PRACTICE Dataset 2013

NTTコミュニケーションズ株式会社
大村優、畑田充弘
2013年6月12日

PRACTICE Dataset 2013 概要

■ データセット概要

- 総務省「国際連携によるサイバー攻撃予知・即応に関する実証実験」（略称：PRACTICE）の挙動観察システムで、マルウェアを長期観測（最大1週間）した際の通信トラヒック（マルウェア感染後の通信挙動）を示すデータ
- 提供するデータの内訳
  - ➢ 検体情報（ハッシュ値、AV4製品によるScan結果）
  - ➢ 通信トラヒックデータ（pcap形式）

■ 対象検体

- PRACTICEで観測中の、通信挙動で特徴的な挙動を示す検体を任意に抽出

■ 対象検体数

- 5検体

■ 観測日時

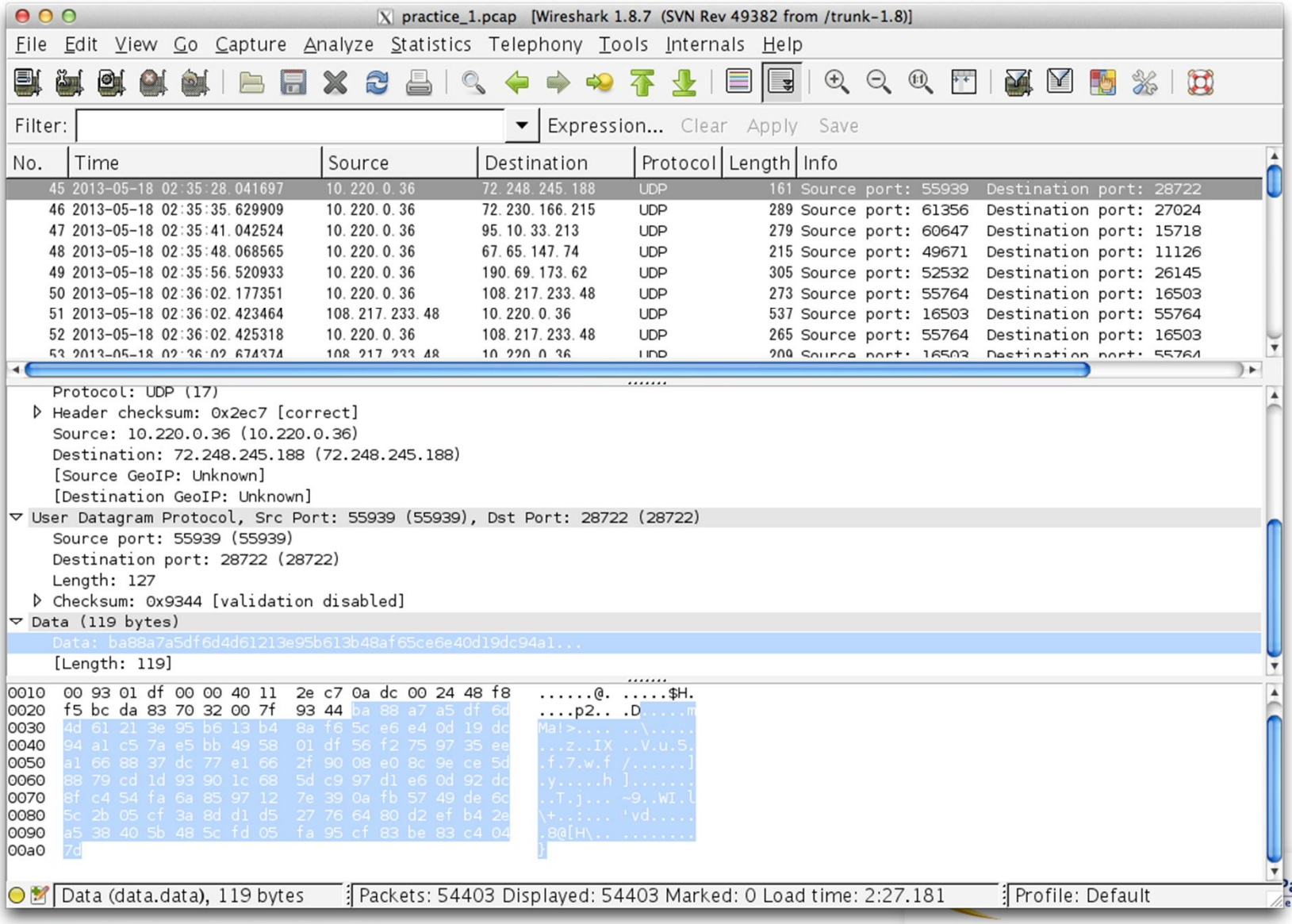
- 観測日時：2013/5/18～2013/5/25（最大1週間分のデータ）

検体情報

■ 提供する長期観測の検体情報　※検体そのものは提供しない

データセット名	検体Hash（SHA-1）	挙動解析IP	AV検知結果		解析時間	ファイルサイズ
practice_1.pcap	5b9f78af4e5609c17fdff4d97e060d1a264b72d3	10.220.0.36	Kaspersky	未検出	start: 2013-05-18 02:35:06 end: 2013-05-25 11:59:53	10MB
			McAfee	PWS-Zbot.gen.alu		
			Symantec	未検出		
			TrendMicro	未検出		
practice_2.pcap	5944b5a106a75a7d0c4b7fe2f4099efb7ba79eae	10.220.0.37	Kaspersky	Backdoor.Win32.VanBot.cx	start: 2013-05-18 02:35:19 end: 2013-05-25 11:35:21	2.6MB
			McAfee	Generic BackDoor		
			Symantec	W32.Spybot.Worm		
			TrendMicro	WORM_MYTOB.IR		
practice_3.pcap	2fec8e24ac3c911955c37ddab6904b2e7db74309	10.220.0.38	Kaspersky	Trojan-Ransom.Win32.PornoAsset.abtn	start: 2013-05-18 02:35:36 end: 2013-05-20 02:00:01	494MB
			McAfee	ZeroAccess.hj		
			Symantec	Trojan.Zeroaccess!g19		
			TrendMicro	TROJ_GEN.RCCC7IT		
practice_4.pcap	12dba89f2c869ff6f12f8005dfb004628e2c983d	10.220.0.39	Kaspersky	Backdoor.Win32.ZAccess.ylb	start: 2013-05-18 02:35:55 end: 2013-05-20 02:00:00	231MB
			McAfee	ZeroAccess.hg		
			Symantec	Trojan.Gen		
			TrendMicro	TROJ_GEN.USBJ05ACN		
