JIT-Scanning: ダークネット観測と連動した即応的スキャンによる感染機器からのバナー収集

奥川 莞多[1,2,a)]　森 好樹[1]　鈴木 宏栄[1]　久保 正樹[1]　毛利 公一[2]　笠間 貴弘[1]

概要：セキュリティ上の不備がある IoT 機器がマルウェアに感染し，サイバー攻撃の踏み台や大規模な DDoS 攻撃の送信源として悪用されている．これらの被害を軽減するためには，マルウェア感染が確認された IoT 機器の機種を特定し，当該機器への感染経路となった脆弱性を明らかにすることが重要であり，機器で稼働する各種サービスから得られる応答（バナー情報）は機種特定のための有力な情報源となる．一方，インターネット空間に対する広域スキャンを通じて多種多様な IoT 機器のバナー情報を収集・分析する取組が多数の組織で行われており，Shodan や Censys などのサービスから得られるバナー情報が機種特定に活用されている．しかしながら，IoT 機器の感染確認とバナー収集との間に時間的なずれが生じる場合，機器の稼働状況や IP アドレスの付け変わりなどの影響により適切なバナー情報が得られない可能性があり，結果として機種特定や脆弱性特定結果の精度が低下する恐れがある．そこで本研究では，ダークネット観測トラフィックに対してストリーム処理による分析を行い，バナー収集対象となる感染機器をリアルタイムに絞り込み，即応的かつ効果的にバナー収集を行うシステムを開発する．当該システムで得られた結果を Shodan や Censys の検索結果と比較することで，バナー収集タイミングによる影響を明らかにする．

キーワード：インターネットスキャン，バナー収集，ダークネット観測

JIT-Scanning: Banner Grabbing from Infected Devices through Reactive Scanning Triggered by Darknet Monitoring

Kanta OKUGAWA[1,2,a)]　Yoshiki MORI[1]　Koei SUZUKI[1]　Masaki KUBO[1]　Koichi MOURI[2]
Takahiro KASAMA[1]

Abstract: IoT devices with security vulnerabilities are being infected with malware and exploited as stepping stones for cyberattacks or as sources for large-scale DDoS attacks. To mitigate these risks, it is crucial to identify the models of IoT devices infected with malware and clarify the vulnerabilities that enabled the infection. Response information (banner information) obtained from various services running on the devices serves as a valuable source for model identification. On the other hand, many organizations are conducting efforts to collect and analyze banner information from a wide range of IoT devices through large-scale scans of the internet, and banner information obtained from services such as Shodan and Censys is being utilized for device identification. However, when there is a time lag between the confirmation of infection of IoT devices and the collection of banner information, there is a possibility that appropriate banner information cannot be obtained due to factors such as the operational status of the devices or changes in IP addresses, which may result in a decrease in the accuracy of device identification and vulnerability identification results. Therefore, this study develops a system that performs stream processing analysis on darknet observation traffic to real-time narrow down infected devices targeted for banner collection, enabling responsive and effective banner collection. By comparing the results obtained from this system with search results from Shodan and Censys, the study aims to clarify the impact of banner collection timing.

Keywords: Internet-Scanning, Banner Grabbing, Darknet Monitoring

1. はじめに

世界の IoT 機器は 2024 年には 420 億台に上り，コンシュマー向け IoT 機器は 142 億台を超えている [1]．それら IoT 機器の一部は，適切なセキュリティ設定や更新が行われない等の要因により脆弱な状態のまま放置され，マルウェアに感染し，サイバー攻撃の踏み台や DDoS 攻撃の送信源として悪用されている．攻撃被害の拡大を防ぐためには，攻撃に悪用されているマルウェア感染機器及びマルウェアの感染経路となった脆弱性を特定し，適切な対処を行うことが必要である．マルウェア感染機器の特定，特にネットワーク経由で感染を拡大するマルウェアに感染した機器を特定する手法としては，ダークネット観測 [2], [3], [4] やハニーポット観測 [3], [5], [6] などの受動的観測手法が用いられる．これらの手法は，マルウェアによるスキャンや攻撃通信を観測することでマルウェア感染機器を特定することが可能だが，受動的な観測情報のみから感染機器の機種特定まで行うことは容易ではない．そのため，当該機器上で稼働する Web 管理機能等の各種サービスから得られる応答（以下，バナー情報という）を能動的に収集・分析することが機種特定において有効な方法となる [6], [7].

一方，ZMap[8] や Masscan[9] といった高速スキャンツールの登場によって IPv4 空間全体を対象としたスキャンが効率的に実施可能となったことで，インターネット上に存在する多種多様な機器のバナー情報を収集・分析する取組が多数の組織で行われている [10]．また，Shodan[11] や Censys[12] などのバナー情報を検索可能な商用サービスも登場しており，セキュリティエンジニアや研究者は自らスキャンを行う代わりに，これらのサービス経由で得られたバナー情報を機種特定等に活用している [13].

