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

横山 絵美里[1]　宗 裕文[1]　山場 久昭[2]　久保田 真一郎[2]　朴 美娘[3]　岡崎 直宣[2]

概要：現在，パケットの通信を盗聴することで利用者がアクセスする Web サイトを特定しようとする行為が問題となっている．この問題解決のため，現在注目されている技術が匿名通信システムである．その中で普及しているシステムの一つとして The Onion Router(Tor) がある．Tor は複数のプロキシを経由することで高い匿名性を実現しているが，これを脅かそうとする攻撃が考案されつつある．その中でも代表的な Tor への攻撃手法として「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することで，利用者のアクセスする Web サイトを特定する手法であり，本論文では，指紋攻撃に対する耐性を持たせるような手法を検討する．そこで本手法では，通信時に読み込まれるトラフィックを，特徴の少ない情報である HTML ファイルと画像コンテンツに分離する．このとき，一方で HTML ファイルのみを，他方で利用者の読み込みたい画像コンテンツをそれぞれ Tor の別の経路を用いて読み込むようにすることで指紋攻撃の脅威を低減することを試みる．そして，この提案手法について実験を行うことで評価し，その有効性について示す．

An examination on countermeasure toward fingerprinting attack upon the Tor anonymity system

Emiri Yokoyama[1]　Hirofumi Sou[1]　Hisaaki Yamaba[2]　Shinichiro Kubota[2]　Mirang Park[3]
Naonobu Okazaki[2]

1. はじめに

近年，インターネットの急速な普及により，私たちは様々な情報を扱うことができるようになってきている．現在では誰もがこの技術を利用するようになっており，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．

現在，この問題解決のために暗号化通信技術 [1] と匿名通信システム [2] 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が可能であるが，誰が誰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普及しているシステムの一つが The Onion Router(Tor)[3], [4] である．今日この技術は，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が考案されつつある．これは利用者の秘匿情報が利用価値のあるものである場合，攻撃者が秘匿情報を特定することでその情報の悪用や Tor 利用者への脅しが考えられるためである．Tor の匿名性を脅かす攻撃の中でも，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」[5] が脅威になっている．この攻撃に対し，本論文では通信時に読み込まれる Web サイトトラフィックを制

[1] 宮崎大学大学院工学研究科
University of Miyazaki
[2] 宮崎大学工学部
University of Miyazaki
[3] 神奈川工科大学
Kanagawa Institute of Technology

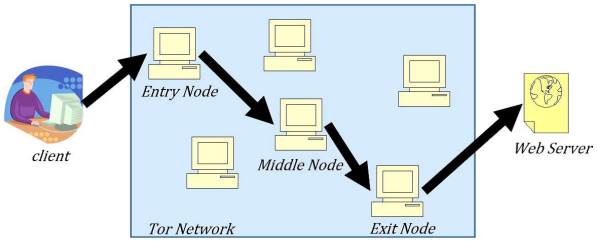

図 1　Tor の仕組み
Fig. 1　Structure of Tor network.

限するような手法を提案する．そしてこの手法がどの程度の耐性を持つのか実験を行うことで評価し，その有効性を示す．

2. The Onion Router

Tor は米海軍調査研究所 [6] が政府の通信保護を目的として開発された技術で，現在のユーザ数は約 200 万人である [12]．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自身のプライバシーを守りつつ情報をやり取りすることである [2]．具体例としては，虐待や深刻な病気など同じ境遇を持つ人々が行う情報共有や検閲の回避などがあげられる．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 は安全に通信を行えるツールとして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 4000 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．このようにして Tor はと匿名化通信を実現させる．

Tor は上記のような処理を行うことにより安全な通信を提供しているが，様々な手段を利用して Tor が実現する匿名性を低下させようとする攻撃が存在する．このような攻撃を行える理由は，Tor ネットワーク内の OR は全てボランティアにより構成されているため，攻撃者がノード群に攻撃を行う OR を含ませることが容易であるためである．以下，攻撃者により占拠された OR を汚染 OR と呼ぶこととする．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しなくてもトラフィックの特徴などから Web サイトを特定する手法も存在する．このような手法を「指紋攻撃」と言うが，これに対する根本的な対策はまだ確立されていない．本論文では実現性が高く，必要な資源が少ない「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 [5]．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．

[5] 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 [7] 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 [5] 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，既存のもの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

表 1　指紋情報の要素
Table 1　Fingerprinting information.

