Discerning User Activity in Extended Reality Through Side-Channel Accelerometer Observations

Tiago Martins Andrade
Security and Cyber Defence
BT Applied Research
Adastral Park, UK
tiago.andrade@bt.com

Max Smith-Creasey
Security and Cyber Defence
BT Applied Research
Adastral Park, UK
max.smith-creasey@bt.com

Jonathan Francis Roscoe
Security and Cyber Defence
BT Applied Research
Adastral Park, UK
jonathan.roscoe@bt.com

***Abstract*—Extended reality technologies such as virtual reality are becoming increasingly common for enterprise applications. They have the potential to create secure multi-user environment in previously less-secure spaces, without the need for privacy filters or secure rooms. In this pilot paper we explore how malicious actors may be able to eavesdrop on a virtual reality session, by tracking the physical movements of a user. This is carried out using a third-party accelerometer, attached to the user. Through initial experimentation, we observe that specific actions and session types can be identified through visual analysis of the accelerometer. We posit there is substantial potential for sophisticated and automatic classification of user activity in VR. We discuss how this may enable eavesdropping by malicious actors, or could serve as a mechanism for improved security.**

***Index Terms*—Extended Reality, XR, Virtual Reality, VR, Augmented Reality, Mixed Reality, Accelerometer, Human-Computer Interaction, User Behaviour, Privacy**

I. Introduction

In the last decade many devices have been designed to enable seamless user interaction, providing access to information and experiences that were difficult to obtain before. The use of extended reality (XR) systems and applications (such as virtual reality (VR) games) is growing as technology develops. This offers users a more immersive computational experience though a headset and interacting with controllers.

However, maintaining the security of the user data remains an important issue in the field of cyber-security. It is ever more prevalent given the potential of side-channel attacks. For example, certain biometrics and sensor data can be used to identify users, behaviours, actions and contexts that may lead to exploitation for malicious actions.

In this paper, we argue that compromised external, third-party devices fitted with an integrated accelerometer (e.g. smartwatches) can collect sufficient data to identify the activity that a user is engaged in and even distinguish between individual users doing the same activity. This is an important security issue that needs to be overcome by studying and creating methods that can protect user identities and actions.

A. XR Technologies: VR, AR and MR

Extended Reality (XR) is an umbrella term that encompasses all the spectrum of technologies combining real and virtual environments such as augmented reality (AR), virtual reality (VR) and mixed reality (MR), and everything in between [1]. All technologies ranging from "the complete real" to "the complete virtual" experience are included [2].

VR makes different cognitive interactions possible in a computer-generated environment which models a 3D virtual space or virtual world. Typically, VR needs a head-mounted display (HMD) to visualise the virtual world, and which enables navigation in the environment as well as manipulation of objects. Rather than creating a completely simulated environment (as in VR) AR preserves the real environment and its surroundings, allowing the user to interact with 3D objects that are placed in the environment. MR is a combination of both VR and AR to produce new environments and visualisations where physical and digital object co-exist so that real or virtual objects can be added to virtual environments and virtual objects can be added to the real world.

B. Side Channel Attacks

Side-channels are the result of how a system is implemented, rather than a result of design. A common example is the examination of acoustic patterns produced by a system for the purpose of eavesdropping [3]. Exploitation of side-channels has been demonstrated on a variety of mechanical, electrical and electronic technology, posing a genuine threat to user safety and privacy. The literature on such attacks applied to XR systems is limited.

Identification of users and activity from a VR environment has been previously demonstrated [4], but this is traditionally from the context of an application developer and requires access to XR system hardware. No investigation has been done on detecting virtual interactions through a side-channel.

II. Experimental Setup

The test subject was required to wear an Android-based mobile device on a right-handed fore-arm apparatus. This mobile device took approximately 3,000 measurements per minute using the on-board triple-axis accelerometer resulting in vectors of $[x, y, z]$. The device is most comparable to wrist-based devices such as the Fitbit, Apple Watch or Garmin range of smartwatches. Three applications were selected:

(a) Manipulating a bottle (b) Smashing a bottle

Fig. 1. Screenshots from Half Life: Alyx, which encourages self-paced interaction with the environment, with many in-world objects offering tactile feedback. Here, the user (a) interacts with an object before (b) smashing it, an evently clearly observable in the accelerometer data in Figure 2c.

- **Beat Saber**[1] - a rhythm game where players "chop" blocks at a variety of different heights and angles. User attention is focused in an approximately 45° field-of-view. Actions may include crouching and leaning.
- **Pistol Whip**[2] - a rhythm game where players shoot targets whilst on a moving walkway. An iconic feature is the "pistol whip" maneuver in which the user must strike a non-player character in its path. Crouching and leaning actions may be performed.
- **Half Life: Alyx**[3] - a first-person shooter with an emphasis on exploration and puzzle aspects. The user has significant freedom of movement and is encouraged to interact with the environment at their own pace.

