

诈骗终结者

一款基于人工智能大模型
的涉诈APK智能识别系统

小组成员：陈秋羽、罗宇航、梁乐怡

小组名：骗局终结者



目录/CONTENTS

01. 项目背景

02. 作品概述

03. 系统设计

04. 系统实现

04. 界面展示

05. 系统测试



01 项目背景

移动设备 上网普遍度高

在近9亿网民中，手机上网的比例高达

99.1%， 移动互联网服务的便捷性、即时性和普惠性在各类应用程序（APP）中得到了充分体现。

APP功能广泛 不可避免

据估计，移动互联网应用商店推广的APP数量接近

400万款，总下载量超过万亿次，APP在推动经济社会发展、服务民生等方面发挥着至关重要的作用。

APP诈骗 案件高发

近年来，电信网络诈骗的作案手法从电话、短信转向利用APP等网络工具，约占整体案发量的

70%。

APP涉诈行为 普遍存在

同时，APP涉诈、涉赌、涉黄以及强制授权、过度索权、超范围收集个人信息的现象普遍存在。

静态分析

静态分析是指**通过反编译来获取控制软件程序的源代码，从代码中分析程序的运行过程**，了解模块中执行命令的一些功能；获取到软件名称、包名等基本信息；了解接入哪些SDK，这些SDK作用；了解到是使用什么语言进行编译或使用什么编译器进行编译的，了解程序是否受到加壳保护。

动态分析

APK动态分析是指**对安卓应用程序（APK）在运行时的行为进行监控和分析**，以发现其中可能存在的漏洞、恶意行为或其他安全问题。

竞品分析

相比市面上的APK安全分析产品，诈骗终结者不仅具有APK反编译等基本功能，还具有**威胁情报、行为异常分析以及网络请求提取**等高级功能，支持恶意代码分析。

支持**静态分析、动态分析及黑白名单管理、人工智能研判以及基于LLM大模型的智能分析**，自动化程度高，功能集成全，普通民众也能轻松分析违法APK！

工具名称	APK反编译	威胁情报检测	网络请求提取	恶意代码分析
诈骗终结者	支持	威胁情报、行为异常	提取DNS、会话信息、HTTP	支持
大狗	支持	匹配已知样本	提取网站请求与响应数据	支持
奇安信-情报沙箱	支持	威胁情报、行为异常	提取DNS、会话信息、HTTP	支持
360-沙箱云	支持	威胁指标	不支持	支持
微步云沙箱	支持	行为检测	不支持	支持
摸瓜	支持	不支持	不支持	不支持



数据来源：元芳科普 | APK取证简单分析及常用APK分析工具对比

02 作品概述

一款集成静态分析、动态分析及黑白名单管理、人工智能研判以及基于LLM大模型的智能分析多种功能的涉诈APP分析系统。

01

技术概览--APK分析&人工智能

综合APK解析技术、APK沙箱分析、机器学习模型、LLMs模型技术构建的先进工具。

02

功能定位--涉诈APK自动化分析工具

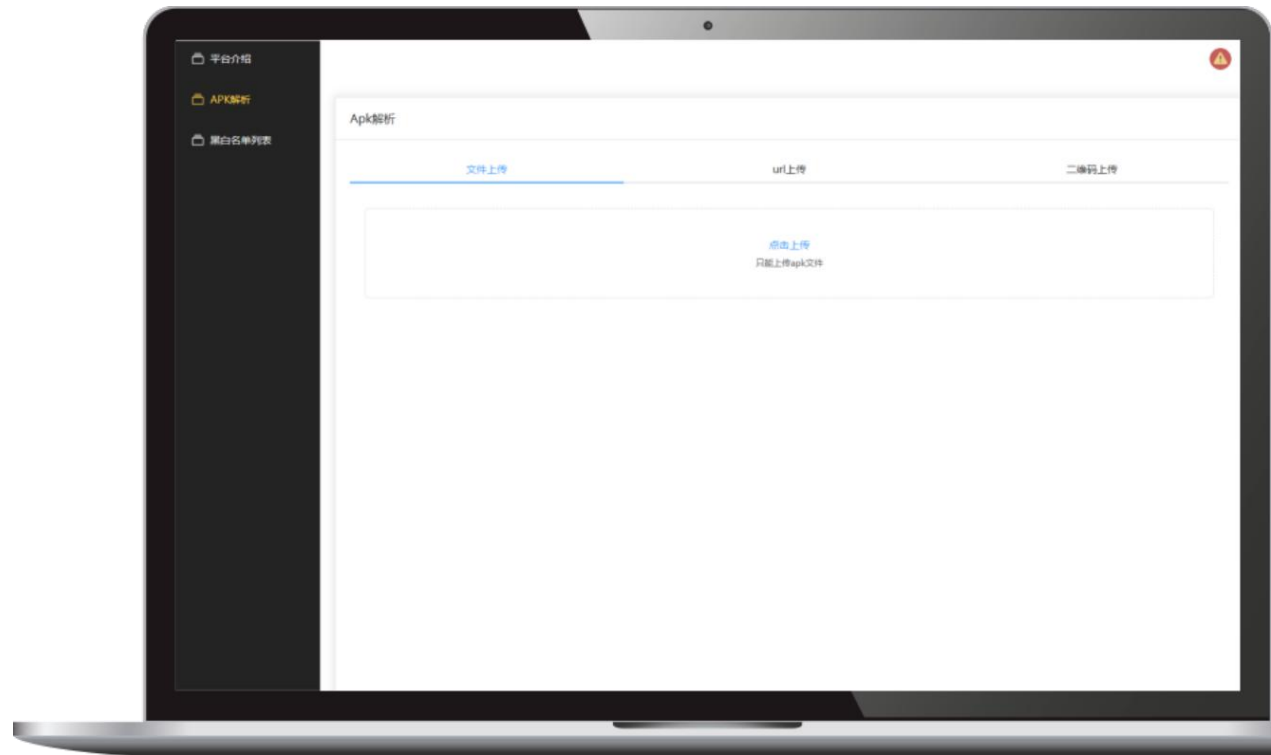
为**业余用户**、**专业用户**、**警方**设计的涉诈APK自动化分析工具，普通民众也能轻松分析违法APK,为诈骗取证和侦查提供有效帮助。

