ドメインの WHOIS 構造を用いた悪性ドメインの判別手法

久山真宏[1]　佐々木良一[1]

概要：近年流行している標的型攻撃では，マルウェアに感染した後に C&C サーバとの間で様々な通信を行う．そのため，出口対策として C&C サーバの通信を監視することにより，被害を発見することが出来る．C&C サーバを判別する手法は複数存在している．しかし，その多くは実際に攻撃者が準備したサーバ類に対して通信を行う必要があり，分析の過程で攻撃者に気づかれてしまう可能性がある．そこで，本研究では，C&C サーバを含めた攻撃者側に検知されにくい WHOIS 情報からメールアドレスと登録期間から特徴点を抽出し，教師あり機械学習を用いて C&C サーバの判別を行う手法を提案する．そして，この手法に実データを適用し C&C サーバの判別を行った結果，約８８％と比較的高い検知率を得ることができ，有効性の見通しを得ることができたので報告する．

Method for detecting malicious domain using WHOIS features

MASAHIRO KUYAMA[1]　RYOICHI SASAKI[1]

1. はじめに

近年，標的型攻撃による被害が問題になっている[1]．標的型攻撃とは，金銭や知的財産等の秘密情報の不正な取得を目的として，特定の企業や組織を標的にしたサイバー攻撃の一種である．ドライブバイダウンロード攻撃や，メールなどに添付されたマルウェアに感染することによって，情報の搾取や破壊活動が行われる．

日本では，国内の大手重工メーカや衆議院，日本年金機構などにおいて，標的型攻撃の被害に遭い，実際にニュースになるほどの重大なインシデントに繋がっている．標的型攻撃の対策が求められており，多層防御として入口対策や出口対策が求められる．

標的型攻撃では，初期侵入段階において，マルウェアに感染させる．その後，C&C サーバ（Command and Control server）と呼ばれる中継サーバを介して，遠隔操作を行うことにより侵入範囲の拡大や，情報摂取等を行う（図 1）．そのため，出口対策として C&C サーバの通信を監視することにより，被害を発見することが出来る[2,3]．しかし，C&C サーバとの通信を監視するには，事前に C&C サーバを特定している必要がある．そのため，特定されていない C&C サーバが用いられた場合には監視対象から外れているため，被害に気づきにくくなる．

図 1　標的型攻撃の流れ

C&C サーバを特定するにはマルウェアおよびその通信を解析する必要があるが，調査・研究の過程で C&C サーバへアクセスを行うことがある．これでは，攻撃者に解析していることを検知され，対策前に回避されてしまう可能性がある．

そこで，本研究では，C&C サーバなどの攻撃者が準備したサーバ類にアクセスせずに得られる情報を用いて，C&C サーバの検知を試みる．C&C サーバにアクセスせずに得られる情報として，ドメインの WHOIS 情報を用いる．特に，WHOIS 情報の中でも偽装が困難だと考えられるドメインの登録日および有効期限，メールアドレスから特徴点を抽出し，機械学習を用いて検知を試みる．

2. 先行研究

2.1 LIFT

現在筆者らは，標的型攻撃に対してインシデント発生時にネットワークログ等のデータを適切に利用し，人工知能

1 東京電機大学
Tokyo Denki University

（AI ともいう）を用いて自動的な応急対応を可能とするとともに，運用者が適切な対策をとれるようにするための LIFT（Live and Intelligent Network Forensic Technologies）システムの開発を行っている[4,5]．さらに，防御側だけではなく，攻撃側にも人工知能を利用して Beyond the Attackers を実現し，Proactive な対策を実現する Supper-LIFT システム[6]を構想している（図 2）．

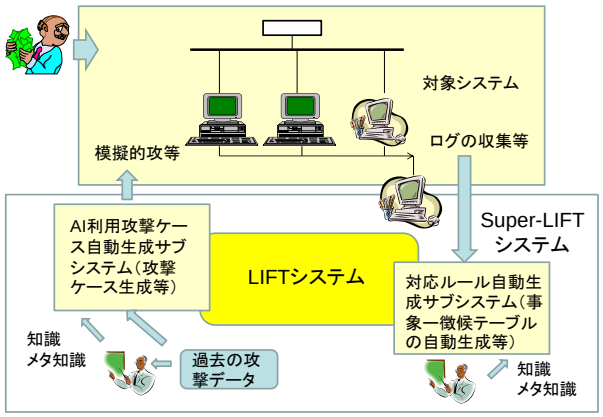

図 2　Supper-LIFT システムの研究

Supper-LIFT では，サイバー攻撃に関する情報の収集・分析の一環として，C&C サーバの分析・調査を行っている[7]．

Supper-LIFT の活動から得られた知見をもとに，既知の C&C サーバに用いられているドメインの WHOIS に紐づく他のドメインを抽出することで，新たな C&C サーバを発見する手法の提案を行った[7]．今回，この手法より得られた知見をもとに C&C サーバの判別手法を提案する．

