OVRseen: Auditing Network Traffic and Privacy Policies in Oculus VR

Rahmadi Trimananda
University of California, Irvine
rtrimana@uci.edu

Athina Markopoulou
University of California, Irvine
athina@uci.edu

ABSTRACT

We study Oculus VR (OVR), the leading platform in the VR space, and we provide the first comprehensive analysis of personal data exposed by OVR apps and the platform itself.[1] We developed OVRseen, a methodology and system for collecting, analyzing, and comparing network traffic and privacy policies on OVR, as well as for checking the consistency between the two. On the networking side, we captured and decrypted network traffic of VR apps, which was previously not possible on OVR, and we extracted data flows, each defined as ⟨*app/platform, data type, destination*⟩. Compared to the mobile and other app ecosystems, we found OVR to be more centralized, and driven by tracking and analytics, rather than by third-party advertising. We show that the data types exposed by VR apps include personally identifiable information (PII), device information that can be used for fingerprinting, and VR-specific data types. Next, by comparing the actual data flows found in the network traffic with the statements made in the apps' privacy policies, we discovered that approximately 70% of OVR data flows were not properly disclosed. Furthermore, we extracted the purpose of data collection from the privacy policies and found that 69% of data flows were sent for purposes unrelated to core functionality; further, we also found apps that do not provide notice or obtain consent. The additional information about ⟨*consistency, purpose, consent*⟩ relates to the transmission principle in the CI framework, and can extend the original ⟨*app/platform, data type, destination*⟩ tuple, so as to better characterize the appropriateness of these data flows. We believe that our approach and analysis generalizes to other platforms beyond VR, where the CI tuple can be used as the basic conceptual building block for auditing data collection practices using a combination of network traffic and privacy policy analysis.

ACM Reference Format:
Rahmadi Trimananda and Athina Markopoulou. . OVRseen: Auditing Network Traffic and Privacy Policies in Oculus VR. In . ACM, New York, NY, USA, 3 pages.

1 OVERVIEW

Among VR platforms, Oculus VR (OVR) is a pioneering, and arguably the most popular one [6, 9]. VR technology enables a number of applications, including recreational games, physical training, health therapy, and many others [20]. Similarly to the other Internet-based platforms (*e.g.*, mobile phones [4, 5], IoT devices [1, 7], and smart TVs [14, 28]), the Virtual Reality (VR) platform introduces privacy risks and some of these risks are unique to VR devices. For example, VR headsets and hand controllers are equipped with sensors that may collect data about such as user's physical movement, body characteristics and activity, voice activity, hand tracking, eye tracking, facial expressions, and play area [11, 13, 15], which may in turn reveal information about our physique, emotions, and home.

In our USENIX Security 2022 paper [26], we present the first large scale, comprehensive measurement and characterization of privacy aspects of OVR apps and platform, using a combined analysis of (1) the network traffic generated by Oculus VR apps and platform, (2) the corresponding privacy policies, and (3) the consistency between network traffic data flows and policy statements. An overview of our approach is depicted in Fig. 1. We characterize and compare how sensitive information is collected and shared in the VR ecosystem, in "theory" (*i.e.,* as stated in the apps' and platform's privacy policies) as well as in practice (as exhibited in the actual network traffic generated by the OVR apps and platform).

The Contextual Integrity (CI) tuple is at the core of our analysis. We were able to extract "data flows", defined as ⟨*app/platform, data type, destination*⟩ from the network packets sent by the OVR apps and platform. We were also able to extract the same information from the statements in the corresponding privacy policies of the apps and platform. Next, we were able to check the consistency between the two, *i.e.,* whether the actual data flows extracted from network traffic agree with the corresponding statements made in the privacy policy. Furthermore, we were also able to extract the *purpose* of data collection from the privacy policies; independently, we also extracted the purpose from the network traffic by checking the destination domains, *i.e.,* whether they are advertising-and-tracking services (ATS) domains based on well-known blocklists. Finally, we found that not all of the apps provided a privacy policy, and even when they did, the apps themselves did not always implement "notice and consent". The additional information about ⟨*consistency, purpose, consent*⟩ relates to the transmission principle in the CI framework, and can extend the commonly studied ⟨*app/platform, data type, destination*⟩ tuple, so as to better characterize the appropriateness of these data flows. The "subject" is typically the user of the device or the app, and can be implicit (denoted as "-") in the CI tuple.

