情報処理学会論文誌 Vol.62 No.9 1523–1535 (Sep. 2021)

推薦論文

偽ショッピングサイトによる攻撃手法の実態解明

小寺 博和[1,†1,a)] 小出 駿[1,2] 千葉 大紀[1] 青木 一史[1] 秋山 満昭[1]

受付日 2020年11月30日, 採録日 2021年6月7日

概要：個人情報や金銭を狙ったフィッシング詐欺が急増している．攻撃者はフィッシングサイトにアカウント情報や個人情報をユーザに入力させることでこれらの情報を窃取する．フィッシング詐欺の 1 つとして，正規のショッピングサイトを模倣した偽のショッピングサイトによる金銭被害が確認されている．ユーザは注文成立後に購入代金を支払うが，注文した商品は発送されず，結果として購入代金を騙し取られる．偽ショッピングサイトには，不自然な日本語表記を用いることや実在しない企業名を名乗る等の特徴があるが，マルウェア感染を起点とする攻撃で用いられる脆弱性の悪用等がないため，アンチウイルスエンジンによる検知が困難である．被害防止の対策検討のためには，偽ショッピングサイトによる攻撃の実態を明らかにすることが必要である．本研究では，偽ショッピングサイトが持つ解析回避機能や，攻撃の実態に関する調査結果を報告する．URL ブラックリストに掲載されたドメイン名や，偽ショッピングサイトのページタイトルを検索エンジンで検索することで，偽ショッピングサイトの疑いのある URL を収集した．偽ショッピングサイトに対して条件を変更した複数の HTTP リクエストを送信し，その応答を分析することで，99.8%が検索エンジンサイト経由のアクセスの場合にのみ偽ショッピングサイトに到達できる解析回避機能を有することを明らかにした．また，改ざんの修正が確認されたウェブサイトのうち 82.9%が 20 日以内に修正されたことが確認された．

キーワード：フィッシング，偽ショッピングサイト，解析回避機能

Understanding Attacks with Fake Shopping Websites

Hirokazu Kodera[1,†1,a)] Takashi Koide[1,2] Daiki Chiba[1] Kazufumi Aoki[1] Mitsuaki Akiyama[1]

Received: November 30, 2020, Accepted: June 7, 2021

Abstract: The number of phishing scams targeting personal information and financial gain has been increasing. Attackers direct victims to a phishing site they prepare and steal the account and personal information of victims. One type of phishing scams is a fake shopping site that mimics a legitimate shopping site. Fake shopping sites trick users into ordering goods and paying for them, although of course the ordered goods are never shipped and the money is never returned. Commonly known fake shopping sites have characteristics of being described in unnatural Japanese or with non-existent company names. However, since there is no trace of malware infection or vulnerability exploitation to be detected robustly, such fake shopping sites cannot be detected by common anti-virus engines. In order to prevent damage caused by such fake shopping sites, we first need to understand the actual situation of attacks by these sites. In this paper, we report the results of investigating the analytical evasion function of fake shopping sites and the actual situation of the attacks. Specifically, we first searched for domain names on the URL blacklist and page titles of fake shopping sites with search engines to efficiently collect the fake shopping site candidates. Next, we sent multiple HTTP requests to the fake shopping site with various conditions and analyzed the responses to reveal that 99.8% of fake shopping sites have the analytical evasion function to redirect only when accessing from search engine results. Also, we found that 82.9% of websites which have been confirmed to be fixed compromising have been fixed within 20 days.

Keywords: phishing, fake shopping websites, analytical evasion function

1. はじめに

フィッシング詐欺は個人情報や金銭をユーザから窃取す

[1] 日本電信電話株式会社
NTT, Musashino, Tokyo 180–8585, Japan
[2] 横浜国立大学
Yokohama National University, Yokohama, Kanagawa 240–8501, Japan
[†1] 現在，NTT セキュリティ・ジャパン株式会社
Presently with NTT Security Japan, Tokyo 101–0021, Japan
[a)] hirokazu.kodera@global.ntt

本論文の内容は 2020 年 3 月の第 88 回 CSEC 研究会にて報告され，同研究会主査により情報処理学会論文誌ジャーナルへの掲載が推薦された論文である．

ることを目的とした攻撃である．Anti-Phishing Working Group（APWG）によると 2019 年第 4 四半期には約 16.2 万件のフィッシングサイトが観測され，2020 年第 1 四半期にも約 16.5 万件のフィッシングサイトが観測されており，フィッシングによる被害は多い状態が継続している [1]．

フィッシング詐欺の 1 つとして，正規のショッピングサイトを模倣した偽のショッピングサイトにユーザを誘導し，ユーザから購入代金やクレジットカード情報を騙し取ることや偽のブランド商品を購入させることを目的とした詐欺行為が確認されている．このような詐欺を目的としたウェブサイトの一種として，日本人をターゲットに購入代金を窃取することを目的した偽のショッピングサイトが多く確認されている．Japan Cybercrime Control Center（JC3）によると，2017 年に 19,834 件の偽ショッピングサイト URL が確認されている [2]．JC3 は偽ショッピングサイトを “Fake Store” と定義し，その攻撃手法には 3 つの特徴があることを明らかにした．犯罪者グループを 3 つの特徴を用いて分類した結果，2017 年に観測された 19,834 件の “Fake Store” による攻撃には 6 つの犯罪者グループが存在したことを明らかにした．

偽ショッピングサイトはユーザがサイトにアクセスする経緯が従来のフィッシングサイトとは異なる．従来のフィッシングサイトは，ユーザが受信したメールに記載された URL をクリックすることで検索エンジン等を経由せずに直接フィッシングサイトにアクセスする．一方で，今回着目する偽ショッピングサイトは，ユーザが商品名や商品型番等を検索エンジンで検索し，検索結果に掲載された URL をクリックすることで偽ショッピングサイトにアクセスする．

Drive-by Download 攻撃を実行するウェブサイトやフィッシングサイトには，アクセス元 IP アドレスや HTTP リクエストヘッダのパラメータをもとにウェブサイトにアクセスしたユーザの環境を判別する手法であるクローキングを用いることで解析を回避することが知られている [3], [4], [5]．これらの悪性なウェブサイトと同様に偽ショッピングサイトもクローキングによる解析回避機能を有する可能性が考えられる．しかしながら，既存研究 [2] では改ざんされた正規のウェブサイト（以降，踏み台サイトとする）を経由して偽ショッピングサイトへ到達することが明らかにされているが，偽ショッピングサイトが有する解析回避機能については明らかにされていない．偽ショッピングサイトが有する解析回避機能を明らかにすることで，解析者はより効率的に調査することが可能になる．

本研究では，偽ショッピングサイトの URL を収集し，偽ショッピングサイトが有する解析回避機能や，偽ショッピングサイトの時間変化を分析することで攻撃の実態調査を実施した．具体的には，まず URL ブラックリストに掲載された URL のドメイン名を対象に，検索エンジンで踏み台サイトとなりうる正規のウェブサイトのドメイン名配下の URL を収集する．さらに，偽ショッピングサイトのページタイトル文字列を検索エンジンで検索し，踏み台サイトとなりうる URL を収集した．次に，収集した URL に対して異なる HTTP リクエストヘッダを設定した複数の HTTP リクエストを送信した際の応答を観測することで踏み台サイトと偽ショッピングサイトが持つ解析回避機能を分析した．収集した URL に対して定期的にアクセスすることで，到達する偽ショッピングサイトの変化や踏み台サイトの改ざんが修正される期間を調査した．最後に，偽ショッピングサイトのドメイン名を対象に Passive DNS のデータを用いて攻撃規模を推定した．

本論文の主な貢献を以下に示す．

- URL ブラックリストに掲載された URL のドメイン名（78,154 件）と偽ショッピングサイトで実際に使われたページタイトル文字列（6,548 件）を検索エンジンで検索した結果から偽ショッピングサイトの踏み台サイトのドメイン名を 2,996 件収集した．また，ドメイン名配下の URL にアクセスした際にリダイレクトされる偽ショッピングサイトのドメイン名を 9,958 件収集した．
- 踏み台サイトのうち 99.8%が検索エンジンサイト経由でアクセスした場合のみ偽ショッピングサイトへリダイレクトさせる解析回避機能を有することを明らかにした．また，94.6%が日本で利用されることが多い Google，Yahoo! JAPAN 等の検索エンジンサイトを経由した場合のみ偽ショッピングサイトへリダイレクトさせることを明らかにした．
- 踏み台サイトに継続的（17 日間）にアクセスし，到達する偽ショッピングサイトのドメイン名が変化した踏み台サイトが 13.7%存在したことを明らかにした．また，改ざんの修正が確認された 579 件の踏み台サイトのうち 82.9%が検索エンジン掲載から 20 日以内に改ざんが修正され，その平均は 16.7 日であったことを明らかにした．
- Passive DNS のデータを用いて偽ショッピングサイトによる攻撃キャンペーンの規模を調査したところ，1 ドメイン名あたり平均 64.9 回，最大 10,921 回の名前解決があり，一定数のユーザが偽ショッピングサイトに到達している可能性があることを明らかにした．

