サイバー攻撃のアトリビューションのための統計学的アプローチ

三村　守

サイバー攻撃における攻撃者の帰属（アトリビューション）は非常に関心の高いテーマであり，その仕組みを理解することは重要である．本稿では，サイバー攻撃において攻撃者のアトリビューションを推定する手法の概要について解説する．まず，サイバー攻撃の種類と仕組みについて簡単に解説し，アトリビューションに関する基本的な情報について説明する．次に，標的型攻撃において統計学的アプローチにより攻撃者のアトリビューションを推定する手法について説明し，多変量解析を用いた具体例を示す．最後に，まとめと機械学習を用いた今後の展望について述べる．なお，この原稿は文献 [1] および文献 [2] を参考として編集し，加筆修正したものである．

キーワード：サイバー攻撃，アトリビューション，多変量解析

1. はじめに

IoT 機器や機械学習の普及に伴い，サイバースペースの重要性はますます増大している．サイバースペースでは，「匿名性が高い」「証拠の改ざんが可能」「地理・時間的な制約がない」「被害が不特定多数に拡大しやすい」などの特性があるとされている．特に，相手が誰であるのかわからない匿名性については，サイバー戦を非対称戦と位置づける重要な特性である．2015 年に生起した年金機構などを狙った大規模なサイバー攻撃では，さまざまなレポートや分析結果が公開されているものの，未だに攻撃者の正体は判明していない [3]．その一方で，2013 年 2 月には，米国に対する一連のサイバー攻撃が人民解放軍の第 61398 部隊の仕業であり，その拠点は上海にあるビルであるとのレポートが公開された [4]．さらに，2014 年 5 月には米司法省が，サイバー攻撃の実行犯として，人民解放軍の将校 5 人を告発している [5]．2018 年には，米大統領選介入に関連する一連のハッキングなどの罪で，ロシア連邦軍参謀本部情報総局のサイバー部隊に所属するとされる 12 名が起訴されている [6]．さらに同年 8 月には，2017 年に大流行したランサムウェアの WannaCry の作成などに関与したとして，北朝鮮の諜報機関である朝鮮人民軍偵察総局に所属する男性が起訴されている [7]．サイバー攻撃において攻撃者の正体が判明した場合には，国家間の関係に重大な影響を与える可能性がある．サイバー攻撃における攻撃者の帰属（アトリビューション）は非常に関心の高いテーマであり，その仕組みを理解することは重要である．そこで本稿では，サイバー攻撃において攻撃者のアトリビューションを推定する手法の概要について解説する．以下，2 節ではサイバー攻撃の種類と仕組みについて簡単に解説する．3 節では，アトリビューションに関する基本的な情報について説明する．4 節では標的型攻撃などにおいて攻撃者のアトリビューションを推定する手法について説明し，5 節では統計学的アプローチである多変量解析を用いた分析の具体例を示す．最後に，まとめと機械学習を用いた今後の展望について述べる．

みむら　まもる
防衛大学校情報工学科
〒 239–8686 神奈川県横須賀市走水 1–10–20

2. サイバー攻撃の概要

サイバー攻撃は，攻撃者が主体的に任意のタイミングで実施できる「能動的攻撃」と，被攻撃者による何らかの動作を必要とする「受動的攻撃」に分類できる．能動的攻撃は，主にインターネットに公開されたサーバを対象としたサイバー攻撃であり，サーバに大量のアクセスを発生させて閲覧を困難にするサービス拒否攻撃や，ホームページのコンテンツの改ざんがこれに該当する．受動的攻撃は，主に利用者の端末を対象としたサイバー攻撃であり，マルウェアを添付したメールを送りつけて利用者の端末の制御を奪う標的型攻撃や，不正なコンテンツを埋め込んだホームページを閲覧させて利用者の端末を攻撃し，最終的にその制御を奪うドライブ・バイ・ダウンロード攻撃がこれに該当する．社会的に重大な影響を及ぼすサイバー攻撃のほとんどは，利用者の端末を対象とした受動的攻撃である．特に標的型攻撃は，機密情報の搾取を目的として隠密に行われることが多く，入念な準備を必要とし，組

