一般社団法人　電子情報通信学会
THE INSTITUTE OF ELECTRONICS,
INFORMATION AND COMMUNICATION ENGINEERS

信学技報
IEICE Technical Report

Telegramにおける初期アクセスブローカー模倣による攻撃観測システムへの攻撃者誘引の実証実験

松村　尚典† 　青砥　　陸† 　インミンパパ†† 　田辺　瑠偉†† 　吉岡　克成†††,††

† 横浜国立大学大学院環境情報学府
†† 横浜国立大学先端科学高等研究院
††† 横浜国立大学環境情報研究院
E-mail: †{matsumura-takanori-tx,aoto-riku-hx}@ynu.jp, ††{yinminn-papa-jp,tanabe-rui-xj,yoshioka}@ynu.ac.jp

あらまし　近年，SNS 上で VPN や RDP などの認証情報が頻繁に取引されており，企業や組織ネットワークへの侵入手段として悪用されている．しかし，SNS 上での認証情報の流通経路やそれに基づくサイバー攻撃の実態は十分に解明されていない. 本研究では，このような認証情報を他者に提供する初期アクセスブローカー（Initial Access Broker: IAB）の SNS 上での活動と IAB に起因するサイバー攻撃を観測・分析する．具体的には架空の囮 IAB が認証情報の販売を行うための囮 Telegram チャンネルを用意し，攻撃観測システムへアクセスするための認証情報を掲載する．さらに，当該チャンネルへの参加を促すため，VPN や RDP などのアクセス情報に興味のある参加者が集まる実在の Telegram グループに対して囮チャンネルの宣伝を投稿する．2025 年 2 月〜2025 年 6 月の期間で 23 個の Telegram グループに対して合計 1,063 件の宣伝投稿を行った結果，囮チャンネルに対して 15 ユーザの登録があり，囮チャンネルに掲載された認証情報を用いた侵入が攻撃観測システムに対して 8 個の IP アドレスから行われた．侵入後には外部への通信の可否を確認する DNS 通信が観測されたが，攻撃観測システムは外部ホストへセッションを確立する通信を許可しておらず，その後の挙動は観測されなかった．また，前述の宣伝の投稿に際して，一部の Telegram グループでは管理者による投稿内容の削除や制限といった反応が観測された．

Experimental Demonstration of Attracting Attackers to Attack Observation Systems by Imitating Initial Access Brokers on Telegram

Takanori MATSUMURA†, Riku AOTO†, Yin Minn Pa Pa††, Rui TANABE††, and Katsunari YOSHIOKA†††,††

† Graduate School of Environment and Information Sciences, Yokohama National University
†† Institute of Advanced Sciences, Yokohama National University
††† Faculty of Environment and Information Sciences, Yokohama National University
E-mail: †{matsumura-takanori-tx,aoto-riku-hx}@ynu.jp, ††{yinminn-papa-jp,tanabe-rui-xj,yoshioka}@ynu.ac.jp

Abstract　In recent years, credentials such as VPN and RDP have been frequently traded on social media and misused as entry points into corporate and organizational networks. However, the distribution channels of these credentials and the actual state of related cyberattacks remain insufficiently understood. This study observes and analyzes the activities of initial access brokers (IABs) who share such credentials on social media, as well as the resulting cyberattacks. To investigate, a fake IAB was created, operating a decoy Telegram channel that posted credentials granting access to an attack observation system. To attract participants, we promoted the channel in real Telegram groups where users interested in VPN and RDP access gather. Between February and June 2025, 1,063 ads were posted in 23 Telegram groups. As a result, 15 users subscribed to the decoy channel, and intrusions using the posted credentials were observed from 8 IP addresses targeting the observation system. Following these intrusions, DNS communications were detected, suggesting attempts to reach external hosts. However, the system blocked such connections, and no further activity occurred. Some Telegram groups also reacted to the posts by deleting or restricting them.

1. はじめに

近年，企業や組織を標的としたサイバー攻撃において顕著な分業化が進展している．特に，初期アクセスブローカー（Initial Access Broker: IAB）と呼ばれる攻撃者が，VPN や RDP などのリモートアクセス認証情報を Telegram をはじめとする SNS 上で販売・仲介し，これらを購入した別の攻撃者が実際のサイバー犯罪を遂行するエコシステムが形成されている [1]．新型コロナウイルス感染症（COVID-19）のパンデミック以降，リモートワークが急速に普及したことに伴い，VPN および RDP を標的としたサイバー攻撃が顕著に増加した．また，これらがランサムウェア攻撃の主要な侵入経路となっていることが複数の調査により報告されている [2]〜[4].

