博士論文

複数サーバログの関係性分析による攻撃検知に関する研究

Detecting Malicious Behavior by Analyzing Relations among Multiple Servers Log

国立大学法人　横浜国立大学
大学院環境情報学府

齊藤　聡美
Satomi SAITO

2018 年 3 月

あらまし

政府や企業といった組織に対するサイバー攻撃が激化・巧妙化の一途を辿っている．特に近年では，様々な種類や規模の組織が攻撃の対象となってきている．そのため，種類や規模を問わず組織に対する通信を監視し，攻撃の発生に早期に気づけるように，セキュリティ監視の導入が求められている．例えば，サイバー攻撃によって引き起こされるインシデント（情報システムの運用におけるセキュリティ上の問題として捉えられる事象，JPCERT コーディネーションセンター（JPCERT CC）より）を発見するための監視体制（SOC, Security Operation Center）の導入などが挙げられる．日本ネットワークインフォメーションセンター（JPNIC）では，SOC とは「情報システムへの脅威の監視や分析などを行う役割や専門組織」を指し,「ファイアウォールや侵入検知システムといったセキュリティ機器，ネットワーク機器や端末のログなどを定常的に監視し，場合によっては起きた事象を分析して，脅威となるインシデントの発見や特定，連絡を行う」といった業務を行う．

こうしたインシデント発生の有無を判断するために，SIEM（Security Information & Event Management System）を導入する場合も考えられる．SIEM では，アクセス履歴やアラートを蓄積したログ（セキュリティログ）を入力し，レコード件数や発生時刻の推移などに着目することで，通常とは異なる「異常」なログの発生を検知できる．しかし，組織内の状況やシステムが目まぐるしく変化するような状況では，異常なログの発生が攻撃事象に結びつかない場合も多い．SIEM では，ログを汎用的に入力対象とできる構成となっており,「異常」よりもセキュリティ的に深い意味を持つアラートの検知を想定していない．そこで，膨大かつ多様なセキュリティログから，インシデントとして対応すべき攻撃事象を抽出するための技術が必要とされている．

本研究では，正規通信，不正通信および意図不明な通信が混在するセキュリティログから，異常に留まらない，インシデントとして対応すべき攻撃事象の抽出を行う．アプローチとして，複数の拠点に存在するサーバから取得できたセキュリティログを対象とし，サーバ間の関係性に着目することで，攻撃事象の抽出を目指す．複数拠点に存在するサーバは，それぞれが異なる利用目的で運用されているため，サーバによってログは異なる特性を持っているはずである．それにも関わらず，こうしたサーバ間で関係性を持つ事象が記録されていたならば，個々のサーバの運用特性から外れた事象であると考えられる．この事象を，攻撃を意図した事象であると判断することで，正規・不正・意図不明な通信が混在するログであっても，攻撃事象を抽出することが可能となる．本研究は，ホスティングサービスやパブリッククラウドサービスといった，サーバ環境を複数の利用者に提供するサービス提供者が，セキュリティ監視を行う場合を想定する．本研究では，一般によく使われるサービスである，Web サイト，SSH（Secure Shell）サービス，RDP（Remote Desktop Protocol）サービスの 3 種類のネットワークサービスを対象に，攻撃事象検知手法を提案する．

まず複数の Web サイトを対象として，単一の送信元から複数の Web サイトに向けて送信されたリクエストに着目することで，脆弱性を持つ Web アプリケーションの探索や，悪意あるコード挿入を行うリクエストで，単一の送信元から送信される際にその URI にパターンが存在する攻撃を検知する手法を提案する．提案手法を，横浜国立大学情報基盤センターが管理している学内向け Web ホスティングサービスより取得できたアクセスログを用いて評価を行った結果，攻撃元となった IP アドレスを誤検知なく抽出できるしきい値が存在することを示した．さらに，既存の攻撃検

知ツールでは悪性と判断しなかったものの，他文献により悪性の可能性が高いリクエストを悪性と判断できることを示した．

次に，様々な意図や規模の攻撃が混在するブルートフォース検知ログから，ログイン試行回数や送信元 IP アドレスの出現回数などの集計では現れない，ステルス性の高い分散型ブルートフォース攻撃事象の抽出および検知を行う手法を提案する．まず，ブルートフォース検知ログから，送信元 IP アドレスと送信先 IP アドレスの関係性を散布図形式で可視化することで，送信元 IP アドレスを変えながら特定の送信先 IP アドレス群に対してブルートフォース攻撃を繰り返す事象を抽出できた．さらに，この攻撃事象を受ける IP アドレスを早期に検知する手法を提案する．この手法を，企業が管理する IP アドレス帯で実際にサービスを運用している複数の SSH サービスに対して記録されたブルートフォース検知ログに適用した結果，検知時刻とログイン試行回数の相関を取る場合に，最も精度良く送信先 IP アドレスを特定できることを示した．

また，ブルートフォース検知ログから，送信元 IP アドレスによるブルートフォース攻撃アラートのインターバルやアラートが検知した攻撃回数の分散を計測することで，送信元 IP アドレスを変えながら，共通の規則性を以ってブルートフォース攻撃を繰り返す事象を抽出できた．さらに，この規則性を有する送信元 IP アドレスを特定することで，この事象でブルートフォース攻撃の発生を遮断する手法を提案する．この手法を，企業が管理する IP アドレス帯で実際にサービスを運用している複数の RDP サービスに対して記録されたブルートフォース検知ログに適用した結果，ドロップ率と過剰遮断時間を最小限に抑えるパラメタを見積もることができた．

Abstract

Attacks on cyber space have become aggressive and sophisticated. In recent years, their targets continue to diversify. Their victims are not only major corporation, government and *etc*, but also small and medium-sized enterprises. Therefore, it is strongly required for organizations to apply security monitoring for detecting incidents caused by cyber attacks. According to Japan Computer Emergency Response Team Coordination Center (JPCERT), the incident is a security problem accident in running on information systems. And according to Japan Network Information Center (JPNIC), Security Operation Center (SOC) is a role for monitoring and analyzing to detect those incidents. The SOC monitors firewalls, intrusion detection system and network devices logs and detect incidents with high security risks.

Security Information and Event Management System (SIEM) is effective to decide incidents from such innumerable alerts, According to the input device and alert logs (security logs), SIEM detects anomalous alerts by focusing the changes in alert records, the frequency in detected date and *etc*. However, as the situation in organizaions systems and network are rapidly changing, only anomalous alerts cannot always indicate incidents directly. It is difficult to have meanings beneficial for incident decision to detected alerts for SIEMs, because those input are supposed to generic. Therefore, it is required new techniques for detecting malicious behavior from innumerous and various security log.

This paper presensts the techniques for detecting not only anomarous but also malicious behavior from security log mixed with normal, malicous and unknown records. We aim to extract malicous behavior by collecting servers log from multiple sites and focusing on the relations among them. Those servers have different purposes and scale each other. The servers log from them have also different feature. Despite that, sharing the same behavor among them indicates the events out of normal servers operations. By judging those events as related attacks, it ensbles extracting malicous behavior from security log mixed with normal, malicous and unknown records. This proposal techniques are for monitoring network service environments by service providers in such as hosting services and public cloud services. This paper's targets are websites, SSH (Secure Shell) services and RDP (Remote Desktop Protocol) services, those are commonly used on the Internet.

First, from website access log collected from multiple sites, we propose a method for detecting malicious requests that are for scanning vulnerable web applications and malicious code insertions. As a result of applying our proposal method to the access log collected from the website hosting service on Yokohama National University, we show that our method can detect malicious requests with no false positives under the suitable paramters. Furthermore, we also show that some malicious requests are cannot be detected by existing tools.

Second, from brute force alert log mixed with various intentions, we propose two methods for detecting stealthy distrbuted brute force attacks events. Those events have much source IP addresses with low login trials that conventional analyzing methods cannot reveal them. The first method applies a scatter-based visualization to brute force alert log. As the scatter-based vizualization clarify the relation among source IP addresses, destination IP addresses and detected date, we extract a distrubuted brute force attacks

event. In this event, specific destination IP addresses have been targets from many source IP addresses with low login trials. We also propose the method for detecting such target destination IP addresses. As a result of applying our detecting method to the brute force alert log against SSH services collected from the network managed by an enterprise, we show that our method can specify the target destination IP addresses with correlating detected date and login trials.

The second method measures the statistics about alert intervals and login trials dispersion from brute force alert log. As a result of the measuring, we extract another distributed brute force attack event. This event have regularities in login trials and intervals shared in different source IP addresses. We also propose the method for detecting and intercepting the login trials related to the event. As a result of applying our detecting and intercepting method to the brute force alert log against RDP services collected from the network managed by an enterprise, we estimated our method parameters that enables minimum dropping ratio and over intercepting time.

目 次

図 目 次

表 目 次

第1章　序論

1.1　背景と目的

政府や企業といった組織に対するサイバー攻撃が激化・巧妙化の一途を辿っている．特に近年では，様々な種類や規模の組織が攻撃の対象となってきている．そのため，種類や規模を問わず組織に対する通信を監視し，攻撃の発生に早期に気づけるように，セキュリティ監視の導入が求められている．例えば，サイバー攻撃によって引き起こされるインシデント（情報システムの運用におけるセキュリティ上の問題として捉えられる事象，文献 [1]より引用）を発見するための監視体制（SOC, Security Operation Center）の導入などが挙げられる．文献 [2]では，SOC とは「情報システムへの脅威の監視や分析などを行う役割や専門組織」を指し,「ファイアウォールや侵入検知システムといったセキュリティ機器，ネットワーク機器や端末のログなどを定常的に監視し，場合によっては起きた事象を分析して，脅威となるインシデントの発見や特定，連絡を行う」といった業務を行う．

これまでに，通信パケットや端末内のログの特徴を用いて，マルウェア特融の挙動を検知するなどにより攻撃通信を検知する技術が多数提案，実用化されている．セキュリティ監視においては，これらの技術により検知された攻撃通信をアラートとして受信し，セキュリティアナリストと呼ばれる専門家によりインシデントとして対応すべきか否かを判断する．しかし今日では，受信するアラートの量が膨大となってきている．さらにその中には誤検知に該当するアラートも存在する．そのためセキュリティ監視においては，膨大なアラートからインシデントとして対応すべき事象を選定する必要がある．

こうしたインシデント発生の有無を判断するために，SIEM（Security Information & Event Management System）を導入する場合も考えられる．SIEM では，アクセス履歴やアラートを蓄積したログ（セキュリティログ）を入力し，レコード件数や発生時刻の推移などに着目することで，通常とは異なる「異常」なログの発生を検知できる．しかし，組織内の状況やシステムが目まぐるしく変化するような状況では，異常なログの発生が攻撃事象に結びつかない場合も多い．SIEM では，ログを汎用的に入力対象とできる構成となっており,「異常」よりもセキュリティ的に深い意味を持つアラートの検知を想定していない．そこで，膨大かつ多様なセキュリティログから，インシデントとして対応すべき攻撃事象を抽出するための技術が必要とされている．

そこで本研究では，正規通信，不正通信および意図不明な通信が混在するセキュリティログから，異常に留まらない，インシデントとして対応すべき攻撃事象の抽出を行う．アプローチとして，複数の拠点に存在するサーバから取得できたセキュリティログを対象とし，サーバ間の関係性に着目することで，攻撃事象の抽出を目指す．複数拠点に存在するサーバは，それぞれが異なる利用目的で運用されているため，サーバによってログは異なる特性を持っているはずである．それにも関わらず，こうしたサーバ間で共通あるいは関係性を持つ事象が記録されていたならば，個々のサーバの運用特性から外れた事象であると考えられる．この事象を，攻撃を意図した事象であると判断することで，正規・不正・意図不明な通信が混在するログであっても，攻撃事象を抽出することが可能となる．本研究は，ホスティングサービスやパブリッククラウドサービスといった，サーバ

環境を複数の利用者に提供するサービス提供者が，セキュリティ監視を行う場合を想定する．本研究では，一般によく使われるサービスである，Web サイト，SSH（Secure Shell）サービス，RDP（Remote Desktop Protocol）サービスの 3 種類のネットワークサービスを対象に，攻撃事象検知手法を提案する．Web サイトは，インターネットに対して一般向けにコンテンツを公開するために使われるサービスである．SSH サービスおよび RDP サービスは，サーバをリモートで管理するために使われるサービスで，SSH サービスは Linux OS，RDP サービスは Windows OS 向けのホストを対象としている．

まず，Web サイトを対象に，Web サイトに対するアクセスログから攻撃事象を検知する手法を提案する．Web サイトでは，インターネットに対して文章や画像，動画などを公開している．そこには，正規ユーザによるアクセスや検索エンジンによる情報収集目的のアクセス，攻撃を目的とした悪意あるアクセスなどが日々到達している．そこで，こうした様々な意図を持つアクセスが混在するアクセスを記録したログから，悪意あるリクエストを検知する手法を提案する．提案手法では，送信元 IP アドレス，送信先ドメイン，URI の関係性に着目することで，Web アプリケーションの脆弱性の探索を目的としたリクエストや悪意あるコード挿入を行うリクエストを抽出する手法を提案する．本手法を横浜国立大学情報基盤センターが管理している学内向け Web ホスティングサービスより取得できたアクセスログを用いて評価を行う．評価の結果，既存の攻撃検知手法では悪性と判断しなかったものの，他文献により悪性の可能性が高いリクエストを悪性と判断できることを示す．

次に，SSH サービスおよび RDP サービスを対象に，当該サービスに対して侵入検知装置（Intrusion Detection System, IDS）がブルートフォース攻撃を検知した記録（ブルートフォース検知ログ）から，分散型ブルートフォース攻撃事象を検知する手法を提案する．SSH および RDP サービスは，どちらもサーバを管理する管理者がリモートでホストを操作することを目的としたサービスである．SSH は Linux 系のホストを，RDP は Windows 系のホストを対象としている．これらのサービスに対して発生したブルートフォース攻撃（サーバへのログインに必要なユーザ名とパスワードの組み合わせの取得を目的として，存在する全ての組み合わせについてログイン可能かどうかを，サーバに対して試す攻撃）を検知するため，IDS をサービスの前段に設置するなどして，ブルートフォース検知ログを蓄積している．このとき，IDS が検知したログイン試行回数が他と比べて大きかったり，同一の送信元 IP アドレスから攻撃が続いている場合，攻撃先のサービスが執拗に狙われているのがすぐにわかる．しかし，ログイン試行回数が小さかったり，同じ送信元 IP アドレスからの攻撃が 1〜数回きりで終わってしまうような場合，他のログに埋もれてしまい，攻撃事象あるいはインシデントとして対処すべきか否かの判断が難しい．そのため，送信元 IP アドレスを変えながらサービスに攻撃を続けていたとしても，その攻撃事象を見逃してしまう場合がある．そこで，様々な意図や規模の攻撃が混在するブルートフォース検知ログから，ログイン試行回数や送信元 IP アドレスの出現回数などの集計では現れない，ステルス性の高い分散型ブルートフォース攻撃事象の抽出および検知を行う手法を提案する．

分散型ブルートフォース攻撃事象抽出に向け，可視化および統計量計測による分析アプローチを適用する．可視化による分析アプローチでは，ブルートフォース検知ログから，送信元 IP アドレスと送信先 IP アドレスの関係性を散布図形式で可視化する．その結果，送信元 IP アドレスを変えながら特定の送信先 IP アドレス群に対してブルートフォース攻撃を繰り返す事象を抽出する．さらに，この攻撃事象を受ける IP アドレスを早期に検知する手法を提案する．これらの提案手法を，企業が管理する IP アドレス帯で実際にサービスを運用している複数の SSH サービスに対して記録されたブルートフォース検知ログに適用し，分散型ブルートフォース攻撃を受ける送信先 IP アドレスを高い精度で抽出できることを示す．送信先 IP アドレスを特定することで，当該 IP アドレスに

対する攻撃を重点監視し，当該 IP アドレス間で共起して発生したログイン試行をブルートフォース攻撃と判断し，通信の遮断が可能となる．

統計量計測による分析アプローチでは，ブルートフォース検知ログから，送信元 IP アドレスによるブルートフォース攻撃アラートのインターバルやアラートが検知した攻撃回数の分散を計測することで，送信元 IP アドレスを変えながら，共通の規則性を以ってブルートフォース攻撃を繰り返す事象を抽出する．さらに，この規則性を有する送信元 IP アドレスを特定することで，この事象でブルートフォース攻撃の発生を遮断する手法を提案する．これらの提案手法を，企業が管理する IP アドレス帯で実際にサービスを運用している複数の RDP サービスに対して記録されたブルートフォース検知ログに適用し，分散型ブルートフォース攻撃の発生を早期に検知し，この攻撃によるログイン試行の到達を最小限に抑えられることを示す．

1.2 本研究で対象とするログ

1.2.1 Web サイトに対するアクセスログ

本研究で対象とする，Web サイトに対するアクセスログについて述べる．アクセスログには，ある送信元 IP アドレスから Web ホスティングサービス管理下のドメインに対して送信されたリクエストが 1 行のレコードとして記録される．アクセスログ中のレコードは，送信元 IP アドレス，送信先となったドメイン（ドメイン），アクセス受信時刻（受信時刻），ドメインに対して送信された URI（URI）の 4 項目から構成される．アクセスログの例を表 1.1 に示す．

表 1.1: アクセスログの例

送信元 IP アドレス	ドメイン	受信時刻	URI
a.a.a.a	ynu	2017/4/1 0:00	GET /index.html
a.a.a.a	ynu	2017/4/1 0:01	GET /logo.png
b.b.b.b	ias	2017/4/1 0:30	GET /robots.txt
:	:	:	:

1.2.2 IDS がブルートフォース攻撃と判断したアラートログ（ブルートフォース検知ログ）

本研究で対象とする，IDS によりブルートフォース攻撃と判断したアラートログ（ブルートフォース検知ログ）について述べる．文献 [3] より，IDS（侵入検知装置）は，セキュリティ侵害を見発見するためにシステムの挙動を監視する目的を持ち，攻撃の可能性がある悪意ある挙動に対してアラートを出す機能を持つ装置である．表 1.2 に，ブルートフォース検知ログの例を示す．IDS に搭載されているブルートフォース攻撃検知アルゴリズムに従って，ブルートフォース攻撃送信元と判断された IP アドレスを送信元 IP アドレス，送信先と判断された IP アドレスを送信先 IP アドレスとする．1 つのブルートフォース検知レコードは，送信元 IP アドレス，送信先 IP アドレス，検知時刻，通信先ポート番号，ブルートフォース攻撃と判断したログイン試行回数（ログイン試行回数）の 5 項目から構成される．

表 1.2: ブルートフォース検知ログの例

送信元 IP アドレス	送信先 IP アドレス	検知時刻	攻撃対象サービス	ログイン試行回数
s.s.s.s	x.x.x.x	2017/4/1 0:00	SSH	100
t.t.t.t	y.y.y.y	2017/4/1 0:01	SSH	92
u.u.u.u	z.z.z.z	2017/1/3 0:20	RDP	56
:	:	:	:	:

1.3 本論文の構成

本論文の構成を示す．まず，第 2 章で本研究に関連する先行研究について紹介する．次に，第 3 章から第 5 章で，本研究のアプローチの有効性を検証するため，複数サーバに対するアクセスログおよびブルートフォース検知ログを題材として攻撃事象抽出のための分析手法を提案し，評価を実施した結果を報告する．第 3 章では，複数の Web サイトを運用する Web ホスティングサービスに対するアクセスログを対象とし，悪意あるリクエストを抽出する手法を提案し，評価した結果を述べる．第 4 章では，企業が管理する IP アドレス帯で実際にサービスを運用している複数の SSH サービスに対するブルートフォース検知ログを対象とし，送信元 IP アドレスを変えながら特定の送信先 IP アドレス群に対してブルートフォース攻撃を繰り返す事象を抽出し，対策する手法を提案し，評価した結果を述べる．第 5 章では，企業が管理する IP アドレス帯で実際にサービスを運用している複数の RDP サービスに対するブルートフォース検知ログを対象とし，送信元 IP アドレスによるブルートフォース攻撃アラートのインターバルやアラートが検知した攻撃回数の分散を計測し，送信元 IP アドレスを変えながら，共通の規則性を以ってブルートフォース攻撃を繰り返す事象を抽出し，対策する手法を提案し，評価した結果を述べる．第 6 章で結論を述べる．

第2章　本研究に関する先行研究

本章では，本研究と関する先行研究および技術について述べる．

2.1　データ分析技術

2.1.1　異常検知

セキュリティログから攻撃事象を抽出するための技術として，異常検知が多く利用されている．文献 [4]によれば，異常とは「予期される挙動に適合しないデータパターン」である．文献 [5]では，異常を「Point Anomalies」「Contextual Anomalies」「Collective Anomalies」の 3 種類に分類している．

分類された各異常について文献 [5]では，それぞれ下記のように説明している．あるデータが残りのデータに対して異常とみなせる場合，このデータを Point Anomaly であるという．例えば図 2.1 では，点 o_1 および点 o_2 および領域 O_3 は，いずれも正常領域（領域 N_1 および領域 N_2）の境界の外側にあり，正常なデータ点とは異なるため，Point Anomalies と判断される．

あるデータが特定の前後関係においてのみ異常とみなせる場合，このデータを Contextual Anomaly であるという．判断基準となる前後関係は，データセットの構造によって決まり，問題設定の一部として設定される必要がある．各データは下記の 2 種類の属性を持つ．(1) 前後関係を示す属性: データの前後関係を定義できる属性．(2) 挙動を示す属性: データの前後関係の定義に関係ない属性．例えば図 2.2 は，過去数年間のある地域の毎月の温度を示す時系列変化の例である．冬の時期の時刻 t_1 と夏の時期の時刻 t_2 での温度そのものは同じ値であるが，時刻 t_2 の方が異常であると判断できる．

ある関連性を持つデータ集団がデータセット全体に対して異常とみなせる場合，このデータ集団を Collective Anomaly であるという．データ集団に属する個々のデータはそれ自体では異常ではないものの，集団としてみた場合には異常と判断できるものを示す．例えば図 2.3 は，人間の心電図の例である．ハイライトされた領域（1000〜1600 間）では，低い値そのものは異常とはみなされないものの，心臓の収縮に対して異常に長い時間低い値が継続して記録されているため，異常であると判断できる．

こうした異常の検知に際し，文献 [6]では，統計的異常検知における基本処理を次の 2 ステップからなると述べている．まず Step1 では，これまでに得られたデータから，データの発生分布の確率モデルを学習する．次に Step2 では，Step1 で学習されたモデルを基に，データの異常度合い，またはモデルの異常な変化度合いをスコアリングする．この基本処理に基づき，統計的異常検知を「外れ値検出」「変化点検出」「異常行動検出」の 3 手法に分類し，適用する確率モデルや検出対象等を，表 2.1 のようにまとめている[1]．

[1]Point Anomalies は外れ値，Contextual Anomalies は変化点，Collective Anomalies は異常行動に，それぞれ対応すると考えられる．

また異常検知の応用先として，セキュリティ分野への応用が述べられている．文献 [6]では，セキュリティ分野への応用における効果として，1) セキュリティ・インシデントに関する知識発見をもたらす．2) 未知のセキュリティ・インシデントを早期に発見できる．3) サイバー犯罪特定のための工数を削減する．の 3 点を挙げている．

表 2.1: 統計的異常検知の分類（文献 [6]より）

機能	入力対象	確率モデル	検出対象	応用
外れ値検出	多次元ベクトル	独立モデル（ガウス混合分布，ヒストグラム）	外れ値	不正検出，侵入検知，故障検知
変化点検出	多次元時系列	時系列モデル（AR モデル，回帰モデルなど）	時系列上の急激な変化，バースト的異常	攻撃検出，ワーム検出，障害予兆検出
異常行動検出	セッション時系列	行動モデル（混合隠れマルコフモデル，ベイジアンネットなど）	異常セッション，異常行動パタン	なりすまし検出，障害予兆検出，不審行動検出

2.1.2 分析ソフトウェア

本節では，セキュリティログの分析を行うためのソフトウェアについて述べる．

Web サイトに対するアクセスログ（Web アクセスログ）を解析するソフトウェアとして，例えば，GoAccess（ [7]）がある．GoAccess は，リアルタイムに Web アクセスログを解析し，アクセス数の時間変化やアクセス数の多かったコンテンツ等を集計し，ターミナル上およびブラウザ上で閲覧できるオープンソースソフトウェアである（図 2.4，図 2.5）．Google は Google Analytics（ [8]）で，解析対象としたい Web サイト上のページにトラッキングコードと呼ばれる JavaScript コードを挿入することで，挿入したページに対するアクセスを Google 側で収集し解析するサービスを提供している（図 2.6）．

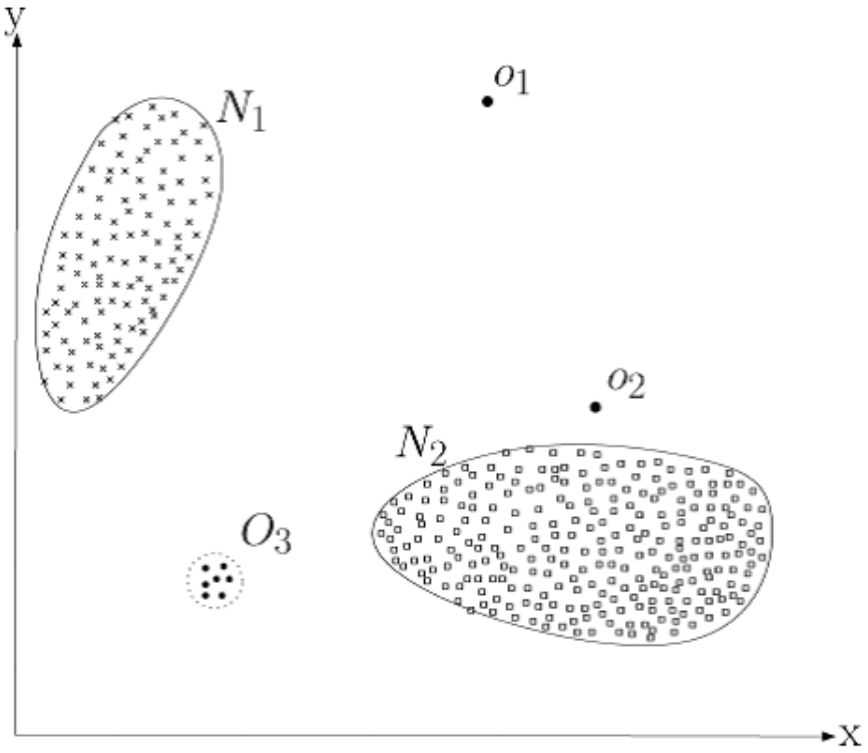

図 2.1: Point Anomalies の例（文献 [5] より）

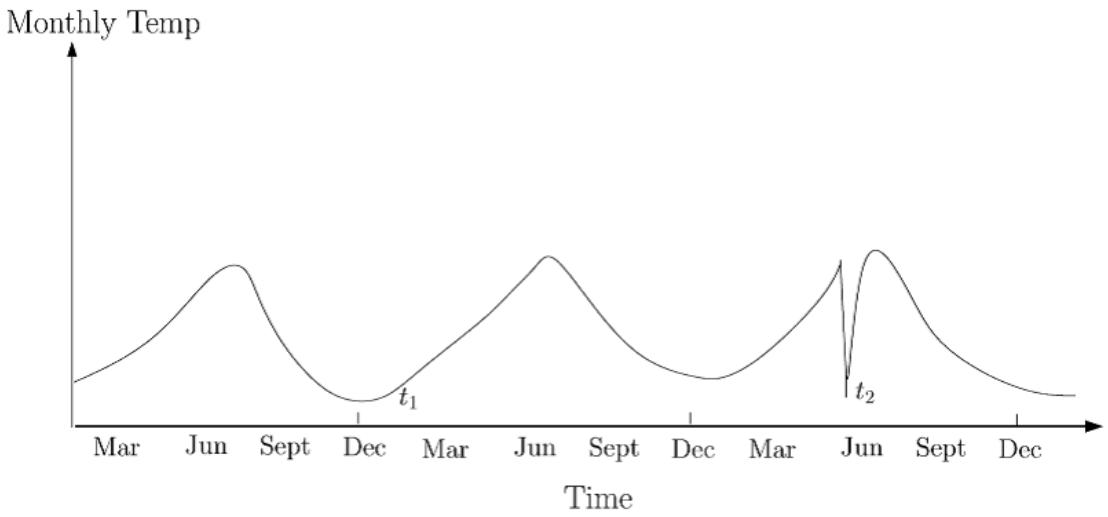

図 2.2: Contextual Anomalies の例（文献 [5] より）

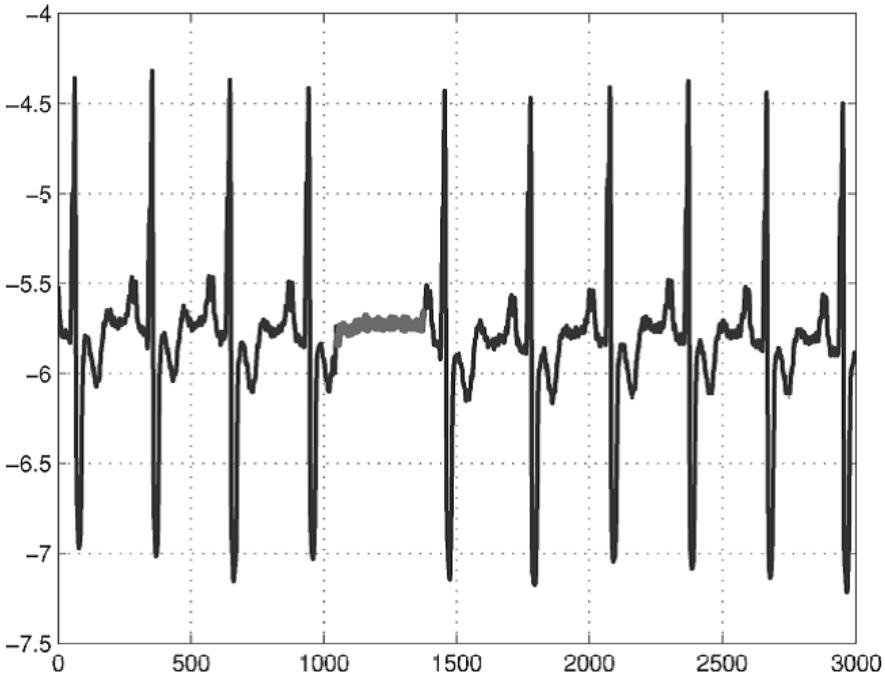

図 2.3: Collective Anomalies の例（文献 [5] より）

```
  3 - Static Requests                                                Total: 82/82

Hits Vis.      %   Bandwidth Mtd Proto    Data
---- ---- ----- ----------- --- -------- ----
35     30 3.08%   37.94 KiB GET HTTP/1.1 /img/image4.jpg
35     31 3.08%   50.35 KiB GET HTTP/1.1 /img/image5.jpg
35     30 3.08%    1.01 MiB GET HTTP/1.1 /img/header-dnj.png
35     34 3.08%   22.87 KiB GET HTTP/1.1 /img/image71.jpg
35     34 3.08%   62.69 KiB GET HTTP/1.1 /img/image14.jpg
34     29 3.00%   55.08 KiB GET HTTP/1.1 /img/image3.jpg
34     30 3.00%   56.45 KiB GET HTTP/1.1 /img/image6.jpg

  4 - Not Found URLs (404s)                                          Total: 9/9

Hits Vis.      %   Bandwidth Mtd Proto    Data
---- ---- ----- ----------- --- -------- ----
27      0 2.38%    5.51 KiB GET HTTP/1.1 /favicon.ico
9       0 0.79%    1.83 KiB GET HTTP/1.1 /robots.txt
2       0 0.18%   418.0    B GET HTTP/1.0 /favicon.ico
2       0 0.18%   460.0    B GET HTTP/1.1 /apple-touch-icon-precomposed.png
2       0 0.18%   436.0    B GET HTTP/1.1 /apple-touch-icon.png
1       0 0.09%   208.0    B GET HTTP/1.0 /robots.txt
1       0 0.09%   221.0    B GET HTTP/1.1 http://www.baidu.com/cache/global/img/
```

