ISC 2019 第七届互联网安全大会

人工智能在漏洞挖掘领域的应用

邹权臣

360集团安全研究院 智能安全研究员

小鹅助理

扫码添加小鹅助理，与数万科技圈人士

分享重量级活动PPT、干货培训课程、高端会议免费

门票


ISC
邹权臣
360集团安全研究院 智能安全研究员


ISC
人工智能在漏洞挖掘领域的应用
360集团安全研究院　邹权臣

ISC

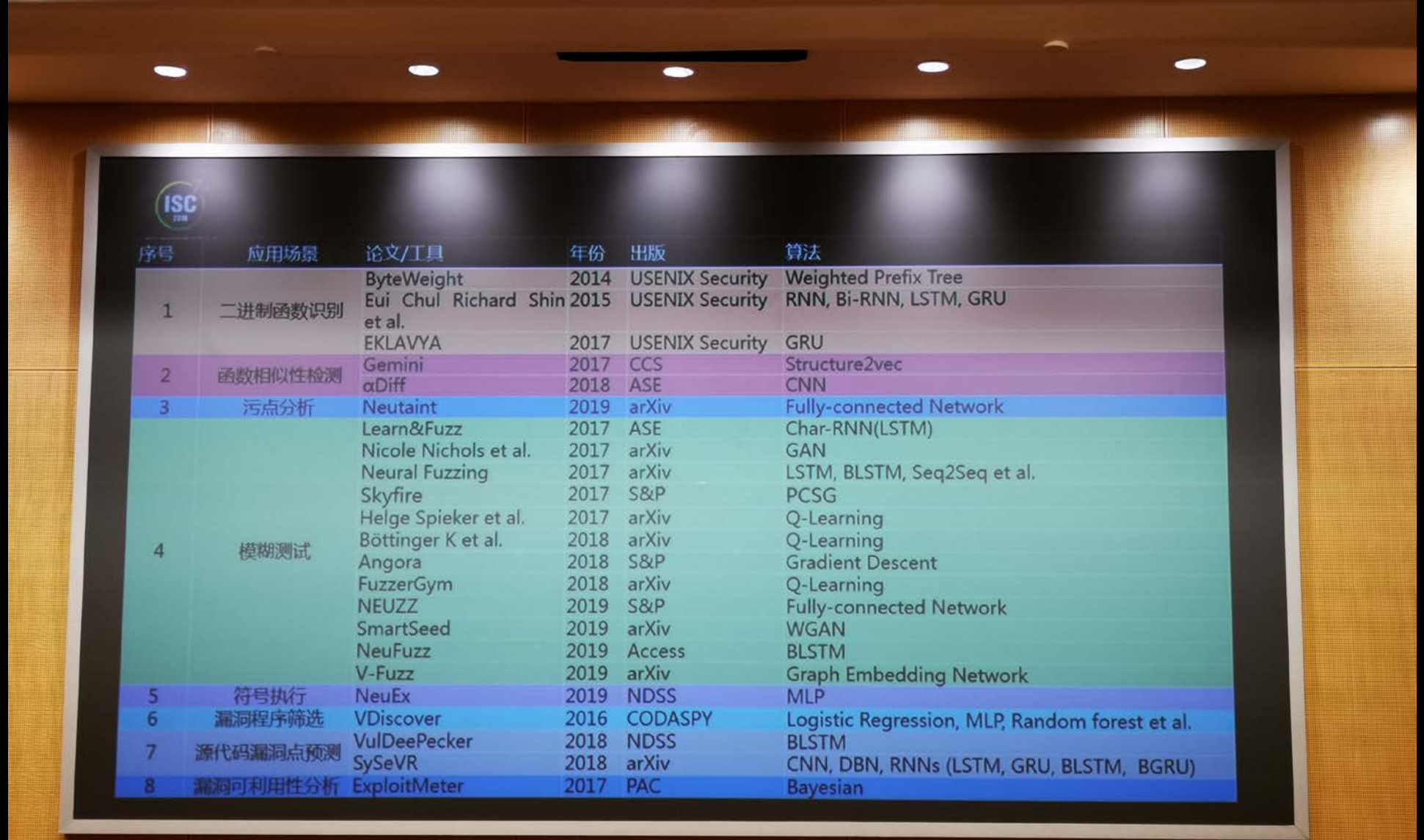
人工智能

CNN、RNN、GAN、DRL等

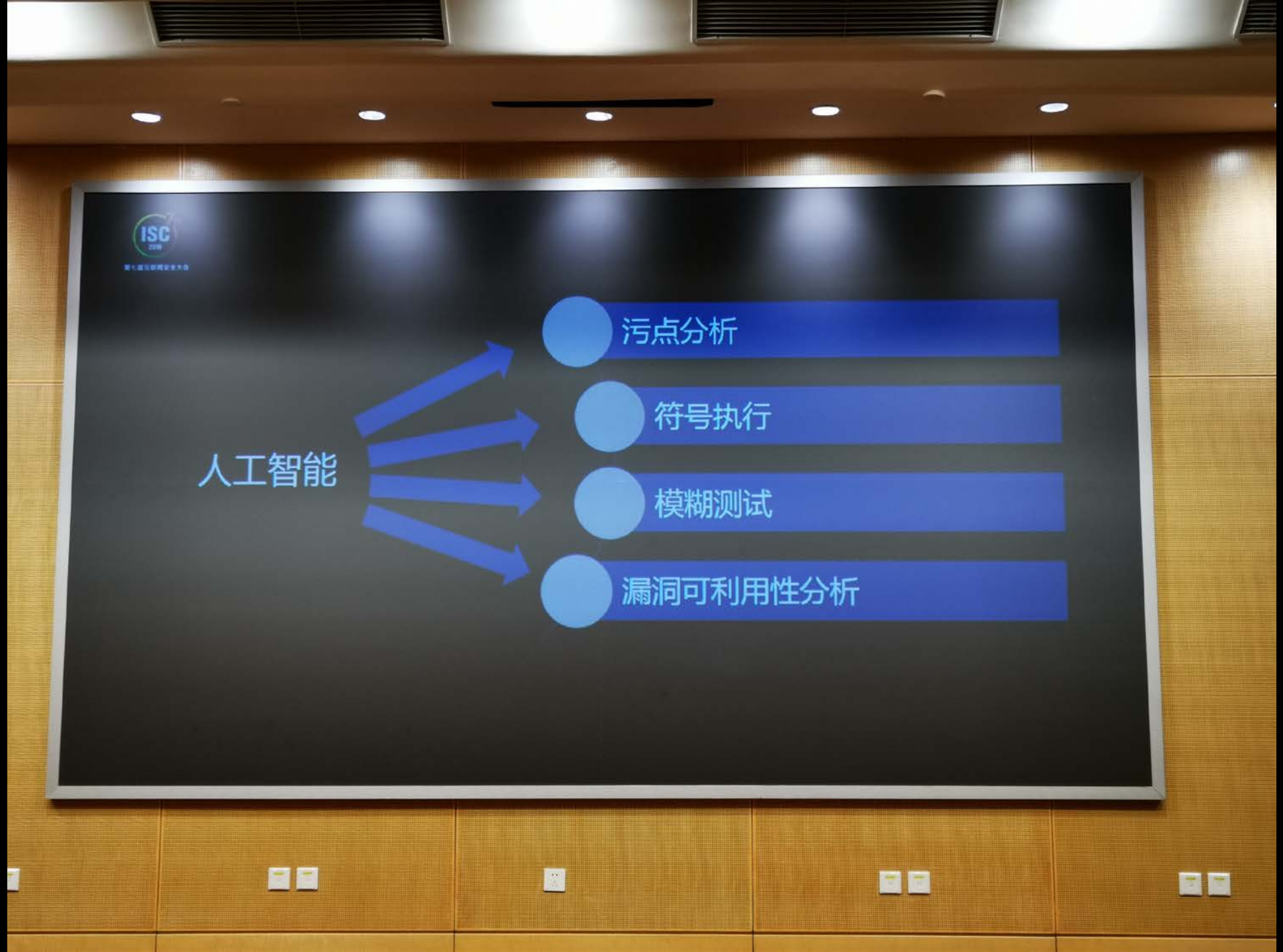
漏洞挖掘

污点分析、符号执行、模糊测试、模型检测、模式匹配等


发表论文数
10
8
6
4
2
0
2014
2015
2016
2017
2018
2019
数据来源：USENIX Security、S&P、CCS、NDSS、ASE、arXiv等
发表论文数呈上升趋势，已经有较多的探索性工作被展开。



序号	应用场景	论文/工具	年份	出版	算法
1	二进制函数识别	ByteWeight	2014	USENIX Security	Weighted Prefix Tree
		Eui Chul Richard Shin et al.	2015	USENIX Security	RNN, Bi-RNN, LSTM, GRU
		EKLAVYA	2017	USENIX Security	GRU
2	函数相似性检测	Gemini	2017	CCS	Structure2vec
		αDiff	2018	ASE	CNN
3	污点分析	Neutaint	2019	arXiv	Fully-connected Network
4	模糊测试	Learn&Fuzz	2017	ASE	Char-RNN(LSTM)