03

功能覆盖--多种实用功能

功能包括**多种APK采集方式**、**静动态分析相结合**、**快速识别涉诈站点及URL**。

核心优势在于**基于XGBoost算法构建的涉诈判别模型**以及**基于LLMs大模型技术实现的智能分析模型**。



多模式APP采集：支持互联网和离线模式下的APP采集，包括**基于链接、二维码和网页按钮**的下载方式，以及直接上传APK安装包进行分析。

01

黑白名单库管理：实现APP黑名单和白名单的动态配置与过滤，以自动化和优化安全分析流程。

02

基本信息提取：自动提取APP的基本信息，包括应用程序名称、版本号、文件大小和MD5等关键指标。

03

静态与动态解析及关键诈骗站点提取：对APP进行静态和动态解析，提取APP中的关键诈骗站点信息，为进一步的安全分析提供数据支持。

04

诈骗终结者

涉诈APK
智能识别系统

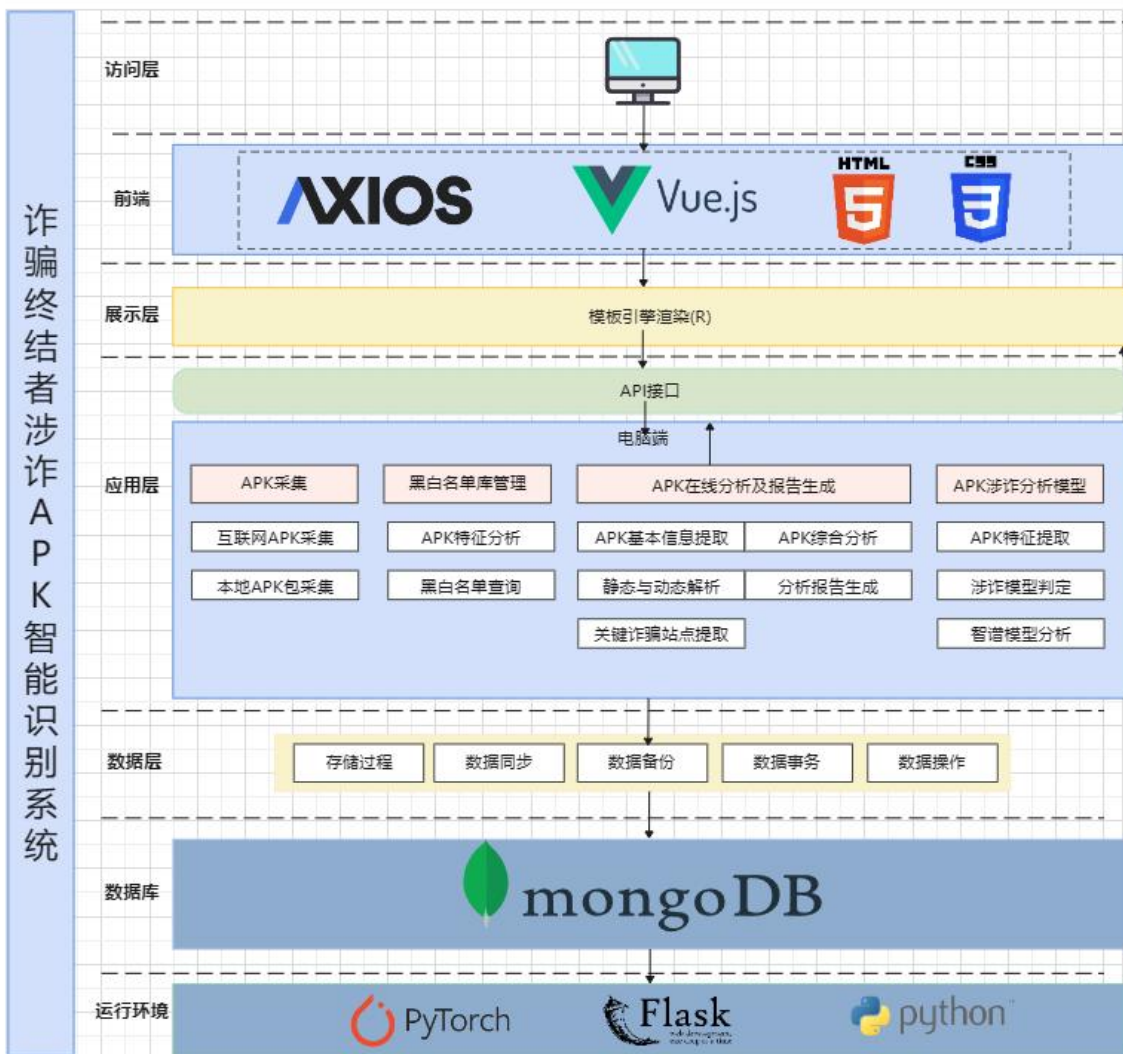
XGBoost涉诈研判模型：将APK特征进行清洗整理并输入到训练好的XGBoost模型中进行涉诈APK二分类

