

# Native Advertisement Selection and Allocation in Social Media Post Feeds

Iordanis Koutsopoulos<sup>(✉)</sup> and Panagiotis Spentzouris

Department of Informatics, Athens University of Economics and Business,  
Athens, Greece  
[jordan@aub.gr](mailto:jordan@aub.gr)

**Abstract.** We study native advertisement selection and placement in social media post feeds. In the prevalent pay-per-click model, each ad click leads to certain amount of revenue for the platform. The probability of click for an ad depends on attributes that are either inherent to the ad (*e.g.*, ad quality) or related to user profile and activity or related to the post feed. While the first two types of attributes are also encountered in web-search advertising, the third one fundamentally differentiates native from web-search advertising, and it is the one we model and study in this paper. Evidence from online platforms suggests that the main attributes of the third type that affect ad clicks are the relevance of ads to preceding posts, and the distance between consecutively projected ads; *e.g.*, the fewer the intervening posts between ads, the smaller the click probability is, due to user saturation.

We model the events of ad clicks as Bernoulli random variables. We seek the ad selection and allocation policy that optimizes a metric which is a combination of (i) the platform expected revenue, and (ii) uncertainty in revenue, captured by the variance of provisionally consumed budget of selected ads. Uncertainty in revenue should be minimum, since this translates into reduced profit or wasted advertising opportunities for the platform. On the other hand, the expected revenue from ad clicking should be maximum. The constraint is that the expected revenue attained for each selected ad should not exceed its apriori set budget. We show that the optimization problem above reduces to an instance of a resource-constrained minimum-cost path problem on a weighted directed acyclic graph. Through numerical evaluation, we assess the impact of various parameters on the objective, and the way they shape the tradeoff between revenue and uncertainty.

**Keywords:** Native advertising · Advertisement allocation · Social media feeds · Mathematical modeling · Optimization · Shortest-path problem

## 1 Introduction

Internet advertising in its early form more than 15 years ago consisted in sponsored search advertising, whereby personalized targeted advertisements were dis-

played in certain order next to web-search results after a search query. The proliferation of social media, micro-blogging and social-networking platforms has created novel opportunities for advertising. Facebook, Twitter, Tumblr, Pinterest and other online platforms aim at user engagement by providing services or showing content of interest, and they leverage the user base for marketing campaigns run by advertisers who pay to have their advertisements displayed. Advertising is performed by inserting sponsored posts in certain positions in the post feed on the user screen as the user scrolls (Fig. 1). A *feed* or *timeline* is a set of posts displayed on a user's screen, such as news, posts, updates, photos or videos. A sponsored post can be an ad adhering to the pay-per-click or the pay-per-impression model, a sponsored news item, or a promoted item such as a video or image.

The term coined for this type of advertising is *native* or *in-stream* advertising, as the ad format is assimilated into that of other content shown and is least intrusive for the user. Native advertising is becoming a multi-billion business with projected spend at \$6.4 billion by 2017 only in USA. Selection, ranking and pricing of ads are realized with the Generalized Second-Price (GSP) auction as in web-search advertising. Advertisers bid an amount to pay per ad click, and ads are selected and displayed in prespecified positions in the post feed. Since the user scrolls on the screen and presumably views ads at a rate of one ad every some tens of seconds, a plausible scenario is that advertisers do not change their bids in between consecutive ad positions. Hence, the auction is run *once* before user scroll, and ads end up in their positions according to their rank as a result of the auction.



Fig. 1. Native advertisements in Facebook and Twitter online post feeds.

Click probability of an ad depends on three types of attributes: those that are inherent to the ad (*e.g.*, ad quality), those related to the user profile and activity, and those related to the post feed itself. The main inherent feature of an ad is its quality, which is reflected into ad design format, nature of the

advertised product, accompanying text, and quality of the landing page. On the other hand, the relevance of the ad to the user profile captures the similarity of an ad to user preferences, search activity, posted items, and so on. For example an ad about a restaurant seems more appealing to a user that usually places food-related posts than it is to a user that posts text about books.

The third type of attributes concern the placement of ads in the user post feed. Based on evidence from online platforms [1–3], in native advertising, the ad click probability may depend on: (i) Relevance (i.e. context similarity) of the ad to preceding posts. For example, an ad about a hotel may be more likely to be clicked if shown directly after a post discussing vacation than after one on politics. (ii) Distance between consecutively shown ads. It is plausible that, the fewer the intervening posts between two ads, the smaller the click probability for the latter ad is, due to user saturation or fatigue effects. (iii) Position of the ad in the stream. If an ad is shown earlier in a feed, it will be clicked with higher probability than if shown later, since the user may quit scrolling.

In web search, the ranked list of ads is created through the GSP auction and is displayed next to search results. The rank of an ad is determined by the product of bid and click probability. On the other hand, in native advertising, the notion of rank is less clear, since native ads are placed in between diverse user posts. Further, in web search, the expected revenue of an ad decreases with its rank in the list. In native advertising, this is not the case, since the ad click probability depends on the precise *placement* of an ad in the post feed. As explained above, ads that appear earlier in the feed are not necessarily more likely to be clicked at, unless they are relevant to preceding posts or they are projected sparsely enough. In addition, in web search, the click probability of an ad depends only on its own rank and not on other ads. In native advertising the situation is more complex. Placing an ad at a certain position in the feed affects the click probability of this ad but also the click probabilities of subsequent ads. If these are placed close to the first ad and close to each other in general, user saturation due to frequent ad projection may lead to reduced click probability.

While the first two types of attributes above are also encountered in web-search advertising, the third one fundamentally distinguishes native from web-search advertising, and it is the one we model and study in this paper.

