Indirect-Reciprocity Data Fusion Game and Application to Cooperative Spectrum Sensing

Biling Zhang, *Member, IEEE*, Yan Chen, *Senior Member, IEEE*, Jung-Lang Yu, *Member, IEEE*, Bin Chen, and Zhu Han, *Fellow, IEEE*

***Abstract*— Data sharing is one critical step to implement data fusion, and how to encourage sensors to share their data is an important issue. In this paper, we propose a reputation-based incentive framework, where the data sharing stimulation problem is modeled as an indirect reciprocity game. In the proposed game, sensors choose how to report their results to the fusion center and gain reputations, based on which they can obtain certain benefits in the future. Taking the sensing and fusion accuracy into account, reputation distribution is introduced in the proposed game, where we prove theoretically the Nash equilibrium of the game and its uniqueness. Furthermore, we apply the proposed scheme to the cooperative spectrum sensing. We show that within an appropriate cost-to-gain ration, the optimal strategy for the secondary users is to report when the average received energy is above a given threshold and keep silence otherwise. Such an optimal strategy is also proved to be a desirable evolutionarily stable strategy. Finally, simulation results are shown to verify the theoretical results and demonstrate that compared with the existing schemes, our proposed scheme achieves better operating characteristic curve and higher system throughput with convincing performance on fairness.**

***Index Terms*— Indirect reciprocity, data fusion, cooperation stimulation, game theory, evolutionarily stable strategy.**

Manuscript received October 17, 2016; revised February 15, 2017 and May 7, 2017; accepted July 4, 2017. Date of publication July 14, 2017; date of current version October 9, 2017. This work was supported in part by the National Natural Science Foundation of China under Grant 61501041 and Grant 61672137, in part by the Open Foundation of State Key Laboratory under Grant ISN16-08, in part by the 111 Project B17008, in part by the Thousand Youth Talents Program of China (to Yan Chen), in part by the U.S. National Natural Science Foundation under Grant CNS-1702850, Grant CNS-1646607, Grant ECCS-1547201, Grant CCF-1456921, Grant CMMI-1434789, Grant CNS-1443917, and Grant ECCS-1405121, and in part by Ministry of Science and Technology Taiwan under Grant MOST-104-2221-E-030-004-MY2. The associate editor coordinating the review of this paper and approving it for publication was S. K. Jayaweera. *(Corresponding author: Biling Zhang.)*

B. Zhang and B. Chen are with the School of Network Education, Beijing University of Posts and Telecommunications, Beijing 100876, China, and also with the State Key Laboratory of Integrated Services Networks, Xidian University, Xi'an 710071, China (e-mail: bilingzhang@bupt.edu.cn; 854671203@qq.com).

Y. Chen is with the School of Electronic Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China (e-mail: eecyan@uestc.edu.cn).

J.-L. Yu is with the Department of Electrical Engineering, Fu Jen Catholic University, New Taipei City 24205, Taiwan (e-mail: yujl@ee.fju.edu.tw).

Z. Han is with the Department of Electrical and Computer Engineering, University of Houston, Houston, TX 77004, USA, and also with the Department of Computer Science and Engineering, Kyung Hee University, Seoul 100744, South Korea (e-mail: zhan2@uh.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/TWC.2017.2725836

I. Introduction

In THE big data era, information about a phenomenon or a system of interest can be obtained through sensors (also data suppliers) of different types with different measurement techniques in different domains. To jointly analyze information from multiple sensors and extract knowledge for various purposes, concepts of data fusion were proposed and corresponding techniques were introduced [1]. With the analytic outcomes of a data fusion process, users are able to obtain a more unified picture and global view of the system, answer specific questions about the system, and make more accurate decisions.

Data fusion methodologies are initially designed for military applications. With the development of technologies, in recent years data fusion are gradually evolved in nonmilitary fields such as cognitive radio networks (CRNs) [2], smart grids [3] and driver assistance systems [4]. To effectively exploit the diversity that multiple sensors offer, a lot of algorithms, based on hard fusion or soft fusion, have been proposed. In the hard fusion based algorithms, each sensor sends a one-bit hard local decision to the fusion center (FC). While in the soft fusion based algorithms, each sensor sends to the FC a quantized version of a local decision statistic such as the log-likelihood ratio or any suitable sufficient statistic.

Although the existing fusion algorithms have been proved to be helpful in achieving results which greatly improve system reliability, most of them are based on the assumption that all the sensors are altruistic and will share their own data for fusion unconditionally. However, such an assumption may not be true in reality. Considering that collecting and reporting data consume energy and bring extra cost, sensors may not cooperate if cooperation cannot bring them benefit.[1] Without a sufficient amount of data, the fusion algorithms cannot work properly. In such a case, how to stimulate different data providers to collaborate and share their data is an important problem.

Two most common categories of cooperation stimulation schemes are payment-based and reputation-based schemes. Payment-based schemes, such as virtual currency and auction, have been proposed to enforce files sharing in P2P networks [5], dynamic spectrum sharing in CRNs [6] and

[1] Since data sharing is regarded as a very important kind of cooperation in a data fusion system, data sharing and cooperation is used interchangeably in our paper.

participatory sensing [7] in wireless communication systems. Although these schemes can achieve promising results, the applications of these schemes are still limited by their requirement of tamper-proof hardware or central banking server(s) for the sake of easy management and security control. Reputation-based cooperation stimulation schemes are extensively discussed in ad hoc networks [8], P2P networks [9] and wireless data networks [10] to ensure reliable communications, or in wireless sensor networks [11] and CRNs [12] for secure data fusion. Nevertheless, until recently few reputation-based schemes is proposed for how to stimulate the sensors to share their sensing results, and there is no theoretical justification about the optimality of such approaches in data fusion.

Game theory has also been used to analyze the problem of cooperation. Specifically, based on cooperative game theory such as bargaining game [13], [14] and coalition game [15]–[17], a number of cooperation schemes have been introduced. Cooperative game theoretical frameworks mainly investigate how to obtain and allocate net income within an alliance. On the other hand, in cooperative games, though players may be willing to accept a bargaining solution that is good enough for both sides, they are eventually interested in maximizing their own outcome. To guarantee cooperation, an enforceable contract or a stimulation mechanism is required, which, however, falls outside the scope of cooperative game theory. Moreover, most of the game theoretic frameworks are based on the direct reciprocity model, where the behaviors of all players will not be evaluated by other players except their opponents. According to the well-known Prisoner's Dilemma and backward induction principle, for two players who are directly played, they play cooperatively only when the game is repeated for infinite times. Nevertheless, due to mobility or changes of environment, players will periodically change their partners to achieve better performances. Then, the only optimal strategy for them is to play non-cooperatively. In such a circumstance, a stimulation mechanism is also very necessary.

Indirect reciprocity, which has been widely applied in the area of social science and evolutionary biology [18], [19], has recently attracted much attention in applications such as packages forwarding [20], cooperative transmission [21], [22], and energy exchange [23]. The basic idea behind indirect reciprocity is "I help you not because you have helped me but because you have helped others". That is, the reason the recipient can get help from the current donor is that the recipient has helped others as a donor before. This means not only the evaluations from the opponents but also the evaluations from other observers will be taken into account. In such a case, players are incentive to play cooperatively even the game will not repeat for infinite times.

In this paper, we propose a new framework to achieve successful data fusion by incorporating indirect reciprocity based incentive mechanism to stimulate the cooperation among sensors, where a sensor is not required to always play with the same group of sensors. Then, we discuss the application of the proposed framework in the CSS system, which relies on the secondary users' (SUs') reporting their sensing signals or local decisions. Considering that SUs are generally selfish, mobile and belong to different organizations, the proposed framework is an ideal tool for such an application scenario.

Our major contributions are summarized as follows.

- We propose an indirect reciprocity game framework to study the data sharing stimulation problem, which is different from previous works that purely focused on direct reciprocity models. In the proposed game, sensors perform sensing and decide how to report the sensing results to the FC, where the sensing results are fused and the reputations of sensors are assigned based on their actions. Later, the sensors will be given the rewards according to their reputations. Indirect reciprocity game here is used not only to stimulate the cooperation of sensors, but also to overcome the cheating behaviors when sensors untruthfully report their sensing results.
- In our model, we consider the centralized scenario where a sensor's reputation value is assigned by and stored in the FC. Although there is no reputation propagation error as considered in the previous works, a sensor may be assigned a reputation value different from its expectation due to the inaccuracy in its sensing and the FC's fusion. Taking the inaccuracy into account, we develop a reputation updating policy to update sensors' reputations, where the concept of reputation distribution is introduced to capture the impact of such inaccuracy. We prove theoretically that under certain conditions, the reputation distributions of both individual sensor and the whole population are unique and stationary. Meanwhile, the Nash equilibrium (NE) of the proposed game and its uniqueness are proved as well.
- We apply the proposed indirect reciprocity game to CSS in CRNs, where the optimal strategy of the sensors (i.e., SUs) is theoretically analyzed. We show that within an appropriate cost-to-gain ration, the optimal strategy for the SUs is to report when the average received energy is above the given threshold and keep silence otherwise, which not only greatly improve the FC's fusion accuracy but also save the SUs' energy. Such an optimal strategy is also proved to be a desirable evolutionarily stable strategy (ESS), which leads to a "good" society where most of the SUs have high reputation. Finally, we extend our discussion on this application from the hard fusion case to the soft fusion case, as well as from the single-channel sensing case to the multi-channel sensing case, which makes our proposed framework more general and practical.

The rest of this paper is organized as follows. Section II describes in details our system model, problem formulation for stimulating data sharing, the reputation updating policy, and the equilibrium of the indirect reciprocity game. In Section III, we apply the proposed indirect reciprocity game to CSS in CRNs, and extend our discussion from the hard fusion case to the soft fusion case as well as from the single-channel sensing case to the multi-channel sensing case. Finally, we show the simulation results in Section IV and draw conclusions in Section V.

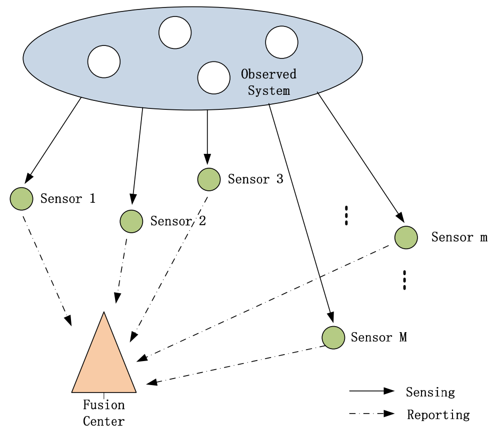


Fig. 1. The system model.

II. Indirect Reciprocity Data Fusion Game

In this section, we will introduce in details the proposed indirect reciprocity data fusion game. We first describe the system model in Section II-A. Then the player's (i.e. sensor's) actions and action rules and the social norm adopted to assign a player's reputation value are defined in Section II-B and Section II-C, respectively. In Section II-D and Section II-E, how the player's reputation value is affected by the inaccuracy existing in its own sensing and the FC's fusion, and how the reputation value is updated are discussed. With the proposed reputation updating policy, we continue to study the player's expected reputation distribution, based on which, the payoff function is defined and the game is proposed in Section II-F. Finally, the equilibrium of the proposed game is studied in Section II-G.

