Vol.2018-CSEC-82 No.41
Vol.2018-SPT-29 No.41
2018/7/26

サイバー攻撃の有無が不明な通信データから サイバー攻撃目的と推定される通信を抽出する手法の提案

鮫島 礼佳[1]　芦野 佑樹[1]　須尭 一志[1]　矢野 由紀子[1]　中村 康弘[2]

概要：近年，サイバー攻撃の手法は複雑化しており，それに伴い様々な対策技術が検討されている．近年の主なサイバー攻撃対策技術として，過去に観測されたサイバー攻撃の分析結果に基づいて，対策を講じる技術がある．しかし，この技術では，過去に観測されていない攻撃や，観測されていても攻撃と識別されなかった攻撃は，分析されないため，今後も攻撃と識別できず，対策を講じることが困難である，という課題がある．この課題を解決するためには，未だサイバー攻撃と識別されていない，攻撃の有無が不明な過去の通信データから，サイバー攻撃目的と思われる通信を見つけ出し，対策を講じられるような分析を行うことが必要である．

そこで，本論文では，筆者らが観測した，サイバー攻撃の有無が不明な過去の通信データから，サイバー攻撃目的と思われる通信を見つけ出すことを目的とした，通信データの分析手法について述べる．併せて，分析の結果見つけることができた，サイバー攻撃目的と思われる通信について考察し，報告する．

キーワード：通信データ分析，センサー，サイバー攻撃

Proposal of Extracting Method for Communication that Aim at Cyber Attack from Communication Data Unknown whether Cyber Attack Exists

Samejima Ayaka[1]　Ashino Yuki[1]　Sugyo Kazushi[1]　Yano Yukiko[1]　Nakamura Yasuhiro[2]

Abstract: In recent years, method of cyber attack to be complicated and many measures are considered. Main measure based on analysis result of observed cyber attack. But, that technology cannot take a measure to be not observed method of cyber attack.

We propose analysis method of communication data to detect communication that aim at cyber attack.

Keywords: Communication Data Analysis, Sensor, Cyber Attack

1. はじめに

近年，インターネットを経由したサイバー攻撃は日々増加しており，その被害と危険性は社会問題となっている [1]．近年の主なサイバー攻撃対策技術として，過去に観測されたサイバー攻撃で用いられたプログラムや，そのプログラムが発生させる通信などを分析し，その分析結果に基づいて，サイバー攻撃の検知やフィルタリングといった対策を講じる技術がある．このような技術によって，分析済みのサイバー攻撃と同様の攻撃手法，もしくは亜種と呼ばれる類似した攻撃手法が用いられたサイバー攻撃は，初期段階のうちに対策を講じることが可能となっている．しかし，この技術では，過去に一度も用いられていない攻撃手法や，過去に観測されなかった攻撃手法，過去に観測されていても攻撃と識別されなかった攻撃手法などを用いた

[1] 日本電気株式会社　ナショナルセキュリティ・ソリューション事業部　サイバーセキュリティ・ファクトリ
Cyber Security Factory, National Security Solution Division, NEC Corporation
[2] 防衛大学校　電気情報学群　情報工学科
Department of Computer Science, School of Electrical and Computer Engineering, National Defense Academy

Vol.2018-CSEC-82 No.41
Vol.2018-SPT-29 No.41
2018/7/26

サイバー攻撃は分析されないため，対策を講じることが困難であるといわれている [2].

実際に，対策が困難であったサイバー攻撃の被害事例として，攻撃を受けてから対策が行われるまでに一年以上の時間を要した事例がある [3]．この事例では，被害端末はマルウェアに感染し，一年以上にわたってユーザが意図していない不審な通信を外部と行っていた．被害端末には，アンチウイルスソフトが導入されていたにも関わらず，このマルウェアは検知されていなかった．この事例から，サイバー空間には，過去にサイバー攻撃と識別されなかったため，一度も分析されておらず，対策が講じられないまま，未だに見逃され続けているサイバー攻撃が存在すると考えられる．

このような，分析されていないサイバー攻撃への対策が困難であるという，従来のサイバー攻撃対策技術の課題を解決するために，筆者らは，サイバー攻撃の有無が不明な過去の通信データの中から，未だにサイバー攻撃と識別されていないが，サイバー攻撃目的と思われる通信を見つけ出す手法を提案する必要があると考えた．

インターネットを経由して行われるサイバー攻撃は，通信プロトコルの性質上必ず通信が発生する [4]．通信には，必ず送信者と受信者がおり，送信者は自身の意図に基づいて伝達したい情報を作成し，その情報を信号に変換することで，受信者に伝達する [5]．つまり，インターネット通信において，送信者と受信者の間で伝達されている信号である通信データは，送信者が伝達したいと意図する情報を含んでいると考えられる．そこで，筆者らは，通信データを分析することで，この送信者の意図を推測することができれば，その意図の中に，サイバー攻撃目的と思われる通信を見つけ出すことができるのではないかと考えた．