しかしながら，外部の情報源から得られたバナー情報を機種特定に用いる場合，感染機器を特定した時刻とバナー収集時刻との間に一定のずれが生じることになる．その結果として，機器の稼働状況の変化や IP アドレスの付け変わり（IP Churn）などの影響により適切なバナー情報が得られず，機種特定の精度に影響を及ぼす恐れがある．例えば，Censys においては約 100 個の主要なポート・プロトコルについては毎日スキャンを実施しているが，その他のポートについては 9 ヶ月かけてスキャンを行っていると報告されている [13]．すなわち，ネットワーク環境によっては短時間で IP Churn が発生する [14] ことを考慮すると，主要なポート・プロトコルであっても適切なバナー情報が得られない可能性があると言える．

[1] 国立研究開発法人情報通信研究機構
National Institute of Information and Communications Technology
[2] 立命館大学
Ritsumeikan University
[a)] kokugawa@nict.go.jp

そこで本研究では，マルウェア感染機器の効果的な機種特定を目的に，ダークネット観測と連動したスキャンシステム JIT-Scanner を開発した．JIT-Scanner はダークネット観測トラフィックに対してストリーム処理による分析を行い，バナー収集対象となる感染機器をリアルタイムに絞り込み，即応的かつ効率的にバナー収集を行う．提案システムを用いて，日本国内に存在する Mirai（亜種含む）感染機器を対象に，攻撃を観測した時点を起点に 10 分後から 5 日後まで繰り返しバナー収集を行い，収集タイミングによるバナー収集結果への影響を明らかにする．また，当該感染機器の IP アドレスを基に Shodan 及び Censys で検索して得られたバナー情報と提案システムによる収集結果を比較することで，提案システムの有効性を検証する．

本稿の構成は以下の通りである．2 章ではインターネットスキャンにおける IP churn の課題について，3 章では提案する観測手法について述べる．4 章ではスキャンタイミングによる応答実験について説明し，5 章で結果について議論を行う．6 章では関連研究を整理し，7 章で本研究のまとめと今後の展望を述べる．

2. インターネット計測における IP Churn の影響

IP アドレス資源の有効活用や管理コストの低減のため，多くの ISP は DHCP（Dynamic Host Configuration Protocol）や PPP（Point-to-Point Protocol）接続時の IPCP（Internet Protocol Control Protocol）などによる動的な IP アドレス割り当てを採用している [15], [16]．このような動的 IP アドレス環境では，リース期間の満了やルータや端末の再起動による回線の切断・再接続などの影響により，同一機器に対して割り当てられる IP アドレスが変動する IP Churn が発生する．特にモバイル回線や PPPoE（PPP over Ethernet）環境では切り替わりの間隔が短くなる傾向にあり，IP アドレスを基にしたインターネット計測に影響を与える．

2.1 受動的観測における IP churn の影響

未使用 IP アドレスに届くトラフィックを観測するダークネット観測や脆弱な機器を模して攻撃活動を観測するハニーポット観測，ネットワークフローの分析を通じて，インターネット上のマルウェア感染機器数を推定する取組が数多く行われている [17]．このとき，典型的には観測されたユニークな IP アドレス数を感染機器数とみなして扱うことが多いが，IP アドレスベースでの推定では IP Churn の影響により同一機器を複数回カウントしてしまい，結果として規模を過大評価してしまう可能性がある [14], [17], [18].

2.2 能動的観測における IP Churn の影響

一方，能動的にインターネット空間に対してアクセスを

行い，IP アドレスの利用状況や機器のバナー情報を収集するインターネットスキャンにおいても，受動的観測と同様に IP Churn の影響を受ける．特に Shodan や Censys といった IPv4 空間全域を巡回的に走査するスキャナでは，大規模なスキャンインフラを整備してスキャンを実施しているものの，同一の IP アドレスに対するスキャン頻度は主要ポートでも 1 日から 1 週間程度であり，それ以外のポートではさらに低頻度でのスキャン実施となっている [13], [19]．これに対し IP churn はネットワーク環境によっては数時間未満の高頻度で発生するため，例えば，同一 IP アドレスの複数のポートから収集したバナー情報が実際には異なる機器からのバナー情報となるなどの不一致が発生する可能性がある．

結果として，これらの外部サービスからオープンポートやバナー情報を取得して分析に利用する場合，参照時点の機器の状態を正確に反映していない可能性があり，本来は得られるはずのバナー情報が欠落したり，過去に同 IP アドレスを割り当てられていた別機器のバナー情報を誤って取得したりするなど，分析に悪影響を及ぼす可能性がある．この問題に対して，Censys では主要ポートやクラウド環境等に対してスキャン周期を短くしたり，72 時間観測されなかったサービスの収集データを削除するなどして，IP Churn の影響を軽減しようとしているが完全には排除できていない [13].

