tool aimed at identifying WeBugs by following three distinct bug patterns [17]. Another study probed into problems such as system resource exposure, subwindow spoofing, and subroutine hijacking within the Mini-Program ecosystem, conducting evaluations on 11 prominent platforms to highlight the pervasive nature of these security issues [12]. Moreover, a novel issue regarding privacy leaks in MiniApps has been explored [22], potentially leading to private data theft by the MiniApp platform, with the researchers detailing an attack process that exploits this vulnerability. Additionally, the discovery of a Cross Mini-Program request forgery vulnerability (CMRF) [20] has been documented, along with the development of the CMRFScanner tool for its detection.

4.2 MiniApp Privacy

A series of studies have emphasized the importance of privacy in MiniApp ecosystem [7, 9, 13, 15, 18, 19, 23, 25]. TaintMini [15] introduced a framework for detecting flows of sensitive data within and across mini-programs using static taint analysis. Another work MiniTracker [9] constructed assignment flow graphs as common representation across different host apps and performed a large-scale study on 150k MiniApps, which revealed the common privacy leakage patterns. Moreover, several studies [7, 13, 25] have focused on taint analysis technique to detect AppSecret leaks. In particular, another work [18] focused on the consistency of data collection and usage in MiniApps. They crawled 2,998 MiniApps and detected 89.4% of them violated their privacy policies. More recently, Zhang et al. introduced SPOChecker and performed the first systematic study of privacy over-collection in MiniApps. Despite these significant contributions to understanding privacy dimensions, there remains a noticeable gap in the literature concerning the privacy compliance and data minimization of Miniapps, indicating a crucial area for further exploration and research.

5 CONCLUSION

This research has illuminated the pressing challenge of data minimization within the complex domain of MiniApps, a challenge that has been compounded by existing coarse-grained privacy measures. In response, we introduced an end-to-end hybrid analysis framework designed to address this issue at a fine-grained level. Comprising three key modules, the framework offers a nuanced, usage-scenario-based approach to data privacy. The planned large-scale study, encompassing 120K MiniApps, will further validate and potentially refine this groundbreaking solution. Our work signifies a vital step towards transparent and responsible data practices in MiniApps, contributing to broader advances in digital security and computer science.

ACKNOWLEDGEMENT

The authors would like to thank all the anonymous shepherd and reviewers for improving this manuscript. We also thank Moxuan Wang, Lizheng Wang, Zhengyang Xiao, Qian Huang for their insightful discussions and feedbacks. This work was supported in part by National Key R&D Program of China (2021YFB2701000), the National Natural Science Foundation of China (grant No.62072046), Knowledge Innovation Program of Wuhan-Basic Research and HUST CSE-HongXin Joint Institute for Cyber Security.