以上のことから，本論文では，サイバー攻撃の有無が不明な過去の通信データから，送信者の意図を推測することができるような，通信データの分析手法を提案する．併せて，分析結果から，送信者の意図を推測し，そこからサイバー攻撃目的と思われる通信を見つ出すことができる可能性について考察する．

本論文の構成は以下のようになっている．2 章では，情報伝達モデルから送信者の意図について説明し，さらに，通信データの分析に関連する研究から，本論文の位置付けを述べる．3 章では，送信者の意図を推測することを目的とした，通信データの分析手法を提案し，4 章で提案した分析手法の評価を行う．5 章では，分析結果から，送信者の意図を推測し，サイバー攻撃目的と思われる通信を見つけることの可能性について考察を述べ，6 章で本論文全体についてまとめる．

2. 関連研究と本研究の位置付け

1 章で述べたように，本論文では，サイバー攻撃の有無が不明な過去の通信データから，送信者の意図を推測することができるような，通信データの分析手法を提案することで，サイバー攻撃目的と思われる通信を見つ出すことを目的としている．

そこで，本章では，まず 2.1 節で本論文で扱う送信者の意図について説明し，次に 2.2 節で通信データの分析に関連する研究について述べ，最後に 2.3 節でこの研究分野における本論文の位置付けについて述べる．

2.1 送信者の意図

本節では，本論文で扱う送信者の意図について，情報伝達モデルを用いて説明する．

インターネットを経由して行われるサイバー攻撃は，通信プロトコルの性質上必ず通信が発生する [4]．サイバー攻撃に限らず，通信には，必ず送信者と受信者がおり，その間で情報が伝達されている．この送信者と受信者の間で行われる情報の伝達は，モデル化されており，そのモデルの一つに，シャノン・ウィーバーモデルという伝達モデルがある (**図 1**)[5].

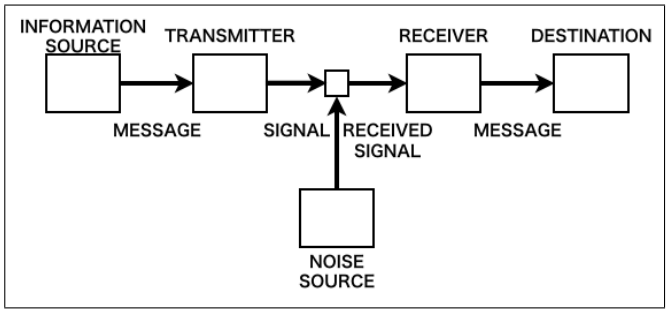

図 1 シャノン・ウィーバーによる情報伝達モデル
Fig. 1 Communication Model from Shannon and Weaver

シャノン・ウィーバーモデルは，電話やラジオなどの通信において，情報を早く正確に伝達することを目的とした研究で発明されたモデルである．シャノンらによると，情報の送信者 (Information Source) は，自身の意図に基づいて伝達したい情報 (Message) を作成し，その意図を送信器 (Transmitter) で信号 (Signal) にエンコードすることで送信する．送信された信号 (Signal) は，搬送中に電波干渉など様々な理由によって発生するノイズ (Noise Source) の影響を受け，受信信号 (Received Signal) となって受信器 (Reciever) で受信される．受信された信号 (Received Signal) は，受信器 (Receiver) によってデコードされ，受信者 (Destnation) によって解釈されることで情報 (Message) となる．このようにして，送信者の意図に基づく情報は，受信者に伝達されている．

本論文では，サイバー攻撃目的と思われる通信を見つ出すことを最終目的としているため，本論文で扱う通信とは，通信プロトコルに則ったインターネット通信のことを指す．

Vol.2018-CSEC-82 No.41
Vol.2018-SPT-29 No.41
2018/7/26

シャノン・ウィーバーモデルをインターネット通信に適用すると，情報の送信者は，自身の意図に基づく情報を，送信器である計算機に入力し，情報を信号にエンコードしているといえる．ここで，送信器である計算機に自身の意図に基づく情報を入力するためには，送信者の意図通りの挙動をするプログラムの存在が必要不可欠である．つまり，計算機上で動くプログラムの挙動は，送信者の意図を表しているといえる．受信者が，受信信号のみから，送信者の意図を推測することは困難であると考える．しかし，送信者の意図によって実行されているプログラムが，どのような挙動をするプログラムなのか，受信信号から読み取ることは，可能であると考える．

例えば，ポートスキャンの場合，送信者の意図として，攻撃を意図して攻撃可能な開いているポートを調査しているのか，防御を意図してペネトレーションテストのように設定上意図せず開いてしまっているポートがないか調査しているのか，学術的な統計調査を意図してサイバー空間で開いている可能性の高いポートを集計しているのか，自身が作成したポートスキャンプログラムの動作確認を意図しているのか，などと様々な意図を推測することが可能である．しかし，推測した意図の中に，送信者の意図が本当に存在するのか，推測した以外の意図が存在するのか，受信した通信データのみの情報から読み取ることは非常に困難であるため，送信者の意図は推測の域をでない．一方，送信者が自身の意図を伝達するために使用しているプログラム，つまり，受信者の計算機の複数のポートに順番にパケットを送信する，というポートスキャンプログラムの挙動は，受信した通信データのみの情報から読み取ることが可能である．

