

ハッシュフラグメント方式によるIPトレースバックの特性分析

塩田茂雄^{†1,†2} 池田 徹^{†1}

DDOS攻撃の攻撃経路を検出するIPトレースバック技術の一つに、ルータが通過パケットにIPアドレスの部分情報を書き込むパケットマーキングがある。本研究では、IPアドレスのハッシュ値の分割（フラグメント）をパケットに記載するハッシュフラグメント方式に着目し、ハッシュフラグメント方式の各種特性（攻撃経路の検出漏れ確率、誤検出確率）を評価し、その最適パラメタ設定について考察する。

Performance Analysis of Hash Fragment Marking Schemes for IP Traceback

SHIGEO SHIODA and TORU IKEDA

Probabilistic packet marking (PPM) is an IP traceback technique that lets routers probabilistically mark packets with partial information of an attack path during packet forwarding. In this paper, we theoretically evaluate the efficiency of *hash fragment marking scheme*, where a router pre-calculates a hash of its IP address and marks a fragment of the hash into the IP header of a packet. Note that the hash fragment marking scheme is a very general PPM model that covers most of existing PPM schemes proposed so far. We derive an explicit representation for false positive ratio (FPR) and false negative ratio (FNR) of the hash fragment marking scheme regarding the attack path detection. We also explain how to set parameters values to satisfy given requirements for the FPR and the FNR.

1. はじめに

インターネットにおいて、大量のパケットをサーバに送信し、正常なサービスの提供を妨害する「サービス妨害攻撃（Denial of Service Attack）」が顕在化している。サービス妨害攻撃の攻撃経路検出技術であるIPトレースバックの一つに、ルータが通過するパケットに確率的にルータ情報を書き込む（マーキングする）パケットマーキングがある。被害ノードは、攻撃パケットに書き込まれたルータ情報から、攻撃経路を推定する。これまでに、マーキング方法が異なる様々なパケットマーキング方式が提案されており^{3),5),8),9)}、これら既存方式の得失を横断的に評価し、パケットマーキングの性能の本質的な理解を得ることが、より高効率なIPトレースバック方式を実現する上で必要である。

パケットマーキングでは、通常、ルータの情報はパケットのIPヘッダのidentificationフィールドに書き込まれる。identificationフィールドは16bitであり、ルータのIPアドレス（IPv4の場合32bit）より小さいため、どのような工夫によりIPアドレス情報をidentificationフィールドに記載するかが効率的なIPトレースバッ

ク方式を実現する上での鍵となる。この点において、多くの既存提案は

- (1) ハッシュ関数を用いてIPアドレスを縮約して記載する手法^{8),9)}
 - (2) IPアドレス（のハッシュ値）を複数の断片に分割し、断片のいずれか一つを記載する手法^{5),9)}
- のいずれか、もしくは両者を組み合わせた方式を採用している。本研究では、この両方の特徴を備え、既存提案の多くを包括する一般的な方式モデル（本稿ではこれをハッシュフラグメント方式と呼ぶ）を用いて、IPトレースバック方式の効率性（攻撃経路の検出漏れ確率、ルータの誤検出確率）を解析的に評価するとともに、その最適パラメタ設定について考察する。

なお、各種IPトレースバック方式の性能評価に関する文献は他にも存在するが^{2),4),10),11)}、パケットマーキングは評価対象でない、攻撃経路数が数個のケースのみ扱っている、シミュレーション評価を用いている、等の点で本研究とは異なる。

以下、本稿の構成を述べる。まず、2章では、ハッシュフラグメント方式について説明する。次いで、3章において、ハッシュフラグメント方式の解析モデルを示し、経路検出に要する受信パケット数の分布関数を導出する。4章では、3章の結果を用いて、ハッシュフラグメント方式の各種検出効率（攻撃経路の検出漏れ確率、誤検出確率）の表式を示し、これら検出効率

†1 千葉大学大学院工学研究科 建築・都市科学専攻

Department of Architecture and Urban Science, Graduate School of Engineering, Chiba University

†2 shioda@faculty.chiba-u.jp

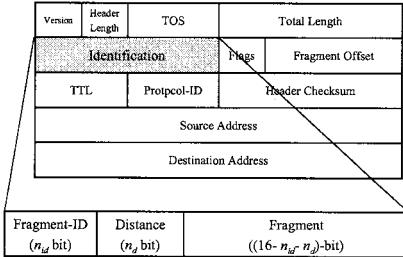


図 1 identification の利用法

が定められた基準を満たすようにパラメタを設定する手法について考察する。5 章では数値結果を示し、6 章で本研究のまとめを行う。

2. ハッシュフラグメント方式

本章では「ハッシュフラグメント方式」の概要を説明する。以下では IPv4 が用いられる場合を仮定する。

2.1 マーキング方法

各ルータは SHA-1 等のハッシュ関数により自分の IP アドレスを 32bit 以上のお情報を (SHA-1 の場合は 160bit) に変換し、その情報を幾つかの断片 (フラグメント) に分割して保持する。さらに、通過する一定の割合のパケットに、いずれかのフラグメントを確率的に書き込む (マーキングする)。

フラグメントは IP ヘッダの identification フィールド (16bit) に書き込まれる。ハッシュフラグメント方式では、identification フィールドの先頭の n_{id} bit (以下、fragment-ID フィールドと呼ぶ) に何番目のフラグメントかを表す情報を、次の n_d bit (distance フィールドと呼ぶ) にマーキングルータの被害ノードからのホップ数を、最後の $(16 - n_d - n_{id})$ bit (fragment フィールドと呼ぶ) にフラグメントの内容を記載する (図 1)。ルータがフラグメントを書き込む際には、必ず distance フィールドの値を 0 に初期化する。また、ルータがパケットにフラグメントを書き込まざるを得ない場合は、distance フィールドの値を 1 増やす。これにより、パケットが被害ノードに到着した際の distance フィールドの値は、マーキングルータと被害ノード間の距離 (ホップ数) から 1 を引いたものに等しくなる。

インターネットの最大ホップ数は 30 程度であるとされている。従って distance フィールド長は 5 程度で充分である ($30 < 2^5 = 32$)。なお、Yaar らは TTL の情報を用いることで、distance フィールドが 1bit であっても被害ノードからの距離を特定できる手法を示しており⁹⁾、彼らの手法もしくは類似の手法を用いれば distance フィールド長は 5 より小さい値に設定可能である。一方、fragment-ID フィールド長は任意である。例えば、fragment-ID フィールドが 3bit、distance

