2-2-3　AI 技術を用いたデータ駆動型サイバーセキュリティ
2-2-3　Data-driven Cybersecurity Empowered by AI Technologies

高橋 健志
TAKAHASHI Takeshi

サイバー脅威がますます深刻化する中、セキュリティ人材の確保が依然として困難な状況が続いている。そのため、セキュリティオペレーションの自動化が喫緊の課題となっている。この自動化を実現するには、AI 技術、分析対象としてのデータ、そしてドメインナレッジを効果的に組み合わせる必要がある。一方で、これらの自動化技術やシステムに組み込まれる AI 自体が、新たな攻撃ベクトルになりつつある。また、AI システムが社会により一層浸透するためには、AI そのものの信頼性を向上させることが不可欠である。そのため、AI のセキュリティ、解釈可能性、データのプライバシーなどを確保する必要がある。本稿では、本領域における我々の活動を紹介する。まず、長年にわたり蓄積してきたデータとその分析経験を基に実施してきた自動化技術の研究開発を紹介する。そして、今後加速すべき AI の信頼性に関する各種の研究開発について議論する。

As cyber threats become increasingly severe, securing sufficient cybersecurity personnel remains a persistent challenge. Therefore, the automation of security operations has become an urgent issue. To achieve this automation, it is necessary to effectively combine AI technology, data as the subject of analysis, and domain knowledge. However, the AI incorporated into these automated technologies and systems is becoming a new attack vector. Moreover, for AI systems to be more widely adopted in society, it is essential to improve the reliability of AI itself. Thus, it is necessary to ensure AI security, interpretability, and data privacy. In this paper, we introduce our activities in this field. First, we present our research and development on automation technologies based on data and analysis experience accumulated over many years. Then, we discuss various research and development efforts that need to be accelerated to enhance AI reliability in the future.

1　まえがき

サイバーセキュリティに対する脅威は深刻化を続けており、例えば攻撃者側は AI 技術などの活用により、その攻撃の効率化及び省力化を実施しつつある。そのため、セキュリティ対策をしっかりと講じていく必要があるが、その担い手であるセキュリティ人材の不足も深刻化してきている。そのため、サイバーセキュリティに対応できる人材の育成を加速するのと同時に、セキュリティを担保するための様々なオペレーションの自動化・省力化を実施していく必要がある。この自動化・省力化を実現するためには、アルゴリズム部分である AI 技術、解析対象となるデータ、そして効果的な分析アプローチを考案するためのドメイン知識が必要となる。

また、実現した自動化技術を実際のオペレーションに活用する際には、AI 自体の信頼性が問題となる。AI が導き出した判断に対する信頼性が担保できなければ、実際のオペレーションへの適用は難しい。実際、AI が導き出した判断を人間が検証する形で設計されているオペレーションが多いのが実態であり、AI そのものの信頼性を高めることによりそれらの自動化の水準を高めていくことが必要となる。そのため、AI が導き出した判断に対する一定レベルの根拠を掲示する技術や、そもそもその AI モジュールのセキュリティを担保する技術の研究開発が求められている。

これらの問題に取り組むべく、我々はこれまで蓄積してきたサイバーセキュリティ関連データと、その分析経験を最大限活用し、各種の AI 技術を活用することにより、様々なサイバーセキュリティオペレーションの自動化を実現する研究開発を実施している。また、取り扱う技術領域も広範囲にわたるため、様々な研究機関との連携研究にも積極的に実施してきている。当該分野では様々な研究開発が求められているものの、現在我々が注力している分野は下記のとおりである。

(1) AI を用いたサイバーセキュリティオペレーションの自動化の研究開発（**2** 参照）
　(ア) インターネット上でのマルウェアの活動検知・分析
　(イ) 組織内ネットワーク上での異常検知・分析

(ウ) マルウェアの検知・分析
(エ) 悪性 web サイトの検知・分析
(オ) 脅威インテリジェンスの生成・アノテーション
(2) AI モジュールの信頼性向上に関する研究開発（**3** 参照）
(ア) AI モジュールへのセキュリティ
(イ) AI 判定結果の説明性向上

