Privacy Assurance for Android Augmented Reality Apps

Xueling Zhang, Rocky Slavin, Xiaoyin Wang, Jianwei Niu
Department of Computer Science, University of Texas at San Antonio
{*xueling.zhang, rocky.slavin, xiaoyin.wang, jianwei.niu*}*@utsa.edu*

***Abstract*—Augmented Reality (AR) is an emerging technique that enriches real environment with virtual information objects. Despite its wide application scenarios, AR techniques also raise concerns on its dependability, especially on the privacy protection of the users and of the people appearing in users' eyesight. In our research, we performed a case study on the mostly popular augmented reality Android app: Google Translate. In this paper, we report our major findings in the case study, and propose potential mechanism to detect unnecessary privacy leaks in Android augmented reality apps.**

1. Introduction

Augmented Reality (AR) is an interactive system that allows users to see the real-world environment where the objects inside are enriched by computer-generated perceptual information. AR is often achieved by a set of hardware and software system, and some well known examples are Microsoft's HoloLens, Google Glass, Sony SmartEyeglass, etc. Although they are still not very popular among normal users nowadays, mainly due to their high price, they are expected to experience a boost in near future [1], [2]. Meanwhile, smartphone AR apps also get more and more popular as they can be directed downloaded from app stores and do not need any extra hardware. Some popular AR apps include Pokemon GO[1], Jurassic World[2], and Google Translate[3], which have achieved high downloads from App stores.

AR apps have raised various new challenges to software dependability assurance, such as the erroneous identification and occlusion of real-world objects. One important new challenge is the assurance of user privacy [3] [4]. Since AR apps often need access to the smartphone's camera all the time, they have access to a much larger interface of the users' privacy. If an app is collecting more than necessary information from the users' camera, or the data collection is not mentioned in the app's privacy policy, users' privacy is undermined.

In our research, we focus on the detection and avoidance of unnecessary privacy data leaks in Android AR apps. In this paper, we first introduce the state-of-the-art works in the area. Then, we report the results of a case study on Google Translation, the most downloaded AR apps in Google Play store. Finally, we propose an general framework to reduce unnecessary privacy leak.

1. https://www.pokemongo.com/en-us/
2. https://play.google.com/store/apps/details?id=com.ludia.jurassicworld
3. https://apps.apple.com/us/app/google-translate/id414706506

2. Background

Security and privacy in AR applications. Augmented Reality(AR) has gained increasing attention from public recent years, the computer security community has recently identified the need to address security and privacy for AR systems. [1] presented the security and privacy challenges for AT technologies, they categorized the challenges related to output, input and data access in Single Application, Multiple Applications and Multiple Systems. [5]conducted a study to find out user's concerns about interact with immersive AR device in single- and multi-user scenarios. [6] investigated the security and privacy vulnerabilities on the AR browsers. They identified multiple architectural flaws in the popular AR browsers and proposed short-term fixes for specific vulnerabilities.

Input and Output Control. To address the problem of AR application having access of sensitive information beyond that application need, [7] introduced a new OS abstraction: the recognizer. Instead of exposing raw sensor data to applications directly, a recognizer takes raw sensor data as input and exposes higher-level objects, such as a skeleton or a face, to applications. To prevent user receiving undesirable content from buggy or malicious application, [8] proposed "Arya" to constrain the output by modifying application behavior according to policies or guidelines from the AR application. SurroundWeb [9], a 3D web browser that allows the AR web applications to project web content onto a room while tackling privacy challenges by considering both AR input and output.

3. A Case Study

As a preliminary work of our project, we first performed a case study on the Google Translate app. Google Translate is an Android app that uses augmented reality technique to help users to read road signs and other texts in the real world. For example, a user may have her smartphone's camera targeting at a white board or exam paper, then all the text will be translated to the preset destination language. Figure 1 shows the screen shots of a white board and AR scene with translated words (i.e., Chinese). Inside the app, when the app is started and the camera is on, a text identification module first identify texts from the video content based on image processing algorithms. Then, the identified texts are sent to the remote server (i.e., the Google Translate Service), and the translated texts in the destination language are returned. Finally, the returned texts are added to the camera scene.

2473-3105/19/$31.00 ©2019 IEEE
DOI 10.1109/PRDC47002.2019.00037

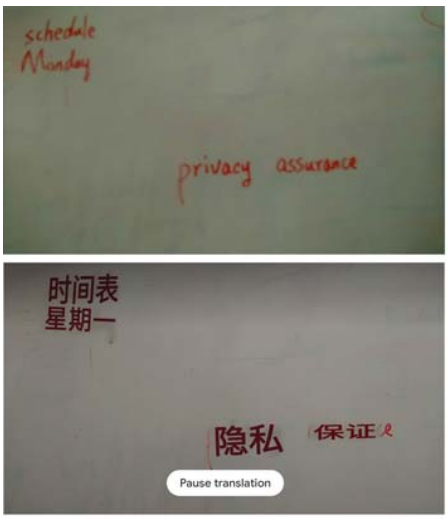

Figure 1. Camera input and VR scene of Google Translate

Looking into Figure 1, we can see that not only the words at the center ('privacy" and "assurance") are translated, the words at the top left corner ("schedule" and "Monday") are also translated. Therefore, some text that the user did not intend to translate (e.g., "schedule" and "Monday" in this example) may also be sent to the remote server and user privacy will be undermined. From our study, we mainly have two findings.

