

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 3 月 15 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, 分散ハッシュテーブル, P2P ネットワーク, DNS Tunneling, ネットワークセキュリティ

*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 3 月 15 日.

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

Shoretsu Takasuka

Abstract

Keywords:

Domain Name System(DNS), Distributed Hash Table(DHT), P2P Network, DNS Tunneling, ネットワークセキュリティ

*Master's Thesis, Graduate School of Information Science, Nara Institute of Science and Technology, March 15, 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	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	9
2.2.1 データ流出	9
2.2.2 データ流入	11
2.3 検知に基づく既存対策手法	12
2.3.1 特徴量	12
2.3.2 閾値推定	12
2.3.3 機械学習に基づくモデル	12
3. 関連研究	13
3.1 P2P ネットワークを利用した名前解決システム	13
3.1.1 GNS	13
3.1.2 SocialNS	13
3.2 Blockchain を利用した名前解決システム	13
3.2.1 Namecoin	13
3.3 課題	13
4. 提案システム	14
4.1 トンネリング抑止指向の DNS : SORES	14
4.2 システム設計	16

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	16
4.2.2	識別子と名前空間	17
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	19
4.2.4	認証基盤に基づく信頼されるレコード情報	21
4.2.5	リソースレコードタイプの選別による DNS Infiltration 対策	22
4.2.6	コンテンツのデータフォーマット	23
4.3	動作メカニズム	24
4.3.1	レコード情報に対する操作	24
4.3.2	名前解決	24
5.	評価	25
5.1	実装	25
5.2	実験環境	25
5.3	DNS Tunneling	25
5.3.1	DNS Exfiltration	25
5.3.2	DNS Infiltration	25
5.4	要素ごとの特性	25
5.4.1	RTT(Round Trip Time)	26
5.4.2	パケットサイズ	26
5.4.3	トラフィック量	27
6.	議論	28
6.1	クエリにおけるハッシュ値計算の最適ノード	28
6.2	今後の課題	28
7.	結論	29
	謝辞	30
	参考文献	31
	付録	34

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	arbitrary-string という任意の文字列が, DNS クエリのラベル部を用いて, 事前に用意した権威サーバ (exfil.com) に転送される様子	10
4	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	11
5	SORES の全体図	15
6	マネージャとプロバイダの関係図	16
7	レコード情報に対する認証プロセス	21
8	コンテンツファイルのデータフォーマット	23

表 目 次

1	主要リソースレコード一覧	8
2	SORES における用語	18
3	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	20
4	DNS Infiltration として利用することができるリソースレコード一覧	23
5	DNS と SORES の特性比較	26

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、高度な手法で執拗に攻撃対象を狙う標的型攻撃において、その攻撃実行者は、悪性通信を通常通信に偽装させるといった秘匿手法を用いることが知られている。この手法を用いることで、攻撃は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 [5] 手法を用いた攻撃に対処するために、通信の監視だけでなく、プロトコル設計段階で悪用されないセキュアな設計にすることが求められる。

攻撃者が使用する秘匿通信の中で、代表的な手法の一つに DNS トンネリングがある。DNS トンネリングは、DNS(Domain Name System) と呼ばれるドメイン名に関連づけられた情報を解決する通信プロトコルを利用した秘匿通信手法である。この DNS トンネリング手法は、DNS が持つ特性のうち、DNS が現在のインターネットにおいて極めて重要な位置づけにあり多くの組織でフィルタリングされることが少ないという特性と、他の秘匿通信手法と比べてより多く情報を注入できるという特性に基づいて機能する。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 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 [7-14]。

DNS トンネリングに対する従来の対策は、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

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 : Command & Control サーバ。

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 [15–20]。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 [21–23] 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。しかし、ツールキットによって生成されるトラフィックは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して使用される実際の攻撃シーンにおける DNS トンネリングとは異なることが予想される。秘匿性を高めた 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 [18]。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 を用いた秘匿通信の発生を抑止することを目的に、新しい名前解決手法を提案する。アーキテクチャを変更にするあたり、既存システムとの互換性をどのように担保しながら、移行させるのかという問題がある。提案手法では、現実の DNS として利用されることを考慮して、フルサービスリゾルバとコンテンツを保持するサーバノードのみに変更点を加えることに留めた。すなわち、コンテンツを保持するサーバ同士を連携させ、リソースレコードを登録したいノードと実際にコンテンツの保持ノードを分離させることによって、DNS トンネリング抑止を実現させる。また、提案手法では、既存システムにおけるルート権威サーバを持たない設計になっており、現在問題となっているルート権威サーバにかかる集中的な負荷を分散することが期待される。さらに、再帰的な問い合わせでなく、コンテンツを保持するサーバが一意に決定される設計による、ネットワーク全体にかかるトラフィック量の削減に貢献することが期待される。本提案手法は、次世代の名前解決システムにおける DNS トンネリングを抑止する際