フィールドが 5bit の場合、ルータは 8 (= 16 - 3 - 5) bit のフラグメントを全部で $8 (= 2^3)$ 種類利用できる。SHA-1 を利用するならば、ルータは自分の IP アドレスを (SHA-1 により) 160bit のハッシュ値に変換し、例えればその上位 $64 (= 8 \times 8)$ bit を (8bit ずつの) 8 つのフラグメントに分割して、いずれかをパケットにマーキングすれば良い。IP アドレスは 32bit であるから、二つ以上の異なる IP アドレスの SHA-1 の上位 64bit が偶然に一致する確率は極めて小さい。従って、被害ノードは 8 つのフラグメントを全て受信することで、通過ルータをほぼ一意に特定できることとなる。なお、この例の場合、被害ノードはルータを特定するために全てのフラグメントを受信する必要はない。例えればいずれか 6 つのフラグメントを受信すれば、通過ルータに関する 48bit の情報が得られるので、やはりほぼ一意に通過ルータを特定できる。この点については後述する。

一般に、fragment-ID フィールドが n_{id} bit、distance フィールドが n_d bit の場合、ルータはの IP アドレスのハッシュ値の上位 $w(n_{id}, n_d) \stackrel{\text{def}}{=} 2^{n_{id}}(16 - n_d - n_{id})$ bit を利用して、パケットマーキングを行う。ただし、 $w(n_{id}, n_d)$ がハッシュ値の出力を上回る場合、利用するフラグメント数は $\lfloor H_{out} / (16 - n_d - n_{id}) \rfloor$ に制限されるとする (H_{out} はハッシュ関数の出力 bit 長、 $\lfloor x \rfloor$ は x 以下の最大の整数)。つまりフラグメント数 f は

$$f = \min\{2^{n_{id}}, \lfloor H_{out} / (16 - n_d - n_{id}) \rfloor\} \quad (1)$$

に等しい*1。

ルータが通過パケットにフラグメントを書き込む際に、当該パケットに既に前段ルータの情報を書き込まれている場合、前段ルータの情報は上書きされて消える。攻撃経路が一つで、マーキング確率がルータによらず一定値 p の場合、被害ノードから距離 h のルータの情報が攻撃パケットにマーキングされている確率は

$$p(1 - p)^{h-1},$$

に等しい。すなわち、被害ノードから遠いルータほど、その情報が受信パケットに書き込まれている確率は減少する。

2.2 攻撃経路検出方法

ハッシュフラグメント方式では、被害ノードは受信したパケットの identification フィールドに書き込まれたフラグメントの情報をから、攻撃経路を推定する。被害ノードはインターネットの地図を持ち、地図にはインターネット内の各ルータについて、その IP アドレス及び自ノードからの距離 (ホップ数) の情報を記載されているものとする。

被害ノードは攻撃パケットを受信すると、identification フィールドに書き込まれた情報を読み取る。今、攻

*1 $w(n_{id}, n_d)$ がハッシュ値の出力を上回る場合でも、ハッシュの出力を元の入力に連結させて再度ハッシュの出力をとり、最初の出力と連結させることで、当初のハッシュ値の整数倍の出力を得ることは可能であるが⁹⁾、本稿ではこのケースは扱わない。

撃パケットに書き込まれた fragment-ID フィールドの値が i , distance フィールドの値が h , fragment フィールドの値が F であったとする。被害ノードは地図から被害ノードから $h+1$ ホップの距離にあるルータを選び出し、その IP アドレスのハッシュ値の i 番目のフラグメントが F に一致するかどうかを調べる。もし、一致するならば、そのルータの i 番目のフラグメントは検出済みとする。この手順をパケットを受信する度に繰り返し、 f_{th} 個以上のフラグメントが検出されたルータを攻撃経路上のルータとして検出する。

既に述べたように、fragment-ID フィールド長が $n_{id}\text{bit}$ の場合、最大 $2^{n_{id}}$ 個のフラグメントを利用できるが、攻撃経路上のルータの検出に全てのフラグメントの検出が必要とは限らない。fragment-ID フィールド長が $n_{id}\text{bit}$ 、distance フィールドが $n_d\text{bit}$ であれば、フラグメント長は $(16 - n_{id} - n_d)\text{bit}$ となるので、 f 個のフラグメントが検出されれば、そのルータについて $f(16 - n_{id} - n_d)\text{bit}$ の情報が入手できることになる。従って $f_{th}(16 - n_{id} - n_d) \gg 2^{32}$ であれば、誤検出の可能性は充分小さいと考えてよい。4 章において、誤検出確率を定められた基準以下に抑えるように f_{th} を定める手法について述べる。

2.3 既存方式との関係

上述したハッシュフラグメント方式は、既存の提案方式の幾つかを包含する。例えば、Song らが提案する（1種類の 11bit ハッシュ関数を用いる）方式⁸⁾ は、 $n_{id} = 0$, $n_d = 5$ としたハッシュフラグメント方式に該当する。Song らは 8 (= 2³) 種類の（8bit ハッシュ関数を用いる）方式も提案しているが⁸⁾、これは $n_{id} = 3$, $n_d = 5$ としたハッシュフラグメント方式に該当する。また Yaar の提案する方式（Fast Internet Traceback: FIT）⁹⁾ は $n_d = 1$ に制限している点を除けば、ハッシュフラグメント方式と等価である。彼らは、 $\{n_{id} = 2, f_{th} = 3\}$, $\{n_{id} = 2, f_{th} = 4\}$, $\{n_{id} = 3, f_{th} = 5\}$ のパラメタの組み合わせを用いた場合の性能をシミュレーションにより比較・評価している⁹⁾。なお、我々は、以前 Savage らが提案するフラグメントマーキング方式⁵⁾、Song らの方式⁸⁾、Yang らの方式⁹⁾を性能を一次評価しており⁷⁾、本研究はこの評価をより精密化したものに相当する。

3. 攻撃経路特定に必要なパケット数

本章では、ハッシュフラグメント方式の性能尺度の一つとして、攻撃経路を特定するまでに受信する必要のある攻撃パケット数 X について考察する。 X は確率変数であり、ここではその補分布を導出する。

3.1 一般的結果

1 つの攻撃ノードから 1 つの被害ノードに対して DoS 攻撃が仕掛けられているケースを考察する。攻撃ノードと被害ノード間を結ぶ経路上には N 台のルータが存在するとし、被害ノードに近い方から順にルータ

タ 1, ルータ 2, … と番号をつける。攻撃経路を構成する全てのルータの検出に要するパケット数を X とする。攻撃パケットを k 個受信したが、ルータ n が未検出である事象を $A_n^{(k)}$ で表す。さらに、 J を $\{1, 2, \dots, N\}$ の部分集合、 $|J|$ を J の要素数とし、

$$S_J^{(k)} = \bigcap_{n \in J} A_n^{(k)},$$

を定義する。

補題 3.1.

