CRACK.ME UP!

AN INTRODUCTION TO

André Baptista

@0xACB

ØxOPOSEC Meetup

Reverse Engineering

- Uncovering the hidden behaviour of a given technology, system, program, protocol or device, by analysing the structure and operation of its components

- Extracting knowledge about any unknown engineering invention

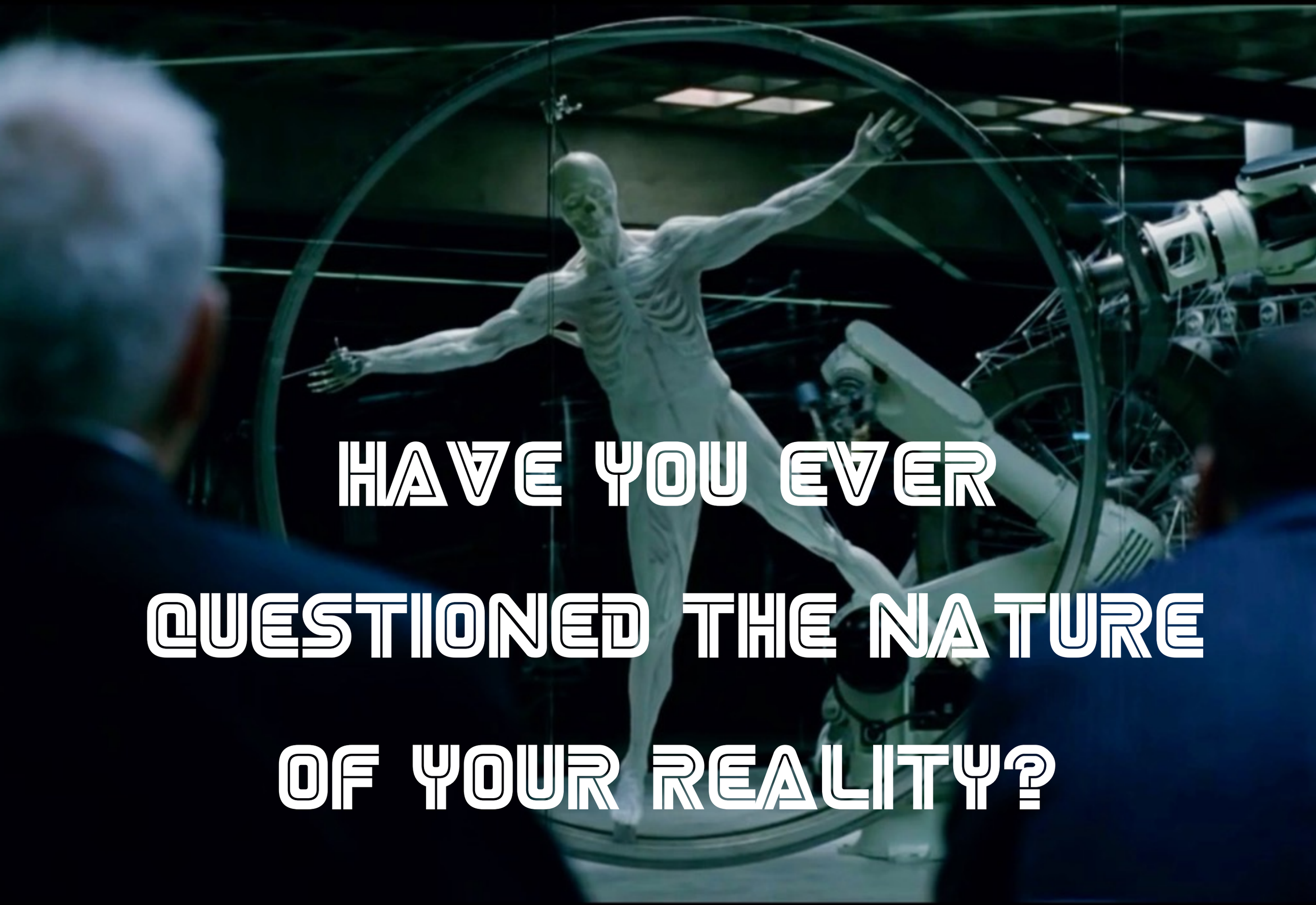
History

ØxOPOSEC Meetup

Reverse Engineering

History

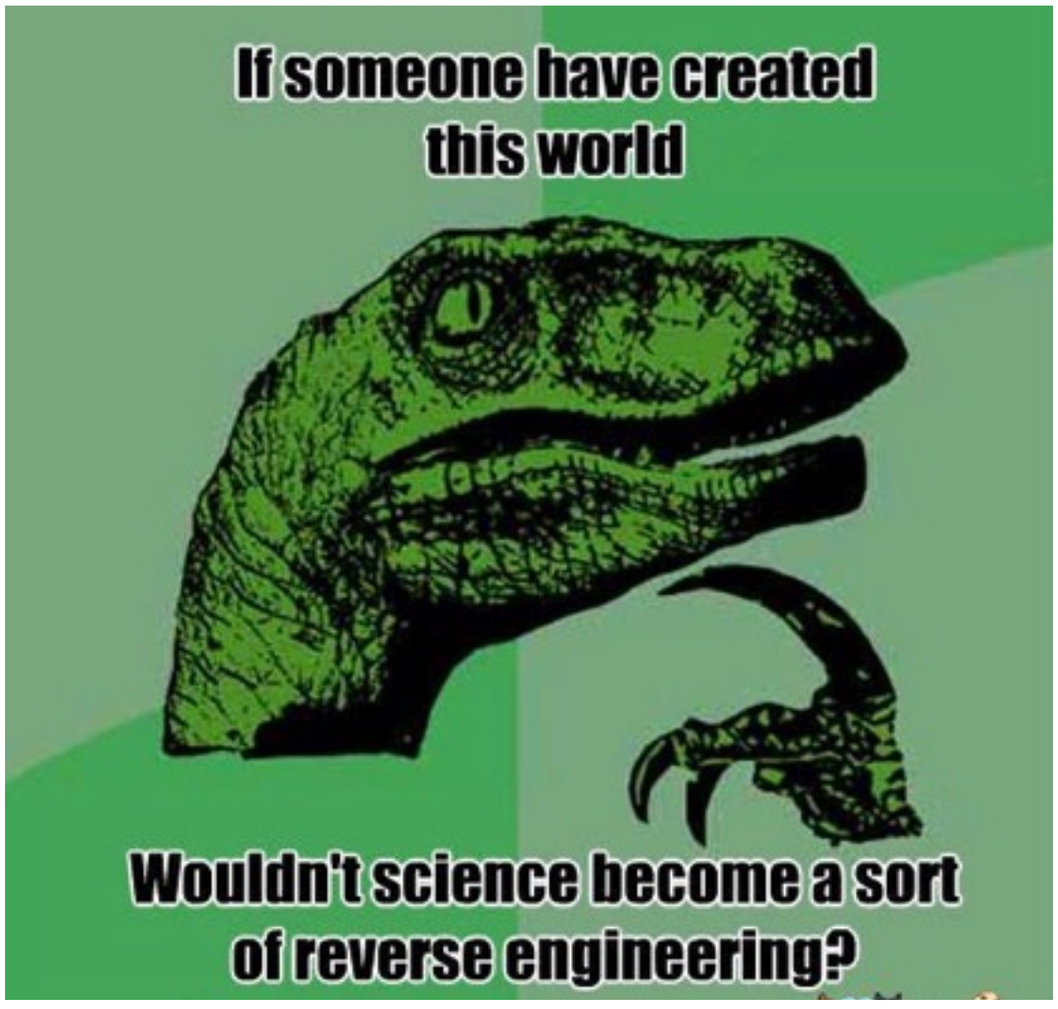
- RE was used to copy inventions made by other countries or business competitors
- Frequently used in the WW2 and Cold war:
  - *Jerry can*
  - *Panzerschreck*


