融合推断分析的移动应用隐私收集分析方法

虞舒甜[1]，史一哲[1]，杨哲慜[1]
[1](复旦大学 计算机科学技术学院 上海 200438)

摘 要 随着移动应用服务多样化，其隐私收集方式日益复杂，除直接收集外，还包括基于已有用户数据进行隐私推断。推断行为虽常用于精准推荐和广告投放，但也带来隐私泄露等安全隐患。现有方法多聚焦于移动应用隐私的直接收集行为，缺乏对移动应用隐私推断能力的系统评估。本文提出一种融合推断分析的隐私收集分析方法，结合程序行为特征定位隐私直接收集行为，并基于大语言模型构建推断预测方法。实验评估表明，本文方法各模块的精确率和召回率均达84%以上。与现有隐私收集分析工具（FlowDroid、ClueFinder、RPNChecker）相比，隐私收集行为的检出提升超过40%。

关键词 隐私推断 隐私收集 大语言模型 静态分析

中图分类号 TP309.2 **文献标志码** A DOI: 10.3969/j.issn.1000-386x.2018.01.001

Privacy Data Collection Analysis for Mobile Applications with Integrated Inference Detection

Yu Shutian[1], Shi Yizhe[1], Yang Zhemin[1]
[1](*School of Computer Science, Fudan University, Shanghai 200433, China*)

Abstract With the increasing diversification of mobile application services, their methods of privacy collection have grown more complex, extending beyond direct collection to include privacy inference based on existing user data. While such inference behaviors are commonly used for targeted recommendations and advertising, they also pose significant privacy leakage risks. Existing approaches primarily focus on direct collection behaviors, lacking systematic evaluation of privacy inference capabilities. This paper proposes a privacy collection analysis method integrating inference analysis, which combines program behavior features to detect direct privacy collection and leverages large language models to build inference prediction mechanisms. Experimental results demonstrate that the proposed method achieves over 84% precision and recall across all modules. Compared to existing privacy analysis tools (FlowDroid, ClueFinder, RPNChecker), our method improves the detection rate of privacy collection behaviors by more than 40%.

Keywords Privacy inference Data collection Large language model Static analysis

0 引 言

随着移动互联网和智能设备的普及，移动应用在提供个性化服务的同时，日益依赖对用户隐私数据的收集与利用。传统的隐私收集行为主要通过系统接口调用、用户输入等方式直接访问敏感信息。但在此基础上，许多应用还会通过对已有用户数据进行分析与关联，推测出用户未主动提供的敏感信息，这一过程被称为隐私推断（Privacy Inference）。虽然推断结果并非直接采集所得，但本质上形成了新的用户特征数据，同样对个人隐私构成实质性影响。

在现实中，隐私推断已广泛存在于各类应用中。例如，电商平台可通过用户浏览和点击记录推测其收入水平和消费偏好；社交类应用则可能基于位置和通讯频率预测用户关系强度。这些推断结果被用于广告投放、内容推荐等自动化决策中，用户往往难以感知，也无法控制这些“派生隐私”的使用过程，从而带来了数据越权使用、隐私泄露等潜在风险。因此，如何识别和评估应用中的隐私推断能力，已成为隐私保护研究中的关键挑战。

当前学术界在隐私收集检测方面的研究工作大多侧重于应用隐私直接收集行为的检测，以辅助应用隐私收集行为合规性的分析。早期工作如 FlowDroid[1]、TaintDroid[2]等通过定位来自安卓系统敏感接口的隐私

收稿日期: 2025-04-xx。国家自然科学基金项目（62172104）。虞舒甜，硕士研究生，主研领域：移动隐私安全。史一哲，博士研究生，主研领域：移动隐私安全。杨哲慜，男，博士，副教授，CCF 会员，主研领域：程序分析技术和数据与隐私安全。

数据流来识别相关的隐私收集行为。后续工作[3-5]则开始关注第三方库的隐私收集行为，通过分析第三方库文档提取敏感接口进行识别。UIPicker[6]、SUPOR[7]、UiRef[8]、GUILeak[9]、IconIntent[10]则通过对交互界面的代码、图像、文本等特征的分析识别应用对用户输入隐私的收集行为。部分工作[11-15]则尝试通过对应用流量的分析识别其中的隐私数据。虽然这些方法在隐私直接收集检测的覆盖面和精度方面不断提升，但仍主要围绕具体接口或数据路径展开，缺乏对应用整体隐私收集能力的系统建模与评估。特别是对于应用通过已有数据推断用户敏感信息的行为，尚未形成通用分析方法。

此外，现有的隐私推断相关研究多数停留在特定推断任务的分析上，难以支持大规模、自动化的能力评估。例如，许多工作[16-18]关注位置信息与用户的人口统计学特征之间的关联，如用户性别、工作地点、家庭住址、教育水平等特征。也有工作[19-22]关注用户的应用使用数据与用户偏好和行为模式的关系，如 Bashir 等人的工作[19]研究 Facebook 等在线广告平台如何基于用户的网页浏览记录等交互数据推断用户偏好以及推断的准确性。此外，也有工作[23-24]依托于知识图谱、机器学习、大语言模型等技术研究更深层的关联关系。