$$P[X > k] = \sum_{n=1}^N (-1)^{n+1} \sum_{J: |J|=n} P[S_J^{(k)}]. \quad (2)$$

Proof. Inclusion-exclusion principle¹⁾ より

$$P\left[\bigcup_{n=1}^N A_n^{(k)}\right] = \sum_{n=1}^N (-1)^{n+1} \sum_{J: |J|=n} P[S_J^{(k)}].$$

従って、

$$\begin{aligned} P[X > k] &= P\left[\bigcup_{n=1}^N A_n^{(k)}\right] \\ &= \sum_{n=1}^N (-1)^{n+1} \sum_{J: |J|=n} P[S_J^{(k)}]. \end{aligned}$$

以上、証明された。 \square

一般に X の分布の数値的な評価は困難であることが知られている。次の結果⁶⁾ は X の補分布 ($P[X > k]$) の数値的に評価可能な上下限を導くときに用いられる。

補題 3.2.

$$\sum_{J: |J|=1} P[S_J^{(k)}] - \sum_{J: |J|=2} P[S_J^{(k)}] \leq P[X > k] \leq \sum_{J: |J|=1} P[S_J^{(k)}]. \quad (3)$$

3.2 $P[S_J^{(k)}]$ の導出

式(3)からわかるように、 X の分布特性の計算には、 $P[S_J^{(k)}]$ の解析的な評価が必要になる。 $S_J^{(k)}$ は、例えば $J = \{1, 2\}$ のとき、 k 個の攻撃パケットを受信した時点でルータ 1 とルータ 2 が未検出である事象に対応する。本節では、 $P[S_J^{(k)}]$ の表式を導出する。

以下では、被害ノードが受信するパケットにルータ n の情報のいずれかのフラグメントが書き込まれている確率を p_n で表す。2 章で述べたように、マーキング確率がルータによらず一定値 p の場合、攻撃経路が一つであれば

$$p_n = p(1 - p)^{n-1}.$$

注. 攻撃経路が複数 (R 本) あり、各攻撃経路の重複が存在しない（複数の攻撃経路上に重複して存在するルータがない）ならば

$$p_n = p(1 - p)^{n-1}/R.$$

3.2.1 ルータ検出に全フラグメントの受信が必要な場合

まず、ルータの検出にルータ情報の全フラグメントの受信が必要な場合を考察する。フラグメント数を $f \stackrel{\text{def}}{=} \min\{2^{n_d}, \lfloor H_{out}/(16 - n_d - n_{id}) \rfloor\}$ で表す。 I を N 個の要素を持ち、各要素が 0 以上 f 以下のいずれかの整

数を取るベクトルとし, \mathbf{l} の 0 以外の値をとる要素の番号からなる集合を $\mathcal{E}(\mathbf{l})$ で表す. 次が成立する.

補題 3.3.

$$P[S_J^{(k)}] = \sum_{\mathbf{l}: \mathcal{E}(\mathbf{l})=J} (-1)^{\sum_{n=1}^N l_n - |J|} \left(\prod_{n=1}^N \binom{f}{l_n} \right) \left(1 - \frac{\sum_{n=1}^N l_n p_n}{f} \right)^k.$$

Proof. 付録 A.1 参照. \square

補題 3.3 より

$$\begin{aligned} P[S_{[n]}^{(k)}] &= \sum_{l=1}^f (-1)^{l-1} \binom{f}{l} \left(1 - \frac{lp_n}{f} \right)^k, \\ P[S_{[m,n]}^{(k)}] &= \sum_{l_m=1}^f \sum_{l_n=1}^f (-1)^{l_m+l_n-2} \binom{f}{l_m} \binom{f}{l_n} \left(1 - \frac{l_m p_m + l_n p_n}{f} \right)^k. \end{aligned} \quad (4)$$

これら結果と補題 3.2 より, $P[X > k]$ の上下限が次のように評価できる.

$$\begin{aligned} P[X > k] &\leq \sum_{n=1}^N \sum_{l=1}^f (-1)^{l-1} \binom{f}{l} \left(1 - \frac{lp_n}{f} \right)^k. \\ P[X > k] &\geq \sum_{n=1}^N \sum_{l=1}^f (-1)^{l-1} \binom{f}{l} \left(1 - \frac{lp_n}{f} \right)^k \\ &\quad - \sum_{m=1}^N \sum_{n>m} \sum_{l_m=1}^f \sum_{l_n=1}^f (-1)^{l_m+l_n-2} \binom{f}{l_m} \binom{f}{l_n} \left(1 - \frac{l_m p_m + l_n p_n}{f} \right)^k. \end{aligned} \quad (5)$$

また, 補題 3.3 の証明 (付録 A.1) の途中で次の関係式が得られる.

$$P[X > k] = \sum_{\mathbf{l} \neq \mathbf{0}} (-1)^{\sum_{n=1}^N l_n + 1} \left(\prod_{n=1}^N \binom{f}{l_n} \right) \left(1 - \frac{\sum_{n=1}^N l_n p_n}{f} \right)^k. \quad (6)$$

従って, X の期待値は

$$\begin{aligned} E[X] &= \sum_{k=0}^{\infty} P[X > k] \\ &= \sum_{\mathbf{l} \neq \mathbf{0}} (-1)^{\sum_{n=1}^N l_n + 1} \left(\prod_{n=1}^N \binom{f}{l_n} \right) \frac{f}{\sum_{n=1}^N l_n p_n}, \end{aligned} \quad (7)$$

であり, 以下の積分表示を有することが確かめられる.

$$E[X] = \int_0^{\infty} \left(1 - \prod_{n=1}^N \left(1 - e^{-p_n t/f} \right)^{l_n} \right)^f dt.$$

積分表示は数値計算の際に便利である.