一連の攻撃の実態を把握するため，認証情報の流通過程に着目した研究が行われている．先行研究 [5], [6] においては，オンラインテキスト共有サービスおよびダークウェブ上に囮の認証情報を意図的に公開し，攻撃者による不正アクセスの挙動を観測・分析することに成功している．しかし，これらの先行研究では，Telegram をはじめとする SNS 上での IAB の活動を模擬した攻撃観測は実施していない.

そこで本研究では，IAB の Telegram 上での活動を簡易的に模擬し，事前に構築した攻撃観測システムへ攻撃者を誘引する実験を実施する．観測実験では，RDP を介したリモートアクセスが可能な Windows 仮想マシンを含む攻撃観測システムを構築した．さらに，認証情報の販売を装った専用 Telegram チャンネル（以下，囮チャンネル）を運営する囮 IAB アカウントを作成し，当該チャンネルにおいて攻撃観測システムへのリモートアクセス情報（デスクトップスクリーンショット，認証情報等）を掲載した．加えて，VPN および RDP などのアクセス情報を取り扱う実在の Telegram グループに対し，当該囮チャンネルへの誘導を目的とした宣伝投稿を実施した.

2025 年 2 月から 2025 年 6 月までの実験期間において，23 個の Telegram グループに対して合計 1,063 件の宣伝投稿を実施した．その結果，囮チャンネルへ 15 ユーザが登録し，掲載した RDP 認証情報を用いた不正アクセスが 8 個の異なる IP アドレスから観測された．侵入後の挙動として，外部ネットワークへの接続性を確認するための DNS 通信等が観測されたが，攻撃観測システムは外部ホストとのセッション確立を禁止していたため，それ以降の攻撃活動は観測されなかった．また，宣伝投稿を実施した 23 個の Telegram グループのうち，12 個のグループにおいて管理者による投稿の削除や投稿権限の剥奪などの制裁措置が確認された．これらの措置は，グループ管理者が不適切な宣伝活動を排除し，グループ内の秩序維持を図るために実施したものと推察される．以上のように，観測実験では囮 IAB チャンネルの作成および宣伝投稿により，攻撃者を攻撃観測システムへ誘引することに成功した．また，攻撃者の挙動や Telegram グループの一部実態明らかになった一方で，外部接続制御の設定や宣伝投稿の手法については，今後さらなる検討が必要であることが分かった.

本論文の貢献は以下の通りである：

- Telegram 上に用意した IAB の活動を模倣した囮チャンネルにより，攻撃観測システムへ攻撃者を誘引する実験を初めて実施した.
- 実在する Telegram グループへの宣伝投稿を通じて囮チャンネルへ攻撃者を誘導し，攻撃観測システムへの RDP 経由での不正侵入を誘引することに成功した.
- Telegram グループへの広告に対する管理者による投稿の削除・制限など，不正活動の実態の一部を明らかにした.

2. 関連研究

文献 [6] では，表層ウェブとダークウェブに対して Gmail アカウントの認証情報を 100 個ずつ配布し，両環境における攻撃者の活動パターンの差異を比較分析した．表層ウェブに配布された認証情報については単発的なアクセスが大半を占めたのに対し，ダークウェブにおいては同一攻撃者による反復的なアクセスや長期的な監視・追跡活動が観測された.

文献 [7] では，国立研究開発法人情報通信研究機構（NICT）が開発した企業ネットワークを模倣したハニーポットである STARDUST を提案している．STARDUST は，企業を標的としたサイバー攻撃の詳細な挙動解析を目的として設計されている．攻撃者の行動観測という基本的な目的は従来のハニーポットと共通するものの，STARDUST は特に二つの要素を重視した設計となっている．第一に，実ネットワークを精巧に模倣した並行ネットワークを使用していること，第二に，外部ネットワークへの影響を防ぐために IDS/IPS を用いることにより悪意のある通信を遮断している．これらの技術的特徴により，攻撃者の活動を大きく制限することなく，かつ安全に観測することが可能となる．その結果，従来のハニーポットでは断片的にしか把握できなかった侵入後の攻撃者の振る舞いを包括的に観測できる．本研究についても今後は STARDUST と連携することで，より高度な標的型攻撃の観測への応用が期待される.

