Eavesdropping on Controller Acoustic Emanation for Keystroke Inference Attack in Virtual Reality

Authors: Shiqing Luo∗, Anh Nguyen∗, Hafsa Farooq†, Kun Sun∗, Zhisheng Yan∗

∗George Mason University

†Georgia State University

Presenter: Anh Nguyen

Diverse Applications

Gaming

Healthcare

Training

For Secure VR Systems

- Protect the sensitive data.

Problems of Existing VR Data Protection Works

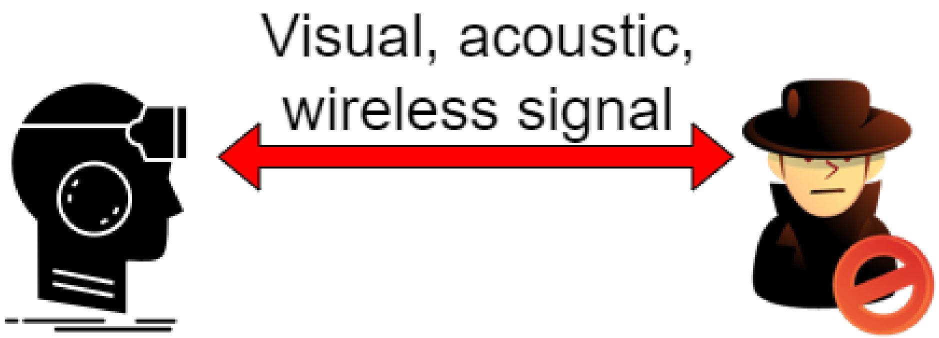
- Existing protection:
  - Verify users' identity using authentication schemes.
  - Prevent unauthorized access when attackers physically possess HMD.

Knowledge-based

Biometric-based

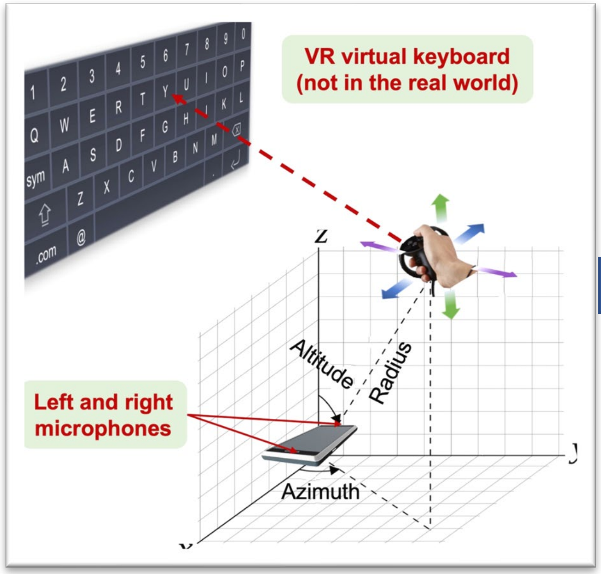
Problems of Existing VR Data Protection Works

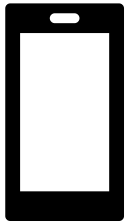
- Interaction between users and HMD is exposed to the public.
  - Attackers can record user interaction through side channels.
    - Video, wireless signal…
  - Attackers can infer sensitive input without possession of the HMD.



Keystroke Inference and Acoustic Signal

- We expose a VR keystroke inference attack exploiting the acoustic emanations from the controller.
  - Microphone records sounds anywhere around the victim.
  - More flexible placement than existing side channels.



