東京電機大学

博 士 論 文

標的型攻撃における外部通信の機械学習を用いた検知手法の研究

Study on detection method using machine learning of external communication in targeted attack

２０１７年１１月

佐々木良一研究室

久山　真宏

目次

図目次

表目次

1. はじめに

コンピュータが身近な存在となり，世界中のコンピュータを接続したインターネットが広く普及した．これにより，インターネットを通して世界中の人とコミュニケーションをとることができるようになった．インターネットはアメリカで軍事目的としてARPANET という名前で開発された．その後，現在のインターネットという形で民間への利用が開始した．日本では JUNET と言われる大学や研究機関を相互接続したネットワークが形成され，1990年代に商用のサービスが開始され一般的に利用できるようになった．当初は通信速度やコンピュータの処理能力も低く，高価であったため一部のユーザへの普及に留まっていた．その後，通信速度の向上や低価格化が進み利用が広まった．さらに，当初はファイルや電子メールの送受信，ネットニュースといった限られたサービスしか提供されていなかったが，速度の高速化，処理能力の向上，価格の低下によりユーザ数が増えるにつれて買い物や金銭のやり取り，テレビ電話としての利用，映画や音楽の再生等の様々なサービスをその場に行かなくても受けられるようになり，便利な社会へと変化した．しかし，利便性が高まる一方で，コンピュータに保存されてあるデータを人質にして金銭を要求したりする問題や情報漏洩といった問題が顕在化してきている [1].

誰もが安心できるサービスを提供するためには，信頼できるインターネット社会としての基盤構造が必要であり，セキュリティの確保が求められている [2].

1.1 セキュリティを脅かす脅威：サイバー攻撃の変容

セキュリティを脅かす脅威の代表格としてサイバー攻撃が存在する．サイバー攻撃は，図 1 に示すように 2000 年ごろまではマス型といわれる攻撃が多かった．これは不特定多数を対象とし、愉快犯などが行う単純な攻撃であった．例えば、マルウェアがコンピュータに感染しても単にコンピュータの画面を見えにくくするものであった．そこから Web サイトの改ざんやサーバの乗っ取りなど

年々手口が巧妙化・多様化していった．インターネットが普及してコンピュータ上でやりとりをする情報の価値が高まり，それにともない目的も愉快犯的なものから金銭や情報を狙ったものへと変容し，社会的な問題になっている．

攻撃の主体も個人から組織へと移り変わっており，近年では国家間で国家機密を狙ったサイバー攻撃が起こり，サイバー空間上では戦争（サイバー戦争 [3]）状態であるとも言われる．

図 1　サイバー攻撃の動向

（「重要インフラ分野における IT 依存度に関する調査」 [4]より抜粋）

サイバー攻撃の中でも，標的型攻撃は特定の組織や会社を対象とする攻撃で、機密情報を盗み出したり，システムの破壊活動を伴うことが多い．

1.2 対策の概要

サイバー攻撃に対する防御策は様々な対策が存在する．大別するとシステムや機器による対策とそれらを用いた運用による対策がある．

1.2.1 システムや機器による防御

サイバー攻撃を防ぐためのシステムや機器として次のようなものがある．なお，各システムや機器によっては複数の機能を持つものもあり，すべての製品を一概に次の分類に分けることはできない．

各システムや機器がネットワーク上のどの部分に該当するかを図 2 に例示する。

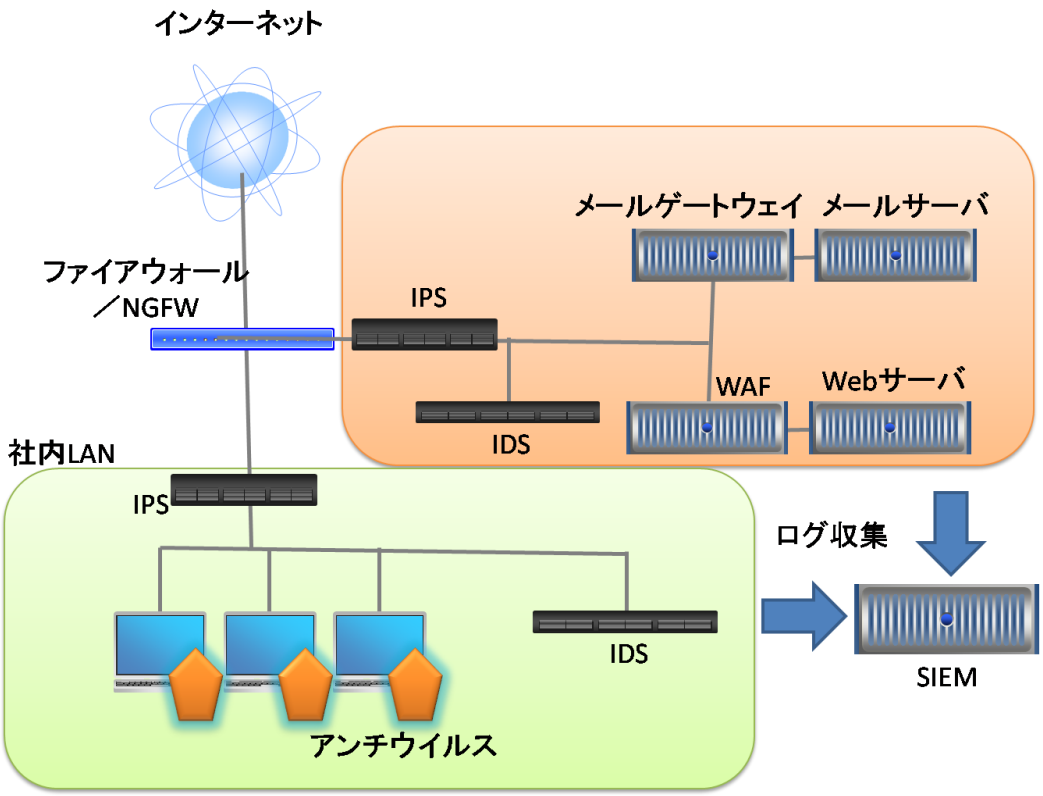

図 2　ネットワーク内におけるシステムや機器の設置箇所

（１）ファイアウォール

ネットワークとネットワークの接続部に設置され，送信元先の IP アドレスやポート番号，プロトコルなどを組み合わせてアクセス制御を行う．アクセス制御を行うためのフィルタリングルールのことを ACL（Access Control List）と言う．防御方法としてはパケットフィルタリング型やアプリケーションゲートウェイ型などのタイプがある．

パケットフィルタリング型は送信元先の IP アドレスとポート番号を用いて通信のフィルタリングを行う．これにより，使用していないポートへの通信を拒否したりすることで不要な通信を遮断することができる．いたってシンプルな設計であるため高速に処理できるが，IP アドレスやポート番号の偽装には弱いといった特徴がある．

アプリケーションゲートウェイ型は通信を行うアプリケーション毎に中継機を経由させることによりフィルタリングを行う．パケットフィルタリング型に比べて実際の通信内容までフィルタリング対象に含めることができるため，より詳細にフィルタリングルールを設定することができる．しかし，使用するアプリケーションプロトコル毎に中継機が必要となる．

ファイアウォールは外部と内部の境界に設置されることが多く，外部からの攻撃を防ぐ代表的な防御手段であるが，システムの処理能力を超えるほどの大量な通信を送りつけてシステムをダウンさせるようなサイバー攻撃だと，パケットが ACL の定義上正常だと判断されると防ぐことができない．

（２）WAF

WAF は Web Application Firewall の略である．

Web アプリケーションに対する攻撃を防ぐのに特化したセキュリティ製品．Web サーバの前に設置され HTTP（S）通信内にクロス

サイトスプリプティングといった Web アプリケーションの脆弱性をねらった攻撃コードが含まれていないか判断を行う製品.

（３）アンチウイルス

事前に不審なコードとして定義されたパターン（シグネチャ）に一致するか確認して，一致したものがあれば警告・駆除する製品である．また，プログラムの動作や振る舞いから不審な点がないかどうかをチェックする機能を持つものもある．不審なプログラムやマルウェアを検知・駆除する製品として広く普及している.

（４）IDS / IPS

IDS は Intrusion Detection System の略であり，IPS は Intrusion Prevention System の略である.

ネットワーク上を流れる通信から，事前に不審な通信として定義されたパターン（シグネチャ）に一致するかを確認して，一致したものがあれば警告を出す製品．IPS は警告を出すとともに該当の通信を遮断する．通信を監視したいネットワーク内に設置する.

IDS は通信を監視したいネットワーク内に設置し，ミラーポートと呼ばれる技術を用いてネットワーク内を流れるすべての通信を収集･監視する.IPS は通信を監視したい箇所のライン上に設置し，そのラインを通過する通信を監視し，不審な通信を遮断する．製品や持ち合わせる機能自体は同じものであるが，ライン上に設置するかどうかで IDS／IPS と役割が変わるものが多い.

（５）SIEM

SIEM は Security Information and Event Management の略であり，セキュリティ情報管理 SIM（Security Information Management）とセキュリティイベント管理 SEM（Security Event Management）を組み合わせた造語である.

セキュリティ製品のログに限らず，端末のログやネットワーク機器のログなどの様々なデバイスのログ情報を収集，相関分析を通してサイバー攻撃を検知する製品．

サイバー攻撃の検知するとともに，ログの収集・集約により事故発生後の調査（フォレンジック）やポリシー管理やコンプライアンス違反の発見にも役立つ．

SIEM は人手では対処しきれないほど多種多様な大量のログを収集・分析することができるとして期待されている．

（６）NGFW

NGFW は Next Generation Firewall の略である．

次世代ファイアウォールとも言われる．従来のファイアウォールにアンチウイルス機能や IDS / IPS 機能など各種様々な機能を持たせたものを言うことが多いが，明確な定義はない．当初はファイアウォールに付随する機能によって統合脅威管理 UTM（Unified Threat Management）などのように呼び名が変わっていたが，現在はどの呼び名であっても様々な機能を有しているため，明確な定義はない状態である．

（７）メールゲートウェイ

メールに関するプロトコルを理解するアプリケーションゲートウェイ型ファイアウォールということができる．メールサーバの前に設置してメールのやり取りを中継する．この時，フィルタリング機能や，迷惑メールやスパムメール対策機能，ウイルス対策機能など様々な機能を用いて不審なメールを削除したりアーカイブすることができる．

ウイルス対策機能には先述したアンチウイルス機能のほかにも，サンドボックスという保護された環境下で実際にプログラムを動作させて，どういった挙動をするのかを監視し，不審な点がないか

どうかを挙動から判断する機能を持った製品もある.

1.2.2 運用での防御

サイバー攻撃対策用のシステムや機器を導入するだけではサイバー攻撃を防ぐことは難しい. 導入後も適切に運用するための仕組みや運用が必要不可欠である. たとえどんなに高価な機器を導入していても使いこなせなければ宝の持ち腐れとなる.

ここでは, 各システムや機器を駆使しながら運用面からサイバー攻撃に対処するために必要となる組織について解説する. 各組織の関係を図 3 に例示する.

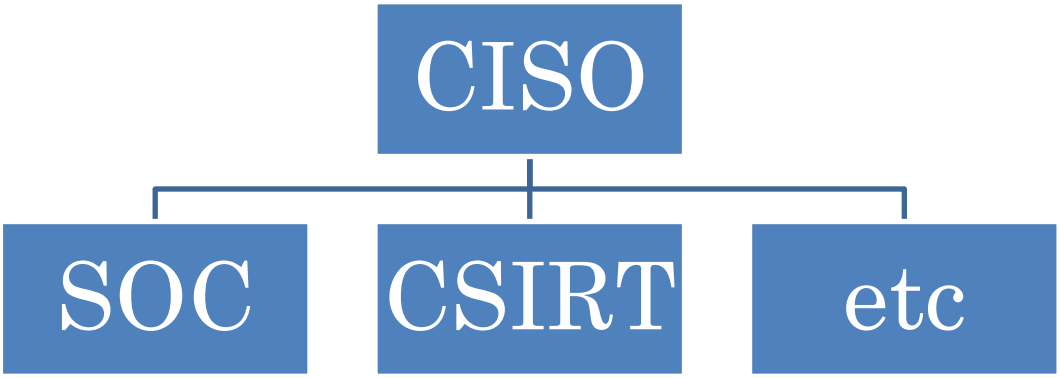

図 3 運用組織のツリー

（１）CISO

CISO は Chief Information Security Officer の略であり，最高情報責任者のことを指す．

会社や組織の経営的な立場から情報セキュリティの推進を行い，情報セキュリティを統括する．情報セキュリティ対策についての現状を把握し，必要となる対策やその維持の推進を行うべく経営会議にて提言を行う．特に後述する CSIRT や SOC を組織内で構築する際には，CISO が推進する立場となり有識者や適任者を選定して構築を進める．

（２）CSIRT

CSIRT は Computer Security Incident Response Team の略である．

コンピュータやネットワークにかかわるインシデントに対処するための組織の総称名である．平時はインシデント関係の情報やソフトウェアの脆弱性情報，攻撃の予兆となる情報などを収集・分析し，対応方針や手順の策定などを行う．ひとたびマルウェアの感染やサイバー攻撃によるコンピュータやネットワークの問題が発生した際に出動する．

