http://www.cisjournal.org

Case Study: Using Video Analytic Data to Target Advertisements

[1] Kalpana Algotar, [2] Sanjay Addicam, [3] Preethi Satishchandra
[1, 2, 3] Intel Corporation, 5000 West Chandler Blvd, Chandler, AZ, 85226

[1] Kalpana.a.algotar@intel.com, [2] Sanjay.v.addicam@intel.com, [3] preethi.satishchandra@intel.com

ABSTRACT

How to get target advertising in digital signage is an innovative and challenging question. This paper summarizes the result of a field trial related to targeted and non-targeted digital signage (DS) with Anonymous Video Analytics (AVA) conducted over 9 month period of time and across 10 different retailers. The goal was to understand the impact of targeted and non-targeted digital signage on sales and offers in a different environment like super market store, electronic store, and general retails. By correlating AVA viewership information with proof-of-sale (POS) data relation can be established between the response time to an ad seen by a certain demographic group and the effect on the sale of the advertised product, so as to provide retailers/advertisers with intelligence to show the right ads to right audience at right time and in the right location. The result of experiment revealed that targeted digital signage has a prominent sale impact on instant food packets, beverages, and beauty products as compared to non-targeted digital signage.

Keywords: *Targeted Advertising, Digital Signage, Anonymous Video Analytics, Data Mining, Intelligent Advertising Framework, Proof of Sale Data, Association Rules, Viewer Event Data.*

1. INTRODUCTION

Digital Signage (DS) [1,2] is a form of electronic display used to describe the term of LCD, LED, plasma, or projected displays to show constantly changing messages. Its applications vary from providing public information to creating an enhanced customer experience to show advertisements of in-store/in-venue products.

Targeted advertising is a type of advertising where advertisements are placed so as to reach certain consumers based on POS data such as purchase history and video analytic (VA) data such as demographics, observed behavior [3]. In our experiment, VA data was of an anonymous nature. We couldn't identify the viewer; we didn't store, capture, or record any images. VA data was used to predict potential consumers and targeting advertising creating a real-time relationship with potential customers; to improve their experience, and providing cross-retail with sell opportunity to boost incremental revenue. This helps to increase sale and improve retailers'/advertisers' Return on Investment (ROI) by placing right advertisement to the potential purchasers rather than playing advertisements in a predetermined order (an offline mode) state as non-targeted advertisement.

This paper propose an Intelligent Advertising Framework (IAF), which integrates AVA data [4], POS data and Data Mining technologies using RapidMiner, SQL technology to achieve targeted and interactive advertising as well as to predict potential audience. Digital signage uses AVA to identify human faces and recognize demographic information of the audiences, then correlate these information with advertisements, proof-of-sale (POS) data as well as other context information, such as time, location etc.[4], and use different data mining algorithm to generate ad category patterns based on the correlation. Moreover, AVA data used to generate a potential viewer patterns using different data mining algorithms. IAF is able to intelligently select and play the most appropriate ads to the right audience in real time using both patterns. The other sections in this paper are organized as follows. Section 2 illustrates the source of data. Section 3 describes our approach. Section 4 describes the process for deploying experiment and result of the process. Section 5 concludes the paper.

2. SOURCE OF DATA

2.1 POS Data

We categorized POS data into two categories. First, we played targeted ads over a 3 month period of time from July to September in 2013 at various stores and collected POS data as an experiment data. Second, we did not play any ads when collected POS data as a control data from April to June in 2013. POS data contains required information such as product code, product name, product category, quantity sold in each store along with month.

2.2 Viewer Event Data

AVA is privacy friendly passive, automated audience measurement technology based on computer vision theory. AVA plus POS data correlation provides digital signage operators a more effective and accurate way to measure marketing ROI by capturing audience data such as number of viewers, average attention span, and even the gender and age of viewers as they pass by a screen [4]. By correlating viewing events with ad playing events which state how effective the displayed ads are in terms of sale of that targeted product and what type of audience is being attracted. VA data contains information such as month, device id such as targeted, or non-targeted digital signage, timeslot such as morning, evening, and afternoon, age such as male and female, gender such as adult, young adult, child, and senior, adid correlated with Ad Info table which gives information about product name, product/ad category, store name, and viewing time such as how long that ads seen by viewers.

2.3 Player Event Data

From VA data we can find two types of gender and 4 groups of ages who have seen ad. By correlating this ad with player event data, we will get information about when this ad played, how long it's played and what type of age-group interested. The required data from both data sources is used in stage 2 to generate Adcategory rules.

