Remote Keylogging in Multi-user VR Applications

Zihao Su*, Kunlin Cai*, Reuben Beeler, Lukas Dresel, Allan Garcia, Ilya Grishchenko, Yuan Tian, Christopher Kruegel, and Giovanni Vigna

*co-first-author

Virtual Reality as a Social Platform

Virtual Reality as a Social Platform

Across the Metaverse: My trip though VR social platforms

I tried out VRChat, Meta Horizon Worlds, and Rec Room. Each one makes a different case for the future of social gameplay.

The Evolution of Social VR Platforms:

The rise of **social VR platforms** is a game-changer, allowing real-time interaction within virtual environments. Users can **host parties, attend virtual concerts, or team up in multiplayer games**. These platforms are becoming increasingly user-friendly, diverse, and community-focused, fostering a more connected and social future of virtual reality.

Virtual Reality as a Social Platform

Across the Metaverse: My trip though VR social platforms

I tried out VRChat, Meta Horizon Worlds, and Rec Room. Each one makes a different case for the future of social gameplay.

FORBES > LEADERSHIP > CMO NETWORK

Social VR, Facebook Horizon And The Future Of Social Media Marketing

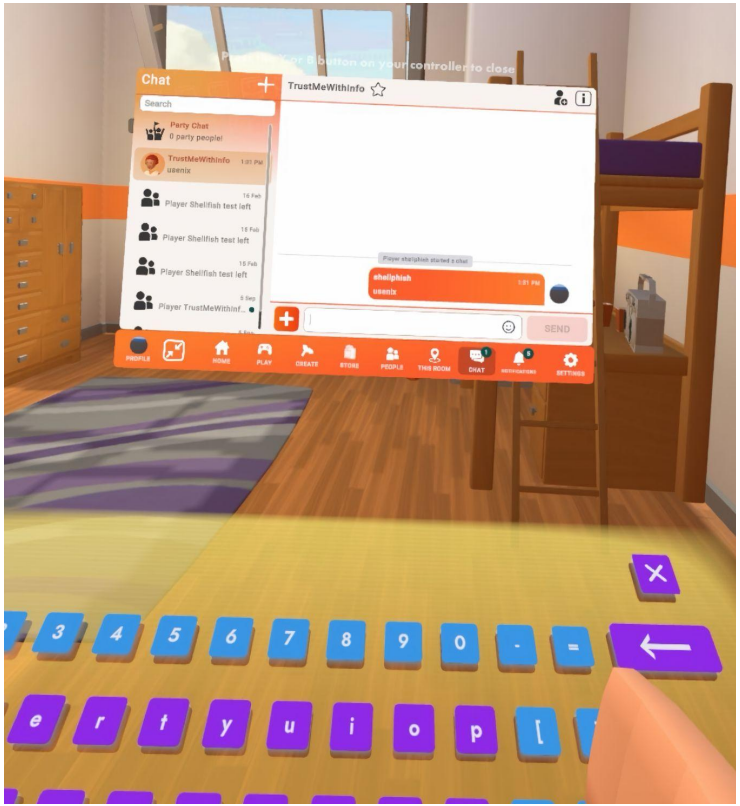
Is social virtual reality the next big thing?

[redacted] | Sep 17, 2021

The Evolution of Social VR Platforms:

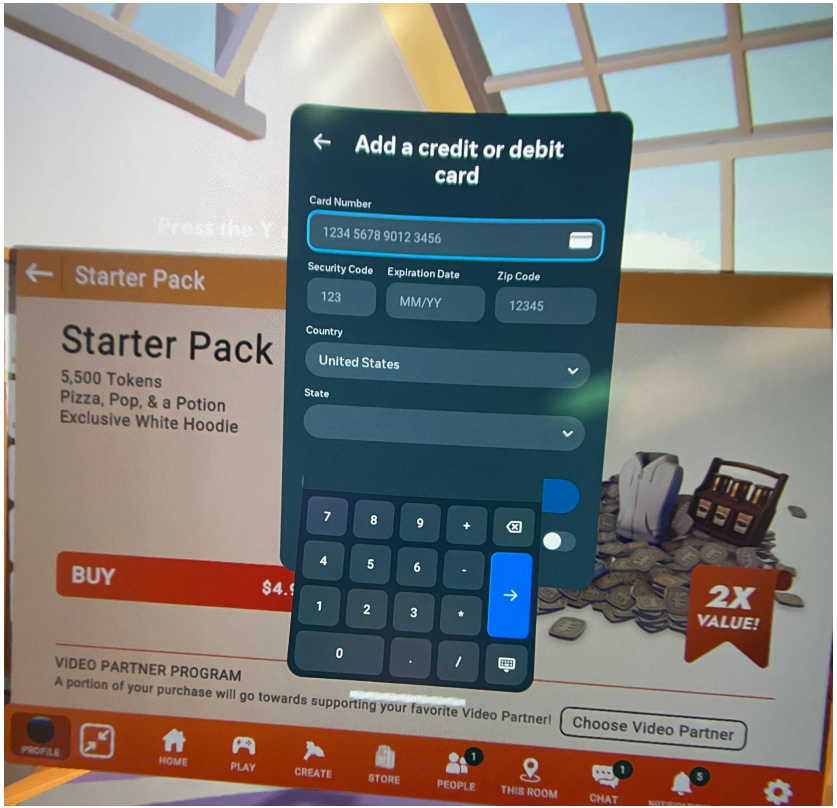
The rise of **social VR platforms** is a game-changer, allowing real-time interaction within virtual environments. Users can **host parties, attend virtual concerts, or team up in multiplayer games**. These platforms are becoming increasingly user-friendly, diverse, and community-focused, fostering a more connected and social future of virtual reality.

Typing In Multi-user VR Applications



Typing In Multi-user VR Applications

Typing In Multi-user VR Applications



Typing In Multi-user VR Applications

Did you know that your typing movement is exposed it to other users in the same virtual room?

Typing In Multi-user VR Applications

Did you know that your typing movement is exposed it to other users in the same virtual room?

