Vol.2018-SPT-31 No.11
Vol.2018-EIP-82 No.11
2018/11/2

複数の Web サイト安全性評価サービスを利用した URL 評価手法の検討

藤根 麻羽[†1]　小倉 加奈代[†1]　Bhed Bahadur Bista[†1]　高田 豊雄[†1]

概要：短縮 URL サービスは，冗長な URL を簡素化するため，文字数制限のあるソーシャルネットワークサービス(SNS)でのメッセージ投稿時にしばしば利用される．短縮 URL は利便性が高い一方で，行先サイトの URL 難読化を悪用したフィッシング詐欺に利用される事例が発生している．そのため，ユーザは短縮 URL ごとにその安全性を判断する必要がある．この問題を解決するために我々は，短縮 URL を展開した URL を，複数の安全性評価サービスにより検査し，その結果を統合した上で安全性を提示する手法を検討している．また，複数の安全性評価サービスの結果を統合する際に，各評価結果の偏りをなくすため，評価サービスごとに重み係数を算出し利用することを検討している．本稿では，良性サイト，悪性サイトのデータベースから集めた URL と，Twitter 等の SNS 投稿文から抽出した短縮 URL の 2 種類の URL に対し，重み係数の算出に SVM やロジスティック回帰等，複数の分類器を用いた統合手法による評価結果と本手法の有効性を検討する．

キーワード：短縮 URL，SNS，安全性評価サービス

A Study on URL Evaluation Method Using Multiple Web Site Safety Services

MAU FUZINE[†1]　KANAYO OGURA[†1]
BHED BAHADUR BISTA[†1]　TOYOO TAKATA[†1]

Abstract: Shortened URL service is often used when posting messages to social network service (SNS) with a limited number of characters in order to simplify lengthy URLs. Although shortened URLs are highly convenient, the damage are occurring due to phishing scams that exploit URL obfuscation of the destination site. Therefore, it is necessary for users to determine the safety of each shortened URL. In order to solve this problem, we examined the expanded URL of the shortened URL using multiple safety evaluation services, and decide the weighting factor for each evaluation service to eliminate the bias of each evaluation result and study the method for integrating the results.

In this paper, we calculated the weighting factor to the two types of URLs collected from the benign site database and malignant site database, and shortened URLs taken from SNS post such as Twitter etc., and applied multiple classifiers such as SVM and Logistic Regression etc., and examine the evaluation results and effectiveness of the integration method of the results.

Keywords: Short URL, SNS, Web Site Safety Services, Classifire

1. はじめに

Twitter[1]や Facebook[2]を代表とするソーシャルネットワークサービス(以下，SNS)は 2000 年代後半に登場し，現在幅広い年代のユーザに利用されている．SNS の投稿サービスには，文字数に制限が設けられているものがあり，URL を含むメッセージ投稿はその制約を受けやすい．そのような文字数制限への対処法として，短縮 URL サービスがしばしば利用される．短縮 URL サービスは対象の URL と対になる簡素な URL を生成するサービスであり，goo.gl[3]や bit.ly[4]がよく知られている．短縮 URL はリンクの冗長性を排除するだけではなく，サービスによってアクセス統計機能や QR コード生成等が可能である．そのため，個人や企業を問わず SNS キャンペーン活動の際に用いられることもある．このような利便性の一方で，行先サイトの URL が短縮 URL サービスのドメインに変換され，行先サイトが不明瞭になる問題点がある．この問題点を悪用した SNS を対象としたアカウントのなりすましやのっとり，フィッシングなどを実行する攻撃である Social Media Attack[5]が発生している．実際，2016 年 5 月に，スパイウェアのダウンロードリンクを短縮し，その URL を Twitter へ投稿・拡散された被害[6]がある．また，SMS メッセージに添付された URL によりフィッシングサイトへ誘導し，クレジットカード番号などの架空請求詐取に用いられる被害[7]も起こっている．さらに．攻撃者自身が短縮 URL サービスを運営し，フィッシングサイトの短縮 URL を生成するケース[8]も実際に起こっている．こういった Social Media Attack の特徴は，SNS のような多くのユーザが目にするメディアで拡散・攻撃活動を行う点，手口が巧妙である点，リダイレクトの複雑化によって引き起こされることが多い点であ

