攻撃の痕跡に着目するサイバー攻撃対策の最前線

岩村 誠 /鐘本 楊
黒米 祐馬 /青木 一史
川古谷 裕平 /折原 慎吾
三好 潤
NTTセキュアプラットフォーム研究所

近年のサイバー攻撃は，巧みに標的を騙すことでマルウェアを企業網内に潜り込ませる傾向が強まっており，感染を未然に防ぎ切ることが困難になりつつあります．一方，外部に公開しているWebサーバなどに対しては，攻撃手法がよく知られるにつれ，攻撃頻度が上昇してアラートが多発し，どれから対処すべきか判断に困るようになりつつあります．本稿では，このような状況に打ち勝つための，攻撃時に残される痕跡に着目したサイバー攻撃対策の最前線の取り組みを紹介します．

エンドポイント防御の現状

企業をねらうサイバー攻撃や，そこで用いられるマルウェア（悪性なソフトウェア）は日々高度化しており，未然に攻撃の侵入を防ぐことが難しくなってきています．そうした中，マルウェアによる侵入を許してしまうことは前提とし，侵入後の対処を考慮したEDR（Endpoint Detection and Response）と呼ばれる技術が注目を集めています．

従来のセキュリティ製品は，マルウェアが実行される前にその外見上の特徴（マルウェアの実行ファイルに含まれるパターン等）をルールとして検知することで感染を未然に防いでいました．しかし，昨今のサイバー攻撃では，その外見上の特徴を変化させ，セキュリティ製品による検知を逃れるマルウェアが用いられるようになってきました．マルウェアの外見上の特徴は，比較的簡単に変化させることが可能です．一方で，感染後の振る舞いはマルウェアが行いたいことと密接に関係しており，外見上の特徴と比較して変化させることが難しいと考えられています．現在注目を集めているEDRは，マルウェアが動き出した後の振る舞いや，それらが残す痕跡を検出することで，こうしたサイバー攻撃に対抗しようとしています．

マルウェアに感染した際に残る痕跡を検出するルールはIOC（Indicator Of Compromise）と呼ばれ，EDR製品によっては利用者が独自につくったIOC（カスタムIOC）によって，マルウェア感染を検知することが可能になっています．以下では，マルウェア感染の痕跡とそれを検出するIOCの生成方法について解説します．

マルウェア感染の痕跡とその検出

ここで，ある端末にマルウェアが感染した際に“mal_a.txt”という名前のファイルが痕跡として残ったとしましょう．この痕跡を検出するには，ファイル名が“mal_a.txt”というIOCを用意すれば良さそうです．ただ，同じマルウェアがほかの端末に感染した際にはファイル名が“mal_b.txt”となる場合，元のIOCでは検出できなくなってしまいます．そこで少し工夫をして，ファイル名が“＊.txt”（＊は任意の文字列）というIOCを用意したとします．すると“mal_a.txt”も“mal_b.txt”も，もしかすると今後出てくるかもしれないほかの痕跡についてもカバーできそうです．ただ，EDRで監視している端末ではマルウェアではない通常のアプリケーションも動いています．もし，ある通常のアプリケーションが“leg.txt”というファイルをつくった場合，“＊.txt”というIOCではそのファイルをマルウェアの痕跡として検出してしまいます．従来技術が着目する外見上の特徴よりは変化しにくくなったとはいえ，IOCにも痕跡の変化に追随できるようにカバー率を上げつつ，誤検出を起こさない表現が求められます．

もう1点，IOCに求められることを考えるにあたって考慮しておくべきことがあります．実際にIOCでマルウェアへの感染が発覚したとしましょう．その後の対策の多くはセキュリティ技術者，つまり人間に委ねられることになります．マルウェアはどういった経路で侵入してきたのか，機密情報を外部に送信していないか，ほかに感染している端末はないか，これらを残されたログなどから解明することになります．時には，IOCが検知したものが何であるかを把握し，検知したIOCを改良してほかの端末を検査する必要性も出てくるでしょう．そのときに必要になるのは，人間が見て解釈しやすいIOCです．ある種の機械学習ではその