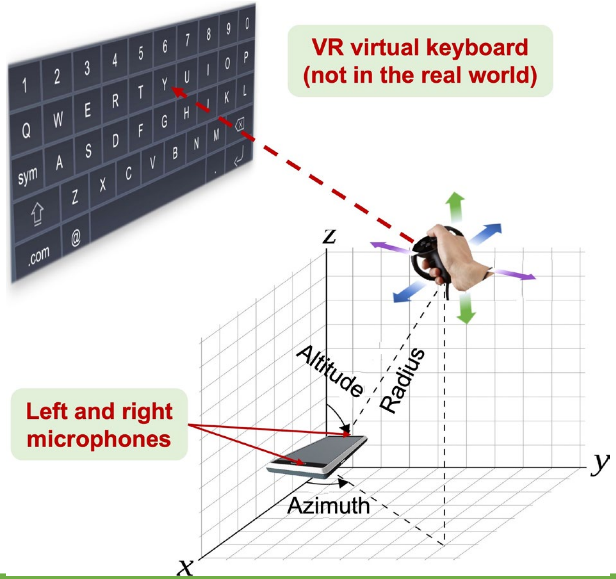


record sound

Top-5 candidates:
jeep
jep
jepp
jeo
jeeo

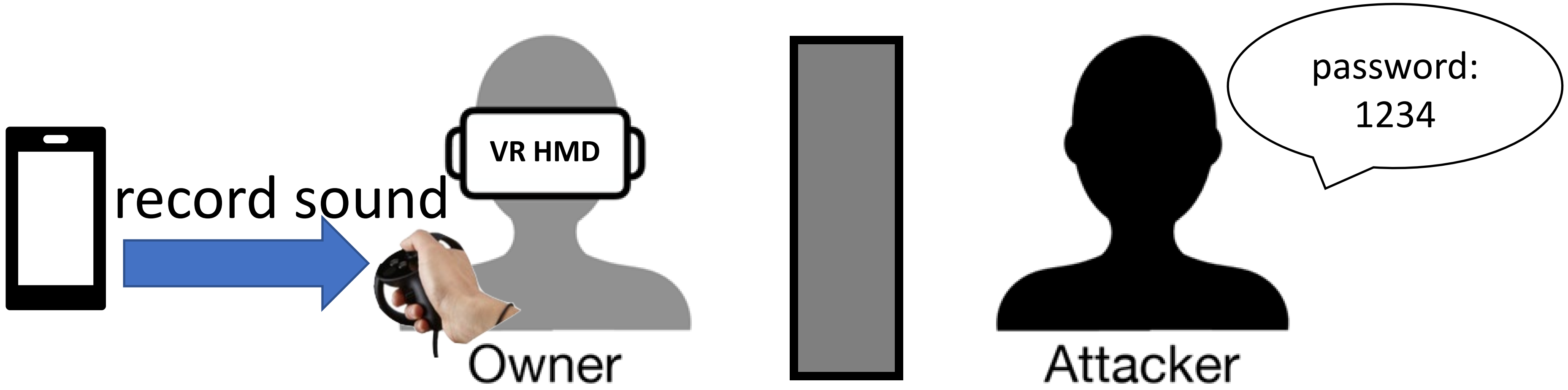
Keystroke Inference and Acoustic Signal

- Controller-based keystroke method in VR:
  - Users move a controller to navigate a virtual cursor through a virtual keyboard.
  - Users click the confirm button to commit keystrokes.
- Attackers place a malicious smartphone nearby.
  - The controller emits keystroke clicking sounds at various locations.
  - The sound arrive the smartphone from different directions.
  - Each keystroke uniquely defines its clicking sound direction.



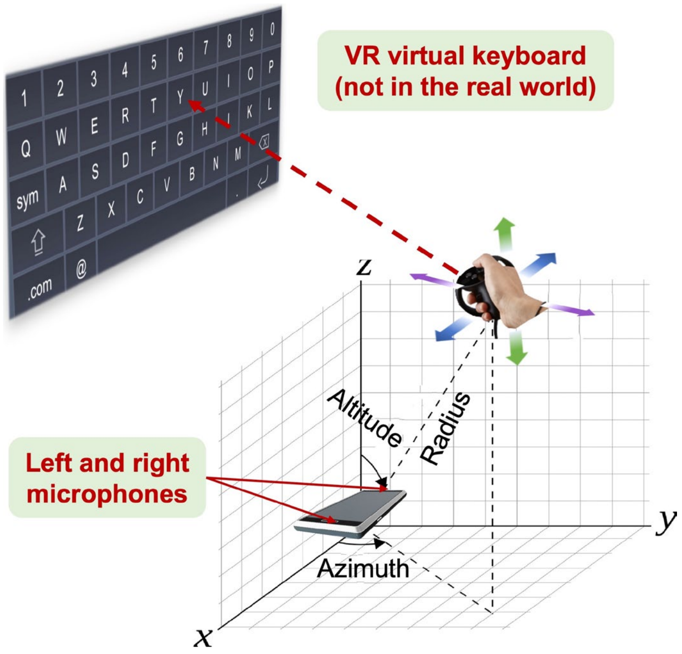
Threat Model

- The user enters sensitive information in VR in a confined setting
  - e.g., a shared table in a library.
- A malicious smartphone is placed nearby.
- The smartphone records the acoustic emanations from the controller.
- The attacker knows the layout of the virtual keyboard.



