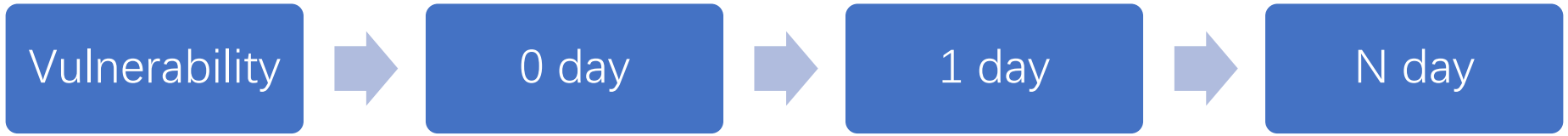
北方工业大学 | 信息安全协会

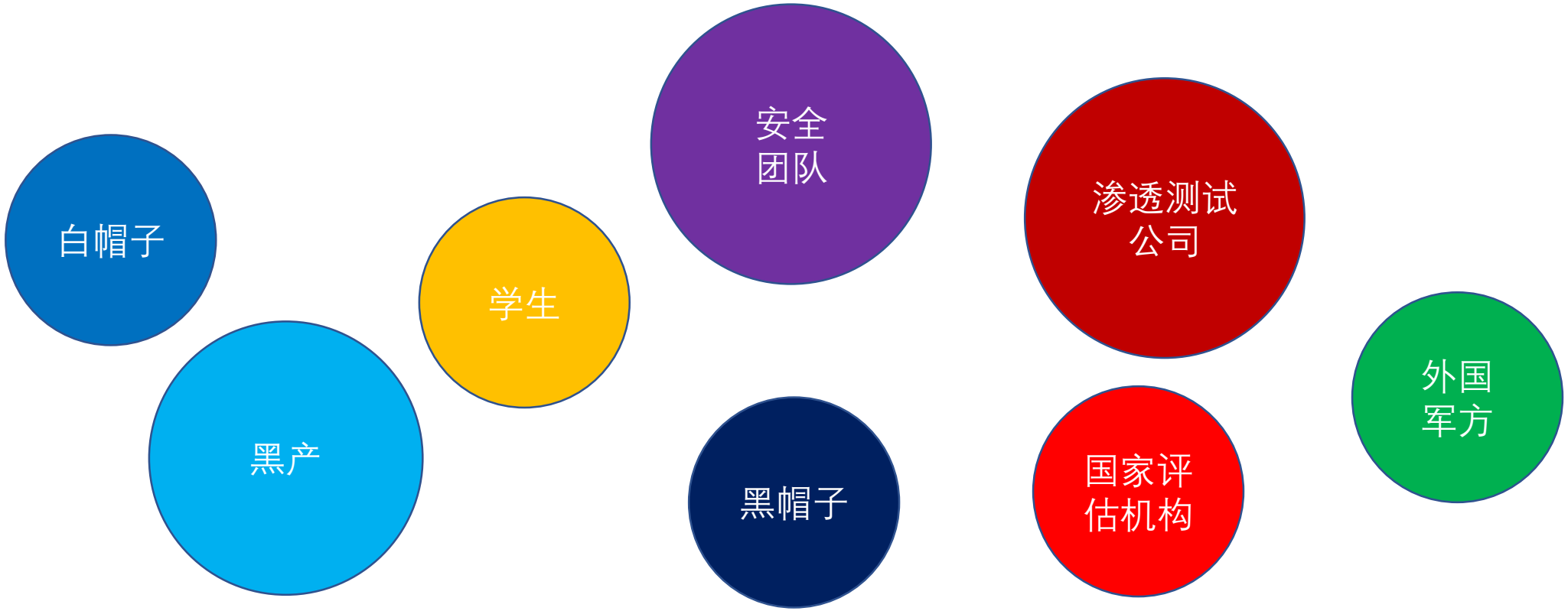
Aspects on Automatic Vulnerability Detection

自 动 漏 洞 挖 掘 面 面 观

谭兴邦（@PolluxAvenger）

About Vulnerability





Kevin David Mitnick

1963 年 8 月 6 日

狮子座

电影：《骇客追缉令》

纪录片：《你瞧，网络世界的幻想》

《线上幽灵：世界头号黑客米特尼克自传》

什么是黑客?

登录克里姆林宫，巡游五角大楼，入侵北美导弹防御网

什么是黑客？黑客就是视铜墙铁壁于无物


NSA Hackers
Shadow Brokers

Shadow Brokers

100万 BitCoin = 568000000$

1 万 BitCoin

200 Zcash = 65000$

16000 Zcash = 4000000$

2016 年 08 月 第一次拍卖

2016 年 10 月 第二次拍卖

2016 年 12 月 第三次拍卖

2017 年 01 月 退出江湖

2017 年 04 月 重出江湖

2017 年 05 月 WannaCry 登场

2017 年 06 月 订阅模式

2017 年 07 月 涨价

2017 年 09 月 再度涨价

National Security Agency

放出 EternalBlue 漏洞

100 Zcash = 23000$

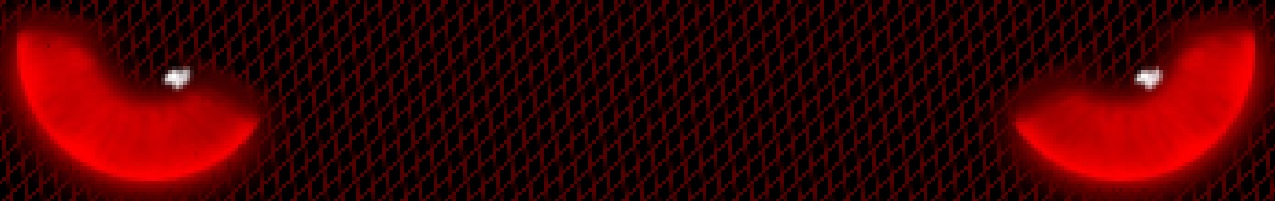
VIP = 400 Zcash = 130000$


Wana Decrypt0r 2.0
What Happened to My Computer?
Your important files are encrypted.
Many of your documents, photos, videos, databases and other files are no longer accessible because they have been encrypted. Maybe you are busy looking for a way to recover your files, but do not waste your time. Nobody can recover your files without our decryption service.
Can I Recover My Files?
Sure. We guarantee that you can recover all your files safely and easily. But you have not so enough time.
You can decrypt some of your files for free. Try now by clicking <Decrypt>.
But if you want to decrypt all your files, you need to pay.
You only have 3 days to submit the payment. After that the price will be doubled.
Also, if you don't pay in 7 days, you won't be able to recover your files forever.
We will have free events for users who are so poor that they couldn't pay in 6 months.
How Do I Pay?
Payment is accepted in Bitcoin only. For more information, click <About bitcoin>.
Please check the current price of Bitcoin and buy some bitcoins. For more information, click <How to buy bitcoins>.
And send the correct amount to the address specified in this window.
After your payment, click <Check Payment>. Best time to check: 9:00am - 11:00am
Payment will be raised on
5/15/2017 16:32:52
Time Left
02:23:59:49
Your files will be lost on
5/19/2017 16:32:52
Time Left
06:23:59:49


