Vol.2020-IOT-48 No.6
2020/3/2

マイニングマルウェアの通信の特徴抽出の試行

村上順也 [1]　山之上卓 [2]

概要：マイニングマルウェアが増加し、感染すると利用者が知らないうちに CPU リソースを奪われることから、これを検知して除去する必要がある。ウイルス対策ソフトでは、ゼロデイアタックは検知できないため、通信の観測によってマイニングマルウェアを検知したい。本稿では、マイニングウェアによってマイニングしているホストの通信を Wireshark でパケットキャプチャし、IP アドレスごとに毎秒の通信量をグラフ化して、マイニングウェアの通信の特徴を抽出したことについて述べる。マイニングウェアの TCP 通信は通常の TCP 通信と比べて単位時間当たりの同じパケット数を継続して確認できることや TCP のストリームの送受信に偏りがあることなどの特徴があった。

キーワード：マイニングマルウェア，パケットキャプチャ，マイニング

Attempt to extract network traffic feature of mining malware

JUNYA MURAKAMI[†1]　TAKASHI YAMANOUE[†2]

Abstract: As the number of mining malware increases and infections take away CPU resources without the user's knowledge, it is necessary to detect and remove this.Since anti-virus software cannot detect zero-day attacks, we want to detect mining malware by observing communications.In this paper, we describe that we capture the communication of the host that is mining by the miningware using Wireshark, and graph the traffic per second for each IP address to extract the characteristics of the miningware communication. Compared with normal TCP communication, miningware TCP communication has features such as the ability to continuously check the same number of packets per unit time and the uneven transmission and reception of TCP streams.

Keywords: mining malware, packet capture, mining

1. はじめに

2018 年からマイニングマルウェアが増加している。仮想通貨の価格の上昇と比例してマイニングマルウェアが増加していることから、仮想通貨の価値の上昇に伴い、マイニングマルウェアを使用した攻撃がこれからも増加する可能性が高いと考えられる。さらに、セキュリティの甘い IoT デバイスを狙ったマイニングマルウェアも増加している。[1]

2017 年に急増し、現在は減少傾向にあるランサムウェアと比較してマイニングマルウェアは低リスク、低コストで稼ぐことができることから攻撃者には都合がいい。攻撃者はいかに見つかることなく感染ホストの不正マイニングを行って仮想通貨を稼ごうとし、日に日に攻撃手法が巧妙化している。トレンドマイクロのセキュリティブログによると正規プロセスの中に不正なプロセスを入れ替えるプロセスハロウイングという手法と不正活動開始にコンポーネントファイルにコマンドライン引数を参照させる手法とを組み合わせる高度な検出回避手法が確認されている。引数が呼び出されなかったときは不正ファイルを見つけ出すことができないことで動的解析を困難化させている。[2]このような攻撃手法でゼロデイ攻撃を受けた場合にアンチウイルスソフトでは検知できない可能性がある。我々は、不正マイニングをされているときの通信の特徴を用いて検知が可能と考え、特徴を基に検知したい。

本稿では、マイニングウェアが動作しているときの通信とマイニングマルウェアが不正マイニングしているときの通信は似た通信の観測結果になると仮定し、マイニングウェアが動作している通信と様々な場面での通信を Wireshark でパケットキャプチャし、それぞれの通信結果を比較してマイニングウェアが動作しているときの特徴を抽出したことについて述べる。

2. 関連研究

2.1 Detection of Bitcoin miners from network measurements

Jordi Zayuelas I Muñoz は CISCO 社のネットワークトラフィックを分析するルータ機能である Netflow を使用してキャプチャした情報の特徴を機械学習によって学習させ、マイニングしている通信の検知を行った[3]。機械学習を用いて検知はできているが、通信そのものがどういう特徴を検出して検知しているのかどうか判断することができない。我々の手法においては、マイニングウェアでマイニングしている通信を Wireshark でキャプチャし、パケットの毎秒の通信量などから特徴を抽出し、その特徴を明らかにしようとしている。