practice_5.pcap	093584d4f63d45fb46beb390ba9c10b73b394a88	10.220.0.40	Kaspersky	Trojan-Spy.Win32.SpyEyes.wb	start: 2013-05-18 02:36:16 end: 2013-05-25 11:59:58	4.3MB
			McAfee	Artemis!1E7C50EACE3D		
			Symantec	Trojan.Gen		
			TrendMicro	TSPY_SPYEYE.SME		

通信トラヒックデータ



解析環境等についての諸注意

■ 解析環境

- NTTセキュアプラットフォーム研究所が開発した動的解析システム
  - ✓ 参考URL：http://www.iwsec.org/mws/2009/paper/A7-3.pdf
- 解析用OS：Microsoft Windows XP SP2
- IPアドレス、デフォルトGW、DNSサーバはDHCPで割当
- 解析環境の動作確認としてpcapに記録されているもの
  - ✓ www.google.co.jpにHTTPの（主に）HEADリクエストを送信
  - ✓ 時刻同期：ntp.jst.mfeed.ad.jp
  - ✓ グローバルIPアドレス確認：checkup.dyndns.org
  - ✓ DNSクエリ：gk-open10の名前解決

■ AV検知結果

- 各検体収集時点で最新の各社パターンファイルに基づく
  - ✓ practice_1: 2012/09/12
  - ✓ Practice_2: 2012/03/22
  - ✓ Practice_3: 2012/09/28
  - ✓ Practice_4: 2012/12/05
  - ✓ Practice_5: 2012/02/27

■ 論文記載上の注意点

#1

```
$ capinfos practice_1.pcap
File name:           practice_1.pcap
File type:           Wireshark/tcpdump/... - libpcap
File encapsulation:  Ethernet
Packet size limit:   file hdr: 4096 bytes
Number of packets:   54403
File size:           10977185 bytes
Data size:           10106713 bytes
Capture duration:    638687 seconds
Start time:          Sat May 18 02:35:06 2013
End time:            Sat May 25 11:59:52 2013
Data byte rate:      15.82 bytes/sec
Data bit rate:       126.59 bits/sec
Average packet size: 185.77 bytes
Average packet rate: 0.09 packets/sec
SHA1:                dafceef264eb9c504ff26b2e81ae779dfb454ba4
RIPEMD160:           64ad22ab0f2c2da7fbdbfe2de04deb32735a36a1
MD5:                 9a590792443587379323b57094cd078b
Strict time order:   True
```

