悪性 IP アドレスの分布特徴に基づく 未知の Web サイトの判別手法

金澤 しほり[†1] 中村 嘉隆[†2] 稲村 浩[†2] 高橋 修[†2]

概要：近年，Drive-by download 攻撃やフィッシングなど Web サイトを介したサイバー攻撃が急増しており，ユーザの個人情報等が不正に取得されるなどした結果，経済的被害も増加している．このような被害を防ぐために，未知の Web サイトに対して正規・不正の判別を行い，未知の不正 Web サイトの検出率を向上させる．本稿では，未知の Web サイトの判別を行うために IP アドレスクラスのネットワークアドレス部の特徴を分析し，交差検定を用いて評価を行った．また，悪性 IP アドレスの特徴に経年変化が存在する IP アドレスクラス B と IP アドレスクラス C に対して年度別の特徴を反映させた結果，判別精度の向上が確認できた．

キーワード：サイバー攻撃，不正 Web サイト，ネットワークアドレス，IP アドレスクラス

Classification method of unknown websites based on distribution information of malicious IP addresses

Shihori Kanazawa[†1] Yoshitaka Nakamura[†2]
Hiroshi Inamura[†2] Osamu Takahashi[†2]

Abstract: In recent years, cyber attacks through websites such as Drive-by download attacks or phishing attacks increase rapidly. The attackers acquire personal information of users illegally by these attacks and inflicts economical damage. We aim to detect malicious websites which cause economic damage. We analyzed features of the network address part of the IP address class to distinguish unknown malicious websites. In addition, we evaluated the proposed system by cross-validation. IP address of address class B and IP address of address class C had aging characteristics of malicious IP addresses. Therefore, features of each fiscal year were reflected in discrimination in training data of the classification. As a result, the classification accuracy could be improved.

Keywords: cyber attack, malicious website, network address, IP address of class

1. 背景

近年，インターネット上で，ウイルスやマルウェアによる攻撃の脅威が年々増加している．その中で，特に Web サイトを利用した攻撃が急増している．2017 年 3 月に IPA(独立行政法人情報処理推進機構)が発表した「2017 年版情報セキュリティ 10 大脅威」によると 1 位と 4 位~6 位の 4 つが Web サイトに対する攻撃となっている[1]．攻撃例として，ユーザが Web サイトを閲覧した際に，ウイルスやマルウェアなどの不正プログラムをパソコンにダウンロードさせる Drive-by download 攻撃や，金融機関を装った偽のサイトへ誘導するフィッシング詐欺が挙げられる．これらの攻撃により，閲覧者のパソコンでマルウェアが活動し，保管されたデータやプログラムを破壊される事件や，暗証番号やクレジットカード番号などの個人情報を不正に取得され，経済的な被害を受ける事件が増加している．図 1 は，2011 年から 2015 年までの警察庁広報資料「インターネットバンキングに係る不正送金事犯発生状況」の発生件数と被害額データを示している[2]．

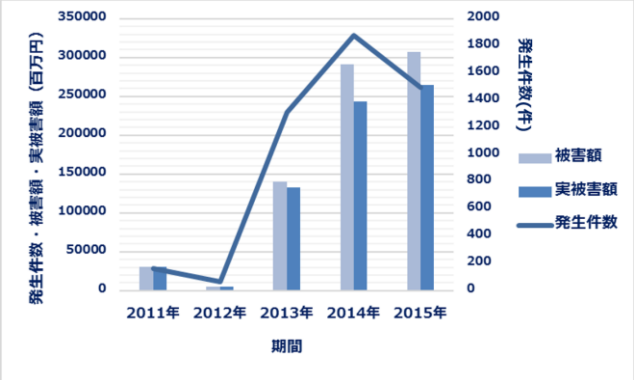


図 1 インターネットバンキングに係る 不正送金事犯発生状況

文献[2]によると，インターネットバンキングを介して，

†1 公立はこだて未来大学大学院 システム情報科学研究科
Graduate School of Systems Information Science, Future University Hakodate
†2 公立はこだて未来大学 システム情報科学部
School of Systems Information Science, Future University Hakodate

個人情報が不正に取得され，経済的被害を被った件数は，2012年まで50件程度であったのに対し，2015年には1400件近くまで急増しており，現在も増加傾向にある．このような被害を防ぐため，ユーザが不正Webサイトにアクセスしないように対策する必要がある．

これに対し，アクセス先の Web サイトが不正 Web サイトである場合に，アクセスを遮断する方法がいくつか存在する．Web レピュテーションシステム[3,4]は，不正 Web サイトブロック機能を持つソフトウェアで実現されている．ユーザによる Web アクセス通信が発生する際に，接続先のドメイン名や Web サイトが不正であると判断された場合には，そのアクセス自体をブロックすることによって，不正プログラムによる感染，およびフィッシングによる被害を防止している．しかし，このとき不正 Web サイトとして判断されるものは，すでにウイルス配信，フィッシング詐欺など，不正行為を行ったことが確認された Web サイトに限られる．また，Intrusion Prevention System（IPS，侵入防止システム）[5]を用いる方法もある． IPS は，ファイアウォールやアンチウイルスソフトウェアのみでは防御が困難とされていた DoS 攻撃やボットによる攻撃など，巧妙かつ高度なセキュリティの脅威にも対応している．Web サイトへアクセス通信が行なわれた際に，通信に含まれる不審な通信パケットを検出して，その通信を遮断する仕組みになっている．しかし，この IPS による検出も，Web アクセス通信に含まれる既知の不審パケットのみに限られる．