図 2.4: GoAccess による解析結果のターミナル画面表示例

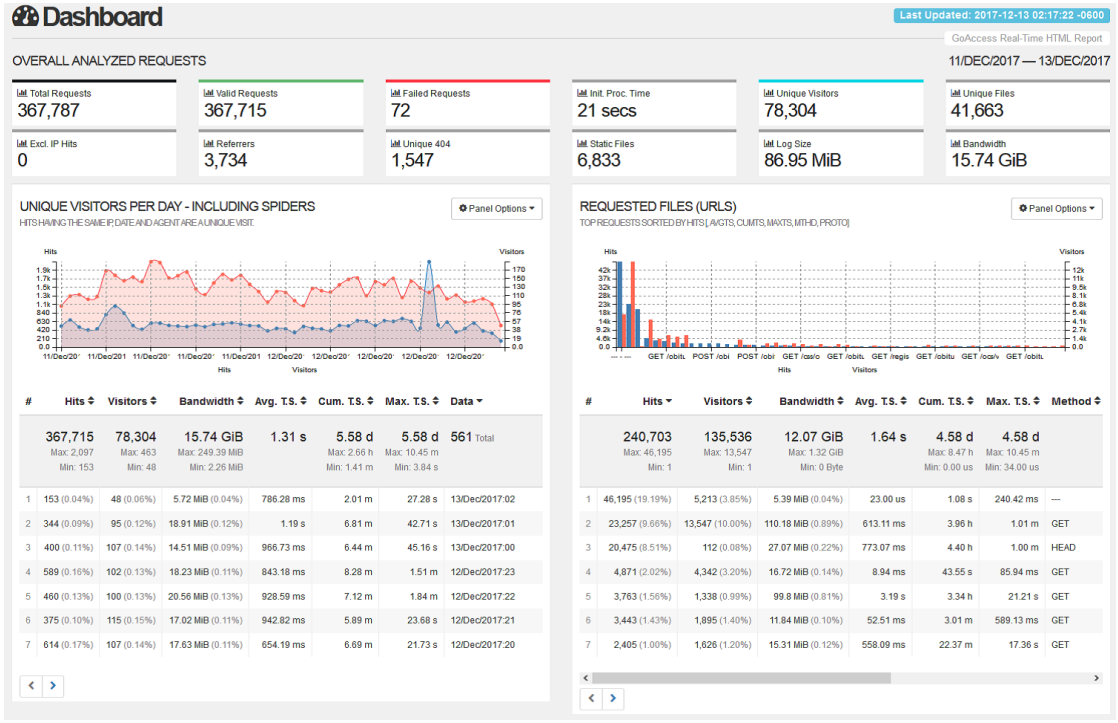

図 2.5: GoAccess による解析結果のブラウザ上での表示例（文献 [7] より）

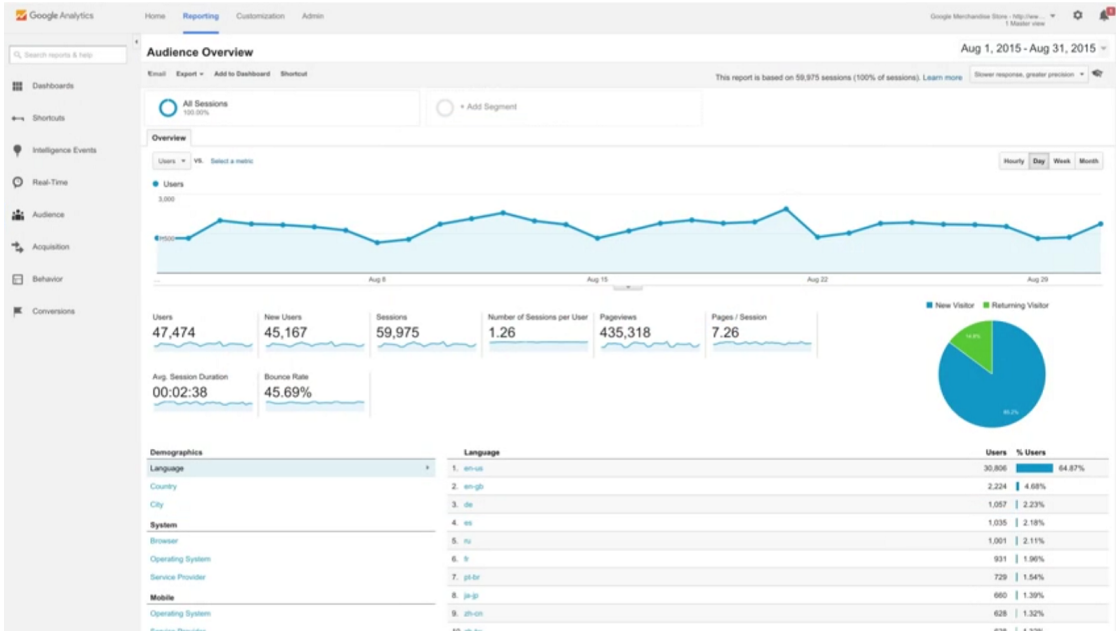

図 2.6: Google Analytics 解析結果のブラウザ上での表示例（文献 [9] より）

2.2 Webサイトに対する悪性リクエストの検知・分析

本節では，Web サイトへの攻撃の分析・検知技術に関する関連研究を紹介する．文献 [10]で Kruegel らは URI の文字列をモデル化し学習することで，Web サイトに対する異常なリクエストを抽出する手法を提案している．文献 [11]では，鐘らは Web アプリケーションの脆弱性や設定ミスに起因する脆弱性を持つ Web サイトの発見を目的としたリクエスト（Web スキャン）を，アクセスログから検知する手法を提案している．この手法では，1 つの送信元 IP アドレスから多数の送信先 IP アドレスへアクセスが発生すること，Web スキャンに該当するリクエストはその URI が文字列として類似することを Web スキャンの特徴としている．しかし，URI が文字列として類似しないリクエストの悪性の有無を判断することが難しい．そのため，文字列としては類似しないものの，送信される URI にパターンが存在するリクエストを抽出できない場合がある．また，文献 [12]では，Sanghyn らはベイズ推定を用いて，ページ遷移の順序性に着目した異常セッションの抽出手法が提案されている．これらの手法により，正規ユーザによる要求が稀な文字列を有するリクエストや複数のリクエストから成るコンテンツアクセスを，攻撃として検知することができる．文献 [13]では，Stevanovic らはアクセスログから抽出した特徴を自己組織化マップに適用し，悪意ある Web クローラの識別を行う手法を提案している．彼らは，1 セッション当たりのリクエスト発生頻度やリクエスト種類，レスポンスのエラー率等を特徴としている．

Web サイトに対する攻撃を防ぐ手段として，Web サイトに対する攻撃を模擬したリクエストを Web サイトに適用することで，その Web サイトが持つ脆弱性を洗い出す手段が存在する．洗い出した脆弱性を公開前に対処することで，Web サイトに対して実際に攻撃が発生するのを防ぐことができる．文献 [14] [15]では，SQL インジェクションや XSS（クロスサイトスクリプティング）といった Web サイトに存在する脆弱性を検査する手法が提案されている．文献 [16]では，Doupe らは Web サイトの内部状態の遷移を考慮した脆弱性検査手法を提案している．また，こうした脆弱性検査技術の評価を行った研究もある．文献 [17]では，Doupe らは著者らが構築した脆弱性を持つ Web サイトに対して 11 種類のブラックボックス型の脆弱性検査ツールを適用し，脆弱性の検知率を計測している．文献 [18]では，Bau らは 8 種類の脆弱性検査製品に対して既知の脆弱性を突く攻撃の検知率を比較した結果を報告している．しかし，パターンマッチングと同様に，ゼロデイや発見されたばかりの脆弱性を突く攻撃であれば，こうした検査が間に合わない場合が存在する．

サーバ型ハニーポットをインターネット上に設置することで，Web サイトに対する攻撃に関する情報を収集できる．ハニーポットに対するアクセスを精査することで，攻撃リクエストの傾向や，ゼロデイといったこれまで認識されていなかった攻撃を知ることができる．文献 [19]では，John らは脆弱性を持つ Web サイトの情報をインターネットから取得し再現するハニーポットシステムを提案している．文献 [20]では，八木らは送信元とインタラクションを行うことで攻撃側の挙動を観測できる Web ハニーポットを提案している．文献 [21]で久世らはハニーポットに対する通信ログから HTTP リクエストおよびレスポンスや通信の特徴を抽出し，悪意あるアクセスを識別する手法を提案している．特に，ハニーポットを複数台設置することで，数値的に連続した IP アドレスに対して順に通信要求を行う特徴を取得可能であると述べている．Davide らは文献 [22]で，Web アプリケーションの脆弱性を模擬した Web サイトを多数設置してリクエストを観測している．Davide らは観測の結果，4 段階の攻撃フェーズと 13 種類の目的に分類できたことを報告している．他にも，インターネット上で入手可能なサーバ型ハニーポットとして，Glastopf（ [23]），Dionea（ [24]），DShield（ [25]），HiHAT（ [26]）などが存在する．ハニーポットで観測できた攻撃情報を IDS 等のシグネチャにフィードバックする技術も提案されている．文献 [27]では，Kreibich らはハニーポットに到達した通信履歴から IDS 用のシグネチャを自動生成する手法を提案してい

る．ハニーポットは実際のサービスが運用されていないため，ハニーポットに到達するリクエストをほぼ悪性とみなして，悪意あるリクエストの収集や分析が実施できる．しかし，囮の Web サイトであることが攻撃側にわかってしまうと，攻撃対象から除外され，悪意あるリクエストが観測できなくなる恐れがある．

2.3 ブルートフォース攻撃の検知・分析

ネットワークサービスに対するブルートフォース攻撃の発生は，各組織より論文や監視レポート等により報告されている．Sperotto らは著者の所属する大学のネットワークから，scanning, brute-force, die-off の 3 種類のフェーズから構成される SSH へのブルートフォース攻撃が発生していたことを報告している［28］．Vykopal は，*srcIP* 数 : *dstIP* 数が 1:1, 1:N, N:1 の 3 種類にブルートフォース攻撃を分類し，大学のネットワーク内でそれぞれの攻撃がどの程度発生していたかを調査している［29］．Mobin らは［30］で，実際の SSH(Secure Shell) サーバアクセスログを分析し，異なる *srcIP* によるブルートフォース攻撃が発生していたことを報告している．IBM SOC，SANS，Dragon Research Group による監視レポート［31］［32］［33］［34］により，1 日に複数の異なる *srcIP* からのブルートフォース攻撃を検知したことが報告されている．また，ブルートフォース攻撃を行う挙動を含むワームの存在も確認されている．Widnows ワークステーションやサーバを狙ったワーム「Morto」は，その挙動にブルートフォース攻撃が含まれることが報告されている［35］．あるマシンが Morto に感染すると，Morto はローカルネットワークをスキャンし RDP(Remote Desktop Protocol) が有効になっている他のマシンをスキャンする．そして RDP が有効になっているサーバを発見すると，ユーザ名を「Administrator」として，「server」「1234qwer」「admin123」等のパスワードでのログインを試みる．文献［36］においても，著者の所属する組織のネットワーク監視ログに Morto によるログイン試行が記録されていたことが報告されている．こうした RDP サービスに対するブルートフォース攻撃事象は増加傾向にあると述べている報告もある（文献［37］，文献［38］など）．また，文献［39］では，侵入可能な RDP サービスが存在するホストや，RDP サービスで頻繁に設定されるアカウントリストが公開されていると報告されている．文献［40］や文献［41］によると，Cisco と Level 3 Comunications は 2015 年 4 月に，ブルートフォース攻撃に悪用されているとされる IP アドレス帯にテイクダウンを実施したことが報告されている．また文献［42］では，Akamai Technologies Inc. により，Linux ホスト上で動く SSH サービスに対してブルートフォース攻撃を行う XOR DDoS Botnet の活動が，2015 年第 4 四半期に観測されなくなった旨が報告され，テイクダウンによるものであると推測している．さらに，2017 年に発表された文献［43］によると，SSH サービスに対してブルートフォース攻撃を行うボット「RuaBot」の存在が確認されており，攻撃元となっているのは IoT デバイスに感染するボットであると報告されている．また，RDP サービスに関しても，同じく 2017 年に発表された文献［44］によると，LockCrypt と呼ばれるランサムウェアは，RDP サービスが有効なホストにブルートフォース攻撃を行いホストに侵入する挙動を持つと報告されている．このように，各組織によりネットワークの監視が行われており，ブルートフォース攻撃事象も報告されている．これらの報告に基づく分析は単一の攻撃対象もしくは同一拠点内の複数の攻撃対象へのブルートフォース攻撃に関するものが主であった．

インターネット上で入手可能なブルートフォース攻撃ツールについて述べる．NCRACK（［45］）と THC-Hydra（［46］）は，オープンソースで提供されているブルートフォース攻撃ツールである．これらのツールは SSH，FTP，HTTP，RDP といったプロトコルをサポートしている．これらのツールの使い方を紹介したり，ベンチマークを行った結果を報告している Web サイトも多く存在す

る．また，Brutik RDP（［47］）や TSGrinder（［48］）のように，RDP サービスに特化したブルートフォース攻撃ツールも存在する．このように，ブルートフォース攻撃が手軽に実行できるようになっているといえる．

ブルートフォース攻撃の発生を検知する研究について述べる．文献［49］では，Najafabadi らは実環境で取得できたデータを対象に，ブルートフォース攻撃の検知に有効な機械学習手法の調査を行っている．文献［52］では，Hellemons らはブルートフォース攻撃が 3 段階のフェーズから構成されるのに着目し，このフェーズの変化をブルートフォース攻撃として検知する手法を提案している．文献［53］では，Satoh らは SSH サービスを対象として，パケットの種類とデータサイズに特定の関係性を持つ通信を辞書攻撃（ブルートフォース攻撃）を検知する手法を提案している．また，このように通信の規則性に着目することで，ブルートフォース攻撃に限らず，悪意あるスキャンを検知する手法も，これまで多数提案されている．文献［54］では，Gu らは IRC プロトコルを対象としてメッセージの送信間隔が一定であることをトリガーとして，ボットネットによる C & C サーバとの通信を検知する手法を提案している．文献［55］でも，Malan らは，送信元と送信先 IP アドレス間の通信間隔を比較し，既知のボットネットからの通信挙動に類似する通信があれば送信元はボットネットに所属すると判断する手法を提案している．文献［56］では，Zhao らはネットワークフロー情報からパケットのペイロード長平均やパケット長の分散などといった統計計測結果から作成した決定木を用いることでボットネットの検知を行う手法を提案している．

こうしたブルートフォース攻撃に該当する通信は，短時間に単一の IP アドレスからの通信が大量に発生したり，ある IP アドレスからの通信が長時間続いていたりすることが多い．このような挙動を検知することでブルートフォース攻撃の発生を検知することが可能である．第 2.1 節でも述べた通り，Lazarevic らは，［4］にて Point Anomalies, Contextual Anomalies, Collective Anomallies の 3 種類について異常検知技術をを挙げ，IDS への適用例を紹介している．ネットワーク監視ログからログの傾向変化を抽出するための分析手法も［57］や［58］などにより提案されている．また，通信元の IP アドレスそのものが攻撃者により用意されたものであるかを判断し，該当する IP アドレスからの通信を遮断する手段も考えられる．現行の IPS 製品には，IP アドレスのレピュテーション（評価）機能が搭載されているものも多く（［60］［59］など），この機能では，例えば通信元 IP アドレスの国情報や各ベンダーが独自にネットワークを監視した結果等から，その通信が不正であるか否かを判断する．しかし，複数の IP アドレスから分散してログイン試行を行ったり，ひとつの IP アドレスからの長時間続いた通信であっても少ない試行回数でのログイン試行を行うようなブルートフォース攻撃の場合，こうした技術では攻撃の発生を検知することが難しく，また検知できたとしても，誤検知であると判断されることも多い．

複数の拠点での監視・分析の実施例として，NICT の nicter［61］によるダークネットトラフィックの観測や警察庁の@plice［63］，JPCERT/CC の TSUBAME［62］によるインターネット上のトラフィック観測がなされている．牧田らは複数拠点に DNS サーバをハニーポットとして設置し DNS サーバを悪用する不正活動の観測・分析を行っている［64］［65］．他にも，ネットワークトラフィック監視・分析に関する研究（［66］［67］［68］［69］など）も多数発表されている．最近では，実際のネットワーク上で，ランダムで低速なポートスキャンが発生していたことが武仲らにより報告されている［70］［71］．特に［71］では，複数の攻撃元から複数の被攻撃先に向けたランダムで低速なポートスキャンの発生が報告されている．

そこで複数の拠点で得られたセキュリティログを分析の対象とすることで，単一の拠点から得られたログや，ネットワークトラフィックだけでは検知することが難しかった攻撃事象の発生を検知することができると考えられる．

第 3 章では，複数の Web サイトを運用する Web ホスティングサービスに対するアクセスログを

対象とし，悪意あるリクエストを抽出する手法を提案し，評価した結果を述べる．第 4 章および第 5 章では，企業が管理する IP アドレス帯で実際にサービスを運用している複数の SSH サービスおよび RDP サービスに対するブルートフォース検知ログを対象とし，送信元 IP アドレスを変えながらブルートフォース攻撃を繰り返す，分散型ブルートフォース攻撃事象を抽出し，さらにこの攻撃の対策手法を提案し，評価を行う．

第3章　Webホスティングサービスに対するアクセスログを用いた悪性リクエストの検知

3.1　はじめに

複数のWebサイトを運用するWebホスティングサービスに対するアクセスログを対象とし，送信元IPアドレス，送信先ドメイン，URIの関係性に着目することで，Webアプリケーションの脆弱性の探索を目的としたリクエストや悪意あるコード挿入を行うリクエストを抽出する手法を提案する．本手法を横浜国立大学情報基盤センターが管理している学内向けWebホスティングサービスより取得できたアクセスログを用いて評価を行う．評価の結果，既存の攻撃検知手法では悪性と判断しなかったものの，他文献により悪性の可能性が高いリクエストを悪性と判断できることを示す．

3.2　本章で対象とする悪意あるリクエスト

インターネットに公開されているWebサイトには正規ユーザによるアクセスや検索エンジンによる情報収集目的のアクセス，攻撃を目的とした悪意あるアクセスなどが日々到達している．

また，今日では多くのWebサイトがWebアプリケーションを利用して運用されている．例えばCMS（Contents Management System，コンテンツ管理システム）では，Webサイト上の管理・編集画面を通して，Webサイトの概観を簡単にカスタマイズしたり，コンテンツを更新したりすることができる．特にWordPress［72］やJoomla!［73］は，オープンソースとして公開されているCMSで，個人・組織を問わず多く利用されている．さらにこれらの機能を拡張するためのプラグインも多く存在する．

しかし，こうしたCMSが攻撃の対象となる事例も報告されている．あるホスティングサービス上においてWordPressを利用したWebサイトが大量に改ざんされた事例（文献［74］）や，CMS中のスクリプトの改ざんが継続して発生していることが報告されている（文献［75］）．CMSはWebサイト毎に異なる形態・規模で運用されているものの，アプリケーションは共通のスクリプトから構成されているため，その脆弱性は多くのWebサイトに影響を与えうる．

Webサイトに対する悪意あるアクセスの検知手段として，リクエストの文字列が特定のパターンに合致すれば攻撃と判断するパターンマッチングや，普段とかけ離れたデータを含むリクエストを攻撃と判断するアノマリ検知が存在する．これらの技術は，通常，各々のWebサイトに対するリクエストあるいはリクエスト群を対象に攻撃を検知する．しかし，Webサイトに対する悪意あるアクセスの中には，一見通常のアクセスと区別が難しいものや，異常性の確認が難しいものも存在する．例えば，ゼロデイや発見されたばかりの脆弱性を突く攻撃であれば，検知パターンが対応していなければシグネチャマッチングによる検知はできない．アノマリ検知についても，不特定多

数からのアクセスを受け付ける Web サイトについては,「正常なリクエスト」の定義が難しい場合がある．

そこで我々は，脆弱性を持つ Web アプリケーションの探索や，Web サイトに対して悪意あるコード挿入を行うリクエストを，悪意あるリクエストとして検知することを考える．我々は，複数の Web サイトを監視対象とし，単一の送信元から複数の Web サイトに向けて送信されたリクエストに着目することで，リクエストの異常性を判断するアプローチを取る．脆弱性を持つ Web アプリケーションの探索や，悪意あるコード挿入を行うリクエストには，単一の送信元から送信される際にその URI にパターンが存在するものがある．例えば，phpMyAdmin に対する脆弱性を探索するとみられるリクエストには，下記 6 種類の URI を複数の Web サイトへ送信する IP アドレスが複数存在する事象が，横浜国立大学で運用している Web ホスティングサービスで観測されている[1]．

- GET //myadmin/scripts/setup.php HTTP/1.1
- GET /muieblackcat HTTP/1.1
- GET //pma/scripts/setup.php HTTP/1.1
- GET //MyAdmin/scripts/setup.php HTTP/1.1
- GET //phpMyAdmin/scripts/setup.php HTTP/1.1
- GET //phpmyadmin/scripts/setup.php HTTP/1.1

また，ShellShock の脆弱性を突く攻撃をリクエストヘッダに含むリクエストには，下記 3 種類の URI パターンで到達する攻撃が存在することが報告されている（文献［77］）．こちらの事象も横浜国立大学で運用している Web ホスティングサービスで観測されている．

- GET / HTTP/1.0
- GET /Ringing.at.your.dorbell! HTTP/1.0
- GET /Diagnostics.asp HTTP/1.0

本章では，Web アプリケーションの脆弱性の探索を目的としたリクエストや悪意あるコード挿入を行うリクエストで，送信される URI 群にパターンが存在するリクエストを悪意あるリクエストとして抽出する．正規ユーザならば，Web サイト管理サービスや検索エンジンによるクローリングを除き，目的・規模の異なる複数 Web サイトに対して全く同じコンテンツをリクエストすることは稀であると考えられる．よって，単一の送信元から複数 Web サイトに対して同一の URI が送信されたならば，当該リクエストは正規目的でなく，攻撃目的の可能性が高いと判断できる．

本章では，複数 Web サイトに対するアクセスログを対象として，複数 Web サイトに送信された悪意あるリクエストを検知する手法を提案する．提案手法は，複数の Web サイトを管理する Web ホスティングサービス管理者が，自身の管理下にある Web サイトの監視を行うため，アクセスログの分析を行う際に適用することを想定する．さらに，本章で提案する手法は，User-Agent や Referer といった送信元の身元やリクエストの起源を示す情報，リクエスト中のヘッダ情報は用いない．User-Agent や Referer の文字列を正規ユーザやクローラを模擬した文字列に設定された場合であっても，提案手法では悪意あるリクエストが抽出できるようにする．

[1]URI 中の文字列「muieblackcat」があるが，これは phpMyAdmin の脆弱性スキャナである Muieblackcat に関係するものと考えられる（文献［76］）．

3.3 提案手法

本章では，複数 Web サイトのアクセスログを対象として，複数 Web サイトに送信された悪意あるリクエストを抽出する手法を提案する．

文献 [11]で提案された鐘らの手法では，1 つの送信元 IP アドレスから多数の送信先 IP アドレスへアクセスが発生すること，Web スキャンに該当するアクセスはその URI が文字列として類似することを特徴として，Web アプリケーションの脆弱性や設定ミスを持つ Web サイト発見を目的としたリクエスト（Web スキャン）の特徴としている．

しかしこの手法では，URI が文字列として類似しないアクセスの悪性の有無を判断することが難しい．例えば，横浜国立大学で運用している Web ホスティングサービスの管理下にある 91 の Web サイトに対して同一の IP アドレスから，ShellShock の脆弱性を突く攻撃をリクエストヘッダ内に含むリクエストが到達していたことが確認されている（ [77]）．このリクエストは下記の 3 種類の URI で構成されている．

- GET / HTTP/1.0
- GET /Ringing.at.your.dorbell! HTTP/1.0
- GET /Diagnostics.asp HTTP/1.0

これらの URI は，同一の送信元 IP アドレスから送信されたこと，複数の Web サイトに対して送信された特徴を持つ．一方で，URI の文字列として見たときに，類似性が高いとはいえない．

これらの悪性 URI は，文字列としては高い類似性を有していないものの，リクエストを受信した Web サイトは，どれも同じ URI を受信していた特徴があった．そこで我々は，1 つの送信元 IP アドレスから多数の Web サイトへアクセスが発生する特徴を鐘らの手法より継承する．その上で，URI の文字列の類似性を比較する代わりに，ある送信元 IP アドレスからある送信先 Web サイトに対して送信された URI をグループ化し，そのグループが複数の Web サイトに対して共通して送信されたことを悪性リクエストか否かを判断する特徴とする．なお，提案手法では URI 間の文字列の類似性を比較しないため，送信元 IP アドレスが攻撃対象とした Web アプリケーションの種別を分類するには，鐘らの手法等で URI を分類する必要がある．

提案手法では，User-Agent や Referer といった送信元 IP アドレスの身元やリクエストの起源を示す情報は入力として用いない．User-Agent や Referer は送信側で任意の文字列を設定できる．そのため提案手法では，User-Agent や Referer の文字列を正規ユーザやクローラを模擬した文字列に設定された場合であっても，提案手法では悪意あるリクエストが抽出できるようにする．

提案手法の全体構成を図 3.1 に示す．提案手法は，複数の Web サイトを管理する Web ホスティングサービスの管理者が，自身の管理する Web サイト群のセキュリティ監視を行うことを目的として使用することを想定する．この Web ホスティングサービスでは，管理下の Web サイト群にはドメインが個別に割り当てられている．提案手法では，複数の Web サイトから取得したアクセスログを入力とし，複数の Web サイトに対して同一の URI を送信した送信元 IP アドレスを攻撃と判断する．

提案手法の入力とするアクセスログについて述べる．アクセスログには，ある送信元 IP アドレスから Web ホスティングサービス管理下のドメインに対して送信されたリクエストが 1 行のレコードとして記録される．

提案手法では，監視対象の Web サイト群に対して発生したアクセスを記録し，一定期間監視した結果得られるアクセスログを入力とする．提案手法の処理手順を述べる．まず，複数送信先ドメ

インに共通する URI を送信した送信元 IP アドレスを抽出する手順を述べる．この抽出手順では，まず複数種類の送信先ドメインに URI を送信した送信元 IP アドレスを抽出し，当該送信元 IP アドレスが送信した URI が送信先ドメイン間で一致するか否かを検証する．入力のアクセスログから送信元 IP アドレスとドメインの項目を取り出し，送信元 IP アドレス毎にドメイン種類数（送信元 IP アドレスに紐づいたドメイン種類数）を集計する．集計の結果，2 種類以上のドメインに URI を送信した送信元 IP アドレスを抽出し，複数送信先ドメインに URI を送信した送信元 IP アドレスと判断する．次に，抽出した送信元 IP アドレスを含むアクセスログを対象に，各送信元 IP アドレスについて，送信先ドメインに同じ URI を送信したか否かを検証する．各送信元 IP アドレスについて，送信先ドメインごとに URI の集合を作成する．作成した URI 集合を比較し，完全一致した場合に，当該送信元 IP アドレスは 2 種類以上の複数種類の送信先ドメインに対して同一の URI を送信したと判断する．なお，各ドメインに対するリクエストの受信時刻が同時刻の場合，リクエストの順序性を判断することが難しい．そのため，本手法では送信先ドメインが受信した URI の順番は考慮しない．

入力のアクセスログが n 行の場合における提案手法の計算量について述べる．複数種類の送信先ドメインに URI を送信した送信元 IP アドレスの抽出処理および送信先ドメインごとの URI 集合の作成処理では，レコードの探索を行うため，計算量はどちらも $O(n)$ となる．しかし，URI 集合の比較処理では，作成した URI 集の組み合わせを比較し，完全一致か否かを判断するため，計算量は $O(n^2)$ となる．従って，提案手法全体での計算量は $O(n^2)$ となる．

上述の抽出処理により，複数送信先ドメインに対して同一の URI を送信した送信元 IP アドレスを抽出できる．これらから，悪性の可能性が高い送信元 IP アドレスを絞り込む．提案手法では，送信元 IP アドレスの送信先ドメイン種類数と，送信元 IP アドレスが送信した URI 集合の種類数の，2 つの観点を悪性送信元 IP アドレスの判断基準とする．このドメイン種類数および URI 集合の種類数について，それぞれしきい値を設ける．設けたしきい値を超えるような多種類のドメインに対して，多種類の共通した URI を送信した送信元 IP アドレスは，悪性の可能性が高いと判断する．

提案手法の適用例を図 3.2 に示す．図 3.2 では，まず Step 1 で入力となるアクセスログ中の送信元 IP アドレスとドメインの 2 項目に着目し，送信元 IP アドレスごとに送信先となったドメイン種類数を集計する．この集計の結果，b.b.b.b の送信先ドメイン種類数が 2 種類であったため，この b.b.b.b を複数ドメインに送信した送信元 IP アドレスと判断する．次に，Step 2 の処理では，複数ドメインに送信した送信元 IP アドレスを含むアクセスログを対象として，ドメインごとに URI 集合を作成する．アクセスログより，b.b.b.b は ynu に対しては GET mal_scan.php および GET mal_insert.cgi を，ias に対しても GET mal_scan.php および GET mal_insert.cgi を送信していた．そのため，b.b.b.b は ynu について URI 集合 { GET mal_scan.php, GET mal_insert.cgi} が，ias についても URI 集合 { GET mal_scan.php, GET mal_insert.cgi} が作成される．そして Step 3 では，1 つの送信元 IP アドレスについて作成された URI 集合を比較する．Step 2 で b.b.b.b について，ynu について作成できた URI 集合と ias について作成できた URI 集合を比較する．これら 2 種類の URI 集合に含まれる要素は一致するため，b.b.b.b は ynu と ias に対して共通の URI 群（GET mal_scan.php, GET mal_insert.cgi）を送信したと判断する．この b.b.b.b の場合では，ドメイン種類数は 2，URI 集合の種類数は 2 となる．

本手法により，検知パターンが存在しなかったり，単一のドメインでは異常性を判断できなかったりする攻撃であっても，しきい値を超える種類のドメインに対して，しきい値を超える種類の共通した URI を送信した送信元 IP アドレスならば，提案手法では悪性として抽出することが可能である．

次節では，提案手法を実際のホスティングサービスより取得できたアクセスログに適用し，ドメ

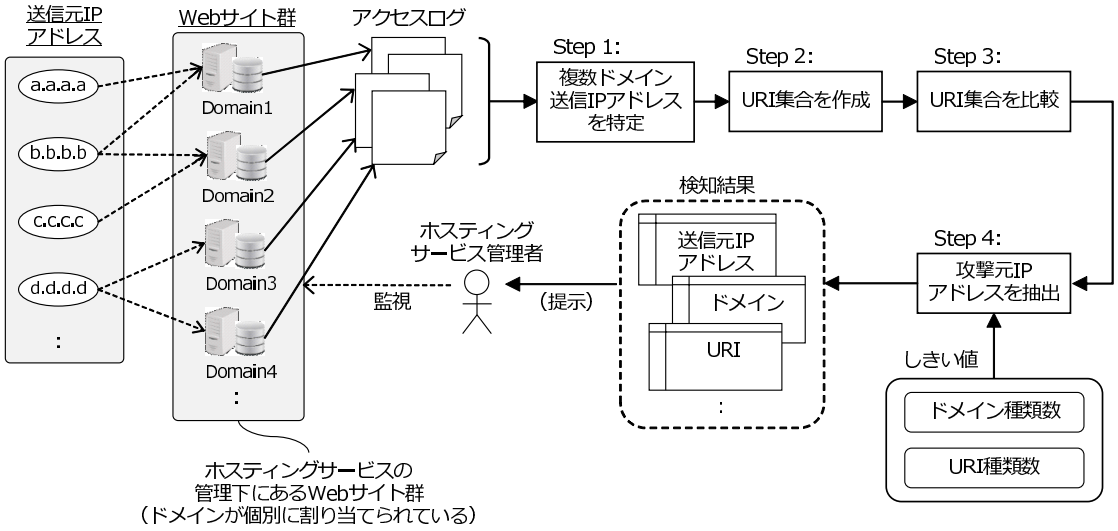

図 3.1: 提案手法の全体構成

イン種類数および URI 種類数のしきい値を変化させたときの，提案手法の検知精度を計算した結果を報告する．

3.4 評価実験

前節にて提案したアクセスログ分析手法を実際のホスティングサービスのアクセスログに適用した結果を報告する．提案手法と既存ツールである IDS の snort（［78］，［79］）と WAF の modsec（［80］）を比較した．本評価実験では，提案手法のパラメタを変化させながら，誤検知率および攻撃見逃し率を計算し，提案手法がどの程度攻撃を検知できるのかを示す．

3.4.1 適用対象とするアクセスログ

横浜国立大学では，学内組織や教職員を対象とした Web サイトホスティングサービスを運用している．このサービスでは，学内の部局や研究室等の Web サイトが運用されており，学内外からのアクセスをアクセスログに記録している．

適用対象としたのは 2015 年 6 月 26 日から 9 月 23 日の 90 日間に発生したアクセスのうち，検索エンジンによるクローラと判断された送信元 IP アドレスによるアクセスを除外したものである．クローラの判断においては，まず Baidu［82］，Google［83］［84］，msn［85］を対象として，逆引きドメイン名が表 3.1 に該当する送信元 IP アドレスをクローラによるものと判断した．

Baidu と msn と同じ URI 集合を送信した送信元 IP アドレスもクローラと判断した．適用対象としたアクセスログの規模は，レコード件数が 44,177,946 件，送信元 IP アドレスが 63,084 種類，ドメイン種類数が 255 種類，URI が 1,062,433 種類であった．

本章の実験では，提案手法により（a）1 日単位で攻撃元を検知する場合，(b）1 時間単位で攻撃元を検知する場合，の 2 通りを実験した．

表 3.1: クローラの判断基準

検索エンジン	ドメイン名における判断基準
Baidu	crawl.baidu.com で終わる．
Google	googlebot.com で終わる．あるいは rate-limited-proxy-で始まり google.com で終わる．
msn	search.msn.com で終わる．

3.4.2 正解データの作成

他ツールにより悪意ありと判断されたリクエストを送信した送信元 IP アドレスを抽出する手順について述べる．本手順は 3 つの手順からなる．

Step 1: 他ツールにより悪性として検知されたリクエストを送信した送信元 IP アドレスの抽出

Step 1 では，URI を既存ツールに適用し，悪性と判断されるリクエストの収集を行った．

まず，実験マシン内でテスト用 Web サイトと，送信対象 URI 一覧に格納されたリクエスト文字列を HTTP リクエストとしてテスト用 Web サイトに対して送信するリクエスト生成器を構築した．構築した Web サイトには文字列「Hello World!」が書かれた index.html のみ設置されている．このリクエスト生成器を用いて，評価対象 URI 一覧へアクセスログ中の URI を登録し，テスト用 Web サイトに対して HTTP リクエストを送信する．テスト環境の構成を図 3.3 に示す．リクエスト生成器では，評価対象 URI 一覧に格納された URI に対してホスト名やヘッダ情報を補完し，テスト用 Web サイト（http:localhost）に対する HTTP リクエストを生成し，送信する．この Web サイトに対するアクセスを，オープンソースの IDS である Snort および WAF の modsec を適用して監視し，これらのツールが HTTP リクエストを検知するかを記録した[2][3]．

各ツールのどちらか一方が攻撃と検知したならば，当該 URI を攻撃と判断した．我々は，このテスト環境を用いて適用対象としたアクセスログに記録された URI に対する snort および modsec による検知結果を得た．

Step 2: 独自シグネチャの作成およびマッチング

Step 1 では Snort および modsec により Web サイトへの攻撃に関する正解データの作成を試みたが，Web サイトへの攻撃は多岐にわたるため，これらの既存ツールにも見逃しが予想される．そこでこれらのツールに加えて，セキュリティサイト等の情報を元に攻撃パターンを正規表現として手動で書き下した独自シグネチャを作成した．独自シグネチャを表 C.1 に示す．

Step 3: 悪意ある送信元 IP アドレスの判断

Step 1 および Step 2 の手順により悪意ありと判断された URI と，送信元 IP アドレスが送信した URI とを比較する．URI が全て攻撃と判断されたならば，当該送信元 IP アドレスは悪意ある URI を送信したと判断した．

[2]Snort ルールセットの取得日は 2016 年 12 月 21 日，modsec ルールセットの取得日は 2017 年 2 月 28 日である．
[3]実験の際には，リクエスト文字列を判断対象とするように調整したルールセットを使用した．

3.4.3 正解データとして抽出できた送信元 IP アドレス

上述の手順に従って正解データを作成した結果，適用対象アクセスログの期間を（a）1 日とした場合では合計 9,364 種類の送信元 IP アドレスが，（b）1 時間とした場合では合計 15,967 種類の送信元 IP アドレスが，それぞれ抽出できた．適用期間 (a) と (b) それぞれについて，期間ごとに抽出できた全送信元 IP アドレスの Web サイト種類数，抽出できた送信元 IP アドレスの変化および URI 種類数の分布を，(a) の場合をそれぞれ図 3.4，図 A.1，図 A.2 に，(b) の場合をそれぞれ図 3.5，図 A.3，図 A.4 に示す．

3.4.4 提案手法の適用結果抽出できた送信元 IP アドレス

アクセスログを提案手法に適用した結果，適用対象アクセスログの期間を（a）1 日とした場合では合計 9,551 種類の送信元 IP アドレスが，（b）1 時間とした場合では合計 15,361 種類の送信元 IP アドレスが，それぞれ抽出できた．第 3.1 節で述べた，文献 [77] で報告された ShellShock の脆弱性を狙った URI 群を送信した IP アドレスも，(a) と (b) それぞれの場合において提案手法で抽出できたことを確認した．表 3.2 に，この URI 群の送信元となった IP アドレス種類数と送信先ドメイン種類数を示す．表 3.2 では例えば，(a) の場合に，7 種類の送信元 IP アドレスが 90 種類のドメインに対してこの URI 群を送信していたことを示す．

適用期間 (a) と (b) それぞれについて，期間ごとに抽出できた全送信元 IP アドレスの Web サイト種類数，抽出できた送信元 IP アドレスの変化および URI 種類数の分布を，(a) の場合をそれぞれ図 3.6，図 B.1，図 B.2 に，(b) の場合をそれぞれ図 3.7，図 B.3，図 B.4 に示す．

