分散型ネーミングサービスHandshakeに登録されたドッペルゲンガードメインの定量的調査

吉田 純一[1]　穐山 空道[1,a)]

概要：TLD が ICANN によって中央集権的に管理されていることを課題視し，ブロックチェーン技術を用いたドメイン管理システムである分散型ネーミングサービスが数々開発されている．特にルートゾーンを管理する分散型ネーミングサービスでは，ICANN による審査を通じた TLD の登録とは対照的に，ユーザが審査なしに自由に TLD を登録できる．しかし，その特徴から通常の DNS に登録されたドメインのドッペルゲンガードメインを登録し，偽の web サイトにアクセスさせるなどのフィッシング攻撃が可能になるという脅威がありえる．本研究では分散型ネーミングサービスである Handshake を対象に，通常ドメインのドッペルゲンガードメインが実際に多く登録されていることを示す．具体的には，約 1200 万個ある Handshake TLD の中から 2900 個のドッペルゲンガードメインを発見し，分析を行った．

キーワード：TLD，分散型ネーミングサービス，Handshake，ドッペルゲンガードメイン

Quantitative Analysis of Doppelgänger Domains Registered to Decentralized Naming Service Handshake

Junichi Yoshida[1]　Soramichi Akiyama[1,a)]

Abstract: Decentralized naming services based on blockchain technologies are actively being developed in order to alleviate the centralized power of ICANN for managing TLDs. Especially in a naming service that manages the root zone, a user can freely register a TLD without any screening from ICANN. However, a malicious user can also register a doppelgänger domain that can be used for a phishing campaign that lead people into fake websites. We analyze the TLDs register to such a naming service, Handshake, and show that many doppelgänger domains are indeed registered. In concrete, we found 2900 doppelgänger domains among approximately 12 million Handshake TLDs, and analyzed the possible threat level of them.

1. はじめに

通常のインターネットで利用されるドメインのうち，TLD (Top Level Domain) は ICANN (The Internet Corporation for Assigned Names and Numbers) によって中央集権的に管理されていると言える．これは新たな TLD の登録には ICANN による審査を経る必要があるためである．本論文では ICANN により審査，登録された TLD を **ICANN TLD** と呼ぶ．

一方で，Handshake [1] と呼ばれる分散型ネーミングサービスでは，ユーザは任意の TLD を審査なしで登録できる．Handshake はブロックチェーンによりルートゾーン情報を管理する分散型ネーミングサービスの一つで，ルートゾーンとルートネームサーバを置き換え TLD 登録の ICANN による中央集権性の解消を目的としている．ただし，Handshake TLD の登録制限として，ICANN TLD および Alexa List 上位 100，000 のドメインを TLD に変換したものが事前に予約されており，これらの TLD はユーザが登録できない．

Handshake には前述のような登録制限があるものの，依然としてドッペルゲンガードメインを登録しフィッシング攻撃へ利用する脅威が考えられる．Handshake では

[1] 立命館大学 情報理工学部
[a)] s-akym@fc.ritsumei.ac.jp

ICANN TLD とは対照的に審査なし TLD を登録できるため，悪意のあるユーザは ICANN TLD や ICANN TLD を含む上位レベルのドメインに酷似したドッペルゲンガードメインの登録が簡単に行える．本論文では ICANN TLD および ICANN TLD を含む上位レベルのドメイン群をまとめて**通常ドメイン**と呼ぶ．例えば `.com` や `.google.com` は通常ドメインである．Handshake が広く利用されるようになると，悪意のあるユーザは Handshake に登録したドッペルゲンガードメインを利用して偽のログインページなどに誘導するフィッシング攻撃を行える危険性があると考えられる．

本研究では Handshake に登録されたドッペルゲンガードメインの定量的調査および分析を行う．具体的なドッペルゲンガードメインの種類として，通常ドメインからドットを除いた文字列の suffix になっているドメイン（例：通常ドメイン `.google.com` に対し `.glecom`）を調査対象とする．以降では，Handshake に登録されたこのようなドメインを**疑悪性ドメイン**と呼ぶ．Handshake に登録された全ドメインを抽出し，その中に含まれる疑悪性ドメインの数や発見された疑悪性ドメインの分析を行う．

本研究の貢献は以下の 4 点である．

(1) Handshake 上に通常ドメインのドッペルゲンガードメインが登録されている可能性をはじめて指摘した．
(2) Handshake に登録された約 1200 万個の TLD を分析し，2900 個の疑悪性ドメインを発見した．
(3) 発見した疑悪性ドメインのうち，2429 個が英語としての意味を持たないことを特定した．これらの疑悪性ドメインは悪意を持って登録された可能性が特に高いと考えられる．
(4) 発見した疑悪性ドメインの登録時期を分析し，英語の意味を持たない疑悪性ドメインの約 9 割が 2020 年から 2022 年の間に登録されたことを発見した．

