IP アドレスを算出することにより攻撃元を特定することができる.

次に被害者が攻撃元を特定するのに必要なパケット数の期待値を示す. 以下では全てのルータは確率 p で通過するパケットにマーキングを行うものとする. d hop 離れたルータでマーキングされたパケットが下流のルータによって上書きされずに被害者に届ける確率は $p(1-p)^{d-1}$ である. そして, d hop 離れたルータから一つのマーキングパケットを受信するのに必要な攻撃パケット数の期待値 Q は $Q = 1/p(1-p)^{d-1}$ である. ここでは簡単化のために, 異なる距離のルータがマーキングパケットを送る確率を同じであるとする. このとき, 攻撃元に近い 2 個のルータのうち, 少なくとも 1 つのルータからマーキングパケットが送られてくる確率は $2p(1-p)^{d-1}$ となる. したがって, 2 個のうち少なくとも一つのルータから一つのマーキングパケットを受信するのに必要なパケット数の期待値 $E(P)$ は $E(P) = 1/2p(1-p)^{d-1}$ となる. さらに, クーポンコレクター問題より, 被害者が受信したマーキングパケットの中から, 何回の試行で 2 種類分のマーキングパケットを集めることができるという数の期待値 $E(C)$ は, k 分割されていることを考慮すると, $E(C) = 2k(ln(2k))$ である. よって, 攻撃元を特定するのに必要なパケット数の期待値 $E(X)$ は $E(P) \times E(C)$ より以下の式で与えられる.

$$E(X) < \frac{k \cdot ln(k)}{p(1-p)^{d-1}} \qquad (d = 0)$$

$$E(X) < \frac{k \cdot ln(2k)}{p(1-p)^{d-1}} \qquad (0 < d < 8)$$

3.4 ノードが 7hop 以上の場合について

今回の提案手法はインターネットの性質である 7hop 進むと 90 % 以上のノードに到達できるという条件のもとに提案された. しかし, 8hop 以

上の場合も少ない確率だが発生するので, 当然考えなくてはいけない. そこで distance フィールドが 7 のパケットを攻撃経路途中のルータが受信した場合, そのルータは以下の作業を追加する.

1. distance フィールドを 0 に初期化し, IP アドレスとハッシュ値の組み合わせのインターリーブ (bit 単位で交互に連結した値) の他に, IP アドレスと別のデータ (例えば全て 0 の bit 列) の組み合わせのインターリーブを用意する.
2. それぞれのインターリーブした値を k 分割する.
3. 攻撃パケットに情報を書き込むときは, 2 種類のインターリーブした値をランダムに選ぶ.

この作業を追加することにより, 8hop 以上離れたルータから被害者に送られるパケットはインターリーブした値の奇数番目が IP アドレスで, 偶数番目が異なる 2 種類のパケットを受信することになる. この 2 つのインターリーブした値を復元することにより, この IP アドレスは偶数番目が異なる 2 つのパケットを持っているので, 8hop 以上離れたルータのアドレスと被害者は判断できる. この distance フィールドが 7 のパケットを攻撃経路途中のルータが受信した場合の作業を図 4 に示す.

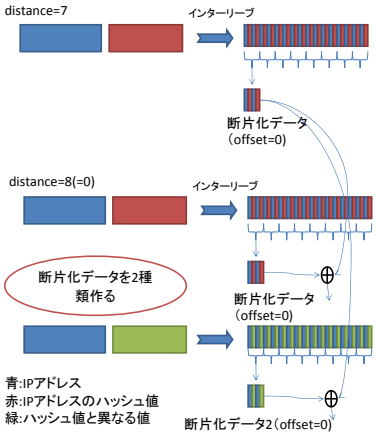

図 4: 攻撃経路が 8hop 以上の場合

また, 8hop 以上離れた時に必要なパケット数の期待値を以下に示す.

$$E(X) < \frac{k \cdot ln(4k)}{p(1-p)^{d-1}} \qquad (d = 8)$$

$$E(X) < \frac{k \cdot ln(6k)}{p(1-p)^{d-1}} \qquad (8 < d < 15)$$

4 シミュレーション

提案した PPM 手法の有効性を示すために, 前章で示した必要パケット数の期待値を用いてシミュレーションで評価を行った.

シミュレーションは identification フィールドの 16bit を使う提案手法 (以下, 16bit の提案手法とする) と identification フィールドと TOS フィールドを使った 24bit を使う提案手法 (以下, 24bit の提案手法とする) の両者を行う. 本実験では攻撃元を特定するのに必要なパケット数に着目し, 比較対象としては Savage ら [4], 岡崎ら [9], Goodrich[8] の手法と比較を行っている. なお, Savage らと岡崎らはマーキングする情報量が違うが, 必要パケット数の期待値は同じなので, グラフでは同一として扱う. シミュレーション条件は, 攻撃者数 $n = 1$, 確率 $p = 1/25$ とし, 攻撃元までの距離を変化させたときの必要パケット数を求めた.

