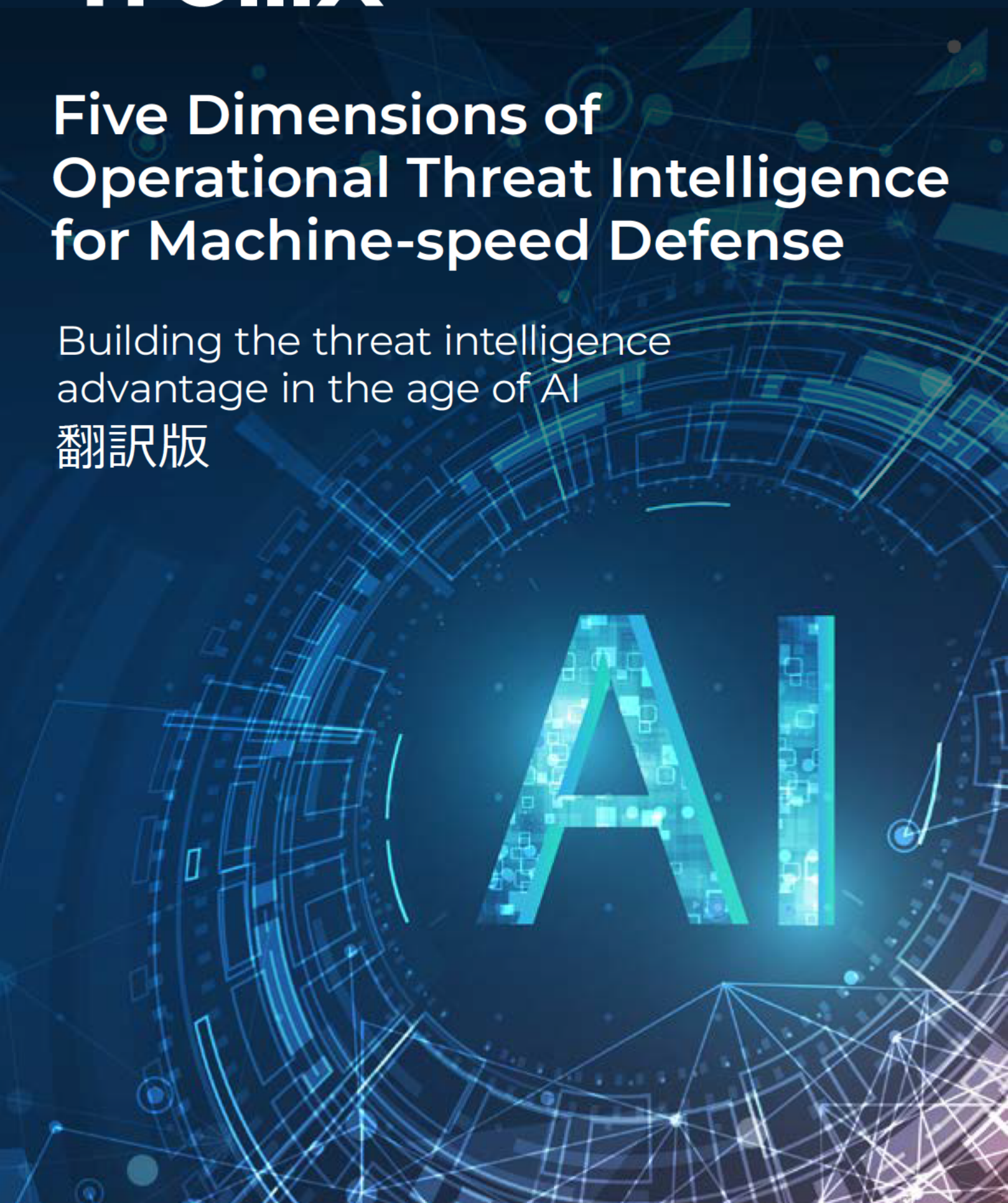
Trellix

Five Dimensions of Operational Threat Intelligence for Machine-speed Defense

Building the threat intelligence advantage in the age of AI

翻訳版



人工知能(AI)は、サイバーセキュリティの運用スピードを根本から変えつつあります。防御側はテレメトリの解釈、調査の加速、脅威ハンティングの支援、そして以前はほぼ人間の専門知識に依存していた分析作業の自動化のために、AIの活用を益々推進しています。同時に攻撃者も、攻撃手法の開発加速、バリエーションの増加、攻撃ライフサイクルの一部の自動化、さらには高度な攻撃の実行に必要な専門知識の要件を下げるために、同様の機能の多くを応用しています。

その結果もたらされるのは、単にサイバーセキュリティの自動化が進むということだけではありません。脅威を理解し、対応するために残された時間が急速に縮小しているのです。