For the virtual reality session, an Oculus Quest (software version 19) was used. For Alyx, Oculus Link (an architecture for streaming input and output between the headset and a PC application) was used.

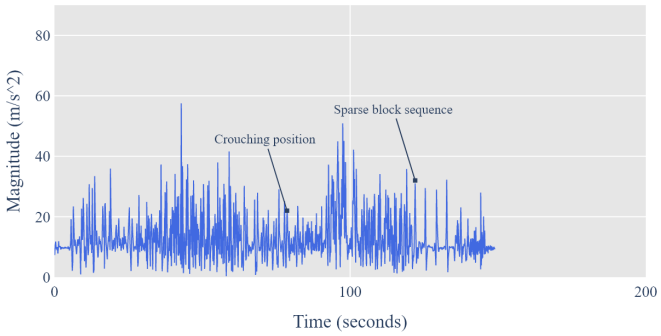
These applications were chosen for their unique modes of interaction as this will most efficiently demonstrate our hypothesis. Beat Saber requires users to strike blocks, it has a finite number of possible interactions and users are required to actively engage in a limited time window. Pistol Whip also falls into the category of a rhythm game, but users are not penalised for missing events and have much more freedom in selecting the order of interactions. Half Life Alyx provides the most freedom to the user and enables them to explore a semi-open world at their own pace. We would expect to see three distinct playstyles evidenced in the accelerometer data.

Analysis is performed using the magnitude for the three-dimensional accelerometer data provided by:

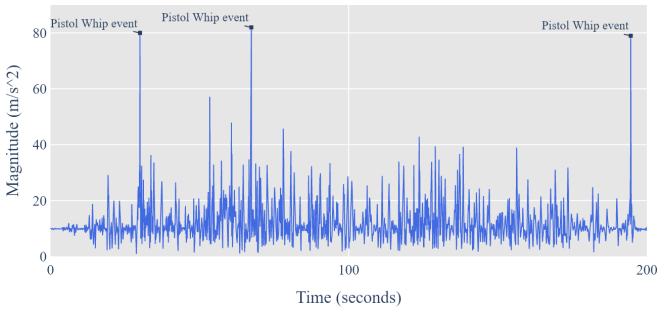
$$m = \sqrt{x^2 + y^2 + z^2}. \quad (1)$$

This provides a summary indicator of the general direction of the readings and the impact that user interaction events have on the measurements recorded by the accelerometer. Key events (such as a "pistol whip" event) in the data were mapped by replaying video recordings of the user session.

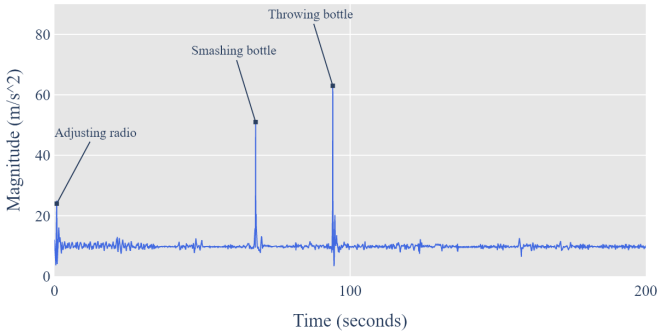
[1]Beat Games, 2019
[2]Cloudhead Games, 2019
[3]Valve, 2020



(a) Beat Saber - Beat Saber (Expert)



(b) Pistol Whip - Black Magic (Normal)



(c) Half Life: Alyx - Entanglement

Fig. 2. Magnitude of 3D accelerometer vectors over time in three virtual environments. The axes have been fixed to enable visual comparison between activities. Annotations are provided to illustrate how certain user interactions have identifiable impact on measured acceleration. Of note is (a) the series of 4 distinct sets of sparsely spaced blocks, (b) the pistol whip maneuver and (c) smashing a bottle as seen in Figure 1.

III. Activity Detection

Accelerometer sensors have been commonly used in related literature to detect user activities [5]. Given the observed differences in different activities in accelerometer readings whilst gaming, it would be feasible to perform activity detection within XR/VR environments. To detect the general activity we see a Long Short-Term Memory (LSTM) Recurrent Neural Network (RNN) as being most appropriate because it has an awareness of historical events within the inputted signals and could use these to aid classification of sections of input

data. Furthermore, specific actions within activities could be detected through methods such as Dynamic Time Warping (DTW), which would require the input of a specific section of a signal where an action occured and matching it, via DTW, to other known actions. DTW is useful here because it can be used to classify time series data where sequences may be of different lengths or contain unique events but at different times in the series.

IV. Security Implications

In this section we discuss how the different activities detected could have a variety of security implications.