指紋情報の要素	説明
トラフィック総量（byte）	パケットサイズの総量
パケット総数	パケットの総数
トラフィック平均（byte）	パケットの平均サイズ
トラフィック分散	パケットの分散
チャンク平均（byte）	チャンクの平均サイズ
チャンク分散	チャンクの分散

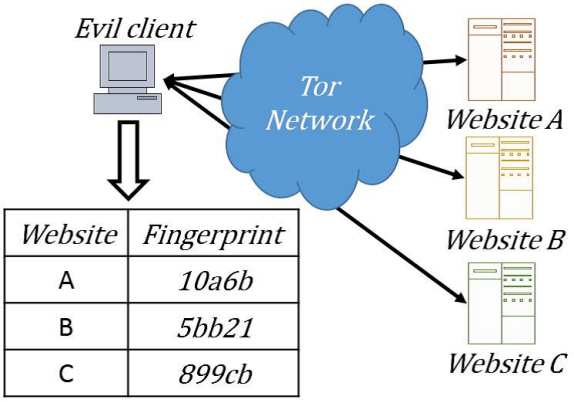

図 2　指紋情報収集フェーズの概略図
Fig. 2　Basic components in fingerprint information collection phase.

るものであり，攻撃者が収集するトラフィックから Web サイトの特徴になり得るものを抽出したもののことである．抽出した指紋情報はベクトル量で表される．本論文の指紋攻撃で設定した指紋情報は表 1 の 6 項目とし，さらにこの 6 項目にトラフィックの入出力を含めた 12 要素とする．ここで，チャンクはパケットの入出力の向き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みなしたものであり，その塊の合計サイズ (byte) とする．

3.3　処理手順

ここでは想定する指紋攻撃における処理手順について記述する．まず，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・Web サイト特定フェーズの二つに分かれる．

(1) 指紋情報収集フェーズ

このフェーズでは，攻撃者は図 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とき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．

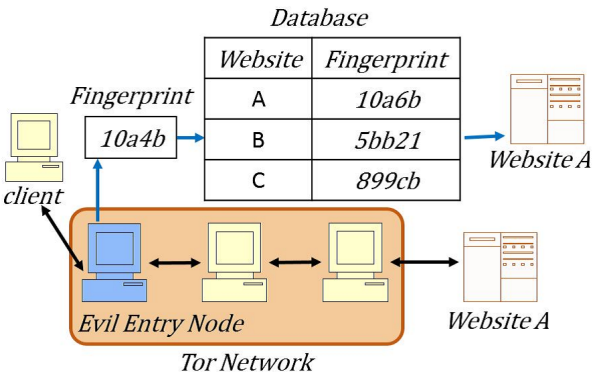

図 3　Web サイト特定フェーズの概略図
Fig. 3　Basic components in website specific phase.

表 2　PC の仕様
Table 2　Spec of PC.

OS	Windows 7 Professional
CPU	Core2 Duo E8400 3.00GHz
Browser	Mozilla Firefox 25.0.1
Tor	v0.2.3.25

(2) Web サイト特定フェーズ

このフェーズでは攻撃者は図 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と指紋情報データベースの比較にコサイン類似度を利用する．

3.4　実験

ここでは想定する指紋攻撃の実現可能性について調査を行い，その脅威について示す．

(1) 実験環境

攻撃者がデータベースを収集する際に使用する 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．このときの仕様を表 2 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 [10] 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だけ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 [11] 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

表 3　特定率 r_i における評価指標
Table 3　Evaluation indicators in specific rate r_i.

