Sequential Chinese Restaurant Game

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

***Abstract*—In a social network, agents are intelligent and have the capability to make decisions to maximize their utilities. They can either make wise decisions by taking advantages of other agents' experiences through learning, or make decisions earlier to avoid competitions from huge crowds. Both these two effects, social learning and negative network externality, play important roles in the decision process of an agent. While there are existing works on either social learning or negative network externality, a general study on considering both effects is still limited. We find that Chinese restaurant process, a popular random process, provides a well-defined structure to model the decision process of an agent under these two effects. By introducing the strategic behavior into the non-strategic Chinese restaurant process, we propose a new game, called the Chinese restaurant game, to formulate the social learning problem with negative network externality. Through analyzing the proposed Chinese restaurant game, we derive the optimal strategy of each agent and provide a recursive method to achieve the optimal strategy. How social learning and negative network externality influence each other under various settings is studied through simulations. We also illustrate the spectrum access problem in cognitive radio networks as one of the application of Chinese restaurant game. We find that the proposed Chinese restaurant game theoretic approach indeed helps users make better decisions and improves the overall system performance.**

***Index Terms*—Chinese restaurant game, Chinese restaurant process, cognitive radio, cooperative sensing, game theory, machine learning, network externality, social learning.**

I. Introduction

HOW agents in a network learn and make decisions is an important issue in numerous research fields, such as social learning in social networks, machine learning with communications among devices, and cognitive adaptation in cognitive radio networks. Agents make decisions in a network in order to achieve certain objectives. However, the agent's knowledge on the system may be limited due to the limited ability in observations or the external uncertainty in the system. This impaired his utility since he does not have enough knowledge to make correct decisions. The limited knowledge of one agent can be expanded through learning. One agent may learn from some information sources, such as the decisions of other agents, the advertisements from some brands, or his experience in previous purchases. In most cases, the accuracy of the agent's decision can be greatly enhanced by learning from the collected information.

The learning behavior in a social network is a popular topic in the literature. Let us consider a social network in an uncertain system state. The state has an impact on the agents' rewards. When the impact is differential, i.e., one action results in a higher reward than other actions in one state but not in all states, the state information becomes critical for one agent to make the correct decision. In most social learning literature, the state information is unknown to agents. Nevertheless, some signals related to the system state are revealed to the agents. Then, the agents make their decisions sequentially, while their actions/signals may be fully or partially observed by other agents. Most of existing works [1]–[4] study how the believes of agents are formed through learning in the sequential decision process, and how accurate the believes will be when more information is revealed. One popular assumption in traditional social learning literature is that there is no network externality, i.e., the actions of subsequent agents do not influent the reward 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. This assumption greatly limits the potential applications of these existing works.

The network externality, i.e., the influence of other agents' behaviors on one agent's reward, is a classic topic in economics. How the relations of agents influence an agent's behavior is studied in coordinate game theory [5]. When the network externality is positive, the problem can be modeled as a coordination game. In the literature, there are some works on combining the positive network externality with social learning, such as voting game [6]–[8] and investment game [9]–[12]. In voting game, an election with several candidates is held, where voters have their own preferences on the candidates. The preference of a voter on the candidates is constructed by how the candidates can benefit him if winning the election. When more voters vote for the same candidate, he is more likely to win the election and thus benefits the voters. In the investment game, there are multiple projects and investors, where each project has different payoff. When more investors invest in the same project, the succeeding probability of the project increases, which benefits all investors investing this project. Note that in both voting and investment games, the agent's decision has a positive effect on ones' decisions. When one agent makes a decision, the subsequent agents are encouraged to make the same decision in two aspects: the probability that this action has the positive outcome increases due to this agent's decision, and the potential reward of this action may be large according to the belief of this agent.

When the externality is negative, it becomes an anti-coordination game, where agents try to avoid making the same

Manuscript received December 15, 2011; revised June 10, 2012, September 14, 2012, and September 24, 2012; accepted September 29, 2012. Date of publication October 16, 2012; date of current version January 11, 2013. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Ignacio Santamaria. This work was supported by the National Science Council of Taiwan under Grant NSC-100-2917-I-002-038.

C.-Y. Wang is with the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742 USA.. He is also with the Graduate Institute of Communication Engineering, National Taiwan University, Taiwan (e-mail: tomkywang@gmail.com).

Y. Chen and K. J. R. Liu are with the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742 USA (e-mail: yan@umd.edu; kjrliu@umd.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2012.2225053

decisions with others [13]–[15]. The combination of negative network externality with social learning, on the other hand, is difficult to analyze. When the network externality is negative, the game becomes an anti-coordination game, where one agent seeks the strategy that differs from others' to maximize his own reward. Nevertheless, in such a scenario, the agent's decision also contains some information about his belief on the system state, which can be learned by subsequent agents through social learning algorithms. Thus, subsequent agents may then realize that his choice is better than others, and make the same decision with the agent. Since the network externality is negative, the information leaked by the agent's decision may impair the reward the agent can obtain. Therefore, rational agents should take into account the possible reactions of subsequent players to maximize their own rewards.

The negative network externality plays an important rule in many applications in different research fields. One important application is spectrum access in cognitive radio networks. In spectrum access problem, secondary users accessing the same spectrum need to share with each other. The more secondary users access the same channel, the less available access time or higher interference for each of them. In this case, the negative network externality degrades the utility of the agents making the same decision. As illustrated in [16], the interference from other secondary users will degrade a secondary user's transmission quality and can be considered as the negative network externality effect. Therefore, the agents should take into account the possibility of degraded utility when making the decisions. Similar characteristics can also be found in other applications, such as service selection in cloud computing and deal selection in Groupon website.

The aforementioned social learning approaches are mostly strategic, where agents are considered as players with bounded or unbounded rationality in maximizing their own rewards. Machine learning, which is another class of approaches for the learning problem, focuses on designing algorithms for making use of the past experience to improve the performance of similar tasks in the future [17]. Generally there exists some training data and the devices follow a learning method designed by the system designer to learn and improve the performance of some specific tasks. Most learning approaches studied in machine learning are non-strategic without the rationality on considering their own benefit. Such non-strategic learning approaches may not be applicable to the scenario where devices are rational and intelligent enough to choose actions to maximize their own benefits instead of following the rule designed by the system designer.

Chinese restaurant process, which is a non-parametric learning methods in machine learning [18], provides an interesting non-strategic learning method for unbounded number of objects. In Chinese restaurant process, there exists infinite number of tables, where each table has infinite number of seats. There are infinite number of customers entering the restaurant sequentially. When one customer enters the restaurant, he can choose either to share the table with other customers or to open a new table, with the probability being predefined by the process. Generally, if a table is occupied by more customers, then a new customer is more likely to join the table, and the probability that a customer opens a new table can be controlled by a parameter [19]. This process provides a systematic method to construct the parameters for modeling unknown distributions. Nevertheless, the behavior of customers in Chinese restaurant game is non-strategic, which means they follow predefined rules without rational concerns on their own utility. We observe that if we introduce the strategic behaviors into Chinese restaurant process, the model can be a general framework for analyzing the social learning with negative network externality. To the best of our knowledge, no effort has been made to bring rationality concerns into such a decision making structure in the literature.

By introducing the strategic behavior into the non-strategic Chinese restaurant process, we propose a new game, called **Chinese Restaurant Game**, to formulate the social learning problem with negative network externality In our previous work [20], we have studied the simultaneous Chinese restaurant game without social learning where customers make decisions simultaneously. In this paper, we will study the sequential Chinese restaurant game with social learning where customers make decisions sequentially. Let us consider a Chinese restaurant with J tables. There are N customers sequentially requesting for seats from these J tables for having their meals. One customer may request one of the tables in number. After requesting, he will be seating in the table he requested. We assume that all customers are rational, i.e., they prefer bigger space for a comfortable dining experience. Thus, one may be delighted if he has a bigger table. However, since all tables are available to all customers, he may need to share the table with others if multiple customers request for the same table. In such a case, the customer's dining space reduces, due to which the dining experience is impaired. Therefore, the key issue in the proposed Chinese restaurant game is how the customers choose the tables to enhance their own dining experience. This model involves the negative network externality since the customer's dining experience is impaired when others share the same table with him. Moreover, when the table size is unknown to the customers, but each of them receives some signals related to the table size, this game involves the learning process if customers can observe previous actions or signals.

In the rest of the paper, we first provide detailed descriptions on the system model of Chinese restaurant game in Section III. Then, we study the sequential game model with perfect information to illustrate the advantage of playing first in Section IV. In Section V, we show the general Chinese restaurant game framework by analyzing the learning behaviors of customers under the negative network externality and uncertain system state. We provide a recursive method to construct the best response for customers, and discuss the simulation results in Section VI. In Section VII, we illustrate how the traditional spectrum access problem can be formulated as a Chinese restaurant game. Finally, we draw conclusions in Section VIII.

II. Related Works

A closely-related strategic game model to our work is the global game [21], [22]. In the global game, all agents, with limited knowledge on the system state and information held by other agents, make decisions simultaneously. The agent's reward in the game is determined by the system state and the number of agents making the same decision with him. The influence may be positive or negative depending on the type of network externality. An important characteristics of global game is that the equilibrium is unique, which simplifies the discussion on the outcome of the game. It draws great attentions in various research fields, such as financial crisis [23], sensor networks [24] and cognitive radio networks [25]. Since all players in the global game make decisions simultaneously, there is no learning involved in the global game.

In recent years, several works [10], [11], [26]–[28] make efforts to introduce the learning and signaling into the global game. Dasgupta's first attempt was investigating a binary investment model, while one project will succeed only when enough number of agents invest in the project in [10]. Then, Dasgupta studied a two-period dynamic global game, where the agents have the options to delay their decisions in order to have better private information of the unknown state in [11].

Angeletos *et al.* studied a specific dynamic global game called regime change game [26], [27]. In the regime change game, each agent may propose an attack to the status quo, i.e., the current politic state of the society. When the collected attacks are large enough, the status quo is abandoned and all attackers receive positive payoffs. If the status quo does not change, the attackers receive negative payoffs. Angeletos *et al.* first studied a signaling model with signals at the beginning of the game in [26]. Then, they proposed a multiple stages dynamic game to study the learning behaviors of agents in the regime change game in [27].

Costain provided a more general dynamic global game with an unknown binary state and a general utility function in [28]. However, the positions of the agents in the game are assumed to be unknown to simply the analysis. Nevertheless, most of these works study the multiplicity of equilibria in dynamic global game with simplified models, such as binary state, binary investment model, or lacking of position information. Moreover, the network externality they considered in their models are mostly positive. By proposing Chinese restaurant game, we hereby provides a more general game-theoretic framework on studying the social learning in a network with negative network externality, which has many applications in various research fields.

III. System Model

Let us consider a Chinese restaurant with J tables numbered $1, 2, \ldots, J$ and N customers labeled with $1, 2, \ldots, N$. Each customer requests for one table for having a meal. Each table has infinite seats, but may be in different size. We model the table sizes of a restaurant with two components: the restaurant state θ and the table size functions $\{R_1(\theta), R_2(\theta), \ldots, R_J(\theta)\}$. The state θ represents an objective parameter, which may be changed when the restaurant is remodeled. The table size function $R_j(\theta)$ is fixed, i.e., the functions $\{R_1(\theta), R_2(\theta), \ldots, R_J(\theta)\}$ will be the same every time the restaurant is remodeled. An example of θ is the order of existing tables. Suppose that the restaurant has two tables, one is of size L and the other is of size S. Then, the owner may choose to number the large one as table 1, and the small one as table 2. The decision on the numbering can be modeled as $\theta \in \{1, 2\}$, while the table size functions $R_1(\theta)$ and $R_2(\theta)$ are given as $R_1(1) = L$, $R_1(2) = S$, and $R_2(1) = S$, $R_2(2) = L$. Let Θ be the set of all possible state of the restaurant. In this example, $\Theta = \{1, 2\}$.

