
simpliFiRE.IDAscope

An IDA Pro extension for easier (malware) reverse engineering

Daniel Plohmann, Alexander Hanel

plohmann@cs.uni-bonn.de
alexander.hanel@gmail.com



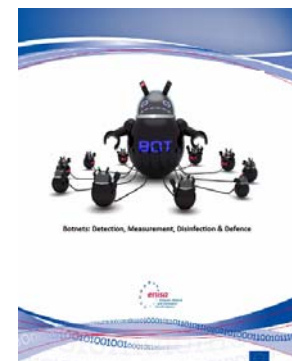
Some words about myself

■ Personal background

- PhD student and researcher at University of Bonn & Fraunhofer FKIE
 - Research focus: Reverse Engineering
 - Work focus: malware analysis and botnet mitigation

■ Projects

- Author of 2011 ENISA Botnet Study [1]
- PyBox [2]
 - Userland-hooking framework (with Felix Leder)
- AntiRE [3]
 - An Executable Collection of Anti-Reversing Techniques



[1] <http://www.enisa.europa.eu/act/res/botnets/botnets-measurement-detection-disinfection-and-defence>

[2] <http://code.google.com/p/pyboxed> [3] https://bitbucket.org/fkie_cd_dare/simplifire.antire

simpliFiRE.IDAscope

Current State

IDAscope

... in a nutshell

- An IDA Pro extension for easier (malware) reverse engineering.
- Motivated by the current workflow of working with IDA Pro.
 - Repeat: „Identify relevant parts of the binary; tear apart; document findings.“
- Common tasks:

1

- Malware RE usually starts with the corner pieces: strings, API calls, signature hits, ...
 - API calls are a good indicator for function semantics.

2

- Reoccurring need for looking up things in MSDN.
 - Switch windows time and time again...

3

- C&C communication schemes are of high interest!
 - Find and understand cryptographic routines used.

Idea:

- Provide automation/integration of „helpers“ that assist with regularly performed tasks.



Hex-Rays Home > Plug-In Contest

Hex-Rays

Plug-In Contest 2012: Hall Of Fame

Contests 2012 2011 2010 2009

This year the plugin contest gathered five contestants. But as you know, there can only be ~~one~~, well, two winners!

Based on the plugin's functionality, robustness, usefulness, ease of use and documentation, we declare the following winners:

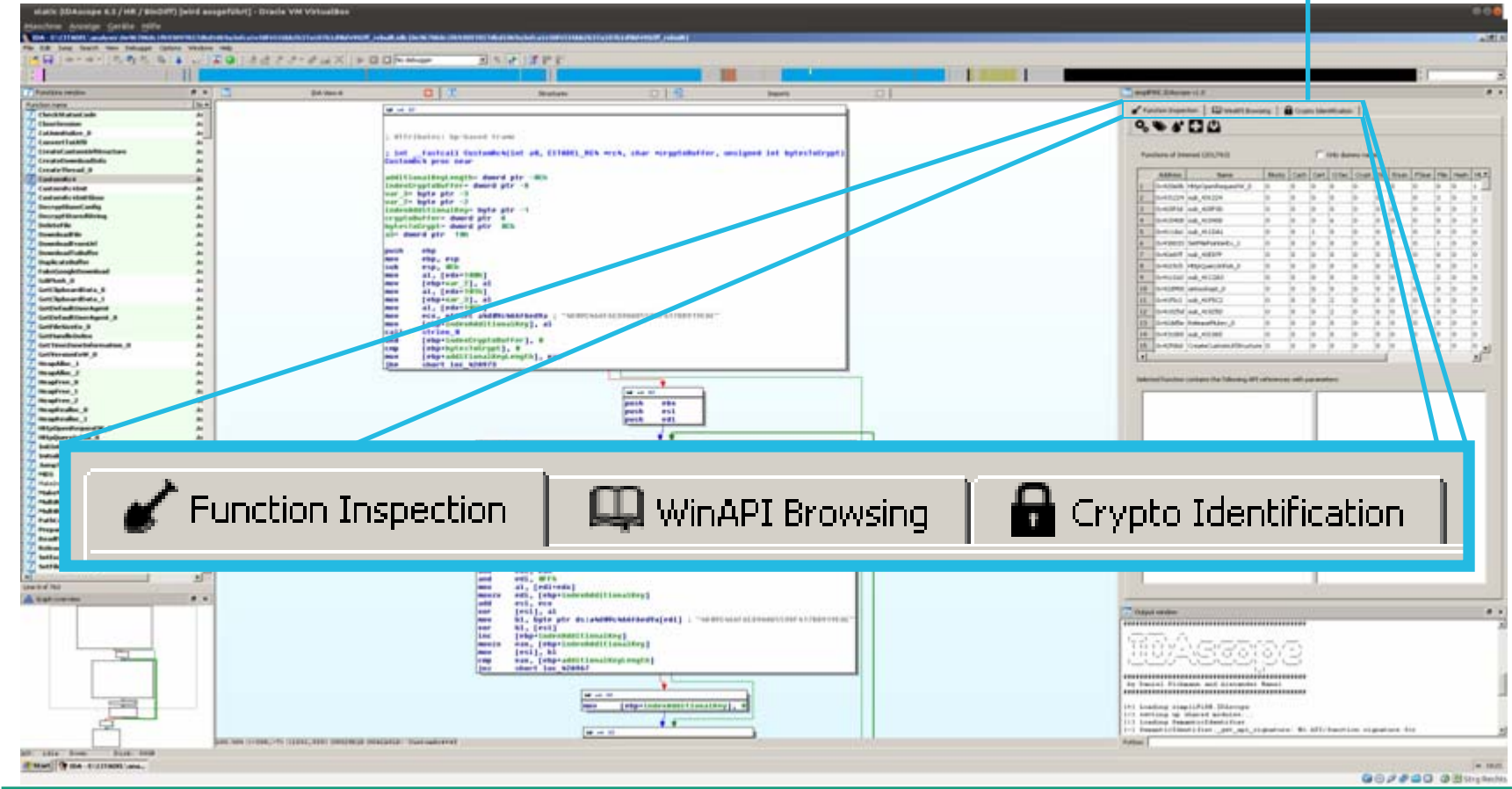
1. [Aaron Portnoy, of Exodus Intelligence, with the IDA Toolbag plugin](#)
2. [Daniel Plohmann, of the Fraunhofer FKIE, with the IDAscope plugin](#)

Congratulations to both! We are pleased with the improved plugin quality and complexity.

Below is the list of all submissions in no particular order. All contest entries are interesting and useful:

IDAscope Overview

- Functionality organized in tabs
- Main window can be dragged around like every other IDA view.



IDAscope: Features

1) Function Inspection



■ Tagging of functions

- Based on API calls
- APIs can be specified via config
- Renaming with tags possible

■ Example

- **DownloadToFile** consists of API calls tagged with File and Network

Functions of Interest (201/763) Only dummy names

	Address	Name	Blocks	Cach	Cert	CrSec	Crypt	Dir	Enum	F5ear	File	Hash	Http	Info	Mod	Mutx	Pipe	Proc	Reg	Url	Virt	WINet	Ws2	
1	0x4209f0	sub_4209f0	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	0
2	0x40dd3e	Thread_MakeInetRe...	0	0	0	0	0	0	0	0	0	0	15	0	0	0	0	0	0	0	0	0	7	0
3	0x42c024	sub_42c024	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0
4	0x41a29e	sub_41a29e	0	0	1	0	0	0	0	0	1	0	9	0	0	0	0	0	0	0	0	0	5	3
5	0x40e0bf	MakeInetRequest	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	4	0
6	0x420541	sub_420541	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0
7	0x42bf6b	InitInternetSession	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0
8	0x42bfd7	CloseSession	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0
9	0x42c56b	sub_42c56b	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0
10	0x40cf0a	sub_40cf0a	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0
11	0x42c515	sub_42c515	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0
12	0x421227	sub_421227	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
13	0x41682e	sub_41682e	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
14	0x42c07f	CheckStatusCode	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	1	0
15	0x421281	sub_421281	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
16	0x42f6bd	CreateCustomUrStr...	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
17	0x42c2c1	DownloadFromUrl	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
18	0x40d77a	sub_40d77a	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
19	0x42110d	sub_42110d	0	0	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
20	0x42c14d	DownloadToBuffer	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
21	0x42117a	sub_42117a	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
22	0x4211a8	sub_4211a8	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
23	0x42c1ed	DownloadToFile	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	1	0
24	0x42e200	sub_42e200	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0
25	0x41a204	sub_41a204	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
26	0x40fc08	sub_40fc08	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0
27	0x420a0b	HttpCreateDefaultW...	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0