文献 [8] では，RDP ハニーポットを用いて攻撃者の侵入経路を体系的に調査した．同研究では，5 種類の異なる侵入経路について詳細な分析を実施した．第一の経路として，ブルートフォース攻撃による認証情報の総当たり突破が確認された．この攻撃手法では，ユーザ名とパスワードの組み合わせを網羅的に試行することで認証システムの突破を試みる．実験では意図的に脆弱な認証情報（ユーザ名：admin，パスワード：password）を設定し，攻撃者による侵入を誘引した．その他の侵入経路として，Pastebin [9]，GitHub Gist [10]，Google Docs [11]，および VirusTotal [12] の各プラットフォームに認証情報を意図的に公開し，攻撃者による発見と侵入を観測した．比較検証のため，ブルートフォース攻撃では突破不可能な複雑な認証情報も同時に設定した．4 つのプラットフォームのうち，Pastebin においては組織外部者による RDP 情報の意図的な配布を模擬し，その他の 3 つのサービスでは偶発的な情報流出を偽装した．しかしながら，同研究では Telegram を用いた認証情報の

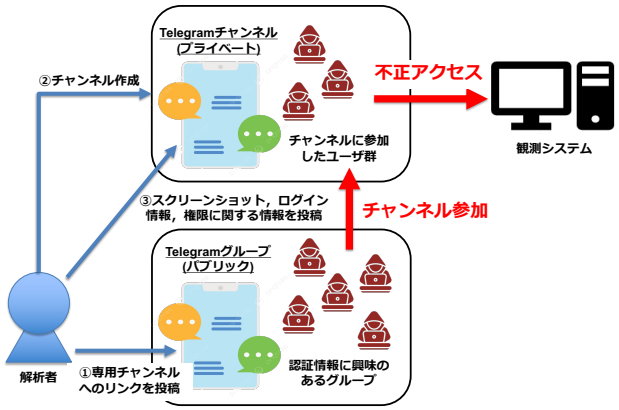


図 1: 本実験の全体像

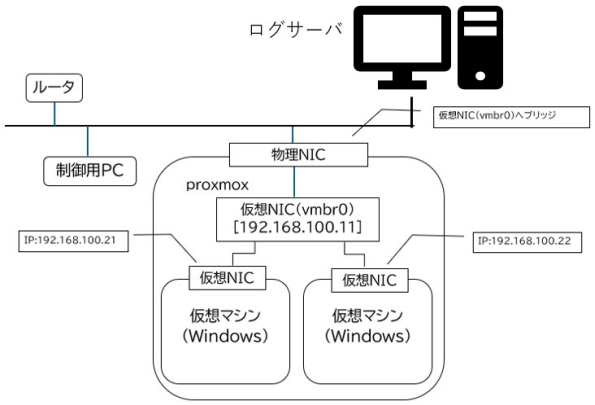


図 2: 攻撃観測システムの全体像

配布は実施されておらず，IAB の活動を模擬した攻撃誘引手法については検討されていない．

文献 [13] では，100 個の Gmail アカウント認証情報をペーストサイト，アンダーグラウンドフォーラム，およびマルウェア経由で配布し，各配布経路における悪用パターンの差異を分析した．分析の結果，マルウェア経由で取得された認証情報へのアクセスは主に Tor 経由で匿名化されていたのに対し，他の経路で取得された認証情報へのアクセスでは匿名化技術の使用率が低いことが判明した．

文献 [14] では，Google Spreadsheet の認証情報をペーストサイト上に意図的に公開し，攻撃者による実際のアクセスおよび利用状況を観測した．

文献 [15] では，Web サーバの認証情報を意図的に公開することで，攻撃者による不正アクセス，悪意のあるコンテンツの注入，マルウェア配布などの攻撃活動を観測した．

以上のように，多様な攻撃誘引手法を用いた研究が実施されてきたが，SNS 上における IAB の活動を模擬した攻撃観測は未だ実現されていない．本研究では，この研究ギャップを埋めるため，Telegram 上で囮 IAB を運用し，攻撃者の誘引と観測を実施する．

3. 観 測 方 法

本実験では，Telegram 上で活動する IAB を模擬し，攻撃観測システムへの侵入を誘引する．図 1 に本実験の全体像を示す．3.1 節では攻撃観測システムの構成を，3.2 節では Telegram 上における囮 IAB を用いた攻撃誘引手法を説明する．

3.1 攻撃観測システム

攻撃観測システムの全体像を図 2 に示す．本システムは，外部からのリモート侵入を許容して攻撃を観測する仮想マシン群と，これらの活動ログを収集・監視するログサーバから構成される．各構成要素の役割と実装について以下に述べる．

