Targeting Television Audiences using Demographic Similarity

Brendan Kitts, Liang Wei, Dyng Au, Stefanie Zlomek, Ryan Brooks, Brian Burdick

Lucid Commerce
159 S. Jackson St, Suite 200, Seattle, WA. 98104. USA.

***Abstract*—Targeting advertising on television is difficult due to limitations around ad tracking and ad delivery. This paper describes a new method of television advertising which can work with today's state of the art broadcast television media. The method works by calculating a match score between historical buyer demographics and television station-program-day-hour demographics. Television media which is very similar to the demographic of the buyer is targeted for advertising. The method is tested in a live media buy and it is shown that the method can significantly increases the performance of television advertising.**

Keywords-television; targeting; advertising; tv; demographics; similarity

I. Introduction: Television's Existential Changes

Television is a market in which existential events are forcing change. In the 1950s there were just 5 television stations. Today there are 2,100 stations. A plethora of stations are catering to increasingly fragmented micro-audiences including Spike, Oxygen, Disney, Speed, Food Network, Do-It-Yourself channel, Home and Garden TV, Cable News Network, Travel Channel, and so on. Advertising placements have also proliferated. From 20,000 advertising placements in 1950 – placed between 30 minute programs across 5 stations, there are now 11.6 million possible placements.

Technological changes have also been dramatic. All United States television broadcasts have switched from analog to digital transmission in 2009. This has made it possible for more stations and higher quality picture. The use of Digital Video Records (DVRs) which allow for time-shifting has grown 37% between 2008 and 2009 (Ferguson and Perse, 2004; Nielsen, 2008 and 2009).

Amidst all of these changes, television has not only survived, but has thrived. The average American spent almost 37 hours a week watching television in 2009. New technologies are making television more – not less attractive to consumers. As a recent Economist article put it, "Far from being cowed by new media, TV is colonizing it." (Economist, 2010).

Television advertising dollars in 2009 were approximately $65 billion. This is compared to around $23 billion for internet. Television is not a medium that is in the process of disappearing – instead it continues to improve and viewers continue to validate the television experience with unprecedented viewership levels.

The problem with television is that it is an immensely challenging medium for tracking and advertising optimization (Dawson, 1996; Kokernak, 2010). In the future, we believe that ad-networks similar to Google and Doubleclick, which are capable of delivering one-to-one advertisements, should emerge in television (Arora, et. al., 2008). IP Television is a technology currently in development that allows for different advertisement to be delivered to each set-top-box (Hart, 2004; Harte, 2009). However such technologies are not yet available.

Advertisers need an answer to ad targeting today.

This article will describe methods for targeting television using with technology that is available today – and so can immediately be rolled into production. In addition, we will show that these technologies can be utilized if or when one-to-one targeting becomes a reality in the future.

II. Previous Work

The state of television technology imposes material constraints on the kinds of targeting that can be performed. For example, because of limitations around tracking it is not possible to simply maintain a cookie of each US TV viewer's activity. It is similarly not possible to deliver personalized one-to-one advertisements to individual viewer television sets (Arora, et. al., 2008). As a result, the space of possibilities in television is constrained and must be based on the current practice of broadcast television.

Based on our analysis of the industry, we believe that there are 5 broad classes of television targeting technologies which can be used today (Fig. 1). We will describe work underway by other authors and research teams in each of these areas.

A. Contextual Television Targeting

Contextual targeting utilizes program information including program listing descriptions, movie credits, and closed captioning transcripts to match advertisements to programs. The canonical version of this technology is to use text descriptions to develop a "bag of words" representation of the program. This is then used to match against a word bag for the advertisement. The strongest vector matches are potential good matches for program-ad selection.

For example, a television program about traveling through Europe might match against advertisements for luggage.

To see how this might work we calculated some real tf-idf scores (Sparck Jones, 1972) from an actual upcoming program broadcast schedule. We used the FUSE music channel from two weeks in April 2010 in the Austin TX market, and then ranked words in order of tf-idf score. Some of the top keywords in order of tf-idf score were "Missy", (ie. "Missy Elliot"), "Gaga" (ie. presumably "Lady Gaga"). In a real contextual system, if a FUSE program featured Lady Gaga, an ad also featuring Lady Gaga ad could be selected to run with the program.

Contextual Advertising has been most successfully employed in online ad serving (Google, 2008), however only a limited amount of work on Contextual advertising has been done in television. We believe this will be a growth area for the future. Google has developed a research prototype microphone which listens to what is being watched on television and offers website recommendations (Fink , Covell and Baluja, 2006, 2008). Mandese (2010) has described work in which potential product placements are tagged in live video, and then advertising either airs at the next commercial break related to the product placement. This is closer to a contextual system, however, relies on manual tagging of product video content.