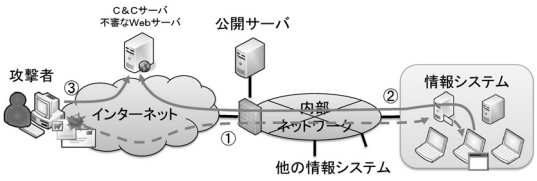

図 1　標的型攻撃の概要

織に深刻な影響を与えることも珍しくない．その中の一部は非常に高度，あるいは執拗であることからAPT (Advanced Persistent Threat) と呼ばれ，大規模な組織や国家の関与が疑われているものもある [8]．そこで本稿では，標的型攻撃における攻撃者のアトリビューションを推定する手法について解説する．

図 1 に標的型攻撃の概要を示す．まず，攻撃者はマルウェアや不審な Web サイトへのリンクをメールで送付する（図 1 ①）．メールの添付ファイルを実行した端末，あるいはリンクをクリックして不審な Web サーバに誘導された端末はマルウェアに感染し，C&C サーバとの通信を開始する（図 1 ②）．攻撃者は C&C サーバを経由して制御を奪った端末を遠隔操作し，機密情報などを搾取する（図 1 ③）．受動的攻撃の攻撃者を特定するための手がかりは，メールの送付に利用されたメールサーバや端末の IP アドレス，端末の遠隔操作に利用された C&C サーバの IP アドレスなどがある．メールサーバや端末の IP アドレスは，端末を調査して発見したメールのヘッダを確認するか，あるいはその利用者のメールサーバのログから調査することが可能である．C&C サーバの IP アドレスなどは，端末やメールサーバから発見したメールの添付ファイルやリンクを専門家が解析することで得られる．この解析では，ほかにもマルウェアの特徴などの攻撃者に関係するさまざまな情報も得られる．端末がマルウェアに感染してしまった場合には，遠隔操作の痕跡が残るため，ファイアウォールやプロキシサーバ のログからも C&C サーバの IP アドレスなどを得ることが可能となる．

3.　アトリビューションに関する基本的な情報

標的型攻撃に用いられた IP アドレスなどの攻撃者のアトリビューションに関する情報は，インディケータと呼ばれており，サイバーインテリジェンスのコミュニティでは情報共有が進んでいる．具体的には，情報処理推進機構 (IPA) が運営するサイバー情報共有イニシアティブ (J-CSIP) や，警察庁を中心としたサイバーインテリジェンス情報共有ネットワークがある．また，CrowdStrike のようにサイバー攻撃のアトリビューションに関するインディケータや，アトリビューションについて分析したレポートを提供するサービスも増えてきている．この節では，アトリビューションに関する基本的な情報であるインディケータについて説明する．

表 1　主な通信先に関するインディケータ

インディケータ	概要
IP アドレス	例：117.103.185.21
FQDN	例：www.mod.go.jp
ドメイン名	例：mod.go.jp
IP アドレスの利用者	IP アドレスの利用者
ドメインのレジストラント	ドメインの利用者
ドメインのレジストラ	ドメインの登録業者

3.1　通信先に関する情報

サイバースペースにおいて攻撃者を特定するための最も基本的なインディケータは，インターネット上の論理的位置を示す IP アドレスなどである．サイバー攻撃に用いられた IP アドレスなどは，Web サーバ，ファイアウォール，プロキシサーバ，メールサーバなどのログ，発見した添付ファイルの解析結果から得ることができる．IP アドレスなどからさらに，インターネットで公開されているオープン情報を活用し，さまざまな通信先に関する情報を調査することが可能である．たとえば，IP アドレスやドメイン名は国ごとの割り当てが決まっているため，どの地域からの攻撃であるかのおおよその目安をつけることができる．これらの最も基本的なインディケータは，あくまでも「サイバー攻撃の最終的な経由地」を示すものであり，「サイバー攻撃の本来の攻撃元」を示しているとは限らない．