2.2 Fine-Grain Feature Extraction from Malware from

1 福山大学
Fukuyama University

Malware's Scan Behavior Based on Spectrum Analysis

江藤らはパケットの宛先 IP アドレスの振動から SPectrum Analysis for Distinction and Extraction of malware features (SPADE) によって離散フーリエ変換し、マルウェアの特徴を抽出するスペクトラム分析手法を提案した。マルウェアの特徴を抽出する部分で類似している。我々の研究では、宛先 IP アドレスの毎秒の通信量をグラフ化した結果自体を扱い、特徴の抽出を試みている。我々の研究でもマイニングウェアが動作している通信をキャプチャし、プール先の宛先 IP アドレスの毎秒の通信量をグラフ化した結果を離散フーリエ変換して特徴抽出を試みたが、有用性のある特徴を見つけることができなかった。

3. マイニングの通信形態

Bitcoin をはじめとする仮想通貨の台帳管理を担うブロックチェーンのシステムを成り立たせるためにはマイニングが欠かせない。ブロックチェーンの未確認のトランザクションの集まりを正しいものと証明するべく、nonce 値、1 つ前のブロックのハッシュ値、未確認のトランザクションデータの 3 つをハッシュ値にしていくが、それぞれのブロックチェーンを採用した仮想通貨はあらかじめ承認するハッシュ値の規則を定めている。規則に合うように承認してくれるハッシュ値を求めて任意の数値である nonce 値を探索することになる。この作業をマイニングと呼ぶ。マイニングをして正しいハッシュ値を手に入れることができた者には報酬として新規発行された仮想通貨が与えられる。報酬を目当てに企業などがマイニングに勤しんでおり、犯罪者も報酬に目を付けた。マイニングは大量の計算能力があればあるほど有利になり、マイニング専用の GPU に需要が高まっている。

マイニングに協力するには自分でマイニングする方法、マイニングプールに参加する方法、クラウドマイニングでマイニングしている組織に投資する方法がある。多数のマイニングマルウェアが使用するマイニングプールについて述べる(図 1)。

マイニングプールに参加した人達の合わせた計算能力を使用して報酬の獲得を狙い、それぞれの計算能力の結果に応じて報酬を山分けする。マイニングプールに参加するメリットは計算能力が低くても報酬を獲得できることにある。マイニングマルウェアは他人のホストからバックグラウンドでマイニングプールに参加させ、CPU リソースを横取りして利益を攻撃者の口座に入るよう設定する。バックグラウンドでマイニングをしているものの、どこかのノードで発生した新しいトランザクションと直前のブロックのハッシュ値をマイニングプールから TCP で受信し、ハッシュ値を計算し、その結果をマイニングプールに TCP で送信する必要がある。正規でマイニングしている通信の特徴が、不正マイニングされている場合の通信にも表れると仮定する。

