IAB Podcast Ad Metrics Guidelines

Released September 6, 2016

This document has been developed for the IAB Audio Committee

The Introduction to Podcast Ad Measurement was created by a working group of volunteers from 23 IAB member companies.

The Podcast Technical Working Group was led by:

- Rockie Thomas, Adswizz
- Ilia Malkovitch, Google

The following IAB member companies contributed to this document:

- AdGear Technologies, Inc.
- AdLarge Media
- Adswizz Inc
- BlogTalk Radio
- CBS Local
- Cox Media
- DGital Media
- ESPN.com
- Google
- Libsyn
- Midroll Media
- NPR

- New York Public Radio
- Nielsen
- PodcastOne
- Podtrac
- RawVoice/Blubrry
- Sizmek
- Panoply/The Slate Group
- Triton Digital
- Westwood One
- WideOrbit
- Wondery

The IAB lead on this initiative was Amit Shetty and documented by Katie Stroud.

https://iabtechlab.com/working-groups/podcast-technical-working-group/

Contact adtechnology@iab.com to comment on this document.

ABOUT THE IAB'S AUDIO COMMITTEE

The Audio Committee's Mission is to establish industry guidance by creating standards and best practices to help bring clarity to the Audio marketplace. The Committee will strive to educate marketers and agencies on the value of Audio as a powerful and effective advertising medium. The IAB Audio Committee is led by Rena Unger

Audio Committee Webpage

This document (and future updates) can be found on the IAB Tech Lab website at:

https://iabtechlab.com/specifications-guidelines/podcast-metrics/

Table of Contents

1. Executive Summary

Podcast audiences represent a growing segment of effective marketable media but inherit challenges of legacy consumption platforms. This restricts the ability for advertisers to obtain the same data parity found in other digital mediums, limiting participation of some advertisers. This document provides an introduction to tracking ad delivery in a podcast and attempts to provide clarity in the marketplace by describing common and existing practices.

Podcasts are downloaded to a device for later listening or for online listening. In most cases the podcast file and any ads included with it are downloaded to a device that doesn't, or can't, send data about the consumption of the podcast and ads. This lack of data beyond ad delivery limits real-time measurement. In contrast, other media are consumed by reading an article and interacting with a site, playing a game, or streaming a video, all of which can be measured in real time. Even audio stations that offer music or news are streamed and measured in real time in today's media marketplace.

Unlike streaming audio, podcast listeners come from an era where files were downloaded. The medium, the distribution, and the platforms used to collect and listen are built around the habit of downloading the file. Tracking content in this time-shifted medium involves filtering server logs to produce meaningful data for measurement. Since podcast technical teams analyze server logs differently, results vary across the industry.

The challenge for podcast producers and distributors is to offer buyers a set of metrics that is consistently defined and measured equally across the podcast medium. While ad measurement will always produce some discrepancy in any medium, the definitions in this document aim to reduce measurement discrepancies. This document is the first attempt to bring IAB members together, describe common practices and provide a common set of metric definitions. With a consistent set of podcast advertising metrics, buyers and sellers can engage in a conversation about campaign strategy with confidence.

Audience

While all professionals in the podcast supply chain can benefit by being familiar with this document, metric definitions are primarily intended for podcast producers and distributors. Specifically, account managers should be familiar with and use metrics as defined in this document when negotiating ad packages with buyers. Additionally, podcast ad operations teams should use the metric definitions in this document to design or adjust the ad measurement technology they use to analyze server logs for podcast ad measurement.

Buyers should also reference this document to better understand how ads in podcast content are counted. This document offers a set of metrics that establish a mutual understanding in podcast advertising negotiations.

2. Overview

Podcast content is an on-demand media format that listeners either download to listen to later or consume online . Unlike the streaming format more common in video, podcasts continue to be downloaded because of the convenience offered by existing platform and application functionality.

Despite the use of the word "streaming" in podcasting, "streamed" podcast files are progressively downloaded via the standard HTTP protocol. True streaming—typically reserved for live events—requires a specialized server and uses an entirely different protocol.

While "streaming" a podcast and true streaming formats appear exactly the same to end users, delivery of a streamed podcast is logged the same way as a downloaded file in the server logs. This important distinction impacts the ability to measure content and ad delivery in real-time without access to client side analytics. Podcast publishers must work around this limitation and track metrics using server log data.

Media delivery via true streaming falls outside of the definition of a "podcast" and is therefore excluded from this document.

Podcast Player Market Share and Tracking Limitations

The ability to track podcast content and ad playback largely depends on the player requesting the file. The native players that operate on iOS systems, namely the Apple Podcasts App and iTunes offer no technology for confirming that a podcast file was played. This lack of client-side response prevents podcast distributors from measuring ad plays at the level expected in other digital media.

