Heterogeneous System Security

Nael Abu-Ghazaleh

UNIVERSITY OF CALIFORNIA
UCRIVERSIDE

What's the role of hardware in security?

Nvidia GPU owners told to update now to patch a range of serious security flaws

News By Sead Fadilpašić published October 29, 2024

Eight vulnerabilities addressed by Nvidia in total

When you purchase through links on our site, we may earn an affiliate commission. Here's how it works.

Patch Tuesday: Intel Publishes 44 and AMD New Advisories

November 2024 Patch Tuesday security advisories to inform customers about vulnerabilities found recently in their

TRENDING

1 Microsoft Confirms Zero-Day Exploitation of Task Scheduler Flaw

2 Ahold Delhaize Cybersecurity Incident Impacts Giant Food, Hannaford

Architecture & Security

Attacks

Vulnerabilities originating in architecture

- Side and covert channels
- Speculation attacks
- Fault injection
- Hardware trojans, …

Defenses

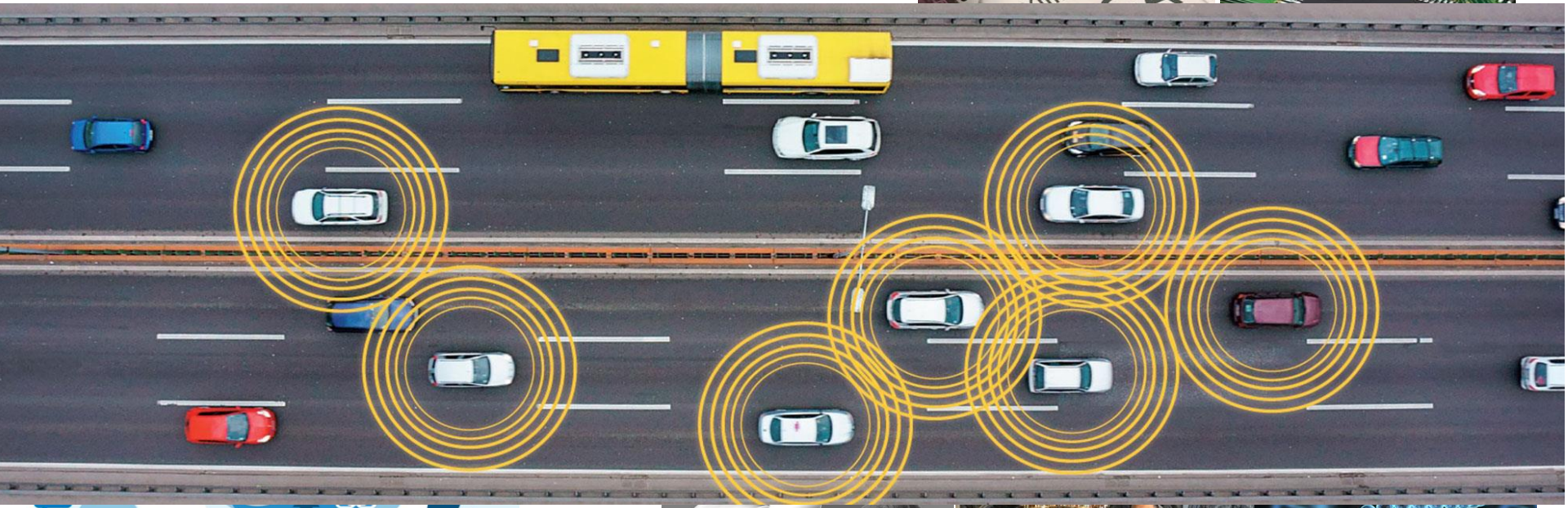
Defenses rooted in architecture

- Do no harm
  - Avoid vulnerabilities in architecture/HW
- Help software
  - Security abstractions/mechanisms
  - Computational support for expensive defenses

