物联网安全
Internet of Things Security

第五章 传感器与执行器安全

冀晓宇
浙江大学

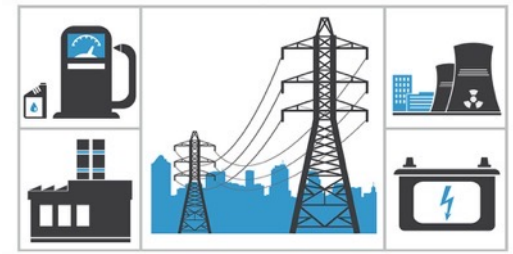
USLAB

目录

USSLAB

5.1 传感器与执行器安全

- 什么是传感器和执行器?
- 传感器和执行器与物联网有什么关系?
- 传感器和执行器安全吗?



Q1：这些设备包含哪些传感器和执行器，你能列举其他的吗?
Q2：一个汽车有多少个传感器?

5.1 传感器、执行器与物联网安全

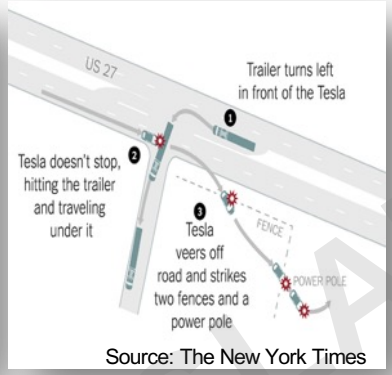
- 传感器在物联网中的作用相当于人体的感觉器官，执行器相当于人体的四肢。

- 传感器、执行器应用于智能制造、车辆控制、海洋探测、环境保护、资源调查、医学诊断、生物工程等领域。

- 随着物联网的快速发展，传感器、执行器技术飞速发展，但其存在的安全问题也不容忽视。

USSLAB

案例：传感器和执行器安全



2016年5月7日，全球第一起自动驾驶的交通事故，美国佛罗里达。

2018年3月5 日，全球首例“自动驾驶”致死车祸发生在中国！新华网报道

2020年6月1日，台湾特斯拉Model 3撞车事故

USSLAB

案例：传感器和执行器安全（续）

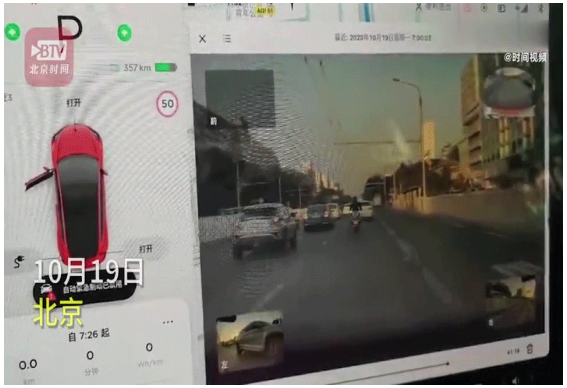
2020年6月16日，江西南昌一辆特斯拉突然从50~60km/h**自动提速**至127km/h，发生撞击事故后起火。

2021年1月，山东，特斯拉新能源汽车**“失控”**撞下了地库的洗手台。

2020年12月，杭州，电动汽车Model 3**失控加速**冲入酒店大堂

2022年10月13日浙江杭州浙大新桥门电车**失速冲撞行人和电动车。**

2022年2月12日，浙江金华一汽车**追尾后失控加速“狂奔”2公里，先后撞击5辆车**



2020年10月19日，北京一特斯拉Model 3在未开启自动辅助驾驶功能的情况下**突然无故向右偏离**

5.2 传感器与执行器安全

- 传感器的定义
- 传感器的组成
- 传感器的分类
- 传感器的特性

USSLAB

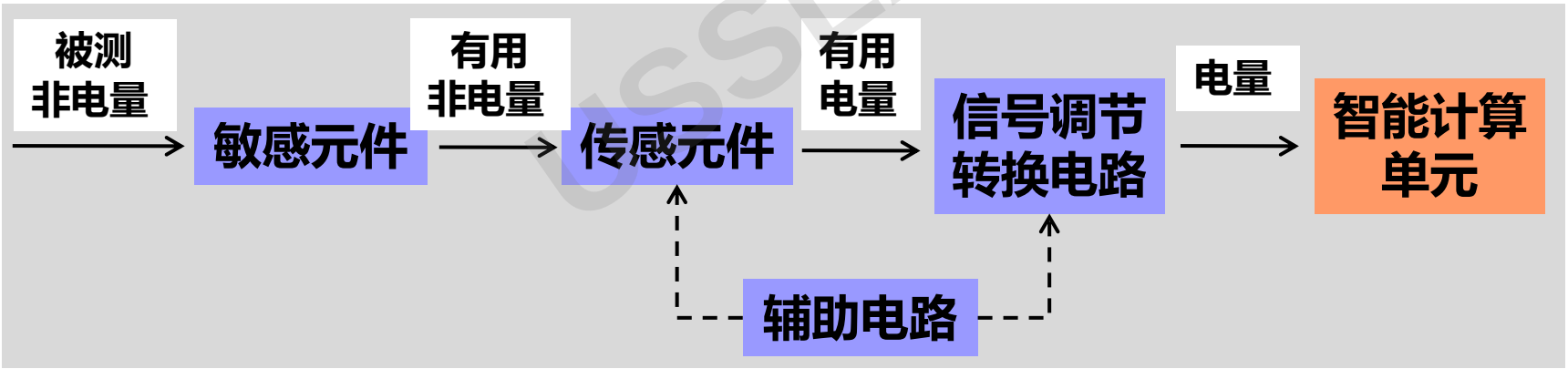
5.2.1 传感器的定义

- **国家标准GB7665-87**定义：能感受规定的被测量并按照一定的规律转换成可用信号的器件或装置，通常由敏感元件和转换元件组成。
- **本书定义**：将被测非电量信号转换为与之有确定对应关系电量输出的器件或装置，也叫**变换器、换能器**。
- 传感器是一种检测装置，能感受到被测量的信息，并能将检测感受到的信息，按一定规律变换成为电信号或其他所需形式的信息输出，以满足信息的传输、处理、存储、显示、记录和控制等要求。
- 被测量的信息可以是光、热、运动、湿气、压力，或其他环境现象中的任何一种。
- 举例：麦克风、加速度计、磁力计、温度计、还有哪些常见的传感器
- 发展趋势：智能传感器，能够对感知信号进行AI处理并输出处理结果

USTAB

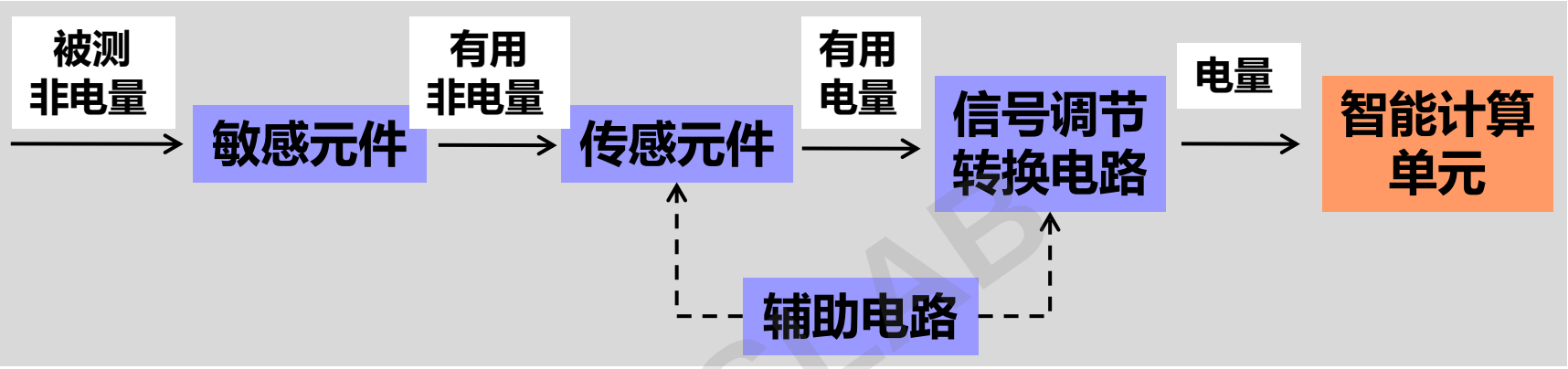
5.2.2 传感器的组成

- 传感器一般由**敏感器件、传感元件、转换电路**和**辅助电路**四部分组成。
- 广义的智能传感器额外具备**智能计算单元**。



USSLAB

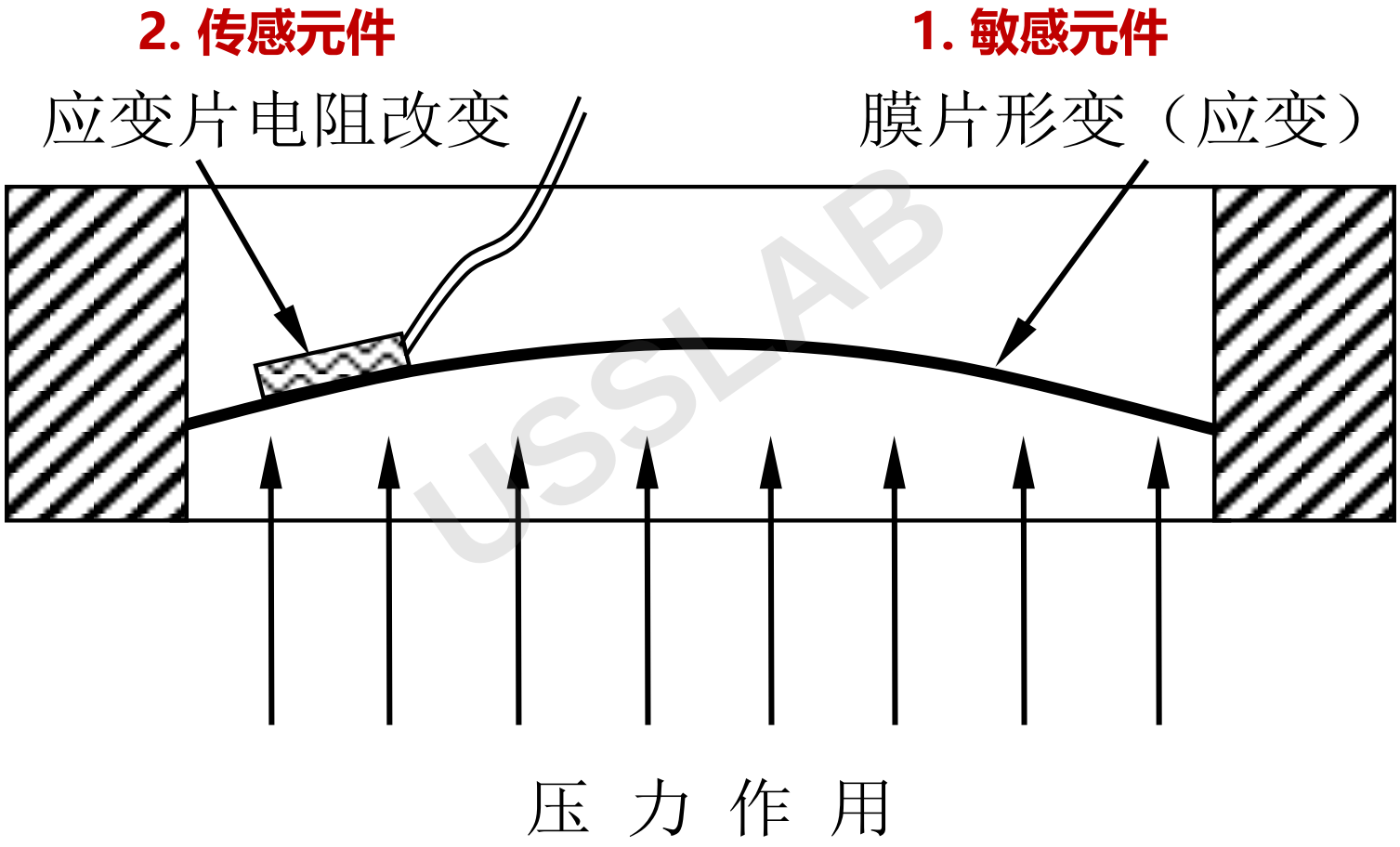
5.2.2 传感器的组成



- **敏感元件**：直接感受被测非电量并按一定规律转换成与被测量有确定关系的其它非电量的元件。
- **传感元件**：又称变换器。能将敏感元件感受到的非电量直接转换成电量的器件，如电阻、电感、电容、或者电流电压等。
- **信号调节与转换电路**：能把传感元件输出的电信号转换为便于显示、记录、处理、和控制的有用电信号的电路。
- **智能计算单元**：执行数字域信号处理、AI算法的计算单元
- 辅助电路：通常包括电源等。

USSLAB

5.2.2 传感器的组成——举例



应变式压力传感器

5.2.3 传感器的分类

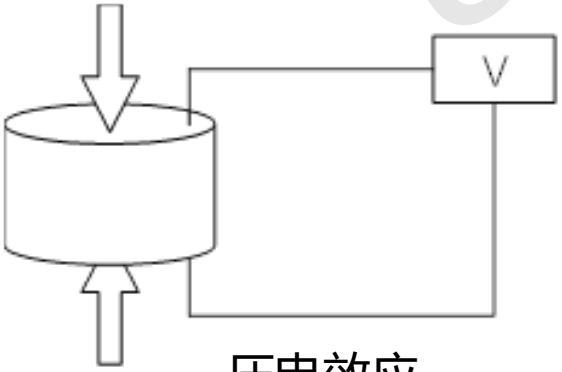
由于传感器应用领域众多，适用范围又广，其品种和规格繁多，根据不同的原则可以将传感器分成不同类型。比较常用的分类如下：

- **按工作原理分类**
  - 物理传感器和化学传感器
- **按构成原理分类**
  - 结构型传感器和物性传感器
- **按能量转换情况分类**
  - 能量控制型传感器和能量转换型传感器
- **按输出信号分类**
  - 模拟传感器、数字传感器、开关传感器

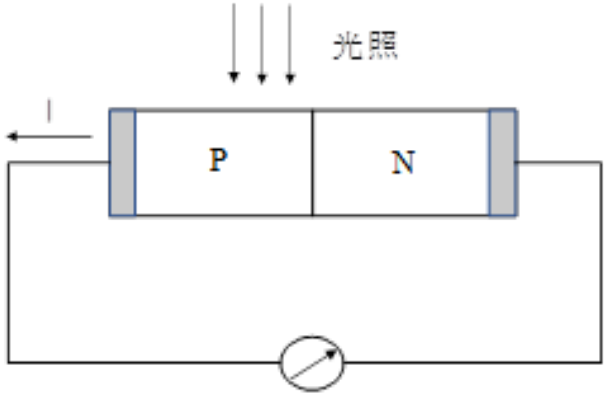
JSSLAB

5.2.3 传感器的分类——工作原理分类

- **物理传感器**
  - 物理传感器可以按其工作原理的物理效应进行分类，诸如**压电效应、磁致伸缩现象、离化、极化、热电、光电、磁电**等效应。
  - 被测信号量的微小变化都将转换成电信号。
  - 大多数传感器是物理传感器。
  - 物理传感器的物理效应大部分存在逆效应，例如逆压电、逆光电等。



压电效应



光电效应

Q：你能列举一些物理传感器逆效应相关的安全问题吗？

USSLAB

5.2.3 传感器的分类——工作原理分类

- **化学传感器**
  - 化学传感器包括以化学吸附、电化学反应等现象为因果关系的传感器，被测信号量的微小变化也将转换成电信号。
  - 化学传感器技术问题较多，例如可靠性问题、规模生产可能性、价格问题等。

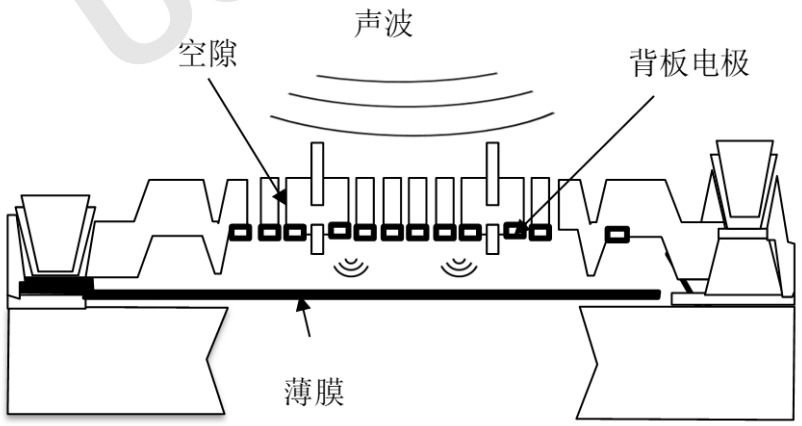
NO_2气体传感器

- 有些传感器既不能划分到物理类，也不能划分为化学类，例如生物传感器，如DNA传感器

USSLAB

5.2.3 传感器的分类——构成原理分类

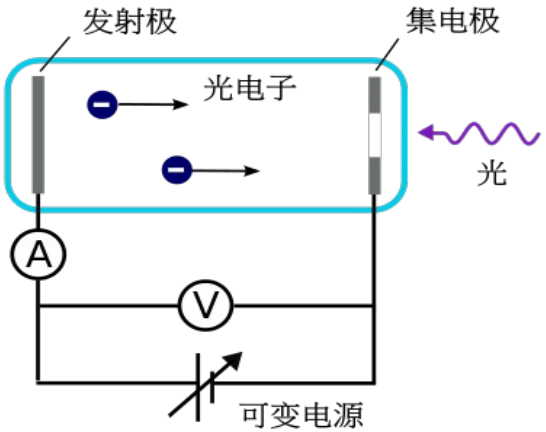
- **结构型传感器：**结构型传感器是**基于物理学中场的定律**构成的，包括动力场的**运动定律**、电磁场的**电磁定律**等。
- 特点：传感器的工作原理是以传感器中元件**相对位置变化**引起场的变化为基础。
- 物理学中的定律一般是以方程式给出，对于传感器来说，这些方程式也就是许多传感器在工作时的数学模型。



MEMS麦克风结构图，薄膜位置随声音而变化

5.2.3 传感器的分类——构成原理分类

- **物性型传感器**：物性型传感器是**基于物质定律**构成的，如**虎克定律、欧姆定律**等，而不发生结构变化。
- 物质定律是表示物质某种客观性质的法则。因此物性型传感器的性能随材料的不同而不同。如光电管就是物性型传感器，它利用了物质法则中的外**光电效应**。
- 物质法则大多数是以物质本身的常数形式给出。这些常数的大小决定了传感器的主要性能。



光电效应

5.2.3 传感器的分类——能量转换分类

- **能量控制型传感器**
  - 能量控制型传感器，通过外部能量输入改变自身电参数，输出信号为阻抗/电容/电感变化量等
  - 如**电阻应变片、热敏电阻、光敏电阻**等传感器

- **能量转换型传感器**
  - 能量转换型传感器，将待测能量直接转换为电信号，输出信号为电压/电荷/电流
  - 自发电特性，无需外部电源
  - 如热电偶将二种金属间的温差直接转换为微小电压

USSLAB

5.2.3 传感器的分类——输出信号分类

- **模拟传感器**
  - 将被测量的非电学量转换成**模拟电信号**。
- **数字传感器**
  - 数字传感器输出信号为数字量(或数字编码)的传感器，一般将传统的模拟式传感器经过加装或改造A/D转换模块，使之**输出信号为数字量(或数字编码)的传感器。**
  - 包括：放大器、A/D转换器、微处理器（CPU）、存储器、通讯接口电路等。
- **开关传感器**
  - 当一个被测量的信号达到某个特定的阈值时，传感器相应地输出一个设定的低电平或高电平信号。

Q：模拟传感器和数字传感器的攻击难易程度如何？

USSLAB

5.2.3 传感器的分类——测量方法分类

- **主动型传感器**
  - 传感器测量过程需要主动向被测对象发射信号并检测其反射信号。发射信号包括声波、电磁波、激光等
  - 举例：超声波避障传感器、激光雷达、毫米波雷达等

- **被动型传感器**
  - 仅通过接收被测量目标发出的信号完成测量过程
  - 举例：摄像头、麦克风、加速度计等

USSLAB

5.2.4 传感器的主要特性——静态特性

- **定义：**

传感器的**静态特性**是指对静态的输入信号，传感器的输出量与输入量之间所具有相互关系。

输入量和输出量都和时间无关，所以它们之间的关系，即传感器的静态特性**可用一个不含时间变量的代数方程**，或以输入量作横坐标、把与其对应的输出量作纵坐标而画出的特性曲线来描述。

- **主要参数：**

表征传感器静态特性的主要参数有：**线性度**、灵敏度、迟滞、重复性、漂移等。

USST

5.2.4 传感器的主要特性——静态特性

- **特性一：线性度**

  传感器的线性度是指其输出量与输入量之间的关系曲线偏离理想直线的程度，又称为**非线性误差**。在不考虑迟滞、蠕变等因素的情况下，其静态特性可用下列多项式代数方程来表示：

$$y = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \qquad (5.1)$$

  其中，y-输出量，x-输入量，a_0-零点输出，a_1-理论灵敏度，a_2，$a_3 \ldots a_n$-非线性项系数。

  由于存在二次、高次项，所以传感器的**输出通常是非线性的**

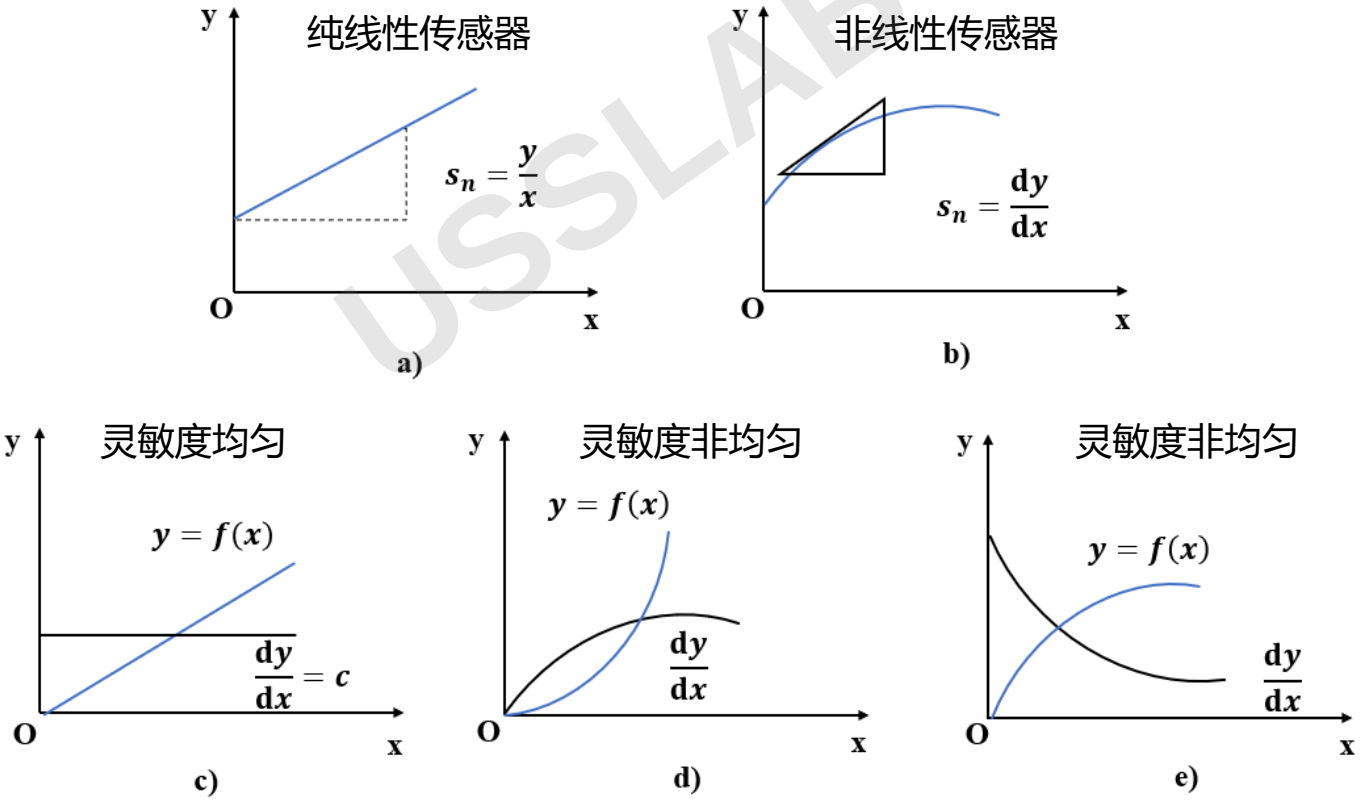
- **安全问题：**非线性特性导致传感器存在输入侧信号注入安全风险，在后续内容中，会对此进行详细介绍（海豚音攻击）。
- 引申阅读：PNN & PFM

USSLAB

5.2.4 传感器的主要特性——静态特性

- **特性二：灵敏度**

灵敏度定义为传感器在稳态信号作用下输出量变化对输入量变化的比值，用S表示灵敏度。



传感器的灵敏度

USSLAB

5.2.4 传感器的主要特性——静态特性

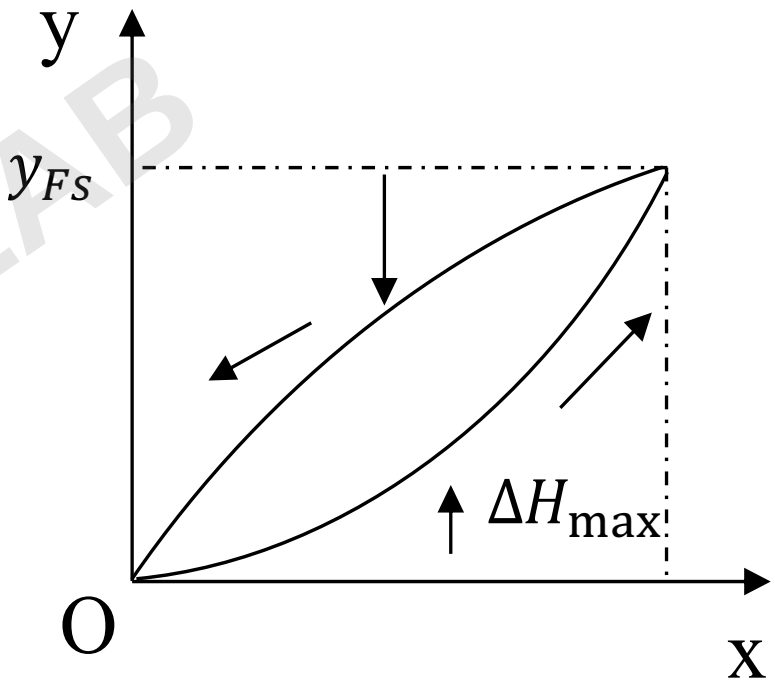
- **特性三：迟滞**

传感器在输入量由小到大（正行程）及输入量由大到小（反行程）变化期间其输入输出特性曲线不重合的现象成为迟滞。

传感器正反行程的输出信号大小差值称为迟滞差值ΔH_max，一般以满量程输出的百分数表示。

原因：机械及结构材料的弱点，如摩擦、松动等。

$$\gamma_H = \frac{\Delta H_{max}}{y_{FS}} * 100\ \% \qquad (5.2)$$



传感器迟滞特性

ESSLAB

5.2.4 传感器的主要特性——静态特性

- **特性四：重复性**

重复误差表征的是传感器在输入按同一方向作全量程连续多次变动时所得特性曲线不一致的程度。其反映的是测量结果偶然误差大小，而并不表示与真实值的差别。

重复误差：$$\gamma_R = \frac{\Delta R_{max}}{y_{FS}} * 100\ \% \qquad (5.3)$$

- **特性五：漂移**

传感器的漂移是指在输入量不变的情况下，传感器输出量**随着时间变化**，此现象称为漂移。

产生漂移的原因有两个方面：一是传感器**自身结构参数**；二是**周围环境**（如温度、湿度等）。

$$温漂 = \frac{\Delta_{max}}{y_{FS}\,\Delta T} * 100\ \% \qquad (5.4)$$

USSLAB

5.2.4 传感器的主要特性——静态特性

- **特性六：分辨率**
  - 分辨率是指传感器能够感知或检测到的**最小输入信号增量**。
  - 分辨率可以用绝对值或与满量程的百分比来表示。
  - 分辨率高是精度高的必要条件而非充分条件。

- **特性七：阈值**

  当传感器的输入从零开始缓慢增加时，在达到某一值后**输出发生可观测**的变化，这个输入值称传感器的阈值电压。

USSLAB

USSLAB

引入案例：海豚音攻击

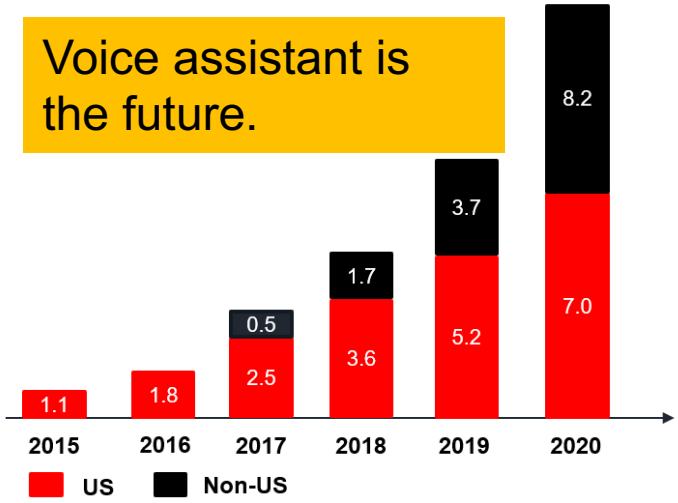
传感器安全案例——海豚音攻击

USSLAB

Siri Google Now Alexa Cortana S Voice Hi Voice

Digital Voice-activated Assistant Device Shipments Worldwide, US vs. Non-us, 2015-2020

millions



Source: Strategy Analytics as cited in press release, Aug,26,2016

智能语音系统安全

语音输入

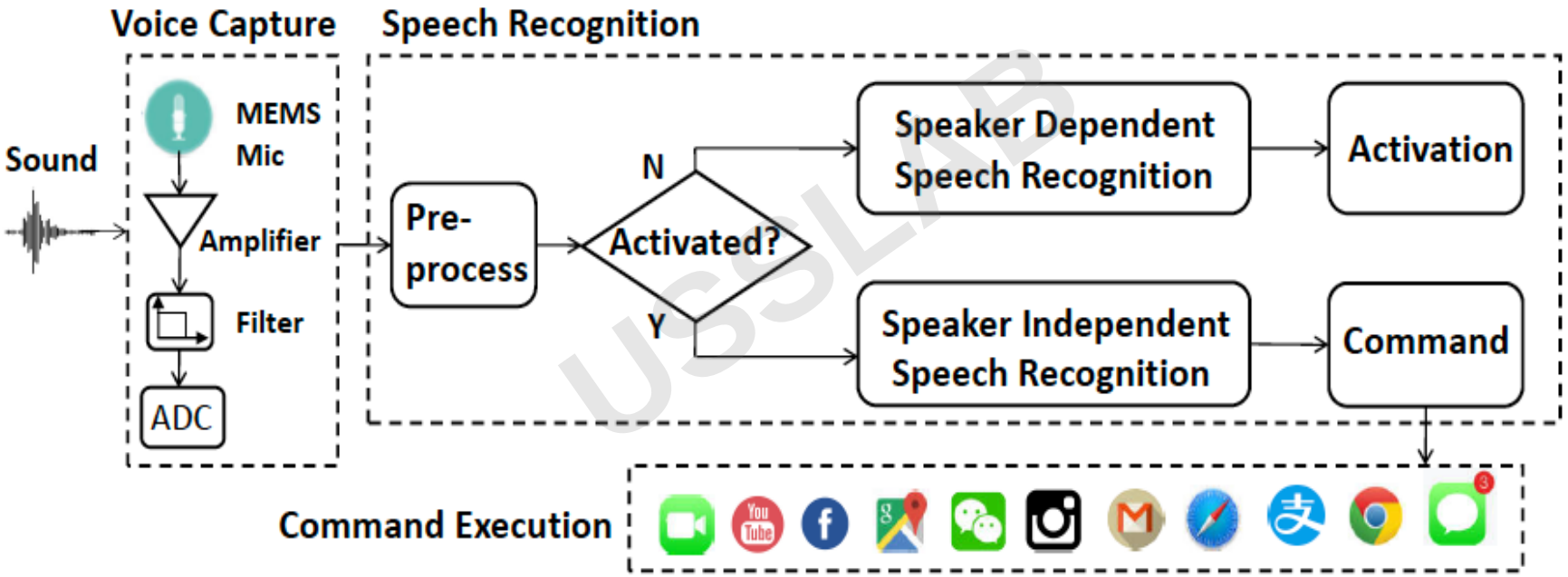
感知

指令执行

实际应用

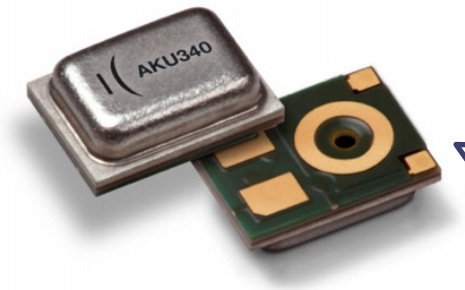
USSLAB

海豚音攻击：智能语音助手

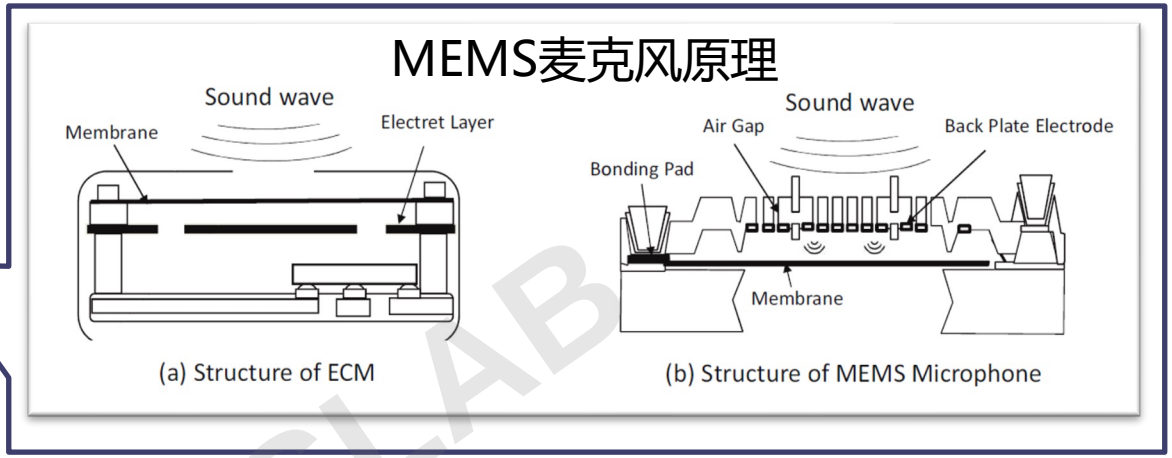


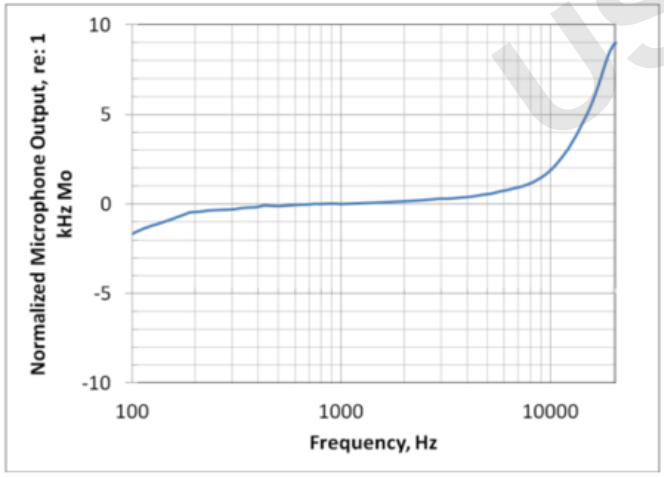
如何针对智能语音助手实现不可听的语音注入攻击呢?

