Computer Security Symposium 2018
22 - 25 October 2018

Rig Exploit Kit を利用した Drive by Download 攻撃における分類評価

野村敬太[†1]　猪俣敦夫[†1]

概要：昨今 Drive-by-Download 攻撃に対する対策として，その攻撃を見分け分類を行う研究は非常に多く行われている．それらの手法では，その Drive-by-Download 攻撃が行われる際に発生するリダイレクトや URL，ドメインなどの情報を用いて Drive-by-Download 攻撃かそうではないか判定を行っている．しかしながら，こういった手法が実際のトラフィックや，新しく発生したトラフィックに対しても有効であるか評価が不十分な場合がある．本研究では，実際のトラフィックに近づけるため多くの利用者がいる SNS サービスである Twitter に注目し，そこで共有されている URL と 2017 年に報告された Drive-By-Download 攻撃のトラフィックデータを用いて，先行研究で示されている特徴がこれらのデータに対して有効に分類できるものであるか評価を行う．

キーワード：Drive-by-Download 攻撃，分類評価

Characteristic evaluation of detection Drive-by-Download Attack using Rig Exploit Kit

Keita Nomura[†]　Atsuo Inomata[‡]

Abstract: Recently, many researchers investigated detecting a Drive-by-Download attack. In these methods, It detects the Drive-by-Download attack using information such as the redirect behaviour, URL, domain. However, these methods of evaluation were not sufficient in real traffics and new Drive-by-Download attack. In this paper, we focus on Twitter that is using many people. These methods of past research evaluate using positive data which is sharing URL on Twitter and negative data which is reporting the Drive-by-Download attack URL.

Keywords: Drive-by-Download Attack, detection

1. はじめに

公開サーバを対象とした攻撃が深刻化している．調査によれば 2018 年 1 月から 3 月に発生したインシデントの半数以上はサーバへの不正侵入である．こういった不正侵入されたサーバは攻撃者の踏み台として利用されるケースが多い．攻撃者の利用に用いられるケースとして，Drive-by-Download 攻撃の踏み台として利用されるケースが存在する．この攻撃は正規の Web サイトの改ざんや広告などを利用し，利用者を攻撃者が用意したサーバへアクセスさせ，数回リダイレクトなどをはさみ，最終的に既存のブラウザやブラウザのプラグインなどの脆弱性を利用しマルウェアに感染させる攻撃である．こういった攻撃は，Exploit Kit と呼ばれる一つのパッケージとして，ダークウェブなどの市場で取引され攻撃への利用が容易となっている．これらの台頭により攻撃者は必ずしも攻撃に必要な知識を持たなくとも攻撃を行うことが可能となり攻撃を行うハードルが著しく下がっている．特に日本では Rig Exploit Kit と呼ばれる Exploit Kit を用いた攻撃が多く報告されている[1].

Rig Exploit Kit は，Drive-by-Download 攻撃で用いられる攻撃をパッケージ化したもので，”ExploitKit as a Service”と表現される場合がある．これは，ビットコインなどの仮想通貨を用いてお金を支払う事により，実際に攻撃に利用でき，その形態が一種のクラウドサービスのように見えるためそのように表現されている．

こういった日々増加する Exploit Kit による Drive by Download 攻撃への対策は急務である．これらの対策は現在でも多くの研究が行われている．しかしながら，現実のトラフィックに対し，提案される手法を採用するには難しい場合が多く，時間とともに劣化してしまい分類精度が落ちてしまう場合や，実際のトラフィックに適用するには誤検知が多く利用者の利便性を大きく残ってしまうケースが少なくない．こういった問題に対して，最新のデータ及び現実に近いデータを用いた検証・評価は急務であり，必要不可欠であるといえる．

本研究では，実際にユーザが利用することを考慮し，現在においても有効に利用できるか，また現実に近いリクエストが発生しても同様に分類ができるかを目的として実験・評価を行う．そこで，Drive by Download 攻撃に対して

† 東京電機大学 17fmi24@ms.dendai.ac.jp
‡ 東京電機大学

先行研究で示されているドメイン情報もしくは URL を用いた特徴に注目し，その特徴を用いて SVM による分類を行い検知率の評価を行う．そして，その値が実際のトラフィックに対しても有効であるかを検証し報告を行う．

2. Drive-by-Download 攻撃

2.1 概要

Drive by Download 攻撃は，ブラウザやブラウザのプラグインの脆弱性を利用し，意図せずマルウェアをダウンロード・実行する脆弱性である．この攻撃には，特徴的な流れが存在する．攻撃は図 1 で示すような流れで行われる．

