2016年度 SWAN Working Group 活動報告

宮本大輔 (daisu-mi@nc.u-tokyo.ac.jp)　　門林雄基 (youki-k@is.naist.jp)

2016年12月15日

目 次

1　はじめに

SWAN (Security for Web 2.0 Application) WG では，悪意あるウェブサイトの動向を観測し検討している．ウェブを介した攻撃にはその攻撃空間が広いという特徴があり，本研究グループはその広い特性に対応した研究を行なっている．これまでの活動としては，エンドユーザの認知能力に合わせたフィッシングサイト解析や脆弱性を持つウェブ2.0のアプリケーションを WIDE メンバーに提供する試み，悪意あるウェブサイトによく見られる難読化された JavaScript の構造に着目した解析，PC だけではなく Android などで動くマルウェアの解析技術のハンズオンなどが挙げられる．

今年度は，認知心理学に着目したサイバー防御システムの開発と評価を行った．この研究領域には様々なチャレンジがあるが，その1つに「人間の観測される情報から，人間の精神状態を推測する」という課題がある．例えば人間の精神的負荷の高まりを眼球運動や脳波，血圧，呼吸，顔表面温度などから分析する研究が行われている．ストレスの増加など精神的負荷の高まりは過失を引き起こしやすいとされており，このような課題は人的要因の解決を行っているとも考えられる．さらに，近年の IoT 技術の普及に伴い，人間から健康に関するデータを取得する様々な生体センサーの技術開発が進んでいる．この情報を認知心理学の観点から活用すれば，情報機器を利用するユーザが今まさに何を考えているのか，どのような情報を入手し，どのような根拠で，どのような意思決定をしようとしているのかを推測できると思われる．

これまでのサイバーセキュリティ対策では，エンドユーザがどのような状態あるかは考えていないが，仮に「まさに騙されそうな状態にあるユーザ」がわかるのであればどうか．普段は，正常なプログラムをマルウェアであると，正規サイトをフィッシングサイトであるといった誤検知を減らすべく工夫されているセキュリティ技術や製品も，この時ばかりは誤検知を承知の上で注意喚起ができるのではないか．このような工夫によって，人的要因によって起こるインシデントは減らせるのではないか．

SWAN WG ではこの観点から研究を行い，以下に報告する活動を行った．

2　フィッシングサイトと視線の相関分析

視線分析の研究は，1880年頃に Javal らが文章を読む際の視線移動を分析したことが最初の事例とされる．Javal らは角膜系を開発して観測を行い，視線は継続的は移動するのではなく，素早い動きと滞留を繰り返すことを発見した．この素早い動きは「サッカード」，滞留は「注視」と呼ばれている．視線分析技術の進化とともに，様々な分野への応用が進められている．従来は研究実験や概念実証のような利用が多くみられてい

たが，視線分析装置の小型化，高精度化，そして低価格化によりより広い範囲の分野に普及しつつある．近年，Lee らの研究グループにより注視の時間と回数からユーザの意図を読み取る研究が行われている．ユーザの意図は何か特別な意図をもって情報を調べている Information 型と，漠然と眺めている Navigation 型に分類され，主に注視の回数と時間からどの意図で分析しているかを識別する．

SWAN WG では，エンドユーザが URL や SSL の鍵アイコンが表示されるアドレスバーを閲覧することの有効性の評価を行った [2]．エンドユーザの視線情報を収集するため，2013 年 11 月から 2014 年 2 月までの期間，東京大学の構内掲示板にて被験者を募集した．実験参加者は 23 人であり，そのうち 20 人が男性，3 人が女性であった．また，22 人が 20 代であり，残りの 1 人が 30 代であった．延べ 331 回のアドレスバーを目視した回数のうち，誤判定があったのは 89 回であった．フィッシングサイトに限定していえば 200 回のアドレスバーを目視した回数のうち 61 回が誤判定であり，残りの 131 回の正規サイトにおいてアドレスバーを目視した場合の誤判定は 28 回であった．従って，エラー率，False Positive 率，False Negative 率はそれぞれ 26.9%，21.4%，30.5%となる．反対にアドレスバーを見ない場合は，41.1%（129 回中 53 回），18.9%（53 回中 10 回），56.6%（76 回中 43 回）であった．False Positive 率ではごくわずかにアドレスバーを見ない場合が低くなっているが，フィッシングサイトのコンテンツは正規サイトと見た目が区別しにくいため，アドレスバーを見ない限り False Positive 率が高くなっている．結論として，アドレスバーを見ることは効果的であると考えられる．

