Computer Security Symposium 2018
22 - 25 October 2018

ドメインを用いない Web アクセスに着目した悪性サイト検知手法の提案

中川 雄太[†1] 上野 航[†1] 田辺 瑠偉[†2] 藤田 彬[†2] 吉岡 克成[†2, 3] 松本 勉[†2, 3]

概要：改ざんされた正規の Web サイトや不正な広告を利用してユーザを悪性 Web サイトに誘導し，Web ブラウザの脆弱性を悪用してマルウェアを強制的にダウンロード・実行させる Drive-by Download 攻撃が問題となっている．このような DBD 攻撃を行う悪性サイトの中には，ドメインを持たずに IP アドレスを直接指定して通信を行うものが存在する．本稿では，エンドユーザから得られる膨大な Web アクセスログの中から，DBD 攻撃などに用いられる悪性サイトの IP アドレスを検知する手法を提案する．具体的には，Web アクセスログの中から IP アドレスを直接指定したアクセスに着目し，IP アドレス空間，ユーザアクセス，URL 文字列を分析して 5 つの特徴量を抽出する．そして，機械学習を適用することで判定器を作成し，悪性サイトの IP アドレスを検知する．評価実験の結果，最も精度の高い判定器が 90.09%の検知率を示した．

キーワード：Drive-by Download 攻撃，Web アクセスログ，悪性サイト検知，IP アドレス，機械学習

Detecting Malicious Websites from Direct-IP Web Accesses

Yuta Nakagawa[†1] Wataru Ueno[†1]
Rui Tanabe[†2] Akira Fujita[†2] Katsunari Yoshioka[†2, 3] Tsutomu Matsumoto[†2, 3]

Abstract: Drive-by download attack has been a threat in recent years. The attack redirects users whom accessed compromised websites or malicious advertisements to malicious websites that exploit vulnerability of web browsers and cause malware infection. Among malicious web sites that trigger DBD attacks, instead of using domains some servers use IP addresses to communicate with each other. In this report, we propose a method to detect IP addresses which are assigned to malicious websites from direct-IP web access log of end users. In our proposal method, we focus on web accesses that use IP addresses instead of domains and extract 5 features concerned with IP address, User access, and URL characters. We create a classifier using machine learning and detect IP addresses of malicious websites. From the evaluation experiment, we show that the rate of detecting malicious IP addresses were 90.09%.

Keywords: Drive-by Download Attack, Web Access Log, Malicious Websites Detection, IP address, Machine Learning

1. はじめに

改ざんされた正規の Web サイトや不正な広告を利用してユーザを悪性 Web サイトに誘導し，Web ブラウザの脆弱性を悪用してマルウェアを強制的にダウンロード・実行させる Drive-by Download 攻撃（以降では，DBD 攻撃と呼ぶこととする）が問題となっている[1][2].

DBD 攻撃を行う Web サイトは，典型的には Exploit Kit と呼ばれるツールを用いて構築される．Exploit Kit の多くは Web ブラウザの脆弱性を突く攻撃コードを複数準備しており，Web サイトにアクセスしてきたユーザの Web ブラウザの脆弱性を悪用して，PC にマルウェアをダウンロード・実行させるまでの一連の攻撃を行う環境基盤を提供する．このような背景から，攻撃者は高度な知識や能力を持たずに DBD 攻撃を行う悪性サイトを構築することができ，DBD 攻撃による被害は増え続けている．

DBD 攻撃への対策として，攻撃に用いられるサーバのドメインや悪性サイトの URL をブラックリスト化することで，DBD 攻撃を未然に防ぐ手法が存在する．しかし，攻撃者はドメインや URL を頻繁に変更することができ，DBD 攻撃を行う Web サイトは増加し続けているため，全ての攻撃をブラックリストで防ぐことは困難である．また，近年では Exploit Kit の一種である RIG Exploit Kit による攻撃事例が多数報告されている．文献[3]の調査報告によると，ドメイン名を持たない悪性サイトが増加しており，対策が求められている．

本研究では，エンドユーザから得られる Web アクセスログの中から IP アドレスを直接指定したアクセスに着目し，悪性サイトの IP アドレスを検知する手法を提案する．IP アドレスを直接指定するアクセスには，不正広告などによってユーザが強制的に悪性サイトへ誘導される場合の他，企業や組織が管理するファイル共用サーバやメールサーバな

†1 横浜国立大学大学院 環境情報学府
Graduate School of Environment and Information Sciences, Yokohama National University.
†2 横浜国立大学 先端科学高等研究院
Institute of Advanced Sciences, Yokohama National University.
†3 横浜国立大学大学院 環境情報研究院
Graduate School of Environment and Information Sciences, Yokohama National University.

