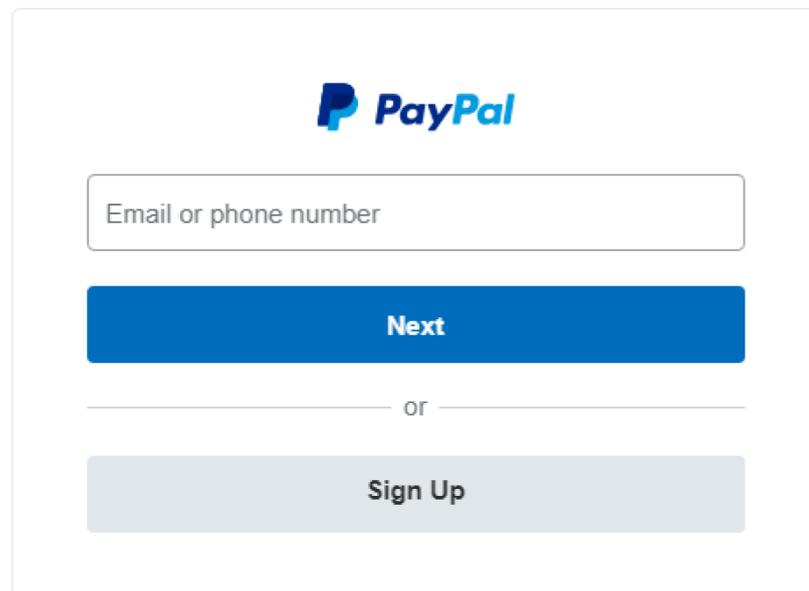


# ドメイン情報とHTTPレスポンスヘッダに 基づくフィッシングサイトの識別と評価

桜井啓多

# 正規サイトとフィッシングサイト

## 正規サイト

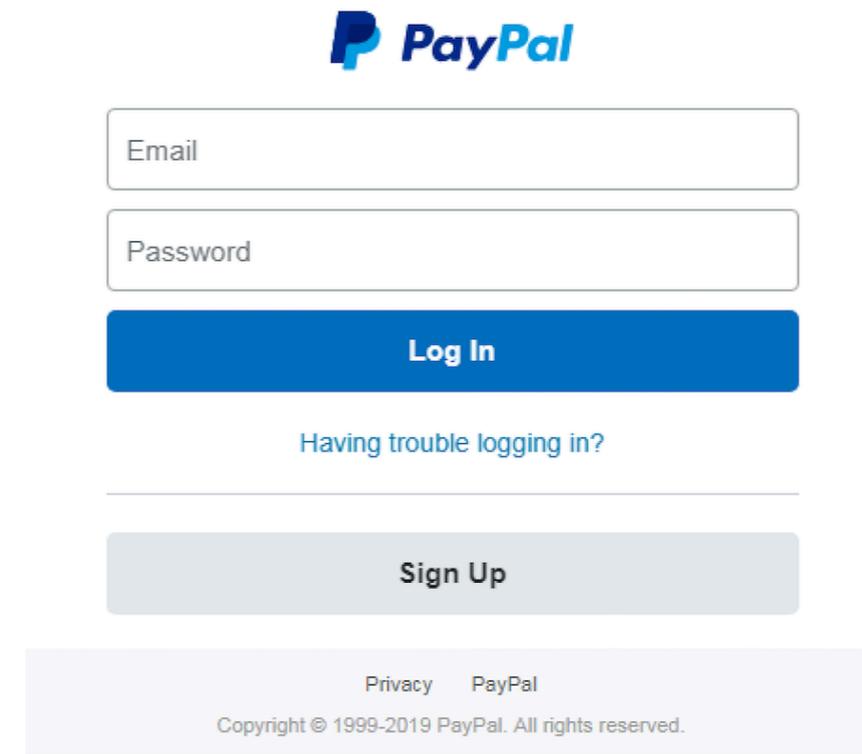


The screenshot shows the legitimate PayPal sign-up page. At the top is the PayPal logo. Below it is a text input field labeled "Email or phone number". Underneath the field is a blue button labeled "Next". Below the "Next" button is a horizontal line with the word "or" in the center. Below the line is a grey button labeled "Sign Up".