仮想マシン:1 台の物理サーバ内に仮想化ソリューションである Proxmox [16] を用いて Windows 仮想マシンを 2 台（VM-A, VM-B）構築した．VM-A は外部からのリモートアクセスによる侵入を許容し，VM-B は侵入後の VM-A からの横展開攻撃を観測する目的で設置した．VM-A には管理者権限アカウントとユーザ権限アカウントを生成し，推測困難な複雑なパスワードを設定した．ユーザ権限アカウントの認証情報は後述する囮 IAB の Telegram チャンネルに掲載し，チャンネル参加者からのリモートアクセスを誘引した．VM-B のすべてのアカウントには推測困難な複雑なパスワードを設定し，非公開とした．攻撃観測システムへのアクセスと侵入後の活動を観測するため，VM-A および VM-B にログ収集システムを導入した．標準の Windows イベントログに加え，Sysmon [17] を導入した．これによりプロセス実行，ネットワーク接続，レジストリ変更等の詳細なシステムイベントを記録し，攻撃者の滞留時間と活動内容の追跡を可能とした．

ログ収集・監視システム: 仮想マシン内で収集されたログは，NXLog [18] を用いて Graylog [19] サーバへ転送し，集約管理を実施した．さらに，ネットワークレベルでの攻撃活動を捕捉するため，物理サーバ上で tcpdump を実行し，外部ネットワークと VM-A，VM-B 間のすべてのパケットを記録した．

通信制御とセキュリティ対策: 実験環境が攻撃の踏み台として悪用されることを防ぐため，物理サーバ上の iptables により通信を厳格に制御した．

物理サーバでは，VM-A の RDP 通信（ポート 3389/tcp）および VM-B の RDP 通信（ポート 13389/tcp）のみを受信可能とした．送信については，時刻同期のための外部 NTP サーバとの通信のみを許可した．ログサーバでは，管理者によるログイン関連通信のみを受信可能とし，送信はその応答のみに制限した．この通信制御により，デコイ環境としての機能を維持しつつ，外部への攻撃拡散を防止する構成を実現した．

3.2 囮 IAB の宣伝

Telegram のメッセージング機能には「グループ」と「チャンネル」の 2 種類の形式が存在する [20]．グループではすべての参加者が発言可能で双方向のコミュニケーションが可能であるのに対し，チャンネルは管理者のみが発言権を持つことで一方向の情報配信に適している．一部の Telegram グループでは，サイバー犯罪者が潜在的顧客の獲得を目的として，自身が運営するチャンネルへの誘導リンクを含む投稿を行うことが確認されている [21]．

本実験では囮 IAB およびその管理する囮 Telegram チャンネ

ルを作成した．このチャンネルに攻撃観測システムへの RDP 認証情報を掲載することで，攻撃者の侵入を誘引した．具体的には，仮想マシンのデスクトップスクリーンショット，ログイン情報，および権限情報を掲載した．チャンネル登録アカウント数を把握するため，第 2 回から第 5 回の実験において Telegram Bot をチャンネルに追加した．Telegram Bot は Telegram アプリケーション内で動作する自動化プログラムであり，様々な処理を自動化できる [22]．本実験では，チャンネル登録アカウントのアカウント名，登録時刻，所属グループの情報を取得する目的で運用した．また，誘引効果を調査するため，招待リンクを通じてのみ参加可能なプライベートチャンネルを使用した．(注1)

さらに，アクセス情報の売買に関心を持つユーザが多く参加していると推測される Telegram グループに対して囮チャンネルの宣伝投稿を行い，チャンネルへの参加を促した．不正アクセスに利用される認証サービスは多岐にわたるが，本実験では RDP 認証情報に関心を持つグループを対象とした．具体的には，「TelegramDB Search Bot [23]」を用いて RDP，Webshell，VPN をクエリとして検索し，発見した 23 個のグループに対して囮チャンネルの宣伝投稿を実施した．

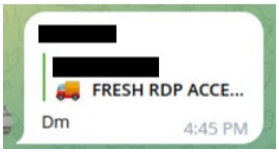
4. 観 測 結 果

前述の通り，「TelegramDB Search Bot」を用いて RDP，Webshell，VPN のクエリで検索して発見した 23 個のグループに対して 2025 年 2 月に 1 回，2025 年 5 月に 4 回，囮チャンネルを宣伝する投稿を行った．表 1 にそれぞれの実験の日時と投稿内容をまとめる．各実験で使用する囮の RDP 認証情報は，いずれも異なるものを用いた．それぞれの実験で対象としたグループについて，同一条件下での実験を行うため，既に実施済みの実験で使用したグループは投稿対象から除外した．多くの投稿が行われるグループに対しては継続的に投稿を行うことで参加者の目に触れる機会が増えるが，そのためには自動的な投稿が必要である．一方，自動投稿に対してペナルティを課するグループも存在するため，宣伝の投稿においては，まずペナルティを課されないように手動で投稿を行い，その後，継続的に投稿を行うために自動投稿を行った．