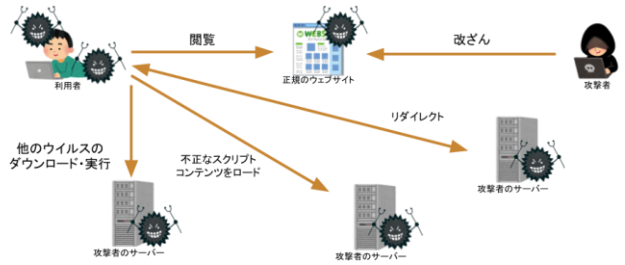

図 1 Drive-by-Download 攻撃の流れ

攻撃者は，改ざんもしくは不正広告を利用し，悪意のある不正なページをロードさせる．ページ改ざんの場合は攻撃者のページをロードする iframe などを正規の Web ページに存在する脆弱性を利用し改ざんを行う．また広告を利用する場合は，正規のサイトでそもそも利用されている広告サービスを用いて，攻撃者のページをロードするような不正な広告を配信する．そのページロードを入り口として，数回リダイレクトしブラウザの脆弱性を利用し攻撃を行うページに到達する．攻撃を行うページでは，Flash Player や VBScript などの脆弱性を利用し，マルウェアのダウンロード及び実行を行っている．

これらの一連の攻撃を一つのパッケージとして，Exploit Kit と呼ばれる形で提供し，ダークウェブ上等で取引されている現状がある．

2.2 Exploit Kit

Exploit Kit は，Drive by Download 攻撃を行うためのパッケージであり，多くは攻撃コード，及びマルウェアのダウンロードを提供する機能を提供するものである．Yin Minn Pa Pa[2]らによる調査によれば，現在でも活動を行っている Rig Exploit Kit は累算でおよそ 767 BTC もの売上を記録しており，日本円に換算すると 2018 年 8 月の交換レートでおよそ 5 億円以上の売上となっている．これは，無視できない規模での攻撃が継続して行われていることを示唆している．

2.3 Rig Exploit Kit

Rig Exploit Kit は，本稿執筆時点においても非常に活発に活動しており，日本国内においても対策が必要とされる Exploit Kit である．報告によれば，2016 年 9 月から非常に活発になり[1]ビットコインの支払い記録からみると 2017 年も継続的に活動を行っていたと考えられる．

この Rig Exploit Kit は，金銭を受け攻撃基盤を提供する “ExploitKit as a Service” の形として提供されている．これは，利用者が行う攻撃のアクセス率などの利用率や，実際の攻撃に用いるサーバの選択など Web からアクセスし管理を容易にするコントロールパネルも含めて提供されており，図 2 のようなログインパネルが存在する．

また Rig Exploit Kit は小池らと RSA 社によるテイクダウンが行われた[3][4]．しかしながら，これらのキャンペーンは，不正利用されているドメインのテイクダウンを目的としたものであり，現在ではドメインではなく IP アドレスを直接指定することで，現在でも攻撃が継続して行われている．

図 2 Rig Exploit Kit のコントロールパネルログイン画面

3. 先行研究

URL に存在する特徴を利用した分類として，尾崎らによる「悪性 URL の強調表示による Drive-by Download 攻撃解析支援手法の提案」があげられる[5]．尾崎らは URL の特徴に着目し，良性データとして有志によって収集された URL420 件，悪性データとして D3M データセット 2010 ~ 2015 から抽出した URL を 420 件用意し，それらのデータから決定木を利用し 87.38%の分類を行えることを示した．また，この研究では URL から分かる特徴量として次に示す情報を利用した．

- URL 全長
- Path 文字数
- 記号文字数
- Query 文字数
- Query の Key 数
- URL の数字の文字数

上述した情報から，URL に存在する特徴から良性・悪性の分類を行うことが可能であることが判明した．

また，類似する研究としてドメイン情報を用いた分類を行った岡安らによる「ボットネットの C&C サーバ特定手法におけるフィルタシステムの提案と評価」が挙げられる[6]．岡安らの研究では，悪性と良性ドメインにおいて公開されているドメイン情報に着目し，DNS と，WHOIS から得られる情報を用いた SVM を利用し，95.7%の分類を行うことに成功した．この研究ではドメインから利用できる特徴量として次に示すものを利用した．

