

学習データに加えられた偽装トラフィックがTor 秘匿サービスへの攻撃に与える影響について

竹之内 玲^{1,a)} 松浦 幹太^{1,b)}

概要: 匿名通信システム Tor は、ユーザとサーバの繋がりを秘匿することで通信に匿名性を持たせる実システムである。Tor を用いてサーバの IP アドレスを隠す Tor 秘匿サービスでは、プロトコルに 4 種類の Tor ネットワーク上のサーキットが含まれる。しかし、機械学習を用いてトラフィック分析をすることで、それぞれを区別出来るということが報告されている。これに対し我々は、Tor 秘匿サービスが発生させるサーバ側のトラフィックと紛らわしいトラフィックである、偽装トラフィックを Tor ネットワークに流すことで、有効なデータセットの割合を下げることを提案している。本論文では、Tor 秘匿サービスの通信データセットの他に一般の Tor 通信も混ぜたデータセットに対し偽装トラフィックを加える。偽装トラフィックの元となるトラフィックの傾向や、加える量によって分類精度にどのような影響があらわれるのか実験し、Tor クライアントが生成するトラフィックのデータ量の 40%程の量の偽装トラフィックが存在すれば分類精度を大きく下げることが可能であることと、偽装トラフィックの生成元となるトラフィックデータの多様さは分類精度に大きい影響を与えないことを明らかにした。さらに偽装トラフィックの運用法について検討し、DHS ノードごとに偽装トラフィックを生成する量を決めることが出来るようにし、また偽装トラフィックの生成元となる Tor クライアントのトラフィックは公開情報にしておくことが適切であると結論付けた。

キーワード: 匿名通信, Tor, 秘匿サービス,

TAKENOUCHI AKIRA^{1,a)} MATSUURA KANTA^{1,b)}

1. はじめに

今日、インターネットにおいて個人の情報を集め、その情報に基づき提供されるサービスが増えている。このような情報はサービス側が個人に合わせたサービスを提供するのに役に立つが、一方でメールアドレスや住所などといった情報は悪用されるおそれがあり、敏感な情報を隠したい人が確かに隠すことが出来るプライバシーエンハンシングの需要が高まってきた。通信の情報も敏感な情報であり、通信の内容を秘匿する技術の一つに暗号化通信がある。暗号化通信では通信の際にその内容を暗号化し、傍受されても攻撃者はその内容を復号出来ずやり取りの内容が漏れないようになるという技術である。しかし暗号化通信においては、ユーザーがどのウェブサイトを開覧しているのかと

いう情報を秘匿することについては考慮されていない。この接続先の情報については、匿名化通信を用いることでこの情報を秘匿することが出来る。実用的な匿名化通信システムの一つに Tor が存在する。

Tor は匿名化通信技術の一つである Onion Routing を実装しており、多くのユーザーが利用、さらには協力している。Tor が提供するシステムはユーザーのブラウジングを匿名化するものの他に、Tor 秘匿サービス (Tor Hidden Service) と呼ばれるものがある。これは、サービスを提供しているサーバの IP アドレスを隠すというシステムである。実際に用いられている Tor 秘匿サービスには人権運動組織や内部告発サイトのようなものがあり、人々の役に立っていると言える。しかしその一方で、麻薬売買サイトや誹謗中傷が多く書き込まれる掲示板などの悪質なサービスも存在し、問題となっている。Tor 秘匿サービスが注目されているなかで、サーバの IP アドレスを暴く様々な攻撃が提案されてきており、Tor 秘匿サービスの匿名性は科

