21 世紀のセキュリティに AI が果たす役割
～既存の価値観からの脱却～

AI と機械学習で防御の仕組みを根本から書き換える
予測型サイバーセキュリティを実現

CYLANCE™

エグゼクティブサマリー

21 世紀の世界を定義する重要なキーワードの一つは、「コネクティビティ」です。ソーシャルネットワークやモバイルデバイス、IoT、組み込みシステムなどにより、この 10 年で社会は大きく変わりました。バーチャルにどの場所にも行くことができ、ほぼ誰とでも繋がることができるようになりました。

ところが、目覚ましい進化を遂げたコネクティビティの影で、IT 業界、特にサイバーセキュリティ分野では、かつてないほどの脅威に直面しています。サイバー攻撃が高度化を続け、頻度を増す中で、確実に対抗できる有用なセキュリティの実現が難しくなってきているからです。2000 万もの個人情報漏洩で話題となった米連邦政府人事管理局の事案をはじめ、日本国内も同様で規模や業種を問わず、データ漏えいが日常的に発生しています。メディアでも取り上げられることが多くなり、セキュリティが企業の経営課題として認識されてきています。

また、市場調査会社ガートナーは、2019 年のセキュリティ産業の売上高が 1,000 億ドルを超えると予想しており、企業の皆さんにとってはセキュリティ対策の選択肢が益々増えることでしょう。

最近ではデジタルイノベーションを背景に、人工知能（AI）と機械学習をベースとした新しいテクノロジーのパラダイムがエンドポイント防御の常識、更にはセキュリティという概念そのものを塗り替えつつあります。予測と防御というテクニックを駆使し、悪意のある攻撃を実行前に防御して阻止する そんなことを実現する人工知能は、レガシーな手法から完全に脱却し、実際にマルウェアの検知率という指標でも驚異的な成果を挙げています。エンドポイントに適用された人工知能と機械学習は、99% の有効率で比類なきレベルの防御を実現しているのです。

古い世界観

「問題を作ったときと同レベルの思考で問題を解決することはできない。」

- アルバート・アインシュタイン

アンチウイルス業界はこれまで長きに渡って、事後対応という考え方に基づいてきました。つまり、攻撃を止めるにはシステムがその攻撃を見ておく必要があり、防御には「生贄の子羊」、つまり最初の犠牲者が必要でした。アンチウイルスベンダーも事後対応モデルをベースとしたソリューションを提供してきました。シグネチャベースの検知、悪用防止、ホワイトリスト、アプリケーション制御、エンドポイントの検知と対応など、高度な技術もこの「最初に犠牲者ありき」というモデルに分類されます。

ところがそうした、事後対応的なアプローチは今では業界最大の弱点になっているのです。

見たこと、経験したこと、または知っていることのみに対応するモデルには限界があります。すべては他の誰かにとっては既知のものが「自社にとっては未知」であるゆえにある程度防御するという考え方で規模を問わず多くの企業がこの手法に黙って従ってきました。

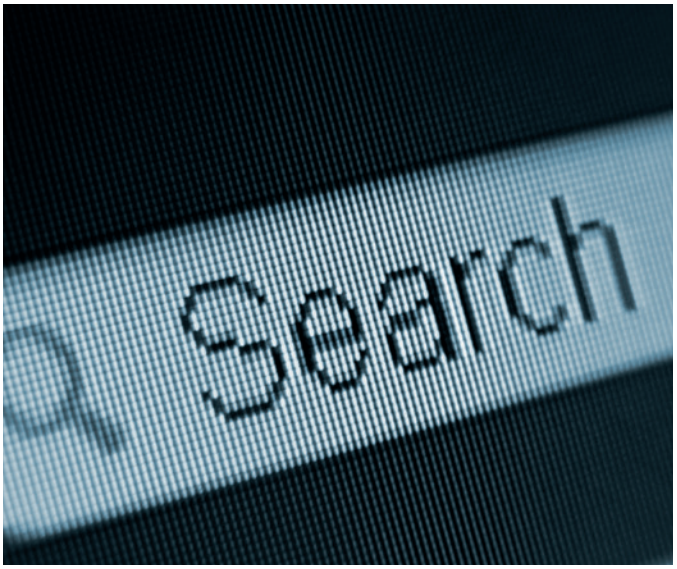
新しい攻撃技術と亜種があまりに多いため、受け身の対策モードから脱却できないまま、できる限り最速で対応できるチーム体制作りに貴重な時間とリソースを注ぎ込んでしまっていたわけです。「起きるかどうかではなく、いつ起きるか」という考え方での成功の判断基準は「復旧」です。