TABLE I. TF-IDF SCORES FOR KEYWORDS ON THE FUSE TELEVISION STATION

calllett ers	word	Term freque ncy	Wor d occu rren ces	Total occurr ences	tfidf
FUSE	superstars	41	4	321505	3295426
FUSE	Missy	23	5	321505	1478923
FUSE	Gaga	23	5	321505	1478923
FUSE	Cent	23	6	321505	1232436
FUSE	Jay-Z	23	8	321505	924326.9
FUSE	Elliott	23	11	321505	672237.7
FUSE	Shakur	4	3	321505	428673.4
FUSE	Tupac	4	3	321505	428673.4

B. *Viewer Response Models*

Viewer response modeling occurs where a panel of volunteer television viewers is maintained who record all of their television viewing habits. Viewer panels were developed by necessity in the 1960s due to television's inherint problems with trackability. The most commonly used panel is the Nielsen panel. This consists of 25,000 households spread around the United States. This is only around 0.2% of the US population. The viewing panels themselves range from fully electronic recording systems, to paper-based diaries.

In reality, the viewing panel is expensive, and advertisers have a diverse range of products. It is rare for advertisers to get direct information about whether the viewers bought their particular product. As a result, direct response modeling is often not possible for certain products. The panel is also extremely small compared to the US population – only 0.2%.

C. *Buyer Profile Matching*

This class of television targeting utilizes information about the product, and the kinds of people who buy it, and then performs a pattern match of this prototype with television media. This class of television targeting is new and we are not aware of other authors having proposed this method. We will discuss more about this technique in this paper.

D. *Historical Television Placement Models*

This form of television targeting uses historical data from previous advertisement airings, and their performance, in order to predict whether buying another airing with the same program-station-day-hour might be effective. Direct Response Academy (2008) discusses how television buying can be performed based on historical spot performance data, and also taking into account seasonality. Tellis et. al. (2005) present an automated system of this kind, which includes lag-terms for ad placements, and responses collected over the past several hours.

This form of ad targeting method is likely to be the most accurate, however, it also has the least amount of data. Buying advertising is extremely expensive, and as a result, there will only be a limited number of historical ad placements that can be drawn upon to predict future performance. Since there are so many programs and television stations, historical data will typically provide information on very little of the television spectrum.

E. *One-to-one Television Targeting*

One-to-one targeting entails delivering a targeted advertisement to a unique set-top-box (Arora, et. al., 2008). At the moment this technology is not available, but a variety of tests and studies of one-to-one targeting for television have been performed.

Personalized television programming has been described by Smyth and Cotter (2000) and Spangler et. al. (2003). Chorianopoulos, Lekakos, Spinellis (2003) and Lekakos and Giaglis (2004) ran experiments which tested the effectiveness of personalized advertising on television. They recruited experimental subjects and had them fill out surveys to classify them into segments. They next used a training set of users who had explicitly indicated their interest in some advertisements to predict interest in the new ads. Although not usable today, this is an area which should eventually become feasible as IP Television infrastructure becomes more prevalent (Pizzuro, 2010).

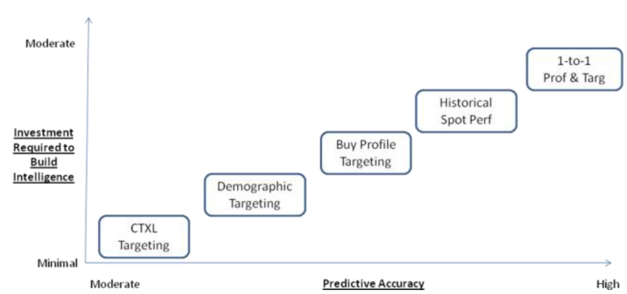

Figure 1. Types of television targeting

III. TARGETING TELEVISION BY DEMOGRAPHIC PROFILING

The basic idea of the buyer similarity targeting method is that the television media for a station-day-hour-program is represented by the demographics of people who are watching. Using this insight, we can create a "fingerprint" for the kind of customer that buys a product of interest. We can then perform a vector match against TV media looking for the closest match. After we find a close match, we recommend buying that television programming.

The method is as follows:

1. Profile a product
2. Create a television station-program profile
3. Calculate the similarity between product and station-program
4. Buy media which is most similar

We are not aware of this approach ever having been used in television media buying and we show experimental validation of the approach.

IV. Product Profile

The first step is to identify which kinds of customers would like to buy the product. In many cases, an advertiser already has a significant number of customers who have bought the product previously. These customers are known because they have had to provide their credit card number, name, address, and phone number as part of the order process.

We can take this database of customers, and enrich the customer records with demographics. For example, using the zip-code of the customer it is possible to infer their household value using US Census data. Using name it is often possible to infer gender and enthicity (eg. “Christine” -> Female, “Bob” -> Male). A variety of third party services for customer demographics exist, and can be used for this purpose. We have used Acxiom, which maintains an extensive database, and enrich to over 400 variables including income, age, gender, interests, and so on.

After enriching the customer records, we can now create an average profile for customers who have bought this product.