のアプローチとして応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、本論に対象とする脅威モデルである DNS トンネリングについて説明する。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 [1, 2]。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。

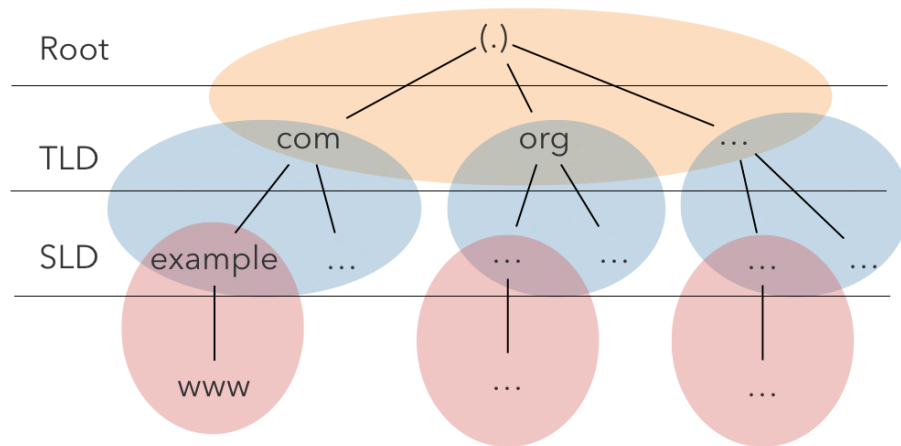


図 1 ドメインにおけるゾーンごとの名前空間

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。

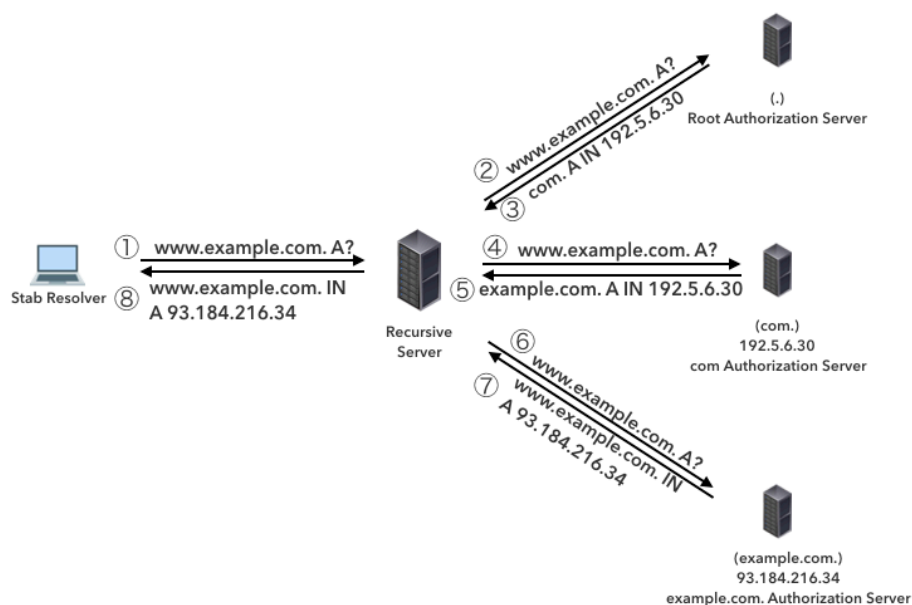


図 2 DNS による名前解決

2.1.2 名前解決

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転送する。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