A. System Model

We consider a data fusion system as shown in Fig. 1. In the data fusion system there are M sensors who perform sensing and obtain local observations. Here, we assume that the local observations are used for application of classic hypothesis testing which compares two exclusive hypotheses, i.e., the null hypothesis H_0, and the alternative hypothesis H_1. However, our proposed framework can also be applied to other applications such as entity identification and attribute estimation. Suppose that the hypotheses can be assessed through the sensor's observation on its received signal energy, i.e., energy detection.[2] Let $S_m = \frac{1}{N}\sum_{t=1}^{N} |r_m(t)|^2$ be the average energy of signal received by sensor m, where N is the number of signal samples, and $r_m(t)$ is the sample of observed signal at the sensor m's receiver which can be defined as

$$r_m(t) = \begin{cases} h_m s(t) + n_m(t), & \text{if } H_1, \\ n_m(t), & \text{if } H_0. \end{cases} \quad (1)$$

[2]Our system model can be easily extended to the systems where other detection techniques such as cyclostationary feature detection and matched filter detection are used, since the hard local decisions or local decision statistics of different detection techniques are similar in the form and the way FC fuses, though they are diverse in their meanings.

In (1), $s(t)$ is the signal appearing at time slot t, h_m is the channel gain, and $n_m(t)$ is the additive white Gaussian noise (AWGN). Then if the energy of the observed signal is lower than the predefined threshold λ, i.e., $S_m < \lambda$, the sensor m will accept H_0 as true. Otherwise, it accepts H_1 as true.

In the data fusion system, a FC is established to fuse data collected from multiple sensors. If all the sensors truthfully report their local data, a more accurate outcome for telling null hypothesis or the alternative hypothesis can be achieved by the FC [4].

B. Action and Action Rule

In the data fusion system, hard fusion is first considered for simplicity. According to hard fusion, sensors are required to report to the FC their local decisions on the hypotheses rather than their original observations. Under such circumstance, we define a sensor's action after sensing as $a \in \mathcal{A}$, where $\mathcal{A} = \{1, 2\}$ is the action set that a sensor can choose its action from. In the action set $\mathcal{A}$, "2" denotes that a sensor reports its local decision of judging hypothesis H_1 to be true. If a sensor's local decision is judging hypothesis H_0 to be true, it chooses action "1" by keeping silence for saving energy.

Note that a sensor can take either action basing on the same average received energy level. Given any average received energy level, what action a sensor will take forms the sensor's strategy. To describe the sensor's strategies, we further introduce the concept of action rule.

Definition 1 (Action Rule): An action rule is an action vector, where the j^{th} element denotes the action a sensor will take under average received energy level j.

Taking into account the sensor's way of observation (i.e., energy detection) and FC's rule of data fusion (i.e., hard fusion), the average energy only needs to be quantized to two levels, i.e., $\mathcal{L} = \{1, 2\}$, where $\mathcal{L}$ is the set of average energy level. Here level 1 means that the average energy the sensor observes is below the given threshold λ, while level 2 means that the average energy the sensor observes is equal to or above the given threshold λ. According to definition 1, there are in total four action rules for a sensor to follow, that is, $\boldsymbol{a}^{(1)} = (2, 2), \boldsymbol{a}^{(2)} = (2, 1), \boldsymbol{a}^{(3)} = (1, 2)$ and $\boldsymbol{a}^{(4)} = (1, 1)$, with the action rule set $\mathcal{A}_l = \{(2, 2), (2, 1), (1, 2), (1, 1)\}$. Here, $\boldsymbol{a}^{(l)}, l = 1, 2, 3, 4$, is used to denote the action rule in $\mathcal{A}_l$ for simplicity. The action rule $\boldsymbol{a}^{(1)}$ ($\boldsymbol{a}^{(4)}$) means that a sensor will (not) report that it considers hypothesis H_1 being true no matter what its average received energy level is at, while the action rule $\boldsymbol{a}^{(2)}$ ($\boldsymbol{a}^{(3)}$) means that a sensor will report that it considers hypothesis H_1 being true only when its average received energy is at level 2 (level 1).

Moreover, when a sensor need to consider how many reputations it has gained before it makes its decision, $a_{i,j} \in \mathcal{A}$ is used to describe the action performed by a sensor with reputation i, $i \in \mathcal{R}$, under average received energy level j, $j \in \mathcal{L}$. Here, $\mathcal{R}$ is the set of reputation.

C. Social Norm: How to Assign Reputation

A social norm $\boldsymbol{\Omega}$ is a matrix used for assigning the immediate reputation value of sensors. In our model, we assume that all sensors in the system share the same norm.

Recall that the FC makes the final decision based on the sensors' local data. Intuitively, the more accurate the local data is the higher probability that the FC can make a correct decision. In such a case, we encourage a sensor with a high average received energy level to send out its local decision indicating H_1 being true. Otherwise, the sensor should keep silence when its average received energy level is low, which is equivalent to reporting H_0 being true. In such a way, the interference to the FC can be successfully avoided and at the same time the sensor's energy can be saved.

However, the FC, who uses the social norm to assign a sensor reputation, has no idea about the actual average received energy level of the sensor when the sensor takes its action. Therefore, in our model, instead of based on the sensor's average received energy level, the social norm is defined according to the FC's final decision as well as the sensor's action. Let the FC's final decision be "1" if the FC's fusion result indicates H_0 being true, and the final decision be "2" otherwise. Then the social norm $\boldsymbol{\Omega}$ is defined as

$$\boldsymbol{\Omega} = [\Omega_{i,j}] = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \tag{2}$$

where $\Omega_{i,j}$ is the reputation value assigned to a sensor who takes action i while the FC's final decision is j, and the set of a sensor's reputation value is $\mathcal{R} = \{1, 2\}$.

From the definition of $\boldsymbol{\Omega}$ we can see, when a sensor's action is consistent with the FC's final decision, it will be given a high reputation value. Otherwise, it will be given a low reputation value. Notice that there is a gap between the FC's final decision and the average received energy level of a sensor due to the inaccuracy of both the FC's fusion and the sensor's sensing. Such a gap definitely affects the sensors' reputation values and their actions. Therefore, in the following subsection, we will discuss the impact in details. Considering that sensors are generally supposed to transmit their data through a dedicated control channel whose error can be minimized and thus is ignorable by Automatic Repeat Request (ARQ) or channel coding such as fountain code, how a sensor's action is distorted by the channel and how its reputation distribution is affected are temporarily not considered in this paper. Readers who are interested in this topic can refer to our previous work [21].

D. Decision Consistence Matrix

As we have discussed in previous subsection, the FC assigns a reputation value to a sensor according to its final decision and the sensor's action. This is because the FC believes that by the fusion result it accurately obtains the true hypothesis, with which it can correctly evaluate the sensor's action. However, a sensor may make a wrong estimation on the hypothesis due to the possibility of miss detection and false alarm in its sensing. Here, the probability of detection $P_{d,m}$ is used to represent sensor m's probability of successfully assessing hypothesis H_1, and the probability of false alarm $P_{f,m}$ is used to represent sensor m's probability of unsuccessfully assessing hypothesis H_0, i.e.,

$$P_{d,m}(\lambda) = Pr(S_m > \lambda | H_1), \tag{3}$$

and

$$P_{f,m}(\lambda) = Pr(S_m > \lambda | H_0). \tag{4}$$

Since all the sensors use the same threshold λ, we have $P_{d,m} = P_d$ and $P_{f,m} = P_f$.

Consequently, the gap between the FC's fusion result and the average energy level of a sensor can be described by the decision consistence matrix $\boldsymbol{D}$ as

$$\boldsymbol{D} = [D_{i,j}] = \begin{bmatrix} D_{1,1} & D_{1,2} \\ D_{2,1} & D_{2,2} \end{bmatrix}, \tag{5}$$

where $D_{i,j}$ is the probability that the energy of the sensor's received signal is at level i while the FC's fusion result indicates that it should be at level j.

Assume that the FC exploits the majority rule to make final decisions. Let $P_s(i|k)$ be the probability that the energy of the sensor's received signal is at level i under hypothesis H_k, $P_{fc}(j|i,k)$ be the probability that the FC considers the energy of the sensor's received signal should be at level j given that the energy of the sensor's received signal is at level i and the hypothesis is H_k, and P_0 and P_1 be the probability that the hypotheses of H_0 and H_1 are true, respectively. Since the FC makes decisions based on the sensors' reports, the sensor's received signal energy level and the level determined by the FC are not independent. Then $D_{1,1}$ can be calculated as

$$\begin{aligned} D_{1,1} &= P_{fc}(1|1,0)P_s(1|0)P_0 + P_{fc}(1|1,1)P_s(1|1)P_1, \\ &= \sum_{k=[\frac{M}{2}]-1}^{M-1} \binom{M-1}{k} [(1-P_f)^{k+1} P_f^{M-k-1} P_0 \\ &\qquad + (1-P_d)^{k+1} P_d^{M-k-1} P_1]. \end{aligned} \tag{6}$$

Note that $P_{fc}(1|1,0)$ ($P_{fc}(1|1,1)$) is calculated as in (6) because there should be more than $[\frac{M}{2}] - 1$ other sensors whose received signal energy is at level 1. Similarly, we have $D_{1,2} = \sum_{k=[\frac{M}{2}]+1}^{M-1} \binom{M-1}{k} [P_f^k (1-P_f)^{M-k} P_0 + P_d^k (1-P_d)^{M-k} P_1]$, $D_{2,1} = \sum_{k=[\frac{M}{2}]+1}^{M-1} \binom{M-1}{k} [(1-P_f)^k P_f^{M-k} P_0 + (1-P_d)^k P_d^{M-k} P_1]$ and $D_{2,2} = \sum_{k=[\frac{M}{2}]-1}^{M-1} \binom{M-1}{k} [P_f^{k+1} (1-P_f)^{M-k-1} P_0 + P_d^{k+1} (1-P_d)^{M-k-1} P_1]$. Without loss of generality, we have $D_{1,1} > D_{1,2}$ and $D_{2,2} > D_{2,1}$.

E. Reputation Updating Policy

In order to establish the reputation of a sensor with the social norm, we develop a reputation updating policy as shown in Fig. 2. Our proposed reputation updating model considers the imperfect sensing of the sensors and the FC's fusion error, both of which influence the sensors' reputation updating. Therefore, $\boldsymbol{D}$, which has been discussed in the previous subsection, is incorporated into the reputation updating policy to describe such an impact.

Due to the influence of sensors' imperfect sensing and the FC's fusion error, a sensor, after taking an action, may be assigned a high or low reputation with a certain probability.

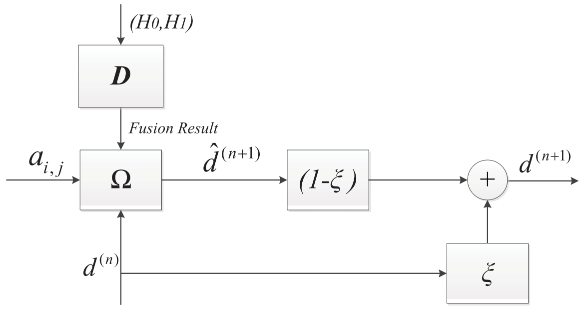


Fig. 2. The reputation updating model.