¹ 東京大学 生産技術研究所
IIS, Meguro, Tokyo 153-8505, Japan

a) takenouc@iis.u-tokyo.ac.jp

b) kanta@iis.u-tokyo.ac.jp

学的評価がまだ十分行われているとは言えず、その存在について社会的な議論が先行することは望ましくない。

Tor 通信の接続先を推定する攻撃で高い精度を挙げている、指紋攻撃を Tor 秘匿サービスの通信に適用し Tor 秘匿サービスプロトコル上のどの段階の通信か推定することで、注目しているノードが秘匿サービスか判定できる、Circuit Fingerprinting Attack (CFA) という攻撃を Kwon らが提案した [1]。今日の多くの指紋攻撃は機械学習に基づいているため、ノイズを加える事で精度が大きく下がると言われている [5]。我々は、CFA に対する防御手法として、Dummy Hidden Service (DHS) を提案する。DHS は秘匿サービスと紛らわしいトラフィックを能動的に生成するノードである。Tor ネットワークに DHS が生成したトラフィックが混ざることによって、機械学習にとって有効といえるデータの割合を下げる事が可能だと考えられる。しかし、Tor の開発陣、Tor project が多くのダミーパケットを挿入することを避けていることから、DHS は Tor ネットワークとノードに小さい負担で指紋攻撃の精度を下げる必要がある。これまでの研究で、Tor 秘匿サービスのトラフィックに単純なダミーパケットを加えるだけでは、CFA の精度を有意に下げられないことが明らかになった。本論文では、クライアントとサーバー両者の側から Tor 秘匿サービスのトラフィックを分析し、サーバー側のトラフィックと紛らわしい、偽装トラフィックを生成する手法を提案し、その効果について評価する。

2. Tor と Tor 秘匿サービス

本章では Tor 及び Tor 秘匿サービスに用いられている技術と実装、さらに利用実態について解説する。

2.1 Tor (The onion router)

Tor は Onion Routing[3] を実装した、非常に大規模な実システムである。ユーザは Tor を用いてウェブサイトへ接続すれば、自分がそのウェブサイトへ接続しているという事実を、ユーザ以外に知られることがない。2016 年 12 月時点で 7000 個程の有志の Relay が存在し *1、利用者数は 160 万人を超えている。

Tor で接続する際、クライアントは Tor Network から中継ノードを 3 つ選択する。それぞれの Relay と鍵共有をした後、メッセージを多重に暗号化して通信を行う。この生成されたリンクは Tor circuit と呼ばれる。Tor circuit においてユーザーに一番近い Relay は Entry Guard と呼ばれる。Entry Guard は、ユーザーが接続したいサーバーのアドレスはわからないもののユーザーの IP アドレスを知っているため、匿名化通信においては非常に重要なノードと言える。

*1 <https://metrics.torproject.org/>

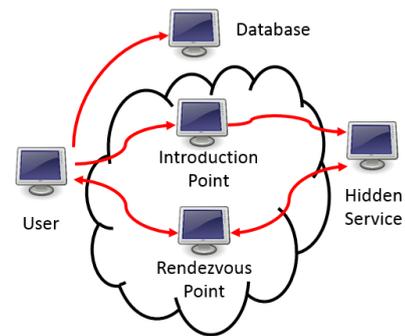


図 1 Tor 秘匿サービスの概要

Fig. 1 Overview of Hidden Service Protocol

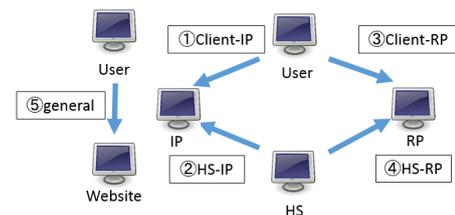


図 2 Tor 秘匿サービスプロトコルにおける 5 種類の Circuit

Fig. 2 Five Circuits on Tor Hidden Service Protocol

2.2 Tor 秘匿サービス

Tor 秘匿サービス (Tor hidden service, HS) は Tor Network を用いて、サービスを提供しているサーバーの IP アドレスを隠すシステムである。メインとなるノードは図 1 の introduction point (IP), rendezvous point (RP) である。IP は HS との Tor Circuit を保持しているノードであり、つまり HS と通信を行うことができる。RP は実際の Client と HS の通信を中継するノードであり、Client は HS と通信を始めようとする際にはこの RP のノードを IP が保持する Tor Circuit を介して HS に送ることで、Client と HS 両側から RP に向かって Tor Circuit を構築する。

3. 関連研究

本章では、機械学習を用いて Tor 秘匿サービスの匿名性を暴く既存の攻撃のうち、受動攻撃のものを一つ説明する。

3.1 Circuit Fingerprinting Attack (CFA)

Kwon らは、トラフィック解析によって秘匿サービスを推定する受動的な攻撃を提案した [1]。攻撃者は攻撃対象と Entry Guard の通信を観測できる攻撃者であり、トラフィックから攻撃対象の IP アドレスを知ることが出来る。Kwon らの攻撃はまず観測した Tor Circuit がプロトコルで現れる複数の Circuit うちのどの Circuit であるかをトラフィックからクラス分類する。Circuit のクラスがわかればエンドノードが秘匿サービスかクライアントか判定できる。そして秘匿サービスと推定したならば、そのノー