海豚音攻击原理——麦克风传感器

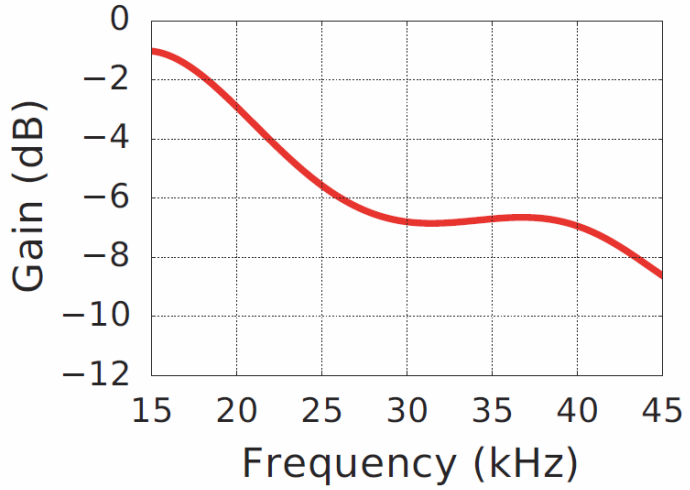


MEMS麦克风





麦克风频率响应（100-10kHz）

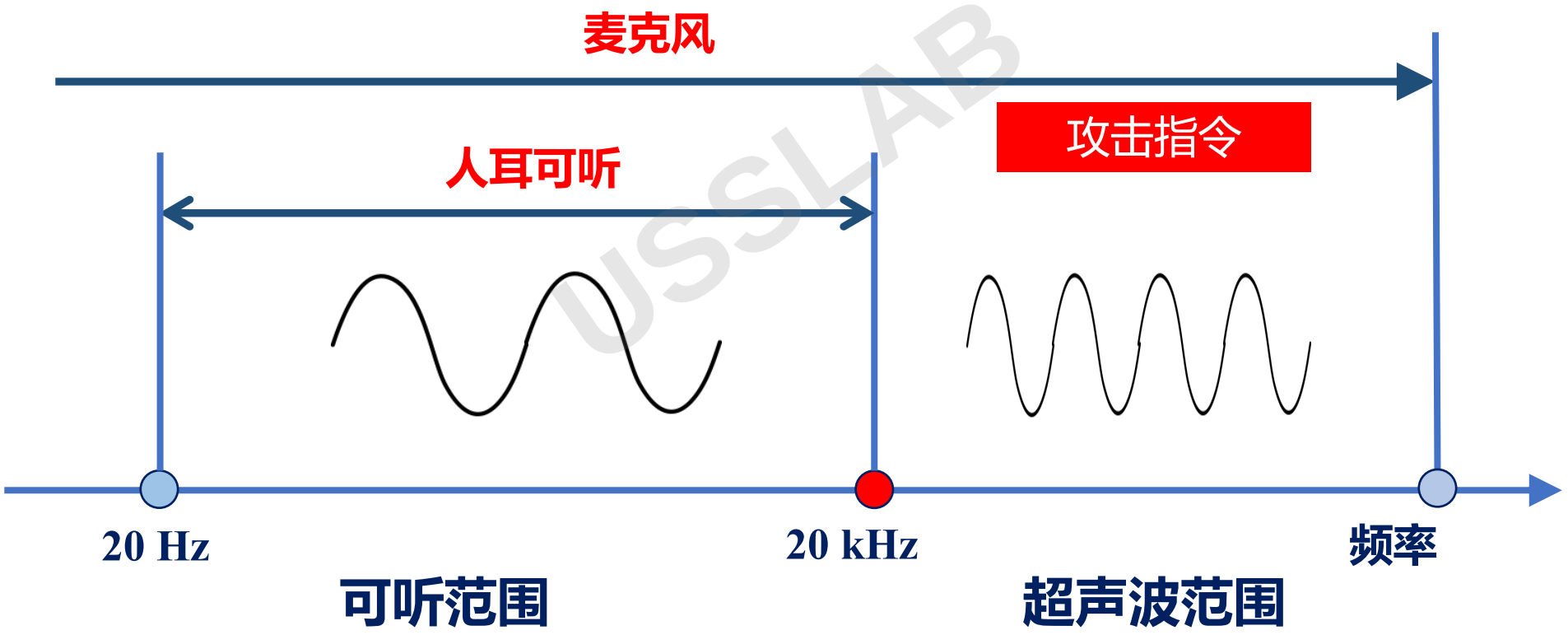


麦克风频率响应（>15kHz）

USSLAB

海豚音攻击原理——超声波频段

- 利用人耳听不到的超声波频段进行攻击！



USSLAB

海豚音攻击原理

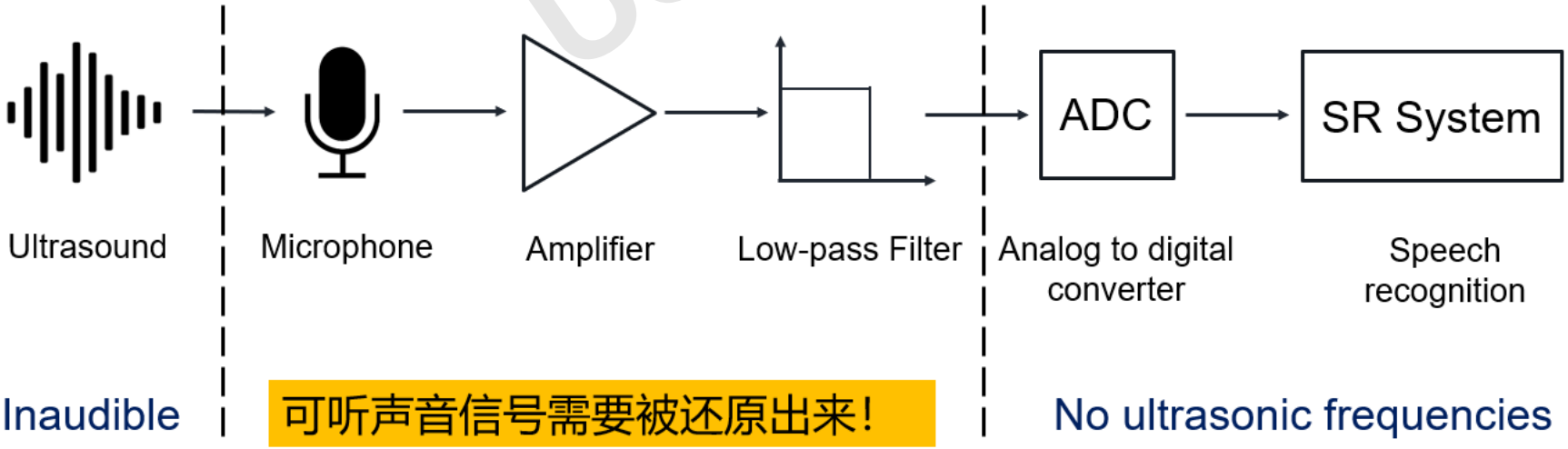
Q1: 超声波信号被低通滤波器滤掉，怎么办?

- 答案：滤波器非完美滤波特性，导致过渡带内的信号不会被滤掉。

Q2：如何实现攻击指令还原?

- 答案：利用调制方法，经过调制的基带信号，会经过麦克风和放大器的非线性作用还原出来！(How?)

USS LAB



补充知识

USSLAB

调制（Modulation）

- 调制：将目标信号进行处理并加到载波上，使其变为适合于信道传输的形式，通常用于无线电波的传播与通信、利用电话线的数据通信等各方面
- 其中，目标信号称为基带信号，载体信号称为载波信号
- 调制方式：
  - 幅度调制：AM
  - 相位调制：PM
  - 频率调制：FM

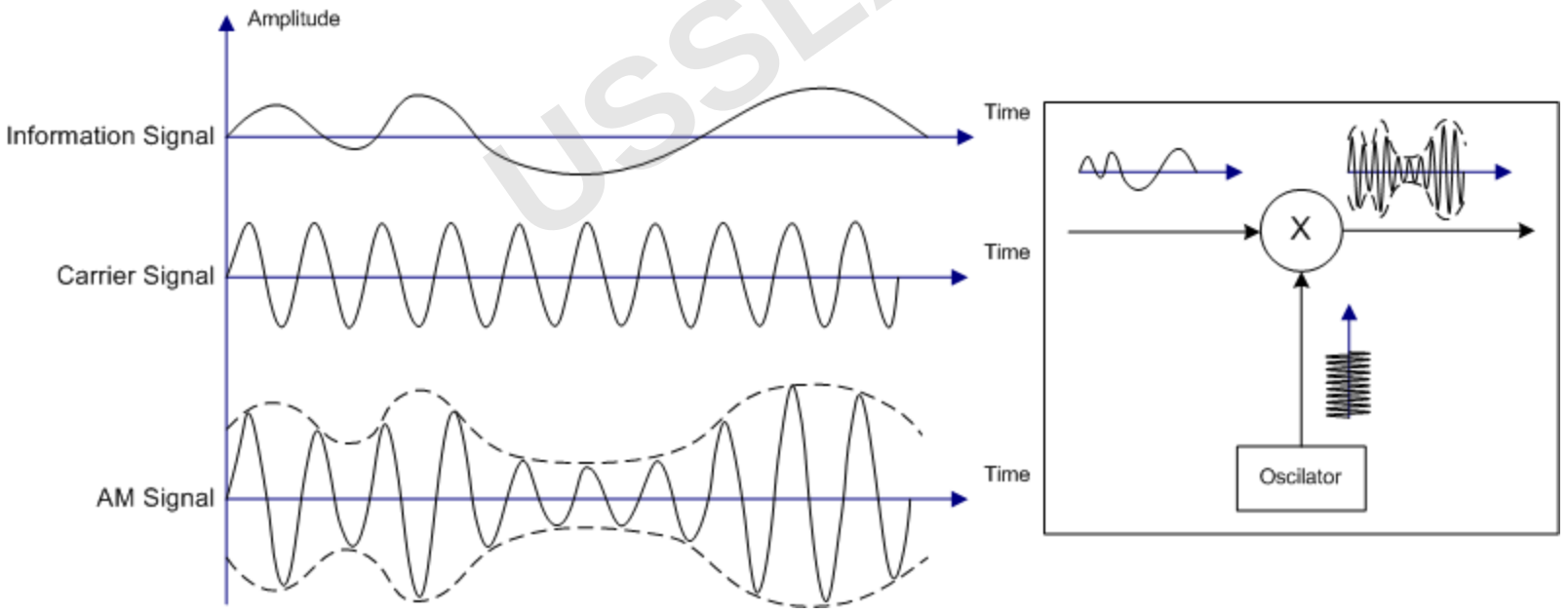
USSLAB

基础知识——调制方式

- **调幅-Amplitude Modulation (AM)**

  载波的幅度（信号强度）是与所发送的基带信号波形成比例变化，调制信号是载波信号的**包络线**。

  图：上为**基带**信号、中为**载波**信号，下为**调制后的信号**。



USSLAB

AM调制原理

- 考虑一个频率为 f_c，幅度为 A 的载波（正弦波）：

$$c(t) = A \cdot \sin(2\pi f_c t)\,.$$

- 令 $m(t)$ 表示基带信号的波形

$$m(t) = M \cdot \cos(2\pi f_m t + \phi)\,,$$

  - 其中 M 是调制的**幅度（深度）**， ***0<M<1 从而使 (1+m(t)) 总是正数。***

- 幅度调制就是用载波 $c(t)$ 乘以 $(1+m(t))$

$$\begin{aligned} y(t) &= [1 + m(t)] \cdot c(t) \\ &= [1 + M \cdot \cos(2\pi f_m t + \phi)] \cdot A \cdot \sin(2\pi f_c t) \end{aligned}$$

- 运用积化和差恒等式，y(t) 可以用三个正弦波的和表示：

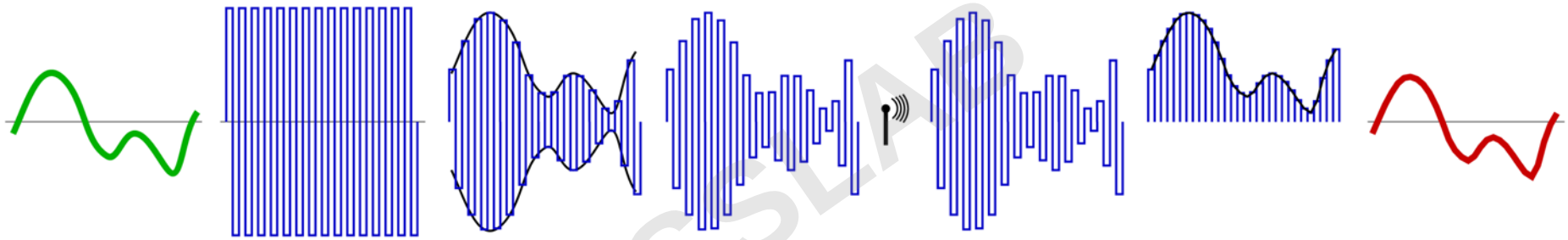
$$y(t) = A \cdot \sin(2\pi f_c t) + \frac{AM}{2}\left[\sin(2\pi(f_c + f_m)t + \phi) + \sin(2\pi(f_c - f_m)t - \phi)\right].$$

边带信号

USSLAB

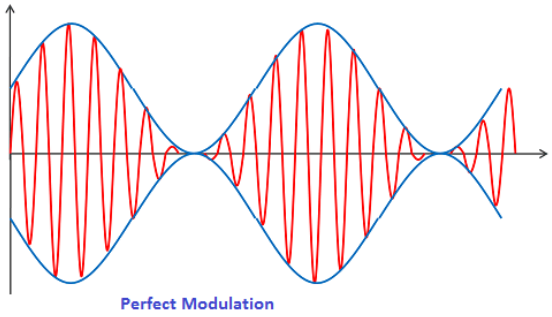
AM调制（续）

- AM调制信号示意图

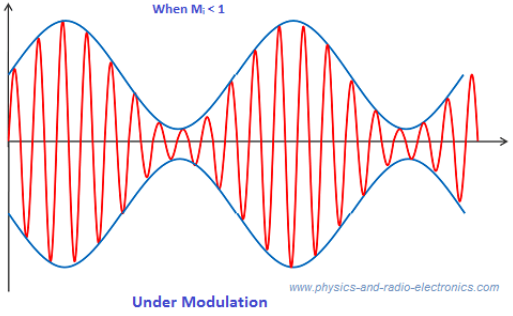


- 调制深度M示意图

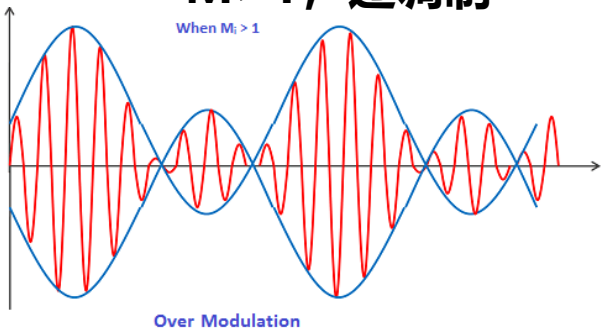
M=1，正常调制



M<1，欠调制



M>1，过调制

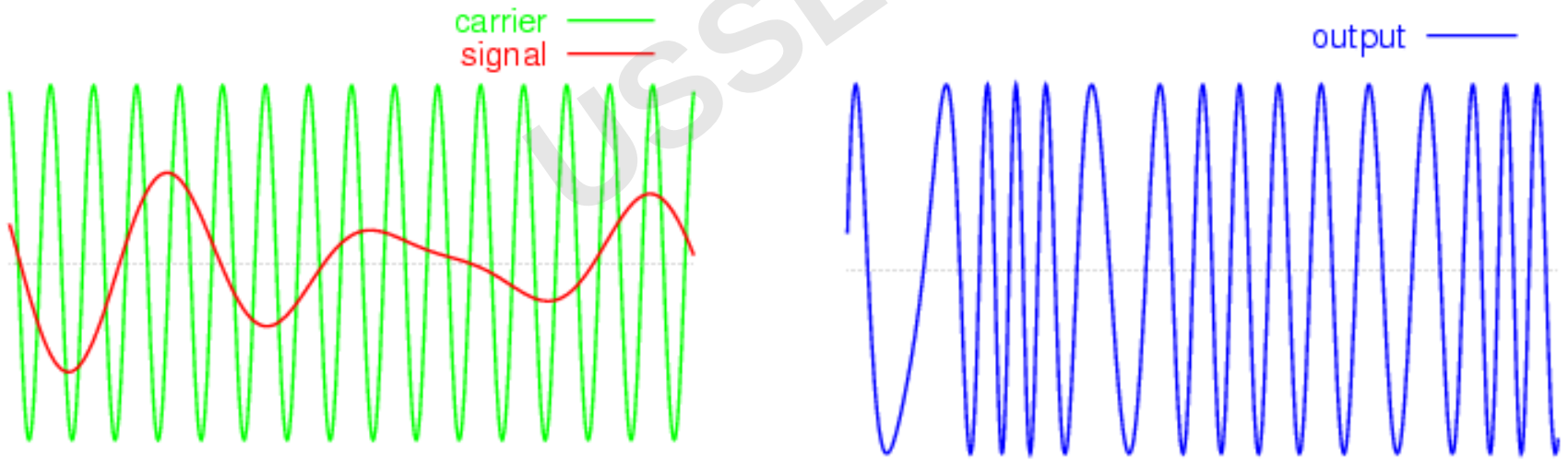


HUSSLAB

调制方式-FM

- **调频-Frequency Modulation (FM)**

  是一种以载波的瞬时频率变化来表示信息的调制方式。调幅方式则是通过载波幅度的变化来表示信息，而其频率却保持不变。

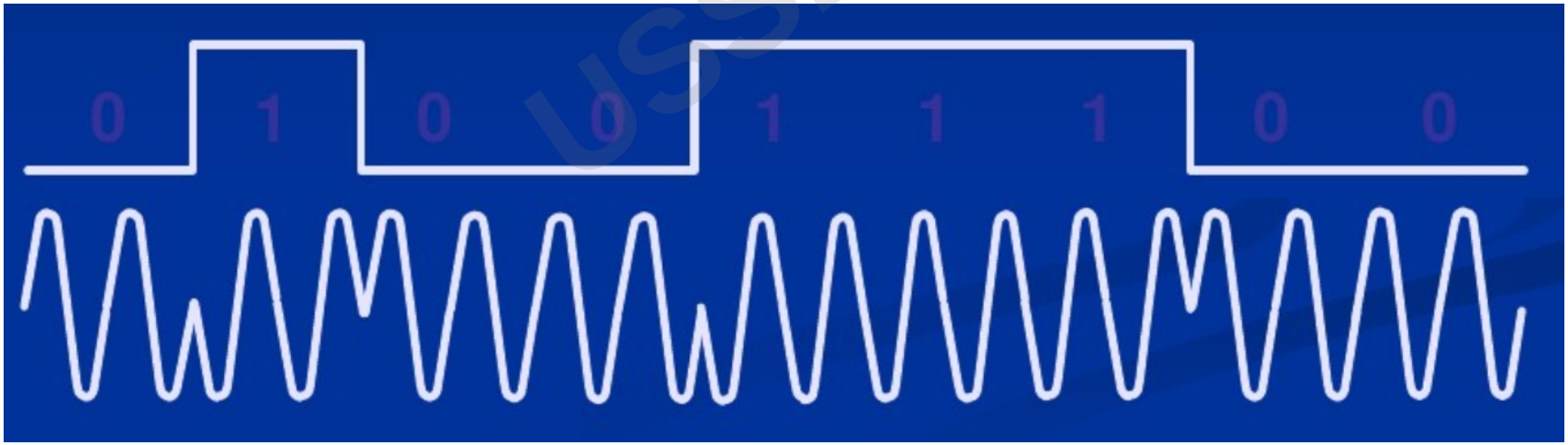


USSLAB

调制方式-PM

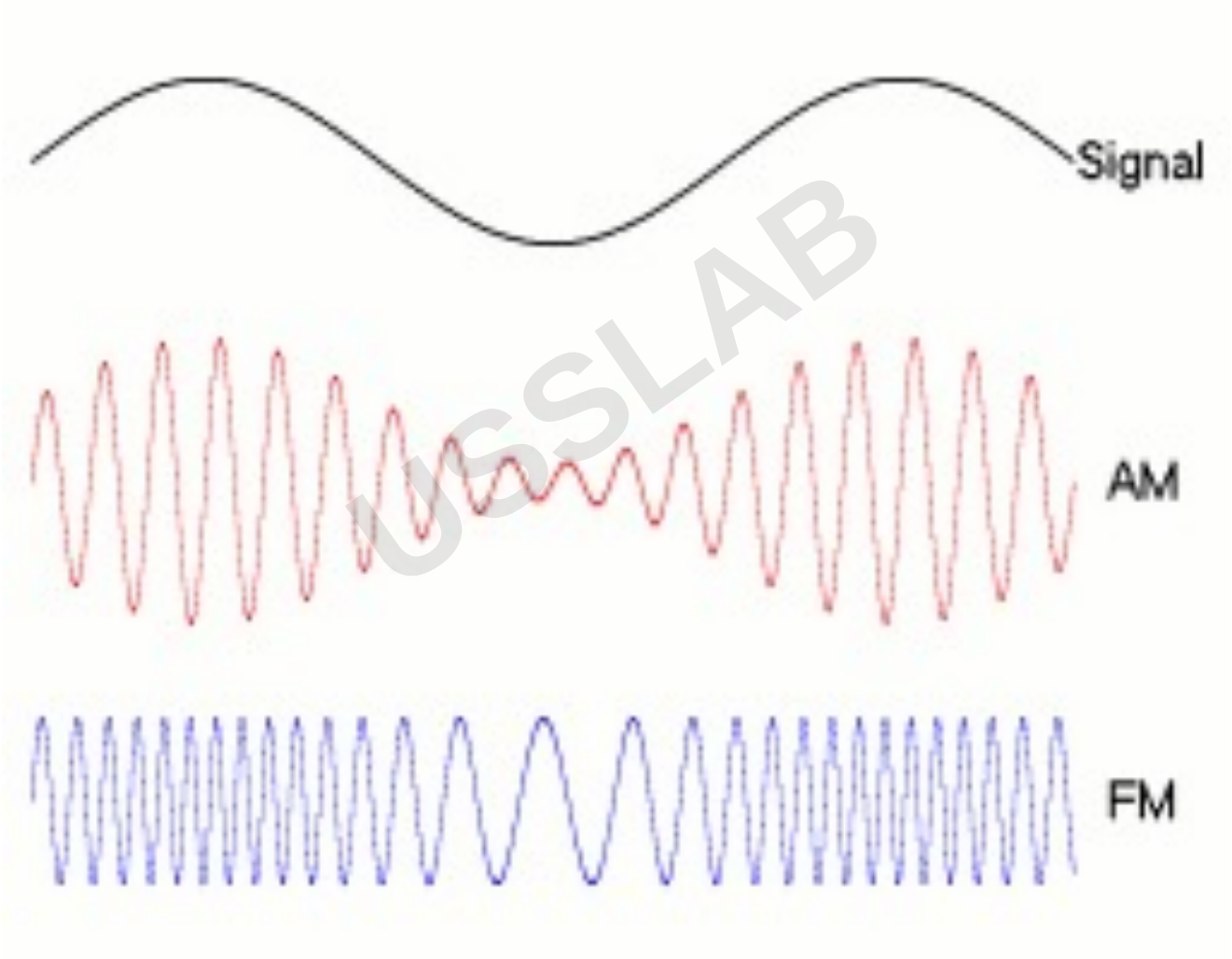
- **调相-Phase Modulation**

  相位调制，是一种以载波的**瞬时相位变化**来表示信息的调制方式，载波的初始相位是随着调制信号而变化的。图中上面为调制信号，下面为载波信号。



USSLAB

AM和FM对比



USSLAB

海豚音攻击

- 海豚音攻击——AM（幅度调制）

USSLAB

海豚音攻击——器件非线性作用

- **非线性作用**
  - **定义**：非线性是指器件的输入输出不是线性相关，电路元器件在本质上是非线性的，由非线性的电路构成或者实现的系统也是非线性的。
  - 麦克风、放大器及滤波器等器件都是输入/输出信号传输特性中具有平方及多次项非线性的组件。
  - **模型**：对麦克风的非线性，假设输入信号为S_{in}，输出信号S_{out}为：

$$s_{out}(t) = As_{in}(t) + Bs_{in}^2(t)$$

  - 其中A是输入信号的增益，B是二次项的增益，以此类推（后面多次项由于能量较小可忽略）。
  - 线性分量正弦输入频率为 f 的信号输出具有相同频率 f 的正弦信号。相比之下，非线性会产生谐波和叉积。
  - **这些非线性特征会带来不应该有的失真，产生新的频率，通过精心设计的输入信号，可以利用这些新的频率恢复出基带信号（攻击指令）。**

USSLAB

海豚音攻击

- **非线性作用影响模型**
  - 假设所需的语音基带信号为$m(t) = cos(2\pi f_m t)$，我们选择中心频率为f_c的载波上的调制信号为：
  - 则：

$$s_{in}(t) = m(t)\cos(2\pi f_c t) + \cos(2\pi f_c t)$$

$$\begin{aligned} \boldsymbol{S_{in}}^{\boldsymbol{2}}(\boldsymbol{t}) &= [\cos(2\pi f_m t)\cos(2\pi f_c t) + \cos(2\pi f_c t)]^2 \\ &= \frac{1+\cos(2\pi 2 f_m)}{2}\frac{1+\cos(2\pi 2 f_c t)}{2} + \frac{1+\cos(2\pi 2 f_c)}{2} \\ &+ 2\cos(2\pi f_m t)\frac{\cos(2\pi 2 f_c)+1}{2} \end{aligned}$$

请大家在纸上推导展开结果

USSLAB

海豚音攻击

- **非线性影响模型**

$$
\begin{aligned}
S_{in}^{2}(t) &= [\cos(2\pi f_m t)\cos(2\pi f_c t) + \cos(2\pi f_c t)]^2 \\
&= \frac{1+\cos(2\pi 2 f_m)}{2}\frac{1+\cos(2\pi 2 f_c t)}{2} + \frac{1+\cos(2\pi 2 f_c)}{2} \\
&\quad + 2\cos(2\pi f_m t)\frac{\cos(2\pi 2 f_c)+1}{2}
\end{aligned}
$$

根据三角函数定理：

$$
\frac{\cos(\alpha+\beta)+\cos(\alpha-\beta)}{2} = \cos\alpha\cos\beta
$$

所以频率 f_c 和 f_m 的余弦波相乘得到的信号包含 f_c+f_m 和 f_c-f_m 两个频率分量，因此可以得到麦克风输出信号包含：

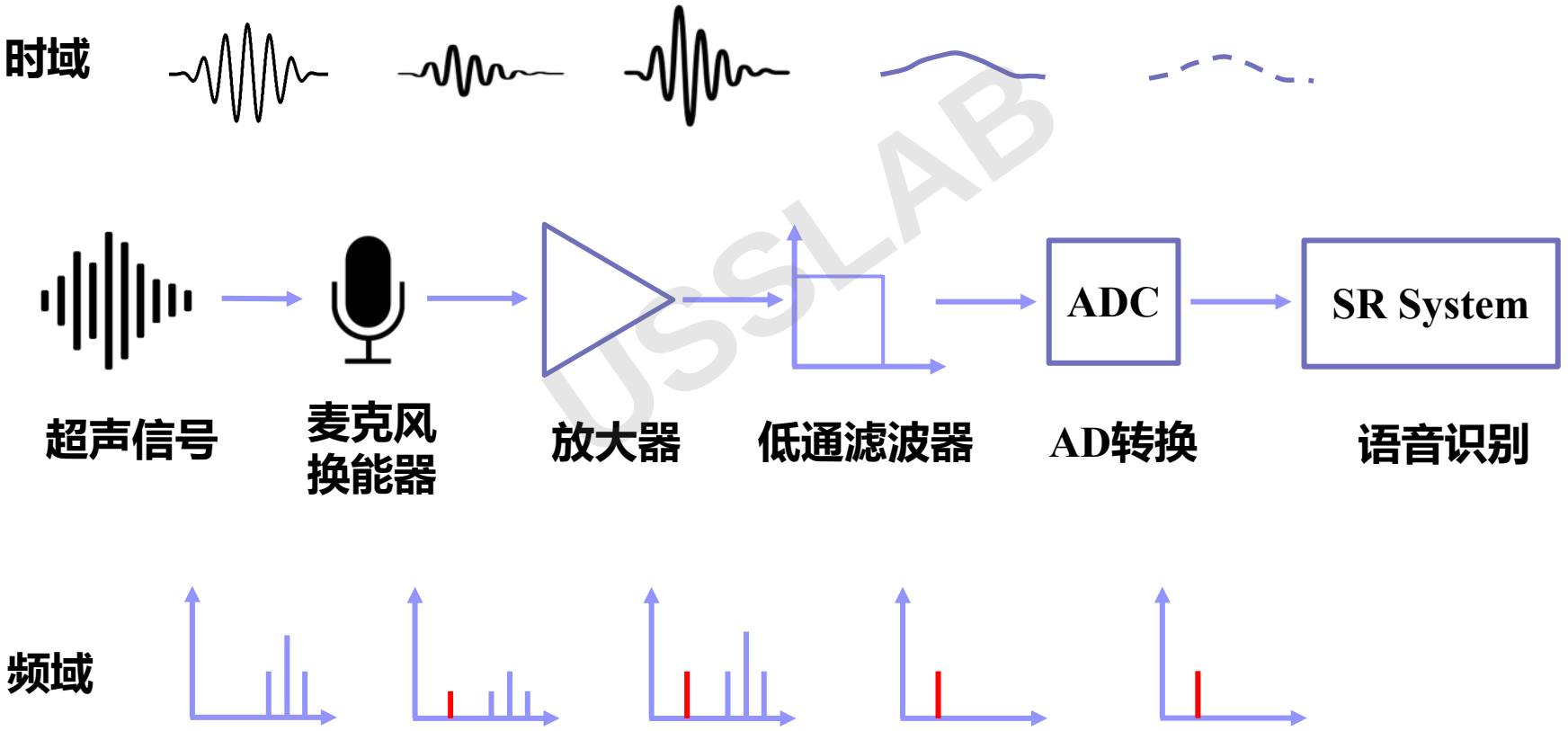
- 预期的频率分量 f_m
- 基本频率分量（即 f_c-f_m，f_c+f_m 和 f_c）
- 谐波和其他交叉乘积（即 $2f_m$，$2(f_c-f_m)$，$2(f_c+f_m)$，$2f_c$，$2f_c+f_m$ 和 $2f_c-f_m$）……

f_m基带信号被成功还原出来！

Quiz: 请推导详细过程

USSLAB

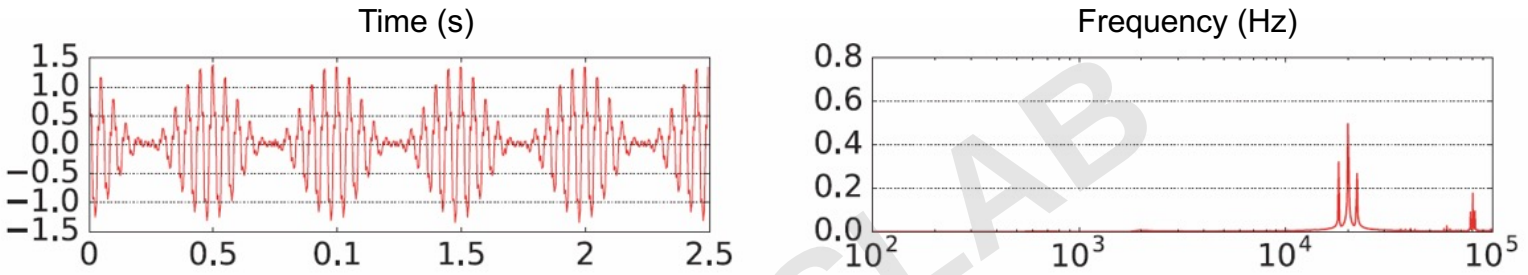
海豚音攻击信号流程图



USSLAB

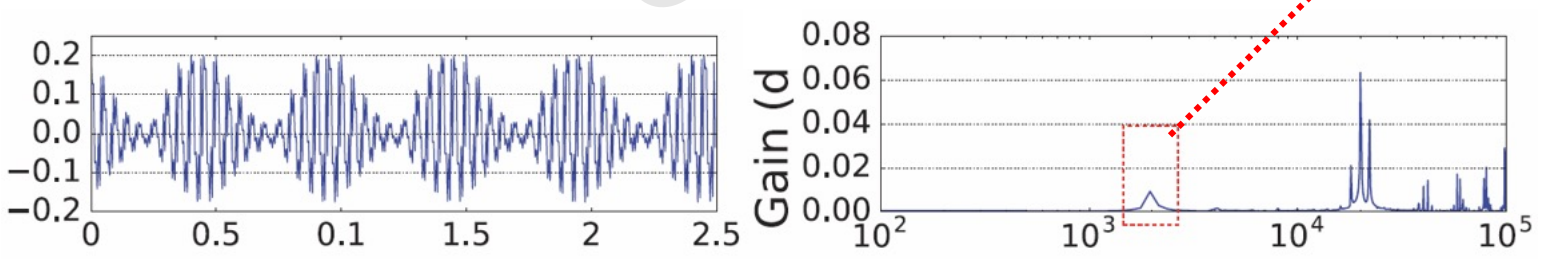
非线性作用验证

$$f_c = 22\ \text{kHz},\ f_m = 2\ \text{kHz}$$



Signals of DolphinAttack（发射信号）

非线性作用!