3. JIT-Scanner: 即応的スキャン方式

本章では，ダークネット観測と連動してマルウェア感染機器に対して即応的にスキャンを行う方式 JIT-Scanner（Just-in-Time Scanner）について述べる．

2 章で述べたとおり，受動的観測によって得られたマルウェア感染機器に対して，巡回型のインターネットスキャンで得られたバナー情報を用いて機種特定を行う場合，IP churn の影響により感染機器の機種特定精度を損なう可能性がある．この課題に対し，本研究が提案する JIT-Scanner は，ダークネット上で検知した感染活動をトリガとして，その時点で稼働し当該 IP アドレスが割り当てられている機器に対して即応的スキャンを実施することで，より正確なバナー収集を可能にする．

NICTER で観測する約 30 万 IP アドレスのダークネットでは，平均して毎秒 2 万パケットを超えるパケットを観測している [10]．JIT-Scanner はこれらのスキャンパケットの送信元の中からバナー収集対象とする機器をリアルタイムに抽出するため，ストリーム処理を用いたフィルタ機能を備えている．ダークネットで観測されたトラフィックをトリガとして，直近の時間窓 N 分に活動した送信元 IP のうちフィルタ条件に合致した対象を即座に抽出し，観測遅延を最小化してスキャンを行うことで，インターネット全域や観測した全送信元アドレスに対してアクセスするなどの不要なトラフィック発生を抑制することができる．

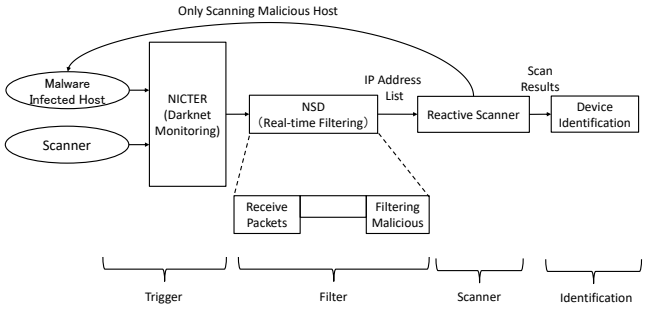

図 1: 即応的スキャンシステム

3.1 システムの構成

図 1 にシステムの全体像を示す．本システムは次の 4 つの機能から構成される．

Trigger　ダークネット観測: ダークネットに到達するトラフィックを観測することでバナー収集対象候補となる送信元 IP アドレスを取得する．

Filter　リアルタイムパケット処理（NSD）: ダークネットトラフィックから調査スキャンやバックスキャッタ，誤送信等の調査対象外の機器を除外し，バナー収集対象 IP アドレスをリアルタイムに抽出する．

Scanner　即応的スキャナ: 抽出した IP アドレスに対して，観測直後に能動的にアクセスを行いバナー情報の収集を行う．

Identification　機器特定・追跡: 取得したバナー情報に対して特徴的な応答（フィンガープリント）を用いた機種特定および追跡を行う．

以下，これらの各機能の役割について詳しく述べる．

3.2 ダークネット観測

本研究では，約 30 万の未使用 IPv4 アドレスを有するダークネット観測システム NICTER[2] を用いる．ダークネット観測では応答を返さないため，ハニーポットとは異なり攻撃ペイロードなどの詳細情報は得られないものの，大規模展開が容易でありネットワーク経由で感染を広めるマルウェアの活動を広範囲に観測するのに適している．

ダークネットに届くトラフィックは概ね次の 4 種類の原因に大別される：

(1) マルウェアによる無差別スキャン活動
(2) セキュリティベンダや研究機関による調査目的のスキャン
(3) DDoS 攻撃の跳ね返りであるバックスキャッタ
(4) 設定ミスや打鍵誤りなどによる誤送信

JIT-Scanner はマルウェア感染機器の機種特定を目的としているため，このうち（1）のマルウェア由来のスキャンに着目するため，後述の NSD において複数のフィルタ

ルールを適用し（2）～（4）を可能な限り除外する．特に新たな攻撃活動や注目すべき特定のマルウェアに限定することで，バナー収集対象となる IP アドレス数は大幅に削減することができる．一例として，IoT マルウェア Mirai に特徴的なスキャントラフィックに限定するフィルタを適用した場合，抽出される IP アドレス数は 2024 年では 1 日あたり約 2.7 万～9 万アドレスであった [10]．これは IPv4 空間全域と比較して非常に少ない数であり，大規模なスキャンインフラを用いずとも高速にスキャンが可能な規模となっている．

3.3 NSD（NICTER Stream Detector）

NSD（NICTER Stream Detector）は，ダークネットトラフィックをリアルタイムに処理し，送信元 IP アドレスに対して国，AS 番号，組織名，DNS 逆引き結果などの属性情報を付加するとともに，パケットの特徴に基づく分析（送信レート，宛先ポートセット，ボット判定，JA4 フィンガープリントなど）を行う．