		Nicole Nichols et al.	2017	arXiv	GAN
		Neural Fuzzing	2017	arXiv	LSTM, BLSTM, Seq2Seq et al.
		Skyfire	2017	S&P	PCSG
		Helge Spieker et al.	2017	arXiv	Q-Learning
		Böttinger K et al.	2018	arXiv	Q-Learning
		Angora	2018	S&P	Gradient Descent
		FuzzerGym	2018	arXiv	Q-Learning
		NEUZZ	2019	S&P	Fully-connected Network
		SmartSeed	2019	arXiv	WGAN
		NeuFuzz	2019	Access	BLSTM
		V-Fuzz	2019	arXiv	Graph Embedding Network
5	符号执行	NeuEx	2019	NDSS	MLP
6	漏洞程序筛选	VDiscover	2016	CODASPY	Logistic Regression, MLP, Random forest et al.
7	源代码漏洞点预测	VulDeePecker	2018	NDSS	BLSTM
		SySeVR	2018	arXiv	CNN, DBN, RNNs (LSTM, GRU, BLSTM, BGRU)
8	漏洞可利用性分析	ExploitMeter	2017	PAC	Bayesian

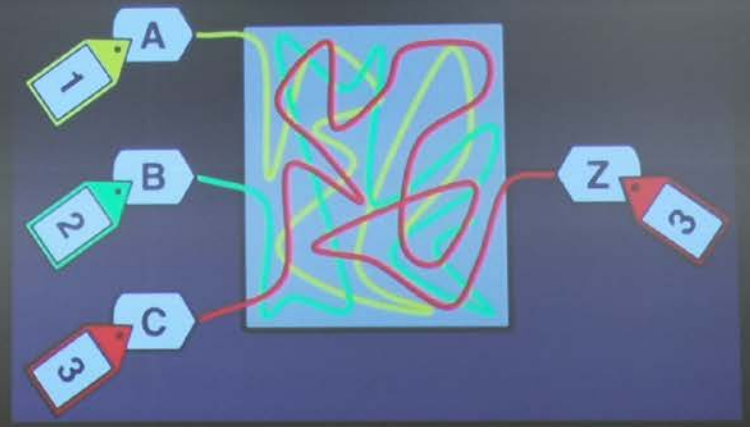

人工智能
污点分析
符号执行
模糊测试
漏洞可利用性分析

人工智能在污点分析中的应用

动态污点分析(Dynamic Taint Analysis ，DTA)