3. OUR APPROACH

To carry out experiment, to collect the data, and to generate rules we installed two types of digital signage one is targeted and another is non-targeted at different types of stores such as retailers, electronics, supermarket, and clothing over 9 period of month. In each store we installed two types of signs in such a way that maximum number of footfalls can be captured. We collected different types of data such as viewer event data, two types of POS data – one is Control Data and another is Experiment Data of each store. We applied different mining technologies such as SQL, RapidMiner on collected data. The targeted sign had VA, and POS based rules to change ads. The non-targeted sign had regular play list. Midway through the experiments, the signs were swapped to minimize the effect of the location of the signs and to average the results.

4. EXPERIMENTS AND RESULT

Data mining technology is the process of analyzing data from different perspectives and summarizing it into useful information – information that can be used to increase revenue [5]. Here we divided experiment into four main stages and compared the result of each stage through graph which is displayed in Fig. 1. We used different data mining algorithm to discover two types of patterns. One is ad category pattern using POS data, and VA data which is used to select and play ads dynamically on DS and another is potential passer pattern from VA and player event data which is used to predict who is next in front of DS.

4.1 Stage 1

In this stage, we considered both digital signages as non-targeted. We generated offline rules using some historical data as reference and deployed those offline rules on non-targeted DS to collect data for generating rules in stage 2.

4.2 Stage 2

In this stage, we used one DS as targeted and another as non-targeted. We calculated two types of rules using data of stage 1, one is passer pattern and another is ad category. Passer pattern rules help to predict who might be next in front of the DS for next timeslot based on current viewer and Adcategory rules help to choose which ad works best for which demographic. Here, we would like to introduce two concepts, weight of previous (WOP) and weight of next (WON) which are used to predict who might be next in front of the screen. To calculate weighted count for each passer, we sliced timeslot T into 10 equal intervals such as $t_0, .., t_9$ and label all type of passer

coming within timeslot T with a position P = 0,1,..9. Repeat the above procedure for multiple timeslots using the Eq. 1 and created training dataset which includes many rows of weighted counts.

$$C = \sum_{n=0}^{9} n \ast \left(1 - \frac{p}{10}\right) \tag{1}$$

Where, n is the number of the passer such as the age group adult and gender as male etc. labeled as position P. We prepared training dataset by selecting the type of the passer with a maximum count as the dominant passer type for each time interval say t_0 of timeslot T, and specified as a WOP and passer with a maximum count for next time interval say t_1 of timeslot T specified as WON for previous time interval t_0 of timeslot T. We used training dataset with required attributes such as time, WOP, WON and trained a prediction model using different types of algorithms such as association rules, decision tree, naïvebayes in SQL, which tells the predicted dominant passer type in the next timeslot. To predict ad category we correlated VA data with player event data and prepared training dataset. We followed the same approach for ad category prediction model as we did for passer pattern prediction model. Among the different algorithm, we choose the decision tree algorithm which is having high accuracy. We deployed those patterns/rules for the time period of April to June in 2013 and collected data for stage 3. In Fig. 2, we compared the number of footfall between stage 1 and stage 2 after deploying rules. As we can see that overall footfall were increased during stage 2 (April to June, 2013) than stage 1 (March, 2103). The end result of this process helps to predict who might be next in front of the DS for next time interval of timeslot based on current viewer and to choose ad which works best for predicted passer. We deployed the both types of rules, collected data for stage 3 and plotted graph as seen in Fig. 1.

4.3 Stage 3

During stage 3, we used the same design concept of DS as stage 2 but we swapped targeted DS as non-targeted and non-targeted DS as targeted DS after some period of time. In this stage we followed same technique of stage 2 for generating passer pattern rules. Here, we generated ad category rules which pointed towards a particular advertisement using association rules of RapidMiner technology. In stage 2, rules were pointed to a particular media id/adid; certain advertisements got pitched because of higher confidence value and got played over and over again. To avoid this we adopted new tactic which pointed towards a media category and once category was selected then one ad within the media category was randomly selected and played.

We deployed the ad category rules, collected data for stage 4 and plotted graph as seen in Fig. 1. The graph shows that overall number of footfall increased during July to September month of year 2013 as compared to previous two stages say time period from March to June of year 2013. It reveals that targeting ads becoming more effective because those ads seen by viewers are more as compared to non-targeted ads. This proves that rules are

useful and effective because targeting ads were generated using those rules.

