Intelligent Digital Signage-based Advertisement Scheduling Algorithms for Impression Maximization Problem

David Wu

Department of Computer Science, National Chiao Tung University

ABSTRACT

In recent years, the digital signage industry has experienced tremendous growth, so the topic of digital signage-based advertisement scheduling become hotter. In this paper, we study a impression maximization problem for advertisement scheduling on digital signages. We train an AD-model to make our system more intelligent to help our scheduling. We proposed two sets of advertisement scheduling algorithms, probability-based scheduling and AD-Pair scheduling. Our probability-based scheduling algorithms help us to save much time to schedule the playing order. And the AD-Pair scheduling algorithms groundbreakingly bind two Ads together as a playing unit. In the situation that there are not many, even any strong Ads, we creatively add some strong Ads (we call them "Attractions") into the advertisements set given by advertisers, to attract more people to watch the signage. We have conducted comprehensive experiments on synthetic datasets, and the experimental results show that the proposed models are suitable for effective impression maximization in digital signage advertising.

1. INTRODUCTION

Digital signage [1] is a sub segment of signage. Digital signages use technologies such as LCD, LED and projection to display news, advertisements, local announcements, and other multimedia content in public venues such as restaurants or malls. In recent years, the digital signage industry has experienced tremendous growth, and it is now only second to the Internet in terms of annual advertising revenue growth [2].

In respect of advertisements, an important problem that advertisers are concerned with is how to improve Return On Investment (ROI). A typical approach is using the concept of targeted advertising. Although targeted advertising has been adopted in many industries [3], such as banking, insurance, telecom marketing, social networks, it is still a new concept for digital signage industry. Quividi [4] and Tru-Media [5] are two popular companies that provide audience measurement solutions. Although their solutions also provide some targeted advertising capabilities, the advertising models are created manually rather than learned from historical viewership data automatically. Immersive Labs developed a prototype that realizes targeted advertising by mining the historical data. But it hasn't yet publicized any details on the used data mining algorithms and targeting accuracy.

The topic of advertisement scheduling has been studied for a long time. There are three types of scheduling patterns [6]: 1, Continuity: advertise throughout the year and evenly throughout the year; 2, Flighting: Advertise only during some months of the year; 3, Pulsing: It is a mix of both continuity and flighting, where you have a base amount of activity and you increase the media activity during some periods. In the era of Internet, the web page advertisement scheduling problem had been studied [7]. And now is the era of IOT, how to schedule advertisements on digital signages is still a new topic to research.

In this paper, we study an impression maximization problem for advertisement scheduling on intelligent digital signages. Given a set of Ads, we are required to let as many people as possible to watch these Ads before their playing deadlines. How can we make it? There are two possible ways: One is making use of the concept, targeted advertising, to improve the playing effectiveness; the other is scheduling—different Ads need different playing time and the playing order make difference. Our research focuses on the latter direction, but not just simple scheduling. We schedule advertisements intelligently! We build an AD-model which can predict an Ad's attraction level. Then we proposed two sets of advertisement scheduling algorithms: One-by-one scheduling which are on the Ads' playing probabilities, and AD-pair scheduling which pair up the Ads first and then schedule the Ad pairs. Our one-by-one scheduling algorithms first assign the playing probability to each advertisement, then generate a random number to decide which advertisement to play next. These algorithms help us to save much time to schedule the playing order. And the AD-pair scheduling algorithms groundbreakingly bind two Ads together as a playing unit. We bind a strong Ad (with high attraction level) and a weak Ad (with low attraction level) together, because we observe that many people won't watch the digital signage if the content is boring, so a strong Ad can attract people's sight to the signage which may help the next playing weak Ad. For the situation that there are not many, even any strong Ads, we creatively add some "strong Ads" (we call them "Attractions") into the advertisements set given by advertisers, to attract more people to watch the signage. In summary, our major contributions are outlined as follows:

- We formulate the impression maximization problem in digital signage-based Ads scheduling problem.

- We build an AD-model to predict the attraction level of an Ad to help our scheduling.
- We propose a set of probability-based scheduling methods and AD-pair scheduling methods. And the AD-pair scheduling is innovative and effective when the Ads pool contains strong Ads and weak Ads.
- We propose an method to actively add "Attractions" into Ads pool as strong Ads if the initial Ads pool lacks attractive Ads.
- We conduct comprehensive experiments on real and synthetic datasets, and the results show that the proposed models are effective for impression maximization in digital signage advertising.

2. RELATED WORK

The digital signage-based advertisement scheduling problem is still a new field to research, so there is no related work about this topic. Therefore, we group some related work into two categories. The first category includes some work on intelligent digital signage. The second category includes related work on scheduling algorithms.

2.1 Intelligent digital signage

In [IAF], The authors proposed an Intelligent Advertising Framework (IAF), which pioneers the integration of Anonymous Viewer Analytics (AVA) and Data Mining technologies to achieve Targeted and interactive Advertising. IAF correlates AVA viewership information with point-of-sale (POS) data, and establishes a link between the response time to an Ad by a certain demographic group and the effect on the sale of the advertised product. With the advertising models learned based on this correlation, IAF can provide advertisers and retailers with intelligence to show the right ads to right audience in right location at right time. Preliminary results indicate that IAF will greatly improve the effect and utility of advertising and maximize the Return on Investment (ROI) of advertisers and retailers.

2.2 Traditional scheduling algorithms

Scheduling problem has been studied for a long time. In Operating System Concepts, several classical CPU scheduling algorithms are proposed [8], like First-Come-First-Served (FCFS) algorithm, Shortest-Job-First (SJF) algorithm, priority-based algorithm, and Round-Robin (RR) algorithm. Most of these CPU scheduling algorithms may not be suitable for advertisement scheduling, because their executing (or playing) time are much different. For example, a process can be executed once in several milliseconds in a CPU, but an advertisement need to be played hundreds of times in a period of time. However, the Round Robin scheduling algorithm can be used in our scheduling problem. So we use it as a comparative algorithm to the algorithms proposed by us.

In 1973, C. L. Liu proposed two priority-based scheduling algorithms in a hard-real-time environment [9]: fixed priority assignment and dynamic priority assignment, which is also called dynamic deadline driven algorithm. And experiments show it is globally optimum and capable of achieving full processor utilization. The deadline driven algorithm is quite similar to our D3 algorithm, which is also dynamic probability assignment.

3. PRELIMINARIES

In this section, three parts are contained. The first part is problem definition. The impression maximization problem will be formulated. In the second part, we propose an objective function based on the impression maximization problem to estimate the efficiency of the scheduling algorithms. The framework of our work is introduced in the last part.