Web サイト特定率 (%)	指紋攻撃に対する耐性
$0 \leq r_i \leq 10$	指紋攻撃に耐性あり
$10 < r_i < 90$	指紋攻撃に耐性なし
$90 \leq r_i \leq 100$	指紋攻撃に脆弱

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．

$$r_i = \frac{Success_i}{Num} \times 100 \quad (1)$$

$$R = \frac{\sum_{i=1}^{Site} Success_i}{Num} \times 100 \quad (2)$$

ここで $Site$，Num，$Success_i$ はそれぞれ，訪問する総 Web サイト数，一つの Web サイトに対し訪問する回数，サイト番号 i の特定成功数である．本実験では $Site = 100$，$Num = 10$ 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 3 のように定義した．

同表で「指紋攻撃耐性あり」は Web サイトが特定される可能性が低いことを表し，「指紋攻撃耐性なし」は Web サイトを特定される可能性があることを示している．そして，「指紋攻撃に脆弱」は指紋攻撃を行われると高確率で特定される Web サイトのことを表している．よって，本研究の目的は「指紋攻撃耐性あり」の Web サイトが増えるような手法を見出すことである．

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．

3.5　実験結果・考察

本実験で Web サイト全体の特定率 56.8%という結果を得た．図 4 は特定率 r_i における Web サイト数を表しており，表 4 は実験による特定率 r_i における評価指標別のサイト数を表している．

図 4 から特定率が上がるほど Web サイト数が増加する傾向にあることがわかる．そして，表 4 から指紋攻撃の危険にさらされている Web サイトが全体の 93%を占めていることがわかる．ここから，現在の Tor ネットワークにおいて指紋攻撃は大きな脅威であり対策が必要であるといえる．

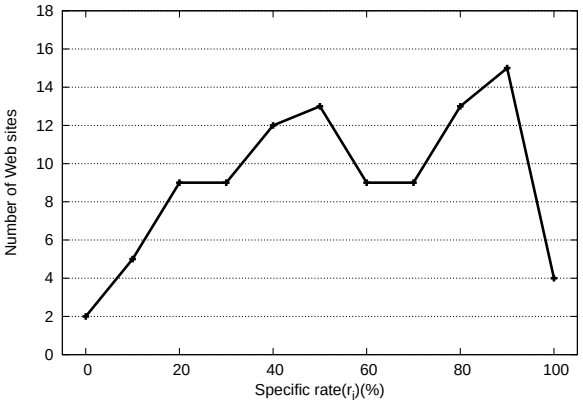

図 4　特定率 r に対する Web サイト数
Fig. 4　Number of website for specific rate r.

表 4　特定率 r における評価指標の結果
Table 4　Result of the evaluation indicators in specific rate r.

指紋攻撃に対する耐性 (%)	Web サイト数
指紋攻撃に耐性あり	7
指紋攻撃に耐性なし	74
指紋攻撃に脆弱	19

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．

[8], [9] 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．[8] の手法ではパケット到着時間上の確率分布に従い，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間隔を均一にして指紋攻撃を防ぐことができるが Tor に対する負荷が大きくなる．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．

[7] 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．例えば図 5 のように利用者が Web サイトへアクセスするときを考える．その際，目的の Web サイトとは別にランダムで Web サイトを選択し，要求を送る．これにより二つの Web サイトのトラフィックが混在して Tor ネットワーク内を通るので，攻撃者は利用者がアクセスする Web サイトを特定することが困難になる．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．