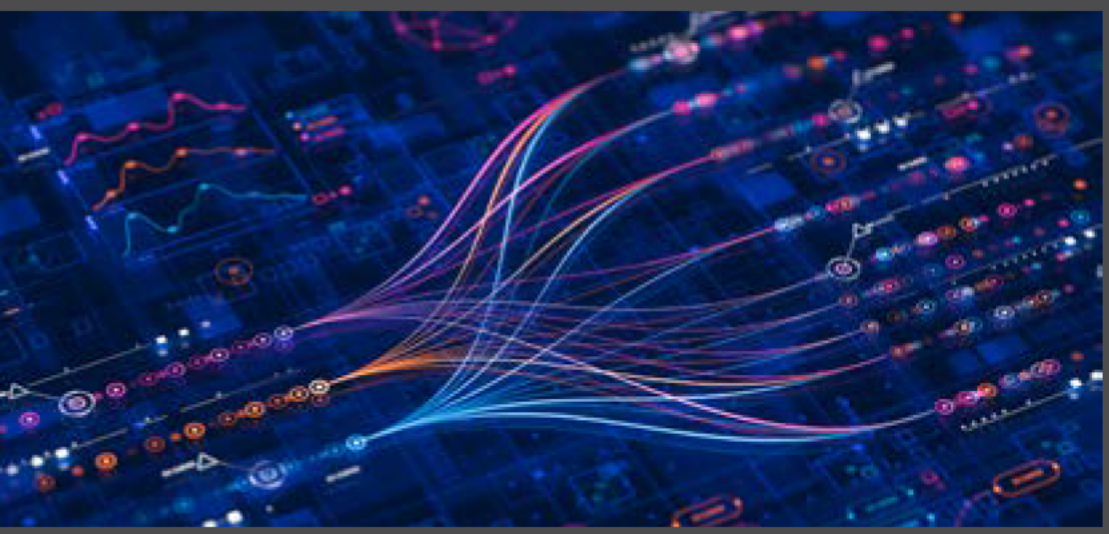
セキュリティ組織にとって、これはマシン・スピードでの検知やマシン・スピードでの修復にとどまらない新たな要件を生み出しています。効果的な防御には、「マシン・スピードでの理解」がこれまで以上に不可欠となります。つまり、何が起きているのかを解釈し、単一の観測結果をより広範な攻撃者の活動と関連付け、自組織にとってどのような意味を持ち得るかを判断し、次に何をすべきかを特定する能力です。

この「理解」を支えているのが、脅威インテリジェンスです。

しかし、脅威インテリジェンス業界はこれまで、人間のアナリストや従来のセキュリティ技術のために設計されたワークフローを中心に構築されてきました。脅威インテリジェンスが、とりわけ規制対象の環境、ソブリン環境、さらには高度なセキュリティ保証が求められる環境において、AI主導のセキュリティのための実質的な基盤となるためには、運用インテリジェンスそのものが進化しなければなりません。

インテリジェンスは、構造化され、コンテキスト(文脈)を備え、マシンリーダブル(機械判読可能)であり、調査や防御アクションを導く能力を持ち、かつAIが動作しているセキュリティ境界内で利用可能である必要があります。

これら5つの次元は全体として、AI対応型運用脅威インテリジェンスが今後どうあるべきかを考える上で有用なフレームワークを提供します。



Part I: 課題—セキュリティはマシン・スピードへと移行

進化への圧力は、脅威情勢から始まります。

AIは、ソフトウェアの分析、悪意あるツールの改変、バリエーションの生成、偵察の自動化、および攻撃テクニックの適応に必要な時間を短縮しています。かつて専門的な知識や手作業に大きく依存していた機能は、ますます加速され、あるいは部分的に自動化できるようになりつつあります。サイバー犯罪におけるAIの兵器化と商品化(weaponization and commoditization of AI in cybercrime)に関する最近のTrellixのリサーチは、これらの機能が実験段階から犯罪エコシステム内での実用化へと移行し始めている状況を浮き彫りにしています。

AIアシストによる脆弱性調査も、同様の動態をもたらします。システムの脆弱性特定能力、悪用可能性(エクスプロイタビリティ)の分析能力、そして修復および攻撃的手法の研究の双方を支援する能力が高まるにつれ、脆弱性の発見から本格的な悪用に至るまでの従来の期間は、今後も縮小し続ける可能性が高いと考えられます。

防御側にとって、これが意味することは非常に重要です。より迅速にパッチを適用することは引き続き不可欠ですが、パッチを適用するだけでは、リスクが解消される前に攻撃者がその脆弱性を悪用したかどうかという問いには答えられません。悪意のあるプロセスをブロックしたとしても、攻撃者が事前に認証情報を取得していたかどうかは判明しません。疑わしいネットワーク接続に対処したとしても、そのアクティビティが単発的な事象だったのか、それともより広範な侵入行為の一部だったのかは判断できません。

したがって、根本的なセキュリティの問いは、個々のアラートが正しく処理されたかどうかにとどまらず、より大きなものとなります。組織は、そのイベントが何を意味しているのか、既知の攻撃者の行動とどのように関連している可能性があるのか、他に何が起きた可能性があるのか、そして他にどのような証拠を探すべきなのかを理解する必要があります。

