Vol.2017-IOT-39 No.3
Vol.2017-SPT-25 No.3
2017/9/29

DNSグラフ分析に基づく悪性ドメインの検知手法の提案

浅井 麻友子[1]　今泉 貴史[2]

概要：DNS (Domain Name System) はドメイン名と IP アドレスの対応付けを管理する階層型分散システムで，インターネット通信に欠かせないものとなっている．一方でフィッシングやボットネット，スパムなどのサイバー攻撃に利用されることが問題となっている．本研究では，DNS グラフを分析することで，攻撃に利用される悪性ドメインとそれに関連する潜在的な悪性ドメインを検知する手法を提案する．この手法により，既存の特徴ベースの検知手法よりも，悪性ドメインの検知率を向上することが期待できる．

Detecting Malicious Doamins through DNS Graph Analysis

Asai Mayuko[1]　Imaizumi Takashi[2]

1. はじめに

DNS はドメイン名と IP アドレスの変換を行う階層型分散システムであり，Web サイトの閲覧や電子メールの送信などインターネット通信のほとんどは DNS 名前解決を必要とする．

一方で，DNS を悪用した様々な攻撃が報告されている．例えば DNS トンネリングは，DNS 通信が 53 番ポートを経由して必要なデータを送信するためにファイアウォールなどを容易に通過する特徴を利用し，ボットネット通信の隠れチャネルとして用いる．その際にドメイン名の TTL を極端に短く設定し，後の調査で追跡することを困難にしている．

また，攻撃に利用するドメインの A レコードを短い TTL で頻繁に変えていく fast-flux という手法がある．A レコードに登録されている Botnet 内の IP アドレスを変化させることで，マルウェアを配信するサイトやフィッシングサイトをより長い時間活動させることを目的としている．

このようにインターネットに欠かせない DNS プロトコルを利用することで，ウィルスに遠隔操作された端末と攻撃者との通信を検知することを困難にしている．また，攻撃者が対象者の DNS 運用状況を十分に把握して攻撃を行う一方，多くの企業や組織でこの脅威に対する対策が十分でないことが，深刻な問題として挙げられる．

本研究では，これらの特徴を基にドメインごとの悪性度を計算し，さらにドメイン間の関連性を考慮して悪性度を精査することで，悪性ドメインを検知する手法を提案する．第一段階での，それぞれのドメインに対して行う悪性度を計算には，複数の特徴を考慮し，網羅的に悪性ドメインを検知する．しかし，攻撃者によってドメインの特徴を操作し，特徴ベースでの検知を回避する例がある．このようなドメインを検知するために，第二段階で悪性度の精査を行う．第二段階では，悪性ドメインは悪性ドメインと関連を持ち，良性ドメインは良性ドメインと関連を持つという特性を利用する．ドメイン間の関連性を考慮して悪性度を精査するとによって，特徴ベースでは検知できなかった悪性ドメインの検知を実現する．これによって攻撃に利用される悪性ドメインを効率的に検知し，且つ検知率を高めることが期待できる．これらの提案手法を，既知の悪性ドメインと良性ドメインを使用して悪性ドメインを求め，悪性ドメインの検知がどの程度できるか実験・評価を行う．

2. 関連研究

攻撃に利用される悪性ドメインには，ドメイン名やドメイン情報において良性ドメインと比較して特徴があることが多い [1]．ドメイン名の WHOIS 情報から得られる有効期間が悪性ドメインは比較的短い [2] ことや，ドメイン名のエントロピーにも差が生じる [3][4].

