匿名通信システムにおける複数経路を用いた指紋攻撃対策手法に関する一検討

横山絵美里[1]　平田木乃美[1]　宗裕文[1]　山場久昭[1]　久保田真一郎[1]
朴美娘[2]　岡崎直宣[1]

概要：現在，注目されている技術として匿名通信システムがある．この技術は，パケットの通信を盗聴し，利用者がアクセスするWebサイトを特定することを防いでいる．その中で最も普及しているシステムの一つとして，多段階プロキシを利用することで，高い匿名性を実現しているThe Onion Router(Tor)があげられる．しかし，これを脅かそうとする攻撃の代表的なものとして，「指紋攻撃」と呼ばれる攻撃が存在する．これは，流れるトラフィックからWebサイトの特徴となるトラフィックを抽出して，利用者のアクセスするWebサイトを特定する攻撃である．本論文では，これに対し，Webサイトの情報を分割して複数の経路で読み込ませるような手法を提案する．これについて実験を行うことでその有効性を示す．

キーワード：匿名通信，The Onion Router，Tor，指紋攻撃

An examination on countermeasure against fingerprinting attack in Tor by downloading contents through two distinct connections

EMIRI YOKOYAMA[1]　KONOMI HIRATA[1]　HIROFUMI SOU[1]
HISAAKI YAMABA[1]　SHINICHIRO KUBOTA[1]　MIRANG PARK[2]
NAONOBU OKAZAKI[1]

Abstract: Tor (The Onion Router) is one of the most widely used anonymous communication systems that realize anonymous web surfing without revealing the user's identity. However, it is known that an onion router that directly communicates with a user can infer which website a user accessed by leveraging site-specific traffic features, e.g., volume and time, and this attack is called fingerprinting attack. In this paper, we propose a countermeasure against the fingerprinting attack by obfuscating site-specific traffic features. The idea is to separately download image-based contents in random balance through two distinct Tor connection which are established using virtual machine. We show the effectiveness of our scheme with experiments.

Keywords: Anonymity system，The Onion Router，Tor，fingerprinting

1. はじめに

近年，インターネットの急速な普及により，私たちはネットワークを介して様々な情報をやり取りできるようになってきている．個人のインターネット利用は年々増加，拡大しており，幅広い利用者に浸透してきている．現在では，だれもがこの技術を利用するようになっており，サイトによっては個人情報をやり取りするなど，インターネットは日常生活にかかせないツールの一つとなっている．しかしこれに伴い，インターネット利用者の通信内容を盗聴する行為や，パケットのヘッダ情報を盗聴することで利用者がアクセスするWebサイトを特定する行為が問題となっている[1].

現在，この問題解決のために暗号化通信技術[2]と匿名通信システム[3]が注目されている．このうち，暗号化通信技術は通信内容を秘匿することで個人情報などの重要な情報を第三者に知られないようにするが，通信内容を秘匿することは可能でも，誰がどのサイトと通信したかといった情報を秘匿することができない．一方，匿名通信システムは自身を特定するような情報を通信相手に知られることなく通信を行うことができる技術である．現在この匿名通信システムの中でも最も普及しているシステムの一つがThe Onion Router(Tor)[4][5]である．Torは，一般人，ジャーナリストなど様々な人々に，自身に迫る脅威から身を守る手段として利用されている．

その一方で，Tor利用者の匿名性を脅かすような攻撃も考案されつつある．その中でも，流れるトラフィックからWebサイトごとのユニークな特徴（指紋）を抽出し，利用者のアクセスするWebサイトを特定する「指紋攻撃[6]」が脅威になっている．本論文ではこれに対し，指紋となるトラフィックの特徴を取得されにくくするために，複数の経路を用いてWebページの情報を取得するような手法を提案する．また，実験により評価を行い，その有効性を示す．

1 宮崎大学
University of Miyazaki
2 神奈川工科大学
Kanagawa Institute of Technology

2. The Onion Router(Tor)

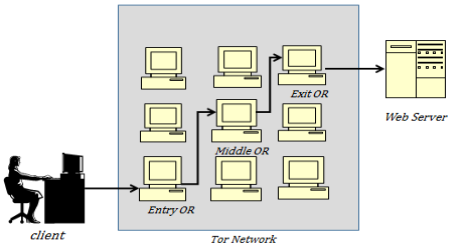

図 1　Tor の概略図

Tor は米海軍調査研究所が政府の通信保護を目的として開発された技術である．2016 年 1 月現在，Tor は一日あたりおよそ 200 万人の人々に利用されており[8]，最も普及している匿名通信システムの一つであるといえる．Tor は複数のプロキシを経由させる仮想回線接続を行うことで，高い匿名性をもつ通信を実現している．