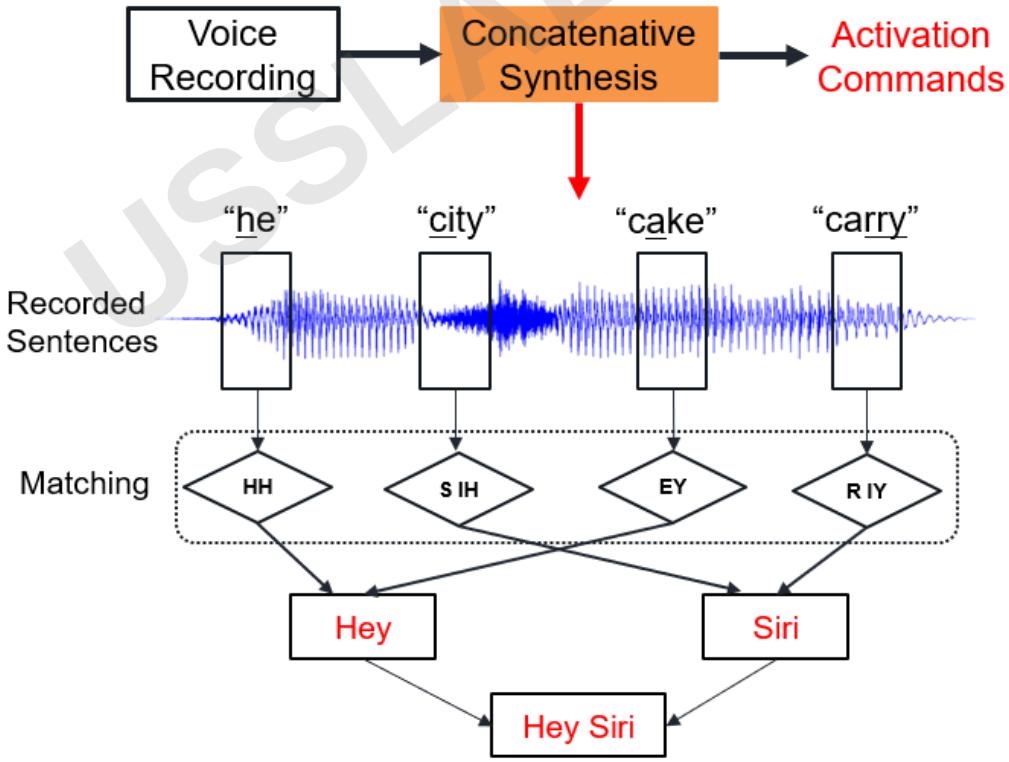


Signals received by a MEMS microphone（麦克风收到信号）

USSLAB

其他方面——绕过声纹解锁

- 语音重放
- 语音合成: (i.e., HH, EY, S, IH, R, IY), 使用包含上述音节的单词如“he, city, cake, “carry



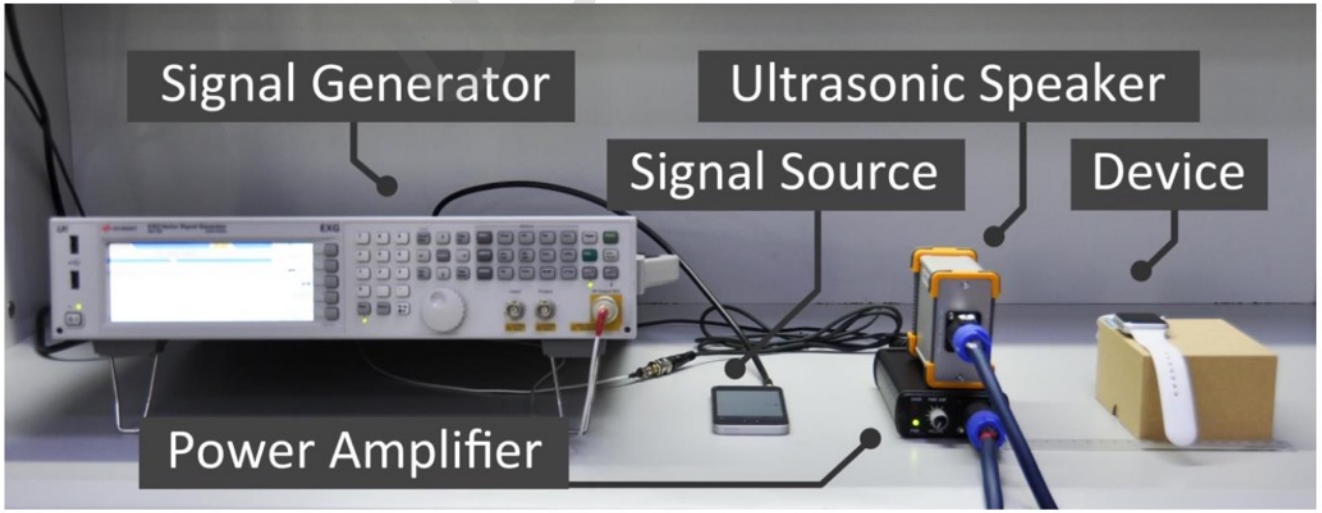
Concatenative synthesis of “Hey Siri”

USSLAB

海豚音攻击

- **攻击环境搭建——**“高级版”实验设备
  - 信号源：手机
  - 信号发生器
  - 功率放大器
  - 发射装置：全频段超声波扬声器(Vifa)

实验一相关！

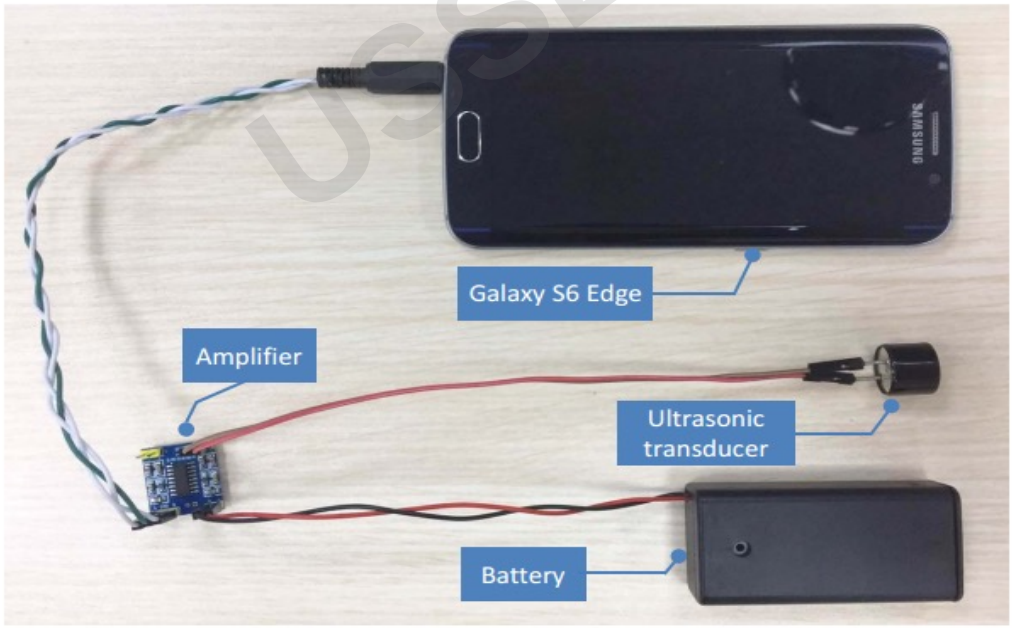


SSLAB

海豚音攻击

- **攻击环境搭建——**“穷人版”攻击实现
  - 信号源：手机
  - 发射装置：超声波探头
  - 低成本功率放大器

实验一相关！

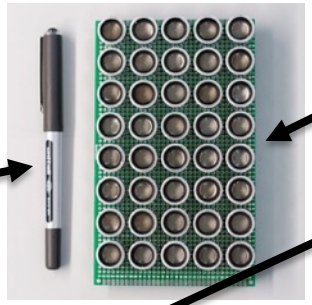


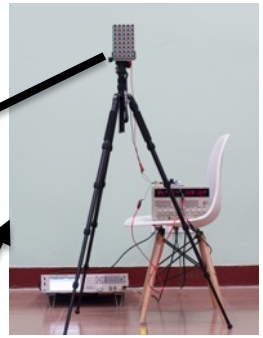
USSLAB

近距离攻击效果演示

远距离攻击效果演示

可以作为探究性实验研究。





10 meters

10 meters

20 meters

USSLAB

全向攻击演示

非相同物理空间环境下如何攻击?

案例：小米汽车发布会小爱同学服务器被DDoS攻击

攻击验证——设备及参数

Manuf.	Model	OS/Ver.	SR System	Attacks		Modulation Parameters		Max Dist. (cm)	
				Recog.	Activ.	f_c (kHz) & [Prime f_c] ‡	Depth	Recog.	Activ.
Apple	iPhone 4s	iOS 9.3.5	Siri	√	√	20–42 [27.9]	≥ 9%	175	110
Apple	iPhone 5s	iOS 10.0.2	Siri	√	√	24.1 26.2 27 29.3 [24.1]	100%	7.5	10
Apple	iPhone SE	iOS 10.3.1	Siri	√	√	22–28 33 [22.6]	≥ 47%	30	25
			Chrome	√	N/A	22–26 28 [22.6]	≥ 37%	16	N/A
Apple	iPhone SE †	iOS 10.3.2	Siri	√	√	21–29 31 33 [22.4]	≥ 43%	21	24
Apple	iPhone 6s *	iOS 10.2.1	Siri	√	√	26 [26]	100%	4	12
Apple	iPhone 6 Plus *	iOS 10.3.1	Siri	×	√	– [24]	–	–	2
Apple	iPhone 7 Plus *	iOS 10.3.1	Siri	√	√	21 24-29 [25.3]	≥ 50%	18	12
Apple	watch	watchOS 3.1	Siri	√	√	20–37 [22.3]	≥ 5%	111	164
Apple	iPad mini 4	iOS 10.2.1	Siri	√	√	22–40 [28.8]	≥ 25%	91.6	50.5
Apple	MacBook	macOS Sierra	Siri	√	N/A	20-22 24-25 27-37 39 [22.8]	≥ 76%	31	N/A
LG	Nexus 5X	Android 7.1.1	Google Now	√	√	30.7 [30.7]	100%	6	11
Asus	Nexus 7	Android 6.0.1	Google Now	√	√	24–39 [24.1]	≥ 5%	88	87
Samsung	Galaxy S6 edge	Android 6.0.1	S Voice	√	√	20–38 [28.4]	≥ 17%	36.1	56.2
Huawei	Honor 7	Android 6.0	HiVoice	√	√	29–37 [29.5]	≥ 17%	13	14
Lenovo	ThinkPad T440p	Windows 10	Cortana	√	√	23.4–29 [23.6]	≥ 35%	58	8
Amazon	Echo *	5589	Alexa	√	√	20-21 23-31 33-34 [24]	≥ 20%	165	165
Audi	Q3	N/A	N/A	√	N/A	21–23 [22]	100%	10	N/A

‡ Prime f_c is the carrier wave frequency that exhibits highest baseband amplitude after demodulation. – No result

† Another iPhone SE with identical technical spec.

* Experimented with the front/top microphones on devices.

WSLAB

如何进行防护

- 软件方案
- 硬件方案

USSLAB

5.3 传感器与执行器安全

- 第三节 传感器测量安全

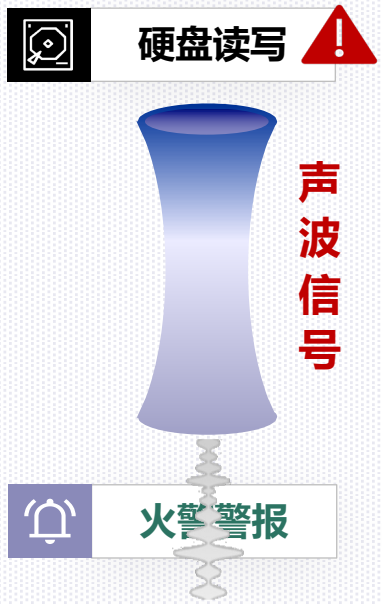
1. 测量安全概述
2. 传感器脆弱性
3. 换能攻击及案例分析
4. 换能攻击传递函数模型
5. 换能攻击防护

USSLAB

传感器安全：声波导致数据中心宕机

数字域

物理域



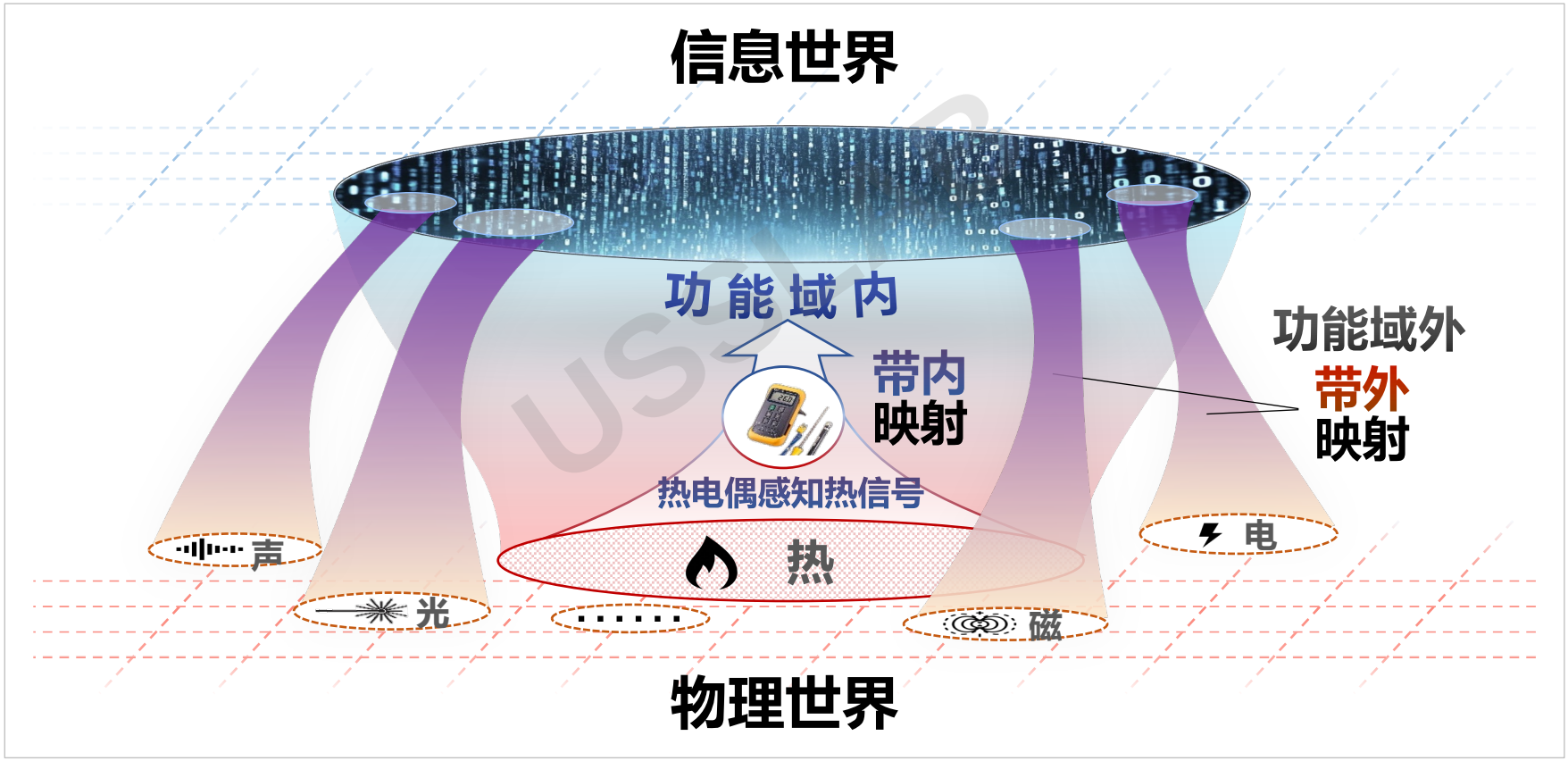
欧洲ING银行：声波 → 数据中心宕机

声波信号 → 硬盘跌落传感器 → 读写错误、系统宕机

官方声称：前所未有(“exceptional”)的技术性灾难问题

5.3.2 传感器脆弱性体系

- 传感器为何存在脆弱性?



真实映射 ≠ 理想映射

5.3.1 传感器测量安全

- **基本问题与意义**

定义：传感器测量安全指**传感器的测量值能否真实反映被测对象**。

- **举例：**在实际温度为20摄氏度的情况下，热电偶测量的温度为 -1000 摄氏度；在一个安静的环境中，麦克风记录到了人耳听不到的声音；静置的加速度计测量到了运动时的加速度等等。

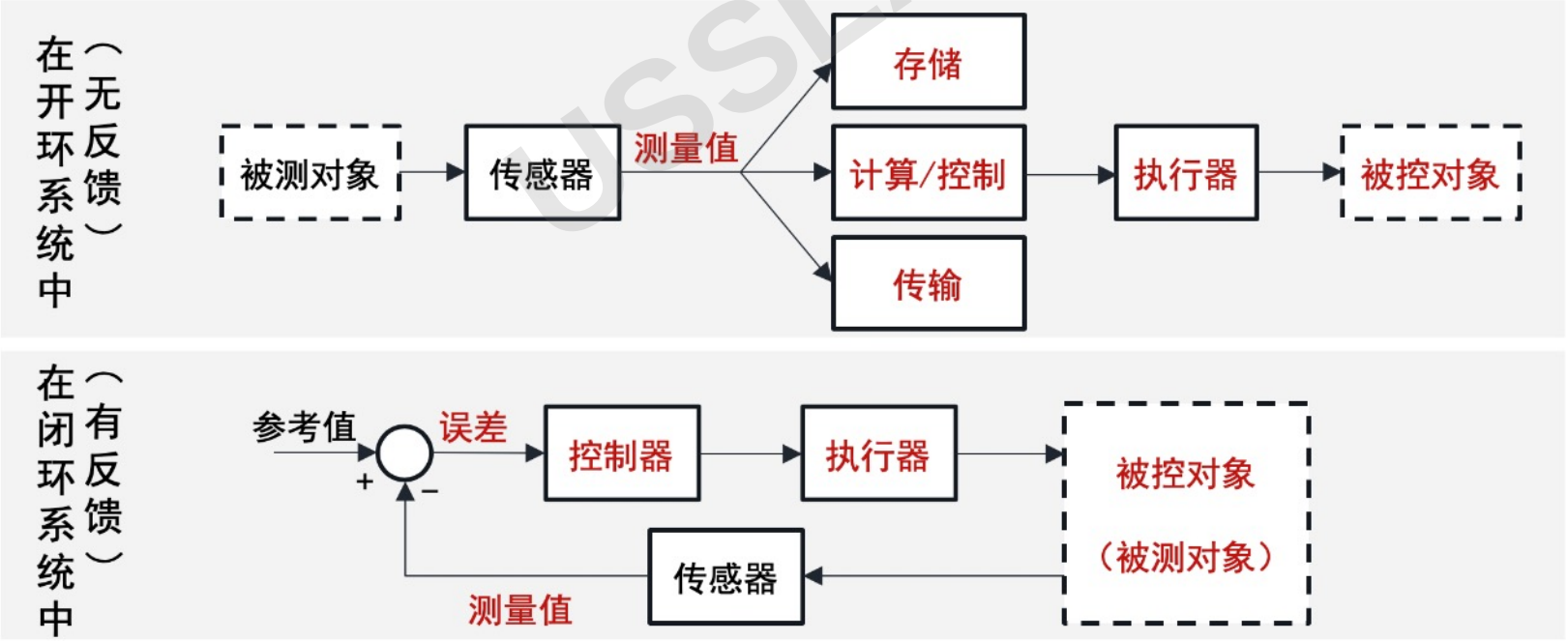
Q：想一想，为什么会出现测量安全问题?

USSLAB

5.3.1 测量安全概述

- **基本问题与意义**

思考：如果传感器的测量值不可信，那么对于使用这些不可信测量值的系统和设备会有什么样的影响？想想，传感器输出影响的单元有哪些？

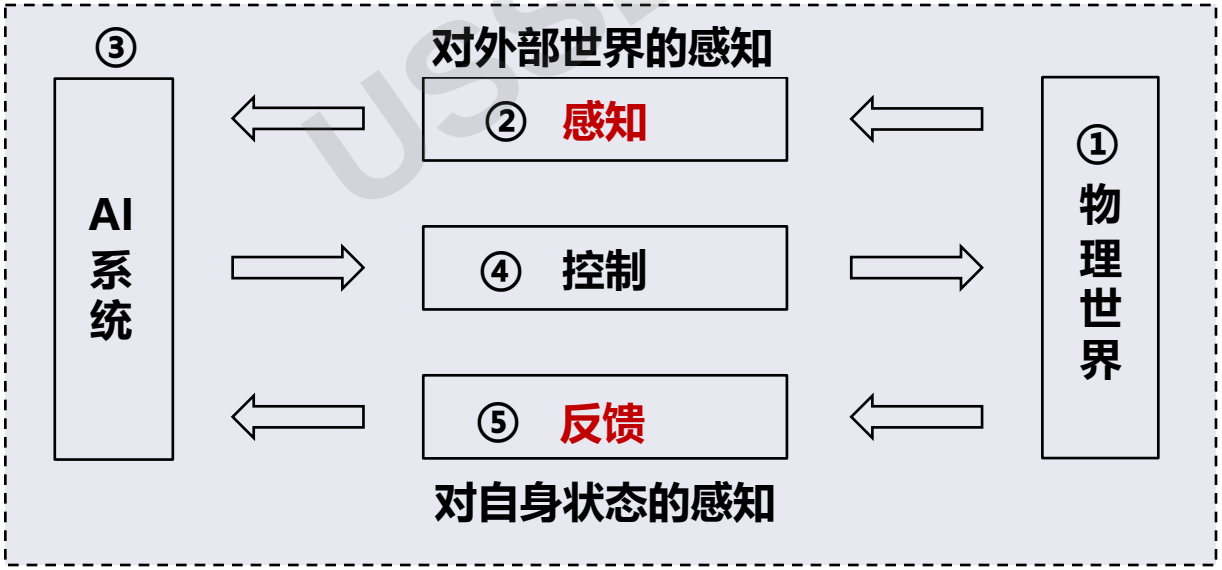


USSLAB

5.3.1 测量安全概述

- **基本问题与意义**

思考：如果传感器的测量值不可信，那么对于使用这些不可信测量值的系统和设备会有什么样的影响？如果传感器用于AI自主系统呢？

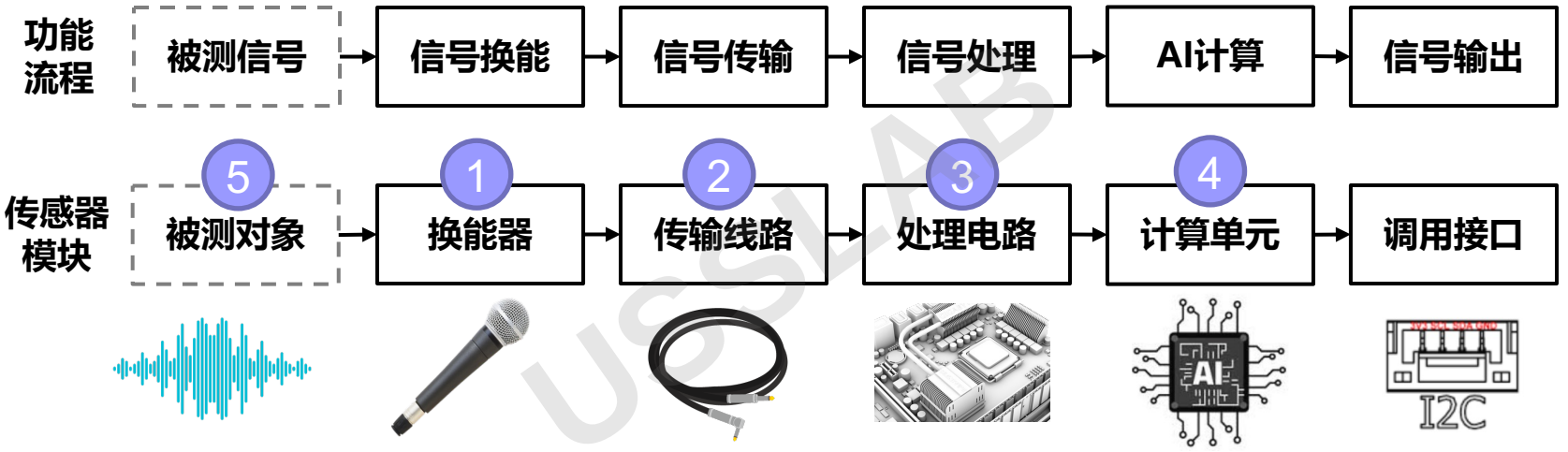


传感器安全对系统安全影响（自动驾驶为例）

USSLAB

5.3.2 传感器脆弱性体系

传感器功能模型



传感器脆弱性分类

1. 信号换能脆弱性
2. 信号传输脆弱性
3. 信号处理脆弱性
4. **感算联动脆弱性**（对于智能传感器而言）
5. 信号鉴权脆弱性（对于主动传感器而言）

……

USSLAB

5.3.2 传感器脆弱性

一、信号换能过程脆弱性

- **定义**：传感器将**设计规范之外**的信号进行换能，包括**超限脆弱性**、**跨场脆弱性**。
- **设计规范**：**信号频率、幅值、类型等**
- **案例**：海豚音攻击、LightoCommand……

USSLAB

- **超限脆弱性**

传感器将设定**量程范围外**（如频率、幅值）的**同类**待测信号转换成电信号。

超声波信号　麦克风　语音

- **跨场脆弱性**

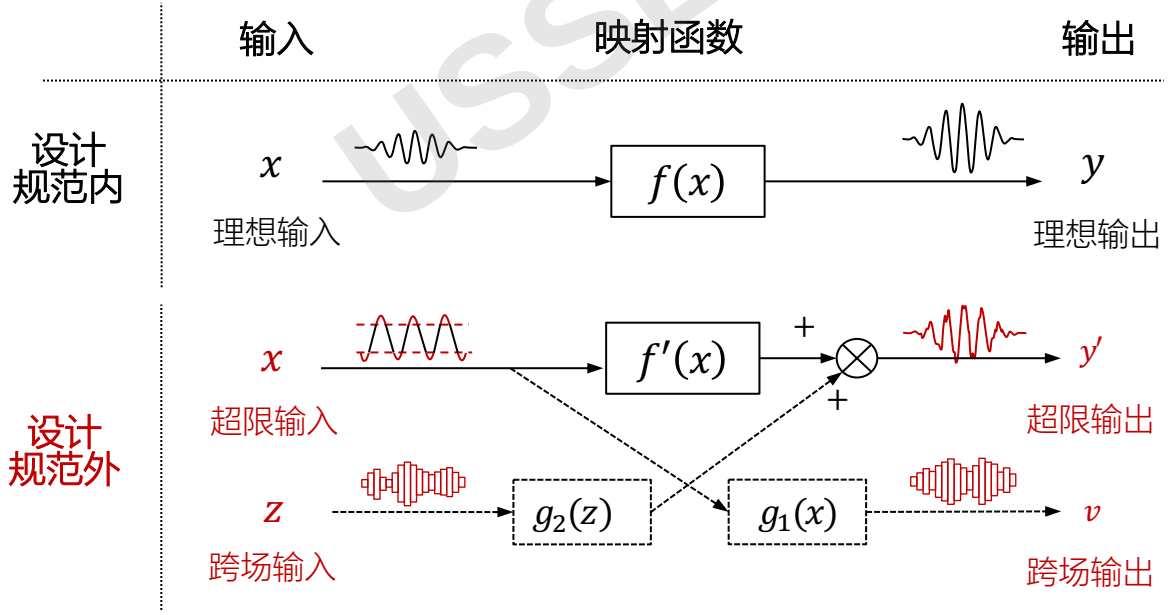
传感器将功能设定之外的**非同类信号**转换为电信号，实现了不同物理场的跨场测量。

声波信号　陀螺仪　加速度

5.3.2 传感器脆弱性

一、信号换能过程脆弱性

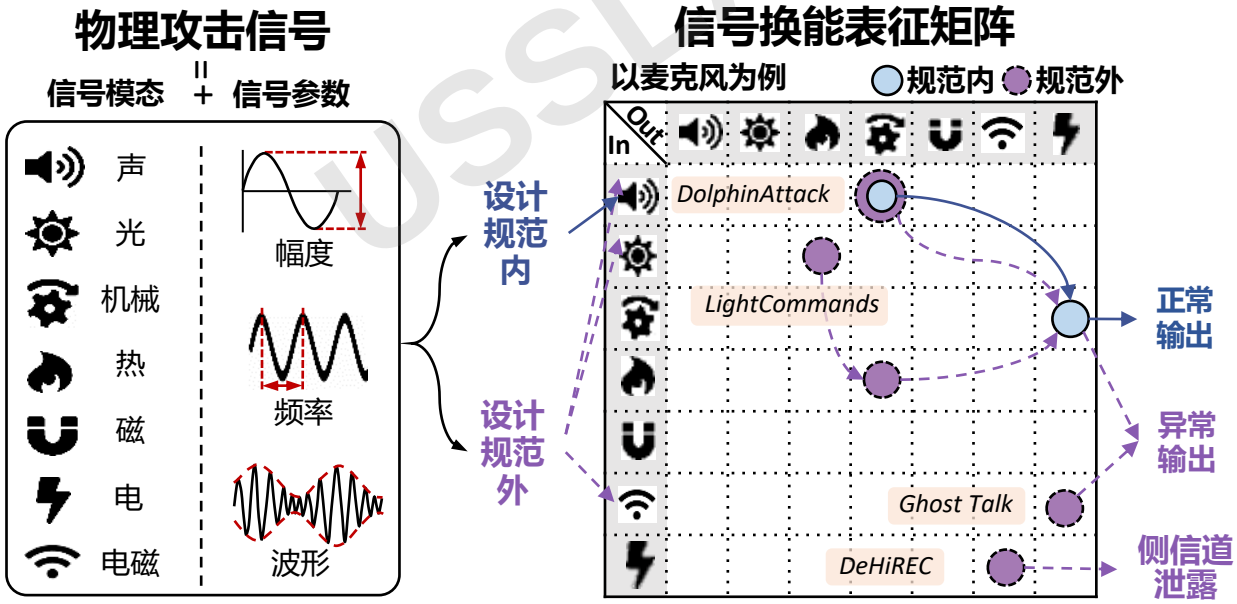
- **定义**：传感器将**设计规范之外**的信号进行换能，包括**超限脆弱性**、**跨场脆弱性**。
- **设计规范：信号频率、幅值、类型等**
- **机理模型：**



USSLAB

5.3.2 传感器脆弱性

- **一、信号换能过程脆弱性**
  - **攻击链路表征：**
    - **信号模态空间：**声、光、力、热、磁、电、电磁等
    - **信号参数空间：**幅度、频率、波形等

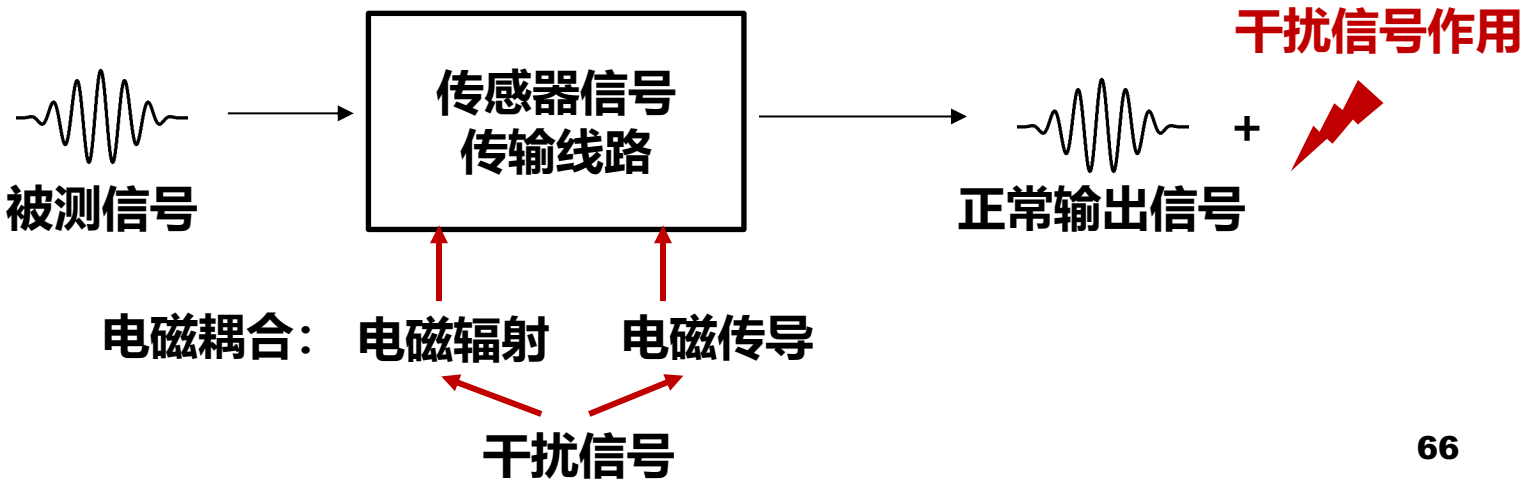


JSSLAB

5.3.2 传感器脆弱性

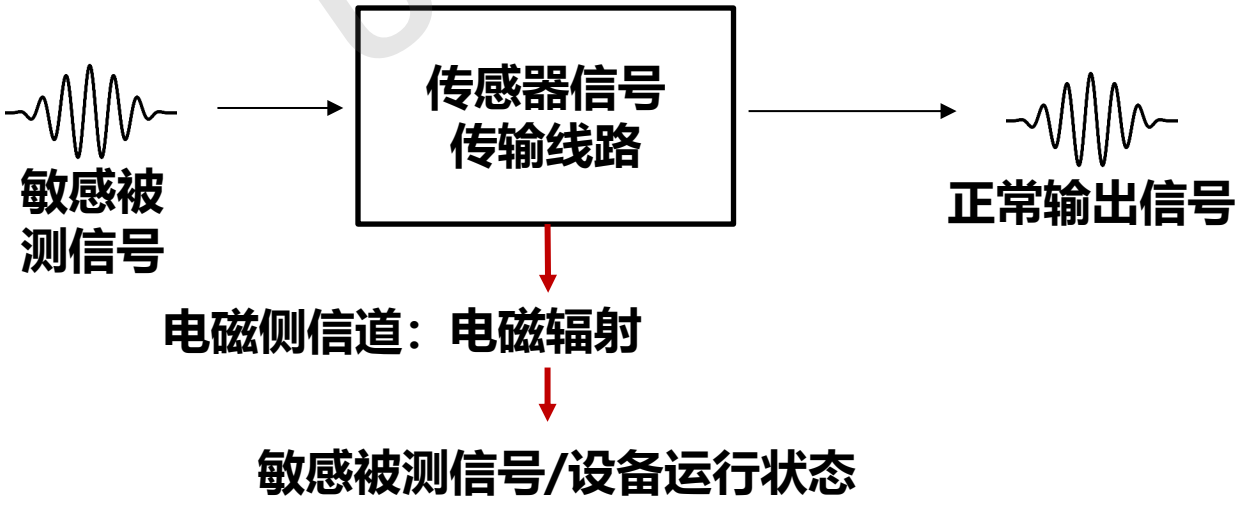
- **二、信号传输过程脆弱性**
  - **定义1**：传感器信号传输线路基于电磁耦合原理接收干扰信号，导致干扰信号在传输线路上作用并改变传感器输出。
  - **电磁耦合：**传感器的传输电路或硬件结构构成潜在的接收天线或传导线路，接收外界的电磁辐射信号或电磁传导信号，从而改变测量电路的模拟电信号。
  - 案例：GhosTalk

USSLAB



5.3.2 传感器脆弱性

- **二、信号传输过程脆弱性**
  - **定义2**：传感器信号传输线路基于电磁感应原理辐射电磁信号，导致传感器内敏感被测信号或设备运行状态被推断。
  - **电磁侧信道：**传感器传输电路或硬件结构构成潜在的发射天线，当天线中有电流经过时，根据电磁感应原理向外界发射电磁辐射信号。
  - 案例：DeHiRec （S&P 23）



USSLAB

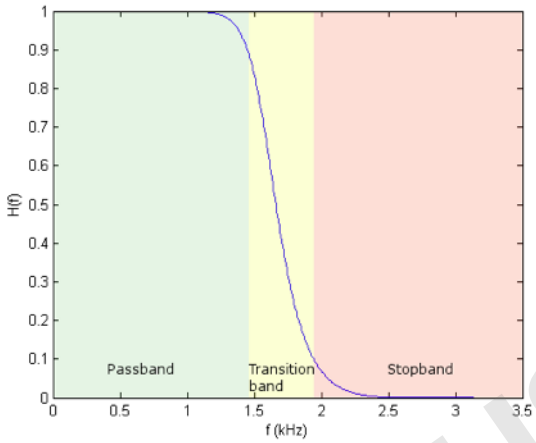
5.3.2 传感器脆弱性

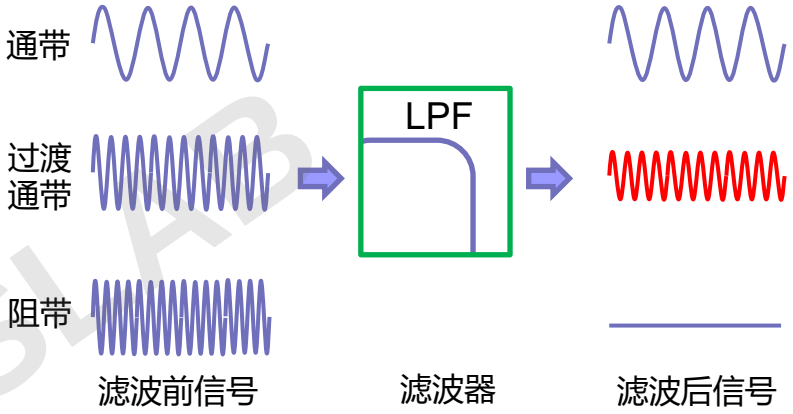
- **三、信号处理过程脆弱性**
  - **定义**：信号处理过程中由于信号处理器件的非理想特性导致传感器测量结果不反映真实情况，通常需要和其他脆弱性配合。
  - 信号处理器件：**滤波器、放大器、模数转换器（ADC）**等
  - 脆弱性实例：
    - **滤波器：非完美截止**
    - **放大器：饱和效应**
    - **放大器：非线性解调**
    - **ADC：混频特性**
    - …
- 器件的各种特性是攻击信号生效的必要条件！

