Impact of Ad Libraries on Ratings of Android Mobile Apps

Israel J. Mojica Ruiz, McAfee

Meiyappan Nagappan, Rochester Institute of Technology

Bram Adams, École Polytechnique de Montréal

Thorsten Berger, University of Waterloo

Steffen Dienst, University of Leipzig

Ahmed E. Hassan, Queen's University

// App developers use several third-party ad libraries to monetize their apps. This article examines the relationship between the number of ad libraries in Android apps and the user ratings of these apps. //

MOBILE APPS ARE software applications for mobile devices, such as smartphones, tablets, and other personal digital assistants. The growing demand for mobile apps has led to rapidly increasing downloads—from 7 billion apps in 2009 to an estimated 102 billion app downloads across all platforms in 2013 (www.gartner.com/newsroom/id/2592315). This rapid growth is attracting many amateur and professional developers, who strive to profit from developing such apps. Developers commonly monetize their apps by displaying advertisements to end users. These advertisements are provided by third-party ad companies such as Google Mobile Ads (https://developer.android.com/google/play-services/ads.html) and Flurry's AppSpot.

The primary method to integrate ads in a mobile app is through the use of one of the ad company's ad libraries. However, the success rate in receiving an ad from an ad company when an ad is requested (*fill rate*) is low when it comes to mobile app ads. In the first half of 2011, the average fill rate for the top 40 ad networks was less than 18 percent. This low fill rate is mainly because the number of ads being requested by apps is increasing faster than the number of ads available in the market.[1] Hence, to achieve a high fill rate, app developers might need to integrate multiple ad libraries from different ad companies. Developers aren't restricted to a single ad company, so connecting with a large number of ad companies helps ensure a higher fill rate and, hence, higher revenues.

In this article, we empirically examine if a relationship exists between the number of ad libraries integrated in an app and the app's user rating. Earlier work looked at the relationship between software economics and software quality as influenced by software evolution[2] and examined how to price a particular app.[3] As far as we know, however, no research study has examined the number of ad libraries commonly integrated in an app or if integrating several different ad libraries impacts an app's rating. This article examines these two issues by empirically examining thousands of Android apps. We strive to give actionable recommendations to app developers about the effective use of ad libraries in apps so they can maximize revenue without affecting an app's rating in the app market.

Our research indicates that app developers indeed often integrate more than one ad library in their apps—in fact, as many as 28 ad libraries—and that the number of ad libraries in an app isn't related to its rating. We did find, however, that integrating certain specific ad libraries can negatively impact an app's rating.

Study Data Collection

Similar to other work on mobile apps,[4–6] we analyzed apps from the popular Android platform. In particular, we analyze 519,739 app versions distributed in 236,245 different Android apps, a superset of the apps used in one of our prior studies (208,601 apps) on reuse in Android apps.[6] Our dataset covers 27 Google Play categories, with the entertainment category having the most apps (more than 20,000) and the weather category having the lowest number of apps (more than 1,000). Here we describe the study design of our work, from data collection to preprocessing, so others can replicate our work.

Crawling Google Play

Our study uses a dataset that we crawled from the official Google Play app store once a day during the first half of 2011 and twice a day during the second half of 2011. As a result, we obtained a set of 625,067 app versions distributed across 281,079 mobile apps. Google Play classifies apps into 27 different categories (including game subcategories).

Adding an extra ad library does not necessarily imply lower app rating.

For each app version that was free to download, our dataset contained two elements:

- the binary in the Android-specific packaging format (APK) and
- app-store-specific metadata information for each app version.

For the present work, we only use each app's average user rating (from one to five stars) metadata. See the technical note on this work for more details.[7]

Extracting App Bytecode

App developers integrate ad libraries into the APK in order to use the ad libraries' APIs. For our study, we needed to identify which ad libraries were integrated in each app. Given that apps are packaged in the APK format, we only have access to the app's Dalvik bytecode, which contains both application and library classes.

We first used an open source tool (dex2jar; http://code.google.com/p/dex2jar) to extract the Java bytecode from the APKs. Then, we used the Apache BCEL library (http://commons.apache.org/bcel/) to extract the fully qualified class names (package or namespace in which a class is contained and the class name) and their corresponding set of method names (APIs) for each class, in each app. We dropped apps with obfuscated class names.