主な通信先に関するインディケータを表 1 に示す．IP アドレスと FQDN (Fully Qualified Domain Name) およびドメイン名は，DNS によって相互に変換することが可能である．IP アドレスと FQDN の対応関係は，1 対 1 となるわけではない．たとえば，複数の FQDN が同じ IP アドレスに対応する場合もあれば，その逆の場合もある．サイバー攻撃の場合には，一般に C&C サーバの IP アドレスを，複数の FQDN やドメイン名を使って使い回す傾向が認められる．また，DNS を利用せず，直接 IP アドレスを利用して C&C サーバとの通信を試みるマルウェアもある．IP アドレスの利用者，ドメインのレジストラントおよびレジストラは，Whois と呼ばれるサービスを利用することで，誰でも容易に調査することが可能である．これらのインディケータには，登録者の名称，住所，電話番号，メール

表 2　主な攻撃手段に関するインディケータ

インディケータ	概要
マルウェアのハッシュ	ファイルの同一性
マルウェアの動作	作成ファイル名や挙動
ツールの種類	ツールやパッケージの名称
メールの内容	メールの件名や送信者
タイムスタンプ	マルウェアなどの作成時刻

アドレスなどが含まれていることがある．これらのインディケータは，より攻撃者との関連が強いものと考えられる．

3.2　攻撃手段に関する情報

攻撃者のアトリビューションに関する別のアプローチとしては，使用されたマルウェア，メールなどの攻撃手段に注目する手法が考えられる．たとえば，メールに添付されたマルウェアが同じ，あるいはメールの件名や本文が同じ内容であれば，それらの攻撃は同じ攻撃者による攻撃である可能性が高いと考えられる．

主な攻撃手段に関するインディケータを表 2 に示す．最も基本的な攻撃手段に関するインディケータは，使用されたマルウェアのハッシュ値である．サイバー攻撃に用いられるマルウェアのほとんどはファイルである．ファイルは同じ名称かつ同じサイズであっても，その中身は異なっている可能性がある．一般にファイルの同一性を確認するためには，ハッシュ値が活用されている．マルウェアがコンピュータの内部で作成するファイル名や挙動も，攻撃手段に関する重要なインディケータである．また，マルウェアはビルダーと呼ばれるパッケージを用いて作成されることも多いため，マルウェアやツールのパッケージの名称もその攻撃の特徴となる．これらのマルウェアに関するインディケータは，専門家の解析によって得ることができる．マルウェアに関するインディケータは，偽装するためにはある程度のスキルを要するため，比較的に信頼性の高いインディケータであると考えられる．ただし，広く出回っており，誰もが容易に利用できるような手段については，攻撃者の特徴とはならないこともある．マルウェアなどが添付されたメールの件名，送信者，添付ファイルの名称なども，攻撃手段に関する重要なインディケータである．これらのインディケータは，メールのヘッダかメールサーバのログから得ることができる．メールの件名，送信者，添付ファイルの名称などは，攻撃者が任意に指定することが可能であるため，その特徴が現れる可能性があるが，偽装も容易である点には注意が必要である．

表 3　主な攻撃者の振舞いに関するインディケータ

インディケータ	概要
目的	対象，搾取された情報など
使用言語	メールやマルウェアの使用言語
活動地域	時刻などから推定される活動地域
能力	手段から推定される能力

3.3　攻撃者の振舞いに関する情報

これまでに示したインディケータは，ログ，マルウェアの解析結果などから直接的に得られるインディケータか，あるいはそれらのインディケータに何らかの機械的な処理を実施することで得られるインディケータであった．サイバー・キル・チェーンという概念を提唱したロッキード・マーティン社のレポートによると，インディケータは「単一で分離不能なもの」「加工により得られたもの」および「振舞いに関するもの」の 3 種類に整理されている [9]．この概念によると，これまでに示したインディケータは，「単一で分離不能なもの」あるいは「加工により得られたもの」となる．次に説明するのは，プロファイリングにより導出する攻撃者の振舞いに関するインディケータである．