検知基準が非常に複雑で，改良することはおろか，理解することすら困難なアルゴリズムも存在します．一連の作業フローの中に人間が存在しているセキュリティの現場では，IOCの解釈性も重要になってきます．

IOCの自動生成

NTTセキュアプラットフォーム研究所では，さまざまな解析妨害機能を持ったマルウェアに対しても，その振る舞いを網羅的に抽出するマルウェア解析技術の研究開発に取り組んでいます．ここで紹介するIOCの自動生成技術[1]は，このマルウェア解析技術により抽出された挙動ログを入力として，高い検出精度とカバー率，解釈性を兼ね備えたIOCの生成を実現しています．具体的には以下の手順によりIOCを生成します（**図1**）．

① マルウェアの収集・選定：IOCによる監視を実施する環境に合わせたマルウェアを収集・選定します．

② マルウェア解析技術により挙動ログを抽出：マルウェア解析専用の仮想環境にてマルウェアを解析し挙動ログを抽出します．これまで培ってきたマルウェア解析技術はここで活用されています．

③ マルウェアの挙動ログから複数の抽象度のIOC候補を生成：IOCの候補となり得るさまざまな抽象度の正規表現を，過去のマルウェア解析ノウハウ等から生成します．

④ 検出精度・解釈の容易さを踏まえた最適なIOCセットの算出：前述のIOC候補について，正規ソフトウェアおよびマルウェアの挙動ログを基に，検出精度・解釈の容易さを踏まえ，各マルウェアファミリに対応する最適なIOCを算出します．

こうして生成されたIOCを市中のEDR製品に対して追加投入することで，これまで発見が難しかったマルウェア感染端末を検知することができるようになります（**図2**）．現在は1週間当り約1万検体を収集・選定し，それらのマルウェアの解析結果から生成したIOCのNTTグループ内への配信を始めています．今後はスクリプト形式など，実行ファイル形式以外のマルウェアへの対応を進めていきます．