[Contact Us](#) [Privacy](#) [Legal](#) [Worldwide](#)

サイトの生存期間: 7134日間  
x-xss-protection: 設定あり  
x-frame-options: 設定あり

## フィッシングサイト



The screenshot shows a phishing site that mimics the legitimate PayPal sign-up page. At the top is the PayPal logo. Below it are two text input fields, one labeled "Email" and one labeled "Password". Underneath the "Password" field is a blue button labeled "Log In". Below the "Log In" button is a link labeled "Having trouble logging in?". Below the link is a grey button labeled "Sign Up". At the bottom of the page is a footer with the text "Privacy PayPal" and "Copyright © 1999-2019 PayPal. All rights reserved."

サイトの生存期間: 189日間  
x-xss-protection: 設定無し  
x-frame-options: 設定無し

# 研究背景

- フィッシングとは、攻撃者がメール等で本物のサイトと同じ様な偽のWebサイトへのリンクを送信し、騙されたユーザがそのサイトに入力してしまった情報を盗む行為である。
- フィッシング対策協議会のフィッシングレポート2018([https://www.antiphishing.jp/report/pdf/phishing\\_report\\_2018.pdf](https://www.antiphishing.jp/report/pdf/phishing_report_2018.pdf))によると、2017年下半期に届け出されたフィッシングサイトのURL件数は約5000件に上る。

(<https://www.antiphishing.jp>)

フィッシング対策協議会  
Council of Anti-Phishing Japan

◁ サイトマップ ▷ プラ  
サイト内を

◻ HOME ◻ ニュース ◻ 報告書類 ◻ 消費者の皆様へ ◻ サービス事業者の皆様へ ◻

Welcome to Council of Anti-Phishing Japan

フィッシング対策協議会は 2005 年 4 月に発足いたしました。フィッシング詐欺に関する事例情報、技術情報の収集及び提供を中心に行うことで、日本国内におけるフィッシング詐欺被害の抑制を目的として活動しております。

「技術・制度検討 WG 報告会」を開催致します。お申し込みは [こちら](#)

◻ フィッシングに関するニュース

緊急情報 ▶ 緊急情報一覧

- ▶ 2019年01月17日 三井住友銀行をかたるフィッシング (2019/01/17)
- ▶ 2019年01月15日 Amazonをかたるフィッシング (2019/01/15)
- ▶ 2018年12月25日 三井住友銀行をかたるフィッシング (2018/12/25)

# 研究方法

- フィッシングサイトの検出を支援するため,対象のサイトが安全か否かを判定するプログラムを作成した.
- 判定を行うための指標として,フィッシングサイトのドメインをphishtankより取得し,そのドメインの情報を100件分と企業のホームページなどの正規サイト100件の計200件のデータを収集した.

フィッシングサイト	100件
正規サイト	100件

先行研究: “プロキシを利用した HTTP リクエスト解析によるフィッシングサイト検出システムの提案“

(“https://www.phishtank.com”)

The screenshot shows the PhishTank website interface. At the top, there is a navigation bar with the PhishTank logo and the tagline "Out of the Net, into the Tank." Below the navigation bar, there is a main content area with a search form and a list of recent submissions. The search form has a text input field with "http://" and a button labeled "Is it a phish?". Below the search form, there is a section titled "Recent Submissions" with a table of data. The table has columns for ID, URL, and Submitted by. The table contains several rows of data, including IDs like 5921357, 5921356, 5921353, 5921352, 5921351, and 5921350, and URLs like http://dfergt.000webhostapp.com/New\_2019/New\_2018/... and http://www.cathayonlineservice.com/. The Submitted by column lists names like Micha and buaya. To the right of the main content area, there are two informational boxes: "What is phishing?" and "What is PhishTank?".