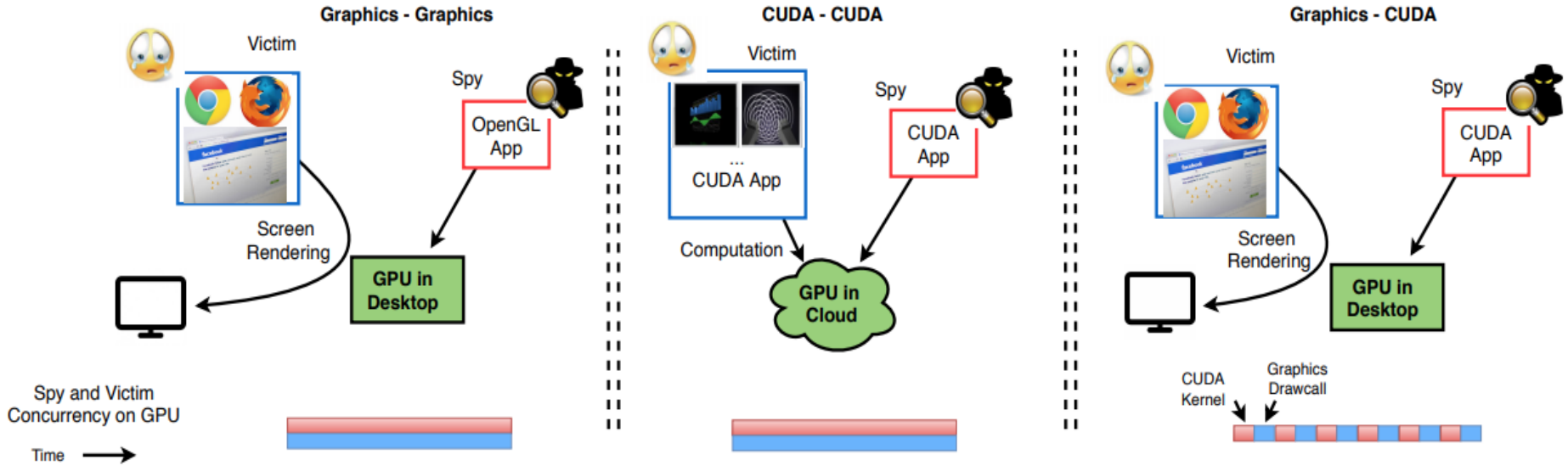
Some Trends (Arch2030)

Acceptance of Moore's law ending

Specialization happening at scale



Started exploring attacks on GPUs



Heterogeneous Systems

Heterogeneity: modern computing platforms are becoming increasingly heterogeneous.

Small-scale **Large-scale**

Enhancing Power Efficiency & Flexibility

Optimizing performance & Cost Efficiency

Vulnerability in small-scale heterogeneous systems

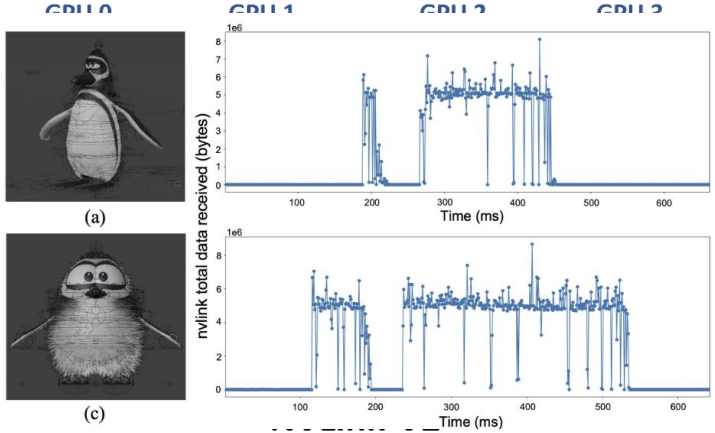
- Side channel attack on AR/VR [5,6]
- Contention attacks
- Shared-state attack in Multi-User Augmented Reality Applications [7]

Vulnerability in large-scale heterogeneous systems

- Vulnerabilities in GPUs, NPUs, Memory
- Attacks on multi-GPUs
- Attacks on memory and interconnect