Wana Decrypt0r 2.0
Ooops, your files have been encrypted!
Chinese (simpl
我的电脑出了什么问题？
您的一些重要文件被我加密保存了。
照片、图片、文档、压缩包、音频、视频文件、exe文件等，几乎所有类型的文件都被加密了，因此不能正常打开。
这和一般文件损坏有本质上的区别。您大可在网上找找恢复文件的方法，我敢保证，没有我们的解密服务，就算老天爷来了也不能恢复这些文档。
有没有恢复这些文档的方法？
当然有可恢复的方法。只能通过我们的解密服务才能恢复。我以人格担保，能够提供安全有效的恢复服务。
但这是收费的，也不能无限期的推迟。
请点击 <Decrypt> 按钮，就可以免费恢复一些文档。请您放心，我是绝不会骗你的。
但想要恢复全部文档，需要付款点费用。
是否随时都可以固定金额付款，就会恢复的吗，当然不是，推迟付款时间越长对你不利。
最好3天之内付款费用，过了三天费用就会翻倍。
还有，一个礼拜之内未付款，将会永远恢复不了。
对了，忘了告诉你，对半年以上没钱付款的穷人，会有活动免费恢复，能否轮
Payment will be raised on
5/16/2017 10:12:00
Time Left
02:22:47:24
Your files will be lost on
5/20/2017 10:12:00
Time Left
06:22:47:24
About bitcoin
How to buy bitcoins?
bitcoin ACCEPTED HERE
Send $300 worth of bitcoin to this address:
12t9YDPgwueZ9NyMgw519p7AA8isjr6SMw
Copy


Ooops, your money have been encrypted!
Chinese (simpl
我的钱包出了什么问题?
您的一些血汗钱被我们拿走了。
现金、存折、信用卡、支付宝、微信钱包等，几乎所有类型的支付手段都已经帮您用过了。
这和一般忘记钱包放在哪有本质上的区别。您大可在网上查找恢复钱包的方法，不过我敢保证，没有我们的平台，就算老天爷来了也不能恢复您的钱包
有没有恢复我的钱包的方法?
当然有恢复的方法。只要获得我们所有的产品即可恢复。我以人格担保，游戏数量是会增加的。
因为这是收费的，所以推迟对你没有好处。
请点击<Free>按钮，就可以获得一些免费游戏。请你放心，是不存在内购的。
但是想要获得全部游戏，需要支付点费用。
是否随时都可以固定金额付款，就会获得的吗，当然不是，推迟付款时间越长对你不利。
最好在打折期间之内付款，过了促销价格就会+107%off。
还有，预购期间之内未付款，将无法获得一些限定道具。
对了，忘了告诉你，对活动内没钱喜加一的穷人，之后可能有限时领取的活动能否轮到你的游戏，就要看你运气如何了。
Payment will be raised on
5/19/2017 15:23:13
Time Left
02:23:59:54
Your money will be lost on
5/23/2017 15:23:13
Time Left
06:23:59:54
About Valve
How to get money?
-25%
Buy your game on this address:
http://store.steampowered.com/
go


我们给予阁下公司最后的让步
Palmiro Panicucci to: zhixuexia, connor
From: "Palmiro Panicucci" <Palmiro.Panicucci@mail.com>
To: zhixuexia <ZhixueXia@hixxionhillschina.com>, connor <cro2@foxmail.com>
History: This message has been forwarded
经过内部沟通，我们现在把赎金降至40个比特币。支付完成后我们将告知解锁密码，网络中的多处漏洞
件，未来也将绝不会再骚扰贵公司。阁下也应该理解我们，不可能白白损失的时间和精力，否则也不会
我们深知中国是个政治社会，如果在中共19大召开之际，制造特殊情况势必会对阁下公司造成巨大打击，
金可以满足贵公司的期望。谢谢。

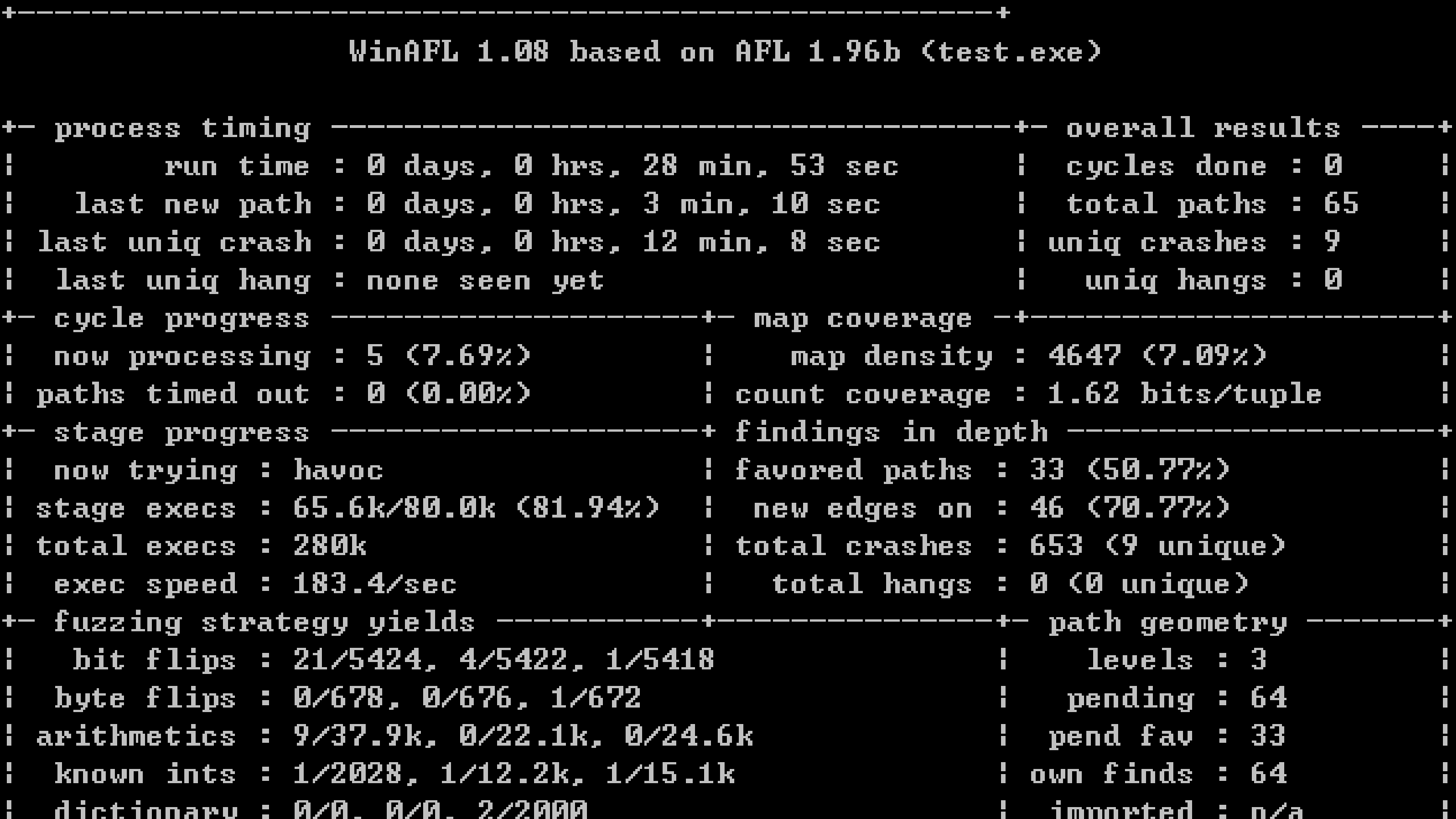


It's Not Over! WannaCry?

