

修士論文

トンネリング抑止を目的とした 分散ハッシュテーブルを利用したDNSに関する研究

高須賀 昌烈

2020 年 1 月 28 日

奈良先端科学技術大学院大学
先端科学技術研究科

本論文は奈良先端科学技術大学院大学先端科学技術研究科に
修士(工学) 授与の要件として提出した修士論文である。

高須賀 昌烈

審査委員：

門林 雄基 教授 (主指導教員)

笠原 正治 教授 (副指導教員)

林 優一 教授 (副指導教員)

妙中 雄三 准教授 (副指導教員)

トンネリング抑止を目的とした 分散ハッシュテーブルを利用した DNS に関する研究*

高須賀 昌烈

内容梗概

キーワード

ドメインネームシステム, DNS トンネリング, 分散ハッシュテーブル, P2P ネットワーク

*奈良先端科学技術大学院大学 先端科学技術研究科 修士論文, 2020 年 1 月 28 日.

A Study of Domain Name System using Distributed Hash Table for Tunneling Deterrence*

Shoretsu Takasuka

Abstract

Keywords:

Domain Name System(DNS), DNS Tunneling, Distributed Hash Table, P2P Network

*Master's Thesis, Graduate School of Information Science, Nara Institute of Science and Technology, January 28, 2020.

目次

1. 序論	1
1.1 背景	1
1.2 目的	2
1.3 本論構成	3
2. 脅威モデル	4
2.1 DNS の概要	4
2.1.1 名前空間	4
2.1.2 名前解決の仕組み	6
2.1.3 ドメイン名とリソースレコード	7
2.2 DNS トンネリング	10
2.2.1 スタブリゾルバから権威サーバ方向	11
2.2.2 権威サーバからスタブリゾルバ方向	13
2.3 検知に基づく既存対策手法	15
2.3.1 特徴量	15
2.3.2 閾値推定	17
2.3.3 機械学習に基づくモデル	17
2.3.4 検知迂回手法	17
3. 関連研究	18
3.1 P2P ネットワークを利用した名前解決システム	18
3.1.1 GNS	18
3.1.2 SocialNS	18
3.2 Blockchain を利用した名前解決システム	18
3.2.1 Namecoin	18
3.3 課題	18
4. 提案システム	19
4.1 トンネリング抑止指向の DNS : SORES	19

4.2 システム設計	21
4.2.1 P2P を組み合わせたハイブリッドなアーキテクチャ	21
4.2.2 識別子と名前空間	23
4.2.3 ソートされたハッシュ空間の範囲に基づいたゾーン	25
4.2.4 認証基盤に基づく信頼されるレコード情報	27
4.2.5 リソースレコードタイプの選別による DNS Infiltration 対策	28
4.2.6 コンテンツのデータフォーマット	29
4.3 動作メカニズム	29
4.3.1 レコード情報に対する操作	29
4.3.2 名前解決	29
5. 評価	30
5.1 実装	30
5.2 DNS トンネリング	32
5.2.1 DNS Exfiltration	32
5.2.2 DNS Infiltration	32
5.3 要素ごとの特性	32
5.3.1 RTT(Round Trip Time)	34
5.3.2 パケットサイズ	34
5.3.3 トラフィック量	34
6. 考察	35
6.1 マネージャノードの最適な数	35
6.2 コンテンツ ID を計算する最適ノード	35
7. 結論	36
謝辞	37
参考文献	38
付録	42

図 目 次

1	ドメインにおけるゾーンごとの名前空間	5
2	DNS による名前解決	6
3	DNS のパケット構造	7
4	DNS のパケットの Answer セクション	8
5	DNS Exfiltration メソッドに基づいて、ドメイン名のラベル部に任意文字列 (“arbitrary-string”) が注入された DNS クエリが、イントラネット内のスタブリゾルバからインターネット上の権威サーバ (“exfil.com”) に転送される様子	11
6	TXT レコードに登録された情報について、DNS クエリで問い合わせることで権威サーバから命令情報を取得している様子	13
7	SORES の概略図	19
8	マネージャとプロバイダの関係図	21
9	レコード情報に対する認証プロセス	27
10	コンテンツファイルのデータフォーマット	29

表 目 次

1	主要リソースレコード一覧	9
2	DNS Infiltration のメソッドとして使用することができるレコードタイプの一覧	14
3	正規 DNS クエリと DNS トンネリングにおける DNS クエリのドメイン名の長さに関する違い	16
4	2017 年 7 月から 8 月までの DNS トラフィックデータセットにおけるリソースレコードのタイプ分布	16
5	SORES における用語	24
6	6 つのマネージャによって管理されるハッシュテーブルにおいて、マネージャの情報とそのマネージャが管理するゾーンが記載された対応表の例	26

7	DNS と SORES の特性比較	33
---	-----------------------------	----

アルゴリズム目次

1	スタブリゾルバからのクエリ処理	31
---	---------------------------	----

1. 序論

1.1 背景

増加のし続けているサイバー攻撃に対して、現在多くの組織は、SIEMのようなネットワークトラフィックを監視するシステムを導入し、そのシステムから発せられるアラートを処理することで脅威に対処している。他方で、機密情報の流出や諜報活動といった目的を実行する者は、目的を成功するため、またはその攻撃を特定・解析されないために、目的達成のために必要な通信を他の通信プロトコルに偽装させるといった秘匿通信手法 [7] を用いることが知られている [1]。そのような目的の実行者は、秘匿通信手法を用いることで、先の監視システムを巧みに迂回しながら、秘密裏に目的を実行することができる。このような秘匿通信手法を用いた攻撃に対処するには、従来のような通信の監視だけではなく、プロトコルの設計において悪用されないようなセキュア指向に基づき設計することが重要である。

しかし、現在のインターネットの根幹に位置づくネットワークプロトコルの中には、当初の設計のまま変更されることなく現在まで使用されているプロトコルがある。DNS(Domain Name System) プロトコルは、ドメイン名に関連づけられた情報を解決するネットワークプロトコルであり、現在インターネットを利用する際に極めて重要な位置にあるプロトコルの一つである。DNS は、現在のインターネットにおける位置づけから多くの組織でフィルタリングされることが少なく、任意の文字列を注入できるという特性を持っているため、秘匿通信手法として都合が良い。事実、DNS トンネリングと呼ばれる秘匿通信手法は、もっとも代表的な秘匿手法である。DNS トンネリングは、DNS クエリを利用することで外部情報に情報を流出させるだけでなく、事前にサーバに情報を登録しておくことで外部から内部方向に情報を転送できるという性質を持つ。このため、DNS トンネリングは、ターゲット組織から取得したデータを外部に流出させる際の手段としてだけでなく、ターゲットネットワークに潜伏しているマルウェアに対する

SIEM : Security Information and Event Management

秘匿通信 : 情報を不正・秘密裏に転送するために使用される回避通信手法

C2 サーバからの命令を送る手段として、実際のインシデントで広く利用されている [9–16].

これまで、DNS トンネリングに対する従来の対策としては、トラフィック特徴およびペイロード特徴に基づいた検知によるアプローチが取られてきた。特にデータマイニングに基づき、閾値を推定する方法や悪性通信モデルを生成する方法は多数考案されている [17, 18, 20–23]. しかし、提案されてきた検知手法では、その多くが検知対象の DNS トンネリングの通信を、github などより一般に入手が可能なツールキット [24–26] によって用意されるトラフィックをサンプルとしている。このようなツールキットによって生成されるトラフィックには、高いエントロピーや長いドメイン名といった顕著な特徴を示す特性があり、検知システムを迂回することを目的とするような実際の攻撃シーンにおける DNS トンネリングとは異なる。実際の攻撃において、使用されるような高い秘匿性を持つ DNS トンネリング手法には、1 回あたりの転送データ量を少なくする手法やパケット間のインターバルを長期間にする手法などがある [21]. このようなツールキットではサンプリングできないより秘匿性を高めた手法に対して、従来の検知手法では検知することが困難である。

検知以外のアプローチには、DNS のアーキテクチャを変更することによる抑止手法が考えられる。これまでに多数の次世代名前解決システムは提案されてきているが、DNS トンネリング抑止を目的としたシステムは筆者が知りうる限り確認されない。

1.2 目的

本研究は、DNS の名前解決の仕組みを悪用した秘匿通信手法の DNS トンネリングを抑止しながら、従来通りドメイン名に関連づけられたリソースレコード情報を解決できる機能を両立する名前解決システムを開発することを目的とする。

既存の DNS による名前解決システムは、現在のインターネットの根幹技術であるため、その名前解決エコシステムに大幅な変更を加えることは、高い導入コ

C2 サーバ : Command & Control サーバ

ストが要求されるため望まれない。例えば、期待されないシステムとしては、以下のようなものが予想される。

1. DNS トンネリングの抑止の機能は実現できるが、リソースレコード情報の解決の機能がない
2. DNS トンネリングの抑止の機能は実現できるが、既存システムからの大幅な変更が必要になり、未対応のコンピュータのインターネット接続に支障をきたす

以上のことを踏まえて、既存システムのエコシステムとの互換性を保ちながら、目的を達成することが重要である。

提案システムは、既存システムにおけるクライアント・サーバアーキテクチャを流用しながら、既存の再帰問い合わせの仕組みのみに改変処理を施すことで、DNS トンネリングを抑止する名前解決システムの実現を目指す。すなわち、提案システムは、スタブリゾルバとフルサービスリゾルバ間における処理はそのままに、フルサービスリゾルバと権威サーバ間における処理にのみ変更を加えることで、DNS のエコシステムとの互換性を保ちながら、DNS トンネリングを抑止することを実現する。このような提案システムは、既存クライアントに変更を加えずに秘匿通信としての機能を抑止するため、セキュアな名前解決システムとして広く一般に利用されることが期待される。

