DARTMOUTH

Institute for Security
Technology and Society
at Dartmouth

Android Security Research

As one of the most widely used mobile operating systems today, Android is consistently the largest target of malware. Nokia's 2017 Threat Intelligence Report states that a staggering 68.5% of all malware targets the Android platform. As a consequence, the massive volume and variants of Android malware often make it difficult for security analysts to detect, classify, and analyze them. Therefore, research into Android malware is crucial for the development of more effective security measures and implementations. To address this, Dartmouth has led a number of research projects into Android security, which have improved the classification and detection of Android malware.

DBank and FARM have both detected two previously unknown Android banking trojans and rooting malware, respectively. These four have been confirmed by the Google Android Security Team.

EC2 (Ensemble Clustering and Classification)

Given the ease with which malware developers use obfuscation to generate numerous variants of the same malware, it is impractical to manually investigate and determine signatures for the increasing numbers of new malware. Dartmouth researchers have developed EC2 as an automated algorithm that combines both classification and clustering techniques to achieve better performance in detecting both small and large Android malware families.

The features extracted from malware in this project mainly fall into two categories: static and dynamic. Specifically, static features consist of author, application components (such as file size or number activities), and permissions while dynamic features consist of a bag of words of operations (such as read, write, and load).

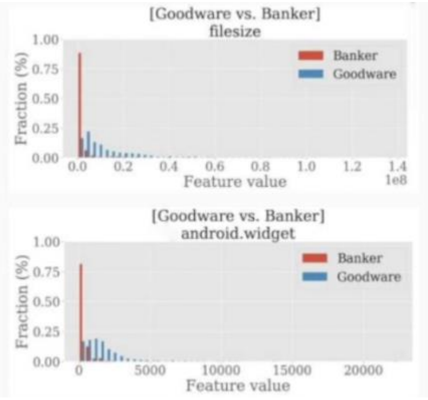
Using these features, EC2 systematically combines supervised classification and unsupervised clustering algorithms. This allows EC2 to minimize classification algorithms' weakness of predicting small families with clustering algorithms' strength in capturing them.

After evaluating the performance of a number of classification and clustering algorithms (the results of which are available on the full paper), EC2's performance was also tested. EC2 was found to outperform several baselines, with an overall Micro and Macro F-Score of 0.76 and 0.97, respectively. EC2 is therefore a highly effective Android malware family detection tool.

DBank

DBank is a framework developed by Dartmouth that detects and characterizes Android Banking Trojans (ABT). DBank invents a novel concept of a Triadic Suspicion Graph (TSG), which has goodware, banking trojans, and Android API packages as nodes. Suspicion scores (SUS) and suspicion ranks (SR) associated with the API packages are also calculated, which takes into account how often goodware and banking trojans invoke the API. After evaluation, DBank was found to have a 99.9% predictive accuracy (AUC) and a 0.3% false positive rate.

DBank was also tested against certain adversary attacks, such as how well an attacker could infer the predictions of the defender model given an intersection of defender and adversary training data. Testing revealed that, even with 80% intersection of data with the adversary, the adversary's error rate is multiple times higher for ABT detection with DBank. Regarding robustness while using a subset of training data, DBank still had a high AUC when using 20% or 30% of the training set, allowing the defender to have a moving defense surface. DBank also proved robust against fake calls to an API to avoid suspicion. Additionally, DBank has identified two previously unknown ABTs, which Google has confirmed. Overall, it is clear that DBank is a reliable and accurate system in detecting ABTs.



General analysis results of features on goodware vs. banking trojans

https://www.cs.dartmouth.edu/~dsail/projects/android.html

Data-Driven Android Spyware Characterization

Spyware in Android devices continue to increase in number and percentage of mobile malware. It is therefore crucial to understand which features differentiate spyware from goodware and other malware. Dartmouth has led a project in using both classical and deep machine learning to find these defining features.

Feature extraction for this project, similar to EC2, consists of static and dynamic features. The static features used here comprise of author, permissions and structure, while the dynamic features comprise of operations such as read, write, or sendsms.

After extracting static and dynamic features, Ensemble Late Fusion (ELF) was used to identify key features used for differentiation of Android spyware. Evaluation of ELF against traditional and deep learning classifiers showed an improvement across all performance metrics when comparing spyware against goodware and spyware against other malware, making it highly qualified in detecting Android spyware.

FARM

FARM, a Feature transformation based Android Rooting Malware detector, is a Dartmouth project that uses machine learning to detect Android rooting malware. Especially since rooting malware is much more difficult to remove than other malware, research into prevention is essential.

Like the previous projects, FARM extracts basic static and dynamic features from Android APKs. Using these features, FARM creates a "basic feature vector" for each APK in the dataset of rooting malware samples and goodware. FARM then transforms these basic feature vectors into a new feature space using three new classes of irreversible feature transformation: landmark based, feature value clustering based, and correlation graph based feature transformations.

Through experimental evaluation, FARM outperformed baselines in detecting rooting malware against goodware and other malware. FARM also proved robust against adversarial attacks, such as fake API call, fake permission, and reduced API feature attacks. In addition, FARM has labeled two malware samples on VirusTotal as rooting malware before this was observed by any of the 61 anti-virus engines on the site. The samples were reported to Google's Android Security Team who have confirmed the findings. Given these results, FARM is a remarkably accurate rooting malware detection system.

Additional Information

References

1. Chakraborty, Tanmoy & Pierazzi, Fabio & Subrahmanian, V.S. (2017). "EC2: Ensemble Clustering and Classification for Predicting Android Malware Families." *IEEE Transactions on Dependable and Secure Computing*.
2. C. Bai, Q. Han, G. Mezzour, F. Pierazzi and V. S. Subrahmanian, "DBank: Predictive Behavioral Analysis of Recent Android Banking Trojans." *IEEE Transactions on Dependable and Secure Computing*.
3. Michele Colajanni, Qian Han, Ghita Mezzour, Fabio Pierazzi, and V.S. Subrahmanian. "A Data-Driven Characterization of Modern Android Spyware." Under review.
4. Qian Han, V.S. Subrahmanian, and Yanhai Xiong. Android Rooting Malware Detection via (Somewhat) Robust Irreversible Feature Transformations. Under review.

Video

https://www.cs.dartmouth.edu/~dsail/projects/android/videos/

Presentation

https://www.cs.dartmouth.edu/~dsail/projects/android/slides/

PARTICIPANTS

Lead: V.S. Subrahmanian

Chongyang Bai, Tanmoy Chakraborty, Haipeng Chen, Michele Colajanni, Qian Han, Ghita Mezzour, Fabio Pierazzi, Yanhai Xiong.

DARTMOUTH
Security and Artificial Intelligence Laboratory