ど，悪性な活動とは無関係なサーバへのアクセスが考えられる．以降では，悪性サイトに割り当てられた IP アドレスを悪性 IP アドレスと呼び，悪性な活動とは無関係なサーバに割り当てられた IP アドレスを良性 IP アドレスと呼ぶこととする．図 5 に悪性 IP アドレスのアクセス URL の例を示す．提案手法では，Web アクセスログから抽出した IP アドレスを直接指定したアクセスに対して，IP アドレス空間，ユーザアクセス，URL 文字列の観点から分析を行い，5 つの特徴量を要素とする特徴ベクトルを抽出する．そして，教師あり学習を行うことで，悪性 IP アドレスと良性 IP アドレスを識別する判定器を作成して悪性 IP アドレスを検知する．

評価実験では，あるセキュリティベンダから提供された Web アクセスログに提案手法を適用し，特徴量の分析指標の組み合わせを変えた 4 種類の判定器を作成した．各判定器の精度評価として 4 分割交差検証を行ったところ，最も精度の高い判定器で 90.09%の検知率を示した．

本稿の構成は以下の通りである．はじめに，第 2 章で関連研究について説明する．そして，第 3 章で提案手法について説明し，第 4 章で評価実験の結果を説明する．最後に，第 5 章で考察を説明し，第 6 章でまとめと今後の課題を説明する．

2. 関連研究

悪性サイトを検知する方法は，悪性サイトの URL を基に悪性判定を行う方法と，悪性サイトの Web コンテンツを基に悪性判定を行う方法に大別することができる．

悪性サイトの URL を基に判定を行う方法には，Exploit Kit を用いて構築された Web サイトの URL 文字列を特徴として用いる検知手法が提案されている[4]．また，論文[5]では Web アクセスログから既知のブラックリストを用いて悪性 URL 群を抽出し，それらの URL 群から悪性サイトに共通するパターンを生成することで悪性サイトの可能性が高い URL 群を抽出する手法が提案されている．本研究では，ドメインを持たない Web アクセスにのみ着目しているため検知の対象とする URL は限られているが，悪性サイトの URL に見られる特徴だけでなく IP アドレス空間，ユーザアクセスの情報を用いて検知を行なっている点が特徴的である．

悪性サイトの Web コンテンツを基に悪性判定を行う方法として，PHP のファイル情報やファイルタイプ，レスポンスの発行時刻情報等が含まれる HTTP ヘッダ情報を用いて悪性サイトを検知する手法が提案されている[6]．論文[7]では，Web コンテンツに含まれる JavaScript に着目し，抽象構文解析木から特徴的な木構造を保持する悪性 JavaScript を抽出する方法が提案されている．また，論文[8]ではブラウザ組込型センサによる Web ブラウジング時のアクセス情報を用いて，Exploit Kit によって構築された DBD 攻撃サイトを検知する手法が提案されている．これらの手法は悪性サイトの特徴を抽出したのち，手動で検知ルールを作成して悪性サイトの検知を行なっているが，抽出した特徴に対して教師あり機械学習を適用することで悪性判定を行う手法も提案されている．論文[9]では，JavaScript の抽象構文木のノードに対して，単純ベイズ分類器を用いることで悪性な JavaScript を検知する手法が提案されている．また，論文[10]では HTML と JavaScript，Web ページの URL から 48 種類の特徴量を抽出し，これらに単純ベイズ分類器やランダムフォレストのようないくつかの典型的な学習モデルを適用することで悪性サイトの可能性の高い Web ページをフィルタリングする手法が提案されている．しかしながら，これらの Web コンテンツを基に悪性判定を行う方法は，検査対象の Web サイトが膨大である場合，Web コンテンツの取得に要するコストが大きくなるため，事前に検索対象の URL を絞り込む必要がある．

