Computer Security Symposium 2019
21 - 24 October 2019

ダークネットトラフィックの分析に基づく継続的な広域ネットワークスキャンの調査

中川 雄太[†1,†2,a)]　韓 燦洙[†1]　島村 隼平[†5]　高橋 健志[†1]　藤田 彬[†3]　吉岡 克成[†1,†3,†4]　井上 大介[†1]

概要：インターネット空間を継続的にスキャンすることで，インターネットに接続されている様々な機器の情報を定常的に収集するサービスもしくは組織の存在が知られている（以降，“スキャンシステム”と呼ぶ）．いくつかのスキャンシステムはスキャンホストの Web サービスでスキャンの目的や利用している IP アドレスなどを公開しているが，その詳細はまだわからない部分が多い．さらに，これらの情報を公開しないまま活動するスキャンシステムが存在すると推測され，その実態の解明が急務である．しかしながら，このような非公開のスキャンシステムの実態を把握するのは容易ではない．そこで我々は，ダークネットで観測される通信について送信元 IP アドレス帯ごとに分析を行い，非公開のスキャンシステムを含む複数の IP アドレス帯から，様々なポートに対して継続的な広域ネットワークスキャンが行われていることを明らかにした．さらに，これらホスト群のスキャン対象ポートとその時系列変化を分析することによって，送信元 IP アドレス帯によってその傾向が異なることを明らかにした．

キーワード：ダークネット, トラフィック分析，ポートスキャン，広域ネットワークスキャン

An Investigation of Constant Internet-Wide Scanning through Analysis of Darknet Traffic

Yuta Nakagawa[†1,†2,a)]　Chansu Han[†1]　Jumpei Shimamura[†5]　Takeshi Takahashi[†1]
Akira Fujita[†3]　Katsunari Yoshioka[†1,†3,†4]　Daisuke Inoue[†1]

Abstract: It is known that some services or organizations have carried on constant Internet-wide scanning to collect information about various devices connected to the Internet. Some of them publicize their activities and objectives of the scan at web services on the scan hosts. However, their details are undisclosed. In this study, we found that multiple hosts actually perform constant Internet-wide scanning through analyzing the darknet traffic from a statistical point of view. Furthermore, we also revealed their scanning patterns are different for each, based on an analysis of scan target ports and transitions of the port.

Keywords: Darknet, Traffic Analysis, Port Scanning, Internet-wide Scanning

1. はじめに

大局的なサイバー攻撃の動向を把握する手段の 1 つとして，ダークネットの観測が挙げられる．ダークネットとは，インターネットから到達可能かつ未使用の IP アドレス空間のことである．本来，未使用の IP アドレスに対して通信は発生しないはずであるが，攻撃者やマルウェアが脆弱な

[†1] 情報通信研究機構 サイバーセキュリティ研究所
Cybersecurity Research Institute, National Institute of Information and Communications Technology.
[†2] 横浜国立大学大学院 環境情報学府
Graduate School of Environment and Information Sciences, Yokohama National University.
[†3] 横浜国立大学 先端科学高等研究院
Institute of Advanced Sciences, Yokohama National University.
[†4] 横浜国立大学大学院 環境情報研究院
Graduate School of Environment and Information Sciences, Yokohama National University.
[†5] 株式会社クルウィット clwit Inc.
[a)] nakagawa-yuta-ck@ynu.jp

サーバの探索や感染拡大を目的としてインターネット空間をスキャンしているため，ダークネットでは多くの通信が観測される．NICTER [1] では，国内外の約 30 万の IP アドレスをダークネットとして観測しており，日々上記のようなサイバー攻撃の観測・分析が行われている．このようにダークネットでは従来，主にマルウェアによるサイバー攻撃が観測されてきた．

しかし近年，大学などの研究組織やセキュリティ関連企業を中心に，インターネット空間を継続的にスキャンすることで，インターネットに接続されている様々な機器の情報を収集する調査目的と見られる活動が急激に増加しており [10], [11], [12], [13], [14]，実際にこのような広域ネットワークスキャンがダークネットでも数多く観測されている [8]．（以降，本稿ではこのような継続的な広域ネットワークスキャンを行う組織もしくはサービスを “スキャンシステム”と呼ぶ．）