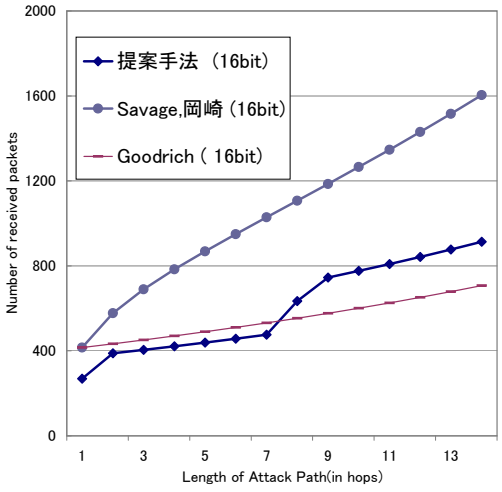

図 5: 攻撃経路の長さを変化させた時の必要パケット数 (16bit)

まず, 16bit の提案手法のシミュレーション結果を図 5 に示す. 比較対象としては同じ 16bit を使う Savage ら (岡崎ら) の手法, 16bit を使う Goodrich の手法である. Dean らの手法も 15bit

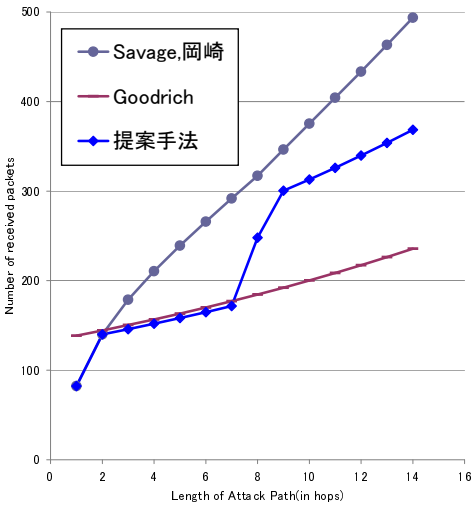

図 6: 攻撃経路の長さを変化させた時の必要パケット数 (24bit)

だが, 提案手法や Savage ら (岡崎ら), Goodrich の手法と比較して多くのパケット数を必要とする手法であるため, ここでは比較の対象から外した. 図 5 を見ると攻撃経路の長さが 7hop までなら提案手法の受信パケット数が一番少ないのがわかる. また, 7hop 以降でも Goodrich の手法を除いた手法より受信パケット数が少ないのがわかる. 次に 24bit の提案手法のシミュレーション結果を図 6 に示す. 比較対象としては 24bit に拡張した Savage ら (岡崎ら) の手法, 25bit を使う Goodrich の手法である. 図 6 を見ると攻撃経路の長さが 7hop までなら提案手法の受信パケット数が一番少ないのがわかる. また, 8hop 以降でも Goodrich の手法を除く手法より受信パケット数が少ないのがわかる. 以上より, 現実のネットワークの特徴であるスモールワールド性を考慮した提案手法が既存の手法より有効であることが確認された.

5 まとめと今後の課題

本稿では, インターネットのトポロジに関する調査結果を元に, インターネットの性質であるスモールワールド性を考慮した PPM 手法を提案した. そして有効性を確認するために, シミュレーション実験を行った結果, どの関連手法よりも受信に必要なパケット数が少ないことを確認できた.

今後の課題としては, DDoS 攻撃を想定したシミュレーションや, 提案手法の改善などがあげられる. また, 現状の確率的パケットマーキング手法ではルータへの適用率を 100 %としているが, 適用率を減少した場合の検討も行う予定である.

参考文献

[1] Caida http://www.caida.org/home/
[2] Dimes http://www.netdimes.org/new/
[3] Guido Caldarelli, Alessandro Vespignani,“ Large scale structure and dynamics of complex networks ”, 2007.
[4] S. Savage, D. Wetherall, A. R. karlin, and T. Anderson,“ Practical network support for IP Traceback ”in Proc. ACM SIGCOMM, pp.295-306, 2000.
[5] D. Song and A. Perrig, “ Advanced and Authenticated Marking Schemes for IP Traceback ”in Proc. IEEE INFOCOM, pp. 876-886, 2001.
[6] D. Dean, M. Franklin, and A. Stubblefield,“ An algebraic approach to ip traceback ”in Proc. Network and Distributed System Security Symp. (NDSS), pp. 3-12, 2001.
[7] T. k. T. Law, D. k. Y. Yau, and J. C. S. Lui, “ An effective statistical methodology to trace back DDOS attackers ”IEEE Trans. Parallel Distrib. Syst., vol. 16, no. 9, pp. 799-813, Sep. 2005.
[8] M. T. Goodrich,“ Probabilistic Packet Marking for Large-Scale IP Traceback ” IEEE/ACM TRANSACTIONS ON NETWORkING, vol 16, no. 1, pp. 15-24, 2008.
[9] 岡崎 直宣, 河村 栄寿, 林 美娘, “ サービス不能攻撃の経路追跡手法の効率化に関する検討 ”情報処理学会論文誌, vol. 44, no. 12, 3197-3201, 2003.