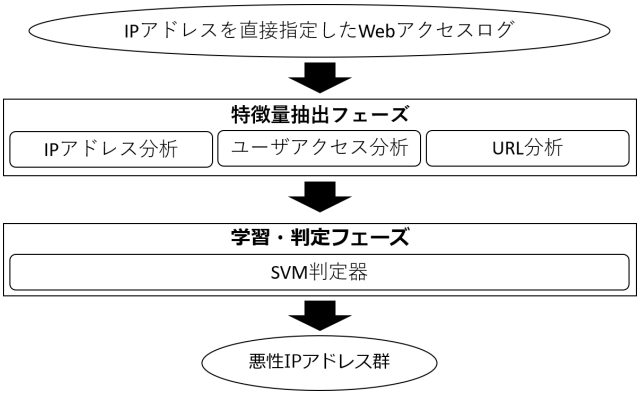

図 1: 提案手法の概要

一方，悪性サイトが運用されている Web サーバを検知する方法に，IP アドレス情報から悪性サイトを検知する手法が存在する．論文[11]では，IP アドレスのネットワークアドレス部のみを用いて不正 Web サイトを検知する手法が提案されている．この手法では，悪性な活動に利用される IP アドレスは一部のネットワークアドレスに密集する傾向にあることを利用して正規サイトと悪性サイトの判別を行なっている[12]．しかしながら，全ての悪性サイトが特定のアドレスレンジに存在するとは言えないため，IP アドレスのみを用いた悪性サイトの検知精度には疑問が残る．そこで，我々は IP アドレスだけでなくユーザアクセスと URL 文字列を用いて悪性サイトの検知を行う．

3. 提案手法

本章では，ドメインを用いない Web アクセスに着目した悪性サイト検知手法を提案する．図 1 に提案手法の概要を示す．提案手法は，特徴量抽出フェーズと学習・判定フェ

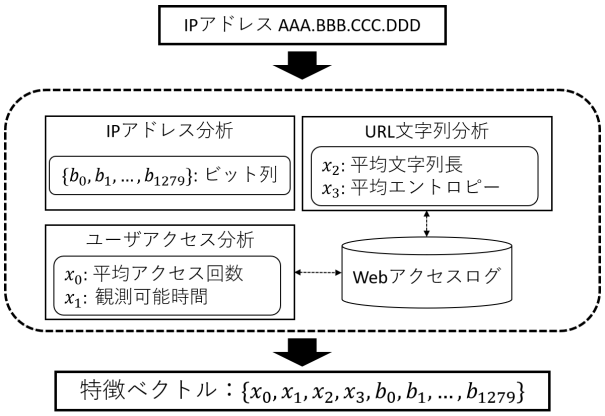

図 2 :特徴ベクトル抽出の流れ

ーズの 2 つのフェーズから構成される．特徴量抽出フェーズでは, IP アドレスを直接指定した Web アクセスを入力として，観測した各 IP アドレスについて IP アドレス空間，ユーザアクセス，URL 文字列の観点から分析を行い，5 つの特徴量を要素とする特徴ベクトルを抽出する．学習・判定フェーズでは, 教師あり学習を適用することで悪性 IP アドレスと良性 IP アドレスを識別する判定器を作成する. 以降では, 3.1 節で Web アクセスログの性質について説明し，3.2 節で抽出する特徴量抽出フェーズについて説明する. そして，3.3 節で学習・判定フェーズについて説明する．

3.1 Web アクセスログ

提案手法の入力となる Web アクセスログは，ユーザがアクセスした URL，アクセス時のタイムスタンプ，ユーザ識別子の計 3 つの情報が含まれていることを前提とする．また，提案手法は悪性サイトの検知を目的としているため，プライベート IP アドレスを指定したアクセスは検査対象外とする．なお，4 章の評価実験では，あるセキュリティベンダから提供された，数十万人規模のユーザに利用されているブラウザセンサを用いて 2018 年 4 月 27 日~2018 年 7 月 25 日の間に収集した 90 日分のアクセスログを用いた．

3.2 特徴量抽出フェーズ

特徴量抽出フェーズでは，IP アドレスを直接指定した Web アクセスから特定した検査対象 IP アドレスを入力として，IP アドレス空間，ユーザアクセス，URL 文字列の観点から分析を行い，5 つの特徴量を要素とする特徴ベクトルを抽出する．図 2 に特徴ベクトル抽出の流れを示す．以降では，それぞれの分析方法について説明する．

3.2.1 IP アドレス空間の分析

文献[12]より，悪性な活動に利用される IP アドレスは特定の IP アドレス空間に偏ることが明らかになっている. そこで，文献[12]で用いられている IP アドレスの第 1~3 オクテットの数値，及び，その組み合わせからなる 1280 のビット列を特徴量とする．