06

LLMs大模型智能分析：对APK解析后生成的报告进行解析，对报告中的信息及条目一一进行分析，帮助用户更好辨别涉诈APK。

05

03 系统设计



架构设计

分为前后端，前端负责用户与界面进行交互，后端负责处理对应的请求。

客户端设计

1. 上传APK文件
2. 查看分析报告
3. 获取研判结果
4. 查看智能分析结果
5. 搜索黑白名单

服务器端设计

1. 分析APK
2. 生成分析报告
3. 计算研判结果
4. 生成智能分析
5. 处理黑白名单搜索请求



数据库

MongoDB非关系型数据库，存储简洁高效



集合

reports集合用于存储详细的报告数据。
list集合用于存储黑白名单的基础条目。

reports

```
_id: objectId
SHA1: string
activities: string[]
application_name: string
architecture: object
  arm64-v8a: bool
  armeabi: bool
  armeabi-v7a: bool
  x86: bool
  x86_64: bool
md5: string
package_name: string
permissions: string[]
static_analysis: object
  data: object
    apkid_metadata []
    basic_info: object
      crc32: string
      detect_et: string
      file_tags: string[]
89 more items...
```

list

```
_id: objectId
packageName: string
apkName: string
md5: string
result: string
```


序号	接口名	请求方法	请求路径	请求参数	返回数据示例
1	文件上传接口	POST	/files/upload	文件 (file)	<pre>{"id": "AZCIHlwQONZSmF3-yCZm"}</pre>
2	应用信息获取接口	GET	/reports/get	id	详见下文
3	名单搜索接口	GET	/lists/search	value, type, md5	name
4	获取白名单接口	GET	/lists/whitelist	无	详见下文



前后端通过接口进行交互



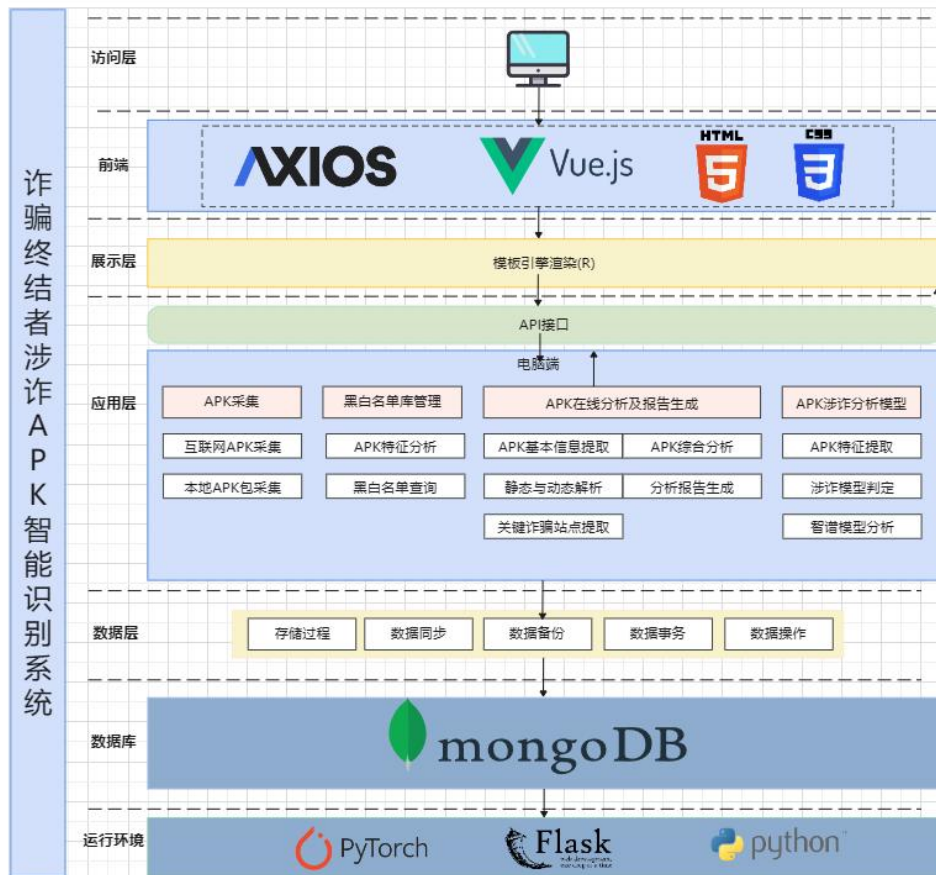
接口遵循Restful设计风格



接口覆盖各个功能：

- 文件上传接口、链接上传apk接口、二维码上传apk接口。
- 应用信息获取接口。
- 名单搜索接口、获取白名单接口、获取黑名单接口、添加到名单接口、删除名单接口

