



USING METADATA TO MAXIMIZE YIELD AND EXPAND INVENTORY IN TV - CONTEXTUAL ADVERTISING

S. Kancherla, R. Warey, Ramki S.

Prime Focus Technologies, India

ABSTRACT

Today's consumers not only have a variety of content choices, they also enjoy the convenience of consuming this content when they want, the way they want. While this has made video viewing increasingly personalized, the ads that play in and around the video are still far from being personalized and relevant. Broadcast networks need to guard against viewer fatigue/loss while brands need to be increasingly conscious of the ROI on every ad dollar spent.

This paper highlights innovation done in contextual advertising using an additional metadata layer of in-video context. An approach that offers broadcasters the opportunity to optimally sell their video inventory and for brands to place their ad in the right program at the right time/place.

The result – not only the opportunities to show an ad increase, but also the yield per ad spot improves considerably. A win-win for both broadcasters and brands.

INTRODUCTION

The way we consume content is changing. Our context drives platform – device choice. Consumption of video over the internet, be it on TV (IPTV) or any other device is the new normal. No wonder terms like “Cord-Cutters”, “Cord-Nevers” etc. are gaining traction and will become part of the mainstream going forward.

In the space of the content (video) consumption on an OTT platform (using IPTV, mobile, tablet), showing contextual ads is a currently a function of global attributes only. From the content perspective, typically parameters like the content type, title, genre, language etc. are considered. While from a user profiling standpoint, parameters like age, gender, and location dominate.

The traditional method of marrying these two sets of parameters to arrive at which ad to show (either predefined or real time) is limiting for both the publisher/ broadcaster and the advertiser. For the publisher, the volume of inventory is limited and yield per unit of inventory is sub optimum. For the advertiser, he/she might be missing out on any many opportunities (relevant points during the video) where serving the ad would make sense. Additionally with limited “context”, ROI on ad spends is not effective – don't know if the right ad is being shown to the right TG at the right time!

The need of the hour is for a solution that increases the ad inventory and relevance without impacting the viewing experience.

CONTEXTUAL ADVERTISING AS IS

Currently, the placement of an ad within content is based on the following process:

- **Step 1** – Brands decide the content they want to place ad(s) in, based on global metadata like title, genre, language etc.
- **Step 2** – The target audience is defined. Mainly on basis of demographics like age, gender, location and so on
- **Step 3** – Marriage of content and user metadata is done to determine a pool of ads that are most relevant for this combination
- **Step 4** – From the pool of ads, one is chosen and shown to the viewer

Steps 3 and 4 above can be either be predefined (in case of offline ad deals) or can happen in real time in the case of programmatic ads.

Key Limitations

The key limitations encountered with the above approach for each of the stakeholders are as below:

PUBLISHER	ADVERTISER	VIEWER
<ul style="list-style-type: none"> • Inventory volume based on global attributes of content/user profile • Limited set of advertisers that can be approached • Large portions of the library remain unsold / undervalued 	<ul style="list-style-type: none"> • Competing for same pool of content that everyone else is after • Weak linkage between ad objectives, content and user profile – low ROI • Not able to see beyond views, impressions and other such global metrics 	<ul style="list-style-type: none"> • Passive viewing of ads • Limited relevance to the viewing context • Turn off

Figure 1 – Limitations of the current approach is contextual advertising

THE NEW APPROACH

Using in video context like the mood of the scene, characters, their emotion & action, each moment within the video is described using certain keywords or metadata. This can be for an exact point or for a given time interval within the video. The additional layer of in-video metadata is extracted using automated tools OR manual tagging OR a combination of the two. The pre-decided factor in these approaches is the data model i.e. rules as to what in-

video metadata needs to be captured. In the combined approach i.e. automated + manual, automated tools analyze frames within the video to extract information. This will not be 100% accurate. Thus a manual effort is added on top of this. The result is filtered, and relevant metadata is curated. These keywords/metadata coupled with already available content (e.g. title, genre etc.) and user (e.g. age, location etc.) metadata becomes highly enriched information. This information is then passed on to the ad decision system. Owing to the additional layer of in-video metadata that has come in, the ad system is able to make a better decision about which ad to show – thus in turn increasing its relevance for that in-video ‘moment’- Better ROI for advertisers. Also, since each frame/time interval is described using certain metadata, the number of such ‘moments’ where an ad can be shown also goes up – potentially more opportunities for a publisher/broadcaster to show an ad.