Let us define customer demographics as $r_{i,Dj,k}$ where D_j is the jth demographic variable for customer response i and product k. The product profile $p_{j,k}$ will be defined as the average of the many customer demographics where that demographic is not a missing value, and the customer in question purchased product $P(i,k)$. This definition assumes that all demographics are ordinal and we have transformed the variables so they meet this condition.

$$p_{j,k} = \frac{1}{I} \sum_{i:P(i,k)} r_{i,D_j,k}$$

After creating the product profile, we can now express each the product profile element as a z-score compared to the population mean and standard deviation:

$$z_{j,k} = \frac{(p_{j,k} - p)}{\sigma}$$

The operation above ensures that each profile value is re-scaled so that if its difference from the mean is scaled to units of standard deviations. Therefore, all of the dimensions are now transformed into the same z-score scale. Higher zscores means more unusually high variable compared to the population.

Table II shows a product profile for a handyman tool product. We learn that customers enjoy woodworking and auto-repair. They buy unusual amounts of “big and tall male apparel” (and so are mainly male), smoke at a higher rate than the population, and engage in outdoor activities and even like fishing.

Table III shows a product profile for a cat product. The highest z-score is “cat owner”. Although this is to be expected, it is an extremely good result – it provides validation that the demographic profile is accurately representing the product in question.

The other z-scores reveal insights about the customers which we may not have otherwise known. This includes the fact that the buyers are older, are interested in environmental issues, and give to charities.

Such insights into the buyer population could be used to optimize the advertising program. For example, because the cat buyers are interested in environmental issues, the advertising creative could be modified to mention that the cat product is bio-degradable. Since the buyers like to give to charities, the advertiser might offer to donate 5% of proceeds to a charity such as an animal shelter.

TABLE II. Demographic Product profile for project 10023 (Handyman Product).

Variable	**z-score**
Woodworking	0.588898
AutoWork	0.491806
Automotive,AutoPartsandAccessories-SC	0.48754
SportsandLeisure-SC	0.470066
HomeImprovement-Do-It-Yourselfers	0.447953
Gardening	0.438412
Camping/Hiking	0.43656
Gardening-C	0.429214
HomeImprovement	0.421209
HomeandGarden	0.417073
SportsGrouping	0.407583
HomeImprovementGrouping	0.39556
AgeinTwo-YearIncrements-InputIndividualID	0.392807
DIYLiving	0.392767
Computers	0.387011
PCOwnerID	0.386974
OutdoorsGrouping	0.385564
TruckOwner	0.369703
CreditCardHolder-UnknownType	0.369275
BankCardHolder	0.36921
HomeFurnishings/Decorating	0.365209
Apparel-Men's-C	0.362816
Income-EstimatedHouseholdID	0.360737
Crafts	0.357646
Income-EstimatedHousehold-NarrowRangesID	0.349372
AgeinTwo-YearIncrements-1stIndividualID	0.342792
Fishing	0.342675

TABLE III. DEMOGRAPHIC PRODUCT PROFILE FOR PROJECT 10019 (CAT PRODUCT).

Variable	z-score
CatOwner	0.72978
OtherPetOwner	0.507782
AgeinTwo-YearIncrements-InputIndividualID	0.497612
AgeinTwo-YearIncrements-1stIndividualID	0.470661
Pets-SC	0.460489
Community/Charities	0.382429
CollectiblesandAntiquesGrouping	0.382163
Value-PricedGeneralMerchandise-SC	0.380277
Collectibles-General	0.378843
Gardening	0.362677
EnvironmentalIssues	0.356035
Gardening-C	0.339698
HomeImprovementGrouping	0.339207
HomeLiving	0.334977
SportsGrouping	0.328068
Cooking/FoodGrouping	0.325231
Movie/MusicGrouping	0.32243
HomeFurnishings/Decorating	0.322362

V. TELEVISION STATION PROFILE

We are able to identify some customers who bought from certain television stations by using "linking keys" in the advertisement. Linking keys can be any of a (a) telephone number, (b) URL, (c) offer, that is uniquely associated with a television airing. When the customer uses the key to buy the product, we can tie that customer back to the unique broadcast where they saw the embedded linking key.

We can create an average profile for customers that we have linked to each television station program. For every television station program S_i we can define the $S_{i,Dj}$ as the jth demographic of the television station program S_i. Each station S_i is equal to the sum of its constituent spot airings and the customers who were linked to those spots. Thus each station demographic profile is an average of the customer demographic vector who purchased from airings on the station.

An example of this kind of Television station demographic profile is shown in Table IV. "Do It Yourself" ("DIY") station watchers have interest in Woodworking, Hunting, Gardening, Sport and Leisure, tend to be male (Big and Tall Male apparel). They also own dogs and smoke at a higher rate than the rest of the population. Please note the significant similarity between DIY profile and that of the handyman tool product. We will use this in the next step.

Table V shows a television station profile for "Animal Planet". The television audience for this station tends to be older, female, owns pets, and so on. There are a variety of traits also in common with the cat product.