USSLAB

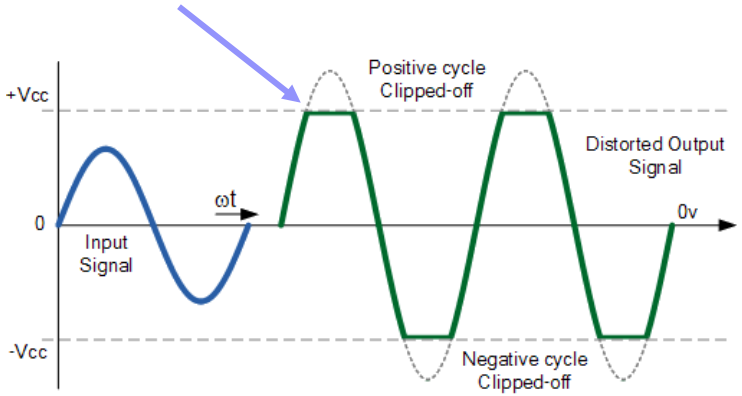
5.3.2 传感器脆弱性：信号处理过程

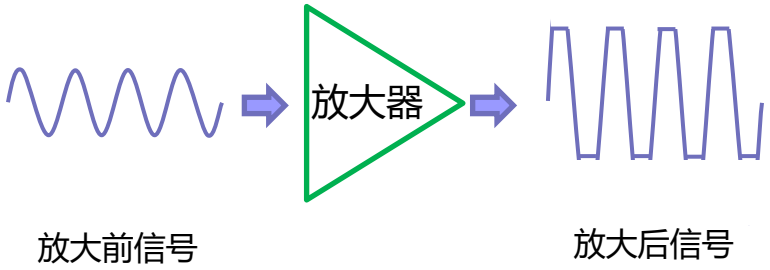
- **1. 滤波器非完美截止特性：**过渡频带内的信号不会完全被去除。





- **2. 放大器饱和效应：**放大器的测量输出量程有限，超出量程的信号可能会被削顶失真。





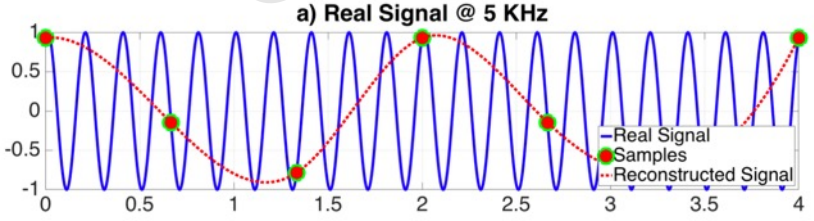
USSLAB

5.3.2 传感器脆弱性：信号处理过程

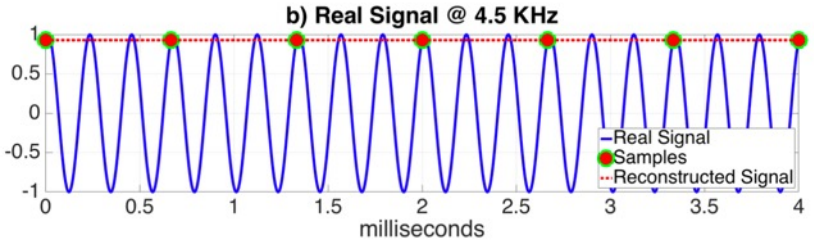
- **3. 放大器的非线性解调：**由于放大器的非线性特性，经过调制的信号被放大器放大后，解调出多个不同频率的信号，如海豚音攻击。

- **4. ADC混频（Aliasing）：**当采样频率不满足香农定理时，原始信号经采样后的信号频率会改变。如果两个被采样的信号频率差为ADC采样频率的整数倍，则两个不信号在ADC采样之后频率相同。

$f_N = 5\ kHz$
$f = 0.5\ kHz$
$f_s = 1.5\ kHz$



$f_N = 4.5\ kHz$
$f = 0\ kHz$
$f_s = 1.5\ kHz$



采样后频率

$$f_N(f) = f + Nf_S$$

原始信号频率　　采样频率

5.3.2 传感器脆弱性

- **四、感算联动脆弱性**
  - **定义**：物理攻击信号使传感器的输出产生**对抗性扰动**，触发后续**AI算法漏洞**。（细节在AI安全介绍）
  - 例如：声波干扰图像防抖模块，导致**目标识别结果错误**，包括目标从有到无、从无到有、从A到B。

X. Ji et al., "Poltergeist: Acoustic Adversarial Machine Learning against Cameras and Computer Vision," 2021 IEEE Symposium on Security and Privacy (S&P), 2021.

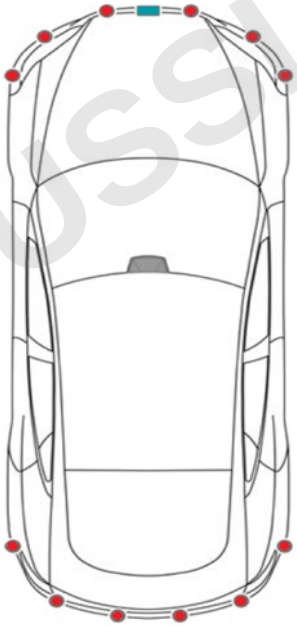
USSLAB

5.3.2 传感器脆弱性

- **五、信号鉴权过程脆弱性**
  - **定义**：主动传感器对外部信号源**缺乏身份合法性认证**过程，接收该信号之后导致感测结果错误，通常对于主动型传感器而言。
  - 例如，攻击者可以主动发射一个传感器的测量信号来进行欺骗，如影响超声波测距、摄像头/激光雷达识别等

攻击信号

毫米波
声波
激光
.....



毫米波雷达

前格栅安装了一个中程雷达。

摄像头

前视摄像头安装在后视镜下的挡风玻璃上。

超声波传感器

超声波传感器位于前后保险杠附近。

.....

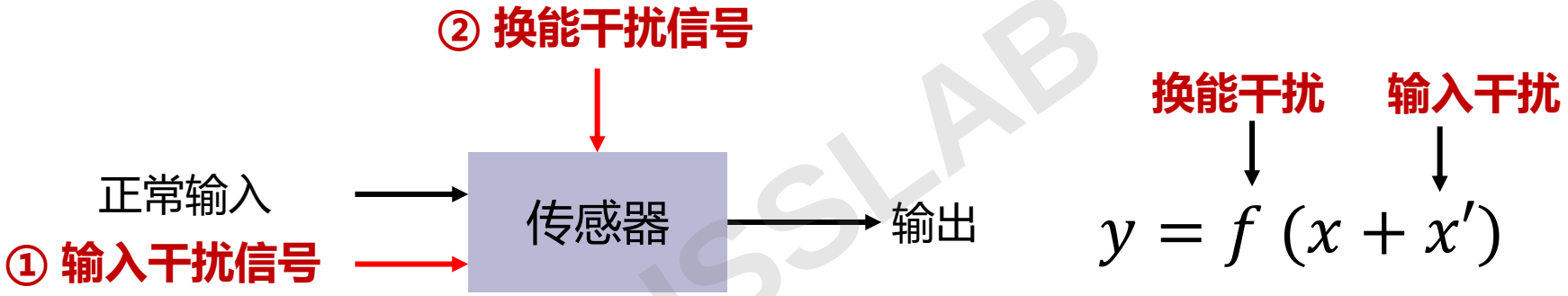
被攻击者发出的信号欺骗，测量错误，做出错误响应！

USSLAB

引申阅读：激光雷达厂商的进化，如禾赛。

5.3.3 传感器换能攻击 (Transduction Attack)

- **定义：**利用传感器**脆弱性**，通过主动构造并注入**攻击信号，干扰传感器正常换能过程**，从而改变传感器测量结果的攻击。



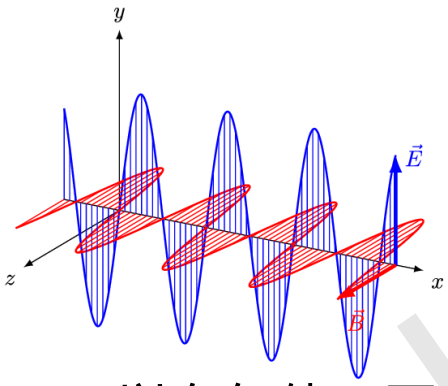
USSLAB

- **解释：**攻击者产生的物理信号可以通过**干扰①测量输入或②换能过程**（也可同时干扰），从而影响传感器的输出测量值

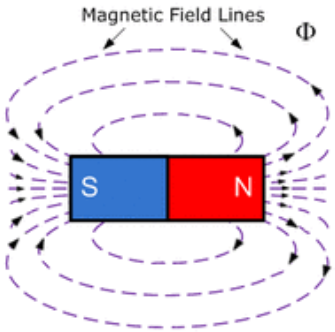
- **换能攻击特点**
  - 不需要和传感器物理接触，因此具有较高的隐蔽性；
  - 在感知过程即注入攻击信号，因此难以防范。

5.3.3 传感器换能攻击 – 攻击信号形态

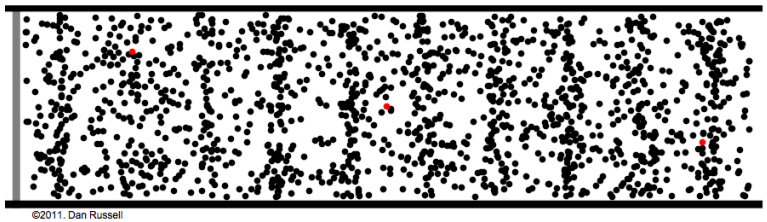
1. 电磁波：同相振荡且互相垂直的电场与磁场，传导、辐射干扰。



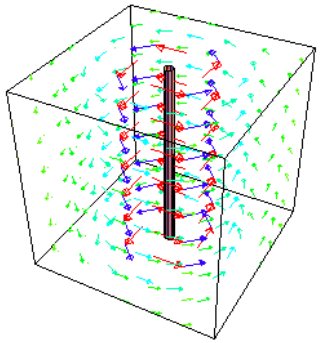
2. 磁场：由磁体、运动的电荷或电场的变化而产生的物理场。



3. 声波：可以在气体、固体、液体中传播的机械波。



4. 电场：存在于电荷周围能传递电荷之间相互作用的物理场。



USSLAB

5.3.3 传感器换能攻击 – 攻击威胁

- **威胁一：拒绝服务（Denial-of-Service, DoS）**
  - **定义**：此类攻击的目的是**阻止传感器输出可用的测量值**，传感器的测量值通常是攻击者无法控制或预测的。
  - **举例**：声音干扰，如果其频率与陀螺仪固有的谐振频率一致，可导致陀螺仪输出一个随机的角速度测量值，导致设备姿态控制功能出错。

- **威胁二：篡改欺骗（Spoofing）**
  - **定义**：此类攻击的目标是**造成传感器输出一个被攻击者操纵的错误测量值**。与拒绝服务攻击的区别是，在此类攻击下攻击者通常可以部分或完全控制传感器的测量值。
  - **举例**：一段特殊制作的声音信号可以欺骗手机中加速度计的输出，从而实现对与其绑定的遥控汽车行为的控制。

USSLAB

5.3.3 传感器换能攻击 – 攻击过程

- **Step 1：信号注入**
  - 定义：恶意干扰信号通过某种方式，转换为传感器信号传输通路上电信号的过程。需要结合传感器的不同环节脆弱性和攻击信号的参数设计，从而将攻击信号注入。
  - 解决信号如何“**进得去**”的问题。

- **Step 2：信号生效**
  - 定义：攻击者对恶意干扰信号进行特殊的设计构造，使得传感器可以输出的预期测量值。主要考虑如何让恶意干扰信号留在传感器的信号传输通路并造成传感器输出攻击者期待的测量值。
  - 解决信号如何“**起作用**”的问题。
  - 综合利用各种信号处理器件的硬件脆弱性。

USSLAB

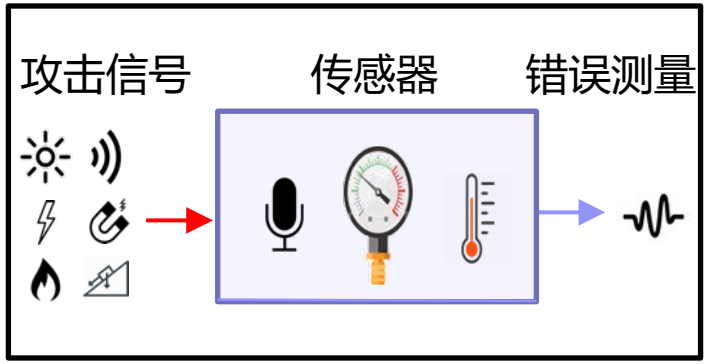
5.3.3 传感器换能攻击

Step 1：信号注入

为向传感器中注入恶意信号，攻击者需要同时考虑信号的注入点和包括信号的**类型、幅度、频率**等可能影响信号注入效率的因素，通常利用的是换能器或者信号传输链路的脆弱性。

考虑因素

1. **信号注入点**和**信号类型**
   a) 换能之前注入：如海豚音
   b) 换能之后注入：如Walnut、GhosTalk
2. **高效注入**因素
   a) 幅度、频率、相位等
3. 同类信号和**跨场**信号
   a) 同场：同类型信号
   b) 跨场：不同类型，如声波到电磁



USSLAB

5.3.3 传感器换能攻击

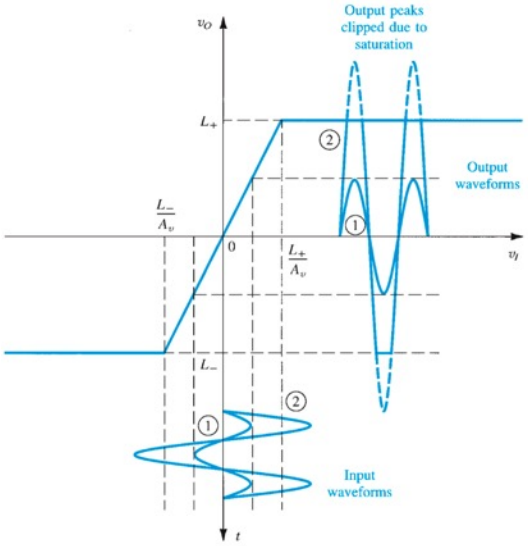
- **Step 2：信号生效**
  - 为使得注入信号可以保留在传感器的信号通路中并篡改传感器输出测量值，攻击者需要对注入的**恶意信号进行优化**，利用传感器信号处理通路软硬件脆弱性，实现对传感器测量值的篡改
  - 回顾传感器信号处理过程脆弱性，常见的信号生效原理包括：
    - **饱和**
    - **交调失真**
    - **包络检测**
    - **混频**
    - **滤波器超频**
    - **……**

回想下第一节传感器基本知识中的各类特性☺

USSLAB

5.3.3 传感器换能攻击

- **Step 2：信号生效——饱和**
  - **定义**：指某种物理量无法超过一定的阈值，是一种模拟电路中常见的现象



如下公式所示，当输入超过一定阈值时放大器就可能进入饱和状态，此时放大器的输出不会继续随着输入线性增长，因而出现了**限幅**：

$$f_i(x_i, n_i + a_i') = \begin{cases} c_i \mathcal{A}(x_i, n_i + a_i'), & \text{当} \mathcal{A}(x_i, n_i + a_i') \le k \\ const, & \text{当} \mathcal{A}(x_i, n_i + a_i') > k \end{cases}$$

其中$\mathcal{A}(x_i, n_i + a_i')$指$x_i$和$n_i + a_i'$叠加后的信号，$x_i$是正常输入，$n_i$是噪声，$a_i'$是攻击信号，$c_i$是放大系数，$k$是保和点。

  - **利用：攻击者可以通过饱和现象向电路中注入可控大小的直流信号，参考WALNUT**

USSLAB

5.3.3 传感器换能攻击

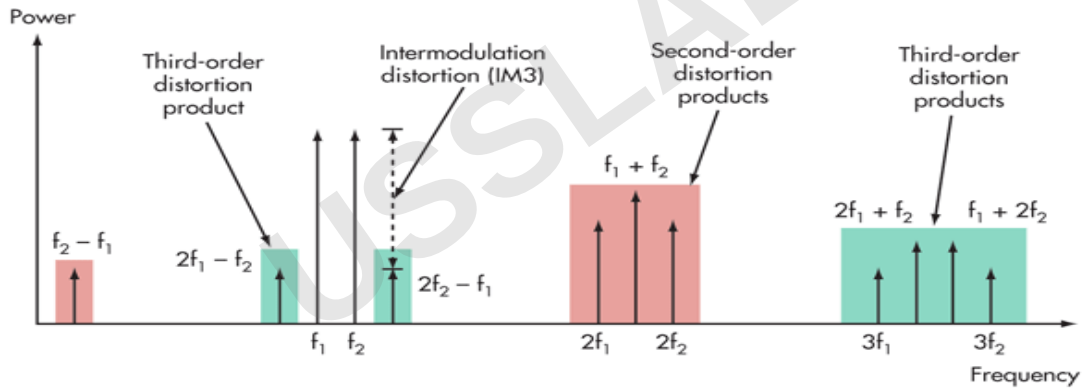
- **Step 2：信号生效——交调失真**
  - **定义**：当一个含有多个频率分量的信号经过一个非线性器件时就有可能发生交调失真（Intermodulation Distortion, IMD）。
  - 常见的非线性器件包括放大器、二极管、换能器等，甚至模数转换器因为其内部存在的放大器也有一定的非线性。
    PS：空气也可以是交调失真信号的载体。
  - **原理**：交调失真会导致输出信号中出现输入信号里不包含的频率分量，它们主要出现在输入信号中频率的和与差以及其倍数。
  - **思考**：对于以下非线性传递函数，假设混合后x_i的包含两个频率，f_1和f_2（$f_1<f_2$），输出x_{i+1}的频率会包括哪些?

$$x_{i+1} = c_0 + c_1(x_i) + c_2(x_i)^2$$

USSLAB

5.3.3 传感器换能攻击

- **Step 2：信号生效——交调失真（续）**
  - 假设混合后的信号x_i包含两个频率，f_1和f_2（$f_1 < f_2$），此时输出x_{i+1}包含频率f_1、f_2、$f_2 - f_1$、$f_1 + f_2$、$2f_1$、$2f_2$……，及直流分量。



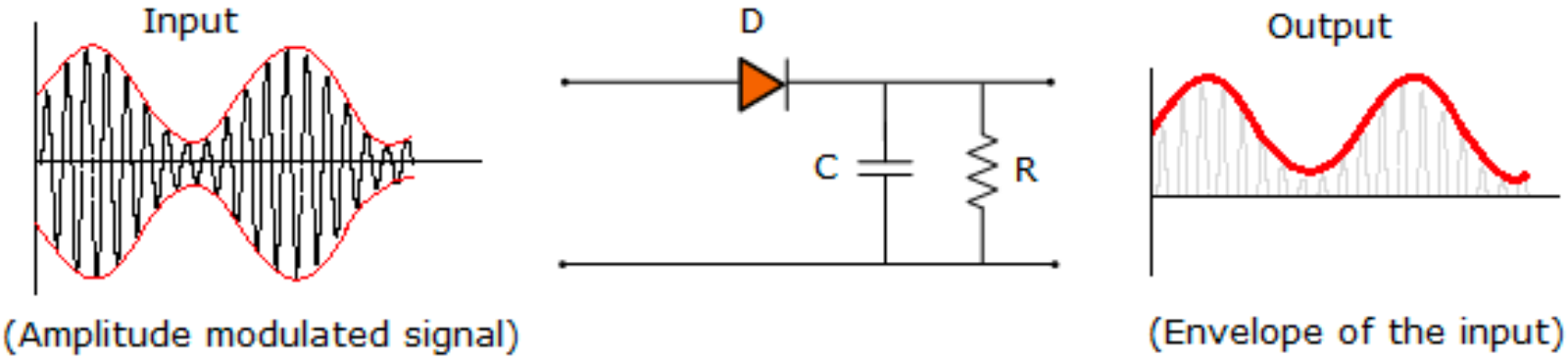
  - **利用：**注意其中$f_2 - f_1$是可能小于f_1和f_2的，因此攻击者可能利用交调失真将恶意的混频信号转化为攻击信号（海豚音攻击）

Q：空气也可以作为非线性介质实现交调失真，如何做海豚音攻击？

USSLAB

5.3.3 传感器换能攻击

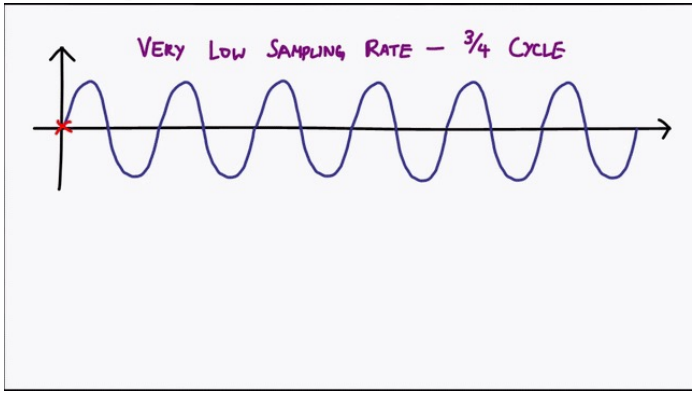
- **Step 2：信号生效——包络检测**
  - 包络检测器可以解调调幅中的基带信号，例如由二极管和电容组成的电路，是一种简单的包络检测器。二极管和电容是模拟电路中非常常见的器件，常用于静电放电保护
  - **案例**：通过将电磁载波信号注入到录音笔，利用录音笔信号处理电路天线效应将信号接收；并利用二极管和电容的包络检测，可以将载波信号解调得到基带语音信号，从而注入了恶意语音到录音笔。



USSLAB

5.3.3 传感器换能攻击

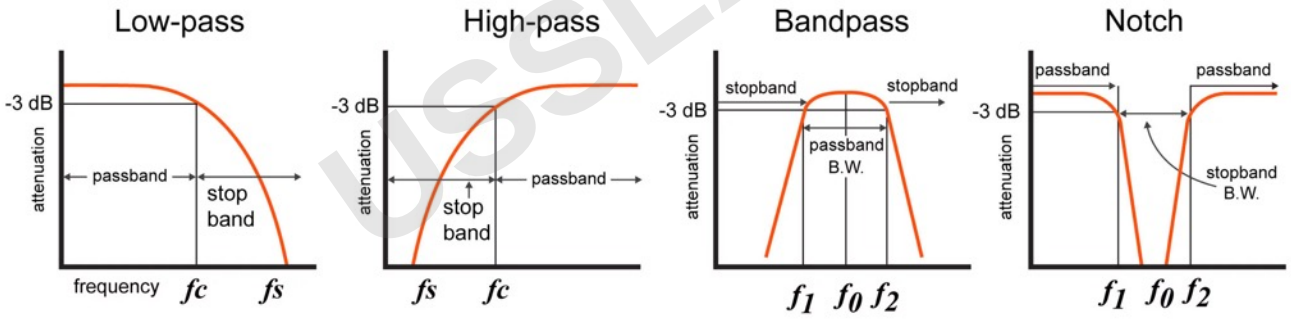
- **Step 2：信号生效——混频**
  - **定义**：根据奈奎斯特-香农采样定律，若一个信号的频率高于采样率的一半，那么这个信号将与其他频率的信号无法区分，这种现象被称为混频（Aliasing）。
  - 举例：如果模数转换器的采样率是F_s，那么频率为f ($f < F_s$)的信号将与频率为$F_s - f$ 的信号无法区分。
  - 利用：攻击者可能利用混频现象影响传感器测量过程。例如Walnut攻击。



USSLAB

5.3.3 传感器换能攻击

- **Step 2：信号生效——滤波超频（非完美滤波特性）**
  - **滤波器的非完美特性**：在理想情况下，一个模数转换器前的滤波器应当去除所有的带外信号并避免混频的发生。然而在现实中大部分的滤波器都有一个截止频率范围，在这个范围中带外信号只会被部分衰减。



  - **利用**：攻击者可以利用滤波器的这一特性构造能够通过滤波器但是会影响其他传感器组件的信号。
  - 例如，MEMS加速度计中的滤波器都有极宽的截止频率范围，因此无法完全去除攻击者注入的高频声音信号，如Walnut攻击。

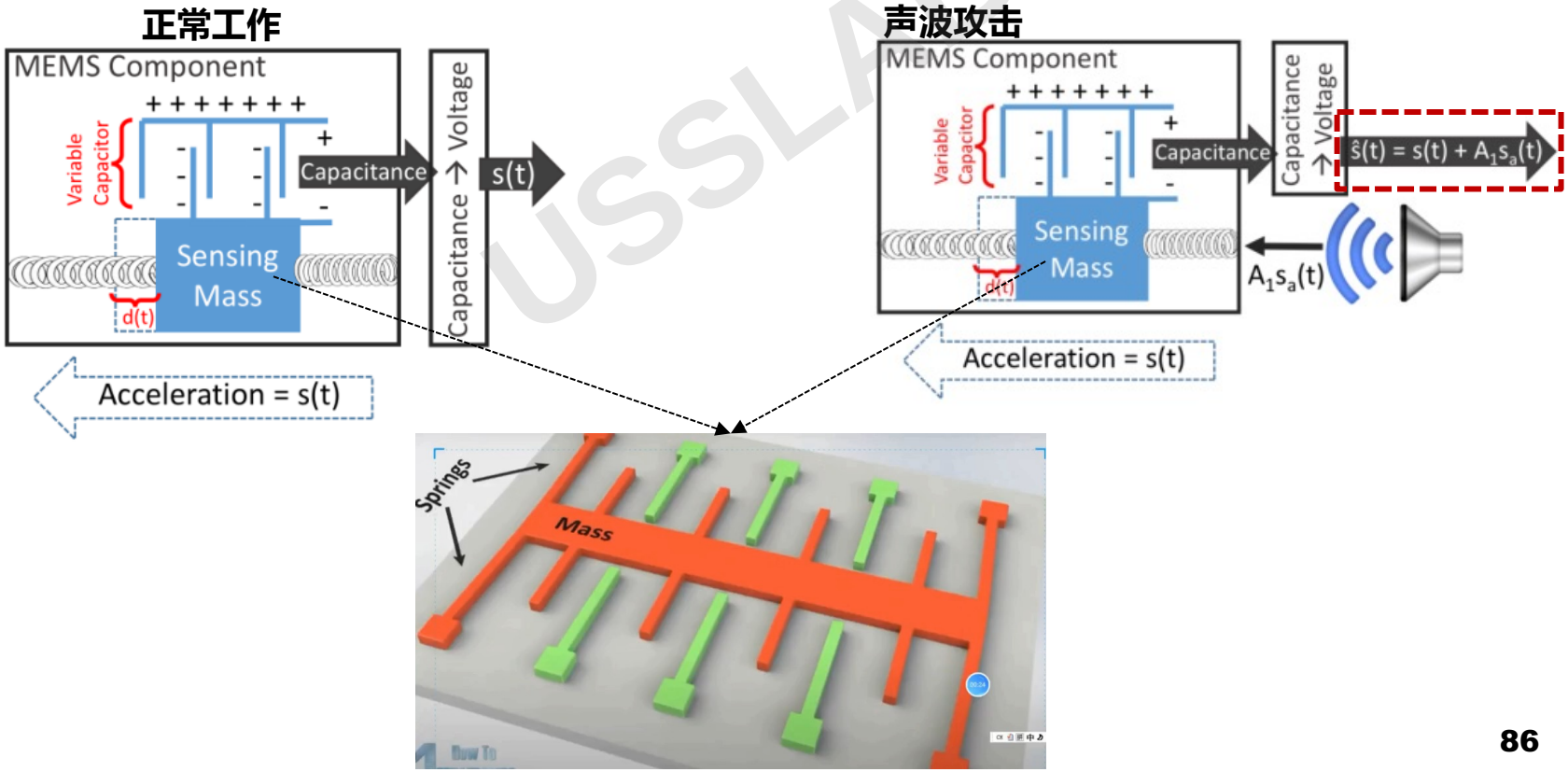
USSLAB

USSLAB

典型传感器安全案例分析

5.3.3 换能攻击案例：WALNUT[1]

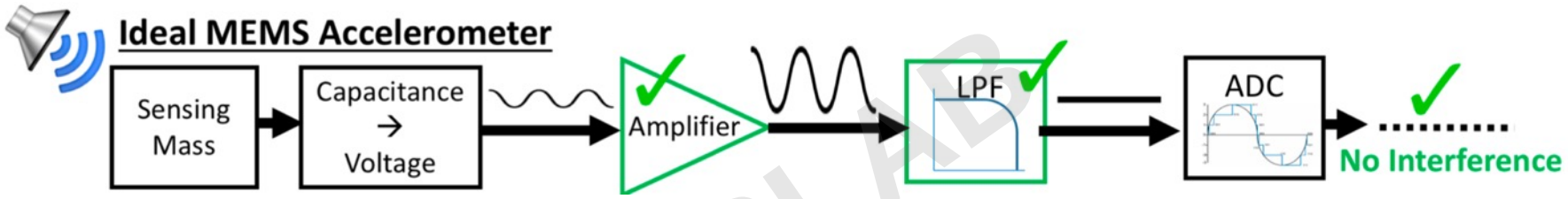
- **脆弱性：放大器饱和、滤波器不完美滤波、ADC混频特性**
- **攻击后果：**利用声波可以控制MEMS加速度传感器输出，例如通过播放音乐文件可以让静止的手机“日行万步”



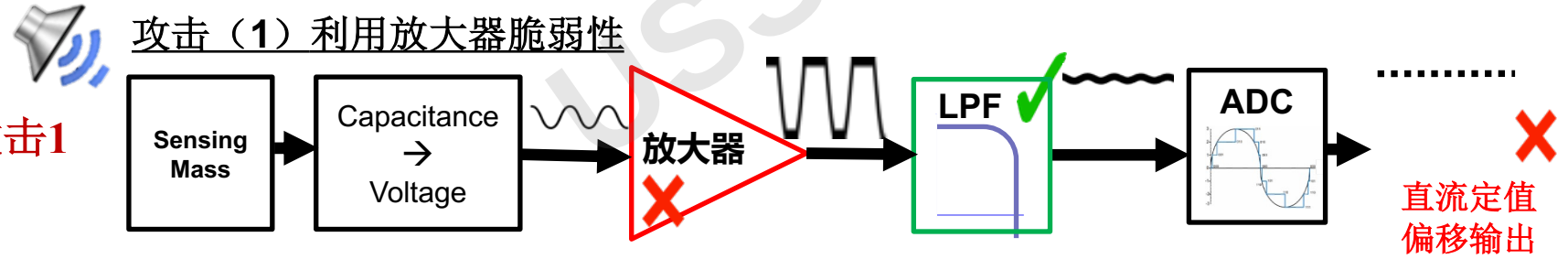
USSLAB

5.3.3 换能攻击案例：WALNUT[1]

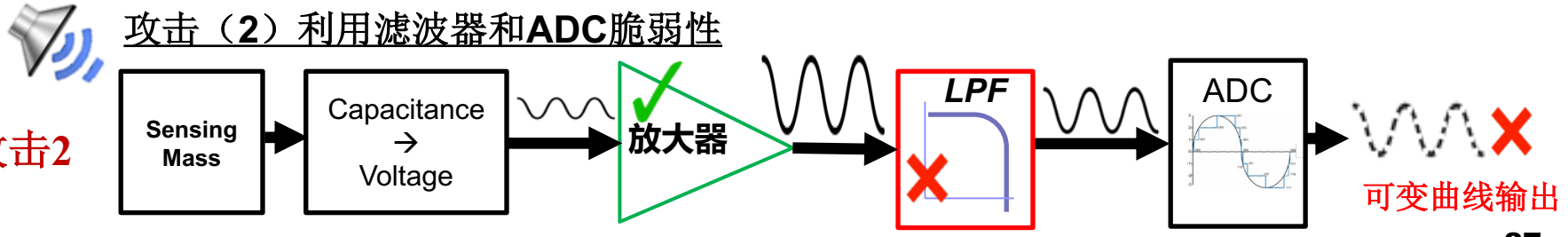
- 根据不同脆弱性，可以实现2种攻击：



攻击1



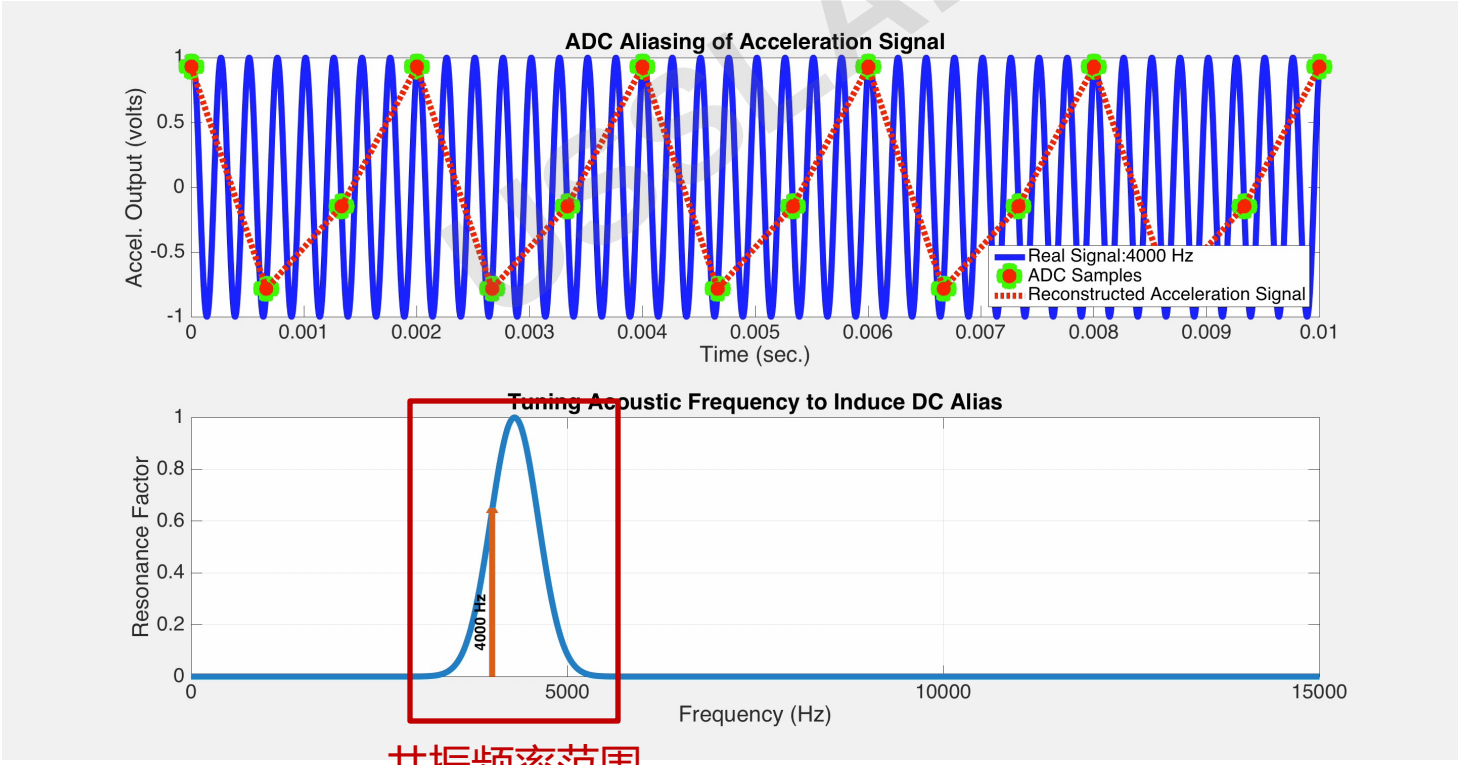
攻击2



USSLAB

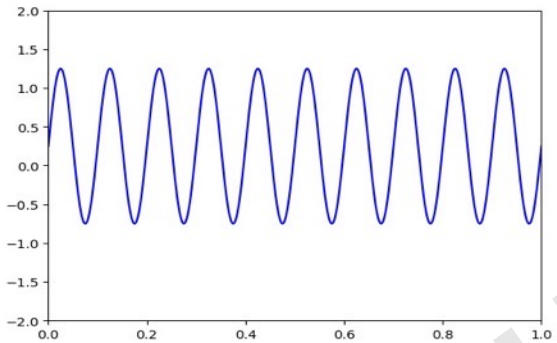
5.3.3 WALNUT攻击：如何找到共振频率

- 通过扫频，改变注入信号的频率，使得输出结果变化（区分于静止）
- ADC倍频：当ADC输出为直流，即ADC采样频率整数倍的 f_s 攻击信号，从而实现后续的精确操控攻击效果
- 注意：共振信号频率通常是一个范围，不是单一频率

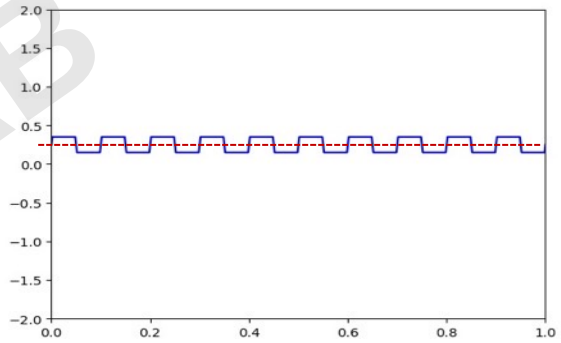


5.3.3 WALNUT攻击1原理分析

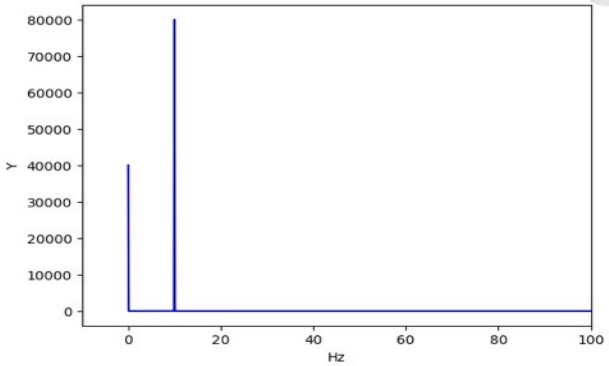
- **脆弱性**：仅利用放大器饱和脆弱性，不利用滤波器和ADC脆弱性
- 假设：由于正常工作需要，放大器设置+0.25偏置（红色虚线）



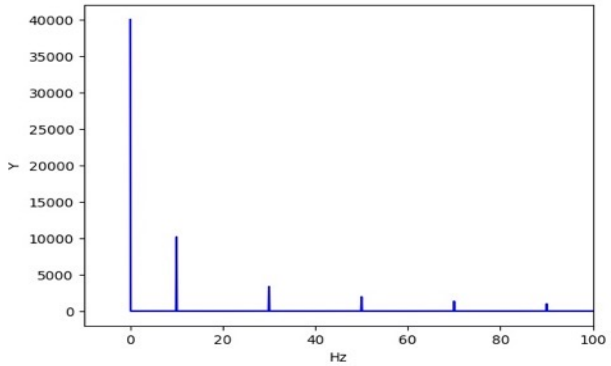
原始信号时域波形



放大器饱和输出信号时域波形



原始信号频谱图



放大器饱和输出信号频谱图

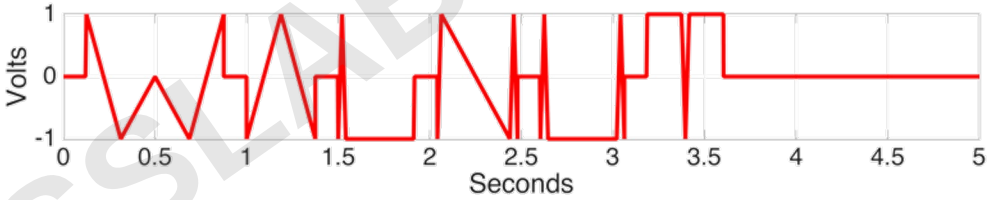
USSLAB

Q: 还有哪些情况可以输出直流?