Tor の主な利用目的は，公共ネットワーク上で組織または個人単位で自分自身のプライバシを守りつつ情報をやり取りすることである．具体例として，虐待や深刻な病気など似たような境遇を持つ人々の間での情報共有や内部告発などがあげられる．このような情報をやり取りする上で通信の匿名化は重要なことであり，Tor はインターネットを利用する上で必要な技術であるといえる．

Tor の通信は図 1 のようにオニオンルータ(以下，OR)と呼ばれる中継プロキシを複数経由することで行われる．Tor を利用する際，利用者は Tor ネットワークから OR を三つ選択し，それぞれの OR と鍵交換を行い，それらを順に経由してサーバへアクセスする．現在 Tor ネットワーク内には約 7000 ノードの OR が存在する．このとき，利用者に近い OR から順に入口 OR，中間 OR，出口 OR と呼ぶこととする．利用者が Web サイトへアクセスする際には，選択した三つの OR がパケットを順に暗号化する．これにより，経路上のどの OR も利用者と利用者がアクセスした Web サイトを特定することが出来ない．

Tor は上記のような処理を行うことにより高い匿名性を持つシステムと言えるが，攻撃者が Tor を使用する利用者の情報を得ようと様々な攻撃を行うため，完全な匿名性を提供しているわけではない．その理由としては，Tor ネットワーク内の OR がすべて一般のボランティアにより構成されていることで，攻撃者がノード群の中に攻撃を行う OR を含ませることが容易なためである．攻撃の例として，Web サイトと直接通信を行う出口 OR では通信内容を暗号化できないことを悪用して通信内容を傍受する手法が存在する．この手法に対しては現在，送信するデータを HTTPS を利用して内容を暗号化することで対処することができる．しかし通信内容を傍受せず，また複数の汚染 OR を必要としない「指紋攻撃」[9][11]と呼ばれる攻撃が存在する．この攻撃は，トラフィックの特徴などから Web サイトを特定する攻撃であり，これに対する根本的な対策はまだ確立されていない．また，他の攻撃に比べて占拠するプロキシの数が少なく済むため，実現可能性が高い．

よって本論文ではこの「指紋攻撃」に注目する．

3. 指紋攻撃

指紋攻撃は，攻撃者が入口 OR となり Tor ネットワーク上を流れるトラフィックを観測することで利用者がアクセスする Web サイトを特定する手法である．Web サイトは様々な画像ファイルやスクリプトファイルから構成されているため，Web サイトごとにファイル数やサイズ，トラフィックの流れなどにユニークな特徴（以下，指紋）が表れる。攻撃者は指紋情報をトラフィックから収集し，Web サイトを特定する．以下に，現在提案されている指紋攻撃についていくつか記述する．

[9]では，指紋情報の分類に Support Vector Machine(SVM)を使用しており, 54%の確率で Web サイトを特定している．この手法の指紋情報にはパケットの総数，HTML のファイルサイズなどトラフィックから抽出できるような情報を用いている．

また[10]では，指紋攻撃に対する対策を行われた場合でも指紋攻撃を可能にする手法について提案している．この手法は，指紋攻撃への対策のためにトラフィックに何らかの処理がなされた場合でも，その処理を打ち消す逆処理を行うことでその対策を無効化するものである．

このように，指紋攻撃は Tor に対して非常に大きな脅威であるといえる．本論文の目的は指紋攻撃に対する対策を考案することであるが，効率的な対策を考えるためには指紋攻撃に関する分析が必要である．そこで，3.1 では[9]の手法を参考に指紋攻撃の実装を行いその脅威の程度を検証する．

3.1 想定する指紋攻撃

本論文で想定する指紋攻撃は，[11]と同様に攻撃者が入口 OR を汚染することで行うものとする．攻撃者は攻撃を行う利用者（以下，ターゲット）のトラフィックから指紋情報を抽出することでターゲットのアクセスしている Web サイトを特定する．指紋情報については後述する．ここでは想定する指紋攻撃について示し，実験によりその脅威の度合いを確かめる．

3.2 指紋情報

ここでは本実験において想定する指紋情報について定義する．指紋情報は本論文で想定する指紋攻撃においてターゲットのアクセスした Web サイトを特定するために用いるものである．攻撃者が収集するトラフィックから Web