本論文の構成は以下のとおりである．2 章で関連研究を示す．3 章で偽ショッピングサイトによる攻撃手法の全体像を述べる．4 章で検索エンジンを用いた偽ショッピングサイト URL の収集手法と，偽ショッピングサイトが持つ解析回避機能の調査手法に関して述べ，5 章で，調査手法の具体例を述べる．6 章で偽ショッピングサイトが持つ解析回避機能の調査結果，偽ショッピングサイトの時系列分析結果，Passive DNS による攻撃規模の推定結果について

述べる．7 章で本研究における制約，今後の課題について述べ，8 章で本論文のまとめを行う．

2. 関連研究

偽ショッピングサイトに関する研究として，偽ブランド商品の EC サイトを検出する手法に関する研究が報告されている [6], [7], [8], [9], [10], [11]．Wadleigh ら [6] は，WHOIS 情報，価格設定，ウェブサイトコンテンツをもとにした偽ブランド商品を販売するウェブサイトの検出手法を提案した．2014 年 1 月から 8 月に収集された検索エンジンでの検索結果のうち 32%が偽ブランド商品を販売するウェブサイトであることを明らかにした．Carpineto ら [7] は，検索エンジンからアクセス可能なウェブサイトのコンテンツから正規か偽のショッピングサイトであるかの判定する手法と，偽造のリスクを分析するための “Counterfeiting Charts” を生成する手法を提案した．39 種の靴のブランドを対象に検索エンジンで検索した結果，3,601 件のウェブサイトから 209 件の偽ショッピングサイトが提案手法により検出された．

フィッシングサイトに関する研究として，フィッシングを行う攻撃者やサーバ環境等の攻撃用インフラに関する研究が報告されている．Han ら [12] は，著者らが設置したハニーポットを攻撃者がフィッシングサイトに改ざんする様子を観測することで，攻撃者の行動分析や被害者数の推定を行った．被害者数を調査した結果，2,438 人のユーザがフィッシングサイトにアクセスし，そのうち 215 人（9%）が認証情報を送信していたことを明らかにした．Oest ら [4] は，フィッシングサイトの構築ツールであるフィッシングキットを解析し，アクセスブロックの対象とされやすい IP アドレスの国別の傾向や，組織別の傾向を調査した．検索エンジンクローラ，セキュリティベンダ，フィッシングの対象となるブランドの企業（Paypal 社，Apple 社等）からのアクセスがブロックされる傾向にあることを示した．小寺ら [5] は，インターネット上に実在するフィッシングサイトを対象にアクセス妨害機能を持つフィッシングサイトを調査した．HTTP リクエストのヘッダ（User-Agent, Referer）を設定してアクセスしてその応答を観測した結果，アクセス妨害機能を持つフィッシングサイトが 10.4%存在することを明らかにした．

偽ショッピングサイトを検出する手法や，一般的なフィッシングサイトに関する攻撃者や攻撃用インフラの分析はされているが，偽ショッピングサイトに関しては解析回避機能のような攻撃インフラに関する調査は報告されていない．

3. 偽ショッピングサイトの概要

本章では偽ショッピングサイトによる攻撃手法の概要について述べる．**図 1** にユーザが検索エンジンを経由して偽ショッピングサイトに誘導されるまでの流れを示す．

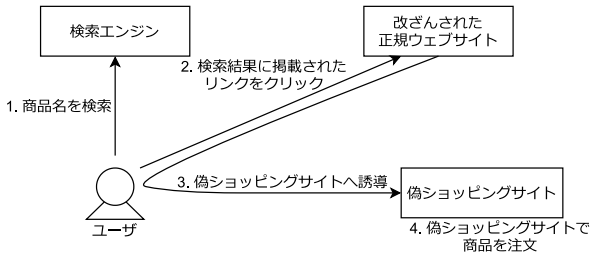


図 1 偽ショッピングサイトの全体概要
Fig. 1 Overview of the attack of a fake shopping site.

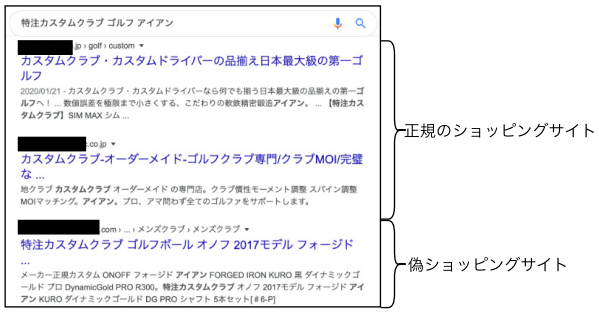


図 2 検索エンジンの検索結果
Fig. 2 A search result by search engines.

まず，ユーザが商品名や商品型番を検索エンジンで検索する．その検索結果には**図 2** のような正規のショッピングサイトへのリンク以外に偽ショッピングサイトへ誘導するためのリンクが含まれる場合がある．検索結果に掲載される URL は偽ショッピングサイトのものではなく，偽ショッピングサイトにリダイレクトさせるために改ざんされた正規のウェブサイト（踏み台サイト）であることが知られている [2]．次に，ユーザが検索結果に掲載された踏み台サイトへのリンクをクリックすると，踏み台サイトを経由して偽ショッピングサイトへ誘導される．偽ショッピングサイトには正規のショッピングサイトと同様に商品の注文機能があり，利用ガイド等のユーザ向けのマニュアルも掲載されていることから見た目の区別がつきにくい．最後に，ユーザが偽ショッピングサイトで商品を注文し，購入代金を支払うことで攻撃が成立し，攻撃者に購入代金を窃取される．

4. 解析回避機能の調査手法

偽ショッピングサイトが有する解析回避機能を調査するために，URL ブラックリストに掲載された URL と偽ショッピングサイトのページタイトル文字列を用いて，検索エンジンで偽ショッピングサイトの URL を収集した．本章では，調査手法の詳細を述べる．

4.1 全体概要

調査の全体概要を**図 3** に示す．

まず，改ざんされやすい正規ウェブサイトドメイン名を収集するために，URL ブラックリストに掲載された URL

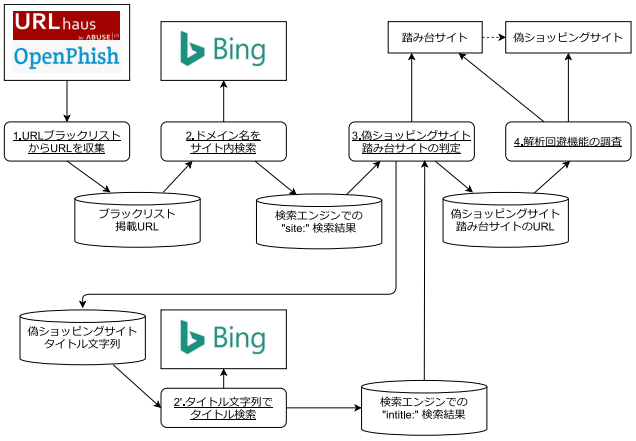


図 3 偽ショッピングサイト収集方法

Fig. 3 The collection method of fake shopping sites.