TABLE IV. DEMOGRAPHIC PRODUCT BUYER PROFILE FOR DO IT YOURSELF CHANNEL

Variable	z-score
Gender-InputIndividualID	-3.45
Reading-FinancialNewsletterSubscribers	3.21
Apparel-Men's-BigandTall-C	2.38
Donation/Contribution-C	1.97
DogOwner	1.62
SeniorAdultinHouseholdID	1.61
Woodworking	1.57
Collectibles-Antiques	1.54
Hunting/Shooting	1.52
Gardening-C	1.49
Golf	1.38
SportsandLeisure-SC	1.32
OutdoorsGrouping	1.29
Arts	1.19
Smoking/Tobacco	1.18

TABLE V.
DEMOGRAPHIC PRODUCT PROFILE FOR ANIMAL PLANET TV CHANNEL

Variable	z-score
Gender-InputIndividualID	-3.08
Pets-SC	2.55
CatOwner	2.35
Veteran	2.30
Donation/Contribution-C	2.25
Gardening-C	2.23
Collectibles-Antiques	2.22
OtherPetOwner	2.15
Woodworking	2.14
Apparel-Children's-C	2.13
History/Military	2.09
Apparel-Women's-C	1.89
PCOperatingSystemID	1.83
EnvironmentalIssues	1.81

An interesting side-effect of being able to track station demographics, is that it is possible to report differences in some of these demographics between stations. For example, Fig. 3 shows the income levels of FOX and CNN compared to Hallmark and Game network. CNN and Fox news have higher-than-normal numbers of customers in the higher-income percentiles. Game and Hallmark have lower-than-normal numbers of customers in the lower-income percentiles.

VI. PRODUCT-TELEVISION STATION SIMILARITY

The disparity δ between the product and television station program can be calculated as below.

$$\delta(p_{i,D_j}, S_{i,D_j}) = \left| p_{i,D_j} - S_{i,D_j} \right|$$

In measuring the disparity between spot and customer response demographics, it is necessary to appropriately scale the variables to maximize the effectiveness of the match. Demographic variables range from ordinal values in the tens (e.g. age ranges from 18..80) to gender which is a two-value binary variable, 0,1. If the variables aren't scaled then in an

L1-distance calculation, the age variable would tend to exert up around 50x more "weight" on the distance match than gender. Yet gender may be just as valuable as age. Because of this, we standardize each disparity to z-scores. The transformation is

$$\Delta(r_i, s_k) = \frac{1}{J}\sum_j Z(r_{i,D_j}, s_{k,D_j})$$

$$Z(r_{i,D_j}, s_{k,D_j}) = \frac{\delta(r_{i,D_j}, s_{k,D_j}) - \frac{1}{MN}\sum_{m,n}\delta(r_{m,D_j}, s_{n,D_j})}{\sqrt{\frac{1}{MN}\sum_{i,k}\left(\delta(r_{m,D_j}, s_{n,D_j}) - \frac{1}{MN}\sum_{m,n}\delta(r_{m,D_j}, s_{n,D_j})\right)^2}}$$

Each demographic is compared against the distribution of its disparities to determine whether it is high or low compared to the norm for disparity.

As a results we now have a similarity score, with the lower value indicating better similarity between the program and station demographics.

Table VI shows the top list of stations that match the handyman product. ESPN, HGTV, DIY, HISTORY channel, Hallmark all are all very similar stations.

TABLE VI. CLOSEST TELEVISION STATIONS FOR PROJECT 10023 (HANDYMAN TOOL PRODUCT)

Project key	**Station natural key**	**Similarity**
10023	KL4	-0.305
10023	ESPNU	-0.302
10023	HGTV	-0.300
10023	DIY	-0.297
10023	DG8	-0.296
10023	DISH_Network	-0.294
10023	HALL	-0.293
10023	HIST	-0.292
10023	ESPN2_Local	-0.291
10023	CRT5	-0.288

VII. SIMILARITY VS REVENUE FOR SEVERAL PROJECTS

In order to test whether this might hold, we analyzed our historical data from previous projects. For each project we calculated the disparities between product and station. We then compared to the historical performance for that station. As a result, for each disparity between station and product, we were able to calculate the average revenue per airing. The result is shown in Fig. 2. All projects show an increase in revenue per airing as demographic similarity increases.

