SECURITY REPORT

2022
MAR

セキュリティ監視知能化の動向及び展望

CyberFortress

このレポートは、株式会社サイバーフォートレスの独自の調査にて収集されたデータに基づき作成しています。

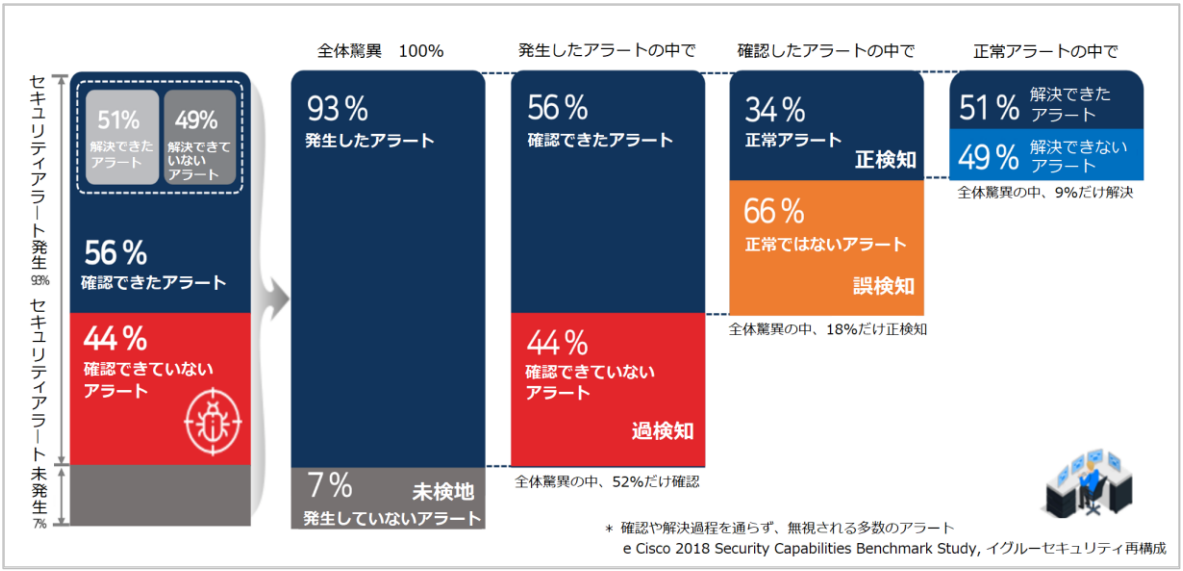
セキュリティ監視知能化の動向及び展望

01. サイバー現況及びトレンドの変化

大勢の人たちの命を奪った新型コロナウィルス感染症(COVID-19)パンデミックで、既存の社会の体系を揺るがす大きな変化が起きている。サイバー環境もその影響を受けている。在宅・リモート勤務及びソーシャルディスタンスの確保が一般的になりITインフラ使用とネットワークのトラフィックが増加していることによって、攻撃者が狙える攻撃対象はさらに広がった。

急激なDXに合わせて巧妙なサイバー攻撃が増加し、膨大なセキュリティアラートを迅速に処理すべきのセキュリティ監視担当者はさらに苦労が増えている。新たな形態のセキュリティ脅威を全て解決できるスーパーマンのような存在が出現しない限り、幾何級数的に生成されるセキュリティイベントに対応するには足りない状況である。

実際、グローバルセキュリティ企業シスコ(Cisco)の資料によると、セキュリティ監視センターに集まるセキュリティ機器から発生したアラートの中で確認及び処理過程を通っていないアラートは多く、処理過程を通したとしても誤検知(False Positives)は少なくない数、発生していることが確認できる。



