セキュリティ製品の妥当性点検に向けた アンチウイルスソフト検知率評価システムの提案及びその有効性の検討

北原 美里, 米谷 雄介, 後藤田 中, 小野 滋己, 青木 有香, 八重樫 理人,
藤本 憲市, 林 敏浩, 今井 慈郎, 最所 圭三, 喜田 弘司
香川大学
s16t225@stu.kagawa-u.ac.jp

Proposal of anti-virus software detection rate evaluation system for validity check of security products and examination of its effectiveness

Misato Kitahara , Yusuke Kometani , Naka Gotoda , Shigemi Ono , Yuka Aoki ,
Rihito Yaegashi , Kenichi Fujimoto , Toshihiro Hayashi , Yoshiro Imai , Keizo Saisho ,
Koji Kida
Kagawa Univ.

概要

近年のセキュリティ事故はマルウェア感染がきっかけになることが多く、マルウェア対策は緊急課題である。マルウェアは日々進化し、多種多様に開発されるためそれら（未種・亜種）全ての対策を同時に実施することは困難である。これに対し、香川大学では、導入したセキュリティ製品の妥当性評価のための調査を行っている。本稿ではまず、3 年に及ぶ手作業による本調査経験をふまえ、妥当性点検の課題を明らかにする。我々は、これら課題を解決するために、自動化するアンチウイルスソフト検知率評価システムを提案・開発した。本稿では、提案システムの試作およびその運用結果について述べ、その有効性を考察する。

1 はじめに

近年のセキュリティ事故はマルウェア感染がきっかけになることが多く、マルウェア対策は緊急課題である。マルウェアは日々進化し、多種多様に開発されるため、それら（未種・亜種）全ての対策を同時に実施することは困難である。一方、アンチウイルスソフトウェアも多くのベンダーから提供されており、各ソフトウェアのパターンファイルも日々更新される中、どのソフトウェアが自分たちの組織に適しているのか不明であり、調べる必要がある。香川大学では、人手でこの作業を行ってきたが、仮にこれが自動化できた場合に、対応したソフトウェア数や調査期間、パターン更新の特性など従来調査できなかった情報を細かく得ることができる。本稿ではこれの自動化を提案する。

2 導入製品に対する点検の課題

2.1 手作業による妥当性点検

本学では、標的型攻撃を受けると、ファイアウォールのサンドボックス機能が検知し、アラートメールがセキュリティ運用者に通知される。妥当性点検のために、アラートメールに記載されているハッシュ値を VirusTotal [1] のフォームに入力する。ここで、妥当性点検とは、製品比較を行い、導入した製品が予算・運用形態などの制約の中で十分な性能を維持できているかを点検することである。VirusTotal では、入力されたハッシュ値に基づき、VirusTotal が検体を提供している各社のアンチウイルスソフトウェアの対応状況を一覧として確認できる。アンチウイルスソフトウェアはベンダーによって、対応期間に差が生じると仮定されるため、上記操作を継続的に繰り返し、対応状況の時間変化を調査する必要がある。

2.2 香川大学における妥当性点検の現状

各セキュリティベンダーの対応状況を以下のルールにより分類を行っている

- 「リアルタイム」: サンドボックスにおいて当該検体が初めて検知されたタイミングで、セキュリティベンダーでは既に検知されているもの
- 「後日更新」: 数週間後に確認した際に対応状況が検知済に変化したもの
- 「検知なし」: マルウェアとして検知されていないもの
- 「偽陽性」: 初期段階では誤検知が多いという仮定を置き、初めて検知されたタイミングで 1 社だけ検知となっているもの

以下の図 1 は、2018 年度における妥当性点検の結果を棒グラフで可視化したものである。縦軸は、各セキュリティ製品において、「リアルタイム」「後日更新」「検知なし」「偽陽性」に分類された検体の数を表す。「リアルタイム」の検知数や「後日更新」を含む総検知数はセキュリティ製品によって異なることがわかる。「リアルタイム」の割合が 100%であったセキュリティ製品はみられず、サンドボックス検出時（本学へ着弾時）にパターン定義で 100%検出可能アンチウイルスソフトは皆無であるといえる。また、アンチウイルスソフトにおいては「後日更新」となっているマルウェアが存在することに対して、サンドボックスでは検知できていることから、サンドボックスの効果は確認可能であるといえる。