主な攻撃者の振舞いに関するインディケータを表 3 に示す．攻撃者の目的は，サイバー攻撃の対象となった組織や搾取された情報など，「攻撃者はサイバー攻撃によって何を得たのか？」あるいは「サイバー攻撃で利益を得たのは誰か？」といった事項に着目することで推定することができる．メールやマルウェアの作成に使用された言語も有益なインディケータである．英語，スペイン語などの広い地域で使用されている言語を除外すると，攻撃者が使用する言語はその活動している地域にも関係する．もちろん使用する言語は偽装することが可能であるが，スペルや文法の誤り，ネイティブは使用しない表現，機械翻訳を使用した形跡などから，攻撃者の母国語ではないことが浮き彫りとなる場合もある．攻撃者が活動している地域は，メールやマルウェアが作成された時刻や時刻帯からも推定することができる．また，攻撃者が使用する言語，攻撃手段などからは，攻撃者の能力を推定することができる．

4.　攻撃者のアトリビューション

4.1　IP アドレスの信頼性

感染した端末の遠隔操作に用いられる C&C サーバ，ドライブ・バイ・ダウンロード攻撃に用いられる Web サーバの FQDN や IP アドレスなどは，ファイアウォール，プロキシサーバなどのログ，発見した不審メール

のリンクおよび添付ファイルの解析結果から得ることができる．これらの FQDN や IP アドレスは，攻撃者が遠隔操作や情報の搾取のために継続して利用する必要があるため，メールの送信経路よりも攻撃者との関連が強いものと考えられている．そのため，C&C サーバの FQDN や IP アドレスは，攻撃者を追跡するための最も主要なインディケータであると考えられる．ただし，組織力がある攻撃者は，C&C サーバなどの攻撃インフラを短期間で入れ換える傾向が認められる．近年ではホスティングサービスのクラウド化が進んでおり，攻撃インフラを短期間で再構築することが容易となってきている．そのため，その攻撃が発生した時点での IP アドレスの利用者，ドメインのレジストラント，レジストラなどの通信先に関するインディケータを蓄積しておくことが重要である．過去の履歴を検索するためには，インターネットにおける通信先に関するインディケータを蓄積しており，過去の履歴を検索することができる DomainTools，PassiveTotal などのサービスも活用できる．C&C サーバなどの攻撃インフラを構築するための手法としては，攻撃者が自前で用意する手法，防弾ホスティングサービスを利用する手法，インターネット上の脆弱性があるサーバを乗っ取るか改ざんして悪用する手法などが考えられる．攻撃者が自前で C&C サーバを用意するか防弾ホスティングサービスを利用する場合，その FQDN や IP アドレスは正規のサイトとは異なる見慣れないものとなるため，通信先に関するインディケータを調査することで見分けることができる場合がある．たとえば，海外の見慣れないサイトでそのドメインが新しく登録されたものであり，かつ登録内容に不審な記述がある場合，攻撃者が新たに構築した攻撃インフラである可能性を考慮すべきであろう．これに対し，攻撃者側も不審に思われるのを防ぐため，対象国のドメインや IP アドレスを取得してその国内に攻撃インフラを構築するケースも目立つようになってきている．中には正規の ISP を用いて構築したプロキシサーバもあり，これらがサイバー攻撃や違法なサービスの中継に使用されることもある．そのため，国内の大手 ISP には悪質なプロキシサーバ業者と契約しないように求めるなどの排除に向けた取り組みも始まっている．さらに，攻撃者が正規のサイトを乗っ取るか改ざんし，C&C サーバを構築するケースも増加している．このような場合には，通信先に関するインディケータを調査したとしても，正規のサイトとの区別は非常に困難である．

