採るべき対策の「なぜ？」に答えるAIの可能性

細見 格

要 旨

サイバー攻撃がより高度かつ組織的に展開されるようになるなか、機械学習型AIによる攻撃検知技術が効果を示しつつある一方で、具体的な対策を決定するための筋立てと十分な証拠の確保は人間のアナリストが持つ知識やノウハウと思考力が必要とされています。ディープラーニングなどの機械学習型AIのみでは困難な、こうした知的作業を定型モデルで自動化するスタートアップ企業が北米で注目されており、今後の増加が予想される非定型で複雑な攻撃にも対応するためのアプローチとして、論理推論型AIの可能性を述べます。

KeyWords　人工知能／論理推論／セキュリティ・オペレーション／サイバーキルチェーン／SIEM／Security Orchestration and Automation

1. はじめに

2002年頃から確認され各国で深刻な被害をもたらした標的型攻撃、2013年頃から急増したランサムウェアなど、サイバー攻撃の脅威は年々深刻化しています。攻撃手段の多様化、標的に応じた個別化、潜伏や痕跡消去などの巧妙化が進展している一方で、対抗手段の1つとしてマルウェアやその振る舞いの自動検知に人工知能（AI）の技術が利用され始め、成果が現れてきています。機械学習を用いたAI技術は、従来のシグネチャやルールによる方法に比べて亜種や未知のマルウェアを検知可能にし、攻撃の多様化への対抗手段として有望です。しかし、機械学習ですべてを解決することは困難であり、人間のアナリストが必要なくなることは当分ないと考えられます。それはなぜでしょうか。本稿では、サイバーセキュリティにAIを活用するうえで考えるべきいくつかの「なぜ？」に焦点をあて、そこから見えてくる課題に対してNECが注目している論理推論型AIの可能性を述べます。

2. なぜ、サイバーセキュリティにAI技術が重要か

2.1 産業化とAI技術で増大するサイバー攻撃の脅威

図1のイメージのように、サイバー攻撃は非常に多様化してきています。マルウェアや詐欺メールの「ばらまき型攻撃」もいまだ数多いなか、攻撃対象を特定した高度な「標的型攻撃」が深刻な脅威となっています。近年のサイバー攻撃の特徴として、その組織化・産業化があります。サイバーキルチェーン（攻撃手順）の初期段階にある「偵察」と「武器化」に関して、闇市場では標的型攻撃に利用される個人情報や各種マルウェアが幅広く、かつ安価で提

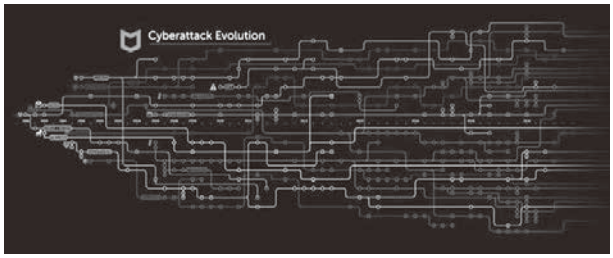
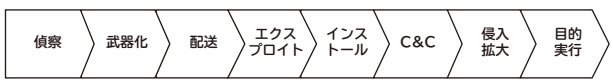


図1　McAfee社によるサイバー攻撃の進化の系譜[1)]



エクスプロイト：セキュリティ脆弱性の悪用
C&C：Command and Control, 外部からの遠隔操作

図2　サイバーキルチェーン

供されていると言われています（**図2**）。

今や、攻撃者は業者やツールを使って短期間に多数のターゲットに標的型攻撃を仕掛けられる状況となりつつあり、「ばらまき型攻撃」と「標的型攻撃」の境界は曖昧になってきています。こうしたサイバー攻撃の激しい変化に、防御側の人材確保や従来の対応速度では間に合わなくなるという認識が、昨今急速な進化を見せるAI技術への期待が高まっている理由の1つと考えられます。

もう1つ、今後は攻撃者もAI技術を駆使してくることが予想され、その対抗手段として防御のためのAI技術が必要となるでしょう。2016年に開催されたDARPAのCyber Grand Challengeでは、参加チーム同士が互いに完全自動でサイバー攻防戦を行い、脆弱性の発見から攻撃、攻撃を防ぐためのパッチ適用がいずれも分単位で行われていました。人間の能力を超えた高速な攻撃には、同様に高速な対抗手段が必要となります。

2.2 防御のためのAI技術の活用

期待されるAI技術をどのように活用すべきでしょうか。既にマルウェア検知には広く利用され、効果が認められています。2017年7月、Googleが運営するマルウェア検査サービスのVirusTotalにCylance社のAI技術が統合されると発表されたことは、その1つの証左と言えるでしょう。また、IBMは、膨大な脅威情報を網羅的に活用する知的検索エンジンとしてのAI技術の有効性をWatson for Cyber Securityで実証しようとしています。

