# 画像局所特徴量を利用した フィッシングサイト検知手法の実装と評価

高橋啓伸[1]　小倉加奈代[2]　Bhed Bahadur Bista[2]　高田豊雄[2]

**概要**：近年，正規のオンラインバンクやオンラインショップのサイトを模倣し，個人情報を窃取するフィッシングサイトが問題となっている．フィッシングサイトはユーザを騙すために模倣元サイトと似たデザイン (ロゴマーク，ボタン等) を使うという特徴がある．そこで我々はフィッシングサイトと模倣元サイトの中から部分的に共通するデザインの画像データベースを作成し，アクセスサイトのキャプチャ画像とデータベース画像の画像局所と特徴量からフィッシングサイトかどうかを判定する手法を提案，実装した．その結果，フィッシングサイトを 88.5 %の精度で正しく検出でき，提案手法の有効性を確認した．

**キーワード**：画像局所特徴量，フィッシングサイト

# Implementation and Evaluation of Phishing Site Detection Method Using Image Local Features

Takahashi Hironobu[1]　Ogura Kanayo[2]　Bhed Bahadur Bista[2]　Takata Toyoo[2]

***Abstract:*** Recent years, to mimic online banks and original sites of shops, phishing sites to steal personal information have become a problem. Phishing sites have a feature that mimics original sites with similar design or the same design to fool users. In this paper, we have proposed and have implemented a method for determining the fishing by creating an image database of partially common designs of fishing sites and the authentic sites and calculate the image local features of captured image and database image of accessed site. A result, the phishing sites can be correctly detected in 88.5 % of accuracy, to confirm the validity of the proposed method.

***Keywords:*** Image local features，Phishing site

## 1. はじめに

近年，正規のオンラインバンクやショップのサイトを模倣し，アカウント ID，パスワード，クレジットカード番号等の個人情報を窃取するフィッシングサイトが問題となっている．フィッシング対策協議会の調査報告 [1][2] によると，2015 年の国内フィッシングサイト報告数は，11,408 件と前年から 2,676 件減少しているが，被害額はわずかに増大している．また，国外では 2015 年上期のフィッシングサイト報告数は過去最高水準となった．依然として被害は拡大傾向にあり，フィッシングサイトへの効率的な対策は急務である．

フィッシング対策技術の代表的な例として URL フィルタリング方式が挙げられる．しかし，Sheng らによって行われた 2009 年の調査研究 [3] によると出現から 1 時間以内のフィッシングサイトが 20 %未満しかデータベースに登録されていないことが示さ

[1] 岩手県立大学大学院
Presently with Iwate Prefectural University Graduate School
[2] 岩手県立大学
Presently with Iwate Prefectural University

れており，新しいフィッシングサイトに効果的でない傾向がある．日々大量に現れるフィッシングサイトに対して逐次データベースを更新していくのは効率的ではない．よって，頻繁なデータベース更新を必要としない，フィッシングサイトの特徴を利用した発見的な手法が必要とされている．

そこで我々はフィッシングサイトの発見的な手法として，サイト画面画像の部分的な特徴を比較する手法を提案する．フィッシングサイトとその模倣元サイトの視覚的な類似性は，フィッシング詐欺として隠ぺいできない特徴であり，これを利用することで高い検出精度が期待できる．フィッシングサイトのサンプルとその模倣元サイトから，共通するデザインやロゴを抜き出した画像，模倣元サイトのドメインを保存したホワイトリストの 2 つを構築し，判定したいサイトの画面画像及びドメイン情報と比較することでフィッシングを判定する．本稿では提案手法を説明し，実装したシステムと実在するフィッシングサイトを利用した評価実験及び実験結果について述べる．

## 2. 既存研究

### 2.1 テキスト特徴による手法

Zhang らはフィッシングサイトとその模倣元サイトの類似性に基づいた検出手法として CANTINA[4] を提案した．この提案は，TF-IDF 法を用いてサイト中のテキストから特徴となる単語を抽出し，それを web 検索した結果を利用してフィッシングを検出する手法である．検出率が 89 %，誤検出が 1 %と高い検出精度を示した．しかし，Web ページを開発する HTML 等のソースコードは，レイアウトの種類や難読化，暗号化手法が多数存在している．フィッシングサイト開発者は，これらを利用してフィッシングサイトを開発することでテキスト特徴による類似度を下げ，検出を避けることが可能である．また，近年は模倣元サイトのスクリーンショット画像を利用し，画像のみで構成されたフィッシングサイト [5] も確認されており，このような方法でも類似度を下げられる可能性がある．

