Vol.2017-IOT-36 No.31
2017/3/4

Torを利用して掲示板書き込みを行うユーザの特定手法

飯田 嘉一郎[1]　吉浦 紀晃[1]

概要：近年では個人情報保護の意識が高まっており，個人情報を伏せたままの通信を可能にする匿名化通信が注目されている．匿名化通信を実現するもので最も普及しているのは Tor(The Onion Router) である．Tor はインターネットにおいて通信経路の匿名化を実現できるが，この匿名化の犯罪利用が非常に多く，早急な対策が必要とされている．本論文では，Tor を利用してインターネット掲示板に悪質な書き込みをするユーザを特定することを目的とする．本論文では，Tor の通信をキャプチャし，そのパケットの特徴を Fingerprint とする．この Fingerprint を利用して，掲示板へ書き込みを行ったユーザを特定するシステムを提案する．さらに，本論文では，この機能を実装し，Tor ネットワーク上で実験を行なった．そして，Fingerprint を利用した従来のユーザ特定手法よりも高い精度でユーザの特定が可能となった．

キーワード：匿名化通信，Website Fingerprint，通信分析，Tor

User identifying method of writing bulletin board using Tor

Kaichiro Iida[1]　Noriaki Yoshiura[1]

Abstract: In recent years, awareness of protection of personal information has been increased, and anonymous communication that enables communication with hiding personal information attracts attention. Tor (The Onion Router) realizes anonymous communication most widely. This anonymization is used for crime very much and urgent measures for it are required. This paper aims to identify users who maliciously write to the Internet bulletin board using Tor. This paper captures Tor's communication packet and consider the feature of that packet as Fingerprint. Using Fingerprint, this paper proposes a system that identifies users who have written to the bulletin board. This paper also implements this function and conducts experiments on the Tor network. As a result of experiment, users can be identified with higher accuracy than the conventional user identification method using Fingerprint.

Keywords: Anonymous Communication，Website Fingerprinting，Traffic Analysis，Tor

1. はじめに

近年では個人情報保護の意識が高まっており，個人情報を伏せたままの通信を可能にする匿名化通信が注目されている．匿名化通信を実現するソフトウェアの中でも特に「Tor(The Onion Router) [9]」が世界中で多く利用されている．このソフトウェアは Web サーバ等と通信をする際，Tor をインストールしたサーバを複数経由して通信することで匿名化通信を実現している．

しかし，Tor の匿名性が犯罪利用される事例は非常に多い [8]．Tor の犯罪利用には違法薬物売買やインターネットバンキングへの不正アクセス等が挙げられるが，本論文ではインターネット掲示板への悪質な書き込みに着目する．理由として，インターネット掲示板への書き込みが，違法薬物売買や犯罪予告等の多くの犯罪に利用されるからである [17]．現在，Tor の犯罪利用に対する取り締まりは通信元を直接特定するものではなく，複数の状況証拠と家宅捜索によって犯人を逮捕する等，間接的な特定にとどまっている．直接的な特定に比べ，間接的な特定は特定作業に膨大な時間と労力を必要とするため，現状のような通信元の

[1] 埼玉大学大学院 理工学研究科
Graduate School of Science and Engineering,
Saitama University

Vol.2017-IOT-36 No.31
2017/3/4

直接的な特定が困難である状況は望ましくない．

本論文では Tor を利用してインターネット掲示板に悪質な書き込みをするユーザを特定することを目的とする．本論文で提案する特定手法は，キャプチャした Tor ネットワークの通信の特徴から，掲示板へ書き込みを行なったユーザを特定する，という手法である．

関連研究として，Web サイトの通信の特徴を利用したユーザ特定を行うものがある [1]．これは Tor ユーザからの通信を最初に経由するサーバにおいて、通信のキャプチャが可能である場合に成立する手法だった．しかし，Tor ネットワークの通信は暗号化とデータ量の変更がされているため特定が難しく，特定率 54%と精度が高くなかった．そこで，本論文では従来のユーザ特定手法とは違い，特定する対象を Web サイト全体からインターネット掲示板への書き込みのみに限定することで，高精度のユーザ特定手法の実現を目指す．

本論文では，「掲示板に書き込みを行なったユーザを特定する」という機能を実装し，実際に Tor ネットワーク上で実験を行なった．実験では，Tor ネットワークを経由して通信をした大量のパケットの中から，指定した掲示板に書き込みを行なったパケットを識別し，その掲示板への書き込みの判定が可能であるか確かめるために実験を行なった．実験の結果，従来のユーザ特定手法とは違い，ユーザが掲示板への書き込みを行なったという判定とそのユーザの特定を特定率 97%で成功させることができた．

本論文の構成は，2 章に Tor の概要や Tor ネットワークについての説明を述べ，3 章では本論文の関連研究について述べる．4 章では本論文の実験で使用したシステムに関する説明を述べる．5 章では本論文の実験の概要を説明し，6 章で実験結果，7 章で実験結果を受けて，考察を述べる．最後に 8 章でまとめを述べる．