†1 岩手県立大学
Iwate Prefectural University

ⓒ 2018 Information Processing Society of Japan

Vol.2018-SPT-31 No.11
Vol.2018-EIP-82 No.11
2018/11/2

る．以上のことから，ユーザは短縮 URL から行先 Web サイトが安全であるか否かを URL ごとに判断する必要がある．しかし，短縮 URL 自体から得られる情報は非常に少ない．ユーザが行先サイトの安全性を確認するには，短縮 URL を展開する Web サイトやサービスを利用し，完全な行先 URL を取得したのち，安全性評価サービスによる検査結果を利用することが有効な対策の 1 つである．

我々は，短縮 URL を展開した URL に，複数の安全性評価サービスに適用し，それらの検査結果を統合する手法を検討している．本稿では，提案手法による短縮 URL の安全性判断結果の精度を確認し，提案手法の有効性を評価する．また，提案手法における重み決定について複数の分類器を利用し，2 種類の評価用データセットに提案手法に適用し，安全性判定精度および，各分類器の違いによる安全性判定精度の違いを評価・検討する．

2. 関連研究

本章では，既存の安全性評価サービスについて述べたのち，複数の安全性評価サービスを利用した URL の安全性提示手法について述べる．

2.1 既存の安全性評価サービス

Web ブラウザやセキュリティソフトに搭載されているものを含め，様々な視点から URL の安全性検査を行う評価サービスが存在する．安全性評価サービスは独自の脅威情報やブラックリストを保持し照会した結果や，Web サイトのクローリングによる検査から安全性評価結果が提供される．各評価サービスは良性および悪性サイトの判断基準が統一されていないことや，各運営者によって検査基準が異なっている等の理由から，同一 URL 検査する場合，評価サービスごとに検査方法とその結果にゆれが生じる可能性がある．そのため本研究では評価結果の偏りを考慮する目的で，複数の評価サービスを利用する．

安全性評価サービスの例として，Google の提供する Google Safe Browsing[9]をあげる．これは，Google Chrome ブラウザに標準搭載されているほか，API，Web サービスとして利用可能である．フィッシングサイトやマルウェアをホストするサイトの URL をリスト化しておき，安全性の検証を行ったうえで当該サイトを表示する．

2.2 検査結果を加算し安全性評価値を提示する手法

坂松らは，サイトの安全性と重要度に応じたパスワード管理ツールに関する研究[10]の中で，サイトの安全度を評価項目としている．この安全度は複数安全性評価サービスを利用した評価項目によって構成されている．具体的には，Virus Total[11]によるウィルス検査，Spamhaus Block List[12]などによるブラックリスト判定，Google Safe Browsing などによる総合安全度の 3 つである．対象の URL が異常またはブラックリストに含まれる場合を 1 として，単純加算による最終的なサイト安全度の算出を行い，「高」「中」「低」の 3 段階評価でユーザに提示する．3 段階評価の閾値は，評価実験で利用した Web サイトの評価値のばらつきによって決定している．坂松らの研究では安全性評価値を単純加算によって求めているが，算出手法の妥当性については検討されていない．また，評価値は選択した評価ツールの特徴や検査結果に影響することが考えられる．

2.3 オンデマンド検査により提示する手法

國分らは，SNS の悪性 URL における分析と防御に関する研究[13]の中で，短縮 URL の最終アクセス先サイトをオンデマンド検査しユーザに提示する手法を提案している．安全性の検査には，Google Safe Browsing などのオンラインで検査結果が得られる複数の評価サービスを用いる．國分らの研究において，実際にシステムで利用する評価ツール群やユーザが必要とする情報の選択については言及されておらず，今後議論するべき課題としてあげられているのみにとどまっている．

