**Dex2jar的测试及优化**

**需求规格说明书**

Version 2.0

小组成员：

卢兴海

王文茹

蒋 波

**版本变更历史**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 版本 | 变更时间 | 修改人 | 审核人 | 备注 |
| 1.0 | 2017/03/27 | 蒋波、卢兴海、王文茹 | 蒋波、卢兴海、王文茹 | 初稿 |
| 1.1 | 2017/03/30 | 蒋波、卢兴海、王文茹 | 蒋波、卢兴海、王文茹 | 二稿 |
| 1.2 | 2017/04/06 | 蒋波、卢兴海、王文茹 | 蒋波、卢兴海、王文茹 | 三稿 |
| 1.3 | 2017/04/21 | 蒋波、卢兴海、王文茹 | 蒋波、卢兴海、王文茹 | 四稿 |
| 2.0 | 2017/04/26 | 蒋波、卢兴海、王文茹 | 蒋波、卢兴海、王文茹 | 修改文档格式与总体概述 |

目 录

[1. 引言 1](#_Toc481003912)

[1.1编写目的 1](#_Toc481003913)

[1.2软件需求分析的重要性 1](#_Toc481003914)

[1.3软件需求分析目标 1](#_Toc481003915)

[1.4参考资料 2](#_Toc481003916)

[1.5术语和缩略语 2](#_Toc481003917)

[2. 总体概述 4](#_Toc481003918)

[2.1产品描述 4](#_Toc481003919)

[2.2运行环境 4](#_Toc481003920)

[2.2.1硬件环境 4](#_Toc481003921)

[2.2.2软件环境 4](#_Toc481003922)

[2.3设计和实现的约束条件 5](#_Toc481003923)

[2.4假设和依赖 5](#_Toc481003924)

[3. 功能性需求 6](#_Toc481003925)

[3.1 用例模型 6](#_Toc481003926)

[3.2 dex文件生成 7](#_Toc481003927)

[3.3 IDEA中Build dex2jar 7](#_Toc481003928)

[3.4 转换dex文件为jar包 8](#_Toc481003929)

[3.5 查看代码 10](#_Toc481003930)

[3.6 获取文件资源 10](#_Toc481003931)

[4. 非功能性需求 12](#_Toc481003932)

[4.1版本兼容性 12](#_Toc481003933)

[4.1.1Android平台版本 12](#_Toc481003934)

[4.1.2操作系统版本等 12](#_Toc481003935)

[4.2可用性 12](#_Toc481003936)

[4.3鲁棒性 12](#_Toc481003937)

[4.4可扩展性 12](#_Toc481003938)

[4.5图形界面 12](#_Toc481003939)

[4.6准确性 13](#_Toc481003940)

[5. 业务需求模型 14](#_Toc481003941)

[5.1反编译apk中的源码 14](#_Toc481003942)

[5.2反编译apk中的资源文件 16](#_Toc481003943)

[6. 设计实现 17](#_Toc481003944)

# 引言

## 1.1编写目的

编写《Dex2jar的测试及优化》需求规格说明书，是为了保证软件开发的质量、需求的完整与可追溯性。通过此文档，以保证规范化的软件开发过程，使之成为整个开发工作的基础。

## 1.2软件需求分析的重要性

软件需求分析是研究用户需求得到的东西，完全理解用户对软件需求的完整功能，确认用户软件功能需求，建立可确认的、可验证的一个基本依据。

软件需求分析是一个项目的开端，也是项目实施最重要的关键点。据有关的机构分析结果表明，设计的软件产品存在不完整性、不正确性等问题80％以上是需求分析错误所导致的，而且由于需求分析错误造成根本性的功能问题尤为突出。因此，一个项目的成功软件需求分析是关键的一步。