2. Tor(The Onion Router)

2.1 Tor の概要

Tor は Web 閲覧の際のプライバシーの保護と匿名化を目的として開発されたオープンソースのソフトウェアであり，現在は Tor Project によって開発が続いている [9]．Tor によって通信を匿名化する OS である Tails も存在する [10]．

通常，Web サイトにアクセスする場合，Web サーバのログファイルに通信元の IP アドレスが記録される．そして，このログファイルに残された IP アドレスと日時によって通信元を特定することができる．プロキシサーバを経由して Web サイトにアクセスした場合，プロキシサーバの IP アドレスが Web サーバに記録され，プロキシサーバに通信元の IP アドレスが記録される．よって，プロキシサーバに記録された IP アドレスを入手することで通信元を特定することができる．一方，Tor を利用して Web サイトにアクセスする場合は通信元の特定が困難になっている．Tor は世界中の Tor ユーザが提供する多数のノードで Tor ネットワークを構築し，Tor を利用した通信は Tor ネットワークを経由する．また，通信が経由するノードの中でもユーザからの通信が最後に経由するノード (Exit ノード) 以外のノードは Web サーバと直接通信しないため，Web サーバはどのノードが通信元なのかが分からない．通信内容も通信が経由するノード間で暗号化されているため，経由したノードも通信元がどこへアクセスしているのかが分からない仕組みになっている．そのため，Web サイトにアクセスする際の匿名化が成立している．

Tor の匿名性は内部告発や政治的な発言に使われるだけでなく，インターネット検閲の回避にも利用される．しかし，中国の The Great Firewall(GFW) [4,5] と呼ばれる巨大なインターネット検閲システムでは高い精度で Tor の通信を検閲している [6]．GFW はインターネット利用者が中国政府に不都合な通信を行おうとした場合，通信を即座に遮断するために導入されたものである．2011 年，GFW には Tor による通信も全て遮断するという機能が追加され，Tor による通信がブロックすることができるようになった [12]．そこで Tor Project では GFW に検知される Tor の通信の特徴をなくすツールの開発等で GFW に対抗している [6]．

2.2 Tor ネットワークについて

Tor ネットワークは世界中の Tor ユーザによって提供されるノードで構築される．Tor ネットワークを構築するノードは大きく 3 つに分類される．

Entry ノード
: Tor ユーザからの通信が最初に経由するノード．Tor ネットワーク内で唯一 Tor ユーザと通信を行うノードのため，デフォルトでは Guard ノード (条件を満たした優良なノード) が選ばれる設定になっており，世界中で約 2500 の Guard ノードが稼働している [11]．

Middle ノード
: Tor ネットワーク内のノードからの通信が他のノードに経由するノード．

Exit ノード
: Tor ネットワーク内のノードから Web サーバに通信を経由をするノード．実際に Web サーバにアクセスする．

Tor を利用してある Web サーバに接続する場合，以上の 3 種類のノードを経由して通信を行う．

2.3 Tor ネットワークの特徴について

Tor を利用して Web サイトにアクセスする際，Tor ネットワークの特徴として，以下が挙げられる．

ログの記録

Tor では図 1 のように通信を経由するが，通信を経由するノードで経由のログを記録しない．ログを記録するノードが存在する可能性もあるが，Web サーバに記録されたログから Tor ユーザを特定するためには，通信を経由した全てのノードのログを入手する必要がある．通信を経由した全てのノードがログを記録している可能性が極めて低いため，Tor ユーザが特定される可能性は極めて低い．

経由ノードの変更

通信を経由するノードは一定時間毎に変更される．通信を経由するノードが不正アクセスを受けたり，不正を行うノードを経由している場合でも，通信を経由するノードが一定時間で変更されるため，Tor ユーザを特定することは困難である．

通信経路の秘匿化

通信を経由するノードの内，Exit ノード以外では通信が暗号化され通信経路が秘匿化される．そのため，通信先だけでなく，Exit ノード以外の通信を経由するノードにも通信内容を知られずに Web 閲覧を行うことが可能である．Exit ノードに通信内容を知られてしまう可能性があるが，IP アドレスは知られないため Tor ユーザが特定される可能性が低い．また，通信経路の秘匿化がされるため，図 1 のノード B は前後のノード A，C しか知ることができない．そのため，Entry ノード以外の経由するノードが Tor ユーザを特定することは困難である．

パケットのパディング

Tor の通信はパケットの盗聴によってユーザの匿名性が損なわれないように，パケットがそれぞれパディングされている [7]．パディングによってパケットのサイズが変更されているため，パケットのサイズの Fingerprint を利用した匿名性を損なう攻撃がしにくくなっている．