We formulate the table selection problem as a game, called **Chinese Restaurant Game**. We first denote $\mathcal{X} = \{1, \ldots, J\}$ as the action set (tables) that a customer may choose, where $x_i \in \mathcal{X}$ means that customer i chooses the table x_i for a seat. Then, the utility function of customer i is $U(R_{x_i}, n_{x_i})$, where n_{x_i} is the number of customers choosing table x_i. According to our previous discussion, the utility function should be an increasing function of R_{x_i}, and a decreasing function of n_{x_i}. Note that the decreasing characteristic of $U(R_{x_i}, n_{x_i})$ over n_{x_i} can be regarded as the negative network externality effect since the degradation of the utility is due to the joining of other customers. Finally, let $\mathbf{n} = (n_1, n_2, \ldots, n_J)$ be the numbers of customers on the J tables, i.e., the **grouping** of customers in the restaurant.

As mentioned above, the restaurant is in a state $\theta \in \Theta$. However, customers may not know the exact state θ, i.e., they may not know the exact size of each table before requesting. Instead, they may have received some advertisements or gathered some reviews about the restaurant. The information can be treated as some kinds of signals related to the true state of the restaurant. In such a case, they can estimate θ through the available information, i.e., the information they know and/or gather in the game process. We assume that all customers know the prior distribution of the state information θ, which is denoted as $\mathbf{g_0} = \{g_{0,l} | g_{0,l} = Pr(\theta = l), \forall l \in \Theta\}$. The signal each customer received $s_i \in \mathcal{S}$ is generated from a predefined distribution $f(s|\theta)$. Notice that the signal quality may vary, depending on how accurate the signal can reflect the state. A simple example is given as follows. Considering a signal space $\mathcal{S} = \{1, 2\}$ and the system state space $\Theta = \{1, 2\}$. Then, we define the signal distribution as follows:

$$Pr(s = \theta|\theta) = p, \quad Pr(s \neq \theta|\theta) = 1-p, \quad 0.5 \leq p \leq 1. \quad (1)$$

In such a case, the parameter p is the signal quality of this signal distribution. When p is higher, the signal is more likely to reflect the true system state.

We introduce **belief**, which is well-known in the Bayesian game literature [3], to describe how a customer estimates the system state θ. Since customers make decisions sequentially, it is possible that the customers who make decisions later learn the signals from those customers who make decisions earlier. Let us denote the signals customer i learned, excluding his own signal s_i, as $\mathbf{h_i} = \{s\}$. With the help of these signals $\mathbf{h_i}$, his own signal s_i, the prior distribution $\mathbf{g_0}$, and the conditional distribution $f(s|\theta)$, each customer i can estimate the current system state in probability with the belief being defined as

$$\mathbf{g_i} = \{g_{i,l} | g_{i,l} = Pr(\theta = l | \mathbf{h_i}, s_i, \mathbf{g_0}), \forall l \in \Theta\} \, \forall i \in N. \quad (2)$$

According to the above definition, $g_{i,l}$ represents the probability that system state θ is equal to l conditioning on the collected signals $\mathbf{h_i}$, received signal s_i, the prior probability $\mathbf{g_0}$, and the conditional distribution $f(s|\theta)$. Notice that in the social learning literature, the belief can be obtained through either non-Bayesian updating rule [1], [2] or fully rational Bayesian rule [3]. For the non-Bayesian updating rule, it is implicitly based on the assumption that customers are only limited rational and follows some predefined rules to compute their believes. Their capability to maximize their utilities is limited not only by the game structure and learned information, but also by the non-Bayesian updating rules. In the fully rational Bayesian rule, customers are fully rational and have the potential to optimize their actions without the restriction on the fixed belief updating rule. Since the customers we considered here are fully rational, they will follow the Bayesian rule to update their believes as follows:

$$g_{i,l} = \frac{g_{0,l} Pr(\mathbf{h_i}, s_i|\theta = l)}{\sum_{l' \in \Theta} g_{0,l'} Pr(\mathbf{h_i}, s_i|\theta = l')}. \tag{3}$$

Notice that the exact expression for belief depends on how the signals are generated and learned, which is generally affected by the conditional distribution $f(s|\theta)$ and the game structure.

IV. Perfect Signal: Advantage of Playing First

We first study the perfect signal case, where the system state θ is known by all customers. Let us consider a Chinese restaurant game with J tables and N customers. Since θ is known, the exact sizes of tables $R_1(\theta), R_2(\theta), \ldots, R_J(\theta)$ are also known by customers.

In sequential Chinese restaurant game, customers make decisions sequentially with a predetermined order known by all customers, e.g., waiting in a line of the queue outside of the restaurant. Without loss of generality, in the rest of this paper, we assume the order is the same as the customer's number. We assume every customer knows the decisions of the customers who make decisions before him, i.e., customer i knows the decisions of customers $\{1, \ldots, i-1\}$. Let $\mathbf{n_i} = (n_{i,1}, n_{i,2}, \ldots, n_{i,J})$ be the current grouping, i.e., the number of customers choosing table $\{1, 2, \ldots, J\}$ before customer i. The $\mathbf{n_i}$ roughly represents how crowded each tables is when customer i enters the restaurant. Notice that $\mathbf{n_i}$ will not be equal to $\mathbf{n}$, which is the final grouping that determines customers' utilities. A table with only few customers may eventually be chosen by many customers in the end.

A strategy describes how a player will play given any possible situation in the game. In Chinese restaurant game, the customer's strategy should be a mapping from other customers' table selections to his own table selection. Recalling that n_j stands for the number of customers choosing table j. Let us denote $\mathbf{n_{-i}} = (n_{-i,1}, n_{-i,2}, \ldots, n_{-i,J})$ with $n_{-i,j}$ being the number of customers except customer i choosing table j. Then, given $\mathbf{n_{-i}}$, the best response of a rational customer i should be

$$BE_i(\mathbf{n_{-i}}, \theta) = \arg\max_{x \in \mathcal{X}} U\left(R_x(\theta), n_{-i,x} + 1\right). \tag{4}$$

Notice that given $\mathbf{n_{-i}}$, $n_j = n_{-i,j} + 1$ if $x = j$. However, the $\mathbf{n_{-i}}$ may not be completely observable by customer i since customers $i+1 \sim N$ make decisions after customer i. Therefore, as shown in the next subsection, customer i should predict the decisions of the subsequent customers given the current observation $\mathbf{n_i}$ and state θ.

A. Equilibrium Grouping

We first study the possible equilibria of Chinese restaurant game. Nash equilibrium is a popular concept for predicting the outcome of a game with rational customers. Informally speaking, Nash equilibrium is an action profile, where each customer's action is the best response to other customers' actions in the profile. Since all customers use their best responses, none of them have the incentive to deviate from their actions. We observe that in Chinese restaurant game, the Nash equilibrium can be translated into the equilibrium grouping [20], which is defined as follows

Definition 1: Given the customer set $\{1, \ldots, N\}$, the table set $\mathcal{X} = \{1, \ldots, J\}$, and the current system state θ, an **equilibrium grouping** $\mathbf{n}^*$ satisfies the following conditions

$$U\left(R_x(\theta), n_x^*\right) \geq U\left(R_y(\theta), n_y^* + 1\right),\ if\ n_x^* > 0, \forall x, y \in \mathcal{X}. \tag{5}$$

Obviously, there will be more than one Nash equilibrium since we can always exchange the actions of any two customers in one Nash equilibrium to build a new Nash equilibrium without violating the sufficient and necessary condition shown in (5). Nevertheless, the equilibrium grouping $\mathbf{n}^*$ may be unique even if there exist multiple Nash equilibria. The sufficient condition to guarantee the uniqueness of equilibrium grouping is stated in the following Theorem.

Theorem 1: If the inequality in (5) strictly holds for all $x, y \in \mathcal{X}$, then the equilibrium grouping $\mathbf{n}^* = (n_1^*, \ldots, n_J^*)$ is unique.

Proof: We would like to prove this by contradiction. Suppose that there exists another Nash equilibrium with equilibrium grouping $\mathbf{n}' = (n'_1, \ldots, n'_J)$, where $n'_j \neq n_j^*$ for some $j \in \mathcal{X}$. Since both $\mathbf{n}^*$ and $\mathbf{n}'$ are equilibrium groupings, we have $\sum_{j=1}^{J} n'_j = \sum_{j=1}^{J} n_j^* = N$. In such a case, there exists two tables x and y with $n'_x > n_x^*$ and $n'_y < n_y^*$. Then, since $\mathbf{n}^*$ is an equilibrium grouping, we have

$$U\left(R_y(\theta), n_y^*\right) > U\left(R_x(\theta), n_x^* + 1\right). \tag{6}$$

Since $n'_x > n_x^*$, $n'_y < n_y^*$, and $U(\cdot)$ is a deceasing function of n, we have

$$U\left(R_x(\theta), n_x^*\right) > U\left(R_x(\theta), n_x^* + 1\right) \geq U(R_x(\theta), n'_x), \tag{7}$$

$$U\left(R_y(\theta), n'_y\right) > U\left(R_y(\theta), n'_y + 1\right) \geq U\left(R_y(\theta), n_y^*\right). \tag{8}$$

Since $\mathbf{n}'$ is also an equilibrium grouping, we have

$$U\left(R_x(\theta), n'_x\right) \geq U\left(R_y(\theta), n'_y + 1\right). \tag{9}$$

According to (7), (8), and (9) we have

$$\begin{aligned} U\left(R_x(\theta), n_x^* + 1\right) &\geq U\left(R_x(\theta), n'_x\right) \\ &\geq U\left(R_y(\theta), n'_y + 1\right) \\ &\geq U\left(R_y(\theta), n_y^*\right), \end{aligned} \tag{10}$$

which contradicts with (6). Therefore, the equilibrium grouping $\mathbf{n}^*$ is unique when the inequality in (5) strictly holds. ■

A concrete example that the equilibrium grouping is and is not unique is as follows. Consider a Chinese restaurant with 3 customers and 2 tables with size R_1 and R_2. When $R_1 = R_2$, we have two equilibrium grouping, which are $\mathbf{n^1} = (1, 2)$ and $\mathbf{n^2} = (2, 1)$. The equilibrium grouping is not unique in this case is because the inequality in (5) does not strictly hold, which means that one customer may have the same utility if he chooses another table given the decisions of others. In contrast, when $R_1 > R_2$ and $U(R_1, 3) < U(R_2, 1)$, we have a unique equilibrium grouping $\mathbf{n^3} = (2, 1)$ since all other grouping cannot be the equilibrium output as we proved in Theorem 1.

The equilibrium grouping can be found through a simple greedy algorithm. In the algorithm, customers choose their actions in the myopic way, i.e., they choose the tables that can maximize their current utilities purely based on what they have observed. Let $\mathbf{n_i} = (n_{i,1}, n_{i,2}, \ldots, n_{i,J})$ with

$\sum_{j=1}^{J} n_{i,j} = i - 1$ be the grouping observed by customer i. Then, customer i will choose the myopic action given by

$$BE_i^{myopic}(\mathbf{n_i}, \theta) = \arg\max_{x \in \mathcal{X}} U\left(R_x(\theta), n_{i,x} + 1\right). \quad (11)$$

We check if the greedy algorithm indeed outputs an equilibrium grouping. Let $\mathbf{n}^* = (n_1^*, n_2^*, \ldots, n_J^*)$ be the corresponding grouping. For a table j with $n_j^* > 0$, suppose customer k is the last customer choosing table j. According to (11), we have

$$\begin{aligned} U\left(R_j(\theta), n_{k,j} + 1\right) &\geq U\left(R_{j'}(\theta), n_{k,j'} + 1\right) \\ &\geq U\left(R_{j'}(\theta), n_{j'}^* + 1\right), \forall j' \in \mathcal{X}. \end{aligned} \quad (12)$$

Note that (12) holds for all $j, j' \in \mathcal{X}$ with $n_j^* > 0$, i.e., $U(R_j(\theta), n_j^*) \geq U(R_{j'}(\theta), n_{j'}^* + 1)$, $\forall j, j' \in \mathcal{X}$ with $n_j^* > 0$. According to Definition 1, the output grouping $\mathbf{n}^*$ from the greedy algorithm is an equilibrium grouping.