4.1 囮チャンネルの登録

全ての実験において，複数のアカウントによる囮チャンネルの登録が確認された．表 2 に各実験の囮チャンネルへの登録アカウント数と攻撃観測システムにアクセスした IP アドレス数をまとめる．なお，第 1 回実験においては，登録したアカウント数の正確な判定が行えなかったため，代わりとして視聴回数を示す．なお，第 2 回実験では 2 アカウントが，5 回実験においては 1 アカウントが一度チャンネル登録を行った後に，これをを解除したことが確認された．なお，第 4 回実験で投稿を行ったアカウントに対して，WebShell の販売を求めるダイレクトメッセージが 3 件確認された．図 3a はその一例である．またグループ内で，囮チャンネルを宣伝する我々の投稿を引用する形で，ダイレクトメッセージでのやりとりを要望する投稿が図 3b にある通り確認された．本実験においては Telegram 上での他のユーザとの会話を想定していないため，ダイレクトメッセージでの応答を行わなかったが，より現実的な IAB の活動を模擬する上で注目すべき事象といえる．

(a) WebShell 販売を求める DM



(b) DM での会話を求める投稿

図 3: 攻撃者による反応

4.2 攻撃観測システムへの侵入

第 1 回と 2 回実験では攻撃観測システムへのアクセスが確認された．表 2 に各実験のチャンネルへの登録アカウント数と攻撃観測システムにアクセスした IP アドレスの数をまとめる．なお，第 2〜5 回実験について，攻撃観測システムに不具合が発生したため，5 月 21 日 20 時頃から 5 月 26 日 17 時頃にかけてログ収集が行えなかった．そのため，第 3 回から第 5 回実験では攻撃観測システムのログサーバが作動していなかった時期に侵入された可能性が考えられる．

第 1 回実験における侵入者: 第 1 回実験において攻撃観測システムへの侵入が観測された攻撃元 IP アドレスは 5 つである．侵入時刻順に A，B，C，D，E としたときのアクセス時間を表 3 にまとめる．また，サイバー攻撃の報告を受け付け公開している AbuseIPDB [24] における調査結果を付録の表 A·1 にまとめる．すべての侵入元 IP アドレスがインドネシアのものであった．攻撃の発生時刻が特定時期に集中していることや送信元の国が同一であることから，これらの攻撃は同一の攻撃者またはグループによるものである可能性がある．

第 2 回実験における侵入者: 攻撃観測システムへの侵入が確認された攻撃元 IP アドレスは 3 つである．侵入順に F，G，H としたときの侵入時間を表 4 にまとめる．また，AbuseIPDB による調査結果を付録の表 A·2 にまとめる．全ての侵入元 IP アドレスがインドのものであった．加えて，都市名についても全 IP アドレスで共通であり，ISP についても表記上は異なっているが，同じ事業会社が提供しているものであった．最初の侵入の 5 分前に新しいアカウントがこの実験に該当する囮チャンネルを登録しており，この登録と直後の攻撃は関連がある可能性がある．イベントログを解析した結果，攻撃元 IP アドレス F，H による侵入中にのみ表 5 に示す通りの DNS 通信が観測された．このうち www.google.com の DNS 通信については両 IP アドレスの侵入期間のうち 2 分の間にのみ観測された．また，accounts.google.com の DNS 通信については両 IP ア

(注1)：Telegram 上のグループおよびチャンネルは，公開範囲の設定として「パブリック」または「プライベート」を選択可能である．パブリックに設定された場合，Telegram 内の検索結果に表示されることで任意のユーザが参加可能となる．一方で，プライベートに設定されたグループやチャンネルは検索対象から除外される．