ドに対して指紋攻撃をしかけ秘匿サービスを特定する。このようにして、秘匿サービスと IP アドレスを紐付け、匿名性を破ることが出来る。Kwon らの攻撃は、Circuit の種類の推定、指紋攻撃の二つのプロセスから成り、論文は両プロセスの提案、実験について述べられた。前半の Circuit の分類を行う攻撃を特に Circuit Fingerprinting Attack という。

Kwon らによれば、図 2 に表した 5 つのリンクは判別可能である。攻撃者はトラフィックを観測することで自分の見ている circuit が、Tor circuit を用いた普通のウェブサイトの閲覧 (general) か、Client-rendezvous Point (Client-RP) か、Client-Introduction Point (Client-IP) か、Hidden Service-rendezvous Point (HS-RP) か、Hidden service-Introduction Point (HS-IP) か推定することが出来る。彼らは推定に用いる特徴量として 3 つの特徴量を用いた。

- 内向き外向きパケット: それぞれの circuit によって内向き、外向きパケットの数に傾向があるが、ここでいう内向きとは図 2 の Client または HS へ向かう向きが内向き、IP または RP へ向かう向きが外向きである。図 2 で IP とつながりがある Client-IP と HS-IP の判別に特に有効である。Client-IP ではそれぞれの向きのパケットの数が同じになるという特徴がある。彼らは今回は通信の最初の 50 個のパケットにおいて、それぞれの向きの数を計算し特徴とした。
- Duration of Activity: 注目している circuit で通信が行われている時間であり、Client-IP と HS-IP またはそれを区別するのに有効である。IP とのやりとりはプロトコル上長く行われるものではない一方で、実際のやりとりである Client-RP, HS-RP, general の 3 つのについてはより長い時間のやりとりとなる。
- サーキット構築シーケンス: 通信の最初パケットは circuit 構築のためのパケットであり、また Circuit の種類によってセルの数やプロセスが異なるため判別する特徴として利用できる。Kwon らは通信の最初の 10 個のセルの向きのシーケンスを特徴とした。

これらの特徴量を用いて機械学習を行う。Kwon らの実験では Tor 秘匿サービスプロトコルの 4 つの Circuit と、Tor を用いた一般のウェブサイトの閲覧の Circuit のトラフィックデータを収集し、交差検定で CFA の性能評価を行った。ある Circuit のクラスに注目した時、注目クラスの Circuit のトラフィックを、正しくクラス分類した場合を True Positive (TP)、間違えた場合を False Negative (FN)、また注目したクラスと別の Circuit のトラフィックを、注目クラスと別のクラスに分類した場合を True Negative (TN)、注目クラスに分類した場合を False Positive と呼ぶ。Kwon らは評価指標に True Positive Rate ($TPR = \frac{TP}{TP+FN}$)、False Positive Rate ($FPR = \frac{FP}{TN+FP}$) を用いた。この指標を全ての検定で平均をとったところ、TPR が 98%、FPR

は 0.1%未満という結果になった。

4. Dummy Hidden Service

CFA はラップトップでも簡単に行うことが出来る上、受動的な攻撃であるためコストが低く、危険性が高い。一方で機械学習を用いた手法はノイズ耐性が低いと Kwon らは述べている。そこで我々は、秘匿サービスの匿名性をより頑強にするため、Dummy Hidden Service (DHS) を提案する。DHS は秘匿サービスと紛らわしいトラフィックである、偽装トラフィックを作り出すノードである。DHS が作ったトラフィックが学習データにまざることで、実際に秘匿サービスであるか推定する精度を下げることを目的とする。DHS は四個の要件を満たす必要がある。

- Low Cost: DHS を運用することは端末にとってなるべく少ない負担である必要がある。さらに、DHS のトラフィックも Tor Network 上には必要最低限である必要がある
- Distribution: 多くの協力を得るために、DHS は簡単に導入出来る実装である必要がある。
- Effective: DHS のトラフィックは機械学習を用いた攻撃によって秘匿サービスと判定される必要がある。
- Indistinctive: ウェブサイト、または秘匿サービスの指紋攻撃によって特定のサーバーと推定されない必要がある。

