挖矿流量分析

笔记本: 应急

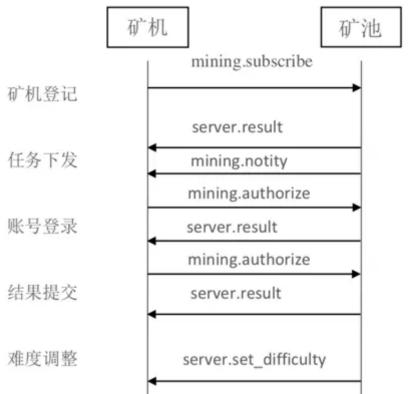
创建时间: 2022/12/8 8:29

一、概述

在对常规的挖矿流量进行分析时,常用的方法是通过抓取流量中的域名、ip地址然后丢到威胁情报平台去分析,查看是否是具有挖矿木马行为标记,但是这里可能存在的问题是假如威胁情报平台更新不及时,就无法及时准确判断出挖矿木马了。最近在参加一个ctf比赛时,正好遇到一个对流量进行分析,识别出挖矿流量,由于给出的流量包已经对一些ip或域名进行模糊,无法通过常规的方法进行筛选。所以只能依靠挖矿木马的一些流量特征进行筛选,总结了下挖矿木马常见的一些特征。

二、常见特征

挖矿木马常用的协议为Stratum协议,而挖矿的流程主要是矿池和矿机之间进行交换,常见的交互过程如下:



矿工网络主要分为矿机、矿池、钱包等几个部分组成,矿机与矿池之间的通讯协议是 stratum,而矿池软件和钱包之间的通讯为bitcoinrpc接口。

stratum协议的原理网上已经有很详细的解读了,具体可以参考如下链接:

https://www.defidaonews.com/article/54210

1、首先矿机是发送subscribe方法进行任务订阅

```
{"id": 1, "method": "mining.subscribe", "params": ["equihashminer", null, "zec-eu.suprnova.cc", "2142"]}
```

矿池会以notify返回订阅号

2、任务分配

该命令由矿池定期发给矿机,当矿机以mining.subscribe方法登记后,矿池应该马上以mining.notify返回该任务。

3、矿机登录

矿机以mining.authorize方法,用某个帐号和密码登录到矿池,密码可空,矿池返回true登录成功。该方法必须是在初始化连接之后马上进行,否则矿机得不到矿池任务。

```
2 22 22 22 22 22 22 22 22 22 22 22 08 00 45 00
0 ba 06 86 40 00 80 06 3c bd c0 a8 3c 4e 25 bb
                                                          ----@--- <---<N%-
                                                         ·I · · · · · · P
4 49 d9 0e 01 bb 98 b5 1d 19 c2 00 00 2d 50 18
0 29 08 8a 00 00 7b 22
                           69 64 22 3a 20 32 2c 20
  6d 65 74 68 6f 64 22 3a 20 22 6d 69 6e 69 6e
                                                          "method" : "minir
                                                          g.author ize",
 7 2e 61 75 74 68 6f 72 69 7a 65 22 2c 20 22 70
1 72 61 6d 73 22 3a 20 5b 22 6b 76 65 74 61 6b
e 77 6f 72 6b 65 72 31  22 2c 22 70 61 73 73 77
                                                          arams": ["kvetak
                                                          .worker1 ","passw
f 72 64 22 5d 7d 0a 7b 22 69 64 22 3a 20 35 2c
0 22 6d 65 74 68 6f 64 22 3a 20 22 6d 69 6e 69
                                                         ord"]}·{ "id": 5,
"method ": "mini
e 67 2e 65 78 74 72 61 Ge 6f 6e 63 65 2e 73 75
                                                          ng.extra nonce.su
2 73 63 72 69 62 65 22  2c 20 22 70 61 72 61 6d
                                                          bscribe" , "param
                                                          s": []}.
  22 3a 20 5b 5d 7d 0a
```

4、结果提交

矿机找到合法share时,就以"mining.submit"方法向矿池提交任务。矿池返回true即提交成功,如果失败则error中有具体原因。

矿机->矿池数据传输

Method字段	功能
mining.authorize	矿机登录认证
mining.extranonce.subscribe	向矿池表面矿机支持set_extranonce方法
mining.get_transactions	获取作业ID
mining.submit	矿机提交挖矿结果
mining.suggest_difficulty	挖矿难度的偏好
mining.suggest_target	挖矿目标的偏好
mining.capabilities(DRAFT)	矿机通知矿池其拥有的能力和可选项

