Link to Article:
https://www.informs.org/Impact/O.R.-Analytics-Success-Stories/More-Effective-Television-Advertising

Part I: The Setting

With the advent of digital media solutions, traditional TV advertising has struggled to compete with the precise targeting and performance reporting advantages inherent in digital advertising. With digital media, every impression is measurable and capable of advanced analysis. In contrast, television advertising lacks sophisticated measureability and precision. An advertiser buys television advertising knowing that they have little control of who will actually be watching the programs their ads air on, and little ability to understand if the person who viewed the ad ended up taking a favorable action toward the advertiser. The headwinds were significant for television advertisers as the digital age created new winners and losers.

Yet, human history has consistently borne proof that those with much to lose find renewed passion to reverse the unfavorable momentum they encounter. The analytic success story I have chosen to explore contains an example of this well-established theme: Turner Media's Innovative Solutions for Television Advertising. Understanding the advantages of digital advertising, Turner Media sought to bring some of the precision and measureability inherent to digital media into the rooted practices of television advertising. To do so, they utilized the disciplines of analytics and data science to create solutions that sought to preserve television's primacy in the media and advertising establishment.

Part II: The Three Models

. Turner Media sought to solve one fundamental problem: digital advertising had greater precision targeting and measureability than television advertising. In digital marketing, an advertiser can choose an exact type of audience and feel confident they will only show their advertisement to that audience with minimal waste. For example, if Georgia Tech's Data Science Masters Program only wants to serve advertising to individuals who express interest in pursuing data science graduate studies, digital media provides a multitude of options to limit your advertising to only the relevant individuals. For instance, you can choose to only serve ads to users who visit the Georgia Tech Data Science website. Or, you can choose to serve ads to individuals 21-40, with college degrees, exhibiting an interest in math and data science. Digital tools have successfully solved the problem of pairing the right advertisement with the correct audience.

In contrast, television advertising had few techniques to ensure Georgia Tech's advertising reached the correct users and only the correct users. Television networks offered ad placements during the commercials of their shows. This technique proved inadequate; pairing a television ad with a commercial spot lacks precision. It both under-delivers against the target audience (those interested in data science have a long tail of television program preferences) and over-delivers to irrelevant audiences (the programs chosen to show an ad on have a much broader audience viewership than just individuals interested in data science). In fact, advertisers buying television spots had no pretense about this. They bought their television ads accepting that they would show their ad to a large portion of irrelevant individuals. Pre Digital Media, this all made sense since advertisers had no other option; post digital media, television advertising became a lot less attractive to many sorts of advertisers.

Turner Media utilized data science models to create fresh and intriguing television advertising solutions that bridged some of the gap created by digital advertising techniques. The first model (1) used external third party data combined with their own first party viewing data to create audience segments and indexes for which audience segments a particular tv spot was likely to contain. The second model (2) utilized an optimization constraint of the advertiser's basic buying parameters, but tried to maximize the delivery of the advertisers television ads against key audiences within that constraint. The third model (3) created a new buying strategy that optimized an advertiser's television ad to only serve during placements that indexed highly to the audience segments the advertiser wished to target.

Model 1: Clustering Model Including Data Inputs Required

Model 1 sought to address the greater precision inherent to digital advertising. Television advertising was typically bought by guaranteeing basic demographic viewership. For example, an advertiser of female hygiene products might ask Turner Media to guarantee delivery against females aged 25-60. Turner Media would then find the programs that had a high likelihood of containing those demographics and serve advertising to those spots. Digital Media, however, offered much more sophisticated and precise targeting. With more signals created in a digital ecosystem, digital advertising had far more dimensions and attributes to create interesting and compelling audiences for advertisers to reach.

Turner Media had to bridge this gap, and they turned to third party data sets and data science modeling. They acquired frequent shopper card data, credit card data, custom surveys, and many other data sets to create more sophisticated demographic maps. When combining this data, they had access to a large and representative data set of the US population with many dimensions and attributes capable of being analyzed. For example, credit card data might provide access to the buying habits of every US resident who uses a credit card. How an individual spends their money has great informative value. Turner combined this credit card data with additional third party datasets that provide geographic information (maps datasets) about those individuals and average income information (census data sets). Thus, you might have a data set with an individual's spending habits, geography, age, frequently visited places, and other characteristics. Finally, television advertisers like Turner likely turned to samples and experiments to directly measure the viewership behavior of many of the individuals they had in their third party data set. Turner likely has a proprietary solution to track the tv viewing behavior of certain individuals who consent to being tracked. Thus, that sampled viewership information could be combined with all of the other dimensions to make a hybrid third party - first party data set with rich insight potential.[1]