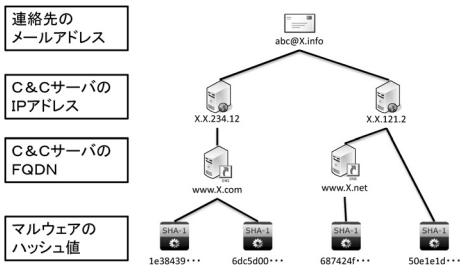

図 2　相互参照の結果を視覚化した例

4.2　インディケータの相関分析

結局のところ，「サイバー攻撃の本来の攻撃元」を追跡するためには，通信経路，サーバ，ドメイン，IP アドレスなどを管理する組織との協力が必要となってくる．これらが国内や同盟国で管理されている場合には，関係機関との協力がある程度可能であるが，国外の場合には追跡が困難であることが多い．そのため，多くの場合に国境を越えた実世界での追跡は現実的ではないのが現状である．そこで，サイバー攻撃のインディケータを集約し，その相関関係を分析することで，攻撃者のプロファイリングを実施する手法が注目されるようになってきた．相関分析の基本は，各インディケータの相互参照（クロス・リファレンス）である．相互参照の結果，ある攻撃とある攻撃のインディケータが一致した場合には，それらの攻撃が同一の攻撃者による可能性が考えられる．

相互参照の結果を視覚化した例を図 2 に示す．この例では，マルウェアのハッシュ値から共通する C&C サーバの FQDN および IP アドレスを参照し，最終的に利用者の登録に用いられたメールアドレスが共通であることが判明している．しかしながら，これまでに示したとおり，インディケータにはさまざまな種類があるため，どのインディケータを重視すべきかを判断することは難しい．重視すべきインディケータを判断するための指標としては，信頼性と関連性が挙げられる．ここで言う信頼性とは，そのインディケータを信頼してよいかを示す指標であり，偽装の困難さを意味する．たとえば，メールの送信者（メールアドレスを含む），件名，本文などは攻撃者が任意に設定することが可能であるため，その信頼性は高いとは言いがたい．これに対し，C&C サーバの FQDN や IP アドレスは，攻撃者が遠隔操作や情報の搾取のために継続して利用する必要があるため，ある程度信頼してよいものと考えられる．もちろん，攻撃者が正規のサイトを乗っ取るか改ざんして構築した C&C サーバはこの限りではない．もう一つの指標である関連性は，そのインディ

ケータと攻撃者の結び付きの強さを示すものである．たとえば，短期的に変化する可能性がある C&C サーバの FQDN や IP アドレスよりは，その登録に利用された情報のほうが攻撃者との関連性は高いものと考えられる．信頼性および関連性の観点から注目されているのは，振舞いに関するインディケータである．振舞いに関するインディケータは，複数のサイバー攻撃の通信先や攻撃手段に関するインディケータを相関分析し，総合的な判断から導き出される高次のインディケータである．この分析は，犯罪心理学におけるプロファイリングに類似している．たとえば，攻撃者の目的はサイバー攻撃の対象となった組織，搾取された情報，サイバー攻撃によりもたらされた被害，組織の利害関係などから推定することが可能である．メールやマルウェアが作成された言語，解析によって得られたタイムスタンプの時刻や時刻帯などからは，攻撃者の使用言語や活動地域が推定できる．マルウェアの種類，使用する脆弱性などからは攻撃者の能力が推定できる．振舞いに関するインディケータを評価するうえで重要となるポイントは，その希少性である．たとえば，マルウェアの開発に使用された言語が希少なものであれば，それは攻撃者の能力を示す重要な特徴となる．修正プログラムが公開されていない未知の脆弱性を用いた攻撃であれば，攻撃者の能力は高く，それなりの組織や資金力を有していることが推定できる．たとえば，スタックスネット[1]のように，複数の未知の脆弱性を用いたマルウェアを作成する能力がある組織は非常に限られている．