これこそが、「マシン・スピードでの対応」と「マシン・スピードでの理解」の違いです。

そしてこれこそが、AIの導入が進むにつれて脅威インテリジェンスがより重要になる理由でもあります。

脅威インテリジェンスの「失われた中間層」

脅威インテリジェンスは従来、戦術的(タクティカル)、運用的(オペレーショナル)、戦略的(ストラテジック)の各層にわたって機能してきました。

戦術的インテリジェンスは、悪意のあるファイル、インフラストラクチャ、IPアドレス、ドメイン、URL、その他の観測可能なデータ(オブザーバブル)に関する、即座に対応可能な情報を提供します。迅速な特定、ブロック、検知、情報付加を支援し、サイバー防御において今なお不可欠な要素であり続けています。



戦略的(ストラテジック)インテリジェンスは、より広範な視野を提供します。これは、どの攻撃者を重視すべきか、地政学的および技術的な進展が脅威環境にどのような影響を与え得るか、どの業界や地域が標的にされているか、そして組織のリスクがどのように変化しているかをセキュリティリーダーが理解するのに役立ちます。

これらの層の間に位置するのが、運用(オペレーショナル)インテリジェンス、すなわち攻撃者が実際にどのように攻撃を実行しているかに関する知識です。

運用的インテリジェンスは、脅威アクターをキャンペーン、ツール、テクニック、動作、インフラストラクチャ、脆弱性、標的選定パターン、過去に観測された攻撃チェーンと結びつけます。これにより防御側は、「何かが疑わしい」と知る段階から、そのアクティビティが何を意味している可能性があるのか、そして攻撃者が次に何を試みる可能性があるのかを理解する段階へと移行することができます。

業界には、この種の知識が膨大に蓄積されています。リサーチャーはキャンペーンがどのように進化するかを理解し、マルウェアアナリストはツールがどのように使用されるかを理解し、脅威ハンターは繰り返し発生する動作を認識し、インシデントレスポンダーは侵入時にどのようなアクティビティが併発しやすいかを理解しています。

歴史的に課題となってきたのは、その集積された知識を、セキュリティシステムが一貫して取り込み運用できる形式へと変換することでした。

その知識の多くは、依然として説明文章によるレポート、脅威評価書、アナリストメモ、リサーチリポジトリ、そして人間の専門知識の中に埋もれたままになっています。それらの洞察を検知、ハンティングロジック、調査手順、あるいは自動ワークフローに変換するには、今なお相当な解釈作業が必要です。

多くの点において、運用的インテリジェンスは「失われた中間層」となってきました。単一のインジケーター(痕跡)のマッチングに還元するには複雑すぎる一方で、レポートや人間のワークフローの中に閉じ込められたままにしておくには貴重すぎる存在なのです。

AIはこの現状を変える機会を提供しますが、それは運用的インテリジェンス自体がマシン(機械)による処理に対応できるよう準備されている場合に限られます。

Part II: AI対応型オペレーショナル脅威インテリジェンスの5つの要件

要件1:構造化—攻撃者に関して把握している知識の体系化(コード化)

第1の要件は「構造化」です。

脅威インテリジェンスは歴史的に、攻撃者に関する広範な知識を文章(散文)で記録してきました。リサーチレポートでは、攻撃者が特定のキャンペーンを展開し、特定のツールを使用し、特定のテクニックに頼り、特定の業界を標的にし、認識可能な順序でアクティビティを実行する傾向があることを説明できます。

人間アナリストはそれらのレポートを読み、頭の中でその関係性を関連付けることができます。

しかしAIシステムは、新たなセキュリティイベントに遭遇するたびに、それらすべての関係性を一から再発見しなければならない状態であってはなりません。

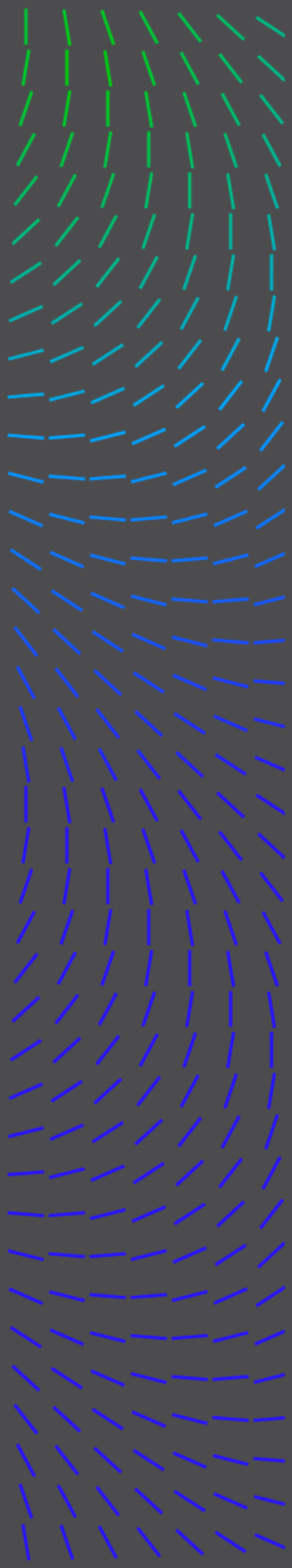
構造化されたインテリジェンスとは、脅威情勢を構成する実体(エンティティ)とそれらの間の関係性を明示的に表現することを意味します。脅威アクターはキャンペーンに関連付けられ、キャンペーンはツールやインフラストラクチャに関連付けられ、ツールはテクニックや動作に関連付けられ、脆弱性は悪用アクティビティに関連付けられ、そして観測結果はそれが確認されたより広範な攻撃者の運用へと連れ戻して関連付けられます。

