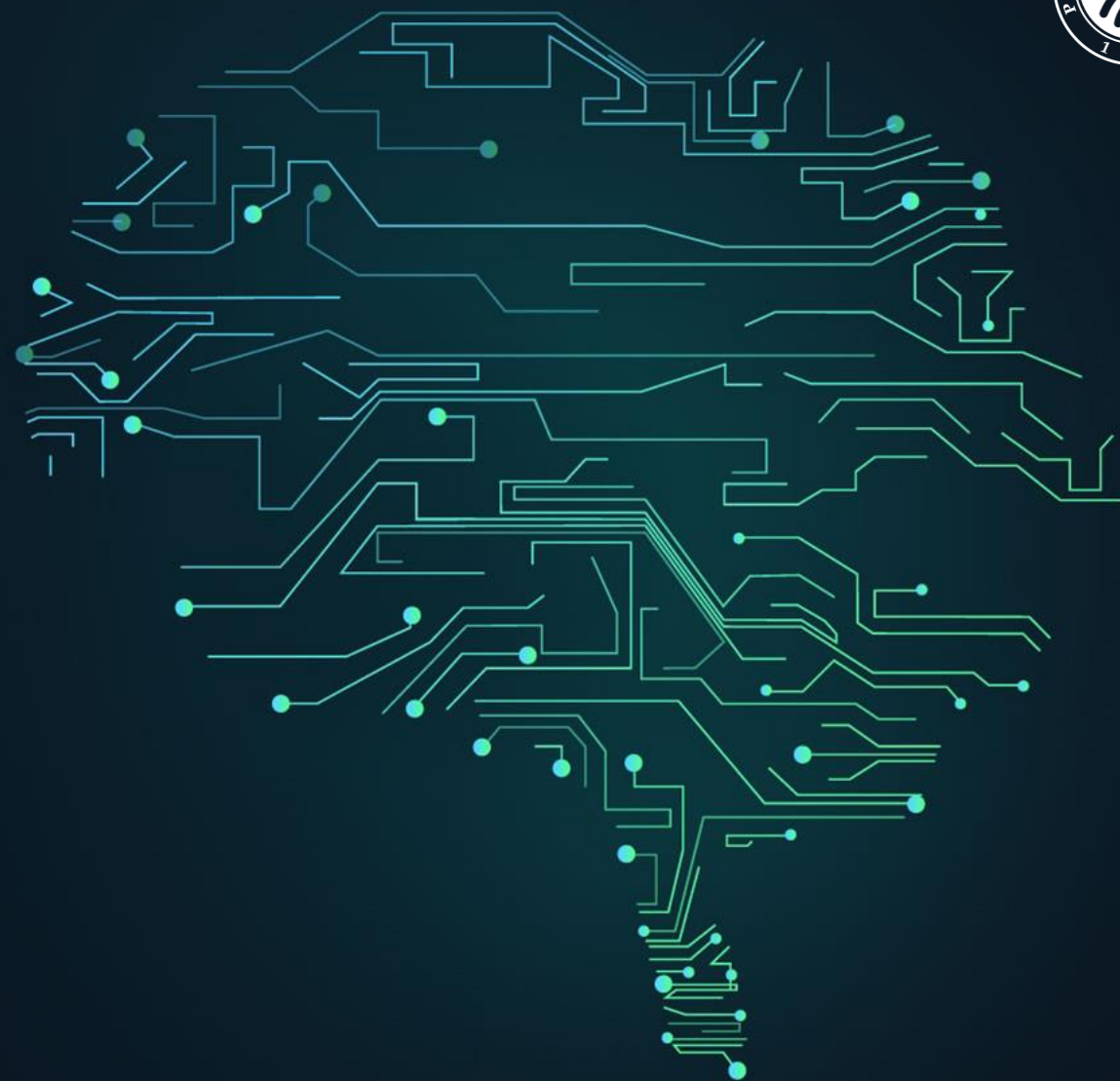




移动应用生态系统： 现状与挑战

郭 耀

北京大学 计算机科学技术系





目录

Contents

01、背景与挑战

02、应用市场生态分析

03、移动开发者生态分析

04、总结与展望

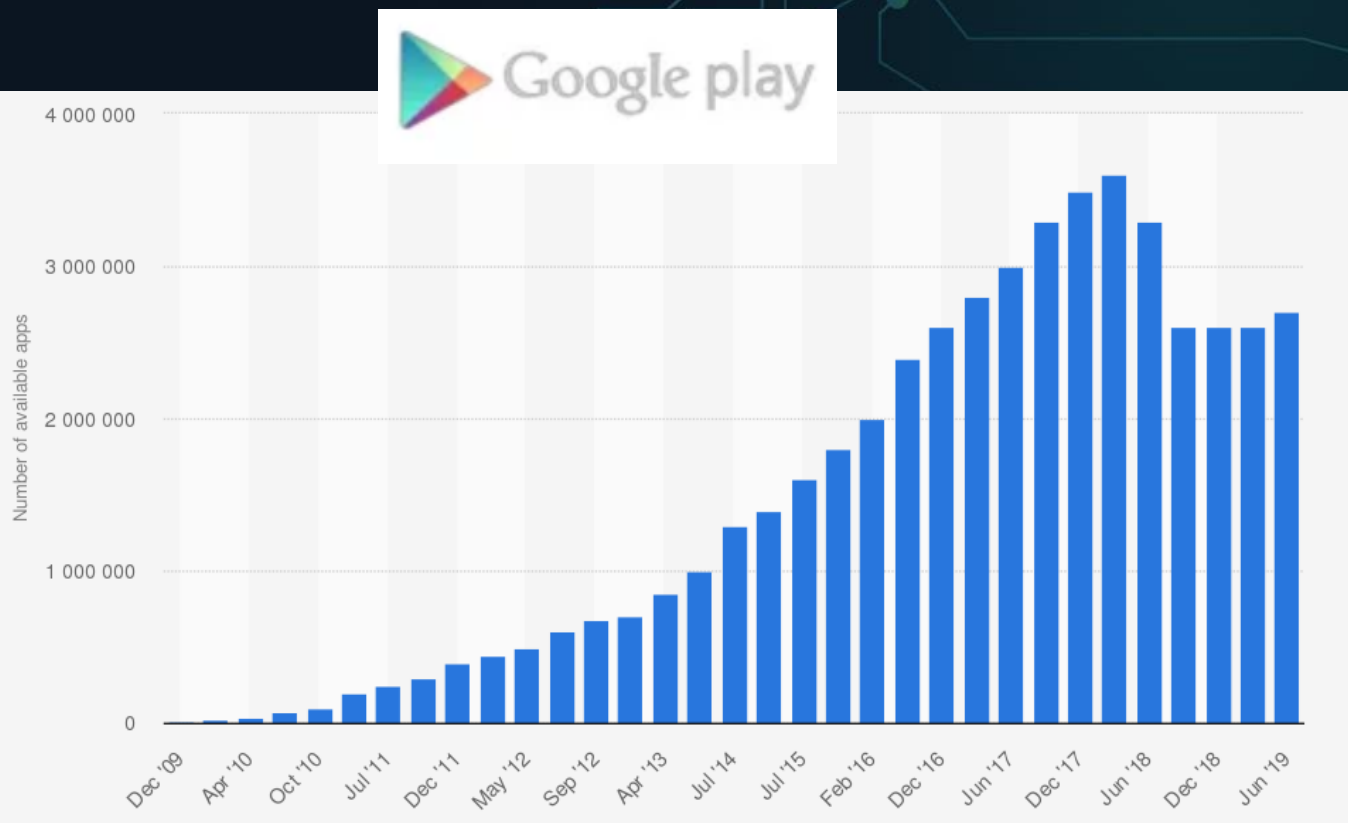


01

章节 PART

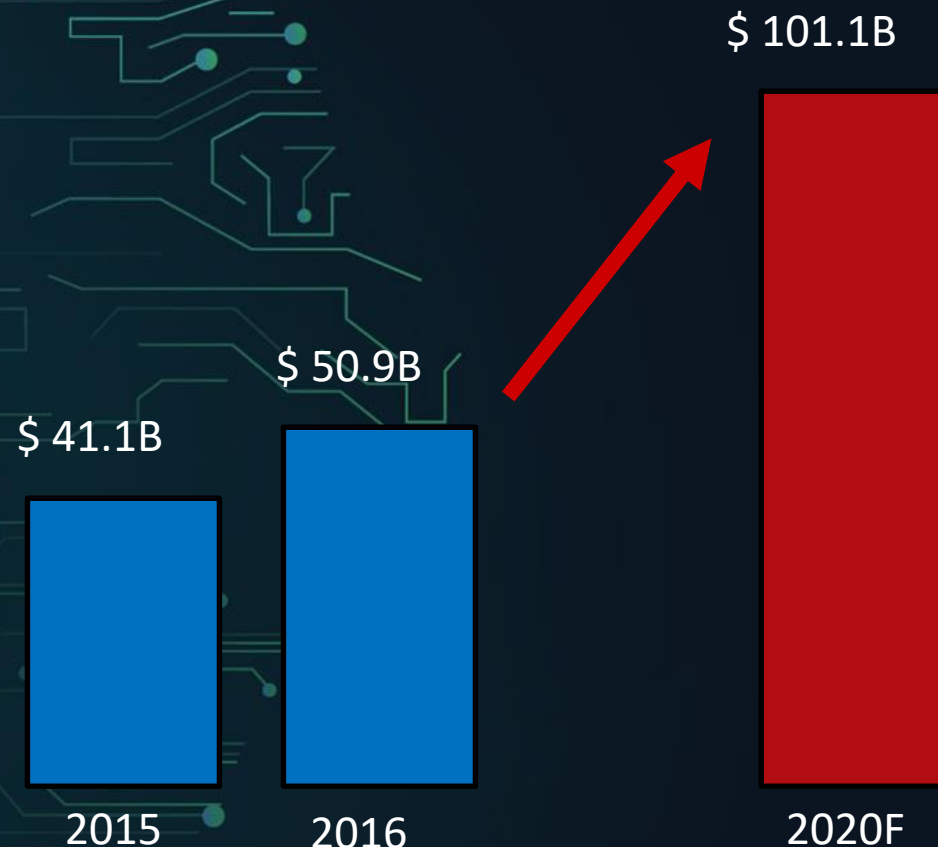
移动应用生态系统： 背景与挑战

移动应用的数量和市场持续增长



Number of Apps in Google Play:
2.7 million by the end of August 2019

Source: AppBrain



App Economy Forecasts:
\$100 Billion Revenue by 2020

Source: App Annie



移动应用生态系统 (Mobile App Ecosystem)

智能手机和移动系统



移动用户



APP开发者



移动APP



安智市场



Google Play



豌豆荚



PP助手



应用宝



乐趣市场



搜狗手机助手



小米应用商店



360手机助手

应用市场

移动应用生态系统 规模和复杂性 与日俱增

NCTS

第二届
中国云测试行业峰会
CHINA CLOUD TESTING INDUSTRY SUMMIT

成千上万的定制移动系统版本！



数十亿移动用户！



超百万APP开发者！



上千万的
移动APP！



安智市场



Google Play



豌豆荚



PP助手



应用宝



历趣市场



搜狗手机助手



小米手机助手



360手机助手

上百个移动应用市场！

(国内) 移动应用生态系统面临严峻挑战!

NCTS

第二届中国云测试行业峰会

大量的移动应用市场鱼龙混杂，用户面临困难选择!

市场中的app没有经过严格的审查，app质量堪忧!

