

Advertising Campaigns Management: Should We Be Greedy?

Sertan Girgin, Jeremie Mary, Philippe Preux and Olivier Nicol

Team-Project SequeL, INRIA Lille Nord Europe, and LIFL (UMR CNRS), Université de Lille, France

Abstract—We consider the problem of displaying advertisements on web pages in the “cost per click” model, which necessitates to learn the appeal of visitors for the different advertisements in order to maximize the revenue. In a realistic context, the advertisements have constraints such as a certain number of clicks to draw, as well as a lifetime. This problem is thus inherently dynamic, and intimately combines combinatorial and statistical issues. To set the stage, it is also noteworthy that we deal with very rare events of interest, since the base probability of one click is in the order of 10^{-4} . We introduce an adaptive policy learning algorithm based on linear programming, and investigate its performance through simulations on a realistic model designed with an important commercial web actor.

Index Terms—Advertisement selection, Optimization, Non-stationary setting, Linear Programming, CTR estimation, Exploration/exploitation trade-off.

I. INTRODUCTION

In this paper, we consider the problem of selecting advertisements in order to maximize the revenue earned from clicks in the “cost per click” economic model (*i.e.* each single click on an advertisement brings a certain revenue) under different settings. Our goal is not to optimize any asymptotic behavior and exhibit algorithms that are able to achieve optimal asymptotic behavior (but perform badly for much too long), but rather to solve efficiently the practical problem that arises on a web site and involves certain degrees of uncertainty originating from various sources. In section II, we formalize the problem we deal with and define a series of problems of increasing complexity, ranging from a static setting in which all information is known, to the dynamic case where key information is missing. Section III presents some experimental results in both settings. Finally, section IV concludes and we briefly discuss the lines of foreseen future works.

II. FORMALIZATION OF THE PROBLEM

In this section, we formalize the problem under study, and introduce the vocabulary and the notation used throughout the paper. We first introduce a static version of this problem, before moving to the general, dynamic case.

This research was supported, and partially funded by Orange Labs, under externalized research contract number CRE number 46 146 063 - 8360, and by Ministry of Higher Edu. and Research, Nord-Pas de Calais Regional Council and FEDER through the “Contrat de Projets Etat Region (CPER) 2007-2013”, and the contract “Vendeur Virtuel Ubiquitaire” of the “Pôle de compétitivité Industries du Commerce”. Simulations were carried out using the Grid’5000 experimental testbed, an initiative from the French Ministry of Research, INRIA, CNRS and RENATER and other contributing partners.

A. The static version of the problem

At a given time t , there is a pool of K advertising campaigns denoted by K^t . Each campaign in the pool $Ad_k \in K^t$ is characterized by a tuple $(status_k, S_k, L_k, B_k, b_k^t, r_k)$ where k is the unique identifier of the campaign. $status_k$, S_k , L_k and B_k are its status, starting time, lifetime and total click budget respectively. The campaign starts at time $t = S_k$, lasts for L_k time steps and expects to receive B_k clicks during its lifetime. The status of an advertising campaign can be either of the following: **scheduled** when the campaign will begin at some time in the future (*i.e.* $t < S_k$) and accordingly, the advertisements of this campaign can not yet be displayed, **running** when the campaign is active (*i.e.* $S_k \leq t < S_k + L_k$) and accordingly, the advertisements of this campaign can be displayed, **expired** when the campaign has ended (*i.e.* $S_k + L_k \leq t$ or $b_k^t = 0$) and accordingly, the advertisements of this campaign can no longer be displayed. $b_k^t \leq B_k$ denotes the remaining budget of the campaign at time t and r_k is the revenue obtained per click on an advertisement of the campaign k . We will use $l_k^t \in [0, L_k]$ to denote the remaining lifetime of Ad_k at time t ; it is defined as $l_k^t = \max(0, S_k + L_k - \max(S_k, t))$.

Now, the problem that we are interested in is as follows:

- The web server receives a continuous stream of visitors, each of which is assumed to be from one of N possible user profiles. The probability that the visitor belongs to a certain profile P_i is R_i with $\sum_{i=1}^N R_i = 1$.
- When a visitor visits the web site, a new “session” begins and we observe one or several iterations of the following sequence of events: (i) the visitor requests a certain page at time t (ii) the requested page is displayed to this visitor with an advertisement Ad_k embedded in it, (iii) the visitor clicks on the advertisement with probability $p_{i,k}$ where i denotes the user profile of the visitor; this probability is usually called the *click-through rate* (CTR), (iv) if there is a click, then the revenue associated with the advertisement r_k is incurred.
- After a certain number of page requests, the visitor leaves the web site and the session terminates.