TCP Endpoints

Address	Port	Packets	Bytes	Tx Packets	Tx Bytes	Rx Packets	Rx Bytes
211.209.241.213	13109	418	209 432	196	194 048	222	15 384
94.137.177.75	26614	1 150	162 281	470	60 931	680	101 350
10.220.0.36	61532	181	137 653	83	5 676	98	131 977
211.75.189.231	20897	1 025	137 241	500	53 727	525	83 514
98.201.1							9 510
46.48.1							85 140
91.12							4 870
99.18							11 852
46.8.1							72 888
188.24							66 354
95.78.							6 620
24.24							6 326

Addres							Bytes
10.220.							22 860
10.220.							12 860
10.220							10 808
194.94.							21 326
94.203.							71 819
75.34.30.							59 007
24.247.216.229	11992	600	133 630	317	80 385	283	53 287
123.110.175.202	15187	618	132 336	312	75 187	306	57 149
75.76.164.189	10788	598	132 125	314	77 960	284	54 165
86.54.238.222	22428	609	128 252	277	65 976	332	62 276

・TCP/UDPともにランダムっぽいhigh port利用
・たまにhttpとか
・名前解決に失敗も多々
・UDPはホスト毎の送受信割合がある程度一定

```
$ tshark -r practice_1.pcap -q -z conv,ip | head -12
================================================================================
IPv4 Conversations
Filter:<No Filter>
                                |      <-      |      ->      |     Total     |  Rel. Start   |   Duration    |
                                | Frames  Bytes | Frames  Bytes | Frames  Bytes |               |               |
10.220.0.100      <-> 10.220.0.36          3340   333668   3344   426240   6684   759908    0.000539000   637201.6798
194.94.127.98     <-> 10.220.0.36           613   123042    602    63335   1215   186377  191.701388000   638470.3955
94.137.177.75     <-> 10.220.0.36           705   106025    495    67150   1200   173175  57003.557062000   307555.2854
46.48.235.250     <-> 10.220.0.36           697    85140    480    46530   1177   131670   69.778915000   637854.7095
46.8.115.89       <-> 10.220.0.36           668    73294    403    42589   1071   115883  273766.309451000   363971.2925
188.242.252.23    <-> 10.220.0.36           583    66354    478    46500   1061   112854  273833.027220000   364027.9457
211.75.189.231    <-> 10.220.0.36           527    83914    502    54559   1029   138473  113839.445968000   193919.9510
```

#2

```
$ capinfos practice_2.pcap
File name:           practice_2.pcap
File type:           Wireshark/tcpdump/... - libpcap
File encapsulation:  Ethernet
Packet size limit:   file hdr: 4096 bytes
Number of packets:   21680
File size:           2704718 bytes
Data size:           2357814 bytes
Capture duration:    637201 seconds
Start time:          Sat May 18 02:35:19 2013
End time:            Sat May 25 11:35:20 2013
Data byte rate:      3.70 bytes/sec
Data bit rate:       29.60 bits/sec
Average packet size: 108.76 bytes
Average packet rate: 0.03 packets/sec
SHA1:                7e70979b73d1058d6fcbf239f963375d77870549
RIPEMD160:           0c7fa9ec5e52ed44c95c64a8eb1fe01b2e6c6a1f
MD5:                 6d8ce180eff7d51ee858dd93a4ef063d
Strict time order:   True
```