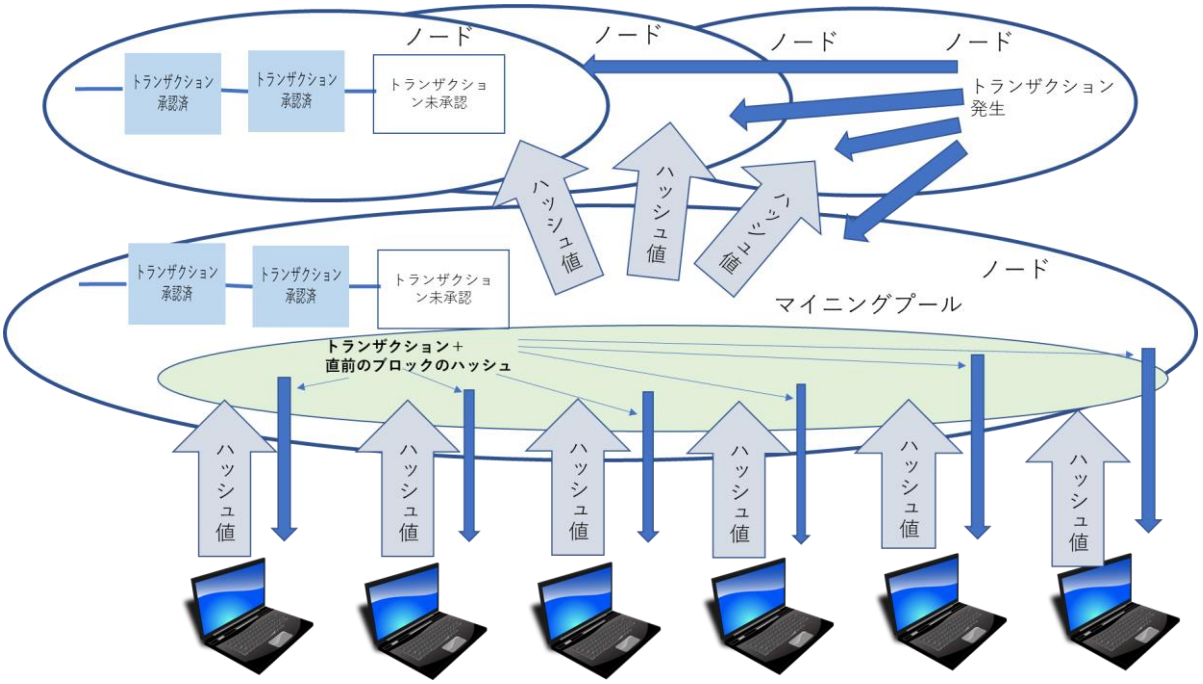

図 1. マイニングプールを使ったマイニングの概要

Ethernet・1 | IPv4・3 | IPv6 | TCP・26 | UDP

Address A	Port A	Address B	Port B	Packets	Bytes	Packets A → B	Bytes A → B	Packets B → A	Bytes B → A	Rel Start	Duration	Bits/s A → B	Bits/s B → A
192.168.11.6	55665	136.243.102.154	443	22	8574	11	1662	11	6912	44.947236	1.8236	7291	30 k
192.168.11.6	55670	136.243.102.154	443	20	8575	9	1422	11	7153	75.172388	2.0815	5465	27 k
192.168.11.6	55666	136.243.102.154	443	15	2219	8	1453	7	766	46.509017	2.2772	5104	2691
192.168.11.6	55169	88.99.142.163	443	10	674	5	342	5	332	24.650996	7.0791	386	375
192.168.11.6	55669	136.243.102.154	443	43	14 k	21	2605	22	11 k	50.495411	7.1647	2908	13 k
192.168.11.6	55171	136.243.102.154	443	9	624	6	399	3	225	16.594537	15.1441	210	118
192.168.11.6	54982	136.243.102.154	443	9	624	6	399	3	225	14.522026	17.4524	182	103
192.168.11.6	55168	136.243.102.154	443	8	570	5	345	3	225	6.522052	25.2088	109	71
192.168.11.6	54980	136.243.102.154	443	9	624	6	399	3	225	6.521745	25.5563	124	70
192.168.11.6	54981	136.243.102.167	443	14	965	7	450	7	515	0.911767	31.1714	115	132
192.168.11.6	55662	136.243.102.154	443	20	8053	9	1141	11	6912	44.912327	60.8161	150	909
192.168.11.6	55660	136.243.102.154	443	21	8834	9	1715	12	7119	44.584912	67.4285	203	844
192.168.11.6	55677	136.243.102.154	443	15	2219	7	1399	8	820	131.591036	67.8755	164	96
192.168.11.6	55659	136.243.102.154	443	30	14 k	13	2497	17	11 k	44.584287	67.9673	293	1363
192.168.11.6	55658	136.243.102.154	443	28	10 k	13	2498	15	7531	44.583544	68.1641	293	883
192.168.11.6	55667	136.243.102.154	443	15	2219	7	1399	8	820	48.525146	68.2662	163	96
192.168.11.6	55678	136.243.102.154	443	53	16 k	23	2753	30	13 k	133.197691	76.4019	288	1458
192.168.11.6	55668	88.99.142.163	45700	255	32 k	132	20 k	123	11 k	49.481922	160.8110	1039	562
192.168.11.6	55663	136.243.102.154	443	40	10 k	17	2017	23	8160	44.936919	161.4898	99	404
192.168.11.6	55664	136.243.102.154	443	38	10 k	16	1964	22	8360	44.937552	161.4943	97	414
192.168.11.6	55023	136.243.102.154	443	21	1596	7	623	14	973	24.754560	181.8704	27	42
192.168.11.6	55190	88.99.142.163	443	21	1596	7	623	14	973	24.555415	181.9766	27	42
192.168.11.6	55197	88.99.142.163	443	25	1970	9	731	16	1239	24.156689	182.3733	32	54
192.168.11.6	55020	136.243.102.154	443	25	1970	9	731	16	1239	24.059515	182.5628	32	54
192.168.11.6	55504	136.243.102.154	443	62	13 k	26	1708	36	11 k	3.510756	206.0686	66	453
192.168.11.6	55160	88.99.142.163	443	62	13 k	26	1708	36	11 k	3.512664	206.0870	66	453