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。DNS ヘッダーの Question セクションの最大長は、255 オクテットである。

$$(Length) + (LabelD) + \dots + (LabelB) + (length) + (LabelA) + 0 \quad (1)$$

$$1 + (Max63) + 1 + (Max63) + 1 + (Max63) + 1 + (Max63) + 1 = (Max255) \quad (2)$$

$$(LabelD).(LabelC).(LabelB).(LabelA) = (Max253) \quad (3)$$

このセクションにドメイン名をそのまま注入するのではなく、はじめにラベル間のドットは省略され、注入するラベルの長さで末尾に終了を示す “0” が含まれる必要がある。すなわち、ラベルの数だけラベル長を表す 1 オクテットと終了の

1 オクテット (octet):8bit

“0”を示すための1オクテットを含めて、253オクテットが実際に使用できるドメイン長の最大である。ラベルの最大長は、63オクテットである。

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。

タイプ	値	意味
A	1	ホストのIPv4アドレス
NS	2	権威サーバ
CNAME	5	別名
SOA	6	権威ゾーンの開始
NULL	10	NULL(実験用)
PTR	12	ドメイン名のポインター(逆引き)
MX	15	メール交換
TXT	16	任意文字列
AAAA	28	ホストのIPv6アドレス
RRSIG	46	リソースレコード署名
NSEC	47	DNSSECのための不在署名
DNSKEY	48	DNSSECのための公開鍵
NSEC3	50	NSECのバージョン3
CAA	257	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

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 5 は、主要なリソースレコードである。

2.2 DNS トンネリング

本節では、DNS トンネリングの仕組みや特定について説明する。

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、分解すると、流出通信を DNS Exfiltration、流入通信を DNS Infiltration という。DNS は、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り、現在のインターネットにおいて極めて重要な役割を担うネットワークプロトコルである。この特性を持つ DNS を悪用することで、秘匿通信を実現するに非常に適している。

このため、DNS のトラフィックは組織においてモニタリングされることが少なく、秘匿通信を実現するのに都合が良い。

さらに、DNS トンネリングは、

しかし、DNS のデザインは、クライアントサーバアーキテクチャに基づき、クライアントのスタブリゾルバと権威サーバが直接やり取りされるため、の問い合わせ情報が権威サーバに転送されるため、転送される仕組みは、データ転送の仕組みでもある。過去のインシデントでは、権威サーバをデータ回収のサーバとして利用するケースや、次 2.2.2 項の流入通信とを組み合わせ、命令サーバとの相互通信の流出方向の通信手段として利用されてきた。

2.2.1 データ流出

DNS Exfiltration は、再帰問い合わせの仕組みによって、DNS クエリがコンテンツを保持する権威サーバに転送される仕組みを利用した手法。これによって、クライアントから権威サーバを宛先とする方向に対して任意のデータを転送することができる。データを転送するためのキャリアは、DNS クエリの Question Section の QNAME 内のラベルである。第 ?? 節で示すように、ドメイン名には一つのラベルあたり 63byte、全体で 253bit という制約がある。すなわち、DNS

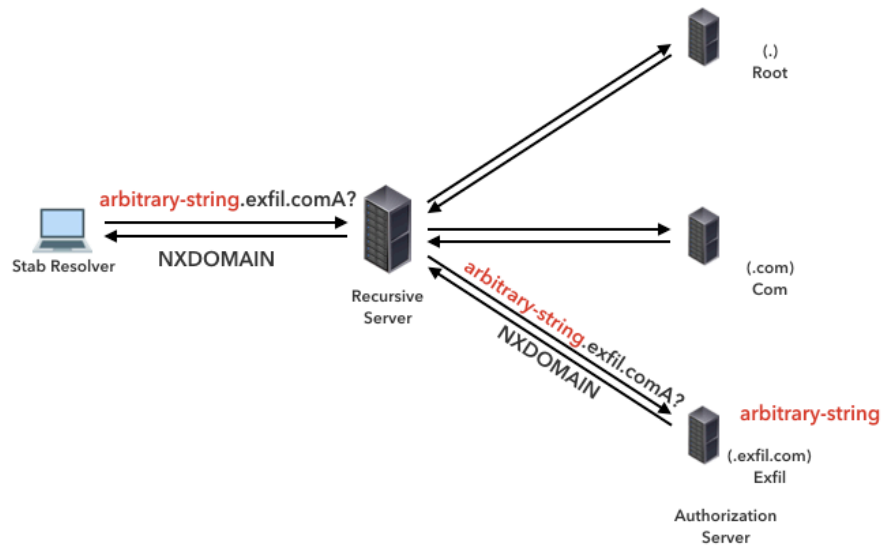


図 3 arbitrary-string という任意の文字列が，DNS クエリのラベル部を用いて，事前に用意した権威サーバ (exfil.com) に転送される様子