Selected function contains the following API references with parameters:

Address	API	Tag	
1	0x42c20d	CreateFileW	File
2	0x42c257	InternetReadFile	WINet
3	0x42c292	FlushFileBuffers	File
4	0x42c272	WriteFile	File

Address	Type	Name	Value	
1	0x42c266	LPOVERLAPPED	lpOverlapped	0x3
2	0x42c26e	LPCVOID	lpBuffer	0xffffffff0
3	0x42c26b	DWORD	nNumberOfBytesToWrite	0xffffffff8
4	0x42c26a	LPDWORD	lpNumberOfBytesWritten	0x0
5	0x42c271	HANDLE	hFile	0x7

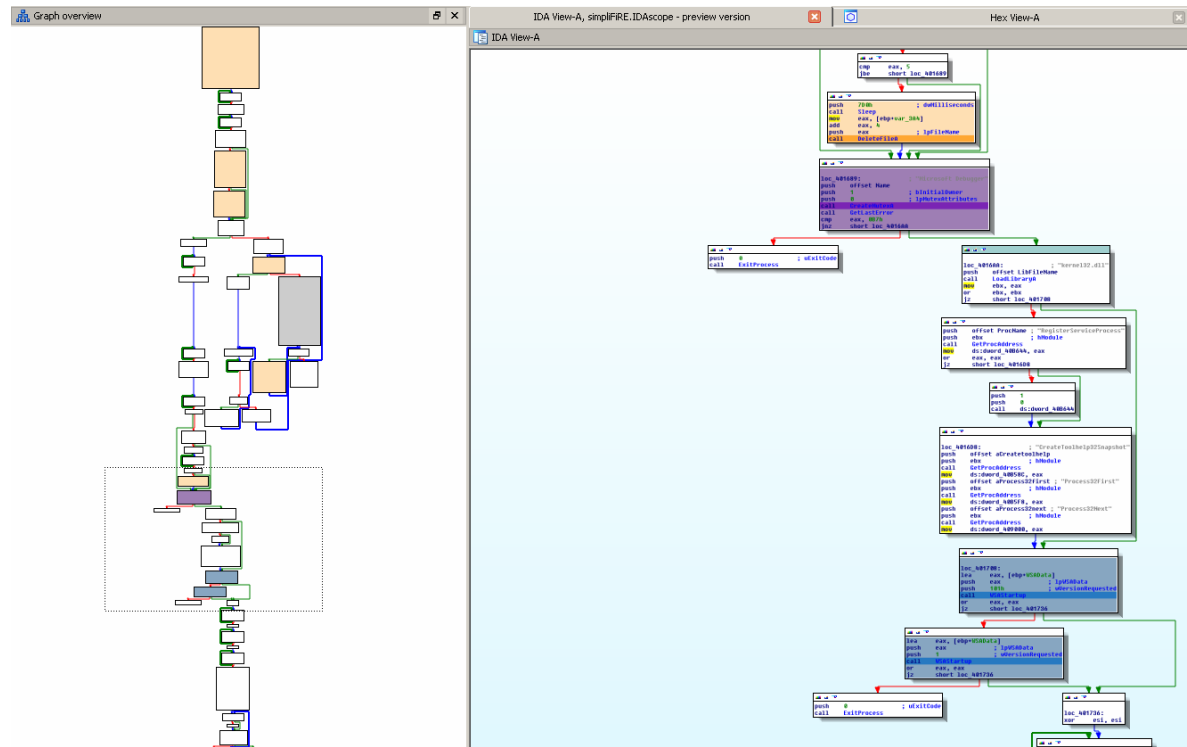
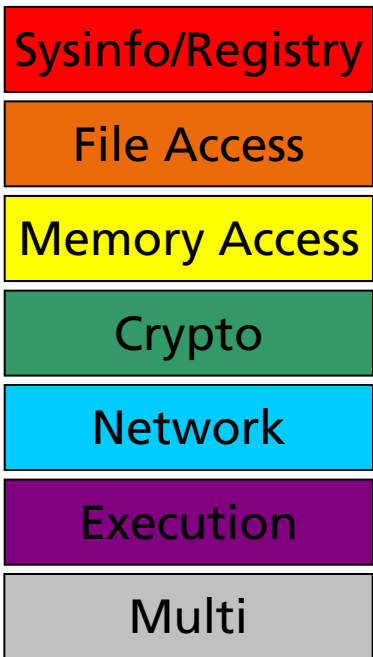
	Address	API	Tag
1	0x42c20d	CreateFileW	File
2	0x42c257	InternetReadFile	WINet
3	0x42c292	FlushFileBuffers	File
4	0x42c272	WriteFile	File

