Monetary Discount Strategies for Real-Time Promotion Campaign

Ying-Chun Lin† Chi-Hsuan Huang† Chu-Cheng Hsieh§
Yu-Chen Shu¶ Kun-Ta Chuang†
Dept. of Computer Science and Information Engineering, National Cheng Kung University†
Slice Technologies, Inc., San Mateo, CA, USA§
Dept. of Mathematics, National Cheng Kung University¶
{yclin, chihsuan}@netdb.csie.ncku.edu.tw chucheng@slice.com
{ycshu, ktchuang}@mail.ncku.edu.tw

ABSTRACT

The effectiveness of monetary promotions has been well reported in the literature to affect shopping decisions for products in real life experience [3]. Nowadays, e-commerce retailers are facing more fierce competition on price promotion in that consumers can easily use a search engine to find another merchant selling an identical product for comparing price.

We study e-commerce data — shopping receipts collected from email accounts, and conclude that for non-urgent products like books or electronics, buyers are price sensitive and are willing to delay the purchase for better deals. We then present a real-time promotion framework, called the RTP system: a one-time promoted discount price is offered to allure a potential buyer making a decision promptly.

To achieve more effectiveness on real-time promotion in pursuit of better profits, we propose two discount-giving strategies: an algorithm based on *Kernel density estimation*, and the other algorithm based on *Thompson sampling* strategy. We show that, given a pre-determined discount budget, our algorithms can significantly acquire better revenue in return than classical strategies with simply fixed discount on label price. We then demonstrate its feasibility to be a promising deployment in e-commerce services for real-time promotion.

1. INTRODUCTION

A successful business model often relies on successful marketing strategies [16] – which are often based on *Product*, *Place*, *Price*, and *Promotion*. Among them, probably the most effective and direct approach is to take advantage of promotion (i.e. providing discounts on label price). For example, popularized by `priceline.com`, the Name-Your-Own-Price (NYOP) [10] has achieved huge success through allowing consumers to ask for a discount on price to facilitate sales [7]. In this research, we extend the strategy to another aspect though empowering businesses to actively adjust their selling price through offering a dynamic discount – we name it the *Real-Time Promotion* (RTP) problem.

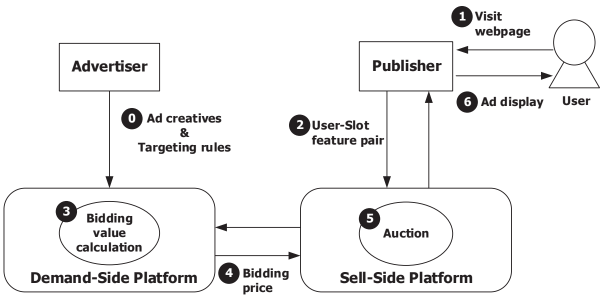


Figure 1: The flow of a RTB system to display an ad in a website.

Akin to the famous *Real-Time Bidding* (RTB) problem [5, 9, 15, 27, 29, 30], which aims at finding the best strategy to spend the budget on placing an advertisement to reach target buyers (Figure 1), the solution of the RTP problem is to optimize the promoting budget in pursuit of the highest revenue. The intuition is to maximize the profit through finding a balance between experimenting how much target audiences are willing to pay and taking some risk of turning away a customer by limiting the discount. The deployment of the RTP strategy is ease of implementation: displaying a one-time promoted discount price to a potential buyer; our goal is to find the "sweet-spot" discount in a way that the competent price is striking a chord with the customer. For simplicity, we assume the discount price is a one-time offer (which also serves as an incentive for an customer to complete a purchase) – a buyer is expected to take this one or none.

We formalize the challenge into a *Discount-Giving Strategy* problem (Section 2). Naturally, no discount or a steep fixed discount are problematic. The no-discount strategy turns away many potential buyers; a steep fixed discount maximizes the sale numbers with a cost of losing an opportunity of getting more profit. It is reasonably viable to determine the discount price in pursuit of the best balance between these two ends. Motivated by this observation, we formalize the challenge as solving the exploration-

WWW 2017, April 3–7, 2017, Perth, Australia.
ACM 978-1-4503-4913-0/17/04.
http://dx.doi.org/10.1145/3038912.3052616

exploitation dilemma [25], and propose RTP-aware strategies, including the optimal profit estimation in the offline manner and stochastic-based *Thompson-sampling* strategy in the online scenario. (Section 3.2).

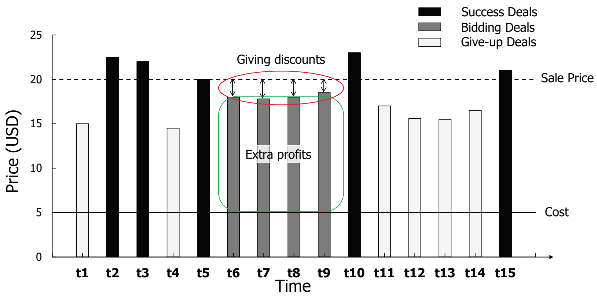


Figure 2: Extra profit gain from Real-Time Promotion.

We use Figure 2 to illustrate the problem setup. Assuming that a business is willing to provide a fixed amount of budget for incentivizing its customers, our goal is to determine how much the discounts should be assigned to each incoming customer. For example, if we are given a budget of \$8, and we lower our sale price to \$18 between the time $t6$ and the time $t9$ through giving out discounts, we accumulate additional profit of \$52, or 4 * (\$18 - \$5). One might argue that if we further lower the price to \$14 starting from $t1$, we will get even more profit. However, in practice, the ideal fixed discount that maximizes the profit is always unknown. Moreover, almost all businesses set some limitation on discounts, because, as discussed in previous research [11, 17], the consistent promotion affects consumer brand choice and may hurt the company image in future.

We conceived that the RTP strategy is well suited for an online marketing promotion. The pop-up advertisements today are often distracting or ineffective. On the contrary, a just for me one-time discount could be an intriguing offer, especially if the price meets a customer's sweet-spot. With a right incentive strategy, it eventually increases the total purchase amount and brings extra profit to the business.

One of the most critical challenges is the design of strategies to determine the best price and to simulate user purchase intent when receiving a discount message. In this paper, we use and observe the online shopping receipts provided by *Slice Technologies*[1]. Slice downloaded and parsed machine-generated receipts for millions of user accounts. Based on the observations, we assume that if a shopper had paid for an item at a price, the shopper is willing to accept a promotion that is lower than the price she paid. In Economics, the maximum price at or below which a shopper accepts to pay, refers to their "willingness to pay (WTP [26]."

For simplicity, we also assume that online shoppers reject a promotion if the prices are higher than the price she/he paid, which are recorded in her/his receipts. That is to say, we use the price labelled on each receipt to be the criterion for accepting an offer or rejecting it at a given time in our experiments, providing us a possible way to evaluate the effectiveness of monetary promotion strategy to online shoppers.

