未知のサイバー攻撃を自動検知する自己学習型システム異常検知技術（ASI）

多賀戸 裕樹　栄 純明　喜田 弘司　朝倉 敬喜

要 旨

年々巧妙化するサイバー攻撃に対する対策が求められています。自己学習型システム異常検知技術（ASI）は、監視ソフトウェアによりPCやサーバの詳細動作ログを収集し、本ログに対する機械学習（AI）を行うことで、監視対象システムの平常状態を生成します。生成した平常状態と現在のシステム動作との比較を行うことで、未知の攻撃も検知します。本技術を適用したセキュリティ監視システムでは、攻撃プロセスの初期・最終フェーズだけでなく、システム内における感染拡大といった、中間フェーズも含めた攻撃プロセスの全体を通して検知を行うことができ、より強固なセキュリティを実現します。

KeyWords　サイバーセキュリティ／サイバー攻撃／監視エージェント／人工知能(AI)／機械学習／異常検知

1. はじめに

昨今、企業や公的機関などの情報システムをターゲットとするサイバー攻撃は巧妙化の一途をたどっています。標的型攻撃や公になっていないソフトウェアの脆弱性を突く攻撃などによる情報漏えいのリスクが高まっています。

サイバー攻撃対策には、アンチウィルスソフトウェアをインストールするなど情報システムの個々のユーザーが行うものと、ファイアウォールやセキュリティゲートウェイの設置・運用などの情報システムの管理者が行うものがあります。現在はいずれも既知のウィルスや攻撃手法の情報に基づいて、対策を行うものが主流です。

したがって、いまだ公になっていないソフトウェアの脆弱性を突く攻撃、あるいは前例のないまったく新しい攻撃手法（以後、本稿ではこのような攻撃を「未知の攻撃」と呼びます）に対しては、攻撃の発見・検知が極めて困難であり、検知するまでに長期間を要したり、外部機関からの通報を受けて初めて攻撃を受けたことが発覚する事例が発生しています。

本稿では、日々新しく生まれ続ける攻撃手法ではなく、攻撃を受ける側の情報システムの平常動作をAI技術の活用により把握し、その動作の変化をリアルタイムに検知することで、未知の攻撃を受けた場合でも、攻撃を受けた事実を迅速に検知し、攻撃による影響範囲を自動で隔離する、自己学習型システム異常検知技術（Automated Security Intelligence：ASI）[1)]を紹介します。

2. 自己学習型システム異常検知技術（ASI）

従来のセキュリティ監視システム、及びASIを用いたセキュリティ監視システムのイメージを**図1**（左：従来、右：ASI）に示します。図1（左）に示す従来システムには以下のような課題があります。

- SIEM（Security Information and Event Management：セキュリティ情報及びイベント管理）ではファイアウォール（FW）やアンチウィルスソフトウェアが出力するログの監視が主であり、前記ソフトウェアの開発者が出力を意図していない情報については監視できない。
- 既存の攻撃手法や脆弱性の情報に基づいた監視であり、未知の攻撃については検知できず、すり抜けてしまう。
- 有事の際にはFWやPCなどの内部状態を含めた詳細分析が必要となるが、動作ログが各所に分散

