セキュリティオペレーションの自動化に向けた機械学習技術の活用

国立研究開発法人情報通信研究機構
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 研究マネージャー　高橋（たかはし）　健志（たけし）

1. 自動化が求められるセキュリティオペレーション

ここ数年、ランサムウェアやIoT機器のボットネット化など、様々なサイバー脅威がインターネット上で猛威を振るっている。そのため、各種のセキュリティオペレーションの実践が求められるが、それに必要な人材が不足しているのが現状である。迅速に本状況に対応するには、人材育成と並行してオペレーションの自動化を実施していく必要がある。

我々は、機械学習を中心とした人工知能（AI）技術を活用することにより、その自動化を推進したい。既に我々は、マルウェア（不正プログラム）の検知や分類、その機能の類推、インターネット上のマルウェア活動の傾向変化検知、組織内通信の異常検知、さらには将来の攻撃予測など、様々な技術革新を目指して研究開発に取り組んできている。オペレーション自動化の実現にあたっては、機械学習以外にも、機械処理可能な形式での情報交換・蓄積の実現や、現場オペレーションの分析に基づく定型処理の自動化などが必要となるが、本稿では機械学習を用いた自動化技術に焦点を当てて、我々の活動を紹介したい。

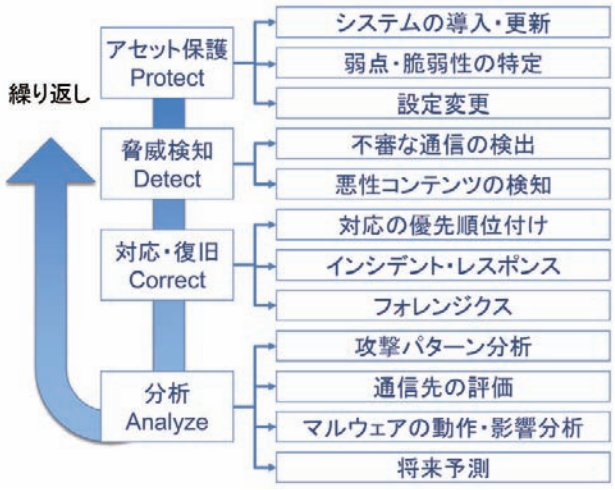
2. セキュリティオペレーションの概観

本章では、本稿における自動化の議論の対象となるセキュリティオペレーションについて概説する。また、本分野における機械学習の活用状況について概論する。

2.1　オペレーションの実施サイクル

図1に、セキュリティオペレーションの概観を示す。本図ではオペレーションを「アセット保護」「脅威検知」「対応・復旧」「分析」の4段階に分けて説明している。

アセット保護の段階では、脅威に対し事前に対策を講じる。具体的には、システムの導入や更新、脆弱性の特定やパッチ適用、設定の変更などを実施する。事前に対策を講じても脅威は存在しており、それを検知するのが脅威検知の段階である。具体的には不審な通信の検出や、マルウェアなどの悪性コンテンツの検知を実施する。脅威を検知すると、その対応を実施する対応・復旧の段階に入る。ここでは各種連絡・連携やトリアージ等のインシデント・レスポンスを実施する。同時にフォレンジックを実施する。なお、検知された脅威によりインシデントが発生しなかった場合には、本段階の手続きは簡易なものになる。また、インシデント対応後に各種の脅威について詳細に分析し、再発防止策などを構築するのが分析の段階である。具体的には、攻撃パターンの分析や、通信先の評価などを実施し、マルウェアのシグネチャやブラックリストURLなどを生成する。また、マルウェアの動作分析・影響度分析、将来予測なども実施する。そして、分析結果に基づき、改めてアセット保護を実施するのが通常であり、この4段階は繰り返し実施される。



■図1.　セキュリティオペレーションの概観

2.2　機械学習の活用領域

上述のオペレーションのうち、機械学習が積極的に用いられている領域は未だ限定的である。機械学習を用いたセキュリティ技術は従来から存在しており、例えばマルウェアの検出やトラフィックパターンの異常検知などの脅威検知の段階に相当する部分については、10年以上も前から積極的に検討がなされてきている。とはいえ、昨今のAIブームを受け、サイバーセキュリティ技術領域でも多数の検討が加速されてきており、その結果、脅威検知技術の深化や、それ以外の段階のオペレーションを対象とした技術検討がなされてきている。