由于应用的隐私推断能力依赖于其通过客户端直接收集的用户隐私，因此本文提出了一种融合推断分析的移动应用隐私收集分析方法，基于应用的直接隐私收集行为预测其可实现的隐私推断行为，综合直接收集行为可构成对应用隐私收集的整体评估。方法首先利用多元程序行为特征主动挖掘隐私源，实现对隐私直接收集行为的准确识别，并引入大语言模型技术针对隐私推断行为进行推理，最后基于现实应用实验和与现有工具的对比实验验证了方法的有效性。本文的主要贡献如下：

(1)本文创新性地结合多维度特征，提出了一种能够主动发现隐私源的识别方法。通过全面识别应用中的隐私获取方法，该方法能够精准定位多样化的隐私源数据，实现对应用隐私收集行为的精准识别。

(2)本文设计了基于大语言模型的隐私推断预测方法，通过整合现有隐私推断研究成果和大语言模型的推理能力，构建了面向隐私推断任务的知识提取方法和基于思维链的检索推理算法。

(3)本文提出了一种综合的隐私收集能力评估方法，实现了对应用整体隐私收集能力的全面分析。实验结果表明，方法在隐私源挖掘部分的精确率和召回率均超过 90%，隐私推断评估部分则分别达到了 84.14%和 86.79%，在与现有工具的对比实验中也表现出较好的性能。

1 方法架构

安卓应用的用户画像构建过程主要涉及两个步骤：1)应用在运行时通过多样的接口获取多来源的隐私数据，并经过一系列处理后发送到服务端进行数据存储；2)在服务端积累了充足的用户数据后，应用基于基本数据进行进一步的隐私推断，丰富用户画像。要实现对应用的隐私收集能力的完整评估，不仅需要对应用中的直接收集行为进行全面识别，覆盖多样的隐私来源，还需要基于应用直接隐私收集行为对应用服务端的隐私推断行为进行评估。

本文提出方法主要包含直接收集检测和推断预测两个模块，整体架构如图 1 所示。

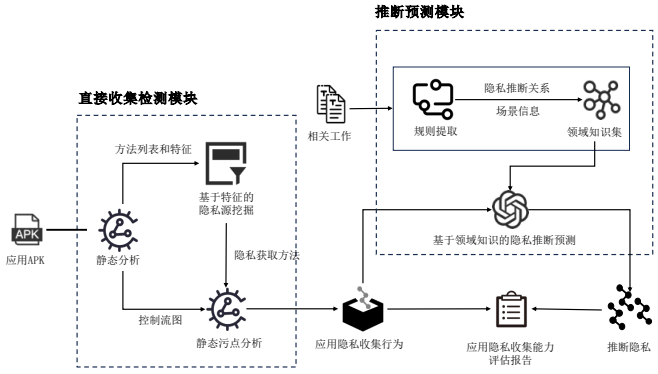

图 1 方法架构图

其中，直接收集检测模块结合静态程序分析技术和大语言模型技术，实现对应用隐私直接收集行为的全覆盖识别；推断预测模块利用大语言模型技术，将隐私推断能力评估问题转化为基于领域知识的模型推理问题。最终综合两个模块的结果，形成应用隐私收集行为评估报告。

2 直接收集检测模块

直接收集检测模块的目标是对给定的应用 APK 文件执行静态分析，获得应用的隐私直接收集行为检测结果。模块主要包括两部分：1)隐私数据源挖掘：通过静态分析提取应用方法的数据流特征和语义特征，识别其中的隐私获取方法和相应的隐私数据源；2)隐私收集行为识别：从隐私获取方法出发，定位从隐私数据源到隐私汇的可达路径，如存在路径能够将隐私数据发送到服务端，则认为存在针对该类型隐私的实际收集行为。

2.1 隐私获取方法特征及判定规则

数据获取方法指的是安卓应用在运行过程中通过系统接口、读取用户输入、解析流量数据等方式，获取设备和用户相关数据的功能性接口。移动应用在运行过程中通过数据的传递来进行功能的实现，例如用户在更新个人档案时，首先在输入框中输入个人信息后，应用通过自定的数据获取方法读取用户输入并生成存储用户输入的对象，随后将数据发送到服务端，进行用户个人档案信息的更新。应用中常见的数据获取形式如图 2 所示。

隐私获取方法是数据获取方法的一个子集，专指用于生成与用户隐私相关的数据变量的数据获取方法。本文针对应用中的隐私数据获取方法进行了调研和分析，总结出了多来源的隐私获取方法中存在的共性特征，分为数据获取特征、隐私获取特征和隐私语义特征三类特征。