本稿では、これらの領域における我々の研究開発活動の現状を紹介する。

2 AI を用いたサイバーセキュリティオペレーションの自動化の研究開発

我々は、これまで蓄積してきたデータを用い、また、NICT 内でのオペレーションから培ってきたノウハウをベースに、サイバーセキュリティオペレーションの自動化技術の研究開発を実施している。その主な活動を以下に紹介する。

2.1　インターネット上でのマルウェアの活動検知・分析

インターネット上で最も頻繁に観察される脅威ベクトルの一つにマルウェア感染が挙げられる。そのマルウェアが発するスキャンを観測することにより、インターネット上でのマルウェア活動の理解を深めることができる。従来、我々はダークネット空間（未使用の IP アドレス空間）に到達するマルウェアスキャンを収集し、オペレータが手動もしくはツールを用いて分析することで、マルウェア活動状況を解明してきたが、この分析プロセスの省力化・自動化と、更なる高度化を目指した研究開発に取り組んでいる。以下、本分野における近年の主な成果を紹介する。

(1) **Dark-Tracer**：インターネット上で新たなマルウェアの登場を自動的に検知する技術 “Dark-Tracer” を構築した [1]。従来は、宛先ポート番号別に送信されているスキャンパケット数が大幅に変化したことを手動もしくはツールにより検知することで、新たなマルウェアの発生を検知していたが、それではある程度の感染活動が拡大してからの検知になってしまうという問題があった。Dark-Tracer では、同時期に同一マルウェアによって感染した端末群のスキャン活動には同期性が観測されることを活用し、感染活動の数的拡大が進行する前に検知することに成功した。

(2) **FINISH**：上述の Dark-Tracer では、世界各地に設置されているダークネットセンサから得られるダークネットトラフィックを一か所に集約して分析することが前提となる。しかしながら、単独組織で収集できる情報には限界があること、そして各組織で収集した情報を組織の枠を超えて共有するのは困難なケースがあることなどから、分散処理可能な形に拡張することが求められている。我々は台湾大学と連携し、各観測地点で収集した情報を、他の組織と共有することなくその場で分析し、その分析結果のみ共有することで、データの秘匿性を確保しつつ高精度・負荷分散を実現する技術 FINISH を構築した [2]。

(3) **Fingerprinting**：スキャンは、その送信元のツールによって、特定のパターン（フィンガープリント）を示すことがある。例えば、同一のマルウェアやスキャンツールからのスキャンであれば、同一のフィンガープリントを観測できるケースがある。我々は、大量のスキャンパケットから、数式で表現可能かつ頻繁に現れるフィンガープリントを抽出する技術を構築した [3]。これにより、これまで知られていなかったフィンガープリントの特定にも成功した。

上記のそれぞれの取組を通じ、インターネット上でのマルウェア活動の早期発見及びその分析を効率化することが可能となった。

2.2　組織内ネットワーク上での異常検知・分析

各組織のネットワークでは、外部からの不正侵入や内部の異常を検知すべく、オペレータやシステム管理者などが、様々なセキュリティアプライアンスを活用したオペレーションを展開している。そのオペレーションで何らかの不正通信の兆候を見逃すことは、組織に甚大な被害を与えかねないリスクを孕（はら）んでいるため、人的ミスは最小化することが望まれる。オペレータは、大量に発行されるセキュリティアラートを確認・検証し、警告されている状況を理解し、必要なアクションを実施する必要がある。我々は、NICT にて実際に発生しているセキュリティアラートを分析し、これらのオペレーションを効果的に実施するための研究を実施している。以下、本分野における近年の主な成果を紹介する。

(1) **アラートスクリーニング技術**：セキュリティアプライアンスから大量に発行されるアラートは、一定のフィルタリングルールに従いその総量を削減した上で、人手による検証作業を実施しているのが現状である。そして、最終的には最大数個の対処が必要なアラートを抽出することとなるが、この作業は大変な労力を要するものであり、疲労等を原因とする人的ミスを誘発しやすい状況にある。本研究では、AI 技術を用いることにより、大量のアラートの中から真にアク