2. 研究背景

2.1 DNS と ICANN TLD

DNS はインターネット上のドメイン名の管理運用，ドメイン名の名前解決を行うシステムでである [2]．そのドメイン名のドットで区切られた内の最後の部分，例えば `exmple.com` ならば `.com` の部分は TLD と呼ばれる．この TLD の管理を行うネームサーバの参照情報の一覧が書かれているルートゾーンは世界中に 13 あるルートサーバで管理されている．この TLD とルートサーバの管理運用は，アメリカで 1998 年に設立された民間の非営利法人である ICANN が行っている．

TLD は ICANN によって中央集権的に管理されている．TLD には国ごとに割り当てられた country code TLD (ccTLD) の他に一般に新規募集が行われる generic TLD (gTLD) が存在するが，新しい gTLD の登録は ICANN による審査を経る必要がある [3]．具体的には，申請された文字列が募集要件を満たしており，その文字列に問題がないと ICANN に判断される必要がある．

2.2 分散型ネーミングサービス

分散型ネーミングサービスとは，ブロックチェーン技術を用いて開発された TLD やルートゾーン，ドメイン名の管理運用と名前解決を行うシステムである．このシステムは，TLD の管理が中央集権性を持つということを問題視し，その非中央集権化を目的として開発された．分散型ネーミングサービスは 2011 年から Namecoin を始めとして数多く開発されている [4]．

代表的な分散型ネーミングサービスを比較したものを表 1 に示す．分散型ネーミングサービスの管理するゾーンは二種類に分類され，特定の TLD の配下，つまりセカンドレベルドメインのゾーンを表す eTLD+1 とルートゾーンがある．前者の eTLD+1 を管理ゾーンとする分散型ネーミングサービスには，Namecoin， ENS， EmerDNS といったものがある．Namecoin は 2011 年に開発され，.bit 配下のドメイン名を管理している [5]．ENS は 2017 年に開発され，.eth 配下のドメイン名を管理している [6]．EmerDNS は 2013 年に開発され，.coin， .emc， .lib， .bazar の 4 個の TLD 配下のドメインのゾーンを管理している [7]．次に後者のルートゾーンを管理している分散型ネーミングサービスには，Handshake， Decentraweb がある [4]．これらは共に任意の TLD を無制限に登録でき，登録された TLD のルートゾーンを管理している．

本研究では，分散型ネーミングサービスの中から Handshake を調査対象に選ぶ．まずルートゾーンを管理する分散型ネーミングサービスを選択する理由は，本研究で扱う脅威の存在可能性がルートゾーンを管理するサービスに限られるためである．従って調査対象は Handshake または Decentraweb に限られる．次に Decentraweb ではなく Handshake を選択する理由は，後者の方が前者よりも TLD の登録数が圧倒的に多い [8] ためである．従って Handshake は調査する脅威が実際に存在する可能性や，将来広く利用され脅威が顕在化する可能性がより高い．

2.3 Handshake

Handshake は既存の DNS を置き換えることではなく，ルートゾーンとルートサーバを置き換えることを目的とした分散型ネーミングサービスである．このシステムではルートゾーンの管理，検証をブロックチェーン技術を用いて参加する全ピアが行う．Handshake のブロックチェーンには Bcoin のフォークが使用されている．Handshake のブロックチェーンネットワークに参加する際に必要なフルノード hsd や Handshake システムにアクセスする際に必要なライトノード hnsd は GitHub 上でオープンソースで

表 1: 各分散型ネーミングサービスの比較

サービス名	Namecoin	Emercoin	ENS	Handshake	Decentraweb
ブロックチェーン基盤	bitcoin	bitcoin	Ethereum	bcoin	Ethereum, Polygon
公開時期	2011 年	2013 年	2017 年	2018 年	2021 年
管理ゾーン	eTLD+1	eTLD+1	eTLD+1	ルートゾーン	ルートゾーン
TLD	.bit	.coin, .emc .lib, .bazar	.eth	無制限	無制限

表 2: ドメインから TLD への変換例

変換前	変換後
google.com	google
bbc.co.uk	bbc

公開されている [9, 10].

Handshake TLD の登録方法とそのコスト: Handshake での TLD の登録方法はオークション形式をとっている．具体的には，OPEN トランザクションによって特定の TLD のオークションが始まり，BID トランザクションを使用してそのオークションに入札する．そして，最も入札額が高かったユーザがその次に高かった入札額を払うことで TLD が登録される．つまり，入札するユーザが多ければ多いほど TLD を登録するコストが上昇していくことになり，逆にもし入札するユーザがいなければ無料で TLD を登録できる．登録した TLD は 2 年ごとに更新する必要があり，RENEW トランザクションを使用して更新する．更新には RENEW トランザクションのマイニング費用のみ発生するため，TLD の維持コストは低い [8, 11].