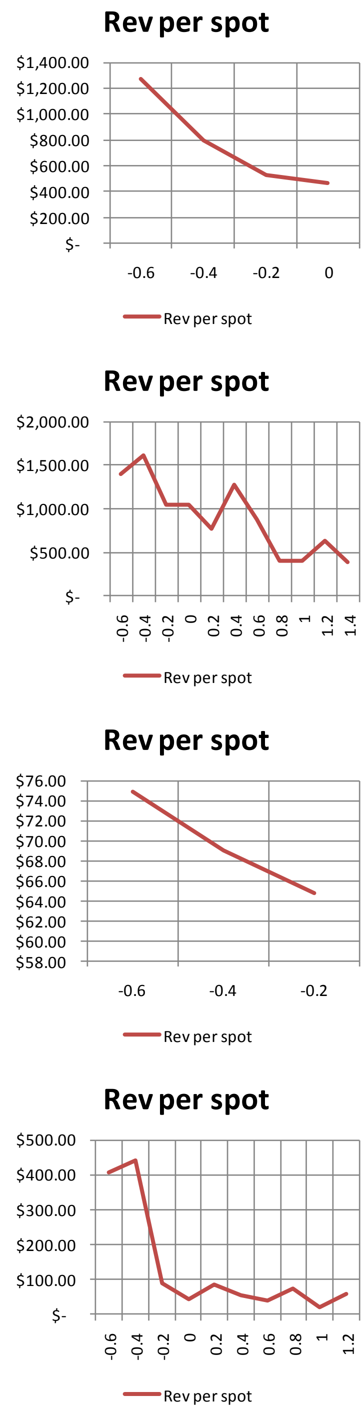

Figure 2. Revenue per spot versus similarity for 4 different projects, based on historical data available in our database on previous television airings. X-axis is similarity. These results suggest that demographic similarity between television station and product is a strong predictor of revenue performance.

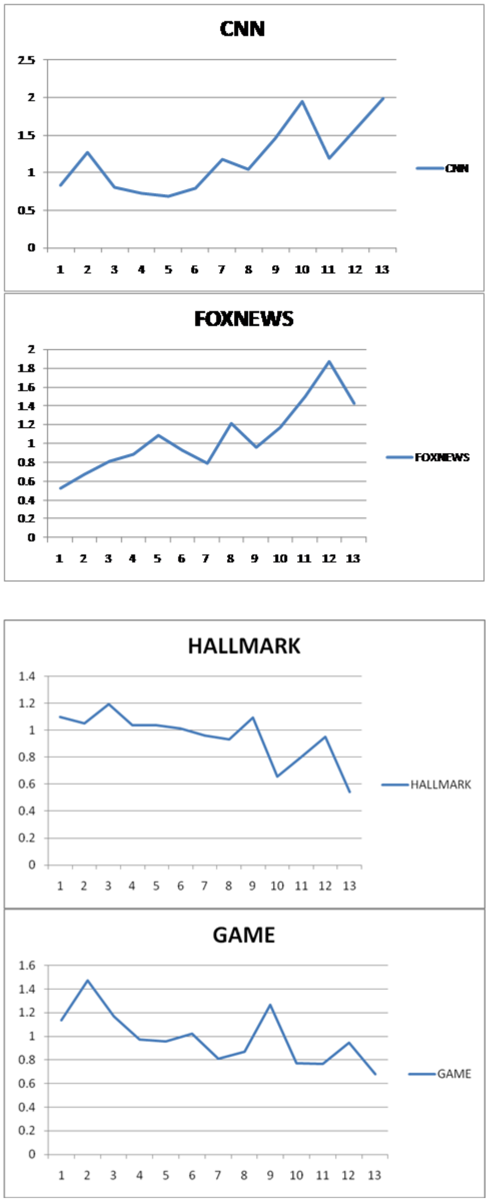

Figure 3. x-axis is income-level (13 levels), and y-axis is percentage of station customers with the income-level in question. 1.0 means that the average number of customers have this income-level. CNN and FOX both have more customers in the higher income bracket (group 10..13) and GAME and Hallmark have fewer. Therefore CNN and FOX tend to have higher income audiences.

VIII. Experiment

A. Design Of Experiment

In order to test the effectiveness of advertising using demographic targeting, we bought television advertising for a product based on demographic similarity. The product that we used is the handyman tool. A 30 minute long-form advertisement was provided by the client for this purpose. We also integrated feeds from web sales, as well as telephone call center, and retail sales, so that we could monitor the sales of the product, and most particularly, monitor sales coming in from a unique 1800 number. Although we used a long-form television program (informercial) we believe the results will hold for short-form also.

We enriched 10,000 customer records to create the product profile. The enrichment was performed in March 2010. The profile for the product is shown in Table II.

Next we created station profiles. These were created using the very large database of historical purchases that we have available – over 1.7 million customer names in linked records of sales that originated from television broadcasts.

Finally we calculated disparity between the product profile and each of the stations. As a result of the disparity calculation we ended up with 1,400 stations ranked in order of similarity. The top 10 stations by similarity are shown in Table VI. We froze these station similarities in a spreadsheet and would only look at them again after the experiment had completed.

In April and May 2010 we bought television advertising for the product. The final set of stations that we bought occupied a large spread of similarity values. 28 stations in total were purchased, with spend levels ranging from $400 to $16,000.

B. Results

The results are shown in Table VII. "Do It Yourself" Channel ("DIY"), "Fox Business" ("FOXB") and "The Outdoor Channel" ("TOC") all performed well. These stations were very close demographic matches to the product.