1.3 本論構成

本稿の構成は次の通りである。第2章では、本論で対象とする脅威モデルであるDNS トンネリングを説明し、これまでの検知に基づくアプローチの限界を示す。第3章では、これまでに提案されてきた名前解決システムについて説明し、DNS トンネリングに対する課題を示す。第4章では、提案手法を説明する。第5章では、提案手法のDNS Tunneling に対する評価を行い、併せて提案手法の特定についても説明する。第6章では、提案手法の課題について議論する。最後に、第7章にて結論を述べる。

2. 脅威モデル

本章では、はじめに DNS の概要について述べる。

次に、本研究における脅威モデルである DNS トンネリングについて説明する。

2.1 DNS の概要

DNS は、IP アドレスをはじめとしたドメイン名に関連づけられたリソースレコードを解決するシステムである [2, 3]。DNS がユーザから問い合わせられたドメイン名の IP アドレスを解決してくれるおかげで、ユーザは識別しづらい IP アドレス (IPv4: “93.184.216.34”, IPv6: “2606:2800:220:1:248:1893:25c8:1946”) を直接入力することなく、サーバにアクセスすることができる。このような利便性を実現する DNS による名前解決の機能は、ユーザがインターネットを利活用する上で極めて重要である。

2.1.1 名前空間

DNS において、各種リソースレコードが関連づけられるドメイン名は、ルートを頂点とする逆ツリー構造で構成されている。ドメインの序列を表記する際には、上位ドメインを右に、下位ドメインを左にする。“example.com.”を例にとると、ドットで表記されるルートが最も右に位置し、ルートの一つ下層に位置する TLD(Top Level Domain) である “com” が後続する。ドメインの区切りには、ドットが使用され、TLD の一つ下層に位置づく SLD(Second Level Domain) として “example” が連結していることを “example.com.” は表しているという具合である。図 1 に示す通り、全てのドメインには、ゾーンと呼ばれる名前空間があり、このゾーンはそれぞれのドメインに配置される権威サーバによって管理される。DNS には、委譲という仕組みを利用することでゾーンを分割できる特徴がある。例えば、“example” と “google” というドメインが連結している “com.” について考える。委譲の仕組みを使用しなかった場合には、“com.” が “www.example.com.” や “mail.google.com.” といった “example.com.” と “google.com.” に連結した全てのド

メイン名を管理する。しかし，“com.”がそれぞれのドメインにゾーンを委譲した場合，先の“www.example.com”は“example.com.”が管理し，“mail.google.com.”は“google.com.”が管理することになる。DNS では，この委譲の仕組みを利用することで，ゾーンの肥大を抑止させ，サーバの負荷分散を実現する設計になっている。

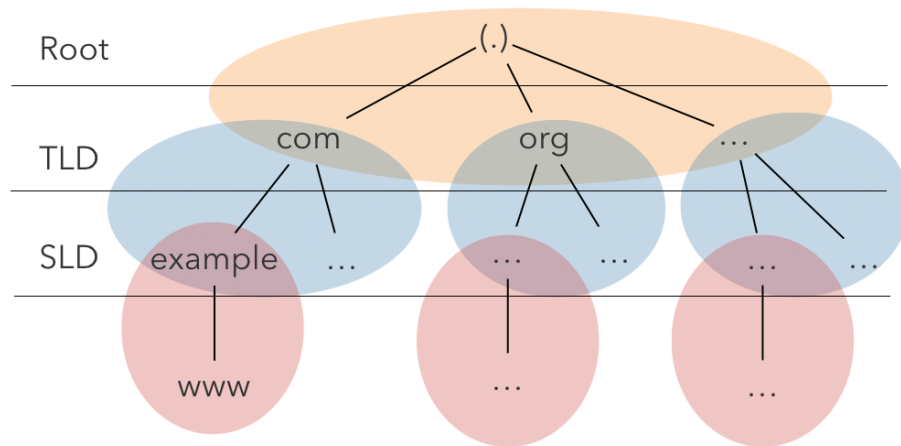


図 1 ドメインにおけるゾーンごとの名前空間

アーキテクチャ

DNS は，クライアント・サーバアーキテクチャで構成され，機能に基づき 3 つのサービスに分類することができる。

- スタブリゾルバ
- フルサービスリゾルバ
- 権威サーバ

スタブリゾルバは，名前解決の問い合わせを依頼するクライアントノードである。フルサービスリゾルバ(キャッシュサーバ・リカーシブサーバとも呼称される)は，スタブリゾルバに代わって，リソースレコードを保持する権威サーバに問い合わせるクライアントノードである。名前解決をする際には，ルートから順に TLD，SLD という具合に権威サーバに再帰的に問い合わせることで，最終的に目的のドメイン名に関するリソースレコード情報を取得する。この時，はじ

めのルート権威サーバのアドレスは “root.hints” と呼ばれるファイルに基づいて問い合わせるが、それより下位のドメインについては、上位の権威サーバが次の権威サーバのアドレスを応答することで名前解決のチェーンを繋げている。すなわち、ルート権威サーバが TLD の権威サーバのアドレスを応答し、TLD の権威サーバが SLD の権威サーバのアドレスを応答していく具合である。権威サーバは、リソースレコードを保持するサーバノードであり、フルサービスリゾルバからの問い合わせ依頼に応答する。

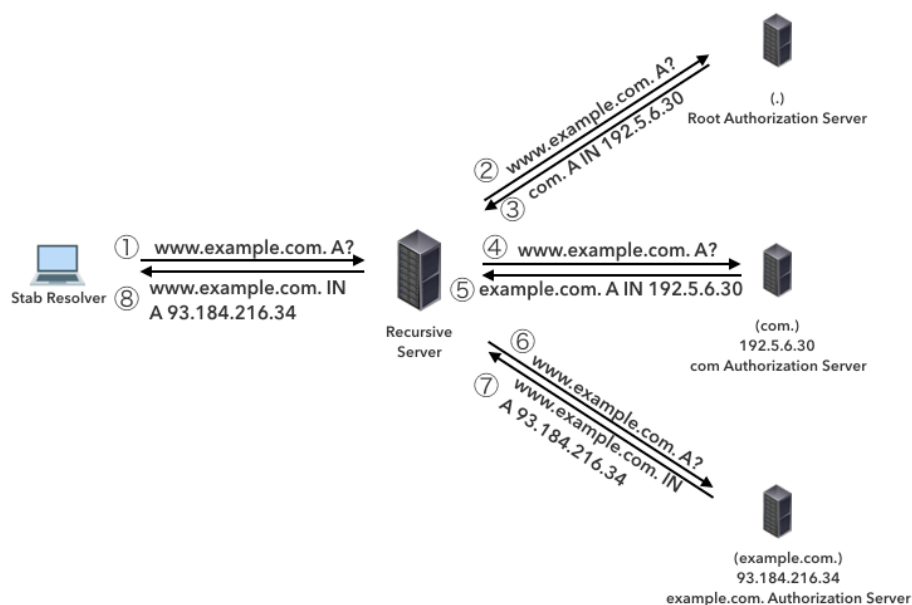


図 2 DNS による名前解決

2.1.2 名前解決の仕組み

いま、クライアントから “www.example.com” の IPv4 アドレスについて問い合わせられた場合を考える。クライアントとなるスタブリゾルバは、スタブリゾルバと同一セグメント内のフルサービスリゾルバもしくは、ネットワークセグメントに依らないパブリックなフルサービスリゾルバ (オープンリゾルバ、パブリックリゾルバとも呼称される) に問い合わせる。フルサービスリゾルバは、その名前解決クエリが過去に解決したものでないかキャッシュデータを確認する。キャッシュに

ヒットした場合にはキャッシュの情報をクライアントに応答され、ヒットしなかった場合には、root.hints ファイルを参照しルート権威サーバにリクエストパケットを転送する。クエリ (問い合わせ) を受け取ったルート権威サーバは、“com” ドメインを委譲した権威サーバのアドレスを応答する。次に、フルサービスリゾルバは、“com” の権威サーバに対し同じクエリを転送する。“com” の権威サーバは、“example.com” ドメインを委譲した権威サーバのアドレスを応答する。フルリゾルバは、“example.com” の権威サーバに同じクエリを転送する。“example.com” の権威サーバは、保持するゾーンファイルからクエリされたドメインのリソースレコードについて探索し、探索の結果としてレコード情報をフルサービスリゾルバに応答する。フルサービスリゾルバは、権威サーバから応答された情報をスタブリゾルバに転送することで、問い合わせられた名前は解決される。DNS による名前解決の一連の動作を図 2 で示す。

2.1.3 ドメイン名とリソースレコード

Header
Question
Answer
Authority
Additional

図 3 DNS のパケット構造

ドメイン名は、ルートから伸びる逆ツリー構造をドメインを表すラベルをドットで区切り表記する。すなわち、ドメイン名は “label3.label2.label1.” という具合にラベルを連結したものである。右端のドットがルートを表現するが、全てのド