のドメイン名を収集した．次に，検索エンジンで入手したドメイン名をサイト内検索することで踏み台サイトとなりうる当該ドメイン名配下の URL を収集した．次に，収集した URL に対してアクセスし，偽ショッピングサイトへリダイレクトされるかを確認することで当該 URL が偽ショッピングサイトの踏み台サイトであるかを判定した．最後に，特定の HTTP リクエストヘッダを設定した複数の HTTP リクエストを送信し，その応答を確認することで偽ショッピングサイトと踏み台サイトが有する解析回避機能を調査した．

さらに，調査規模を拡大するために，偽ショッピングサイトと判断したページタイトル文字列を検索エンジンで検索することで，踏み台サイトとなりうる URL を収集した．この方法で収集した URL に対しても同様に偽ショッピングサイトへリダイレクトされるかの判定と，解析回避機能の調査を実施した．

4.2 URL ブラックリストから URL の収集

3 章で述べたように，ユーザは検索エンジンの検索結果に掲載された踏み台サイトを経由して偽ショッピングサイトに到達する．また，正規のウェブサイトが改ざんされて踏み台サイトとされていることが知られている．したがって，検索エンジンで改ざんされる可能性が高いウェブサイトのドメイン名に対してあるドメイン名配下の URL を検索する機能であるサイト内検索（`site:` 検索コマンド）を実施することで，踏み台サイトとなりうる URL を効率的に収集できる．

改ざんされやすいウェブサイトは，マルウェアの配布サイトや C&C サーバとして悪用されることが知られている [13]．フィッシングサイトも同様に正規のウェブサイトの改ざんによるケースが存在することが知られている [14]．本研究では，マルウェアやフィッシングの URL ブラックリストに掲載された URL を用いて後続の調査を実施した．マルウェアに関する URL ブラックリストとして，URLhaus [15]，VX Vault [16]，Malware Domain List [17] が知られている．本研究ではその中でも更新頻度が高く，悪用された URL を多く入手可能な URLhaus を用いた．フィッシングに関するブラックリストとして，OpenPhish [18]，PhishTank [19] が知られている．本研究では悪用されたブランドの特定がされており，フィッシングサイトである確度がより高い URL を配布していると考えられる OpenPhish を用いた．

図 4 検索エンジンの検索結果の各部説明

Fig. 4 Description of each part of a search result by search engines.

4.3 検索エンジンを用いた踏み台サイト URL の収集

4.2 節で入手した URL のドメイン名を対象に，検索エンジンの `site:` 検索コマンドでそれぞれのドメイン名配下の踏み台サイトの候補となる URL を収集した．さらに，調査規模を拡大するために，後続の 4.4 節で偽ショッピングサイトと判定したウェブページ HTML の `title` タグを，指定したキーワードがウェブサイトのタイトルに含まれる URL を検索する機能である検索エンジンの `intitle:` 検索コマンドで踏み台サイトの候補となる URL を収集した．本研究では，Bing Web Search API を検索エンジンとして用いた．

(1) URL ブラックリストのドメイン名を用いた URL 収集

URL ブラックリストに掲載された URL のドメイン名に対して検索エンジンの `site:` 検索コマンドを実行し，当該ドメイン名配下の URL を収集した．検索エンジンによる検索結果は，図 4 のようにタイトル，URL，スニペットから主に構成される．今回対象とする偽ショッピングサイトは日本人を攻撃対象とするため，タイトルやスニペットに日本語で記述された商品説明が含まれる．そこで，本研究ではタイトルに日本語が含まれる検索結果を踏み台サイトの疑いのある URL として収集した．

(2) 偽ショッピングサイトの `title` タグ検索による URL 収集

偽ショッピングサイトの `title` タグの文字列には，商品名だけではなく，「激安」「お得」のようなユーザを引き付けるための単語も含まれている．したがって，偽ショッピングサイトの `title` タグの文字列を検索エンジンで検索することで，商品名だけで検索するよりも効率的に踏み台サイトの疑いのある URL が収集できると考えられる．そこで，本研究では偽ショッピングサイトの `title` タグの文字列を検索エンジンで `intitle:` 検索を実施し，得られ

た検索結果の URL を踏み台サイトの疑いのある URL として収集した．

この方法を実施した場合，当該商品を取り扱う正規のショッピングサイトも調査対象に含まれてしまうという問題が考えられる．偽ショッピングサイトの特徴として，図 1 のように検索エンジンに掲載された URL とは異なるドメイン名にリダイレクトされることが知られている．また，正規のショッピングサイトにはそのような挙動を示すものは少ないと考えられる．そこで，検索エンジンに掲載されたドメイン名とは異なるドメイン名の URL にリダイレクトされたことが確認されたもののみを調査対象とするヒューリスティックな手法をとった．

4.4 偽ショッピングサイトの判定

4.3 節で収集した踏み台サイトの疑いのある URL にアクセスし，踏み台サイトであるかを判定した．さらに，踏み台サイトからリダイレクトされる偽ショッピングサイトの URL を収集した．

踏み台サイトから偽ショッピングサイトへのリダイレクトには，HTTP ステータスコード 300 番台によるものだけではなく，JavaScript や HTML の meta タグを用いる場合も想定される．HTTP ステータスコードによるリダイレクトは curl や wget 等でも実行されるが，JavaScript や HTML の meta タグによるリダイレクトを発生させることができない．そこで，Web ブラウザ自動化ツールである Selenium [20] を用いて踏み台サイトにアクセスすることで，JavaScript や HTML の meta タグによるリダイレクトを観測した．

検索エンジンサイト経由でユーザが踏み台サイトへアクセスした場合にのみ偽ショッピングサイトにユーザを誘導する挙動を有する事例があることが知られている [2]．そこで，本研究では検索エンジンサイトの URL を Referer ヘッダに設定した状態で踏み台サイトにアクセスすることで偽ショッピングサイトへのリダイレクトを発生させる方式をとった．特定の検索エンジンを経由した場合のみリダイレクトが発生することも想定されるため，本研究では**表 1** の 4 種類の Referer ヘッダをそれぞれ設定して踏み台サイトに計 4 回アクセスした．

Selenium の仕様として，HTTP リクエストに任意の Referer ヘッダを設定することができない．そこで，まず Referer ヘッダに設定する検索エンジンサイトにアクセスし，検索エンジンサイト上で Listing 1 の JavaScript コードを実行することで，Referer ヘッダを設定した状態で踏み台サイトにアクセスした．また，Referrer-Policy のデフォルト値である `no-referrer-when-downgrade` により，踏み台サイトが HTTPS ではない場合，Referer ヘッダが設定されない．そこで，Listing 1 のコードの実行前に Listing 2 の JavaScript コードを実行することで HTTPS ではない踏み台サイトにアクセスする場合でも Referer ヘッダを設定できるようにした．

表 1 Selenium に設定した Referer ヘッダ
Table 1 Referer header set to HTTP request header when using Selenium.

No.	Referer
1	https://www.google.co.jp/
2	https://www.bing.com/
3	https://www.yahoo.co.jp/
4	https://www.yahoo.com/

Listing 1 Referer ヘッダを付与するための JavaScript コード

```
1 window.location.href = "<Compromised Website URL>";
```

Listing 2 Referer ポリシー制御のための JavaScript コード

```
1 var head = document.getElementsByTagName("head")[0];
2 var child = document.createElement("meta");
3 child.setAttribute("name","referrer");
4 child.setAttribute("content","origin");
5 head.appendChild(child);
```

Selenium で表 1 の Referer ヘッダを設定した状態で踏み台サイトへアクセスし，異なるドメイン名の URL へリダイレクトされた場合はリダイレクト先のコンテンツを手動で確認した．リダイレクト先が偽ショッピングサイトであった場合は，それぞれ踏み台サイトと偽ショッピングサイトの URL であると判定した．

4.5 解析回避機能の調査

4.4 節で踏み台サイトと判定した URL に対して異なる HTTP ヘッダを設定した HTTP リクエストを複数回送信することで，踏み台サイトが有する解析回避機能を調査した．踏み台サイトに対して，**表 2** に示す User-Agent と Referer のパターンを HTTP ヘッダに設定して踏み台サイトにアクセスすることで，踏み台サイトが有するアクセス回避機能を調査した．

No.1–6 でそれぞれアクセスすることで，検索エンジンボットによるアクセスの場合にのみ検索エンジンにインデックスさせるためのコンテンツを応答するかの判定ができる．No.6–13 でそれぞれアクセスすることで，検索エンジン経由でユーザがアクセスした場合にのみ偽ショッピングサイトへリダイレクトされるかの判定ができる．調査対象とする検索エンジンサイトとして，国内で利用されることが多い Google，Yahoo! JAPAN，Bing に加え，海外で利用されることが多い Yandex，Yahoo!，Baidu をそれぞれ用いた．Google については海外で利用される検索エンジンサイトとして，TLD が．uk である google.co.uk も調査対象として用いた．これ以降，google.co.jp，google.co.uk をそれぞれ Referer としてアクセスした場合の表記を Google

表 2 解析回避機能調査のための HTTP リクエストヘッダ
Table 2 HTTP request headers for analyzing evasion function.