（３）SOC

SOC は Security Operation Center の略である．

コンピュータやネットワークを監視してサイバー攻撃の検知や分析，対応策のアドバイスを行う組織であり，CSIRT がインシデント発生後に重点が置かれているのに対して，SOC はインシデントの発見に重点が置かれている．

従来は IDS / IPS が検知したイベントについて，誤検知・攻撃の成功有無を調査して，攻撃が成功していた場合はその影響度を分析し，対応策のアドバイスを行うのが主流であった．IDS / IPS の設置個所としては監視対象のネットワーク内であったり，そのネットワ

ークの出入り口に設置する．しかし，サイバー攻撃がより高度化するとともに出入り口の監視での検知が困難になり，最近は SIEM を用いて各種セキュリティ製品からのイベントやネットワーク機器などのログからサイバー攻撃を検知するところが多くなっている．

1.3 既存対策とマルウェアを用いたサイバー攻撃

既存の対策を適切に運用することで，サイバー攻撃の多くは防ぐことができる．例えば，マルウェアを添付したメールが送信された場合，まずは（７）メールゲートウェイがメールに添付されているファイルがマルウェアでないかをチェックしている．それをすり抜けても，次はマルウェアが添付されたメールを受信したユーザの端末にある（３）アンチウイルスがマルウェアかどうかをチェックする．さらに，それすらすり抜けてマルウェアを実行（感染）してしまっても，（４）IDS / IPS により C&C サーバとの通信がないかをチェックしたり，（５）SIEM により各種ログの相関分析して不審な挙動がないかどうかをチェックできる．このように，一つの攻撃に関して各種様々な機器により多層的・重層的にチェックしているため，どこかの機器でサイバー攻撃を見逃しても，他の機器によりサイバー攻撃を検知できるようになっている．しかし，標的型攻撃のような高度な攻撃は，標的となるネットワークのことを事前に調査し，既存の対策では検知・防御できない攻撃手段を用いてくる．そのため，既存の対策だけで防ぎきることは難しい．

実際に日本年金機構などで用いられた標的型メール攻撃と言われる種類のマルウェアを送られてから情報を抜き取るまでの流れを次に示す（図 4 参照）．

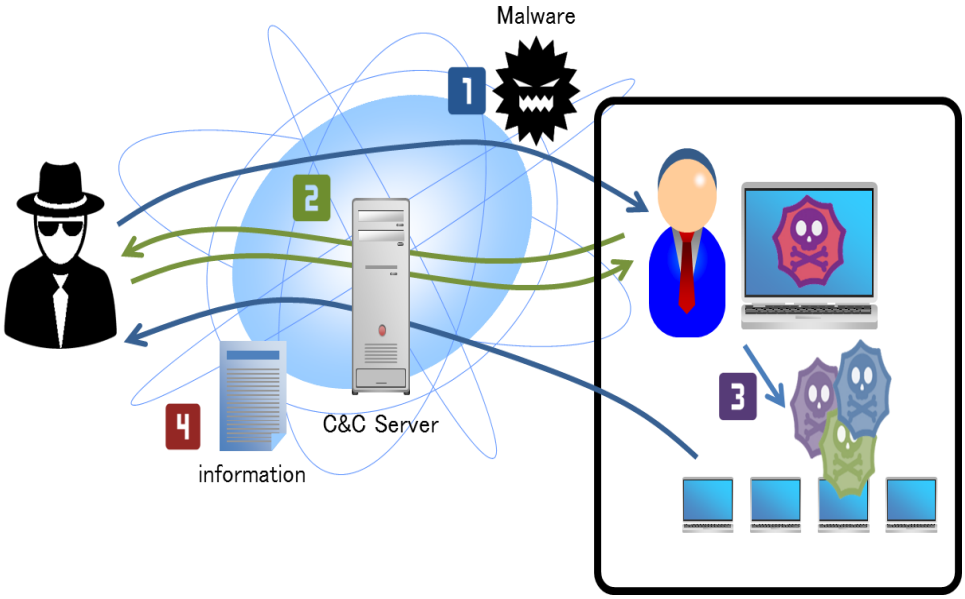

図 4　標的型メール攻撃の流れ

ステップ 1: メールなどを用いて標的となる組織の LAN 内にある端末にマルウェアを感染させる．

ステップ 2：マルウェアに感染した端末は，C&C サーバと通信する．そして，目的を達成するためにより適切なマルウェアなどが端末にダウンロードされる．

ステップ 3：マルウェアは，LAN 内の他の PC やサーバに侵入範囲を拡大しようとする．

ステップ 4：重要な情報，機密情報や組織の個人情報といった目的とする情報を見つけると，攻撃者に送信される．

ステップ1で攻撃者から送られてくるマルウェアを検知して防衛できるのが良いが，標的型攻撃では標的となる組織毎に検知されにくいようにカスタマイズされたマルウェアが用いられるため，検知するのは難しい．その後の一連の流れに着目すると，標的となった端末がマルウェアに感染してから攻撃者とやり取りを行うにあたり，C&C サーバは攻撃者が目的を達成するための司令塔という重要な役割を担っている．そのため，C&C サーバとの通信を検知することにより，被害を早期に発見することが出来る．

1.4 研究目的

本研究では標的型攻撃におけるマルウェアの通信に着目し，C&C サーバとの接続を遮断するための手法を提案する．我々の研究グループでは，これまでに，メールアドレスの構造を用いる検知手法[5]や検索エンジンから得た情報を組み合わせて検知する手法 [6]，これらの方式を統合した手法 [7] [8]を提案するとともに，実験およびその評価結果について報告する．

1.5 関連研究

C&C サーバの特定を目的とした研究は，次の 2 種類に大別される．

1.5.1 C&C サーバとの通信に着目した研究

C&C サーバとマルウェア間で行われる通信に着目し，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]，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 [11]などがある．

これらの手法は，実際の通信内容から検証するため，十分な検証により高い検出精度で特定することができる．しかし，ゼロデイ攻

撃などの未検証な検体への対応に不十分な問題がある.

なお，既存の対策である（４）IDS / IPS での検知も C&C サーバとの通信の内容から検知している.

1.5.2 C&C サーバのドメインに着目した研究

C&C サーバのドメインに着目し，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 [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]などがある.

これらの手法は，活動中の C&C サーバに対して高い検出精度で特定することができる．しかし，C&C サーバや C&C サーバのドメインを管理する DNS サーバといった攻撃者の関与するサーバ類へリクエストが飛んでしまい，攻撃者に解析していることを検知され，攻撃者に対策されてしまう問題がある.

1.5.3 先行研究

当研究室では 2009 年より攻撃を受けた端末から攻撃元を追跡していき，最終的には攻撃者を特定することを目的とした多段追跡システムの研究を行っている [16]．その中で，数量化理論 2 類 [17] を用いてボットネットの C&C サーバを判別する手法を 2009 年に提案した．2009 年当時は 96.5%の精度でボットネットの C&C サーバを検出できていたが，継続的に調査を行ったところ，検出精度は年々下がり，2011 年には 76.5%まで検出精度が下がった [18]．これはボットネットの C&C サーバの特徴が時間経過とともに変化していることが原因である [19] [20]．そのため，一定期間ごとに最新

のデータを用いて判別モデルの見直しを行ってきた．継続調査による検出精度の変化を表 1 に，各モデルにおいて検出に用いた特徴を表 2 に示す．

表 1　継続調査による検出精度の変化

モデル	検知率（年）				
	2009	**2010**	**2011**	**2013**	**2014**
2009	96.5%	85.0%	76.5%	-	-
2011	-	-	95.2%	42.5%	-
2013	-	-	-	80.3%	80.8%
2014	-	-	-	-	96.7%

表 2　各モデルで使用した特徴

用いた特徴		モデル			
		2009	**2011**	**2013**	**2014**
DNS	逆引き	○	○	○	
	TTL				○
	minimum	○	○		○
	A レコード		○	○	
	MX レコード				
	NS レコード				○
	CNAME レコード			○	
	TXT レコード				○
WHOIS	登録期間	○	○	○	○
個数		3	4	4	5

これまでの取り組みでは，ボットネットの C&C サーバを検知する研究を行っている．ボットネットの C&C サーバは 1 台の C&C サーバで複数の感染端末を同時に操作するのに対して，標的型攻撃に用いられる C&C サーバは 1 台の C&C サーバで感染端末を個別に操作する特徴がある．今回は標的型攻撃に用いられた C&C サーバの検知についてもドメインの特徴から検知可能かどうかを研究する．

2. 標的型攻撃

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 [21]. 標的型攻撃はサイバー空間のみに限らず，現実世界で実際に標的となる組織の人員に会いソーシャルエンジニアリングを利用するなど，多様な手段を用いて情報を盗み出す．一般的には標的となる企業や組織に対して有効となるように作り込まれたマルウェアを感染させる攻撃が多い [22].

2.1 標的型攻撃の事例

実際に国内の大手重工メーカや衆議院，日本年金機構といった様々な組織が標的型攻撃の被害に遭い，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．

表 3　標的型攻撃の歴史

年月	内容
2009 年 11 月	世界中のエネルギー関連企業や製薬会社へのサイバー攻撃
2010 年 1 月	Google といった米国企業へのサイバー攻撃
2010 年 6 月	イランの核燃料施設へのサイバー攻撃
2011 年 4 月	ソニーの米国子会社へのサイバー攻撃
2011 年 9 月	三菱重工へのサイバー攻撃
2011 年 10 月	衆議院へのサイバー攻撃
2012 年 5 月	原子力安全基盤機構での情報漏洩
2012 年 7 月	財務省での情報漏洩
2012 年 11 月	宇宙関連事業所のマルウェア感染
2013 年 1 月	農林水産省での情報漏洩
2013 年 2 月	外務省での情報漏洩
2013 年 5 月	Yahoo!JAPAN へのサイバー攻撃
2014 年 8 月	日本の ISP や学術機関へのサイバー攻撃
2015 年 6 月	日本年金機構へのサイバー攻撃
2015 年 11 月	厚生労働省へのサイバー攻撃
2016 年 1 月	複数省庁へのサイバー攻撃

2.2 標的型攻撃の流れ

標的型攻撃では攻撃に至るまで複数のプロセスを得て攻撃される．攻撃手順のプロセス化された代表的なものとして，Lockheed Martin 社の Mike Cloppert 氏らにより提唱された「サイバーキルチェーン」という考え方がある．

図 5　サイバーキルチェーン

（PwC「サイバー演習の種類と概要」 [23]より引用）

（１）　偵察

事前に標的となる組織を偵察して情報を得る.

（２）　武装

偵察によって得た情報を用いてマルウェアを作成.

（３）　配送

作成したマルウェアを標的となる組織に送付.

（４）　標的型攻撃

標的組織で送付されたマルウェアを実行.

（５）　インストール

マルウェアの感染が拡大.

（６）　コマンド＆コントロール

攻撃者がマルウェアへ指示を出す.

（７）　目的の実行

情報を盗み取る.

攻撃者は事前に標的となる企業や組織のことを（１）で調査し，調査によって得た情報を用いて，標的となる企業や組織ごとに見つかりにくく，より目的を達成できるように（２）でカスタマイズされたマルウェアが作成される．そのため，（３）でマルウェアが配送されたときに既存の防御策であるアンチウイルス製品などでは検知できないマルウェアが用いられることが多い．つまり，マルウェアを検知して未然に感染を防ぐ入口対策だけでは不十分であり，多層的・重層的な対策が求められる．

3. 外部リポジトリと機械学習を用いた C&C ドメイン検知手法の提案 [7]

本提案手法は，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，外部リポジトリとして DNS と WHOIS，検索エンジンを用いて C&C サーバのドメインの検知を試みる. DNS, WHOIS, 検索エンジンから各特徴点を抽出し，機械学習を用いて C&C サーバのドメインであるかどうかを判別する.

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．次に，各ドメインの WHOIS，DNS，検索情報から特徴点を抽出する.

抽出した特徴点を機械学習で学習させ，訓練モデルを構築する．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.

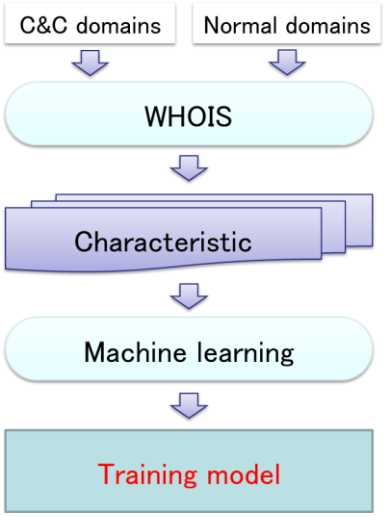

図 6　機械学習を用いた訓練モデルの構築

3.1 評価ドメイン

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.” [24]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を解析して抽出したドメインを利用した [25]. マルウェアの収集にあたっては, VirusTotal [26]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.

VirusTotal を使用して実際に収集したマルウェアの内訳を表 4 に示す.