メインがルートを頂点としているため，一般にルートを表す右端のドットは省略される．図 3 は，DNS のパケット構造を表す．



図 4 DNS のパケットの Answer セクション

解決したいドメイン名は，図 4 における Qname フィールドである．各フィールドはサイズが決まっており，Qname フィールドが最大長 255 バイト，リソースレコードのタイプを表す Qtype フィールドとクエリクラスを表す Qclass フィールドがそれぞれ 2 バイトとなっている．

$$(\text{LabelD}).(\text{LabelC}).(\text{LabelB}).(\text{LabelA}). \quad (1)$$

$$(\text{Length}) + (\text{LabelD}) + \dots + (\text{length}) + (\text{LabelB}) + (\text{length}) + (\text{LabelA}) + 0 \quad (2)$$

$$1 + (\text{Max63}) + \dots + 1 + (\text{Max63}) + 1 + (\text{Max63}) + 1 = (\text{Max255}) \quad (3)$$

(1) は，複数のラベルで構成されたドメイン名の例である．(2) は，Question ヘッダーに注入される際のそのドメイン名を表すデータである．Question セクションでは，ドメイン名を表す際にドットは省略され，ラベルの長さとラベル名，そしてドメイン名の終わりを意味する “0” で表現される．(3) は，ラベルの長さとラベル名のサイズを表す．ラベルの長さは，1 バイトのサイズで表現され，ラベル自体の最大長は 63 バイトである．Question セクションの最大長 255 バイトは，ラベルの長さとラベル，そしてドメイン終了を表す “0” を含めた長さである．このため，最初のラベル長を表す 1 バイトとドメイン名の終了を意味する “0” を表すための 1 バイトを差し引いた 253 バイトが，実際のドメイン長の最大長である．

ラベルには，数字とアルファベットおよびハイフン (“-”) を使用することができ，ラベル中に大文字・小文字の区別はない．他方で，アルファベットなどの ASCII

タイプ	意味
A	ホストの IPv4 アドレス
NS	権威サーバ
CNAME	別名
SOA	権威ゾーンの開始
NULL	NULL(実験用)
PTR	ドメイン名のポインター (逆引き)
MX	メール交換
TXT	任意文字列
AAAA	ホストの IPv6 アドレス
SRV	ドメイン名に対するサービスの場所
RRSIG	リソースレコード署名
NSEC	DNSSEC のための不在署名
DNSKEY	DNSSEC のための公開鍵
NSEC3	NSEC のバージョン 3
CAA	証明書を発行する認証局の指定

表 1 主要リソースレコード一覧

以外にも、国際化ドメイン名 (IDN: Internationalized Domain Name) を使用すると日本語やアラビア語なども使用することができる。IDN は、Punycode [6]などのエンコーディング手法に基づき、DNS クエリする際には ASCII コードに変換される。

ドメイン名に関連づけられる情報であるリソースレコードには複数のタイプが定義されており、目的ごとに使い分けられる。例えば、A というレコードタイプは、ドメイン名に対する IPv4 アドレスを対応づけるために用いられる。クライアントがあるドメイン名の IPv4 アドレスを知りたいとき、スタブリゾルバは、ドメイン名と A のレコードタイプを指定することを希望の IPv4 アドレスを取得することができる。表 1 は、主要なリソースレコードである。

Punycode: Unicode 文字列を一意かつ可逆的に ASCII 文字列に変換する符号化方式

2.2 DNS トンネリング

DNS トンネリングは，DNS をデータ転送のメディアとした秘匿通信手法の総称であり，転送元と転送先の方向性によって二つに分類することができる．一つ目が DNS Exfiltration と呼ばれる，スタブリゾルバから権威サーバ方向へのデータ転送手法であり，二つ目が DNS Infiltration と呼ばれ，権威サーバからスタブリゾルバ方向へのデータを転送手法である．DNS トンネリングは，以下に示す DNS の特性に基づいた秘匿手法である．

- 通常のインターネットの利活用において名前解決は必要な機能であるため，多くの組織で DNS のサービスポートが閉ざされることがない
- 名前解決のトラフィックはほとんどのサービスに先立って発生するため，クエリログが肥大化しやすく長期のログ保存が困難である
- パケットフォーマットの構造において，任意の文字列を注入できるフィールドを保持する

DNS トンネリングがデータ転送のキャリアとするフィールドは，クエリの Question セクションの Qname と，Answer セクションの Rdata である．Question セクションの Qname フィールドを利用することで，スタブリゾルバから権威サーバ方向にデータを転送できる．この方向の通信は，ビーコン通信やターゲットから取得した情報を外部に漏えいさせるといった攻撃の最終目的を達成するのに使われる．また，Answer セクションの Rdata フィールドを利用することで，権威サーバからスタブリゾルバ方向にデータを転送することができる．この方向の通信は，ターゲットネットワーク内のホストに潜伏したマルウェアなどへの命令コードを送信するのに使われる．さらに，この二つのキャリアを利用することが双方向の通信路を確保できるため，C2 通信を実施することも可能である．

DNS トンネリングの手法は，ポートスキャンで知られる Nmap のメーリングリストにて初めて一般に公開された (1998 年) [42]．その後，多くの DNS トンネリング実装 [24–38] が公開されながら，実際の攻撃に転化されるまでに至っている．

2.2.1 スタブリゾルバから権威サーバ方向

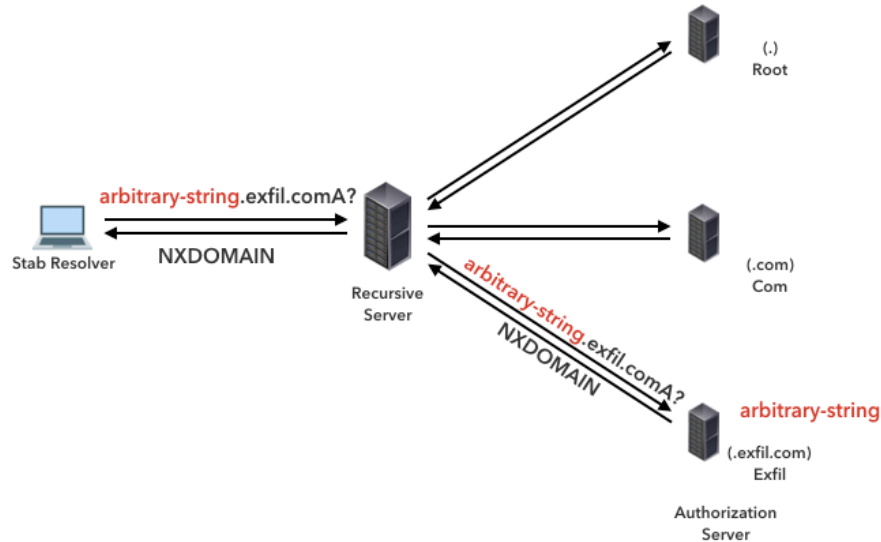


図 5 DNS Exfiltration メソッドに基づいて、ドメイン名のラベル部に任意文字列 (“arbitrary-string”) が注入された DNS クエリが、イントラネット内のスタブリゾルバからインターネット上の権威サーバ (“exfil.com”) に転送される様子

本項では、スタブリゾルバから権威サーバ方向にデータを転送する手法である DNS Exfiltration の詳細について説明する。

DNS Exfiltration は、名前解決として問い合わせられるドメイン名が、そのドメインのゾーンを管理する権威サーバに転送される仕組みを利用した手法である。DNS では、ドメイン名に関連づけられるリソースレコードの情報は、そのドメインをゾーンとする権威サーバが保持しており、ルートから再帰的に問い合わせていくことでその権威サーバからの応答を受け取る。このため、問い合わせられたドメイン名が実在しない場合でも、再帰問い合わせの仕組みに従って、そのドメイン名の最後の権威サーバまで転送されることになる。権威サーバでは通常、クエリされたドメイン名の実在有無に寄らず、問い合わせられたクエリ情報をログとして管理する。このような特性に踏まえて DNS を利用すると、DNS クエリのドメイン名のラベルに組織外ネットワークに転送したいの文字列を注入することで、組織外ネットワーク上に設置された権威サーバにそのデータを転送すること

ができる。これが DNS Exfiltration の動作原理である。

このような仕組みで動作する DNS Exfiltration を動作させるには、宛先となる権威サーバを用意する必要があるため、グローバルなドメインを取得することを前提としている。第 2.1.2 項で述べるように、ドメイン名の最大長は 253 バイトであり、その内ラベルの最大長は 63 バイトまでという制約がある。そのため、DNS Exfiltration 手法を用いてデータを転送する際には、TLD のラベルと宛先権威サーバのラベルもしくは SLD ラベルと権威サーバのラベルを差し引いたサイズが実際に転送できる最大長となる。また、任意の文字列を DNS Exfiltration メソッドを用いて外部に転送するにあたり、転送キャリであるドメイン名における文字列制約を満たすように転送したいデータに前処理を施す必要がある。ドメイン名に使用できる文字列は、第 2.1.2 項で述べるように、“a”から“z”までのアルファベットと“0”から“9”までの数字と先頭以外のハイフン“-”記号である。この文字列制約については、転送したいデータをバイナリデータに変換し、そのバイナリデータをラベルとして印字可能な ASCII コードに変換することでその制約を満たすことができる。この前処理について、既存の DNS トンネリング実装の多くが Base Encoding [5] を用いている。この処理によって、転送データがバイナリデータである際にも転送効率上げたり、ラベルの文字列制約を満たさないデータも転送することができる。また、メッセージの意味抽出を困難にするためのトリックとして用いられる。