Version 2.0 Arrived


Google

白帽团队	谷歌		微软		苹果	致谢次数
	Android	Chrome	漏洞公告	赏金项目		
谷歌安全团队	22	0	104	0	88	214
360	126	2	48	4	29	209
趋势科技	47	1	48	0	39	135
腾讯	24	10	31	0	36	101
微软安全团队	0	0	63	0	0	63
阿里巴巴	11	0	0	0	5	16
FireEye	0	0	12	1	0	13
苹果安全团队	0	0	0	0	12	12
长亭科技	0	0	1	0	8	9
Zimperium zLabs Team	0	0	0	0	8	8
百度	3	0	0	0	5	8


WinAFL 1.08 based on AFL 1.96b (test.exe)
process timing
run time : 0 days, 0 hrs, 28 min, 53 sec
last new path : 0 days, 0 hrs, 3 min, 10 sec
last uniq crash : 0 days, 0 hrs, 12 min, 8 sec
last uniq hang : none seen yet
overall results
cycles done : 0
total paths : 65
uniq crashes : 9
uniq hangs : 0
cycle progress
now processing : 5 (7.69%)
paths timed out : 0 (0.00%)
map coverage
map density : 4647 (7.09%)
count coverage : 1.62 bits/tuple
stage progress
now trying : havoc
stage execs : 65.6k/80.0k (81.94%)
total execs : 280k
exec speed : 183.4/sec
findings in depth
favored paths : 33 (50.77%)
new edges on : 46 (70.77%)
total crashes : 653 (9 unique)
total hangs : 0 (0 unique)
fuzzing strategy yields
bit flips : 21/5424, 4/5422, 1/5418
byte flips : 0/678, 0/676, 1/672
arithmetics : 9/37.9k, 0/22.1k, 0/24.6k
known ints : 1/2028, 1/12.2k, 1/15.1k
path geometry
levels : 3
pending : 64
pend fav : 33
own finds : 64

Aspects on Automatic Vulnerability Detection

是否让程序运行? ⇨ Static / Dynamic Analysis

是否进行软件测试? ⇨ Smart Fuzzing

是否进行逻辑推理? ⇨ Symbolic Execution

Aspects on Automatic Vulnerability Detection

是否让程序运行? ⇨ Static / Dynamic Analysis

是否进行软件测试? ⇨ Smart Fuzzing

是否进行逻辑推理? ⇨ Symbolic Execution

Static Analysis

优势

分析程序所有可能运行的基本块

不需要运行目标程序

劣势

指针/引用/别名分析

加密分析

误报率高、漏报率高

Dynamic Analysis

优势

得到信息准确，执行表现的功能真实

可解决静态分析无力的情况

劣势

执行恶意样本导致环境风险不可控

执行与监控资源消耗大

误报率低、对输入依赖高

Dynamic Taint Analysis 动态污点分析

Taint Source

污染源头？开始污染的位置在哪？

Taint Policy

污染怎样在数据中传递？

Taint Sink

关键污染检查点？

Questions

受污染的地址 / 控制依赖 / 去污染化

e.g. TaintCheck

```
x = input()
y = 10 + x
goto y

y = load(base + input)

if(input == 1)
      y = 2;
z = 3

input ^ input
```

污染源是 x

污染通过赋值传递给 y

关键污染检查点 goto

Instrumentation　插桩分析技术

Static Source Instrumentation

静态源码插桩

使用影子内存（Shadow map）跟踪 addressable bytes

Detect Vulnerabilities

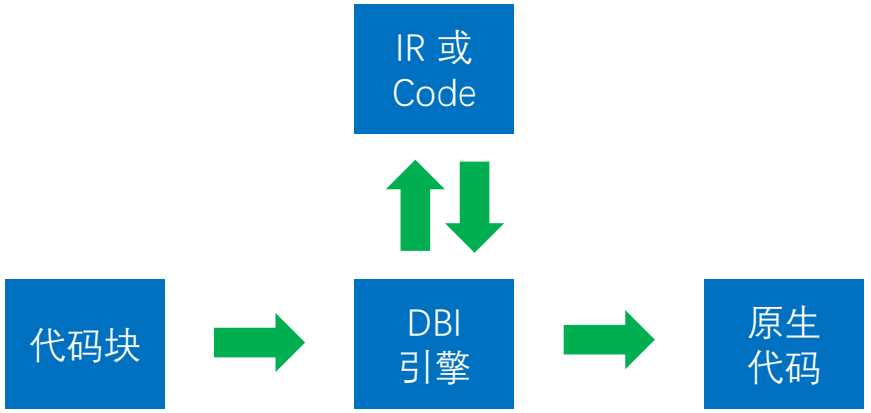
基于空间：buffer overflow(heap/stack)