前述した 2 つの方法は，既知の不正 Web サイトや不正 Web サイトへのアクセス通信に含まれる既知の不審パケットなど，既知の情報を用いているため，既知の不正 Web サイトに関しては検出率が高いという利点がある．しかし，未知の不正 Web サイトに対応した検出が困難であり，仮に検出できた場合であっても，十分な精度が得られるか不明である．このような問題点を解決するために，未知の不正 Web サイトを含めた検出が可能な検出条件を設定し，検出された Web サイトを既知の Web サイトか未知の Web サイトのどちらかに分類した上で，未知の Web サイトに対して正規 Web サイトか不正 Web サイトを判別する手法を提案する．

本論文は，以下のように構成されている．第 2 章では，本研究に関連する研究について述べる．第 3 章では，提案手法について述べる．第 4 章では，提案手法の評価実験及び実験結果の考察について述べる．最後に，第 5 章では，今後の課題について述べる．

2. 関連研究

不正 Web サイトの検出に関する研究として，URL の特徴に基づいた検出手法を用いるものや，ドメイン名の特徴に基づいた検出手法を用いるもの，IP アドレスの特徴に基づいた検出手法を用いるものなどがある．

例えば，URL の特徴に基づいた検出手法には，J. Ma ら[6]の，URL の字句構造に基づいて，URL を良性または悪性に分類するための教師あり学習手法を取り入れたシステムがある．

ドメイン名の特徴に基づいた検出手法には，劉ら[7]の，不正 Web サイトに見られるドメイン名の特徴を検出条件として用いている手法や，L. Bilge ら[9]の，DNS 分析技術を用いて悪質な活動に関与するドメイン名を検出する手法，田中ら[10]の，マルウェアが通信を行う際の特徴を利用して，DNS 通信の観測によって未知の不正 Web サイトを検出する手法などがある．文献[7]では，不正 Web サイトに見られるドメイン名の特徴を検出条件として用いている．不正 Web サイトのドメイン名は，英数字がランダムに混在するものが多い傾向にあるため，英数字が混在するドメイン名を利用しているかどうかを 1 つ目の検出条件としている．また，不正 Web サイトのドメイン名は，ボットに感染したパソコン群(ボットネット)を利用してフィッシングやウイルス配布などを行う Fast-Flux[8]などの攻撃手法を用いて自動生成されることが多いため，人間にとって扱いにくい 10 文字以上の長い文字列で構成されているものが多い．そのため，10 文字以上で構成されるドメイン名であることを 2 つ目の検出条件としている．また，文献[10]は，マルウェアが通信を行う際の特徴を利用して，DNS 通信の観測によって未知の不正 Web サイトの検出を行っている．通信を行うマルウェアに感染しているクライアントは複数の不正 Web サイトにアクセスを行う傾向にあるため，不正 Web サイトにアクセスを行ったクライアントは，他の不正 Web サイトにもアクセスを行っている可能性が高い．そのため，DNS 通信において既知の悪性ドメインにアクセスを行っていたクライアントから名前解決要求のあるドメインは，マルウェアとの関連が深いドメインであると考えられるため，未知の不正 Web サイトとして検出している．これらの手法は，ドメイン名のブラックリストを使用するため，既知の Web サイトの検出に有効である．しかし，条件に該当しない不正 Web サイトを検出できない．また，ドメイン名は，容易に生成することができるため，頻繁に変更されやすく，ブラックリストを最新に維持することが難しい．次に，IP アドレスの特徴に基づいた研究について述べる．

IP アドレスの特徴に基づいた検出手法には，千葉ら[11,12]の，不正 Web サイトに見られる IP アドレスの特徴を不正 Web サイトを見つけるための判別条件として用いる手法がある．不正 Web サイトは，特定の IP アドレス群を使用する傾向があり，この特徴を用いて，良性 IP アドレスと悪性 IP アドレスを判別する．ドメイン名と比べ，IP アドレスは容易に変更できる情報でないため，判別に用いる情報として適している．しかし，この手法では特徴を取得するのに利用できる IP アドレスが限られているため，判別

が可能な IP アドレスの範囲が狭い．また，IP アドレスを千次元以上の特徴ベクトルに変換して判別に利用するため判別の負荷が大きい．

3. IP アドレスを用いた不正 Web サイトの判別手法

3.1 アプローチ

本研究では，ドメイン名の特徴と IP アドレスの特徴を併用した不正 Web サイトの判別手法を提案する．ドメイン名の特徴を用いた検出時には，既存手法等で用いられている複数の検出条件を併用して検出条件を拡張することで，ブラックリスト型の欠点の解消をめざす．IP アドレスの特徴を用いた判別時には，IP アドレスのすべての範囲において Web サイト分布特徴を取得し，各 IP アドレスの特徴について最低限の情報量に制限して表現することで，判別コストを軽減しながら高い判別制度の維持をめざす．

