現実的なインターネットトポロジを考慮した確率的パケットマーキング手法の提案と評価

後藤 成聡[1]　金岡 晃[2]　岡本 栄司[1]　岡田 雅之[3]

概要：サービス不能攻撃は現代のインターネット社会において重大な脅威である．その対策技術である IP トレースバックの 1 つに確率的パケットマーキング (PPM) があり，他の IP トレースバック手法に対し多くの利点を有している．しかし実用化を検討する上では PPM 未対応ルータへの対応，パケット収集の効率性，攻撃経路再構成の高速化などの課題が依然残されている．また既存の PPM 手法としてマーキングを付与する確率を静的に設定するものと動的に設定するものとがあるが，いずれの手法の評価も理想的なトポロジを用いた効率性の評価であった．そこで本研究では課題のうちの 1 つである攻撃経路再構成に必要なパケット収集の効率性に焦点を当て，既存手法の課題の解決を目的とし，現実的な実現性をもつと考えられる PPM 手法を提案する．本提案手法ではユーザとインターネットの境界に位置するルータでのマーキング確率を任意に設定し，得られるマーキング情報のバラつきを減少させる．また CAIDA のインターネットルータトポロジを利用し，マーキングされたパケットが被害者へ到着する確率の期待値と，同トポロジ上でのシミュレーションによって既存の手法と提案手法の評価を行う．

Proposition and Evaluation of Probabilistic Packet Marking Methods Considering Realistic Internet Topology

Nasato Goto[1]　Akira Kanaoka[2]　Eiji Okamoto[1]　Masayuki Okada[3]

1. はじめに

現在のインターネット社会においてサービス不能攻撃 (Denial of Service Attack, 以降 DoS 攻撃) は非常に重大な脅威となっている．DoS 攻撃はサーバやネットワークを構成する機器に対し大量のパケットを送信することで対象の処理能力や通信帯域を占有し，本来提供されるサービスの利用を困難にするサイバー攻撃である．また近年はボットネット等を利用し DoS 攻撃を多数のホストから一斉に行う，分散型 DoS 攻撃 (Distributed DoS 攻撃, 以降 DDoS 攻撃) が主流となり，単にサーバや機器の増強といった対策では凌ぎきれないほどの強力な攻撃として問題となっている．

このような攻撃への対策技術として IP トレースバックがあり，DoS 攻撃の攻撃元を特定することを目的として盛んに研究が行われている．IP トレースバックはその手法によっていくつかの種類に分けられ，確率的パケットマーキング (Probabilistic Packet Marking, 以降 PPM) はその 1 つに分類される．PPM は他の IP トレースバック手法に対しネットワークを監視する機構が不要であること，攻撃元を特定するためだけに利用する余計なパケットを発生させないこと，そして現在のパケット通信に利用されている IPv4 ヘッダに情報を書き込むことで拡張情報が不要であることといったいくつかの点において優れている．

PPM は 2000 年に Savage らによって提案され [1]，それをもとに様々なアプローチが研究されてきた [2]〜[10]．Savage らの手法は攻撃者から被害者へ至る経路上に存在するルータが静的な確率を用いて確率的にパケットへ自身の情報をマーキングし，被害者がマーキングされたパケットを収集することで攻撃経路を再構成する．この手法には

[1] 筑波大学, 〒 305-8577 茨城県つくば市天王台 1 丁目 1-1, {goto@cipher. , okamoto@}risk.tsukuba.ac.jp
[2] 東邦大学, 〒 274-8510 千葉県船橋市三山 2-2-1, akira.kanaoka@is.sci.toho-u.ac.jp
[3] 社団法人日本ネットワークインフォメーションセンター, 〒 101-0047 東京都千代田区内神田 3-6-2 アーバンネット神田ビル 4F, okadams@nic.ac.jp

マーキング確率が静的であるが故に，あるルータでマーキングされたパケットが被害者に到着する確率 (以降到着確率と呼ぶ) がそのルータから被害者までの距離 (中継されたルータ数) の増加に伴い減少するという問題がある．また他にも経路上でマーキングされなかったパケット (以降，非マークパケットと呼ぶ) が発生し，攻撃経路の再構成に必要なハッシュ計算の回数を大幅に増加させるという問題もある．これらの問題に対し Liu らは 2007 年に動的な確率でマーキングを行う Dynamic PPM を提案し [11]，その後それをもとにした手法がいくつか提案された [12][13]．これらによると動的な確率を利用することでマーキングを行うルータがいずれの距離にあっても到着確率を一定にすることが可能になる．しかしながらこれらの手法の評価は理想的なトポロジ上で行われており，実際のインターネットに即した現実的な評価は行われていなかった．

そこで本研究でははじめに攻撃経路再構成に要するマークパケットの収集効率について，現実的なネットワークトポロジを考慮した立場から再評価を行う．次に既存手法の課題の克服を目的とし，Fruit-structured PPM という手法を提案する．CAIDA のインターネットルータトポロジをもとにして攻撃ケースを多数作成し期待値評価，シミュレーション評価を行った結果，Fruit-structured PPM は既存の Savage らの PPM や Liu らの Dynamic PPM より各距離における到着確率の差が小さく，必要パケット数のバラつきが小さい点で優れていることが分かった．