図 2. MinerGate の IP アドレス側の 45700 ポートの通信が長時間継続して確認されやすいことを表した対話

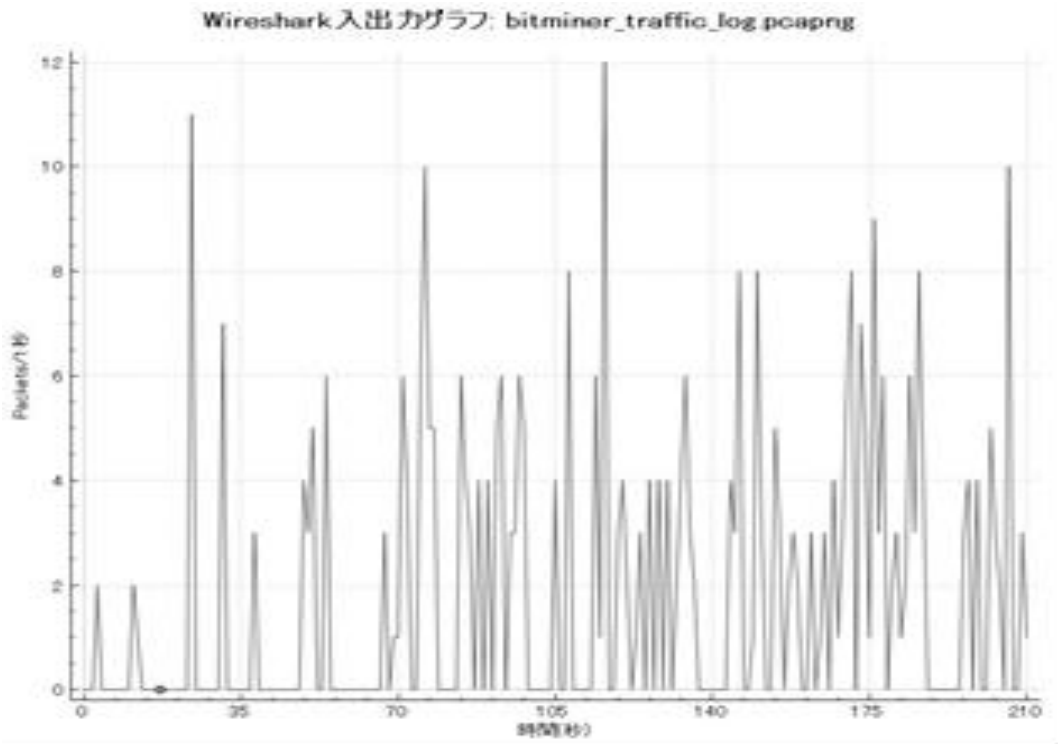

図 3. MinerGate を実行しているときのパケット数の変化

4. パケット観測方法

本章では、マイニングウェアの通信の特徴を抽出するためのデータを収集するための観測方法について述べる。

ホストの通信をパケットキャプチャするツールは Wireshark を使用した。Wireshark とは、ネットワークプロトコルアナライザであり、リアルタイムでネットワーク上の流れているパケットを取得できる。Wireshark の入出力グラフを表示できる機能や IP アドレスのペアごとの情報が比較できる機能があり分析の効率が上がる。

また、マイニングウェアは MinerGate の MinerGate Mobile Miner と呼ばれるアプリケーションを使用した。MinerGate とは、様々な仮想通貨をマイニングできるマイニングプールである。MinerGate でマイニングする際の仮想通貨は Monero を選定した。Monero は匿名性が高く、2018 年ごろに価格が上昇したこともあり、攻撃者によって不正マイニングされやすい。paloalto NETWORKS によると、Monero はマイニングマルウェアの標的となる仮想通貨の中で大部分を占めていることが確認されている。[]この原因として、マイニングする際に他の仮想通貨と比較して計算の難易度が低く、一般的な PC の CPU でもマイニング可能であることから不正マイニングされる仮想通貨の中で上位に位置付けていると考えられる。