基于时间：use-after-free, double free

e.g. AddressSanitizer

Dynamic Binary Instrumentation

动态二进制插桩



e.g. PIN / DynamoRIO / Valgrind

Aspects on Automatic Vulnerability Detection

是否让程序运行? ⇨ Static / Dynamic Analysis

是否进行软件测试? ⇨ Smart Fuzzing

是否进行逻辑推理? ⇨ Symbolic Execution

Symbolic Execution　符号执行

- Concrete Execution
- Symbolic Execution

约束求解器

Satisfiability Modulo Theories　可满足性模理论

Question 1：Solver Limitation

① 用摘要代替库函数的求解

② 混合执行

e.g. Z3 / STP / Yices

Question 2：Path Explosion

① 路径搜索策略 DFS / BFS

② 使用代码覆盖度做指导

Symbolic Execution 符号执行

代表工作

e.g. EXE / DART / KLEE / S^2E / SAGE / Mayhem

Industry VS Research

70% 30%

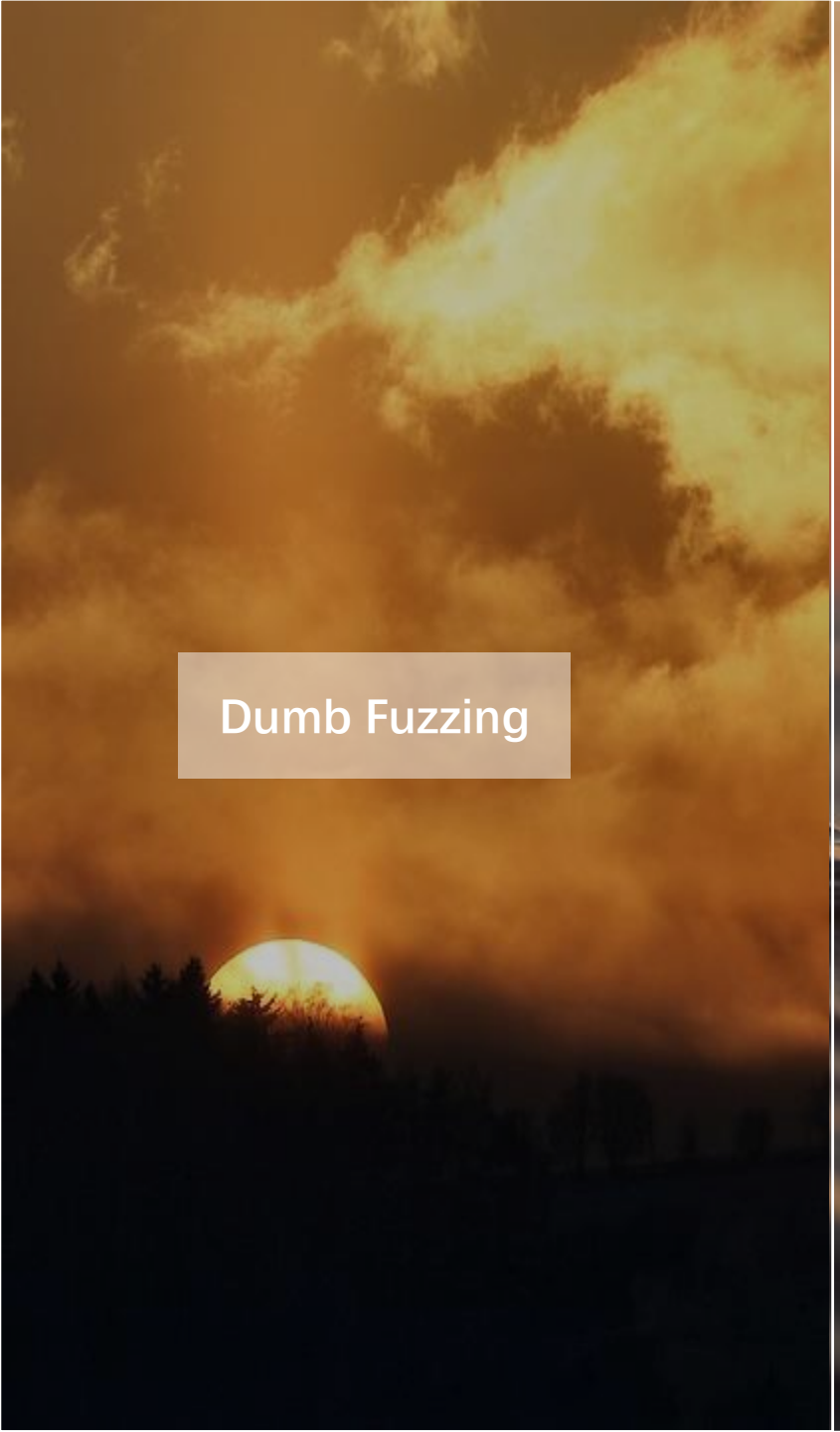
The Future of Symbolic Execution

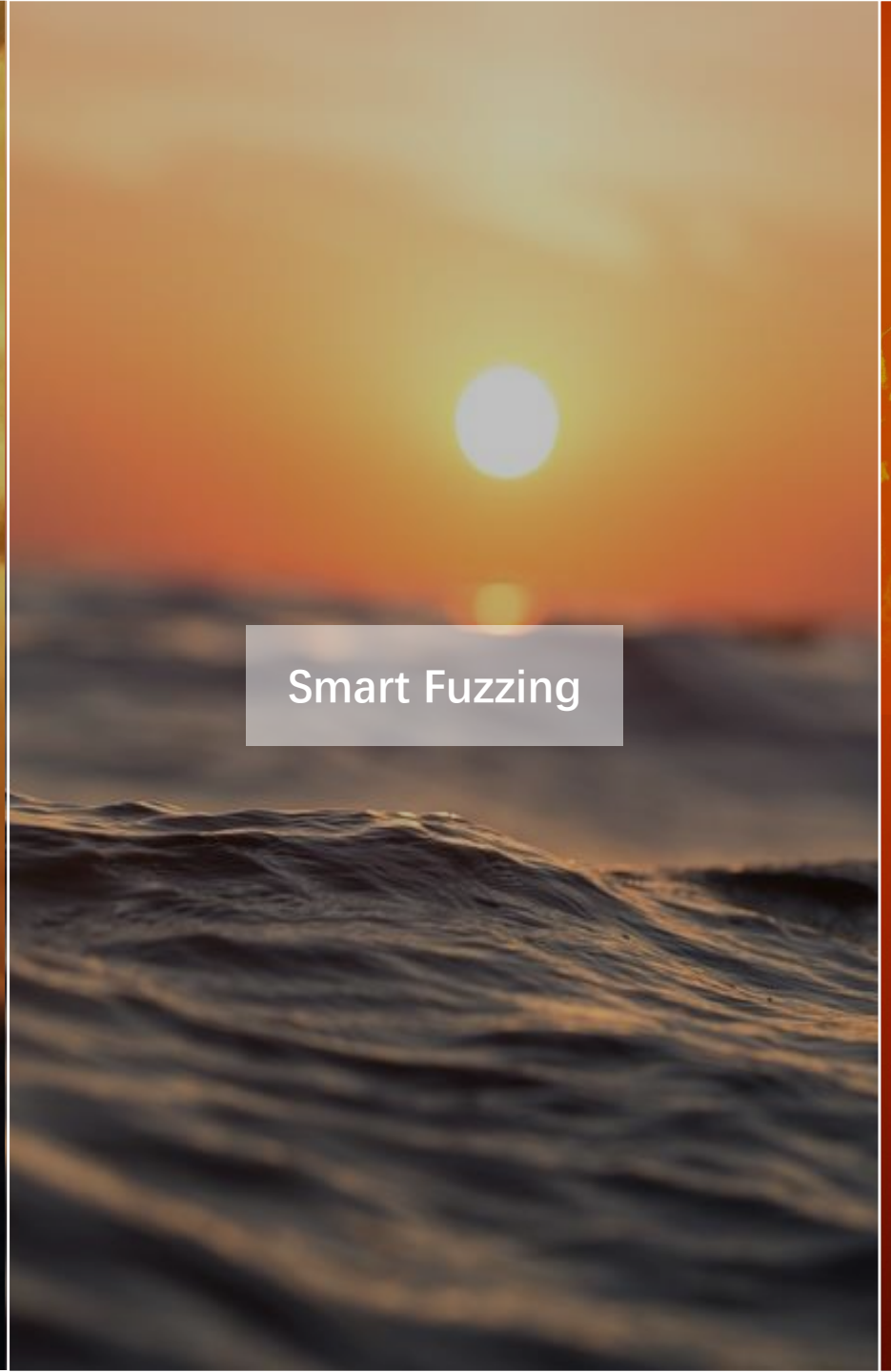
Aspects on Automatic Vulnerability Detection