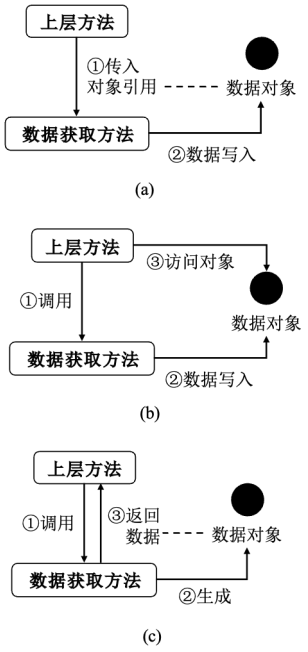

图 2 应用的主要数据获取形式

(1)数据获取特征：作为数据获取方法的子集，隐私获取方法具有典型的数据获取特征。由于数据获取方法通过各类方式将应用所需的数据转换为应用中的对象，以支持数据在各个组件间传递并服务于应用功能的实现，其方法必定涉及对外部对象的读取和写入，可总结为以下两种情况。

•方法将生成的数据存储到调用参数指向的对象或者全局变量中，如图 2(a)、(b)所示。

•方法读取外部数据并生成新的数据存储在新建对象中，并作为返回值返回，由调用该方法的上级方法赋值给对应对象，如图 2(c)所示。

(2)隐私获取特征：在数据获取方法中，其数据来源主要来自于从外部对象获取的值或者通过系统调用生成的值。而隐私获取方法中数据来源很有可能具有隐私属性，主要包括调用具有隐私属性的系统接口和读入具有隐私属性的字段两种情况。

(3)隐私语义特征：隐私获取方法中也会存在一定的语义信息表征其隐私属性，常体现在方法签名、方法调用栈、方法内使用的常量中。

根据上述特征，本文根据特征所对应的代码特性进一步将特征划分为数据流特征和语义特征两类。

数据流特征主要用于判断方法是否满足数据获取特征，属于一个数据获取方法。语义特征包含方法调用的系统接口、读取的外部对象的字段名、方法签名、使用的常量等特征，对应上述的隐私获取特征和隐私语义特征，用于判断一个数据获取方法是否与隐私相关。

2.2 隐私数据源挖掘

基于以上特征总结和规则定义，隐私数据源挖掘模块结合静态程序分析与大语言模型技术，从应用中识别潜在的隐私获取方法，并生成包含方法信息、外部变量写入情况及对应隐私类型的结果列表。该模块包括三个阶段：首先提取所有方法的特征信息，如读写对象、调用点及语义特征；其次基于数据流规则筛除不符合条件的方法，获得数据获取方法集合；最后根据语义规则识别隐私相关性并确定其类型。

在语义特征的分析上，本文利用大语言模型技术实现隐私相关性的确定和隐私类型的识别，解决此前方法[13-15,25]依赖关键词列表、难以处理复杂方法语义的局限。然而，通用大语言模型在隐私语义特征识别这一特定任务上的性能表现有待提升，本文考虑在现有的大模型基础上进行二次训练，以支持隐私获取方法的识别。具体而言，二次训练在 Llama-3-8B-Instruct 模型[26]基础上进行预训练和微调，其中预训练阶段通过大规模隐私相关文本注入知识，包含隐私法律法规和隐私政策等文本数据；微调阶段则采用专用问答模板和 LoRA（Low-Rank Adaptation）[27]技术优化模型，利用 500 个随机选取的数据收集方法进行人工标注后构成微调数据集进行训练，使其在隐私语义识别任务中具有更强的适应能力。

完成预训练和微调两个步骤后，本文使用 500 个随机选取且与训练集不重合的数据收集方法检验模型的准确率，准确率达到 86.37%。

2.3 隐私收集行为识别

隐私收集行为识别基于前述隐私源挖掘结果，采用静态污点分析技术追踪隐私数据的传播路径，判断其是否通过网络接口发送至服务端，从而构成隐私收集行为。

针对每个隐私获取方法，本文根据其隐私数据的输出方式（返回值或外部对象写入）标记相应污点源，并从程序入口出发，结合控制流图与调用图筛选可达调用点，确保传播路径真实可达。随后，以网络请求类 API 为污点汇，执行前向污点分析，识别数据从源到汇的完整路径，并据此提取收集的隐私类型。

3 推断预测模块

推断预测模块的目标是基于直接收集检测模块确定的收集隐私类型，预测应用可以实现的推断行为。大语言模型具有强大的文本理解和知识推理能力，在隐私推断领域也有较多的落地实践，可以基于应用的隐私收集行为和相关的领域知识推理其隐含的关联关系与推断逻辑。

本文设计模块主要包含两个环节：1)大语言模型驱动的知识提取：该部分针对隐私推断知识特性利用大语言模型对文本进行知识提取和结构化；2)基于思维链的检索推理：该部分以前序生成的隐私推断知识集作为基础知识，基于应用收集隐私类型在知识集中进行检索，形成思维链，与其他信息一并提供给大模型进行推理。

3.1 大语言模型驱动的知识提取

该环节的主要任务是理解隐私推断相关工作的文本内容，提取出需要的隐私推断规则信息。

本文工具利用大语言模型的语言理解能力提取文章中的关键信息，总结推断规则信息。然而，尽管大语言模型在处理短文本生成、语言翻译和自然语言理解等任务中表现出色，当面对学术论文或技术文档这样的长文本时，其理解能力往往受到输入长度限制和模型架构的影响，难以全面捕捉文档的整体逻辑结构，这种局限性可能降低提取知识的准确性。