应用

- 自动化漏洞挖掘
- 模糊测试制导
- 信息泄露检测
- 恶意程序行为分析等

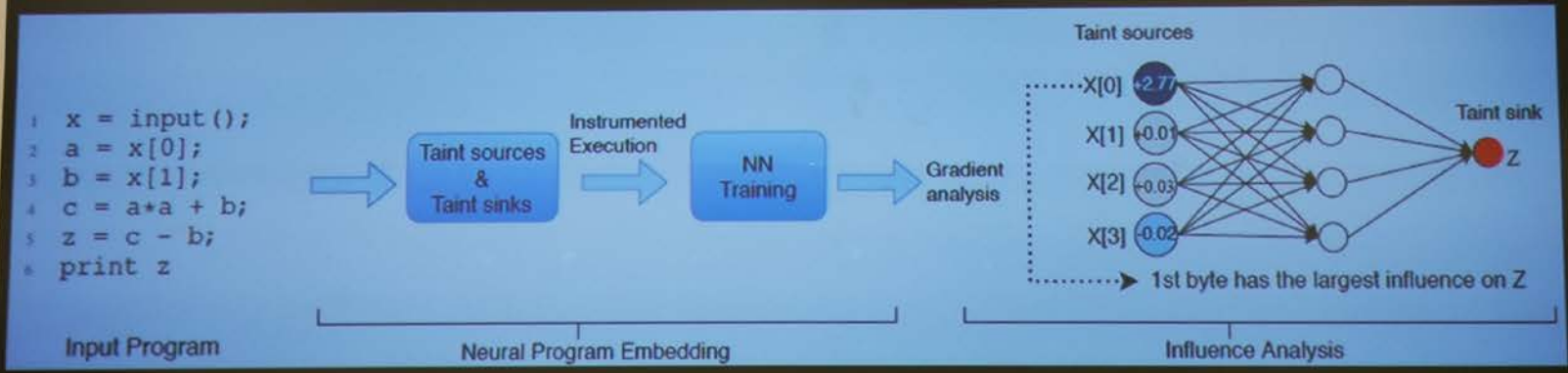


Dytan: A Generic Dynamic Taint analysis Framework (ISSTA 2007)

人工智能在污点分析中的应用——Neutaint

Neutaint

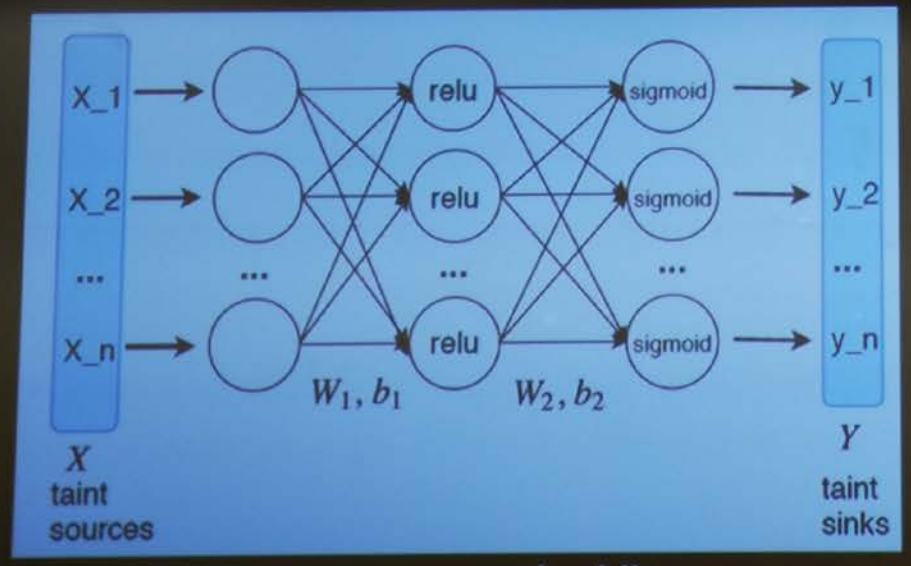
arXiv, 2019年7月， 哥伦比亚大学Suman Jana教授团队



- 提出基于neural program embedding、gradient analysis的信息流跟踪技术
- 提高污点分析的准确率，缓解传播错误累积问题，降低运行时开销

人工智能在污点分析中的应用——Neutaint

Neutaint



dynamic program embeddings

数据收集

- 运行AFL收集变异样本；
- LLVM插装记录CMP指令；

Saliency maps

- **粗粒度信息流**：单个source点对所有sink点的影响
- **细粒度信息流**：单个source对单个sink点的影响

人工智能在污点分析中的应用——Neutaint

Programs	File Format	Edge coverage NEUTAINT	Libdft	DFSan	Triton
readelf-2.30	ELF	5540	4164	2489	440†
harfbuzz-1.7.6	TTF	5395	3796	n/a	11†
mupdf-1.12.0	PDF	399	248	192	48†
libxml2-2.9.7	XML	918	428	n/a	236†
libjpeg-9c	JPEG	649	n/a	n/a	n/a
zlib-1.2.11	ZIP	200	131	n/a	54†

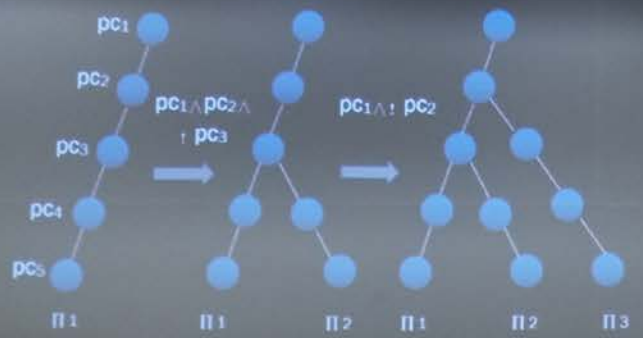
†indicates cases where Triton analyzed partial inputs from dataset.