表 1: 実験期間と Telegram 上での投稿先と投稿内容の一覧

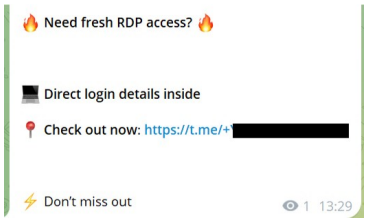
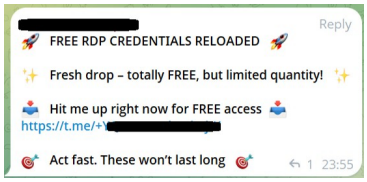
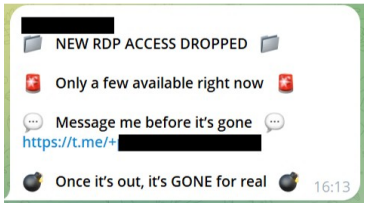
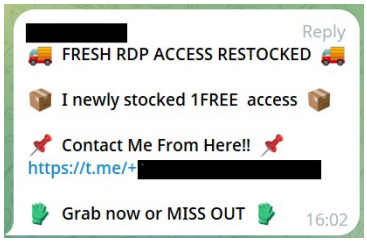
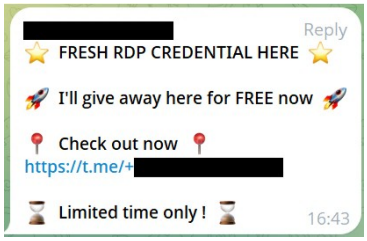
実験回数	実施期間	対象グループ数	グループ参加のべ人数	投稿内容	備考
第 1 回実験	2025 年 2 月 8 日 - 2025 年 2 月 10 日	7	25,353		2 月 9 日,10 日は Telegram API を用いて自動投稿を行った.
第 2 回実験	2025 年 5 月 16 日 - 2025 年 6 月 1 日	4	6,850		5 月 23 日-6 月 1 日は Telegram API を用いて自動投稿を行った.
第 3 回実験	2025 年 5 月 16 日 - 2025 年 6 月 1 日	4	5,582		5 月 23 日-6 月 1 日は Telegram API を用いて自動投稿を行った.
第 4 回実験	2025 年 5 月 16 日 - 2025 年 6 月 1 日	4	5,769		5 月 23 日-6 月 1 日は Telegram API を用いて自動投稿を行った.
第 5 回実験	2025 年 5 月 16 日 - 2025 年 6 月 1 日	4	8,647		5 月 23 日-6 月 1 日は Telegram API を用いて自動投稿を行った.

表 2: 実験毎の観測数の内訳

実験回数	チャンネル登録数	攻撃観測システムへの侵入
第 1 回	不明 (視聴回数 4 回)	5 (IP アドレス)
第 2 回	4 (アカウント)	3 (IP アドレス)
第 3 回	3 (アカウント)	0 (IP アドレス)
第 4 回	3 (アカウント)	0 (IP アドレス)
第 5 回	1 (アカウント)	0 (IP アドレス)

ドレスの侵入期間のうち 3 分の間にのみ観測された．このことと，F，H による DNS 通信数が同程度であったことを踏まえると，これらの通信は自動的に行われた可能性が考えられる．

表 3: 攻撃元 IP アドレス別のアクセス時間（実験 1）

IP アドレス	攻撃開始時刻	攻撃終了時刻	攻撃持続期間
A	2 月 10 日 1 時 20 分	2 月 10 日 1 時 34 分	14 分
B	2 月 10 日 1 時 36 分	2 月 10 日 1 時 56 分	20 分
C	2 月 10 日 23 時 25 分	2 月 10 日 23 時 33 分	8 分
D	2 月 11 日 20 時 45 分	2 月 11 日 21 時 15 分	30 分
E	2 月 13 日 6 時 05 分	2 月 13 日 6 時 15 分	10 分

表 4: 攻撃元 IP アドレス別のアクセス時間（実験 2）

IP アドレス	攻撃開始時刻	攻撃終了時刻	攻撃持続期間
F	6 月 1 日 17 時 02 分	6 月 1 日 17 時 14 分	12 分
G	6 月 1 日 17 時 14 分	6 月 1 日 17 時 15 分	1 分
H	6 月 1 日 17 時 14 分	6 月 1 日 17 時 33 分	19 分

表 5: 攻撃元 IP アドレス F, H による DNS 通信

試行された DNS 通信	通信の観測回数	
	F	H
www.google.com	17	16
accounts.google.com	14	14
play.googleapis.com	0	3
www.googleapis.com	0	1
safebrowsingohttpsgateway.googleapis.com	4	5
optimizationguidepa.googleapis.com	4	4

これらの通信先ドメインはインターネットへの導通確認等に典型的に用いられるものであるため，攻撃観測システムに侵入した攻撃者は，外部ホストへの通信が可能であるかを確認していた可能性がある．攻撃観測システムは外部への攻撃の流出を防ぐため，外部ホストに対するセッションの確立を許可していなかったが，今後，侵入後の攻撃者の振る舞いを観測するためには，十分に注意しつつ外部への通信を一部許可するといった対応が必要になると思われる．

