Optimization Levers for Promotions Personalization Under Limited Budget

DMITRI GOLDENBERG, Booking.com, Tel Aviv, Israel
JAVIER ALBERT, Booking.com, Tel Aviv, Israel
GUY TSYPE, Booking.com, Tel Aviv, Israel

Modern e-commerce platforms make use of promotional offers, such as discounts and rewards, to encourage customers to complete purchases. As expected, revenue is also affected by promotions, and a dedicated budget usually limits monetary losses. In order to allocate promotions efficiently within budget constraints, a marketer can use causal machine learning based personalization along with constrained optimization tools. In this paper we study four decision levers of promotional campaigns, allowing optimal and personalized offers allocation within budget constraints. We demonstrate the optimization problems in real-life promotional campaigns and formulate them as variations of the Knapsack problem, allowing us to introduce efficient applied solutions. We demonstrate that such solutions have a significant impact on promotional campaigns at Booking.com - a world leading online travel platform.

CCS Concepts: • **Information systems** → **Personalization**; **Recommender systems**; • **Applied computing** → *Multi-criterion optimization and decision-making*.

Additional Key Words and Phrases: Uplift Modeling, Causal Inference, Promotions Personalization, Online Optimization, Knapsack Problem

1 INTRODUCTION

E-commerce applications use promotional campaigns to offer more value to customers and help grow their customer base [10]. Online travel platforms frequently offer discount incentives on a variety of products (Figure 1 demonstrates travel promotions examples). Although promotional campaigns increase the likelihood of purchase completion, it can also result in incremental monetary loss when examining net revenue. [13, 21, 26]. This incremental revenue loss is usually limited by a dedicated budget. The promotional campaign can remain sustainable as long as overall net revenue loss is within this budget [7].

Customer purchase prediction and expected net revenue losses may vary from customer to customer, even when they are shown the same promotion. A number of machine learning techniques have been developed to estimate the *Conditional Average Treatment Effect* (CATE) of a specific treatment on an individual [2], and the field is commonly known as *Uplift Modeling*. Uplift modeling is gaining popularity among web and e-commerce businesses, such as Facebook and Amazon, [15] Criteo [5], Uber [27], and Booking.com [24]. Uplift modeling is also used in cost-aware decision making on synthetic [20] and real [6, 17] promotional marketing use cases.

Marketers have more options than just allocating personalized promotions [8]. Along with choosing which promotion to assign to which customer, the marketer can also change the amount of the discount, the reward conditions of the promotion, and even the timing and persistence of the promotion itself. By dynamically changing these parameters, one can create a complex optimization system that maximizes incremental sales within budget constraints. In this paper, we present the application of four decision levers used:

(1) Campaign eligibility modeling,
(2) Multiple-choice promotion recommendation,
(3) Reward conditions optimization,

(4) Campaign persistence tuning.

Relying on our previous work, we aim to demonstrate how to use the different levers, showing real-life examples of our approach. The paper is outlined as follows: Section 2 describes campaign eligibility modeling; Section 3 focuses on multiple-choice promotion recommendation; Section 4 showcases reward conditions optimization; Section 5 presents campaign persistence optimization and the final section concludes the paper.

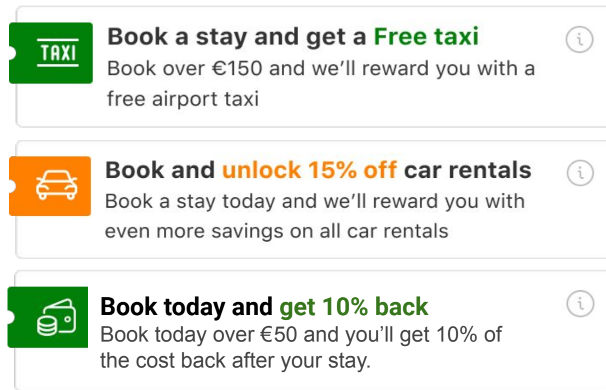


Fig. 1. Examples of travel promotions.

2 CAMPAIGN ELIGIBILITY MODELING

In this section, we focus on maximizing the overall number of customers completing a purchase by deciding whether to offer a specific promotion, while considering global budget limitations. Note the random variable $Y \in \{0, 1\}$ representing the completion of a purchase. Additionally, the random variable R, we consider to represent the net monetary revenue associated with the purchase (including the cost associated with the promotion). We briefly cover the optimization method and its results, which are fully described in a previous work [7].

Using the *Potential Outcomes framework* [11], we express the causal effects of the promotion: for each customer $i \in U$, we denote $Y_i(1)$ as a potential purchase if i received the promotion ($T = 1$), and $R_i(1)$ as the potential revenue (with deducted cost of the promotion). $Y_i(0)$ and $R_i(0)$ represent the potential outcomes if the promotion is not offered. Given customer's pre-promotion covariates x (features such as origin country or search dates) the conditional average treatment effect (CATE) of these variables are:

$$CATE_Y(x) = \mathbf{E}(Y_i(1) - Y_i(0)|X_i = x)$$
$$CATE_R(x) = \mathbf{E}(R_i(1) - R_i(0)|X_i = x)$$

The net revenue loss is defined as $\mathcal{L}_i = -R_i$ (we use positive loss values as a measure for weight in the knapsack problem defined in Equation 1). We aim to model the decision whether to offer a promotion to a customer with given covariates x, in order to maximize the total incremental purchases subject to budget constraints.