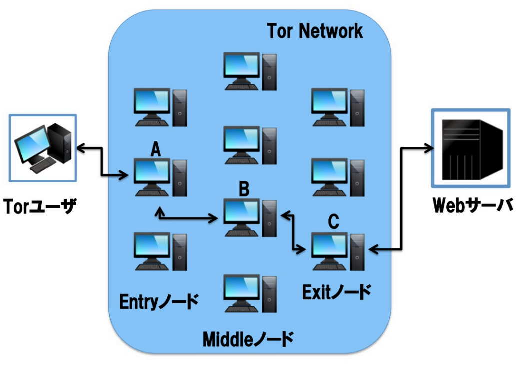

図 1　Tor ネットワークを経由する通信
Fig. 1　Communication via Tor network

以上の特徴によって，Tor ユーザを特定することが困難になり，Web サイトにアクセスした際の匿名性が保護される．

Tor には利点だけでなく欠点も存在する．Tor を利用した場合，Tor ユーザは Tor ネットワークを経由して Web サイトにアクセスする．そのため，Tor ネットワークの通信速度は提供されるノードに依存する．したがって，一つでも通信速度や処理速度が遅いノードを経由してしまうと通信速度が遅くなる．

2.4　Tor Browser

現在，Tor の利用は Tor Browser [9] が主流である．Tor Browser は Tor Project によって提供される Tor を利用した通信を行う Web ブラウザであり，特別な設定を必要としないため簡単に利用することができる．

3.　関連研究

3.1　Tor ネットワーク内での Web サイトの Fingerprint を利用した攻撃

Andriy らの研究 [1] では，Tor ネットワーク内で暗号化されている通信に対して Web サイト固有の Fingerprint を利用し，Web サイトを特定する手法を提案した．この手法では通信しているパケットの「パケットの送受信の順序」や「通信パケット長」，「日時」から特定する手法である．実際に Tor ネットワークの Entry ノードからこの手法で Web サイトを特定することができれば，Entry ノードから Tor ユーザと Web サイトの双方が分かるため Tor の匿名性を損なわせることが可能になる．Andriy らの研究ではこれらの Fingerprint を利用し，Web サイトへのアクセスデータから実際にアクセスした Web サイトを特定した．この実験での特定率は 54%ほどであった．

本論文では Andriy らの研究の Fingerprint を利用した手法をベースとし，Web サイト全体の特定をするのではなく，掲示板への書き込みだけを特定することを目的としている．掲示板への書き込みは Web サイトの閲覧とは違い，データを Tor ユーザから送信する．よって Fingerprint が特徴的であるため，特定率の向上が期待できる．

3.2　特定率の高い Tor ネットワーク内での Fingerprint 攻撃

Xiaogang らの研究 [2] と Z.Ling らの研究 [3] では高い特定率の Fingerprint 攻撃を成功させている．これらの二つの手法に言えるのは，より Fingerprint を検知しやすいものに細工することで特定率を向上させたことである．

Xiaogang らの研究では攻撃者が管理する Entry ノードと偽コンテンツ，Exit ノードを用意する．管理下の Exit ノードから Web サイトに要求があった場合，管理下の Exit ノードは偽コンテンツを混入させてパケットを経由する．偽コ

ンテンツにはリンク先のないイメージタグが挿入されているため，Web ブラウザは画像を取得しようとする．管理下の Entry ノードはこの時に発生する通信の Fingerprint を利用して Tor ユーザと Web サイトを特定する手法である．

一方 Z.Ling らの研究 [3] では攻撃者が管理する Entry ノードと Exit ノードを用意する．Tor ユーザが管理下の Entry ノードと Exit ノードを利用している場合，Exit ノードから Tor ユーザに向かっていく通信に細工を行うことで攻撃が成立する．管理下の Exit ノードが Tor の通信で発生するパケットの数を少し増やすことで，通常の Tor の通信では発生しない Fingerprint を作り出す．その Fingerprint を利用して Tor ユーザと Web サイトを特定する手法である．

これらの研究は強力な攻撃を成功させているが，管理下の Entry ノードと管理下の Exit ノードを Tor ユーザが同時に利用してなければならないため，実現性が低い．本論文では Entry ノードを Tor ユーザが利用している時に成立するものであるため，実現性が高くなっている．

4. システム

4.1 システム概要

本論文では，キャプチャしたパケットのデータを元に，特定の掲示板への書き込みを判定するシステムを提案する．このシステムは，Entry ノードが管理下にあるという条件と Entry ノードでキャプチャしたパケットのログが残っているという条件を満たしている場合，任意の書き込みとそれを書き込んだ Tor ユーザの特定ができるシステムである．

このシステムは Andriy らの研究 [1] で提案している手法のように，キャプチャしたログの「パケットの送受信の順序」や「通信パケット長」，「日時」の 3 つの Fingerprint から特定を行う．異なる点として，Fingerprint に掲示板書き込みの際に発生する Tor ユーザからの送信パケットが含まれるため，特定率の向上が期待できる．