IDAscope: Features

1) Function Inspection



- Coloring of basic blocks
 - Based on API semantics
 - Colors can be adjusted
- More an experiment :)



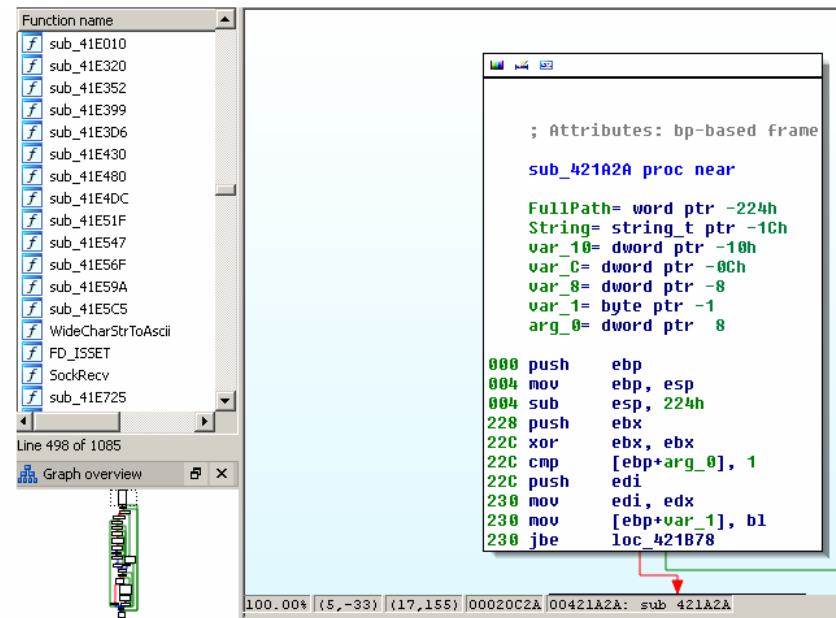
IDAscope: Features

1) Function Inspection



- Code to function conversion
 - Function prologues get handled first
 - Then remaining undefined areas
 - Opens these code sections to further analysis

```
:00421A2A      push    ebp
:00421A2A      mov     ebp, esp
:00421A2B      sub     esp, 224h
:00421A2D      push    ebx
:00421A33      xor     ebx, ebx
:00421A36      cmp     dword ptr [ebp+8], 1
:00421A3A      push    edi
:00421A3B      mov     edi, edx
:00421A3D      mov     [ebp-1], b1
:00421A40      jbe    loc_421B78
:00421A46      mov     eax, [edi+4]
:00421A49      cmp     [eax], bx
:00421A4C      jz     loc_421B78
:00421A52      push   esi
:00421A53      push   eax
:00421A54      call   ds:PathIsURLW
:00421A5A      cmp     eax, 1
:00421A5D      jnz    loc_421AF4
:00421A63      push   dword ptr [edi+4]
:00421A66      lea    esi, [ebp-1Ch]
:00421A69      or     eax, 0FFFFFFFh
:00421A6C      call   CreateStringStruct
```



Function name

- sub_41E010
- sub_41E320
- sub_41E352
- sub_41E399
- sub_41E3D6
- sub_41E430
- sub_41E480
- sub_41E4DC
- sub_41E51F
- sub_41E547
- sub_41E56F
- sub_41E59A
- sub_41E5C5
- WideCharStrToAscii
- FD_ISSET
- SocketRecv
- sub_41E725