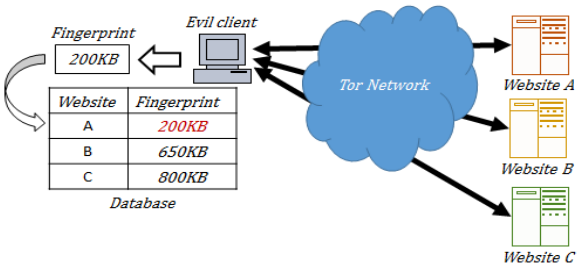

図 2　指紋情報収集フェーズ概略図

サイトの特徴になりえるものを抽出したものが指紋情報となる．本論文の指紋攻撃で設定した指紋情報となり得る特徴は以下の項目である．

- $S_{total}^{p\to q}$：p から q へ向かうパケットの総量(byte)
- $N_{total}^{p\to q}$：p から q へ向かうパケットの総数
- $S_{avg}^{p\to q}$：p から q へ向かうパケットの平均サイズ(byte)
- $S_{var}^{p\to q}$：p から q へ向かうパケットの分散(byte)
- $C_{avg}^{p\to q}$：p から q へ向かうチャンクの平均サイズ(byte)
- $C_{var}^{p\to q}$：p から q へ向かうチャンクの分散(byte)

項目中の p と q はそれぞれ利用者 c と Web サイト w を表している．指紋情報は上記の 6 項目に加えてトラフィックの向きも含めた$12(=6\times 2)$要素とする．またチャンクは，パケットの向き$(p\to q,\ q\to p)$が，前回向きが変わったときから次に向きが変わる直前までを一つの塊としてみたものであり，その塊の合計サイズ(byte)とする．

3.3　処理手順

本論文で想定する指紋攻撃は大きく分けると指紋情報の収集フェーズと Web サイト特定フェーズの二つに分かれる．

(1)　指紋情報の収集フェーズ

このフェーズでは，攻撃者は指紋情報を収集するために図 2 のように Tor 利用者としてターゲットのアクセスしそうな Web サイトに定期的にアクセスする．そして Web サイトにアクセスした際のトラフィックから指紋情報を収集し，データベース化する．このように，定期的に Web サイトへのアクセスすることでニュースサイトやショッピングサイトのような時間とともに変化するサイトにも対応することができる。

(2)　Web サイト特定フェーズ

このフェーズでは，図 3 のように入口 OR としてターゲットが接続してくるのを待つ．攻撃者はターゲットの接続を確認すると，(1)と同様にしてターゲットのトラフィックから指紋情報を収集する．指紋情報の抽出が終わると指紋情報データベースと比較を行う．このとき類似度が最も高かった Web サイトをターゲットがアクセスしたサイトと推定する．本論文では収集したターゲットの指紋情報（識別データ）と指紋情報データベース（学習データ）の比較

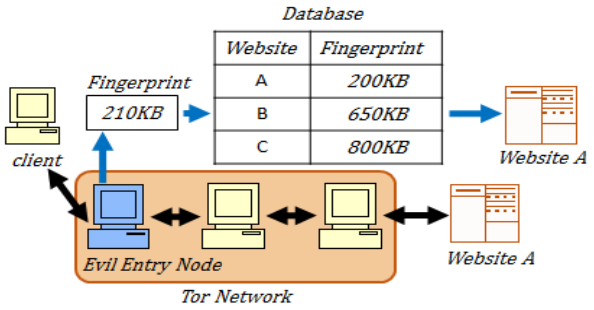

図 3　Web サイト特定フェーズ概略図

表 1　PC の仕様

OS	Windows 7 Professional
CPU	Corei7-4770 3.40GHz
Browser	Mozilla Firefox 42.0
Tor	v0.5.16.3
Perl	v0.2.3.25
R	v3.2.2

3.4　実験

ここでは指紋情報データベースとターゲットの指紋情報を比較することで Web サイトが識別される度合いを調査し，その脅威を示す．

(1)　実験環境

攻撃者がデータベースを収集するのに使用する PC とターゲットからトラフィックを収集する PC は同一のものを用いる．この時の仕様を表 1 に示す．本実験で使用する Web サイトは全て実在するものを用いる．Web サイトはアクセスランキングサイトである Alexa[12] から上位 100 サイトを選択した．このとき，ランキングには国別トップレベルドメインが異なるのみの同一サイトも含まれているため，これを除く Web サイトを選択することで重複のないようにした．指紋情報の収集は通信トラフィックを Wireshark[13]によりパケットキャプチャすることで行う．

(2)　実験方法

まず，Alexa より選択した Web サイトへ Tor を利用してアクセスし，攻撃者用の指紋情報データベースを作成する．同様にしてターゲットの指紋情報も収集し，データベースと比較を行うことでターゲットがアクセスした Web サイトを特定する．このとき各 Web サイトの特定率r_iおよび全体の特定率 R をそれぞれ以下のように定義する．

$$r_i = \frac{Success_i}{Num}\times 100$$

$$\mathrm{R} = \frac{\sum_{i=1}^{Site} r_i}{Site}$$

ここで $Site$，Num，$Success_i$は，それぞれ訪問する Web サイトとその回数，サイト番号 i の特定成功率である．本実

表 2　評価指標

Web サイトの識別率(%)	指紋攻撃に対する耐性
$0 \leq r_i \leq 10$	指紋攻撃への耐性が高い
$10 \leq r_i \leq 100$	指紋攻撃への耐性が低い