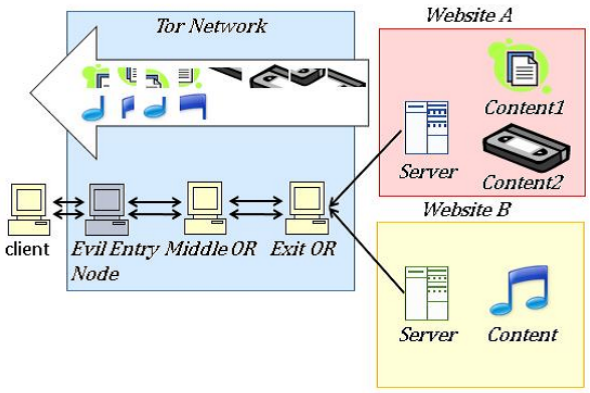

図 5　Web サイトトラフィックを利用した対策
Fig. 5　Fingerprint attack measures using traffic.

5. 提案手法

5.1 概要

3.5 の考察から指紋攻撃が Tor に対し，脅威となる攻撃であることがわかった．そのため，指紋攻撃に耐性を持たせるような手法を提案することを目的とする．そこで本論文では Web サイトの情報を分離し，異なる経路で読み込むことで指紋情報に差がつきにくくなるような手法を提案する．

5.2 前提条件

指紋攻撃は入口 OR を汚染する必要があるが，本提案手法では経路を 2 本用意するため，攻撃者が完全に指紋情報を収集するには二つの入口 OR を汚染する必要がある．しかし，利用者の経路上にある二つの入口 OR を同時に汚染できる可能性は非常に低い．そのため本論文では二つの入口 OR が同時に汚染される場合を考えず，HTML ファイルを読み込む側の経路が占拠された場合と，画像コンテンツを読み込む側の経路が占拠された場合に分けて考える．

5.3 提案手法とその考え方

本手法は，通常一度に読み込まれるはずの Web サイトの情報を HTML ファイルと画像コンテンツに分離してそれぞれ Tor 上の別の経路で読み込ませるようにすることで本来の Web サイトの情報を出にくくする手法である．複数の経路作成については後述する．このように Web サイトの情報を分離したのは，HTML ファイルは文字のみで構成されているため，画像コンテンツに比べ指紋情報に差が出ず，画像コンテンツについては利用者が自身のタイミングで読み込みたい画像コンテンツを選択するため，Web サイトに対する指紋情報が出にくくなると考えたためである．ここから，まず利用者はアクセスしたい Web サイトの HTML ファイルのみを読み込む．そして，もし利用者が読み込みたい画像コンテンツがある場合は，Tor 上の別の経路を使って当該コンテンツのみを読み込む．このとき，利用者の読み込みたい画像コンテンツを判別する必要がある．それに関しては画像コンテンツの表示される場所の上にボタンを作成することで対応する．ボタン作成方法についても後述する．

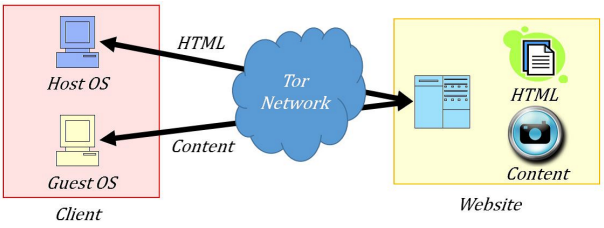

図 6　提案手法の仕組み
Fig. 6　Structure of the proposed method.

- 経路作成
  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．だが，提案手法の導入を容易にするためには同一 PC 内で同時に TorBrowser を起動させることが望ましい．そこで PC 内に仮想 OS を設け，同時に Tor を起動させる実験を行ったところ TorBrowser を同時に起動できることを確認した．
  本手法では利用者が Web サイトへアクセスする際にはホスト 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．このとき，図 6 のようにホスト OS 側では HTML ファイルのみを読み込み，ゲスト OS 側では画像コンテンツを読み込むこととする．
- ボタン作成
  ここではボタン作成の手順について記述する．まず利用者が HTML ファイルを読み込む際，画像を示す拡張子をさがす．画像拡張子を見つけるとその上にボタンを表示させる HTML 文を追加する．この動作については実現可能かどうか実験を行うことで確認を行った．このボタン作成の処理が完了すると図 7 の (a) のようにブラウザが表示される．利用者がブラウザを閲覧している際に読み込みたい画像コンテンツがある場合，そのコンテンツの上にあるボタンを押すとゲスト OS 側の Tor Browser が当該コンテンツのみを読み込み，ブラウザに表示する．そのときのブラウザの状態が図 7 の (b) である．