B. Subgame Perfect Nash Equilibrium

In a sequential game, we will study the subgame perfect Nash equilibrium. Subgame perfect Nash equilibrium is a popular refinement to the Nash equilibrium under the sequential game. It guarantees that all players choose strategies rationally in every possible subgame. A subgame is a part of the original game. In Chinese restaurant game, any game process begins from player i, given all possible actions before player i, could be a subgame.

Definition 2: A subgame in Chinese restaurant game is consisted of two elements: 1) It begins from customer i; 2) The current grouping before customer i is $\mathbf{n_i} = (n_{i,1}, \ldots, n_{i,J})$ with $\sum_{j=1}^{J} n_{i,j} = i - 1$.

Definition 3: A Nash equilibrium is a subgame perfect Nash equilibrium if and only if it is a Nash equilibrium for any subgame.

We would like to show the existence of subgame perfect Nash equilibrium in Chinese restaurant game by constructing one. Basically, as a rational customer, customer i should predict the final equilibrium grouping according to his current observation on the choices of previous customers $\mathbf{n_i}$ and the system state θ. Then, he may choose the table with highest expected utility according to the prediction. Following from this idea, we derive the best response of customers in a subgame.

We first implement the prediction part through two functions as follows. First, let $EG(\mathcal{X}^s, N^s)$ be the function that generates the equilibrium grouping for a table set $\mathcal{X}^s$ and number of customers N^s. The equilibrium grouping is generated by the greedy algorithm shown in previous section with $\mathcal{X}$ being replaced by $\mathcal{X}^s$ and N being replaced by N^s. Notice that $\mathcal{X}^s$ could be any subset of the total table set $\mathcal{X} = \{1, \ldots, J\}$, and N^s is less or equal to N.

Then, let $PC(\mathcal{X}^s, \mathbf{n^s}, N^s)$, where $\mathbf{n^s}$ denotes the current grouping observed by the customer, be the algorithm that generates the set of available tables given $\mathbf{n^s}$ in the subgame. The algorithm removes the tables that already occupied by more than the expected number of customers in the equilibrium grouping. This helps the customer remove those unreasonable choices and correctly predict the final equilibrium grouping in every subgame. The basic flow of this algorithm is shown as follows 1) calculate the equilibrium grouping $\mathbf{n^e}$ given the table set $\mathcal{X}^s$ and number of customers N^s, 2) check if there is any overly occupied table by comparing $\mathbf{n^s}$ with $\mathbf{n^e}$. If so, 3) remove these tables from $\mathcal{X}^s$ and the customers occupying these tables from N^s, and go back to 1). Otherwise, the algorithm terminates. The procedures of implementing $PC(\mathcal{X}^s, \mathbf{n^s}, N^s)$ are described as follows:

1) Initialize: $\mathcal{X}^o = \mathcal{X}^s$, $N^t = N^s$
2) $\mathcal{X}^t = \mathcal{X}^o$, $\mathbf{n^e} = EG(\mathcal{X}^t, N^t)$, $\mathcal{X}^o = \{x | x \in \mathcal{X}^t, n_j^e \geq n_j^s\}$, $N^t = N^s - \sum_{x \in \mathcal{X}^s \backslash \mathcal{X}^o} n_x^s$.
3) If $\mathcal{X}^o \neq \mathcal{X}^t$, go back to step 2.
4) Output $\mathcal{X}^o$.

Now, we propose a method to construct a subgame perfect Nash equilibrium. This equilibrium also satisfies (5). For each customer i, his strategy in a subgame is

$$BE_i^{se}(\mathbf{n_i}, \theta) = \arg \max_{x \in \mathcal{X}^{i,cand}, n_{i,x} < n_x^{i,cand}} U\left(R_x(\theta), n_x^{i,cand}\right), \quad (13)$$

where $\mathcal{X}^{i,cand} = PC(\mathcal{X}, \mathbf{n_i}, N)$, $N^{i,cand} = N - \sum_{x \in \mathcal{X} \backslash \mathcal{X}^{i,cand}} n_{i,x}$, and $\mathbf{n^{i,cand}} = EG(\mathcal{X}^{i,cand}, N^{i,cand})$. The proposed best response $BE_i^{se*}(\mathbf{n_i}, \theta)$ chooses the table with the highest utility according to the predicted equilibrium grouping $\mathbf{n^{i,cand}}$ and candidate table set $\mathcal{X}^{i,cand}$. The equilibrium grouping $\mathbf{n^{i,cand}}$ is obtained by $EG(\mathcal{X}^{i,cand}, N^{i,cand})$, where the candidate table set $\mathcal{X}^{i,cand}$ is derived by $PC(\mathcal{X}, \mathbf{n_i}, N)$. In Lemma 2, we show that the above strategy results in the equilibrium grouping in any subgame.

Lemma 2: Given the available table set $\mathcal{X}^s = PC(\mathcal{X}, \mathbf{n^s}, N)$, $N^s = N - \sum_{x \in \mathcal{X} \backslash \mathcal{X}^s} n_x^s$, the proposed strategy shown in (13) leads to an equilibrium grouping $\mathbf{n}^* = EG(\mathcal{X}^s, N^s)$ over $\mathcal{X}^s$.

Proof: We prove this by contradiction. Let $\mathbf{n} = (n_j | j \in \mathcal{X}^s)$ be the final grouping after all customers choose their tables according to (13). Suppose that $\mathbf{n} \neq \mathbf{n}^* = EG(\mathcal{X}^s, N^s)$, then there exists some tables j that $n_j > n_j^*$. Let table j be the first table that exceeds n_j^* in this sequential subgame. Since $n_j > n_j^*$, there are at least $n_j^* + 1$ customers choosing table j. Suppose the $n_j^* + 1$-th customer choosing table j is customer i. Let $\mathbf{n_i} = (n_{i,1}, n_{i,2}, \ldots, n_{i,J})$ be the current grouping observed by customer i before he chooses the table. Since customer i is the $n_j^* + 1$-th customer choosing table j, we have $n_{i,j} = n_j^*$. Since table j is the first table exceeding $\mathbf{n}^*$ after customer i's choice, we have $n_{i,x} \leq n_x^* \forall x \in \mathcal{X}^s$.

According to the definition of $PC(\cdot)$, none of the tables will be removed from candidates. Thus, $\mathcal{X}^{i,cand} = \mathcal{X}^s$ and $N^{i,cand} = N^s$. We have

$$\mathbf{n^{i,cand}} = EG(\mathcal{X}^{i,cand}, N^{i,cand}) = EG(\mathcal{X}^s, N^s) = \mathbf{n}^*. \quad (14)$$

However, according to (13), the customer i should not choose table j since $n_{i,j} = n_j^* = n_j^{i,cand}$. This contradicts with our assumption that customer i is the $n_j^* + 1$-th customer choosing table j. Thus, the strategy (13) should lead to the equilibrium grouping $\mathbf{n}^* = EG(\mathcal{X}^s, N^s)$. ■

Note that Lemma 3 also shows that the final grouping of the sequential game should be $\mathbf{n}^* = EG(\mathcal{X}, N)$ if all customers follow the proposed strategy in (13). In the following Lemma, we show that $PC(\mathcal{X}^s, \mathbf{n^s}, N^s)$ removes the tables that are dominated by other tables if all customers follow (13).

Lemma 3: Given a subgame with current grouping $\mathbf{n^s}$, if table $j \notin \mathcal{X}^s = PC(\mathcal{X}, \mathbf{n^s}, N)$, then table j is never the best response of the customer if all other customers follow (13).

Proof: Let $\mathbf{n}' = EG(\mathcal{X}, N)$, and $\mathbf{n}^*$ be the final grouping. We first show that for every table under the final grouping $\mathbf{n}^*$, there always exists a table providing a less or equal utility under the grouping $\mathbf{n}'$. According to Lemma 2, the final grouping $\mathbf{n}^*$ is an equilibrium grouping over $\mathcal{X}^s$ if all customers follow (13). Additionally, $n_j^* = n_j^s$ since no customers will choose table j. Assuming that there exists a table $k \in \mathcal{X}^s$ with $n'_k < n^*_k$. Since $n_j^* = n_j^s > n'_j$, we have $\sum_{x\in\mathcal{X}\setminus\{j\}} n_x^* < \sum_{x\in\mathcal{X}\setminus\{j\}} n'_x$. Therefore, $\exists k' \in \mathcal{X}^s$ that $n'_{k'} > n^*_{k'}$. Since $\mathbf{n}'$ and $\mathbf{n}^*$ are equilibrium groupings over $\mathcal{X}^s$, similar to (10), we have

$$\begin{aligned} U(R_k(\theta), n'_k+1) &\geq U(R_k(\theta), n^*_k) \geq U(R_{k'}(\theta), n^*_{k'}+1) \\ &\geq U(R_{k'}(\theta), n'_{k'}) \geq U(R_k(\theta), n'_k+1) \quad (15) \end{aligned}$$

The first and third inequalities are due to $n'_k < n^*_k$ and $n'_{k'} > n^*_{k'}$, and the second and fourth ones come from the equilibrium grouping condition in (5). The equation is valid only when all equalities hold. Thus, if $n'_k < n^*_k$, $\exists k' \in \mathcal{X}^s$ that $U(R_k(\theta), n^*_k) = U(R_{k'}(\theta), n'_{k'})$, which means that we can always find a table k' providing the same utility as $U(R_k(\theta), n^*_k)$ under grouping $\mathbf{n}'$. When $n'_k \geq n^*_k$, we have $U(R_k(\theta), n^*_k) \geq U(R_k(\theta), n'_k)$. Therefore, $\forall k \in \mathcal{X}^s$, $\exists k' \in \mathcal{X}^s$ that $U(R_k(\theta), n^*_k) \geq U(R_{k'}(\theta), n'_{k'})$.

Then, we show that table j is dominated by all other tables under $\mathbf{n}^*$. Since table j is removed by $PC(\mathcal{X}, \mathbf{n}^\mathbf{s}, N)$, we have $n^s_j > n'_j$. Therefore, according the above discussion and the fact that $\mathbf{n}'$ is an equilibrium grouping, we have $\forall k \in \mathcal{X}^s$,

$$\begin{aligned} U\left(R_k(\theta), n^*_k\right) &\geq \min_{k'\in\mathcal{X}^s} U\left(R_{k'}(\theta), n'_{k'}\right) \\ &\geq U\left(R_j(\theta), n'_j+1\right) > U\left(R_j(\theta), n^s_j+1\right). \quad (16) \end{aligned}$$

Since $U(R_j(\theta), n^s_j+1)$ is the highest utility that can be offered by table j, it is dominated by all other tables in $\mathcal{X}^s$ under the final grouping $\mathbf{n}^*$. So, table j is never the best response of the customer. ■

Theorem 4: There always exists a subgame perfect Nash equilibrium with the corresponding equilibrium grouping $\mathbf{n}^*$ satisfying (5) in a sequential Chinese restaurant game.

Proof: We would like to show that the proposed strategy in (13) forms a Nash equilibrium. Suppose customer i chooses table j in his round according to (13). Then, customer i's utility is $u_i = U(R_j(\theta), n^*_j)$ since based on Lemma 2, the equilibrium grouping $\mathbf{n}^*$ will be reached at the end.