2. 関連研究

本章では既存手法の中で代表的な Savage らの PPM と Liu らの Dynamic PPM の概要とその課題について説明する．

2.1 PPM

2.1.1 Savage らの手法

2000 年に Savage らが提案した PPM 手法では，攻撃者から被害者に至る経路上のルータが通過するパケットに確率的に自身の情報をマーキングする．マーキングを行う確率は静的に定められるため，後に説明する Dynamic PPM と比較し Static PPM(以降 S-PPM) と呼ぶこととする．マーキングの内容について，経路上のあるルータを X とした時，X は事前に自身の IP アドレス (32 ビット) とそのハッシュ値 (32 ビット) をビットインタリーブした 64 ビットの情報を 8 つに分割し，保持しておく．X が静的な確率 p でマーキングを行う際，8 分割のうちの 1 つ (8 ビット) がランダムに選択され，そこにオフセット情報 (3 ビット)，被害者から X までの距離情報 (5 ビット) を初期値 0 にセットしたものを合わせ，計 16 ビットが IPv4 ヘッダの Identification フィールド (16 ビット) に書き込まれる．

Savage らの手法の特長はパケットがどの 2 ルータ間を

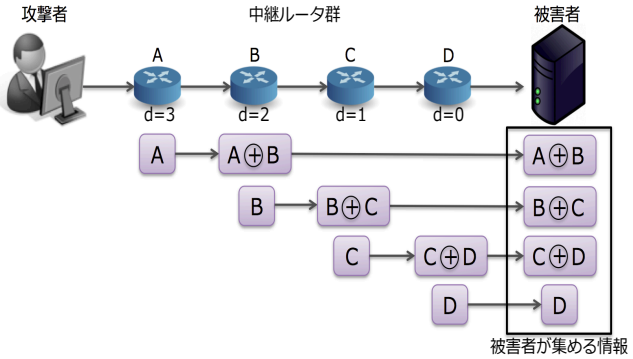

図 1 Savage らの手法におけるマーキング方法

通過したのか，という情報を記録する点である．ルータ Y においてマーキングが行われなかった場合 (確率 $1-p$ で発生)，書き込まれている距離情報が 0 であった場合には既に書き込まれているマーキング情報とルータ Y 自身の情報の排他的論理和 (XOR) で上書きする．また距離情報が 0 であるないに関わらず距離情報をインクリメントする．Savage らはこの手法を Edge Sampling と呼んでいる．図 1 は経路上の各中継ルータがパケットに対し XOR 処理を行う様子を表している．被害者は AxorB，BxorC，CxorD という情報をもったパケットと，$d=0$ である D のルータのマーキング情報を収集する．

被害者が攻撃経路の再構成を行うためにはマーキングされたパケットを収集する必要がある．まず収集したパケットのうち距離情報が 0 のパケットは XOR 演算が行われていない直近のルータのマーキング情報をもつパケットということになる．図 1 では D と書かれたパケットがそれに該当する．ここで距離情報が 0 である 8 種類のオフセットのマーキング情報を合体し，ビットインタリーブを解くことで d=0 のルータ D の IP 情報を復元することができる．この IP 情報と，図 1 で C⊕D とかかれている，距離情報 1 をもつ 8 種類の組み合わせとの排他的論理和をとることで d=1 のルータ C の IP 情報が復元される．これを繰り返すことで攻撃経路上にあるルータ群の IP 情報を復元することができる．

2.1.2 課題

Savage らの手法における課題は 2 つある．まず 1 つは各距離ごとにパケットの到着確率に差が生じてしまうことである．パケットの到着確率とは経路上に存在するあるルータのマーキング情報を被害者が得られる確率であり，静的なマーキング確率 p を用いた場合の被害者から距離 d だけ離れた位置にあるルータの到着確率は $p(1-p)^d$ となる．この値は d が増加し，攻撃者に近づくほど低くなってしまう．これはすなわち被害者からの距離が遠くなるにつれ，そのルータ情報が得られにくくなることを意味し，再構成に必要なパケットの収集に影響を及ぼすこととなる．

2 つ目は非マークパケットの存在である．静的な確率 p

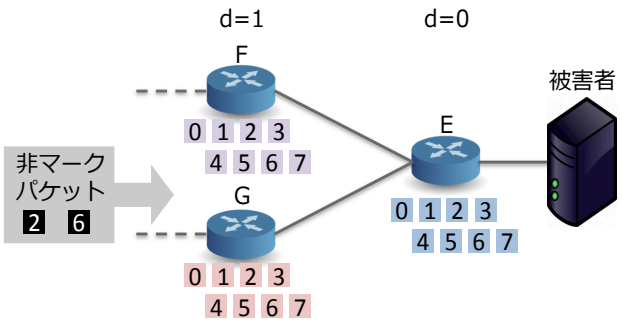

図 2 再構成のためのハッシュ検証

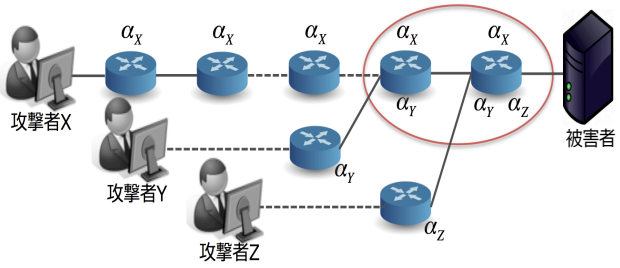

図 3 DDoS 攻撃に対する Dynamic PPM の到着確率

を用いてマーキングを行うと，攻撃経路を辿る過程でどのルータにもマーキングされないパケットが生じる．攻撃経路の再構成のためには同じ距離情報をもつパケットをオフセットの数である 8 だけ集めて合体させ，ビットインタリーブを解きハッシュの再計算による検証を行う必要がある．ここにルータの IP 情報を全く持たない非マークパケットが入り込むことでハッシュの再計算量が増大してしまう．これは再構成に要する時間を増加させ，攻撃元の発見を遅らせることにつながる．