HAVE YOU EVER
QUESTIONED THE NATURE
OF YOUR REALITY?

ØxOPOSEC Meetup

Reverse Engineering

"Is biology reverse engineering?"



Binary Reverse Engineering

- It’s the process of getting knowledge about compiled software, in order to understand how it works and how it was originally implemented.

ØxOPOSEC Meetup

Binary Reverse Engineering

- It's the process of getting knowledge about compiled software in order to understand how it was originally implemented.

WITHOUT THE SOURCE CODE

Binary Reverse Engineering

Motivation

- Software and hardware cracking
- Malware analysis - botnet clients, spyware, ransomware
- Finding bugs in compiled software
- Creating or improving docs
- Interpreting unknown protocols
- Academic purposes
- Industrial or military espionage
- Software interoperability

Who knows how to do this stuff?

- Hackers in general
- Some intelligence agencies
- Antivirus companies
- Students and curious people

ØxOPOSEC Meetup

Binary Reverse Engineering

Formats of compiled software

- **ELF** (Linux & UNIX like)
- **Mach-O** (OSX)
- **PE** (Windows)
- **Class** (Java bytecode)
- **DEX** (Android - Dalvik bytecode)
- **PYC** (Python bytecode)
- …


REQUIRED SKILLS

Required skills

- **Debugging** (GDB, Valgrind, WinDbg, OllyDbg)
- **Assembly** (x86, x64, ARM, MIPS and many others)
- **Programming** (C, C++, Java, Python, Ruby, etc)
- **Software architecture**
- **Logic, math, crypto, protocols, networks,**
- **Don’t giving up**

ØxOPOSEC Meetup

Awesome tools

- **Disassemblers**
- **Debuggers**
- **Decompilers**
- **Patchers**

Disassemblers

These programs translate machine code to assembly.

```
            main:
080483e4        push    ebp                                     ; XREF=_start+23
080483e5        mov     ebp, esp
080483e7        sub     esp, 0x18
080483ea        and     esp, 0xfffffff0
080483ed        mov     eax, 0x0
080483f2        add     eax, 0xf
080483f5        add     eax, 0xf
080483f8        shr     eax, 0x4
080483fb        shl     eax, 0x4
080483fe        sub     esp, eax
08048400        mov     dword [ss:esp+0x18+var_18], 0x8048528   ; argument "format" for method j_printf
08048407        call    j_printf
0804840c        mov     dword [ss:esp+0x18+var_18], 0x8048541   ; argument "format" for method j_printf
08048413        call    j_printf
08048418        lea     eax, dword [ss:ebp+var_4]
0804841b        mov     dword [ss:esp+0x18+var_14], eax
0804841f        mov     dword [ss:esp+0x18+var_18], 0x804854c   ; argument "format" for method j_scanf
08048426        call    j_scanf
0804842b        cmp     dword [ss:ebp+var_4], 0x149a
08048432        je      0x8048442
--------------------------------------------------------------------------------------------------
08048434        mov     dword [ss:esp+0x18+var_18], 0x804854f   ; argument "format" for method j_printf
0804843b        call    j_printf
08048440        jmp     0x804844e
--------------------------------------------------------------------------------------------------
08048442        mov     dword [ss:esp+0x18+var_18], 0x8048562   ; "Password OK :)\\n", argument "format"
08048449        call    j_printf
--------------------------------------------------------------------------------------------------
0804844e        mov     eax, 0x0                                ; XREF=main+92
08048453        leave
08048454        ret
                ; endp
08048455        nop
```