表 4　収集したマルウェアの内訳

Malware type	Samples
Emdivi	50
PlugX	63
PoisonIvy	50

収集したマルウェアを LastLine [27]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．

なお、VirusTotal や LastLine について詳しくは Appendix B Lastline&VirusTotal を参照願いたい。

3.2 WHOIS の特徴

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 日刻みで図 7 に示す．

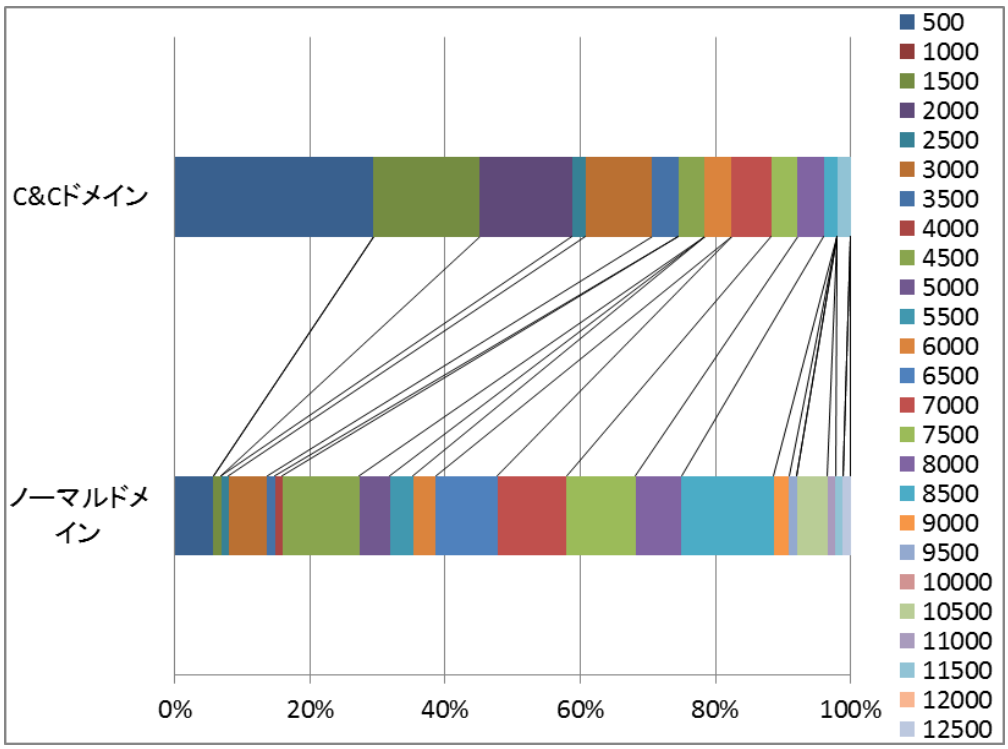

図 7　ドメインの有効日数

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” [28]のテキストマイニングツールを用いて，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 8 参照）と C&C ドメイン（図 9 参照）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，一文の中から出現する単語のパターンをネットワーク構造化して関係性を図として表したものであり，テキストマイニングの一種である．これにより，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.

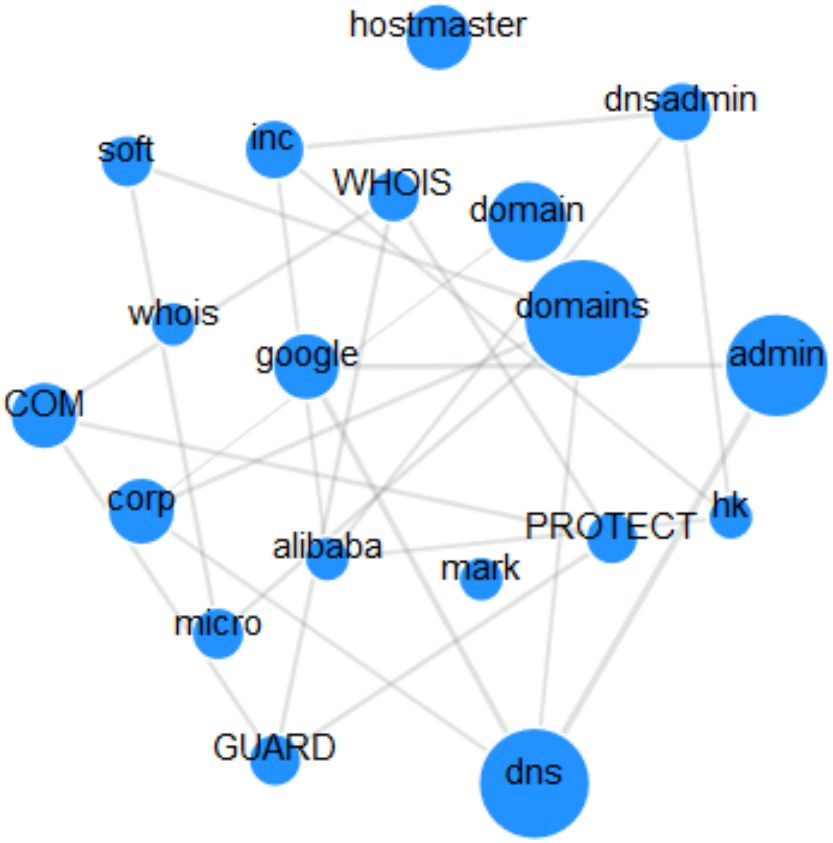

図 8　メールアドレス共起ネットワーク（ノーマルドメイン）

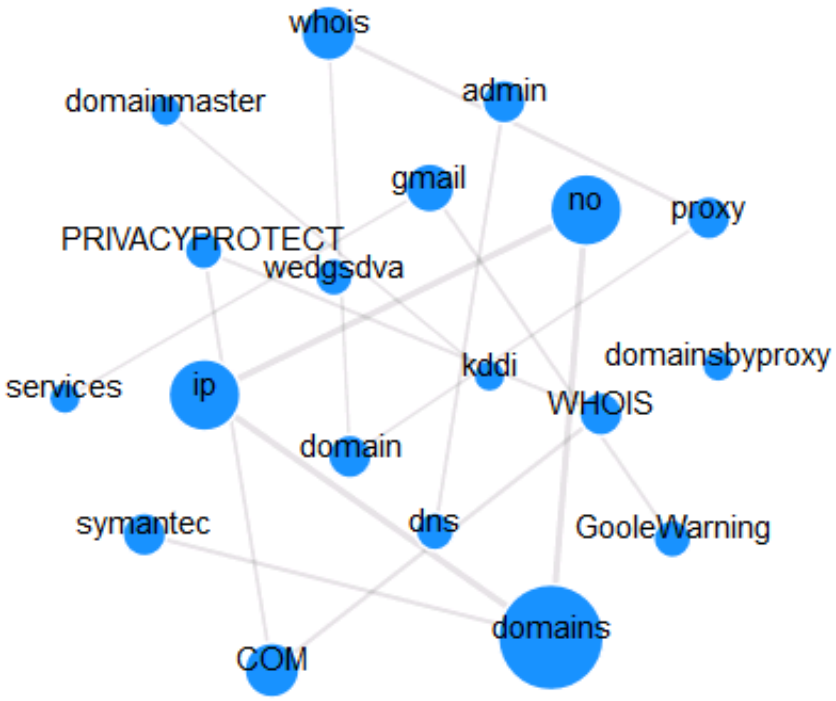

図 9　メールアドレス共起ネットワーク（C&C ドメイン）

比較すると，図 8 の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，図 9 の 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表 5 に示す．

表 5　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若干高い結果となった．

以上の結果より，WHOIS より得られたメールアドレス，さらにドメインの有効年月日からドメインの登録年月日を引いて算出されたドメイン有効日数を特徴点として選択した．

3.3 DNS の特徴

DNS は，ドメイン名を IP アドレスに変換するシステムである．DNS の技術仕様と運用ルールは RFC1034 [29]と RFC1035 [30]に定められている．

DNS からは次のレコードを取得することができる．

a) Address(A) record
b) Start of authority (SOA) record
c) Host information (HINFO) record
d) MX record
e) NS record
f) Canonical name (CNAME) record
g) Well-known services (WKS) record
h) Text (TXT) record

この中でも，NS レコードと MX レコードの登録レコード数には顕著な違いがある．ノーマルドメインと C&C ドメインの NS レコード数を図 10 に，MX レコード数を図 11 に示す．

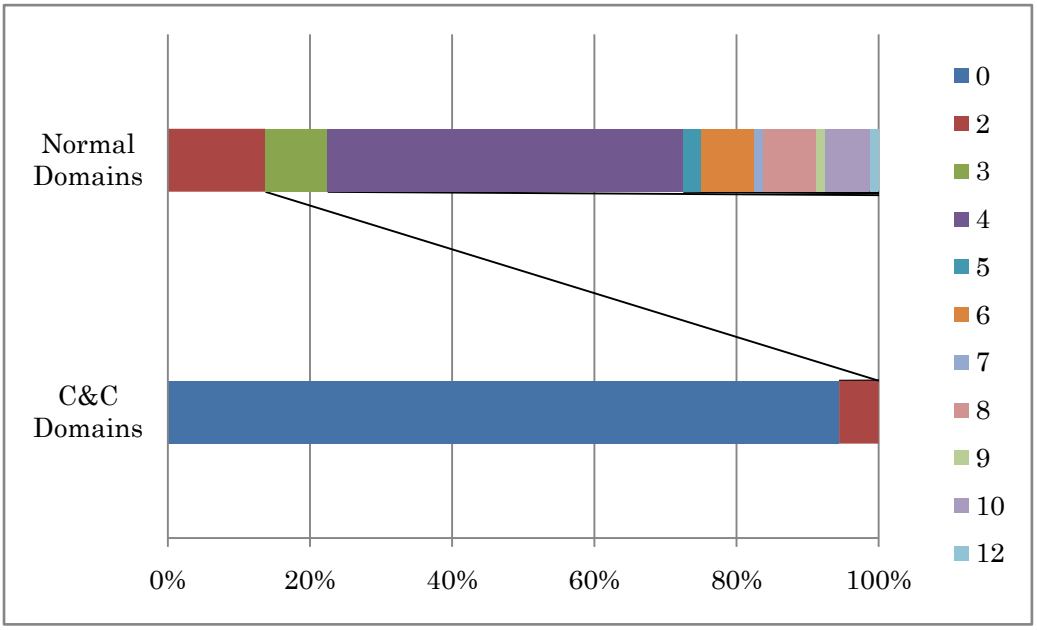

図 10　NS レコード数の比較

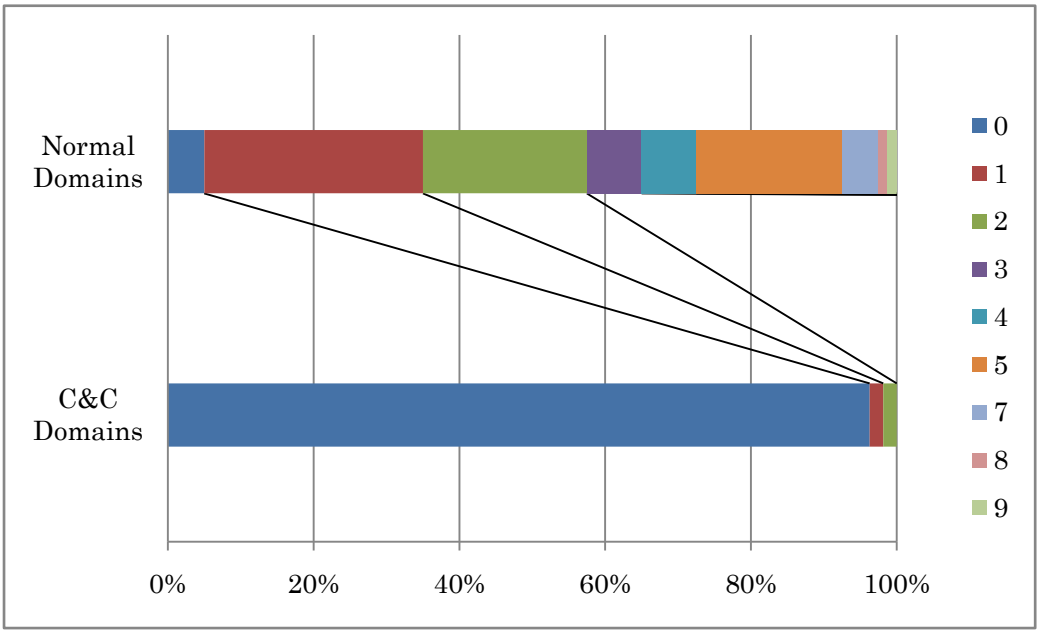

図 11　MX レコード数の比較

NS レコードおよび MX レコードともに C&C ドメインにはほとんど登録されてないが，ノーマルドメインでは複数登録されているのがわかる．

したがって，DNS 情報からは NS レコードの数と MX レコードの数を特徴点として選択した．

3.4 検索エンジンの特徴

関連研究 [15]では，検索エンジンを使用して，ドライブごとの既知の C&C サーバの特性を検索して，未知の C&C サーバを検出している．

ドライブバイダウンロード攻撃 [31]では，PC は Web サイトをブラウズすることでマルウェアに感染する．マルウェアに感染している悪意のある Web サイトは，検索エンジン最適化（SEO）を行い，より多くのマルウェア感染を引き起こしている．SEO の目的は顧客を引き付けることであると考えられる．