### 2.2 画像的な特徴による手法

原らは web サイト中のテキストによる情報を使わず，画像的な類似度を用いた手法がよって提案した [6][7]． フィッシングサイトとその模倣元サイトが画像的に類似していることを明らかにし，正規サイトの画面画像とドメイン情報を保存したデータベースを利用してフィッシングを判別する手法を提案した．検索対象のサイト画面画像をキーとしてデータベースから類似した画像を検索し，その結果類似した画像があり，かつドメインが正規のものと一致しなかった場合にフィッシングであると判断する．画像の類似判定には類似画像検索ソフト ImgSeek[8] を利用した．正規サイト 521 件とフィッシングサイト 200 件をそれぞれ 2 分割し，一方を画像データベースとし，他方を評価する実験を相互に行った結果，検出率が 82.5 %，誤検出 が 21.5 %となった．検出できなかったフィッシングサイトはデータベースに模倣元サイトが少なかったサイトや，ページ内広告等の部分的な表示画像の違いによって類似度が下がったものだった．図 1 は広告画像や記事のサムネイル画像の違いによって類似度が下がり，検出できなかった例である．また，同じサイトを模倣したフィッシングサイトであっても全体の構成やデザインは大きく違う場合もあり，差分ごとにデータベースへスクリーンショットを追加しなければ，そのようなサイトも検出できない可能性が考えられる．図 2 は同じ模倣元サイトを持ちながらボタン等の UI やレイアウトの違うフィッシングサイトの例である．この程度のデザイン差異があるフィッシングサイトは無数に存在し，画像的な特徴を抽出するためにはある程度のロバスト性が必要であると考える．

図 1 原らの研究で検出できなかった例

**Fig. 1** Example of could not detected in the study of the Hara et al.

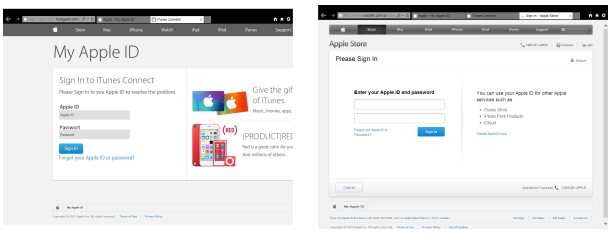


図 2 同じ模倣元を持ちつつ異なるデザインのフィッシングサイトの例

**Fig. 2** Examples of phishing sites of different design and same imitation source

## 3. 提案手法

本提案手法は，Web ブラウザ上の当該サイトのキャプチャ画像から画像局所特徴量を算出し，事前に構築したデータベース画像及びドメインホワイトリストと比較することでフィッシングサイトか否かを判定する．提案手法の処理手順を図 3 に示す．