これらの中で特に重要であるのが Low cost と Effective である。

4.1 偽装トラフィック

前述の通り、偽装トラフィックとは秘匿サービスが生成するトラフィックと紛らわしいトラフィックである。これは、分類器が偽装トラフィックと秘匿サービスが生成するトラフィックを分類した際に、有意に分類できないことを意味する。我々はこの偽装トラフィックについて検討することで、DHS の Effective と Low cost を達成する手法を提案する。今回の実験においては、偽装トラフィックのトラフィックの性質のみに注目し、その実装方法については考慮していない。

偽装トラフィックの生成手法

[2] で我々は、分類器にとって紛らわしいと言える偽装トラフィックの生成手法を提案した。その評価実験について簡単に説明する。

偽装トラフィックの生成手法として我々は Rev-NPN という手法を提案した。偽装トラフィックは Tor クライアントが生成するトラフィックを元に生成する。Rev-NPN はまず元のトラフィックの 1 番目のパケットの向きを反転させる。そして 2 番目から 4 番目までのパケットの向きはそれぞれ予め決めた確率でパケットの向きを外向き、内向き、外向きとする。また、そうしなかった場合にはそれぞ

れ内向き, 外向き, 内向きとする. つまり 2 番目から 4 番目のパケットは元のクライアント側のトラフィックに関わらず, 確率によって向きが決まる. この確率についてはヒューリスティックに求まると考えられる. それ以降のパケットは, 一定確率で外向きとするかまたはパケットの向きを入れ替える処理を行う. これは, 1つのクライアントのトラフィックから多様な偽装トラフィックを生成することを期待している.

クライアントのトラフィックから Rev-NPN で生成した偽装トラフィック (dummy)220 個と秘匿サービスのトラフィック (HS)195 個からなるデータセットに対し, 分類を行った. 偽装トラフィックを生成する際に 5 番目以降のパケットを外向きにする確率 p をパラメータとし, 評価指標としては, 下式 1, 2 に示した FPR と FNR を用いた. ここで, 偽装トラフィックを謝って秘匿サービスのトラフィックと判定することを False Positive, 秘匿サービスのトラフィックを謝って偽装トラフィックと判定することを False Negative としている.

$$FPR = \frac{\# \text{ of False Positive}}{\# \text{ of dummy}} \quad (1)$$

$$FNR = \frac{\# \text{ of False Negative}}{\# \text{ of HS}} \quad (2)$$

この実験では, 5 番目以降のパケットを外向きにする確率が 30% の際, FPR, FNR が 35% 程度まで上昇した. この手法では偽装トラフィック単体で評価した際には秘匿サービスのトラフィックと紛らわしいという性質を達成したと言える. 次章では, 生成した偽装トラフィックの運用について実験, 検討した成果について述べる.

5. 偽装トラフィックの運用に関する実験

本章では偽装トラフィックを生成する量や生成候補が分類器の精度に与える影響について実験した結果を述べる.

5.1 攻撃者モデル

攻撃者は Tor circuit の中で, エンドノードと Entry Guard の通信を観測することが出来る. 攻撃者は通信を観測し, そこからエンドノードが秘匿サービス, Tor クライアント, DHS のどれか分類する. この推測によってエンドノードの IP アドレスと秘匿サービスであるという情報が紐付けられ, 秘匿サービスの匿名性が失われる. 攻撃者は事前に各エンドノードのトラフィックのデータを集めておき, 分類器に学習させてトラフィックをクライアントのトラフィックか秘匿サービスのトラフィックか偽装トラフィックかに分類する. また, 攻撃者の行動はパケットを観測するだけの受動的な攻撃に限定する. Kwon らによれば, トラフィック分類器を用いた攻撃は能動的な攻撃と比べ個人の端末で簡単に実行できる.

5.2 実験データセット

[2] の実験で用いた, 秘匿サービスに接続した際のクライアントのトラフィック (Client)220 個, 秘匿サービスのトラフィック (HS)195 個を 44 種の秘匿サービスから集めた. また, 一般 Website に Tor を介して接続した際のクライアントのトラフィック (general) を 502 個用いる. これらのデータはページの読み込み始めから読み込み終わりまでを tcpdump コマンドによって記録し, パケットの向きと時間の情報だけ抜き出して整形したパケットのシーケンスが 1 つのデータである. 実験の際にはこのデータセットに偽装トラフィック (dummy) を加える.