Handshake TLD の登録規則: Handshake TLD に使用できる文字列に制限は無い．しかし，既存の ICANN TLD および Alexa ドメインランキング上位 100,000 のドメインが表 2 のように TLD に変換され，事前に Handshake のブロックチェーン上に予約されている．予約された TLD はユーザが取得できない [12] ため，既存のドメインに類似するドメインの登録をある程度は防いでいる．

名前解決プロセス: Handshake は DNS のルートゾーンとルートサーバのみを置き換え，ルートゾーン以下のゾーンの管理には通常の DNS の仕組みを利用する．Handshake による名前解決プロセスの例を図 1 に示す．クライアントが web サイト `example.tld` にアクセスする時，まずフルサービスリゾルバと同じ役割を担うライトノードに `example.tld` の名前解決を依頼する．ライトノードはブロックチェーンのブロックを全て持つフルノードに `.tld` を管理する権威サーバの IP アドレスを問い合わせる．問い合わせを受けたフルノードはブロックチェーン上に保存されたルートゾーンを参照し，`.tld` の権威サーバの IP アドレスを応答する．次に，ライトノードは `.tld` の権威

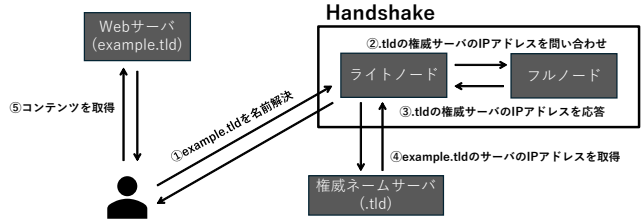

図 1: Handshake の名前解決プロセス

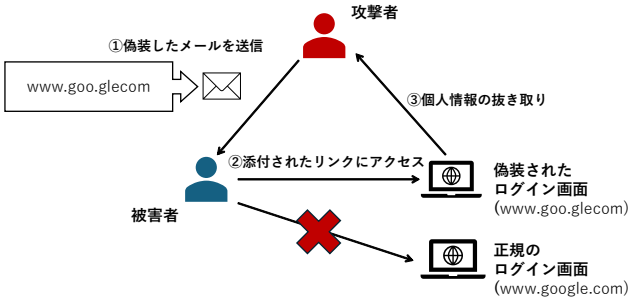

図 2: Handshake に登録されたドッペルゲンガードメインを使った攻撃の例

サーバから `example.tld` の A レコードを取得し，その情報をクライアントに応答する．最後に，クライアントは `example.tld` の web サーバからコンテンツを取得する．

3. 考えられる脅威

Handshake には通常ドメインのドッペルゲンガードメインを簡単に登録できる問題がある．ここで通常ドメインとは，ICANN TLD および ICANN TLD を含む上位レベルのドメイン群のことである．例えば `.com` や `google.com` は通常ドメインである．

考えられるドッペルゲンガードメインを利用した攻撃の流れを図 2 に示す．悪意のある攻撃者はまず Handshake に通常ドメイン `.google.com` のドッペルゲンガードメインである `.glecom` を登録し，正規の web サイト `www.google.com` を模倣したフィッシングサイトである `www.goo.glecom` を作成する．そして，Google 社の関係者を装い `www.goo.glecom` へのリンクを添付したメールを被攻撃者へ送信する．被攻撃者は偽のログイン画面に誘導され，入力したメールアドレスやパスワードなどが攻撃者

に奪われる．

この脅威は Handshake が中央集権性を持たないことに起因する．通常の TLD の登録には ICANN による審査が必要であり，.glecom という gTLD はこの審査を通過しないと予想される．実際に ICANN ではユーザの混乱を防ぐためのルール整備を行っており，例えば国際化ドメインの有効性を決定する仕組みである Root Zone Label Generation Rules (RZ-LGR) [13] では平仮名の「へ」と片仮名の「へ」は「適切に扱われるべき (*should be handled adequately*)」ペアであると定義されている．Handshake ではこのような中央集権性がない反面，ユーザに混乱をきたすような TLD も登録できてしまう．

