Game-Theoretic Cross Social Media Analytic: How Yelp Ratings Affect Deal Selection on Groupon?

Chih-Yu Wang, *Member, IEEE*, Yan Chen, *Senior Member, IEEE*, and K. J. Ray Liu, *Fellow, IEEE*

Abstract—Deal selection on Groupon is a typical social learning and decision making process, where the quality of a deal is usually unknown to the customers. The customers must acquire this knowledge through social learning from other social medias such as reviews on Yelp. Additionally, the quality of a deal depends on both the state of the vendor and decisions of other customers on Groupon. How social learning and network externality affect the decisions of customers in deal selection on Groupon is our main interest. We develop a data-driven game-theoretic framework to understand the rational deal selection behaviors cross social medias. The sufficient condition of the Nash equilibrium is identified. A value-iteration algorithm is proposed to find the optimal deal selection strategy. We conduct a year-long experiment to trace the competitions among deals on Groupon and the corresponding Yelp ratings. We utilize the dataset to analyze the deal selection game with realistic settings. Finally, the performance of the proposed social learning framework is evaluated with real data. The results suggest that customers do make decisions in a rational way instead of following naive strategies, and there is still room to improve their decisions with assistance from the proposed framework.

Index Terms—Social learning, game theory, network externality, social media

1 Introduction

Deal selection on Groupon is a complex learning and decision making process. Groupon illustrates a new possibility of e-commerce business model [1]. As shown in Fig. 1, it offers small businesses, especially local restaurants, a platform to promote their products with significant discounted deals. These deals are mostly effective only in a limited time or even in a limited amount for advertising purpose. Customers who purchase deals on Groupon may also promote the deal to other social networks like Facebook or Twitter through the built-in tools provided by Groupon. What attracts the customers to buy certain deals most is the possibility that they can purchase high quality products or services with a bargain price. Nevertheless, the quality of a deal is usually unknown to the customers at the first sight. The information provided on Groupon alone is limited and potentially biased. It usually requires further efforts to learn the knowledge from external sources.

- C.-Y. Wang is with the Research Center for Information Technology Innovation, Academia Sinica, Taiwan and the Institute of Information Science, Academia Sinica, Taipei City 11529, Taiwan. E-mail: cywang@citi.sinica.edu.tw.
- Y. Chen is with the School of Information and Communication Engineering and the Center for Future Media, University of Electronic Science and Technology of China, Chengdu, Sichuan 611731, China. E-mail: eecyan@uestc.edu.cn.
- K.J.R. Liu is with the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742. E-mail: kjrliu@umd.edu.

Manuscript received 22 Jan. 2016; revised 18 Nov. 2017; accepted 24 Nov. 2017. Date of publication 4 Dec. 2017; date of current version 30 Mar. 2018. (Corresponding author: Chih-Yu Wang.)
Recommended for acceptance by X. Lin.
For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TKDE.2017.2779494

The potential customers may acquire this knowledge through learning from other information sources, such as the experiences shared by their friends on Twitter [2], rating on Groupon given by previous customers, or reviews on third-party websites like Yelp (Fig. 1). All these information sources are regarded as social medias where the information is generated by typical social entities and delivered through online social networks and/or Internet. The information on different social medias could be correlated. For instance, one may find the rating of a Groupon deal provider in Yelp, and one may find the popularity of a vendor in Yelp by checking the sold quantity of the deal provided by the vendor on Groupon. In the deal selection problem, potential customers may survey the reviews and ratings of certain deals on the corresponding vendor records on Yelp, and construct their own estimation on the quality of deal. Then, based on their own estimation on the deal quality, price, and other factors, customers purchase the deals that fit their needs most.

Notably, the quality of a deal may not only depend on the state of the vendor but also the decisions of potential customers on Groupon, that is, the network externality. The network externality, which describes the mutual influence among agents, plays an important role in numerous network-related applications [3], [4], [5]. When the network externality is negative, i.e., the more agents make the same decision, the lower utilities they have in the network, agents tend to avoid making the same decision with others in order to maximize their utilities. When the externality is positive, agents tend to make the same decision to increase their utilities. This phenomenon has been observed in many applications in various research areas, such as storage service

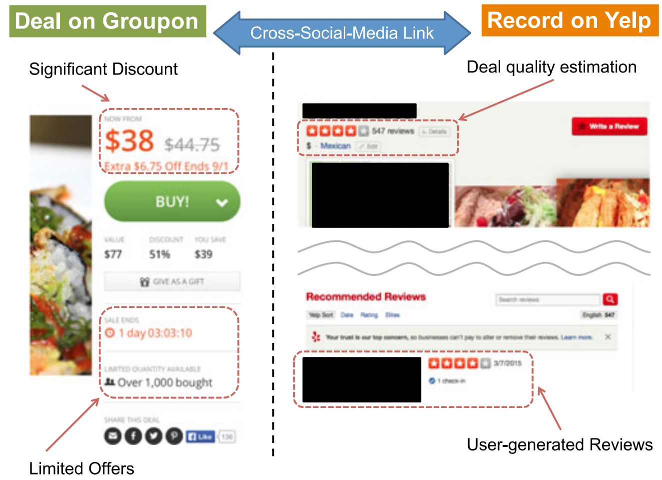


Fig. 1. Deals on Groupon versus record on Yelp.

selection in cloud computing [6] and deal selection on Groupon [1]. It has been observed that when one product or service has been successfully promoted on Groupon, it is likely to receive negative responses and low reputation due to degraded qualities of services or over-expectation on the products [7], [8]. For example, a local restaurant may provide a 50 percent-off deal for advertising purpose. However, this deal may be sold in thousands, which means that the restaurant needs to serve a huge number of customers in months with little profit. In such a case, the quality of meal and service will be degraded due to limited resources and incentives, otherwise the restaurant may not be able to survive [9]. This effect is contributed by the decisions of other customers. A rational agent should not only estimate the quality of the deal but also predict this effect through the information learned with social learning. How these two effects, network externality and social learning, affect the decisions of customers in deal selection on Groupon, with information from external social medias such as Yelp, is the main interest of this paper.

1.1 Complex Social Learning and Chinese Restaurant Game

Social learning is a typical solution for the agents to expand their knowledge in the network. For instance, we consider an agent seeking for the best juice from thousands of choices in a market. The agent may learn from advertisements provided by some brands, his own experience from previous purchases, or the real experience such as reviews or discussions shared in social medias. All the information can help the agent to construct a belief on the quality of the juice, due to which the accuracy of the agent's decision can be greatly enhanced. The last information source, which is the experience shared in social medias, is the one that receives significant attentions recently in researches on online social network, social media, and data science. The information of other agents is shared through the links constructed by the social relation in social medias. Since each agent may have different social relations with others and make decision at different time, the information one agent received may be different from others. The process of an agent learning from such information is called Social Learning [10], [11], [12], [13].

Besides social learning, agents also need to predict other agents' decisions since their rewards generally are determined by not only their own decisions but also others'. A well-known example is the choice of cellular service provider. A user with more friends using the same service provider is more likely to choose the same service provider in order to enjoy lower in-network call fee. Both behaviors, learning and predicting, can be concurrently observed in various social medias and play important roles in the deal selection problem.

Chinese Restaurant Game [14], [15] is a general framework for modeling strategic learning and decision processes in the social learning problem with negative network externality. In a Chinese restaurant game, there is a Chinese restaurant with multiple tables and customers, where customers request seats from these tables for having meals. Customers prefer bigger space for a comfortable dining experience. Thus, one may be delighted if he has a bigger table. However, when multiple customers request for the same table, they need to share the table. In such a case, the customer's dining experience is impaired due to the dining space reduction. How customers rationally choose the tables to enhance their own dining experiences in the restaurant is the main focus of Chinese restaurant game. In sequential Chinese restaurant game, the sizes of tables are unknown to the customers but each customer receives some signals related to the table sizes. Then, the customers make their requests sequentially, while the latter ones know the decisions and signals of the early ones. Under this scenario, the game additionally involves the social learning effects since a rational customer will learn the unknown table sizes from the observed actions and signals. By jointly considering both network externality and social learning effects, Chinese restaurant game provides a general framework for analyzing how agents make rational decisions under these effects.

The original Chinese Restaurant Game focuses on snapshot-based scenarios, that is, the system state and customers are assumed given and fixed in the process of the game. Dynamic Chinese Restaurant Game (D-CRG) is proposed to study how a user in a dynamic social network learns the uncertain state and makes optimal decision [16]. A user should not only consider the immediate utility he received at a given time, but also the influence from the subsequent users' decisions and the transitions of the system state. In D-CRG, Multi-dimensional Markov Decision Process model is proposed to describe the decision process in the dynamic system. The system is similar to traditional Markov decision process except that there exist multiple reward functions and transition probability matrices. It is shown that the system could reach social welfare optimal when the service provider imposed a proper pricing strategy to regulate these rational users. Dynamic Chinese Restaurant Game has been successfully applied in cognitive radio networks [16], wireless network access [17], and SVC multicasting system [18].

1.2 Contributions

Nevertheless, D-CRG still relies on few assumptions which could be unrealistic for real world scenarios and make it impractical to the deal selection problem we consider here. One critical assumption is that the available choices of tables remain unchanged over time. This could be impractical since social medias such as Groupon are highly dynamic and stochastic. The available choices for agents change frequently in such a system. The deals offered on Groupon, for instance, change even hourly. Each customer may observe different set of deals on sale and therefore face different deal selection problems. The proposed model should also address this phenomenon by modeling the dynamics in available choices in order to reflect the dynamics in Groupon. Additionally, the externality considered in D-CRG is pure negative, which is not general enough to cover real cases in real world. In Groupon, according to our observations in the collected dataset, the externaltiy is more general: it could be positive for low-quality deals but negative for high-quality deals. In sum, an extension to D-CRG, based on characteristics we found in real data, is necessary here.