3.2.2 一部のフラグメントを受信すれば良い場合

次に, f 個のフラグメントのうちいずれか f_{th} 個受信すればリンクの検出ができるケースを考える. 次が成立する.

$$P[X > k] = P\left[\bigcup_{n=1}^N F_n; |F_n|=f-f_{th}+1 \right] \bigcup_{l \in F_n} A_{F_n}^{(k)},$$

ここで

$$A_{F_n}^{(k)} \stackrel{\text{def}}{=} \bigcap_{l \in F_n} A_{nl}^{(k)}.$$

であり, $A_{nl}^{(k)}$ は攻撃パケットを k 個受信したが, リンク n の l 番目のフラグメントが未検出である事象を表

す. 次を定義する

$$\mathcal{F}(i) \stackrel{\text{def}}{=} \{F_1 \times \cdots \times F_N; |F_n| \geq i \text{ or } |F_n| = 0, n = 1, \dots, N\},$$

$$\mathcal{L}(i) \stackrel{\text{def}}{=} \{(l_1, \dots, l_n); l_n \geq i \text{ or } l_n = 0, n = 1, \dots, N\},$$

$$\mathcal{F}_\uparrow(i) \stackrel{\text{def}}{=} \mathcal{F}(i) \setminus \emptyset, \quad \mathcal{L}_\uparrow(i) \stackrel{\text{def}}{=} \mathcal{L}(i) \setminus \emptyset.$$

補題 3.4.

$$\begin{aligned} P[S_J^{(k)}] &= \sum_{\mathbf{l} \in \mathcal{L}_\uparrow(f-f_{th}+1), \mathcal{E}(\mathbf{l})=J} (-1)^{|J|} \\ &\quad \times \prod_{n=1; l_n \geq f-f_{th}+1}^N \left\{ (-1)^{l_n-f+f_{th}} \binom{l_n-1}{f-f_{th}} \binom{f}{l_n} \right\} \left(1 - \frac{\sum_{n=1}^N l_n p_n}{f} \right)^k. \end{aligned}$$

Proof. 付録 A.2 参照. \square

特に,

$$P[S_{[n]}^{(k)}] = \sum_{l=f-f_{th}+1}^f (-1)^{l-f+f_{th}+1} \binom{l-1}{f-f_{th}} \binom{f}{l} \left(1 - \frac{lp_n}{f} \right)^k,$$

及び

$$\begin{aligned} P[S_{[m,n]}^{(k)}] &= \sum_{l_m=f-f_{th}+1}^f \sum_{l_n=f-f_{th}+1}^f (-1)^{(l_n-f+f_{th}+1)+(l_m-f+f_{th}+1)} \\ &\quad \times \binom{l_n-1}{f-f_{th}} \binom{l_n-1}{l_n} \binom{f}{l_n} \left(1 - \frac{l_m p_m + l_n p_n}{f} \right)^k. \end{aligned}$$

これら結果と補題 3.2 の結果より, $P[X > k]$ の上下限が評価できる. また, 補題 3.4 の証明 (付録 A.2) の途中で得られた X の補分布の表式

$$\begin{aligned} P[X > k] &= \sum_{\mathbf{l} \in \mathcal{L}_\uparrow(f-f_{th}+1)} (-1) \prod_{n=1; l_n \geq f-f_{th}+1}^N \left\{ (-1)^{l_n-f+f_{th}} \binom{l_n-1}{f-f_{th}} \binom{f}{l_n} \right\} \\ &\quad \times \left(1 - \frac{\sum_{n=1}^N l_n p_n}{f} \right)^k. \end{aligned}$$

を利用すると, X の期待値が以下のように求まる.

$$\begin{aligned} E[X] &= \sum_{\mathbf{l} \in \mathcal{L}_\uparrow(f-f_{th}+1)} (-1) \\ &\quad \times \prod_{n=1}^N \left\{ (-1)^{l_n-f+f_{th}} \binom{l_n-1}{f-f_{th}} \binom{f}{l_n} \right\} \frac{f}{\sum_{n=1}^N l_n p_n}. \end{aligned}$$

4. ハッシュフラグメント方式の検出効率

本章では, ハッシュフラグメント方式の検出効率を評価するため, 検出漏れ確率および誤検出確率の評価式を導出し, 検出効率を最適化するためのパラメタ設定について考察する.

4.1 検出漏れ確率

R 本の攻撃経路から k 個の攻撃パケットを受信した時点で, 未検出の攻撃経路数の期待値を $R_{ud}(k)$ で表す. これを用いて検出漏れ確率 (False Negative Ratio: FNR) を,

$$FNR \stackrel{\text{def}}{=} \frac{R_{ud}(k)}{R},$$

により定義し, $P_{fn}(k)$ で表す. 攻撃経路 r の特定に必要なパケット数を X_r とすると

$$P_{fn}(k) = \frac{\sum_{r=1}^R P[X_r > k]}{R}.$$

もし, $\{X_1, \dots, X_R\}$ が互いに独立で同一の分布に従えば $P_{fn}(k) = P[X_1 > k]$.

表 1 パラメタの組み合わせの例 ($n_d = 1, P_{fp} < 0.001$)

Fragment-ID (n_{id})	2	3	4	5	6	7	8	9	10
# of fragments (f)	4	8	14	16	17	20	22	26	32
Fragment length	13	12	11	10	9	8	7	6	5
f_{th}	4	5	5	6	7	8	9	11	15

すなわち、3章で求めた X の補分布から検出漏れ確率が計算できる。

注. 上で述べたように、検出漏れ確率は各攻撃経路の検出に要する攻撃パケット数の補分布から求まる。攻撃経路数は

$$p_n = p(1-p)^{n-1}/R,$$

を通して検出漏れ確率に影響する。

4.2 誤検出確率

次に、誤検出確率 (False Positive Ratio: FPR) を「検出した総ルータ数 (誤検出分も含む)」に対する「誤検出した総ルータ数」の比で定義し、 P_{fp} で表す。攻撃経路上のルータ (ルータ A) の IP アドレスのハッシュ値の i 番目のフラグメントと、任意に選んだルータ (ルータ B) の IP アドレスのハッシュ値の i 番目のフラグメントが一致する確率は i によらず

$$\frac{1}{2^{(16-n_{id}-n_d)}},$$

に等しい。ルータ B がルータ A と誤認識されるためには、少なくとも f_{th} 個のフラグメントが一致しなければならない。この確率は

$$p_0 \stackrel{\text{def}}{=} \sum_{i=f_{th}}^f \binom{f}{i} \left(\frac{1}{2^{(16-n_{id}-n_d)}} \right)^i \left(1 - \frac{1}{2^{(16-n_{id}-n_d)}} \right)^{f-i},$$