5.3.3 WALNUT攻击2原理分析

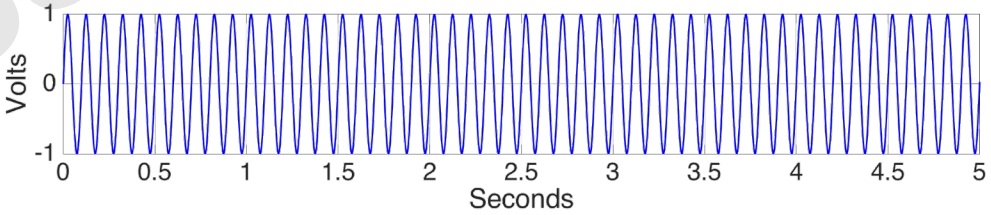
- **脆弱性**：利用滤波器非完美滤波、ADC混频脆弱性
- 基于找到的共振频率点 f_s，利用AM构造特定信号，改变加速度计ADC的输出结果

期望的加速度计输出信号



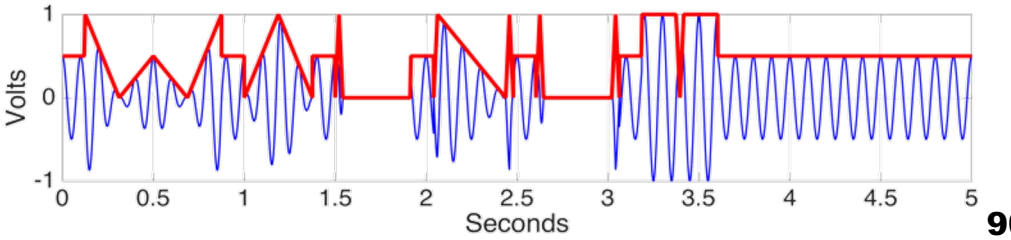
\+

MEMS 共振频率(载波信号)



=

AM调制后攻击信号



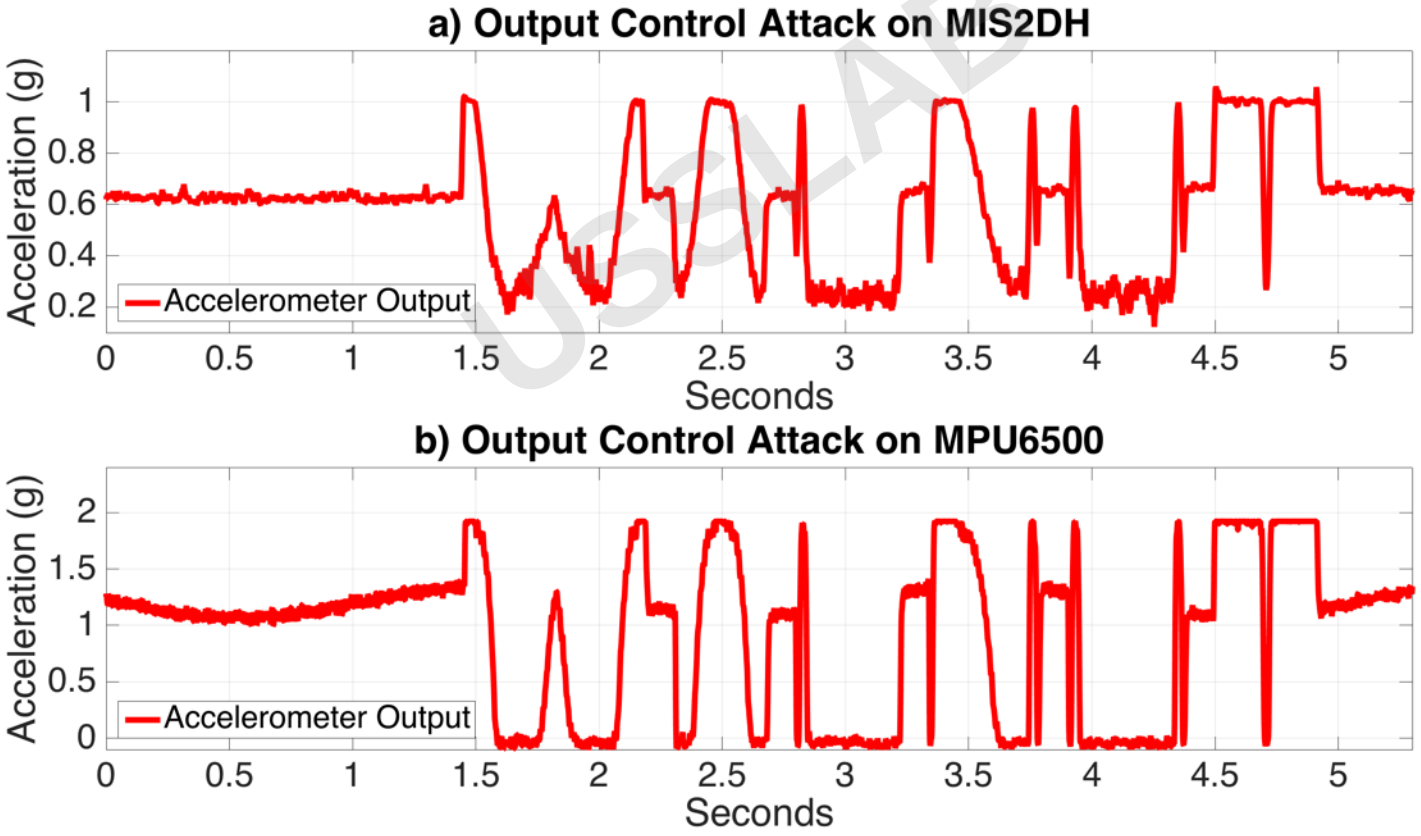
“WALNUT”单词

USSLAB

5.3.3 WALNUT攻击2原理分析

- **脆弱性**：放大器非完美滤波、ADC混频脆弱性

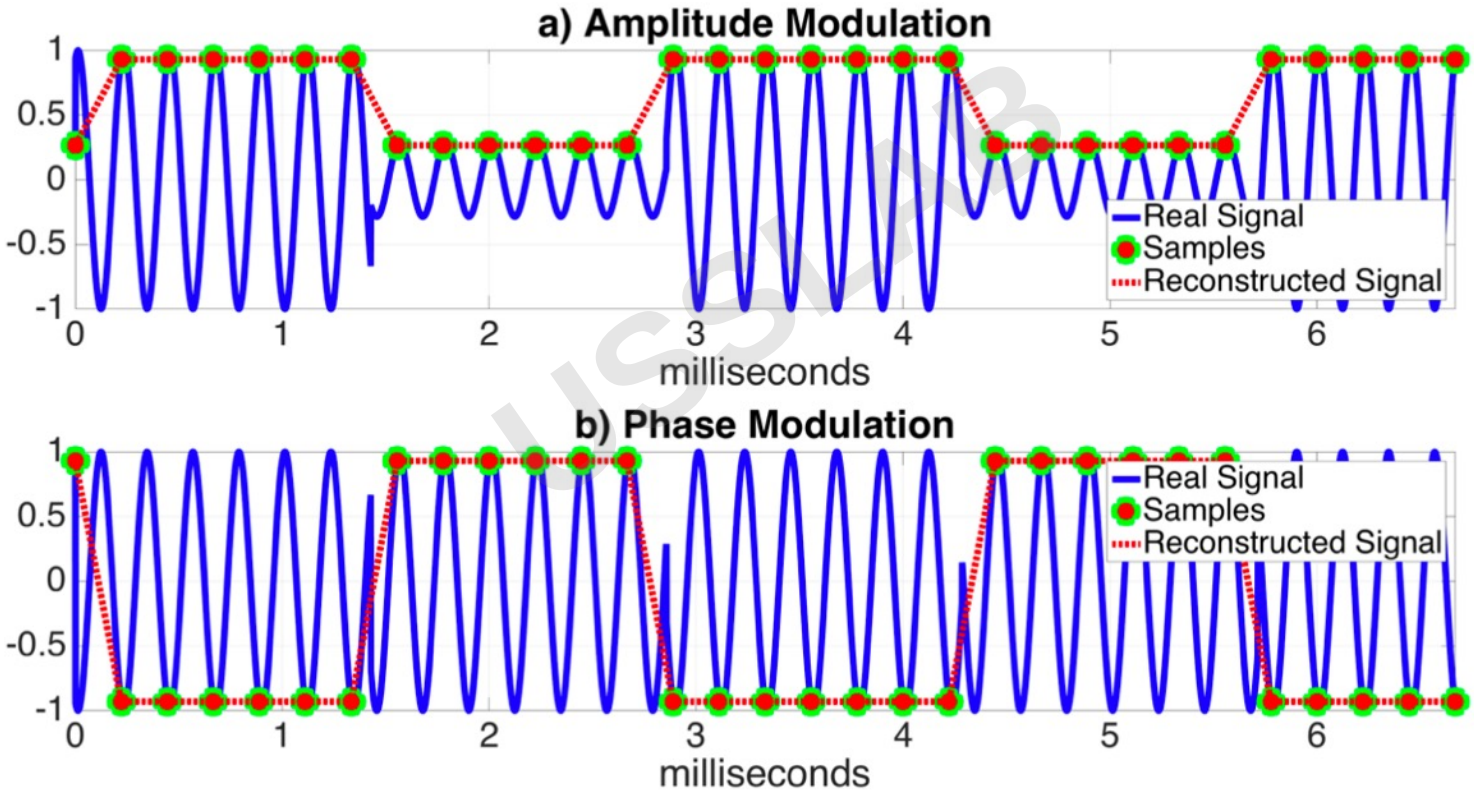
攻击后果：精准操控加速度计输——写出“WALNUT”单词



USSLAB

5.3.3 WALNUT攻击2原理分析

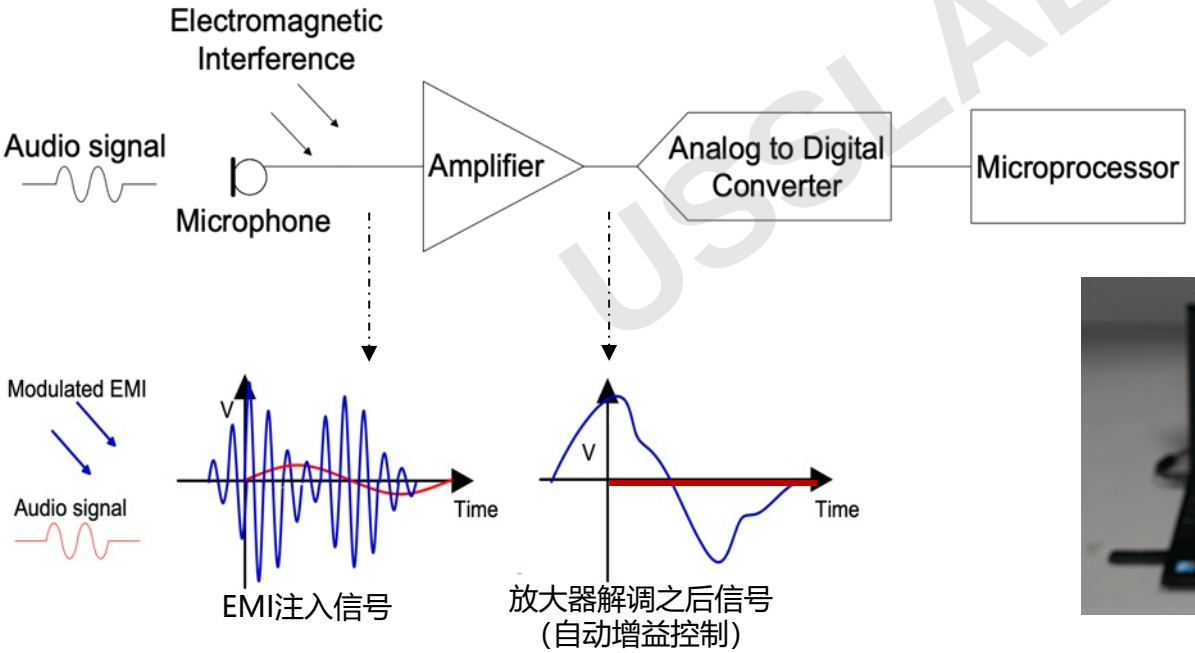
- **思考：**除了AM调制，还可使用PM调制，构造攻击信号



Q：PM调制相对AM，有什么优点?

5.3.3 换能攻击案例 - GhosTalk[1]

- **脆弱性**：信号传输电路耦合电磁干扰（EMI）信号、放大器非线性脆弱性，导致操控传感器输出
- **攻击后果**：向录音笔中无声注入声音、防窃听等



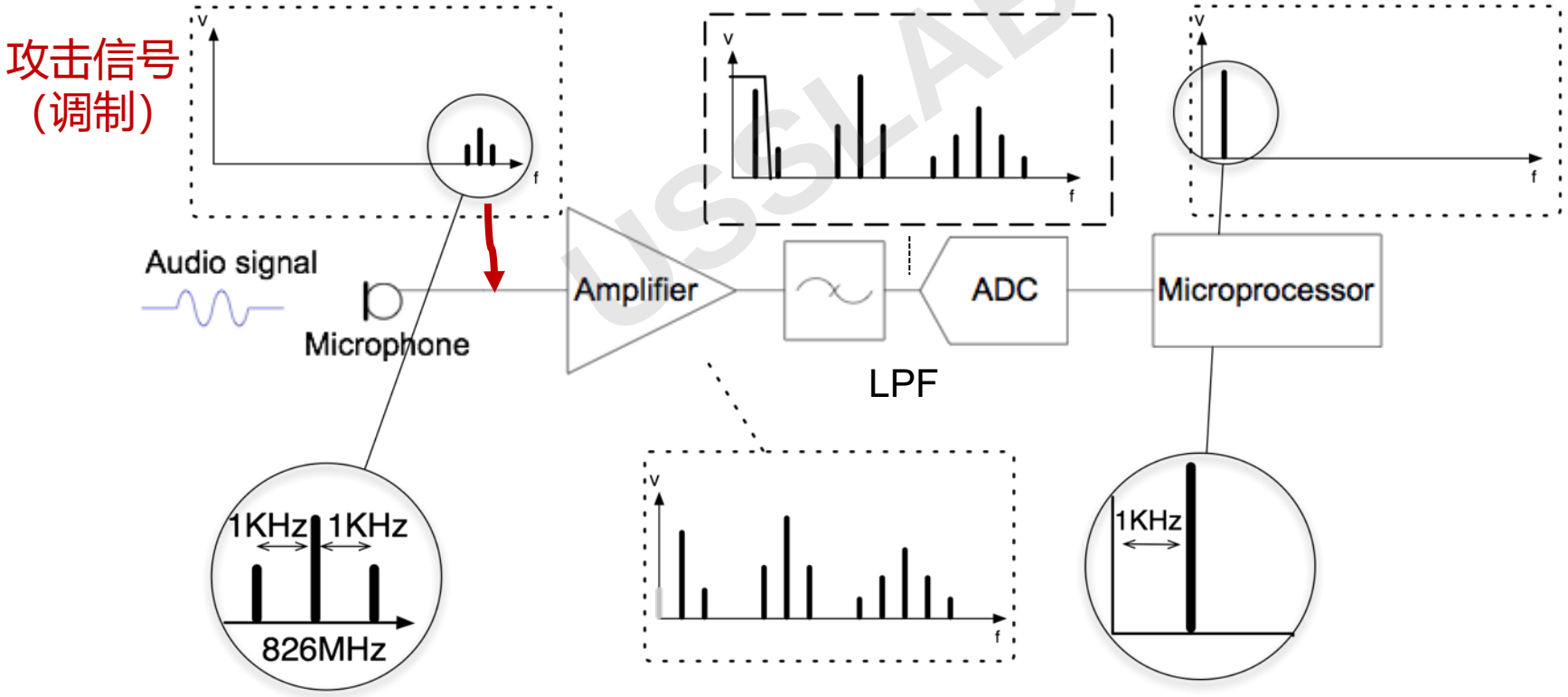
攻击后果：防录音

USSLAB

[1] Kune, Denis Foo, et al. "Ghost talk: Mitigating EMI signal injection attacks against analog sensors." *2013 IEEE Symposium on Security and Privacy*. IEEE, 2013.

5.3.3 换能攻击案例 - GhosTalk[1]

- 1. 信号传输线充当天线，耦合环境EMI干扰信号（调制信号）
- 2. 干扰信号利用放大器的非线性实现解调，还原出攻击信号



USSLAB

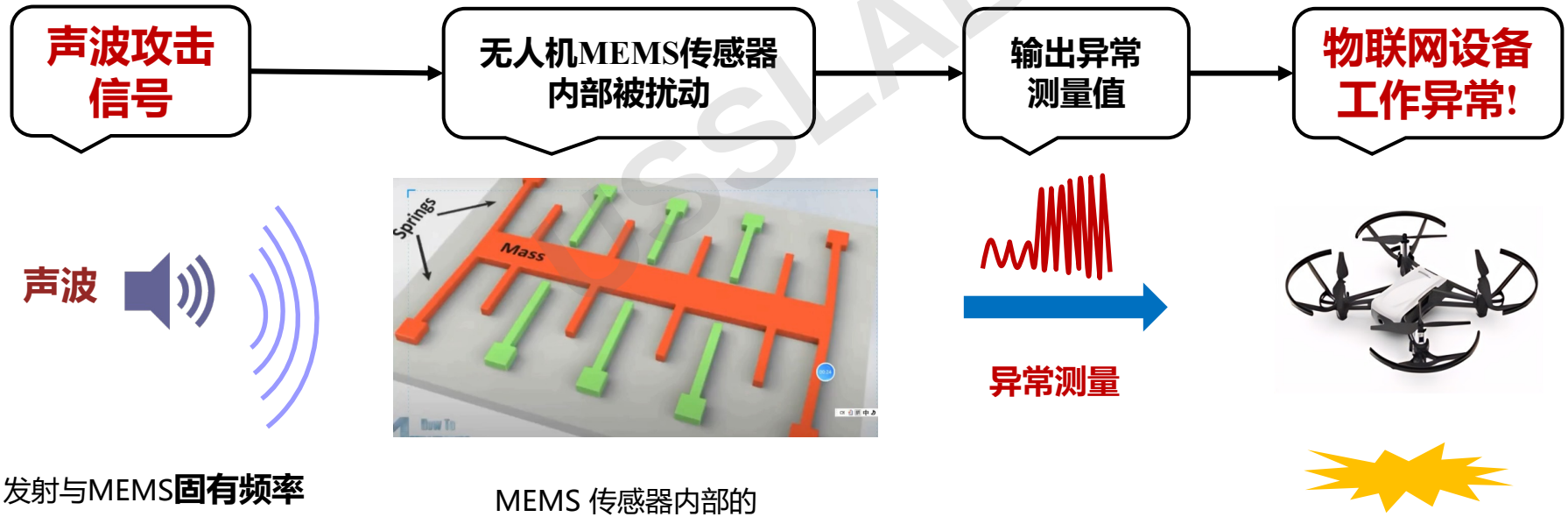
关于GhosTalk类攻击的思考

- 问题1：都有哪些非天线介质可以充当天线?
- 问题2：攻击信号频率和哪些因素有关?
- 问题3：攻击信号是否可以在数字信号线上注入?

USSLAB

延伸案例：声波反无人机

- **MEMS声波共振机理：**指MEMS器件在受到固有频率下的**机械波（如声波）**干扰时，会呈现共振，输出异常值的现象。



SSLAB

延伸案例：声波反无人机

- **MEMS声波共振机理：**指MEMS器件在受到固有频率下的**机械波（如声波）**干扰时，会呈现共振，输出异常值的现象。

延伸案例：声波反机器人

- **MEMS声波共振机理：**指MEMS器件在受到固有频率下的**机械波（如声波）**干扰时，会呈现共振，输出异常值的现象。

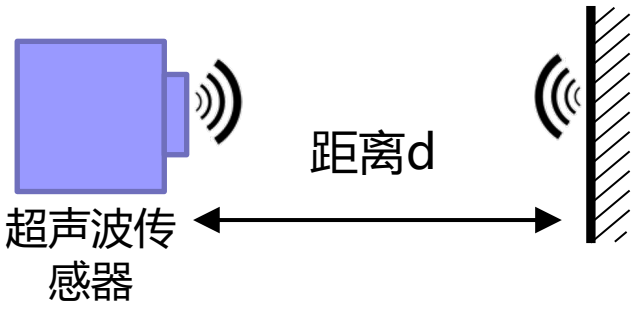
5.3.3 换能攻击案例 – 超声波雷达传感器[1]

- **脆弱性：超声波避障传感器回波信号无鉴权**[3]
- **攻击后果：**通过欺骗传感器可以导致自动驾驶汽车在遇到障碍物不能停车或者在没有障碍物的情况下紧急制动，从而造成事故的发生。

USSLAB

超声波避障传感器的工作方式

- 1) 发射超声波和接收回声。
- 2) 测量传播时间t_e(飞行时间)。
- 3)计算距离



$$d = 0.5 \cdot t_e \cdot c$$
(c是声速)

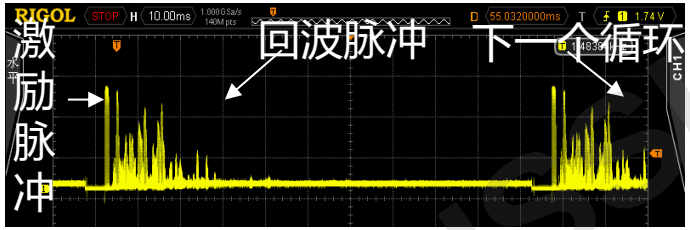
[1] Wenyuan Xu, etc., “Analyzing and enhancing the security of ultrasonic sensors for autonomous vehicles”, *IEEE Internet of Things Journal, 2018*

5.3.3 换能攻击案例 – 超声波雷达传感器

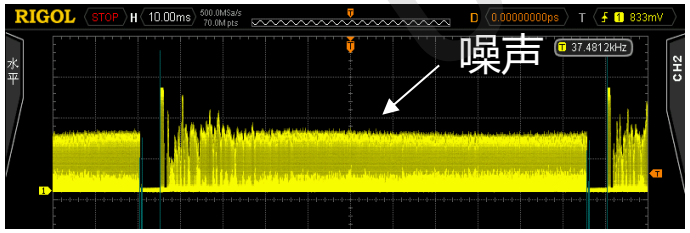
- **特斯拉超声波避障传感器攻击**

攻击方法：1）阻塞攻击（Jamming）

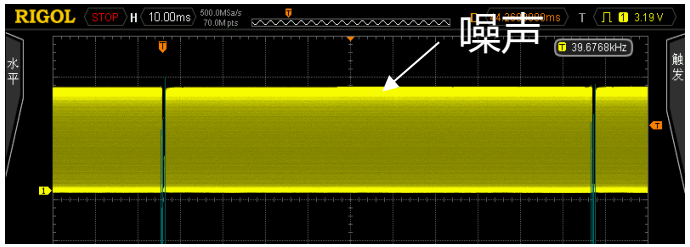
无jamming



弱Jamming



强Jamming



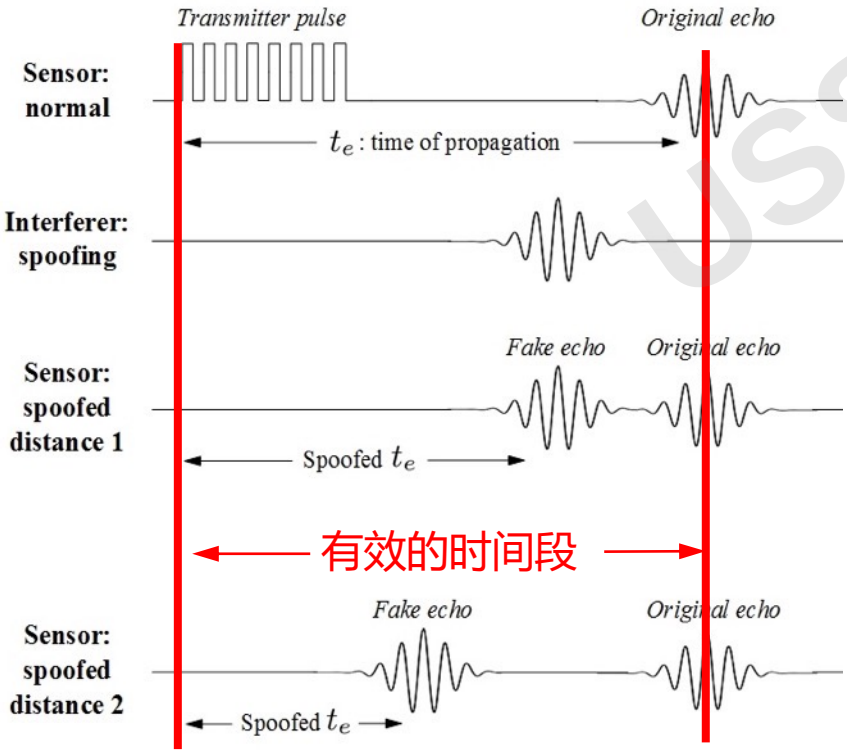
结果：该停停不下来

HSLAB

5.3.3 换能攻击案例 – 超声波雷达传感器

- **特斯拉超声波避障传感器攻击**

  攻击方法：2）欺骗攻击（Spoofing）



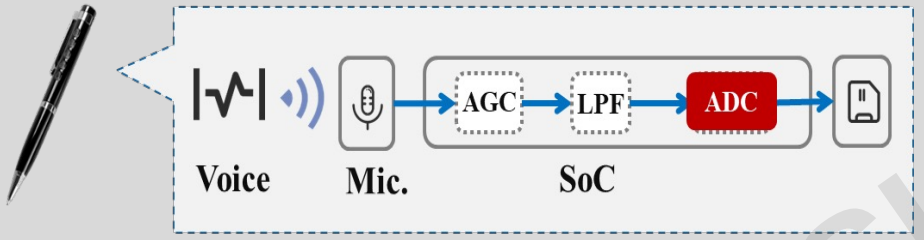
结果：不该停却停下来

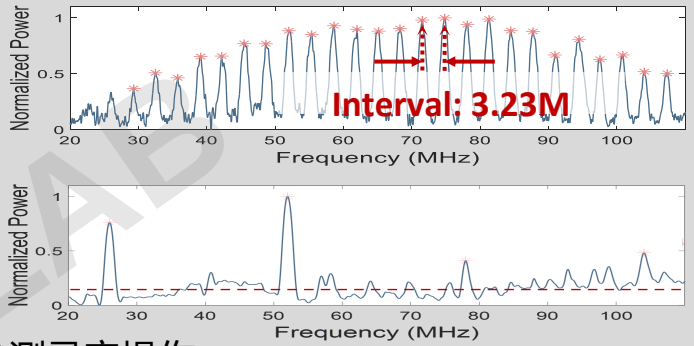
USSLAB

5.3.3 传感器脆弱性 – 传感器侧信道

- 案例：DeHeRec [1], EMeye [2]

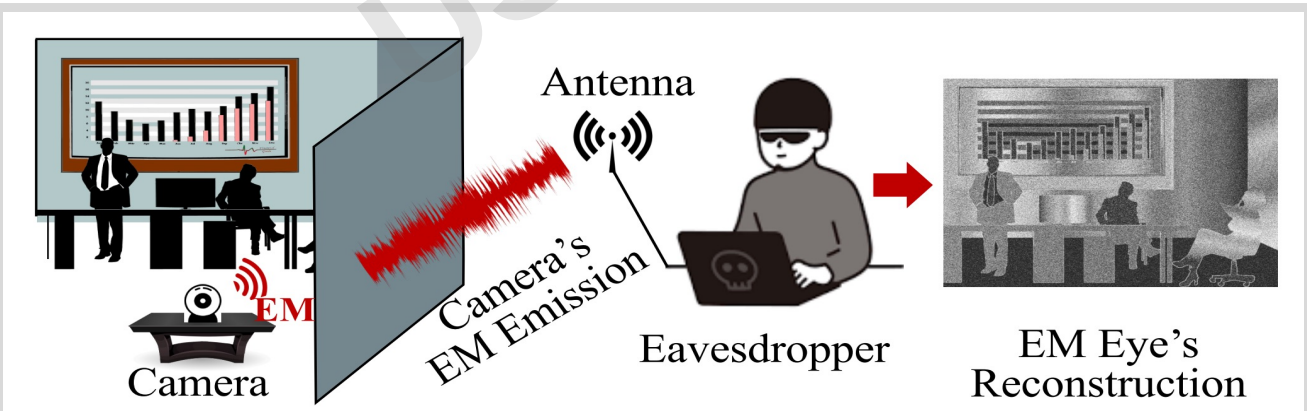
DeHeRec





利用录音笔ADC电磁侧信道检测录音操作

EMeye



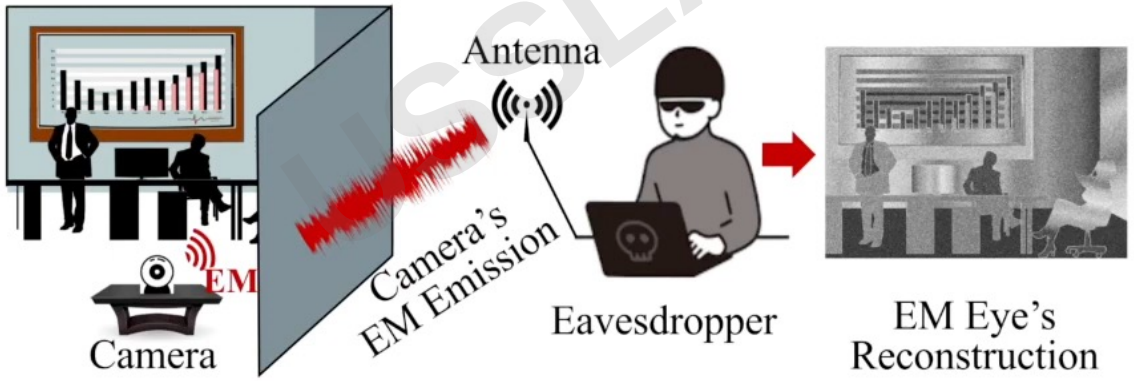
利用摄像头Camera sensor - CPU传输线电磁侧信道重建拍画面

USSLAB

[1] Zhou, Ruochen, et al. "DeHiREC: Detecting Hidden Voice Recorders via ADC Electromagnetic Radiation." IEEE S&P 2023.
[2] Long, Yan, et al. EM Eye: Characterizing Electromagnetic Side-channel Eavesdroping on Embedded Cameras. NDSS 2024.