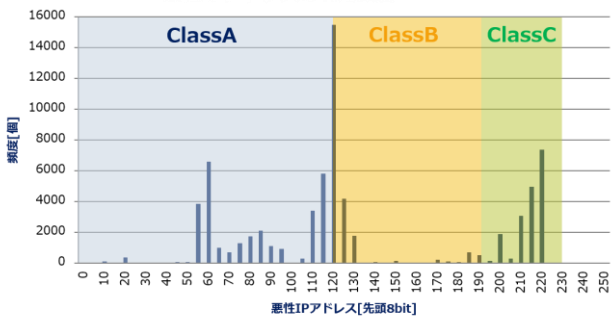


図 2 悪性 IP アドレスの利用頻度

図 2 は，悪性 IP アドレスの利用頻度を表す分布である[14]．サイバー攻撃は，特定の IP アドレス群を使用する傾向がある[11,12,13]．また，各 IP アドレスクラスにおいて悪性 IP アドレスの利用頻度に差が表れている．千葉ら[11]の研究では，主に IP アドレスクラス C の利用頻度に着目していたが，他の IP アドレスクラスにおいても，特定の悪性 IP アドレス群が集中的に利用されているため，IP アドレスクラス C に限定せず，全ての IP アドレスの範囲の特徴を取得して用いる．また，文献[11]では，判別時に，IP アドレスを最大 1000 以上の次元数に変換した特徴ベクトルが必要となる．そこで，IP アドレスそのものを 2 進数変換したうえで，ネットワークアドレス空間の特徴に基づき，各 IP アドレスのネットワークアドレス部のみを用いることによって，判別に必要な特徴ベクトルの次元数を高精度かつ低コストな判別を可能とする．

3.2 提案手法

図 3 に，提案手法の概要を示す．

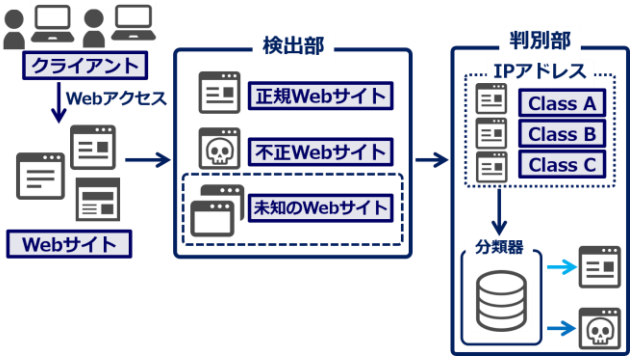


図 3 提案手法の概要

クライアントがアクセスする Web サイトは，正規 Web サイト，不正 Web サイト，および未知の Web サイトの 3 つに分類できる．このうち，既知である正規 Web サイト・不正 Web サイトに関してはブラックリスト方式で対応可能であるため，特に未知の Web サイトに対し，IP アドレスのみを用いて低コストで正規・不正に判別可能なシステムを提案する．

図 4 に，提案システムの全体像を示す．

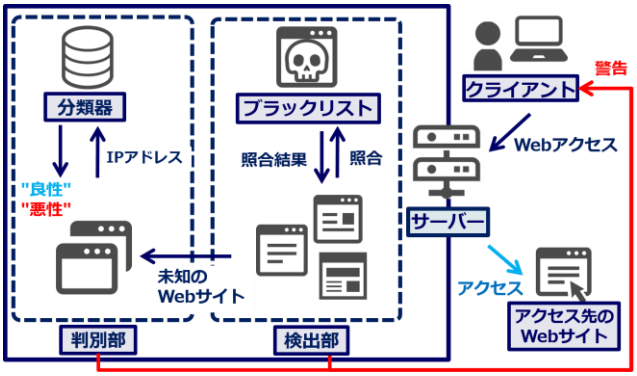


図 4 提案システムの全体像

提案システムは，検出部と判別部の 2 つのフェーズで構成されている．

検出部では，ドメイン名に関するブラックリストを用いて正規 Web サイトを検出対象から除外し，未知の Web サイトを検出する．まず，クライアントからある Web サイトへの通信が DNS サーバを通過する際に，検出部は，アクセス先の Web サイトをブラックリストと照合する．アクセス先の Web サイトが未知の Web サイトであると判断された場合，検出部は，判別部に該当 Web サイトの IP アドレスを送信する．判別部は，検出部より送信された IP アドレスの特徴を用いて，未知の Web サイトが正規 Web サイトであるか不正 Web サイトであるか判別する．判別結果が正規 Web サイトである場合，クライアントが Web サイトにアクセスすることを許可する．一方，判別結果が不正 Web サイトである場合，クライアントに警告することによって通信

の中断を促し，該当する不正 Web サイトをブラックリストに追加して最新の状態に保つ．

3.3 未知の Web サイトの検出方法

図 5 は，検出部の詳細を示している．

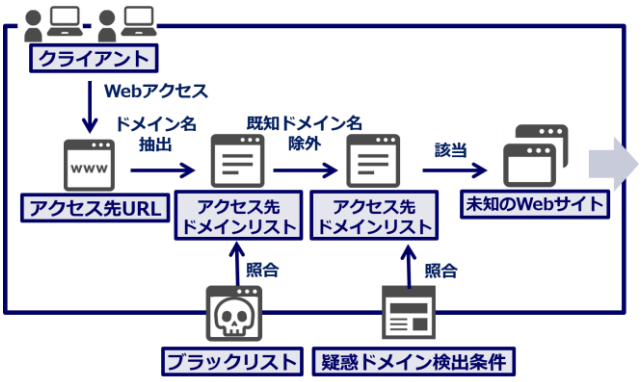


図 5 検出部の詳細

検出部では，ドメイン名の特徴を用いて未知の Web サイトを検出する．ドメイン名は，Web サイトの URL から取得することができる．既知の悪性ドメインを除外するために，ドメイン名に関するブラックリストと照合する．Web サイトのドメイン名がブラックリストに存在しない場合，ドメイン名を，ドメイン名の特徴に基づく検出条件と照合する．ドメイン名の特徴には，10 文字以上のドメイン名，英数字がランダムに混在するドメイン名，マルウェアに感染されたクライアントからアクセスされたドメイン名の 3 つの条件を設定する．