### 1.1 Our Contribution

We study optimal native ad selection and placement in social media post feeds, and the way it impacts the platform revenue. A set of ads emerge out of a GSP auction. Each ad comes with its apriori budget. In the prevalent pay-per-click model, each ad click entails a given amount of revenue for the platform and a corresponding amount of reduction for ad budget. The budget of each ad is renewed after a certain time interval. The product of click probability and revenue per ad click is the expected revenue from the ad. In our model, the ad click events are represented by Bernoulli random variables, and the ad click probability depends on the *relevance* of the ad to the preceding post, and on the *distance* between consecutive ads in the feed.

We seek the ad selection and allocation policy that optimizes a metric which combines (i) platform expected revenue, and (ii) uncertainty in revenue, captured by the variance of provisionally consumed budget of selected ads. The constraint is that the expected revenue attained for each selected ad should not exceed its apriori set budget. Uncertainty in revenue should be minimum, since this translates into reduced profit or wasted advertising opportunities for the platform. An ad selection and allocation policy that would lead to few clicks and small expected consumed budget is not preferable, since the revenue of the platform is smaller that it could potentially be. On the other hand, an ad allocation policy that would involve too many ad clicks while the ad budget is exhausted is also not desirable, since it leads to wasted advertising opportunities for the platform that are provided for free. That is, ad clicks that correspond to ad budget beyond the apriori one do not incur additional revenue for the platform. Furthermore, the expected revenue from ad clicking should be maximum.

To the best of our knowledge, both the problem and the model are novel in the literature. For clarity purposes, we study the problem for the post feed of one user. The model is amenable to multiple users as well as to extensions that include the other two types of features *i.e.*, those inherent to the ad or related to user profile. The contributions of our work are as follows.

- We provide a model and mathematical formulation for the problem of minimizing a combined metric of (i) uncertainty in ad-click generated revenue, which is quantified as the variance of provisionally consumed budget, and (ii) total expected revenue. The constraint is that the expected consumed budget for each selected ad should be no more than its apriori budget.
- We showcase how the relative positioning of ads in the post feed alters the ad click probabilities and therefore the revenue and uncertainty about it, and we specify the way in which the joint positioning of ads needs to be engineered so as to achieve the optimization objective above.
- We show that the problem above is an instance of a resource-constrained minimum-cost path one on an appropriately defined directed acyclic graph. The solution path reflects the policy of selecting which ads to show in the feed (out of a given set of ads), and in which positions to place them.

Through numerical evaluation results, we verify the tradeoff between revenue and uncertainty. The work in [27] is the most relevant to ours. Compared to that work, we consider the relevance of ads to posts as a factor that shapes ad click probability. Further, rather than adhering to the cumulative effect of projected ads on click probability as in [27], we assume that click probability of an ad depends on the distance from the previously shown ad. This essentially translates to a dependence of click probability on the average ad projection rate *i.e.*, the average number of posts elapsed until an ad is shown, which is a more plausible scenario that captures user annoyance. The work [21] is also relevant, in the sense that the measure of “regret” could be seen as similar to variance. In that work, the emphasis is on controlling social-network diffusion through user targeting without considering post feed aspects, while our work considers ad placement in the post feed.

The paper is organized as follows. In Sects. 2 and 3 we present the model, problem formulation and solution. Numerical results are provided in Sect. 4, literature overview is provided in Sect. 5, and the paper is concluded in Sect. 6. In the sequel, we use the words “ad” and “advertisement” interchangeably.

## 2 Model

We consider a set  $\mathcal{T}$  of  $T$  posts displayed on a user screen in a social-media platform. Posts are displayed as a stream in a certain order e.g. most recently occurred, and the user scrolls through the posts. There is also a set  $\mathcal{A}$  of  $N$  native ads with  $N < T$ . Typically,  $N = \beta T$ ,  $0 < \beta < 1$ , with  $\beta$  in the range  $10^{-2}$  to  $10^{-1}$ . The set of ads  $\mathcal{A}$  is the outcome of a bidding process among competing advertisers, and the social media platform selects the ones to display and their positions. Non-selected ads may be placed in subsequent feeds.

We assume the pay-per-click payment model; each time the user clicks on a displayed ad  $a$ , the advertiser is committed to pay the platform an amount  $b_a$  which emerges from the auction. Each ad comes with an apriori budget  $B_a$ , and each time an ad is clicked by a user, its budget is reduced by  $b_a$ . The budget of each ad is renewed after a certain time interval, and we focus our attention in studying the ad allocation policy in such a time interval.

For each ad  $a \in \mathcal{A}$  and post  $t \in \mathcal{T}$ , let  $r_{at} \in [0, 1]$  be the *relevance* of ad  $a$  and post  $t$ . Relevance quantifies context similarity between the ad and the post. For example, an ad about a hotel is more relevant to a post on vacation than to one on politics. Relevance may be computed through cosine similarity [4, Chap. 9] or other metrics on vectors of words that are representative of the post and the ad. These may be defined e.g., with the Term-Frequency-Inverse Document Frequency (TF-IDF) metric from information retrieval [4, Chap. 1].

We consider two main determinants of probability of click for an ad: ad relevance to the preceding post (after which the ad is placed), and distance (in elapsed posts) from the previously displayed ad. We assume that we learn from historical data the probability  $p(r, d)$  that an ad is clicked if placed at distance  $d$  posts from the previous ad,  $d = 1, \dots, T$ , and if it is displayed after a post of relevance  $r$  to it. Logistic regression or other machine-learning tools can be used to train a model and learn  $p(r, d)$ . Recall that the ad click probability may depend on other attributes such as user profile or ad quality, or its position (early/late) on the stream. However, we choose not to include them here for clarity of presentation, and because we wish to focus on attributes that are peculiar to native advertising.

The nature of native advertising implies that, besides the style and layout of the ad, the content of the ad should be assimilated to that of other content shown on the platform e.g., the post feed. Hence, we consider the relevance of the ad to the preceding post as an attribute that affects the ad click probability. Further, the distance from the previously shown ad is essentially mapped to ad projection rate i.e., average number of posts elapsed until an ad is shown. Our rationale for selecting these attributes to map to ad click probability is spurred

by realistic marketing principles in online social networks and media. Platforms aim at high user experience while ensuring that ads obtain substantial attention. For instance, Facebook does not show too many ads and too frequent so as to prevent negative impact on user experience and engagement. It shows adequate number of ads so as to consume their budget and have advertisers satisfied.