Wireshark でパケットキャプチャする時間は 256 秒当たりで打ち切る。MinerGate でのマイニングする際の設定は CPUcore 数が 3 で Reword method は PPS とする。

MinerGate によってマイニングされている通信の Wireshark の気になる点を挙げる。

- MinerGate の IP アドレスからマイニングしているホストへのポート使用数が格段に多くなる時があること
- 単位時間あたり、同じパケット数を継続して確認できること
- MinerGate の IP アドレス側の 45700 ポートの通信が長時間継続して確認されやすいこと
- ポートそれぞれのバイト数の偏り方
- 双方向の通信のバイト数を合わせた大きさ

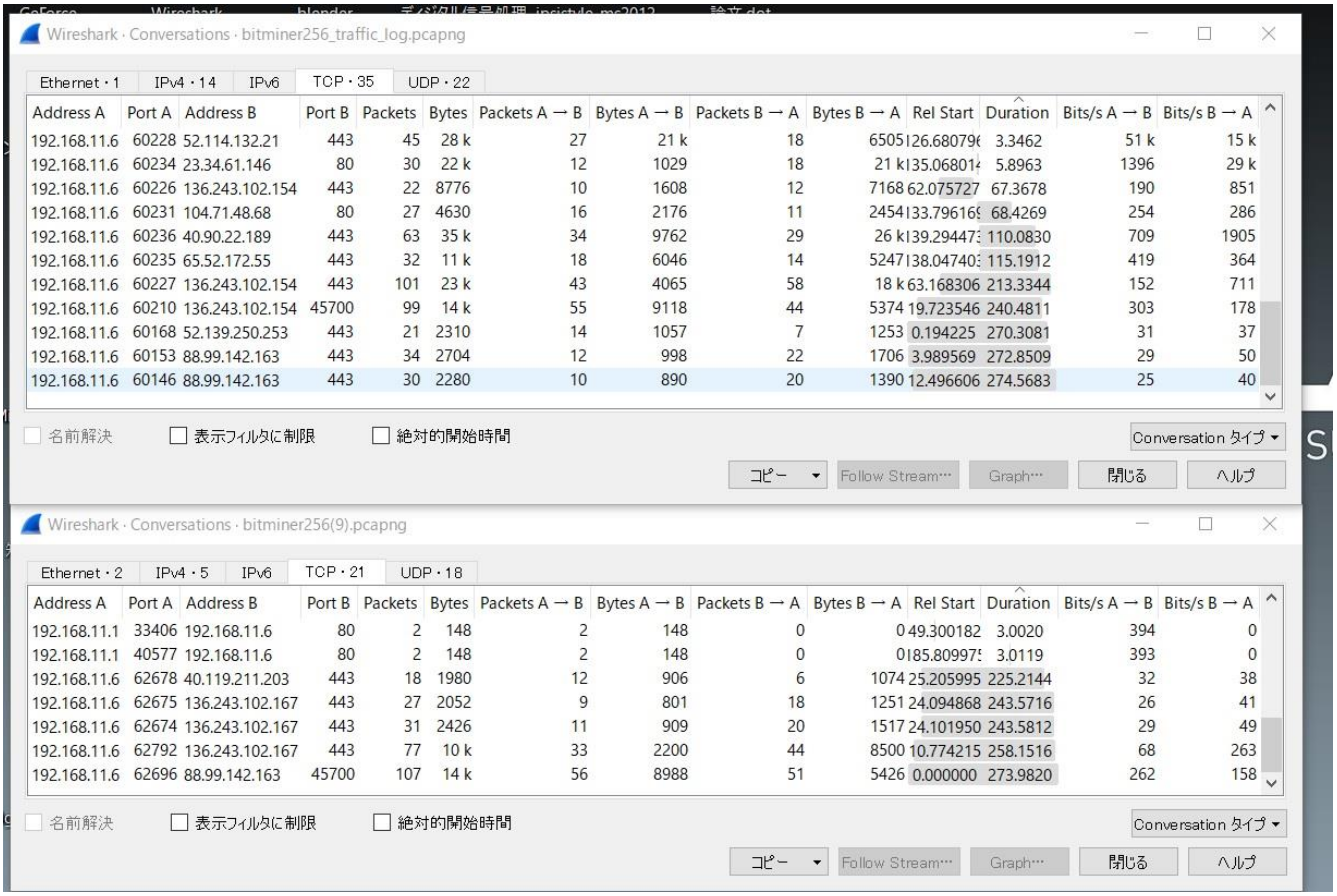

図 4. MinerGate の通信を Wireshark でパケットキャプチャした際の対話の画面

気になる点に着目し、下記の日常で PC を使用している

ときの PC の状況を再現した通信を Wireshark でパケットキャプチャして比較した。

- Skype でマイク確認テストを数回行ったとき
- ネットサーフィンをしているとき
- Youtube を視聴しているとき
- 何もせず放置したとき

Skype でのマイク確認テストは 1 分ほどで終わるので、何回も繰り返しテストをした結果をパケットキャプチャした。同時に、Skype の通話自体のパケットキャプチャ中は切れていることになることから、Skype の通話自体のパケット数の継続は見込めない。