マルウェアに感染したクライアントの検出方法について，図 6 に詳細を示す．

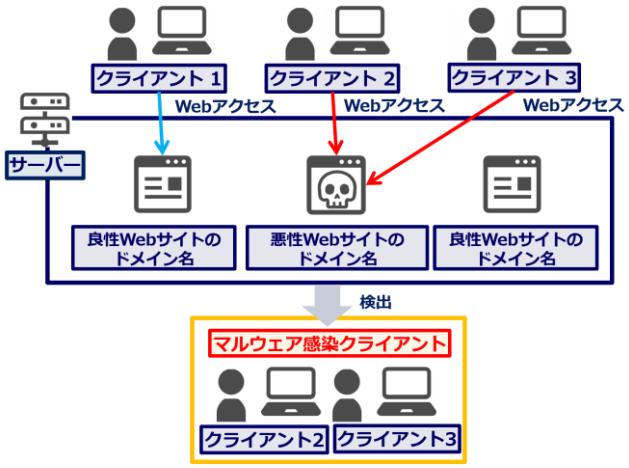


図 6 マルウェアに感染したクライアントの検出方法

一般にマルウェアは，感染を拡大させるために多数の不正 Web サイトへアクセスを試みる．そのため，不正 Web サイトは，同時に複数のマルウェア感染クライアントからアクセスが行われている可能性が高い[10]．文献[10]は，悪性のドメイン名を持つ Web サイトにアクセスしているクライアントを検出する方法を提案している．検出されたクライアントをマルウェア感染クライアントと呼ぶ．マルウェア感染クライアントが頻繁にアクセスしている Web サイトは，不正 Web サイトであると推定できる．提案方法では，このマルウェア感染クライアントの挙動を確認し，マルウェア感染クライアントがアクセスしている Web サイトのドメインを検出条件に追加して用いることで，既知の不正 Web サイトを特定する．

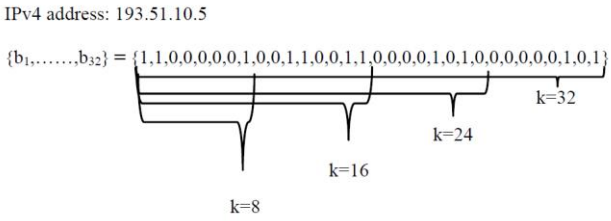
ブラックリスト，ドメイン名の特徴に基づいた検出条件，及びマルウェア感染クライアントのアクセス先のいずれにも該当しない Web サイトを未知の Web サイトと定義し，これらを判別部において不正・正規の判別対象とする．

3.4 未知の不正 Web サイトの判別手法

提案システムの判別部は，2 つのフェーズから構成される．既知の不正 Web サイトを蓄積したブラックリスト，既知の正規 Web サイトを蓄積したホワイトリストのデータから特徴ベクトルを生成し，それらを教師データとして分類器を構築する．2 つ目に，構築された分類器を用いて未知の Web サイトを判別する．

3.4.1 教師データセットを用いた分類器の構築

既知の不正 Web サイトおよび正規 Web サイトの特徴を特徴ベクトル化して，判別のための分類器を生成する．この時の特徴ベクトルの次元数が判別のコストに影響するため，次元数を削減した特徴数を用いて，Web サイトを判別する手法を提案する．図 7 に，特徴ベクトルの生成手法を示す．



k	Feature vector
8	b_k=1 (k=1, 2, 8) b_k=0 (otherwise)
16	b_k=1 (k=1, 2, 8, 11, 12, 15, 16) b_k=0 (otherwise)
24	b_k=1 (k=1, 2, 8, 11, 12, 15, 16, 21, 23) b_k=0 (otherwise)
32	b_k=1 (k=1, 2, 8, 11, 12, 15, 16, 21, 23, 30, 32) b_k=0 (otherwise)

図 7 特徴ベクトルの生成

まず，ホワイトリスト，ブラックリストのデータから構成

された教師データセットに含まれるすべての IP アドレスを 2 進数表現のビット列に変換する. すべてのビット列は, k 次元ベクトル$\{b_1, \ldots\ldots, b_k\}$として表される. 各 IP アドレスの IP アドレスクラスに応じて 3 種類の特徴ベクトルとして生成する. IP アドレスクラス A の場合は, 8 次元の特徴ベクトル, IP アドレスクラス B の場合は, 16 次元の特徴ベクトル, IP アドレスクラス C の場合は, 24 次元の特徴ベクトルとして生成する. 悪性 IP アドレスから生成された特徴ベクトルには「1」, 良性 IP アドレスから生成された特徴ベクトルには「0」とラベルを付ける. 表 1 は, 教師データセットの特徴ベクトルにラベルを付ける例を示している.

