APT 攻撃を検知するための ファイルアクセスの記録と比較手法の研究

園田 真人 (1310037)

北陸先端科学技術大学院大学 情報科学研究科

2016 年 2 月 10 日

キーワード: APT, Security, Tree Structure, Detection, File System.

Advanced Persistent Threat(APT) 攻撃は対象の情報を窃取することを目的として明確な犯意を持って知能的な活動を行う標的型攻撃の一種である。APT 攻撃の特徴として 1. 特定の相手を対象として執拗な攻撃を行う 2. 相手に合わせて侵入方法を選択する 3. 侵入後に長期的な探索を行うことが挙げられる。この 3 つの特徴は、情報の機密性を維持するために高度なサイバーセキュリティ対策を実施している組織に対して情報窃取を行うためであると考えられる。APT 攻撃による被害例として、2010 年には google や yahoo!などの 30 以上の Web サイトを対象とした Operation Aurora や 2011 年には 70 以上の企業や政府機関などを対象とした Operation Shady RAT(Remote Access Trojan) による大規模な情報流出が起きている [1] [2] [3] [4]。APT 攻撃の情報の窃取により発生する大規模な情報流出は、損害賠償による有形資産の減少の可能性や無形資産におけるコーポレートブランドの低下といった甚大な被害を組織にもたらすものである。

APT 攻撃の活動段階は侵入段階、内部活動段階、情報持出段階の 3 つに大別される。侵入段階ではエキスプロイトコードの実行を目的としたスピアフィッシングが用いられる。スピアフィッシングは、脆弱性が報告されてから修正されるまでに侵入を行うゼロデイ攻撃またはヒューマンエラーから侵入を行うソーシャルエンジニアリングを利用するため、APT 攻撃における侵入を完全に防ぐことは困難である。内部活動段階では、侵入したシステムの環境を調べた後に、目的の情報を窃取するためにファイル/ディレクトリの探索が少ない頻度で長期的に行われる。この特性により IDS によるアノマリ検知を回避するため、内部活動を検知することは困難であるとされている。情報持出段階は、踏み台のサーバを経由して窃取した情報を外部に持ち出す段階である。情報の暗号化を行い通信プロトコルを偽造するため、ネットワークトラフィックから情報持出活動を検知することが困難である。また、情報持出段階では攻撃者が目的の情報にアクセスしていることから、この段階に到達した時点で情報流出が始まっている。よって、APT 攻撃による情報流出を防ぐためには、情報持出の段階以前で APT 攻撃を検知し阻止しなければならない。

本研究では、ファイルアクセスログ（FAL）から Tree Structured Log（TSL）を構築し、その TSL を記録し比較を行うことでアノマリ検知を行う fspeek を提案する。fspeek における APT 攻撃の検出手法として、ファイルアクセスのアクセス範囲を尺度としてアノマリ検知を行う File Access Scope T-test（FAST）を提案する。システムを常用する正規のユーザと攻撃者の間では、システム内におけるファイルアクセスの活動の傾向が異なり、活動傾向の差異がファイルアクセスログにおけるアクセスの範囲で表現できることを利用する。FAST では、ファイルアクセスのアクセス範囲を TSL の面積として定量化し、指定した区間における TSL の面積を時系列でから 2 つのグループに分けて、対応のある T 検定を行う。このとき攻撃者が内部で長期間活動している期間があれば、どちらかのグループの面積が大きくなることで T 検定での有意差が生じるためアノマリ検知が可能となる。fspeek は、FAST において長期間の TSL を比較するために、ユーザが利用しているシステム上で常時動作させることを前提としており、長期的に TSL を記録できるように記録するデータ量は FAL よりも小さい。

fspeek を利用して、2015 年 2 月から 11 月の 10ヶ月分の TSL に対して提案手法をもとに実験を行った。2 月から 6 月の read only TSL をグループ A、7 月から 11 月の read only TSL をグループ B とする。そして、2 月から 6 月の read only TSL に探索的なファイルアクセスを混ぜた TSL をグループ A_n とする。また、7 月から 11 月の read only TSL に探索的なファイルアクセスを混ぜた TSL をグループ B_n とする。ここで、B_n 及び A_n の n は read only TSL に混ぜる 1 日当たりのファイルアクセスの回数を示す。n=(1,2,3,4,6,8,12) として各 TSL の面積を計算し、帰無仮説にグループ間に差はないとおいて T 検定を行った。結果として、有意水準 α が 10%のときの t 検定 (B,A_3),(B,A_4),(B,A_6),(B,A_8),(B,A_{12}),(A,B_6),(A,B_8),(A,B_{12}) において有意差があることから、長期間な活動が行われたときに、その活動を検知することが可能であることが分かった。よって、ユーザのソフトウェア操作によるファイルシステムへのアクセスの範囲からアノマリ検知が可能となった。

APT 攻撃の内部活動における長期的で探索的なファイルアクセスが検知可能となった。よって、機密情報が攻撃者に取得される前に攻撃者の長期的な内部活動を検知した場合、攻撃者が侵入しているシステムをネットワークから隔離することで、情報流出を防ぐことが可能となった。また、システムに fspeek を導入した時点で、そのシステムが攻撃者に侵入されていた場合、その攻撃者の活動が停止したときに、蓄積された TSL の比較によって攻撃者が活動停止以前に行っていた内部活動の検知が可能となった。ゆえに、一度システムに侵入した攻撃者による 2 回目以降の内部活動による情報窃取を未然に防ぐことが可能である。以上より、APT 攻撃における内部活動を検知することで、情報窃取による情報流出の被害を軽減させることが可能である。また、本研究の提案した fspeek は、FAL をもとにした TSL を年単位で記録することが可能であり、システムを操作しているユーザに普段行われない活動が含まれていないかどうかを検知するものである。ゆえに、fspeek を稼働させることで、サイバーセキュリティにおける真正性および責任追及性の向上も期待できる。

参考文献

[1] 「Google」 https://www.google.com/

[2] 「Yahoo!」 https://www.yahoo.com/

[3] McAfee Labs, McAfee Foundstone Professional Service 「Protecting Your Critical Assets -Lesson Learned from "Operation Aurora"-」 McAfee, 2010

[4] Dmitri Alperovitch, Vice President, Threat Research 「Revealed: Operation Shady RAT」 McAfee, 2011