《信息安全创新综合实验》设计任务书

一、设计目的

1.培养学生在综合运用各种理论知识的基础上，结合具体案例并动手进行工程应用设计。

2.使学生直接体验实验的方案制订、方案实施、设计实现和综合测试等，提高分析问题解决问题的能力，达到综合应用的目的，进一步加深学生对信息安全相关理论和概念的理解和综合应用。

3.使学生能掌握综合运用专业基础理论知识、技术方法和实践技巧进行信息安全系统设计、信息安全基础部件的研究与构造、信息安全各环节综合分析设计、信息安全系统评估能力和信息安全系统的运行与维护的能力。

二、设计题目及任务

题目：基于图神经网络的安卓恶意软件检测模型

简介：

安卓操作系统作为移动设备上最广泛使用的开放源代码操作系统，具有着丰富的应用市场与客户人群，但其普及率与快速的增长也为恶意软件发展提供了温床。为此，现有研究者通过提取程序特征来实现恶意软件检测机制。但是这些方法往往忽视了程序语义信息，导致恶意软件检测精度不理想。为此，该系统设计题目需要针对安卓应用软件，引入图卷积神经网络，能够将程序语义关系转换为图模型，进而采用图分析方法，实现精确度高、检测速度快的恶意软件检测方法。

基本方法：

相关工作：现有基于深度学习的安卓恶意软件检测方法主要可以分为基于语法特征和基于语义特征的检测方法。基于语法特征的方法主要关注于代码的语法结构，通过从语法结构中提取出代表性的特征与恶意行为标签，来训练模型从而实施检测。该方法的优点在于可以快速准确地识别简单和常见的恶意代码，但是对于复杂和高级的恶意软件则表现出相对较差的准确性和效率。基于语义特征的方法通常通常会采用程序分析技术。程序分析是一种将程序的行为抽象为数学模型的技术，以便进一步展示和分析，相比于基于语法特征的方法，该方法可以更好地识别复杂和高级的恶意软件，以及从中提取有用的信息。通过程序分析，研究者可以更好地提取恶意软件执行过程中的语义信息，以理解其行为和特征，这类特征往往会通过图和数据流的方法进行表达。

可行方法：该系统可行的方法是对安卓应用程序首先进行静态分析，提取应用程序的函数调用图，并对利用现有敏感 API 数据集（https://apichecker.github.io/）对函数调用图进行自动标注。然后设计图卷积神经网络模型，并通过 One-Hot 编码实现对函数调用图的特征编码，对图卷积神经网络模型每个节点特征向量进行初始化，最后利用现有数据集对模型进行训练。其中安卓应用程序静态分析工具

可以采用 Androguard（https://github.com/androguard/androguard）。恶意安卓应用可以从 https://virusshare.com/ 平台上下载，正常安卓应用可以从 AndroZoo（https://androzoo.uni.lu/）平台上下载。

任务：

1. 查阅相关资料，了解恶意软件检测的原理；
2. 设计实现安卓应用程序的函数调用图提取及其敏感 API 调用自动标注方法；
3. 设计实现基于图卷积神经网络模型的安卓恶意软件模型训练方法；
4. 设计实验分析方法，要求测试恶意应用和正常应用均不少于 50 个，评估模型的准确率、召回率和 F1 得分；
5. 根据小组分工，撰写个人课程报告，形成整体作品报告。

三、应提交的作品

1. 个人课程报告，内容包括：

（1）实验背景和目的；

（2）实验内容，包括实验方案、实验过程、实验结果；

（3）体会和收获。

2. 小组作品报告，内容包括：

（1）作品概述；

（2）作品设计与实现；

（3）作品测试与分析；

（4）创新性说明；

（5）总结。

3. 小组汇报 PPT（10 分钟）。

四、评分标准

构成最终成绩由课程报告(30%),作品报告及汇报答辩(60%),平时成绩(10%)三部分组成（参见教学大纲）。各部分所占比例如下：

考核方式	考核/评价环节	权重	备注
课程报告	规范性	15	是否覆盖了实验报告要求

	原创性	15	有无原创性、有无个人的实质性工作
作品报告及汇报答辩	准备充分度	30	实验、作品报告及答辩准备是否充分
	表达能力	30	语言表达流畅度、缜密程度
平时成绩	参与度	10	是否积极参与实验作品的设计开发

五、时间安排

周次	工作内容
第一周	布置设计任务及分工，设计方案并开展实验
第二周	完成实验内容并撰写报告，课程设计答辩