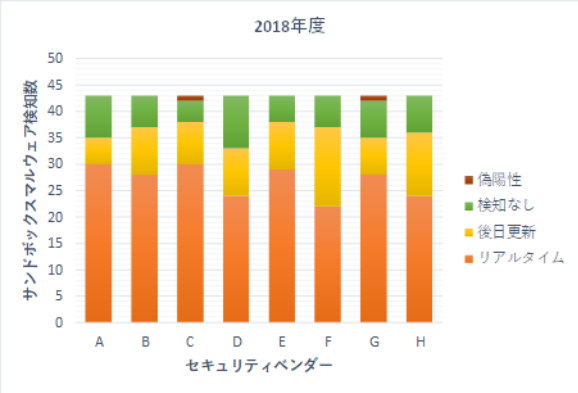

図 1　2018 年度の各社の対応状況

2.3 妥当性点検の課題

前節の妥当性点検は、セキュリティ運用者が他の業務を兼ねながら負担のない範囲で一定の期間をおいて継続的に調査している[2]。このため以下の課題があげられる。

課題 1: 現状の人手による方法では、休日に通知されたサンドボックスからのアラートについては休日明けに調査することになるので、どのアンチウイルスソフトウェアが最初に対応できたかが判断できない。例えば、1 回目の調査が遅れた場合、1 番最初に対応したアンチウイルスソフトウェアがどれなのか見逃してしまうことがありうる。

課題 2: 現状は後日確認のタイミングが着弾時から約 2 週間置いて実施されている。この間の各セキュリティベンダーの対応の早さについては評価できておらず、対応した順位が判断できない。これは人手による調査であるので細かく調査できず対応状況の時間変化が正確に判定できないためである。

課題 3: VirusTotal の各パターンファイルがまだ対応していないことにより、判定結果が不正確な場合があり、これを利用したアンチウイルスソフトウェアの検知精度の評価結果も不正確になる場合がある。現状は偽陽性の判断を課題 2 で述べた通り、2 週間の期間を経て再調査により行っているが、最終的により細かく（毎日）、より期間を 2 週間よりも長く（1 か月）調査することで、より正確な評価結果が得られる。

課題 4: 近年、サンドボックスの性能向上により、メールに添付されている URL も検査対象となり、URL にアクセスしたサイトにマルウェアが仕込まれている場合も検知してくれる。そのようなフィッシングサイトについては VirusTotal に検体が上がっていない場合があり、VirusTotal で判定できない。ただし、例えば後日、フィッシングサイトへ誘導するメールが本学以外の組織に着弾することをきっかけに、VirusTotal の判定基準が更新され、対応する可能性がある。このため VirusTotal でのチェックを継続的に行う必要がある。

3 課題解決のアイデア

以下の 3 つの機能で、前節の課題を解決する。

- リアルタイム初期調査機能: 一定期間毎にアラートメールが届いたかどうかを確認し、届い

ていれば調査する。香川大学では、1 日に最大 10 件程度のアラートメールが来ることもある。ただし、VirusTotal の更新間隔は 1 日単位で管理されることから、本システムも 1 日毎の調査に設定した。より具体的には、1 日の区切りとして 0 時に調査を行う設定としている。

・継続調査のタイミング調整機能：検知件数の推移が見られないまま一定期間が経過すると判定を行わない設定とすることを考えている。本設定を行わない場合、すべてのハッシュ値を判定することになり、時間がかかってしまうからである。我々の試用では 4 時間かかったこともある。

・偽陽性判定機能：複数のアンチウイルスソフトウェアのうち 1 社だけがマルウェアと判定し、継続調査の後でも結果が変わらない場合は、その 1 社の判定結果を誤りとする。