CylanceのCylancePROTECTやNECの自己学習型システム異常検知技術（ASI）は、それぞれマルウェア

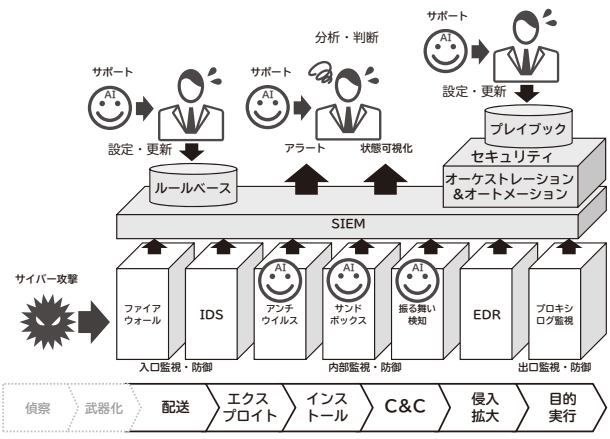


図3　セキュリティ・オペレーションにおけるAI活用

IOC: Indicators of Compromise, 攻撃の痕跡を示す情報

図4　プレイブックの大まかなステップの例

などを検知するセンサーに相当しますが、これらを含む複数のセンサーのデータから総合的に攻撃を判定する手段としてSIEMの導入が進んでいます。現在のSIEMはほぼルールベースで機能しており、その効果的運用にはタイムリーなルール更新が鍵となります。SIEMの運用効率化や総合的判定の支援も、AI活用の機会となるでしょう（**図3**）。

インシデント対応の効率化・迅速化という点では、「セキュリティ・オーケストレーション＆オートメーション（Security Orchestration and Automation、以下SOA）」と呼ばれるタイプのソリューションを提供するスタートアップ企業が、2014年頃から米国で注目を集めています。米Phantom Cyberや米Demisto、イスラエルのHexaditeなどが代表的です。SOAは、プレイブックと呼ばれる防御側の対応手順をあらかじめ定義し、その手順に従って情報収集・解析・対策実行までを自動で行うことができます（**図4**）。ここでも、AIによるプレイブックの適応的な策定・更新が重要になると考えられます。

3. なぜ、機械学習型AIだけでは不十分なのか

3.1 大量の学習データが必要

機械学習には、人手で知識を与えなくてよい替わりに大量の学習用データが必要になります。特にディープラーニングは、他の機械学習方式に比べても多くの学習用データを要します。すなわち、従来と大きく異なる新種のマルウェアなどを検知するには、その特徴を表す大量のデータが集まるまでの時間が必要となります。ターゲットが絞り込まれた標的型攻撃では、十分なデータが収集できない可能性もあります。

3.2 出た結果の理由が分からない

ディープラーニングを含む機械学習型AIが持つ他の特徴として、学習用データのどれとも異なるデータについても判定可能なことが挙げられます。すなわち、マルウェア

の亜種が出てきても、過去のマルウェアとの共通性をとらえて検知できる可能性があります。ただし、検知したものが既知のいずれのマルウェアの特徴とも一致しない場合、なぜそれをマルウェアと判定したのかを、現在の機械学習型AIは人間に分かりやすく説明することができません。

3.3 判別とは異なる種類の技術課題が存在

現在の機械学習は、基本的に何かの判別や分類を得意としているため、マルウェアや異常な状態変化の検知に適しています。しかし、過去のコンテキストに応じた判断をすべき場合や複数の手順を踏んだ対策案を導き出す必要がある場合、従来の機械学習では対応困難です。囲碁で人間を圧倒したAlphaGoも、局面ごとの人の指し手の学習にディープラーニングを用いましたが、自身の指し手を考えるところにはモンテカルロ法という別の技術を使っています。

4. なぜ、将来的にも人間のアナリストが必要なのか

4.1 人間にしかできないこと

前述したように、機械学習型AIは対処手順の計画といった系統だった知的処理が苦手です。社会常識や攻撃者の心理を踏まえた判断は、まだ当分人間の専門家にしかできないでしょう。しかも、法的にも文化的にも最終的に判断の責任を取れるのは人間だけです。また、常時監視している対象以外から情報を取得する必要がある場合、その許可を得るために人間同士の交渉が必要になります。したがって、攻撃の多様化や高速化が進んだとしても、人間の介在が依然必要です。