2.2 多段追跡システム

2009 年，ボットネットによる被害が増加したことによる問題に対して，ボット PC や C&C サーバ，攻撃者の特定を目的とした多段追跡システムの研究を行っている[8]．本研究における C&C サーバの推定手法として，C&C サーバと正規のサーバの DNS 情報をもとに数量化理論 2 類を用いて推定を行っている．

本手法は，2009 年当時の検知率は 96.5%であった．しかし，継続的調査を行ったところ，年々検知率が下がり，その度に手法の見直しを行い，改善させた（表 1）．

表 1　多段追跡システム継続的調査による検知率

モデル	検知率（%）				
	2009	2010	2011	2013	2014
2009	96.5	85.0	76.5	-	-
2011	-	-	95.2	42.5	-
2013	-	-	-	80.3	80.8

検知率が下がる原因として，C&C サーバが解析されていることを攻撃者が知ることにより，攻撃者が解析対策を行っていることが考えられる．特に，本手法では，推定に用いる情報として DNS 情報を用いるため，C&C サーバに直接接続することはなくとも，C&C ドメインを管理している DNS サーバ（C&C サーバや C&C ドメインを管理している DNS サーバなど，攻撃者が直接もしくは間接的にでも管理している可能性があるサーバ郡のことを攻撃者の準備したサーバ郡とする）へは接続を行うため，攻撃者に対して解析していることを知らせてしまう危険性がある．

2.3 関連研究

C&C サーバを判別するための先行研究として，制御通信のペイロードに含まれる文字列などの特徴を分析することで検知を行う手法[9][10]，ドメイン情報や外部リポジトリから取得した情報を併用して，RIPPER と呼ばれるデータマイニング手法を用いて検知を行う手法[11]，テイント解析技術を応用したマルウェア解析を実施することで通信データの改ざんを検知し，C&C サーバを特定する手法[12]，WHOIS と DNS の情報から未知の悪性ドメインを推定する手法[13]，URL の特徴や DNS，WHOIS，地理的な情報から機械学習を用いて検知する手法[14]，既知の悪性 Web サイトのコンテンツや WHOIS などの情報から検索エンジンを利用して未知の悪性ドメインを推定する手法[15]等がある．しかし，これらは C&C サーバの特定に際して，攻撃者が準備したサーバ類に対して実際に通信を行う必要があり，攻撃者に解析していることを知られる恐れがある．それにひきかえ，本手法では特徴抽出に用いるのが WHOIS 情報であるため，攻撃者が準備したサーバ類と通信する必要が無く，攻撃者に解析を行っていることを知られにくく出来るという利点がある．

3. 提案手法

3.1 提案概要

提案手法は，C&C サーバのドメインに着目した検知手法である．本手法は， WHOIS 情報から特徴点を抽出し，機械学習を用いて C&C サーバかどうかを判別する．用いる情報は，WHOIS 検索により容易に取得可能であり，実際に攻撃者が準備したサーバ類と通信をせずに取得できる情報である．

今回，悪性かどうかの 2 クラスのパターン識別として機械学習を利用する．そのため，事前準備として，機械学習における訓練モデルを構築する．

3.2 訓練モデルの構築

訓練モデルの構築にあたり，まず悪性ドメインとして C&C サーバのドメイン（C&C ドメイン）と，通常の無害なドメイン（ノーマルドメイン）を準備する．そこから，各ドメインの WHOIS 情報を取得し，特徴を抽出する．

抽出した特徴を機械学習で学習させ，訓練モデルを構築する（図 3）．実際にアクセスする際に訓練モデルを用いてドメインの評価を行い，C&C サーバであるかどうか判別する．