Let us index post positions by  $t = 1, \dots, T$ , and ads by  $a = 1, \dots, N$ . Define binary variables  $x_{at}$ , for  $a = 1, \dots, N$ , and  $t = 1, \dots, T$ , with  $x_{at} = 1$ , if ad  $a$  is placed after post  $t$ , and  $x_{at} = 0$  otherwise. An *ad selection and allocation policy* is a  $NT \times 1$  *ad allocation vector*  $\mathbf{x} = (x_{at} : a \in \mathcal{A}, t \in \mathcal{T})$ . For ad  $a$ , let  $t_{\mathbf{x}}(a)$  be the post after which ad  $a$  is placed according to policy  $\mathbf{x}$ . For ad  $a$ , define the distance from the previous ad,  $d_a(\mathbf{x})$ , for allocation policy  $\mathbf{x}$ , as

$$d_a(\mathbf{x}) = \begin{cases} t_{\mathbf{x}}(a), & \text{if } t_{\mathbf{x}}(a) = \min_{a' \in \mathcal{A}} t_{\mathbf{x}}(a'), \\ t_{\mathbf{x}}(a) - \max_{\substack{a' \in \mathcal{A}: \\ t_{\mathbf{x}}(a') < t_{\mathbf{x}}(a)}} t_{\mathbf{x}}(a'), & \text{otherwise.} \end{cases} \quad (1)$$

If an ad is the first one placed in the feed, its distance is just the index of the preceding post (after which the ad is placed). Otherwise, the distance is the difference between the index of the preceding post of that ad, and that of the preceding post of the immediately previously placed ad. We define  $d_a(\mathbf{x}) = 0$  if ad  $a$  is not allocated in the feed, i.e. if  $x_{at} = 0$  for all  $t$ .

Given an ad allocation policy  $\mathbf{x}$ , each allocated ad  $a$  has a click probability  $p(r_{at}, d_a(\mathbf{x}))$ . For notational simplicity, and with a little abuse of notation, let us in the sequel use notation  $p_{at}(\mathbf{x})$  to denote  $p(r_{at}, d_a(\mathbf{x}))$ . The event of click of each allocated ad  $a$  when placed after post  $t$  according to allocation policy  $\mathbf{x}$  may be represented by a Bernoulli random variable  $X_{at}(\mathbf{x})$  having as parameter the probability  $p_{at}(\mathbf{x})$ . Thus,

$$X_{at}(\mathbf{x}) = \begin{cases} b_a, & \text{w.p. } p_{at}(\mathbf{x}), \\ 0, & \text{w.p. } 1 - p_{at}(\mathbf{x}), \end{cases} \quad (2)$$

with expectation  $\mathbb{E}[X_{at}(\mathbf{x})] = b_a p_{at}(\mathbf{x})$  and variance

$$\text{var}[X_{at}(\mathbf{x})] = b_a^2 p_{at}(\mathbf{x}) \cdot (1 - p_{at}(\mathbf{x})). \quad (3)$$

Denote by  $R(\mathbf{x})$  the random variable that shows platform revenue as function of the allocation policy  $\mathbf{x}$ , with

$$R(\mathbf{x}) = \sum_{a=1}^N \sum_{t=1}^T X_{at}(\mathbf{x}) x_{at}. \quad (4)$$

The total expected revenue for the platform for an ad allocation policy  $\mathbf{x}$  is

$$\mathbb{E}[R(\mathbf{x})] = \sum_{a=1}^N \sum_{t=1}^T b_a p_{at}(\mathbf{x}) x_{at}. \quad (5)$$

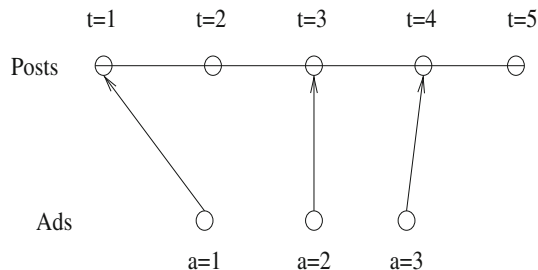
The underlying assumption in (5) is that the Bernoulli random variables that show click events of different ads in the feed are independent from each other. This is a plausible assumption, since the probability that an ad is clicked does not seem to depend on whether a previous ad was clicked or not. On the other hand, the model recognizes that the probability of ad click depends on *how many posts elapse* since the appearance of the most recent ad in the feed, and thus it captures in a sense the annoyance caused to the user; this is reflected through the dependence of click probability on distance. The variance of revenue is

$$\text{var}[R(\mathbf{x})] = \sum_{a=1}^N \sum_{t=1}^T \text{var}[X_{at}(\mathbf{x})]x_{at} = \sum_{a=1}^N \sum_{t=1}^T b_a^2 p_{at}(\mathbf{x})(1 - p_{at}(\mathbf{x}))x_{at}. \quad (6)$$

### 3 Problem Formulation and Solution

We are interested in the ad selection and allocation policy  $\mathbf{x}^*$  that minimizes a metric of the form,  $(\text{var}[R(\mathbf{x})] - \lambda \mathbb{E}[R(\mathbf{x})])$ , where  $\lambda \geq 0$  is a calibration parameter that determines the relative emphasis on total expected revenue and its variance. The rationale for selecting this metric is that it arises in the constrained optimization problems of maximizing expected revenue subject to keeping variance of revenue less than a given value, and that of minimizing variance of revenue subject to keeping expected revenue larger than a value. The optimal policy selects a number of ads to place in the feed and may place an ad several times in the feed. The number of selected ads may be small or large, depending on what is better for the objective above. The optimal policy may place an ad more times after posts that induce small click probability or fewer times after posts that induce larger click probability. An allocation policy  $\mathbf{x}$  changes the expected values and the variances of individual random variables that correspond to placed ads in the feed, and hence it affects  $\mathbb{E}[R(\mathbf{x})]$  and  $\text{var}[R(\mathbf{x})]$ . There exists a non-trivial coupling in the problem. The probability associated with a certain ad placed depends on the specific position (post) through the ad-post relevance, but it also depends on the post distance from the previously placed ad.