The objective is to maximize the total revenue by choosing the advertisements to be displayed “carefully”. Since page requests are atomic actions, in the rest of the paper we will take a page request as the *unit of time* to simplify the discussion, *i.e.* a time step will denote a page request and vice versa. Note that in the real-world, some of the parameters mentioned above may not be known with certainty in advance. These and

other issues that we will address throughout the paper make this problem a *non-trivial* one to solve¹. In order to better understand the problem and derive our solution, we will first start with and investigate the simplest setting in which all the information is available, and subsequently move to the setting in which only a part of the information is available.

1) *Static setting with full information*: In this setting, we assume that there is a fixed time horizon T and all parameters are known; to be more precise, (a) the pool of advertising campaigns at each time step $0 \leq t < T$ is given, (b) the visit probabilities of user profiles, R_i , and their click probabilities for each campaign, $p_{i,k}$ are known, and (c) there is no uncertainty in the actual profiles of the visitors, *i.e.* we know the profile of each visitor. Note that, even if we have full information, the visitor at time t and whether he/she will click on the displayed advertisement or not are still unknown.

Under this setting, given a visitor from profile P_i at time t , one possible and efficient way to choose an advertising campaign to display would be to pick the running campaign with the highest expected revenue per click among K^t , that is $\arg\max_{Ad_k \in K^t} r_k p_{i,k}$; we will call this particular method the *highest expected value* (HEV) policy. Alternatively, we can employ a stochastic selection method where the selection probability of a running advertising campaign is proportional to its expected revenue per click. This variant will be called the *stochastic expected value* (SEV) policy.

As both policies exploit advertising campaigns with possibly high return and assign lower priority to those with lower return, one expects them to perform well if the lifetimes of the advertising campaigns are “long enough” to ensure their total click budgets. However, even under some very simple situations they may perform inferior to choosing an advertisement randomly at each step (see the example in [2] Sec. II.A.1). In order to do better, it is compulsory to take into consideration the interactions between the advertising campaigns over the entire timeline and determine which campaign to display accordingly, *i.e. do planning*. One can observe that the interactions between the advertising campaigns materialize as *overlapping time intervals* over the timeline (Fig. 1); the problem then becomes finding an *optimal allocation* of the number of advertising campaign displays in each interval.

Let s_k^t be the relative starting time of a *non-expired* campaign Ad_k at time t defined as $s_k^t = \max(0, S_k - t)$ and $e_k^t = S_k + L_k - t$ be its relative ending time. In general, given a pool of campaigns $K^t = \{Ad_1, \dots, Ad_K\}$ at time t , the time intervals during which the campaigns overlap with each other can be found from the set of their relative starting and ending times. Let $[t_0, t_1, \dots, t_M]$, $M \leq 2 \times K$, be the sorted list of elements of the set $\{x | x = s_k^t \text{ or } x = e_k^t, k \in K^t\}$. With a slight abuse of notation, we will use $k \in K^t$ and $Ad_k \in K^t$ interchangeably. By definition, the M intervals defined by $I_j = [t_{j-1}, t_j]$, $1 \leq j \leq M$ cover the entire

¹We may formulate this problem as a Markov decision problem (MDP). From a practical point of view, the state space would be huge, making its resolution very computationally demanding. However, the fact that it may be formulated as an MDP provides a proof that the problem has a solution.

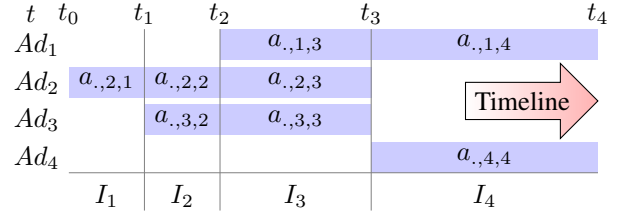


Fig. 1. The timeline divided into intervals and parts. $Ad_{1,2}$ are in scheduled state at time t_1 , and $Ad_{2,3}$ expire after t_3 . I_j denotes the j^{th} interval $[t_{j-1}, t_j]$ and $a_{k,j}$ denotes the allocation for Ad_k in interval I_j .