正解データとして抽出できた送信元 IP アドレスと提案手法の適用結果抽出できた送信元 IP アドレスを対象に，単位期間当たりの送信元 IP アドレス数，ドメイン種類数，URI 種類数について，最小値，最大値および平均値を比較した結果を表 3.3 に示す．まず，提案手法で抽出できた送信元 IP アドレスと正解データとして抽出した送信元 IP アドレスを比較する．単位期間当たりに抽出できた送信元 IP アドレス数については，(a) と (b) のどちらの場合でも，提案手法よりも正解データの方が多かった．また，URI 種類数についても，提案手法よりも正解データの方が最大値・平均値ともに多かった．URI 種類数の分布に着目すると，特に (b) の場合に，正解データでは 2〜3 種類の URI を送信した送信元 IP アドレスが全体の約 19.4%存在した．これに対し，提案手法では 2〜3 種類の URI を送信した送信元 IP アドレスが全体の約 78.5%であった．このことから，1 種類のみのドメインに URI を送信した送信元 IP アドレスが多く，それらが送信した URI 種類数も多かったことがわかる．図 3.6 および図 3.7 に示したドメイン種類数の分布からも，URI が共通か否かに関わらず，2 種類以上のドメインに URI を送信した送信元 IP アドレスは正解データ全体と比較すると少ないことが分かる．次に，提案手法で抽出できた送信元 IP アドレスについて，(a) と (b) の場合を比較すると，ドメイン種類数の平均値の差は約 0.99 であった．ドメイン種類数の分布を比較すると，(a) よりも (b) の場合の方が，Web サイト種類数が 2 だった送信元 IP アドレスが占める割合が大きかった．一方で，URI 種類数に関しては，平均値の差は約 0.06 であった．このことから，複数種類のドメインに対して共通した URI を送信した送信元 IP アドレスの多くは，1 時間以内に URI を送信し終えていたといえる．

表 3.2: (a) と (b) の場合において，提案手法で抽出できた ShellShock の脆弱性を狙った URI 群に紐づく送信元 IP アドレスと送信先ドメイン

	送信元 IP アドレス種類数	送信先ドメイン種類数
(a)	7	90
	1	4
	1	91
(b)	6	90
	2	4
	1	91
	1	87

表 3.3: 評価対象となった送信元 IP アドレスの基本統計量

			最小値	最大値	平均値
送信元 IP アドレス数	(a)	提案手法	30	454	106.12
		正解データ	17	560	104.04
	(b)	提案手法	2	26	7.17
		正解データ	1	117	7.87
ドメイン 種類数	(a)	提案手法	2	91	4.97
		正解データ	2	91	4.37
	(b)	提案手法	2	91	3.98
		正解データ	2	91	3.36
URI 種類数	(a)	提案手法	1	304	1.33
		正解データ	1	96	1.78
	(b)	提案手法	1	337	1.27
		正解データ	1	96	2.17

3.4.5 提案手法の実行時間

適用期間 (a) と (b) それぞれについて，期間ごとに提案手法の入力としたアクセスログのレコード件数と実行時間の関係を，(a) の場合を図 3.8 に，(b) の場合を図 3.9 に，それぞれ示す．提案手法は Python2.7.10 にて実装し，Intel(R) Xeon(R) CPU E3-1240 V2 3.40GHz および RAM 16GB を搭載したマシン上で実行した．

計測の結果，入力となるアクセスログのレコード件数は，(a) の場合に最小で 281,128 件，最大で 2,383,006 件，平均で 1,063,119 件であり，(b) の場合に最小で 4,490 件，最大で 163,432 秒，平均で 42,995 件であった．これに対し，提案手法の実行時間は，(a) の場合に最小で 1,603 秒，最大で 17,180 秒，平均で 6,030 秒であり，(b) の場合に最小で 15 秒，最大で 1,312 秒，平均で 191 秒であった．また，図 3.8 および図 3.9 から，入力となるアクセスログのレコード件数が大きくなるほど，実行時間も大きくなっていることがわかる．

従って，本章での実験環境では，処理の実行時間の平均からは，(a) と (b) のどちらの場合においても，アクセスログを提案手法で処理した結果は，次の提案手法の処理開始までに得られることがわかり，毎日あるいは毎時間を想定した定期的な分析を行う運用が実現できることを確認できた．

3.4.6 FP および FN の計算結果

提案手法を適用した結果について，ドメイン種類数のしきい値 S および URI 種類数のしきい値 R を変化させたときの誤検知率（False Positive，FP）と見逃し率（False Negative，FN）を，それぞれヒートマップ形式で可視化した.FP および FN は次式により計算した．

- FP :（Snort，modsec，独自シグネチャのいずれかの手段が検知しなかったものの，提案手法が検知した送信元 IP アドレス数）/（提案手法が検知した送信元 IP アドレス数）
- FN :（Snort，modsec，独自シグネチャのいずれかの手段が検知したものの，提案手法が検知しなかった送信元 IP アドレス数）/（Snort，modsec，独自シグネチャのいずれかの手段が検知した送信元 IP アドレス数）

(a) 1 日で区切った場合における FP の変化を図 3.10 に，FN の変化を図 3.11 に，(b) 1 時間で区切った場合における FP の変化を図 3.12 に，FN の変化を図 3.13 に，それぞれ示す．ただし，ヒートマップ内の白色の領域は，該当する送信元 IP アドレスが存在しなかったことを示す．

ヒートマップ中の色の分布から，(a) と (b) のどちらの場合においても，総じて誤検知率は小さいものの，見逃し率は高い結果となった．FP の変化に関しては，(a) と (b) の両方において，S も R もどちらも小さい値に設定した場合には，FP が大きくなった．FN の変化に関しては，(a) と (b) の両方において R が 84 以上のときには見逃し率が 1 となった．これは，提案手法では URI 種類数が 84 以上であった送信元 IP アドレスを抽出できなかったことを示している．

次に，(a) と (b) の場合において FP および FN がそれぞれ 0 となる S と R の範囲を表 3.4 に示す．特に FP が 0 となった範囲は，(a) と (b) の場合どちらも同じ範囲で，次の 3 種類であった．

- 領域 FP_1: S=91, $4 \leqq$ R $\leqq 6$
- 領域 FP_2: $5 \leqq$ S $\leqq 90$, $16 \leqq$ R $\leqq 44$
- 領域 FP_3: $3 \leqq$ S $\leqq 90$, $45 \leqq$ R $\leqq 83$

特に領域 FP_2 および領域 FP_3 については，$88 \leqq$ S $\leqq 90$，$29 \leqq$ R $\leqq 83$ の範囲においては，FP だけでなく，FN も 0 となった．

したがって，攻撃を誤検知なしに検知可能な範囲として，下記の 2 種類の範囲が設定可能であるといえる．まず，送信対象のリクエスト種類数が少なくとも，ドメイン種類数が多かった範囲である．本評価実験では，領域 FP_1 である，ドメイン種類数が 91 以上且つ URI 種類数が 4 以上に該当する．次に，ドメイン種類数が少なくとも，送信対象の URI 種類数が多かった範囲である．本評価実験では，領域 FP_2 および領域 FP_3 が重なる領域である，ドメイン種類数が 5 以上且つ URI 種類数が 45 以上に該当する．

FP=0 となった領域について，独自シグネチャにより悪性と判断できた送信元 IP アドレス（独自シグネチャ含有送信元 IP アドレス）の割合を，表 3.5 に示す．この表から，(a) と (b) どちらの場合も，領域 FP_1 および領域 FP_3 は既存ツールで検知できた攻撃のみが含まれていた．一方，領域 FP_2 には，既存ツールでは検知できなかった，独自シグネチャでのみ検知できた攻撃も含まれていた．この表の結果から，FP=0 となるしきい値を設定したとき，既存のツールでは検知できなかった攻撃を，提案手法では検知できたことが確認できた．

独自シグネチャでのみ検知できた攻撃について述べる．領域 FP_2 には 2 種類の攻撃を行った送信元 IP アドレスが含まれていた．まず，Asset Manager を通して任意のファイルのアップロードが可

能な assetmanager モジュールを見つけるため（[86]），assetmanager.asp，assetmanager.aspx のように拡張子を変えながら URI を送信した送信元 IP アドレスである．次に，FCKeditor を通して任意のファイルのアップロードが可能な connecter モジュールについて，connector.asp，connector.aspx のように拡張子を変えながら URI を送信した送信元 IP アドレスである．Snort には拡張子”.asp”を含む URI を検知するシグネチャが存在したため，これらの送信元 IP アドレスが送信した URI の中には Snort でも攻撃と判断した URI が含まれていた．しかしそれ以外の拡張子を含む URI は攻撃と判断されなかった．また，connector モジュールに対して URI を送信した送信元 IP アドレスには，Snort がマルウェアによるファイルのアップロード試行を行うために攻撃と検知される URI を含む送信元 IP アドレスも存在した．以上の結果から，作成したシグネチャに不足がある場合であっても，提案手法では悪意ある URI を送信した IP アドレスを抽出できることを確認できた．

表 3.4: (a) および (b) の場合における FP=0，FN=0 となる S と R の範囲

		S の範囲	R の範囲	FP		FN	
				最小値	最大値	最小値	最大値
(a)	FP=0	S=91	4 ≦ R ≦ 6	0	0	0.714	0.714
		5 ≦ S ≦ 90	16 ≦ R ≦ 44	0	0	0	0.833
		3 ≦ S ≦ 90	45 ≦ R ≦ 83	0	0	0	0.75
	FN=0	S=90	9 ≦ R ≦ 28	0	0.6	0	0
		88 ≦ S ≦ 90	29 ≦ R ≦ 83	0	0	0	0
(b)	FP=0	S=91	4 ≦ R ≦ 6	0	0	0.333	0.429
		5 ≦ S ≦ 90	16 ≦ R ≦ 44	0	0	0	0.857
		3 ≦ S ≦ 90	45 ≦ R ≦ 83	0	0	0	0.8
	FN=0	S=90	9 ≦ R ≦ 28	0	0.6	0	0
		88 ≦ S ≦ 90	29 ≦ R ≦ 83	0	0	0	0

表 3.5: FP=0 の領域における独自シグネチャにより悪性と判断できた送信元 IP アドレスの割合

	領域	IP アドレス含有割合
(a)	領域 FP_1	0
	領域 FP_2	1
	領域 FP_3	0
(b)	領域 FP_1	0
	領域 FP_2	0.5
	領域 FP_3	0

3.4.7 リクエストの収集に必要な時間

提案手法が送信元 IP アドレスの悪性の有無を判断するため，リクエストの収集に必要な時間を評価する．適用対象アクセスログの期間 (a) および (b) の場合において，提案手法で抽出できた送信元 IP アドレスがアクセスログに出現した最後の時刻から最初の時刻を引いた値をリクエスト送信時間として計算した．

リクエスト送信時間の分布を，(a) の場合，(b) の場合をそれぞれ図 3.14，図 3.15 に示す．各図中の黒色の棒は提案手法で抽出できた送信元 IP アドレスを，灰色の棒は提案手法で抽出でき，且つ正解データに該当する送信元 IP アドレスを示す．また，(a)，(b) それぞれの場合において累積度数が 50%，75%，95%以上の下限となるリクエスト送信時間を表 3.6 に示す．表 3.6 では，例えば，(a) の場合に提案手法で抽出できた送信元 IP アドレスが累積度数 50%以上となる，つまり送信元 IP アドレスの占める割合が半分以上となる下限は，リクエスト送信時間が 3,500 秒であることを示す．

まず (a) の場合について，リクエスト送信時間が 3,000 秒以下であった送信元 IP アドレスは，提案手法で抽出できたものが約 55.8%，正解データに該当するものが約 55.5%であった．どちらの送信元 IP アドレスも半数強が，リクエスト送信時間が 3,000 秒以下であったといえる．正解データに該当する送信元 IP アドレスを 95%以上含むのは，リクエスト送信時間が 69,000 秒以下の範囲である．

次に (b) の場合について，リクエスト送信時間が 100 秒以下だった送信元 IP アドレスは，提案手法で抽出できたものが約 49.7%，正解データに該当するものが約 76.7%であった．さらに，提案手法で抽出できた送信元 IP アドレスは 3,400 秒より大きく 3,500 秒以下の範囲にも多く集中していた．正解データに該当する送信元 IP アドレスを 95%以上含むのは，リクエスト送信時間が 2,300 秒以下の範囲である．

なお，(a) の場合においてリクエスト送信時間が長かった送信元 IP アドレスには，下記のように，Web ホスティングサービス下の Web サイトにリクエストが断続的に到達したものが存在した．

この事例では，設定された Web サイト一覧に対して順番にリクエストを送信しており，その一覧に今回評価対象とした Web ホスティングサービス管理下の Web サイトが含まれていたと考えられる．これらのリクエストを送信した IP アドレスを提案手法で判断できるのは，1 行目のリクエストと，それが到達してから 4 時間 12 分後に到達した 2 行目のリクエストを読み込んだ時点となる．あるいは，数時間情報を保持しない場合であっても，2 行目のリクエストと，それが到達してから 7 秒後に到達した 3 行目のリクエストを読み込んだ時点でも，提案手法では IP アドレスの判断が可能である．ただしこの場合，1 行目のリクエストは提案手法では判断の対象とはならない．

表 3.6: (a) および (b) の場合における累積度数が 50%，75%，95%以上の下限となるリクエスト送信時間

		累積度数		
		50%以上	75%以上	95%以上
(a)	提案手法で抽出できた	3,000 秒	21,000 秒	69,000 秒
	正解データに該当	3,000 秒	21,000 秒	69,000 秒
(b)	提案手法で抽出できた	500 秒	3,400 秒	3,600 秒
	正解データに該当	100 秒	100 秒	2,300 秒

3.4.8 提案手法で観測すべきドメイン種類数

提案手法適用時に観測すべきドメインの規模について評価する．提案手法では観測対象とするドメインの種類が多ければ多いほど，より高い精度で悪意あるリクエストを送信した IP アドレスを抽出できる可能性が大きくなる．しかし，ドメインの種類が多くなれば提案手法が読み込むべきアクセスログの量や処理内で行われる URI 集合のドメイン間の比較回数が増加するため，提案手法の処理に必要なメモリや時間も増加する．

適用対象アクセスログの期間 (a) および (b) の場合において，送信先となったドメインそれぞれについて，提案手法で抽出できた送信元 IP アドレス数を，提案手法で抽出できた送信元 IP アドレスおよび提案手法で抽出でき，且つ正解データに該当する送信元 IP アドレスについてそれぞれ数え上げた．数え上げた結果を，提案手法で抽出できた送信元 IP アドレスについて昇順で並び替えた結果を図 3.16，図 3.17 に示す．各図中の三角形のマーカーは提案手法で抽出できた送信元 IP アドレスを，四角形のマーカーは提案手法で抽出でき，且つ正解データに該当する送信元 IP アドレスを示す．

まず (a) の場合について，提案手法で抽出できた送信元 IP アドレスの送信先ドメインを数え上げた結果，上位 5 種類のドメインは 1,000 を超える IP アドレスから URI が送信されていた．さらに，上位 94 番目と 95 番目のドメインを送信先とした IP アドレス数には 196 の差があり，それ以降のドメインには 1 個の送信元 IP アドレスのみが紐づく状況となっている．正解データに該当する送信元 IP アドレスの場合についても，同様の箇所で折れ線グラフが落ち込んでいる．

次に (b) の場合について，提案手法で抽出できた送信元 IP アドレスの送信先ドメインを数え上げた結果，上位 5 種類のドメインは 2,400 を超える IP アドレスから URI が送信されていた．さらに，上位 90 番目と 91 番目のドメインを送信先とした IP アドレス数には 104 の差があり，それ以降のドメインには 1 個の送信元 IP アドレスのみが紐づく状況となっている．正解データに該当する送信元 IP アドレスの場合についても，同様の傾向がある．

以上より，本実験の対象とした Web ホスティングサービス管理下の Web サイト（ドメイン）は 255 種類であったが，提案手法に関与したドメインは (a) の場合で 236 種類，(b) の場合で 231 種類であった．つまり，実際は Web ホスティングサービス管理下のドメイン全てに悪意あるリクエストが送信されていなかった．さらに，(a) の場合では 95 番目，(b) の場合では 91 番目以降のドメインに対して URI を送信した IP アドレス数は，それら以前と比較すると 200 規模の差が開いていた．このことから，これらのドメインは観測対象から除外するといった，観測すべきドメインの選定も可能である．

3.5 議論

3.5.1 提案手法が誤検知と判断する URI

本節では，誤検知に該当した送信元 IP アドレスについて事例を挙げ，提案手法が誤検知と判断する URI について考察する．

誤検知に該当した送信元 IP アドレスが送信した URI 集合を集計し，該当する送信元 IP アドレスが多かった URI 集合の上位 5 を表 3.7 に示す．ただし，表中の「null」は URI の値がログに存在しなかったことを示す．表 3.7 中の，#4 については文献 [87] より Microsoft Internet Explorer 11 等によるサムネイル画像の要求を含むリクエスト，#5 については Web サイトの情報を収集するクローラによるリクエストにより送信された URI である可能性が高いと判断できる[4]．また，#3 については，複数の Web サイトを人間が閲覧するためにアクセスしたケースも含まれると考えられる．本実験の評価対象は学内の部局や研究室等の Web サイトに対するアクセスログである．そのため，数種類程度ならば人間が手動でトップページをアクセスして回った，といったシナリオも考えられる．

[4] robots.txt や sitemap.xml は，Web サイトの情報を収集する目的で攻撃者が利用する場合も考えられるため，クローラの挙動として除外せず，本章では提案手法の適用対象とした．

以上より，提案手法の適用結果，誤検知となる URI は主に下記の 3 種類の場合であると考えられる．

1. Web ブラウザ等送信元側の環境で自動的に送信されるリクエスト
2. 複数 Web サイトの情報収集を目的とするリクエスト
3. 複数 Web サイト間で共通の名前を持つページを要求するリクエスト

こうした誤検知に対して，(1) および (3) の場合に関しては，提案手法の処理の対象外とする URI をホワイトリスト化によって，誤検知を減らせる場合がある．

まず，(1) の場合に関しては，送信元側の環境で自動的に送信される URI であることが明らかになれば，これらの URI をホワイトリストに追加できる．次に，(3) の場合に関しては，今回の実験で利用したアクセスログは大学内の Web サイトを運用する Web ホスティングサービスであった．そのため，例えば学内の部局の Web サイトにアクセスした際に，大学のロゴ画像や部局間で共通して利用されるレイアウトに関する JavaScript や CSS ファイルをリクエストしたとみられるケースがあった．こうしたコンテンツをリクエストする際に送信される URI も，ホワイトリスト化が可能であると考えられる．

ただし，上述でホワイトリスト化した URI が Web サイトの脆弱性発見に有用であることがわかったなど，攻撃に用いられることがわかった場合に，攻撃目的で URI が送信されたとしても，提案手法では処理の対象外となり，攻撃事象の抽出ができない．また，今回の Web ホスティングサービスのようにサービス利用者や利用目的が限定されていない，不特定多数の利用者が各々異なる目的でサービスを利用する形態の Web ホスティングサービスでは，特に (3) の場合のホワイトリスト化は難しいと考えられる．

表 3.7: 誤検知に該当する送信元 IP アドレスが送信した URI 集合上位 5

#	URI 集合	送信元 IP アドレス数	Web サイト種類数		
			最小値	最大値	平均値
1	GET:/favicon.ico	1,005	2	14	2.33
2	HEAD:/	258	2	91	8.36
3	GET:/ GET:/favicon.ico	200	2	10	2.82
4	null GET:/browserconfig.xml	179	2	4	2.27
5	GET:/ GET:/robots.txt GET:/sitemap.xml	77	2	9	2.44

3.5.2 提案手法の限界

提案手法の限界について考察する．提案手法では，2 以上の複数種類のドメインに対して，共通の URI を送信した送信元 IP アドレスを悪意ありと判断する．そのため，1 種類のドメインにしか URI を送信しなかった送信元 IP アドレス，複数種類のドメインに送信した URI が共通でなかった

送信元 IP アドレスは，提案手法では抽出の対象から除外される．よって，提案手法では抽出が難しい悪性 IP アドレスは，下記の 2 種類であるといえる．

1. 特定の 1 種類の Web サイトに対して悪意あるリクエストを送信した送信元 IP アドレス
2. 複数種類の Web サイトに対してそれぞれ異なる悪意あるリクエストを送信した送信元 IP アドレス

提案手法を拡張することで，2. に該当する送信元 IP アドレスも悪性と判断できる場合がある．提案手法では，Web サイトごとに送信した URI 集合を比較する際に，各 URI 集合が完全に一致した送信元 IP アドレスを抽出していた．この処理を，完全一致ではなく，類似度を計算したうえであるしきい値以上の類似度を持つ送信元 IP アドレスを抽出するような処理に拡張することで，同一の URI を送信していなくても，ある程度共通した URI 群を送信していた送信元 IP アドレスも抽出結果に含めることが可能である．ただしこのような手段をとった場合であっても，Web サイトごとにまったく異なる URI を送信した送信元 IP アドレスは抽出の対象とすることはできないといった限界が存在する．例えば Web サイトによって GET リクエスト中のパラメータを変化させながら URI を送信した送信元 IP アドレスは，同じ意図を持った URI を送信したとしても，提案手法では抽出できないという限界がある．

3.6 まとめと今後の課題

本章では，複数ドメインに対するアクセスログを対象として，複数ドメインに送信された悪意あるリクエストを抽出する手法を提案した．提案手法では，一定期間収集したアクセスログから，送信元 IP アドレス，送信先ドメイン，URI の関係性を分析し，複数のドメインに対して同一の URI を送信した IP アドレスを抽出する．本手法により，単一のドメインへのアクセスを分析することでは異常性の判断が難しいリクエストであっても，しきい値を超える種類のドメインに対して，しきい値を超える種類の共通した URI を送信した送信元 IP アドレスならば，提案手法では悪性として抽出することが可能である．

提案手法を，実際の Web ホスティングサービスより取得できたアクセスログに適用した結果，攻撃元となった IP アドレスを誤検知なく抽出できるしきい値が存在することを示した．さらに，既存のオープンソースの IDS および WAF ではシグネチャが登録されておらず検知できない攻撃についても検知できる事例を確認できた．

今後の課題として次の 2 点がある．まず 1 点目として，リクエスト群の時系列的な関係性にも着目することで，例えばあるホストからスキャンを受けた後に別のホストから攻撃を受けるといった，複数端末を連携させた攻撃事象の抽出が挙げられる．

次に 2 点目として，提案手法の処理の高速化が挙げられる．本章での実験環境における処理時間の計測結果からは，毎日あるいは毎時間を想定した定期的な分析を行う運用が実現できることを確認できた．しかし，提案手法の計算量は入力となるアクセスログのレコード件数 n に対して $O(n^2)$ であるため，レコード件数が増加すると処理時間も増加する可能性がある．そのため，全体の処理時間を短縮できるよう，並列化に対応させる等の処理の改良が必要である．