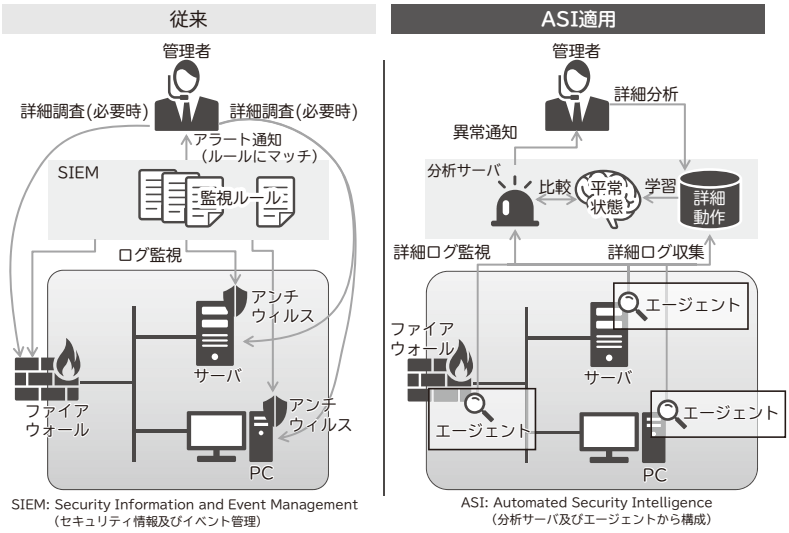


図1　従来及びASIを適用したセキュリティ監視システムのイメージ

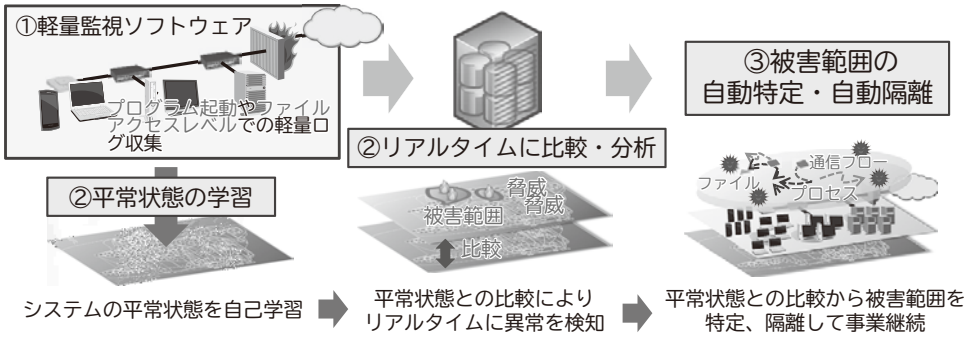


図2　ASIの技術的特長

しており分析に手間が掛かる場合や、そもそも詳細分析に必要なデータが得られない場合がある。

また、情報処理推進機構により公開されている「『高度標的型攻撃』対策に向けたシステム設計ガイド」[2)]では、従来重点的に対策が行われてきた入口・出口対策（ファイアウォールによるウィルスの検知・遮断や、端末でのアンチウィルスソフトウェアによる検知・駆除）に加え、今後はウィルスが前記の対策をすり抜けて内部に侵入してしまうことを前提にした内部対策の強化が必要であると指摘しています。

ASIは、先に述べた課題を解決し、攻撃者による内部システムへの侵入、感染範囲の拡大、及び重要情報の窃取のすべての段階において、攻撃を検知し、システムの防御を実現するために開発された技術であり、以下のような特長を持っています（**図2**）。

(1) 軽量な監視ソフトウェアで詳細なログ情報を収集

システム動作を監視する従来のソフトウェア（エージェント）は、PCやサーバの動きを遅延させるなどの悪影響を与えることがあります。ASIでは、システムに掛かる負荷を常に考慮して、監視処理のタイミングなどを適宜制御する機能を持つ軽量なエージェントを開発しました。これにより、システム動作を遅延させずに、プログラムの起動、ファイルへのアクセス、及びネットワークへのアクセスなどの詳細な動作ログの収集を実現しました。

また、収集した動作ログはデータベースとして一元的に管理するため、有事の際の詳細ログ分析においても、各所に散らばったログデータを収集するなどの手

間を掛ける必要がなく、セキュリティ管理者のスピーディなインシデント対応が可能です。

(2) AIを活用してリアルタイムに異常を検知

PCやサーバなどシステム全体の複雑な動作状態（プログラムの起動、ファイルアクセス、ネットワークアクセスなど）から平常状態を機械学習し、本平常状態と現在のシステムの動きをリアルタイムに比較し、平常状態から外れた場合、異常として自動検知します。異常を検知した場合、原因となるシステムの一連の動作を自動で特定し、ネットワークから自動隔離できます。これにより、システム全体を止めることなく被害範囲の拡大を最小限に抑える防御を実現します。