Using previously collected data on randomly assigned promotions, we evaluate the optimal re-assignment of the promotions given to customers set U, with the following optimization formulation:

$$
\begin{aligned}
\text{Maximize } & \sum_{i \in U} z_i \cdot CATE_Y(x_i) \\
\text{subject to: } & \sum_{i \in U} z_i \cdot CATE_{\mathcal{L}}(x_i) \leq 0
\end{aligned}
\tag{1}
$$

Here $z_i \in \{0, 1\}$ is the assignment variable indicating whether customer i is offered the promotion or not. The target function is maximizing the total incremental purchases, and the constraint limits the incremental budget to zero. This formulation resembles the *Binary Knapsack Problem* with possible negative weights and utility values. In our setting, $CATE_Y \leq 0 \implies CATE_{\mathcal{L}} \geq 0$, due to the fact that the incremental revenue is generated only by incremental purchases. We apply a transformation on negative values [25] converting our problem into a standard binary Knapsack Problem for customers with $CATE_Y > 0$ and $CATE_{\mathcal{L}} > 0$. The new budget constraint constant is achieved by $\sum_j CATE^j_{\mathcal{L}}$ for all customers where $CATE^j_Y > 0$ and $CATE^j_{\mathcal{L}} \leq 0$.

We can approximate our problem to the *Fractional Knapsack Problem*, and use the greedy algorithm with sorting customers U by descending $utility_i/weight_i$. Here, this sorting criteria is calculated by $CATE^i_Y/CATE^i_{\mathcal{L}}$. Relying on a decision threshold θ, the assignment strategy is to provide a promotion to all customers with a ratio above the threshold. The threshold θ is calibrated on historical data such that the total budget of the customers above the threshold is not exceeding the constraint. In an online setup, the threshold θ can be further tuned according to new data behaviors and target shifts.

The quantity $CATE^i_Y/CATE^i_{\mathcal{L}}$ is estimated by modeling $CATE^i_Y$ and $CATE^i_{\mathcal{L}}$ separately. In [7] we propose an extended method, called *Retrospective Estimation* to learn the fractional quantity directly, relying solely on positive examples' data. We demonstrate that the suggested method is more effective at solving the constrained optimization problem than other estimators. Similarly, in [6], the authors suggest another direct learning approach to estimate the fractional quantity for such a problem.

2.1 Experimental Results

We compared the *Fractional Approximation* and *Retrospective Estimation* [7] with *Two-models* [10] and *Transformed Outcome* [24] estimating only $CATE^i_Y$ in an experimental study. The study comprised over 100 million website visits. Figure 2 presents the comparison of models' uplift performance; Figure 3 presents the comparison of the models' ROI profile such that:

$$ROI = \frac{\Delta Revenue - \Delta Investment}{\Delta Investment}$$

The stars on each curve represent the best operating points (in terms of total treatment effect) for each method, subject to $ROI \geq 0$ constraint (zero-budget).

In terms of unconstrained uplift effect, the *Transformed Outcome* method achieves the highest *Area Under the Uplift Curve* ($AUUC = 0.912$), and the *Two-Models* method has the maximal potential treatment effect (109%). However, those methods have a limited feasible solutions space ($ROI \geq 0$), resulting in at most 38.6% of the potential effect.

At the same time, the *Fractional Approximation* method achieves the highest population coverage (38%) at $ROI \geq 0$ and reaches 65.8% of the potential treatment effect. The *Retrospective Estimation* method provides the maximal treatment effect achieving 79.7% of the possible uplift with $AUUC = 0.805$. The solution allows to run a self-sponsored promotional campaign, with zero-budget, while achieving up to 80% of the possible impact.

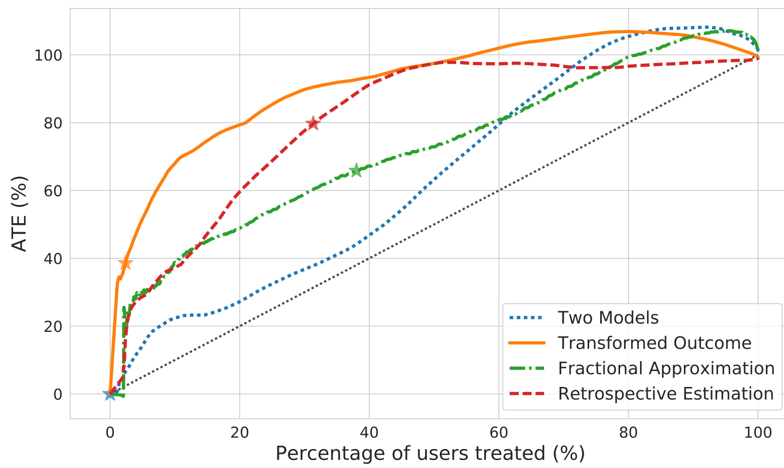


Fig. 2. Qini curves comparison [4]

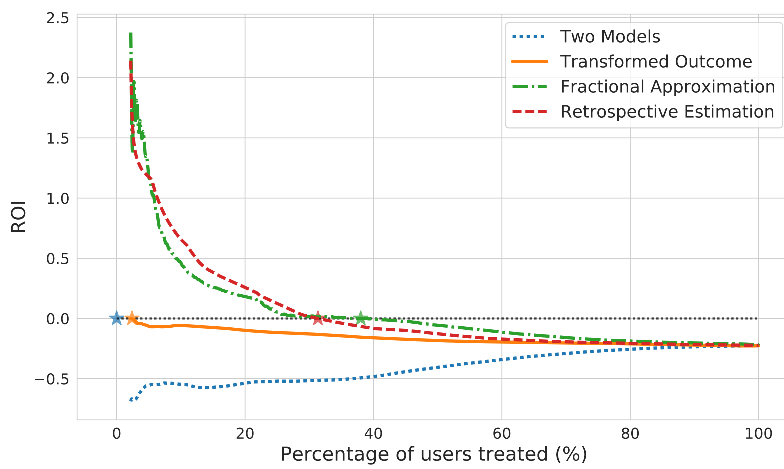


Fig. 3. Qini-ROI curves comparison [7]