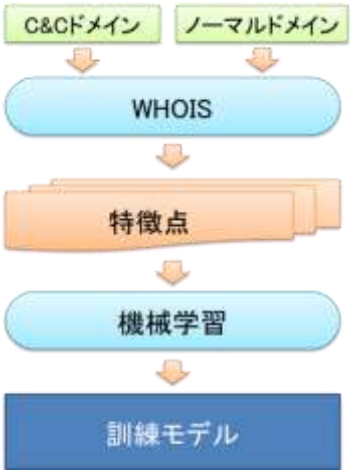

図 3　提案手法概要

- **各ドメインの準備**

まず，ノーマルドメインと C&C ドメインの 2 種類のドメインを準備する．

ノーマルドメインには，安全性が高いドメインが最適であるため，世界のアクセスランキングトップ 500 を掲載している Alexa の” The top 500 sites on the web.”[16]に載っているドメインを利用した．また，C&C ドメインには，実際のマルウェアから抽出したドメインが最適であるため，標的型攻撃での使用率の高い Emdivi，PlugX，PoisonIvy と呼ばれる 3 種類のマルウェア群[17]を解析して抽出したドメインを利用した．

マルウェアの収集にあたっては，VirusTotal[18]を用いて，キーワードに Emdivi，PlugX，PoisonIvy の種別名で検索を実施し，計 163 件のマルウェアを収集した（表 2）．

表 2　収集したマルウェアの検体数

マルウェア種別	検体数
Emdivi	50
PlugX	63
PoisonIvy	50

収集したマルウェアを LastLine[19]と呼ばれる Sandbox を用いて解析を実施．解析結果より，マルウェアが通信を行う接続先のドメイン 54 件を利用した．

- **WHOIS からの特徴抽出**

WHOIS からは一般的に以下の情報を得ることが出来る．

a)　登録ドメイン名
b)　レジストラ名
c)　ドメインが登録されている DNS サーバ名
d)　ドメインの登録年月日
e)　ドメインの有効期限
f)　ドメイン名登録者の連絡先
g)　技術的な連絡の担当者連絡先
h)　登録に関する連絡の担当者連絡先
i)　登録者への連絡窓口の連絡先

この中でも，改ざんが困難なものとして a)～e)があげられる．通常のサーバであれば，長期的に運用することからドメインの登録期間は長く，逆に標的型攻撃における C&C サーバは，標的となる組織において目的が達成されればドメインを放棄するため登録期間が短い[13,14,15]．このことに着目し，登録期間を割り出すため，d)の日数から e)の日数を引いた値を用いることとした．実際に求めた日数を 500 日刻みで示す（図 4）．

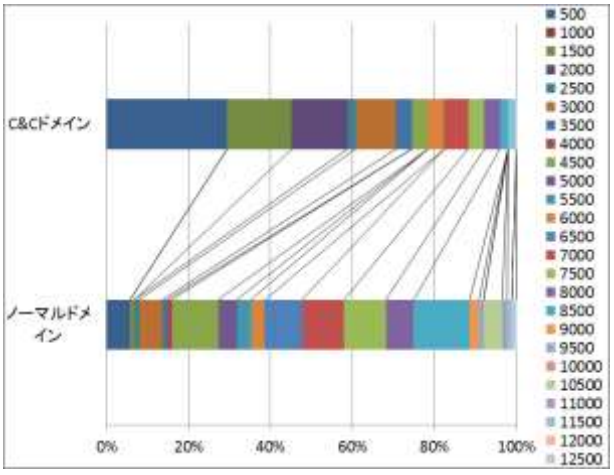

図 4　ドメインの有効日数

比較すると，ノーマルドメインより，C&C ドメインの方が値は小さい．

次に，f)～i)は各担当の連絡先が記載されており，以下の

情報を得ることができる．

j) ID
k) 名前
l) 組織名
m) 住所
n) 郵便番号
o) 電話番号
p) 国名
q) FAX 番号
r) メールアドレス

これらは，比較的容易に秘匿や改ざんすることができる．特に C&C サーバの多くは，身元を特定されないためにドメイン登録時に WHOIS の登録を代行してくれるサービス（WHOIS 登録代行サービス）を利用して登録情報を隠蔽していたり，でたらめな情報が登録されていることが多い．しかし，でたらめな情報が登録されている場合でも，r)のメールアドレスについては，実際に連絡を行ううえで必要なものであるため，偽装されていない可能性が高いことが考えられる．そこで，まずメールアドレスを対象に特徴点の抽出を行った．まず，ノーマルドメインと C&C ドメインの WHOIS に登録されてあるメールアドレスをデータマイニングにかけて，構造の特徴を抽出した．