Exfiltration によって転送できるデータ量の最大は 253byte である．他方で，DNS の名前解決の仕組みはインターネットの利活用において必要不可欠であるため，多くの組織において DNS のトラフィックがフィルタリングされることはほとんどない．このような性質をもつ DNS を利用することで，悪意を持つユーザがなんらかの方法で潜伏した組織内部ネットワークから情報を流出させる際の手法として，他のネットワークプロトコルに比べて都合が良い．

DNS Exfiltration において，データを転送する際の宛先ノードは特定のゾーンを管理する権威サーバである．例えば，あるローカルネットワーク内に位置づくノードからグローバルネットワーク内の “exfil.com” という権威サーバにデータを転送することを考える．一般に，転送するデータを QNAME におけるラベルに埋め込む際には，Base Encoding [4] のような変換規則によるエンコード処理が施される．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．変換された文字列は，

この権威サーバは，“exfil.com”以下の全ての名前空間をゾーンとして管理することができる。データを転送する際のキャリアとなるのは，DNS クエリの“exfil.com”以下のラベルである。DNS のラベルは，数字・アルファベット・先頭以外でハイフン(“-”)という文字列の制約があるため，任意の文字列をそのまま注入することは困難である。任意の文字列を文字列エンコーディングの手法で変換させ，用意できた文字列を QNAME に含め，適当なりソースレコードタイプを指定すれば，宛先となる権威サーバにデータが転送される。最後に，ラベルを同一の文字列エンコーディングアルゴリズムでデコードすることで，元のデータを受け取ることができるという具合である。このようにして，DNS の名前解決の仕組みを応用することで，任意の権威サーバに任意の文字列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 3 に，DNS Exfiltration のメカニズムを図解した様子である。

2.2.2 データ流入

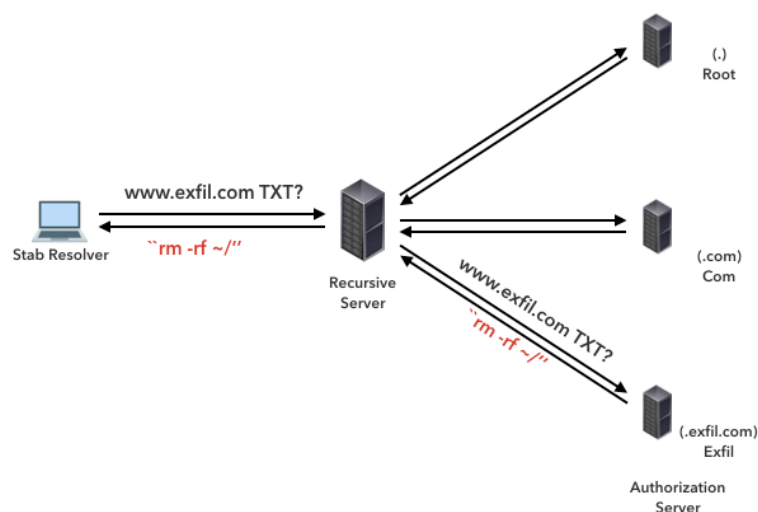


図 4 TXT レコードに登録された情報について，DNS クエリで問い合わせることで権威サーバから命令情報を取得している様子

DNS Infiltration は，DNS Exfiltration と逆方向の権威サーバからクライアント方向にデータを転送する手法である。この手法は，DNS のリソースレコードに

任意の文字列が含まれる設計になっていることを利用したものである。リソースレコードのタイプは多種多様であるが、ゾーンファイルを自由に編集できる現在の DNS エコシステムでは、ドメイン名との関わりに関係なく任意の情報をドメイン名に関連づけることが可能である。DNS Infiltration を実施するには、送信元となる権威サーバのドメイン名に、適当なリソースレコードのタイプ (E.g. TXT) に転送したい文字列を登録しておくことで、そのドメイン名を QNAME としてそのレコードタイプを問い合わせることで、クライアントのもとでデータを回収することができるという具合である。DNS のリソースレコードを転送キャリアとする流入通信のメカニズムを図解した様子が、図 4 である。