Challenge 1 – Sound Source in 3D Space

- It is difficult to differentiate the clicking sounds.
  - Sound source locations are distributed in 3D space rather than 2D plane*.
  - The omnidirectional smartphone microphone cannot differentiate keystroke sounds.



*Compagno, Alberto, Mauro Conti, Daniele Lain, and Gene Tsudik. "Don't skype & type! acoustic eavesdropping in voice-over-ip." In *Proceedings of the 2017 ACM on Asia Conference on Computer and Communications Security*, pp. 703-715. 2017.

Challenge 2 – Various User Microphone Placement

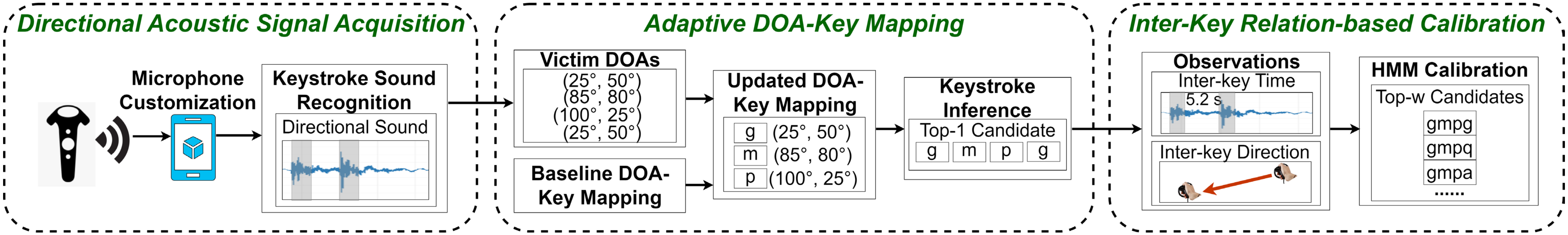
- The relative position and orientation between the user and the microphones vary in different attack scenarios.
  - The direction of arrival (DOA) of the same key's keystroke sounds varies.
  - The mapping between DOAs and keys varies.

Challenge 3 – Mapping Error

- No strict one-to-one mapping between keystroke sound and keys.
  - Users may rotate the controller to navigate the cursor.
  - Keystroke sound with the same DOA results in different keys.

Heimdall System Architecture

- Module 1: record differentiable keystroke sound using directional microphones.
- Module 2: adapt DOA-Key mapping to the attack case.
- Module 3: correct the mapping errors using a Hidden Markov Model (HMM).



