Proposed Topic/Title of Research:

Information flow analysis for mobile applications

Peter Zhang

Background:

The pervasive usage of mobile phone applications is now changing the way people use traditional software. Smartphone apps generated an impressive USD 45 billion in full-year 2016, and in total 138 billion apps were downloaded in the year. The last few years have seen an unprecedented number of people rushing to develop mobile apps. Android which occupies more than 80% of the mobile market, has already been an essential part of our daly life. However, the safety issues of Android apps also stand out. Privacy leaks, code hijacking, and system invasion become common yet severe problems.

Traditional static value-flow analysis method uses pointer/alias analysis for modeling program control and data dependence. For precise analysis of Android apps, traditional value-flow analysis has some major limitations in analyzing complicated Android features, such as message sending/receiving through the internet, callbacks for system-event handling, UI interaction and components with distinct life cycles, etc.

Aims:

To overcome the above limitations, the project aims to design and implement a new tool to accurately locate potential security problems for Android applications and help improve the existing permission system in the Android OS. It will automatically detect critical security bugs, such as information leakage, for modern Android applications.

Approach:

The proposed project will achieve the above goal through the following techniques:

1) Value-flow based information flow analysis. The new information leak analysis will use dynamic analysis to obtain a new permission vulnerability matrix of Android and the test cases based on the vulnerable program parts determined by static analysis. The performance of the new tool will be tested based on different Android applications source code. The implementation will be based on the existing tools, such as SVF (https://github.com/SVF-tools/SVF), developed by Dr. Yulei’ s research team.

2) A machine-learning guided static analysis approach that bridges the gap between the existing static analysis and dynamic analysis by capturing the correlations between complicated Android program features and the state-of-the-art program analysis tools. The approach aims to learn and predict complicated likelihood information leakage bugs by steering program analysis using Support Vector Machine (SVM) or TensorFlow.

Expected Outcomes:

The expected outcomes of the project are 1) an open-source tool for automatically detecting privacy data leakage of the Android applications; 2) high quality publications in the area of software engineering and artificial intelligence.