- There are no existing mechanisms in Android system checking whether information collected by AR apps are necessary or not (e.g., in Google Translate, whether only extracted texts are collected and sent to remote servers).
- AR apps may extract all information from the camera and thus cause unintentional and accidental data extraction and collection.

Although our study is just on the Google Translate app, considering that Google AR Core [4] is the most popular mobile AR framework, our study results can be largely applied to many other Android AR apps.

4. Proposed Framework

In our research, we propose a framework to enhance users' control on their privacy data, and solve the two major issues found in our case study. In particular, our framework will (1) detect whether sending certain types of information to remote servers is necessary, and (2) allow users to set an intention area on the screen and only data extracted from the intention area can be sent to remote servers. Figure 2 illustrate the overview of our framework. To achieve the first goal, we will check whether the responses of corresponding network requests are used in app's output to users. For example, in Google Translate, since the extracted texts are sent to remote servers through network requests and their responses (translated texts) are used in the app's output to

4. https://developers.google.com/ar

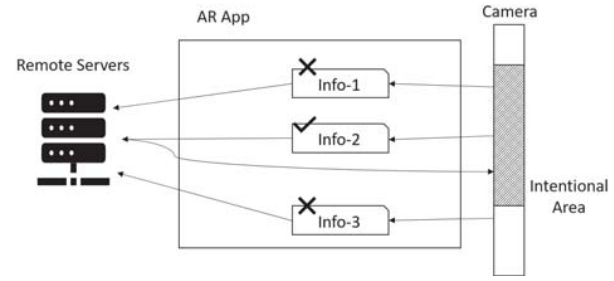

Figure 2. Overview of proposed framework

the user (AR scene), these network requests are considered necessary. In contrast, if other information extracted from camera (e.g., pictures or color information) are sent to network, they will be considered unnecessary as their responses will not be used in the AR output. To achieve the second goal, we will use taint analysis [10] [11] to track all the information extracted from the camera with coordinates describing where on the screen they are extracted. Then our framework will forbid information not from intentional area to be sent to the network.

References

[1] F. Roesner, T. Kohno, and D. Molnar, "Security and privacy for augmented reality systems." *Commun. ACM*, vol. 57, no. 4, pp. 88–96, 2014.

[2] I. Rodriguez and X. Wang, "An empirical study of open source virtual reality software projects," in *Proc. ESEM*. IEEE, 2017, pp. 474–475.

[3] X. Wang, X. Qin, M. B. Hosseini, R. Slavin, T. D. Breaux, and J. Niu, "Guileak: Tracing privacy policy claims on user input data for android applications," in *Proc. ICSE*. ACM, 2018, pp. 37–47.

[4] X. Xiao, X. Wang, Z. Cao, H. Wang, and P. Gao, "Iconintent: automatic identification of sensitive ui widgets based on icon classification for android apps," in *Proc. ICSE*. IEEE Press, 2019, pp. 257–268.

[5] K. Lebeck, K. Ruth, T. Kohno, and F. Roesner, "Towards security and privacy for multi-user augmented reality: Foundations with end users," in *Proc. IEEE S&P*. IEEE, 2018, pp. 392–408.

[6] R. McPherson, S. Jana, and V. Shmatikov, "No escape from reality: Security and privacy of augmented reality browsers," in *Proc. WWW*. International World Wide Web Conferences Steering Committee, 2015, pp. 743–753.

[7] S. Jana, D. Molnar, A. Moshchuk, A. Dunn, B. Livshits, H. J. Wang, and E. Ofek, "Enabling fine-grained permissions for augmented reality applications with recognizers," in *Proc. USENIX Security Symposium*, 2013, pp. 415–430.

[8] K. Lebeck, K. Ruth, T. Kohno, and F. Roesner, "Securing augmented reality output," in *Proc. IEEE S&P*. IEEE, 2017, pp. 320–337.

[9] J. Vilk, D. Molnar, B. Livshits, E. Ofek, C. Rossbach, A. Moshchuk, H. J. Wang, and R. Gal, "Surroundweb: Mitigating privacy concerns in a 3d web browser," in *Proc. IEEE S&P*. IEEE, 2015, pp. 431–446.

[10] X. Wang, L. Zhang, T. Xie, H. Mei, and J. Sun, "Transtrl: An automatic need-to-translate string locator for software internationalization," in *Proc. ICSE*. IEEE Computer Society, 2009, pp. 555–558.

[11] ——, "Locating need-to-translate constant strings for software internationalization," in *Proc. ICSE*. IEEE, 2009, pp. 353–363.