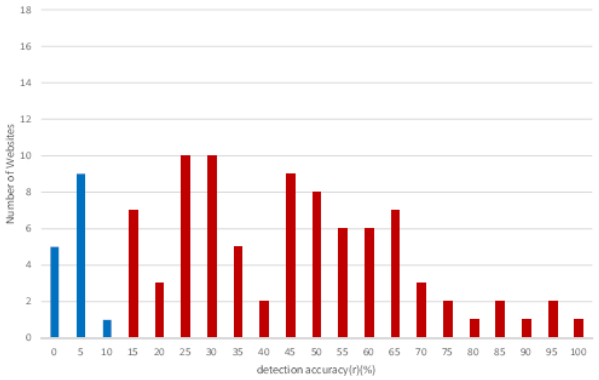

図 4　識別率r_iに対する Web サイト数

験では，は *Site* = 100，*Num* = 20 とした．また本実験では特定率r_iに対する指紋攻撃耐性の判定について表 2 のように定義した．同表で「指紋攻撃への耐性が高い」は指紋攻撃によって Web サイトが特定される可能性が低いことを表し，「指紋攻撃への耐性が低い」は指紋攻撃によって Web サイトを特定される可能性があることを示している．よって，本研究の目的は「指紋攻撃への耐性が高い」Web サイトを増やすことである．本実験では簡略化のために，ターゲットが Web サイトへアクセスする際に閲覧するページはトップページのみとする．また，閲覧時間は 2 分間とした．閲覧時間の設定については，Tor を利用して Web サイトへアクセスする際に通常の接続より時間がかかるため余裕を持たせた値にした．ここで，あらかじめ代表的な Web サイトに対して実験を行ったところ，2 分間時間を与えることでどの Web サイトでも全てのコンテンツを受信することができた．

(3) 実験結果・考察

本実験で Web サイト全体の識別率 39.65%という結果を得た．図 4 は識別率r_iに対する Web サイト数を示している．図 4 と表 2 の評価指標から「指紋攻撃への耐性が高い」Web サイトが 15%，「指紋攻撃への耐性が低い」Web サイトが 85%を占めていることがわかった．この結果から，現在の Tor ネットワークにおいて指紋攻撃は脅威であり対策が必要であるといえる．

4. 既存研究

3. の実験により，指紋攻撃は Tor に対して脅威となる攻撃であることを示した．以下に，指紋攻撃に対する対策をいくつか記述する．

[10]では，Web サイトのトラフィックを利用した対策について提案している．この手法は複数の Web サイトへ同時にアクセスすることで指紋情報を隠すというものである．しかし，この手法はトラフィック量も通常の倍に増加するため，Tor ネットワークへの通信負荷も大きくなる．

[15], [16]ではトラフィックにダミー情報を含ませることで指紋攻撃への防御策を提案している．[15]の手法ではパケット到着時間上の確立に従い，中間ノードでダミー情報をパディングする．これにより，Web サイトの指紋情報となりうるパケットのサイズと感覚を均一にして指紋攻撃を防ぐことができるが Tor に対する負担が大きくなる．さらに，この手法は Web サイトから送信されるデータに大きな差がある場合，総トラフィック量や総パケット数などから Web サイトが特定されてしまう恐れがある．

[11]では，攻撃者に指紋情報を取得しにくくさせることを狙いとしており，利用者が Web サイトにアクセスする際，Web サイトの情報を HTML ファイルと画像コンテンツに分けてそれぞれ別の経路で取得する．まずアクセスする Web サイトの HTML ファイルのみを読み込む．この時点で利用者には画像の一切ないテキストのみのページが表示される．もし利用者が表示させたいと思う画像コンテンツがある場合，利用者は当該コンテンツの要求を行う．画像コンテンツの近くには利用者が要求をできるように，画像コンテンツの付近にボタンを設置することで対応している．これにより，通信するトラフィック量を増やすことなく Web サイトの指紋情報を隠すことができる．しかしこの手法は利用者に必要な画像コンテンツを要求させるといった不便さや，指紋情報になりやすい画像コンテンツが片方の経路に集中してしまうことによって Web サイトが特定される可能性がある．

5. 提案手法

5.1 目的

4. で説明したように複数経路を利用した手法にはユーザに必要なコンテンツを選択させるという不便さや，片方の経路に指紋情報となりやすい画像が集中するという欠点があった．そこで本提案手法では，複数経路を利用した手法と同様に通常読み込まれるはずの Web サイトの指紋情報を読み込ませにくくするという特性を保持しつつ，既存の問題点を解決するような手法を提案する．