(3) 被害範囲を特定し、ネットワークから自動的に隔離

システムの動きを詳細に把握しているため、異常検知から本検知に至るまでのシステムの一連の動作を時系列で自動追跡できます。これにより、従来の人手による作業に比べ、1/10の時間で被害範囲の特定が可能です。また、将来的にはシステム管理ツールやSDN（Software-Defined Networking）との連携により、特定した被害範囲をネットワークから切断することで、自動的な隔離も可能です。これらにより、情報漏えいやシステム破壊の被害の拡大を最小限に抑え、システム全体の停止回避を実現します。

3. AI技術を用いた平常状態の生成

前章で述べたように、ASIではエージェントと呼ぶ監視ソフトウェアをシステム内のPCやサーバにインストールし、当該マシンの詳細動作ログをリアルタイムで収集します。収集した動作ログに対して学習（AI）処理を行うことにより、ASIが監視する対象のシステムにおける平常状態を生成します。

平常状態を生成するに当たっての基本的な考え方は、監視対象のシステムの動作は安定であるということです。ここでいう動作とは、PCやサーバ内におけるプログラムの起動、プログラムからのファイルへのアクセス、プログラムからのネットワークへのアクセスを指します。これに関して、図を用いてもう少し詳しく説明します。

図3は、ある企業のネットワークシステムと、そのシステムにおける平常状態のイメージを示すものです。例示したシステムでは、DNSサーバ及びWebプロキシサーバといった企業内で共通的に使用される共用サーバが存在するサブネットワーク、開発用サーバやPCなどが設置され開発部門により使用される開発部門サブネットワーク、及び事務部門により使用される事務部門サブネットワークがあります。

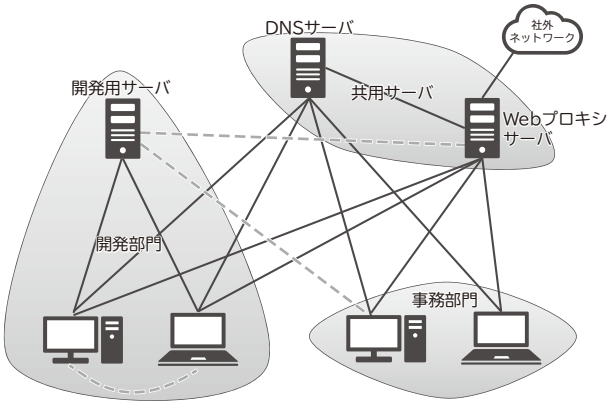


図3　ASIにおける平常状態のイメージ

共用サーバは企業内のすべてのPCから共通的にアクセスされます。また、開発部門における開発用サーバは開発部門のPCからアクセスされます。事務部門のPCが開発部門の開発用サーバにアクセスすることは、一般的には起こりません。このように、図中の実線で示したマシン間の関係を平常状態と呼びます。

平常状態の学習が完了した後、平常状態に含まれないネットワークアクセスが行われたことを検知すると、これを異常として報告します。例えば、図中に点線で示すように、開発部門内のPC同士が直接通信を行った場合、事務部門のPCから開発部門の開発用サーバへのアクセスが行われた場合、また通常は社外ネットワークへの通信を行わない開発用サーバがWebプロキシサーバへの通信を行った場合を異常として検知します。

部門内端末同士の直接の接続は、攻撃プロセスにおける感染拡大フェーズ（ラテラルムーブメント：Lateral Movement）でよく見られます。また、開発用サーバからWebプロキシサーバへの接続は、攻撃プロセスにおける目的遂行フェーズ（重要情報の窃取）で行われるものです。

従来のサイバー防御システムでは、入口・出口対策が主であり、本対策を突破されると攻撃検知が非常に困難であるという問題があります。ASIでは、攻撃プロセスの初期潜入・目的遂行のフェーズのみでなく、攻撃者がシステム内において最終攻撃目標を探して感染を拡大する中間