いくつかのスキャンシステムは，スキャンホストの Web サービスでスキャンの目的や利用している IP アドレス帯を記載したオンラインドキュメントを公開しているが，具体的にどのようなポートを対象にスキャンを行っているかなど，その詳細は公開されていないケースが多い．さらに，これらの情報を公開しないまま活動する非公開のスキャンシステムが存在すると推測され，その実態の解明が急務である．しかしながら，このような非公開のスキャンシステムの実態を把握することは容易ではない．

そこで我々は，IP アドレス帯を基準とするスキャンシステムのモデルを仮定し，ダークネットで観測される通信について送信元 IP アドレス帯ごとに分析を行うことで，上記のようなオンラインドキュメントが確認できないホストを含む多数の IP アドレス帯から，継続的な広域ネットワークスキャンが行われていることを明らかにした．さらに，これらの IP アドレス帯のスキャン対象ポートとその時系列変化を分析することで，IP アドレス帯によってその傾向が異なることを示した．

本稿の構成は次の通りである．まず，2 章で調査対象となるスキャンシステムの提案モデルと広域ネットワークスキャンの定義について説明する．3 章でダークネットの観測データとその分析結果について述べ，4 章で考察を行う．5 章で関連研究について説明し，6 章でまとめと今後の課題を述べる．

2. 調査対象のモデル化

本章では，調査対象となるスキャンシステムのモデルを仮定し，さらに同システムが行う広域ネットワークスキャンの定義とその継続性の評価手法について説明する．

2.1 提案モデル

ダークネットで観測される主なスキャンの 1 つとして，ボットネットによる侵入先の探索が挙げられる．例えば，2016 年に猛威を振るった Mirai ボットネットはランダムに生成した IP アドレスの 23/tcp や 2323/tcp に対してスキャンを行い，侵入を試みる事が知られている [4]．新たに感染したボットネットも同様にスキャンを行うため，感染が拡大したボットネットから発生するスキャンは特定の IP アドレス帯によらず，インターネット空間の広い範囲から発生する場合が多い．一方，スキャンシステムはスキャンを専門に行う比較的少数のホストを用意し，継続的に広域ネットワークスキャンを行なっていると考えられる [5]．

そこで我々は，スキャンシステムが比較的少数のスキャンホストから構成される場合，それらホスト群が特定の IP アドレス帯に集中していると仮定し，ダークネットで観測される送信元 IP アドレス帯ごとに，観測される広域ネットワークスキャンとその継続性の評価を行う．なお，3 章のダークネットトラフィックの分析では送信元 IP アドレスの第 3・第 4 オクテットを無視し，送信元 IP アドレス帯をクラス B ネットワーク単位で扱った．

2.2 広域ネットワークスキャンの定義と継続性の評価

観測に用いるダークネットの全 IP アドレスの集合を $D = \{d_0, d_1, \ldots\}$，またダークネットの観測期間を適切な時間で分割した観測期間の集合を $T = \{t_0, t_1, \ldots\}$ で表す．

広域ネットワークスキャンの定義

本稿における広域ネットワークスキャンの定義について述べる．期間 t_i において，ある IP アドレス帯 A からポート p にパケットが届いたダークネットの IP アドレスの部分集合を $D_{t_i}(A \to p)$ とする．$D_{t_i}(A \to p)$ が閾値 $\alpha(0 < \alpha \leq 1)$ に対し

$$\frac{|D_{t_i}(A \to p)|}{|D|} \geq \alpha$$

を満たすとき，p を期間 t_i における A のスキャン対象ポートとみなす．

期間 t_i における A のスキャン対象ポートの集合（以降，ポートセットと呼ぶ）を $P_{t_i}(A) = \{p_0, p_1, \ldots\}$ で表す．$P_{t_i}(A)$ が閾値 n に対し