そこで，本論文では，送信者が使用しているプログラムの挙動を読み取ることで，送信者の意図を可能な限り推測し，その中からサイバー攻撃目的と思われる通信を見つけ出すことの可能性について考察していく．

2.2 関連研究

サイバー攻撃の有無が不明な過去の通信データから，送信者の意図を推測することができるような，通信データの分析手法を提案するにあたって，以下の 3 つの観点から関連研究を調査する必要があると考えた．

(a) 通信データにサイバー攻撃の特徴が表れる可能性があること．

(b) 通信データからプログラムの挙動を読み取れる可能性があること．

(c) サイバー攻撃の有無が不明な過去の通信データから，サイバー攻撃と思われる通信を見つけ出せる可能性があること．

本節では，(a)，(b)，(c)3 つの観点から調査した関連研究について以下に述べる．

(a) 通信データからサイバー攻撃を分類，識別している研究

ここでは，マルウェア実行中に発生した攻撃通信データを分析し，その通信の特徴を抽出することで，攻撃の種類の分類や，攻撃を識別する手法を提案している研究について述べる．

桑原ら [6] は，ハニーポットで観測した通信データから特徴量を抽出することで，感染しているマルウェアの種類や，感染の可否を判別する手法を提案している．この研究では，ペイロードに出現する文字列の種類や，パケット数等を特徴量とすることで，感染種類の判定を行っている．

田中ら [7] は，マルウェア実行時の通信データを分析することで，通信データから悪性通信を特定できる通信の特徴を検討している．この研究では，定期的に発生するパケットや，パケットサイズ，接続先 IP アドレスの数等を特徴とすることで，正常通信との差異を表している．

水野ら [8] は，マルウェア実行時の通信データのうち，HTTP パケットのヘッダに含まれる情報に機械学習を適用することで，悪性通信と良性通信を判別する手法を提案している．

(b) 通信データをアプリケーション (プログラム) ごとに分類している研究

ここでは，アプリケーション実行時の通信データから，アプリケーションごと，つまり送信者が使用しているプログラムごとに，通信データを分類する手法を提案している研究について述べる．

佐藤ら [9] は，アプリケーションの通信に利用されているパケットの観測順に，ペイロード長の遷移を表現することで，通信データからアプリケーションを識別する手法を提案している．

葛野ら [10] は，Android アプリケーションが行う通信のうち，HTTP 通信に着目し，通信データからアプリケーションを分類する手法を提案している．この研究では，アプリケーションが行う通信のうち，HTTP パケットのヘッダに含まれる，送信先情報と送信内容情報を利用することで，アプリケーション毎の通信を分類している．また，分類したアプリケーションが，ユーザの意図していない端末上の情報を外部に送信しているか否かの判別も行っている．

グェンら [11] は，SSH 暗号化トンネル経由で MMS，HTTP，HTTPS，POP，FTP，SMTP 通信を行った通信データを独自に作成し，その通信の送受信パケット長とパケットの時間感覚を特徴とし，機械学習を適用することで，フロー分割の困難な暗号化トンネル内通信から，アプリケーションを識別する手法を提案している．

(c) サイバー攻撃の有無が不明な過去の通信データから，サイバー攻撃と思われる通信を見つけ出している研究

ここでは，各研究の研究者らが，独自のセンサーで長期間観測したサイバー攻撃の有無が不明な通信データから，

サイバー攻撃と思われる通信を見つけるための分析手法を提案している研究について述べる．

芦野ら [12], [13] は，サイバー空間上に設置したセンサーで観測した通信データを用いて，サイバー攻撃の初期段階の活動と推定される通信の発信源の分類や，初期活動で使用されていると推定されるプログラムを分類する手法を提案している．この研究では，サイバー攻撃には初期段階に情報収集を行う偵察活動が存在し，その偵察活動を捉えることのサイバー攻撃対策における重要性を述べている．

梶川ら [14], [15] は，サイバー空間上に設置したセンサーで観測した通信データを用いて，分散型走査活動とその発信源の検出，攻撃ペイロードを分類する手法を提案している．この研究では，分散型走査活動から，攻撃者の意図や，攻撃グループの推定が行える可能性についても考察している．

2.3 本論文の位置付け

2.2 節で述べた関連研究から，本論文の位置付けを述べる．

(a) の研究により，サイバー攻撃であると識別済みの通信データを分析することは，分析済みのサイバー攻撃と同様の攻撃手法，もしくは亜種と呼ばれる類似した攻撃手法が用いられたサイバー攻撃の分類，識別に有効であることが示されている．このことから，通信データには，サイバー攻撃の特徴を表す情報が含まれているといえる．