We believe that our approach and analysis generalize to other platforms beyond VR [26] and their apps, including website, mobile, smart TV, IoT, *etc.*. The CI tuple, ⟨*app/platform, data type, destination,* - , ⟨*consistency, purpose, consent*⟩⟩, can be used as the basic data structure for auditing data collection practices using a combination of network traffic and privacy policy analysis.

2 OVRSEEN

Our methodology and system, OVRseen, is depicted on Fig. 1. It consists of two parts: network traffic and privacy policy analyses.

Network Traffic Analysis. We experimented with 150 popular, paid and free, OVR apps and we used the best known practices to explore them to generate rich network traffic. We use OVRseen to decrypt and collect network traffic generated by these apps. We then extracted data flows from the collected network packets.

[1]This work is supported by NSF Awards 1815666 and 1956393.

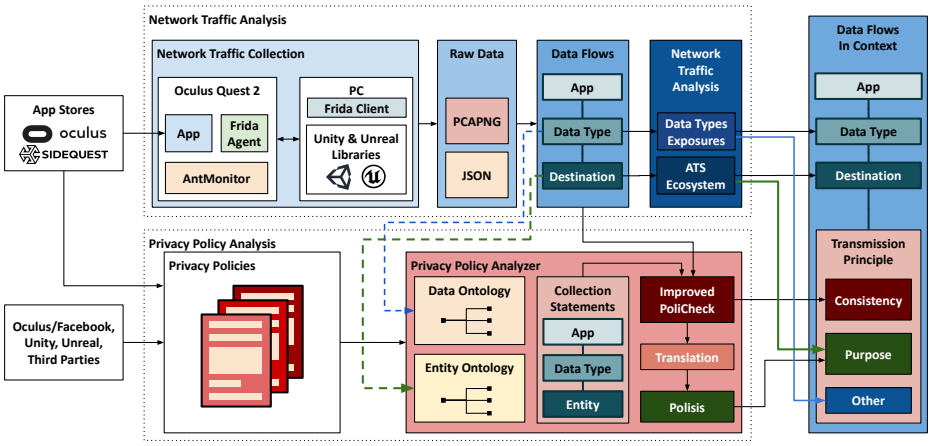

Figure 1: Overview of OVRseen. On one hand, we experiment with the most popular OVR apps and analyze their network traffic: we extract data flows ⟨*app/platform, data type, destination*⟩ and analyze them *w.r.t.* data types and ATS ecosystem. On the other hand, we also analyze the apps' privacy policies: network-to-policy consistency analysis for each app using PoliCheck [3] and collection purpose extraction from the app's privacy policy using Polisis [8]. The end result is that data flows, extracted from network traffic, are augmented with additional attributes related to the transmission principle, such as consistency with the statement made in the privacy policy and purpose of collection.

First, the sender of information is the VR *app* or platform. Second, we found 21 *data types* including personally identifiable information (PII such as device ID, user ID, android ID, *etc.*), device information that can be used for fingerprinting, and VR sensor data (*e.g.*, physical movement, play area). Third, w.r.t. the recipient of the information, we extracted the *destination* domain and we further categorize it into entity or organization, first *vs.* third party *w.r.t.* the sending app, and ATS. We find that OVR exposes data primarily to tracking and analytics services, but not to advertising services (yet!); and that existing blocklists block only 36% of these exposures.

Privacy Policy Analysis. We use an NLP-based methodology for analyzing the privacy policies that accompany VR apps. OVRseen maps each data flow found in the network traffic to its corresponding data collection statement found in the text of the privacy policy, and checks the *consistency* of the two. Furthermore, it extracts the *purpose* of data flows from the privacy policy, as well as from the ATS analysis of destination domains. Consistency, purpose, and additional information (*e.g.*, about notice and consent) relate to the *transmission principle*, which can help assess the appropriateness of the information flow [16]. Our methodology builds on, combines, and improves state-of-the-art tools originally developed for mobile apps: PolicyLint [2], PoliCheck [3], and Polisis [8]. We curated VR-specific ontologies for data types and entities, guided by both the network traffic and privacy policies. We also interfaced the different NLP models of PoliCheck and Polisis to extract the purpose behind each data flow.