Now we show that table j is indeed customer i's best response. Let's assume that customer i is the last customer, i.e, $i = N$, and chooses another table $j' \neq j$ in his round, then his utility becomes $U(R_{j'}(\theta), n^*_{j'}+1)$. However, according to (5), we have

$$u^*_j = U\left(R_j(\theta), n^*_j\right) \geq U\left(R_{j'}(\theta), n^*_{j'}+1\right). \quad (17)$$

Thus, choosing table j is never worse than choosing table j' for customer N.

For the case that customer i is not the last customer, we assume that he chooses table j' instead of table j in his round. Since all customers before customer i follow (13), we have $n_{i,j} \leq n^*_j \forall j \in \mathcal{X}$. Otherwise, $\mathbf{n}^*$ cannot be reached, which contradicts with Lemma 2.

If $n_{i,j'} < n^*_{j'}$, we have $n_{i+1,j'} \leq n^*_{j'}$. In addition, we have $n_{i+1,j} = n_{i,j} \leq n^*_j \forall j \in \mathcal{X} \setminus \{j'\}$, since other tables are not chosen by customer i. Thus, $\mathcal{X}^{i+1,cand} = PC(\mathcal{X}, \mathbf{n_{i+1}}, N)$ and $N^{i,cand} = N$. According to Lemma 2, the final grouping should be $\mathbf{n}^* = EG(\mathcal{X}, N)$. Thus, the new utility of customer i becomes $u'_i = U(R_{j'}(\theta), n^*_{j'})$. However, according to (13), we have

$$\begin{aligned} u_i &= U\left(R_j(\theta), n^*_j\right) \\ &= \arg \max_{x\in\mathcal{X}, n_{i,x}<n^*_x} U\left(R_x(\theta), n^*_x\right) \geq U\left(R_{j'}(\theta), n^*_{j'}\right) \\ &= u'_i. \end{aligned} \quad (18)$$

Thus, choosing table j' never gives customer i a higher utility.

If $n_{i,j'} = n^*_{j'}$, and the final grouping is $\mathbf{n}' = (n'_1, n'_2, \ldots, n'_J)$. Since customer i chooses table j' when $n_{i,j'} = n^*_{j'}$, we have $n'_{j'} \geq n_{i+1,j'} = n_{i,j'} + 1 = n^*_{j'} + 1$. Thus, we have

$$\begin{aligned} u_i &= U\left(R_j(\theta), n^*_j\right) \geq U\left(R_{j'}(\theta), n^*_{j'}+1\right) \geq U\left(R_{j'}(\theta), n'_{j'}\right) \\ &= u'_i, \forall j' \in X, \end{aligned} \quad (19)$$

where the first inequality comes from the equilibrium grouping condition in (5), and the second inequality comes from the fact that $U(R,n)$ is decreasing over n and $n'_{j'} \geq n^*_{j'}+1$. Thus, under both cases, choosing table j' is never better than choosing table j. We conclude that $\{BE^{se}_i(\cdot)\}$ in (13) forms a Nash equilibrium, where the grouping being the equilibrium grouping $\mathbf{n}^*$.

Finally, we show that the proposed strategy forms a Nash equilibrium in every subgame. In Lemma 3, we show that if the table j is removed by $PC(\mathcal{X}, \mathbf{n}^\mathbf{s}, N)$, it is never the best response of all remaining customers. Thus, we only need to consider the remaining table candidates $\mathcal{X}^s = PC(\mathcal{X}, \mathbf{n}^\mathbf{s}, N)$ in the subgame. Then, with Lemma 2, we show that for every possible subgame with corresponding $\mathcal{X}^s$, the equilibrium grouping $\mathbf{n}^* = EG(\mathcal{X}^s, N^s)$ will be achieved at the end of the subgame. Moreover, the above proof shows that if the equilibrium grouping $\mathbf{n}^\mathbf{s}$ will be achieved at the end of the subgame, $BE^{se}_i(\cdot)$ is the best response function. Therefore, the proposed strategy forms a Nash equilibrium in every subgame, i.e., we have a subgame perfect Nash equilibrium. ■

In the proof of Theorem 4, we observe that the sequential game structure brings advantages for those customers making decisions early. According to (13), customers who make decisions early can choose the table providing the largest utility in the equilibrium. When the number of customers choosing that table reaches equilibrium number, the second best table will be chosen until it is full again. For the last customer, he has no choice but to choose the worst one.

V. Imperfect Signal Model: How Learning Evolves

In Section IV, we have showed that in the sequential Chinese restaurant game with perfect signal, customers choosing first have the advantages for getting better tables and thus higher utilities. However, such a conclusion may not be true when the signals are not perfect. When there are uncertainties on the table sizes, customers who arrive first may not choose the right tables, due to which their utilities may be lower. Instead, customers who arrive later may eventually have better chances to get better tables since they can collect more information to make the right decisions. In other words, when signals are not perfect, learning will occur and may result in higher utilities for customers choosing later. Therefore, there is a trade-off between more choices when playing first and more accurate signals when playing later. In this section, we would like to study this trade-off by discussing the imperfect signal model.

In the imperfect signal model, we assume that the system state $\theta \in \Theta = \{1, 2, \ldots, L\}$ is unknown to all N customers. The sizes of J tables can be expressed as functions of θ, which are denoted as $R_1(\theta), R_2(\theta), \ldots, R_J(\theta)$. The prior probability of θ, $\mathbf{g_0} = \{g_{0,1}, g_{0,2}, \ldots, g_{0,J}\}$ with $g_{0,l} = Pr(\theta = l)$, is assumed to be known by all customers. Moreover, each customer receives a private signal $s_i \in \mathcal{S}$, which follows a p.d.f. $f(s|\theta)$. Here, we assume $f(s|\theta)$ is public information to all customers. When conditioning on the system state θ, the signals received by the customers are uncorrelated.

In Chinese restaurant game with imperfect signal model, the customers make decisions sequentially with the decision orders being their numbers. After a customer i made his decision, he cannot change his mind in any subsequent time and his decision and signal are revealed to all other customers. Since signals are revealed sequentially, the customers who make decisions later can collect more information for better estimations of the system state. We assume customers are fully rational, which means they should apply Bayesian learning rule in their decision making process [3]. Therefore, when a new signal is revealed, all customers follow the Bayesian rule to update their believes based on their current believes. Derived from (3), we have the following belief updating function

$$g_{i,l} = \frac{g_{i-1,l} f(s_i|\theta = l)}{\sum_{w \in \Theta} g_{i-1,w} f(s_i|\theta = w)}. \tag{20}$$

A. Best Response of Customers

Since the customers are rational, they will choose the action to maximize their own expected utility conditioning on the information they collect. Let $\mathbf{n_i} = (n_{i,1}, n_{i,2}, \ldots, n_{i,J})$ be the current grouping observed by customer i before he chooses the table, where $n_{i,j}$ is the number of customers choosing table j before customer i. Then, let $\mathbf{h_i} = \{s_1, s_2, \ldots, s_{i-1}\}$ be the history of revealed signals before customer i. In such a case, the best response of customer i can be written as

$$BE_i(\mathbf{n_i}, \mathbf{h_i}, s_i) = \arg\max_j \mathbb{E}\left[U\left(R_j(\theta), n_j\right)|\mathbf{n_i}, \mathbf{h_i}, s_i\right]. \tag{21}$$

From (21), we can see that when estimating the expected utility in the best response function, there are two key terms needed to be estimated by the customer: the system state θ and the final grouping $\mathbf{n} = (n_1, n_2, \ldots, n_J)$. The system state θ is estimated using the concept of belief denoted as $\mathbf{g_i} = \{g_{i,1}, g_{i,2}, \ldots, g_{i,L}\}$ with $g_{i,l} = Pr(\theta = l|\mathbf{h_i}, s_i)$. Since the information on the system state θ in $\mathbf{n_i}$ is fully revealed by $\mathbf{h_i}$, given $\mathbf{h_i}$, $\mathbf{g_i}$ is independent with $\mathbf{n_i}$. Therefore, given the customer's belief $\mathbf{g_i}$, the expected utility of customer i choosing table j becomes

$$\begin{aligned}&\mathbb{E}\left[U\left(R_j(\theta), n_j\right)|\mathbf{n_i}, \mathbf{h_i}, s_i, x_i = j\right]\\ &\quad= \sum_{w \in \Theta} g_{i,w} \mathbb{E}\left[U\left(R_j(w), n_j\right)|\mathbf{n_i}, \mathbf{h_i}, s_i, x_i = j, \theta = w\right].\end{aligned} \tag{22}$$

Note that the decisions of customers $i+1, \ldots, N$ are unknown to customer i when customer i makes the decision. Therefore, a close-form solution to (22) is generally impossible and impractical. In this paper, we purpose a recursive approach to compute the expected utility.

B. Recursive Form of Best Response

Let $BE_{i+1}(\mathbf{n_{i+1}}, h_{i+1}, s_{i+1})$ be the best response function of customer $i+1$. Then, according to $BE_{i+1}(\mathbf{n_{i+1}}, \mathbf{h_{i+1}}, s_{i+1})$, the signal space $\mathcal{S}$ can be partitioned into $\mathcal{S}_{i+1,1}, \ldots, \mathcal{S}_{i+1,J}$ subspaces with

$$\begin{aligned}&\mathcal{S}_{i+1,j}\left(\mathbf{n_{i+1}}, \mathbf{h_{i+1}}\right)\\ &= \left\{s|s \in \mathcal{S}, BE_{i+1}\left(\mathbf{n_{i+1}}, \mathbf{h_{i+1}}, s\right) = j\right\}, \forall j \in \{1, \ldots, J\}.\end{aligned} \tag{23}$$

Based on (23), we can see that, given $\mathbf{n_{i+1}}$ and $\mathbf{h_{i+1}}$, $BE_{i+1}(\mathbf{n_{i+1}}, \mathbf{h_{i+1}}, s_{i+1}) = j$ if and only if $s_{i+1} \in \mathcal{S}_{i+1,j}$. Therefore, the decision of customer $i+1$ can be predicted according to the signal distribution $f(s|\theta)$ given by

$$Pr\left(x_{i+1} = j|\mathbf{n_{i+1}}, \mathbf{h_{i+1}}\right) = \int_{s \in \mathcal{S}_{i+1,j}(\mathbf{n_{i+1}}, \mathbf{h_{i+1}})} f(s) ds. \tag{24}$$

Let us define $m_{i,j}$ as the number of customers choosing table j after customer i (including customer i himself). Then, we have $n_j = n_{i,j} + m_{i,j}$, where n_j denotes the final number of customers choosing table j at the end of the game. Moreover, according to the definition of $m_{i,j}$, we have

$$m_{i,j} = \begin{cases} 1 + m_{i+1,j}, & x_i = j; \\ m_{i+1,j}, & \text{else.} \end{cases} \tag{25}$$

The recursive relation of $m_{i,j}$ in (25) will be used in the following to get the recursive form of the best response function. We first derive the recursive form of the distribution of $m_{i,j}$, i.e., $Pr(m_{i,j} = X|\mathbf{n_i}, \mathbf{h_i}, s_i, x_i, \theta)$ can be expressed as a function of $Pr(m_{i+1,j} = X|\mathbf{n_{i+1}}, \mathbf{h_{i+1}}, s_{i+1}, x_{i+1} = j, \theta = l), \forall l \in \Theta$, $0 \le j \le J$, as in (26) (see equation at bottom of next page) where $\mathbf{h_{i+1}}$ and $\mathbf{n_{i+1}}$ can be obtained using

$$\mathbf{h_{i+1}} = \{h_i, s_i\} \quad \text{and} \quad \mathbf{n_{i+1}} = (n_{i+1,1}, \ldots, n_{i+1,J}), \tag{29}$$

with

$$n_{i+1,k} = \begin{cases} n_{i,k} + 1, & \text{if } x_i = k, \\ n_{i,k}, & \text{otherwise.} \end{cases} \tag{30}$$

Based on (26), $Pr(m_{i,j} = X|\mathbf{n_i}, \mathbf{h_i}, s_i, x_i, \theta = l)$ can be recursively calculated. Therefore, we can calculate the expected utility $\mathbb{E}[U(R_j(\theta), n_j)|\mathbf{n_i}, \mathbf{h_i}, s_i]$ by (27). Finally, the best response function of customer i can be derived by (28).