No.	User-Agent	Referer
1	Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)	なし
2	Mozilla/5.0 (compatible; bingbot/2.0; +http://www.bing.com/bingbot.htm)	なし
3	Mozilla/5.0 (compatible; YandexBot/3.0; +http://yandex.com/bots)	なし
4	Mozilla/5.0 (compatible; Yahoo! Slurp; http://help.yahoo.com/help/us/ysearch/slurp)	なし
5	Mozilla/5.0 (compatible; Baiduspider/2.0; +http://www.baidu.com/search/spider.html)	なし
6	Mozilla/5.0 (Windows NT 10.0; WOW64; Trident/7.0; rv:11.0) like Gecko	なし
7	Mozilla/5.0 (Windows NT 10.0; WOW64; Trident/7.0; rv:11.0) like Gecko	https://www.google.co.jp/
8	Mozilla/5.0 (Windows NT 10.0; WOW64; Trident/7.0; rv:11.0) like Gecko	https://www.google.co.uk/
9	Mozilla/5.0 (Windows NT 10.0; WOW64; Trident/7.0; rv:11.0) like Gecko	https://www.bing.com/
10	Mozilla/5.0 (Windows NT 10.0; WOW64; Trident/7.0; rv:11.0) like Gecko	https://yandex.com/
11	Mozilla/5.0 (Windows NT 10.0; WOW64; Trident/7.0; rv:11.0) like Gecko	https://www.yahoo.co.jp/
12	Mozilla/5.0 (Windows NT 10.0; WOW64; Trident/7.0; rv:11.0) like Gecko	https://www.yahoo.com/
13	Mozilla/5.0 (Windows NT 10.0; WOW64; Trident/7.0; rv:11.0) like Gecko	https://www.baidu.com/

（JP），Google（UK）とする

5. 調査手法の具体例

本章では，4 章の手法による調査の具体例を述べる．例として，URL ブラックリストで入手した URL が http://mal.example.com/malicous/content.exe であった場合の調査手法を説明する．

まず，この URL のドメイン名である `mal.example.com` を検索エンジンの `site:` 検索コマンドで検索を実行し，当該ドメイン名配下の URL を収集する．収集した URL のうち，検索結果のタイトルに日本語が含まれる URL を踏み台サイトの疑いのある URL（例：http://mal.example.com/redirect.php）として後続の処理を実行する．

次に，収集した URL が踏み台サイトの URL であるかを調査するために，表 1 の Referer を設定して Selenium でアクセスする．Selenium でのアクセス時に異なるドメイン名の URL（例：http://fakeshop.example.net/productid-2001.php）へリダイレクトされた場合は，到達先のコンテンツを取得し，手動でコンテンツを確認して偽ショッピングサイトであるかを判定する．収集した URL にアクセスしたときに，偽の懸賞ページ等に誘導するソーシャルエンジニアリング攻撃サイト [21] や他の正規のウェブサイトに到達する場合も確認された．まず，これらのサイトではないことを確認するために，商品画像や価格等のショッピングサイトに特有の情報が掲載されているかを確認した．次に，正規のショッピングサイトではないことを確認するために，価格の 3 割を超える値引きの有無や，当該ドメイン名が Alexa Top Sites [22] の上位 10,000 件に含まれるかを確認した．これらで判別できない場合は，検索エンジンで当該ドメイン名を検索し，その検索結果に含まれる第三者による情報や当該ドメイン名配下の別の URL にアクセスした結果をもとに判断した．

表 3 解析回避機能調査の実施例
Table 3 Example of analyzing evasion function.

No.	User-Agent	Referer	到達先
1	GoogleBot	なし	**検索インデックスページ**
2	BingBot	なし	**検索インデックスページ**
3	YandexBot	なし	正規ウェブサイト
4	YahooBot	なし	正規ウェブサイト
5	BaiduBot	なし	正規ウェブサイト
6	Windows10 IE11	なし	正規ウェブサイト
7	Windows10 IE11	Google（JP）	**偽ショッピングサイト**
8	Windows10 IE11	Google（UK）	正規ウェブサイト
9	Windows10 IE11	Bing	**偽ショッピングサイト**
10	Windows10 IE11	Yandex	正規ウェブサイト
11	Windows10 IE11	Yahoo!Japan	**偽ショッピングサイト**
12	Windows10 IE11	Yahoo!	正規ウェブサイト
13	Windows10 IE11	Baidu	正規ウェブサイト

最後に，踏み台サイトの解析回避機能を調査するために，踏み台サイト URL に表 2 の HTTP ヘッダを設定してアクセスする．**表 3** にアクセスした結果の例を示す．簡単のために User-Agent と Referer には通称を記述する．表 3 は，GoogleBot，BingBot でのアクセス時に検索インデックスに掲載されるための商品説明ページを応答し，Google（JP），Bing，Yahoo! JAPAN を経由したアクセス時に偽ショッピングサイトへリダイレクトさせることを示している．この結果より，この踏み台サイトは検索エンジンボットや検索エンジンからのアクセスの場合のみ踏み台サイトとしての挙動を示す解析回避機能を持つことが分かる．

さらに，偽ショッピングサイトの HTML の `title` タグに用いられた文字列を対象に検索エンジンで `intitle:` 検索コマンドを実行し，同じタイトルを持つ URL を収集する．この方法で収集した URL に対しても同様に，踏み台サイト URL であるかの調査を実施し，踏み台サイトであると判明した場合は解析回避機能の調査を実施する．

6. 調査結果

本章では，4 章の手法による偽ショッピングサイトの収集結果と解析回避機能の分析結果について述べる．次に，偽ショッピングサイトの構築ツールの分析結果について述べる．次に，偽ショッピングサイトに定期的にアクセスしたときの時系列変化の分析結果について述べる．最後に，Passive DNS のデータをもとにした偽ショッピングサイトの攻撃キャンペーン規模の推定結果を述べる．

6.1 偽ショッピングサイト URL の収集結果

4.3 節の手法による収集結果を述べる．踏み台サイト収集に用いた URL ブラックリストに掲載された URL と偽ショッピングサイトのタイトル文字列と，収集した踏み台サイトから到達した偽ショッピングサイトの結果を説明する．

(1) ブラックリスト掲載 URL による踏み台サイト候補のドメイン名収集

URLhaus，OpenPhish に 2020/5/11–2020/9/11 に掲載された URL をもとに収集した偽ショッピングサイトの収集結果を**表 4** に示す．URL ブラックリストとして用いた OpenPhish と URLhaus から合計 256,574 件の URL を取得した．そのうち，ユニークなドメイン名を 78,154 件を Bing Web Search API による `site:` 検索コマンドを実施した．`site:` 検索コマンドによる検索結果のうち，タイトルに日本語が含まれていたものは 5,174 件あった．

(2) 偽ショッピングサイト掲載名による踏み台サイト候補のドメイン名収集

偽ショッピングサイトの `title` タグの文字列をもとに収集した踏み台サイトの可能性のあるドメイン名の収集結果を表 4 に示す．6,548 件の偽ショッピングサイトのタイトル文字列を検索エンジンで `intitle:` 検索コマンドを実施し，9,292 件のドメイン名を収集した．

(1)，(2) で収集したドメイン名 14,466 件に対して 4.4 節の手法でアクセスし，偽ショッピングサイトの踏み台サイトとして動作していると確認されたドメイン名の数は 2,996 件だった．踏み台サイトの可能性のあるドメイン名に対して 20.7%と少なくなった理由の 1 つとして，すでに改ざんが修正されたことや閉鎖されたウェブサイトが多く存在したことが考えられる．

(1)，(2) で収集した踏み台サイトのドメイン名配下の URL にアクセスした結果，9,958 件のドメイン名の偽ショッピングサイトに到達した．踏み台サイトが検索エンジンボットに応答するコンテンツには同一ドメイン名配下 URL へのリンクが掲載されている場合があり，これらの URL にもアクセスした．同一ドメイン名の踏み台サイトから複数の偽ショッピングサイトへリダイレクトされるということが確認され，最も多いもので 144 のドメイン名の偽ショッピングサイトにリダイレクトされることが確認された．また，異なる踏み台サイトから共通の偽ショッピングサイトにリダイレクトされる場合も確認され，最も多いもので 71 のドメイン名の踏み台サイトからリダイレクトされる偽ショッピングサイトが確認された．

表 4 偽ショッピングサイトの収集結果
Table 4 Collection result of fake shopping sites.