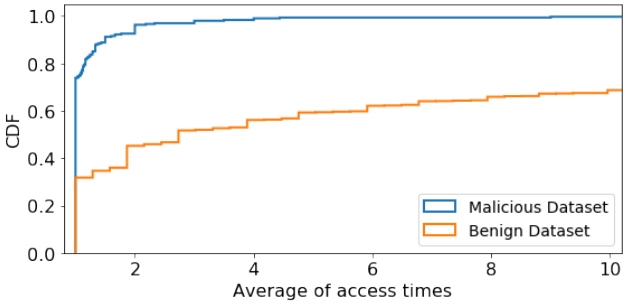

図 3: 良性・悪性 IP アドレスの
平均アクセス回数の累積分布

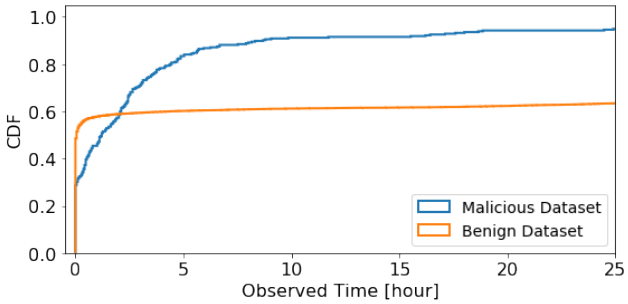

図 4: 良性・悪性 IP アドレスの
観測可能時間の累積分布

3.2.2 ユーザアクセスの分析

Web アクセスログから検査対象 IP アドレスに関するアクセスを特定し，検査対象 IP アドレスに対する(1) ユーザ 1 人当たりの平均アクセス回数，(2) IP アドレスの観測可能時間の 2 つを特徴量とする．以降では，それぞれの特徴について説明する．

(1) ユーザ 1 人当たりの平均アクセス回数

IP アドレスを直接指定したアクセスには，不正広告などによってユーザが強制的に攻撃サーバへ誘導される場合の他，企業や組織が管理するファイル共用サーバやメールサーバ，Web カメラ等 IoT 機器の Web UI，個人が運営するブログの Web サーバへのアクセスなど, 悪性な活動とは無関係なサーバへのアクセスが考えられる．

前者の場合，攻撃者はなるべく多くのユーザによるアクセスを目標とする．通常，そのような攻撃サーバへ積極的にアクセスするユーザは存在しないと考えられることから，悪性 IP アドレスへのユーザ 1 人あたりの平均アクセス回数は少なくなると考えられる．後者の場合，企業や組織が管理するファイル共有サーバは，原則その組織に属する特定のユーザ群のみアクセスが許されており，かつ，ファイル共有サーバなどは同一ユーザの定期的な利用が予想されることから，ユーザ 1 人あたりの平均アクセス回数が多くなると考えられる．このため，ユーザ 1 人あたりのアクセス頻度を分析することで, 悪性である可能性が高い IP アドレスの識別が可能だと考えられる．図 3 に 4 章の評価実験で用いたデータセットの良性・悪性 IP アドレスに対する平

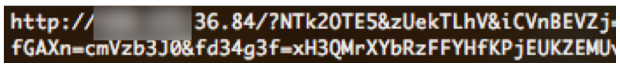

図 5: 悪性 IP アドレスのアクセス URL 例

均アクセス回数の累積分布（CDF）を比較した結果をまとめる．これより，多くの悪性 IP アドレスの平均アクセス回数が 2.0 以下となっていることが確認できる．そこで，ユーザ 1 人当たりの平均アクセス回数を特徴量とする．

(2) **IP アドレス観測可能時間**

悪性な活動を行うサーバやそのドメインは短命であることが知られている．ドメインを用いない攻撃サーバの IP アドレスも同様の特徴を持つことが予想されるため，IP アドレスの観測期間を分析する．

観測期間中におけるある IP アドレスへのk回目のアクセス時のタイムスタンプをt_kとし，各アクセス時のタイムスタンプからなる系列を次の式で定義する．

$$T = \{t_0, t_1, \ldots, t_m\}$$

さらに，k 回目のアクセスの前後のタイムスタンプの時間差を次の式で定義する．

$$\Delta t\ (= t_{k+1} - t_k)$$

本研究では，ある閾値τについて$\Delta t < \tau$を満たすタイムスタンプの部分系列を「観測可能期間」と定義する．観測可能期間は IP アドレス毎に 1 つ以上定義することができる．さらに，観測可能期間の始めのタイムスタンプと最後のタイムスタンプの時間差を「観測可能時間」と定義する．ただし，1 つの IP アドレスにつき 1 つ以上定義できる観測可能時間のうち，最も長いものを当該 IP アドレスの観測可能時間とする．

図 4 に 4 章の評価実験で用いたデータセットの良性・悪性 IP アドレスに対し，観測可能時間の累積分布を比較した結果をまとめる．これより，悪性 IP アドレスの多くが 24 時間以内に観測できなくなっていることが確認できる．そこで，IP アドレスの観測可能時間を特徴量とする．