また，Lee らの研究成果を用い，Internet Explorer に表示されるウェブサイトの真贋判定における意図抽出を試みた [3]．この実験の概略を図 1 に示す．

まず，関心領域（Area of Interest, AoI) を図 2 のように定義し，その画面における注視時間と回数を調べ，Support Vector Machines によるパターン認識を試み，その画面を閲覧する意図が Information 型か Navigation 型かの推測を行った．この結果，アドレスバーを見たユーザが URL に基づいた判定を行ったかをアンケートで調べた所，およそ 85.5% の確率で予想できることがわかった．また，アドレスバーでの目視を確認するだけで 75.9% の確率でこのユーザが正しく判定できるか否かを予想できることがわかった．詳しくは文献 [3] を参照されたい．

3 視線に基づいたサイバー防御を行うシステムの研究開発

この解析で得られた知見を元に，エンドユーザがどの程度危険を認識しているかを判別することは可能と考え，また，その認識の程度によってサイバー防御を切り分けることを考えた．

SWAN WG は文献 [1] において，様々なヒューリスティクスを用いて調査し，その結果を Random Forest により組み合わせてフィッシングサイトか否かを識別するシステムを実装した．ヒューリスティクスとはフィッシングサイトらしさを計算するためにもちいられる情報であり，例として「ドットの数」という経験則が知られている．フィッシングサイトの FQDN に含まれる数のドットは正規サイトに比べ多くなる傾向にある．このようなヒューリスティクスを機械学習により組み合わせることによりパターン認識の精度が高くなる一方で，識別にかかる時間は増加する．一方で，ドットの数だけに頼るアルゴリズムは機械学習によるアルゴリズムと比較すると，識別の精度は低いものの，識別に必要な時間は短い．また，正規サイトをフィッシングサイトであると間違えて判定する可能性が極端に低い．検知の目安である適合率と再現率は，機械学習による方式ではそれぞれ 0.86，0.76 であったのに対し，ドットの数による方式は 1.00, 0.06 であった [4]．

そこで，一定時間以内にアドレスバーを閲覧したエンドユーザは正しく判定できる可能性が高いと考え，明白にフィッシングサイトと思われるもののみを除外することが可能な「ドットの数による方式」によるフィッシングサイト対策を提供する．反対に，アドレスバーを閲覧しなかったエンドユーザは誤った判定を行う可能性が高いと考え,「機械学習による方式」を適用する．また，フィッシングサイトと判定された場合，防御システムはページの閲覧を完全にブロックするのではなく，文献 [2] に述べられていた，個人情報を入力可能なコンテンツを無効化する方式を採用した．

防御システムのアーキテクチャを図 3 に示す．システムは，視線追跡カメラを制御することによりエンドユーザの注視箇所を計測する Eye-Tracking モジュー

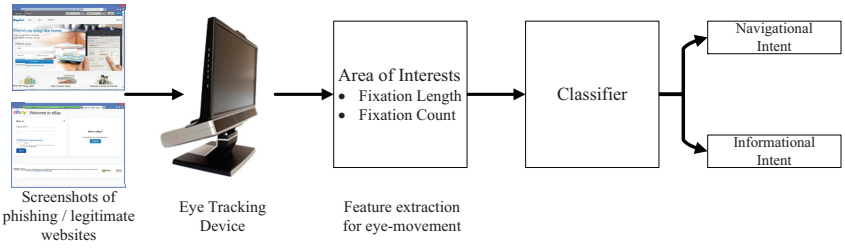
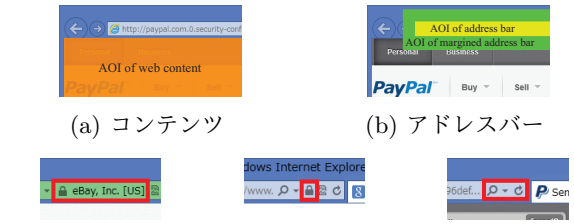