		件数
サイト内検索	URL ブラックリスト掲載 URL 数	256,574
	URLhaus	104,496
	OpenPhish	152,078
	サイト内検索実施ドメイン名数	78,154
	検索結果タイトルに日本語を含むドメイン名数	5,174
タイトル検索	タイトル検索実施数	6,548
	収集したドメイン名数	9,292
踏み台サイト候補のドメイン名数		14,466
偽ショッピングサイトの踏み台サイトドメイン名数		2,996
偽ショッピングサイトドメイン名数		9,958

表 5 踏み台サイトドメイン名の TLD
Table 5 Top-level domains of domain name of landing sites.

No.	TLD	件数
1	.com	1,638
2	.org	122
3	.ru	105
4	.net	100
5	.in	99
6	.br	84
7	.uk	54
8	.de	37
9	.vn	34
10	.au	33
11	.jp	28
-	その他（113 種）	662

踏み台サイトと偽ショッピングサイトのドメイン名の TLD の集計結果をそれぞれ示す．攻撃に用いられる TLD の分析結果は，7.3 節に示す偽ショッピングサイトが正規ショッピングサイトの画像ファイルを読み込む性質を利用した偽ショッピングサイトのテイクダウン等の対処を早期に実施するために有用な情報である．踏み台サイトとされたウェブサイトのドメイン名の TLD の割合を**表 5** に示す．.com の TLD を持つドメイン名が最も多く 1,638 件確認され，合計 124 種の TLD を持つドメイン名が確認された．踏み台サイトには．jp の TLD を持つドメイン名も 28 件確認された．偽ショッピングサイトのドメイン名の TLD の割合を**表 6** に示す．.asia，.xyz，.icu の順に多く，TLD の種類も 16 種と踏み台サイトと比較して少ない．

6.2 偽ショッピングサイトの解析回避機能

偽ショッピングサイトと踏み台サイトが有する解析回避機能の分析結果について述べる．

表 6 偽ショッピングサイトドメイン名の TLD
Table 6 Top-level domains of domain name of fake shopping sites.

No.	TLD	件数
1	.asia	2,932
2	.xyz	2,752
3	.icu	1,037
4	.club	760
5	.top	607
6	.fun	487
7	.site	391
8	.online	365
9	.shop	310
10	.wang	83
-	その他（6 種）	234

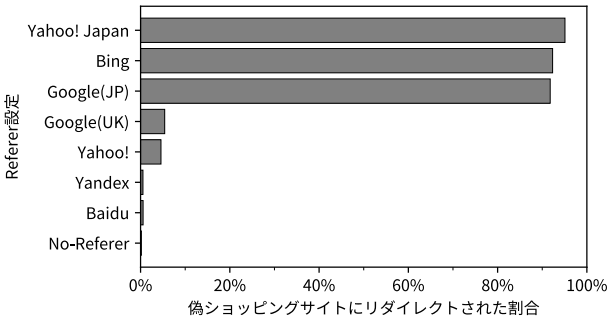


図 5 解析回避機能（Referer 設定によるアクセス）の調査結果
Fig. 5 Analysis result of the evasion function (HTTP Referer header).

（1）踏み台サイトの解析回避機能

表 2 の HTTP リクエストヘッダを用いて踏み台サイトにアクセスし，解析回避機能を持つ踏み台サイトがどの程度存在したかを述べる．**図 5** に検索エンジンサイトの URL を Referer に設定してアクセスした場合の調査結果を示す．Referer に検索エンジンに設定せずにアクセスした場合に偽ショッピングサイトへリダイレクトさせる踏み台サイトは 2,996 件中 5 件のみだった．残りの 2,991 件の踏み台サイトは表 2 の 7–13 の検索エンジンの URL を Referer に設定しない場合は偽ショッピングサイトにリダイレクトされなかった．その中でも，国内で利用されることが多い Google（JP），Bing，Yahoo! JAPAN のうちいずれかからのアクセスでないと偽ショッピングサイトへリダイレクトされない踏み台サイトが 94.6%確認された．偽ショッピングサイトにリダイレクトされない場合は，踏み台サイトが持つ正規のコンテンツやエラーページが応答された．

このようになった結果の理由として，2 つの理由が考えられる．まず，今回の調査対象である偽ショッピングサイトの攻撃対象が日本のユーザであるためである．国内での利用が多い Google（JP），Bing，Yahoo! JAPAN 経由のアクセスのみを偽ショッピングサイトへリダイレクトすることで攻撃者の目的が達成される．次に，国外のユーザが

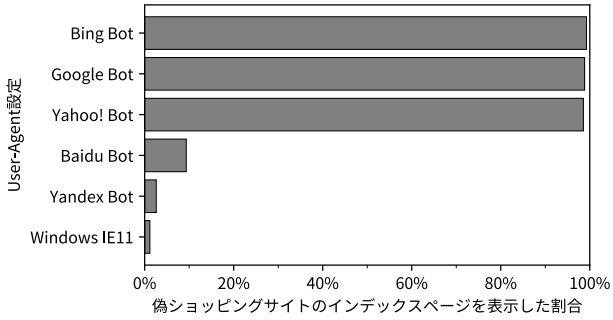


図 6 解析回避機能（User-Agent 設定によるアクセス）の調査結果
Fig. 6 Analysis result of the evasion function (HTTP User-Agent header).

利用する検索エンジンからのアクセスや直接のアクセスによる解析回避のためである．国外のセキュリティ事業者やウェブサイト管理者がアクセスした場合に，改ざんされる前の元のウェブサイトを表示することで踏み台サイトとなっていることを気づくことができない．このようにすることで踏み台サイトの修正を遅らせることができる．

図 6 に検索エンジンボットの User-Agent を設定してアクセスした場合の調査結果を示す．表 2 の 1–5 の検索エンジンボットの User-Agent を設定した場合でも 2,996 件中 429 件は検索エンジンにインデックスさせるための商品紹介を記載したページの応答が確認できなかったため，当該ページを応答した 2,567 件について述べる．User-Agent に Internet Explorer11 を設定してアクセスした場合に検索エンジンにインデックスさせるための商品紹介を記載したページを応答する踏み台サイトは 30 件のみだった．残りの 2,537 件の踏み台サイトは表 2 の 1–5 のいずれかの検索エンジンボットの User-Agent を設定することで商品紹介を記載したページが表示された．また，検索エンジンボットの回避機能が有効に動作しているかを確認するために，解析回避機能が異なるいくつかの踏み台サイトのドメイン名を Yandex で `site:` コマンドによる手動検索を実施した．YandexBot に対する解析回避のない踏み台サイトは Yandex の検索結果に掲載され，解析回避機能のある踏み台サイトは Yandex の検索結果に掲載されなかった．この結果より，攻撃者の意図した挙動となっているものと考えられる．

表 7 にフィッシングサイトと偽ショッピングサイトの踏み台サイトが有する解析回避機能の比較結果を示す．一般的に，フィッシングサイトはメールに記載されたリンクをユーザがクリックしてアクセスすることを想定している．そのため，フィッシングサイトは直接アクセスのみ許可し，検索エンジン経由や検索エンジンボットのアクセスは不許可としているものが多く存在する [5]．一方で，偽ショッピングサイトは検索エンジン経由でユーザがアクセスすることを想定しており，検索エンジン経由のアクセスを許可し，直接アクセスの場合は改ざん対象のウェブサイト自体

表 7 フィッシングサイトとの解析回避機能の比較
Table 7 Comparison of the evasion function with phishing sites.