3.2.3 URL 文字列の分析

Web アクセスログから検査対象 IP アドレスに関するアクセスを特定し，検査対象 IP アドレスを含む Web アクセスに対する(1) URL 文字列の長さ，(2) URL のエントロピーの 2 つを特徴量とする．なお，いくつかの IP アドレスは対応するユニークな URL が複数存在する場合があるため，提案手法ではユニークな URL 毎に特徴量を抽出し，その平均値を当該 IP アドレスの特徴量とする．以降では，それぞれの特徴について説明する．

(1) **URL 文字列の長さ**

悪性サイトに利用される URL は，良性 Web サイトの URL に比べて比較的長い URL が見受けられることが知られている[13]．そこで，検査対象 IP アドレスを含む Web アクセス毎の URL 文字列長を計算して特徴量とする．

表 1: 入力した Web アクセスログのデータ

観測期間		90 日
Web アクセスログ件数		2,215,802
ユニーク IP アドレス	良性 IP アドレス	5,281
	悪性 IP アドレス	232
	総 IP アドレス	5,513

表 2: 判定器作成における特徴量分析の組み合わせ

判定器	IP 分析	ユーザアクセス分析	URL 分析
A	✔	✔	✔
B	✔		
C		✔	
D			✔

(2) **URL 文字列のエントロピー**

URL の長さに加え，悪性サイトに利用される URL のドメインやクエリ文字列は，ランダムな文字列を含む傾向が見られる．図 5 に悪性 IP アドレスのアクセス URL 例を示す．そこで，検査対象 IP アドレスを含む Web アクセス毎の URL パスと URL クエリからなる部分 URL 文字列のエントロピーを計算して特徴量とする．ここで，n文字の部分 URL 文字列$X = \{x_0, x_1, \ldots, x_{n-1}\}$のエントロピーを次の式で定義する．

$$E = -\sum_{i=0}^{n-1} p(x_i) \log p(x_i)$$

$p(x_i)$は部分 URL 文字列Xにおける文字x_iの出現確率を表す．

3.3 判定器の作成

全ての検査対象 IP アドレスに対して特徴ベクトルを抽出し，教師あり学習を適用することで悪性 IP アドレスと良性 IP アドレスを識別する判定器を作成する．提案手法は，悪性 IP アドレスの検知を目的とした「２値分類問題」とみなせるため，教師あり学習手法として SVM（Support Vector Machine）を用いる．SVM は教師あり学習を用いる２クラスパターン認識モデルの一種である．本研究のように訓練データの特徴量分布に線形性が期待できない場合においても，カーネルトリックにより十分な分離性能での学習が可能である．

4. 評価実験

本章では，提案手法の精度評価のために行った実験について説明する．評価実験では，まず，あるセキュリティベンダから提供された Web アクセスログに提案手法を適用

表 3: 各判定器で設定した最適パラメータ

判定器	C	γ
A	1.000	0.001953
B	16.00	0.015625
C	64.00	0.062500
D	1.000	0.000244

表 4: 各判定器の Recall, Precision, F 値

判定器	Recall	Precision	F 値
A	0.8879	0.9406	0.9135
B	0.9009	0.9858	0.9414
C	0.7456	0.8122	0.7775
D	0.8922	0.9764	0.9324

し，特徴量分析指標の組み合わせを変えた 4 種類の判定器を作成した．作成した各判定器で 4 分割交差検証を行い，その結果から得られる検知率（Recall），適合率（Precision），F 値を算出して，提案手法を評価した．以降，4.1 節で評価実験の概要について説明し，4.2 節で作成した判定器について説明する．そして，4.3 節で各判定器から得られた評価結果を説明する．

4.1 実験概要

評価実験では，2018 年 4 月 27 日~2018 年 7 月 25 日の間に収集した IP アドレスを直接指定した Web アクセスログを入力として使用する．なお，観測期間中 1 度しかアクセスがなかった IP アドレスは，観測可能時間が定義できないため除外した．また，プライベート IP アドレスを指定した Web アクセスも検知対象外であるため除外した．観測可能期間の定義に用いる閾値τは 168 時間（1 週間）とし，各 IP アドレスの教師データの作成には Virus Total[15]の URL 検査サービスを利用した．その際，入力した Web アクセスログに含まれる URL 群を加工し，各 IP アドレスと検査対象 URL が対応するよう紐付けた．そして，Virus Total の検知エンジンの内，2 件以上が悪性判定を示した場合，当該 IP アドレスを悪性とみなした．表 1 に実験に使用したデータセットに関する情報をまとめる．