Common oversight by developers: avatar keeps being rendered when typing

Typing In Multi-user VR Applications

Typing In Multi-user VR Applications

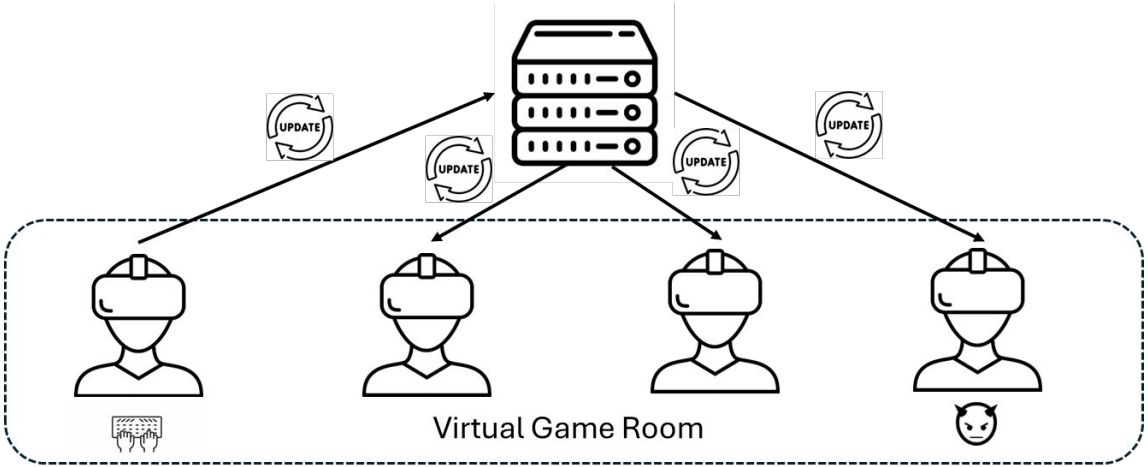
Can we accurately reconstruct keys remotely based on the typing movement?