4.3 相関分析の限界

通信経路，サーバ，ドメイン，IP アドレスなどの追跡は現実的ではないという現状を考慮すると，攻撃者のプロファイリングは唯一の現実的な選択肢であると考えられる．しかしながら，プロファイリングにより得られた攻撃者の特徴は，あくまでも状況証拠を積み重ねて分析した結果に過ぎないという点には留意する必要がある．したがって，高度な攻撃者であれば，そのプロファイルすら偽装しているという可能性も考えられる．また，作成したプロファイルを最終的にどうやって攻撃者個人に結び付けるかという課題もある．この課題を解決するために注目されているのは，SNS (Social Networking Service) を活用するアプローチである．SNS には，Facebook のように実名での利用を前提としているものがある．また，個人の実生活に関する書き込み，写真やその中に埋め込まれている GPS (Global Positioning System) の位置情報，交友関係など，個人を特定するための情報の宝庫でもある．これらの情報の中には，インターネットから誰でも閲覧可能なものもある．これらの SNS から得られた情報とインディケータの相互参照により，攻撃者と個人が結び付けられる場合がある．たとえば，人民解放軍の第 61486 部隊の仕業とされるサイバー攻撃のレポートでは，ドメインの取得に用いられたメールアドレスに含まれる「CPYY」という文字列を個人のブログと紐づけ，その個人の活動を追跡することで第 61486 部隊の特定に至っている [10]．このように，サイバースペースにおいて得た状況証拠を積み重ねることにより，プロファイリングの精度を高めることは可能である．また，自組織に対するサイバー攻撃のインディケータと，このようなレポートで公開されているインディケータを相互参照することにより，自組織に対するサイバー攻撃のアトリビューションというパズルの最後の 1 ピースを埋めるという手段もありうる．しかしながら，この場合にも第三者によるなりすましであることを主張することで，攻撃者の言い逃れを可能とする余地が残る．

5. 統計学的アプローチの具体例

5.1 前処理

ここまでに説明したさまざまなインディケータを総合的に分析する手法として，統計学的アプローチである多変量解析が挙げられる．しかしながら，分析の対象となる多くのインディケータは名義尺度であり，そのままでは多変量解析を実施することはできない．そこで，インディケータを数値に変換する前処理が必要となる．