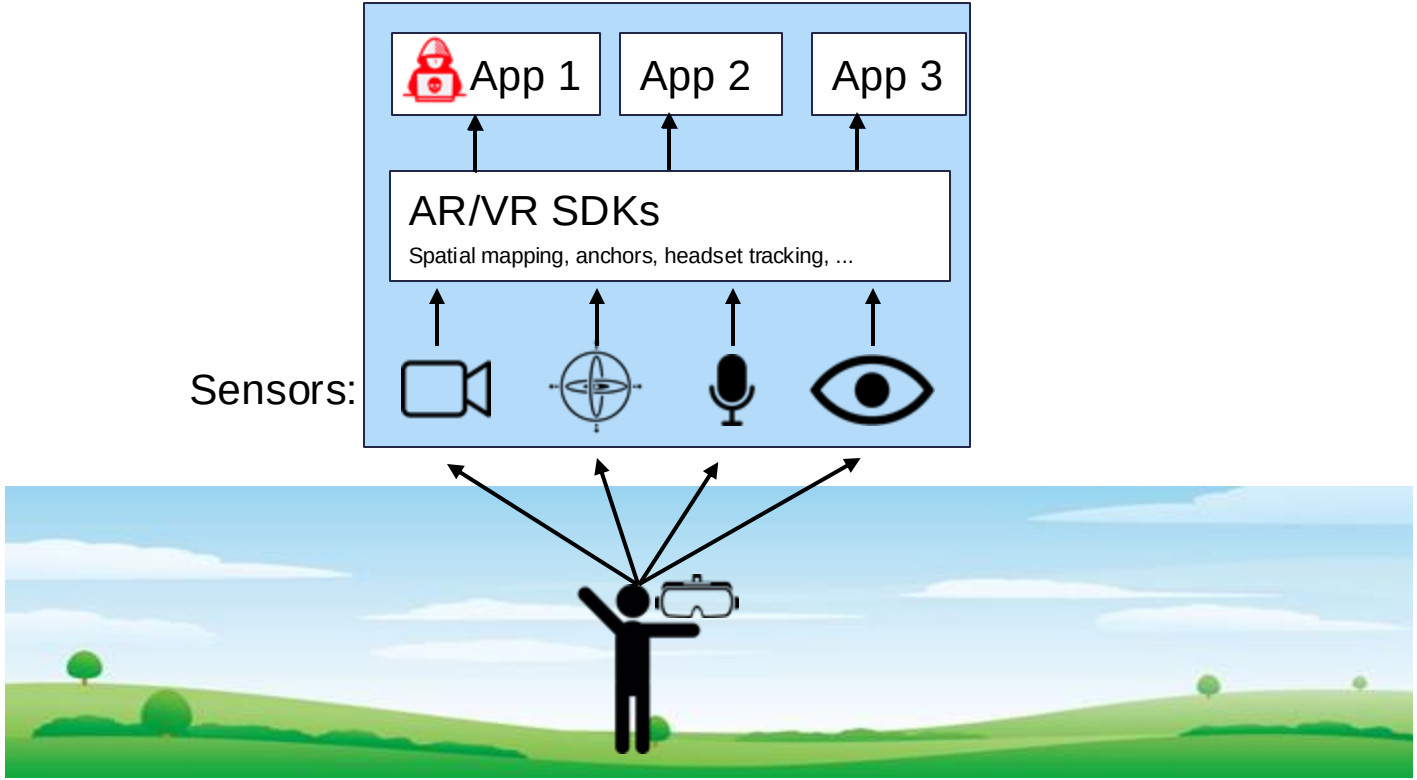
Security of heterogeneous systems

- Every system is different
  - How to think about security?
- Started exploring attacks
  1. Large scale heterogeneity
     - Data centers with accelerators
     - Multi-GPU systems
  2. Small scale heterogeneity
     - SoCs, Mobile systems, Embedded/IoT
     - XR security



Where can the attacker come from?

Side channel attacks across apps



Threat model 1:
Attacks across apps within the same headset

Its all in your headset: Side channel attacks on AR/VR systems, USENIX Security'23

Going through the motions: AR/VR Keylogging from user head motion, Usenix Security'23

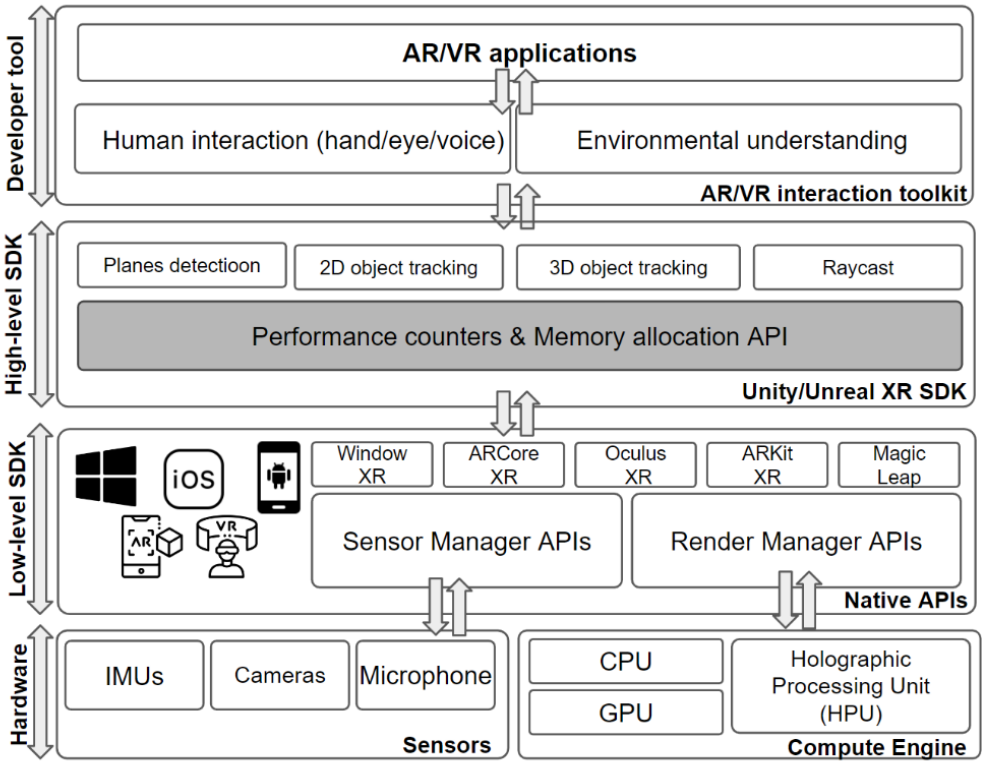
Scenario: Multiple applications running simultaneously

- Many AR/VR devices support multiple apps simultaneously augmenting environment
  - Ex: Microsoft Hololens 2, Meta Quest 2, and Magic Leap One

- Permissions management on AR/VR systems allows access to common lower-level SDKs
  - Even though apps can be visually isolated

