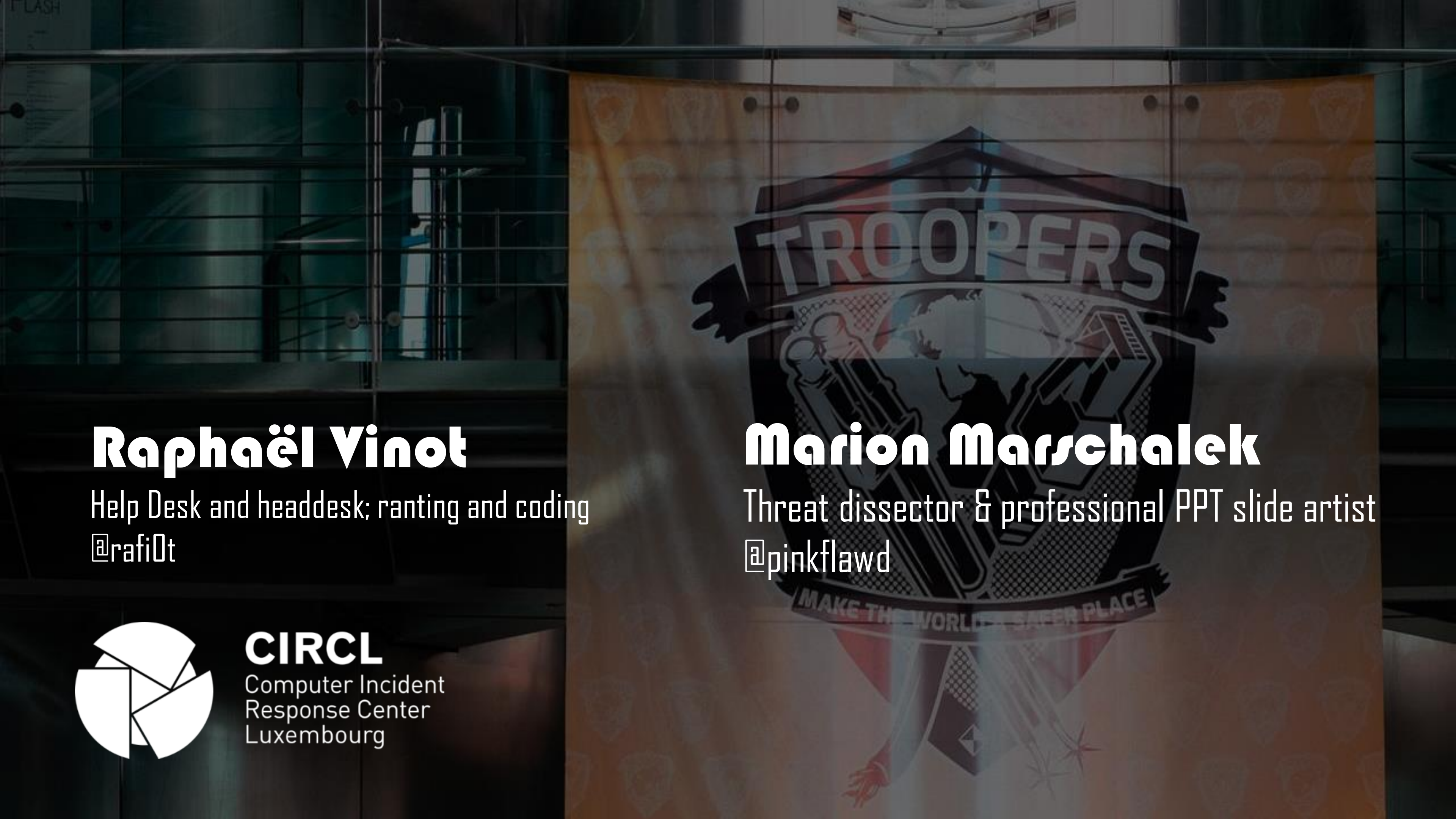
Graph me, I'm famous!

Indicator Wars

FLASH
TROOPERS
MAKE THE WORLD A SAFER PLACE



Raphaël Vinot

Help Desk and headdesk; ranting and coding
@rafi0t

Marion Marschalek

Threat dissector & professional PPT slide artist
@pinkflawd

CIRCL
Computer Incident
Response Center
Luxembourg

Cyber, Cyber & Sharing

Every vendor sells the best feed ever,
only sometimes, they contain new info.

The Cloud is where all your indicators go to die,

so your vendor can resell them :)

those glassy leaflets are expensive y'know

Difficult to compare

Depending on a single vendor,

… or a format that may turn out to be incompatible

… because sharing means caring

The WHYs of Information Sharing

MISP Threat Sharing.

-2015-2545: overview of current threats

3865

57460863-76dc-4272-8116-4ea302de0b81

CIRCL

CIRCL

alexandre.dulaunoy@circl.lu

tlp:white x circl:osint-feed x Type:OSINT x estimative-language:likelihood-probability="very-likely" x +

2016-05-25

Medium

Completed

All communities

OSINT - CVE-2015-2545: overview of current threats

Yes

0 (0)

Related Events

2016-05-27 (3883)

2016-05-23 (3844)

2016-05-06 (3828)

Orgc: CIRCL

Date: 2016-05-23

Info: OSINT - Operation Ke3chang Resurfaces With New TidePool Malware

212.7.217.10

webconncheck.myfw.us

be35b7882469ae4d9de233f75e7bebf211fddc2c

Almost no chance - remote - 01-05% 0 estimative-language:likelihood-probability="almost-no-chance"

Event 4425

reg.

MISP threat sharing platform is a free and open source software helping information sharing of threat and cyber security indicators.

METADATA

Metrics. Moarrrr metrics.

Wishlist

gimme all binaries that call LoadLibrary/GetProcAddress on multiple occasions

gimme all binaries that listen to a C&C command named "listprocesses"

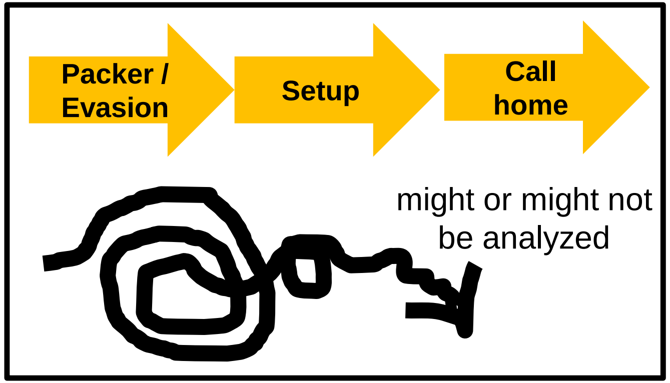
gimme all binaries with a code section entropy between 6.56778 and 6.60000

gimme binaries that call CryptEncrypt and contain the string www.maldomain.com

gimme all binaries that are able to list running processes, contain the string „Babar“, and were compiled before 2011

What my customer thought the malware does

What my sandbox thought the malware does



What I thought the malware does

What the malware REALLY does

```
Encrypting files
Keylogging
Screenshots
Screen captures
DDoS
Downloading more malware
```

We want

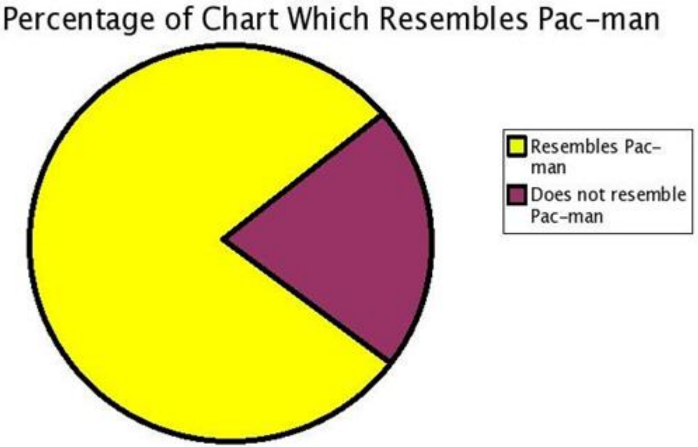
- Way to statically extract behavior information
- And general metrics
- Which are easily shared

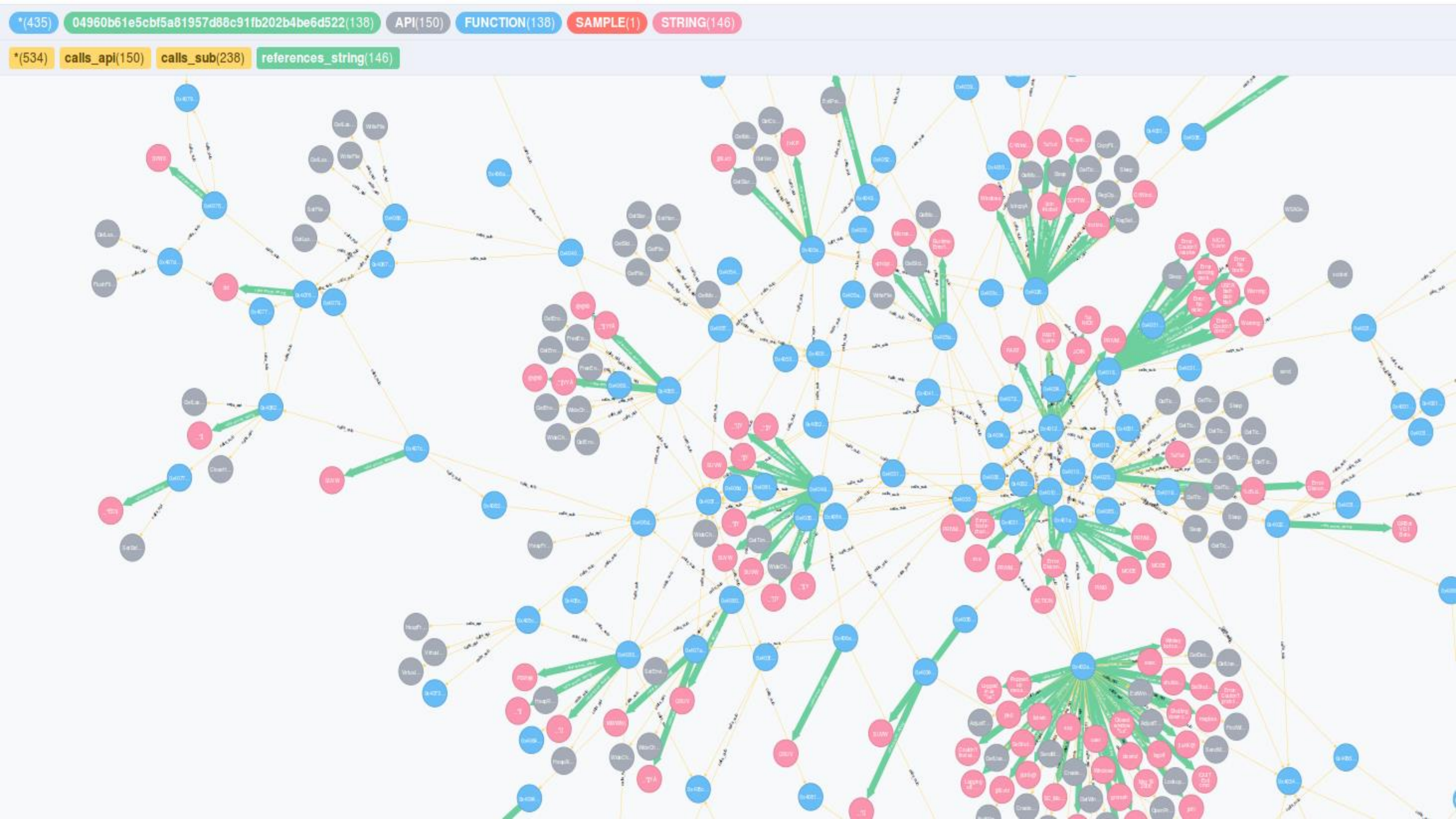
We did

- Plug a call graph generation tool into MISP
- Based on radare2
- Find and evaluate a _lot_ of indicators

- Function call graphs
  - Function cross references within code section
    - References to function offsets
    - Outside executable section(s)
- Nodes: functions
  - => Offset, size, calling convention
- Edges: calls, handler functions

Static Call Graphs




*(435)
04960b61e5cbf5a81957d88c91fb202b4be6d522(138)
API(150)
FUNCTION(138)
SAMPLE(1)
STRING(146)
*(534)
calls_api(150)
calls_sub(238)
references_string(146)