この手法により，通常の Web サイトのアクセス時に比べ，読み込むコンテンツ量が減少するので Tor に対する負荷が減少する．さらに，Tor 側ではなくクライアント側での実装が可能なので導入が容易である．

6. 評価実験

ここでは HTML ファイルのみを読み込む経路において 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

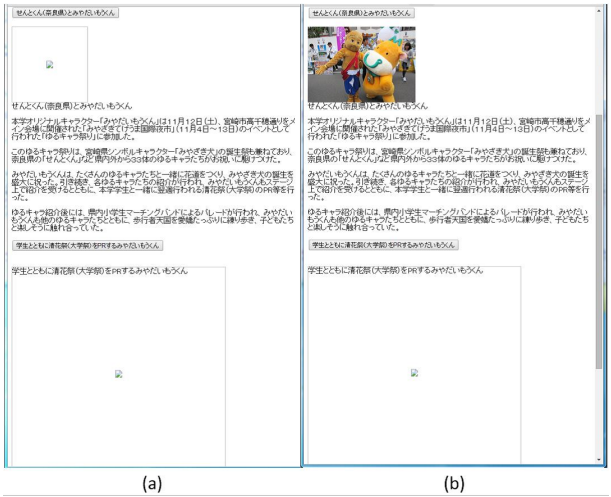

図 7　提案手法適用時のブラウザ画面
Fig. 7　An example of proposed method.

表 5　PC の仕様
Table 5　Spec of PC.

OS	Windows 7 Professional
CPU	Core(TM) i7-4770 CPU 3.40GHz
Browser	Mozilla Firefox 28.0
Tor	v0.2.3.25

をもつのか，そして画像コンテンツのみを読み込む経路において同様に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と 3. で示した既存手法とする．

6.1　実験環境

実験環境は 3. のときに加え，画像取得実験の際に攻撃者がデータベースを収集する PC とターゲットからトラフィックを収集する PC は表 5 に示される仕様の PC を利用する．また，実験で使用した Web サイトは 3. で使用したものと同様のものとする．

6.2　実装

本論文における提案システムの動作手順を図 8 に示す．本提案システムはサブシステム A およびサブシステム B の二つのサブシステムから構成される．動作は以下のようになる．

(a)　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 ファイルの要求を行う．
(b)　(a) で読み込んだ HTML に 5.3 で述べた手順でボタンを作成する HTML 文を追加する．
(c)　利用者が読み込みたい画像のボタンを押した際，要求された画像コンテンツの URL をテキストファイルに書き込み，共有フォルダに保存する．
(d)　共有フォルダから要求された画像コンテンツの情報を受け取る．

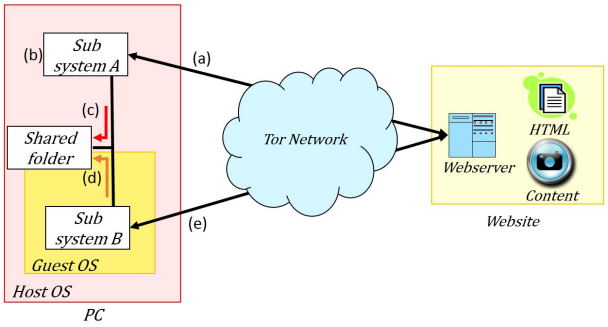

図 8　提案システムの概略図
Fig. 8　Basic components in the proposed system.