$$|P_{t_i}(A)| \geq n$$

を満たすとき，期間 t_i において A から広域ネットワークスキャンが行われたと判断する．

継続性の評価

ある IP アドレス帯 A から広域ネットワークスキャンが行われた期間の集合を $T(A)$ で表す．$T(A)$ が閾値 m に対し

$$|T(A)| \geq m$$

を満たすとき，A から継続的な広域ネットワークスキャンが行われたと判断する．

表 1　各ダークネットの IP アドレスの規模

ダークネット	I	II	III	IV	合計
規模	899	505	256	253	1913

本稿では，このような継続的な広域ネットワークスキャンを行う IP アドレス帯を調査対象として扱う．

3. ダークネットトラフィックの分析

本章では，分析に用いたダークネットトラフィックの観測データおよびその分析結果について述べる．

3.1 ダークネットの観測

2018 年 7 月 1 日から 2019 年 6 月 30 日の期間中に，国内および国外のクラス A が異なる 4 か所のダークネットで観測された TCP の SYN パケットを対象に分析を行った．それぞれのダークネットの IP アドレスの規模を表 1 に示す．

本観測データに 2 章の提案モデルを適用し，継続的な広域ネットワークスキャンを行う IP アドレス帯を調査した．その際，センサ I から IV の全ダークネット の IP アドレスを D，観測期間 T は月ごとに 12 分割し，送信元 IP アドレス帯は IP アドレスの第 3・第 4 オクテットを無視したクラス B ネットワークを基準とした．また，広域ネットワークスキャン判定時に用いる閾値 α，n 及び継続性評価時に用いる閾値 m は，ヒューリスティクスに基づきそれぞれ $\alpha = 0.8$，$n = 30$，$m = 6$ とした．

3.2 分析結果

観測期間において，継続的に広域ネットワークスキャンを行う 34 のクラス B ネットワークを確認した．

3.2.1 送信元ネットワークの国情報

上記 34 のクラス B ネットワークが属する国を調査し，整理した．その際，1 つのクラス B ネットワークに複数の国が割り当てられている場合が考えられる．そこで，観測期間 T における各クラス B ネットワークから観測された総パケット数のうち，最も多くのパケット数を占めるクラス C ネットワークを特定し，それらに割り当てられている国を当該ネットワークの国情報として扱った．なお，国情報の調査には MaxMind 社の GeoIP2 [15] を用いた．図 1 に国ごとのクラス B ネットワークの数を示す．アメリカ，ロシアのクラス B ネットワークが最も多く，続いて中国，オランダのネットワークが多く観測された．

3.2.2 オンラインドキュメントの有無

文献 [2] は，広域ネットワークスキャンを行う際のガイドラインを提案している．その中の 1 つに，スキャンホストの Web サービスでスキャンの目的や使用する IP アドレス，連絡先などを記載したオンラインドキュメントの公開を推奨している．上記 34 のクラス B ネットワークのうち，