3 MULTIPLE-CHOICE PROMOTIONS RECOMMENDATION

In this section, we extend the previously presented problem, such that we can offer each customer at most one promotion from a finite set of promotions that are eligible. In other words, rather than deciding whether to give or not to give an offer, the marketer determines which promotion to assign to each user. Similar to the previous section, a global budget constrains the overall incremental net revenue loss generated by the promotions.

A general case like this no longer allows for trivial incremental comparisons. Examining the incremental uplifts, we question whether the comparison be made with the no-promotion baseline or with the second-best option. Recent research tackled the multiple-promotion uplift problem using meta-learning, taking conversion and cost into account [27]. Another promotion recommendation study recommended an offline-constrained optimization strategy [15], based on a constant promotion cost assumption. However, in our online e-commerce setup, these solutions are not feasible, since the decision about promotion selection needs to be made in real-time under the current budget constraints. We proposed a solution method relying on the Online Multiple-Choice Knapsack Problem [29] that is fully described in [1].

3.1 Solution Approach

We express the causal effects of the promotions with two variables: $Y_i(k)$ represents the potential purchase if customer i is offered the promotion k, while $R_i(k)$ represents the potential net revenue if customer i is offered the promotion k. Likewise, $Y_i(0)$ and $R_i(0)$ represent the potential outcomes if no promotion is offered to customer i, an option that is always available for every customer. We define the conditional average treatment effect on Y_i and R_i for a customer with pre-promotion covariates x as follows:

$$CATE_Y(i,k) = \mathbf{E}(Y_i(k) - Y_i(0) \mid X = x_i)$$

$$CATE_R(i,k) = \mathbf{E}(R_i(k) - R_i(0) \mid X = x_i)$$

Both quantities $CATE_Y(i,k)$ and $CATE_R(i,k)$ can be positive or negative. The CATE on the expected net revenue loss $\mathcal{L}$ is:

$$CATE_{\mathcal{L}}(i,k) = -CATE_R(i,k)$$

We estimate $CATE_Y(i,k)$ and $CATE_{\mathcal{L}}(i,k)$ per each customer i and each promotion $k \in K_i$ using uplift modeling trained on past treatment data. Item $k = 0$ for which $CATE_Y(i,0) = 0$ and $CATE_{\mathcal{L}}(i,0) = 0$ represents the base treatment of not giving a promotion, and therefore has no incremental effect. The target of the optimization problem is for every customer i picking a single item $k^* \in K_i$ in order to maximize the total incremental impact $CATE_Y(i,k^*)$, while the total selected $CATE_{\mathcal{L}}(i,k^*)$ would not exceed the budget constraint C.

This problem is formalized as the *Multiple-Choice Knapsack Problem* (MCKP) [22] in Equation 2, where, $Z_{i,k}$ is a binary assignment variable indicating if a customer i should be offered the item k.

$$\begin{aligned}
&\textbf{Maximize} \sum_{i \in U} \sum_{k \in K_i} CATE_Y(i,k) \cdot Z_{ik} \\
&\textbf{subject to:} \\
&1.\ \sum_{i \in U} \sum_{k \in K_i} CATE_{\mathcal{L}}(i,k) \cdot Z_{ik} \leq C \qquad \forall i \in U, k \in K_i \\
&2.\ \sum_{k \in K_i} Z_{ik} = 1 \qquad \forall i \in U \\
&3.\ Z_{ik} \in \{0,1\} \qquad \forall i \in U, k \in K_i
\end{aligned} \tag{2}$$

In our case, the value of each item (a promotion offered to a specific customer) v_{ik} is $CATE_Y(i,k)$ and the weight of each item w_{ik} is $CATE_{\mathcal{L}}(i,k)$. Different from the classical setup, we allow the weights and values of the items to be negative. For practical applications, we investigate the Online-MCKP [28], a variation of the MCKP where customers arrive one-by-one. Here, we need to decide which promotion to offer each customer in an online manner.

3.2 Optimization Framework

We address the problem with a two-phased approach: an estimation phase and an optimization phase. The first phase estimates $CATE_Y(i,k)$ and $CATE_{\mathcal{L}}(i,k)$. In the second phase, we then use these estimations and solve the MCKP and the Online-MCKP. CATE estimation for both Y and R are achieved using uplift modeling and data from previous randomized controlled trials. We consider different estimators such as two-models [10], transformed outcome [2] and X-learner [12]. Common evaluation metrics such as Qini Curves and Qini Score (Area under the Uplift Curve [4] - $AUUC$) are used for model selection. It results in estimations of $CATE_Y(i,k)$ and $CATE_{\mathcal{L}}(i,k)$ for every customer i and every promotion k. These quantities will serve as the input item sets K_i for the following optimization phase.