[1]`https://www.slice.com/`

In this work, we propose two strategies: an algorithm based on kernel density estimation, named *optimal estimation* (Section 3.1), and the other algorithm based on *Thompson sampling* strategy (Section 3.2). Figure 3 shows the flow diagram of the proposed RTP system. The RTP system can be easily integrated into e-commerce websites. When a customer searches for a product, our algorithm(s) will determine the discount based on the information learned in the offline/online scenario.

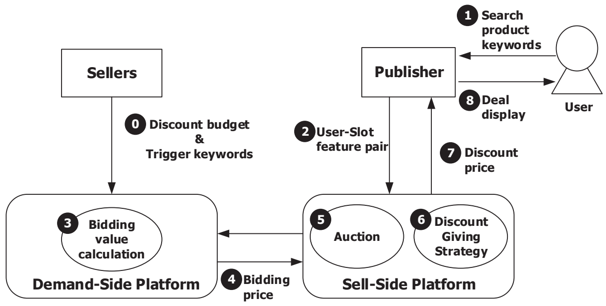


Figure 3: The flow of a RTP system to give discounts in any e-commerce platform.

In our experimental studies, the feasibility is evaluated by using return on investment (defined in Section 4.2.1). We conducted experiments across different categories of items based on real-world transactions instead of synthetic data. Our results show that, with the help of our proposed algorithms, businesses will easily gain much more profit than putting a fixed discount on label price.

2. PROBLEM FORMULATION

In this section, we define the Discount-Giving Strategy (DGS) in the RTP system. In an offline setting, the RTP system is devised to reach the maximum profits in light of the selling history. We also consider the situation that, in the online environment, the customers arrive the system in random order during the promotion period. In light of this, it is difficult for the online RTP system to achieve maximum profits similar to the case in the offline scenario. As such, the DGS problem is presented as a stochastic problem in the online environment. Based on the foregoing, we turn to maximize the profits during the promotion period instead of pursuing offline result. The frequently used symbols are summarized in Table 1 for reference. We give necessary definitions at first.

Definition 1. ***Profit***: The profit of a product y is the final sale price s_y subtracting the discount $k(\cdot, y)$ and the cost c_y, such as direct labor unit costs, direct material unit costs, and bidding costs, and so on:

$$r(\cdot, y) = s_y - k(\cdot, y) - c_y.$$

Definition 2. ***Deal***: An agreement between a seller and a buyer for trading a product at the certain price. If a deal is successful, it means that a buyer agrees to buy a product at the given discount. The deal fails otherwise. Note that it's possible that a RTP system chooses not to offer a deal — it implies that the provider is not interested in offering an deal to a buyer.

Symbol	Description
B	Discount budget
$\vec{u_x}$	Feature vector of the customer x
s_y	Sale price for product y
$S_y(\vec{u_x})$	predicted WTP price (with features from a customer $\vec{u_x}$ for the product y)
c_y	Cost for product y
$r(x, y)$	Profit of produce y for customer x
$k(x, y)$	Discount of product y for customer x
η	Random variable to signify WTP prices
$f_y(\eta)$	The PDF distribution of price η given a product y
$\xi(x, y)$	WTP price (customer x for the product y)
$p(x, y, k)$	Probability of successful selling y to customer x with discount k

Table 1: Symbols.

Definition 3. ***Discount-Giving Strategy***: Determine the discount amount $k(x, y)$ for each potential customer x to

$$\text{maximize} \sum_{j=1}^{N} r(x_j, y)$$

$$\text{subject to} \sum_{j=1}^{N} k(x_j, y) \leq B,$$

where N is the number of customers.

The DGS problem can be modeled as the knapsack(KP) problem. The items in the KP problem are the deals in the DGS problem. Each item has its weight w_i and its value v_i. Given that a knapsack has capacity W, the goal is to maximize the total profits of the items in the knapsack. Similarly, a deal in the DGS problem has its discount $k(x, y)$ and its profit $r(x, y)$. An RTP system has a discount budget B in total during the promotion period, and the goal of RTP is to maximize the total profits of the promotion. Different from the KP problem, the DGS problem has an optimal solution as long as we know the distribution of the WTP price for product y. In this paper, the highest price that a customer will buy the product y is referred to be the WTP price. The following section is an instance for calculating the optimal solution.

2.1 Offline Optimal Discount Giving

Suppose N customers (x_1, x_2, $\cdots$, x_N) visit the RTP system during the promotion and the system offers different discounts for each deal. We assume that the product y will be specified by the customers (they might already be browsing the page of y at the moment). Note that no customer considers every product and vice versa; therefore each product y are visited by different customers x, and each of them has corresponding features $\vec{u_x}$). For simplicity, we assume that y is fixed. As aforementioned, the expected profit will be

$$\sum_{j=1}^{N} (s_y - k(x_j, y) - c_y) p(x_j, y, k(x_j, y)). \tag{1}$$

Our goal is to maximize the expected profit by giving the customer real time discount $k(x, y)$ under the budget B.

The difficulty is to determine the value of $k(x, y)$ and the probability $p(x, y, k)$ by using the feature vector $\vec{u}_x$. To simplify the problem of determining the probability $p(x, y, k)$, the WTP price of the product y is predicted by $S_y(\vec{u_x})$ for each customer. Then, $k(x, y)$ is obtained from $k(x, y) = s_y - S_y(\vec{u_x})$. Suppose that the WTP price distribution $f_y(\eta)$ of customers who will buy the product y at price η is known, which can be estimated from the selling history. Then the probability that the product will be sold at the discount k is

$$p(y, k) = p(\cdot, y, k) = \int_{s_y - k}^{\infty} f_y(\eta) d\eta.$$

If we give a fixed discount k, the profit will be

$$\sum_{j=1}^{N} (s_y - k - c_y) p(y, k).$$

Since we have budget B, the optimal fixed discount k_{opt} can be determined from the equation

$$B = N \int_{s_y - k_{\text{opt}}}^{s_y} f_y(\eta) \cdot (s_y - \eta) d\eta.$$

It means that we try to attract most people who will buy the product under the limited budget. And the optimal profit is

$$N \left(\int_{s_y - k_{\text{opt}}}^{\infty} f_y(\eta) \left(\min\{s_y, \eta\} - c_y\right) d\eta \right).$$

This solution can be applied when we have a long-time promotion or we can reach almost all the customers. Note that for real cases, we still need a real-time strategy without the assumption of looking all purchase beforehand.

2.2 Online Stochastic Discount Giving

When the RTP system is deployed in an online environment, customers visit the system in a random order. This means that the customers whose willingness-to-pay (WTP) price is greater than price $s_y - k_{\text{opt}}$ may not appear during the limited promotion period (the information is not able to be sensed). Different from the optimal result in an offline setting, it is more viable to deploy the online strategy of the RTP system to achieve better profits as possible.