4 開発中のシステム

前節の機能の実現に向けて、まずはハッシュ値に対応する各社アンチウイルスソフトウェアの対応状況を収集する基本システムを開発しテストした。本システムの機能を目標とするシステムとの違いを示すため、便宜的に検知情報抽出機能、対応状況収集機能と呼ぶ。図 2 に検知情報抽出機能のアルゴリズム（検知情報抽出アルゴリズム）、図 3 に対応状況収集機能のアルゴリズム（対応状況収集アルゴリズム）を示す。

4.1 検知情報抽出アルゴリズム

事前にサンドボックスからアラートメールが届く専用のアカウントを用意しておく。まず、このアカウントにログインし（STEP-1）、新規のアラートメールが届いていればメールの本文を受信する（STEP-2）。本メールには、マルウェアであると判定した根拠の情報や、ファイルのハッシュ値や、詳細な分析結果を記したウェブページへのリンクなどさまざまな情報が記述されている。この中からハッシュ値をテキスト解析により抽出する（STEP-3、4）。ここで、サンドボックスからのアラートメールは、同じハッシュ値でも繰り返し送られてくることがあり注意が必要である。これに対し、同じファイルが繰り返し検知される場合に備えて、すでに受信したことがあるハッシュ値を CSV ファイルに保存しておき、本 CSV ファイルを使って一度届いたことがあるかどうかをチェックすることで対応する（STEP-5）。本処理を定期的に繰り返す（例えば 1 時間毎）ことによりサンドボックスで検知されたファイルのハッシュ値が重複することなく CSV ファイルに保存される（STEP-6）。

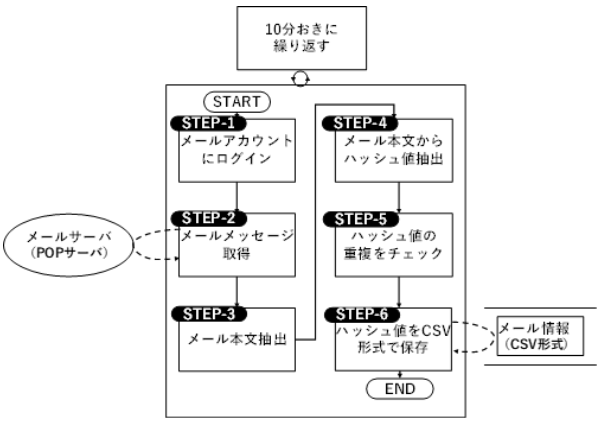

図 2　検知情報抽出の処理フロー

4.2 対応状況収集アルゴリズム

前記の CSV ファイルからハッシュ値を読み込み(STEP-1)、VirusTotal で判定する(STEP-2)。ただし、この判定は VirusTotal の API を用いるが、この API は、1 分間に 4 回しか使用できない制限があるため、毎回 API を利用するごとに適宜スリープを入れることで対応する。VirusTotal では、アンチウイルスソフトウェアごとに検知の有無やパターンファイルのアップデート日時が json 形式で返ってき、CSV ファイルに保存する(STEP3、4)。なお、アンチウイルスソフトウェアのバージョンや判定結果の検出名も返ってくるが、使用しないためその項目は削除する。

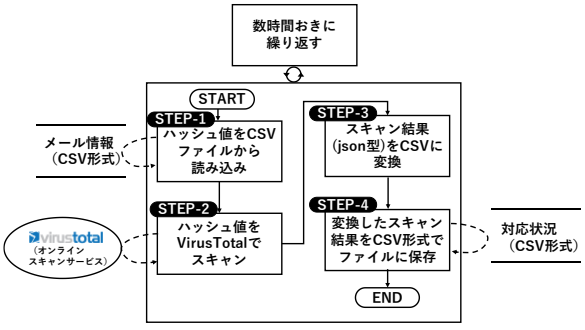

図 3　対応状況収集の処理フロー

4.3 対応状況の時間変化を分析

判定結果を保存した CSV ファイルは Excel ファイルと関連付けられており、運用担当者が蓄積された対応状況情報を任意のタイミングで Excel 上にインポートすることができる。読み込んだ判定結果は、ピボットテーブルの集計表と連携しており、クリックのみで集計結果を更新することができる。担当者は、集計結果を別のワークシートにコピーし、分類を行ったり、図 4 のような時系列でグラフ化したりするなど、対応状況の変化を分析することができる。本システムは 2019 年 5 月 23 日から稼働を開始している。