为了弥补这一不足，本文设计了针对学术论文的分解-推理-汇总-精炼的自底向上模式多层提示工程技术（图 3），采用分段解析和多轮提示工程（Prompt Engineering），逐章提取信息并保留关键文本，同时通过设计层级化提示语引导模型生成符合要求的输出。具体而言，本文首先提示模型根据每个章节的文本数据提取需要的信息和对应的关键文本；随后将相邻段落关联起来，基于已提取的信息和关键文本进行再次总结，根据关键的原文信息和章节之间的关系解决段落之间可能出现的冲突信息；最后再根据所有的信息进行汇总，得到最终的结果。在各个步骤之间，提示词也会引导模型生成结果的置信度，以辅助后续的判断。

在初步提取隐私相关语句后，本文进一步结构化文本信息，提取源数据类型、收集场景、推断类型及其准确率。该阶段结合大语言模型的语义理解能力与预设先验知识，将文本映射到标准化标签，最终生成结构化的隐私推断四元组。

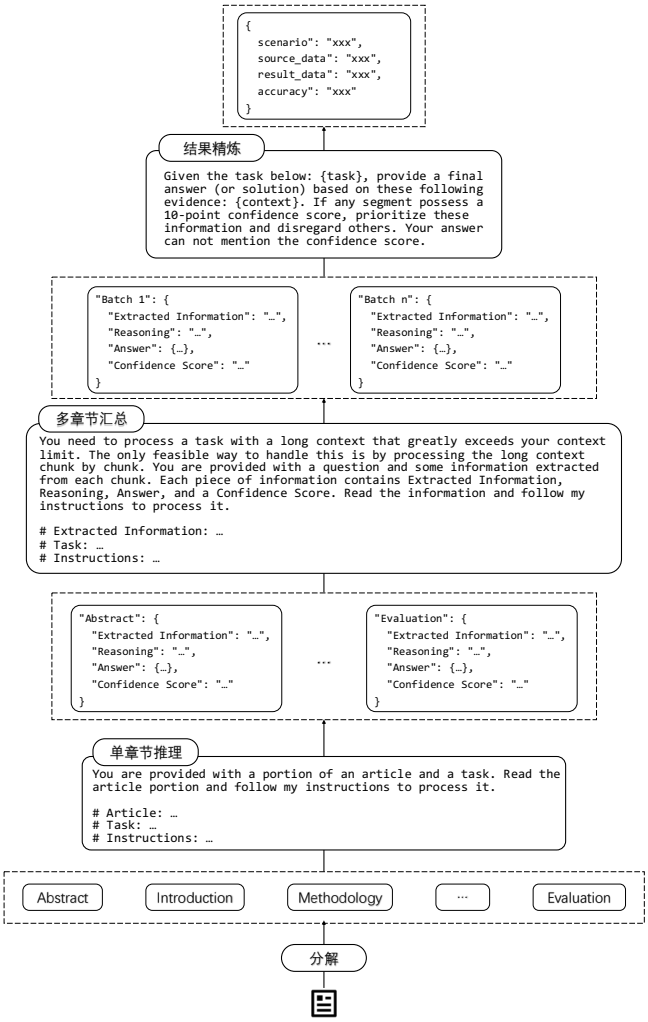

图 3 文本规则提取提示工程

本文从计算机和数据科学领域学术会议和期刊中选取 7 个相关会议和期刊，并爬取过去六年的所有文章，作为隐私推断领域的核心文献来源。为了确保论文数据集和隐私推断的相关性，本文通过标题筛选和摘要分析，剔除和隐私推断无关的研究，最终保留了 118 篇相关文章，如表 1 所示。

表 1 隐私推断相关论文数据集

会议名称	2019-2024 文章总数	相关论文数
IEEE S&P	898	4
NDSS	578	11
ACM CCS	1577	9
USENIX Security	1607	7
ICWSM	594	36
IWMUT	793	30
SigKDD	1505	21
总数	7552	118

本文使用 Deepseek-2.5[28]模型完成整个知识提取的流程，最终从 118 篇相关文章中抽取了 140 条规则。

通过对 50 篇文章提取规则进行人工检查后，四元组各元素的提取准确率均都达到 80%以上，如表 2 所示。

表 2 隐私推断知识提取评估结果

指标	应用场景	初始隐私类型	推断隐私类型	推断准确率
准确率	82.14%	90.48%	88.09%	79.76%

3.2 检索推理算法

本文通过引入隐私推断知识增强大语言模型在隐私推理任务中的理解能力，整体流程包括知识检索与模型推理两个阶段。

首先，根据输入的隐私类型和可选的应用场景，从知识库中检索源数据匹配的推断知识，优先选择场景一致的内容，最多保留 10 条用于构建提示词。随后，方法结合应用收集的隐私类型、运行场景及检索到的推断知识构建提示，引导大模型完成两轮推理。第一轮生成初步的隐私推断结果，包括类型、原因与置信度，提示词模板如图 4 所示；第二轮则对推理结果进行优化，完成类型映射和错误筛除。

为提升推理准确性，优化阶段引入多项筛选规则，如源类型存在性、已收集类型排除、置信度阈值限制和合理性审查，并通过提示词（图 5）明确表达，引导模型剔除低可信度或逻辑不符的推理结果。

Task

You are tasked with analyzing the privacy inference capabilities of a mobile application. Based on the provided collected privacy list, application context, and a partial inference knowledge graph, determine which types of privacy can be inferred. Each inference is defined by one or more source privacy type, a relationship, and a target privacy type with a confidence score. Please return the answer in JSON format.