Our network-to-policy consistency analysis reveals that about 70% of data flows from VR apps were not disclosed or consistent with their privacy policies. Furthermore, 38 apps did not have privacy policies, including apps from the official Oculus app store. Many app developers also tend to neglect declaring data collected by the platform and third parties. We also found that 69% of data flows have purposes unrelated to the core functionality, and only a handful of apps are explicit about notice and consent (see Fig. 2).

Due to lack of space, we defer the details of our methodology and results to [26], and the software and datasets to [27].

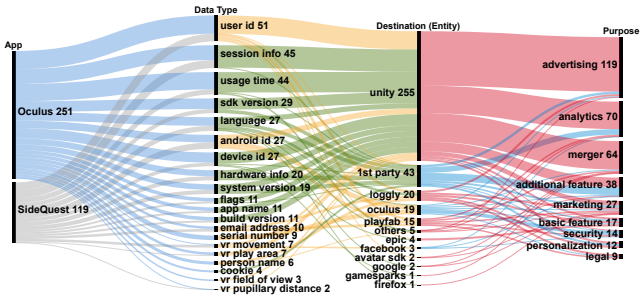

Figure 2: We consider the 370 data flows with consistent disclosures between network traffic and privacy policies. We also extract their *purpose* from the privacy policy and depict the augmented tuples: ⟨*app, data type, destination, purpose*⟩.

3 WORKS CITED

Privacy of Various Platforms. The research community has looked into privacy risks in various platforms, including Android [4, 5, 17, 19, 21–23], smart TVs [14, 28], and IoT s [12, 18]. Our work is the first to perform network traffic analysis on VR.

Privacy Policy Analysis. Privacy policy and consistency analysis in various app ecosystems [2, 3, 8, 25, 29–31] is becoming increasingly automated. Wang *et al.* applied similar techniques to check for privacy leaks from user-entered data in GUI [29]. We leveraged two state-of-the-art tools, namely PoliCheck [3] and Polisis [8], to perform data-flow-to-policy consistency analysis and extract data collection purposes—as detailed in the previous section.

CI applications. Most prior work on network traffic analysis characterized only destinations (first *vs.* third parties, ATS, *etc.*) and/or data types exposed without additional context for the CI-tuple. One exception is MobiPurpose [10], which inferred data collection purposes of mobile (not VR) apps, using network traffic and app features (*e.g.*, URL paths, app metadata, domain name, *etc.*). The authors stated that their "purpose interpretation can be subjective and ambiguous", while we extracted purpose explicitly stated in the

privacy policies and/or indicated by the destination domain matching ATS blocklists. Shvartzshnaider *et al.* applied the CI framework to analyze privacy policies (not network traffic), with a case study of Facebook, and conducted manual inspection (not NLP) [24].

REFERENCES

[1] Omar Alrawi, Chaz Lever, Manos Antonakakis, and Fabian Monrose. Sok: Security evaluation of home-based iot deployments. In *2019 IEEE Symposium on Security and Privacy (SP)*, pages 1362–1380. IEEE, 2019.

[2] Benjamin Andow, Samin Yaseer Mahmud, Wenyu Wang, Justin Whitaker, William Enck, Bradley Reaves, Kapil Singh, and Tao Xie. Policylint: Investigating internal privacy policy contradictions on google play. In *28th USENIX Security Symposium (USENIX Security 19)*, pages 585–602. USENIX Association, August 2019.

[3] Benjamin Andow, Samin Yaseer Mahmud, Justin Whitaker, William Enck, Bradley Reaves, Kapil Singh, and Serge Egelman. Actions speak louder than words: Entity-sensitive privacy policy and data flow analysis with policheck. In *29th USENIX Security Symposium (USENIX Security 20)*, pages 985–1002. USENIX Association, August 2020.