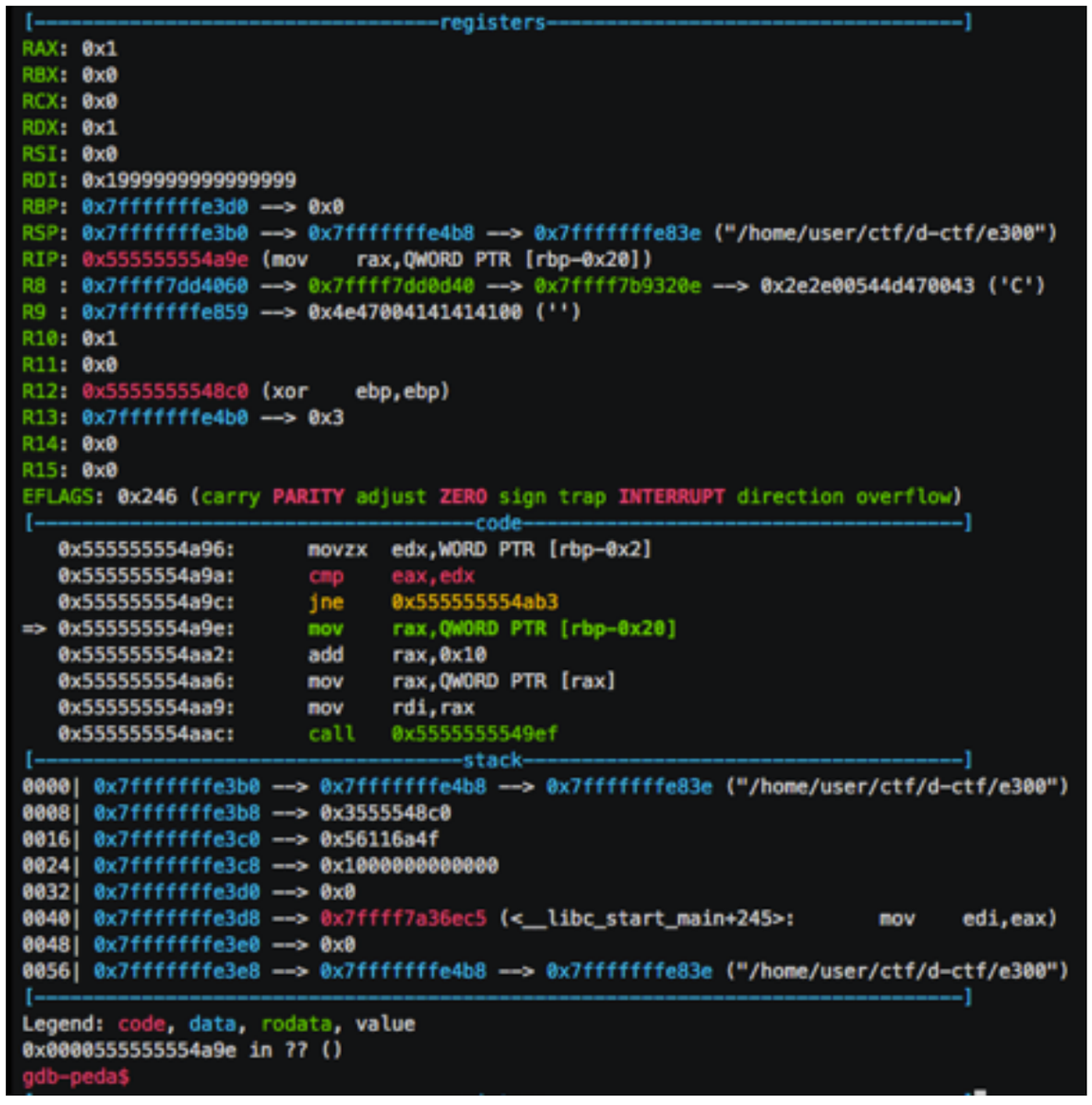
ØxOPOSEC Meetup

Debuggers

These programs are used to test other programs.

Debuggers allow us to inspect memory and CPU registers, modify of variables in runtime, set breakpoints and call functions outside the program flow.

In reverse engineering they are widely used for *dynamic analysis.*



ØxOPOSEC Meetup

Decompilers

These programs try to achieve the *near-impossible* task of translating compiled software to the original source code.

Sometimes, the generated code is enough to perform reversing tasks.

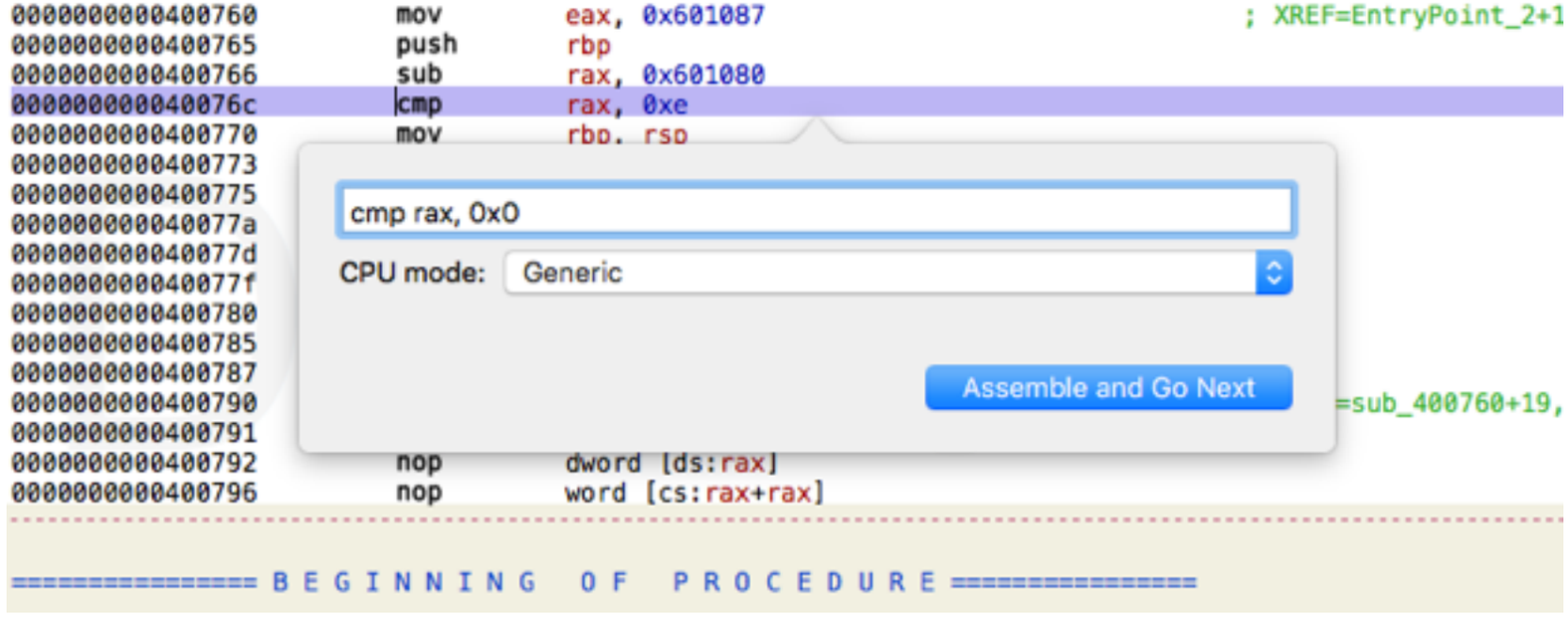

```
signed __int64 __fastcall sub_400962(__int64 a1)
{
 signed int v2; // [sp+14h] [bp-Ch]@6
 int v3; // [sp+18h] [bp-8h]@1
 int i; // [sp+1Ch] [bp-4h]@1

 *(_BYTE *)(strlen((const char *)a1) - 1 + a1) = 0;
 v3 = 1;
 for ( i = 0; i <= dword_600F30; ++i )
 {
  if ( !*(_BYTE *)(i + a1) )
  {
   v3 = i;
   break;
  }
 }
 v2 = 0;
 if ( dword_600F30 - 1 < v3 )
 {
  while ( 1 )
  {
   i *= v3 + 1 / dword_600F30;
   if ( i * i * v3 < i )
    break;
   v2 ^= v3 * i;
  }
 }
 return sub_400616(a1, v3, v2);
}
```