Skype を使用すると必ず IP アドレス(20.189.78.37)を使用し、その IP アドレスは名前解決がされており、Google 関連の安全な IP アドレスであった。この IP アドレスは MinerGate の気になる点であるポート使用数と単位時間当たりに同じパケット数を継続して通信するところが被る。だが、ポートそれぞれのバイト数の偏りと双方向の通信を合わせたバイト数の大きさに差があるので区別することができる。

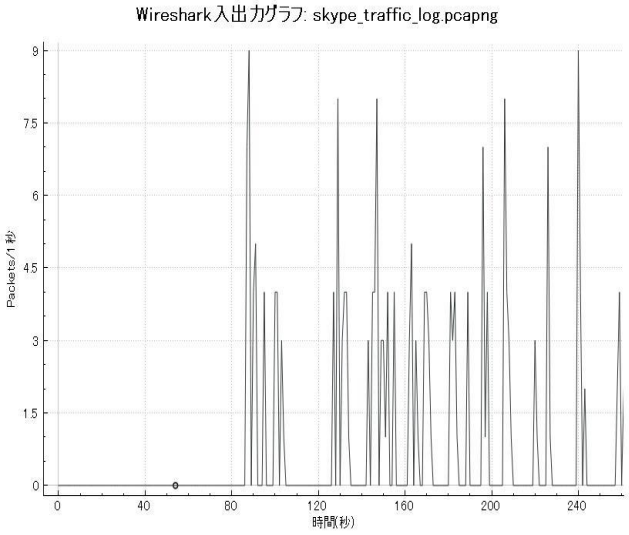

図 5. Skype のマイク確認テスト時のパケット数の変化

ネットサーフィンの通信は、**（図）**のような通信ばかりであり、単位時間当たりの同じパケット数を継続して確認することはなかった。このことから、安全にネットサーフィンをすることができれば、MinerGate との通信の特徴は明白に区別できる。

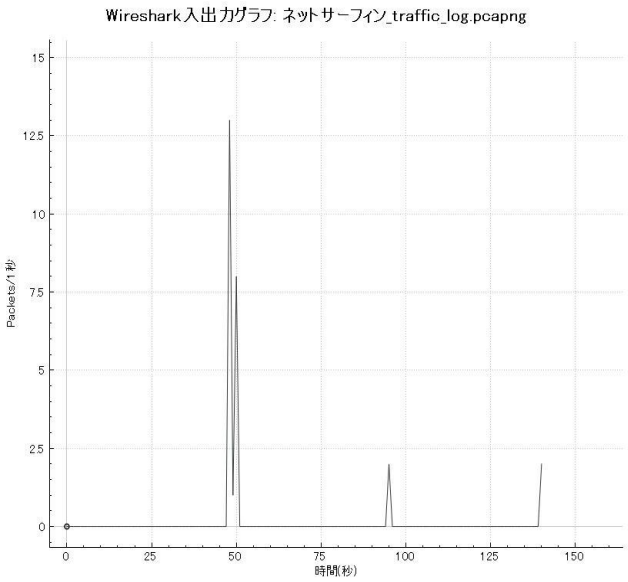

図 6. ネットサーフィン時のパケット数の変化

何もせず放置した通信は何もしていないがたびたび通信があった。パケットを見た限りだと、Windows の OS により通信をしていると思われる。通信自体があまり見られないことから、MinerGate が動作した際には通信を見れば確実にわかる。