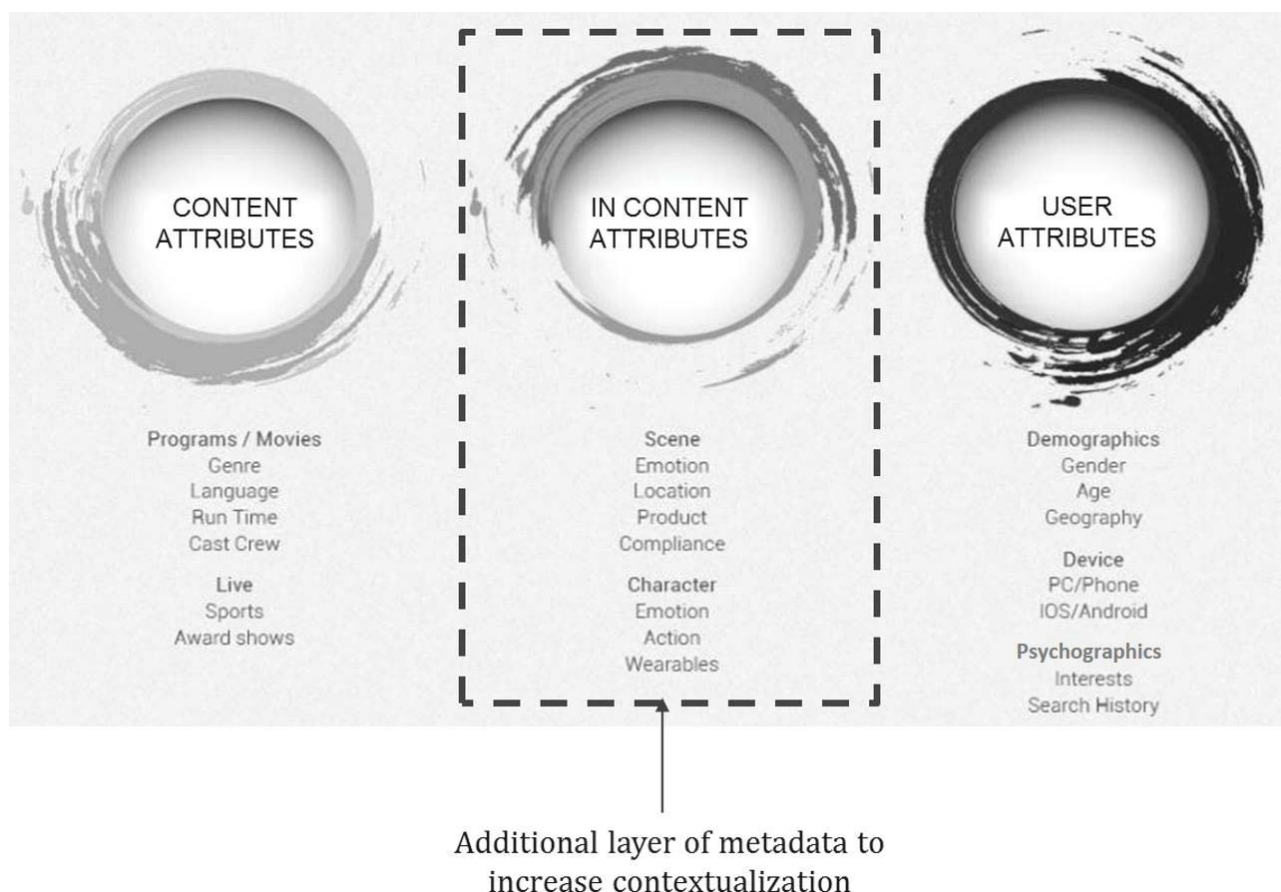


Figure 2 – Set of parameters used to fetch a contextual ad

THE SOLUTION

In the typical IP based content and ad ecosystem, the key components involved and their roles are as below:

- Media asset management (MAM) – Repository for content and metadata

- Content management system (CMS) – Backend to manage a content platform
- Ad systems – Repository for ad assets and defining ad targeting rules
- End user device – View content, pass on user information to content, ad systems

The interplay between these components is as per the figure below. A brief explanation of the same is also provided.