図 2 を用いてこれを説明する．ルータ E のオフセットが全て揃い IP 情報が既に再構成できているとした時，ルータ F,G の再構成でもそれぞれ 8 種類のオフセットをもつパケットが必要になる．すなわち距離情報とオフセットが全く同じで，分割データのみが異なるパケットが 8 対必要になり，ビットインタリーブを解くパケットの組み合わせとして $2^8 = 256$ 通りを試す必要がある．ここで仮にパケットのヘッダフィールドに記載されている情報を PPM のデータとして扱った際に偶然 $d = 1$ でオフセットが 2 と 6 の 2 つの非マークパケットが被害者に到達した場合，$3^2 \times 2^6 = 576$ 通りの組み合わせのハッシュ検証が必要になり，これは非マークパケットがない場合の 2 倍になる．

2.2 Dynamic PPM

2.2.1 Liu らの手法

マークパケット到着確率の減少，非マークパケットの発生という 2 つの課題を解決するため，Liu らは Dynamic PPM(以降 D-PPM と呼ぶ) の提案を行った．この手法ではマーキングの確率としてパケットの経過ホップ数の逆数を利用することを提案し，攻撃者から見て i 番目に経由するルータは確率 $p = 1/i$ でマーキングを行う．Liu らはこれにより攻撃経路のパス長が D である経路上に存在する全てのルータの到着確率は $1/D$ という一定の値になることを示した [11]．さらに攻撃者からのパケットが最初に届いたルータの経過距離は 1 であるため，そのルータでは確率 1 で必ずマーキングすることになり，非マークパケットを発生させないという特徴をもつ．

ルータが位置する場所ごとに異なるこの動的なマーキング確率は IPv4 ヘッダの TTL(Time To Live) フィールドの減少幅を各ルータが参照することで実現する．そのためには TTL の初期値を知る必要があるが, Liu らは初期 TTL は OS やプロトコルごとに異なるものの，$\tau = \{32, 64, 128, 255\}$ の 4 つのうちどれかに定められていることを前提としている．そのため到着したパケットの TTL の値を t とした時の初期 TTL は集合 τ のうち，t 以上で最小のものと推定される．例えば $t = 47$ の時は初期 TTL は 64 であり，経過ホップ数 $= 64 - 47 = 17$ となる．

2.2.2 課題

ここで D-PPM の課題を 3 つ挙げる．第 1 に攻撃元が複数ある DDoS 攻撃に対しては到着確率が一定にならないという予想が考えられる．DDoS 攻撃のトポロジは図 3 で描かれているように被害者の直近やその数ホップ先のルータで攻撃経路が集約された木構造となる．このため攻撃者 X から被害者へ至る経路上の各ルータの到着確率 α_X は一定となるものの，図中において丸で囲まれた被害者に近いルータの到着確率は攻撃者 Y，攻撃者 Z の攻撃経路の到着確率である α_Y，α_Z も含むこととなり結果として値が上昇する．これは D-PPM の本来の目的の実現を大きく損なわせる．また Liu らは一直線のトポロジ上での数式による評価において S-PPM より優れているという結果を得ているが，実際に行われている DDoS 攻撃のトポロジに対する評価では必要パケット数などを算出しているわけではない．

第 2 にパケットの経過距離数は TTL フィールドより算出するとしているが，TTL の値に完全な信頼を置くことの疑問が生じる．これは攻撃者によって初期 TTL が偽装されたパケットによる DoS 攻撃が行われる可能性があること，そして初期 TTL の値が集合 τ に含まれていることについての保証はなく，例外や OS，プロトコルのバージョンアップなどにより TTL から正しい経過距離数が算出できなくなる恐れがあることが理由である．

第 3 にネットワークの根幹に位置するバックボーンルータにおいてパケットごとにマーキング確率を変えることが可能なのかという疑問がある．現在の IPv4 パケット通信においても TTL フィールドはパケット毎に参照されデクリメントされているが，この処理は FPGA などのハードウェア実装によって実現されていることが多い．パケッ

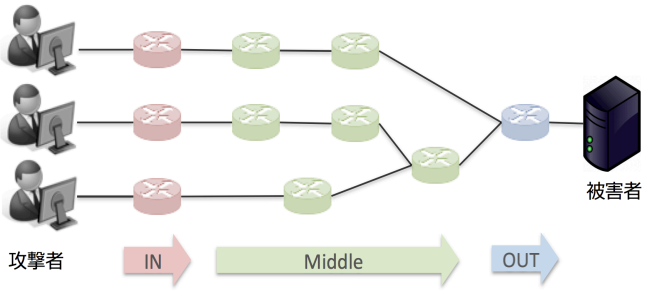

図 4 Fruit-structured PPM の物理モデル

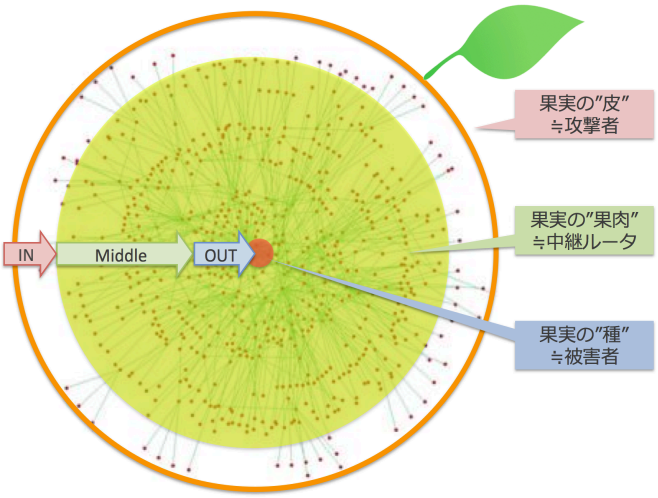

図 5 Fruit-structured PPM の論理モデル

トごとに TTL の値からマーキング確率を算出するという D-PPM の仕組みもハードウェア化することで実現は可能であるが，既に利用されているルータのハードウェア変更のコストを考慮すると困難であると考えられる．