4.2 掲示板書き込み判定プログラムについて

実際に実験に使用した掲示板書き込み判定プログラムについて記述する．暗号化されている通信からどの掲示板に書き込みがされたかを特定する際に，通信パケットの Fingerprint を利用して判定を行なっている．その Fingerprint として「パケットの送受信の順序」や「通信パケット長」，「日時」を利用して掲示板書き込みの有無を判定する．

掲示板書き込み判定プログラムの作成方法を述べる．掲示板書き込み判定プログラムを作る場合は以下の手順を踏むことで掲示板書き込み判定プログラムを作る．

(1) 対象の掲示板へ 10 回程度書き込みを行い，そのパケットをキャプチャする．
(2) キャプチャしたパケットから，Fingerprint になるような，同じパケットの並びを手動で探す．
(3) 同じパケットの並びを Fingerprint とし，Fingerprint を見つけたら「掲示板書き込み有り」と判定するプログラムにする．

Wireshark でキャプチャしたログファイルを，掲示板書き込み判定プログラムでは利用する．Wireshark でキャプチャしたログファイルには，「パケット概略部，パケット詳細部，パケット詳細の 16 進数表示部」の 3 つの情報が記録してある．パケット概略部にはパケットの送受信のアドレスや時間，通信パケット長などの情報が 1 行で記載してある．パケット詳細部には，パケット概略部にある情報が数行に渡って記載しており，パケット詳細の 16 進数表示部にはこれらの情報が 16 進数で記載してある．

掲示板書き込み判定プログラムはそのパケット概略部を読み込むことでパケットの Fingerprint を判定する．図 2 に掲示板書き込み判定プログラムのフローチャートを示す．このプログラムの処理時間はログファイルのデータ量に比例して増加していき，120MB で 2.8 秒程度，1.5GB で 34 秒程度であった．以上の手法を利用して様々な掲示板に対して，掲示板書き込み判定プログラムを作ることが出来る．

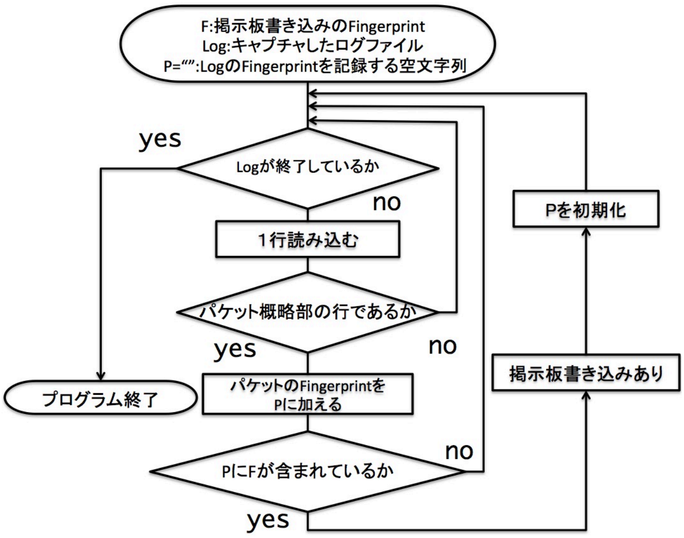

図 2 掲示板書き込み判定プログラムのフローチャート
Fig. 2 Flowchart of determination program for bulletin board writing

4.3 システム構成

実験を行なったシステム構成を図 3 に示す．図 3 の構成は Entry ノードを固定し，他の Middle ノードと Exit ノードは Tor の動作にしたがって経路変更がされるという構成である．これは，実際に運用する場合に Entry ノードが管理下にあるため変更されず，他のノードは管理下にないため変更されるという状況に近づけるためこのように構成した．実験では，Tor によって暗号化された通信をキャプ

チャするために Tor ユーザから Entry ノードへの通信をキャプチャする．

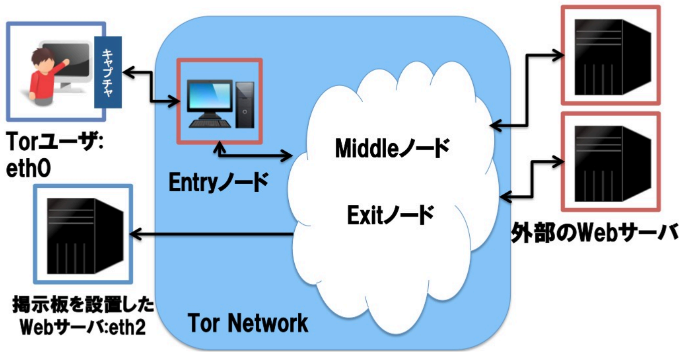

図 3　システム構成図
Fig. 3　System configuration

4.4　実装環境

この実験を行う上で使用したものは，図 3 のように，Tor ユーザの PC と Tor ネットワーク内のノード，Web サーバである．これらの中で筆者の管理下にあるものは Tor ユーザの PC と実験で使用する掲示板 CGI プログラムを設置した Web サーバのみで，他は管理下ではない．