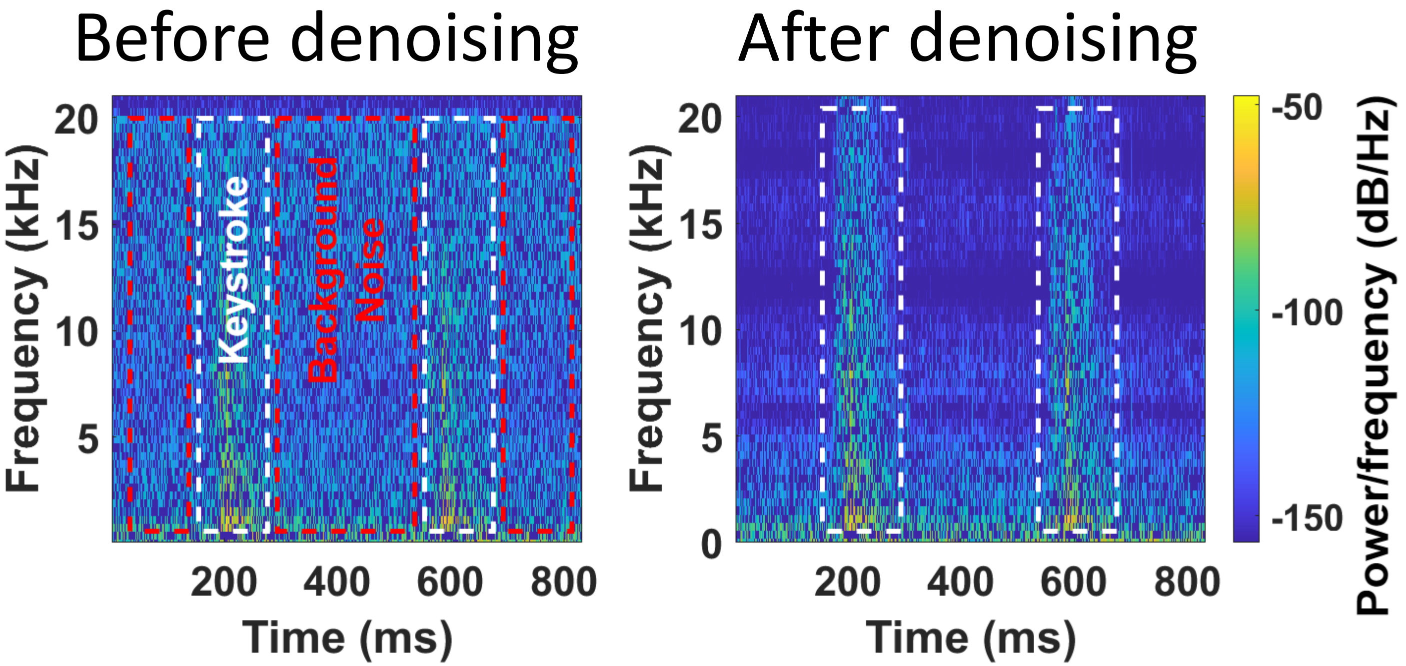
Module 1 – Hardware Design

- Convert the omnidirectional microphones into directional ones.
  - Inspired by a shotgun directional microphone.
  - Two tubes with side slots are attached to the microphones.
  - The microphones change the intensity and phase of recorded sounds based on their DOA.

Mason, W. P., and R. N. Marshall. "A tubular directional microphone." *The Journal of the Acoustical Society of America* 10, no. 3 (1939): 206-215.

Module 1 – Acoustic Signal Acquisition

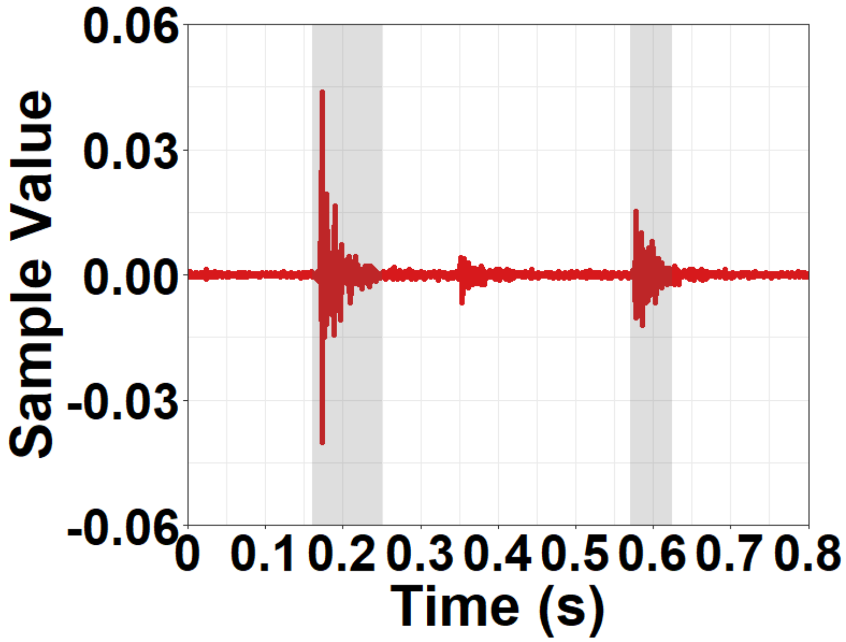
- Remove ambient noise using wavelet denoising*.



