1. Should You Consider Adware as Malware in Your Study?

Jun Gao (1); Li Li (2); Pingfan Kong (1); Bissyande, T.F. (1); Klein, J. (1)

Source: *2019 IEEE 26th International Conference on Software Analysis, Evolution and Reengineering (SANER). Proceedings*, p 604-8, 2019; **ISBN-13:** 978-1-7281-0591-8; **DOI:** 10.1109/SANER.2019.8668010; **Conference:** 2019 IEEE 26th International Conference on Software Analysis, Evolution and Reengineering (SANER), 24-27 Feb. 2019, Hangzhou, China; **Publisher:** IEEE, Piscataway, NJ, USA

Author affiliation: (1) University of Luxembourg, Luxembourg (2) Monash University, Melbourne, VIC, Australia

Abstract: Empirical validations of research approaches eventually require a curated ground truth. In studies related to Android malware, such a ground truth is built by leveraging Anti-Virus (AV) scanning reports which are often provided free through online services such as VirusTotal. Unfortunately, these reports do not offer precise information for appropriately and uniquely assigning classes to samples in app datasets: AV engines indeed do not have a consensus on specifying information in labels. Furthermore, labels often mix information related to families, types, etc. In particular, the notion of “adware” is currently blurry when it comes to maliciousness. There is thus a need to thoroughly investigate cases where adware samples can actually be associated with malware (e.g., because they are tagged as adware but could be considered as malware as well). In this work, we present a large-scale analytical study of Android adware samples to quantify to what extent “adware should be considered as malware”. Our analysis is based on the Androzoo repository of 5 million apps with associated AV labels and leverages a state-of-the-art label harmonization tool to infer the malicious type of apps before confronting it against the ad families that each adware app is associated with. We found that all adware families include samples that are actually known to implement specific malicious behavior types. Up to 50% of samples in an ad family could be flagged as malicious. Overall the study demonstrates that adware is not necessarily benign. (0 refs)

Inspec controlled terms: Android (operating system) - invasive software - mobile computing
Uncontrolled terms: adware app - specific malicious behavior types - empirical validations - Android malware - online services - app datasets - AV engines - Android adware samples - anti-virus scanning - label harmonization tool
Classification Code: C6190V Mobile, ubiquitous and pervasive computing - C6130S Data security
IPC Code: G06F9/44 - G06F9/46 - G06F21/00
Treatment: Practical (PRA)
Database: Inspec
Data Provider: Engineering Village