30 万アドレスのダークネットで観測されるトラフィックは年間で約 6882 億パケット（平均で約 2 万パケット/秒以上）となっているが，これらのパケットにはマルウェアによる無差別スキャン以外にも様々な種類が含まれる．特に近年では，セキュリティベンダや研究機関が実施する調査目的のスキャンが増加しており，ダークネットで観測されるパケットの半数以上を占めるに至っている [20]．これらのパケット全ての送信元 IP アドレスに対して能動的にリアルタイムスキャンを行うことはネットワーク負荷が大きく，不要なトラフィックを発生してしまう．そこで NSD では，付加情報やパケット特徴に基づくフィルタリングを組み合わせることで，調査対象とする攻撃活動やマルウェアに起因するトラフィックのみに絞り込みスキャン対象の IP アドレスを抽出する．この処理により，能動的スキャンの対象を効果的に絞り込み，不要な負荷を回避しつつバナー収集精度を高めることができる．

3.4 即応的なスキャン

NSD から抽出された IP アドレスに対して，ダークネット観測でトラフィックが観測された時刻からの遅延を最小限に抑えてスキャンを行う．受動的観測をトリガにして即応的にスキャンを行うことで，送信元機器の IP アドレスが付け変わる前にバナー収集を行うことができ，IP churn の課題を解決する．

NSD からスキャン対象アドレスを抽出する際には，直近 N 分以内に観測したフィルタ条件に合致する送信元 IP アドレスリストのように，受動的観測の時刻と能動的スキャンの時間間隔を適切に調整することで，スキャン処理のオーバヘッドを考慮しつつ効率的なスキャンを実地する．

また，スキャン対象ポートについてもスキャン負荷を考慮し，全ポートを対象としたスキャンは行わず，基本的な Web アクセスのポートに加えて，日次で更新される動的リスト（NICTER への宛先ポートセット）や既知 IoT 機器の開放ポート，受動的観測で見られたスキャンポートセットなどを用いて，スキャン対象を効率的に選択する．

3.5 機種特定・追跡

収集したバナー情報を基に機種を特定するための特徴的な応答（フィンガープリント）を用いて機種特定を行う．具体的なフィンガープリントは例えば以下に示すような情報から他の応答とは異なる値の自動的抽出や，解析者による実機調査を通じて生成する．

- TCP/IP ヘッダ: TCP のオプションやウィンドウサイズなどヘッダ情報
- アプリケーションバナー: HTTP ヘッダやボディ，TLS 証明書，SSH 鍵のフィンガープリントなど

これらのフィンガープリントを拡張していくことでマルウェア感染している機種の特定が可能になる．加えて，当該機種を調査することで悪用されている脆弱性を特定したり，実機をハニーポットとして用いて攻撃活動の詳細を調査するといった活動につなげることができる．また，正確な機種特定が実現すれば，例えば IP Churn が発生したと推測されるタイミング前後の近傍 IP アドレスの機器から得られたバナー情報を基に，IP アドレスの変更先を特定し機器の追跡が行える可能性が高まる．

4. 評価実験

本章では，3 章で提案した即応的なスキャンの有効性を検証する．具体的には，ダークネットで観測された感染活動から能動的スキャン実施までの時間経過とスキャン応答率の影響を分析した上で，大規模調査スキャナによるスキャン結果と提案手法によるスキャン結果を比較することで本手法の有効性を示す．

4.1 実験方法

本実験では，多様な感染機器が存在することが想定される IoT マルウェアの感染機器に焦点を当て，以下のフィルタリング条件に当てはまる送信元 IP アドレスを抽出した．

(1) TCP SYN パケット: TCP Flag が SYN=1，ACK=0 であること
(2) Mirai のスキャン特徴: 初期シーケンス番号が送信元 IP の値に一致
(3) 送信元地域が日本国内: GeoIP Database により判定

それぞれの条件は，（1）はバックスキャッタなどを除外し TCP によるスキャン活動を行う感染機器に限定するため，（2）は IoT マルウェアの代表例である Mirai（亜種含む）の感染機器に限定するため，（3）はアクセス範囲を制限するためのフィルタリングである．