Attack Steps

1. Be a legitimate user of the app (e.g., download client, register account)

2. Perform experiments to prepare for the attack (e.g., observe keyboard layout)

3. Join a public room with any victim.

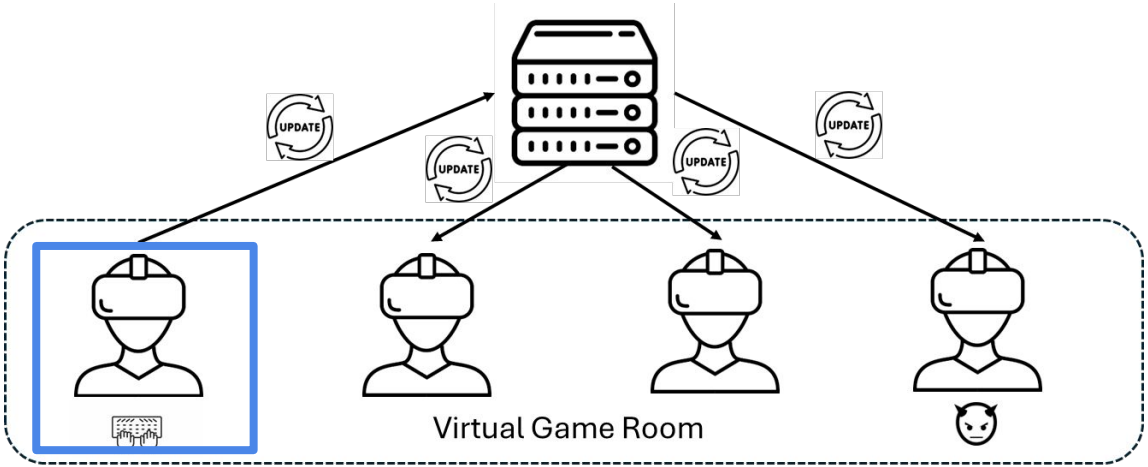


Attack Steps

1. Be a legitimate user of the app (e.g., download client, register account)

2. Perform experiments to prepare for the attack (e.g., observe keyboard layout)

3. Join a public room with any victim.

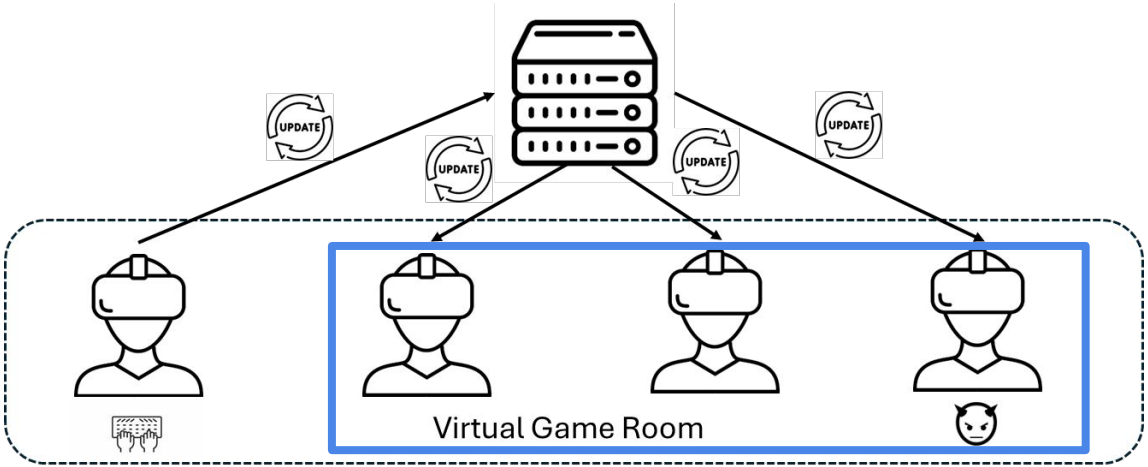


Attack Steps

1. Be a legitimate user of the app (e.g., download client, register account)

2. Perform experiments to prepare for the attack (e.g., observe keyboard layout)

3. Join a public room with any victim.

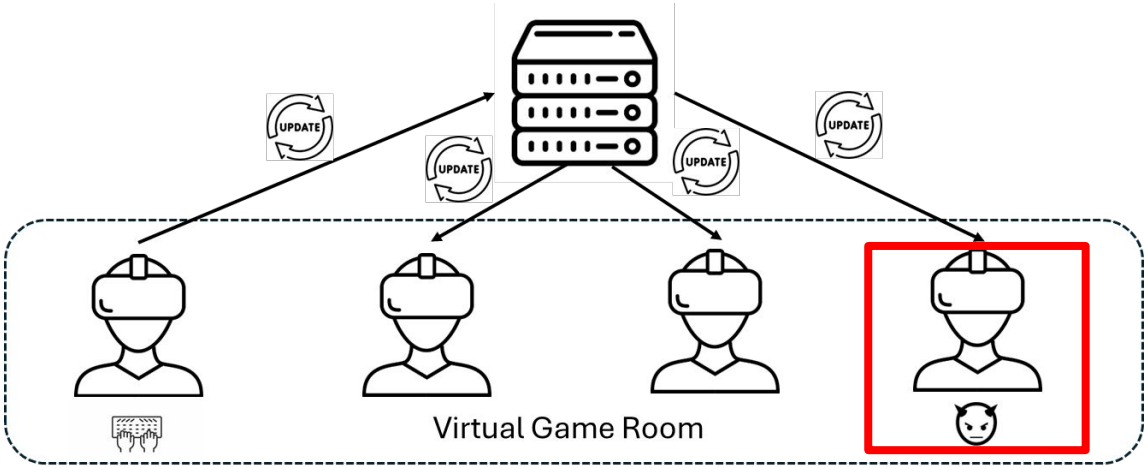


Attack Steps

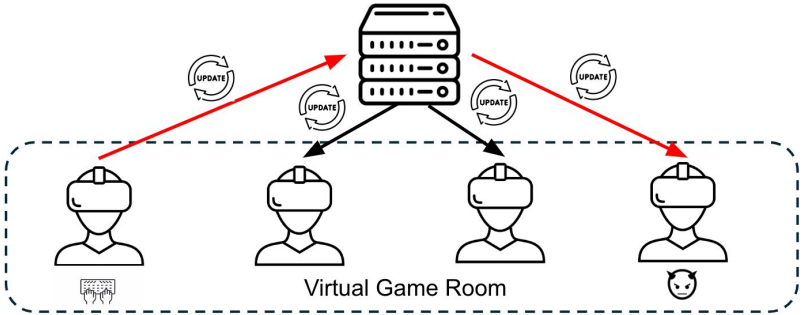
1. Be a legitimate user of the app (e.g., download client, register account)

2. Perform experiments to prepare for the attack (e.g., observe keyboard layout)

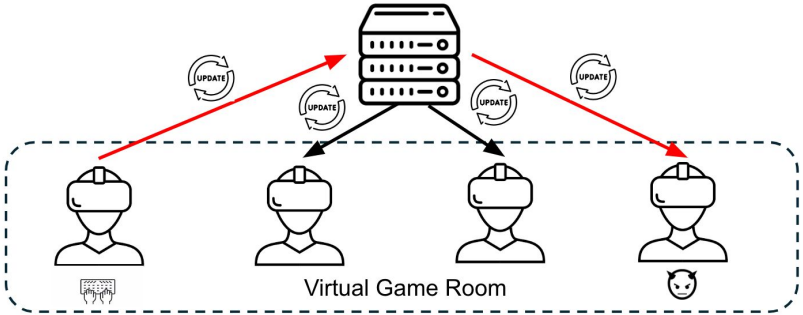
3. Join a public room with any victim.