timeline of the pool of the advertisement campaigns. Let $AI_j = \{Ad_k | s_k < t_i \leq e_k\}$ be the set of running campaigns in interval I_j . Note that for some of the intervals, this set may be empty; these intervals are not of our interest as there will be no campaign to display during such intervals and without loss of generality we can ignore them. Let $\mathcal{A}^t = \{I_j | AI_j \neq \emptyset\}$ be the set of remaining intervals, $l_j = t_j - t_{j-1}$ denote the length of interval I_j , and $IA_k = \{I_j | Ad_k \in AI_j\}$ be the set of intervals that cover Ad_k . We can define the optimization problem that we want to solve as follows where $a_{i,k,j}$ denotes the number of displays allocated to Ad_k in interval I_j for P_i :

$$\text{maximize } \sum_{I_j \in \mathcal{A}^t} \sum_{Ad_k \in AI_j} r_k p_{i,k} a_{i,k,j} \quad (1)$$

$$\text{subject to } \sum_{Ad_k \in AI_j} a_{i,k,j} \leq R_i l_j, \forall 1 \leq i \leq N, I_j \in \mathcal{A}^t \quad (2)$$

$$\sum_{i=1}^N \sum_{I_j \in IA_k} p_{i,k} a_{i,k,j} \leq b_k^t, \forall Ad_k \in K^t \quad (3)$$

The objective function aims to maximize the total expected revenue, the first set of constraints ensures that for each interval we do not make an allocation for a particular user profile that is over the capacity of the interval (*i.e.* the portion of the interval proportional to the visit probability of the user profile), and the second set of constraints ensures that we do not exceed the remaining total click budgets. This corresponds to the maximization of a *linear objective function* ($a_{i,k,j}$ being the variables), subject to *linear inequality constraints*, which is a *linear programming* problem and can be solved efficiently.

The solution of the linear program at time t , *i.e.* the assignment of values to $a_{i,k,j}$, indicates the number of displays that should be allocated to each campaign for each user profile and in each interval, but it does not provide a specific way to choose the campaign to display to a particular visitor from user profile P_i at time t . For this, we need a method to calculate the display probability of each running campaign from their corresponding allocations, *i.e.* that maps allocations to probabilities. It is easy to see that if the first interval I_0 is not of the form $[0, l_0]$ then this means that there is no running campaign to display at time t . Otherwise, let $\bar{a}_{i,j} = \sum_{Ad_k \in AI_j} a_{i,k,j}$ be the total allocation of advertising campaign displays for the user profile P_i in interval I_j and $\hat{p}_{k,0} = a_{i,k,0} / \bar{a}_{i,0}$ be the ratio of displays allocated to the campaign Ad_k in the first interval. One can either pick the

campaign having the highest ratio, which we will call the *highest LP policy* (HLP), or employ a stochastic selection method in which the selection probability is proportional to its ratio, which will be called the *stochastic LP policy* (SLP)².

By defining and solving the linear program at each time step $0 \leq t < T$ for the current pool of non-expired campaigns (which depends on the observed visitors and their responses to the displayed advertisements), and employing one of the policies mentioned above, advertising campaigns can be displayed in such a way that the total expected revenue is maximized, ignoring the uncertainty in the predictions of the future events. In this case, the performance of HLP and SPL policies will be similar, due to the fact that the preference will gradually shift toward campaigns that have initially lower ratios as those with high ratios eventually receive more clicks reducing their remaining budgets and therefore ratios.

When the number of campaigns, and consequently the number of variables and constraints, is high, or there is a need for fast response time, solving the optimization problem at each time step may not be feasible. An alternative approach would be to solve it with regular periods and/or intermittently (such as, when the budget of a campaign is met and hence it becomes expired), and use the resulting allocation to determine the campaigns to be displayed until the next problem instance is solved, *i.e.* iterations of planning followed by multiple steps of execution. This can be accomplished by updating the allocated number of advertisement campaign displays as we move along the timeline and reducing the allocation of the chosen campaigns in the corresponding intervals. The complete algorithm is presented in the extended version of the paper ([2] Fig. 2). Note that, in practice the planning and execution steps can be asynchronous as long as the events that have occurred from the time that planning has started until its end are reflected properly to the resulting allocation.