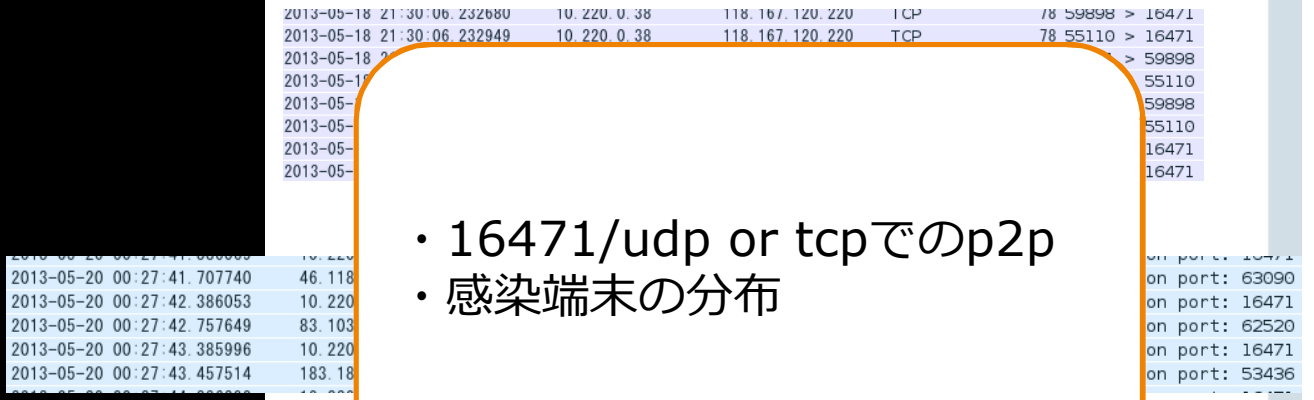
```
4f f7 3a 69 72 63 2e 66  6f 72 63 65 2e 66 6f 20   O.:irc.f orce.fo
4e 4f 54 49 43 45 20 4a  50 4e 7c 30 30 7c 58 50   NOTICE J PN|00|XP
7c 53 5[illegible]                                  |S [illegible] 80351 :*
2a 2a                                               re perma
6e 65                                               anned fr
6f 6d                                               (no rea
73 6f                                               ROR :Clo
73 69                                               k: JPN|0
30 7c                                               |L|38035
31 5b                                               617.toky
6f 2d                                               .ne.jp]
28 55                                               s been p
65 72                                               ly banne
64 20                                               orce (no
20 72                                               )..
```

・繰り返しIRC接続を試みるも、特定のUSERでブロックされてるので、、、の繰り返し

```
$ tshark -r practice_2.pcap -q -z conv,ip | head -12
================================================================================
IPv4 Conversations
Filter:<No Filter>
                                |      <-      |  |      ->      |  |     Total     |  Rel. Start  |   Duration   |
                                | Frames  Bytes |  | Frames  Bytes |  | Frames  Bytes |              |              |
78.24.188.201       <-> 10.220.0.37          12067   925574    8216   1127118   20283   2052692    14.401138000    605123.8127
10.220.0.100        <-> 10.220.0.37            666   135808     672    162138    1338    297946     0.000500000    637200.9945
74.125.235.88       <-> 10.220.0.37             10      812      10      2116      20      2928     1.022166000    604797.1264
210.173.160.27      <-> 10.220.0.37              8      720       8       720      16      1440     0.525294000    604797.3533
216.146.43.70       <-> 10.220.0.37              6      503       6       652      12      1155     1.707067000         6.7889
216.146.38.70       <-> 10.220.0.37              5      437       4       532       9       969  604798.558769000         0.6297
255.255.255.255     <-> 10.220.0.37              2      684       0         0       2       684     0.000000000    604797.2456
```

#3

```
$ capinfos practice_3.pcap
File name:           practice_3.pcap
File type:           Wireshark/tcpdump/... - libpcap
File encapsulation:  Ethernet
Packet size limit:   file hdr: 4096 bytes
Number of packets:   1160063
File size:           517963595 bytes
Data size:           499402563 bytes
Capture duration:    170664 seconds
Start time:          Sat May 18 02:35:36 2013
End time:            Mon May 20 02:00:00 2013
Data byte rate:      2926.23 bytes/sec
Data bit rate:       23409.82 bits/sec
Average packet size: 430.50 bytes
Average packet rate: 6.80 packets/sec
SHA1:                86efa5943323aec7c874ad413efe9d2ddbfca7a2
RIPEMD160:           d923294e579107d975eadd766ad82fcf1fd01618
MD5:                 8f41893919cdb2c8f94668d4919fba8c
Strict time order:   True
```