入口/出口対策としてアラートの真偽に日々アナリストの判断を要していますが、侵入が認められた際には何人ものアナリストが数日から時には1カ月以上も調査と対処に追われ、数千万円ものコストが掛かる場合もあります[2)]。マルウェア侵入の検知率低下が深刻化していると言われるなか、侵入後の攻撃をいち早く検知し隔離・復旧させることが重要です。その一手段として期待のあるSOAでは、プレイブックに沿って情報収集や対策を自動化できますが、プレイブック自体は人間が攻撃の実態に応じて策定しなければなりません。また、IBMやHPEなどのセキュリティ大手は、攻撃手段の多様化・個別化が拡大するなかで既定のプレイブックのみに基づく対応は困難であり、将来もアナリストによる判断が必要だと主張しています。

4.2 なぜ、AIが「なぜ？」に答えなければならないか

従来の機械学習では結果の理由を人間に分かりやすく説明できないと述べましたが、なぜ人はその理由を理解する必要があるのでしょうか。対策を決定し、その判断の責任を取るのは人間です。また、被害や対策によって直接の損失を被る可能性がある部門や顧客に対して、攻撃の発生や取るべき対策の根拠を説明する必要があります。

5. 論理推論型AIの可能性

NECは、セキュリティのインシデント対応のように、状況を周囲のコンテキストとともに大局的にとらえ、人と連携しながら問題を解決するAIの実現を目指しています。そのためには、AIと人がそれぞれ考えていることを互いに理解できる表現方法で共有できなければなりません。また、目の前の問題に対する人の意見を、AIの知識にもその場で取り込んで分析を進められる必要があります。

このようなAIをインシデント対応に適用する場合、SOAで用いられるようなプレイブックを攻撃の状況に応じてオンデマンドで設計し、自動で調査・確認できることはAIが行い、人間の手を借りるべきことはアナリストに適切に依頼する仕組みが必要になると考えています。プレイブックの策定には、攻撃がどのように進行していくかを表すサイバーキルチェーンの具体化がまず必要です。

以上の要件を満たすため、NECは機械学習型AIとはルーツの異なる、論理推論型AIの可能性を検討しています。機械学習型AIが、データの特徴を数学的にモデル化した人間には読めない知識を用いるのに対し、論理推論型AIでは記号や論理式で表された人間にも読める知識を用います。かつて、厳密に定義された知識を用いるエキスパートシステムがその柔軟性の欠如や運用の困難さから下火になりましたが、近年では論理的な説明性と柔軟な推論能力を併せ持つ技術が開発されてきています。

論理推論型AIには、結果を導き出した経緯を人間にも分かる形で提示できる他、推論のための知識を部分的に追加・修正することが容易という利点があります。この技術により、センサーで検知した異常を起点に関連する事実としてのデータを可能な限り集め、サイバー攻撃に関する知識を用いてマルウェアや攻撃者自身によりどのような操作が行われているのかを推定します。最終的に攻撃目的が達成された状態までの流れを導出し、その流れを検証するた

めに、推定した各操作や途中の状態を、得られている事実を用いて確認していきます。概ね攻撃の流れが正しいと判断できたら、今度はその流れを止めて対策が完了するまでの手順を導出し、アナリストと連携しながら検証します。

以上のような技術の確立にはまだ多くの研究開発が必要ですが、簡単なイメージを示すために類例としてセキュリティのCTF（Capture The Flag、旗取りゲーム）問題を論理推論で解いてみた一例を挙げます。**図5**は、通信パケットのデータからフラグ（特定の文字列）を見つけよ、という問題が与えられた場合のフラグ発見までのありうる手順を論理推論で導出した例です（ただし、実際の内部表現はそれぞれプログラムの関数のようになっています）。図2のサイバーキルチェーンや図4のプレイブックの例と同様に、一連の操作や状態からなるものであることが分かると思います。パケットデータからは、テキストファイルが見つかる場合もあれば、フラグが直接見つかるかも知れません。そうした可能性を各種のツールやコマンドを使って自動で検証し、間違っていれば手順を修正して、最終的にフラグが見つかるまで進めていきます。技術的課題としては、さまざまな状態において起きうる可能性を網羅できる十分な知識の構築や、できるだけ少ない検証と修正で正解に到達するための優位な手順選択方式の確立などがあります。

この例のように、推論によって手順の候補を出し、検証を進めて問題を解くアプローチを、攻撃の全体像（サイバーキルチェーン）及び対応手順（プレイブック）の策定にまで発展させたいと考えています。ただし、CTFでは明確な答えがあり、必要な証拠となるデータがすべてそろっていますが、実際のサイバー攻撃では攻撃者しか正解を知らず、必要な証拠がすべてそろうことも期待しづらい点は、考慮しなければなりません。推論のための知識は、IBM Watsonと同様に脅威情報やデータから機械学習を利用して獲得・更新していきますが、前述のように、アナリストの意見からも即座に知識を獲得する技術が重要になると考えています。