ØxOPOSEC Meetup

Patchers

Patchers can change machine code in order to modify the software behaviour. Hex editors can also be used for patching but there are better tools out there that allow us to patch assembly instructions.



Badass tools

- IDA Pro - https://www.hex-rays.com/products/ida
- Hopper Disassembler - http://www.hopperapp.com
- binary.ninja - https://binary.ninja
- Radare 2 - http://rada.re
- ODA - http://www.onlinedisassembler.com
- OllyDbg - http://www.ollydbg.de
- ILSpy - http://ilspy.net
- Linux tools: objdump, ltrace, strace, readelf, gdb
- Apktool - https://ibotpeaches.github.io/Apktool

Decompilers

- IDA Pro - https://www.hex-rays.com/products/ida (x86, x64, ARM, MIPS, etc)
- Hopper Disassembler - http://www.hopperapp.com (x86, x64, ARM)
- Retargetable Decompiler (AVG) - https://retdec.com (x86, ARM, MIPS, Power PC)
- JADX - https://github.com/skylot/jadx (DEX)
- JetBrains dotPeek - https://www.jetbrains.com/decompiler (.NET)
- ILSpy - http://ilspy.net (.NET)
- uncompyle2 - https://github.com/Mysterie/uncompyle2 (Python bytecode)

ØxOPOSEC Meetup

Static Analysis

- **Do not** execute the program
- Read the spooky assembly/decompiled code
- Inspect flow charts
- Take lots of notes
- Translate procedures to the programming language of your choice
- It’s a pain in the ass to reverse obfuscated of very complex programs