Line 498 of 1085

Graph overview

```
; Attributes: bp-based frame
sub_421A2A proc near
FullPath= word ptr -224h
String= string_t ptr -1Ch
var_10= dword ptr -10h
var_C= dword ptr -0Ch
var_8= dword ptr -8
var_1= byte ptr -1
arg_0= dword ptr 8
000 push    ebp
004 mov     ebp, esp
004 sub     esp, 224h
228 push    ebx
22C xor     ebx, ebx
22C cmp     [ebp+arg_0], 1
22C push   edi
230 mov     edi, edx
230 mov     [ebp+var_1], b1
230 jbe    loc_421B78
```

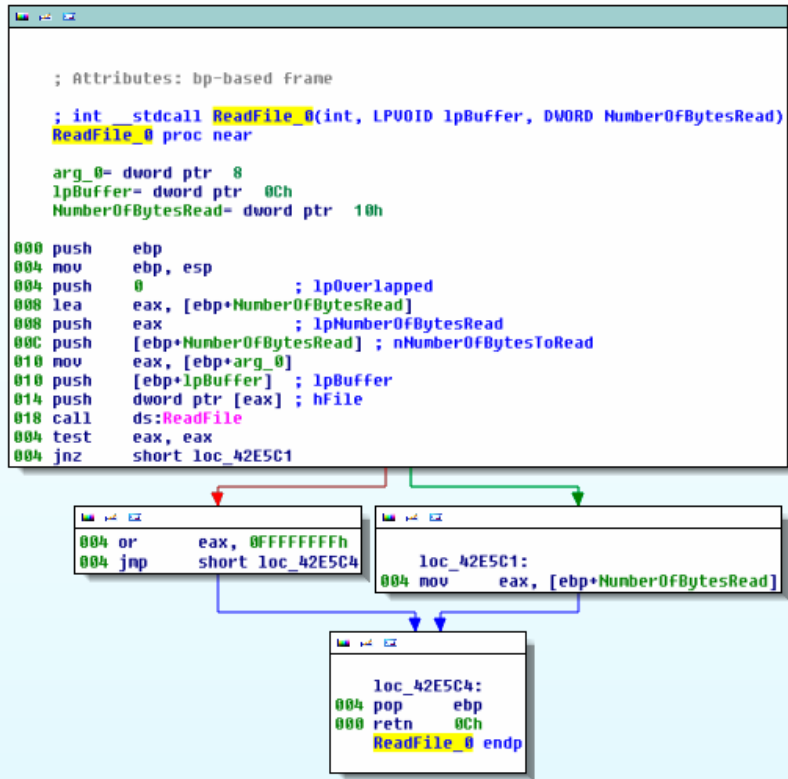
100.00% | (5,-33) | (17,155) | 00020C2A | 00421A2A: sub_421A2A

IDAscope: Features

1) Function Inspection



- Automatic renaming of wrapper functions
 - Credits go to **Branko Spasojevic** (author of Optimice) for providing the code!



```
Identified and renamed potential wrapper @ [0040931a] to [memcpy_0]
Identified and renamed potential wrapper @ [00409595] to [InitializeCriticalSection_0]
Identified and renamed potential wrapper @ [00409c31] to [memcpy_0]
Identified and renamed potential wrapper @ [00409c67] to [memcpy_1]
Identified and renamed potential wrapper @ [00410f50] to [CreateFileMappingW_0]
Identified and renamed potential wrapper @ [0041ca4d] to [VirtualQueryEx_0]
Identified and renamed potential wrapper @ [0041da77] to [CoInitializeEx_0]
Identified and renamed potential wrapper @ [0041daa8] to [CoUninitialize_0]
Identified and renamed potential wrapper @ [0041dae3] to [CoCreateInstance_0]
Identified and renamed potential wrapper @ [00420e1c] to [StrCmpNIA_0]
Identified and renamed potential wrapper @ [00422b7b] to [GdiFlush_0]
Identified and renamed potential wrapper @ [00422b7b] to [HeapAlloc_0]
Identified and renamed potential wrapper @ [0042282dc] to [LoadLibraryW_0]
Identified and renamed potential wrapper @ [00422blbc] to [PathMatchSpecW_0]
Identified and renamed potential wrapper @ [0042b535] to [CreateEventW_0]
Identified and renamed potential wrapper @ [00422bf1] to [WaitForSingleObject_0]
Identified and renamed potential wrapper @ [0042d8ae] to [__imp_memset_0]
Identified and renamed potential wrapper @ [0042e53d] to [SetFilePointerEx_0]
Identified and renamed potential wrapper @ [0042e57f] to [SetFilePointerEx_1]
Identified and renamed potential wrapper @ [0042e59e] to [ReadFile_0]
Identified and renamed potential wrapper @ [00436fe0] to [SetUnhandledExceptionFilter_0]
```