In order to provide insight into the limits on tracking podcast content, the IAB Podcast Measurement Working Group was asked to provide reports on the market share for platforms that request podcast files. Podtrac, Blubrry/RawVoice, WideOrbit, Libsyn, and PodcastOne submitted reports used to produce the following table, which aggregates the resulting market share percentages for the month of April 2016.

**Aggregate Report on Podcast Player
Market Share (April 2016)**

Platform requesting podcast file	Range of market share %
iOS - Apple Podcast App	45-52%
iTunes	8-13%
Browsers	6-14%
Stitcher	2-7%
Everything else	12-30%

Reports for non-iOS and iTunes market share varied from one report to another, but the results reported for the iOS Apple Podcast app were consistent across all reports, with a mean of about 49%.

For about half of the podcast ads served to browsers (6-14%), some servers may be able to distinguish ad delivery from a probable ad "play." Using browser plugins or other technology, a specialized tag used to request the ad file can indicate that the player accessed the ad. While this technique offers valuable tracking data, half of the 6-14% means that only about 3-7% of podcast ads can be tracked this way.

Another key insight for this report (not represented in the table) is that less than 3% of the market share enables client-side tracking as it exists in other forms of digital advertising. Only one of the five participants reported a small percentage (3%) of market share for host-branded players (players owned by the podcast producers). Since the producer controls the player and the content, they can request ads and trigger tracking beacons based on ad play. Most podcast distributors see almost zero activity in this market. Even of the 3% of reported host-branded players, many may not be equipped to find tracking beacons and use them.

The data provided for this report shows that half of the market share for podcast players belongs to the Apple Podcasts app, which prevents any client-side tracking or even the ability to count a "play." Podcast distributors must turn to server log analysis and report on ad delivery. Some distributors count an ad once it's been served. This count offers a valuable metric, but is out of scope for this document because an ad served doesn't indicate whether the ad file was downloaded.

Despite the limitations, podcasts audiences are growing and offer valuable exposure for marketers. In order to offer this value to buyers, metrics must be consistently defined across the industry. IAB collaborated with members in the podcasting community to establish metric definitions that can be used consistently in the podcast marketplace.

Establishing consensus and clarity for podcast reporting metrics improves communication and establishes trust and accountability with buyers.

Scope

This document defines content, ad, and audience metrics in the context of downloaded podcasts whether saved for later listening or listened to while being downloaded. In this context, both formats are typically pre-recorded and available on demand whenever the listener is ready to access the files.

Podcasts that use true streaming technology to deliver the ad offer the ability to track activity in real-time or near real-time and one metric used to measure "client-confirmed ad delivery" is covered in this document. However, the percentage of market share for applications that support true streaming is currently too small to account for any meaningful campaign measurement. Additional measurement guidance for true streaming audio is covered in the MRC Audio Measurement Guidelines currently in development as of the release of this document. Those guidelines do not cover Podcast measurement – mainly due to the limitations around client side measurement in the industry.

Ad measurement in podcasting presents the industry with many challenges. For the sake of establishing common ground in tracking podcast ads, the definitions presented in this document address counting ad delivery. This count comes from analyzing server log files to determine what was actually delivered.

3. The Podcast Medium – Content Delivery

Podcast listeners acquire podcast files in one of two ways: either by downloading the file for later listening (downloaded), or by listening while the file is downloaded (online listening). To a lesser degree, some podcasts may also be played while a persistent connection to the server is maintained (streamed), but the market share for applications that support this format is insignificant for campaign measurement and excluded from discussion here.

Delivery methods for downloaded files, whether listened to later or during download, offer valuable inventory to advertisers, but content and ad delivery are handled differently in both environments. An overview of each format is explained below. Despite different tracking capabilities in each environment, a few baseline metrics should be able to offer similar reports for both podcast types.

Downloaded Podcasts

Podcast downloading allows the audience to download full episodes of content that can be played at a later date and time. Listeners may subscribe to select programs, and platforms like iTunes continue to support full downloads to a personal library for listening offline at anytime in the future. The convenience of this system makes downloaded podcasts a continued preference among listeners.

Online Podcasts

Online podcasts appear to be streamed, but the file is actually being downloaded while the listener is listening to the file. The downloaded file is stored in a temporary location rather than to a library as with a downloaded podcast. Since online files are typically downloaded the same way as the files stored for later listening, delivery for these two formats are recorded the same way in the server logs. The only difference between the two are whether the listener is actively playing the file as it's downloaded or being saved for later listening - which can only be discerned by the player.