5.3 評価指標

今回はデータマイニングツール Weka[4] に実装された SVM, そして最近傍法を用いた分類器で学習, 評価を行う. 学習, 評価データセットの分け方は定めず 5 回の交差検定を行った. 偽装トラフィックがどれだけ紛らわしいかを評価する指標は, 偽装トラフィックを正しく偽装トラフィックと分類することを True Negative(TN), 秘匿サービスのトラフィックと誤って分類することを False Positive(FP) とし, False Positive Rate ($FPR = \frac{FP}{TN+FP}$) を用いた. また偽装トラフィックを学習データセットに用いることで秘匿サービスの分類精度にどのような影響が出るかを評価する指標は, 秘匿サービスのトラフィックを正しく分類することを True Positive(TP), 誤って偽装トラフィックと分類することを False Negative(FN) とし, False Negative Rate ($FNR = \frac{FN}{TP+FN}$) を用いた.

5.4 偽装トラフィックの生成量に関する実験

偽装トラフィックは 4.1 に述べたように, Rev-NPN でクライアントのトラフィックから生成する. この生成する量を制限するが, クライアントのトラフィックの割合をパラメータとし, 分類器の精度の変化を調べる. このパラメータを Rate とすると, 偽装トラフィックの生成量は下式 3 のとおりである.

$$\# \text{ of dummy} = (\# \text{ of client} + \# \text{ of general}) \cdot \text{Rate} \quad (3)$$

Rate は 0 から 1.0 まで 0.1 刻みに動かす.

図 3 は kNN による分類結果であり, 割合が 60% になるまでは徐々に FNR が上昇するが, それ以降は FNR が 25% 程で安定する. 一方, FPR は 5% 程を保っている. 最近傍法にとって重要なのは先頭 10 パケットの向きであることを [2] で明らかにしたが, 偽装トラフィックを増やしても秘匿サービス同士で近い先頭 10 パケットのシーケンスを持つものは増えずに, 今回の実験では FPR が上昇せず, 秘匿サービスの多様さによって上昇するのだと考えられる.

一方, 右図 4 は SVM による分類結果を示しているが, 偽

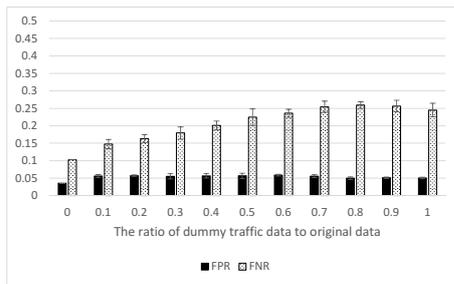


図 3 偽装トラフィックの数を変化させた際の kNN 分類器の FPR と FNR

Fig. 3 FPR and FNR of classification with knn. The ratio of dummy traffic varies.

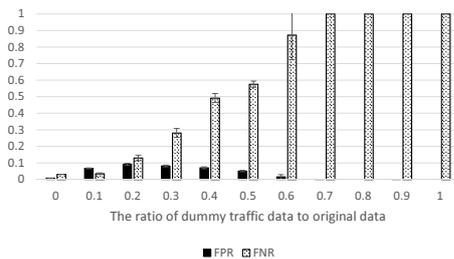


図 4 偽装トラフィックの数を変化させた際の SVM 分類器の FPR と FNR

Fig. 4 FPR and FNR of classification with svm.

装トラフィックを増やすにつれ FNR が上昇するが、FPR が低下していくことがわかる。特に、Client, general クラスのトラフィックの 70%の量以上に偽装トラフィックを混ぜると、HS クラスのトラフィックを HS クラスと分類することができなくなる。ここで、実際には HS クラスのトラフィックを HS クラスと分類できない割合である FNR は 50%が最高である。これは、HS クラスであるかそうでないかの二つをランダムで決めるという手法が存在するからである。偽装トラフィックの割合が 50%を超える際には、攻撃者は HS クラスでないと分類されたトラフィックを HS クラスのトラフィックと分類してしまえば、ランダムな推定よりかえってより当たりやすくなる。そのため、FNR が 50%に達する、40%のときに効率よく防御出来ていると言える。