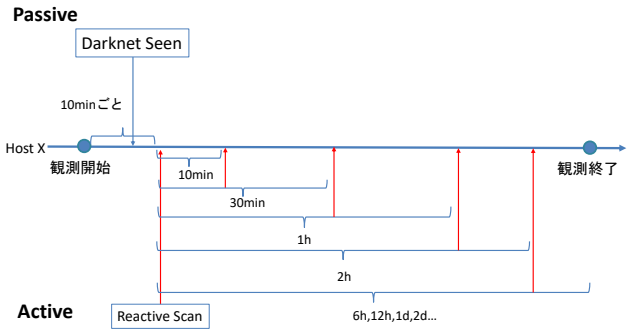

図 2: 実験の観測フロー

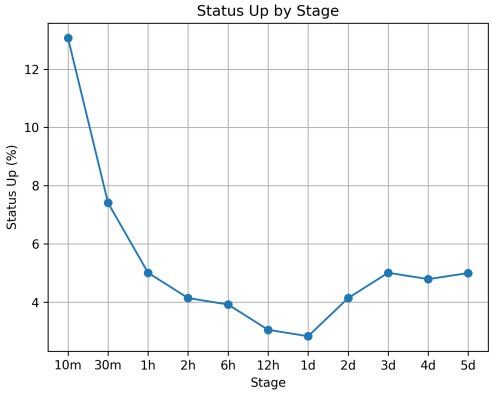

図 3: 攻撃観測からバナー収集までの経過時刻によるアクセス成功率

図 2 に実験の観測フローを示す．ダークネット観測から 10 分ごとに更新される直近の感染活動を起点として，その後 {0-10 分，10 分〜30 分，30 分〜1 時間，1〜2 時間，2〜6 時間，6〜12 時間，12 時間〜1 日，1〜2 日，2〜3 日，3〜4 日，4〜5 日 } の範囲中の各タイミングで nmap[21] を用いたスキャンとバナー収集を行うことで，短期から長期にわたる応答性の変動を測定する．ただし，継続したスキャンが観測される送信元 IP アドレスについては，ダークネットで観測される度に直近 10 分から再度計測をはじめることとした．スキャンを行う対象のポートは典型的に Web アクセスで使われる {80 8080 443 8443} に限定し，取得する情報はポートの開放状態と HTTP バナー情報のみとした．

4.2 実験結果

実験期間（2025 年 8 月 14 日〜20 日）において，観測されたユニークなホストは 1,586 IP アドレスであった．観測対象ホストのうち，HTTP 系のポート（80,8080,443,8443）からバナーが収集できたホスト数はそれぞれ順に 186 件（11.7%），9 件（0.6%），46 件（2.9%），4 件（0.3%）であった．Web アクセスが可能なユニークなホスト数は 203 件（全体の約 12.8%）であった．

一方，これらの IP アドレスについて実験期間中のダークネットでの観測状況の推移に着目すると，継続して 4 日以上の感染活動が観測されなかったホスト（グループ 1 とする）が 459 IP アドレス，継続して観測されたホスト（グループ 2 とする）が 1,127 IP アドレスであった．

まず，IP chrun が発生した可能性が高いグループ 1 について分析を行った．**図 3** はグループ 1 における最後の感染活動からスキャンの遅延時間ごとの応答率を示したグラフである．グラフを見ると観測直後に実施したスキャンの応答率は 13.1%であったが，応答率は数時間以内で顕著に低下し，1 日後まで時間経過とともに応答率が減少している．一方で 2 日以降になると応答率は情報傾向を示しているが，これは IP Churn により当該 IP アドレスが別の機器に割り当てられたことによって観測直後とは異なる機器からの応答が収集されていると考えられる．また，**図 4** はポートごとの割合変動を示している．時間が経つにつれ

図 4: スキャンタイミングごとのバナー収集成功率

てポート 80 へのアクセス成功数がある程度低下していく一方，24 時間経過したあとは 8080 や 443 では一部で応答率が増加しており，特に 8080 は観測直後には応答率が 0%だったことからも，当該 IP アドレスが別の機器に再度割り当てられた可能性を強く示唆している．さらに，**表 1** にグループ 1 におけるポート 80 の状態遷移の頻度を示す．特に，10 分〜1 時間以内に `open/filtered/closed` →応答なし へ移行するケースが多く，感染活動から短時間での観測がバナー収集に不可欠であることがわかる．一方で 2 日以降に 応答なし→ `open` となるケースが存在し，接続切れや再起動による動的 IP 割当や機器変化が影響している可能性がある．このように，グループ 1 のような感染活動が継続的に見られない機器については即応的スキャンが効果的であると言える．

一方で，持続的な感染活動を行っているグループ 2 についてのみで応答率を計算した場合，直近 10 分では 25%を超えていた．これは，グループ 2 のホストが安定的に稼働していることを反映しており，感染活動をトリガにスキャンを行うことで，高い確率でその IP アドレスに到達できることを示している．

表 1: ポート 80 におけるステージごとの変化率とその数
o=open, f=filtered, c=closed, d=応答なしとする

期間	変化率	o↓f	o↓c	o↓d	f↓o	f↓c	f↓d	c↓o	c↓f	c↓d	d↓o	d↓f	d↓c
10m→30m	8.28			9	1	1	7			15	1	1	3
30m→1h	3.92	1		5			5			4		1	2
1h→2h	1.74			1			3			2	2		
2h→6h	0.22						1						
6h→12h	1.31			2						3			1
12h→1d	0.22									1			
1d→2d	1.31												6
2d→3d	1.74									2	1	1	4
3d→4d	0.65									2			1
4d→5d	0.22												1