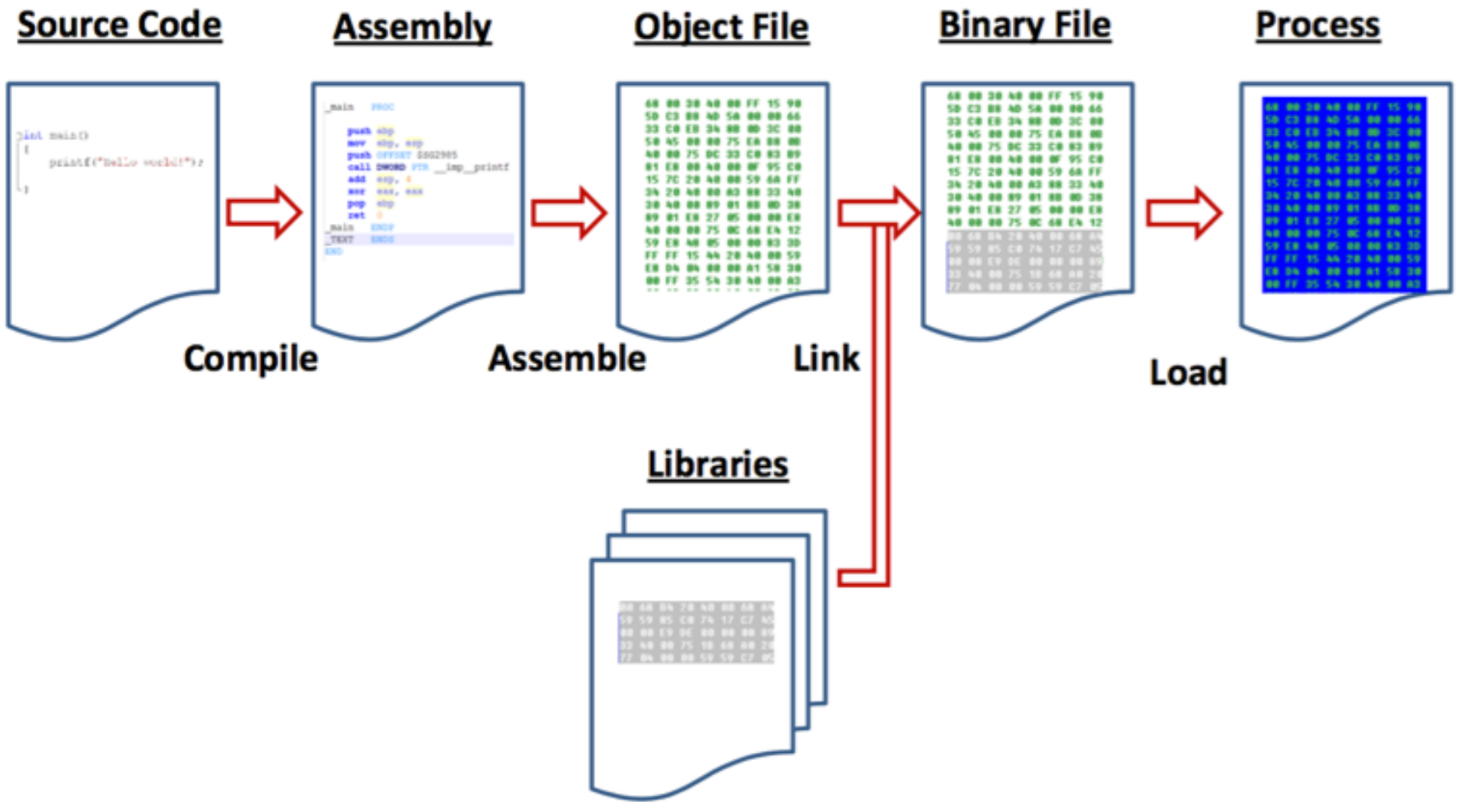
ØxOPOSEC Meetup

Dynamic Analysis

- **Execute** the program
- Inspect the program behaviour
  - Use a debugger to understand the values of the CPU registers, memory (*stack*, *heap*), what values are being returned and the arguments passed in function calls, inspect specific states of execution
- It's difficult to achieve if any anti-debugging protections exist (Some even crash common RE tools on purpose)

ØxOPOSEC Meetup

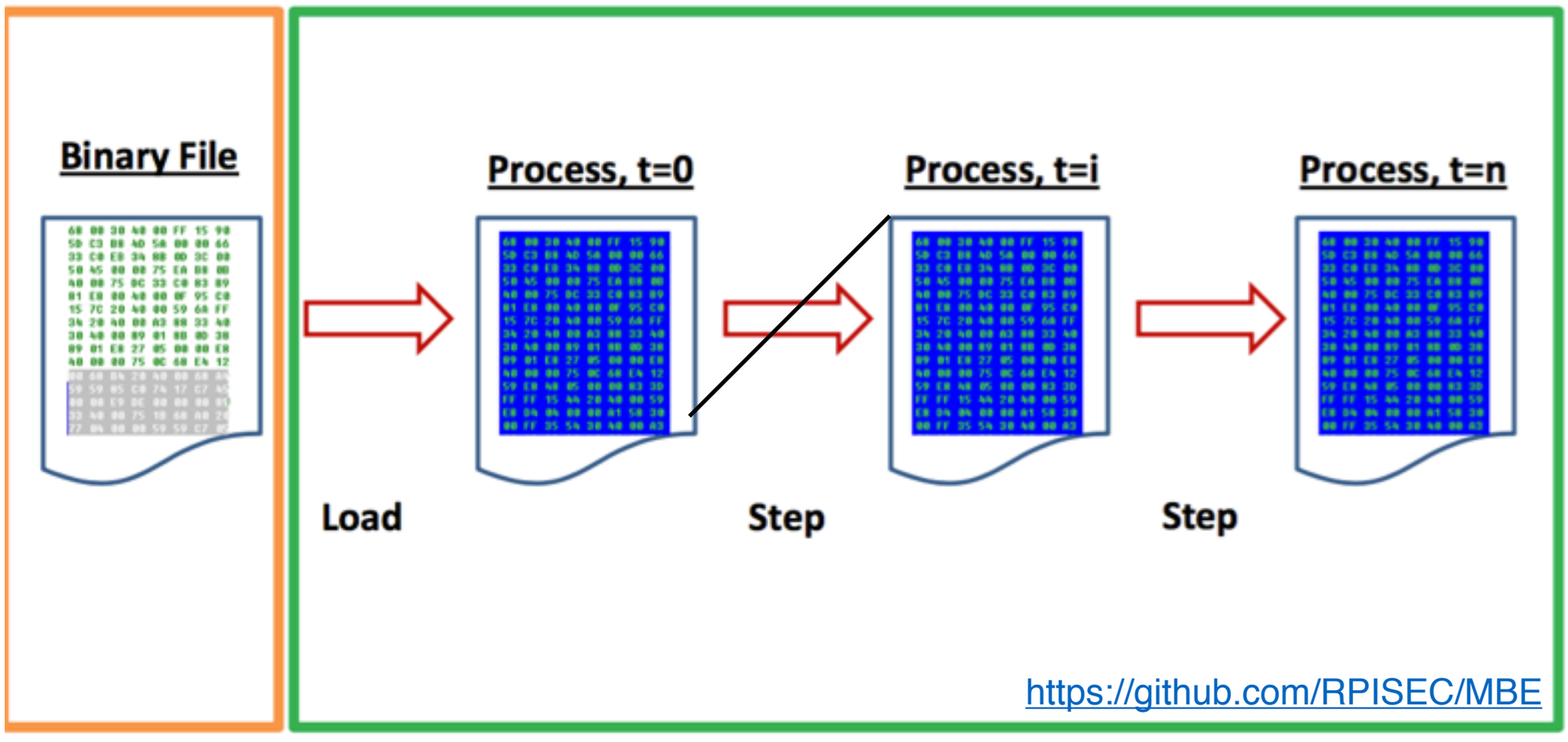
Binary RE



https://github.com/RPISEC/MBE

ØxOPOSEC Meetup

Binary RE



Static

Dynamic

Demo 1 - Static vs Dynamic analysis

https://goo.gl/XXoPCV

Cracking

- This demo was a very simple cracking example
- The real stuff involves much more complex tasks (static analysis, dynamic analysis, concolic analysis, taint analysis)
- E.g. If you want to create a keygen you need to fully understand the serial number validation algorithm
- You need to be a patching ninja to remove anti-debugging protections (usually triggered in runtime)

ØxOPOSEC Meetup

Cracking Games

- In the good old times: to crack a game you just needed to patch code to bypass PC-CDROM identification checks
- Then, virtual drive tools became a thing. But games started to be compiled with additional protections: Anti-debugging, obfuscation and virtual emulators detection (DAEMON Tools/ Generic SafeDisc emulator).