5.5 偽装トラフィックの生成元候補に関する実験

本実験では、偽装トラフィックが生成する候補の数によって分類器にどのような影響があるか調べる。5.4 で結論づけたように、偽装トラフィックはクライアントのトラフィックの量の 40%混ぜる。偽装トラフィックの生成候補はクライアントのトラフィックの P 倍であるとし、本実験では P を 1%, 2%, 5%, 10%, 20%, 50%, 100%で変化させた際の分類器の精度を調べた。

左図 5 は kNN による分類結果を示す。生成元候補の割合が全体の 1%, 2%のあたりでは FPR が 3%程、FNR が

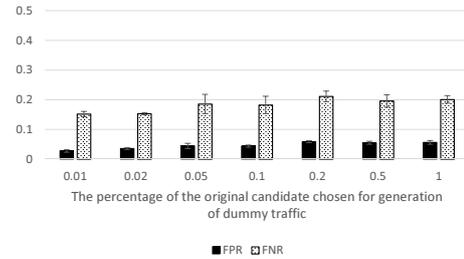


図 5 偽装トラフィック生成元候補の数を変化させた際の knn 分類器の FPR と FNR

Fig. 5 FPR and FNR of classification with knn. The percentage of the original traffic used to generate dummy traffic varies.

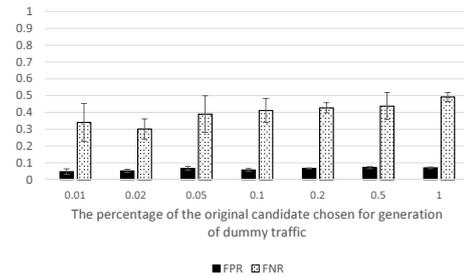


図 6 偽装トラフィック生成元候補の数を変化させた際の SVM 分類器の FPR と FNR

Fig. 6 FPR and FNR of classification with svm. The percentage of the original traffic used to generate dummy traffic varies.

15%程であり、割合が 5%以上のところでは、FPR が 5%程、FNR は 20%程である。前述の通り、kNN では先頭 10 パケットの向きの特徴が重みとして強く、候補があまりにも少ない状態では今回用いた生成手法では先頭 10 パケットの多様さを作り出せず、kNN が誤ることが少なかったのだと考えられる。一方、割合が 5%以上では FPR, FNR に有意な差はみられない。Rev-NPN で生成した偽装トラフィックは、生成元候補が増えても先頭 10 パケットの向きの多様さがそこまで増えないと言える。

一方、右図 6 は SVM による分類結果を示す。一方 FNR は上下するものの決まった傾向はないが、生成元候補が少ないときに分散が大きくなっている。SVM での分類結果に対しては、Rev-NPN は候補元の選び方が偏ると分類精度が低くならないのだと考えられる。

5.6 運用に関する考察

偽装トラフィックの生成候補

偽装トラフィックの生成候補元について、5.5 項の結果から、全体の 5%ほどで Rev-NPN で生成した偽装トラフィックの多様さが限度に達するが、選び方によっては偏った偽装トラフィックになるといえる。また CFA の特徴量の多様さにはそもそも限度が存在し、実際の Tor 環境の規模では 5%より少ない割合で十分な多様さが産まれると考

えられる。よって、候補元として、Client, general のトラフィックを公開情報としておきこれを DHS に用いてもらえばよい。また攻撃者この候補が知られても本節の実験の仮定の範囲内であり、分類精度に与える影響は変わらないと言える。偽装トラフィックを各 DHS ノードが作るということは、自由に生成して良いということであるため、攻撃者が各ノードを占拠し、偽装トラフィックを生成させず秘匿サービスのトラフィックを流すかもしくは DHS を停止させることが考えられるが、DHS は Tor のエンドノードであり、また多数存在するため、攻撃者の労力は非常に高いと言える。

偽装トラフィックの生成割合

次に偽装トラフィックをどの程度 Tor Network に流すかについて考える。ここでは、Client, general クラスのデータに対する偽装トラフィックのデータの割合を Rate とし、これを動かす問題として考える。

