膨大なサイバー攻撃、効率的に対処するには？ ～マルウェアを自動的に判別するAI技術～

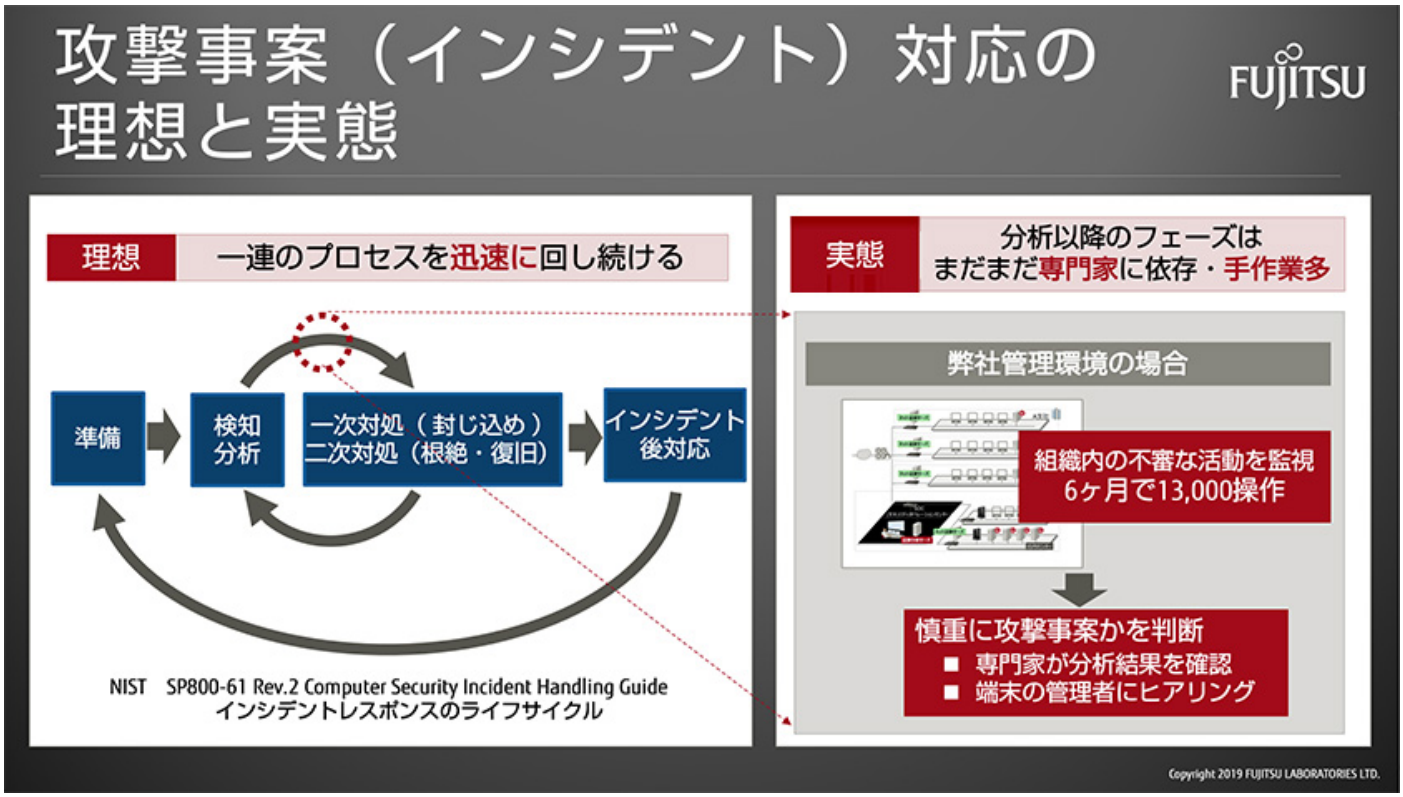
FUJITSU JOURNAL / 2019年7月12日

マルウェアの不審な動きへの対処をAIが自動判別

近年、企業や組織、個人を狙ったサイバー攻撃が増え続けています。中でも、ある特定の企業や組織、個人を標的にして、執拗に重要な情報を盗み取ろうとしたり、システムを破壊したりする「標的型攻撃」は大きな脅威となっています。

標的型攻撃では、攻撃者は遠隔操作が可能なマルウェアを巧妙な手口で企業や組織内のネットワークに送り込み、マルウェアに感染したパソコンなどの端末を遠隔操作して機密情報を盗み出そうとします。こうした攻撃に対し、企業や組織では様々なセキュリティ対策を実施しています。マルウェアの「不審な活動」を検知した時には、その活動が「危険な攻撃」であるかどうかをセキュリティ専門家が手作業で調査・確認し、対処が必要かどうか時間をかけて判断しています。

