機械学習手法を用いた ネットワークトラフィックによるマイニング検出方式の検討

1215080　合田 亮登　【 セキュリティシステム研究室 】

Study on Mining Detection of Network Traffic based on Machine learning approach

1215080　Ryoto GODA　【 Security Systems Lab. 】

1　はじめに

Bitcoin に代表される暗号通貨には，マイニングと呼ばれるインターネットを介した各端末での分散処理によって取引を検証することで，検証者が報酬を得る仕組みが存在する．

現在，マルウェア感染端末に対して不正にマイニングさせることによって攻撃者が収益を得るマイニングマルウェアが出現している．マイニングマルウェアを検出するアンチウイルスソフトウェアが存在するが，端末にインストールするアンチウイルスソフトウェアでは，攻撃者が持ち込んだ端末やインストールが困難な IoT 機器からのマイニングを検知することはできない．

そこで本稿では，トラフィックパターンを用いることにより，実行端末に依存しないマイニング検出手法を検討する．

また，1 台のハードウェアを複数人で共有して使用する VPS(Virtual Private Server) において，基本的に全ユーザがコンピュータの資源を最大限使用しないことを前提としているため，CPU 負荷が高いマイニング行為を検知する手法が求められる．

2　MONERO

暗号通貨の MONERO は取引内容が追跡できないため匿名性が高く，CPU でのマイニングに最適化されていることなどから，マイニングマルウェアに多く利用される．

MONERO をマイニングする代表的なスクリプトとして Coinhive がある．Coinhive は，Web サイト閲覧者に暗号通貨の MONERO をマイニングさせることで，Web サイト運営者が収益を得ることができるサービスである．JavaScript で記述され，Web ブラウザ上で動作することから，複数のプラットフォームで動作する．

そこで今回検知するマイニング通信として Coinhive を用いる．

3　マイニング通信の構成

他のマルウェアと比較して，継続的なインターネット通信を必要とする．

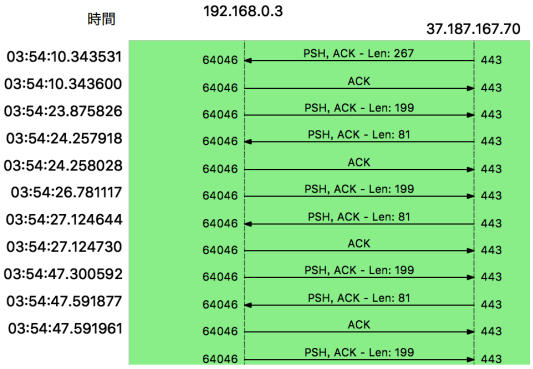

図 1　Coinhive のマイニングにおける通信内容

マイニング通信は初回の認証を除いて，図 1 のように主に 2 つの通信から成り立っている．また，利用者の環境や送信間隔によっては Ack が送信される．

1. サーバが利用者に検証する値を送信
2. 利用者がサーバに検証結果を送信

4　研究目的

Coinhive を実行することによる消費電力を計測する．Coinhive において CPU 性能を 50%使う (throttle 0.5) ように設定した時の消費電力とハッシュレートを計測した結果，表 1 のようになった．測定に当たり，システム及びアイドルを無効化し，1 時間計測した時の平均値を取得した．なお，ブラウザを開いた状態の待機電力は共に 8.2W であった．GoogleChrome の場合，マイニング時の消費電力は待機電力の 3.7 倍となる一方で，YouTube の動画を HD 画質でストリーミング再生した場合の消費電力 23W に対して 1.26 倍になった．

消費電力の監視だけでは，不正な電力の消費は検知できるが，不正行為の特定は困難である．そこで，検討方式では，ネットワークトラフィックを監視することで，マイニングを検知する．

ブラウザ	消費電力	ハッシュレート
GoogleChrome 71.0.3578.98	30.4W	15.8$hash/s$
Firefox Quantum 64.0.2	25.6W	18.3$hash/s$

表 1　マイニング時のシステムの消費電力
(MacBook Pro 13-inch Early 2013)

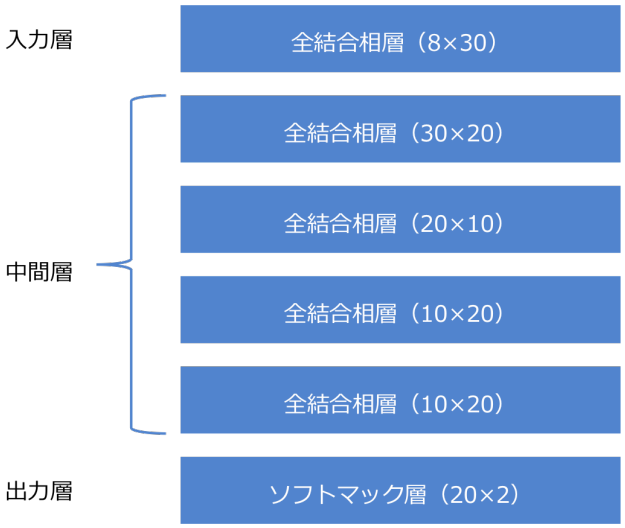

図 2　提案方式で利用するモデル

5　検討方式

ネットワークトラフィックを tcpdump により抽出し，機械学習モデルに与えることで，マイニングトラフィックが否かを判別する．

機械学習モデルの入力に使う情報は，ip ヘッダ情報から，時間，ttl，id，offset，flag，プロトコル，長さである [1].

機械学習モデルには，図 2 に示す活性化関数を relu とする 7 層のマルチレイヤパーセプトロン (MLP) による 2 値分類モデルを利用する．入力層にはパケットデータから抽出した 8 次元のデータを与え，出力層はマイニングマルウェアである確率を出力する [2][3].

6　今後の展望

本稿では，ネットワークトラフィックを利用したマイニング検出方式を検討した．検討手法は，マイニング実行端末に依存せず，マイニングトラフィックを検出できる利点がある．検討手法の詳細な有用性は，評価実験により示す予定である．

今後は，検討方式におけるマイニング通信の検知率を評価するために，Keras を用いてモデルを構築する予定である [4].

参考文献

[1] Ryo Yamada, ”Using abnormal TTL values to detect malicious IP packets”, Yamada, R., & Goto, S. (2013). Using abnormal TTL values to detect malicious IP packets. Proceedings of the Asia-Pacific Advanced Network, 2012.

[2] 小出 駿, 鈴木 将吾, ”通信プロトコルのヘッダの特徴に基づく不正通信の検知・分類手法,” Computer Security Sympsium, October, 2014.

[3] 中野和俊, ”ディープラーニングによる IP ネットワーク上のストリーミングトラヒック識別の検討”, 北陸先端科学技術大学院大学 修士学位論文, 2017.

[4] Michael Collins, 中田 秀基 (監訳), 木下 哲也 (訳者),”データ分析によ るネットワークセキュリティ,” オライリー・ジャパン, 2016.