[1] 千葉大学融合科学研究科
[2] 千葉大学統合情報センター

悪性ドメインを検知する手法として，DNS レコードや DNS クエリ・レスポンスから抽出される特徴に基づいて分類・検知する方法がある．攻撃者は DGA(Domain Generation Algorithm) を利用し，動的にドメインを生成する．これは，ランダムな文字列としてドメイン名を生成するために，人が考える文字列とは異なる特徴が表れる．このランダム性を利用し，ドメインの特徴を N-gram を用いて抽出するアノマリ型の検出手法 [6] がある．しかしこの手法は，単語を連結したドメインなどには対応できず，検知手法として十分ではない．悪性ドメインの検知には，ドメインのランダム性以外にも注目する必要がある．

DNS メッセージのペイロードから得られた特徴を複数用いてクラスタ分析を行い，fast-flux や DNS-tunneling のようなボットネット検知から逃れるための手法を見つける研究 [5] がある．これは，ドメイン名の長さ，ドメイン名と DNS メッセージのエントロピー，TTL，DNS の uncommon レコードの使用など多くの特徴を用い，複数の攻撃手法を検知する．この論文では，それぞれの特徴に基づいて，ボットネットで用いられる手法である cycling of IP mapping, domain flux, faset-flux, DNS-tunneling にクラスタ分析をしている．攻撃手法によって計算に考慮する特徴を変化させ，重みづけた結果をラベルに反映させている．

一方、既知の悪性ドメインを利用して未知の悪性ドメインを検出する手法 [7][8][9] がある．DNS 通信におけるドメイン，IP アドレス，ホストの関係性から，ドメインをノードで表現した DNS グラフを作成し，解析して悪性ドメインを検知する．

[7][8] では，DNS トラフィックデータから A レコードの問い合わせドメイン名とその回答である Resolved IP アドレスを用いて，2 部グラフである DNS グラフを構築する．確率伝搬法によって事前情報が既知であるノードの確率から未知のノードの周辺確立を計算し，ドメイン間の関連を調べることで，特徴ベースでは検知することができなかった悪性ドメインを検知することが可能である．これらの手法は，既知の悪性ドメインと強い関連性があるドメインは悪性である可能性が高いことを利用して，悪性ドメインとどの程度関連があるか計算し，閾値以上のドメインを悪性ドメインとして検知する．この手法では，特徴ベースで悪性ドメインを検知する手法に比べ，誤検知を減少することができる．また，攻撃者がドメイン間の関連度を操作して，検知を回避することは困難である．[9] は，ドメイン間でそのドメインに解決される IP アドレスが共通する数が多いほど，関連度が高くなるとし，既知の悪性ドメインと関連性が高いドメインを悪性度する手法である．しかしこれらの検知手法では，事前情報を準備する必要性があり，その情報の信用度を確保しなければならない．

3. 提案手法

本研究では，特徴ベースで悪性ドメインと検知する手法とドメイン間の関連度を調べて検知する手法を組み合わせた手法を提案する．まず、特徴ベースでドメインごとに初期悪性度を計算する．そしてドメインと対応する DNS サーバの IP アドレスを用いて DNS グラフを構築・分析し，関連性を考慮して悪性度を精査することで悪性ドメインを検知する．

複数の特徴を考慮して，初期悪性度を計算することで，さまざまな攻撃に対応するようにする．ドメイン名の乱雑性のみに注目すれば，あらかじめ単語を登録してたリストからランダムに単語を組み合わせてドメインを生成するリスト型の DGA に対応することができない．DNS レコード以外にも WHOIS 情報の登録日や有効期限といった複数の情報源から取得したほうが，検知精度が増すと考えられる．

初期悪性度の計算については，0 から 1 までの悪性度を求める必要がある．[5] などの複数の特徴を用いてクラスタ分析を行うい攻撃を検知している先行研究の手法を用いれば，初期悪性度の求めた段階で，かなり精度の高い結果が得られると考えられる．しかし，今回は初期悪性度を数値化する必要性があり，[5] で扱った攻撃手法以外に用いられる悪性ドメインについても検知できるよう，本研究では，独自の手法で初期悪性度を求める．

初期悪性度の精度を高めるために，ドメイン間の関連性を踏まえた，初期悪性度を精査して，悪性度を求める．ドメイン間の関連性は，攻撃者によって操作することが難しい要素である．特徴ベースの検知を回避することのできない方法をとることによって，検知の精度を高める．