ションが必要なアラートを抽出する技術を構築している [4]–[6]。現時点において、従来比 50 % の分析対象アラートの削減を実現できている。

(2) **信頼性の高い IDS 技術の構築**：トラフィックの異常検知に関する研究は、これまでも様々な報告がなされてきたが、我々は、従来の異常通信の検知精度を向上する研究開発 [7] に加え、後述の検知した結果の解釈可能性を示す技術 [8] 及び AI 技術の検知を免れる攻撃に関する研究開発を実施している（**3** 参照）。

2.3　マルウェアの検知・分析

インターネットや組織内の通信を分析することで、様々なセキュリティ対策を実施することができる一方で、攻撃者側が用いているマルウェアそのものを直接分析することができれば、セキュリティ対策をより効果的に講じることができる。マルウェアは、その対象に応じて、例えば IoT マルウェアや Android マルウェアなど、様々なものが存在し、そのそれぞれを対象とした分析を、アナリストがツールを用いて実施しているのが現状である。これらのツールには AI 技術が使われているものも多数存在するが、分析の信頼性や深さの更なる追及が必要であり、我々はそのための研究開発を実施している。以下、本分野における近年の主な成果を紹介する。なお、下記のそれぞれの技術は、データセットの都合上、Android と IoT に特化した内容になっているものの、マルウェアの検知・分析技術として様々なタイプのマルウェアに対して適用可能である。

(1) **Android 向けマルウェアの検知技術**：Android アプリがマルウェアか否かの判定を自動的に実施するための技術を提案した。提案技術では、大量の Android アプリとそれらのメタ情報（アプリのカテゴリや説明文）を収集し、それらから特徴情報を抽出して学習させることにより、Android アプリがマルウェアか否かを高精度に判定可能となる。抽出した特徴の次元数を、コンテクスト情報を考慮して低減することにより、従来方式より高精度化及び処理時間の大幅低減を実現した [9]。また、Android の特徴量間の関係を考慮して特徴量を調整することにより、マルウェア検知精度を向上した [10]。

(2) **IoT 機器を攻撃対象とするマルウェア群の機能分析技術**：IoT 機器を標的として攻撃を実施するマルウェアを効果的に分析するための研究開発を実施している。マルウェア感染端末数が最も多いのが IoT 機器であり、その被害を低減することは喫緊の課題である。様々な研究を実施してきているが、マルウェアのバイナリの類似度が高いものは機能も類似することを利用し、類似する機能を有するマルウェア群を実時間でグループ化・系統樹生成技術を構築 [11] した。これにより、分析の手間を低減する研究を実施している。また、より直接的にマルウェアの機能を特定すべく、マルウェアバイナリの構造をグラフ化して機械学習に用いる特徴情報として活用可能にする技術 [12]、グラフ化せずにバイナリ自体のバイトシーケンスを特徴情報として抽出する技術 [13]、さらにはマルウェアの機能を特定するためのシグネチャ構築技術 [14] も開発している。

2.4　悪性 web サイトの検知・分析

サイバーセキュリティを担保するためには、マルウェア感染拡大のためのスキャンや不正侵入など、攻撃者からのアクションに対応することも重要であるが、それ以外にも、ユーザが自ら被害に遭いに行くケース、すなわち、悪性 Web サイトにアクセスしてしまうケースにも対応する必要がある。これまでも、多数のアナリストが様々なツールを駆使して悪性 Web サイトを見つけ出し、その URL をブラックリストに登録し、そのブラックリストに登録された URL へのアクセスを監視・遮断することで、ユーザを守る取組が実施されて生きているが、日々多数の悪性 Web サイトが登場する中で、それらをより効果的かつ自動的に検出していく技術が必要である。以下、本分野における我々の主な研究成果を紹介する。