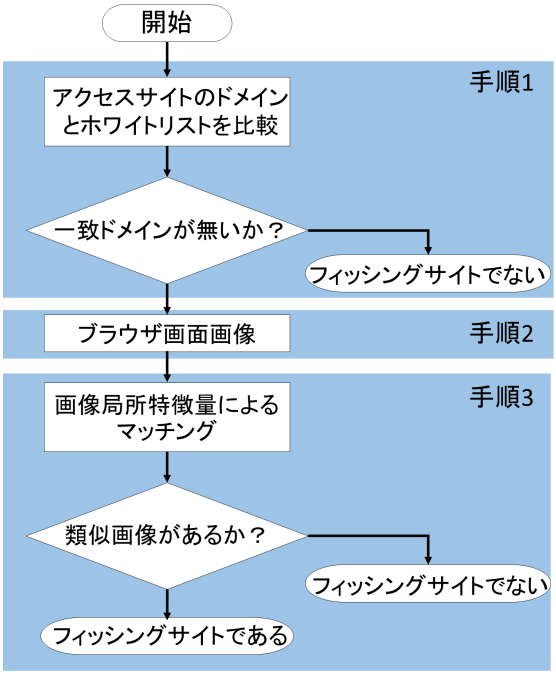


図 3　提案手法フローチャート
**Fig. 3**　Flowchart of proposed method

### 3.1 事前準備：ホワイトリストと画像データベースの構築

フィッシング判定処理を行う前に，アクセスサイトとの比較対象となる共通デザイン画像データベースとドメインホワイトリストを構築する．共通デザイン画像データベースには，複数のフィッシングサイトとその模倣元サイトから共通するデザインを抜き出し，画像形式で保存したものが格納されている．ドメインホワイトリストは，共通デザイン画像データベースに格納されている画像の模倣元オリジナルサイトのドメイン，及び正規のサイトである事が確認されているサイトのドメインを格納したリストである．複数のドメインで運営されている場合にはそのすべてのドメインを保存する．この 2 つはいずれも手動で構築する．

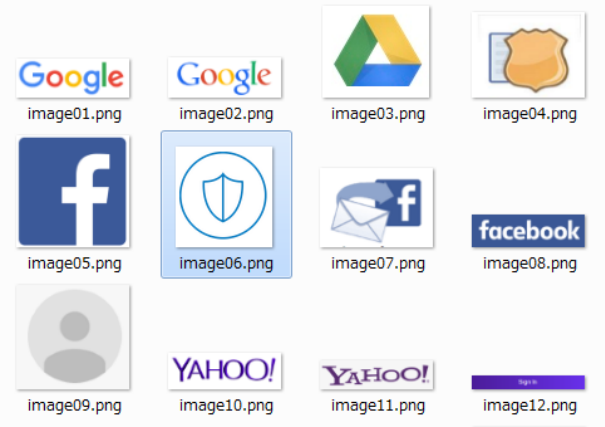


図 4　共通デザイン画像データベースのイメージ
**Fig. 4**　Image of common design image database

### 3.2 手順 1：ドメインの照合

アクセスサイトのドメインとホワイトリスト内のドメインを比較する．一つでも一致があった場合はフィッシングサイトでないと判断する．一致しない場合は手順 2 へ進む．

### 3.3 手順 2：サイト画面画像保存

アクセスしたページのブラウザ表示画面から表示部のみを抜きだし，キャプチャ画像として png 形式で保存する．アクセス完了時にブラウザに表示されている範囲をキャプチャするため，スクロールなどによって確認するページ下部等は含まない．

### 3.4 手順 3：画像局所特徴量を利用した画像マッチング

本提案では局所的な画像の特徴を抽出するため，SURF 特徴量抽出アルゴリズム (Speeded-Up Robust Features)[9] を使用する．特徴量抽出アルゴリズムは複数種類存在するが，本提案では実装時のオーバーヘッドを考慮し，高速処理手法である SURF を採用する．

特徴量抽出アルゴリズムを利用することで画像中の特徴となる点を抽出し，その周囲の領域から特徴量を算出する．そして算出した特徴量を比較することで画像間の類似した箇所を検出することができ，これを利用してキャプチャ画像と共通デザイン画像データベースとの類似箇所を検出する．画像局所特徴量は画像の回転，明度変化，サイズ変化に一定の不変性があり，少ないデータベース画像で効率的な類似検出が可能となる．特徴点から類似画像を検索するイメージを図 5 に示す．

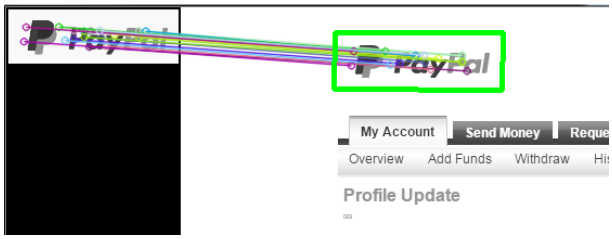


図 5　局所特徴量による画像マッチングのイメージ図
**Fig. 5**　Image view of image matching by local features