Overview of a typical XR system

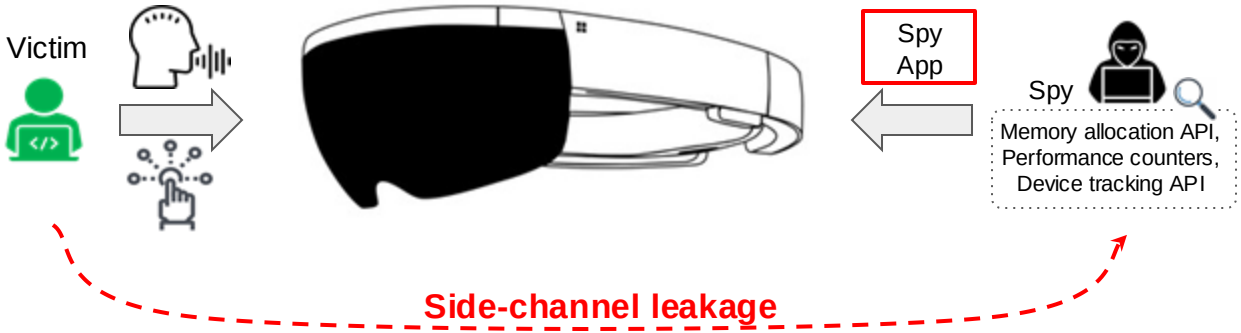
Performance counters accessible to user applications



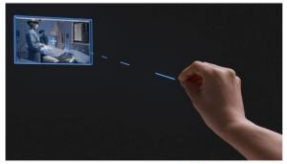
Threat model: Software side-channel

- A malicious application runs in the background with normal permissions
- Probe Unity/Unreal APIs to leak side-channel information
  - Memory allocation API: Memory usage/limit/peak, etc.
  - Rendering performance counters: CPU/GPU frame rate, etc.
  - Device tracking API: headset acceleration, gyroscope, etc.

Normally provided to developers to optimize app performance



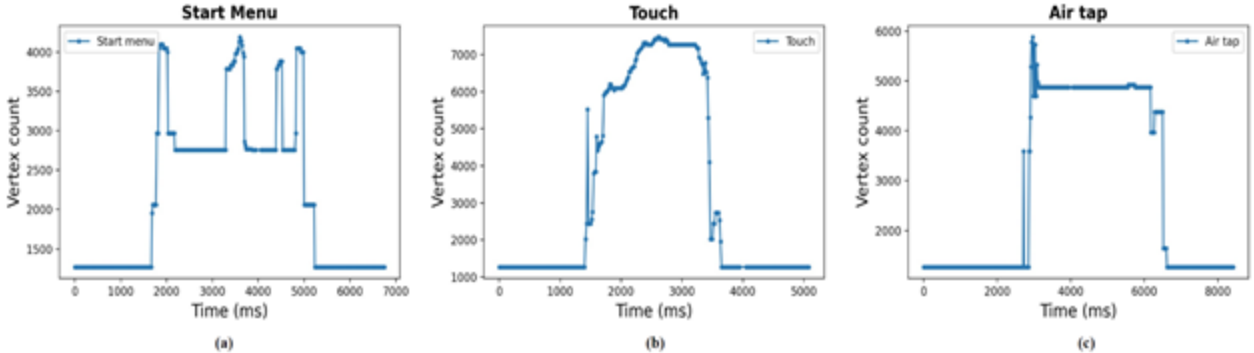
User Interaction Attack—Hand Gestures



(a) Air tap [45]

(b) Start gesture [45]

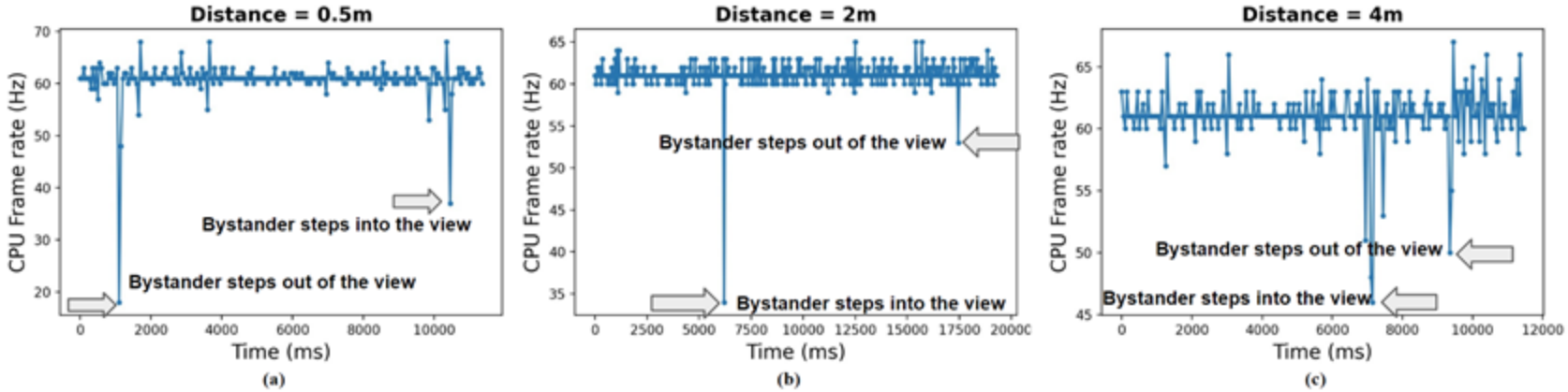
- Recall: malicious program in background samples pertormance counters
  - Performance counters have unique time series signatures