In this paper, we aim to develop a data-driven social learning model to understand the deal selection behaviors in real world with data from two specific social medias: Groupon and Yelp. Specifically, a stochastic learning model based on Chinese Restaurant Game is proposed to understand how rational customers select the deal based on their knowledge from collected external reviews with concerns on the externality caused by other customers. The model shares a similar structure with D-CRG with additional supports on dynamic available deal sets, stochastic review generating process, and general externality in utility functions.

We develop a set of social media data collection tools to construct the required dataset of targeting social medias. A year-long experiment is conducted to trace the competitions among deals on Groupon and the influences on the corresponding rating in Yelp records. We utilize the dataset to extract the required information for the proposed stochastic learning model, such as regressions of arrival process, departure process, and utility functions. Based on the learning model and dataset, we analyze the deal selection game on a deal website with realistic settings. The sufficient condition of the Nash equilibrium in the deal selection game is identified. A value-iteration algorithm is proposed to find the optimal deal selection strategy. Furthermore, the performance of the proposed social learning framework is evaluated with real data in terms of social welfare and customer utility. A further discussion on the rationality of customers in deal selection by comparing the results from simulations with real data is provided. Finally, we draw our conclusions.

2 Related Works

The information cascade in a social network is a popular topic in social learning. Users in a social network usually do not have a global knowledge on the system state, but may receive some local and noisy signals about it. These users may exchange their information or observe others' revealed signal or actions in order to gain a consensus on the state/opinion/decision. Information cascade happens when some locally revealed signals, even only consists of a small portions of total signals, are strong enough to dominate any of other signals which are not revealed yet. In such a scenario, the global consensus will be determined by few signals revealed, which is likely to be a biased one since a large portion of signals are ignored. Social learning studies how agents reach the consensus (or not) through social learning in the sequential decision process. A significant part of existing works [10], [12], [13], [19] relies on the assumption that there is no network externality, i.e., the actions of subsequent agents do not change the payoff of the former agents. In such a case, agents will make their decisions purely based on their own believes without considering the actions of subsequent agents.

For the case that externality does play a role in the payoff of the users, the challenge of information cascade is more significant since users are more likely to be influenced by others in their decisions. An extreme case is that negative externality effect is so strong that all users would completely give up the potential choices and even threaten others not to choose it. An application is ad-free website: all companies may pay the website for not publishing any advertisements from its rivals [20].

One important research direction in this area is to determine the scheduling, that is, the sequence to trigger the nodes to make decisions, in order to avoid undesired information cascade [21]. An updated study suggests that an opinion which is expected-to-be-popular is usually difficult to be reversed in information cascade, even if the scheduling can be determined in advance [22]. Most related works in this field consider myopic users, which follow a simple myopic strategy without considering the expected future outcome and externality from subsequent users. When the strategic users are considered, that is, the users are now aware of the expected outcome in the final stage, the outcome of the network could be either better or worse than the myopic user case, depending on the graph structure [23]. For the case of dynamic state, it requires further efforts to guarantee that the network will reach the correct consensus. How the users cooperate with each others to actively trace the changing state in an incomplete graph is discussed in [24]. For the case that the graph is difficult to maintain the correct consensus, they discuss how the error can be reduced by adding new link or relationship into the graph. There are also different approaches to information cascade in machine learning. For instance, active learning can also be applied in the system from a global viewpoint. A trainer first define an objective function or goals of all users. Users in the network cooperatively pass their information to maximize the objective function. A practical system of this approach is intruder detection in sensor network, where sensors in an incomplete graph are required to reach a consensus about whether an intruder is detected or not [25].

It is still very challenging to understand how rational human beings in real world make decisions or respond to external issues. Game theory provides a well-established mathematical framework but usually is too complex for capturing all important characteristics featured in real world. Machine learning can help capture the hidden relations between strategic decisions and inputs using relatively simpler decision making models through training. The received models can help verify whether human beings do behave in a rational way as we assumed in game theory. Additionally, the trained models can also feedback into game-theoretic models, such as a refinement to the original utility function. Some empirical studies has been conducted in this area. It has been shown that human being may behave more cooperatively when they are in a repeated game and have previous experiences in similar scenarios. Nevertheless, they

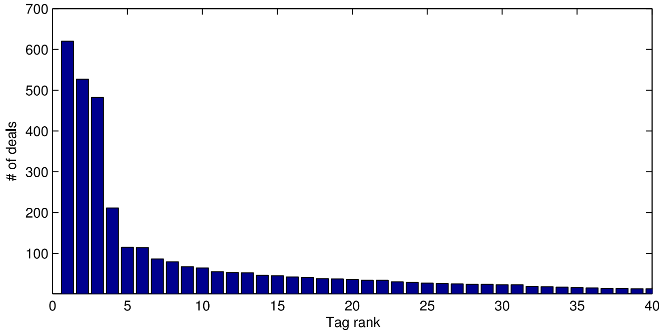


Fig. 2. Tag distribution.

may also behave selfishly if they have been betrayed in their previous experiences [26]. It also has been shown that normal people have very limited memory and cannot do complex strategy planning. Such limitations lead to suboptimal decisions [27]. The limited memory is also confirmed by another experiment where they observe that humans tend to ignore old signals in decision making [28]. Nevertheless, when it comes to a large-scale system such as Taobao shopping system, the accuracy to predict the decisions made by real human is still low if only traditional machine learning algorithms are applied [29].

3 Cross Social Media Dataset

We first introduce the constructed Groupon and Yelp dataset using our Python-based social network data collection and analysis toolbox. The toolbox provides several key functions:

1) Standard RESTful API calls support
2) Scheduling-support crawler
3) Cross-social network identification and matching
4) Feature extraction
5) Distribution regression

We implement the toolbox on a Linux-based machine and collect the data from Groupon and Yelp. Specifically, our target is the Groupon deals offered in Washington D.C. area and the corresponding Yelp records to each deal. For three times in each day, the Groupon crawler will first collect the deals offered on the Groupon in Washington D.C. area through Groupon API. Then, a social element matcher will identify each deal's corresponding Yelp record(s) through Yelp API. Notice that one deal may match to mulitple Yelp records when the deal can be used in several stores. When valid records are identified, the ids of the Yelp records will be stored in a database marked as targets. Another Yelp crawler will collect the data of all marked Yelp records in a fixed interval, independently. Due to the limitation of the RESTful API provided by Yelp, we can only access the latest three reviews of the records. This short-coming is overcome through high-frequency data crawling (three times a day)

The data collection process is executed for 19 months (Dec. 2012 to July. 2014). We collected 6,509 deals, where 2,389 deals have valid Yelp records. We have established a Groupon-Yelp relational dataset featuring:

- Groupon deals
  - Start/end time, expire time, tags
  - Options: Price, discount, location
  - Sold quantity tracking (3 times per day)
- Yelp records
  - Basic information (location, phone, ...)
  - Linking to Groupon deal(s)
  - Rating tracking (3 times per day)
  - Reviews: rating, content, author, time

It should be noted that unlike most deal datasets we can found in public for academic usage, our dataset provides the tracking of sold quantity of each deal. Through this feature, we have the information about the competition among deals, that is, we can answer the question like "which deal is the best-selling given a set of these deals available?". This is also the main interest of this work and our social learning framework as we plan to utilize the framework to analyze and predict the competition effects in social media.

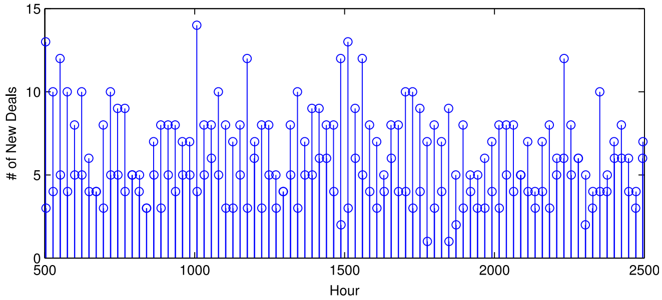
Through analyzing the dataset, we have extracted some interesting characteristics of Groupon deals. At each day, new deals get online in two batches. Specifically, we check the exact online time of each deal through the collected data since the collected data of each deal has a "startAt" tag which indicates the exact start time of the deal. The first batch contains on average 6.73 new deals, while the second batch contains only 4.72 deals (Fig. 3a). The available duration of a deal, that is, the number of days a deal is available for purchase at Groupon, is very diverse. A deal could be available for purchase within duration from few days to months. On average the duration is 9.21 days, with a standard deviation of 20.14 days. The diversity comes from the fact that there are different types of deals (marked with tags) offered in Grouopn. One may see that the distribution is closely matched to the exponential distribution as we illustrated in Fig. 4a.

The valid duration of the deal, that is, the number of days that a customer may utilize the deal after the deal goes online, also varies significantly with an average of 126.38 days and a standard deviation of 47.96 days. Notice that on average the valid duration of deals is significantly larger than the average available duration of deals.

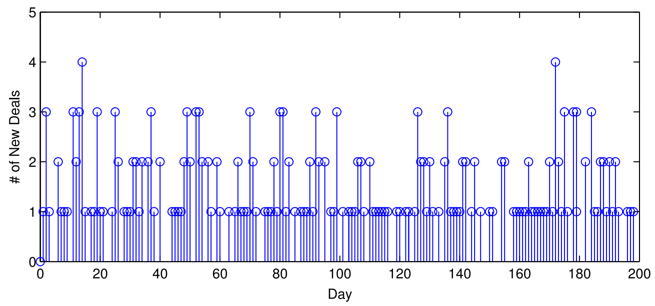
A Groupon deal may be marked with multiple tags, while the most popular tags are 1) Beauty and Spas, 2) Restaurants, and 3) Arts and Entertainment. These three tags includes more deals than all other tags combined (Fig. 2). When it comes to a specific type of deals, say, Restaurants, the arrival of new deals is much predictable. The restaurant-type deals still arrive in batches, while the interval could be 24 hours or 48 hours. On average 1.65 deals are arrives in each batch (Fig. 3b).