3.1 Problem definition

In this part, we give a formal definition of the Impression Maximization (IM) problem in digital signage-based advertisement scheduling. First we need to explain the meaning of impression.

Figure 1: An example of the impression of an Ad

Definition 1 (Impression): In the context of digital signage advertising, each time an Ad is fetched and watched by one person, it is counted as one impression. For example, in Figure 1, two people are standing in front of the signage and only one person is watching, so the impression that the advertisement gets is one.

Definition 2 (Impression Maximization, IM): Given a set of advertisements AD = ad1, ad2, ... , adn, and each advertisement has its Impression Request (IR) and Due Day (DD), (IR1, DD1), (IR2, DD2), ... , (IRn, DDn), design an advertisement playing schedule for a day to maximize the impression those advertisements can get.

3.2 Objective function

An objective function of Impression Maximization problem is given in this part. Before we look at the objective function, we will explain some nouns first.

Definition 3 (Daily Impression Request, DIR): DIR is the impression request of an advertisement for each day, which can be calculated by this formula: DIR = IR / LT , where Life Time (LT) means how many days the advertisement can be played. (LT = DD - CD. where CD means Current Day.)

Definition 4 (Daily Impression, DI): DI is the actual impression an advertisement gets in one day. And DI of an Ad is supposed to be higher than DIR, otherwise we cannot reach its impression request by its due day.

If not all Ads can reach their daily impression request, our objective is to maximize the number of successful Ads. If all Ads can reach their DIR, the objective function can be formally described as below:

$$S = \arg\max \sqrt[N]{\prod_{i=1}^{N}(|DI_i - DIR_i| + 1)^{sign(DI_i - DIR_i)}} - 1 \tag{1}$$

Where N is the number of Ads to play. This function express the geometric mean of ΔDI of N Ads. We use geometric mean rather than arithmetic mean because we expect the ΔDI of Ads are as even as possible.

3.3 Framework

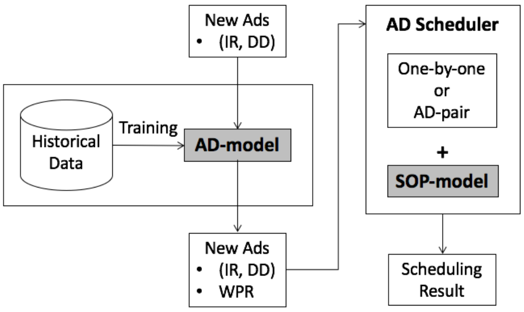


Figure 2: Framework for intelligent digital signage-based advertisement scheduling

Figure 2 shows the framework of intelligent digital signage-based advertisement scheduling. The framework mainly consists of two modules: AD-model module and scheduling module. AD-model is a predicting model which can predict the WPR of an Ad. The higher WPR represents the higher attraction level of an Ad. We use WPR to help our scheduling. We propose two sets of scheduling algorithms, one-by-one scheduling and pair-wise scheduling. One-by-one scheduling is traditional and easy-understanding. AD-pair scheduling is an innovative way to help to improve the WPR of Ads which is less attractive. And we also make use of the SOP-model to help our scheduling. SOP-model provides the stream of people data.

4. AD-MODEL

In this section, we introduce the AD-model module. Based on the concept of targeted advertising, we know that different types of people like different types of Ads (or content). For example, boys like watching NBA, so those Ads related to NBA will attract boys' attention. And girls like fashion and beauty, so those Ads about fashion or makeup will attract girls' attention. (More examples?) Ideally, we want to know the features (gender, age, and expression) of the audiences, and the features of the Ads, which can help us with the scheduling tremendously. However, due to our technical restriction, we haven't found a way to extract the features of audience from the recorded video in an enough short time. Only thing we know about the audience is their number with time. How can we make use of this information? We propose a term called fuzzy targeted advertising, which means the target is not a certain person or certain type of people but the number of people. On the other hand, we can extract some features of the Ads, and these features can influence the level of attraction of Ads.

4.1 Construction

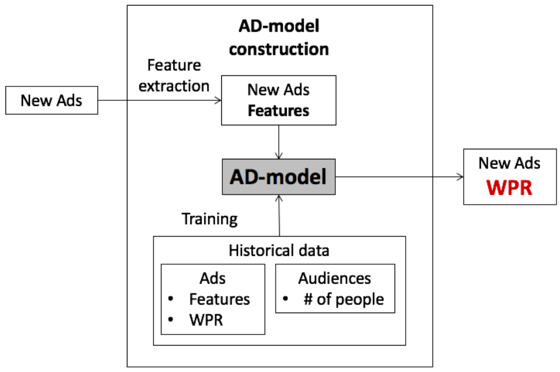


Figure 3: AD-model construction

Figure 3 shows the construction of AD-model. Given new Ads, we handcraftedly extract the features of Ads. Then we input the features to the trained AD-model based on the historical data, it will output the predicted label called WPR which represents the attraction level of an Ad.

4.2 Watch Pass Rate

In this part, we introduce an very important notion called Watch Pass Rate (WPR). The definition of WPR is as follows:

Definition 5 (Watch Pass Rate, WPR): As the term suggests, the watch pass rate of an Ad is the number of people who watch the Ad divided by the number of people who pass by the signage (or stand in front of the signage). WPR can be calculated by Formula(2). And it indicates the attraction level of the Ad. The higher WPR means the more people watching the signage when the number of passing people is the same, which means the Ad is more attractive.

$$WPR = \frac{\#\ of\ Watch}{\#\ of\ Pass} \tag{2}$$

In the example of Figure 1, the WPR is 0.5. According to Formula(2), WPR's value is between 0 and 1, owing to the number of people watching the signage is always less than passing by the signage. Because WPR is also the predicted label in our AD-model, we transform WPR from continuous value to discrete value in the following way: see Table 1.

Table 1: WPR transformation

Continuous WPR	[0, 0.2]	(0.2, 0.4]	(0.4, 0.6]	(0.6, 0.8]	(0.8, 1]
Discrete WPR	1	2	3	4	5

4.3 Feature extraction

What will influence the attraction level of an Ad? Content type? Media type? Color? We formulate a feature schema for Ads, and it contains five features: Category, Subcategory, Media, Text-percentage and Color-scheme. Media has two option: Picture and Video. Text-percentage and Color-scheme are only for picture-type Ads. Text-percentage means the percentage of text area the Ad contains, and it has ten options: 0, 0.1, 0.2, ... , 0.9. Color-scheme has two types: unicolor and multicolor (colorized). Figure 4 shows two examples of feature extraction for Ads.