04 系统实现



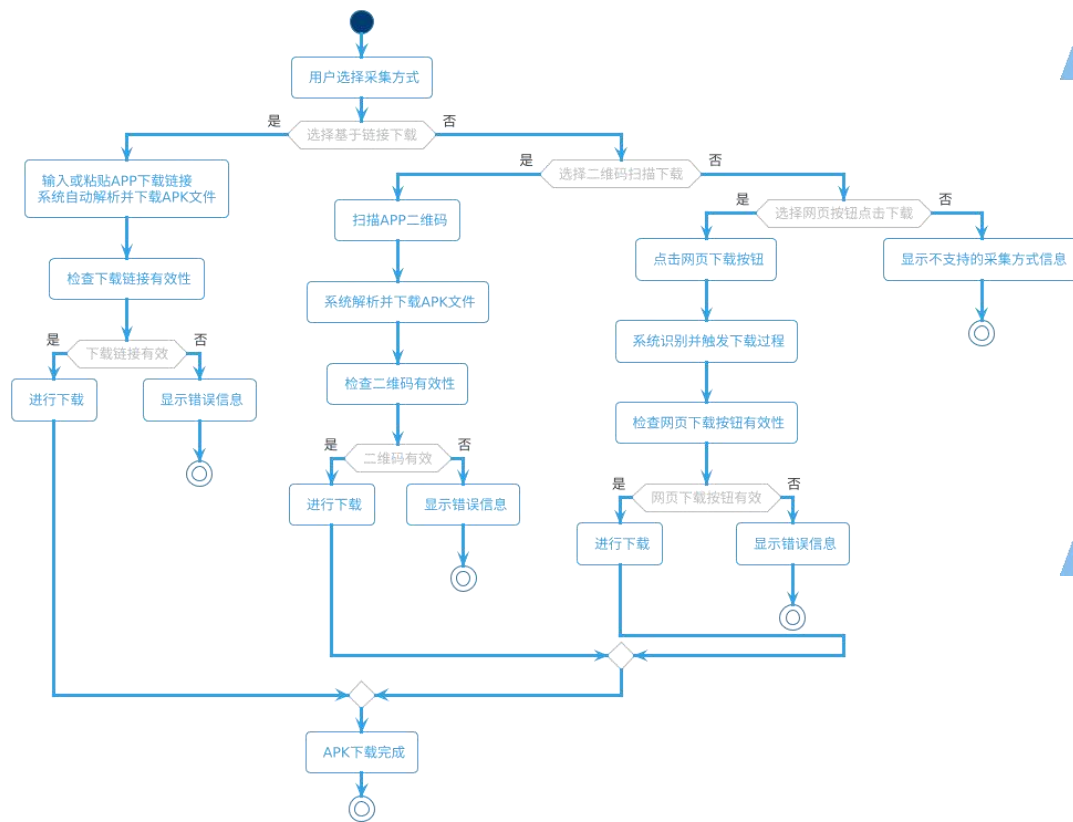
诈骗终结者-前端

- html + CSS + Javascript
- ES6: HTML5语言
- Vue: 渐进式框架
- Vue Router: VUE路由插件
- Vuex: VUE状态管理
- Axios: 一个基于 promise 的 HTTP 库, 用于 GET/POST 请求
- Node.js+webpack: 项目构建工具



诈骗终结者-后端

- Flask: 项目基础框架
- Pytorch: 模型训练框架
- MongoDB: 非关系型数据库服务
- Androguard: APK反编译工具
- 智谱清言: LLMs大模型