前処理の例を図 3 に示す．この例では，名義尺度であるユニークな IP アドレスに 1 から連続する整数を割

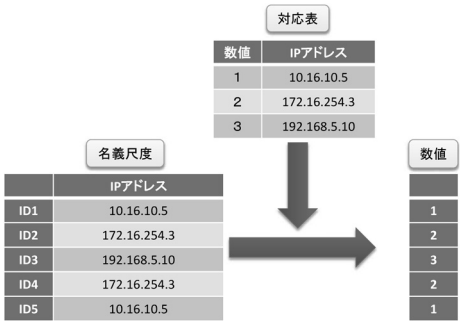

図 3 前処理の例

[1] 米国とイスラエルがイランの核開発を妨害するために開発したとされるマルウェアであり，2010 年に発見された．

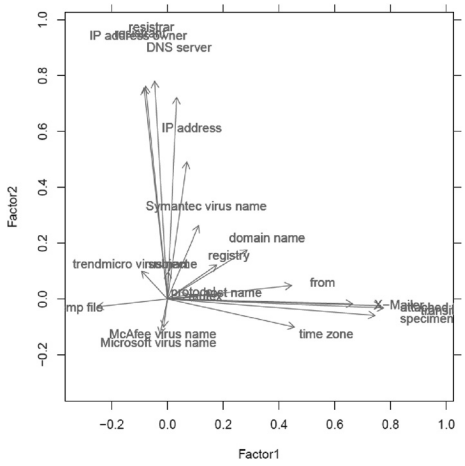

図 4　因子分析の結果の視覚化の例

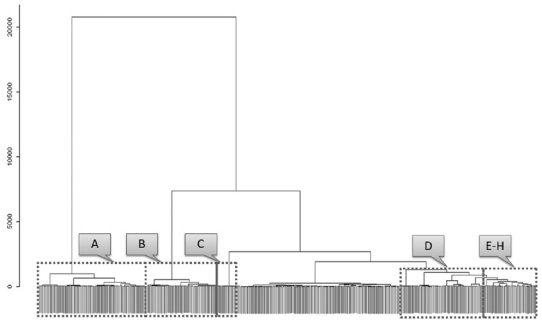

図 5　階層的クラスター分析の例

り当てた対応表を作成し，単純に IP アドレスを対応する整数に変換している．完全一致する IP アドレスは，同じ整数に変換される．このように，名義尺度であるインディケータを何らかの数値に変換することで，多変量解析を適用することが可能となる．この前処理では，単純に元のインディケータの意味を考慮せずに数値に変換している．そのため，各インディケータの相関分析を実施した場合，完全に一致する場合のみ相関があると判定される．しかしながら現実的には，ある二つのインディケータは完全一致していないが，部分的に一致あるいは似ている場合も考えられる．そのため，インディケータの類似性を考慮して数値に変換することで，相関分析の精度を向上できる可能性がある．

5.2　因子分析

次に，数値化した複数のインディケータを使用した因子分析の具体例を示す．因子分析は多変量解析の分析手法の一つであり，数値化した各データの相関からデータの構造を明らかにし，何らかの共通の因子を探すための手法である．ここでは，数値化したサイバー攻撃に関する各インディケータの相関から，その構造を明らかにすることを試みる．

因子分析の結果を視覚化した例を図 4 に示す．この例では横軸を第 1 因子，縦軸を第 2 因子とし，2 次元空間に各インディケータの相関を視覚化している．各矢印は，インディケータの各因子に対する相関（向きと強さ）を示している．第 1 因子に関しては，IP アドレスなどの通信先に関するインディケータとの強い相関が認められる．しかしながら，ドメイン名については通信先に関するインディケータであるにもかかわらず，第 1 因子との相関はそれほど高いとは言えない．この原因の一つとしては，攻撃者がさまざまなドメインを使い回し，IP アドレスとの相関を隠そうとしている可能性が考えられる．このように，因子分析を実施することで，各インディケータの傾向を調査することが可能である．多変量解析においては，分析に必要なパラメータの設定やその結果の解釈は基本的に分析者の裁量に依存する．そのため，結果の解釈についてはさまざまな見方がある点には注意が必要である．

5.3　クラスター分析

次に，標的型攻撃をクラスター分析で，攻撃者ごとに分類することを試みる．クラスター分析も多変量解析の分析手法の一つであり，複数の項目から構成される多数のデータを，互いに項目が類似するデータを同一のクラスターに集約する分類手法である．ここでは数値化した標的型攻撃の各インディケータから，攻撃者との相関が高いと思われるものを抽出し，クラスター分析を実施する．

階層的クラスター分析の例を図 5 に示す．この例では，図中の横軸の一番下の部分が約 500 件の各攻撃を示しており，縦軸は各クラスター間の距離を示している．同一のクラスターについては，トーナメント表のように横線で結ばれており，その横線に対応する縦軸の目盛りがクラスター間の距離を示している．ここで，図中の破線で示す距離で同一のクラスターを解釈すると，A〜H の八つのクラスターに攻撃を分類することができる．これらの攻撃は，各々同一の攻撃者による標的型攻撃である可能性が考えられる．このように，階層的クラスター分析では，距離によって構成されるクラスターの数が異なる．このクラスターの数やその結果の解釈については，やはり分析者の裁量に依存する点には注意が必要である．

