	イナヨシ　ヒロキ
氏　　名	**稲吉　弘樹**
学位の種類	博士（工学）
学位記番号	博第1326号
学位授与の日付	2024年3月31日
学位授与の条件	学位規則第４条第１項該当　課程博士
学位論文題目	A Study on Taint Analysis with Runtime Data for Tracking Information Flows in Android Apps (Android アプリ内情報フローを追跡する実行時データを用いたテイント解析に関する研究)
論文審査委員	主査　教授　齋藤　彰一 教授　松尾　啓志 准教授　打矢　隆弘 教授　吉田　敦 （南山大学）

論文内容の要旨

Android OS occupies 70% of the total mobile OS market share in 2023. The official market, Google Play Store, currently provides 2.6 million apps downloaded and used in users' day-to-day activities on their devices, always connected to the internet. On the other hand, the need to protect user privacy has been increasing. Data protection regulations, such as COPPA, CCPA, and GDPR, have been put into operation in the past two decades. Google has also made changes to data protection mechanisms on the OS and policies in the Google Play Store. It is also essential to uncover how well app developers and third-party SDK providers follow the rules to protect user privacy. Researchers have investigated real-world apps and found many non-compliant, policy-violating, and protection-circumventing behaviors. Taint analysis techniques have been actively developed and utilized to detect such suspicious behaviors.

Toward a reliable analysis of real-world apps, this paper addresses two issues. First, taint analysis can be circumvented by anti-taint-analysis (ATA) techniques. A series of ATA techniques has been demonstrated on the Android platform. They are only a few lines of code each and could be introduced into apps with obfuscator tools by app

developers to defend their apps against a taint analysis. However, there are only a few counter approaches against ATA techniques, which are only partially effective against ATA techniques. Second, the community needs a reliable taint tracker for analyzing real-world apps. Researchers recently tested popular static taint analyzers and concluded that the tools are inaccurate and cannot be used for analyzing real-world apps dependably. On the other hand, researchers examined a famous dynamic taint tracker, TaintDroid, and pointed out that TaintDroid is the most difficult to set up compared to the static analysis tools they audited. Also, TaintDroid depends on specific devices and versions of Android OS released in 2013, narrowing down the scope of analyzable apps. Other dynamic analyzers are not effortlessly usable.

This paper proposes two approaches named VTDroid and T-Recs based on the idea of utilizing the app's runtime data to improve the taint analysis. Chapter 3 explains VTDroid, designed to make it difficult for apps to evade taint tracking by neutralizing such uncomplicated techniques not specific to a particular ATA technique. This paper characterizes the ATA techniques by four types of information flow and proposes value logging and matching that propagate taint among registers based on their data values, in addition to the traditional bytecode-level tracking. VTDroid is evaluated with newly created test suites and real-world apps compared with TaintDroid, CTT, and FlowDroid. The results demonstrate that VTDroid tracks more information flows resulting from the ATA techniques and generates fewer FPs than CTT.

Chapter 4 describes T-Recs, a taint tracker that solves the current situation of no tracker that can analyze apps reliably. It records and reconstructs the app execution and performs taint analysis on an ordinary computer, not depending on Android OS. T-Recs' accuracy, analysis time, and success rate are evaluated in privacy leak detection compared to currently available taint analyzers, which are FlowDroid (w/ and w/o IC3), Amandroid, DroidSafe, DroidRA, IccTA, and TaintDroid (w/ and w/o IntelliDroid). The evaluation involves 158 test cases in DroidBench, 254 popular apps from Google Play in 2016 and 2021, and 39,480 SDK-version-varied apps from Google Play and Anzhi. The results show that T-Recs outperforms the compared tools in detection accuracy. T-Recs also achieves reasonable analysis time, app-runtime overhead, and success rate. VTDroid and T-Recs have been made available to the community.

The regulations, market policies, and protection mechanisms will be reformed in the future, and researchers should keep examining apps and libraries. VTDroid and T-Recs should be promising tools that empower researchers to analyze apps in the future.

論文審査結果の要旨

Android OSはモバイルOSシェアの70%を占め広く利用される一方で，アプリによる規則違反やストアポリシー違反，OSのデータ保護メカニズムの回避といった問題が実態調査により発見されてきた．実態調査では，アプリ解析にテイント解析が用いられるため，信頼できるテイント解析システムが求められている．

より信頼できるテイント解析システムの実現に向けて，申請者は2つの問題に取り組んだ．1つ目は，テイント解析を回避するanti-taint-analysis (ATA) である．公開されているものとして，16種類のATAテクニックを集めたScrubDroidがある．ATA搭載アプリがアプリストアで流通して実態調査に影響することが懸念される．一方で，ATAへの対応に取り組んだ研究は少なく，どれも部分的に有効なもののみである．2つ目は，実世界アプリの解析において信頼できるテイント解析システムそのものが不足している問題である．近年の研究で，代表的な静的テイント解析システムは実世界アプリの解析において信頼できる結果を得られないと指摘された．代表的な動的テイント解析システムであるTaintDroidが要求する端末は入手性が悪く寿命も短く，新しいアプリを解析できず，性能が弱いため解析時間もかかる．また，その他の既存の動的テイント解析システムは，調査した限りでは実利用に足るものはない．

申請者は，テイント解析にアプリの実行時データを活用するというアイディアで，VTDroidとT-Recsの2つを提案している．VTDroidは， ScrubDroidが示す既知のATAテクニックを全体的に無力化することで，アプリによるテイント解析回避をより困難にしている．提案手法は，アプリ実行時のレジスタの値を用いて情報維持判定を行うことで，誤追跡を削減する．新たに作成したテストスイートと実世界のアプリを用いて，既存の3システムと比較した結果，VTDroidはScrubDroidに全体的に対応でき，CTTよりも大幅に誤検知が少なく，より実利用に導入しやすいことを示した．

T-Recsは，記録したアプリ実行時データを元に，Android OSに依存しない通常の計算機上でアプリ実行を再構築してテイント解析を行うことで，正確性と端末非依存性を両立する新たなテイント解析システムである．新旧幅広いOSバージョンに対応でき，1度のアプリ実行で複数回のテイント解析が可能である．T-Recsを現在利用可能な10のシステムと幅広いアプリを用いて評価した結果，T-Recsは正確性や解析成功率が高く，解析時間とアプリ実行時オーバヘッドは現実的な範囲内であることを示した．

アプリストアポリシーなどは今後も変更されると予想され，引き続き実態調査が必要である．VTDroidとT-Recsは申請者によって公開されており，今後の実態調査での活用が期待できる．

これらの研究成果は，コンピュータセキュリティ分野およびソフトウェア工学分野での優れた評価を得て，2編の学術雑誌論文と3編の国際会議論文（いずれも審査あり）として発表されている．以上の研究成果，論文内容の審査ならびに公聴会当日の質疑応答により，博士の学位を授与するに足りる成果であると判断する．