Rate を固定する場合は、本節の実験の仮定と同様である。この時は、攻撃者はあらかじめ本節と同じ実験を行っておき、分類器の Rate ごとの精度を調べておく。Rate が固定されているということは Rate は公開情報であるため、攻撃者は分類器の精度を知ることが出来、エンドノードの予測に役立たせることは出来る。

Rate が時間によってランダムに変動する場合、こちらも攻撃者はどう動くか、どう動いたのかを知ることが出来る。Tor の通信が多い時間について考えると、HS クラスのトラフィックの先頭 10 パケットの向きが多様になり、kNN によって偽装トラフィックを誤って HS クラスと分類することが増えると考えられる。よって、この時間では SVM による分類が有効であると言える。Tor 通信が多い時間に偽装トラフィックが少ないと、その時間に学習データセットを作った攻撃者は SVM によって高い精度で攻撃が成功する。一方で、Tor の通信が多い時間に偽装トラフィックを多くすると、その分 Tor Network に対する負荷が大きくなってしまう。また、Tor 通信が少ない時間に偽装トラフィックが多いと、SVM の分類精度を大きく下げることが出来る。kNN は HS クラスのトラフィックの多様性が少ないため、False Positive が少なくなると考えられるが、FNR には一定の効果がある。この時間に偽装トラフィックが少ないと、SVM も kNN も高い精度で攻撃が成功する一方で、Tor Network には余力があると言えるため、不適切な状況である。よって、ランダムな遷移では攻撃者にとって有利な状況を産んでしまう。

各 DHS がランダム、または各ノードが自由に偽装トラフィックを生成する場合、Rate は攻撃者にもわからなくなる。つまり、攻撃者は学習データセットにどの程度偽装トラフィックを混ぜればいいかわからない。実際の偽装トラフィックの Rate 通りならば本節で述べた結果と同じになるが、実際の偽装トラフィックの Rate が攻撃者の想定

より少ないと、攻撃者の分類器は本来秘匿サービスのトラフィックが持つ特徴を偽装トラフィックの特徴と学習してしまっているの、と考えられ、秘匿サービスを謝って偽装トラフィックと判定する False Negative が増えると考えられる。実際の偽装トラフィックの Rate が攻撃者の想定より多いと、分類器は偽装トラフィックの特徴を十分に学習できず、偽装トラフィックは秘匿サービスに近い特徴を持つので、偽装トラフィックを誤って秘匿サービスとして判定する False Positive が増えるよって、前後で FNR の変動が大きい Rate を基準としてユーザに提示しておき、後に各ノードが各々の設定で DHS を利用することが有効な手段であると考えられる。なお、本研究では図 4 から分かる通り、Rate = 0.3 が適していた。

6. 結論

本研究の実験では、偽装トラフィックをどのトラフィックから生成し、どの程度生成するかを検討するための実験を行った。その結果、Tor クライアントのトラフィックの 40% の量の偽装トラフィックが適切であることと、生成元となるトラフィックの候補の数は、分類精度に大きい影響を与えないが、偏った特徴をもった候補だけでは、分類精度を有意に下げることが出来ないことが分かった。さらに、この実験の結果から DHS を運用する際偽装トラフィックをどのように生成するべきか検討を行った。DHS ノードごとに偽装トラフィックを生成する量を決めることが出来るようにし、また偽装トラフィックの生成元となる Tor クライアントのトラフィックは公開情報にしておくことが適切であると結論付けた。

参考文献

- [1] Albert Kwon, Mashaal AlSabah, David Lazar, Marc Dacier and Srinivas Devadas, "Circuit Fingerprinting Attacks: Passive De-anonymization of Tor Hidden Services," In Proceeding of the 24th USENIX Security Symposium, 2015.
- [2] 竹之内 玲, 松浦 幹太, "Tor 秘匿サービスへの攻撃に対抗する偽装トラフィック生成," 2017 年 暗号と情報セキュリティシンポジウム, 2017.
- [3] David M. Goldschlag, Michael G. Reed and Paul F. Syverson, "Hiding Routing Information," In Proceeding of the First International Workshop on Information Hiding, 1996.
- [4] Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, Ian H. Witten, "The WEKA data mining software: an update," In ACM SIGKDD Explorations Newsletter, 2009.
- [5] Tao Wang and Ian Goldberg, "On Realistically Attacking Tor with Website Fingerprinting," Technical report, 2015.