5.2 概要

本提案手法はアクセスする Web サイトの情報を HTML ファイルと画像コンテンツに分け，二つの経路を使って全て取得する1．各経路で取得する Web サイト情報は，片方の経路で HTML ファイルと画像コンテンツを取得し，も

1　保存する Web サイトの情報をコンテンツに限定しているが，その他に JavaScript や Cascading Style Sheetts などもある

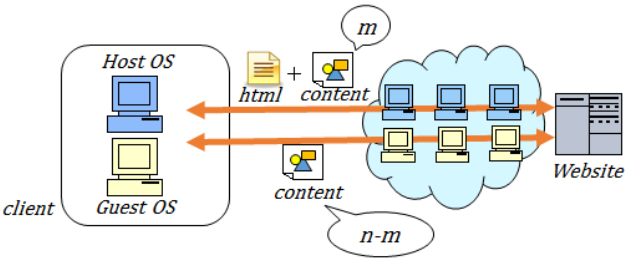

図 5　提案手法の概略図

う片方の経路で残りの画像コンテンツを取得する．本論文では各経路で取得する画像コンテンツの振り分け方について，アクセスする Web ページの総画像数を基にした手法を 2 つ提案する．この 2 つの画像振り分け法については後述し，実験を行うことで評価する． また、2 つの経路の確立方法については 4. の手法と同様に PC 内に仮想 OS を設けて実現させる．本提案手法は 4. の手法とは異なり，指紋情報となりやすい画像コンテンツを複数の経路を通じてランダムに全て取得している．これにより利用者にコンテンツを選択させるという不便さや，画像コンテンツによる指紋情報を分散させることが可能となる．

5.3 前提条件

指紋攻撃は，攻撃者が利用者となって様々な Web サイトにアクセスし，その際に利用者と入口 OR 間を行き来するパケットをキャプチャすることでデータベースを作成する．よって本提案手法のように二つの経路を利用する場合であったとしても，攻撃者はそれぞれの経路に対するデータベースを作成することが可能である．また，本手法では二つの経路を確率する必要がある．しかし，現在の Tor の仕組みでは一つのホスト内で起動することが出来る Tor Browser は一つのみである．そこで PC 内に仮想 OS を設けることで二つの経路を実現する．

5.4 等数分割法

本手法はアクセスする Web ページの総画像数が n のとき，それぞれの経路で保存する画像コンテンツ数 m を最も単純な分け方である $m=n/2$ と固定して保存する．まず HTML から抽出した画像 URL のリストを乱数を用いて並べ，並べ替えた画像リストの上から順に 0 か 1 の値を1/2の確率で生成した乱数に従ってどちらの経路で保存するか決定する．その操作を行っている間にどちらかの経路の保存する画像コンテンツ数が1/2の値に達した場合，並べ替えた画像リストの残りの画像コンテンツ全てをもう片方の経路で保存するようにする．

この手法は各経路で保存する画像コンテンツ数を固定しているため，指紋情報を紛らわせることに限界があることも考えられた．そこで各経路で保存する画像コンテンツ数の比率に自由度を持たせるような手法も考案した．その手法について 5.5 で説明する．

5.5 確率的重み付け配分法

本手法はアクセスする Web ページの総画像数が n のとき，片方の経路で保存する画像コンテンツ数 m を $1 \leq m < n$の範囲の乱数で決定し，もう片方の経路で$n-m$の画像コンテンツを保存する．m の範囲は指紋情報となりやすい画像コンテンツが片方の経路のみで保存されることのないようにするため，この値とした．まず，HTML から抽出した画像 URL のリストを乱数を使って並べ替える．次に，各経路で保存する画像コンテンツ数を設定するための値mを決定する．そして，並べ替えた画像リストの上から順に 0 か 1 の値をn/mの確率で生成した乱数に従ってどちらの経路で保存するか決定する．その操作を行っている間にどちらかの経路の保存する画像コンテンツ数がmの値に達した場合，並べ替えた画像リストの残りの画像コンテンツ全てをもう片方の経路で保存するようにする．

5.6 処理手順

本提案手法の概略図を図 5 に示す．提案手法ではホスト OS 側とゲスト OS 側それぞれで Tor を起動し，経路を 2 本用意する．利用者が Tor を介して Web サイトへアクセスする場合，まずホスト OS 側の経路で HTML ファイルを保存する．次に，保存した HTML ファイルから画像を示す拡張子を探し，画像コンテンツをどちらの経路で保存するか振り分ける．最後に画像コンテンツをそれぞれの決められた経路で保存し，保存した HTML ファイルを書き換える．これらの動作を行うことで，利用者にアクセスする Web サイトを表示させる．