4.3 大規模調査スキャナとの比較

次に，バナー収集タイミングによる影響を明らかにするために，大規模調査スキャナ（Shodan, Censys）で得られる検索結果と提案手法による収集結果の比較を行った．観測終了時点で両サービスの API を用いて，各 IP アドレスに対する最新のデータを取得し，本手法で観測した 1,586 IP のうちポート {80 8080 443 8443} でバナーを収集できたホストと照合した．**表 2** にそれぞれのポート数と重複を示す．本手法ではポート 80 で 186 件の応答を確認できていたのに対し，Censys で 98 件，Shodan で 97 件という結果であった．すなわち，本手法は大規模調査スキャナの約 2 倍のバナー情報を収集でき，短期的に変化するホストの把握に優れていることがわかる．また，Censys で得られたホストは本手法で得られたホストにすべて含まれており，Shodan については 20 件は本手法ではバナーを取得できなかったホストであったが，最終更新日時が 1 週間以上前の古い情報が多かった．それ以外の Web アクセス（8080, 443, 8443）についてもそれぞれの収集結果は Censys で 14，29，1 件，Shodan で 16，25，0 件となり，8080 以外については我々の手法がアクセス成功率として優れている．Censys と Shodan いずれの重複以外の IP アドレスのバナー収集結果については 本実験よりも古い結果が提供されているものが多かった．

さらに，提案手法では観測したタイミングで応答が存在したバナー情報を取得できた一方，Shodan や Censys では過去のスキャン結果が保持されており，実際の状態との乖離があることを確認した．Censys と Shodan で取得できていなかったバナーのうち本手法でバナーを取得できたホストは，直近の 10 分から 30 分程度のみ open であったホストが多かった．さらに，特定の IP アドレスでは，古いデータが提供されており，我々が取得した結果と異なるものがあった．これは一度スキャンしたのち，再度スキャンされるまで長期間空いている可能性があり，受動的観測をトリガとすることでスキャンタイミングが向上されている

表 2: 各ポートにおけるスキャン成功数の比較
JIT=JIT-Scanner, C=Censys，S=Shodan とする

Port	JIT	C	S	JIT∩C	JIT∩S
80	186	98	97	98	77
8080	9	14	16	6	5
443	46	29	25	27	17
8443	4	1	0	0	0

と言える．一方，本手法では応答なしであったが，Censys と Shodan では取得したバナーが記録されているホストが一部存在したが，このようなホストについては今後の課題とする．

以上の結果から，大規模スキャナと比較して本提案手法は，(i) 応答率が高い，(ii) バナー情報が常に新鮮なデータである，という 2 点で優位性を持つことが示した．

4.4 ケーススタディ

最後に，GeoIP データベースからそれぞれのホストに AS や回線情報を付与し，通信環境の特性を考慮した分析を行った．プロバイダの実装に依存するため，正確にどのタイミングで IP churn が発生するかを決定づけることは困難であるが，本実験において IP churn が発生する際に見られる複数の事例を観測した．

事例 1: 特定の回線における別ユーザーの機器への再割り当て

PPPoE 接続を提供する特定の法人向け ISP では，接続直後は応答がなかったが，数日後に応答ありに変化するケースが多くみられた．観測中にマルウェア感染端末に割り当てられていた IP アドレスが解放されたのち，すぐに応答を返す別のユーザの機器へ再割り当てとなった可能性がある．

事例 2: 観測期間中感染活動を続けていたホスト

特定の ISP に関わらず，多数のホストで観測期間中ダークネットで長期間感染活動を行っていたホストについては，比較的安定してバナーを収集可能であった．また，一部のホストで数時間から数日間で継続的にダークネットで観測されていたが，感染活動が観測されなくなった途端，応答がなくなった事例があった．これは，特定の機器がセッション切断や電源オフなどの原因でホストが一時接続不可となったホストのうち，IP アドレスを短期間で使いまわさないもしくは，リース期間があるため再割り当てとなっていない環境に接続されている可能性がある．

以上の事例から，IP churn の発生頻度や影響は回線種別や AS に依存することがわかる．特に，日本国内において一部 ISP 環境では，なんらかの理由で接続切れが発生した直後に別ユーザの機器へ再割り当てが発生する IP churn が観測されており，従来の定期スキャンでは正確なホスト追跡が難しいことがわかる．また，持続的な感染活動を

行っているホストについては，比較的安定してバナーの収集が可能となっているが，感染活動の観測ができなくなってからは順次応答が少なくなっている．これらの結果は，提案手法のように「その時点でアクティブな機器」を対象に即応的にスキャンする方式が，IP churn 環境における正確な感染機器追跡に有効であることを示している．

5. 議論

本章では，前章の評価実験で得られた結果を踏まえ，提案手法 JIT-Scanner の有効性と制限について議論する．5.1 節で実験結果から得られた観測特性を整理し，5.2 節で本手法の課題や制約について述べる．

5.1 実験結果の考察

実験では，提案手法による即応的スキャンが，大規模調査スキャナ（Shodan, Censys）ではバナー情報が得られなかった IP churn が発生した可能性が高いホストを観測できることを示した．