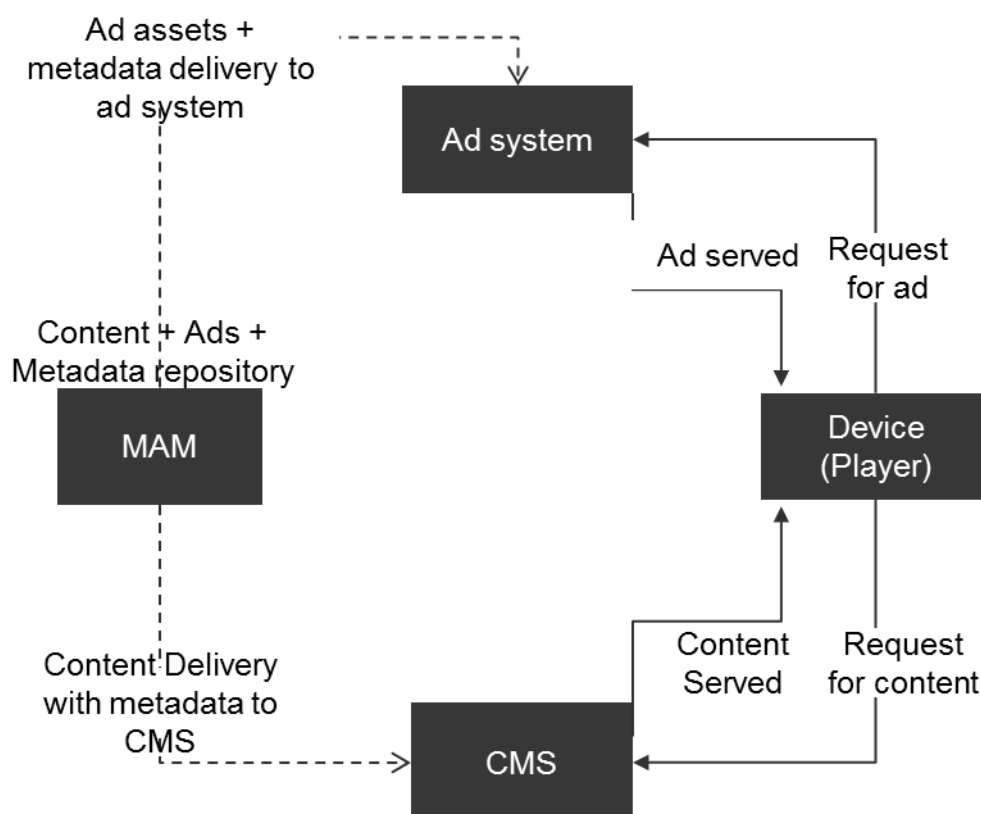


Figure 3 – Existing content & ad serving ecosystem

- Content assets, ad assets and metadata associated with them is stored in the media asset management system (MAM)
- Content assets are prepared and published to the content management system (CMS). Here preparation refers to the video being compliant to technical and metadata specs of the platform it is being published to
- Ad assets are prepared and published to ad server
- When the end user clicks play on his device, the content (video) is served by the CMS. Similarly, on the trigger of an ad, the ad (video) is served by the ad server. The trigger for the ad can be predefined

For this new approach of using in-video metadata, two additional steps were introduced within the above ecosystem.

Rich Metadata Extraction

This referred to capturing in-video metadata. As a first step, the data model i.e. what in-video information needs to be captured was defined. To extract this metadata a combination of automated tools and manual effort was used. The automated tools analysed key frames within the video. Variables like facial expression, objects, background/foreground, colour depth and more were taken into account while doing this. The extracted information was then compared against an existing database of information (images, text etc.) for a match. The end result is descriptors/keywords for a given in-video moment. The automated tools are not 100% accurate. Also, they would throw a lot of metadata which may not be needed in the current context. A layer of manual metadata curation and tagging was thus added on top to get the final, cleaned up and relevant metadata. This manual effort was either borne by the content owner or a third party metadata service provider. While each video is a different case, on an average the split of automated to manual effort was in the range of 60:40.

The sources for the metadata could be anything that is time code dependent. E.g. subtitles, audio tracks, scene level details like – characters, objects etc. The typical in-video metadata captured would be the presence of characters – their emotion and actions, as well as the overall scene mood, location and products.

Interpretation Engine

The extracted metadata would then be passed onto an interpretation engine. The interpretation engine would process the rich metadata to:

- Generate markers to show potential opportunities within the video where an ad could be shown. This information would be passed on to the CMS
- Map the 'moment' to a tag within the ad system for custom targeting. E.g. If the metadata extractor has described a scene as a "car accident", the interpretation engine would map that information to an ad tag like "insurance" within the ad system

Revised Ecosystem

With the addition of the rich metadata extraction and interpretation engine, the revised ecosystem looked something like the illustration below.