- MX
- TXT
- CNAME
- Minimum
- 登録期間

ここまでの先行研究で示された手法などの違いを表 1 に示した．

表 1 先行研究の比較

	尾崎ら[5]	岡安ら[6]
データソース	URL	ドメイン
データセット	D3M データセット	Alexa Top Site
利用する特徴	URL 全長，Path 長，記号文字数，Query 文字数，Query Key 数，URL に含まれる数字の数	MX レコード数，TXT レコード数，CNAME レコード数，Minimum 値，登録期間
分類手法	決定木	SVM
正答率	87.38%	95.7%

4. データセット

4.1 概要

本研究では，適切な評価を目指し実際の環境に近いデータを利用する．そこで，悪性データは 1 年以内のものを収集し，また良性データは，アクセス数の多いものかつ利用者のアクセスとして一般的なものを利用した．

4.2 悪性データ

悪性データは nao_sec[7]よって報告された URL 及びドメインを用いる．nao_sec は日本における非営利のセキュリティチームであり，Drive-by-Download 攻撃を継続的に観測し研究を行っている．これらのデータは D3M と比較し最新かつ件数も多い．そこで本研究では，nao_sec らによって発見された Rig Exploit Kit に利用されている URL 及びドメインのリストを採用し，2017 年に報告されたドメイン及び URL を対象として実験を行った．

4.3 良性データ

良性データとして，実験で利用するデータは URL の特徴を用いた実験では，Twitter のツイートに含まれるリンクを抽出し，短縮 URL などを解除した URL を用いた．これは，一般ユーザの多くは SNS を多く利用していると考え採用した．また，収集したデータはスパムなどに利用される URL も含むと考えられるが，本実験では Drive-by-Download 攻撃とそれ以外での判定を行うため，その URL を用いた．

また，ドメイン情報は OpenDNS のアクセス上位 100 万件の情報を持つ Cisco Umbrella による Domain List[8]を利用する．これは，Alexa が提供しないデータを利用するためこのデータセットの利用をおこなった．

5. 実験

5.1 概要

実験では，先行研究で用いられている特徴が，現在においても有効であるか，また実際を模したトラフィックに対しても有効であるか評価を行う．実際を模したトラフィックとしては，前述の Twitter 先行研究で示された URL の特徴から得られる情報と Cisco Umbrella Domain List を用いて DNS および WHOIS から得られる情報を用いて SVM を利用した分類実験を行い評価した．評価は，正答率と False Positive，True Negative の値を算出し，結果を示すこととする．

5.2 URL の特徴を用いた分類実験

Twitter から得られた URL 情報及び，nao_sec から得た URL 情報を用い，悪性・良性の URL から 2500 件ずつを無作為に選択した．それら混合した 5000 件のデータを元に先行研究で示された尾崎らの研究で利用された特徴を利用し SVM による学習を行い，残りの 5000 件で評価を行った．データは悪性・良性を各 4 分割し交差検証としてテストデータと教師データを入れ替え，実験結果としてその平均値を計測する．

先行研究では，決定木による分類を行っていたためすべての特徴を利用していたが，本実験では最も有効に働いた特徴の組み合わせを用いて分類の組み合わせによる結果を報告する．

5.3 URL の特徴を用いた分類実験結果

実験として先行研究で示された特徴からどの組み合わせが最も検知率が高くなるか実験を行い，特徴として URL 全長，Path 長，記号の数を用いたときが有効に働くことが分かった．結果として，正答率は 99.90%と非常に高い精度で

の分類が可能であることを示した．このときの実験結果を表 2 に示す．

表 2　URL の特徴を用いた分類結果

正答率	99.90 %
False Positive	0.10%
True Negative	0.00%

5.4 ドメイン情報を用いた分類実験

ドメイン情報は時間により変わるものであるため，直近ドメインが存在する 3 ヶ月のもののみを対象として実験を行う．実験では，Cisco Umbrella が公開する Cisco Umbrella List から得られるドメインと，nao_sec から得られた Rig Exploit Kit のドメインを良性・悪性ドメインとして 500 件ずつを選択し，それらを混合した 1000 件のデータを元に，岡安らによる研究で示されている特徴を用いて 2 分割し，SVM により分類を行った．