【▲ セキュリティ監視センター(SOC)から発生したアラートの現況(2008) (参考：シスコ)】

セキュリティ監視知能化の動向及び展望

このようなサイバー環境の難しさを解決するためには、なにが一番必要なのか。CES 2021とRSAC 2021、そして市場調査機関であるガートナーから共通して示している要素がある。それは外部の力によって変形された状態から元に戻ろうとする「回復力(Resilience)」である。現在のサイバーセキュリティ体系にはいつでも攻撃と侵入の可能性があるため、このような危機状況から一早く回復できる力が必ず必要である。

Rohit Ghai RSA CEOは2021年RSAC 2021の基調演説から「現在のサイバー環境では回復力(Resilience)が必要である。転んだ時、再び立ち上がるだけでは十分ではない。よく回復されるためにはよく転んで転んだ時にこれを耐えて、毎度さらに強くなった状態で立ち上げなければならない。」と回復力の重要性を強調したことがある。

それでは回復力(Resilience)を強化するためには、なにが必要なのだろうか。様々な解決策があるが、その中で意思決定のための有用な情報を意味する「インテリジェンス」の重要性が強調されている。既知もしくは未知の脅威が発生した時、よく検知して分析し、対応するためには各段階ごとに正しい意思決定が必ず要求されるためである。

02. セキュリティ監視にインテリジェンスを実現するために必要な3つの方法

セキュリティ脅威が知能化され、セキュリティ脅威と検知されるセキュリティイベントの量が急増することで、これまでよりセキュリティ監視インテリジェンスの重要性が話題になっている。以前からサイバーセキュリティ監視分野のインテリジェンスは様々な方法で適用されていたが、本文では大きく▲データ拡張による知能を加える方法、▲マシンラーニングによる監視者経験を知能化する方法、▲セキュリティオーケストレーション及び自動化・対応(SOAR)による監視プロセスを知能化する方法、この3つで区分し、各方法について説明する。

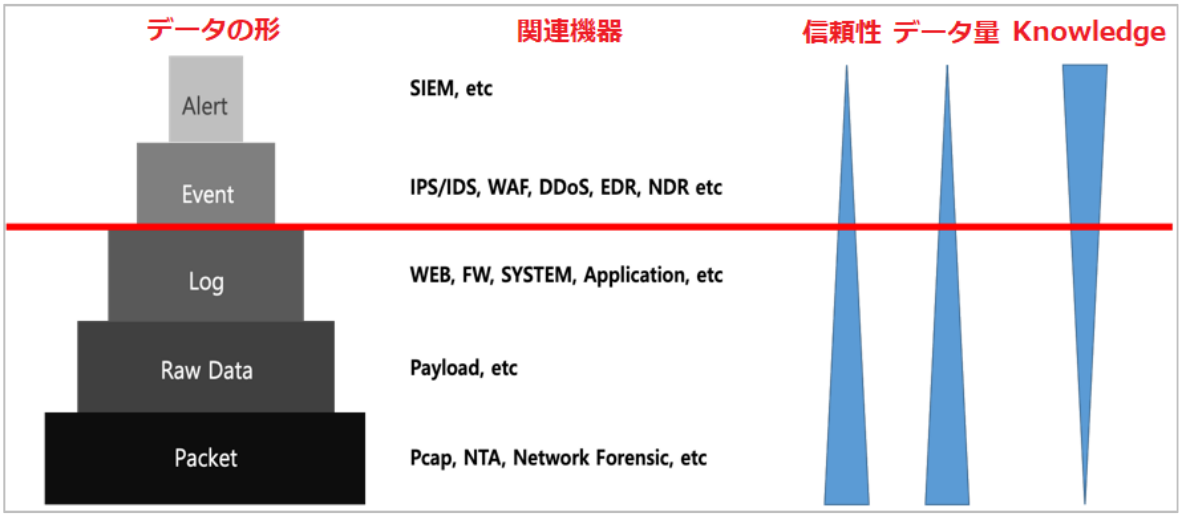
1) データ拡張で知能を加える(eXtended)