図 1: 実験のブロックダイアグラム図



(a) コンテンツ　(b) アドレスバー

(c) EV-SSL 証明書　(d) SSL 証明書　(e) 証明書がない場合の領域

図 2: 関心領域の設定

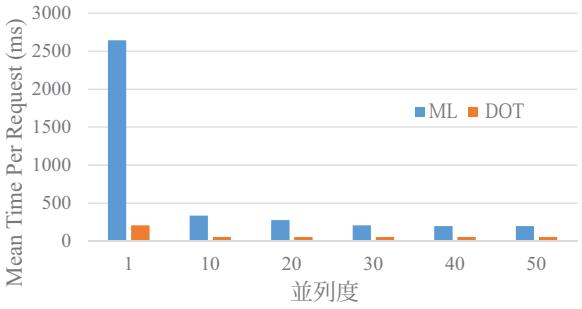


図 4: 防御システムの応答速度

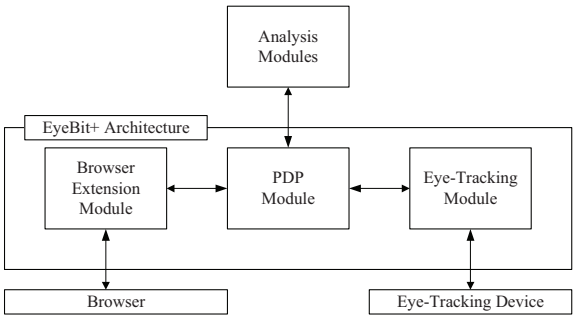


図 3: 防御システムのアーキテクチャ

ルと，この結果からどの検知アルゴリズムを使うかを決定する Policy Decision Point (PDP) モジュール，そして実際のサイバー防御を行うためにエンドユーザが閲覧しているウェブサイトのコンテンツを無効化する Browser Extention モジュールによって構築した．Browser Extention モジュールは Google Chrome ウェブブラウザの拡張機能として実装し，Eye-Tracing モジュールは視線追跡カメラ EyeTribe に用意されている SDK を用いて開発した．また，Analysis モジュールとして実装された機械学習及びドットの数による検知アルゴリズムは，PDP モジュールとはサイバー脅威交換フォーマットである n6 [1]を用いて情報交換を行った．具体的には，PDP モジュールが検知依頼を送信し，Analysis モジュールが検知結果を返信した．なお，実装はオープンソースで公開 [2]で公開されている．

2015 年 12 月から 2016 年 3 月までの期間，東京大学の構内掲示板にて被験者を募集した．応募があった被験者には，実験の目的として，「セキュリティ技術の研究開発を目的としたウェブサイトを閲覧した際のエンドユーザの挙動の観測」であることを説明し，作業内容として「ウェブサイトの画面を閲覧してもらい，正規のサイトか，あるいは偽サイトかを判定していただきます．またその際に　判定基準をアンケート形式でお答えいただきます」と説明した．この実験に関する個人に属する情報として，性別，年代（10 代，20 代，30 代……60 代以上），ウェブサイトを見た際の判定結果（正規サイト，偽サイト），視線の動き（判定時の目線の動き），アンケートによる回答を取得し，個人を特定可能な情報は記録しないことを説明した．実験参加者 21 名のうち 13 人が男性，8 人が女性であっ

[1]CERT Polska, n6 -network security incident exchange. Available at `https://n6.cert.pl`