3. 安全性評価結果の統合手法

本研究では，短縮 URL により不明瞭になった行先サイトを複数の評価サービス（本稿では 66 件）を用いて検査し，その検査結果をユーザに提示する手法を提案する．

本提案手法の全体構成を図 1 に示す．Web サイトの安全性評価は事前準備および手順 1～3 で構成される．以下よりそれぞれの手順について説明する．

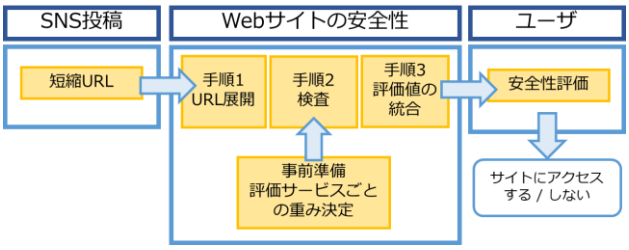

図 1　提案手法全体構成図
Figure 1 System Configuration

事前準備：評価サービスごとの重み決定

事前準備として，66 件の各評価サービスに適用する重みを決定する．各サービスの重みを決定する理由は，それぞれの評価サービスによる評価結果の偏りを考慮するためである．重みは，あらかじめ用意したデータセットに対する各評価サービスの検査結果をもとに統計手法を適用することで決定される．重み決定に利用する統計手法については 4.1 節で説明する．

手順 1：URL 展開

SNS 投稿に含まれる短縮 URL を対象として URL を展開し，行先サイトの URL を取得する．対象の URL が多重に短縮されている場合は，複数回展開する．

ⓒ 2018 Information Processing Society of Japan

手順 2：検査

安全性判断には，事前準備と同一の 66 件の安全性評価サービスを利用する．具体的には，ESET[14]や Yandex Safe Browsing[15]などである．行先サイトの URL を対象として 66 件の評価サービスにより検査する．なお，本提案では，オンラインで検査結果が得られる評価サービスを利用する．

検査対象となる URL は評価サービスにより悪性サイトの疑いがないかを検査する．良性と判断した場合は 1，悪性と判断した場合は 0 として評価サービスごとに検査結果を収集する．事前準備であらかじめ決定した重みを各検査結果に適用し，最終的な検査結果とする．

手順 3：評価値の統合

手順 2 で得たそれぞれの評価サービスの最終的な評価結果を合算した値を「安全」,「危険」のいずれかで最終的な評価結果としてユーザに提示する．

4. 評価実験

評価実験では，良性および悪性サイトのデータベースおよび，Twitter から収集した短縮 URL に対し，前章で述べた提案手法を適用し，安全性を正しく判定できるかを検証する．また，前章の事前準備処理における重み決定について，利用する分類器による精度の違いについて検討する．なお本稿では，正規サイトを良性サイト，フィッシングサイトやマルウェア配布サイトのようなユーザが当該サイトにアクセスした際に不利益を被るサイトを悪性サイトと定義する．

図 2 に本稿で実施する評価実験の概要を示す．全ての実験で使用する訓練用データと分類器は同じである．評価データについては，実験 1 では良性および悪性サイトのデータベースをもとにした評価用データ，実験 2 では実運用を想定し，Twitter の投稿文中の短縮 URL をもとにした評価用データを用いる．なお，実験 2 において，短縮 URL 群の抽出方法により評価用データの構成が異なるため，実験 2-1，実験 2-2 と分けて取り扱う．

図 2　評価実験の概要

Figure 2 Outline of Experiments

4.1 重み決定に利用する統計手法

3 章で述べたように事前準備として，訓練用データとして良性および悪性サイトデータベースから収集した URL（4.2.1 項にて説明）に対し，66 件の評価サービスによる検査を実施し，その結果をもとに統計手法を適用し，各評価サービスの重みを決定する．本稿では，適用する統計手法として線形回帰，SVM，ロジスティック回帰，ナイーブベイズ, ランダムフォレスト, 多層パーセプトロンを用いる．決定した重みは，実験 1 および実験 2 ともに同じ重みを利用する．なお，実装には開発言語として Python3.6.3 を利用した．各種分類器は機械学習ライブラリを利用し，線形回帰のみ TensorFlow[20]，その他のアルゴリズムは scikit-learn[21]を利用した．

4.2 実験 1：良性および悪性データベース URL の検査

実験 1 では，提案手法により良性サイトと悪性サイトをどの程度正しく判定できるか，重み決定に利用する分類器の違いによる判定結果の違いの 2 つを検証する. そのため，重み係数を適用しない場合（以下，重み未適用）と前節で述べた各種分類器ごとの重み係数を適用した場合について良性サイトか悪性サイトかを判定する．