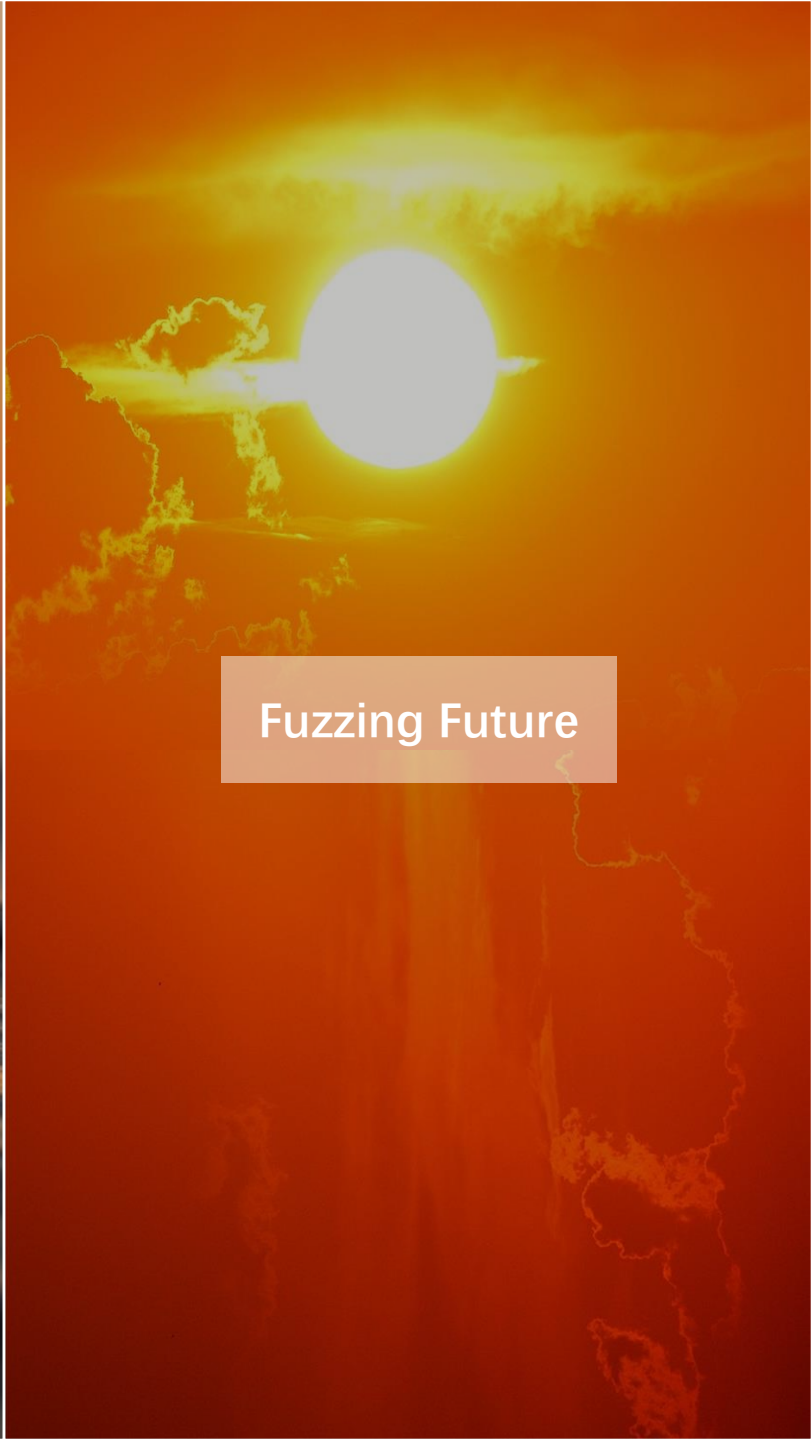
是否让程序运行? ⇨ Static / Dynamic Analysis

是否进行软件测试? ⇨ Smart Fuzzing

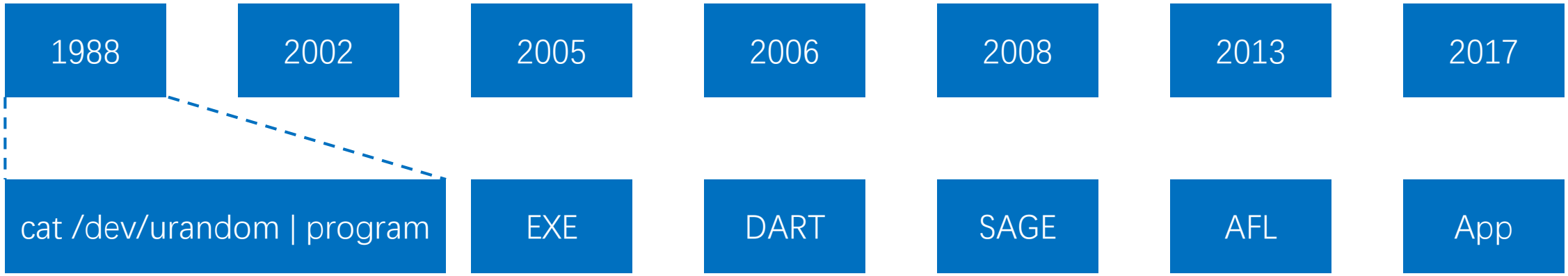
是否进行逻辑推理? ⇨ Symbolic Execution


Dumb Fuzzing


Smart Fuzzing


Fuzzing Future

Evolution of Smart Fuzzing



Smart Fuzzing

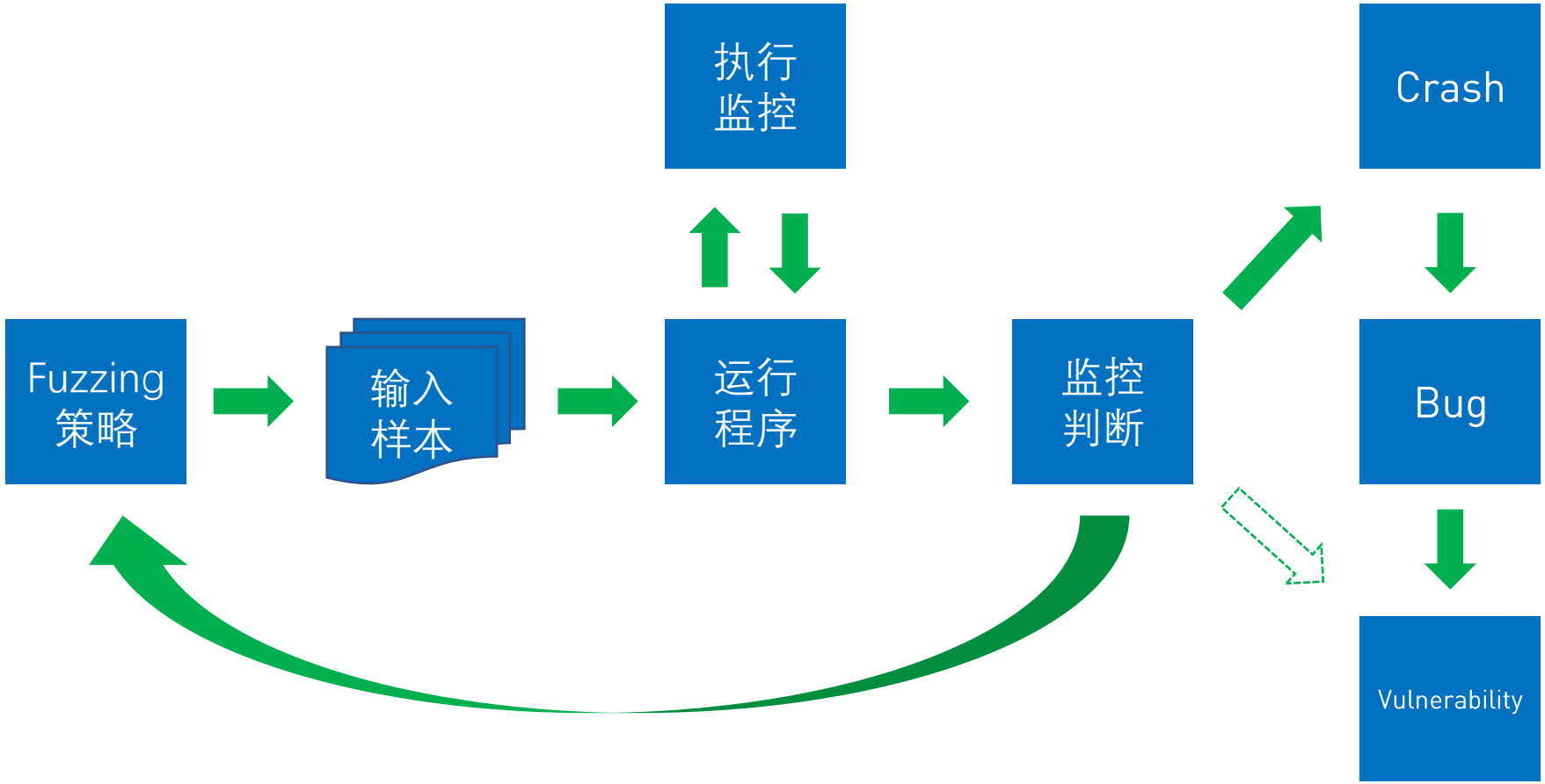
	BlackBox	GreyBox	WhiteBox
基于畸变	Miller	AFL Driller VUzzer TaintScope	SAGE libfuzzer
基于生成	SPIKE Peach Sulley		