With this data set constructed, Turner likely built a clustering model to create specific audience segments. The k-means algorithm would find the optimal number of cluster centers, but Turner also had some ability to influence the clustering based on their knowledge of the datasets and what advertisers wanted. For example, information from the credit card data can show which users are heavy clothes shoppers vs which users are avid sports enthusiasts. You

[1] Will explain more about how they collected first party data in the 'Data Collection' section.

can begin to build clusters of audiences that are more rich than basic demographic dimensions and start to resemble what is possible in a digital ecosystem. When you have the clusters properly built out, you can also use some of the collected first party viewership data to understand which television spots certain clusters overindex against. Thus, you will have created both rich audience segments and an understanding of which television spots the segments are most likely to be watching.

Importantly, this data set could be continually revisited and re-analyzed. Perhaps, Turner initially built out a set of clusters and labeled the clusters in a way that advertisers were used to, but what if an advertiser wanted a specific audience not already mapped? With this macro data set, Turner would likely be able to instead create a classification model where they tried to figure out (via credit card data or any of the other dimensions) which users fit into that audience the advertiser wanted to target and include that audience in their model. They could then use their sample data about viewership to create a heat map / index of television spots that this custom audience shows greater and lesser interest in. Finally, they could tell the advertiser that the specific audience they want to target is most likely to be watching these specific programs, so it would be most optimal to serve your television ads during those spots.

Model 2: Maximize Delivery Against Target Audience with Constraint of Standard Demographic Buying

Once Turner completed their first model, it opened new possibilities for innovative advertising products. The first model did not yet create a product, it created a more powerful mapping solution that told advertisers which unique audiences over-indexed to which particular TV spots. Pre-Model 1, Turner had basic demographic understanding (M/F, age, etc.) of audiences; Post-Model 1, Turner had rich custom audiences ('car enthusiasts', 'outdoorsy', 'music aficionado', etc.) and a mapping of which audiences were watching which spots. This type of audience-placement pairing better aligned with how advertisers were buying media in digital ecosystems.

The second model had to take the outputs of the first model, and productize it. Turner knew that many advertisers were used to buying television a certain way and getting them to abandon that buying format, even with the compelling outputs of Model 1, would be jarring and difficult. Thus, Turner created a hybrid solution for advertisers. An advertiser could buy their tv placements against basic demographic guarantees (as they were used to), but they could have a secondary optimization that tried to maximize delivery to strategic target audiences the advertiser cared about. Thus, it was an optimization-constraint model like we studied in class.

- Maximize delivery against strategic target audiences that were created by Model 1 (as selected by advertiser)
- Constraint: ensure basic demographic guarantees are hit (similar to how previous television buying would work)

Advertisers could ensure their business decision makers used to standard demographic buying received results as they were used to, while still optimizing as much as possible against specific audiences that the advertiser knows are useful to their business. The outputs of Model

1 were used as a maximization function, but the constraint of the standard buying model ensured that advertisers incorporated these advanced data science techniques in ways they were comfortable with.

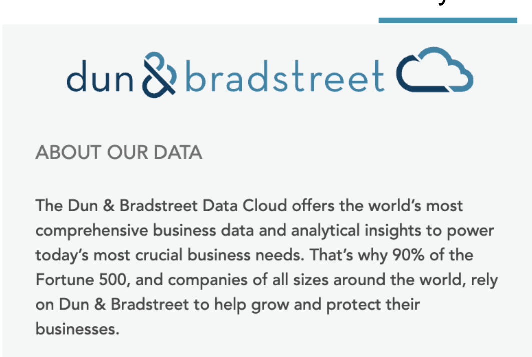
Model 3: Maximize Delivery Against Target Audience without Constraint

The third model Turner created attempted to provide advertisers with the most innovative solution possible. They removed the constraint in their optimization formula and allowed advertisers to buy media exclusively against their preferred target audiences (output of Model 1). Thus, an advertiser could select the custom audiences they wanted to target, and the optimization model would find the spots that indexed highest with those audiences and serve ads on those spots. While costs were not discussed in the article, it is very likely that Turner charged a premium for this product.

Interestingly, more constraints get introduced when you try to deploy both models 2 and 3 in real-time. If many advertisers all buy spots using different buying formats, optimizing who gets what ad becomes even more complicated. We will return to this problem on the section about combining models and will go deeper into how the optimization model might work when you introduce the challenge of multiple advertisers competing for limited spots.