IDAscope: Features

2) WinAPI Browsing

■ Seamless integration of MSDN in IDA Pro

- accessible via shortcut on highlighted elements
- Now also with online lookup!
- But not multi-threaded / no backgrounded lookups yet

The screenshot displays the IDA Pro interface with the WinAPI Browsing window open. The main window shows assembly code for a function, with a call to `InternetReadFile` highlighted. The WinAPI Browsing window shows the details for the `InternetReadFile` function, including its syntax, parameters, return value, and remarks. The syntax is: `BOOL InternetReadFile(_In_ HANDLE hFile, _Out_ LPVOID lpBuffer, _In_ DWORD dwNumberOfBytesToRead, _Out_ LPDWORD lpNumberOfBytesRead)`. The parameters are: `hFile` (HANDLE), `lpBuffer` (LPVOID), `dwNumberOfBytesToRead` (DWORD), and `lpNumberOfBytesRead` (LPDWORD). The return value is `BOOL`. The remarks describe the function's behavior and its use in retrieving data from a handle.

IDAscope: Features

3) Crypto Identification

■ Identification of cryptographic / compression routines

- Based on ratio of arithmetic / logic instructions to all instructions in a basic block
- Approach described in „Dispatcher: Enabling Active Botnet Infiltration using Automatic Protocol Reverse-Engineering“ by Juan Caballero et al.

The screenshot displays the IDA Pro interface. On the left, the assembly code for the `DecryptStoredString` function is shown, including instructions like `push esi`, `mov edx, [eax+4]`, `xor dx, [eax+2]`, and `jmp short loc_4180C7`. The right pane shows the `Crypto Identification` settings, including the `Arithmetic/Logic Heuristic` section with sliders for `ArithLog Rating` (30.00), `Basic Blocks size` (8), and `Allowed calls` (0). A table below the settings lists 14 blocks from a total of 9750 blocks, with columns for `Address`, `Name`, `Block Address`, `# Instr`, and `Arithmetic/Logic Rating`. The table includes entries for `DecryptStoredString`, `sub_42D135`, `sub_42C622`, `sub_42AABC`, `sub_42AC65`, `sub_42A77E`, `sub_42A7C2`, `sub_42AD1F`, `sub_42AACD`, `sub_423376`, `_SomeRandomFunction`, `sub_418BD0`, and `sub_42DA35`.

Address	Name	Block Address	# Instr	Arithmetic/Logic Rating
0x418b9a	DecryptStoredString	0x418baf	9	33.33
0x42d135	sub_42d135	0x42d20e	8	37.50
0x42c622	sub_42c622	0x42c65a	26	42.31
0x42aabc	sub_42aabc	0x42aaf9	31	45.16
0x42ac65	sub_42ac65	0x42b109	128	64.06
0x42a77e	sub_42a77e	0x42a7c2	9	55.56
0x42ac65	sub_42ac65	0x42ad1f	127	64.57
0x42aacb	sub_42aacb	0x42aacd	15	53.33
0x423336	sub_423336	0x423376	8	50.00
0x42a708	_SomeRandomFunction	0x42a729	11	54.55
0x42a708	_SomeRandomFunction	0x42a5e6	14	57.14
0x42a708	_SomeRandomFunction	0x42a599	14	57.14
0x418bd0	sub_418bd0	0x418bbe	12	33.33
0x42da35	sub_42da35	0x42da7f	8	50.00

Found Crypto Signatures

- [-] CRC 32 Generator
- [-] AES Inverse box
- [-] AES Forward box

IDAscope: Features

3) Crypto Identification

■ Identification of cryptographic / compression routines

- Based on ratio of arithmetic / logic instructions to all instructions in a basic block
- Approach described in „Dispatcher: Enabling Active Botnet Infiltration using Automatic Protocol Reverse-Engineering“ by Juan Caballero et al.