With the recursive form, the best response function of all customers can be obtained using backward induction. The best response function of the last customer N can be found as

$$BE_N(\mathbf{n_N}, h_N, s_N) = \arg\max_j \sum_{l \in \Theta} g_{N,l} u\left(R_j(l), n_{N,j} + 1\right). \tag{31}$$

Note that $Pr(m_{N,j} = X|\mathbf{n_N}, \mathbf{h_N}, s_N, x_N, \theta)$ can be easily derived as follows:

$$Pr(m_{N,j} = 1|\mathbf{n_N}, \mathbf{h_N}, s_N, x_N, \theta) = \begin{cases} 1, & \text{if } x_N = j, \\ 0, & \text{otherwise.} \end{cases} \tag{32}$$

As of the convergence of the recursive best response, which is based on the traditional backward induction technique, it definitely converges since this game has finite players. As a Chinese restaurant game with N players, only N recursive calls are required to derive all the best responses.

VI. Simulation Results

In this section, we verify the proposed recursive best response and corresponding equilibrium. We simulate a Chinese restaurant with two tables $\{1,2\}$ and two possible states $\theta \in \{1,2\}$. When $\theta=1$, the size of table 1 is $R_1(1)=100$ and the size of table 2 is $R_2(1)=100r$, where r is the ratio of table sizes. When $\theta=2$, $R_1(2)=100r$ and $R_2(2)=100$. The state is randomly chosen with $Pr(\theta=1)=Pr(\theta=2)=0.5$. The number of customers is fixed. Each customer receives a randomly generated signal s_i at the beginning of the simulation. The signal distribution $f(s|\theta)$ is given by $Pr(s=1|\theta=1)=Pr(s=2|\theta=2)=p$, $Pr(s=2|\theta=1)=Pr(s=1|\theta=2)=1-p$, where $p \geq 0.5$ can be regarded as the quality of signals. When the signal quality p is closer to 1, the signal is more likely to reflect the true state θ. With the signals, customers make their decisions sequentially. After the i-th customer makes his choice, he reveals his decision and signal to other customers. The game ends after the last customer made his decision. Then, the utility of the customer i choosing table j is given by $U(R_j(\theta),n_j)=\frac{R_j}{n_j}$, where n_j is the number of customers choosing table j in the end.

A. Advantage of Playing Positions vs. Signal Quality

We first investigate how the decision order and quality of signals affect the utility of customers. We fix the size of one table as 100. The size of the other table is $r\times 100$, where r is the ratio of the table sizes. In the simulations, we assume the ratio $r \in [0,1]$. When the ratio $r=1$, two tables are identical, but the utility of choosing each table may be different since we may have odd customers. When $r=0$, one table has a size of 0, which means a customer has a positive utility only when he chooses the correct table.

Due to the complicated game structure in Chinese restaurant game, the effect of signal quality and table size ratio is generally non-linear. As shown in Fig. 1(a), when the number of customers is 5, customer 5 has the largest utility when the signal quality is high and the table size ratio is low, while customer 1 has the largest utility when the signal quality is low and the table size ratio is high. This phenomenon can be explained as follows. When the table size ratio is lower, all customers desire the larger table since even all of them select the larger one, each of them still have a utility larger than choosing the smaller one. In such a case, customers who choose late would have advantages since they have collected more signals and have a higher probability to identify the large table. Nevertheless, when the signal quality is low, even the last customer cannot form a strong belief on the true state. In such a case, the expected size of each table becomes less significantly, and customers' decisions rely more on the negative network externality effect, i.e., how crowded of each table. In such a case, the first customer has the advantage to choose the table with fewer customers in expectation.

However, we observe that in some cases, customer 3 becomes the one with largest utility. The reasons behind this phenomenon is as follows. In these cases, the expected number of customers in the larger table is 3, and this table provides the customers a larger utility at the equilibrium. Therefore, customers would try to identify this table and choose it according to their own believes. Since customer 3 collects more signals than customers 1 and 2, he is more likely to identify the correct table. Moreover, since he is the third customer, this table is always available to him. Therefore, customer 3 has the largest expected utility in these cases.

Note that the expected table size is determined by both the signal quality and the table size ratio. Generally, when the signal quality is low, a customer is less likely to construct a strong belief on the true state, i.e., the expected table sizes of both tables are similar. This suggests that a lower signal quality has a similar effect on the expected table size as a higher table size ratio. Our arguments are supported by the concentric-like structure shown in Fig. 1(a). The same arguments can be applied to the 10-customer scheme, which is shown in Fig. 1(b). We can observe the similar concentric-like structure. Additionally, we observe that when the table size ratio increases, the order of customer who has the largest utility in the peaks decreases from 10 to 5. This is

$$
\begin{aligned}
&Pr\left(m_{i,j}=X|\mathbf{n_i},\mathbf{h_i},s_i,x_i,\theta=l\right)\\
&\quad=\begin{cases} Pr\left(m_{i+1,j}=X-1|\mathbf{n_i},\mathbf{h_i},s_i,x_i,\theta=l\right), & x_i=j,\\ Pr\left(m_{i+1,j}=X|\mathbf{n_i},\mathbf{h_i},s_i,x_i,\theta=l\right), & x_i\neq j,\end{cases}\\
&\quad=\begin{cases} \sum_{u\in\{1,\ldots,J\}}\int_{s\in\mathcal{S}_{i+1,u}(\mathbf{n_{i+1}},\mathbf{h_{i+1}})} Pr\left(m_{i+1,j}=X-1|\mathbf{n_{i+1}},\mathbf{h_{i+1}},s_{i+1}=s,x_{i+1}=u,\theta=l\right)f(s|\theta=l)ds, & x_i=j,\\ \sum_{u\in\{1,\ldots,J\}}\int_{s\in\mathcal{S}_{i+1,u}(\mathbf{n_{i+1}},\mathbf{h_{i+1}})} Pr\left(m_{i+1,j}=X|\mathbf{n_{i+1}},\mathbf{h_{i+1}},s_{i+1}=s,x_{i+1}=u,\theta=l\right)f(s|\theta=l)ds, & x_i\neq j,\end{cases}
\end{aligned}
\tag{26}
$$

$$
\begin{aligned}
&\mathbb{E}\left[U\left(R_j(\theta),n_j\right)|\mathbf{n_i},\mathbf{h_i},s_i\right]\\
&\quad=\sum_{l\in\Theta}\sum_{x=0}^{N-i+1} g_{i,l}Pr\left(m_{i,j}=x|\mathbf{n_i},\mathbf{h_i},s_i,x_i=j,\theta=l\right)U\left(R_j(l),n_{i,j}+x\right).
\end{aligned}
\tag{27}
$$

$$
\begin{aligned}
&BE_i(\mathbf{n_i},\mathbf{h_i},s_i)\\
&\quad=\arg\max_j\sum_{l\in\Theta}\sum_{x=0}^{N-i+1} g_{i,l}Pr(m_{i,j}=x|\mathbf{n_i},\mathbf{h_i},s_i,x_i=j,\theta=l)U\left(R_j(l),n_{i,j}+x\right).
\end{aligned}
\tag{28}
$$

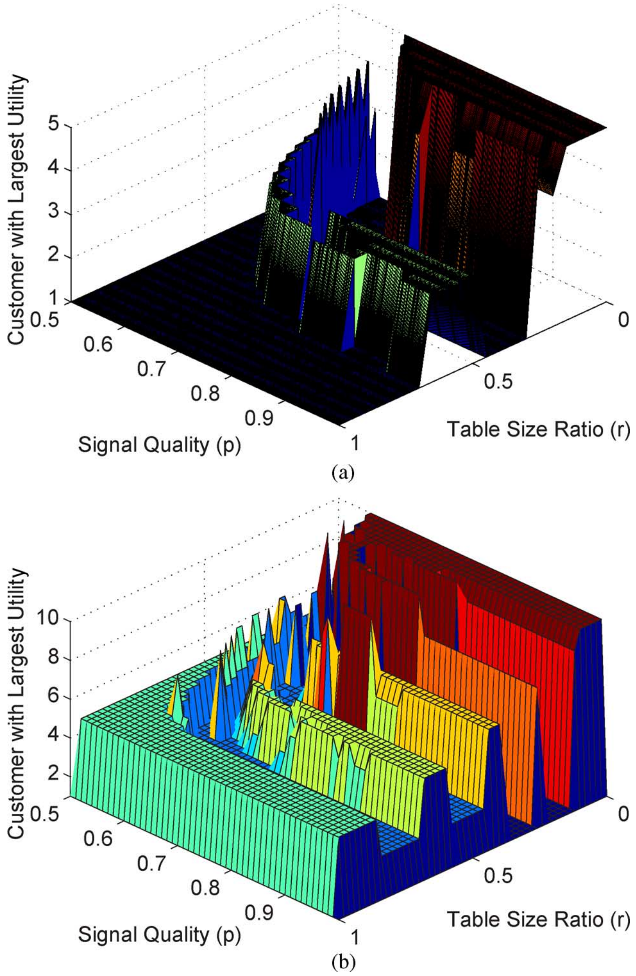


Fig. 1. The effect of different table size ratio and signal quality. (a) 5 customers; (b) 10 customers.

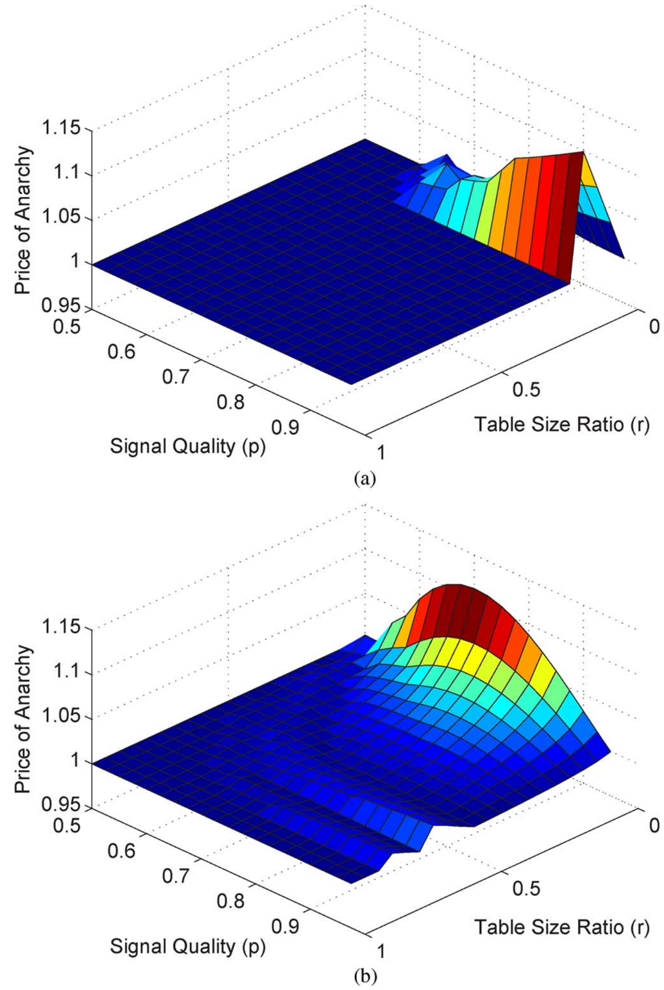


Fig. 2. Price of anarchy with different utility functions. (a) $U(R,n)=\frac{R}{n}$; (b) $U(R,n)=log\left(\frac{R}{n+10}\right)$.

consistent with our arguments since when the table size ratio increases, the equilibrium number of customers in the large table decreases from 10 to 5. This also explains why customer 1 does not have the largest utility when the table size ratio is high. In this case, the equilibrium number of customers in the large table is 5, and the large table provides higher utilities to customers in the equilibrium. Since customer 5 can collect more signals than previous customers, he has better knowledge on the table size than customer 1 to 4. Moreover, since customer 5 is the fifth one to choose the table, he always has the opportunity to choose the large table. In such a case, customer 5 has the largest expected utility when the table size ratio is high.

