COLUMBIA | ENGINEERING
The Fu Foundation School of Engineering and Applied Science

Boyu Liu, Duanyue Yun, Xin Guo, Xiao Ji, Huiyu Song | Data Science Institute
Shirish Singh, Gail Kaiser | Dept. of Computer Science

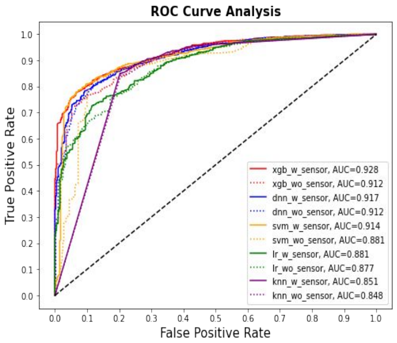
Detecting Sensor-Based Repackaged Malware

Introduction:

- The privacy implications of zero-permission sensors have previously been studied. The sensor data can be utilized in the app code or via third party libraries.
- In this study, we analyzed 15,297 app pairs and found evidence that repackaged malware utilize additional sensors to perform malicious activities. We incorporated sensor-related features into classification models for detecting malware.

Results:

- We built a pipeline to extract sensor features from raw APK files.
- We trained SVM, KNN, Logistic Regression, XGBoost and DNN to detect malware.



Statistical analysis:

- We conducted a two-proportion z-test on the null hypothesis that malicious apps are as likely to use sensors as benign apps.
- At 1% significance level, we found statistically significant evidence that the proportion of malicious apps using sensors is higher than that of benign apps.

Conclusion:

- The proportion of malicious apps using sensors is higher than that of benign apps.
- We trained five classifiers to detect repackaged malware using 892 features extracted from the APK files, achieving a detection rate of 95%.
- Sensor-related features improve model performance.