The average available duration of the deal is exactly 7 days, with a standard deviation of 14 days. One may see that the distribution is closely matched to the exponential distribution as we illustrated in Fig. 4b. The valid duration, on the other hand, has an average of 109.44 days and a standard deviation of 28.31 days. In the following analysis, we limited the targeting dataset to the deals tagged as Restaurants in order to remove the heterogeneity among different type of deals.

The quantity of deals sold at each day is shown in Figs. 5a and 5b. On average, 36.35 deals are sold for all deals and 13.66 deals are sold for restaurant type deals per hour. Here we assume that each deal is purchased and utilized by



(a) Deal Arrival (All deals)



(b) Deal Arrival (Restaurant)

Fig. 3. Deal arrival process.

exactly one user. The purchase record then represents the effective user (users who actually purchase a deal) arrival distribution. Notice that the distributions again closely match to the exponential distribution, as we illustrated in Figs. 5a and 5b.

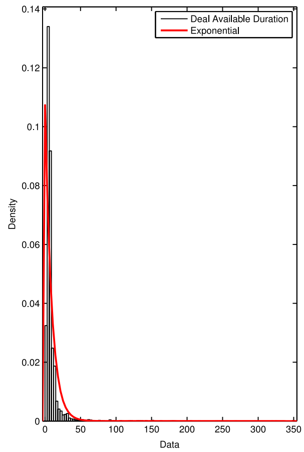
Through linking the deal records to Yelp records, we have more observations on the cross-effect of Yelp rating and Groupon deal sold quantity. First, there are significantly more deals offered by vendors with 4-star rating in Yelp, while no deals are offered by vendors rated by 1 or 2 stars. Additionally, lots of deals are sold at most 1,000 quantities, which suggests that this is a common capacity for a typical vendor to afford (Fig. 6a). We define the new reviews as all reviews posted within the duration between the days a deal becomes available and expires. Through tracking the new reviews generated by customers after the deal is online, we observed that there is a significant but complex relation between the sold quantity and rating distribution of new reviews. The relation, which can be explained with network externality, is positive when the original Yelp rating of the vendor is 3-star, but negative when the original rating is 5-star. The results suggest that the externality effects are correlated to the original rating of the vendor. A simple negative or positive assumption cannot properly capture the complex externality effect. We therefore use the curve fitting toolbox in Matlab using Polynomial model with degree of 5. The data has been smoothed using moving average method (Fig. 6b).[1]

4 System Model

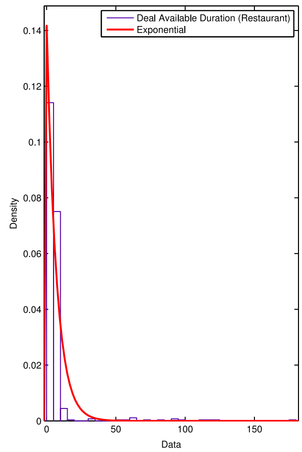
We now formulate the system model based on our observations from the Groupon—Yelp dataset. We consider a deal website providing multiple deals denoted by the deal set $\mathcal{D}[t] = \{d_1, d_2, \ldots\} \subset \mathcal{D}^{all}$, where $\mathcal{D}^{all}$ is the universal deal set and $\mathcal{D}[t]$ is the available deal set at time t. We assume that the system time t is discrete. There is a one-to-one relation between vendor and offered deal: a vendor offers exactly one deal at a time, which belongs to the universal deal set $\mathcal{D}^{all}$. All deals shown up in the online deal set $\mathcal{D}[t]$ are drawn from the universal set. New deals arrive and are added into the deal set $\mathcal{D}[t]$ following a Poisson arrival distribution with parameter $\bar{\lambda}$. Specifically, the opportunity for all offline deals to get online comes following the Poisson arrival process, and each deal may go online independently in this batch with a probability ρ_d if it is offline in previous slot. Each deal goes offline and is removed from the deal set $\mathcal{D}[t]$ following an exponential distribution with parameter $\bar{\mu}$. Each deal d has a price of p_d, which is controlled by the vendor. The quality of the deal d is r_d, which is unknown to the customers.

Customers arrive the website stochastically following a exponential arrival process with parameter $\bar{\lambda}_u$. One customer will purchase exactly one deal when he arrives at time t and see the available deals in $\mathcal{D}[t]$. The customer may utilize the deal when the deal goes offline. This reflects our observation from the dataset that the deal valid duration is significantly

1. The dataset will be published using privacy preserving data publishing (PPDP) process in respect to our agreement with the websites.

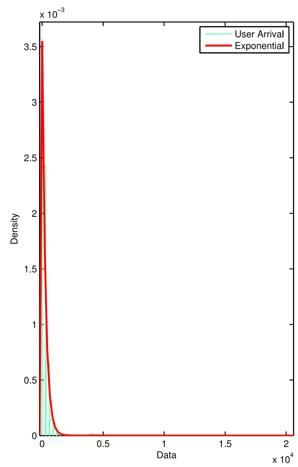


(a) Deal Available Duration (All deals)

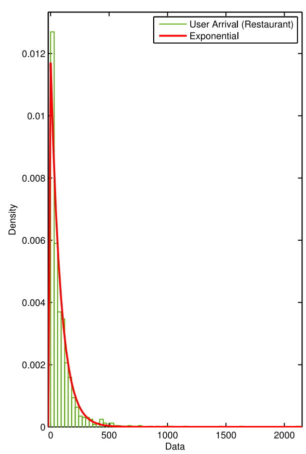


(b) Deal Available Duration (Restaurant)

Fig. 4. Deal available duration.

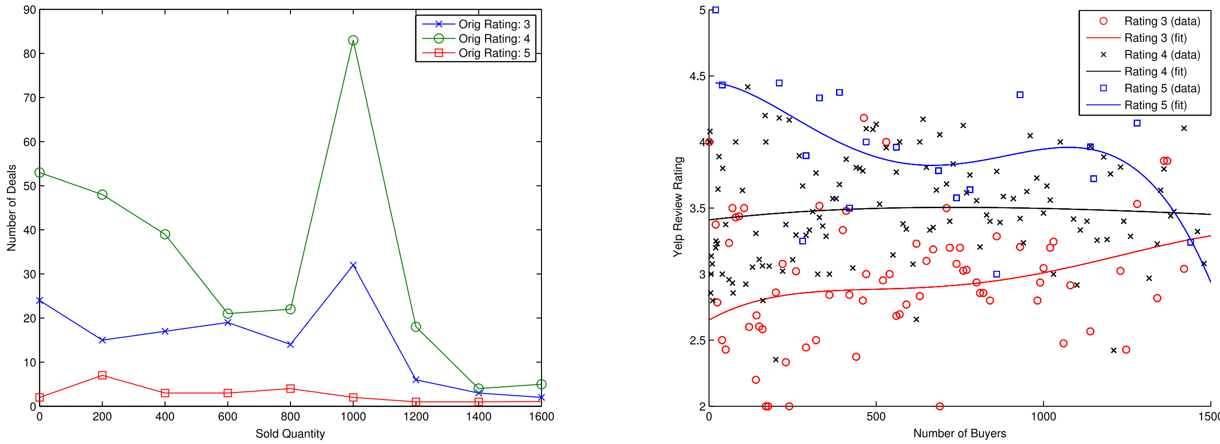


(a) Sold Quantity per Hour (All deals)



(b) Sold Quantity per Hour (Restaurant)

Fig. 5. User arrival.



(a) Deal's corresponding Yelp original rating (b) New review rating distribution

Fig. 6. Yelp review rating versus Groupon deal.

longer than sell duration and therefore the probability of using deal within the sell duration can be neglected.

The valuation of the deal depends on the quality of the deal, which could be known or unknown to the customer before purchase. It is also influenced by the number of deals sold before it goes offline. This externality effect could be positive or negative: it depends on the characteristic of the deal. Based on the observations from our dataset, we assume that the externality is not only a function of sold quantity but also the original quality of the deal. The utility of customer is therefore given by

$$U(r_d, n_d^*, p_d) = V(r_d, n_d^*) - \alpha p_d, \tag{1}$$

where r_d is the original rating of the deal d, n_d^* is the number of customers purchasing deal d before the deal goes offline, p_d is the price of the deal, and α is the weight of the payment. The exact form of the valuation function $V(r_d, n_d^*)$ is referred to the real data we extracted from the dataset. (Fig. 6b). Roughly speaking, when $r_d \leq 3$, the valuation of the deal increases with n_d^*. For the case that $3 < r_d \leq 4$, the valuation function is slightly concave in n_d^*. When $r_d > 4$ the valuation function is an decreasing function.

Notice first that the quality of the deal is unknown to the customers. It requires an extra effort to learn this knowledge through social learning. Additionally, the utility of a customer is realized when the deal goes offline, which happens *after* the customer chooses the deal. It is possible that the sold quantity of the deal increases before the deal goes offline and therefore has a positive or negative impact on the utility of the customer. A rational customer should not only consider the current state of the available deals but also predict the increases in sold quantity of the deals in order to estimate the expected utility if she chooses certain deals.

4.1 External Information from Social Media

Customers learn the unknown deal quality r_d from personal experiences, reviews, or rating shared on third-party social medias such as Yelp and Facebook. According to our observations from the dataset, we assume that these information arrive the system stochastically. The information are regarded as signals and become the knowledge of the customers on the unknown quality r_d. We assume that the current estimation on the expected rating of deal d at time t is denoted as $r_d[t]$, which is a random variable. New review w_d on the quality of deal d arrives following Poisson arrival process with parameter $\bar{\lambda}_{w_d}$. The value of review w_d is a random variable with conditional probability function $Pr(w_d|r_d)$, which describes the accuracy of the review in reflecting the true quality of the deal. The reviews are independent when conditioning on r_d. Since the reviews arrive stochastically, customers who arrive at different time may have observed different sets of reviews and therefore have different estimations on r_d.