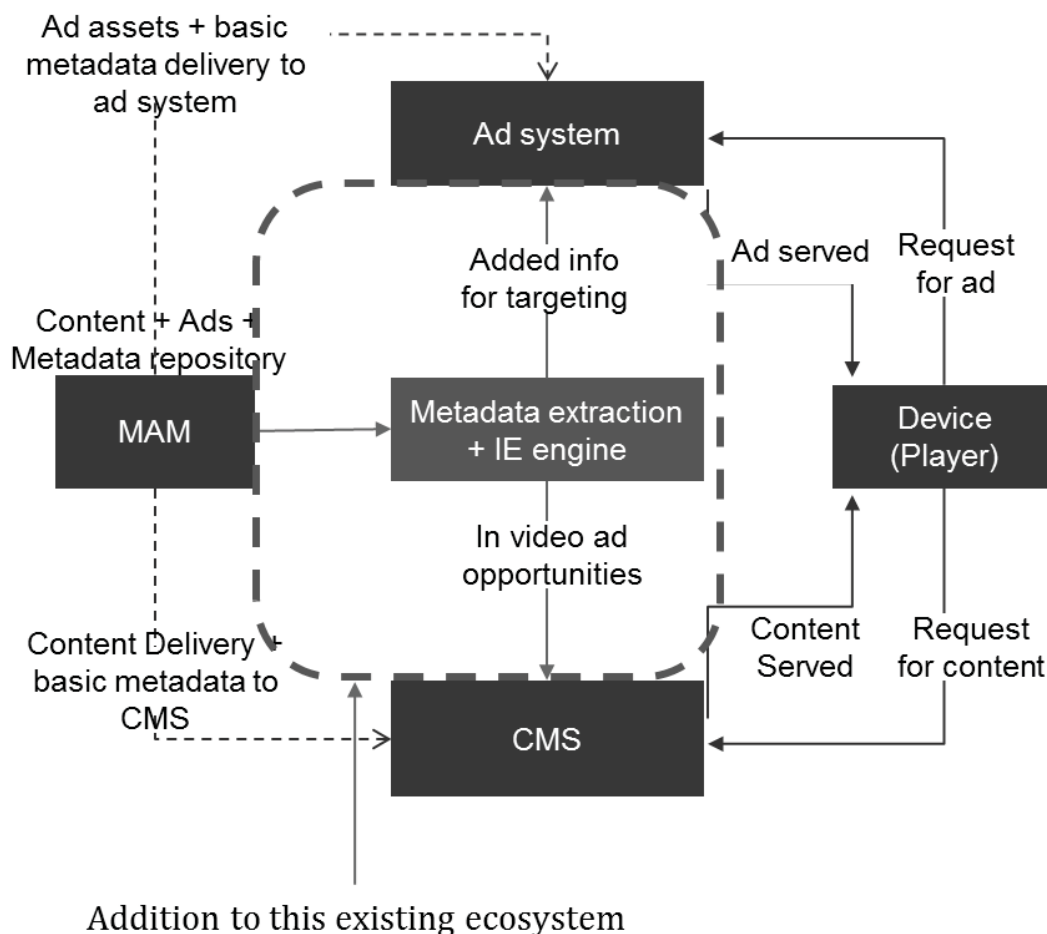


Figure 4 – Revised ecosystem

As the user started watching a video, at a given moment (time code), the system would already know how that moment is described (using metadata). It would pass this along with the user profile to the ad system. The ad system in turn had certain targeting rules defined. The targeting rules were defined based on a combination of content metadata, user profile metadata and in-content or in-video scene specific metadata. Once a rule was met, a contextually relevant ad would play based on predefined offline ad deals (done directly between publisher and advertiser) and/or programmatic ad buys (done in real time). Not just ads, these 'moments' could be filled with contests, trivia, polls etc.

