2E-02

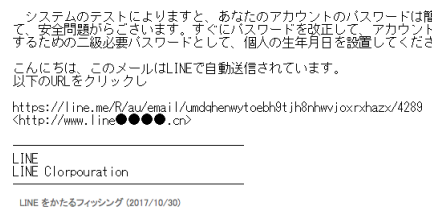
ドメインの特徴によるフィッシング詐欺サイト探索の可能性

松田　健[1], 加藤雅彦[1], 唐沢勇輔[2], 丹京真一[3], 中村智史[4], 林　憲明[5], 加藤孝浩[6]

[1] 長崎県立大学, [2] ソースネクスト株式会社, [3] 日立システムズ, [4]LINE 株式会社,
[5] トレンドマイクロ株式会社, [6] トッパン・フォームズ株式会社

1　はじめに

金融機関や有名企業が提供しているウェブサービスを装うことで，ユーザーの個人情報を詐取するフィッシング行為はますますエスカレートしている．フィッシングの具体的な手口は，ユーザーに本物と区別がつかない程巧妙に作成されたメールを送りつけてフィッシングサイトにユーザーを誘導し，犯罪者の目的となる情報をユーザーに入力させるというものである．図はフィッシング対策協議会が公開している，実際のフィッシングメール [1] である．


システムのテストによりますと、あなたのアカウントのパスワードは簡単しすぎて、安全問題がございます。すぐにパスワードを改正して、アカウントをログインするための二級必要パスワードとして、個人の生年月日を設置してください

こんにちは、このメールはLINEで自動送信されています。
以下のURLをクリックし

https://line.me/R/au/email/umdqhenwytoebh9tjh8nhwvjoxrxhazx/4289
<http://www.line●●●●.cn>

LINE
LINE Clorpouration

LINE をかたるフィッシング (2017/10/30)

図 1: フィッシングメールの例

フィッシング対策協議会は，これまでにコンテンツベースによるフィッシング検知方法を提案するなど，フィッシング撲滅のための活動を行っている [2]．現状では，図 1 のようなフィシングメールが出回ってからフィッシングサイトが特定され，そのサイトを停止させるため，この期間の間に実際に被害に会ってしまう可能性があるため，未然にフィッシング被害を防ぐための有効な手法の開発が急がれる．本研究では，実際のフィッシングサイトに使用されたドメイン情報を収集してその特徴について考察し，フィッシングサイトの早期発見に繋げるための情報収集方法について検討する．

2　関連研究

フィッシングに関する最近の研究として，詐欺に騙されないための訓練を実施する手法 [3] が提案されている．文献 [4] では，提示されたウェブサイトを見たときのユーザーの眼球運動を観察することで，認知心理学の知見からフィッシングサイトをユーザーが見分ける訓練のサポートが出来ないかどうか検討がなされている．また過去に観測されたフィッシングサイトの url の情報から，類似性の高い url を探索する手法 [5] についても研究されている．本研究では，過去に観測された様々な種類のフィッシングサイト url から一定の規則が認められるデータを整理し，新しく発生する可能性があると考えられる url について，特にドメイン名に着目することで検討する．

3　ドメイン特徴

実際にフィッシングサイトに使用された url のデータは，いくつかのパターンに着目することで大まかに分類することができる．その中には，ユーザーに本物と信じ込ませるためと思えるような文字列で構成されているものもあれば，ランダムな文字列で構成されているものもあり，url の情報だけでフィッシングサイトを特定することは困難であるように思える．特に，ランダムな文字列で構成されているフィッシングサイトの場合，未知のフィッシング url を被害が出る前に発見することは非常に困難であると考えられる．ある程度規則的な文字列で生成される url である場合は，想定される url の範囲は狭くなるため，ランダムな文字列で生成されるフィッシング詐欺 url よりも特定は容易に思えるが，このような場合においても，実際に未知のフィッシング詐欺 url を未然に発見することは容易ではない．本稿では，ある程度，一定の規則が見られるフィッシング詐欺 url のドメイン部分の情報について紹介し，そのようなドメインにどのような特徴が見られるかということについて報告する．本研究では，line をかたるフィッシング詐欺 url のドメインについて焦点を当て，これまで観測された url のドメインにどのような特徴が見られるかということについて紹介する．line をかたるフィッシングサイトは非常に多く観測されており，2017 年 4 月 1 日から 12 月 22 日までの間に 160 個以上のフィッシングサイトが観測されている．その中には，line**.cn のように，line という文字列を含むフィッシングサイトは 120 個以上含まれており，これは非常に大きな特徴であると言える．ドメインの*の部分には

[1] Takeshi MATSUDA, Masahiko KATOH, Yusuke KARASAWA, Shinichi TANKYO, Tomofumi NAKAMURA, Noriaki HAYASHI, Takahiro KATOH
[1] University of Nagasaki, [2] SOURCENEXT CORPORATION, [3] Hitachi Systems, Ltd
[4] LINE Corporation, [5] Trend Micro Incorporated, [6] Toppan Forms co. ltd