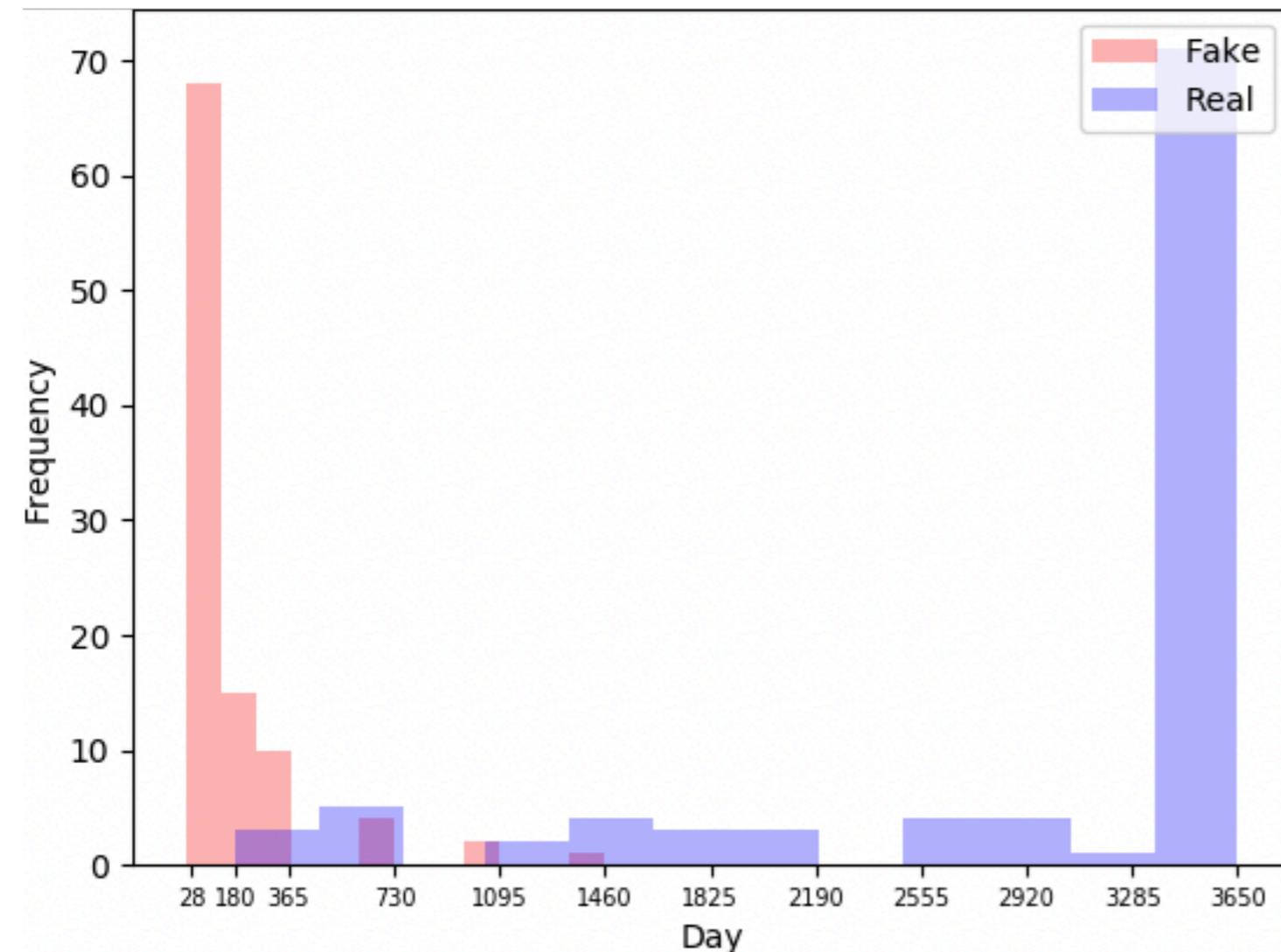
ID	URL	Submitted by
5921357	http://dfergt.000webhostapp.com/New_2019/New_2018/...	Micha
5921356	http://www.cathayonlineservice.com/	Micha
5921353	http://raptorsshield.com/reba/vad/var.html	buaya
5921352	http://raptorshelda.com/var/crotes/Rev.html	buaya
5921351	http://durgapublishers.org.in/jl/00/00/00/980014/i...	buaya
5921350	http://durgapublishers.org.in/jl/00/00/00/980014/i...	buaya