To capture such a possibility, we assign for the sensor a reputation distribution $\boldsymbol{d} = (d_l, d_h)$ with $d_l + d_h = 1$, where d_l indicates the probability of the sensor being assigned a low reputation, while d_h indicates the probability that the sensor being assigned with a high reputation value.

In a simple case, a sensor's reputation at time index n is i and its received energy is at level j. After the sensor determines and performs action $a_{i,j}$, at time index $n+1$ its immediate reputation distribution will be

$$\hat{\boldsymbol{d}}^{(n+1)}(a_{i,j}) = \sum_{l=1}^{2} \boldsymbol{e}_{\Omega_{a_{i,j},j}} \cdot D_{j,l}, \tag{7}$$

where $\boldsymbol{e}_i$ is the standard basis vector. Here, $\Omega_{a_{i,j},j}$ is the reputation value assigned to the sensor after it takes action $a_{i,j}$ under the received energy level j, and $e_{\Omega_{a_{i,j},j}}$ is the corresponding reputation distribution, if the influence of the sensor's imperfect sensing and the FC's fusion error is not considered. When such an influence is taken into account, however, the sensor's reputation value $\Omega_{a_{i,j},j}$ may be changed to the low one or the high one with probability of $D_{j,1}$ and $D_{j,2}$, respectively. Hence, we have the sensor's immediate reputation distribution as in (7).

Then, according to our reputation updating policy in Fig 2, after the FC collects all the information from sensors it updates (and then broadcasts) the sensor's reputation distribution at time index $n+1$, i.e., $\boldsymbol{d}^{(n+1)}(a_{i,j})$, using a linear combination of the sensor's original reputation distribution $\boldsymbol{d}^{(n)}$ and the immediate reputation distribution $\hat{\boldsymbol{d}}^{(n+1)}$ with a weight ξ as

$$\boldsymbol{d}^{(n+1)}(a_{i,j}) = (1-\xi)\hat{\boldsymbol{d}}^{(n+1)}(a_{i,j}) + \xi \boldsymbol{d}^{(n)}. \tag{8}$$

Here, the weight ξ can be treated as a discounting factor of the past reputation.

With the reputation distribution $\boldsymbol{d}$, a sensor is able to pursue a certain interest/benefit according to the mechanism designed by a specific application. Based on the reputation distribution $\boldsymbol{d}$, we can also evaluate a sensor's overall performance. For example, if a sensor has a high value of d_l, the sensor must have always made decisions that are inconsistent with the FC's fusion results, indicating the sensor is in a very bad sensing environment or it always reports its decisions incorrectly. On the other hand, if a sensor has a high value of d_h, the sensor must have always made decisions that are consistent with the FC's fusion results, indicating the sensor's decisions are accurate and it always truthfully reports its sensing decisions. In such a case, the FC is able to tell whether a sensor is a malicious sensor and overcome the cheating behavior by discarding the information a malicious sensor reports.

At time index n, suppose sensor m takes action $a_m \in \mathcal{A}$ and obtains reputation distribution $\boldsymbol{d}^{(n)}(a_m)$, the reputation distribution of the whole data fusion system can be defined as

$$\boldsymbol{\eta}^{(n)}(a_1, \ldots, a_M) = \frac{1}{M} \sum_{m=1}^{M} \boldsymbol{d}^{(n)}(a_m). \tag{9}$$

From (9) we can see that $\boldsymbol{\eta}^{(n)}(a_1, \ldots, a_M)$ is a function of actions of all the sensors, and will be denoted as $\boldsymbol{\eta}^{(n)}(\cdot)$ for simplicity.

F. Payoff Function

In this subsection, we are going to define the sensor's payoff function. Before that, we first give the definition of expected reputation distribution, which is used to describe the reputation distribution for a given action rule.

Definition 2 (Expected Reputation Distribution): For any given action rule $\boldsymbol{a}$, a sensor's expected reputation distribution is $\boldsymbol{d}(\boldsymbol{a}) = P_{L_1}\boldsymbol{d}(\boldsymbol{a}(1)) + P_{L_2}\boldsymbol{d}(\boldsymbol{a}(2))$, where P_{L_1} (P_{L_2}) is the probability that the sensor's average received energy is at level 1 (level 2).

According to the reputation updating policy, we have the updating policy of the expected reputation distribution as shown in Lemma 1.

Lemma 1: $\forall \boldsymbol{a} \in \mathcal{A}_l$ and $\boldsymbol{d}^{(n)}$, the sensor's expected reputation distribution at time index $n+1$ is updated as $\boldsymbol{d}^{(n+1)}(\boldsymbol{a}) = (1-\xi)\hat{\boldsymbol{d}}^{(n+1)}(\boldsymbol{a}) + \xi\boldsymbol{d}^{(n)}$, where $\boldsymbol{d}^{(n)}$ is the sensor's reputation distribution at time index n, and $\hat{\boldsymbol{d}}^{(n+1)}(\boldsymbol{a})$, which follows Definition 2, is the sensor's immediate expected reputation distribution with action rule $\boldsymbol{a}$ at time index $n+1$.

Proof: See proof in Appendix A. ■

Under certain conditions, $\boldsymbol{d}^{(n+1)}(\boldsymbol{a})$ has a nice property, which is characterized in Lemma 2 for later analysis.

Lemma 2: $\forall \boldsymbol{a} \in \mathcal{A}_l$, $\boldsymbol{d}^{(n+1)}(\boldsymbol{a}) = \hat{\boldsymbol{d}}(\boldsymbol{a})$ is the steady expected reputation distribution if $\boldsymbol{a}$ is always followed. Here, $\hat{\boldsymbol{d}}(\boldsymbol{a})$ is the immediate expected reputation distribution after the sensor follows action rule $\boldsymbol{a}$.

Proof: See proof in Appendix B. ■

From Lemma 2, we can see that a sensor's expected reputation distribution with a certain action rule is approaching to the corresponding immediate expected reputation distribution and remains steady, if it keeps following the action rule. Further, the stability of individual sensor's expected reputation distributions will lead to a steady reputation distribution of the whole system, which will be proved in Theorem 4.

Next, we are going to define the sensor's payoff function for following a certain action rule at any given time index. For simplicity, here we omit the superscript of time index.

As we all know, reporting local information consumes energy. When a sensor takes a certain action, a cost may incur. Hence, we usually define a cost function $f_c(a)$ which is used to measure a sensor's cost after it takes action $a \in \mathcal{A}$. On the other hand, through action a, the sensor gains a corresponding reputation distribution $\boldsymbol{d}(a)$, with which the sensor can achieve certain benefit $f_g(\boldsymbol{d}, \boldsymbol{\eta})$. Considering that in a fair society the resource is always allocated in a fair manner, we assume that $f_g(\boldsymbol{d}, \boldsymbol{\eta})$ is an increasing function of $\boldsymbol{d}$ while a decreasing function of the reputation of the whole population $\boldsymbol{\eta}$. Therefore, the utility function of a sensor, say sensor m who takes action a_m while other sensors take actions a_{-m}, should be

$$W(a_m, a_{-m}) = f_g(\boldsymbol{d}(a_m), \boldsymbol{\eta}(a_m, a_{-m})) - f_c(a_m). \quad (10)$$

Similarly, we define the sensor m's expected utility with action rule $\boldsymbol{a}_m \in \mathcal{A}_l$ while other sensors take actions rules $\boldsymbol{a}_{-m}$ by revising (10) as

$$W(\boldsymbol{a}_m, \boldsymbol{a}_{-m}) = f_g(\boldsymbol{d}(\boldsymbol{a}_m), \boldsymbol{\eta}(\boldsymbol{a}_m, \boldsymbol{a}_{-m})) - f_c(\boldsymbol{a}_m). \quad (11)$$

From (10) and (11) we can see that a sensor's (expected) utility function is determined not only by its action (or action rule), but also by other sensors' actions (or action rules) which have been incorporated into the entire population distribution $\boldsymbol{\eta}(\cdot)$. In such a case, the interactions among sensors form a game called indirect reciprocity data fusion game.

G. Equilibrium of Indirect Reciprocity Data Fusion Game

To discuss how a sensor chooses the optimal action in the game, we first give the formal definition of NE, and then derive the NE for the proposed game.

Definition 3 (NE): NE is the action rule profile $\boldsymbol{a}^* = \{\boldsymbol{a}_1^*, \ldots, \boldsymbol{a}_M^*\}$ where $\forall m \in \mathcal{M} = \{1, \ldots, M\}$, $\boldsymbol{a}_m^* = \arg\max_{\boldsymbol{a}_m \in \mathcal{A}_l} W(\boldsymbol{a}_m, \boldsymbol{a}_{-m}^*)$. Here, $\boldsymbol{a}_m$ is the action rule adopted by sensor m, $\boldsymbol{a}_{-m}^*$ is the optimal action rules of all sensors except sensor m, and $W(\boldsymbol{a}_m, \boldsymbol{a}_{-m}^*)$ is sensor m's expected utility defined in (11).

With the definition, we show in Theorem 1 that there exists a NE of our proposed game.

Theorem 1: Given the set of sensors $\mathcal{M}$ and the set of action rule $\mathcal{A}_l$, there exists a NE of the proposed game, i.e., $\boldsymbol{a}^* = \{\boldsymbol{a}_1^*, \ldots, \boldsymbol{a}_M^*\}$, which satisfies

$$W(\boldsymbol{a}_m, \boldsymbol{a}_{-m}^*) \geq W(\boldsymbol{a}_m', \boldsymbol{a}_{-m}^*), \quad \forall m \in \mathcal{M}, \quad (12)$$

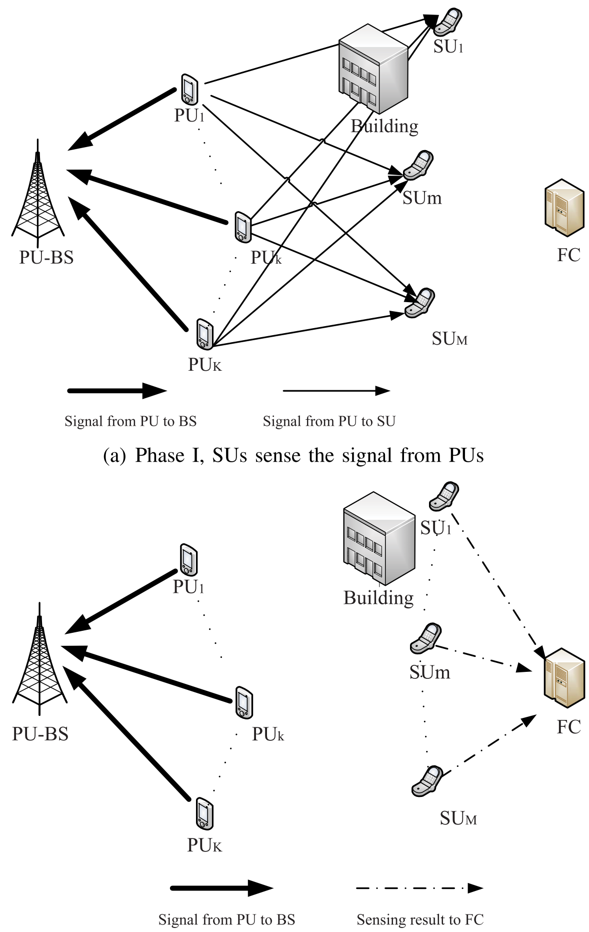
where $\boldsymbol{a}_m' \in \mathcal{A}_l \setminus \boldsymbol{a}_m^*$ and $W(\cdot)$ is defined in (11).