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

DNS トンネリングの手法を用いる場合、いくつかの特徴が現れることがある。Born ら [15] は、DNS トンネリングにおけるクエリ内のドメイン名に出現する文字の出現頻度を着目することで、正規のドメイン名と DNS Exfiltration メソッドによって生成されたドメイン名とを分類のに有用であることを明らかにした (2010)。著者らは、正規のドメインであれば英語のような自然言語における文字の出現頻度と高い相関があることに加えて、エントロピーと文字列分布に高い相関があることも明らかにした。

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。

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．

4.1 トンネリング抑止指向の DNS : SORES

本節では、システム概念と目的について説明する．

SORES(Symbol Oriented Records rEsolution System) は、全てのレコード情報をシンボルとなる識別子が付与し、フラットな名前空間で管理することによって、問い合わせの際にシンボル指向でレコード情報を解決する名前解決システムである．シンボルには、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使用する．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、

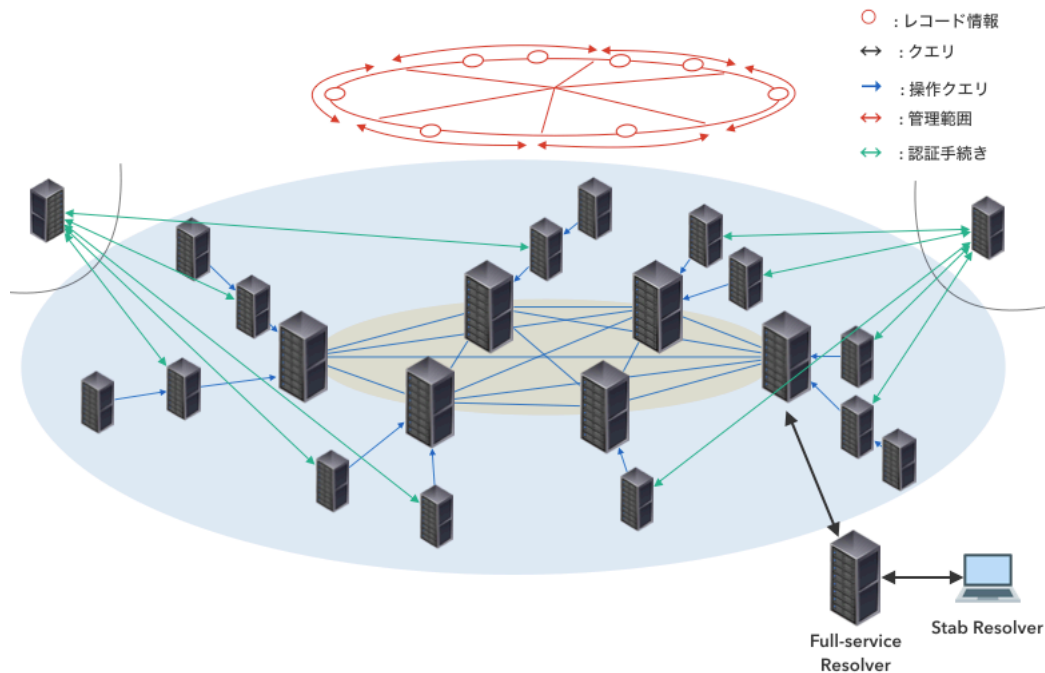


図 5 SORES の全体図

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

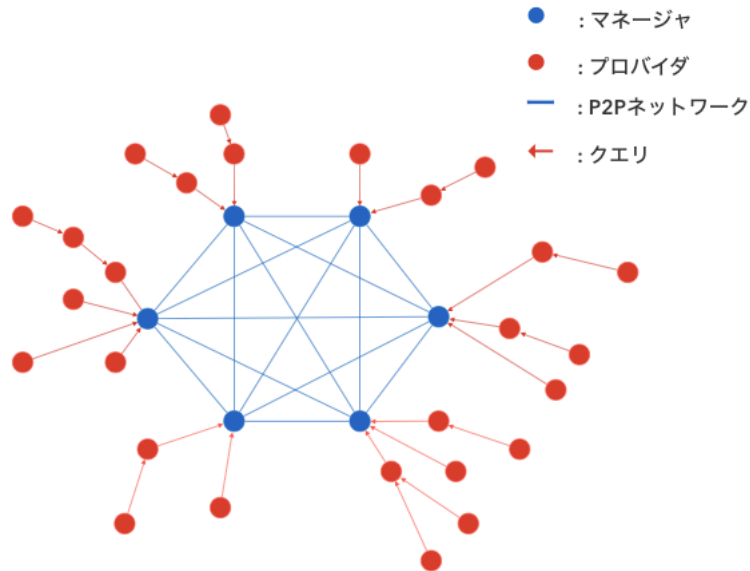


図 6 マネージャとプロバイダの関係図

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と同様に動作するノードであり、DNSクライアントとして名前解決を依頼する主体として位置づくノードである。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。

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に応答するサーバノードである。また、マネージャは、既存のDNSにおけるTLDに位置づく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。