表 1 教師データセットの例

IP address	Feature vector	Label
193.51.10.5	1,1,0,0,0,0,0,1,0,0,1,1,0,0,1,1	1
10.10.10.10	0,0,0,0,1,0,1,0,0,0,0,0,0,1,0,1	1
203.4.12.89	1,1,0,0,1,0,1,1,0,0,0,0,0,1,0,0	0
…	…	…

本稿では, パターン識別法の 1 つである SVM（Support Vector Machine）を用いた判別を行っている. J. Ma らは, SVM を用いることにより, 不正 Web サイトの高精度な検出が可能であることを示している[5]. 提案システムは, 上記の次元の異なる特徴ベクトルに基づいて, 各 IP アドレスクラスの分類器を構築する.

3.4.2 分類器を用いた未知 IP アドレスの判別

検出部から渡された未知 Web サイトの IP アドレス（未知 IP アドレス）は, フェーズ 1 で構築された分類器によって良性と悪性に判別される. 図 8 に, 3 つのステップをもつ判別手法の概要を示す.

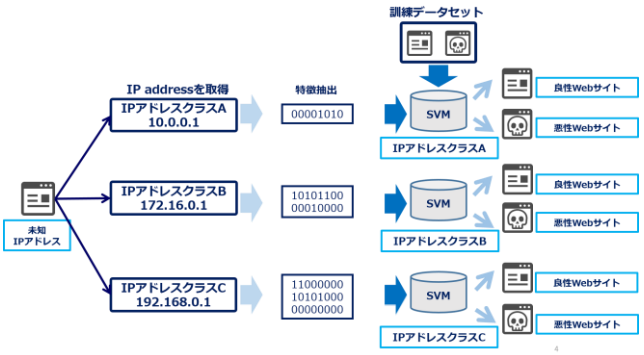


図 8 判別手法

判別部は, 検出部から渡された IP アドレスから特徴ベクトルを生成する. この特徴ベクトルを, 判別部のフェーズ 1 で構築された分類器によって良性または悪性として判別する.

4. 評価実験

4.1 評価実験の概要

提案されたシステムの目的は, 未知の Web サイトを悪性と良性に判別することである. 提案システムで, IP アドレスクラス A, B, C に属する各 IP アドレスが, それぞれのアドレスクラスのネットワークアドレス部である上位 8 ビット, 16 ビット, 24 ビット部分を用いて判別したときに高い精度を示すことを確認する. 提案手法の有効性を確認するために, 判別部のフェーズ 1 で構築された分類器を精度, 適合率, 再現率の 3 つを評価指標として評価した. 本稿では, 悪性 IP アドレスを正しく悪性 IP アドレスと判別した数を表す真陽性（TP）, 良性 IP アドレスを誤って悪性 IP アドレスと判別した数を偽陽性（FP）, 良性 IP アドレスを正しく良性 IP アドレスと判別した数を真陰性（FN）, 悪性 IP アドレスを誤って良性 IP アドレスと判別した数を偽陰性（TN）とする. このときの精度, 適合率, 再現率をそれぞれ下記の計算式で求める.

$$精度 = (TP + TN)/(TP+ TN + FP+FN)$$
$$適合率 = TP/(TP+ FP)$$
$$再現率 l = TP/(TP+ FN)$$

良性および悪性データセットは, Malware Workshop Datasets[14]から取得した IP アドレスを用いている. 悪性 IP アドレスと良性 IP アドレスの比率は 8 : 2, 5 : 5, 2 : 8 の 3 パターンを作成した. 悪性の教師データセットは, マルウェア検体を収録したボット観測データ群 CCC DATAset（2008 年～2011 年）と Web 感染型マルウェアデータ D3M（2010 年～2015 年）のデータをもとに作成した. また, 良性の教師データセットは, Alexa Top Global Sites[15]の 50,000 件（2016）のデータとホワイトデータセット NCD in MWS Cup(2014)のデータをもとに作成した. IP アドレスクラス A の IP アドレスには, 悪性 IP アドレス 49164 個との良性 IP アドレス 40667 個, IP アドレスクラス B の IP アドレスには, 悪性 IP アドレス 3523 個と良性 IP アドレス 10735 個, IP アドレスクラス C の IP アドレスには, 悪性 IP アドレス 75000 個と良性 IP アドレス 14288 個を用いた. 実験は 5 分割交差検定を用いて評価した.

4.2 実験結果

IP アドレスのうち判別に用いる部分として, Case1 を上位 8 ビット, Case2 を上位 16 ビット, Case 3 を上位 24 ビ

ット，Case4 を上位 32 ビットと定義し，それぞれの場合について判別結果を評価した．IP アドレスクラス A の判別結果を表 2 に示す．

表 2 IP アドレスクラス A の実験結果

	精度	適合率	再現率
Case1(k=8)	84.06079	89.86656	90.25327
Case2(k=16)	83.74358	89.79705	89.89365
Case3(k=24)	83.76079	89.8188	89.89058
Case4(k=32)	83.87882	89.95386	89.8875

精度と再現率は IP アドレスクラス A のネットワークアドレス部のみを特徴ベクトルとして用いた Case1 で最高値を示した．適合率は，Case4 で最高値を示した．関連研究[11]の精度が 74.7~75.1%であったのに対し，提案手法の精度は 84%であり，関連研究の精度を上回る結果が得られた．次に，IP アドレスクラス B の実験結果を表 3 に示す．