しかし，これらの研究では，過去に一度も用いられていない攻撃手法や，過去に観測されなかった攻撃手法，過去に観測されていても攻撃と識別されなかった攻撃手法などを用いたサイバー攻撃は分析されないため，分類や識別が困難であることも示されている．したがって，サイバー攻撃であると識別済みの通信データを分析することは，本論文の目的である，サイバー攻撃の有無が不明な過去の通信データから，サイバー攻撃目的と思われる通信を見つ出すためには，不適当な分析手法であると考える．

(b) の研究により，アプリケーション実行時の通信データを分析することで，通信データをアプリケーションごとに分類できることが示されている．このことから，通信データにはアプリケーション，つまり，送信者が使用しているプログラムを識別できる情報が含まれているといえる．したがって，通信データを分析することで，送信者が自身の意図を伝達するために使用しているプログラムの挙動を読み取ることが，可能であると考える．

(c) の研究により，サイバー空間上に設置したセンサーで観測した，サイバー攻撃の有無が不明な通信データには，サイバー攻撃，もしくはサイバー攻撃の初期活動と思われる通信が，存在していることが示されている．したがって，サイバー攻撃の有無が不明な通信データを分析することで，サイバー攻撃目的と思われる通信を見つ出せる可能性があると考える．

以上のことから，本論文では，サイバー攻撃であると識別済みの通信だけに限らず，サイバー空間上に設置したセンサーで観測した，サイバー攻撃の有無が不明な通信データから，送信者が使用しているプログラムの挙動を読み取れるような，通信データの分析手法を提案する．また，分析結果から，送信者の意図を可能な限り推測し，その中からサイバー攻撃目的と思われる通信を見つけ出すことの可能性について考察する．

3. 提案手法

本章では，サイバー攻撃の有無が不明な過去の通信データから，送信者が使用しているプログラムの挙動を読み取れるような，通信データの分析手法を提案する．3.1 節で，通信データの分類手法を提案するにあたって，本論文でとったアプローチについて述べ，3.2 節で，提案手法の手順について述べる．

3.1 アプローチ

1 章で述べた通り，近年では，過去に観測されたサイバー攻撃の分析結果に基づいて，対策を講じる技術が主流となっている．しかし，この技術では，分析されていないサイバー攻撃への対策が困難であるという，課題がある．この課題を解決するために，本論文では以下のアプローチをとる．

- サイバー攻撃の有無が不明な過去の通信データを扱う．
- 通信データから，送信者が自身の意図を伝達するために使用しているプログラムの挙動を読み取ることを目的とした分析を行う．
- 分析結果からプログラムの挙動を読み取り，送信者の意図を推測する．
- 推測した送信者の意図から，サイバー攻撃目的と思われる通信を見つけ出すことの可能性について考察する．

3.2 提案手法の手順

本節では，3.1 節で述べたアプローチに則り，サイバー攻撃の有無が不明な過去の通信データから，送信者が使用しているプログラムの挙動を読み取れるような，通信データの分析手順について述べる．

(1) 通信データから，プログラムの挙動を読み取るために有効な特徴を選定し，抽出する

(2) 抽出した特徴を用いて通信データを分析し，プログラムの挙動を読み取れる特徴を持つ通信データを，全体のサイバー攻撃の有無が不明な過去の通信データの中から再度抽出する．