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
レコード情報	・ リソースレコードの具体的な値 (E.g. 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 (E.g. A, AAAA, MX)
オブジェクト	・ 問い合わせる対象 (E.g. 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 2 SORES における用語

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る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. 最大 63byte の最大ラベル長を満たさなくてはならない

ここで、既存のハッシュ関数の特性をまとめた table を示す。

上記の内容から、SORES では、224bit 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 と、“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 3 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 3 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．

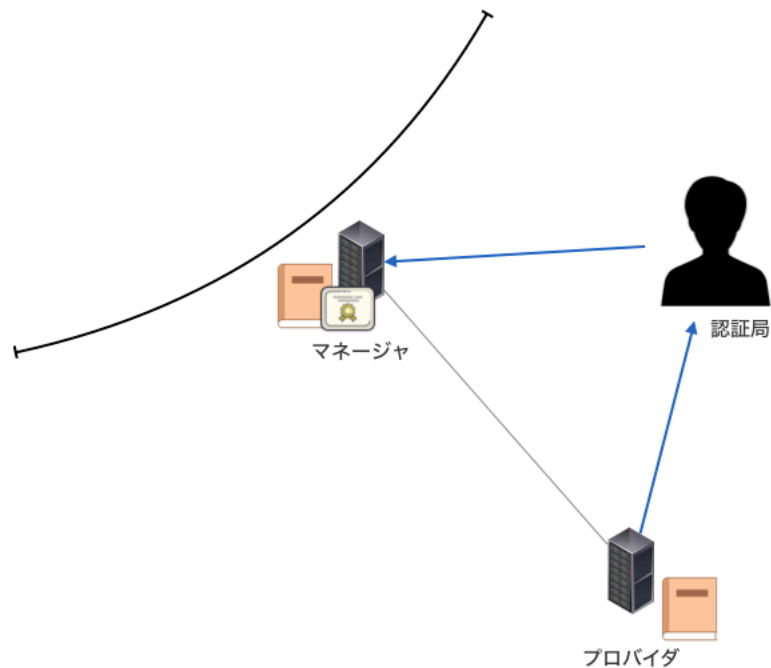


図 7 レコード情報に対する認証プロセス

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 [3] 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。

上記の 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”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 [2]。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

タイプ	意味	Infiltration の手法
A	ホストの IPv4 アドレス	
NS	権威サーバ	
MF	メール転送サーバ	
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	
DNSKEY		