Similar to the 0-1 knapsack problem, the optimization phase relies on an approximation solution to overcome the computational limitations and fit for online environments. In [1] we extend previously known solutions [29] of MCKP Lueker's algorithm [14] and allow for negative values, negative weights, and possible negative budget constraints, which are essential for our business case and problem formulation. We suggest a four-step optimization solution:

(1) Eliminate dominated items from the solution space.
(2) Calculate incremental values and weights to allow comparison between items.
(3) Transform the incremental quantities to efficiency angles to allow sorting of positive and negative quantities.
(4) Select a single item according to an efficiency angle threshold, designed to meet the capacity constraints.

The outline of the solution is presented in algorithm 1. The proposed algorithm addresses the assignment problem in an online manner, and is able to tune the strategy according the up-to-date remaining budget status.

3.3 Solution Example

We demonstrate the optimization flow by illustrating a toy example on Figure 4. In this example, we observe various promotional offers for four different customers (green-squares, red-circles, blue-triangles, and purple-rhombuses) given a budget of zero. Namely, we need to pick one promotion per customer, such that the total weight (expected net revenue loss $\mathcal{L}$) will not be positive and the total value will be the maximal. For each promotion k we present its value

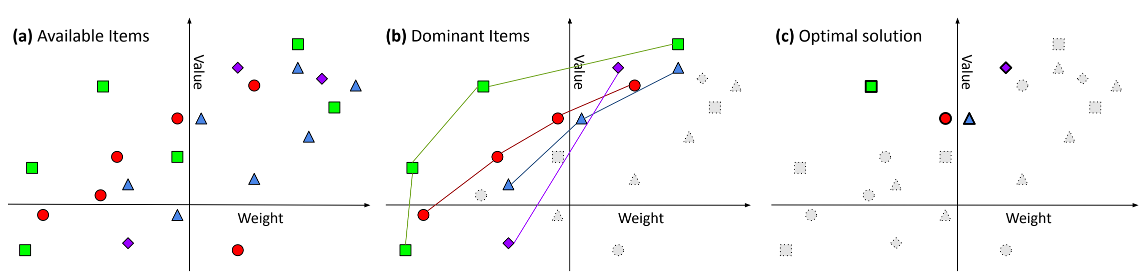


Fig. 4. Example of the solution method. (a) all available promotional options of 4 customers (colors and shapes) on value and weight axes; (b) the dominant items per-customer (colored); (c) the optimal assignment given a threshold angle.

Algorithm 1 Online MCKP [1]

Input:
- Customer set U
- Item sets K_i
- Knapsack capacity C

$P \leftarrow \emptyset$
for $(i \in U \mid 1 \leq i \leq |U|)$ **do**
  $D_i \leftarrow$ dominant items of K_i sorted by increasing weight
  for $d \in D_i$ **do**
    Compute incremental values and weights $(\overline{v}_{id}, \overline{w}_{id})$:
    if d=0: $\overline{w}_{i0} \leftarrow w_{i0}$; $\overline{v}_{i0} \leftarrow v_{i0}$
    else: $\overline{w}_{id} \leftarrow w_{id} - w_{id-1}$; $\overline{v}_{id} \leftarrow v_{id} - v_{id-1}$
    Compute efficiency angle θ_{id}:

$$\theta_{id} = \begin{cases} \frac{3\pi}{2} & if \;\; \overline{v}_{id} = 0 \;\wedge\; \overline{w}_{id} = 0 \\ 2\pi + atan2(\overline{v}_{id}, \overline{w}_{id}) & if \;\; \overline{v}_{id} < 0 \;\wedge\; \overline{w}_{id} \leq 0 \\ atan2(\overline{v}_{id}, \overline{w}_{id}) & else \end{cases}$$

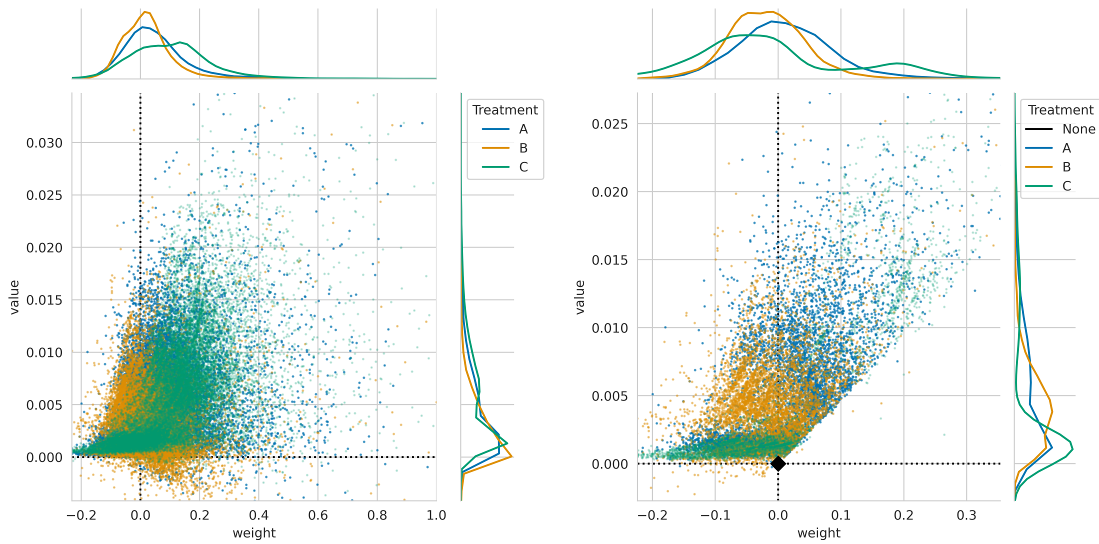
    $P \leftarrow P \cup (\theta_{id}, w_{id})$
  end for
  Sort P by decreasing angle θ_p
  for $p \in P$ **do**:
    if p=0: $f(\theta_0) = w_0/|P|$
    else: $f(\theta_p) = f(\theta_{p-1}) + w_p/|P|$
  end for
  Update efficiency threshold θ^*:

$$\theta^* \leftarrow min_{p \in P} \left\{ \theta_p \mid f(\theta_p) \leq \frac{C}{\frac{|P|}{i} \cdot (|U| - i + 1)} \right\}$$