移动开发者没有掌握足够的技能，开发的app存在问题

移动用户对技术疏于了解，对移动应用的功能权限不理解

- 基于大规模移动应用分析平台，理解移动应用生态系统面临的问题与挑战，检测app的恶意行为。

大规模移动应用安全与隐私分析平台



国内外Android应用市场生态的分析

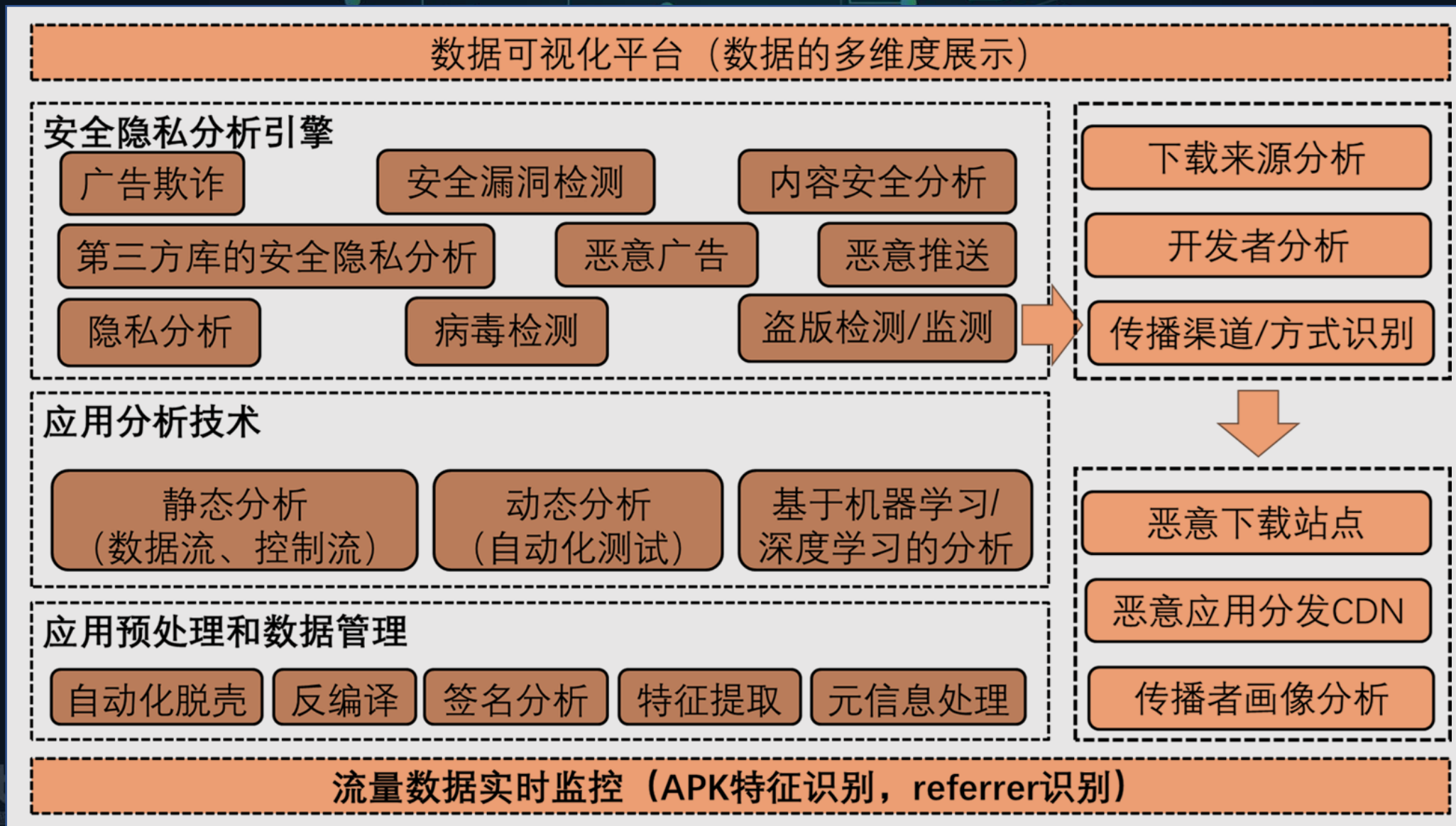


Android开发者生态的分析



利用app测试技术检测恶意行为

大规模移动应用分析平台





02 章节 PART

移动应用市场生态

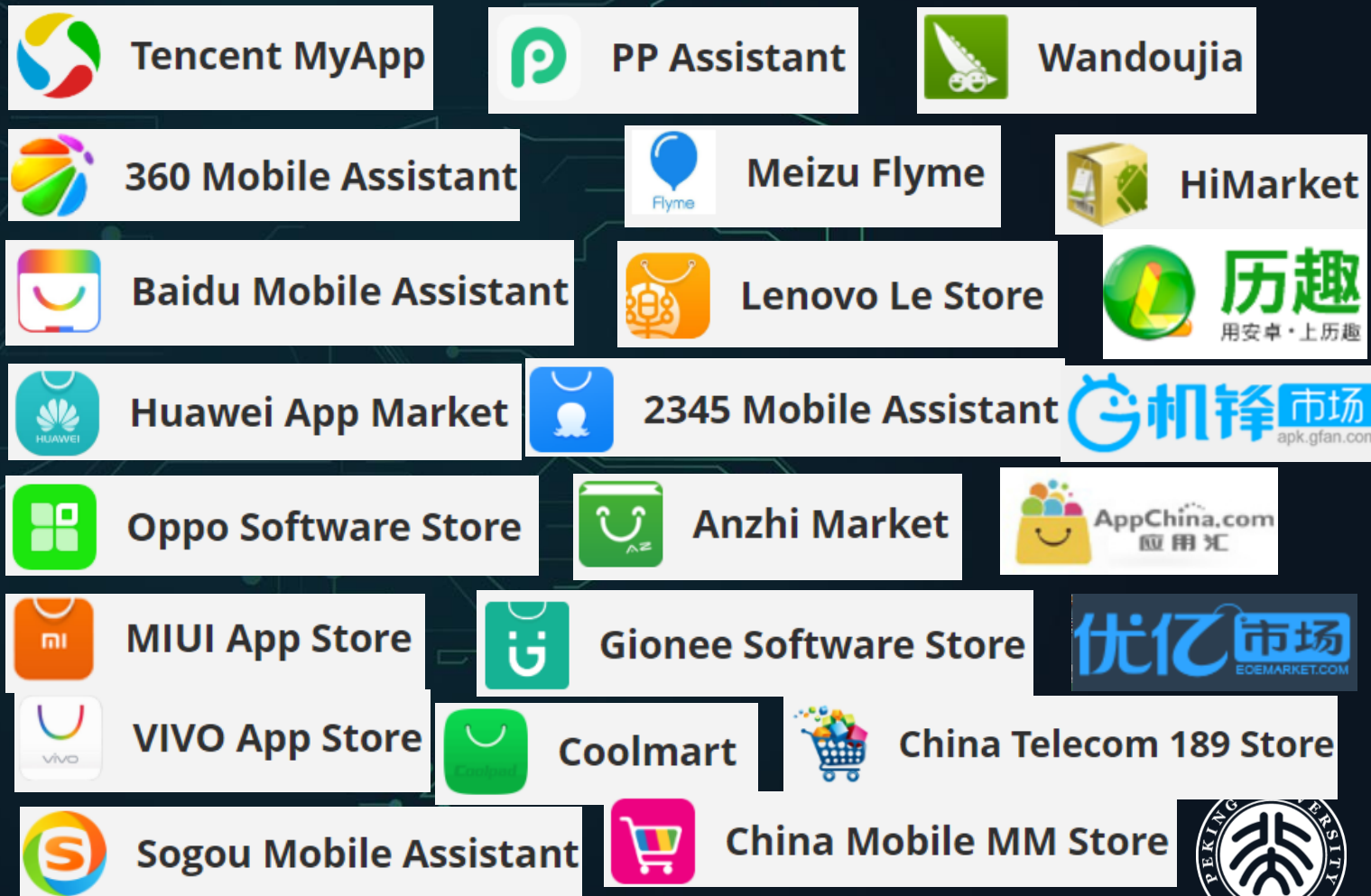
- Beyond Google Play: A Large-Scale Comparative Study of Chinese Android App Markets, IMC'18

Beyond Google Play ...