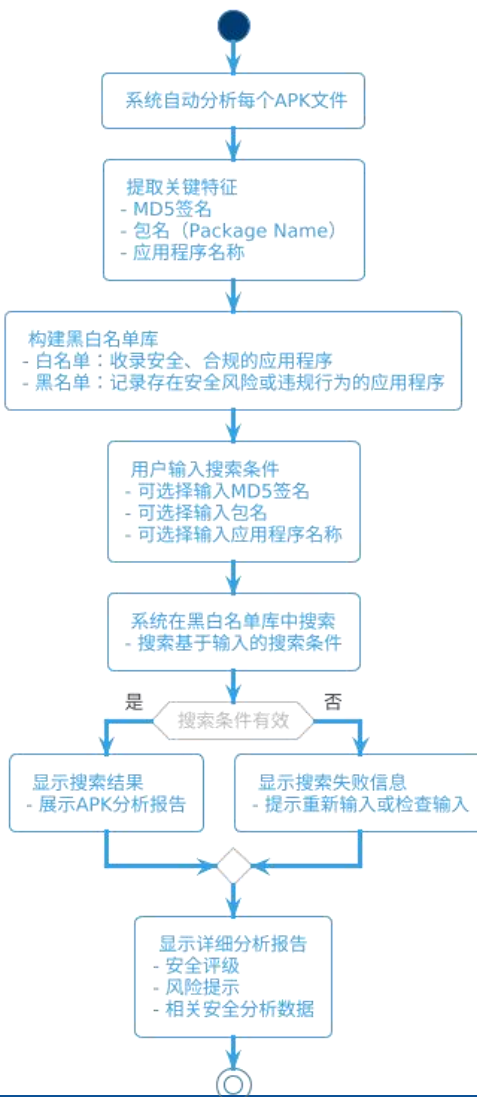


前端

实现**二维码扫描、URL输入和文件选择**界面。

后端

处理文件接收、验证、存储和安全检查，确保数据安全和一致性。
提供API接口供前端调用，完成数据传输和处理。



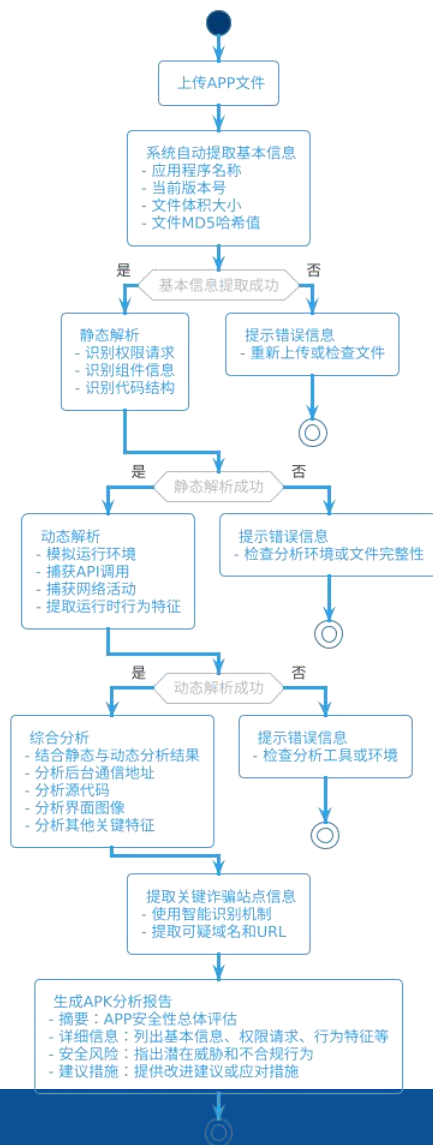
前端模块

- 设计直观的用户界面，允许用户通过**MD5、包名或APK名称**进行查询。
- 展示清晰的搜索结果列表，包括APK的基本信息和状态（如是否在黑名单中）。

后端模块

- 负责存储和管理黑白名单数据库，确保数据的安全性和准确性。
- 实现**基于MD5、包名或APK名称的模糊搜索算法**，快速响应用户查询请求，返回相关的APK记录和其黑白名单状态。

APK分析及报告生成开发



前端

设计报告展示界面，提供清晰直观的数据分析结果。

后端

- 接受APK包、对APK进行反编译。
- 调用奇安信/API接口获取**APK的基本信息、静态和动态解析，生成详细报告、涉诈URL及站点。**
- 利用**智谱大模型对报告进行深度解读。**

数据集与 数据清洗

01

摸瓜平台爬取获取白名单
APK，通过CICMaIDroid
数据集扩充涉诈APK，提
高模型鲁棒性。

特征选择

02

基于文献研究基于APK获取的
权限作为特征进行分析，通
Androguard获取APK对应权
限特征构造数据集。

涉诈研判模型

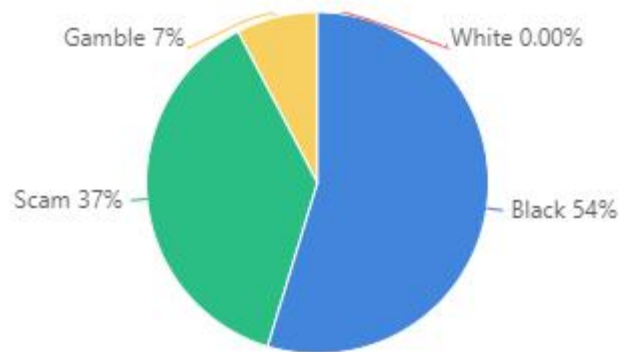
03

基于APK获取的特征训练
XGBoost判别模型，二
分类准确率高达0.976。

模型部署及调用

04

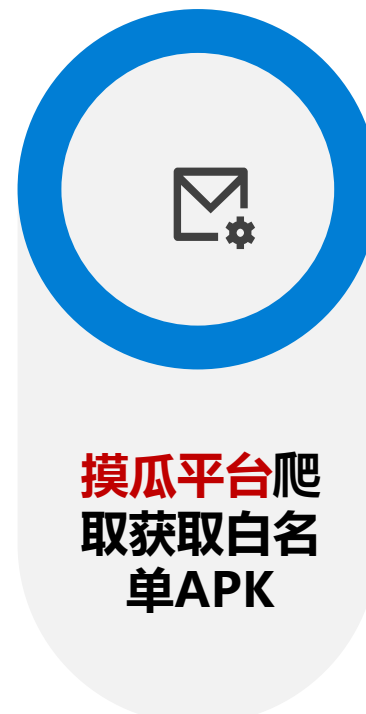
通过Pytorch框架导出模型
文件，并基于Flask框架进
行模型部署调用。



赛题提供的APK数据量较少，数据量不足以进行正常的模型构建。**需要进行数据扩充。**



- 1、最新安卓恶意软件数据集
- 2、收集了超过 **17,341** 个 Android 样本。
- 3、跨越**五个不同类别**。
- 4提高模型的准确性、鲁棒性。

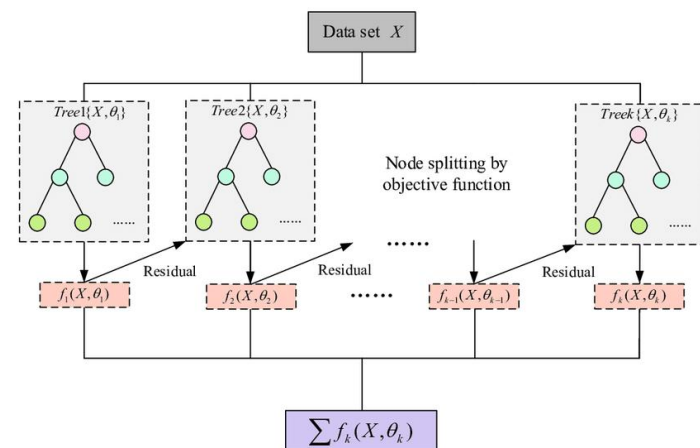


- 1、专业的移动应用分析工具
- 2、**集成分析工具、数据接口**
- 3、高效收集白名单APK

基于**APK获取的权限**作为特征进行分析，通过**Androguard**获取APK对应权限特征构造数据集。

特征	含义
Class	APK对应的类别
android.permission.INTERNET	允许应用程序访问网络
android.permission.WRITE_EXTERNAL_STORAGE	允许应用程序写入外部存储（如SD卡）
android.permission.READ_EXTERNAL_STORAGE	允许应用程序读取外部存储（如SD卡）
android.permission.CAMERA	允许应用程序使用手机的摄像头
android.permission.ACCESS_NETWORK_STATE	允许应用程序访问网络状态信息，例如是否有网络连接