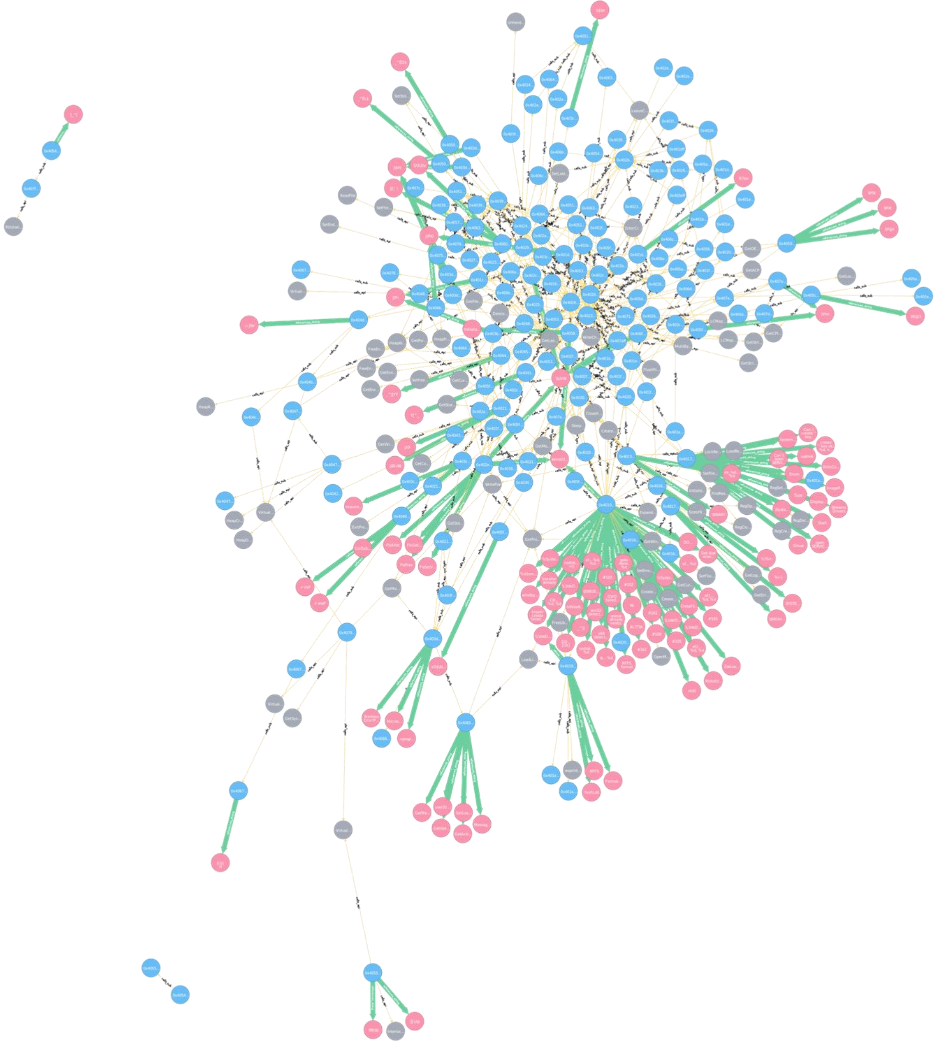
Neo4j & r2graphity

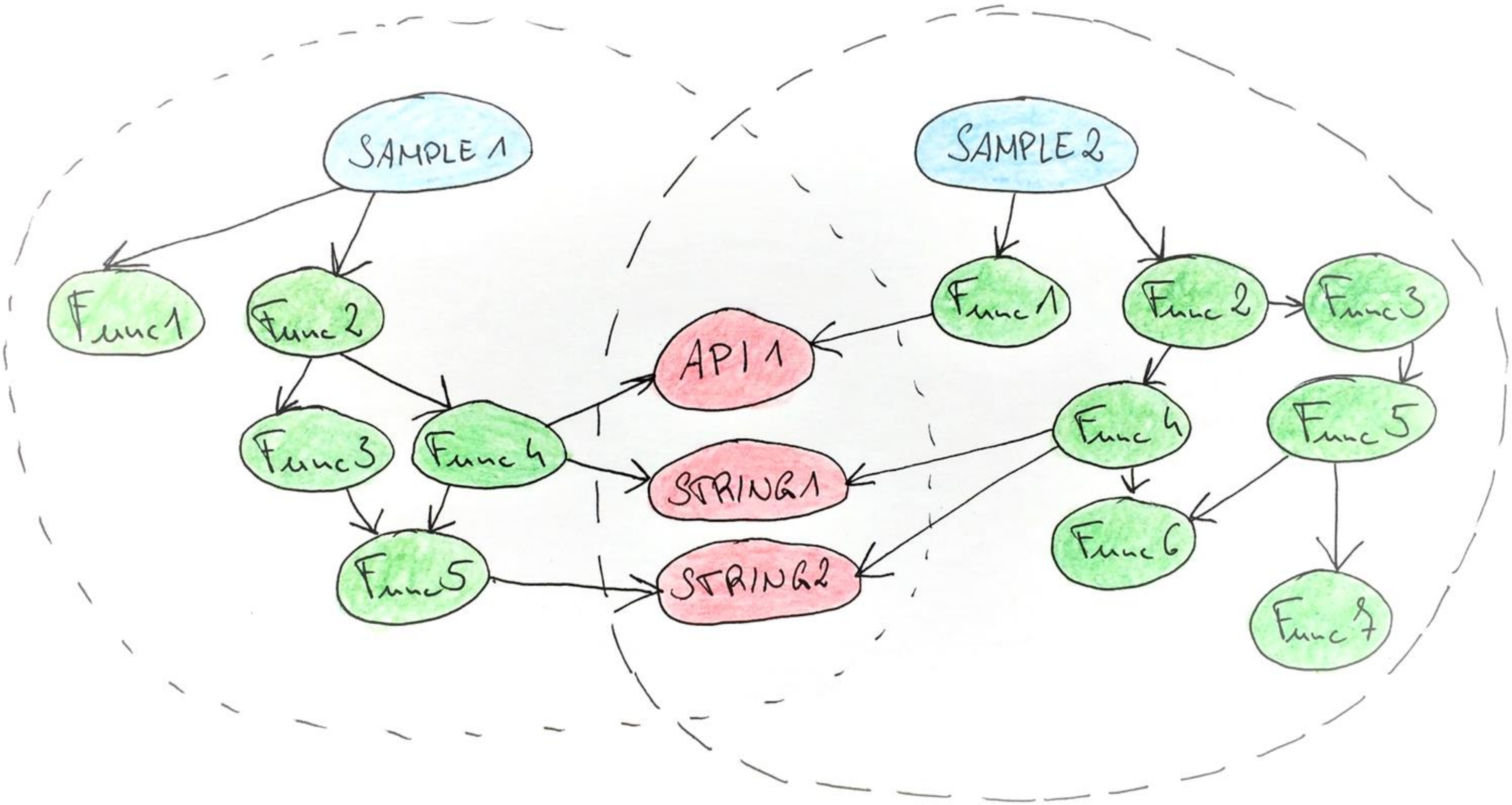
Parsing Windows PE call graphs to Neo4j
Functions, strings, API calls

“Fake” super node for graph separation & distance measuring

https://github.com/pinkflawd/r2graphity

Sofacy / APT28 as a test case
So many samples and incidents to pick from..




SAMPLE 1
Func1
Func 2
Func3
Func 4
Func5
API 1
STRING1
STRING2
SAMPLE 2
Func 1
Func 2
Func3
Func 4
Func 5
Func 6
Func 7

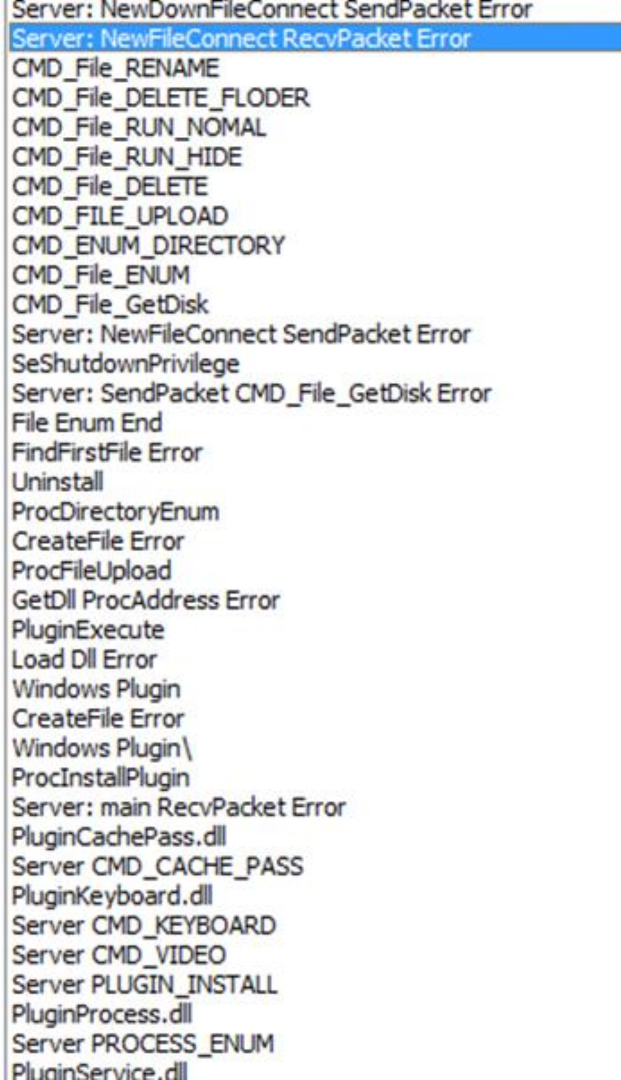
Strings

Strings with code cross references

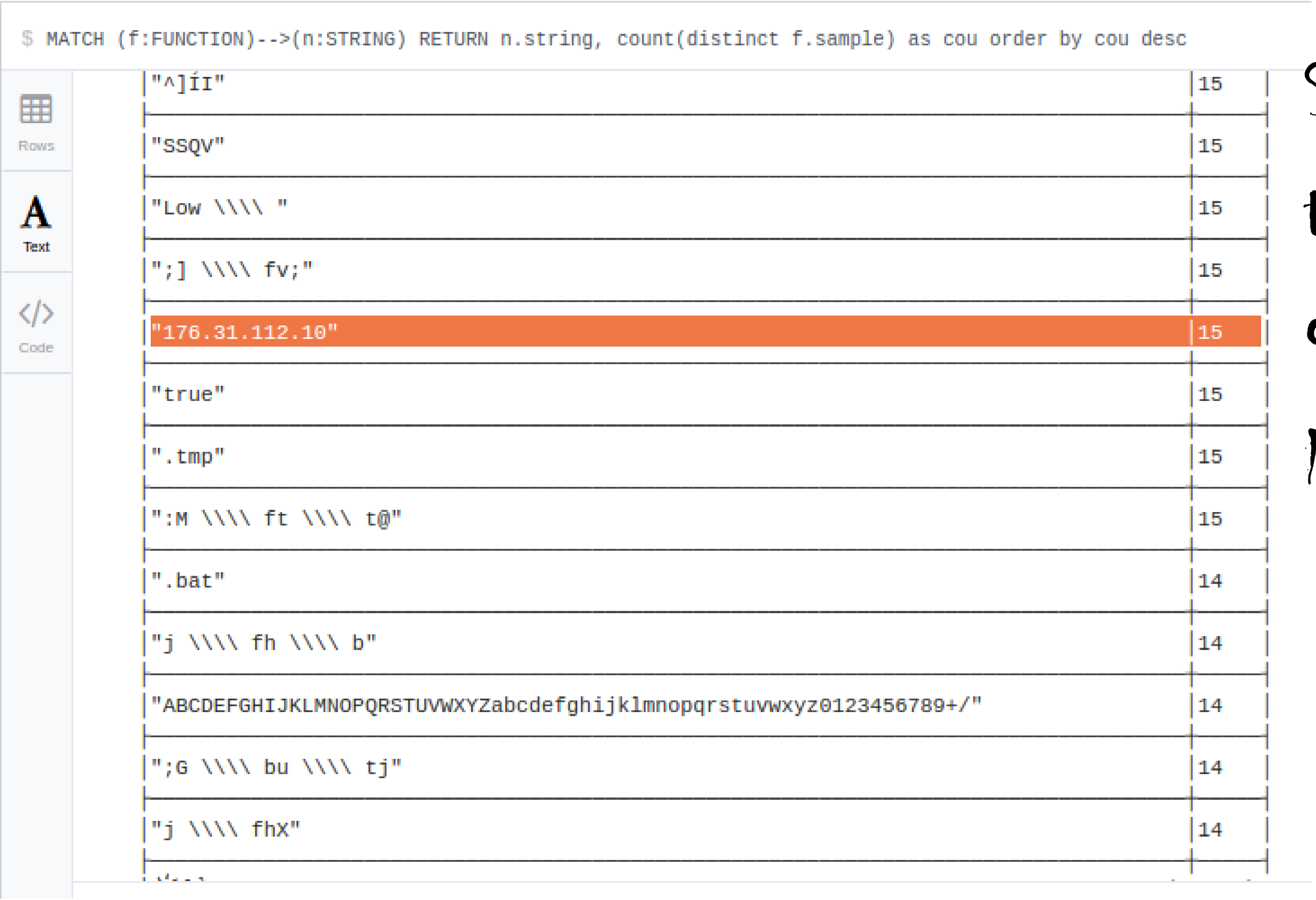
String list detection

- length + alignment
- following strings w/o cross references

Evaluation: ASCII, cross references, experimental character frequency test



```
Server: NewDownFileConnect SendPacket Error
Server: NewFileConnect RecvPacket Error
CMD_File_RENAME
CMD_File_DELETE_FLODER
CMD_File_RUN_NOMAL
CMD_File_RUN_HIDE
CMD_File_DELETE
CMD_FILE_UPLOAD
CMD_ENUM_DIRECTORY
CMD_File_ENUM
CMD_File_GetDisk
Server: NewFileConnect SendPacket Error
SeShutdownPrivilege
Server: SendPacket CMD_File_GetDisk Error
File Enum End
FindFirstFile Error
Uninstall
ProcDirectoryEnum
CreateFile Error
ProcFileUpload
GetDll ProcAddress Error
PluginExecute
Load Dll Error
Windows Plugin
CreateFile Error
Windows Plugin\
ProcInstallPlugin
Server: main RecvPacket Error
PluginCachePass.dll
Server CMD_CACHE_PASS
PluginKeyboard.dll
Server CMD_KEYBOARD
Server CMD_VIDEO
Server PLUGIN_INSTALL
PluginProcess.dll
Server PROCESS_ENUM
PluginService.dll
```

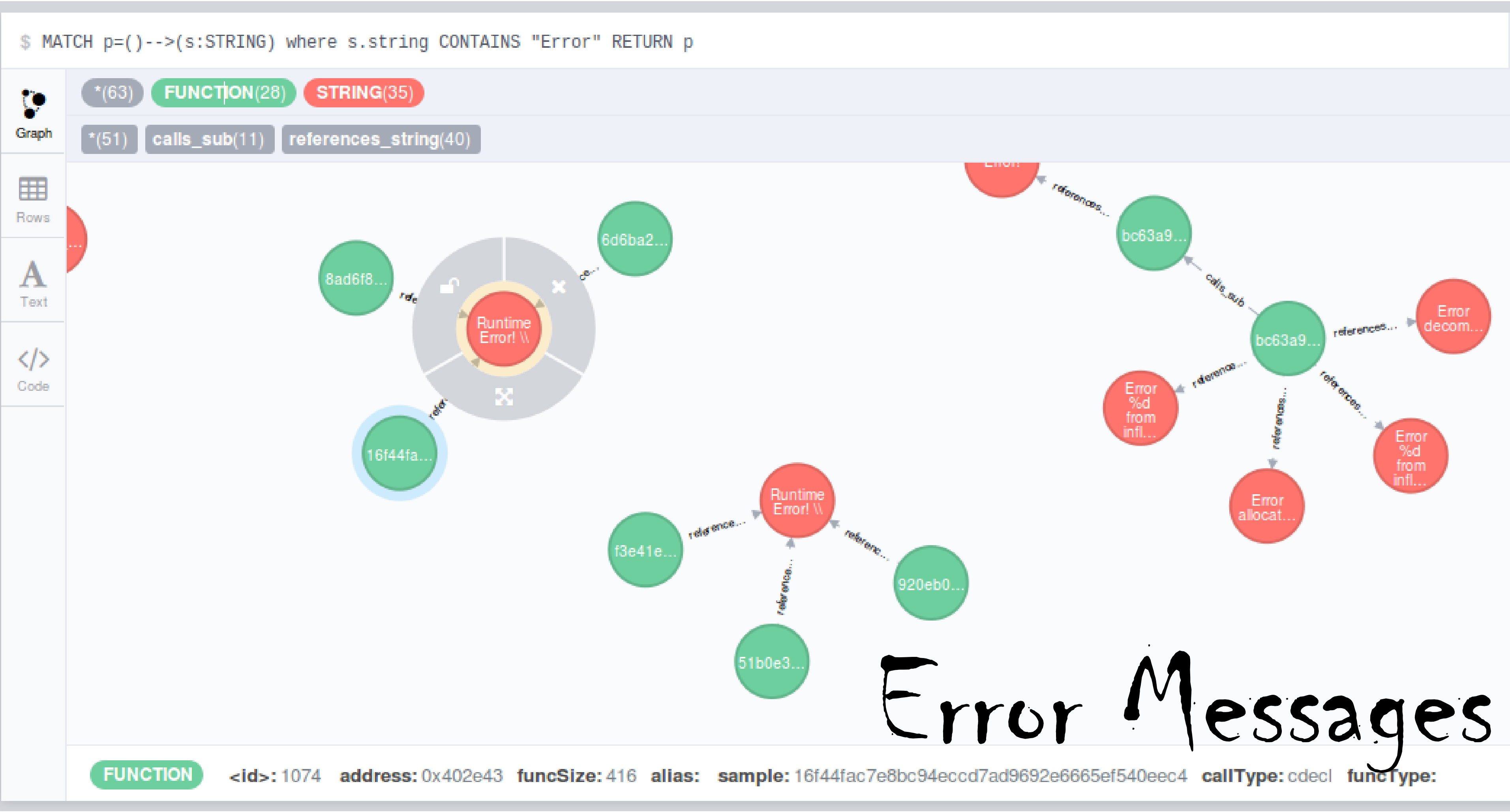


```
$ MATCH (f:FUNCTION)-->(n:STRING) RETURN n.string, count(distinct f.sample) as cou order by cou desc
```

```
│"^]ÍI"                                                               │15  │
├─────────────────────────────────────────────────────────────────────┼────┤
│"SSQV"                                                               │15  │
├─────────────────────────────────────────────────────────────────────┼────┤
│"Low \\\\ "                                                          │15  │
├─────────────────────────────────────────────────────────────────────┼────┤
│";] \\\\ fv;"                                                        │15  │
├─────────────────────────────────────────────────────────────────────┼────┤
│"176.31.112.10"                                                      │15  │
├─────────────────────────────────────────────────────────────────────┼────┤
│"true"                                                               │15  │
├─────────────────────────────────────────────────────────────────────┼────┤
│".tmp"                                                               │15  │
├─────────────────────────────────────────────────────────────────────┼────┤
│":M \\\\ ft \\\\ t@"                                                 │15  │
├─────────────────────────────────────────────────────────────────────┼────┤
│".bat"                                                               │14  │
├─────────────────────────────────────────────────────────────────────┼────┤
│"j \\\\ fh \\\\ b"                                                   │14  │
├─────────────────────────────────────────────────────────────────────┼────┤
│"ABCDEFGHIJKLMNOPQRSTUVWXYZabcdefghijklmnopqrstuvwxyz0123456789+/"   │14  │
├─────────────────────────────────────────────────────────────────────┼────┤
│";G \\\\ bu \\\\ tj"                                                 │14  │
├─────────────────────────────────────────────────────────────────────┼────┤
│"j \\\\ fhX"                                                         │14  │
├─────────────────────────────────────────────────────────────────────┼────┤
```