Proof: See proof in Appendix C. ∎

Note that the NE may not be unique. For example, if the equality of (12) holds for certain sensors, then each of such sensors has at least two action rules that can bring it maximal expected utility. Choosing any one of such rules will lead to a different NE. However, when (12) strictly holds for all sensors, the NE is unique.

III. Application to Cooperative Spectrum Sensing

In this section, we study the application of the proposed indirect reciprocity data fusion game in CSS, which is proposed to counter the hidden terminal problem and improve the sensing accuracy in CRN. In the literature, many aspects of CSS approaches have got thoroughly studied, among which, sensing schemes for an individual SU [24], fusion algorithms for the FC [25], and cooperation strategies [13]–[17], [26] have been comprehensively investigated. Although the existing CSS approaches have assumed that SUs will share their spectrum sensing results with each other, SUs in practice are generally selfish and belong to different organizations, due to which they will not take part in CSS unconditionally.



(a) Phase I, SUs sense the signal from PUs

(b) Phase II, SUs send their results to the FC

Fig. 3. The cooperative sensing system.

The sensing result sharing problem in CSS thereby can be ideally modeled as an indirect reciprocity data fusion game, where the sensors are the SUs, the FC is a dedicated sever or cognitive radio station [27], and the hypotheses are whether the primary users (PUs) are active in the channels or not. Within the proposed game framework, how the SUs make their decisions on results sharing and how they are stimulated to cooperate are investigated.

A. System Model

As shown in Fig. 3, we consider a system consisting one primary network and one secondary network. Assume that there are K PUs in the primary network, each of which owns a licensed channel. Meanwhile, M SUs in the secondary network are sensing these K channels to find out spectrum holes for access opportunity. To implement CSS, the whole procedure is divided into two phases. During the first phase, the SUs sense the signal from PUs. During the second phase, the SUs send their local sensing information to the FC. After the

FC collects all the local sensing results or individual decisions from SUs, it makes final decisions for them. To encourage SUs to participate in CSS, the FC assigns a corresponding reputation value to the SU after it takes a certain action, i.e., reports its sensing result. Based on the reputation, the SU can apply for the usage of vacant channels in the future.

To detect the PUs, energy detection, the most common method for spectrum sensing without knowing the prior information of PUs, is adopted by SUs. Let H_1 and H_0 be hypotheses of PU being present or absent, respectively. Given the energy threshold λ, the SU's probability of detection $P_{d,m}$ and probability of false alarm $P_{f,m}$ can be calculated using (3) and (4), respectively. In practice, the threshold λ is firstly determined from the desired false alarm probability, and the probability of detection is then calculated as $P_{d,m} = 1-\frac{e^{-\sqrt{z}C_m}E(e^{B_m})}{e^{-\sqrt{\lambda}}}$, where z is the groups of observation samples, C_m is a constant, and B_m is a limiting distribution, as defined in [28, eq. (18)].

In the following, we first apply the proposed game to the CSS system where there is only one primary channel, i.e., $K = 1$, and the FC makes the final decision with hard fusion in Section III-B. For this scenario, we derive theoretically the NE of the game and the stable condition of the NE. Then in Section III-D, we extend our discussion on the single-channel scenario from hard fusion case to soft fusion case. Finally, how to apply our scheme to the systems where there are multiple primary channels, i.e., $K > 1$, is discussed in Section III-E.

B. Proposed Game for the Single Channel ($K = 1$) and Hard Fusion case

For the CSS system where there is only one primary channel and the FC makes the final decision with hard fusion, the proposed game can be directly applied. That is, given the threshold λ, a SU's average energy is quantized to 2 levels, i.e., $\mathcal{L} = \{1, 2\}$. Under each level, a SU has two choices: reporting or keeping silence. The SU's selection, i.e., its action, thereby can be defined as $a \in \mathcal{A} = \{1, 2\}$. $a_{i,j} \in \mathcal{A}$ is then used to describe the action performed by a SU with reputation i under average energy level j. The social norm used by the FC to assign SUs immediate reputation value has been defined as $\boldsymbol{\Omega}$ in (2). To characterize the gap between the fusion result and SU's actual average energy level, we have the decision consistence matrix $\boldsymbol{D}$ for this CSS scenario by following the definition in (5), where P_0 and P_1 are the prior probabilities of PU's being absent and present, respectively. Then the reputation updating policy in Fig. 2 is adopted. In the game, each SU is assigned a reputation distribution $\boldsymbol{d} = (d_l, d_h)$. According to the policy, a SU's reputation distribution with action a at time index $n+1$, i.e., $\boldsymbol{d}^{(n+1)}(a)$, is updated by (8), while its expected reputation distribution of adopting a certain action rule follows $\boldsymbol{d}^{(n+1)}(\boldsymbol{a}) = (1-\xi)\hat{\boldsymbol{d}}^{(n+1)}(\boldsymbol{a}) + \xi\boldsymbol{d}^{(n)}$, as has been proved in Lemma 1.

With reputation distribution $\boldsymbol{d}$, a SU can apply for a corresponding amount of vacant channels whenever it has data to forward. Suppose the FC allocates the vacant channels to SUs in a proportionally fair manner. Then the access time of the vacant channel allocated to the SU with reputation distribution $\boldsymbol{d}$ will be

$$t(\boldsymbol{d}, \boldsymbol{\eta}) = (1-P_F)\frac{N_s}{M}\frac{E(\boldsymbol{d})-1}{E(\boldsymbol{\eta})-1}, \tag{13}$$

where $P_F = \sum_{k=[\frac{1+M}{2}]}^{M} \binom{M}{k} P_f^k(1-P_f)^{M-k}$ is FC's probability of false alarm, $\boldsymbol{\eta}$ is the current reputation distribution of the CSS system defined in (9), N_s is the number of SUs a vacant channel allows to access, $(1-P_F)\frac{N_s}{M}$ is the average available channels for each SU, and $E(\cdot)$ is the expected operation on the reputation distribution. Let g be the gain of unit accessing time of vacant channels, the benefit the SU achieves thereby is

$$f_g(\boldsymbol{d}, \boldsymbol{\eta}) = g t(\boldsymbol{d}, \boldsymbol{\eta}). \tag{14}$$

On the other hand, all the SUs sense the channel and part of them will report their sensing results. Here, we omit the sensing cost since it is the same to all SUs. Instead, we consider the cost incurred by power consumption when the SU sends its sensing result to the FC. Let c be the unit price of power, the cost function of a SU's taking action a is

$$f_c(a) = c(a-1), \tag{15}$$

Then as the action rule is considered, the expected utility of SU m who takes action rule $\boldsymbol{a}_m \in \mathcal{A}_l$ while other SUs take action rules $\boldsymbol{a}_{-m}$ is defined from a statistical point of view as

$$\begin{aligned} W(\boldsymbol{a}_m, \boldsymbol{a}_{-m}) &= f_g(\boldsymbol{d}(\boldsymbol{a}_m), \boldsymbol{\eta}(\boldsymbol{a}_m, \boldsymbol{a}_{-m})) - f_c(\boldsymbol{a}_m) \\ &= g(1-P_F)\frac{N_s}{M}\frac{E(\boldsymbol{d}(\boldsymbol{a}_m))-1}{E(\boldsymbol{\eta}(\boldsymbol{a}_m, \boldsymbol{a}_{-m}))-1} \\ &\quad - c(E(\boldsymbol{a}_m)-1), \end{aligned} \tag{16}$$

where $E(\boldsymbol{a}_m) = P_{L_1}\boldsymbol{a}_m(1) + P_{L_2}\boldsymbol{a}_m(2)$ is the SU's expected action.

Based on $W(\cdot)$, we derive theoretically the SU's optimal action rule. The analytical result is stated in Theorem 2, where the uniqueness of the optimal action rule can be easily seen.

Theorem 2: The optimal action rule for SUs, i.e., $\boldsymbol{a}^*$, is given as follows

$$\boldsymbol{a}^* = \begin{cases} \boldsymbol{a}^{(2)} = (1,2), & \text{if } 0 < c/g < (1-\xi) \\ & \quad (1-P_F)\frac{N_s}{M}\frac{D_{2,2}-D_{2,1}}{E(\boldsymbol{\eta}^*)-1}, \\ \boldsymbol{a}^{(4)} = (1,1), & \text{otherwise}, \end{cases} \tag{17}$$

where $\boldsymbol{\eta}^*$ is the stationary reputation distribution of the system as defined in (20).

Proof: See proof in Appendix D. ■

Obviously, $\boldsymbol{a}^{(4)}$ is an uncooperative strategy since, with such a strategy, the SUs will not report their local decisions to the FC under any circumstance, and thus the FC's performance will degrade. While $\boldsymbol{a}^{(2)}$ is the desired equilibrium where SUs report their local decisions of hypothesis H_1 being true only when their average received energy is above the given threshold and keep silence otherwise, which greatly improve the FC's fusion accuracy.

During the analysis of finding the optimal action rule, we do not include the perturbation effect. Nevertheless, due to uncertainty of the system and/or the incorrect (noisy) parameters, it is possible for the SUs to take a non-optimal action rule in practice. Taking the perturbation effect into account, we need to evaluate the stability of the optimal action rule, that is, whether a SU will deviate the optimal action rule while others always take it, and under what condition will the SU keeps following the optimal action rule.

To discuss the stable condition for optimal strategy $\boldsymbol{a}^{(2)}$, we adopt the concept of ESS. The ESS can be explained as "a strategy such that, if all members of the population adopt it, then no mutant strategies could invade the population under the influence of natural selection". According to [29], an optimal strategy $\boldsymbol{a}^*$ is an ESS if and only if, $\forall \boldsymbol{a} \neq \boldsymbol{a}^*$, $\boldsymbol{a}^*$ satisfies

- equilibrium condition: $U_m(\boldsymbol{a}, \boldsymbol{a}^*) \leq U_m(\boldsymbol{a}^*, \boldsymbol{a}^*)$, and
- stability condition: if $U_m(\boldsymbol{a}, \boldsymbol{a}^*) = U_m(\boldsymbol{a}^*, \boldsymbol{a}^*)$, $U_m(\boldsymbol{a}, \boldsymbol{a}) < U_m(\boldsymbol{a}^*, \boldsymbol{a})$,

where $U_m(\boldsymbol{a}, \boldsymbol{a}^*)$ is the utility of SU_m's deviation from the optimal action rule $\boldsymbol{a}^*$ and $U_m(\boldsymbol{a}^*, \boldsymbol{a}^*)$ is SU_m's utility when all SUs follow action rule $\boldsymbol{a}^*$.