Knowledge Graph

The inference knowledge graph contains the following relationships:
- From <Source List> acquired in the context of <Scene>, it is possible to infer <Target List>.
- From <Source List>, it is possible to infer <Target List> with a confidence score of <accuracy>.

Input

```
{
    "Application Scenarios": [ "<<<Scene1>>>", "<<<Scene2>>>" ],
    "Collected Privacy Data Types": ["<<<Privacy Data>>>", "<<<Privacy Data>>>",
"<<<Privacy Data>>>" ]
}
```

Output

```
{
    "<<<Privacy Data>>>":{
        "Source": ["<<<Privacy Data>>>", "<<<Privacy Data>>>"]
        "Confidence": 0.xx,
        "Reasoning": "xxx"
    },
    "<<<Privacy Data>>>":{
        "Source": ["<<<Privacy Data>>>", "<<<Privacy Data>>>"]
        "Confidence": 0.xx,
        "Reasoning": "xxx"
    }
}
```

图 4 推理提示词模板

4 实　验

本章首先对直接收集检测和隐私推断预测两个模块的效果进行评估，随后通过与 FlowDroid、ClueFinder、RPNChecker 的整体性能比较，体现本文方法的有效性。

Task

The Task is to refine the privacy inference results.Each inference includes one or more source privacy type, a relationship, and a target privacy type with a confidence scoreand reasoning details. You need to determine if the inference is feasible based on the given information and refine it based on the given rules.

Instructions

1. All privacy types must belong to the predefined privacy type list: ["<<<Privacy Data>>>", "<<<Privacy Data>>>", "<<<Privacy Data>>>"…]
2. Any inferred types that are already included in the provided information should be removed.

Input

```
{
    "Application Scenarios": [ "<<<Scene1>>>", "<<<Scene2>>>" ],
    "Collected Privacy Data Types": ["<<<Privacy Data>>>", "<<<Privacy Data>>>",
"<<<Privacy Data>>>" ]
    "Answer": {
        "<<<Privacy Data>>>":{
            "Source": ["<<<Privacy Data>>>", "<<<Privacy Data>>>"]
            "Confidence": 0.xx,
            "Reasoning": "xxx"
        },
        "<<<Privacy Data>>>":{
            "Source": ["<<<Privacy Data>>>", "<<<Privacy Data>>>"]
            "Confidence": 0.xx,
            "Reasoning": "xxx"
        }
    }
}
```

Output

```
{
    "<<<Privacy Data>>>":{
        "Source": ["<<<Privacy Data>>>", "<<<Privacy Data>>>"]
        "Confidence": 0.xx
    },
    "<<<Privacy Data>>>":{
        "Source": ["<<<Privacy Data>>>", "<<<Privacy Data>>>"]
        "Confidence": 0.xx
    }
}
```

图 5 优化提示词模板

4.1 数据集

为了测试在真实应用场景中的工具效果，本文收集了一批真实移动应用的 APK 文件，构建了实验所需的数据集。此实验数据集的构建来自于谷歌应用商店（Google Play Store）。本文从 33 个应用类别在德国、英国、美国三个国家 2024 年 9 月的榜单中选取排名前 50 的应用形成初始数据集，最终收集到 4656 个应用。同时为了对工具效果进行评估，本文从初始数据集中随机选取 50 个应用构成验证数据集。

4.2 评估指标

本文选择了四个评估指标：准确率（Accuracy）、精确率（Precision）、召回率（Recall）以及综合指标 F1-score。计算方式如下所示：

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{1}$$

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{4}$$

4.3 隐私源挖掘能力评估

由于依赖人工无法完成对应用内所有方法的人工标注和验证，因此本文从验证应用的所有方法中随机选取 500 个方法人工标注其隐私收集属性，形成最终

的基准数据集，其中包含 384 个隐私源方法和 116 个非隐私源方法。在实验验证过程中，本文对以上 500 个方法按照工具设计进行数据流特征和语义特征的抽取和隐私相关性的判断，最终将所得结果与基准数据集的结果进行比较得出评估结论。

表3展示了本文方法在基准数据集上的评估结果。隐私源挖掘模块共识别了 359 个隐私收集方法，其精确率和召回率分别是 97.77%和 91.41%，F1-score 为 94.48%，体现了隐私源挖掘技术在真实应用分析中的良好性能。