(e)　受け取った情報をもとにゲスト OS 側から Tor を利用して利用者の要求する画像コンテンツの読み込みを行う．

サブシステム A は (a) から (c)，サブシステム B は (d), (e) の動作を行う．本論文では HTML ファイルのみを読み込むことによる評価実験，そして画像コンテンツのみを読み込むことによる評価実験を行うため (a)，(e) の実装を行った．

6.3　実験方法

(1) HTML 取得実験
まず，攻撃者の指紋情報データベースの作成は，3. で選択した 100 サイトへ Tor を経由して 10 回ずつアクセスし，HTML ファイルのみを読み込むことで行う．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．パケットキャプチャの方法や特定率 r および R の求め方は 3. の実験と同様である．

(2) 画像コンテンツ取得実験
本実験では，利用者が Web サイトへアクセスした際に Web サイト総画像数の 50% の画像数を読み込むものと考える．実験簡略化のため，実験対象とする Web サイトを 3. で選択した 100 サイトのうち総画像数が 1 以上 150 以下の 77 サイトとした．攻撃者データベース作成時と指紋情報収集時に読み込む 50% の画像はランダムに選択し，互いに同じものとする．画像を読み込む順番は攻撃者データベース作成時と指紋情報収集時でランダムに決定する．さらに画像を読み込むタイミングについても攻撃者データベース作成時と指紋情報収集時で異なるものとする．攻撃者の指紋情報データベースの作成，指紋情報収集時には Tor を経由して対象の Web サイトへ各 1 回ずつアクセスする．パケットキャプチャの方法や特定率 R の求め方は 3. の実験と同様である．

表 6　実験による Web サイト全体の特定率 R

Table 6　Specific rate R of the entire Websites by experiment.

	既存手法	提案手法 (HTML のみ)	提案手法 (画像のみ)
特定率 R	56.8 %	34.7 %	37.7

表 7　提案手法適用による評価指標の結果

Table 7　Results of evaluation indicators by the proposal method aplied.

	既存手法	提案手法 (HTML のみ)
指紋攻撃に耐性あり	7%	24%
指紋攻撃に耐性なし	74%	73%
指紋攻撃に脆弱	19%	3%

表 8　脆弱な Web サイトに対する特定率の結果

Table 8　Results of specific rate in the vulnerable web sites.

	既存手法	提案手法 (HTML のみ)	提案手法 (画像のみ)
特定率 R	92.5 %	31 %	53.3 %

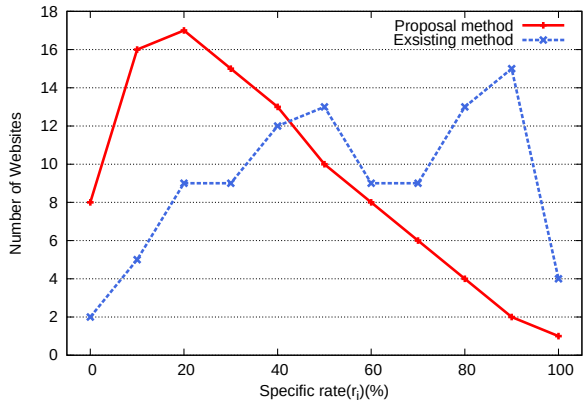

図 9　特定率 r に対する Web サイト数の変化

Fig. 9　Number of websites for specific rate r.

6.4　実験結果と考察

本実験により HTML ファイル取得時の全体特定率 34.7%，画像コンテンツ取得時の全体特定率 37.7% という結果を得られた．本実験で得られた結果を表 6 から表 8 および図 9 から図 11 に示す．表 6 は既存手法，提案手法 (HTML のみ)，提案手法 (画像のみ) での特定率，表 7 は提案手法 (HTML のみ) による評価指標の結果について示している．また，表 8 は脆弱な Web サイトに対する既存手法，提案手法 (HTML のみ)，提案手法 (画像のみ) の特定率について表している．図 9 は特定率 r における Web サイト数を示している．また，図 9 は本実験結果に加え，3. の実験結果も含めて示す．図 10 は 3. の実験により，指紋攻撃に脆弱と判断された Web サイトにのみに注目した提案手法による効果について，図 11 は Web サイトの画像コンテンツ数から 5 つに分類し，それに対する特定率の平均を表している．