図 3.2: 提案手法における送信元 IP アドレス抽出処理の例

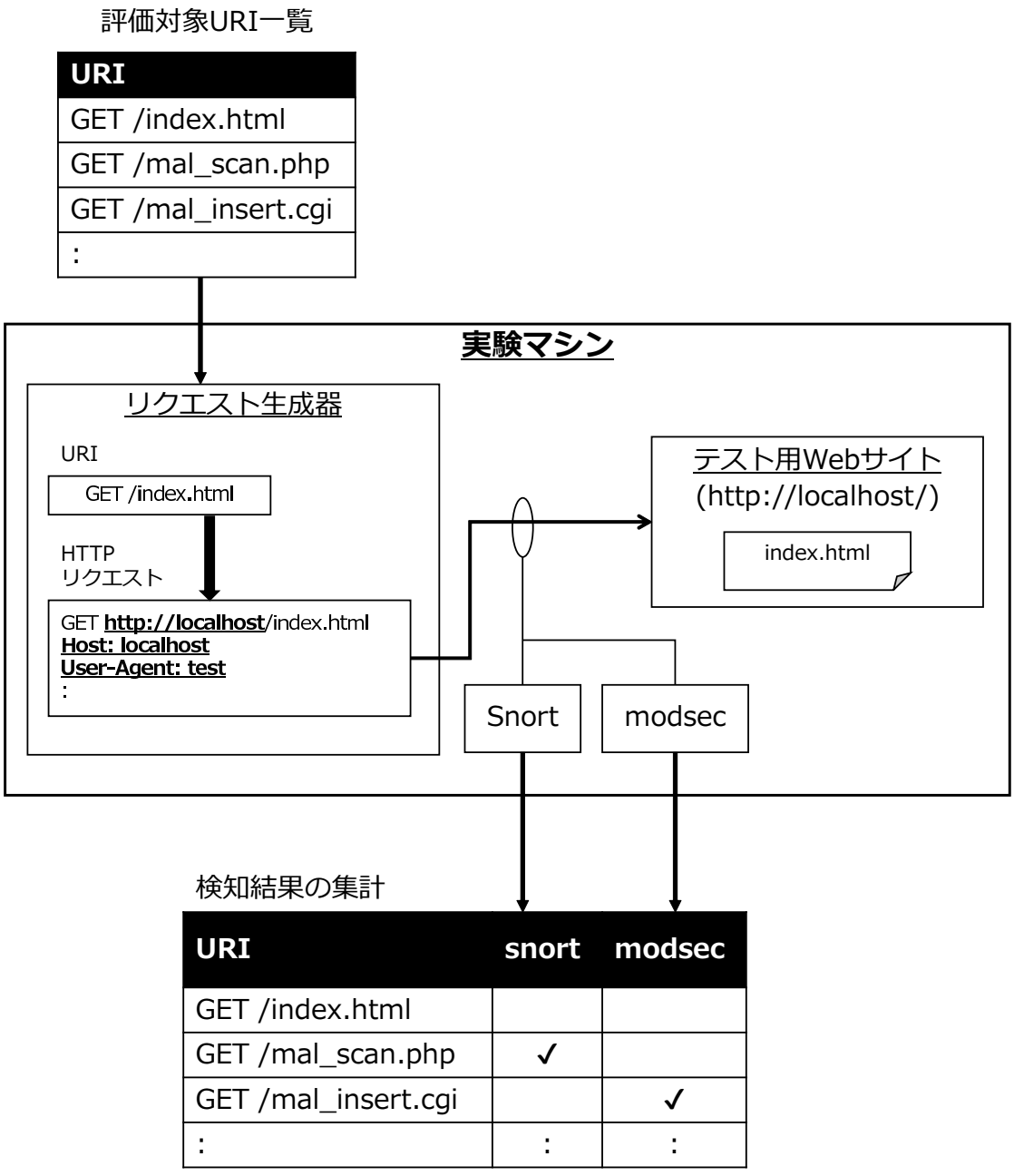

図 3.3: 既存ツールとの比較を行った環境

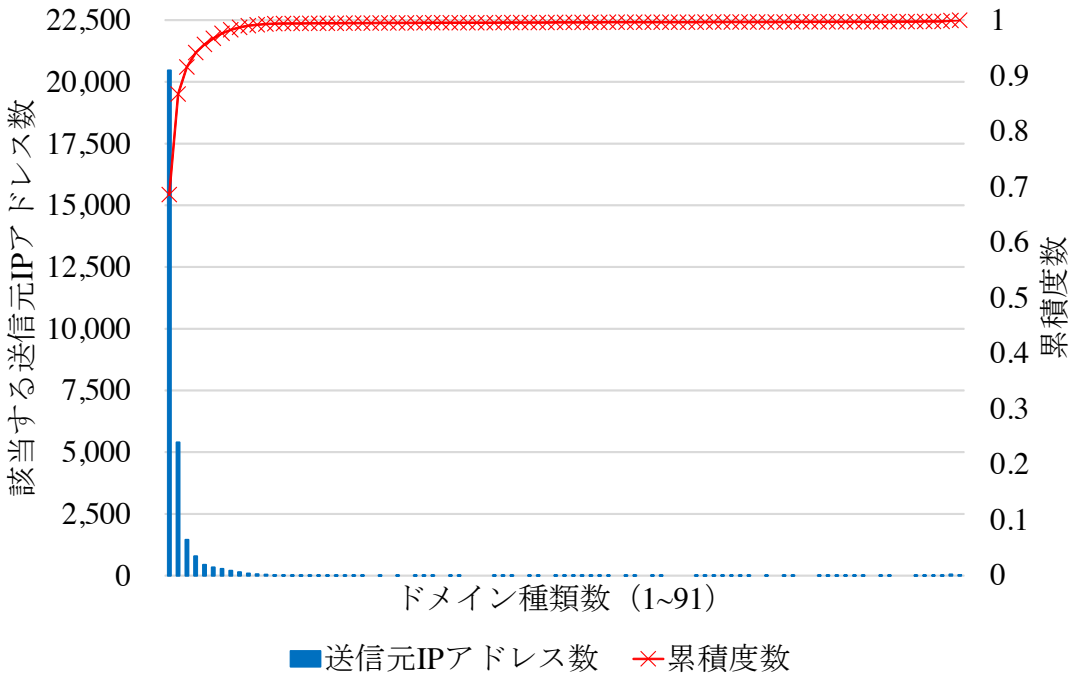

図 3.4: (a) の場合に正解データと判断した送信元 IP アドレスのドメイン種類数の分布

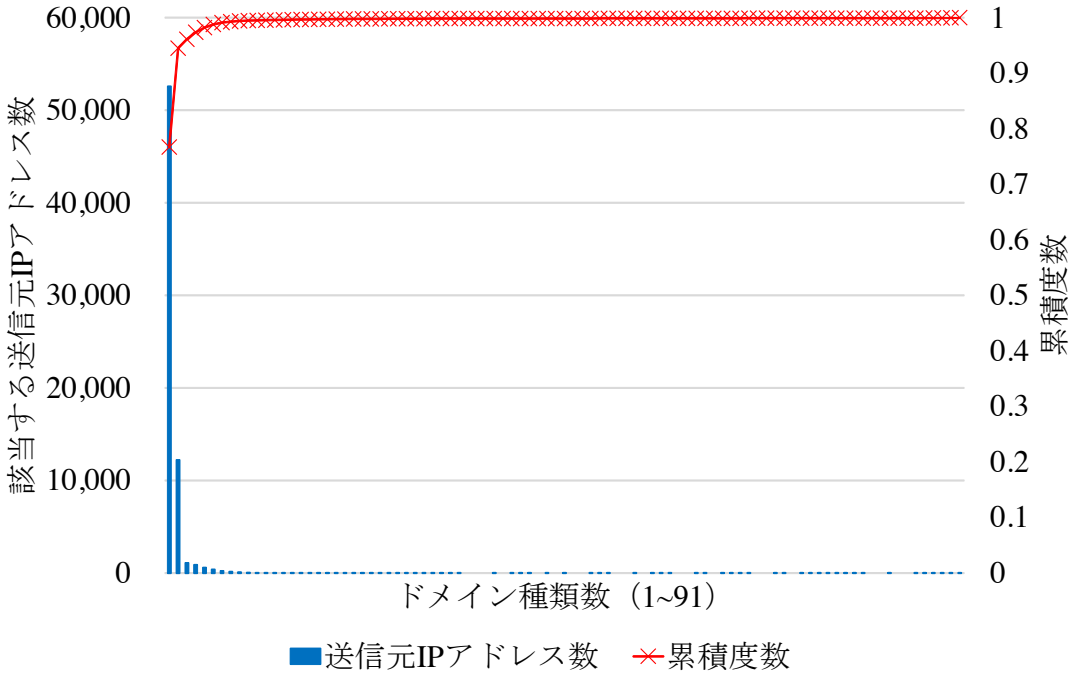

図 3.5: (b) の場合に正解データと判断した送信元 IP アドレスのドメイン種類数の分布

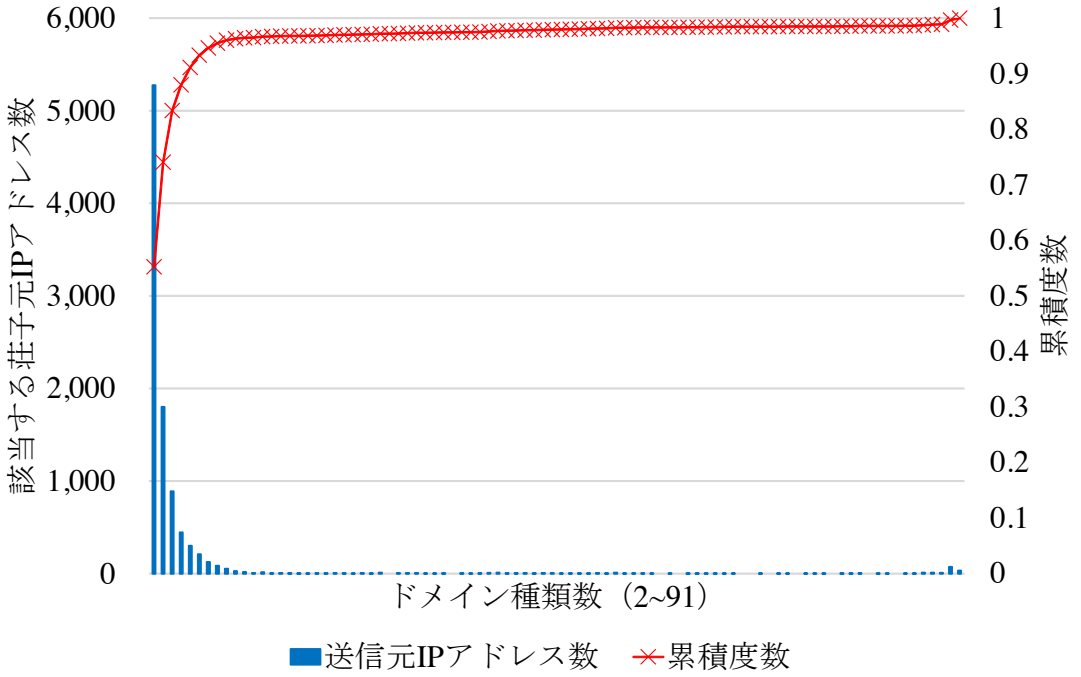

図 3.6: (a) の場合に提案手法が抽出した送信元 IP アドレスのドメイン種類数の分布

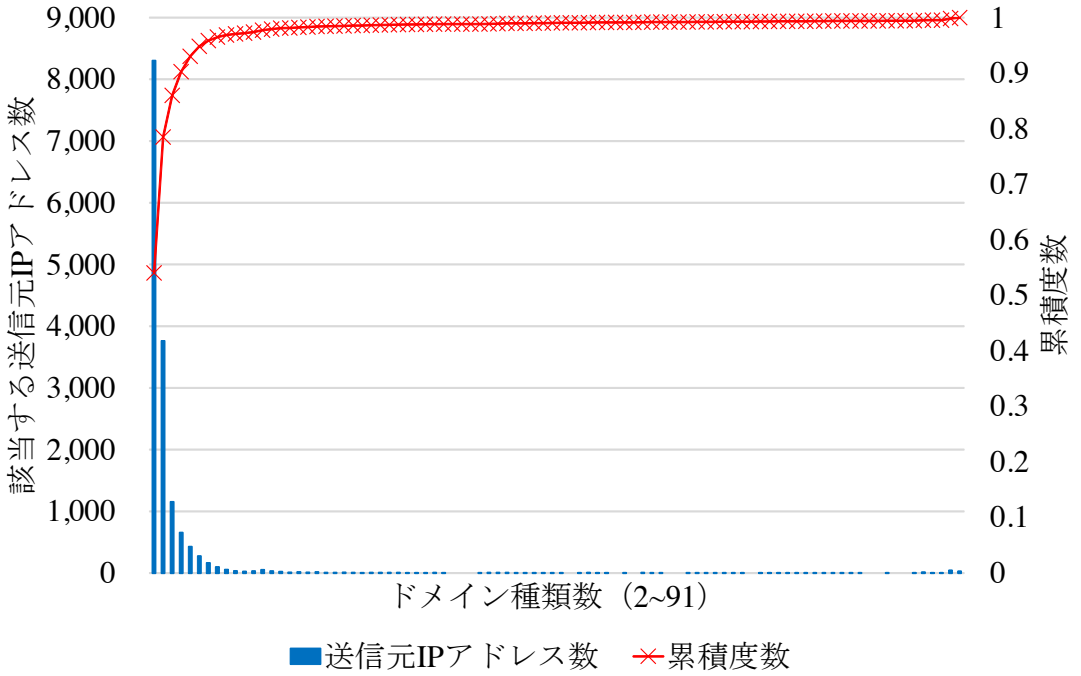

図 3.7: (b) の場合に提案手法が抽出した送信元 IP アドレスのドメイン種類数の分布

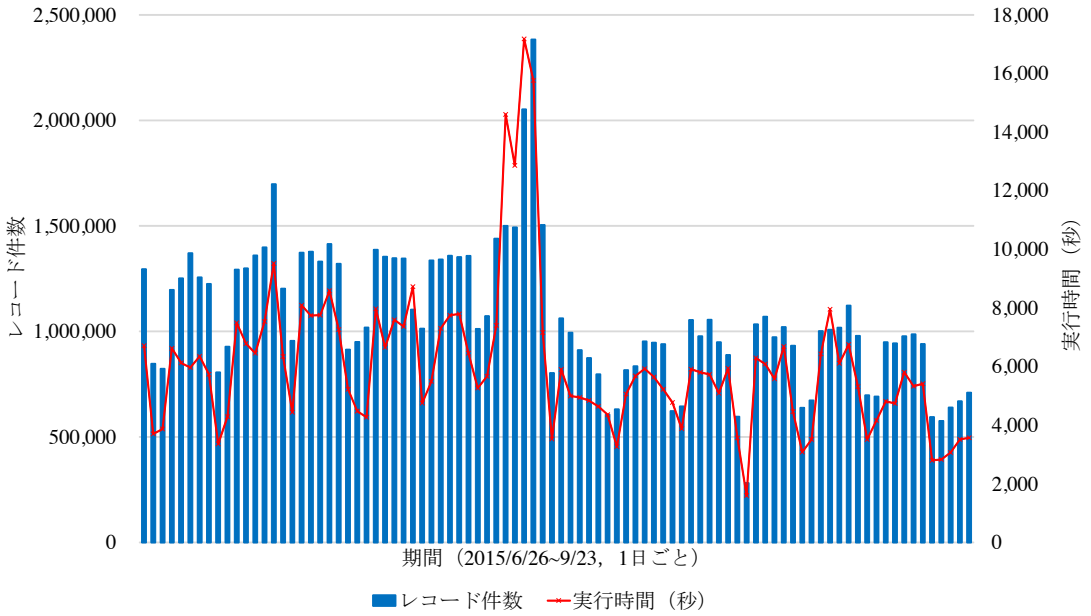

図 3.8: (a) の場合における期間ごとのアクセスログのレコード件数と処理の実行時間

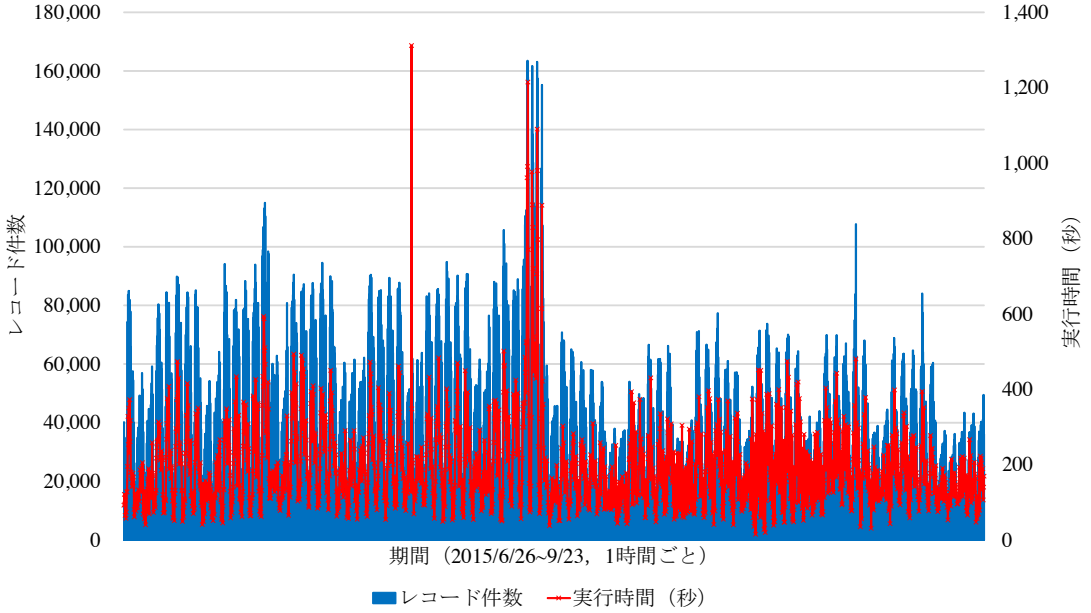

図 3.9: (b) の場合における期間ごとのアクセスログのレコード件数と処理の実行時間

(1) 誤検知率 (False Positive, FP)

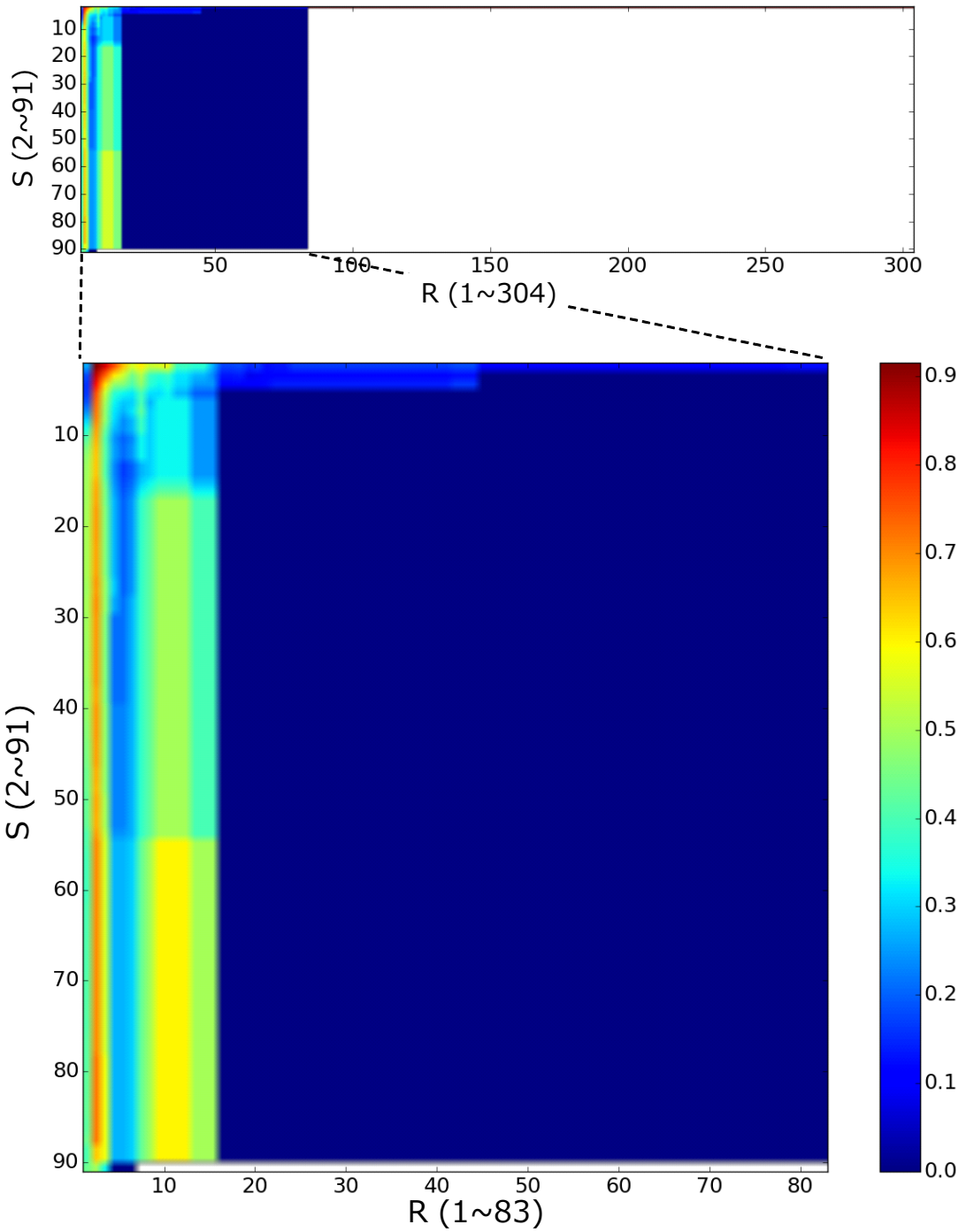

図 3.10: (a) の場合における FP の変化

(2) 見逃し率 (False Negative, FN)

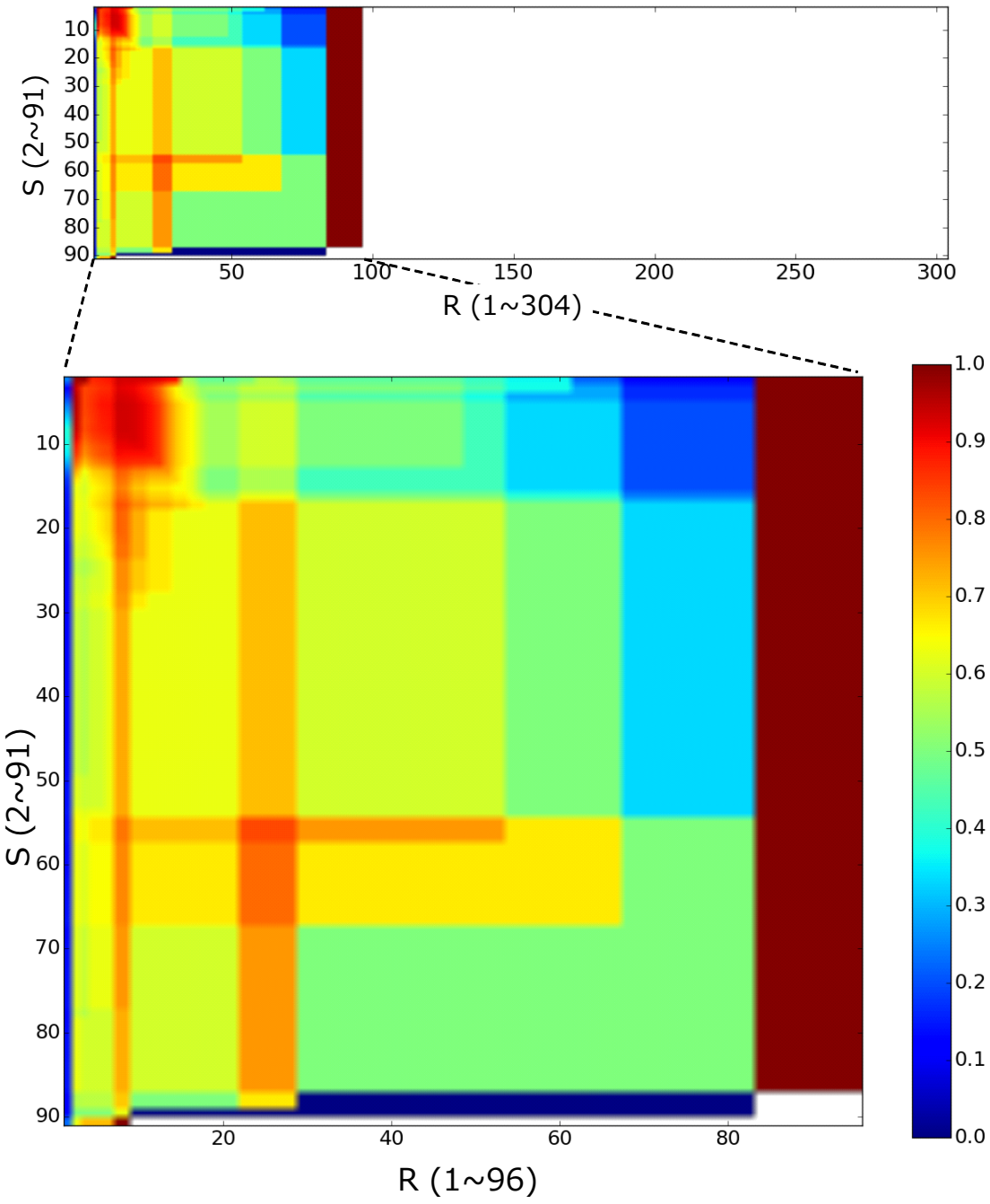

図 3.11: (a) の場合における FN の変化

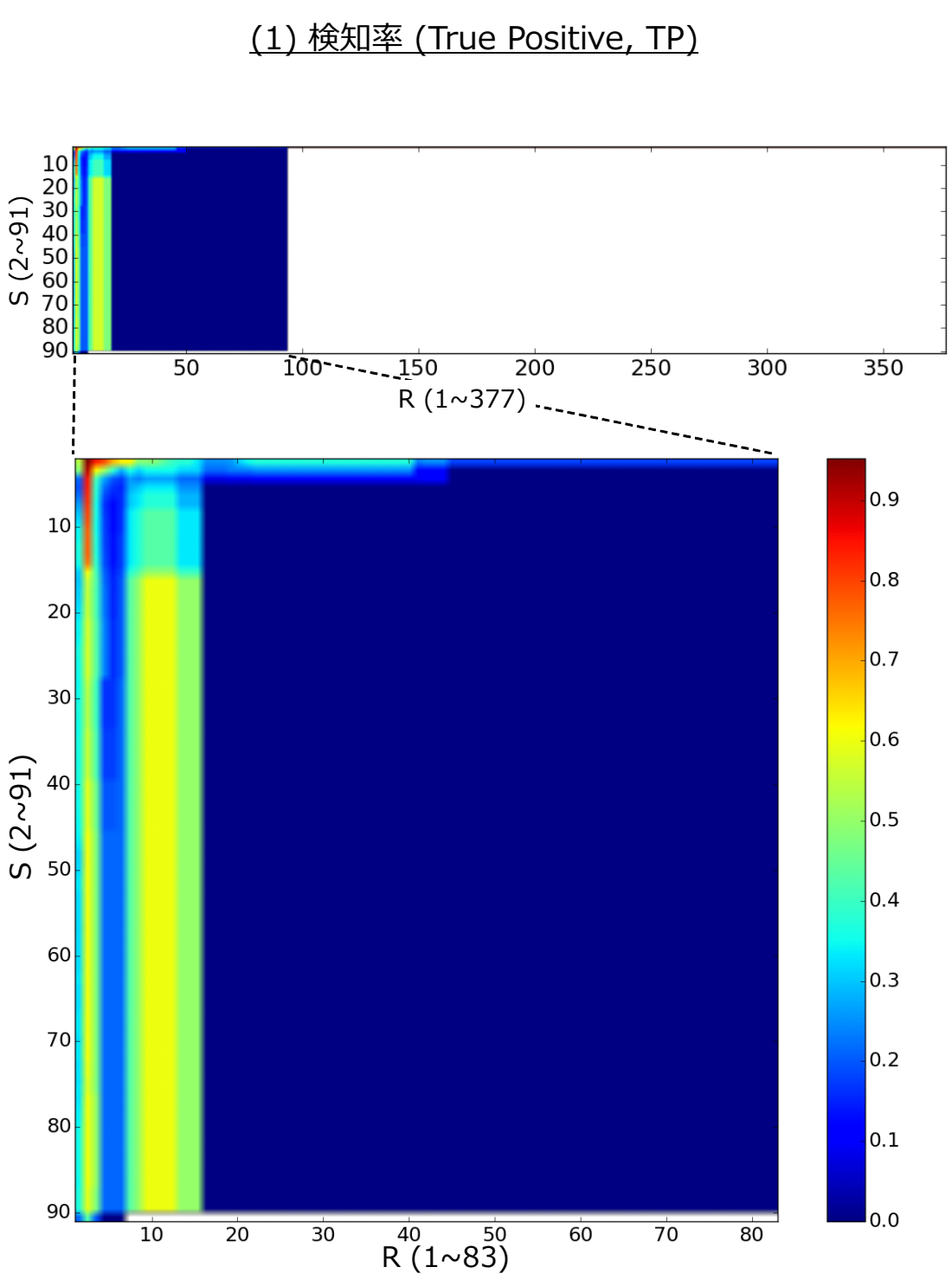

図 3.12: (b) の場合における FP の変化

(2) 見逃し率 (False Negative, FN)

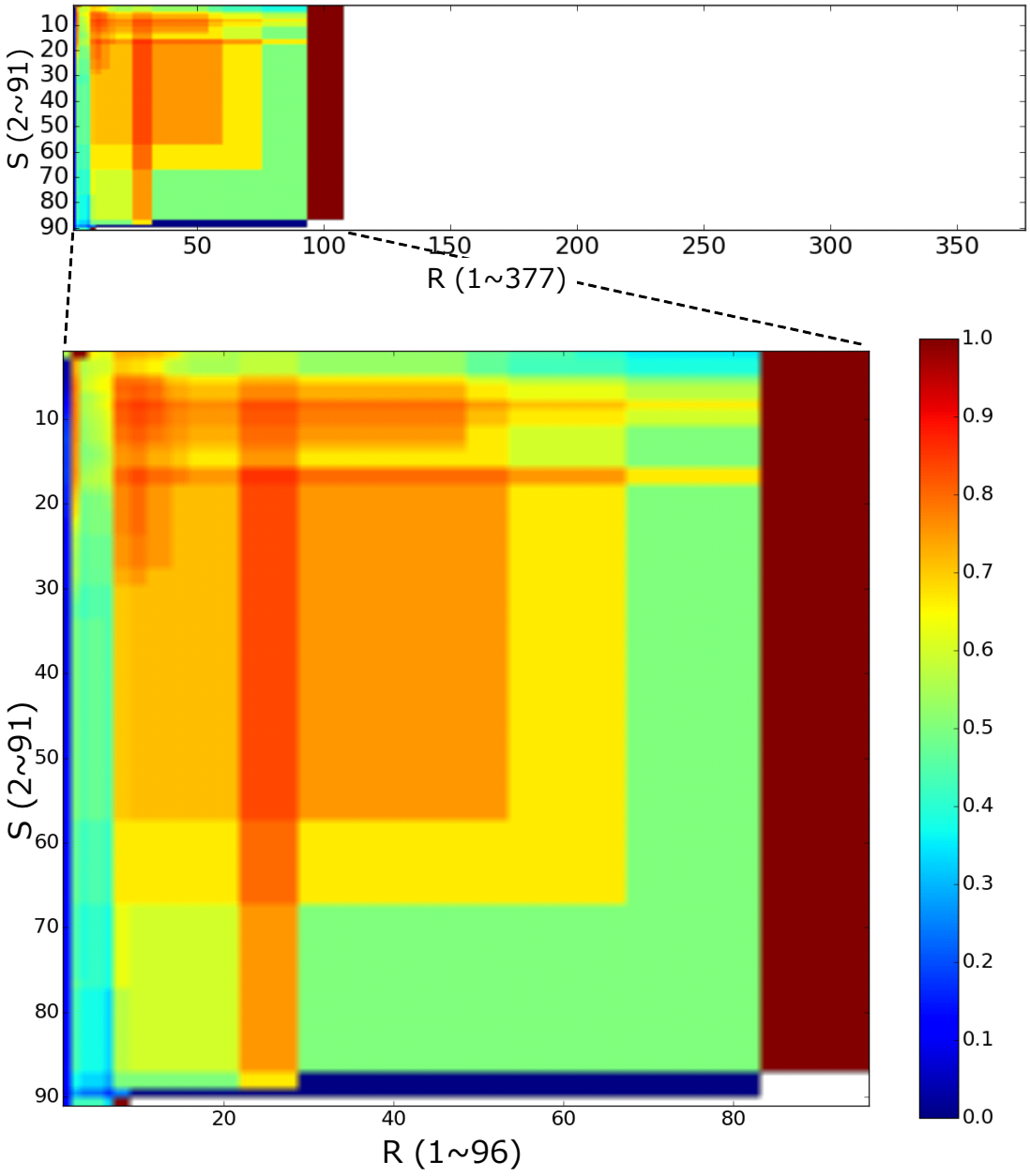

図 3.13: (b) の場合における FN の変化

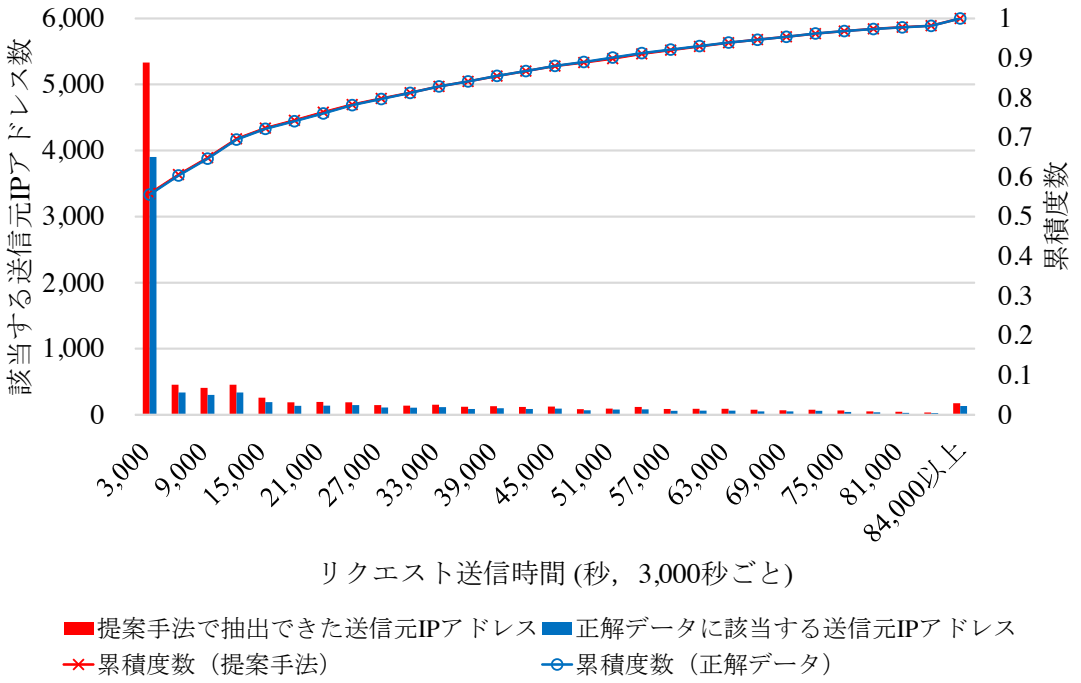

図 3.14: (a) の場合におけるリクエスト送信時間と送信元 IP アドレスの分布

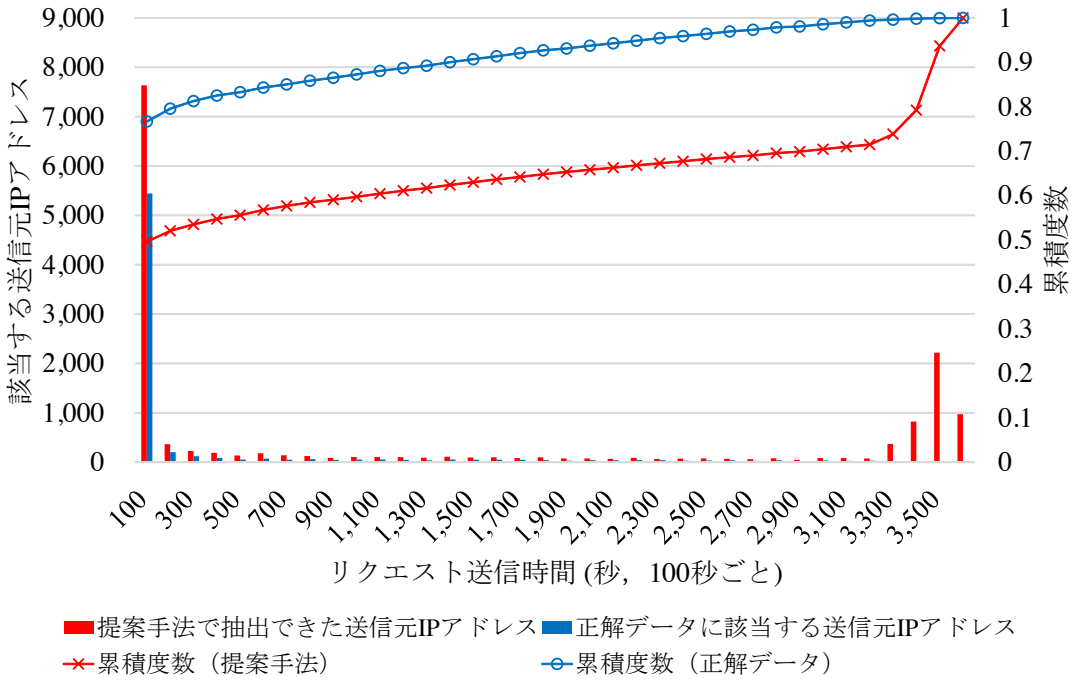

図 3.15: (b) の場合におけるリクエスト送信時間と送信元 IP アドレスの分布

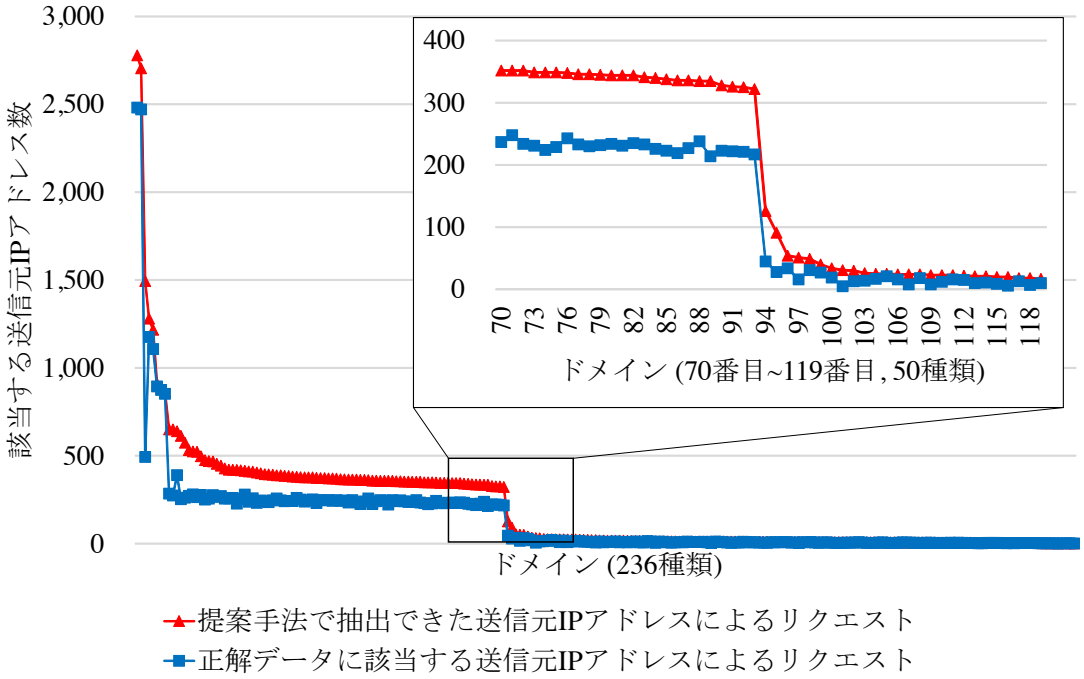

図 3.16: (a) の場合におけるドメインにリクエストを送信した IP アドレス数の分布

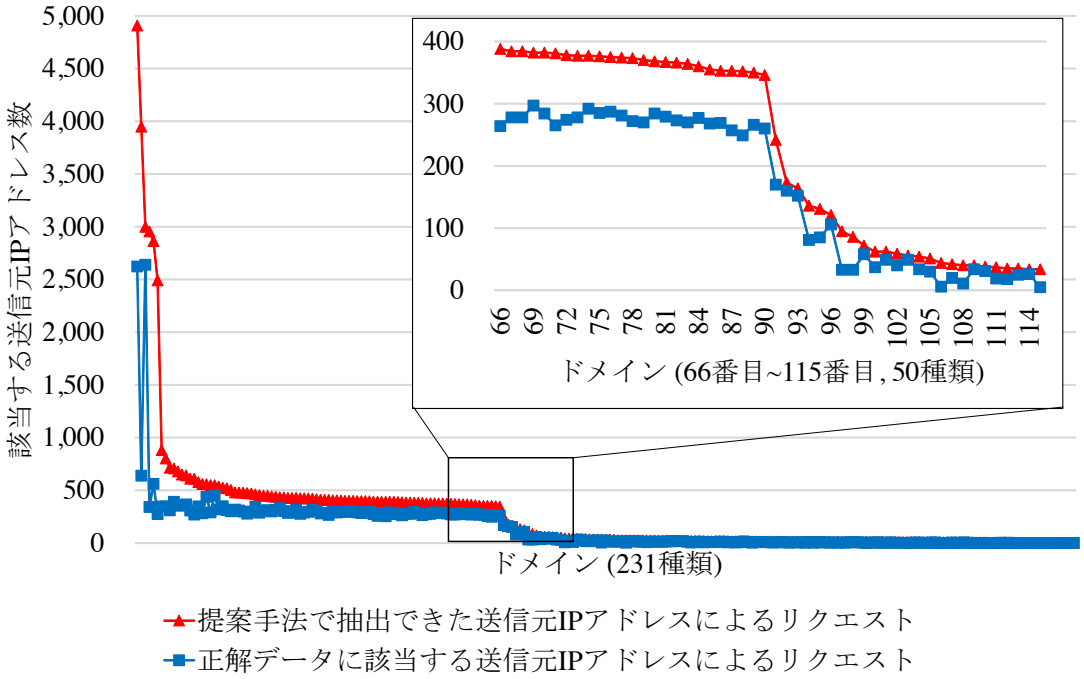

図 3.17: (b) の場合におけるドメインにリクエストを送信した IP アドレス数の分布

第4章　SSH（Secure Shell）サービスに対するブルートフォース検知ログを用いた分散型ブルートフォース攻撃の抽出

4.1　はじめに

ネットワークセキュリティ分野において，ブルートフォース攻撃とは，ネットワークサービスの利用に必要なユーザ名とパスワードの組合せの取得を目的として，存在する全ての組合せについて利用可能かどうかをサービスに対して試す攻撃を指す．例えば，ユーザ名を「root」に固定し，パスワードを「0000」から「9999」まで順番に変えながら，サービスのログインに成功する組合せが存在するかどうかを試す．他にも辞書に収録されている単語をユーザ名あるいはパスワードの候補として探す辞書攻撃や，システムに初期設定される値を使うといった手段も存在する．さらには，他サービスから漏えいしたと考えられるユーザ名・パスワードの組合せを別のサービスのログイン試行に使用する攻撃（パスワードリスト攻撃とも呼ばれている）も報告されている（文献 [88]）．

近年では，攻撃の発生をネットワークサービス管理者に知られないような手段を伴うブルートフォース攻撃が報告されている．例えば，少ないログイン試行回数でのログイン試行を長期間続ける，複数の異なる IP アドレスからログイン試行を行うといった手段が複数のネットワーク監視組織より報告されている．専門組織による報告だけではなく，2013 年 4 月にはコンテンツ管理システム「WordPress」が，同年の 11 月にはソースコード管理システム「GitHub」が，どちらも大規模なブルートフォース攻撃を受けていたことが報告されている．WordPress は 9 万以上のサーバから 100 万回以上のログイン試行を受けていたこと（文献 [89]，文献 [90]）が，GitHub では約 4 万の異なる IP アドレスからログイン試行を受けていたこと（文献 [91]）がそれぞれ報告されている．

上記のようなブルートフォース攻撃は，侵入検知システム（Intrusion Detection System, IDS）だけでは攻撃の発生を検知することも，効果的な対策を適用することも難しい．IDS は通信の挙動等からブルートフォース攻撃である可能性が高い通信を検知することができる．例えば，短時間に単一の IP アドレスからの通信で，大量のログイン失敗を伴っていたものなどである．一方で，IDS は正規通信を「攻撃」と判断することもあり得る（誤検知ともいう）．正規ユーザによる数回のログイン失敗もブルートフォース攻撃と判断されることもある．この誤検知に対し，ある通信が IDS によりブルートフォース攻撃と検知されても，ログイン失敗回数がある値より少ない場合は何もしない，といったように，誤検知の可能性が高い検知結果は無視するようにサービス管理者等により設定されている場合も多い．そのため，前述のような複数 IP アドレス且つ少ログイン試行を伴うブルートフォース攻撃の発生を IDS が検知できたとしても，通信の遮断といった対策が適用されない．

本章では，実際にサービスを運用しているサーバ（拠点ともいう）から取得されたブルートフォース検知ログより，複数の攻撃元 IP アドレスによる，検知時刻，ログイン試行回数について拠点間で互いに強い相関を持つブルートフォース攻撃の発生を，我々の知る限り初めて突き止めたので報

告する．単一の拠点だけでなく，複数の拠点で IDS により検知されたブルートフォース攻撃ログを横断的に分析し，攻撃元として検知された IP アドレスと検知時刻に着目した可視化を適用した．この可視化により，既知のブルートフォース攻撃事象に加えて，複数 IP アドレスが攻撃元となったブルートフォース攻撃が継続して検知されていたことがわかり，本事象の分析の起点とすることができた．本章では，この事象を分散型ブルートフォース攻撃事象 1 とする．

さらに我々は，ブルートフォース検知ログに残る分散型ブルートフォース攻撃事象 1 の挙動を分析することで，本事象が検知された IP アドレスを短期間に取得されたブルートフォース検知ログからでも抽出できる手法を提案する．提案手法ではログイン試行回数の相関に着目し，この攻撃を受け続ける可能性の高い IP アドレスを抽出する．検知した結果を用いることで，抽出した IP アドレスに対する通信を限定して監視することや，ブルートフォース攻撃に関する注意喚起を行うことができる．

提案手法について実際のブルートフォース検知ログログを用いて評価実験を行った．評価の結果，検知時刻とログイン試行回数が抽出精度に大きく影響することがわかった．

本稿の貢献は主に次の 3 つである．

- 拠点横断分析により，分散型ブルートフォース攻撃事象 1 の発生を検知した．
- 分散型ブルートフォース攻撃事象 1 が検知された IP アドレスを短期間に取得されたブルートフォース検知ログからでも抽出できる手法を提案・比較した
- 我々の持つブルートフォース検知ログを手法に適用し，抽出に影響するパラメタの評価を行った