4. 調査手法

4.1 調査の概要と目的

第 3 章で記した脅威の実在性を示すため，本研究では以下の 4 つの点について Handshake に登録されたドメインの定量的調査および分析を行う．

(1) Handshake 上の全 TLD を収集する．これは Handshake TLD の一覧が存在しないためである．
(2) 収集した Handshake TLD のうち，通常ドメインのドッペルゲンガードメインであるものを抽出する．抽出されたドメインを疑悪性ドメインと呼ぶ．疑悪性ドメインとは悪意を持って登録されたことが疑われるドメインという意味である．
(3) 抽出された疑悪性ドメインそれぞれが英語としての意味を持つかを調査する．これは各疑悪性ドメインの分析を行うためである．英語の意味を持たずかつ文字数が長い疑悪性ドメインは，悪意を持って登録された可能性がより高いと予想される．通常ドメインの suffix と疑悪性ドメインが偶然一致する確率は 5 文字の疑悪性ドメインの場合，約 6000 万分の 1 の確率と非常に低い確率となる．そのため，ここでは 5 文字以上の疑悪性ドメインは文字数が長いと判断する．
(4) 抽出された疑悪性ドメインの登録時期を調査する．これは疑悪性ドメインの登録がある時期に一斉に行われたのか，長期に亘り継続的に行われているのかを調べるためである．

具体的なドッペルゲンガードメインの種類として，本研究では「通常ドメインからドットを除いた文字列」の suffix になっている Handshake TLD を対象とする．例えば .om や .glecom という Handshake TLD は疑悪性ドメインである．前者は ICANN TLD である .com からドットを除いた文字列の suffix，後者は ICANN TLD とセカンドレベルドメインの組み合わせである .google.com からドットを除いた文字列の suffix である．


```
{
  "txid": "9eeer33..."
  “hash”: “11aabb…”,
   ・
   ・
  “vout”: [
    {
      “value”: 2000000,
      “address”: “・・・”
      “covenant”: {
        “type”: 6,
        “action”: “REGISTER”,
        “items”: [
          "eb13199d62...",
           ・
           ・
         ]
      }
    }
   ・
}
```


図 3: トランザクションデータの例

4.2 Handshake TLD の収集

Handshake TLD はブロックチェーン上に保存されているトランザクションデータを分析することで収集する．具体的には，フルノードを構築するためのソフトウェアである hsd を使用しブロックチェーンデータを同期後，提供されている API [11] を用いてブロック高 0 から特定のブロック高まで以下の手法を繰り返す．ブロック高とは，始めのブロックから何番目のブロックかを表す．なお本手法は Ito らの調査手法 [8] と同一である．

(1) getblockbyheight コマンドを使用し，指定したブロック高のブロック情報を取得する．
(2) 取得したブロック情報から図 3 のようなトランザクションを抽出する．
(3) 抽出したトランザクションから action が REGISTER であるもに着目し，その items から TLD のハッシュ値を取得する．これらは TLD の新規登録を表すトランザクションである．
(4) getnamebyhash コマンドを使用し，TLD のハッシュ値から TLD の文字列を取得する．

4.3 疑悪性ドメインの抽出

収集した全 Handshake TLD の中から疑悪性ドメインを抽出する．具体的には以下のような方法を行う．

(1) 調査対象となる通常ドメインのリストを作成する．これは ICANN TLD から英数字以外の文字列（例：日本語）で構成された国際化トップレベルドメインを除く約 1300 個のドメイン [14] と，Cloudflare 社の提供するドメインランキング [15] のうち TOP 1000 のドメインを結合して得る．Handshake で使われている

Alexa List は公開停止になっているため利用しない．
(2) 通常ドメインのリスト内の全てのドメインに対し，「ドメインからドットを除いた文字列」の全 suffix リストを作成する．例えば .com に対しては {com, om, m} というリストを作成する．
(3) 収集した Handshake TLD を逐次探索し，作成した suffix リスト内の要素と一致したものを疑悪性ドメインとして抽出する．

4.4 抽出した疑悪性ドメインの分析

抽出した疑悪性ドメインを英語の意味を持つかどうかで分類するために，自然言語処理で用いられる英語の語彙データベースである WordNet [16] を使用する．具体的には，以下のような方法で分析を行う．
(1) WordNet を使用するための Python ライブラリ nltk 内の関数である nltk.download() を実行し，その際に立ち上がる専用のダウンロード画面上で WordNet の辞書データをダウンロードする．
(2) ダウンロードされた辞書データが入った nltk-data ディレクトリ内の corpora ディレクトリの中には words ディレクトリがあり，その中の en という英単語の一覧が書かれたテキストファイルを抽出する．
(3) 各疑悪性ドメインに対し，同じ文字列がそのテキストファイル内にあるかどうかを探索する．

4.5 Handshake TLD の登録時期の取得

Handshake TLD の登録時期は，それぞれの TLD が登録されたブロック高のブロック情報を分析することで取得する．ここでの登録時期とは，TLD のオークションが始まったブロックの生成日時のことを指す．具体的に，提供されている API [11] を使用して Handshake TLD それぞれに対して以下のような方法を行う．
(1) getnameinfo コマンドを使用し，指定した TLD の情報を取得する．
(2) 取得した TLD の情報からオークションが始まったブロック高を抽出する．
(3) getblockbyheight コマンドを使用し，抽出したブロック高のブロック情報を取得する．
(4) 取得したブロック情報に含まれる UNIX 時刻を日付に変換する．