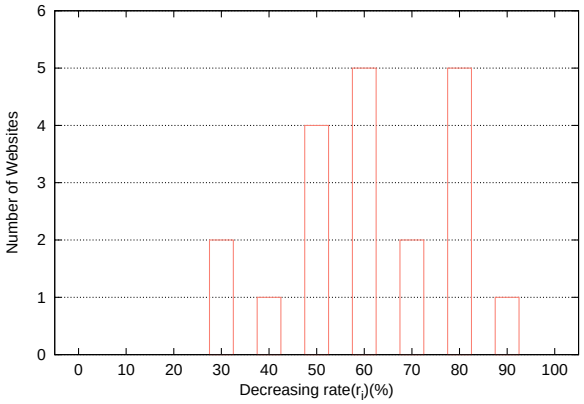

図 10　脆弱な Web サイトに対する特定率

Fig. 10　specific rate R in the vulnerable web sites.

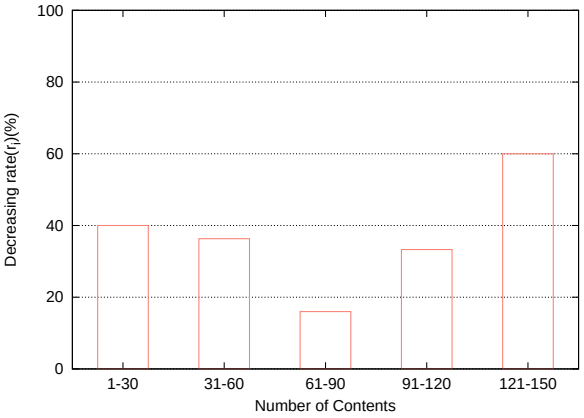

図 11　画像コンテンツごとの特定率

Fig. 11　Specific rate r of each image content.

表 6 と図 9 から本提案手法を適用することで多くの Web サイトに，指紋攻撃に対する耐性を持たせることができたと考えられる．また図 9 より，特定率が上がるほど Web サイトの数が減少していることがわかる．表 7 では，提案手法を適用させることで「指紋攻撃耐性あり」の Web サイトを 7%から 24%に増加させ，「指紋攻撃に脆弱」な Web サイトを 19%から 3%に減らすことができた．しかし，「指紋攻撃に耐性なし」についてはあまり変化がみられなかった．Web サイトの特定率を個別に見ると提案手法が全ての Web サイトに耐性を持たせているのではなく，逆に提案手法を適用することで匿名性を低下させている場合があった．それらの事例の共通点を調べると，特定率の上がった Web サイトに共通することは全ての Web サイトで，使用されている画像ファイルの更新が頻繁に行われていることがわかった．Web サイトの画像ファイルが頻繁に変化すると，そのときどきで読み込まれる指紋情報に変化が生じ，指紋情報データベースと収集した指紋情報が一致しなくなる．しかし，本提案手法では HTML ファイルのみを読み込むため，頻繁に画像ファイルが変化したとしても指紋情報は変化しない．これにより提案手法を適用した場合の方が特定率が高くなったと考えられる．表 8，図 10 は脆弱な

Web サイトに注目したものである．ここから本提案手法が脆弱な Web サイトに対して高い耐性を提供していることがわかる．また，表 8 から HTML のみ読み込む場合，画像のみを読み込む場合でそれぞれ特定率を抑えていることがわかる．そして，図 10 では脆弱な Web サイトの特定率を最大で 90%，平均すると 61.76%低下させることができた．さらに，全ての脆弱な Web サイトの特定率については 0%から 60%の範囲に抑えることができた．最後に，図 11 からは画像コンテンツが多いと特定率が高くなることがわかる．ここから大量の画像コンテンツは指紋情報になるといえる．また，画像コンテンツ数が 1 から 30 の特定率が高い理由として画像コンテンツ数が少ないとその分トラフィックにノイズが入る余地がなくなり，Web サイトごとの特徴が少なくなることが原因だと考えられる．