以上のように，DNS メッセージをキャプチャしたデータから抽出した情報や，レコードの内容，WHOIS 情報を使用するため，事前に悪性・良性ドメインリストなどを準備する必要がない．更に特徴ベースでの検知を，ドメイン間の関連性を用いて精査することにより，検知率の向上を実現する．

検知の手順は以下の通りである．

3.1 DNS リゾルバで DNS メッセージをキャプチャ

ローカルの DNS サーバで DNS メッセージをキャプチャしたものを，集積ポイントに転送してデータベース化して Passive DNS データを作成する．悪性ドメインには，ドメイン名のエントロピーが高い，ドメイン名が長いなどの特徴がある．データベースではドメイン d ごとに，考慮する特徴それぞれの値を保持する．

また，DNS グラフを作成する際に必要な情報も保持する．ドメイン d に対して，ドメイン間の関連性を計算するための情報もここで抽出する．

Vol.2017-IOT-39 No.3
Vol.2017-SPT-25 No.3
2017/9/29

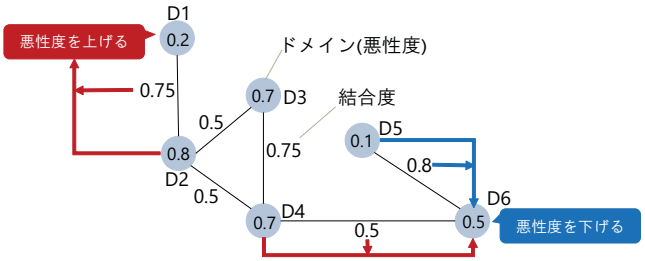

図 1　DNS グラフ

3.2　初期悪性度を計算

DNS メッセージを解析した特徴量をもとに，ドメインごとに初期悪性度を計算していく．

3.3　DNS グラフを作成

DNS グラフは，ノードをドメインとし，関連があるドメイン同士を結んだ無向グラフである．ノード間の関連の度合を数値化した値をエッジに重みづけしていく．ノードには，そのドメインの悪性度を保持している．

3.4　ドメイン間の結合度によって悪性度を精査

あるノードに対し，隣接しているノードの悪性度と結合度から新たな悪性度へ値を精査する．

悪性度の精査の例を図 1 に示す．例えば，悪性ドメインと検知するための閾値を 0.45 とおくとする．このとき D1 の悪性度を計算すると，式 (3) より，0.46 となり，閾値を超え悪性ドメインと判断する．D6 の場合，悪性ドメイン D4 と良性と思われる D5 の 2 つのドメインと関連があり，悪性度は 0.40 になる．この場合，閾値を超えないため良性ドメインと判断する．

設定した閾値より高い悪性度をもつドメインを，悪性ドメインとして検知する．

4.　実験

提案手法でどの程度検知率が向上するのか，実験を行う．今回，良性ドメインとして Majestic Million[10] のトップアクセスドメイン，悪性ドメインとして，DNS-BH[11] から，それぞれ 7115 個のドメインを使用した．またすべて TLD が”com”であるドメインを使用した．

本実験では，ドメインの長さ，文字の種類数，エントロピー，NS レコードと WHOIS 情報におけるネームサーバの一致度，ドメインの有効期間の特徴を調べ，悪性度を計算するための要素としての有効性を調べた．

調べた特徴量をもとに，初期悪性度を求め，さらにドメインごとに管理するネームサーバの IP アドレスリストからドメイン間の関連度を計算し，悪性度を再計算する．最終的な悪性度をもとに良性または悪性を判断し，悪性ドメインが検知されるかを確認する．