国外的应用市场

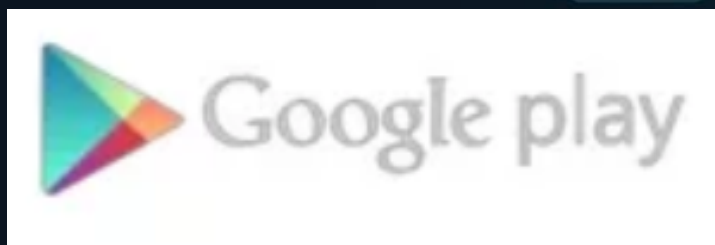


国内的应用市场（们）



数据集：超过620万Android应用

■ 官方应用市场



2.03M

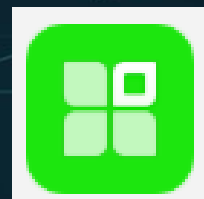
■ 设备厂商应用市场

Xiaomi



91K

Oppo



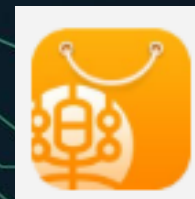
426K

Huawei



51K

Lenovo



37K

Meizu



80K

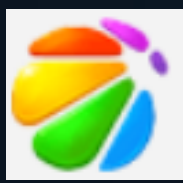
■ 大公司应用市场

Tencent
Myapp



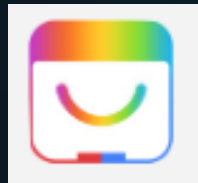
636K

360



163K

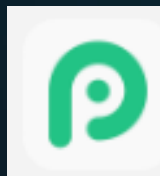
Baidu



227K

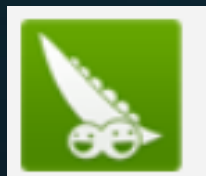
■ 专门的应用市场服务

25PP



1.01M

Wandoujia



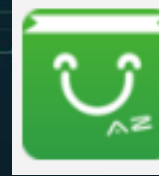
554K

HiAPK



246K

AnZhi



223K

LIQU



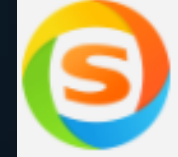
179K

PC Online



134K

Sogou

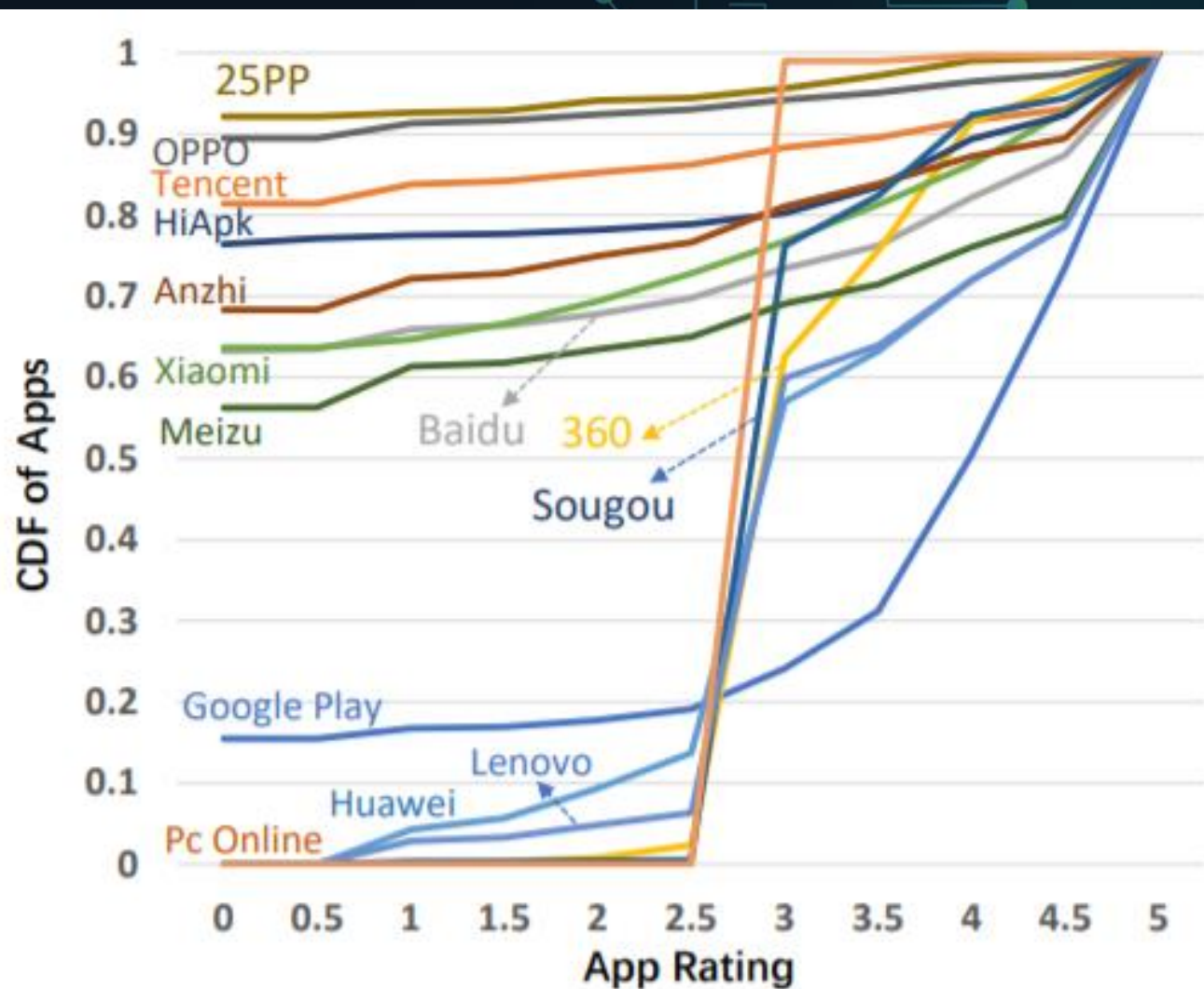


128K

AppChina



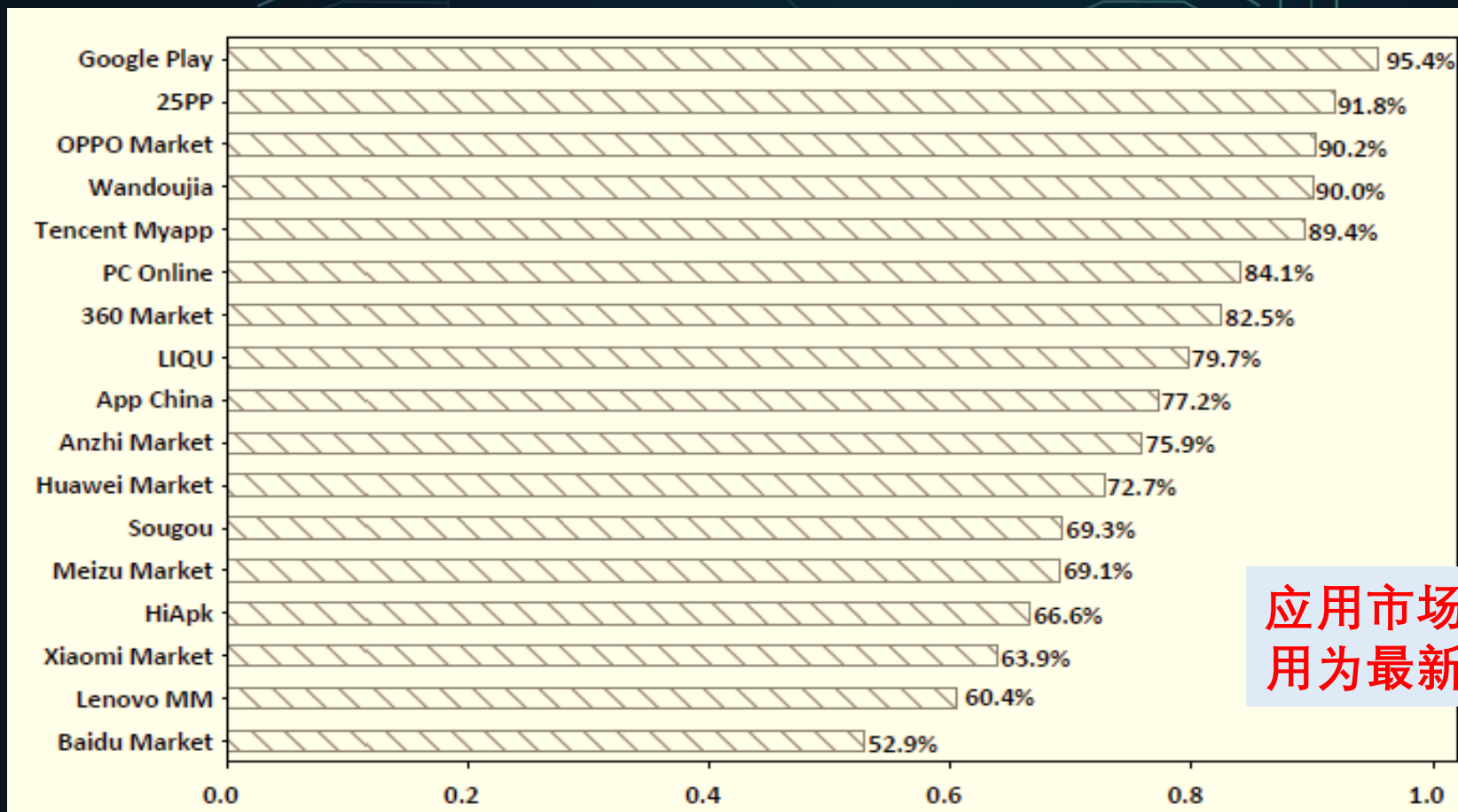
结果概况：每个应用市场的应用评分分布



- 与Google Play相比，国内市场的评分呈现出非常不同的分布
- **模式 #1:** 国内市场中约80%的应用没有任何用户评分
- **模式 #2:** 有些应用市场的缺省评分不是0（比如缺省分数为3）

市场及时性：新版本应用的比例

- 同一个应用在不同市场中存在大量不同版本，很多应用市场版本更新非常不及时！



应用市场中所包含应用为最新版本的比例

- 虚假或仿冒应用检测
