Task 1: Topic Review

Introduction: SoK: Data Privacy in Virtual Reality

As technology continues to progress towards the so-called “metaverse,” user’s personal information is at risk just from feeding VR systems a few minutes of telemetry data. Data harvesting has been happening long before VR, but defenses such as Tor, VPNs, proxies, and incognito mode exist for current internet platforms. Meanwhile, VR systems have many vulnerabilities that have not had defense systems implemented yet.

In the context of this paper, a VR threat exists when an individual's sensitive data attributes are at risk of exposure. The paper describes four types of threats: hardware adversaries, client adversaries, server adversaries, and user adversaries. Hardware adversaries control the hardware and firmware of the VR device and have access to the raw input data. Client adversaries are the developers of client-side VR applications and can manipulate user’s through outputs to the VR device and have access to input data via system APIs. Server adversaries control external servers and thus, can process received data before sending it to other devices. User adversaries are other users of the same VR application that receive data from a server and can interact with the target user.

In Task 2 we will explore an attack through a client adversary, in which we design a VR application that infers if a user is male or female based on VR sensor data recorded while the user is walking. This information could build towards a user profile that could reveal the identity of the user and be misused by attackers.

Article 1: FaceMic: Inferring Live Speech and Speaker Identity via SUbtle Facial Dynamics Captured by AR/VR Motion Sensors (Dingyi Sun)

This article is motivated by privacy risks associated with using voice interfaces while wearing AR/VR headsets, which could expose highly sensitive information, including speaker gender, identity, and speech content.

The article presents a new eavesdropping attack, called Face-Mic, which leverages speech-associated subtle facial dynamics captured by zero-permission motion sensors in AR/VR headsets to infer such information. The headset's close proximity to the user's face can capture underlying facial dynamics, including movements of facial muscles and bone-borne vibrations, that encode private biometrics and speech characteristics, and how an attacker can infer sensitive information by analyzing the captured facial dynamics.

The critical challenge is to identify relevant and useful features from all types of body movements. The article proposes a signal source separation technique to separate speech-associated facial dynamics from other types of body movements. It presents a deep learning-based framework to extract representative features with respect to two types of facial dynamics. Finally, the article validates the attack's generalizability, effectiveness, and high accuracy using four mainstream VR headsets.

The article discusses several defenses. Disabling the motion sensors is not a practical solution, as they are necessary for head movement tracking in most AR/VR apps. Instead, the article suggests using sensory noises or ductile materials to obfuscate facial vibrations and limit the sampling rate of the accelerometer and gyroscope. However, these solutions may have usability or functionality issues, and low-frequency facial muscle movements can still be used to derive sensitive information.

Article 2: You Can't Hide Behind Your Headset: User Profiling in Augmented and Virtual Reality

In this article, the researchers demonstrate the feasibility of user profiling in AR and VR settings using machine learning. The research was conducted in a recent rapid increase of usage of VR and AR while little being known how users interaction behavior and data could be used for other purposes. The authors demonstrate how interaction behavior and data could be used for user profiling; in this research, the researchers develop an attack model where sensor data from AR and VR could lead to user identification and private information inference. The attack model collected users' behavioral data when users engage with the AR/VR environment. Then, potential bias from the data was removed for data cleaning purposes. The raw temporal data was then processed for meaningful insights, which were then processed into a machine learning model. Three models were considered, Logistic Regression, Decision Tree and Random Forest. Total 45 young adults participated in the research, where each participant performed various AR/VR tasks, such as walking, looking or typing. Their head position, head rotation, eyes, controller position and controller rotation were monitored, and then these features were used to train the machine learning models to predict users' personal information. The researchers evaluated the efficacy of the attack model by measuring the F1-Score of the machine learning model in predicting users' age and gender. AR setting scored 60% while VR setting scored above 90% in predicting users' age and gender. Among three ML models, Random Forest turned out with the highest accuracy. Overall, the researchers concluded that users can be identified and profiled both in AR and VR, but VR accuracy being much higher due to the incorporation of eye tracking sensors, and thus the profiling threat is more threatening in VR settings. While the researchers did not propose a direct defense method in the paper, the researchers propose alternative ways to protect privacy, such as supporting privacy by default, incorporating incognito metaverse or privacy-preserving measures on daily usage systems such as authentication could be helpful, according to the authors.

Article 3: Behavioural Biometrics in VR: Identifying People from Body Motion and Relations in Virtual Reality

This paper demonstrates the feasibility of identifying individuals through VR motion data recorded during generic tasks performed in VR. The paper discusses this in the context of a feature rather than an attack. Identifying a user through VR motion data could be used as a form of authentication or to adapt a VR environment to a user's preferences.