5. 調査結果

5.1 Handshake TLD の収集結果

Handshake に登録された全 TLD の収集では，調査を行った時点でブロック高 224,000 までのブロックが存在した．またそれらのブロックから TLD を抽出した結果，12,268,598 個の TLD が抽出できた．

なお，2023 年 10 月に投稿された Ito らの調査 [8] では Handshake TLD の登録数は 11,595,404 個だと報告されている．従って Handshake TLD は約半年間でおよそ 67 万個と約 6 %増加した．

5.2 疑悪性ドメインの抽出結果

約 1200 万個の Handshake TLD から 2900 個の疑悪性ドメインを抽出できた．Handshake TLD の個数と疑悪性ドメインの個数をまとめたものを表 3 に示す．

表 3: Handshake TLD と疑悪性ドメインの個数

Handshake TLD	12,268,598 個
疑悪性ドメイン	2900 個
疑悪性ドメイン割合	0.024 %

表より全 Handshake TLD のうち疑悪性ドメインの割合は 0.024 % であり，これは非常に多いとは言えない．一方で調査対象とした通常ドメインは約 2300 個（ICANN TLD が約 1300 個，Cloudflare ドメインランキングが 1000 個）であるため，平均して通常ドメイン 1 個に対し 1.26 個の疑悪性ドメインが存在している．

5.3 分析結果

抽出した 2900 個の疑悪性ドメインを英語の意味があるかどうかで分類すると，英語の意味を持つものが 471 個，持たないものが 2429 個であった．この結果から約 16 % の疑悪性ドメインは英語の意味を持つ，つまり悪意のないドメインがたまたま通常ドメインの suffix となってしまった可能性がある．一方で残る約 84 % についてはより詳細な分析が必要である．

そこで英語の意味を持たない疑悪性ドメインのみを対象に，(1) 文字数の分布の調査，(2) Cloudflare ドメインランキング TOP 100 に対応するものの目視での確認，を追加で行う．まず図 4 に文字数の分布を示す．図より 2 文字から 5 文字の疑悪性ドメインが他の文字数と比べて多いことが分かる．次に表 4 に，ドメインランキング TOP 100 に対応かつ文字数が 5 文字以上の疑悪性ドメインを全て示す．TOP 100 ドメインは左上から右下に順位が下がるよう配置されており，一行目の google.com はランキング 1 位，root-servers.net はランキング 2 位である．また表中の N/A は対応する疑悪性ドメインがなかったことを示す．表を見ると googlecom，root-servesnet，amazonawscom，instagramcom のように，非常に文字数が多くかつ上位の通常ドメインとドットを除き完全一致する疑悪性ドメインが存在する．これらのドメインは悪意を持って登録された可能性が特に高いと予想される．なお，これらの疑悪性ドメインは Handshake による登録制限のために予約されたものではない．第 2.3 節で示したように，登録制限のための予約では google.com は google に変換されるため