現在は、1 日毎に調査する設定だがこれをより短い時間にしていくことでリアルタイム初期調査機能が実現できる。また、図 4 に示したように、あるファイルに対する検知件数は時間経過後に一定値に収束することが分かる。このような変化をもとに調査を継続すべきか判断でき、継続調査判定機能も実現できる。これに加えて、最終的に陰性と判定されたファイルについて遡って検知時点の各社の判定結果と突き合わせることでマルウェア検知当初にどの製品が偽陽性と判定したかを知ることができ、偽陽性判定機能も実現できる。

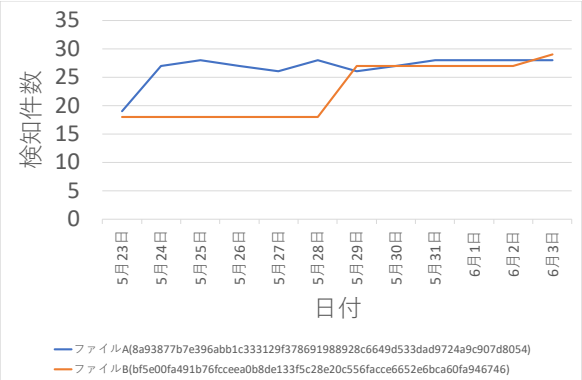

図 4　時系列グラフ

5 有効性の考察

2 章の課題 1 から課題 4 に関して、本システムの効果を以下に考察する。

課題 1:システム化したことにより、現状の人手による方法とは異なり、休日に関わらず 1 日毎に調査が実施され、休日に通知されたサンドボックスからのアラートについても解析が可能となった。これにより、どのアンチウイルスソフトウェアが最初に対応できたかが判断できる。

課題 2:本システム導入により、人手による方式とは異なり、毎日のチェック結果が蓄積される。このため、各セキュリティベンダーの対応の早さを評価でき、対応し順位が判断できる。

課題 3: 本システム導入により、人手による方式とは異なり、偽陽性の判断を毎日かつ 2 週間以上の長期に渡り、調査をすることができる。このため、偽陽性の判断をより正確に実施できる。

課題 4: 本システム導入により、人手による方式とは異なり、VirusTotal で即時判定できないようなフィッシングサイト等のケースにおいても、VirusTotal で対応できるまで調査を継続することにより、本ケースも対応可能である。

6 おわりに

本システムはマルウェアが届いた瞬間に調査を開始でき、継続調査によりアンチウイルスソフトの対応状況の時間変化が分かるだけでなく検知結果の誤りも評価できることが特徴である。これらの情報はアンチウイルスソフトの妥当性点検の判断材料として活用する。また、今回はこれまでの妥当性点検の手作業における課題を解決するために VirusTotal から返された「検出有無」のデータを用いた。他にもアンチウイルスソフトのバージョンやサンドボックスによるマルウェアの分類結果の情報を得ることができる。これらの情報を活用することでバージョンや分類ごとの検知率の違いなどより詳細な分析が可能である。今後はこれらの情報を含めた可視化についても検討したい。

参考文献

[1] VirusTotal 、 https://www . virustotal . com/ja/ （参照日:2018 年 09 月 03 日).

[2] 小野 滋己、後藤田 中、米谷 雄介、青木 有香、八重樫 理人、藤本 憲市、林 敏浩、今井 慈郎、最所 圭三、“パターン定義に要する対応期間の調査に基づくセキュリティ製品の妥当性点検 “、大学 ICT 推進協議会 2018 年度年次大会（AXIES2018）論文集、WA1-3、2018 年 11 月 21 日.