実験で使用した掲示板 CGI プログラムについて説明する．掲示板 CGI プログラムは KENT-WEB [13] という Web サイトからダウンロードしたもの 5 つと，59bbs [14] という Web サイトからダウンロードした ThreadPlus を使用した．また，これらの掲示板 CGI プログラムに加え，KENT-WEB からダウンロードした ASKA BBS を改造した ASKA.V2 も用意した．KENT-WEB は CGI/Perl プログラムを無料で配布する日本最大の Web サイトであり，59bbs も同様にオープンソースの CGI を扱っている Web サイトとして多く利用されている．よってこれらの掲示板 CGI プログラムは一般に利用されているものといえる．

実験で使用した掲示板は ASKA BBS，COM BOARD，GateWayBoard，LIGHT BOARD，ThreadPlus，Webforum，ASKA.V2 の 7 つであり，全て Perl のプログラムである．

ASKA BBS
　名前，メールアドレス，タイトル，メッセージ，参照先，削除キー，画像認証が付いているシンプルな掲示板システムである．投稿への返信機能も付いている．

COM BOARD
　フレーム機能を使った掲示板であり，コメント表示部と投稿部が分かれている．ASKA BBS とほぼ同機能を持った掲示板システムである．

GateWayBoard
　ログインしなければアクセスできない掲示板であり，投稿フォームは上記の掲示板と同様である．

LIGHT BOARD
　ASKA BBS とほとんど同機能の掲示板であり，投稿記事を管理者の承認後に表示させることも可能である．

ThreadPlus
　スレッド型の掲示板である．フォームは ASKA BBS 等と同様の項目であり，パスワードを設定することで，パスワードを知っているユーザにだけ自身の投稿の閲覧を許可することもできる．

Webforum
　ツリー表示式のスレッド型掲示板であり，他の掲示板と違い，画像アップロード機能がある．

ASKA.V2
　ASKA BBS と全く同機能の掲示板である．掲示板への書き込みの際にやりとりする，Web サーバ側からの送信パケットで無駄な文字列を送信することで，パケットのやり取りに特徴づけがしてある．

5.　実験

5.1　実験概要

本論文では以下の実験を行なった．

実験 1
　提案手法の有効性評価のため，掲示板書き込み判定プログラムの精度を評価する．実験では， ASKA BBS と ASKA.V2 への書き込みをしたパケットを収集し，収集したパケットを掲示板書き込み判定プログラムで判定した．実際に書き込みを行なった回数と，掲示板書き込みの判定回数を比較し，掲示板書き込み判定プログラムの精度を評価した．

実験 2
　掲示板への書き込みの誤判定率の調査のため，複数の掲示板を対象に掲示板書き込み判定プログラムの精度の評価を行う．そして，掲示板書き込み判定プログラムによって 7 つの掲示板をそれぞれ特定をする実験を行なった．

実験 3
　様々な Web サイトへのアクセスを行なった大量のパケットから，掲示板への書き込みを判定することで掲示板書き込みの判定精度を評価を行う．キャプチャした大量のパケットの中から特定の掲示板への書き込みの判定をする実験を行なった．

実験 4
　Tor ユーザが複数の通信を同時に行なっている場合の，掲示板書き込み判定プログラムの精度の変化について評価を行う．実験では，掲示板への書き込みと他の Web サーバとの通信を同時に行い，Tor の通信パケットをキャプチャした．そして，キャプチャしたパケットの中から特定の掲示板への書き込みを判定する実験を行なった．

6. 実験結果

6.1 実験 1

この手法の有効性を評価するために，Web サーバに掲示板を用意し，掲示板書き込み判定プログラムによって判定を行う実験を行なった．実際に設置した掲示板は ASKA BBS と ASKA.V2 である．掲示板への書き込みを 100 回行い，その書き込みのうち，どれだけの書き込みを判定することができたか確かめた．

表 1 ASKA BBS への書き込みを判定プログラムで判定した結果
Table 1 The result of judging programming by writing to ASKA BBS

書き込み回数	判定結果 (判定回数)	判定率
100 回	書き込みを判定 (99)	99%

表 2 ASKA.V2 への書き込みを判定プログラムで判定した結果
Table 2 As a result of judging write to ASKA.V2 by the judgment program

書き込み回数	判定結果 (判定回数)	判定率
100 回	書き込みを判定 (97)	97%

実験結果は表 1 の通りである．判定できなかった書き込みは，掲示板書き込みの際のパケットロスや，パケットが通常と異なるサイズで送られてくる事象が稀に発生したため，判定できなかった．

