Problem proposal for the
31st Annual Workshop on
Mathematical Problems in Industry

Proponent: clypd, Inc.
                    212 Elm St. 4th. Floor
                    Somerville, MA 02144
                    www.clypd.com

Representative: Dr. Marco A. Montes de Oca
                    marco@clypd.com
                    http://montes-de-oca.com

Project: **Programmatic TV**

Keywords: Data Science, Optimization, Statistics, Probability.

Introduction

Why did you buy that smartphone you have in your pocket? What made you decide to purchase the laptop with which you are reading this document? Regardless of your answers to these questions, at the very origin of the chain of events that led to your final purchase, you will find *advertising*.

Advertising is fundamental to the economy. Without advertising, no one would know what goods or services are available in the marketplace, and therefore, no one would know how to (best) satisfy their needs. Advertisers are always looking for channels to spread their message, but they are particularly interested in efficient channels. One such channel is TV. In 2013, for example, it was estimated that there were more than 115 million homes in the US

with at least one TV set, which meant that about 294 million people could watch TV at home[1]. Thus, in principle, an advertiser[2] could reach approximately 93% of the US population with relatively few advertisements.

Nevertheless, not everything in advertising is about reach. What purpose would it serve to, say a car dealer, to advertise its services to a 10 year-old?[3] Or, why should a household with $50,000 yearly income care about a car that costs five times as much? When irrelevant advertisements (ads) air on TV, we tend to switch channels. This phenomenon is not just annoying to viewers, it represents a waste of money to the advertiser, and a business risk to the TV network or cable company because the advertiser might not buy again ad placements through that network/cable operator.

Ideally, the ads that people watch on TV should be about goods or services that they need (or desire) at this moment, or will need in the (near) future, and when those needs are satisfied, people should stop watching previously relevant ads and start watching new relevant ads. Our goal at clypd, Inc. is to change the TV advertising industry, so that TV ads reach the right people at the right time. This idea is called *Programmatic TV*.

Problem statement

Let us again state the goal of programmatic TV: to deliver the right advertisements to the right people at the right time. This is an easy to state, difficult to solve problem for various reasons, some of which will be described in the next few paragraphs.

Context.

To understand the challenges underlying programmatic TV, one must understand where programmatic TV sales platforms are placed in the TV ecosystem (see Fig. 1). In its simplest form, a programmatic TV sales platform receives campaign information from one or more buyers (typically advertising agencies). Each campaign consists of an actual video clip that contains the advertisement, its duration (in seconds), the start and end dates of the campaign, the target audience, the budget the advertiser is willing to spend, and an indicator of how much the advertiser is willing to spend on a per-placement basis, which we will call bid.