To maximize the profits during the online process, we convert the DGS problem into a Multi-Armed Bandit (MAB) problem [2, 12]. In the MAB problem, a gambler standing in front of several slot machines has to decide a machine to play at any given time to maximize total expected rewards. However, the gambler does not know the underlying reward distribution for each machine. In every round, she/he has to decide to choose a machine to gain rewards or to understand more about its reward distribution. The decision problem is referred to as the exploration-exploitation dilemma in the literature [25].

For a stochastic MAB problem, given a discounts/actions set A ($A = \{a_1, a_2, \cdots, a_M\}$),the reward distribution of each action is assumed to be independently and identically distributed [14]. In each trial t, a learner chooses an action or a discount $a_t \in A$ and the reward r is drawn with the probability $P(r|a_t)$. The expected reward for the trial t is

$$\mathbb{E}_t(r) = \int_0^{\infty} r \cdot P(r|a_t) dr. \tag{2}$$

The expected total rewards after $T+1$ trials are

$$\sum_{t=0}^{T} \mathbb{E}_t(r).$$

The goal of the learner is to maximize the total rewards.

In the DGS problem, the action set can be viewed as the discrete discount prices. Every time t that a customer visits the RTP system, the system has to choose a discount based on the past experience. The reward r of each trial t is

$$r = s_y - a_t - c_y. \tag{3}$$

Since the final reward for the discount offering is known, $P(r|a_t)$ becomes the purchase probability of a deal with discount price a_t. The expected total profits are

$$\sum_{t=1}^{T} r \cdot P(r|a_t).$$

The goal of the online DGS is to maximize the total profits during the promotion period. The way to estimate $P(r|a_t)$ will be discussed in the next session.

3. REAL-TIME PROMOTION STRATEGY

A good RTP strategy can solve DGS problem by making two important decisions. First, a strategy should decide whether to give discounts to a customer or not. In Figure 2, we have demonstrated that spending discount budget on the competitive customers can gain extra profits in return. Second, once the system decides to give a discount to a customer, the strategy needs to determine further the value of discounts such that the customer will buy a product and the seller can win more profits. However, the customer distribution f_y(PDF of the WTP given a product y), the price at which a customer is willing to pay for a product $\xi(x, y)$, and the purchase probability $P(r|k)$ are all unknown. In this section, we apply several machine learning methodologies and sampling techniques to learn those unknown distributions in both offline and online manners.

3.1 RTP with Optimal Estimation Strategy

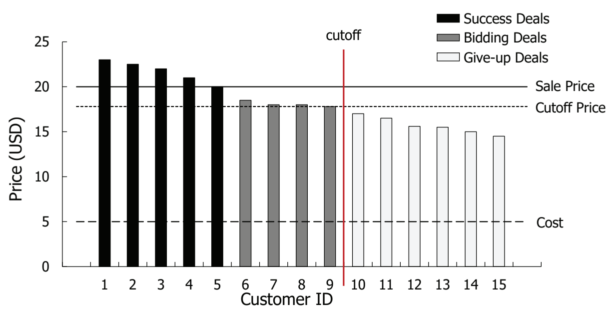


Figure 4: Illustration of offline optimal strategy.

In Section 2.1, we have shown that the offline DGS problem has an optimal solution. Here we discuss how the proposed optimal estimation (RTP$_{\text{oe}}$) strategy obtains the maximum profits.

Suppose that we know the distribution f_y of the price that a customer will buy for the product y, in short, the distribution of the WTP prices. The distribution f_y can be obtained from the purchase history before the promotion period. Then, we sampled N customers in f_y and they will all buy the product y during the promotion period. Suppose that we know the WTP price of the sampled customers. The customers are sorted in the decreasing order according to their WTP prices, shown in Figure 4.

Originally, if the system does not give discounts to the customers, only customers who have the WTP price higher than the s_y will agree on the deal. For example, the customer 1 to 5 will buy the product in Figure 4. To achieve the optimal total profits, the RTP$_{\text{oe}}$ strategy prefers to convince the customers with higher WTP prices to buy the product because the profit decreases as long as the WTP price decreases. The strategy provides a discount $k(x, y) = s_y - \xi(x, y)$ for those high WTP price customers, where the $\xi(x, y)$ is the WTP price of the customer x for a product y. Therefore, the strategy consumes less discount budget and gains as many profits as possible. After deciding how much a discount is given to each customer, the strategy decides who should receive the discount considers the discount budget B. Intuitively speaking, B can be viewed as a spending quota, signifying the willingness of risking company image for reaping a short-term profit gain.

The total discounts have a limit set by the discount budget B, and the optimal total profits are achieved by giving discounts to customers with higher WTP price. According to the distribution of the WTP prices f_y, the maximum discount price k_{opt} has to satisfy the following equation

$$B = N \int_{s_y - k_{\text{opt}}}^{s_y} f_y(\eta) \cdot (s_y - \eta) d\eta, \tag{4}$$

where N is the number of customers, and $\int_{s_y - k_{\text{opt}}}^{s_y} f_y(\eta) \cdot (s_y - \eta) d\eta$ is the average discount offered to customers. Since the RTP$_{\text{oe}}$ strategy now predicts the maximum discount price, it gives discounts to the customer who has the WTP price between $s_y - k_{\text{opt}}$ and s_y. Clearly, the customers who have the WTP price within $[s_y - k_{\text{opt}},\ s_y]$ will be the set of potential discount receivers, and the value of $s_y - k_{\text{opt}}$ is called as the cutoff price in the RTP$_{\text{oe}}$ strategy. For example, the discounts will be given to customers 6~9 in Figure 4, since $\sum_{x=6}^{9} s_y - \xi(x, y) \le B < \sum_{x=6}^{10} s_y - \xi(x, y)$. The cutoff price is the WTP price of customer 9.

The optimal total profits are known after we obtain the cutoff price. The maximum total profits that can be obtained during the promotion period is

$$N\left(\int_{s_y - k_{\text{opt}}}^{s_y} f_y(\eta)(\eta - c_y) d\eta + \int_{s_y}^{\infty} f_y(\eta)(s_y - c_y) d\eta\right). \tag{5}$$

Suppose that all WTP price are known, as shown in Figure 4, the optimal total profit is $\sum_{x=1}^{5}(s_y - c_y) + \sum_{x=6}^{9}(\xi(x, y) - c_y)$.

3.1.1 Cutoff Price Estimation

In practice, the distribution of f_y is unknown, and the distribution f_y varies depending on the product y, which is shown in Section 4.3. To estimate the distribution f_y, we employ a non-parametric statistical method, namely Kernel Density Estimation (KDE) [21] to learn the distribution from the previous purchase history.