5.7 期待される効果

本提案手法を適用することによって得られる効果について述べる．4. で記述した複数経路を利用した手法の問題点としてクライアントが利用する際の不便さや画像コンテンツの集中によって指紋情報を収集される可能性が挙げられていた．しかし，本提案手法はアクセスする web サイトの画像コンテンツをクライアントの要求に頼ることなく全て保存ことができる．また指紋情報になりやすい画像コンテンツを複数の経路でランダムに取得させることによって，攻撃者にデータベースを作成させにくくすることが可能である．実行環境としては，4. と同様に仮想環境を入れる必要はあるが Tor の改造せずに導入することができる．

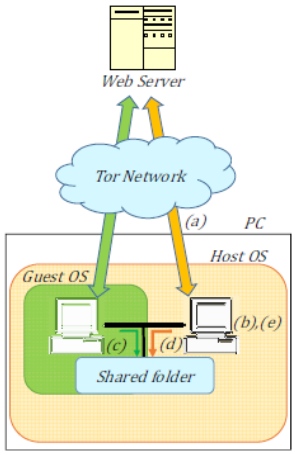

図 6　提案システムの概略図

6. 評価

3. で示した指紋攻撃に対して本提案手法がどの程度の耐性を持つのか評価する．比較対象は提案手法を適用した場合の実験結果と 3.4 で示した実験結果とする．

6.1 実装

本提案システムは，ホスト OS 側とゲスト OS 側それぞれで動く 2 つのサブシステムから構成される．また，本提案システムは HTML ファイルや画像コンテンツを保存するが，ホスト OS とゲスト OS で連携することができるよう共有フォルダーを利用することで対応した．

本実験における提案システムの動作手順について説明する．まず，利用者が Web サイトにアクセスする際，ホスト OS 側から Tor を利用して目的の Web サイトへ HTML のみを要求し，保存する．(図 6 (a)) 次にその保存した HTML から画像コンテンツを示す URL をテキストに書き出し，どちらの経路で保存するか決定する(同図(b))．画像コンテンツを振り分ける方法として 5.4 と 5.5 の手法を提案している．そして各経路は 5.4 と 5.5 の手法で決めた画像をホスト OS 側とゲスト OS 側それぞれで取得し，共有フォルダーに保存する（同図(c),(d)）．画像の保存名は画像コンテンツの URL と同じものにする．最後に HTML ファイル内の画像コンテンツを示している部分を保存した画像の保存場所に書き換える．（同図(e)）

以上の 4 つの手順を踏むことで目的の Web サイトを閲覧することができる．具体例として本提案手法を適用して Tor を介した場合と適用せずに Tor を介した場合のブラウザを図 7 に示す．本提案で取得する Web ページの情報は画像を示す拡張子を持ったコンテンツと HTML ファイルのみである．そのため，その他のコンテンツは表示されていない．

提案手法適用前　　提案手法適用後

図 7　提案システム適用前後のブラウザ画面

6.2 実験環境

実験環境はゲスト OS 側も含め，実験で使用した Web サイトも 3. で使用したものと同様のものを使用した．

6.3 想定する状況

本論文では，攻撃者によるデータベースの作成方法について 2 通りの状況を想定して実験を行った．

case1 攻撃者は提案手法の適用されていない Tor を使用してデータベースを作成し，学習データ（データベース）として単一経路での通信から指紋情報を収集し，識別データ（ターゲット指紋情報）として複数経路の通信から指紋情報を収集する場合

case2 攻撃者は提案手法を適用した状態で Tor を使用してデータベースを作成し，学習データと識別データ共に複数経路の通信から指紋情報を収集する場合

case1 は，攻撃者は指紋情報収集フェーズにおいて Web サイト本来の指紋情報を収集することが可能だが，Web サイト特定フェーズにおいて Web サイト本来の指紋情報を収集することが出来ない．一方 case2 の場合は，攻撃者は指紋情報収集フェーズ，Web サイト特定フェーズのどちらの場合においても Web サイト本来の指紋情報を収集することが出来ない．このような実験を行った理由として前者現状における現実的な環境を想定したものであり，後者は本提案手法にとって厳しい環境を想定したものであり，実験を行うことで本提案手法の有効性を示そうと考えたためである．

6.4 実験方法

本実験は，3.4 と同様の Web サイトを使用して攻撃者は利用者として Tor プロキシ経由で Tor ネットワークへ接続する．トラフィックをパケットキャプチャする方法についても 3.4 同様 Wireshark を用いることとする．攻撃者が指紋情報データベースを作成する際は前述した 2 通りの方法を適用し，さらにそれに対し，2 通りの画像保存方法(5.4,5.5)を適用して実験を行う．サイトへの訪問回数は 3.4 と同様に Alexa のトップ 100 の Web サイトへ 20 回行う．