4.2 ブルートフォース検知ログの拠点横断分析

本章では，ネットワークサービスを実際に運用している複数の拠点から得られたブルートフォース検知ログに対して拠点横断分析に基づく可視化を適用した結果を報告する．拠点横断分析では，次の 2 項目を満たすような分析方法を適用することが必要である．まず，複数の拠点間でいつ，どのような攻撃が，どの攻撃元から検知されたのかを把握できることが必要である．拠点ごとに分析するのではなく，各拠点で発生した事象を俯瞰することで，分析対象としているネットワーク全体の攻撃傾向を検知することができる．次に，ブルートフォース検知ログを構成する項目（フィールド）について，着目すべきフィールドを限定することも必要である．ブルートフォース検知ログには複数種類のフィールドが含まれる．顕著に現れる必要最低限のフィールドを分析の対象とすることで，より迅速且つ明確に攻撃事象を検知することができる．

4.2.1 拠点横断分析に基づく可視化

我々は，送信元 IP アドレス，送信先 IP アドレス，攻撃検知時刻に着目した散布図による可視化をブルートフォース検知ログに対して適用した．送信先 IP アドレスを縦軸，検知時刻を横軸として，ブルートフォース攻撃の検知レコードをドットとして描画する．送信元 IP アドレスによりドットの色・形が異なる．この散布図により，送信元 IP アドレス毎に，どの送信先 IP アドレスに対して，いつ攻撃が検知されたかのかを把握することができる．送信元 IP アドレス毎にドットの色・形を変化させることで，同一あるいは異なる送信元 IP アドレスからの攻撃であったのかを区別することができる．

4.2.2 ブルートフォース検知ログ可視化適用結果

前項にて述べた散布図を，2011 年〜2012 年の間に取得された 8ヶ月間のブルートフォース検知ログの，ブルートフォース攻撃を検知したブルートフォース検知ログに適用した．22 番ポートに対する SSH サービスへ向けたブルートフォース攻撃を検知したブルートフォース検知ログ（22 番ポートログ）に適用した結果を図 4.1 に示す．図 4.1 から，大きく 3 種類のブルートフォース攻撃検知事象が確認できた．

4.2.3 確認できたブルートフォース攻撃検知事象

ブルートフォース検知ログ可視化適用結果から，主に 3 種類のブルートフォース攻撃検知事象が確認できた．1 種類目は，ある 1 つの送信元 IP アドレスから複数の送信先 IP アドレス群に対してブルートフォース攻撃が継続して検知されていた事象である．2 種類目も，ある 1 つの送信元 IP アドレスから複数の送信先 IP アドレス群に対してブルートフォース攻撃が継続して検知されていた事象であるが，描画されたドットに規則性があり，検知時刻に規則性を持っていた．これら 2 種類のブルートフォース攻撃検知事象は従来知られているブルートフォース攻撃事象である．

3 種類目は，複数の送信元 IP アドレス群から，複数の送信先 IP アドレス群に対してブルートフォース攻撃が継続して検知されていた事象である．しかも，同じ色・形のドットが縦の列に並んで描画されていた特徴から，ある検知時刻におけるブルートフォース攻撃は 1 つの送信元 IP アドレスから複数の送信先 IP アドレス群に同時に検知されていたことがわかる．同一の送信元 IP アドレスが検知されていない特徴から，この事象を分散型ブルートフォース攻撃事象 1 と呼ぶ（図 4.2）．

4.3 分散型ブルートフォース攻撃事象 1 の分析

分散型ブルートフォース攻撃事象 1 を検知した送信先 IP アドレスの集合を送信先 IP アドレス，送信先 IP アドレスに対するブルートフォース攻撃が検知された送信元 IP アドレスの集合を送信元 IP アドレスとする．送信元 IP アドレスに含まれるある $srcIP_1$ が一定期間送信先 IP アドレスに対してブルートフォース攻撃を検知され続けたとき，検知が開始された時刻を t_{start}，検知された最後の時刻を t_{fin} とする．送信元 IP アドレスに含まれる送信元 IP アドレスである $srcIP_1$ は，時刻 t_{start} から t_{fin} の間 IDS によりブルートフォース攻撃が検知されていたと表現できる．

22 番ポートログから分散型ブルートフォース攻撃事象 1 に該当するレコードを抽出し，送信元 IP アドレス，ログイン試行回数，検知時刻に関する統計を計測した．

まず送信元 IP アドレスに関して，送信元 IP アドレスに含まれる 1 つの送信元 IP アドレスは，一定期間ブルートフォース攻撃を検知された後は，再度同じ送信元 IP アドレスが検知されることがほとんどなかった．ブルートフォース攻撃攻撃が検知された約 1〜2ヶ月に同じ送信元 IP アドレスが登場したケースはあったものの，そのような送信元 IP アドレスはごくわずかであった．

送信元 IP アドレスがブルートフォース検知ログに登場した日数の集計結果を表 4.1 に示す．送信元 IP アドレス数の括弧内の数字は日付をまたいでいた送信元 IP アドレス数を示す．約 80%の送信元 IP アドレスが約 8ヶ月間の中で 1 日しか登場していなかった．5 日，8 日登場した送信元 IP アドレスについて日付に規則性はなかった．

次にログイン試行回数について，1 つの送信元 IP アドレスにおける t_{start} から t_{fin} の間に検知されたブルートフォース攻撃における検知時刻におけるログイン試行回数の最小値・最大値・平均を図 4.3 に示す．検知時刻の最小単位は 1 分である．横軸は 1 つの送信元 IP アドレスによるブルー

表 4.1: ログに登場した日数

登場日数	送信元 IP アドレス数
1	304(0)
2	53(14)
3	15(1)
4	4(3)
5	1(0)
8	1(0)

トフォース攻撃を，縦軸はログイン試行回数を示す．線分の上側の終端がログイン試行回数の最大値，下側の終端が最小値，三角が平均をそれぞれ示す．分析対象のブルートフォース検知ログ全体のログイン試行回数の平均が約 72.18 回であったことと比較するとログイン試行回数が少ないことがわかる．

検知時刻に関して，ある 1 つの送信元 IP アドレスから送信先 IP アドレスに対してほぼ同じ時刻でブルートフォース攻撃が検知されていた．送信先 IP アドレスによっては，1 分間のみブルートフォース攻撃が検知されたものもあれば，数時間検知され続けたものもあった．送信元 IP アドレスに含まれる送信元 IP アドレスによるブルートフォース攻撃が検知された時間の分布を図 4.4 に示す．データ区間を 30 分とし，0〜1440 分の範囲について作成した縦軸は該当する送信元 IP アドレスの数を示す．ヒストグラムの山は左側に偏っており，攻撃の約 9 割は 3 時間以内に終了していた．

以上から，分散型ブルートフォース攻撃事象 1 の挙動に関して次の 3 点が推測できる．

- ブルートフォース攻撃の発生をサービス管理者から隠ぺいするため，少ないログイン試行回数でブルートフォース攻撃を行っていた．
- 送信先 IP アドレスへブルートフォース攻撃を試みる攻撃者が存在し，その攻撃者によって送信元 IP アドレスが用意された．
- 攻撃者は攻撃が発生していることをカモフラージュするため，あるいは自身の身元を隠すために，送信元 IP アドレスに含まれる送信元 IP アドレスを使い捨てながらブルートフォース攻撃を行い続けていた．

4.4 分散型ブルートフォース攻撃事象 1 の送信先となった送信先 IP アドレスの抽出

分散型ブルートフォース攻撃事象 1 に該当するブルートフォース攻撃を検知された送信先 IP アドレス群（分散型ブルートフォース攻撃事象 1 被検知送信先 IP アドレス）をブルートフォース検知ログから抽出することで，早期に分散型ブルートフォース攻撃事象 1 被検知送信先 IP アドレスを検知する手法を検討する．分散型ブルートフォース攻撃事象 1 被検知送信先 IP アドレスは，長期間分散型ブルートフォース攻撃事象 1 によるブルートフォース攻撃を検知され続けていた．早期に送信先 IP アドレスを抽出することで，抽出した送信先 IP アドレスに対する通信を限定して監視することや，ブルートフォース攻撃に関する注意喚起を行うことができる．

4.4.1 利用できる挙動と検討対象とする抽出手法

送信元 IP アドレスに含まれる 1 つの送信元 IP アドレスによるブルートフォース攻撃の挙動を利用してブルートフォース検知ログを分析することで，送信先 IP アドレスに該当する送信先 IP アドレス群を抽出することができる．分析に利用できる挙動として，次の 3 つが挙げられる．これらの挙動がブルートフォース検知ログ中に存在するかを分析することで，該当するレコードから送信先 IP アドレス群を抽出する．

a) ある 1 つの送信元 IP アドレスから複数の送信先 IP アドレス群 (送信先 IP アドレス) へブルートフォース攻撃が検知される

b) 同じ送信元 IP アドレスからは同一の時刻にブルートフォース攻撃が検知される

c) 同じ送信元 IP アドレス，同じ時刻に検知されたブルートフォース攻撃におけるログイン試行回数は同一である

抽出手法として，本稿では次の 3 種類の手法を検討対象とする．

抽出手法 1: 送信元 IP アドレス，送信先 IP アドレス，検知時刻，ログイン試行回数の相関を計算

挙動 a)，b)，c) に基づき，srcIP，送信先 IP アドレス，ある時刻に検知された単位時間当たりのログイン試行回数について互いに相関が高い送信先 IP アドレス群を抽出する．入力は送信元 IP アドレス，送信先 IP アドレス，検知時刻，ログイン試行回数をフィールドとして持つブルートフォース検知ログである．入力されたブルートフォース検知ログから，送信元 IP アドレス，検知時刻におけるログイン試行回数について互いに相関が高い送信先 IP アドレス群を抽出する．

抽出された送信先 IP アドレス群はいずれも，同じ時刻に，同じ送信元 IP アドレスから，同じログイン試行回数によるブルートフォース攻撃を検知された特徴を持つ．しかし IDS の検知アルゴリズム等により，異なる時刻やログイン試行回数で検知された場合，この抽出手法 1 では送信先 IP アドレス群を正確に抽出できない可能性もある．

抽出手法 2: 送信元 IP アドレス，送信先 IP アドレスの相関を計算

送信元 IP アドレス，送信先 IP アドレスについて互いに相関が高い送信先 IP アドレス群を抽出する．入力は送信元 IP アドレス，送信先 IP アドレスをフィールドとして持つブルートフォース検知ログである．入力されたブルートフォース検知ログから，送信先 IP アドレス毎に，各送信元 IP アドレスからブルートフォース攻撃を検知された回数を数え上げる．各送信元 IP アドレスからのブルートフォース攻撃を検知された回数について互いに相関が高い送信先 IP アドレス群を抽出する．

抽出された送信先 IP アドレス群はいずれも，同じ送信元 IP アドレスからブルートフォース攻撃を検知された特徴を持つ．しかし抽出手法 2 では，抽出された送信先 IP アドレス群が検知されたログイン試行回数の相関を考慮することができない．

抽出手法 3: 送信先 IP アドレス，検知時刻，ログイン試行回数の相関を計算

送信先 IP アドレス，ある時刻に検知された単位時間当たりのログイン試行回数について互いに相関が高い送信先 IP アドレス群を抽出する．入力は送信先 IP アドレス，検知時刻，ログイン試行

回数をフィールドとして持つブルートフォース検知ログである．入力されたブルートフォース検知ログから，検知時刻におけるログイン試行回数について互いに相関が高い送信先 IP アドレス群を抽出する．

抽出された送信先 IP アドレス群はいずれも，同じ検知時刻に同じログイン試行回数によるブルートフォース攻撃を検知された性質を持つ．しかし抽出手法 3 では，抽出された送信先 IP アドレス群が検知された送信元 IP アドレスが同一であったか否かを考慮することができない．

4.4.2 処理の流れ

抽出手法の処理の流れを示す．まず，入力されたブルートフォース検知ログから，送信先 IP アドレスを行とする 2 次元データ列を作成する．2 次元データ列の行はそれぞれ，抽出手法 1 では送信元 IP アドレスと検知時刻，抽出手法 2 では送信元 IP アドレス，抽出手法 3 では検知時刻である．データ列中の数値はそれぞれ，抽出手法 1 では送信元 IP アドレスと検知時刻におけるログイン試行回数，抽出手法 2 では送信元 IP アドレスにおけるブルートフォース検知ログ中のレコード件数，抽出手法 3 では検知時刻におけるログイン試行回数である．

次に，作成された 2 次元データ列中の各行をデータ列とみなし，相関の高低を計算することで，互いに相関が高い送信先 IP アドレス群を計算する．互いに相関が高い送信先 IP アドレス群を送信先 IP アドレスとして出力する．

各抽出手法の処理において作成される 2 次元データ列の例を図 4.5 に示す．

4.5 分散型ブルートフォース攻撃事象 1 の送信先となった送信先 IP アドレス抽出手法の比較

前節にて述べた分散型ブルートフォース攻撃事象 1 の被検知送信先 IP アドレス抽出手法を比較するため，我々の所持するブルートフォース検知ログに対して適用し，分散型ブルートフォース攻撃事象 1 被検知送信先 IP アドレスの抽出精度を比較した．抽出精度を次の 2 点から評価した．

- **抽出正確性** : 分散型ブルートフォース攻撃事象 1 被検知送信先 IP アドレスを誤検知あるいは見逃しなく正確に抽出できること．
- **抽出安定性** : 手法の入力となるブルートフォース検知ログに依存せずに分散型ブルートフォース攻撃事象 1 被検知送信先 IP アドレスを抽出できること．

相関係数のしきい値を変化させ抽出手法 1，抽出手法 2，抽出手法 3 を適用し，分散型ブルートフォース攻撃事象 1 被検知送信先 IP アドレスの抽出精度を比較した．

4.5.1 比較手順

実際の運用時には 1 日に 3 回抽出手法を適用することで分散型ブルートフォース攻撃事象 1 被検知送信先 IP アドレスの有無を調査することを想定し，ある 8 時間の間に検知されたレコード（レコード集合）を各抽出手法の入力であるブルートフォース検知ログとした．2011〜2012 年の 8ヶ月間に取得された，分散型ブルートフォース攻撃事象 1 が検知された 22 番ポートに対する SSH へ向けたブルートフォース攻撃検知ログに含まれるレコードから，無作為抽出により 100 のレコード集

合を作成し，各抽出手法に適用した．作成されたレコード集合に含まれるレコード件数は，最小で 2 件，最大で 3,371 件，平均が 436.37 件であった．

前節に述べた処理の流れに沿って，相関係数のしきい値を 0.1 から 0.9 まで 0.1 ずつ増加させながら 100 のレコード集合に対して 3 種類の抽出手法を適用し，送信先 IP アドレスを抽出した．なお，レコード集合における検知時刻は 2 分単位とした．なお，相関の計算には，ピアソンの相関係数を用いて，データ列 $\{x_i\}$，$\{y_i\}$ の相関係数を次式により計算した．$\bar{x}$，$\bar{y}$ はそれぞれデータ列 $\{x_i\}$，$\{y_i\}$ の平均を示す．

$$\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i-\bar{y})^2}}$$

また，互いに相関が高い送信先 IP アドレスは最大クリークを抽出する手法（文献 [94]）を用いて抽出した．送信先 IP アドレスをノード，相関の高い送信先 IP アドレスのペアをノード同士がエッジで結ばれているとみなし，最大クリークに該当する送信先 IP アドレス群を抽出した．

第 4.3 節で統計の計測のために抽出されたレコードに基づき分散型ブルートフォース攻撃事象 1 被検知送信先 IP アドレスを作成し，各抽出手法により抽出できた送信先 IP アドレスと比較した．送信先 IP アドレスの誤検知の割合（False Positive, FP），見逃しの割合（False Negative, FN）を計算した．FP は分散型ブルートフォース攻撃事象 1 被検知送信先 IP アドレスに該当しないものの，手法により分散型ブルートフォース攻撃事象 1 被検知送信先 IP アドレスと判断された攻撃の割合を示す．FN は分散型ブルートフォース攻撃事象 1 被検知送信先 IP アドレスに該当するものの，手法により分散型ブルートフォース攻撃事象 1 被検知送信先 IP アドレスと判断されなかった攻撃の割合を示す．

4.5.2 比較結果

しきい値毎の FP，FN の平均を折れ線グラフとしてそれぞれ図 4.6，図 4.7 に示す．FP においては，抽出手法 1 はしきい値を 0.3 あるいは 0.4 に，抽出手法 2 はしきい値を 0.6 に，抽出手法 3 はしきい値を 0.4 にそれぞれ設定したとき，小さい FP で分散型ブルートフォース攻撃事象 1 被検知送信先 IP アドレスを抽出することができた．FN においては，どの抽出手法においてもしきい値を 0.1 に設定したとき，小さい FN で分散型ブルートフォース攻撃事象 1 被検知送信先 IP アドレスを抽出することができた．グラフの変化から，抽出手法 1・3 と比較して，抽出手法 2 では相関係数のしきい値が大きくなっても FP，FN に大きな差が確認できず，どのしきい値を設定した場合においても低い FP・FN で分散型ブルートフォース攻撃事象 1 被検知送信先 IP アドレスを抽出することができた．一方で，検知時刻におけるログイン試行回数を考慮した抽出手法 1・抽出手法 3 では低い FP・FN で分散型ブルートフォース攻撃事象 1 被検知送信先 IP アドレスを抽出することが難しかった．また，抽出手法 1 と抽出手法 3 の場合における FP，FN の変化が類似しており，検知時刻におけるログイン試行回数を考慮することで FP・FN に大きな差が出ることがわかる．

次に，しきい値毎の FP，FN の標準偏差を折れ線グラフとしてそれぞれ図 4.8，図 4.9 に示す．標準偏差が大きいほど，各レコード集合における FP，FN のばらつきが大きいことを示す．グラフの変化から，抽出手法 1・3 と比較すると，抽出手法 2 ではどのしきい値を設定した場合においても，どのレコード集合に対しても変わらない FP・FN で分散型ブルートフォース攻撃事象 1 被検知送信先 IP アドレスを抽出することができた．一方で，抽出手法 1・3 を適用した場合は，設定するしきい値の値によっては FP の標準偏差の値が大きい場合もあり，FP の値のばらつきが大きかったことがわかる．

4.5.3 考察

以上の結果から，抽出正確性，抽出安定性の 2 点から 3 種類の抽出手法を評価する．相関係数のしきい値を変化させた場合において，いかなるしきい値を設定した場合でも，小さい FP 平均，FN 平均で分散型ブルートフォース攻撃事象 1 の送信先となった送信先 IP アドレスを検知できた抽出手法を「抽出正確性を満たす」とし，いかなるしきい値を設定した場合でも，小さい FP 標準偏差，FN 標準偏差で分散型ブルートフォース攻撃事象 1 の送信先となった送信先 IP アドレスを検知できた抽出手法を「抽出安定性を満たす」とする．図 4.6〜図 4.9 より，抽出手法 2 が相関係数のしきい値を変化させた場合において抽出正確性と抽出安定性の両方を満たす．

検知時刻，ログイン試行回数の抽出手法への影響を考察する．3 種類の抽出手法を適用した結果，抽出手法 1 と抽出手法 3 は図 4.6〜図 4.9 のいずれのグラフにおいても，グラフの変化が類似していた．これらのことから，検知時刻とログイン試行回数が FP・FN の低下に大きく影響していた．理由として，分散型ブルートフォース攻撃事象 1 に該当するブルートフォース攻撃であっても，検知されたログイン試行回数や検知時刻が，IDS の検知アルゴリズムや送信先 IP アドレスが割り当てられているサーバの設定によって送信先 IP アドレス毎に異なる検知時刻・ログイン試行回数が検知された可能性が考えられる．そのため，今回の実験では，検知時刻とログイン試行回数のどちらも考慮しなかった抽出手法 2 が 3 種類の中で抽出正確性，抽出安定性を最も満たす手法であると判断された．もしも，分散型ブルートフォース攻撃事象 1 被検知送信先 IP アドレスをブルートフォース検知ログより抽出するには，検知時刻とログイン試行回数について値の切り下げ・切り上げ等の処理を適用する必要がある．

4.6 まとめ

本章では，サービスを実際に運用している複数の拠点から得られたブルートフォース検知ログに対して，拠点横断分析に基づく可視化を適用した結果，既知のブルートフォース攻撃事象に加えて，複数 IP アドレスが攻撃元となり少ログイン試行回数を伴うブルートフォース攻撃事象（分散型ブルートフォース攻撃事象 1）を検知することができた．分散型ブルートフォース攻撃事象 1 が持つ挙動を利用してブルートフォース検知ログを分析することで，分散型ブルートフォース攻撃事象 1 が検知された送信先 IP アドレスをブルートフォース検知ログから抽出することができる．評価実験の結果，検知時刻とログイン試行回数が分散型ブルートフォース攻撃事象 1 を受けた送信先 IP アドレスの抽出精度に大きく影響することがわかった．

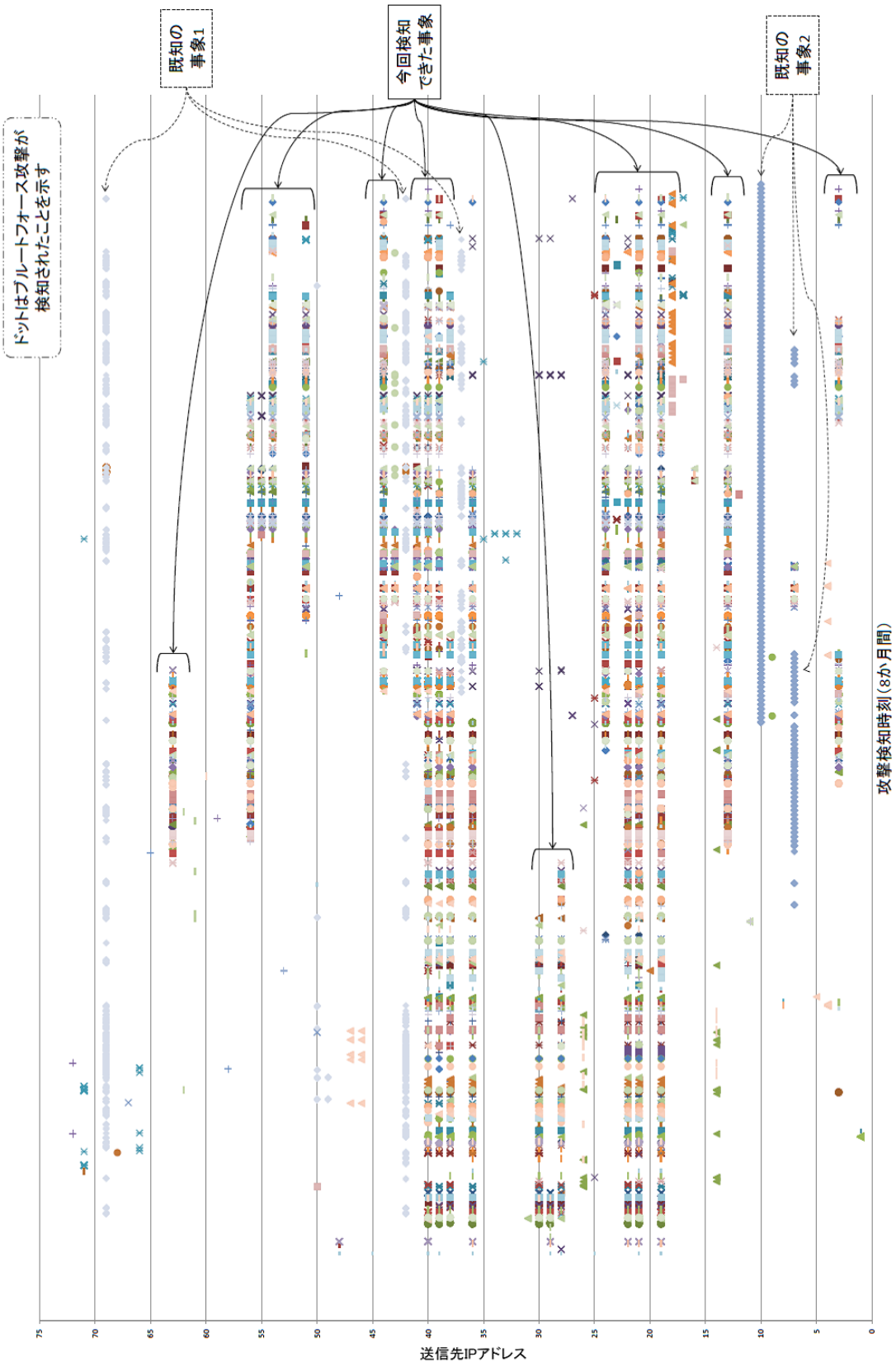

図 4.1: 22 番ポートに対するブルートフォース検知ログの可視化結果

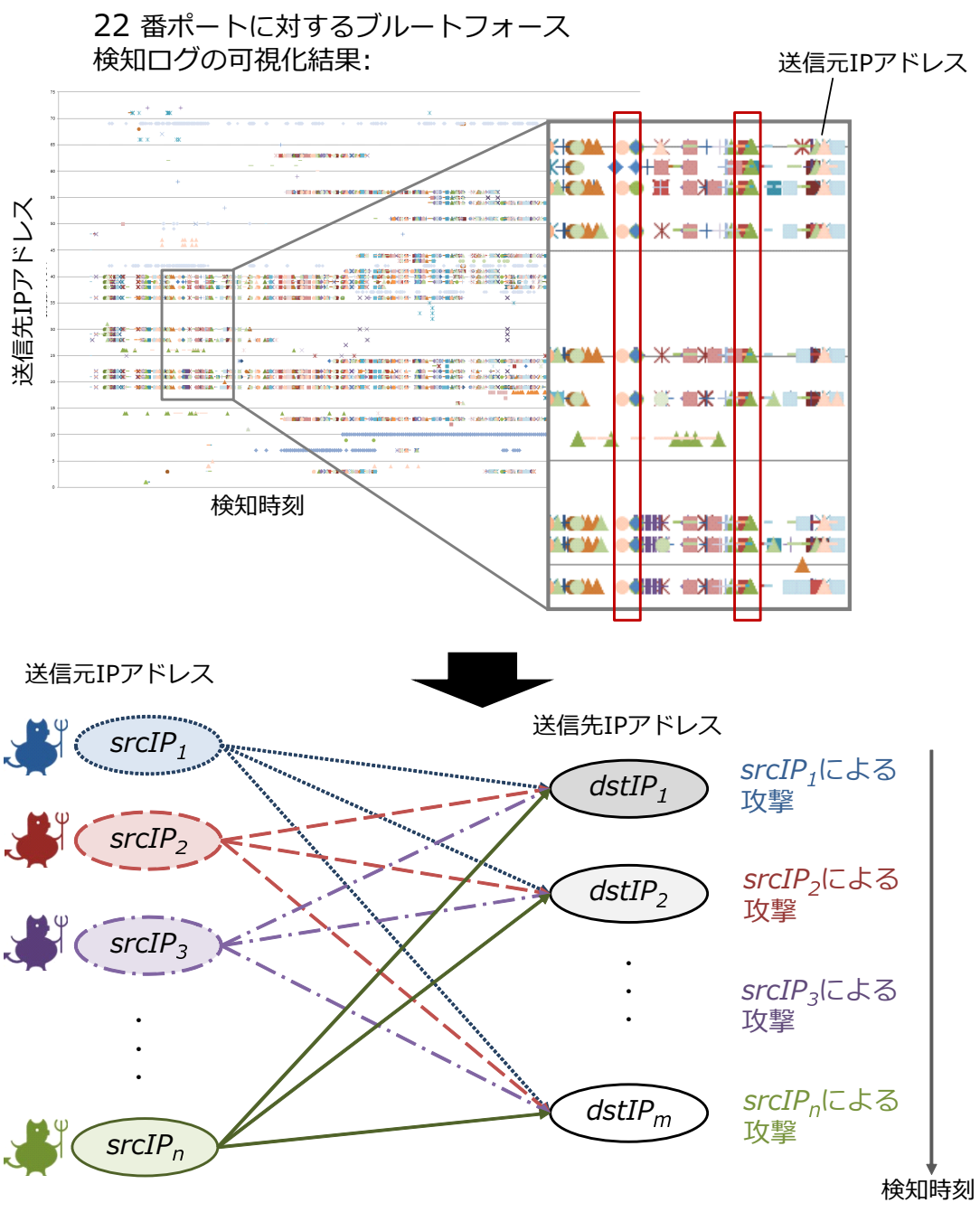

図 4.2: 散布図表現により抽出できた分散型ブルートフォース攻撃事象 1 の挙動

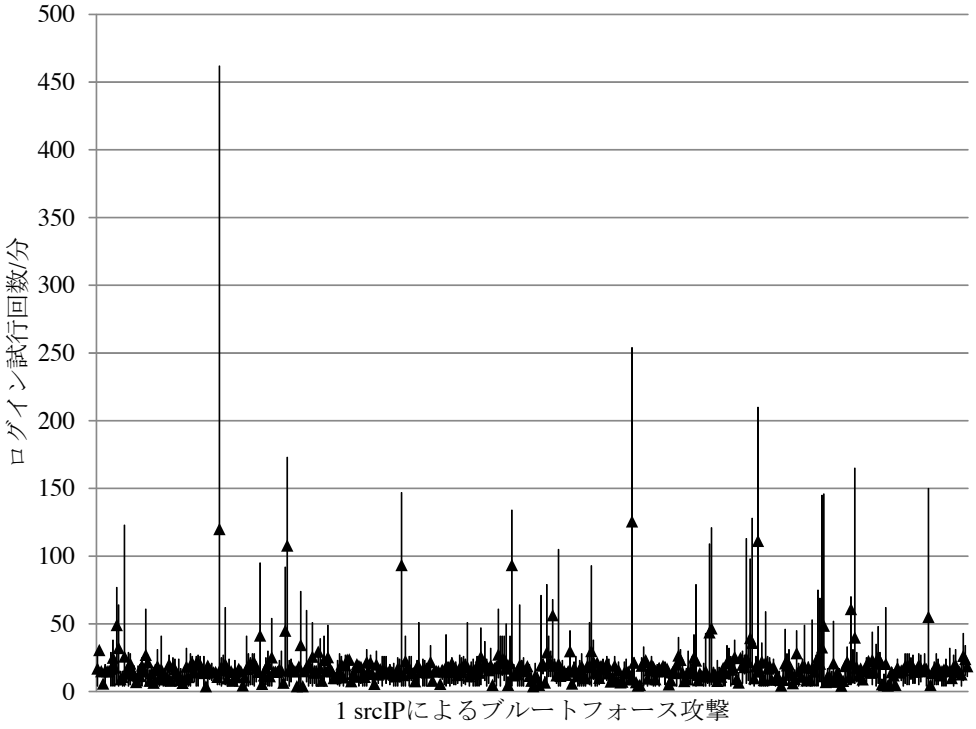

図 4.3: 送信元 IP アドレス毎のログイン試行回数

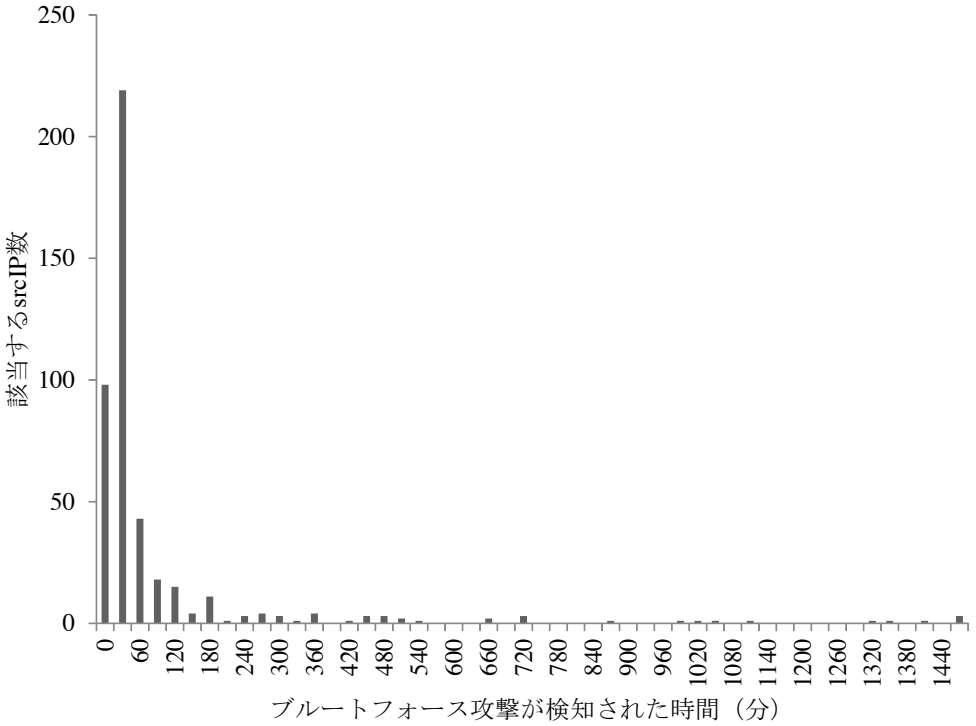

図 4.4: 30 分毎，0～1440 分における攻撃検知時間の分布

抽出手法1

ログイン試行回数		$dstIP_1$	$dstIP_2$	$dstIP_3$	$dstIP_4$	…
$srcIP_1$	1/1 0:00	10	10	10	0	…
	1/1 0:01	11	10	9	0	…
	1/1 0:02	13	12	12	0	…
$srcIP_2$	1/1 0:00	0	0	5	5	…
	1/2 2:01	0	0	0	3	…
:	:	:	:	:	:	…

抽出手法2

レコード件数	$dstIP_1$	$dstIP_2$	$dstIP_3$	$dstIP_4$	…
$srcIP_1$	3	3	3	0	…
$srcIP_2$	0	0	1	2	…
$srcIP_3$	10	10	10	0	…
$srcIP_4$	5	5	5	0	…
$srcIP_5$	0	0	0	20	…
:	:	:	:	:	…

抽出手法3

ログイン試行回数	$dstIP_1$	$dstIP_2$	$dstIP_3$	$dstIP_4$	…
1/1 0:00	10	10	15	5	…
1/1 0:01	11	10	9	0	…
1/1 0:02	13	12	12	0	…
1/1 0:03	0	0	20	0	…
1/1 0:04	0	0	0	3	…
:	:	:	:	:	…