KDE is an approach to estimate the probability density function of a random variable. In our $\mathrm{RTP}_{\mathrm{oe}}$ strategy, the random variable is the WTP price ξ for the customer x_j and product y. Let the WTP prices $(\xi(x_1, y), \xi(x_2, y), \cdots, \xi(x_N, y))$ in the purchase history be an independent and identically distributed sample. The samples are drawn from some distributions with an unknown density f_y. The kernel density estimator is

$$\widehat{f_y}(\xi) = \frac{1}{Nh} \sum_{j=1}^{N} K\left(\frac{\xi - \xi(x_j, y)}{h}\right),$$

where the $K(\cdot)$ is the kernel, which satisfies $\int K(u)du = 1$ and $K(u) \geq 0$. The parameter h controls the smoothness of the estimated $\widehat{f_y}(\xi)$. A large h heads toward an over-smoothed distribution, and a small h refers to a possible under-smoothed outcome. In our experiments, we set $h = 0.75$.

After obtaining the distribution $\widehat{f_y}(\xi)$, the cutoff price of the product y is estimated by Equation (4). It is the sale price of y subtracting the maximum discount k_{opt}. The maximum total profits after investing the discount budget B are calculated as Equation (5).

3.1.2 Discount Price Estimation

Algorithm 1 $\mathrm{RTP}_{\mathrm{oe}}$

Input: s_y : sale price, $S_y(\vec{u})$: price range classifier, $\vec{u}_x$: user feature vector, k_{opt}: maximum discount price, B_{left}: left budget;
Output: $k(x, y)$:discount, B_{left}:left budget;
$\xi(x, y) = S_y(\vec{u}_x)$
if $\xi(x, y) \geq s_y$ **then**
  $k(x, y) = 0$
else
  $k(x, y) = s_y - \xi(x, y)$
  if $k(x, y) > k_{\mathrm{opt}}$ or $k(x, y) > B_{left}$ **then**
    $k_{x,y} = 0$
return $k(x, y)$, B_{left}

The $\mathrm{RTP}_{\mathrm{oe}}$ strategy needs to know the WTP price of each customer to achieve maximum total profits as well. If the probability $p(x, y, k)$ is known, the $\mathrm{RTP}_{\mathrm{oe}}$ strategy can find the potential discount receivers and give them adequate discounts to gain more profits. However, it is difficult to estimate $p(x, y, k)$, especially when the purchase history of a customer is not sufficient for precise distribution estimation.

Instead of estimating the probability $p(x, y, k)$, we transform the problem into a classification problem. Since the discount usually can be logically represented as some integers, such as \$1, \$5 and so on, the customers can be classified into different discount prices, $k(x, y)$. As such, potential discount receivers are these customers who are classified into the discount prices not exceeding the maximum discount.

Specifically, based on the purchase history of a product y, which is richer than a customer purchase history, a model $S_y(\vec{u}_x)$ takes the customer feature vectors $\vec{u}_x$ as input and classifies the customers into different price ranges. $S_y(\vec{u}_x)$ can be any multiclass classifier, such as SVM [4] or Decision Tree[19]. The model only classifies those customers who search the product y during the promotion period. Based on the foregoing, the $\mathrm{RTP}_{\mathrm{oe}}$ strategy takes the upper bound of the price range as the WTP price of a customer. Similarly, the discount price $k(x, y)$ is s_y subtracting the WTP price.

In the offline process, namely the RTP system with the $\mathrm{RTP}_{\mathrm{oe}}$ strategy, both maximum discount k_{opt} and the price range classifier $S_y(\vec{u})$ are learned from the purchase history of a product y. The discounts are given by the RTP system once it is triggered either by searching or viewing the web page of a product y. The flow of the $\mathrm{RTP}_{\mathrm{oe}}$ strategy offering a discount for each potential buyer is outlined in Algorithm 1. The algorithm makes sure that offering such a discount for the buyer can help to earn additional profits.

3.2 RTP with Thompson-sampling Strategy

Essentially, the maximum optimal profits, which are calculated by the $\mathrm{RTP}_{\mathrm{oe}}$ strategy, are difficult to achieve during the online process. Since the order of customers access the RTP system is unpredictable, it is possible that quite a few potential discount receivers do not visit the RTP system during the promotion period. As a result, the budget spending becomes over conservative. Also, it is possible that our estimation are suffering from other bias. In this section, we propose another strategy to regulate the momentum of spending of budget.

The RTP system using the $\mathrm{RTP}_{\mathrm{oe}}$ strategy leads to optimal profits if the distribution f_y is independent and identically distributed all around the year. In practice, seasonal effect appears almost everywhere. For example, the price during the holiday season is often lower than other period. Most of the WTP prices should go further below the cutoff price, or customers are unlikely to agree upon a deal during holiday sales.

Algorithm 2 $\mathrm{RTP}_{\mathrm{oe+tps}}$

Input: s_y : sale price, $S_y(\vec{u})$: price range classifier, $\vec{u_x}$: user feature vector, k_{opt}: maximum discount price, B_{left}: left budget, $\{S_a\}$: the number of success deals for each action, $\{F_a\}$: the number of fail deals for each action, α, β: the prior parameters;
Output: $k(x, y)$:discount, B_{left}:left budget;
$\xi(x, y) = S_y(\vec{u}_x)$
if $\xi(x, y) \geq s_y$ **then**
  $k(x, y) = 0$
else
  $k(x, y) = s_y - \xi(x, y)$
  if $k(x, y) > k_{\mathrm{opt}}$ **then**
    for $a \in A$ **do**
      $\widehat{\theta_a} \sim Beta(S_{a,t} + \alpha, F_{a,t} + \beta)$
    $k(x, y) = \arg\max_a \mathbb{E}_a(r)$
  if $k(x, y) > B_{left}$ **then**
    $k(x, y) = 0$
return $k(x, y)$, B_{left}

To address the trend seasonal effect and to ensure our regulation on the spending momentum of the budget, we propose the Thompson-sampling ($\mathrm{RTP}_{\mathrm{tps}}$) strategy. The $\mathrm{RTP}_{\mathrm{tps}}$ strategy tackles the online DGS problem through treating it as a multi-arm bandit(MAB) problem (Section 2). The given discount prices are a set of actions. The $\mathrm{RTP}_{\mathrm{tps}}$ strategy utilizes Thompson sampling [1], which is usually applied to address the MAB problem, to decide which action should be taken. The action is the amount of the discount price.

Thompson sampling is a heuristic for choosing actions. It determines the action based on the expected reward. The action which has maximum expected total profits will be selected. The expected reward in each trial is estimated as Equation (2). The reward distribution $P(r|a)$ is learned from the experience.