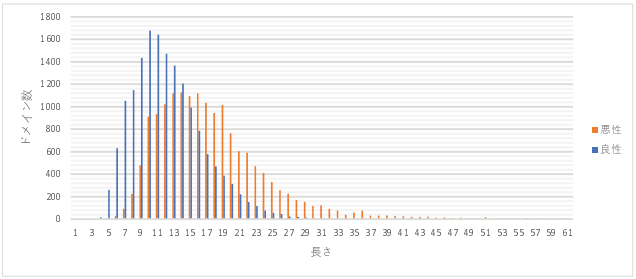

図 2　ドメインの長さ

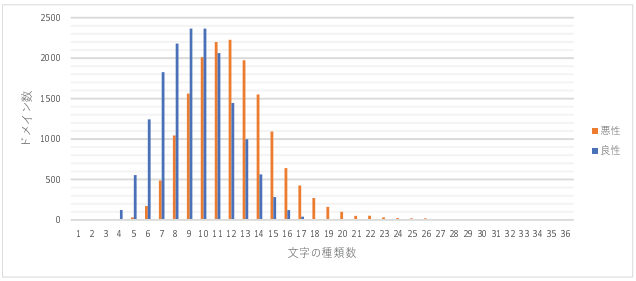

図 3　ドメインに用いられる文字の種類数

4.1　特徴分析

まず良性ドメイン，悪性ドメインごとに，ドメインの長さ，文字の種類数，エントロピー，NS レコードと WHOIS 情報におけるネームサーバの一致度，ドメインの有効期間について調べる．

4.1.1　ドメインの長さ

今回使用したデータでのドメインの長さの分布は図 2 の通り，悪性ドメインほど，ドメインが長くなることが分かる．

4.1.2　文字の種類数

DGA で動的に生成されるドメインは，子音母音関係なく文字をランダムに並べていくため，ドメインに使われる文字の種類数も増加する．その特徴を調べた結果を図 3 に示す．

4.1.3　エントロピー

悪性ドメインにはランダム性が高い文字の組み合わせになることが多いため，エントロピーが高くなる傾向がある．本研究では，$P(x_I)$ をドメイン $X = \{x_1, x_2, ..., x_n\}$ においてある文字列 x_i が出現した確率としたときの式 (1) を用いて求める．本データでのエントロピーの分布の様子を図 4 に示す．

$$E = -\sum_{i=1}^{b} P(x_i) log P(x_i) \qquad (1)$$

4.1.4　NS レコードと WHOIS 情報におけるネームサーバの一致度

NS レコードのネームサーバ一覧が WHOIS 情報から取得できるネームサーバ一覧とどれくらい一致するのかを調べた．一致率は，(NS レコードと WHOIS 情報で共通するネームサーバの数) / (NS レコードのネームサーバの数) で求めた．図 5 のように，良性ドメインは 80%以上のドメイ