提案手法でスキャンに成功したが，Shodan や Censys では応答がないものについては，IP churn に起因する可能性が高い．ケーススタディで確認されたように，ホストの状態が観測の途中で変化する事例や，2 日目以降に新たに応答が得られる事例では，同一 IP アドレスが別ユーザの機器に再割当された可能性がある．このような短期的な変動を考慮するには，受動的観測に連動した即応性があるスキャンで得た結果をもとに感染機器を分析することが有効である．Shodan や Censys は 24 時間以内にスキャンを行うポートやアドレスレンジが存在する [13], [19] が，今回発見した一時的に出現するようなホストに対しては観測できていない可能性がある．JIT-Scanner においてもスキャンの頻度やキャッシュを調整するためには，それぞれの ISP ごとの通信特性を意識した分析が必要であり，これについては今後の課題とする．

一方で，提案手法では応答がないが，Shodan と Censys では Web ポート以外のスキャン結果が存在することも考慮すべきである．今回の実験では Web ポートに絞ってスキャンを実地したが，Shodan や Censys などの大規模調査スキャナはより広範囲なポートセットを対象としており，ポート開閉に関する情報量という点では劣る．また，IoT 機器は well-known ポート以外の固有のハイポートが使用されている事例も多くある．そのため，マルウェアがスキャンしたポートセットがその機器の感染した原因である仮定し，そのポートに対してスキャンする方式や，直近の大規模な調査スキャナの結果からスキャンするポートを決める方式など検討の余地がある．しかし，機器特定という観点からはアプリケーションバナーや Web 画面の情報が必要となってくるため，いずれにしてもスキャン対象ポートをどのような基準で設計するかについては議論が必要である．

5.2 本手法の課題と制限

提案手法にはいくつかの制約が存在する．能動的観測においてもフルポートスキャンを行わない場合，応答率は 3 割程度にとどまる [7]．また，IP アドレス変動を考慮した場合で考えると，能動的観測と受動的観測で 3 日間ほど期間が開いた場合，応答があったホストが正確かどうかは確認することはできない．そのため，本研究では，直近に動作したということに基づいて観測することには一定の効果が得られている．一方で，マルウェアが動作するホストはポートの開閉もする場合が多く，それ起因して今回の調査では観測率自体は落ちている可能性があるが，マルウェアがポートを開閉するタイミングについては別途検討が必要である．

また，最近ではスキャン活動をせず，C2 サーバからの命令を受ける際のみに通信を行うような，サイレントなマルウェアも存在する．そのような感染端末は従来のスキャンでは把握しづらいため，そもそもダークネットでは観測できず，感染ホストの特定はできない．そのような端末を捉えるためにはフローデータ分析や C2 通信監視との連携が不可欠である．

5.3 研究倫理

本研究では，広域な IP アドレス空間をスキャンせず，マルウェアに感染した機器のみに制限しており，一般的な Web アクセスのポートのみに限定してスキャンを実地している．また，既存研究でも使用されているツールを使用しており，不正アクセスになり得る通信は行っていない．適切なスキャンレートに制限しており，パケット量も大きくならず，最短でも 1 ホストにつき 10 分に 1 回のみのスキャンであるため増幅はしない設計となっている．マルウェアに感染する機器を特定できるような情報は公開せず，特定の個人の紐づけも行っていない．

6. 関連研究

インターネットにおける脅威観測や感染機器の特定に向けて，これまでにもさまざまな観測手法が提案されてきた．

文献 [18] は受動的観測に基づき，Mirai に感染したホストが IP churn によりどのように変動するかを明らかにしている．しかし，この手法ではパッシブデータを参照しているため個別の感染機器を直接的に特定することは難しい．また，受動的観測と能動的観測を組み合わせた研究が存在する．文献 [7],[6] では，受動的観測のデータを元に能動的観測を行っているが，スキャン方法やタイミングについては述べられていない．

Censys の研究結果では [13]，ML-based なスキャンポート選定の効率化などが考慮されており，フルポートスキャ

ンではない特定の局所性を持たせたスキャンが検討されている．また，スキャン地点による依存性を明らかにした調査では [22]，スキャナの地理的・ネットワーク的な位置によってスキャン成功率に与える影響を調査している．文献 [23] では，あるホストの生存性をネットワークのレイヤごとに接続した際の動作によって確認する手法が提案されいる．以上のように，いずれの既存研究においてもどのようにスキャンするかといった観測対象の選択や観測方法に重点を置いており，スキャンのタイミングが観測結果に与える影響については十分に検討されていないこれに対し，提案する JIT-scanner は，フルスキャンに比べて効率的かつ，ハニーポットよりも大規模な観測を可能にする．特に「いつスキャンするか」という時間的要素を取り入れることで，IP churn が激しい IoT 機器を含む環境において，効果的なバナー収集を実現している．