Demo 2 - So… Let’s crack a game for fun (and profit)

Educational purposes only

0xOPOSEC Meetup

CTFs

- Reverse engineering is one of the main categories in Security CTFs

- In CTFs, the contestants are typically challenged to solve cracking problems

- The simplest case is just like the last demo. Find the correct input 

Cracking

Advanced techniques

```
for (i=0;i<10;i++) {
  if (input[i] != password[i]) {
    puts("Wrong!");
    return;
  }
}
puts("Correct!");
```

What's wrong? 🤔

ØxOPOSEC Meetup

Cracking

Advanced techniques

- **Timing attacks**
  - When a char is correct: one more cycle is executed, i.e. more instructions
  - It's possible to launch a timing attack, char by char
  - The attack complexity is reduced from 256^{length} to $256 \times length$
  - **How can we prevent this kind of attacks? Constant time algorithms**
  - Tools for local binary timing attacks: Pin tool, GDB scripts

ØxOPOSEC Meetup

Cracking

Advanced techniques

- **Solvers**

  - Serial number validation algorithms are usually composed by complex verifications, whose components are for e.g. the values of certain indexes of the serial number.

    E.g. $sn[17] == sn[21] \oplus sn[34] - sn[5] \bmod (sn[14] * sn[43]^2)$

  - These verifications can be translated to systems of equations, that can be easily solved by powerful stuff like Z3 Theorem Prover, Sage, Maple, Matlab

  - Z3 supports both arithmetic and bitwise operators, and custom functions as well.

ØxOPOSEC Meetup

Cracking

Advanced techniques - Z3

```python
from z3 import *

s = Solver()

x = BitVec("x", 32)
y = BitVec("y", 32)
z = BitVec("z", 32)
a = BitVec("a", 32)
b = BitVec("b", 32)
c = BitVec("c", 32)

def shiftRight(y, c):
    return y >> c

s.add(x != 0)
s.add(x == a ^ b * z * shiftRight(y, c))

while (s.check() == sat):
    print(s.model())
    s.add(x != s.model()[x], y != s.model()[y])
```

Python script

```
[z = 0,
 b = 1,
 a = 33587201,
 y = 173155,
 x = 33587201,
 c = 17]
[z = 0,
 b = 1,
 a = 34635777,
 y = 173159,
 x = 34635777,
 c = 17]
[z = 0,
 b = 1,
 a = 34766849,
 y = 173158,
 x = 34766849,
 c = 17]
[z = 0,
 b = 1,
 a = 33718273,
 y = 238694,
 x = 33718273,
 c = 17]
```

Solutions

ØxOPOSEC Meetup

What about the future?

Predicting the future using **Naive Mayes**

Let’s get to the powerful stuff


CYBER
HEARTBLEED
CYBER
DARPA
DARPA
splunk

CGC

CYBER

GRAND_CHALLENGE

ØxOPOSEC Meetup

DARPA CGC

- A very important mark in the history of infosec
- It was the first-ever all-machine hacking tournament
- These machines were able to automatically find and patch vulnerabilities in binaries
- The Mechanical Phish project, from the Shellphish team, was able to identify vulnerabilities using both **fuzzing** and **symbolic execution** techniques. It's open source 😎

ØxOPOSEC Meetup

DARPA CGC

Mechanical Phish - Driller

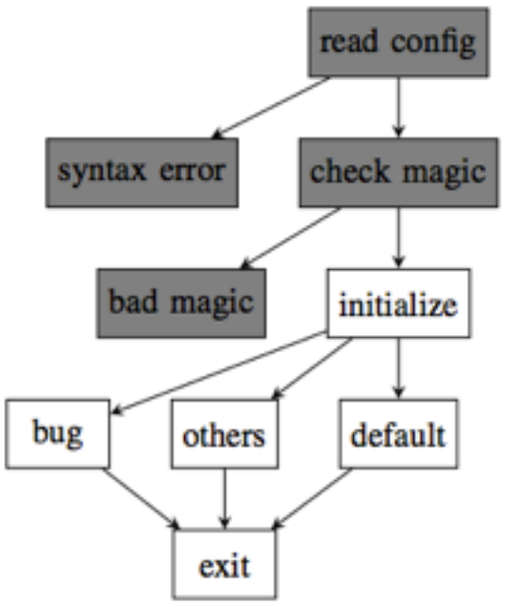


Fig. 1. The nodes initially found by the fuzzer.

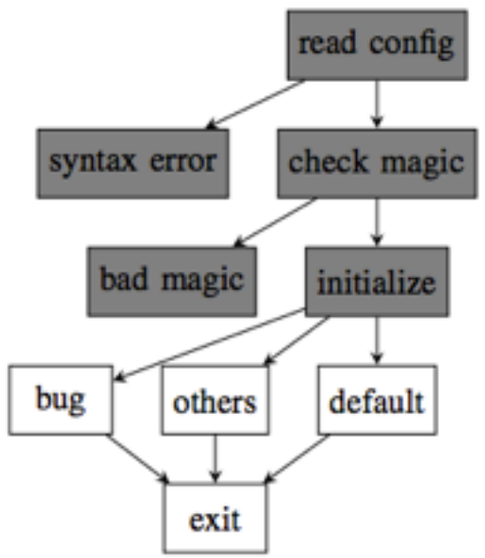


Fig. 2. The nodes found by the first invocation of concolic execution.

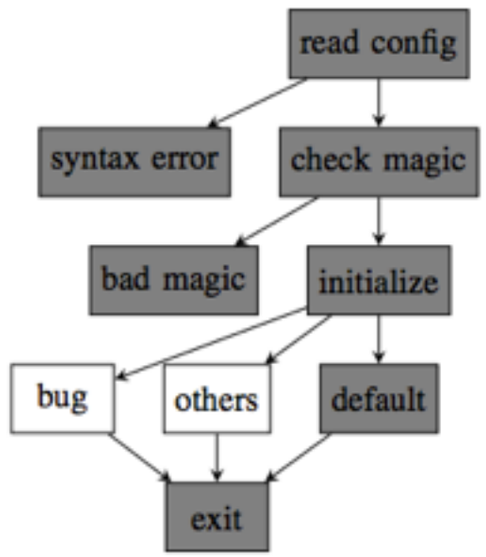


Fig. 3. The nodes found by the fuzzer, supplemented with the result of the first Driller run.