(1) **フィッシングサイトの検知技術**：Web ユーザにとって、大きな脅威の一つに、フィッシングサイトが存在する。フィッシングサイトにアクセスすると、情報漏えいやマルウェア感染につながることがあるため、そういったサイトは事前に検知し、ブラックリストに登録しておくことが望ましい。しかしながら、フィッシングサイトはその生存期間が数時間しかないため、ブラックリストの有効性も限定的である。そのため、アクセスしようとしているページを効果的かつ瞬時に分析できる技術が求められている。様々なアプローチが存在するが、我々は、サイト内で利用されている JavaScript コードをグラフ化する技術を構築した [15]。これにより、そのグラフ情報を特徴量として機械学習に活用可能となり、高精度でのフィッシングサイト検知を実現した。ページのコードを分析するアプローチがある一方で、我々はページを画像として捉え、その画像中のフォームの存在とそのフォームに表示される説明文を分析し、特徴情報化することで、より高いフィッシングサイト検知精度を実現した [16]。

(2) **危険なサイトを検知する技術**：フィッシングサイトなどの悪性サイトを検知する研究と並行し、そのようなサイトにアクセスするユーザのアクセスログを分析する研究も実施してきている。その結果、ユーザが悪性 URL に到達するまでの経路情報が明らかになり、それらの情報を重ね合わせることにより、悪性サイトに到達する可能性が高い危険なドメインを特定することに成功した [17]。これらの危険なドメイン自体には、悪性コンテンツは存在しないものの、これらのドメインにアクセスしたユーザはその後悪性 URL に到達する可能性が高いため、そのようなドメインへのアクセスにアラートを発行することで、悪性 URL に到達するユーザ数を低減することが可能となった。

2.5　脅威インテリジェンスの生成・アノテーション

サイバーセキュリティを担保するためには、サイバー攻撃や、悪性サイト、マルウェアの検知・分析などの技術発展と並行して、サイバーセキュリティに関する状況や対策方法を広く共有し、世界規模で知識を共有する取組も必要不可欠である。そのための手段として、脅威インテリジェンス（セキュリティに関する情報がまとめられたもので、通常は精査されており、一定以上の品質が担保されているもの）が注目されているが、現状では質の良い脅威インテリジェンスが十分に供給されている状況とは言いがたい。また、インテリジェンスを効率的に生成する技術や、さらには機械処理しやすい形に昇華するための技術が必要不可欠である。以下、本分野における我々の主な研究活動を紹介する。

(1) **ダークウェブからのインテリジェンス抽出**：通常のインターネット空間とは隔離されたダークウェブ空間は、違法ドラッグの取引を含む様々な犯罪のプラットフォームとなっている。サイバー攻撃に関しても同様に、マルウェアやエクスプロイトキットなどの販売や、攻撃キャンペーンなどの議論が本空間にて行われている。そこで我々は、ダークウェブ上の情報を収集･分析することで、サイバー攻撃の準備段階の情報や攻撃予告などを抽出し、表層ウェブ上のセキュリティレポートや脆弱性情報などからでは得られない脅威インテリジェンスを生成する技術を研究している。一例として、ダークウェブ上での DaaS（DDoS as a Service）と実際に発生したインシデントのレポートとを紐づけることに成功している [18]。

(2) **非構造化データから構造化されたインテリジェンスを生成**：様々なサイバーセキュリティ情報が Web にて提供されるようになってきた一方で、それらの情報は構造化されていない、もしくはその構造に統一性がないため、解析処理の自動化が困難なケースが多く、セキュリティオペレーターが人手により処理しているケースが多い。そこで我々は、複数の情報源から定常的にサイバーセキュリティ関連情報のクローリングを行う収集基盤を構築するとともに、構造化されていない情報に対してマルチラベルを付与して、必要なインテリジェンスを構造化して抽出する技術を構築した [19][20]。さらに、Web のレポート群から LLM を用いることで、必要なカテゴリの情報をサマリーとして抽出し、インテリジェンス化する技術を構築している（現在論文投稿中）。