図 4.5: 抽出手法の処理により作成された 2 次元データ列の例

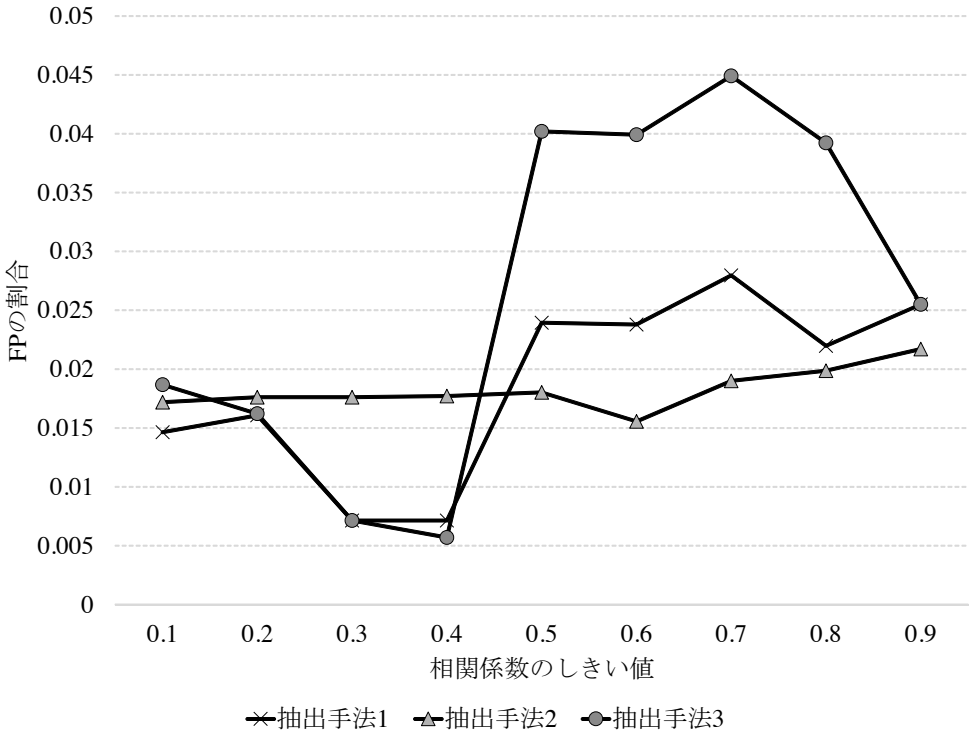

図 4.6: FP 平均変化

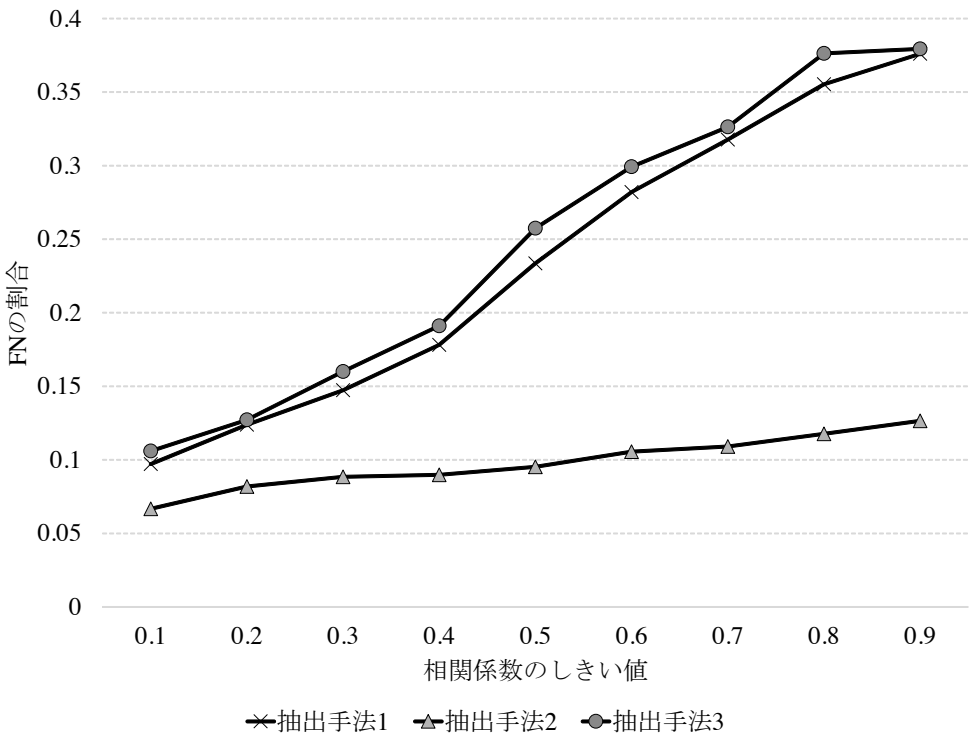

図 4.7: FN 平均変化

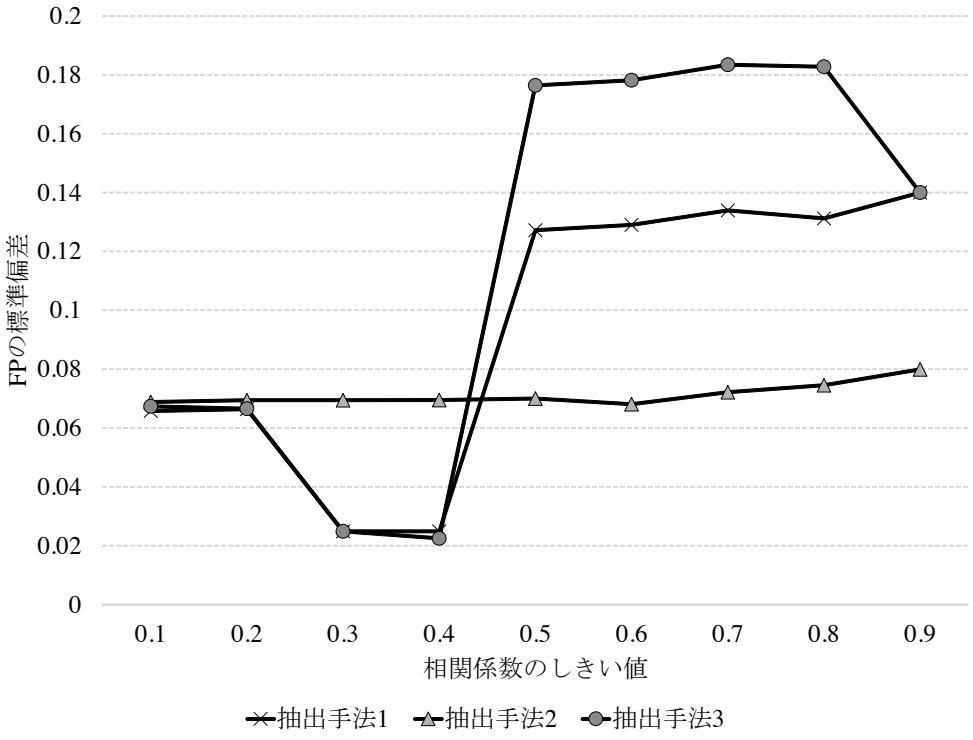

図 4.8: FP 標準偏差変化

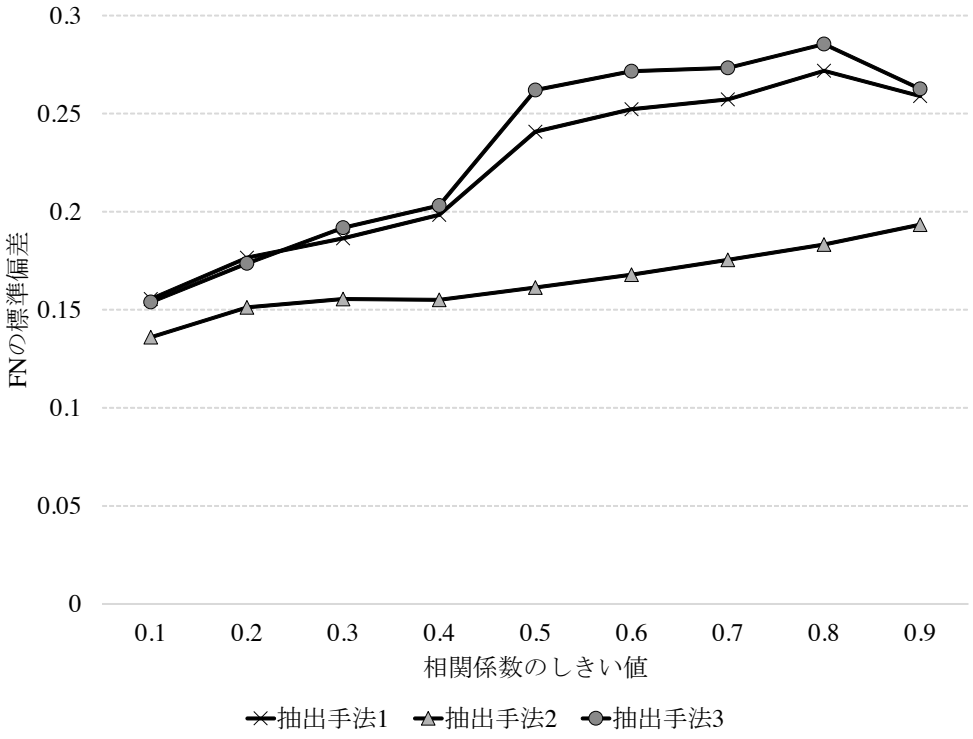

図 4.9: FP 標準偏差変化

第5章　RDP（Windows Remote Desktop）サービスに対するブルートフォース検知ログを用いた分散型ブルートフォース攻撃の抽出

5.1　はじめに

企業が管理するIPアドレス帯で実際にRDPサービスを運用している複数のサーバに対して，IDSがブルートフォース攻撃と判断したアラートログを対象とする．このブルートフォース検知ログから，送信元IPアドレスによるブルートフォース攻撃アラートのインターバルやアラートが検知した攻撃回数の分散を計測し，送信元IPアドレスを変えながら，共通の規則性を以ってブルートフォース攻撃を繰り返す事象（分散型ブルートフォース攻撃事象2）を抽出する．さらに，この規則性を有する送信元IPアドレスを特定することで，分散型ブルートフォース攻撃事象2でブルートフォース攻撃の発生を遮断する手法を提案し，提案手法が題材とするブルートフォース検知ログに対して提案手法が当該事象による被害を抑えられることを示す．

5.2　本章で対象とするブルートフォース攻撃事象（分散型ブルートフォース攻撃事象2）

本章では，企業が管理するIPアドレス帯で実際にRDP（Remote Desktop Protocol）サービスを運用しているサーバ群に対し，IDS（侵入検知装置，Intrusion Detecting System）がブルートフォース攻撃と判断したアラートログ（ブルートフォース検知ログ）から抽出できたブルートフォース攻撃事象について述べる．

5.2.1　分散型ブルートフォース攻撃事象2の構造

本章で対象とするブルートフォース攻撃事象（分散型ブルートフォース攻撃事象2）の構造を述べる．この攻撃事象では，複数の送信元IPアドレスが対象となった複数の送信先IPアドレスに，一定期間，同じ回数だけログイン試行を実施する．送信先IPアドレスに着目すると，この事象では図5.1に示すように，複数の送信元IPアドレスがある特定の1送信先IPアドレスを攻撃先と定め，ブルートフォース攻撃を実施する様子がブルートフォース検知ログより読み取れる．1送信元IPアドレスからは，1送信先IPアドレスに対してブルートフォース攻撃が継続して発生している．以降では，ある送信元IPアドレスからある送信先IPアドレスに対して継続的に発生したブルートフォース攻撃を，攻撃シーケンスとする．図5.1では，送信元IPアドレス群（送信元IPアドレス $srcIP_1$，送信元IPアドレス $srcIP_2$，送信元IPアドレス $srcIP_3$，...，送信元IPアドレス $srcIP_n$）

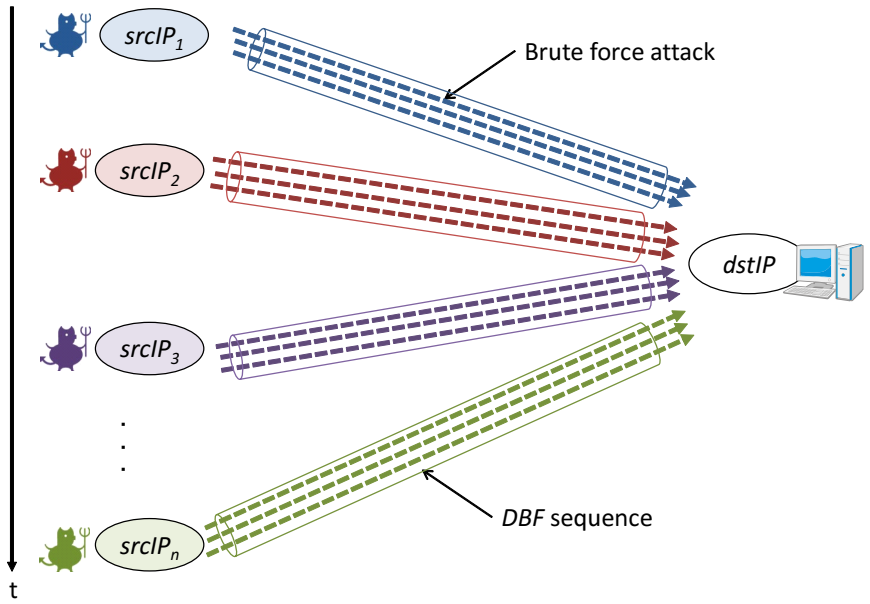

図 5.1: 第 5 章で対象とするブルートフォース攻撃事象の構造

が，送信先 IP アドレス $dstIP_1$ を攻撃対象としている．この図には，$srcIP_1$ から $dstIP_1$ に対する攻撃シーケンス，$srcIP_2$ から $dstIP_1$ に対する攻撃シーケンス，$srcIP_3$ から $dstIP_1$ に対する攻撃シーケンス，...，$srcIP_n$ から $dstIP_1$ に対する攻撃シーケンスといったように，計 n 個の攻撃シーケンスが含まれている．これらの攻撃シーケンスは，$srcIP_1$ による攻撃シーケンスが終了したのちに $srcIP_2$ による攻撃シーケンスが発生する，といったように，1 送信先 IP アドレスに対して同時に発生するのではなく，1 攻撃シーケンスずつ発生する．

第 4 章で述べた分散型ブルートフォース攻撃事象との相違点を述べる．第 4 章で述べた事象は，送信元 IP アドレスが 1 度に複数の送信先 IP アドレス群に対してブルートフォース攻撃を実施していたことを攻撃事象の主な特徴としていた．しかも，送信先となった IP アドレス群にはログイン試行回数に相関がみられた．一方で，分散型ブルートフォース攻撃事象 2 では，1 送信元 IP アドレスが送信先とする IP アドレスは 1 つである．そのため，送信先 IP アドレス間のログイン試行回数の相関を計測することができない．そこで分散型ブルートフォース攻撃事象 2 は，1 送信元 IP アドレスと 1 送信先 IP アドレスの間に存在するブルートフォース攻撃の規則性を主な特徴とする．この事象では，送信元 IP アドレスはある送信先 IP アドレスに対して，同じログイン試行回数を以って一定期間，ブルートフォース攻撃を実行し続けていた．

この攻撃事象は，多数の IP アドレスリソースを操作可能な攻撃者により引き起こされたものであると推測できる．しかし，個々の攻撃シーケンスについては下記に述べるように，IP アドレスリソースを操作する攻撃者が原因とはならない場合が存在する可能性がある．例えば，正規ユーザが間違えたパスワードを送信し続けた結果，IDS によりブルートフォース攻撃と判断される場合が考えられる．また，悪意がある場合であっても，サーバに侵入したい意図を持った個々のマルウェアや人間によって引き起こされたブルートフォース攻撃も存在する可能性がある．

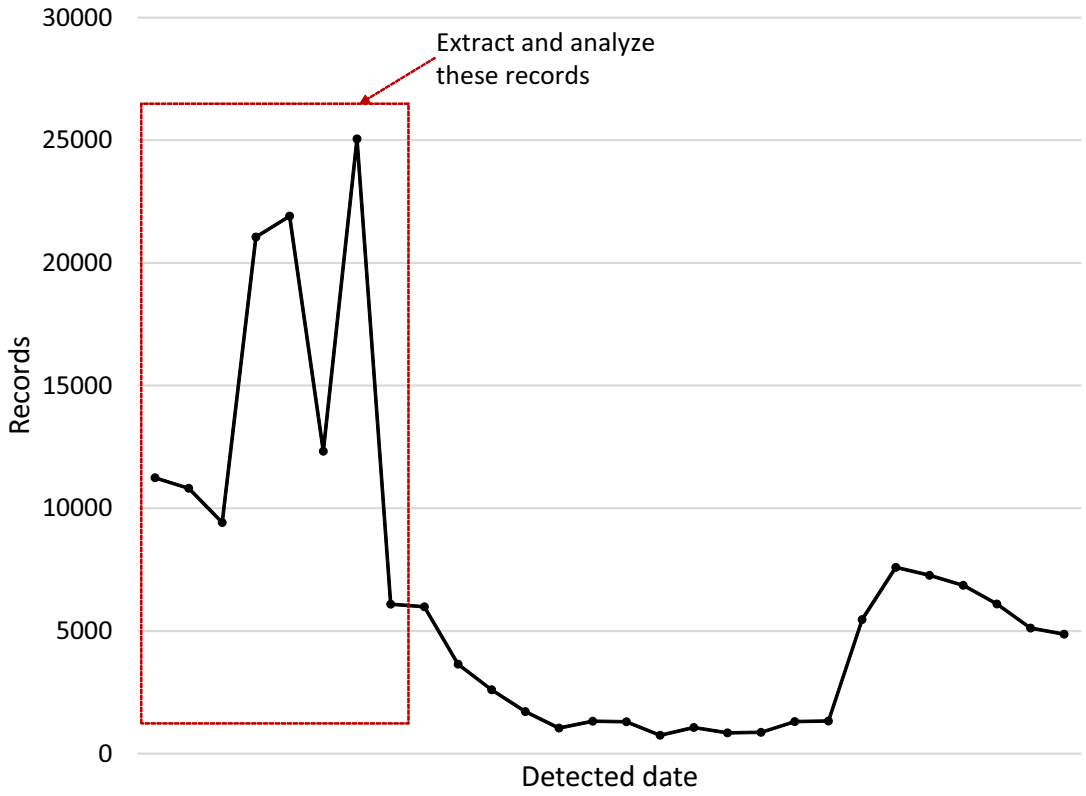

図 5.2: 調査対象のブルートフォース検知ログのレコード件数の計測結果

5.2.2 攻撃事象の統計量調査

本攻撃事象の対策に有益な特徴を得るため，企業が管理する IP アドレス帯に対する通信に対して取得できたブルートフォース検知ログを対象に，本攻撃事象に関する統計量の調査を行った結果を述べる．

対象のブルートフォース検知ログは，2011 年から 2014 年の 28 か月間において，IDS が検知した RDP サービスに対するブルートフォース攻撃を記録したレコードで構成されている．図 5.2 に，1 か月ごとのレコード件数を計測した結果を示す．この 28 か月間から最初の 8 か月分に該当する 117,924 件のブルートフォース検知ログを抽出し，攻撃事象の統計量調査を実施した．抽出したブルートフォース検知ログには，3,260 種類の送信元 IP アドレス，53 種類の送信先 IP アドレスが含まれている．

ログイン試行

まず，IDS がブルートフォース攻撃と判断したアラートにおけるログイン試行回数に着目した．図 5.3 に，1 送信元 IP アドレスから 1 送信先 IP アドレスに対して発生したログイン試行回数の合計に関する分布を示す．図 5.3 中の横軸はログイン試行回数を，縦軸は各ログイン試行回数に該当する送信元 IP アドレスと送信先 IP アドレスのペア数を示す．図 5.3 から，ログイン試行回数が 600

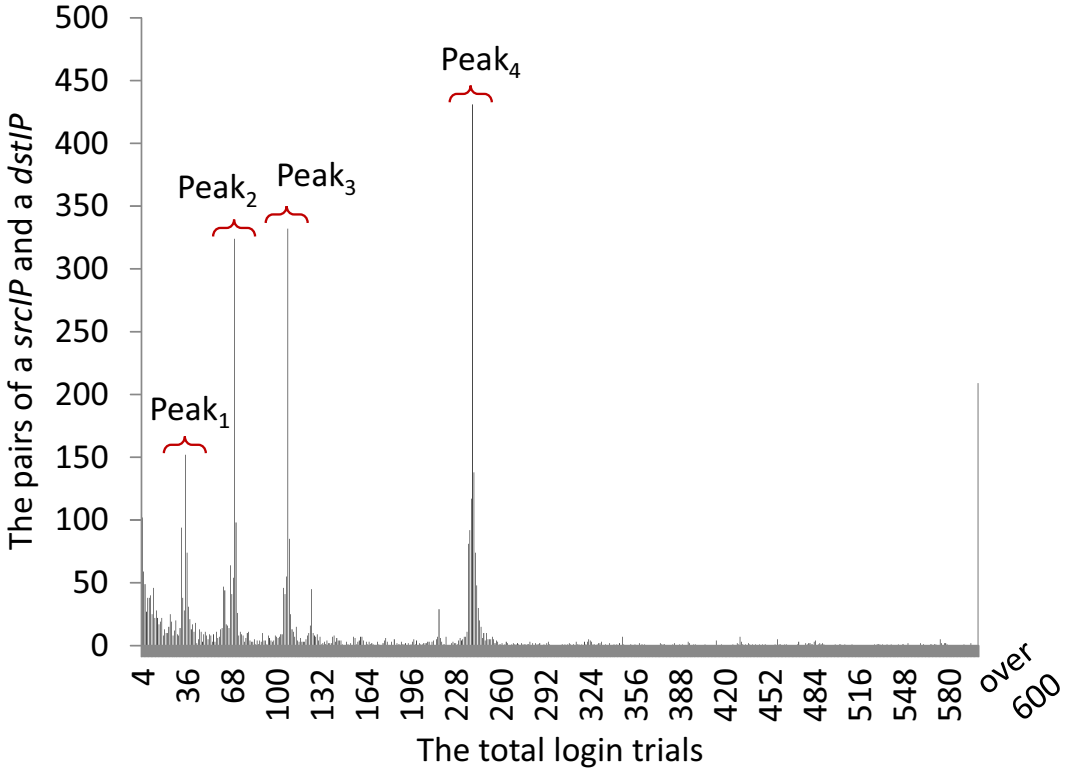

図 5.3: 1 送信元 IP アドレスから 1 送信先 IP アドレスに対して発生したログイン試行回数の合計に関する分布

回以上に該当する場合を除くと，この分布には，ピークが 4 つ存在する．以降では，この 4 つのピークを，小さい順に，Peak1，Peak2，Peak3，Peak4 とする．各ピーク値では，1 送信先 IP アドレスに対するログイン試行回数の合計値を共有する送信元 IP アドレスが集中していたことがわかる．次に，各ピークに該当する送信元 IP アドレスについて，1 送信先 IP アドレスに対する攻撃シーケンスにおける規則性について計測を行った．図 5.4 に，各ピークに該当する送信元 IP アドレスと送信先 IP アドレスのペア毎に，ログイン試行回数の平均と標準偏差（Standard deviation, STD）の分布を示す．図 5.4 から，どのピークにおいても，多くのペアが同じログイン試行回数平均と標準偏差を共有していることがわかる．ログイン試行回数の平均は 5 から 10 の範囲に存在し，標準偏差は 0 から 2 の範囲に存在している．各ピークに該当する送信元 IP アドレスと送信先 IP アドレスのペア毎に，攻撃継続時間の分布を計測した結果を図 5.5 に示す．図 5.5 より，特に Peak1 では多くのペアが攻撃継続時間を共有していることがわかる．一方で，Peak2，Peak3，Peak4 になるにつれ，同じ攻撃継続時間を共有するペアが分散していっていることもわかる．

以上より，今回分析対象としたブルートフォース検知ログでは，同じログイン試行の挙動を共有する送信元 IP アドレスと送信先 IP アドレスのペアが存在することがわかった．送信元 IP アドレスにおいては，同じログイン試行回数でブルートフォース攻撃を継続していることがわかった．また，各ペアにおけるログイン試行回数の合計の違いは，攻撃継続時間の長短に由来する．図 5.3 中の各ピークに該当するペアは，分散型ブルートフォース攻撃事象 2 に該当すると判断できる．

送信元 IP アドレスと送信先 IP アドレス

次に，このログイン試行回数に関する統計量計測結果から，分散型ブルートフォース攻撃事象 2 に該当するレコード 11,982 件を抽出した．この抽出したレコードについて，送信元 IP アドレスと送信先 IP アドレスの観点から，送信元 IP アドレス，送信先 IP アドレスとの関係，およびブルートフォース攻撃以外の IDS 検知アラートとの共起性について調査を行った．

図 5.6 に，送信元 IP アドレス，送信先 IP アドレス，IDS による検知時刻の関係を散布図として表現した結果を示す．図 5.6 において，横軸は検知時刻，縦軸は異なる送信先 IP アドレスを，散布図中の異なるドットは異なる送信元 IP アドレスによる攻撃があったことを示す．図 5.6 から，送信先 IP アドレスのいくつかは，大量の送信元 IP アドレスから代わる代わるブルートフォース攻撃を受け続けていたことがわかる．また，送信元 IP アドレスの約 93%は，1 つの攻撃シーケンスについてのみ攻撃元となっていた．

ブルートフォース攻撃以外の IDS 検知アラートとの共起性について調査を行った結果を示す．表 5.1 に，送信元 IP アドレスで，ブルートフォース攻撃以外のアラートを集計した結果を示す．分散型ブルートフォース攻撃事象 2 に特有の共起事象か否かを検証するために，ブルートフォース攻撃が検知された送信元 IP アドレス全体に対して他アラートを集計した結果（all）と，抽出したレコードに含まれる送信元 IP アドレスに対して集計した結果（分散型ブルートフォース攻撃事象 2）を示す．表 5.1 から，ブルートフォース攻撃を検知された送信元 IP アドレスが他に検知された攻撃は TCP3389 番ポートが開いているホストを検索する行為（Host sweep (3389/tcp)）が一番多かったものの，該当する送信元 IP アドレスは全体の 1.81%に過ぎなかった．よって，ブルートフォース攻撃と共起して発生していた攻撃は存在しなかったといえる．

また，送信先 IP アドレスについても，同様にブルートフォース攻撃以外の IDS 検知アラートを集計した結果を表 5.2 に示す．集計の結果，送信元 IP アドレスの場合とは対照的に，ブルートフォース攻撃が検知された送信先 IP アドレス全体の集計結果と，本節で抽出したレコードに含まれる送信先 IP アドレスに対して集計した結果と大きな差は得られなかった．

表 5.1: 送信元 IP アドレスが分散型ブルートフォース攻撃事象 2 と共起して検知されたアラート（上位 5 位）

	Alert	*srcIP*s(%)
all	Host sweep (3389/tcp)	1.81
	Netbios scan (137/udp)	0.77
	Host sweep (1433/tcp)	0.31
	SMB(Netbios) service scan (445/tcp)	0.18
	SMB service connect (445/tcp)	0.12
Distributed brute force attacks	Netbios scan (137/tcp)	1.76
	Host sweep (3389/tcp)	0.35
	Malicious scanning (40235/tcp)	0.35
	DNS ETC type query flooding (53/udp)	0.18
	Host Sweep (1433/tcp)	0.18

表 5.2: 送信先 IP アドレスが分散型ブルートフォース攻撃事象 2 と共起して検知されたアラート（上位 5 位）

Alert	*dstIP*s(%)
Netbios Scan (137/udp)	100
MSSQL Server2000 Resolution Service DOS (1434/udp)	100
Slamer Worm (1434/udp)	100
SMB Service sweep (445/tcp)	79.25
Trace Route (0/tcp)	77.36

5.3 攻撃検知・対策システムの提案

本章では，分散型ブルートフォース攻撃事象 2 を検知し，ブルートフォース攻撃の発生を最小限に抑える手法を提案する．

5.3.1 提案手法の概要

提案手法の全体構成を図 5.7 に示す．提案手法は，複数のサーバに対する通信を IDS が監視し，IDS によりブルートフォース攻撃と検知されたアラートからなるブルートフォース検知ログを入力とする．

提案手法の処理は抽出ステップと遮断ステップから構成される．まず最初に抽出ステップでは，入力としたブルートフォース検知ログからブルートフォース攻撃事象の規則性に基づき，ブルートフォース攻撃を受けている送信先 IP アドレスを抽出する．ここで抽出できた送信先 IP アドレスは，次の遮断ステップで使用する．遮断ステップでは，前の抽出ステップで抽出できた送信先 IP アドレスに対する IDS アラートを監視する．監視対象となった送信先 IP アドレスに対するブルートフォース攻撃を検知する IDS によるアラートが，攻撃シーケンスの開始と判断されたならば，攻撃シーケンスの送信元 IP アドレスによる通信を一定時間遮断する．

提案手法では，ブルートフォース攻撃事象に該当する攻撃シーケンスを開始時に検知し遮断することにより，ブルートフォース攻撃が送信先 IP アドレスに到達するのを防ぐことががができる．また，攻撃シーケンスの終了を見込んだ時間だけ遮断対象とすることで，ブラックリストを作成して送信元 IP アドレスを追加していくのに比べ，リストサイズの爆発を抑えることができる．

5.3.2 提案手法の処理手順

提案手法は，抽出ステップと遮断ステップの 2 ステップから構成される．

抽出ステップでは，入力となるブルートフォース検知ログからブルートフォース攻撃事象の対象となっている送信先 IP アドレスを抽出する．このブルートフォース攻撃事象では，各攻撃シーケンスは同じログイン試行の挙動を共有する．そこで本ステップでは，この攻撃シーケンスを見つけるため，入力のブルートフォース検知ログに含まれる送信元 IP アドレスと送信先 IP アドレスの組み合わせについて統計量を計測し，統計量を共有する送信元 IP アドレスと送信先 IP アドレスの組み合わせを抽出する．

まず，送信元 IP アドレスと送信先 IP アドレスの組み合わせについてログイン試行回数の合計を計算し，ログイン試行回数の合計を共有する組み合わせを抽出する．図 5.3 で実施したように，ロ

グイン試行回数の合計に関する分布を取り，互いに多くのログイン試行回数の合計を共有する組み合わせを抽出する．

次に，ログイン試行回数の合計を共有する組み合わせから，攻撃シーケンスに該当する組み合わせを絞り込む．ログイン試行回数の合計を共有する組み合わせから，1 レコード当たりのログイン試行回数の平均，標準偏差を計測する．これらの統計量を特徴量ベクトルとみなし，クラスタリングすることで，類似する統計量を持つ組み合わせを特定する．前節で実施した分析により，攻撃シーケンスに該当する組み合わせは，ログイン試行回数の平均が 10 付近に，標準偏差が 1 付近に集中することがわかっている．攻撃シーケンスに該当する組み合わせを特定するには，ログイン試行回数の平均および標準偏差が上記の値に最も近いクラスタに所属する組み合わせを抽出する．抽出できた組み合わせから，送信先 IP アドレスを抽出し，ブルートフォース攻撃事象の対象となっている送信先 IP アドレスとする．

遮断ステップでは，抽出ステップで抽出できた送信先 IP アドレスに対する通信を重点的に監視する．監視対象となった送信先 IP アドレスに対して発生したブルートフォース攻撃について，そのログイン試行やログイン試行間の間隔が攻撃シーケンスの持つ挙動と一致したならば，そのブルートフォース攻撃元となった送信元 IP アドレスからの通信を一定期間遮断する．

この遮断ステップでは，提案手法のユーザは次の 4 種類のパラメタを設定する．

1. ログイン試行回数と攻撃シーケンスの開始を判断するまでの範囲
2. 攻撃シーケンスと判断するためのログイン試行回数の標準偏差
3. 通信遮断を開始するまでの IDS アラート件数
4. 攻撃遮断を終了するまでの時間

上記のしきい値の中で 1)，2)，4) については，前節で分析できた結果からパラメタを決定する．

5.4 提案手法の評価

ブルートフォース検知ログを対象として，提案手法の遮断ステップにより遮断できるブルートフォース攻撃を見積もることで評価を行う．本章で評価対象とする項目としてドロップ率，過剰遮断時間の 2 項目を用意する．まずドロップ率は，攻撃シーケンスに含まれるブルートフォース攻撃のうち，提案手法が遮断できないブルートフォース攻撃の割合を示す．このドロップ率が小さいほど，提案手法が攻撃シーケンスに含まれるブルートフォース攻撃を遮断できることを示す．過剰遮断時間では，当該事象を検知してから攻撃を遮断する時間のうち，当該事象が終了したにも関わらず遮断を継続している時間を示す．過剰遮断時間が小さいほど，遮断対象となる送信元 IP アドレスを手法の内部に保持すべき時間が短くなるため，提案手法の処理負荷を軽減できる[1]．

本評価実験では，図 5.2 から 5 番目，8 番目，17 番目，23 番目の月に収集できたブルートフォース検知ログを対象とする．前者の 2 か月に該当するブルートフォース検知ログは，第 5.2 節で分析したログの範囲に含まれているが，後者の 2 か月に該当するブルートフォース検知ログは，第 5.2 節で分析したログの範囲に含まれていない．そのため，ブルートフォース攻撃事象に該当するレコードを第 5.2 節と同じ手順で抽出した．表 5.3 に，評価対象とするブルートフォース検知ログの

[1]本評価実験にて評価項目として設定した過剰遮断時間は提案手法の誤検知を計測するものではない．提案手法の誤検知を計測するには，ネットワークサービスの管理者がブルートフォース検知ログに対して，正規ユーザによる正当なログイン操作か，悪意あるログイン試行に由来するものかを区別できる環境で提案手法を評価する必要がある．

レコード件数，送信元 IP アドレス種類数，送信先 IP アドレス種類数を示す．これら 4 種類のブルートフォース検知ログに対して，遮断開始までのレコード数（S）と遮断時間（P）を変えながら，ドロップ率および過剰遮断時間を計測する．第 5.2 節での分析結果に従い，ログイン試行回数を 4〜10，ログイン試行回数の標準偏差を 0〜2 の範囲に設定し，評価実験を行う．

S を 1, 2, 3 および P を 20, 40, 60 に設定した上で，提案手法を対象とした 4 カ月間に適用した結果を比較する．図 5.8，図 5.9，図 5.10，図 5.11 に，ドロップ率と過剰遮断時間の平均を示す．ドロップ率について結果を述べる．提案手法が 1 アラートに相当するブルートフォース攻撃を遮断しなかった場合，平均してブルートフォース攻撃事象の約 12.3%が送信先 IP アドレスへ到達した．ドロップ率を提案手法を適用した 4 カ月間で比較すると，1 番目の月が最も低い．これは，1 番目の月には他の月よりも長期間継続したブルートフォース攻撃事象が多く含まれていたからである．過剰遮断時間について結果を述べる．1 番目の月を除くと，P が 20 から 40 に変化した場合，過剰遮断時間は平均 19.4 分増加した．P が 40 から 60 に変化した場合，過剰遮断時間は平均 19.9 分増加した．この結果から，遮断期間の増加分がほぼ過剰遮断時間となったことがわかる．

この評価実験では，ドロップ率および過剰遮断時間を最小に抑えられるパラメタは，S=1，P=20 に設定したときであるといえる．P を長く設定すると，過剰遮断時間が長くなり，監視対象とするネットワークサービスユーザの利便性が低下する．提案手法では，正規ユーザが特定の回数サービスに対するログインに失敗すると，IDS ではこの挙動をブルートフォース攻撃として検知する．その後に，提案手法でこのブルートフォース検知ログを受信し分散型ブルートフォース攻撃事象が発生したと判断し，通信遮断を開始してしまうと，この正規ユーザは通信遮断が終わるまでそのサービスにログインできない．設定した遮断時間よりも長く続く分散型ブルートフォース攻撃事象が発生した場合，通信遮断終了後であっても，ブルートフォース攻撃を検知したアラートの発生を契機として，再度通信遮断を開始する．この場合，最初の通信遮断終了から次の通信遮断開始までに発生したブルートフォース攻撃を遮断できない．