2) Uncertainty in the static setting with full information:

The static setting with full information has two sources of uncertainty: (a) the user profiles of visitors are drawn from a categorical distribution, and (b) each campaign display is a Bernoulli trial with a certain probability, which is known, and the result is either a success (*i.e.* click) or a failure (*i.e.* no click). The aforementioned linear programming solution of the optimization problem focuses on what happens in the expectation. Following the resulting policy in different instances of the same problem, that is a certain realization of the random problem, may lead to different total revenue that vary from its expected value (see the example in [2] Sec. II.A.2)³. In reality, reducing this variability may also be important and could be considered as a secondary objective to obtaining a high total

²Note that, as we are planning for the entire timeline, the solution of the linear program at time t may not allocate any advertising campaigns to a particular user profile i , *i.e.* it may be the case that $\bar{a}_{i,0} = 0$, simply suggesting not to display any advertisement to a visitor from that user profile. In practice, when the current user is from such a user profile, choosing an advertising campaign with a low (or high) expected revenue per click instead would be a better option and likely to increase the total revenue at the end.

³Since the number of visitors from each user profile and the number of clicks on the displayed campaigns will not exactly match their expected values.

revenue. This leads to the question of how to incorporate risk-awareness to our formulation of the optimization problem.

When we look closely at Eq. 1-3, we can identify two sets of expressions of the form $R_i l_j$ and $p_{i,k} a_{i,k,j}$; the first one denotes the expected number of visitors from user profile P_i during the timespan of interval I_j , which can be considered a random variable having a Poisson distribution with parameter $\lambda = R_i t$, and the second one denotes the expected number of clicks that would be received if the campaign Ad_k is displayed $a_{i,k,j}$ times to the visitors from user profile P_i , which can be considered a random variable having a Poisson distribution with parameter $\lambda = p_{i,k} t$. Let $\mathbf{Po}(\lambda)$ denote a Poisson-distributed random variable with parameter λ . Replacing $R_i l_j$ and $p_{i,k} a_{i,k,j}$ with the corresponding random variables, we obtain the following stochastic optimization problem:

$$\max \sum_{I_j \in \mathcal{A}^t} \sum_{Ad_k \in AI_j} r_k \mathbb{E}[\mathbf{Po}(p_{i,k} a_{i,k,j})] \quad (4)$$

$$\text{s.t.} \quad \sum_{Ad_k \in AI_j} a_{i,k,j} \leq \mathbf{Po}(R_i l_j), \forall 1 \leq i \leq N, I_j \in \mathcal{A}^t \quad (5)$$

$$\sum_{i=1}^N \sum_{I_j \in IA_k} \mathbf{Po}(p_{i,k} a_{i,k,j}) \leq b_k^t, \forall Ad_k \in K^t \quad (6)$$

The summation of independent Poisson-distributed random variables also follows a Poisson distribution whose parameter is equal to the sum their parameters. Assuming that $\mathbf{Po}(p_{i,k} a_{i,k,j})$ are independent, the budget constraints in equation (6) can be written as $\mathbf{Po}(\sum_{i=1}^N \sum_{I_j \in IA_k} p_{i,k} a_{i,k,j}) \leq b_k^t, \forall Ad_k \in K^t$ which is equivalent to its linear program counterpart in expectation. The rationale behind this set of constraints is to bound the total expected number of clicks for each campaign (while at the same time trying to stay as close as possible to the bounds due to maximization in the objective function). Without loss of generality, assume that in the optimal allocation the budget constraint of campaign Ad_k is met. This means that the expected total number of clicks for Ad_k will be a Poisson-distributed random variable with parameter b_k^t and in any particular instance of the problem the probability of realizing this expectation (our target) would be 0.5. In order to increase the likelihood of reaching the target expected total number of clicks, a possible option would be to use a higher budget limit in the constraint. Let α_k be our risk factor and $\mathbf{Po}(\lambda_k)$ be the Poisson-distributed random variable having the smallest parameter λ_k such that $Pr(\mathbf{Po}(\lambda_k) > b_k^t - 1) \geq \alpha_k$ which is equivalent to $1 - \alpha_k \geq \mathbf{F}_{\mathbf{Po}(\lambda_k)}(b_k^t - 1)$ where $\mathbf{F}_{\mathbf{Po}(\lambda_k)}$ is the cumulative distribution function of $\mathbf{Po}(\lambda_k)$. Note that b_k^t and α_k are known, and λ_k can be found using numerical methods. If we replace b_k^t with λ_k in the budget constraint and solve the linear optimization problem again, the expected total number of clicks for Ad_k based on the new allocation would be greater than or equal to b_k^t and will have an upper bound of λ_k . Following the same strategy, one can derive new bounds for the user profile constraints and replace $R_i l_j$ terms in equation (5) with the corresponding parameters of the random variables. In this case, an additional set of constraints

will be necessary to ensure that for each interval the sum of allocations for all user profiles do not exceed the length of the interval: $\sum_{i=1}^N \sum_{Ad_k \in AI_j} a_{i,k,j} \leq l_j, \forall I_j \in \mathcal{A}^t$