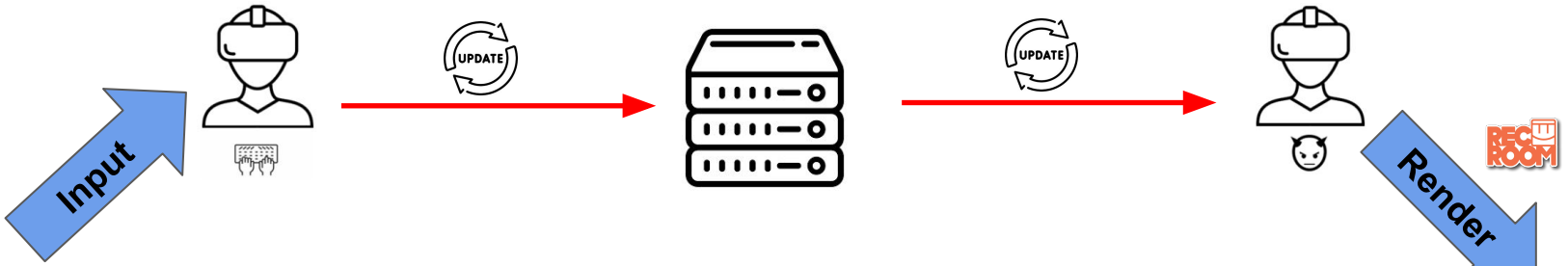


Method

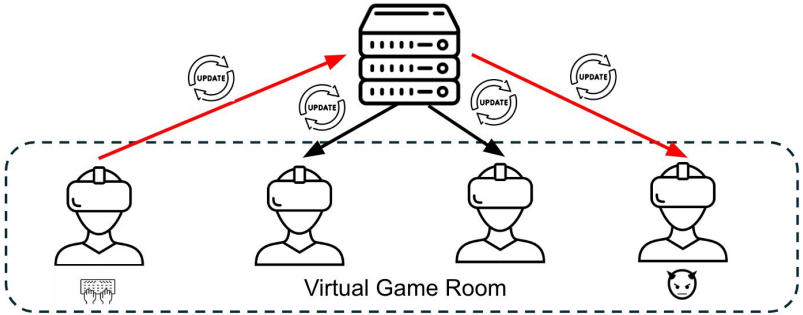


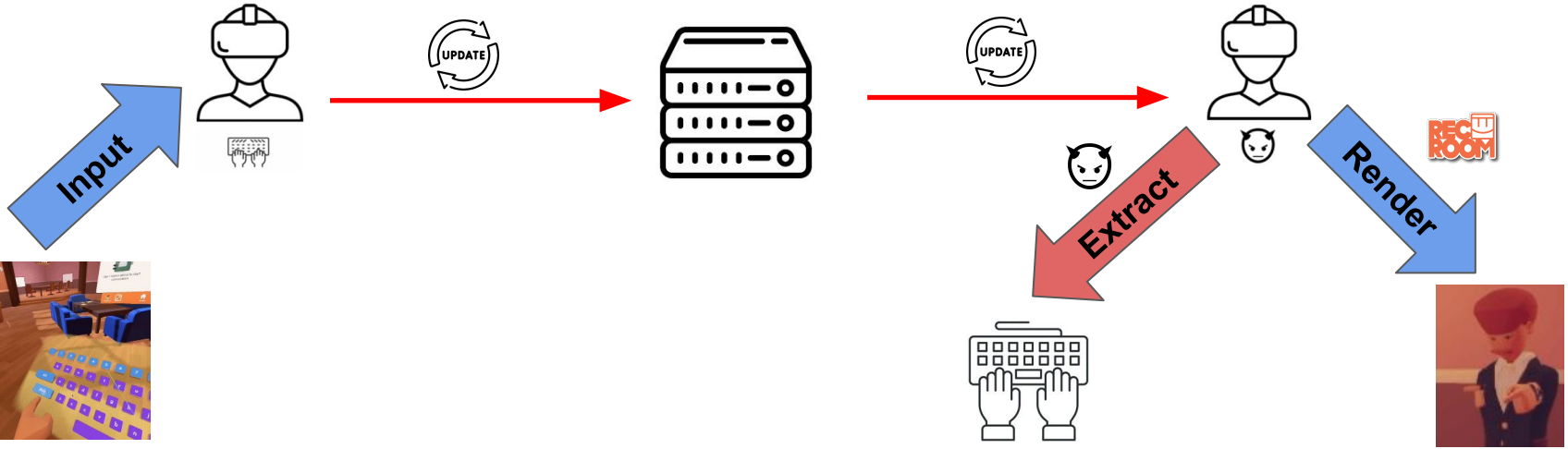
Method


UPDATE
UPDATE
UPDATE
UPDATE
Virtual Game Room

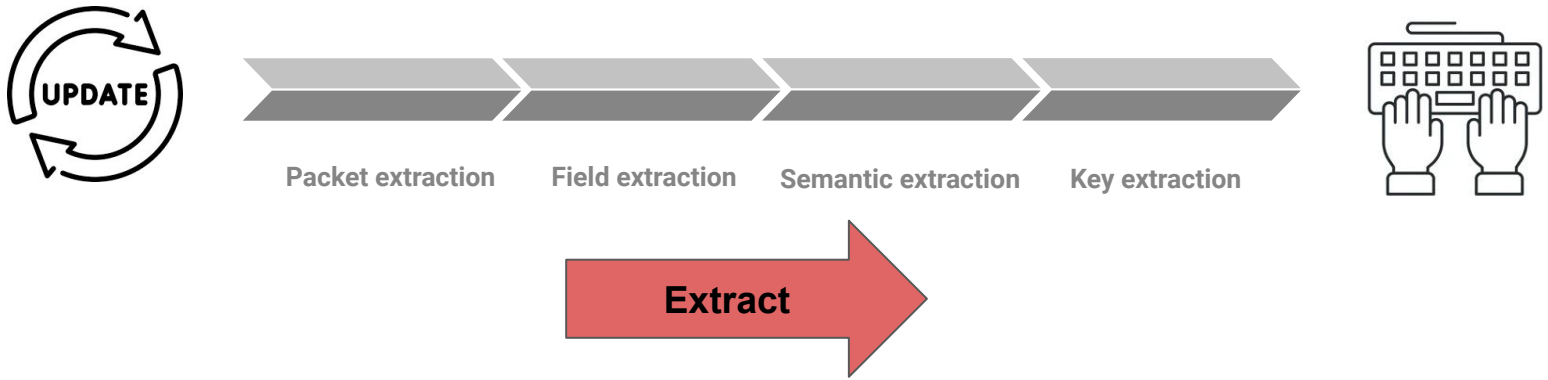

Input
UPDATE
UPDATE
Render
REC ROOM

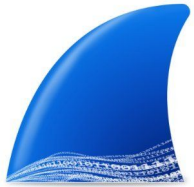
Method


UPDATE
UPDATE
UPDATE
UPDATE
Virtual Game Room


Input
UPDATE
UPDATE
Extract
Render
REC ROOM

Method


UPDATE
Packet extraction
Field extraction
Semantic extraction
Key extraction
Extract



Packet 1

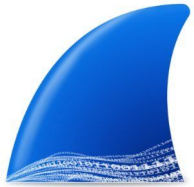
Packet 2

Packet 3

...

Packet 10035

Packet 10036



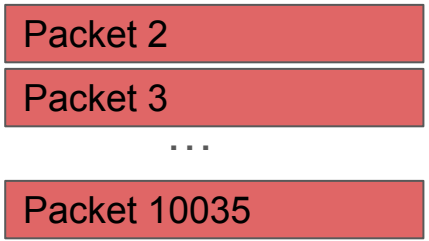
Packet 1
Packet 2
Packet 3
...
Packet 10035
Packet 10036