一方，標的とする攻撃のマルウェア感染には Web サイトは使用されていないことが多い．攻撃者は標的とされた攻撃を検知されないようするために C&C サーバであることがばれないようにしている．また，短命な C&C サーバは，Web 検索エンジンのクローラに発見される前にサーバが停止するため，Web 検索エンジンで C&C サーバが見つからないことが考えられる．

そこで，Google 検索エンジンを使用して評価ドメインが検索エンジンによって見つけることができるか調査した．その結果を図 12 に示す．

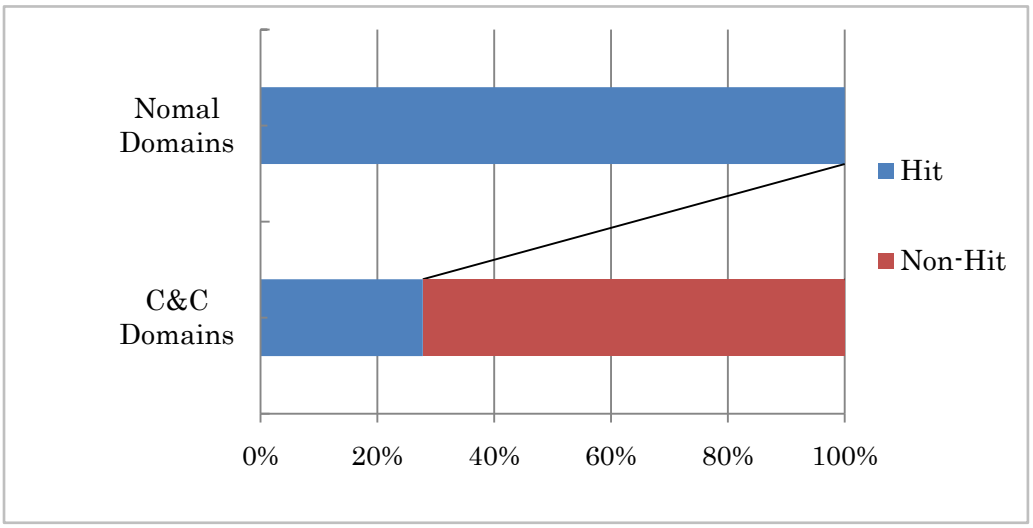

図 12　検索エンジンの結果

多くの C&C ドメインは検索サイトにヒットしなかった．いくつかのヒットした C&C ドメインは，ハイジャックされたサーバの可能性が高いものであった．そのため，検索エンジンからはヒットの有無を特徴点として選択した．

3.5 機械学習アルゴリズム

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を用いた訓練モデルを構築する．

SVM は，パターン認識により 2 つのクラスに分類する機械学習である [32].

ニューラルネットワークは，教師付き学習方法の一種である．

人間の脳機能に見られる特徴のいくつかの数学的モデリングによって，入力と出力との間の関係を表現することが可能である．標準的な方法は，2 つの層を使用する階層型ニューラルネットワークである [33].

WHOIS からの電子メールアドレスと有効な用語，NS レコードの数，DNS からの MX レコードの数，および検索サイトからのヒット数を含むニューラルネットワークを使用してトレーニングモデルを構築する．

機械学習で用いる特徴点を表 6 に示す．

表 6　使用する特徴点

Input		# Type
Label		Normal or C&C
Domain		String
WHOIS	Admin mail address	String
	Registered mail address	String
	Technical mail address	String
	Valid term	Number
DNS	NS record	Number
	MS record	Number
Search site		Hit or Non-Hit

3.6 評価

評価にあたって，Alexa から 80 のノーマルドメインおよび収集したマルウェアから抽出できた 54 の C&C ドメインを評価ドメインとして使用した．

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 [34]を用いて評価を行う．

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（図 13 参照）[35]．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．

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

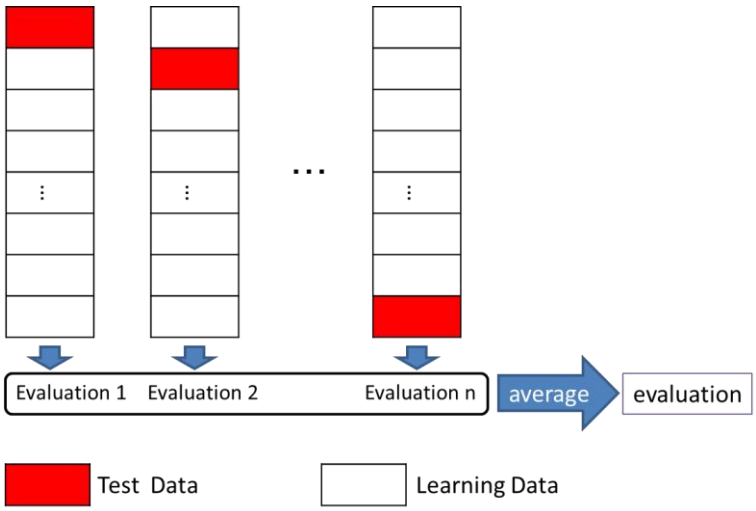
Evaluation 1
Evaluation 2
Evaluation n
average
evaluation
Test Data
Learning Data

図 13　交差検証法

相互検証法を用いた SVM およびニューラルネットワークによる WHOIS，DNS および検索サイトの各組み合わせの評価結果を表 7 に示す.

表 7　評価結果一覧

Combination	SVM	Neural network
WHOIS only	88.8%	88.8%
DNS only	96.3%	95.5%
Search site only	88.8%	88.8%
WHOIS + DNS	97.8%	98.5%
WHOIS + Search site	91.8%	92.5%
DNS + Search site	99.3%	99.3%
WHOIS + DNS + Search site	99.3%	99.3%

結果，SVM とニューラルネットワークは 99.3%の優れた検出率を達成した．

WHOIS のみまたは検索サイトは，88.8%の検出率であった．しかしながら，DNS はより高い検出率を達成することができた．これは，DNS が重要な要素であることを示している．

WHOIS と DNS に検索サイトを追加した結果，検出率も向上した．したがって，検索サイトを使用することは有効であることがわかる．

攻撃者によって構築された C&C サーバは，検索サイトにヒットしなかった．したがって，検索サイトは C&C サーバの検出に効果的であると言える．ただし，攻撃者がサーバをハイジャックした場合，ヒット有無に対しては無効である．そこで，サーバをハイジャックする攻撃者は，WHOIS と DNS の組み合わせによって検出される．

3.7 まとめ

本稿では，C&C ドメインに使用する電子メールアドレスの特徴点を収集し，WHOIS，DNS，Web 検索エンジンの結果などのよく知られた情報を用いた機械学習を用いて C&C サーバを決定する方法を提案した．

共起ネットワークにおける抽出された電子メールアドレスの単語の関係を示すことにより，C&C ドメインに使用される WHOIS 登録エージェンシーサービスの特徴を明らかにした．さらに，C&C サーバの検出のための検索サイトの使用が効果的であることが判明した．

最後に，ドメイン名と電子メールアドレスを評価した．WHOIS からメールアドレス，DNS からは NS レコードの数と MX レコードの数，検索エンジンからはヒットした数を機械学習に入力した．その結果，99.3%の高い検出率で C&C サーバを識別することができた．今後の作業では，機械学習アルゴリズム，入力値，および前処

理を改訂することで精度を向上を図る.

4. 攻撃者に察知されにくい情報を用いた C&C サーバの検知手法の提案 [8]

第 3 章で提案した手法は DNS の情報を用いるため，解析していることを攻撃者に気付かれる危険性がある．そこで本章では，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することなく収集できる情報を用いて C&C サーバのドメインを高精度に検出する手法を提案する．第 2 章でも説明した節もあるが，齟齬のないよう本章における定義として改めて説明する．

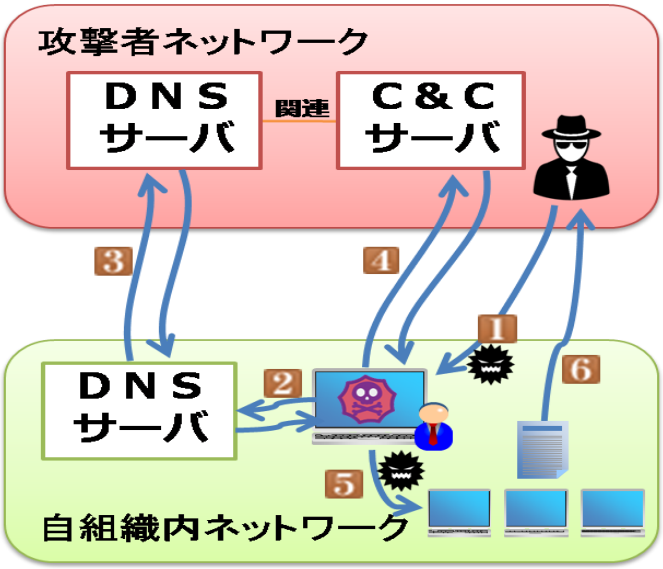

図 14　標的型攻撃における DNS

ステップ 1：標的型攻撃を行うために，LAN 内の端末にマルウェアを感染させる．

ステップ 2：マルウェアに感染した端末は，C&C サーバと接続を行うために，自組織内の DNS サーバに C&C サーバのドメインの名前解決を要求する．

ステップ 3：自組織内の DNS サーバ内に要求されたドメインに対応した IP アドレスが不明の場合，さらに上位の DNS サーバに名前解決を要求して，回答のあった IP アドレスをマルウェアに感染した端末に返す．

ステップ 4：マルウェアに感染した端末は，回答のあった IP アドレスをもとに C&C サーバと通信する．そして，目的を達成するためにより適切なマルウェアなどが端末にダウンロードされる．

ステップ 5：マルウェアは，LAN 内の他の PC やサーバに侵入範囲を拡大しようとする．

ステップ 6：重要な情報，機密情報や組織の個人情報といった目的

とする情報を見つけると，攻撃者に送信される．

本研究ではステップ 3 の通信に着目し，攻撃者が準備した C&C サーバや DNS サーバなどにアクセスすることなく収集できる情報を用いて，自組織内の DNS サーバで C&C サーバのドメインを検知し，遮断する手法を提案する．

本手法を自組織内にある DNS サーバに実装することで，自組織内の DNS サーバへ名前解決のクエリからドメインを抽出し，提案手法で悪性かどうかを判別する．判別の結果，悪性と判断されたものは名前解決要求を拒否することで，その後の通信を遮断することができる．また，これにより自組織外の DNS サーバへ通信を発生する前に通信を遮断することができ，攻撃者が関与する DNS サーバに飛ぶ前に通信を遮断することができる．

C&C サーバの判別には，ドメインの WHOIS 情報と Google の検索エンジンを用いる．WHOIS とは，ドメインの登録に関する情報を管理・提供するサービスであり，RFC812 [36]および RFC3912 [37]に技術仕様や運用規則が定められている．トップレベルドメイン（TLD）のレジストラごとに特定の組織のみが運用を許可されており，WHOIS に登録されてある情報は一般公開されていることから，WHOIS に登録されてある情報を利用しても攻撃者は自身のドメインが WHOIS で参照されているのかどうか気づきにくい．また，WHOIS はドメインに関する連絡先や管理母体を示す情報であり，内容を変更するには運用元のレジストリに対して申請を行う必要がある．DNS よりも特徴が変化しにくく，情報を参照しても攻撃者に察知されにくいといったメリットがある．

Google [38]も運用元が Google.inc.と特定された企業であり，提供されている情報も一般的に公開されているため，同義の理由から Google の検索エンジンから得られる情報を利用しても攻撃者は気づきにくく，さらに攻撃者の意思で特徴を変更することが困難なことから特徴が変化しにくいといったメリットがある．以上より，WHOIS と Google の検索エンジンから得られる情報を利用すること

とした．

WHOIS と Google の検索エンジンから得られた情報から特徴点を抽出し，機械学習を用いて C&C サーバの判定を行う．今回，悪性かどうかの 2 クラスのパターン識別として教師あり機械学習であるサポートベクタマシン（SVM）とニューラルネットワークを用いる．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．

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．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．

4.1 評価ドメインの準備

第 2 章の提案手法と同様に，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[25]を収集・解析し，抽出でき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 を用いて，キーワードに Emdivi，PlugX，PoisonIvy の種別名で検索を行い，2015 年 1 月～2016 年 8 月の間に投稿された計 464 件のマルウェアを収集した．

表 8　収集したマルウェアの検体数

マルウェア種別	検体数
Emdivi	78
PlugX	311
PoisonIvy	75

収集したマルウェアは仮想環境上にてマルウェアを実際に動作させて解析（動的解析）を行う Lastline で解析した．Lastline の解析結果から接続先として抽出されたドメインから，重複および WHOIS の引けないドメイン，マルウェアがインターネットに接続可能な環境かどうかを調査する目的の通信先ドメイン（例えば Yahoo! [39]や Google など）を排除した結果，計 89 件のドメインを得ることができたため，これを今回の実験における C&C ドメインの評価データとして用いた．