3. 提案手法

図 4 は本論文の提案手法である Fruit-structured PPM(以降 F-PPM) の物理的なモデルを表している．2.2.2 節でも見たように，DDoS 攻撃の攻撃トポロジは木構造となる．F-PPM ではユーザとインターネットとの接点に着目し，より簡易に実現でき，かつ効率的なパケット収集を目的としている．

まず攻撃者からインターネットへパケットを送るルータを IN ルータ，インターネットから被害者へパケットを送るルータを OUT ルータ，さらにインターネット内部でパケット転送を行うルータを Middle ルータと呼ぶこととする．F- PPM ではこの 3 種のルータのそれぞれで異なるマーキング確率を設定することで，これまでのマーキング手法を改善できると考える．確率の定め方としては全ての Middle ルータでは従来の Savage らの PPM と同様に一定の確率 p_{mid} を使うのに対し，IN ルータでは p_{mid} よりも大きい確率 p_{in} を，OUT ルータでは p_{mid} よりも小さい確率 p_{out} を設定させる．これは攻撃者に近いルータにおける到着確率を高め，被害者に近いルータにおける到着確率を低くするためであり，これにより全ての距離で一定とはいかないものの各距離における到着確率の差を他の手法と比較して小さくすることができると考えられる．さらに $p_{in} = 1$ と定めることで IN ルータで必ずマーキングを行うようにし，再構成に支障をきたす非マークパケットが出現しないようにすることが可能である．

そして被害者を中心に据え，周囲に攻撃者が配置されていると考えることで図 5 のような論理的なモデルが得られる．この図の円全体を果実 (Fruit) と捉えると，攻撃者は“皮”，中継ルータは“果実”，被害者は“種”と考えることができる．

2.2 節で述べたように D-PPM はネットワーク上のルータに，全てのパケットの TTL からマーキング確率を算出する仕組みが必要となる．この実現性の問題については 2.2.2 項で述べた通りであるが，それと比較すると F-PPM は IN，Middle，OUT でのマーキング確率はそれぞれ静的に定めるため，実装上の負荷は Savage らの PPM 手法と変わらない．特に岡田らにより Linux ルータにおけるマーキング処理の負荷はマーキング確率の大きさと無関係であることが示されており [7]，任意にマーキング確率を設定することの負荷については考慮する必要がない．以上を踏まえ実用化への障壁という観点から考えた場合，F-PPM は D-PPM よりも現在のインターネットに導入しやすいと考えられる．

4. 評価

第 2 章で既存手法の評価の現実性についての問題を指摘した．本章では 2003 年の CAIDA のインターネットルータトポロジデータ [14] を利用して PPM の各手法の評価を行う．

4.1 評価方法

評価をするにあたり以下の手順で CAIDA のデータセットの処理を行う．CAIDA のデータは 192244 個のノードと 636643 本のリンクから構成されている．そこからまず被害者となるノードをランダムに 500 通り選択する．それぞれの被害者ノードにおいて攻撃者数を 7 パターン {10, 20, 50, 100, 200, 500, 1000} 設定し，各被害者ノードの各攻撃者数パターンにおいてランダムな攻撃者の選択を 100 回行う．ここまでで各攻撃者数につき 5 万ケース，合計で 35 万ケースの攻撃トポロジを得ることができる．

この 35 万ケースの攻撃トポロジについて期待値評価とシミュレーション評価の 2 通りの評価を行う．まず期待値評価では 35 万ケースの 1 ケースごとに攻撃経路上の各ノードの到着確率を求める．次に各ケースにおいて被害者から