表 4 DNS Infiltration として利用することができるリソースレコード一覧

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 Infomation", "Certificate Signature"]
},
{
  "contentID" : "47d87f616f550758706e1633f4de3bf28f049814eb0a52da6b554cb6",
  "message" : ["86ff", "www.example.com", "A", 3800, "93.184.216.34", '']
},
}
```

図 8 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 Tunneling に対する機能性について評価した結果を示す．また，ネットワークに対する影響を測るために，3つの特性 (RTT・パケットサイズ・トラフィック量) について既存システムと比較した結果を述べる．評価にあたり，提案システムを実装した．はじめに，システム実装で採用したアルゴリズムに本章では，提案システムの DNS Tunneling に対する機能性評価とネットワークへの影響度について各要素ごとに関する特性評価を示す．

5.1 実装

本節では，提案システムの機能性および特性評価を目的として実装した

5.2 実験環境

5.3 DNS Tunneling

5.3.1 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

5.3.2 DNS Infiltration

5.4 要素ごとの特性

本節では，SORES の特性について既存の DNS と比較に基づき評価する．

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。以上のことから、SORES では、以下のような特性を持つことがわかる。

	DNS	SORES
RTT	全ての権威サーバとの RTT の総和	マネージャとの RTT
ドメイン長 (パケットサイズ)	変長 (最大 253byte)	固定長 (56byte)
トランザクション回数	ラベル数依存	1
トラフィック量	線形増加	線形増加
リクエスト速度	N/A	ハッシュ計算分遅延

表 5 DNS と SORES の特性比較

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。

5.4.1 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。

5.4.2 パケットサイズ

SORES では、224bit 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．

5.4.3 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. 議論

6.1 クエリにおけるハッシュ値計算の最適ノード

6.2 今後の課題

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] P.V. Mockapetris. “Domain names - concepts and facilities. RFC 1034 (INTERNET STANDARD),” November 1987. Updated by RFCs 1101, 1183, 1348, 1876, 1982, 2065, 2181, 2308, 2535, 4033, 4034, 4035, 4343, 4035, 4592, 5936.
- [2] P.V. Mockapetris. “Domain names - implementation and specification. RFC 1035 (INTERNET STANDARD),” November 1987. Updated by RFCs 1101, 1183, 1348, 1876, 1982, 1995, 1996, 2065, 2136, 2181, 2137, 2308, 2535, 2673, 2845, 3425, 3658, 4033, 4034, 4035, 4343, 5936, 5966, 6604.”
- [3] R. Arends, R. Austein, M. Larson, D. Massey, S. Rose. “DNS Security Introduction and Requirements, ” March 2005. Updated by RFCs 6014, 6840.
- [4] S. Josefsson. “The Base16, Base32, and Base64 Data Encodings, ” October 2006. Updated by RFC 3548.
- [5] 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).
- [6] S. Bortzmeyer. “DNS Privacy Considerations RFC 7626 (INTERNET STANDARD), ” August 2015.
- [7] KrebsOnSecurity. “Deconstructing the 2014 Sally Beauty Breach,” May 2015. <https://krebsonsecurity.com/2015/05/deconstructing-the-2014-sally-beauty-breach/>. (accessd 2019-11-30).
- [8] IronNet. “Chirp of the PoisonFrog,” February 2019. <https://ironnet.com/blog/chirp-of-the-poisonfrog/>. (accessd 2019-11-30).
- [9] Nick Hoffman. “BernhardPOS,” July 2015. <https://securitykitten.github.io/2015/07/14/bernhardpos.html>. (accessd 2019-11-30).
- [10] Fireeye. “MULTIGRAIN – Point of Sale Attackers Make an Unhealthy Addition to the Pantry,” April 2016. https://www.fireeye.com/blog/threat-research/2016/04/multigrain_pointo.html. (accessd 2019-11-30).
- [11] 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).

- [12] Kaspersky. “Use of DNS Tunneling for C&C Communications,” April 2017. <https://securelist.com/use-of-dns-tunneling-for-cc-communications/78203/>. (accessed 2019-11-30).
- [13] CISCO Talos. “Spoofed SEC Emails Distribute Evolved DNSMessenger,” October 2017. <https://blog.talosintelligence.com/2017/10/dnsmessenger-sec-campaign.html>. (accessed 2019-11-30).
- [14] Cylance. “Threat Spotlight: Inside UDPOs Malware,” February 27 2018. https://threatvector.cylance.com/en_us/home/threat-spotlight-inside-udpos-malware.html. (accessed 2019-11-30).
- [15] 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.
- [16] 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.
- [17] 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 Trust-com/BigDataSE/ICISS, Sydney, NSW, 2017, pp. 339-346.
- [18] 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.
- [19] 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.
- [20] 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.
- [21] “OzymanDNS - Tunneling SSH over DNS,” <https://room362.com/post/2009/2009310ozymandns-tunneling-ssh-over-dns.html/>, (accessed 2019-11-20).

- [22] “iodine,” <http://code.kryo.se/iodine/>, (accessed 2019-11-20).
- [23] “DNScat2,” <https://github.com/iagox86/dnscat2>, (accessed 2019-11-20).
- [24] 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.
- [25] 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.
- [26] “Namecoin, ” <https://www.namecoin.org/>, (accessed 2019-12-30)

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.