表 3 隐私源挖掘能力评估结果

TP	FP	TN	FN	精确率	召回率	F1-score
351	8	108	33	97.77%	91.41%	94.48%

在本轮评估中，误报和漏报主要由两类问题引起：一是应用中广泛存在的代码混淆行为，使得方法签名、常量命名和对象字段缺乏可识别的语义特征，干扰了模型对方法隐私相关性的准确判断，进而导致误报与漏报；二是部分隐私收集逻辑依赖异步调用链或数据持久化路径，涉及复杂的跨函数、跨对象交互，超出了当前模型对数据流和控制流的建模能力，导致漏报。

除此之外，本文还发现一类基于静态数据混淆的漏报特例：出于安全保障的原因，OPPO 软件商店等应用会将一些数据以自定义编码的形式存储在内存中，需要使用时将先进行解码，这导致静态分析时无法获取其语义信息，导致漏报；然而，这种基于简单编码的本地数据混淆行为在应用运行时无法提供安全保障，仅仅用于增加安全分析人员的逆向难度。

4.4 直接隐私收集行为识别能力评估

直接收集模块的作用是检测应用程序中所触发的隐私收集行为，本文通过半自动的动态分析对直接隐私收集行为进行评估。

具体而言，本文采用下述步骤收集隐私数据收集行为形成基准数据集：1)使用 Frida 工具启动分析目标应用，并插桩用于传输数据的网络请求 API；2)手动探索界面以触发尽可能多的隐私收集行为，对于被触发的网络请求 API，利用插桩记录器调用栈和参数信息作为数据流事实。3)重复上述步骤直到不再触发新的页面。

在该流程中，本文仅能够收集到真实可触发数据流的一部分，一些触发条件困难（例如基于应用崩溃或长时间运行后定期上传的数据流）则无法被追踪到。本文以人工逆向的方式尽可能补全无法动态触发的数据流，最终结合两部分构建了直接隐私收集行为的基准数据集。最终，本文在 50 个 APK 中共抽取了 362 条隐私数据流用于构建基准数据集。

本文将被检出的隐私收集数据流作为正样本进行评估，由于无法判断检出的不在基准数据集中的隐私数据流是否误报，因此本文评估中缺少负样本的数据，仅评估了真阳性（TP）、假阴性（FN），最终也仅针对召回率进行评估。具体实验结果如表 3 所示，直接隐私收集检测模块的实验召回率为 93.10%，说明直接隐私收集模块具有良好的分析性能。

表 4 直接收集行为识别评估结果

TP	FN	召回率
337	25	93.10%

4.5 隐私推断预测能力评估

本文基于验证数据集中 50 个 APK 的直接隐私收集数据和其应用场景列表，结合专家知识人工标注隐私类型之间的推断关系共 159 条，形成基准数据集。这一基准数据集是对应用可能具备的隐私推断能力的保守下界（lower-bound）。由于某些复杂推断关系涉及隐私类型间的多跳间接关联或依赖上下文的语义推理，人工标注过程中可能存在漏标，但本文工作在标注过程中通过两名专家同时标注并交叉验证来提升数据集质量，尽可能保证数据集的可靠性。

对于每个应用，推断可得隐私类型同时存在于工具分析结果和专家标注结果中，则被认为是 TP；仅存在于工具分析结果中则为 FP；仅存在于专家标注结果中则为 FN。

最终实验结果如表 5 所示，整体而言方法达到了 F1-score 为 85.4%的性能。部分误报主要源于模型推理结果中个别隐私类型的置信度接近设置的阈值边界，尽管其推理依据较弱，但仍被纳入最终结果。为了在准确性和覆盖率之间取得合理平衡，本文将置信度阈值设定为 0.75。该阈值是基于验证集上多轮调试后选定的最优值，在保持较低误报率的同时，尽可能覆盖潜在的推断隐私类型，支持对应用推断能力的全量评估。此外，本文也对引入领域知识的有效性进行了验证，在提示词中删去领域知识部分后对以上50款应用进行了重复实验，并进行了结果统计。结果如表 6 所示，可以看到删去领域知识后大模型推理在各项指标上均出现了较大幅度的性能下降，主要原因是因为领域知识的引入可以引导大模型进行合理的推理，一方面通过领域知识引导大模型进行推理，减少了 FN；另一方面利用领域知识避免大模型进行过于泛化的推理，从而减少了 FP。

表 5 隐私推断预测评估结果

TP	FP	FN	精确率	召回率	F1-score
138	26	21	84.14%	86.79%	84.45%

表 6 去除领域知识的隐私推断预测评估结果

TP	FP	FN	精确率	召回率	F1-score
93	54	66	63.26%	58.49%	60.78%

4.6 相关工作对比实验

为验证所提出方法在隐私收集能力识别方面的有效性，本文选取了 FlowDroid[1]、ClueFinder[25] 和 RPNChecker[29] 三项工具开展对比实验。

FlowDroid 基于静态污点分析识别敏感数据从隐私源到泄露点的传播路径；ClueFinder 通过分析语句语义信息，挖掘潜在的隐私收集操作；RPNChecker 则融合了对系统敏感接口的动态追踪与网络流量分析，以识别运行时的隐私收集行为。本文对上述工具均进行了针对本文实验的适配性调整，对于 FlowDroid，本文给定了一系列安卓系统敏感接口和常见第三方敏感接口作为隐私源进行实验；由于 ClueFinder 未开源，本文基于其方法描述实现了一个原型工具以还原其关键逻辑。