(3) 抽出した通信データを詳細分析することで，送信者の意図を考察する．

この提案手法の手順を，以下の**図 2** に示す．この提案手

Vol.2018-CSEC-82 No.41
Vol.2018-SPT-29 No.41
2018/7/26

法によって，サイバー攻撃の有無が不明な過去の通信データから，送信者が使用しているプログラムの挙動を読み取ることが可能となる．

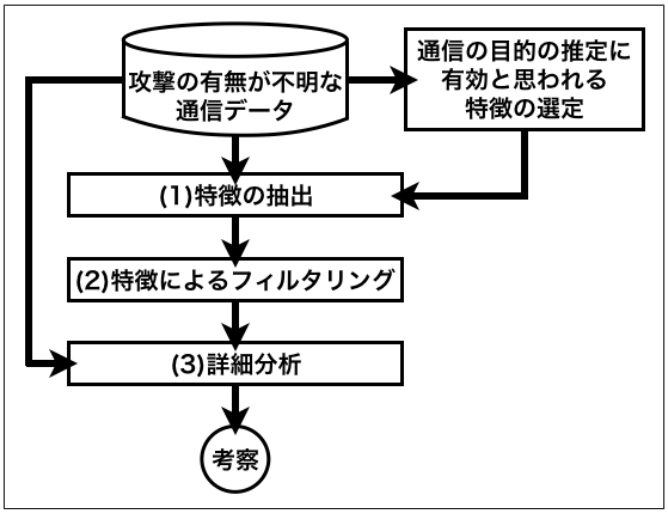

図 2 提案手法の手順
Fig. 2 Procedure of Propoal Technique

4. 評価実験

本章では，筆者らがサイバー空間上に設置したセンサーで観測した，サイバー攻撃の有無が不明な過去の通信データを用いて，3 章で述べた提案手法の評価を行う．

4.1 使用データ

評価実験に使用した通信データの概要を，以下の**表 1** に示す．このデータは，筆者らがサイバー空間上に設置した，約 1500 個の IP アドレスを観測しているセンサーを用いて収集したものである．センサーは，SYN フラグの立った TCP/IP パケットを受信した際に，SYN フラグと ACK フラグを立てた TCP/IP パケットと，RST フラグと ACK フラグを立てた TCP/IP パケットを返送するように設定されている．

表 1 使用データの概要
Table 1 About Using Data

観測期間	2016/10/20〜2016/10/29
データサイズ	約 186GB
TCP/IP パケット数	約 9.66 億パケット
UDP/IP パケット数	約 298 万パケット
合計パケット数	約 9.72 億パケット

4.2 通信データの分類に用いる特徴の選定

本節では，通信データから送信者が使用しているプログラムの挙動を読み取る際に有効な，特徴の選定について述べる．

4.2.1 IP アドレス

まず，送信元 IP アドレスについて，近年のサイバー攻撃対策の活動により，ボットネットやマルウェア配布サーバなど，攻撃の発信元になっていると知られている IP アドレスがブラックリスト化されて，対策に活用されている [16]．これは，過去の分析により，既にサイバー攻撃と識別できている攻撃に対する対策としては有効であると言える．しかし，本論文では 1 章で述べた通り，攻撃の有無が不明な通信データの中からサイバー攻撃と推定される通信を抽出することを目的としている．また，送信元 IP アドレスは，環境次第で自由に変えることができるサービスも存在する [17]．そのため，送信元 IP アドレスをサイバー攻撃と推定される通信の抽出の特徴として利用するのは困難であると考える．

次に，宛先 IP アドレスについて，本論文では宛先となるセンサーのアドレスは約 1500 個となっている．宛先 IP は，何箇所のセンサーで観測されているか，どのような順番で遷移しているか，などの情報から，サイバー攻撃の規模や目的の推定に活用できる可能性がある．

4.2.2 ポート番号

まず，送信元ポート番号について，送信元ポート番号は，送信側で任意に変更可能であったり，セッションを貼り直すたびに遷移していくという性質がある．そのため，送信元ポート番号は，サイバー攻撃と推定される通信の抽出の特徴として利用するのは困難であると考える．

次に，宛先ポート番号について，TCP プロトコルの性質上、特定のサービスで利用するために予約されている well-known ポートが存在する．また，送受信 IP アドレスの同じ，同一セッション内の宛先 IP アドレス数の情報は，送信元が接続したいポートの範囲の推測や，走査活動の検知 [15], **?**に活用利用できる可能性がある．そのため，宛先ポート番号の情報は，送信元が利用したいプロトコルかた通信の目的を推測するのに活用できる可能性がある．

4.2.3 プロトコルとフラグ

プロトコルについて，4.1 節で述べたように，本論文で利用しているセンサーは，TCP/IP の SYN フラグが立ったパケットに対して，SYN フラグと ACK フラグを立てたパケットを返送している．これにより，TCP の 3Way-HandShake のコネクションが確立する場合がある．そこで，本論文では OSI 参照モデル [18] のネットワーク層以上のプロトコル，つまり ARP などを除いた，ICMP，TCP/IP，UDP/IP プロトコルを対象として分析を行っていく．

TCP/IP の場合，プロトコルの性質上，SYN や ACK といったフラグが存在する．そこで，本論文では，TCP/IP を用いた通信の場合は，同一セッション内で受信した TCP/IP パケットの，プラグの順番も特徴になりうるのではないかと考えた．本論文では，この同一セッション内で TCP/IP パケットのフラグの受信した順番を，フラグパターンと

呼ぶ.

4.2.4 パケットの受信時間間隔

受信したパケットの受信時間間隔について，送信元で意図的にパケットの送信時間間隔を操作している場合，センサー側で受信したパケットの受信時間間隔から通信の目的の推定に活用できると考えた．そこで，実際に同一送受信 IP アドレスの通信において，受信時間間隔を測定してみたところ，**表 2** のようになった.