Raw Server Logs

In a downloaded file, segments of the file are collected on the listener's device, or progressively downloaded. These progressively downloaded files result in a server log with several requests to the server, which must then be analyzed and filtered from other server requests in order to represent how many files were downloaded and to what audiences. When podcast ad operations do this using a consistent process, metrics can be reported and trusted with a higher level of confidence.

4. The Podcast Medium – Ad Delivery

Podcast ads can be delivered and tracked in a variety of different ways, but in general two different methods are used with variations on each.

Integrated Ads

Historically podcasting ad campaigns often involve ads that are read by the podcast host or a familiar voice. A static ad or jingle may be also included as part of the file. These ads are part of the content and included, or "baked-in," with the file that is downloaded. Targeting is limited because everyone who downloads the file gets the same ads. When the ad is added to the podcast at the time of request, it is dynamically inserted, as described below.

Dynamically Inserted Ads

In recent years, ad technology has allowed for ads to be targeted and dynamically inserted at the time of file request. The ad server determines the best ad to serve to the listener at the time of request. In a podcast consumed online, ads may be inserted into a file that is being progressively downloaded at designated ad breaks. Some publishers may count this dynamic ad serve as an "impression" without confirming ad delivery. The metrics in this document focus on confirming that the ad was delivered. Server logs can confirm that the entire ad file was downloaded, but the process for counting a served ad can only determine that ad file was sent.

5. Podcast Ad Tracking

In digital display advertising, ad tracking is done using beacons that are triggered in the web browser, or client, which verifies that the ad was viewed or at least had an opportunity to be viewed. In podcasting, client-side tracking is usually only possible when the client player passes ad data back to podcast producer or distributor. In this set-up, the player is the client and is programmed to notify the server when an ad has been played. While this set-up offers the most accurate ad delivery counts, it represents a very small percentage of the podcast industry–less than 3% according to member reports on market share in the industry (see table 1).

Filtering Server Logs

In order to produce accurate counts for podcast ads, technical staff in podcast ad operations must analyze server logs. These server logs may include file requests for a combination of downloaded podcast files, dynamically inserted ads, and any content requested by the web page or application hosting the player. A number of factors are used to analyze log files.

HTTP GET requests, and under certain circumstances HTTP HEAD requests, may be processed that contain the following data.

- **IP Address** - The IP address is one of the factors that may be used to determine if the request is unique or a duplicate. (Exceptions are shared locations such as corporate offices, dorms etc., that have a large number of people sharing the external IP Address) It may also be used to determine geographical information of the media consumer.
- **Time Stamp** - The date and time may be used to determine if the request should be counted.
- **HTTP Status Code** - The appropriate HTTP status code is examined to determine if the request should be counted.
- **Bytes Served** - The value may be used to determine if the media was completely downloaded or if not, how much was downloaded. (Note: This information is only available from native server log files.)
- **Referrer** - The origin of the download may be used to determine if the request should be counted. e.g. media that is auto played upon loading a web page may be removed or reported.
- **User Agent** - The identifier of the application or service consuming the media may be analyzed to determine if the request is unique.
- **Byte Range** - The range of bytes requested in a given request may be used to determine what portion of the media is requested.

When analyzed across multiple requests, the information may offer data that represent ad delivery. Since media technology is always changing, no specific combination of factors or techniques will offer the most accurate count indefinitely. However, meeting some minimum requirements will help produce more consistent results. Suggestions for server log analysis are offered in the appendix.

Podcast producers and distributors may include additional metrics beyond the ones defined here, but such additional metrics should be labeled separately from the core list of metrics described in this document.

Content Metric Definitions

Since podcast ads are so closely integrated with podcast content, metrics that measure content are vital to ad measurement in podcasting. The following metrics are used to describe content downloads. Server log analysis for content delivery should filter data to produce metrics as defined below.

Unique file request: the number of file requests for a single file originating from a single user or client within a specified time frame. Techniques for deriving the counts for this metric should be disclosed as part of an agreement between parties.

Complete file download: a unique file request that was completely downloaded. A complete file download confirms that all bytes of the file were sent. A last byte sent marker is not a sufficient way to determine a complete file download.

Partial file download: a unique file request that was less that 100% downloaded. This metric may be used to help determine ad delivery for ads that were included in the downloaded portion of the ad for files that were not completely downloaded. Partial downloads also help determine audience engagement.

While these metrics define content downloads and exclude ads, the counts are needed to determine whether ads were delivered, as described below.