公開サーバ防御の現状

続いて，ここからは話題が変わり，外部公開サーバに対するサイバー攻撃対策について解説します．

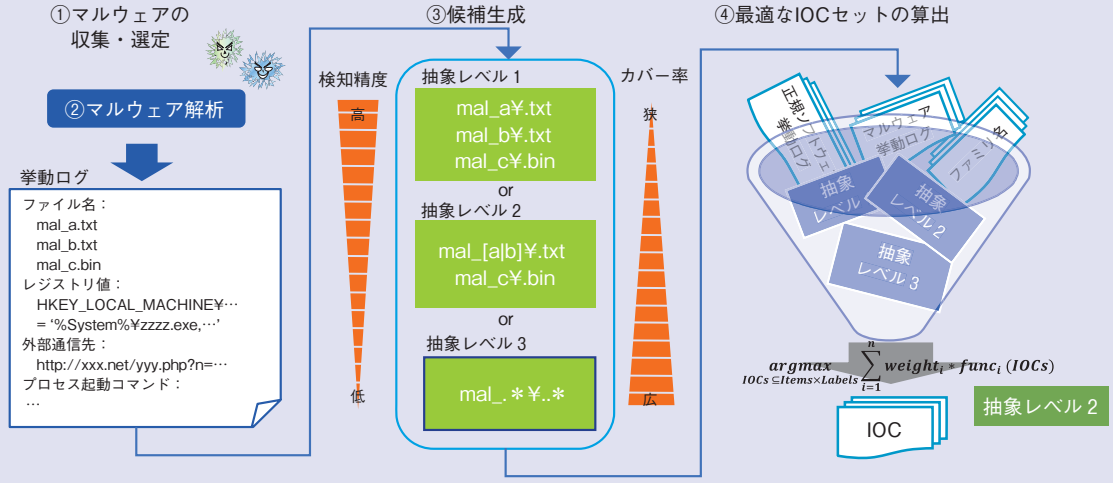


図1　IOC生成手順

図2　カスタムIOCの活用

新たな脆弱性がサーバやアプリケーションで発見されるたび，その脆弱性をねらうサイバー攻撃が発生します．サーバやアプリケーションの脆弱性を悪用するサイバー攻撃は世界で1000万件／日を超える規模となりました．これらの攻撃を検知・遮断するためにIPS（Intrusion Prevention System）[*1]やWAF（Web Application Firewall）[*2]といったセキュリティ機器を導入することが一般的となってきています．これらのセキュリティ機器によってすべての攻撃を正しく検知し，遮断できることが理想ですが，現実にはなかなか難しいのです．理由は誤遮断によるサービス品質低下というリスクがあるからです．セキュリティ機器が運用者によって十分にチューニングされていない場合，正常な通信を攻撃として誤って検知して遮断してしまう可能性があります．このリスクゆえに，すべての攻撃を遮断することは運用者によるチューニングがなければ難しいという問題があります．そして，攻撃を遮断するにしても誤検知でないことが確実な一部の限られた攻撃のみであったり，IDS（Intrusion Detection System）[*1]のように検知のみを行い通知するアラートを人手で対応しているのが実態です．

より効率的なセキュリティオペレーションの必要性

企業や組織でサイバー攻撃の対応にあたるのがCSIRT（Computer Security Incident Response Team）と呼ばれるセキュリティインシデントを専門に対処するチームや，アラートの通知に対応することを専門に行うSOC（Security Operation Center）アナリストです．CSIRTやSOCアナリストは日々，セキュリティ機器から通知されるアラートを基にセキュリティ侵害が発生していないか分析を行います．特にサーバに対する攻撃を検知するWAFやIDSは日々数千・数万のアラートを通知するため，侵害を分析する際はアナリストの知識や経験を基に，より重要なアラートを絞り込んで処理していかなければ発生するアラートの量に到底追いつくことができません．この絞り込みは知識や経験を持った数少ないアナリストのみができることであり，誰しもができることではないため，攻撃が大規模化している現在では完全に人手のみによってすべての攻撃を分析することは現実的ではありません．また，攻撃者もその実状を把握しているため，大量の無意味な攻撃を仕掛けつつ，攻撃の目的を達成する本命となる攻撃はたった一瞬だけ，といった戦術で挑むことも可能です．企業のセキュリティ監視を麻痺させ，CSIRTやSOCアナリストが気付いたときにはすでに時遅しということになってしまいます．そのため，CSIRTやSOCアナリストは常に不利

*1　IPS/IDS：脆弱性を悪用した攻撃からアプリケーションを保護するシステム．IDSは検知のみを行う利用形態を指し，IPSは検知した攻撃を遮断する利用形態を指します．
*2　WAF：IPS/IDSと同様に攻撃からアプリケーションを保護するシステム．Webアプリケーションに特化した検知能力を有します．

な戦いを強いられています．

アラートトリアージ技術

NTTセキュアプラットフォーム研究所では，ネットワーク通信からサーバに対する攻撃の成否を攻撃の痕跡から自動的に判定し，その攻撃に関連付くアラートが優先的に対応すべきものか否かを決めるアラートトリアージ技術[(2),(3)]を開発しました．攻撃の成否に着目してトリアージ（優先度付け）を行う技術は世界初です．この技術により，喫緊の対応が必要な攻撃のみに人手による分析を集中させることができるようになります（**図3**）．