Strings and their occurence per sample

Error Messages

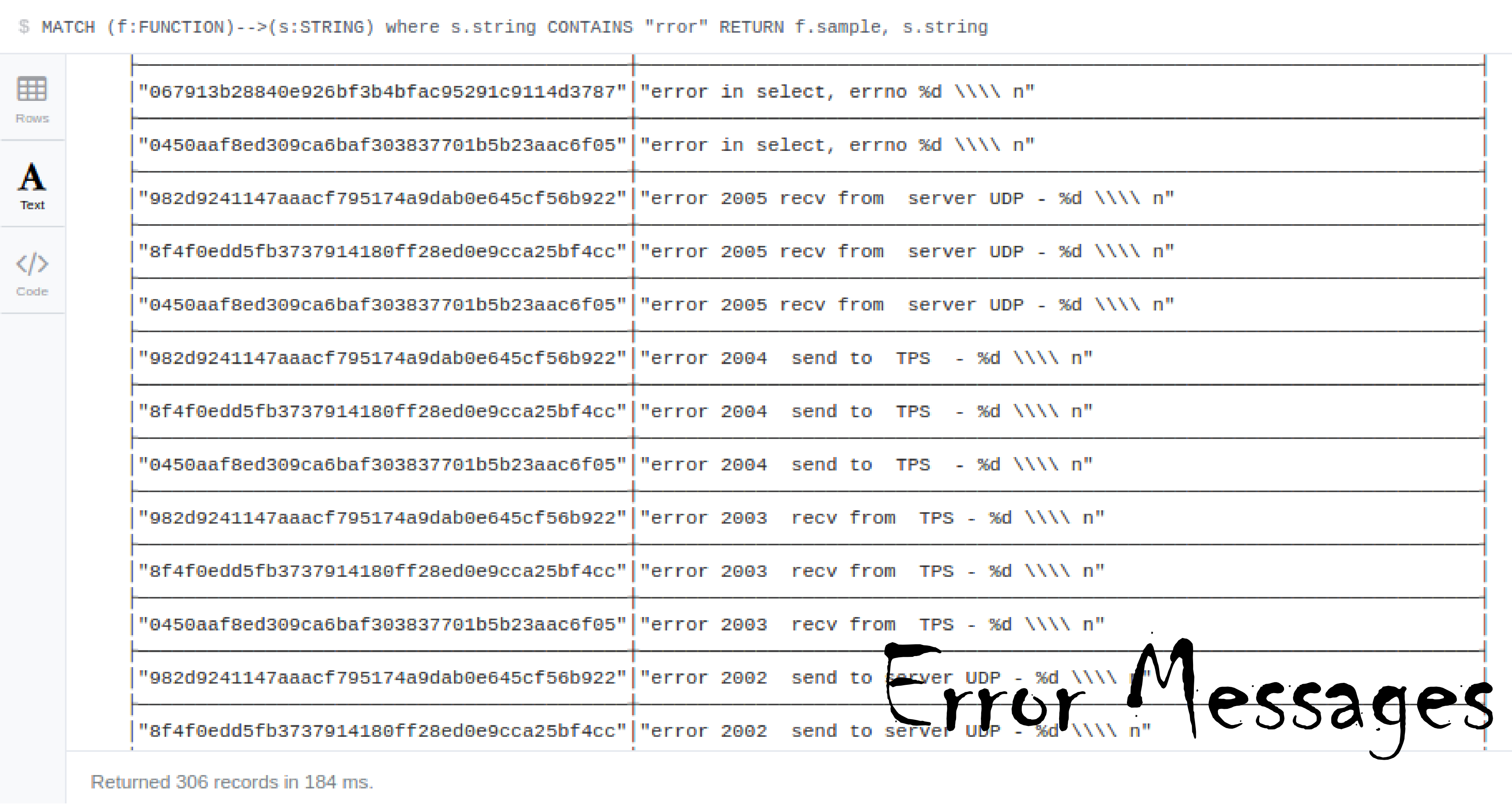


```
$ MATCH p=()-->(s:STRING) where s.string CONTAINS "Error" RETURN p
```

*(63) FUNCTION(28) STRING(35)

*(51) calls_sub(11) references_string(40)

FUNCTION <id>: 1074 address: 0x402e43 funcSize: 416 alias: sample: 16f44fac7e8bc94eccd7ad9692e6665ef540eec4 callType: cdecl funcType:

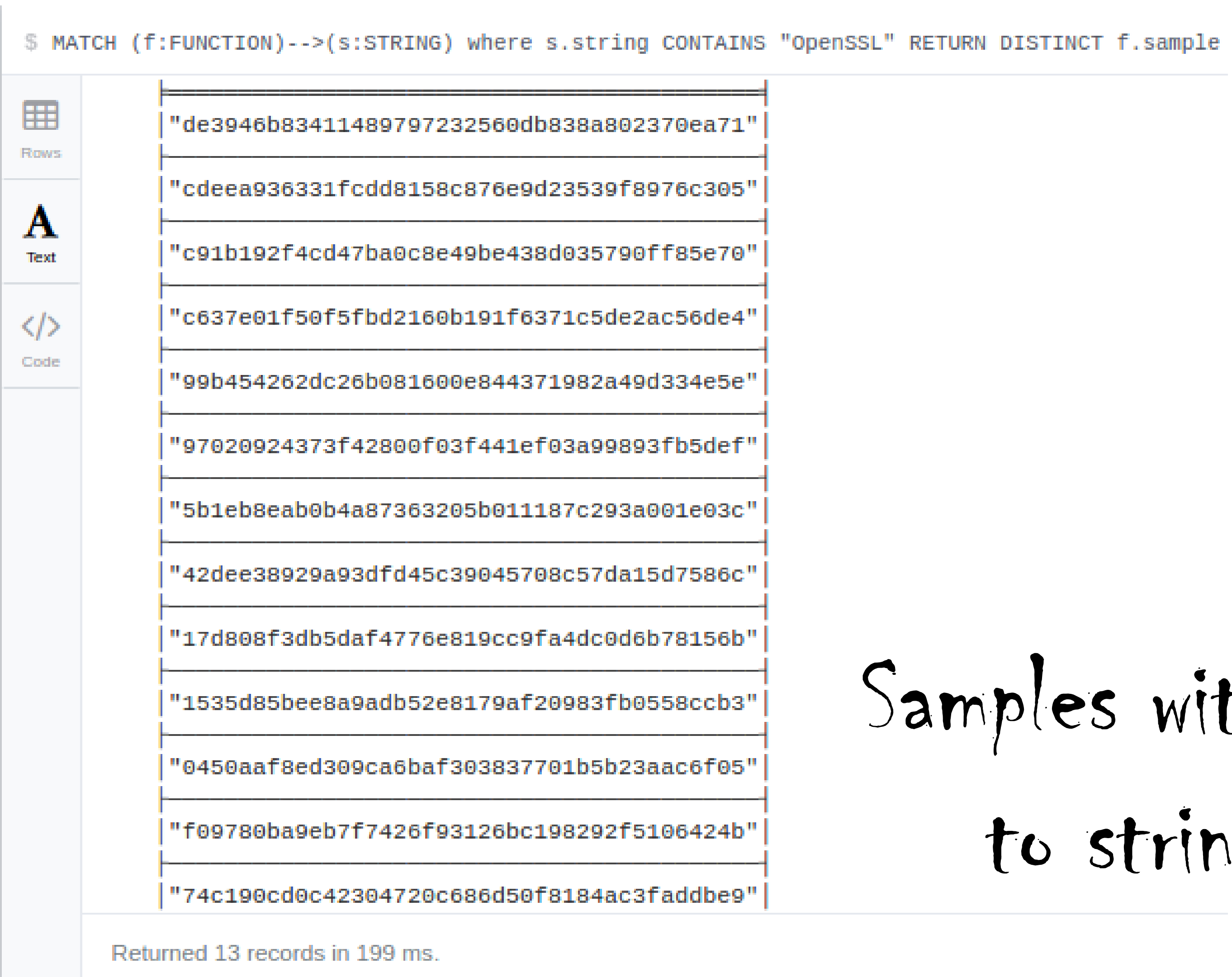


```
$ MATCH (f:FUNCTION)-->(s:STRING) where s.string CONTAINS "rror" RETURN f.sample, s.string
```

```
|"067913b28840e926bf3b4bfac95291c9114d3787"|"error in select, errno %d \\\\ n"                 |
|"0450aaf8ed309ca6baf303837701b5b23aac6f05"|"error in select, errno %d \\\\ n"                 |
|"982d9241147aaacf795174a9dab0e645cf56b922"|"error 2005 recv from  server UDP - %d \\\\ n"     |
|"8f4f0edd5fb3737914180ff28ed0e9cca25bf4cc"|"error 2005 recv from  server UDP - %d \\\\ n"     |
|"0450aaf8ed309ca6baf303837701b5b23aac6f05"|"error 2005 recv from  server UDP - %d \\\\ n"     |
|"982d9241147aaacf795174a9dab0e645cf56b922"|"error 2004  send to  TPS  - %d \\\\ n"            |
|"8f4f0edd5fb3737914180ff28ed0e9cca25bf4cc"|"error 2004  send to  TPS  - %d \\\\ n"            |
|"0450aaf8ed309ca6baf303837701b5b23aac6f05"|"error 2004  send to  TPS  - %d \\\\ n"            |
|"982d9241147aaacf795174a9dab0e645cf56b922"|"error 2003  recv from  TPS - %d \\\\ n"           |
|"8f4f0edd5fb3737914180ff28ed0e9cca25bf4cc"|"error 2003  recv from  TPS - %d \\\\ n"           |
|"0450aaf8ed309ca6baf303837701b5b23aac6f05"|"error 2003  recv from  TPS - %d \\\\ n"           |
|"982d9241147aaacf795174a9dab0e645cf56b922"|"error 2002  send to server UDP - %d \\\\ n"       |
|"8f4f0edd5fb3737914180ff28ed0e9cca25bf4cc"|"error 2002  send to server UDP - %d \\\\ n"       |
```

Returned 306 records in 184 ms.

Error Messages

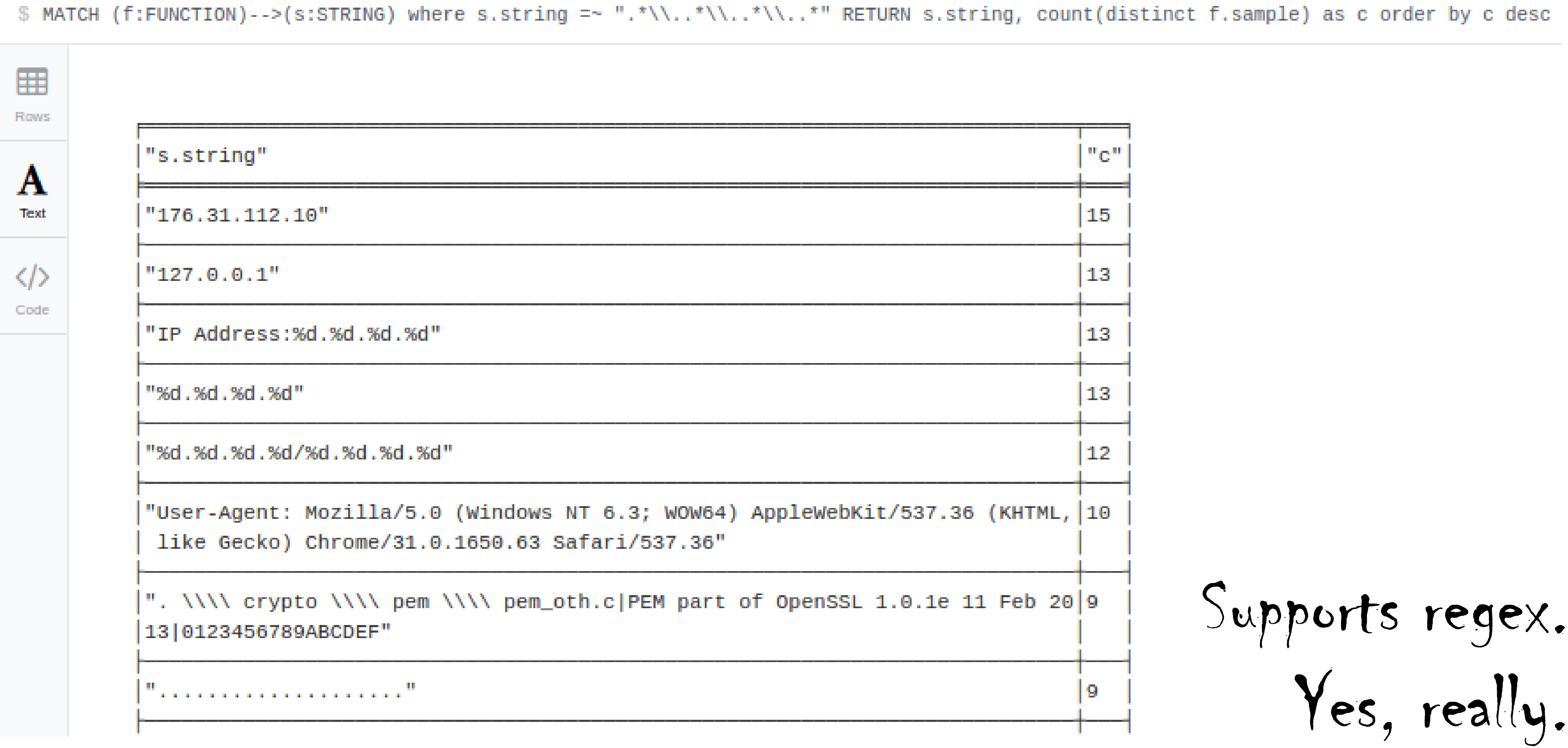


```
$ MATCH (f:FUNCTION)-->(s:STRING) where s.string CONTAINS "OpenSSL" RETURN DISTINCT f.sample
```

```
╞══════════════════════════════════════════╡
│"de3946b83411489797232560db838a802370ea71"│
├──────────────────────────────────────────┤
│"cdeea936331fcdd8158c876e9d23539f8976c305"│
├──────────────────────────────────────────┤
│"c91b192f4cd47ba0c8e49be438d035790ff85e70"│
├──────────────────────────────────────────┤
│"c637e01f50f5fbd2160b191f6371c5de2ac56de4"│
├──────────────────────────────────────────┤
│"99b454262dc26b081600e844371982a49d334e5e"│
├──────────────────────────────────────────┤
│"97020924373f42800f03f441ef03a99893fb5def"│
├──────────────────────────────────────────┤
│"5b1eb8eab0b4a87363205b011187c293a001e03c"│
├──────────────────────────────────────────┤
│"42dee38929a93dfd45c39045708c57da15d7586c"│
├──────────────────────────────────────────┤
│"17d808f3db5daf4776e819cc9fa4dc0d6b78156b"│
├──────────────────────────────────────────┤
│"1535d85bee8a9adb52e8179af20983fb0558ccb3"│
├──────────────────────────────────────────┤
│"0450aaf8ed309ca6baf303837701b5b23aac6f05"│
├──────────────────────────────────────────┤
│"f09780ba9eb7f7426f93126bc198292f5106424b"│
├──────────────────────────────────────────┤
│"74c190cd0c42304720c686d50f8184ac3faddbe9"│
```

Returned 13 records in 199 ms.