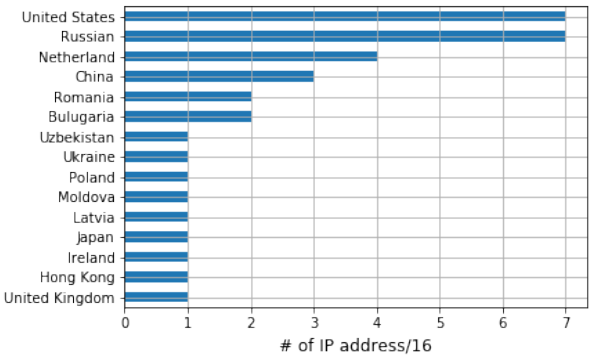

図 1　継続的に広域ネットワークスキャンを行うクラス B ネットワークの国ごとの数

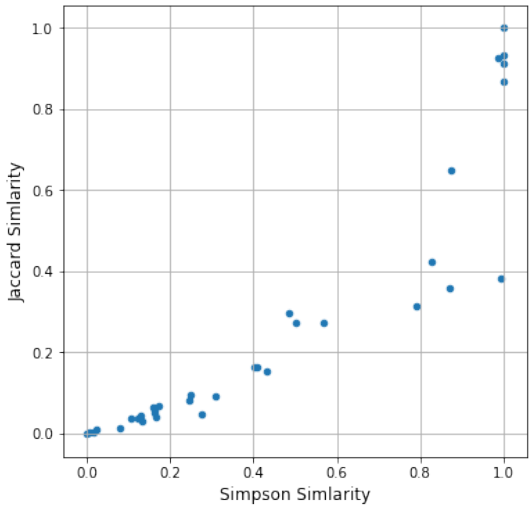

図 2　各クラス B ネットワークのポートセットの類似度

2019 年 8 月 9 日時点で 80/tcp 上でスキャン活動を行う旨を記載したオンラインドキュメントを公開しているホストを調査した．その結果，4 つのクラス B ネットワークでのみ，このようなオンラインドキュメントを公開しているホストが確認できた．

3.2.3 ポートセットの変化

上記 34 のクラス B ネットワークについて，各ネットワークの時系列に対するポートセットの変化を調査するため，クラス B ネットワークごとの各ポートセット $P_{t_i}(x)$ 間の Jaccard 係数と Simpson 係数をそれぞれ求めた．（t_i は $t_i \in T(x)$ をみたす．）Jaccard 係数と Simpson 係数は集合間の類似度を表す指標であり，以下の式で与えられる．

$$\mathrm{Jaccard}(i,j) = \frac{|P_i \cap P_j|}{|P_i \cup P_j|}$$

$$\mathrm{Simpson}(i,j) = \frac{|P_i \cap P_j|}{\min(|P_i|, |P_j|)}$$

ここで i, j は $i, j \in T(x)$ かつ $i \neq j$ である．それぞれの係数の中央値を当該クラス B ネットワークの代表値として用いた．図 2 に縦軸を Jaccard 係数，横軸を Simpson 係数とする各クラス B ネットワークの分布を示す．図より，

各クラス B ネットワークが大きく分けて（1）Jaccard 係数 > 0.8 かつ Simpson 係数 > 0.7 を満たす**ルーティンスキャン**を行うネットワーク，（2）Jaccard 係数 ≤ 0.8 かつ Simpson 係数 > 0.7 を満たす**準ルーティンスキャン**を行うネットワーク，（3）Simpson 係数 ≤ 0.7 を満たす**アドホックスキャン**を行うネットワークの 3 種類の傾向に分類できることがわかる．

図 3, 4, 5 にそれぞれの傾向を示すクラス B ネットワークの観測事例を示す．ルーティンスキャンを行うクラス B ネットワークは，時間経過によらず常にほぼ一定のポートセットに従ってスキャンを行っていることが確認できる（図 3）．また，準ルーティンスキャンを行うクラス B ネットワークは，ポートセットの規模が多少変化するものの，ポート番号の分布は比較的安定しており，時間経過に対するポートセット間の類似性が確認できる（図 4）．一方，アドホックスキャンを行うクラス B ネットワークは，時間経過によってポートセットが大きく変化し，またポート番号の分布が特定の範囲に偏りやすい傾向が確認できる（図 5）．

3.2.4 スキャン対象ポート

スキャン対象になりやすいポート番号を調査するため，ルーティンスキャン及び準ルーティンスキャンを行うクラス B ネットワークのポートセット $P_{t_i}(x)$ において，ポート番号ごとに全ポートセット $P_{t_i}(x)$ あたりに含まれる頻度を計算した．表 2 に頻度が最も高かった上位 10 ポートを示す．表より，Well-known ポートに加え，HTTP，HTTPS の代替ポートやデバイスの遠隔操作に関わるポートがスキャン対象となりやすいことが確認できる．

表 2　スキャン対象ポートの出現頻度

22 (SSH)	95.0%	8080 (HTTP-Alternate)	89.2%
3306 (MySQL)	95.0%	23 (Telnet)	88.2%
80 (HTTP)	91.1%	3389 (Microsoft RDP)	85.2%
21 (FTP)	91.1%	110 (POP3)	83.3%
443 (HTTPS)	89.2%	8443 (HTTPS-Alternate)	83.3%