实验结果

- 与Libdft、Triton、DFSan相比，准确率提升10%，运行时开销降低40倍以上；
- 用于taint-guided fuzzing时，多覆盖61%的分支。


ISC
人工智能在符号执行中的应用
S²E
Mayhem
angr
Kite
Manticore
SAGE
TRITON
Dynamic Binary Analysis
Pex

人工智能在符号执行中的应用



状态空间爆炸问题

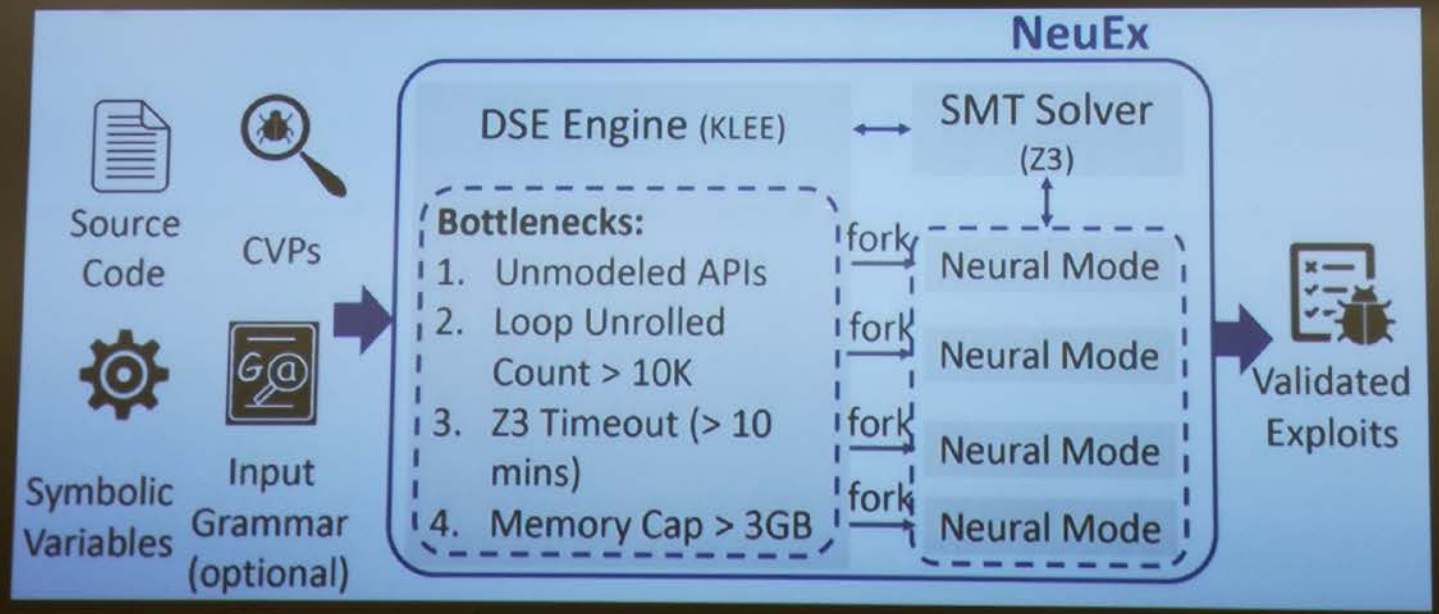
约束求解问题

其他问题：内存建模、环境交互、浮点运算、并行计算

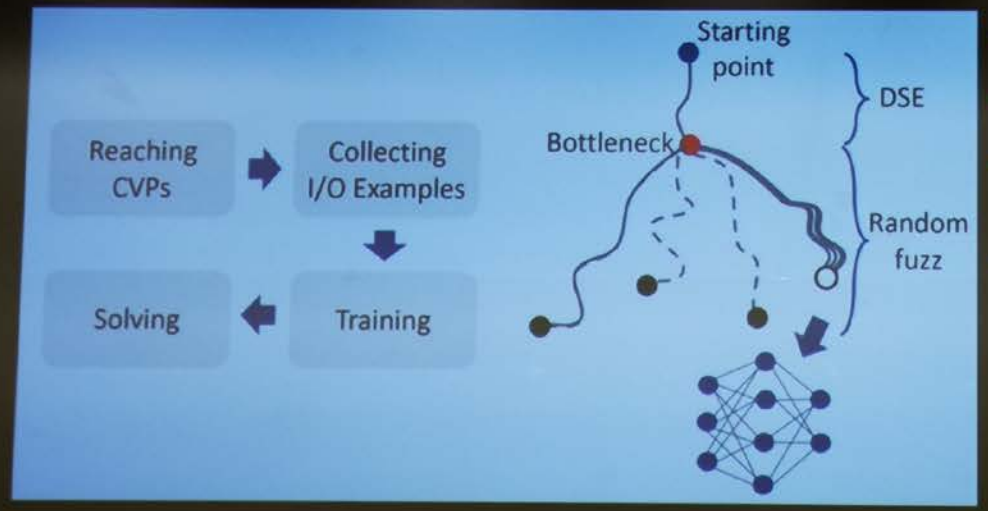
人工智能在符号执行中的应用——NeuEx

NeuEx， NDSS 2019，新加坡国立大学, Prateek Saxena教授团队

- 缓解符号执行中的约束求解问题



人工智能在符号执行中的应用——NeuEx



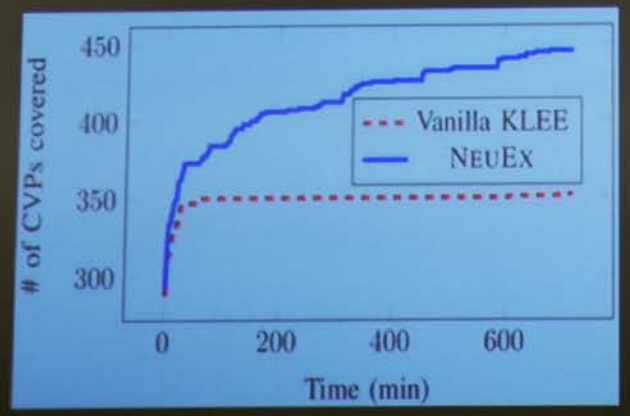
Neural Mode

- 神经网络：MLP
- 激活函数：Relu
- 数据收集：SMT求解+随机变异
- 约束类型：
  - symbolic constraints
  - neural constraints
  - mixed constraints