Samples with references
to string 'OpenSSL'



```
$ MATCH (f:FUNCTION)-->(s:STRING) where s.string =~ ".*\\..*\\..*\\..*" RETURN s.string, count(distinct f.sample) as c order by c desc
```

```
╒═══════════════════════════════════════════════════════════════════════════╤═══╕
│"s.string"                                                                 │"c"│
╞═══════════════════════════════════════════════════════════════════════════╪═══╡
│"176.31.112.10"                                                            │15 │
├───────────────────────────────────────────────────────────────────────────┼───┤
│"127.0.0.1"                                                                │13 │
├───────────────────────────────────────────────────────────────────────────┼───┤
│"IP Address:%d.%d.%d.%d"                                                   │13 │
├───────────────────────────────────────────────────────────────────────────┼───┤
│"%d.%d.%d.%d"                                                              │13 │
├───────────────────────────────────────────────────────────────────────────┼───┤
│"%d.%d.%d.%d/%d.%d.%d.%d"                                                  │12 │
├───────────────────────────────────────────────────────────────────────────┼───┤
│"User-Agent: Mozilla/5.0 (Windows NT 6.3; WOW64) AppleWebKit/537.36 (KHTML,│10 │
│ like Gecko) Chrome/31.0.1650.63 Safari/537.36"                            │   │
├───────────────────────────────────────────────────────────────────────────┼───┤
│". \\\\ crypto \\\\ pem \\\\ pem_oth.c|PEM part of OpenSSL 1.0.1e 11 Feb 20│9  │
│13|0123456789ABCDEF"                                                       │   │
├───────────────────────────────────────────────────────────────────────────┼───┤
│"...................."                                                     │9  │
├───────────────────────────────────────────────────────────────────────────┼───┤
```

Supports regex.

Yes, really.

APIs

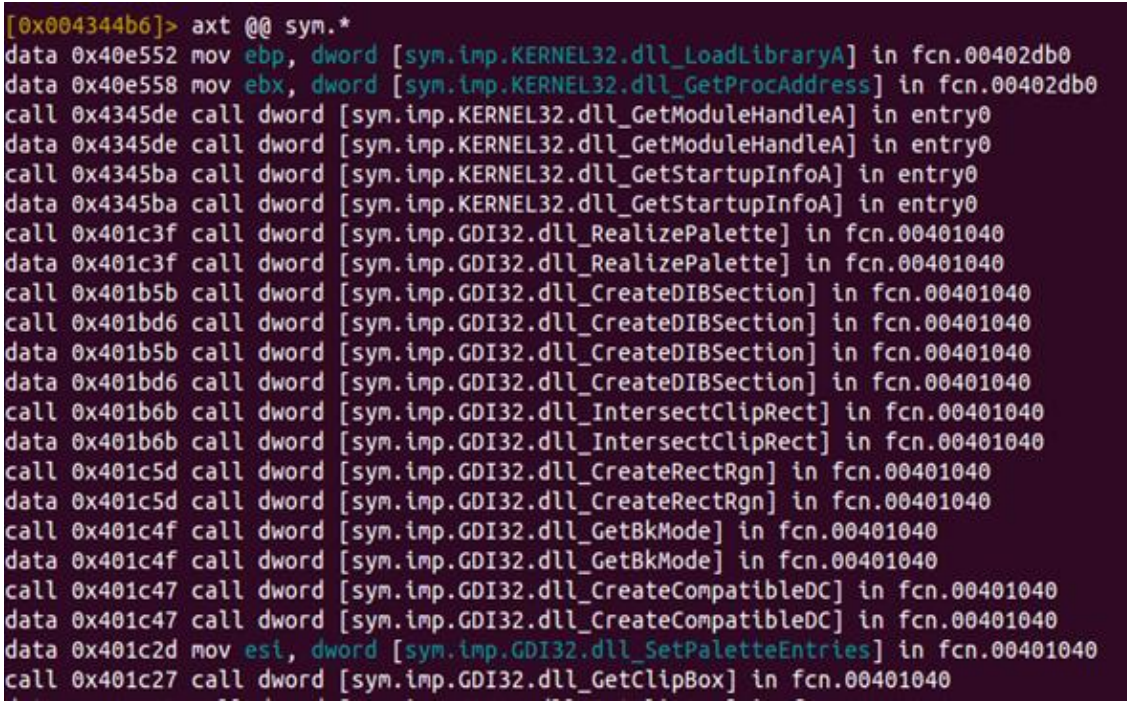
Cross references on symbols

Indirect calls

- parsing for mov/lea
- disassembling further
- call and jmp considered xref

Thunk pruning

Dynamic loading



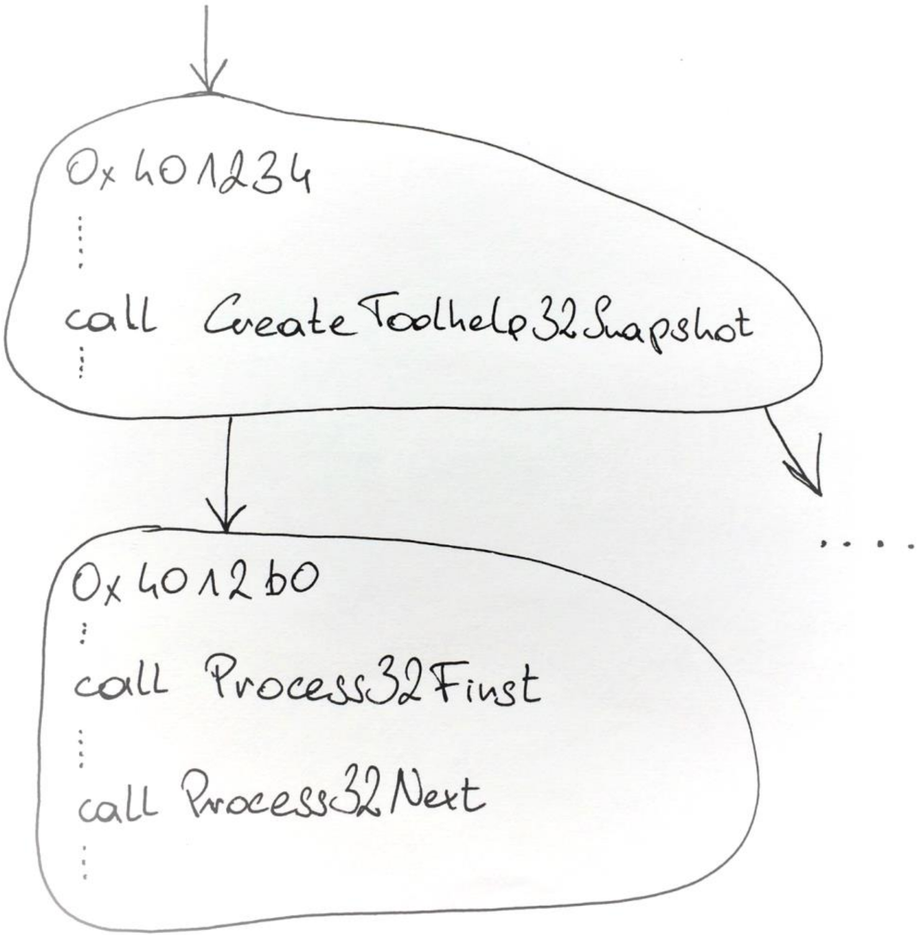
```
[0x004344b6]> axt @@ sym.*
data 0x40e552 mov ebp, dword [sym.imp.KERNEL32.dll_LoadLibraryA] in fcn.00402db0
data 0x40e558 mov ebx, dword [sym.imp.KERNEL32.dll_GetProcAddress] in fcn.00402db0
call 0x4345de call dword [sym.imp.KERNEL32.dll_GetModuleHandleA] in entry0
data 0x4345de call dword [sym.imp.KERNEL32.dll_GetModuleHandleA] in entry0
call 0x4345ba call dword [sym.imp.KERNEL32.dll_GetStartupInfoA] in entry0
data 0x4345ba call dword [sym.imp.KERNEL32.dll_GetStartupInfoA] in entry0
call 0x401c3f call dword [sym.imp.GDI32.dll_RealizePalette] in fcn.00401040
data 0x401c3f call dword [sym.imp.GDI32.dll_RealizePalette] in fcn.00401040
call 0x401b5b call dword [sym.imp.GDI32.dll_CreateDIBSection] in fcn.00401040
call 0x401bd6 call dword [sym.imp.GDI32.dll_CreateDIBSection] in fcn.00401040
data 0x401b5b call dword [sym.imp.GDI32.dll_CreateDIBSection] in fcn.00401040
data 0x401bd6 call dword [sym.imp.GDI32.dll_CreateDIBSection] in fcn.00401040
call 0x401b6b call dword [sym.imp.GDI32.dll_IntersectClipRect] in fcn.00401040
data 0x401b6b call dword [sym.imp.GDI32.dll_IntersectClipRect] in fcn.00401040
call 0x401c5d call dword [sym.imp.GDI32.dll_CreateRectRgn] in fcn.00401040
data 0x401c5d call dword [sym.imp.GDI32.dll_CreateRectRgn] in fcn.00401040
call 0x401c4f call dword [sym.imp.GDI32.dll_GetBkMode] in fcn.00401040
data 0x401c4f call dword [sym.imp.GDI32.dll_GetBkMode] in fcn.00401040
call 0x401c47 call dword [sym.imp.GDI32.dll_CreateCompatibleDC] in fcn.00401040
data 0x401c47 call dword [sym.imp.GDI32.dll_CreateCompatibleDC] in fcn.00401040
data 0x401c2d mov esi, dword [sym.imp.GDI32.dll_SetPaletteEntries] in fcn.00401040
call 0x401c27 call dword [sym.imp.GDI32.dll_GetClipBox] in fcn.00401040
```

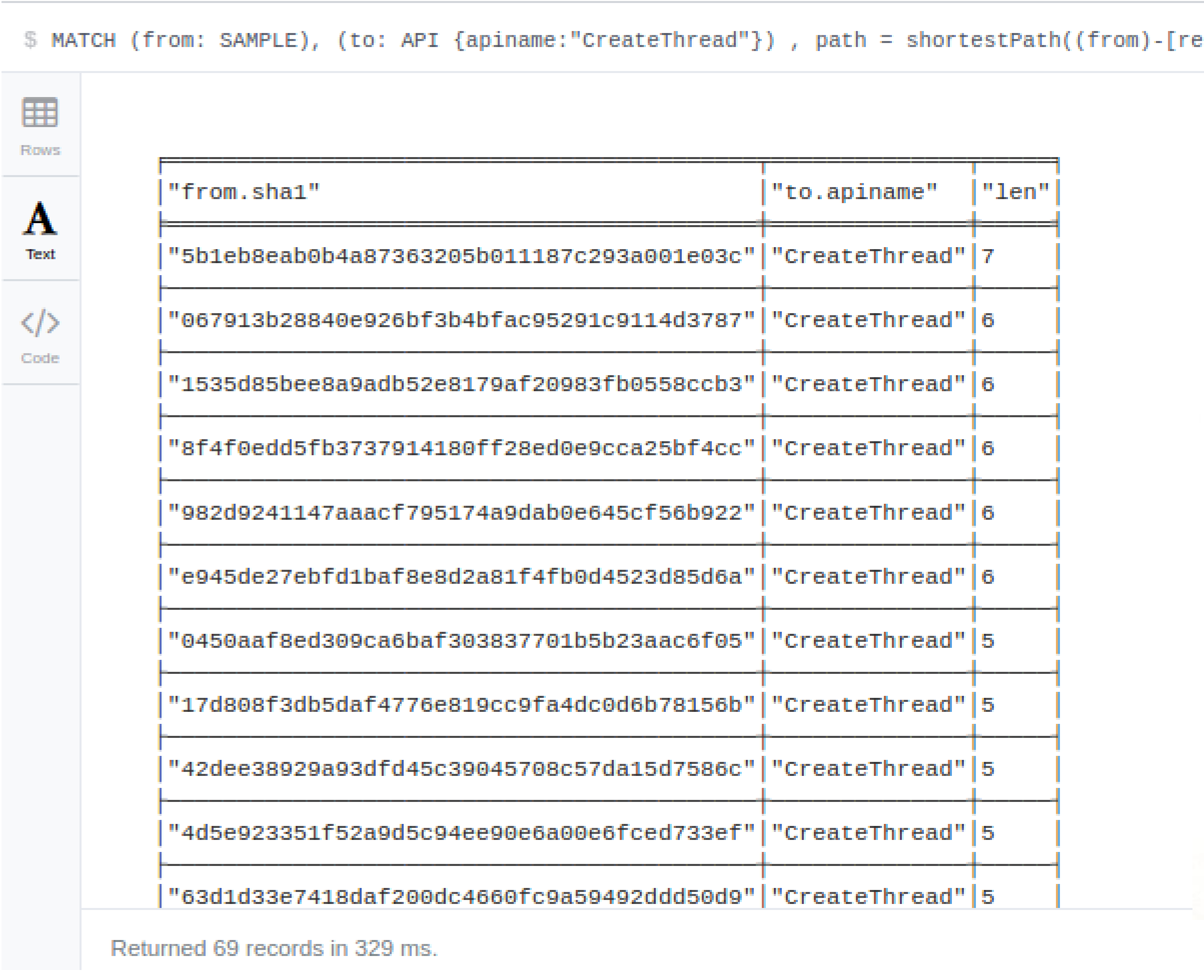
"Behavior" Gadgets

```
For APILOADING found {'GetProcAddress': '0x1000def8', 'LoadLibrary': '0x1000def8'}
For APILOADING found {'GetProcAddress': '0x10014e88', 'LoadLibrary': '0x10014e88'}
For READFILE found {'ReadFile': '0x100032a0', 'CreateFile': '0x100032a0'}
For READFILE found {'ReadFile': '0x1000d6b0', 'CreateFile': '0x1000d6b0'}
For APILOADING2 found {'GetModuleHandle': '0x1000fbd3', 'GetProcAddress': '0x1000fbd3'}
For APILOADING2 found {'GetModuleHandle': '0x1000f8ef', 'GetProcAddress': '0x1000fbd3'}
For APILOADING2 found {'GetModuleHandle': '0x10012552', 'GetProcAddress': '0x10012552'}
For SHELLEXEC found {'ShellExecute': '0x1000d330'}
For FILEITER found {'FindClose': '0x1000d330', 'FindFirstFile': '0x1000d330', 'FindNextFile':
'0x1000d330'}
For CREATETHREAD found {'CreateThread': '0x1000ebc2'}
For CREATETHREAD found {'CreateThread': '0x10009b10'}
For CREATETHREAD found {'CreateThread': '0x10002190'}
For CREATETHREAD found {'CreateThread': '0x1000a050'}
For CREATETHREAD found {'CreateThread': '0x10001820'}
For CREATETHREAD found {'CreateThread': '0x10001000'}
For WRITEFILE found {'WriteFile': '0x1000d880', 'CreateFile': '0x1000d880'}
For WRITEFILE found {'WriteFile': '0x1000a4f0', 'CreateFile': '0x1000a4f0'}
For WRITEFILE found {'WriteFile': '0x10001f80', 'CreateFile': '0x10001f80'}
For RECV found {'recv': '0x1000b290', 'send': '0x1000b290'}
For SCREENSHOT found {'GetDeviceCaps': '0x100094d0', 'CreateCompatibleBitmap':
    '0x100094d0', 'BitBlt': '0x100094d0', 'CreateCompatibleDC': '0x100094d0'}
For REGQUERY found {'RegOpenKey': '0x10001000', 'RegQueryValue': '0x10001000'}
```