4.2 SVM による学習及び判定器の作成

特徴量分析の組み合わせが異なる A~D の 4 パターンの判定器を昨成した．表 2 に作成した判定器毎に用いた特徴量分析指標の組み合わせを示す．判定器の作成には Python の scikit-klearn[14]モジュールを使用して SVM による学習を行った．学習には SVM のカーネルメソッドとして RBF（ガウス）カーネルを用いた．また，学習時に設定が必要になるCとγの 2 種類のパラメータはグリッドサーチによる網羅的な探索を行い，最適値を決定した．表 3 に本実験で用いた判定器毎のパラメータを示す．

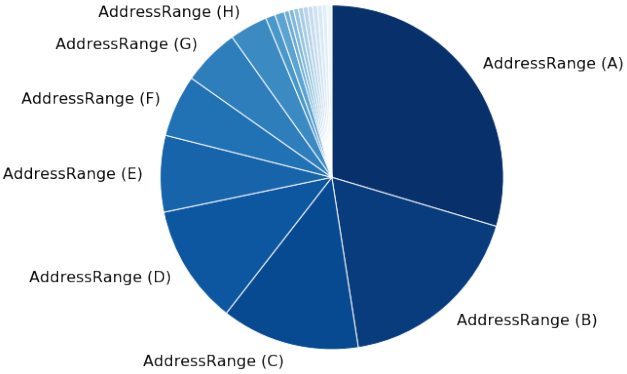

図 6: データセットに含まれる
悪性 IP アドレスのネットワーク分布

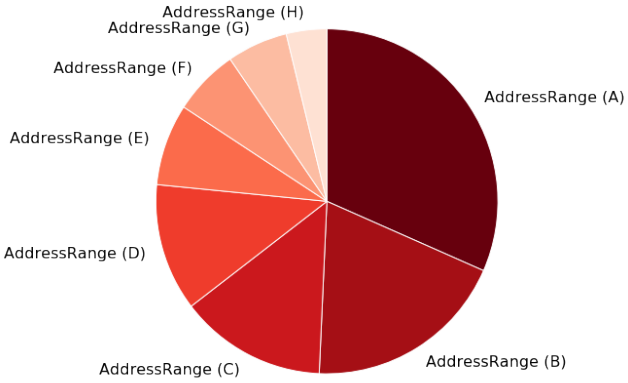

図 7: 判定器 B で検知できた
悪性 IP アドレスのネットワーク分布

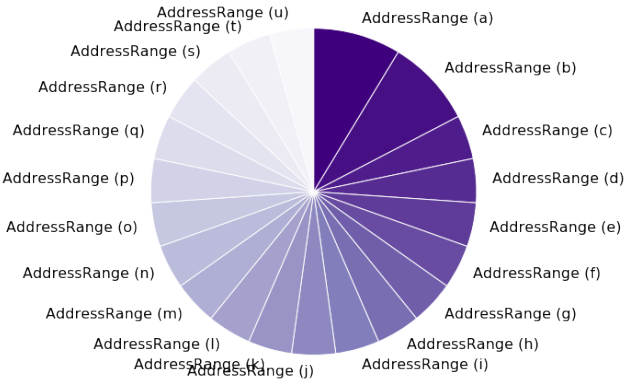

図 8: 判定器 B で検知できなかった
悪性 IP アドレスのネットワーク分布

4.3 実験結果

作成した各判定器に対し，4 分割交差検証を行った．検証結果から算出した各判定器の検知率（Recall），適合率（Precision），F 値を表 4 に示す．実験の結果，判定器 A,