人工智能在符号执行中的应用——NeuEx

实验结果

- 测试程序: cURL、SQLite、libTIFF、libsndfile、BIND、Sendmail、WuFTP
- 约束复杂性：Complex loops、Floating-point variables、Unmodeled APIs



触发CVP的数量比KLEE多25%

与KLEE的BFS、RAND模式相比，在12个小时内，多发现94%、89%的bug

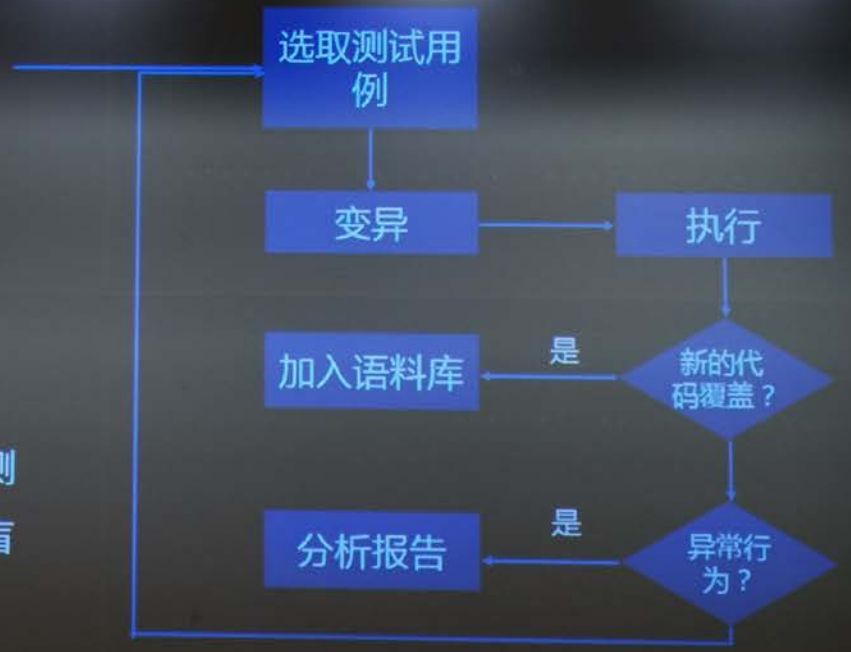
人工智能在模糊测试中的应用

模糊测试（Fuzzing）

- 分类

1) 黑盒、灰盒、白盒
2) 基于生成、基于变异
3) 文件类、网络协议类、内核类

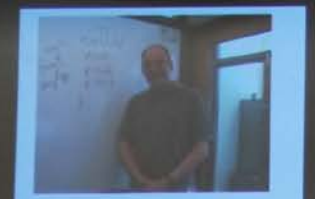
- 优点：可扩展性好
- 缺点：依赖于种子输入的质量、测试冗余、攻击面模糊、测试路径盲目性较高等。



代码覆盖制导的模糊测试流程

人工智能在模糊测试中的应用(1)——Learn&Fuzz

Microsoft Research



Patrice Godefroid

Email: pg AT microsoft.com
Mail: Microsoft Research, One Microsoft Wa

Learn&Fuzz:
Machine Learning for Input Fuzzing

Patrice Godefroid
Microsoft Research, USA
pg@microsoft.com

Hila Peleg
Technion, Israel
hilap@cs.technion.ac.il

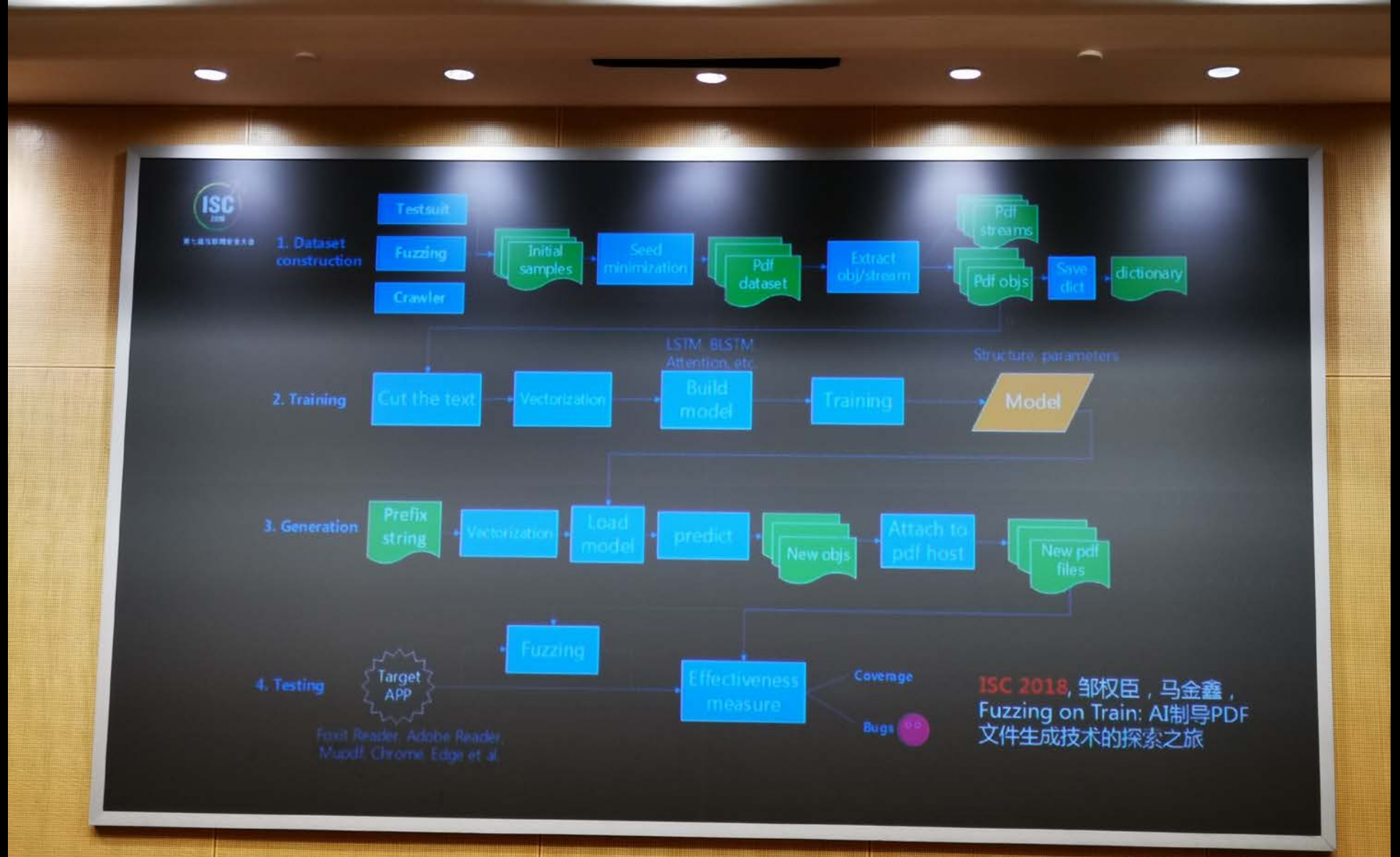
Rishabh Singh
Microsoft Research, USA
risin@microsoft.com