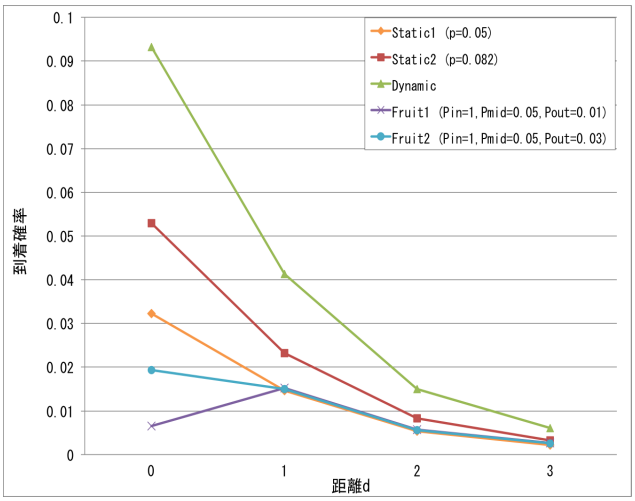

図 6 攻撃者数 100 の時の $0 \leq d \leq 3$ における到着確率

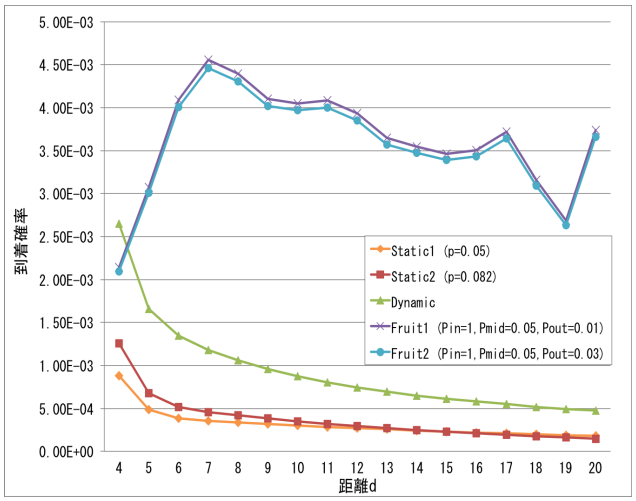

図 7 攻撃者数 100 の時の $4 \leq d$ における到着確率

の距離ごとに到着確率の平均を計算し，それをさらにランダムな攻撃者の選択回数である 100 回分で平均をとる．最後に被害者の選択数である 500 通りの平均を出すことで，最終的に 7 パターンの攻撃者数ごとに到着確率の平均を得ることができる．シミュレーション評価では Java を用いて PPM 用のシミュレータを作成し，入力として与えられた攻撃トポロジ上で仮想的にパケットにマーキングを行い，被害者で収集する様子を再現した．出力として各ノードの IP 情報再構成のために必要なオフセット数を揃えるまでに収集した全体のパケット数，すなわち必要パケット数を得る．

評価の対象とするのは Savage らの S-PPM，Liu らの D-PPM，本論文の提案手法である F-PPM の 3 手法とする．この 3 手法のうち D-PPM はパケットの経過距離の逆数という一律のマーキング確率を利用するが，S-PPM と F-PPM については各ノードの到着確率を算出するためにマーキング確率を設定する必要がある．まず S-PPM では $\{0.05, 0.082\}$ という 2 つの確率を利用する．前者は Savage らが利用した 1/25 に近い値であり，後者は岡田ら [7] によって求められた S-PPM における最適マーキング確率である．次に F-PPM では $p_{in} = 1$，$p_{mid} = \{0.05, 0.082, 0.1\}$，$p_{out} = \{0.01, 0.03, 0.05\}$ という確率を利用する．p_{in} を 1 とするのは前述のとおり非マークパケット出現させないためである．また p_{mid} は S-PPM と同様の二つの確率にそれらより少し大きい 0.1 を追加しており，p_{out} については p_{mid} 以下の確率として 3 種類を用いている．これによりマーキング確率の組み合わせは 9 通りとなる．

またパケットに書き込む情報，分割数，Edge Sampling 手法については Savage らの手法をそのまま利用する．

4.2 期待値評価結果

本節では期待値評価の結果を示す．4.1 節で述べたように F-PPM のマーキング確率は 9 通りの組み合わせがあるため，事前に F-PPM 手法内において評価を行い，結果が優れている確率の組み合わせである $(p_{mid}, p_{out}) = (0.05, 0.01), (0.05, 0.03)$ の 2 通りを利用する．

図 6,7 は攻撃者数が 100 の時の各距離ごとの到着確率の平均値を表す．グラフの X 軸は被害者からの距離を表し，Y 軸はその距離におけるマークパケットの到着確率の平均値を表す．また距離 d は最大 20 まであるが全ての距離を同じグラフに表現すると $4 \leq d$ の視覚的な状況把握が難しいため，$0 \leq d \leq 3$ と $4 \leq d$ の 2 つのグラフに分けて掲載する．

グラフの評価の 1 つの指標は到着確率の高さである．この値が高ければ単純にその距離に位置するルータのマークパケットが得られやすいことを意味するためである．そして到着確率はただ値が高いだけでなく，各距離における差が小さくグラフ全体として平坦であることが重要となる．各距離ごとの到着確率の差が大きければそれだけ集まるパケットにばらつきが生じ，再構成のためのパケット収集に時間を要する．

上記を踏まえてグラフを見る．まず図 6 は攻撃者数が 100 で $0 \leq d \leq 3$ の到着確率のグラフである．Dynamic の到着確率は $d = 0$ と $d = 3$ との間に 0.1 程度の差があり Static1 と Static2 は 0.03〜0.04 の差が生じている．これに対し Fruit は到着確率の差が 0.01〜0.015 程度と取りうる値の幅が小さいことが分かる．図 7 より d が 4 以上になると Dynamic と Static の到着確率は距離の増加に伴い減少しているのに対し，Fruit の到着確率は平均しておよそ 4 倍程度の値を保っている．Fruit 方式は他 2 手法と比較し取りうる値の幅は大きいものの，距離の増加が到着確率へ与える影響が小さい点に注目したい．