また、構造化においては、確信度(コンフィデンス)、由来(プロベナンス)、裏付けとなる証拠、時間的コンテキストといった性質も保持される必要があります。2つのエンティティが関連していると知ることも有用ですが、なぜその関連性が存在するのか、その評価の強さはどの程度か、そして最後に観測されたのはいつかを知ることで、インテリジェンスの価値は大幅に高まります。

これこそが、攻撃者のアクティビティの「記録」から、攻撃者に対する理解の「体系化(コード化)」への進化です。

これを実現するには、リサーチャーが把握した知識を再現可能な方法で捉えられるよう、インテリジェンス組織が一貫したデータモデル、タクソノミー(分類法)、関係性、分析基準を開発する必要があります。MISPのようなオープンなプラットフォームや標準は重要な基盤を提供できますが、個別のテクノロジー以上に、その根底にある規律(ディシプリン)こそが重要となります。

目的は、構造化それ自体ではありません。人間のインテリジェンス分析の結果を保持し、アナリスト、製品、自動ワークフロー、そしてAIシステムが同じ関係性を何度も再構築することなく再利用できるようにすることにあります。



要件2:コンテキスト(文脈)化—動作に意味を与える

構造はインテリジェンスがどのように表現されるかを説明します。コンテキスト(文脈)は観測結果がなぜ重要なのかを説明します。

この区別は、攻撃者が攻撃に関連するアーティファクト(形跡)を変更する能力を強めるにつれて、ますます重要になります。マルウェアは改変され、インフラストラクチャはローテーションされ、コマンドは変更され、個々のインジケーターの運用寿命はより短くなる可能性があります。

一方、侵入の根本的な目的はより限定的です。攻撃者は依然として、アクセス権の獲得、環境の把握、認証情報の取得、権限の昇格、永続性の維持、システム間の横展開、情報へのアクセス、あるいは運用的インパクトの創出を行う必要があります。

現代のエンドポイントおよびネットワークセキュリティ技術は、すでにこうした動作の多くを特定しています。インテリジェンスにおける機会とは、それらの動作をコンテキストの中に位置づけることにあります。

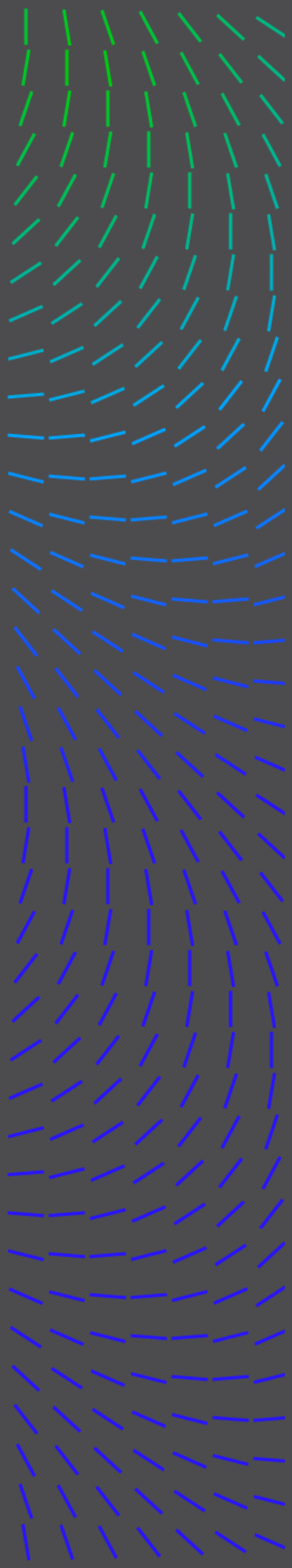
従来のエンリッチメント(情報付加)では、特定のハッシュ、ドメイン、IPアドレスが過去に観測されたことがあるかを問うかもしれません。一方、コンテキストを備えた運用インテリジェンスは、より広い問いを投げかけます。そのアクティビティが過去のキャンペーンで確認された動作に類似しているか、過去にどの攻撃者が同等の手法を用いていたか、どのような他のアクティビティがそれに伴う傾向があったか、そしてそれらのパターンが防御対象の組織にとってどれほど関連性があるか、といった問いです。

したがって、優れたコンテキスト型インテリジェンスには、単なるデータの蓄積以上のものが求められます。動作、攻撃者の意図、キャンペーンの履歴、標的設定、業界、地理的地域、そしてミッション(使命)の間の関係性を理解することが必要なのです。

また、不確実性を保持することも不可欠です。動作の類似性が自動的に「攻撃者の特定(アトリビューション)」に結びつくべきではありません。攻撃者は頻繁にテクニック、ツール、TTPs(戦術・技術・手順)を共有します。コンテキストは、証拠が裏付けていない確信を無理に作り出すものではなく、調査における仮説を強固にするものでなければなりません。