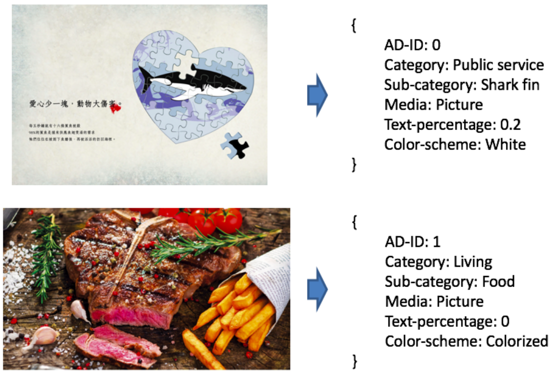


Figure 4: Two examples of feature extraction for Ads

4.4 Training and predicting

Table 2: An example of training table

AD-ID	Category	Sub-C	Media	Text-P	Color	WPR
0	Sport	Basketball	Video	-	-	5
1	Entertainment	Music	Video	-	-	4
2	Quote	-	Picture	0.3	Black	3
3	Scenery	Sea	Picture	0	Blue	2
4	Announcement	Award	Picture	0.9	White	1
...	...	...	...	...	...	...

Table 2 shows an example of training table of the AD-model. The leftmost column is the ID of Ads. The rightmost column is the label: WPR. And the middle columns are the features of Ads. We can calculate the WPR of Ads which have been played. First we get the number of passing people and the number of watching people from the video which records the behavior of people in front of the digital signage. And then we calculate the WPR using formula(2). We use these historical AD information as training data to train the AD-model. And when the new Ads' features are inputted into the AD-model, the AD-model will output their predicted WPR. Then we can use the new Ads' WPR to help our scheduling.

5. PROBABILITY-BASED SCHEDULING ALGORITHMS

In this section, we proposed three probability-based algorithms, Due Day Driven scheduling (D3), DD and IR driven scheduling (DDIR), as well as DD, IR and WPR driven scheduling (DIW). The probability-based algorithms calculate the playing probability of each Ad first, using different methods. Then we generate a series of random numbers to decide the Ads to play. Those Ads with higher playing probability will get more playing time and those with lower playing probability will get less, if the time of ?dicing? is large enough, according to the Law of Large Numbers.

5.1 Due Day Driven scheduling

Most of us do scheduling. We have all kinds of tasks to finish: homework, exams, reports, meetings, papers, etc. Most of them have a deadline, and the deadline driven scheduling method is quite common. C.L.Liu also propose a deadline driven algorithm in 1973. [C.L.Liu] In our impression maximization problem, each Ad also has a playing due day. So a possible solution, a deadline driven algorithm, is proposed. And we call it Due Day Driven (D3) scheduling algorithm. First of all, we assume that all advertisements' durations are the same. In this paper, we set them 10 seconds. Given the due day of advertisements, we calculate the Life Time (LT) of each advertisement. LT = DD - CD, where DD means Due Day and CD means Current Day. Then we pick up the advertisement with the shortest life time as the initialization, the first Ad to play in the schedule. And we calculate the playing probability of each advertisement according to the following formula:

$$P_i = \frac{\frac{1}{LT_i}}{\sum_{k=1}^{n} \frac{1}{LT_k}} \tag{3}$$

In Formula(3), we know that the shorter life time an Ad has, the higher playing probability the Ad will have, which means the more playing time the Ad will have. According to the Ads' playing probability, we form the intervals between 0 and 1. And then we generate a series of random numbers between 0 and 1, which decide the advertisements to play.

Table 3: An example of D3

(a) Given LT of Ads

	ad0	ad1	ad2	ad3	ad4
LT	6	8	9	16	11

(b) Probability matrix

Pij	ad0	ad1	ad2	ad3	ad4
ad0	0	0.32	0.29	0.16	0.23
ad1	0.39	0	0.26	0.14	0.21
ad2	0.37	0.28	0	0.14	0.21
ad3	0.34	0.25	0.23	0	0.18
ad4	0.36	0.27	0.24	0.13	0

(c) Interval matrix

Iij	ad0	ad1	ad2	ad3	ad4
ad0	-	[0, 0.32]	(0.32, 0.61]	(0.61, 0.77]	(0.77, 1]
ad1	[0, 0.39]	-	(0.39, 0.65]	(0.65, 0.79]	(0.79, 1]
ad2	[0, 0.37]	(0.37, 0.65]	-	(0.65, 0.79]	(0.79, 1]
ad3	[0, 0.34]	(0.34, 0.59]	(0.59, 0.82]	-	(0.82, 1]
ad4	[0, 0.36]	(0.36, 0.63]	(0.63, 0.87]	(0.87, 1]	-

(d) Decide the schedule by random numbers

Random numbers		0.33	0.81	0.44	0.67	...
Schedule	ad0	ad2	ad4	ad1	ad3	...

Table 3 shows an simple example of the procedure of D3 method. In Table 3(a), the life time of each Ad are given. Table 3(b) shows the probability matrix calculated using Formula (3). For example, when the last playing Ad is ad0, the probability of playing ad1 next is 0.32. Table 3(c) shows the interval matrix generated according to the probability matrix. Finally, Table 3(d) shows an series of random number and the playing schedule. The first Ad to play is ad0 because ad0 has the shortest life time, 6 days. Then the first random number is 0.33, which falls in the interval, (0.32, 0.61], so we choose ad2 as the next Ad to play. Then 0.81 falls in (0.79, 1], so we choose ad4 to play, and so on.

5.2 DD and IR driven scheduling

The deadline driven algorithm is intuitive and practical, the task with the shortest deadline will be executed first. But it there any other factor influencing the playing priority? In the real life, we try to compare two tasks: writing a paper and eating dinner. If we just consider the deadline, obviously the deadline of eating dinner is much urgency. However, we usually spend several hours to write a paper and we just spend less than one hour to have dinner each day. Because writing a paper requires much more time than eating dinner, even though the deadline of writing a paper is much longer. In our impression maximization problem, each Ad also has a impression request. It is not hard to know that the advertisements who have higher impression request should own more playing time. So the DD and IR driven algorithm (DDIR) is proposed.

The main process of DDIR is the same as D3. The only different part is the probability calculating formula:

$$P_i = \frac{\frac{IR_i}{LT_i}}{\sum_{k=1}^{n} \frac{IR_k}{LT_k}} \tag{4}$$

In Formula (4), we can see the Ads with higher impression request will own higher playing probability, and the Ads with shorter life time will own higher playing probability. Furthermore, we observe that IR/LT actually has a physical meaning, Daily Impression Request (DIR), which means that the Ads with higher DIR will have higher playing probability.