表 2 同一送受信 IP アドレス，送受信ポート番号におけるパケットの受信時間間隔

Table 2 Receipt Time Interval of Packet

プロトコル	フラグ	受信時間間隔 A	受信時間間隔 B
TCP/IP	ACK	約 0.2442 秒	約 0.5559 秒
TCP/IP	FIN/ACK	約 4.053e-06 秒	約 2.981e-03 秒
TCP/IP	FIN/ACK	約 0.7534 秒	約 1.857 秒

表 2 において，A は 2016/10/20 に，B は 2016/10/29 に観測された通信である．この結果により，同一送受信 IP アドレス，同一プロトコル，同一フラグパターンの通信であっても，観測日によって，約 2 倍以上も変動することがわかった．この変動の原因として，経路情報が更新され，通信経路の変化が発生したり，経路上のネットワークのどこかで遅延が発生していたり，という可能性が考えられる．また，同一の目的を持った，同一のプログラムによって発生した通信であったとしても，そのプログラムを実行している送信元の設備の処理能力や OS などの環境によって，送信されるパケットの時間間隔に影響を及ぼす可能性も考えられる．したがって，パケットの受信時間間隔は，サイバー攻撃と推定される通信の抽出の特徴として利用するのは困難であると考える.

4.2.5 有効な特徴

以上の検討により，本論文では，宛先 IP アドレス，宛先ポート番号，プロトコル，TCP/IP パケットのフラグの受信した順番 (フラグパターン) を，通信データの分類における特徴と定める.

4.3 評価実験の手順

本節では，実際に 4.1 節で述べたデータを分析した手順の一例について述べる.

(1) 特徴の抽出

4.1 節で述べたデータの各パケットのヘッダから，受信時刻，送受信 IP アドレス，送受信ポート番号，プロトコル，フラグの情報を抽出した.

抽出した情報を元に，各パケットの情報をセッションごとにまとめた．ここで，セッションとは，SYN フラグが立った TCP/IP パケットを受信した時刻から 30 秒間の間に，同一の送受信 IP アドレス，同一送受信ポート番号で行われる通信のことを指す.

作成したセッションごとに，4.2.5 節で選定した 4 つの特徴 (宛先 IP アドレス，宛先ポート番号，プロトコル，フラグパターンを抽出した.

(2) 抽出した特徴を用いたフィルタリング

(1) で抽出した 4 つの特徴を用いて，4.1 節で述べた通信データを分析した.

ここでは，分析に利用した特徴の一例として，同一の送信元 IP アドレスから，複数の宛先 IP アドレスの同一宛先ポートに向けて，複数のフラグパターンを送信している，送信元 IP アドレスの通信を抽出した.

(3) 詳細分析

(2) で抽出した通信データの詳細分析を行った．この際，(1) で抽出した特徴だけではなく，分析対象となる分類のパケットを，改めて 4.1 節で述べた pcap から抽出することで，(1) では扱わなかった，ペイロードの情報やセッション前後の通信なども含めて，詳細に分析を行った.

詳細分析の結果から，通信の目的やサイバー攻撃目的と推定される通信について，考察を行う．考察の内容は，5 章で述べる.

4.4 評価実験の結果

本節では，4.3 節で述べた実験手順に則って，実験を行った結果について述べる.

4.4.1 セッションごとに応答を変化させる通信

4.1 節で述べた通り，本実験で使用しているセンサーは，SYN フラグの立った TCP/IP パケットを受信した際は，SYN フラグと ACK フラグを立てた TCP/IP パケットと，RST フラグと ACK フラグを立てた TCP/IP パケットを返送している．このセンサーの返送に対し，自身の送信元ポートを変えつつ，複数回 SYN を送信してセッションを張り直し，その度に応答を変化させている送信元が見つかった．応答の内容は 6 種類あり，TCP/IP の再送を除いて，1 パケットしか応答しないものから最大 3 パケット送ってくる場合もあった．この応答内容の種類を**表 3** にまとめる.

表 3 同一送信元による応答の種類

Table 3 Kinds of Response from Common Source

1 パケット目	2 パケット目	3 パケット目
応答無し	ー	ー
RST	ー	ー
FIN/ACK	ー	ー
ACK	FIN/ACK(再送あり)	ー
ACK	FIN/ACK(再送あり)	RST
ACK	FIN/ACK(再送あり)	RST/CWR

4.4.2 特定のフラグパターンの通信

4.1 節で用いたデータを，フラグパターンごとに，送信元 IP アドレスの種類数を算出したところ，一つの送信元 IP アドレスからしか観測できないフラグパターンが存在した．そのフラグパターンの送信元は，SYN を送信し，センサーが SYN/ACK と RST/ACK を返送したのち，ACK，PSH/ACK(再送あり)，FIN/ACK，FIN/RST/ACK の順に TCP/IP パケットを送信するセッションを，複数回行なっていた．セッションごとに，送信元ポートは変化していたが，宛先ポートは 23 か 2323 のどちらかに絞られたいた．

5. 考察

本章では，4.4 節で述べた評価実験の結果について 5.1 節で考察し，**??**節で今後について述べる．

5.1 評価実験の結果に関する考察

5.1.1 セッションごとに応答を変化させる通信の考察

本節では，4.4.1 節で述べた，評価実験の結果について考察する．

4.1 節で述べた通り，筆者らが設置しているセンサーは，SYN フラグの立った TCP/IP パケットを受信した際は，SYN フラグと ACK フラグを立てた TCP/IP パケットと，RST フラグと ACK フラグを立てた TCP/IP パケットを返送している．センサーが，毎回同じパターンのパケットを返送しているにも関わらず，この送信元は，セッションの度に応答が異なっている．このことから，この送信元で実行されているプログラムは，受信者になんらかの応答を期待して，数種類のパターンの応答を試している可能性が考えられる．仮に，送信者の期待する応答を，センサーで返送することができていたら，今回観測されたパターンとは別のフラグパターンが観測できた可能性があると考えられる．