(3) **ソースコードの脆弱性検知・分析・レポート作成補助**：脅威インテリジェンスの代表的なものに脆弱性情報があるが、それを自動で生成するための技術の構築に取り組んでいる。我々は、プログラムのソースコードからプログラムのグラフ構造を抽出し、それを学習することにより、脆弱性存在の可能性と、そのプログラム中での該当箇所を特定し [21]、さらにはその特定された脆弱性がどの種類の脆弱性であるのかを自動的に判定する技術を構築した [22][23]。また、脆弱性情報から、その深刻度を示す CVSS スコアを自動的に算出する技術を構築した [24]。これらの技術を発展していくことにより、ソースコードを検査し、脆弱性の存在を検知し、そのレポートを自動生成することにより、現場の負荷を低減していきたい。

3　AI モジュールの信頼性を向上する技術に関する研究開発

AI 技術が一定のレベルまでコモディティ化しつつある現在、AI を用いたセキュリティ関連技術も多数登場してきている。しかしながら、これらの技術の利用には、慎重にならざるを得ないケースも多い。実際、セキュリティインシデントが発生した際に、その責任を AI が負うことはできず、それを利用している者の責任になるためである。そのため、AI を活用した技術は、人間が確認・検証等を実施するプロセスを設けた上で利用されているケースが多い。

今後、AI の利活用を更に進めていくためには、AI に対する信頼性の向上が不可欠である。本分野における我々の活動はまだ日が浅く、多くの成果を挙げてい

くのはこれからという状況ではあるが、今後の少子高齢化社会における AI 技術の利活用は必要不可欠であり、それを促進するためには必要不可欠な研究開発であると確信しており、今後、重点研究開発領域として研究開発に注力していく。現時点で我々が取り組んでいるのは、主に下記の 2 項目である。

3.1　AI モジュールのセキュリティ

AI モジュールには、いくつかの脆弱性が知られており、その一つに敵対的サンプルがある。この攻撃手法を用いると、本来陽性と判断されるべきところが陰性と判断されるようは、誤判定を誘発することが可能となる。様々なサイバーセキュリティシステムが AI 技術を採用する中、それらのシステムに対する攻撃を実施し、その実現可能性と影響度、またその対策について検討していく必要がある。

異常通信の検知、マルウェア分析、そしてフィッシングサイト検知の領域において、検知・分析の正常な判定を阻害する攻撃とその対策の研究開発を実施しているが、AI ベースの IDS 技術のセキュリティ検討については、一定の成果が創出されている。AI 技術を利用した IDS は、その AI 技術の検知を免れる攻撃が実現可能であることが知られている。我々はその攻撃を再現した結果、AI 技術の検知を免れても元々の攻撃性が損なわれてしまう可能性が高いことを確認し、元々の攻撃性を維持した形で AI 技術の検知を免れる攻撃手法を検討している [25]。

3.2　AI 判定結果の説明性向上技術

AI モジュールの判定結果を盲目的に信じたサイバーセキュリティオペレーションの実施がなされるケースは多くなく、実際には人間のオペレータがその判定結果を検証し、状況を理解した上で、次のアクションが実施されている。しかしながら、その判定結果を検証する際に、AI 技術の解釈可能性が問題となっている。AI 技術はその判定の根拠が分かりにくく、特に高精度を誇る深層学習などを活用した際には、その処理は人間にとっては解釈できないブラックボックスとなっている場合が多い。そのため、その解釈可能性を向上するための試みが、主に人工知能分野にて行われているが、その解釈可能性の向上はいまだ発展途上にある。そして、サイバーセキュリティ領域で AI ツールを利用するオペレータの視点からすると、たとえ AI 領域での解釈可能性が向上しても、彼らが知りたいレベルの解釈可能性は提供される見込みがないのが現状である [26]。そこで我々は、AI の判定結果を、AI の挙動の解釈可能性というレベルではなく、サイバーセキュリティの観点から生じている事象の解釈可能性を提供するための研究開発を実施している。