いま、DNS Exfiltration を用いて、あるイントラネット内のホストからイントラネット外のホストにデータを転送することを考える。転送される宛先となるイントラネット外のホストには、“exfil.com”より下位の全ての名前空間をゾーンとする権威サーバ (“exfil.com”) を指定する。転送したい文字列にエンコーディング前処理を施した後、“`j用意した文字列j.exfil.com`”という具合に文字列をラベルとして含めることで、ドメイン名が用意できる。適当なりソースレコードタイプを指定し、DNS クエリとして転送すると、その権威サーバにはログとして、文字列を含んだドメイン名を取得する。最後に、受け取ったサーバサイドは、前処理と逆のデコード処理を施すことで、オリジナルのデータを取得できる。以上のように再帰問い合わせとラベルという転送キャリア、エンコーディング処理を組み合

わせることで、イントラネット内のホストから外部ネットワークに任意の情報を転送することができる。これが、DNS Exfiltration の動作メカニズムである。図 5 に、DNS Exfiltration のメカニズムを図解した様子である。

2.2.2 権威サーバからスタブリゾルバ方向

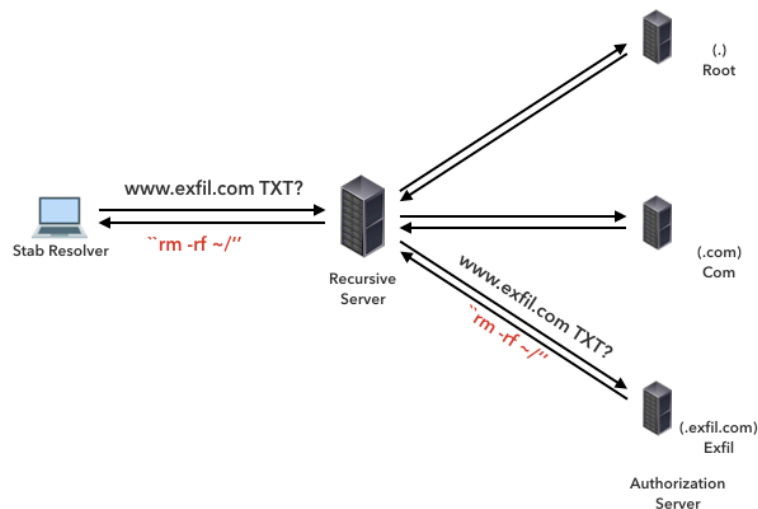


図 6 TXT レコードに登録された情報について、DNS クエリで問い合わせることで権威サーバから命令情報を取得している様子

本項では、権威サーバからスタブリゾルバ方向にデータを転送する手法である DNS Infiltration の詳細について説明する。

DNS Infiltration は、DNS における幾つかのリソースレコードが任意の文字列を記述できる設計を利用したデータ転送手法である。ドメイン名に関連づけられた情報を管理・提供する権威サーバは、ゾーンファイルに関連づけたい情報を記述する。リソースレコードには、レコード情報を検証する機構が備わっていないため、任意の文字列を登録することができる。特に、記法が決まっていない TXT タイプや NULL タイプなどもあり、DNS Infiltration ではこのようなレコードタイプに転送したいデータを登録しておく。このようにして登録されたレコード情報について、名前解決問い合わせることによって、インターネット (権威サーバ)

からイントラネット (スタブリゾルバ) にデータを転送することができる。DNS Infiltration として利用され得るレコードタイプについて、これまでのトンネリング実装で使用されたものに基づいてまとめたのが、表 4 である。

タイプ	最大サイズ (byte)	説明	実装
A	4	ホストの IPv4 アドレス	
NS	4	権威サーバ	[26]
CNAME	253	別名	[25], [26], [32], [38]
NULL	255	NULL(実験用)	[25]
PTR	4	ドメイン名のポインター (逆引き)	
MX	253	メール交換	[25], [26]
TXT	255	任意文字列	[25], [26], [29], [30], [32], [33], [35]
AAAA	32	ホストの IPv6 アドレス	
SRV	180	ドメイン名に対する サービスの場所	[25]
DNSKEY	40	DNSSEC のための公開鍵	[37]

表 2 DNS Infiltration のメソッドとして使用することができるレコードタイプの一覧

NS・CNAME・MX レコードでは、DNS Exfiltration と同じ要領でドメイン名のラベルに転送したい文字列を注入できる。また、NULL・TXT・SRV・DNSKEY を用いる場合には、レコード構文の指定がないため文字列をそのまま注入できる。最後に、A・AAAA・PTR レコードを用いる場合には、転送したい文字列を数字に変換させた後に、ドット (.) 区切りもしくはコロン (:) 区切りで注入できる。

いま、TXT レコードタイプを用いて DNS Infiltration することを考える。

(4) で示すように、権威サーバは、転送したい文字列をゾーンファイルの TXT

レコードとして登録する。

`www.exfil.com IN TXT "rm - rf /"` (4)

次に、スタブリゾルバは、“www.exfil.com”の“TXT”レコードタイプを通常通り問い合わせる。後は、再帰問い合わせの仕組みに基づいて、そのDNSクエリは“exfil.com”まで転送され、ゾーンファイルの“TXT”レコードタイプの値がフルサービスリゾルバを経由したのち、スタブリゾルバまで応答される。DNS Infiltrationの流入通信を図解した様子が、図 6 である。このようにして、DNS Infiltrationでは、正規の名前解決の方法を用いて、インターネットから組織内へとデータを取得することができる。

2.3 検知に基づく既存対策手法

本節では、DNS トンネリングに対するこれまでの対策手法について説明する。

2.3.1 特徴量

本項では、DNS トンネリングメソッドを使用した際に出現する傾向のある特徴量について説明する。

一つ目が、表 3 で示すように、Qname フィールドのドメイン名の長さやパケットサイズである。例えば、DNS Exfiltration において、一回あたりのデータ転送量を増加させる場合、Qname フィールドのドメイン長もそれに比例して長くなり、結果としてパケットサイズも増加するという具合である。DNS Infiltration においても同様で、一回あたりのデータ転送量を増加させる場合、Rdata フィールド内のデータ量も大きくなり、応答パケットのサイズが増加する。このことから、ドメイン名の長さやパケットサイズは DNS トンネリングの特徴といえる。

二つ目が、同一ドメインに対するトラフィック頻度についてである。例えば、転送したいデータが大きい場合、一回につき転送できるデータ量には限界があるため、複数のパケットに分割することによって転送する手法が考えられる。このため、同一ドメインに対して、ラベルの異なる DNS クエリが多数発生することと

種類	ドメイン名
正規	www.example.com
トンネリング	arbitrary-long-text-utill-it-reaches-253byte.example.com

表 3 正規 DNS クエリと DNS トンネリングにおける DNS クエリのドメイン名の長さに関する違い

なる．同様にして，大きなデータサイズの情報をクライアントに転送する際には，複数のトラフィックが発生することになる．以上のことから，特定の時間あたりで同ドメインに対して，トラフィックが急激に増加するものを DNS トンネリングと判断できる可能性があることがわかる．

三つ目は，リソースレコードのタイプについてである．この特徴は，DNS Infiltration メソッドを使用する際に出現する．第 2.2.2 で述べるように，DNS Infiltration では，リソースレコードをデータ転送のキャリアとして利用する．しかし，通常のインターネットの利活用において使用されるリソースレコードのタイプには，偏りが存在することが知られている．Tatang ら [19] が，2017 年 7 月 30 日から 9 月 1 日までの期間にて，取得した

タイプ	パケット数	割合 (%)
A	1,121,025,638	54.90
AAAA	197,388,865	9.67
MX	682,948	00.3
NS	7,662,147	0.38
CNAME	156,708,021	7.68
TXT	41,593,164	2.04
NULL	432,232,574	21.17
Other	84,371,709	4.13

表 4 2017 年 7 月から 8 月までの DNS トラフィックデータセットにおけるリソースレコードのタイプ分布

による調査では，表 ??で示す通り

最後に，Born ら [17] および Cheng ら [18] によって明らかにされた文字列の分布とエントロピーの特徴がある．著者らは，ドメイン名に出現する文字の出現頻度に着目することで，DNS トンネリングメソッドをしようした場合にランダムなドメイン名となる傾向にあることを明らかにした．著者らは，正規のドメイン名と自然言語である英語，および DNS トンネリングにおけるドメイン名に文字の出現頻度をそれぞれ評価した．結果として，正規のドメイン名が英語の自然言語における文字の出現頻度に高い相関関係を示すのに対して，DNS トンネリングにおけるドメイン名と英語には相関がないことから，DNS トンネリングメソッドをしようした場合，ランダムな文字列で構成されるドメイン名となる傾向であることを示している．

これは，本来 DNS では，人が認知しやすくするためにドメイン名を用いるため，正規のドメイン名を作成するとき自然言語に近いラベルを作成する傾向にあることが推測できる．これに対して，DNS トンネリングメソッドを使った場合には，文字列を注入する際にラベルの文字列制約を満たすためにバイナリおよび ASCII コードに変換する必要がある，このためラベルに意味を持たせる

2.3.2 閾値推定

2.3.3 機械学習に基づくモデル

DNS トンネリングメソッドを使用した時の DNS クエリは，で述べるような特性が出現する．この性質に基づき，これまでに多数の検知手法が提案されてきた．

2.3.4 検知迂回手法

3. 関連研究

本章では、次世代 DNS として提案されている名前解決システムのうち提案手法に関連する P2P アーキテクチャによるもの紹介し、提案手法との差分および DNS トンネリングに対する課題を示す。