Then, according to the one shot deviation principle and the definition of ESS, $\boldsymbol{a}^{(2)}$ is an ESS if the following inequality holds,

$$U_m(\boldsymbol{a}^{(2)}, \boldsymbol{a}^{(2)}) > U_m(\boldsymbol{a}, \boldsymbol{a}^{(2)}), \quad \forall \boldsymbol{a}. \tag{18}$$

With (18), we can derive the stable condition for $\boldsymbol{a}^{(2)}$ as described in Theorem 3.

Theorem 3: The optimal action rule $\boldsymbol{a}^{(2)}$ is an ESS if the cost-to-gain ratio c/g satisfies

$$0 < c/g < (1-\xi)(1-P_F)\frac{N_s}{M}\frac{D_{2,2} - D_{2,1}}{E(\boldsymbol{\eta}^*) - 1}. \tag{19}$$

Proof: See proof in Appendix E. ■

Further, when all SUs always follow the optimal rule $\boldsymbol{a}^{(2)}$ without deviation, the CSS system has a unique stationary reputation distribution of the entire population, as stated in Theorem 4.

Theorem 4: In the CSS system where (19) is satisfied, there exists a unique stationary reputation distribution of the entire population which is

$$\begin{aligned}\boldsymbol{\eta}^* &= \underset{n\to\infty}{\boldsymbol{\eta}^{(n)}}(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_M) \\ &= (P_{L1}D_{1,2} + P_{L2}D_{2,1}, P_{L1}D_{1,1} + P_{L2}D_{2,2}), \end{aligned} \tag{20}$$

and the condition for $\boldsymbol{a}^{(2)}$ to be an evolutionary stable and optimal strategy thereby is

$$0 < c/g < (1-\xi)(1-P_F)\frac{N_s}{M}\frac{D_{2,2} - D_{2,1}}{P_{L1}D_{1,1} + P_{L2}D_{2,2}}. \tag{21}$$

Proof: See proof in Appendix F. ■

C. Proposed Game for the Single Channel (K = 1) and Soft Fusion Case

In the previous subsection, the characteristic of the proposed game is analyzed based on the assumption that the FC adopts hard fusion where a SU only needs 1 bit to indicate its detection of the PU's presence. However, the proposed scheme is also applicable for the scenario where the FC adopts soft fusion. According to soft fusion, SUs are required to report their original observed signal strength instead of their local decisions. In the following, analysis for the soft fusion case is conducted under the scenario where there is only one primary channel ($K = 1$).

Assume equal gain combining (EGC), one of the approaches to implement soft fusion, is exploited in this scenario. Then the average energy collected by the FC from M SUs will be

$$S_{fc} = \frac{1}{M}\Big(\sum_{m\in\mathcal{M}_1} S_m + |\mathcal{M}_2|\Big), \tag{22}$$

where $\mathcal{M}_1$ is the set of SUs who report their sensing observations, S_m is the sensing observation that SU_m reports, $\mathcal{M}_2$ is the set of SUs who keep silence. For those SUs in $\mathcal{M}_2$, their received energy is supposed to equal to the average energy of AWGN which is always set to 1 [30]. Therefore, $|\mathcal{M}_2|$ is total energy of SUs in $\mathcal{M}_2$. Then by comparing S_{fc} to the predefined threshold λ, the FC achieves the final decision.

Different from the case where hard fusion is used, in the soft fusion case, we quantize the SU's average received energy into J ($J > 2$) levels. Level 1 represents the SU's average received energy is lower than the given threshold λ, while level 2 to level J represent that the SU's average received energy is equal to or higher than the given threshold λ. Accordingly, the action of a SU is defined as $a \in \mathcal{A}_s = \{1, \ldots, J\}$. Here $\mathcal{A}_s$ is the action set. If a SU would like to indicate that its average received energy is lower than the given threshold λ, it will keep silence and thus chooses action $a = 1$. Otherwise, it will choose action $a > 1$ reporting the a^{th} level received energy of the PU to the FC.

Then, the social norm for the soft fusion case is extended from (2) as $\boldsymbol{\Omega}' = [\Omega'_{i,j}]$ with

$$\Omega'_{i,j} = \begin{cases} 2, & \text{if } i = j = 1 \text{ or } 1 < i \leq J \text{ and } j = 2, \\ 1, & \text{otherwise,} \end{cases} \tag{23}$$

where $\Omega'_{i,j}$ is the reputation value assigned to a SU who takes action i while the FC's final decision is j.

Accordingly, the decision consistence matrix is extended from (5) as

$$\boldsymbol{D}_s = [D_{i,j}], \quad i = 1, \ldots J, j = 1, 2. \tag{24}$$

Here, the level 2 to level J can be equally quantized, such that we have $D_{j,1} = D_{j',1}$ and $D_{j,2} = D_{j',2}$ when $j, j' > 1$ and $j \neq j'$. Without loss of generality, we assume $D_{j,1} > D_{j,2}$ when $j = 1$ and $D_{j,2} > D_{j,1}$ when $j > 1$.

From (23) we can see that when the SU keeps silence while the FC's final decision is that the PU is not active, or when the SU reports its average received energy and the FC's final decision is that the PU is active, the SU is given a high reputation value. Otherwise, the SU is given a low reputation value. In such a way, SUs will be encourage to take part in the CSS and truthfully report their sensing results, which will be proved in the Theorem 5.

Theorem 5: The optimal action rule for SUs in the soft fusion case is $\boldsymbol{a}^* = (1, 2, \ldots, J)$ under condition $0 < c/g < (1-\xi)(1-P_F)\frac{N_s}{M}\frac{D_{2,2}-D_{2,1}}{E(\boldsymbol{\eta}^*)-1}$.

Proof: See proof in Appendix G. ■

D. Proposed Game for the Multiple Channel ($K > 1$) Case

In the previous subsections, we focus our discussion on the single-channel case, i.e., $K = 1$. However, our discussion can be easily extended to the multi-channel case where $K > 1$.

In the literature, it is common to assume that channels in the primary system are independent [31]. Therefore, to apply our indirect reciprocity scheme, the action of a SU only need to be redefined as $a_{i,j} = (a_{i,j,1}, \ldots a_{i,j,k}, \ldots, a_{i,j,K})$, where $\mathcal{A}_k$ is the SU's action set toward channel k, and $a_{i,j,k} \in \mathcal{A}_k$, $1 \le k \le K$, is the action that the SU will take toward channel k when its reputation is i and the average received signal energy on channel k is at the j^{th} level.

To assign the SU reputation, the FC exploits the same social norm defined in (2) for the hard fusion scenario or (23) for the soft fusion scenario. Then for a SU, say SU$_m$, its reputation distribution of all K channels is defined as $\boldsymbol{d}_m = (\boldsymbol{d}_{m,1}, \ldots, \boldsymbol{d}_{m,k}, \ldots, \boldsymbol{d}_{m,K})$, with $\boldsymbol{d}_{m,k} = (d_l, d_h)$ being its reputation distribution of channel k.

The reputation updating policy for each channel is also follows the one described in Fig. 2. With the updating policy, SU$_m$'s reputation distribution of taking action a for channel k at time index $n + 1$, i.e., $\boldsymbol{d}_{m,k}^{(n+1)}(a)$, can be updated by linearly combining the SU's original reputation distribution $\boldsymbol{d}_{m,k}^{(n)}$ and the immediate reputation distribution $\hat{\boldsymbol{d}}_{m,k}^{(n+1)}(a)$ with a weight ξ as

$$\boldsymbol{d}_{m,k}^{(n+1)}(a) = (1 - \xi)\hat{\boldsymbol{d}}_{m,k}^{(n+1)}(a) + \xi \boldsymbol{d}_{m,k}^{(n)}. \tag{25}$$

By following action rule $\boldsymbol{a}_m = (\boldsymbol{a}_{m,1}, \ldots, \boldsymbol{a}_{m,k}, \ldots, \boldsymbol{a}_{m,K})$, SU$_m$ obtains expected reputation distribution $\boldsymbol{d}_m = (\boldsymbol{d}_{m,1}(\boldsymbol{a}_{m,1}), \ldots \boldsymbol{d}_{m,k}(\boldsymbol{a}_{m,k}), \ldots, \boldsymbol{d}_{m,K}(\boldsymbol{a}_{m,K}))$, with which the SU can apply for a corresponding amount of vacant channel whenever it has data to forward. In a proportionally fair manner, the access time of the vacant channel k allocated to the SU with its reputation distribution of this channel, i.e., $\boldsymbol{d}_{m,k}$, will be

$$t(\boldsymbol{d}_{m,k}, \boldsymbol{\eta}_k) = (1 - P_F)\frac{N_s}{M}\frac{E(\boldsymbol{d}_{m,k}) - 1}{E(\boldsymbol{\eta}_k) - 1}. \tag{26}$$

The benefit the SU achieves from K channels thereby is

$$f_g(\boldsymbol{d}_m(\boldsymbol{a}_m), \boldsymbol{\eta}) = \sum_{k=1}^{K} g t(\boldsymbol{d}_{m,k}, \boldsymbol{\eta}_k), \tag{27}$$

where $\boldsymbol{\eta} = (\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_k, \ldots, \boldsymbol{\eta}_K)$ is the reputation distribution of the entire population for K channels, and $\boldsymbol{\eta}_k = \boldsymbol{\eta}_k(\boldsymbol{a}_{1,k}, \ldots, \boldsymbol{a}_{M,k})$ following (9) is the reputation distribution of channel k.

On the other hand, the SU's cost caused by following action rule $\boldsymbol{a}_m$ will be

$$f_c(\boldsymbol{a}_m) = \sum_{k=1}^{K} c(E(\boldsymbol{a}_{m,k}) - 1). \tag{28}$$

Then let $W(\boldsymbol{a}_m, \boldsymbol{a}_{-m})$ denote the expected utility of a SU when it takes action rule $\boldsymbol{a}_m$ while other SUs take action rules $\boldsymbol{a}_{-m}$, we have

$$\begin{aligned} W(\boldsymbol{a}_m, \boldsymbol{a}_{-m}) &= f_g(\boldsymbol{d}_m(\boldsymbol{a}_m), \boldsymbol{\eta}) - f_c(\boldsymbol{a}_m) \\ &= \sum_{k=1}^{K} g t(\boldsymbol{d}_{m,k}, \boldsymbol{\eta}_k) - \sum_{k=1}^{K} c(E(\boldsymbol{a}_{m,k}) - 1) \\ &= \sum_{k=1}^{K} W(\boldsymbol{a}_{m,k}, \boldsymbol{a}_{-m,k}), \end{aligned} \tag{29}$$

where $W(\boldsymbol{a}_{m,k}, \boldsymbol{a}_{-m,k})$, which is defined in (11), is SU$_m$'s expected utility obtained from channel k by following action rule $\boldsymbol{a}_{m,k}$.

Since $W(\boldsymbol{a}_{m,k}, \boldsymbol{a}_{-m,k})$, $k = 1, \ldots, K$, are independent with each other, to maximize W, the SU only needs to choose respectively the optimal action rule for channel k. In such a case, we have Theorem 6 as follows.