 - 利用基于聚类的方式严格匹配应用的名称
 - 利用heuristic规则来删除合法的聚类
- 应用克隆（重打包）检测
 - 基于签名的克隆检测
 - 相同包名（package name），不同签名
 - 基于代码分析的克隆检测
 - 基于我们之前的工作：WuKong [ISSTA 2015]
 - 组合了粗粒度和细粒度检测的两阶段克隆检测技术

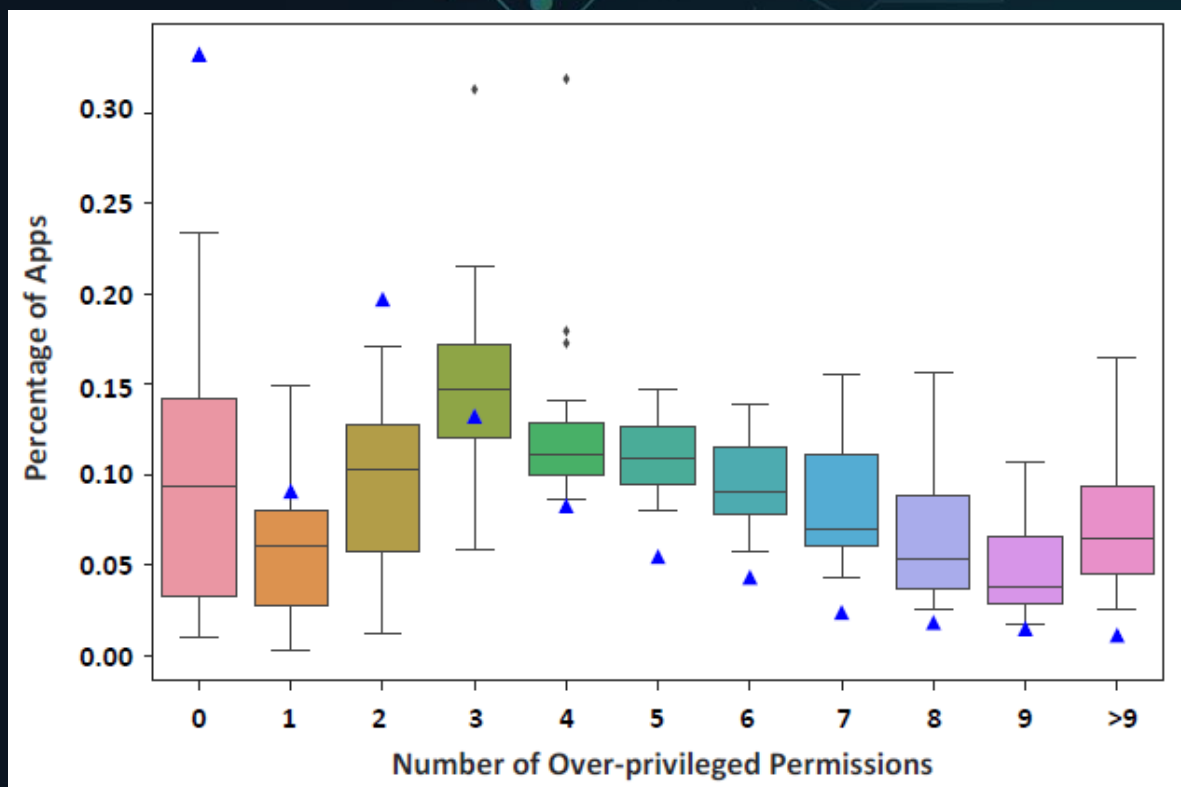
虚假和克隆应用检测(2/2)

| Market | Fake (%) | Clones | |
|---------------|-------------|--------------|--------------|
| | | SB (%) | CB (%) |
| Google Play | 0.03 | 4.01 | 17.82 |
| Tencent Myapp | 0.53 | 8.24 | 22.73 |
| Baidu Market | 0.48 | 10.98 | 17.38 |
| 360 Market | 0.50 | 5.43 | 23.26 |
| Huawei Market | 0.33 | 11.54 | 18.76 |
| Xiaomi Market | 0.0 | 8.00 | 20.11 |
| Wandoujia | 0.39 | 5.98 | 21.23 |
| HiApk | 0.64 | 7.51 | 20.08 |
| AnZhi Market | 0.57 | 4.92 | 20.71 |
| OPPO Market | 0.38 | 5.85 | 20.94 |
| 25PP | 0.35 | 7.16 | 24.08 |
| Sougou | 1.83 | 4.86 | 18.28 |
| MeiZu Market | 1.14 | 6.65 | 18.42 |
| LIQU | 0.40 | 5.32 | 16.68 |
| App China | 0.0 | 10.17 | 13.23 |
| Lenovo MM | 0.67 | 7.81 | 16.37 |
| PC Online | 1.89 | 8.60 | 23.34 |
| Average | 0.60 | 7.24 | 19.61 |