AIにとって、この区別は極めて重要です。目的は、すべてのイベントに確信を持ってラベルを貼れるだけの情報をモデルに与えることではありません。その観測結果が何を意味し得るかをモデルが推理し、どのような証拠があればその評価が強まるか、あるいは弱まるかを判断できるだけの十分なコンテキストを提供することにあります。

これこそが、動作インテリジェンスが現代のエンドポイント検知・対応(EDR)、ネットワーク検知・対応(NDR)、エンドポイント、サードパーティのテレメトリ、そして脅威リサーチャーの蓄積された知識の架け橋となり得る領域なのです。



要件3:マシンリーダブル(機械判読可能)—
別の変換層を介さずにインテリジェンスを取り込み可能にする

構造化されたインテリジェンスとマシンリーダブル(機械判読可能)なインテリジェンスは関連していますが、同一のものではありません。

インテリジェンスが良く整理されていたとしても、セキュリティ製品や自動化ワークフローがそれを利用できるようになる前に、人間のアナリストによる解釈を必要とする場合があります。

マシンリーダブルなインテリジェンスは、そこからさらに一歩踏み込みます。これは、セキュリティシステム、API、自動化ワークフロー、およびAIエージェントが、その意味合い(ニュアンス)を損なうことなく直接取り込めるよう、十分な一貫性をもってインテリジェンスを表現し配信することを意味します。

次世代の脅威インテリジェンスの利用対象(コンシューマー)に、マシン(機械)が含まれる割合が益々増加していくため、この違いが重要となります。



歴史的に、リサーチャーが詳細なレポートを公開し、そのレポートのどの部分を運用上のアクションに変換すべきかの判断を、脅威ハンター、検知エンジニア、またはアナリストに依存することが可能でした。AI主導の環境では、その変換ははるかに直接的なものになる必要があります。

AIシステムは、別のアナリストが説明文章からすべての関係性を手動で抽出することをあらかじめ要することなく、どの攻撃者がキャンペーンに関連しているか、どのテクニックや動作が彼らの運用を特徴づけているか、それらの関係性を裏付ける確信度はどれほどか、どのセクターが関連しているか、そして調査中にどの観測結果が有用であるかを特定できる必要があります。

これを実現するには、一貫したスキーマ、識別子(ID)、タクソノミー(分類法)、関係性、確信度モデル、そして多様な利用対象(コンシューマー)にインテリジェンスを配信する方法が必要となります。また、インテリジェンスの作成者が、ダウンストリーム(後続)のマシンを人間のアナリストと並ぶファーストクラス(同格・最優先)の利用者として考慮することも求められます。

これは、説明文章によるレポートを排除することを意味するものではありません。戦略的レポート、分析、そして人間による説明は、引き続き不可欠です。重要なのは、そうしたレポートを裏付ける根本的なインテリジェンスが、製品、API、ハンティング・ワークフロー、そしてAIシステムへも独立して流れていけるようにすることです。

こうして脅威インテリジェンスは、同じ信頼できる基盤から人間とマシンの双方に役立てられるようになります。

要件4:アクション指向—次に何を行うべきかの指針を示す

第4の次元は、有用なマシンリーダブル(機械判読可能)なインテリジェンスと、受動的な知識リポジトリを分かつものです。

運用インテリジェンスは、アクション(対処)の指針となる必要があります。

従来のエンリッチメント(情報付加)は、アナリストの目の前にあるイベントに関する問い、すなわち「このファイル、このインフラストラクチャ、またはこのアラートについて何が分かっているか?」に答えるにとどまることが多く見られました。

AI対応の運用インテリジェンスは、「次に何を行うべきか」を決定する上でも役立つ必要があります。

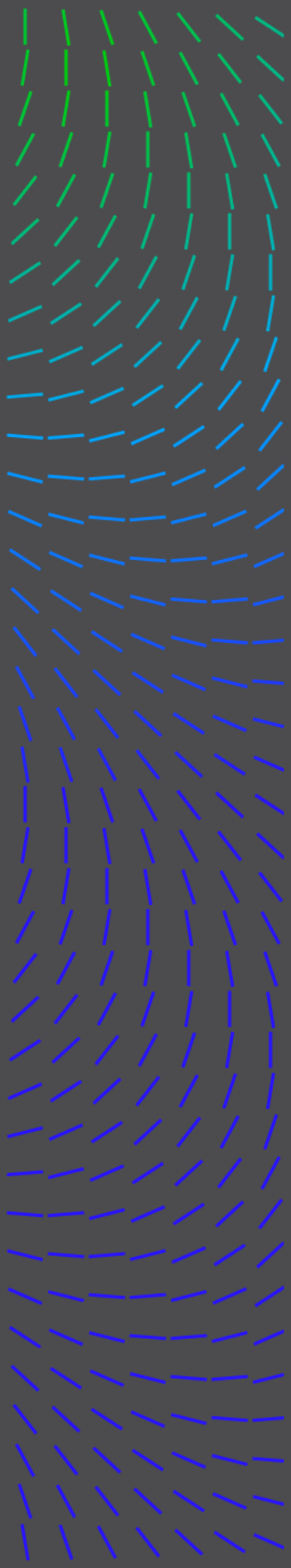
観測された動作が過去に分析されたキャンペーンに関連するパターンと類似している場合、インテリジェンスは、それらの攻撃運用において過去にどのような他の動作が見られたか、どのシステムやアイデンティティをさらに詳細に調査すべきか、そして初期の仮説を補強または弱める観測結果には何があるかについてのコンテキストを提供できます。