Extract

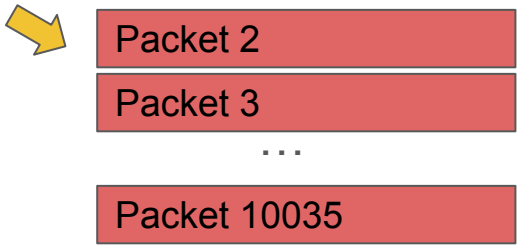
Rec Room packets

Packet 2
Packet 3
...
Packet 10035

Rec Room packets



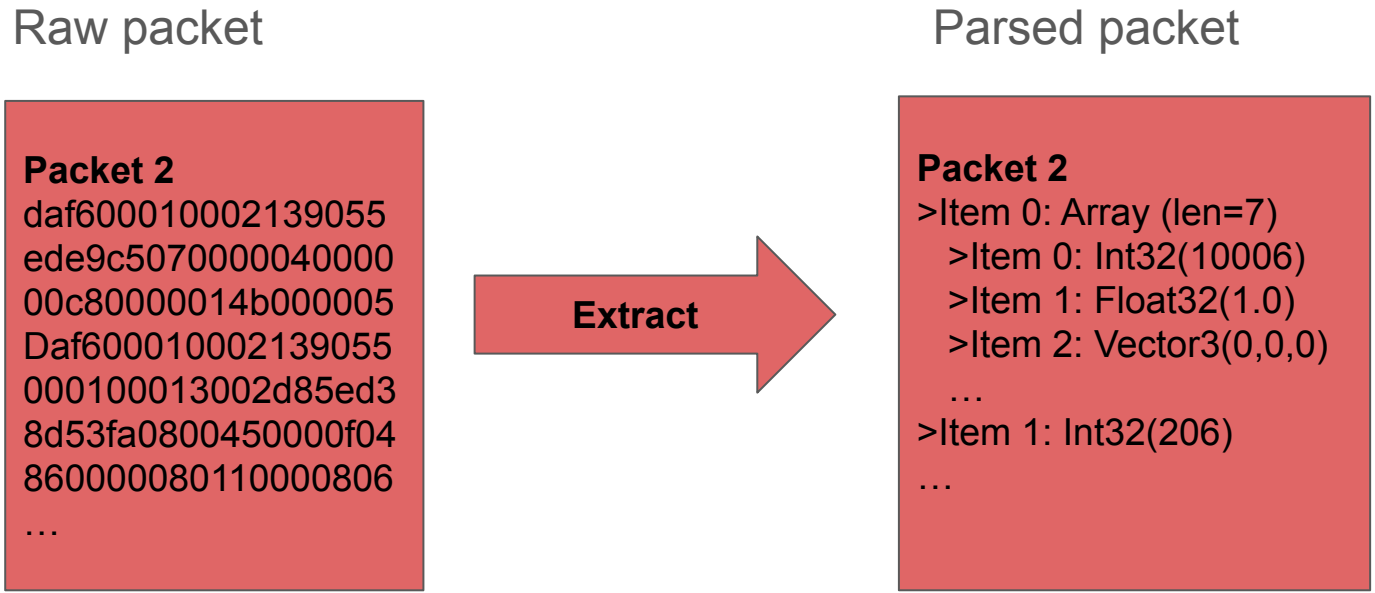
Rec Room packets



Raw packet

Packet 2
daf600010002139055
ede9c5070000040000
00c80000014b000005
Daf600010002139055
000100013002d85ed3
8d53fa0800450000f04
860000080110000806
…

Packet extraction | Field extraction | Semantic extraction | Key extraction


Raw packet
Parsed packet
Packet 2
daf600010002139055
ede9c5070000040000
00c80000014b000005
Daf600010002139055
000100013002d85ed3
8d53fa0800450000f04
860000080110000806
…
Extract
Packet 2
>Item 0: Array (len=7)
>Item 0: Int32(10006)
>Item 1: Float32(1.0)
>Item 2: Vector3(0,0,0)
…
>Item 1: Int32(206)
…

Parsed packet

Packet 2
>Item 0: Array (len=7)
  >Item 0: Int32(10006)
  >Item 1: Float32(1.0)
  >Item 2: Vector3(0,0,0)
  …
>Item 1: Int32(206)
…

Parsed packet

Packet 2
>Item 0: Array (len=7)
  >Item 0: Int32(10006)
  >Item 1: Float32(1.0)
  >Item 2: Vector3(0,0,0)
  …
>Item 1: Int32(206) ?
…

Packet extraction | Field extraction | Semantic extraction | Key extraction

Parsed packet

Packet 2
>Item 0: Array (len=7)
  >Item 0: Int32(10006)
  >Item 1: Float32(1.0)
  >Item 2: Vector3(0,0,0)
  …
>Item 1: Int32(206)
…

Extract

Parsed packet with semantics

Packet 2
>Item 0: Array (len=7)
  >Object ID: Left hand
  >Click down: Float32(1.0)
  >Position: Vector3(0,0,0)
  …
>UserID: Int32(206)
…

Packet extraction | **Field extraction** | **Semantic extraction** | Key extraction

Parsed packet

Packet 2
>Item 0: Array (len=7)
  >Item 0: Int32(10006)
  >Item 1: Float32(1.0)
  >Item 2: Vector3(0,0,0)
  …
>Item 1: Int32(206)
…

Extract

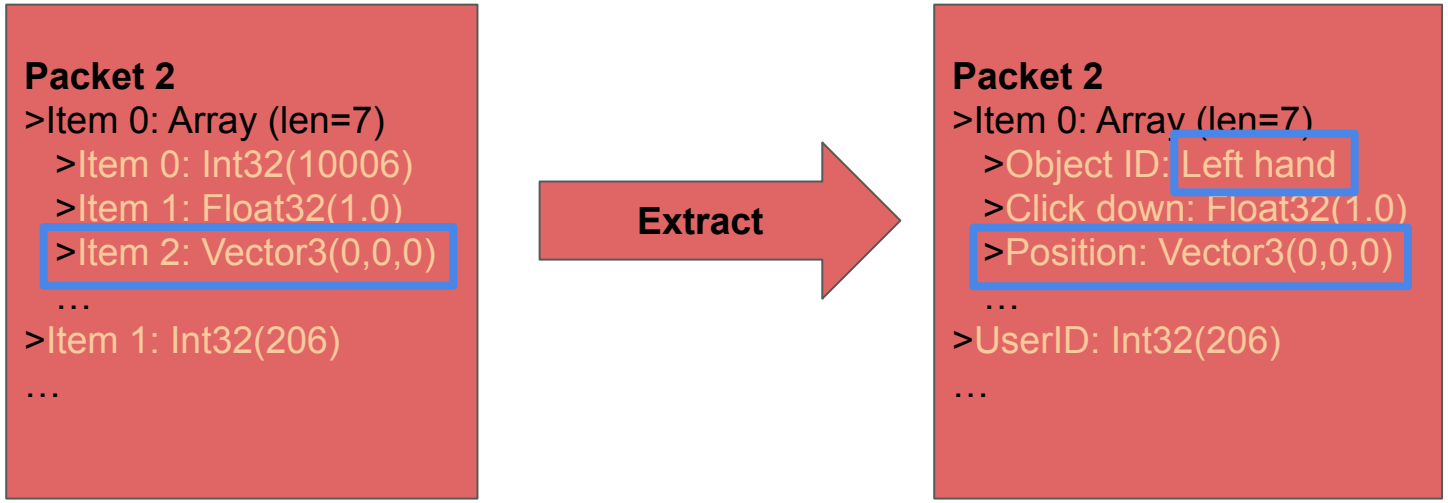
Parsed packet with semantics

Packet 2
>Item 0: Array (len=7)
  >Object ID: Left hand
  >Click down: Float32(1.0)
  >Position: Vector3(0,0,0)
  …
>UserID: Int32(206)
…

Packet extraction | Field extraction | Semantic extraction | Key extraction

Parsed packet

Parsed packet with semantics



Packet extraction | Field extraction | Semantic extraction | Key extraction

Parsed packets with semantics

Packet 2

Packet 3

...