- 虚假应用在所有的应用市场中都存在，包括Google Play
- 基于代码的克隆检测技术更加准确和全面，检测到**约20%的克隆应用**，基于签名的方法只能检测到不到10%的克隆应用
- 克隆应用的流向
 - Google Play是克隆应用的主要来源
 - 25PP应用市场拥有数量最多的克隆应用



- 越权应用(Over-privileged apps): 申请权限超过其必需权限
- 基于代码分析和PScout (Permission map)进行检测。



- 65% 的Google Play应用会申请过多权限
- 在国内市场中越权应用的比例大约是 82%

恶意应用 (malware) 检测

| Market | AV-rank (% apps) | | |
|---------------|------------------|-------|-------|
| | >= 1 | >= 10 | >= 20 |
| Google Play | 17.03 | 2.09 | 0.32 |
| Tencent Myapp | 34.15 | 11.16 | 3.45 |
| Baidu Market | 42.77 | 12.24 | 3.30 |
| 360 Market | 41.40 | 12.35 | 3.10 |
| OPPO Market | 42.97 | 16.43 | 6.00 |
| Xiaomi Market | 55.11 | 9.12 | 1.82 |
| MeiZu Market | 51.40 | 10.70 | 3.14 |
| Huawei Market | 57.48 | 4.71 | 0.57 |
| Lenovo MM | 54.20 | 7.53 | 1.52 |
| 25PP | 32.36 | 8.26 | 2.06 |
| Wandoujia | 31.99 | 7.98 | 2.19 |
| HiApk | 41.89 | 11.12 | 2.72 |
| AnZhi Market | 55.32 | 11.37 | 2.41 |
| LIQU | 45.91 | 13.00 | 4.27 |
| PC Online | 55.93 | 24.01 | 8.37 |
| Sougou | 52.41 | 16.53 | 4.59 |
| App China | 48.55 | 14.13 | 4.27 |
| Average | 36.49 | 12.30 | 3.69 |

- **恶意应用检测**: 利用VirusTotal
 - VirusTotal集成了约60个检测引擎
- **AV-Rank**: 被多少个检测引擎检测为malware
- 以超过10个引擎报告为准:
 - **Google Play中约~2%** 的app为标记为恶意应用malware
 - 16个国内市场中, 超过11个市场包含**超过10%**的malware



恶意应用有没有被及时移除?

- 在8个月之后, 对所有市场中被检测到malware进行了爬取

| Market | %Malware Removed | #Overlapped with GPRM | %Removed |
|---------------|------------------|-----------------------|----------|
| Google Play | 84% | - | - |
| Tencent MyApp | 8.75% | 7,157 | 3.1% |
| Baidu Market | 23.99% | 1,422 | 34.53% |
| 360 | 43% | 1,198 | 34.22% |
| Xiaomi | 32.50% | 636 | 31.13% |
| Meizu | 29.18% | 668 | 26.20% |
| Huawei | 26.92% | 169 | 23.08% |
| Lenovo MM | 22.75% | 263 | 16.35% |
| 25PP | 19.63% | 7,804 | 17.31% |
| Wandoujia | 34.51% | 5,289 | 44.74% |
| AnZhi | 27.61% | 632 | 25.78% |
| LIQU | 14.08% | 1,878 | 11.18% |
| PC Online | 0.01% | 1,117 | 0.00% |
| Sougou | 24.24% | 1,082 | 22.00% |
| App China | 20.51% | 546 | 30.24% |

- Google Play中 ~84% 的 malware 已经被移除
- 在国内市场中, malware 被移除的比例: 从0.01% 到 34.51%
- 针对Google Play已经移除的 AV-rank ≥ 10 的malware, 超过 70% 依然能够在至少一个国内市场中找到

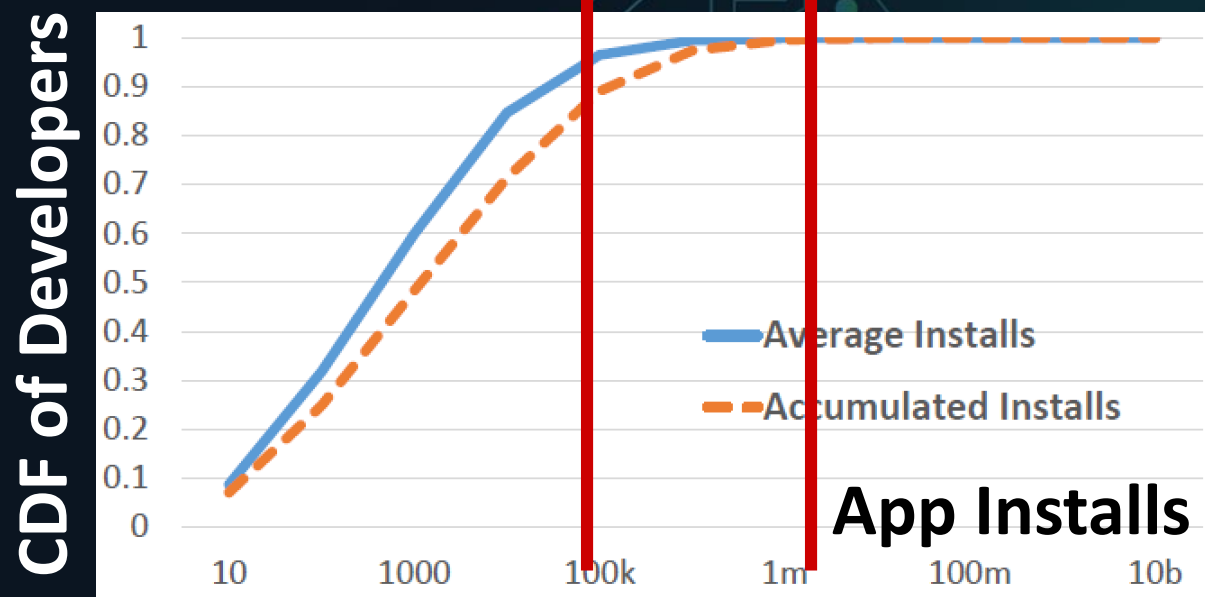


03 章节 PART

移动开发者生态分析

- An Explorative Study of the Mobile App Ecosystem from App Developers' Perspective, WWW'17
- Characterizing the Global Mobile App Developers: A Large-scale Empirical Study, MobileSoft'19
- Characterizing Android App Signing Issues, ASE'19