EM Eye: Electromagnetic Side-channel Eavesdropping on Embedded Cameras

Network and Distributed System Security (NDSS) 2024



Contact Authors

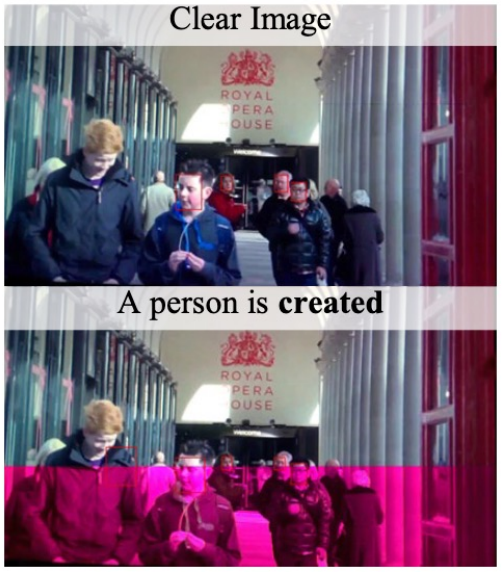
USSLAB

摄像头：图像篡改

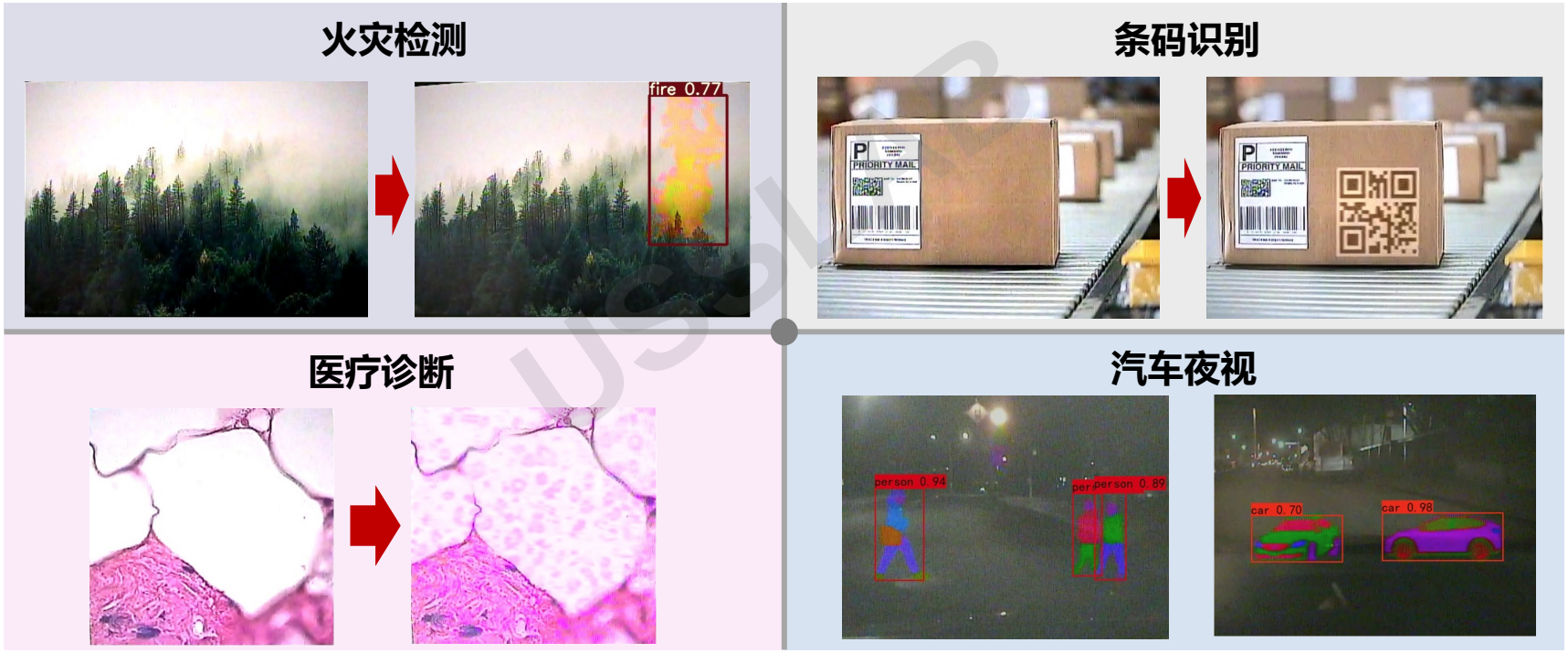
- 分析**相机成像工作链路中的电磁耦合漏洞**，通过电磁干扰非接触实现任意图案的**可控注入**，并欺骗计算机视觉系统

"GhostShot: Manipulating the Image of CCD Cameras with Electromagnetic Interference", NDSS 2025

摄像头：图像篡改

隐藏物体	隐藏人脸	创建人脸
		

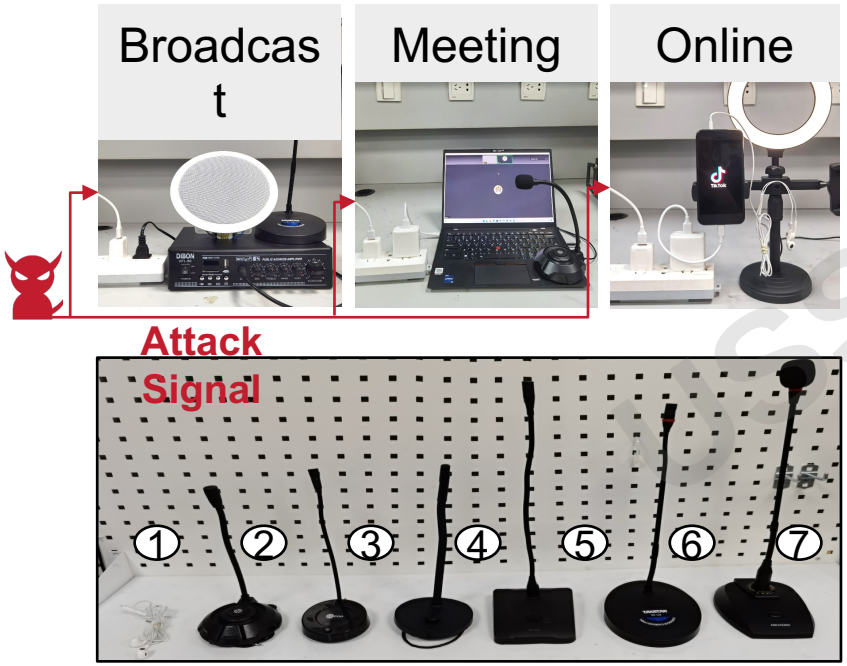
摄像头：图像篡改



USST

PowerRadio：有线注入

- Attack against broadcast system to inject malicious audios.



Experimental Setup

In-room attack

Across-room attack

Jiang, Y., Ji, X., Jiang, Y., Wang, K., Xu, C., & Xu, W. (2024). Powerradio: Manipulate sensor measurementvia power gnd radiation. *NDSS'25*

传感器安全模型

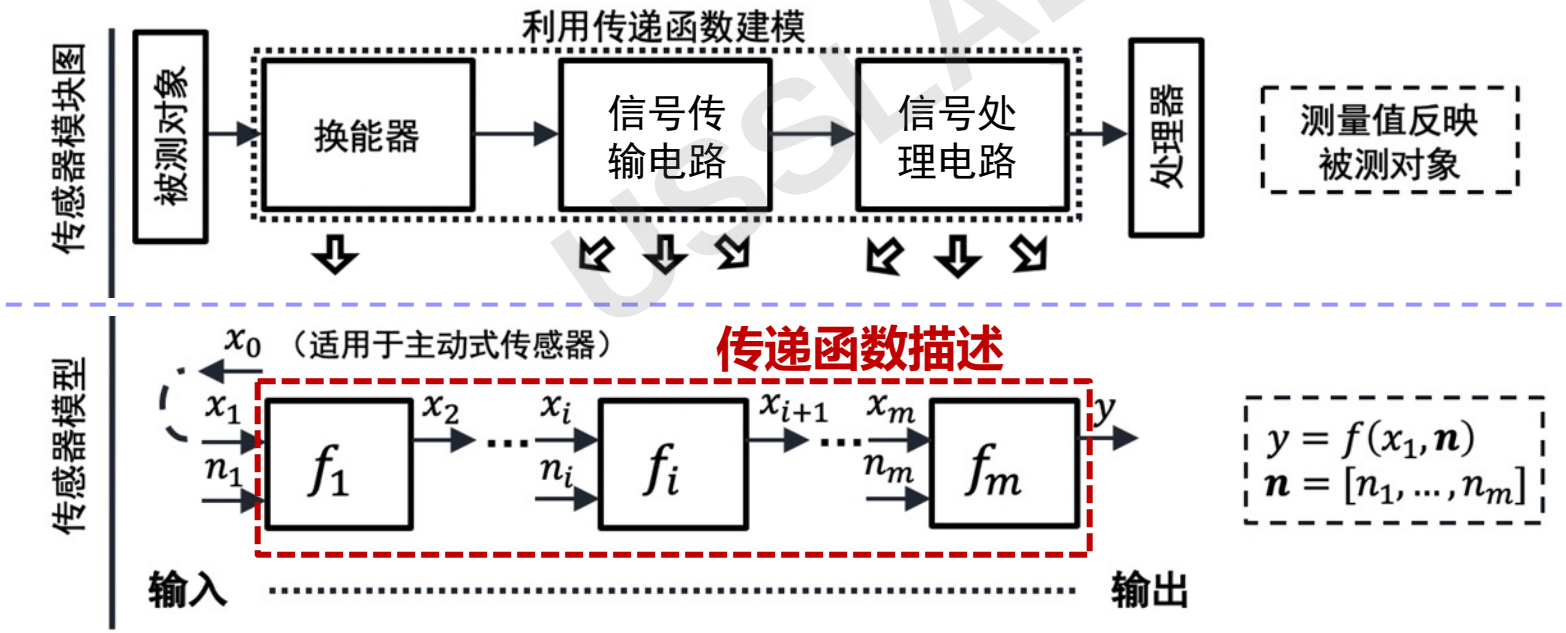
传感器换能攻击如何用统一模型描述?

©SSLAB

5.3.4 传感器安全模型

- **传感器模型：基于传递函数形式化描述**

传递函数：是用来拟合或描述系统的输入与输出之间关系的数学表示。



注： x_0是针对主动式传感器而言，工作需要发射信号，如超声波雷达传感器

USSLAB

5.3.4 传感器安全模型

- **传感器传递函数**

单级传递函数输入：正常待测信号 x_i 和噪声信号 n_i

单级传递函数输出：信号为 x_{i+1}。本课程将传感器简化为一个由构成组件组成的级连模型，x_{i+1} 是下一个（第 $i+1$ 个）组件的输入，因此：

$$\boldsymbol{x_{i+1} = f_i(x_i, n_i)}$$

所以，**传感器系统级**传递函数表示为：

$$y = f_m(\cdots f_2(f_1(x_1, n_1), n_2)\cdots, n_m)$$

其中，y 是传感器的测量值，m 是级连的传感器组件数量，传感器的输入包括换能器处接收的物理激励信号 x_1 以及每个组件处接收的噪声n。定义所有组件收到的噪声表示为一个向量：

$$\boldsymbol{n} = [n_1, n_2, \ldots, n_m]$$

将整个**传感器的传递函数**表示为 $\boldsymbol{f}$，那么传感器传递函数可以简化为：

$$y = f(x_1, \boldsymbol{n})$$

USSLAB

5.3.4 传感器安全模型

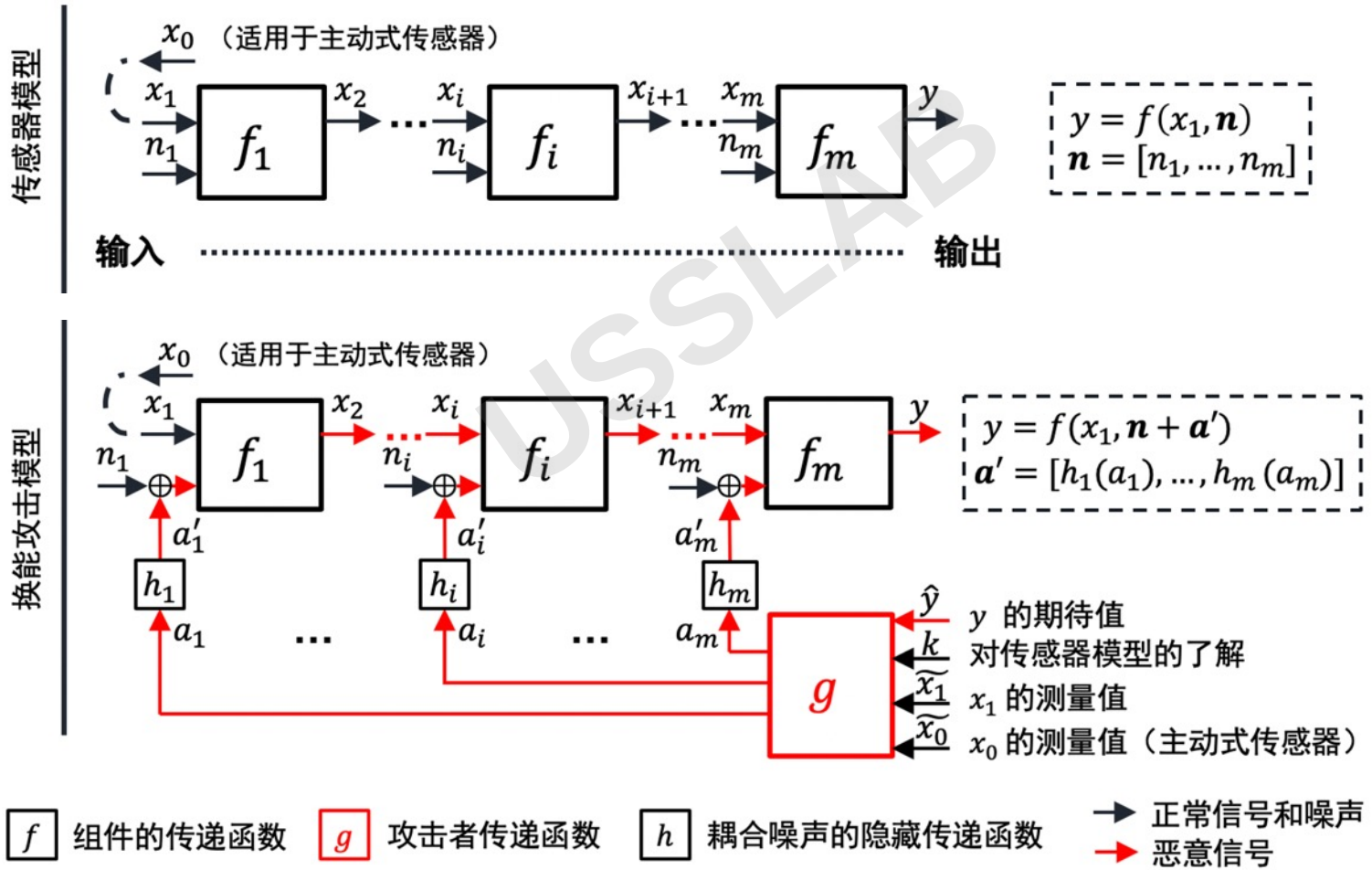
- **传感器模型传递函数（续）**

- 一些传感器常用的传递函数，以最常用的加性噪声为例:
  - **线性传递函数：**可表示为 $x_{i+1}= c_0 + c_1(x_i + n_i)$，其中 c_0 是函数的截距，c_1 是斜率。线性的传递函数被用于描述部分传感器组件的理想或简化状态，例如线性放大器。
  - **非线性传递函数：**适用于描述大部分传感器
    - 光电二极管：对数函数 $x_{i+1}= c_0 + c_1 \ln(x_i + n_i)$
    - 热敏电阻：指数函数 $x_{i+1}= c_1 e^{k(x_i+n_i)}$
    - 硅电阻：幂函数 $x_{i+1} = c_0 + c_1(x_i + n_i) + c_2(x_i + n_i)^2$
  - 其他情况：可以用更高阶的幂级数描述。

DSSLAB

5.3.4 传感器安全模型

- **换能攻击模型**



USSLAB

5.3.4 传感器安全模型

传感器正常测量传递函数

$$y = f(x_1, \boldsymbol{n})$$
$$\boldsymbol{n} = [n_1, n_2, \ldots, n_m]$$

- **换能攻击模型**

在换能攻击下，传感器测量值可以表示为：

$$y = f(x_1, \boldsymbol{n} + \boldsymbol{a}')$$
$$\boldsymbol{n} + \boldsymbol{a}' = [n_1 + a_1', n_2 + a_2', \ldots, n_m + a_m']$$
$$a_i' = h_i(a_i)$$

其中，$\boldsymbol{n}$是正常测量噪声，$\boldsymbol{a}'$是攻击者向传感器注入的恶意噪声向量。

a_i'是向第 i 个传感器组件注入的噪声，它来自于攻击者产生的物理信号 a_i 并经过该组件传递函数h_i 转换得到。

h_i 一般同时描述来自正常噪声信号 n_i 和恶意噪声信号a_i' 的耦合过程，是一个**隐藏传递函数，可以表示跨场也可以是超限利用。**

USSLAB

5.3.4 传感器安全模型

- **换能攻击模型**

- **注入信号**：注入的信号 a_i' 和正常的噪声信号 n_i 结合（通常为叠加）后的信号影响被干扰组件的输出和传感器的测量值。
- 攻击者产生的攻击信号可以用一个传递函数$g(*)$表示：

$$a = [a_1, a_2, \ldots, a_m] = g(\hat{y}, k, \widetilde{x_1}, \widetilde{x_0})$$

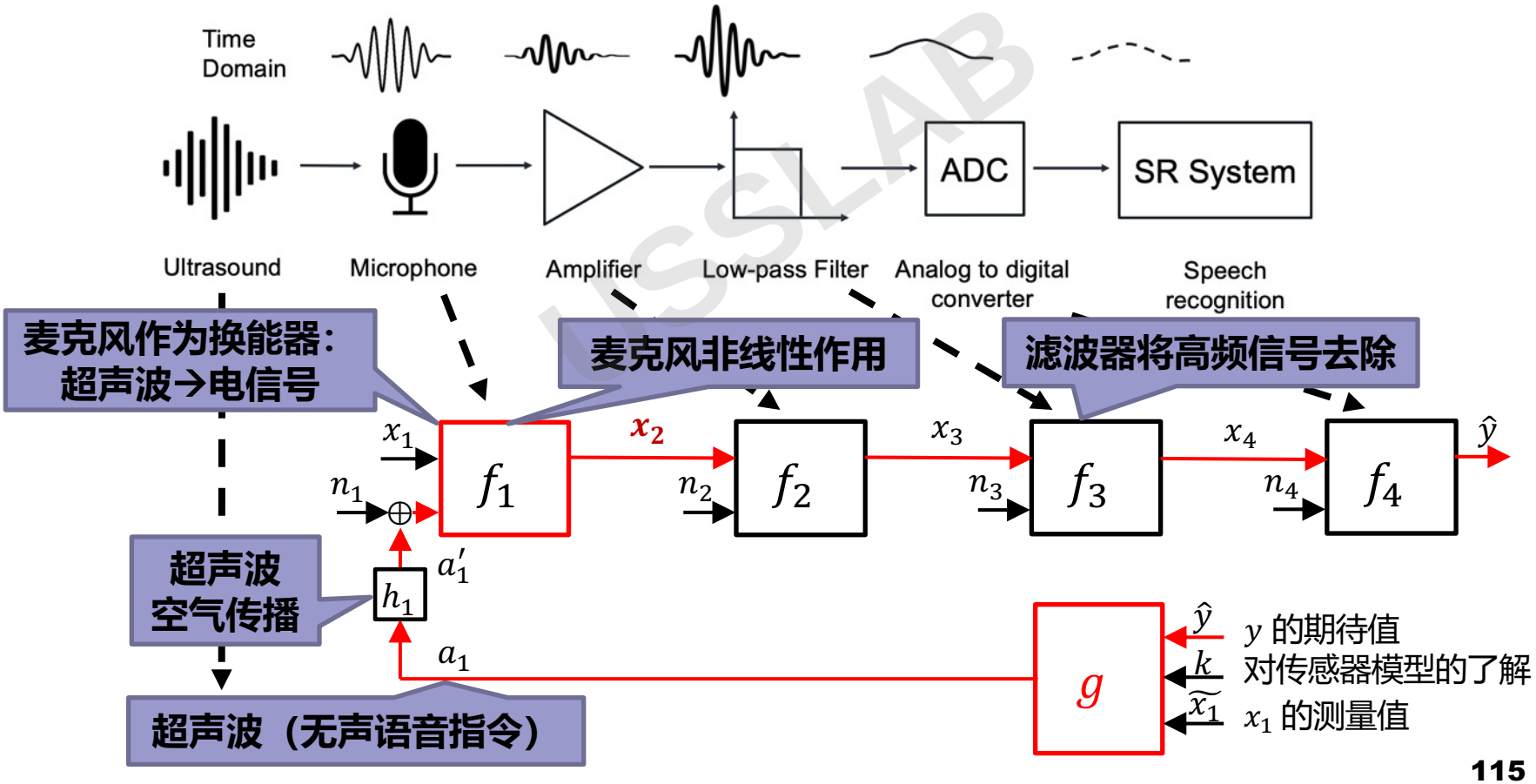
  - $\hat{y}$ 是攻击者想要制造的传感器**期望测量值**
  - k 代表攻击者对目标传感器模型的了解知识
  - $\widetilde{x_1}$ 是攻击者对目标传感器所测量的物理量的测量值
  - 当目标传感器为主动式传感器时，攻击者还需要测量目标传感器发出的物理激励 $\widetilde{x_0}$ 以实现对其测量值的完全控制。