- Example: Different patterns from user hand gestures

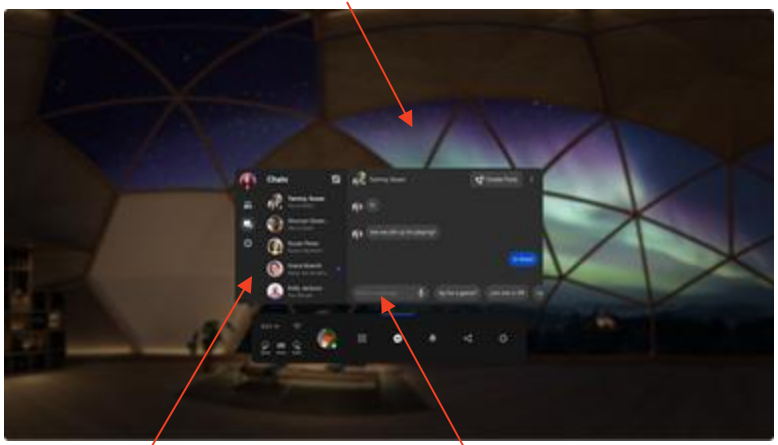
Environment Inference attack – Bystander ranging

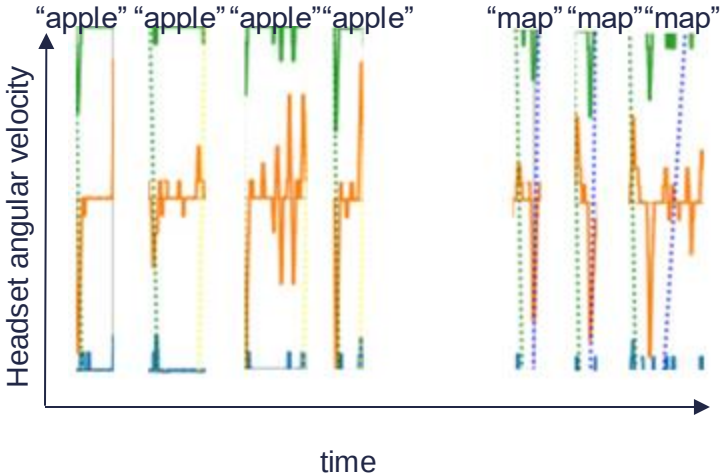
- Yes! Performance counters have unique time series signatures



Can tracking APIs/sensors reveal typing in another app?

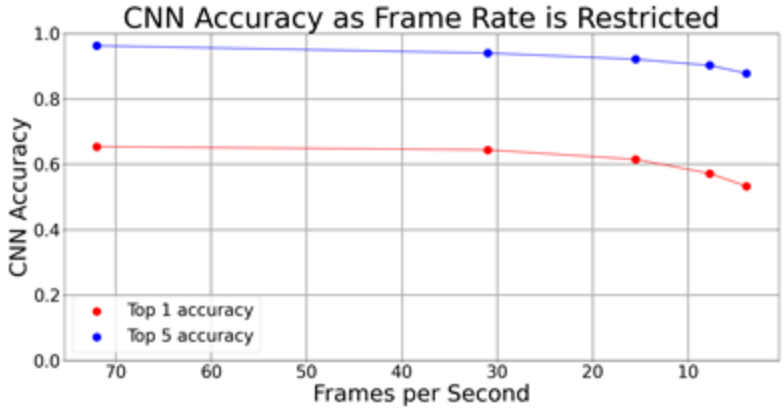
- Yes! Head pose has unique time series signatures for different characters