B. Price of Anarchy

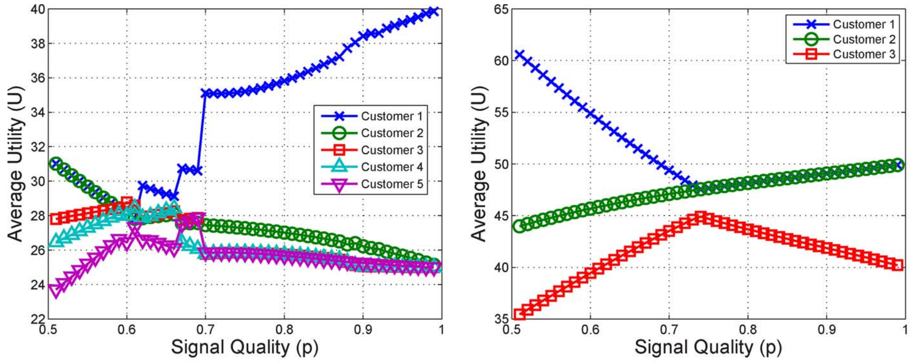
We then investigate the efficiency of the equilibrium grouping in Chinese restaurant game using price of anarchy, which is a popular measurement in game theory on the degradation of the system efficiency due to rational behaviors of players. Basically, the price of anarchy in a game-theoretic system is defined as the ratio of the social welfare under worst equilibrium in the system to the one under the centralized-optimal solution. Therefore, when the price of anarchy is close or equal to 1, the rational behaviors generally do not incur efficiency loss to the system.

We first define the social welfare function $W(\mathbf{B})$ in Chinese restaurant game as the sum of customers expected utilities, that is, $W(\mathbf{B})=\mathbb{E}[\sum_{i=1}^{N}U(R_{x_i}(\theta),n_{x_i})|\mathbf{B}]$, where $\mathbf{B}$ denotes the strategies of customers applied in Chinese restaurant game. Let $\mathcal{B}^{\mathcal{U}}$ be the universal set of all possible strategies and $\mathcal{B}^{\mathcal{E}}$ be the set of all equilibria in Chinese restaurant game, then the price of anarchy is defined as follows:

$$PoA=\frac{\max_{\mathbf{B}\in\mathcal{B}^{\mathcal{U}}}W(\mathbf{B})}{\min_{\mathbf{B}'\in\mathcal{B}^{\mathcal{E}}}W(\mathbf{B}')}. \tag{33}$$

We simulate a 5-customer restaurant with two tables and two states. All other settings are the same as the ones in Section IV-A except the utility function. In this simulation, we apply two utilities functions: $U(R,n)=\frac{R}{N}$ and $U(R,n)=log\left(\frac{R}{n+10}\right)$. The former represents the case that the resource is equally shared, while the latter roughly represents the SINR-throughput in wireless networks. The centralized-optimal solution is found through exhaustive search. The prices of anarchy under all combinations of signal quality and table size ratio are shown in Fig. 2.

As shown in Fig. 2(a), when the utility function is set as $\frac{R}{n}$, the price of anarchy is equal to one under most combinations



(a) (b)

Signals	Best Response	
s_1, s_2, s_3	$p = 0.9$	$p = 0.6$
2,2,2	2,2,1	1,2,2
1,2,2	1,2,2	2,1,2
2,1,2	2,1,2	1,2,2
1,1,2	1,1,2	2,1,1
2,2,1	2,2,1	1,2,2
1,2,1	1,2,1	2,1,1
2,1,1	2,1,1	1,2,1
1,1,1	1,1,2	2,1,1

(c)

Fig. 3. Average utility of customers in resource pool scenario when $r = 0.4$. (a) 5 customers; (b) 3 customers; (c) best response when $\mathrm{N} = 3$.

except when the table size ratio is close to 0. The reason that the price of anarchy is larger than 1 at these points is that the smaller table is so small that all customers have a higher utility even sharing the larger one. In such cases, the small table will not be chosen, and the resource provided by this table is lost due to the rational behaviors of customers. For the scenario that the utility function is set as $U(R,n) = log\left(\frac{R}{n+10}\right)$, the price of anarchy never exceeds 1.06 (Fig. 2(b)). This is because in such a scenario, a proper balance in loadings on tables will greatly increase the social welfare, which is automatically achieved by the rational choices of customers due to their concerns on negative network externality. Therefore, the rational behaviors in Chinese restaurant game generally does not harm much on the system efficiency, and the equilibrium we found is efficient even compared with the centralized-optimal solution.

C. Case Study: Resource Pool and Availability Scenarios

Finally we discuss two specific scenarios: the resource pool scenario with $r = 0.4$ and available/unavailable scenarios with $r = 0$. In resource pool scenario, the table size of the second table is 40. In available/unavailable scenario, the second table size is 0, which means that a customer has positive utility only when he chooses the right table. For both scenarios, we examine the schemes with $N = 3$ and $N = 5$.

From Fig. 3, we can see that in the resource pool scenario with $r = 0.4$, customer 1 on average has significant higher utility, which is consistent with the result in Fig. 1(a). Using 5-customer scheme shown in Fig. 3(a) as an example, the advantage of playing first becomes significant when signal quality is very low ($p < 0.6$), or the signal quality is high ($p > 0.7$). We also find that customer 5 has the lowest average utility for most signal quality p. We may have a clearer view on this in the 3-customer scheme. We list the best response of customers given the received signals in Fig. 3(c). We observe that when signal quality p is large, both customer 1 and 2 follow the signals they received to choose the tables. However, customer 3 does not follows his signal if the first two customers choose the same table. Instead, customer 3 will choose the table that is still empty. In this case, although customer 3 may know which table is larger, he does not choose that table if it has been occupied by the first two customers. The network externality effect dominates the learning advantage in this case.

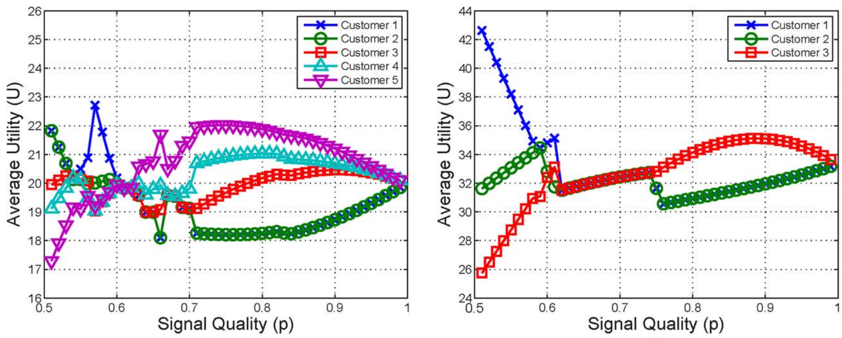
However, when p is low, the best response of customer 1 is opposite, i.e., he will choose the table that is indicated as the smaller one by the signal he received. At the first glance, the best response of customer 1 seems to be unreasonable. However, such a strategy is indeed customer 1's best response considering the expected equilibrium in this case. According to Theorem 4, if perfect signals ($p = 1$) are given, the large table should be chosen by customer 1 and 2 since the utility of large table, which is $\frac{100}{2} = 50$, is larger than the that of the small table, which is $\frac{40}{1} = 40$, in the equilibrium. However, when the imperfect signals are given, customers choose the tables based on the expected table sizes. When signal quality is low, the uncertainty on the table size is large, which leads to similar expected table sizes for both tables. In such a case, customer 1 favors the smaller table because it can provide a higher expected utility, compared with sharing with another customer in the larger table.

In the available/unavailable scenario, as shown in Fig. 4, the advantage of customer 1 in playing first becomes less significant. Using 5-customer scheme shown in Fig. 4(a) as an example, when signal quality p is larger than 0.6, customer 5 has the largest average utility and customer 1 has smallest average utility. Such a phenomenon is because customers should try their best on identifying the available table when $r = 0$. Learning from previous signals gives the later customers a significant advantage in this case.

Nevertheless, we observe that the best responses of later customers are not necessary always choosing the table that is more likely to be available. We use the 3-customer as an illustrative example. We list the best response of all customers given the received signals in Fig. 4(c). When the signal quality is pretty low ($p = 0.55$), we have the same best response as the one in resource pool scenario, where the network externality effect still plays a significant role. Using $(s_1, s_2, s_3) = (2, 2, 1)$ as an example, even customer 3 finds that table 2 is more likely to be available, his best response is still choosing table 1 since table 2 is already chosen by both customer 1 and 2, and the expected utility of choosing table 1 with only himself is higher than that of choosing table 2 with other two customers. As the signal quality p becomes high, e.g., $p = 0.9$, customer 3 will choose the table according to all signals s_1, s_2, s_3 he collected since the belief constructed by the signals is now strong enough to overcome the loss in the network externality effect.

VII. Application: Cooperative Spectrum Access in Cognitive Radio Networks

We would like to illustrate an important application of Chinese restaurant game: cooperative spectrum access in cognitive radio networks. Traditional dynamic spectrum access methods



Signals	Best Response		
s_1, s_2, s_3	$p = 0.9$	$p = 0.7$	$p = 0.55$
2,2,2	2,2,2	2,2,2	1,2,2
1,2,2	1,2,2	1,2,2	2,1,2
2,1,2	2,1,2	2,1,2	1,2,2
1,1,2	1,1,1	1,1,2	2,1,1
2,2,1	2,2,2	2,2,1	1,2,2
1,2,1	1,2,1	1,2,1	2,1,1
2,1,1	2,1,1	2,1,1	1,2,1
1,1,1	1,1,1	1,1,1	2,1,1

(c)

Fig. 4. Average utility of customers in available/unavailable scenario when $r = 0$. (a) 5 customers; (b) 3 customers; (c) best response when $N = 3$.

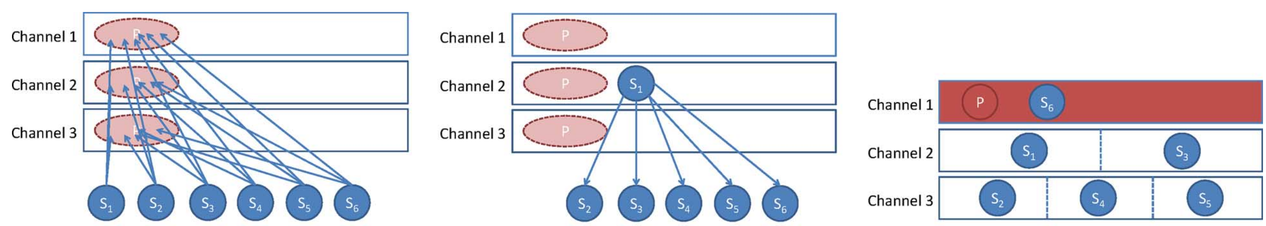


Fig. 5. Sequential cooperative spectrum sensing and accessing. (a) Channel sensing; (b) channel selection and signal broadcast; (c) data transmission.

focus on identifying available spectrum through spectrum sensing. Cooperative spectrum sensing is a potential scheme to enhance the accuracy and efficiency of detecting available spectrum [29]–[31]. In cooperative spectrum sensing, the sensing results from the secondary users are shared by all members within the same or neighboring networks. These secondary users then use the collected results to make spectrum access decisions collaboratively or individually. If the sensing results are independent from each other, the cooperative spectrum sensing can significantly increase the accuracy of detecting the primary user's activity. Secondary users can learn from others' sensing results to improve their knowledge on the primary user's activity. After the available spectrum is detected, secondary users need to share the spectrum following some predetermined access policy. In general, the more secondary users access the same channel, the less available access time for each of them, i.e., a negative network externality exists in this problem. Therefore, before making decision on spectrum access, a secondary user should estimate both the primary user's activity and the possible number of secondary users accessing the same spectrum.