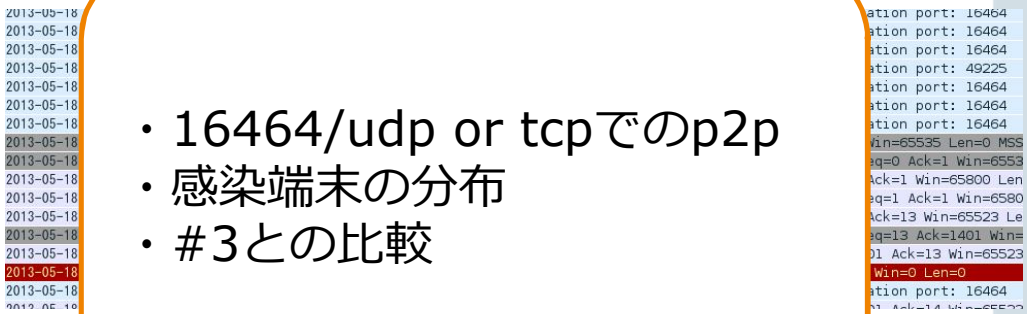


- 16471/udp or tcpでのp2p
- 感染端末の分布

```
$ tshark -r practice_3.pcap -q -z conv,ip | head -12
================================================================================
IPv4 Conversations
Filter:<No Filter>
                                |      <-      |  |      ->      |  |     Total     |  Rel. Start  |   Duration   |
                                | Frames  Bytes |  | Frames  Bytes |  | Frames  Bytes |              |              |
219.80.142.21       <-> 10.220.0.38          3452   218172     20     9958   3472   228130  3399.178702000    167228.7679
120.201.89.250      <-> 10.220.0.38          1135    70554   1154   417898   2289   488452   761.741767000    169177.2880
158.254.253.254     <-> 10.220.0.38          1869   113420      0        0   1869   113420   289.677549000    170294.6172
134.254.253.254     <-> 10.220.0.38          1824   110630      0        0   1824   110630   285.677644000    170295.6173
166.254.253.254     <-> 10.220.0.38          1822   109320      0        0   1822   109320   291.677416000    170291.6174
113.254.253.254     <-> 10.220.0.38          1820   109200      0        0   1820   109200   293.677414000    170291.6173
206.254.253.254     <-> 10.220.0.38          1806   108360      0        0   1806   108360   280.677486000    170295.6175
```

#4

```
$ capinfos practice_4.pcap
File name:           practice_4.pcap
File type:           Wireshark/tcpdump/... - libpcap
File encapsulation:  Ethernet
Packet size limit:   file hdr: 4096 bytes
Number of packets:   631453
File size:           241793733 bytes
Data size:           231690461 bytes
Capture duration:    170645 seconds
Start time:          Sat May 18 02:35:55 2013
End time:            Mon May 20 02:00:00 2013
Data byte rate:      1357.73 bytes/sec
Data bit rate:       10861.84 bits/sec
Average packet size: 366.92 bytes
Average packet rate: 3.70 packets/sec
SHA1:                460ab56703e98edc648f8616605986808927eb37
RIPEMD160:           7957c1ea566c010c25a5fdd0013e73638c279b3f
MD5:                 47a9d4cb1496a243edda34a94a58aac0
Strict time order:   True
```