[4] William Enck, Peter Gilbert, Byung-Gon Chun, Landon P. Cox, Jaeyeon Jung, Patrick McDaniel, and Anmol N. Sheth. Taintdroid: An information-flow tracking system for realtime privacy monitoring on smartphones. In *9th USENIX Symposium on Operating Systems Design and Implementation (OSDI 10)*. USENIX Association, October 2010.

[5] William Enck, Damien Octeau, Patrick McDaniel, and Swarat Chaudhuri. A study of android application security. In *20th USENIX Security Symposium (USENIX Security 11)*, pages 315–330. USENIX Association, August 2011.

[6] Facebook. Facebook to acquire oculus. https://about.fb.com/news/2014/03/facebook-to-acquire-oculus/, March 2014. [Accessed: 06 June 2021].

[7] Earlence Fernandes, Jaeyeon Jung, and Atul Prakash. Security analysis of emerging smart home applications. In *2016 IEEE Symposium on Security and Privacy (SP)*, pages 636–654. IEEE, 2016.

[8] Hamza Harkous, Kassem Fawaz, Rémi Lebret, Florian Schaub, Kang G. Shin, and Karl Aberer. Polisis: Automated analysis and presentation of privacy policies using deep learning. In *27th USENIX Security Symposium (USENIX Security 18)*, pages 531–548. USENIX Association, August 2018.

[9] Scott Hayden. Oculus quest 2 surpasses original quest in monthly active users. https://www.roadtovr.com/oculus-quest-2-monthly-active-users/, January 2021. [Accessed: 06 June 2021].

[10] Haojian Jin, Minyi Liu, Kevan Dodhia, Yuanchun Li, Gaurav Srivastava, Matthew Fredrikson, Yuvraj Agarwal, and Jason I Hong. Why are they collecting my data? inferring the purposes of network traffic in mobile apps. In *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, volume 2, pages 1–27. ACM, 2018.

[11] Ben Lang. Where to change quest 2 privacy settings and see your vr data collected by facebook. https://www.roadtovr.com/oculus-quest-2-privacy-facebook-data-collection-settings/, October 2020. [Accessed: 06 June 2021].

[12] Christopher Lentzsch, Sheel Jayesh Shah, Benjamin Andow, Martin Degeling, Anupam Das, and William Enck. Hey alexa, is this skill safe?: Taking a closer look at the alexa skill ecosystem. In *28th Annual Network and Distributed System Security Symposium (NDSS 2021)*. The Internet Society, 2021.

[13] Mark Roman Miller, Fernanda Herrera, Hanseul Jun, James A Landay, and Jeremy N Bailenson. Personal identifiability of user tracking data during observation of 360-degree vr video. *Scientific Reports*, 10(1):1–10, 2020.

[14] Hooman Mohajeri Moghaddam, Gunes Acar, Ben Burgess, Arunesh Mathur, Danny Yuxing Huang, Nick Feamster, Edward W Felten, Prateek Mittal, and Arvind Narayanan. Watching you watch: The tracking ecosystem of over-the-top tv streaming devices. In *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security*, pages 131–147. ACM, 2019.

[15] Mozilla Corporation and Individual mozilla.org contributors. Privacy & security guide: Oculus quest 2 vr headset. https://foundation.mozilla.org/en/privacynotincluded/oculus-quest-2-vr-headset/, November 2020. [Accessed: 06 June 2021].

[16] Helen Nissenbaum. *Privacy in Context - Technology, Policy, and the Integrity of Social Life*. Stanford University Press, 2010.

[17] Abbas Razaghpanah, Rishab Nithyanand, Narseo Vallina-Rodriguez, Srikanth Sundaresan, Mark Allman, Christian Kreibich, and Phillipa Gill. Apps, trackers, privacy, and regulators: A global study of the mobile tracking ecosystem. In *25th Annual Network and Distributed System Security Symposium (NDSS 2018)*. The Internet Society, 2018.