5.3 DD, IR and WPR driven scheduling

It seems we have already considered all the factors, due day and impression request, which will influence the playing probability of of Ad. But we forget the AD-model, which helps us to predict the Watch Pass Rate (WPR) of an Ad. In other word, the AD-model can predict the level of attraction of an Ad. Does the attraction level of an Ad also has something to do with its playing probability? The answer is yes. Take doing homework as an example, David and Jesse are asked to finished the same amount of homework. Finally Jesse get the homework done in one hour and David has to spend two hours, because Jesse is more clever and productive. In a similar way, to satisfy the same impression request, the more attractive Ad need less time to play. So we propose DD, IR and WPR driven scheduling (DIW). The probability calculating formula is as follows:

$$P_i = \frac{\frac{IR_i}{LT_i \cdot WPR_i}}{\sum_{k=1}^{n} \frac{IR_k}{LT_k \cdot WPR_k}} \tag{5}$$

In Formula (5), we can see that the Ad with higher WPR will have less playing probability, when their due day and impression request are the same.

Table 4 is an very simple example which shows that using different methods will lead to different results. Table 4(a) is three Ads which are given the due day (life time), impression request and watch pass rate. Table 4(b) shows the playing probability of each Ad when we use D3, DDIR and DIW respectively. When we use D3, we can see ad0 has the highest playing probability. But under DDIR, the result changes, DDIR has the highest probability. And if we use DIW, the result changes again, ad2 has the highest probability. Theoretically, DIW is supposed to beat D3 and DDIR, because it consider all the three influencing factors.

Table 4: Comparison of playing probability with different methods

(a) AD info

AD-ID	0	1	2
LT	6	8	10
IR	60	96	70
WPR	2	3	1

(b) Playing probability

AD-ID	0	1	2
$\frac{1}{LT}$	0.43	0.32	0.25
$\frac{IR}{LT}$	0.35	0.41	0.24
$\frac{IR}{LT \cdot WPR}$	0.31	0.25	0.44

5.4 Scheduling with SOP-model

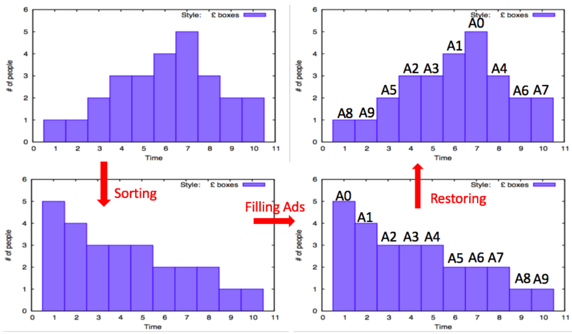


Figure 5: An toy example of scheduling with SOP-model

Stream-Of-People model (SOP-model) is a model predicting the number of people with time in the future days, according to the historical data recorded by the camera. Figure 5 shows a toy example of how we make use of the SOP-model to help our scheduling algorithms.

6. AD-PAIR SCHEDULING ALGORITHMS

Traditional advertisement scheduling methods are one-by-one playing. However, based on some experimental observation, we find that the impression of an advertisement can be influenced by the last playing advertisement. For example, an attractive advertisement (or other content) like a NBA video can improve the impression of the next advertisement which is less attractive, such as an announcement of the college. Because the NBA video has attracted people's sight, and when the signage change to play the announcement, their sight will habitually stop on the signage. Based on this phenomenon, we try to put the weak (less attractive) Ad right after the strong (more attractive) Ad when we do the scheduling, to improve the impression of the weak Ads. That is how the idea of AD-pair scheduling is born.

In this section, three innovative AD-pair scheduling algorithms will be proposed, which are Probability-based AD-Pair scheduling (PBAP), Proportion-based Slicing AD-Pair scheduling (PPSAP) and Proportion-based Slicing AD-Pair with Attraction scheduling (PPSAPA).

6.1 AD-pair model

Before we introduce AD-pair scheduling algorithms, we need to introduce our innovative AD-pair model.

Table 5: An example of one-by-one playing model and AD-pair playing model

(a) One-by-one

AD-ID	WPR
0	3
1	4
2	1
3	5
4	3
5	2
...	...

(b) AD-pair

Pair-ID	Higher WPR	Lower WPR
0	5	1
1	5	2
2	4	1
3	4	2
4	3	2
5	3	3
...	...	...

Table 5(a) shows the traditional one-by-one advertisement playing model, whose scheduling unit is one Ad. Table 5(b) shows the AD-pair playing model, whose scheduling unit is a pair of Ads. The advertisements pairing principle is a strong AD (higher WPR) plus a weak AD (lower WPR). For example, an Ad with 1 WPR is paired with an Ad with 5 WPR, and an Ad with 2 WPR can be paired with an Ad with 4 WPR. After we pair the Ads, we don't separate them during the scheduling of the day. In other word, we regard the two Ads as one to schedule and play. There are many different methods to realize the AD-pair scheduling, and three algorithms are given in this section as follows:

6.2 Probability-based AD-pair scheduling

AD-pair scheduling algorithms consist of two parts: Ads pairing and scheduling. The simplest way to pair the Ads is to sort the advertisements by their WPR, from high to low, and then we orderly pair the Ad with the highest WPR and the Ad with the lowest WPR. After pairing the Ads, we continue to use the probability-based scheduling methods to schedule the AD-pairs. That is how the Probability-based AD-pair scheduling (PBAP) is proposed.

Table 6 shows an example of PBAP. Table 6(a) is given 9 Ads with their WPR. Table 6(b) shows the action of sorting the Ads by their WPR, from high to low. The pairing method is shown in Table 6(c). We pair the Ad with the highest WPR and the Ad with the lowest WPR orderly. If the number of Ads is odd like this example, one Ad will be left at last, and we pair the left Ad with itself, which means it will be played twice at one time. In Table 6(d), we use DDIR algorithm to schedule, so we have to assign the new LT and IR to the AD-pair. There are three possible ways to assign LT and IR to the AD-pair: 1, Use the strong Ad's. 2, Use the weak Ad's. 3, Use the average value of the strong Ad and the weak Ad. In this paper, we choose the second assigning way in experiments for comparison.