上述の4段階は、処理の流れの観点から便宜的に定義したものであり、その境界線に技術的な断絶が存在するわけではない。セキュリティのオペレーションに機械学習を用いるということが一つの研究領域として確立してきた結果、これまでは注目されていなかったオペレーションがユースケースとして取りあげられ、積極的に研究されるようになってきたのではないかと感じている。例えば、マルウェア検知（脅威検知段階）に関する研究は、現在ではマルウェアの機能推定を実現する分類手法（分析段階）の研究へと進化してきている。同様に、DDoS攻撃の発生を捉える研究（検知段階）は、現在ではDDoS攻撃に参画しているボットを特定する研究（分析段階）へとそのスコープを広げてきている。

本分野の研究は、保持しているデータセットにより、その方向性が大きく左右される。そのため、以下ではまず、我々の保持しているデータセットについて簡単に紹介し、それを踏まえて我々の研究活動をいくつか紹介したい。

3. 研究競争力の源泉となるデータセット

サイバーセキュリティ分野において、データセットは研究開発の競争力の源泉となる。機械学習は大量のデータを学習・分析する技術であるため、どれだけ多く、良質なデータを収集・蓄積できるかが非常に重要になる。表のとおり、我々は様々なデータを蓄積しており、例えばダークネット（未使用IPアドレス）トラフィックの観測データについては約10年間、収集・分析・蓄積を継続してきている。

■表．NICTが保有するデータ例

カテゴリ	蓄積データの具体例
ダークネット関連データ	未使用IPアドレスへの攻撃関連通信データ。pcapファイル、統計情報、悪性ホスト情報、など。
ライブネット関連データ	NICT内の通信データ。pcapファイル、フローデータ、セキュリティアプライアンスが生成するアラート、など。
ハニーポットデータ	高対話型/低対話型のハニーポット/クライアントハニーポットで収集したデータ、など。
マルウェア関連データ	マルウェア検体、静的・動的解析結果、など。
Android APK関連データ	APKファイル、マーケットのアプリ情報、など。
Webページデータ	URLリスト、Webコンテンツ、コンテンツの評価結果、など。
ブログ・記事	TwitterのTweet、セキュリティベンダーブログ、など。
スパムメール関連データ	スパム（ダブルバウンス）メールデータ、統計情報、など。
商用インテリジェンス情報	各社から購入したマルウェアをホストしているサイトの情報、ボットやC&Cのリスト、ドメイン履歴データ、検体、脅威レポート、など。

4. 機械学習の活用

機械学習を活用することにより、各種のセキュリティオペレーションの効率化・自動化が実現可能になる。もちろん、すべてのオペレーションが自動化されることは少なくとも直近では考えにくいが、機械学習を適用することで効率化を実現できる部分は確実に存在する。我々は上述の収集データを活用し、機械学習を用いた分析・自動化技術を検討しており、本章ではそのいくつかを紹介する。

4.1　パッカーの特定

多くのマルウェアは、分析を難しくするため、パッカーと呼ばれる難読化ツール（実行形式ファイルをその機能を損なうことなく暗号化するツール）を利用している。パッカーはソフトウェアを圧縮し、解凍プログラムと共に1つのファイルに再構成する。そのため、マルウェアの分析をする際には、どのパッカーが利用されているのかを特定する必要がある。本研究は、その特定を自動化することにより、マルウェア分析のオペレーションの効率化に貢献する。

シグネチャを用いたパッカー特定手法が従来から存在しているが、その手法ではスケーラビリティが低く、パッカー特定ができないケースが増加している。そこで我々は、シグネチャによらない、機械学習に基づくパッカー特定手法を提案している。復号ルーチンに利用したパッカーの特徴が表れることが多いという特性を考慮し、提案手法ではパッカーの復号ルーチンのバイナリ列の先頭から一定のバイト数を特徴情報とし、サポートベクタマシン（SVM）を用いてそれを分類する。評価実験では、99.46%の精度でパッカーを自動的に特定できることを確認した[1]。

4.2　DDoS攻撃の発生検知

我々は、インターネット上のDDoS攻撃（分散型サービス不能攻撃）発生を早期検知するのにも機械学習を活用している。ダークネットではバックスキャッタ*が観測されるが、それを分析することで、DDoS攻撃の発生を早期に検知できる。具体的には、ダークネットトラフィックについて、IPアドレスごとに17種類の特徴を定め、その特徴に基づいて機械学習にて分類処理を実施することにより、DDoS攻撃

＊　送信元IPアドレスが詐称されたDoS攻撃（SYN-flood攻撃）を受けているサーバからの応答（SYN-ACK）パケットのこと

に関係するバックスキャッタか否かを判定する。この17種類の特徴とは、一定時間内のパケット数、パケット送信間隔（平均値/標準偏差）、ソースポート番号別送信パケット数（平均値/標準偏差）、宛先ポート番号別パケット数（平均値/標準偏差）、ペイロードサイズ（平均値/標準偏差）、などである。