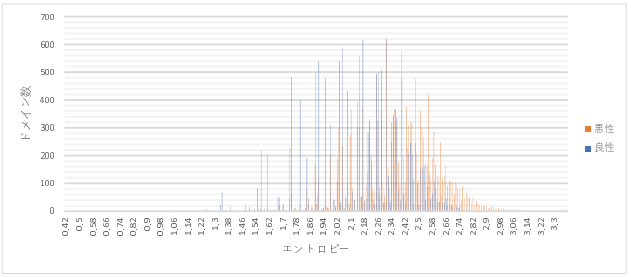

図 4　エントロピー

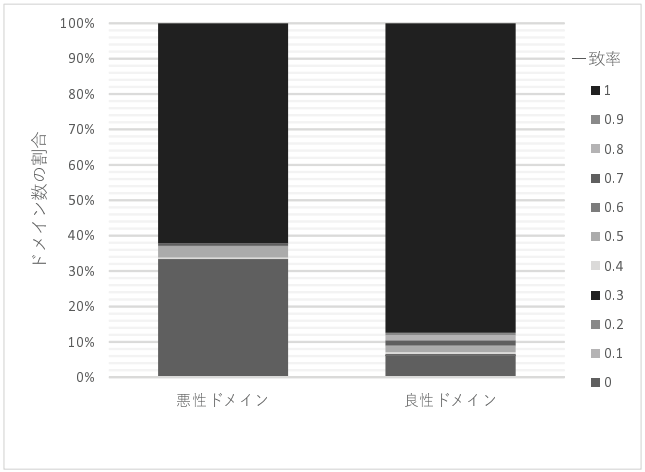

図 5　ネームサーバの一致率

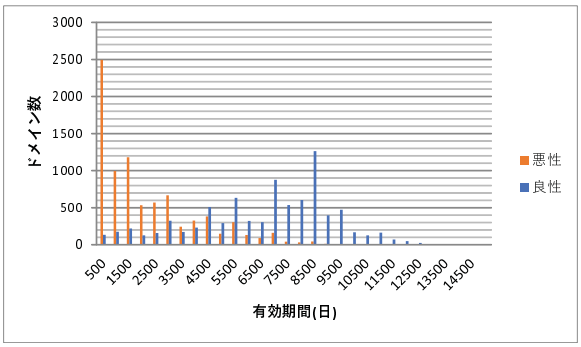

図 6　ドメインの有効期間

ンが NS レコードと WHOIS 情報のネームサーバが一致してるが，悪性ドメインは完全一致するドメインは約 60%である．

4.1.5　ドメインの有効期間

WHOIS 情報から，ドメインの登録日と有効期限を取得し，それらの差を有効期間 (日) とした．結果を図 6 に示す通り，悪性ドメインほど，ドメインの有効期間は短い．

4.2　結合度の求め方

ある 2 つのドメインに対して，共通するネームサーバの IP アドレスがあると，そのドメイン間に関連性があると考えた．本実験では，管理する DNS サーバの IP アドレスの共有する数が多いほどドメイン間の結合度が強くなり，エッジに重みづけする．ドメイン d に対して，そのドメインの管理するネームサーバのリスト $L_{ip}(d)$ に基づいてドメイン間の結合度求めていく．ドメイン d_1 とドメイン d_2 の結合度 $w(d_1, d_2)$ は，以下の式 (1) で求める．結合度が 0 であれば，共通するネームサーバの IP アドレスがないため，ドメイン同士のエッジは存在しない．共通する IP アドレスが多いほど，結合度の重みは増していく．

$$w(d_1, d_2) = \begin{cases} 1 - \frac{1}{1+|L_{ip}(d_1) \wedge L_{ip}(d_2)|} & (d_1 \neq d_2) \\ 1 & (\text{otherwise}) \end{cases} \quad (2)$$

4.3　悪性度の精査方法

ノード i に対し，隣接ノードの集合を N_i とする．ノード i における悪性度の初期値を ϕ_i とする．本実験では，悪性度を精査するために，ノード i の再計算した悪性度 mal_i を式 (3) で求める．

$$mal_i = \frac{1}{\sum_{j \in N_i} w(d_i, d_j)} \sum_{j \in N_i} \phi_j w(d_i, d_j) \quad (3)$$

ノードは全て各ノードの初期悪性度を用いて精査し，グラフ全体を更新していく．ノード自身との結合度を 1 と定義し，そのノードと関連するノードの初期悪性度を用いて悪性度を求める．

5.　結果

それぞれ特徴別の悪性度 = (悪性数) / (悪性数 + 良性数)$(0 < \phi \leq 1)$ を求める．求めた特徴別の悪性度はドメインの相乗平均をとり，ドメインの初期悪性度とした．初期悪性度の散布図を図 7 に示す．

次にドメイン間の結合度をもとに悪性度を精査する．精査した後の悪性度の散布図を図 8 に示す．

閾値を 0.5 とし，悪性，良性と判断した結果が表 1 である．この結果より，ドメイン間の結合度を考慮して悪性度

表 1　悪性度の精査結果

		良性ドメイン	悪性ドメイン
初期悪性度	良性と判断	40.2%	12.1%
	悪性と判断	9.8%	37.9%
精査後	良性と判断	42.0%	9.4%
	悪性と判断	8.0%	40.6%