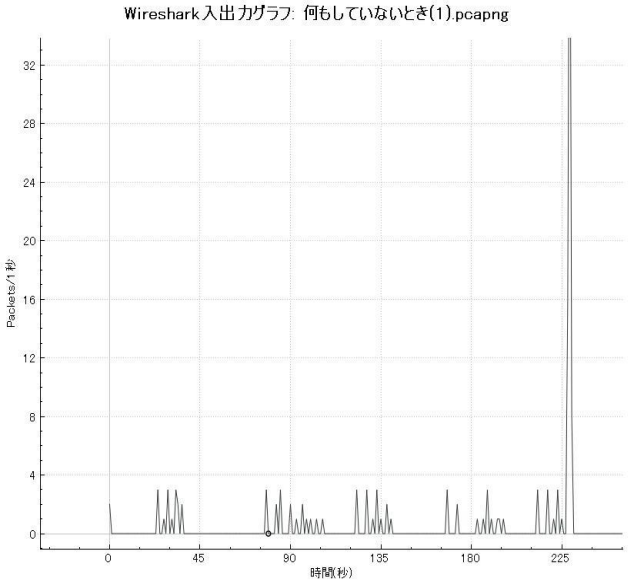

図 7. 何もせず放置しているときのパケット数の変化

おわりに

MinerGate でマイニングしているホストの通信を Wireshark でパケットキャプチャし、IP アドレスごとに毎秒の通信量をグラフ化して、MinerGate の通信に特徴があることを示した。MinerGate と 4 つの日常を想定した通信をそれぞれ比較して MinerGate の通信のパケットキャプチャから得られた特徴が他の通信に現れず、区別できる特徴であった。十分なデータではないことから、偶然類似した

特徴が見つからなかった可能性がある。まだ見つけられていない特徴があり、本稿で、疑っている特徴よりも有用性のある特徴が見つけられるよう比較するデータを増やすことが必要である。特徴が類似していそうな通信を推測し増やしていくと新たな特徴を発見できるカギになると考える。また、マイニングマルウェアとマイニングウェアの TCP ストリームが実際に類似しているかどうか確かめる必要がある。

参考文献

[1] “仮想通貨マイニングマルウェアが 71%増、マカフィー2018 年第 3 四半期の脅威レポート”. https://ascii.jp/elem/000/001/798/1798580/, (参照 2018-6-14).
[2] “巧妙な検出回避手法「プロセスハロウイング」を使用する不正コインマイナー攻撃“. https://blog.trendmicro.co.jp/archives/23376, (参照 2019-12-27).
[3] “科学技術情報流通技術基準 参照文献の書き方(SIST 02)”. http://jipsti.jst.go.jp/sist/pdf/SIST02-2007.pdf, (参照 2018-12-02).
[4] “Microsoft Office”. https://office.microsoft.com/ja-jp/, (参照 2018-12-02).
[5] “Office 製品“. https://office.microsoft.com/ja-jp/products, (参照 2018-12-02).
[6] 桜井貴文. 直観主義論理と型理論. 情報処理, 1999, vol. 30, no. 6, p. 626-634.
[7] 野口健一郎, 大谷真. OSI の実現とその課題. 情報処理, 1990, vol. 31, no. 9, p. 1235-1244.
[8] 田中正次, 村松茂, 山下茂. 9 段数 7 次陽的 Runge-Kutta 法の最適化について. 情報処理学会論文誌. 1992, vol. 33, no. 12, p. 1512-1526.
[9] Itoh, S. and Goto, N.. An Adaptive Noiseless Coding for Sources with Big Alphabet Size. IEICE Transactions. 1991, vol. E74-A, no. 9, p. 2495-2503.
[10] Foley, J. D. et al.. Computer Graphics: Principles and Practice in C. 2nd ed., Addison-Wesley Professional, 1990, 1200p.
[11]千葉則茂, 村岡一信. レイトレーシング CG 入門. サイエンス社, 1990, 282p.
[12] Chang, C. L. and Lee, R. C. T.. Symbolic Logic and Mechanical Theorem Proving. Academic Press, 1973, 331p.