矿池->矿机的数据传输

Method	功能
client.get_version	获取矿机版本信息
client.reconnect	等待指定时间(s)后重连
client.show_message	矿机展示信息
mining.notify	响应矿机的mining.subscribe请求
mining.set_difficulty	矿池更新难度
mining.set_extranonce	矿池更新extranonce
mining.set_goal(DRAFT)	通知矿机未来的工作目标(尚未使用)

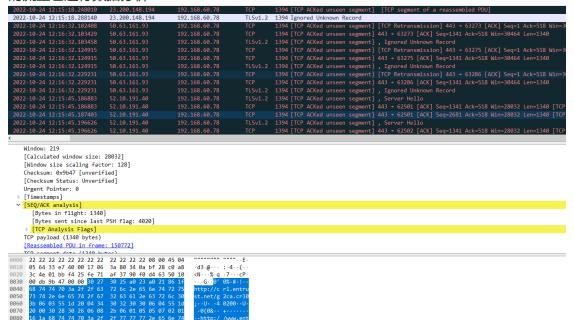
给出的题目主要内容如下:

近期,校园网开展了整治虚拟货币"挖矿"专项活动。选手拿到的是一段时间的校园网网关部分流量数据,其中目标域名、IP已经过匿名化处理。请选手审计所给的数据包,实现一个不依赖于IOC的挖矿流量检测分析引擎,并找出数据包中的所有挖矿流量。

判题脚本说明

选手需要上传—个txt文件,文件的每一行代表挖矿流量在pcap文件包中的编号数字,如图所示:

下载相关流量数据包,进行分析,第一个感觉是无从下手,因为确实还没从没有这么细致的流量包进行数据分析

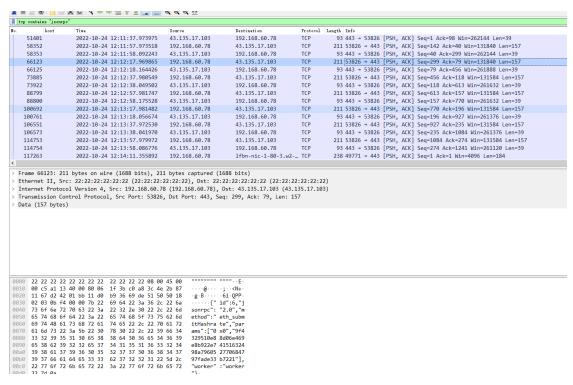


从网上找的部分挖矿流量特征:

```
'id', 'method', 'jsonrpc', 'params', 'result', 'login', 'pass', 'agent', 'job_id', 'seed_b
```

3、特征筛查

1、有了特征就会想到怎么从流量中去筛选,可直接在wireshak 可以直接利用wireshark的搜索语法进行搜索,比如tcp contains "jsonrpc" 、 http contains等内容



2、还可尝试利用脚本对pcap文件进行分析,脚本主要是利用 pyshark函数,部分代码如下:

```
capture =
pyshark.FileCapture('F:\ctf\\cryptomining.pcap', tshark_path="E:\ProgramFiles\Wire

for packet in capture:
    # Print the packet info
    print("Packet number:", packet.number)
    print("Timestamp:", packet.sniff_time)
    print("Protocol:", packet.transport_layer)
    print("Source:", packet.ip.src)
```

```
print("Destination:", packet.ip.dst)

#提取tcp data部分

if packet.transport_layer == 'TCP':
    payload_hex=str(packet.tcp.payload).replace(':','')
    payload_bytes = bytes.fromhex(payload_hex)
    payload_str = payload_bytes.decode('utf-8')
```

通过这几行代码基本就能分析出pcap包中的tcp源端口、目的端口、编号、tcp的数据等部分内容,然后在根据挖矿的流量特征进行匹配就可以了,打印出部分结果如下:

```
| Compacted in capture | unpacted in unpacted
```

这只是提供了一个针对挖矿流量脱离IOC的识别思路,可能搜集的特征库还不够全。看网上还有一种简单方法就是直接导入安全检测设备,看设备的检测结果,如果具备条件的可以一试。