4.2.1 訓練用/評価用データセットの作成

実験 1 および実験 2 で使用する訓練用データセットと，実験 1 で使用する評価用データセットは，良性および悪性サイトデータベースから収集した URL を利用し，作成した．良性サイトは ALEXA[16], dmoztool.com[17]より 250 件ずつ，悪性サイトは Malware Domain List[18], Phishtank[19]より 250 件ずつ収集し，合計 1,000 件の URL を取得した．さらに 1000 件の URL をランダムに並べ替え，700 件を訓練用，300 件を評価用に分割し，それぞれのデータセットとした．各データセットの良性サイトと悪性サイトの内訳を表 1 に示す．上位 700 件を訓練用，残り 300 件を評価用としたため，良性サイトおよび悪性サイトの件数比率は異なる．

表 1　各データセット内訳

Table 1 The details of Datasets

	良性サイト	悪性サイト	合計
訓練用データ	343	357	700
評価用データ	157	143	300

4.3 実験 2：SNS 投稿から収集した URL の検査

実用に向けた評価として Twitter の投稿文に含まれる短縮 URL を評価用データセットとし，これに対し，提案手法を適用し，実験 1 との判定結果違いと，重み決定に利用する分類器の違いによる判定結果の違いの 2 つを確認する．

ⓒ 2018 Information Processing Society of Japan

4.3.1 本実験の評価用データセットの作成

対象となる URL は Twitter Streaming API を利用し, Twitter の全投稿のうち 1%をランダムに収集し，投稿に含まれる短縮 URL を抽出した．収集した短縮 URL は行先サイトの URL に展開し，重複する URL を取り除いた 2,347 件である．また，全ての URL に良性か悪性かの正解ラベルを付与するため全 2,347 件の URL はあらかじめ Google Safe Browsing により検査した．その結果，2,347 件中，良性サイトは 2,334 件，悪性サイトは 13 件となり，この結果を正解ラベルとした．なお，正解ラベルの作成に Google Safe Browsing を用いた理由は, 本評価サービスは同社の Web ブラウザ Google Chrome だけでなく Firefox や Safari といった Web ブラウザに搭載され，利用デバイスを問わず多くのユーザに利用されているためである．

さらに，本実験で使用する評価用データセットとして，全 2,347 件の URL から実験 1 と同様に 300 件を抽出した．この際，全 2,347 件をランダムに並び替え，そのうちの 300 件を抽出した評価用データセット（実験 2-1），全 2,347 件のうち悪性ラベルが付与された全 13 件と，良性ラベルが付与された 2,334 件をランダムに並び替え，287 件を抽出した評価用データセット（実験 2-2）の 2 種類を用意した．それぞれの評価用データセットの良性と悪性の内訳を表 2 に示す．

表 2　実験 2 のデータセット内訳
Table 2 The details of Evaluation Datasets in Experiment 2

	良性サイト	悪性サイト	合計
実験2-1	296	4	300
実験2-2	287	13	300

5. 結果と考察

本章では, 実験 1 と実験 2 の結果について述べる. なお, 実験 2 については，評価用データベースの作成方法により実験 2-1，実験 2-2 と 2 つに分けて結果を述べる．

それぞれの実験結果は, 4.1 節で説明した訓練データを用いた重み付けの際に利用した 6 種類の分類器ごとに正解率, 適合率，再現率を算出し，実験 1 のみ重みを適用しない場合（重み未適用）の正解率，適合率，再現率も算出した．

全評価結果の正解率，適合率，再現率の算出方法は以下である．小数点第 2 位未満切り捨てとする．

正解率：(正しく悪性と判断した件数+正しく良性と判断した件数) / 評価用データ件数

適合率：正しく良性と判断した件数 /(正しく良性と判断した件数+誤って良性と判断した件数)

再現率：正しく良性と判断した件数 /(正しく良性と判断した件数+誤って悪性と判断した件数)

5.1 実験 1：良性および悪性データベース URL 評価結果

実験 1 の評価結果を表 3 に示す．正解率は，各種分類器による結果が重み未適用の場合を全て上回る結果となり，URL の検査結果に対して各種分類器による統計手法を適用することは有効であると考えられる．また，多層パーセプトロンの 99.66%が最も高い正解率であり, 適合率に関しても 98.68%と最も高い結果となった. 本実験での全体の正解率が 90%以上になった理由として, 2 つ考えられる．1 つ目の理由は, 重み未適用の場合の正解率が高いことである．各評価サービスの検査結果がある程度一致しており，評価値の統合をする以前に，まとまった結果が得られた可能性がある．2 つ目の理由は，良性・悪性サイトのデータベースからデータセットを作成したため，データセットに含まれる URL の生存期間が比較的長いことである．それにより，URL が掲載されてから本実験のため URL を収集するまでにある程度の期間があったと考えられ，各評価サービスによる検査方法の変更・検査結果の更新が行われた可能性が考えられる．