Belief is the customer's estimation on the probabilistic distribution of $r_d[t]$ after collecting all available reviews. A belief on deal d at time t is denoted by $\mathbf{b}_d[t] = \{b_{d,X}[t]|X\}$, where $b_{d,X}[t] = Pr(r_d = X|\{w_d\}[t])$, $\sum_X b_{d,X}[t] = 1$. In a stochastic system, the belief on deal d can be updated when new review w'_d arrives using Bayesian update rule as follows:

$$b_{d,X}[t] = \begin{cases} \frac{Pr(\{w'_d\}|r_d=X)b_{d,X}[t-1]}{\sum_{X'=1}^{5} Pr(\{w'_d\}|r_d=X')b_{d,X'}[t-1]}, & \text{new review } w'_d. \\ b_{d,X}[t-1], & \text{else.} \end{cases} \tag{2}$$

We then denote $\mathbf{b}[t] = \{\mathbf{b}_d[t]\}$ as the common belief of all customers on all deals at time t. Nevertheless, $\mathbf{b}[t]$ is intractable when it comes to real data since its dimension increases exponentially with the number of deals. An approximation is necessary to reduce its complexity. Observing the collected Yelp data set, we find that the review ratings received by a vendor follow Gaussian distributions in general (Table 1). The variance of the received ratings depends on the average rating of the deal. Based on this observation, we assume that $b_{d,X}[t]$ follows Gaussian distribution with the average rating $b_{d,avg}[t]$ and variance $b_{d,var}$. In such a case, we only need to track $b_{d,avg}[t]$ of each deal instead of whole $\mathbf{b}_d[t]$, and the

TABLE 1
Review Rating Distribution versus Original Rating

Original	1-star	2-star	3-star	4-star	5-star
3-star	350	322	374	397	280
4-star	480	507	736	1,302	1,032
5-star	27	27	45	92	154

computation complexity is significantly reduced. Finally, we denote $\mathbf{b}_{avg}[t] = \{\mathbf{b}_{d,avg}[t]\}$ as the belief in average sense.

5 Stochastic Deal Selection Game

We propose a game-theoretic stochastic learning model for the deal selection problem. The model shares a similar structure with Dynamic Chinese Restaurant game [16] in state, profile, and belief. Nevertheless, the original model is not applicable for this problem due to the following reasons: 1) the available deals (tables) are no longer fixed but should be described by a stochastic process, 2) the reviews (signals) are not always generated in sync with the customer arrival process, and 3) the externality could be either positive, negative, and quality-dependent. A new game-theoretic stochastic model is necessary to address above new characteristics in the deal selection problem.

We consider a deal selection game with players being the customers who arrive the system stochastically. A customer may select exact one deal $d \in \mathcal{D}[t]$ when he arrives at time t. He leaves the game when the deal goes offline, and the utility is realized at this moment. The utility is given by the utility function $U(r_d, n_d^*, p_d)$ which takes the number of customers choosing the same deal (n_d^*), deal quality r_d, and the price p_d as inputs. A rational customer should seek to select the deal that maximizes his expected utility in the stochastic game. Our objective is to derive the optimal deal selection strategy of each customer in the game.

Following the system model we formulated in previous section, customers arrive the system following a Poisson arrival process with parameter $\bar{\lambda}_u$. A customer's action space is the available deals when he arrives at time t, that is, a customer's action is denoted by $a \in \mathcal{D}[t]$. One customer will not change his action in the remaining of the time until he leaves the system. Customers leave the system only after the selected deals go off-line. Therefore, their departure follows the deal departure process, which is assume to be an exponential distribution with parameter $\bar{\mu}$. The utility of the customer is realized as soon as he leaves.

5.1 Multi-Dimensional Markov Decison Process

A rational customer seeks to maximize the utility in the game. Since the system is stochastic, the exact utility of each deal cannot be obtained. The customers must estimate the expected utility he can receive from each deal before selecting the deal according to his knowledge on the system. Here we propose to apply Multi-dimensional Markov Decision Process (M-MDP) [17], which is also the foundation of D-CRG [16], to estimate the expected utility of the customers.

State in M-MDP represents the current observations and knowledge of the customers on the deals in the game. Here we denote state as

$$\mathbf{s}[t] = \{\mathcal{D}[t], \mathbf{n}[t], \mathbf{b}_{avg}[t]\} \in \mathcal{S}, \tag{3}$$

where $\mathcal{D}[t]$ is the deals available for purchase and $\mathbf{n}[t] = \{n_d[t] | d \in \mathcal{D}^{all}\}$ is the number of customers choosing each deal d at time t. We assume $n_d[t]$ is rounded to hundreds or even thousands. This is a common practice in deal websites such as Groupon, where only rounded figure is revealed publicly. The $\mathbf{b}_{avg}[t]$ is the beliefs on each deal d's quality as we mentioned in previous section. Finally, $\mathcal{S}$ is the universe set of all possible states. A rational customer makes use of the state information when he arrives in order to maximize his expected utility. Therefore, we expect the optimal deal selection strategy be related to the arrival state of the system. Notice that the state describes the knowledge of the customers on the system based on their limited observations on public information. It may not accurately reflect the hidden information of the system, such as the exact sold number and quality of the deals.

Strategy profile, which is denoted by a function $\pi(\mathbf{s}[t]) \in \mathcal{D}[t]$, describes the actions of each player in every possible state of the system. It can be viewed as the behavior predictions on the players in a given M-MDP. In the deal selection game, the strategy profile describes which deal will be selected by next customer under the current state $\mathbf{s}[t]$. Due to the effect of externality on the utility in (1), a rational customer should also be aware of the deal selection strategy of other customers in the system.

5.2 State Transition

The state of the system may transit from one to another when events occur. The transition probability depends on the viewpoint of the customer [17]. Specifically, a customer may observe the system state changes when new customers arrive and choose deals, and some deals go online or offline. Nevertheless, his utility is realized when the deal he selects goes offline. This event is different from others as when it occurs, the customer also departs from the system. Therefore, no future state transition will influence his utility. For the customers who are in the system and select the deal d, the observed state transition probability is denoted by

$$Pr(\mathbf{s}'|\mathbf{s}, \pi, d) = Pr(\{\mathbf{n}', \mathbf{b}'_{avg}, \mathcal{D}'\}|\mathbf{s}, \pi, d). \tag{4}$$

Notice that the transition probability is not only determined by the current state $\mathbf{s}$ but also by the strategy profile π. In general, the state changes when specific events occur, for which the probability of occurrence is known. In the proposed system, three independent events may change the states:

1) A new customer arrives
2) A new review w_d on a specific deal d arrives
3) A deal d' goes online or offline

5.2.1 New Customer Arrives

A new customer arrives following a Poisson distribution with parameter $\bar{\lambda}_u$. When the time interval T is sufficient small, the probability that a customer arrives can be denoted by $\lambda_u = T\bar{\lambda}_u$ [17], [30]. When a new customer arrives, only the grouping state $\mathbf{n}$ will change. Nevertheless, the new grouping state $\mathbf{n}'$ depends on the strategy profile π, that is, which deal is selected by the new customer. The state transition probability conditioning on new customer arrival event is as follows:

$$\begin{aligned} &Pr(\mathbf{n}'|\mathbf{n}, \pi, d, (\text{new customer})) \\ &= \begin{cases} \Lambda(\lambda_u), & \mathbf{n}' = \mathbf{n} + \mathbf{e}_{\pi(\mathbf{s})}; \\ 1 - \Lambda(\lambda_u), & \mathbf{n}' = \mathbf{n}; \\ 0, & \text{else.} \end{cases} \end{aligned} \tag{5}$$

where $\mathbf{e}_x$ is the unit vector in dimension x, and $\Lambda(\lambda_u)$ is the probability that (rounded) n_d increased by 1 (100 more amounts are sold, for instance) after this purchase.

5.2.2 New Review Arrives

Recalling that new review w_d on the quality of deal d arrives following Poisson arrival process with parameter $\bar{\lambda}_{w_d}$. With the assumption that the time slot is very short so at most one review will be generated within a slot, we have the review arrival probability as $\sum_{d\in\mathcal{D}_{all}} \lambda_{w_d}$ where $\lambda_{w_d} = T\bar{\lambda}_{w_d}$. The belief state of the specific deal will change with the new review. Recalling that the belief follows Gaussian distribution with a known variance, we may transform the average belief $\mathbf{b}_{avg}$ recorded in the state back to exact belief $\mathbf{b}$ and then apply Bayesian update function (2) to update the belief to $\mathbf{b}'$. The updated average belief can be transformed back to average belief $\mathbf{b}'_{avg}$ in the new state $\mathbf{s}'$. The state transition probability conditioning on this event is as follows:

$$Pr(\mathbf{b}'|\mathbf{b}, \pi, d, \text{ new review } w_{d'} \in \{1,2,3,4,5\}) = \begin{cases} \frac{\lambda_{w_{d'}} \sum_X Pr(w_{d'}|r_{d'}=X) b_{d',X}}{\sum_{d\in\mathcal{D}_{all}} \lambda_{w_d}}, & \text{updated by } w_{d'}; \\ 0, & \text{else.} \end{cases} \tag{6}$$

5.2.3 Deal Online or Offline

Finally, the state also changes when a new deal goes online or an available deal goes offline. Given the batch arrival property we observed in the dataset, we assume that the opportunity that deals go online follows the Poisson arrival process with $\bar{\lambda}$. Given the short slot time assumption, we have the deal arrival probability per slot as $\lambda = T\bar{\lambda}$. Each deal may go online independently when the opportunity comes. Specifically, let ρ_d be the probability that deal d goes online at this slot if it is offline in previous slot at this batch. The state transition probability conditioning on a set of deals δ arrives is

$$Pr(\mathbf{s}'|\mathbf{s}, \pi, d, \text{ new deal set } \delta) = \begin{cases} \prod_{d'\in\delta} \rho_{d'} \prod_{d''\in\mathcal{D}^{all}\setminus(\delta\cup\mathcal{D})} (1-\rho_{d''}), & \mathcal{D}' = \mathcal{D}\cup\delta, \mathcal{D}\cap\delta = \emptyset; \\ 0, & \text{else.} \end{cases} \tag{7}$$

This formulation is inspired by two observations from the Groupon dataset: deals arrive in batches in a periodic pattern and repeated availability of popular deals. These special patterns in the dataset reduce the complexity in the calibration of the proposed model to the dataset. Specifically, the arrival rate of batches can be estimated by measuring the average period of the batches in the dataset. The probability of deal d goes online in each batch can then be approximated by calculating the average number of times the deal d goes online in all batches in the collected dataset.