USSLAB

5.3.4 传感器安全模型

- 海豚音攻击（DolphinAttack）换能攻击传递函数模型

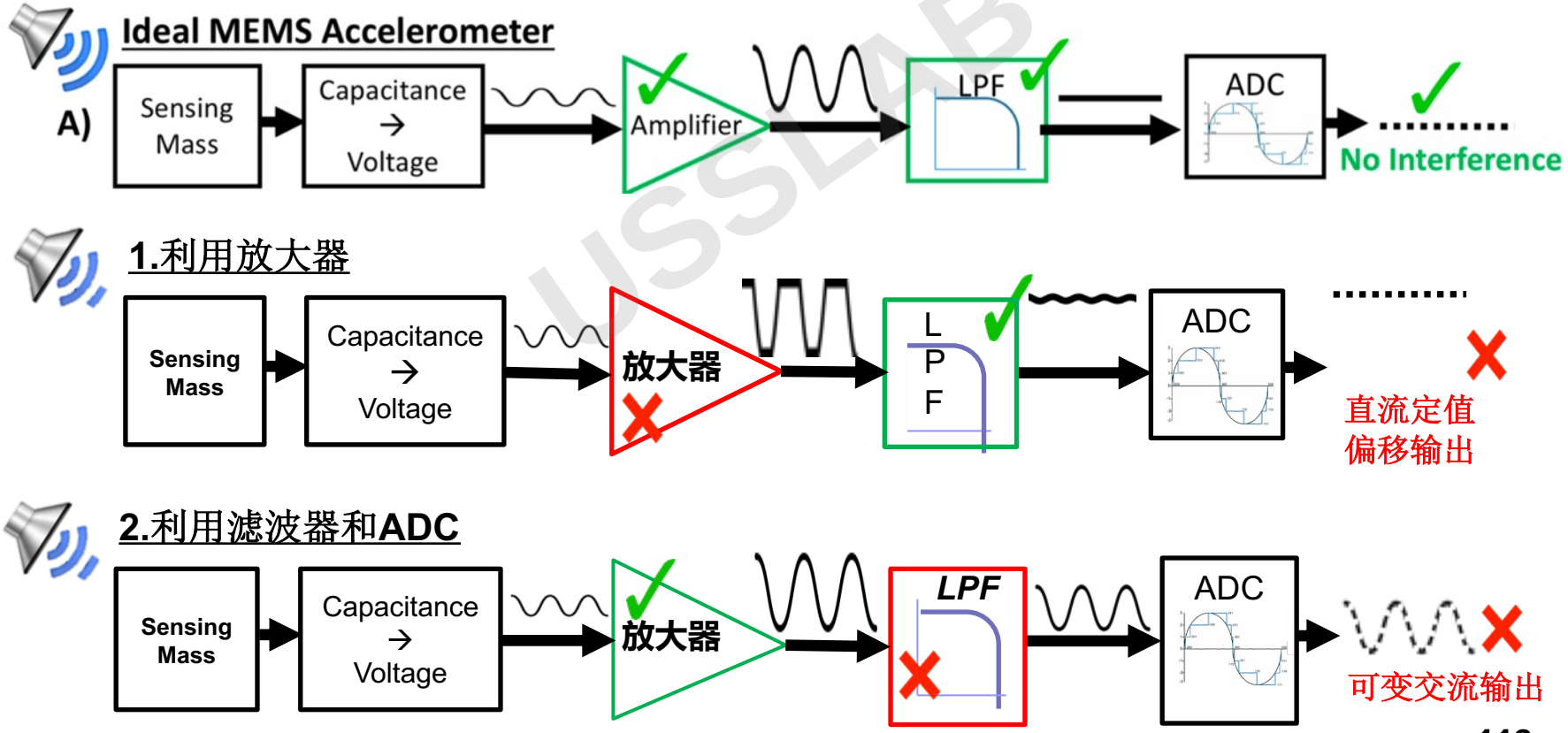
换能攻击模型描述



5.3.4 传感器安全模型

- Walnut换能攻击传递函数模型

攻击原理：（1）利用放大器饱和效应，（2）滤波器非完美滤波+ADC混频

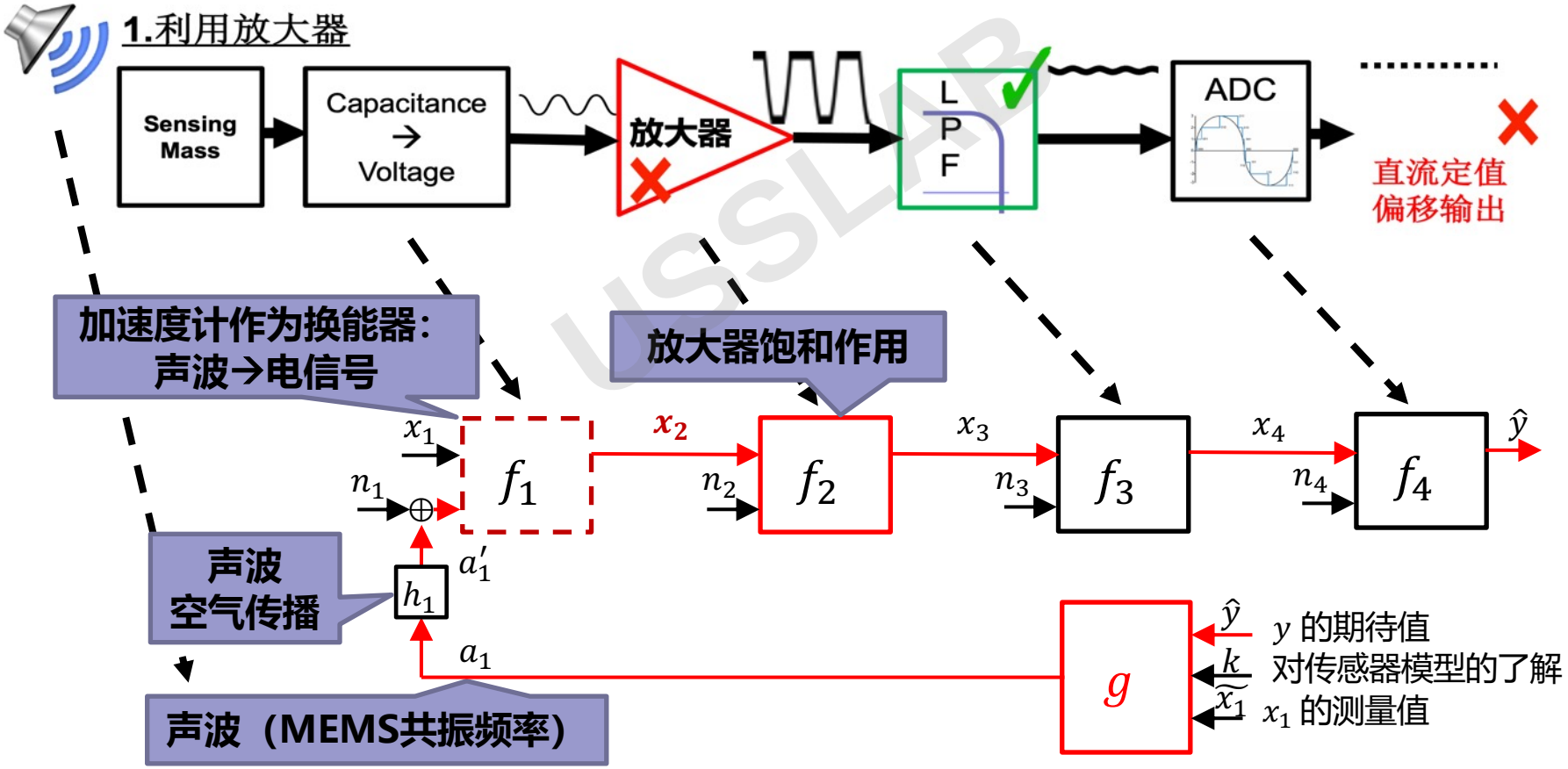


USSLAB

5.3.4 传感器安全模型

- **Walnut换能攻击传递函数模型**

换能攻击模型描述——方法1：利用放大器饱和效应

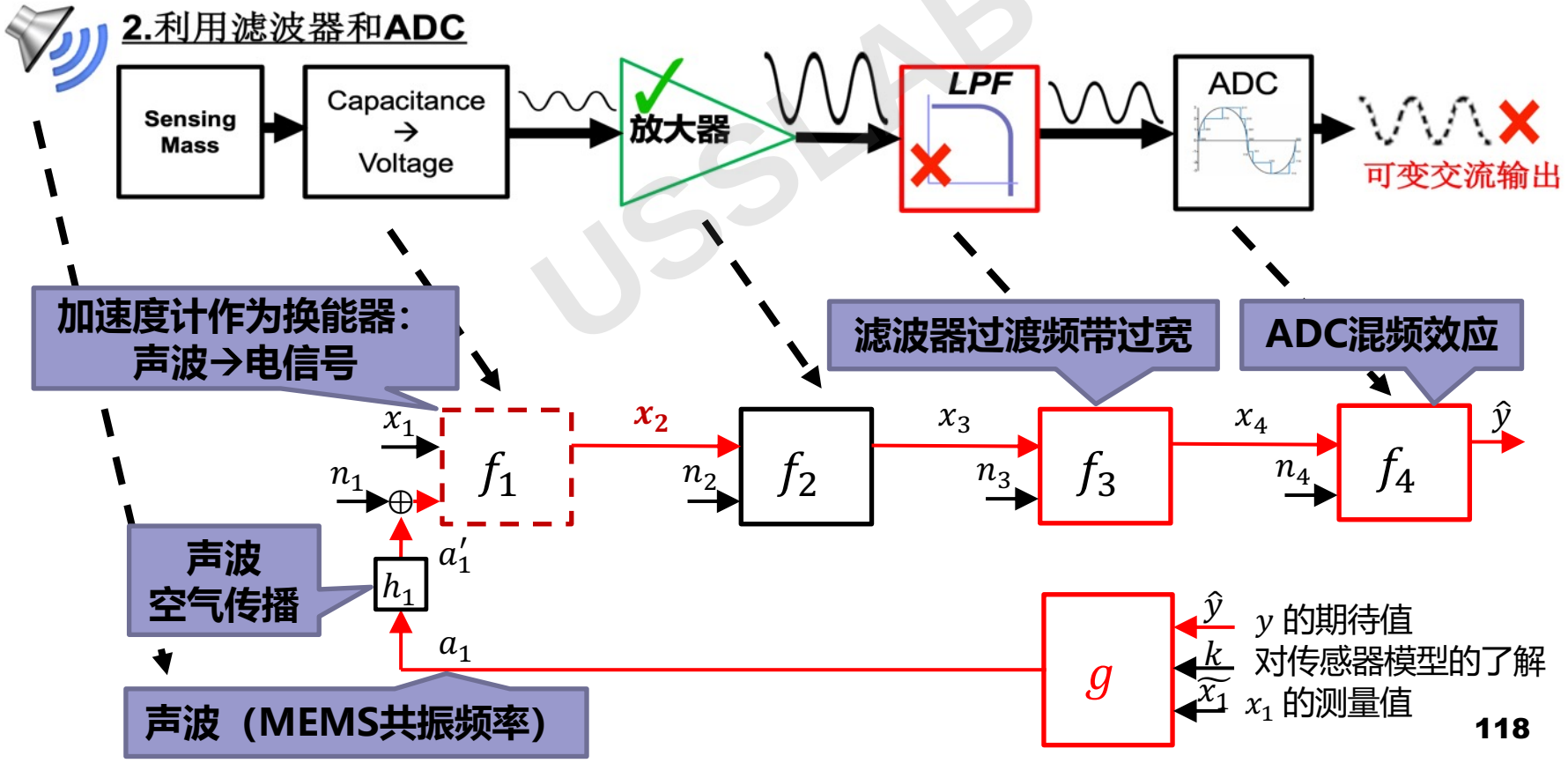


USSLAB

5.3.4 传感器安全模型

- **Walnut换能攻击传递函数模型**

换能攻击模型描述——方法2：滤波器不完美特性+ADC混频特性

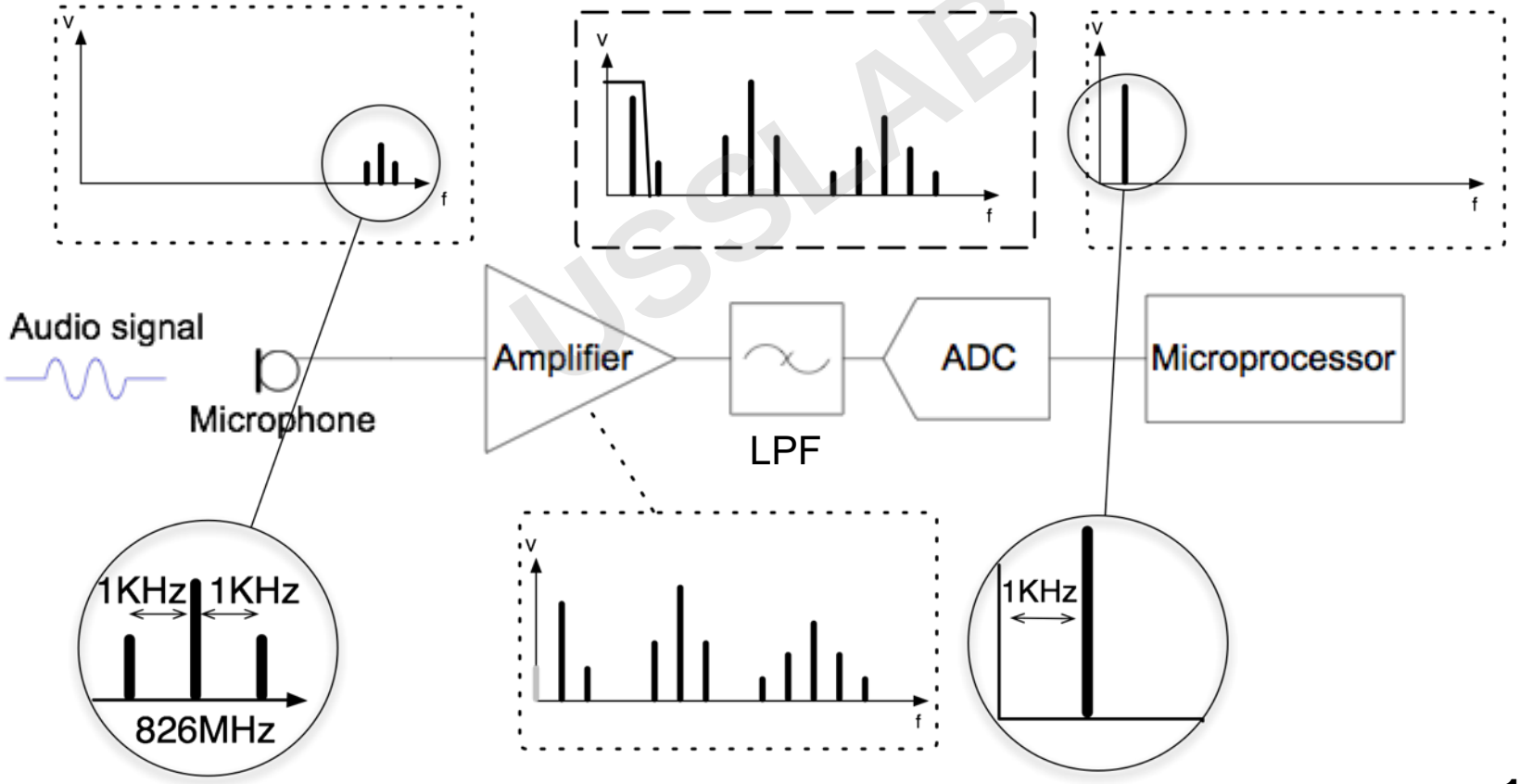


USSLAB

5.3.4 传感器安全模型

- GhosTalk换能攻击传递函数模型

攻击方法：**利用电路EMI耦合特性、放大器的非线性解调**

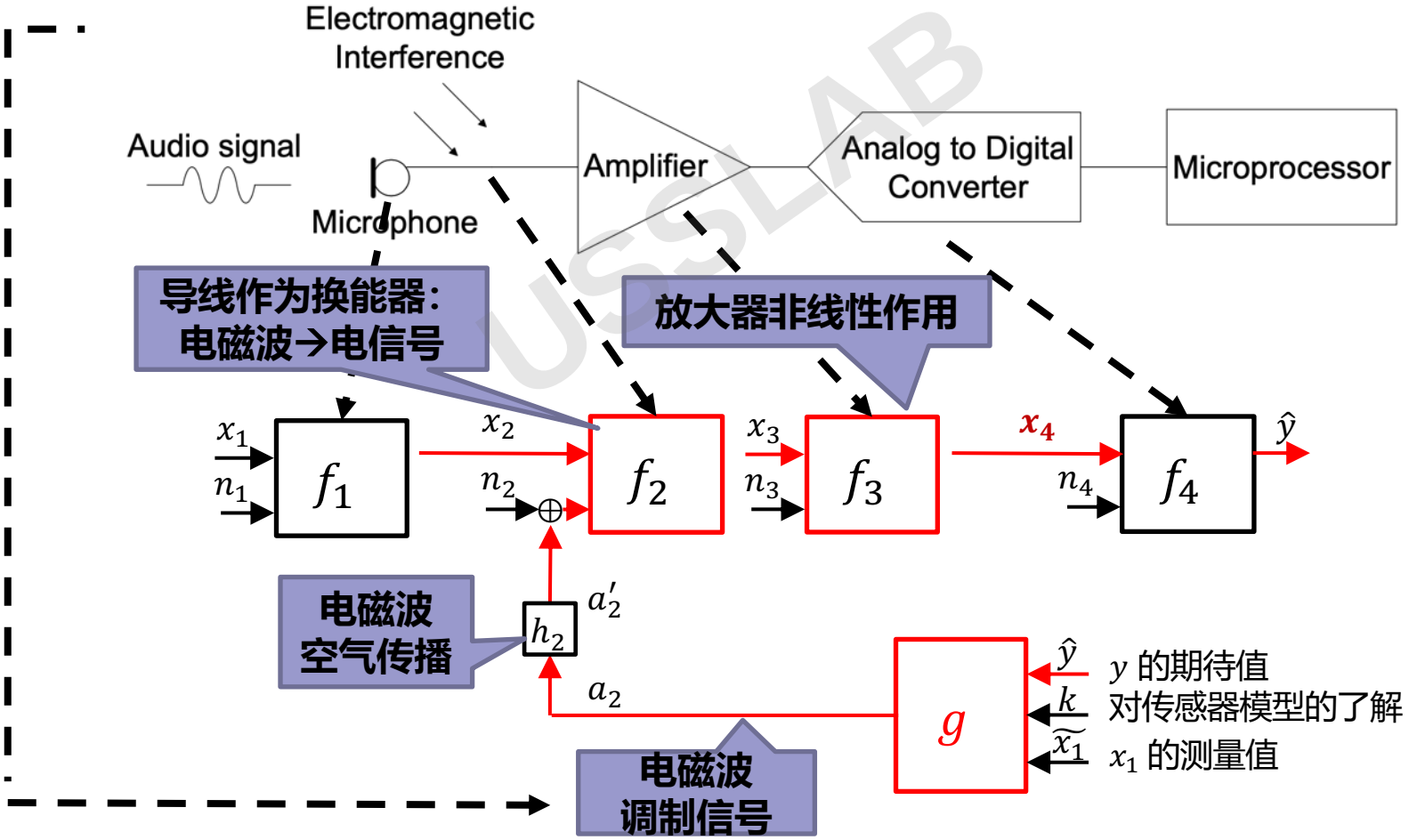


USSLAB

5.3.4 传感器安全模型

- GhosTalk换能攻击传递函数模型

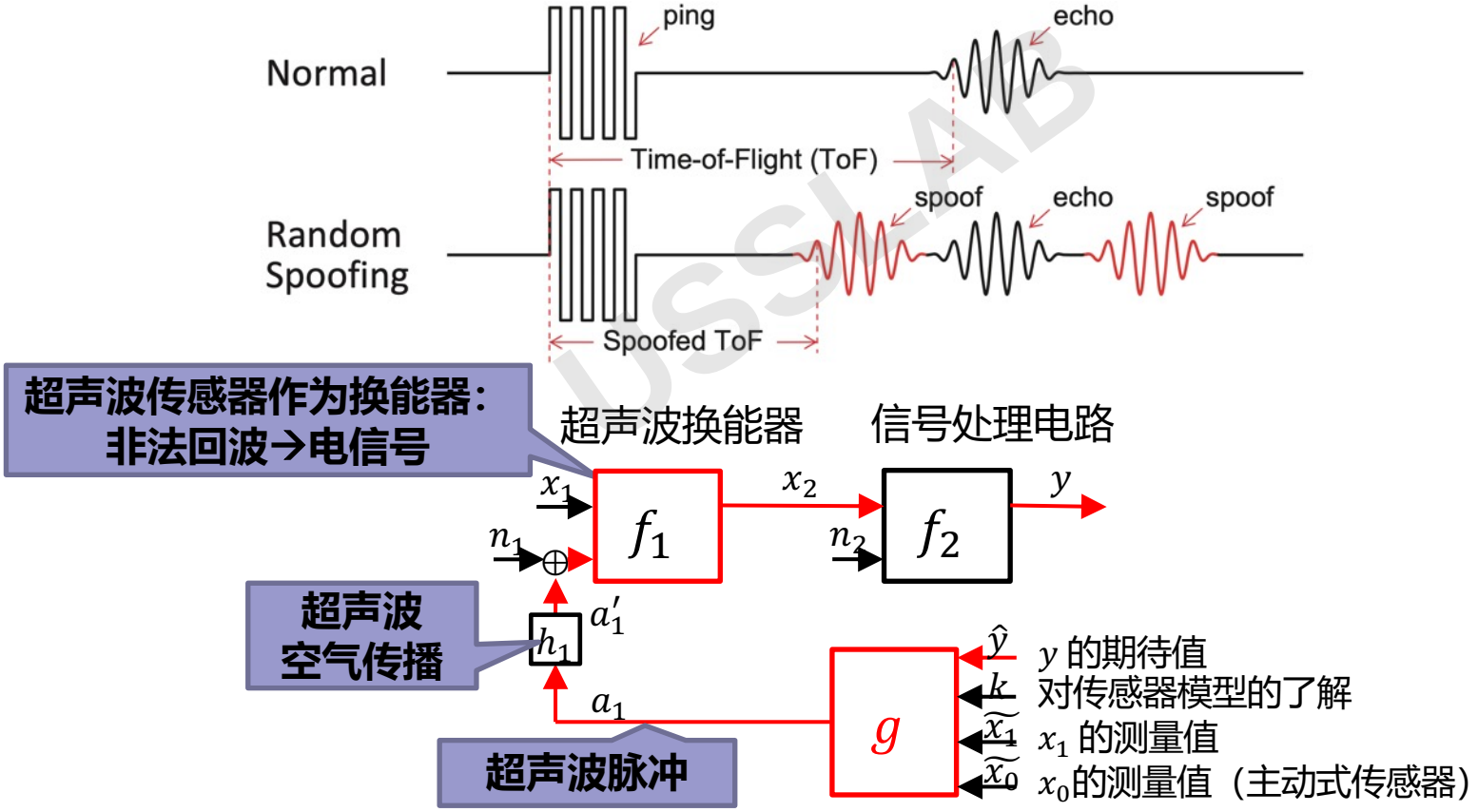
换能攻击模型描述——利用电路导线EMI耦合+放大器非线性解调



5.3.4 传感器安全模型

- **超声波雷达传感器换能攻击传递函数模型**

换能攻击模型描述



USSLAB

5.3.5 传感器换能攻击防护

- 防护方式：**攻击检测**与**攻击抵御**
  - 攻击检测旨在检测到换能攻击的存在；
  - 攻击抵御则是为了抵御攻击对传感器测量值的影响，让传感器即使在攻击发生时也可以有可信的输出。

- **攻击检测方法**
  - 检测信号注入步骤
    - 带外信号检测
    - 验证执行
    - 输出随机
  - 检测信号生效步骤
    - 饱和检测
    - 交调失真的特征检测
    - ……

- **攻击抵御方法**
  - 屏蔽：如物理隔离等
  - 滤波
  - 随机化
  - 改进组件质量
  - 多传感器融合

USSLAB

5.3.5 传感器换能攻击防护

- **攻击检测方法**
  - **检测信号注入步骤。**可以利用额外的换能器有针对性地对注入的带外信号检测。例如，使用额外的麦克风就可以检测到攻击MEMS加速度计和陀螺仪的共振频率声音。
  - **检测信号生效步骤。**可以利用换能信号和正常信号的差异进行检测。如利用交调失真进行信号解调的攻击可能会在模拟信号中留下可识别的特征。由于这些特征是由交调失真引入，海豚音攻击防护可以通过检测500 Hz – 1 kHz语音信号强度来识别利用交调失真解调的无声语音指令。（例如海豚音攻击文章中防护部分）

USTLAB

5.3.5 传感器换能攻击防护

- **换能攻击防护方法**
  - **屏蔽——物理隔离。**防御者可根据实际的应用场景增加物理隔离以衰减进入传感器的外部物理信号，例如使用法拉第笼来屏蔽电磁波。常见的物理隔离包括屏蔽导线、隔音、光屏蔽。
  - **输出随机化。**当主动式传感器的输出波形变得随机化之后，传感器就可以集中搜寻输入信号中与输出的随机性匹配的有用信号，而由于攻击者并不了解该如何产生与之匹配的随机信号，攻击对传感器的影响变得有限。

USSLAB

传感器与执行器安全

第四节 传感器隐私安全

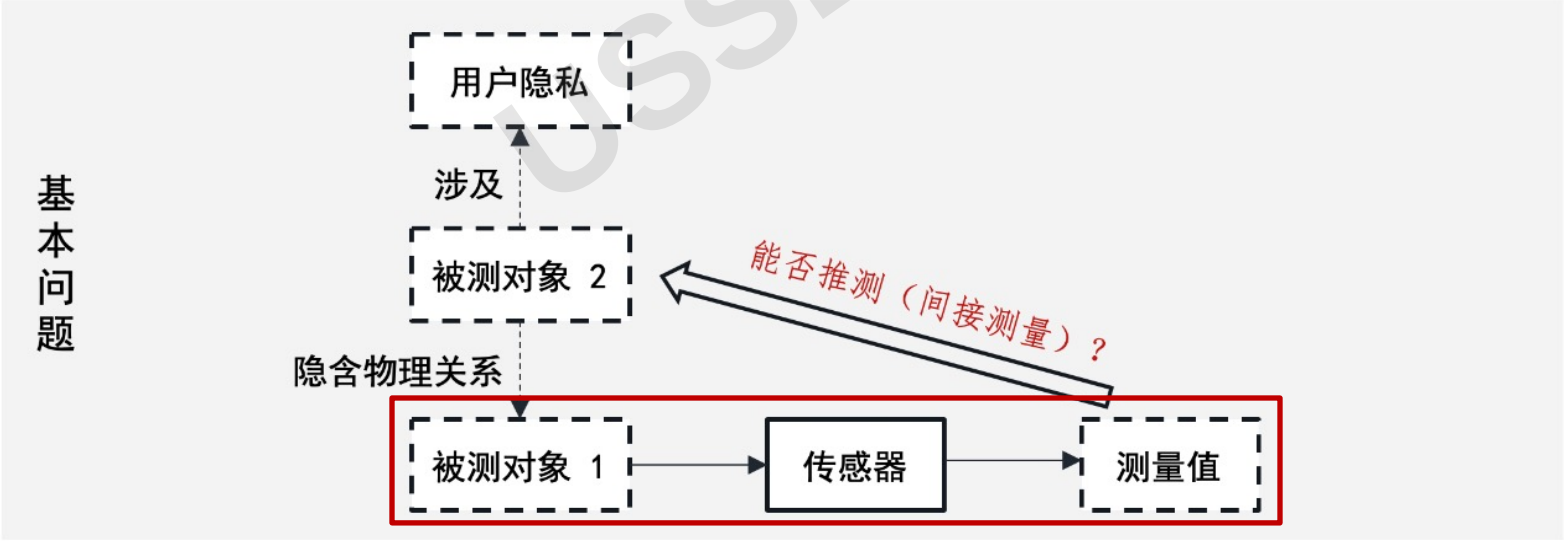
1. 传感器隐私安全定义
2. 案例分析
   - 用户按键推断
   - 用户位置追踪
   - 用户任务推断
   - 用户语音窃听

USSLAB

5.4 传感器隐私安全

- **定义**：攻击者利用低敏感权限传感器推测用户高敏感信息，或其他非直接可测量信息，这些信息导致用户**隐私泄露**

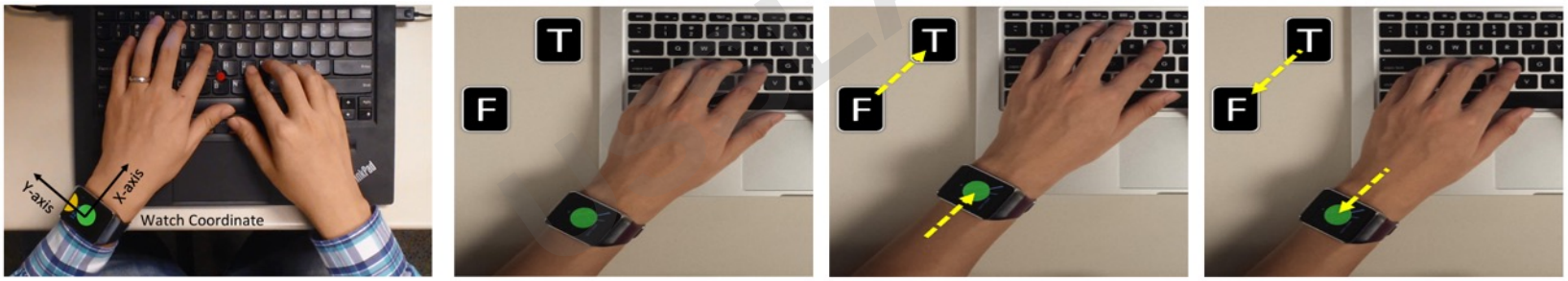
- 低敏感权限传感器数据获取权限容易获得，能被攻击者利用



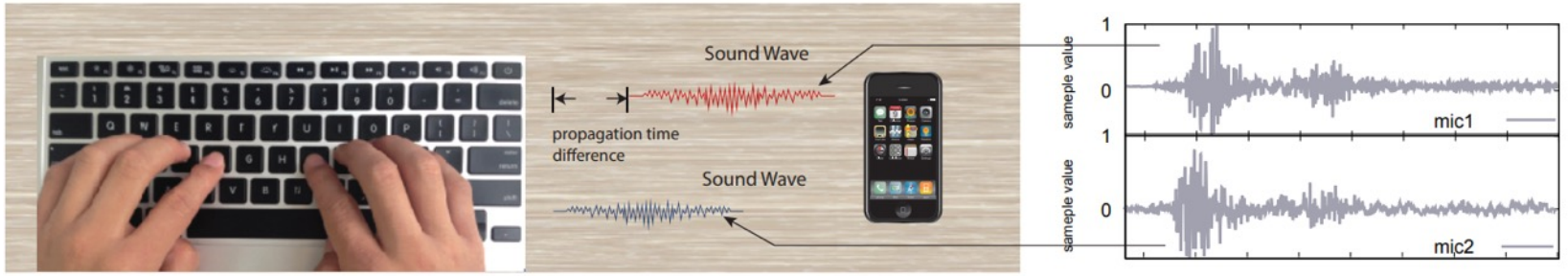
USSLAB

5.4.1 用户按键推断

- **原理：**当用户按键操作时，会产生特定声音或设备位移，从而被各类传感器（加速度计、陀螺仪、麦克风、光传感器等）记录数据，可用于推断智能设备的按键操作。
- **攻击案例：**Mole [1]



利用手部动作推测键盘输入



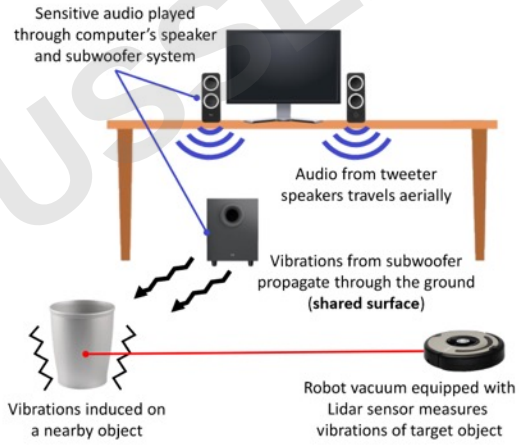
利用打字声音推测键盘输入

[1] Wang, He, Ted Tsung-Te Lai, and Romit Roy Choudhury. "Mole: Motion leaks through smartwatch sensors." MobiCom 2015.

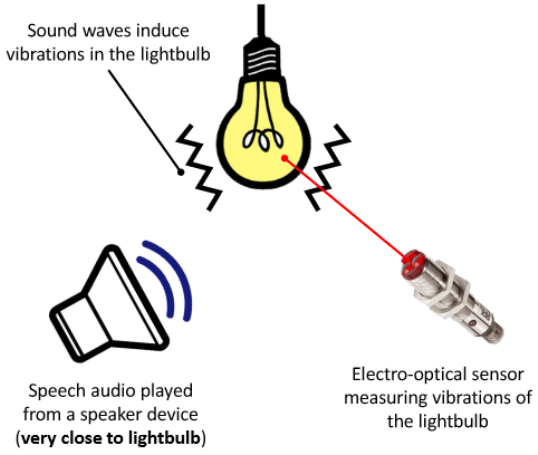
5.4.2 用户语音窃听

- **原理：**语音可以在物体上引起机械振动，被MEMS加速度计、陀螺仪直接感知，或者通过激光雷达、光电传感器间接测量。通过信号重建算法可以重建用户敏感语音信息。
- **攻击案例：**Gyrophone[1], Lidarphone[2], Lamphone[3]

案例1：运动传感器
直接感知音频振动



案例2：激光雷达
间接测量物体共振



案例3：光电传感器
间接测量灯泡振动

USSLAB

[1] Michalevsky, Yan, et, al. "Gyrophone: Recognizing speech from gyroscope signals." *USENIX Security 2014*.
[2] Sami, Sriram, et al. "LidarPhone: acoustic eavesdrop** using a lidar sensor." *Sensys 2020*.
[3] Nassi, Ben, et al. "Lamphone: Passive sound recovery from a desk lamp's light bulb vibrations." *USENIX Security* 2022.

案例：“金唇”窃听

- 金唇：一种利用被动技术传输声音信号的窃听器。
- 它被藏在一个苏联送给美国驻莫斯科大使的礼物中。
- 金唇由外界电磁波激活和驱动，不需安装电池或外接电源，所以被认为是射频识别技术的前身。

OSSLAB

传感器与执行器安全

第五节 控制/执行器安全

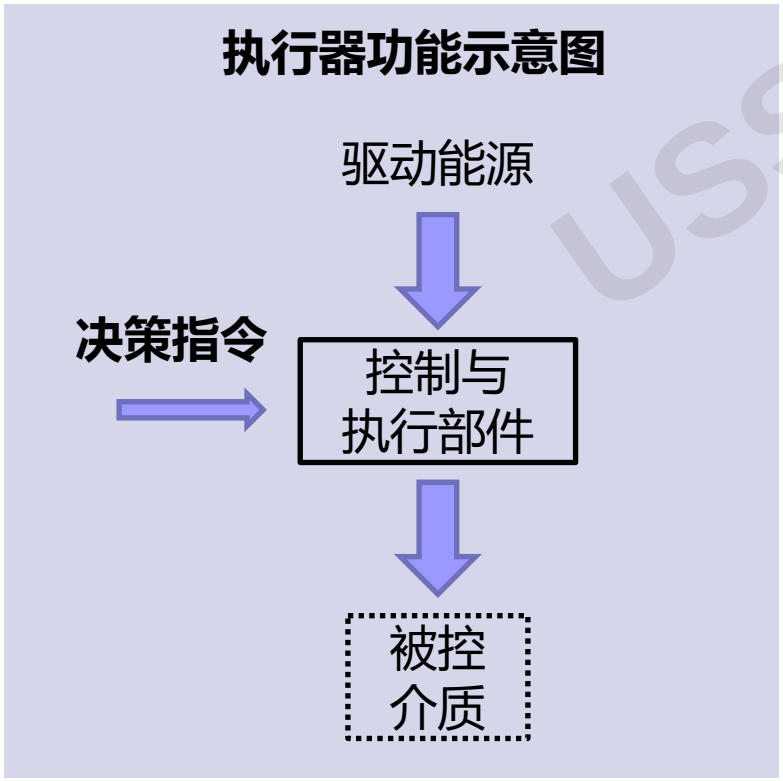
1. 控制/执行器的定义和组成

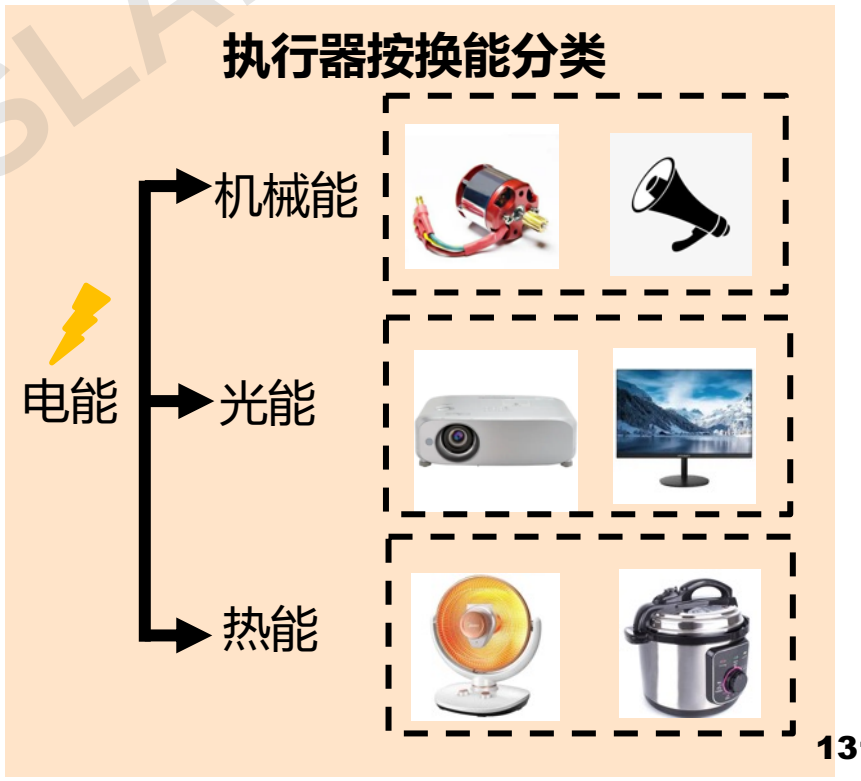
2. 控制/执行器安全问题

USSLAB

5.5.1 定义和组成

- **定义**：可以根据指令，利用某种**驱动能源**（如电能），改变**被控介质**（如电机转子）的机械性状或其他物理特性，从而将被控量（如转速）维持在所要求数值上或一定范围内的装置。

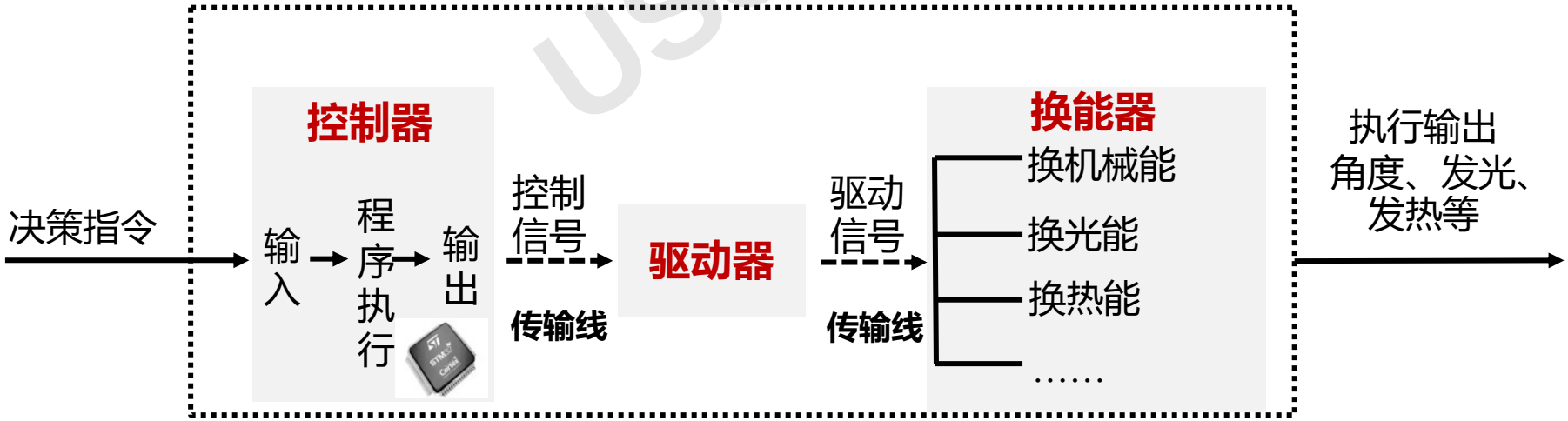




USSLAB

5.5.1 定义和组成

- **组成：包括控制器、驱动器、换能器及其传输线**
  - **控制器：**根据指令执行控制程序，输出控制信号，如单片机、FPGA(可编程逻辑阵列），DSP(数字信号处理器)，PLC(可编程逻辑控制器)等
  - **驱动器：**将**控制信号**转换为**驱动信号**，例如电机驱动电路
  - **换能器：**在驱动信号驱动下完成**换能的装置**，例如电动机、扬声器等

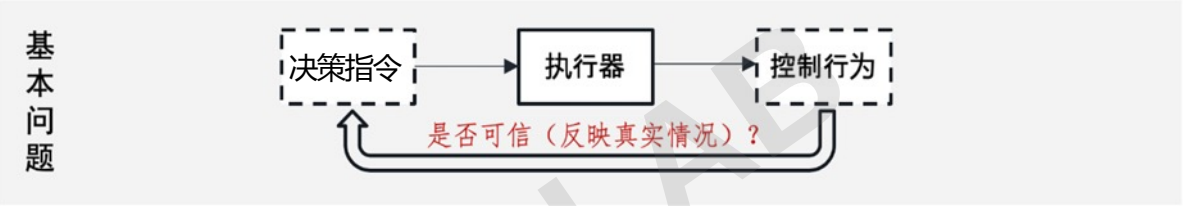


控制/执行器的组成

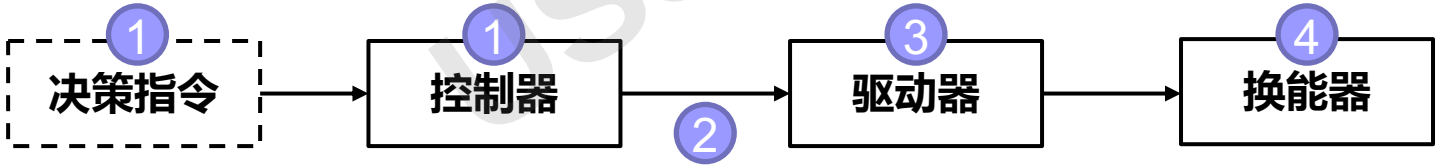
USSLAB

5.5.2 控制与执行安全问题

- **定义**：控制和执行在决策指令的作用下，将被控量维持在要求数值或规定范围内



- **脆弱性及攻击**

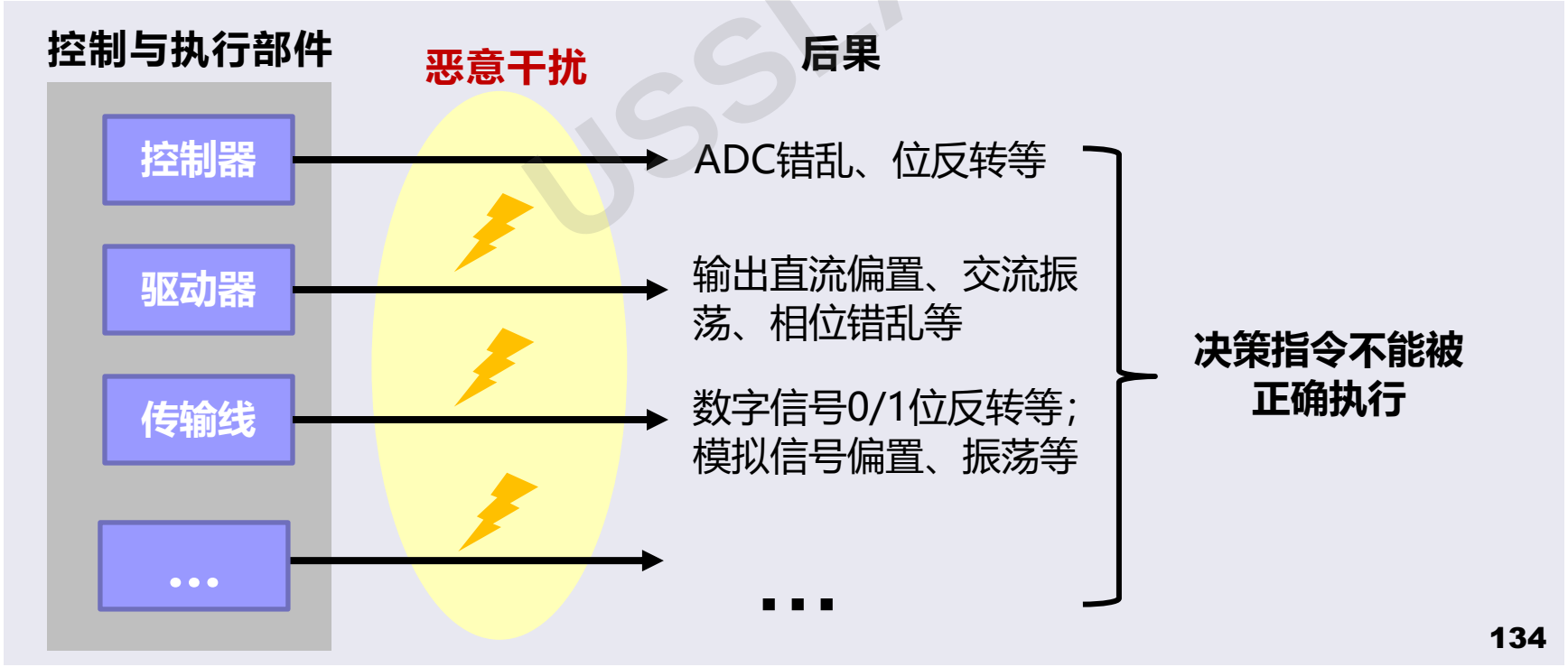


1. **决策、控制指令篡改**：恶意固件篡改指令、错误数据注入等
2. **控制信号篡改**：例如利用EMI改变传输线上的控制信号
3. **针对驱动器硬件的攻击**：利用驱动器放大电路特性进行攻击
4. **执行器换能攻击**：改变本身换能过程、或者利用非设计功能内的换能

USSLAB

5.5.2 控制与执行安全问题

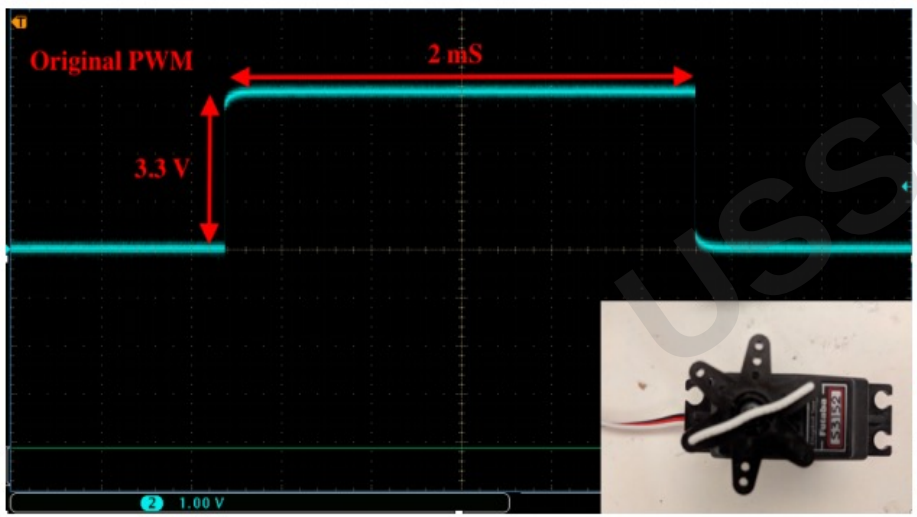
- **执行超限脆弱性**
  - **定义**：由于受外部信号干扰，执行器没有正确执行决策指令或者执行错误指令引发的执行器安全问题。
  - 案例：电机PWM控制信号被EMI干扰



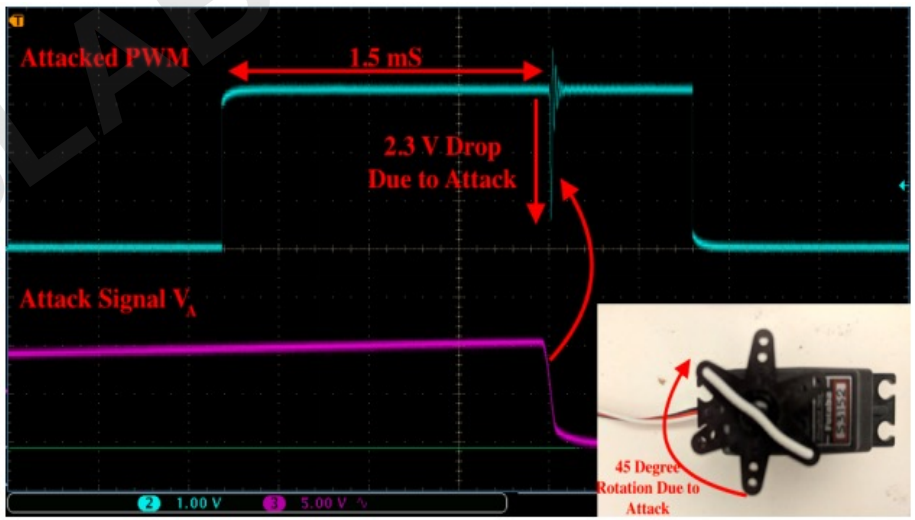
USSLAB

5.5.2 控制与执行安全问题

- **案例： 无人机PWM控制信号电磁注入攻击**
  - **原理**：利用电磁信号耦合到PWM信号传输线，改变PWM信号，实现对执行器的控制，进而恶意控制舵机角度



(a) 原始PWM 信号和对应的伺服电机的位置



(b) 攻击信号VA （紫色）； 被攻击PWM信号（蓝色） 和对应的伺服电机位置

舵机PWM信号攻击实验图

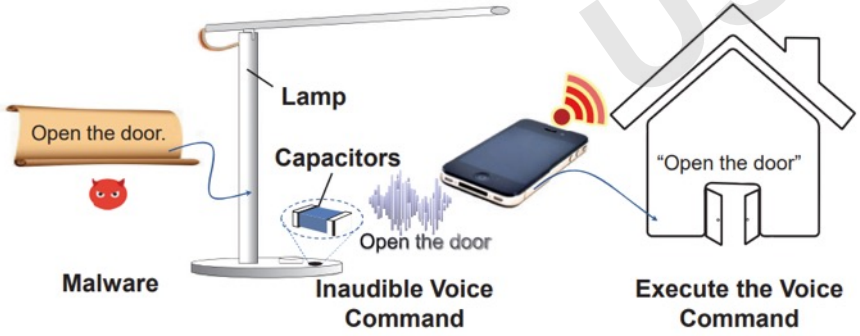
USSLAB

J. Selvaraj, et al., “Electromagnetic induction attacks against embedded systems”, ACM AsiaCCS 2018

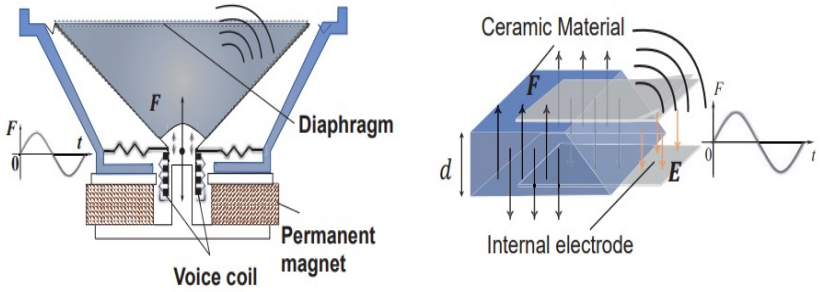
5.5.2 控制与执行安全问题

- **隐蔽执行脆弱性**
  - **定义**：执行器执行了非设计功能之外的功能。例如，电容逆压电效应产生声音；一些显示、通信部件在设计功能之外存在的隐蔽通道。

- 案例1：CapSpeaker [1] 利用电容的**逆压电效应**，通过控制电容两端的电压振动发声。将恶意语音命令调制到高频声波中，利用麦克风的非线性特性解调出恶意指令，欺骗控制智能语音助手



CapSpeaker示意图



电容 vs. 扬声器

[1] Xiaoyu Ji, Juchuan Zhang, Shui Jiang, Jishen Li, Wenyuan Xu, CapSpeaker: Injecting Voices to Microphones via Capacitors, ACM CCS 2021.

