反编译工具Dex2jar的优化与改进

软件项目计划书

小组成员：王文茹 SY1606415

卢兴海 SY1606414

蒋 波 SY1606412

# 1 引言

## 1.1背景

计算机软件反向工程（Reverse engineering）也称为计算机软件还原工程，是指通过对他人软件的目标程序（比如可执行程序）进行“逆向分析、研究”工作，以推导出他人的软件产 品所使用的思路、原理、结构、算法、处理过程、运行方法等设计要素，某些特定情况下可能推导出源代码。反编译作为自己开发软件时的参考，或者直接用于自己 的软件产品中。

随着软件技术的不断发展、对现有软件的学习、理解、改造、维护和复用日益变得重要，在不侵犯软件版权或者经授权的情况下，引进软件技术的消化、吸收和汉化也具有巨大的经济效益。因此，反编译显得越来越重要，归纳起来，它具有如下一些作用：

维护和改造拥有使用权的程序。按目前的软件交货方式，软件制造者一般只向用户提供软件的机器代码程序和有关软件使用操作的文档，由于技术保护等原因，通常不提供源代码，这给用户维护和改进程序造成了严重的困难。

恢复自己开发的程序的源代码的丢失了的部分。这种情况并不罕见，例如意外的事故，保管不善，被离职的雇员携走等。

探测外来软件，检查外来软件是否剽窃了别的软件中部分，为了国家安全和军事的需要，检查来自他国的软件中是否含有潜在的危险的部分。

获取软件可重用知识与部件。软件重用已经引起软件界的广泛重视。但可重用知识与部件的获取是其中的困难之一，从已有软件中获取可重用知识与部件是解决这一困难的途径之一，对于机器代码程序，如果能将它翻译成高级语言，将有利于可重用知识与部件的获取。

在安全关键性的系统中，假设编译器得不到足够信任，因此需要反编译可执行代码，并且由于编译通常不经过优化，用反编译技术进行程序验证也是可能的。

用于代码的质量评价中。例如，发现其中结构怎样，或者满足其他要求的各种各样的特征。

为了辅助调试过程，现行的调试器通常都将机器代码反汇编到较高级的汇编语言表示，这有助于获得结构化的高级程序形式的更高级的代码表示，反编译能自动搜索代码中的结构，向软件工程师提供最合适的程序，一边帮助他理解代码的操作。

特别是，随着Internet上的可免费软件和部件的急剧增加，为用户提供了有利条件。但是，这些软件和构件的安全性是不可回避的严重问题，特别是为了构造安全重要的软件系统，在这种情况下，反编译系统将会发挥不可替代的重要作用。

上述都是合法的使用，与其他逆向工程工具一样，反编译也面临是否合法使用的问题。使用的合法与否由使用者的动机决定，不由工具本身决定。

反编译的过程通常如下：

目标语言代码→反汇编→语言和编译器版本识别→库函数或类库识别→预处理→中间语言→数据库分析、控制流分析→编译器→高级语言

## 1.2 定义

编译：转换高级语言或目标代码的程序。

反编译：一种试图从机器代码转换回到与源程序相似的某种程序。

## 1.3参考资料

《Android Dalvik.pdf》

《Android\_DEX\_文件格式详解.doc》

《ClassFile.doc.html》

以上资料已上传小组github。

# 2.项目概述

## 2.1Dex2jar简介

dex2jar 是一个能操作Android的dalvik(.dex)文件格式和Java的(.class)的工具集合，包含以下几个功能dex-reader/writer: 用于读写 Dalvik Executable (.dex) 文件格式. 包含一个简单的API(与ASM相似)d2j-dex2jar: 执行dex到class的文件格式转换smali/baksmali: 与smali工具功能一致，但是对中文更友好。

## 2.2基本操作流程

1 将Android apk格式改为zip并解压，得到classes.dex

2 在命令行中执行Dex2jar d classes.dex –o 输出目录可得到jar包

3 最后用jd-gui这种工具可看到jar包中class的源码

## 2.3目前尚不完善的地方及优化方向

1 反编译出的class与实际执行后产生的class有出入。

2 对混淆后的apk反编译不是很成功。

3 对资源文件解码无力。

优化方向

将dex2jar的解码机制完善，使其反编译出的.class与原class一致。

## 2.4项目开发环境

此次项目软硬件开发环境如下：

软件环境

Windows 7/Windows 10/Ubuntu

硬件环境

4G内存 + 320G硬盘

开发工具

Dex2jar、Eclipse 、apktool、jd-gui

## 2.5验收方式与依据

(1)验收方式

项目验收方式采用中期课堂评审和终期交付验收。

(2)验收依据

验收依据“应交付成果”清单进行验收。

# 3.项目成员分工

## 3.1项目成员

王文茹、卢兴海、蒋波

所有成员了解dex,class字节码格式，了解dex2jar源码。

王文茹：优化编译后的class文件的格式，使其与原class一致。

蒋波：完善dex2jar中错误结构的设计，尽可能准确得到错误产生的原因。

卢兴海：对dex2jar的编译过程进行优化，尽可能复原class

一起进行测试分析。

具体分工详见《项目计划v1.mpp》

## 3.2小组合作与沟通方式

小组成员平时通过微信交流，每周组织1到2次小组会议，需详细记录会议时间与内容，微信需保留聊天记录。

# 4.实施计划

## 4.1项目开发过程计划

确定项目计划及人员分工 2017.03.10-2017.03.17

软件需求分析，建模，与验证方法 2017.03.18-2017.03.24

软件需求评审及复评审 2017.03.25-2017.04.7

软件产品设计，改进及展示 2017.04.8-2017.05.05

软件测试 2017.05.06-2017.05.12

软件测试演示及评审 2017.05.13-2017.06.02

分析和总结 2017.06.03-2017.06.09

## 4.2应交付成果

项目应交付的产品如表1所示。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 编号 | 名称 | 交付时间 | 要求 | 备注 |
| 1 | 项目计划书 | 2017/3/17 | 完整、明确 | 暂无 |
| 2 | 项目进度表 | 2017/3/16 | 完整、详细 | 暂无 |
| 3 | 项目工作日志 | 待定 | 暂无 | 暂无 |
| 4 | 软件需求规格说明书 | 待定 | 暂无 | 暂无 |
| 5 | 需求检查单 | 待定 | 暂无 | 暂无 |
| 6 | 测试需求规格说明书 | 待定 | 暂无 | 暂无 |
| 7 | 软件测试分析报告 | 待定 | 暂无 | 暂无 |
| 8 | 软件测试评审报告 | 待定 | 暂无 | 暂无 |
| 9 | 项目工作量统计报告 | 待定 | 暂无 | 暂无 |
| 10 | 项目方法总结报告 | 待定 | 暂无 | 暂无 |
| 11 | 项目个人总结 | 待定 | 暂无 | 暂无 |

表1 应交付成果列表

# 5.关键问题

Dex2jar的优化问题，要将dex2jar的解码机制完善，使其反编译出的.class与原class一致。