# 収集した特徴量の概要

特徴	内容	例	重み
country	サイトを運用しているIPアドレスの国コード	US	0.4
domain interval	アクセス日時 - ドメイン作成日時	98	0.45
domain lifetime	ドメイン期限日 - ドメイン作成日時	6	0.45
x-xss-protection	HTTPレスポンスヘッダ.XSS攻撃を防止するための設定	TRUE	0.33
x-frame-options	HTTPレスポンスヘッダ.クリックジャッキング攻撃を防止する設定	TRUE	0.33
x-content-type-options	HTTPレスポンスヘッダ.コンテンツの内容を見ない様にする事でXSS攻撃のリスクを減らす設定	TRUE	0.33
content-security-policy	HTTPレスポンスヘッダ.UAが読み込みを許可されたリソースを管理出来る様にするための設定	TRUE	0.33

# ドメインに関するデータ①

”domain interval”: ドメイン作成日 - アクセス日

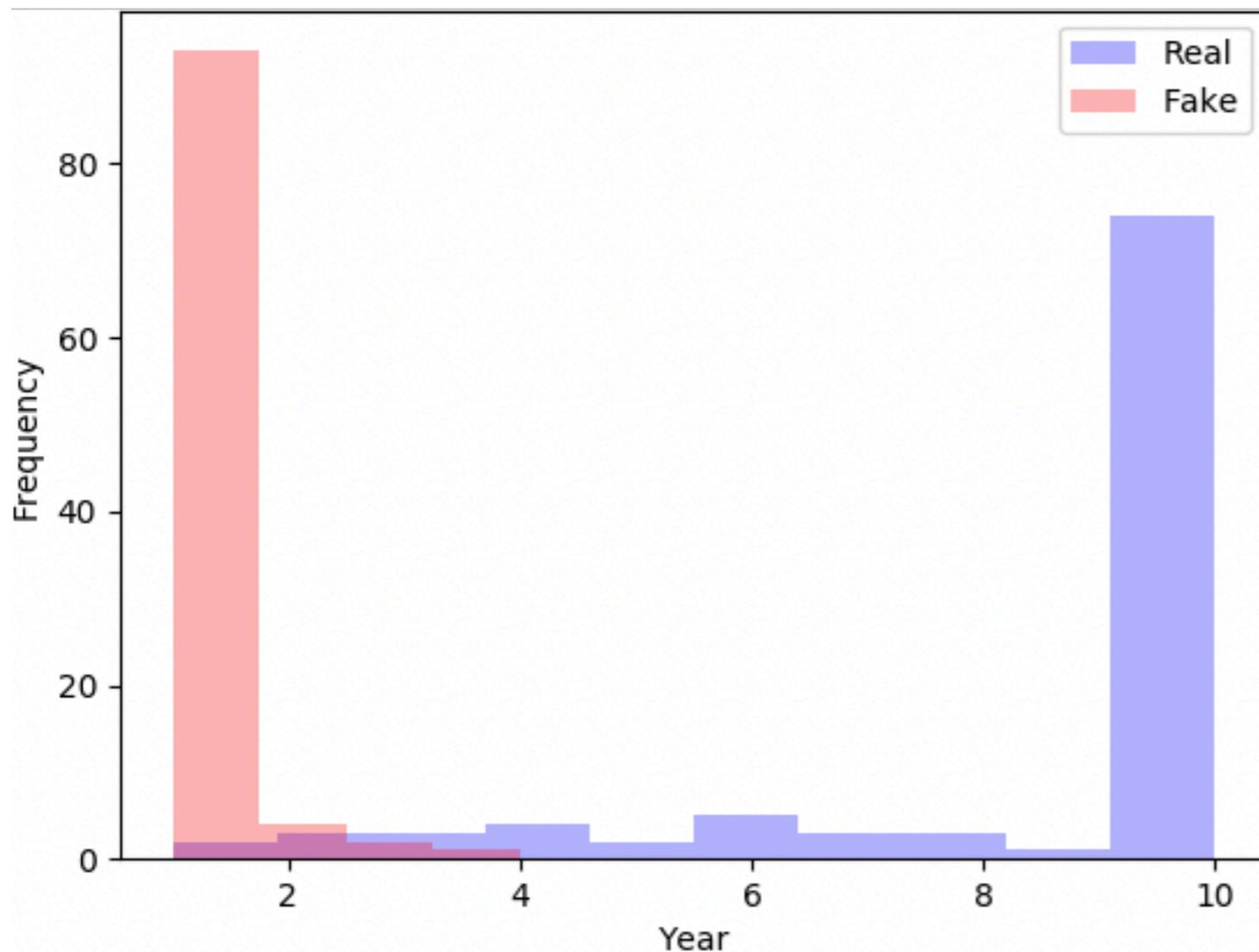


	正規サイト [日]	フィッシング サイト[日]
平均値	3089	135