3) *Static setting with partial information:* So far, we have assumed that the visit probabilities of user profiles and their click probabilities for each campaign are known. In reality these probabilities are hardly known in advance and have to be estimated based on observations. The simplest way to estimate unknown probabilities would be to use maximum likelihood estimation. In our problem, the profile of a visitor can be considered a categorical random variable \mathbf{R} with profile P_i having an estimated probability of \hat{R}_i , and the click of a visitor from user profile P_i on an advertisement from campaign Ad_k can be considered a Bernoulli random variable $\mathbf{p}_{i,k}$ with success probability $\hat{p}_{i,k}$. Let $visit_i^t$ denote the total number of visitors from user profile P_i that have visited the web site at time $0 \leq t$, then the maximum likelihood estimate of \hat{R}_i will be $visit_i^t / (t + 1)$, and similarly the maximum likelihood estimate of $\hat{p}_{i,k}$ at time t will be $click_{i,k}^t / display_{i,k}^t$ where $click_{i,k}^t$ is the number of times that visitors from user profile P_i clicked on advertisement Ad_k and $display_{i,k}^t$ is the number of times Ad_k had been displayed to them. Alternatively, we can employ Bayesian maximum a posteriori estimates using the conjugate priors. The conjugate priors of the categorical and Bernoulli distributions are Beta and Dirichlet distributions, respectively. If $\mathbf{Beta}(\alpha_{i,k}, \beta_{i,k})$ is the Beta prior with hyper-parameters $\alpha_{i,k}$ and $\beta_{i,k}$ for $\mathbf{p}_{i,k}$, then the posterior at time t is the Beta distribution $\mathbf{Beta}(\alpha_{i,k} + click_{i,k}^t, \beta_{i,k} + display_{i,k}^t)$. $\mathbf{Beta}(1, 1)$ corresponds to having a uniform prior. At time t , the posterior of the prior Dirichlet distribution with hyper-parameters v_i for \mathbf{R} will have hyper-parameters $v_i + visit_i^t$. The initial hyper-parameters can be guessed or determined empirically based on historical data. As we will see later in the experiment section, choosing good priors (i.e. hyper-parameters) may have a significant effect on the outcome.

By estimating probabilities at each step (or periodically) and replacing the actual values with the corresponding estimates, we can use the previous algorithm presented for the full information setting to determine allocations (optimal up to the accuracy of the estimations) and choose advertising campaigns to display. For maximum a posteriori estimates, the mode of the posterior distribution can be used as a point estimate and a single instance of the problem can be solved, or several instances of the problem can be generated by sampling probabilities from the posterior distributions, solved separately and then the resulting allocations can be merged (for example taking their mean; note that, in this case the final allocations will likely be not bound to the initial constraints). As in many online learning problems, one important issue that arises in this approach is the need for balancing the exploitation of the current estimates and exploration, i.e. estimation of the unknown or less-known (e.g., with higher variance) parameters. See [2] Sec. II.A.3 for a discussion.

B. Dynamic Setting

In this more general and realistic setting, the time horizon is no longer fixed, and furthermore new campaigns may appear with time. We will consider two main cases in which either we have a *generative model* or not; given a set of parameters and the current state, a generative model can (stochastically) generate a continuous stream of advertisement campaigns during a specified time period.

When a generative model is not available, what we have is an incomplete and uncertain image of the timeline; we know only about campaigns that have been revealed, and new advertisement campaigns may appear periodically or randomly according to a model which is unknown. In this setting, at any time step t the known pool of campaigns imposes a maximum time horizon H_{max} . Although, it is possible to apply the aforementioned methods and calculate the allocations for the known campaigns, doing so would ignore the possibility of the arrival of new campaigns that may overlap and intervene with the existing ones; the resulting *long-term* policies may perform well if the degree of dynamism in the environment is not high. On the contrary, one can focus only on short or medium-term conditions omitting the known campaigns that start after a not-too-distant time H in the future, i.e. do planning for the campaigns within the chosen planning horizon. The resulting policies will be greedier as H is smaller and disregard the long-time interactions between the existing campaigns; however, they will also be less likely to be affected by the arrival of new campaigns (see the example in [2] Sec. II.B). For such policies, choosing the optimal value of the planning horizon is not trivial due to the fact that it strongly depends on the unknown underlying model. One possible way to overcome this problem would be to solve for a set of different planning horizons $H_1, \dots, H_u = H_{max}$ (as the planning horizons are different, the structure of the optimization problems would also be different from each others) and then combine the resulting probability distributions of advertising campaign displays (such as by majority voting).