Arithmetic/Logic Heuristic

ArithLog Rating: Exclude Zeroing

Basic Blocks size: Looped Blocks only

Allowed calls: Group by Functions

```
loc_418BAF:
mov     edx, [eax+4]
movzx  esi, cx
mov     dl, [edx+esi]
xor     dl, [eax]
xor     dl, cl
inc     ecx
mov     [esi+edi], dl
cmp     cx, [eax+2]
jnb    short loc_418BAF
```

Example: Citadel string decryption.

- 1) 3 ArithLogInstructions / 9 Instructions = 33% rating
 - 2) 9 instructions
 - 3) 0 calls
 - 4) Is a looped basic block
- => Matches above parameters

5	0x42ac65	sub_42AC65	0x42b109	128	64.06
6	0x42a77e	sub_42A77E	0x42a7c2	9	55.56
7	0x42ac65	sub_42AC65	0x42ad1f	127	64.57
8	0x42aabc	sub_42AABC	0x42aacd	15	53.33
9	0x423336	sub_423336	0x423376	8	50.00

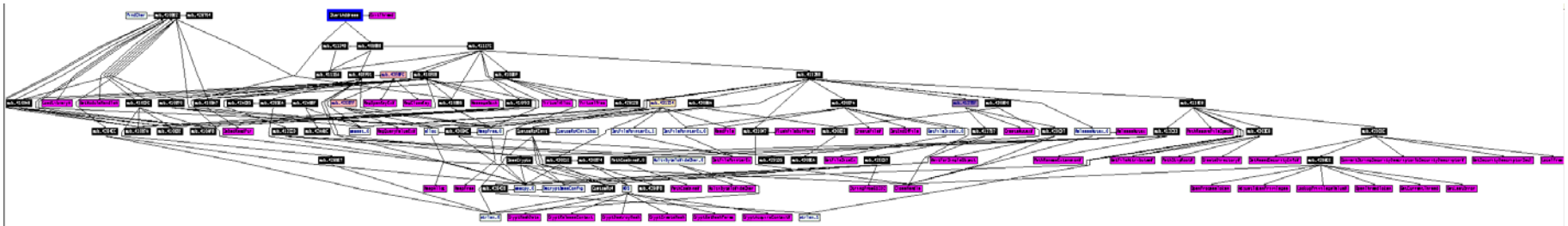
simpliFiRE.IDAscope

Future Plans

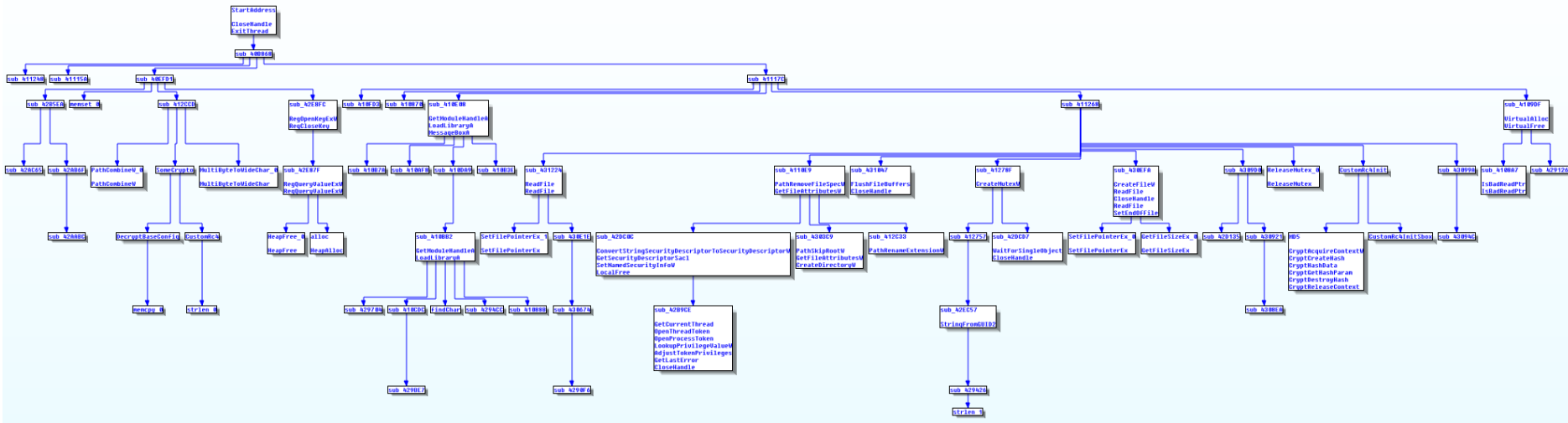
IDAscope: Future Plans

4) Threads / Function Relationship

- Threads and function call chains are a good indicator of functionality
 - A „big picture“ would be very helpful.
 - My opinion: We need something better than this (WinGraph) or step by step navigation via xrefs.



- Same function scope as IDA graph (IDAPython API has limited graph support), not much better...

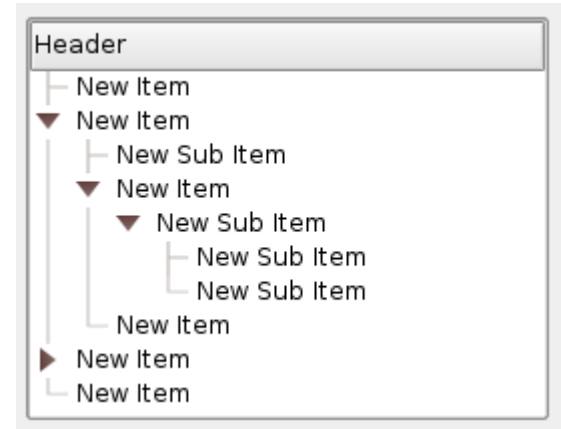
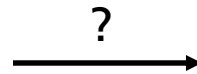


IDAscope: Future Plans

4) Threads / Function Relationship

- Threads and function call chains are a good indicator of functionality
 - Same displayed as tree, generated with Alex' script [4]

```
CreateThread Call 0x40bc39      sub_41278F
StartAddress (IpStartAddr)    sub_412757
sub_40B868                    sub_42EC57
sub_40EFD1                    sub_429426
memset_0                      strlen_1
sub_412CCD                    * Call StringFromGUID2
SomeCrypto                   * Call CreateMutexW
