Research Proposal

Information flow analysis for mobile applications

STUDENT NAME:	**Yanxin Zhang**
COURSE NAME:	**Doctor of Philosophy**
DEPARTMENT:	**Faculty of Engineering and Information Technology**
COURSE CODE:	**Course code C02029**
SUPERVISOR:	**Yulei Sui**
DATE OF SUBMISSION:	**19/10/2017**

1. Introduction:

The pervasive use of mobile phone applications is now changing the way people use the 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 daily life. However, the safety issues of Android apps also stand out. Privacy leaks, code hijacking, and system invasion become common yet severe problems.

2. Problem statement

Due to high overheads incurred in bounds checking at runtime, many program inputs cannot be exercised, that will cause some input-specific spatial errors to go undetected in today' s commercial software.Spatial errors (e.g., buffer overflows) continue to be one of the dominant threats to software reliability and security in C/C++ programs. Presently, the software industry typically enforces spatial memory safety by instrumentation.

3. Objectives

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.

4. Preliminary Literature Review:

A preliminary literature review shows that traditional static value-flow analysis method uses pointer/alias analysis for modeling program control and data dependence. For the 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. What is missing from the past studies is a comprehensive and structured method in information flow analysis for mobile applications.

5. Methodology:

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.

6. 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.

7.Research Timeline

Months 1-6	Conduct continuous, through literature review to identify gaps in knowledge and experts Identify specific aims of project based on research vision,plan, data results literature review results
Months 7-11	Target on experts based on their objective ans how they relate to my specific project
Months 12-16	Obtain advice/guidance from colleagues and sponsor sources(e.g., research supervisor) Know potential reviewers(my audience)
Months 17-27	Write research paper draft Put research paper draft aside for a time,then edit Ask external reviewers to review draft and provide comments
Months 28-33	Rewrite research paper based on external reviewers comments(this process may continue until close to proposal submission)
Months 33-36	Submit Ph.d paper to UTS graduate school and review

References:

- Yulei Sui and Jingling Xue. On-Demand Strong Update Analysis via Value-Flow Refinement，[slide] ACM SIGSOFT International Symposium on the Foundation of Software Engineering.

- Hua Yan, Yulei Sui, Shiping Chen, and Jingling Xue. Automated Memory Leak Fixing on Value-Flow Slices for C Programs，31st ACM/SIGAPP Symposium on Applied Computing

- Jieyuan Zhang, Yulei Sui and Jingling Xue. Incremental Analysis for Probabilistic Programs ,24th International Static Analysis Symposium

- Yulei Sui, Ding Ye, and Jingling Xue. Detecting Memory Leaks Statically with Full-Sparse Value-Flow Analysis，IEEE Transactions on Software Engineering