トラフィックの異常検知に関する研究領域においては、これまで十分な解釈可能性を掲示する技術は確立されていない。我々は、この領域における実態調査を徹底的に実施したのち [26]、AI モデル自体の解釈可能性と、サイバーセキュリティ領域でのオペレータに有意な解釈可能性の間に大きなギャップがあることを示し、そのギャップを埋めるための研究開発を実施している [8]。同様に、フィッシングサイト検知に関する研究についても、画像情報を入力としてフィッシングサイトを検知する手法について、画像のどの部分に着目して判定を下しているのかを特定する技術を構築した [27]。

4　今後の研究開発の方向性

冒頭に述べたとおり、我々はセキュリティオペレーションの自動化を加速すべく、AI × Cybersecurity の研究開発を実施しており、これまではその中の AI を用いたサイバーセキュリティオペレーションの自動化に焦点を当てた研究開発に注力してきた。これらの研究を通じ、AI 技術を用いたサイバーセキュリティオペレーションの自動化を促進できることを検証してきた。また、それらの技術を実際の社会で実装してもらうためには、AI 技術自体の信頼性を高めていく必要があるという認識の下、我々は AI モジュールのセキュリティに関する研究と、AI モジュールの判定結果の解釈可能性に関する研究を実施している。前者については、敵対的サンプルを用いた攻撃可能性やデータポイズニング攻撃の攻撃可能性・影響度の分析を実施してきており、後者についても、AI モデルが下した判断に対する解釈可能性を高め、AI モデルの評価をより安心･信頼して利用できるようにしていくべく、検討を進めている。

AI × Cybersecurity の研究開発は、今後の安心安全なサイバー社会を構築・維持していく上で、最重要となる研究開発課題の一つであると認識している。特に近年急速な発展を遂げている LLM や生成 AI 技術については、十分な技術的評価を実施の上、安心して安全かつセキュアに活用できる基盤技術を構築していく必要がある。また、この課題認識は日本に閉じるものではないため、研究開発を推進するにあたっては、世界中の研究機関と切磋琢磨（せっさたくま）しつつも連携し、グローバルな競争力を獲得していくべきであると考えている。これらの活動を通じ、安心・安全で、かつ信頼されるサイバー社会の構築に貢献していきたい。

謝辞

本稿の研究成果は、サイバーセキュリティ研究所サイバーセキュリティ研究室内 AICS チームにおいて創出されたもので、班涛主任研究員、古本啓祐研究員、韓燦洙研究員、田中智研究員、宮本耕平研究員、Rozi Muhammad Fakhrur 研究員、Ndichu Samuel 研究員、Tanuwidjaja Harry Chandra 研究員、倍味幸平研究技術員らと共に実施した研究成果である。