A. Authentication

It is common for XR devices to be shared by multiple users for entertainment, development or commercial purposes. Hence, a system that authenticates users is needed to constrain access to sensitive applications or data. As with a user's gait movement, a user's movement during VR interaction can also be used to authenticate them [6]. Accelerometer data could also be fused with other data available from a VR system such as eye saccadic movements, button-press analysis, and physical movements on the VR controller. These could allow the creation of a multi-modal biometric profile and use it as a more secure authentication method (similar studies on mobile devices have shown the benefits of multi-modal biometric schemes [7]). Based on this, it would be feasible for movement data collected via a side-channel to be involved in user authentication.

B. Contextual Awareness

Contextual awareness is a growing field of work that can adapt the user's experience based on the detected context. For example, if a smartwatch were to detect that the user is playing a game (or performing another VR activity) that required strenuous exercise it could record this as an exercise and register calories burned. Similarly, if such activities were detected, the smartwatch could silence notifications due to the context so as not to disrupt the immersive VR experience.

Contextual awareness can also help prevent unauthorised users or insecure devices from being able to access sensitive data. This could include detecting that a user is engaged in a confidential activity and preventing outside viewing (e.g. locking the computer screen or obscuring password entry).

C. Privacy Issues

Possibly the most important security implication is the potential for privacy leakage. A user may be unaware that their smartwatch could detect the specific activity they are performing whilst interacting with a VR system. If such information was egressed to an attacker it would enable them to know the applications the user is engaging with (which may be adult/private in nature). Such information may even allow attackers to note when in the data the user will play VR and attempt robberies when the user is likely to be occupied.

V. Discussion

Figure 2 shows accelerometer data collected from three user sessions of approximately 3 minutes each.

Beat Saber (Figure 2a) has a relatively constant flow of activity, with a small combination of interactions (the user swiping at particular angles/positions). Peaks in this data are rare, but certain events are noticable, such as a series of 4 sparse pairs of blocks towards the end of the session.

In contrast, whilst Pistol Whip (Figure 2b) has a similar pattern of constant activity, it has optional "pistol whip" events. These require a significantly different range of motion from the user and are easily identifiable as peaks.

Lastly, Half Life: Alyx (Figure 2c) has a low-level of activity, due to the lack of significant movement required for interaction at this stage. Interactions with significant extension or force are very easily identifiable as they are distinct from more common activity such as ambulation and observation.

The visualisations verify that specific actions performed by the user are recognisable. As this data is the magnitude of three-dimensional vector, actions which involve the user exert themselves are shown as a peak in the graph.

VI. Future Work & Conclusions

We recognise that this pilot paper uses limited data which will need to be expanded for future work. This should include repetition of specific tasks by a cohort of users and additional sensors. There may also be other side-channels to explore, such as environmental vibration and motion capture. With a superior dataset it should be possible to explore automatic classification of data via methods discussed in Section III.

In conclusion, we demonstrated that user movement through external accelerometers is a possible side-channel for identifying activity in XR sessions. This could have implications for privacy and security and requires further investigation.

References

[1] T. Andrade and D. Bastos, "Extended Reality in IoT scenarios: Concepts, Applications and Future Trends," in *2019 5th Experiment International Conference (exp.at'19)*, 2019, pp. 107–112.

[2] T. Leland, "Extended Reality Convergence, Qualcomm." 2017, accessed July 2020. [Online]. Available: https://www.qualcomm.com/news/onq/2017/05/31/ extended-reality-convergence

[3] J. F. Roscoe and M. Smith-Creasey, "Acoustic Emanation of Haptics as a Side-Channel for Gesture-Typing Attacks," in *2020 International Conference on Cyber Security and Protection of Digital Services (Cyber Security)*, 2020, pp. 1–4.

[4] K. Pfeuffer, M. J. Geiger, S. Prange, L. Mecke, D. Buschek, and F. Alt, "Behavioural Biometrics in VR," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019, pp. 1–12.

[5] N. Ravi, N. Dandekar, P. Mysore, and M. L. Littman, "Activity recognition from accelerometer data," in *Aaai*, vol. 5, no. 2005, 2005, pp. 1541–1546.

[6] A. Kupin, B. Moeller, Y. Jiang, N. K. Banerjee, and S. Banerjee, "Task-Driven Biometric Authentication of Users in Virtual Reality (VR) Environments," in *MultiMedia Modeling*, I. Kompatsiaris, B. Huet, V. Mezaris, C. Gurrin, W.-H. Cheng, and S. Vrochidis, Eds. Cham: Springer International Publishing, 2019, pp. 55–67.

[7] M. Smith-Creasey and M. Rajarajan, "A novel scheme to address the fusion uncertainty in multi-modal continuous authentication schemes on mobile devices," in *2019 International Conference on Biometrics (ICB)*, 2019, pp. 1–8.