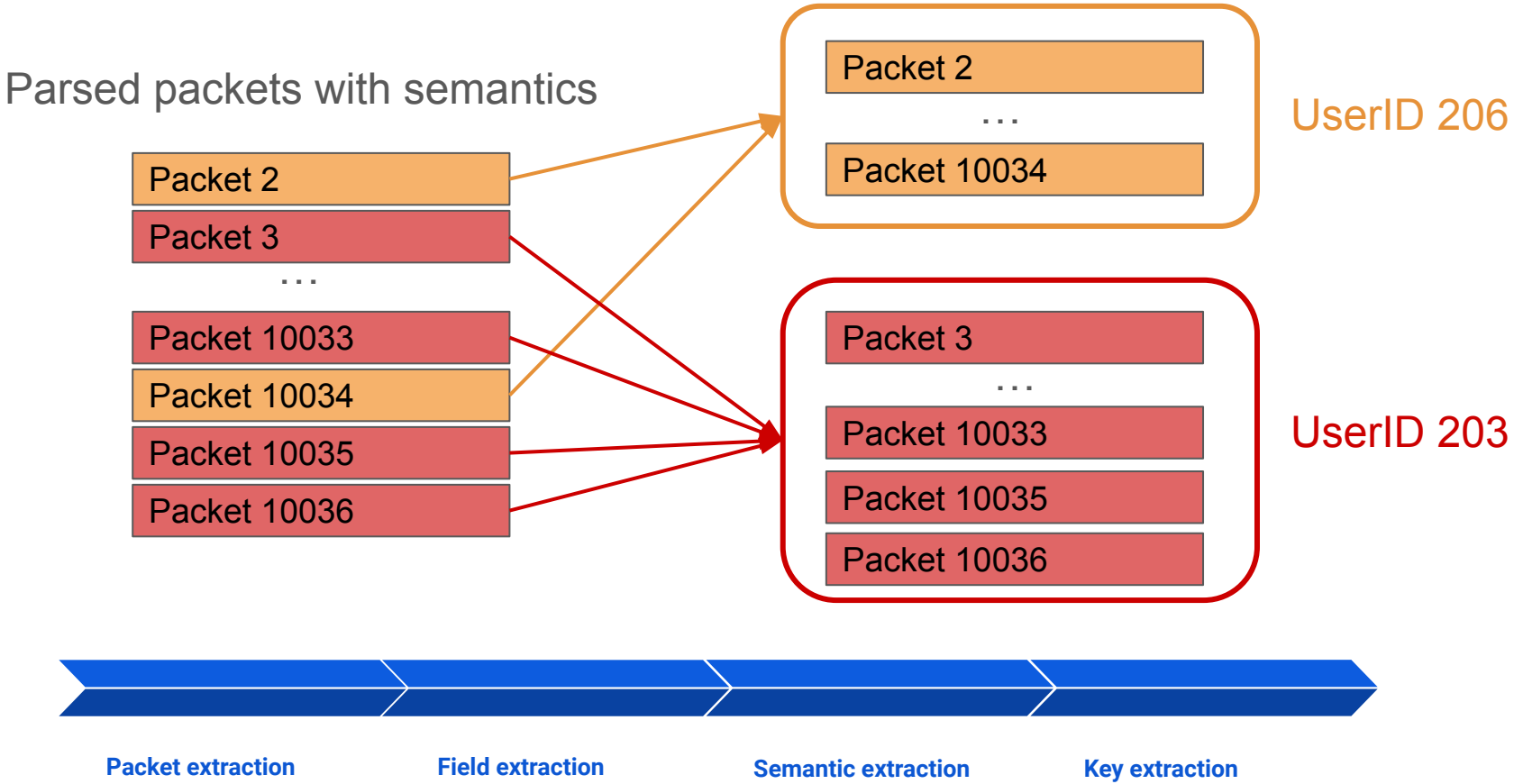
Packet 10033

Packet 10034

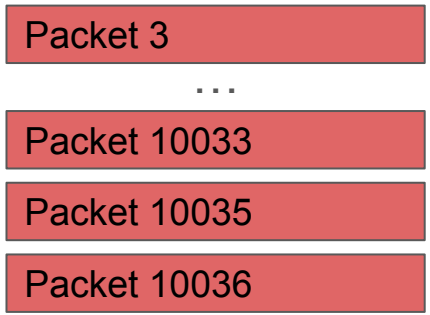
Packet 10035

Packet 10036

Packet extraction | Field extraction | Semantic extraction | Key extraction

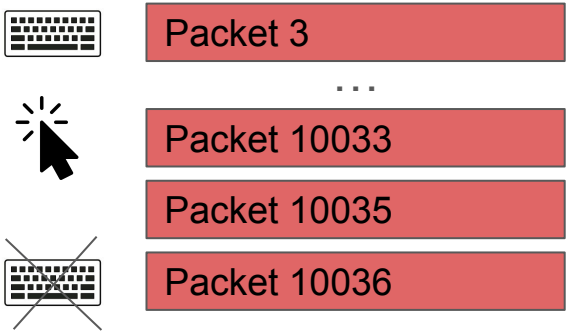

Parsed packets with semantics
Packet 2
Packet 3
...
Packet 10033
Packet 10034
Packet 10035
Packet 10036
Packet 2
...
Packet 10034
UserID 206
Packet 3
...
Packet 10033
Packet 10035
Packet 10036
UserID 203
Packet extraction
Field extraction
Semantic extraction
Key extraction