Theorem 6: $\forall m \in \mathcal{M}$, SU$_m$'s optimal action rule for K channels, i.e., $\boldsymbol{a}_m^* = (\boldsymbol{a}_{m,1}^*, \ldots, \boldsymbol{a}_{m,K}^*)$, is given as

$$W(\boldsymbol{a}_{m,k}^*, \boldsymbol{a}_{-m,k}^*) \ge W(\boldsymbol{a}_{m,k}', \boldsymbol{a}_{-m,k}^*), \quad \forall k \in \{1, \ldots, K\}, \tag{30}$$

where $\boldsymbol{a}_{m,k}' \in \mathcal{A}_k \setminus \boldsymbol{a}_{m,k}^*$ and $W(\cdot)$ is defined in (11).

Theorem 6 can be proved in a similar way as Theorem 1 for the hard fusion case or as Theorem 5 for the soft fusion case. We skip the proof here due to page limit.

If $\boldsymbol{a}_k^* = (\boldsymbol{a}_{1,k}^*, \ldots, \boldsymbol{a}_{M,k}^*)$ is also an ESS, we can approve that there exists a unique stationary reputation distribution for channel k in the CSS system, as stated in Lemma 3.

Lemma 3: There exists a unique stationary reputation distribution for channel k in the CSS system which is $\boldsymbol{\eta}_k^* = \boldsymbol{\eta}(\boldsymbol{a}_{1,k}^*, \ldots, \boldsymbol{a}_{M,k}^*)$.

Lemma 3 can be proved in a similar way as Theorem 4 and we also skip the proof here.

IV. Simulation

In this section, we will evaluate the effectiveness and efficiency of the proposed scheme in CSS of CRNs. Specifically, we verify the optimal action rule of the proposed game and its evolutionary stability. Then we evaluate the performance of the CSS system under our incentive mechanism by comparing it with some other CSS schemes. Finally, we exam at how much an extent our proposed scheme can overcome the cheating behaviors when the SUs untruthfully report their sensing results.

A. The Optimal Action Rule and Its Evolutionary Stability

We consider a scenario in which there are $M = 40$ SUs in the secondary system. we assume the PU locates at coordinate (0,0), and the SUs are randomly distributed in a circle centered at coordinate (1,0) with a radius of 0.3. The channels between the PU and SUs are Rayleigh fading channels where the channel coefficient is complex Gaussian, i.e., $h_m \sim \mathcal{CN}(0, \sigma_m^2)$ with $\sigma_m^2 = 1/r_m^{\theta}$, r_m being the distances between the PU and SU$_m$, and the path loss factor θ being set to 4. The transmission power of the PU is set to 5, and the noise variance of each channel is set to $\sigma_n^2 = 1$. The pattern of the PU's activity is assumed to be $P(H_0) = P(H_1) = 0.5$. The reputation distribution of each SU is initialized as $(0.5, 0.5)$,

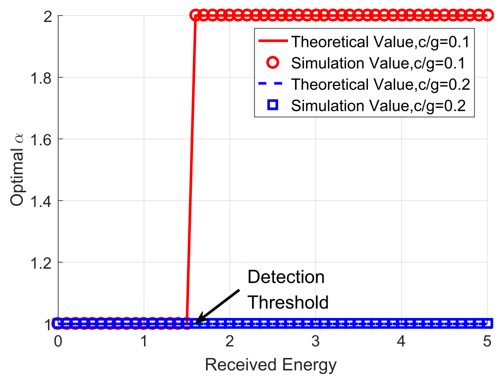


Fig. 4. The optimal action rules of SUs.

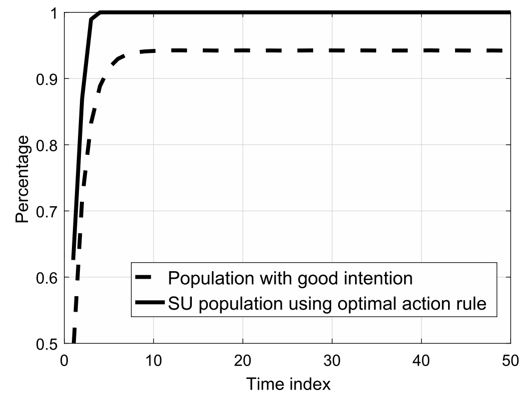


Fig. 5. The percentage of the SU population when c/g=0.1.

N_s is set to 5, and the discount weight ξ is set to 0.5. The P_d and P_f of the SU are set to 0.9 and 0.1, respectively. The threshold for energy detection is thereby set to 1.57 according to [30].

We first verify the optimal action rule of SUs when there is no perturbation effect. Taking hard fusion for example, we assume that all SUs in the system use the social norm defined in (2) and agree on the reputation updating policy shown in Fig. 2. Since the performance of the multi-channel case is similar to that of the single-channel case, the simulation is conducted only on the single-channel case and the results are shown in Fig. 4.

Noting that according to the simulation setting and (21), the critical cost-to-gain ratio is equal to 0.12. From Fig.4 we can see that in the case of $c/g = 0.1$ (i.e., $c/g < 0.12$), the SU chooses action $a^* = 1$ when its received energy is lower than the given threshold, and chooses action $a^* = 2$ otherwise. This is because in this case, the SU's cooperation gain will be greater than the cooperation cost, and it thus is willing to cooperate. While in the case of $c/g = 0.2$ (i.e., $c/g > 0.12$), the SU always chooses action $a^* = 1$ no matter what the received energy is. This is because in this case, the SU's cooperation gain will be smaller than the cooperation cost, and it thus refuses to participate in the game. In both cases, the curve of optimal action rules obtained by simulations matches the one that derived by the theoretical analysis as in Theorem 2, implying that our analysis is correct and solid.

Next, we will verify the evolutionary stability of the optimal action rule when $c/g = 0.1$, where the cost-to-gain ratio is lower than the critical value and the optimal action rule is $\boldsymbol{a}^{(2)}$.

Due to the perturbation effect such as noise, not all SUs will take the optimal action at the beginning. Therefore, we assume that in our system, only half SUs will take the optimal strategy $\boldsymbol{a}^{(2)}$ under the same condition as (21) while the other choose strategies $\boldsymbol{a}^{(1)}$, $\boldsymbol{a}^{(3)}$ or $\boldsymbol{a}^{(4)}$ randomly. We assume the SU stays in the system with probability of 95% and leaves the system with probability of 5%. For every SU who leaves, a new SU enters the system to keep the total population size constant. Here, we adopt the Wright-Fisher model [32] to investigate how the SUs update their actions. According to the Wright-Fisher model, the percentage of the population using action rule $\boldsymbol{a}^{(l)}$ at time index $n+1$, i.e., $y_l^{(n+1)}$, is proportional to the total payoff of the users using $\boldsymbol{a}^{(l)}$ in time index n, i.e.

$$y_l^{(n+1)} = \frac{y_l^{(n)} U_l^{(n)}}{\sum_{l=1}^{4} y_l^{(n)} U_l^{(n)}} \tag{31}$$

where $U_l^{(n)}$ is the total payoff of the users using $\boldsymbol{a}^{(l)}$ at time index n.

The results of percentage of population using the optimal strategy $\boldsymbol{a}^{(2)}$, and the results of percentage of the population with good intention are shown in Fig. 5. From Fig. 5, we can see that the optimal strategy spreads over the whole system within 10 iterations. After the whole population adopt the optimal strategy, there would be no mutation for SUs to deviate this state, which verifies that the optimal strategy $\boldsymbol{a}^{(2)}$ is the ESS under the simulation condition. From Fig. 5 we can also see, the distribution of the whole population with good intention converges within 10 iterations, which proves that strategy $\boldsymbol{a}^{(2)}$ is a desirable strategy since it leads to a good society with almost 95% of SUs having good intention.

According to the results we have obtained above, in the following, we only evaluate the performance of our system when it is in the steady stage where most of SUs take the optimal strategy $\boldsymbol{a}^{(2)}$ and have good reputation.

B. System Performance

In this subsection, we will evaluate the performance of our system in terms of the receiver operating characteristic curve (ROC), total throughput, and fairness. In the first two simulations, our proposed indirect reciprocity based CSS scheme (IRCS) is compared with three other existing schemes, i.e., random spectrum sensing (RCS), cooperative spectrum sensing (CCS) [33] and evolutionary cooperative spectrum sensing (ESSCS) [34]. In the RCS scheme, SUs send their sensing results to the FC in a random way, i.e., report and do not report with a probability of 0.5. In the CCS scheme, all the SUs send out their sensing results. While in the ESSCS scheme, the probability of a SU taking action a, $(a \in \{1, 2\})$ at instant $t + 1$ can be calculated as

$$P_a(t+1) = P_a(t) + \triangle[\bar{U}(a, \boldsymbol{a}_{-1}) - \bar{U}(a')]P_a(t), \tag{32}$$

where $\bar{U}(a, \boldsymbol{a}_{-1})$ is the SU's average payoff taking action a while other SUs taking actions $\boldsymbol{a}_{-1}$ at instant t, $\bar{U}(a')$ is the

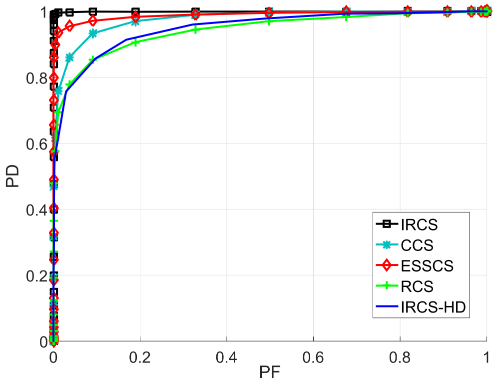


Fig. 6. The ROC curves of four different schemes.

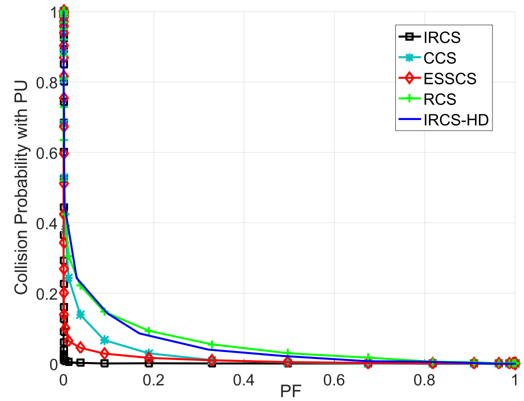
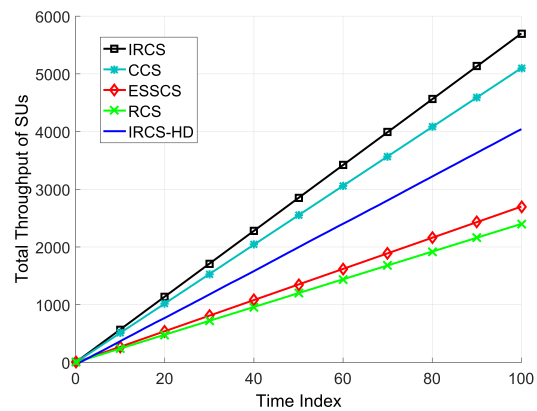
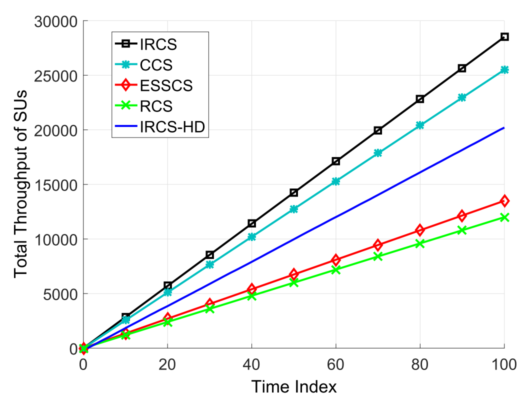


Fig. 7. Collision probability with PU under different schemes.