いろいろな意味で、従来のアンチウイルス業界はこの考え方を推進してきました。数十年にわたって多数の新しいソフトウェアやソリューションが開発されてきましたが、まったく新しい概念や考え方を市場にもたらしたものは、つい最近まで皆無でした。

しかし、最近になって脚光を浴びている AI とその数学的部分集合である機械ベースの学習は、サイバーセキュリティの仕組みと考え方を根本から変えつつあります。

長年、金融や医薬品、保険などの業界では問題解決に AI を適用してきました。機械学習と数学を利用した文明の利器には、オンラインコンテンツの提案やドローンの操縦、音声認識などがありますが、最も広く利用されている機械学習の例は、単純なオンライン検索でしょう。検索語の入力を始めると、キーワードや文を予測表示してくれるのが機械学習のアルゴリズムです。

このテクノロジーがセキュリティの分野に適用されてこなかったのはなぜでしょうか？



サイバーセキュリティへの機械学習適用の試みが失敗してきたのには、さまざまな理由があります。まず、データサンプルが不足している、アルゴリズムが曖昧すぎる、コストが高すぎる、というのが主な理由として挙げられます。ただし、ここ数年で状況は変わってきました。IoT とクラウドにより利用可能なデータが爆発的に増えたこと、低コストで高い計算能力が得られるようになったこと、そして高度なアルゴリズムが開発されたこと。それらによりセキュリティへの AI の適用が可能になったのです。これには調査会社も注目しており、Gartner が最近公開した「Top 10 Strategic Technology Trends for 2017: Advanced Machine Learning（戦略的テクノロジトレンドのトップ 10：高度な機械学習）」によれば、現在の機械学習と AI は飛躍的に進化していることが確認されています。

重要度が増すサイバーセキュリティ

一番脆弱なポイントであるエンドポイントに AI で予測防御のアプローチ

ランサムウェア

ランサムウェアの脅威が拡大を続けています。攻撃者はあなたのファイルを暗号化し、身代金を支払うまで、データをアクセス不能な状態にします。多くの場合、ランサムウェアはソーシャルエンジニアリングを利用して足掛かりを得ます。犠牲者の多くは、突然のことに慌て、安易な方法でデータを元に戻そうと身代金を支払う傾向にあります。今後 12 か月でランサムウェア攻撃は更に大幅に増加すると予測されています。

マルウェア

最近のマルウェアは、追跡が難しく、多くの場合はデータを盗み出すようにカスタマイズされています。また、マルウェアは「デジタルの落書き」から標的型攻撃に変化しています。組織犯罪、国家組織、ハクティビストなどが「ダーク Web」上にあるオンラインコミュニティとリソースから利益を得ており、「サービスとしてのサイバー犯罪」が広がることで、ますます複雑化した標的型攻撃が増えてきています。

規制の順守

地域や国、国際的な規則で、情報プライバシーとデータセキュリティの強化の優先度が高くなっています。違反の場合、組織はこれまで以上に重い法的義務と懸案を抱えることになり、複雑化する法規制に準拠するための技術ソリューションを必要としています。たとえば、欧州の一般データ保護規則（General Data Protection Regulation）の下では、個人データの適切な処理と保護に対する企業の責任が増します。新しい規制を遵守できない場合、データを漏洩した企業には数百万ドルの罰金が課せられるか、最大で年間売上高の 4% を失う可能性があります。