The paper had 22 participants perform pointing, grabbing, walking, and typing tasks and recorded their head, hand, and eye motion data. Notably, dominant hand was considered as an identifying feature, so three left handed participants were filtered out. Random Forest and Support Vector Machine classifiers were trained to differentiate between individuals. Compared to the baseline guessing accuracy of 5.26% (1 out of 19 participants), the paper was able to achieve overall accuracies of about 40%. It was also found that the best features to identify individuals are their head motions, and the distances between devices. Interestingly, pointing had higher accuracy than grabbing hand movements because in grabbing, users have to move to the same absolute positions, while pointing allowed users more freedom in movement. Additionally, it was found that head movement can better differentiate users in the walking task than hand movements. This is surprising considering head movement is not normally a defining feature to determine whether a person is walking.

Because the paper discusses user identification as a feature rather than an attack, it does not present defense systems against identification through body motion and relations. However, the defense systems described in article 2 could also apply to defense against user identification; namely incorporating incognito metaverse or privacy-preserving measures on daily usage systems.

Task 2: Attack Implementation of

YOU CAN'T HIDE BEHIND YOUR HEADSET:
USER PROFILING IN AUGMENTED AND VIRTUAL REALITY

Part 1: Data Collection

The goal of our attack was to infer if a user is male or female based on VR motion tracking data captured while a user is walking. We asked nine males and nine females to walk in a clockwise circle within the largest VR boundary we could create. Each participant recorded three trials of walking, with each trial being ten seconds long. We controlled all the other confounding variables.

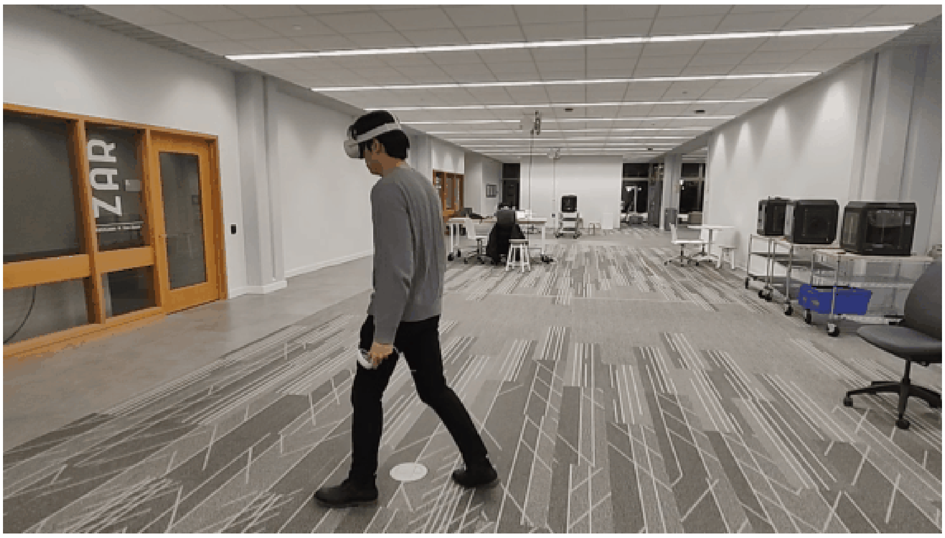


Figure 1: Example Trial

Part 2: Preprocessing and the Classifier

The raw input data format is the same as that of Lab1 and Lab2, but we have modified the data collection Unity scene for easier data collection.

The paper specifies the following raw data to take:

1. Head Pos Norm:	norm(x2 - x1)
2. Head Vertical Oscillation:	h2- h1
3. Head rot:	3D ang_speed
4. Controller Pos Norm:	norm(x2 - x1)
5. Cotroller Ang	3D ang_speed

We first did some data-cleaning by removing some data and ensuring that every data set has exactly 10 seconds of measurements. After preparing the raw data in this form in a pandas data frame, we extract four features from each of the time series of the attributes:

1. mean

2. standard deviation
3. quantile (average of top 25% data)
4. entropy

At last, we flatten the matrix represented by the data frame into 1D-array as the input format and associate it with a label specified by its filename (M for male and F for female).

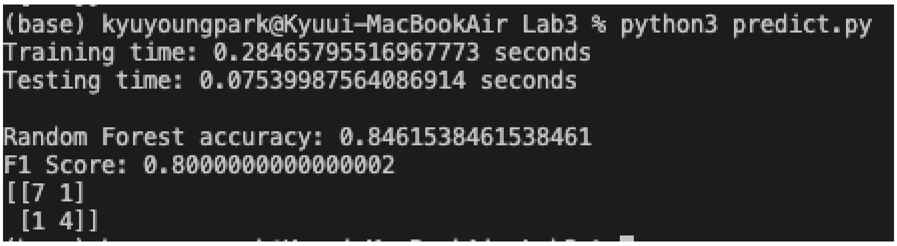
The paper suggested the random forest algorithm as the best algorithm, and likewise we have selected random forest as our AI model.

