Online Advertising with Periodic Budgets

Introduction. Due to the tremendous growth in online personalized data and advertising channels, advertisers have gained the ability to create advertisement campaigns that target advertisements to specific customer segments through a variety of advertising channels. Each combination of advertisement, customer segment, and advertising channel forms a different target. For example, in search advertising, a target can be to advertise a website link to people searching for related keywords. In general, publishers of online advertisements have a limited supply of advertising slots when compared to the large demand for advertisements. For this reason, many publishers run real-time auctions to determine which advertiser gets to show their ad. The market associated with these online advertising auctions is large and growing rapidly. According to a 2017 report by PricewaterhouseCoopers, just in the United States, online advertising revenues were on the order of $40.1 billion in the first half of 2017 and increased by 22.6% when compared to the first half of 2016. Thus, it is important for the publishers to create the right auction platform and for the advertiser to place the right bids.

Previous studies have focused on the auction and ad allocation problem of the publishers. In contrast, we focus on the bidding problem that advertisers face in practice. In each time period (e.g. hourly, daily), the advertiser needs to decide which targets to bid on to maximize expected advertising revenue. However, in each period, the advertiser has a limited advertising budget to spend, and outspending the budget leads to lost revenues. This is caused by the advertiser being unable to participate in any further ad auctions once a period’s advertising costs exceed the budget. The advertising revenue of a target comes from conversions, i.e., revenues are gained when customers click on ads and generate a purchase. At the same time, the advertising cost of a target is dictated by the auctions run during the period, i.e., costs are incurred when auctions are won and result in customer clicks. The advertising revenues and costs are random to the advertiser, as they depend on the conversion behavior of customers and the bidding of competing advertisers. Even the expected revenue and cost of a target is unknown to the advertiser, as it cannot be estimated from the limited data available for a large number (millions) of targets. We model the learning of the revenue and cost of targets (exploration) and bidding on the most valuable targets (exploitation) as a multi-armed bandit (MAB) problem, and propose a new optimistic-robust learning (ORL) algorithm.
Literature Review. In the original multi-armed bandit (MAB) model, the advertiser bids on targets, each delivering a random revenue from unknown distributions, in order to maximize revenues. Previous literature on MAB models with budgets has focused on the budget-limited MAB model (Ding et al., 2013; Tran-Thanh et al., 2014; Xia et al., 2015). In the budget-limited MAB, the advertiser sequentially selects a single target that maximizes the expected advertising revenue over the time that it takes to deplete the advertising campaign budget. In our MAB with periodic budgets, the advertiser periodically selects a set of targets to maximize the expected advertising revenue of a period subject to the period’s advertising cost not exceeding the period’s budget. Thus, there are two important differences between our MAB model with periodic budgets and the budget-limited MAB model. First, under periodic budgets, the advertiser selects a set of targets to bid on during a longer time period, while a budget-limited advertiser selects targets one by one as they come in. Secondly, the advertiser with periodic budgets needs to satisfy a budget constraint in every period, instead of the single campaign budget that is slowly depleted in the budget-limited MAB. In practice, advertisers propose an advertising campaign that runs over a fixed time horizon with a fixed campaign budget, and then split this campaign into time periods each with a fraction of the campaign budget (Pani et al., 2018). Their goal becomes to select, at the beginning of a period, a set of targets to bid on if the publisher offers them during the period. Our MAB model with periodic budgets is able to mimic this practice, whereas the budget-limited MAB could only capture the case whether or not to bid on the current target offered by the publisher.

Contributions. Our main contributions are the following:

- We model the online advertising problem as a multi-armed bandit (MAB) problem with periodic budgets. In comparison to most online advertising literature from the perspective of online advertisement publishers, we analyze how an advertiser should bid on online auctions. In each time period, the advertiser’s goal is to bid on a set of targets that maximizes advertising revenues under the constraint that advertising costs are not allowed to exceed the period’s advertising budget. In the MAB context, this means that we wish to pull a set of arms that maximizes revenue while keeping the cost of pulling those arms below the period’s budget. The reason for periodic budgets is that, in reality,
advertisers split their advertising campaigns into time periods (often days) with separate budgets. These periodic budgets and the pulling of multiple arms per period create a considerably different and more complex model than the typical MAB models.

- We propose an optimistic-robust learning (ORL) algorithm. We show that the MAB with periodic budgets is difficult to solve, even for an oracle that knows the expected revenue and cost of each arm. Therefore, we use techniques from robust optimization to create an oracle policy that approximates the optimal policy. Robust optimization methods help to avoid the case where the cost of pulling arms exceeds the budget. As the expected revenue and cost are unknown, the oracle policy is not implementable, but we devise the ORL algorithm based on both UCB algorithms and robust optimization. This algorithm is able to explore the possible arms (learn the expected revenue and cost), and exploit the valuable arms (generate revenue while staying within budget).

- We prove that the ORL algorithm incurs a bounded expected regret compared to an oracle policy. We show how the parameters of the ORL algorithm can be tuned to achieve a bound on expected regret. Under certain conditions, we observe that the regret is tightly bounded by a logarithmic function of the number of time periods. Finally, the bound is polynomial in the number of arms. This is in contrast with the MAB model whose arms represent a collection of the original set of arms (targets), whose expected regret bound is exponential in the number of arms (targets).

- We benchmark the performance of the ORL algorithm through computational experiments. To benchmark the ORL algorithm, we develop a passive learning approach that does not actively explore or account for budget violations, but is able to update its estimate of revenue and cost over time. Through simulations, we observe that using the parameters proposed by our expected regret bound leads to a significant reduction in regret by already more than 20% over short time horizons.

- We show that the ORL algorithm increases revenues on real-world online advertising data. Working together with a large online advertising intermediary, we test the algorithm on instances based on client advertising portfolios and observe an improvement of 5-8% over current algorithms that learn passively.