アルファベットが入るため，その組み合わせは 676 通り存在することになる．そのため，これら 676 のドメイン情報を用いて監視を行えば，比較的簡単にフィッシングサイトを発見できるのではないかと考えられる．しかしながら，実際に調査を進めると，アクセスする端末に依存してウェブページの見え方が異なったりするため，現状では，未然にフィッシングサイトを発見することは困難である．しかしながら，多様な考察の末に有益な情報が見出せる可能性を否定することはできないため，現状のデータを整理して必要な情報と不要な情報の整理をすることも重要であると考えられる．図 2 は line**.cn における名前空間のマトリックスであり，色がついている部分はフィッシングサイトとして使用された実績があるドメインに対応する．左端の文字列が 1 文字目，上段の文字列が 2 文字目に対応している．ここから読み取れるフィッシングサイトの特徴について，紙面の都合上 1 つだけ考察を行う．図 2 のマトリックスは 2017 年 12 月 22 日までのデータを用いて生成したものであるが，line**o*の部分の 1 文字目として最も多く使用されているものはアルファベットの i であり，2 文字目として最も多く使用されているものは e であることが分かる．これにより，完全にランダ

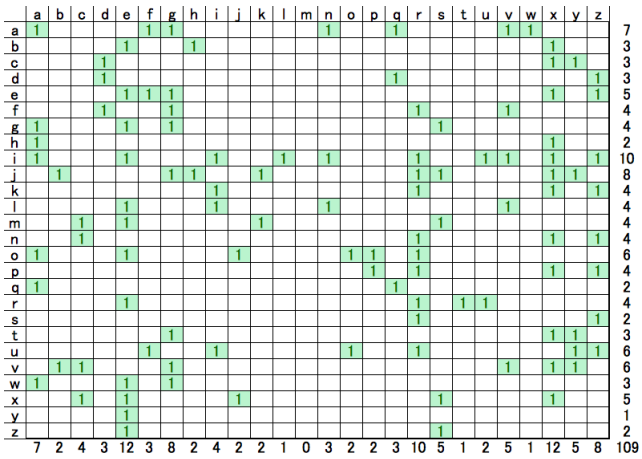


図 2: フィッシング詐欺 url に含まれるドメイン特徴

ムにドメイン名が生成されている訳でないことは確認できるが，しかし一方ではっきりとしたドメイン名の生成ルールも存在しないことが分かる．また，図 3 のように，2017 年 11 月と 12 月にそれぞれ生成したマトリックスを比較すると，ある程度固まった場所で新たなフィッシングドメインが観測されていることが分かる．これらの情報は直接的にフィッシングサイトの発見に役に立つとは限らないが，その経過を観測することで有益な情報が見出される可能性があることについて否定することもまた困難である．

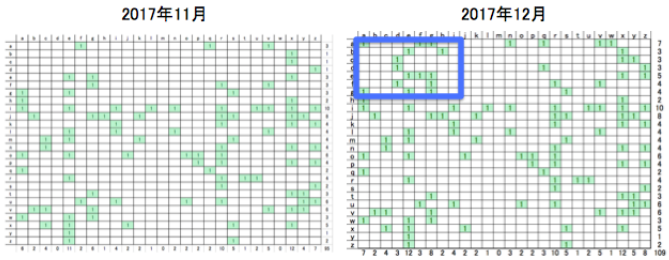


図 3: ドメイン特徴の時系列的変化

4 今後の課題

被害が出る前に先回りしてフィシングサイトを発見することは依然として困難なままであるが，今後も多角的に情報収集を続け，自動的にフィッシングサイトを発見するための方法について検討する．

参考文献

[1] フィッシング対策協議会, “LINE をかたるフィッシング (2017/10/30),” https://www.antiphishing.jp/news/alert/line_20171030.html (2017 年 12 月 29 日確認)

[2] フィッシング対策協議会, “コンテンツベースフィッシング検知手法の実用化に向けた評価と改良” https://www.antiphishing.jp/report/other/content_based_phishing_detection_2011.html (2017 年 12 月 29 日確認)

[3] 東野 正幸，川戸 聡也，大森 幹之，川村 尚生, “仮名化による個人情報の保護に配慮したパブリッククラウド型フィッシングメール対応訓練システムの開発 (マルチメディア情報ハイディング・エンリッチメント),” 電子情報通信学会技術研究報告:信学技報 117(128), pp.231-234 (2017)

[4] 宮本 大輔, “認知心理学に基づいたサイバーセキュリティ研究に関する一考察,” 日本認知心理学会発表論文集, p.33 (2016)

[5] 孫 博，秋山 満昭，八木 毅，森 達哉, “既知の悪性 URL 群と類似した特徴を持つ URL の検索,” コンピュータセキュリティシンポジウム 2014 論文集 2014, pp.1-8, (2014)