上記 34 のクラス B ネットワークの概要を表 3 にまとめる．

4. 考察

本章では，ダークネットトラフィックの分析結果およびアプリケーションレイヤへのスキャンについて考察する．

4.1 Mirai ボットネットの影響

観測期間において，スキャンシステム以外に広域ネットワークスキャンを行なっていた可能性が高いエンティティとして，Mirai ボットネットが挙げられる．そこで，上記 34 のクラス B ネットワークごとに，観測された全パケットに含まれる Mirai ボットネットの特徴をもつパケットの割合を調査した．その結果，34 全てのクラス B ネットワークにおいて，Mirai ボットネット特徴をもつパケットの割合が 1%を下回っていることを確認した．本稿の分析結果において，Mirai ボットネットによる広域ネットワークスキャンはほとんど影響しなかったと考えられる．

4.2 ポートセットの変化

3.2 節の分析結果では，上記 34 のクラス B ネットワークが行う継続的な広域ネットワークスキャンが，大きく分けて 3 種類に分類できることを確認した．準ルーティンスキャンを行うクラス B ネットワークでは，図 4 のように，

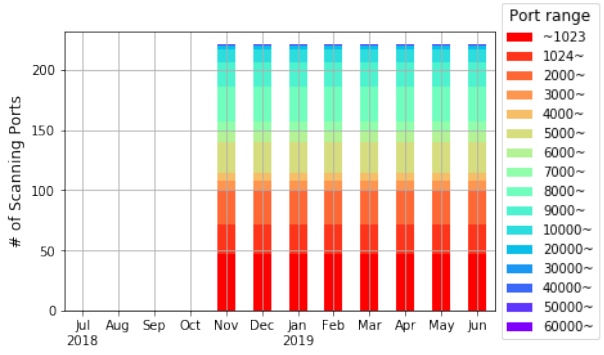

図 3　ルーティンスキャンを行うクラス B ネットワークの観測事例

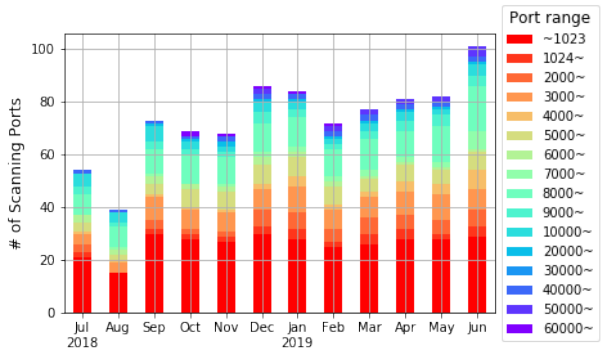

図 4　準ルーティンスキャンを行うクラス B ネットワークの観測事例

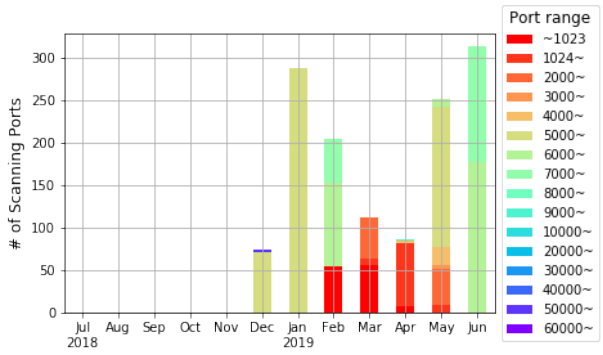

図 5　アドホックスキャンを行うクラス B ネットワークの観測事例

表 3 各クラス B ネットワークの概要．✓ が付いているクラス B ネットワークは，スキャン活動を明記するオンラインドキュメントが確認できたネットワークである．