Explore Strategy

① 导向型

② 基于覆盖度型

Smart Fuzzing



Smart Fuzzing

先验知识	难易度	覆盖度	校验和
不需要 先验知识	开始简单 容易 自动化	初始样本 相关度 极大	无法通过 校验和与 复杂检查
对协议的 理解是 先验知识	生成器 依赖 协议的 复杂度	完备的 Completeness	可应对 复杂检查 与校验和

American Fuzzy Lop (AFL)

基于覆盖度的，基于畸变的，灰盒的 fuzzer

Smart ?

保留好的输入样本，加以畸变

Good ?

执行触发新基本块的为好样本

Efficient ?

几乎没有性能损失

Challenges

★ 校验和检查

★ 样本间的优先级

★ 畸变位置

★ 畸变策略

★ 依赖初始种子文件

Dumb VS Smart in real world ?

The Future of Fuzzing

35M $ | 2M $ | 1M $

Mayhem

Defcon CTF last but two

Cyber Grand Challenge

DARPA

32bit x86 + 7 system call

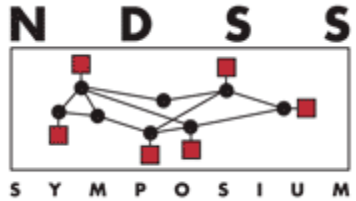
Heartbleed

New Trend in Vulnerability Discovery

融合

智能

Top 4 Conferences in Information Security



Network and Distributed System Security Symposium

USENIX Security

ACM Symposium on Computer and Communications Security

IEEE Symposium on Security and Privacy

TaintScope (S&P 2010, Tielei Wang)

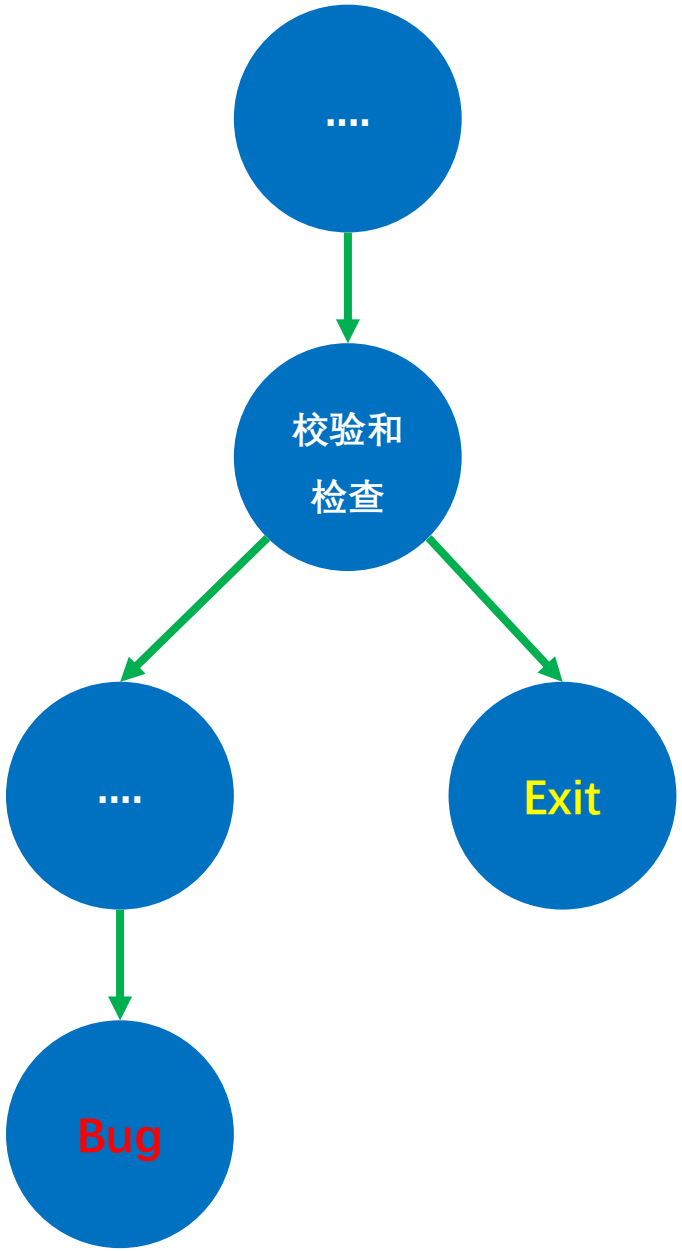
Challenge

Fuzzing 过程中的校验和检查

Solution

利用动态污点分析解决校验和检查

通常会走 false 分支



Driller (NDSS 2016, Shellphish)