に等しい。ルータ A と誤認識されるルータ数は最大 $p_0 2^{32}$ であるから、

$$P_{fp} \leq \frac{p_0 2^{32}}{1 + p_0 2^{32}}.$$

4.3 パラメタ設定法

一般に、検出漏れ確率と誤検出確率はトレードオフの関係にあり、一方を小さくすると他方は大きくなる。そこで、本節では誤検出確率を基準値以下に抑えながら、検出漏れ確率を最小化するパラメタ設定について検討する。まず、誤検出確率の基準値を P_{fp}^h とすると

$$\frac{p_0 2^{32}}{1 + p_0 2^{32}} \leq P_{fp}^h \quad (8)$$

を満たすように f_{th} を決定すれば、誤検出確率は基準値 P_{fp}^h を下回ることが保障される。 $n_d = 1$ と $n_d = 5$ の場合それぞれについて、 $n_{id}+n_d < 16$ を満たすように n_{id} を選び、次いで(1)により f を決定し、さらに(8)を満たす最小の f_{th} を選んだ結果を、表 1 ($n_d = 1, P_{fp} < 0.001$)、表 2 ($n_d = 1, P_{fp} < 0.0001$)、表 3 ($n_d = 5, P_{fp} < 0.001$)、表 4 ($n_d = 5, P_{fp} < 0.0001$) に示す (SHA-1 の利用を仮定し、 $H_{out} = 160$ とした)。これらパラメタ候補のうち、検出漏れ確率を最小化するパラメタの組が、誤検出確率・検出漏れ確率の点で最適なパラメタ設定となる。次章では最適なパラメタの組を数値的に調査する。

表 2 パラメタの組み合わせの例 ($n_d = 1, P_{fp} < 0.0001$)

Fragment-ID (n_{id})	2	3	4	5	6	7	8	9	10
# of fragments (f)	4	8	14	16	17	20	22	26	32
Fragment length	13	12	11	10	9	8	7	6	5
f_{th}	4	5	6	6	7	8	10	12	15

表 3 パラメタの組み合わせの例 ($n_d = 5, P_{fp} < 0.001$)

Fragment-ID (n_{id})	3	4	5	6	7	8	9
# of fragments (f)	8	16	26	32	40	53	80
Fragment length	8	7	6	5	4	3	2
f_{th}	6	8	11	15	20	29	52

表 4 パラメタの組み合わせの例 ($n_d = 5, P_{fp} < 0.0001$)

Fragment-ID (n_{id})	3	4	5	6	7	8	9
# of fragments (f)	8	16	26	32	40	53	80
Fragment length	8	7	6	5	4	3	2
f_{th}	7	9	12	15	21	31	53

5. 数値評価

5.1 検出漏れ確率

5章の表 1, 2, 3, 4 で示したパラメタの組み合わせの幾つかについて検出漏れ確率を評価した結果を図 2 ($n_d = 1, P_{fp} < 0.001$)、図 3 ($n_d = 1, P_{fp} < 0.0001$)、図 4 ($n_d = 5, P_{fp} < 0.001$)、図 5 ($n_d = 5, P_{fp} < 0.0001$) に示す。攻撃経路数は 1、各ルータのマーキング確率は 0.04、経路上のルータ数は 20 とした。各図において、実線は補題 3.2 を用いて $P[X > k]$ の上限を評価したもの、点線は下限を評価したものを表す。図から確認できるように、上下限の差は無視できるほど小さく、上限式 (もしくは下限式) により充分高い精度で検出漏れ確率を評価できることがわかる。

検出漏れ確率はパラメタの組み合わせに大きく依存し、検出漏れ確率の点で有利な組み合わせが存在する。例えば、 $n_{id} = 1$ の場合 (図 2 と図 3) は、 $\{n_{id} = 5, f_{th} = 6\}$ の組み合わせが比較した中では最も有利であった。

5.2 検出漏れ確率を x 以下とする受信パケット数

最適なパラメタの組を抽出するため、検出漏れ確率が x 以下に収まる受信パケット数 (以下、 $P_{fn}^{(-1)}(x)$ で表す) を算出し、それらを比較した結果を、表 5 ($n_d = 1, P_{fp} < 0.001$)、表 6 ($n_d = 1, P_{fp} < 0.0001$)、表 7 ($n_d = 5, P_{fp} < 0.001$)、表 8 ($n_d = 5, P_{fp} < 0.0001$) に示した。

表からわかるように、 $P_{fn}^{(-1)}(0.01)$ を最小化するパラメタの組み合わせは以下のとおりになる。

(1) $n_d = 1, P_{fp} < 0.001$ の場合: $n_{id} = 4, f_{th} = 5$

(2) $n_d = 1, P_{fp} < 0.0001$ の場合: $n_{id} = 5, f_{th} = 6$

(3) $n_d = 5, P_{fp} < 0.001$ の場合: $n_{id} = 4, f_{th} = 8$

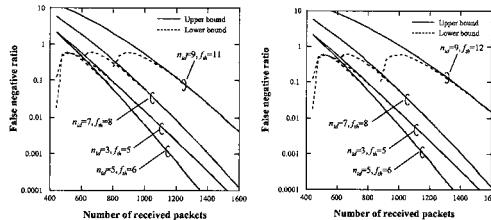


図 2 検出漏れ確率
($R = 1, n_d = 1, P_{fp} < 0.001$)

図 3 検出漏れ確率
($R = 1, n_d = 1, P_{fp} < 0.0001$)

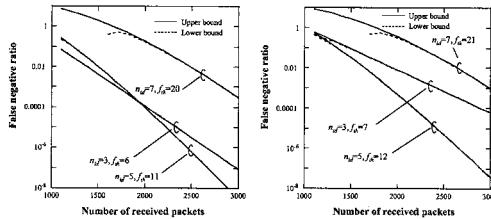


図 4 検出漏れ確率
($R = 1, n_d = 5, P_{fp} < 0.001$)

図 5 検出漏れ確率
($R = 1, n_d = 5, P_{fp} < 0.0001$)

表 5 $P_{fn}^{(-1)}(0.01)$ ($R = 1, n_d = 1, P_{fp} < 0.001$)

n_{id}	2	3	4	5	6	7	8	9	10
f_{th}	4	5	5	6	7	8	9	11	15

表 6 $P_{fn}^{(-1)}(0.01)$ ($R = 1, n_d = 1, P_{fp} < 0.0001$)

n_{id}	2	3	4	5	6	7	8	9	10
f_{th}	4	5	6	6	7	8	10	12	15

表 7 $P_{fn}^{(-1)}(0.01)$ ($R = 1, n_d = 5, P_{fp} < 0.001$)

n_{id}	3	4	5	6	7	8	9
f_{th}	6	8	11	15	20	29	52

表 8 $P_{fn}^{(-1)}(0.01)$ ($R = 1, n_d = 5, P_{fp} < 0.0001$)

n_{id}	3	4	5	6	7	8	9
f_{th}	7	9	12	15	21	31	53

(4) $n_d = 5, P_{fp} < 0.0001$ の場合: $n_{id} = 4, f_{th} = 9$

一方、表 9, 10, 11, 12 には $P_{fn}^{(-1)}(0.001)$ を比較して示した。 $P_{fn}^{(-1)}(0.001)$ を最小化するパラメタの組み合わせは、 $P_{fn}^{(-1)}(0.01)$ を最小化するパラメタの組み合わせに等しい。