表 3 IP アドレスクラス B の実験結果

	精度	適合率	再現率
Case1(k=8)	83.54143	87.85521	92.15756
Case2(k=16)	81.69694	88.16298	89.07026
Case3(k=24)	82.94552	88.61024	90.27679
Case4(k=32)	83.25766	88.76131	90.5252

精度と再現率は，Case1 で最高値を示した．一方，適合率は，Case4 で最高値を示した．したがって，IP アドレスクラス B のネットワークアドレス部のみを特徴ベクトルとして用いた Case2 で最高値を達成することができていない，また，関連研究[11]の精度が 84.6~86.2%であったのに対し，提案手法の精度は 81.6~83.5%であり，関連研究の精度を下回る結果が得られた．最後に，IP アドレスクラス C の実験結果を表 4 に示す．

表 4 IP アドレスクラス C の実験結果

	精度	適合率	再現率
Case1(k=8)	81.78191	87.19137	90.52493
Case2(k=16)	81.20101	86.9819	89.965
Case3(k=24)	81.243	87.00101	90.0
Case4(k=32)	81.222	86.56467	90.58618

精度と適合率は，Case1 で最高値を示した．一方，再現率は，Case4 で最高値を示した．したがって，IP アドレスクラス C のネットワークアドレス部のみを特徴ベクトルとして用いた Case3 で最高値を達成することができていない，また，関連研究[11]の精度が 85.1~88.5%であったのに対し，提案手法の精度は 81.2~81.7%であり，関連研究の精度を下回る結果が得られた．表 5 に，IP アドレスクラス C において，悪性 IP アドレスと良性 IP アドレスの比率が 5:5 であるときの実験結果を示す．

表 5 IP アドレスクラス C の実験結果(5:5)

	精度	適合率	再現率
Case1(k=8)	67.25224	65.24805	73.82419
Case2(k=16)	67.25224	65.24805	73.82419
Case3(k=24)	67.28024	65.31068	73.71221
Case4(k=32)	66.03443	65.12211	69.05095

精度と適合率は，IP アドレスクラス C のネットワークアドレス部のみを特徴ベクトルとして用いた Case3 で最高値を示した．再現率は，Case1 と Case2 で最高値を示した．また，Case4 で，精度，適合率率および再現率が最低値を示した．

4.3 考察

IP アドレスクラス A は、ネットワークアドレス部のみの判別で十分な精度を保つことができるため，提案手法は有効であるといえる．一方，IP アドレスクラス B と IP アドレスクラス C では，IP アドレスクラス A と比較して，全体的に精度が低い結果となった．考えられる要因として，2 つ挙げられる．まず，教師データセットの IP アドレス数が極端に少ないため，特徴がはっきり表れていない可能性がある点が挙げられる．次に，悪性 IP アドレスの特徴に経年変化が起きている可能性がある点が挙げられる．そこで，悪性 IP アドレス利用頻度のデータを元に，IP アドレスクラスの利用状況を分析した．

図 9 に，CCC DATAset から収集した IP アドレス数が比較的多い 2008 年から 2011 年を対象に悪性 IP アドレスの利用頻度をグラフ化したものを示す．

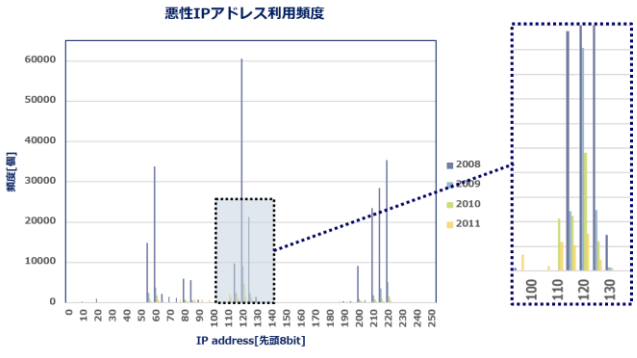


図 9 IP アドレス分布

(IP アドレス上位 8 ビットの 120 付近拡大)[2008-2011]

IP アドレス上位 8 ビットの 120 付近に着目すると, 2008 年から 2011 年にかけて悪性 IP アドレスの利用頻度が減少していることが確認された.

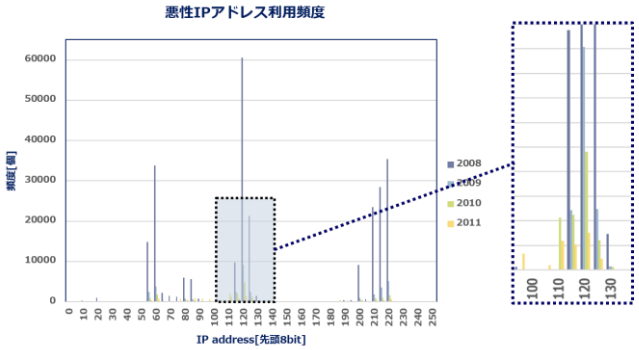


図 10 IP アドレス分布

(IP アドレス上位 8 ビットの 110 付近拡大)[2008-2011]

一方，図 10 は，IP アドレス上位 8 ビットの 110 付近に着目したものである．これによると，2008 年から 2009 年まで悪性 IP アドレスの利用頻度が 0 件であるのに対し, 2010 年から利用頻度が増加していることが確認された.