これは、経験豊富な脅威ハンターが既に行っている作業方法を反映しています。彼らが完全な情報から調査を始めることはめったにありません。仮説を立て、攻撃者の動作に対する理解を活用して存在するはずの証拠を特定し、その証拠を捜索し、得られた発見に基づいて評価を絶えず更新していきます。

AIは、その手法の要素をより広範な調査へと適用する機会を提供します。

これにより、アラートは単に情報を付加してクローズすべき隔離されたイベントではなく、インテリジェンスに基づいた調査の出発点となります。運用インテリジェンスはAIやアナリストが次に発すべき問いの指針を示し、セキュリティのテレメトリはその問いに答えるために必要な証拠を提供します。

目的は、自律的なアクションそれ自体ではありません。調査における「確実性(アシュアランス)」を高めることにあります。



セキュリティリーダーにとって、この違いは極めて重要です。1つのアラートを解決したからといって、必ずしも侵入が終了したことが証明されるわけではないからです。

インテリジェンスに基づいたハンティングは、関連する証拠が環境内の他の場所に存在するかどうか、そして元のイベントが単発の発生なのか、それともより大きな攻撃の一部なのかを判断するのに役立ちます。

同じ原則がプロアクティブ(能動的)な予防策にも当てはまります。運用インテリジェンスは、自組織の業界、地域、ミッション、およびリスクの露出(エクスポージャー)に最も関連性の高い攻撃者が誰であるかを組織が理解し、その理解を準備状況に関する問いへと変換するのに役立つべきです。適切な統制(コントロール)が存在しているか？ 関連するテレメトリは利用可能か？ 自組織の環境はそれらの脅威に関連する動作を検知できるか？ アラートが表示される前に標的を絞ったハンティングを実行できるか？

マシン・スピードの環境において、インテリジェンスは脅威を説明するだけで終わることはできません。その後に続く防御の意思決定を導く支援をする必要があります。

要件5:セキュリティ境界内での利用可能性—ソブリンAIへインテリジェンスをもたらす

第5の次元は、規制対象の組織、ミッションクリティカルな組織、高度なセキュリティ保証が求められる組織にとって特に重要となります。それは「AIの動作が許可されている場所でインテリジェンスを利用できなければならない」という点です。

今日のAIアーキテクチャの多くは、接続性を前提としています。クラウドベースのシステムは、不確実性が生じるたびに追加情報を取得したり、他のサービスに照会したり、外部ソースと比較したり、追加のコンテキストを要求したりします。

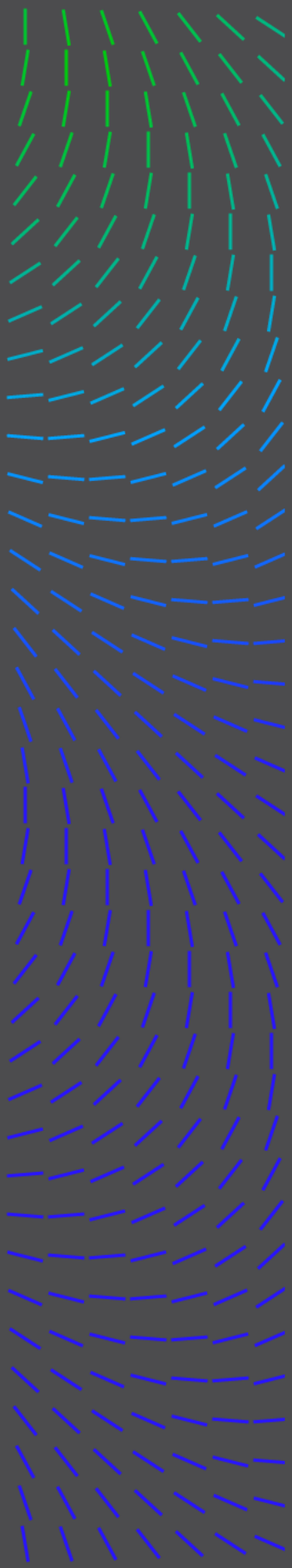
こうした反復的な情報の取得および拡張(エンリッチメント)サイクルは、あいまいさを解消し、確信度を高めるための追加の機会を提供するため、非常に価値があります。

しかし、ソブリン環境やエアギャップ(孤立)環境では、この前提が一変します。

未接続の環境で運用されている海軍艦艇を考えてみましょう。そのエンドポイントやネットワークは引き続き保護が必要であり、セキュリティ統制はテレメトリを生成し続け、防御側は疑わしいアクティビティを調査できなければなりません。艦艇がインターネットに接続できないからといって、攻撃者が低度化するわけではないのです。