REFERENCES

[1] 2022. Act on the Protection of Personal Information. https://www.ppc.go.jp/.
[2] 2022. California Consumer Privacy Act. https://oag.ca.gov/privacy/ccpa.
[3] 2022. General Data Protection Regulation. https://commission.europa.eu/law/law-topic/data-protection_en.
[4] 2022. Personal Data Protection Act. https://www.pdpc.gov.sg/.
[5] 2023. WECHAT POLICY. https://www.wechat.com/en/privacy_policy.html.
[6] Benjamin Andow, Samin Yaseer Mahmud, Justin Whitaker, William Enck, Bradley Reaves, Kapil Singh, and Serge Egelman. 2020. Actions speak louder than words:{Entity-Sensitive} privacy policy and data flow analysis with {PoliCheck}. In *29th USENIX Security Symposium (USENIX Security 20)*. 985–1002.
[7] Supraja Baskaran, Lianying Zhao, Mohammad Mannan, and Amr Youssef. 2023. Measuring the Leakage and Exploitability of Authentication Secrets in Super-apps: The WeChat Case. *arXiv preprint arXiv:2307.09317* (2023).
[8] Duc Bui, Yuan Yao, Kang G Shin, Jong-Min Choi, and Junbum Shin. 2021. Consistency analysis of data-usage purposes in mobile apps. In *Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security*. 2824–2843.
[9] Wei Li, Borui Yang, Hangyu Ye, Liyao Xiang, Qingxiao Tao, Xinbing Wang, and Chenghu Zhou. 2023. MiniTracker: Large-Scale Sensitive Information Tracking in Mini Apps. *IEEE Transactions on Dependable and Secure Computing* (2023).
[10] Yuxi Ling, Kailong Wang, Guangdong Bai, Haoyu Wang, and Jin Song Dong. 2022. Are they toeing the line? diagnosing privacy compliance violations among browser extensions. In *Proceedings of the 37th IEEE/ACM International Conference on Automated Software Engineering*. 1–12.
[11] Yi Liu, Jinhui Xie, Jianbo Yang, Shiyu Guo, Yuetang Deng, Shuqing Li, Yechang Wu, and Yepang Liu. 2020. Industry practice of javascript dynamic analysis on wechat mini-programs. In *Proceedings of the 35th IEEE/ACM International Conference on Automated Software Engineering*. 1189–1193.
[12] Haoran Lu, Luyi Xing, Yue Xiao, Yifan Zhang, Xiaojing Liao, XiaoFeng Wang, and Xueqiang Wang. 2020. Demystifying resource management risks in emerging mobile app-in-app ecosystems. In *Proceedings of the 2020 ACM SIGSAC conference on computer and communications Security*. 569–585.
[13] Shi Meng, Liu Wang, Shenao Wang, Kailong Wang, Xusheng Xiao, Guangdong Bai, and Haoyu Wang. 2023. WeMinT: Tainting Sensitive Data Leaks in WeChat Mini-Programs. In *Proceedings of the 38th IEEE/ACM International Conference on Automated Software Engineering*. IEEE.
[14] Faysal Hossain Shezan, Zihao Su, Mingqing Kang, Nicholas Phair, Patrick William Thomas, Michelangelo van Dam, Yinzhi Cao, and Yuan Tian. 2023. CHKPLUG: Checking GDPR Compliance of WordPress Plugins via Cross-language Code Property Graph.. In *NDSS*.
[15] Chao Wang, Ronny Ko, Yue Zhang, Yuqing Yang, and Zhiqiang Lin. 2023. Taintmini: Detecting flow of sensitive data in mini-programs with static taint analysis. In *2023 IEEE/ACM 45th International Conference on Software Engineering (ICSE)*. IEEE, 932–944.
[16] Chao Wang, Yue Zhang, and Zhiqiang Lin. 2023. Uncovering and Exploiting Hidden APIs in Mobile Super Apps. *arXiv preprint arXiv:2306.08134* (2023).
[17] Tao Wang, Qingxin Xu, Xiaoning Chang, Wensheng Dou, Jiaxin Zhu, Jinhui Xie, Yuetang Deng, Jianbo Yang, Jiaheng Yang, Jun Wei, et al. 2022. Characterizing and detecting bugs in WeChat mini-programs. In *Proceedings of the 44th International Conference on Software Engineering*. 363–375.
[18] Yin Wang, Ming Fan, Junfeng Liu, Junjie Tao, Wuxia Jin, Qi Xiong, Yuhao Liu, Qinghua Zheng, and Ting Liu. 2023. Do as You Say: Consistency Detection of Data Practice in Program Code and Privacy Policy in Mini-App. *arXiv preprint arXiv:2302.13860* (2023).
[19] Yuqing Yang, Chao Wang, Yue Zhang, and Zhiqiang Lin. 2023. SoK: Decoding the Super App Enigma: The Security Mechanisms, Threats, and Trade-offs in OS-alike Apps. *arXiv preprint arXiv:2306.07495* (2023).
[20] Yuqing Yang, Yue Zhang, and Zhiqiang Lin. 2022. Cross Miniapp Request Forgery: Root Causes, Attacks, and Vulnerability Detection. In *Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security*. 3079–3092.
[21] Jianyi Zhang, Leixin Yang, Yuyang Han, Zixiao Xiang, and Xiali Hei. 2023. A Small Leak Will Sink Many Ships: Vulnerabilities Related to mini-programs Permissions. In *2023 IEEE 47th Annual Computers, Software, and Applications Conference (COMPSAC)*. IEEE, 595–606.
[22] Lei Zhang, Zhibo Zhang, Ancong Liu, Yinzhi Cao, Xiaohan Zhang, Yanjun Chen, Yuan Zhang, Guangliang Yang, and Min Yang. 2022. Identity Confusion in {WebView-based} Mobile App-in-app Ecosystems. In *31st USENIX Security Symposium (USENIX Security 22)*. 1597–1613.
[23] Xiaohan Zhang, Yang Wang, Xin Zhang, Ziqi Huang, Lei Zhang, and Min Yang. 2023. Understanding Privacy Over-collection in WeChat Sub-app Ecosystem. *arXiv preprint arXiv:2306.08391* (2023).
[24] Yue Zhang, Bayan Turkistani, Allen Yuqing Yang, Chaoshun Zuo, and Zhiqiang Lin. 2021. A measurement study of wechat mini-apps. *ACM SIGMETRICS Performance Evaluation Review* 49, 1 (2021), 19–20.
[25] Yue Zhang, Yuqing Yang, and Zhiqiang Lin. 2023. Don't Leak Your Keys: Understanding, Measuring, and Exploiting the AppSecret Leaks in Mini-Programs. *arXiv preprint arXiv:2306.08151* (2023).