开发者生态：Google Play

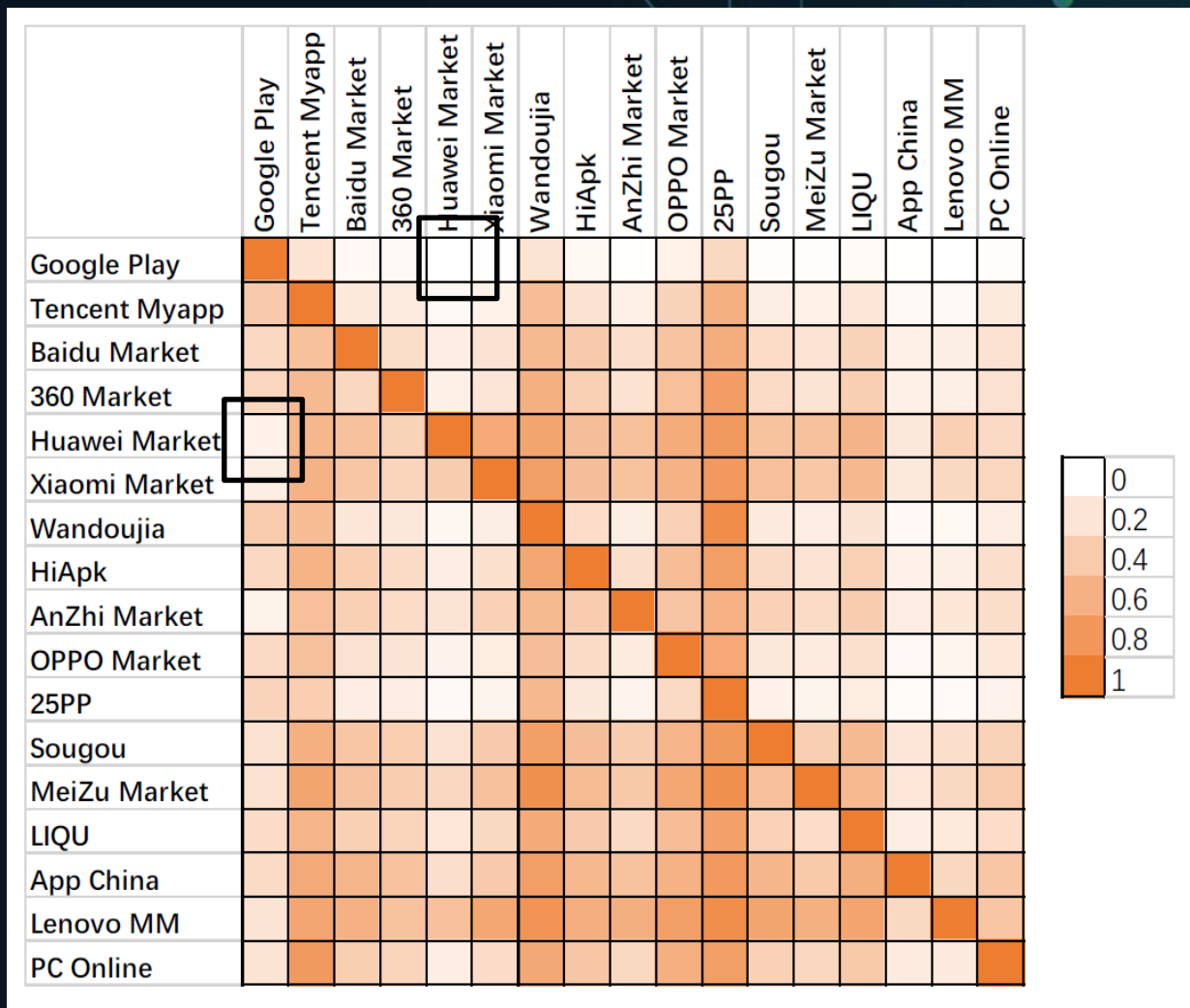


Popular developer: top 1% of developers
(accumulated installs > 3 million)

The top 1% of the developers with the most app installs account for 80% of the total installs

| Developer | #Apps | Total Installs |
|------------|-------|----------------|
| Google | 128 | 13.2 Billion |
| Facebook | 11 | 1.5 Billion |
| Samsung | 20 | 1.2 Billion |
| WhatsApp | 2 | 1.1 Billion |
| Outfit7 | 25 | 817 Million |
| Rovio Ent. | 21 | 591 Million |
| Gameloft | 79 | 570 Million |
| Instagram | 4 | 560 Million |
| Skype | 5 | 511 Million |
| Twitter | 3 | 510 Million |

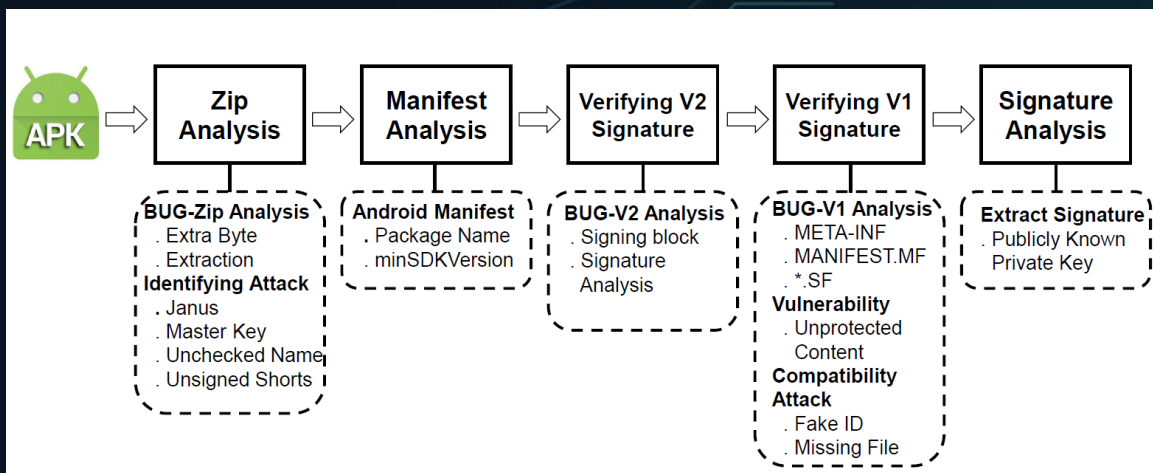
开发者生态：Google Play vs. 国内市场



- 约50%的开发者在Google Play发布app
- 超过70%的开发者在国内市场发布app
- **Google Play和国内市场的开发者之间的重叠比较低**
 - 例如：Huawei 和 Google Play 的开发者之间的重叠度大约只有 1%

开发者在使用签名机制中存在很多问题

Android提供了三个不同版本的app签名机制，
后续版本（V2）修复了之前签名机制（V1）存在的漏洞



- 分析总结了21种与app签名有关的安全漏洞或潜在问题
- 构建了自动化工具识别签名问题

- ~93.7%的app只采用了V1签名，安装在Android 7.0之前的系统会被攻击。
- 在研究的25个市场中，7~45%的app存在不同程度的签名问题。
- 成千上万的app（超过65K）采用公开的密钥进行签名，其中包括许多下载量上亿的app。
- 大约9万的app拥有兼容性问题，在有些设备上无法安装，其中包括下载量超过十亿的app。