[1] Source: Nielsen (http://www.nielsen.com/us/en/insights/news/2013/nielsen-estimates-115-6-million-tv-homes-in-the-u-s---up-1-2-.html)
[2] We are going to use the term "advertiser" to refer to any entity interested in reaching audience via TV.
[3] Let us ignore brand-name awareness campaigns that could influence purchasing behavior in the distant future. For example, a young boy today might want to buy a BMW by the time he is an adult because he saw BMW commercials today.
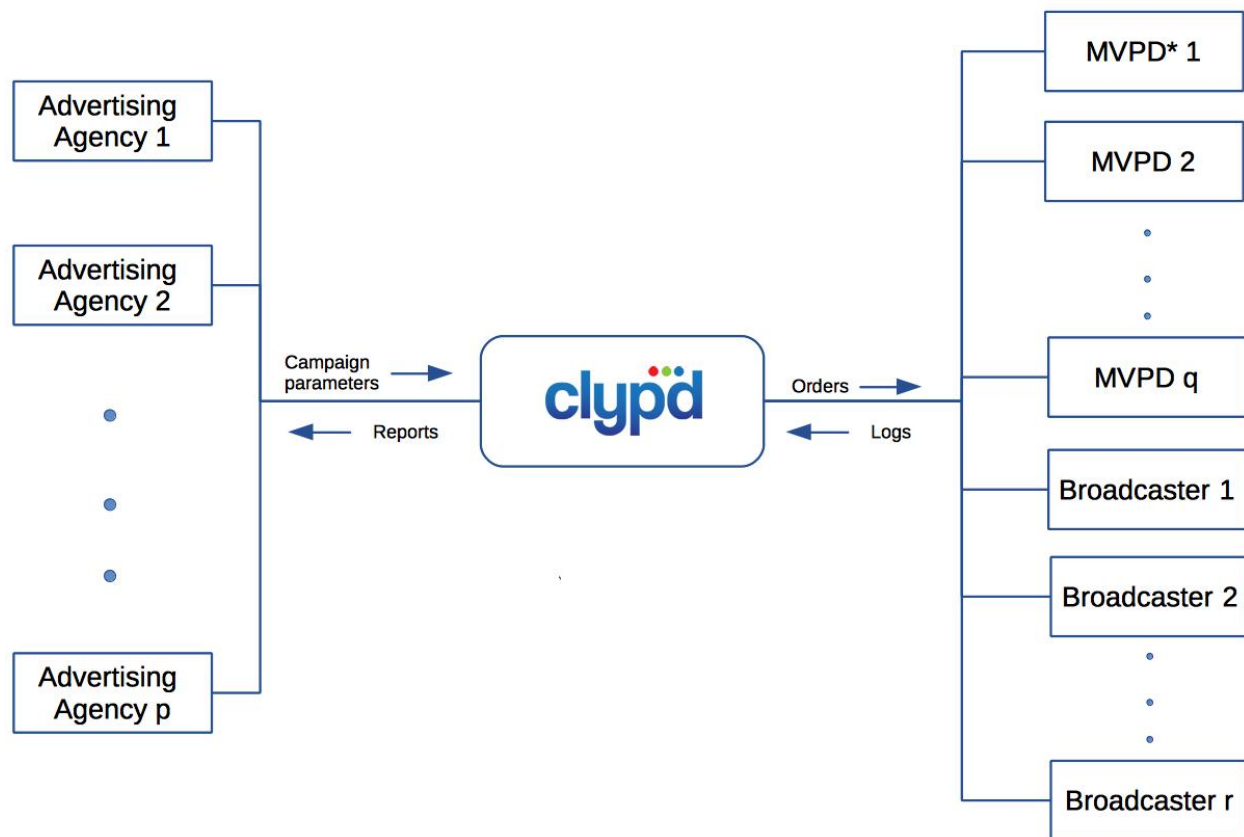
Fig. 1. A programmatic TV sales platform sits between advertisers (buyers) and TV networks/cable operators (inventory owners). *MVPD stands for multichannel video programming distributor, which is the technical term to refer to companies that deliver video programming services for a fee.

As campaigns arrive, the programmatic TV sales platform should find the right audience across all media owners' (MVPDs, broadcasters (e.g., ABC), and programmers (e.g., Discovery)) inventory (TV breaks), and schedule ads to the appropriate breaks so that the goals of the campaigns are satisfied and the media owners' revenue is maximized. Then, orders are sent to each media owner. An order's parameters are similar to a campaign's parameters in that it contains information about the ad, such as its duration, but it also specifies the exact date and time, the TV network, and geographic region the associated ad should air on, as well as the bid's amount.

Media owners' receive the programmatic TV platform's orders and process them together with other orders it might receive from other platforms or their own sales teams. This means that the orders generated by programmatic TV platforms compete with other orders for placement and therefore, they are not guaranteed to run as specified. Therefore, the programmatic TV platform must process the logs sent to it by the media owners so that it can reschedule ads based on which orders cleared, and which orders did not.

Finally, the programmatic TV platform sends buyers a report containing a summary of the campaign's performance based on rating points, or some other similar audience measurement benchmarks.