重要情報の保護

今日の市場では、機密データ保護のコストは史上かつてないほど増大しています。機密データには、業種や業界に応じて個人情報、会社情報、知的財産（IP）、財務情報、患者情報、クレジットカードデータなどが含まれます。

戦略イニシアチブとしてのサイバーセキュリティ

インテリジェントなサイバーセキュリティ戦略の導入がビジネスに付加価値を与えることを認識する企業が増えてきています。IT 幹部は、セキュリティを強化するだけでなく、企業の使命に組み込み、包括的な計画やプロセス、手順を策定する責任を負っています。同様に、企業の社会的責任が経営層の課題となり、今やサイバーセキュリティは経営課題として扱われています。

新しい考え方

「もし当時欲しいものを尋ねたら、人々は速く走る馬が欲しいと答えていただろう」

- ヘンリー・フォード

業界の優先順位が大きく変わり、セキュリティの重要性が増大する今日、単に個別のツールや技術、ソリューションを組み合わせて最適化しただけで対応していくことはもはや不可能です。業界を再定義するための革新的な考え方が必要です。

そこで登場するのが、機械学習を使用したプロアクティブな予測型防御です。AI、機械学習では、アルゴリズムを使用してパターンを発見するモデルを構築し、このモデルを学習機能で継続的に改良していきます。機械学習を利用することで、人間の能力をはるかに上回るスピードと規模で的確な意思決定を下すことができます。これは、過去の経験に基づく予測が可能だからです。

能動的に防ごうとする考え方の中では、攻撃の対象となりやすいエンドポイントに焦点を当てる必要があります。単体技術のパッチワークをネットワークに追加するのではなく、最前線にいるユーザーに注力するべきだからです。

機械学習では、マルウェアを包括的かつ微細な方法で、コードの DNA レベルで分析して防御することができます。ヒトの DNA の場合は複雑な指揮系統があり、それがブロック（遺伝子）の中で他のブロックと相互作用して、生命体を形成するためのパターンが作られます。機械学習では、遺伝子と同様に相関するコードとファイルの特性のブロックを、人による手動分析とは比較にならない速さと規模で分析することができます。システムやワークフローのパフォーマンスが低下することもありません。

機械学習の驚くべき能力の 1 つが、人による分析とは異なり、マルウェアの構造が解析されると、統計的に類似するコードブロックのビューが分析されて、ファイルを最初に実行せずに悪意のあるコード（つまり「不良遺伝子」）の存在を特定できることです。AI は、観察やパターン認識、予測分析を使用して、悪意のあるファイルを判断できます。これにより、既知、未知を問わず、マルウェアの 99% を回避することが可能になります。

従来のサンドボックス、ヒューリスティック、または振る舞い型（すべて事後対応型の実行後アプローチ）に対して機械学習の価値の理解が深まるにつれて、実行前の予測防御として機械学習に投資する企業が増えています。

機械学習アルゴリズムをエンドポイントに配置すれば、コードが実行される前にファイルに悪意があるかファイルが良性かをすばやく判断できます。クラウド問い合わせ型の分析技法とは対称的に、エンドポイントにはアルゴリズムが継続的に存在するため、クローズドなネットワークにあっても同レベルの防御力を得ることができます。クラウドに日常的に再接続する必要があるシグネチャルックアップなどの技術とは異なり、機械学習アルゴリズムは、オフライン環境下でも高い検知率を維持できます。

トレンドを示す興味深いデータとして、Enterprise Strategy Group の「Enterprise Adoption of Next-Generation Endpoint Security: How Enterprises are Evaluating, Testing, and Deploying Security Products（次世代エンドポイントセキュリティの企業導入：セキュリティ製品を企業が評価、テスト、および導入する方法）」によれば：