## 1.3软件需求分析目标

软件需求分析的主要实现目标：

1)对实现软件的功能做全面的描述，帮助用户判断实现功能的正确性、一致性和完整性，促使用户在软件设计启动之前周密地、全面地思考软件需求；

2)了解和描述软件实现所需的全部信息，为软件设计、确认和验证提供一个基准；

3)为软件管理人员进行软件成本计价和编制软件开发计划书提供依据；

需求分析的具体内容可以归纳为六个方面：软件的功能需求，软件与硬件或其他外部系统接口，软件的非功能性需求，软件的反向需求，软件设计和实现上的限制，阅读支持信息。

软件需求分析应尽量提供软件实现功能需求的全部信息，使得软件设计人员和软件测试人员不再需要需求方的接触。这就要求软件需求分析内容应正确、完整、一致和可验证。此外，为保证软件设计质量，便于软件功能的休整和验证，软件需求表达无岔意性，具有可追踪性和可修改性。

## 1.4参考资料

1.《软件工程》  齐治昌 谭庆平 宁洪 高等教育出版社

2.《PowerBuilder》 崔巍 高等教育出版

3.《实用软件工程》第三版 殷人昆 清华大学出版社

4. dex2jar源码

## 1.5术语和缩略语

术语和缩略语表如表1所示。

|  |  |  |  |
| --- | --- | --- | --- |
| 编号 | 术语 | 英文 | 说明 |
| 1 | UCM | UCM | 用例建模 |
| 2 | RUCM | RUCM | 限制性用例模型 |
| 3 | Dalvik | Dalvik | 一段Dalvik汇编代码由一系列Dalvik指令组成，指令语法由指令的位描述与指令格式标识来决定。 |
| 4 | DexFileReader | DexFileReader | 从dex文件中读取内容，并调用里面的相应的函数处理读取到的数据。 |
| 5 | DEX | DalvikVM executes | Android Dalvik执行程序。 |
| 6 | APK | AndroidPackage | Android应用程序安装包。通过将APK文件直接传到Android模拟器或Android手机中执行即可安装应用。 |
| 7 | Jar | Jar | 可简单理解为一些.class文件打包后的结构 |
| 8 | Github | Github | GitHub 是一个面向开源及私有软件项目的托管平台 |
| 9 | IR指令 | IR指令 | 寄存器指令 |
| 10 | IDE | Integrated Development Environment | 集成开发环境 |

表1 术语和缩略语表

# 总体概述

## 2.1产品描述

反编译技术是通过对低级语言代码（二进制代码或者汇编代码等）进行分析转化，得到等价的高级语言（不限制语言类型，本书的描述主要以C语言为例）代码的过程。它涉及指令系统，可执行文件格式，反汇编技术，数据类型分析技术，控制流分析技术和高级代码生成技术等。

反编译的本质是编译的逆过程。从二十世纪五十年代第一个编译器出现开始，将机器码转换成为高级语言的期望就引发了人们广泛的兴趣。编译器同反编译器都将程序从一种形式转换到另外一种形式，而且在转换的步骤中，都使用了类似的中间表示。差别只是在于编译器的总体方向是从源程序到机器码，而反编译器的总体方向则是从机器码到源程序。尽管在整体方向上两者是相反的，但编译器和反编译器往往在分析阶段使用类似的技术，例如数据流分析。

dex2jar这个源码包，是用来将dex文件转换为Java文件,在github上面的地址是https://github.com/pxb1988/dex2ja。但我们分析这个源码可以看到，里面的功能不只是将dex转换为jar包，也可以转换dex为smail文件等，dex2jar就是将dex文件转换为class文件，然后将class文件打成jar包，dex转换为class文件过程中会先转换为一种IR的中间指令格式。

## 2.2运行环境

### 2.2.1硬件环境

CPU：Inter1.0Ghz 及以上

内存：2G

### 2.2.2软件环境

操作系统：Windows 7 / 8.1 / 10 / Vista、Linux

jdk版本：jdk 1.8.0\_121

开发IDE：IntelliJ IDEA 2016.3.4

## 2.3设计和实现的约束条件

无

## 2.4假设和依赖

无

# 功能性需求

## 用例模型

Dex2jar的需求较为单一，主要是实现对apk文件的反编译，获取源代码和资源文件，用例包括生成.dex文件、在IDEA中buildddex2jar、运行dex2jar并获取jar包、用所写工具校验代码、第三方工具查看jar包、获取资源文件，dex2jar用例图如图3.1所示。