  Find dominant item d^*:
    $d^* \leftarrow \arg\min_{d \in D_i} \{\theta_{id} \mid \theta_{id} \geq \theta^*\}$
  Update capacity:
    $C \leftarrow C - w_{id^*}$
  Pick item d^*
end for

$CATE_Y(i,k)$ and weight $CATE_{\mathcal{L}}(i,k)$ on a two-dimensional chart, as shown in sub-figure (a). We observe promotions in all four quadrants of the axes, representing both positive and negative expected value (y-axis) and weight (x-axis). Sub-figure (b) presents a partial solution of the problem, by eliminating dominated items for each customer - resulting in a value-weight Pareto efficient frontier. Given the frontier we remain with the question - which of the dominant items should we pick. This would be determined by comparing the incremental values of each item to the efficiency-angle threshold learned from past data.

Sub-figure (c) presents the optimal solution to the problem. In this case, we also pick promotions with positive weight for the blue and purple customers since their weight is compensated with the negative weight of the selected green and red promotions. We can observe that the total value (overall position of selected options on the y-axis) is higher than any other possible combination within the budget constraints.



(a) Input dataset

(b) Optimal Solution

Fig. 5. Distribution of value and weight on the discounts dataset (a) and the optimal On-MCKP assignment solution (b).

3.4 Experimental Results

An assessment of the potential impact of various uniformly assigned promotions was conducted before we optimized the promotion assignment. The treatment groups in our randomized controlled trial received three different levels of discounts on Booking.com products. The target metrics - completion of purchases, promotion costs, and incremental revenues - were aggregated and compared between control and treatment groups, yielding an average treatment effect on net revenue and purchase completion per promotion. Despite conclusively positive treatment effects, each promotion resulted in a loss in net revenue. Ideally, a zero-budget promotional campaign will be able to operate for a long time using a solution including personalized promotions.

More than 20 million entries were generated in the study. For each data point, the binary variable Y was used to represent the completion of a sale, the continuous variable R was used to represent the total net revenue (including the promotional costs), and covariates X were used to represent customer characteristics. We selected the best model from a number of uplift modeling techniques to obtain $CATE_Y(i,k)$ and $CATE_{\mathcal{L}}(i,k)$ and picked the best model according to the highest Qini Score on the test set.

Next, for each customer $i \in U$ and each potential promotion k we calculated the expected value ($CATE_Y(i,k)$) and the expected weight ($CATE_{\mathcal{L}}(i,k)$) based on the models predictions. This resulted in $|U| \times K_i$ rows, where the number of treatments per customer is $m_i = 4 \ \forall i \in U$ (three promotion levels and a no-discount treatment). Figure 5(a) displays the joint distribution of value and weight across three discount levels (A,B,C) on a normalized scale. We observe items in all four quadrants of the chart, with a vast majority in the first quadrant, meaning that we predict that the promotions have a positive value and a positive weight for most customers.

We evaluated the *On-MCKP* solution, which simulates the real-world scenario where customers arrive one at a time, and the decision of which promotion to offer is made live, at each time-step. At the beginning of the process, we have no information about the general weights and values distributions. The method adapts the promotion assignment decision to the remaining budget and the updated efficiency angle function as described in Algorithm 1.

The outcome of the promotions assignment is depicted in Figure 5(b). The solution demonstrates a near-optimal performance with a maximal optimality gap of 0.245% (compared to the optimal Integer Linear programming results) on the real discounts dataset. The suggested solutions widely outperform the greedy benchmarks and play a game-changing role in allowing the promotional campaign to become self-sponsored, with a significant improvement compared to simple *Greedy* (+51%) solutions. We observe an interesting phenomenon, where the selected promotions follow a linear value/weight trend. This forms a separation bound by excluding the solutions below the line, which is equivalent to the efficiency threshold used by the method. It consists of a blend of all three available promotions, using items from three possible quadrants.

4 REWARD CONDITIONS OPTIMIZATION

Similar to the previous methods, here we aim to maximize incremental sales given budget (or incremental ROI) constraints. However, instead of allocating different promotions or deciding whether to reward a promotion to a customer based on its cost-effectiveness, we focus on tuning the reward conditions. More specifically, while the same promotion is offered on all supply and all customers, the qualification conditions (such as minimal spend requirement) might vary. A good example for such condition is presented in Figure 1, where a customer needs to complete a purchase above a certain amount in order to qualify for the promotion. This method allows the marketer to control for the efficiency of the promotion on the targeted audience, while at the same time maintaining a consistent promotional offer to all customers. The optimization of the reward condition can be done on various resolutions, ranging from a global decision (uniform minimal spend requirement), through a geography based decision (different conditions in different cities), to individual product or even customer-level decisions.