ノーマルドメインには安全性の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.” [24]に載っているドメイン 500 件を用いることとした．また，人気サイトはサイト規模が大きい傾向にあるため，特徴量に偏りが生じる可能性がある．そこで,”IR サイトランキング” [40]に載っているドメイン 200 件と”FORTUNE” [41]に載っている 100 件のドメインも用いた．3 サイトに載っている計 800 件のドメインの中から重複ドメインを排除して，C&C ドメインと同数となるようにランダムに 89 件のドメインを抽出し，ノーマルドメインの評価ドメインとして用いた．

4.2 特徴抽出

評価ドメインであるノーマルドメインと C&C ドメインから WHOIS と Google の検索エンジンから得られる情報を抽出する．

4.2.1 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．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有効日数として用いることとした．評価ドメインの有効日数の比較を図 15 に示す．

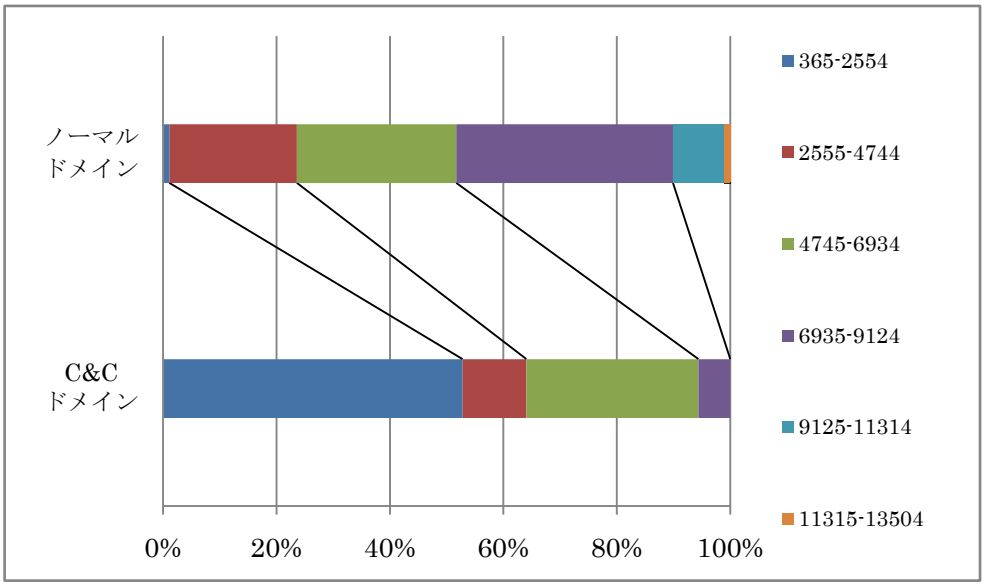
ノーマル
ドメイン
C&C
ドメイン
0%
20%
40%
60%
80%
100%
365-2554
2555-4744
4745-6934
6935-9124
9125-11314
11315-13504

図 15　有効日数の比較

ノーマルドメインと比較して，C&C ドメインは有効日数が短いことがわかる．

他方，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．

a)ID
b)名前
c)組織名
d)住所
e)郵便番号
f)電話番号
g)国名
h)FAX 番号
i)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たりすることが多い．しかし，でたらめな情報が登録されている場合でも，r)メールアドレスは，実際に連絡を行ううえで必要なことが多いため，偽装されていない可能性が高いと考えられる．そのため，まずメールアドレスを対象に特徴点の抽出を行った．比較結果を次に示す．この時，「フリー」はメールアドレスのドメインが無料でメールサービスを提供しているドメインと一致しているものとし，「関係有」は評価ドメインとメールアドレスのドメインが同一もしくは TDL が異なるだけのもの，「登録代行」は WHOIS 登録代行サービスを提供しているドメイン提供元のドメインと一致しているもの，「不明」はそれ以外として判定した（図 16

参照).

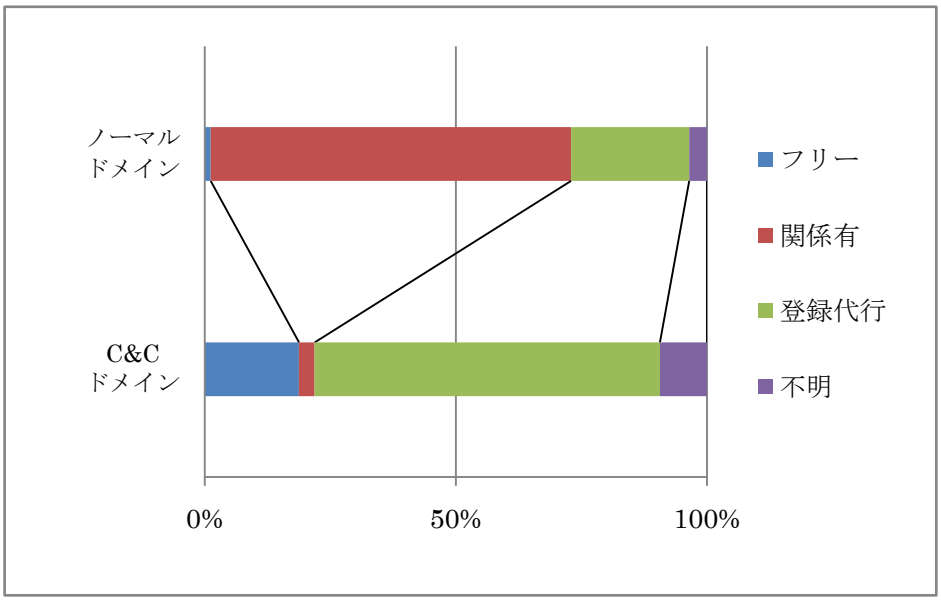

図 16　紐づくメールアドレスの比較

「フリー」と「登録代行」の割合は C&C ドメインが高く,「関係有」の割合はノーマルドメインが高くなる傾向にあり，ノーマルドメインと C&C ドメインにおけるメールアドレスの差異があることが判明した.

以上の結果より,ノーマルドメインと C&C ドメインの WHOIS における特徴の差を図 17 に示す．なお，倫理的な観点より実ドメインの例示は避け，特徴の差を極端に表した例として示した.

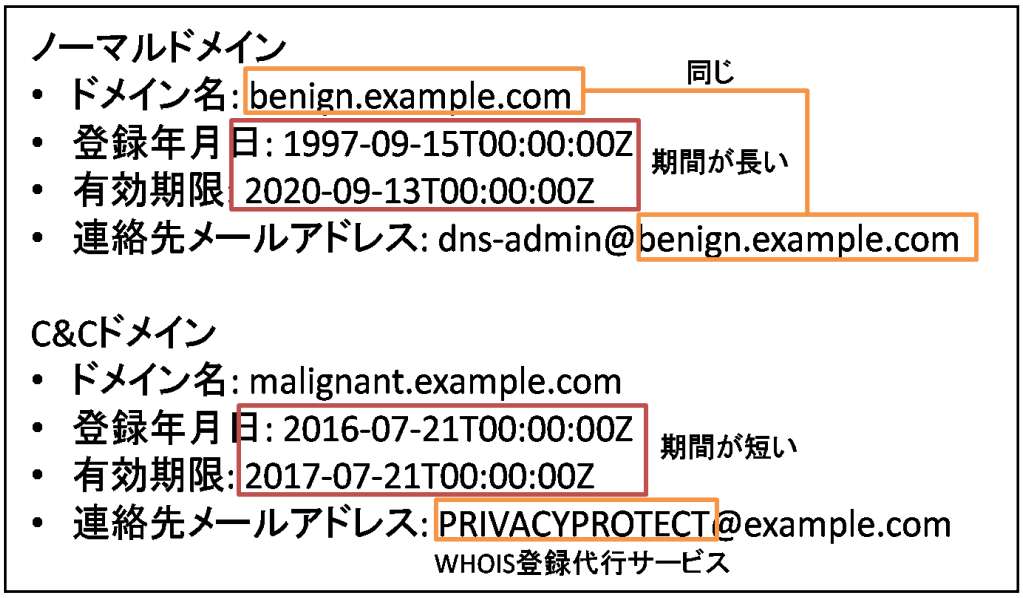

図 17　ドメインの WHOIS 例

ノーマルドメインと C&C ドメインには特徴差があるため，WHOIS よりこれらの有効日数およびメールアドレスを特徴として用いることとした．

4.2.2 検索エンジンからの特徴抽出

既知の悪性ドメインのコンテンツなどから特徴を抽出し，抽出した結果から検索エンジンを用いて新たな悪性ドメインの発見を行っている研究がある．この研究では，ドライブ・バイ・ダウンロード攻撃における悪性ドメインの発見手法を提案しており，ドライブ・バイ・ダウンロード攻撃は Web 閲覧によって攻撃が成功する性質から集客を行うために検索エンジン最適化（SEO）を行っていることが予想される．また，閲覧させるために，正規の Web サイトを改ざんし，悪性ドメインへリダイレクトするスクリプトを仕込んでいることもある．そのため，悪性ドメインもしくは悪性ドメインへのリダイレクト元となる Web サイトは Google 検索にヒットする可能性は高いことが予想される．しかし，標的型攻撃における C&C サーバは短命であり，検索エンジンのクローラにドメインが発見される前にドメインが停止される．さらに，C&C サーバは通信を遮断されないためにも発見されないように隠れていることが予想され，検索にヒットしないことが考えられる．そこで，Google の検索エンジンを用いて評価ドメインが検索でヒットしたかどうか調査した．Google での検索にあたっては，評価ドメインのサイト以外のサイトがヒットしないように「site:」コマンドを用いて「site: 評価ドメイン」となるように検索を行い，検索結果の件数が 0 件であれば「無」，1 件以上であれば「有」とし，その結果を図 18 に示す．

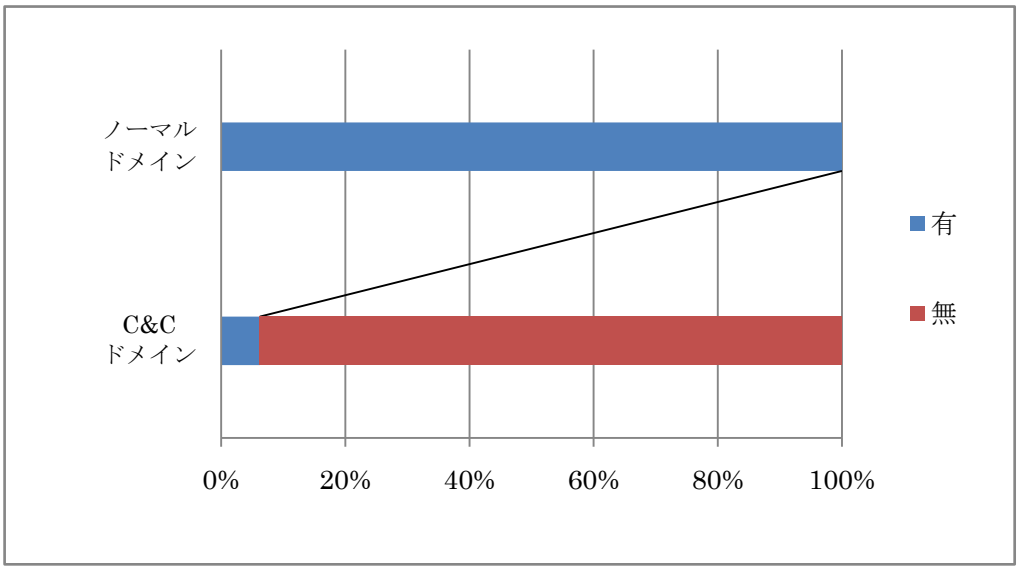

図 18　Google 検索ヒット有無

C&C ドメインにおいては検索にヒットしないものが多数であった．検索にヒットした C&C ドメインにおいては，改ざんもしくはサーバを乗っ取られていた可能性の高い正規のサーバであった．これは，標的型攻撃においては，正規のサーバが乗っ取られて C&C サーバ化されるよりも攻撃者が自前で用意した C&C サーバが用いられているためと考えられる．

以上の結果より，Google での検索結果を特徴として用いることとした．なお，検索のヒット有無だけで差がみられるため，特徴をより抽象化させて精度を向上させるためにヒット件数ではなくヒットの有無を用いることとした．

4.2.1 機械学習アルゴリズム

手法の有効性を検討するため，機械学習アルゴリズムのパラメータチューニングは行わずに実験を行うこととした．また，性能差を検討するため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，与えられたデータからパターン認識を用いて 2 クラスの分類を行う教師あり学習の一種である．高い識別精度で判別を行うことができ，解析を行うデータ量が増加しても高速に識別することができる [42]．そのため，機械学習アルゴリズムの一つとして SVM を選択した．

ニューラルネットワークとは，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教師あり学習の一種である [33]．音声や文字などの識別にも使用されており，誤差逆電伝播法 [43]を用いることで入力と出力のあいだにどういった関係があるのかを表現することが出来る [44]．そのため，単なる数値での識別ではなく，WHOIS 情報と Google の検索エンジンからの特徴と C&C サーバとの関係を学習して識別さ

れることに期待して，機械学習アルゴリズムの一つとしてニューラルネットワークを選択した．

各アルゴリズムにおける訓練モデルを構築する前段階として，メールアドレスは"@"で区切って前半部分のローカルパートと後半部分にドメインに分割，評価ドメインとメールアドレスのドメインとの間での関係の有無，フリーメールアドレスの使用有無，WHOIS登録代行サービスの使用有無を調査する．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と，Google の検索エンジンを用いて評価ドメインが検索にヒットするか調査しておく．これらの情報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（表 9 参照）．