After this process, we ended up with 519,739 app versions of 236,245 different apps. We stored the fully qualified class names and their corresponding methods in a database. We then used this data to identify the ad libraries in each app.

Identifying Ad Libraries

We filtered the fully qualified class names of all the apps to look for the regular expression [aA][dD] (such as com.packageAdlibraryName.AdclassName) across all the apps in the database. Notice that the regular expression [aA][dD] is simple. Hence, it matched many class names across the database, even class names that weren't necessarily part of an ad library.

We grouped and sorted the fully qualified classes according to their popularity in the database. The most frequent, fully qualified class was com.google.ads.AdActivity, with 149,321 repetitions. We performed the following manual process to find the ad libraries:

1. For each fully qualified class name, perform a Web search for the package name to find the website of the ad library provider

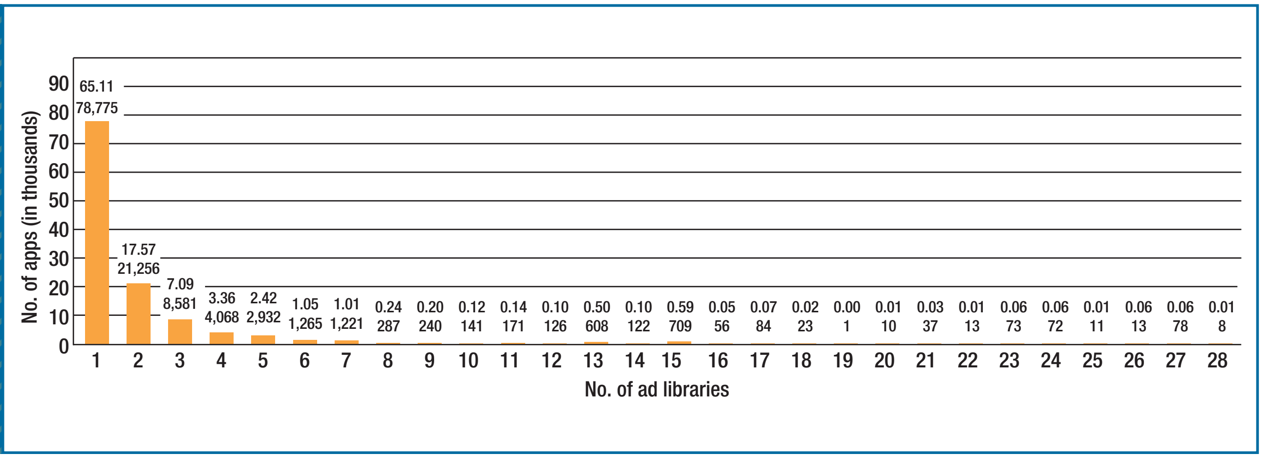

FIGURE 1. Breakdown of the percentage (top) and total number (bottom) of apps (*y*-axis) with ad libraries based on the number of integrated ad libraries per app (*x*-axis).

(the ad company) and verify that the name of the library found is a real ad company.

2. If an ad library website is found, add the package name to the set of ad libraries. Otherwise, discard it.
3. Filter out from the database all the class names that belong to the package name identified in step 1.
4. Repeat this process for each remaining class name.

For step 1, we assumed that information about an ad library would exist online if it was a real, trusted ad library because app developers must sign contracts with the ad companies to receive ads and later payments.

We repeated this process until a class name was repeated no more than 200 times in the database. This means that any library not integrated would only occur in fewer than 200 apps (0.08 percent of the studied apps). We stopped at this point because there were still thousands (802,012) of fully qualified class names to verify that wouldn't necessarily be part of an ad library. If they were, the presence of such an ad library would be in a minuscule set of apps.

More than half (51.21 percent) of the studied Android apps (236,245) had at least one ad library. We identified 72 different ad libraries that were used by apps in our study and obtained the unique package name that identifies each of these ad libraries.

Integrating Multiple Ad Libraries

Because the fill rate for ads in mobile apps is so low,[1] developers may choose to integrate more than one ad library in an app. Thus, our first step was to determine the number of ad libraries that are integrated in each app in our dataset.

We found that most apps only integrate one ad library. However, at least 42,206 (34.88 percent) of the apps with ad libraries in them have two or more ad libraries. Figure 1 shows a breakdown of the number (and percentage) of apps based on the number of integrated ad libraries. We found eight apps that integrate as many as 28 ad libraries.