表 3　実験 1 の評価結果
Table 3 Evaluation results of Experiment 1

分類器	正解率(%)	適合率(%)	再現率(%)
重み未適用	94.33	97.20	91.44
線形回帰	95.33	93.45	100.00
SVM	98.66	95.54	100.00
ロジスティック回帰	96.33	93.16	100.00
ナイーブベイズ	97.66	95.54	100.00
ランダムフォレスト	99.00	98.03	100.00
多層パーセプトロン	99.66	98.68	100.00

5.2 実験 2：SNS 投稿から収集した URL の評価実験

実験 2 では前述の通り，評価用データセットの作成方法により実験 2-1（評価用データセットのもととなる URL 群の全 2,347 件のうち良性・悪性問わずランダムに 300 件抽出），と実験 2-2（評価用データセットのもととなる URL 群全 2,347 件中の悪性全 13 件と良性サイト 287 件をランダムに抽出）に分けて述べる．

実験 2-1 の評価結果を表 4 に示す．結果は，線形回帰が 95.00%と最も高い正解率となり，次いでロジスティック回帰の 94.33%となった. 適合率に関しては, 線形回帰, SVM, ロジスティック回帰が 99.29%の最も高い結果となった．

ⓒ 2018 Information Processing Society of Japan

表 4　実験 2-1 の評価結果
Table 4 Evaluation results of Experiment 2-1

分類器	正解率(%)	適合率(%)	再現率(%)
線形回帰	95.00	99.29	95.60
SVM	87.33	99.29	87.83
ロジスティック回帰	94.33	99.29	94.93
ナイーブベイズ	92.00	99.27	92.25
ランダムフォレスト	87.33	99.22	88.85
多層パーセプトロン	83.00	99.19	81.75

実験 1 と実験 2-1 を比較して，線形回帰を除いた 5 つの分類器は正解率が低くなる傾向がみられ，適合率に関してはロジスティック回帰以外の分類器で高くなった．実験 2-1 で用いた評価用データセットのうち 296 件が良性サイトと偏りがあったため，適合率が高くなった可能性が考えられる．

正解率の高い線形回帰とロジスティック回帰について，混同行列を用いた評価結果を表 5 および表 6 に示す．線形回帰とロジスティック回帰のどちらも悪性を誤って良性と判断した偽陽性(False Positive，以下 FP)よりも，良性を誤って悪性と判断した偽陰性(False Negative，以下 FN)が多いことがわかる．線形回帰およびロジスティック回帰で FP に分類された 2 件の URL は同一のものであった．また，2 件の URL の検査結果も同一の評価サービスが悪性と判断していた．こういった FP に分類される URL を削減するために，利用する評価サービスの取捨が対策として考えられる．評価サービスの選定を行うことで，分類器による精度の向上をはかることが可能であると考えられる．

表 5　実験 2-1 の評価結果(線形回帰)
Table 5 Classification results of Experiment 2-1 (Linear Regression)

	予測 Negative	予測 Positive
実際 Negative	2	2
実際 Positive	13	283

表 6　実験 2-1 の評価結果(ロジスティック回帰)
Table 6 Classification results of Experiment 2-1 (Logistic Regression)

	予測 Negative	予測 Positive
実際 Negative	2	2
実際 Positive	15	281

実験 2-2 の結果を表 7 に示す．実験 2-1 と同様に，実験 1 と比較すると正解率は低いが，適合率が高くなった．悪性サイトの件数は 13 件であり，実験 1 より増加したが，正解率に大きな差は見られなかった．正解率は線形回帰が最も高く，98.66%となった．そのうち，線形回帰について，混同行列を用いた評価結果を表 8 に示す．実験 1 と比較して，データセットの違いによる大きな正解率の差は見られなかった．

