

Siren Song: Acoustic Attacks on Pose Estimation in XR Headsets

Zijian Huang¹, Yicheng Zhang^{2,3}, Sophie Chen¹,
Nael Abu-Ghazaleh², Jiasi Chen¹

yzhang95@gmu.edu

¹*University of Michigan, Ann Arbor*

²*University of California, Riverside*

³*George Mason University*

Motivation

- XR has applications across many areas:

But...what if pose estimation is under attack?

- Importance of pose estimation:



Rendering



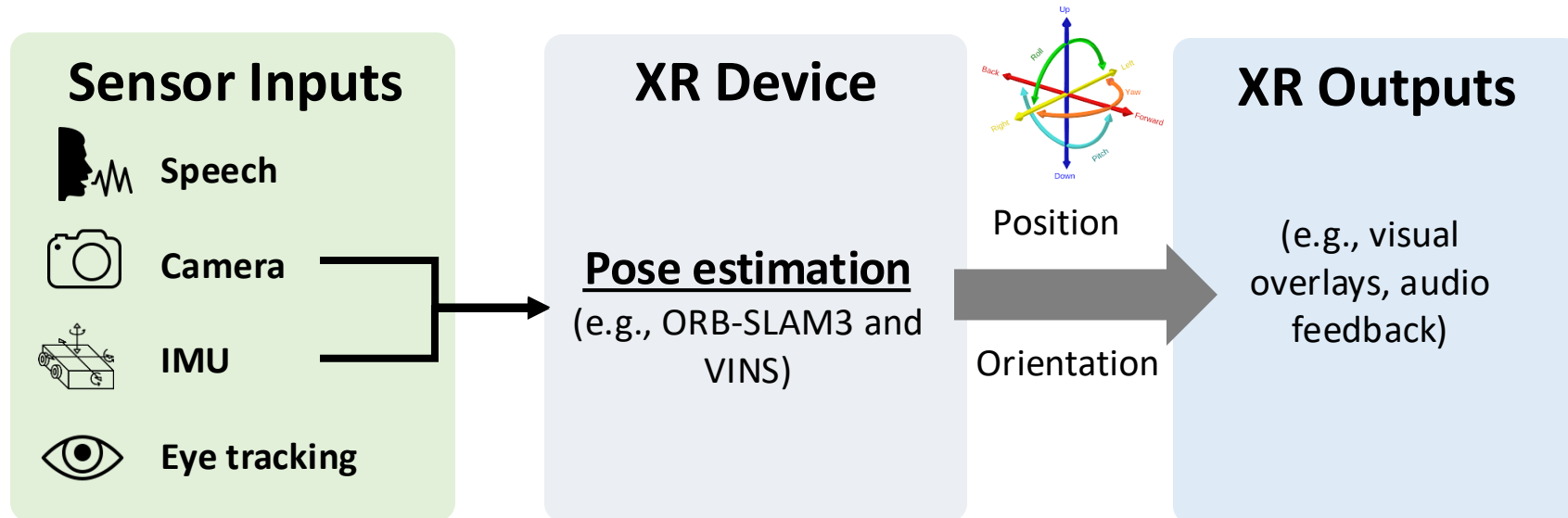
User Interaction



VR sickness

Background: Pose estimation in XR

- XR devices fuse IMU and camera data to estimate the headset's location and orientation.



Background: Acoustic Attack

- Plenty of acoustic attacks on drones, vehicles, and phones.
- **Key Challenges:**
 - Noninvasive attack setup.
 - XR-specific pipeline.
 - Impact on user experience.

<https://www.usenix.org/conference>

This paper is in
24th USENIX
August

ready known to cause malfunctions by disrupting gyroscopes. However, an open question remains beyond denial of service attacks to achieve control of sensor outputs. Our work investigates acoustic injection attacks can damage the popular type of sensor: the capacitive MEMS. Spoofing such sensors with intentional digital values to microprocessors and enables an out-of-spec pathway for attack. Our contributions include (1) modeling the acoustic interference on MEMS accelerometers, the circuit-level security flaws that cause measuring acoustic injection attacks on MEMS as well as systems that employ on these sensors, and (3) two software-only defenses that mitigate many of the risks to the integrity of MEMS accelerometer outputs.



vulnerabilities, but fewer methodologies exist in the hardware domain.