その艦艇に搭載されたAI機能は、機密性の高いテレメトリを常に外部サービスへ送信したり、不確実性に直面するたびに追加のインテリジェンスを取得したりできるとは仮定できません。

インテリジェンスは、システムとともに持ち運ばれ(移動し)、存在していなければならないのです。



これにより、関連する2つの要件が生じます。すなわち「可用性」と「品質」です。

関連するインテリジェンスは、セキュリティ境界内に存在している必要があります。しかし、単に大量のデータ集合を環境内に持ち込むだけでは不十分です。システムには、不十分で不完全な、あるいはあいまいに記述されたインテリジェンスを、外部からのさらなる取得サイクルによって補う機会が少なくなる可能性があるためです。

その結果、最初から利用可能なインテリジェンスは、より高い基準を満たしている必要があります。

それは精選・検証され、由来（プロベナンス）や確信度が明確に把握されているべきです。攻撃者、キャンペーン、動作、標的選定、攻撃パターンの間の関係性は、モデルの推論に全面的に委ねるのではなく、明示的である必要があります。また、インテリジェンスは単に可能な限り最大のデータ量を表すのではなく、組織のミッションや脅威環境に関連するものであるべきです。

これらの品質は、AIの誤りやハルシネーション（幻覚）を完全に無くすものではありません。しかし、推論のためのより強固な基盤を提供することで、モデルが行わなければならない裏付けのない解釈作業を減らすことができます。

これこそが、ソブリンAI（主権型AI）を「単に別の場所にデプロイされたクラウドAI」と見なすべきではない理由です。

モデルをローカルで実行することは「AIがどこで実行されるか」に対する解答にはなりますが、「AIが何を認知しているか」「その知識がどれほど信頼できるか」「外部世界が利用できない場合にどれほど効果的に推理できるか」に対する自動的な回答にはなりません。

接続されたアーキテクチャでは、AIは多くの場合、より多くのコンテキストが必要になった際、インテリジェンスにアクセスして反復処理を行うことができます。一方、ソブリン・アーキテクチャでは、インテリジェンスをAIの元へ届ける必要があり、かつ最初から役立つだけの十分な構造、関連性、コンテキスト、確信度を備えた状態で届かなければなりません。

したがって、高度なセキュリティ保証が求められる環境においては、信頼できる脅威インテリジェンス自体が、ソブリンAIのアーキテクチャの一部となるのです。



Part III: フレームワークの適用— インテリジェンスからインフラストラクチャへ

これら5つの次元は、AI主導のセキュリティ・アーキテクチャにおいて、脅威インテリジェンスがどのように見なされるべきかを変革します。

目的は、もはや単により多くの脅威データを生成することや、新たなエンリッチメント(情報付加)サービスを構築することではありません。目的は、人間、製品、自動化ワークフロー、およびAIシステム全体にわたってセキュリティ上の意思決定を支援できる、信頼性の高いインテリジェンス基盤を構築することにあります。

これは、Trellix Threat Intelligenceのより広範なミッション、すなわち「グローバルな可視性と最前線の専門知識を、人間およびAI主導のセキュリティ意思決定を支援する信頼性の高いインテリジェンスへと変化させ、可視性を人々、製品、マシン(機械)のための実行可能なアクションへと変えること」と密接に合致しています。

Trellixは、グローバルな脅威の可視性と、セキュリティ・テレメトリ、最前線のリサーチ、マルウェア分析、脅威ハンティング、パートナーシップ、そして攻撃者の活動に関する蓄積された知識を融合させています。その基盤は戦術的、運用的、戦略的インテリジェンスにわたりますが、運用インテリジェンスは「脅威を理解すること」と「それに基づいてアクションを起こすこと」を結ぶ重要な架け橋を提供します。

私たちの機会は、その運用上の知識を益々構造化および体系化(コード化)し、セキュリティシステムを直接支援できるようにすることにあります。

Trellix Threat Intelligenceにおいて、広範な可視性はインテリジェンスが開発される元となる観測結果を提供します。リサーチャーや脅威ハンターは、それらの観測結果を解釈し、なぜそれが重要なのかを明らかにします。構造化されたインテリジェンスは、攻撃者、キャンペーン、動作、ツール、標的選定、攻撃チェーンの間の関係性を保持します。マシンリーダブル(機械判読可能)な配信により、その知識が製品やAIシステムを支援できるようになり、一方でローカルなテレメトリは、それらのグローバルな観測結果が特定の環境内に関連しているかどうかを判断するために必要な証拠を提供します。

単一のコンポーネントだけで優位性が生み出されるわけではありません。

テレメトリは観測を提供します。人間の専門知識は理解を提供します。構造化は関係性を保持します。AIは大規模な推論を可能にします。セキュリティ統制は防御アクションのためのメカニズムを提供します。