A. System Model

We consider a cognitive radio system with J channels, N secondary transmitter-receiver pairs, and one primary user. We assume that the spectrum access behavior of secondary users is organized by an access point through a control channel. Through the organization, the secondary users can synchronize their channel sensing and selection time. Suppose that the primary user is always active and transmitting some data on one of the channels. In addition, the primary user's access time is slotted. At each time slot, each channel has equal probability of $\frac{1}{J}$ to be selected by the primary user for transmission. The secondary users' activities are shown in Fig. 5. At the beginning of each time slot, secondary users (transmitters) individually perform sensing on all channels $1 \sim J$. Then, they follow a predefined order to sequentially determine which channel they are going to access in this time slot. Without loss of generality, we assume they follow the same order as their indices. When making a decision, a secondary user i reports his decision and the sensing result to the access point through a pre-allocated control channel. At the same time, all secondary users also receive this report by overhearing. After all secondary users have made their decisions, the access point announces the access policy of each channel through the control channel: secondary users choosing the same channel equally share the slot time. However, if the channel is occupied by the primary user, their transmission will fail due to the interference from primary user's transmission.

Such a cognitive radio system can be modeled as a sequential Chinese restaurant game. Let H_j be the hypothesis that channel j is occupied by the primary user. Then, let the sensing results of secondary user $i \in \{1, 2, \ldots, N\}$ on channel $j \in \{1, 2, \ldots, J\}$ be $s_{i,j}$. We use a simple binary model on the sensing result in this example, where $s_{i,j} = 1$ if the secondary user detected some activities on channel j and $s_{i,j} = 0$ if no activity is detected on channel j. For secondary user i, his own sensing results are denoted as $\mathbf{s_i} = (s_{i,1}, s_{i,2}, \ldots, s_{i,J})$. In addition, the results he collected from the reports of previous users are denoted as $\mathbf{h_i} = \{\mathbf{s_1}, \mathbf{s_2}, \ldots \mathrm{s}_{\mathrm{i}-1}\}$.

We define the belief of a secondary user i on the occupation of channels as $\mathbf{g_i} = \{g_{i,1}, g_{i,2}, \ldots, g_{i,J}\}$, where $g_{i,j} = Pr(H_j|\mathbf{h_i}, \mathbf{s_i})$. Let the probability of false alarm and miss detection of the sensing technique on a single channel be p_f and p_m, respectively. The probability of $\mathbf{s_i}$ conditioning on H_j is given by

$$Pr(\mathbf{s_i}|H_j) = p_m^{1-s_{i,j}}(1-p_m)^{s_{i,j}} \prod_{k\in\{1,\ldots,J\}\setminus\{j\}} p_f^{s_{i,k}}(1-p_f)^{1-s_{i,k}}. \quad (34)$$

Thus, we have the following belief updating rule

$$g_{i,j} = \frac{g_{i-1,j} Pr(\mathbf{s_i}|H_j)}{\sum_{k=1}^{J} g_{i-1,k} Pr(\mathbf{s_i}|H_k)}. \quad (35)$$

With this rule, the belief of secondary user i is updated when a new sensing result is reported to the access point. The available access time of a channel j within a slot is its slot time, which is denoted as T. However, if the channel occupied by primary user, its access time becomes 0. Thus, we define the access time of channel j as

$$R_j(H_k) = \begin{cases} 0, & j = k. \\ T, & \text{otherwise}. \end{cases} \quad (36)$$

Then, let x_i be secondary user i's choice on the channels, and n_j be the number of secondary users choosing channel j. We define the utility of a secondary user i as

$$u_i = U(x_i) = \frac{Q_{x_i} R_{x_i}(\theta)}{n_{x_i}}, \quad (37)$$

where $\theta \in \{H_j\}$ is the hypothesis to be true and Q_{x_i} is the channel quality of channel x_i. Here we assume that the secondary users are close to each other and share the similar channel conditions that are mainly determined by the external interference and background noise. The differences in channel gains are mainly influenced by the frequency or time-dependent external interference. If the channel has higher quality, the secondary users choosing the channel have higher data rates, and thus higher utility. Then, the best response of secondary user i is as follows,

$$BE_i(\mathbf{n_i}, \mathbf{h_i}, \mathbf{s_i}) = \arg\max_x \sum_{k \in \{1,2,\ldots,J\} \setminus \{x\}} g_{i,k} E\left[\frac{Q_x T}{n_x} | \mathbf{n_i}, \mathbf{h_i}, \mathbf{s_i}, H_k\right]. \quad (38)$$

This best response function can be solved recursively through the recursive equations in (26) and (28).

B. Simulation Results

We simulate a cognitive radio network with 3 channels, 1 primary user, and 7 secondary transmitter-receiver pairs. When the channel is not occupied by the primary user, the available access time for secondary users in one time slot is 100 ms. Secondary users (transmitters) sense the primary user's activity in all three channels at the beginning of the time slot. We assume that the primary user has equal probability to occupy one of three channels. Conditioning on the primary user's occupation of the channel, the probabilities of miss detection and false alarm in sensing one channel are 0.1. The channel quality factor of channel 1 is $Q_1 = 1$, while channel 2 and 3 are $1 - d$ and $1 - 2d$. The d is the degraded factor, which is within [5%, 50%] in the simulations.

We compare our best response strategy in (28) with the following four strategies: random, signal, learning, and myopic strategies. In the random strategy, customers choose their strategies randomly and uniformly, i.e., all J tables have equal probability of $\frac{1}{J}$ to be chosen under the random strategy. In the signal strategy, customers make their decisions purely based on their own signal. Information from other customers, including the revealed signals and their choices on tables, is ignored. The objective of signal strategy is to choose the largest expected table size conditioning on his signal given by

$$x_i^{signal} = \arg\max_x \sum_{l \in \Theta} Pr(\theta = l|\mathbf{s_i}, \mathbf{g_0}) Q_x R_x(l). \quad (39)$$

The learning strategy is an extension of the signal strategy. Under this strategy, the customer learns the system state not only by his own signal but also by the signals revealed by the previous customers. Therefore, the learning strategy can be obtained as

$$x_i^{learn} = \arg\max_x \sum_{l \in \Theta} g_{i,l} Q_x R_x(l), \quad (40)$$

where $g_{i,l} = Pr(\theta = l|\mathbf{h_i}, \mathbf{s_i}, \mathbf{g_0})$ is the belief of the customer on the state.

Finally, the myopic strategy simulates the behavior of a myopic player. The objective of a customer under myopic strategy is maximizing his current utility, i.e., the customer makes the decision according to his own signal, all revealed signals, and the current grouping $\mathbf{n_i}$ as follows,

$$x_i^{myopic} = \arg\max_x \sum_{l \in \Theta} g_{i,l} \frac{Q_x R_x(l)}{n_{i,x} + 1}. \quad (41)$$

From (41), we can see that the myopic strategy is similar to the proposed best response strategy except the Bayesian prediction of the subsequent customers' decisions. The performance of all these four strategies will be evaluated in all simulations in this application . They will be treated as the baseline of the system performance without fully rational behaviors of customers.

The simulation results are shown in Fig. 6. From Fig. 6(a), 6(b), and 6(d), we can see that secondary users have different utilities under different orders and schemes. For both the myopic and the proposed best response schemes, secondary user 3 has a larger utility than secondary user 1 when the degraded factor is low. This is because secondary user 3 has the advantages in collecting more signals than secondary 1 to identify the channel occupied by the primary user. Moreover, the loadings of the other two channels are still far from their expected equilibrium loadings since only two secondary users have made choices. Therefore, secondary user 3 has a larger utility than secondary user 1. Nevertheless, when the degraded factor is high, we can see that secondary user 1's utility is larger than that of secondary user 3. This is because when the degraded factor increases, the quality difference among channels increases. In such a case, even secondary user 3 successfully identify the occupied channel, and the channel that offers a higher utility in the equilibrium is usually the one with fewer number of secondary users. The expected number of secondary users accessing such a channel is generally 2 or even 1, and secondary user 3 can no longer freely choose those channels. For secondary user 7, he usually has no choice since there are six secondary users making decisions before him. Therefore, he has the smallest utility.

Generally, the myopic scheme provides an equal or lower utility than the best response scheme for secondary users

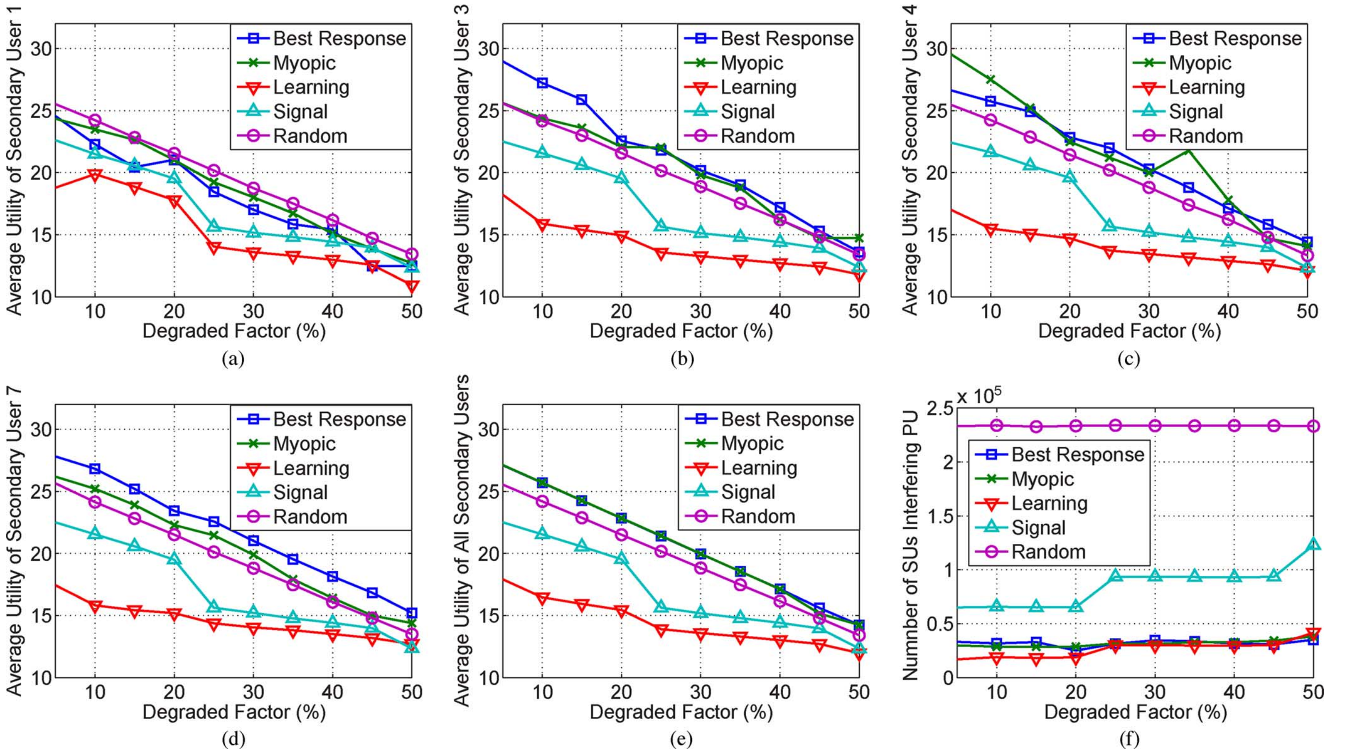


Fig. 6. Spectrum accessing in cognitive radio network under different schemes. (a) Secondary user 1; (b) secondary user 3; (c) secondary user 4; (d) secondary user 7; (e) average utility; (f) SUs interfering PU.

making decisions early, such as secondary user 1, since secondary users in the myopic scheme do not predict the decisions of subsequent users. However, some secondary users eventually benefit from the mistakes made by early secondary users. We can see from Fig. 6(b) and Fig. 6(c) that for some cases, customer 3 and 4 has a higher utility under the myopic scheme than under the best response scheme due to the mistakes made by customer 1 and 2. We can also see from Fig. 6(e) that both best response and myopic schemes provides the same average utilities of all secondary users. In such a case, the utility loss of some secondary users in the myopic scheme will lead to the utility increase of some other secondary users. For random and signal schemes, there is no difference among the average utilities of secondary user 1, 3, and 7 since secondary users do not learn from other agents' actions and signals under these two schemes. For the learning scheme, we can see that secondary user 1 has a significantly larger utility than secondary user 3 and 7. This is because in the learning scheme, secondary users do not take the negative network externality into account when making decisions on the channel selection. Since secondary users who made decisions later are more likely to identify the primary user's activity, they are more likely to choose the same channels, and their utilities are degraded due to the negative network externality.

