Cylance Japan株式会社

従来製品の限界を破った「AIアンチウイルス」の革新的な防御力の秘密

ランサムウェアを含む未知のマルウェアにも防御率99.7%――そのわけとは？

近年、使い捨てるように新種のマルウェアを作り続ける攻撃者を前に、各社アンチウイルス製品は動的な検知技術を取り入れるなど改良を重ねてきた。だが、そうした対策も限界が訪れてきているほか、運用も複雑化・肥大化し、ユーザーも負担を強いられている。しかし、この状況を一変させる新たな技術が登場した。人工知能（AI）を用いたアプローチである。

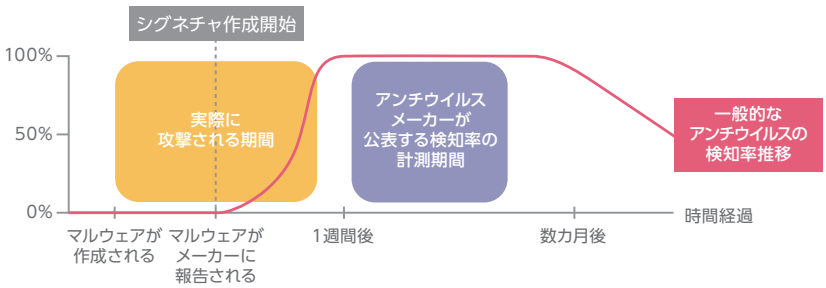
使い捨てマルウェアの台頭 従来の対策はもはや限界に

アンチウイルスソフトウェアは20年以上の長い歴史を持つが、検知の根本的な仕組みは長年に渡って大きく変わっていない。マルウェアのパターンファイルを用いた「シグネチャ方式」である。これはいわばブラックリスト方式であり、登録されているマルウェアをほぼ確実に識別できる反面、登録されていないマルウェアは見落としてしまう。つまり、新種のマルウェアが出てきた場合、アンチウイルスのベンダーが検体（実際のマルウェア）を入手してパターンファイルを作成・登録、そしてユーザーに配信するまでは検知できず、数日のタイムラグが生じてしまうことになる。

10年ほど前まであれば、それでも大半のマルウェアを防ぐことが可能だった。マルウェアを一から作って拡散させるには、それなりのスキルや時間が必要だったため、新種のマルウェアはそれほど頻繁には登場してこなかったのだ。おかげでアンチウイルスソフトウェアのベンダーやそのユーザーにも時間的な猶予があった。

しかし、近年では新種のマルウェアが登場するペースは桁違いに早まっている。手軽にマルウェアを作成できるキットやツールが広く出回り、攻撃者はこれを利用して新種のマルウェアを次々に作り出しては使い捨て、アンチウイルスベンダーがパターンファイルを作成する前にすべての仕事を終える。つまり、攻撃、感染、情報の奪取、そして場合によってはビットコインでの回収までを終えている。数日後にはシグネチャが配布され検知率が上昇するものの、その頃になると攻撃者は別の新たなマルウェアを武器としているはずで、実質的な防御力はほとんどなくなっているのである（図1）。

一般的なシグネチャ方式のアンチウイルス製品における検知率の推移



シグネチャベースで未知のマルウェアを止めることができないのは構造的な問題。

実際に攻撃で使われている期間に防御できるかどうかが一番のポイント。

図1：マルウェアが作成されてからアンチウイルスベンダーが検知できるようになるまではタイムラグがあるため、検知率が高くなる頃には、そのマルウェアは攻撃には使われなくなる。

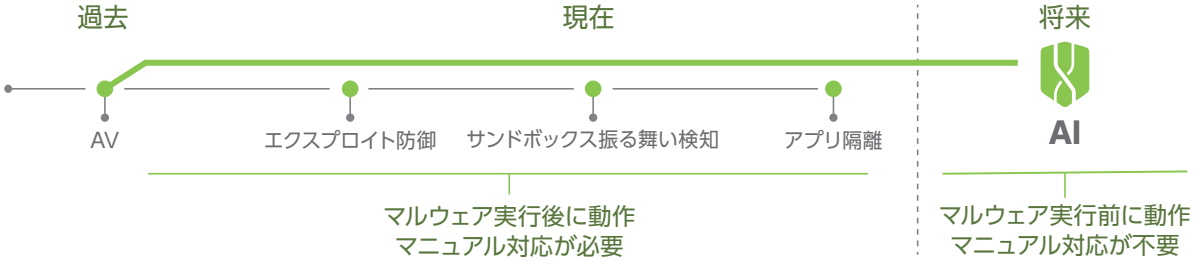


図2：アンチウイルスの手法は変遷してきたが、サイランスは安心できる静的分析を確立することで、複雑化・肥大化したアンチウイルスの仕組みを再びシンプルにすることができるようになると提唱する。