これらが一体となることで、信頼できるインテリジェンスの基盤が創出されるのです。

TRELLIX WISE™:
マシン・スピードでのオペレーショナル脅威インテリジェンスの適用

Trellix Wiseは、これらの機能がいかに相乗的に統合され得るかを象徴しています。

目的は、単に汎用言語モデルをセキュリティ・オペレーション・センター(SOC)の傍らに配置することではありません。システムの推論能力が顧客の環境と、そこへ最も強く関連する攻撃者の双方に向けられるよう、AIの推論力、信頼できる脅威インテリジェンス、セキュリティ・テレメトリ、そして防御機能を組み合わせることにあります。

運用インテリジェンスはそれら攻撃者がどのように行動するかという知識を提供し、ローカル・テレメトリは顧客環境内部で何が起きているかという証拠を提供します。AIはこれら2つを結びつけ、アナリストによるアクティビティの解釈を支援し、調査を誘導し、インテリジェンスをハンティングや防御ワークフローへと変換する手助けをします。

したがって、モデルは「推論能力」を提供するものの、その推論の価値はモデルを取り巻く要素に依存します。

脅威インテリジェンスは理解を提供し、テレメトリはローカルな観測データを提供し、セキュリティ統制はアクション(対処)を提供します。

ソブリン環境やエアギャップ(孤立)環境において、このアーキテクチャは特に重要な意味を持ちます。AI主導の調査を支援するために必要なインテリジェンスを、外部のエンリッチメント(情報付加)サービスへの継続的なアクセスに依存することなく、組織のセキュリティ境界内で提供できるためです。

これこそが、5つの次元が集約されるポイントです。インテリジェンスは、攻撃者に関する知識を保持できるほど十分に「構造化」され、観測結果がなぜ重要かを説明できるほど十分に「コンテキストを備え」、AIが直接取り込めるほど十分に「マシンリーダブル」であり、調査やハンティングを導けるほど十分に「アクション指向」であり、そしてAIの動作が期待される環境内で「利用可能」でなければなりません。

インフラとしてのインテリジェンス

長期的な機会は、より迅速なエンリッチメント(情報付加)や、個々のAIアシストによる調査の枠を超えて広がります。それは、グローバルな脅威の可視性と、ローカルな防御アクションとの間に継続的な関係性を構築することです。

脅威アクティビティは、テレメトリ、リサーチ、ハンティング、マルウェア分析、パートナーシップ、およびインテリジェンス収集を通じて観測されます。リサーチャーは、それらの観測結果が何を意味するかを解明します。その理解は、人間とマシンの双方を支援できる構造化された運用インテリジェンスとなります。AIは、ローカルのセキュリティ・アクティビティを解釈し、次にどのような証拠を探すべきかを決定する際に、その知識を適用します。そして新たな観測結果は、インテリジェンス、ハンティングロジック、検知ルール、そして将来の防御上の意思決定へと還元されていくのです。

これにより、可視性、インテリジェンス、AI、ハンティング、検知、調査、そして対応の間に強化し合う(相乗的な)関係が生まれます。

また、これは脅威インテリジェンス自体の役割をも変化させます。

脅威インテリジェンスは歴史的に、主に「コンテンツ」(レポート、フィード、インジケーター、評価書、ブリーフィング)として扱われてきました。そうした成果物は今後も重要であり続けます。しかし、AI主導のセキュリティ・アーキテクチャにおいて、インテリジェンスはセキュリティシステムが攻撃者を理解するための「インフラストラクチャ」の一部へとますます変化していきます。

それこそが、次世代のサイバーセキュリティにおける最も重要な差別化要因の1つとなる可能性があります。

AIモデルは向上し続け、エージェントのフレームワークは成熟し、今日差別化されているように見える機能の多くは、最終的には広く利用可能になるでしょう。永続的な優位性は、モデルを取り巻く要素(モデルが何を見ることができるか、何を知っているか、その知識がどれほど信頼できるか、ミッションにどれほど関連しているか、そして攻撃者の活動に関する新たな理解がどれほど迅速に防御アクションへと変換されるか)に益々依存するようになります。

最も高いリスクを伴う組織にとって、これらの要件はさらに重要になります。そうした組織に必要なのは、単なるAIではありません。信頼できるインテリジェンスに裏打ちされ、自らが制御するアーキテクチャ内にデプロイされ、外部への接続が制限されているか利用できない状況であっても意思決定を支援できるAIなのです。

したがって、AI主導の防御の未来は、モデルがどれほどインテリジェントになるかだけで決まるわけではありません。モデルを取り巻く運用インテリジェンスの品質、およびセキュリティが現在要求しているスピードでその知識を利用可能にする私たちの能力にも同様に依存するのです。

モデルのインテリジェンス(知能)は重要です。しかし、私たちがモデルに与えるインテリジェンス(脅威情報)は、それ以上に重要となるかもしれません。

関連するTRELLIXのリサーチと展望

Weaponized AI: The Commoditization of Cybercrime
Trellix Advanced Research Center

Accelerating Trellix Engineering Transformation with AI
Trellix

Discover more cybersecurity threat insights from the Trellix Advanced Research Center, and subscribe to the ARC newsletter on LinkedIn.

Copyright © 2026 Musarubra US LLC