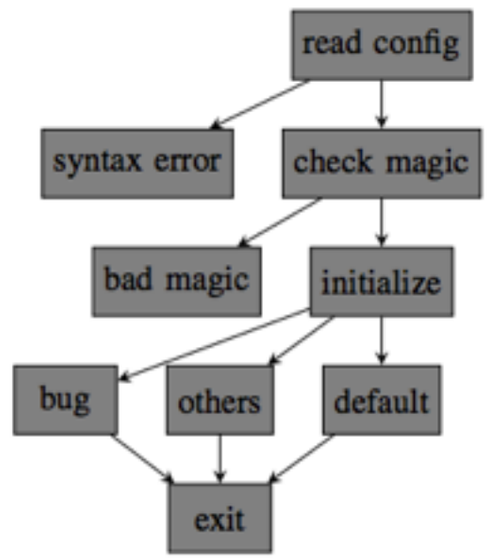


Fig. 4. The nodes found by the second invocation of concolic execution.

ØxOPOSEC Meetup

DARPA CGC

Mechanical Phish - ANGR

- ANGR is a very powerful binary analysis framework. It was implemented mostly by the Shellphish team and is one of the main components of Driller

- It's one of the most recent open source technologies to perform reversing/cracking tasks

- We can easily accomplish *control-flow* analysis, i.e., realize the damn conditions that make the program reach a specific state of execution

- First, it translates the binary in VEX Intermediate Representation. Then, simulates instructions in a simulation engine -> symbolic execution: SimuVEX

- Finally, they use a custom Z3 wrapper. It is called claripy: *"a abstracted constraint-solving wrapper"*

Demo 3 - Angr

https://goo.gl/42T4mi

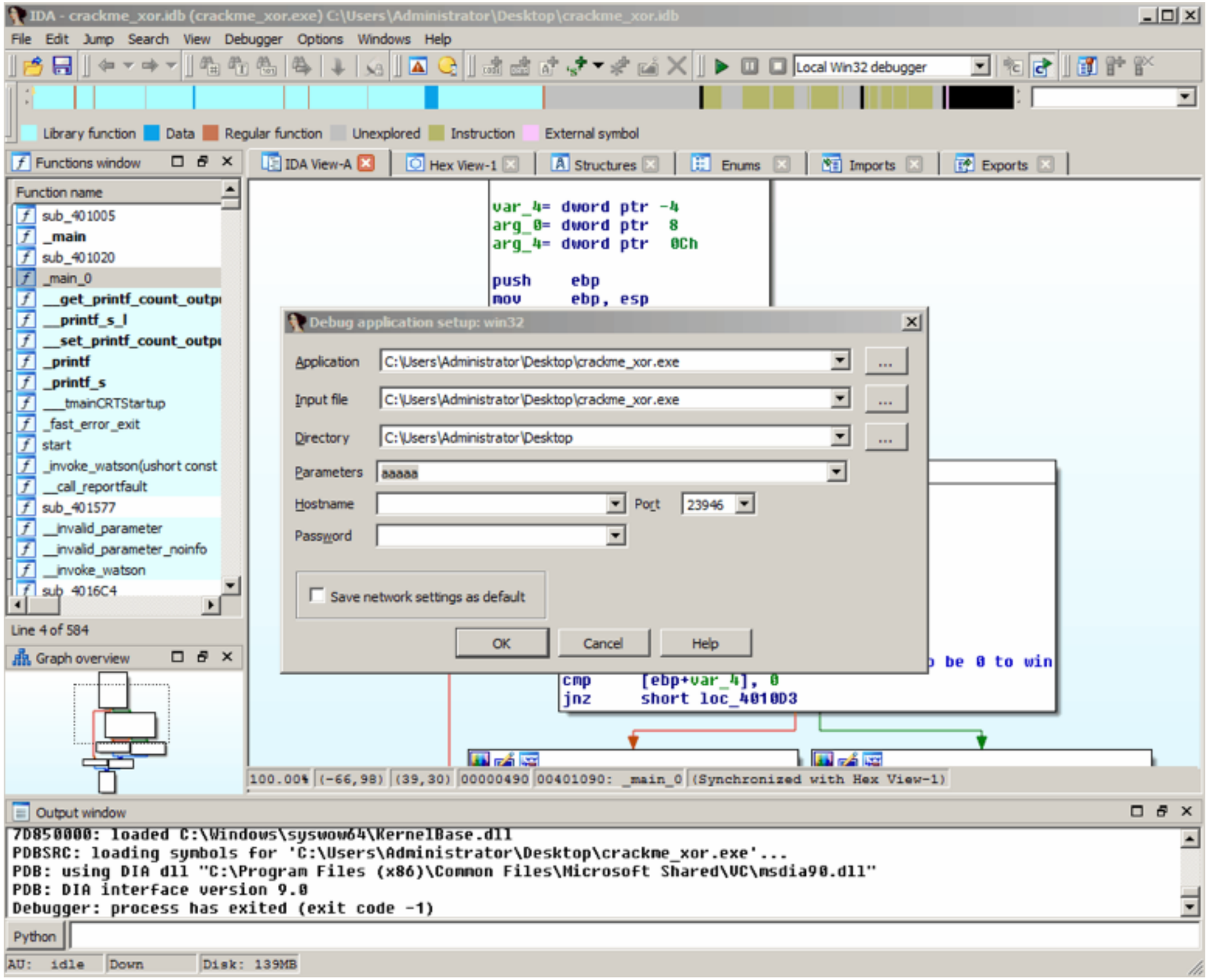
ØxOPOSEC Meetup

Ponce

IDA plugin contest - 2016

- **Taint analysis:** this mode is used to easily track “where” a user input occurs inside a program and observe all the propagations related with the given input