3.1 P2P ネットワークを利用した名前解決システム

これまでに、DNS における～の課題に対して、P2P に基づいた名前解決システムは数多く提案されてきた。

3.1.1 GNS

3.1.2 SocialNS

3.2 Blockchain を利用した名前解決システム

3.2.1 Namecoin

3.3 課題

本節では、先行研究および新しいアーキテクチャに基づく名前解決システムにおける DNS トンネリングへの課題を示す。(大筋) 検知に基づく手法は、誤検知が避けられない。ペイロードアナリシスに対しては、一回あたりの転送量を調整することでバイパスすることができる。分析の対象がトラフィックの場合は、頻度を長期間に延長することで、ログファイルの肥大によるストレージの過去のログファイルとの分析コストを重くなり分析の隙間をバイパスできる。既存のアーキテクチャに基づくアプローチは、既存の DNS と根本から異なるアーキテクチャを採用しており、マイグレーションが考慮されていない。また、ピュアの P2P アーキテクチャおよびブロックチェーンのアーキテクチャでは、スタブリゾルバからのクエリが権威サーバとブロックチェーンをそれぞれ介することで、依然として、DNS Tunenling として機能し、発生を抑止するメソッドではない。

4. 提案システム

本章では，提案システムについて説明する．はじめに設計目的とシステム概要について説明したのち，具体的なアーキテクチャ・動作メカニズムおよびプロトコルについて説明する．

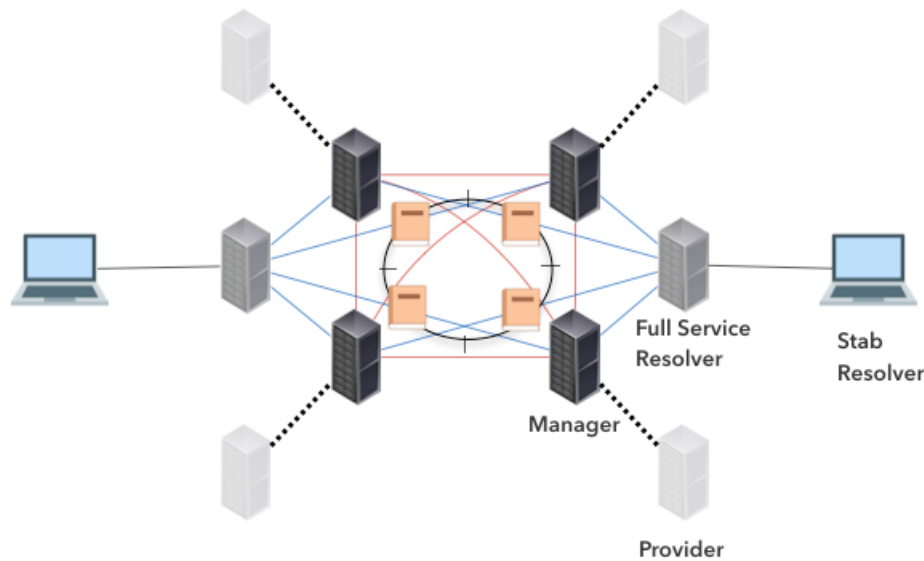


図 7 SORES の概略図

4.1 トンネリング抑止指向の DNS : SORES

本節では，システムの目的と概念について説明する．

第 2.2 で述べたように，DNS を利用することによって任意のデータをイントラネット内からインターネットに転送することができる．また，DNS トンネリングは，DNS の特性を巧みに利用することで検知システムの網を迂回できるポテンシャルがあり，精度を高めると取りこぼすリスクが高まり，取りこぼしを減らす場合誤検知率が高まり，SOC オペレータへの負荷が高まるといった課題がある．このため，DNS トンネリング手法を抑止しながら，既存の名前解決エコシステムを実現するようなセキュアな設計の名前解決システムが求められる．そこで提案

するのが、トンネリング発生抑止を目的とした名前解決システム SORES(Symbol Oriented Records rEsolution System) である。

SORES では、フラットな名前空間をソートされたハッシュ値の範囲に基づきゾーンを分割することで、スタブから任意のコンテンツサーバへのデータ転送を抑止する機能を実現する。提案手法は、リソースレコードのタイプごとにコンテンツの識別子を付与し、この識別子に基づき名前解決やコンテンツを操作する。識別子は、224bit の名前空間をもつハッシュ関数によって算出されるメッセージダイジェストを利用する。具体的なレコード情報は、ゾーンを管理するサーバノードに階層的に連結したノードのよって作成・更新・破棄される。SORES の名前空間は、ソートされたハッシュ値の範囲に基づき分割され、既存システムの TLD によって分割された範囲をゾーンとして分割的に管理される。SORES におけるクライアントは、シンボルを算出することによって、レコード情報を保持するノードを一意に特定することができる。また、シンボルからレコード情報を保持するノードを一意に特定の名前解決プロセスでは、レコード情報を作成したエンティティは介在されない。すなわち、レコード情報を作成する機能とレコード情報を管理する機能を異なるサービスノードに分離させることによって、SORES では DNS Exfiltration が発生することを抑止する。

SORES では、DNS Infiltration に対して、認証基盤の導入と使用できるリソースレコードを制限するメソッドを採用する。認証基盤では、ドメインに関連づけるレコード情報との関連性を評価することで不審なレコード情報がハッシュテーブル上にストアされることを防止する。Qname にシンボルとして 224bit のメッセージダイジェストを使用することによって、副次的効果として、偽装 DNS 応答パケットの作成が困難にできるという性質が期待される。また、既存システムにおける DNSSEC によって実現されてきた送信元のトレーサビリティについて、SORES では認証基盤とシンボルに基づく DNS 応答パケットの偽装困難性によって実現され、DNSSEC の必要性はなくなる。さらに、任意の文字列を注入することができるレコードタイプ “NULL” と “TXT” について、実験目的の NULL タイプの使用制限と TXT タイプは機能をシンタックスの限定している SPF に回帰させることで対処する。

SORES の特徴を以下に示す.

- シンボルに基づく名前解決システム
- フラットな名前空間
- ソートされた範囲に基づくゾーン分割
- レコード情報の操作機能と管理機能の分離
- 認証されたレコード情報

4.2 システム設計

4.2.1 P2P を組み合わせたハイブリッドなアーキテクチャ

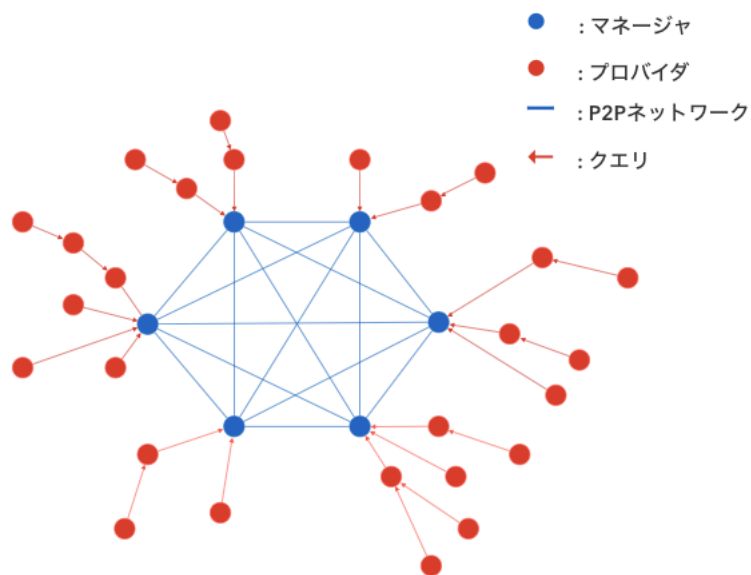


図 8 マネージャとプロバイダの関係図

全体のアーキテクチャに関して, SORES では, 既存の DNS 同様にクライアントサーバアーキテクチャで構成される.

サーバは, フルメッシュネットワークで相互に接続される固定ノードによる P2P ネットワークアーキテクチャで構成される.

SORES におけるサービスノードは、機能に基づいて4つに分けることができる。スタブリゾルバは、既存の DNS と変わらず、DNS クライアントとして名前解決を依頼するノードである。フルサービスリゾルバもスタブリゾルバ同様、1つ追加機能を除いて変更点はなく、スタブリゾルバからの問い合わせに対してリソース情報を保持する主体に代理的に問い合わせ機能と、問い合わせた情報を一定期間キャッシュするキャッシュサーバとして機能するノードである。変更された点は、以降で説明するように、SORES ではレコード情報に対する識別子をキーとしてアクセスする。フルサービスリゾルバは、スタブリゾルバからの従来のフォーマットの DNS クエリについて、ドメイン名とレコードタイプからハッシュ値の算出を行う機能を担う。算出したハッシュ値をキーとして、フルサービスリゾルバは、サーバに転送する。SORES では、権威サーバが機能に基づき二つサービスノードに分割される。

SORES では、既存の DNS における権威サーバの機能を二つのサービスノードで動作させることによって、DNS トンネリングの抑止を図る設計になっている。権威サーバの機能と性質を列挙すると以下ようになる。