04

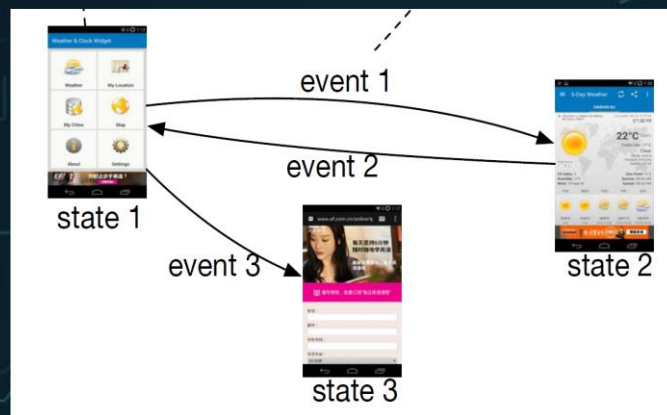
章节 PART

总结与展望

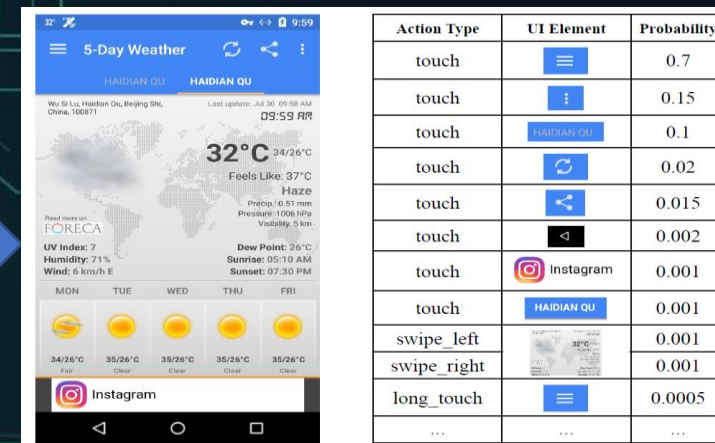
利用UI分析和强化学习，提高App测试输入生成的效率



Monkey:
随机生成点击



Droidbot:
根据UI转换图
提高点击效率



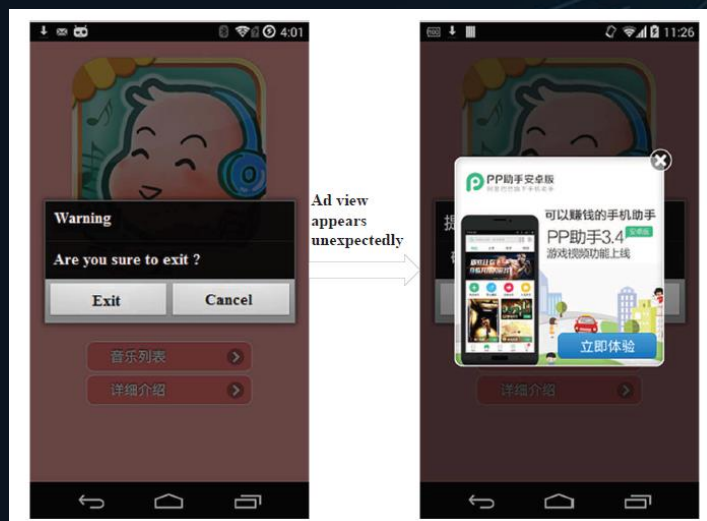
| Action Type | UI Element | Probability |
|-------------|------------------|-------------|
| touch | [Menu Icon] | 0.7 |
| touch | [Info Icon] | 0.15 |
| touch | [HAIDIAN QU] | 0.1 |
| touch | [Refresh Icon] | 0.02 |
| touch | [Share Icon] | 0.015 |
| touch | [Back Arrow] | 0.002 |
| touch | [Instagram Icon] | 0.001 |
| touch | [HAIDIAN QU] | 0.001 |
| swipe_left | [Weather Card] | 0.001 |
| swipe_right | [Weather Card] | 0.001 |
| long_touch | [Menu Icon] | 0.0005 |

Humanoid:
采用深度学习
模仿人的操作

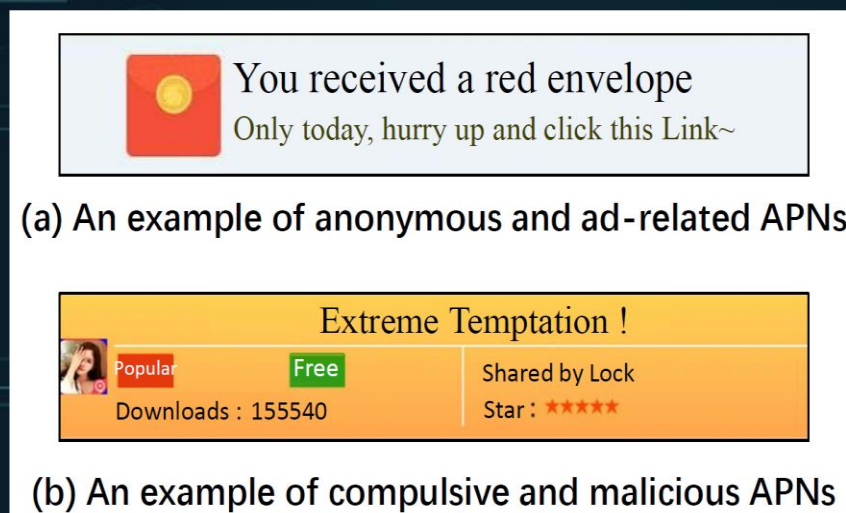
- Droidbot: A Lightweight UI-Guided Test Input Generator for Android, ICSE'17
- Humanoid: A Deep Learning-based Approach to Automated Black-box Android App Testing, ASE'19



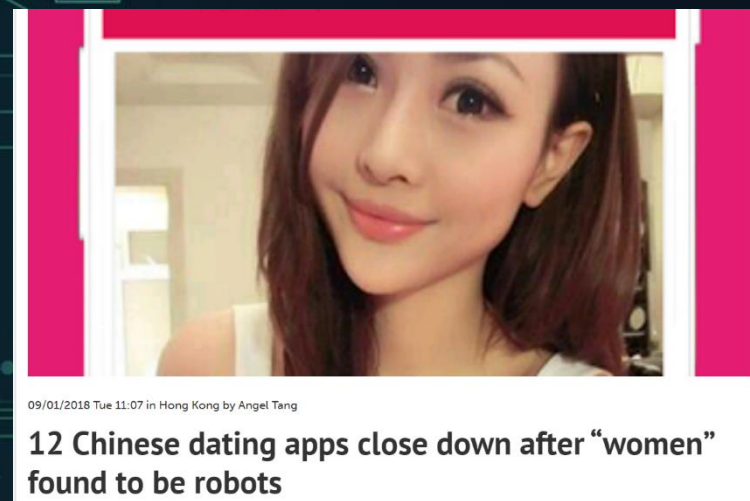
利用App静态分析和动态测试技术查找App中的恶意行为



广告欺诈行为



恶意下推通知



虚假约会App

- DaPanda: Detecting Aggressive Push Notification in Android, ASE'19
- FraudDroid: Automated Ad Fraud Detection for Android Apps, FSE'18
- Dating with Scambots: Understanding the Ecosystem of Fraudulent Dating Applications, TDSC'19



共同努力、改进国内移动应用生态!





THANK YOU

感谢聆听

Contact: yaoguo@pku.edu.cn