Using Twitter Data to Improve Android Malware Detection

Jordan DeLoach and Doina Caragea
Kansas State University
Kansas State University
Computer Science

Abstract

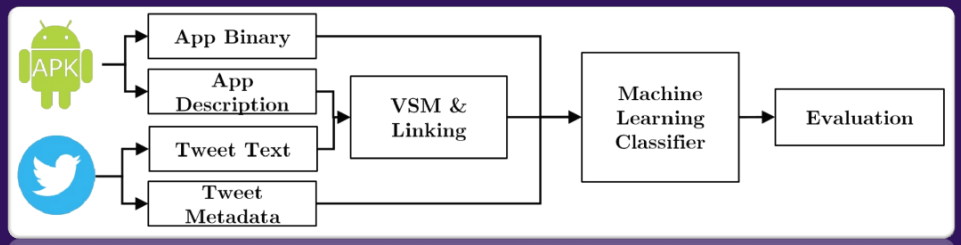
Existing machine learning approaches to Android malware detection have nearly exclusively used app contents to extract features for classification. We seek to understand if auxiliary Twitter data can be used to improve the performance of existing approaches. Throughout the course of our research, we collected over 50 million tweets potentially related to Android apps. We propose to link tweets with apps using approaches inspired from the standard vector space model, and subsequently study the usefulness of the linked tweets in malware detection. We find that Twitter data accurately linked to apps through HTTP links can be used to improve the machine learning classifier performance across a variety of common malware detection classifiers.

Problem

We study the usefulness of Twitter data in detecting Android malware, under the assumption that the text and/or metadata associated with a tweet referring to an Android app can be indicative of benign versus malicious behavior.

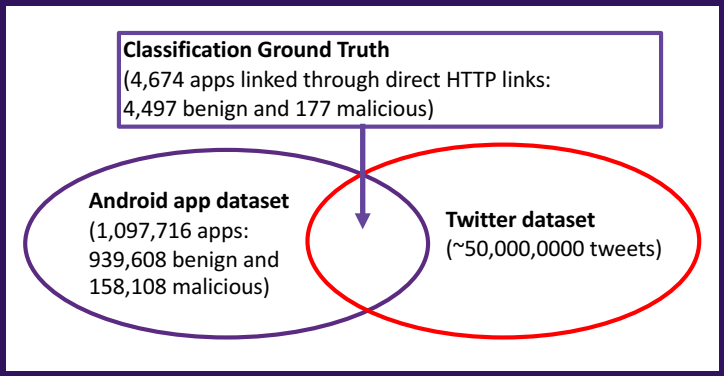
Approaches

- We link tweets to apps using two main approaches: 1) direct HTTP links in the tweet text to apps in Google Play Store; 2) matches between tweet text and app descriptions using approaches inspired from the standard vector space model (VSM).
- We use information extracted from the linked tweets (e.g., metadata about the tweet or about the user who posted the tweet) to augment the feature vectors provided as input to machine learning classifiers trained to detect Android malware.



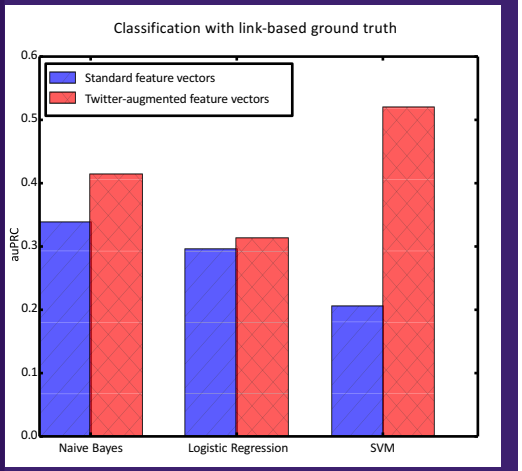
Experimental Setup

- We assemble a classification ground truth dataset by selecting Android apps for which tweets with direct links to the apps exist.
- We augment the feature vectors constructed from app binaries with features extracted from tweets, and compare the two sets of features using standard machine learning algorithms: Naïve Bayes, Logistic Regression and support vector machines (SVM).



Results

- We find that Twitter-augmented feature vectors perform better than the feature vectors constructed solely from app binaries, regardless of the learning algorithm used.



Conclusions

- Twitter data can be used to improve the results of the machine learning algorithms for Android malware detection.
- Information indicative of spam tweets or spam users contributes more to the classification as compared to the text or the sentiment of the tweet.
- Classification experiments with Twitter data automatically linked to apps using VSM-type approaches reveal the need for more robust linking approaches.

Acknowledgements: This research is supported by National Science Foundation under Grants No. 1429316 and 1717871. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessary reflect the views of National Science Foundation.