表 9　機械学習への入力値

特徴		入力値
ラベル		ノーマル C&C
ドメイン		(文字列)
メール	ローカルパート	(文字列)
	ドメイン	(文字列)
	タイプ	フリー 関係有 登録代行 不明
有効日数		(数値)
検索エンジン		有 無

第 3 章では，WHOIS に登録されてあるメールアドレス 3 種類を用いていたが,ここでは管理者のメールアドレス 1種類だけとした．WHOIS に登録されてある情報は，情報を管理・保持するレジストラ毎に内容が異なるため，用いるドメインによっては管理者のみのメールアドレスであることも多くあった．その際は取得できない情報はブランクとして第3章では機械学習で分類を行っていた.また，3 種類登録されてあっても，その多くは同一のメールアドレスとなっており，登録のないドメインがブランクとして認識されてしまう分だけ精度を下げてしまう要因になってしまっていることが懸念された．そこで，ここでは管理者のメールアドレスのみを対象にした．さらにメールアドレスそのままデータとして用いるのではなく，メールアドレスの持つ意味に着目した．メールアドレスは"@"の前半部分はローカルパート，後半部分はドメインを表しており，それぞれで意味が異なっているため分割して入力値とした．さらに，メールアドレスの各タイプにおいて，ローカルパートおよびドメインに特徴があることが考えられる．例えば C&C サーバで用いられる登録代行サービスやフリーメールアドレスにも攻撃者にとって都合の良い特定の業者が用いられことが考えられる．そのため，メールアドレスはタイプとあわせてローカルパートおよびドメインをセットで入力値とした．

4.3 評価

今回，評価に用いるデータ量が少ないため，実際に訓練モデルを構築するための訓練データと評価に用いるテストデータを準備する評価手法では，テストデータの選び方によって精度に大きな差が生じる可能性がある．そこで，評価に用いる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．

交差検証法 [34]は,第 3 章でも説明したが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 19） [35]．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．

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

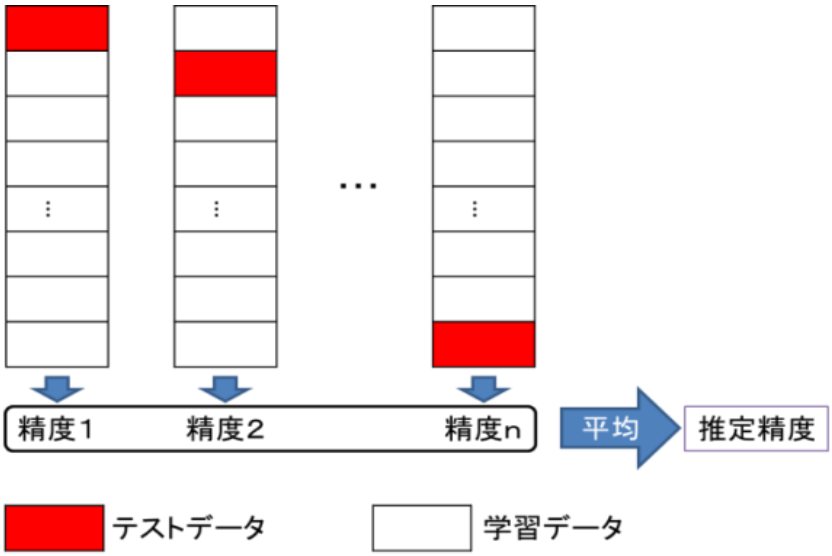

図 19　交差検証法

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結果を表 10 に示す．今回，評価データを 10 分割し，そのうちの一つをテストデータ，残りを学習データとして10回評価を行い導き出された精度の平均を推定精度とした．また，ノーマルドメインと正しく識別された精度を TPR，C&C ドメインと正しく識別された精度を TNR として算出した．

$$\mathrm{TPR} = \frac{\text{正しくノーマルドメインと識別された数}}{\text{ノーマルドメインの総数}}$$

$$\mathrm{TNR} = \frac{\text{正しく C\&}C\text{ドメインと識別された数}}{\text{C\&}C\text{ドメインの総数}}$$

ノーマルドメインの選定による偏りが発生していないかどうかを調査するため，収集した 800 件のドメインの中から重複ドメインを排除して，ランダムに 89 件のドメインを抽出し，交差検証法を用いた評価を複数回行った．その結果，SVM，ニューラルネットワークともに推定精度に大きな変化は見受けられなかったため，本手法を採用し，ノーマルドメイン 89 件で実験を行った．

入力値はラベルとドメイン，メールアドレス，有効日数，検索エンジンの 5 種類に分類される．この中で重みづけをおこなうために，種類ごとに組み合わせを変えて評価を行った．なお，ラベルとドメインは識別を行うのに必須であるため，どの組み合わせにおいても入力値として用いている．

表 10　評価結果（交差検証法）

入力値の組み合わせ	SVM				ニューラルネットワーク			
	推定精度	TPR	TNR	処理時間	推定精度	TPR	TNR	処理時間
①メール	90.4%	91.0%	89.9%	0.03 sec	88.8%	91.0%	89.9%	117.42 sec
②有効日数	70.2%	76.4%	64.0%	0.02 sec	68.5%	71.9%	65.2%	26.19 sec
③検索エンジン	96.6%	100%	93.3%	0.03 sec	96.6%	100%	93.3%	25.91 sec
④メール + 有効日数	93.8%	94.4%	93.3%	0.03 sec	93.3%	93.3%	93.3%	108.25 sec
⑤メール + 検索エンジン	97.8%	100%	95.5%	0.03 sec	97.8%	100%	95.5%	108.09 sec
⑥有効日数 + 検索エンジン	96.6%	100%	93.3%	0.02 sec	96.6%	100%	93.3%	26.41 sec
⑦メール + 有効日数 + 検索エンジン	**98.9%**	**100%**	**97.8%**	**0.02 sec**	**98.9%**	**100%**	**97.8%**	**126.33 sec**

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大きな差異はなく比較的高い推定精度を導き出せた．これは，多段追跡システムの 2014 年モデルの検知率 96.7%を上回る結果である．また，従来は時間経過とともに検出精度は下がる傾向にあったが，今回，新たに解析したマルウェアから得られた C&C サーバのドメインを加えた結果，過去の検知精度である 97.3%を上回る精度であった．これは，時間経過による検出率低下への耐性に期待できる．

SVM とニューラルネットワークの推定精度に大きな差異はなかったものの，処理時間では SVM が高速に処理することができた．特に SVM は入力値が増えても当初の期待通り，比較的高速に安定した速度で処理できたと言える．他方，ニューラルネットワークはどの入力値の組み合わせにおいても SVM を超える推定精度は出ておらず，当初期待していた各種の特徴と C&C サーバとの関係を学習することで単なる数値の識別よりも高い精度を出せたとは言い難い．

表 10 から入力値として検索エンジンの結果が C&C サーバの識別を行うにあたり最も重要な役割を担うことが明らかになった．次いでメールアドレス，有効日数の順に有効であることがわかった．

検索エンジンの結果に有効日数を加えても検知率に変化はないが，メールアドレスと検索エンジンの結果に有効日数を加えたところ，検知率が改善した．これは，単体での識別では検知できなかった C&C ドメインを 3 種類の特徴を組み合わせることで検知可能になったことを意味する．つまり，メールアドレスと有効日数，検索エンジンの結果を組み合わせることが有効であることが示された．

2009 年から C&C サーバの検知における継続調査では，有効日数は経年変化に耐性のあるとても重要な役割を持っていた．これは，C&C サーバは基本的に短命であるため，ドメインの有効日数は短い傾向が変化していないことに由来される．C&C サーバが短命で

あれば，検索エンジンにドメイン情報が収集される前段階でドメインが無効となり検索にヒットしないため，有効日数と同様に経年変化へ耐性があるものと考えられる．

WHOIS に登録されているメールアドレスにも特徴があった．これは，C&C ドメインにおいて，特定の WHOIS 登録代行サービスへの偏り，フリーメールアドレスが使用されていることが大きな要因として考えられる．攻撃者が C&C サーバを準備する際により容易に安価に構築可能なサービスを用いているために特徴が出ているものと推測され，攻撃者が攻撃にかける費用や労力が増えなければ変動しにくい．

4.4 考察

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することなく収集できる情報である WHOIS 情報と検索エンジンから特徴を抽出し，機械学習にかけることにより，C&C サーバの判別ができることを示した．これにより，攻撃者に解析していることを知られずに C&C サーバを検知できたものと考える．さらに，マルウェアから抽出した実データを用いて評価した結果，従来手法より高い検知率で C&C サーバを判別することができた．

従来は C&C サーバを特定するために通信内容や通信先である C&C サーバに直接アクセスまたは接続の際に名前解決により攻撃者が管理するサーバ類へアクセスすることで，攻撃者に解析されていることに気づかれ，対策される危険性があった．本論文では，自組織内の DNS サーバに提案手法を適用することで攻撃者が管理する DNS サーバへの名前解決を行う前に通信を遮断することができる．そのため，攻撃者に知られることなく C&C サーバを判別できる本手法は有効な手法であると言える．

実際に標的型攻撃で用いられるマルウェアを収集して解析した結果では，正規のサーバが乗っ取られて C&C サーバ化している例

は少なかった．しかし，正規のサーバが乗っ取られて C&C サーバ化した場合，WHOIS に正規のユーザの情報が登録されており，さらに，Google 検索にヒットしやすくするため SEO が行われていることが多いため，WHOIS 情報や Google の検索結果からでは差異が出にくく，誤検知が多くなることが懸念される．そのため，正規サーバが乗っ取られたケースでは，他の手法を組み合わせて検知するといった対策が必要になると考えられる．

今後は標的型攻撃以外で用いられている C&C サーバや，経年経過による特徴変化に対しても提案手法が有効であるか調査するとともに，正規サーバが乗っ取られたケースなどに対処できる手法との組み合わせについて検討する．さらに，これらを実装し，実環境での運用を通して処理時間や分析性能といった実用面からの検討を行う．

5. まとめ

第１章ではインターネットの普及からサイバー攻撃および既存の防御方法について概説した．インターネットが普及したことにより利便性が高まる一方で，コンピュータに保存されてあるデータを人質にして金銭を要求したりする問題や情報漏洩といったサイバー攻撃による問題が顕在化してきている。当初、サイバー攻撃は愉快犯などが行う単純な攻撃であったが、年々手口が巧妙化・多様化していき，標的型攻撃といわれる特定の組織や会社を対象として、機密情報の盗み出し，システムの破壊活動を目的とした執拗な攻撃が問題となっていることを述べた．また，そうしたサイバー攻撃に対する防御策についてシステムや機器による防御と企業や組織の運用での防御に大別して，それぞれの方法について述べた．

第２章では標的型攻撃について，攻撃の流れや過去の事例について概説した．標的型攻撃は標的となる企業や組織のことを事前に調査し，調査によって得た情報を用いて，標的となる企業や組織ごとに見つかりにくく，より目的を達成できるようにカスタマイズしたマルウェアを用いてくる．そのため，マルウェアが配送されたときに既存の防御策であるアンチウイルス製品などだけでは検知できないマルウェアが用いられることが多く，多層的・重層的な対策が求められる．

第３章では外部リポジトリから C&C サーバを検知する手法を提案した．C&C サーバのドメインに着目し， DNS と WHOIS，検索エンジンから得られた特徴から機械学習を用いて C&C サーバのドメインの検知を試みた．評価に際して，事前に VirusTotal を用いて実際に標的型攻撃で用いられたマルウェアを収集し，収集したマルウェアを Lastline を用いて C&C サーバのドメインを収集した．そこから，収集した C&C サーバのドメインから交差検証法を用いて本提案手法を評価してみた．その結果，C&C ドメインを 99.3%という高い精度で検知することが可能であった．

第４章では使用する外部レポジトリによっては攻撃者に解析していることをが知られる危険性があることを概説するとともに，第３章で提案した手法をベースに攻撃者に解析していることを察知されないように改良した手法を提案した．評価に際して，第３章同様に事前に VirusTotal を用いて実際に標的型攻撃で用いられたマルウェアを収集し，収集したマルウェアを Lastline を用いて C&C サーバのドメインを収集した．そこから，収集した C&C サーバのドメインから交差検証法を用いて本提案手法を評価してみた．また，第３章よりもより多くのマルウェアを収集し，実験を行った．その結果，攻撃者に解析されていることを察知されずに C&C ドメインを 98.9%という高い精度で検知することが可能であった．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することなく収集できる情報である WHOIS 情報と検索エンジンから特徴を抽出し，機械学習にかけることにより，C&C サーバの判別ができることを示した．

第３章では WHOIS から抽出した特徴点であるメールアドレスを入手できた分だけ入力値として扱っていたが，第 4 章では対象となるメールアドレスを一つに絞り，さらにメールアドレスの持つ意味に着目してメールアドレスを意味のあるパート毎にわけ，さらにメールアドレスを一定の枠組みで分類した結果も付与して学習データとした．その結果，第 3 章では第 4 章で用いた特徴点である WHOIS と検索エンジンだけだと 92.5%だったのに対して，第 4 章では 98.9%まで検知率が向上した．また，ノーマルドメインを C&C ドメインと誤分類したものはなかったため，実際に本提案手法を導入してもノーマルドメインとの通信をブロックする危険性は低いと考える．そのため，第 4 章の手法は誤検知を許容できない箇所に対してとても有効に機能することが期待される．