表 3　等数分割法における識別率

想定する状況	識別率
case1	1%
case2	28.5%

表 4　確率的重み付け配分法における識別率

想定する状況	識別率
case1	1%
case2	20.1%

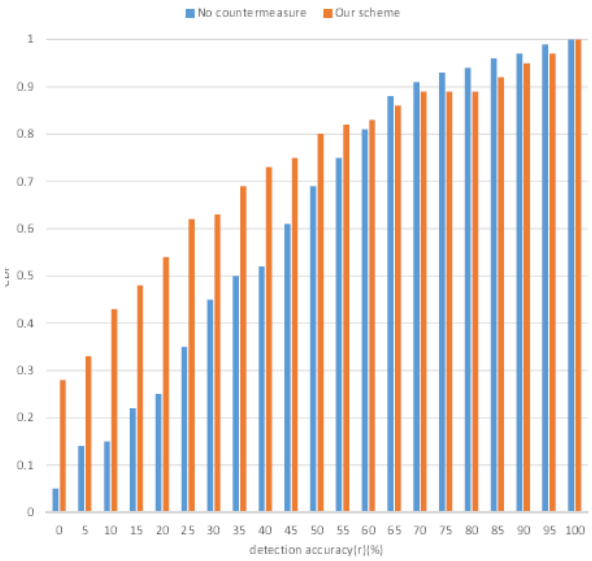

図 8　識別率に対する累積分布関数（頭数分割法）

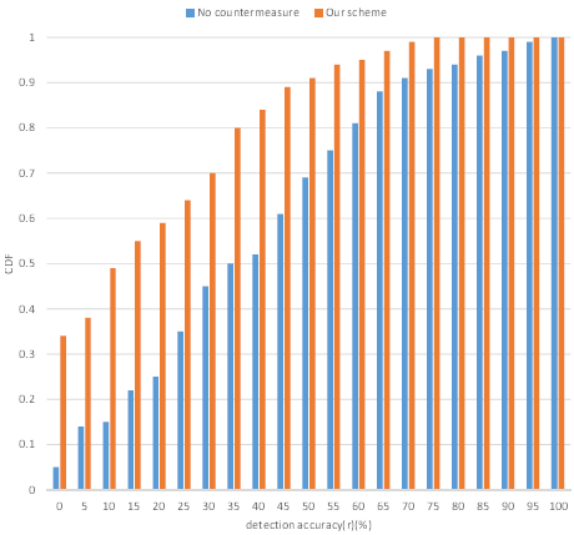

図 9　識別率に対する累積分布関数（確率的重み付け配分法）

6.5 実験と考察

表 3 と表 4 は，それぞれ等数分割法と確率的重み付け配分法において 6.3 で想定した case1，case2 それぞれの状況における識別率を示している．また図 8，図 9 はそれぞれ表 3 と表 4 の case2 における結果を累積分布関数（CDF）で表したものである．また，図 8 と図 9 については本実験結果に加え，3.4 の実験結果も含めて示している．

3.4 の実験による識別率が 39.65% だったのに対し，表 3，表 4 から攻撃者の行動を想定したいずれの環境においても識別率を低下させることができたことがわかる．

その中でもまず case1 の状況に注目すると，どちらの画像振り分け方法においても識別率が 1%という結果が得られた．例えば，利用者がアクセスした Web サイトを無作為に Web サイトを選んで識別しようとした場合，1%の確率でどの Web サイトにアクセスしたか識別できる．このことを考慮して，提案手法が攻撃者に対して全く指紋情報を与えていないといえる．

次に case2 の状況に注目すると，表 3 と表 4 の結果と 3.4 の実験結果を比較することで，等数分割法では 11.15 ポイント，確率的重み付け配分法では 19.55 ポイント識別率を抑えることが出来たとわかる．また図 8 において評価指標である 10%の値を見ると，「指紋攻撃への耐性が高い」Web サイトが 43%，「指紋攻撃への耐性が低い」Web サイトが 57%と提案手法を適用させることによって実験で使用した Web サイトの約 4 割の識別率を 10%以下に抑えることが出来た．しかし，識別率 65% ～95%の範囲を見てみると，提案手法を適用しない場合の方が若干だが良い結果になったことがわかった．この原因については解明できていない．次に，図 9 において評価指標である 10%の値を見ると，「指紋攻撃への耐性が高い」Web サイトが 49%，「指紋攻撃への耐性が低い」Web サイトが 51%と提案手法を適用させることによって実験で使用した Web サイトの約 5 割の識別率を 10%以下に抑えることが出来た．また，識別率 80%～100%の累積分布の値が 1 になっていることから全ての Web サイトの識別率を 75%以下に抑えることができたとわかる．