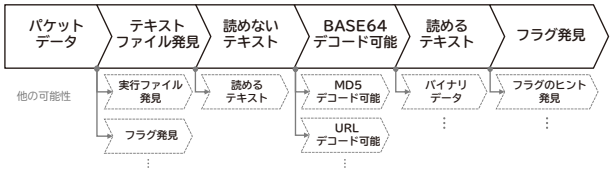


図5　CTFの問題を解く手順の生成イメージ

6.「なぜ」に答え、人と協力し合えるAIへ

NECは、論理推論型AIの研究開発とサイバーセキュリティへの適用を通じて、「なぜ今インシデント対応が必要なのか」という問いには推定した攻撃手口を示し、「なぜその対応が適切なのか」という問いには推定した対策を示し、それらの妥当性については自らの調査・分析とアナリストの判断に基づく説明ができるようなAIの実現を目指しています。これにより、従来の機械学習型AIが支援できていないアナリストによる総合的判断の時間と労力を大幅に削減し、今以上に多様化し頻発するサイバー攻撃にも対応できるソリューションの実現に寄与したいと考えています。

参考文献

1) Brian Dye：Changing Cybersecurity for a New Era, 2017.4.7
https://www.slideshare.net/scoopnewsgroup/brian-dye-changing-cybersecurity-for-a-new-era
2) Ponemon Institute：2016 Cost of Cyber Crime Study & the Risk of Business Innovation, 2016.10
https://www.ponemon.org/local/upload/file/2016%20HPE%20CCC%20GLOBAL%20REPORT%20FINAL%203.pdf

執筆者プロフィール

細見 格
セキュリティ研究所
主任研究員

NEC 技報のご案内

NEC 技報の論文をご覧いただきありがとうございます。
ご興味がありましたら、関連する他の論文もご一読ください。

NEC技報WEBサイトはこちら

NEC技報（日本語）　NEC Technical Journal（英語）

Vol.70 No.2　ビジネスの安全・安心を支えるサイバーセキュリティ特集
～Futureproof Security　安心の先へ。～

ビジネスの安全・安心を支えるサイバーセキュリティ特集によせて
増大するサイバー犯罪の根源的な解決へ～“産学官”のオールジャパンで立ち向かう第三者機関の姿とは？～
サイバーセキュリティを取り巻く社会動向とNECの取り組み

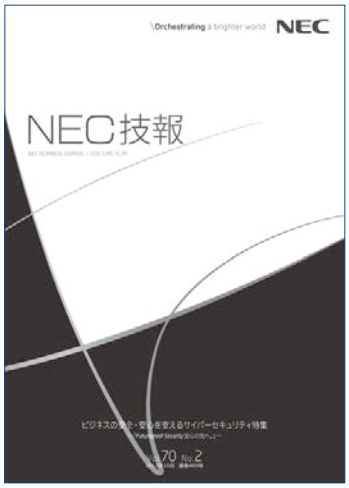
◇ 特集論文

社会の動向とNECの取り組み
重要インフラに対するサイバー攻撃の実態と分析
2017 サイバー攻撃最新動向 NEC Cyber Threat Intelligence活用モデル
サイバーセキュリティ対策の社内事例

サイバーセキュリティソリューション
複雑化 / 高度化するサイバー脅威に対抗するSOCとセキュリティ監視サービス
攻撃被害を極小化するためのインシデント対応支援ソリューション
サイバー演習によるインシデントハンドリング能力の強化
セキュリティ統合管理・対処ソリューション「NEC Cyber Security Platform」
クラウド型ファイル暗号化サービス －ActSecureクラウドセキュアファイルサービス－
セキュリティLCMサービス
EMMを活用したセキュアなモバイルワークソリューション
IoT 時代の経営を支援するサイバーセキュリティコンサルティング

サイバーセキュリティへのAI技術の活用
AI（人工知能）を活用した未知のサイバー攻撃対策
採るべき対策の「なぜ？」に答えるAIの可能性
オープンソースインテリジェンスを活用したサイバー脅威の検出と自動分析
犯罪捜査支援のためのサイバー・フィジカル統合分析技術

お客様に安全・安心を届けるための社内の取り組み
安全・安心な製品・サービスをご提供するための取り組み ～セキュア開発・運用～
サイバーセキュリティ人材のタレントマネジメント



Vol.70 No.2
（2017年10月）

特集TOP