クラス B ネットワーク	国	観測期間	スキャン対象ポートの選択傾向	単位期間あたりのスキャン対象ポート数の最大値	単位期間あたりのスキャン対象ポート数の最小値
a	IE	2018/07 - 2019/06	アドホック	757	130
b	RU	2018/07 - 2019/02	アドホック	330	50
c（✓）	US	2018/07 - 2019/06	準ルーティン	101	39
d	BG	2018/07 - 2019/02	アドホック	317	65
e	BG	2018/07 - 2019/02	アドホック	1411	227
f（✓）	NL	2018/07 - 2019/06	準ルーティン	330	31
g	PL	2018/09 - 2019/03	アドホック	423	48
h	NL	2018/08 - 2019/06	アドホック	869	40
i	RU	2018/09 - 2019/05	アドホック	222	102
j	RU	2018/07 - 2019/06	アドホック	4043	135
k（✓）	US	2018/11 - 2019/06	ルーティン	221	221
l	US	2018/07 - 2019/06	アドホック	352	130
m	UZ	2018/07 - 2019/02	アドホック	311	40
n	CN	2019/01 - 2019/06	アドホック	297	198
o	CN	2018/07 - 2019/06	準ルーティン	284	35
p	CN	2018/09 - 2019/06	準ルーティン	480	50
q	JP	2018/09 - 2019/06	ルーティン	35	30
r	US	2018/07 - 2019/06	準ルーティン	172	32
s	HK	2018/09 - 2019/06	アドホック	41	33
t	UA	2018/08 - 2019/02	アドホック	10430	182
u	MD	2018/07 - 2019/05	アドホック	497	117
v	RU	2018/12 - 2019/06	アドホック	5745	1759
w	NL	2018/07 - 2019/06	アドホック	896	529
x	LV	2018/12 - 2019/06	アドホック	313	75
y	RU	2018/11 - 2019/04	アドホック	939	237
z	GB	2018/07 - 2019/06	アドホック	271	30
aa	RU	2018/12 - 2019/06	アドホック	1017	252
ab	RU	2018/07 - 2019/02	アドホック	826	66
ac	RO	2018/07 - 2019/06	アドホック	121	31
ad	RO	2019/01 - 2019/06	アドホック	426	111
ae	NL	2018/09 - 2019/06	アドホック	192	37
af	US	2018/07 - 2019/06	ルーティン	97	83
ag（✓）	US	2018/11 - 2019/06	ルーティン	108	84
ah	US	2018/08 - 2019/06	ルーティン	82	43

ポートセットの一部の内容と規模に変化が見られた．この要因として，次のような可能性が考えられる．

(1) ネットワークスキャンサービスを提供する一部の組織などでは，利用者のリクエストに応じて特定のポートやネットワーク帯にスキャンを行う機能を提供している．このような機能の影響によって，特定のネットワーク帯やポートへのスキャンが集中した．

(2) 昨今のソフトウェアやアプリケーションの急速な発展に伴い，日々様々な脆弱性情報が公表されている．ネットワークスキャンを行う者もしくは組織がこうした脆弱性をもつ機器に強い関心をもち，関連するポートへのスキャンが集中した．

(3) 3 章において，IP アドレスの第 3・第 4 オクテットを無視した/16 を基準に分析を行ったが，観測期間中に同じクラス B ネットワークから研究者や攻撃者などが突発的にネットワークスキャンを行なった．

また，アドホックスキャンを行うクラス B ネットワークでは，図 5 のようにポートセットの内容と規模に大きな変化が見られた．さらに，これらのポートセットにおけるスキャン対象ポートの分布などもクラス B ネットワークごとに異なり，よりミクロな視点での分析が必要であると考えられる．

4.3 アプリケーションレイヤへのスキャン

ダークネットで観測された TCP の通信は 3-way ハンドシェイクの SYN パケットしか観測できないため，観測したスキャンが SYN パケットのみを送っているのか，それともセッション確立後アプリケーションレイヤまで踏み

込んだスキャンを行なっているのか，把握できない．そこで，全てのポートにおいて TCP のハンドシェイクを行いセッションの確立を行う簡易なハニーポットを運用・観測することで，3 章で継続的な広域ネットワークスキャンを確認した 34 のクラス B ネットワークがアプリケーションレイヤまで踏み込んだスキャンを行なっているか，調査を行った．具体的には，2019 年 7 月 1 日から 30 日の 1 ヵ月間（便宜上，本観測期間を τ とする），国内のある IP アドレス/24 に属する 31 アドレスで，全てのポートで TCP のハンドシェイクを行いセッションの確立を行う簡易なハニーポットの運用・観測を行った．さらに同期間中，3 章で用いたダークネットセンサで観測を行ない，上記 34 のクラス B ネットワークから広域ネットワークスキャンが行われているか確認した．