  DecryptBaseConfig          sub_42DCD7
  memcpy_0                   * Call WaitForSingleObject
  CustomRc4                  * Call CloseHandle
  strlen_0                   sub_4110E9
MultiByteToWideChar_0        sub_412C33
  * Call MultiByteToWideChar * Call PathRenameExtensionW
PathCombineW_0                * Call PathRemoveFileSpecW
  * Call PathCombineW        sub_4303C9
sub_42E8FC                    * Call PathSkipRootW
  * Call RegOpenKeyExW       * Call GetFileAttributesW
sub_42E87F                    * Call CreateDirectoryW
  * Call RegQueryValueExW   sub_42DC0C
  alloc                      sub_42B9CE
  * Call HeapAlloc          * Call GetCurrentThread
  * Call RegQueryValueExW   * Call OpenThreadToken
  HeapFree_0                 * Call OpenProcessToken
  * Call HeapFree           * Call LookupPrivilegeValueW
  * Call RegCloseKey        * Call AdjustTokenPrivileges
sub_42B5EA                    * Call GetLastError
sub_42AB6F                    * Call CloseHandle
  sub_42AABC                 * Call ConvertStringSecurityDescriptorToSecurityDescriptorW
  sub_42AC65                 * Call GetSecurityDescriptorSacl
sub_41115A                    * Call SetNamedSecurityInfoW
sub_41117C                    * Call LocalFree
sub_411268                    * Call GetFileAttributesW
```



Use a TreeWidget for rendering?

[4] <http://hooked-on-mnemonics.blogspot.com/2012/08/ida-thread-analysis-sript.html>

IDAscope

Conclusion

- Start using it! :)
 - Repository at
 - <http://idascope.pnx.tf>
(points to: https://bitbucket.org/daniel_plohmann/simplifire.idascope)
 - I report about updates
 - in my blog: <http://blog.pnx.tf>
 - on twitter [@push_pnx](https://twitter.com/push_pnx)
 - Alex has a blog, too: <http://hooked-on-mnemonics.blogspot.com>
- Send **feedback** or **ideas** for improvement!
 - idascope@pnx.tf