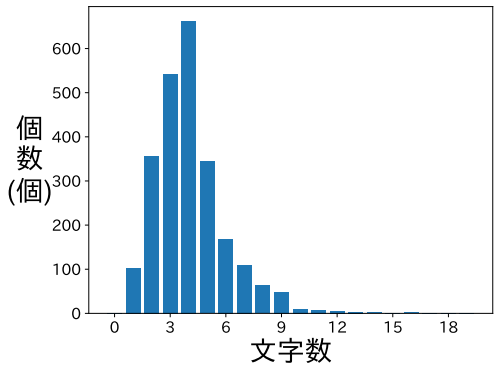

図 4: 英語の意味を持たない疑悪性ドメインの文字数分布

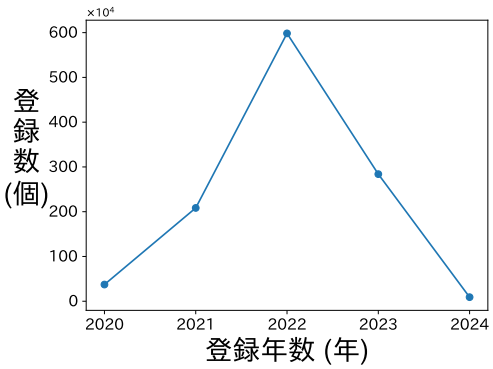

図 5: Handshake TLD の登録時期

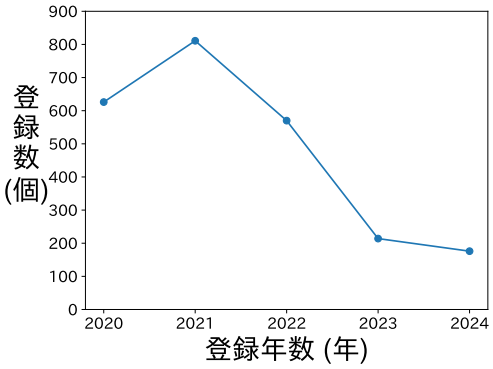

図 6: 英語の意味を持たない疑悪性ドメインの登録時期

googlecom は予約されていない．

5.4 登録時期の取得結果

図 5 と図 6 に，全 Handshake TLD および英語の意味を持たない疑悪性ドメインの年ごとの登録件数に示す．図 5 より，2022 年に Handshake TLD が約 600 万個も登録されたことが確認できる．これは調査時点の Handshake TLD の登録数の 5 割を占めている．また図 6 より，英語の意味を持たない疑悪性ドメインは 2020 年から 2022 年の間に全体の約 8 割が登録されたことが確認できる．逆に 2023 年からは登録数が減少し，それまでの年間 600 件以上から年間 200 件程度になっている．

登録時期の調査結果より以下の二点が分かる．

(1) 全 Handshake TLD の登録数と疑悪性ドメインの登録数は傾向が異なる．前者は 2022 年にピークだが，後者はそれより早い 2021 年がピークである．

(2) 疑悪性ドメインは短期間で全てが登録されておらず，長期間に亘って登録され続けている．2023 年，2024 年は登録数が減っているものの，依然として数百件の疑悪性ドメインが登録されている．

6. 関連研究

6.1 通常の DNS 上の悪性ドメインの研究

Garrett ら [17] はドッペルゲンガードメインを用いたメールでの攻撃の影響度を，Fortune 500 の各企業を対象に調査し，151 社がドッペルゲンガードメインの影響を受けやすいことを発見した．また，ドッペルゲンガードメインを用いた攻撃の対策として二つの対策を提案した．一つ目は事前にドッペルゲンガードメインを登録して攻撃者の利用を防ぐこと，二つ目はドッペルゲンガードメインを登録した攻撃者を特定し統一ドメイン名紛争処理方針に基づく申し立てを行い利用を停止させることである．本研究で提唱した脅威の対策には，一つ目は有効であると考えられ，実際に今回発見した疑悪性ドメインが通常ドメインの所有者により登録された可能性もある．しかし二つ目の対策は，一度登録されたデータは消去できないというブロックチェーンの特性上有効でないと考えられる．

Simpson ら [18] は中堅および大企業の持つ .com ドメインの約 7% に対し一見しただけでは区別がつかないように正規のドメイン名内の文字列を変更したドメインが登録されていることを発見した．このようなドメインはビジネスメールを用いた詐欺に利用するのに適していると指摘し，10 年間の VIDN の登録時期，割合，動きの三点を分析した．その結果，特に 2015，2016 年に VIDN が大量に登録されており，加えてその 2 年間で VIDN を使用した攻撃を受けた企業の数も極めて多いことを特定した．

6.2 分散型ネーミングサービス上のドメインの研究

Isobe ら [19] は，ICANN TLD から辿れるセカンドレベルドメインと Handshake TLD の間の類似性を定量調査した．具体的には Punycode を用いる Handshake TLD のうち，一文字変更すると ICANN TLD から辿れるセカンドレベルドメインと一致するものを対象とした．本稿の疑悪性ドメインは Punycode を含まず，Isobe らの調査対象とは相補的である．なお本稿は [19] の出版前に投稿された．

Ito ら [8] はルートゾーンを管理する二つの分散型ネーミングサービス，Handshake と Decentraweb の間の TLD の衝突および ICANN TLD との衝突を分析した．その結果，Handshake TLD と Decentraweb TLD の間には 6,975 個の名前衝突があること，ICANN TLD との間には Handshake では 10 個，Decentraweb では 2 個の名前衝突があ