It is already known that acoustic interference cause denial of service (DoS) attacks against

images, image sensors with internal sensors alleviate image blurring caused by camera such a trend opens a new attack surface. This a system-level vulnerability resulting from the emerging image stabilizer hardware susceptible manipulation and the object detection algorithms adversarial examples. By emitting deliberate acoustic signals, an adversary can control the output sensor, which triggers unnecessary motion compensation results in a blurred image, even if the camera blurred images can then induce object misclassification safety-critical decision making. We model of such acoustic manipulation and design an attack that can accomplish three types of attacks, including, and altering objects. Evaluation results demonstrate the effectiveness of our attacks against four academic object detectors (YOLO V3/V4/V5 and Fast R-CNN), and one commercial detector (Apollo). We further introduce the concept of AMPLe attacks, a new class of system-level security vulnerabilities resulting from a combination of adversarial machine learning and physics-based injection of information-carrying signals into hardware.

This paper is included in the Proceedings of the
27th USENIX Security Symposium.

August 15–17, 2018 • Baltimore, MD, USA

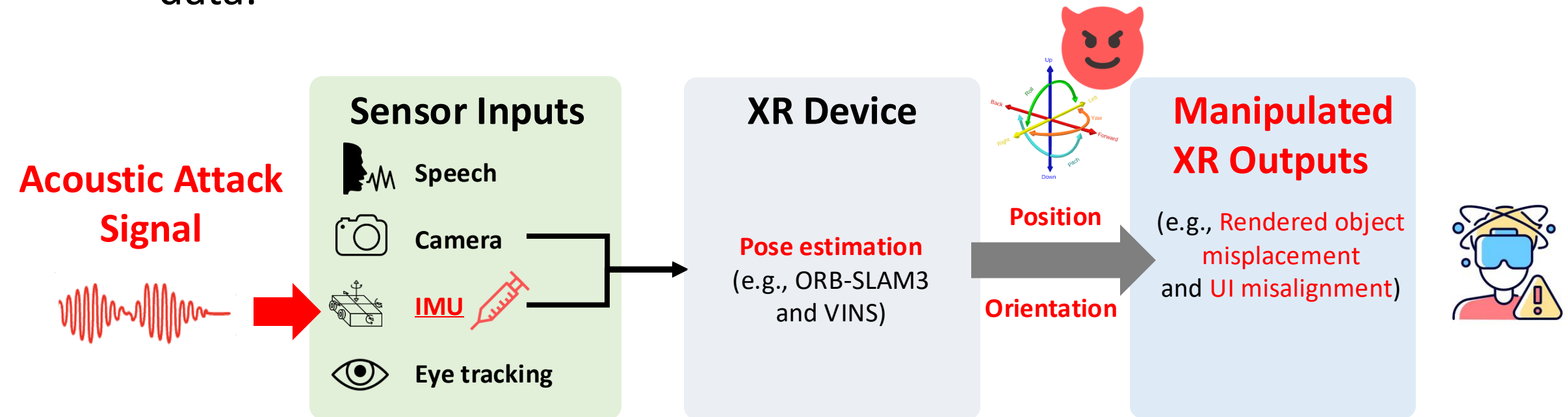
978-1-939133-04-5

of the emergent image stabilization hardware susceptible to acoustic manipulation and the object detection algorithms subject to adversarial examples, and design Poltergeist attacks (in short, PG attacks) that exploit such vulnerabilities.

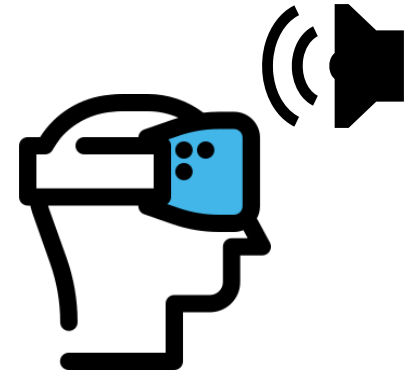
Unlike existing work that focused on altering what the main sensors (e.g., CMOS sensors) perceive by changing the visual appearance of an object [12], [38], [59], [27] or by

Background: Acoustic Attack

- Acoustic signals trigger MEMS resonance, corrupting IMU motion data.