4.4 Stage 4

This stage is a combination of stage 2 and stage 3. In this stage, we collected VA data from stage 3 and POS data was provided at a monthly cadence. We selected product for which we had corresponding ads and used the POS data across the control and experiment phase to generate dataset. We prepared ad category rules using FP-Growth algorithm of RapidMiner technology on training dataset. For passer pattern rules, we used new tactic, because we had good source of VA data which is used to predict who will be next based on who is currently present by calculating WOP and WON. To calculate WOP, we sliced each timeslot T into 20 minute time intervals such as t0, t1 etc. and counted all type of passer coming within each interval of timeslot T. We calculated the confidence value for each type of passer passing by the screen using the equation (2).

$$(G\_A\_Count \ / \ Number \ Of \ interval \ in \ T)* (G\_A\_ViewingTime/ \ Number \ Of \ interval \ in \ T) \quad (2)$$

Where, G_A_Count is combination of each type of gender and age count for example number of male adult in each time interval of timeslot T. G_A_ViewingTime is total viewing time in each time interval of timeslot T. Number Of intervals in T is for example 10 means timeslot T is divided into 10 intervals of each 20 minutes.

For each time interval t of timeslot T, the higher confidence value considered as WOP and higher confidence value of the next time interval t+1 of timeslot T is considered as WON of previous time interval t of timeslot T. We deployed those rules and going to collect data for stage 5.

4.5 Stage 5

In this stage, both digital signages is used as targeted and going to use data collected in stage 4 to generate new rules using same tactic of stage 4.
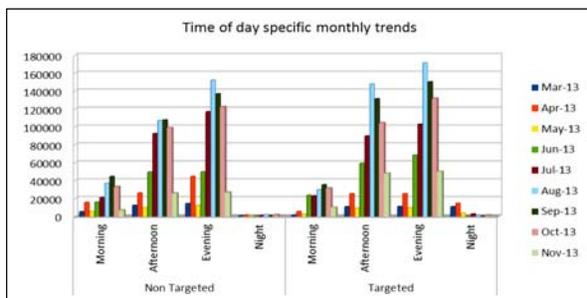

Fig 1: Time of Day Specific Monthly Trends of Different Stages (562=Targeted, 566=Non-Target)

5. CONCLUSION

This paper proposed data mining technology to generate the ad category patterns using VA and POS data to predict targeted advertising. By analyzing AVA viewership data collected from previous audience in front of a display, some viewing patterns can be discovered with data mining technology [4]. These viewing patterns can be deployed to the digital signs and then used to choose specific ads based on ad category patterns from the inventory of available content to intelligently target future audiences.

The experiment result in Fig. 2 shows the sale comparison between targeted and non-targeted digital signage as well as insight into the impact of content design. The Fig. 2 reveals that targeted ads having more sales as compared to non-targeted ads. From our experiment, we conclude that "A Targeted sign attracts on an average 1 hour of daily viewership as oppose to 23 minutes of daily viewership from a non-targeted sign and results in an average monthly sales increase of 16%".
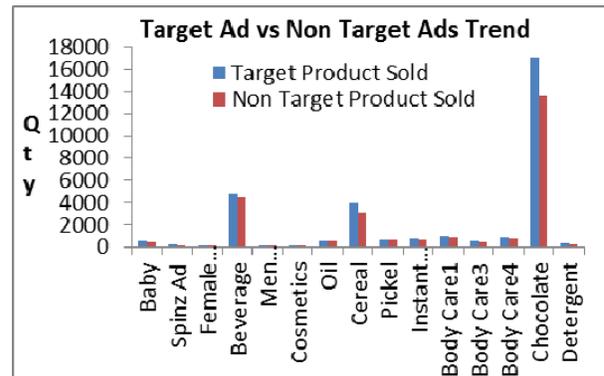
Fig 2: Comparison of Sales between Targeted and Non-Targeted DS

REFERENCES

[1]   http://en.wikipedia.org/wiki/Digital_signage, a web page on concept, standards and relevant technology of digital signage from Wikipedia.

[2]   J.Schaeffler, Digital Signage: Software, Networks, Advertising, and Displays: A Primer for Understanding the business. Focal press, Burlington, MA USA, 2008

[3]   http://en.wikipedia.org/wiki/Targeted_advertising, a web link to two principal forms of targeted advertising, behavioral targeting and contextual advertising

[4]   T. Phil et al., "Intelligent Advertising Framework for Digital signage", in Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining, 2012, pp.1532-1535

[5] http://www.anderson.ucla.edu/faculty/jason.frand/te acher/technologies/palace/datamining.htm   Data Mining

AUTHOR PROFILES

Kalpana Algotar received the degree in Computing Science Technology from Arizona State University, in 2012. She is a Software Engineer at Intel.