Scanning for Gadgets

Pre-defined API patterns
Searching the graph for anchor
Scanning nodes in close vicinity

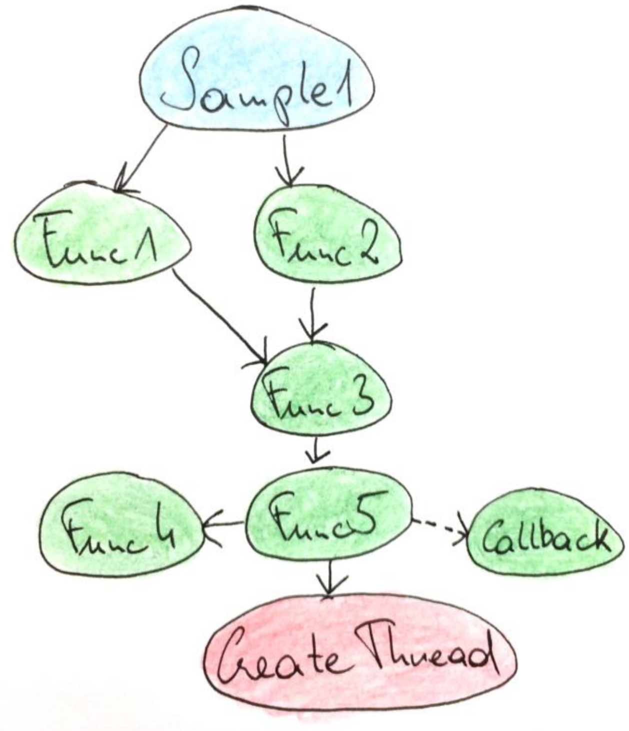




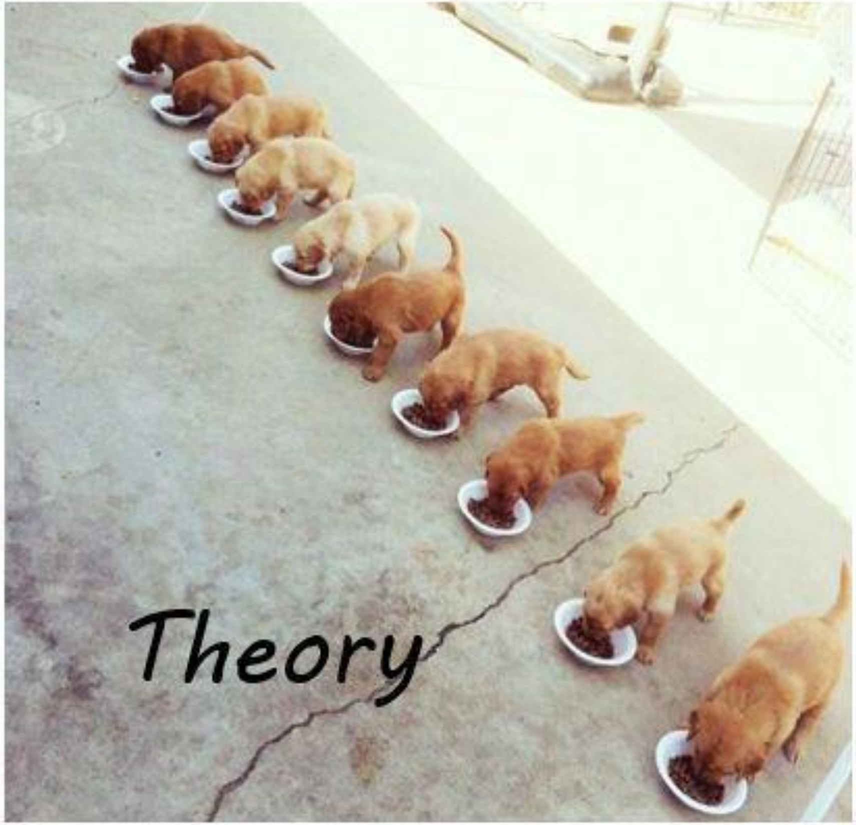
```
$ MATCH (from: SAMPLE), (to: API {apiname:"CreateThread"}) , path = shortestPath((from)-[rels*]->(to)) RETURN from.sha1, to.apiname, length(path)
```

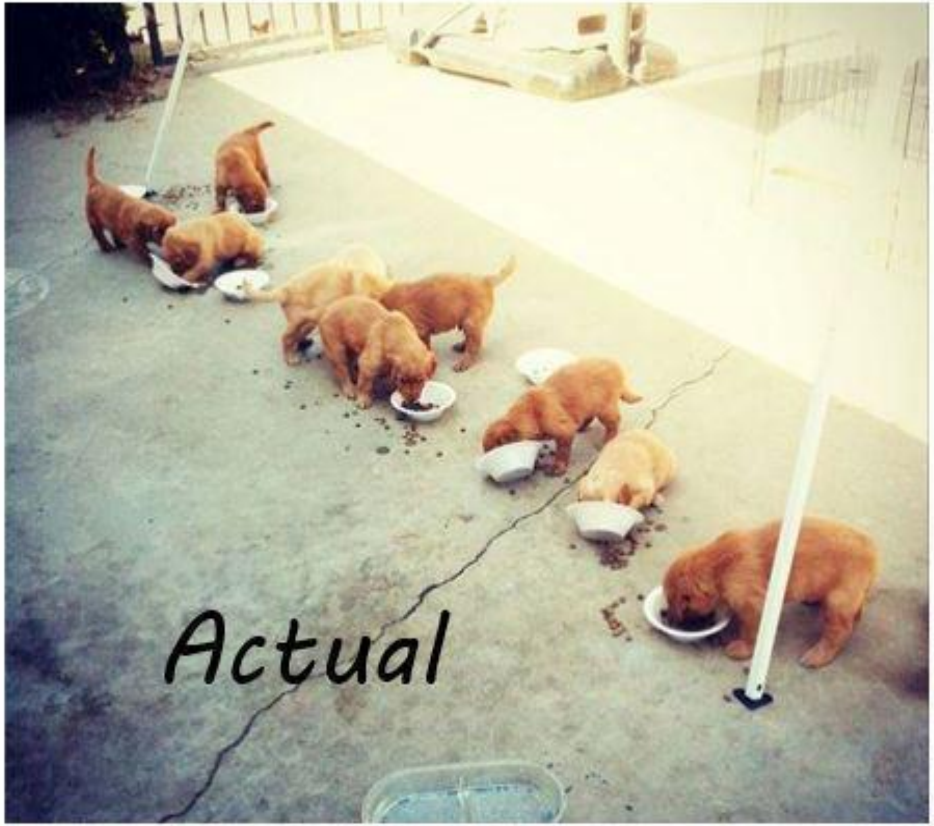
```
"from.sha1"                                 "to.apiname"    "len"
"5b1eb8eab0b4a87363205b011187c293a001e03c"  "CreateThread"  7
"067913b28840e926bf3b4bfac95291c9114d3787"  "CreateThread"  6
"1535d85bee8a9adb52e8179af20983fb0558ccb3"  "CreateThread"  6
"8f4f0edd5fb3737914180ff28ed0e9cca25bf4cc"  "CreateThread"  6
"982d9241147aaacf795174a9dab0e645cf56b922"  "CreateThread"  6
"e945de27ebfd1baf8e8d2a81f4fb0d4523d85d6a"  "CreateThread"  6
"0450aaf8ed309ca6baf303837701b5b23aac6f05"  "CreateThread"  5
"17d808f3db5daf4776e819cc9fa4dc0d6b78156b"  "CreateThread"  5
"42dee38929a93dfd45c39045708c57da15d7586c"  "CreateThread"  5
"4d5e923351f52a9d5c94ee90e6a00e6fced733ef"  "CreateThread"  5
"63d1d33e7418daf200dc4660fc9a59492ddd50d9"  "CreateThread"  5
```

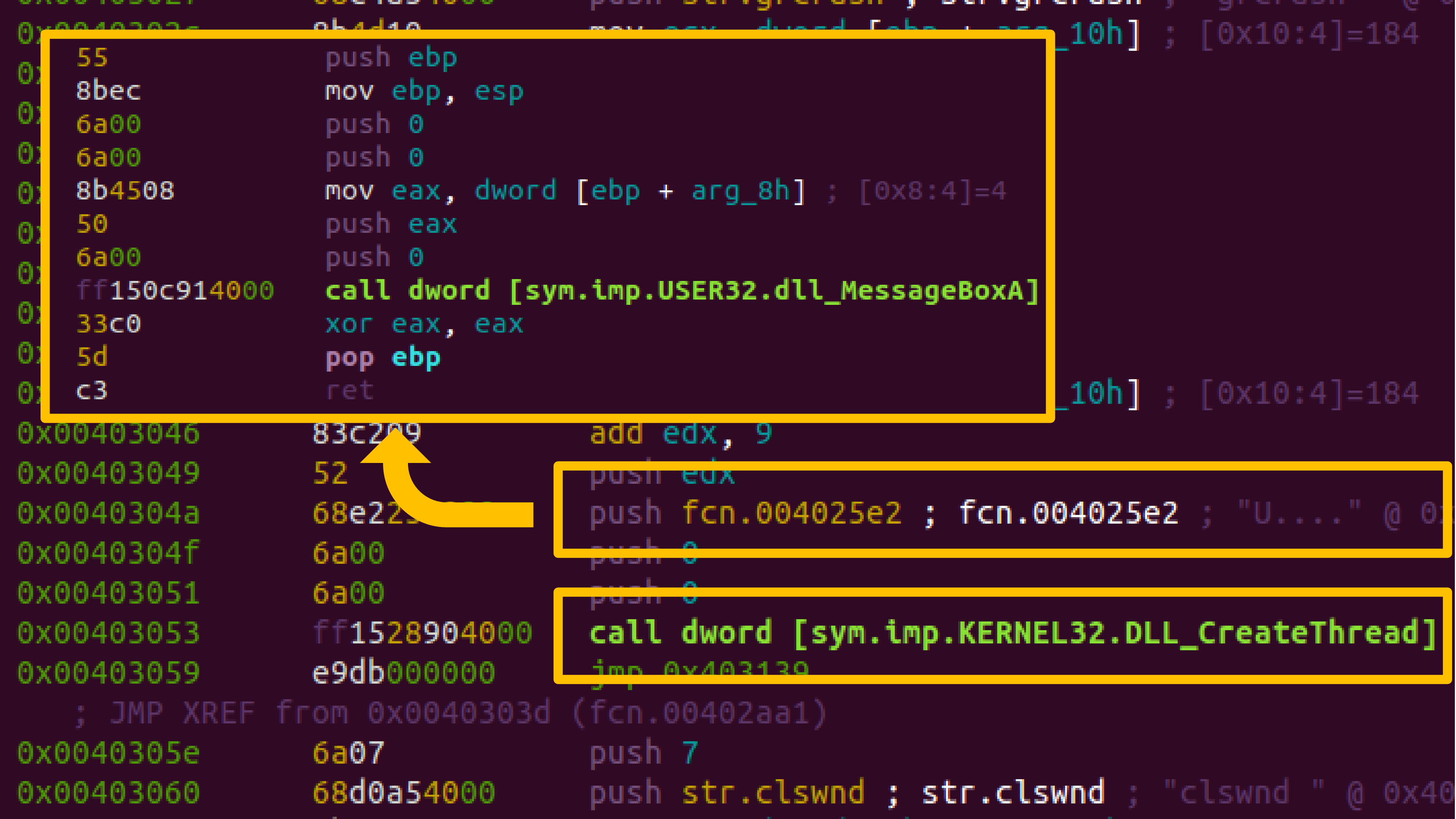
Returned 69 records in 329 ms.



Multithreaded programming







```
0x0040302c   8b4d10        mov ecx, dword [ebp + arg_10h] ; [0x10:4]=184
```

```
55                push ebp
8bec              mov ebp, esp
6a00              push 0
6a00              push 0
8b4508            mov eax, dword [ebp + arg_8h] ; [0x8:4]=4
50                push eax
6a00              push 0
ff150c914000      call dword [sym.imp.USER32.dll_MessageBoxA]
33c0              xor eax, eax
5d                pop ebp
c3                ret
```

```
                                 _10h] ; [0x10:4]=184
0x00403046   83c209        add edx, 9
0x00403049   52            push edx
0x0040304a   68e22         push fcn.004025e2 ; fcn.004025e2 ; "U...." @ 0x
0x0040304f   6a00          push 0
0x00403051   6a00          push 0
0x00403053   ff1528904000  call dword [sym.imp.KERNEL32.DLL_CreateThread]
0x00403059   e9db000000    jmp 0x403139
   ; JMP XREF from 0x0040303d (fcn.00402aa1)
0x0040305e   6a07          push 7
0x00403060   68d0a54000    push str.clswnd ; str.clswnd ; "clswnd " @ 0x40
```