Threat model

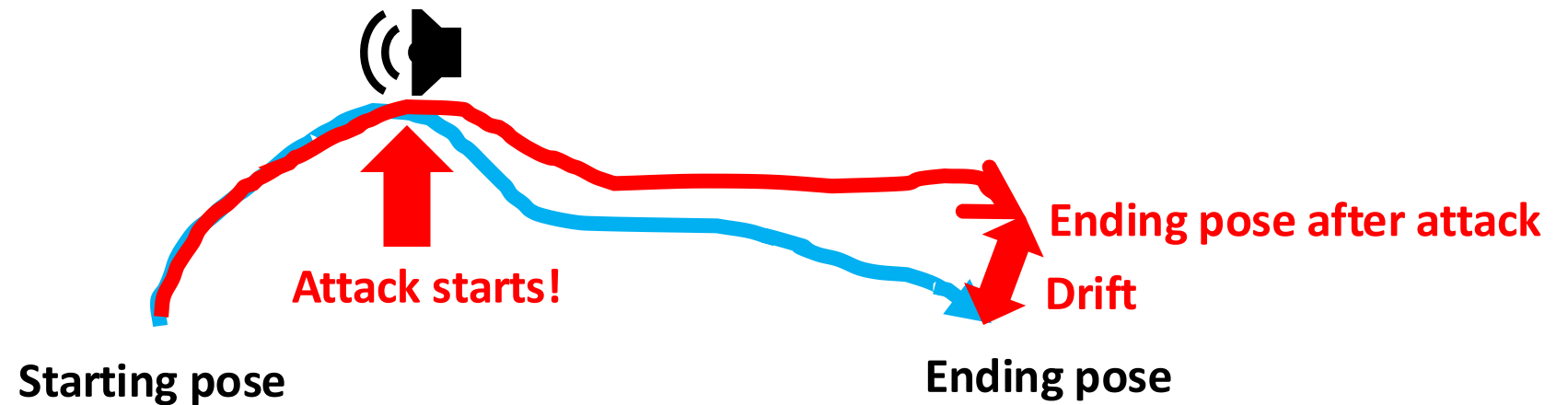


- **Attacker's capability:**
 - Play sound near victim — no physical or sensor access.
- **Attacker's resources:**
 - Profile same-model sensors to find resonant frequencies.
- **Attacker's goal:** benefit attacker (cheat) or harm victim (disrupt / induce errors).



Attacks on open-source systems

- Tested systems: ILLIXR [1] and ORB-SLAM3 [2].
- **Slow drifting attack** — when IMU bias is small.
 - Small and slow pose drift.

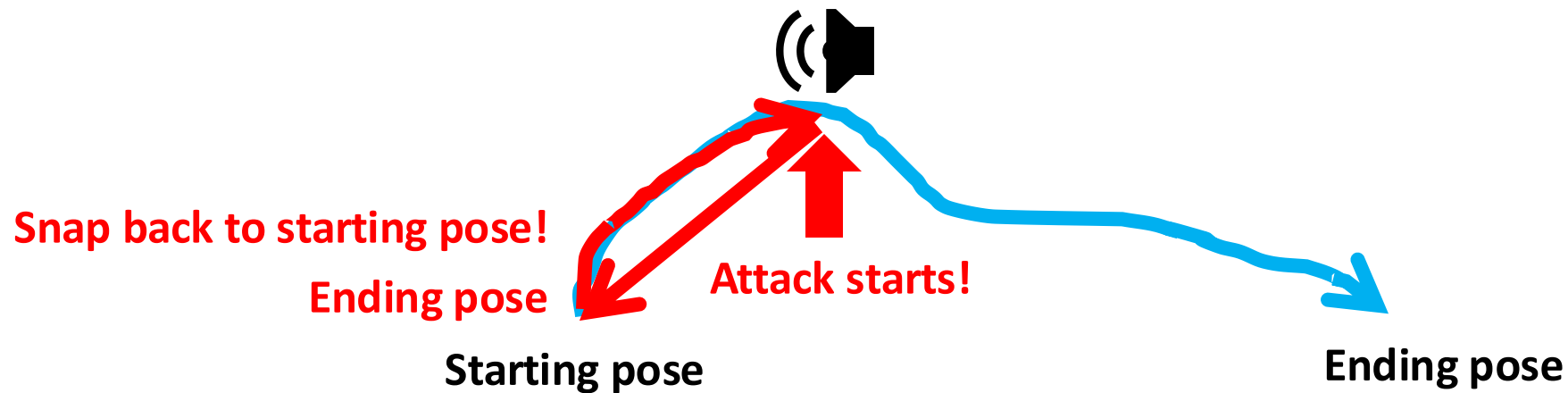


[1] Huzaifa, Muhammad, et al. "ILLIXR: Enabling end-to-end extended reality research." *2021 IEEE International Symposium on Workload Characterization (IISWC)*. IEEE, 2021.