When a generative model of advertising campaigns is available, it can be utilized to compensate for the uncertainty in future events. In this case, in addition to the known pool of campaigns, the model allows us to generate a set of *hypothetical* campaigns (for example, up to H_{max}), simulating what may happen in future, and include them in the planning phase. By omitting allocations made for these hypothetical campaigns from the (optimal) allocation scheme found by solving the optimization problem, display probabilities that inherently take into consideration the effects of future events can be calculated. Note that, this would introduce bias to the resulting policies which can be reduced by running multiple simulations and combining their results as discussed before.

III. EXPERIMENTS

Our approach was tested on a toy-model designed with experts from Orange Labs⁴ to fit the real-world problem.

⁴The research division of an important commercial web actor with tens of millions of page views per day over multiple web sites.

We took care that each advertisement campaign has its own characteristics that more or less appeal to the different visits. The model assumes that each campaign A_k has a *base click probability* p_k that is sampled from a known distribution (e.g. uniform in an interval, or normally distributed with a certain mean and variance). As clicking on an advertisement is in general a rare event, the base click probabilities are typically low (around 10^{-4}). The click probability of a visitor from a particular user profile is then set to $p_{i,k} = p_k \gamma^{d-1}$ where $\gamma > 1$ is a predefined multiplicative coefficient and the random variable \mathbf{d} is sampled from the discrete probability distribution with parameter n that has the following probability mass function $Pr[\mathbf{d} = x] = 2^{n-x}/(2^n - 1), 1 \leq x \leq n$. When n is small, all campaigns will have similar click probabilities that are close to the base click probability; as n increases, some campaigns will have significantly higher click probabilities for some but not all of the user profiles⁵. In the experiments we used two values for the γ parameter, 2 and 4; experts recommended use of the latter value, but as we will see shortly having a higher γ value may be advantageous for the greedy policy. The value of n is varied between 2 and 6.

Similar to the way that we introduce the proposed method, in the experiments we will also proceed from simpler settings to more complex ones. Due to the space limitations, we opted to focus on core measures and therefore omit some of the extensions that have been discussed in the text. We begin with the static setting with full information, and consider a fixed time horizon of one day (assumed to be equivalent to 4 million page visits). The distribution of user profiles is uniform and the budget and lifetime of campaigns are also sampled uniformly from fixed intervals. In order to determine the starting times of campaigns, we partitioned the time horizon into M equally spaced intervals (in our case 80) and set the starting time of each advertisement to the starting time of an interval chosen randomly such that the ending times do not exceed the fixed time horizon. The base click probability is set to 0.0001. We solved the optimization problem every 10000 steps.

Fig. 2 shows the relative performance of HLP policy with respect to the HEV policy for different values of the click probability generation parameter n and budget for the case in which there is a single user profile and 40 campaigns with an average lifetime of $1/10^{th}$ of the time horizon; all campaigns have the same budget. We can make two observations, all other parameters being fixed HLP is more effective with increasing budgets, and the performance gain depends largely on the value of γ . For $\gamma = 4$, which is considered to be a realistic value by experts, and reasonable budgets the greedy policy would perform well. A similar situation also arises when the number of campaigns is low, whereas increasing the number of user profiles favors planning (Fig. 3).

Next, we tried longer static settings of over one week period with and without full information. The campaign lifetimes

⁵Note that, the number of such assignments will be exponentially low; for fixed γ , the number of campaigns with click probability p will be twice that of with click probability γp . This allows us to model situations in which a small number of campaigns end up being popular in certain user profiles.