表 7　実験 2-2 の評価結果
Table 7 Evaluation results of Experiment 2-2

分類器	正解率(%)	適合率(%)	再現率(%)
線形回帰	98.66	99.29	99.30
SVM	84.00	98.76	84.21
ロジスティック回帰	92.66	97.47	94.73
ナイーブベイズ	89.00	99.21	89.12
ランダムフォレスト	84.33	98.77	78.51
多層パーセプトロン	82.66	98.74	85.26

表 8　実験 2-2 の評価結果(線形回帰)
Table 8 Classification results of Experiment 2-2 (Linear Regression)

	予測 Negative	予測 Positive
実際 Negative	11	2
実際 Positive	2	285

ⓒ 2018 Information Processing Society of Japan

5.3 実験全体の考察

実験で明らかにする点の 1 つである，提案手法により良性サイトと悪性サイトの判定精度については，良性および悪性サイトのデータベースから作成した評価用データセットに対する評価（実験 1），Twitter の投稿文から抽出した短縮 URL 群から作成した評価用データセットに対する評価（実験 2）のいずれの場合も，正解率で最も結果の良かった分類器を用いた場合, 95%前後の結果を示した. これは，提案手法が，サイトの安全性評価に有効であることを示すといえる.

また，提案手法の事前準備処理である全評価サービスの重み決定時に利用する分類器による精度の違いについては, 良性および悪性サイトのデータベースから作成した評価用データセットに対する評価（実験 1）では，多層パーセプトロンが, 実運用を想定した Twitter を用いた評価（実験 2）では, 線形回帰が最も高い正解率, 適合率を示した. 特に，線形回帰の実験結果に関して，実験 1 では他の分類器よりも正解率は低くなったが，3 つの実験全てで 94%以上の正解率を示しており，訓練用データと評価用データで種類が異なる場合でも，ある程度の正解率が見込めると考えられる．また，今回の実験では正解率が低かった他の分類器についても，学習回数や閾値の調整など，より詳細な設定によって精度向上の余地が十分にあると考えられる.

さらに，本研究における安全性評価の結果はユーザの判断に直結する可能性が考えられるため FP を重視するべきである. ユーザが実際に利用する場合, FN が頻出する評価結果では，使用感を損ねる可能性も考えられる．そのため決定閾値や提示方法を考慮する必要がある.

データセットの作成に関しては，実験 1 で収集したサンプル URL は海外のサイトから，実験 2 は Twitter の日本語のメッセージ投稿に含まれる短縮 URL を対象とし収集した. そのため, 実験 1 では英語で記述されたサイトが多く，実験 2 では日本語で記述されたサイトが多く見られた．本稿の実験で使用した訓練用データには，海外のサイトを多く含まれていたが，日本語で記述されたサイトの URL を訓練用データセットに用いることで精度が向上する可能性がある.

6. まとめ

本稿では，Twitter や Facebook など SNS 投稿メッセージに含まれる短縮 URL に対し，複数の安全性評価サービスを用いて安全性を検査し，各評価結果の偏りをなくすため評価サービスごとに重みを決定し，結果を統合する手法を提案した．また，提案手法における重み決定について複数の分類器を利用し，2 種類の評価用データセットに提案手法に適用し，安全性判定精度および，各分類器の違いによる安全性判定精度の違いを評価・検討した.

提案手法によるサイトの安全性判定精度については，良性および悪性サイトのデータベースから作成した評価用データセット，および，Twitter の投稿文から抽出した短縮 URL 群から作成した評価用データセットに提案手法を適用したいずれの場合ももっとも高い正解率が約 95%以上であった．この事から，提案手法は，サイトの安全性評価に有効であるといえる.

使用する分類器による精度の違いについては，実験 1 と実験 2 で高い結果を示す分類器は異なっていたが，いずれの実験でも線形回帰を利用した場合，94%以上の正解率を示しており，現状，本提案手法に対しては安定して利用できる分類器であると考えられる.