を計算するとによって，false positive rate, false negative rate 共に減少し，検知の精度が上がっていることが明らかである．悪性度の精査に，ドメイン同士の関連度を考慮することで，検知率が向上した．よって，悪性ドメイン同士，または良性ドメイン同士で，共通するネームサーバが多いことが分かる．本実験でのドメイン間の結合を表した様子の一部が図 9 の通りである．青い点は既知の良性ドメインで，オレンジの点は悪性ドメインである．同じ色通りの点が集合となった塊がいくつか見えるとおり，良性ドメイン

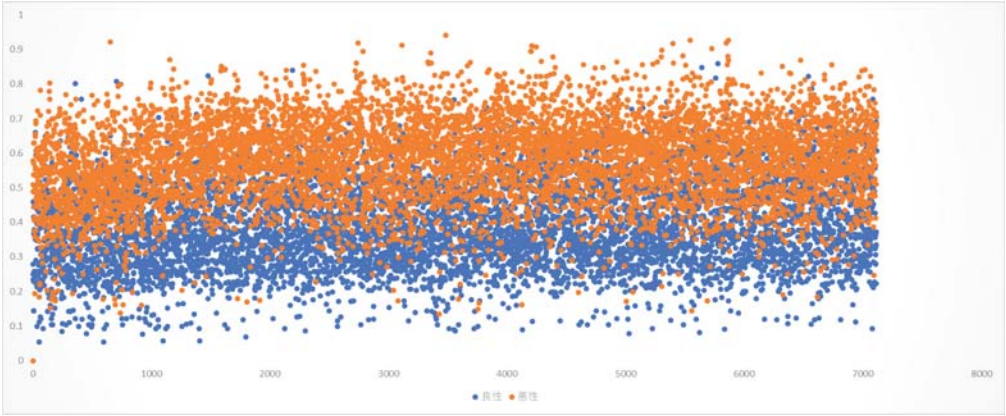

図 7　初期悪性度

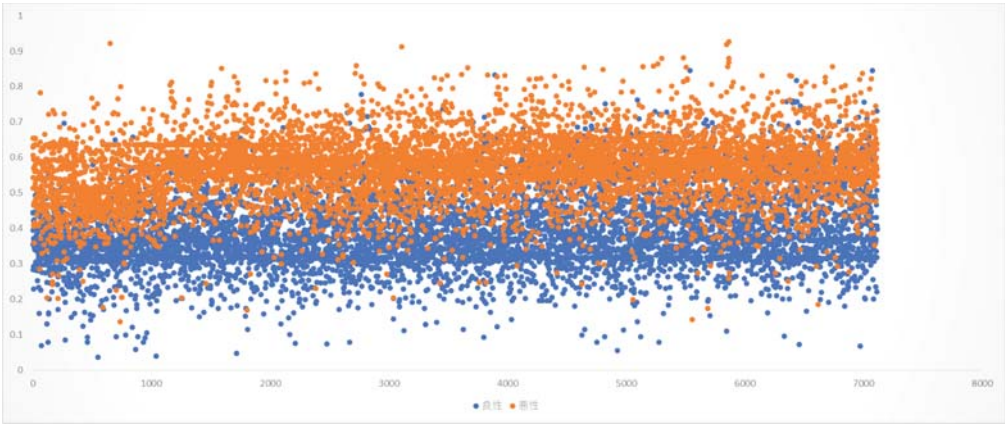

図 8　精査後の悪性度

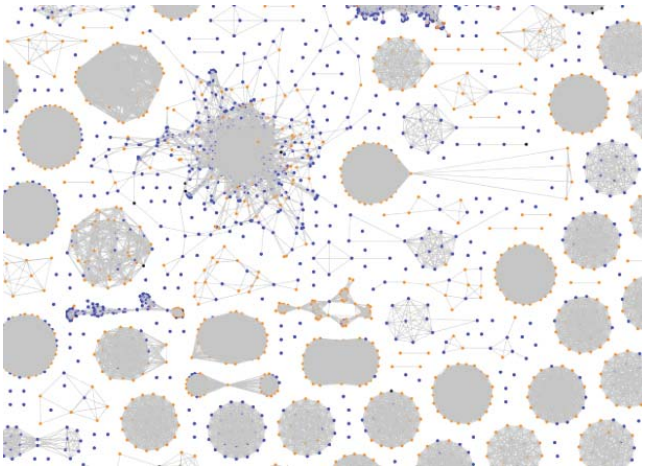

図 9　本実験でのドメイン間の結合

は良性ドメイン同士，悪性ドメインは悪性ドメイン同士で関連性があることが分かる．また，悪性ドメインと良性ドメインで結合している部分もみられる．悪性ドメインを多く結びつきがある良性ドメイン，またはその反対の例で誤検知・検知漏れが発生しやすいと考えられる．

悪性度の精査で，ドメインがどの判断したかの結果を表 2 に示す．特徴ベースで初期悪性度を求めだ段階では，閾値以下で良性と判断されていた悪性ドメインが，精査後悪性ドメインと検知することができている例が 10.6%確認できた．またその反対に，精査前悪性ドメインと判断されていた良性ドメインが，精査後に良性ドメインと判断されていた例も 18.0%確認できた．しかし，これらに下回る数ではあったが，初期悪性度の段階で良性と判断された良性ドメインが精査後に悪性ドメインとなったドメインが 12.0%，また初期悪性度の段階で悪性と判断された悪性ドメインが精査後に良性ドメインとなったドメインが 5.1%存在する．