とはいえ、セキュリティベンダーもただ手をこまねいているわけではない。近年ではシグネチャ方式の欠点を補う形で、多種多様な技術を取り入れてきた。

その主な方向性は、「挙動からマルウェアかどうかを判定する」というものだ。例えば「スパイウェア通信検知」「振る舞い検知」「サンドボックス」などは、いずれもコンピュータそのものに害を及ぼさないような形で対象プログラムの挙動を調べることで、悪質なものかどうかを判定している。「エクスプロイト保護」も、OSやアプリケーションの脆弱性を突こうとするのを防ぐものだが、やはり主にプログラムが実行されて該当の挙動が確認されてはじめて検知ができる。

これら動的分析と呼ばれる検知手法は、静的分析であるシグネチャ方式よりもコンピュータに負荷を与えがちになることが課題だ。また、正常なプログラムをマルウェアと判定してしまう誤検知も生じやすいということあり、現場のオペレーションにおいては、誤検知で業務に支障がないようにチューニングしたり、アラートが出た際にセキュリティ技術者が1件ずつ確認したりするなど、運用面の負担も生じている。

マシンに負荷の少ない静的ファイル解析かつ高精度の検知を可能にした方法とは

そもそも数年前にサンドボックスを代表とする動的分析に基づくマルウェア検知手法が流行したのは、ひとえに静的分析であるシグネチャ方式の検知率が著しく低いことに起因する。逆に言えば、負荷が少ない静的分析のアプローチで、信頼できる高精度な検知技術があれば動的分析に頼る必要はなくなるはずだ。

それを実現することはもちろん簡単なことではない。しかし最近では、「AI」を利用して従来とは全く異なる技術でエンドポイントのアンチウイルスを提供する企業がある。この“AIアンチウイルス”を手掛けるのは、大手セキュリティベンダーで活躍していた技術者らが2012年に米国で創業した「Cylance（サイランス）」である。

サイランスの主力製品「CylancePROTECT」は、エンドポイントすなわちPCやサーバをマルウェアから保護するためのソフトウェアだが、パターンファイルを使用していない点が既存のアンチウイルスと大きく異なっている。1つのマルウェアに1つの処方箋をあてがう代わりに、AIの技術を用いてマルウェアファイルの概念モデルを作成、そしてファイルがそのモデルに近いかどうかをスコアで算出し、予測して防御するという。どのような仕組みで脅威を検知しているのか、以下にその仕組みを見ていこう。

まず、CylancePROTECTのマルウェア検知の要となっているのは、AIの一分野である機械学習技術を用いて実際のマルウェアの特徴を学習して作り出した、「データモデル」という数理モデルである。機械学習技術というのは、一例を挙げると、さまざまな写真の中から「猫」「犬」「車」「人」など特定の種類を見つけ出すような画像認識の分野でも活用されている。学習の中で、形・色・パーツなどの「特徴点」を基にあらかじめラベルが貼られた教師データを与えて学習していくことで、それを識別するための数学的な概念を作り出していく。そこで作られたモデルを基に、見たことがない新しい写真に対しても推論による予測判定ができるという仕組みだ。

サイランスでは、「Infinity」という名の機械学習システムをクラウド上に構築・運用しており、そこに教師データとして、マルウェアファイルを含む10億個ものファイルを投入し、マルウェアファイルとそうでないファイルの構造を学習させている。こうすることで、マルウェアファイルを判断するためのデータモデルを作っていくのである。

学習によってできあがったデータモデルの内容は、ひとまとまりの数式群のようなものである。そこに必要なパラメータを投入して計算することで総合的なスコアを算出し、マルウェアファイルかどうかの判定を行う（図3）。サイランスではファイルデータから最大700万もの特徴点を分析し脅威度を判定しているが、これらはすべて対象ファイルを実行することなく静的なスキャンにより得ることができる。データモデルを用いた計算を合わせてもコンピュータの負荷は軽く、動的分析はもちろんのことシグネチャ方式のスキャンと比較しても、高速に実行できるという。