6.2 実験 2

Web サーバに掲示板を 7 つ用意し，実際に判定をする実験を行なった．実際に設置した掲示板は ASKA BBS，COM BOARD，GateWayBoard，LIGHT BOARD，ThreadPlus，Webforum，ASKA.V2 の 7 つである．4.2 節に述べた手法で，それぞれの掲示板に対して表 3 のように掲示板書き込み判定プログラムを作った．

表 3 掲示板 (略名) と掲示板書き込み判定プログラム
Table 3 bulletin board (short name) and bulletin board write determination program

掲示板 (略名)	プログラム	掲示板 (略名)	プログラム
ASKA BBS(AS)	aska-id	ThreadPlus(TP)	thpl-id
COM BOARD(CB)	combbs-id	Webforum(Wf)	wforum-id
GateWayBoard(GWB)	gwbbs-id	ASKA.V2(A2)	askav2-id
LIGHT BOARD(LB)	light-id		

実験では，掲示板に 10 回書き込みを行なったパケットをキャプチャし，そのパケットを掲示板書き込み判定プログラムで判定した．この実験は，全く同じ「日時」で掲示板に書き込みがあった場合に誤判定をしてしまうかを実験しているため，Fingerprint の一つである「日時」は無視して実験を行なった．

表 4 判定プログラムの掲示板誤判定率
Table 4 Misjudgment rate of judgment program

プログラム	掲示板の略名						
	AS	CB	GWB	LB	TP	Wf	A2
aska-id	-	100%	100%	0%	60%	0%	0%
combbs-id	0%	-	100%	60%	70%	0%	0%
gwbbs-id	100%	100%	-	100%	100%	0%	0%
light-id	0%	100%	100%	-	100%	0%	0%
thpl-id	0%	100%	100%	80%	-	0%	0%
wforum-id	0%	0%	0%	0%	0%	-	0%
askav2-id	0%	0%	0%	0%	0%	0%	-

表 4 は縦軸の掲示板に対して横軸の掲示板書き込み判定プログラムがどのくらいの割合で誤った判定をしたかを表している．この誤った判定とは，別の掲示板に対して別の掲示板書き込み判定プログラムが「掲示板への書き込みがあった」という判定をすることである．つまり，0%の項目が多いほど，特定しやすい掲示板となる．この時，特定しやすい掲示板は，「掲示板書き込み判定プログラムが他の掲示板に対して誤った判定をせず，他の掲示板書き込み判定プログラムから誤った判定をされないもの」である．表 4 から ASKA.V2 と Webforum であると言える．

表 4 より，「Webforum，ASKA.V2」以外の 5 つの掲示板はパケットの Fingerprint から誤判定してしまう場合があることが確認できた．これらの掲示板は，メッセージ投稿の際に発生するパケットの送受信パターンが全く同じであり，かつパケット数が 8 個から 12 個と非常に少なかったためである．やりとりを行うパケット数が少ないことで，Fingerprint が同一になってしまう場合があるが，インターネット上の掲示板ではもっとパケット数が多い．実際に Yahoo!掲示板 [16] という Web サイトに書き込みをした時のパケットのやり取りが数百であったことから，実験で使用した掲示板のやり取りするパケット数が少ないことがわかる．パケット数が多いことで，この手法で利用している Fingerprint が同一になることが少なくなるため，この手法の有効性がさらに高くなると推測できる．

一方，Web サイトのソースコードを編集し，通信に特徴づけをすることで誤判定を回避できることも確認できた．これは ASKA BBS と ASKA.V2 の結果から言える．ASKA BBS を判定するプログラムである aska-id は COM BOARD，GateWayBoard，ThreadPlus を誤判定している．この結果から，ASKA BBS のパケットの Fingerprint は COM BOARD，GateWayBoard，ThreadPlus と似ていたということがわかる．改造後の ASKA.V2 を判定する askav2-id は他の掲示板を誤判定せず，他の掲示板書き込み判定プログラムも ASKA.V2 を誤判定しなかった．これは掲示板を改造することで他の掲示板とパケットの Fingerprint が被らないようにできることがわかる．

6.3 実験 3

Tor Browser で様々な Web サイトにアクセスし，キャプチャした大量のパケットの中から任意の掲示板への書き込みを判定できるかという実験を行なった．掲示板や SNS などのフォーム入力を含めた，ランダムな 1000 個の Web サイトにアクセスを行なった．1000 個の Web サイトにアクセスをしながら，実験 2 で利用した ASKA.V2 に書き込みを行い，実際に掲示板への書き込みがあったという判定が可能か実験した．実験では Web サイトの読み込みが終わってから次の Web サイトにアクセスを行なったため，2 つ以上の Web サイトに同時にアクセスはしていない．Web サイトの中には Tor の通信をブロックする Web サイトが複数確認された．また，Web サイトはアクセスできるが掲示板に書き込むことが出来ない掲示板も複数確認された．これらの掲示板は Web サーバに記録される，リファラというデータが記録されない時に書き込みを拒否していることがわかった．Tor Browser はこのリファラを Web サーバに記録しないため，書き込みが出来なかった．