模型训练



- 基于APK获取的特征训练**XGBoost判别模型**
- 模型在十折交叉验证集上二分类准确率达**0.976**。

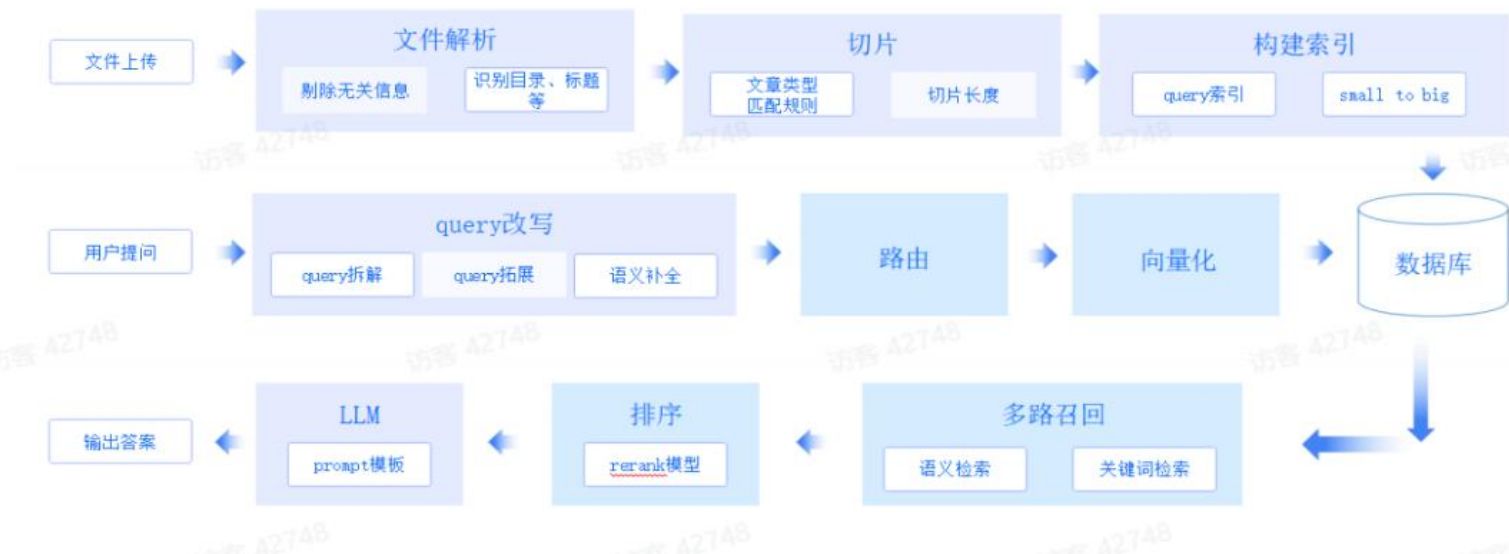
研判

涉诈APK

正常APK



- GLM-4: 智谱AI推出的新一代基座大模型，整体性能大幅提升，接近GPT-4。
- 知识库挂载RAG：通过从大规模的知识库中检索相关信息，并将其与生成模型相结合，生成更准确、更丰富的文本输出。