これらの実験結果から画像振り分け方法を確率的重み付け配分法にすることで特に指紋攻撃への耐性を発揮させることができるとわかった．

7. おわりに

本論文では，匿名通信システム Tor に対して匿名性を低下させる指紋攻撃についてその脅威を測り，それに対して複数の経路を用いた対策手法を提案した．そのためにまず，指紋攻撃が Tor にどの程度の脅威なのか実験を行うことにより検証を行った．その結果，Tor に対して十分な脅威となり，対策が必要であることを示した．また，現在の対策手法では根本的な解決にはなっていないことから，複数経路を利用した手法の特性を保持しつつ，既存の問題点を解決するような手法を提案した．提案した手法は，2 本の経路を用意し，一方の経路で HTML ファイルと画像コンテンツを取得し，もう一方の経路で残りの画像コンテンツを

取得する手法である．これにより，クライアント側での実装が可能なため，導入が容易になる．そしてこの提案手法が指紋攻撃に対してどの程度有効であるか実際に Tor ネットワークを利用することで評価を行った．その結果，全体の識別率を約半分に抑えることができた．また識別率 80%以上の Web サイト数を 0 にすることができ，実験で使用した半分の Web サイトの識別率を 0%にすることができた．本提案手法の目的は「指紋攻撃耐性あり」の Web サイトを増やすことであり，上記の実験結果からその目的を達成できたといえる．本提案はコンテンツの数を基に経路を振り分ける手法を提案したが，場合によっては指紋情報を分散させることができないほど片方の経路に特徴が表れる可能性がある．そこで今後の課題として，コンテンツのサイズを基準に重み付けを行うことで取得する経路を振り分ける手法を考え，検証する．

参考文献

[1] ITpro by 日経コンピュータ: NSA が Google および Yahoo!のデータセンターから通信傍受，米紙が報道, (online), available from(http://itpro.nikkeibp.co.jp/article/NEWS/20131031/515126/)(2014).
[2] Whiteld Diffie and Martin E. Hellman: NEW DIRECTIONS IN CRYP-TOGRAPHY, IEEE Transactions on Information Theory, Volume 22 Issue 6,pp.644-654 (1976).
[3] David Chaum, Communications Of The Acm, R. Rivest, David L. Chaum:Untraceable electronic mail, return addresses, and digital pseudonyms. Communications of the ACM Volume 24 Issue 2, pp.84-90 (1981).
[4] Tor Project Anonymity online, (online), available from (https://www.torproject.org/)(2014).
[5] Roger Dingledine, Nick Mathewson, and Paul Syverson. Tor: The SecondGeneration Onion Router, In Proceedings of the 13th USENIX Security Symposium (2004).
[6] U.S. Naval Research Laboratory: U.S. Naval Research Laboratory, (online),available from (http://www.nrl.navy.mil/)(2014).
[7] Ryan Pries, Wei Yu, Xinwen Fu andWei Zhao: A New Replay Attack AgainstAnonymous Communication Networks, In Proceedings of the 16th ACM conference on Computer and communications security, pp.578-589 (2009).
[8] Tor Metrics Portal:Directly connecting users, (online), available from (https://metrics.torproject.org/users.html), (2014.05.16).
[9] Panchenko, A, Niessen, L, Zinnen, A, Engel, T: Website Fingerprinting in Onion Routing Based Anonymization Networks, In Proceedings of the 10th annual ACMworkshop on Privacy in the electronic society, pp.103-113, (2011).
[10] 横手健一，松浦幹太: 匿名通信システム Tor の安全性を低下させるトラフィック逆加工, Computer Security Symposium 2012, Vol.3, pp.624-631, (2012).
[11] 横山絵美里，宗裕文，山場久昭，久保田真一郎，朴美娘，岡崎直宣：匿名通信システム Tor に対する指紋攻撃とその対策に関する検討，マルチメディア，分散，協調とモバイル (DICOMO2014) シンポジウム，pp. 498-505 (2014).
[12] Alexa: Alexa, The top 500 sites on the web, (online), available from , (http://www.alexa.com/topsites)(2014)
[13] Wireshark: Wireshark, (online), available from , (http://www.wireshark.org/)(2014).
[14] Andriy Panchenko, Lukas Niessen, Andreas Zinnen, and Thomas Engel: Website Fingerprinting in Onion Routing Based Anonymization Networks, In Proceedings of the 10th annual ACM workshop on Privacy in the electronic society, pp.103-114 (2011).
[15] Vitaly Shmatikov and Ming-Hsui Wang: Timing analysis in low-latency mixnetworks: Attacks and defenses, Computer Security ESORICS 2006, 11thEuropean Symposium on Research in Computer Security, pp.18-33 (2006).
[16] Andrew Hintz: Fingerprinting Websites Using Traffic Analysis, Privacy Enhancing Technologies, Lecture Notes in Computer Science Vol.2482, pp 171-178, (2003)