実験結果は表 5 のようになった．実験結果では 100%の精度で掲示板書き込みを判定することができた．つまり，掲示板書き込み判定プログラムが，様々な他の Web サイトのパケットを「ASKA.V2 への書き込みパケットである」と誤判定しなかったことがわかる．パケットをキャプチャした Web サイトには掲示板サイトや SNS，ブログへのコメント投稿を含んでおり，テキストを送信する他の Web サイトと ASKA.V2 との判別が可能であることが確認できた．

ここで表 2 の askav2-id の判定精度と表 5 の結果から，大量のパケットの中から askav2-id が ASKA.V2 を判定する精度を推測できる．表 2，5 の二つの結果より、大量のパケットの中から ASKA.V2 の書き込みを 97%の精度で判定が可能であると推測できる．

表 5 1000 個の Web サイトのパケットを askav2-id で判定した結果
Table 5 The result of judging the packet of 1000 Web site by askav2-id

Web サイトの数	書き込み回数	判定結果 (判定回数)	判定率
1000	4 回	書き込みを判定 (4)	100%

6.4 実験 4

Tor Browser が他の Web サーバと同時に通信している場合に，キャプチャしたパケットの中から任意の掲示板への書き込みの判定ができるか実験を行なった. 今回の実験では 3 種類の別の通信を行いながら掲示板書き込みを行なった．別の通信は YouTube [15] のライブストリーミング再生と YouTube の動画再生，そして，Web サイトの読み込みの 3 種類で実験した．Web サイトの読み込みは，データ量の多い Web サイトを選んで別の通信を行なった．理由として，データ量の多い Web サイトは読み込み時間が長いため，確実に掲示板書き込みと同時の通信を発生させることができるためである．

実験では Tor Browser を 2 ウインドウで起動し，片方のウインドウでは ASKA.V2 に書き込みを行い，もう片方のウインドウでは別の通信を行なった．

結果は表 6 の通りである．実験結果から，この手法であると他の通信をしながら掲示板に書き込みを行なった場合にその書き込みを判定できないことがわかる．

表 6 実験 4 の結果
Table 6 Result of Experiment 4

掲示板以外の通信	判定結果 (判定回数)	書き込み回数
YouTube 動画再生	判定なし (0)	20 回
YouTube ライブ再生	判定なし (0)	20 回
Web サイトの読み込み	判定なし (0)	20 回

7. 考察

本論文の実験では，高精度で掲示板書き込みを判定することができた．課題として，他の掲示板への書き込みを誤判定してしまうことや，掲示板書き込みと同時に他の通信があると特定ができなくなることが挙げられる．本論文の実験では他の掲示板への書き込みを誤判定することはあったが，掲示板への書き込みをしていないパケットを誤判定することはなかった．これより，掲示板書き込みの際に発生する通信が特徴的であることがわかる．考察では，提案システムの課題を解決できるような掲示板書き込みシステムの考察を行う．

まず，他の掲示板への書き込みを誤判定してしまうという課題について考察する．他の掲示板への書き込みを誤判定してしまう理由は，通信パケットのやり取りが全く同じであるためであった．そのため，Fingerprint が全く同一となっていた．実験に使用した掲示板の中で Webforum が唯一他の掲示板とパケットの Fingerprint が同一にならなかった. 他の掲示板と Webforum を比較して大きく異なっている点は入力フォームである. 他の掲示板では，タイトル，名前，コメント，削除キーなどの文字のみをアップロードする入力フォームであったが，Webforum では画像をアップロードできるフォームがあり，この機能の影響でパケットの Fingerprint が他の掲示板と異なっていると推測できる. よって，掲示板の入力フォームに，他の掲示板が利用しないような機能を加えることで特徴づけができるのではないかと考える. また，掲示板を改造することで他の掲示板と被っていた Fingerprint を被らない Fingerprint に変化させることができるという結果も得られた. 実際に掲示板サイトを運用する Web サーバ管理者は自身の掲示板の書き込みをする際のパケットに Fingerprint を持たせ

ることで犯罪に巻き込まれた場合でも対応できる可能性が高くなる．

次に，他の通信パケットが掲示板の通信パケットに割り込むと判定できなくなるという課題について述べる．それは，Fingerprint としているパケットとパケットの間に他の通信のパケットが割り込むことによって，Fingerprint が発見できずに書き込みの判定ができなかった．そこで，同時に他の通信をしている場合でも掲示板書き込みを判定する手法として考えられるものを提案する．これは，Fingerprint としているパケットとパケットの間に割り込んでくる，別の通信のパケットを無視することで掲示板への書き込みの判定を行うという手法である．掲示板書き込みの際に発生するパケットは短い時間で一定のパケットの送受信を行う．そのため，Fingerprint を受け付ける制限時間を設け，その時間内に Fingerprint になるようなパケットを待ち続ける．制限時間内は他の通信パケットを無視し，Fingerprint となるパケットだけを受け付ける．この時，他の通信パケットは Web サイトやデータをロードしているため，出来るだけ長い通信パケット長で通信をしている．しかし，掲示板への書き込みを行うパケットの長さは，出来るだけ長い通信パケット長で通信をしていないことが多いため，この手法で割り込みパケットをフィルタリングできると考えられる．考慮しなければいけない点として，制限時間を設けた場合に他の通信を誤判定してしまう可能性が高くなるため，制限時間は出来るだけ短くする必要がある．