Given $T{+}1$ observed data $D_T = \{(a_0, r_0), (a_1, r_1) \cdots (a_T, r_T)\}$ from trial 0 to T and the likelihood function $P(r|a, \theta)$, the posterior distribution is

$$P(\theta|D_T) \propto P(D_T|\theta)P(\theta|\cdot) = \prod_{t=0}^{T} P(r|a_t, \theta)P(\theta).$$

In the trial $T + 1$, a reward parameter θ is sampled from the posterior distribution, $\widehat{\theta}_{T+1} \sim P(\theta|D_T)$. Based on the sampled $\widehat{\theta}_{T+1}$, it can determine the next action by

$$a_{T+1} = \arg\max_{a} \int_{0}^{\infty} r \cdot P(r|a, \widehat{\theta}_{T+1})dr$$

In light of the Thompson sampling algorithm, the $\mathrm{RTP_{tps}}$ strategy chooses the discount price which has maximum expected profits. Once the $\mathrm{RTP_{tps}}$ strategy decides the action to take, which is the amount of the discount price, the reward of a trial t is determined as Equation (3). The reward distribution $P(r|a, \theta)$ becomes the purchase probability when the discount price a_t and the reward $r_{y,t}$ are given. Specifically, the reward distribution can be interpreted as the probability that the bidding deal is successful. This can be modeled by a Bernoulli distribution with an unknown parameter θ. We use a Beta distribution as the conjugate prior, $P(\theta) = Beta(\theta; \alpha, \beta)$. The posterior distribution then becomes

$$P(\theta|D_T) = Beta(S_{a,t} + \alpha, F_{a,t} + \beta),$$

where $S_{a,t}$ is the number of success deals with action a and $F_{a,t}$ is the number of fail deals. Then, the next action is determined by

$$a_{T+1} = \arg\max_{a}\ \mathbb{E}_a(r) = r \times (\frac{B_{left}}{a}) \times P(r|a, \widehat{\theta}_{T+1}).$$

In Algorithm 2, we outline a mixed strategy, called $\mathrm{RTP_{oe+tps}}$, which combines advantages of $\mathrm{RTP_{oe}}$ and $\mathrm{RTP_{tps}}$. Specifically, the $\mathrm{RTP_{oe+tps}}$ strategy is used for giving a discount when a customer is not a potential discount receiver, which is shown in lines 6˜9 in Algorithm 2. Note that the parameter S_a and F_a are updated once a customer x accepts or rejects the deal with discount a. Through the mixed strategy, we can guarantee that the RTP system can help a platform provider to gain a lot more extra profits in a stochastic online environment.

4. PERFORMANCE EVALUATION

We now will evaluate the performance of our proposed strategies and show the feasibility of the RTP-aware strategies. To prove that our strategies can optimize the promotion budget, we run our online simulation on a real dataset for several different product categories. The sensitivity of the RTP system is shown by how much the system spends on training the data of the various data sizes.

4.1 Experimental Setup

The dataset is the online shopping receipts provided by *Slice Technologies*. The purchase history is acquired by parsing machine-generated receipts, from Amazon, Ebay, Walmart and so on, in 2013 (in total ˜38,600,000 purchase transactions). Each record in the purchase history contains the user id, product description, purchase date, purchase price, merchant, and product category. In the classification model (described in Section 3), user features include the most recent purchase date and price [13], the purchase merchant, and the purchase month.

We run the online simulations to show how much is the additional profit gain. For each product, the online simulation is conducted in the product purchase history. To ensure randomness, we randomly sample testing data from the product purchase history for each simulation. The size of the testing data is 30 percent of the data. The rest of the data is for offline training, such as the cutoff price estimation and the discount price estimation. In each round of the simulation, the RTP system offers a one-time promoted discount price for each test customer to simulate the promotion period.

The online simulation is run 100 times for each product to ensure the robustness of our RTP system. The experiment results are the outcomes of 100 online simulations in average. The discount budget is set according to the size of testing data. The total profit gain is the final result of a simulation.

We compare the profit gain with the no-discount strategy where customers do not receive any discounts. For each simulation, we set the sale price of a product. The no-discount strategy offers a deal with the sale price s_y to each customer. If the price on a receipt of a customer is no less than the offered price, a deal is successful, and a seller can win the profit from the customer. We can obtain the total profits of the no-discount strategy after a round of the simulation.

We also compare the result with fixed discount strategies. To ensure the fixed discount strategy in its best performance price, we find the optimal discount price for the comparing strategies (worst fixed and opt fixed as depicted in experimental figures). Both two fixed discount strategies are developed under the assumption that we know the actual price distribution of the product. With the given sale price, we find the optimal discount price k for them in the offline process.

$$\arg\max_{k} N \int_{s_y - k}^{s_y} f_y(\eta) \cdot (s_y - k)d\eta$$

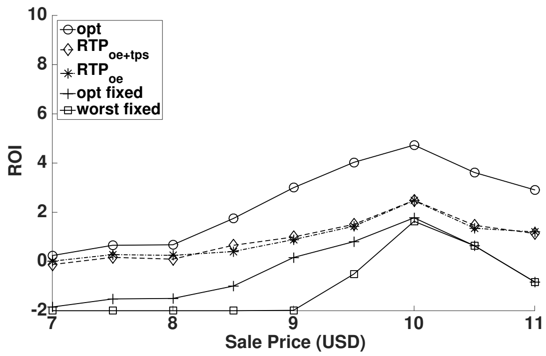
If these two strategies use the optimal discount price, they can earn the maximum profits under the given price distribution. The optimal fixed strategy offers a deal with the optimal discount price $s_y - k$ for each customer. The worst fixed strategy demonstrates the worst case of the optimal fixed strategy.

Finally, we compare all the strategies with the optimal (opt) result of the DGS problem. Recall in Section 2; we have discussed that the DSG problem has an optimal solution if we know the price distribution of a product and the WTP price for each customer. The opt case in experimental figures represents this optimal result.

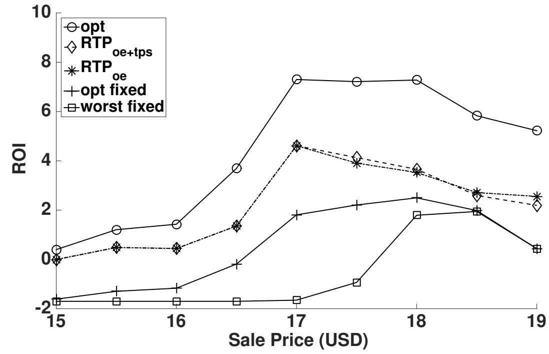
4.2 Strategy Performance Analysis

The performance analysis is discussed in terms of three factors: return on investment (ROI), success rate, and model training time. The four popular product categories of books, electronics, health, and food are examined in our experiments. Due to the space limitation, we only pick one product in each category for demonstrating the result.