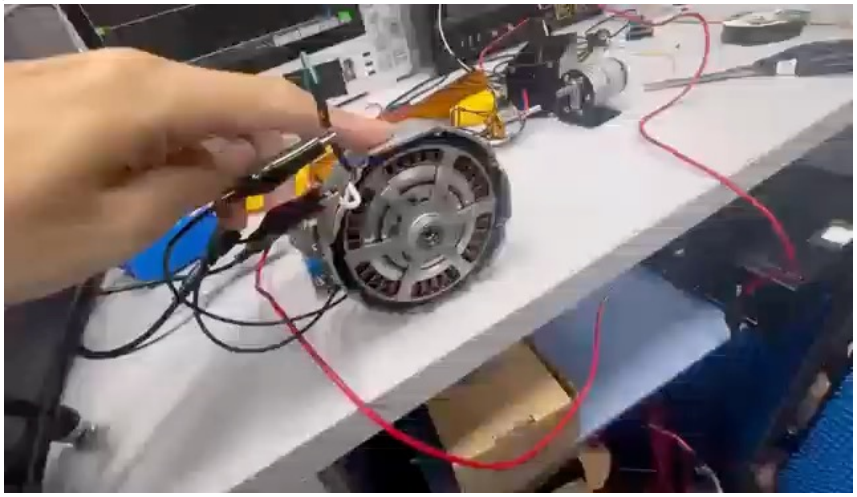
5.5.2 控制与执行安全问题

- **隐蔽执行脆弱性**

- 案例2：电机发声。电动机具备较好的“电能-机械能”换能特性，可以高效地发特定宽频带的声音

USSLAB

四旋翼无人机电动机播放
“Ok, Google”



宇树机器人电动机播放
“What’s the temperature outside”

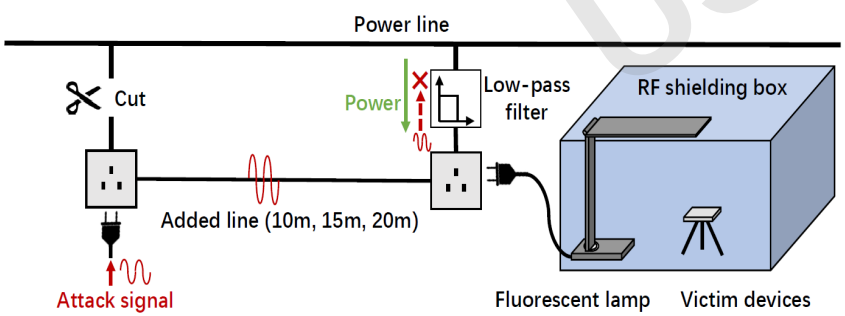
5.5.2 控制与执行安全问题

- **隐蔽执行脆弱性**

- 案例3：**荧光灯发射磁信号。**从公共电源线注入特殊调制的高频信号，可以激励荧光灯产生可利用的EMI信号，操控附近容易被EMI干扰的电子设备

USSLAB

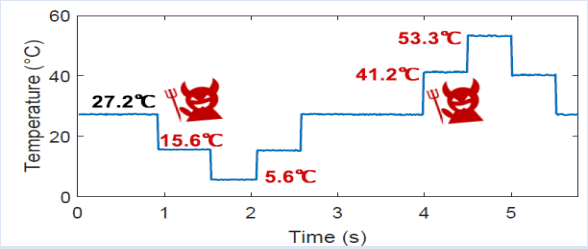
从电源端注入信号，利用荧光灯产生EMI操控目标传感器



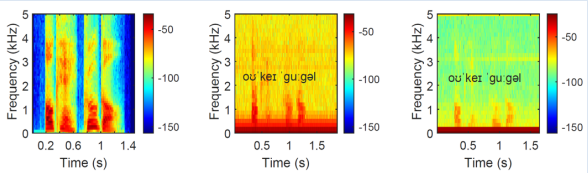
从公共电源端注入信号能实现隐蔽、远距离的攻击，突破空间的限制

基于荧光灯电磁辐射的传感器攻击

操控温度传感器



操控麦克风

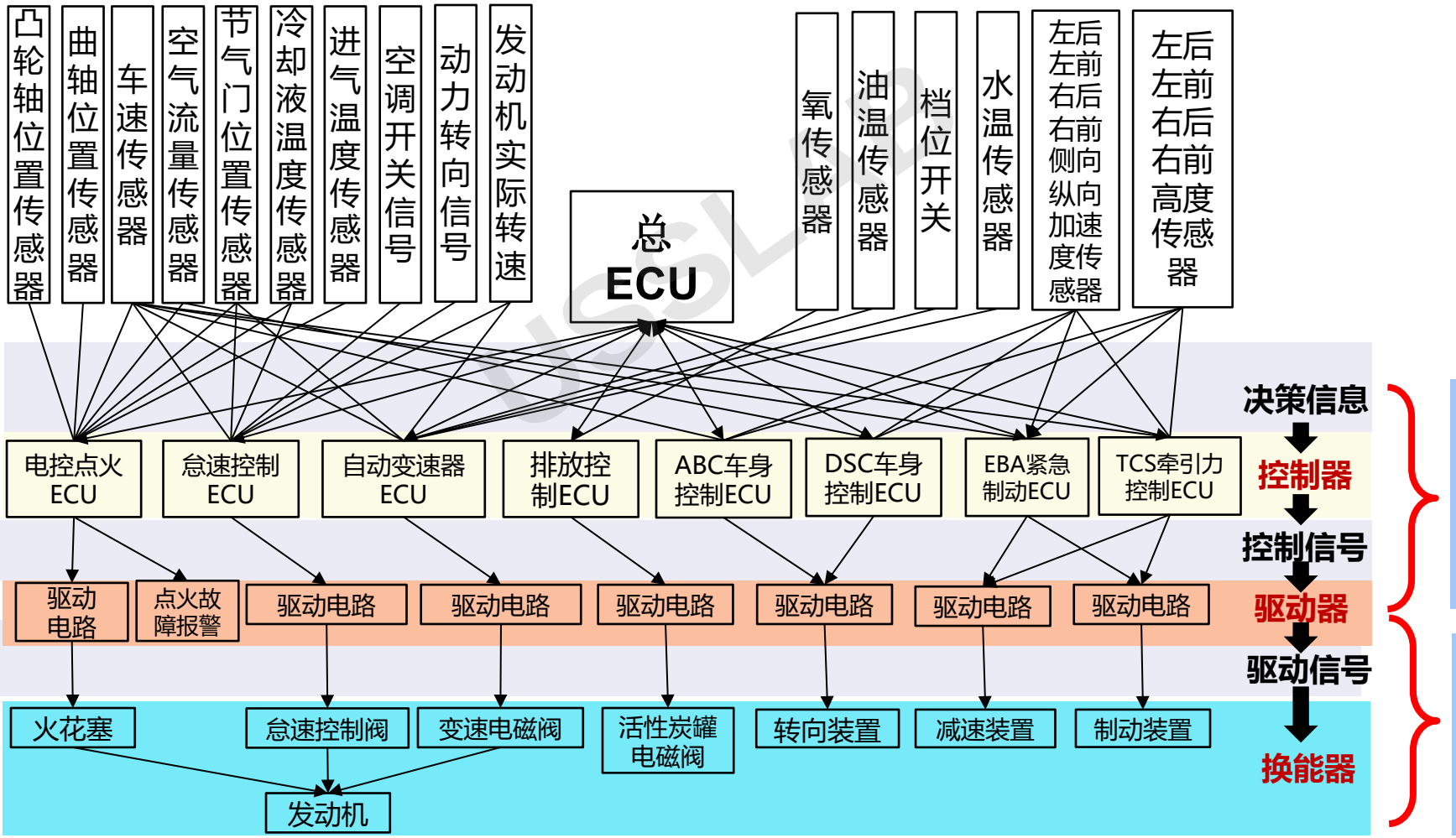


原始音频　天线注入　荧光灯注入

Fengchen Yang, Wenze Cui, Xinfeng Li, Chen Yan, **Xiaoyu Ji**, Wenyuan Xu. "LightAntenna: Characterizing the Limits of Fluorescent Lamp-Induced Electromagnetic Interference", accepted by the Network and Distributed System Security Symposium (**NDSS**) 2025.

5.5.2 控制与执行安全问题

- 汽车电控系统中的控制、执行安全问题



传感器与执行器安全

- 第六节 传感器安全发展趋势

1. 传感器发展趋势

2. 未来传感器案例

USSLAB

5.6.1 传感器发展趋势

- 传感器发展带来数据洪流
  - 传感器**数据生成速度** >> 全人类**数据处理速度**

Figure 2: The brain's ability to perceive and reason is based on ultra-compressed sensing capabilities with 100,000 data reduction and a low operation power

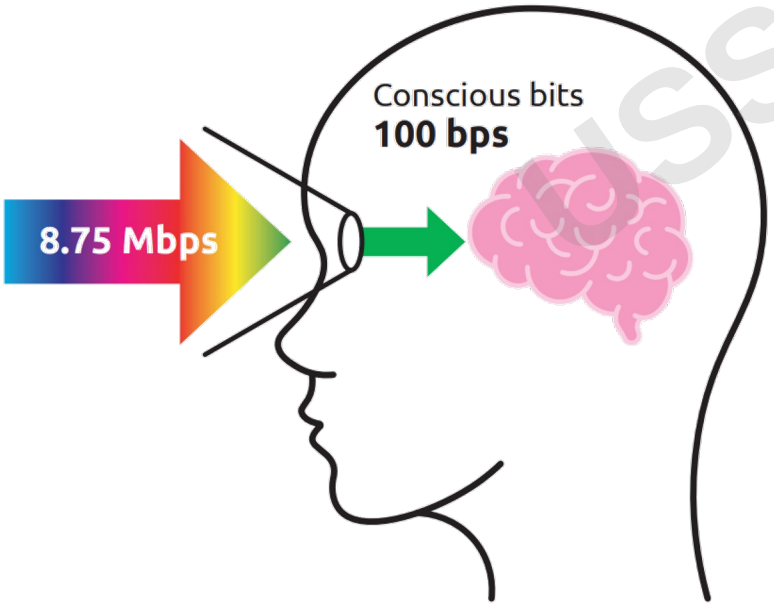
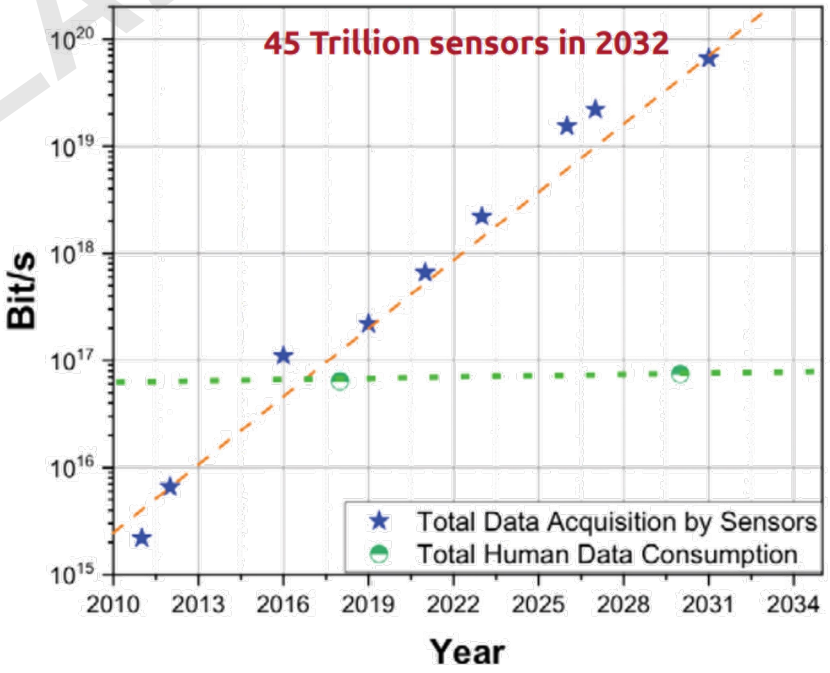


Figure 3: Trend in world's installed sensing capacities



USSLAB

Decadal plan for semiconductors [2021]

5.6.1 传感器发展趋势

- 从“感算分离”到“感算一体”
  - **传感器职责：**感知 → 感知 + 计算
  - **目的：**加快处理速度；节省功耗；降低时延

图1：传统传感器计算架构

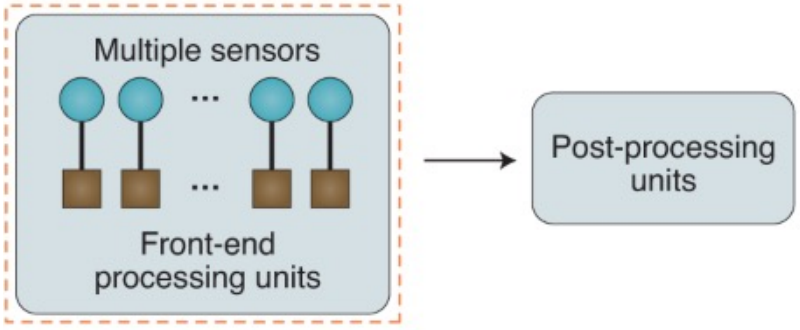


图2：未来近感计算架构

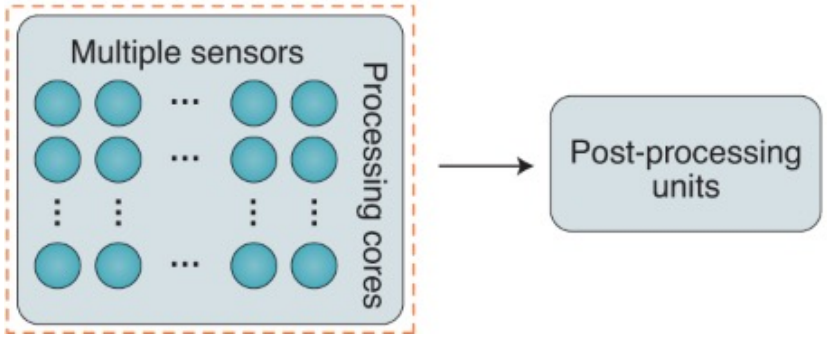
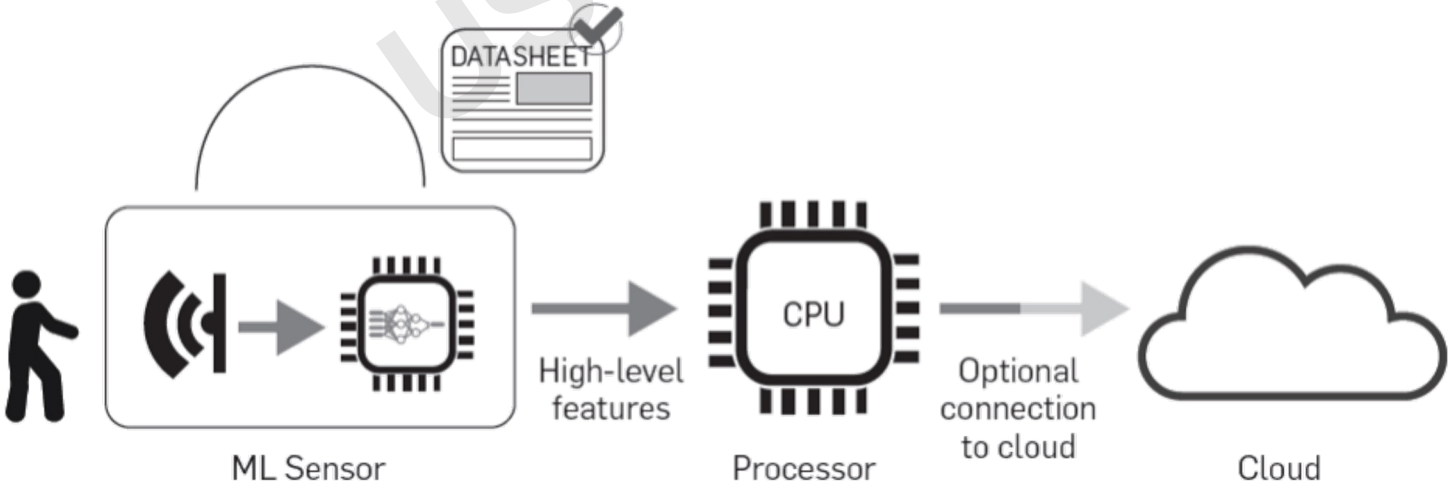


图3：未来感内计算架构

Near-sensor and in-sensor computing [2020, Nature]

5.6.2 未来传感器案例

- **机器学习传感器[1]**
  - 在硬件层面上，将传感器输入数据和机器学习处理进行封装，并提供一个简洁接口
  - 例如，人体检测传感器仅提供1位输出信号(0代表无人，1代表有人)

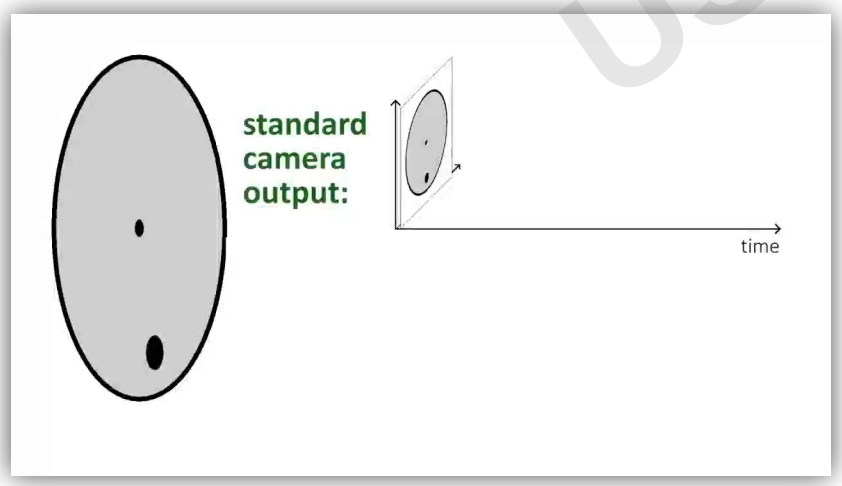


USSLAB

[1] Machine learning sensors. [2023, Communications of the ACM]

5.6.2 未来传感器案例

- **事件相机[1]**
  - 基于生物视觉启发的传感器。其工作原理与传统帧式相机有本质区别，每个像素独立配备光强检测电路，当检测到亮度变化超过设定阈值时（如ΔL ≈ 10-20%），触发事件并记录（x坐标，y坐标，时间戳，亮度变化极性）
  - 应用：自动驾驶（动态目标检测）、工业检测（高速生产线）和无人机避障等

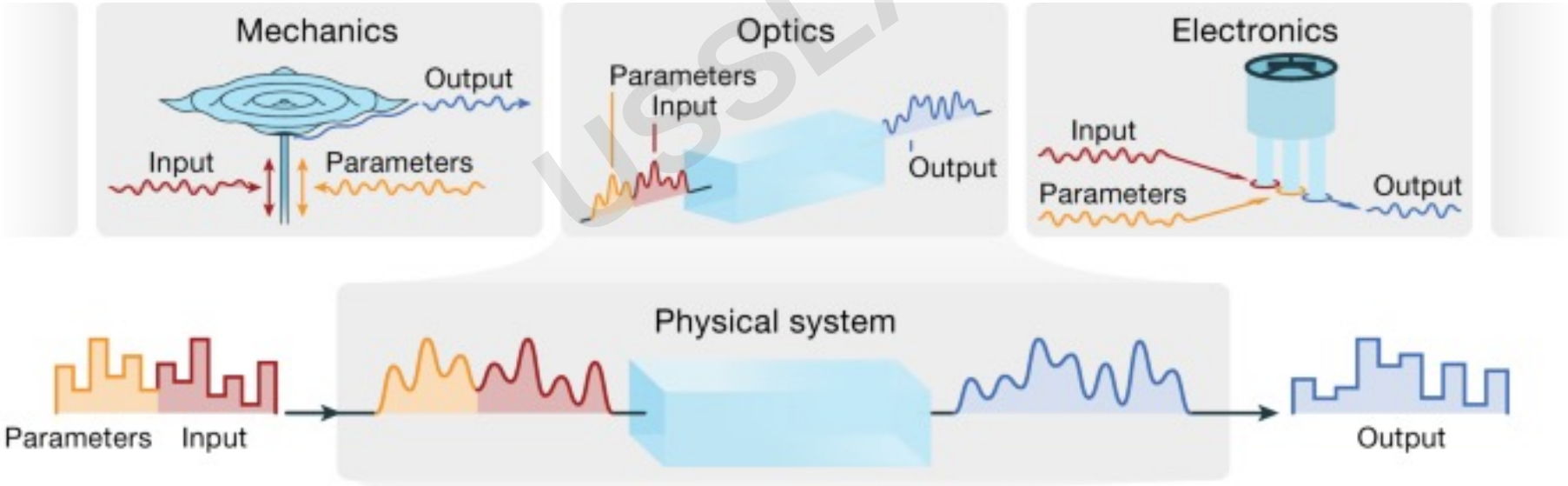


USSLAB

[1] Event-based, 6-DOF Pose Tracking for High-Speed Maneuvers. [2014, IROS]

5.6.2 未来传感器案例

- **物理神经网络[1]**
  - 利用(机械/光学/电子)器件的物理性质
  - 物理非线性替代神经网络中的激活函数



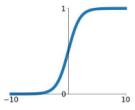
[1] Deep physical neural networks trained with backpropagation. [2022, Nature]

5.6.2 未来传感器案例

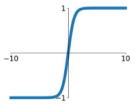
- **物理神经网络**
  - 数字激活函数：例如Sigmoid、ReLU等
  - 物理器件非线性：类似激活函数，提供非线性变换

USSLAB

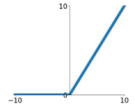
Sigmoid
$\sigma(x) = \frac{1}{1+e^{-x}}$



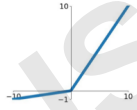
tanh
$\tanh(x)$



ReLU
$\max(0, x)$

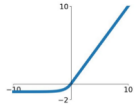


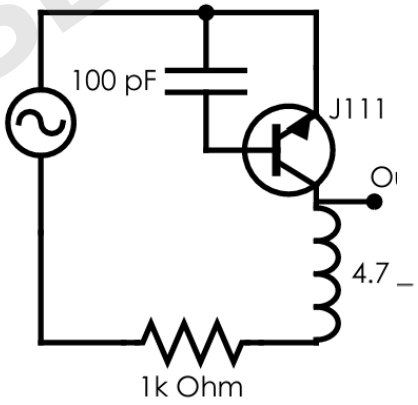
Leaky ReLU
$\max(0.1x, x)$



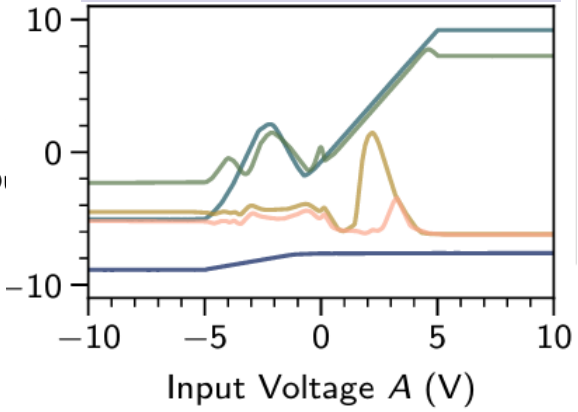
Maxout
$\max(w_1^T x + b_1, w_2^T x + b_2)$

ELU
$\begin{cases} x & x \geq 0 \\ \alpha(e^x - 1) & x < 0 \end{cases}$





输出电压响应非线性曲线



深度神经网络中常用的激活函数　　v.s.　　模拟电路提供的物理激活函数

Deep physical neural networks trained with backpropagation. [2022, Nature]

In-sensor computing and CIA

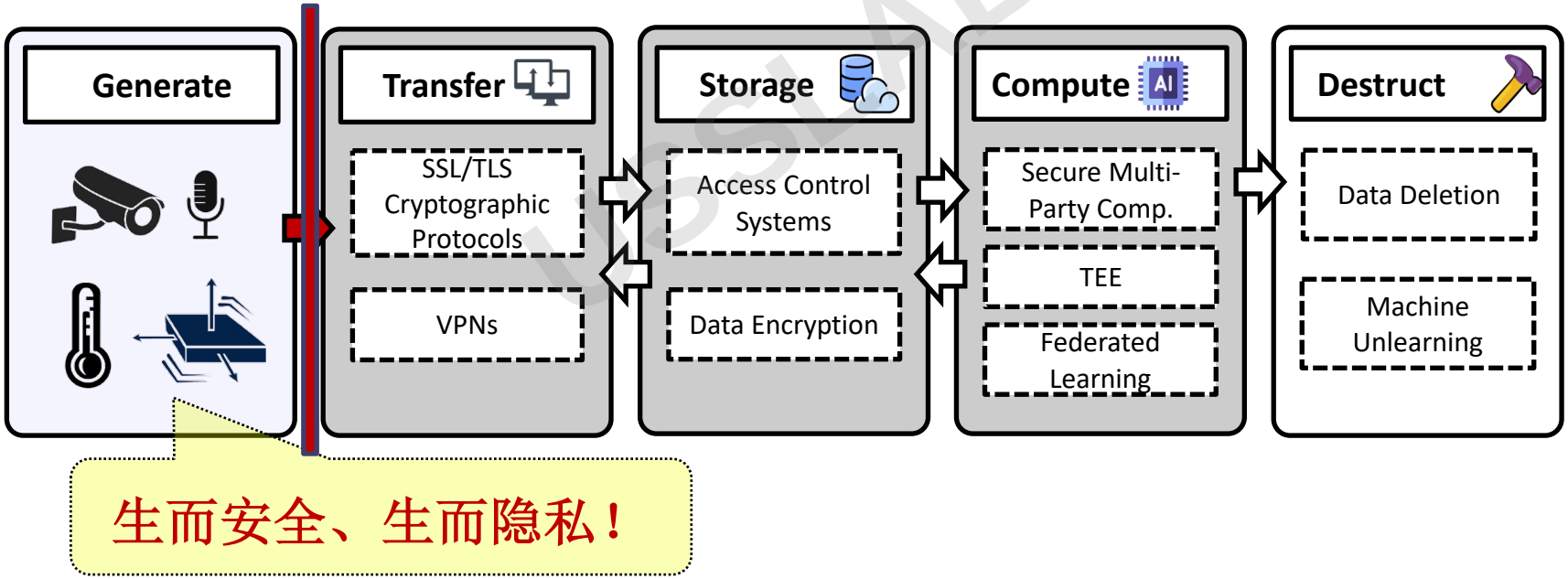
- MicPro
- CamPro
- Sensor-oriented deepfake detection

USSLAB

Wenjun Zhu, Yuan Sun, Jiani Liu, Yushi Cheng, Xiaoyu Ji, Wenyuan Xu, "CamPro: Camera-based Anti-Facial Recognition", ***NDSS*** 2024
Shilin Xiao, Xiaoyu Ji, Zhicong Zheng, Chen Yan, Wenyuan Xu, "MicPro: Microphone-based Voice Privacy Protection", ***CCS*** 2023.

传感器安全：自源保护（By birth）

- 传感器应该承担攻击防护第一道防线，保护测量可信、数据隐私等（CIA属性）

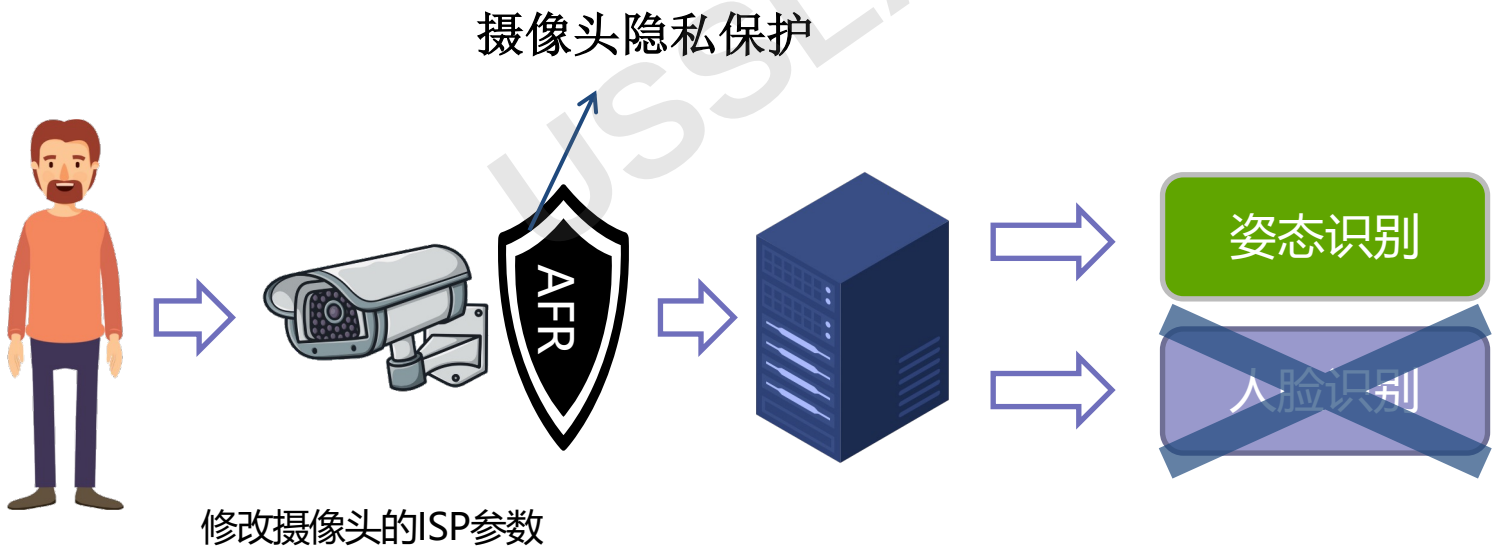


Xiaoyu Ji, et. al, Sensor-based IoT data privacy protection, Nature Review Electrical Engineering

CamPro: 摄像头自动人脸打码

- 防护效果：在相机传感器中实现人脸匿名化。
- 优点：

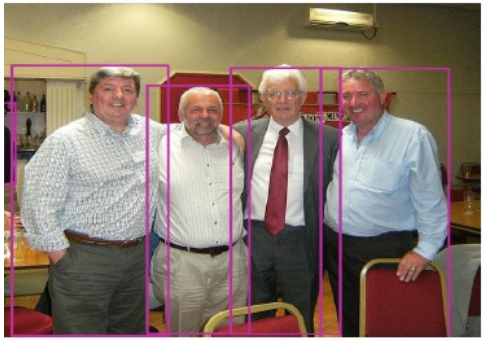
  （1）不修改相机的硬件组成；（2）保持拍照功能可用性



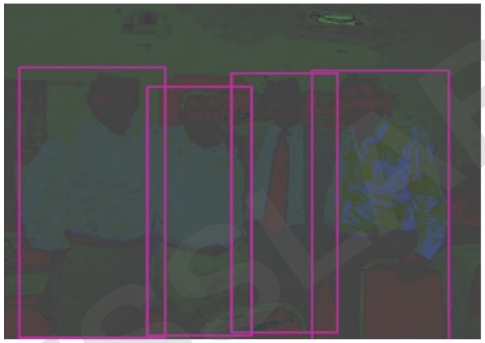
USSLAB

Wenjun Zhu, Yuan Sun, Jiani Liu, Yushi Cheng, Xiaoyu Ji, Wenyuan Xu, "CamPro: Camera-based Anti-Facial Recognition", ***NDSS*** 2024

CamPro: 摄像头“生而隐私”



原始图像



捕获的图像

增强后的图像

摄像头拍摄的照片
不包含人脸信息

本章总结

- 了解传感器的基本定义、组成和静态特性
- 掌握传感器测量安全
  - 传感器脆弱性
  - 换能攻击
  - 典型攻击案例：Dolphinattack、Walnut、GhosTalk等
  - 传感器安全模型
- 掌握执行器的安全攻击定义，了解常见的攻击方法

IISSLAB

Reference

- **海豚音攻击：** Dolphinattack: Inaudible voice commands, ACM CCS’17
- **Walnut攻击：** WALNUT: Waging doubt on the integrity of mems accelerometers with acoustic injection attacks, IEEE Euro S&P’17
- **GhostTalk攻击：** Ghost talk: Mitigating EMI signal injection attacks against analog sensors, IEEE S&P’13
- **DeHiREC**: Detecting Hidden Voice Recorders via ADC Electromagnetic Radiation, IEEE S&P’23
- **智能传感器：** Machine learning sensors, Communications of ACM, 2023
- **物理神经网络：** Deep physical neural networks trained with backpropagation, Nature, 2022
- Physical Foundation Models: Fixed hardware implementations of large-scale neural networks, https://arxiv.org/abs/2604.27911

USSLAB