6.3 Proportion-based slicing AD-pair scheduling

In PBAP, after pairing Ads, we use DDIR algorithm to schedule. But a very important problem is that the playing probability of two Ads in a pair are generally different. And we just bind two Ads according to their WPR without considering their playing probability, then assign a new probability to the AD-pair, which means there must be at least one Ad get higher or lower estimated on its playing probability. To avoid this problem, we expect to bind two Ads not just according to their WPR, but also according to their playing probability. We hope the bound Ads' playing probability are very close. This concept is like having similar value is an important factor to make a marriage happy.

Table 6: An example of probability-based AD-pair scheduling

(a) Ads with WPR

AD-ID	0	1	2	3	4	5	6	7	8
WPR	2	4	3	3	3	2	5	1	1

(b) Sort by WPR

AD-ID	6	1	2	3	4	5	0	7	8
WPR	5	4	3	3	3	2	2	1	1

(c) Ads pairing

Pair-ID	AD-ID(WPR)	AD-ID(WPR)
0	6 (5)	8 (1)
1	1 (4)	7 (1)
2	2 (3)	0 (2)
3	3 (3)	5 (2)
4	4 (3)	4 (3)

(d) Assign IR and LT

Pair-ID	0	1	2	3	4
IR	200	250	220	300	280
LT	6	8	9	16	11

However, in most cases, we can see there is a trade-off between finding opposite WPR and close probability. To solve this problem, we propose a method called Proportion-based Slicing AD-Pair scheduling (PPSAP). PPSAP replaces the playing probability of the Ads with the playing proportion of the Ads. It is like dividing a cake, and we divide the time of a day and assign them to the Ads. Some Ads get bigger proportion of playing time and some get the smaller, according to the due day, impression request and WPR. And the key process is coming here: Firstly we evenly slice the "cake" (time) to 100 proportions, and then assign them to the Ads. Some Ads get more proportions and some get less. For example, ad0 gets 7 proportions, ad1 gets 5 proportions. Secondly we evenly "slice" the Ads according to how many proportions they have. For example, ad0 will be "sliced" into 7 parts, or we can say we duplicate ad0 for 6 times (so we have 7 ad0). After doing two time of "slicing", we can pair the Ads by their WPR in spite of their proportion because their proportion are all the same, 0.01. An example of PPSAP is as follows: see Table 7.

6.4 Proportion-based slicing AD-pair with Attractions scheduling

PBAP and PPSAP make sense only if there are enough strong Ads. When the Ads pool doesn't have enough strong Ads, the AD-pair methods will lose their theory foundation and they will expectedly perform worse than one-by-one scheduling methods. To get out of this trap, we propose a method called Proportion-based Slicing AD-Pair with Attractions scheduling (PPSAPA), which actively adds some interesting content into the pool as strong Ads. We call them "Attractions", and their WPR is 6!

The disadvantages of Attractions is they occupy time and human resource to prepare them. The advantages of Attractions is to improve Ads' impression individually and globally. The former means an Attraction can improve the WPR of the next Ad to play, which is the same as the strong Ads in PPSAP. The latter influence is not so direct but can be pro-

Table 7: An example of proportion-based slicing AD-Pair scheduling

(a) Ads with WPR and P

AD-ID	0	1	2	3	4	5	6	7	8
WPR	2	4	3	3	3	2	5	1	1
P	0.081	0.110	0.100	0.090	0.119	0.120	0.110	0.130	0.140

(b) Sorted by WPR and correct P

AD-ID	6	1	2	3	4	5	0	7	8
WPR	5	4	3	3	3	2	2	1	1
P	0.11	0.11	0.10	0.09	0.12	0.12	0.08	0.13	0.14

(c) Ads slicing by P

AD-ID	6*11	1*11	2*10	3*9	4*12	5*12	0*8	7*13	8*14
WPR	5	4	3	3	3	2	2	1	1
P	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01

(d) Ads pairing

Pair-ID	AD-ID(WPR)	AD-ID(WPR)
0	6 (5)	8 (1)
1	6 (5)	8 (1)
...	...	...
11	1 (4)	8 (1)
12	1 (4)	8 (1)
13	1 (4)	8 (1)
14	1 (4)	7 (1)
...	...	...

found. Through a long-term experimental observation, we find that almost no one will watch the signage if the content on it are boring, like announcements of our college. However, after we add some Attractions, like NBA video and MTV, into the playing list, we find that much more people begin to watch the screen. After enough times of being attracted, people will form a habit of watching the screen when they pass by.

7. EXPERIMENTS

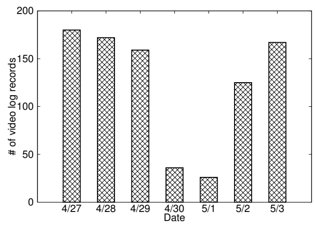
In this section, some experiments are conducted to evaluate the effectiveness of our proposed methods. There are three parts for this section. One is the evaluation of one-by-one scheduling algorithms. The second part is the evaluation of AD-pair scheduling algorithms. In the third part, we compare the one-by-one scheduling with the AD-pair scheduling to see which type is better. We implemented all the models in Python.

7.1 AD-model experiments

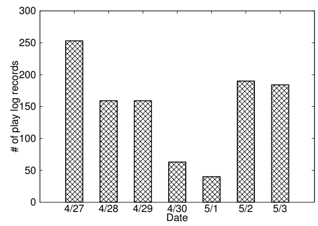
7.1.1 Dataset description

Even though we haven't found a way to convert the video log to data automatically, we do it by ourselves in order to get some real inside. Because we have months of video log and each day records 24 hours, and converting the video log to data is quite time-consuming, we choose one week's video log and from 11:00 to 13:00 in each day for our experiments. The week we choose is from April 27th (Wednesday) to May 3rd (Tuesday) in 2016.

Figure 6 shows the number of records distribution with different days for video log and Ads play log. A record in the video log is based on a person passing the digital signage. It records the person's coming-time and leaving-time, start-watch-time and end-watch-time if the person watches the signage. According to the above information, we can calculate the staying-time and watching-time (if he/she watches). We even record the gender (male or female) and age (young or senior) of people. A record in the play log is based on an Ad. It includes the start-playing-time and end-playing-time, the number of people passing the signage and the number of people watching the signage. Using the number of people passing and watching, we can calculate the WPR of the Ad.