サイバー環境に侵入を検知することはネットワークパケットから悪意的な行為を認知し、これをイベント化することと定義できる。以前からセキュリティ監視の分野では悪意ある行為を検知するために多くの努力をしてきた。それはデータを取集することから始まり、検知システム及び分析ソリューションの高度化でより多くのデータ種類と連関づけられ、より進化した分析ができるようになった。

セキュリティ監視知能化の動向及び展望

1-1) より多く収集する

ガートナーは「eXtended」を「拡張する」または「全て」の意味で定義している。つまり、「データの拡張」の概念で解釈できる。使用されるデータの量も重要だが、データの変数を通じて質を向上することもとても重要である。特にマシンラーニングのような目的で使用される場合は、良質なデータを用意しなければならない。このような背景からネットワーク検知及び対応(NDR)、端末検知及び対応(EDR)などを活用しネットワークと使用者エンドポイント領域からの収集及び分析とデータ領域が拡大されている。



【▲ データの形にようる分類及び特徴】

上記の図のようにセキュリテデータの形を五段階で分ける場合、セキュリティ監視専門家は、どのデータを確認するだろうか。一般的にアラート(Alert)またはイベント(Event)を確認し、詳細分析や事後分析を行う場合はログ(Log)、原本データ(Raw Data)、パケット(Packet)などを詳しく確認するだろう。

図のように上の段階に上がるほど信頼性とデータの量は急減するが、そのデータが持っている意味(Knowledge)は高くなっていることが確認できる。一方、下の段階に下がるほどデータが収集される機器とデータのタイプは多くなる。下の段階のデータを確認すると異常行為を見逃すことが最小化できる。しかし、意味ないデータ(Garbage)が多かったり、膨大なデータを処理するシステムを整えてなければむしろデータの中から確認すべきものを見逃すことも起きる。

セキュリティ監視知能化の動向及び展望

1-2) 脅威インテリジェンス(外部の脅威情報も取集する)

データはネットワークを通すトラフィックだけではなく、外部から発生する脅威情報である「脅威インテリジェンス(Threat Intelligence)」または「サイバー脅威インテリジェンス(Cyber Threat intelligence)」まで拡張できる。ガートナーは「現在存在したり発生しうる脅威に対応するための決定のために当該脅威に対する脈絡(Context)、メカニズム、指標、予想結果及び実行可能な条件などを含む証拠基盤の知識」で脅威インテリジェンスを定義している。

セキュリティ監視分野ではどのように脅威インテリジェンスを活用しているだろうか。以前はどれだけ多くの情報を収集するかを中点にしていた。しかし今は各サイトに付合する脅威をどれだけ正確に収集するか、そしてこれを自動的にシステムに適用して活用できるのかに重点を置いている。つまり「実行可能なインテリジェンス(Actionable intelligence)」の重要性が高まっていると言える。多様な収集先 、多様なデータの形式で共有される脅威インテリジェンスにどのような重要情報が含まれているかによってセキュリティ担当者の意思決定が違うためである。

1-3) 内部資産情報及び脆弱性を収集する

ネットワークトラフィック、外部脅威情報と共に内部資産情報もとても重要である。保護すべき資産に対して、より詳細で正確な情報の収集がされていれば、より迅速な判断及び対応が行える。コンピューターの歴史上、最悪の脆弱性とも呼ばれている「Log4j」を活用した攻撃も資産に対する情報把握か不足していたせいでさらに混乱が発生したことがわかる。

2) マシンラーニングによる監視者経験を知能化する方法(Machine Learning)

今までセキュリティ監視の分野の中で一次分析及び詳細分析に多くの時間と人力を使っていた。それでも、目まぐるしく進化している攻撃方法に追いつくには限界があった。監視者ごとに経験及び知識のレベルに差があり、限定された数の監視者が判断するには、とても多いイベントが発生しているためである。このような問題を解決するためにマシンラーニングで監視者の経験をモデリングしようとする試みは続いている。