*Rathore, Aditya Singh, Weijin Zhu, Afee Daiyan, Chenhan Xu, Kun Wang, Feng Lin, Kui Ren, and Wenyao Xu. "Sonicprint: A generally adoptable and secure fingerprint biometrics in smart devices." In *Proceedings of the 18th International Conference on Mobile Systems, Applications, and Services*, pp. 121-134. 2020.

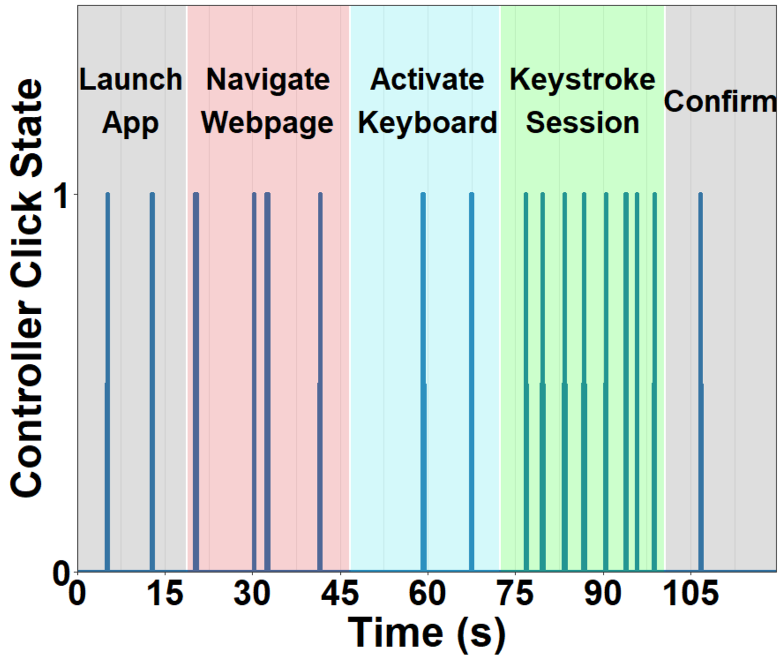
Module 1 – Signal Processing

- Signal peaks exceeding a threshold are extracted as keystroke clicking sounds.
  - Optimal threshold value is 0.028.



Module 1 – Keystroke Session Recognition

- We identify keystroke session based on the controller click frequency*.



* Wu, Yi, Cong Shi, Tianfang Zhang, Payton Walker, Jian Liu, Nitesh Saxena, and Yingying Chen. "Privacy Leakage via Unrestricted Motion-Position Sensors in the Age of Virtual Reality: A Study of Snooping Typed Input on Virtual Keyboards." In *2023 IEEE Symposium on Security and Privacy (SP)*, pp. 3382-3398. IEEE Computer Society, 2023.

Module 2 – Keystroke Mapping

- We derive a baseline DOA-Key mapping table.
  - The mapping stores DOA and their corresponding keys (26 English letters and 10 digits).
- We update the baseline DOA-key mapping to match the DOAs in the attack.

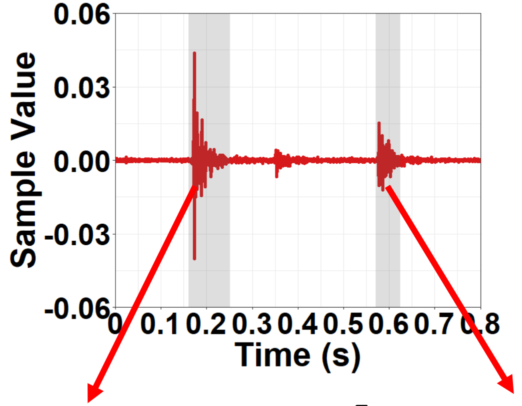
DOA [azimuth altitude]	Key
[-80° 70°]	1
[-60° 70°]	2
…	…
[40° -20°]	N
[60° -20°]	M

update

DOA [azimuth altitude]	Key
[-60° 40°]	1
[-40° 40°]	2
…	…
[20° -50°]	N
[40° -50°]	M

Module 2 – Keystroke Mapping

- We infer keystrokes in attack by looking up their DOA in the updated mapping.



