Enhancements in Audio Ads Delivery and Effectiveness

Rj Harry

Independent Researcher
New York, NY, USA
June 10, 2019

harry.t.raymond@gmail.com

Abstract

This work details progress in audio ads - delivery and effectiveness. With the increased use of smart speakers and voice-based commands, a new ad avenue in the form of audio ads is rapidly emerging. Audio ads however are different in several aspects from traditional banner/image/video/text (email) ads. For one they're richer in terms of contextual information as the user interaction with voice is more personal and can be used to determine various aspects from user mood to activity being preformed.

1 Developments in Measuring Effectiveness of Audio Ads

Most smart home speakers or voice-enabled computing devices ranging from smartphones to smart wearables have wakewords configured. The utterance of the wakeword by the user typically signals the device to start processing user conversations. Most devices are always active listening into user conversations although processing of commands/user requests begins only once the wakeword is said by the user. Audio ads are served to the users of these devices as they converse with the smart device. The served audio ads are typically relevant to the user's interests, contextual usage (ie: based on the requests/suggestions the user is making of the device at that time), demoographics (like location, etc) and other related data. Once these audio ads are served on the user, the impression is recorded and reported to the sponsor of the ad. In this section we present further tracking avenues to measure the effectiveness of such ads.

1.1 Tracking User Sentiment on Advertised Product

This requires adding new words to the already existing wakeword/s. These words may be ad related words, either chosen by the smart device or set by the sponsor of the advertisement. If the words are ad-related words chosen by the smart device, this may be accomplished by means of machine learning algorithms that detect high impact/subject in a sequence of words (ie: those part of the ad, played by the smart device in the audio). Several works in literature are available that demonstrate the extraction of topics of the ad or identifying hot terms.

Before proceeding further, we will define τ, which may be defined as the window of monitoring/measuring ad effectiveness, starting from the time the audio ad was served. τ may be decided by the sponsor or chosen by the smart device. τ must be chosen such that it gives an accurate representation of the effect of the ad, and as such, τ serves as the monitoring window period of the effect of the ad.

Once the ad-related words is added to the list of wakewords of the smart device, monitoring begins. The smart devices may also be configured such that the ad-related wakewords are treated differently when compared with the usual wakewords. The smart device may then:

- count the number of times the ad-related wakeword is uttered in the set time period τ.
- use machine learning algorithms to detect the sentiment (positive, negative, etc) around the advertised product within the timeframe,τ, when the ad-related wakewords were spoken. For this purpose of detecting sentiment within the monitoring window of τ, additional inputs such as user voice tone, tempo of speech, accents, tones, etc could also be used as inputs.

- track household activity being performed at the time of utterance of the ad-related wakewords.
- detect and record the number of recipients involved in the conversations when the wakeword was uttered.
- detect and record the number of recipients possibly present in the room/within earshot when wakeword was uttered.

The above may be presented then to the sponsors to underline the effectiveness of their ads. An example table of such a presntation may look like is provided in Table 1. In Table 1, the sentiments row mentions P, N and D next to the numbers. These are Positive, Negative and Doubtful. The classification of contexts ain which the ad-related utterances were made is done by means of machine learning algorithms. A plethora of such sentiment analysis algorithms employing machine learning are readily available in literature. Similarly, the 'Activity' row presents H or C/GD. These are abbreviations for Heavy household activity (something generaing noise) and Conversation/Group Discussion respectively. Activity may be detected by means of motion sensors, vibration detectors on the smart device, apart from the sound input that the device receives. Again there are plenty of machine learning algorithms that are able to classify activity based on sound, motion and vibration inputs.

2 Conclusion

In this study we presented metrics that may be a reliable moniker to the effectiveness of audio ads served on recipients. Note that the additional insights that may be gathered from the metrics are mainly due to the highly personal nature of voice interactions, versus say key input, mouse clicks or other means of machine input. Smart devices that accept voice/sound inputs thus allow us to gain deeper insights on the effectiveness of the advertising on the recipient.

Similar metrics can be extended to other format apart from audio ads. For instance the same smart device may also have means to deliver image/banner/text/video ads by means of some output means connected to the smart device. Wakewords for each of these ad formats may be obtained either directly from the sponsor or by running machine learning content analysis algorithms to track hot terms and topics in each of these content formats. Smart devices can then use their smart listening functionality to arrive at the metrics above.

Going further, effectiveness of ads in different formats on different devices that have been delivered to the recipient can also be monitored provided there are means to uniquely identify the recipient delivered the ad in one device, but also using the smart evice at the same time. This could be in form of a single login, common email, username or some related common user artifact

Table 1. Audio Ads Effectiveness

$\tau->$	**2 Hours**	**2-6 Hour**	**6-14 Hour**
Number of Utterances of Ad-Related Wakewords	6	3	0
Sentiment around Utterances of Ad-Related Wakewords	45% P, 25% N, 10% D	25% P	NA
Number of Recipients Involved	1	3	NA
Number of Recipients Present	1	5	NA
Activity at the Time	H	C/GD	NA