今回，”UserLocal”のテキストマイニングツール[20]を用いて，各ドメイン別に特徴の抽出を行った．まず，ノーマルドメイン（図 5）と C&C ドメイン（図 6）におけるメールアドレスに用いられた単語の出現パターンの関係性を共起ネットワークで示す．

共起ネットワークとは，テキストの中で用いられた単語のパターンを構造化して関係性を示したネットワークであり，出現パターンが類似している単語同士を線で結んでいる．テキストマイニングの一種であり，これを用いて，メールアドレスの構造を明らかにするとともに，特徴の抽出を試みる．

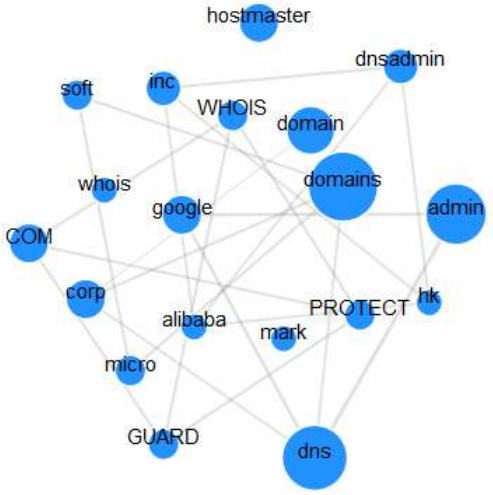

図 5　メールアドレス共起ネットワーク
（ノーマルドメイン）

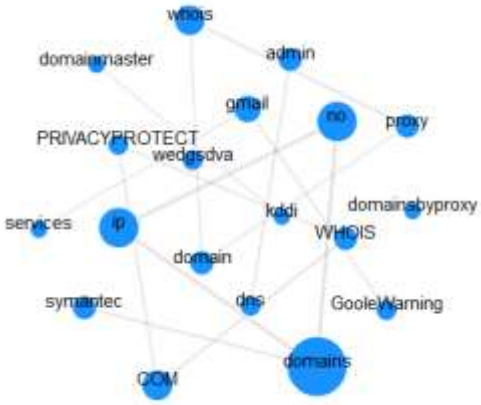

図 6　メールアドレス共起ネットワーク
（C&C ドメイン）

比較すると，ノーマルドメインの共起ネットワークは複数種類の単語が繋がっている大きな一つのかたまり，他に 4 種類の単語が相互に関係性を持っているパターンが 2 通り出現しており，C&C ドメインの共起ネットワークは大きなかたまりは見受けられないものの，3 種類の単語がお互いに関係性を持っているパターンが 3 通り出現している．

小さなかたまり一つ一つに注目すると，各かたまりの中に”no”や”PROTECT”，”proxy”といった WHOIS 登録代行サービスに用いられやすい単語が含まれていた．このことより，ノーマルドメインおよび C&C ドメインにおいてよく用いられる WHOIS 登録代行サービスに違いがあるのではないかと考えられる．

また，フリーメールアドレスが用いられている割合を示す（表 3）．

表 3　フリーメールアドレスの割合

ノーマルドメイン	C&C ドメイン
13.6%	17.5%

フリーメールアドレスの割合に大きな差は見られないものの，C&C ドメインの割合の方が高い結果となった．

以上の結果より，ドメインおよびメールアドレス，さらにドメインの有効年月日よりドメインの登録年月日を引いた有効日数の 3 種類の特徴を機械学習にかけて判別を行うこととした．

- **訓練モデルとアルゴリズム**

機械学習のアルゴリズムとして SVM（support vector machine）とニューラルネットワークの 2 種類を用いて訓練モデルの構築を行う．

SVM とは，教師あり学習の一種であり，パターン認識により 2 クラスの分類を行う[21]. 2 クラス以上の分類を行い際は複数の SVM を組み合わせることによって実現される．

ニューラルネットワークとは，教師あり学習の一種であり，脳機能にみられるいくつかの特性を数学モデル化することで，入力と出力の関係性を表現することができる[22].