Part III: The Data Needed + Data Collection

Much has already been said about the data needed to construct the three models. Model 1 was constructed with the use of third party data sets like credit card data and census data and some explanation was already provided about how these data sets were combined to create the basis for audience profiling. It may be helpful, though, to elaborate on how companies can acquire a third party data set with valuable user level information. Private companies can buy many valuable data sets and whole industries have developed to service this data need. Credit card companies sell their data to private companies for great profit. Other companies – like Dun and Bradstreet (image below)– entirely specialize on collecting data and creating valuable datasets to sell to other companies. Turner Media likely built out insights teams that developed an expertise on which datasets existed for purchase and advocated for budget to explore the datasets and assess whether they were suitable for the audience solution they were trying to create with Model 1. Constructing the macro data set used for Model 1 likely took years and quite a bit of financial investment before any of the modeling even took place.



Another aspect of the data collection that I want to elaborate on further is the 1st party data collected via surveys and experiments. This data was crucial to creating an index on which audiences viewed particular spots. This process also likely involved quite interesting data collection techniques. While I am no expert on the industry, I know that companies like Nielsen have developed solutions where households install proprietary tech into their televisions, and the tech is capable of tracking every show that those households actually watch. Turner likely built a similar solution to sample how users actually watch their programs. They likely ensured that they had a representative sample so that they could scale the insights they find to the whole US population. Thus, Turner had a representative sample of the US population and all of the tv shows that that sample actually watches. When they married that viewership data with the third party data sets they bought, this allowed them to understand which audience profiles indexed highly against particular TV spots.

Part IV: How the Models were Combined

Much has already been said about how the models were combined when describing the three models. Model 1 (clustering model) was used to feed into Model 2 and Model 3 (optimization models). You need the custom audiences created from Model 1 to optimize to them with Model 2 and Model 3. Rather than rehashing that process further, I would like to elaborate on what happens when the models go live with multiple advertisers. It is easy to optimize when there is only one advertiser and infinite spots, but when you introduce multiple advertisers buying television spots and a limited number of slots, this creates the need for another optimization model. Below is a very simplified way in which it could work.

> There are likely a few variables in this model:
> k= advertiser
> i= audience profile
> j= tv spot
> h=demographic (m/f, 25-45, etc.)
> Aik = {0,1} if audience profile i is used by advertiser k
> Bij = index score of audience profile i to spot j
> Ckj = {0,1} if spot j served to advertiser k
> Dk = binary whether advertiser k using Model 1 or Model 2
> Ejh = the demographic viewership in spot j
> Fj = cost of spot j
> Hk = total budget advertiser k

While oversimplified, this can give some idea of how the models could interact. Variable i (audience profile) and index score Bij (the index score of that audience to a particular television spot) will be outputs of Model 1. To optimize, you can then first maximize sum Aik* Ckj * Bij for all advertisers using Model 3, constrained by the advertisers total budget (Fj*Ckj <= Hk). Aik * Ckj * Bij will tell you for advertiser k who buys spot j, what is the index score of the audience profile k that to that spot (Aik will be 1 only if Advertiser k selects that audience as valuable).

Model 3 would likely be optimized first, since this was the likely more expensive product and is used for advertisers willing to exclusively try and buy their specific audiences.

Then, you can apply the same maximization function for advertisers using Model 2, but also have the constraint that they receive their fully guaranteed demographic buy (Ejh). There will likely be other constraints, like if an advertiser specifies that they wish to buy a very specific spot, that could be a constraint in the model as well.

Part V: How Often to Refresh the Model

Knowing that individuals change their preferences and attitudes over time, Model 1 will likely need to be updated quite frequently. It should not be too difficult for Turner to get refreshes on the third party data sets that they buy, and it's very likely that they have ongoing contracts with these data providers to receive refreshed data at a regular cadence. Additionally, for their first party data about viewership, they likely have this data being collected in real time so it would be possible to update this as frequently as necessary as well. Additionally, tv programs often change and new shows are introduced with little data about who watches them. You will always have to update the indexes about which audiences prefer which types of shows. Luckily, once you update Model 1, the optimization Models 2 and 3 will continue to work.

Part VI: Conclusion

Turner Media needed to innovate to keep up with the advances that were occurring in digital media. They used data science models to build richer audience profiles for advertisers to target and understood which of their programs paired well with these audiences. They then created multiple optimization models to figure out how best to productize and monetize this new audience targeting capability. While TV advertising still struggles to compete with the capabilities of digital advertising, companies like Turner Media have created compelling solutions which attempt to bridge as much of the gap as possible, often to much success.