Convenience and Hidden Concerns Coexist: Voice Eavesdropping Based on IMU

2024040415 Yichen Dai

Abstract

This article delves into the dual aspects of convenience and privacy risks associated with the rapid development of inertial measurement units (IMUs) in smartphones. IMUs, which encompass accelerometers and gyroscopes, are extensively utilized to detect device motion and position. However, these sensors can also be exploited to steal users' voice information, posing significant privacy threats.

The article outlines the comprehensive process of IMU-based eavesdropping, from the initial stages of audio propagation and sensor response to the subsequent steps of data collection, uploading, and analysis. This process can lead to the unauthorized extraction of personal information, including sensitive data such as home addresses, financial details, and personal identifiers. The challenges involved in data preprocessing, audio recognition, and reconstruction are thoroughly discussed, along with the innovative techniques employed by attackers to circumvent sampling rate limitations.

Experimental results demonstrate that deep learning systems can achieve high accuracy in recognizing and reconstructing speech content from acceleration signals, even in noisy environments. These findings underscore the potential vulnerabilities in current sensor technology and highlight the need for robust defense mechanisms.

The article concludes with a discussion on potential defense measures, such as limiting sampling rates, generating resonance noise, and reducing sensor accuracy, to protect user voice privacy. Despite these measures, the article emphasizes the necessity for ongoing research and the development of more effective defense strategies to safeguard against the evolving techniques of eavesdropping. The continuous advancement of IMU technology necessitates a proactive approach to ensure that user privacy is not compromised while maintaining the convenience and functionality of modern smartphones.

Keywords: IMU; Voice Eavesdropping; Attack Process;

1 Introduction

The rapidly development of mobile motion sensors has indeed brought a lot of convenience to people's lives. For example, by sensing the position and status changes of mobile phones, sensors record all information of mobile phone users, enhancing the user experience. However, behind these conveniences are also hidden security risks, especially privacy breaches.

Mobile phone accelerometers can collect voice information, which means attackers can steal various private data from users' phones. The attacker may extract the user's home address, credit card information, ID number, user name and password, so far, destination, hobbies and other important information from the voice message. The inertial measurement unit (IMU) in smartphones can transform into a "bug", stealing user data and endangering user privacy.

IMU is typically composed of accelerometers and gyroscopes, widely used in smartphones and other devices to detect device motion and position. In order to protect user privacy, Google has limited the frequency of sampling data from IMU for Android applications to 200 times per second. However, research has discovered a vulnerability called STAG, which can bypass these protection measures and increase the actual sampling rate of applications from 200 times per second to 400 times per second, breaking through the aforementioned protection measures.

The malicious program created by the attacker disguises itself as a regular program and is downloaded by the victim. The application then collects real-time data from the victim's IMU without attracting the user's attention and uploads it to the backend. Finally, the attacker analyzes and identifies this IMU information to obtain the user's privacy.

2 Background and Related Work

With the popularity of smartphones, built-in inertial measurement units (IMUs) such as accelerometers and gyroscopes have been used for eavesdropping on audio due to their sensitivity to vibration. Although these sensors are generally considered low-risk, research [1, 3, 5, 4, ?, 6, 7] has shown that they can be used for so-called ”zero authorization” attacks, which involve accessing built-in accelerometers without user permission or attention, thereby stealing voice privacy.

Research has shown [1] that motion sensors in smartphones can capture speech signals that propagate through solid media, thereby revealing speech privacy. The study also found that airborne vibrations of speech do not affect the motion sensors of mobile phones. The Spearphone [3] attack method utilizes the speech reverberation generated by the built-in speaker of a smartphone, captured by an accelerometer, and successfully performs gender classification (with an accuracy rate of over 90%) and speaker recognition (with an accuracy rate of over 80%). The InertiaEAR [5] attack method utilizes the side channel from the speaker to the IMU to eavesdrop on audio from the top and bottom speakers on the smartphone. This method achieves high recognition accuracy (78.8%) by handling the consistency between accelerometer and gyroscope readings and supports cross-device attacks. Watch the Rhythm [7] shows that by analyzing the accelerometer data from mobile phone motion sensors and combining their rhythm and time-frequency features, users' voice privacy can also be compromised at an extremely low sampling rate of 5Hz.