[20° -50°] [-40° 40°]

search →

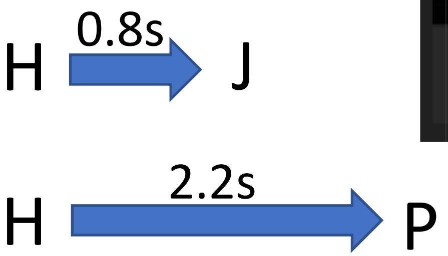
DOA [azimuth altitude]	Key
[-60° 40°]	1
[-40° 40°]	2
…	…
[20° -50°]	N
[40° -50°]	M

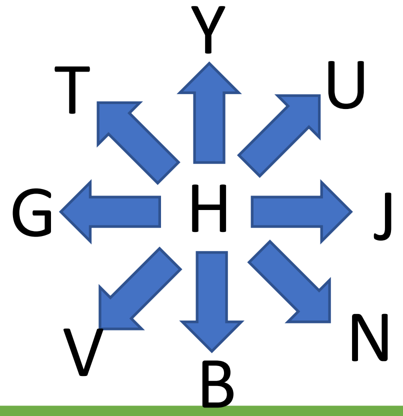
infer →

N,2…

Module 3 – Transition between Keystrokes

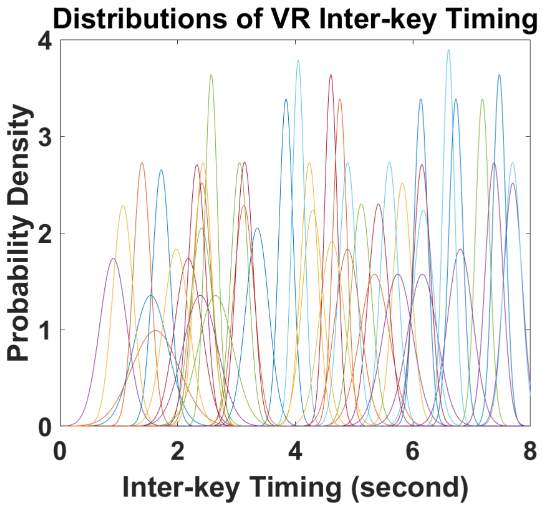
- Transition between keystrokes contains information on the keys.
  - Time interval between two successive keystrokes depends on their inter-distance.
  - Controller moving direction committing two keystrokes depends on keys location on the keyboard.



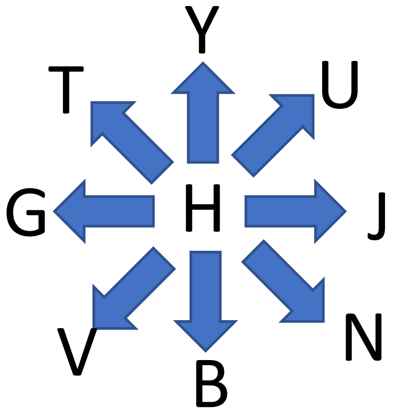


Module 3 – Keystroke Mapping Correction

- A Hidden Markov Model (HMM) models the keystroke transition.
  - The inter-key time and controller moving direction are observations.
  - The keys are hidden state.



Derive inter-key time from acoustic signals.



Derive inter-key directions based on DOA.

Module 3 – Keystroke Mapping Correction

- We use the HMM to correct mapping errors.
  - The model generates a list of key sequences given on the inter-key time and direction.
  - We select the sequences with high probability and similarity with the original mapping result.