Recalling that a deal may goes offline following the deal departure process, which is assume to be an exponential distribution with parameter $\bar{\mu}$. We therefore have per slot deal departure probability as $\mu = T\bar{\mu}$. The corresponding state transition is

$$Pr(\mathbf{s}'|\mathbf{s}, \pi, d, \text{ remove deal } d') = \begin{cases} 1, & \mathcal{D}' = \mathcal{D}\setminus\{d'\}; \\ 0, & \text{else.} \end{cases} \tag{8}$$

5.2.4 State Transition Probability

Combining all the events above and the corresponding probability, we may derive the state transition probability given the current state. The state changes when any of the events above occurs. So a complete form of state transition probability will be as follows:

$$\begin{aligned} Pr(\mathbf{s}'|\mathbf{s}, \pi, d) = &(\text{Prob. of Customer Arrival from } s \text{ to } s') \\ &(\text{Prob. of Review Arrival from } s \text{ to } s') \\ &(\text{Prob. of Deal Arrival/Departure from } s \text{ to } s'). \end{aligned} \tag{9}$$

To simplify the form and improve the tractability of the model, we assume that the interval T of the discrete system time t is sufficient small that at most one event will happen at each time slot t. This assumption comes from the fact that the probability of multiple independent events occurring within the same time slot t will diminish to zero when the probability of each event decreases, according to the Taylor expansion of (9). The individual occurrence probabilities of each type of event are proportional to the interval T when T is sufficent small [17] and will decreases when the interval T decreases. With this assumption, we only need to consider the state transitions involving single event occurrence.[2]

The probability that state remains unchanged is

$$\begin{aligned} Pr(\mathbf{s}|\mathbf{s}, \pi, d) = &(1 - Pr(\text{Customer Arrival})) \\ &(1 - Pr(\text{Review Arrival})) \\ &(1 - Pr(\text{Deal Arrival/Departure})) \\ = &1 - Pr(\text{Customer Arrival}) - Pr(\text{Review Arrival}) \\ &- Pr(\text{Deal Changes}) + \delta(T), \end{aligned}$$

according to the Taylor expansion. When T is sufficient small, the $\delta(T)$ will diminish to zero due to the decrease of multiple event occurrence probability within a time slot.

In sum, the overall state transition probability observed by a customer selecting deal d is shown

$$Pr(\mathbf{s}'|\mathbf{s}, \pi, d) = \begin{cases} \lambda_u \Lambda(\lambda_u), & n'_d = n_d + 1; \\ \lambda_{w_j} \sum_X Pr(w_j|r_j = X) b_{j,X}, & \mathbf{b}'_j \text{ is updated with } w_j; \\ \lambda \prod_{d'\in\delta} \rho_{d'} \prod_{d''\in\mathcal{D}^{all}\setminus(\delta\cup\mathcal{D})} (1-\rho_{d''}), & \mathcal{D}' = \mathcal{D}\cup\delta, \mathcal{D}\cap\delta = \emptyset; \\ \mu, & \mathcal{D}' = \mathcal{D}\setminus\{d'\}, d' \neq d; \\ 1 - \lambda_u\Lambda(\lambda_u) - \lambda - \sum_{j\in\mathcal{D}_{all}} \lambda_{w_j} - |D|\mu, & \mathbf{s}' = \mathbf{s}; \\ 0, & \text{else.} \end{cases} \tag{10}$$

One may notice that the sum of the transition probability from s to $s' \in \mathcal{S}$ in (10) is $1 - \mu$. This is because we ignore the event that the deal d which is selected by the customer goes offline (The fourth cases in (10)). We will show that this formation is very useful for the following analysis.

5.3 Expected Utility and Strategy Profile

A rational customer will select the deal with highest expected utility conditioning on the state $\mathbf{s}[t^a]$ he observes at

2. Notice that this assumption can be easily relaxed without changing any insight we derived in this paper.

arrival time t^a. The expected utility is influenced by not only the initial observed state of the system but also all the future states until the deal he selected goes offline. Given a specific strategy profile $\pi(\mathbf{s})$, the state transition probability $Pr(\mathbf{s}'|\mathbf{s}, \pi, d)$ is known. Assuming that customer i enters the system at state $\mathbf{s}[t^e]$ and chooses deal d_i, his utility will be realized when deal d_i goes offline. The expected utility of customer i conditioning on the entering state $\mathbf{s}[t^e]$, which is denoted by $E[u(d_i)|\mathbf{s}[t^e]]$, is given by

$$E[u(d_i)|\mathbf{s}[t^e], \pi] = \sum_{t=t^e}^{\inf} (1-\mu)^{t-t^e} \mu E[U(r_{d_i}, n_{d_i}[t], p_{d_i})|\mathbf{s}[t], \pi]. \tag{11}$$

Notice that the exact realized state at each time slot is stochastic, while the state transition probability is described by (10) and is determined by the applied strategy profile π.

The expected utility of each deal can be derived in a closed form by Bellman equations. We denote $W(\mathbf{s}, d)$ as the expected utility a customer may receive if he selects deal d at state $\mathbf{s}$, which is the sum of expected utilities when the deal d goes offline at this moment (with probability μ) and remains online (with probability $1-\mu$), respectively. When deal d goes offline, the expected utility is calculated according to the number of customers selecting the same deal n_d, deal price p_d, and the belief on the quality of the deal $b_{d,avg}$ at the state. For the case that the deal remains online, the state may transit according to the state transition probability observed by the customer who selects this specific deal d conditioning on the fact that deal d stays online. The Bellman equation of $W(\mathbf{s}, d)$ conditioning on a given strategy profile π therefore can be given as follows:

$$\begin{aligned} W(\mathbf{s}, d) = E[u(d)|\mathbf{s}, \pi] &= \mu E[U(r_d, n_d, p_d)|\mathbf{s}] \\ &+ (1-\mu) \sum_{\mathbf{s}' \in \mathcal{S}} \frac{Pr(\mathbf{s}'|\mathbf{s}, \pi, d)}{(1-\mu)} W(\mathbf{s}', d). \end{aligned} \tag{12}$$

The first part of (12) is the utility a customer will receive when the deal d he selects goes offline. The second part of (12) shows the expected utility in the future, if the deal d doesn't go offline in next slot. This part is conditioning on the fact that deal d doesn't go offline so conditional probability is applied. Notice that the state transition probability $Pr(\mathbf{s}'|\mathbf{s}, \pi, d)$ we introduced in (10) already remove the event that deal d goes online therefore can be directly applied here after normalization.

5.4 Nash Equilibrium and Value-Iterative Algorithm

We then analyze the necessary and sufficient condition of Nash equilibrium in the proposed deal selection game. Nash equilibrium represents the rational outcome of a game, in which each players has applied their optimal strategy in responses to the strategies applied by other players. Mathematically speaking, a Nash equilibrium is a strategy profile π^* defined as follows:

Definition 1 (Nash Equilibrium). *A strategy profile π^* is a Nash equilibrium if and only if*

$$E[u(\pi^*(\mathbf{s}))|\mathbf{s}, \pi^*] \geq E[u(d')|\mathbf{s}, \pi^*], \quad \forall d' \in \mathcal{D}, \forall \mathbf{s} \in \mathcal{S}.$$

On the other hand, the best response for each customer who arrives at certain state s should be based on the expected utility of each deal at that state, which is given by $W(\mathbf{s}, d)$. All customers should select the deal which gives them the highest utility from all available deals. Therefore, the optimal strategy profile π^* given the currently applied strategy profile π is given as follows:

$$\pi^*(\mathbf{s}[t]) \in \arg\max_{d \in \mathcal{D}[t]} E[u(d|\mathbf{s}[t], \pi) = \max_{d \in \mathcal{D}[t]} W(\mathbf{s}[t], d). \tag{13}$$

Clearly, the Nash equilibrium of the deal selection game is achieved when expected utility (12) and optimal strategy profile (13) match with each other, which are denoted as equilibrium conditions.