Relationship between Ad Libraries and App Rating

One key distinction between ad libraries and other libraries (such as utility or interface libraries) is that ad libraries aren't required for an app to function correctly: they are exclusively for monetizing the app. From the previous section, we know that thousands of apps integrate more than one ad library. Integrating more ad libraries in an app increases the maintenance effort for the developers, as does adding any other piece of code. Thus, an increased number of ad libraries could impact an app's quality. Consequently, we examined the relationship between the number of ad libraries in an app and its rating. Past research has shown that ratings are highly correlated with the number of app downloads,[8] which in turn is a concrete measure of success.

In Google Play, app users can rate an app from one to five stars (where five is the highest value). Because an app's rating is subject to rater bias,[9] the number of raters is an important factor to consider. Hence, to minimize this rating bias, we limited our analysis to ad-supported app ver-

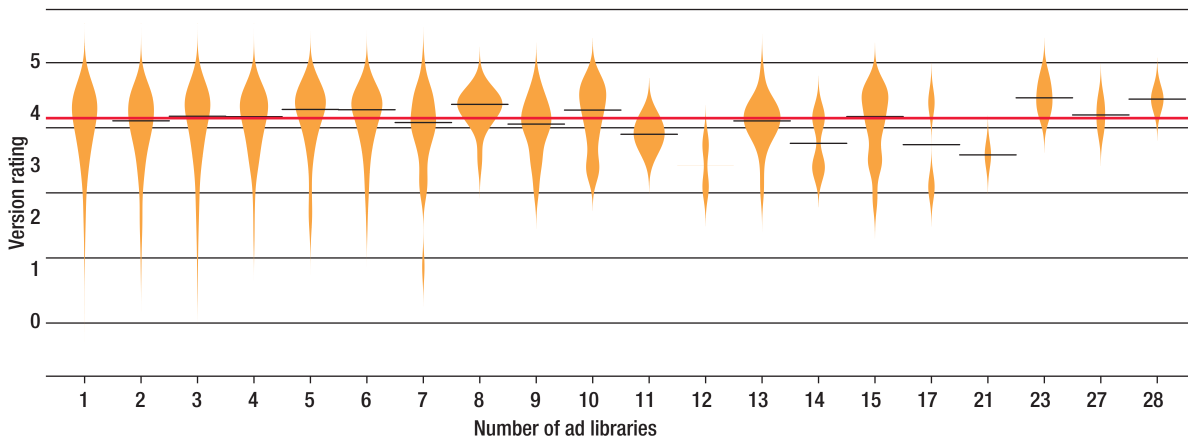

FIGURE 2. Relation between the number of ad libraries and the corresponding version rating. The beanplots are sorted in ascending order by the number of ad libraries. The horizontal red line represents the median of the app version rating of all the apps taken together.

sions with at least 10 raters. We also chose to study apps with at least two versions in 2011 because they represent apps that are actively maintained. Our study criteria of apps with at least one ad library, more than one version, and at least 10 raters per version resulted in 13,983 versions distributed across 5,937 different apps.

Finally, given that some apps have more versions than others, using every version of an app could affect our results. Because each app should have the same importance, we used only the latest version of each app. We present the results of the app ratings based on the number of ad libraries within each app.

The Spearman rank correlation between the number of ad libraries in an app and its rating is 0.016. Such a weak correlation illustrates that there is no relationship between the number of ad libraries in an app and its rating. In Figure 2, we break down the data and show more details by means of a beanplot of the apps' ratings grouped by the number of integrated ad libraries. Beanplots show both the median value for the rating (indicated by the solid horizontal line) and the actual distribution (the width of the curve at any point in each beanplot indicates the number of apps with a particular rating). The beanplots are ordered by the number of ad libraries contained in each group of apps (x-axis). The y-axis indicates the app's rating.

We can see from Figure 2 that the median rating for all apps taken together is more than four stars. Apps with a higher number of ad libraries (the beanplots at the right end) tend to have only a slightly lower median rating, except for the apps with 23, 27, and 28 ad libraries, which have a rating higher than the median. However, these groups represent only a small percentage of all apps. Most app versions with between seven and 21 ad libraries have a rating below the median.