Thread Model Modelling

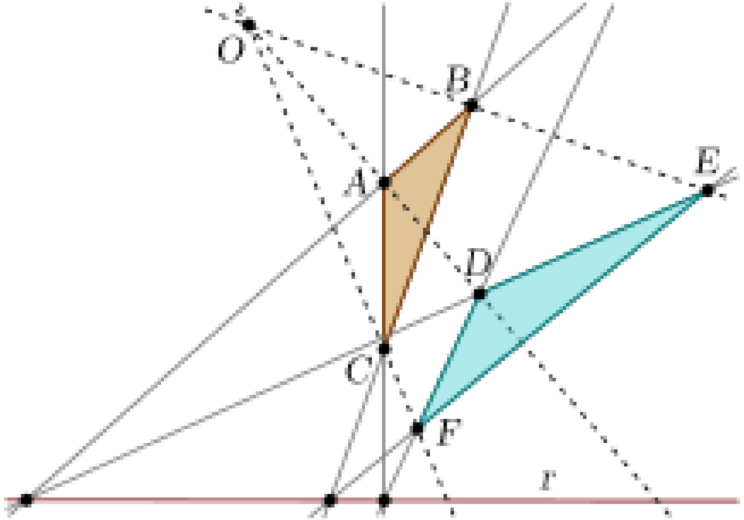
Number of calls to CreateThread

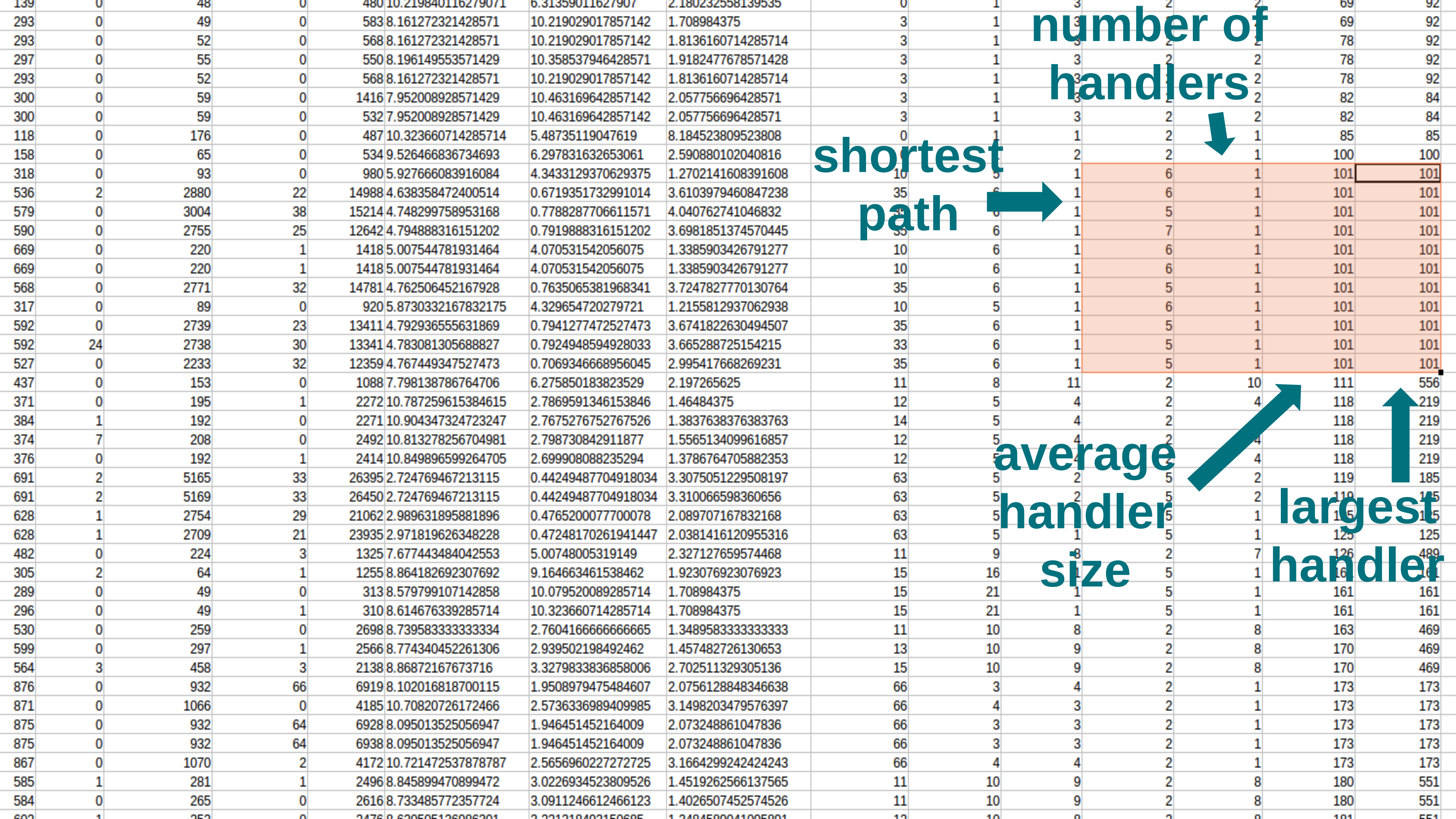
Shortest path to CreateThread

Number of handler functions

Average size of handler functions

Size of biggest handler function





139	0	48	0	480	10.219840116279071	6.31359011627907	2.180232558139535	0	1	3	2	2	69	92
293	0	49	0	583	8.161272321428571	10.219029017857142	1.708984375	3	1	3	2	2	69	92
293	0	52	0	568	8.161272321428571	10.219029017857142	1.8136160714285714	3	1	3	2	2	78	92
297	0	55	0	550	8.196149553571429	10.358537946428571	1.9182477678571428	3	1	3	2	2	78	92
293	0	52	0	568	8.161272321428571	10.219029017857142	1.8136160714285714	3	1	3	2	2	78	92
300	0	59	0	1416	7.952008928571429	10.463169642857142	2.057756696428571	3	1	3	2	2	82	84
300	0	59	0	532	7.952008928571429	10.463169642857142	2.057756696428571	3	1	3	2	2	82	84
118	0	176	0	487	10.323660714285714	5.48735119047619	8.184523809523808	0	1	1	2	1	85	85
158	0	65	0	534	9.526466836734693	6.297831632653061	2.590880102040816	0	1	2	2	1	100	100
318	0	93	0	980	5.927666083916084	4.3433129370629375	1.2702141608391608	10	5	1	6	1	101	101
536	2	2880	22	14988	4.638358472400514	0.6719351732991014	3.6103979460847238	35	6	1	6	1	101	101
579	0	3004	38	15214	4.748299758953168	0.7788287706611571	4.040762741046832	35	6	1	5	1	101	101
590	0	2755	25	12642	4.794888316151202	0.7919888316151202	3.6981851374570445	35	6	1	7	1	101	101
669	0	220	1	1418	5.007544781931464	4.070531542056075	1.3385903426791277	10	6	1	6	1	101	101
669	0	220	1	1418	5.007544781931464	4.070531542056075	1.3385903426791277	10	6	1	6	1	101	101
568	0	2771	32	14781	4.762506452167928	0.7635065381968341	3.7247827770130764	35	6	1	5	1	101	101
317	0	89	0	920	5.8730332167832175	4.329654720279721	1.2155812937062938	10	5	1	6	1	101	101
592	0	2739	23	13411	4.792936555631869	0.7941277472527473	3.6741822630494507	35	6	1	5	1	101	101
592	24	2738	30	13341	4.783081305688827	0.7924948594928033	3.665288725154215	33	6	1	5	1	101	101
527	0	2233	32	12359	4.767449347527473	0.7069346668956045	2.995417668269231	35	6	1	5	1	101	101
437	0	153	0	1088	7.798138786764706	6.275850183823529	2.197265625	11	8	11	2	10	111	556
371	0	195	1	2272	10.787259615384615	2.7869591346153846	1.46484375	12	5	4	2	4	118	219
384	1	192	0	2271	10.904347324723247	2.7675276752767526	1.3837638376383763	14	5	4	2	4	118	219
374	7	208	0	2492	10.813278256704981	2.798730842911877	1.5565134099616857	12	5	4	2	4	118	219
376	0	192	1	2414	10.849896599264705	2.699908088235294	1.3786764705882353	12	5	4	2	4	118	219
691	2	5165	33	26395	2.724769467213115	0.44249487704918034	3.3075051229508197	63	5	2	5	2	119	185
691	2	5169	33	26450	2.724769467213115	0.44249487704918034	3.310066598360656	63	5	2	5	2	119	185
628	1	2754	29	21062	2.989631895881896	0.4765200077700078	2.089707167832168	63	5	1	5	1	125	125
628	1	2709	21	23935	2.971819626348228	0.47248170261941447	2.0381416120955316	63	5	1	5	1	125	125
482	0	224	3	1325	7.677443484042553	5.00748005319149	2.327127659574468	11	9	8	2	7	126	489
305	2	64	1	1255	8.864182692307692	9.164663461538462	1.923076923076923	15	16	1	5	1	161	161
289	0	49	0	313	8.579799107142858	10.079520089285714	1.708984375	15	21	1	5	1	161	161
296	0	49	1	310	8.614676339285714	10.323660714285714	1.708984375	15	21	1	5	1	161	161
530	0	259	0	2698	8.739583333333334	2.7604166666666665	1.3489583333333333	11	10	8	2	8	163	469
599	0	297	1	2566	8.774340452261306	2.939502198492462	1.457482726130653	13	10	9	2	8	170	469
564	3	458	3	2138	8.86872167673716	3.3279833836858006	2.702511329305136	15	10	9	2	8	170	469
876	0	932	66	6919	8.102016818700115	1.9508979475484607	2.0756128848346638	66	3	4	2	1	173	173
871	0	1066	0	4185	10.70820726172466	2.5736336989409985	3.1498203479576397	66	4	3	2	1	173	173
875	0	932	64	6928	8.095013525056947	1.946451452164009	2.073248861047836	66	3	3	2	1	173	173
875	0	932	64	6938	8.095013525056947	1.946451452164009	2.073248861047836	66	3	3	2	1	173	173
867	0	1070	2	4172	10.721472537878787	2.5656960227272725	3.1664299242424243	66	4	4	2	1	173	173
585	1	281	1	2496	8.845899470899472	3.0226934523809526	1.4519262566137565	11	10	9	2	8	180	551
584	0	265	0	2616	8.733485772357724	3.0911246612466123	1.4026507452574526	11	10	9	2	8	180	551
602	1	252	0	2476	8.620505126086201	3.221318403150685	1.2494580041005801	12	10	9	2	8	181	551

number of handlers

shortest path

average handler size

largest handler

A Feature Factory

“Build it simple, then scale it up.” - Smart guy from Google

Performance?
Scalability?
Robustness?

No, we don't do machine learning

Yes, its built on top of radare2

The feature "flattening" process

Its fast, but not extremely fast

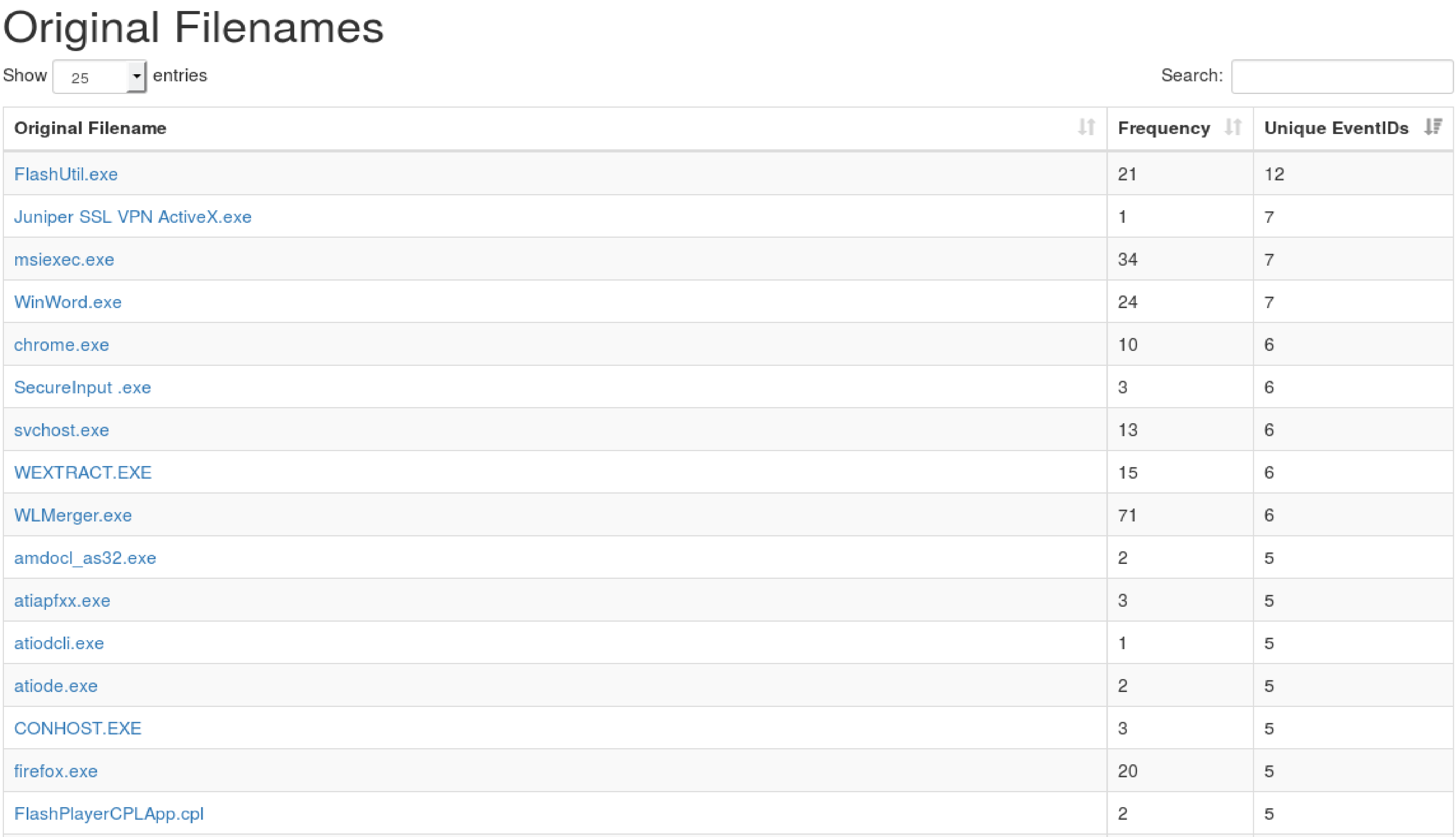
Step back in time:
I know what you did last summer

