KCon

2019

启明星辰ADLab

智能语音设备安全研究

演讲人：王启泽

目录
CONTENTS

PART 01
背景

语音正成为人与设备交互的方式

语音包括相当丰富的信息

- 性别.年龄
- 环境.健康
- 想法.情绪
- 籍贯

我们研究的对象

我们关注的点

安全及隐私

PART 02
网络安全

音箱网络架构

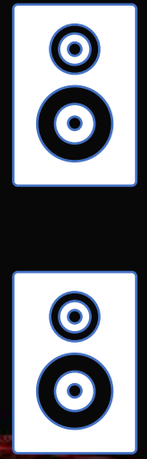
云平台的特权命令

智能音箱攻击演示-特权命令

智能音箱攻击演示-设备间通信

多音箱场景

当房间里存在多个音箱时，
多个音箱之间需要协商，
决定由哪个音箱来响应用户的语音命令



多音箱场景

攻击者（IP：192.168.1.3）

伪造语音广播
（能量：99.00,
IP：192.168.1.2）

能量列表里最大的IP

接受语音命令

目标音箱（IP：192.168.1.2）

设备伪造

智能音箱

设备广播

访问设备网址，确认设备是目标设备

伪电视

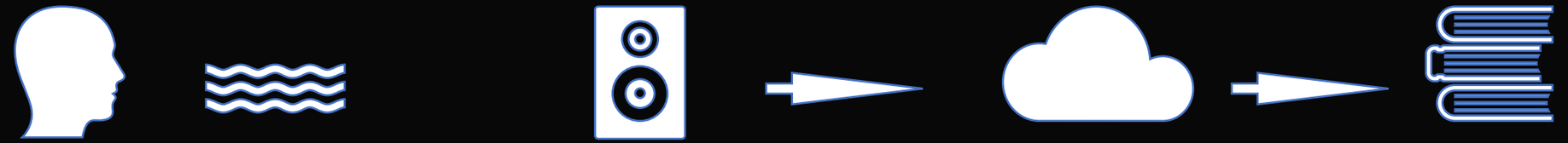
设备伪造

漏洞列表

漏洞编号	危害
CNVD-2019-13611、CNVD-2019-06254 CNVD-2019-05625、CNVD-2019-05626	远程命令执行、远程代码执行
CNVD-2019-07688、CNVD-2019-15526	播放恶意音频
CNVD-2019-12111、CNVD-2019-13278	敏感信息泄漏
CNVD-2019-12775	语音窃听

PART 03
语音及内容安全

语音安全



声音传播　　语音唤醒/声纹识别　　语音识别/语义理解　　内容

超声波攻击

超声波

识别

次声波 | 人耳可听声 | 超声波

20hz 20Khz

听不到

超声波攻击-演示

1. 笔记本电脑喇叭性能的提升使得攻击者无需额外硬件即可发出超声波信号。

2. 新型的智能手机依然存在问题

语音识别攻击-基于发音模型的攻击

1.基于韵母的攻击

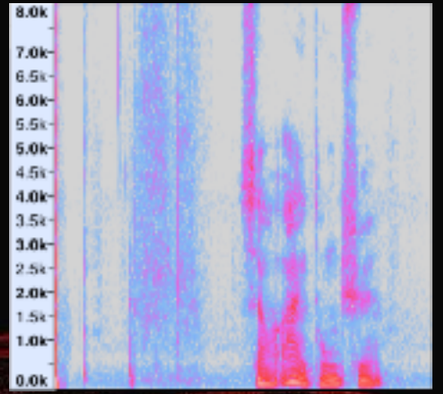
每个汉字的发音都是由声母、韵母两部分构成的。
声母部分发音时间短,信号变化剧烈;
而韵母部分发音时间长，是声带共鸣产生，携带了音节的大部分能量。
韵母是由元音或元音加辅音组成。

某唤醒词算法主要根据元音来判断

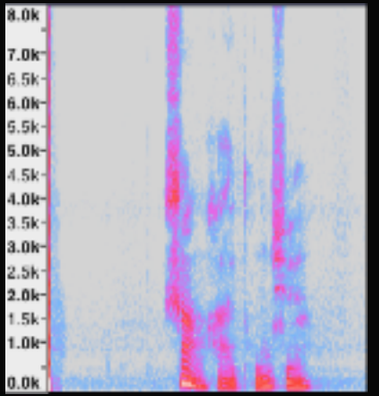
唤醒词识别攻击示意

1.上海同学(shang hai tong xue)

