6V-09

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提案

高橋 啓伸 †　　小倉 加奈代 †　　Bhed Bahadur Bista†　　高田 豊雄 †

† 岩手県立大学

1　はじめに

金融サイトやオンラインショップを模倣し，個人情報を窃取するフィッシング詐欺は昨今問題となっている．

既存のフィッシング対策手法の例として，URL フィルタリング方式が挙げられる．これはフィッシングサイトのアドレスをデータベースに登録し，アクセスしたサイトのアドレスと比較することでフィッシングを判定する．しかし Sheng らによって行われた 2009 年の調査研究 [1] 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れており，新しいフィッシングサイトに弱い傾向がある．よって，手動ではなく自動的にフィッシングサイトを検出する手法が必要とされている．

そこで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[2] が提案された．この提案はサイト中の文言から特徴語を抽出し，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，HTML 等の，ソースコードのレイアウトの種類や難読化手法は多数存在しており，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文言による類似度を下げることが可能である．

原らは Web サイトの文言による情報に頼らず，表示的な類似度を用いた手法を提案した [3][4]．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，正規サイトのブラウザ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検出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，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．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であった．

そこで我々はサイトの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と，模倣元サイトのドメインを保存したデータベースを構築し，フィッシングサイトの画面画像及びドメイン情報と比較してフィッシングを検出する．本稿では提案手法を説明し，有効性の検証実験とその結果について述べる．

2　提案手法

2.1　提案手法の概要

本稿では画面画像から特徴を抽出し，データベース画像と比較を行うことでフィッシングサイトを検出するアルゴリズムを提案する．提案手法の処理の流れを図 1 のフローチャートに示す．画像マッチング処理と，事前に構築される画像データベースについては次節から説明する．

手順 1 アクセスしたページのブラウザ画面から表示部のみを抜きだし，png 形式で保存する．

手順 2 データベース画像と類似した箇所が無いか SURF による局所特徴量に基づいて走査する．類似箇所が検出された場合は手順 3 へ，検出されなかった場合はフィッシングサイトでないと判定する．

手順 3 アクセスサイトのドメインと，検出したデータベース画像の類似元サイトのドメインを比較し，一致しない場合はフィッシングサイトと判定する．

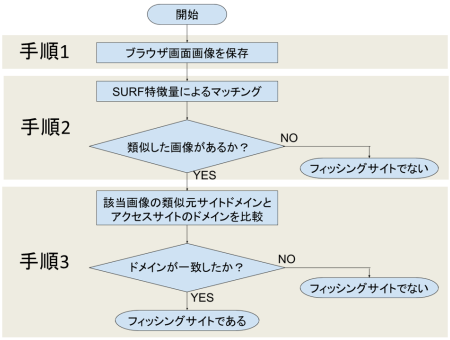

図 1: 提案手法フローチャート

2.2　SURF 特徴量による画像マッチング

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)[5] を使用する．特徴量抽出アルゴリズムには複数種類があるが，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，これを利用してフィッシングサイトの画面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．

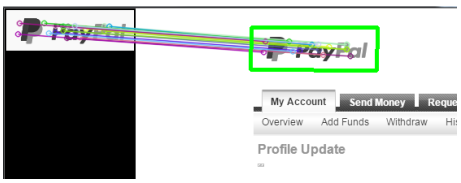

図 2: 局所特徴量による画像マッチングの例

2.3　共通デザイン画像データベース

複数のフィッシングサイトとその模倣元サイトの共通デザイン画像と，模倣元サイトドメイン情報を保存したデータベースを処理前に手動で構築する．複数のドメインで運営している場合にはそのすべてのドメインを保存する．このデータベース中の画像をアクセスサイトのブラウザ画面画像中から走査した結果と，ドメイン情報の比較によってフィッシングを判定する．

2.4 期待される効果

サイトのブラウザ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，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てサイトを構築しなければならないため，フィッシングサイト構築が難化すると考えられる．

3 有効性の検証

SURF 特徴量による局所画像特徴量を利用した画像マッチングが，フィッシングサイト検出に十分有効な精度であるかを評価する．実験を行うため SURF 特徴量に基づき，任意の検索画像からデータベース内の画像と画像マッチングするプログラムを開発した．

3.1 実験 1: 検出率の評価実験手順

SURF 特徴量に基づく画像マッチングが，フィッシングサイトを表示したブラウザの画面画像から，データベース内の画像と類似した部分をどれほど正確に検出するかを調べた．画面画像に使ったサイトは Alexa[1] による閲覧上位サイトで，PhishTank[2] によるカテゴリ登録がされている Google，Yahoo，Facebook，Apple，及びフィッシングサイト数の多い Paypal の５つを選出した．以下に実験の手順を示す．

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 実験対象サイトの画面画像を取得する

手順 3 集めた画面画像に対して，画像マッチングを行う．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，検出成功とする．

3.2 実験 2: 誤検出率の評価実験手順

SURF 特徴量に基づく画像マッチングが正規サイトを表示したブラウザの画面画像に対してどの程度の誤検知が生じるか調べた．Alexa による閲覧上位サイトから，実験 1 対象サイト以外のサイト 100 件を対象とする．

手順 1 画像データベースの構築を行う．内容は実験 1 で使用したデータベースと同様である．

手順 2 実験対象サイトの画面画像を取得する．

手順 3 集めた画面画像に対して，画像データベースを使って画像マッチングを行う．

4 実験結果と考察

3.1 節，3.2 節で説明した実験の結果については，検出精度は全体的に高く，100 件のフィッシングサイトのうち 91 件を検出し，誤検出も正規サイト 100 件中 8 件に抑えられた．検出ができなかったサイトは，画像データベースに類似の画像があるにも関わらずできなかったものが 8 件，模倣元のサイトと著しくデザインが変わったものが 1 件であった．計算時間は平均約 8.8 秒という結果となった．計算時間は画像サイズの数に比例しているため，画像自体の軽量化とデータベース画像の削減，特徴点抽出数の制限が処理の高速化に有効であると考えられる．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．

また，今回の実験では明らかにならなかったが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誤検出であるため，実装の際には Pagerank 等の Web ページのソースコードに依存しない特徴と組み合わせる必要があると考えられる．

5 まとめ

本稿ではフィッシングサイトのブラウザ画面画像の局所的な特徴量を利用した検出手法を提案した．そして局所的な特徴量を利用した画像マッチングがフィッシング検出に対して有効かを実験によって検証し、データベース内画像との類似箇所を 91 %の精度で検出し，類似の無い箇所を 8 %の割合で誤検出することを明らかにした．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検出の情報源として十分な精度であると考えられる．今後は画像マッチングの高速化を目指した更なる分析と，システム全体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 の助成を受けたものである．

参考文献

[1] 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)

[2] Yue Zhang, Jason Hong, and Lorrie Cranor: CANTINA : A Content-Based Approach to Detect Phishing Web Sites, *In Proceedings of the 16th World Wide Web Conference*, pp. 639–648, (2007).

[3] 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)

[4] 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)

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

[1] http://www.alexa.com/

[2] https://www.phishtank.com/