本技術の根幹は次の3つの機能から成ります（**図4**）．

① 攻撃が成功した際に攻撃者がそのサーバに実行したいコードあるいはコマンドを抽出する機能

② 抽出した攻撃コードあるいはコマンドをさまざまなサーバを模擬したエミュレータ内で実行し，IOC と呼ばれる攻撃の痕跡を抽出する機能

③ エミュレータから抽出したIOCが実際の通信に発生しているかを確認し，発生していれば攻撃成功，発生していなければ攻撃失敗と判断する機能

本技術により，アラートが発生した際にそのアラートに対して，攻撃が成功している，攻撃が失敗している，攻撃の成否が不明といった情報を付加することができます（**図5**）．もしア

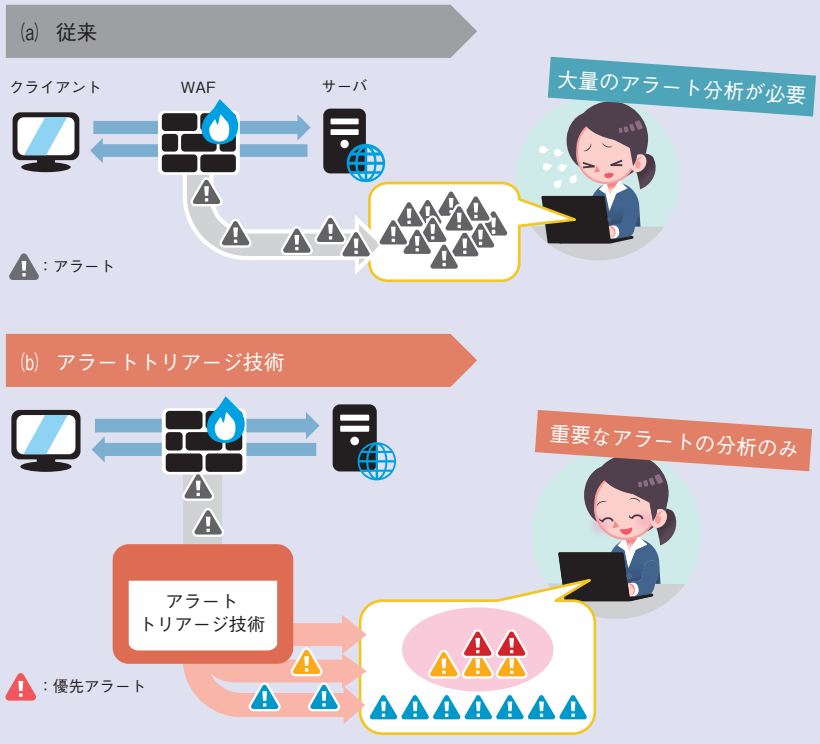


図3　アラートトリアージ技術

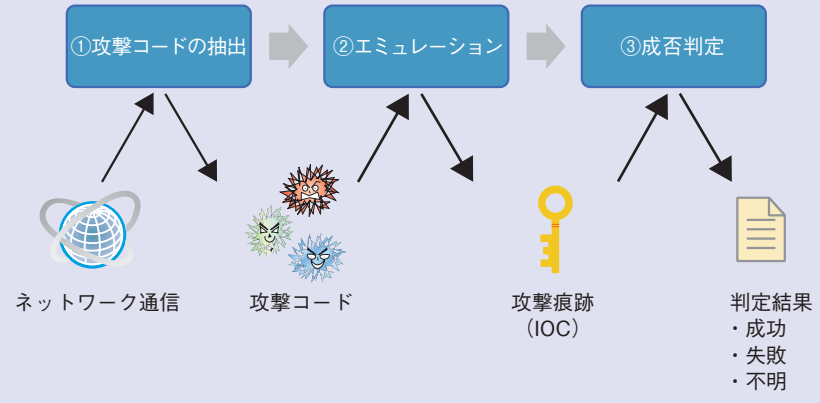


図4　アラートトリアージ技術の仕組み

ラートに攻撃が成功していると情報が付加されている場合，対応優先度は高く，ほかのアラートの確認を後回しにしてでも先にこのアラートを確認すべきです．逆にアラートに攻撃が失敗していると情報が付加されている場合，対応優先度は低く，ほかの攻撃の成否が不明なアラートを先に確認すべきということになります．本技術により優先的に対応すべきアラートが一目瞭然になり，アラート数が大量に増加する大規模なイベントやフォーラム，あるいは攻撃者によるキャンペーンがある場合，アラートトリアージ技術による効果はより顕著に現れると考えています．