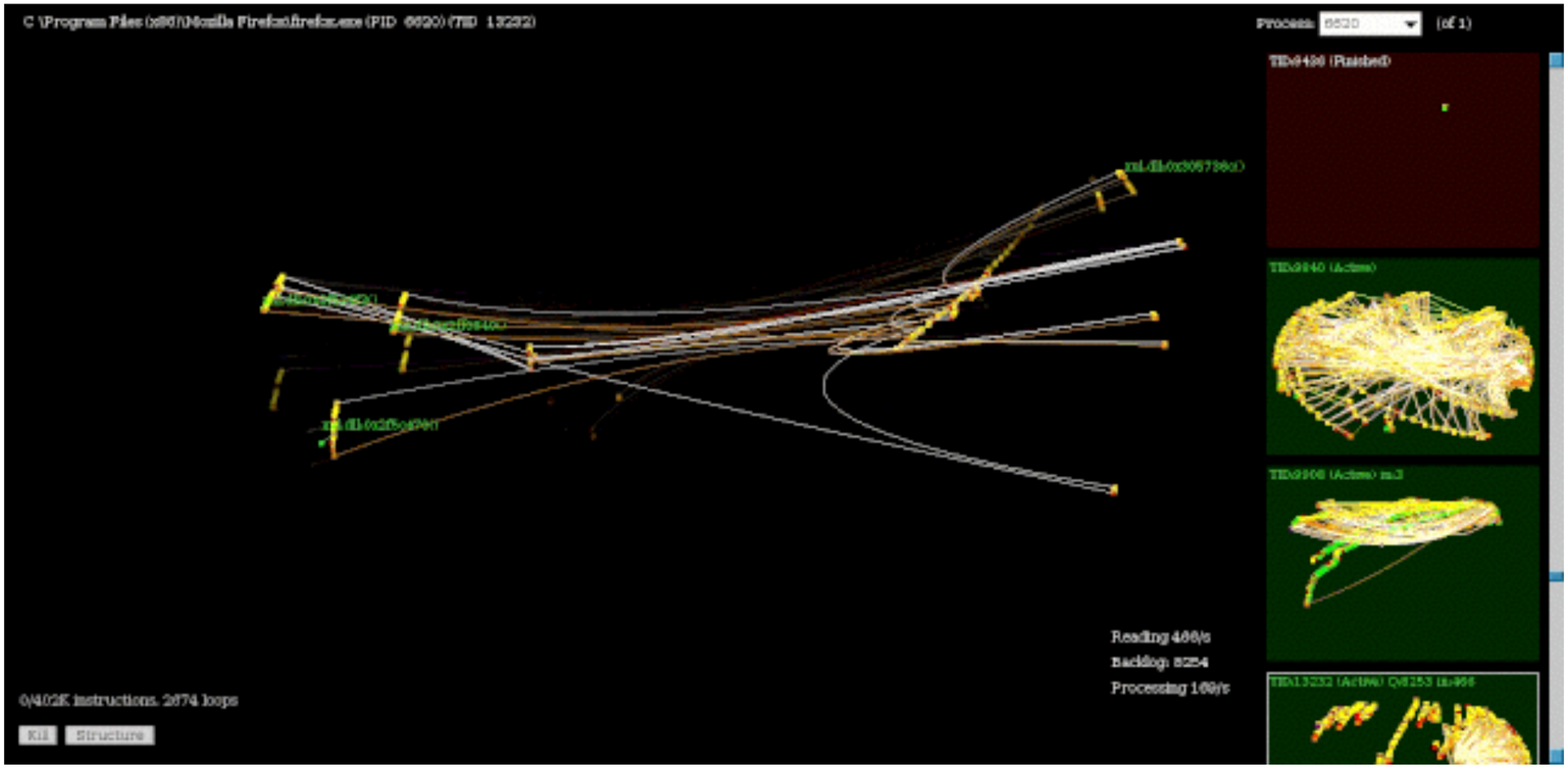
- **Symbolic analysis:** in this mode, the plugin maintains a symbolic state of registers and memory at each step in a binary’s execution path, allowing the user to solve user-controlled conditions to do manually guided execution


IDA - crackme_xor.idb (crackme_xor.exe) C:\Users\Administrator\Desktop\crackme_xor.idb
File Edit Jump Search View Debugger Options Windows Help
Local Win32 debugger
Library function Data Regular function Unexplored Instruction External symbol
Functions window
Function name
sub_401005
_main
sub_401020
_main_0
__get_printf_count_outp
__printf_s_l
__set_printf_count_outp
_printf
_printf_s
___tmainCRTStartup
_fast_error_exit
start
_invoke_watson(ushort const
__call_reportfault
sub_401577
__invalid_parameter
__invalid_parameter_noinfo
__invoke_watson
sub_4016C4
Line 4 of 584
Graph overview
IDA View-A
Hex View-1
Structures
Enums
Imports
Exports
var_4= dword ptr -4
arg_0= dword ptr 8
arg_4= dword ptr 0Ch
push ebp
mov ebp, esp
Debug application setup: win32
Application C:\Users\Administrator\Desktop\crackme_xor.exe
Input file C:\Users\Administrator\Desktop\crackme_xor.exe
Directory C:\Users\Administrator\Desktop
Parameters aaaaa
Hostname
Port 23946
Password
Save network settings as default
OK
Cancel
Help
be 0 to win
cmp [ebp+var_4], 0
jnz short loc_4010D3
100.00% (-66,98) (39,30) 00000490 00401090: _main_0 (Synchronized with Hex View-1)
Output window
7D850000: loaded C:\Windows\syswow64\KernelBase.dll
PDBSRC: loading symbols for 'C:\Users\Administrator\Desktop\crackme_xor.exe'...
PDB: using DIA dll "C:\Program Files (x86)\Common Files\Microsoft Shared\VC\msdia90.dll"
PDB: DIA interface version 9.0
Debugger: process has exited (exit code -1)
Python
AU: idle Down Disk: 139MB

ØxOPOSEC Meetup

RGAT

An instruction trace visualisation tool



Useful links to fry your brain

(Over 1337 ºC)

- **Chill**
  - https://github.com/RPISEC/MBE (lectures 2 and 3)
  - Reddit
    - https://reddit.com/r/reverseengineering
    - https://reddit.com/r/netsec
- **Practice**
  - http://reversing.kr
  - https://ringzer0team.com
  - http://crackmes.de
  - https://ctftime.org (Read CTF writeups and try to solve some available challenges)
  - Play CTFs!

ØxOPOSEC Meetup

THE END

Security through obscurity

André Baptista

@0xACB