If  $\lambda = 0$  or it is very small, the objective is to minimize deviation of revenue from the expected one. As the total expected revenue has no or little weight in the objective, the platform is selective in choosing a subset of ads to allocate in the feed such that the uncertainty in revenue is minimized. A large uncertainty translates to potentially reduced revenue or to wasted advertising opportunities for the platform. Specifically, an ad allocation policy that would lead to an ad clicking profile with few clicks would result in a smaller expected consumed budget than the one that could potentially be consumed. On the other hand, an ad allocation policy with an ad clicking profile with too many clicks is also not desirable, since it would translate into wasted advertising opportunities and advertising service that the platform would provide for free. That is, ad clicks that correspond to ad budget beyond the apriori defined one do not incur revenue until the budget is renewed. If  $\lambda = \infty$  or it is very large, the aim is to



**Fig. 2.** Example of a feasible ad placement in  $T = 5$  posts. Ad 1 is placed after post 1, ad 2 is placed after post 3, and ad 3 is placed after post 4. The expected revenue for this assignment is  $b_1p(r_{11}, 1) + b_2p(r_{23}, 2) + b_3p(r_{34}, 1)$ . For example, for ad 2, click probability is  $p(r_{23}, 2)$  since the ad is placed after post 3 (hence the relevance factor  $r_{23}$ ) and at distance  $d_2 = 2$  posts from the previous ad, ad 1. Here, each ad is placed in exactly one position in the feed; in our case, an ad may be repeated.

maximize total expected profit. In that case, the platform does not place emphasis on revenue uncertainty, and hence it is more tolerant to wasted advertising opportunities.

The problem of ad selection and allocation so as to minimize the combined metric above is formulated as follows:

$$\min_{\mathbf{x}} (\text{var}[R(\mathbf{x})] - \lambda \mathbb{E}[R(\mathbf{x})]) = \min_{\mathbf{x}} \sum_{a=1}^N \sum_{t=1}^T \left( b_a^2 p_{at}(\mathbf{x})(1 - p_{at}(\mathbf{x})) - \lambda b_a p_{at}(\mathbf{x}) \right) x_{at} \quad (7)$$

subject to:

$$\sum_{t=1}^T b_a p_{at}(\mathbf{x}) x_{at} \leq B_a, \quad \forall \text{ ad } a \in \mathcal{A}, \quad (8)$$

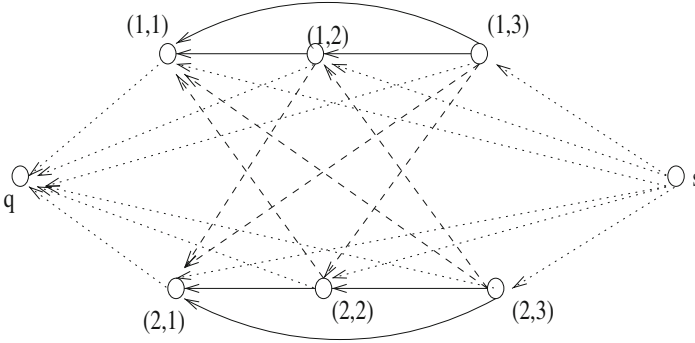
and

$$\sum_{a=1}^N x_{at} \leq 1, \quad \forall \text{ post } t \in \mathcal{T}, \quad (9)$$

with  $x_{at} \in \{0, 1\}$ . Constraint (8) says that for each ad, the expected revenue from policy  $\mathbf{x}$  should be no more than its apriori budget,  $B_a$ . Further, constraint (9) says that at most one ad is placed after a post. Figure 2 depicts an example allocation policy of ads to posts, where each ad is displayed exactly once.

Problem (7)–(9) is a non-standard one. If ad positions were known, distances  $d_a(\cdot)$  between ads would also be known, and the problem would be to select the ads to place in these positions. Even in that case, the problem would be a generalized assignment (GAP) one on the bipartite graph of nodes  $\mathcal{A} \cup \mathcal{T}$  with link weights  $b_a^2 p_{at}(1 - p_{at}) - \lambda b_a p_{at}$ , for each link connecting ad  $a$  and post  $t$ . The GAP problem is already NP-Hard [5]. The need to determine distances  $d_a(\cdot)$  further complicates the problem, since the decision on distance  $d_a(\cdot)$  of an ad  $a$





**Fig. 3.** A graph  $G$  corresponding to  $N = 2$  ads and  $T = 3$  posts that illustrates the mapping of the ad selection and placement problem to a shortest-path one. For example, if the shortest path is  $s \rightarrow (1, 3) \rightarrow (2, 1) \rightarrow q$ , this means that placing ad 1 after post 3 and ad 2 after post 1 minimizes the value of objective (7).

would affect the weights of links emanating from ad  $a$  but also the weights of links for the ad to be placed after ad  $a$ .

### 3.1 Graph Model and Solution

We construct the following directed graph  $G$ . For each pair of ad  $a \in \mathcal{A}$  and post  $t \in \mathcal{T}$ , we define a node  $(a, t)$ . The easiest way to visualize it is if we place nodes  $(a, t)$  in rows and columns; for each ad  $a$ , nodes  $(a, t)$ ,  $t = 1, \dots, T$  are in one row, and nodes corresponding to different ads are in different rows. Node  $(a, t)$  represents the tentative placement of ad  $a$  after post  $t$ . There also exist two other nodes  $s$  and  $q$ .