この最適パラメタ構成は攻撃経路数によらない。参考までに、攻撃経路数が 100 のときの $P_{fn}^{(-1)}(0.001)$ を表 13, 14, 15, 16 に示す。

5.3 既存方式との比較

最後に、前節で得られた最適なパラメタの組み合わせを用いた場合の検出性能を、既存の提案方式と

表 9 $P_{fn}^{(-1)}(0.001)$ ($R = 1, n_d = 1, P_{fp} < 0.001$)

n_{id}	2	3	4	5	6	7	8	9	10
f_{th}	4	5	5	6	7	8	9	11	15

表 10 $P_{fn}^{(-1)}(0.001)$ ($R = 1, n_d = 1, P_{fp} < 0.0001$)

n_{id}	2	3	4	5	6	7	8	9	10
f_{th}	4	5	6	6	7	8	10	12	15

表 11 $P_{fn}^{(-1)}(0.001)$ ($R = 1, n_d = 5, P_{fp} < 0.001$)

n_{id}	3	4	5	6	7	8	9
f_{th}	6	8	11	15	20	29	52

表 12 $P_{fn}^{(-1)}(0.001)$ ($R = 1, n_d = 5, P_{fp} < 0.0001$)

n_{id}	3	4	5	6	7	8	9
f_{th}	7	9	12	15	21	31	53

表 13 $P_{fn}^{(-1)}(0.001)$ ($R = 100, n_d = 1, P_{fp} < 0.001$)

n_{id}	2	3	4	5	6	7	8	9	10
f_{th}	4	5	5	6	7	8	9	11	15

表 14 $P_{fn}^{(-1)}(0.001)$ ($R = 100, n_d = 1, P_{fp} < 0.0001$)

n_{id}	2	3	4	5	6	7	8	9	10
f_{th}	4	5	6	6	7	8	10	12	15

表 15 $P_{fn}^{(-1)}(0.001)$ ($R = 100, n_d = 5, P_{fp} < 0.001$)

n_{id}	3	4	5	6	7	8	9
f_{th}	6	8	11	15	20	29	52

表 16 $P_{fn}^{(-1)}(0.001)$ ($R = 100, n_d = 5, P_{fp} < 0.0001$)

n_{id}	3	4	5	6	7	8	9
f_{th}	7	9	12	15	21	31	53

比較した。まず、 $n_d = 1$ の場合について、Yaar の方式⁹と、 $\{n_{id} = 5, f_{th} = 6\}$ のパラメタの組み合わせ ($P_{fp} < 0.0001$ の条件では最適) の検出性能 (誤検出確率、 $P_{fn}^{(-1)}(0.001)$) を比較した結果を表 17 に示す。各ルータのマーキング確率は 0.04、経路上のルータ数は 20、攻撃経路数は 100 とした。

Yaar が論文中で用いた 3 種類のパラメタの組 ($\{n_{id} = 2, f_{th} = 3\}, \{n_{id} = 2, f_{th} = 4\}, \{n_{id} = 3, f_{th} = 5\}$) と、 $\{n_{id} = 5, f_{th} = 6\}$ との差はさほど顕著ではない。誤検出確率の点では $\{n_{id} = 3, f_{th} = 5\}$ が最良であり、 $P_{fn}^{(-1)}(0.001)$ を低減化する (検出漏れを減らす) という意味では $\{n_{id} = 2, f_{th} = 3\}$ が最良である。しかし、 $\{n_{id} = 5, f_{th} = 6\}$ の組み合わせは、誤検出確率と $P_{fn}^{(-1)}(0.001)$ の低減化をバランス良く満たすパラメタの組み合わせとなっている。

次に、 $n_d = 5$ の場合について、Song の方式⁸と、 $\{n_{id} = 4, f_{th} = 9\}$ のパラメタの組み合わせ ($P_{fp} < 0.0001$ の条件では最適) とを比較した結果を表 18 に

表 17 既存の提案方式との比較 ($n_d = 1$)

	Yaar	Optimal Setting
$\{n_{id}, f_{th}\}$	{2, 3} {2, 4} {3, 5} {5, 6}	
FPR	3.03e-2 9.54e-7 2.08e-7 2.96e-5	
$P_{fn}^{(-1)}(0.001)$	105692 203368 129333 116622	

表 18 既存の提案方式との比較 ($n_d = 5$)

	Song	Optimal Setting
$\{n_{id}, f_{th}\}$	{0, 1} {3, 8} {4, 9}	
FPR	1.0 2.3e-10 5.07e-6	
$P_{fn}^{(-1)}(0.001)$	43906 434724 175260	

示す。Song の方式 ($\{n_{id} = 0, f_{th} = 1\}, \{n_{id} = 3, f_{th} = 8\}$) のうち, $\{n_{id} = 0, f_{th} = 1\}$ は誤検出確率が大きすぎ, また $\{n_{id} = 3, f_{th} = 8\}$ は $P_{fn}^{(-1)}(0.001)$ が大きすぎる結果となる。これに比べて, $\{n_{id} = 4, f_{th} = 9\}$ は, やはり誤検出確率と $P_{fn}^{(-1)}(0.001)$ の低減化をバランス良く満たすパラメタの組み合わせであることが確認できる。

6. まとめ

本稿では, 既存の多くのパケットマーキング方式を「ハッシュフラグメント方式」により統一的にモデル化し, このハッシュフラグメント方式の効率性を解析的に分析するとともに, その最適パラメタ設定について論じた。

ハッシュフラグメント方式はパラメタ値を調整することにより, 誤検出確率と検出漏れ確率の両方を実用上充分小さい値に抑えることが可能である。また, Distance field に 5bit を利用しても, Distance field が 1bit と場合と比べて, さほど遜色ない性能を得ることができることも確認された(表 17, 表 18 参照)。本稿では解析モデルにより性能評価を行ったが, 今後, シミュレーションを利用して解析モデルの結果の正当性を検証する予定である。