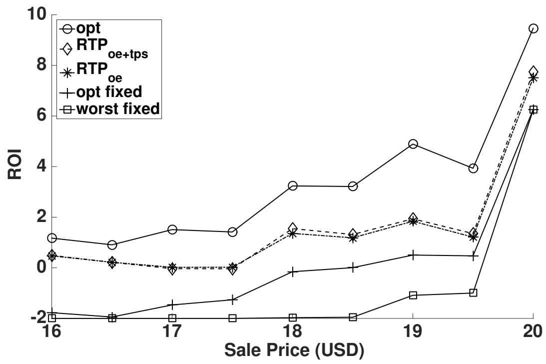
4.2.1 *Return on Investment*



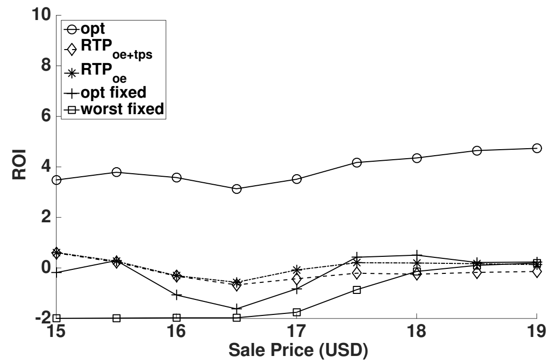
(a) The book "Quiet: The Power of Introverts in a World That Can't Stop Talking" with discount budget USD$600.



(b) The headphones "JVC HAFX1X Headphone" with discount budget USD$100.



(c) The product "Dropps Laundry Pacs, Fresh Scent, 80-Load Pouch" with discount budget USD$30.



(d) The food "Cheese Thin Crust Pizza" with discount budget USD$50.

Figure 5: The ROI of the simulation result in average.

The ROI is usually used in the literature of business management. The purpose of the ROI metric is to evaluate the performance of an investment. The basic definition of ROI is $\frac{gain - cost}{cost} = \frac{S_y - k(x,y) - cost}{cost}$, in which $gain - cost$ is the profit comparing against the no-discount strategy. We use the lowest price among all receipts of a product as the baseline (cost). Note that the cost is subtracted from the gain when we calculate the total profits. The higher the ROI value corresponds to the better result.

Intuitively speaking, a negative ROI implies the profit we reaped is a bad investment, meaning that the budget we spent, i.e. a discount or $k(x, y)$, is lower than the profit we reap from a deal. On the contrary, a high ROI means that we use a small discount to exchange a huge profit.

We can compare the performance of different strategies in Figure 5. Figure 5 shows the arithmetic average ROI of each strategy given a hypothetical list price. The x-axis refers to a hypothetical retail price (before the discount), and the y-axis corresponds to the ROIs of different strategies. Each sale price of a receipt becomes a WTP such that we can run the simulation for computing ROIs. The center of x-axis refers to the average receipt price, e.g. \$9 in Figure 5(a). We assume the average price reported on receipts is possibly close to the prevailing list price of a product; therefore, we zoom-in the graph to display x close to the average price.

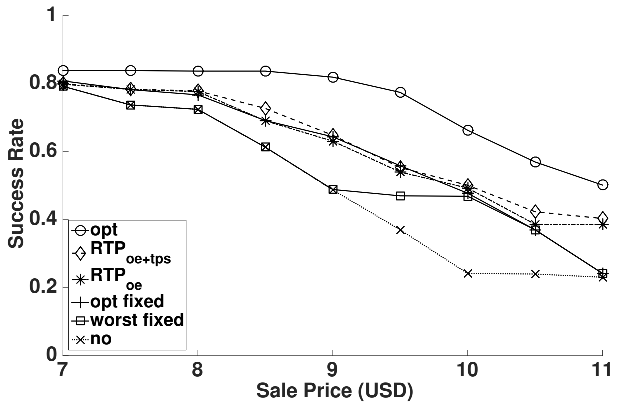
In general, the RTP_{oe} and $\text{RTP}_{\text{oe+tps}}$ perform better than the two fixed discount strategies. Specifically, the RTP system has more extra profits gain in the book, electronics, and health product categories. In Figure 5c, the extra profits gained even reaches eight times more than the discount budget. This means that with USD\$30 discount budget, the RTP system can earn USD\$240 extra profits. In fact, whether an RTP system can gain extra profits is highly related to the customer behavior of a category. The details are discussed in Section 4.3.

The ROI value of the two fixed discount strategies is often below zero, which means that the promotion strategy fails. It even suggests that the promotion strategy leads to the profit loss, comparing to the no-discount strategy. The profit loss is observed from Figure 5. We will discuss what causes the risk of profit loss in Section 4.3.

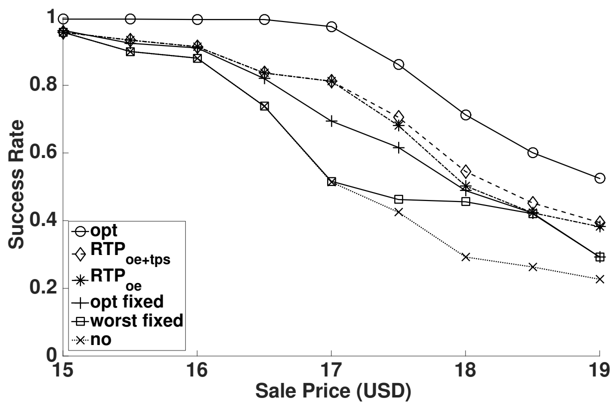
The performance of $\text{RTP}_{\text{oe+tps}}$ is comparable with RTP_{oe}. In general, the ROI value of $\text{RTP}_{\text{oe+tps}}$ is higher when the sale price is higher. Most of the customers do not accept the price exceeding the estimated cutoff price during the online simulation. Therefore, RTP_{oe} seldom gives discounts due to the high cutoff price. On the other hand, $\text{RTP}_{\text{oe+tps}}$ can detect the change of the WTP price distribution during the online simulation. We also observed that most users do not accept the price with small discounts from the reward distribution, and this increases the discount amount to maximize the total profits.

4.2.2 *Success Rate*

The success rate can demonstrate how many product units



(a) The book "Quiet: The Power of Introverts in a World That Can't Stop Talking".



(b) The headphones "JVC HAFX1X Headphone".

Figure 6: The success rate of a deal during the promotion period.

a system can sell during the promotion period. This value is formalized as the number of the success deals divided by the number of all deals, and is an important performance index for the inventory management. We can now observe how the proposed strategies balance between the number of the success deals and the additional profit gain.

In Figure 6, an obvious trend for all strategies is that the success rate decreases when the sale price increases. Because the sale price is high, the number of customers who are willing to buy the product at the price decreases. The success rate drops accordingly.

While comparing results between Figure 5a and Figure 6a, as well as 5b and Figure 6b, we find that the RTP system can make a better decision on when to bid on or give up the customer. The RTP system does not sacrifice the number of success deals to achieve more additional profits in return. On the contrary, it achieves the better performance both on the ROI and the success rate comparing to the fixed discount strategies.