Next, we add links as follows. We add a link from each node  $(a, t)$  to nodes  $(a', t')$  with  $t' < t$ . That is, for each ad  $a$ , we add links between nodes  $(a, t)$  and  $(a, t')$  for  $t' < t$ . We also add links between  $(a, t)$  and  $(a', t')$ , for  $a' \neq a$  and  $t' < t$ . The weight of each link that points from node  $(a, t)$  to  $(a', t')$  is

$$w_{(a,t),(a',t')} = b_a^2 p(r_{at}, t - t') [1 - p(r_{at}, t - t')] - \lambda b_a p(r_{at}, t - t'). \quad (10)$$

We also add a link from node  $s$  to each node  $(a, t)$  in the graph with weight 0 and a link from each node  $(a, t)$  to  $q$  with weight

$$w_{(a,t),q} = b_a^2 p(r_{at}, t) [1 - p(r_{at}, t)] - \lambda b_a p(r_{at}, t). \quad (11)$$

For each node  $(a, t)$ , let  $\mathcal{O}_{(a,t)}$  be the set of outgoing links of node  $(a, t)$ , i.e. the set of links that originate from  $(a, t)$ . The resulting graph is a weighted directed acyclic graph (DAG).

**Main observation.** A path from  $s$  to  $q$  in the weighted graph  $G$  corresponds to an ad selection and placement policy. A minimum-cost path from  $s$  to  $q$  corresponds to a policy that leads to minimum value in the objective (7). Nodes

$(a, t)$  that are part of the minimum-cost path correspond to ads  $a$  that are assigned in positions  $t$ . An example graph  $G$  for  $N = 2$  ads and  $T = 3$  posts is shown in Fig. 3.

First, consider the problem with the optimization objective (7) subject only to constraint (9), while the ad budget constraint (8) is relaxed. From the discussion above, we deduce that finding a policy that minimizes the objective (7) subject to constraint (9) is equivalent to finding a minimum-cost path from  $s$  to  $q$  in the graph above. The minimum-cost path from  $s$  to  $q$  can be found by running the Bellman-Ford (BF) algorithm, which also applies to graphs with negative weights, as long as there are no negative cycles. In our case, the graph is a DAG with possibly negative link weights, but with no cycles. In fact, a variant of the BF algorithm can find the minimum-cost path for a DAG in  $\Theta(|V| + |E|)$  time, where  $|V|$  is the number of nodes and  $|E|$  is the number of links of the graph [6, Sect. 24.2]. In  $G$ , there exist  $(NT + 2)$  nodes and  $O(NT^2)$  number of links, thus the algorithm runs in  $O(NT^2)$  time.

Now, consider our problem with the optimization objective (7) subject to constraints (8) and (9). In our formulation, the expected consumed budget for each ad should not exceed  $B_a$ . We associate each ad  $a$  with a “resource type”  $a$ . For link  $e \in \mathcal{O}_{(a,t)}$  from node  $(a, t)$  to node  $(a', t')$  let ad  $a$  have budget consumption  $d_a^e = b_a p(r_{at}, t - t')$  while other ads  $a' \neq a$  have consumption  $d_{a'}^e = 0$  for that link. Furthermore, link  $e$  from node  $(a, t)$  to node  $q$  has consumption level  $d_a^e = b_a p(r_{at}, t)$  for ad  $a$ , and 0 for all other ads. Links from node  $s$  to nodes in the graph have consumption level equal to 0 for all ads. The total consumption level for ad  $a$  in a path  $p$  is  $d_a(p) = \sum_{\ell \in p} d_a^\ell$ .

A resource-constrained path  $p$  (where “constrained” refers to the total consumed budget of a resource type, *i.e.*, ad in it) from  $s$  to  $q$  is *feasible*, if and only if  $d_a(p) \leq B_a$  for all ads  $a \in \mathcal{A}$ , *i.e.* if it comprises links such that the total consumed budget for each ad included in the path is more than  $B_a$ . For ads that are not included in the path, the inequality trivially holds since these ads do not consume budget. The problem in this case is equivalent to a resource-constrained shortest-path one, which is NP-Hard [5]. There exist several heuristics proposed in the literature for solving the problem, see *e.g.*, [7, 8].

## 4 Numerical Evaluation

### 4.1 Setup and Data

We approach user ad click behavior as an instance of the two-class probabilistic classification problem, where the two classes  $C_0$  and  $C_1$  correspond to the alternatives of not clicking and clicking an ad respectively. The way the user weighs the attributes associated with an ad  $a$  (namely the relevance  $r_a$  to the post, and the distance  $d_a$  from the previous ad) so as to reach a decision is modeled through a logistic regression model. Given a vector of values  $\mathbf{x}$  for the two ad attributes, the ad is clicked with probability

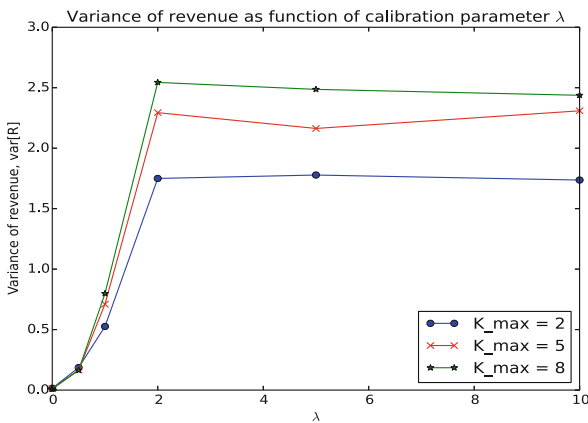
$$\Pr(C_1|\mathbf{x}) = \frac{1}{1 + e^{-\mathbf{w} \cdot \mathbf{x}}} = \sigma(\mathbf{w} \cdot \mathbf{x}), \quad (12)$$

where  $\sigma(y) = (1 + e^{-y})^{-1}$  denotes the logistic sigmoid function, while  $\mathbf{w} \cdot \mathbf{x}$  denotes vector dot product, and  $\mathbf{w}$  is the vector of attribute weights. These weights are learned from historical data and capture the significance that the user places on the two different attributes and their values in reaching a decision. Similarly, ad  $a$  is not clicked with probability  $\Pr(C_0|\mathbf{x}) = 1 - \Pr(C_1|\mathbf{x})$ . An important property of logistic regression is that the objective for learning weights  $\mathbf{w}$  is convex, so there are no local optima involved.