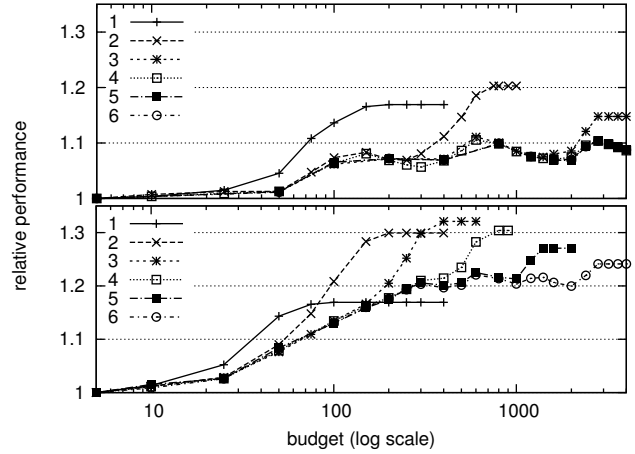


Fig. 2. The relative performance of the HLP policy with respect to the HEV policy for different values of the click prob. generation parameter n under the static setting with one profile and 40 campaigns. γ is 2 (bottom) and 4 (top).

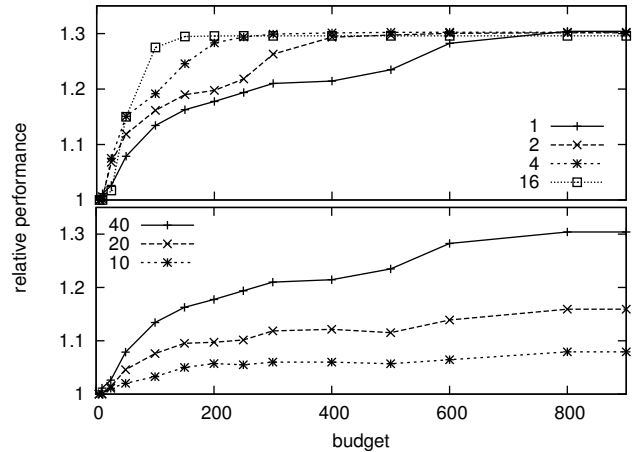


Fig. 3. The effect of the number of user profiles (top) and campaigns (bottom) when $n = 2$, $\gamma = 4$ and other parameters are kept constant.

and their budget were more realistic (2-5 days, 500-4000 clicks). The campaigns are generated on a daily basis at the beginning of a run, i.e. a set of 7-9 new advertisement arrives at every 4 million steps. We tested different values for the click probability generation parameters. There were 8 user profiles with equal visit probabilities. As presented in Fig. 4 (a), in this setting although HLP policy performs better than the greedy policy, the performance gain stays limited. While the greedy policy quickly exploits and consumes new advertisements as they arrive, HLP tends to keep a consistent and uniform click rate at the beginning and progressively becomes more greedy towards the end of the period (Fig. 5). Fig. 4 (b) shows the effect of the planning horizon, i.e. when we focus on near future and ignore or do not have information about distant events; since we are not in the dynamic setting, using less information than available hinders the performance. Note that, this prominently depends on the degree of interaction between the campaigns and in this and other experiments we observed

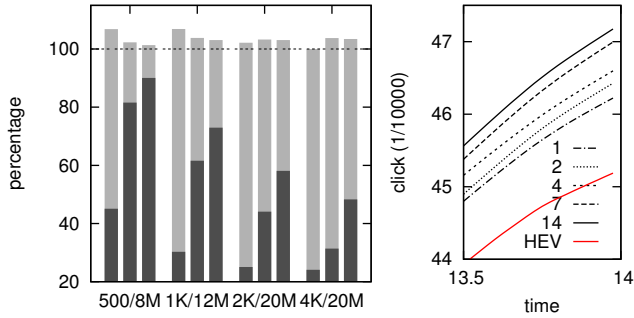


Fig. 4. (a) The performance of the random (dark gray) and the HLP (light gray) policies with respect to the HES policy under the 7 days static setting for different budget (500 to 4000), lifetime (2-5 days) and n values. The three sets of bars in each group corresponds to the case where $n = 2, 4$ and 6 in that order. (b) The effect of horizon (1, 2, 4, 7 and 14 days) in the 14 days static setting with full information.