Fig. 4 and Fig. 5 show the shape of the similarity versus orders per spot curve. Closer similarities result in higher orders per spot. A media buyer should be able to use these similarities to buy media is they would tend to select stations which are close.

Table IX shows all possible similarity thresholds for a "simulated media buy". Using this graph we can gauge if there are any convenient similarity thresholds at which we might decide to buy / not-buy.

We find that up to a similarity of around -0.06, the lower similarity stations out-perform the others with around 4.2x better performance in terms of orders per spot. This difference is also statistically significantly as measured using a Wilcoxon rank sum test (this test looks at the comparative ranks of station performance and notes the probability that ranks from one group would appear consistently ahead of the other). A surprising amount of television station inventory is available at this threshold. In the campaign we found that we spent 60% of our advertising spend on stations this good or better!

A company looking for maximum performance could achieve even better results however. It could restrict its media campaign to only the most similar television stations to the product. If this strategy is pursued, then we could propose that similarities < -0.20 are used. In that case a 7.3x performance gain might be achieved in terms of orders per spot. At this significantly higher performance, 36% of the media budget can be spent on this higher performing inventory. This is still a very high amount of spend, given that the television assets are so much higher performing.

TABLE VII. STATIONS AND EXPERIMENTAL RESULTS

Station	Similarity	Responses Per Spot	Indexed Media Cost	Phone Orders Per Spot
DIY	-0.297	9.50	1.275	4.50
TOC	-0.265	2.83	16.299	0.45
FOXB	-0.247	20.71	9.225	0.50
TCN	-0.176	2.18	5.674	1.33
GMC	-0.165	1.50	0.595	0.00
WMBC	-0.131	1.79	5.737	0.25
WKMG	-0.127	0.00	0.425	0.64
KJZZ	-0.103	0.00	0.000	0.50
WBPX	-0.099	1.50	3.740	0.11
TVG	-0.068	1.75	1.402	0.00
DT25	-0.015	1.15	3.183	0.00
WCIA	0.007	0.08	1.908	0.38
WNOL	0.010	0.50	0.842	0.32
WCVB	0.010	1.81	4.930	0.00
KOB	0.019	1.33	0.323	0.47
WVIT	0.021	0.77	2.720	0.17
KDVR	0.022	1.86	1.632	0.00
KAYU	0.024	0.00	0.000	0.14
KGW	0.026	1.86	1.615	0.00
KHQ	0.030	0.00	0.000	0.00
KIFI	0.034	0.00	0.000	0.00
WTVD	0.039	1.33	0.383	0.00
DOC	0.043	11.25	2.176	0.52
WVBT	0.064	0.00	0.000	0.31
WOLF	0.065	4.42	2.010	0.00
WGNO	0.076	2.00	0.191	0.00
KOLN	0.083	1.11	3.634	0.00
USHOP2	0.114	0.68	1.547	0.57
KUPX	0.132	0.27	1.913	0.76

TABLE VIII. STATION DESCRIPTIONS

Station	Station name
DIY	Do It Yourself
TOC	The Outdoor Channel
FOXB	Fox Business Network
TCN	TCN
GMC	Gospel Music Channel
WMBC	WMBC
WKMG	CBS - ORLANDO,FL
KJZZ	IND - KJZZ - SALT LAKE CITY,UT
WBPX	ION - BOSTON,MA
TVG	TVG
DT25	DT25
WCIA	CBS - CHAMPAIGN,SPRINFIELD,DECATUR,IL
WNOL	CW - NEW ORLEANS,LA
WCVB	ABC - BOSTON,MA
KOB	NBC - ALBUQUERQUE,NM
WVIT	NBC - HARTFORD,CT
KDVR	FOX - DENVER,CO
KAYU	FOX - SPOKANE,WA
KGW	NBC - PORTLAND,OR
KHQ	NBC - SPOKANE,WA
KIFI	ABC - IDAHO FALLS,ID
WTVD	ABC - RALEIGH-DURHAM,NC
DOC	Documentary Channel
WVBT	FOX - NORFOLK,VA
WOLF	FOX - WILKES BARRE,PA
WGNO	ABC - NEW ORLEANS,LA
KOLN	CBS - LINCOLN,NE
USHOP2	USHOP2
KUPX	ION - SALT LAKE CITY,UT

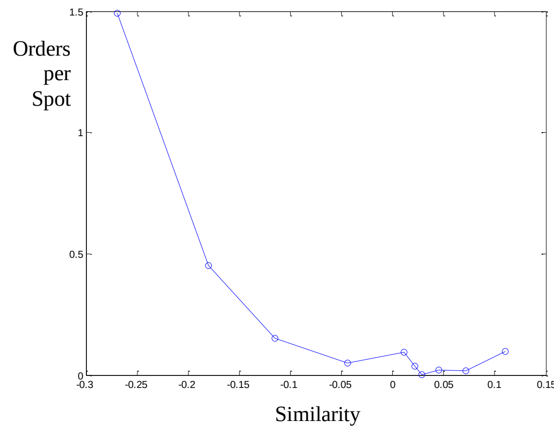

Figure 4. Phone orders per airing. Each point represents 10% of the data. Each point is a weighted average of media cost.