图3.1 用例模型

## 3.2 dex文件生成

要进行反编译apk文件，首先要从已获取的apk文件中，生成classes.dex文件，从而用此文件来进行反编译，获取需要的jar包。此步骤很简单，只需要将apk文件的后缀改为zip，解压缩后即可得到需要的classes.dex文件，若更改后缀后名字重复，则取消更改并报错。dex文件生成的RUCM图如图3.2所示。



图3.2 dex文件生成RUCM图

## 3.3 IDEA中Build dex2jar

获取classes.dex文件后，在集成环境IDEA中调用编译dex2jar工具，构建…。该构建过程，如图3.3所示。

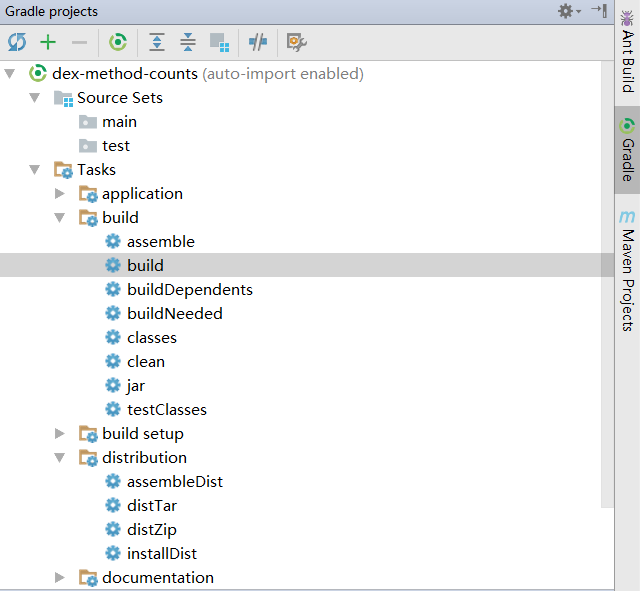


图3.3 Build过程

## 3.4 转换dex文件为jar包

此用例即是将dex文件转换为jar包的步骤，先将生成的dex文件拷贝到dex2jar工具所在的目录下，然后打开命令提示符，使用cd指令进入dex2jar工具所在目录，执行d2j-dex2jar.bat classes.dex命令，若工具正常执行，则会在目录下生成classes-dex2jar.jar文件。转换dex文件为jar包的用例说明参见图3.4所示的RUCM图。生成工具所在目录如图3.5所示。



图3.4 转换dex文件为jar包RUCM图

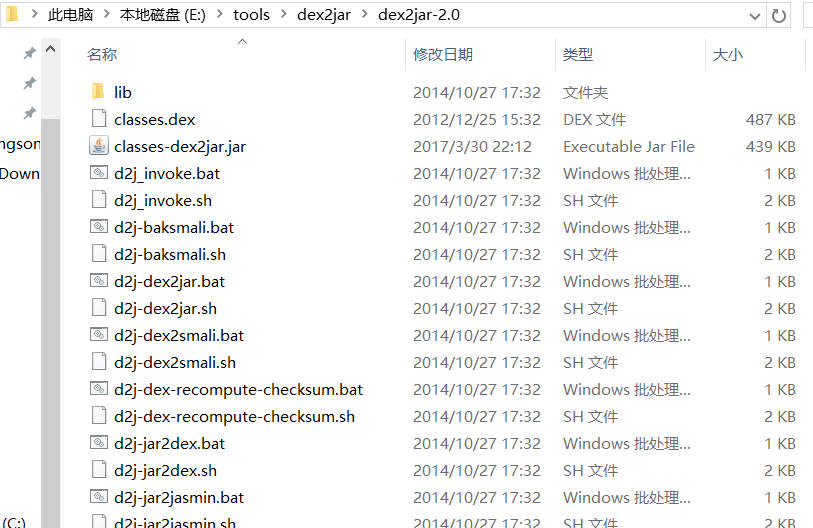


图3.5 生成工具所在目录

## 3.5 查看代码

反编译完成后，可通过JD-GUI工具打开jar文件，并查看其中代码，对比与源码的差距，测试反编译的效果。查看代码的RUCM图如图3.6所示。



图3.6 查看代码RUCM图

## 3.6 获取文件资源

Dex2jar编译主要针对源码，通过可反编译产生出jar包和java文件，但是无法获取其中的图片文件和布局文件等资源文件，我们想在dex2jar原有基础上，增加资源文件获取的功能。获取文件资源的RUCM图如图3.7所示。