[18] Jingjing Ren, Daniel J. Dubois, David Choffnes, Anna Maria Mandalari, Roman Kolcun, and Hamed Haddadi. Information exposure from consumer iot devices: A multidimensional, network-informed measurement approach. In *Proceedings of the Internet Measurement Conference (IMC 19)*, pages 267–279. ACM, 2019.

[19] Jingjing Ren, Ashwin Rao, Martina Lindorfer, Arnaud Legout, and David Choffnes. Recon: Revealing and controlling pii leaks in mobile network traffic. In *Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services (MobiSys 16)*, page 361–374. ACM, 2016.

[20] Sol Rogers. Virtual reality for good use cases: From educating on racial bias to pain relief during childbirth. https://www.forbes.com/sites/solrogers/2020/03/09/virtual-reality-for-good-use-cases-from-educating-on-racial-bias-to-pain-relief-during-childbirth/, March 2020. [Accessed: 06 June 2021].

[21] Anastasia Shuba, Anh Le, Emmanouil Alimpertis, Minas Gjoka, and Athina Markopoulou. Antmonitor: A system for on-device mobile network monitoring and its applications. *arXiv preprint arXiv:1611.04268*, 2016.

[22] Anastasia Shuba and Athina Markopoulou. Nomoats: Towards automatic detection of mobile tracking. In *Proceedings on Privacy Enhancing Technologies*, volume 2020, pages 45–66. Sciendo, 2020.

[23] Anastasia Shuba, Athina Markopoulou, and Zubair Shafiq. Nomoads: Effective and efficient cross-app mobile ad-blocking. In *Proceedings on Privacy Enhancing Technologies*, volume 2018, pages 125–140. Sciendo, 2018.

[24] Yan Shvartzshnaider, Noah Apthorpe, Nick Feamster, and Helen Nissenbaum. Going against the (appropriate) flow: a contextual integrity approach to privacy policy analysis. In *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing*, volume 7, pages 162–170, 2019.

[25] Rocky Slavin, Xiaoyin Wang, Mitra Bokaei Hosseini, James Hester, Ram Krishnan, Jaspreet Bhatia, Travis D. Breaux, and Jianwei Niu. Toward a framework for detecting privacy policy violations in android application code. In *Proceedings of the 38th International Conference on Software Engineering (ICSE 16)*, pages 25–36. ACM, 2016.

[26] Rahmadi Trimananda, Hieu Le, Hao Cui, Janice Tran Ho, Anastasia Shuba, and Athina Markopoulou. OVRseen: Auditing Network Traffic and Privacy Policies in Oculus VR. In *Proceedings of the 31st USENIX Security Symposium (USENIX Security 22)*, 2022.

[27] UCI Networking Group. OVRseen project page. https://athinagroup.eng.uci.edu/projects/ovrseen/.

[28] Janus Varmarken, Hieu Le, Anastasia Shuba, Athina Markopoulou, and Zubair Shafiq. The tv is smart and full of trackers: Measuring smart tv advertising and tracking. In *Proceedings on Privacy Enhancing Technologies*, volume 2020, pages 129–154. Sciendo, 2020.

[29] Xiaoyin Wang, Xue Qin, Mitra Bokaei Hosseini, Rocky Slavin, Travis D. Breaux, and Jianwei Niu. Guileak: Tracing privacy policy claims on user input data for android applications. In *2018 IEEE/ACM 40th International Conference on Software Engineering (ICSE)*, pages 37–47. IEEE, 2018.

[30] Sebastian Zimmeck and Steven M. Bellovin. Privee: An architecture for automatically analyzing web privacy policies. In *23rd USENIX Security Symposium (USENIX Security 14)*, pages 1–16. USENIX Association, August 2014.

[31] Sebastian Zimmeck, Ziqi Wang, Lieyong Zou, Roger Iyengar, Bin Liu, Florian Schaub, Shomir Wilson, Norman M. Sadeh, Steven M. Bellovin, and Joel R. Reidenberg. Automated analysis of privacy requirements for mobile apps. In *24th Annual Network and Distributed System Security Symposium (NDSS 2017)*. The Internet Society, 2017.