まず，前処理として，ドメインとメールアドレスは，データマイニングを用いて構造化し，さらにドメインの有効期限年月日と登録年月日から有効日数を算出，この 3 種類をテストデータとして各アルゴリズムに学習させて訓練モデルを構築する（図 7）．

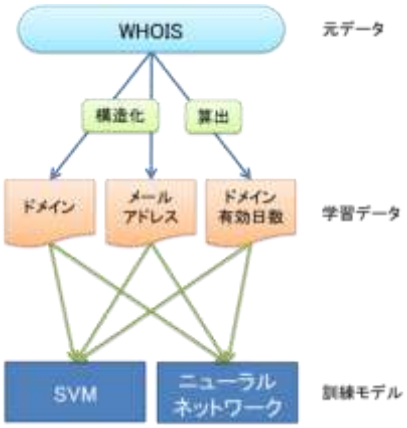

図 7　訓練モデル構築フェーズ

実際に構築した訓練モデルを用いて検知を行う際は，接続を試みるドメインの WHOIS より，訓練モデルを構築する際に利用した 3 種類の情報を抽出し，それをもとに訓練モデルを用いて判定を行う（図 8）．

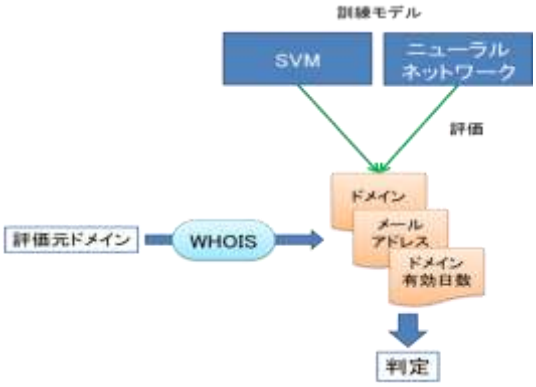

図 8　判別法

本手法をブラウザや Proxy などに組み込むことによって，C&C サーバの検知を行うことができる．

3.3 評価

今回，用いるデータ量が少ないため，実際に評価に用いるテストデータを準備しての評価では，テストデータの選び方によって精度に大きな誤差が生じる可能性がある．特に，標的型攻撃に用いられるドメインは，提供データが少なく，不足するため，データ量が少なくても比較的誤差を少なくできる手法である交差検証法を用いて評価を行う[23].

交差検証法とは，学習データとなる元のデータを一定のブロック単位に分割し，一つのブロックをテストデータ，その他のブロックを学習データとして評価を行う．分割したブロックごとに評価を行い，各評価結果の平均を推定精度として算定する手法である（図 9）．この方法を用いることにより，データ量が少なくても，推定される精度の誤差を少なくすることができ，以下の数式において求めることができる．この時，テストデータの総数はN^{ts}，正確に分類された総数はt^{ts}，n 回目の評価精度は$A^{ts}(d^n) = \frac{t^{ts}}{N^{ts}}$，求めたい推定精度は$A^{CV}(d)$とする．

$$A^{CV}(d) = \frac{1}{n}\sum_{i}^{n} A^{ts}(d^i)$$

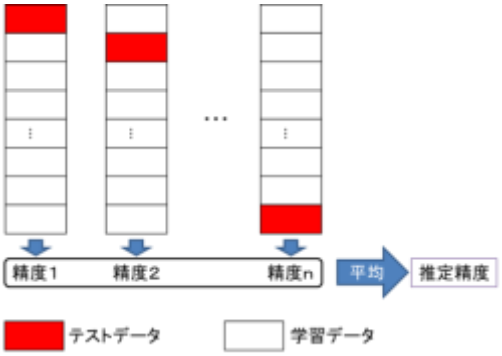

図 9　交差検証法

交差検証法において SVM およびニューラルネットワークで構築した訓練モデルを評価した（表 4）.

表 4　評価結果（交差検証法）

	SVM	ニューラルネットワーク
推定精度	88.3%	87.6%

評価結果より, SVM およびニューラルネットワークどちらにおいても比較的高い検知率を導き出せた．これは，多段追跡システムの 2013 年モデルの検知率を上回る結果である.

4. おわりに

提案手法により, C&C ドメインに用いられるメールアドレスの特徴点を明らかにし，ドメインの有効日数と組み合わせて機械学習することにより, C&C サーバの判別ができることを示した．また，抽出したメールアドレスに用いられている単語の関係性を共起ネットワークで示すことにより，C&C ドメインに利用されやすい WHOIS 登録代行サービスにおいても，特徴があることを示した.

本手法は攻撃者の準備したサーバ郡へアクセスすることなく，C&C ドメインの推定が可能である．これにより，攻撃者に解析していることを検知されにくくすることができた.