ASE 2017

解决模糊测试中的高结构化样本生成合法性弱或人工参与度高的问题

贡献

- ✓首次把模糊测试中的**高结构化样本**生成问题转换成了**NLP**领域的**文本生成**问题
- ✓采用了**Char-RNN**模型实现对PDF文件中的obj语法的学习
- ✓提出了**基于生成模型**的采样算法**SampleFuzz**，生成obj对象和PDF文件


1. Dataset construction
Testsuit
Fuzzing
Crawler
Initial samples
Seed minimization
Pdf dataset
Extract obj/stream
Pdf streams
Pdf objs
Save dict
dictionary
2. Training
LSTM, BLSTM, Attention, etc
Cut the text
Vectorization
Build model
Training
Structure, parameters
Model
3. Generation
Prefix string
Vectorization
Load model
predict
New objs
Attach to pdf host
New pdf files
4. Testing
Fuzzing
Target APP
Foxit Reader, Adobe Reader, Mupdf, Chrome, Edge et al
Effectiveness measure
Coverage
Bugs
ISC 2018, 邹权臣，马金鑫，
Fuzzing on Train: AI制导PDF
文件生成技术的探索之旅

人工智能在模糊测试中的应用(1)——Learn&Fuzz

	优化环节	我们的工作
1	数据收集	通过**公开测试集**、**模糊测试工具**生成等多种渠道收集数据集
2	学习对象	首次采用了面向**stream obj**语法的学习方案
3	模型设计	首次采用了**BLSTM**、**Attention**等模型
4	采样算法	首次采用了**可动态调整概率分布**的多项分布采样算法
5	样本生成方式	**并行**加载模型生成obj，**并行**生成PDF样本
6	测试和验证	生成**大批量**样本，对**多款**软件进行了测试