3 Attack Model

3.1 Eavesdropping Process

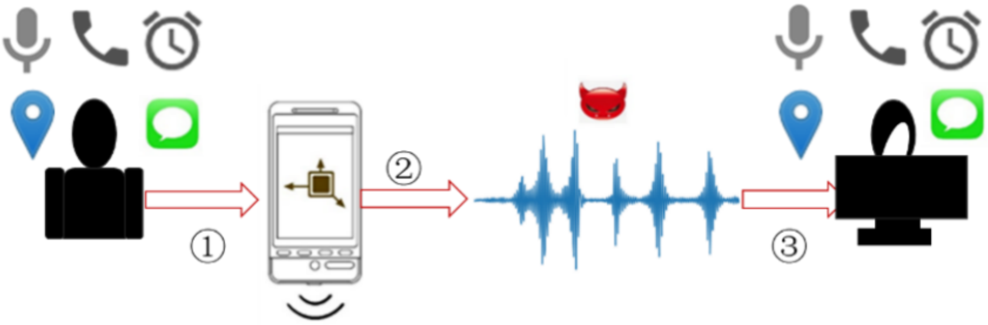

Figure 1: IMU Eavesdropping Process

IMU eavesdropping technology is a method of stealing voice information through the built-in inertial measurement unit (IMU) of smartphones. The IMU attack process is illustrated in Figure 1, and the following is a detailed description and improvement of this process:

3.1.1 Audio Propagation and Sensor Response

When the victim makes a call using a smartphone, whether through the speaker or earpiece, the phone outputs an audio signal. These audio signals propagate in the form of sound waves and affect the IMU on the same motherboard, particularly the accelerometer and gyroscope, causing corresponding vibrations in these sensors.

3.1.2 Data Collection and Uploading

Malicious software installed on the victim's smartphone can monitor and collect data from the accelerometer and gyroscope in real-time. This data is then secretly uploaded to cloud servers, providing attackers with a basis for analysis.

3.1.3 Data Analysis and Identification

Attackers can identify and reconstruct audio signals by analyzing data from the accelerometer and gyroscope, in order to obtain the victim's personal information.

3.2 Attack Method

IMU eavesdropping technology leverages the built-in inertial measurement unit (IMU) of smartphones to steal voice information. The IMU typically includes an accelerometer and a gyroscope, which are designed to detect motion and orientation. However, these sensors can also capture subtle vibrations caused by sound waves, making them potential targets for eavesdropping attacks. The attack process involves several steps, each designed to covertly collect and analyze data to reconstruct audio signals and extract sensitive information.

3.2.1 Data Preprocessing

The attacker first preprocesses the uploaded data to enhance its quality and usability. This step includes noise and interference removal, as well as data interpolation. The noise can be intrinsic or environmental, and the primary interference is motion interference, given that motion sensors are highly sensitive to movement.

3.2.2 Audio Recognition and Reconstruction

The preprocessed data is then used for audio recognition and reconstruction. By analyzing the time-frequency characteristics or other features of the accelerometer and gyroscope data, attackers can reconstruct speech signals and obtain the victim's private information.

3.2.3 Challenges and Responses to Sampling Rate Limitations

According to Nyquist's theorem, the original signal can only be fully restored if the sampling frequency is greater than twice the signal frequency; otherwise, aliasing may occur, leading to signal distortion. To protect user privacy, mobile phone manufacturers have reduced the sampling frequency of motion sensors. Despite this, attackers can still extract key information by combining data from accelerometers and gyroscopes or by analyzing the time-frequency and rhythm characteristics of accelerometer data. This approach allows attackers to identify critical information even at lower sampling rates.

3.3 Experimental Results and Effects

The AccelVe attack method uses deep learning systems to recognize and restore speech content from acceleration signals, achieving a recognition accuracy of 78% even in high noise environments. Spearphone experiments have shown that even at lower volume levels, speech leakage from high-quality speakers can be more destructive, leading to serious privacy breaches. Watch the Rhythm demonstrates that the recognition rate of keywords at an extremely low frequency of 5Hz is as high as 78.8%, and it has good robustness even in different scenarios, states, volumes, and voice assistants.

4 Defense Measures and Countermeasures

4.1 Sampling Rate Limitation and Secure Filtering

A common defense strategy against IMU-based eavesdropping is to limit the sampling rate of inertial measurement units (IMUs) to prevent the overlap between the inertial measurement range and the sound baseband. However, recent research has shown that even under these restricted conditions, IMUs can still capture and reconstruct private voice data. For instance, the "Watch the Rhythm" study [7] demonstrated that it can breach user privacy even at an extremely low sampling rate of 5Hz. Notably, 5Hz is the maximum allowable offline sampling rate for mobile phones, primarily used to detect whether the phone screen has been switched. Reducing the sampling rate below 5Hz would negatively impact user experience, potentially leading users to disable the motion sensor altogether.

4.2 Resonance Noise

As a non-hardware modified solution, it is recommended that users actively utilize onboard speakers to generate resonance noise while speaking. This approach aims to interfere with the IMUs, disrupt coherent signal segmentation, and confuse the recognition algorithms. By introducing controlled noise, the integrity of the captured signals is compromised, making it difficult for eavesdropping systems to accurately reconstruct the spoken words.

4.3 Reduce Sampling Accuracy

Another effective measure is to reduce the sampling accuracy of mobile phones. The "Watch the Rhythm" study [7] suggests that Android systems or phone manufacturers should impose limitations on sensor accuracy, similar to the restrictions on sampling rates. Without explicitly declaring high-precision permissions, the system should automatically switch to a low-precision mode to mitigate the risk of user privacy leakage. This approach ensures that even if an application attempts to access the IMU data, the reduced accuracy limits the potential for precise signal reconstruction.