UserID 203



Packet extraction | Field extraction | Semantic extraction | Key extraction

UserID 203



Packet extraction | Field extraction | Semantic extraction | Key extraction

UserID 203

Packet 3

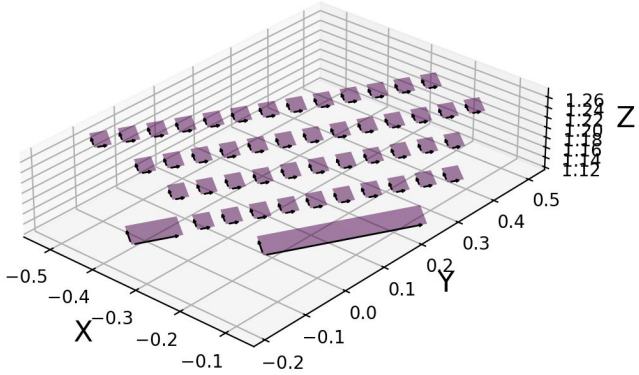
...

Packet 10033

Packet 10035

Packet 10036

Extract



Packet extraction | Field extraction | Semantic extraction | Key extraction

UserID 203

Packet 3

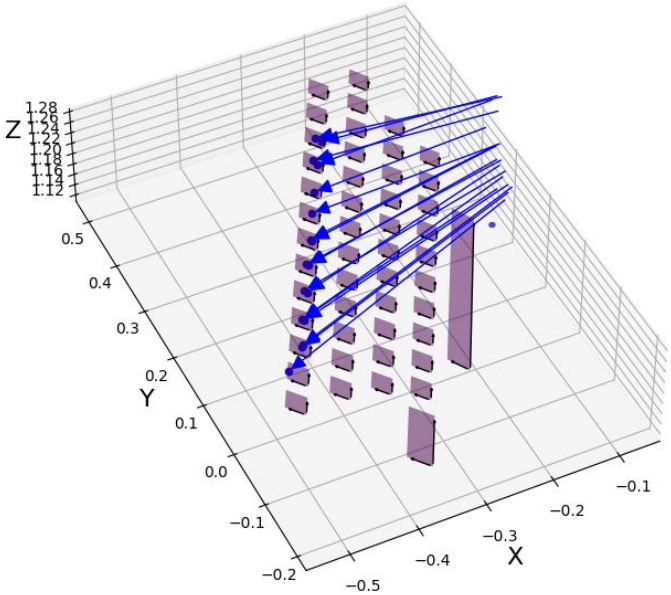
...

Packet 10033

Packet 10035

Packet 10036

Extract



Packet extraction

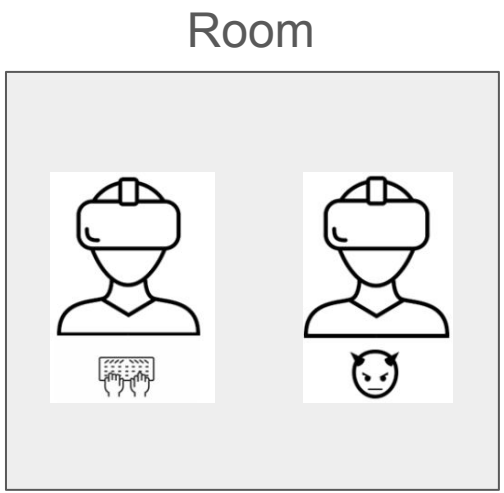
Field extraction

Semantic extraction

Key extraction

RQ 1: How Effective Is Our Attack in Inferring Keystrokes?

- 20 participants
- In Rec Room, each participant typed:
  - 30 trials on numbers
  - 20 trials on passwords
  - 15 trials on sentences



Our Attack is Highly Effective in Inferring Keys

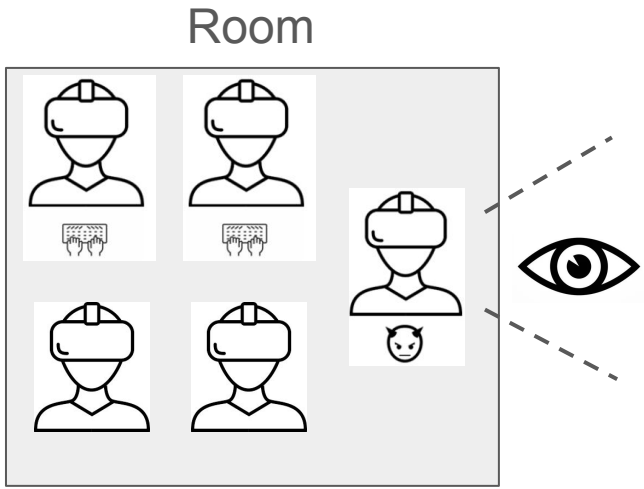
	App	Top 1	Top 3	Top 5
RQ 1	Rec Room	**97.62%**	**98.15%**	**98.34%**

RQ 2: Does the Attack Work in Practical Scenarios?

- Does the attack work when there are **multiple users** in the room? Can the attacker distinguish keys from different users?
- Does the attack work when the attacker **cannot see** the users typing?

RQ 2: Does the Attack Work in Practical Scenarios?

- 5 users in the same room
  - 1 attacker
  - 2 participants typing concurrently
  - 2 dummy players
- Attacker faces the wall



Our Attack is Practical

	App	Top 1	Top 3	Top 5
RQ 1	Rec Room	97.62%	98.15%	98.34%
RQ 2	Rec Room	**97.53%**	**99.51%**	**99.59%**

RQ 3: Is the Attack Generalizable Across Applications?