Potential Mitigations ineffective

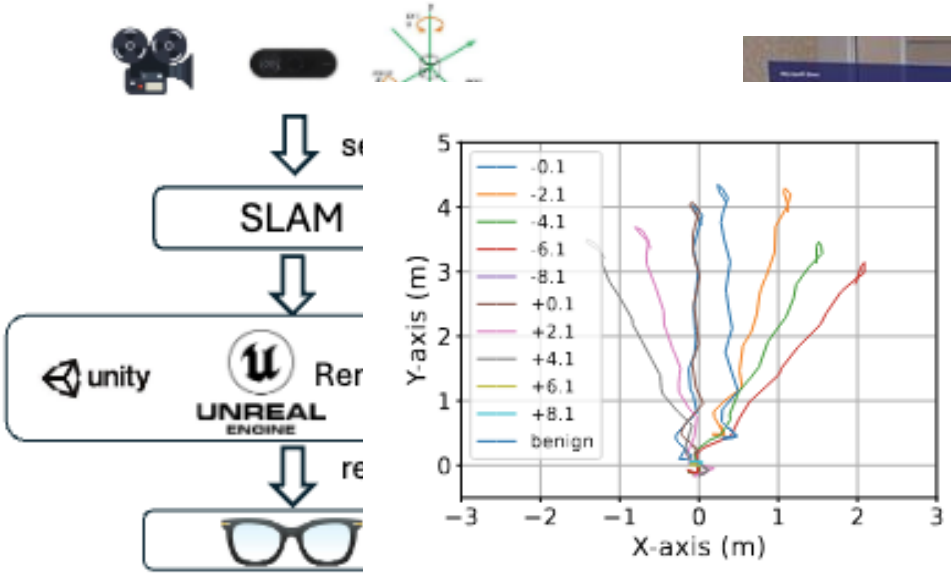
- Reduce available sampling rate of performance counters



→ Sampling rate reduction ineffective!

- Block access to performance counters completely → user suffers
- User permissions → Informed consent difficult
- …??