5.5 ドメイン情報を用いた分類実験結果

表 3 で示す結果となった．

表 3 ドメイン情報を用いた分類結果

正答率	92.60%
False Positive	6.87%
True Negative	1.03%

6. 考察

6.1 実験結果

本実験では，一般ユーザが Web ブラウジングすることを想定し，非常に多く利用される SNS である Twitter のツイートを収集し，より現実的な URL を収集した．またドメインは Cisco Umbrella が提供するドメインリストを用いて実験を行った．実験の結果から URL を用いた分類に関して非常に高いレートでの分類に成功することが分かった．これは，選択した特徴での結果が明確に分かれているためと考えられる．

6.2 データセット

より高い検知率を記録した URL の特徴について実験で示された結果から実際のデータに着目し，URL データにどのような違いが存在するのかを明らかにする．Rig Exploit Kit で用いられる URL は図 3 で示すようなものである．

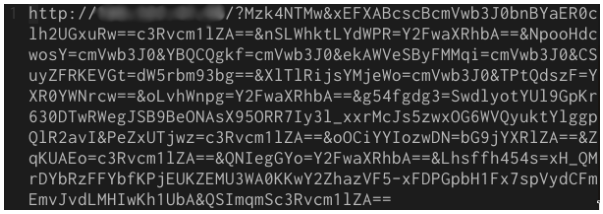

図 3 Rig Exploit Kit で用いられる URL

また，Twitter から収集した URL は図 4 で示すような URL である．

図 4 Twitter から収集した URL

図 3，図 4 で示された画像から，通常の URL とは明らかに違う形となっているのが伺える．実験で用いた特徴点である，Path の文字列長，URL 文字列長，記号文字数において，明らかに違いがあるのが見て取れる．

そこで，これらの特徴について，どのように差が存在するか特徴量とそれに対応する総数をヒストグラムとして，図 5，図 6，図 7 で示す．これらの比較から，Rig Exploit Kit で利用される URL の文字列には明確に実験で用いた URL 全長，Path 文字列全長，記号数の 3 つにおいて色濃く差が存在することが示された．本研究では Rig Exploit Kit に着目し分類を行ったが，今後他の Exploit Kit を用いた分類を行う必要があると考えられる．

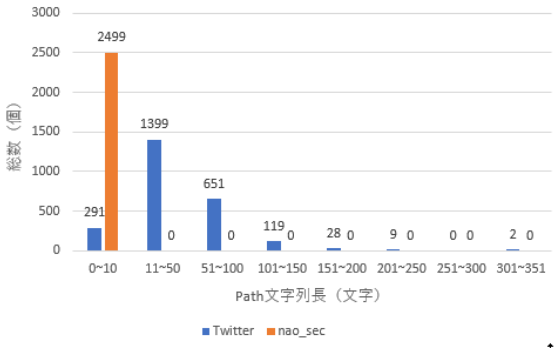

図 5 Path 文字列長

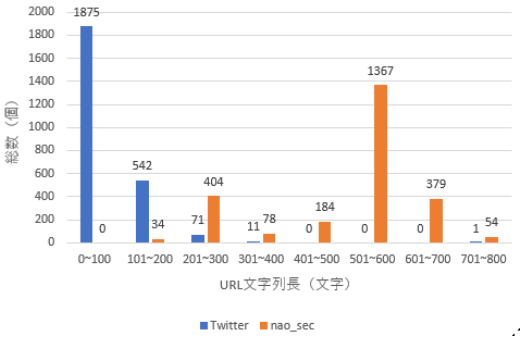

図 6 URL 文字列長

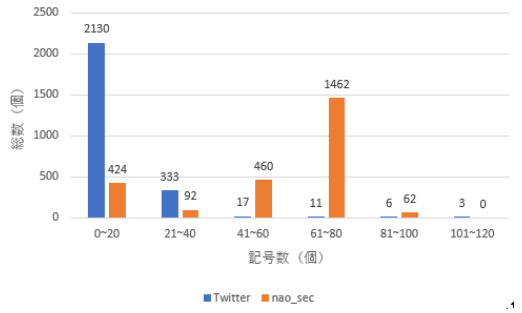

図 7 記号文字数

6.3 先行研究との比較

本実験は，Rig Exploit Kit のみに注目し先行研究によって示された特徴を用いて，それが現在において，また実際のトラフィックを模したものに対しても有効であるかを評価するものである．先行研究で検知対象としているデータに大きな違いがあり，単純に比較できるものではないが本研究では SVM を利用し，特徴選択によって Rig Exploit Kit の非常に高い検知率を記録することに成功した．検知率などの先行研究との比較を，表 4 に示す．