### 3.5　期待される効果

ブラウザが表示したサイト画面画像を用いることで，攻撃者がユーザを騙す上で不可欠な模倣元サイトとの視覚的類似という隠蔽できない情報を使った検出ができる．広告画像の変更等によって画像全体の類似度が変わった場合も，局所的な特徴を利用しているため影響を受けない．また，仮にフィッシングサイト製作者が本提案を深く理解している場合でも，検出されないサイトを構築するためには，データベースに使われている正規のサイトやフィッシングに良く使われるデザインを避けければならない．提案手法をすり抜けるサイトは模倣元サイトから大きくかけはなれたデザインとなるため，フィッシングサイト構築が難化すると考えられる．

## 4.　評価実験

本章では，提案手法の有効性を確認するため，実在のフィッシングサイトと正規サイトを用い，正誤検出率の評価実験，データベースと類似するデザインを含むサイトの誤検出の 2 つの評価実験を行った．

### 4.1　実験 1：フィッシングサイトの正誤検出

提案手法が，ブラウザ表示画面画像をもとにフィッシングサイトをどの程度正確に判別できるかを評価するために，正誤検出実験を行う．

#### 4.1.1　正誤検出評価の実験手順

正検出実験に使用する模倣元サイトは Alexa トラフィックランキング [*1] による上位サイトで，かつ PhishTank[*2] による対象ブランド分類がされている Google[*3]，Yahoo[*4]，Facebook[*5]，Apple[*6]，及びフィッシングサイト数の多い Paypal[*7] の 5 つを選出した．使用するサイト件数は，それぞれのサイトを模倣したフィッシングサイトを 40 件ずつ，合計 200 件である．実験手順は以下のとおりである．

**手順 1**　画像データベースの構築を行う．実験対象フィッシングサイトは別に，各サイトからフィッシングサイトを 5 件ずつ選出し，模倣元サイトのデザインと比較して共通画像データベースを作成する．

**手順 2**　ホワイトリストに，実験対象フィッシングサイトの模倣元サイトドメイン情報を登録する．

**手順 3**　PhishTank に報告されたサイトの中から，対象ブランド分類によって対象模倣元サイトに分類されているものを，最新順に 40 件ずつ収集する．

**手順 4**　収集したフィッシングサイトに対し，提案手法を使用してフィッシング検出処理を行う．データベース中から一つでも正しく類似箇所を検出できた場合，正検出とする．

#### 4.1.2　誤検出評価の実験手順

誤検出実験に使用する正規サイトは，Alexa によるトラフィックランキング上位サイトから 4.1.1 項の正検出実験で用いる対象模倣元サイト以外の 200 件を対象とする．なお，フィッシング判定に使用する画像データベースおよびホワイトリストは前項 4.1.1 の正検出実験と同じものを使用する．その他の手続きは，まず，Alexa によるトラフィックランキング上位サイトから，前項 4.1.1 の正検出実験の対象模倣元サイト以外のサイトを 200 件に達するまで収集し，収集した正規サイトに対し，提案手法を適用しフィッシング検出処理を行う．データベース中からひとつでも類似を検出した場合，誤検出と見なす．

### 4.2　実験 2：データベース内と類似するデザインを含むサイトによる誤検出

本提案は正規サイトと類似するデザインが含まれ，かつフィッシングサイトではないサイトを誤検出する可能性が極めて高い．画像マッチングの検出精度が高いほど避けられない課題であるため，対応の重要性と方向性を明らかにする必要がある．そこで前節の正誤検出実験で使用した共通デザイン画像データベースに格納されている画像と類似，もしくは一致した画像を含むサイトに提案手法を適用し，どの程度誤検出が生じるかを調査する．なお，使用する画像データベースおよびホワイトリストは，前節の正誤検出実験と同じものを使用する．その他の実験

*1　http://www.alexa.com/
*2　https://www.phishtank.com/
*3　https://www.google.com/
*4　https://www.yahoo.com/
*5　https://www.facebook.com/
*6　http://www.apple.com/
*7　https://www.paypal.com/

手続きは，まず，共通デザイン画像データベースに格納されている画像をキーに Google 画像検索で検索を行い，検索結果として表示されたサイト 65 件を収集する．次に，収集したサイトに対し，提案手法を適用しフィッシング検出処理を行う．データベース中からひとつでも類似を検出した場合，誤検出と見なす．