7. まとめ

本論文では Tor に対する攻撃の中で利用者の匿名性を低下させる指紋攻撃に注目した．そのためにまず，指紋攻撃が Tor に対し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，指紋攻撃が Tor に対して十分な脅威となり，対策が必要な攻撃であることを示した．そして実験を考慮して，利用者に読み込まさせる情報を最小限に抑えるような防御手法を提案した．この手法は Tor 側や Web サイト側で実装を行う必要がなく，クライアント側で即時に導入ができるため実現性の高い防御手法であるといえる．そしてこの提案に対し，防御手法の実現可能性と効果を示すために HTML を読み込む実験，画像コンテンツのみを読み込む実験を行うことで検証を行った．その結果，HTML を読み込む実験では全体の特定率を約半分に抑えることができ，脆弱な Web サイトに対しては平均 60%，最大で 90%特定率を下げることができた．画像コンテンツを読み込む実験でも全体特定率を約半分に抑えることができ，脆弱な Web サイトに対しては HTML を読み込む実験ほどではないが高い効果を示した．このことから，本提案手法は特に指紋攻撃に脆弱な Web サイトに対して最も効果を発揮できることがわかった．そして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ったが，HTML を読み込む実験により 3%から 24%にまで増加させることができた．今後はさらに「指紋攻撃耐性あり」の Web サイトを増やすための手法，提案手法を適用することで起こる特定率上昇に対する改善策の考案について検討する予定である．

参考文献

[1] Whitfield Diffie and Martin E. Hellman: NEW DIRECTIONS IN CRYPTOGRAPHY, IEEE Transactions on Information Theory, Volume 22 Issue 6, pp.644-654 (1976).

[2] David Chaum, Communications Of The Acm, R. Rivest, David L. Chaum: Untraceable electronic mail, return addresses, and digital pseudonyms. Communications of the ACM, Vol.24, pp.84-88, (1981).

[3] Tor Project: Anonymity online, (online), available from (`https://www.torproject.org/`), (2014.01.29).

[4] Roger Dingledine, Nick Mathewson, and Paul Syverson: Tor: The Second-Generation Onion Router, In Proceedings of the 13th USENIX Security Symposium Volume13, pp.303-320, (2004).

[5] Panchenko, A, Niessen, L, Zinnen, A, Engel, T: Website Fingerprinting in Onion Routing Based Anonymization Networks, In Proceedings of the 10th annual ACM workshop on Privacy in the electronic society, pp.103-113, (2011).

[6] U.S. Naval Research Laboratory: U.S. Naval Research Laboratory, (online), available from (`http://www.nrl.navy.mil/`)(2014).

[7] 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).

[8] Vitaly Shmatikov and Ming-Hsui Wang: Timing analysis in low-latency mix networks: Attacks and defenses, Computer Security ESORICS 2006, 11th European Symposium on Research in Computer Security, pp.18-33, (2006).

[9] Andrew Hintz: Fingerprinting Websites Using Traffic Analysis, Privacy Enhancing Technologies, Lecture Notes in Computer Science Vol.2482, pp 171-178, (2003)

[10] Alexa: Alexa, The top 500 sites on the web, (online), available from , (`http://www.alexa.com/topsites`), (2014.01.29)

[11] Wireshark: Wireshark, (online), available from , (`http://www.wireshark.org/`), (2014.01.29)

[12] Tor Metrics Portal:Directly connecting users, (online), available from (`https://metrics.torproject.org/users.html`), (2014.05.16).