表 4: 英語の意味を持たないかつ 5 文字以上の疑悪性ドメインの一覧

通常ドメイン	対応する疑悪性ドメイン	通常ドメイン	対応する疑悪性ドメイン
google.com	googlecom, lecom	root-servers.net	ersnet, root-serversnet, rsnet, serversnet
googleapis.com	iscom	apple.com	applecom, lecom
gstatic.com	iccom	facebook.com	bookcom, facebookcom, okcom
tiktokcdn.com	dncom	microsoft.com	ftcom, microsoftcom, softcom
amazonaws.com	amazonawscom, wscom	googlevideo.com	videocom
cloudflare.com	cloudflarecom, recom	fbcdn.net	cdnnet
doubleclick.net	cknet, clicknet	youtube.com	becom, tubecom, youtubecom
icloud.com	cloudcom, icloudcom, udcom	apple-dns.net	dnsnet
amazon.com	amazoncom	instagram.com	amcom, gramcom, instagramcom, ramcom
whatsapp.net	appnet	akadns.net	dnsnet
googleusercontent.com	N/A	akamai.net	ainet, mainet
ntp.org	N/A	tiktokv.com	kvcom
googlesyndication.com	googlesyndicationcom, oncom	gvt2.com	N/A
cloudfront.net	frontnet, ntnet	akamaiedge.net	edgenet
cloudflare-dns.com	dnscom, nscom	cdn77.org	77org
netflix.com	ixcom, netflixcom	ytimg.com	mgcom
aaplimg.com	mgcom	cdninstagram.com	amcom, gramcom, instagramcom, ramcom
live.com	livecom, vecom	bytefcdn-oversea.com	eacom, seacom
bing.com	bingcom, ingcom, ngcom	spotify.com	fycom, spotifycom
app-analytics-services.com	cescom, escom	gvt1.com	t1com
snapchat.com	atcom, chatcom	google-analytics.com	cscom, icscom
one.one	neone, oneone	yahoo.com	hoocom, oocom, yahoocom
googleadservices.com	cescom, escom, googleadservicescom	unity3d.com	3dcom
twitter.com	ercom, twittercom	digicert.com	rtcom
app-measurement.com	N/A	fastly.net	lynet
dns.google	dnsgoogle, oogle	ttlivecdn.com	dncom
ui.com	uicom	applovin.com	incom, vincom
office.com	cecom, officecom	amazon-adsystem.com	emcom
msftncsi.com	csicom, sicom	roblox.com	oxcom, robloxcom
ggpht.com	htcom	samsung.com	ngcom, samsungcom
googletagmanager.com	ercom, gercom	baidu.com	baiducom, ducom
rocket-cdn.com	dncom	trafficmanager.net	gernet, managernet
azure.com	recom	wikipedia.org	diaorg, wikipediaorg
appsflyersdk.com	dkcom	lencr.org	N/A
bytefcdn-ttpeu.com	eucom	xiaomi.com	micom
rbxcdn.com	dncom	skype.com	skypecom
qq.com	qqcom	msn.com	msncom, sncom
sentry.io	N/A	tiktokcdn-us.com	uscom
android.com	androidcom, idcom	criteo.com	N/A
linkedin.com	incom, linkedincom	gmail.com	gmailcom, ilcom, mailcom
windows.net	windowsnet	3gppnetwork.org	rkorg
taboola.com	lacom, olacom	microsoftonline.com	inecom, linecom, necom, onlinecom
qlivecdn.com	dncom	windows.com	wscom
cdn-apple.com	applecom, lecom	doubleeverify.com	fycom, verifycom
miui.com	uicom	office365.com	365com, 65com
2mdn.net	N/A	windowsupdate.com	tecom
crashlytics.com	cscom, icscom	mzstatic.com	iccom
pangle.io	N/A	vungle.com	lecom
taobao.com	aocom, taobaocom	ampproject.org	ctorg
mtgglobals.com	alscom, lscom	casalemedia.com	diacom, iacom, mediacom

ることを発見した．Ito らの研究では TLD の完全一致のみを対象としているが，本研究では通常ドメインに類似のドメインを調査した点が異なる．

Babakian ら [20] は異なるネーミングサービス間の一貫性を取るため，各サービスに一意な suffix の付与が有効だと述べている．一方で統一的な suffix の合意方法という課題も指摘している．Babakian らの述べるような suffix が導入され，例えば Handshake 上に登録された `glecom` がユーザから見れば `glecom.hnd` と表現されるようになれば，本研究で指摘した脅威はある程度緩和される．しかし `www.google.com` と `www.goo.glecom.hnd` は依然として紛らわしく，完全な解決にはならない．

7. 今後の課題

今後の課題の一点目は，同様の調査をより広範に行うことである．本研究ではドッペルゲンガードメインとして通常ドメインの suffix になっているもののみに着目した．ドッペルゲンガードメインには他にも例えばホモグリフを悪用したものがあり，その存在の調査も必要である．ホモグリフとはアルファベットの `l` と数字の `1` のような見た目が紛らわしい文字ペアのことである．また，Handshake と同じくルートゾーンを管理する Decentraweb においても本研究で提唱した脅威があり得るため，Decentraweb も対象に調査を行う必要がある．

今後の課題の二点目は，疑悪性ドメインの登録意図をより詳細に分析，理解し疑悪性ドメインによる脅威の防止に資することである．本研究で抽出された疑悪性ドメインはフィッシング攻撃に利用するために登録された可能性以外にも以下の可能性があり，より詳細な分析が必要である．