4.3 Telegram グループの反応

今回の実験では，合計 23 グループに対して宣伝の投稿を行ったが，半数を超える 12 グループにおいて投稿に対する反応が見られた．これらの反応は，投稿に対して敵対的な返答を行うもの，投稿を削除するもの，投稿を数日間禁止するもの，グループから永久追放を行うもの，自動投稿に対して投稿を永久に禁止するものの 5 つに大別できる．ただし，各グループが複数の性質を持つ対応を行うこともあった．各グループの反応を付録の表 A·3 にまとめる．また，今回配布を行った 23 グループ中 7 グループがプライベートチャンネルのリンクを含んだ投稿を制限した一方で，3 グループはパブリックチャンネルのリンクを含んだ投稿を許可したことが確認された．

投稿に対して反応を示したグループの多くは敵対的な反応であった．特に，認証情報を販売することを謳い，対価を窃取するような詐欺行為を未然に防止することを目的とした行動が多く見られた．一方で，永久追放を行ったグループは，我々の投稿に対してだけでなく，同様の投稿を試みた他のアカウントも追放していたことが確認されたことから，RDP 認証情報の配布を独占的に行うことを目的とした対処であるとみられる．また，自動投稿後に投稿を剥奪する行為は，自動投稿による連続投稿を防止し，グループの秩序を維持することを目的とした処置であると考えられる．ただし該当するグループに対して短期間に連続した投稿は行っていないため，Telegram API を用いた自動投稿をトリガーとしていると考えられる．

5. まとめと今後の課題

RDP 認証情報に興味があると想定される Telegram グループに対して宣伝を投稿することによって，囮の Telegram チャネルを介して攻撃観測システムへの攻撃の誘引を行うことが可能であることを実証した．今回は簡易的な囮 IAB を用意することで誘引の可能性を検証したが，今後は，より多様な標的型攻撃の観測を実際に行うために，チャンネル上に記載する情報や，認証情報を配布する手段を精査する．また，広告が投稿された Telegram グループの半数以上が投稿に対して防御的または敵対的な措置を取った．このため，囮 IAB の宣伝や囮チャンネルへの誘導は，投稿先のグループのルールやポリシーを確認の上，慎重に実施する必要があるといえる．

6. 研 究 倫 理

本研究では，SNS 上における IAB の活動と，それに起因するサイバー攻撃の実態を明らかにすることを目的とし，Telegram 上に囮チャンネルを設置し，架空の認証情報を掲載することで攻撃の誘発と観測を行った．このような手法は，他者の行動を観察対象としつつ，欺瞞的要素を含むため，倫理的な配慮が不可欠である．本研究の実施にあたっては，所属大学の研究倫理審査委員会の審査を受け，研究計画が承認されている．具体的な対応として，第一に，囮チャンネルや投稿内容において明らかに虚偽の情報や現実の被害を誘発する誇張表現を避け，情報の提示は控えめかつ限定的な範囲にとどめた．これにより，SNS 上の他者に対する誤解や不安の助長を最小限に抑えるよう配慮した．第二に，観測期間を必要最小限に限定することで，長期的な影響の抑制と対象グループへの負荷軽減を図った．また，宣伝投稿の頻度や内容にも注意を払い，無差別なスパム行為とならないように設計しているが，グループの管理者から投稿禁止等の反応があったことから，今後はさらに十分な配慮が必要といえる．第三に，掲載した認証情報は，攻撃観測専用に構築された閉域の仮想環境へのアクセス情報とし，当該システムでは外部への通信を遮断することで，実社会における二次的な被害が発生しないよう制御している．また，攻撃に使用された IP アドレス等のログ情報はすべて研究目的に限定して利用し，個人や組織を特定可能な情報は記録していない．本論文において公開する結果も，統計的処理や抽象化を通じて匿名性を確保している．本研究は，サイバー攻撃の予兆となり得る活動を観測し，セキュリティ上の脅威に対する対策を強化するための基盤知見を提供することを目的としており，社会的意義は大きいと考えられる．一方で，欺瞞的手法や SNS 上での観察という性質上，倫理的リスクを伴うことは否定できず，今後も透明性と慎重さをもって研究に取り組む必要がある．

謝辞　本研究の一部は国立研究開発法人新エネルギー・産業技術総合開発機構（ＮＥＤＯ）の委託事業「経済安全保障重要技術育成プログラム／先進的サイバー防御機能・分析能力強化」（JPNP24003）によるものである．

文　　献

[1] Cyberint, “A deep-dive into initial access brokers: Trends, statistics, tactics and more,” https://cyberint.com/blog/research/a-deep-dive-into-initial-access-brokers-trends-statistics-tactics-and-more/s/.

[2] IPA, “コンピュータウイルス・不正アクセスの届出状況［2023年（1月～12月）］,” https://www.ipa.go.jp/security/todokede/crack-virus/ug65p9000000nnpa-att/2023-report.pdf.