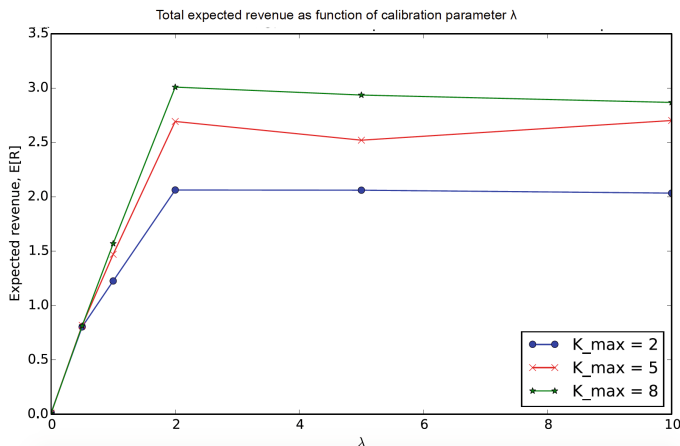
Real datasets for native advertising are scarce to find, and studies on native ads are either purely theoretical with no data experiments, *e.g.*, [23, 27], or they use company proprietary data, *e.g.*, [24, 25] [26, Chap. 7]. Hence in this work, we employ synthetic datasets to justify our claims and test our model. The training dataset consists of triads of the form  $(r_a, d_a, c_a)$ , where each triad represents an ad  $a$ . In each triad,  $r_a \in [0, 1]$  is the relevance to the post after which the ad was placed, while  $d_a \in [0, 1]$  is the distance from the previous ad, normalized with a defined maximum possible distance between two ads, and  $c_a \in \{0, 1\}$  denotes whether the ad  $a$  was clicked or not. To set the value of  $c_a$  to 0 or 1, we calculate metric  $0.5 \times r_a + 0.5 \times d_a$ , and if this is greater than a configurable threshold, which we take here to be equal to 0.75, then we set  $c_a = 1$ , else we set  $c_a = 0$ .

The classical loss function minimization approach with regularization was used to train our algorithm [9]. For training, we generate 50 triads and for testing we generated 1,000  $(r, d)$ -pairs. Given the trained logistic-regression model, we estimate the ad click probabilities through the sigmoid function. We create a pool of 50 ads and in each experiment, we select  $T = 20$  posts and we draw from the pool a certain maximum number of ads,  $K_{\max}$  to include in the feed. For simplicity we take the bid  $b_a = 1$  and the apriori budget  $B_a = 5$  for each ad.

The resource-constrained minimum-cost problem (7)–(9) was solved with the Lagrangian relaxation heuristic from [7], and the parameters were selected so that problem feasibility was not an issue. Based on [7], we can show that the



**Fig. 4.** Variance of revenue as a function of  $\lambda$  for different values of ad pool size,  $K_{\max} = 2, 5, 8$ .



**Fig. 5.** Expected revenue as function of  $\lambda$  for different ad-post relevance.

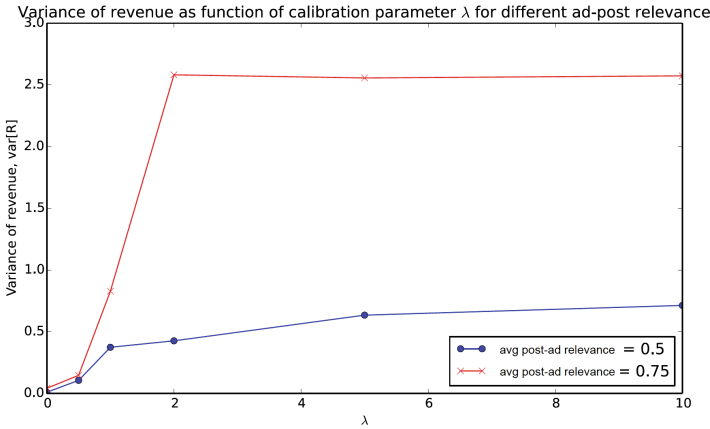
algorithm runs in  $O(N^2 T^4 \log^4(NT^2))$  time. Note that the algorithm does not need to solve the problem in an online fashion, but rather it pre-computes the position of ads in a post feed, hence the algorithm requirements in execution time (and thus, complexity) are not so stringent. Although the scale of the problem will be larger in practice, our conjecture is that the trends will remain the same, as only the associated parameters of the optimization problem will change.