Next, we examine the eight apps that have 28 ad libraries. Five out of these eight apps come from the same app developer, and the other three are by two different app developers. Three of the apps were removed from Google Play for undisclosed reasons. We verified the rating and rater numbers for the remaining five apps to determine if their users are unhappy because of the high number of ad libraries. Surprisingly, those apps still have similar ratings (a high average rating) and continue accumulating positive feedback. For example, the app with the largest number of raters of these five apps has 1,489 raters and an average rating of 4.3 out of five, with 946 users giving five stars. Interestingly, the number of users continues to increase in spite of having a large number of ad libraries. Note that these apps specialize in displaying pictures and are classified as "mature only for 18+ viewing."

One possible explanation for our finding that the number of ad libraries

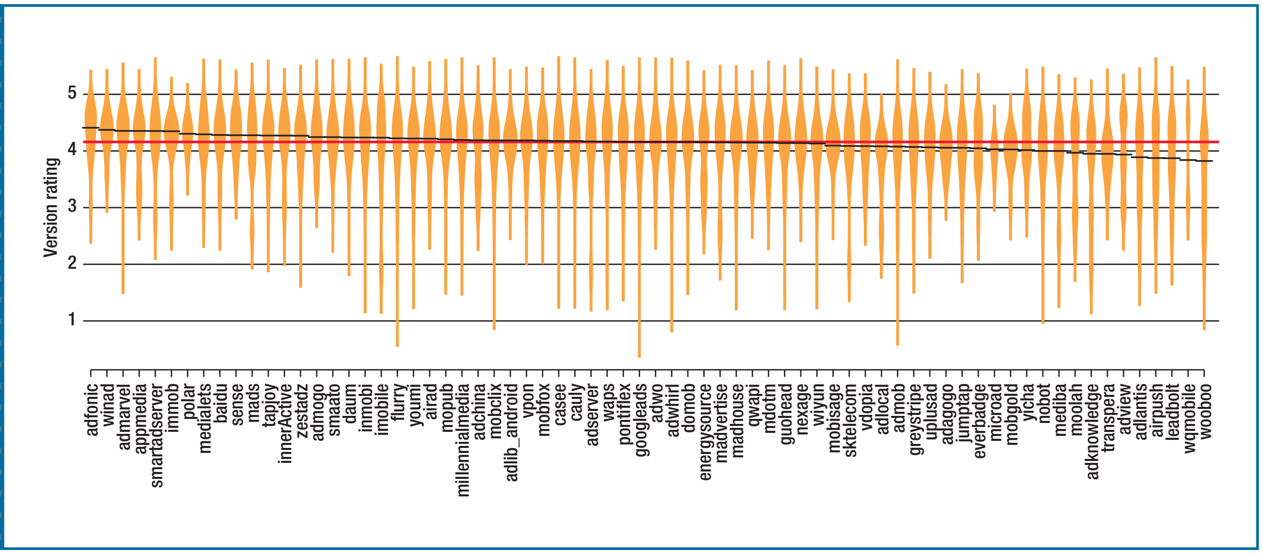

FIGURE 3. Distribution of the ratings of the apps that contain each of the ad libraries. The beanplots are sorted descending by the median app rating. The red and black lines across each beanplot represent the median rating of all the apps and the median rating of the apps that integrate a particular ad library, respectively.

doesn't impact an app's perceived quality is that in practice a high number of ad libraries doesn't mean that an app displays more than one ad at the same time. Instead, the use of a large number of ad libraries is simply a monetization strategy used to achieve a high fill rate.

Relationship between Specific Ad Libraries and Ratings

A key goal in advertising is to attract user attention. As such, ads are often part of an app's GUI. Hence, app developers must consider how users react to the displayed ads to keep their app users happy. Otherwise, app users will rate their apps poorly. App developers might have to be cautious about the specific ad libraries that they integrate in their apps. Thus, we studied the relationship between particular ad libraries and the ratings of apps that integrate the particular ad library.

To do so, we used the same ad-supported multi-app-version dataset, which we've grouped by the type of ad libraries that the apps use and queried for the rating of each app in its last version. For example, consider a three-star app *X* that integrates *ad1* and *ad2* ad libraries, a four-star app *Y* that integrates *ad1* and *ad3* ad libraries, and a 3.5-star app *Z* that integrates *ad1*, *ad2*, and *ad3* ad libraries. In this case, the ratings for each ad library are *ad1* = {3.0, 4.0, 3.5}, *ad2* = {3.0, 3.5}, and *ad3* = {4.0, 3.5}.