中央値	3650	14
最頻値	3650	7

正規サイトは長く,フィッシングサイトは短い

# ドメインに関するデータ②

”domain lifetime”: ドメイン期限日 - ドメイン作成日



	正規サイト [年]	フィッシング サイト[年]
平均値	8.6	1.1
中央値	10	1
最頻値	10	1

フィッシングサイトは短い運用  
正規サイトは長い運用

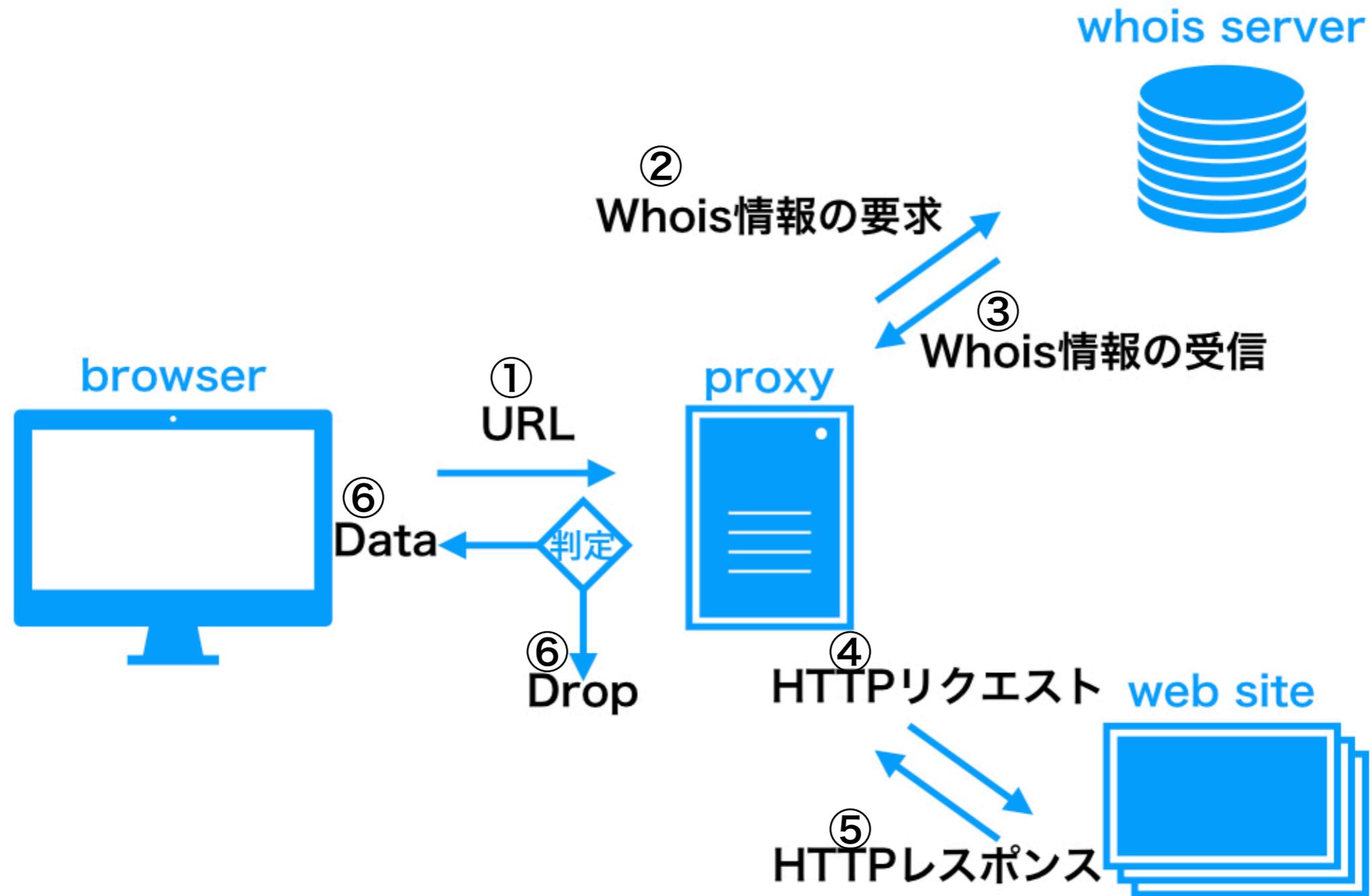
# HTTPレスポンスヘッダに関するデータ

HTTPレスポンスヘッダ名	正規サイト	フィッシングサイト
x-frame-options	70	4
x-content-type-options	58	4
x-xss-protect	56	4
content-security-policy	25	2

# 国に関するデータ

国名	正規サイト	フィッシングサイト
アメリカ	56	40
日本	40	2
アイルランド	2	0
台湾	0	20
オランダ	0	6
イギリス	0	4
パナマ	0	4
ロシア	0	4
カナダ	0	3
韓国	0	2
モロッコ	0	2
フランス	0	2
その他(1件のみ)	2	9
不明	0	2

# システム概要



# 判定方法

安全か判定する式

$$\sum_{i=0}^7 w_i * d_i > \theta$$

$w_i$ : 特徴量に与えた重み

$d_i$ : カテゴリ化して割り当てた定数

$\theta$ : 閾値

(例) domain\_intervalの定数 $d_i$

期間	定数
1週間以下	0
1ヵ月以下	10
2年以下	50
4年以下	70
6年以下	80
7年以上	100

# 判定結果

FN	5.8
FP	2.2

- FN: 正規サイトをフィッシングサイトと判断した件数
- FP: フィッシングサイトを正規サイトと判断した件数

$$Accuracy: \frac{TP + TN}{TP + TN + FP + FN}$$

全体正答率Accuracy: 95%

# まとめ

- フィッシングサイトのドメインとHTTPレスポンスヘッダに関するデータは正規サイトのデータと異なり,特有の傾向がある事が分かった.
- ドメイン,HTTPレスポンスヘッダ,IPアドレスの割り当て国を用いた識別は正答率95%となり,有用性があった.
- フィッシングサイトを作成するためのキット等に焦点を当て,その特徴を求める事でより検出精度を向上する事を今後の課題とする.