機能

- レコード情報を保持する機能
- スタブリゾルバからの名前解決問い合わせに応答する機能
- レコード情報を作成・修正・削除する機能

性質

- ドメイン名

DNS トンネリングは、これら機能と性質を組み合わせることで発生する手法である。例えば、スタブリゾルバから権威サーバ方向のデータ転送手法である DNS Exfiltration は、“スタブリゾルバからの名前解決問い合わせに応答する機能”と“”を組み合わせることで機能する。また、権威サーバからスタブリゾルバ方向のデータ転送手法である DNS Infiltration は、“スタブリゾルバからの名前解決問い合わせに応答する機能”と“ゾーン内のレコード情報を作成・修正・削除する機能”

が組み合わさることで機能するとそれぞれ捉えられる。そこで、SORESでは、権威サーバの機能を二つのサービスノード(マネージャとプロバイダ)に分けることで、既存の名前解決機能とDNS トンネリング抑止機能を実現する。

マネージャは、マネージャは、実際のレコード情報を保持し、クライアントからの問い合わせがあった際に応答するサービスノードである。SORES では、マネージャは、既存のDNSにおけるTLDに相当する権威サーバが担当する。続いて、4つ目のサービスノードとなるプロバイダは、レコード情報を作成・更新および消去といった操作を担当する。プロバイダは、既存のDNSにおけるSLD以降の権威サーバに相当し、ドメインの階層構造を従いマネージャに接続される。プロバイダは、マネージャを介在することで、レコード情報を操作することができる。例えば、example.com プロバイダが“www”のIP アドレス情報を作成することを考える。example.com プロバイダは、“www.example.com”とレコードタイプ“A”およびその値“93.184.216.34”を含むデータを接続先のcom マネージャにリクエストする。com マネージャは、リクエストされたドメイン名とそれに関連づけるレコードタイプから識別子を算出し、担当のマネージャにストアリクエストを転送するという具合で動作する。SORESにおける用語については、表 4.2.1で示す。

4.2.2 識別子と名前空間

本項では、レコード情報にアクセスするために用いる識別子、コンテンツIDとその名前空間について説明する。既存のDNSの名前解決システムでは、正引きをする際、ドメイン名とレコードタイプの二つの情報をキーとして、サーバは保持するゾーンファイルから該当するレコード情報が求まる。他方、SORESでは、ドメイン名とレコードタイプに基づき算出されるコンテンツIDを識別子として指定することでレコード情報にアクセスする。

におけるレコード情報へのアクセス方法は、識別子をキーとすることでデータを取得することができる。コンテンツIDと呼ぶ、この識別子は、ドメイン名とレコードタイプの順番でその文字列の和を引数とするメッセージダイジェストで表現することができる。例えば、ドメイン名を“www.example.com”，レコードタイ

表記	意味もしくは機能
コンテンツ	・ 識別子に関連づけられたレコード情報の実体
コンテンツ ID	・ 識別子
ドメイン ID	・ 識別子 (コンテンツ ID が重複した際に使用)
レコード情報	・ リソースレコードの具体的な値 (例 IP アドレス)
リソースレコードタイプ	・ オブジェクトに関連づけるリソースレコードの型 (例 A, AAAA, MX)
オブジェクト	・ 問い合わせる対象 (ドメイン名もしくは IP アドレス)
スタブリゾルバ	・ 名前解決クライアント
フルサービスリゾルバ	・ スタブリゾルバからのクエリハンドリング ・ 識別子の作成
マネージャ	・ フルサービスリゾルバからのクエリハンドリング ・ ゾーン管理 ・ コンテンツの保持
プロバイダ	・ コンテンツの作成・更新・削除操作

表 5 SORES における用語

プを“A”とした場合、引数となるのは“www.example.comA”という具合である。

続いて、SORES が採用するハッシュアルゴリズムについて説明する。SORES では、全てのコンテンツ ID がフラットな名前空間上にマップされる。従来では、異なるリソースレコードタイプのレコード情報について一つのゾーンファイルに記述することで、管理することができた。他方で、SORES では、ドメイン名とリソースレコードタイプの組をハッシュ値の引数とするため、コンテンツ ID は、レコード情報の数に比例して増加する特性がある。また、識別子の引数の一つにドメイン名が含まれていることから、識別子から元のメッセージが導き出くことが困難な性質を備えていなくてはならない。この性質を満たすことで、なんらか

の方法で識別子を悪意の第三者が取得した際に DNS Exfiltration として機能してしまうことを抑止することができる。最後に、SORES では、既存の DNS のプロトコルフォーマットを流用する設計になっているため、識別子が格納される DNS の Question Section の Qname のサイズ制約を満たさなくてはならない。Qname は、ラベルの最大長が 63byte とする任意長の領域である。以上のことを踏まえて、本システムにおけるハッシュアルゴリズムの要件は、以下の通りである。

1. 名前空間は不足を無視できる程度に大きくなくてはならない
2. アルゴリズムは一方方向性の性質を備えなくてはならない
3. ラベル長は最大 63byte, ドメイン長は最大 253byte である

上記の内容から、SORES では、56byte の名前空間をもつ sha3 のアルゴリズムを採用する。

続いて、分離連鎖法と 2 重ハッシュ法によるコリジョン対策方法について説明する。SORES では、コンテンツのストアリングフェーズで ID にコリジョンが発生した場合、分離連鎖法に基づきストアされるハッシュテーブルに連結リストという形式でコンテンツがストアされる。リスト構造で延長するコンテンツの識別には、ドメイン ID を識別子として利用する。ドメイン ID は、コンテンツ ID と同様のハッシュアルゴリズムを用いて算出されるメッセージダイジェストの先頭 32bit で表現される、ドメイン名を引数として生成される識別子である。例えば、ドメイン名を “www.example.com” とする場合、そのメッセージダイジェストが “86ff20100c058b857bae9785bf0267e6c6afb740c18b8e9a87258485” であるとする、 “86ff” がドメイン ID となる具合である。このように算出されたドメイン ID は、DNS の Question Section のうち、それぞれ 16bit 分の領域を持つタイプとクラスの領域に埋め込まれる。上記の仕組みによって、コリジョンが発生した際には、ドメイン ID をキーとしてコンテンツを識別する。

4.2.3 ソートされたハッシュ空間の範囲に基づいたゾーン

本項では、ゾーンの分割方法およびマネージャノードのアドレスとそのゾーンの範囲に関する対応表について説明する。はじめに比較のために、従来のシステ

ムの場合について説明する。従来のシステムでは、ドメインの階層構造に従い、ドメインの管理ノードを下位のドメイン管理ノードに委譲することでゾーンが分割される。この仕組みでは、ゾーン内の全てのレコード情報はゾーンファイルに画一的にまとめられ、そのゾーンを管理する権威サーバがレコード情報の保持機能とクライアントから応答するという二つの機能を担う。このゾーン分割メソッドでは、レコード情報の帰属が明確であり、ドメインの管理ノードがトラストアンカーとしての役割を同時に担うことができるメリットがある。一方の SORES では、識別子を算出する際に使用するハッシュ関数によって構成される名前空間に基づき、ソートされたハッシュの名前空間の連続した範囲で分割する。この分割された連続した範囲に基づきゾーンがマネージャに割り当てられることで、既存システム同様にレコード情報全体を分散的に管理する。

上記で説明するように、マネージャが管理するゾーンは、ハッシュの名前空間の連続した一部の範囲である。従って、レコード情報は、ハッシュの名前空間上で識別子をソートした際に、帰属する範囲を管理するマネージャによって保持される。マネージャのアドレスを解決する方法には、ゾーンとしてハッシュ値の範囲とそのマネージャおよびマネージャのアドレスに関する対応表 6 によって解決される。SORES では、全てのサービスノードがこの対応表を保持できることを想定しており、ノードは識別子に基づきどのマネージャがコンテンツを保持しているのかを一意に特定する。

ゾーン	マネージャアドレス	ドメイン
(00...00, 4z...zz)	192.35.51.30	com.
(50...00, az...zz)	192.5.6.30	net.
(b0...00, gz...zz)	199.249.112.1	org.
(h0...00, mz...zz)	213.248.216.1	uk.
(n0...00, sz...zz)	199.254.31.1	info.
(t0...00, zz...zz)	194.0.0.53	de.