本研究において，以下 3 つが今後の課題である．まず 1 点目に, 悪性 URL の収集方法である. 本実験では多くの分類器で 80%～90%以上の精度となったが，悪性サイトの件数が極端に少ないことが要因として考えられる．良性および悪性サイトの件数比率を固定したデータセットの作成，他種の SNS や掲示板サイトから URL を収集などサンプル URL の収集方法を検討することで各分類器を再評価する必要がある．2 点目に，さらなる精度の向上である．本研究では 66 件評価サービスと 6 種類の分類器により URL の安全性を評価したが，使用する評価サービスの選定と分類器の学習回数や統計手法の再検討，アンサンブル学習などの方法を検討することによって，安全性評価の精度がさらに向上することが可能であると考えられる. 3 点目に, URL とそれに付随する情報との関係性の検討である．SNS 投稿には, URL の他に添付画像や投稿メッセージが含まれている場合が多いため，間接的に URL に関する情報が得られる可能性がある. 特に FP や FN に該当する URL について, URL が含まれていた投稿内容との関係について考察し, その特徴の抽出方法について検討することで, URL の安全性判断に関する新たなアプローチの可能性があると考える.

謝辞

本研究は JSPS 科研費 16K01025 の助成を受けた.

参考文献

[1] Twitter (online), available from <https://twitter.com/> (accessed 2017-12-24).

[2] Facebook (online), available from <https://www.facebook.com/> (accessed 2017-12-24).

[3] Google URL Shortener (online), available from <https://goo.gl/> (accessed 2017-02-01).

[4] Bitly | URL Shortener and Link Management Platform (online), available from <https://bitly.com/> (accessed 2017-04-30).

[5] Anatomy Of A Social Media Attack (online), available from <https://www.darkreading.com/analytics/anatomy-of-a-social-media-attack/a/d-id/1326680> (accessed 2018-02-

ⓒ 2018 Information Processing Society of Japan

17).

[6] Trojanized Propaganda App Uses Twitter to Infect, Spy on Terrorist Sympathizers | McAfee Blogs (online), available from <https://securingtomorrow.mcafee.com/mcafee-labs/trojanized-propaganda-app-uses-twitter-to-infect-spy-on-terrorist-sympathizers/> (accessed 2017-05-28).

[7] Apple Credentials | McAfee Blogs (online), available from <https://securingtomorrow.mcafee.com/mcafee-labs/active-ios-smishing-campaign-stealing-apple-credentials/> (accessed 2017-05-28).

[8] スパマー自身が URL 短縮サービスを運営し、スパムに悪用する例も ー シマンテックレポート | マイナビニュース, 入手先(オンライン) <https://news.mynavi.jp/article/20111128-symantec10/> (参照 2018-02-21).

[9] セーフ ブラウジング - 透明性レポート - Google , 入手先(オンライン) <https://www.google.com/transparencyreport/safebrowsing/?hl=ja> (参照 2017-04-30).

[10] 坂松春香, 小倉加奈代, ベッドバハドゥールビスタ, 高田豊雄: サイトの安全性と重要度に応じたパスワード管理ツールに関する研究, 2016 年暗号と情報セキュリティシンポジウム(SCIS2016）, 1F1-3, 2016.

[11] VirusTotal - Free Online Virus Malware and URL Scanner (online), available from <https://www.virustotal.com/en/> (accessed 2017-07-02).

[12] The Spamhaus Project (online), available from <https://www.spamhaus.org/> (accessed 2017-09-20).

[13] 國分佑太朗, 中村章人: SNS における悪性 URL の分析と防御, 社会情報学会(SSI)2017, 入手先(オンライン) <http://gmshattori.komazawa-u.ac.jp/ssi2017/wpcontent/uploads/2017/03/30.pdf> (参照 2017-12-24).

[14] セキュリティソフト「ESET」 | 検出率・軽さ・満足度 No.1 ウィルス対策ソフト 入手先(オンライン) <https://www.eset-smart-security.jp/> (参照 2018-02-02).

[15] Yandex (online), available from <https://www.yandex.com/> (accessed 2017-08-15).

[16] Top Sites in Japan - Alexa (online), available from <https://www.alexa.com/topsites/countries/JP> (accessed 2017-12-15).

[17] The Directory of the Web (online), available from <http://dmoztools.net/> (accessed 2017-12-15).

[18] MDL (online), available from <https://www.malwaredomainlist.com/> (accessed 2017-12-15).

[19] PhishTank | Join the fight against phishing (online), available from <https://www.phishtank.com/> (accessed 2017-09-20).

[20] TensorFlow (online), available from <https://www.tensorflow.org/> (accessed 2018-08-17).

[21] scikit-learn: machine learning in Python (online), available from <http://scikit-learn.org/stable/> (accessed 2018-08-17).

ⓒ 2018 Information Processing Society of Japan