Lemma 1 (Equilibrium Condition). *The strategy profile π^* is Nash equilibrium in the deal selection game if and only if*

$$\begin{aligned} W^*(\mathbf{s}, d) &= \mu E[U(X, n_d, p_d)|\mathbf{s}] \\ &+ (1-\mu) \sum_{\mathbf{s}'} \in \mathcal{S} \frac{Pr(\mathbf{s}'|\mathbf{s}, \pi^*, d)}{1-\mu} W^*(\mathbf{s}', d). \end{aligned} \tag{14}$$

$$\pi^*(\mathbf{s}) \in \arg\max_{d \in \mathcal{D}} W^*(\mathbf{s}, d). \tag{15}$$

Proof. For a customer who arrives at state $\mathbf{s}$ assuming all other customers follow the strategy profile π^*, his expected utility if he chooses deal d is given by $W^*(\mathbf{s}, d)$ described in the lemma and from (12). According to (15), we have

$$W^*(\mathbf{s}, \pi^*(s)) \geq W^*(\mathbf{s}, d), \forall d \in \mathcal{D}.$$

Thus, Definition 1 is satisfied, and π^* is a Nash equilibrium. □

When the equilibrium conditions are satisfied, no customer has the incentive to deviate from the strategy profile π since all customers maximize their expected utilities, which are derived conditioning on π. The equilibrium conditions represent the necessary conditions of the rational outcome of the deal selection game. We may identify and predict the outcome of the game through checking the equilibrium conditions iteratively. Nevertheless, the fixed-strategy Nash equilibrium may not always exist due to the complex interactions among customers in the D-CRG model [17]. In light of this, we propose to derive $\epsilon-$ optimal Nash equilibrium, a relaxed version of Nash equilibrium defined as follows:

Definition 2 ($\epsilon-$ optimal Nash Equilibrium). *A strategy profile π^* is a $\epsilon-$ optimal Nash equilibrium if and only if*

$$E[u(\pi^*(\mathbf{s}))|\mathbf{s}, \pi^*] \geq E[u(d')|\mathbf{s}, \pi^*] - \epsilon, \ \forall d' \in \mathcal{D}, \forall \mathbf{s} \in \mathcal{S}.$$

The $\epsilon-$ optimal Nash equilibrium of deal selection problem can be found through multi-dimensional version of value iteration algorithm [18], which is illustrated in Algorithm 1:

TABLE 2
Simulation Settings

Parameter	Value
Maximum # of deals	4
Maximum available deals	2
Maximum quantity of each deal	1,500
Average Customer Arrival Rate	13.66 per hour
Average Deal Online Duration	168 hours
Price factor α	0.05

Algorithm 1. Multi-Dimensional Value Iteration

1: Initialize π^o, W^o;
2: **while** 1 **do**
3:   **for all s**, d **do**
4:     $\pi^n \leftarrow (13)$;
5:     $W^n \leftarrow (12)$;
6:   **end for**
7:   $W^d \leftarrow W^n - W^o$
8:   **if** $\max W^d - \min W^d < \epsilon$ **then**
9:     Break
10:   **else**
11:     $W^o \leftarrow W^n$
12:   **end if**
13: **end while**
14: Output π^n and W^n

Basically, Algorithm 1 is an improved value-iteration algorithm for the proposed CRG. The π^o and W^o are the original policy and expected rewards before the update process in Step 4 and 5, and π^n and W^o are the updated ones. Specifically, Step 4 and 5 in Algorithm 1 use (13) and (12) to update the policy and expected rewards. The choice of utility function determines the exact form of (13), while (13) determines the exact from of (12). W^d is the difference in expected rewards before and after the updates. When the difference in expected reward is less than the threshold ϵ, the algorithm terminates.

Theorem 1. *when Algorithm 1 terminates at round n, the ϵ-Nash equilibrium of the proposed deal selection game is given by π^n.*

Proof. When Algorithm 1 terminates at round n, we have $W^d = W^n - W^o$ and $\max W^d - \min W^d < \epsilon$. We also have

$$\pi^n(\mathbf{s}) = \arg\max_{d\in\mathcal{D}} W^o(\mathbf{s}, d).$$

Given that $\max W^d - \min W^d < \epsilon$, we have

$$W^n(\mathbf{s}, \pi^n(s)) \geq W^n(\mathbf{s}, d) - \epsilon, \forall d \in \mathcal{D},$$

which satisfies Definition 2. Therefore, π^n is an $\epsilon-$ optimal Nash equilibrium. □

It is still possible to have no ϵ-optimal Nash Equilibrium policy exists when ϵ is small. This phenomenon is due to the competitive nature of deal selection in this problem. A larger ϵ can guarantee the existence of equilibrium policy but may result in instability since some choices are suboptimal for certain customers. A smaller ϵ also suggests a longer convergence time. It depends on the choices of system administrator according to his priority on stability or convergences.

6 Simulation Results

We first evaluate the performance of proposed learning model with simulations using Groupon and Yelp dataset we constructed in Section 3. Specifically, all deals tagged as "Restaurants" and the corresponding Yelp records in our dataset are used to extract the necessary parameters for the simulations, including user arrival rate, deal online distribution, and the utility function. We simulate a deal website with 4 potential vendors. The price of the deals are \$10, \$10, \$60, and \$60 respectively. The maximal sold quantity of each deal is 1,500, while the sold quantity revealed to customers is rounded by 250. The rating of each deal are randomly drawn from $[3, 4, 5]$. The review w_d on each deal d is generated following the above probability density function

$$Pr(w_d|r_d) = \begin{cases} p, & w_d = r_d; \\ \frac{1-p}{4}, & \text{else}, \end{cases} \tag{16}$$

where p is the accuracy of the review. A higher p means more likely the review represents real rating of the deal.

The utility function $U(\cdot)$ we applied in the simulations is as follows:

$$U(d) = U(r_d, n_d^*, p_d) = f(r_d, n_d^*) - \alpha * p_d, \tag{17}$$

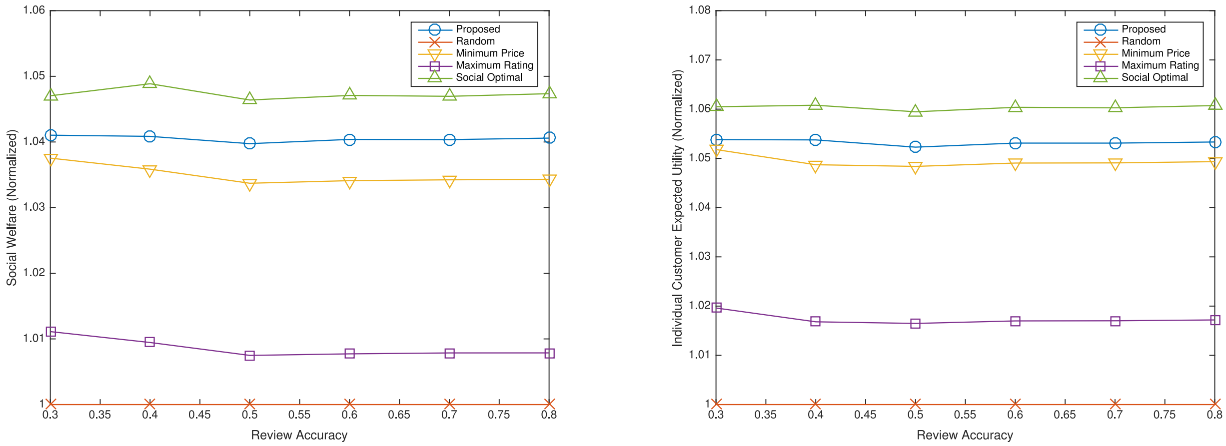
where $f(r_d, n_d^*)$ is the expected rating when n_d^* quantity of deal d is sold, and α is the adjustable price factor. The exact function of $f(\cdot)$ is regressed from the Yelp dataset we introduced in Section 3. All other parameters, such as deal arrival process, departure process, and customer arrival process, are also regressed from the dataset by calculating the average and variance of the targeting process. The detailed parameter settings are listed in Table 2.

In all simulations, we compare the proposed deal selection strategy with five strategies: Random, Minimum Price, Maximum Rating strategies, and Social Optimal strategy. The first three strategies are naive and are treated as baselines. Customers applying random strategy randomly select one deal from available deals when they arrive, regardless of the states. For Minimum Price and Maximum Rating strategies, customers always select the deal with minimum price and maximum rating, respectively

$$\pi^{maximum\_rating}(\mathbf{s}[t]) = \arg\max_{d\in\mathcal{D}[t]} E[r_d|\mathbf{s}[t])$$
$$\pi^{minimum\_price}(\mathbf{s}[t]) = \arg\min_{d\in\mathcal{D}[t]} p_d.$$

The Social Optimal strategy denotes the optimal strategy that maximizes the overall social welfare, which is defined as the sum of all customer's expected utilities. The social optimal strategy can be derived by solving the following Bellman equation and strategy profile equation sets

$$\begin{aligned} W^{social}(\mathbf{s}) = & \sum_{d\in\mathcal{D}} \mu E[U(X, n_d, p_d)|\mathbf{s}] \\ & + (1-\mu) \sum_{\mathbf{s}'\in\mathcal{S}} \frac{Pr(\mathbf{s}'|\mathbf{s}, \pi^{social})}{(1-\mu)} W^{social}(\mathbf{s}'). \end{aligned} \tag{18}$$



(a) Social Welfare (b) Individual Utility

Fig. 7. Influence of review accuracy.

$$\pi^{social}(\mathbf{s}[t]) \in \arg\max_{d\in\mathcal{D}[t]} Pr(\mathbf{s}'|\mathbf{s},\pi^{social})W^{social}|_{n'_d=n_d+1}. \quad (19)$$

Such a strategy can be derived through traditional MDP algorithms with social welfare at each state as reward [18]. This strategy represents the optimal strategy preferred by the society that the sum of utilities of all customers is maximized. This strategy relies on the assumption that all customers will follow the strategy regardless whether they will be sacrificed in order to achieve a better social utility. This means that some customers may choose not to follow the social optimal strategy when they are selfish. It takes some external efforts, such as penalty, to implement this strategy. Therefore, we will show the performance of both strategies. The performance gap between Nash equilibrium and social optimal strategies is the price of anarchy, or the performance degradation due to the selfish players in the system.

6.1 Review Accuracy

We evaluate the social welfare of the system under different deal selection strategies. We adjust the review accuracy from 0.2 to 0.9 in the simulations, and the results are shown in Fig. 7. Notice that the values of each strategy are normalized to the random strategy.