(a) Single-channel



(b) Multi-channel

Fig. 8. The total throughput of SUs.

SU's average payoff using mixed strategy a' at instant t, and $\triangle$ is the step size of adjustment.

Unless otherwise stated, the soft fusion approach EGC is used by the FCs in these four schemes. To demonstrate the gap of performance between different fusion approaches, hard fusion is also used in our proposed scheme which is denoted as "IRCS−HD".

We first evaluate the performance of different schemes in terms of ROC. The scenario is the same as that described in Section IV-A, and the results are demonstrated in Fig. 6. From Fig. 6 we can see that with the same fusion approach, the RCS scheme has the worst performance among all the schemes since SUs in the secondary system just share part of information to the FC. With the report of all SUs, the CCS scheme increases the fusion data in the FC, and therefore increases the fusion accuracy. Nevertheless, due to unstable environment condition, quite a part of SUs will suffer seriously from fading or shadow effect, which cause the SUs even cannot received the primary signal when the PU is active, these SUs then will send a certain degree of incorrect information to the FC, leading the FC to make wrong judgements. Taking this into consideration, the ESSCS scheme restrains the impact of sensing results of poor quality and further improve the sensing accuracy by decreasing (increasing) the sending probability when the estimated average energy is low (high). In our proposed system, unreliable SUs are completely constrained to silence and only the ones with good sensing quality send out their results, and the FC thereby can make the most accurate decisions. From Fig. 6 we can also see that for the same scheme, e.g. IRCS, the soft fusion outperforms the hard fusion in that it retains more information for the FC to make more accurate decisions.

From the view of CRNs, we also provide here the performance of the above schemes in terms of SUs' collision probability with PU, which, as shown in Fig. 7, can be easily derived from the results of the previous simulation.

Next, we study the performance of the schemes in terms of the total throughput of the secondary system, both in the single-channel scenario and the multi-channel scenario. While other settings are the same for these two scenarios, in the multi-channel scenario, we assume there are 5 PUs randomly distributed in a circle centered at coordinate (0,0) with a radius of 0.4, each owning one channel. The pattern of each PU's activity is assumed to be $P(H_0) = P(H_1) = 0.5$, and the results are shown in Fig. 8.

As we can see, the trends in the two scenarios are similar. This is because in our model, a SU will be selected to use the vacant channel with certain probability proportional to its expected reputation. Moreover, SUs are allocated the vacant channel in a way that can maximize the throughput of the second system, which is formulated into a maximum weight bipartite graph matching problem and solved by the

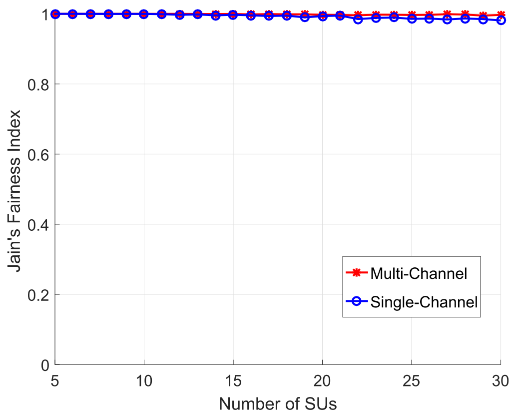


Fig. 9. The fairness index of our proposed scheme.

Kuhn-Munkres algorithm. In such a case, SUs in our proposed system is motivated to gain a good reputation by sending accurate information and thus in return have high throughput. However, in the CCS scheme, vacant channel is allocated through auction where the SU is assigned a channel that can only guarantee the SU's individual transmission rate but not the optimal one for the whole system's throughput, while with schemes RCS and ESSCS the vacant channel is just randomly allocated to SUs. As a result, the throughput of system with the CCS scheme is lower than that with our proposed scheme, and the throughput of system with RCS and ESSCS is the lowest. From 8 we can also see that from the perspective of throughput, the soft fusion obtains a better performance than the hard fusion.

In the third simulation, we study the fairness issue of the proposed scheme. To evaluate the fairness performance, we use the Jain's fairness index, i.e.,

$$F = \frac{(\sum_{m=1}^{M} T_m)^2}{M \sum_{m=1}^{M} T_m^2}, \tag{33}$$

where T_m is the access time of SU_m. Note that the fairness index varies between 0 and 1, with a larger value meaning more fairness. From Fig. 9 we can see that in both the single-channel and multi-channel cases, the Jain's Fairness Index of our proposed scheme is approaching to 1, and is almost unchanged along with the increase of the number of SUs, indicating that our scheme can achieve a good performance on fairness.

C. Anti-Cheating

According to the social norm, a SU gets a high reputation value only when its report is consistent with the decision of the FC, i.e., the decisions of most SUs. In such a case, none of the SUs has incentive to cheat if most SUs in the system are honest. However, it may be another case in a network where most SUs are malicious. Therefore, in this section, we are going to exam at how much an extent our proposed scheme can overcome the cheating behaviors when the SUs untruthfully report their sensing results.

For those 40 SUs, we change the percentage of malicious SUs from 0 to 100%, and the results of detection rate of cheating behaviors vs. the percentage of malicious SUs are

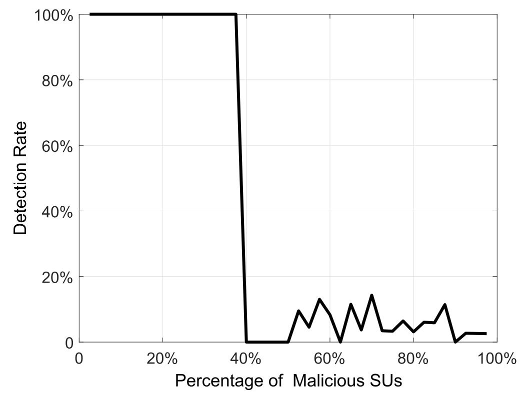


Fig. 10. The detection rate of malicious SUs.

shown in Fig. 10. From Fig. 10 we can see that when the percentage of malicious SUs is lower than 38%, the detection rate of cheating behaviors is near 100%, indicating the cheating actions of those SUs are successfully restrained by the proposed scheme.

V. Conclusion

In this paper, we propose a novel data sharing incentive scheme based on the indirect reciprocity game modeling. Such a scheme provides a general framework for analyzing the uniqueness and stationary of the reputation distribution as well as the optimal action rule for the sensors in the system. We further discuss the application of the proposed scheme into CSS in CRNs. Through theoretical analysis, we show that the SUs are stimulated through reputation to use the optimal action rule, i.e., to report the sensing result when the average received energy is above the given threshold and keep silence otherwise. With such an optimal action rule, the FC's fusion accuracy is greatly improved and at the same time the SU's energy is saved. Moreover, we derive the cost-to-gain ration under which such an optimal strategy is stable. Simulation results show that the optimal action rule will lead to a "good" society where most of the SUs have high reputation. The simulation results also demonstrate that compared with the state-of-the-art direct reciprocity based CSS schemes, our proposed scheme achieves better ROC, higher total throughput of the secondary system with convincing performance on fairness. By our proposed scheme, the cheating actions of malicious SUs can be successfully restrained as well.

Appendix A
Proof of Lemma 1

Proof: According to (8), $\forall \boldsymbol{a} \in \mathcal{A}_l$ and $\boldsymbol{d}^{(n)}$, we respectively have $\boldsymbol{d}^{(n+1)}(\boldsymbol{a}(1)) = (1-\xi)\hat{\boldsymbol{d}}^{(n+1)}(\boldsymbol{a}(1)) + \xi \boldsymbol{d}^{(n)}$ and $\boldsymbol{d}^{(n+1)}(\boldsymbol{a}(2)) = (1-\xi)\hat{\boldsymbol{d}}^{(n+1)}(\boldsymbol{a}(2)) + \xi \boldsymbol{d}^{(n)}$. Then we have $\boldsymbol{d}^{(n+1)}(\boldsymbol{a}) = P_{L_1}\boldsymbol{d}^{(n+1)}(\boldsymbol{a}(1)) + P_{L_2}\boldsymbol{d}^{(n+1)}(\boldsymbol{a}(2)) = (1-\xi)$ $\hat{\boldsymbol{d}}^{(n+1)}(\boldsymbol{a}) + \xi \boldsymbol{d}^{(n)}$. ■

Appendix B
Proof of Lemma 2

Proof: If the sensor keeps following action rule $\boldsymbol{a}$ at time index n, then according to Lemma 1, its expected reputation at that time index is updated as $\boldsymbol{d}^{(n)}(\boldsymbol{a}) = (1-\xi)\hat{\boldsymbol{d}}^{(n)}(\boldsymbol{a}) + \xi \boldsymbol{d}^{(n-1)}(\boldsymbol{a})$ where $\hat{\boldsymbol{d}}^{(n)}(\boldsymbol{a}) = (P_{L_1}\boldsymbol{e}_{\Omega_{\boldsymbol{a}(1),L_1}} + P_{L_1}\boldsymbol{e}_{\Omega_{\boldsymbol{a}(1),L_2}} + P_{L_2}\boldsymbol{e}_{\Omega_{\boldsymbol{a}(2),L_1}} + P_{L_2}\boldsymbol{e}_{\Omega_{\boldsymbol{a}(2),L_2}})\boldsymbol{D}$. We notice that $\hat{\boldsymbol{d}}^{(n)}(\boldsymbol{a})$ is deterministic and independent with n when $\boldsymbol{a}$ is given. Hence, we drop the subscript n in $\hat{\boldsymbol{d}}^{(n)}(\boldsymbol{a})$ and re-written it as

$$\begin{aligned}\boldsymbol{d}^{(n)}(\boldsymbol{a}) &= (1-\xi)\hat{\boldsymbol{d}}(\boldsymbol{a}) + \xi \boldsymbol{d}^{(n-1)} \\ &= (1-\xi^{n-1})(\hat{\boldsymbol{d}}(\boldsymbol{a}) - \boldsymbol{d}^{(1)}) + \boldsymbol{d}^{(1)}. \end{aligned} \tag{34}$$

Since $0 < \xi < 1$, when $\boldsymbol{a}$ is always followed, i.e., $n \to \infty$, we have a steady reputation distribution

$$\boldsymbol{d}^{(n)}_{n\to\infty}(\boldsymbol{a}) = (1 - \xi^{n-1}_{n\to\infty})(\hat{\boldsymbol{d}}(\boldsymbol{a}) - \boldsymbol{d}^{(1)}) + \boldsymbol{d}^{(1)} = \hat{\boldsymbol{d}}(\boldsymbol{a}). \tag{35}$$

■

APPENDIX C
PROOF OF THEOREM 1

Proof: Noting the player set $\mathcal{M} = \{1, \ldots, M\}$ and the set of action rule $\mathcal{A}_l$ are finite, according to [35], there exists a NE of the proposed game.