## 5. 実験結果と考察

### 5.1 実験 1: 正誤検出評価の実験

正検出実験結果を表 1 に，誤検出実験結果を表 2 に示す．正検出については，平均検出率は，88.5 %であり，画像のみで構成されたサイトや部分的に同じデザインを使用しながら全体の配置が違うサイトも正しく判別できた．誤検出については，4.5 %と低い誤検出率に抑えることができた．

正検出，誤検出それぞれについて，正しく検出できなかったサイトを分析した．その結果，正検出については，多くは，模倣元サイトとかけ離れたデザインのため，共通デザイン画像データベースとの類似で検出できないものであった．実際の例を図 6 に示す．特に Facebook において類似性の低いフィッシングサイトが多く，他のサイトと比較して大きく検出精度が低かった．このようなサイトは人間の目から見ても模倣元との違いが明らかであるため，提案手法を改良するのではなく，ユーザ側で確実に判断させることで対処可能であると我々は考える．この点について，今後，該当サイトが不正なサイトであるか否かをユーザに判断させる実験を行い，改善策を検討する．また，正検出について，模倣元サイトと同様のデザインを使用しながら，正しく検出できなかったフィッシングサイトも少数あった．ロゴや UI の背景が，模倣元サイトは無色であったのに対し，該当フィッシングサイトは写真を背景として使用していたことで類似性を検出することができなかった．実際の例を図 7 に示す．スクリーンショットを利用しているため，背景とデザインは同じレイヤの画像として扱われる．よって，ロゴや UI を透過した場合や，隙間から背景が見えている場合に正しく検出できない可能性がある．

誤検出については，正しく判別できなかったサイトは，その全てが共通デザイン画像データベース内の画像とは全く異なる箇所を類似として検出していることが確認できた．これは偶然類似した局所特徴量が集中し，その集中箇所の相対的な位置もある程度一致したものであると考えられる．

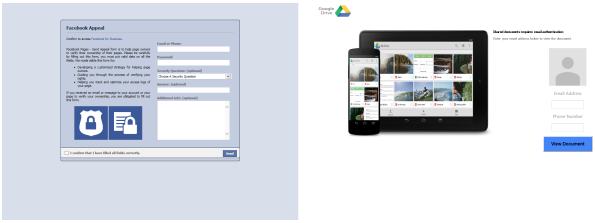

図 6 模倣元 (図左) とかけはなれたデザインのフィッシングサイト (図右)

**Fig. 6** Phishing website of design that far removed from imitation source


図 7 背景の変更によって検出できなかった例

**Fig. 7** Example that could not be detected by the change of the background

### 5.2 実験 2: データベース内と類似するデザインを含むサイトによる誤検出実験

実験 2 の結果を表 3 に示す．32.3 %の誤検出が生じており，実験 2 の結果に比べて高い比率で誤検出が生じた．データベース内と類似するデザインを含み，かつフィッシングサイトでないものに対して対策を講じる必要がある．Zhang らは本研究と同じく模倣元サイトとの類似性を利用しているが，複数のヒューリスティクスを組み合わせてフィッシングサイトを検出することで誤検出を低く抑えている．本提案においても，ドメイン登録期間や Pagerank 等の Web ページのソースコードに依存しない特徴を利用したヒューリスティクスと組み合わせる必要があると考えられる．それぞれの検出精度からスコアを算出し，重みづけを行った総合的な基準による判別手法を提案する予定である．

表 1 実験 1 の結果

**Table 1** Results of Experiment 1

| | サイト数 | 検出数 | 検出率（%） |
|---|---|---|---|
| Google | 40 | 37 | 92.5 |
| Apple | 40 | 36 | 90 |
| Paypal | 40 | 38 | 95 |
| Yahoo | 40 | 38 | 97.5 |
| Facebook | 40 | 28 | 67.5 |
| 全体 | 200 | 176 | 88.5 |