Total 54 data samples were randomly distributed into training and test sets. We set the ratio of training and test set to be 80%. The random distribution ended up assigning 5 female data sets and 8 male sets as test sets. The random forest model was trained with training sets, and we evaluated the feasibility of the attack model by measuring the accuracy and F1 score of the gender prediction of the model.

Part 3: Results

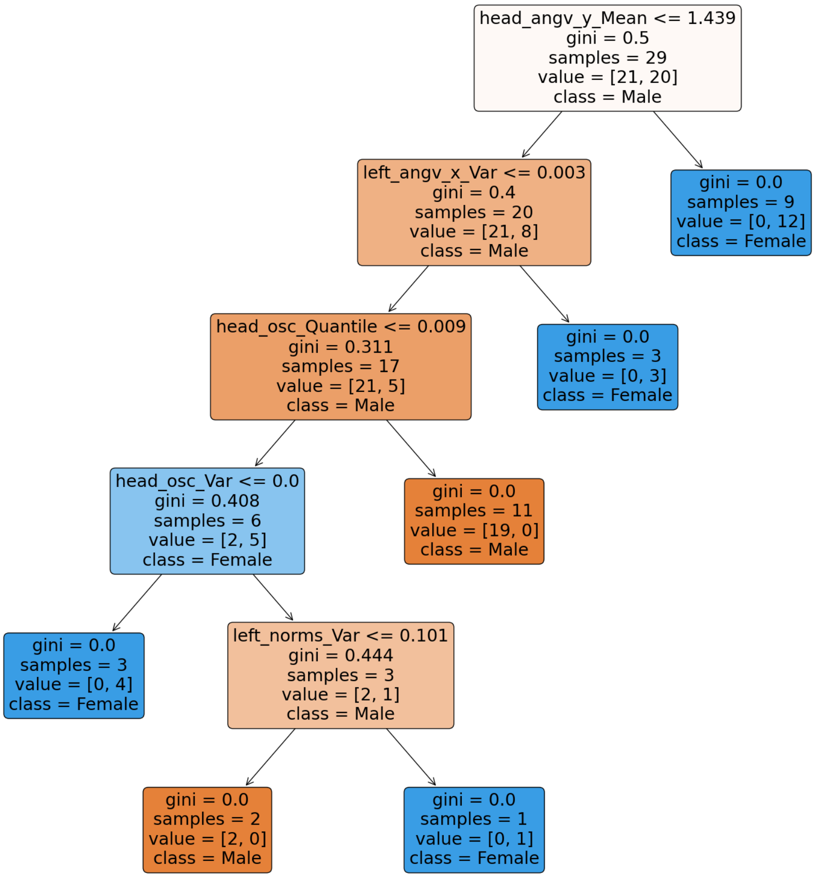
Accuracy: 85%
F1 Score: 80%
Confusion Matrix

	True Class		
Predicted Class		Male	Female
	Male	7 (TP)	1 (FP)
	Female	1 (FN)	4 (TN)



```
(base) kyuyoungpark@Kyuui-MacBookAir Lab3 % python3 predict.py
Training time: 0.28465795516967773 seconds
Testing time: 0.07539987564086914 seconds

Random Forest accuracy: 0.8461538461538461
F1 Score: 0.8000000000000002
[[7 1]
 [1 4]]
```



Above is one of the trees in the random forest model. It appears that head_angv_y_mean, left_angv_x,_var, head_osc_quantile, head_osc_var, left_norms_var were significant features that the model used to determine one's gender. However, some other trees in the forest may have used other features differently.

Just like the paper, we were able to demonstrate the feasibility of the attack model, predicting one's gender from one's VR activity. However, we also acknowledge some limitations in our implementation of the attack model, such as having low number of sample (18 people) and their lack of representativeness (all UChicago students); unlike the paper, we only used one ML algorithm of Random Forest method instead of using all four of Logistic Regression, Ridge, Decision Tree and Random Forest; unlike the paper, we only focused on one simple task of walking, and only predicted one's gender; and finally, unlike the paper, we could not leverage the eye tracking sensor (Oculus Quest 2 does not have eye tracking sensor).

Contributions:

- Kyu: Article 2 review, modify Unity scene, data collection and cleaning, some python code, identify significant data features
- Dingyi: Article 1 review, data collection, implement ML model
- Kendrick: Introduction, Article 3 review, modify Unity scene, data collection and cleaning