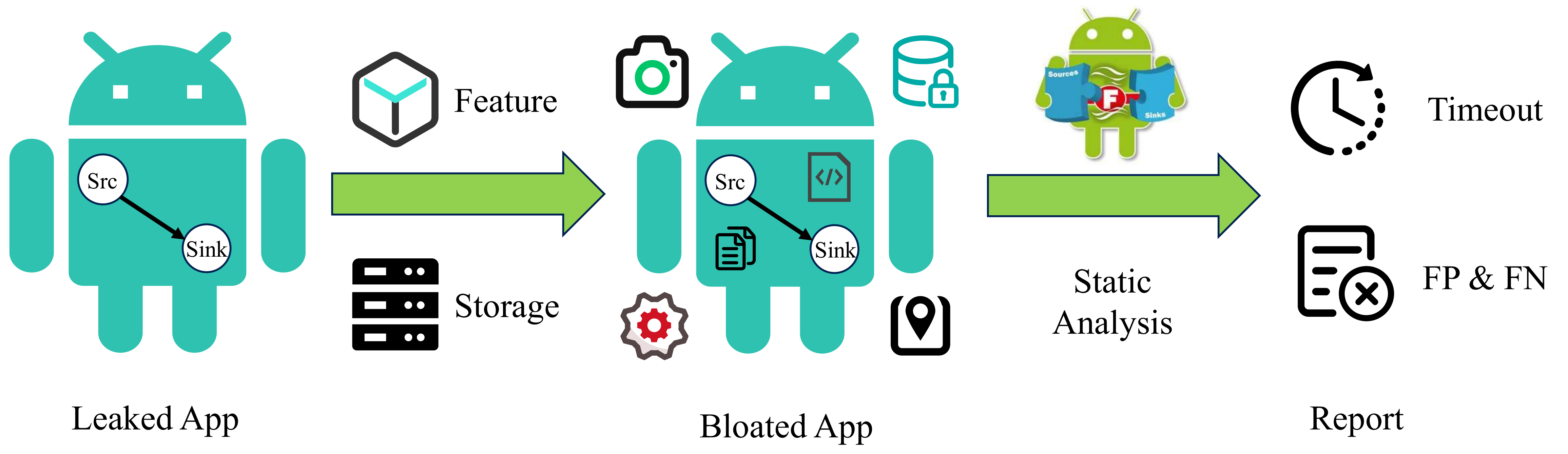
ISCAS 中国科学院软件研究所学术年会 2025 第十届
暨重点实验室科技活动周
学术论文

An Efficient Android App Debloating Approach Based on Multi-layer Dependence Graph
一种基于多层依赖图的高效安卓应用降肿方法

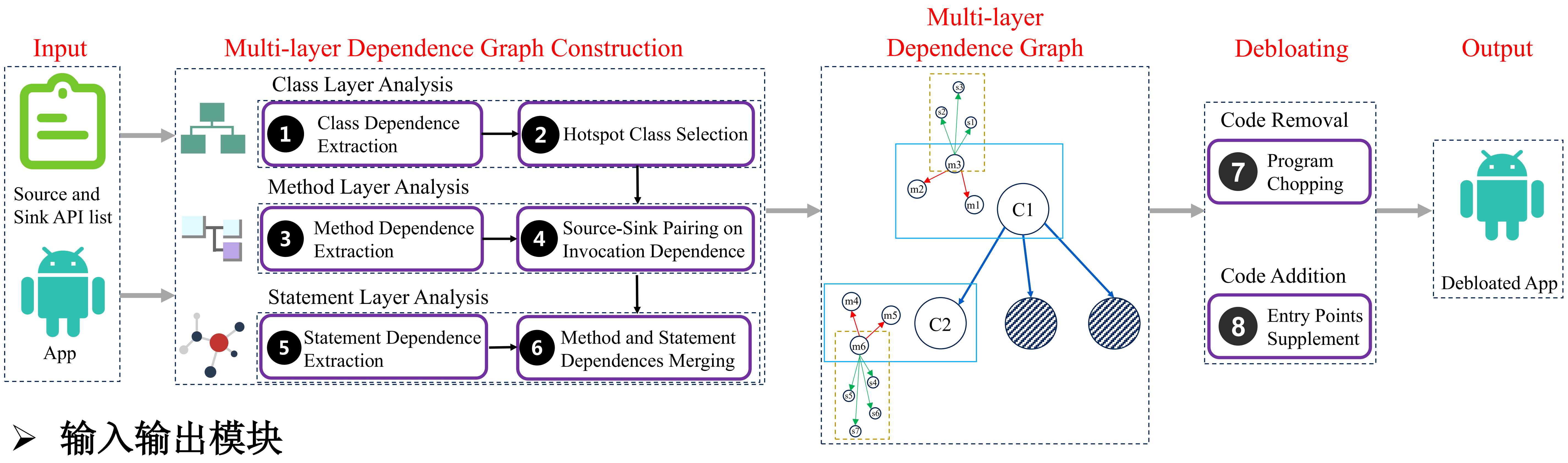
杨恒钦，燕季薇，严俊，梁彬，张健
The International Conference on Software Maintenance and Evolution (ICSME'25)
基础软件与系统重点实验室 · 软件工程技术研究开发中心
联系人：杨恒钦，燕季薇　联系方式：{yanghq, yanjw}@ios.ac.cn