5.1.2 特定のフラグパターンの通信の考察

本節では，4.4.2 節で述べた，評価実験の結果について考察する．

4.4.2 節で述べたフラグパターンの通信は，4.1 節で述べたデータの中にはこの一つしか含まれていなかった．この通信は，TCP/IP のコネクションを確立しようと ACK を送信し，センサーが RST/ACK を返送しているにも関わらず，PSH/ACK を送信してきていたことから，仮にセンサーとコネクションが確立していたら，何らかのデータを送信しようとしていた可能性が考えられる．ここで，もう一つの特徴として，宛先ポートに着目した．この通信では，9 割ほどの通信が 23 番のポート宛にパケットを送っていたが，残りの 1 割ほどのパケットは，2323 番ポート宛だった．そこで，このような特徴を持つ通信について調べた見た所，ちょうどこの通信が存在してた時期と同時期に，IoT 機器用マルウェアである”Mirai”による不正アクセスが増加していたと，報告されていた [19]．このレポートによると，「「23/TCP」に対するスキャンは「2323/TCP」に対するスキャンの 9 倍多く行われる」と報告されていた．したがって，今回の評価実験で見つけた通信は，この Mirai の活動の一部であったと考えられる．このことから，今回の実験で用いたフラグパターンは，サイバー攻撃目的と思われる通信を見つけ出すのに有効であった可能性が高いといえる．

6. まとめ

本論文では，分析されていないサイバー攻撃への対策が困難であるという，近年のサイバー攻撃対策技術の課題を解決するために，サイバー攻撃の有無が不明な過去の通信データから，サイバー攻撃目的と思われる通信を見つ出すことの必要性を述べた．そして，サイバー攻撃の有無が不明な過去の通信データから，サイバー攻撃目的と思われる通信を見つ出すことを目標とし，送信者の意図を推測することができるような，通信データの分析手法を提案した．併せて，提案した分析手法を用いて，筆者らがサイバー空間上に設置したセンサーで観測したサイバー攻撃の有無が不明な過去の通信データを分析することで，サイバー攻撃目的と思われる通信を見つ出し，提案手法の有効性を示した．　今後は，送信者の意図から，サイバー攻撃における攻撃者の攻撃目的を推測できるような分析手法に発展させていく．

参考文献

[1] JPCERT/CC：インシデント報告対応レポート (オンライン)，入手先 ⟨http://www.jpcert.or.jp/ir/report.html⟩，(参照 2018-06-24)．
[2] JPCERT/CC：JPCERT/CC の活動とサイバー攻撃解析協議会への期待 (オンライン)，入手先 ⟨http://www.meti.go.jp/committee/kenkyukai/shoujo/cyber_attack/pdf/001_08_00.pdf⟩，(参照 2018-06-24)．
[3] 情報処理推進機構：サイバーレスキュー隊 (J-CERT) 分析レポート 2016 長期間線の実態　1 台の感染 PC に残された攻撃痕跡の分析 (オンライン)，入手先 ⟨https://www.ipa.go.jp/files/000057175.pdf⟩，(参照 2018-06-24)．
[4] JPNIC：IP(Internet Protocol) とは (オンライン)，入手先 ⟨https://www.nic.ad.jp/ja/basics/beginners/ip.html⟩，(参照 2018-06-24)．
[5] Claude E. Shannon，Warren Weaver：*THE MATHEMATICAL THEORY OF COMMUNICATION*，The University of Illinois(1949)．
[6] 桑原和也，菊池浩明，寺田真敏，藤原将志：パケットキャプチャーから感染種類を判定する発見的手法について，コンピュータセキュリティシンポジウム 2009(2009)．
[7] 田中恭之，畑田充弘，稲積孝紀：パケットキャプチャから見た悪性通信に関する特徴の考察，コンピュータセキュリティシンポジウム 2013(2013)
[8] 水野翔，畑田充弘，森達哉，後藤滋樹：マルウェア感染ホストが生成する通信の弁別手法，電子情報通信学会信学技報 ICSS2015-66(2016)．