Mapping result: H2GM

HMM generate

Top-5 candidate list
HWGM
H2HM
H3GM
H2GN
HWGN

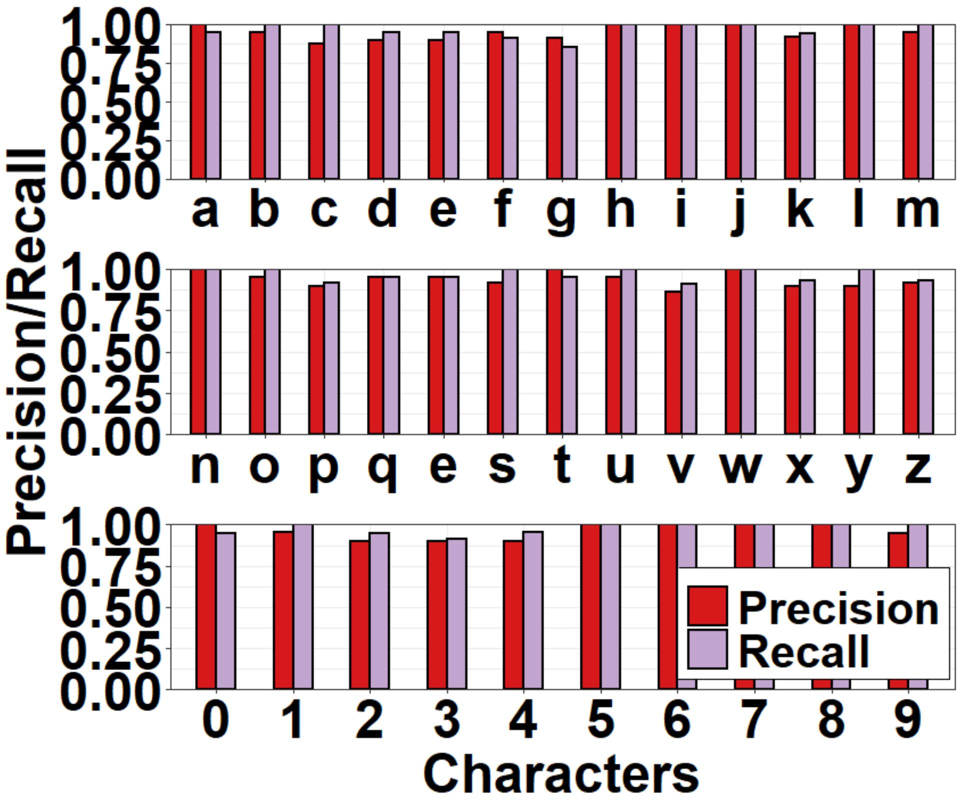
Experiment Setup

- 30 participants.
  - They first take a practice session: input several keys as instructed.
  - Each of them inputs 45 passwords for testing.
- Length of passwords ranges from 4 to 8 characters.
  - Passwords are the most common ones*.

*SecLists. 1000 most common passwords. https: //github.com/danielmiessler/SecLists/ blob/master/Passwords/Common-Credentials/ 10-million-password-list-top-1000.txt, 2020. Accessed: 2021-09-15.

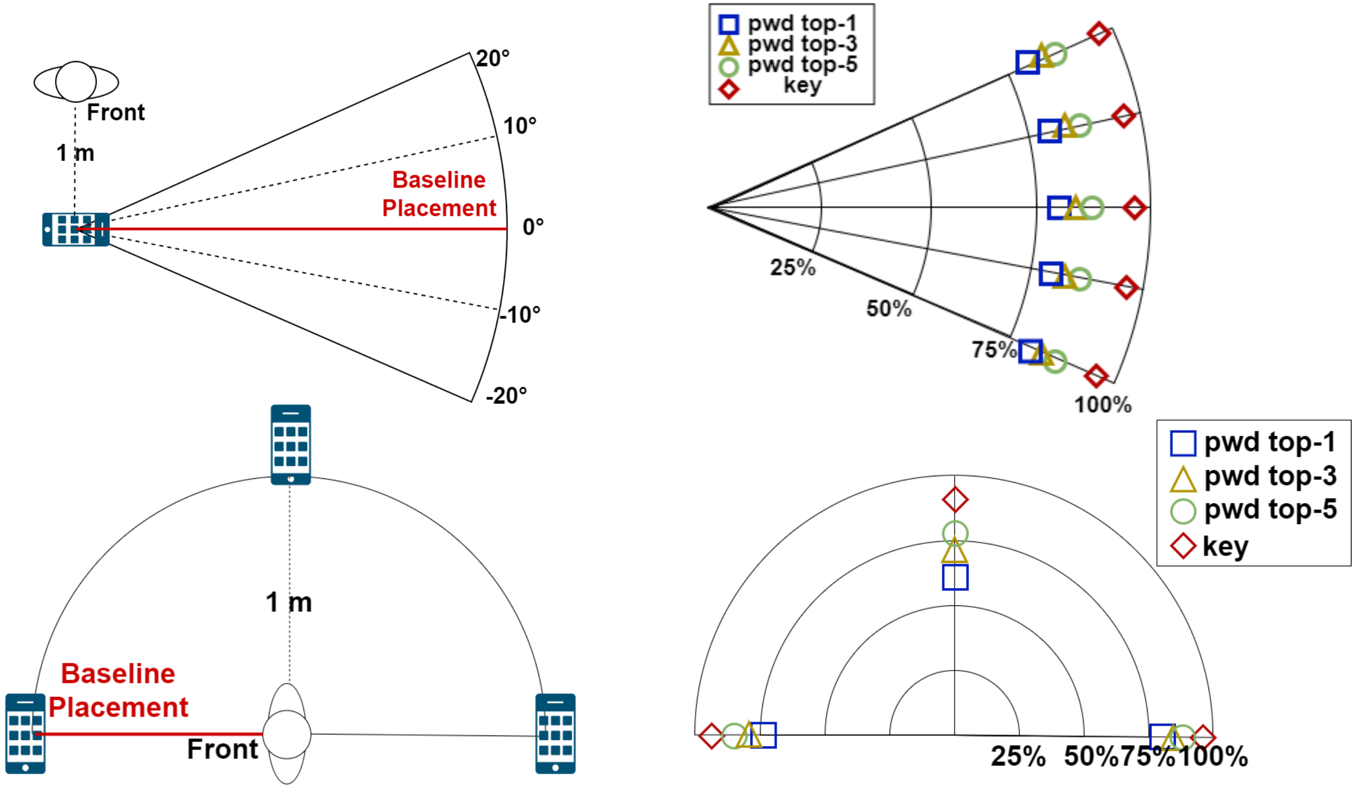
Experiment – Benchmark for All Characters