Such method relies on a simple assumption that a strict reward condition (such as a very high minimum spend requirement) will affect smaller population but result in cost-effective budget allocation, while loose reward conditions (or no conditions at all) will have a great impact on the sales, but result in a significant budget loss. Therefore, the

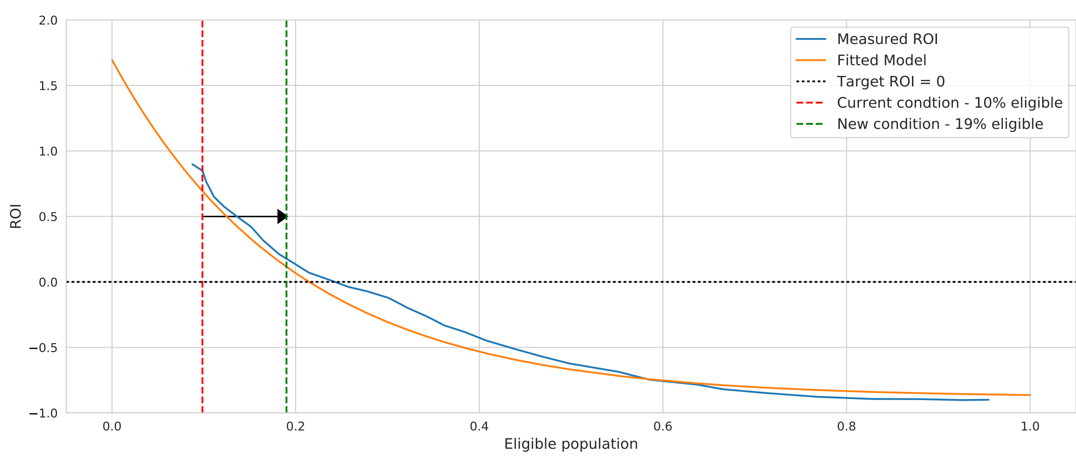


Fig. 6. An example of eligibility criteria tuning, based on ROI-Model

optimization method is seeking for an operating point, in terms of reward condition, which maximizes the size of the exposed population, subject to the budget constraints. Moreover, while such decisions can be achieved separately in each sub-segment (for example destination city), a global budget allocation policy can reach higher sales results.

By using empirical data of purchases as a response to different reward conditions, we can build an ROI-elasticity curve, mapping between the portion of the exposed population (θ) according to reward condition and the associated ROI. The blue line in Figure 6 represents the empirical ROI-elasticity curve gathered from one of the targeted geographies. As expected, the curve is decreasing with the portion of exposed population, reducing the cost efficiency of the promotion when exposed to a bigger audience. It is important to note that the curve has an intersection with the target ROI value (in our example - $ROI = 0$), allowing us to update the operating point threshold to the desired value θ^*.

However, during an online campaign, the general traffic distribution might differ and therefore the optimal threshold θ^* can shift accordingly. We propose a simple curve fitting technique to address this problem. As presented in the orange curve in Figure 6, we assume a non-linear monotonic relation between the portion of the exposed population θ and the resulted ROI. We suggest to model this relation as an exponential curve, by learning the parameters a,b and c using the Levenberg-Marquardt least-squares curve fitting algorithm [19] to select a new θ^* such that:

$$\widehat{ROI}(\theta) = ae^{b\theta} + c$$
$$\widehat{ROI}(\theta^*) = 0$$

Using this parametric curve fitting method allows incorporating historical and up-to-date data in ROI estimation, by refitting the curve every time-period. The update step $\theta^*_{t-1} \rightarrow \theta^*_t$ (moving from the red to green operating thresholds in Figure 6) can be re-evaluated on the new data, allowing to build a new elasticity curve in the next iteration.

5 PROMOTIONAL OFFERS PERSISTENCE

In this section we survey the persistence lever, as described in details in [9]. This lever implies how long and under which conditions, a personalization model should keep offering the same promotion to the customer.

Given a wide variety of offers $A \in \mathcal{A}$, our personalized solution aims to pick the best treatment A^* to maximize the expected business impact. As mentioned in previous sections, we can model the conditional treatment effect, or any other desired outcome of a treatment A with a learner function $\mathcal{L}$. The context of the input would use customer pre-treatment covariates $\hat{X}$ at time t. Our goal is to pick the best treatment A^* such that:

However, a customer might interact with our platform multiple times, generating repeated browsing sessions and additional model calls. The model decision about the optimal treatment may change, due to context change, model parameters update or even due to an intrinsic exploration strategy within the model [3]. Such change can have a direct impact on customer browsing experience with an unexpected change of a promotional offer in the middle of the purchase process. This may result in a negative effect on customer satisfaction, purchase completion, and introduce instability into effect measurement. Therefore, sometimes it is necessary to maintain the initial decision on the treatment, to allow a consistent customer experience.

Theoretically, the optimization model may determine itself whether to persist with the previously selected decision. Previous customer interactions' data can be fed into the learner model $\mathcal{L}$, by using sequential [18] or reinforcement [23] learning methods such that:

$$A^*(\hat{X}, t) = \arg\max_A \mathcal{L}(A, \hat{X}, t, A[t-1])$$

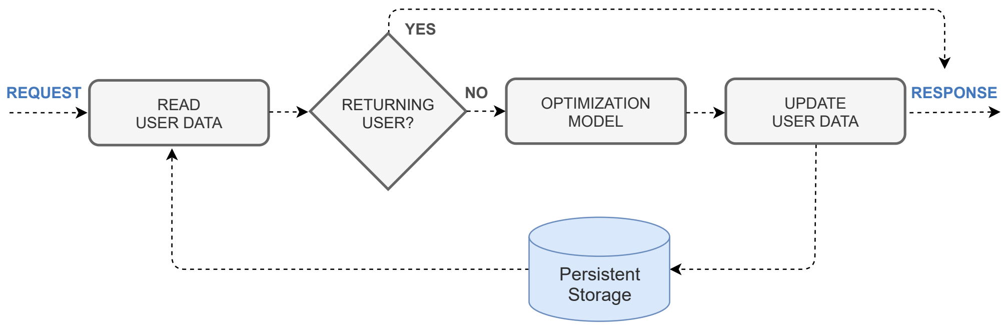


Fig. 7. Persistence mechanism flowchart

In practice, online applications rely on delayed, sparse and biased feedback [16]. Therefore, learning $\mathcal{L}(A, \hat{X}, t, A[t-1])$ is challenging, and while a holistic optimization method can achieve a global optimal strategy, it still might harm the experience consistency of individuals.

Using an explicit persistency mechanism will ensure consistent model behavior and allow more control and clarify about the assignment of the offers. At the same time, to allow flexibility, setting the right expectations with the customer such as "*This offer is valid for the next three days*" is crucial. Therefore, while the persistence challenge might sound purely technical, it is important to involve user experience, copywriting, and business stakeholders in the process.

$$A^*(\hat{X}, t) = \arg\max_A \mathcal{L}(A, \hat{X}, t)$$

We suggest implementing a flexible storage-based system, to ensure a consistent user journey while continuously learning and optimizing offers allocation. A distributed key-value database will be used to keep track of the personalized recommendations. At each point of customer-system interaction, the system decides whether to stick with the previously offered treatment or recalculate a new up-to-date offer (as shown in Figure 7). The mechanism is specified by two hyper-parameters:

- *Persistence Key* - a set of attributes defining the persistence.
- *Persistence Timeout* - treatment persistence time span.

The *Persistence Key* defines the set of context properties that require to stick to the same treatment. To guarantee the same treatment, the key can be constructed from customer and session characteristics (for example, stick for the same decisions give the same user identifiers and search destination). The *Persistence Timeout* defines how long the decision remains valid. Parameters for the key, as well as conditions for timeouts, are often defined based on business use cases. Examples include session length, coupon expiration time, and customer expectations. Parameters can be tuned to achieve optimal performance.

In our previous work [9] we presented a simulated study demonstrating the optimality gaps of different persistence strategies compared to the best, up-to-date treatment. In this study, we observe the cost of the persistence in our scenario, in the form of sub-optimal treatments, which should be evaluated against the long-term benefit of a consistent user experience. In order to estimate the cost of harming the consistent user experience, we recommend testing the various persistence strategies via online experimentation for different personalization use cases and suggest examining how methodological persistence keys and timeouts work, and whether include persistence within optimization models.

6 CONCLUSION

In this paper we showcased four levers that can be adjusted for the purpose of maximizing promotion allocation while remaining within budget constraints. In addition to modifying one of these levers per campaign, one can also try multiple combinations of different levers. For example, combining campaign eligibility modeling with promotion reward conditions optimization. Of course, the more complex the campaign will be, the harder it will be to accurately measure and optimize. To enable such multi-dimensional decision making, there needs to be a robust pipeline or system in place where all these decisions can be made. This system will take in a campaign with all its viable possibilities dictated by the business, and outputs the optimized campaign with the right benefit, reward condition, and persistence setting on the instance level.

It is important to note that even while all these levers could be optimized, the success of a campaign is still strongly dependent on the overall customer experience. This includes the customer-facing design and copywriting, and how clear and attractive the benefit is communicated to the customer or how explainable the campaign is for customer support to answer related incoming inquiries Further, one can achieve optimal campaign configuration with the benefit offered per customer, the right reward condition, and the right persistence threshold, but the campaign will fail due to poor customer-facing user interface. Moreover, such promotions tend to have long term effects on customer loyalty, and given a dynamic business environment, the optimisation decisions need to be re-calibrated periodically. Introducing advanced methods such as exploration via bandits and cross-campaign learning may help with fast iterations and dynamic updates, but at the same time create an additional level of complexity and evaluation challenges. We look forward to investigating the different usages of the levers discussed, and validating this approach as a one-stop-shop for personalized promotions.

ACKNOWLEDGMENTS

We thank Thomas Bosman, Nerie Ohana and Itsik Adiv for their help in editing this paper; to *Booking.com Value Research Group* for fruitful ideation and discussions; and to Igor Spivak, Amir Tzur and Ola Czajkowska for making these concepts work in real-life systems.

REFERENCES

[1] Javier Albert and Dmitri Goldenberg. 2021. E-Commerce Promotions Personalization via Online Multiple-Choice Knapsack with Uplift Modeling.

[2] Susan Athey and Guido W Imbens. 2015. Machine learning methods for estimating heterogeneous causal effects. *stat* 1050, 5 (2015), 1–26.

[3] Lucas Bernardi, Pablo Estevez, Matias Eidis, and Eqbal Osama. 2020. Recommending Accommodation Filters with Online Learning. In *Proceedings of the Workshop on Online Recommender Systems and User Modeling (ORSUM @ RecSys 2021).*

[4] Floris Devriendt, Darie Moldovan, and Wouter Verbeke. 2018. A literature survey and experimental evaluation of the state-of-the-art in uplift modeling: A stepping stone toward the development of prescriptive analytics. *Big data* 6, 1 (2018), 13–41.