ただし、次々に仕掛けられるマルウェアの全てに時間をかけていたのでは、対処が間に合わなくなる可能性があります。また、サイバー攻撃への「準備」「検知」「分析」「対処」「インシデント後対応」といった一連のプロセスへの対処を全てセキュリティ専門家の手作業に依存していては、迅速な対応が難しくなることがあります。



サイバー攻撃への対処は「準備」「検知」「分析」「対処」「インシデント後対応」を迅速に回し続けることが大切

また、マルウェアに対処する場合には、攻撃を受けたパソコンなどの端末をネットワークから遮断する必要がありますが、対処が必要かどうかには慎重な判断が求められます。しかし、日本国内では、こうした高度な判断ができる専門的な技術者（セキュリティ人材）が不足しています。経済産業省が2016年に発表した「IT人材の最新動向と将来推計に関する調査結果」によると、日本では2020年にセキュリティ人材が19万3,000人も不足することが示されています。

こうした課題を解決するため、富士通と富士通研究所では、マルウェアの不審な活動を検知した時に、AI（人工知能）がセキュリティ専門家と同等レベルで対処が必要かどうかを自動で判断する技術を開発しました。

「不審な活動は危険な攻撃か」をAIにどう判別させるか

このAI技術の開発には、いくつかの課題がありました。まずは、不審な活動はマルウェアの活動であり「企業や組織のネットワークや機器に損害を与えるか否か」を、AIにいかに学習させるかということです。

サーバや端末、ネットワーク機器には、正常な動作の記録とマルウェアによる攻撃の記録が混在し、大量に蓄積されています。「この活動はマルウェアによる攻撃である」とAIが適切に判別できるようにするには、膨大な動作記録の中から攻撃性のある動作の記録を正確に選別し、それをAIに学習させる必要があります。

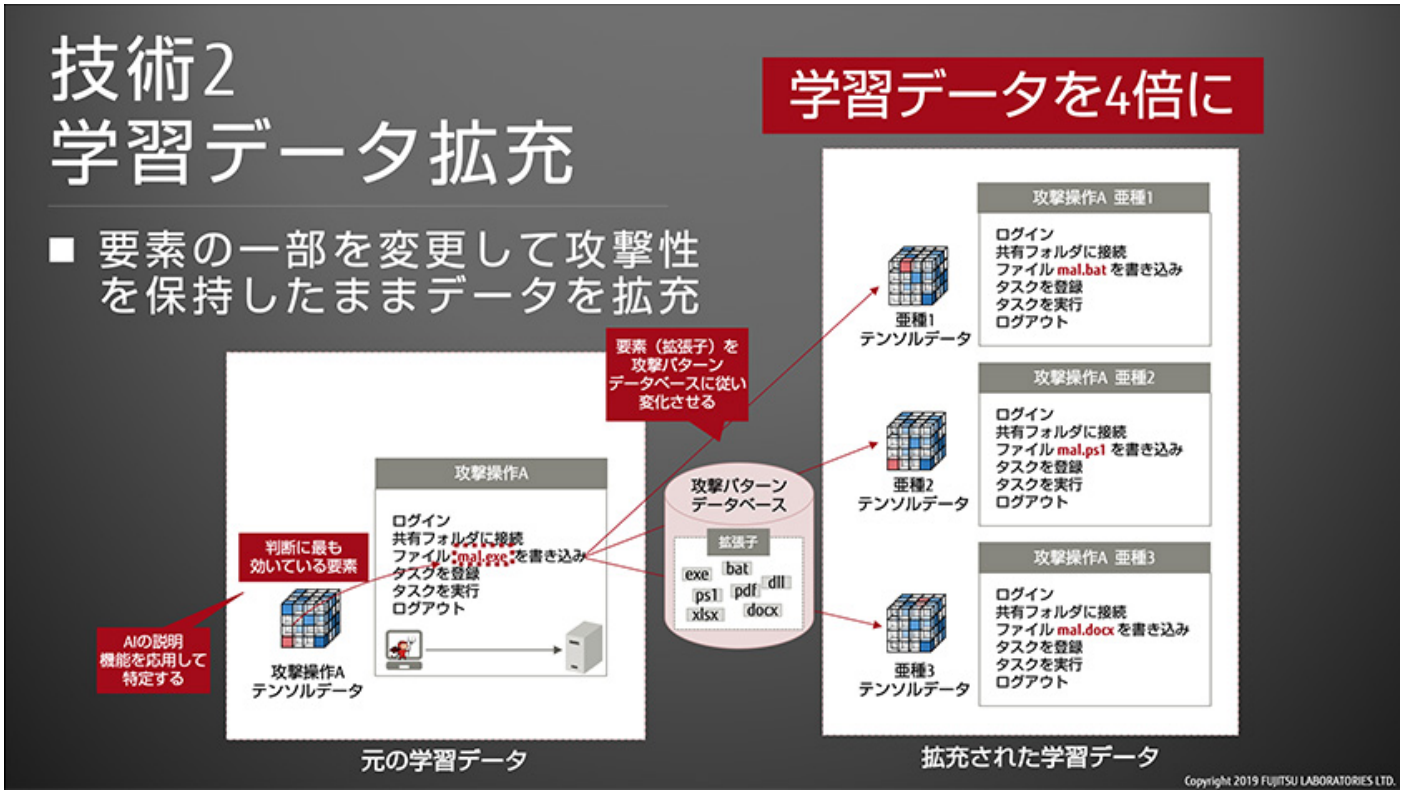
もう1つの課題は、AIに学習させるデータの「量」です。攻撃データの量はもともと少ないため、少量の学習データを加工して疑似データを作成し、学習するためのデータ量を拡張しています。しかし、標的型攻撃の学習データは単純な加工処理では攻撃性が失われる場合があり、拡張が難しいことが課題でした。