- Replicate Experiment for RQ1 on 3 additional apps

Galaxity

Sing Together: VR Karaoke

oVRshot

- 3 participants per app

Our Attack is Generalizable Across Apps

	App	Top 1	Top 3	Top 5
RQ 1	Rec Room	97.62%	98.15%	98.34%
RQ 2	Rec Room	97.53%	99.51%	99.59%
RQ 3	Galaxity	**98.25%**	**99.71%**	**99.73%**
	Sing Together: VR Karaoke	**98.27%**	**99.97%**	**99.97%**
	oVRshot	**99.07%**	**99.61%**	**99.61%**

Machine Learning Approach

With keystroke labels on partial data, using machine learning to skip manual reversing steps and recovering keystrokes is ***possible***

(Even if it is from raw bytes extracted from packets)

	Top 1	Top 3	Top 5
Random Guess	2.13%	6.38%	10.64%
SVM	44.87%	64.47%	71.57%
LightGBM	46.49%	66.24%	71.61%
MLP	61.99%	79.81%	85.34%
CNN	**68.07%**	**85.96%**	**90.28%**

Defense

Common defense mechanisms cannot solve this problem:

1. Encrypting network traffic is not enough
2. Adding noise to all movement comes with utility trade-offs

Defense

Common defense mechanisms cannot solve this problem:

1. Encrypting network traffic is not enough
2. Adding noise to all movement comes with utility trade-offs

Need better defense mechanism:

1. Full blockage of hand motion updates during sensitive typing activities
2. From both applications (e.g., Rec Room) and OS level (e.g., SteamVR)

Realistic Impact

- Attack acknowledged by **SteamVR**, **Rec Room**, **Sing Together: Karaoke**

Realistic Impact

- Attack acknowledged by SteamVR, Rec Room, Sing Together: Karaoke
- Defense implemented by **SteamVR**, **Rec Room**

Zihao,

Today's SteamVR beta includes an update that "restricts client applications from seeing controller/tracker positions while the Steam keyboard is visible." We believe that will address the issue you called out in your original email.

SteamVR Beta Updated - 2.7.1

If you encounter issues with this update, please post in the SteamVR Bug Report forum. If possible, please include a system report to aid in tracking down your issue. **Replies to this post are not tracked for bug reporting purposes. Please use the forum linked above to report issues.**

The Steam Link for Meta Quest FAQ page is available here.

Anyone can opt into the SteamVR Beta. Instructions are available here.

SteamVR:

- Restrict applications from seeing controller/tracker positions while the Steam keyboard is visible.
- Panels dragged by a grab handle now face the user and move more smoothly.

Hi jerrysu,

My apologies for the late response.

We recently shipped a change to not sync VR hands when typing into fields that are marked as sensitive.

As seen in the latest update in https://recroom.com/ship-notes:

> Hands will no longer sync if you are typing into a sensitive text field in VR (passwords, personal info, private messaging, etc.). Does not apply for insensitive text fields, though. They wouldn't really care anyways!

##General Improvements & Bug Fixes

- Fixed a bug where the Holotar scale did not match the player avatar size when scale was 1.
- Text may be a little sharper on some platforms (but shouldn't be too noticeable).
- Hands will no longer sync if you are typing into a sensitive text field in VR (passwords, personal info, private messaging, etc.). Does not apply for insensitive text fields, though. They wouldn't really care anyways!
- Restored the avatar snapshot as your default profile photo for new players.

Realistic Impact

- Attack acknowledged by SteamVR, Rec Room, Sing Together: Karaoke
- Defense implemented by SteamVR, Rec Room
- **Rec Room** rewarded us a bounty

Realistic Impact

- Attack acknowledged by SteamVR, Rec Room, Sing Together: Karaoke
- Defense implemented by SteamVR, Rec Room
- Rec Room rewarded us a bounty
- More apps may still be vulnerable today
  - 18/30 multi-user VR apps we found have typing functionalities
  - All of the 18 apps share typing motions with remote users

Realistic Impact

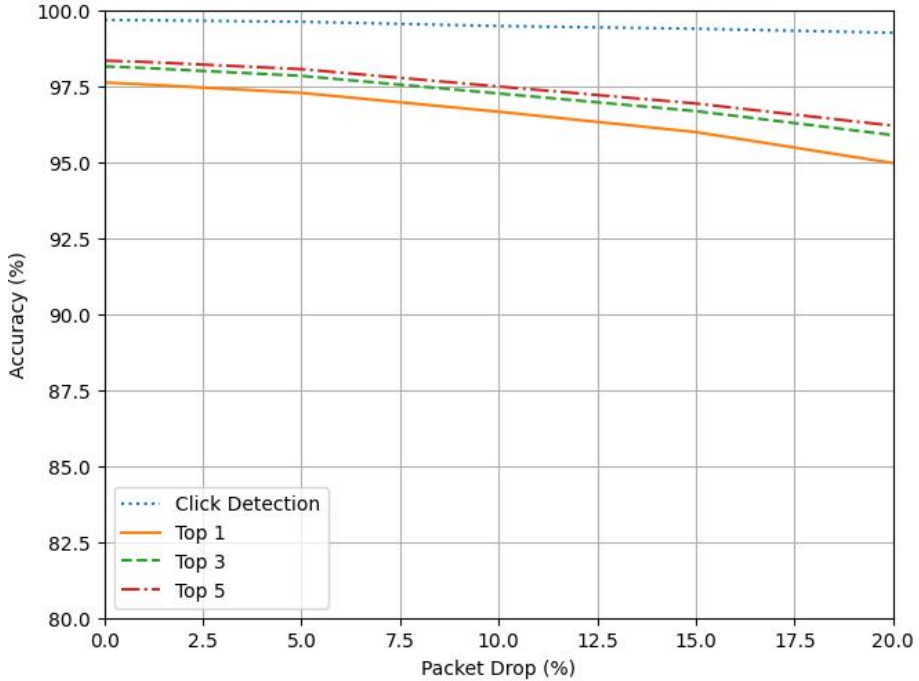
- **Attack acknowledged by SteamVR, Rec Room, Sing Together: Karaoke**
- **Defense implemented by SteamVR, Rec Room**
- **Rec Room rewarded us a bounty**
- **More apps may still be vulnerable today**
  - 18/30 multi-user VR apps we found have typing functionalities
  - All of the 18 apps share typing motions with remote users

zihaosu@ucsb.edu

SEC LAB

UCLA

Our Attack is Robust Against Packet Loss



Our attack achieves a top-1 accuracy of **94.97%** even when 20 percent of the packets are dropped