7. おわりに

本研究では，マルウェア感染機器の効率的なバナー収集を目的として，ダークネット観測と連動した即応的スキャンシステム JIT-Scanner を提案した．JIT-Scanner はダークネット観測で検出した感染活動をトリガとして，観測直後に対象ホストに対して能動的にバナー収集を行うことで， IP Churn 環境における従来手法では見逃しえるホストの観測が可能となる．日本国内に存在する Mirai 感染機器を対象とした実験により，観測直後のスキャンで高い応答率を得られること，また時間経過とともに応答率が低下することを明らかにした．さらに，Shodan や Censys と比較した結果，提案手法は大規模調査スキャナではバナー収集が困難なホストを特定でき，新鮮なバナー情報を取得できる点で優位性を持つことを確認した．

今後の方針として，観測規模の拡大とスキャン対象ポートについての議論を行う．日本国内の Mirai の特徴を持つスキャンホストに制限して観測したため，ユニークホスト数が千件程度であり，より多くのバナーを収集するためには世界規模で観測を行う必要がある．各国の異なる ISP やネットワーク環境における IP Churn の特性を比較分析することで，今後は観測規模の拡大を検討する．また，どのようなポートをスキャンの対象とするかを検討し，より効率的なバナー収集を目指す．さらに，収集したバナーから機器を特定するためのフィンガープリンティングの方法についても検討していく．

参考文献

[1] 総務省：情報通信白書令和 7 年版 データ集，`https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/r07/html/datashu.html` (2025).

[2] Inoue, D. and et al.: An incident analysis system NICTER and its analysis engines based on data mining techniques, *The 15th International Conference on Advances in Neuro-Information Processing*, p. 579–586 (2008).

[3] Antonakakis, M. and et al.: Understanding the Mirai Botnet, *The 26th USENIX Security Symposium*, pp. 1093–1110 (2017).

[4] Griffioen, H. and Doerr, C.: Examining Mirai's Battle over the Internet of Things, *The 2020 ACM SIGSAC Conference on Computer and Communications Security*, pp. 743–756 (2020).

[5] Pa Pa, Y. M. and et al.: IoTPOT: Analysing the Rise of IoT Compromises, *The 9th USENIX Workshop on Offensive Technologies* (2015).

[6] 森 博志， 他：能動的観測と受動的観測による IoT 機器のセキュリティ状況の把握，研究報告コンピュータセキュリティ（CSEC），Vol. 76, No. 27, 情報処理学会，pp. 1–6 (2017).

[7] 笠間 貴弘，井上 大介：大規模ダークネット観測と能動的スキャンによるマルウェア感染 IoT 機器の分類，情報処理学会論文誌， Vol. 58, pp. 1388–1398 (2017).

[8] The ZMap Project: ZMap: The Internet Scanner, `https://github.com/zmap/zmap` (2025).

[9] robertdavidgraham: MASSCAN: Mass IP port scanner, `https://github.com/robertdavidgraham/masscan` (2025).

[10] 国立研究開発法人情報通信研究機構：NICTER 観測レポート 2024 (2024).

[11] Shodan: Shodan Search Engine, `https://www.shodan.io/` (2025).

[12] Censys: Censys |The Authority for Internet Intelligence and Insights, `https://censys.io/` (2025).

[13] Durumeric, Z. and et al.: Censys: A Map of Internet Hosts and Services, *The ACM SIGCOMM 2025 Conference*, pp. 1–17 (2025).

[14] Moura, G. C. M. and et al.: How dynamic is the ISPs address space? Towards internet-wide DHCP churn estimation, *The 2015 IFIP Networking Conference*, pp. 1–9 (2015).

[15] Padmanabhan, R. and et al.: Reasons Dynamic Addresses Change, *The 2016 Internet Measurement Conference*, p. 183–198 (2016).

[16] Richter, P. and et al.: Beyond Counting: New Perspectives on the Active IPv4 Address Space, *The 2016 Internet Measurement Conference*, p. 135–149 (2016).

[17] Böck, L. and et al.: How to Count Bots in Longitudinal Datasets of IP Addresses, *The Network and Distributed System Security Symposium*, pp. 1–16 (2023).

[18] Griffioen, H. and Doerr, C.: Quantifying autonomous system IP churn using attack traffic of botnets, *The 15th International Conference on Availability, Reliability and Security*, pp. 1–10 (2020).

[19] Bennett, C. and van Oorschot, P. C.: Empirical scanning analysis of Censys and Shodan, *The Measurements, Attacks, and Defenses for the Web Workshop 2021* (2021).

[20] Kasama, T. and et al.: Please Stop Knocking on My Door: An Empirical Study on Opt-Out of Internet-Wide Scanning, *IEEE Access*, Vol. 13, pp. 48416–48430 (2025).

[21] Lyon, G.: Nmap: the Network Mapper - Free Security Scanner, `https://nmap.org` (accessed 2025-08-22).

[22] Wan, G. and et al.: On the Origin of Scanning: The Impact of Location on Internet-Wide Scans, *The ACM Internet Measurement Conference*, p. 662–679 (2020).

[23] Bano, S. and et al.: Scanning the Internet for Liveness, *SIGCOMM Comput. Commun. Rev.*, Vol. 48, No. 2, p. 2–9 (2018).