6.　おわりに

本稿では，代表的なサイバー攻撃である標的型攻撃

において，攻撃者のアトリビューションを推定する手法について解説した．サイバー攻撃における攻撃者のアトリビューションを調査するためには，通信経路，サーバ，ドメイン，IP アドレスなどを追跡するアプローチは現実的ではない．現状では，サイバー攻撃に関するインディケータを継続的に蓄積し，相互参照によりその相関関係を分析し，攻撃者のプロファイリングを実施することが最良のアプローチであると考えられる．プロファイリングを効果的に実施するためには，C&C サーバの FQDN や IP アドレスなどのインディケータ，サイバー攻撃に用いられたマルウェア，その分析レポートなどの共有が不可欠である．サイバー攻撃を受けた個々の組織がそれらを共有しなければ，各サイバー攻撃の相関関係や全体像を把握することが困難となり，攻撃者のプロファイリングは不可能となってしまう．しかしながら，サイバー攻撃に関する情報の共有には慎重な姿勢の組織が多いのが現状である．その主な原因は，自組織がサイバー攻撃を受けたことを他者に知られるのではないかという疑念であると考えられる．そのため，サイバー攻撃に関する情報の共有は難しい．サイバー攻撃の攻撃者は，このような被害者側のジレンマを巧みに利用している．これに対抗するためには，一部の情報を削除するなどの被害者の匿名性を考慮した処置を実施し，サイバー攻撃に関する情報を積極的に共有することが重要である．

本稿では，統計学的アプローチの具体例として，多変量解析を用いた具体例を示した．ここで用いた因子分析やクラスター分析は，機械学習の分野では教師なし学習モデルに分類される．教師なし学習モデルは，一般に正解がないデータに対して規則性を分析するための手法である．他方，サイバー攻撃においては，攻撃者のアトリビューションが特定される場合もありうることを本稿では示した．もしこのように正解が付与される事例が増加すれば，アトリビューションの正解をデータに付与し，教師あり学習モデルを訓練することで，サイバー攻撃のアトリビューションを推定することも可能となりうる．また，近年注目を集めている深層学習では，分類のための特徴を自動的に学習することが可能である．これにより，分析者が攻撃者との相関が高いと思われるインディケータを抽出しなくても，自動的に機械学習モデルが攻撃者の特徴を抽出する効果も期待できる．

参考文献

[1] 三村守，“サイバー戦入門（その 3）―攻撃者の正体を暴く―，”波涛，**43**, pp. 13–33, 2017.

[2] 三村守，田中英彦，“多変量解析による標的型攻撃の分類，”情報処理学会論文誌，**54**, pp. 2461–2471, 2013.

[3] 内閣サイバーセキュリティセンター，「日本年金機構における個人情報流出事案に関する原因究明調査結果」，https://www.nisc.go.jp/active/kihon/pdf/incident_report.pdf（2019 年 6 月 13 日閲覧）

[4] Mandiant, “APT1: Exposing One of China’s Cyber Espionage Units,” https://www.fireeye.com/content/dam/fireeye-www/services/pdfs/mandiant-apt1-report.pdf （2019 年 6 月 13 日閲覧）

[5] U.S. Department of Justice, “U.S. charges five Chinese military hackers for cyber espionage against U.S. corporations and a labor organization for commercial advantage,” https://www.justice.gov/ （2019 年 6 月 13 日閲覧）

[6] U.S. Department of Justice, “Grand Jury indicts 12 Russian intelligence officers for hacking offenses related to the 2016 election,” https://www.justice.gov/（2019 年 6 月 13 日閲覧）

[7] U.S. Department of Justice, “North Korean regime-backed programmer charged with conspiracy to conduct multiple cyber attacks and intrusions,” https://www.justice.gov/ （2019 年 6 月 13 日閲覧）

[8] Kaspersky Lab, “Equation group: The crown creator of cyber-espionage,” https://www.kaspersky.com/（2019 年 6 月 13 日閲覧）

[9] E. M. Hutchins, M. J. Cloppert and R. M. Amin, “Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains,” https://www.lockheedmartin.com/content/dam/lockheed-martin/rms/documents/cyber/LM-White-Paper-Intel-Driven-Defense.pdf（2019 年 6 月 13 日閲覧）

[10] CrowdStrike Intelligence Report, “PUTTER PANDA,” https://www.crowdstrike.com/blog/hat-tribution-pla-unit-61486/