- 2020 年には、投資の優先度トップ 5 に AI(人工知能) を含める CIO が 30% を超える
- 2020 年には、CFO は AI によるスマートなデータとアルゴリズム的ビジネスから導出される企業価値に対応する必要がある
- 2018 年末には、AI から得られたデータを顧客向けのセールス、請求、サービスのワークフローに利用する耐久消費財メーカーが 25% に達する
- 2018 年末には、AI の導入に向けた R&D ベースのエンドユーザーアプローチが、IT プロジェクトベースのアプローチの 3 倍以上のビジネス価値を生み出すと予想される
- 2018 年には、「ロボットの上司」に管理される労働者が世界で 300 万人を超えると予想される

新しい働き方：機械学習の仕組み

「数が宇宙を支配する ...」

- ピタゴラス

AI と聞くと、科学というより SF 小説が連想されますが、概念自体は非常に単純です。機械学習は次の 4 つのフェーズに分類されます：

収集

機械学習の第一段階は、できるだけ多くのデータを収集することです。数億本のファイルが複数のソースから集められます。これには、ライブデータフィードや政府機関のデータベース、プロプライエタリなリポジトリに加え、公開された研究や科学的調査などが含まれます。クラウドにデータを収集して保存するだけでなく、モバイルや IoT、組み込みシステムからデータを抽出することも可能であるため、数年前までは考えられなかったような新しい水準でのデータ収集が実現します。これらすべてのソースからデータを拾い集めることで、極めて広範なファイルタイプと作成者を表す妥当なサイズのサンプルを入手できます。

抽出

特徴の抽出プロセスでは、1 つのファイルを数百万ものさまざまな特性に分解します。各属性は、他のファイルから導出された数百万の特性と比較して分析されます。このようにして、収集フェーズで集められた数百万のレコードが数百万の変数に個々に分解されてから、ベクトル化（Vectorization）されます。

機械学習は、一種の自律的なネットワークまたはエコシステムとして考えることができます。

学ぶには、観察が必要だ。

観察には、何を探すべきかを知る必要がある。

何を探すべきかを知るには、事前に学習している必要がある。

これは一見すると、循環論法のように見えますが、まずパズルのピースをバラバラにしてから、抽出と解析でパズルを組み立てていく、というインテリジェンスの古典的なやり方なのです。

抽出は、科学とエンジニアリングをプロセスに適用し、それを数百万倍に拡大するという役割を果たします。最終的に、AI を訓練する観察結果が生み出されます。この学習されたパターンを通じて、AI は新しいファイルが良性か悪意かを判断することができるわけです。

学習

データを収集し、各ファイルから特徴を抽出して得られた数百万の属性を、学習プロセスに利用することができます。属性はベクトルという形式で数値に変換され、モデルの訓練に使用されます。予測の正確性を確保するために、測定値を使用していくつものモデルが作成されます。また、テストプロセスでも無効なモデルを特定することができます。モデルのテストと検証には、膨大な数のファイルが使用されます。テストされ、改良され、アクションが可能になると、最終モデルが読み込まれます。

学習プロセスでファイルとファイル属性が解析されていく中で、モデルはサンプルの意図についての理解を積み上げていきます。この理解を予測的手法に適用すれば、新しいファイルを実行することなく、そのファイルの潜在的なリスクを判断することができます。

分類

より高度な統計モデルを構築することで、高度にチューニングされた分類とクラスタリングが可能になります。機械学習の要は、コンテンツの正確なカテゴリ分類に尽きるからです。エンドポイント防御の場合、悪意か、良性かという分類になります。また、分類では、たとえば、ファイルの目的がキーロガーなのか、トロイの木馬などかなど、目的に基づいてファイルが整理されます。

AI は、人の目から見ると無害に思われたり、見逃してしまったりするような、わずかな統計的な関連性を検知できます。この分析は数ミリ秒で完了します。また、ファイルとファイル特性が幅広く分析されるため、極めて正確です。

分析では、分類プロセスの一部として「スコア」が提供されます。このスコアによって追加の洞察が得られ、それを 1 つのファイルに対する判断の重み付けに使用できます。たとえば、ファイルをブロックするか、隔離するか、監視するか、またはさらに分析するかについて判断することができます。