LLMs报告智能分析模型开发



05 界面展示

平台介绍
APK解析
黑白名单列表

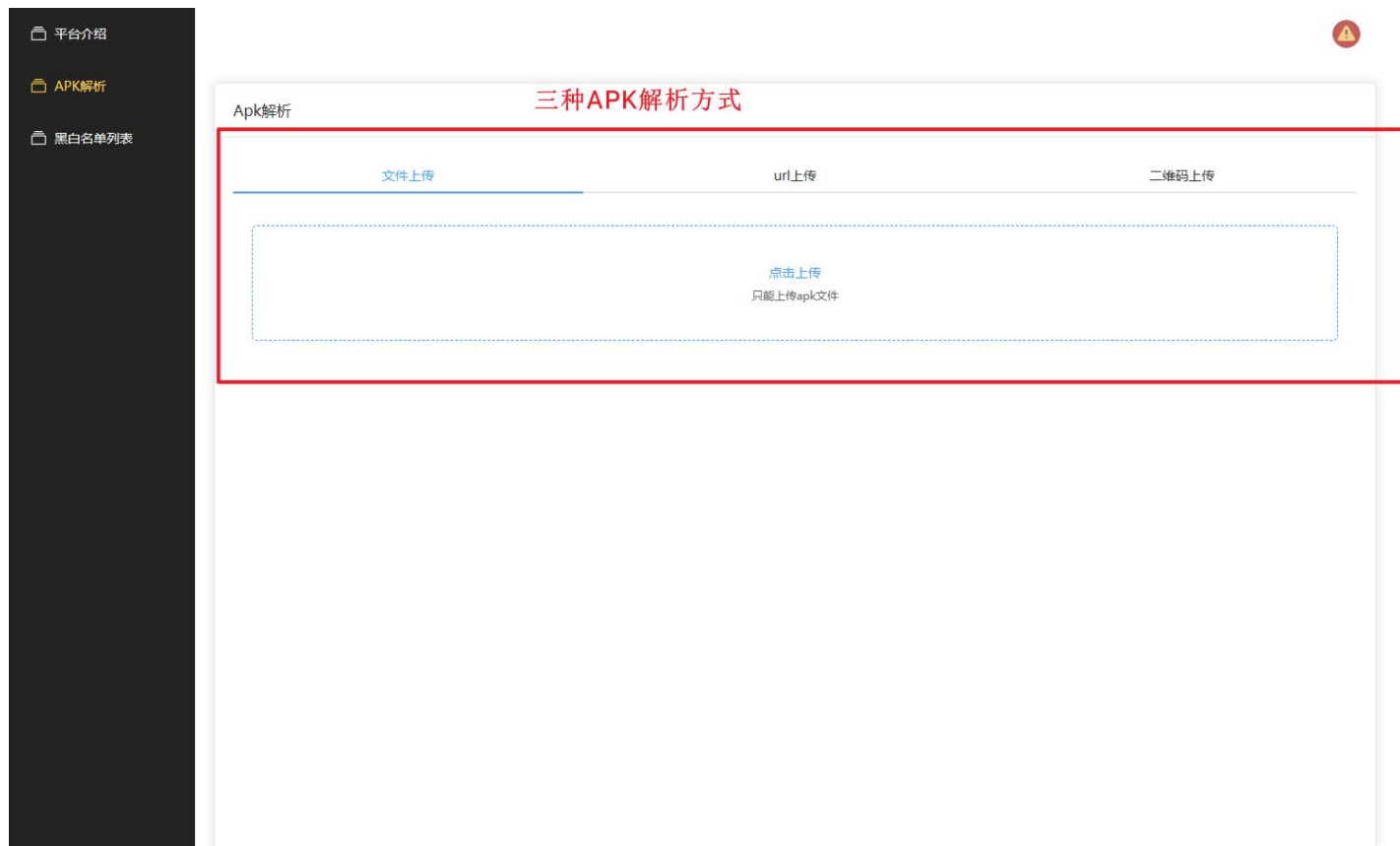
可以选择搜索类别（md5、下载包名、软件名称）

黑白名单列表

搜索类别: app名称 搜索: 蚂蚁 搜索 重置

下载包名	软件名称	可能的结果md5	apkid	类别
ant0531	蚂蚁加速器	912d306c6fe3efb5728 5505af9f2a654	6690e60207c4229a71c95846	black
com.antfortune.wealth	蚂蚁财富	e927ef30d412cd09af8 19ab9dd273676	6690e60207c4229a71c967d0	zj_white
com.antfortune.wealth	蚂蚁财富	59dcb97671e3e8cc5c8 d12dccc80a374	6690e60207c4229a71c9694e	zj_white
com.antfortune.wealth	蚂蚁财富	51634c87e2a2e7d5588 4c197ede6174d	6690e60207c4229a71c973cd	zj_white
com.antfortune.wealth	蚂蚁财富	b9f722ac80f00394779 e221bcca2b308	6690e60207c4229a71c9770a	zj_white
com.antfortune.wealth	蚂蚁财富	f970259c8dab9274f6a 7ad5602aedfbf	6690e60207c4229a71c9800a	zj_white
com.sinyee.babybus.ant	宝宝认蚂蚁	11ba0c6e78de29d2c1a dd5668525fd88	6690e60207c4229a71c985cf	zj_white
com.ants.idlegame.xsl	蚂蚁殖民地	a334ad2aaa592d019f8 f94ce93adfc08	6690e60207c4229a71c9892a	zj_white
com.Company.ants.xsl	蚂蚁世界模拟器	81320f82b494d5cd3e7 42662e1d50dcc	6690e60207c4229a71c991af	zj_white
com.FinalNorth.FinallyAnts.mt	最后的蚂蚁	56496fd2d09e0ef056a 32c7cad99f40a	6690e60207c4229a71c9930f	zj_white
com.lisdw.maYIMng.vivo	蚂蚁世界模拟器	1776b70f8606374b507	6690e60207c4229a71c994a8	zj_white

- 黑白名单管理界面可以根据md5、包名、名称搜索对应的APK。
- 搜索后展示APK的基本信息及黑白名单类别。



- 在APK上传界面可以通过点击三个按钮选择不同的上传方式。
- 通过拖拽APK或者URL或者二维码的方式上传APK。
- 上传时会显示选择的按钮表示上传中。
- 上传成功后会显示分析报告。

平台介绍

APK解析

黑白名单列表

诈骗终结者——涉诈APK智能识别系统

分析结果

AI检测结果

根据提供的信息，该软件“在线娱乐”（com.pro.stocktradeandroid.zxyl）的MD5和SHA1值表明它是一个安卓应用程序。以下是对该软件的分析：

- 权限：应用程序请求了互联网权限（android.permission.INTERNET），这是许多应用程序为了正常工作而必需的，但同时也可能用于收集用户数据或进行远程操作。
- 架构：该应用程序不依赖于特定的CPU架构，这意味着它可以在大多数安卓设备上运行。
- 分析信息：静态分析和威胁分析均显示无异常，没有发现行为异常或潜在威胁。

基于以上信息，没有明显的迹象表明该软件有害。然而，以下几点需要注意：

- 虽然没有发现异常行为，但应用程序的权限可能涉及敏感数据访问，用户应确保该应用来自可信来源。
- 应用程序的行为可能会随着版本更新而变化，因此建议用户在安装后保持对应用程序的监控。
- 应用程序的详细行为（如数据收集和隐私政策）没有在提供的信息中说明，用户在安装前应查阅相关信息。

综上所述，根据当前提供的信息，没有理由认为该软件有害。但用户仍应保持警惕，并确保其来源和目的符合个人安全标准。

该Apk分析得出

apkid	AZDJHtajONZSmf3-yOuN	应用名称	在线娱乐
安装包名	com.pro.stocktradeandroid.zxyl	md5	53f8860bffa3b3f9370040bd10022bad
版本号	102	版本名称	1.02
目标SDK版本号	29		
SHA1指纹	5F 44 D5 09 3C 6D A9 C7 43 00 97 67 6B 9E F6 5D 04 09 84 B5		
活动列表	com.pro.stocktradeandroid.MainActivity		
权限列表	android.permission.INTERNET		

- 在报告分析页面可以看到APK分析的所有结果。
- 包括**AI检测结果**、APP基本信息、APK静态解析结果、APK动态解析结果、**涉诈模型分析结果**等等。