図4　ASIによる異常検知の例

フェーズでも、これを検知することができるため、より攻撃検知のチャンスを増やすことができます。

本章では、PCやサーバ間のネットワーク通信に関する平常状態のみについて述べましたが、プログラムの起動やプログラムからのファイルアクセスなどについても平常状態を学習しており、ネットワークの通信がないPCやサーバ内部のみにおける平常状態からの逸脱に関しても、検知することを可能としています。

図4にASIにおける異常検知の画面例を示します（見やすくするため画像を一部加工しています）。右側中央の円状のグラフが、監視対象のネットワークを示しており、実線がPCやサーバ間の平常のネットワーク接続、太い実線が平常でない（すなわち異常と見なした）ネットワーク接続を示します。

4. おわりに

本稿では、攻撃を受ける側の情報システムの平常動作をAI技術により把握し、その動作の変化をリアルタイムに検知することで、未知のサイバー攻撃を受けた場合でも、攻撃をリアルタイムに検知し、その影響範囲を自動で隔離する、自己学習型システム異常検知技術を紹介しました。

参考文献

1）NECプレスリリース：NEC、AI（人工知能）を活用し未知のサイバー攻撃を自動検知する「自己学習型システム異常検知技術」を開発，2015.12
http://jpn.nec.com/press/201512/20151210_01.html
2）情報処理推進機構：「『高度標的型攻撃』対策に向けたシステム設計ガイド」の公開，2014.9
https://www.ipa.go.jp/security/vuln/newattack.html

執筆者プロフィール

多賀戸 裕樹
セキュリティ研究所
主任研究員

栄 純明
セキュリティ研究所
主任研究員

喜田 弘司
サイバーセキュリティ戦略本部
エキスパート

朝倉 敬喜
セキュリティ研究所
研究部長

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.69 No.1　AIによる社会価値創造 ～NEC the WISEの世界～

AIによる社会価値創造特集によせて
AI時代における社会ビジョン ～人々の働き方、生き方、倫理のあり方～
NECが目指すAIによる社会価値創造

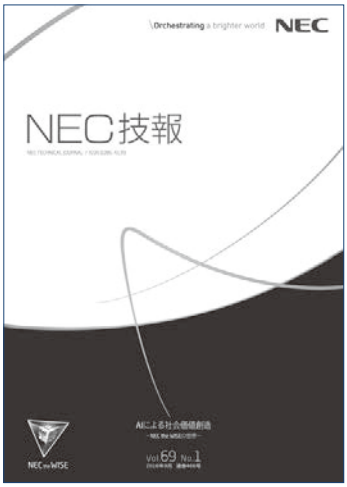
◇ 特集論文

NECが目指す社会価値創造像
都市空間の安全・安心を支えるセーフティ・オペレーション
新たな消費エクスペリエンスを提供するリテール産業オペレーション
都市交通サービスにおける「NEC the WISE」
第四次産業革命を支えるインダストリー・オペレーション

NECが誇る最新のAI技術
リアルタイム監視を実現する動画顔認証技術
社会インフラの保全を効率化する光学振動解析技術
IoTの活用を広げる物体指紋認証による個体識別
未知のサイバー攻撃を自動検知する自己学習型システム異常検知技術（ASI）
防犯カメラ映像から未登録の不審者を見つけ出す時空間データ横断プロファイリング
きめ細かなマーケティングの実現に向けた顧客プロフィール推定技術
要因分析エンジンを用いた工場・プラントでの品質管理
予測から意思決定へ ～予測型意思決定最適化～
REFLEXによるバス運行の動的最適化

最先端のAI技術開発におけるNECのオープンイノベーション活動
脳の「ゆらぎ」を応用した超低消費電力のコンピュータで「おもろい社会」を実現
アナログ回路の活用により本物の脳を再現する「ブレインモルフィックAI」とは
AIとシミュレーションを組み合わせ、データに乏しい状況でも意思決定を可能に

AI技術ブランド「NEC the WISE」



Vol.69 No.1
（2016年9月）

特集TOP