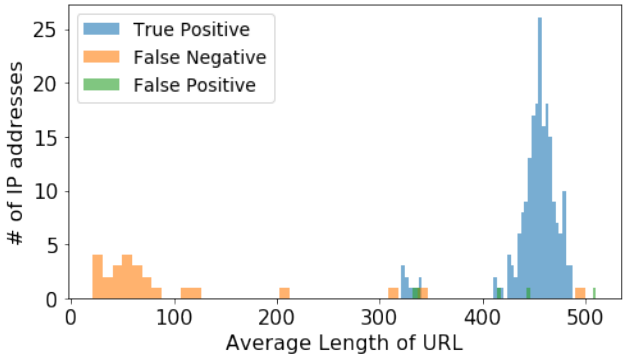

図 9: URL 文字列長のヒストグラム

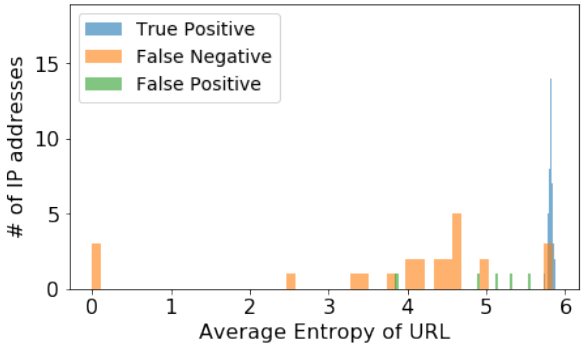

図 10: URL 文字列エントロピーのヒストグラム

B，D が高い検知率を示した．判定器 B，D はそれぞれ IP アドレス空間，URL 文字列のみを特徴量分析指標とした判定器である．これは， IP アドレス空間と URL 文字列の特徴が悪性 IP アドレスの検知に有効であることを示している．また，判定器 A は上記特徴量を包括しており，判定器 B，D に並ぶ検知率を示した．

5. 考察

評価実験の結果，悪性 IP アドレスの検出率が最も高い判定器は B の判定器であり，その検出率は 90.09%と高い精度を示した．B の判定器は IP アドレス空間から得られる特徴量のみを用いた判定器である．そこで我々は，評価実験で用いたデータセットに含まれる悪性 IP アドレスと，判定器 B で検知できた悪性 IP アドレス（=True Positives），判定器 B で検知できなかった悪性 IP アドレス（=False Negatives）の計 3 種類の IP アドレス群について，IP アドレス空間の偏りを比較した．図 6~8 にその割合を比較した円グラフ示す．図 6 より，データセットに含まれる悪性 IP アドレスの多くが，いくつかの特定の IP アドレス空間に集中していることが確認できる．また，図 6 と図 7 で示された IP アドレス空間の偏りが一致していることが確認できる．一方，図 8 に示された IP アドレス群では図 7 のような偏りは確認できなかった．これらの比較から，実験に用いたデータセットに含まれる悪性 IP アドレスの多くが特定の IP アドレス空間に偏っており，判定器 B はこうした偏りを学習して検知を行なったと考えられる．このような理由から，判定器 B は特定の IP アドレス空間に存在する悪性 IP アドレスは検知できるが，そのような空間に属さない悪性 IP アドレスを検知することはできないことが示された．これは IP アドレス空間を特徴量とした検知手法の限界と考えられる．

次に，判定器 D で検知できた悪性 IP アドレス（=True Positives），検知できなかった悪性 IP アドレス（=False Negatives），誤って検知してしまった良性 IP アドレス（=False Positives）の計 3 種類の IP アドレス群を対象に，URL 文字列分析から得られる 2 種類の特徴量について，ヒストグラムを比較した．図 9，図 10 にその結果を示す．これより，判定器 D で検知できた悪性 IP アドレスが各特徴量の特定の値域に集中していることが確認できる．一方，これらの値域から外れた悪性 IP アドレスは検知できておらず，IP アドレス空間から得られる特徴量を用いた検知手法と同様の限界を示していると考えられる．

最後に，判定器 C の結果について考察する．判定器 C は他の判定器と比べると検知率・適合率共に高くなかった．しかし，判定器 C で検知できた悪性 IP アドレス（=True Positives）について IP アドレス空間の偏りを調査したところ，判定器 B で検知できなかった IP アドレス空間に存在する悪性 IP アドレスを 3 件検知できたことを確認した．このことから，IP アドレス空間や URL 文字列から得られる既存の特徴量では検知できない悪性 IP アドレスが存在し，そのような悪性 IP アドレスに対する新たな検知指標として有用となる可能性がある．

6. まとめと今後の課題

本稿では，エンドユーザから得られる Web アクセスログの中から IP アドレスを直接指定した Web アクセスに着目し，悪性 IP アドレスを検知する手法を提案した．提案手法は観測した各 IP アドレスについて，IP アドレス空間，ユーザアクセス，URL 文字列の観点から特徴ベクトルを抽出し，教師あり学習を適用することで，悪性 IP アドレスを検知する判定器を作成した．また，提案手法の精度評価のため，実際の Web アクセスログを用いて判定器を作成し，悪性 IP アドレスの検知を行なった．その結果，最も精度の高い判定器で 90.09%の検知率を示した．また，本稿の評価実験では十分な精度を示せなかったが，悪性 IP アドレス検知のための新たな指標として，ユーザアクセスから得られる特徴量が有用となる可能性を示した．