2.小爱同学(xiao ai tong xue)



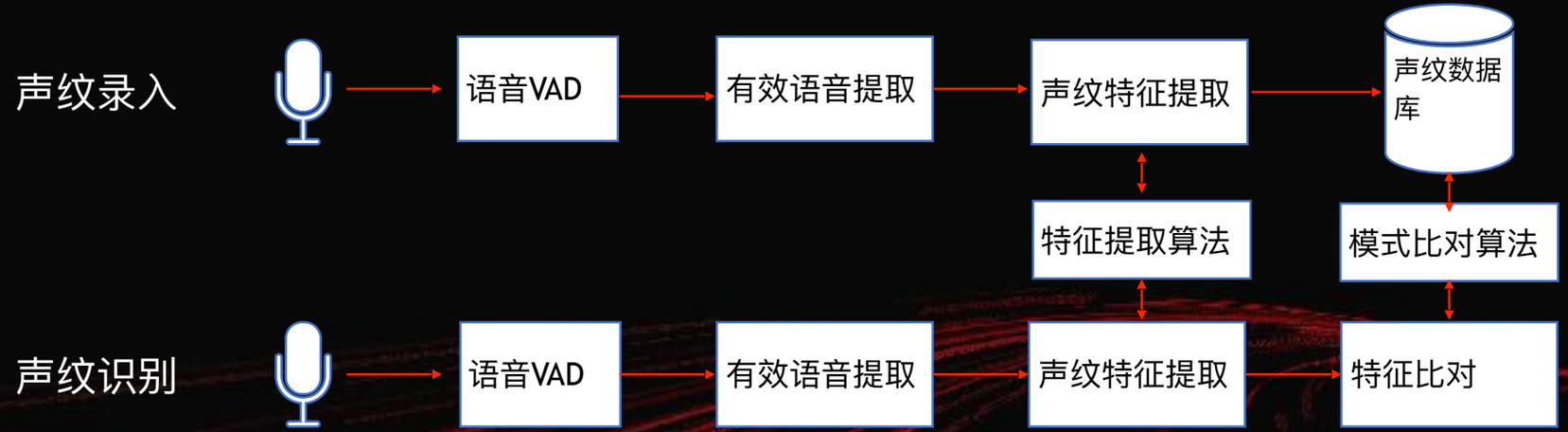
小爱同学



上海同学

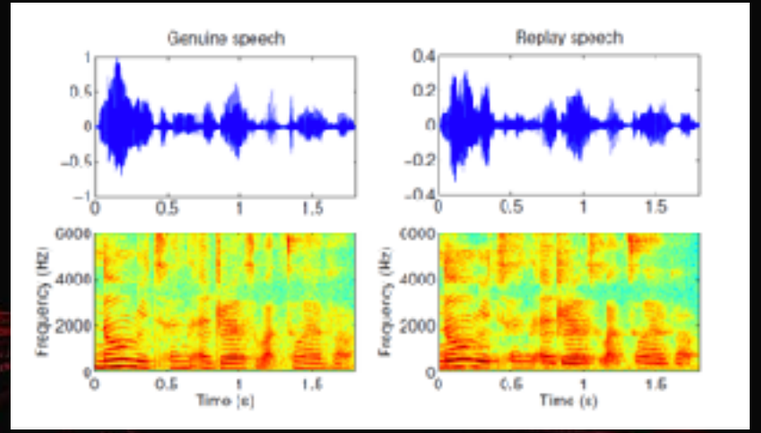
唤醒词识别攻击-演示

声纹识别


声纹录入
语音VAD
有效语音提取
声纹特征提取
声纹数据库
特征提取算法
模式比对算法
声纹识别
语音VAD
有效语音提取
声纹特征提取
特征比对

常见的声纹算法攻击

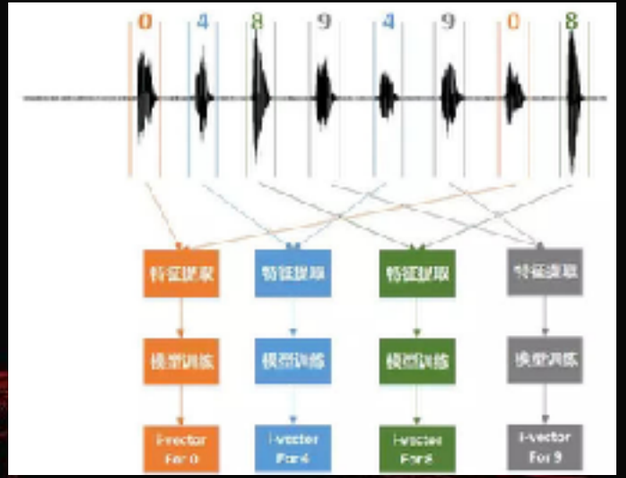
- 拼接合成攻击
- 样本攻击
- 持续语速变化攻击
- 端到端攻击
- 录音攻击



声纹识别攻击-声纹比对算法

(D1+D2+D3+...D8)/8

(80+80+80+40+40+40+40+40)/8=55

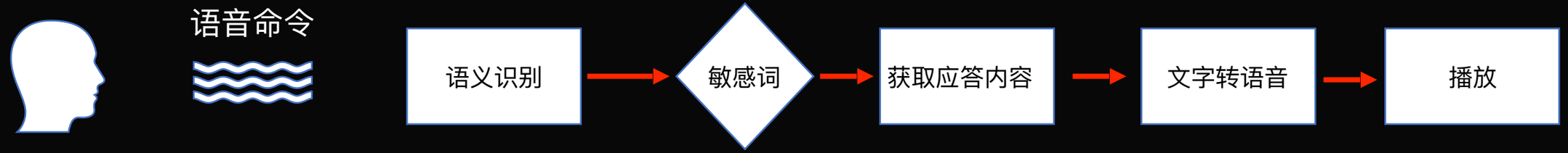


声音识别算法攻击-声纹比对算法

- 声纹识别没有错误次数限制
- 持续变化语速导致评分标准浮动大
- 中性的声音得分较高

声纹识别算法-攻击演示

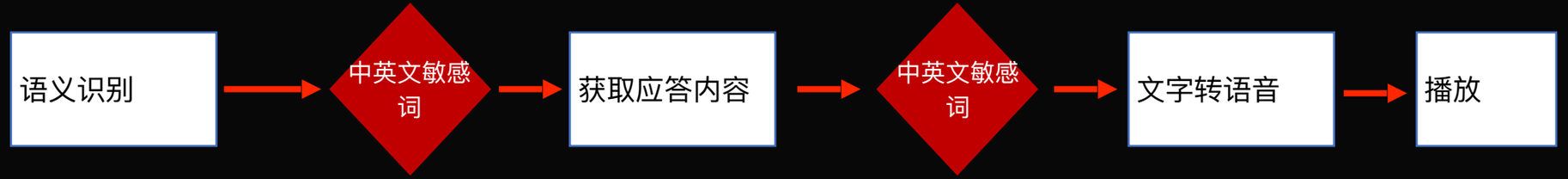