图3.7 获取文件资源RUCM图

# 非功能性需求

## 4.1版本兼容性

### 4.1.1Android平台版本

Dex2jar需要实现对所有的apk文件可以正常的反编译，得出原始的jar包，因此要求兼容各个版本的android平台。

### 4.1.2操作系统版本等

Dex2jar要使更多的用户可以方便的使用此工具，至少需兼容主流的windows平台（Windows 7 / 8.1 / 10 / Vista）和linux平台。

## 4.2可用性

dex2jar要实现可用性，即应当使得dex2jar对于不同大小的.apk文件的处理时间均在合理的范围内，在一定时间内可以成功反编译出需求的文件。

## 4.3鲁棒性

Dex2jar需要在反编译过程中保持稳定性，在异常情况下能正确处理反编译的文件。

## 4.4可扩展性

Dax2jar还可新增从jar编译到.java，以及从apk反编译出其中的资源文件的功能。

## 4.5图形界面

目前dax2jar是在命令行运行的，我们想首先在windows操作系统环境下，为它增加图形化的处理界面，使dax2jar的操作更加简单方便，若顺利实现，则继续为其增加linux操作系统环境下的图形界面。具体来说，就是抛弃dex2jar现有的命令行操作模式，转为将apk包拖拽到bat批处理命令上，然后让其自动执行反编译工作。

## 4.6准确性

返编译出jar包中的.class文件显然是经过dex2jar优化后的，那么准确性主要体现在要与原.java产生出的.class文件在运行下效果上一致。

# 业务需求模型

反编译就是把目标代码转为汇编代码的过程，也可以说是把机器语言转换为汇编语言代码、低级转高级的意思，常用于软件破解（例如找到它是如何注册的，从而解出它的注册码或者编写注册机）、外挂技术、病毒分析、逆向工程、软件汉化等领域。

随着安卓市场的不断发展，apk程序的应用也越来越广泛。在不侵犯产权的基础上，若要想对apk程序进行修改以更符合自己的需求，则需要用到反编译工具，来将apk文件反编译为源码，从而进行适当的修改。有许多程序可以进行逆向操作即反编译以求修改，例如Flash的文件生成的SWF文件，可以被反汇编成Flash原码。

## 5.1反编译apk中的源码

使用dex2jar，可以将apk文件反编译为源码，下面介绍下dex2jar反编译源码的过程。

dex2jar就是将dex文件转换为class文件，然后class文件打成jar包，dex转换为class文件过程中会先转换为一种IR的中间指令格式，大概的转换流程如图5.1所示。