機械学習アプローチと従来の脅威調査アプローチの間には重要な違いがあります。

機械学習の場合は、モデルを構築して、ファイルが悪性か良性かを判断するか、スコアに基づいてファイルを疑わしいものとして特定します。悪意あるファイルと良性のファイル（そして、さらなる調査が必要になるファイル）を特定することで、AI の強力なメリットを得ながら、何らかの手動分析が必要と考えられるわずかな異常値を隔離することも可能になります。つまり、プロセスを自動化、加速化、および劇的に向上する強力なテクノロジーに加えて、必要に応じて人間の専門知識を適用するという選択肢が与えられるわけで、両方の長所をうまく利用できます。セキュリティスタッフは重大度が異なる数千もの警告への対応から解放され、その専門知識を他のアクティビティに投入できる、というメリットが得られるのです。加えて、従来のアンチウイルス方式に影響する可能性のある偏りをプロセスから取り除くこともできます。

未来の新しい方法

「アイデアの価値はそれを実際に使うことにある。」

- トーマス・エジソン

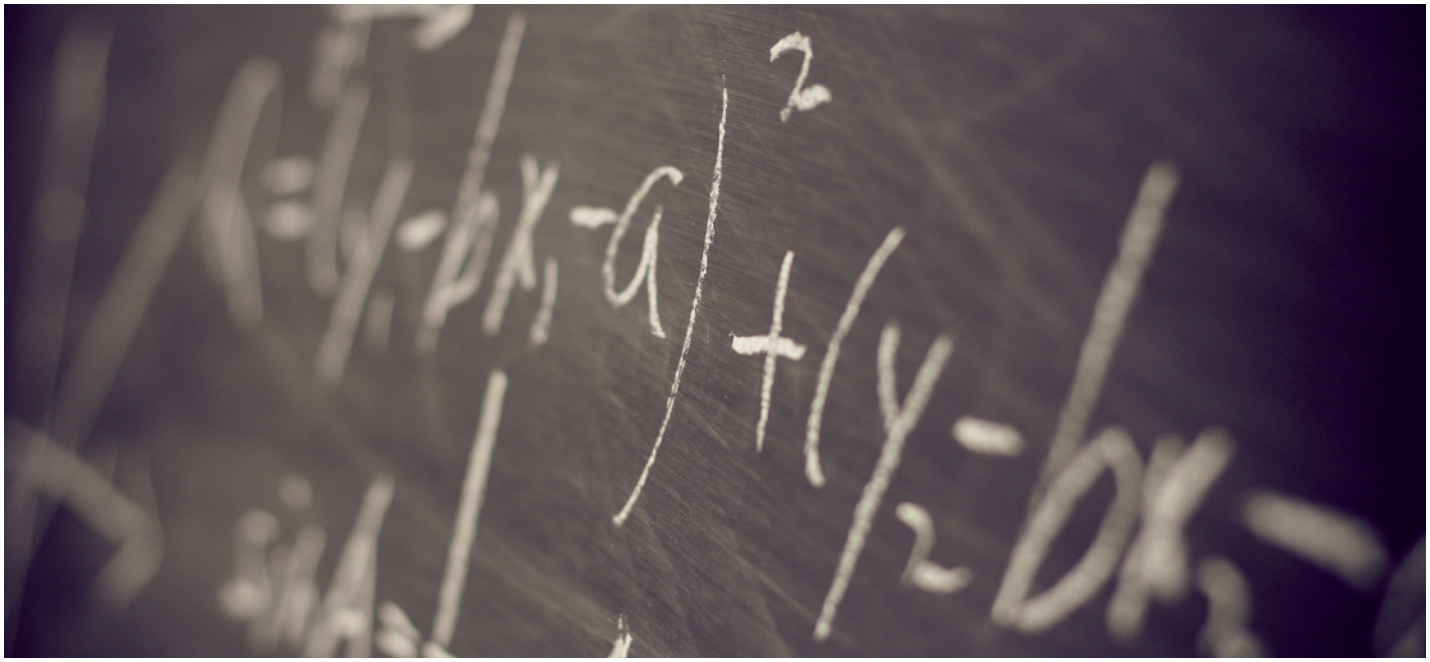
AI と機械学習によって、サイバーセキュリティの未来への扉が開かれました。従来のアンチウイルス、サンドボックス、封じ込め、実行後の対応、クリーンアップはもう過去のものです。エンドポイント防御戦略を適用することで、次のようなメリットが得られます。