	偽ショッピングサイト	フィッシングサイト
検索エンジン経由（国内）	○	●
検索エンジン経由（海外）	▲	●
対象 URL への直接接続	▲	○
検索エンジンボット（国内）	○	●
検索エンジンボット（海外）	▲	●

（○：アクセス許可 ●：アクセス不許可 ▲：正規コンテンツを応答）

が持つコンテンツを応答する．また，検索エンジンボットによるアクセスの場合は検索エンジンにインデックスさせるための商品紹介のページを応答する．ユーザの流入経路がそれぞれ異なるため，解析回避機能もそれぞれ異なることが確認された．

（2）偽ショッピングサイトの解析回避機能

偽ショッピングサイトの中には，Proxy サーバを経由する際に HTTP リクエストヘッダに付加されることがある `X-Forwarded-For` ヘッダの有無に応じて偽ショッピングサイトへのアクセスを制御する機能を有するものが確認された．`X-Forwarded-For` ヘッダを持つ HTTP リクエストの送信時点で HTTP ステータスコード 302 が応答され，その多くが https://www.yahoo.co.jp/へリダイレクトされており，サーバサイドでのクローキングがされていた．今回検証したショッピングサイトのうち 82.7%が本機能を有することが確認された．この機能を有している理由の 1 つとして，解析者等が Proxy を経由したアクセスで調査を試行する場合に解析を回避するためであると考えられる．

6.3 偽ショッピングサイト構築ツール分析

収集した偽ショッピングサイトと踏み台サイトのコンテンツの分析結果をそれぞれ述べる．

（1）プロキシを利用した画像の読み込み

偽ショッピングサイトには多くの商品の画像が掲載されている．調査の結果，偽ショッピングサイトでは**図 7** に示す 2 つの手法で正規のショッピングサイトで使われる商品画像を表示することが確認された．

図 7 (a) に `img` タグの `src` 属性に正規のショッピングサイトの画像の URL を直接指定した場合を示す．この場合，外部の URL が Referer ヘッダに設定された HTTP リクエストが正規のショッピングサイトの Web サーバのアクセスログに残る．そのため，ショッピングサイト事業者は画像の URL のアクセスログの Referer を確認することで，画像の参照元となる偽ショッピングサイトの URL を知ることができると考えられる．

図 7 (b) に `img` タグの `src` 属性に画像を代理で取得するための画像プロキシサーバの URL を指定した場合を示す．画像プロキシサーバは，URL のクエリストリングに対象の画像の URL を指定することで画像を代理で取得することを実現している．この場合，画像プロキシサーバの

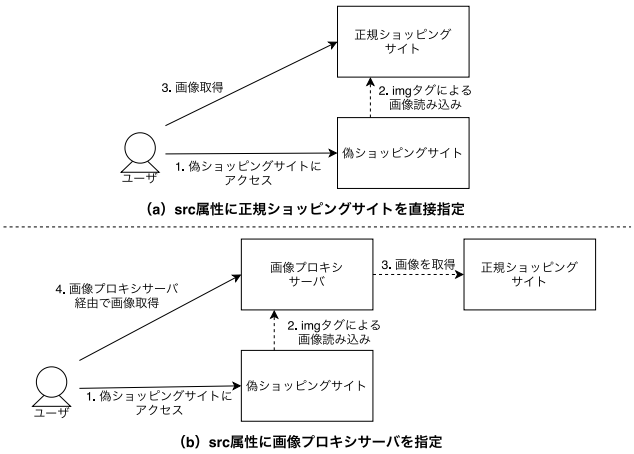


図 7 偽ショッピングサイトの画像の読み込み手法
Fig. 7 Image loading method of fake shopping sites.

IP アドレスからの HTTP リクエストが正規のショッピングサイトの Web サーバのアクセスログに残る．画像プロキシサーバは偽ショッピングサイトと同サーバに配置される場合と，別のサーバに配置される場合が確認された．そのため，参照元の偽ショッピングサイトの URL を特定することが (a) と比較して困難になる．

（2）踏み台サイトの検索エンジン掲載手法

改ざんされる可能性の高い正規のウェブサイトのドメイン名と “販売” 等のショッピングサイトで用いられるキーワードを合わせて検索エンジンで検索すると，数万件の検索結果が得られる場合がある．通常，検索エンジンに URL を掲載するためには検索エンジンのクローラにすべての URL を巡回させる必要がある．しかしながら，攻撃者が正規のウェブサイトを改ざん後に数万件の新規の URL を個別に検索エンジン事業者に巡回要求することは現実的ではない．そこで，攻撃者は検索エンジンクローラに効率的に巡回させるために sitemap.xml と，踏み台サイトの内部リンクを利用しているものと考えられる．

sitemap.xml とは検索エンジンのクローラに巡回させる URL を指示するために用いられる．sitemap.xml を踏み台サイトから取得したところ，510 件の踏み台サイトから収集できた．**図 8** にそれぞれの sitemap.xml に定義された URL 数を示す．最も多いもので 80,000 件の URL を定義した sitemap.xml も確認された．攻撃者が配置した sitemap.xml を用いて検索エンジンクローラにアクセスさせ，検索エンジン結果に掲載させることを試行したものと考えられる．

踏み台サイトには**図 9** のように検索エンジンのスニペットに掲載するための商品の説明と，同一ドメイン名内に存在する他の商品説明のページの URL への内部リンクが多数含まれている．検索エンジンクローラは内部リンクをたどるため，多数の内部リンクを同一ドメイン名の異なる URL に相互に掲載することで効率的に検索エンジン結果

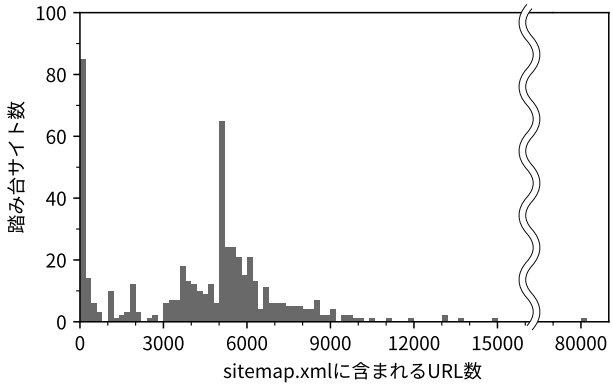


図 8 sitemap.xml に含まれる URL 数（15,000URL 以下）
Fig. 8 The number of URLs included in sitemap.xml (15,000 URLs or less).

図 9 踏み台サイトの内部リンク
Fig. 9 Internal links of landing sites.

に掲載させているものと考えられる．

（3）偽ショッピングサイトへのリダイレクト機能

6.2 節で示したように，踏み台サイトから偽ショッピングサイトへリダイレクトされるかはユーザのアクセスが検索エンジン経由であるかをもとに判断している．そのため，JavaScript コードを実行しない curl コマンドや wget コマンドで踏み台サイトのコンテンツを収集し，HTML を解析することでも偽ショッピングサイトの URL を特定することができる．しかし，踏み台サイトのコンテンツを収集するだけでは偽ショッピングサイトの URL を特定することが困難な踏み台サイトの事例が確認された．

複数の踏み台サイトにおいて Listing 3 のような JavaScript コードを含むコンテンツが応答された．Listing 3 の JavaScript コードが実行されると，名前が `beget`，値が `begetok` となる Cookie を設定した状態で reload される．Cookie が設定された状態で踏み台サイトへアクセスすると，偽ショッピングサイトへのリダイレクトするためのコンテンツが応答されるという挙動が確認された．この場合，curl コマンド等で踏み台サイトのコンテンツを取得しただけでは偽ショッピングサイトの URL を特定できないため，Selenium 等の JavaScript が実行できる環境での調査が必要となる．

Listing 3 JavaScript の実行を要するリダイレクトコード

```
<script>
    function set_cookie(){
        var now = new Date(); var time = now.getTime();
        time += 19360000 * 1000; now.setTime(time);
        document.cookie="beget=begetok"+"; expires="+now.
            toGMTString()+"; path=/";
    }
    set_cookie(); location.reload();;
</script>
```

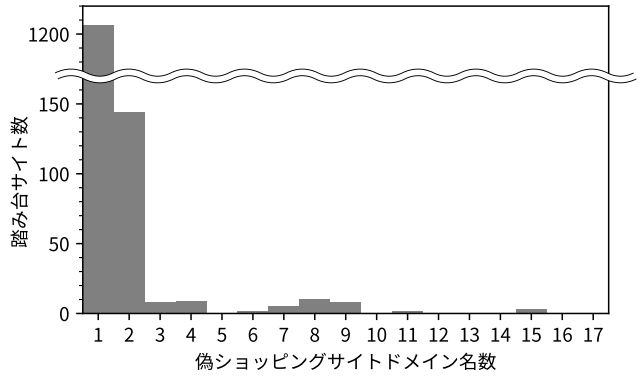


図 10 各踏み台サイトから到達した偽ショッピングサイトのドメイン名数
Fig. 10 The number of fake shopping site domain names redirected from landing sites.