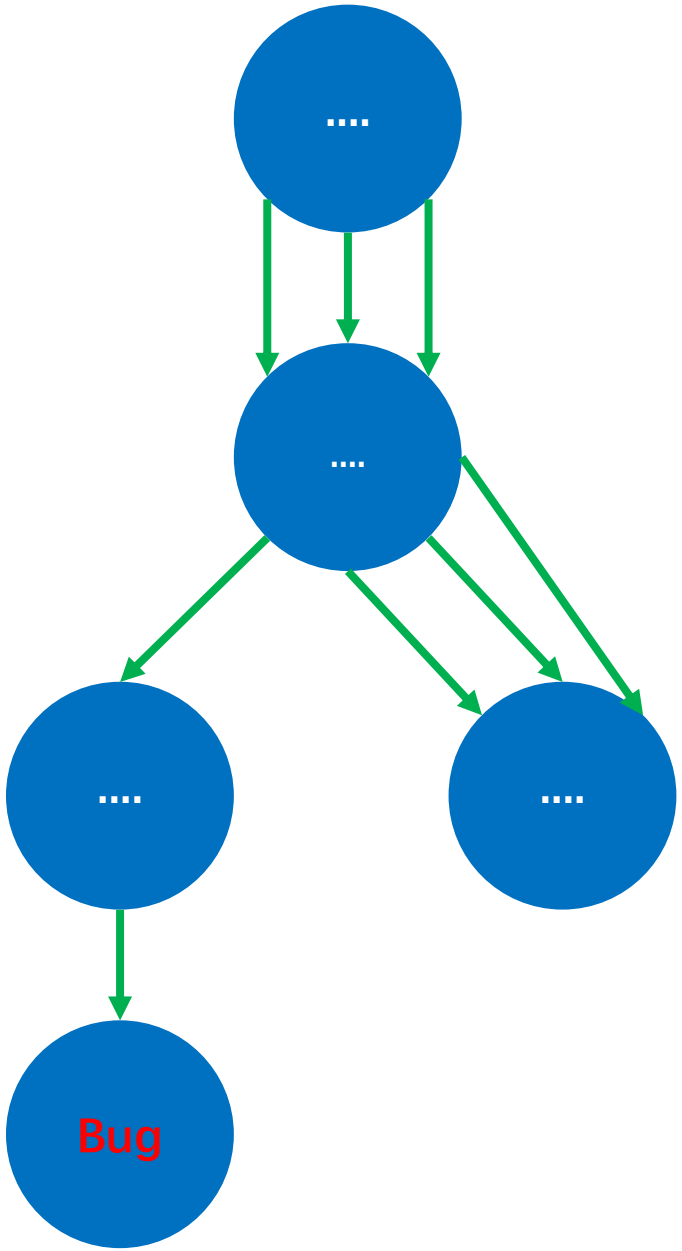
Challenge

Fuzzing 过程中被卡在某一点

Solution

利用符号执行在卡住时探索另一条路

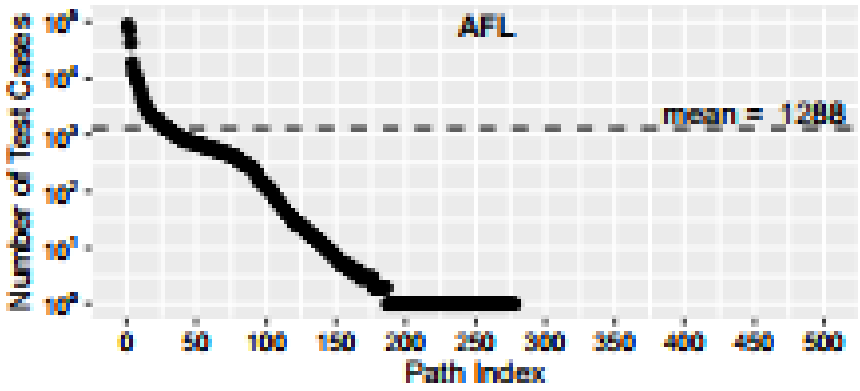
通用型 Fuzzer 可以联合符号执行



AFLfast (CCS 2016, Marcel Bohme)

Challenge

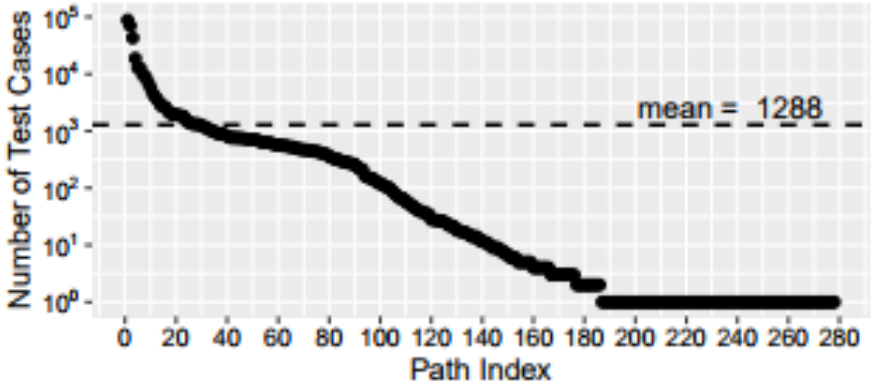
AFL 将大量时间浪费在执行高频路径上



Solution

同频路径分配不同的优先级

利用马尔科夫链为低频路径分配更高优先级



VUzzer (NDSS 2017, Sanjay Rawat)

Challenge

选择样本的优先级、畸变的位置、畸变的策略

Solution

利用静态与动态分析辅助

数据流解决畸变的位置、畸变的策略

控制流解决选择样本的优先级

```
1 int main(int argc, char **argv){
2     unsigned char buf[1000];
3     int i, fd, size, val;
4     if ((fd = open(argv[1], O_RDONLY)) == -1)
5         exit(0);
6     fstat(fd, &s);
7     size = s.st_size;
8     if (size > 1000)
9         return -1;
10    read(fd, buf, size);
11    if (buf[1] == 0xEF && buf[0] == 0xFD) // notice the order of CMPs
12        printf("Magic bytes matched!\n");
13    else
14        EXIT_ERROR("Invalid file\n");
15    if (buf[10] == '%' && buf[11] == '@') {
16        printf("2nd stop: on the way...\n");
17        if (strncmp(&buf[15], "MAZE", 4) == 0) // nested IF
18            ... some bug here ...
19        else {
20            printf("you just missed me...\n");
21            ... some other task ...
22            close(fd); return 0;
23        }
24    } else {
25        ERROR("Invalid bytes");
26        ... some other task ...
27        close(fd); return 0;
28    }
29    close(fd); return 0;
30 }
```

New Trend in Vulnerability Discovery

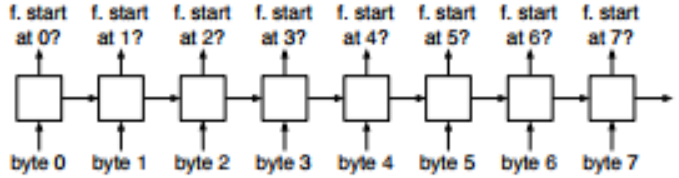
融合

智能

Recognize functions with RNN (USENIX Security 2015, Eui Chul Richard Shin)

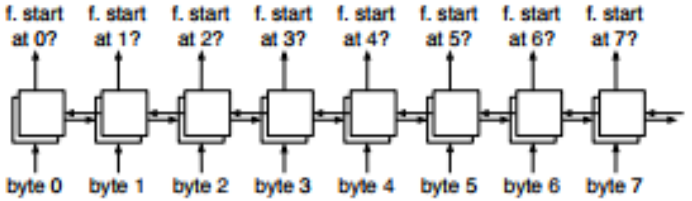
Challenge

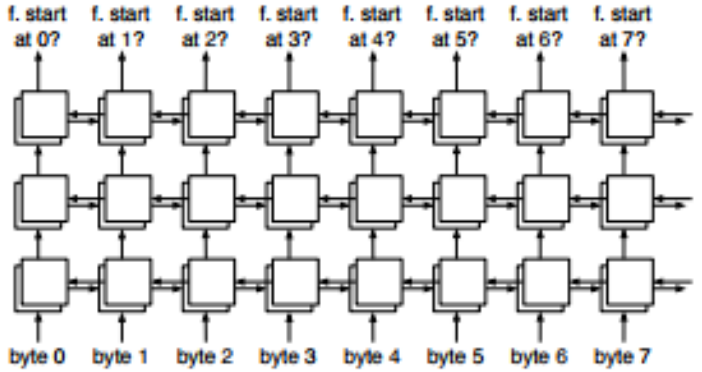
程序中哪一个字节是函数的开始?