セキュリティ監視知能化の動向及び展望

2-1) マシンラーニングを介した分析

経験をモデリングするためには、良質の学習データとセキュリティに最適化されたアルゴリズムが要求される。マシンラーニングはセキュリティ監視者・分析家者の経験がデータ化された学習データを利用してセキュリティを学習する。

セキュリティ分析者は検知パターン、インシデント履歴、対象の脆弱性現況、個人的な経験などを活用して不正侵入防止システム(IPS)などのセキュリティ機器から検知されたセキュリティイベントを析してきた。下記の図はIPSから発生したイベントの中のペイロードの一部分である。マシンラーニングはこのような特徴とセキュリティ分析者の経験を基に判断をする。

【▲ セキュリティ監視時、判断をするために分析するセキュリティイベントのペイロードの例】

セキュリティ分析者は「夜明けに発生したイベントなのか」、「脅威国家からアクセスしたユーザーなのか」、「攻撃に使用されたパターンが多いのか」などをインシデント特徴と統計的な特徴を導き出し、これをマシンラーニングが理解できるように前処理してデータ化している。マシンラーニングの分野からはこれを「フィーチャ(Feature)」と定義している。一つのセキュリティ機器から提供されるドメイン情報を基に情報を組み合わせて攻撃の特徴を抽出する概念である。

【▲ フィーチャ定義及び前処理されたデータの例】

セキュリティ監視知能化の動向及び展望

前処理データは大きく数字型とテキスト型で分類されるが、マシンラーニングはこれを基に正検知と誤検知を区別するモデルを作る。

このような教師あり学習と共に多様なセキュリティイベントとログなどの相関分析でセキュリティ機器から検知されない異常行為を見つけ出す教師なし学習も行われている。

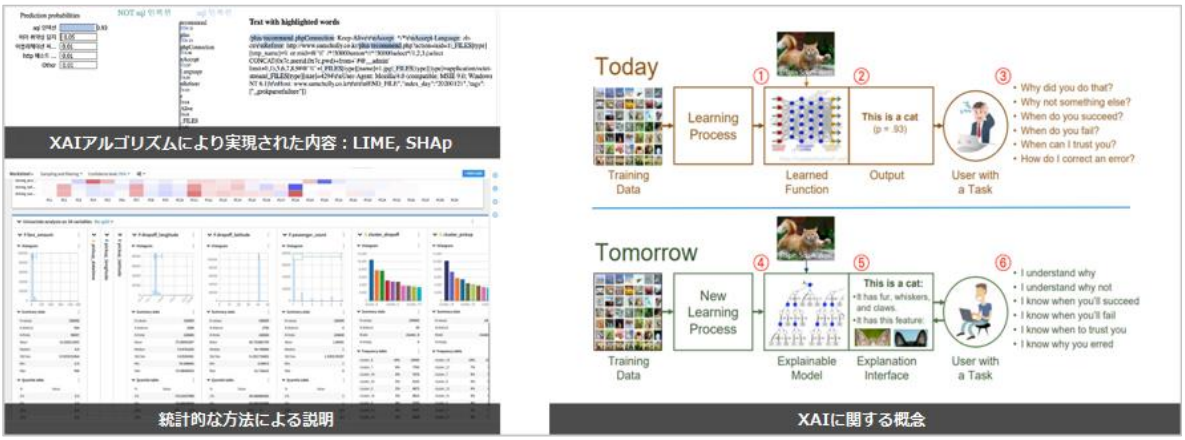
しかし、上記のようにセキュリティ分析者の経験をマシンラーニングにうまく適用するためには必ず考慮すべきの事がある。まず、分析者の経験を知能化する明確な目的が必要である。つまり、明確な目的からデータサイエンス方法論を通じてモデルを作るとき、始めて意味あるモデルが作られる。

また、このようなモデルをさらに精巧に標準化するために学習データが要求される。セキュリティ分析者がスノートルール(Snort Rule)などの各種ハッキングパターンを基に分析していた以前とは違って、マシンラーニングモデルは学習データを基盤に自ら認知し、判断基準を作って分析すると言える。インシデントに対する学習データを、実データの基準で分析し特徴を表すことである。