Figure 3 presents beanplots for 70 out of the 72 ad libraries in our studied dataset. (The adHUBS and Mobus ad libraries weren't integrated in the last version of any of the studied apps.) The beanplots present the ratings for all the apps that integrate each specific ad library, sorted in descending order by the median rating. The dotted and solid lines across each beanplot represent the median rating of all the apps and the median rating of the apps that integrate a particular ad library, respectively. The beanplots show a slight decrease, depending on the type of ad libraries. Most of the apps in the dataset with a specific ad library are higher or slightly lower than the overall median rating (4.15).

We examined three ad libraries from Figure 3 that have apps with lower ratings to determine the possible causes of the low ratings. The first, Wooboo (www.wooboo.com.cn), is an ad network based in China. This company also develops its own apps. We found that past research has flagged the Wooboo ad library as spyware.[7,9] One app user complained about an app with this ad library: "This is a direct copy of another app, it displays ads and now your password belongs to someone in China."

The second ad network, Leadbolt (www.leadbolt.com), is based in Australia. This ad library uses an intrusive technique called *push notification*, where the ad library pushes ads into the notification bar. This

ABOUT THE AUTHORS

ISRAEL J. MOJICA RUIZ is a software engineer at McAfee. His research interests include mobile software and empirical software analysis. Mojica received an MS in computer science from Queen's University. Contact him at israel_mojica@mcafee.com.

MEIYAPPAN NAGAPPAN is an assistant professor in the Rochester Institute of Technology's software engineering department. His research interests include deriving solutions for software system stakeholders and using large-scale software engineering data to address the concerns of software operators, build engineers, and project managers. Nagappan received a PhD in computer science from North Carolina State University and was a postdoctoral fellow in the Software Analysis and Intelligence Lab (SAIL) at Queen's University. Contact him at mei@se.rit.edu.

BRAM ADAMS is an assistant professor at the École Polytechnique de Montréal, where he heads the Maintenance, Construction, and Intelligence of Software (MCIS) lab. His research interests include software release engineering, software integration, software build systems, software modularity, and software maintenance. Adams received a PhD in computer science engineering from Ghent University. He was an organizer of the First International Workshop on Release Engineering (RELENG 13) and is a member of IEEE. Contact him at bram.adams@polymtl.ca.

THORSTEN BERGER is a postdoctoral fellow in the Generative Software Development Lab at the University of Waterloo. His research interests include model-driven development, variability modeling for software product lines and software ecosystems, variability-aware static analyses of source code, and mining software repositories. Berger received a PhD in computer science from the University of Leipzig. Contact him at tberger@gsd.uwaterloo.ca.

STEFFEN DIENST is a doctoral student in the Chair of Business Information Systems at the University of Leipzig. His research interests include using machine learning in the renewable-energy domain. Other interests comprise static analysis, software reverse-engineering, and mining software repositories. Dienst received an MS (Dipl.-Inf.) in computer science from the University of Leipzig. Contact him at sdienst@informatik.uni-leipzig.de.

AHMED E. HASSAN is the Canada Research Chair (CRC) in Software Analytics and the NSERC/BlackBerry Software Engineering Chair at the School of Computing at Queen's University, Canada. His research interests include mining software repositories, empirical software engineering, load testing, and log mining. Hassan received a PhD in computer science from the University of Waterloo. He spearheaded the creation of the Mining Software Repositories (MSR) conference and its research community, and he serves on the editorial boards of *IEEE Transactions on Software Engineering*, *Empirical Software Engineering*, and *Computing*. Contact him at ahmed@cs.queensu.ca.

results in ads being displayed even when the app isn't running (www.leadbolt.com/blog/earn-more-with-notifications). In 2012, Leadbolt added a new type of ad called *app icon* that installs new icons on the Android device even without the app user's authorization. In addition, app stores have problems with this ad library because it sends the device ID in plaintext.[4]

The third ad network, Airpush (www.airpush.com), is based in the US. Similar to Leadbolt, this ad library uses push notification and app icon ads to serve advertisements. In fact, this ad library has become controversial among app users for its intrusive behavior.[10]

We observed that the apps with highly intrusive ad network behavior resulted in complaints from app users, and in some occasions, such behavior ruined the app users' perceived quality of the app. We found comments such as "Good game, bad ads. I was loving the game until I noticed it put a new shortcut called 'Apps' on my launcher. Sorry, but if your idea of advertising is putting sh*t in launcher pages or notifications then I'm not interested. Keep ads INSIDE the app." Another user complained, "Good game, bad ads. This game is awesome but the ads … embarrising (sic). The ads totally ruin it and there isn't a way to take them off : (."