内容安全



内容安全-攻击演示

内容安全

语音命令



PART 04
隐私审计

隐私权

权利	内容
用户有权力决定	哪些他本人的信息可以被收集，什么时候收集、什么地点收集。
用户有权力了解和决定	这些数据是如何被收集的，这些数据将对谁共享，为什么要对他人分享，如何对他人分享。

语音设备使用场景的隐私泄露

阶段	敏感信息
设备注册阶段	地理位置，周边WIfi信息，路由器MAC地址，WI-FI密码，已安装应用情况，短/彩信
设备使用阶段	音频信息，通信录，设备使用情况，业务的使用情况，日志， 音视频信息标号

几种常见的隐私风险

问题点	危害
误唤醒	泄漏通话内容
APP权限	泄漏通信录等敏感信息
日志收集	泄漏Wifi密码等信息
API接口	合作厂家可以获得非业务需要的敏感信息
明文通信	泄漏用户账户及密码等信息

总结

物物间通信场景更多：　需要关注设备间认证的安全

声纹识别算法：　大多商用算法还不成熟

公众对隐私的关注度越来越高：　需要关注设备及数据的隐私保护

感谢

ADLab小伙伴
KCon组委会

谢谢观看

演讲人：王启泽
Email：wang_qize@venustech.com.cn