[9] 佐藤剛，和泉勇治，田中和之：ペイロード長遷移パターンの順序性評価によるネットワークアプリケーション識別，電子情報通信学会信学技報 IA2010-23(2010)．
[10] 葛野弘樹：HTTP トラフィックを利用したクラスタリングによる Android アプリケーションの分類，情報処理学会研究報告 Vol.2012-CSEC-56 No.19(2012)．
[11] グェン タイン トゥアン，中村康弘：Multi-class SVM による暗号化トンネル内通信のアプリケーション識別，電子情報通信学会論文誌 学生論文特集 Vol.J97-D No.3 pp.488-495(2014)．
[12] 芦野佑樹，山根匡人，矢野由紀子，島成佳：長期間に渡るインターネットノイズの観測に基づいたサイバー攻撃の初期活動と推定される通信の発信源を分類する手法の提案，情報処理学会研究報告 (2017)．
[13] 芦野佑樹，中村康弘，矢野由紀子，島成佳：サイバー攻撃の初期段階と推定される活動で使用されるプログラムの分類手法の提案と評価，コンピュータセキュリティシンポジウム 2017(2017)．
[14] 梶川慶太，中村康弘：分散型走査グループの検知と攻撃ペイロードの分類，第 16 回情報科学技術フォーラム (2017)．
[15] 梶川慶太，中村康弘：宛先変化数に着目した分散走査活動の検出，電子情報通信学会総合大会 2018(2018) ．
[16] 一般財団法人インターネット協会：ホワイトリスト、ブラックリストって何ですか?(オンライン)，入手先 ⟨http://salt.iajapan.org/wpmu/anti_spam/universal/measure/whitelist-blacklist/⟩，（参照 2018-06-24）．
[17] Amazon：Amazon EC2 インスタンスの IP アドレッシング (オンライン)，入手先 ⟨https://docs.aws.amazon.com/ja_jp/AWSEC2/latest/UserGuide/using-instance-addressing.html⟩，（参照 2018-06-24）．
[18] JPNIC：OSI 参照モデルとは (オンライン)，入手先 ⟨https://www.nic.ad.jp/ja/basics/terms/osi.html⟩，（参照 2018-06-24）．
[19] Trend Micro：「Mirai」に感染した IoT 機器による不正アクセスが増加　警察庁調べ (オンライン)，入手先 ⟨https://www.trendmicro.com/jp/iot-security/news/42⟩，（参照 2018-06-24）．

2018年 第82回CSEC研究会投稿論文　正誤表

1. 訂正する論文

第82回コンピュータセキュリティ・第29回セキュリティ心理学とトラスト合同研究発表会
2018/07/26(木)午前 C-4:CSEC(4)09:30 - 11:10 発表
タイトル：「サイバー攻撃の有無が不明な通信データからサイバー攻撃目的と推定される通信を抽出する手法の提案」
著者：鮫島礼佳・芦野佑樹・須尭一志・矢野由紀子(日本電気株式会社ナショナルセキュリティ・ ソリューション事業部サイバーセキュリティ・ファクトリ),
中村康弘(防衛大学校電気情報学群情報工学科)

2. 訂正箇所

p7 4.4.2 特定のフラグパターンの通信

[誤]	4.1 節で用いたデータを，フラグパターンごとに，送信元IPアドレスの種類数を算出したところ，一つの送信元IPアドレスからしか観測できないフラグパターンが存在した．そのフラグパターンの送信元は，SYNを送信し，センサーがSYN/ACKとRST/ACKを返送したのち，ACK, PSH/ACK(再送あり)，FIN/ACK，FIN/RST/ACKの順 にTCP/IPパケットを送信するセッションを，複数回行なっていた．セッションごとに，送信元ポートは変化していたが，宛先ポートは23か2323のどちらかに絞られたいた．
[正]	4.1 節で用いたデータを，フラグパターンごとに集計したのち，送信元 IP アドレスごとに宛先ポート番号の種類とパケット数を算出したところ，同じフラグパターンで発生する通信の中で宛先ポート番号に偏りのある送信元IPアドレスの集団が存在した．その送信元 は，SYN を送信し，センサーがSYN/ACKとRST/ACK を返送したのち，ACK，FIN/ACKの順にTCP/IPパケットを送信しており，宛先ポート番号は23番と2323番のみとなっていた．宛先ポート番号ごとのパケット数の内訳は，どの送信元も23番宛が2323番宛の9~11倍のパケット数となっていた．

p7 5.2 特定のフラグパターンの通信の考察

[誤]	4.4.2節で述べたフラグパターンの通信は，4.1節で述べたデータの中にはこの一つしか含まれていなかった．この通信は，TCP/IPのコネクションを確立しようとACKを送信し，センサーがRST/ACKを返送しているにも関わらず，PSH/ACKを送信してきていたことから，仮にセンサーとコネクションが確立していたら，何らかのデータを送信しようとしていた可能性が考えられる．ここで，もう一つの特徴として，宛先ポートに着目した．この通信では，9割ほどの通信が23番のポート宛にパケットを送っていたが，残りの1割ほどのパケットは，2323番ポート宛だった．
[正]	4.4.2 節で述べた特徴をもつ通信は，複数の送信元IPアドレスが存在した．4.4.2節で述べたように，この通信では，どの送信元も23番宛が2323番宛の9~11倍のパケット数となっていた．