Ad libraries with security issues and/or intrusive behavior may result in apps getting low ratings, disapproval from users,[11] and even rejection from different app stores. Recently, Google Play updated its policies to encourage app developers to be more conscious about ad libraries that have highly intrusive behavior. Thus, app developers now need to perform ad maintenance tasks in order to align their apps with the new policies in Google Play.[12]

Our case study shows that app developers integrate many ad libraries in their apps, some of which can have a negative impact on the app's rating. The negative impact doesn't seem to be related to the actual number of ad libraries in an app. Indeed, when considering the possible motivation of app developers for adding certain ad libraries that could potentially hurt their ratings, it's obvious that although such libraries are intrusive, they also provide the largest amount of payout for each ad clicked or viewed. For example, we found that ad libraries such as Airpush offer a larger payout than Admob and other conventional ad networks.[13] Hence, it's important for developers to perform a thorough cost-benefit analysis to determine whether the increased revenue per click is worth the possible low rating (and possibly low future downloads of their app). If so, it might make sense to integrate ad libraries such as Leadbolt and Airpush that utilize more intrusive strategies. Otherwise, app developers should avoid those libraries.

Given a certain real estate space that the ads will occupy on the screen of a device (control variable), app developers can add as many ad libraries as needed to increase the fill rate without impacting their ratings. However, developers need to be careful and selective about the specific ad libraries that they choose to integrate.

References

1. L. Niemann, "Smaato Releases Q2 2011 Mobile Metrics Report," 2011; www.smaato.com/metricsq22011-2/.
2. M. Godfrey and D. German, "The Past, Present, and Future of Software Evolution," *Proc. Frontiers of Software Maintenance* (FoSM 08), 2008, pp. 129–138.
3. E. Platzer, "Opportunities of Automated Motive-Based User Review Analysis in the Context of Mobile App Acceptance," *Proc. Central European Conf. Information and Intelligent Systems* (CECIIS 11), 2011, pp. 309–316.
4. W. Enck et al., "A Study of Android Application Security," *Proc. 20th Usenix Conf. Security* (SEC 11), 2011, p. 21.
5. M.C. Grace et al., "Unsafe Exposure Analysis of Mobile In-App Advertisements," *Proc. 5th ACM Conf. Security and Privacy in Wireless and Mobile Networks* (WISEC 12), 2012, pp. 101–112.
6. I.J. Mojica et al., "A Large-Scale Empirical Study on Software Reuse in Mobile Apps," *IEEE Software*, vol. 31, no. 2, 2014, pp. 78–86.
7. S. Dienst and T. Berger, "Static Analysis of App Dependencies in Android Bytecode," version 2.0, tech. note, 2014; www.informatik.uni-leipzig.de/~berger/tr/2012-dienst.pdf.
8. M. Harman, Y. Jia, and Y. Zhang, "App Store Mining and Analysis: MSR for App Stores," *Proc. 9th IEEE Working Conf. Mining Software Repositories* (MSR 12), 2012, pp. 108–111.
9. M. Chen and J.P. Singh, "Computing and Using Reputations for Internet Ratings," *Proc. 3rd ACM Conf. Electronic Commerce* (EC 01), 2001, pp. 154–162.
10. C. Ionescu, "Airpush Begins Obfuscating Ad Modules," blog, 2012; www.symantec.com/connect/blogs/airpush-begins-obfuscating-ad-modules.
11. E. Chien, "Madware Makers Mixing Up Their Code," 2012; www.mobilesecurity.com/articles/285-madware-makers-mixing-up-their-code.
12. "Google Play Developer Program Policies," Google, 2014; https://play.google.com/about/developer-content-policy.html
13. Airpush, "How Is Airpush Different from Admob and the Other Conventional Mobile Ad Networks?," 2014; www.airpush.com/resources/faqs.