表 6 6つのマネージャによって管理されるハッシュテーブルにおいて、マネージャの情報とそのマネージャが管理するゾーンが記載された対応表の例

4.2.4 認証基盤に基づく信頼されるレコード情報

本項では，レコード情報の信頼性担保のための認証基盤について説明する．

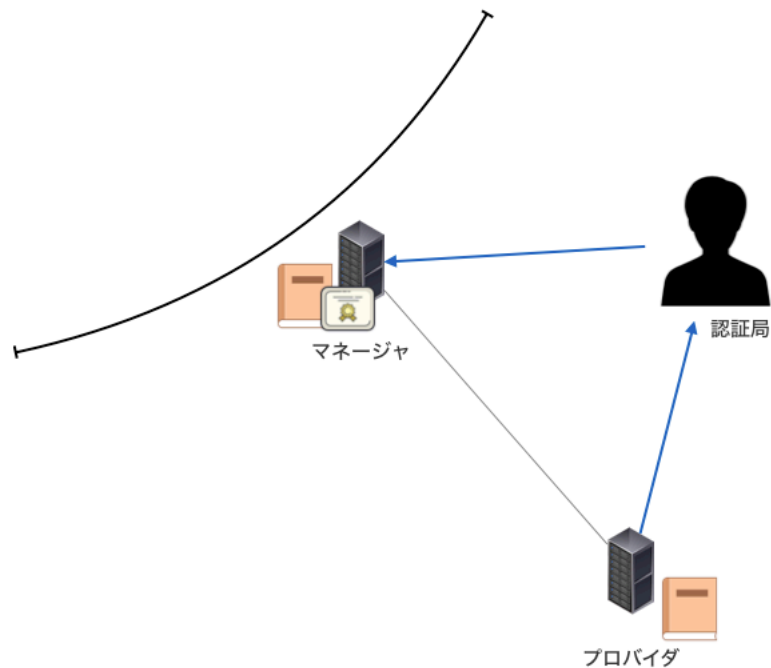


図 9 レコード情報に対する認証プロセス

SORES では，全てのコンテンツについて，信頼される第三者からストアしても良いと認可されていることを前提としている．すなわち，ハッシュテーブル上のコンテンツへの操作，またはコンテンツをハッシュテーブル上にストアするなどの操作処理をする際，プロバイダは，信頼される第三者からのレコード情報に操作することを認可してもらう必要がある．認可の証明書を発行する認証局は，プロバイダの基本情報とレコード情報に基づき証明書の発行を決定する．いま，ドメイン名 “www.example.com” のリソースレコードタイプ A として “93.184.216.34” というレコード情報を関連づけるとする．プロバイダは，認証局に対して証明書発行リクエストを転送する．認証局は，リクエストされたレコード情報について IP アドレスの到達性と不審な文字列が含まれていないこと，利用目的について評価を施す．認可された場合には，その証にデジタル証明書を発行し，マネージャにストアリクエストを転送する．マネージャは，デジタル証明書に付与さ

れた署名に基づきコンテンツの完全性を評価し、認証された場合コンテンツ ID を計算し、担当マネージャにストアリクエストを転送する。上記の手続きを経たコンテンツがハッシュテーブル上にストアされる。

4.2.5 リソースレコードタイプの選別による DNS Infiltration 対策

本項では、SORES で使用するリソースレコードのタイプについて説明する。

はじめに、SORES における DNSSEC の位置づけについて述べる。DNSSEC [4] は、権威サーバからの応答パケットの偽装を検知することを目的として、データの作成元の確認とデータの完全性および、不在情報応答情報の証明する DNS の拡張仕様である。これは、主として DNS の応答パケットを偽装できる程度のパラメータであることに起因する。他方で、SORES では、応答パケットに 224bit のメッセージダイジェストが含まれるため、悪意のある応答パケットをフルサービスリゾルバに意図的にキャッシュすることは極めて困難である。以上の理由から、SORES では DNSSEC の目的にそぐわないため、リソースレコードとして使用されない。

次に、DNSSEC 以外のリソースレコードについて説明する。第 2.2.2 項で示すように、既存の名前解決システムでドメインに関連づけることができるリソースレコードのいくつかのタイプは、DNS Infiltration として機能することができる。DNS Infiltration を抑止するリソースレコードであることの必要条件は、ドメインに関連のない任意の文字列がレコード情報に含められないことである。既存の DNS のリソースレコードのタイプのうち、任意の文字列を含めることができるのタイプは以下の通りである。

表の DNS Infiltration として機能する可能性のあるリソースレコードのタイプのうち、IP アドレスを偽装して情報を転送するものについては、第 4.2.4 項で述べた認証基盤によってレコード情報の正当性評価でスクリーニングすることができる。他方で、任意の文字列を注入できるのが、NULL・TXT・CNAME タイプである。

NULL について考える。NULL タイプの目的は、実験用と定義されている [3]。TXT について考える。CNAME について考える。ホスト名に対する別名で関連

づけることができる CNAME は、一つのサーバにおいてサービスごとにサーバの名前を変更させるために使用される。

4.2.6 コンテンツのデータフォーマット

本項では、マネージャにて管理されるコンテンツのフォーマットについて説明する。

```
{
  "ContentID" : "xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx",
  "Value" : ["DomainID", "QNAME", "RType", "TTL", "Record Information", "Certificate Signature"]
},
{
  "contentID" : "47d87f616f550758706e1633f4de3bf049814eb0a52da6b554cb6",
  "message" : ["86ff", "www.example.com", "A", 3800, "93.184.216.34", '']
},
}
```

図 10 コンテンツファイルのデータフォーマット

4.3 動作メカニズム

4.3.1 レコード情報に対する操作

本項では、SORES において使用されるリソースレコードのタイプと

4.3.2 名前解決

5. 評価

本章では、提案システムの DNS トンネリングに対する抑止機能とシステムの特長について評価した結果を示す。

はじめに、DNS トンネリングへの抑止機能の評価するために、提案システムのプロトタイプを実装した。このプロトタイプ上で、既存のトンネリング実装を実行させた時の動作に基づいて、DNS トンネリングへの抑止機能の評価する。また、提案システムの特長については、既存の DNS による名前解決システムとの比較に基づく評価を行う。

DNS は、現在のインターネットの利活用するにあたり、極めて重要なプロトコルであるために、現実のネットワーク環境に導入した時の影響は極めて大きいことが推察することができる。評価する要素は、以下の 2 点である。

- トラフィック量
- 名前解決速度

5.1 実装

本節では、抑止機能をするために実装したプロトタイプについて説明する。提案システムは、第 4.1 で述べるように、既存の DNS におけるフルサービスリゾルバと権威サーバに修正が加えている。はじめに、提案システムのフルサービスリゾルバが満たす機能の要素は、以下の通りである。

- スタブリゾルバからの名前解決クエリに対して、コンテンツ ID とドメイン名 ID を算出する機能
- コンテンツ ID に基づき、マネージャを探索する機能
- スタブリゾルバからの名前解決クエリを SORES のフォーマットに変換させ、マネージャに転送する機能
- マネージャからの応答パケットを DNS のフォーマットに変換させ、送信元のスタブリゾルバに転送する機能

以上の機能を提供するサービスのアルゴリズムを以下のように設計した.

アルゴリズム 1. スタブリゾルバからのクエリ処理

qname ← *client.data.qname*

rtype ← *client.data.rtype*

start ← *map.range.start*

end ← *map.range.end*

Calculate the Content ID and the Domain ID

calculate_id(*qname*, *rtype*):

content_id ← *hash.sha3_224(qname + rtype)*

domain_id ← *hash.md5(qname)/2*

 return *content_id*, *domain_id*

Find a manager who has zone includes the Content ID

find_manager(*start*, *end*):

 for *i*, *j* in *map.range.start*, *map.range.end* :

 if *i* ≤ *content_id* ≤ *j* :

p ← *map.range.start.index(i)*

manager_addr ← *map.addr[p]*

 break

Query the content to the manager

answer ← *query(manager_addr, content_id, domain_id)*

Transfer the answer to Client

transfer((*client.addr*, *client.port*),

client.data.qname, *answer.header.rcode*, *answer.rdata*)

マネージャ

- プロバイダからのコンテンツストアリクエストに対して，コンテンツ ID を探索し担当マネージャノードにコンテンツ保持をリクエストする機能
- コンテンツを管理するためのデータベースを保持する機能
- フルリゾルバからの名前解決クエリに対して，そのコンテンツ ID に一致するコンテンツを探索し応答する機能

プロバイダ

- コンテンツに対して操作する機能

5.2 DNS トンネリング

5.2.1 DNS Exfiltration

本項では，提案システムの DNS Exfiltration への機能性について評価した結果を述べる．評価には，実装したシステムを

5.2.2 DNS Infiltration

5.3 要素ごとの特性

本節では，DNS との比較評価に基づいた提案手法の特性について説明する．提案システムと既存システムの名前解決における差分には，表 7 のようなものが予想される．

RTT: 名前解決の速度

トラフィック量: ネットワーク全体への影響度

上記の特性に関する評価にあたり，従来の DNS トラフィックのデータセットと提案システムにおける名前解決時のトラフィックデータセットを用意した．これらデータセットについて，統計アプローチに基づいて，評価を行った結果を示す．

	DNS	SORES
ドメイン長	変長 (最大 253byte)	固定長 (72byte) (コンテンツ ID(56) & ドメイン ID(16))
トランザクション 試行回数	ラベル数回	1
RTT	全ての権威サーバ との RTT 総和	マネージャとの RTT のみ
ゾーンファイル	ファイル	インメモリデータベース
その他		・ ハッシュ計算の発生 ・ マネージャ探索

表 7 DNS と SORES の特性比較

SORES では、スタブリゾルバからの名前解決クエリのうち、そのヘッダー情報に基づき識別子を算出することで、この識別子に関連づけられたレコード情報を操作する仕組みに基づき名前解決機能が動作する。DNS では、ゾーンをドメインごとに分割し、フルサービスリゾルバはルート権威サーバから目的の権威サーバに向かって再帰的に問い合わせ、権威サーバのアドレスを解決しながら、最終的にコンテンツを保持する権威サーバからコンテンツを取得する。既存の DNS と SORES の特性を表 7 に示す。

から識別子に基づき、コンテンツを保持するサーバに名前解決リクエストを投げる。SORES は、フルサービスリゾルバにてスタブリゾルバから問い合わせられた名前解決命令に基づきレコード情報の識別子を算出する。この識別子に基づき、フルサービスリゾルバからは、1 ホップでコンテンツを保持するサーバに問い合わせる仕組みで動作する。

5.3.1 RTT(Round Trip Time)

本項では、SORES が識別子から一意にレコード情報にアクセス可能である点について、DNS との比較評価を行う。SORES の名前解決では、従来の DNS では、ラベルごとにゾーンが移譲されている場合、レコード情報を保持するノードまでのホップ数はラベル数 n に比例する。それに対して、SORES では、識別子から一意にレコード情報を保持するマネージャノードを特定できるので、常にホップ数は 2 である。このため、既存の名前解決システムより速度の向上が期待される。

5.3.2 パケットサイズ

SORES では、56byte を固定長とするコンテンツ ID をシンボルとすることによって、レコード情報にアクセスする。この仕組みの影響で、SORES のパケットは従来のパケットと比較して肥大する特性がある。このため、送信元を目的ホストと偽装することで目的ホストの計算リソースを圧迫する DDoS 攻撃に対して、脅威を高める可能性が予想される。

5.3.3 トラフィック量

本項では、提案手法におけるトラフィックについて評価する。評価において、SORES では、シンボル志向の名前解決メソッドによって、既存の再帰問い合わせによるメソッドよりも少ないトラフィックに抑えることが期待される。SORES では、クエリ数とトラフィック量は比例関係にある。他方で、新たにマネージャ間通信という従来にはないトラフィックが発生する。本項では、これトラフィックがネットワーク全体にどの程度影響を及ぼすのかについて評価する。

6. 考察

6.1 マネージャノードの最適な数

6.2 コンテンツ ID を計算する最適ノード

7. 結論

謝辞

ご指導ご鞭撻賜りありがとうございました.