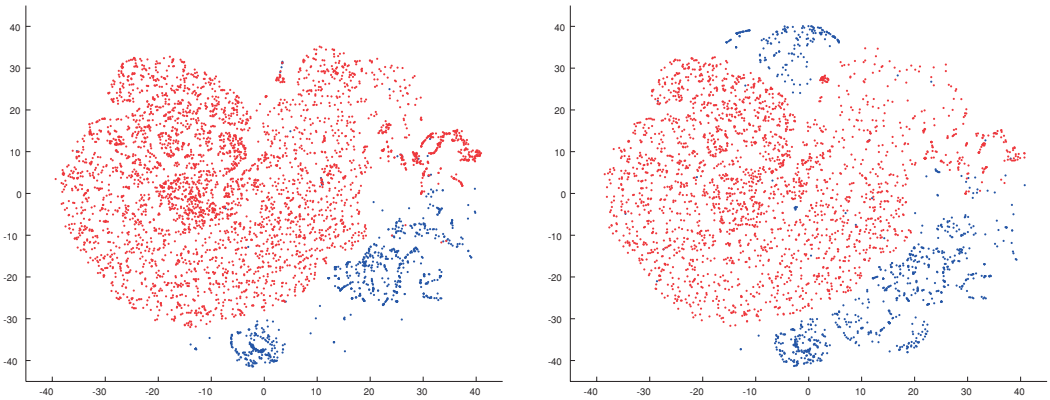
ここで、インターネット上の実トラフィックを分析する際には、その傾向が刻一刻と変化していくことを考慮し、学習を追加で繰り返していく必要がある点に留意されたい。**図2**は実験開始月（**左図**）及びその翌月（**右図**）の特徴ベクトルをt-SNEと呼ばれる次元圧縮アルゴリズムを用いて2次元化した結果を示している。この図のうちの赤い部分がDDoS攻撃に関連する特徴ベクトルをプロットしたもので、そうでないものが青い部分である。両図を比較すると、この赤と青の領域が少しずつ変化している様子が分かる。[2]

4.3　Androidアプリ分析

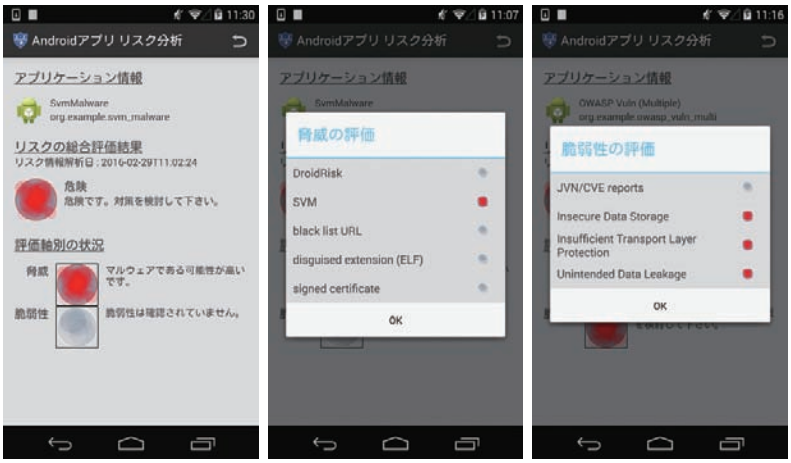
Androidアプリ はGoogle Playなどの各種オンラインマーケットにて配布されているが、これらの中にはマルウェアが混在することがある。そのため、アプリがマルウェアであるかどうかを判定する必要がある。我々は、アプリの静的分析結果及びオンラインマーケット記載のメタ情報から特徴を抽出し、SVMに入力することで、マルウェアか否かを判定する手法を構築した。

アプリの静的解析では、APKファイル内に存在するmanifestファイルからパーミッション情報を、Javaバイトコードから API呼び出しの一覧情報を抽出する。また、オンラインマーケットからはアプリのカテゴリ情報及び説明文に関する情報を収集しており、これらの情報から特徴情報を構築している。また、その収集した特徴情報の重要度を評価し、マルウェア検知に貢献する特徴のみを抽出することで、その検知精度のさらなる向上を実現している。

評価実験では、我々は94.07%の精度でマルウェアを判定することに成功している[3]。同時に、**図3**のとおり、そのマルウェアか否かの判定結果を可視化するツールも構築している。本ツールでは、脅威と脆弱性の両視点からアラート（赤信号）の有無を可視化しているが、その中の脅威情