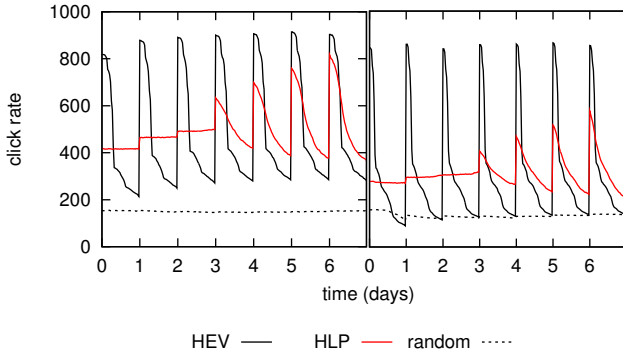


Fig. 5. The moving average of click rate for different policies under the 7 day static setting; the lifetime of campaigns is 5 days (20 million time steps) and their budgets are either 4000 (left) or 2000 (right).

that being very far-sighted may not be necessary.

Finally, we conducted experiments in the dynamic setting with partial information where the probabilities are not known in advance but estimated online. We employed ϵ -greedy exploration mechanism with different values of ϵ and maximum a posteriori estimation with Beta priors. The results show that HLP can perform better than HEV, however for both policies the chosen set of parameters influences the outcome (Fig. 6).

IV. CONCLUSION AND FUTURE WORK

In this paper, we have shown that optimizing advertisement display, handling finite budgets and finite lifetimes in a dynamic and non-stationary setting, is feasible within realistic computational time constraints. We have also given some insights in what can be gained by handling this constraint, depending on the properties of the advertisements to display. Our experimental results indicate that if there are few overlapping advertisements, or many advertisements with long lifetimes and good click rates, then we should be greedy. Between these two extreme solutions, one should consider the constraints associated to each advertisement campaign. In particular, the lifetime of campaigns have an important impact.

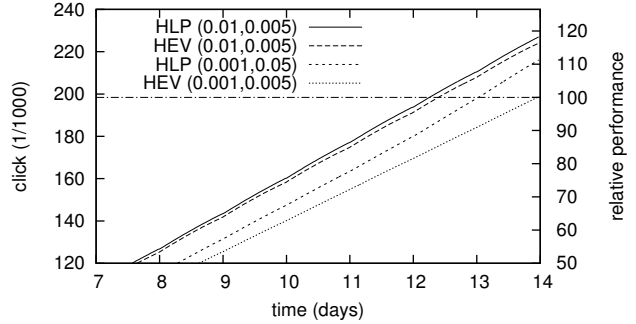


Fig. 6. The performance of HEV and HLP algorithms in the dynamic setting with partial information using ϵ -greedy exploration. The numbers in parenthesis denote the values of the parameter of the Beta prior and ϵ .

The oldest reference we were able to spot is Langheinrich et al. [3] who mixed a linear program with a simple estimation of CTR to select advertisements to display. No attention is paid to the exploration/exploitation trade-off and the problem of the estimation of the CTR is very crudely addressed. Then, Abe and Nakamura [1] introduce a multi-arm bandit approach to balance exploration with exploitation. Their work is based on display proportions, that is unlimited resources; they also deal with a static set of advertisements. They also consider multi-impression of advertisements on a single page [4]. Aiming at directly optimizing the advertisement selection, side information is used to improve the accuracy of prediction in several recent papers. However, all these works do not consider finite budget and finite lifetime constraints, as well as the continuous creation of new campaigns; they also do not consider the CTR estimation problem. See [2] Sec. III for detailed discussion of these and other related works.

This work calls for many further developments. A possibility is to solve the problem from the perspective of the advertiser, *i.e.* help them to set the value of a click, and adjust it optimally with respect to the number of visitors (equivalent to a local sensitivity analysis of the LP problem). A more difficult issue is that of handling multiple advertisements on the same page; in this case, the correlation between the advertisements becomes important. Finally, we are also willing to draw some theoretical results on how far from the optimal strategy we are. Dealing with finite resources, under finite time constraints, in a dynamic setting makes that kind of study quite difficult.

REFERENCES

- [1] N. Abe and A. Nakamura, "Learning to Optimally Schedule Internet Banner Advertisements," in *Proc. of the 16th International Conference on Machine Learning*, 1999, pp. 12–21.
- [2] S. Girgin, J. Mary, P. Preux, and O. Nicol, "Advertising campaigns management: Should we be greedy?" Sep. 2010, INRIA Research Report RR-7388.
- [3] M. Langheinrich, A. Nakamura, N. Abe, T. Kamba, and Y. Koseki, "Unintrusive customization techniques for web advertising," *Computer Networks*, vol. 31, Jan. 1999.
- [4] A. Nakamura and N. Abe, "Improvements to the linear programming based scheduling of web advertisements," *Electronic Commerce Research*, vol. 5, no. 1, pp. 75–98, Jan. 2005.