Let us take a deeper look at the average utility of all secondary users shown in Fig. 6(e). On one hand, we can see that both best response and myopic schemes achieve highest average utilities of all secondary users. The network externality effects in spectrum access force strategic secondary users to access different channels instead of accessing the same high quality channels. On the other hand, learning and signal schemes lead to poor average utilities since they do not consider the network externality in their decision processes. All secondary users tend to access the same available high quality channel, and therefore the spectrum resource in other available channels is wasted. This also explains the phenomenon that learning scheme leads to poorer performance than signal scheme. Under the learning scheme, secondary users are more likely to reach a consensus on the primary user's activity and make the same choice on the channels, which degrades the overall system performance.

Finally, we show the number of secondary users causing interference to the primary user in Fig. 6(f). We can see that those schemes involving learning, which are best response, myopic, and learning schemes, have low interference to the primary user. Secondary users who learn from others' signals efficiently avoid the channel occupied by the primary user.

VIII. Conclusion

In this paper, we proposed a new game, called sequential Chinese restaurant game, by combining the strategic game-theoretic analysis and non-strategic machine learning technique. The proposed Chinese restaurant game can provide a new general framework for analyzing the strategic learning and predicting behaviors of rational agents in a social network with negative network externality. By conducting the analysis on the proposed game, we derived the optimal strategy for each agent and provided a recursive method to achieve the equilibrium. The tradeoff between two contradictory advantages, which are making decisions earlier for choosing better tables and making decisions later for learning more accurate believes, is discussed through simulations. We found that both the signal quality of the unknown system state and the table size ratio affect the expected utilities of customers with different decision orders. Generally, when the signal quality is low and the table size ratio is high, the advantage of playing first dominates the benefit from learning. On the contrary, when the signal quality is high and the table size ratio is low, the advantage of playing later for better knowledge on the true state increases the expected utility of later agents. Our simulations also showed that the price of anarchy under Chinese restaurant game is close to one, which suggests that the efficient loss due to the rational behaviors of customers is close to zero. The small price of anarchy is achieved by the loading balance among tables, which is automatically achieved in Chinese restaurant game. Finally, we illustrated a specific application of Chinese restaurant game in wireless networking: the cooperative spectrum access problem in cognitive radio networks. We showed that the overall channel utilization can be improved by taking the negative network externality into account in secondary users' decision process. The interference from secondary users to the primary user can also be reduced through learning from the sensing results of others.

References

[1] V. Bala and S. Goyal, "Learning from neighbours," *Rev. Econ. Studies*, vol. 65, no. 3, p. 595, 1998.

[2] B. Golub and M. O. Jackson, "Naive learning in social networks and the wisdom of crowds," *Amer. Econom. J.: Microeconomics*, vol. 2, no. 1, pp. 112–149, 2010.

[3] D. Acemoglu, M. A. Dahleh, I. Lobel, and A. Ozdaglar, "Bayesian learning in social networks," *Rev. Econom. Studies*, vol. 78, no. 4, pp. 1201–1236, 2011.

[4] D. Acemoglu and A. Ozdaglar, "Opinion dynamics and learning in social networks," *Dyn. Games Appl.*, vol. 1, pp. 3–49, 2011.

[5] R. W. Cooper, *Coordination Games: Complementarities and Macroeconomics*. Cambridge, U.K.: Cambridge Univ. Press, 1999.

[6] J. Wit, "Social learning in a common interest voting game," *Games Econ. Behav.*, vol. 26, no. 1, pp. 131–156, 1999.

[7] M. Battaglini, "Sequential voting with abstention," *Games Econ. Behav.*, vol. 51, no. 2, pp. 445–463, 2005.

[8] S. Nageeb Ali and N. Kartik, *Observational Learning With Collective Preferences. Manuscript*. New York: Columbia Univ., 2010.

[9] D. Gale, "Dynamic coordination games," *Econ. Theory*, vol. 5, pp. 1–18, 1995.

[10] A. Dasgupta, "Social learning with payoff complementarities," Working Paper, 2000 [Online]. Available: http://personal.lse.ac.uk/DASGUPT2/research.html

[11] A. Dasgupta, "Coordination and delay in global games," *J. Econ. Theory*, vol. 134, no. 1, pp. 195–225, 2007.

[12] S. Choi, D. Gale, S. Kariv, and T. Palfrey, "Network architecture, salience and coordination," *Games Econ. Behav.*, vol. 73, no. 1, pp. 76–90, 2011.

[13] M. L. Katz and C. Shapiro, "Technology adoption in the presence of network externalities," *J. Politic. Econ.*, pp. 822–841, 1986.

[14] W. H. Sandholm, "Negative externalities and evolutionary implementation," *Rev. Econ. Studies*, vol. 72, no. 3, pp. 885–915, 2005.

[15] 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.

[16] S.-J. Kim and G. B. Giannakis, "Optimal resource allocation for MIMO ad hoc cognitive radio networks," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 3117–3131, May 2011.

[17] T. M. Mitchell, *Machine Learning*. New York: McGraw-Hill, 1997, ISBN 0070428077.

[18] D. Aldous, I. Ibragimov, J. Jacod, and D. Aldous, "Exchangeability and related topics," in *Lecture Notes in Mathematics*. Berlin, Germany: Springer, 1985, vol. 1117, pp. 1–198.

[19] J. Pitman, "Exchangeable and partially exchangeable random partitions," *Probab. Theory Related Fields*, vol. 102, no. 2, pp. 145–158, 1995.

[20] C.-Y. Wang, Y. Chen, and K. J. R. Liu, "Chinese restaurant game," *IEEE Signal Process. Lett.*, vol. 19, no. 12, pp. 898–901, 2012.

[21] H. Carlsson and E. Van Damme, "Global games and equilibrium selection," *Econometrica: J. Econometric Soc.*, pp. 989–1018, 1993.

[22] S. Morris and H. Shin, "Global games: Theory and applications," presented at the Cowles Foundation Discussion, 2001, Paper No. 1275R.

[23] G. M. Angeletos and I. Werning, "Crises and prices: Information aggregation, multiplicity, and volatility," *Amer. Econ. Rev.*, pp. 1720–1736, 2006.

[24] V. Krishnamurthy, "Decentralized activation in sensor networks—Global games and adaptive filtering games," *Digit. Signal Process.*, vol. 21, no. 5, pp. 638–647, 2011.

[25] V. Krishnamurthy, "Decentralized spectrum access amongst cognitive radios: An interacting multivariate global game-theoretic approach," *IEEE Trans. Signal Process.*, vol. 57, no. 10, pp. 3999–4013, Oct. 2009.

[26] G. M. Angeletos, C. Hellwig, and A. Pavan, "Signaling in a global game: Coordination and policy traps," *J. Pol. Econ.*, vol. 114, no. 3, pp. 452–484, 2006.

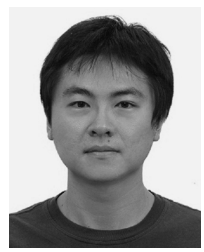
[27] G. M. Angeletos, C. Hellwig, and A. Pavan, "Dynamic global games of regime change: Learning, multiplicity, and the timing of attacks," *Econometrica*, vol. 75, no. 3, pp. 711–756, 2007.

[28] J. S. Costain, "A herding perspective on global games and multiplicity," *BE J. Theoretic. Econ.*, vol. 7, no. 1, p. 22, 2007.

[29] S. M. Mishra, A. Sahai, and R. W. Brodersen, "Cooperative sensing among cognitive radios," in *Proc. IEEE Int. Conf. Commun.*, 2006, vol. 4, pp. 1658–1663.

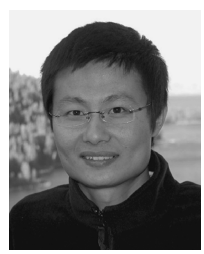
[30] B. Wang, K. J. R. Liu, and T. C. Clancy, "Evolutionary cooperative spectrum sensing game: How to collaborate?," *IEEE Trans. Commun.*, vol. 58, no. 3, pp. 890–900, 2010.

[31] K. J. R. Liu and B. Wang, *Cognitive Radio Networking and Security: A Game-theoretic View*. Cambridge, U.K.: Cambridge Univ. Press, 2010.



Chih-Yu Wang (S'97) received the B.S. degree in electrical engineering from the National Taiwan University, Taipei, Taiwan. in 2007.

He has been a visiting student in the University of Maryland, College Park, in 2011. He is currently working toward the Ph.D. degree in the Graduate Institute of Communication Engineering, National Taiwan University. His research interests mainly are applications of game theory in wireless networking and social networking.



Yan Chen (S'06–M'11) received the Bachelor's degree from the University of Science and Technology of China in 2004, the M.Phil. degree from the Hong Kong University of Science and Technology (HKUST) in 2007, and the Ph.D. degree from the University of Maryland, College Park, in 2011.

He is currently a Research Associate in the Department of Electrical and Computer Engineering at the University of Maryland, College Park. His current research interests are in social learning and networking, smart grid, cloud computing, crowd sourcing, network economics, multimedia signal processing, and communication.

Dr. Chen received the University of Maryland Future Faculty Fellowship in 2010, the Chinese Government Award for outstanding students abroad in 2011, and the University of Maryland ECE Distinguished Dissertation Fellowship Honorable Mention in 2011, and was the Finalist of the A. James Clark School of Engineering Dean's Doctoral Research Award in 2011.

K. J. Ray Liu (F'03) received the B.S. degree from the National Taiwan University in 1983 and the Ph.D. degree from the University of California, Los Angeles (UCLA), in 1990, both in electrical engineering.

He was named a Distinguished Scholar-Teacher of University of Maryland, College Park, in 2007, where he is the Christine Kim Eminent Professor of Information Technology. He leads the Maryland Signals and Information Group, conducting research encompassing broad areas of signal processing and communications with recent focus on cooperative communications, cognitive networking, social learning and networks, and information forensics and security.

Dr. Liu is the recipient of numerous honors and awards, including the IEEE Signal Processing Society Technical Achievement Award and Distinguished Lecturer. He also received various teaching and research recognitions from the University of Maryland, including the university-level Invention of the Year Award; and the Poole and Kent Senior Faculty Teaching Award and Outstanding Faculty Research Award, both from the A. James Clark School of Engineering. An ISI Highly Cited Author, he is a Fellow of AAAS. He is President of the IEEE Signal Processing Society, where he has served as Vice-President Publications and Board of Governor. He was the Editor-in-Chief of the *IEEE Signal Processing Magazine* and the founding Editor-in-Chief of the *EURASIP Journal on Advances in Signal Processing*.