[2] Campos, Carlos, et al. "Orb-slam3: An accurate open-source library for visual, visual-inertial, and multimap slam." *IEEE transactions on robotics* 37.6 (2021): 1874-1890.

Attacks on open-source systems

- **Slow drifting attack** — when IMU bias is small.
 - Steady and small pose bias.
- **Snapback attack** — when IMU bias is large.
 - Sudden pose “snap” to starting point.



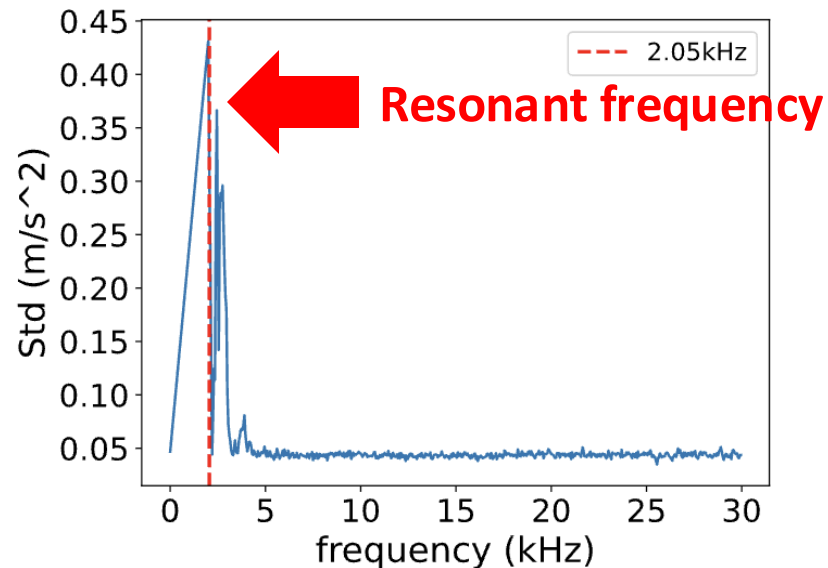
[1] Huzaifa, Muhammad, et al. "ILLIXR: Enabling end-to-end extended reality research." *2021 IEEE International Symposium on Workload Characterization (IISWC)*. IEEE, 2021.

[2] Campos, Carlos, et al. "Orb-slam3: An accurate open-source library for visual, visual-inertial, and multimap slam." *IEEE transactions on robotics* 37.6 (2021): 1874-1890.

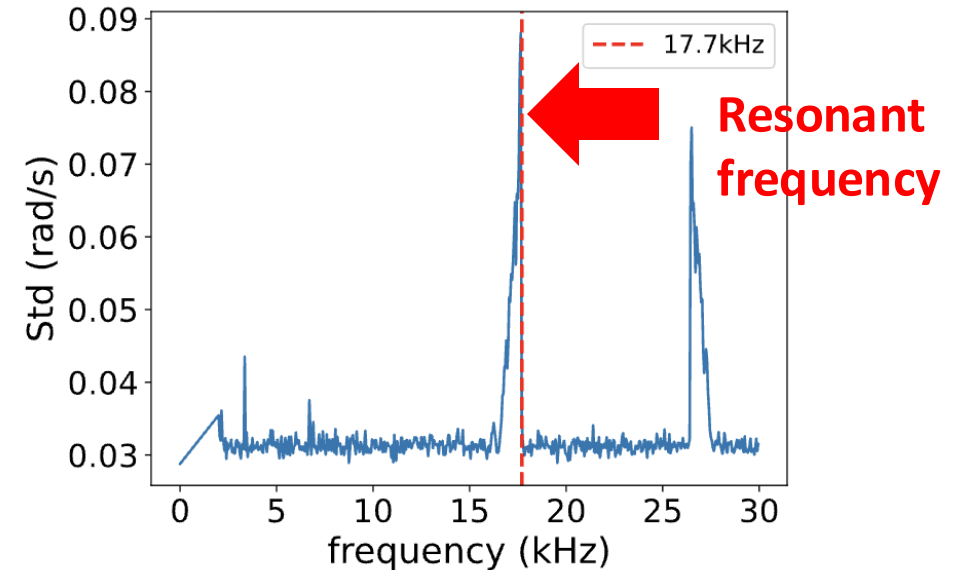
Attacks on Microsoft HoloLens

- Resonant frequencies of the HoloLens IMU.
- Snapback attack (Success rate: > 90%).