- The average precision and recall across characters reach 95.14% and 96.29%.



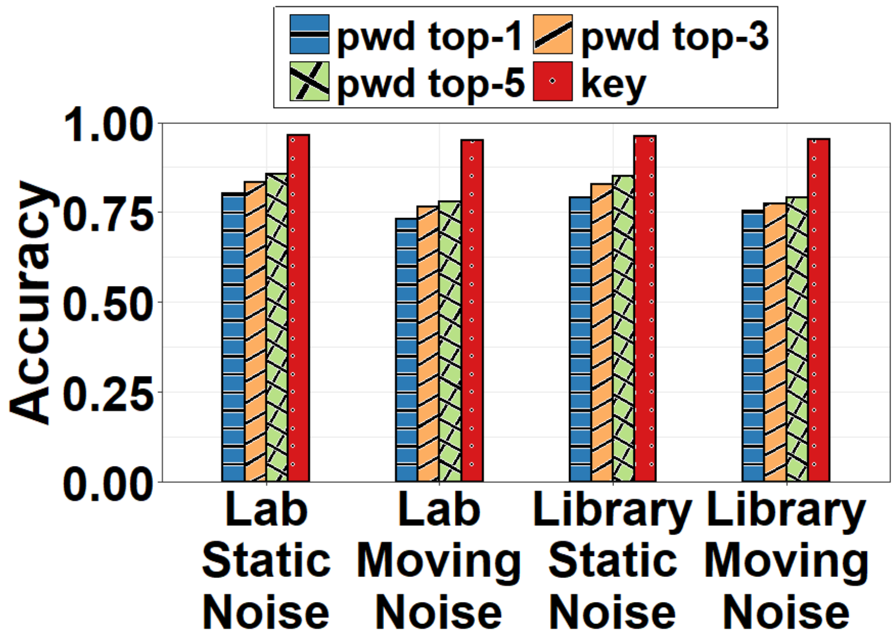
Experiment – Smartphone Placement

- The top-w and key accuracy are consistent across different user-smartphone placements.



Experiment – Different Environment

- The attack can be generalized to different scenarios.
  - Static noise: desktop fans and air conditioners.
  - Moving noise: people talking and walking.



Conclusion

- We expose a security threat in VR systems that allow attackers to infer input by analyzing the acoustic emanations from the controller.
- We propose Heimdall, a placement-flexible acoustic keystroke inference attack in VR.
- We extensively evaluate the Heimdall system in terms of keystroke inference accuracy and its robustness.

Thank you