本文实验在包含 50 个 APK 的数据集上运行上述工具，并统计其识别的隐私类型数量。经人工验证结果准确性（表 7），本文方法在隐私收集行为识别上的精确率均优于相关工作，提升至 90%以上。去除假阳性后，各工具的最终识别结果如表 8 所示。

表 7 本文工具与相关工作的实验精确率

本文工具	FlowDroid	ClueFinder	RPNChecker
91.60%	89.47%	87.5%	88.35%

表 8 本文工具与相关工作检出的隐私收集数

本文工具		FlowDroid	ClueFinder	RPNChecker
含推断	不含推断	476	728	675
1036	885			

本文工具在直接收集行为的检测上（885 项）高于 FlowDroid（476 项）、RPNChecker（675 项），略高于 ClueFinder（728 项），表明在直接收集行为分析方面具备更强覆盖能力。此外，本文工具通过引入隐私推断分析，在整体隐私识别数量上达到 1036 项，相较 ClueFinder 实现了超过 42.3% 的总检出能力提升。

现有方法主要聚焦于客户端的直接隐私收集路径，无法识别基于数据关联的服务端推断行为。相比之下，本文工具不仅在直接收集能力上取得更优结果，同时还能覆盖额外的推断隐私类型。

5 总结与反思

应用隐私收集行为一直以来是移动应用安全中的重要研究内容。移动应用的隐私收集行为日益复杂，已从传统的接口调用与用户输入等直接收集方式拓展至隐私推断等间接方式。然而，现有评估方法多聚焦于直接收集，难以全面反映应用的实际隐私收集能力。为此，本文提出了一种融合推断分析的评估方法，以实现对应用隐私直接收集和推断行为两方面能力的综合评估。

本文提出的方法实现了对应用整体隐私收集能力的全面分析，但是本文在各个模块的实现中仍然存在一定的局限性。目前的检索推理算法的精确率和召回率尽管达到了 84.14%和 86.79%，但仍存在进一步优化的空间。例如，目前推理主要依赖静态分析生成的全局性描述，缺乏对隐私使用上下文的细粒度建模，可能导致部分推断关系被遗漏。未来可尝试将隐私类型与具体收集场景结构化结合，提升推理准确性，同时引入动态分析手段，弥补静态方法的信息缺失。

此外，本文方法分析得到的隐私推断能力仅反映潜在可能性，尚不能作为实际隐私推断行为的直接证据。未来研究可探索如侧信道分析等技术，从客户端行为中捕捉推断行为的间接信号，提升对服务端隐私推断的可验证性和解释力。

参 考 文 献

[1] ARZT S, RASTHOFER S, FRITZ C, et al. FlowDroid: Precise context, flow, field, object-sensitive and lifecycle-aware taint analysis for Android apps[J]. ACM SIGPLAN Notices, 2014, 49(6): 259-269.

[2] ENCK W, GILBERT P, HAN S, et al. TaintDroid: An information-flow tracking system for real-time privacy monitoring on smartphones[J]. ACM Transactions on Computer Systems (TOCS), 2014, 32(2): 1-29.

[3] ZIMMECK S, WANG Z, ZOU L, et al. Automated analysis of privacy requirements for mobile apps[C]// Proceedings of the 2016 AAAI Fall Symposium Series. 2016.

[4] LIU X, LIU J, ZHU S, et al. Privacy risk analysis and mitigation of analytics libraries in the Android ecosystem[J]. IEEE Transactions on Mobile Computing, 2019, 19(5): 1184-1199.

[5] WANG J, XIAO Y, WANG X, et al. Understanding malicious cross-library data harvesting on Android[C]// Proceedings of the 30th USENIX Security Symposium (USENIX Security 21). 2021: 4133-4150.

[6] NAN Y, YANG M, YANG Z, et al. UIPicker: User-input privacy identification in mobile applications[C]//Proceedings of the 24th USENIX Security Symposium (USENIX Security 15). 2015: 993-1008.

[7] HUANG J, LI Z, XIAO X, et al. SUPOR: Precise and scalable sensitive user input detection for Android

apps[C]//Proceedings of the 24th USENIX Security Symposium (USENIX Security 15). 2015: 977-992.

[8] ANDOW B, ACHARYA A, LI D, et al. UIRef: Analysis of sensitive user inputs in Android applications[C]//Proceedings of the 10th ACM Conference on Security and Privacy in Wireless and Mobile Networks. 2017: 23-34.

[9] WANG X, QIN X, HOSSEINI M B, et al. GuiLeak: Tracing privacy policy claims on user input data for Android applications[C]//Proceedings of the 40th International Conference on Software Engineering. 2018: 37-47.

[10] XIAO X, WANG X, CAO Z, et al. IconIntent: Automatic identification of sensitive UI widgets based on icon classification for Android apps[C]//Proceedings of the 2019 IEEE/ACM 41st International Conference on Software Engineering (ICSE). IEEE, 2019: 257-268.