研究背景

- **软件膨胀现象及其对静态分析工具带来的挑战**
  - 随着功能模块不断增加，软件程序占用更多内存、磁盘空间、CPU等资源的现象，被称为软件膨胀。
  - 待分析应用的膨胀现象会为静态分析工具带来分析效率减低、误报漏报数量上升等负面影响，导致静态分析工具实用性降低。
- **面向Source-Sink流的安卓应用切分场景**
  - 以静态污点分析工具为例，工具检测应用中是否存在从Source(隐私数据收集点)到Sink(隐私数据泄露点)的数据流，进而判断待分析应用中是否存在Leak。基于这类数据流求解问题，我们提出了面向Source-Sink流的安卓应用切分场景，旨在通过切除与污点传播无关的代码元素，加速静态污点分析。
  - 现有的安卓应用降肿工具和相关技术在面对这样的切分场景时，往往由于缺失对数据流依赖的精确建模或难以处理复杂的真实应用而无法得到可分析的降肿应用。



降肿流程设计

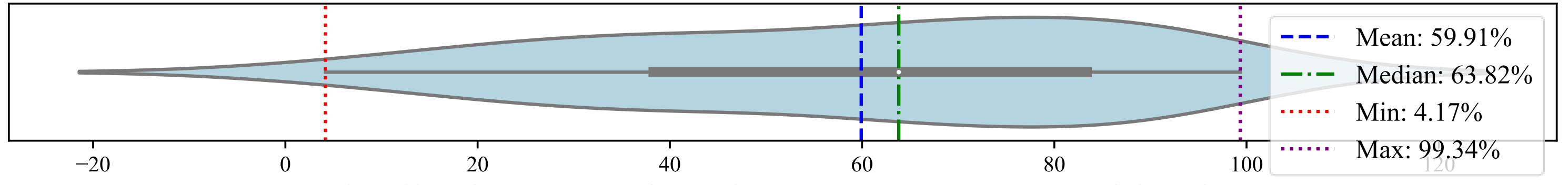


- **输入输出模块**
  - 输入：安卓 App 和污点分析配置的Source-Sink API 列表。
  - 输出：降肿后的安卓App。
- **多层依赖图构建模块**
  - 根据类层次、方法层次、语句层次的依赖关系构建多层依赖图(MDG)，对应用进行建模和表示。
  - 类依赖提取模块：动态识别热点类(Hotspot Classes)，缩小提取依赖项的范围，缓解依赖图规模限制。
  - 方法依赖提取模块：执行额外的源-汇配对分析以排除孤立的源点和汇点，提升分析效率。
  - 语句依赖提取模块：合并提取到的方法-语句层依赖，汇总得到过程间的依赖分析结果。
- **应用降肿模块**
  - 采用切片技术，在多层依赖图上执行可达性查询，去除不包含在切片中的代码片段。在切片的基础上补充有关程序入口的相关代码，从而确保降肿后应用的鲁棒性、可分析性。

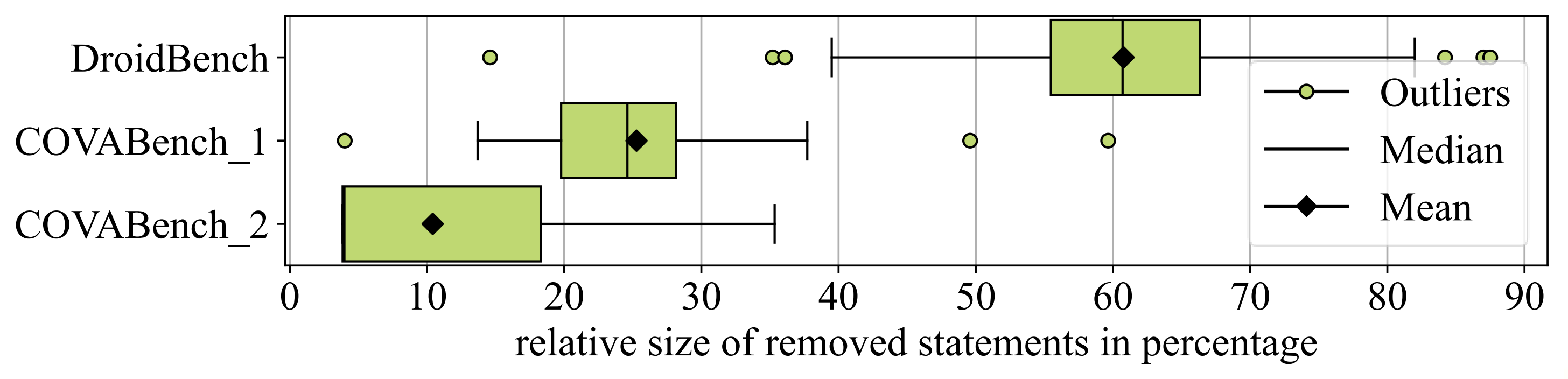
实验评估及工具仓库

Benchmark	App 数量	输入 App	检测出的Leak 数	检测出的TP	检测出的FP	准确率Prec.
DroidBench	119	原 App	94	73	21	0.777
		降肿 App	86	73	13	0.849

Percentage Reduction in FlowDroid Time After Debloating



Distribution of Reduced Percentage After Debloating



- 工具有效性评估。降肿后分析结果可以保留全部TP，FP的数量减少38.09%，准确率提升约7%。在真实应用上可以多检测出212个leak。
- 工具效率评估。所有应用在降肿后需要的分析时间均减少，时间减少量的中位数为63.82%，平均值为59.91%。
- 工具代码切分量评估。数据集中多数应用的代码切除率在55.48%-66.33%。部分真实应用的代码切除率在10%-25%。
- 工具github仓库链接及二维码如下：
  https://github.com/SQUARE-RG/FlowSlicer