参考文献

- 1) Feller, W.: *An Introduction to Probability Theory and Its Applications*, 2nd ed., Vol. 1, New York:Wiley (1966).
- 2) Kuznetsov, V., Sandstrom, H. and Simkin, A.: An evaluation of different IP traceback approaches, *ICICS*, pp.37-48 (2002).
- 3) Ogawa, T., Nakamura, F. and Wakahara, Y.: Branch label based probabilistic packet marking for counteracting DDoS attacks, *IEICE Trans. Commun.*, Vol.E87-B, No.7, pp.1900-1909 (2004).
- 4) Park, K. and Lee, H.: On the effectiveness of probabilistic packet marking for IP traceback under Denial of Service attack, *IEEE INFOCOM* (2001).
- 5) Savage, A., Wetherall, D., Karlin, A. and Anderson, T.: Network support for IP traceback, *IEEE/ACM Trans. Networking*, Vol.9, No.3, pp.226-237 (2001).
- 6) Shiota, S.: Some upper and lower bounds on the coupon collector problem, *Journal of Computational and Applied Mathematics*, Vol.200, No.1, pp.154-167 (2007).
- 7) Shiota, S. and Wang, H.: A Comparative Study on Different Probabilistic Packet Marking Schemes for IP Traceback, *IEEE TENCON* (2006).
- 8) Song, D. and Perrig, A.: Advanced and authenticated techniques for IP traceback, *IEEE INFOCOM* (2001).
- 9) Yaar, A., Perrig, A. and Song, D.: FIT: Fast Internet traceback, *IEEE INFOCOM* (2005).
- 10) 鈴木彩子, 大森圭祐, 松嶋竜, 川端まり子, 大室学, 甲斐俊文, 西山茂: IP トレースバックシステムの信頼性の特性分析, 情報処理学会 CSEC 研究発表会 (12月) (2004).
- 11) 澤井裕子, 大江将史, 飯田勝吉, 門林雄基: IP トレースバック逆探知パケット方式のトラヒック量と攻撃経路再構成時間の性能解析, 電子情報通信学会インターネットアーキテクチャー研究会, IA2002-10, pp.7-13 (2002).

Trans. Networking, Vol.9, No.3, pp.226-237 (2001).

- 6) Shiota, S.: Some upper and lower bounds on the coupon collector problem, *Journal of Computational and Applied Mathematics*, Vol.200, No.1, pp.154-167 (2007).
- 7) Shiota, S. and Wang, H.: A Comparative Study on Different Probabilistic Packet Marking Schemes for IP Traceback, *IEEE TENCON* (2006).
- 8) Song, D. and Perrig, A.: Advanced and authenticated techniques for IP traceback, *IEEE INFOCOM* (2001).
- 9) Yaar, A., Perrig, A. and Song, D.: FIT: Fast Internet traceback, *IEEE INFOCOM* (2005).
- 10) 鈴木彩子, 大森圭祐, 松嶋竜, 川端まり子, 大室学, 甲斐俊文, 西山茂: IP トレースバックシステムの信頼性の特性分析, 情報処理学会 CSEC 研究発表会 (12月) (2004).
- 11) 澤井裕子, 大江将史, 飯田勝吉, 門林雄基: IP トレースバック逆探知パケット方式のトラヒック量と攻撃経路再構成時間の性能解析, 電子情報通信学会インターネットアーキテクチャー研究会, IA2002-10, pp.7-13 (2002).

付 錄

A.1 補題 3.3 の証明

次が成立する。

$$P[X > k] = P\left[\bigcup_{n=1}^N A_n^{(k)}\right] = P\left[\bigcup_{n=1}^N \bigcup_{l=1}^f A_{nl}^{(k)}\right].$$

今, F_n ($n = 1, \dots, N$) をそれぞれ $\{1, \dots, f\}$ の任意の部分集合とすると, $F \stackrel{\text{def}}{=} F_1 \times \dots \times F_N$ は直積空間 $\{1, \dots, f\}^N$ の部分集合となる。Inclusion-exclusion principle¹⁾ より

$$P\left[\bigcup_{n=1}^N \bigcup_{l=1}^f A_{nl}^{(k)}\right] = \sum_{F=F_1 \times \dots \times F_N, F \neq \emptyset} (-1)^{\sum_{n=1}^N |F_n|+1} P[S_{F_1 \times \dots \times F_N}^{(k)}].$$

ここで

$$S_{F_1 \times \dots \times F_N}^{(k)} \stackrel{\text{def}}{=} \bigcap_{n=1}^N \bigcap_{l \in F_n} A_{nl}^{(k)},$$

である。なお, $F = \emptyset$ の場合, $\bigcap_{l \in F} A_{nl}^{(k)} = \Omega$ と定義する。

$$P[S_{F_1 \times \dots \times F_N}^{(k)}] = \left(1 - \frac{\sum_{n=1}^N |F_n| p_n}{f}\right)^k,$$

に注意すると, $|F_n| = |F_n|$ とすれば

$$P\left[\bigcup_{n=1}^N \bigcup_{l=1}^f A_{nl}^{(k)}\right] = \sum_{\mathbf{l} \neq \emptyset} (-1)^{\sum_{n=1}^N l_n + 1} \left(\prod_{n=1}^N \binom{f}{l_n}\right) \left(1 - \frac{\sum_{n=1}^N l_n p_n}{f}\right)^k.$$

式(2)と上式を比較することにより,

$$(-1)^{|J|+1} P[S_J^{(k)}]$$

$$= \sum_{\mathbf{l}: \mathcal{E}(\mathbf{l})=J} (-1)^{\sum_{n=1}^N l_n + 1} \left(\prod_{n=1}^N \binom{f}{l_n}\right) \left(1 - \frac{\sum_{n=1}^N l_n p_n}{f}\right)^k.$$

従って,

$$P[S_J^{(k)}] = \sum_{\mathbf{l}: \mathcal{E}(\mathbf{l})=J} (-1)^{\sum_{n=1}^N l_n - |J|} \left(\prod_{n=1}^N \binom{f}{l_n}\right) \left(1 - \frac{\sum_{n=1}^N l_n p_n}{f}\right)^k.$$

A.2 補題 3.4 の証明

まず次を証明する.