(a) Video log

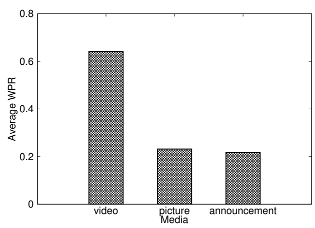


(b) Play log

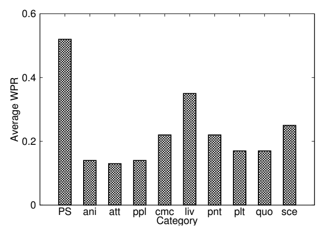
Figure 6: Number of video log and play log records with different days in a week

7.1.2 WPR distribution of different features

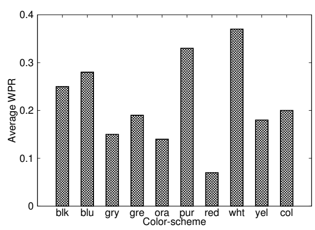
We design a feature schema for Ads to help us to predict the attraction level of Ads. Figure 7 shows how the features affect the WPR of Ads. For Media, it is very obvious that videos are much more attractive than pictures and announcements. Pictures slightly beat announcements. For Category, we have 10 kinds: Public Service Ads, Animal, Architecture, People, Comics, Living, Plant, Quote and Scenery. We can see the top 3 categories: Public Service, Living and Scenery. Public Service is real Ads which are well designed by some professional people, and the other 9 kinds of categories are just pictures we find to be substitution of Ads. That is why they are the most attractive I think. For Color-scheme, the top 3 are white, purple and blue. No obvious rule to find. For Text-percentage, it is interesting to find that 0.1 and 0.9 are the highest.



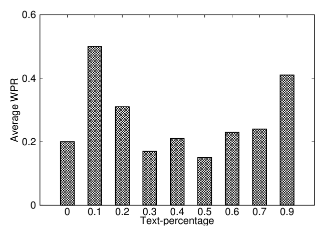
(a) Media



(b) Category



(c) Color-scheme



(d) Text-percentage

Figure 7: Average WPR distribution of different features

7.1.3 Comparison of AD-model using different algorithms

In this part, we will check out the effectiveness of our AD-model. We use the former 6 days' data as training data, and the last day's data as testing data. Though the days of training data are much longer than testing data, the factually items of training data is the same as the testing data. Because we have 56 Ads totally, and we will aggregate the records by Ad's ID first and then calculate their WPR. After aggregation, the items of training data and testing data are both 56.

Four algorithms are used for comparison to train the AD-model. They are Naive Bayes (NB), Support Vector Machine (SVM), Decision Tree (DT) and Linear Regression (LR). We just use Scikit-Learn to implement it. Figure 9 shows the comparison of effectiveness of these four algorithms. We have two evaluation criterias: accuracy and mean absolute error (MAE). Accuracy is the percentage of correctly predicted items, and it is the higher, the better. And MAE is the lower, the better. In Figure 8, we can see SVM gets the best performance: the highest accuracy and the lowest MAE.

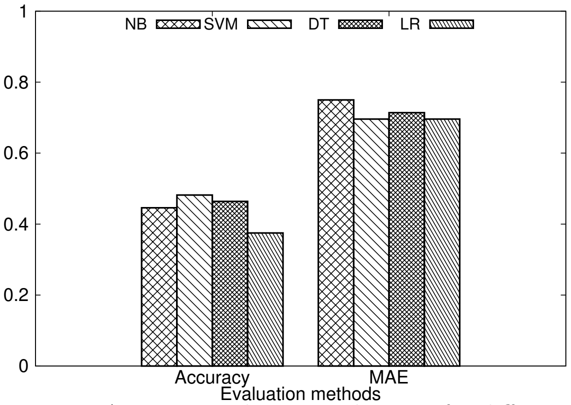


Figure 8: Accuracy and average distance for different training algorithms

7.2 Scheduling algorithms experiments

7.2.1 Dataset description

In the scheduling experiments, we use synthetic data because we cannot get enough real data yet. As we introduced in section 6 and section 7, our scheduling algorithms consider ads' IR, DD and WPR. So we directly generate these data regardless of their features. Among these three factors, WPR is the most important one. To let the synthetic data be as real as possible, we first generate a simulation datasets which is made by simulating the distribution of WPR in our field experiment. Then we consider other three possible situation of datasets, which are respectively strong, middle and weak datasets. The WPR's distribution of different datasets are shown in Figure 8.

Table 8: WPR's distribution of different synthetic datasets

WPR	1	2	3	4	5
Simulation	0.5	0.3	0.16	0.04	0
Strong	0	0	0.33	0.33	0.33
Middle	0.2	0.2	0.2	0.2	0.2
Weak	0.33	0.33	0.33	0	0

7.2.2 Simulation scheduling experiments

Table 9: Number of failed Ads with different scheduling algorithms for simulation datasets

# of Ads	40	42	44	46	48	50	52	54	56	58	60
RR	7	10	11	12	15	13	17	12	20	21	**19**
D3	7	10	17	19	17	19	21	15	25	23	28
DDIR	7	7	15	24	20	18	23	20	27	24	30
DIW	1	**0**	**0**	**5**	**6**	4	**9**	**2**	**13**	22	22
RRAP	7	10	11	12	15	13	15	12	19	**20**	**19**
PBAP	16	19	16	26	24	30	26	30	29	31	35
PPSAP	1	4	4	9	11	**2**	12	17	17	21	22
PPSAPA	**0**	3	4	7	11	**2**	14	16	16	26	24

In this part, we want to find out which scheduling algorithm is the best using the simulation datasets. We compare the the eight algorithms together, six algorithms designed by us and two are the Round Robin methods. We use the number of failed Ads to evaluate the effectiveness of these methods. Table 9 shows the result, and the best results are marked in bold. We can see DIW wins at the most time. And the followings are PPSAP and PPSAPA, which show their advantage when the number of Ads is less than 52. As the number of Ads grows, RR and RRAP turn to become the winners.

7.2.3 Comparison of one-by-one scheduling algorithms

For one-by-one AD scheduling algorithms, we propose three heuristic algorithms: Due Day Driven (D3), DD and IR driven (DDIR), and DD, IR and WPR driven (DIW). We will use the classical scheduling method, Round Robin (RR), as our baseline.

Table 10(a) shows the number of failure Ads for four one-by-one algorithms in the strong datasets. We can see all number is 0, which means these four methods all satisfy the most basic request. In this situation, we will look at the Geometric Mean of ΔDI. Figure 9(a) shows a very close result among the four methods, but we still can tell RR is slightly better and D3 is just a little bit worse. For the middle Ads datasets, Table 10(c) shows DIW is the best method because it has the most 0 with different number of Ads. DDIR is the second best, and D3 and RR are the worst. In Figure 9(b), the lines are tangly, but we do not need to get information here because Table 10(c) has distinguished the four methods. For the weak Ads datasets, in Table 10(e), we can see the former 3 methods are almost failed with all the number of Ads. Only DIW still makes it when the number of Ads is less than 22. Obviously, DIW beats others in the weak Ads datasets. Then we still look at the Geometric Mean of ΔDI. In Figure 9(c), DIW is also the best when the number of Ads varies from 16 to 22.