Ongoing Work -- Decorrelating Virtual and Physical Worlds



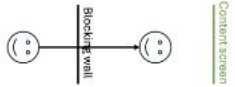
Attack the sensors

Selective Contention

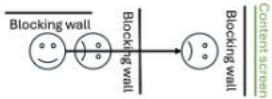
(a) no sound, frame 1

(b) no sound, frame 2

(c) no sound, frame 3



(d) user trace without sound



(a) no sound, frame 1

(b) no sound, frame 2

(c) no sound, frame 3

(d) no sound, frame 4

(e) no sound, frame 5

(f) with sound, frame 1

(g) with sound, frame 2

(h) with sound, frame 3

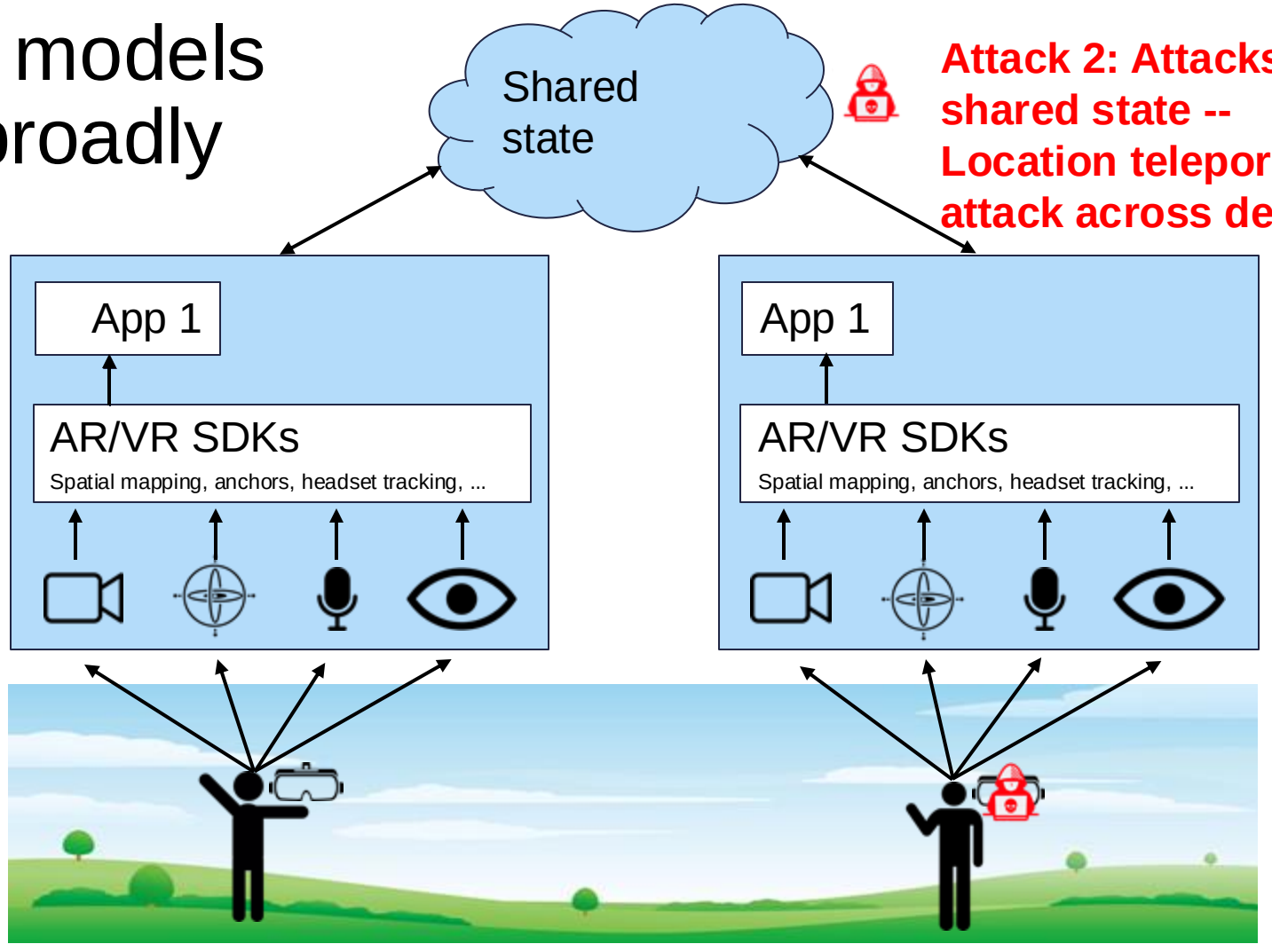
(i) with sound, frame 4

(j) with sound, frame 5

Threat model 2: Shared virtual state

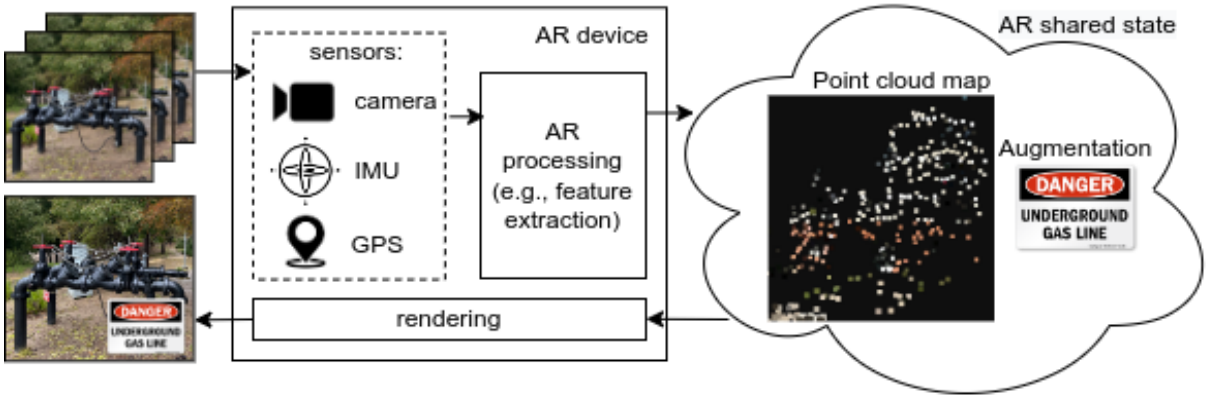
That doesn’t go there:Attacks on Shared State in Multi-User Augmented Reality Applications, Usenix Security’24

Threat models more broadly

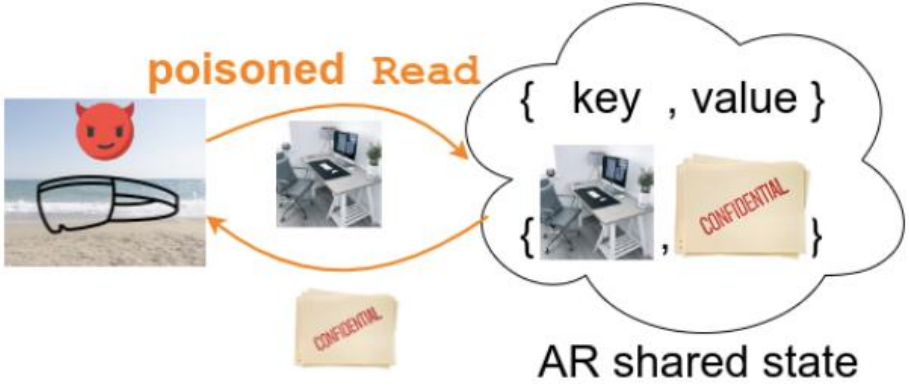


Attack 2: Attacks on shared state -- Location teleportation attack across devices

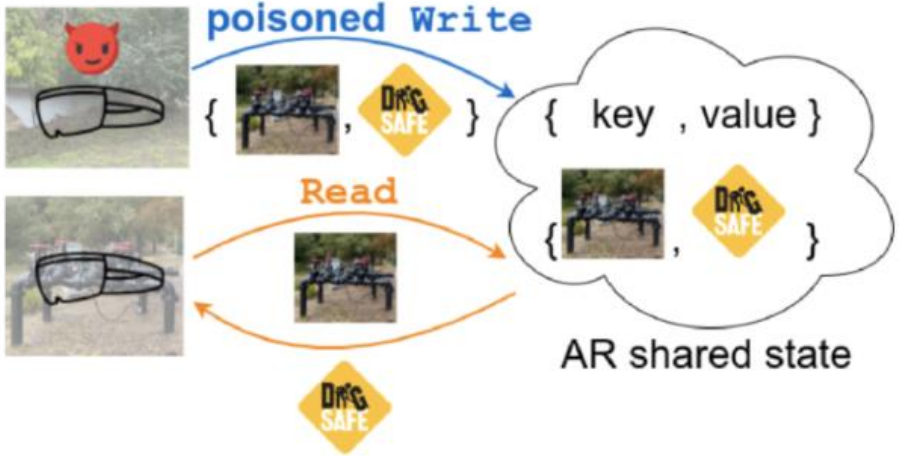
Multi-user AR experiences



- Most AR experiences leverage shared state
  - Sometimes for optimization such as shared SLAM
- Location as an example target