【参考文献】

1 C. Han, et al., “Dark-TRACER: Early Detection Framework for Malware Activity Based on Anomalous Spatiotemporal Patterns,” IEEE Access, 2022.
2 YW.Chang, et al., “FINISH: Efficient and Scalable NMF-Based Federated Learning for Detecting Malware Activities,” IEEE Trans. Emerg. Top. Comput., 2023.
3 A. Tanaka, C. Han, and T. Takahashi, “Detecting Coordinated Internet-Wide Scanning by TCP/IP Header Fingerprint,” IEEE Access, 2023.
4 T.Ban, et al., “Breaking Alert Fatigue: AI-Assisted SIEM Framework for Effective Incident Response.” Appl. Sci. 2023.
5 S.Ndichu, et al., “AI-Assisted Security Alert Data Analysis with Imbalanced Learning Methods,” Appl. Sci. 2023.
6 ME. Aminanto, et al., “Threat Alert Prioritization Using Isolation Forest and Stacked Auto Encoder With Day-Forward-Chaining Analysis,” IEEE Access, 2020.
7 K. Miyamoto, et al., “Consolidating Packet-Level Features for Effective Network Intrusion Detection: A Novel Session-Level Approach,” IEEE Access, 2023.
8 HC.Tanuwidjaja, et al., “Hybrid Explainable Intrusion Detection System: Global vs. Local Approach,” Workshop on Recent Advances in Resilient and Trustworthy ML Systems in Autonomous Networks, 2023.
9 B.Sun, et al., “Detecting Android Malware and Classifying Its Families in Large-scale Datasets,” ACM Trans. Manage. Inf. Syst., 2022.
10 YC.Chen, et al., “Impact of Code Deobfuscation and Feature Interaction in Android Malware Detection,” IEEE Access, 2021.
11 T. He, et al., “Scalable and Fast Algorithm for Constructing Phylogenetic Trees With Application to IoT Malware Clustering,” IEEE Access, 2023.
12 CY. Wu, et al., “IoT malware classification based on reinterpreted function-call graphs,” Computers & Security, 2023.
13 TL.Wan et al., “Efficient Detection and Classification of Internet-of-Things Malware Based on Byte Sequences from Executable Files,” IEEE Open J. Comput. Soc., 2020.
14 R.Kawasoe, et al., “Investigating behavioral differences between IoT malware via function call sequence graphs,” ACM Symposium on Applied Computing, 2021.
15 M. F. Rozi et al., “Detecting Malicious JavaScript Using Structure-Based Analysis of Graph Representation,” IEEE Access, 2023.
16 SC. Lin, et al., “SenseInput: An Image-Based Sensitive Input Detection Scheme for Phishing Website Detection,” IEEE International Conference on Communications, 2022.
17 T.Takahashi, et al., “Tracing and Analyzing Web Access Paths Based on User-Side Data Collection: How Do Users Reach Malicious URLs?,” International Symposium on Research in Attacks, Intrusions and Defenses, 2020.
18 K.Furumoto et al., “Extracting Threat Intelligence Related IoT Botnet From Latest Dark Web Data Collection,” IEEE International Conferences on Internet of Things, 2021.
19 R. Nagasawa et al., “Partition-Then-Overlap Method for Labeling Cyber Threat Intelligence Reports by Topics over Time (Letter),” IEICE Transactions on Information and Systems, vol.E104-D, no.05, pp.556-561, May 2021.
20 Y.Osada et al., “Multi-Labeling with Topic Models for Searching Security Information,” Annals of Telecommunications, 2022.
21 MF. Rozi, et al., “Securing Code with Context: Enhancing Vulnerability Detection through Contextualized Graph Representations,” IEEE Access, in press.
22 M. Aota, et al., “Multi-label Positive and Unlabeled Learning and its Application to Common Vulnerabilities and Exposure Categorization,” IEEE International Conference on Trust, Security and Privacy in Computing and Communications, 2021.
23 M.Aota, et al., “Automation of Vulnerability Classification from its Description using Machine Learning,” IEEE Symposium on Computers and Communications, 2020.
24 S.Nakagawa et al., “Character-Level Convolutional Neural Network for Predicting Severity of Software Vulnerability from Vulnerability Description,” IEICE Trans. Inf. & Syst., 2019.
25 M.Mbow et al., "Evading IoT Intrusion Detection Systems with GAN," Asia Joint Conference on Information Security, 2024.
26 F.Charmet, et al., “Explainable artificial intelligence for cybersecurity: a literature survey,” Ann. Telecommun., 2022.
27 F. Charmet, et al., “2024. VORTEX : Visual phishing detectiOns aRe Through EXplanations,” ACM Trans. Internet Technol., 2024.

高橋 健志（たかはし たけし）
サイバーセキュリティ研究所
シニアマネージャー
博士（国際情報通信学）
サイバーセキュリティ、人工知能

【受賞歴】
2020 年　The 19th IEEE International Conference on Trust, Security and Privacy in Computing and Communications, Best Paper Award
2018 年　コンピュータセキュリティシンポジウム 2018（CSS2018）最優秀論文賞
2012 年　日本 ITU 協会賞（奨励賞）