实验结果

✓代码覆盖率最高提升约为**Learn&Fuzz**方案和**PinAFL**的**3**倍

✓挖掘出包括foxit reader在内的**6款**PDF阅读器中的bug **164个**，其中可利用漏洞**1个**

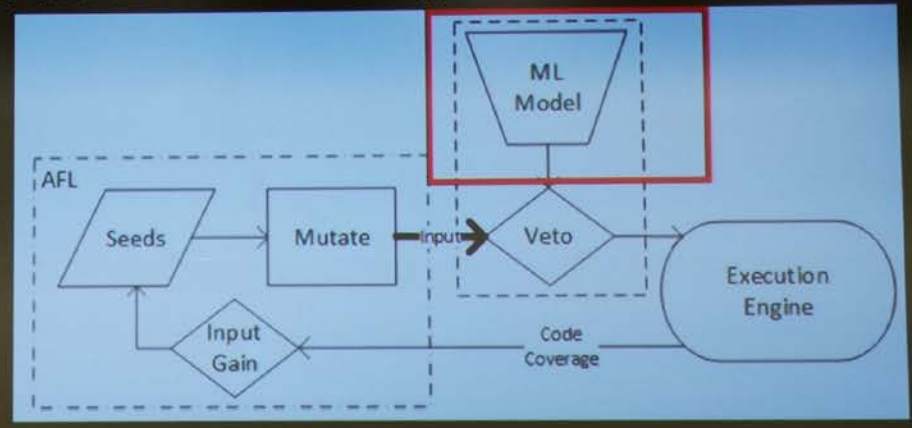
人工智能在模糊测试中的应用(2)——Neural Fuzzing

Neural Fuzzing，2017，arXiv, 美国微软研究院， Mohit Rajpal等


Neural fuzzing: applying DNN to software security testing

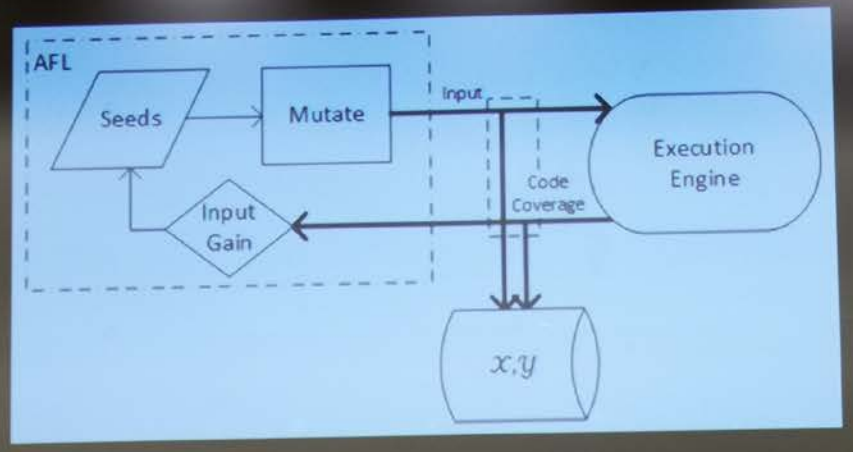
Microsoft researchers have developed a new method for discovering software security vulnerabilities that uses machine learning and deep neural networks to help the system root out bugs better by learning from past experience. This new research project, called neural fuzzing, is designed to augment traditional fuzzing techniques, and early experiments have demonstrated promising results.

Software security testing is a hard task that is traditionally done by security experts through costly and targeted code audits, or by using very specialized and complex security tools to detect and assess vulnerabilities in code. We recently released a tool, called Microsoft Security Risk Detection, that significantly simplifies security testing and does not require you to be an expert in security in order to root out software bugs. The Azure-based tool is available to Windows users and in preview for Linux users.



- 增强模糊测试中样本变异的导向性
- 用神经网络对测试输入的关键字节做标注形成热力图，对变异生成的文件判决

人工智能在模糊测试中的应用(2)——Neural Fuzzing



- 数据收集：input, code coverage pairs
- 神经网络：LSTM、BLSTM、Seq2Seq、Seq2Seq+Attn

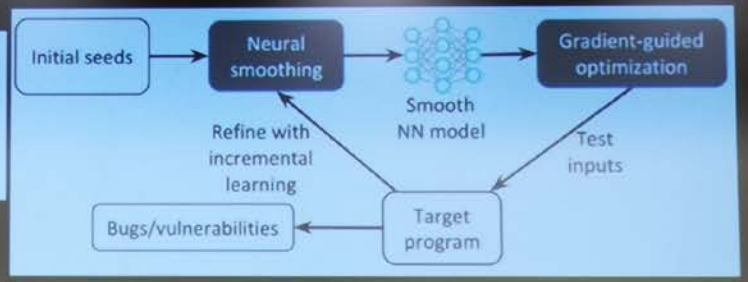
实验结果

✓ 相比于AFL，在readelf, readpng, mupdf, libxml中发现更多路径和crash

人工智能在模糊测试中的应用(3)——Neuzz

NEUZZ: Efficient Fuzzing with Neural Program Smoothing

Dongdong She, Kexin Pei, Dave Epstein, Junfeng Yang, Baishakhi Ray, and Suman Jana
Columbia University



S&P (Oakland) 2019，哥伦比亚大学Suman Jana教授团队

用神经网络学习目标程序中的**分支**行为，然后采用**梯度制导**的方式指导样本生成

实验结果

- 测试程序：mupdf、readelf、libjpeg、libxml、zlib等**10个**常用程序
- 对比工具：AFL、AFLFast、Driller、Vuzzer、KleeFL、Steelix等
- 测试结果：分支覆盖**3倍**以上，并挖掘出更多漏洞

人工智能在漏洞可利用性分析中的应用

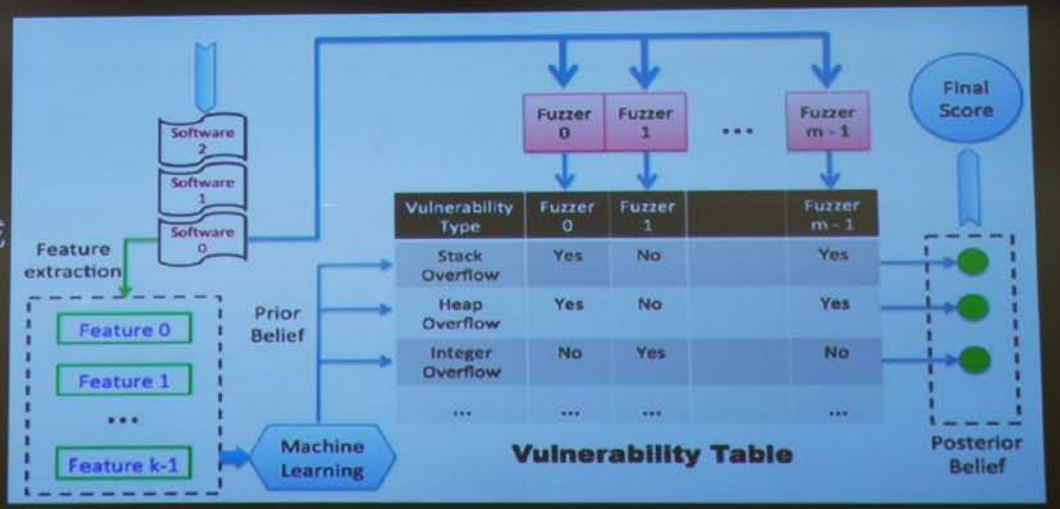
漏洞可利用性分析

- !exploitable、gdb-exploitable、Asan
  - exploitable、probably exploitable、probably not exploitable、unknow
  - 具有误报率高的缺陷，仍依赖专家逆向分析确认，并编写POC
- APEG、 AGE 、Mayhem、 FlowStitch等
  - 限制条件较多
  - 支持漏洞类型有限
  - 难以绕过缓解措施

人工智能在漏洞可利用性分析中的应用——ExploitMeter

量化漏洞可利用性

1) 先验信念：静态分析提取多种特征（hexdump、objdump等），训练分类模型，预测漏洞类型；
2) 后验信念：多种fuzzer生成crash，采用贝叶斯方法计算后验信念；
3) 可利用性评分计算：基于概率论，组合不同类型漏洞的可利用性得分，得到最终评分。