参考文献

- [1] Mitre Att&ck., “Custom Command and Control Protocol, ” <https://attack.mitre.org/techniques/T1094/>, (accessed 2020-1-5).
- [2] Mockapetris, P., “Domain names - concepts and facilities”, STD 13, RFC 1034, November 1987, <https://www.rfc-editor.org/info/std13>.
- [3] Mockapetris, P., “Domain names - implementation and specification”, STD 13, RFC 1035, November 1987, <https://www.rfc-editor.org/info/std13>.
- [4] Arends, R., Austein, R., Larson, M., Massey, D., and S. Rose, “DNS Security Introduction and Requirements”, RFC 4033, March 2005, <https://www.rfc-editor.org/info/rfc4033>.
- [5] Josefsson, S., “The Base16, Base32, and Base64 Data Encodings”, RFC 4648, October 2006, <https://www.rfc-editor.org/info/rfc4648>.
- [6] Costello, A., “Punycode: A Bootstring encoding of Unicode for Internationalized Domain Names in Applications (IDNA)”, RFC 3492, March 2003, <https://www.rfc-editor.org/info/rfc3492>.
- [7] ICANN, “What Is an Internet Covert Channel?,” August 2016. <https://www.icann.org/news/blog/what-is-an-internet-covert-channel>. (accessed 2019-12-12).
- [8] Bortzmeyer, S., “DNS Privacy Considerations”, RFC 7626, August 2015, <https://www.rfc-editor.org/info/rfc7626>.
- [9] KrebsonSecurity. “Deconstructing the 2014 Sally Beauty Breach,” May 2015. <https://krebsonsecurity.com/2015/05/deconstructing-the-2014-sally-beauty-breach/>. (accessd 2019-11-30).
- [10] IronNet. “Chirp of the PoisonFrog,” February 2019. <https://ironnet.com/blog/chirp-of-the-poisonfrog/>. (accessd 2019-11-30).
- [11] Nick Hoffman. “BernhardPOS,” July 2015. <https://securitykitten.github.io/2015/07/14/bernhardpos.html>. (accessd 2019-11-30).
- [12] FireEye. “MULTIGRAIN – Point of Sale Attackers Make an Unhealthy Addition to the Pantry,” April 2016. <https://www.fieeye.com/blog/threat-research/2016/04/multigrain.pointo.html>. (accessd 2019-11-30).

- [13] Palo alto Networks. “New Wekby Attacks Use DNS Requests As Command and Control Mechanism,” May 2016. <https://unit42.paloaltonetworks.com/unit42-new-wekby-attacks-use-dns-requests-as-command-and-control-mechanism/>. (accessd 2019-11-30).
- [14] Kaspersky. “Use of DNS Tunneling for C&C Communications,” April 2017. <https://securelist.com/use-of-dns-tunneling-for-cc-communications/78203/>. (accessd 2019-11-30).
- [15] CISCO Talos. “Spoofed SEC Emails Distribute Evolved DNSMessenger,” October 2017. <https://blog.talosintelligence.com/2017/10/dnsmessenger-sec-campaign.html>. (accessd 2019-11-30).
- [16] Cylance. “Threat Spotlight: Inside UDPOs Malware,” February 27 2018. https://threatvector.cylance.com/en_us/home/threat-spotlight-inside-udpos-malware.html. (accessd 2019-11-30).
- [17] K. Born and D. Gustafson, “NgViz: detecting DNS tunnels through n-gram visualization and quantitative analysis,” Proceedings of the Sixth Annual Workshop on Cyber Security and Information Intelligence Research, Oak Ridge, Tennessee, 2010, pp. 1-4.
- [18] Cheng Qi, Xiaojun Chen, Cui Xu, Jinqiao Shi, Peipeng Liu, “A Bigram based Real Time DNS Tunnel Detection Approach,” Procedia Computer Science, Volume 17, 2013, Pages 852-860.
- [19] Dennis Tatang and Florian Quinkert and Nico Dolecki and Thorsten Holz., “A Study of Newly Observed Hostnames and DNS Tunneling in the Wild, ” arXiv, 2019cs.CR. <https://arxiv.org/abs/1902.08454>.
- [20] J. Liu, S. Li, Y. Zhang, J. Xiao, P. Chang and C. Peng, “Detecting DNS Tunnel through Binary-Classification Based on Behavior Features,” 2017 IEEE Trustcom/BigDataSE/ICISS, Sydney, NSW, 2017, pp. 339-346.
- [21] Asaf Nadler, Avi Aminov, Asaf Shabtai, “Detection of malicious and low throughput data exfiltration over the DNS protocol,” Computers & Security, Volume 80, 2019, Pages 36-53.

- [22] J. Steadman and S. Scott-Hayward, “DNSxD: Detecting Data Exfiltration Over DNS,” 2018 IEEE Conference on Network Function Virtualization and Software Defined Networks (NFV-SDN), Verona, Italy, 2018, pp. 1-6.
- [23] J. Ahmed, H. H. Gharakheili, Q. Raza, C. Russell and V. Sivaraman, “Monitoring Enterprise DNS Queries for Detecting Data Exfiltration from Internal Hosts,” in IEEE Transactions on Network and Service Management.
- [24] “OzymanDNS - Tunneling SSH over DNS,” <https://room362.com/post/2009/2009310ozymandns-tunneling-ssh-over-dns-html/>, (accessed 2020-1-5).
- [25] “iodine,” <http://code.kryo.se/iodine/>, (accessed 2020-1-5).
- [26] “DNScat2,” <https://github.com/iagox86/dnscat2>, (accessed 2020-1-5).
- [27] “TCP-over-DNS,” <http://analogbit.com/software/tcp-over-dns/>, (accessed 2020-1-5).
- [28] “dnscat,” <https://wiki.skullsecurity.org/Dnscat>, (accessed 2020-1-5).
- [29] “DeNiSe,” <https://github.com/mdornseif/DeNiSe>, (accessed 2020-1-5).
- [30] “DNS-shell,” <https://github.com/sensepost/DNS-Shell>, (accessed 2020-1-5).
- [31] “DNSBotnet,” <https://github.com/magisterquis/dnsbotnet>, (accessed 2020-1-5).
- [32] “DNScapy,” <https://github.com/FedericoCeratto/dnscapy>, (accessed 2020-1-5).
- [33] “dohtunnel,” <https://github.com/jansect/dohtunnel/blob/master/doh/doh.go>, (accessed 2020-1-5).
- [34] “goDoH,” <https://github.com/sensepost/goDoH>, (accessed 2020-1-5).
- [35] “DoHC2,” <https://github.com/SpiderLabs/DoHC2>, (accessed 2020-1-5).
- [36] “magicTunnelAndroid,” <https://github.com/MagicTunnel/MagicTunnelAndroid>, (accessed 2020-1-5).
- [37] “dns2tcp,” <https://github.com/alex-sector/dns2tcp>, (accessed 2020-1-5).
- [38] “TUNS,” <https://github.com/lnussbaum/tuns>, (accessed 2020-1-5).

- [39] Ramasubramanian, Venugopalan and Sirer, Emin Gün, “The Design and Implementation of a next Generation Name Service for the Internet, ” Association for Computing Machinery, NY, USA, SIGCOMM Comput. Commun. Rev, number 4 Volume 34, pp.331–342.
- [40] P. Danielis, V. Altmann, J. Skodzik, T. Wegner, A. Koerner and D. Timmermann, ”P-DONAS: A P2P-Based Domain Name System in Access Networks,” in ACM Transactions on Internet Technology (TOIT), vol. 15, no. 3, pp. 1-21, September 2015.
- [41] “Namecoin, ” <https://www.namecoin.org/>, (accessed 2019-12-30)
- [42] “Bugtraq mailing list archives DNS Tunnel - through bastion hosts, ” <https://seclists.org/bugtraq/1998/Apr/79>, (accessed 2019-1-4)

付録

A. 発表リスト (国内研究会)

1. 高須賀 昌烈, 妙中 雄三, 門林 雄基, “非実在ドメインに対するネガティブキャッシュの拡張と再帰問い合わせハッシュ化の提案”, 電子情報通信学会情報ネットワーク研究会, 2019-10-ICTSSL-IN, 2019 年 10 月.