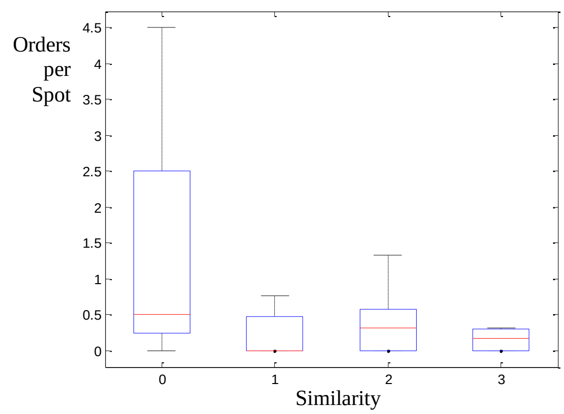

Figure 5. Box-plot of Phone orders per airing by quartile of similarity score. The closest similarity score has the largest average orders per airing, although there is also widest in terms of preformance.

TABLE IX. SIMILARITY THRESHOLD VS ADVERTISING PERFORMANCE

d	mean(orders) | similarity < d	mean(orders) | similarity > d	p-value orders (Wilcoxon)	n | similarity < d	Ratio of means	Cost | similarity < d (prop of total)
-0.2646	2.475	0.2581	0.0986	2	9.59	0.24
-0.2469	1.8167	0.2488	0.0534	3	7.30	0.37
-0.1764	1.6958	0.2054	0.0105	4	8.26	0.44
-0.1651	1.3567	0.2139	0.0678	5	6.34	0.45
-0.1313	1.1722	0.2124	0.0785	6	5.52	0.53
-0.1267	1.0966	0.1928	0.0246	7	5.69	0.53
-0.1035	1.022	0.1782	0.0114	8	5.74	0.53
-0.0988	0.9202	0.1818	0.0189	9	5.06	0.59
-0.0682	0.8281	0.1914	0.0601	10	4.33	0.60
-0.0153	0.7529	0.202	0.1484	11	3.73	0.65

0.0069	0.7214	0.1918	0.1079	12	3.76	0.67
0.0097	0.6902	0.1841	0.0839	13	3.75	0.69
0.0099	0.6409	0.1964	0.1817	14	3.26	0.75
0.0192	0.6297	0.1766	0.108	15	3.57	0.76
0.0208	0.6008	0.1773	0.1063	16	3.39	0.79
0.0219	0.5654	0.1921	0.2149	17	2.94	0.82
0.0238	0.542	0.1966	0.2254	18	2.76	0.82
0.0265	0.5134	0.2162	0.4049	19	2.37	0.84
0.0298	0.4878	0.2403	0.6599	20	2.03	0.84
0.034	0.4645	0.2703	0.9798	21	1.72	0.84
0.0389	0.4434	0.3089	0.7113	22	1.44	0.84
0.0432	0.4467	0.274	0.9555	23	1.63	0.87
0.0639	0.4409	0.2672	0.8574	24	1.65	0.87
0.065	0.4233	0.334	0.7678	25	1.27	0.90
0.0756	0.407	0.4454	0.3342	26	0.91	0.90
0.0833	0.3919	0.6681	0.0549	27	0.59	0.95

IX. Conclusion

We have presented a method for targeting television station inventory by matching the demographics of those television stations to the demographics of the product. We have also shown experimentally that stations which have closer demographics out-perform those that do not. We have shown that performance gains of 7x can be achieved, just by selecting close stations as measured by demographic similarity.

This opens up the prospect of using automated methods to select station inventory, and automatically targeting those ads to the millions of advertising placements that are currently available on the television spectrum.

We have shown this result using television stations – which are themselves an aggregate of programs. The natural extension of this work is to match programs – which should have stronger demographic profiles than the stations. Using this method, our automated system could detect upcoming programs, detect that they will have a demographic similarity that is close to the product, and recommend a buy for those stations.

We believe this method represents a significant advance in television advertising. Moreover the method is usable today using current television technology, and is experimentally re-producable. In addition, when one-to-one technologies become available, the demographic matching technology will comprise a useful feature vector that will improve the accuracy of the one-to-one targeting.

References

[1] Ardissono, L., Console, L, and Torre, I. (2001), "An adaptive system for the personalized access to news", AI Communications, Vol. 14, No. 3, pp. 129-147.

[2] Arora, N., Hess, J., Joshi, Y., Neslin, S., Thomas, J. (2008), "Putting one-to-one marketing to work: Personalization, customization, and choice", Marketing Letters, Vol. 19, pp. 305-321.

[3] Bozios, T. and G. Lekakos and V. Skoularidou and K. Chorianopoulos (2001). "Advanced Techniques for Personalized Advertising in a Digital TV Environment: The iMEDIA System", Proceedings of the eBusiness and eWork Conference pp..