この掲示板書き込み判定プログラムは，掲示板にある特定の書き込みに対して，誰が書き込みを行なったのかを特定することを想定している．条件が揃っている場合，この手法は掲示板書き込みをしたユーザを高精度で特定することができるが，課題はまだまだ多い．また，今回は実際にインターネット上で多くのユーザが利用している掲示板を対象に実験が出来なかったため，今後は実際にインターネット上の他の掲示板に対してこの手法が有効であるか実験する必要がある．ただ，この手法の有効性が高いと考えている．理由として，実際のインターネットにある掲示板は掲示板書き込みの際にやり取りを行うパケットの数が非常に多いため，掲示板書き込みのパケットの Fingerprint を判別することは難しくないと考えられるからである．

8. まとめ

本論文では，Tor を利用して掲示板書き込みを行なったユーザを特定するシステムの提案を目的とした．実際に Web サイトにアクセスし，キャプチャしたパケットの中から特定の掲示板へ書き込みを行なったパケットを判定することが出来た．また，複数の通信が同時に発生している場合に判別が正しく出来ないことを示した．今後の課題として，複数の通信が同時に発生している場合でも正しく判定できるような手法の提案と，インターネット上で多くのユーザに利用されている掲示板での実験が挙げられる．

参考文献

[1] Andriy Panchenko, Lukas Niessen, Andreas Zinnen, Thomas Engel: Website FingerPrinting in Onion Routing Based Anonymization Networks, ACM workshop on Privacy in the electronic society, pp. 103-114, (2011).
[2] Xiaogang Wang, Junzhou Luo, Ming Yang, and Zhen Ling: ”A potential HTTP-based application-level attack against Tor”, Future Generation Computer Systems, pp. 67-77, (2011).
[3] Z.Ling, J.Luo, W.Yu, X.Fu, D.Xuan, and W.Jia: A New Cell-Counting-Based Attack Against Tor, Networking, IEEE/ACM Transactions on, Vol.20, Issue.4, pp.1245-1261,(2012).
[4] Daniel Anderson: ” Splinternet Behind the Great Firewall of China”, queue - Web Security, pp. 1-10, (2012).
[5] Richard Clayton, Steven J. Murdoch, and Robert N. M. Watson: ”Ignoring the Great Firewall of China”, International conference on Privacy Enhancing Technologies, pp. 22-35, (2006).
[6] PhilippWinter, StefanLindskog: How the Great Firewall of China is Blocking Tor，Karlstad University，(2012).
[7] MASHAEL ALSABAH, IAN GOLDBERG: ”Performance and Security Improvements for Tor: A Survey” pp.32:1-pp.32:36, ACM Computing Surveys, Vol. 49, No. 2, Article 32, Publication date: September (2016).
[8] 30% of Tor Web Browser Transactions Found to Be Fraudulent: 入手先 ⟨http://www.infosecurity-magazine.com/news/30-of-tor-web-browser-transactions-found-to-be⟩ (2017.01.18).
[9] Tor Project:Anonymity Online 入手先 ⟨https://www.torproject.org⟩ (2017.1.18).
[10] Tails - Privacy for anyone anywhere 入手先 ⟨https://tails.boum.org⟩ (2017.1.18).
[11] Welcome to Tor Metrics 入手先 ⟨https://metrics.torproject.org⟩ (2017.1.18).
[12] Wilde，T:Great Firewall Tor Probing Circa 09 DEC 2011，入手先 ⟨https://gist.github.com/twilde/da3c7a9af01d74cd7de7⟩ (2017.1.18).
[13] KENT WEB - CGI/Perl フリーソフト 入手先 ⟨http://www.kent-web.com⟩ (2017.1.18).
[14] 掲示板 CGI ThreadPlus（スレッドプラス）入手先 ⟨http://59bbs.org/threadplus/⟩ (2017.1.18).
[15] You Tube 入手先 ⟨https://www.youtube.com/⟩ (2017.1.18).
[16] Yahoo! 掲示板 入手先 ⟨http://textream.yahoo.co.jp/⟩ (2017.1.18).
[17] サイバー自警団 - 必ず捕まる！ Tor 犯罪予告のパソコン遠隔操作事件まとめ 入手先 ⟨http://koshikien.co/tor-remote⟩ (2017.1.23).