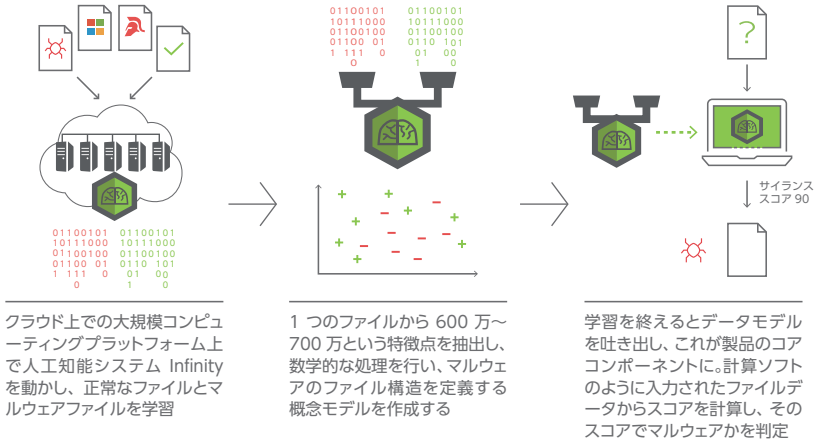


図3：Infinityでのデータモデル作成と、CylancePROTECTでの検知プロセス。

未知のマルウェアにも99.7%の防御率を実現

CylancePROTECTが持つデータモデルは、いわば膨大なファイルをInfinityが学習して得た「マルウェアというものの概念」といったところだろう。学習を手助けするのは世界でトップクラスのデータサイエンティストと呼ばれる数学者のチームだ。Infinityは大量の教師データに加えて、優秀なデータサイエンティストのガイドにより、高精度にマルウェアかどうかを判断できるデータモデルを作成しているのである。どんなに優秀なマルウェア解析官でもファイルを分析してマルウェアかどうかを判断するにはある程度の時間を要するが、データモデルによる推論判定は一瞬でその判断のスピードと正確さは人間の能力を超える。

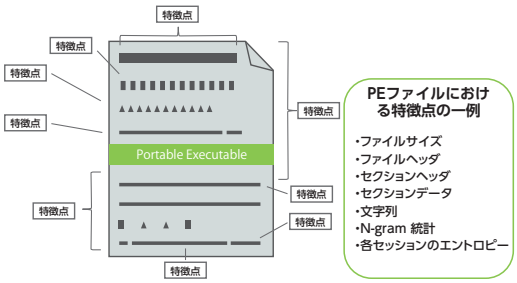


図4：CylancePROTECTのデータモデルは、現在のところ実行可能ファイル（WindowsではPortable Executable：PEファイルと呼ばれるものが主流）を対象としており、ファイルのさまざまな特徴点を解析に用いる。

しかも「マルウェアの概念」であるデータモデルは、既知のマルウェアだけしか識別できないシグネチャ方式とは大きく異なり、未知のマルウェアに対しても安定して高い精度を発揮する。サイランスでは99.7%※の防御率を謳っているが、これは未知であろうと既知であろうと差がない。その自信は、同社が各地で開催する巡回デモイベント「Unbelievable Tour」にも表れている。毎回イベントでは、24時間以内に取得したばかりのマルウェアの検体をCylancePROTECTおよび競合アンチウイルスでスキャンして比較し、未知のマルウェアにも強さを発揮することをライブデモで実施して話題となっている（最新のイベント情報：https://www.cylance.com/jp）。

未知のマルウェアでも高い防御力を示すにもかかわらず、CylancePROTECTのデータモデルの更新頻度は約半年から10カ月に1度ほどだという。IT部門にとっては、社内の各端末に最新パターンファイルが適用されてなおかつフルスキャンが行われているかどうかを日々チェックする負担がなくなるほか、インターネットから隔離されたネットワーク内やオフライン状態のエンドポイントについても配布管理の手間を省けるのが大きなメリットだ。仮想デスクトップ環境でしばしば問題になる「スキャンストーム」の心配もない。