今後の課題として，ユーザアクセス分析から得られる特徴量の拡張及び改善と，本提案手法適用後の Web アクセスログの詳細な分析の 2 つが挙げられる．本稿の評価実験及び考察でも言及したが，ユーザアクセスに基づいた検知指標はまだ十分な精度を示せていない．本提案手法で導入したユーザ 1 人あたりの平均アクセス回数と観測可能時間に

加え，さらなる特徴量の抽出を目指していきたい．また，本提案手法はエンドユーザから得られる Web アクセスログが前提にあるため，検知した悪性 IP アドレスを起点として，次のような分析への発展が期待できる．

(1) **悪性サイト到達フローの分析**

DBD 攻撃など多くの悪性サイトへのアクセスは，不正広告や改ざんされた正規サイトを閲覧することによって発生する場合が多い．提案手法で検知した悪性 IP アドレスを起点に前後のアクセスログを分析することで，前述のような入り口サイトや，ブラウザやソフトウェアの脆弱性を突く攻撃コードを送り込む，攻撃サーバの発見が期待できる．

(2) **悪性サイト到達ユーザの分析**

文献[5]では，頻繁に悪性サイトへアクセスするユーザの存在が報告されている．本提案手法で検知した悪性 IP アドレスを起点に前述のようなユーザを発見し，ユーザ単位でアクセスログを分析することで，悪性サイトの効率的な発見が期待できる．

謝辞 本研究成果の一部は、国立研究開発法人 情報通信研究機構(NICT)の委託研究「Web 媒介型攻撃対策技術の実用化に向けた研究開発」により得られた．本研究の一部は文部科学省国立大学改革強化推進事業の支援を受けて行われた．

参考文献

[1] “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 2018-08-20)

[2] “Rig Exploit Kit によるドライブ・バイ・ダウンロード攻撃の検知状況”. https://www.ibm.com/blogs/tokyo-soc/rig-exploit-kit/, (参照 2018-08-20)

[3] “Rig エクスプロイトキットの調査”. https://www.jpcert.or.jp/present/2018/JSAC2018_05_ikuse.pdf, Japan Security Analyst Conference 2018.

[4] 佐藤祐磨，中村嘉隆，高橋修，“エクスプロイトキットで利用される文字列特徴を用いた悪性 URL 検出手法の提案”，情報処理学会研究報告 Vol.2016-CSEC-72 No.25.

[5] 森島周太，中野弘樹，吉岡克成，松本勉，藤原礼征，“多数のユーザの Web アクセスログから効率的に悪性サイトを抽出する手法”，コンピュータセキュリティシンポジウム 2017 論文集, Vol.2017, No.2 (2017)

[6] 酒井 裕亮，佐々木 良一，“Drive By Download 攻撃に対する HTTP ヘッダ情報に基づく検知手法の提案”，情報処理学会研究報告 Vol.2013-CSEC-60 No.29.

[7] 神薗雅紀，西田 雅太，小島 恵美，星澤 裕二，“抽象構文解析木による不正な JavaScript の特徴点抽出手法の提案”，情報処理 学会論文誌 Vol.54 No.1 349-456.

[8] 笠間 貴弘，神薗 雅紀，井上 大介，Exploit Kit の特徴を用いた悪性 Web サイト検知手法の提案，コンピュータセキュリティシンポジウム 2013 論文集, Vol.2013, No.4, pp.603-610(2013)

[9] C. Curtsinger, B. Livshits, B. Zorn, and C. Seifert. “Zozzle: Fast and precise in-browser JavaScript malware detection.”, In Proc. of USENIX Security Symposium, 2011.

[10] D. Canali, M. Cova, G. Vigna, and C. Kruegel. “Prophiler: A fast filter for the large-scale detection of malicious webpages.”, In Proceedings of the International World Wide Web Conference, Mar. 2011.

[11] 金澤 しほり，中村 嘉隆，稲村 浩，高橋 修，“未知の不正 Web サイト判別のための IP アドレスクラスの特徴分析”，コンピュータセキュリティシンポジウム 2016 論文集, Vol.2016, No.2, pp.777-783(2017)

[12] D. Chiba, T. Mori, S. Goto. “Detecting Malicious Websites by Learning IP Address Features”, 2012 IEEE/IPSJ 12th International Symposium on Applications and the Internet

[13] L. Xu, Z. Zhan, S. Xu, and K. Ye. “Cross-layer detection of malicious websites”, Proceedings of the third ACM conference on Data and application security and privacy(CODASPY’13), pp.141-152, 2013.

[14] “scikit-learn: Machine Learning in Python”, http://scikit-learn.org/stable/, (参照 2018-08-20)

[15] “Virus Total”, https://www.virustotal.com/, (参照 2018-08-20)