本研究では，既知の悪性ドメインが悪性ドメインと検知された確率は 81.2%であり，既知の良性ドメインが良性ドメインだと判定される確率が 84.0%であった．

表 2　検知結果

	良性→悪性	悪性→良性
良性ドメイン	12.0%	18.0%
悪性ドメイン	10.6%	5.1%

6. 関連研究との比較

[5] では，DNS メッセージから多くの特徴を抽出し，ボットネット検知を回避する手法のクラスタ分析を高い精度で実現している．本手法では，高い検知率でなかった原因の一つとして初期悪性の計算方法が単純であり，精度が低

かったことが挙げられる．本実験では，使用した特徴は 5 つで，その特徴ごとに重みづけすることなく相乗平均とり，初期悪性度を求めた．本研究では，ドメイン間の関連度を考慮した悪性度の精査が，検知率に現れるかどうか確認するかを目的としているため，初期悪性度の計算は単純な方法をとった．先行研究などの精度の高い手法を用いることで，精度の特徴ベースでの検知手法をさらに向上することが期待できる．攻撃手法によって，初期悪性度の計算で用いる要素を変え，複数の悪性度から一番高い悪性度を採用することで，検知率は向上すると考えられる．また，より大きなデータセットを用いることで，さらに高い精度が期待できる．[5] では，指定したクラスタのみの分析を行う．本提案手法では，複数の特徴を用いて初期悪性度を求めるため，攻撃手法を限定せずに悪性ドメインを検知することができる．実験でも，フィッシングやボットネットなど複数の攻撃に利用される悪性ドメインを検知することができた．また，ドメイン間の関連性を考慮することで，新しい攻撃に利用される悪性ドメインも検知が可能だと考えられる．