■図2.　時間推移による特徴領域の変化（左図：実験開始月、右図：その翌月）



■図3.　Androidアプリのリスク分析・可視化ツール

報の分析に上述の手法を実装している。

4.4　その他の研究分野

上記の事例に限らず、我々は様々な研究開発を実施している。現在注力している研究を2つここで簡単に紹介する。

1つ目は、セキュリティアプライアンスから出される**アラートのスクリーニング**技術の検討である。セキュリティアプライアンスは各種のアラートを出すが、その数は膨大であり、現場のオペレータは対応に窮している。現在は固定ルールを定義し、その結果抽出されるアラートを手作業にて精査することで、真に対応すべきアラートを抽出しているが、更なる自動化が求められている。そこで我々は、機械学習を用いることにより、これらのアラートの中で特に重要なものを抽出し、優先順位付けする技術を検討している。

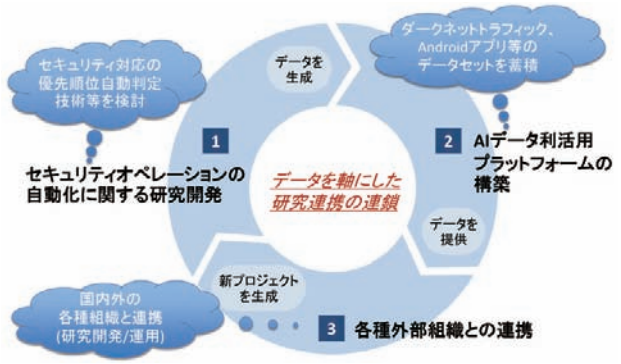
2つ目は、**脆弱性の深刻度評価**技術の検討である。組織内のソフトウェア資産には各種の脆弱性が存在することがあるが、これらの脆弱性がどれくらい深刻であるかを評価することにより、対処すべき脆弱性の優先順位付けをする必要がある。現在はその深刻度の評価は手作業により実施されている。そこで我々は、機械学習を用いた脆弱性の深刻度評価の自動化を検討している。

これらの研究は、どちらも現場のオペレータがセキュリティ対策を考える際に、対応の優先順位を具体的に明示してくれるものであり、実際の現場での課題をまさに機械学習により解決しようとしている。

5. 研究連携を加速するための仕組み

セキュリティオペレーションの自動化については、その他にも多数の研究が必要であり、それらの研究を実施するにはデータセットやコンピュータリソース、そして人的リソースが必要となる。そこでNICTでは、国内の研究開発機関を中心に、お互いに連携して取り組める仕組み作りにも注力している。その活動を強力に推進しているのが、知能化学融合研究開発センター（AIS）である。AISは2017年度にNICT内に設立され、サイバーセキュリティ領域におけるAI活用は本センターの重点課題の一つとなっている。

AISでは、これらの自動化に関する研究開発やデータ収集を通じ、AIデータテストベッドなどの研究プラットフォームを整備し、我々の競争力の源泉を拡充していく。同時に、それらのデータを基に各種外部機関と連携し、更なる自動化研究の推進・加速を実現していくという、データを軸にした研究連携の連鎖を実践していく方針である（図4参照）。



■図4.　AISのデータを軸にした研究連携

元々、AI技術とサイバーセキュリティ技術は異なるコミュニティで発展してきた経緯があるため、両領域の技術の効果的な融合には時間と労力を要する。しかしながら、その両者を幅広い領域で融合させることができれば、セキュリティの自動化は大きく進展する。AISでは、その融合を加速し、双方の領域に精通した人材を発掘・育成し、サイバーセキュリティ技術を進化させるのに重要な役割を果たそうとしている。

6. おわりに

セキュリティ分野での機械学習の適用範囲は今後急速に拡大していくことが予想される。我々は、競争力の源泉となるデータを継続的に収集・蓄積していくのと同時に、解決すべき課題を再定義し、それに適用可能な機械学習技術自体の高度化にも取り組んでいく。また、機械学習技術を実際のセキュリティオペレーションの現場に適用するためには、リアルタイム性の担保など、より一層の技術革新が必要である。東京2020オリンピック・パラリンピック競技大会に向けて、オペレーションの自動化を進めるべく、研究開発を加速していきたいと考えている。

参考文献

[1] R. Isawa, et al., “An Accurate Packer Identification Method using Support Vector Machine,” IEICE Trans. Fundamentals, Vol.E97-A, No.1, pp.253-263, 2014.

[2] I. Skrjanc et al., “Evolving Cauchy Possibilistic Clustering and Its Application to Large-Scale Cyberattack Monitoring,” IEEE Symposium Series on Computational Intelligence, Hawaii, 2017

[3] T. Takahashi, et al., “The Usability of Metadata for Android Application Analysis,” The 23nd International Conference on Neural Information Processing, Kyoto, 2016