静态分析结果

静态分析	
评分	6
SHA1哈希值	7d06bd1e64f6b9a021d200ff883a88bfae86e016
SHA256哈希值	a1a2628a1779f85c9a4937381f3766fc99b7c2c4e115ec2d84650dc564b67944
SHA512哈希值	4f922e1f1a688c42374206b6a8f8bb4c7d5ee39e5bfafb52eb69a75ffa616647790bde4d6fc2b840e1d62fc5c9619237964432a8c65bc3f
大小	3430789
ssdeep	98304J8ZxSBCzpbQeKzY11/4U8N4PU7qDpZjfrRvN:52kzDKzY11/4ZN4PUB
类型	Zip archive data, at least v?[0] to extract
域名	["schemas.android.com"]
所涉及到的邮箱	[]
IP	[]
URL	["http://schemas.android.com/apk/res/android"]
主要活动	com.it.app.MainActivity
	android.permission.ACCESS_NETWORK_STATE

- 在报告分析页面可以看到APK分析的所有结果。
- 包括**AI检测结果**、APP基本信息、APK静态解析结果、APK动态解析结果、**涉诈模型分析结果**等等。

名称	Name	严重程度
应用请求危险的权限 (基于静态分析)	android_dangerous_permissions	3
应用查询设备信息	android_queried_device_information	3
应用执行动态加载的代码	android_dynamic_code	3
文件操作	android_file_operation	3
请求internet操作	android_internet	3
启动时注册广播接收者	android_registered_receiver_runtime	3
触摸操作	android_touch_action	3
检测加解密操作	android_encrypt_and_decrypt	0
加载dex文件操作	android_load_dex	0
加载so文件操作	android_load_so	0
设置可见不可见展示	android_set_view	0
对话框窗体显示或者隐藏	android_show_hide_dialog	0
显示操作	android_view_operation	0

敏感权限列表

敏感权限	android.permission.CAMERA
	android.permission.READ_EXTERNAL_STORAGE
	android.permission.WRITE_EXTERNAL_STORAGE

- 在报告分析页面可以看到APK分析的所有结果。
- 包括**AI检测结果**、APP基本信息、APK静态解析结果、APK动态解析结果、**涉诈模型分析结果**等等。

ip	122.51.36.234	1
ip	122.51.53.230	1
ip	124.221.170.221	1
ip	124.222.185.242	1
ip	124.222.248.183	1
ip	172.217.163.46	1
ip	18.238.192.30	1
ip	18.238.192.4	1
ip	18.238.192.50	1
ip	18.238.192.90	1
ip	202.112.29.82	1
ip	23.63.242.91	1
ip	23.63.243.99	1
ip	39.100.76.72	1
ip	42.192.113.159	1
ip	49.235.183.51	1
domain	a1951.w16.akamai.net	1
domain	cbec-file.wo-shop.net	1
domain	cdn.dcloud.net.cn	1
domain	cfshopee.com.my	1

- 在报告分析页面可以看到APK分析的所有结果。
- 包括**AI检测结果**、APP基本信息、APK静态解析结果、APK动态解析结果、**涉诈模型分析结果**等等。

06 系统测试

本次测试环境



- (1) 操作系统: Ubuntu x64
- (2) CPU: : Intel(R) Xeon(R) CPU E5-2680 v3 @ 2.50GHz
- (3) 内存: 512M 及以上
- (4) 硬盘空间: 40GB 及以上
- (5) 服务器: 阿里云服务器



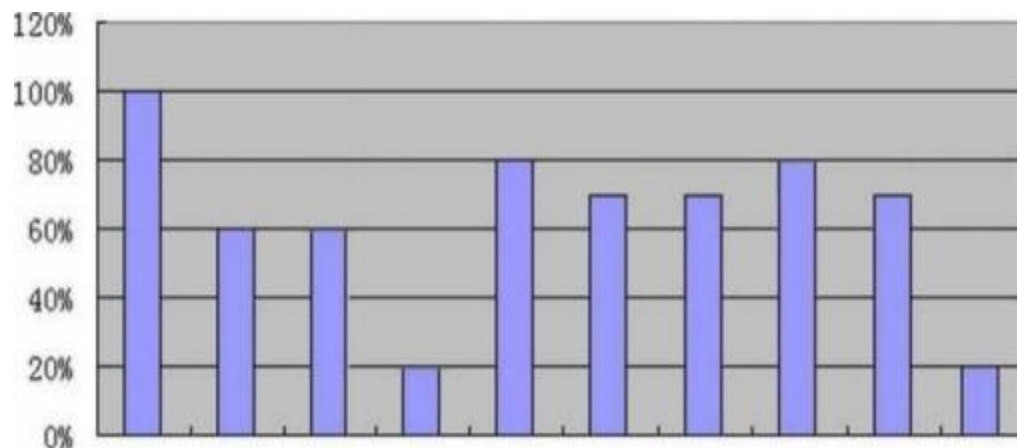
- (1) RAM 512 MB 及以上
- (2) 操作系统: Windows7及以上
- (3) 浏览器: Chrome 浏览器

- 功能测试: 对常规功能进行逐个测试并撰写测试报告。
- 白盒测试: 根据代码逻辑编写完善的测试用例并测试。
- 接口测试: 使用Postman或任何支持HTTP请求的工具执行测试用例。

功能测试及接口测试基本通过。

测试覆盖率基本符合测试标准，在基本功能上可以保证测试的有效性和正确性。本次测试的各指标覆盖如图所示。

**更多测试细节
请在测试文档内查看**



请各位评审老师批评指正

汇报人：陈秋羽

汇报时间：2024-7-19