Main Challenges

1. Definition of "right." How do we determine who is the appropriate audience for a given advertisement? This question is typically answered by the advertiser itself, or perhaps by an agency. For many years, the answer to this question has been presented in terms of very broad demographic information about the target audience. For example, the target audience for a home improvement advertisement could be stated as "females aged between 24 and 54." More recently, audience profiles have been enriched with more information than just age and gender. For example, the target audience for an all-electric car could be stated as "males 18-45, not married, with an annual income greater than $150K." The underlying assumption is that people with similar demographic profiles will have similar TV viewing behaviors[4]. Despite its shortcomings, the TV advertising industry has used, and continues to use, these demographic characterizations of TV audiences as proxies for TV viewing and purchasing behavior.

TV networks and cable companies, as well as advertising agencies, rely on third-party companies to measure and characterize TV audiences. These measurements are effectively the currency in the TV industry (that is, they are exchanged for dollars). One well-known example of these measurements are the so-called Nielsen ratings, which in their most basic form are the percentage of TV households (or other group, such as adults) in the US watching a certain TV program.

Today's advertisers are in general not satisfied with targeting their campaigns at the age and gender level. They look for very specific audiences and the industry as a whole is opening to new ways to find the right audiences. For instance, one can envision experiments with ad schedules to see their effect on purchases via credit card usage in a certain geographic region. With these results, one could obtain demographic information of those households and re-target but now with a more refined profile in mind such as "households in Hartford, CT with presence of children and whose head of household is hispanic." Or one could use machine learning techniques to measure the emotional reaction of people to advertisements (e.g., using something like http://www.affectiva.com/) to determine who would respond to the advertised product/service more favorable.

---

[4] Which, of course, is akin to stereotyping.

2. Delivery. Once we have identified who watches what (see point 1 above), how do we determine where and when to schedule an ad in order to make the ad relevant to the correct audience and maximize the media owner's revenue? When no programmatic TV platform is used, this problem is tackled "by hand" with agencies typically ordering ad placements in the most popular shows in hopes that their real target audience watched their ads as well. This practice leaves breaks in not-so-popular shows, and times unsold. However, a campaign's target audience might actually watch those shows. Thus, unsold breaks might be a missed opportunity for the advertiser, and are definitely lost revenue for the media owner, who either does not air the unsold break, or places its own advertising in them, which results in absurd situations such as a Comcast customer watching a Comcast commercial.

The second challenge is therefore to associate ads with breaks with two goals in mind: (a) matching the campaign's target audience with the break's audience, and (b) maximize the media owner's revenue. From the media owner's point of view everything is about money, that is, the more money the schedule makes, the better. From the media buyer's point of view, the more "target impressions" the schedule gives them for their money the better.

3. Scale. The third challenge is related to the size of the problem. A typical MVPD has two to four minutes to place ads in per hour per network. Broadcast and cable networks can vary the ad loads on a day-to-day basis, with 15-18 minutes per hour being typical. Consider also that each media owner has presence in different geographic regions, which can be controlled independently. As the planning horizon increases, so does the number of available breaks. Very rapidly, for even short planning horizons, one has several million breaks available per media owner. Finally, the number of orders, though typically smaller than the inventory size, can be in the order of hundreds of thousands in a time window of a few months. We need to process those orders in a timely fashion as they arrive, or maybe process them in batches. The choice depends on the nature of the solution mechanism.

4. Uncertainty. The fourth and final challenge in this proposal is related to the uncertainty associated to all the measurements used throughout the programmatic TV process. Uncertainty is associated with (a) the age and gender measurements (which are national projections based on a small sample of households), (b) the derivation of more refined demographic targets, (c) the geographic projection of national estimates, (d) the forecasting of the number of viewers of a future show, and (e) the clearance of orders sent to media owners. We need a method that meets the goal of programmatic TV while guaranteeing minimum audience and revenue levels to the advertisers and the media owners, respectively.

Goals

Programmatic TV is not so much an engineering endeavor, as it is a data science effort. In this sense, we are looking for a sound mathematical framework to address the challenges described in this document.

The participants will not only learn about TV advertising, but also will have the opportunity to learn about how their mathematical skills are highly relevant in a typical data science project.

We also hope to learn about alternative mathematical approaches that could be used in a programmatic TV platform.