[11] SONG Y H, HENGARTNER U. Privacyguard: A vpn-based platform to detect information leakage on android devices[C]//Proceedings of the 5th Annual ACM CCS Workshop on Security and Privacy in Smartphones and Mobile Devices. 2015: 15-26.

[12] VALLINA-RODRIGUEZ N. Lumen Privacy Monitor [EB/OL]. 2016. [2024-12-20]. https://www.icsi.berkeley.edu/icsi/projects/networking/haystack.

[13] SHUBA A, LE A, GJOKA M, VARMARKEN J, LANGHOFF S, MARKOPOULOU A. Antmonitor: Network traffic monitoring and real-time prevention of privacy leaks in mobile devices[C]//Proceedings of the 2015 Workshop on Wireless of the Students, by the Students, & for the Students. 2015: 25-27.

[14] REN J, RAO A, LINDORFER M, LEGOUT A, CHOFFNES D. Recon: Revealing and controlling pii leaks in mobile network traffic[C]//Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services. 2016: 361-374.

[15] SHUBA A, BAKOPOULOU E, MEHRABADI M A, et al. Antshield: On-device detection of personal information exposure[J]. arXiv preprint arXiv:1803.01261, 2018. DOI: 10.48550/arXiv.1803.01261.

[16] BARON B, MUSOLESI M. Where you go matters: A study on the privacy implications of continuous location tracking[J]. Proceedings of the ACM on Interactive,Mobile, Wearable and Ubiquitous Technologies, 2020, 4(4): 1-32.

[17] WANG P, SUN F, WANG D, et al. Inferring demographics and social networks of mobile device users on campus from ap-trajectories[C]//Proceedings of the 26th international conference on world wide web companion. 2017: 139-147.

[18] BACKES M, HUMBERT M, PANG J, et al. walk2friends: Inferring social links from mobility profiles[C]//Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security. 2017: 1943-1957.

[19] BASHIR M A, FAROOQ U, SHAHID M, et al. Quantity vs. quality: Evaluating user interest profiles using ad preference managers.[C]//NDSS. 2019.

[20] ZHAO S, PAN G, ZHAO Y, et al. Mining user attributes using large-scale app lists of smartphones[J]. IEEE Systems Journal, 2016, 11(1): 315-323.

[21] ZHAO S, XU Y, MA X, et al. Gender profiling from a single snapshot of apps installed on a smartphone: An empirical study[J]. IEEE Transactions on Industrial Informatics, 2019, 16(2): 1330-1342.

[22] ZHAO S, XU F, LUO Z, et al. Demographic attributes prediction through app usage behaviors on smartphones[C]//Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers. 2018: 870-877.

[23] NARAIN S, VO-HUU T D, BLOCK K, et al. Inferring user routes and locations using zero-permission mobile sensors[C]//2016 IEEE Symposium on Security and Privacy (SP). IEEE, 2016: 397-413.

[24] ZUFFEREY N, HUMBERT M, TAVENARD R, et al. Watch your watch: Inferring personality traits from wearable activity trackers[C]//32nd USENIX Security Symposium (USENIX Security 23). 2023: 193-210.

[25] YUHONG N, ZHEMIN Y, YUAN Z, et al. Finding clues for your secrets:Semantics-driven, learning-based privacy discovery in mobile apps[C]//2018 Network and Distributed Systems Security (NDSS) Symposium. 2018.

[26] Meta. Meta-Llama-3-8B-Instruct [EB/OL]. 2024. [2024-12-20]. https://huggingface.co/meta-llama/Meta-Llama-3-8B-Instruct

[27] HU E J, SHEN Y, WALLIS P, et al. Lora: Low-rank adaptation of large language models[A]. 2021.

[28] DeepSeek. DeepSeek-V2.5 [EB/OL]. 2024. [2024-12-20]. https://github.com/deepseek-ai/DeepSeek-V2

[29] LI S, YANG Z, NAN Y, et al. Are we getting well-informed? an in-depth study of runtime privacy notice practice in mobile apps[C]//Proceedings of the 2024 on ACM SIGSAC Conference on Computer and Communications Security. 2024:1581-1595.

联系方式

虞舒甜（本文负责人） 硕士研究生

主研领域：移动隐私安全
身份证号：330481199902283044
联系电话：17621192815
就读单位：复旦大学计算机科学技术学院
通信地址：上海市杨浦区淞沪路 2005 号复旦大学计算机科学技术学院
邮政编码：200438
电子邮箱：yushutian@fudan.edu.cn

史一哲 博士研究生

主研领域：移动隐私安全
身份证号：410223199909133538
联系电话：18021005851
就读单位：复旦大学计算机科学技术学院
通信地址：上海市杨浦区淞沪路 2005 号复旦大学计算机科学技术学院
邮政编码：200438
电子邮箱：yzshi23@m.fudan.edu.cn

杨哲慜 博士、副教授

主研领域：程序分析技术和数据与隐私安全
身份证号：xxx
联系电话：xxxx
就读单位：复旦大学计算机科学技术学院
通信地址：上海市杨浦区淞沪路 2005 号复旦大学计算机科学技术学院
邮政编码：200438
电子邮箱：yangzhemin@fudan.edu.cn