「Deep Tensor」で学習データ量を4倍に拡張

そこで富士通研究所は、攻撃の記録を正確に抽出できる「学習データの抽出技術」と、十分な量の標的型攻撃に関する学習データを確保する「学習データの拡張技術」を開発しました。

学習データを正確に抽出する技術では、富士通研究所がこれまでのセキュリティ関連業務と研究で培ってきたノウハウをもとに、攻撃分析で得た約7年間の実績データから「攻撃パターンデータベース」を構築しました。攻撃の記録をもとに攻撃属性と危険度を判別しながら相関の高い記録を検索し、一連の攻撃記録を抽出。こうして構築したデータベースを利用することで、膨大な記録から一連の情報搾取の活動を正確に特定・抽出することが可能になりました。

一方、学習データの拡張技術では、抽出した標的型攻撃の攻撃パターンに対して、攻撃と他の動作との一連のつながり（相関関係）を示すグラフ構造のデータを高度に解析し、新たな知見を導き出す富士通のAI技術「Deep Tensor（ディープテンソル）」を活用しました。攻撃パターンごとに、攻撃と他の動作との一連のつながりを分析。攻撃パターンを構成する要素の一部を変更することで、攻撃性を失うことなく、攻撃パターンの学習データを4倍に拡張できました。



攻撃の動作で書き込むファイル名を変更するなど、要素の一部を変更することで攻撃の亜種を作成

数時間から数日間かかっていた判断が、AIの活用で数十秒から数分に短縮

富士通研究所では、今回生成した学習データによって学習させた判定モデルの評価実験を実施しました。1万2000件に上る約4カ月分のデータを活用してシミュレーションした結果、セキュリティの専門家が手動で分析した結果との一致率は約95%、要対処の見逃しはゼロという結果が得られました。

また、国立研究開発法人情報通信研究機構（NICT）が運用しているサイバー攻撃誘引基盤「STARDUST（スターダスト）」で、企業を狙った実際のサイバー攻撃を使用した実証実験も実施。その結果、対処が必要な攻撃であることをAI技術で自動判断できるという有用性を確認しました。さらに、今回開発した技術によって、これまで数時間から数日間かかっていた専門家による対処の要否判断を、数十秒から数分で高精度に自動判断することも可能にしました。

富士通研究所では、標的型攻撃の被害状況の全貌を短時間で分析する「高速フォレンジック技術」と組み合わせることで、攻撃の分析から対処指示までの一連の対応を自動化し、サイバー攻撃への即時対処と被害の最小化への貢献が期待できると考えています。

サイバー攻撃から人々の暮らしを守り、安心安全な世の中の実現を目指す

近年は、大企業だけではなく、関連企業や取引先、従業員個人などへの攻撃を足掛かりに、本来の標的である企業や組織に攻撃を仕掛ける「サプライチェーン攻撃」も増えています。つまり、標的型攻撃は誰でもターゲットになる可能性があるのです。

また、企業や組織が標的型攻撃を受け、重要情報が流出してしまうといった事態になれば、社会的信用を失墜することにもつながりかねません。加えて、個人を標的とした愉快犯的なものから組織や重要インフラ、国家を標的とした経済犯および組織犯的なものに移行するなど、次第に高度化・複雑化しているのが現状です。

今回開発した技術を活用することで、要対処と判断されたサイバー攻撃に対してすぐに対策を実施でき、企業や組織の損失防止や業務継続に寄与します。さらに、多くの人々の日々の暮らしの安心安全を支える基盤にもなり得ます。

富士通では、今回開発した技術をサイバー攻撃の対処基盤として、マネージドセキュリティサービスなどでの活用を視野に入れています。また、高度なセキュリティ技術を持つ専門家を発掘・育成する「セキュリティマイスター認定制度」を設けています。今後も富士通は、お客様のICT運用を支えながら、安心安全な世の中を目指します。

FUJITSU JOURNAL / 2019年7月12日