[4] Chickering, D. and Heckerman, D. (2000), "Targeted Advertising with inventory management", Proceedings of the ACM Conference on Electronic Commerce, ACM Press, NY. Pp. 145-149.

[5] Chickering, D. and Heckerman, D. (2000), "A decision-theoretic approach to targeted advertising", Proceedings of the 16th Conference on Uncertainty in Artificial Intelligence, Morgan Kaufman, SF. Pp. 82-88.

[6] Chorianopoulos, K. and G. Lekakos and D. Spinellis (2003). "The Virtual Channel Model for Personalized Television", Proceedings of the 1st EuroITV conferece: From Viewers to Actors pp.59-67.

[7] Dawson, C. (1996), "Television Advertising: In Need of Reinvention?", International Journal of Marketing, Vol. 16, No. 4.

[8] Direct Response Academy, (2008), Advanced DRTV Media: Direct Response Television Media Management, Training Course materials on Direct Television media buying, http://www.directresponseacademy.com

[9] http://techcrunch.com/2006/06/08/google-research-prototypes-ambient-audio-contextual-content/.

[10] Google Blog, (2010), "Introducing Google TV", Google Blog, May 2010, http://googleblog.blogspot.com/2010/05/announcing-google-tv-tv-meets-web-web.html

[11] Ferguson, D. and Perse, E. (2004), "Audience satisfaction among TiVo and ReplayTV users", Journal of Interactive Advertising, Vol. 4, No. 2.

[12] Kokernak, M. (2010), "What's Television's Next Business Model?" Media Post Daily News, Wednesday, March 17, 2010, 1:15 PM http://www.mediapost.com/publications/?fa=Articles.showArticle&art_aid=124424

[13] Leaders, (2010), "In Praise of Television: The great survivor", The Economist, April 2010.

[14] Hart, J. (2004), "Will IPTV be more of the same, or different?" In Noam, E., Groebel, J., Gerbarg, D. (eds), Internet Television, pp. 205-214, Lawrence Erlbaum Associates.

[15] Harte, L. (2009), "Addressable TV Advertising", IPTV: IP Television Magazine, Nov 2009. http://www.iptvmagazine.com/IPTV_Magazine_Featured_Article.html

[16] Hawkins, I., Best, R., Coney, K. (1998), "Consumer behavior: Building marketing strategy", Irwin/McGraw-Hill, NY.

[17] Krulwich, B. (1997), "Lifestyle finder: Intelligent user profiling using large-scale demographic data", AI Magazine, pp. 37-45.

[18] Mandese, J. (2010), "Context Rules New TV Audience Targeting System", Media Post Daily News, Friday, June 4, 2010, 9:00 AM http://www.mediapost.com/publications/?fa=Articles.showArticle&art_aid=129490

[19] Michael Fink , Michele Covell and Shumeet Baluja (2008), "Social- and Interactive-Television: Applications Based on Real-Time Ambient-Audio Identification", http://www.mangolassi.org/covell/pubs/euroITV-2006.pdf

[20] Michele Covell & Shumeet Baluja (2006), "Interactive TV and Best Paper", Google Research Blog, Tuesday, June 06, 2006 at 6/06/2006 10:13:00 AM http://googleresearch.blogspot.com/2006/06/interactive-tv-conference-and-best.html#links

[21] Nielsen Corporation, (2008), "A2/M2 Three Screen Report: Television, Internet and Mobile Usage in the United States", 4th Quarter 2008.

[22] Nielsen Corporation, (2009), "Americans watching more TV than ever", Nielsen Wire, May http://blog.nielsen.com/nielsenwire/online_mobile/americans-watching-more-tv-than-ever/

[23] Peltier, W., Schibrowsky, A., Schultz, E. and Davis, J. (2002), "Interactive psychographics: Cross-selling in the banking industry", Journal of Advertising Research, Vol. 42, No. 2, pp. 7-22.

[24] Pizzurro, C. (2010), "The State of Addressable Television Advertising", Media Post News Daily, May 28, 2010, http://www.mediapost.com/publications/?fa=Articles.showArticle&art_aid=129149

[25] Smyth, B. and Cotter, P. (2000), "A Personalized Television Listings Service", Communications of the ACM, Vol. 43, No. 8, August 2000, pp. 107-111.

[26] Spangler, W., Gal-Or, M., May, J. (2003), "Using data mining to profile TV viewers", Communications of the ACM archive, Volume 46 , Issue 12 (December 2003), pp. 66 - 72

[27] Sparck Jones, K. (1972), "A statistical interpretation of term specificicity and its application in retrieval", Journal of Documentation.

[28] Tellis, G., Chandy, R., MacInnis, D., Thaivanich, P. (2005), "Modeling the Microeffects of Television Advertising: Which Ad Works, When, Where, for How Long, and Why?", Marketing Science, Marketing Science 24(3), pp. 351–366, INFORMS. http://www-rcf.usc.edu/~tellis/AdMicro.pdf