## 4.2 Numerical Results

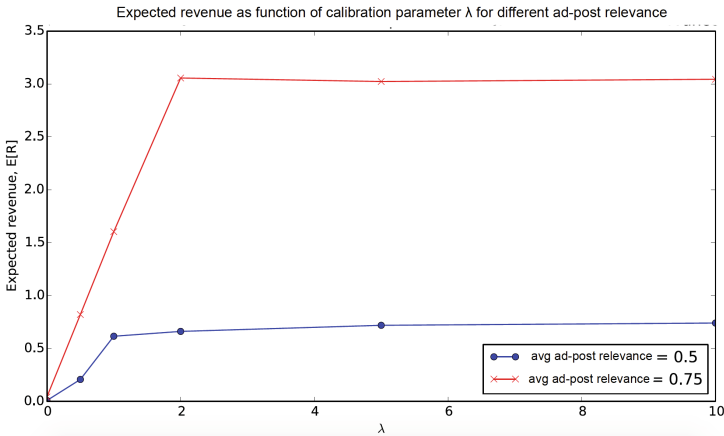
In the first set of numerical experiments, we assess the impact of calibrating parameter  $\lambda$ . In Figs. 4 and 5 we depict  $\text{var}[R]$  and  $\mathbb{E}[R]$  respectively as a function of  $\lambda$  for a maximum number of ads to be placed in the feed  $K_{\max} = 2, 5$  and 8. For each value of  $\lambda$ , we solve the resource-constrained shortest-path problem, and for the solution path we measure  $\text{var}[R]$  and  $\mathbb{E}[R]$  by summing the corresponding link costs over path links. Each value in the plot corresponds to the average value over ten experiments, where for each experiment the pool of ads and the set of  $T = 20$  posts are varied. Both the variance of revenue and the expected revenue are seen to increase as  $\lambda$  increases up to a certain value *i.e.*,  $\lambda = 2$ , while for values  $\lambda > 2$ , the respective values of variance and expected revenue are almost stabilized, or they change slightly. This fact demonstrates the tradeoff that, if the platform wishes to increase revenue, it would have to tolerate higher variance, *i.e.*, higher uncertainty in revenue. As  $\lambda$  increases, link costs decrease, and therefore more ads tend to be placed in principle in the feed. This results both in higher  $\text{var}[R]$  and  $\mathbb{E}[R]$ .

The second observation from Figs. 4 and 5 is that as the value of  $K_{\max}$  increases, both the expected revenue and the revenue variance increase, albeit the difference in the increase of these metrics decreases, as  $K_{\max}$  increases.

In a second set of experiments, in Figs. 6 and 7, we plot  $\text{var}[R]$  and  $\mathbb{E}[R]$  respectively as function of  $\lambda$  for different values of average ad-post relevance, which were produced by changing the parameters of a uniform distribution with which we generated different ads. The value of  $K_{\max}$  was 5. We observe that a higher ad-post relevance results in higher  $\text{var}[R]$  and  $\mathbb{E}[R]$  as expected, because of the raise in click probability. A moderate change in relevance seems capable of making a difference both in expected revenue and revenue variance. The same behavior as that in Figs. 4 and 5 is observed with respect to  $\lambda$ .



**Fig. 6.** Variance of revenue as function of  $\lambda$  for different ad-post relevance.



**Fig. 7.** Expected revenue as function of  $\lambda$  for different ad-post relevance.

## 5 Related Work

In sponsored-search auctions that are used in web search advertising, ads are ranked based on expected profit, which is the product of bid and click-through-ratio (CTR), *i.e.*, the probability that the ad will be clicked. When the user clicks on an ad at the  $k$ -th position with bid  $b_k$  and  $\text{CTR}_k$ , the advertiser pays  $b_{k+1} \times \text{CTR}_{k+1}$  according to the Generalized Second-Price (GSP) auction [10, 11]. In [12], a variation of GSP is presented, where each ad bid undergoes a fine, equal to a metric of negative impact of the ad on user experience. Advertisers that are charged with large fines are less willing to enter the competition, and thus the platform becomes more attractive to other advertisers. Under certain conditions, the winners' gains in this new less crowded setting may supercede payments due to fines, and thus overall winners are benefited.

A recent thread concerns ad allocation through stochastic optimization. In [13], the optimal-auction framework is used for the single-slot revenue maximization. The optimal policy is to allocate the slot to users in decreasing order of  $q_i \nu_i$  where  $q_i$ ,  $\nu_i$  are the selling probability and valuation of user  $i$ . In [14], the authors use Lyapunov optimization for the problem of maximizing long-term average revenue for a web-search service provider by dynamically allocating ads to webpage slots in the presence of dynamic keyword query arrivals, subject to a long-term average budget constraint. The work [15] studies the problem of allocating budget-constrained advertisers in each keyword auction round so as to maximize the likelihood of ad click or to reduce advertiser cost per click. Dynamic actions under limited budget over the entire horizon are studied in [16] through the lens of multi-armed bandit theory.

Another thread that relates to advertising is *social influence*, which involves positive externalities, namely that the benefit from influencing a user comes also from indirect influence of that user onto others. The seminal paper [17] formulates the problem of influence maximization as one of selecting a subset of users (seeds) to advertise to so that the cascading effect in the graph reaches the maximum number of users. They show that the problem is NP-Hard and propose a greedy algorithm with constant-factor approximation guarantees based on sub-modularity properties of the set function of anticipated influence. Various extensions have been considered, *e.g.*, on optimal marketing strategies that include pricing and the sequence of offers to social-network users that may or may not respond strategically [18, 19], and on user targeting for global opinion maximization under a game-theoretic model for opinion formation [20]. The work [21] studies ad allocation to social-network users under a certain diffusion model and topic-based influence. The host aims on leveraging virality to improve advertising efficacy, while avoiding giving away free service due to uncontrolled virality. The problem is to allocate ads to minimize regret, defined as the absolute difference between the expected revenue and the budget of each ad. Social diffusion in ads is also considered in [22], where the joint problem of targeting ad impressions to users and of scheduling them in time is studied with the aim to maximize expected number of clicks. The work [23] considers the problem of selecting a set of ads and the number of times to display each ad so as to minimize the error in estimating the true CTR of ads.

Native ads have spurred interest of the research community in the last two years or so, aiming at improved user experience, see *e.g.*, [24, 25] [26, Chap. 7]. The works [24, 25] aim to predict ad quality by focusing respectively on post-click user behavior on the ad landing page, and on user feedback about offensive ads. In a related work [27], the problem of ad placement in a stream is addressed. The model involves a probability that the user will reach to the ad, which is decreasing function of the number of ads shown previously. Given a set of ads, a reward and a set of candidate positions to place each ad, the objective is to find an ad placement that maximizes total reward. An approximation algorithm is proposed, albeit the computational complexity of the problem is not characterized. The authors also use the optimal-auction framework to design a mechanism that is truthful and approximately optimal in terms of revenue.