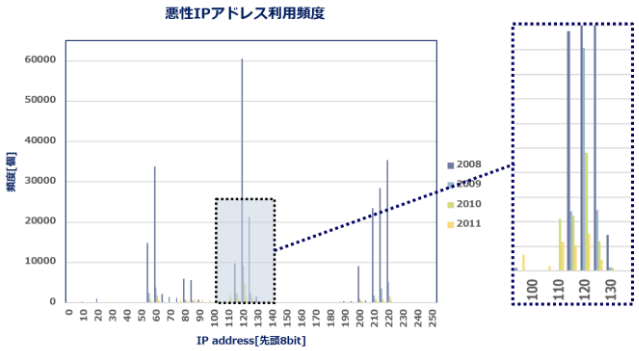


図 11 IP アドレス分布

(IP アドレス上位 8 ビットの 200~220 付近拡大)[2008-2011]

また，図 11 のように IP アドレス上位 8 ビットの 200 から 220 付近に着目すると，2008~2011 年にかけて悪性 IP アドレスの利用頻度が減少していることが確認された.

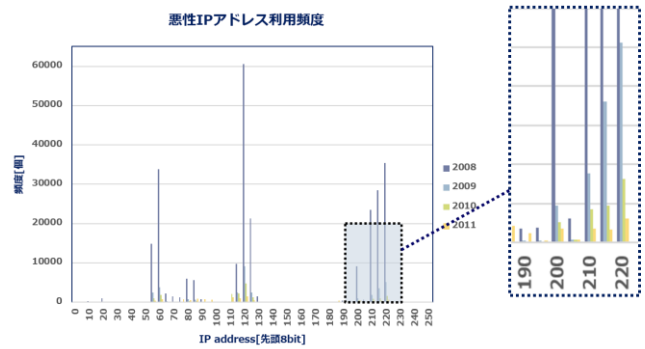


図 12 IP アドレス分布

(IP アドレス上位 8 ビットの 170~180 付近拡大)[2008-2011]

さらに, 図 12 のように IP アドレス上位 8 ビットの 170~180 付近に着目すると，2010 年，2011 年から悪性 IP アドレスの利用頻度が増加している傾向が見られ，全体的に利用頻度に変化が生じていると考えられる.

4.4 追加実験

4.3 の考察より，各 IP アドレスクラスにおける特徴を年度別に抽出して判別に用いることで，利用頻度の経年変化に対応した高精度な判別が実現できるかどうかについて確認するために追加実験を行った．4.2 節で，精度が低い結果となった IP アドレスクラス B と IP アドレスクラス C の IP アドレスを実験対象とする. 実験に用いるデータセットは，IP アドレス数が比較的多い CCC DATAset の 2008 年と 2009 年から取得した IP アドレスから作成した.

4.4.1 追加実験結果

図 13 に，IP アドレスクラス B における追加実験結果を示す.

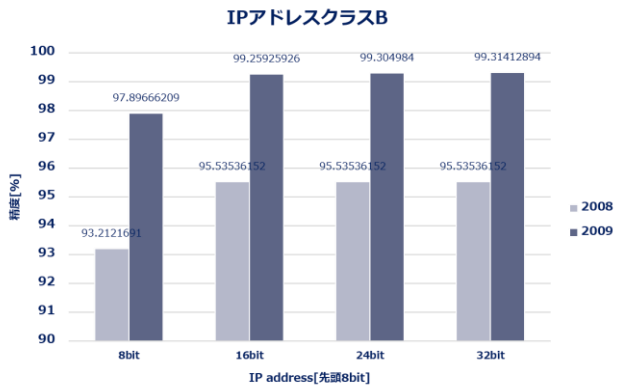


図 13 IP アドレスクラス B の追加実験結果

判別精度を見ると, 2008 年から 2009 年までの判別精度が，90%超まで向上した．図 14 に，IP アドレスクラス C における追加実験結果を示す.

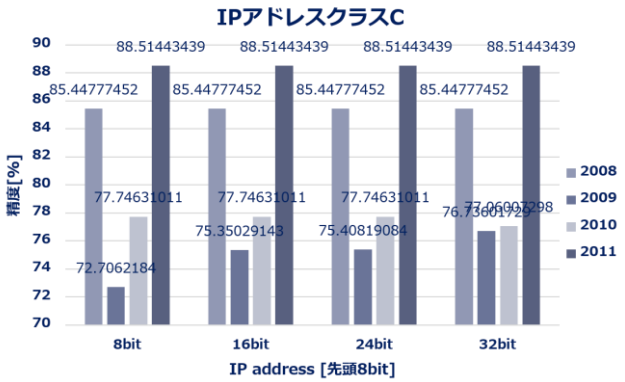


図 14 IP アドレスクラス C の追加実験結果

2008 年と 2011 年の判別精度が 85%割まで向上したことが確認できた.

4.4.2 追加実験の考察

追加実験により, 悪性 IP アドレスの特徴が経年変化していることが判明した. これらの情報を踏まえて, 悪性 IP アドレスの特徴の経年変化を考慮した判別を行うことにより, 判別精度を向上させることができた. これらの結果から, 悪性 IP アドレスデータ群（ブラックリスト）をそのまま蓄積して使用する方法では, 精度を保つことが困難であったが, IP アドレスクラスごとに教師データの状態を変えることで, 年度別に異なる特徴を維持することができ, 判別精度が保たれたと考えられる. したがって IP アドレスクラスの中でも, 変動のあるネットワークアドレス群の範囲を見つけることで, 未知の Web サイトに対応していくことが可能であると考えられる.