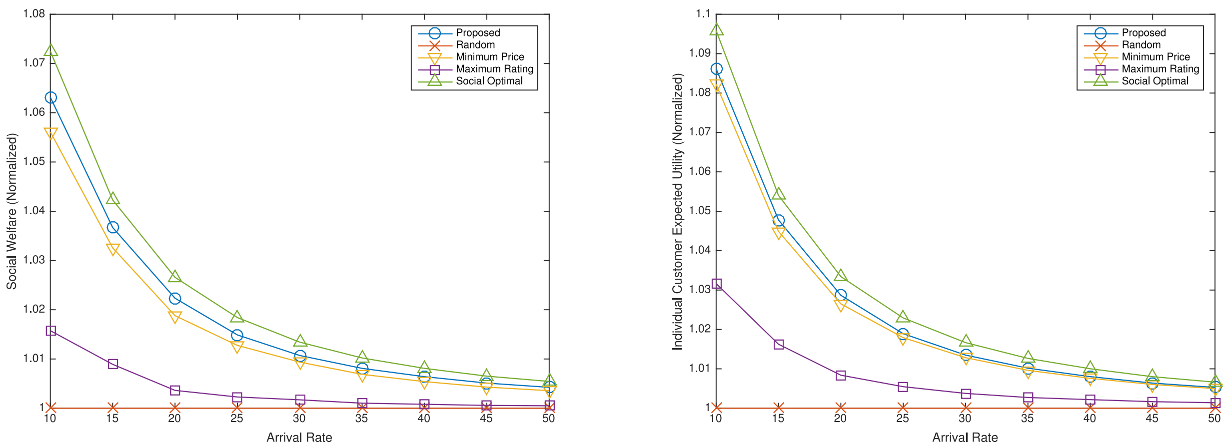
We observed that in Fig. 7a the proposed strategy significantly outperforms all other naive strategies including Random, Minimum Price, and Maximum Rating. The reason behind this enhancement is that proposed strategy further considers the externality effect and the potential state transition. Nevertheless, there exists a minor welfare degradation in the proposed strategy comparing with the Social Optimal strategy, which is the price of anarchy due to the non-cooperative nature in the deal selection game.

We also evaluate the performance of all strategies in terms of individual customer utility with results shown in Fig. 7b. The trend is almost the same as the social welfare, in which the proposed strategy significantly outperforms all naive strategies. Notice that customers clearly benefit from the fully-rational strategy provided by the proposed CRG.

6.2 Arrival Rate

We then evaluate the social welfare of the system under different arrival rates. The arrival rate is adjusted from 10 to 50 customers per hours. The review accuracy is set to 0.8 in this simulation. The results are shown in Fig. 8. Notice that the values of each strategy again are normalized to the random strategy.

The trend in Fig. 8 shows that the improvement from different strategy over the random strategy decreases with the



(a) Social Welfare (b) Individual Utility

Fig. 8. Influence of arrival rate.

TABLE 3
Deal Selection Prediction

Strategy	Accuracy
Random	0.2777
Maximum Rating	0.2789
Minimum Price	0.3147
Social Optimal	0.3240
Proposed (Myopic)	0.2867
Proposed (Fully Rational)	0.3273

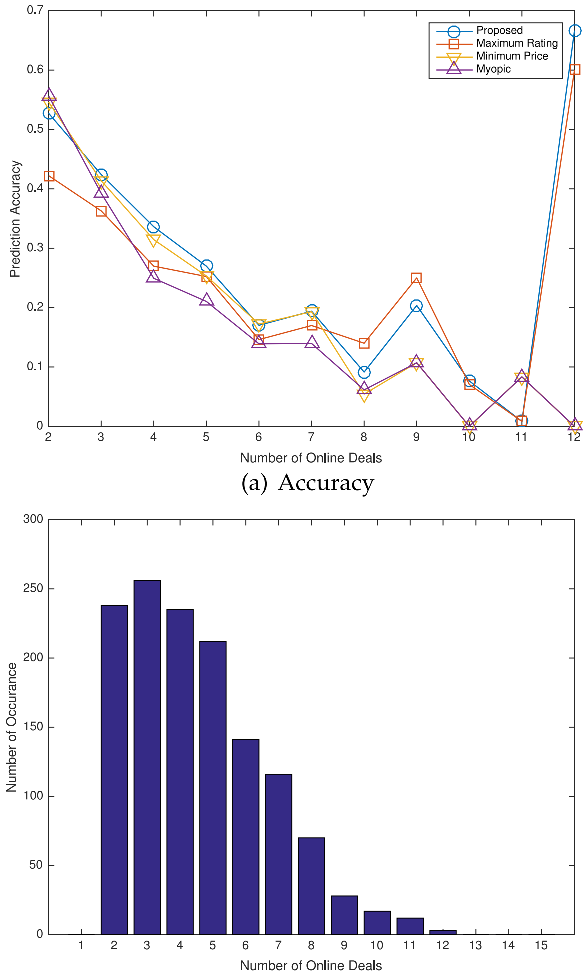
increment of arrival rate. This is due to the fact that all deals have limited quantities. A higher arrival rate means that customers are more likely to have less deals available for purchase when they arrive the website. In such a case, the differences among strategies diminishes. Nevertheless, the proposed strategy still outperforms all other naive strategies and performs closely with the social optimal one under all arrival rates.

7 Experiments - Are Customers Rational?

Finally, we would like to check whether customers in real world behave in a fully-rational way in deal selection problem. We apply several deal selection strategies on each snapshot of the available deals tagged with Restaurant in our Groupon dataset. Given a snapshot, the applied strategy selects a deal from the available ones accordingly. We define the prediction as correct when the strategy correctly predicts the deal that sold most quantities at next snapshot. When applying the proposed strategy, we reuse the deal selection model in previous simulations. We apply a pair-wise comparison method to find the best deal among all available deals. We first randomly treat the deals into pairs. For each pair, one deal is selected according to the trained deal selection model, while the other deal is removed from the set. The process repeats until exactly one deal remains. In this experiment, an interesting case on the rationality of customers, the myopic cases, is also considered. The proposed (myopic) strategy is a variant to the proposed deal selection strategy with limited rationality. Under this strategy, customers select the deal that maximizes their myopic utilities, that is, the utility if the deal goes offline immediately after his selection. This strategy represents the case that customers do aware of their utilities but lacking the capability to estimate the influence on the externality from the future.

We compare the prediction accuracy of the proposed strategy, both myopic and fully-rational versions, along with the ones under Random, Maximum Rating, and Minimum Price strategies. The results are shown in Table 3. We observed that Proposed (Fully-Rational) strategy has the highest accuracy among all the strategies, which suggests that customers are more likely to select the deals in a fully-rational way comparing to naive strategies. Additionally, the Minimum Price strategy has the highest accuracy among all naive strategies, which suggests that customers are more aware of the price of the deals than the rating on the other websites.

A detailed analysis on the prediction accuracy given different number of simultaneously online deals is provided in Fig. 9a. We observe that the proposed strategy performs best among all other strategies in most cases. In addition,



(a) Accuracy

(b) Number of Occurrences in Dataset

Fig. 9. Accuracy of predictions.

the minimum pricing strategy performs similar to the proposed Nash equilibrium strategy when the number of simultaneous available deals is fewer than 8. Notice that according to the dataset, it is also much common to have fewer than 8 Restaurant type deals available on the website, as we illustrated in Fig. 9b. Nevertheless, when the number of online deals increases, the maximum rating strategy performs better than the minimum pricing strategy. This may suggest that when the number of online deals increases, users may tend to find more information to compare the quality of deals. In contrast, when the available deals is limited, customers may prefer to just select the deal according to the price. The proposed strategy can utilize the advantage of both naive strategies and prediction on the externality of the deals and therefore performs best on average.

Nevertheless, the accuracy of the proposed strategy is still low, which also suggests that there is plenty of room for the customers to improve their decisions with proper assistance. The proposed deal selection strategy can be served as a deal suggestion tool to help customers correctly identify the price, rating, and externality effects on the experienced quality of the deal. With its assistance, not only the utility of the customer will increase but also the overall social welfare will be improved. Additionally, the proposed strategy is Nash equilibrium, which is compatible to the selfish nature of the customers who seek to maximize their utilities.

The reason that Nash equilibrium is a desired outcome is that it represents a predicted outcome when all customers behave selfishly, given that all customers have maximized their utility. This means that a customer, even if he is selfish, will be willing to follow this strategy. In other words, it takes less effort to guide the customers to follow the strategy.

One limitation of this experiment is that many deals in the dataset are not included in the experiment due to the lacking of corresponding Yelp records. This may cause bias in the experiment result since those deals with valid Yelp records suggest that the vendors have spent some effort in promotions. It may be helpful if we can include more information sources, such as reviews on Google maps or discussion on Twitter. Nevertheless, some deals may still have no records on any third-party website if the vendors are new. In such a case, we do not have objective information sources to analyze how customers collect information and their experience on the deal. It is still an open question to analyze how customers react to the deals when no public review is available.

8 Conclusions

In this paper, we develop a game-theoretic social learning framework to model the rational deal selection behaviors in Groupon with external information from Yelp. A stochastic learning model based on Chinese Restaurant Game is proposed to understand how rational customers select the deal with knowledge from external reviews and concerns on the externality caused by other customers. A year-long social media experiment is conducted to trace the competitions among deals on Groupon and the influences on the corresponding rating in Yelp records. Based on the learning model and social media dataset, we analyze the proposed deal selection game on a deal website with realistic settings. The performance of the proposed social learning framework is evaluated with simulations. We showed that the proposed framework significantly improves the social welfare and customer utility comparing to naive strategies. A further discussion on the rationality of customers in deal selection by comparing the results from simulations with real data is provided. The results suggest that customer do make decisions in a rational way instead of following naive strategies, but there is still room to improve the accuracy of their decisions with assistance from the proposed framework.