4.4 Additional Considerations

While these measures provide a robust defense against IMU-based eavesdropping, they must be implemented in conjunction with other security practices. Users should also be educated about the risks associated with IMU-based attacks and encouraged to adopt additional protective measures, such as using secure communication channels and regularly updating their devices' software to patch known vulnerabilities. Furthermore, regulatory bodies and industry standards organizations should collaborate to establish guidelines and best practices for securing IMU data, ensuring that user privacy remains a top priority.

5 Conclusion

The advent of advanced mobile motion sensors, particularly within inertial measurement units (IMUs), has undeniably introduced a new era of convenience, enhancing user experience through the precise detection of mobile phone movements. However, this technological boon has an underbelly, revealing significant vulnerabilities in the form of privacy breaches. The ability of mobile phone accelerometers to inadvertently capture voice information presents a critical risk, potentially exposing a trove of sensitive user data to malicious entities.

IMUs, comprising accelerometers and gyroscopes, are pivotal in the functioning of smartphones and other devices, detecting motion and orientation. Despite measures like Google's restriction on IMU data sampling to 200 times per second for Android applications, vulnerabilities such as STAG have been discovered. This allows for an increase in sampling rate to 400 times per second, effectively circumventing privacy protections.

The threat is further exacerbated by the development of malicious software that can covertly collect IMU data in real-time, uploading it to remote servers for analysis. This presents a stark contrast to the intended use of IMUs, transforming them into tools for eavesdropping. The Spearphone and InertiaEAR attacks, along with the Watch the Rhythm method, demonstrate the alarming potential of IMUs to compromise voice privacy, even at low sampling rates.

The eavesdropping process is a multifaceted one, involving the propagation of audio signals that influence the IMU, leading to detectable vibrations. Malicious software can then collect and preprocess this data, using sophisticated methods to recognize and reconstruct audio signals, thereby extracting personal information.

Experimental results have shown that attacks like AccelVe can achieve high recognition accuracy even in noisy environments, highlighting the robustness of these methods. The ability to recognize speech at low sampling rates, as demonstrated by Watch the Rhythm, underscores the persistent risk to privacy, regardless of efforts to limit data sampling.

In response to these threats, a variety of defense measures have been proposed. Limiting the sampling rate and implementing secure filtering are common strategies, but they have been shown to be insufficient against determined attackers. Generating resonance noise and reducing the sampling accuracy of mobile phones offer alternative approaches, but they too have limitations and may impact user experience.

The conclusion drawn from this discourse is that while IMUs have significantly contributed to the convenience of mobile technology, they also pose a real and present danger to user privacy. The ongoing development of more sophisticated attack methods necessitates a reevaluation of current defense strategies. There is an urgent need for innovative solutions that can effectively safeguard against IMU-based eavesdropping, ensuring that the benefits of mobile motion sensors do not come at the expense of user privacy. This calls for a collaborative effort between technology developers, security researchers, and policymakers to create a secure environment where convenience and privacy can coexist harmoniously.

References

[1] Anand, S. A., & Saxena, N. (2018). Speechless: Analyzing the threat to speech privacy from smartphone motion sensors. In *2018 IEEE Symposium on Security and Privacy (SP)* (pp. 1000-1017). IEEE.

[2] (2024). SensorManager. `https://developer.android.google.cn/reference/android/hardware/SensorManager`.

[3] Anand, S. A., Wang, C., Liu, J., Saxena, N., & Chen, Y. (2021). Spear phone: a lightweight speech privacy exploit via accelerometer-sense dreverberations from smartphone loudspeakers. In *Proceedings of the 14th ACM Conference on Security and Privacy in Wireless and Mobile Networks* (pp. 288-299).

[4] Ba, Z., Zheng, T., Zhang, X., Qin, Z., Li, B., Liu, X., & Ren, K. (2020). Learning-based Practical Smartphone Eavesdropping with Built-in Accelerometer. In *Network and Distributed System Security Symposium (NDSS)* (Vol. 2020, pp. 1-18).

[5] Gao, M., Liu, Y., Chen, Y., Li, Y., Ba, Z., Xu, X., Han, J., & Ren, K. (2022). Device-independent smartphone eavesdropping jointly using accelerometer and gyroscope. *IEEE Transactions on Dependable and Secure Computing*, 20(4), 3144-3157.

[6] Han, J., Chung, A. J., & Tague, P. (2017). Pitchln: eavesdropping via intelligible speech reconstruction using non-acoustic sensor fusion. In *Proceedings of the 16th ACM/IEEE International Conference on Information Processing in Sensor Networks* (pp. 181-192).

[7] Yao, Q., Liu, Y., Sun, X., Dong, X., Ji, X., & Ma, J. (2024). Watch the Rhythm: Breaking Privacy with Accelerometer at the Extremely-Low sampling Rate of 5Hz. *CCS '24*, October 14-18, 2024.