今回，ドメイン名および WHOIS に登録されてあるメールアドレス、ドメインの有効日数を入力値として，機械学習を用いて検証を行った．今後は入力値や入力値の前処理手法，機械学習アルゴリズムを見直すことで，より高い精度での判定を目指す.

参考文献

[1]　標的型攻撃等の脅威について
http://www.nisc.go.jp/conference/suishin/ciso/dai18/pdf/2.pdf
[2]　標的型攻撃　対策指南書（第 1 版）
http://www.lac.co.jp/anti-apt/guidebook/pdf/anti-apt_guidebook_ver1.pdf
[3]　「高度標的型攻撃」対策に向けたシステム設計ガイド
https://www.ipa.go.jp/files/000046236.pdf
[4]　比留間裕幸，橋本一紀，上原哲太郎，松本隆，佳山こうせつ，柿崎淑郎，八槇博史，佐々木良一：標的型攻撃に対する知的ネットワークフォレンジックシステム LIFT の開発（その１）－予兆検知と対策指示方法の提案－, DICOMO2015, pp.29-37(2015).
[5]　橋本一紀，比留間裕幸，上原哲太郎，松本隆，佳山こうせつ，柿崎淑郎，八槇博史，佐々木良一：標的型攻撃に対する知的ネットワークフォレンジックシステム LIFT の開発（その２）－プロトプログラムの開発と評価－, DICOMO2015, pp.38-43(2015).
[6]　佐々木良一，八槇博史：標的型攻撃に対する知的ネットワークフォレンジックシステム LIFT の開発（その３）－今後の研究構想－, DICOMO2015, pp.44-50(2015).
[7]　久山真宏,佐々木良一：標的型攻撃に用いられるドメインの WHOIS を基にした被害の早期発見手法の提案，CSEC71, pp.1-5(2015).
[8]　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数量化理論と機械学習での評価と提案, DICOMO2015, pp.991-917(2015).
[9]　D. I. Jang, M. Kim, H. C. Jung, B. N. Noh：Analysis of HTTP2P Botnet：Case Study Waledac, 2009 Ieee 9th Malaysia International Conference on Communications (Micc), pp. 409-412(2009).
[10]　Wei. Lu, M. Tavallaee, Ali. A. Ghorbani：Automatic Discovery of Botnet Communities on Large-Scale Communication Networks. ASIACCS '09 Proceedings of the 4th International Symposium on Information, Computer, and Communications Security(2009).
[11]　M. H. Tsai, K. C. Chang, C. C. Lin, C. H. Mao, H. M. Lee：C&C Tracer: Botnet Command and Control Behavior Tracing,　in IEEE International Conference on Systems, Man and Cybernetics (SMC), Anchorage, AK, pp.1859-1864(2011).
[12]　幾世知範，青木一史，八木毅，針生剛男：改ざんデータの出自確認に基づいた C&C サーバ特定手法の提案, 2014 年電子情報通信学会ソサイエティ大会　通信(2), pp.6-16(2014).
[13]　M. Felegyhazi, C. Kreibich, and V. Paxson：On the Potential of Proactive Domain Blacklisting, USENIX Conference on Large-scale Exploits and Emergent Threats, pp.6 (2010).
[14]　J. Ma, L. K. Saul, S. Savage and G. M. Voelker：Beyond Blacklists: Learning to Detect Malicious Web Sites from Suspicious URLs, ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp.1245–1254(2009).
[15]　L. Invernizzi, S. Benvenuti, P. M. Comparetti, M. Cova, C. Kruegel, and G. Vigna：EvilSeed: A Guided Approach to Finding Malicious Web Pages, IEEE Symposium on Security and Privacy, pp.428–442(2012).
[16]　Alexa Top 500 Global Sites
http://www.alexa.com/topsites
[17]　国内標的型サイバー攻撃分析レポート 2015 年版
http://www.trendmicro.co.jp/jp/about-us/press-releases/articles/20150409062703.html
[18]　VirusTotal
https://www.virustotal.com/
[19]　LastLine
https://www.lastline.com/
[20]　User Local
http://textmining.userlocal.jp/
[21]　John C. Platt: Fast Training of Support Vector Machines using Sequential Minimal Optimization. In B. Scholkopf,C. J. C. Burges, & A. J. Smola (Eds.).Advances in kernel methods—support vector learning.
[22]　Multilayer Perceptron
http://deeplearning.net/tutorial/mlp.html
[23]　モデルの精度を推定する
http://musashi.osdn.jp/tutorial/mining/xtclassify/accuracy.html