5. 今後の課題

まず, IP アドレスクラス B は極端に IP アドレス数が少ないため, 教師データが少ない状態でも判別精度を保つ方法を検討する. 次に, IP アドレスクラス C の更なる特徴分析を行う. 単純な年度比較ではなく, IP アドレスの配布ポリシーの変化があったタイミングなどでさらに細かく区切り, IP アドレスクラス C に適する判別方法を検討する. 最後に, 悪性 IP アドレスと良性 IP アドレスの特徴を最新に保つ手法の検討することにより, 将来的には, データを全て差し替えることなく, 限られた範囲のデータのみ更新できるようになると考えている.

6. まとめ

本稿は, 未知の Web サイトを検出し, ネットワークアドレス部のみを用いることで, 判別に必要なコストを軽減しながら, 未知の Web サイトを判別する手法を提案する. 5 分割交差検定により, 各 IP アドレスクラスについて提案された分類器の判別精度を評価した. 評価実験の結果, IP アドレスクラス A では, 高精度な判別ができ, 提案した判別手法の有効性を確認できた. 一方で, IP アドレスクラス B および IP アドレスクラス C における判別精度はそれほど高くなかった. また, 悪性と判別された IP アドレスは, 二度と良性 IP アドレスに復帰することができないという問題点がある.

今後は, より有効な判別を行うために, 各 IP アドレスクラスにおけるデータの更なる分析を行い, より効果的な教師データの生成手法の検討を行う必要がある. また, IP アドレスクラスごとにデータ数に差がある場合でも判別精度を保つ方法を検討する. 最後に, IP アドレスを用いた判別手法において, 一度悪性と判別された IP アドレスが破棄され, 新たにそのアドレスに正規 Web サイトが構築された場合に, どのような手続きで良性 IP アドレスとみなすように判別を変更するかを考慮する必要がある. 以上の項目を検討した上で, 提案システムの精度向上が達成されるかどうかを確認する必要がある.

参考文献

[1] 独立行政法人情報処理推進機構:2016 年版 情報セキュリティ 10 大脅威, <https://www.ipa.go.jp/files/000051691.pdf> [参照 2017-5-31].

[2] 警察庁広報資料: 平成 26 年中のインターネットバンキングに係る不正送金事犯の発生状況等について <https://www.npa.go.jp/cyber/pdf/H270212_banking.pdf> [参照 2017-1-17].

[3] トレンドマイクロ: Web レピュテーション, <http://www.trendmicro.co.jp/why-trendmicro/spn/features/web/index.html> [参照 2016-8-11].

[4] M. A. Rajab, L. Ballard, N. Jagpal, P. Mavrommatis, D. Nojiri, N. Provos, and L. Schmidt,"Trends in circumventing web-malware detection, Google Technical Report, 2011.

[5] シスコシステムズ: 侵入防御システム（IPS：Intrusion Prevention System） <https://www.cisco.com/c/ja_jp/about/technology-commentary/tech-2006/intrusion-prevention-system-ips-intrusion-prevention-system.html>[参照 2017-8-27].

[6] J. Ma, L. K. Saul, S. Savage, and G. M. Voelker, :Beyond blacklists: learning to detect malicious web sites from suspicious urls, Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD'09), pp. 1245–1254, 2009.

[7] 劉亦晨: DNS 情報による悪意のあるサイトの検出法, 2012 年度 早稲田大学大学院 基幹理工学研究科 情報理工学専攻 修士論文, 2012.

[8] 日立ソリューションズ: 情報セキュリティブログ, <http://securityblog.jp/words/2898.html> [参照 2017-1-17].

[9] L. Bilge, E. Kirda, C. Kruegel, and M. Balduzzi,:Exposure Finding Malicious Domains Using Passive DNS Analysis, Proceedings of the 18th Annual Network & Distributed System Security Symposium (NDSS Symposium 2011), 2011.

[10] 田中晃太郎, 長尾篤, 森井昌克: DNS ログからの不正 Web サイト抽出についてー解析手法とその匿名化ー, 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.132-138 (2013).

[11] D. Chiba, K. Tobe, T. Mori, and S. Goto, :Detecting Malicious Websites by Learning IP Address Features, Proceedings of the IEEE/IPSJ 12th International Symposium on Applications and the Internet(SAINT2012), pp.29-39, 2012.

[12] 千葉大紀,森達哉, 後藤滋樹: 悪性 Web サイト探索のための優先巡回順序の選定法, コンピュータセキュリティシンポジウム 2012 論文集, Vol.2012, No.3, pp.805-812 (2012).

[13] 千葉大紀,森達哉, 後藤滋樹: 悪性 Web サイト探索のための優先巡回順序の選定法, コンピュータセキュリティシンポジウム 2012 論文集, Vol.2012, No.3, pp.805-812 (2012).

[14] 高田雄太, 寺田真敏, 村上純一, 笠間貴弘, 吉岡克成, 畑田光弘: マルウェア対策のための研究用データセット～MWS Datasets 2016～, 情報処理学会研究報告, Vol.2016-CSEC-74, No.17, pp. 1-8, 2016.

[15] Alexa Internet, Inc.: The top 500 sites on the web," <http://www.alexa.com/topsites> [参照 2017-8-27].