[3] IPA, “コンピュータウイルス・不正アクセスの届出状況［2019年（1月～12月）］,” https://www.ipa.go.jp/security/todokede/crack-virus/ug65p9000000nnpa-att/000080224.pdf.

[4] 警視庁, “令和５年におけるサイバー空間をめぐる脅威の情勢等について,” https://www.npa.go.jp/publications/statistics/cybersecurity/data/R5/R05_cyber_jousei.pdf.

[5] D. Fraunholz, D. Schneider, J. Zemitis, and H.D. Schotten, “Hack my company: An empirical assessment of post-exploitation behavior and lateral movement in cloud environments,” Proceedings of the central European cybersecurity conference 2018, pp.1–6, 2018.

[6] D.A. Bermudez Villalva, J. Onaolapo, G. Stringhini, and M. Musolesi, “Under and over the surface: a comparison of the use of leaked account credentials in the dark and surface web,” Crime Science, vol.7, no.1, pp.1–11, 2018.

[7] 津田侑他, “サイバー攻撃誘引基盤 stardust,” コンピュータセキュリティシンポジウム 2017.

[8] R. Middelweerd, “Defining who is attacking by how they are hacking,” https://www.cs.ru.nl/masters-theses/2019/R_Middelweerd___Defining_who_is_attacking_by_how_they_are_hacking.pdf.

[9] Pastebin, https://pastebin.com/.

[10] GitHub, https://gist.github.com/.

[11] Google, https://docs.google.com/.

[12] VirusTotal, https://www.virustotal.com/.

[13] J. Onaolapo, E. Mariconti, and G. Stringhini, “What happens after you are pwnd: Understanding the use of leaked webmail credentials in the wild,” Proceedings of the 2016 internet measurement conference, pp.65–79, 2016.

[14] M. Lazarov, J. Onaolapo, and G. Stringhini, “Honey sheets: What happens to leaked google spreadsheets?,” 2016.

[15] M. Akiyama, T. Yagi, T. Hariu, and Y. Kadobayashi, “Honeycirculator: distributing credential honeytoken for introspection of web-based attack cycle,” International Journal of Information Security, vol.17, pp.135–151, 2018.

[16] PROXMOX, “Proxmox virtual environment,” https://www.proxmox.com/en/products/proxmox-virtual-environment/overview.

[17] Microsoft, “Sysmon v15.15,” https://learn.microsoft.com/ja-jp/sysinternals/downloads/sysmon.

[18] NXLog, https://nxlog.co/.

[19] Graylog, https://graylog.org/.

[20] Telegram, “Channels, supergroups, gigagroups and basic groups,” https://core.telegram.org/api/channel.

[21] TrendMicro, “Telegram（テレグラム）とは？サイバー犯罪に悪用される理由”.

[22] Telegram, “Bots: An introduction for developers,” https://core.telegram.org/bots.

[23] TelegramDB, https://www.telegramdb.org/.

[24] AbuseIPDB, https://www.abuseipdb.com/.

付　　録

表 A· 1: AbuseIPDB による 5 つの IP アドレスの調査結果

IP アドレス	IP アドレスで判定した国名	ASN	ISP	AbuseIPDB で既出か	AbuseIPDB における報告数
A	インドネシア	AS9341	PT Indonesia Comnet Plus	Yes	16
B	インドネシア	AS24203	Asia Pacific Network Information Centre	Yes	1
C	インドネシア	AS23700	PT. First Media, Tbk	No	0
D	インドネシア	AS23693	PT. Telekomunisaki Selular (Telekomsel) Indonesia	No	0
E	インドネシア	AS131111	PT.Mora Telemakita Indonesia	Yes	3

表 A· 2: AbuseIPDB による 3 つの IP アドレスの調査結果

IP アドレス	IP アドレスで判定した国名	ASN	ISP	AbuseIPDB で既出か	AbuseIPDB における報告数
F	インド	AS24560	BHARTI AIRTEL LTD.	No	0
G	インド	AS45609	Bharti Airtel Limited	No	0
H	インド	AS24560	PT. CG 7/7 Airtel main office, Vijaipur Colony, Vibhuti Khand, Gomti Nagar, Lucknow, Uttar Pradesh 226010	No	0

表 A· 3: 広告配布によるペナルティ概要

グループ	敵対的な返答を行うグループ	投稿を削除するグループ	投稿を 3 日間禁止するグループ	グループから永久追放を行うグループ	自動投稿をトリガーとして投稿を永久に禁止するグループ
1	○				
2		○			
3		○	○		
4		○		○	
5	○				
6		○			○
7	○				
8					○
9	○				
10	○	○			
11		○			
12		○	○		
計	5	7	2	1	2