6.4 偽ショッピングサイトの時系列分析

2020/9/3–2020/9/19 において，2020/9/3 時点で踏み台サイトであると確認された 1,397 件のドメイン名に 1 日 1 回アクセスした際に到達した偽ショッピングサイトのドメイン名を観測した．**図 10** に観測期間中に各踏み台サイトからリダイレクト先となった偽ショッピングサイトのドメイン名数を示す．観測対象期間において，リダイレクト先となる偽ショッピングサイトのドメイン名に変化がなかった踏み台サイトは 1,206 件（86.3%）確認された．その一方で，2 つ以上の偽ショッピングサイトへリダイレクトさせた踏み台サイトは 191 件（13.7%）あり，最も多いもので 17 日間で 15 件の偽ショッピングサイトのドメイン名へリダイレクトさせる踏み台サイトが確認された．このことから，攻撃者は複数のドメイン名を偽ショッピングサイトとして運用し，定期的に到達する偽ショッピングサイトを変更することでブラックリストによる偽ショッピングサイトへのアクセス遮断を回避している可能性が考えられる．

図 11 に観測対象期間中に改ざんが修正された踏み台サイトが検索エンジンに掲載されてから修正されるまでの日数を示す．観測期間中のある時点で踏み台サイトであることが確認され，ある時点以降偽ショッピングサイトへリダイレクトさせる挙動を示さなくなったものを改ざんが修正された踏み台サイトと定義した．検索エンジンに掲載された日時は検索エンジンクローラがウェブページを最後にクロールした時刻を利用した．そのため，検索エンジンク

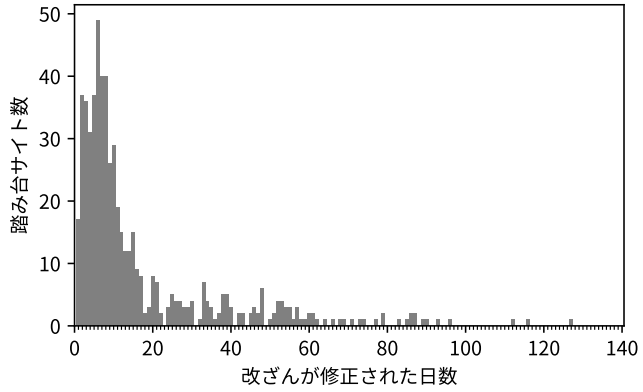


図 11 検索エンジン掲載から踏み台サイトの復旧までの日数
Fig. 11 Days that recovered landing sites from listed on search engines.

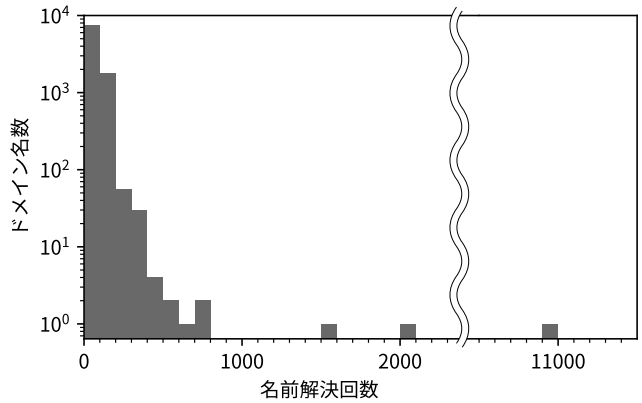


図 12 偽ショッピングサイトドメイン名の名前解決数
Fig. 12 Number of fake shopping site domain name resolutions.

ローラが複数回クロールした場合は実際に踏み台サイトとされていた日数より長くなることを制約として考慮する必要がある．なお，今回確認された踏み台サイトのドメイン名は Alexa Top Sites [22] の上位 50 万件には含まれておらず，頻繁にクロールされる人気のあるドメイン名ではないことから，本制約の影響を受けた可能性は低いと考えられる．今回収集した 2,996 件の踏み台サイトのうち，観測期間中に 579 件の踏み台サイトで改ざんの修正が確認された．そのうち 82.9%の踏み台サイトは図 11 より 20 日以内にその改ざんが修正されていることが確認された．また，平均 16.7 日で改ざんが修正されていることが確認された．

6.5 偽ショッピングサイトの攻撃規模推定

DNS での名前解決要求回数がアクセスしたユーザ数であると仮定し，偽ショッピングサイトによる攻撃規模をインターネット上の DNS サーバ間等の DNS クエリを収集したデータベースである Passive DNS のデータを用いて推定した．Passive DNS のデータは FarSight 社の DNSDB [23] を用いた．調査結果の制約として，キャッシュ DNS の影響や観測範囲の影響を考慮すると，実際のアクセス数よりも少ない結果である可能性がある．なお，本攻撃はユーザの検索エンジンでの検索が攻撃の起点となり，フィッシングのようなフィッシングメールをばらまいた直後がアクセスのピークとなる攻撃とは異なるため，同時に多くのユーザがアクセスする可能性は少なく，本制約による影響を大きく受けていないものと考えられる．また，本研究を含む調査通信によって発生した名前解決も含まれていることも制約として考慮する必要がある．なお，DNSDB がデータを取得しているサーバは公表されていないが，我々が調査の過程で確実に DNS クエリを送出したドメイン名のうち 710 件（7.1%）は DNSDB にいっさいその記録がないため，本調査の DNS クエリは DNSDB のデータ取得対象に含まれていない可能性が考えられる．

図 12 に DNSDB で名前解決が確認できた 9,344 件の偽ショッピングサイト名前解決数の集計結果を示す．最も多いドメイン名で 10,921 回の名前解決があり，当該ドメイン名は今回観測された踏み台サイトのうち 4 つのドメイン名から到達可能だった．全体で見ると，1 ドメイン名あたり平均 64.9 回の名前解決がされており，一定数のユーザが偽ショッピングサイトに到達している可能性があることが分かった

7. 議論

本章では本研究での制約，今後の課題に関して述べる．

7.1 解析回避機能調査の制約

本調査では，HTTP リクエストヘッダ（User-Agent，Referer）を変更することで解析回避機能の調査を実施した．フィッシングサイトには IP アドレスでの解析回避機能があることが知られており [4]，偽ショッピングサイトにも同様の機能がある可能性が考えられる．たとえば，検索エンジンのクローラであるかの判断をアクセス元 IP アドレスで制御している場合，本調査では検出できていないものと考えられる．なお，本調査では国内の一般ユーザと同様のアクセス条件となる国内の ISP 回線を用いた踏み台サイトへのアクセスによる調査を実施した．そのため，踏み台サイトであるかの判定結果に大きな影響はないと考えられる．

7.2 海外ユーザをターゲットとした偽ショッピングサイト

本調査では，4.3 節で述べたようにタイトルに日本語が含まれている検索結果を踏み台サイトの疑いのある URL としている．既存研究 [6], [7], [8], [9], [10], [11] において海外ユーザをターゲットとした偽ブランド商品の販売サイトが存在することが明らかになっているが，海外ユーザをターゲットとした偽ショッピングサイトも存在することが考えられる．しかしながら，本調査では海外ユーザをターゲットとした偽ショッピングサイトは調査対象に含まれていない．今後の課題として，日本国内だけではなく海外ユーザ

をターゲットとした偽ショッピングサイトの実態を明らかにする必要がある．

7.3 偽ショッピングサイト対策への提言

本調査の結果から考えられる偽ショッピングサイトに対する対策方法を提言する．

ショッピングサイト事業者：6.3 節において，偽ショッピングサイトの画像は正規のショッピングサイトから読み込まれることがあることを示した．ショッピングサイト事業者が Web サーバのアクセスログを確認し，表 6 に掲載された TLD を含むドメイン名が画像読み込みリクエストの Referer に含まれている場合，その URL が偽ショッピングサイトである可能性が高いと判断して発見することができ，テイクダウン等の対処を早期に実施できる．

セキュリティ事業者：偽ショッピングサイトは攻撃コードやマルウェアを利用しないため，シグネチャによる検知や攻撃時の振舞い検知では発見することが困難である．よって本研究の発見方法もしくは前述のショッピングサイト事業者による発見方法に基づいて，偽ショッピングサイトのブラックリストを作成し，セキュリティ製品・サービスに活用することで，エンドユーザが偽ショッピングサイトにアクセスすることを抑制できる．

検索エンジン事業者：6.3 節において，sitemap.xml を用いて検索エンジンの検索結果に踏み台サイトを掲載させる手法があることを示した．通常，sitemap.xml によって巡回対象の URL 数が急増することは考えにくい．そこで，検索エンジン事業者はあるドメイン名の過去の巡回対象 URL 数と比較して大幅に増加している等の sitemap.xml に掲載される URL 数の傾向が異なる場合に，踏み台サイトの疑いのあるドメイン名と判断して巡回を抑制することで，踏み台サイトが検索エンジンの検索結果に掲載されることを防ぐことができる．

