中国科学院软件研究所学术年会’2019
暨计算机科学国家重点实验室开放周

学术论文

# 安卓应用中误暴露组件的分析与检测

**Characterizing and Identifying Misexposed Activities in Android Apps (ASE 2018: 691-701)**

燕季薇，邓茜，王平，吴添勇，严俊，张健
软件工程技术研究开发中心，计算机科学国家重点实验室
Email: yanjw@ios.ac.cn Tel: 13261537532

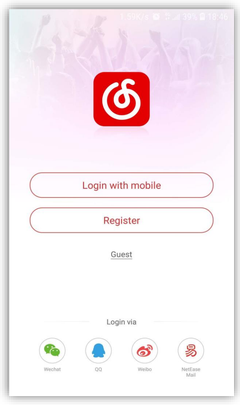
## 背景

暴露活动（EA）是安卓应用程序中可被其他应用中组件直接调用的活动，是实现多个应用程序之间交互和协作的重要的组件间通信机制。 每个暴露活动都是安卓应用的一个隐藏入口，恶意或畸形的调用将带来潜在的安全隐患，如权限泄露攻击等。开发人员应根据组件暴露的必要性谨慎使用这类组件，避免提供不必要的程序入口。然而，我们的研究表明，在商业应用中，暴露活动被开发者广泛使用，活动的误暴露现象也大量存在。在这个工作中，我们首次发现并提出了安卓组件的误暴露问题，提取出多种典型误暴露模式，并实现轻量级工具对大规模安卓应用进行检测。实验表明，暴露活动的误用问题在真实应用中广泛存在！

## 正常的EA

正常的暴露活动用于提供交互性功能，如共享，显示，登录，注册或支付。

Share & Display ✔

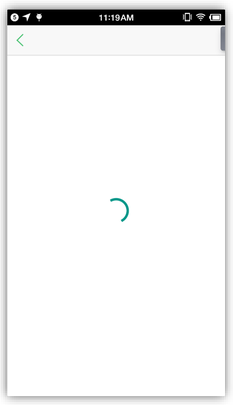


Login & Register ✔

Payment ✔

## 误暴露的EA

误暴露活动常提供无效页面，调试，或提供不完整的内部功能等。



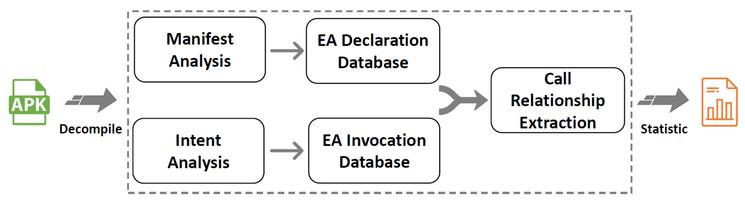
Invalid Page ✘

Debugging ✘

Internal Interface ✘

## 特征分析

- 收集了13,873 个真实Android 应用
- 静态代码分析 + 安卓文档分析 + 数据集特征对比分析 + 反编译代码人工审查
- 提取了6类典型的活动误暴露模式




## 模式提取

❖ **开发人员不知情的误暴露**
- P1：异常暴露比例
- P2：复制粘贴模式
- P3：不匹配的action和data
- P4：不正确的category设置
- P5：隐式内部调用意图

❖ **开发人员知情的误暴露**
- P6：未隐藏的调试界面


## 误用识别

❑ 根据暴露活动特征总结出10种启发式分类规则

❑ 基于提取的规则设计和实现了工具**Mist** (**MIS**exposure iden**T**ification for Android)，工具下载链接：
https://github.com/AndroidMist/Mist

**数据集&实验结果:**

- **Dataset AL:** 收集的全部应用程序，包括来自三个应用市场的13,873个安卓程序。
- **Dataset AR:** 暴露活动异常率最高的前50个应用程序，其活动更可能被错误暴露。
- **Dataset MD:** 最热门的50个应用程序，其EA更有可能被正确使用。

误暴露活动在不同数据集上的分布

暴露活动的误用在真实应用中广泛存在！