従来の検知手法では，C&C サーバを特定するために通信内容や通信先である C&C サーバに直接アクセスまたは接続の際に名前解

決などにより攻撃者が管理するサーバ類へアクセスすることで，攻撃者に解析されていることに気づかれる危険性があった．本研究での大きな成果として，攻撃者に気付かれにくい情報のみを用いて検知させることにより，攻撃者に気付かれることなく C&C サーバを判別することができる．

今後は標的型攻撃以外で用いられている C&C サーバや，経年経過による特徴変化に対しても提案手法が有効であるか調査するとともに，正規サーバが乗っ取られたケースなど標的型攻撃にとらわれずより汎用的に応用可能かどうかを検討する．さらに，これらを実装し，実環境での運用を通して処理時間や分析性能といった実用面からの検討を行い，実用化に向けて研究を行う．

謝辞

本論文は，筆者が東京電機大学大学院　先端科学技術研究科　情報工学メディアネットワーク専攻博士課程在学中に行った研究をまとめたものです．本研究をまとめるにあたり終始懇切なるご指導ご鞭撻と格別のご配慮を賜りました本学教授佐々木良一先生，副査を務めて頂いた本学教授絹川博之先生，本学教授齊藤泰一先生，本学教授八槇博史先生に心より感謝致します．

本研究を進めるにあたってご指導ご鞭撻いただきました本学教授猪俣敦夫先生，創価大学名誉教授勅使河原可海先生，本学助教柿崎淑郎先生には謹んで深謝致します．

筆者の所属する佐々木良一研究室の皆様をはじめ，業界関係者の皆様には本研究を進めるにあたり様々な面でご配慮を頂いた．ここに記して御礼申し上げる．

最後になりますが，ありとあらゆる場面で私を温かく見守り続けてくれた家族および親族，友人の皆様に深く深く感謝致します．ありがとうございました．

参考文献

1. 佐々木良一. インターネットセキュリティ入門. 岩波新書, 1999.
2 *ネット社会と本人認証—原理から応用まで*—電子情報通信学会 2010 年
3. 伊東寛. サイバー戦争論　ナショナルセキュリティの現在. 原書房, 2016.
4. 重要インフラ分野における IT 依存度に関する調査. (オンライン) https://www.nisc.go.jp/inquiry/pdf/it_izon_honbun.pdf.
5. 久山真宏, 佐々木良一. ドメインの WHOIS 構造を用いた悪性ドメインの判別手法. DICOMO2016, 2016.
6. 久山真宏, 柿崎淑郎, 佐々木良一. 攻撃者に察知されにくい情報を用いた C&C サーバの判別手法. コンピュータセキュリティシンポジウム 2016 論文集, 2016.
7. M.Kuyama, Y.Kakizaki, R.Sasaki. Method for detecting a malicious domain by using only well-known information. International Journal of Cyber-Security and Digital Forensics 5(4), 2016.
8. 久山真宏, 柿崎淑郎, 佐々木良一. 攻撃者に察知されにくい情報を用いた C&C サーバの検知手法の提案と評価. 情報処理学会論文誌 58(9), 2017.
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), 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. 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), 2014.

12. M.H.Tsai, K.C.Chang, C.C.Lin, C.H.Mao, H.M.Lee. C&C Tracer: Botnet Command and Control Behavior Tracing. IEEE International Conference on Systems, Man and Cybernetics (SMC), 2011.
13. M.Felegyhazi, C.Kreibich, V.Paxson. On the Potential of Proactive Domain Blacklisting. USENIX Conference on Large-scale Exploits and Emergent Threats, 2010.
14. J.Ma, L.K.Saul, S.Savage, G.M.Voelker. Beyond Blacklists: Learning to Detect Malicious Web Sites from Suspicious URLs. ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2009.
15. L.Invernizzi, S.Benvenuti, P.M.Comparetti, M.Cova, C.Kruegel, G.Vigna. EvilSeed: A Guided Approach to Finding Malicious Web Pages. IEEE Symposium on Security and Privacy, 2012.
16. 三原元, 佐々木良一. 数量化理論と攻撃データ（CCCDATAset2009）を利用したボットネットの C&C サーバ特定手法の提案と評価. 情報処理学会論文誌 51(9), 2010.
17. 林知己夫. 数量化―理論と方法 (統計ライブラリー). 朝倉書店, 1993.
18. 中村暢宏, 佐々木良一. 累積データを用いたボットネットの C&C サーバ特定手法の評価. コンピュータセキュリティシンポジウム 2011 論文集, 2011.
19. 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, 2015.
20. S.Okayasu, R.Sasaki. Proposal and Evaluation of Methods Using the Quantification Theory and Machine Learning for Detecting C&C Server Used in a Botnet. 2015 IEEE 39th Annual Computer Software and Applications Conference (COMPSAC), 2015.
21. 標的型攻撃等の脅威について. (オンライン) http://www.nisc.go.jp/conference/suishin/ciso/dai18/pdf/2.pdf.

22. 標的型攻撃対策指南書（第 1 版）. (オンライン) http://www.lac.co.jp/anti-apt/guidebook/pdf/anti-apt_guidebook_ver1.pdf.
23. サイバー演習の種類と概要. (オンライン) https://www.pwc.com/jp/ja/services/cyber-security/red-team-exercises.html.
24. Alexa Top 500 Global Sites. (オンライン) http://www.alexa.com/topsites.
25. 国内標的型サイバー攻撃分析レポート 2015 年版. (オンライン) http://www.trendmicro.co.jp/jp/about-us/press-releases/articles/20150409062703.html.
26. VirusTotal. (オンライン) https://www.virustotal.com/.
27. LastLine. (オンライン) https://www.lastline.com/.
28. UserLocal. (オンライン) https://textmining.userlocal.jp/.
29. DOMAIN NAMES - CONCEPTS AND FACILITIES. (オンライン) https://www.ietf.org/rfc/rfc1034.txt.
30. DOMAIN NAMES - IMPLEMENTATION AND SPECIFICATION. (オンライン) https://www.ietf.org/rfc/rfc1035.txt.
31. 「ウェブサイトを閲覧しただけでウイルスに感染させられる"ドライブ・バイ・ダウンロード"攻撃に注意しましょう！」. (オンライン) http://www.ipa.go.jp/files/000008801.pdf.
32. V.Vapnik, A.Lerner. Pattern recognition using generalized portrait method. Automation and Remote Control.24, 1963.
33. Multilayer Perceptron. (オンライン) http://deeplearning.net/tutorial/mlp.html.
34. R.Kohavi. A study of cross-validation and bootstrap for accuracy estimation and model selection. Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence 2 (12), 1995.
35. モデルの精度を推定する. (オンライン)

http://musashi.osdn.jp/tutorial/mining/xtclassify/accuracy.html.
36. NICNAME/WHOIS. (オンライン)
https://www.ietf.org/rfc/rfc812.txt.
37. RFC3912WHOIS Protocol Specification. (オンライン)
http://www.ietf.org/rfc/rfc3912.txt.
38. Google. (オンライン) https://www.google.co.jp/.
39. Yahoo! JAPAN. (オンライン) http://www.yahoo.co.jp/.
40. IR サイトランキング. (オンライン)
http://www.gomez.co.jp/ranking/ir/index.html.
41. FORTUNE. (オンライン) http://fortune.com/.
42. P.John. Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines. Technical Report MSR-TR-98-14, 1998.
43. Rumelhart,D.E., Hinton,G.E., Williams,R.J. Learning representaions by backpropagating errors. Nature Vol.323-9, 1986.
44. −. Parallel Distributed Processing: Explorations in the Microstructure of Cognition: Foundations. MIT Press, 1986.
45. 人工知能研究. (オンライン)
https://www.ai-gakkai.or.jp/whatsai/AIresearch.html.

研究業績

A. 査読論文

[1]**M. Kuyama**, Y. Kakizaki, and R. Sasaki. Method for detecting a malicious domain by using only well-known information, International Journal of Cyber-Security and Digital Forensics 5(4), 166-174.

[2]久山真宏, 柿崎淑郎, 佐々木良一. 攻撃者に察知されにくい情報を用いたC&Cサーバの検知手法の提案と評価, 情報処理学会論文誌 58(9), 1410-1418.

B. 学会発表

[1]久山真宏. 情報漏えい発生時のレピュテーショナルリスクと危機管理～システム監査の視点から, システム監査学会第29回研究大会 (2015).

[2] 木村裕一, 赤尾嘉治, 久山真宏, 桜井由美子, 西澤利治. 中小企業へのサイバー攻撃を防御するためのCSIRT導入の研究, システム監査学会第29回研究大会 (2015).

[3] 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, CSEC-71, (2015).

[4]木村裕一, 赤尾嘉治, 久山真宏, 桜井由美子, 西澤利治. 中小企業へのサイバー攻撃を防御するためのCSIRT導入の考察, システム監査学会第30回研究大会 (2016).

[5] 久山真宏, 佐々木良一. ドメインのWHOIS構造を用いた悪性ドメインの判別手法, DICOMO2016 (2016).

[6] 渋谷健太, 久山真宏, 佐藤信, 三村聡志, 松本隆, 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－標的型攻撃マルウェアの解析と亜種の予測－, DICOMO2016 (2016).

[7] 島川貴裕, 久山真宏, 佐藤信, 名和利男, 高倉弘喜, 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－模擬 C&C サーバを用いたマルウェアの挙動解析－, DICOMO2016 (2016).

[8]久山真宏, 柿崎淑郎, 佐々木良一. 攻撃者に察知されにくい情報を用いた C&C サーバの判別手法, コンピュータセキュリティシンポジウム 2016, (2016).

[9] **M. KUYAMA**, Y. KAKIZAKI, R. SASAKI. Method for Detecting a Malicious Domain by using WHOIS and DNS features, The Third International Conference on Digital Security and Forensics (DigitalSec2016), (2016).

[10]島川貴裕, 佐藤信, 久山真宏, 佐々木良一. 標的型攻撃に対する侵害範囲特定ツールの開発と評価, CSEC-76, (2017).

[11] 渋谷健太, 久山真宏, 松本隆, 八槇博史, 佐々木良一. 標的型に対する知的ネットワークフォレンジックシステム LIFT の開発と機能拡張（その 4）－将来起こりうる攻撃方法の推定－, CSS2017 (2017).

C. 解説記事

[1] 久山真宏. 無線 LAN にまつわるセキュリティの課題を再確認しよう, @IT.http://www.atmarkit.co.jp/ait/articles/1504/22/news002.html

[2] 久山真宏, 林侑輝. 八槇博史先生インタビュー「セキュリティの視点から考える,AI の強みと盲点」, 人工知能 30(6).

[3] 土斐崎龍一, 久山真宏. 佐藤浩先生インタビュー「楽観的に, あきらめ悪くやり続ける」, 人工知能 31(6).

D. 受賞

[1] 優秀プレゼンテーション賞：久山真宏, 佐々木良一. ドメインの WHOIS 構造を用いた悪性ドメインの判別手法, DICOMO2016 (2016).