Samples and indicators, sorted and tagged

Clustering of samples

Adding a web interface

https://github.com/MISP/misp-workbench

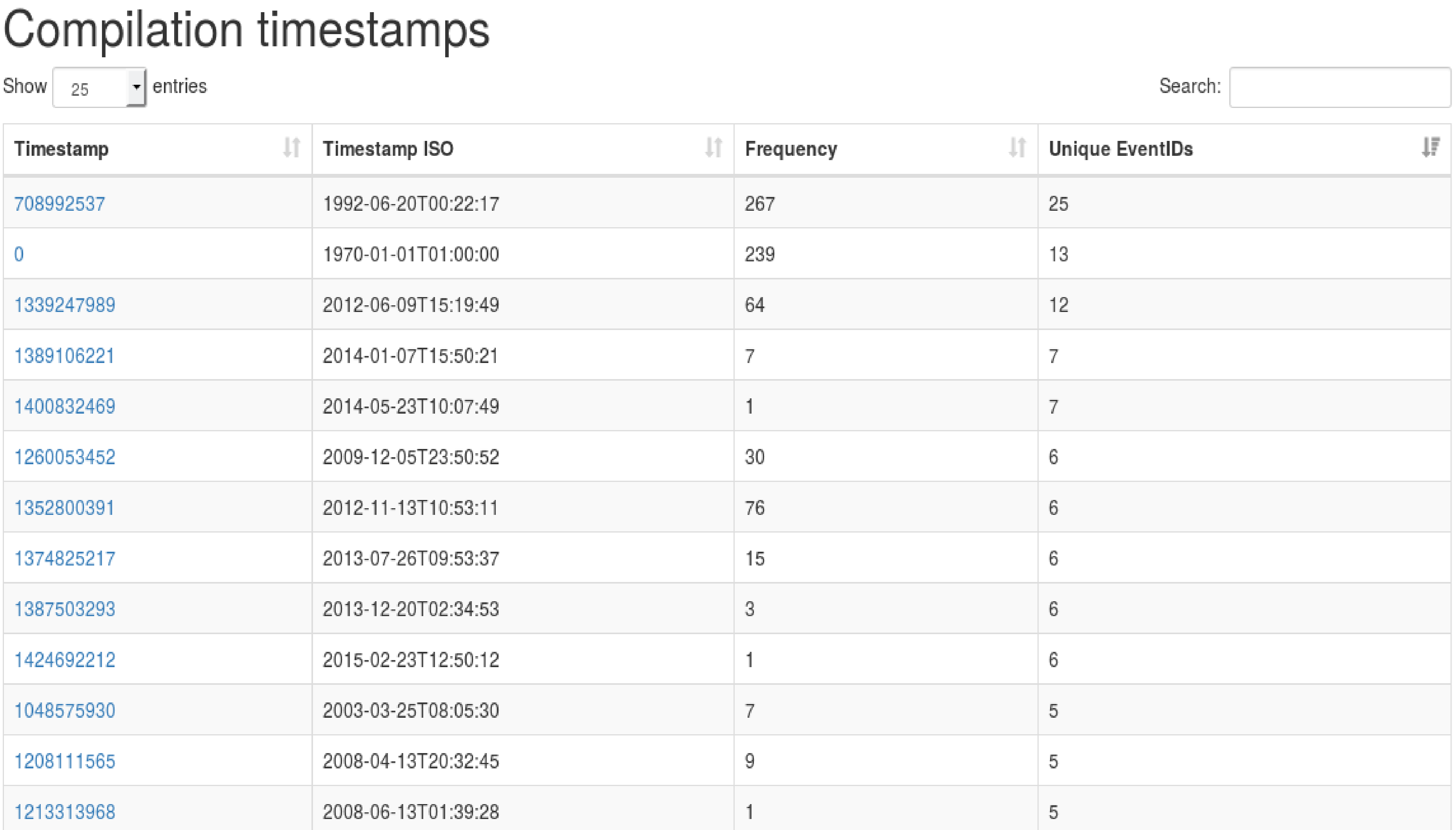


Original Filenames

Show 25 entries

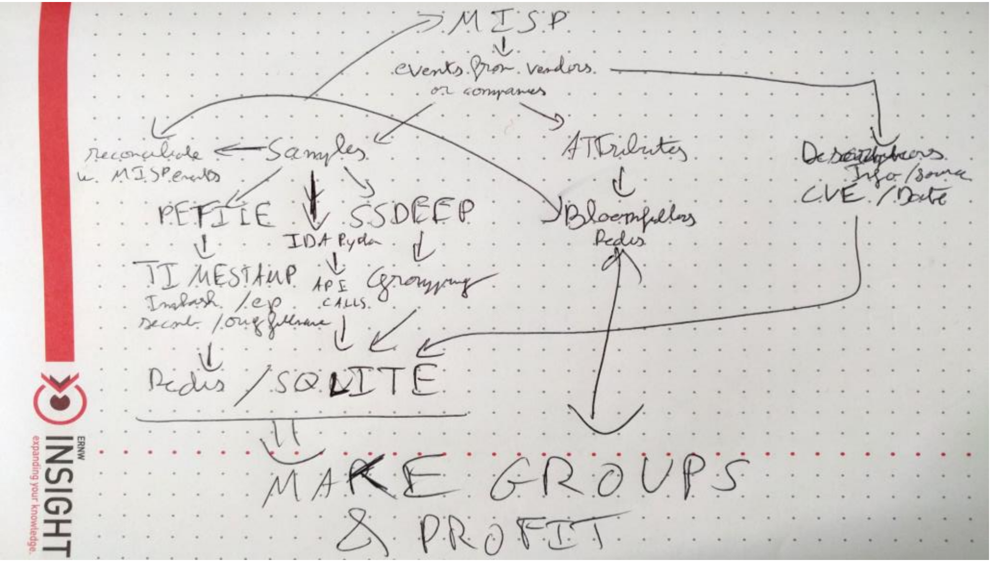
Search:

Original Filename	Frequency	Unique EventIDs
FlashUtil.exe	21	12
Juniper SSL VPN ActiveX.exe	1	7
msiexec.exe	34	7
WinWord.exe	24	7
chrome.exe	10	6
SecureInput .exe	3	6
svchost.exe	13	6
WEXTRACT.EXE	15	6
WLMerger.exe	71	6
amdocl_as32.exe	2	5
atiapfxx.exe	3	5
atiodcli.exe	1	5
atiode.exe	2	5
CONHOST.EXE	3	5
firefox.exe	20	5
FlashPlayerCPLApp.cpl	2	5



Compilation timestamps

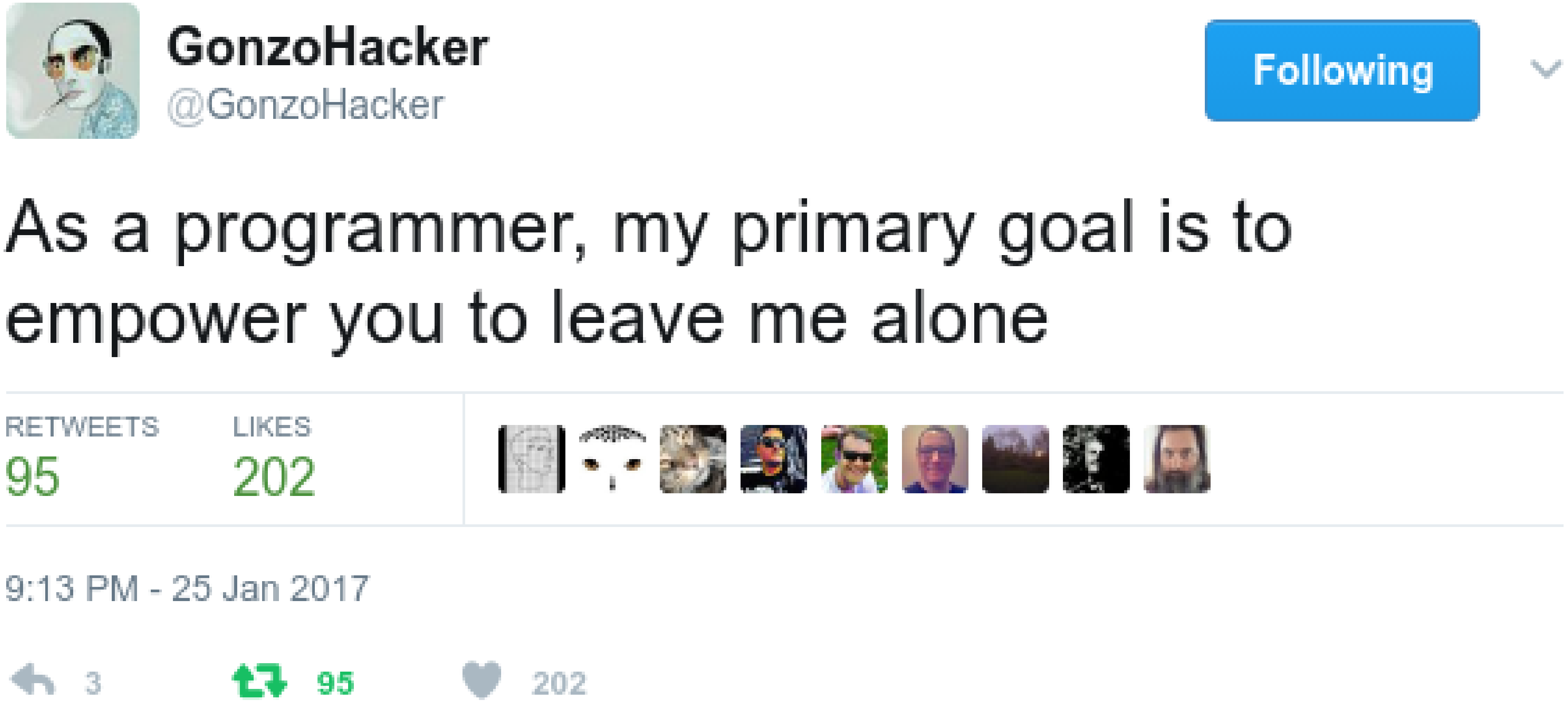
Timestamp	Timestamp ISO	Frequency	Unique EventIDs
708992537	1992-06-20T00:22:17	267	25
0	1970-01-01T01:00:00	239	13
1339247989	2012-06-09T15:19:49	64	12
1389106221	2014-01-07T15:50:21	7	7
1400832469	2014-05-23T10:07:49	1	7
1260053452	2009-12-05T23:50:52	30	6
1352800391	2012-11-13T10:53:11	76	6
1374825217	2013-07-26T09:53:37	15	6
1387503293	2013-12-20T02:34:53	3	6
1424692212	2015-02-23T12:50:12	1	6
1048575930	2003-03-25T08:05:30	7	5
1208111565	2008-04-13T20:32:45	9	5
1213313968	2008-06-13T01:39:28	1	5



So… Workbench.

The obstacles:

- Have root on your MISP server of choice
- Run 5 scripts in the right order to have a standalone interface
- Understand my trail of thought, because open source, yay
- And anyway, works on my machine



As a programmer, my primary goal is to empower you to leave me alone

How do we integrate <new feature> in MISP?

Which solutions exist?

Which of them are actually useable?

Can we base our implementation on an existing standard?

Is that standard sane??

MISP Threat Sharing

Requirements

1. Objects to group indicators as one entity
2. Feasible way to extract the indicators from binaries & graphs
3. Organise, store & display everything
4. Means for object interconnection & correlation
5. Flexibility & scalability & buzzwordbuzzword

MASTERPLAN

Object definition which can be plugged into MISP

PE & graph feature extraction

Mapping of features to object definition

Generate a JSON file in MISP Object format

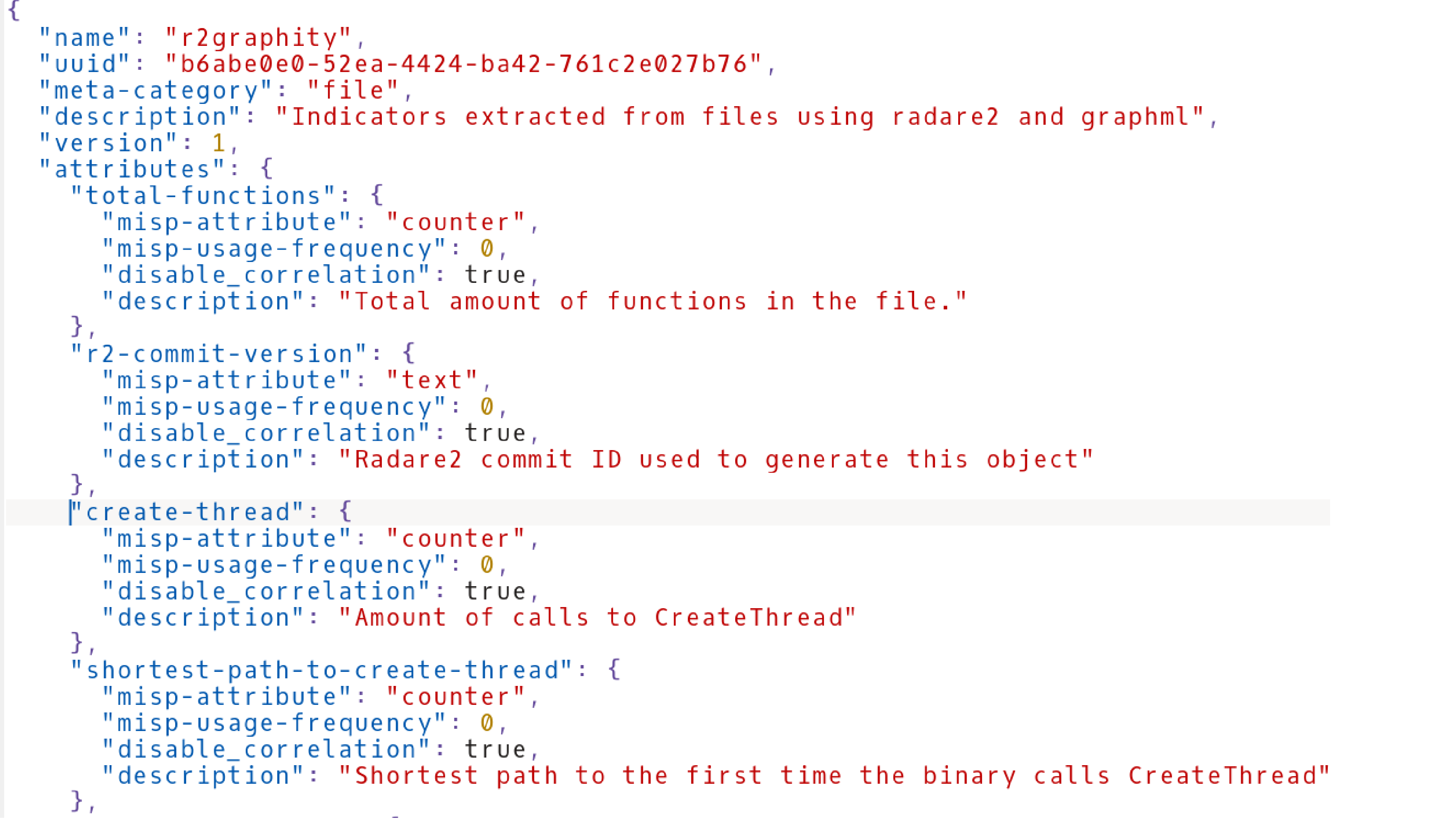
Implementation of objects in MISP core

Objects for other file formats

Integration of the feature generator in the STL

Soon-ish: string search, automatic correlation on per-instance basis

Later-ish: behaviour gadget search, straight from the graphs



```
{
  "name": "r2graphity",
  "uuid": "b6abe0e0-52ea-4424-ba42-761c2e027b76",
  "meta-category": "file",
  "description": "Indicators extracted from files using radare2 and graphml",
  "version": 1,
  "attributes": {
    "total-functions": {
      "misp-attribute": "counter",
      "misp-usage-frequency": 0,
      "disable_correlation": true,
      "description": "Total amount of functions in the file."
    },
    "r2-commit-version": {
      "misp-attribute": "text",
      "misp-usage-frequency": 0,
      "disable_correlation": true,
      "description": "Radare2 commit ID used to generate this object"
    },
    "create-thread": {
      "misp-attribute": "counter",
      "misp-usage-frequency": 0,
      "disable_correlation": true,
      "description": "Amount of calls to CreateThread"
    },
    "shortest-path-to-create-thread": {
      "misp-attribute": "counter",
      "misp-usage-frequency": 0,
      "disable_correlation": true,
      "description": "Shortest path to the first time the binary calls CreateThread"
    },
```