※2017 NSS Labs Advanced Endpoint Protection Testingの結果より

WannaCryも18カ月も前から防げていた

CylancePROTECTの実力は、2017年5月に爆発的に流行し、世界各地で大きな混乱を引き起こしたランサムウェア「WannaCry」の一件を見

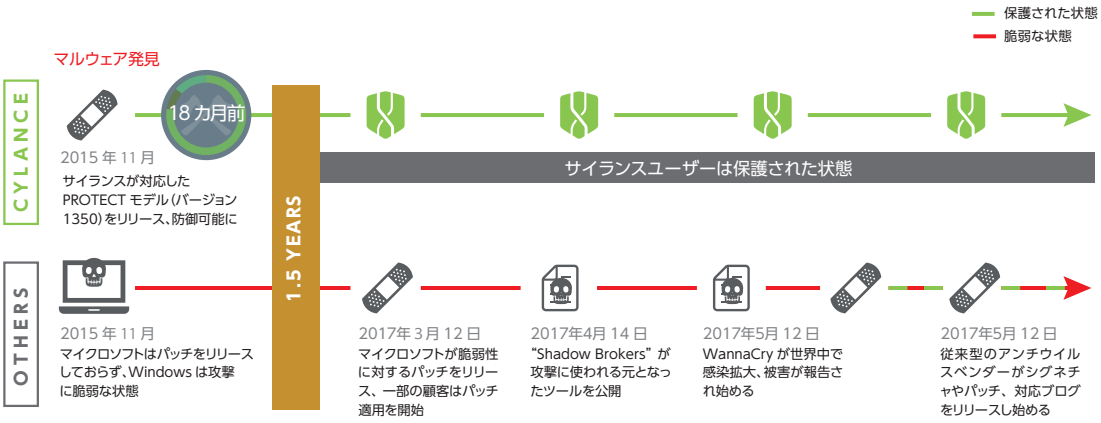


図5：WannaCryへのCylancePROTECTの対応

ても明らかだろう。他社アンチウイルスでは感染拡大を受けてからベンダーが対応を開始していたような状況だったのに対し、CylancePROTECTは即座にWannaCryをマルウェアと判定していたのだ。驚くべきことに、18カ月前の2015年11月にリリースした同社のデータモデル（バージョン1350）を使った実験でも、同じくWannaCryを検知できたという（図5）。つまり未来の脅威を予測して阻止していたことになる。

同じような事例は、米連邦政府の人事管理局（Office of Personnel Management：OPM）における情報流出事件の報告書にも見受けられる。OPMでは、2015年にサイバー攻撃を受けていたことが判明し、最終的には2000万人超もの個人情報が流出したと言われているが、CylancePROTECTはたまたま評価導入されたOPMの環境で問題のマルウェアを検出し、その後プロフェッショナルチームを派遣してインシデントレスポンスを行うなどの事後対策に貢献しており、その実績は米国議会の調査報告書により明らかになっている。

もちろんインシデントの渦中に限らず、平時での検証でもCylancePROTECTの評価は高い。さまざまな第三者機関による評価レポートでも、特に検体の鮮度が高いマルウェアや亜種を使った検証ほど他を圧倒する結果を出し、数々の賞も獲得している。大手アンチウィルスメーカーの最新バージョンや、高い検知率を謳う新進気鋭のエンドポイント製品でも、実際の検知試験では検知率が軒並み低くなってしまうのに対し、CylancePROTECTは厳しい条件の試験でも抜群の安定した実績を記録している。

稼働実績は世界で1000万台超
業界問わず国内ユーザーも増加中

CylancePROTECTは、2014年に出荷が開始されたばかりのまだ新しいセキュリティ製品だが、早くも稼働実績1000万台超と業績は急拡大中だ。2016年には日本オフィスを設立、その有用性が話題となり国内でも有力ディストリビュータが次々に取り扱いを開始しており、セキュリティ意識の高い国内大手企業でも導入が進んでいる。官公庁、金融、製造、サービス、教育と特定の業界によらず、幅広い分野で採用されているのも特徴だ。

ちなみに、アンチウイルス製品の多くはOS上で競合し合うことから同一エンドポイント上に共存できないが、CylancePROTECTは既存アンチウイルスと異なる仕組みで検知するため、他のアンチウイルス製品と共存させることができる。既存の製品を入れたまま、CylancePROTECTを追加して並行稼働させ、その効果を検証しつつ移行するといったプロセスも可能だ。いきなりすべてをリプレースすることが不安であれば、このような形でまず試しに使ってみるのも得策であろう。

CYLANCE
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
お問い合わせ　TEL. 03-6386-0061（代）
www.cylance.co.jp