## 6 Conclusion

We study native advertisement selection and placement in the post feed of a user, and we optimize a metric that combines expected platform revenue and revenue uncertainty. Ad click probabilities are derived with a machine-learning model that maps the key attributes of post-ad relevance and distance from the previous ad to click probability. Next, these ad click probabilities are engineered and adapted through ad selection and placement in the feed to achieve the objective. We showed that the problem becomes a resource-constrained minimum-cost path one.

To the best of our knowledge, both the problem and the model are novel. The model is amenable to various extensions such as the one for multiple user feeds, thus encapsulating the relevance of an ad to personalized user profile as a factor that shapes click probability. In that case, personalized user models for ad click behavior would be needed. Other attributes that shape ad click probability could be included in the model as well, such as ad quality. We are currently in the process of designing a larger-scale real-life experiment with training and test data that come through some hundreds of real users on tentative Facebook feeds and ads presented to them through a mobile app.

## References

1. <https://www.bigfin.com/blog/facebook-to-provide-relevant-news-feed-ads/>
2. <http://www.buzzfeed.com/mattlynley/this-is-how-an-ad-gets-placed-in-your-facebook-news-feed#.gq48arOBY>
3. <http://marketingland.com/facebook-raises-limits-daily-frequency-news-feed-ads-96517>
4. Leskovec, J., Rajaraman, A., Ullman, J.: Mining Massive Datasets. Cambridge University Press, New York (2014)
5. Garey, M.R., Johnson, D.S.: Computers and Intractability: A guide to the Theory of NP-Completeness. Freeman, New York (1979)
6. Cormen, T., Leiserson, C.E., Rivest, R.L., Stein, C.: Introduction to Algorithms, 3rd edn. MIT Press, Cambridge (2009)

7. Juttner, A., Szviatevski, B., Mecs, I., Rajko, Z.: Lagrange relaxation based method for the QoS routing problem. In: *Proceedings of IEEE INFOCOM* (2001)
8. Boland, N., Dethridge, J., Dumitrescu, I.: Accelerated label setting algorithms for the elementary resource constrained shortest path problem. *Oper. Res. Lett.* **34**(1), 58–68 (2006)
9. Flach, P.: *Machine Learning: The Art and Science of Algorithms that Make Sense Out of Data*. MIT Press, Cambridge (2012)
10. Varian, H.: Online ad auctions. *Am. Econ. Rev.* **99**(2), 430–434 (2009)
11. Narahari, Y., Garg, D., Narayanam, R., Prakash, H.: *Game Theoretic Problems in Network Economics and Mechanism Design Solutions*. Advanced Information and Knowledge Processing. Springer, London (2009)
12. Stourm, V., Bax, E.: Pigovian taxes can increase platform competitiveness: the case of online display advertising. arXiv preprint [arXiv:1411.0710](https://arxiv.org/abs/1411.0710)
13. Menache, I., Ozdaglar, A., Srikant, R., Acemoglu, D.: Dynamic online-advertising auctions as stochastic scheduling. In: *Proceedings of NetEcon* (2009)
14. Tan, B., Srikant, R.: Online advertising, optimization and stochastic networks. *IEEE Trans. Autom. Control* **57**(11), 2854–2868 (2012)
15. Karande, C., Mehta, A., Srikant, R.: Optimizing budget constrained spend in search advertising. In: *Proceedings of ACM Web Search and Data Mining (WSDM) Conference* (2013)
16. Wu, H., Srikant, R., Liu, X., Jiang, C.: Algorithms with logarithmic or sublinear regret for constrained contextual bandits. In: *Proceedings of the Neural Information Processing Systems Conference* (2015)
17. Kempe, D., Kleinberg, J.M., Tardos, E.: Maximizing the spread of influence through a social network. In: *Proceedings of ACM Knowledge Discovery Data Mining (KDD)* (2003)
18. Hartline, J.D., Mirrokni, V.S., Sundararajan, M.: Optimal marketing strategies over social network. In: *Proceedings of ACM World Wide Web (WWW) Conference* (2008)
19. Candogan, O., Bimpikis, K., Ozdaglar, A.: Optimal pricing in the presence of local network effect. In: *Proceedings of Conference on Web and Internet Economics (WINE)* (2010)
20. Gionis, A., Terzi, E., Tsaparas, P.: Opinion maximization in social networks. In: *Proceedings of SIAM International Conference on Data Mining (SDM)* (2013)
21. Aslay, C., Lu, W., Bonchi, F., Goyal, A., Lakshman, L.V.S.: Viral marketing meets social advertising: ad allocation with minimum regret. In: *Proceedings of Very Large DataBases (VLDB) Conference* (2015)
22. Abbassi, Z., Bhaskara, A., Misra, V.: Optimizing display advertising in online social networks. In: *Proceedings of WWW Conference* (2015)
23. Gopalakrishnan, R., Bax, E., Chitrapura, K.P., Garg, S.: Portfolio allocation for sellers in online advertising. arXiv preprint [arXiv:1506.02020](https://arxiv.org/abs/1506.02020)
24. Lalmas, M., Lehmann, J., Shaked, G., Silvestri, F., Tolomei, G.: Promoting positive post-click experience for In-stream Yahoo Gemini users. In: *Proceedings of ACM Knowledge Discovery Data Mining (KDD)* (2015)
25. Zhou, K., Redi, M., Haines, A., Lalmas, M.: Predicting pre-click quality for native advertisements. In: *Proceedings of ACM World Wide Web (WWW) Conference* (2016, to appear)
26. Lehmann, J.: *From site to inter-site user engagement: fundamentals and applications*. Ph.D. thesis, Universitat Pompeu Fabra (2014)
27. Jeong, S., Mahdian, M., Vassilvitskii, S.: Advertising in a stream. In: *Proceedings of ACM WWW Conference* (2014)