[7][8][9] では，本提案手法の同じように，ドメイン間の関連性を考慮している．このような牽連性を利用することで，攻撃者の手でドメイン間の関連性を簡単に操作することは困難なため，攻撃者によって特徴ベースでの攻撃が回避された場合でも，検知を行うことが可能である．ドメイン名に解決される IP アドレスが，他のドメインにも紐づけられている繋がりを利用して，既知の悪性ドメインとの結合の度合いを求めている．本実験では，ドメイン間で共通するネームサーバの IP アドレスの数の多さを，結合度と定義した．実験から，これらの関連性が，実際に悪性ドメイン同士，または良性ドメイン同士を関連する要素となることが明らかとなった．実際の攻撃では連続して攻撃者が DNS 通信することがある．その場合，攻撃に使われるドメインのネームサーバが，攻撃者が管理するネームサーバに集中すれば，悪性ドメイン間の関連性も増すことが考えられる．さらに悪性ドメインの検知も精度も向上すると期待できる．一方で，本提案手法の検知率の向上に向けて，結合度の定義を検討する必要はある．

本実験では，既知の悪性ドメインまたは良性ドメインのリスト事前に準備する必要がない．そのため，先行研究で必要であった事前情報の信頼性の確認や情報の更新を行う手間を減らすことが可能である．

7. おわりに

悪性ドメイン間で関連性が強くなることを利用して，DNS グラフを解析し，ドメインごとに特徴ベースで計算した悪性度をドメイン間の関連性を考慮して精査することで，特徴ベースで計算した悪性度より精度が増すことを示した．これにより，事前情報を用意することなく，悪性ドメインを検知することができる．

この手法によって，複数の特徴を考慮して悪性度を計算するため，様々な攻撃に利用される悪性ドメインを検知することができる．また，悪性ドメインの特徴ベースでの検知に加え，ドメイン間の関連性を考慮することによって誤検知を減らし，事前情報を用意することなく効率的により多くの悪性ドメインの検知が期待できる．

本研究では，考慮した特徴数は 5 つと少なく，初期悪性度も計算方法も単純であった．TXT レコードの使用有無や DNS メッセージの長さなど特徴量を考慮し，初期悪性度の計算方法も，攻撃手法に応じたものに工夫すれば，検知率の向上も期待できる．

今後，ドメインの悪性度やドメイン間の結合度の決め方，関連するドメインの悪性度と結合度を用いた悪性度の精査方法について検討する必要がある．

参考文献

[1] Farnham, Greg, and A. Atlasis. ”Detecting DNS tunneling.” SANS Institute InfoSec Reading Room (2013): 1-32.
[2] 久山真宏, and 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.” マルチメディア, 分散協調とモバイルシンポジウム 2016 論文集 2016 (2016): 1711-1716.
[3] Karasaridis, Anestis, Kathleen Meier-Hellstern, and David Hoeflin. ”Nis04-2: Detection of dns anomalies using flow data analysis.” Global Telecommunications Conference, 2006. GLOBECOM’06. IEEE. IEEE, 2006.
[4] 千葉大紀, 森達哉, and 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 2012.3 (2012): 805-812.
[5] Lysenko, Sergii, et al. ”DNS-based anti-evasion technique for botnets detection.” Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), 2015 IEEE 8th International Conference on. Vol. 1. IEEE, 2015.
[6] 井原栄. ”DNS クエリを用いたネットワーク監視によるボットネット検知手法”, 千葉大学大学院融合科学研究科, 2015
[7] ZOU, Futai, et al. ”Detecting malware based on DNS graph mining”. International Journal of Distributed Sensor Networks, 2015.
[8] 風戸雄太, 福田健介, and 菅原俊治. ”DNS グラフ上でのグラフ分析と脅威スコア伝搬による悪性ドメイン特定.” コンピュータ ソフトウェア 33.3 (2016): 3_16-3_28.
[9] Khalil, Issa, Ting Yu, and Bei Guan. ”Discovering malicious domains through passive DNS data graph analysis.” Proceedings of the 11th ACM on Asia Conference on Computer and Communications Security. ACM, 2016.
[10] Majestic Million，`https://majestic.com` (2017.07.24).
[11] DNS-BH - Malware Domain Blocklist by RiskAnalytics，`http://www.malwaredomains.com` (2017.07.12).