(1) Handshake が普及した際に通常ドメインの所有者にそれと酷似した疑悪性ドメインを売却するために登録された可能性．
(2) フィッシング攻撃などの脅威を防ぐため，通常ドメインの所有者（例えば Google 社）が事前に自社の所有する通常ドメインに類似したドメインを Handshake に登録した可能性．

8. 結論

TLD 管理の非中央集権化を目的とするブロックチェーン技術を活用した分散型ネーミングサービスが近年数多く開発されている．特にルートゾーンを管理する分散型ネーミングサービスでは TLD をユーザが自由に登録できるメリットがあるが，それに伴いセキュリティリスクも生じる．本研究では，Handshake を対象にドッペルゲンガードメインを登録できてしまう問題を挙げ，疑悪性ドメインについて調査および分析を行った．その結果，約 1200 万個ある Handshake TLD の中に疑悪性ドメインが 2900 個登録されていること，既存の人気ドメインとドットを除いて完全一致する TLD も登録されていることが分かった．

参考文献

[1] Handshake: Handshake Documentaion, 入手先〈`https://hsd-dev.org/`〉. (参照 2024-08-01).
[2] JPNIC: DNS とは，入手先〈`https://www.nic.ad.jp/ja/basics/beginners/dns.html`〉. (参照 2024-08-01).
[3] JPNIC: 新 gTLD の募集, 入手先〈`https://www.nic.ad.jp/ja/dom/new-gtld/recruitment.html`〉. (参照 2024-07-29).
[4] 株式会社日本総合研究所先端技術ラボ: ブロックチェーンを用いた名前解決 分散型ネームサービスの概要 , 入手先〈`https://www.jri.co.jp/MediaLibrary/file/advanced/advanced-technology/pdf/14872.pdf`〉. (参照 2024-08-13).
[5] namecoin: Namecoin, 入手先〈`https://www.namecoin.org`〉. (参照 2024-08-16).
[6] ENS: Welcome to the New Internet, 入手先〈`https://ens.domains/`〉. (参照 2024-08-16).
[7] EMERCOIN: EmerDNS, 入手先〈`https://emercoin.com/en/emerdns/`〉. (参照 2024-08-16).
[8] Ito, D., Takata, Y., Kumagai, H. and Kamizono, M.: Investigations of Top-Level Domain Name Collisions in Blockchain Naming Services, *Proceedings of the ACM on Web Conference (WWW)*, p. 2926–2935 (2024).
[9] handshake-org: hsd, 入手先〈`https://github.com/handshake-org/hsd`〉. (参照 2024-07-31).
[10] handshake-org: hnsd, 入手先〈`https://github.com/handshake-org/hnsd`〉. (参照 2024-07-31).
[11] Handshake: API Docs, 入手先〈`https://hsd-dev.org/api-docs/`〉. (参照 2024-07-31).
[12] handshake-org: handshake-names, 入手先〈`https://github.com/handshake-org/hs-names`〉. (参照 2024-08-06).
[13] JGP (Japanese Generation Panel): Proposal for a Japanese Script Root Zone LG, 入手先〈`https://www.icann.org/en/system/files/files/proposal-japanese-lgr-20dec21-en.pdf`〉. (参照 2024-10-20).
[14] ICANN: List of Top-Level Domains, 入手先〈`https://www.icann.org/resources/pages/tlds-2012-02-25-en`〉. (参照 2024-08-19).
[15] Cloudflare Radar: Domain Rankings, 入手先〈`https://radar.cloudflare.com/domains`〉. (参照 2024-08-19).
[16] Princeton University: WordNet, 入手先〈`https://wordnet.princeton.edu/`〉. (参照 2024-08-19).
[17] Gee, G. and Kim, P.: Doppelganger Domains, 入手先〈`https://godaigroup.net/wp-content/uploads/doppelganger/Doppelganger.Domains.pdf`〉. (参照 2024-08-12).
[18] Simpson, G., Moore, T. and Clayton, R.: Ten years of attacks on companies using visual impersonation of domain names, *APWG Symposium on Electronic Crime Research (eCrime)*, pp. 1–12 (2020).
[19] Isobe, K., Eisenbarth, J.-P., Kondo, D., Cholez, T. and Tode, H.: A Deeper Grasp of Handshake: A Thorough Analysis of Blockchain-based DNS Records, *Conference on Blockchain Research Applications for Innovative Networks and Services (BRAINS)*, pp. 1–10 (2024).
[20] Babakian, A., Huston, G., Braun, R. and Lipman, J.: Internet Identifiers: A Survey of History, Challenges, and Future Perspectives, *IEEE Access*, Vol. 12, pp. 51919–51941 (2024).