Solution

基于后面的字节判断前面的字节是不是函数的开始





VDiscover (Codaspy 2016, Gustavo Grieco)

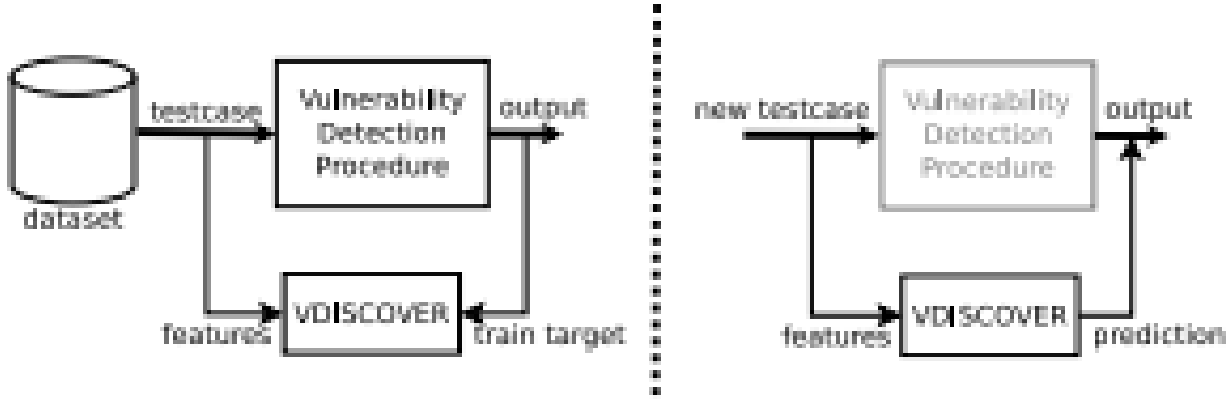
Challenge

如何根据给定的测试集和程序预测漏洞的位置

Solution

提取 API 调用序列和参数作为特征

训练：逻辑回归、随机森林、多层神经网络



Learn&Fuzz (2017, Patrice Godefroid)

Challenge

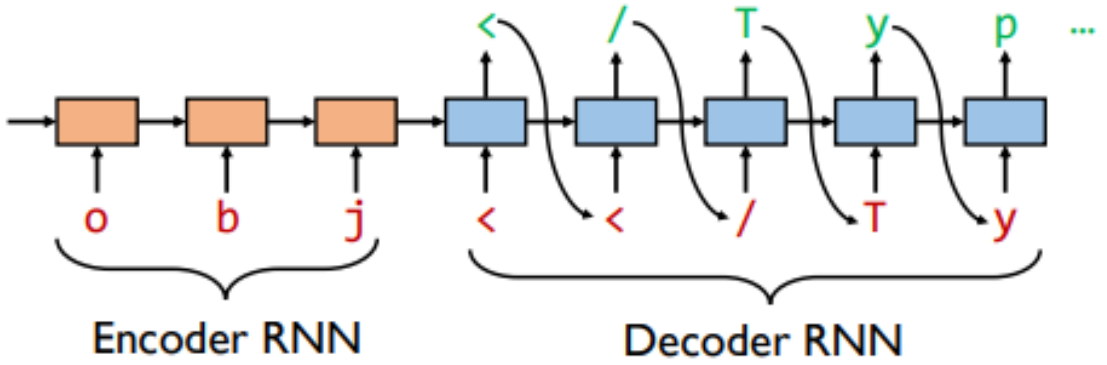
Fuzzer 在格式不良好的输入样本情况下的代码覆盖率太低

Algorithm	Coverage	Pass Rate
SampleSpace+Random	563,930	36.97%
baseline+Random	564,195	44.05%
Sample-10K	565,590	78.92%
Sample+Random	566,964	41.81%
SampleFuzz	567,634	68.24%

Solution

使用 RNN 预测 token 的可能值

在执行过程中学习到输入样本的语法格式



Cross the Wall - Bypass All Modern Mitigations of Microsoft Edge (BlackHat Asia 2017, Xiaoning Li)

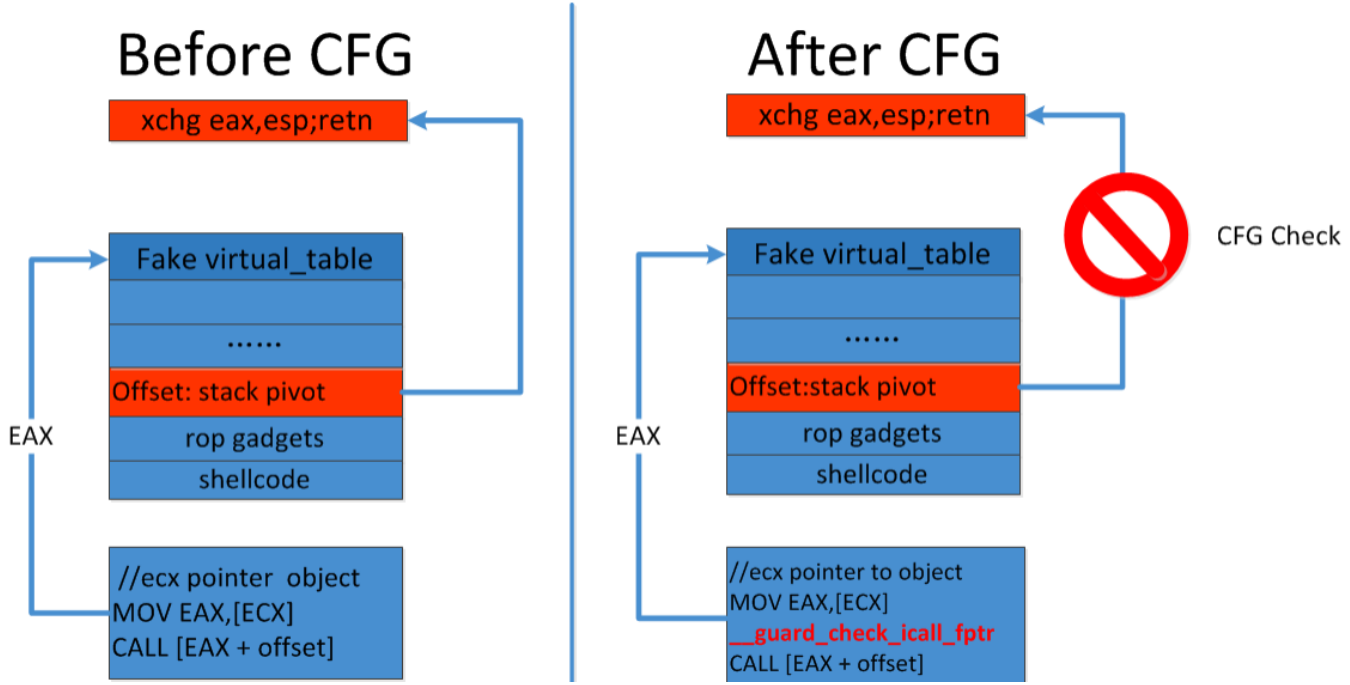
Control Flow Guard　控制流防护

跟踪合法的目标实体，验证运行时的目标实体

Solution

利用 CPU 的 PMI 来收集执行的运行时信息

在上千万的记录中发现了两个语句并不遵循安全机制



北方工业大学 | 信息安全协会

Next:
内核&内存之战

谭兴邦（@PolluxAvenger）

北方工业大学 | 信息安全协会

Aspects on Automatic Vulnerability Detection

自 动 漏 洞 挖 掘 面 面 观

Q & A

谭 兴 邦（@PolluxAvenger）