Lane Gramling
EECS 700
Reading 11

Summary - What Mobile Ads Know About Mobile Users

This paper concerns itself with exploring the structure of the modern ad serving ecosystem that dominates advertising on Android. In their exploration, they provide a number of details on possible exploitations within in the realm of permissions afforded to advertisement SDKs in Android. Of these, they emphasize an issue where advertisement libraries can infer information about the user by accessing clues about external data, all from within the confines of what is afforded to these libraries for the purpose of serving media-rich advertisements.

Contributions

The authors begin by outlining the structure of the most popular advertising libraries in use on the Google Play store. Of these, they select AdMob, MoPub, AirPush, and AdMarvel as their ad SDKs of choice for studying the mobile advertising ecosystem and Android's advertising software stack. Most crucial to the authors' key idea in their outline is that the power to set file access for ad-containing WebViews lies under the arbitration of the AdSDK, because the AdSDK services also request that developers using their services request the READ_EXTERNAL_STORAGE permission. This permission is what places such power in the AdSDK's hands.

This information is what leads the paper to its first major contribution, which is the establishment a number of inference mechanisms enabling advertisements to gather potentially sensitive data simply by exploiting the standard permissions afforded to advertisers. Firstly, the authors note how AdSDKs can change the default setting of ad-containing WebViews using setAllowFileAccessFromFileURLs or setAllowUniversalAccessFromFileURLs without user permission, enabling an ad to download potentially malicious HTML files capable of stealing files through a series of JavaScript shenanigans. A second notable inference mechanism that was shown is the capability to reverse engineer location data of a user to link old Google Advertisement IDs with new GAIDs, effectively making the ability to change GAIDs useless. Other demonstrated inferences enable finding information about user medications, dating partner information, browsing history, and social groups.

Limitations

This paper, in my opinion, does not contribute much in terms of exploring potential defenses against their inference mechanisms - hence why defenses are left out of the contributions section entirely. As noted by the authors, there is little an app developer can do if they wish to keep using advertisements (and thus keep making money). Additionally, the realm of possibility for their inference mechanisms seems to exist at the intersection of a number of conditions, which at first glance feel far less likely to be true simultaneously than the paper implies. Granted, this could be a common set of conditions, but little to no information on the prevalence of this issue is given, so I am left unsure of whether this is a widespread issue, or if the paper has found itself exploring a niche case only shrinking in probability. More data on this would have considerably substantiated how severe of an issue their findings are.

Future Work

While the limitations of this paper feel glaring, it does leave plenty of room for further exploring the issues enumerated above, such as gaining insight into how common this problem is, or how to come up with better defenses. Personally, the defense outlined on the OS layer feels the most feasible, and perhaps worth expounding upon in future work. Regardless, this paper provides great clarity on what kind of things malicious advertisers may be capable of learning about or otherwise exploiting users. w