しかし、マシンラーニングで学習する良質の学習データを作るのは簡単ではないため、長い間セキュリティデータを分析し攻撃者と戦った経験と知識が要求される。

2-2) セキュリティ監視専門家とマシンラーニングの差(gap)を減らす

まセキュリティ監視専門家の観点とマシンラーニングとの差(gap)を減らす必要はある。今までのマシンラーニングから導き出された予測値は分析者の観点からみると、少なくない差があることが確認できる。これを解決するためにマシンラーニングのモデルがイベントをどのように正検知または誤検知として判断したのか、その根拠を提供する「説明可能なAI(eXplainable Artificial Intelligence, XAI)」技術が適用されている。これでマシンラーニングの予測過程を理解し、マシンラーニングのモデルから導き出された結果の改善ができる。



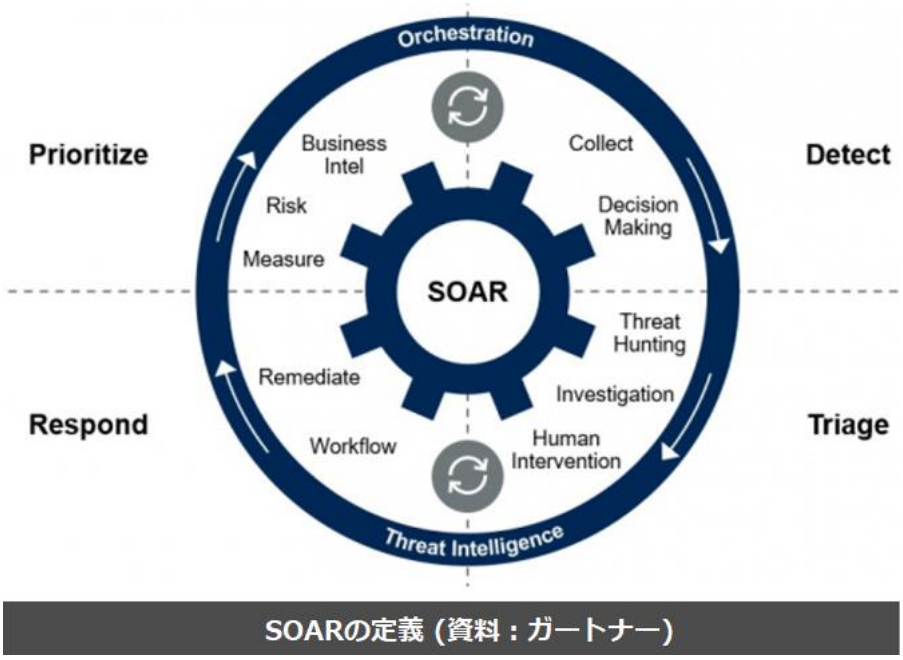
【▲ 説明可能なAI(XAI)】

セキュリティ監視知能化の動向及び展望

3) セキュリティオーケストレーション及び自動化・対応(SOAR)による監視プロセスを知能化する方法(Orchestration)

この方法は最近「セキュリティオーケストレーション及び自動化・対応Security Orchestration, Automation & Response, SOAR)」で実現されている。多様なセキュリティ状況ごとに対応のためのシナリオと手順をデータ化した「プレイブック(Playbook)」を基に単純な業務は自動処理できるようにする形である。

ガートナーは「多様なサイバー脅威と関連して対応レベルを自動で分類し、標準化された業務プロセスに従ってセキュリティ業務担当者とソリューションが有機的に協力できるようにサポートするプラットフォーム」でSOARを定義している。SOARの適用でセキュリティ専門家は単純に繰り返す業務から離れ、セキュリティ専門家の判断が必ず必要で複雑な業務に集中できる時間が得られる。