表 5.3: 評価対象とするブルートフォース検知ログのレコード件数，送信元 IP アドレス種類数，送信先 IP アドレス種類数

Month	Records	unique *srcIP*	unique *dstIP*
5th	21,905	709	21
8th	6,095	366	21
17th	1,068	111	21
23rd	7,591	528	19

5.5 議論

5.5.1 分散型ブルートフォース攻撃事象 1 との相違点

第 4 章で報告した分散型ブルートフォース攻撃事象と，本章で対象とする事象の相違点を述べる．第 4 章で報告した事象では，ある送信元 IP アドレスが複数の送信先 IP アドレスに対して同時にブルートフォース攻撃を行う．この攻撃の対象となった送信先 IP アドレスには，ログイン試行回数の相関も存在する．しかし，送信元 IP アドレスは攻撃を継続する場合とそうでない場合がある．さらに，攻撃を継続した場合であっても，ログイン試行回数は同一ではない．一方で，本章で対象とする事象では，ある送信元 IP アドレスは 1 つの送信先 IP アドレスにしかブルートフォース攻撃

を行わない．しかし，1 送信元 IP アドレスと 1 送信先 IP アドレス間に発生したブルートフォース攻撃は継続して発生しており，攻撃頻度およびログイン試行回数には規則性が存在する．

第 4 章で提案したブルートフォース攻撃検知・遮断手法を，本章で対象とするブルートフォース攻撃事象に対して適用したとしても，本事象を対策することは難しい．また，本章で述べた提案手法を，ブルートフォース攻撃事象 1 を適用する場合について議論する．ブルートフォース攻撃事象 1 では，送信元 IP アドレスは複数の送信先 IP アドレスに対して，同時刻に同じログイン試行回数でブルートフォース攻撃を実行する．さらに，送信元 IP アドレスは攻撃を継続しない場合もある．この攻撃の対策として，第 4 章では検知時刻とログイン試行回数の相関に着目し，攻撃対象となる送信先 IP アドレスを抽出する手法を提案した．一方で，本章で対象とするブルートフォース攻撃の対策では，継続して発生するブルートフォース攻撃のにおけるログイン試行回数の規則性に基づき，攻撃を受け続ける送信先 IP アドレスを抽出する．そのため，この本章で述べた提案手法では送信先 IP アドレス間の相関関係を分析しないため，ブルートフォース攻撃 1 を受ける送信先 IP アドレスを抽出することができない．また，本章で述べた提案手法における通信遮断時の判断は，ログイン試行回数の規則性に基づいており，攻撃先となる送信先 IP アドレスが既知であるか，攻撃が複数送信先 IP アドレス間で同期して発生しているか，といった条件には基づいていない．そのため，ブルートフォース攻撃事象が規則正しく攻撃を継続しない限り，本章で述べた提案手法は攻撃の発生を検知できない．

5.5.2 提案手法を認知している攻撃者に向けた対策

本章で述べた提案手法を認知している攻撃者に向けた対策について述べる．提案手法による検知を回避するため，攻撃者が個々のブルートフォース攻撃のログイン試行回数やブルートフォース攻撃間の頻度をランダムに設定する手段が考えられる．その場合，提案手法ではこうしたランダムな頻度とログイン試行回数を持つブルートフォース攻撃を検知できないという限界がある．

この限界を緩和するため，到達したブルートフォース攻撃に対して，文献 [95] で述べられているような，ランダム性を検証する処理を提案手法の抽出処理に追加する手段が挙げられる．しかしこの場合，ブルートフォース攻撃のランダム性を判断するため，分散型ブルートフォース攻撃事象の発生から通信遮断を開始するまでにより多くの時間が掛かるというデメリットが発生する．

5.6 結論

本章では，企業が管理する IP アドレス帯で実際に RDP サービスを運用しているサーバ群に対するブルートフォース検知ログから，送信元 IP アドレスを変えながら共通の規則性を持ってブルートフォース攻撃を繰り返す事象を報告した．さらに，この規則性を有する送信元 IP アドレスを特定することで，この事象によりブルートフォース攻撃を遮断する手法を提案した．提案手法をこのブルートフォース検知ログに適用し，ドロップ率と過剰遮断時間を最小限に抑えるパラメタを見積もることができた．

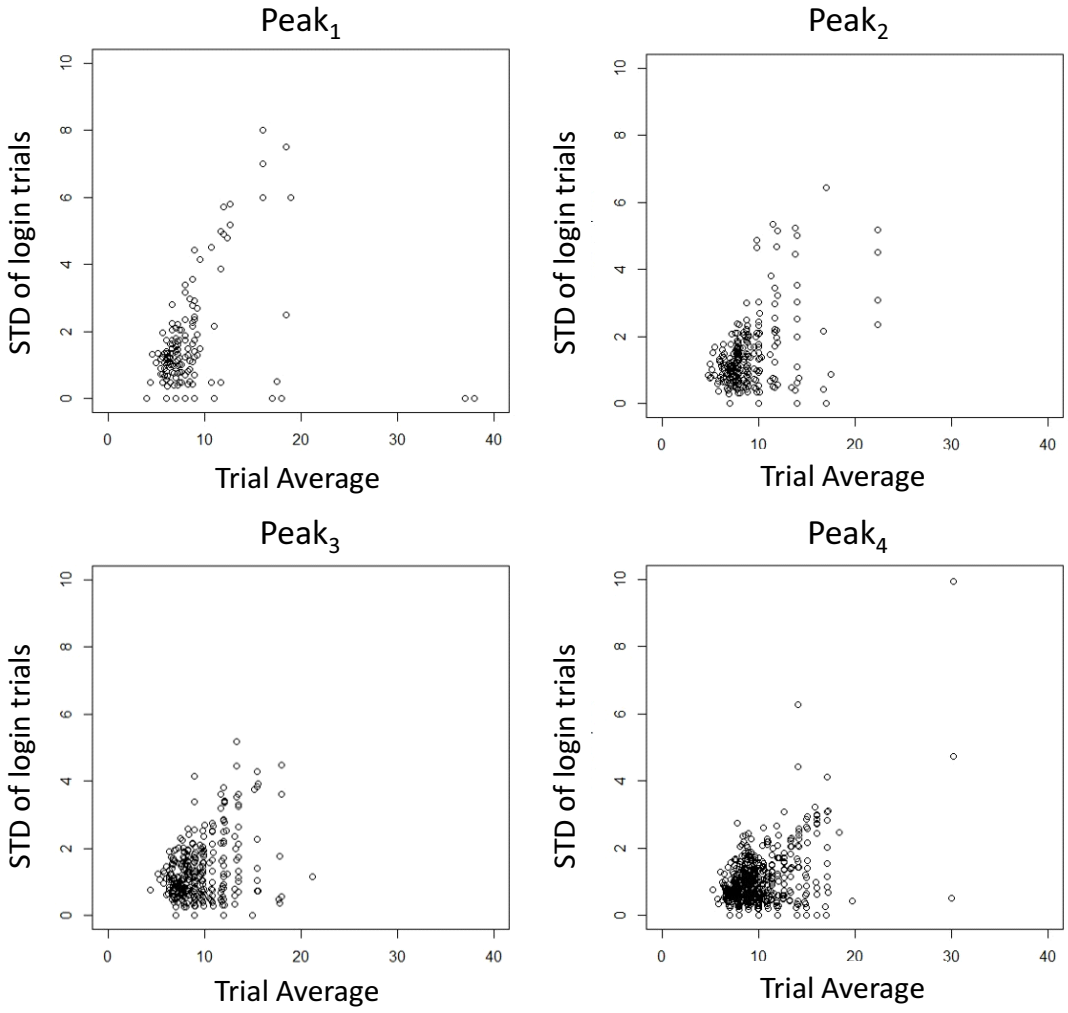

図 5.4: 各ピークに該当する送信元 IP アドレスと送信先 IP アドレスのペアにおけるログイン試行回数の平均と標準偏差の分布

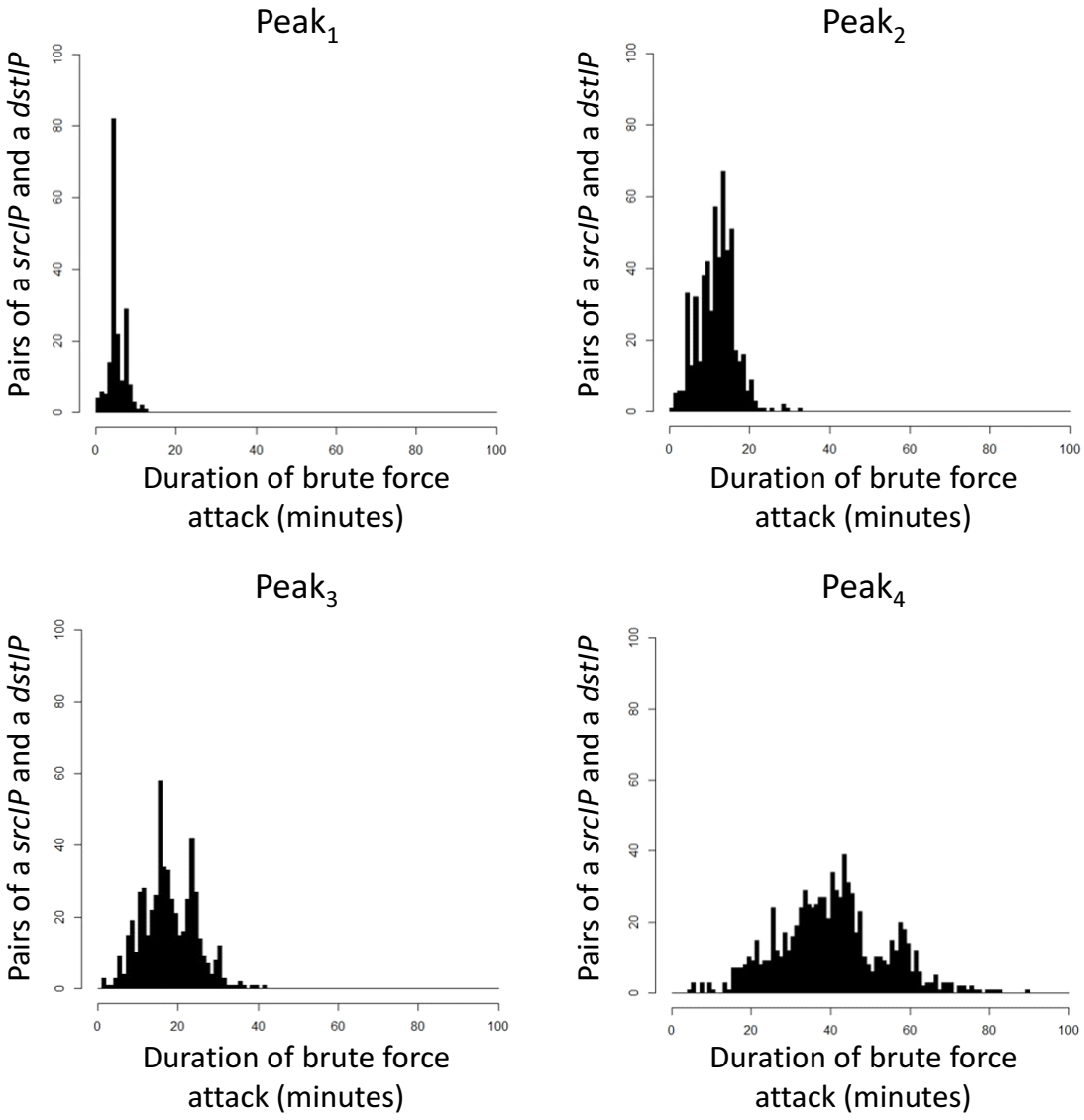

図 5.5: 各ピークに該当する送信元 IP アドレスと送信先 IP アドレスのペアにおける攻撃継続時間の分布

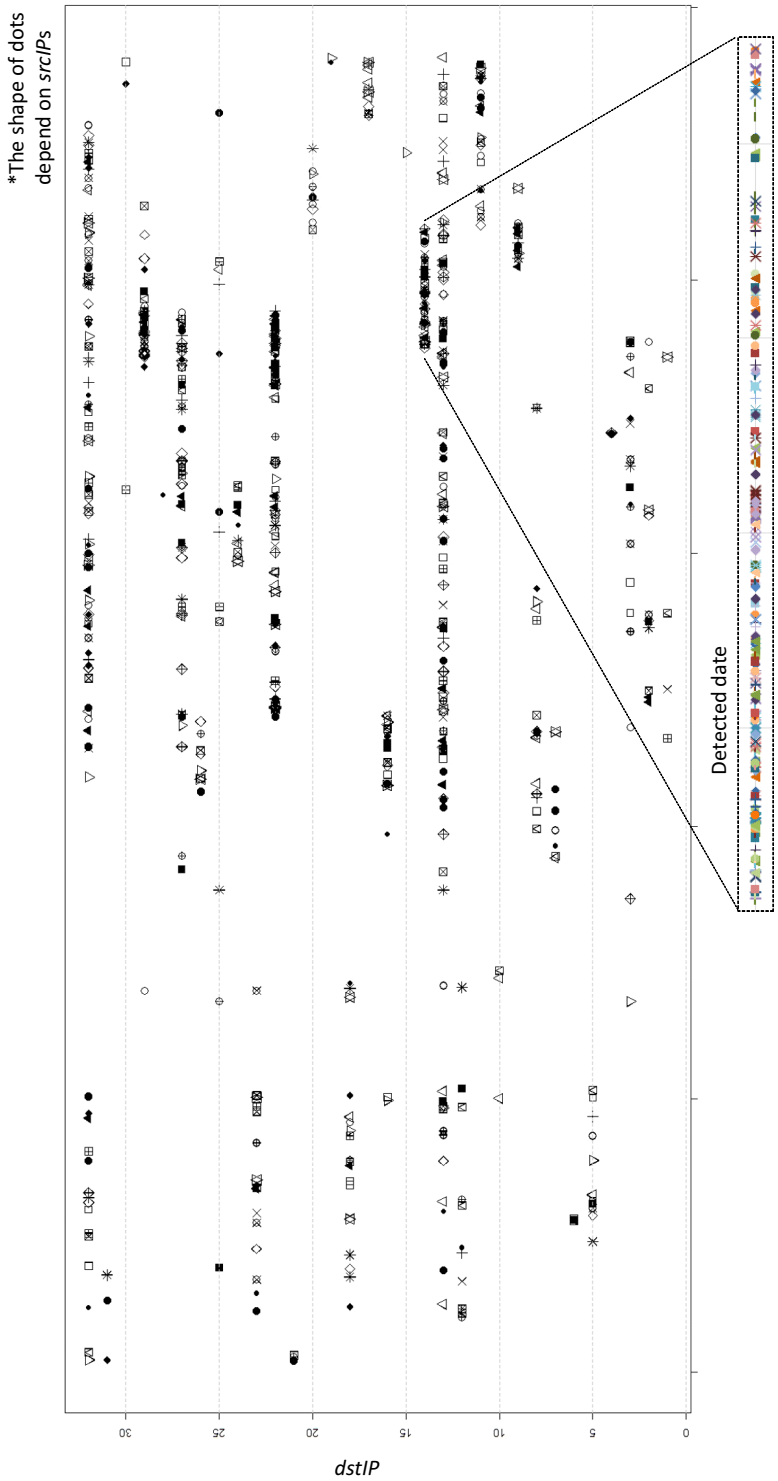

図 5.6: 分散型ブルートフォース攻撃事象 2 に該当するブルートフォース検知ログにおける送信元 IP アドレス・送信先 IP アドレス・検知時刻の関係

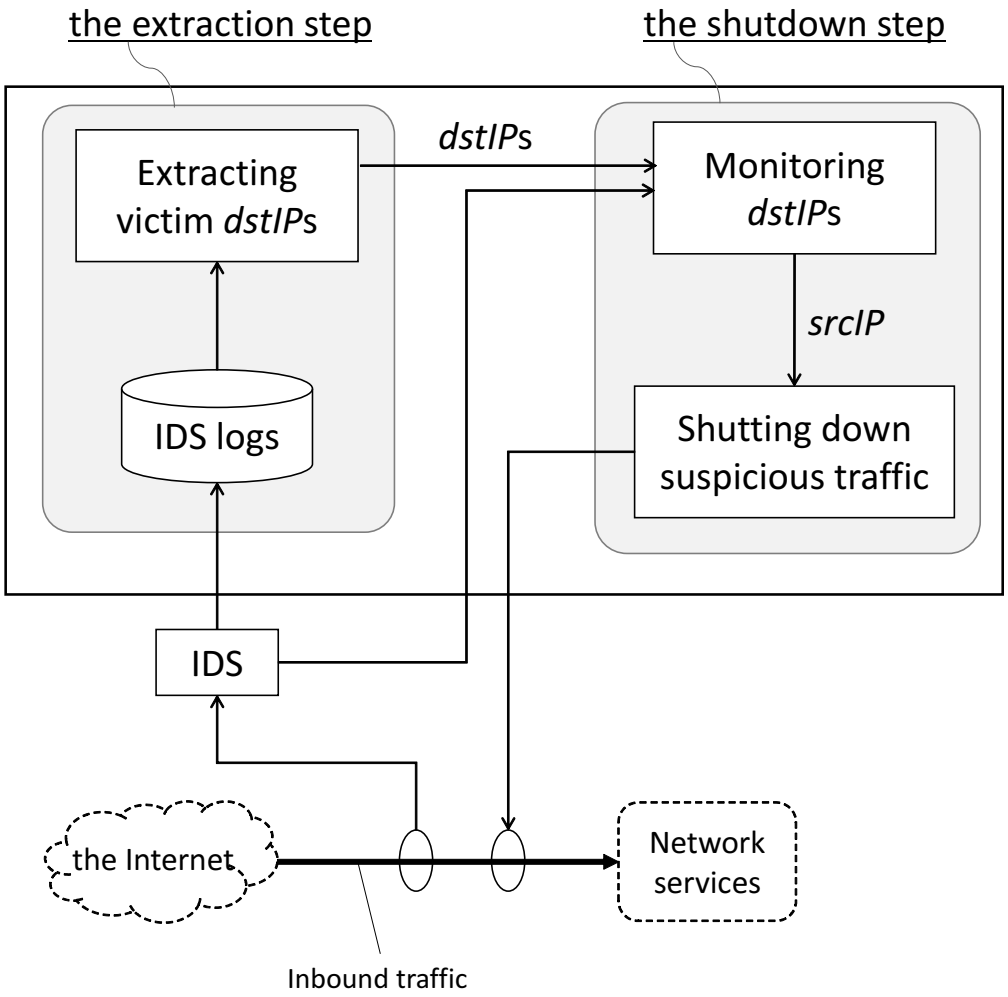

図 5.7: 提案手法の全体構成

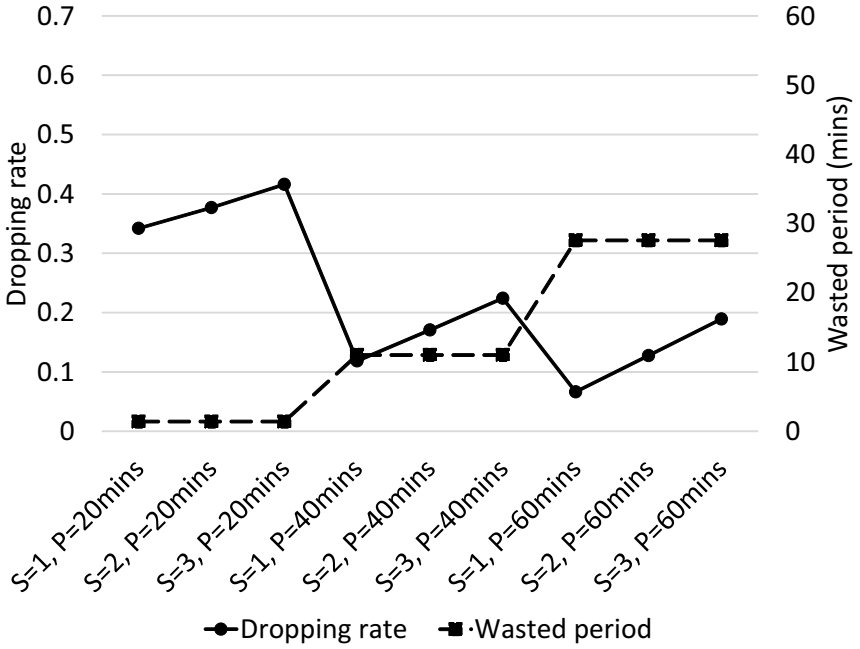

図 5.8: 5 番目の月におけるドロップ率および過剰遮断時間の平均の変化

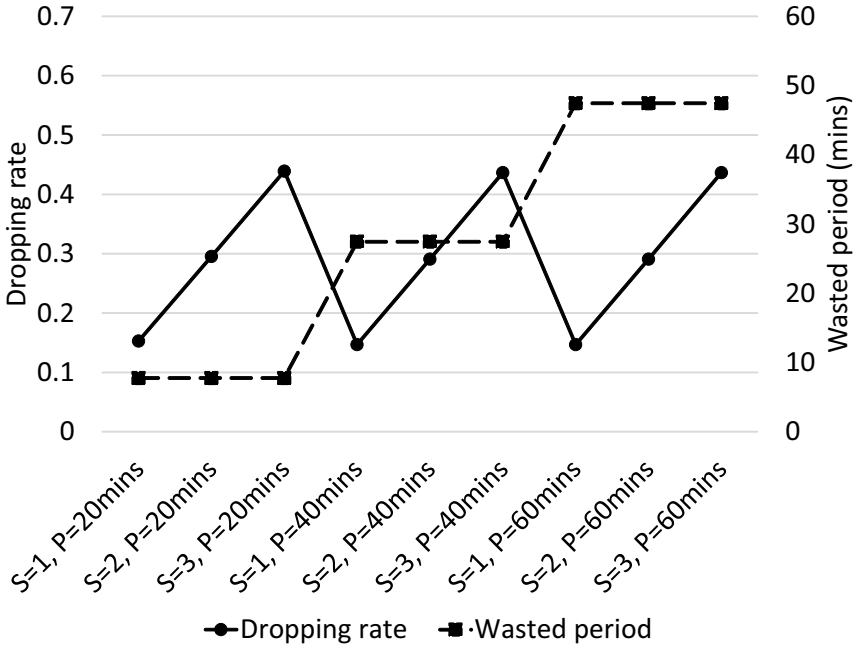

図 5.9: 8 番目の月におけるドロップ率および過剰遮断時間の平均の変化

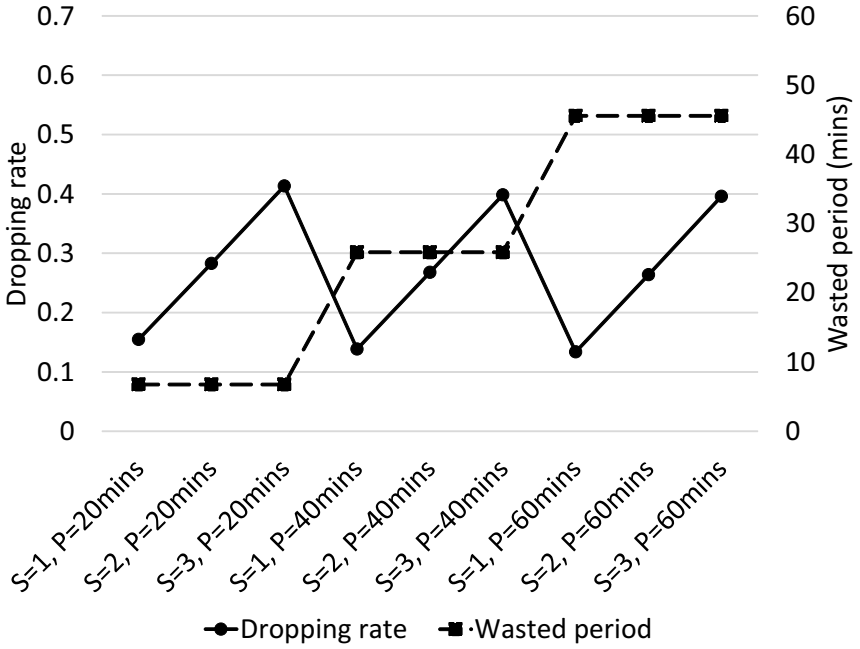

図 5.10: 17 番目の月におけるドロップ率および過剰遮断時間の平均の変化

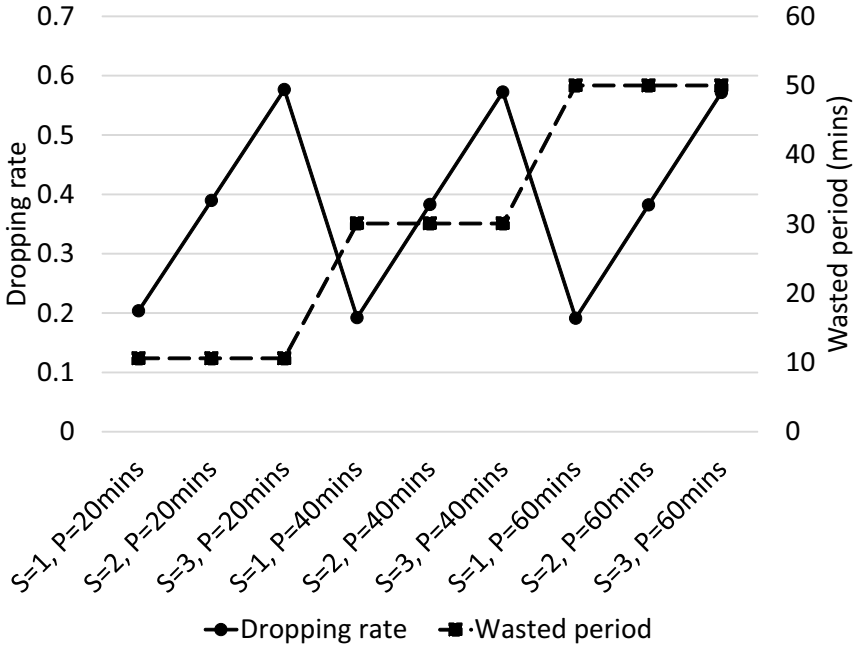

図 5.11: 23 番目の月におけるドロップ率および過剰遮断時間の平均の変化

第6章　結論

6.1　本研究で得られた成果

本研究では，正規通信，不正通信および意図不明な通信が混在するセキュリティログから，インシデントとして対応すべき攻撃事象の抽出を目的として，分析技術の提案および評価を行った．本研究のアプローチとして，複数拠点に存在するサーバから取得できたセキュリティログを対象に，サーバ間の関係性から共通性を見つけることで，攻撃事象の抽出を行う．本研究では，一般によく使われるサービスである，Web サイト，SSH サービス，RDP サービスの 3 種類のネットワークサービスを対象に，攻撃検知手法を提案し，実環境より取得できたサーバログを用いて評価を行った．

まず，複数の Web サイトを監視対象とし，単一の送信元から複数の Web サイトに向けて送信されたリクエストに着目することで，脆弱性を持つ Web アプリケーションの探索や，悪意あるコード挿入を行うリクエストで，単一の送信元から送信される際にその URI にパターンが存在する攻撃を検知する手法を提案した．提案手法を，横浜国立大学情報基盤センターが管理している学内向け Web ホスティングサービスより取得できたアクセスログを用いて評価を行った結果，攻撃元となった IP アドレスを誤検知なく抽出できるしきい値が存在することを示した．さらに，既存の攻撃検知ツールでは悪性と判断しなかったものの，他文献により悪性の可能性が高いリクエストを悪性と判断できることを示した．

次に，様々な意図や規模の攻撃が混在するブルートフォース検知ログから，ログイン試行回数や送信元 IP アドレスの出現回数などの集計では現れない，ステルス性の高い分散型ブルートフォース攻撃事象の抽出および検知を行う手法を提案した．まず，ブルートフォース検知ログから，送信元 IP アドレスと送信先 IP アドレスの関係性を散布図形式で可視化することで，送信元 IP アドレスを変えながら特定の送信先 IP アドレス群に対してブルートフォース攻撃を繰り返す事象を抽出できた．さらに，この攻撃事象を受ける IP アドレスを早期に検知する手法を提案した．この手法を，企業が管理する IP アドレス帯で実際にサービスを運用している複数の SSH サービスに対して記録されたブルートフォース検知ログに適用した結果，検知時刻とログイン試行回数の相関を取る場合に，最も精度良く送信先 IP アドレスを特定できることを示した．

また，ブルートフォース検知ログから，送信元 IP アドレスによるブルートフォース攻撃アラートのインターバルやアラートが検知した攻撃回数の分散を計測することで，送信元 IP アドレスを変えながら，共通の規則性を以ってブルートフォース攻撃を繰り返す事象を抽出できた．さらに，この規則性を有する送信元 IP アドレスを特定することで，この事象でブルートフォース攻撃の発生を遮断する手法を提案した．この手法を，企業が管理する IP アドレス帯で実際にサービスを運用している複数の RDP サービスに対して記録されたブルートフォース検知ログに適用した結果，ドロップ率と過剰遮断時間を最小限に抑えるパラメタを見積もることができた．

以上の結果から，複数の拠点間にまたがる共通性を見つけることで，既存のツールでは検知が難しい，あるいは認知されていなかった攻撃事象を抽出できた．

6.2 今後の課題

今後の課題として，異なる種類のログを対象とした分析と，大規模データ対応のためのスケールアウトの 2 点が挙げられる．

本研究では，アクセスログやブルートフォース検知ログといった，単一種類のログを手法の対象としていた．しかし，アクセスログとブルートフォース検知ログや，別のセキュリティログを組み合わせたものを対象として分析することで，攻撃事象の発生を高い精度で検知できたり，攻撃事象を多面的に把握できるといったことができると考えられる．

また，分析対象とするログの規模が大きくなると，攻撃事象抽出にかかる時間が長くなり，インシデントとしての対応が遅れる恐れもある．そのため，分析をスケールアウトできるようにし，分散して処理を行うことで攻撃抽出にかかる時間を短縮する，さらには攻撃事象の局所性を判断することで，不要な抽出処理を省略するといった機能の追加が考えられる．

謝辞

本研究を進めるにあたり，多大なご指導とご助言を頂きました，横浜国立大学大学院環境情報研究院　松本勉教授，吉岡克成准教授に深く感謝致します．未熟な私に，親身に温かくご指導頂きましたおかげで，本研究をまとめることができました．

本論文をまとめるにあたり，論文審査をお引き受け頂き，貴重なご指導，ご助言を頂きました，横浜国立大学大学院環境情報研究院　森辰則教授，四方順司教授，白川真一講師に深く感謝致します．特に四方教授には，IPS 研究会の場にても多大なご指導とご助言を頂き，また私の学部生時代のコンタクト教員としても，長くご指導頂きました．

本研究を進めるにあたり，貴重なご議論，ご助言を頂きました横浜国立大学先端科学高等研究院情報・物理セキュリティ研究ユニット　Christopher Kruegel 上席特別教授，Engin Kirda IAS 連携教授，Michel van Eeten IAS 連携教授，William Robertson IAS 連携助教に深く感謝致します．先生方と直接ご議論ができ，自分の研究を改めて国際基準で捉えることも体験できました．

本研究を進める上で日頃より議論に協力して頂きました，松本研究室，四方研究室，吉岡研究室の皆様に深く感謝致します．また，研究活動に際し多大なご援助を頂きました秘書の成松美央氏，技術補佐員の石舘知子氏に深く感謝致します．

本論文の第 3 章では，提案手法の評価に横浜国立大学情報基盤センターより提供頂きました Web ホスティングサービスのアクセスログを利用致しました．アクセスログを提供頂きました横浜国立大学情報基盤センターの皆様に深く感謝致します．

私に研究の機会を与えてくださり，大学院通学を応援して頂きました，株式会社富士通研究所セキュリティ研究所　鳥居悟特任研究員，武仲正彦所長および研究員の皆様，富士通株式会社の技術者の皆様に深く感謝致します．

最後に，家族として，また博士課程後期卒業生の先輩として，私の研究活動を応援頂きました，夫の齊藤航太氏に深く感謝致します．

参考文献

［1］https://www.jpcert.or.jp/ir/, ”JPCERT コーディネーションセンター インシデント対応とは？“, last visited 2017/11/02.

［2］https://www.nic.ad.jp/ja/basics/terms/soc.html, ”インターネット用語 1 分解説～SOC とは～ - JPNIC”, last visited 2017/11/02.

［3］Tim Crothers, ”Implementing Intrusion Detection Systems: A Hands-On Guide for Securing the Network.” John Wiley & Sons, pp.1-pp.2, 2002.

［4］Aleksandar Lazarevic, Arindam Banerjee, Varun Chandola, Vipin Kumar, Jaideep Srivastava, ”Data Mining for Anomaly detection.” Tutorial at the European Conference on Principles and Practice of Knowledge Discovery in Databases, 2008.

［5］Chandola Varun, Arindam Banerjee, and Vipin Kumar, ”Anomaly detection: A survey.” ACM Computing Surveys, Vol.41, No.3, Article 15, 2009.

［6］山西健司, ”データマイニングによる異常検知.” pp.3-pp.5 and pp.8-pp.9, 共立出版株式会社, 2009.

［7］https://goaccess.io/, ”GoAccess - Visual Web Log Analyzer.” last visited 2017/12/13.

［8］https://analytics.google.com, ”Google Analytics.” last visited 2017/12/13.

［9］https://analytics.google.com/analytics/academy/course/6, ”Google アナリティクス初級者向けコース.” last visited 2017/12/13.

［10］Christopher Kruegel, Giovanni Vigna, and William Robertson. ”A multi-model approach to the detection of web-based attacks.” Computer Networks 48.5 (2005): 717-738.

［11］鐘場, 折原慎吾, 谷川真樹, 嶋田創, 村瀬勉, 高倉弘喜, 大嶋嘉人, ”URI の共起性に基づく Web スキャンの実態調査.” 信学技報 IEICE Technical Report, ICSS2015-51(2016-03), 2016.

［12］Cho Sanghyun, and Sungdeok Cha. ”SAD: web session anomaly detection based on parameter estimation.” Computers & Security 23.4 (2004): 312-319.

［13］Stevanovic Dusan, Natalija Vlajic, and Aijun An, ”Detection of malicious and non-malicious website visitors using unsupervised neural network learning.” Applied Soft Computing 13.1 (2013): 698-708, 2013.

［14］Adam Kieyzun, Philip J. Guo, Karthick Jayaraman, and Michael D. Ernst. ”Automatic creation of SQL injection and cross-site scripting attacks.” 2009 IEEE 31st International Conference on Software Engineering. IEEE, 2009.

［15］Prithvi Bisht, Timothy Hinrichs, Nazari Skrupsky, Radoslaw Bobrowicz, and V.N. Venkatakrishnan. ”NoTamper: automatic blackbox detection of parameter tampering opportunities in web applications.” Proceedings of the 17th ACM conference on Computer and communications security. ACM, 2010.

［16］Adam Doupe, Ludovico Cavedon, Christopher Kruegel, and Giovanni Vigna. ”Enemy of the State: A State-Aware Black-Box Web Vulnerability Scanner.” Presented as part of the 21st USENIX Security Symposium (USENIX Security 12). 2012.

［17］Adam Doupe, Marco Cova, and Giovanni Vigna. ”Why Johnny can’t pentest: An analysis of black-box web vulnerability scanners.” International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment, 2010.