Appendix A 人工知能と機械学習

１．人工知能

人工知能（AI）とは学習や学習した情報から推理したりできる機械のことをいう．人間のように知能を持ち行動を行えるようにする機械を作るのが目的であり，これを強い AI と言う．しかし，実際の研究は人間の知識・行動のごく一部を実現する弱い AI と言われる研究が主流である．人工知能の定義として，人間にできてコンピュータにできない部分のギャップを人工知能と捉える考え方もある．

なお，人工知能について様々な考え方があるが，厳密な定義は存在しない．

２．人工知能の種類

弱い AI と言われる人工知能研究には以下のような様々な技術が存在する（図 20 参照）．

図 20　人工知能の技術

（「人工知能研究」 [45]より引用）

これらの技術を応用したり組み合わることで特定の分野に特化した人工知能の研究が盛んである．このような特定の処理に特化したAIのことを特化型人工知能とも言われ，囲碁をはじめとした一部の分野では，特化型人工知能は既に人を超えた処理が可能である．

本研究で用いた機械学習は人工知能研究における一分野であり，人間の学習過程を実現するための技術である．複数のデータの中から規則性を見つけ出しデータを分類したり，そこから新たなデータを予想したりする．複雑なデータの分類や大量のデータ処理に関して期待される技術である．

３．機械学習の種類

機械学習には大きく分類すると（１）教師あり学習，（２）教師なし学習，（３）強化学習の３種類が存在する．

（１）教師あり学習

事前に評価データとして情報とそれに対応するラベルを紐付けたデータをアルゴリズムで学習させておき，学習した内容から未分類のデータに適切なラベルを振り分ける．過去のデータから特徴を見つけ出し，新たなデータに適用することで判別を行うことができる．

過去の経験から推測することができるため高い精度を得ることができるが，事前に学習させるためのデータが必要である．

（２）教師なし学習

事前にデータを学習しておく必要がなく，与えられたデータから規則性を見つけ出し，データの判別を行う．複数のデータの中から規則的なデータ集団の特徴をみつけだし，特徴のあるグループとそれ以外に判別することができる．新たなデータがどのグループに特徴が近いかを判別することで分類を行う．

教師あり学習に比べて事前にデータを準備する必要はないが，正解がわからないなかから分類を行うため精度を上げにくい特徴がある．

（３） 強化学習

行動により得られる結果を学習して最適な行動を選択する．規定された環境のなかで行動を行うと確率が変化する状況下で，AI が自ら行動し，確率の変化を学習していき最適な行動を行う．規定された環境下で AI 自らが学習するデータを作り出して学習していく特徴がある．自ら行動して学習するため，環境の変化にも対応できるが，行動の結果を観測できない場面での適応は難しい．

４．本研究における機械学習

本研究では機械学習の中でも，より高い精度を求めて教師あり学習を使用した．今回，Appendix B Lastline&VirusTotal で解説する手法を用いることで，事前に学習するためのデータを収集することができ，かつ学習データがノーマルドメインか C&C ドメインかを事前に判断できるため，教師あり学習が最適であると考えた．また，教師あり学習にも様々なアルゴリズムが存在する（図 21 参照）．

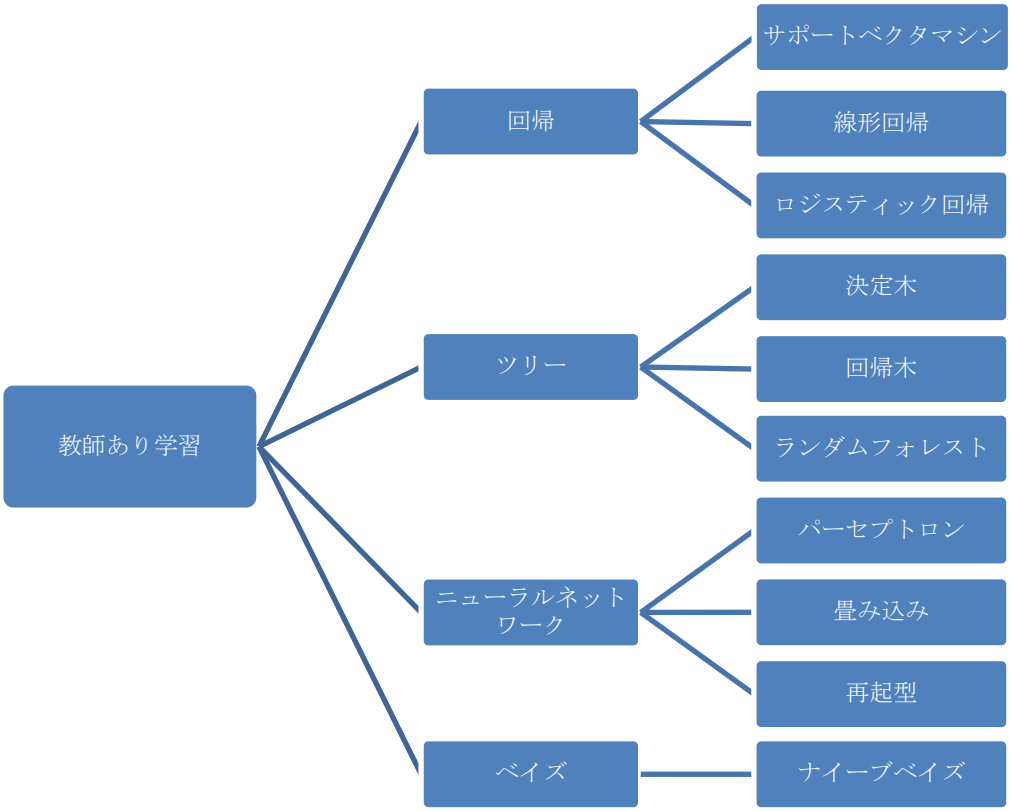

図 21　教師あり学習の分類

数あるアルゴリズムの中でも，本研究ではサポートベクタマシン（SVM）とニューラルネットワークを教師あり学習のアルゴリズムとして選択した．

５．サポートベクタマシン（SVM）

SVM は与えられたデータからパターン認識を用いて 2 クラスの分類を行う教師あり学習の一種である．学習するデータを点としてとらえ，点と点の距離が最大となる基準を求めて分類を行う．例えば図 22 では事前に青点と赤点を学習させ，青点と赤点の各々の距離が最大となる位置に線を引き，この線を基準として学習データを分類する．

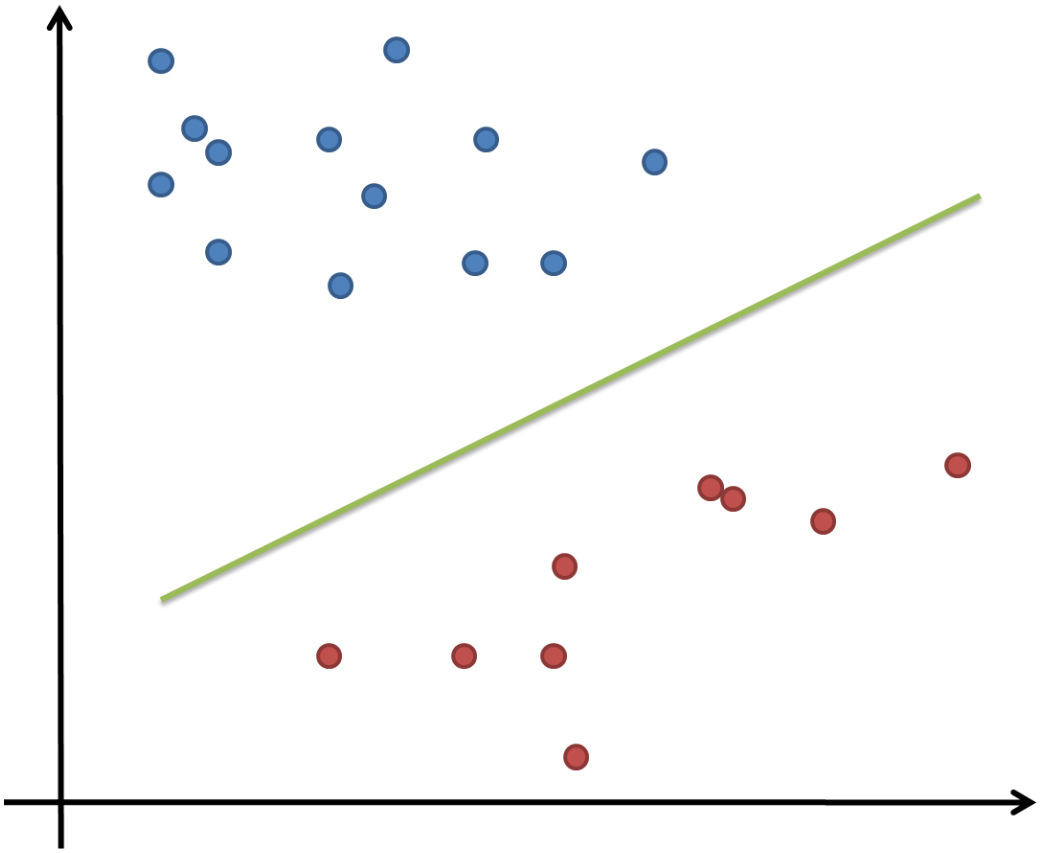

図 22 SVM

2 分類する分類機の中でも高い識別精度で判別を行うことができ，解析を行うデータ量が増加しても高速に識別することができる．また，広く一般的に用いられており，実装するツールも豊富に存在する．そのため，本研究では SVM を使用した．

６．ニューラルネットワーク

ニューラルネットワークとは，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教師あり学習の一種である．単純な計算を行うニューロンと言われるユニットを用いる．ニューロン同士を接続して，接続部分に重みづけを行う．入力に対してニューロンは決められた単純な計算を行い，接続しているニューロンの重みづけを加えて出力を行う．ニューロンの接続部分の重みづけによって分類される．狭義には多層パーセプトロンといわれる3層以上の層にニューロンを分類して層毎に各ニューロンを接続して重みづけを行うもののことを差し，入力層，中間層，出力層の 3 層からなる 3 層パーセプロトンが広く利用されている（図 23 参照）．

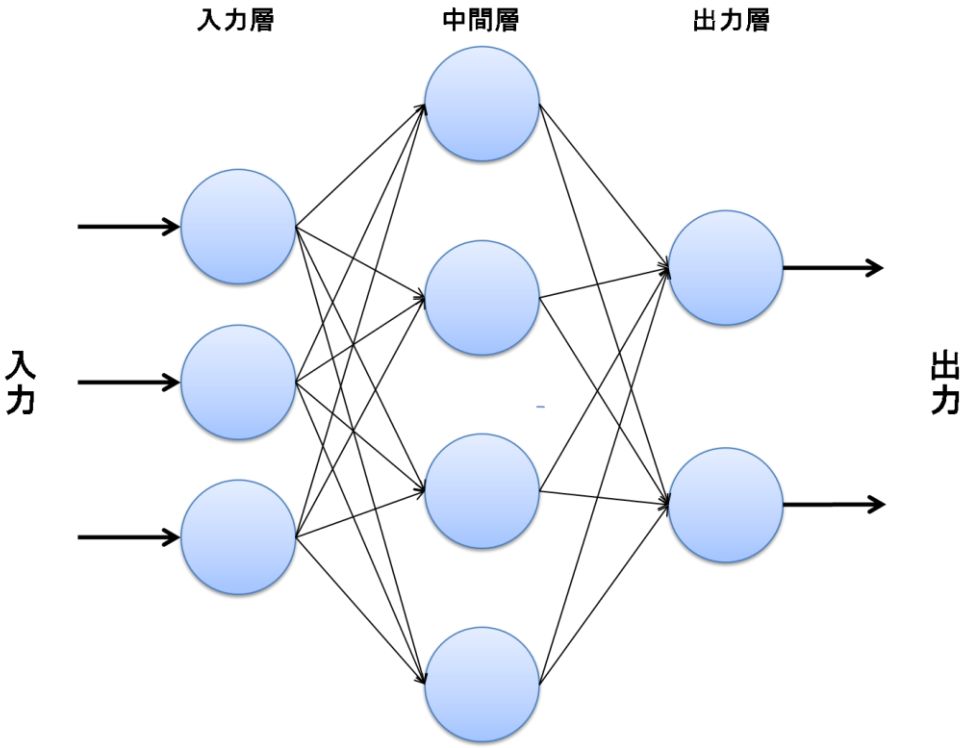

図 23　ニューラルネットワーク

本来は学習データの入力層から出力層に対して重みづけを行う．しかし，誤差逆電伝播法を用いることで入力から出力を求めた後に，望ましい出力との誤差が最小となるように出力層から入力層にかけて重みづけを修正することができる．また，ニューロン同士の重みづけによって各ニューロン同士の関係を求めることができる．そのため，単なる数値での識別ではなく，入力と出力の関係を重みづけから判断することができる．今回，WHOIS 情報と Google の検索エンジンからの各特徴と C&C サーバとの関係を明らかにし，より高い精度で識別されることに期待して，本研究ではニューラルネットワークを使用した．

Appendix B Lastline&VirusTotal

１．VirusTotal

複数のアンチウイルスベンダの検知技術を用いてファイルやwebサイトの解析を行うwebサービス．

ファイルをアップロードすることで，複数のアンチウイルスベンダの検知結果を一覧で表示することができる．

有償サービスであるVirusTotal Intelligenceに契約することで，過去にVirusTotalにアップロードされたデータを検索・ダウンロードすることができる．そのため，個人情報や機密情報が含まれたデータがアップロードされると世界中の人に閲覧される危険性がある．

本研究では，実際に標的型攻撃に使用された検体からC&Cサーバのドメインを入手するためにVirusTotal Intelligenceと契約を行い，検体の入手して研究に用いた．

なお，VirusTotalは当初，Hispasec社により開発・運営されていたが，2012年にGoogle社に買収されて以降，Google社により管理・運営されている．

２．Lastline

Lastline社の提供する商用のサンドボックス．特に標的型攻撃に用いられるマルウェアの検知・解析に特化しており，マルウェアの挙動から検知を行う．また，解析する対象のマルウェアがC&Cサーバへの接続を行った場合に，偽のC&Cサーバと通信を行わせてその挙動を解析することができ，これによりC&Cサーバを独自の技術で特定することができる．

標的型攻撃に用いられるマルウェアは，耐解析機能と言う検知されないようにする技術が用いられていることが多い．そのため，従来のサンドボックスでは解析できない検体も多くある．Lastlineは独自の技術により耐解析機能を持ったマルウェアの解析を可能と

しており，より多くの検体から C&C ドメインを抽出できると期待して使用することにした．

３．本研究での活用

本研究では，VirusTotal から検体をダウンロードし、これを Lastline で解析した（図 24 参照）．解析結果から得られた C&C サーバのドメインとして研究に用いた．

図　24　VirusTotal & Lastline

①の条件で入手したい検体名を指定して検索する．今回は「Emdivi」「PlugX」「PoisonIvy」を指定して検体をダウンロード(②)してきた．

ダウンロードできた検体は Lastline に送信（③）して，サンドボックスで動作した挙動より C&C ドメインを抽出（④）した．