SOARの定義 (資料：ガートナー)

【▲ SOARの定義 (参考：ガートナー)】

但し、プロセスに知能を加えるためには必ず複雑なセキュリティ状況ごとの明確な業務プロセス化が行わなければならない。また、RestFul APIと脅威情報の共有規格(STIX/TAXII)などを基盤した製品間の緊密な連携が必要である。

また、データとプロセスに「サイバーキールチェーン(Cyber Kill Chain)」、「マイターアタック(MITRE ATT&CK)」、「ディフェンド(D3FEND)」などの攻撃戦術及び技術が加えれば正しい攻撃及び防御の判断をするのに役に立つだろう。最近マイターアタックマトリックスをセキュリティ監視に適用しようとしているが、まだ関連イベントと件数をマッピングするぐらいのレベルになっている。これから最新攻撃戦術及び技術をセキュリティ業務により活用するためのシステムとソリューションの導入が増えると予想される。

セキュリティ監視知能化の動向及び展望

03. 結論及び展望

2000年代前半から約20年に渡って、セキュリティ分野は絶え間ない発展をしてきた。サイバー攻撃を迅速に検知、分析して対応するために多くのデータが収集され、セキュリティ監視専門家の分析を経て、定型化されたプロセスによる対応が行われていた。そしてさらにセキュリティ監視に「インテリジェンス」、すなわち知能を加えるため、さらに改善が行われる時だである。セキュリティ監視の知能化によって得られる期待効果、知能化実現のための必須要件、今後の展望を以下にまとめる。

・ セキュリティ監視専門家はデータ拡張により、侵入に対するより明確な判断ができるようになる。つまり、検知したイベントの誤検知率と未検知率が下がるだろう。しかし、これを検知するイベントが減っている(サイバー攻撃が減っている)と解釈してはいけない。これまで確認できなかったイベントが検知できることによって、セキュリティ専門家が処理すべきのイベントはさらに多くなるかもしれないが、セキュリティの死角は明らかに減るだろう。

・ 分析者の経験を知能化してマシンラーニングに適用することで、フィルターや相関分析、統計分析を通じて検知および分析を行ったこれまでの方法に比べてより迅速かつ正確な結果を得ることができるようになる。但し、このような効果を極大化するためには、各分析方法についての明確な理解が行われ、お互いのメリットが緊密に結ばれなければならない。また、マシンラーニングを適用するためにはドメインの情報、最適なアルゴリズム、良質の学習データが必要となる。マシンラーニングの導出過程を明確に説明できる「説明可能な人工知能(XAI)」も必要な機能の一つである。

・ プロセスを知能化するためのプロセスが確立される必要がある。何を自動化すべきか分からない組織では、自動化の導入はかえってもう一つの業務を発生させかねない。アラートの受付から自動遮断までの一連のプロセスが水が流れるようにうまくいくためには、セキュリティ業務についての明確な定義、脅威状況から発生しうる全てのをプロセス化したセキュリティ監視方法論が裏付けられる必要がある。

最後に協業と連携の重要性はいくら強調しても足りないだろう。これを1904年に発生したボルチモア大火にたとえてみることができる。迅速な火災認知により、十分な消防隊員と装備の支援が行われ、完全に近い対応プロセスが設けられていたにもかかわらず、多くの人命被害と甚大な経済的被害が発生した。原因は何だろうか。給水が供給される消火栓の規格が違ったためである。セキュリティ分野も同じことである。セキュリティ監視の知能化を実現するためには、各構成要素とシステム、プロセス間の連携及び関連専門家間の緊密な協業が必要である。

これまでセキュリティ監視を知能化するための3つの方法について調べてみた。データ、マシンラーニング、プロセス自動化、全て急激に発展をしている。これらの要素間の緊密な連携による真の意味でのセキュリティ監視知能化が行われれば、日々高度化するサイバー脅威にもより柔軟に対応できるのではないだろうか。