［18］Jason Bau, Elie Bursztein, Divij Gupta, and John Mitchell. ”State of the art: Automated black-box web application vulnerability testing.” 2010 IEEE Symposium on Security and Privacy. IEEE, 2010.

［19］John P. John, Fang Yu, Yinglian Xie, Arvind Krishnamurthy, and Martin Abadi.

”Heat-seeking honeypots: design and experience.” Proceedings of the 20th international conference on World wide web. ACM, 2011.

［20］Takeshi, Yagi, Naoto Tanimoto, and Takeo Hariu,

”Intelligent high-interaction web honeypots based on URL conversion scheme.” IEICE transactions on communications 94.5 (2011): 1339-1347, 2011.

［21］久世尚美, 石倉秀, 八木毅, 千葉大紀, 村田正幸, ”複数のハニーポットにおいて観測された情報に基づく通信のネットワーク上の特徴を考慮したぜい弱性スキャン識別.” 信学技報, vol. 115, no. 488, ICSS2015-55, pp. 47-52, 2016 年 3 月.

［22］Davide Canali, and Davide Balzarotti, ”Behind the scenes of online attacks: an analysis of exploitation behaviors on the web.” in Proceedings of Network and Distributed System Security Symposium (NDSS), 2013.

［23］https://github.com/mushorg/glastopf, ”GitHub - mushorg/glastopf: Web Application Honeypot.” last visited 2017/4/10.

［24］http://dionaea.carnivore.it/, ”dionaea, catches bugs.” last visited 2016/6/15.

［25］https://sites.google.com/site/webhoneypotsite/, ”DShield Web Honeypot Project.”

［26］Michael Muter, Felix Freiling, Thorsten Holz, and Jeanna Matthews, ”A generic toolkit for converting web applications into high-interaction honeypots.” http://people.clarkson.edu/ jnm/publications/honeypot-raid2007.pdf.

［27］Kreibich Christian, and Jon Crowcroft. ”Honeycomb: creating intrusion detection signatures using honeypots.” ACM SIGCOMM computer communication review 34.1 (2004): 51-56., 2004.

［28］Anna Sperotto, Ramin Sadre, Pieter-Tjerk de Boer, Aiko Pras, “ Hidden Markov Model modeling of SSH brute-force attacks.” International Workshop on Distributed Systems: Operations and Management 2009 (DSOM 2009): Integrated Management of Systems, Services, Processes and People in IT , pp164-176, 2009.

［29］J Vykopal,“A Flow-Level Taxonomy and Prevalence of Brute Force Attacks.”American Control Conference 2011(ACC 2011) Part II, pp666-675, 2011.

［30］Mobin Javed, Vern Paxson, ”Detecting Stealthy, Distributed SSH Bruteforcing.” 2013 ACM SIGSAC conference on Computer & communications security, pp85-96, 2013.

［31］SANS Internet Storm Center, ”ISC Diary — Distributed SSH Brute Force Attempts on the rise again.” https://isc.sans.edu/diary/Distributed+SSH+Brute+Force+Attempts+on+the+rise+again/9031, 2010.

［32］Dragon Research Group, ”SSH Brute Force Attack Source Insight（2011-04-29）.” http://www.dragonresearchgroup.org/ 2011/04/29/,2011.

［33］IBM, ”Tokyo SOC Report 2010 年下期.” https://www-304.ibm.com/connections/blogs/tokyo-soc/, 2010.

［34］IBM, ”Tokyo SOC Report 2013 年上期.” https://www-304.ibm.com/connections/blogs/tokyo-soc/,2010.

［35］F-Secure, ”Windows Remote Desktop Worm ”Morto” Spreading - F-Secure Weblog : News from the Lab.” http://www.f-secure.com/weblog/archives/00002227.html, last visited 2014/4/8.

［36］Vizvary Martin, Jan Vykopal, ”Flow-based detection of RDP brute-force attacks.” 7th International Conference on Security and Protection of Information (SPI 2013), 2013.

［37］Security Affairs, http://securityaffairs.co/wordpress/26247/cyber-crime/kaspersky-lab-revealsincrease-rdp-bruteforce-attacks.html, ”Kaspersky Lab reveals an increase in RDP bruteforce attacks - Security Affairs.” last visited in 2014/7/9.

［38］Alert Logic, ”CLOUD SECURITY REPORT- SPRING 2014.” pp.3-pp.7, 2014.

［39］http://krebsonsecurity.com/2013/12/hacked-via-rdp-really-dumb-passwords/, ”Hacked Via RDP: Really Dumb Passwords ? Krebs on Security.” last visited in 2015/3/9.

［40］https://blogs.cisco.com/security/talos/sshpsychos, ”Threat Spotlight: SSHPsychos.” last visited 2017/12/14.

［41］https://threatpost.com/group-behind-ssh-brute-force-attacks-slowed-down/112095/, ”Group Behind SSH Brute Force Attacks Slowed Down — Threatpost — The first stop for security news.” last visited 2017/12/14.

［42］Akamai Technologies Inc., ”Case Study: FastDNS Infrastructure battles Xor Botnet.” akamai’s [state of the internet]/security, 2015.

［43］https://www.hackread.com/new-ssh-brute-force-lua-bot-shishiga-detected-in-the-wild/, ”New Linux SSH Brute-force LUA Bot Shishiga Detected in the Wild.” last visited 2018/01/02.

［44］https://www.scmagazineuk.com/rdp-brute-force-attacks-used-to-spread-lockcrypt-ransomware/article/706856/, ”RDP brute force attacks used to spread LockCrypt ransomware.” last visited 2018/01/03.

［45］http://nmap.org/ncrack/,”Ncrack - High-speed network authentication cracker.” last visited in 2015/3/9.

［46］https://www.thc.org/thc-hydra/,”THC-HYDRA - fast and flexible network login hacker.” last visited in 2015/3/9.

［47］http://pk6chat.blogspot.jp/2012/08/brutik-rdp-special-edition-cracked.html, ”Brutik RDP - Special edition [CRACKED] - Dont Mess With Your Best Die Like The Rest.” last visited in 2015/3/9.

［48］http://www.hammerofgod.com/downloads.php,”HoG Downloads.” last visited in 2015/3/9.

［49］Maryam M. Najafabadi, Taghi M. Khoshgoftaar, Clifford Kemp, Naeem Seliya, Richard Zuech, ”Machine Learning for Detecting Brute Force Attacks at the Network Level.” IEEE International Conference on Bioinformatics and Bioengineering (BIBE), pp.379-pp.385, 2014.

［50］P. Garcia-Teodoro, J. Diaz-Verdejo, G. Macia-Fernandez, E. Vazquez, ”Anomaly-based network intrusion detection: Techniques, systems and challenges.” ELSEVIER Computers & Security, Volume 28, Issues 1-2, pp.18-pp.28, 2009.

［51］Mohammad Sazzadul Hoque, Md.Abdul Mukit, Md.Abu Naser Bikas, ”An Implementation of Intrusion Detection System Using Genetic Algorithm.” International Journal of Network Security & Its Applications, Volume 4, Number 2, pp.109-pp.120, 2012.

［52］Laurens Hellemons, Luuk Hendriks, Rick Hofstede, Anna Sperotto, Ramin Sadre, Aiko Pras, ”SSHCure: A Flow-Based SSH Intrusion Detection System.” 6th IFIP WG 6.6 International Conference on Autonomous Infrastructure, Management, and Security(AIMS 2012), pp.86-pp.97, 2012.

［53］Akihiro Satoh, Yutaka Nakamura, Takeshi Ikenaga, ”SSH Dictionary Attack Detection Based on Flow Analysis.” IEEE/IPSJ 12th International Symposium on Applications and the Internet (SAINT2012), pp.51-pp.59, 2012.

［54］G Gu, J Zhang, W Lee, ”BotSniffer: Detecting botnet command and control channels in network traffic.” 15th Annual Network and Distributed System Security Symposium (NDSS2008), 2008.

［55］Choi, Hyunsang, Heejo Lee, and Hyogon Kim. ”BotGAD: detecting botnets by capturing group activities in network traffic.” 4th International ICST Conference on COMmunication System softWAre and middlewaRE. ACM, 2009.

［56］David Zhao, Issa Traore, Bassam Sayed, Wei Lu, Sherif Saad, Ali Ghorbani, Dan Garant, ”Botnet detection based on traffic behavior analysis and flow intervals.” ELSEVIER Computers & Security 39, Part A, pp.2-pp.16, 2013.

［57］竹森敬祐, 三宅優, 中尾康二,“IDS ログ分析支援システムの提案.” 情報処理学会研究報告 2003-CSEC-21, 2003.

［58］仲小路博史、寺田真敏,“周波数分析に基づくインシデント傾向検知手法に関する検討.”情報処理学会研究報告 2005-CSEC-30, 2005.

［59］Cisco Systems, ”Cisco IPS.” http://www.cisco.com/web/JP/product/hs/security/ids4200/index.html

［60］McAfee, ”McAfee Network Security Platform.” http://www.mcafee.com/japan/enterprise/nsp/

［61］Daisuke Inoue, Masashi Eto, Katsunari Yoshioka, Shunsuke Baba, Kazuya Suzuki, Junji Nakazato, Kazuhiro Ohtaka, and Koji Nakao, ”nicter: An Incident Analysis System Toward Binding Network Monitoring with Malware Analysis.” WOMBAT Workshop on Information Security Threats Data Collection and Sharing, pp58-66, 2008.

［62］JPCERT コーディネーションセンター, ”TSUBAME（インターネット定点観測システム）.” http://www.jpcert.or.jp/tsubame/

［63］”警察庁セキュリティポータルサイト@police-インターネット定点観測.” http://www.npa.go.jp/cyberpolice/detect/observation.html

［64］牧田大祐, 吉岡克成, 松本勉, ”DNS ハニーポットによる不正活動観測.” 情報処理学会研究報告 2013-CSEC-62, 2013.

［65］牧田大佑, 吉岡克成, 松本勉, 中里純二, 島村隼平, 井上大介,“DNS アンプ攻撃の早期対策を目的とした DNS ハニーポットとダークネ.ットの突合分析.”暗号と情報セキュリティシンポジウム（SCIS2014）, 2014.

［66］Cliff Changchun Zou, Lixin Gao, Weibo Gong, Don Towsley, ”Monitoring and early warning for internet worms.” ACM Conference on Computer and Communications Security (CCS), pp190-199, 2003.

［67］Vinos Yegneswaran, Paul Barford, Dave Plonka, ”On the Design and Use of Internet Sinks for Network Abuse Monitoring.” 7th International Symposium on Recent Advances in Intrusion Detection（RAID）, 2004.

［68］Michael Bailey, Evan Cooke, Farnam Jahanian, Jose Nazario, David Watson, ”The Internet Motion Sensor: A Distributed Blackhole Monitoring System.” Network and Distributed Security Symposium（NDSS）, 2005.

［69］Chenfeng Vincent Zhou, Christopher Leckie, Shanika Karunasekera, ”A Survey of coodinated attacks and collaborative intrusion detect.” Computer & Security on ELSEVIER, pp124-140, 2010.

［70］武仲正彦, 鳥居悟, 清水聡,“ランダムで低速なポートスキャンの検知についての検討.”コンピュータセキュリティシンポジウム（CSS2012）, 2012.

［71］武仲正彦, 鳥居悟, 古川和快, 清水聡,“ランダムで低速なポートスキャンの検知についての検討 2.”暗号と情報セキュリティシンポジウム（SCIS2013）, 2013.

［72］https://wordpress.org/, ”Blog Tool, Publishing Platform, and CMS - WordPress.” last visited in 2016/11/14.

［73］https://www.joomla.org/, ”Joomla! The CMS Trusted By Millions for their Websites.” last visited in 2016/8/9.

［74］https://lolipop.jp/info/news/4149/, ”第三者によるユーザーサイトの改ざん被害に関するご報告-2013 年 08 月 29 日 10 時 57 分/ 新着情報/ お知らせ- レンタルサーバーならロリポップ！.” last visited 2016/8/1.

［75］http://www.jpcert.or.jp/magazine/acreport-cms.html, ”改ざんの標的となる CMS 内の PHP ファイル (2016-02-25).” last visited 2016/8/1.

［76］http://eromang.zataz.com/2011/08/14/suc027-muieblackcat-setup-php-web-scanner-robot/, ”Muieblackcat setup.php Web Scanner/Robot.” last visited 2017/08/30.

［77］http://www.skepticism.us/2015/05/new-in-your-face-malware-attacks-me-ringing-at-your-dorbell/, “New, in your face, malware attacks me:/Ringing.at.your.dorbell!-Dog Is My Copilot.” last visited 2017/8/23.

［78］Roesch Martin, ”Snort: Lightweight intrusion detection for networks.” Proceedings of LISA, 13rd Systems Administration Conference, Vol. 99, No. 1, 1999.

［79］https://www.snort.org/,”Snort - Network Intrusion Detection & Prevention System.” last visited 2017/4/10.

［80］https://modsecurity.org/,”ModSecurity: Open Source Web Application Firewall.” last visited 2017/4/10.

［81］Apache, http://httpd.apache.org/, ”Welcome! - The Apache HTTP Server Project.”

［82］http://www.baidu.com/,”百度.”

［83］https://www.google.co.jp/,”Google.”

［84］https://support.google.com/webmasters/answer/80553?hl=ja, ”Googlebot かどうかの確認 - Search Console ヘルプ.” last visited 2017/4/10.

［85］http://www.msn.com/ja-jp/,”MSN Japan - Hotmail, Outlook.com, Skype, OneDrive, Bing.”

［86］https://www.exploit-db.com/exploits/12133/,”Asset Manager 1.0 - Arbitrary File Upload.” last visited 2017/08/31.

［87］https://msdn.microsoft.com/ja-jp/library/dn455106(v=vs.85). aspx,”IE11 での Web サイト用カスタム タイルの作成 (Windows).” last visited 2017/4/19.

［88］IPA, ”コンピュータウィルス・不正アクセスの届け出状況［2012 年 6 月分］.” http://www.ipa.go.jp/security/txt/2012/07outline.html, 2012.

［89］SUCRI Blog, ”Mass WordPress Brute Force Attacks?? Myth or Reality.” http://blog.sucuri.net/2013/04/mass-wordpress-brute-force-attacks-myth-or-reality.html,2013.

［90］SUCRI Blog, ”The WordPress Brute Force Attack Timeline.” http://blog.sucuri.net/2013/04/the-wordpress-brute-force-attack-timeline.html,2013.

［91］PCWorld, ”GitHub bans weak passwords after brute-force attack results in compromised accounts.” http://www.pcworld.com/article/2065340/ github-bans-weak-passwords-after-bruteforce-attack -results-in-compromised-accounts.html, 2013.

［92］US-CERT, ”Risks of Default Passwords on the Internet.” http://www.us-cert.gov/ncas/alerts/TA13-175A, 2013.

［93］Ars Technica, ”Mass-login attack on Nintendo fan site hijacks 24,000 account.” http://arstechnica.com/security/2013/07/mass-login-attack-on-nintendo-fan-site-hijacks-24000-accounts/, 2013.

［94］Kazuhisa Makino, Takeaki Uno, ”New Algorithms for Enumerating All Maximal Cliques”, 9th Scandinavian Workshop on Algorithm Theory, 2004.

［95］Wu Jiang, Vangala Sarma, Gao Lixin, ”An Effective Architecture and Algorithm for Detecting Worms with Various Scan.” 14th Annual Network and Distributed System Security Symposium (NDSS2004), 2004.

研究業績リスト

論文

1. 本多聡美, 海野由紀, 丸橋弘治, 武仲正彦, 鳥居悟, ”拠点横断分析による IP 使い捨て型ブルートフォース攻撃の検知とその抽出手法.” 情報処理学会論文誌 Vol56 No.3 pp.911-pp.920, Mar. 2015.
2. Satomi Saito, Koji Maruhashi, Masahiko Takenaka, Satoru Torii, ”TOPASE: Detection and Prevention of Brute Force Attacks used Disciplined IPs from IDS Logs.” Journal of Information Processing Vol.24 No.2 pp.217-pp.226, Mar. 2016.
3. 齊藤聡美，吉岡克成，松本勉, ”多数の Web サイトを対象とした攻撃の共起性に基づく攻撃アクセス検知手法とその評価.” 情報処理学会論文誌 Vol59, No.2, pp.1-p.17, Feb. 2018.

国際会議発表

1. Satomi Honda, Yuki Unno, Koji Maruhashi, Masahiko Takenaka, Satoru Torii, “Detection of Novel-Type Brute Force Attacks used Ephemeral Springboard IPs as Camouflage,” International Conference on Information and Network Security(ICINS2014), 2014.
2. Satomi Honda, Yuki Unno, Koji Maruhashi, Masahiko Takenaka, Satoru Torii, ”Poster Abstract: CITRIN: Extracting Adversaries Strategies Hidden in a Large-Scale Event Log,” International Symposium on Research in Attacks, Intrusions and Defenses(RAID2014), 2014.(Poster)
3. Satomi Honda, Yuki Unno, Koji Maruhashi, Masahiko Takenaka, Satoru Torii, “TOPASE: Detection of Brute Force Attacks used Disciplined IPs from IDS Log,” 1ST IEEE/IFIP Workshop on Security for Emerging Distributed Network Technologies (DISSECT), 2015. (Short paper)
4. Satomi Saito, Satoru Torii, Katsunari Yoshioka, Tsutomu Matsumoto, ”Wamber: Defending Web Sites on Hosting Services with Self-Learning Honeypots.” The 11th Asia Joint Conference on Information Security(AsiaJCIS), 2016.

国内学会発表

1. 本多聡美, 海野由紀, 丸橋弘治, 武仲正彦, 鳥居悟, “使い捨て IP による新型ブルートフォース攻撃の検出,” コンピュータセキュリティシンポジウム (CSS2013), 2013.
2. 本多聡美, 海野由紀, 丸橋弘治, 武仲正彦, 鳥居悟, “使い捨て IP によるブルートフォース攻撃検出手法の評価,” 暗号と情報セキュリティシンポジウム (SCIS2014), 2014.

3. 本多聡美, 海野由紀, 丸橋弘治, 武仲正彦, 鳥居悟,“RDP サービスへの分散型ブルートフォース攻撃”, マルチメディア，分散，協調とモバイルシンポジウム (DICOMO2014), 2014.

4. 本多聡美, 海野由紀, 丸橋弘治, 武仲正彦, 鳥居悟,“ネットワークサービスに対するブルートフォース攻撃の傾向比較,”コンピュータセキュリティシンポジウム (CSS2014), 2014.

5. 齊藤聡美, 武仲正彦, 鳥居悟,“SSH ログインセンサによる STBF(Brute Force attacks with Single Trials) の観測.”コンピュータセキュリティシンポジウム (CSS2015), 2015.

6. 齊藤聡美，吉岡克成，松本勉,”多数の Web サイトを対象とした攻撃の共起性に基づく検知手法.” コンピュータセキュリティシンポジウム (CSS2016), 2016.

7. 齊藤聡美, 武仲正彦, 鳥居悟,”新しいタイプの分散型 SSH ログインブルートフォース攻撃 STBF の攻撃元数の推定.” コンピュータセキュリティシンポジウム (CSS2016), 2016.

8. 丸橋弘治, 齊藤聡美, 鳥居悟, 武仲正彦, ”外れ構造検知：大規模離散値データの圧縮による集団アノマリの発見.” データ工学と情報マネジメントに関するフォーラム (DEIM2017), 2017.

9. 西野琢也，菊地亮太，丸橋弘治，福田大輔，齊藤聡美，鳥居悟，伊豆哲也，”テンソル分解に基づくグラフ分類による組織内ネットワーク攻撃活動検知技術．”コンピュータセキュリティシンポジウム (CSS2017), 2017.

受賞

1. Excellent paper awarded (対象論文: Satomi Honda, Yuki Unno, Koji Maruhashi, Masahiko Takenaka, Satoru Torii,“Detection of Novel-Type Brute Force Attacks used Ephemeral Springboard IPs as Camouflage,”International Conference on Information and Network Security(ICINS2014), 2014.)

2. 優秀プレゼンテーション賞・優秀論文賞受賞 (対象論文: 本多聡美, 海野由紀, 丸橋弘治, 武仲正彦, 鳥居悟,“RDP サービスへの分散型ブルートフォース攻撃”, マルチメディア，分散，協調とモバイルシンポジウム (DICOMO2014), 2014.)

3. Best poster award (対象論文: Satomi Honda, Yuki Unno, Koji Maruhashi, Masahiko Takenaka, Satoru Torii,“TOPASE: Detection of Brute Force Attacks used Disciplined IPs from IDS Log,” 1ST IEEE/IFIP Workshop on Security for Emerging Distributed Network Technologies (DISSECT), 2015.(short paper))

付録A　第3章において正解データとして抽出できたIPアドレスに関する統計

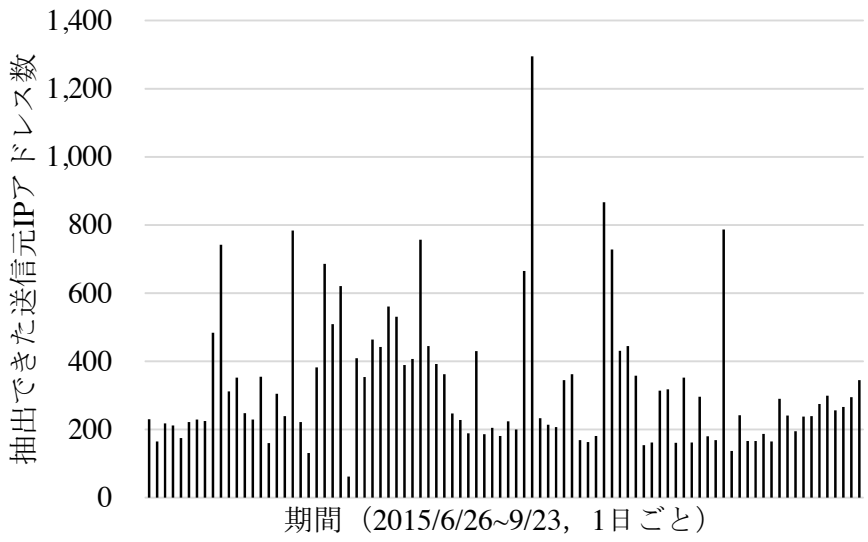

図 A.1: (a) の場合に期間ごとに正解データと判断した送信元 IP アドレス

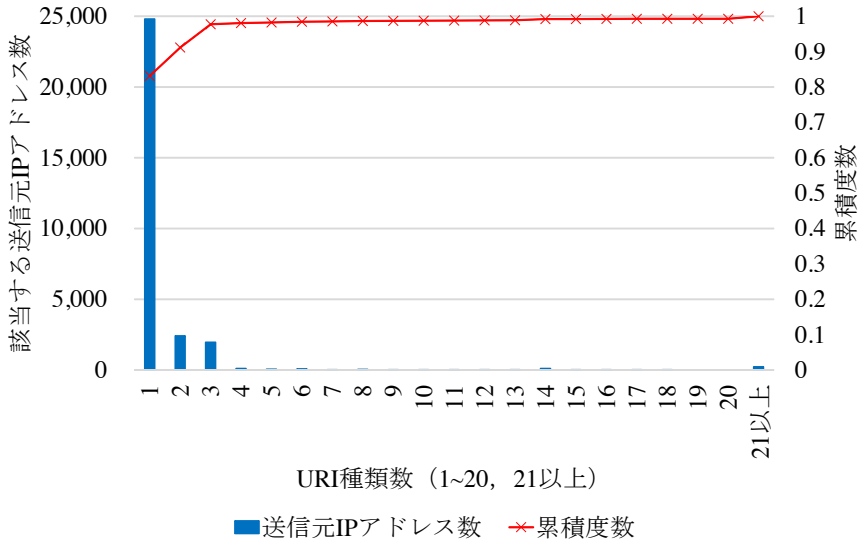

図 A.2: (a) の場合に正解データと判断した送信元 IP アドレスの URI 種類数の分布

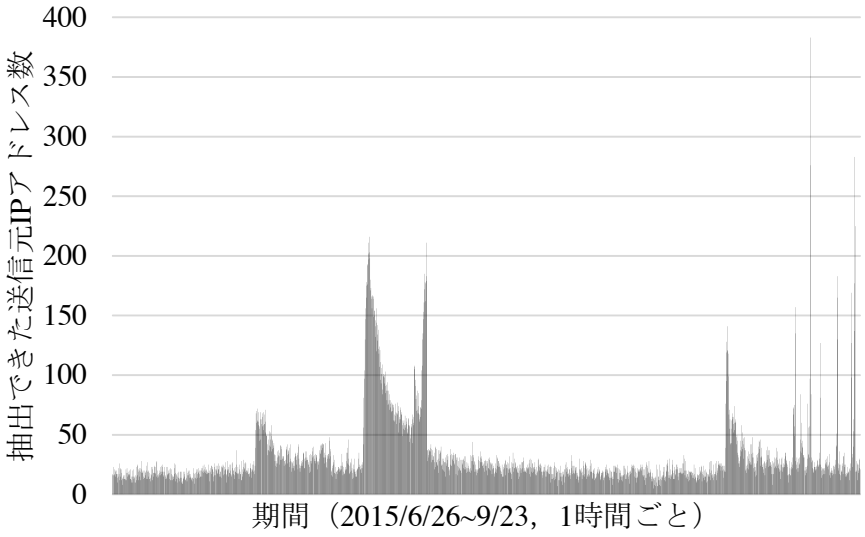

図 A.3: (b) の場合に期間ごとに正解データと判断した送信元 IP アドレス

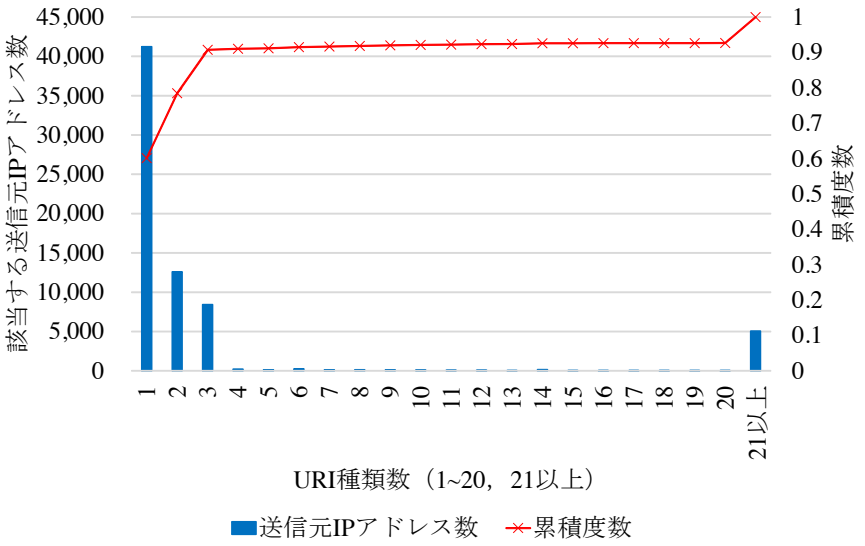

図 A.4: (b) の場合に正解データと判断した送信元 IP アドレスの URI 種類数の分布

付録B　第3章おいて提案手法の適用結果抽出できたIPアドレスに関する統計

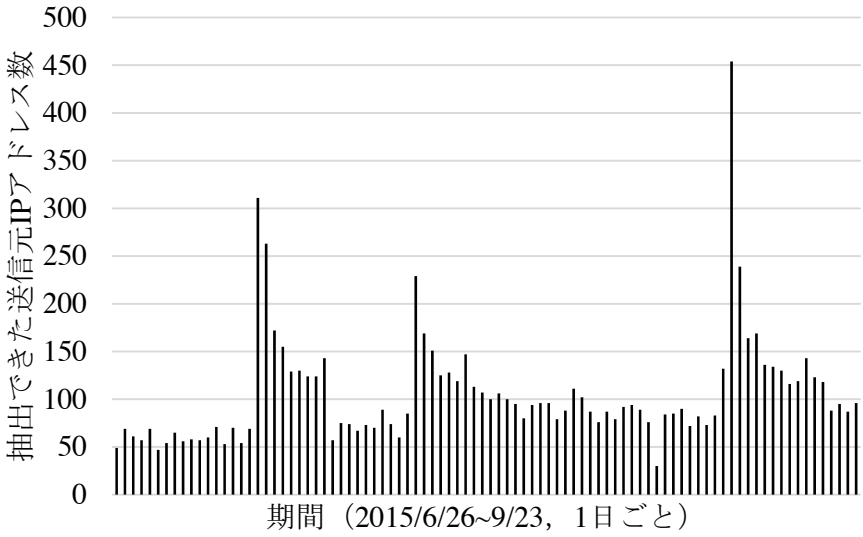

図 B.1: (a) の場合に期間ごとに提案手法が抽出した送信元 IP アドレス

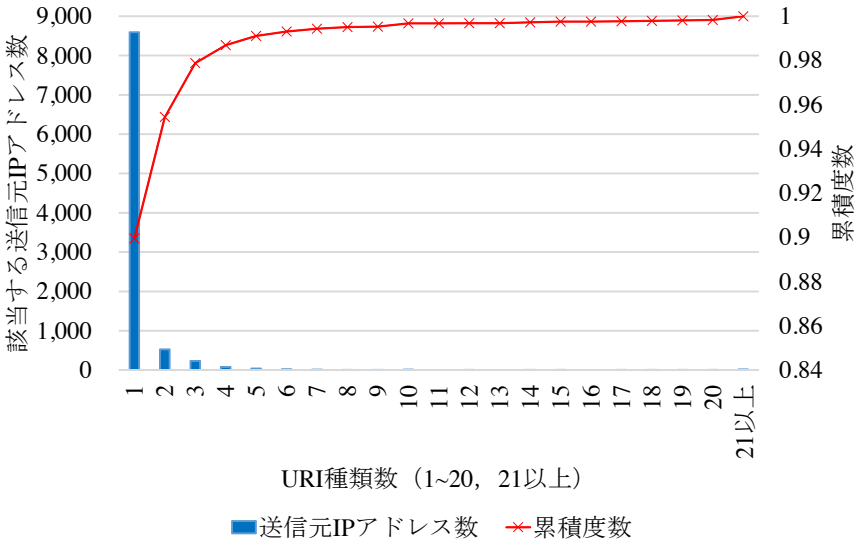

図 B.2: (a) の場合に提案手法が抽出した送信元 IP アドレスの URI 種類数の分布

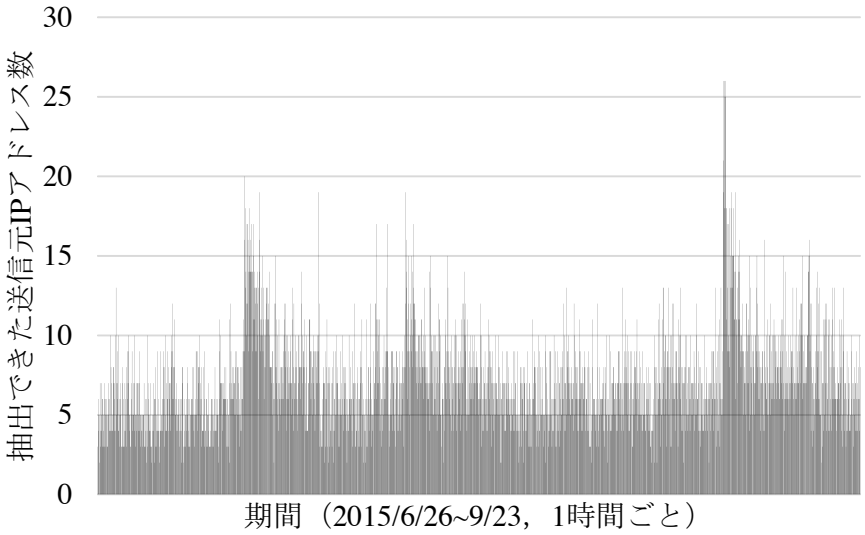

図 B.3: (b) の場合に期間ごとに提案手法が抽出した送信元 IP アドレス

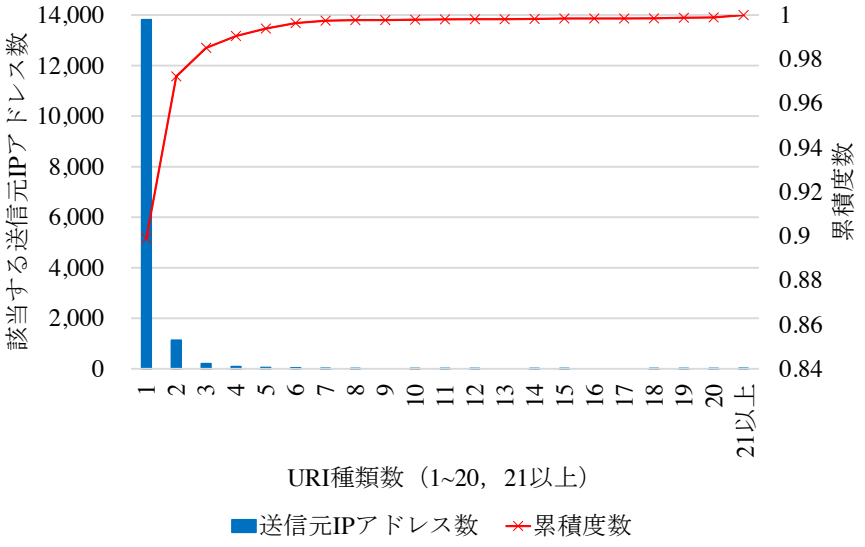

図 B.4: (b) の場合に提案手法が抽出した送信元 IP アドレスの URI 種類数の分布

付録C　第3章において評価に用いた独自シグネチャの一覧

第 3 章では Snort および modsec により Web サイトへの攻撃に関する正解データの作成を試みたが，Web サイトへの攻撃は多岐にわたるため，これらの既存ツールにも見逃しが予想される．そこでこれらのツールに加えて，セキュリティサイト等の情報を元に攻撃パターンを正規表現として手動で書き下した独自シグネチャを作成した．

表 C.1: 独自シグネチャの一覧

カテゴリ	正規表現
CMS	.*wp-login.php.*
	.*wp-content.*.php
	.*wp-admin.*.php
	.*xmlrpc.php
	.*administrator.*joomla.*
	.*assetmanager.aspx
	.*soapCaller.bs
	.*open=1&arrs1[]=99&arrs1[]=102&arrs1[]=103&arrs1[]=95&arrs1[].*
FCKeditor	.*fckeditor.*.+.html
	.*fckeditor.*.+.php
	.*fckeditor.*.+.aspx
	.*ckeditor.*.+.php
	.*editor.*.+.html
	.*editor.*.+.php
	.*editor.*.+.aspx
	.*fckeditor.*$
	.*fckeditor.*
	.*editor.*$
	.*editor.*
不正中継ホストの探索	^CONNECT.+:25$
	^CONNECT.+:80$
	^CONNECT.+:443$
	.+:http:.+
	.+:https:.+
IoT 機器の存在確認	.*:rom-0$
	.*:tmUnblock.cgi$
	.*:onvifsnapshot$
	.*:rtpd.cgi
Struts	.*login.action$
	.*LoginPage.do$
MongoDB	.*moadmin.php$
SQL injection を含む	.*%20union%20select%20.*
	.*unionselect.*from.*
PHP-CGI 攻撃	.*:cgi-binphp
改ざんサイトの存在確認	.*:nyet.+$
ShellShock 脆弱性	.*:Ringing.at.your.dorbell!$