[2]NECOMA - GitHub, Available at: `https://github.com/necoma`

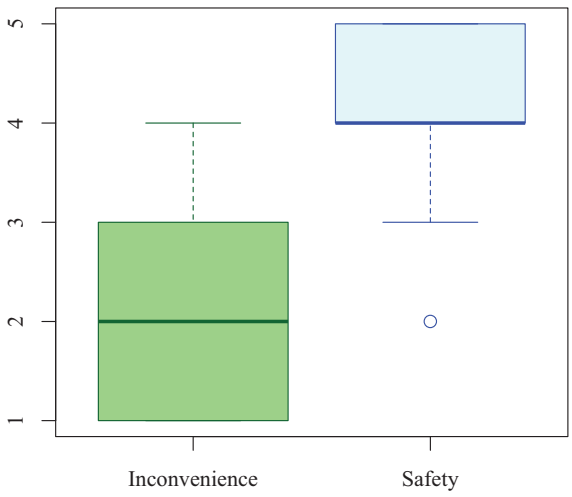


図 5: 防御システムのアンケート調査

た．また，20 代が 13 人，30 代が 6 人，40 代・50 代が 1 人ずつであった．なお実験につかった表示装置はタブレット型の PC である Microsoft Surface 3 Pro であった．

アルゴリズムを実装したシステムの性能としては，1 回あたりのリクエストにかかる時間を評価した．ここでは，HTTPS プロトコルにより n6 でのリクエストを 100 回送信する場合について，並列度を 1,10,20,30,40,50 に変更しながら評価を行った．結果を図 4 に示す．URL に基づいた調査を行うドットによる判定のアルゴリズムの場合は最大で 0.23 秒のオーバヘッドが，機械学習によるアルゴリズムの場合は最大で 2.6 秒程度のオーバヘッドが観測された．

この遅延はフィッシングサイトを誤って判定しそうなユーザに対してのみ行われるものであり，その場合もウェブコンテンツは入力を受け付けないだけで，画像や文言は表示されている．このため，利便性の低下は限定的であると考えられる．この仮説を確認するため，防御システムを利用した被験者に対してアンケートを行い,「不便に思うか」という質問を行い，この解答を「全く思わない（１）」「あまり思わない（２）」「どちらでもない（３）」「ややそう思う（４）」「強くそう思う（５）」の五段階指標により解答させた．図 5 に示す結果の通り,「あまり思わない」が大半であり，残りは「全く思わない」「どちらでもない」がほぼ同じという結果が得られた．一般に，利便性と安全性はトレードオフの関係にるため,「安全だと思うか」という質問も行った．結果は,「ややそう思う」が大半で，ついで「強くそう思う」が多いという結果が観測された．

4　おわりに

SWAN Working Group は悪性ウェブサイト対策技術についての研究活動を行っており，近年は代表的な悪性ウェブサイトであるフィッシングサイトに焦点を当てて研究を行っている．来年度も幅広い種類の悪性ウェブサイトについて，技術的側面にとどまらず，様々な側面からの分析を行い，対策技術の研究開発を行う．

参考文献

[1] Daisuke Miyamoto, Hiroaki Hazeyama, Youki Kadobayashi, *“An Evaluation of Machine Learning-based Methods for Detection of Phishing Sites,”*, *Australian Journal of Intelligent Information Processing Systems*, Vol. 10, No. 2, pp.54-63, November 2008.

[2] Daisuke Miyamoto, Hiroaki Hazeyama, Youki Kadobayashi, *“EyeBit: Eye-Tracking Approach for Enforcing Phishing Prevention Habits,”*, *In Proceedings of the 3rd International Workshop on Building Analysis Datasets and Gathering Experience Returns for Security*, September 2014.

[3] Daisuke Miyamoto, Gregory Blanc, Youki Kadobayashi, *“Eye Can Tell: On the Correlation between Eye Movement and Phishing Identification,”*, *In Proceedings of the 22nd International Conference on Neural Information Processing of the Asia-Pacific Neural Network Assembly*, November 2015.

[4] Adam Kozakiewicz, Romain Fontugne, *“Deliverable D2.1Threat Analysis,”*, *NECOMA Project*, Available at: `http://www.necoma-project.jp/ja/deliverables/necoma-d21`, November 2014.