The accelerometer is vulnerable at 2.05 kHz.



Gyroscope is vulnerable at 17.7 kHz.



Attack 1: Manipulating user input (harm to user)

- **Attack motivation:** Many XR games map head motion to game controls.

Attack 1: Manipulating user input (harm to user)

- **Attack motivation:** Many XR games map head motion to game controls.
- **Method (attack design):** We place a HoloLens 2 on a remote-controlled car and inject acoustic signals targeting the IMU.



Attack 1: Manipulating user input (harm to user)

- **Outcome (attack result):**

- The car initially moves but then shifts slightly.
- Due to snapback, its estimated position resets to the start — the virtual car is stuck and the game is unusable for the victim.

Benign



(a) Benign, time=0



(b) Benign, time=1



(c) Benign, time=2

Car moving
direction

Attack 2: Clickjacking (harm to user)

- **Motivation:** World-anchored UI (e.g., virtual keyboard) drifting breaks input and can destroy user actions.

Benign:
User can correctly
type words on
virtual keyboard.



**The board is
anchored in the
fixed location.**

Attack 2: Clickjacking (harm to user)

- **Outcome:** The virtual keyboard drifts or disappears

Attack:
user mis-clicks
or cannot type.



**Virtual keyboard
drifts up and
down, making
user mis-click.**

Attack 2: Clickjacking (harm to user)

- Small user study: goals & setup.
 - **RQ1 (Usability degradation)** : How does the acoustic attack affect usability and task performance?
 - **RQ2 (Impact of the acoustic signal)**: Under what conditions is the acoustic signal more comfortable?
- **Participants & environment**: 5 volunteers, HoloLens 2, click-jacking app, 30-minute sessions.
 - RQ1: A timed typing task (three 3-letter words).
 - RQ2: Earmuffs or listening to music.

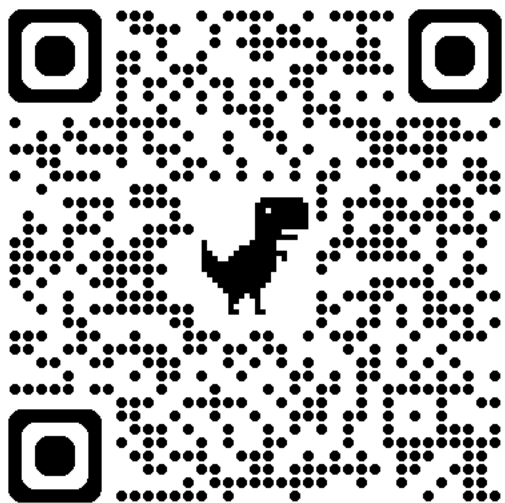


Attack 2: Clickjacking (harm to user)

- **Performance drop:** task completion fell from 100% (**benign**) to ~46.7% (**attack**).
- **Slower & harder:** average completion time rises (increasing from **6s** to **16s**); all participants said typing was harder under attack.
- **Practical note:** listening to music made the attack more tolerable and still effective — stealthier attack vector.

Conclusion

- First to study **acoustic attacks on XR pose estimation**.
- Identified **two novel effects** on XR pose estimation pipelines.
- Designed and demonstrated **four proof-of-concept attacks** on XR headsets.
- Conducted **user study** showing real usability degradation.
- Discussed **defense strategies**: hardware filtering, software re-localization.



Paper

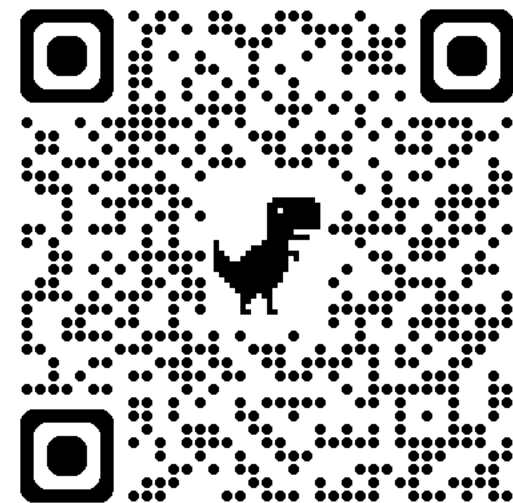
Thank you!

Any questions?

Yicheng Zhang

yzhang95@gmu.edu

<https://yichez.site>



Personal Page