An Example of the targeting rule for a programmatic ad buy is as below

```

IF {content type = 'Movie' & content genre = 'Drama' & content language = 'English'}
  AND
  {user gender = 'Male' , user age range = '25-40' , user location = 'India'}
  AND
  {product in scene = 'Cigarette', character action in scene = 'Smoking'}
THEN
  Play an 'Anti - smoking campaign' banner ad for the duration of the scene
  
```

MEASURING EFFECTIVENESS

To test the effectiveness of this solution, a survey was done internally (conducted by Prime Focus Technologies). The research design was as below:

- Two groups of people were formed
 - Control group - Shown content with no method/logic to the ad placement. Ads were placed randomly
 - Experimental group - Shown content with an ad contextual to the content and user parameters. Ads were placed at the apt moment
- A representative sample was chosen – Participants were internal (Prime Focus group employees) and external (no association with Prime Focus), who were located across geographies and were users of various devices (Tablets, mobiles, IPTV)
- Users either individually or within a group were exposed to the video and the ad

The following parameters were measured

- General recall – Did you see any ad at all?
- Theme recall – What was the theme of the ad that you saw?
- Brand recall – Which brand did you see?
- Message recall – What was the specific message of the ad?
- Purchase intent - Given the need, will you buy/ use this product / service in the near future?

The results of the survey were as below:

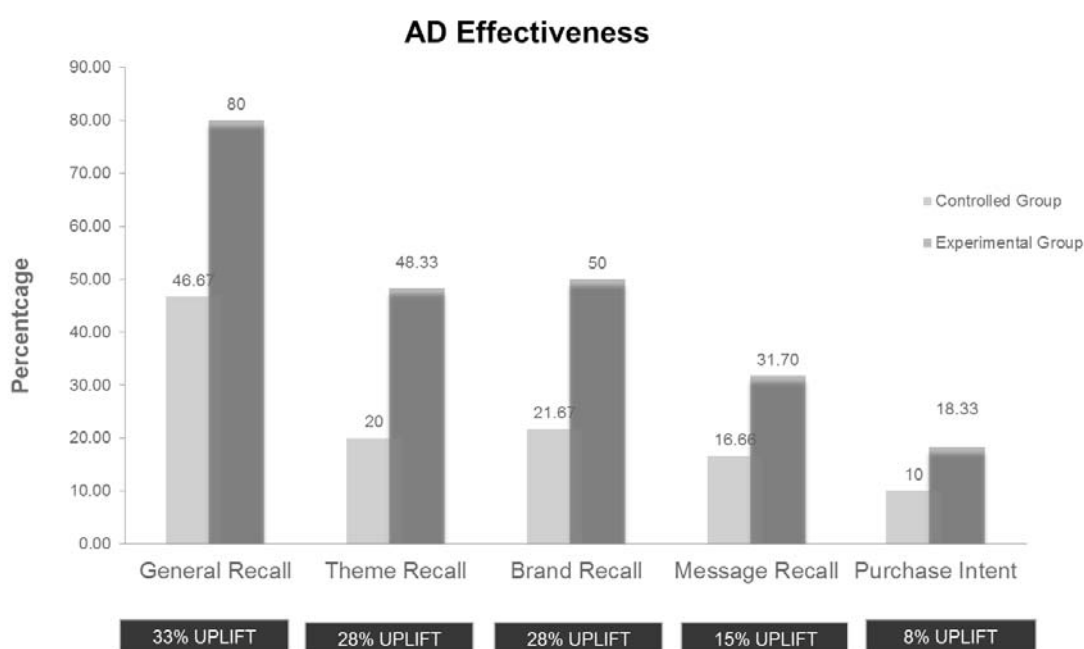


Figure 5 – Results of the sample survey

On all parameters measured, the ads when shown using the additional metadata layer of in-video context performed much better.

BENEFITS

The key benefits of this innovation are:

- For the Publisher / Broadcaster
 - Ability to slice and dice the content based on in-video context – opportunities to show an ad are increased
 - Create packages using this in-video context that can be sold in a targeted manner to multiple advertisers
 - Command higher ad rates since the inventory being made available is richer in context

- For the Advertiser
 - Relevance increases since the ad is shown at the apt 'moment'
 - Improved Brand Affinity due to relationship with the context of the content
 - Better ROI

CONCLUSION

In an era of tremendous revenue and margin pressures, brands are increasingly questioning the effectiveness of every dollar spent in advertising. Broadcasters (publishers), on the other hand, are facing an uphill task trying to sell all of their inventory and justifying it's true worth to the advertisers.

The way forward is to adapt innovative methods in advertising that is beneficial to both parties. Using in-video context, not only does the ad inventory increase, but the yield and ROI per ad spot goes up too. More importantly though, the end consumer gets a seamless experience of content and ads bundled together that is non-intrusive and does not disrupt the overall viewing experience.