観測の結果，期間 τ において，上記 34 のクラス B ネットワークのうち，25 のネットワークで広域ネットワークスキャンが行われていることを確認した．これら 25 のクラス B ネットワークのポートセット $P_\tau(x)$ と，簡易ハニーポットのアクセスログの突合を行い，ポートセット $P_\tau(x)$ においてアプリケーションレイヤまで踏み込んだスキャンが行われているポートの割合 γ（$0 \leq \gamma \leq 1$）を算出した．なお，各ポート p について，簡易ハニーポットを運用する IP アドレスの半数以上である 16 以上の IP アドレスで TCP セッションが確立した場合，当該ネットワークからポート p に対してアプリケーションレイヤまで踏み込んだスキャンが発生したと判断した．

突合の結果，19 のクラス B ネットワークのポートセット $P_\tau(x)$ において，ポートセットに含まれる 10%以上のポートでアプリケーションレイヤまで踏み込んだスキャンが行われていることがわかった．各クラス B ネットワークにおけるアプリケーションレイヤスキャンの割合を表 4 に示す．

4.4 提案モデルの限界

本稿では非公開のスキャンシステムも含めて調査を行うため，スキャンホストのドメイン名やオンラインドキュメントの情報を用いず，IP アドレス帯を基準とする分析を行った．そのため，次のような要因によって実態とは異なる分析結果を得ている可能性が考えられる．

(1) 複数のスキャンシステムが同じ IP アドレス帯を利用していた．
(2) 同じ IP アドレス帯から研究者や攻撃者などが突発的にネットワークスキャンを行っていた．
(3) スキャンシステムが複数の IP アドレス帯を使い分けていた．

(3) は，例えば IP アドレス帯ごとに異なるポートセットを用いてスキャンを行う場合や，IP アドレス帯ごとに異なるネットワークを対象にスキャンを行う場合が挙げられる．そのようなスキャンシステムが存在する場合，提案モデルを正しく適用することができない．

表 4 アプリケーションレイヤまで踏み込んだスキャンの実態

クラス B ネットワーク	スキャン対象ポートの数 $\lvert P_\tau(x) \rvert$	アプリケーションスキャンの割合 γ
a	142	0.7%
c	119	73.1%
f	733	15.8%
h	1892	2.8%
i	47	23.4%
j	77	5.2%
k	221	99.5%
l	271	100.0%
n	287	10.8%
o	283	100.0%
p	389	99.7%
r	218	100.0%
s	34	97.1%
u	149	21.5%
v	4547	8.2%
w	933	0.0%
x	404	33.9%
z	216	29.6%
aa	99	87.9%
ac	121	24.8%
ad	174	0.6%
ae	188	93.6%
af	87	100.0%
ag	92	98.9%
ah	76	22.4%

5. 関連研究

IPv4 空間を高速にスキャンするツールとして，Masscan と Zmap が挙げられる [2], [16]．どちらもオープンソースで公開されており，大規模なネットワークスキャンを可能にしている．

文献 [3] は，ダークネットで観測された通信を分析し，スキャンの送信元や対象ポート，その当時公開された脆弱性情報との関係性など，包括的な調査を行なった．さらに，スキャンを行なっている組織の多くが，大学などの研究機関であることを明らかにし，また近年の大規模スキャンツールが研究者と攻撃者の両者にとって有効なツールとなることを指摘した．しかしながら，近年，非公開の Scan System を含む多様な組織からのスキャンが急増しており，その当時発生していなかったスキャンの存在が考えられる．

文献 [6] は，ライブネットで観測された長期的なトラフィックのデータセットを分析し，近年の大規模スキャンツールの使用状況や，SSH などの代表的なポートについて，年時経過に対する通信量の変化を示した．さらに，短期間における送信元ホストからの宛先 IP アドレスの範囲と宛先ポートの傾向から，発生したスキャンのプロファイリン