[5] Eustache Diemert, Artem Betlei, Christophe Renaudin, and Massih-Reza Amini. 2018. A large scale benchmark for uplift modeling. In *KDD.*

[6] Shuyang Du, James Lee, and Farzin Ghaffarizadeh. 2019. Improve User Retention with Causal Learning. In *The 2019 ACM SIGKDD Workshop on Causal Discovery*. PMLR, 34–49.

[7] Dmitri Goldenberg, Javier Albert, Lucas Bernardi, and Pablo Estevez. 2020. Free Lunch! Retrospective Uplift Modeling for Dynamic Promotions Recommendation within ROI Constraints. In *Proceedings of the 14th ACM Conference on Recommender Systems*.

[8] Dmitri Goldenberg, Kostia Kofman, Javier Albert, Sarai Mizrachi, Adam Horowitz, and Irene Teinemaa. 2021. Personalization in Practice: Methods and Applications. In *Proceedings of the 14th International Conference on Web Search and Data Mining*.

[9] Dmitri Goldenberg, Guy Tsype, Igor Spivak, Javier Albert, and Amir Tzur. 2021. Learning to Persist: Exploring the Tradeoff Between Model Optimization and Experience Consistency. In *Companion Proceedings of the Web Conference 2021*. Association for Computing Machinery, New York, NY, USA, 527–529. https://doi.org/10.1145/3442442.3452051

[10] Behram J Hansotia and Bradley Rukstales. 2002. Direct marketing for multichannel retailers: Issues, challenges and solutions. *Journal of Database Marketing & Customer Strategy Management* 9, 3 (2002), 259–266.

- [11] Guido W Imbens and Donald B Rubin. 2010. Rubin causal model. In *Microeconometrics*. Springer, 229–241.
- [12] Sören R Künzel, Jasjeet S Sekhon, Peter J Bickel, and Bin Yu. 2019. Metalearners for estimating heterogeneous treatment effects using machine learning. *Proceedings of the national academy of sciences* 116, 10 (2019), 4156–4165.
- [13] Ying-Chun Lin, Chi-Hsuan Huang, Chu-Cheng Hsieh, Yu-Chen Shu, and Kun-Ta Chuang. 2017. Monetary discount strategies for real-time promotion campaign. In *Proceedings of the 26th International Conference on World Wide Web*. 1123–1132.
- [14] George S Lueker. 1998. Average-case analysis of off-line and on-line knapsack problems. *Journal of Algorithms* 29, 2 (1998), 277–305.
- [15] Rahul Makhijani, Shreya Chakrabarti, Dale Struble, and Yi Liu. 2019. LORE: a large-scale offer recommendation engine with eligibility and capacity constraints. In *Proceedings of the 13th ACM Conference on Recommender Systems*. 160–168.
- [16] Themis Mavridis, Soraya Hausl, Andrew Mende, and Roberto Pagano. 2020. Beyond algorithms: Ranking at scale at Booking. com. In *Proceedings of the Fourth Workshop on Recommendation in Complex Scenarios. CEUR-WS*.
- [17] Alex Miller and Kartik Hosanagar. 2020. Personalized discount targeting with causal machine learning. (2020).
- [18] Sarai Mizrachi and Pavel Levin. 2019. Combining Context Features in Sequence-Aware Recommender Systems.. In *RecSys (Late-Breaking Results)*. 11–15.
- [19] Jorge J Moré. 1978. The Levenberg-Marquardt algorithm: implementation and theory. In *Numerical analysis*. Springer, 105–116.
- [20] Diego Olaya, Wouter Verbeke, Jente Van Belle, and Marie-Anne Guerry. 2021. To do or not to do: cost-sensitive causal decision-making. *arXiv preprint arXiv:2101.01407* (2021).
- [21] Alon Sela, Erez Shmueli, Dima Goldenberg, and Irad Ben-Gal. 2016. Why spending more might get you less, dynamic selection of influencers in social networks. In *2016 IEEE International Conference on the Science of Electrical Engineering (ICSEE)*. IEEE, 1–4.
- [22] Prabhakant Sinha and Andris A Zoltners. 1979. The multiple-choice knapsack problem. *Operations Research* 27, 3 (1979), 503–515.
- [23] Csaba Szepesvári. 2010. Algorithms for reinforcement learning. *Synthesis lectures on artificial intelligence and machine learning* 4, 1 (2010), 1–103.
- [24] Irene Teinemaa, Javier Albert, and Dmitri Goldenberg. 2021. Uplift Modeling: from Causal Inference to Personalization. (2021).
- [25] Paolo Toth and Silvano Martello. 1990. *Knapsack problems: Algorithms and computer implementations*. Wiley. 14–15 pages.
- [26] Kui Zhao, Junhao Hua, Ling Yan, Qi Zhang, Huan Xu, and Cheng Yang. 2019. A Unified Framework for Marketing Budget Allocation. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 1820–1830.
- [27] Zhenyu Zhao and Totte Harinen. 2019. Uplift modeling for multiple treatments with cost optimization. In *2019 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*. IEEE, 422–431.
- [28] Yunhong Zhou, Deeparnab Chakrabarty, and Rajan Lukose. 2008. Budget constrained bidding in keyword auctions and online knapsack problems. In *International Workshop on Internet and Network Economics*. Springer, 566–576.
- [29] Yunhong Zhou and Victor Naroditskiy. 2008. Algorithm for stochastic multiple-choice knapsack problem and application to keywords bidding. In *Proceedings of the 17th international conference on world wide web*. 1175–1176.