シンプル

組織は、わずかなリソースとインフラストラクチャで対応できる数学をベースとしたアプローチを使用して、社内すべてのエンドポイントを防御できます。従来のアンチウイルスソリューションやエンドポイント検知ソリューションとは異なり、シグネチャー更新やスキャンスケジュールが不要なため、AI アンチウイルスにはクラウドサービスへの接続が必須ではありません。管理者は、インターネットへの接続や、シグネチャーファイルの頻繁な更新が不要な環境でも脅威の予測防御ができます。

静寂性

過去の技術で次々と生成されていた警告が大幅に減少し、組織のセキュリティ運用の全体的な効率が向上して、ユーザーへの影



響も軽減されます。さらに、アルゴリズム的なエンドポイントセキュリティを使用した場合、メモリや CPU はほとんど消費されません。

サイバーセキュリティのアプローチを実行前防御のアプローチに変更すると、継続的な対応に必要だったテクノロジーやリソース、人員を段階的に減らしていくことができます。つまり、コストが大幅に低減されます。

拡張性

AI で使用されるプロアクティブな予測型の予防技術は、すべてのプラットフォーム、オペレーティングシステム、ファイルタイプ、デバイスに適用できます。また、Microsoft Windows と Mac OS 、Linux を保護し、既存の SIEM プラットフォームと容易に統合することも可能です。ソリューションは、テクノロジーパートナーの OEM や組み込みバージョンでも利用できるため、ファイルまたはメモリベースの攻撃、悪意のあるドキュメント、ゼロデイマルウェア、権限昇格、スクリプト攻撃に対応でき、潜在的に望ましくないプログラムなどの脅威を防御できる柔軟なセキュリティソリューションを実現することができます。

結論

AI、機械学習、数学は、エンドポイント防御のパラダイムシフトを実現する強力な組み合わせです。これまでの事後対応型アプローチは、プロアクティブな新しいアプローチに全面的に置き換えられ、予防が防御の手段となります。

以前のパラダイムでは、問題の核の部分を解決できませんでした。いくつものセキュリティ対策を多層に積み重ねたとしても、どこかのポイントで「生贄の子羊」がなければ未知の脅威を検出することは不可能でした。

規模に関係なく、組織がエンドポイントで実行を制御できれば、真の意味での予防策となります。進化した AI のおかげで、以前には想像もできなかった成功率で予測型のエンドポイントの防御が可能になったのです。

少ない人員でも効果的にエンドポイントを防御できるため、IT やセキュリティ運用部門は人材をより有益なポジションに投入して、最終的にコストの大幅な削減を実現することができます。

サイバーセキュリティ業界の多くのベンダーは自社製品の技術やメリットを強調するかも知れませんが、今こそ数学と科学に信頼を託すという姿勢に踏み出すときです。

百聞は一見にしかず

自社の環境でテストし、有効性を確認ください。

デモのお申し込み、お問い合わせはこちらからどうぞ。

www.cylance.co.jp

Cylance Japan 株式会社
〒100-6510
東京都千代田区丸の内 1-5-1　新丸の内ビルディング 10F
www.cylance.co.jp
03-6386-0061 (代表)

©2018 Cylance Inc. Cylance®および CylancePROTECT®、ならびに関連するログはすべて Cylance Inc.の商標または登録商標です。その他の登録商法または商標の所有権はそれぞれの所有者に帰属します。

201709-WP001