PAC 2017, 纽约州立大学, Guanhua YAN

人工智能在漏洞可利用性分析中的应用——ExploitMeter

实验结果

Test order	Application	Score	E	PE	PNE	U
5	vlc	0.811	1	0	0	0
13	mediainfo	0.937	1	1	2	0
18	qpdfview	0.647	0	1	1	0
19	xpdf.real	0.824	1	0	1	0
22	evince	0.930	1	1	0	1
25	odt2txt	0.806	1	0	0	1
31	objcopy	0.986	2	1	1	3
35	xine	0.994	3	0	2	1
36	jpegtran	0.999	4	1	0	1
39	abiword	1.000	5	3	1	3
40	size	0.995	2	2	1	3
46	catdoc	0.828	1	0	1	2
49	pdfseparate	0.825	1	0	1	0
66	pdftk	0.824	1	0	1	0
67	avplay	0.841	1	0	2	0
74	pdftohtml	0.965	2	0	1	1
76	qpdf	0.961	2	0	0	0
82	ar	0.972	1	2	1	3
91	mpv	0.994	2	2	1	3
100	mencoder	0.989	2	1	3	1

小结

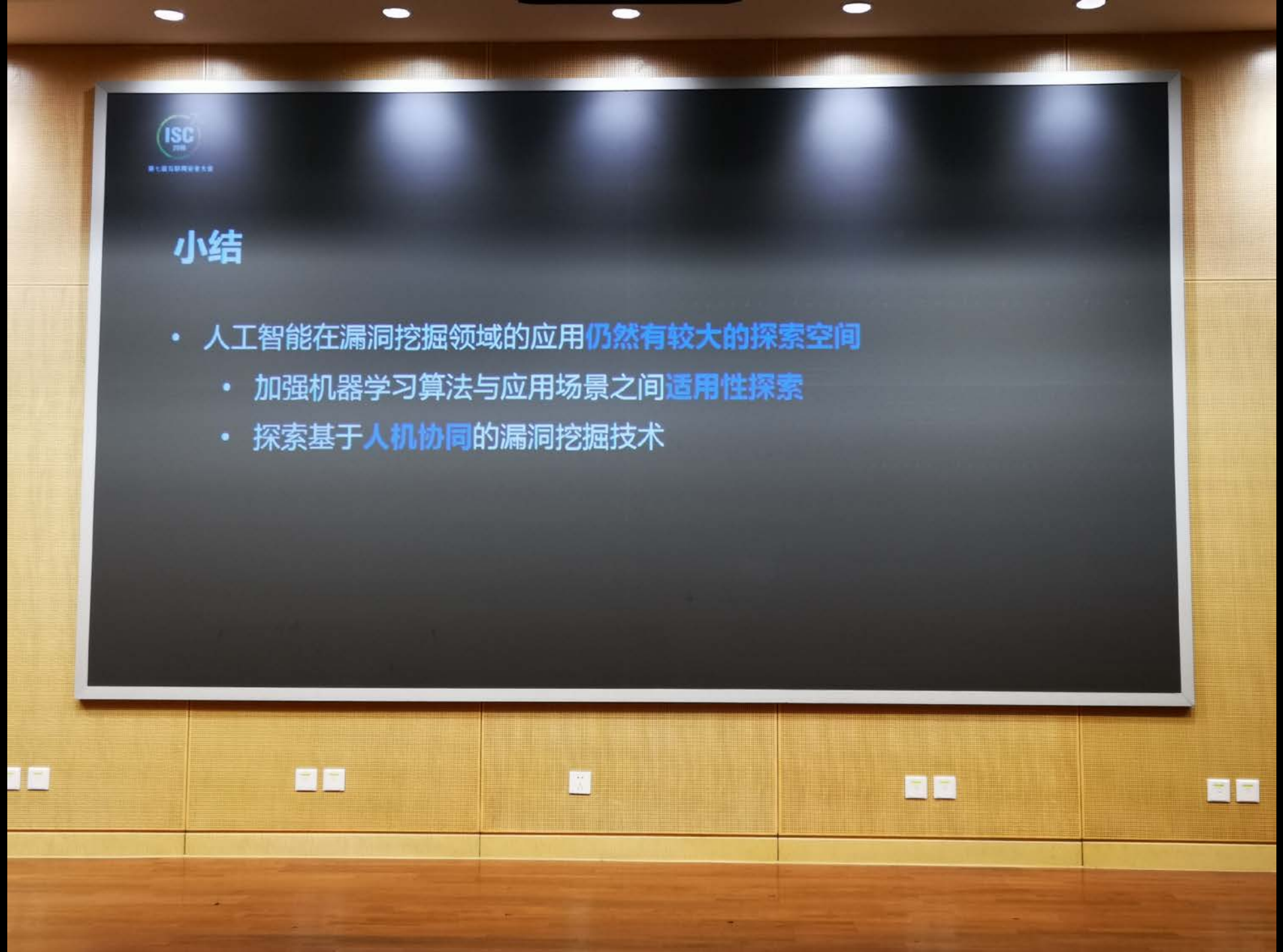
- 人工智能为解决传统漏洞挖掘技术的**瓶颈问题**提供了**新的思路**
- 基于人工智能的漏洞挖掘技术本身也面临着一系列的**挑战**：

人工智能

- 算法不够健壮安全
- 不可解释性
- 模型场景化
- 数据饥饿

基于人工智能的漏洞挖掘

- 算法选择依赖经验
- 特征选择依赖专家知识
- 漏洞数据收集困难


ISC
小结
• 人工智能在漏洞挖掘领域的应用仍然有较大的探索空间
• 加强机器学习算法与应用场景之间适用性探索
• 探索基于人机协同的漏洞挖掘技术

小鹅助理

谢谢！

扫码添加小鹅助理，与数万科技圈人士

分享重量级活动PPT、干货培训课程、高端会议免费门票