Ad Metric Definitions

The following metrics represent the first step toward improved ad measurement in podcast advertising. These metrics are derived using the content metrics defined above. As these metrics become adopted in the industry, additional steps can be made toward an improved podcasting ecosystem.

ad delivered: an ad that was delivered as determined by server logs that show either all bytes of the ad file were sent or the bytes representing the portion of the podcast file containing the ad file was downloaded.

For example, if an ad was included within the first 25% of a podcast and at least 25% of the podcast file was downloaded, then the ad can be counted as delivered. When ads are dynamically inserted into the podcast file or within an ad break within the podcast, 100% of the ad content (all bytes) must be downloaded before it may be counted as delivered. These measurement guidelines do not support counting served ads without confirming that the ad was delivered.

client-confirmed ad play: counts an ad that was able to prompt a tracking beacon from the client when the file was played. Whenever possible, metric should include information about how much of the ad was played using the markers: ad start, first quartile (25%), midpoint (50%), third quartile (75%), and complete (100%).

While the client-confirmed ad play metric represents the most accurate count for ad plays in a podcast, it requires client-side tracking. As discussed earlier, the platforms used to download, store, and play podcast files lack or prevent the technology needed for client-side counting. Aggregate reports on player market share among podcast publishers estimate that less than 3% of players are capable of providing client-side tracking data (see table 1).

Podcast Audience Metric

Podcast listeners usually download more than one episode and usually from more than one podcast. A measure of how many people downloaded episodes can be useful by describing the reach of the podcast or group of podcasts.

Website metrics of reach usually depend on the use of cookies or JavaScript to identify individual devices or users. Downloading an MP3 file does not usually trigger the reading of cookies or the running of JavaScript, so alternative methods must be used. HTTP requests include the IP address of the recipient of the file and usually include a user agent that, while not unique to the user, provides some ability to distinguish multiple users behind one IP address.

> **listener:** data that represents a single user who downloads content (for immediate or delayed consumption) from the podcast publisher. Listeners may be represented by a combination of IP address and user agent ID and must be specified within a stated time frame (day, week, month, etc.). This metric should also be filtered by podcast group. For example, one listener may download a collection of podcasts. The listener should only be counted once across a collection of podcasts.

6. Summary

Podcast audiences represent a growing segment of marketable media, and are considered to be some of the most engaged. However, the medium is asynchronous and in most cases severed from data collection once delivered, which presents advertisers with measurement challenges. In addition, measurement practices have been fragmented and ad-hoc so far. This document offers the first step in an improved environment where buyers and sellers can start to use the same language with clearly defined meaning. As communication improves, producers will be able to scale their operations and invest in technology that brings tracking closer to the standards available in other media options.

7. Appendix - Examples

Metric definitions provide a goal for consistent vocabulary in podcast advertising. The technical details for achieving those metrics are dependent on each company's abilities, techniques, and business models. Until some baseline technical standards can be established, the following appendices offer examples for how metrics may be established as defined in this document. Your organization may or may not be able to duplicate these sample techniques for filtering and analyzing server logs for content and ad delivery.

1. Example formula for measuring unique media file requests

One or multiple server requests for a single file within a 24-hour period and identified as unique by a combination of IP address, browser/application user agent, and the file being requested. A maximum of 2 unique file requests within a 24-hr period may be counted. Unique file requests filters log files for user access to files whether completely downloaded or partially downloaded.

The 24 hour time period minimum prevents the possibility of fraud by limiting the number of duplicate requests from a unique *IP Address*. Vendors who use a smaller time period should deploy additional filtering techniques to determine that the additional media file consumption is legitimate.

The 2 unique *IP Address* and *User Agent* combination maximum prevents the possibility of fraud by *User Agent* spoofing. Vendors who exceed this limit should deploy additional filtering techniques to determine that the additional media file consumption is legitimate.

2. Example formula for determining a partial download

One or more requests with total bytes served less than file size identified by the initial request meeting the requirements of a unique file download. Bytes served total may be calculated by using byte range request and bytes served data. Range request data may be used to determine which portions of the file were downloaded. Bytes served and byte range data may be used to determine the total bytes delivered.

3. Example formula for determining a completed download

One server request meeting the requirements of a *unique file download* with all bytes served. A *Partial download* may be counted as a completed download if the bytes served total equals or exceeds file size.

Partial file downloads are common occurrence in podcast media consumption. Below is a brief list of the common reasons why they occur.

- Only portions of file progressively played from within a web browser or application.
- Network connection or application issues causing downloads to resume or restart at a later time and/or on a different network.
- Download acceleration technique of splitting downloads into chunks.
- A download in progress is canceled by the user.