表 2　実験 2 の結果
**Table 2**　Results of Experiment 2

| | サイト数 | 検出数 | 検出率（%） |
|---|---|---|---|
| 全体 | 200 | 9 | 4.5 |

表 3　実験 3 の結果
**Table 3**　Results of Experiment 3

| | サイト数 | 検出数 | 検出率（%） |
|---|---|---|---|
| 全体 | 65 | 21 | 32.3 |

### 5.3 考察: オーバーヘッド

1 件のサイトに対してフィッシングサイト判定を行う時間から，サイトへのアクセス時間を除いた数値を本提案手法によるオーバーヘッドと考える．オーバーヘッドは平均すると 1 サイトごとに 12438 ミリ秒 (CPU: Core i7 2.93GHz，HDD: WD5000AAKS ) 要しており，現時点ではユーザがブラウザでアクセスしたサイトすべてに適用するといった用途に組み込むのは難しいと考えられる．計算時間は対象サイトキャプチャ画像のサイズに比例しているため，キャプチャ画像及びデータベース画像の軽量化や，特徴点検出数の制限，特徴量抽出アルゴリズムの変更などが処理の高速化に有効と考えられる．いずれも画像マッチングの精度に影響を及ぼすと考えられるため，今後これらの最適な値を検討する必要がある．

## 6. まとめ

本稿ではブラウザ画面画像の局所的な特徴量とドメイン情報を利用したフィッシングサイト検出手法を提案し，その性能を実在の Web サイトを用いた実験によって評価した．実験の結果，フィッシングサイトを 88.5 %の精度で正しく検出し，正規サイトを 4.5 %の割合で誤検出することを確認した．この結果から局所的な特徴量を利用した画像マッチングはフィッシングサイト判別手法としてある程度の精度が保証でき有効であると言える．また，検出できなかったサイトの分析から，検出できなかったサイトは大部分は人間の目から見ても模倣元との違いが明らかであることがわかった．この点について，今後，ユーザによる検出実験を進め，ユーザ側からの確実な判別可能性を検討する．さらに，本提案手法の問題点として，データベース内の画像と類似点があり，かつ正規のサイトに対して高い割合で誤検出を生じる点，本手法の処理時間として，1 サイト平均 12438 ミリ秒のオーバーヘッドが生じている点の 2 点があげられる．今後は処理の高速化と, 画像処理以外のヒューリスティクスを組み合わせた総合的な判別手法の開発に取り組む予定である.

**謝辞**　本研究は一部，JSPS 科研費 26330159 および，16K01025 の助成を受けたものである．

### 参考文献

[1] フィッシング対策協議会:フィッシングレポート 2016 の掲載 ～ 世界に広がるフィッシング対策の輪 ～(online), 〈https://www.antiphishing.jp/report/pdf/phishing_report_2016.pdf〉, (2016.07.22)

[2] フィッシング対策協議会:フィッシングレポート 2016 の掲載 ～ 進む対策、利用者としてできること ～(online), 〈https://www.antiphishing.jp/report/pdf/phishing_report_2015.pdf〉, (2016.07.22)

[3] Steve Sheng, Brad Wardman, Gary Warner, Lorrie Faith Cranor, Jason Hong, and Chengshan Zhang: An Empirical Analysis of Phishing Blacklists, *In Proceedings of the 6th Conference on Email and Anti-Spam* (online), available from 〈http://repository.cmu.edu/cgi/viewcontent.cgi?article=1286&context=hcii〉, (2009)

[4] Yue Zhang, Jason Hong, and Lorrie Cranor: CANTINA : A Content-Based Approach to Detect PhishingWeb Sites, *In Proceedings of the 16thWorld Wide Web Conference*, pp. 639–648, (2007)no

[5] ITmedia Inc. : 国勢調査の“偽サイト”作った意図は？ 総務省から削除依頼……「騒ぎになり深く反省」と制作者, 〈http://www.itmedia.co.jp/news/articles/1509/15/news083.html〉, (2016.07.23)

[6] 原正憲，山田明，三宅優: ブラウザ表示を利用した悪意あるサイト検知方式の提案, 情報処理学会研究報告コンピュータセキュリティ 2008, pp. 49–54, (2008)

[7] 原正憲，山田明，三宅優: ブラウザ表示を利用したフィッシングサイト検知方式の評価, 電子情報通信学会ソサイエティ大会講演論文集 2008 年通信 (2), p. 84, (2008)

[8] imgSeek(online):〈https://sourceforge.net/projects/imgseek/〉, (2016.07.29)

[9] Herbert Bay, Andreas Ess, Tinne Tuytelaars, and Luc Van Gool: Speeded-Up Robust Features (SURF), *Computer Vision and Image Understanding*, pp. 346–359, (2008)