Are These Ads Safe: Detecting Hidden Attacks Through the Mobile App-Web Interface

JAN 15TH, 2016

论文下载：http://pages.cs.wisc.edu/~vrastogi/static/papers/rscpzr16.pdf

主要针对从APP里引出到web（所谓的App-Web）的虚假和欺诈广告问题，做出检测和现状调查说明。同时本文也开发了一个框架能够大规模的对APP进行动态测试并记录广告链接和最终链接（广告内容或者下载的文件等），进行恶意链接或者文件的判断。

主要贡献点：

- 开发了框架能够分析从APP引出到WEB的接口。主要分成触发接口，检测恶意内容和对这些恶意内容的追根溯源。整个系统相当程度的自动化。
- 触发系统可以和UI控件交互并且有基于计算机图形学的算法辅助查找能够点击的元素。
- 查找广告网络对应的广告代码包名，并找到了201个。

- 系统异常稳定运行了两个月在两所大学对中美的市场做了超过60万个APP的检测。

方法

触发系统

基于动态分析，他们搞了模拟器跑。使用了AppsPlayground里使用的启发式算法，抽取UI的特征并建立APP里UI转换的状态机，并识别避免重复的窗口或者activity。另外处理webview无法使用上面方法，借助Selendroid这个项目可以获取webview页面里的内容，因此作者使用了基于计算机图形学的算法来识别出按键，先识别出图像中的边，然后找到轮廓线，找到封闭的凸曲线，计算边界之类的。目的就是找到可以点击的图片或者按键。他们改进后的算法可以比Selendroid更加有效，测试发现能够多触发5倍的链接。

检测系统

记录访问的链接，重定向链，以及最终导向的广告页。由于广告链接会有联盟转包等等所以通常会有很长的重定向链并有多种实现，如js里，HTML meta tags，HTTP301/302等等，所以作者修改并自己实现了个基于webview的浏览器去记录整个重定向链。最终广告页里作者也各种点击来保证如果有文件下载就能得到。然后把这一系列的链接和得到的文件送到virustotal里去检测，链接检测有多个服务商，文件检测也是多个杀毒软件厂商。

追根溯源

如果最终页是恶意的，那么重定向链中导向最终页的那个链接就是有问题的。APP里也会发一个含链接的intent让浏览器去访问广告页，作者修改并记录这个intent来保证如果有恶意URL的话，能够追踪到APP。如果是APP里，那么需要进一步

识别是否是APP开发者自身代码，还是APP引入的广告库造成的。因此需要有个广告代码库和广告网络的对应关系。识别广告代码首先利用APP的相似性找到大家都会包含的包名，一般就是广告库。进一步的利用代码相似性的方法去识别广告库，找到APP里松耦合的部分做特征然后匹配。通过这两种方式找到了201个广告库，第二个方式还能抗类名混淆。

具体实现

Python写的。AppsPlayground用来触发系统，做了些许修改来适配现在版本的系统，如用UIAutomator。为了加快动态分析，利用了KVM-accelerated virtualization。基于X86平台所以native APP不管了大概是全部的30%不分析。任务管理调度基于celery，后台数据库用MySQL。另外浏览器自动点击和触发最终广告页的行为基于Chromium并使用了Watir和Selenium Webdriver框架，所有处理都是无需GUI的，使用了Xvfb，可以不需要屏幕输出。每个APP最多跑5分钟，平均少于2两份，最终页测试最多跑15分钟。

实验结果

492534个play上的APP以及422505个中国市场APP从91，安智，AppChina和木蚂蚁，爬play用了PlayDrone，中国的市场随便爬。美国西北大学服务器专门来爬play的，中国浙大爬中国市场，全自动从去年4月到6月跑了两个月。美国跑出1百万个链接，948个恶意URL，来自64个域名。中国跑出41万5千个，1475个恶意URL来自139个域名。美国下载了468个文件其中271个是恶意的，其中主要是一个搓比假冒伪劣杀毒软件。除去这个特殊案例，6分之一的最终页下载的APP是恶意的，当然这些文件不包括广告库导向Play的那些链接。导向Play要求下载的达到了433000个，集中到19000个要求被下载

的APP，其中5%是恶意的。其中大多数是被认为是adware。中国这边1097个下载文件其中435个恶意，102个是那个假AV软件。

案例

假杀毒软件 Armor for Android通过Tapcontext广告商。APP里显示这个AV的广告，然后有时候广告页就开始显示帮你扫描病毒了，然后点了如果下载，，广告页会变成非常类似本地Android系统安装的界面来诱骗你安装，但每次下载MD5值不一样，这个APP貌似付费，被很多杀毒软件公司认为是有问题。一开始play上还有后来下架了。

免费送ipad欺诈。很多导向送iphone或者ipad的页面，但要求你填入个人信息，最后还要你下载APP或者浏览器工具栏什么的。其实并不会真的送（作者猜的，因为他没填真实的信息）。但这些URL并不被virustotal认为是有问题的。但其中重定向链中某些URL被virustotal检测出有问题，但作者人工验证发现这些域名下的内容基本没啥大问题，广告库是Mobclix和Tapfortap，都是那种广告联盟什么的，广告交换啊，分包什么的所以中间有很多不可控因素，整个产业链中从开发者到广告商都没法保证内容的安全可靠。

直接连接的欺诈。有些欺诈链接由直接写在APP里的URL导向而非APP里的动态广告链接，APP里链接导向某些友好链接，但友好链接里有重定向的广告导向欺诈。这个链接是直接写在APP代码里而非APP里的广告库。

恶意软件。诱使用户下载播放器软件，而软件是木马。中国百度和Nobot都有直接连接让用户下载木马，另外这两个广告网络还发布一针见效类似的虚假医药广告。

一些google的广告也有些可疑但不能确定的案例。