グ手法を提案している．しかしながら，分析対象をデータセットの提供元がスキャンと判断したトラフィックに限定しており，分析結果が提供元の判断ありきのものになっていることが問題だと考えられる．

6. まとめと今後の課題

本稿では，継続的な広域ネットワークスキャンを行うスキャンシステムのモデルを仮定し，ダークネットで観測された通信について統計的な観点から分析を行うことで，その実態の調査を試みた．分析の結果，観測期間において，スキャン活動を明示するオンラインドキュメントが確認できない多数のネットワークを含む 34 のクラス B ネットワークから，継続的な広域ネットワークスキャンが行われていることを明らかにした．また，クラス B ネットワークごとのポートセットを分析することで，その特徴と傾向が異なることを示した．さらに，3-way ハンドシェイクを行い TCP のセッションを確立させる簡易なハニーポットを運用・観測することで，複数のクラス B ネットワークからアプリケーションレイヤへのスキャンが行われていることを明らかにした．

今後の課題として，次のことが挙げられる．まず，4.3 節で示したように，多数のクラス B ネットワークからアプリケーションレイヤまで踏み込んだスキャンが確認された．ダークネットの観測データだけでなく，セッション確立後のペイロードを分析することで，アドホックスキャンを行うクラス B ネットワークを中心とするハイポートへのスキャンが，具体的にどのようなサービスを想定して行われているのか，より詳細な分析を行いたい．また，いくつかのスキャンシステムは収集した情報の一部を誰でも閲覧可能な状態で公開しており，公開された情報が攻撃者に悪用される可能性が指摘されている [9]．収集した情報の公開に関わらず，スキャンシステムによって収集された情報が攻撃に悪用される，もしくはそもそも攻撃を前提にスキャンシステムが活動している可能性が考えられるため，スキャンシステムによって収集された情報がどのように利用されているのか，その実態の把握に取り組んでいきたい．

参考文献

[1] D. Inoue, et al.: nicter: An Incident Analysis System Toward Binding Network Monitoring with Malware Analysis. In 2008 WOMBAT Workshop on Information Security Threats Data Collection and Sharing, Sep 2008.

[2] Z. Durumeric, et al.: ZMap: Fast Internet-Wide Scanning and its Security Applications. In Proc. 22nd USENIX Security Symposium, Aug. 2013.

[3] Z. Durumeric, et al.: An Internet-Wide View of Internet-Wide Scanning. In Proc. 23nd USENIX Security Symposium, Aug. 2014.

[4] M. Antonakakis, et al.: Understanding the Mirai Botnet. In Proc. 26th USENIX Security Symposium, Aug. 2017.

[5] Z. Durumeric, et al.: A search engine backed by Internet-wide scanning. In 22nd ACM Conference on Computer and Communications Security, Oct. 2015.

[6] J. Mazel, et al.: Profiling Internet Scanners: Spatiotemporal Structures and Measurement Ethics. In 2017 Network Traffic Measurement and Analysis Conference (TMA).

[7] 牧田 大佑, 他. 全ポート待受型の簡易ハニーポットによるサイバー攻撃観測. 暗号と情報セキュリティシンポジウム 2019.

[8] サイバーセキュリティ研究所サイバーセキュリティ研究室. NICTER 観測レポート 2018. Technical report, 国立研究開発法人情報通信研究機構, 2019.

[9] SHODAN を悪用した攻撃に備えて -制御システム編-. https://www.jpcert.or.jp/ics/report0609.html. 参照 2019 年 8 月 18 日.

[10] Shodan. The search engine for Internet of things. https://www.shodan.io/

[11] Censys. https://censys.io/

[12] ShadowServer. Open Resolver Scanning Project. https://dnsscan.shadowserver.org/

[13] Rapid7 Labs. Project Sonar. https://opendata.rapid7.com/about/

[14] BinaryEdge. https://www.binaryedge.io/

[15] MaxMind GeoIP2 Database & Services. https://www.maxmind.com/en/geoip2-services-and-databases

[16] MASSCAN: Mass IP port scanner. https://github.com/robertdavidgraham/masscan