On the other hand, $\mathrm{RTP}_{\mathrm{oe+tps}}$ has higher success rate than $\mathrm{RTP}_{\mathrm{oe}}$ because $\mathrm{RTP}_{\mathrm{oe+tps}}$ adjusts the discount price according to the learned success rate for each discount price, which approaches the reward distribution. $\mathrm{RTP}_{\mathrm{oe+tps}}$ calculates the expected total profits considering both success rate and the profit. As a result, $\mathrm{RTP}_{\mathrm{oe+tps}}$ can make a better balance between the total profits and success rate.

4.2.3 Sensitivity Analysis

Before a product promotion is initiated in the RTP system, the classification model, and the cutoff price should be ready in the offline process. The time spent on the online process should be within a millisecond. In this study, we will show that both $\mathrm{RTP}_{\mathrm{oe}}$ and $\mathrm{RTP}_{\mathrm{oe+tps}}$ are viable strategies regarding model training time.

Category	Size	KDE Time (ms)	Class. Time (ms)
book	4,261	0.026	1.208
electronics	443	0.0030	0.0317
food	245	0.0024	0.0191
health	170	0.0015	0.0049
game	144	0.0011	0.0084

Table 2: Training time for items in different categories.

In Table 2, the model training time is shown regarding the number of the purchase transactions in the log (denoted by "Size" in Table 2). The "KDE time" represents the execution time that the $\mathrm{RTP}_{\mathrm{oe}}$ strategy uses to estimate the cutoff price of a product. The "Class. time" represents the time spent on training a classification model. We use SVM as our classification model. The classification model determines the discount price for a bidding deal. All offline processes are finished within a second. The turnaround time for the offline process is extremely short, as compared to the number of days during the product promotion.

4.3 Discussions

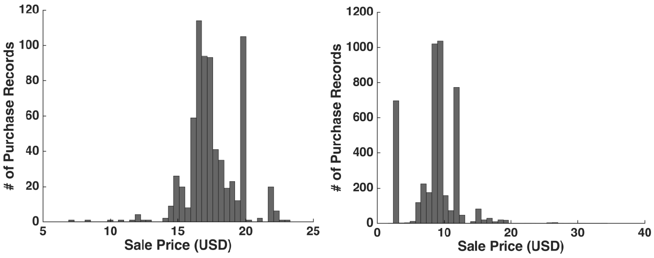
In this section, we discuss some observations concluded from the experiment results. First, we identify that the RTP system can work better in the non-urgent categories. Then, we elaborate on why the ROI is highly related to the distribution. Finally, we demonstrate why the fixed discount strategies may have the risk of profit loss.

4.3.1 The effect of the product categories

After several experiments on other product categories, we found that the RTP system is most suitable for the product categories in which the buyers of the product can delay their purchase. This means that the buyers can wait until the price comes down to the range where they are willing to pay. As can be seen in Figure 5, the RTP system has a high ROI value in most of the products, except the food product. The result is related to customer behavior when buying a product in a food category. Customers tend to buy food immediately since it belongs to living essentials, despite the fact that the sale price is not relatively economical. On the other hand, they are willing to wait until the price of a smartphone goes down. We have undertaken several experiments and consistently observe the same result. The result shows that the RTP system performs well in these categories that buyers are willing to wait for the competent discount.

4.3.2 The effect of the price distribution

The ROI is highly related to the price distribution of a product. The relation can be observed from Figure 5 and Figure 7, which are come from the logs of the same book and the same headphone. In Figure 7a, the number of customers buying the electronic product suddenly increase around USD$16.5. The increase causes the ROI rise around USD$17 in Figure 5b. The similar result is shown around USD$10 in Figure 7b and USD$9.5 Figure 5a. Such results support that if an RTP system gives small discounts, it is possible to have more extra profit gain.



(a) The headphones.   (b) The book.

Figure 7: Price distribution.

4.3.3 The risk of profit loss

We can easily observe that a fixed discount strategy has the risk of loss of profits, especially when the sale price is low. We still observe the risk even when we find the optimal discount price for a fixed discount strategy. The cause of the profit loss is due to the price distribution. Because a fixed discount strategy gives the same discount price for every customer, it can gain the extra profits from the population who are willing to pay the price between the cutoff price and the sale price. However, the profit gain is obtained by investing a discount budget. If the gain cannot recover the cost of the discount budget, it becomes negative. Therefore, the total profits of the fixed discount strategies are less than that of the no-discount strategy.

In our experiments, we have shown that the RTP system with $\mathrm{RTP}_{\mathrm{oe}}$ or $\mathrm{RTP}_{\mathrm{oe+tps}}$ is capable of gaining more extra profits from observing the ROI. Both RTP-aware strategies also have higher success rates due to the better discount setting tactics. The RTP system is a feasible method for providing real-time discounts during the promotion period. Further, we observe that the RTP system has better ROI on the products where the customers can delay the purchase to wait for a better price. In summary, the experiment results demonstrate that this new paradigm can be a promising marketing strategy.

5. RELATED WORK

Our work emphasizes on how the RTP system gives discounts to maximize the extra profit gain. The amount of discount price is subject to a discount budget. The RTP system dynamically gives discounts according to the customer features or the learned the purchase probability for different discount prices from the market. Essentially, the problem is similar to the dynamic pricing strategy.

The dynamic pricing problem is one of the problems in revenue management [24]. It is a pricing strategy in which businesses set flexible prices for products or service based on current market demands. The pricing strategy aims at maximizing the revenue. Many businesses, such as airlines or hotels [18, 23], have deployed the dynamic pricing strategies. Schlosser *et al.* [20] used logistic regression model for predicting the purchase probability of a product. The model dynamically adjusts the price based on the sale probability. In our work, we study the dynamic pricing problem in another perspective. Our expected profits are subject to a fixed budget which is generally ignored in aforementioned studies.

Some dynamic pricing strategies are subject to the amount of the resources. Babaioff *et al.* [20] studied the pricing strategy which maximizes the revenue under the given number of items and potential customers. Singer *et al.* [22] investigated the similar problem in the crowdsourcing markets. They studied how to allocate the tasks to the workers such that they can maximize the number of finished works and minimize the total expense. The problem is also subject to a fixed budget, but the proposed solution is orthogonal to the design for our motivation. Some other works [6, 8, 15] studied the similar problems and resolved the problem by the optimization methods or the machine learning techniques. Though our work similarly focuses on the problem of maximizing the target and subject to some constraints, the relation between the discount price and the profits cannot be correspondingly considered in the previous problem settings.

Several studies discussed online promotion systems. The most popular topic among them is the RTB system. The research topics related to the RTB system is about Click-Through Rate estimation [32], winning price prediction [28] or bidding strategies [15, 31]. They focused on increasing CTR instead of the total profits. Another topic is to predict the WTP of a product. Zhao *et al.* [33] used linear regression model to predict WTP price for a customer. The model can be part of our RTP system for estimating discount price. However, these works are different from the DGS problem in nature.