また他の攻撃者数の場合にも攻撃者数 100 の時と同様の傾向が見られたため，本論文では割愛する．

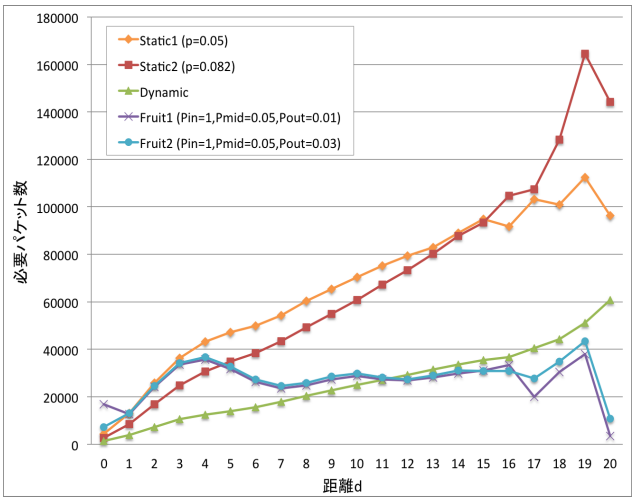

図 8 攻撃者数 100 の必要パケット数

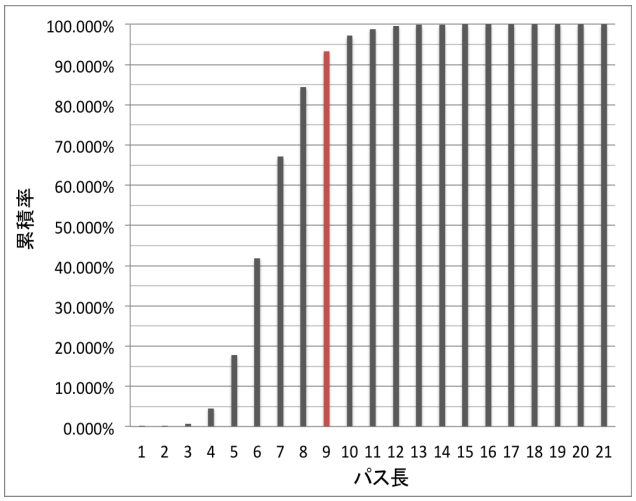

図 9 攻撃者数 100 のパス長の累積率

4.3 シミュレーション評価結果

本節ではシミュレーション評価の結果を示す．図 8 は攻撃者が 100 の時のシミュレーションの結果である．グラフの Y 軸はその距離のノードの IP 情報を再構成するために，そのノードで集めるべきオフセットをもつパケットをそれぞれ最低 1 つ得られた時の，全収集パケット数を表す．例えば Savage らの手法ではマーキング情報を 8 つに分割しているため，8 つそれぞれのパケットを最低 1 つずつ集めるために全体で何パケット収集したか，ということになる．この値が小さいことは必要パケット数が少ないことを意味し，優れた手法であるといえる．

結果を見ると Static は他 2 手法と比較し非常に多くのパケットを必要とすることが分かる．その一方で Static2 は $d = 0 \sim 4$ の距離においては Fruit を下回っている．また傾向として Static と Dynamic は距離の増加に比例し必要パケット数が上昇している．Fruit は $d = 1 \sim 6$ では Dynamic より約 2 倍程度のパケットを必要とし，$d = 7 \sim 11$ では Dynamic のラインに肉薄し，$d = 12 \sim 20$ では下回っている．また Fruit の $d = 0$ の必要パケット数は p_{out} に大きく左右されていることが分かる．

期待値評価結果と同様に，他の攻撃者数の場合でも多少のバラつきはあるものの大方同様の傾向が見られたため，本論文では割愛する．

4.4 考察

4.2, 4.3 節で得られた結果について考察を行う．まず 4.2 節の期待値評価について，図 6,7 の到着確率の平均値のグラフにおける D-PPM と S-PPM の結果を見ると，それぞれ $d = 0, 1$ で高い到着確率となり，距離が増加するに連れてその値は減少している．2.2.2 項で述べたように D-PPM は S-PPM の課題である距離の増加に伴う到着確率の減少を防ぐため，動的なマーキング確率を利用することで全ての距離での到着確率を一定に保つことがその提唱動機の 1 つに挙げられていた．しかし今回の結果を見ると複数の攻撃者が存在する DDoS 攻撃のトポロジに対しては，D-PPM も S-PPM 同様に到着確率は距離の増加に伴い減少し，一定の値にならないということが分かった．

それに対し F-PPM の到着確率は距離の増加に大きく影響を受けないことが分かった．これにより，特に被害者から攻撃者までの距離が離れたトポロジにおいて D-PPM よりも効率的にマークパケットが収集可能であると考える．

また図 9 は攻撃者が 100 の攻撃トポロジの全パス長を調べ，累積率で表したグラフであり，これを見ると長さが 1〜9 のパス数の合計が全体の 90%を占めることが分かる．つまり 9 ホップのルータを経由することでインターネットの 90%以上のノードに到達可能であることを意味する．図 7 と合わせて見ると，特に $d = 8$ において F-PPM の到着確率は他 2 手法に比べ高く，より効率的なパケット収集が期待できる．

続いて 4.3 節のシミュレーション評価のグラフ，図 8 を見ると S-PPM と D-PPM の必要パケット数は距離の増加に比例して増えていることが分かり，これは期待値評価の結果とも一致していると考えられる．必要パケット数が距離の増加に比例する傾向はあるものの，D-PPM は S-PPM と比較しその必要パケット数はかなり抑えられている．

また F-PPM は距離の増加に伴う必要パケット数の変化は他の 2 手法と比較すると小さく，全体として平坦なグラフの形状を得ることができた．これは F-PPM を提案した動機の 1 つである，距離ごとのパケット収集効率の差を小さくするという目的をある程度果たせていると考えることができる．

しかし $d = 12$ 以降で D-PPM の必要パケット数を確実に下回っているものの，それ以前の距離では D-PPM や S-PPM を上回っている箇所も見られる．これは期待値評価の結果とも乖離しており，期待値評価をもとに $d = 5$ 以降は F-PPM は安定して D-PPM を下回るとした予想とは