In summary, we can say DIW is the best algorithm among these one-by-one scheduling methods, even though RR beats others according to the Geometric Mean of ΔDI in the strong Ads datasets. Because maximizing the number of successful Ads is out first concern. As for the excessive part of the impression, we don't know how much the advertisers will pay, so it is not the biggest concern. In the middle and weak Ads datasets, the advantage of DIW is completely shown. That is to say we can handle the highest number of Ads when we are using DIW algorithm, so we can make the highest profit.

7.2.4 Comparison of AD-pair scheduling algorithms

Table 10: Comparison of # of failed Ads for one-by-one and AD-pair scheduling algorithms with different datasets

(a) One-by-one, Strong Ads

# of Ads	16	18	20	22	24	26	28	30
RR	0	0	0	0	0	0	0	0
D3	0	0	0	0	0	0	0	0
DDIR	0	0	0	0	0	0	0	0
DIW	0	0	0	0	0	0	0	0

(b) AD-pair, Strong Ads

# of Ads	16	18	20	22	24	26	28	30
RRAP	0	0	0	0	0	0	0	0
PBAP	0	0	0	0	0	0	5	1
PPSAP	0	0	0	0	0	0	0	0
PPSAPA	0	0	0	0	0	0	0	0

(c) One-by-one, Middle Ads

# of Ads	16	18	20	22	24	26	28	30
RR	0	2	1	3	3	6	8	13
D3	0	2	1	4	4	3	6	10
DDIR	0	0	0	5	6	3	7	10
DIW	0	0	0	0	0	0	2	12

(d) AD-pair, Middle Ads

# of Ads	16	18	20	22	24	26	28	30
RRAP	0	0	0	1	0	0	0	7
PBAP	0	1	0	2	2	2	3	9
PPSAP	0	0	0	0	0	0	3	16
PPSAPA	0	0	0	0	0	0	3	18

(e) One-by-one, Weak Ads

# of Ads	16	18	20	22	24	26	28	30
RR	3	2	2	7	8	11	10	17
D3	3	2	2	9	10	9	12	13
DDIR	2	2	4	9	9	8	9	13
DIW	0	0	0	1	2	11	17	27

(f) AD-pair, Weak Ads

# of Ads	16	18	20	22	24	26	28	30
RRAP	3	2	2	7	8	11	10	17
PBAP	3	4	1	11	11	12	17	18
PPSAP	0	0	0	1	3	9	19	28
PPSAPA	0	0	0	2	3	9	18	27

For AD-pair scheduling algorithms, we propose three algorithms: Probability-based AD-Pair (PBAP), Proportion-based Slicing AD-Pair (PPSAP), and Proportion-based Slicing AD-Pair with Attraction (PPSAPA). We will use the Round Robin AD-Pair (RRAP) as our baseline.

For the strong Ads datasets, in Table 10(b), we see that RRAP, PPSAP and PPSAPA are all 0. In Figure 9(d), PPSAP and PPSAPA are very close, but RRAP slightly beats them. And PBAP is the worst at any time. For the middle Ads datasets, in Table 10(d), we can see RRAP, PPSAP and PPSAPA are still the best because they have the most 0. In Figure 9(e), RRAP still performs the best, while unexpectedly PPSAP and PPSAPA perform the worst. For the weak Ads datasets, in Table 10(f), RRAP has no 0, but PPSAP and PPSAPA still get 0 when the number of Ads is less than 22. And Figure 9(f) also shows PPSAP and PPSAPA are the highest at the former three points.

In summary, we can say PPSAP and PPSAPA are the best methods, which are in our expectation, because the slicing technique overcomes the shortcoming of PBAP, which are the probability of new AD-pair cannot satisfy both the strong Ad and the weak Ad, and probability-based methods are uncertain. PPSAP and PPSAPA perform very close in our experiments. In fact, the proportion of Attraction influence the performance of PPSAPA. In our experiments, we set the proportion of Attraction is 1/101. More experiments can be done to see how the proportion of Attraction affects.

7.2.5 Comparison of one-by-one and AD-pair scheduling algorithms

Even though AD-pair scheduling methods are new and they seem to be better than one-by-one scheduling methods, we still need to verify it through experiments. To examine if AD-pair methods can beat one-by-one methods, we choose two representative methods in one-by-one, DIW and RR, and the two representative methods in AD-pair, PPSAP and RRAP, totally four methods to compare.

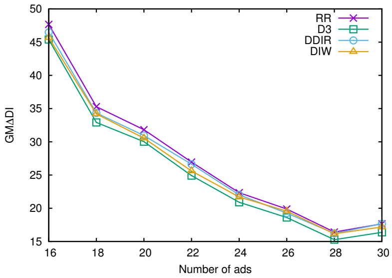
For the strong Ads datasets, we see Table 10(a) and 10(b) first. The four methods are all 0. Then we look at Figure 9(g), RR and RRAP are the same and slightly beat others. For the middle Ads datasets, if we just look at the number of failure Ads, DIW, RRAP and PPSAP have the most 0 and we cannot distinguish. Look at Figure 9(h), RRAP obviously beats the other three methods. For the weak Ads datasets, just consider the number of failure Ads, DIW and PPSAP are very close, so are RR and RRAP. And DIW and PPSAP beat RR and RRAP. In Figure 9(i), DIW beats PPSAP slightly when the number of Ads is less than 23.

In summary, AD-pair methods beat one-by-one methods in the middle Ads datasets, but in the strong and weak Ads datasets, AD-pair methods don't show any superiority and even perform worse. The results are expected, because the idea of paring Ads is based on we have both enough strong Ads and weak Ads.

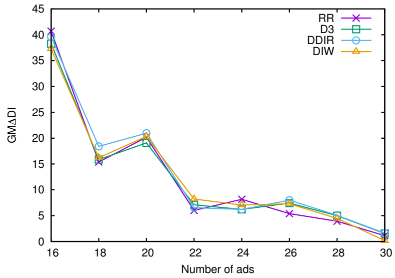
8. CONCLUSION

In this paper, we formulate an impression maximization problem for advertisement scheduling on intelligent digital signages. We propose an AD-model to predict the WPR (representing of attraction level of the Ad) of Ads, which can help our scheduling. We design a feature schema of Ads, like category, media, etc. And these features decide the attraction level of the Ad. We have a field experiment for more than a half year to record information. Even though we can not convert the records to flat data, we sample some records to convert by ourselves for getting real data. And the experiments show that the accuracy of our AD-model can be close to 0.5.