Acknowledgments

This work was partly supported by the Ministry of Science and Technology under Grant MOST 103-2218-E-001-002-MY2, MOST 105-2221-E-001-003-MY3 and the Academia Sinica under Thematic Research Grant (to Chih-Yu Wang). This work was also partly supported by the 111 Project No. B17008, the National Natural Science Foundation of China under Grant 61672137 and 61602090, and the Thousand Youth Talents Program of China (to Yan Chen).

References

[1] U. Dholakia, “How Effective are Groupon Promotions for Businesses?” Working Paper, Sep. 2010. [Online]. Available: https://ssrn.com/abstract=1696327
[2] F. Chen, D. Joshi, Y. Miura, and T. Ohkuma, “Social media-based profiling of business locations,” in *Proc. 3rd ACM MM Workshop Geotagging Appl. Multimedia*, 2014, pp. 1–6. [Online]. Available: http://doi.acm.org/10.1145/2661118.2661119
[3] M. Katz and C. Shapiro, “Technology adoption in the presence of network externalities,” *J. Political Economy*, vol. 94, pp. 822–841, 1986.
[4] W. Sandholm, “Negative externalities and evolutionary implementation,” *Rev. Econ. Stud.*, vol. 72, no. 3, pp. 885–915, 2005.
[5] G. Fagiolo, “Endogenous neighborhood formation in a local coordination model with negative network externalities,” *J. Econ. Dyn. Control*, vol. 29, no. 1/2, pp. 297–319, 2005.
[6] D. Oppenheimer, A. Ganapathi, and D. A. Patterson, “Why do Internet services fail, and what can be done about it?” in *Proc. 4th Conf. USENIX Symp. Internet Technol. Syst.*, 2003. [Online]. Available: http://dl.acm.org/citation.cfm?id=1251460.1251461
[7] J. Byers, M. Mitzenmacher, and G. Zervas, “Daily deals: Prediction, social diffusion, and reputational ramifications,” in *Proc. 5th ACM int. conf. Web search data mining (WSDM '12)*, pp. 543–552, Feb. 2012. [Online] Available: https://dl.acm.org/citation.cfm?id=2124361
[8] J. W. Byers, M. Mitzenmacher, and G. Zervas, “The groupon effect on yelp ratings: A root cause analysis,” in *Proc. 13th ACM Conf. Electron. Commerce*, Jun. 2012, pp. 248–265.
[9] T. A. H., “Groupon CEO apologizes to Japanese customers,” *Associated Press*, Jan. 2011. [Online]. Available: http://abcnews.go.com/Technology/wireStory?id=12630377
[10] V. Bala and S. Goyal, “Learning from neighbours,” *Rev. Econ. Stud.*, vol. 65, no. 3, 1998, Art. no. 595.
[11] B. Golub and M. Jackson, “Naıve learning in social networks: Convergence, influence, and the wisdom of crowds,” Working Paper, 2007.
[12] D. Acemoglu, M. Dahleh, I. Lobel, and A. Ozdaglar, “Bayesian learning in social networks,” *Rev. Econ. Stud.*, vol. 78, no. 4, pp. 1201–1236, Jan. 2011.
[13] D. Acemoglu and A. Ozdaglar, “Opinion dynamics and learning in social networks,” *Dyn. Games Appl.*, vol. 1, pp. 3–49, 2011.
[14] C.-Y. Wang, Y. Chen, and K. J. R. Liu, “Chinese restaurant game,” *IEEE Signal Process. Lett.*, vol. 19, no. 12, pp. 898–901, Dec. 2012.
[15] C.-Y. Wang, Y. Chen, and K. J. R. Liu, “Sequential Chinese restaurant game,” *IEEE Trans. Signal Process.*, vol. 61, no. 3, pp. 571–584, Feb. 2013.
[16] C. Jiang, Y. Chen, Y.-H. Yang, C.-Y. Wang, and K. J. R. Liu, “Dynamic chinese restaurant game: Theory and application to cognitive radio networks,” *IEEE Trans. Wireless Commun.*, vol. 13, no. 4, pp. 1960–1973, Apr. 2014.
[17] Y.-H. Yang, Y. Chen, C. Jiang, C.-Y. Wang, and K. J. R. Liu, “Wireless access network selection game with negative network externality,” *IEEE Trans. Wireless Commun.*, vol. 12, no. 10, pp. 5048–5060, Oct. 2013.
[18] C.-Y. Wang, Y. Chen, H.-Y. Wei, and K. J. R. Liu, “Scalable video multicasting: A stochastic game approach with optimal pricing,” *IEEE Trans. Wireless Commun.*, vol. 14, no. 5, pp. 2353–2367, May 2015.
[19] B. Golub and M. O. Jackson, “Naive learning in social networks and the wisdom of crowds,” *Amer. Econ. J.: Microeconomics*, vol. 2, no. 1, pp. 112–149, 2010.
[20] C. Deng and S. Pekec, “Money for nothing: Exploiting negative externalities,” in *Proc. 12th ACM Conf. Electron. Commerce*, Jun. 2011, pp. 361–370.
[21] F. Chierichetti, J. Kleinberg, and A. Panconesi, “How to schedule a cascade in an arbitrary graph,” in *Proc. 13th ACM Conf. Electron. Commerce*, Jun. 2012, pp. 355–368.
[22] M. Hajiaghayi, H. Mahini, and D. Malec, “The polarizing effect of network influences,” in *Proc. 15th ACM Conf. Econ. Comput.*, Jun. 2014, pp. 131–148.
[23] T. Martin, G. Schoenebeck, and M. Wellman, “Characterizing strategic cascades on networks,” in *Proc. 15th ACM Conf. Econ. Comput.*, Jun. 2014, pp. 113–130.
[24] S. Shahrampour, A. Jadbabaie, and A. Rakhlin, “Online learning of dynamic parameters in social networks,” in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2013, pp. 2013–2021.
[25] M.-F. F. Balcan, C. Berlind, A. Blum, E. Cohen, K. Patnaik, and L. Song, “Active learning and best-response dynamics,” in *Proc. Int. Conf. Neural Inf. Process. Syst.*, Dec. 2014, pp. 2222–2230.
[26] W. Mason, S. Suri, and D. J. Watts, “Long-run learning in games of cooperation,” in *Proc. 15th ACM Conf. Econ. Comput.*, Jun. 2014, pp. 821–838.
[27] D. Fudenberg and A. Peysakhovich, “Recency, records and recaps: Learning and non-equilibrium behavior in a simple decision problem,” in *Proc. 15th ACM Conf. Econ. Comput.*, Jun. 2014, pp. 971–986.

[28] S. Zhang and A. J. Yu, "Forgetful Bayes and myopic planning: Human learning and decision-making in a bandit setting," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2013, pp. 2607–2615.
[29] S. Guo, M. Wang, and J. Leskovec, "The role of social networks in online shopping: Information passing, price of trust, and consumer choice," in *Proc. 12th ACM Conf. Electron. Commerce*, Jun. 2011, pp. 157–166.
[30] E. Parzen, *Stochastic Processes*, vol. 24. Philadelphia, PA, USA: SIAM, 1999.

Chih-Yu Wang (S'07-M'13) received the BS and PhD degrees in electrical engineering and communication engineering from National Taiwan University (NTU), Taipei, Taiwan, in 2007 and 2013, respectively. He has been a visiting student with the University of Maryland, College Park, in 2011. He is currently an assistant research fellow with the Research Center for Information Technology Innovation, Academia Sinica, Taipei, Taiwan. His research interests include game theory, wireless communications, social networks, and data science. He is a member of the IEEE.

Yan Chen (SM'14) received the bachelor's degree from the University of Science and Technology of China, in 2004, the MPhil degree from the Hong Kong University of Science and Technology, in 2007, and the PhD degree from the University of Maryland, College Park, Maryland, in 2011. He was with Origin Wireless Inc. as a founding principal technologist. Since Sept. 2015, he has been a full professor with the University of Electronic Science and Technology of China. His research interests include multimedia, signal processing, game theory, and wireless communications. He was the recipient of multiple honors and awards, including the best student paper award at the PCM in 2017, best student paper award at the IEEE ICASSP in 2016, the best paper award at the IEEE GLOBECOM in 2013, the Future Faculty Fellowship and Distinguished Dissertation Fellowship Honorable Mention from the Department of Electrical and Computer Engineering in 2010 and 2011, respectively the Finalist of the Dean's Doctoral Research Award from the A. James Clark School of Engineering, University of Maryland in 2011, and the Chinese Government Award for outstanding students abroad in 2010. He is a senior member of the IEEE.

K. J . Ray Liu (F'03) was named a distinguished scholar-teacher of the University of Maryland, College Park, in 2007, where he is Christine Kim Eminent Professor of Information Technology. He leads the Maryland Signals and Information Group conducting research encompassing broad areas of information and communications technology with recent focus on smart radios for smart life. He was a recipient of the 2016 IEEE Leon K. Kirchmayer Technical Field Award on graduate teaching and mentoring, IEEE Signal Processing Society 2014 Society Award, and IEEE Signal Processing Society 2009 Technical Achievement Award. He is recognized by Thomson Reuters as a Highly Cited researcher. His invention of Time-Reversal Machine by Origin Wireless won the 2017 CEATEC Grand Prix award. He is IEEE vice president, Technical Activities Elect. He was president of the IEEE Signal Processing Society, where he has served as vice president Publications and Board of Governor, and a member of the IEEE Board of Director as Division IX Director. He has also served as the editor-in-chief of the *IEEE Signal Processing Magazine*. He also received teaching and research recognitions from the University of Maryland including university-level Invention of the Year Award; and college-level Poole and Kent Senior Faculty Teaching Award, Outstanding Faculty Research Award, and Outstanding Faculty Service Award, all from A. James Clark School of Engineering. He is a fellow of the IEEE and AAAS.