補題 A.2.1.

$$\begin{aligned} & P\left[\bigcup_{n=1}^N \bigcup_{F_n:|F_n|=m} A_{F_n}^{(k)}\right] \\ &= \sum_{F_1 \times \cdots \times F_N \in \mathcal{F}_1(m)} (-1) \prod_{n=1:|F_n| \geq m}^N \left\{ (-1)^{|F_n|-m+1} \binom{|F_n|-1}{m-1} \right\} \\ &\quad \times P[S_{F_1 \times \cdots \times F_N}^{(k)}]. \end{aligned} \quad (9)$$

Proof. $\bigcup_{n=1}^N \bigcup_{F_n:|F_n|=m} A_{F_n}^{(k)}$ 内の任意の標本点を E とする. 標本点 E は, (一般性を損なうことなく) リンク 1 からリンク i まではそれぞれ m 個以上のフラグメントが未検出であり, リンク $i+1$ からリンク N までは m 個未満のフラグメントが未検出な標本点と仮定する. 特に, リンク n ($1 \leq n \leq i$) については 1 番目から $f_n (\geq m)$ 番目のフラグメント, またリンク n ($n > i$) については 1 番目から $f_n (< m)$ 番目のフラグメントのみが未検出と仮定する. E の定義より, $P[E]$ は (9) の左辺に (一回分だけ) 寄与する. 一方, E が $S_{F_1 \times \cdots \times F_N}^{(k)}$ に含まれるのは, $F \stackrel{\text{def}}{=} F_1 \times \cdots \times F_N \in \mathcal{F}_1(m)$ が

$$F_n \subset \{1, \dots, f_n\}, n = 1, \dots, i,$$

かつ

$$F_n = \emptyset, n = i+1, \dots, N,$$

を満たす場合のみである. このうち $|F_1| = l_1, |F_2| = l_2, \dots, |F_i| = l_i$ となる F は, (9) の右辺に全部で

$$(-1) \prod_{n=1:|F_n|=m}^i (-1)^{l_n-m+1} \binom{f_n}{l_n} \binom{l_n-1}{m-1} P[E],$$

の寄与をする. これららの総和を取ると

$$(-1) \sum_{j=1}^i \sum_{I:|I|=j} \prod_{n \in I} \left\{ \sum_{l_n=m}^{f_n} (-1)^{l_n-m+1} \binom{f_n}{l_n} \binom{l_n-1}{m-1} \right\} P[E].$$

ここで I は $\{1, \dots, i\}$ の任意の部分集合である.

$$\begin{aligned} & \sum_{l_n=m}^{f_n} (-1)^{l_n-m} \binom{f_n}{l_n} \binom{l_n-1}{m-1} \\ &= \sum_{v=0}^{f_n-m} (-1)^v \binom{f_n}{m+v} \binom{m+v-1}{m-1} = 1, \end{aligned} \quad (10)$$

であることから (付録 A.2.1 参照),

$$\begin{aligned} & (-1) \sum_{j=1}^i \sum_{I:|I|=j} \prod_{n \in I} \left\{ \sum_{l_n=m}^{f_n} (-1)^{l_n-m+1} \binom{f_n}{l_n} \binom{l_n-1}{m-1} \right\} \\ &= (-1) \sum_{j=1}^i \sum_{I:|I|=j} \prod_{n \in I} (-1) = (-1) \sum_{j=1}^i \sum_{I:|I|=j} (-1)^j \\ &= (-1) \sum_{j=1}^i (-1)^j \binom{i}{j} = 1. \end{aligned} \quad (11)$$

よって E は (9) の右辺にも $P[E]$ の寄与をする. 以上証明された. \square

補題 A.2.1 より

$$\begin{aligned} P[X > k] &= P\left[\bigcup_{n=1}^N \bigcup_{F_n:|F_n|=f-f_{th}+1} A_{F_n}^{(k)}\right] \\ &= \sum_{F_1 \times \cdots \times F_N \in \mathcal{F}_1(f-f_{th}+1)} (-1) \\ &\quad \times \prod_{n=1:|F_n|\geq f-f_{th}+1}^N \left\{ (-1)^{|F_n|-f+f_{th}} \binom{|F_n|-1}{f-f_{th}} \right\} P[S_{F_1 \times \cdots \times F_N}^{(k)}] \\ &= \sum_{\mathbf{l} \in \mathcal{L}_1(f-f_{th}+1)} (-1) \prod_{n=1:|F_n|\geq f-f_{th}+1}^N \left\{ (-1)^{l_n-f+f_{th}} \binom{l_n-1}{f-f_{th}} \binom{f}{l_n} \right\} \\ &\quad \times \left(1 - \frac{\sum_{n=1}^N l_n p_n}{f} \right)^k. \end{aligned} \quad (12)$$

ここで, $\mathbf{l} = (l_1, \dots, l_N)$ である. 式 (2) と上式を比較することにより,

$$\begin{aligned} & (-1)^{|J|+1} P[S_J^{(k)}] \\ &= \sum_{\mathbf{l} \in \mathcal{L}_1(f-f_{th}+1); \mathcal{E}(\mathbf{l})=J} (-1) \\ &\quad \times \prod_{n=1:|F_n|\geq f-f_{th}+1}^N \left\{ (-1)^{l_n-f+f_{th}} \binom{l_n-1}{f-f_{th}} \binom{f}{l_n} \right\} \left(1 - \frac{\sum_{i=1}^N l_n p_n}{f} \right)^k. \end{aligned}$$

従つて

$$\begin{aligned} P[S_J^{(k)}] &= \sum_{\mathbf{l} \in \mathcal{L}_1(f-f_{th}+1); \mathcal{E}(\mathbf{l})=J} (-1)^{|J|} \\ &\quad \times \prod_{n=1:|F_n|\geq f-f_{th}+1}^N \left\{ (-1)^{l_n-f+f_{th}} \binom{l_n-1}{f-f_{th}} \binom{f}{l_n} \right\} \left(1 - \frac{\sum_{i=1}^N l_n p_n}{f} \right)^k. \end{aligned}$$

A.2.1 (10) の証明

$$f(i) \stackrel{\text{def}}{=} \sum_{v=0}^{i-m} (-1)^v \binom{m+v-1}{m-1} \binom{i}{m+v},$$

とする. まず $f(m) = 1$ は明らか. さらに

$f(i+1)$

$$\begin{aligned} &= \sum_{v=0}^{i-m} (-1)^v \binom{m+v-1}{m-1} \binom{i+1}{m+v} + (-1)^{i+1-m} \binom{i}{m-1} \\ &= \sum_{v=0}^{i-m} (-1)^v \binom{m+v-1}{m-1} \left(\binom{i}{m+v-1} + \binom{i}{m+v} \right) \\ &\quad + (-1)^{i+1-m} \binom{i}{m-1} \\ &= f(i) + \sum_{v=0}^{i-m} (-1)^v \binom{m+v-1}{m-1} \binom{i}{m+v-1} + (-1)^{i+1-m} \binom{i}{m-1} \\ &= f(i) + \left(\binom{i}{m-1} \sum_{v=0}^{i-m} (-1)^v \binom{i+1-m}{v} \right) + (-1)^{i+1-m} \binom{i}{m-1} \\ &= f(i) + \left(\binom{i}{m-1} \sum_{v=0}^{i-m+1} (-1)^v \binom{i+1-m}{v} \right) = f(i) \end{aligned}$$

従つて, 全ての $i \geq m$ に対して $f(i) = 1$ である.