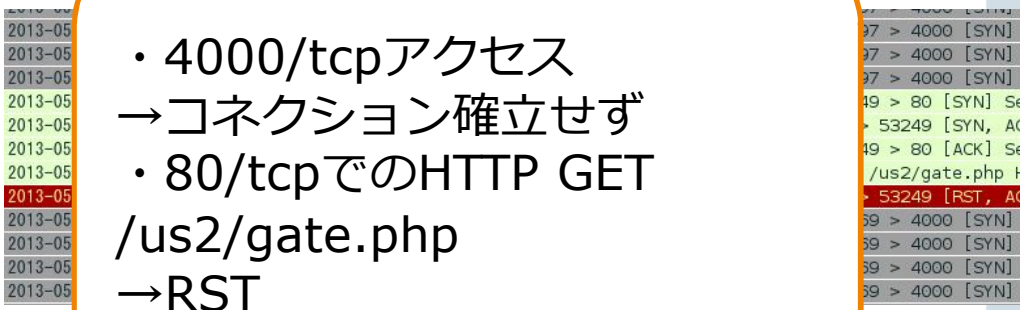


- 16464/udp or tcpでのp2p
- 感染端末の分布
- #3との比較

```
$ tshark -r practice_4.pcap -q -z conv,ip | head -12
================================================================================
IPv4 Conversations
Filter:<No Filter>
                                |      <-      |  |      ->      |  |     Total     |  Rel. Start  |   Duration   |
                                | Frames  Bytes |  | Frames  Bytes |  | Frames  Bytes |              |              |
189.95.79.14       <-> 10.220.0.39          2317    146190       8     2036    2325    148226  3272.591404000    167166.3274
219.55.222.3       <-> 10.220.0.39          2086    131556       7     1030    2093    132586  23290.745905000    147313.0884
69.170.93.61       <-> 10.220.0.39          2081    131214       5      894    2086    132108  23176.702321000    147292.2011
158.254.253.254    <-> 10.220.0.39          1715    104090       0        0    1715    104090   282.086861000    170298.3829
134.254.253.254    <-> 10.220.0.39          1659    100640       0        0    1659    100640   278.086807000    170299.3830
197.254.253.254    <-> 10.220.0.39          1650     99078       3      258    1653     99336   277.086893000    170310.3827
190.254.253.254    <-> 10.220.0.39          1645     98700       0        0    1645     98700   279.086790000    170307.3828
```

#5

```
$ capinfos practice_5.pcap
File name:           practice_5.pcap
File type:           Wireshark/tcpdump/... - libpcap
File encapsulation:  Ethernet
Packet size limit:   file hdr: 4096 bytes
Number of packets:   49137
File size:           4559205 bytes
Data size:           3772989 bytes
Capture duration:    638622 seconds
Start time:          Sat May 18 02:36:16 2013
End time:            Sat May 25 11:59:58 2013
Data byte rate:      5.91 bytes/sec
Data bit rate:       47.26 bits/sec
Average packet size: 76.79 bytes
Average packet rate: 0.08 packets/sec
SHA1:                482b3ce0ab3073e08ea4b0c5606ad697d404487e
RIPEMD160:           62cc339cecdb0edda31eb7ff1724fb086dd6a9ad
MD5:                 97665e29bfd42d9d83aba44f663231df
Strict time order:   True
```



・4000/tcpアクセス
→コネクション確立せず
・80/tcpでのHTTP GET /us2/gate.php
→RST

```
$ tshark -r practice_5.pcap -q -z conv,ip | head -12
================================================================================
IPv4 Conversations
Filter:<No Filter>
                                |      <-      |  |      ->      |  |     Total     |  Rel. Start  |   Duration   |
                                | Frames  Bytes |  | Frames  Bytes |  | Frames  Bytes |              |              |
89.149.253.239      <-> 10.220.0.40          37467   2522778      0        0   37467   2522778    15.659111000    638535.6609
131.253.18.11       <-> 10.220.0.40           4935    527934   3234   221780    8169    749714    17.986709000    638603.6925
131.253.18.12       <-> 10.220.0.40           1245    133083    816    55926    2061    189009    17.997790000    637994.3859
10.220.0.100        <-> 10.220.0.40           687    137866    693   166380    1380    304246     0.000439000    637195.8529
91.198.22.70        <-> 10.220.0.40            12      1006     10     1184      22      2190     1.485325000    604798.8333
74.125.235.87       <-> 10.220.0.40            10       812     10     2116      20      2928     0.798676000    604792.7204
210.173.160.27      <-> 10.220.0.40             4       360      4      360       8       720  604793.075706000        0.0734
```

期待する研究テーマ

- マルウェア感染後の通信挙動分析
  - 通信先、通信内容、通信タイミング等の分析から、通信先評価、悪性通信判定技術へ応用
    - ✓ リアルタイムブラックリスト
    - ✓ C&C通信判定（将来的なC&Cの自動トラッキング等）
    - ✓ 悪性通信判定技術、悪性サイトの自動判別
  - 通信アルゴリズムを解明し、マルウェアの拡散予測、攻撃予測へ応用
- マルウェア感染後の通信検知手法の評価
  - マルウェア感染後の通信トラヒックの特徴を把握し、検知技術に応用
  - トラヒックパターン分析により、検知シグネチャ、パターンファイルの精度向上、フィルタリングパターン生成等
  - 既存対策技術の検知評価、課題の抽出による改善提案
- ネットワークを含む解析環境の課題、改善提案
  - 動的解析環境の環境要因、擬似環境、エミュレーション技術の改善
- その他