Metrics Engineering

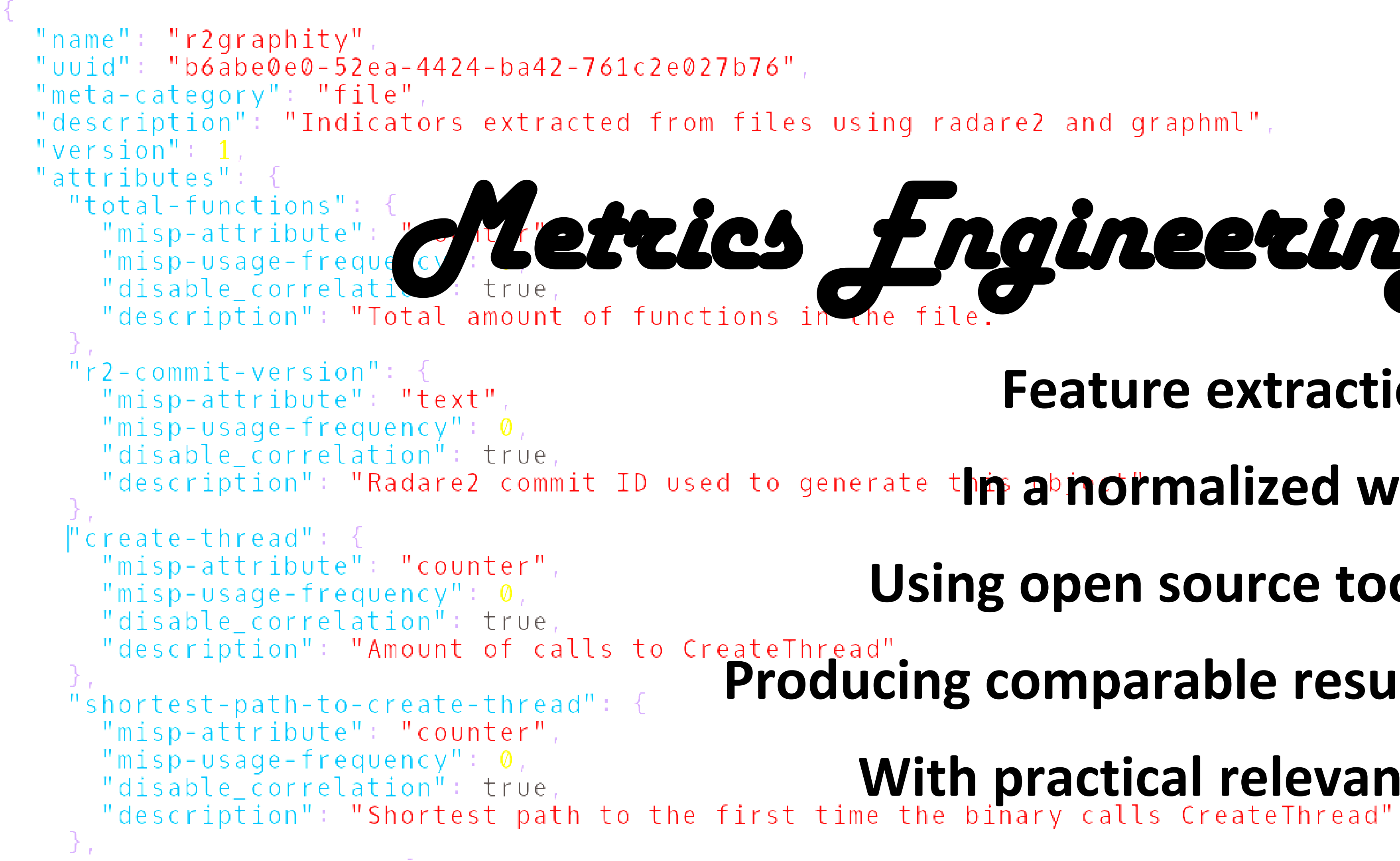
Feature extraction

In a normalized way

Using open source tools

Producing comparable results

With practical relevance

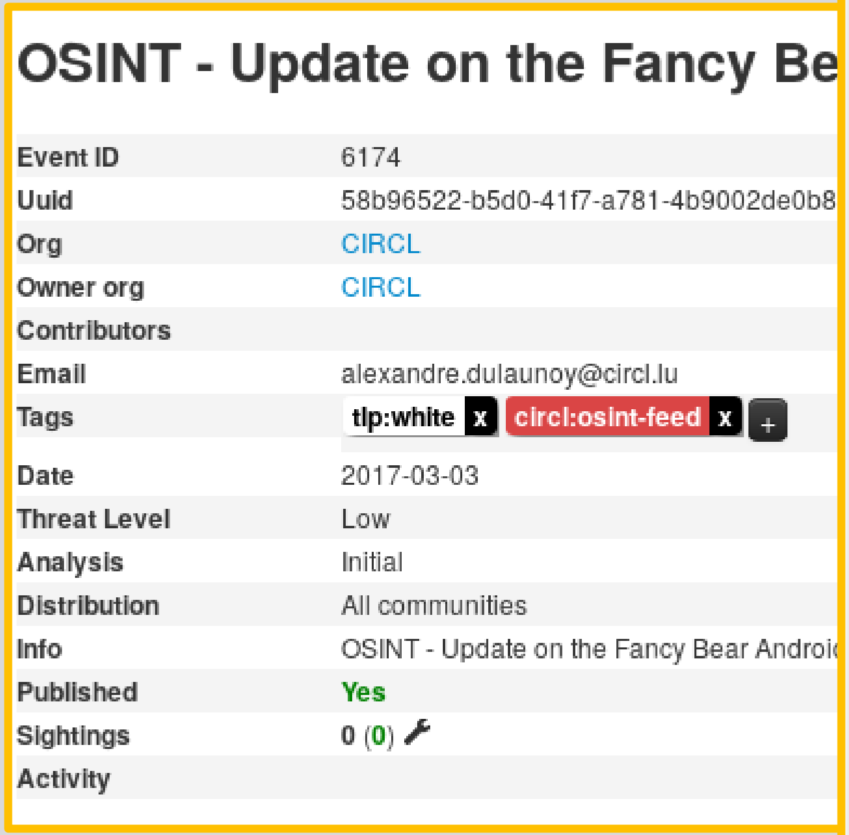


```
{
  "name": "r2graphity",
  "uuid": "b6abe0e0-52ea-4424-ba42-761c2e027b76",
  "meta-category": "file",
  "description": "Indicators extracted from files using radare2 and graphml",
  "version": 1,
  "attributes": {
    "total-functions": {
      "misp-attribute": "counter",
      "misp-usage-frequency": 0,
      "disable_correlation": true,
      "description": "Total amount of functions in the file."
    },
    "r2-commit-version": {
      "misp-attribute": "text",
      "misp-usage-frequency": 0,
      "disable_correlation": true,
      "description": "Radare2 commit ID used to generate this object"
    },
    "create-thread": {
      "misp-attribute": "counter",
      "misp-usage-frequency": 0,
      "disable_correlation": true,
      "description": "Amount of calls to CreateThread"
    },
    "shortest-path-to-create-thread": {
      "misp-attribute": "counter",
      "misp-usage-frequency": 0,
      "disable_correlation": true,
      "description": "Shortest path to the first time the binary calls CreateThread"
    },
```

Chicken & Egg Problem

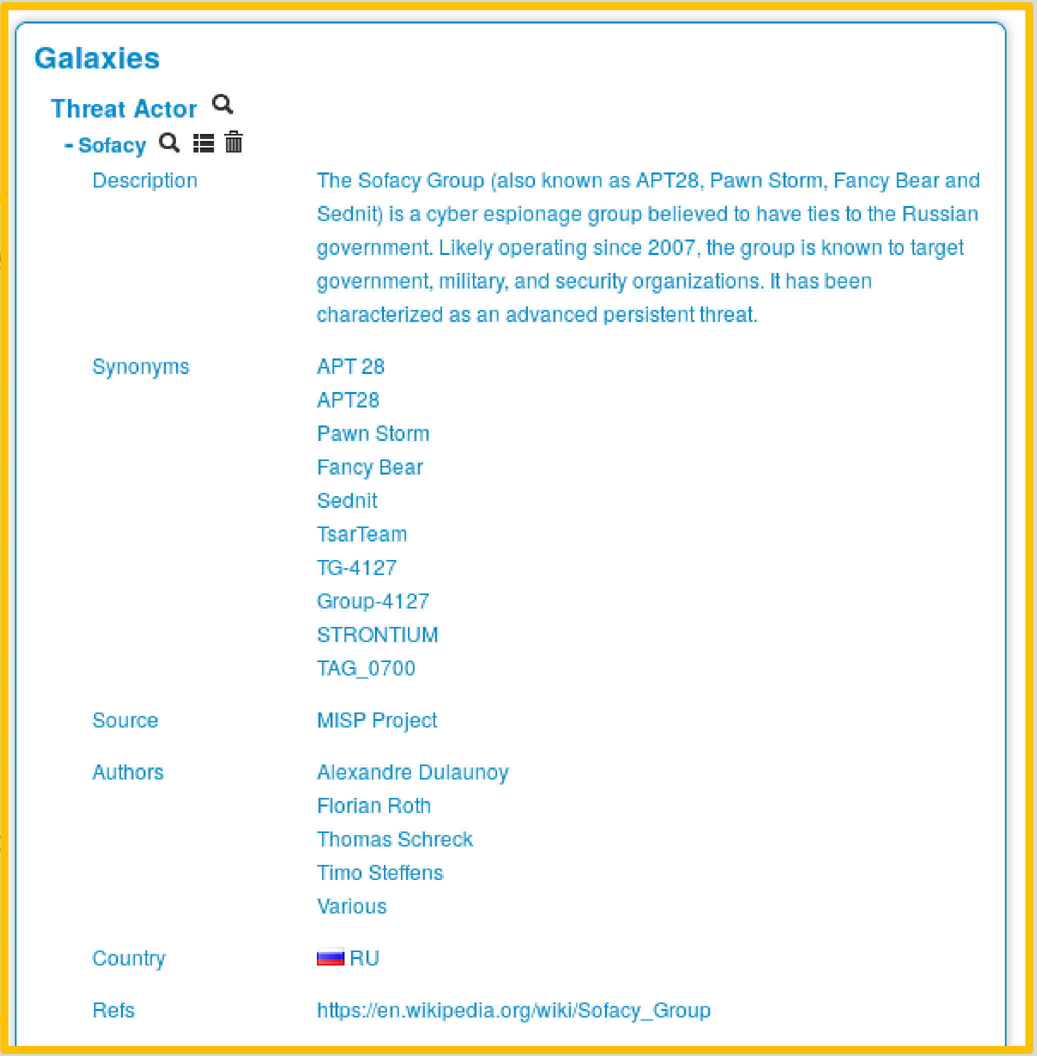
1. You can't identify good indicators if they aren't stored, accessible, and easy to generate

2. It doesn't make sense to rely on indicators if every other research project creates new ones



OSINT - Update on the Fancy Be

Event ID	6174
Uuid	58b96522-b5d0-41f7-a781-4b9002de0b8
Org	CIRCL
Owner org	CIRCL
Contributors	
Email	alexandre.dulaunoy@circl.lu
Tags	tlp:white x circl:osint-feed x +
Date	2017-03-03
Threat Level	Low
Analysis	Initial
Distribution	All communities
Info	OSINT - Update on the Fancy Bear Androi
Published	Yes
Sightings	0 (0)
Activity	

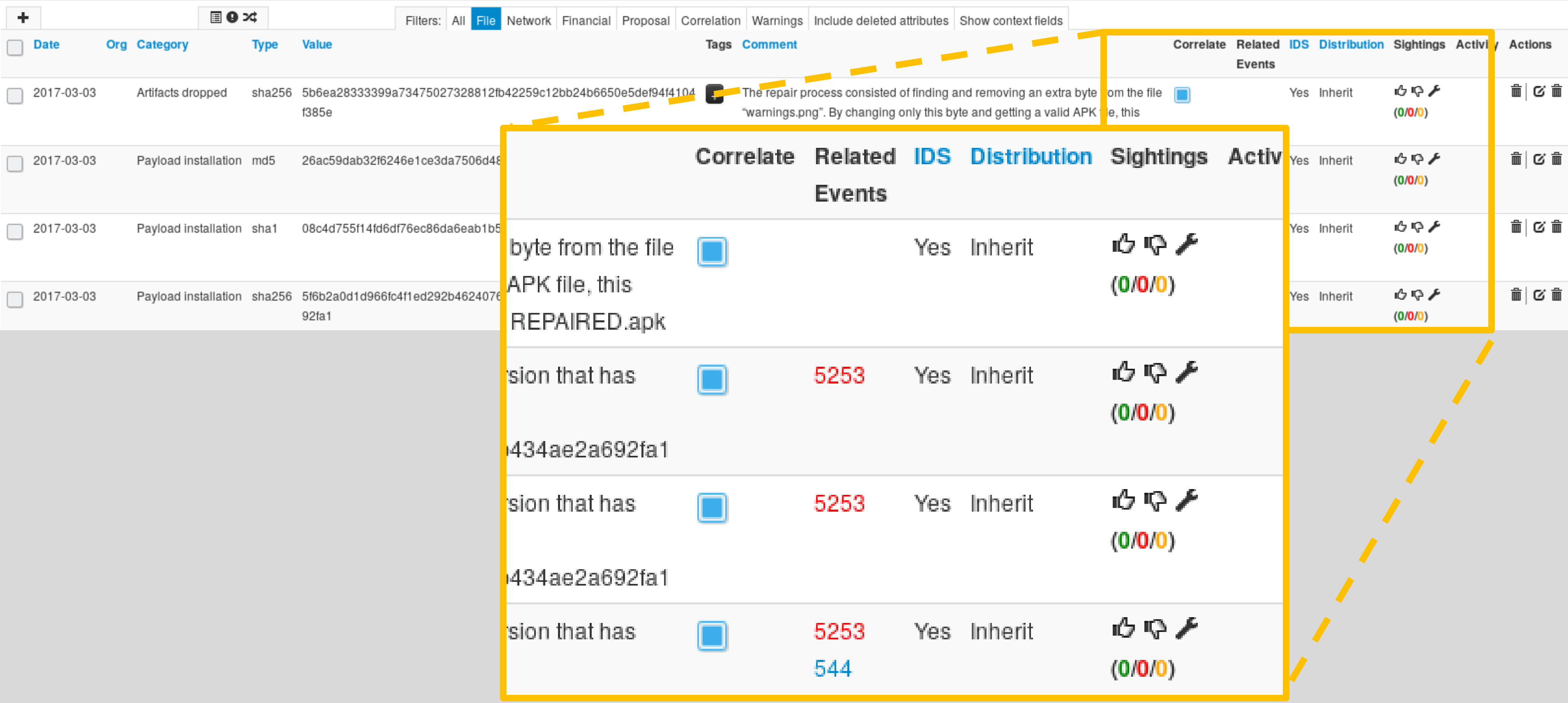


Galaxies

Threat Actor

- Sofacy

Description	The Sofacy Group (also known as APT28, Pawn Storm, Fancy Bear and Sednit) is a cyber espionage group believed to have ties to the Russian government. Likely operating since 2007, the group is known to target government, military, and security organizations. It has been characterized as an advanced persistent threat.
Synonyms	APT 28 APT28 Pawn Storm Fancy Bear Sednit TsarTeam TG-4127 Group-4127 STRONTIUM TAG_0700
Source	MISP Project
Authors	Alexandre Dulaunoy Florian Roth Thomas Schreck Timo Steffens Various
Country	RU
Refs	https://en.wikipedia.org/wiki/Sofacy_Group


+
Filters: All File Network Financial Proposal Correlation Warnings Include deleted attributes Show context fields

□	Date	Org	Category	Type	Value	Tags	Comment	Correlate	Related Events	IDS	Distribution	Sightings	Activity	Actions
□	2017-03-03		Artifacts dropped	sha256	5b6ea28333399a73475027328812fb42259c12bb24b6650e5def94f4104 f385e	+	The repair process consisted of finding and removing an extra byte from the file "warnings.png". By changing only this byte and getting a valid APK file, this	☑		Yes	Inherit	(0/0/0)		
□	2017-03-03		Payload installation	md5	26ac59dab32f6246e1ce3da7506d48					Yes	Inherit	(0/0/0)		
□	2017-03-03		Payload installation	sha1	08c4d755f14fd6df76ec86da6eab1b5					Yes	Inherit	(0/0/0)		
□	2017-03-03		Payload installation	sha256	5f6b2a0d1d966fc4f1ed292b4624076 92fa1					Yes	Inherit	(0/0/0)		

	Correlate	Related Events	IDS	Distribution	Sightings	Activ
byte from the file APK file, this REPAIRED.apk	☑		Yes	Inherit	(0/0/0)	
sion that has 434ae2a692fa1	☑	5253	Yes	Inherit	(0/0/0)	
sion that has 434ae2a692fa1	☑	5253	Yes	Inherit	(0/0/0)	
sion that has	☑	5253 544	Yes	Inherit	(0/0/0)	

Wrappn' it up

Graphs

Tons of metrics

MISP objects

Exchange platform & infrastructure

Feature Marxism

All the features
- by default,
- on all samples,
- shared with everyone,
- constantly, integrated, automatic

Historical data

De-facto standards

Implicit feedback loops

Thank you!

Marion Marschalek
@pinkflawd

Raphaël Vinot
@rafi0t

CIRCL
Computer Incident
Response Center
Luxembourg

http://karmadecay.com/