表 4 先行研究との比較

	検知率	対象
岡安らによる研究 [5]	87.38%	C&C サーバ
尾崎らによる研究 [6]	95.7%	D3M データセットに含まれる Drive-by-Download トラフィック
本研究	99.90%	Rig Exploit Kit

本研究ではより多くのデータ数と比較的新しいデータセットを用いて，評価・分類を行った．尾崎らによる先行研究では，2010 年から 2015 年のデータを用いており，また悪性 URL の件数は 420 件であった．本実験では，2017 年に報告された 2500 件の悪性 URL を用いて評価を行った．したがって，より十分な評価が行えたのではないかと考えられる．

また，岡安らによるドメイン情報を用いた分類では，92.60%の検知率であった．ドメイン情報はサーバが削除された際や，テイクダウンされた際などに変わる恐れがあり，実際にその場での特徴量抽出が必要となるといえる．先行研究の対象としている通り悪性トラフィックに対し，広範囲での適用が望まれるが，その分類を行うモデルの作成は先行研究で示されている 1 年間スパンの悪性・良性ドメインの情報を保持し続けなければならないため難しいと考えられる．本研究では，URL から得られる情報を用いているが，URL 自体ログに載っているデータそのままを利用でき，時間経過による変更がないため，モデル作成はより簡単であると考えられる．

7. 今後の展望

本研究では，Rig Exploit Kit に注目し，先行研究で示された URL の特徴から SVM を利用し悪性・良性のデータ 2500 件ずつを混合したデータから分類と，ドメイン情報を用いた SVM 分類両者の評価を行った．特に URL の特徴を用いた分類では 99.90%の正答率を出すことに成功した．この数値はかなりの精度であり，実際のトラフィックに対しても十分に検証が可能であると考えられる．

Exploit Kit は流行によって形を大きく変えるものである．実験時では，高精度であるが今後の新たな攻撃キャンペーンの台頭により，特徴選択等も変更する必要があると考えられる．今後は，さらなる Exploit Kit が出現した際に同様の評価を行い，従来手法が有効であるかなどを調査する必要があると考えられる．また，同じくこの特徴選択などを自動化し，より楽に検証ができないかなど研究を進める必要があると考えられる．

今後の展望として，本研究における評価を Rig Exploit Kit だけではなく，別の Exploit Kit に対し，Rig Exploit Kit のデータを用いて作成されたモデルでの検知率評価，新規 Exploit Kit のデータを含めた新規モデルの作成と評価を行う必要がる．さらに，実際のトラフィックに対して実験成果を適用するため，フィルタリングシステムの開発を行い，評価を行っていく必要があるといえる．

参考文献

[1] NTT セキュリティ, RIG エクスプロイトキット 解析レポート
https://www.nttsecurity.com/docs/librariesprovider3/default-document-library/jp-rigek-analysis-report , (参照 2018-8-1)
[2] Yin Minn Pa Pa , Hiroshi Kumagai, Masaki Kamizono, Takahiro Kasama “Counter Infiltration: Future Proof Counter Attacks Against Exploit Kit Infrastructure”, Black Hat Asia 2018 Briefings, https://www.blackhat.com/docs/asia-18/asia-18-papa-Future-

Proof%20Counter%20Attacks%20Against%20Exploit%20Kit%20Infrastructure-WP.pdf , (参照 2018-8-1)
[3] 小池 倫太郎, 菊池 浩明 “Drive-by Download 攻撃における RIG Exploit Kit の解析回避手法の調査”, コンピュータセキュリティシンポジウム 2017, pp364-389, 2017 年 11 月
[4] RSA, “SHADOWFALL”, https://www.rsa.com/en-us/blog/2017-06/shadowfall , (参照 2018-8-1)
[5] 尾崎 幸也, 上山 真也, 小西 達也, 山崎 雅斗, 坂東 翼, 小林 孝史” 悪性 URL の強調表示による Drive-by Download 攻撃解析支援手法の提案”, コンピュータセキュリティシンポジウム 2017, pp817-822, 2017
[6] 岡安翔太, 佐々木良一” ボットネットの C&C サーバ特定手法におけるフィルタシステムの提案と評価”, 情報処理学会論文誌,Vol58,No.1, pp249-257, 2017
[7] nao_sec
http://www.nao-sec.org/ , (参照 2018-8-1)
[8] Cisco Umbrella, “Umbrella Popularity List”, http://s3-us-west-1.amazonaws.com/umbrella-static/index.html , (参照 2018-8-1)