実ネットワーク環境を用いた私たちの評価では約52%のアラートを正しく攻撃失敗として判断し，アラートの対応優先度を下げることを実現しました．また，大量のアラートに紛れたわずか0.1%の攻撃成功に関する重要なアラートの優先度を上げることを実現できました．攻撃被害がまだ少ない偵察段階で攻撃が成功していることに気付き，運用者に通知することで攻撃被害が拡大する前に対処を行うことができたという良好な結果を得ることができました．

時刻	アラート内容	送信元	送信先	成否判定※	対応優先度※
2019/12/2 21:38:12	Remote code execution attack detected	x.x.x.x	y.y.y.y	失敗	低
2019/12/2 21:43:03	SQL injection attack detected	x.x.x.x	y.y.y.y	成功	高
2019/12/2 21:43:15	Cross-site scripting attack detected	x.x.x.x	y.y.y.y	不明	中

※本技術により付加される情報

図5　アラートトリアージ技術の効果

今後の展開

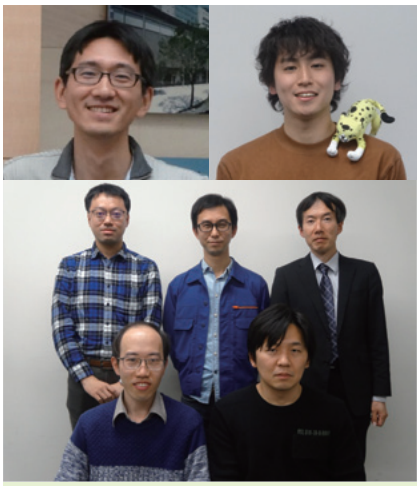
サイバー攻撃を未然に防ぎ切ることは現実的に難しいという現状を踏まえ，本稿では，エンドポイント端末のマルウェア感染や公開サーバへの攻撃の成否を，攻撃時に残される痕跡に着目して判定する技術を紹介しました．今後は，検知後の対応まで自動化する技術の研究に取り組み，ますますの高度化と増加が見込まれるサイバー攻撃に立ち向かっていきます．

■参考文献

⑴ Y. Kurogome, Y. Otsuki, Y. Kawakoya, M. Iwamura, S. Hayashi, T. Mori, and K. Sen: “EIGER: Automated IOC Generation for Accurate and Interpretable Endpoint Malware Detection,” ACSAC 2019, San Juan, U.S.A., Dec. 2019.

⑵ 鐘本・青木・三好・嶋田・高倉：“攻撃コードのエミュレーションに基づくWebアプリケーションに対する攻撃の成否判定手法,” 情処学論, No.60, Vol.3, pp.945-955, 2019.

⑶ Y. Kanemoto, K. Aoki, M. Iwamura, J. Miyoshi, D. Kotani, H. Takakura, and Y. Okabe: “Detecting Successful Attacks from IDS Alerts Based on Emulation of Remote Shellcodes,” Proc. of COMPSAC 2019, Vol.2, pp.471-476, Milwaukee, U.S.A., July 2019.



（上段左から）青木　一史/ 黒米　祐馬
（下段後列左から）折原　慎吾/
川古谷　裕平/
三好　潤
（下段前列左から）鐘本　楊/ 岩村　誠

2020年代を迎え，サイバー攻撃のさらなる激化が予想されます．NTTセキュアプラットフォーム研究所では，攻撃者優位の状況を抜本的に覆すことを目標に最先端の研究開発に取り組んでまいります．

◆問い合わせ先
NTTセキュアプラットフォーム研究所
企画担当
E-mail　scpflab@hco.ntt.co.jp