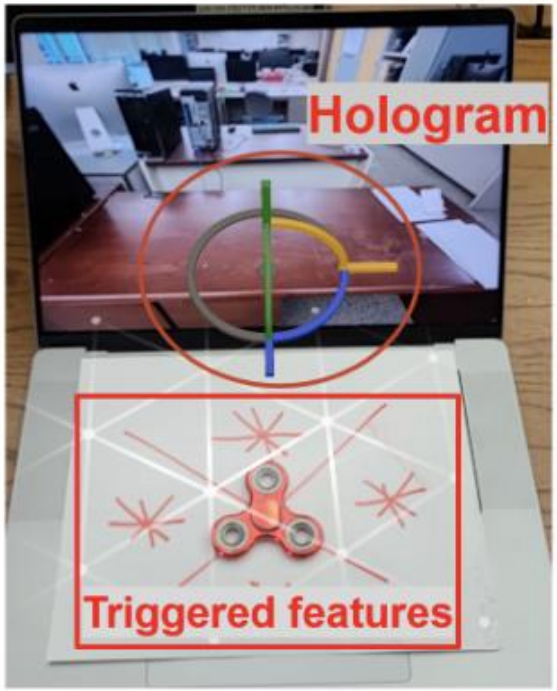


(a) Read attack.



(b) Write attack.

Write attack

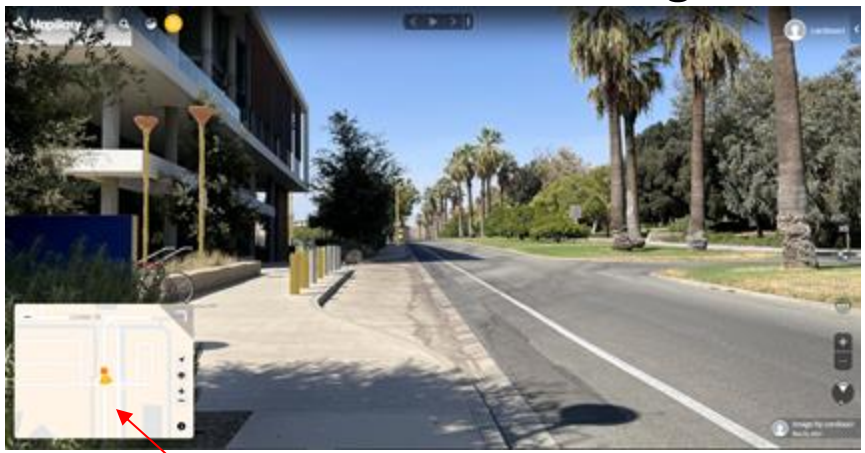


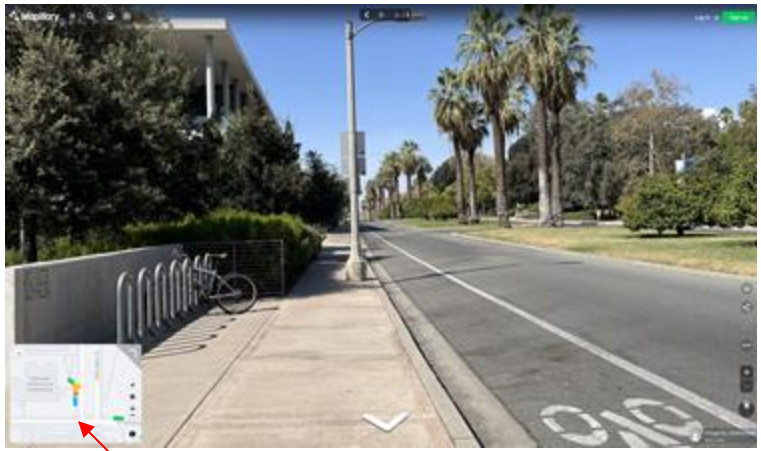
Write hologram at remote location with **triggered** features

Read hologram at physical location with **triggered** features

Write attack on Mapillary

1. Assume that shared state is crowd-sourced 3D maps (e.g., Mapillary)
2. Attacker writes false images to the shared state[1]





Same image present at two different map locations!

3. TO DO: Victim AR device reads bad information from the shared state

→ teleport and wrong holograms displayed

[1] We received permission from Mapillary to upload false images in a geofenced area

Location A

Location B

No Attack

With Attack



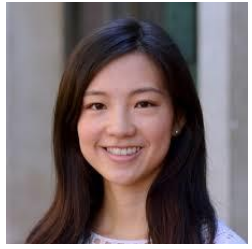













THANKS – QUESTIONS?