8. まとめ

本研究では，URL ブラックリストに掲載された URL や偽ショッピングサイトのページタイトルと検索エンジンを用いて偽ショッピングサイトとその踏み台サイトを収集し，偽ショッピングサイトの解析回避機能や攻撃者が用いる構築ツールの実態を明らかにした．偽ショッピングサイトの踏み台サイトが有する解析回避機能を調査した結果，検索エンジンを経由したアクセスの場合のみ偽ショッピングサイトにリダイレクトさせる踏み台サイトが 99.8%あった．また，94.6%が日本で利用されることが多い Google，Yahoo! JAPAN 等の検索エンジンサイトを経由した場合のみ偽ショッピングサイトへリダイレクトさせた．既存研究 [2] では，検索エンジン経由の場合に偽ショッピングサイトへリダイレクトされることが示されていたが，本研究ではその中でも日本国内のユーザをターゲットとしていることを明らかにした．Passive DNS のデータを用いて偽ショッピングサイトによる攻撃キャンペーンの規模を調査したところ，1 ドメイン名あたり平均 64.9 回の名前解決があったことが確認された．今後は海外ユーザをターゲットとした偽ショッピングサイトの調査や，検索エンジンの検索結果等から偽ショッピングサイトへ到達するリンクを検知するための技術確立が必要である．

参考文献

[1] Anti-Phishing Working Group: Phishing Activity Trends Report 1st Quarter 2020 (2020), available from ⟨https://docs.apwg.org/reports/apwg_trends_report_q1_2020.pdf⟩.

[2] Japan Cybercrime Control Center: Revealed Threat of Fake Store, available from ⟨https://www.jc3.or.jp/about/pdf/JC3_APWG_Revealed_Threat_of_Fake_Store.pdf⟩.

[3] Akiyama, M., Yagi, T., Yada, T., Mori, T. and Kadobayashi, Y.: Analyzing the ecosystem of malicious URL redirection through longitudinal observation from honeypots, *Computers & Security*, Vol.69, pp.155–173 (2017).

[4] Oest, A., Safei, Y., Doupé, A., Ahn, G.-J., Wardman, B. and Warner, G.: Inside a phisher's mind: Understanding the anti-phishing ecosystem through phishing kit analysis, *APWG Symposium on Electronic Crime Research* (*eCrime*), pp.1–12 (2018).

[5] 小寺博和，芝原俊樹，千葉大紀，青木一史，波戸邦夫，秋山満昭：動的解析を利用したフィッシングサイトのアクセス妨害機能の実態解明，情報処理学会論文誌，Vol.61, No.3.

[6] Wadleigh, J., Drew, J. and Moore, T.: The E-Commerce Market for "Lemons" Identification and Analysis of Websites Selling Counterfeit Goods, *Proc. 24th International Conference on World Wide Web*, pp.1188–1197 (2015).

[7] Carpineto, C. and Romano, G.: Learning to detect and measure fake ecommerce websites in search-engine results, *Proc. International Conference on Web Intelligence*, pp.403–410 (2017).

[8] Cheung, M., She, J., Sun, W. and Zhou, J.: Detecting online counterfeit-goods seller using connection discovery, *ACM Trans. Multimedia Computing, Communications, and Applications* (*TOMM*), Vol.15, No.2, pp.1–16 (2019).

[9] Mostard, W., Zijlema, B. and Wiering, M.: Combining Visual and Contextual Information for Fraudulent Online Store Classification, *IEEE/WIC/ACM International Conference on Web Intelligence*, pp.84–90 (2019).

[10] Carpineto, C., Re, D.L. and Romano, G.: Using Information Retrieval to Evaluate Trustworthiness Assessment of Eshops., *IIR*, pp.1–8 (2017).

[11] Maktabar, M., Zainal, A., Maarof, M.A. and Kassim, M.N.: Content based fraudulent website detection using supervised machine learning techniques, *International Conference on Health Information Science*, pp.294–304, Springer (2017).

[12] Han, X., Kheir, N. and Balzarotti, D.: Phisheye: Live monitoring of sandboxed phishing kits, *Proc. 2016 ACM SIGSAC Conference on Computer and Communications Security* (*CCS*), pp.1402–1413 (2016).

[13] Invernizzi, L., Comparetti, P.M., Benvenuti, S., Kruegel, C., Cova, M. and Vigna, G.: Evilseed: A guided ap-

proach to finding malicious web pages, *2012 IEEE Symposium on Security and Privacy*, IEEE, pp.428–442 (2012).
[14] Corona, I., Biggio, B., Contini, M., Piras, L., Corda, R., Mereu, M., Mureddu, G., Ariu, D. and Roli, F.: Deltaphish: Detecting phishing webpages in compromised websites, *European Symposium on Research in Computer Security*, pp.370–388, Springer (2017).
[15] abuse.ch, available from ⟨https://urlhaus.abuse.ch/⟩.
[16] VX Vault, available from ⟨http://vxvault.net/ViriList.php⟩.
[17] Malware Domain List, available from ⟨https://www.malwaredomainlist.com/⟩.
[18] OpenPhish, available from ⟨https://openphish.com/⟩.
[19] PhishTank, available from ⟨https://www.phishtank.com/⟩.
[20] Selenium, available from ⟨https://selenium.dev/⟩.
[21] Koide, T., Chiba, D. and Akiyama, M.: To Get Lost is to Learn the Way: Automatically Collecting Multi-step Social Engineering Attacks on the Web, *Proc. 15th ACM Asia Conference on Computer and Communications Security*, pp.394–408 (2020).
[22] Amazon Web Services, Inc.: Alexa Top Sites, available from ⟨https://aws.amazon.com/alexa-top-sites/⟩.
[23] Farsight Security, Inc., available from ⟨https://www.dnsdb.info/⟩.

推薦文

本論文は，偽ショッピングサイトへの対策検討を行う前段として，踏み台サイトの挙動や攻撃者が用いる構築ツールの実態を明らかにする試みを行っている．数千件にのぼる踏み台サイトと偽ショッピングサイトのドメイン名の収集に成功し，またそれらの分析より，多くの偽ショッピングサイトが検索エンジン経由のユーザのみを偽ショッピングサイトにリダイレクトする解析回避機能を有すること，および一定数のユーザが実際に偽ショッピングサイトに到達している可能性があることを明らかにし，対策に対する提言を行っている．社会に大きな影響を与える結果であることから研究会推薦に資する論文と判断した．

（コンピュータセキュリティ研究会主査　山内 利宏）

小寺 博和　（正会員）

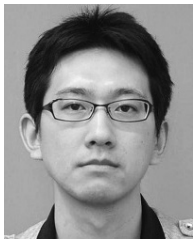
1989 年生．2011 年早稲田大学基幹理工学部情報理工学科卒業．2013 年同大学大学院修士課程修了．同年エヌ・ティ・ティ・コミュニケーションズ（株）入社．2017 年日本電信電話（株）．サイバー攻撃対策技術の研究開発に従事．現在，NTT セキュリティ・ジャパン（株）主査．

小出 駿

2014 年横浜国立大学工学部電子情報工学科卒業．2016 年同大学大学院環境情報学府情報メディア環境学専攻博士課程前期修了．同年日本電信電話（株）入社．以来，サイバー攻撃対策技術の研究開発に従事．現在，同社研究員．博士（情報学）．

千葉 大紀

1988 年生．2011 年早稲田大学基幹理工学部情報理工学科卒業．2013 年同大学大学院修士課程修了．同年日本電信電話（株）入社．以来，サイバー攻撃対策技術の研究開発に従事．現在，同社研究主任．博士（工学）．電子情報通信学会，IEEE 各会員．



青木 一史　（正会員）

1981 年生．2004 年東北大学工学部情報工学科卒業．2006 年同大学大学院情報科学研究科修士課程修了．同年日本電信電話（株）入社．以来，サイバー攻撃対策技術の研究開発に従事．現在，同社主任研究員．電子情報通信学会会員．

秋山 満昭　（正会員）

2005 年立命館大学工学部卒業．2007 年奈良先端科学技術大学院大学情報科学研究科修士課程修了．同年日本電信電話（株）入社，NTT 情報流通プラットフォーム研究所にてマルウェア対策技術の研究開発に従事．2016 年 NTT セキュアプラットフォーム研究所特別研究員．2019 年同所上席特別研究員．主としてサイバー攻撃対策技術の研究開発に従事．博士（工学）．電子情報通信学会，IEEE 各会員．