6. CONCLUSION

In this work, we proposed a novel discount-giving system, called Real-Time Promotion, for e-commerce services to determine the real-time discount in pursuit of high revenue. We proposed various strategies, including $\mathrm{RTP}_{\mathrm{oe}}$ which uses KDE and a classification model to estimate proper discounts for the bidding deals and $\mathrm{RTP}_{\mathrm{oe+tps}}$ which further considers the randomness of the online process. In addition to evaluating the effectiveness of $\mathrm{RTP}_{\mathrm{oe}}$ for offline optimal revenue, we also demonstrated the feasibility of $\mathrm{RTP}_{\mathrm{oe+tps}}$ in the online scenario. As validated in studies on real data, the RTP framework has been recognized to be a promising promotion strategy for e-commerce services.

7. ACKNOWLEDGMENT

This paper was supported in part by Ministry of Science and Technology, R.O.C., under Contract 105-2221-E-006-140-MY2. We also thank Joshua Borden for providing valuable editorial feedback to the paper.

8. REFERENCES

[1] S. Agrawal and N. Goyal. Analysis of thompson sampling for the multi-armed bandit problem. In *COLT 2012 - The 25th Annual Conference on Learning Theory*, 2012.

[2] P. Auer, N. Cesa-Bianchi, and P. Fischer. Finite-time analysis of the multiarmed bandit problem. *Machine Learning*, 47(2-3), 2002.

[3] P. Chandon, B. Wansink, and G. Laurent. A benefit congruency framework of sales promotion effectiveness. *Journal of marketing*, 64(4), 2000.

[4] C.-C. Chang and C.-J. Lin. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2, 2011. `http://www.csie.ntu.edu.tw/~cjlin/libsvm`.

[5] Y. Chen, P. Berkhin, B. Anderson, and N. R. Devanur. Real-time bidding algorithms for performance-based display ad allocation. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2011.

[6] A. V. den Boer and B. Zwart. Dynamic pricing and learning with finite inventories. *Operations Research*, 63(4), 2015.

[7] M. Ding, J. Eliashberg, J. Huber, and R. Saini. Emotional bidders an analytical and experimental examination of consumers' behavior in a priceline-like reverse auction. *Management Science*, 51(3), 2005.

[8] G. Gallego and G. van Ryzin. Optimal dynamic pricing of inventories with stochastic demand over finite horizons. *Manage. Sci.*, 40(8), 1994.

[9] S. C. Geyik, S. Faleev, J. Shen, S. O'Donnell, and S. Kolay. Joint optimization of multiple performance metrics in online video advertising. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016.

[10] I.-H. Hann and C. Terwiesch. Measuring the frictional costs of online transactions: The case of a name-your-own-price channel. *Management Science*, 49(11), 2003.

[11] M. U. Kalwani and C. K. Yim. Consumer price and promotion expectations: An experimental study. *Journal of marketing Research*, 1992.

[12] M. N. Katehakis and A. F. V. Jr. The multi-armed bandit problem: Decomposition and computation. *Math. Oper. Res.*, 12(2), 1987.

[13] F. Kooti, K. Lerman, L. M. Aiello, M. Grbovic, N. Djuric, and V. Radosavljevic. Portrait of an online shopper: Understanding and predicting consumer behavior. In *Proceedings of the 9th ACM International Conference on Web Search and Data Mining*, 2016.

[14] T. L. Lai and H. Robbins. Asymptotically efficient adaptive allocation rules. *Advances in applied mathematics*, 6(1), 1985.

[15] C.-C. Lin, K.-T. Chuang, W. C.-H. Wu, and M.-S. Chen. Combining powers of two predictors in optimizing real-time bidding strategy under constrained budget. In *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management CIKM*, 2016.

[16] E. J. McCarthy. *Basic marketing: A managerial approach.* RD Irwin, 1978.

[17] C. F. Mela, S. Gupta, and D. R. Lehmann. The long-term impact of promotion and advertising on consumer brand choice. *Journal of Marketing research*, 1997.

[18] D. F. Otero and R. Akhavan-Tabatabaei. A stochastic dynamic pricing model for the multiclass problems in the airline industry. *European Journal of Operational Research*, 242(1), 2015.

[19] L. Rokach and O. Maimon. *Data Mining with Decision Trees: Theroy and Applications.* World Scientific Publishing Co., Inc., River Edge, NJ, USA, 2008.

[20] R. Schlosser, M. Boissier, A. Schober, and M. Uflacker. How to survive dynamic pricing competition in e-commerce. In *Proceedings of the 10th ACM Conference on Recommender Systems RecSys*, 2016.

[21] B. W. Silverman. *Density estimation for statistics and data analysis*, volume 26. CRC press, 1986.

[22] Y. Singer and M. Mittal. Pricing mechanisms for crowdsourcing markets. In *22nd International World Wide Web Conference, WWW*, 2013.

[23] S. Sun, R. Law, M. Schuckert, and L. H. N. Fong. An investigation of hotel room reservation: What are the diverse pricing strategies among competing hotels? In *Information and Communication Technologies in Tourism 2015.* 2015.

[24] K. T. Talluri and G. J. Van Ryzin. *The theory and practice of revenue management*, volume 68. Springer Science & Business Media, 2006.

[25] S. B. Thrun. Efficient exploration in reinforcement learning. Technical Report CMU-CS-92-102, Carnegie Mellon University, 1992.

[26] H. R. Varian. *Microeconomic analysis.* Norton & Company, 1992.

[27] J. Wang and S. Yuan. Real-time bidding: A new frontier of computational advertising research. In *Proceedings of the 8th ACM International Conference on Web Search and Data Mining*, 2015.

[28] W. C. Wu, M. Yeh, and M. Chen. Predicting winning price in real time bidding with censored data. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2015.

[29] S. Yuan, J. Wang, and X. Zhao. Real-time bidding for online advertising: Measurement and analysis. *CoRR*, 2013.

[30] W. Zhang, Y. Rong, J. Wang, T. Zhu, and X. Wang. Feedback control of real-time display advertising. In *Proceedings of the 9th ACM International Conference on Web Search and Data Mining*, 2016.

[31] W. Zhang, S. Yuan, and J. Wang. Optimal real-time bidding for display advertising. In *The 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2014.

[32] W. Zhang, T. Zhou, J. Wang, and J. Xu. Bid-aware gradient descent for unbiased learning with censored data in display advertising. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016.

[33] Q. Zhao, Y. Zhang, D. Friedman, and F. Tan. E-commerce recommendation with personalized promotion. In *Proceedings of the 9th ACM Conference on Recommender Systems RecSys*, 2015.