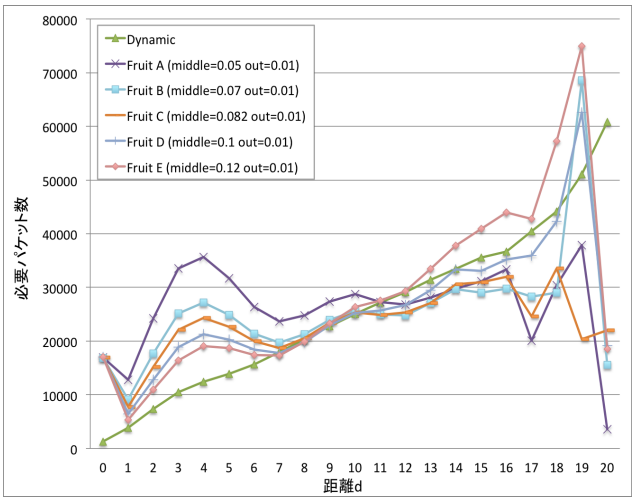

図 10 Fruit-structured PPM における p_{mid} の調節

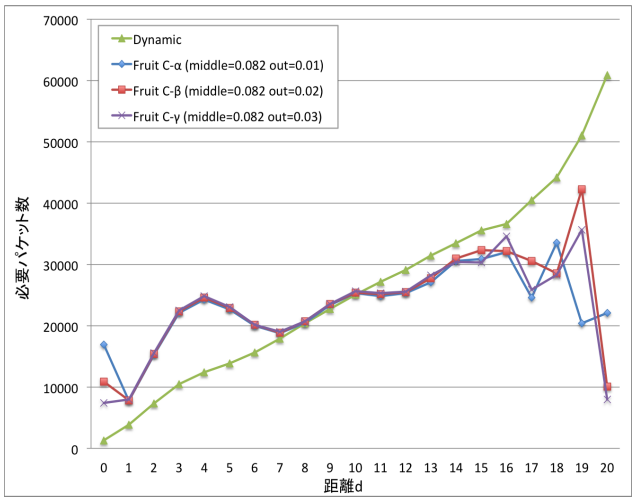

図 11 Fruit-structured PPM における p_{out} の調節

異なる結果となった．原因としては実際にパケットにマーキング行い収集する場合にはクーポンコレクタ問題が影響し，全てのオフセットをもつパケットを揃えるには予想外にパケット数が必要であるためと考える．

また，特に距離が小さい位置における必要パケット数は期待値評価の結果からも分かるように S-PPM を上回ることは十分に考えられるものの，もう少し距離が小さい段階で D-PPM の値を下回ることが望ましい．この評価において利用した F-PPM のマーキング確率は S-PPM での確率を参考にしており，F-PPM のマーキング確率を調節し，より適した確率を見つけることでさらに必要パケット数の減少が図れると考える．

5. 追評価

4 章の考察で F-PPM のマーキング確率のチューニングの必要性を述べた．そこで本章ではマーキング確率を調節した F-PPM で再度シミュレーション評価を行い，その結果について考察する．

5.1 追評価結果

4.1 節では F-PPM のマーキング確率を $p_{in} = 1$，$p_{mid} = \{0.05, 0.082, 0.1\}$，$p_{out} = \{0.01, 0.03, 0.05\}$ とし評価を行った．F-PPM のマーキング確率のうち p_{in} は 1 とすることで非マークパケットが発生しないという利点があるため，確率として動かすことができるのは p_{mid} と p_{out} になる．チューニングを行う方法としては，図 8 の結果から全体として大きく必要パケット数に影響するのは p_{mid} と推測し，まず p_{mid} の調節を行い，その後 p_{out} を調節する流れとした．

p_{mid} についていくつかのマーキング確率を試した中で $p_{mid} = \{0.05, 0.07, 0.082, 0.1, 0.12\}$ とした時の結果を図 10 に示す．Dynamic と Fruit A は図 8 と共通のものである．これを見ると p_{mid} を上昇させるにつれて $d = 1 \sim 7$ の必要パケット数が減少していることが分かる．逆に $d = 12$ 以降の値は上昇し，Dynamic を上回っている箇所も見られる．その中でも $p_{mid} = 0.082$ の Fruit C は d が小さい位置では Fruit A や Fruit B を下回り，$d = 12$ 以降も大幅な値の上昇は見られず，より平坦であるといえる．

続いて p_{out} の結果を示す．p_{mid} の値の中で比較的優れているとした $p_{mid} = 0.082$ について $p_{out} = \{0.01, 0.02, 0.03\}$ とした時の結果を図 11 に示す．このグラフより p_{out} が大きく影響するのは $d = 0$ と $d = 15$ 以降であり，中間部ではほとんど影響しないことが再確認された．$d = 0$ についてはマーキング確率が大きいほど必要パケット数は少ない．また $d = 15$ 以降については今回シミュレーションを行った 3 種の確率では優位な差は見られなかった．

5.2 考察

5.1 節の評価結果について考察を行う．まず p_{mid} の調節について距離 d が小さい段階において p_{mid} を上げると必要パケット数は減少するが，中間の $d = 7 \sim 10$ では F-PPM のグラフが D-PPM のグラフに肉薄するものの，下回ることはなかった．原因としては図 9 で見たように中間の長さのパスは数が多く，全オフセットを揃えるために必要なパケット数を平均した際に 2〜2.5 万パケットが下限になるためと考えられる．しかしながら必要パケット数がなるべく平坦であるという観点において，F-PPM の中でも特に $p_{mid} = 0.082$ については D-PPM よりも優れていると言うことができる．

次に p_{out} の調節について PPM ではあるルータの IP 情報を再構成するためにはその 1 つ前のルータの IP 情報との XOR を解く必要があり，特に $d = 0$ は攻撃経路再構成の起点となるためそのパケットは早く収集される必要がある．今回シミュレーションを行った $p_{out} = \{0.01, 0.02, 0.03\}$ では $d = 15$ 以降にも有意な差は見られなかったため，なるべく $d = 0$ の必要パケット数が少なくなるよう $p_{out} = 0.03$