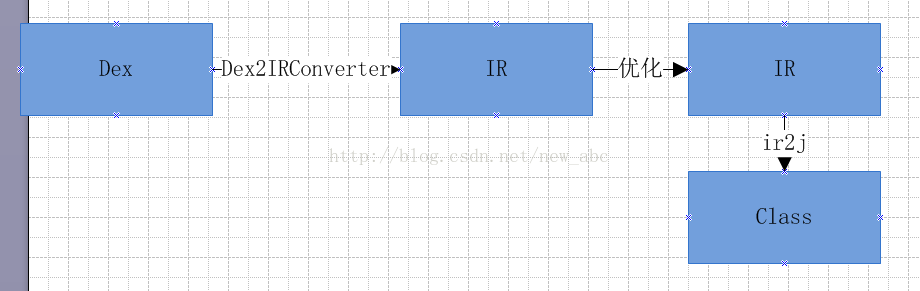


图5.1 dex文件转换流程

中间主要调用的函数接口如图5.2所示。

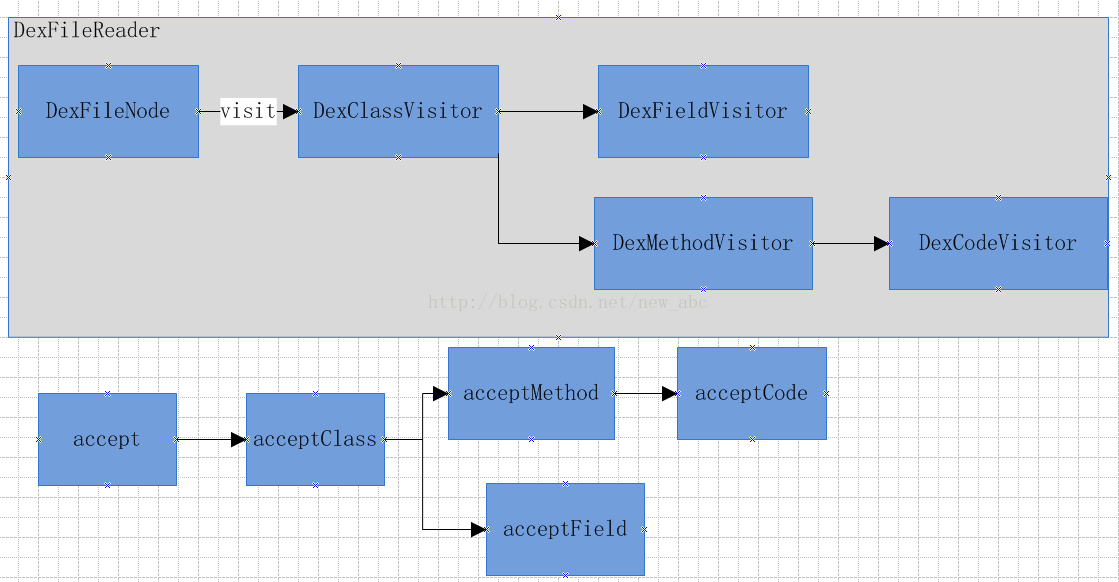


图5.2 调用函数接口

反编译的本质是编译的逆过程，图5.3揭示了编译器同反编译器之间的关系。

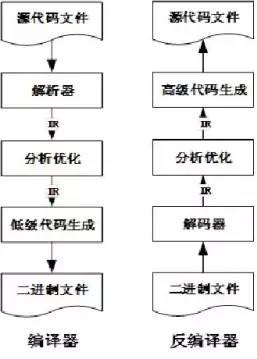


图5.3 编译器与反编译器

可以看到编译器同反编译器都将程序从一种形式转换到另外一种形式，而且在转换的步骤中，都使用了类似的中间表示。差别只是在于编译器的总体方向是从源程序到机器码，而反编译器的总体方向则是从机器码到源程序。尽管在整体方向上两者是相反的，但编译器和反编译器往往在分析阶段使用类似的技术，例如数据流分析。

编译器通过对源代码进行解析得到中间表示（IR，intermediate representation），反编译器通过对指令进行解码得到中间表示。类似的，编译器的低级代码生成同反编译器的高级代码生成恰好对应。

## 5.2反编译apk中的资源文件

Dex2jar的侧重点在于反编译出apk文件的源码，而有时候用户需求的是apk文件中的资源文件，例如图片文件、布局文件等。所以我们想要给dex2jar扩展功能，使其可以获取到apk中的资源文件，满足用户的各种需求。具体来说，就是获得原应用的res文件夹下的图片资源，字符串资源、color资源以及layout资源。

# 设计实现

本小组准备在dex2jar的基础上，为其扩展新功能，主要在两个方面：

1、Dex2jar编译主要针对源码编译，可反编译出jar包和java文件，但是无法获取到图片文件和布局文件等资源文件，我们想在dex2jar原有基础上，增加资源文件获取的功能。

2、目前dax2jar是在命令行运行的，我们想首先在windows操作系统环境下，为它增加图形化的处理界面，使dax2jar的操作更加简单方便，若顺利实现，则继续为其增加在linux操作系统下的图形界面。具体来说，就是抛弃dex2jar现有的命令行操作模式，转为将apk包拖拽到bat批处理命令上，然后让其自动执行反编译工作。