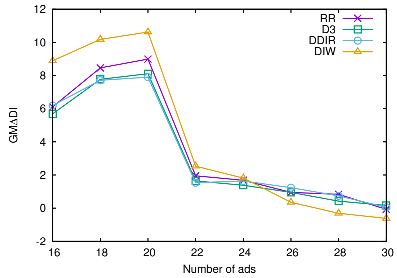
We proposed two sets of Ad scheduling algorithms: 1. probability-based (one-by-one) scheduling, including D3, DDIR and DIW; 2. AD-pair scheduling, including PBAP, PPSAP and PPSAPA. In probability-based methods, we calculate the playing probability of each Ad and then randomly decide which Ad to play. These methods are lazy and tricky because they throw a part of scheduling work to the machine. But the shortcoming of probability-based methods is the uncertainty. The fewer of the times of "dicing" (random number generated), the further the result get from our expectation. Among these probability-based algorithms, DIW, which considers all the influencing factors, has the best per-



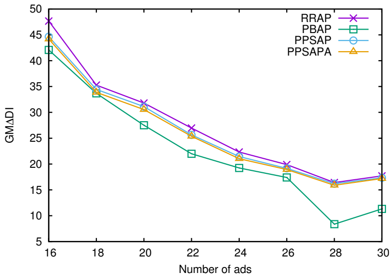
(a) One-by-one, Strong Ads



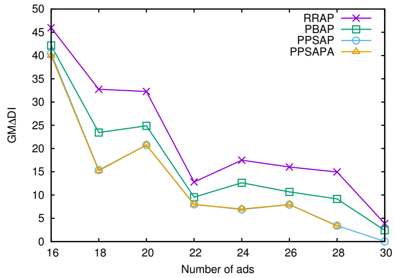
(b) One-by-one, Middle Ads



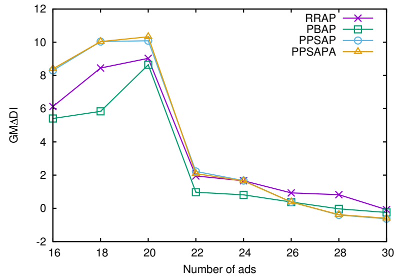
(c) One-by-one, Weak Ads



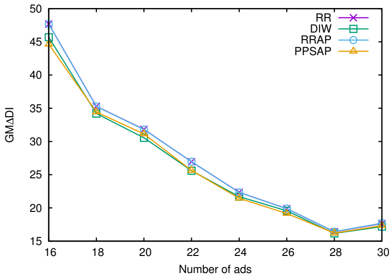
(d) AD-pair, Strong Ads



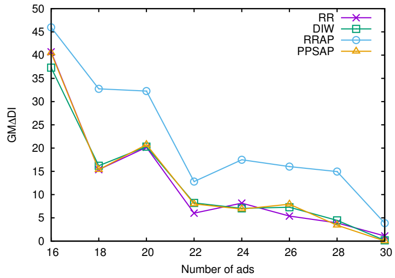
(e) AD-pair, Middle Ads



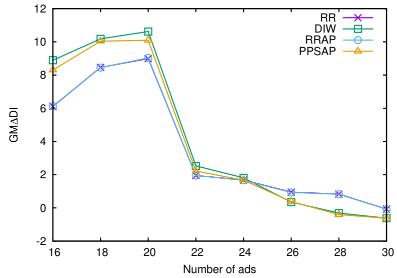
(f) AD-pair, Weak Ads



(g) One vs Pair, Strong Ads



(h) One vs Pair, Middle Ads



(i) One vs Pair, Weak Ads

Figure 9: Comparison of the Geometric Mean of ΔDI for one-by-one and AD-pair scheduling algorithms with different datasets

formance. Based on the observation that a weak Ad's impression can be improved by a strong Ad which is played before the weak Ad, we propose the AD-pair methods treating a pair of Ads as a scheduling unit. AD-pair methods are very useful when the Ads pool contains both many strong Ads and weak Ads. In the case lacking of strong Ads, we groundbreaking add "Attractions" as strong Ads to improve not only the weak Ad's impression, but also the global impression. Among AD-pair methods, PPSAP and PPSAPA are the best.

In the future, more experiments can be done. For example, comparison of using SOP-model and not using SOP-model, and how the proportion of Attraction influences the effectiveness of PPSAPA. Furthermore, we will study a more interesting work. We will consider some new features of people who appear in the camera, such as watching time, age, gender and even expression. [10][11][12]

9. REFERENCES

[1] "Digital signage." http://en.wikipedia.org/wiki/Digital_signage.

[2] J. Muller, D. Wilmsmann, and J. E. et al, "Displaying blindness: the effect of expectations on attention towards digital signage," *LNCS*, vol. 5538, pp. 1–8, 2009.

[3] "Targeted advertising," https://en.wikipedia.org/wiki/Targeted_advertising

[4] "Quividi," http://www.quividi.com/.

[5] "Tru-media," http://www.tru-media.com/.

[6] "Scheduling strategy," https://brandalyzer.wordpress.com/2012/05/21/sche strategy-for-advertisements/.

[7] S. Kumar, V. S. Jacob, and C. Sriskandarajah, "Scheduling advertisements on a web page to maximize revenue," *EJOR*, vol. 173, pp. 1067–1089, 2005.

[8] A. Silberschatz, P. B. Gallvin, and G. Gagne, *Operating System Concepts*. Wiley, 2010, ch. 5.

[9] C. L. Liu and J. W. Layland, "Scheduling algorithm for multiprogramming in a hard-real-time environment," *JACM*, vol. 20, pp. 46–61, 1973.

[10] P. Tian, A. V. Sanjay, K. Chiranjeevi, and S. M. Malik, "Intelligent advertising framework for digital signage," in *Proceedings of the 2012 ACM SIGKDD international conference on Knowledge Discovery and Data mining*. ACM, 2012, pp. 1532–1535.

[11] J. Muller, J. Exeler, M. Buzeck, and A. Kruger, "Reflective signs: digital signs that adapt to audience attention," in *Proceedings of the 2009 international conference on Pervasive Computing*. Springer, 2009, pp. 17–24.

[12] S. Clinch, N. Davies, A. Friday, and C. Efstratiou, "Reflection on long-term use of an experimental digital signage system," in *Proceedings of the 2011 international conference on Ubiquitous Computing*. ACM, 2011, pp. 133–142.