が適していると考える．

今回追評価を行った中で F-PPM のマーキング確率として適しているのは $(p_{mid}, p_{out}) = (0.082, 0.03)$ であると考える．この値であれば必要パケット数は全体として平坦となり，また攻撃者から距離が離れても D-PPM より必要パケット数が少なくてすむ．

6. まとめ

本論文では IP トレースバック技術の 1 つである確率的パケットマーキング (PPM) について，既存の手法の課題とその評価における問題点を指摘した上で新たな PPM 手法として Fruit-structured PPM の提案を行い，また実際のインターネットルータトポロジを利用した現実的な評価手法によって評価を行った．

既存手法の課題である到着確率の距離ごとのバラつきを抑えるため，Fruit-structured PPM ではインターネットとユーザとの接点に着目し，そこに位置するルータでのマーキング確率を工夫した．また被害者を中心に，攻撃者を被害者の周囲に配置した論理的なモデルは果実のように見える．インターネットの内部 (果実) で中継されるパケットへマーキングする確率 (p_{mid}) に対し攻撃者からインターネットへパケットを転送するルータ (皮に位置する) のマーキング確率 (p_{in}) を高め，インターネットから被害者へパケットを転送するルータ (種に位置する) のマーキング確率 (p_{out}) を低くする．これにより Savage らが提案した PPM と同様の実装上の負荷のもと，被害者から遠くに位置するルータにおける到着確率を高く保つことができ，結果として再構成に必要なパケット数を少なくすることができる．また $p_{in} = 1$ とすることで再構成に大きな影響を及ぼす非マークパケットが発生しないようにすることが可能である．

また CAIDA のトポロジデータを用いた評価として，35 万の攻撃トポロジを用意し期待値評価とシミュレーション評価を行った．前者では Dynamic PPM の到着確率は一定にならないことが確認できたと同時に，Fruit-structured PPM が高い到着確率を有していることが分かった．また後者では期待値評価ほどの優位性は見られなかったものの，Fruit-structured PPM の必要パケット数が Static PPM，Dynamic PPM と比較し平坦なことが示された．

さらに Fruit-structured PPM のマーキングの調整を行い，再度シミュレーション評価を行った．その結果から今回行ったシミュレーションの中では $(p_{mid}, p_{out}) = (0.082, 0.03)$ とすることで必要パケット数は攻撃者からの距離全体においてより平坦になり，また距離が離れた位置においても Dynamic PPM を下回る結果となった．

参考文献

[1] Savage, S., Wetherall, D., Karlin, A. and Anderson, T. :Practical network support for IP Traceback, *Proc. ACM SIGCOMM '00*, pp.295-306 (2000).

[2] Goodrich, M.T. :Probabilistic Packet Marking for Large-Scale IP Traceback, IEEE/ACM TRANSACTIONS ON NETWORKING, Vol.16, No.1, pp.15-24 (2008).

[3] Song, D. and Perrig, A. :Advanced and Authenticated Marking Schemes for IP Tracebask, *Proc. IEEE INFOCOM*, pp.876-886, (2001).

[4] Dean, D., Franklin, M. and Stubblefield, A. :An Algebraic Approach to IP Traceback, *Proc. Network and Distributed System Security Symp. (NDSS)*, pp.3-12 (2001).

[5] Law, T.K.T., Yau, D.K.Y. and Lui, J.C.S. :You can run, but you can't hide: An effective statistical methodology to trace back DDoS attackers, *IEEE Trans. Parallel Distrib. Syst*, Vol.16, No.9, pp.799-813 (2005).

[6] 金岡 晃, 岡田 雅之, 勝野 恭治, 岡本 栄司 :DoS 攻撃経路を効率的に再構成するため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, 情報処理学会論文誌 Vol.52, No.9, pp.2718-2728 (2011).

[7] 岡田 雅之, 金岡 晃, 勝野 恭治, 岡本 栄司 :確率的パケットマーキングにおける最適マーキング確率の推定, 情報処理学会論文誌, Vol.52, No.3, pp.929-939 (2011).

[8] 金岡 晃, 岡田 雅之, 岡本 栄司 :確率的パケットマーキングの実用化検討, *Proc.* コンピュータセキュリティシンポジウム 2011 年 論文集, Vol.2011, No.3, pp.618-623 (2011).

[9] Okada, M., Kanaoka, A., Katsuno, Y. and Okamoto, E. :32-bit AS Number Based IP Traceback, *Proc. In Proceeding of the 5th International Workshop on Advances in Information Security (WAIS 2011)*, pp.628-633, (2011).

[10] 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, pp.3197-3201 (2003).

[11] Liu, J., Lee, Z-J., and Chung, Y-C. :Dynamic probabilistic packet marking for efficient IP trace- back, *Proc. Computer Networks*, Vol.51, Issue.3, pp. 866-882 (2007).

[12] Yan, Q., He, X. and Ning, T. :An Improved Dynamic Probabilistic Packet Marking for IP Traceback, *Proc. I.J.Computer Network and Information Security*, Vol.2, No.2, pp.47-53 (2010).

[13] Tian, H., Bi, J. and Jiang, X. :An adaptive probabilistic marking scheme for fast and secure traceback, *Proc. Networking Science*, Vol.2, Issue.1-2, pp.42-51 (2010).

[14] CAIDA: The Cooperative Association for Internet Data Analysis, available from ⟨http://www.caida.org/home/⟩, (2003).