$\forall m \in \mathcal{M}$, given the optimal action rules of other sensors $\boldsymbol{a}^*_{-m}$, if sensor m's action rule $\boldsymbol{a}^*_m$ satisfies $W(\boldsymbol{a}^*_m, \boldsymbol{a}^*_{-m}) \geq W(\boldsymbol{a}'_m, \boldsymbol{a}^*_{-m})$ where $\boldsymbol{a}'_m \in \mathcal{A}_l \setminus \boldsymbol{a}^*_m$, then we have $\boldsymbol{a}^*_m = \arg\max_{\boldsymbol{a}_m \in \mathcal{A}_l} W(\boldsymbol{a}_m, \boldsymbol{a}^*_{-m})$, i.e., $\boldsymbol{a}^*_m$ is sensor m's best response to $\boldsymbol{a}^*_{-m}$. According to Definition 3, the action rule profile $\boldsymbol{a}^* = \{\boldsymbol{a}^*_1, \boldsymbol{a}^*_2, \ldots, \boldsymbol{a}^*_M\}$ constructed by $\boldsymbol{a}^*_m$, $m \in \mathcal{M}$, is the NE of the proposed game. ■

APPENDIX D
PROOF OF THEOREM 2

Proof: Let the current reputation distribution of a SU be $\boldsymbol{d}'$, and the stationary reputation distribution of the system be $\boldsymbol{\eta}^*$. After following action rule $\boldsymbol{a}$, the SU's immediate expected reputation distribution and the expected reputation distribution should respectively be

$$\begin{aligned}\hat{\boldsymbol{d}}(\boldsymbol{a}) = (&P_{L_1}\boldsymbol{e}_{\Omega_{\boldsymbol{a}(1),L_1}} + P_{L_1}\boldsymbol{e}_{\Omega_{\boldsymbol{a}(1),L_2}} \\ &+ P_{L_2}\boldsymbol{e}_{\Omega_{\boldsymbol{a}(2),L_1}} + P_{L_2}\boldsymbol{e}_{\Omega_{\boldsymbol{a}(2),L_2}})\boldsymbol{D}, \end{aligned} \tag{36}$$

and

$$\boldsymbol{d}(\boldsymbol{a}) = (1-\xi)\hat{\boldsymbol{d}}(\boldsymbol{a}) + \xi \boldsymbol{d}'. \tag{37}$$

Moreover, the expected action of action rule $\boldsymbol{a}$ can be calculated as

$$E(\boldsymbol{a}) = P_{L_1}\boldsymbol{a}(1) + P_{L_2}\boldsymbol{a}(2). \tag{38}$$

Given other SU's optimal action rules $\boldsymbol{a}^*_{-1}$, and substituting (37) and (38) into (16), we have the SU's expected utilities under the four action rules, i.e., $W(\boldsymbol{a}^{(1)}, \boldsymbol{a}^*_{-1})$, $W(\boldsymbol{a}^{(2)}, \boldsymbol{a}^*_{-1})$, $W(\boldsymbol{a}^{(3)}, \boldsymbol{a}^*_{-1})$ and $W(\boldsymbol{a}^{(4)}, \boldsymbol{a}^*_{-1})$.

Since $D_{1,1} > D_{1,2}$, we have

$$\begin{aligned} &W(\boldsymbol{a}^{(2)}, \boldsymbol{a}^*_{-1}) - W(\boldsymbol{a}^{(1)}, \boldsymbol{a}^*_{-1}) \\ &\quad = W(\boldsymbol{a}^{(4)}, \boldsymbol{a}^*_{-1}) - W(\boldsymbol{a}^{(3)}, \boldsymbol{a}^*_{-1}) \\ &\quad = g(1-P_F)(1-\xi)\frac{N_s}{M}\frac{P_{L_1}(D_{1,1}-D_{1,2})}{E(\boldsymbol{\eta}^*)-1} \\ &\quad\quad + cP_{L_1} > 0. \end{aligned} \tag{39}$$

That is, $W(\boldsymbol{a}^{(2)}, \boldsymbol{a}^*_{-1}) > W(\boldsymbol{a}^{(1)}, \boldsymbol{a}^*_{-1})$ and $W(\boldsymbol{a}^{(4)}, \boldsymbol{a}^*_{-1}) > W(\boldsymbol{a}^{(3)}, \boldsymbol{a}^*_{-1})$.

To find the optimal action rule, we just need to compare $W(\boldsymbol{a}^{(2)}, \boldsymbol{a}^*_{-1})$ and $W(\boldsymbol{a}^{(4)}, \boldsymbol{a}^*_{-1})$. If $W(\boldsymbol{a}^{(2)}, \boldsymbol{a}^*_{-1}) > W(\boldsymbol{a}^{(4)}, \boldsymbol{a}^*_{-1})$ holds, i.e.,

$$0 < c/g < (1-\xi)(1-P_F)\frac{N_s}{M}\frac{D_{2,2}-D_{2,1}}{E(\boldsymbol{\eta}^*)-1}, \tag{40}$$

then the action rule $\boldsymbol{a}^{(2)}$ is the optimal action rule. Otherwise, $\boldsymbol{a}^{(4)}$ is the optimal action rule.

In summary, we have the optimal action rule $\boldsymbol{a}^*$ as shown in (17). ■

APPENDIX E
PROOF OF THEOREM 3

Proof: From Lemma 2, if all SUs always follow action rule $\boldsymbol{a}^{(2)}$, a SU's expected reputation distribution at time index $n+1$ should be

$$\boldsymbol{d}^{(n+1)}(\boldsymbol{a}^{(2)}) = \hat{\boldsymbol{d}}(\boldsymbol{a}^{(2)}). \tag{41}$$

However, if a SU deviate from action rule $\boldsymbol{a}^{(2)}$ and choose to follow action rule $\boldsymbol{a}^{(l)}$ at time index $n+1$, $i \neq 2$, its expected reputation distribution will be

$$\boldsymbol{d}^{(n+1)}(\boldsymbol{a}^{(l)}) = (1-\xi)\hat{\boldsymbol{d}}^{(n+1)}(\boldsymbol{a}^{(l)}) + \xi \boldsymbol{d}^{(n)}(\boldsymbol{a}^{(2)}). \tag{42}$$

Moreover, we have

$$E(\boldsymbol{a}^{(l)}) = P_{L_1}\boldsymbol{a}^{(l)}(1) + P_{L_2}\boldsymbol{a}^{(l)}(2). \tag{43}$$

Substituting (41) and (43) into (16), we have the SU's utility when it keep follow the optimal action $\boldsymbol{a}^{(2)}$, i.e., $U(\boldsymbol{a}^{(2)}, \boldsymbol{a}^{(2)})$. Substituting (42) and (43) into (16), we have the SU's utility when it deviates from the optimal action rule and chooses action rule $\boldsymbol{a}^{(l)}$, $i \neq 2$, i.e., $U(\boldsymbol{a}^{(l)}, \boldsymbol{a}^{(2)})$.

Since $D_{1,1} > D_{1,2}$ and $D_{2,2} > D_{2,1}$, we have $U(\boldsymbol{a}^{(2)}, \boldsymbol{a}^{(2)}) > U(\boldsymbol{a}^{(1)}, \boldsymbol{a}^{(2)})$ and $U(\boldsymbol{a}^{(4)}, \boldsymbol{a}^{2}) > U(\boldsymbol{a}^{(3)}, \boldsymbol{a}^{(2)})$. According to (18), for $\boldsymbol{a}^{(2)}$ to be an ESS, $U(\boldsymbol{a}^{(2)}, \boldsymbol{a}^{(2)}) > U(\boldsymbol{a}^{(4)}, \boldsymbol{a}^{(2)})$ must hold, which results in (19). ■

APPENDIX F
PROOF OF THEOREM 4

Proof: According to the definition in (9), the reputation distribution of the whole system in time index n should be

$$\boldsymbol{\eta}^{(n)}(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_M) = \frac{1}{M}\sum_{m=1}^{M}\boldsymbol{d}^{(n)}(\boldsymbol{a}_m). \tag{44}$$

If sensor m always follows action rule $\boldsymbol{a}_m$, then according to Lemma 2, $\boldsymbol{\eta}^{(n)}(\cdot)$ turns to be

$$\begin{aligned}\boldsymbol{\eta}^* &= \boldsymbol{\eta}^{(n)}_{n\to\infty}(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_M) = \frac{1}{M}\sum_{\substack{m=1 \\ n\to\infty}}^{M}\boldsymbol{d}^{(n)}(\boldsymbol{a}_m) \\ &= \frac{1}{M}\sum_{m=1}^{M}\boldsymbol{d}^{(n)}_{n\to\infty}(\boldsymbol{a}_m) = \frac{1}{M}\sum_{m=1}^{M}\hat{\boldsymbol{d}}(\boldsymbol{a}_m), \end{aligned} \tag{45}$$

which is a unique and stationary reputation distribution of the entire system.

In the CSS system where (19) is satisfied, SUs always follows action rule $\boldsymbol{a}^{(2)}$. In such a case, we have

$$\boldsymbol{\eta}^* = \frac{1}{M}\sum_{m=1}^{M}\hat{\boldsymbol{d}}(\boldsymbol{a}^{(2)}) = (P_{L1}D_{1,2} + P_{L2}D_{2,1}, P_{L1}D_{1,1} + P_{L2}D_{2,2}). \quad (46)$$

From Theorem 2 and Theorem 3, the condition for $\boldsymbol{a}^{(2)}$ to be an evolutionary stable and optimal strategy thereby is (21). ■

APPENDIX G
PROOF OF THEOREM 5

Proof: Let the current reputation distribution of a SU be $\boldsymbol{d}'$, and the stationary reputation distribution of the system be $\boldsymbol{\eta}^*$. By contradiction, we suppose that there exists another optimal action rule $\boldsymbol{a}' \neq \boldsymbol{a}^*$, in which there is at least one element being not equal to that in $\boldsymbol{a}^*$, say $\boldsymbol{a}'(j) \neq \boldsymbol{a}^*(j)$ and $\boldsymbol{a}'(j) = j'$.

If $j = 1$, then $\boldsymbol{a}'(j) = j' > 1$. The SU's immediate average reputation distribution should be

$$\hat{\boldsymbol{d}}(\boldsymbol{a}'(j)) = \boldsymbol{e}_{\Omega_{j',1}}D_{j,1} + \boldsymbol{e}_{\Omega_{j',2}}D_{j,2} = (D_{1,1}, D_{1,2}). \quad (47)$$

On the other hand, the SU can truthfully reports its observation, i.e., takes the action $\boldsymbol{a}(j)$. Then its immediate average reputation distribution should be

$$\hat{\boldsymbol{d}}(\boldsymbol{a}(j)) = \boldsymbol{e}_{\Omega_{j,1}}D_{j,1} + \boldsymbol{e}_{\Omega_{j,2}}D_{j,2} = (D_{1,2}, D_{1,1}). \quad (48)$$

Given other SU's optimal action rules $\boldsymbol{a}^*_{-1}$, and substituting (47), (48) and $\boldsymbol{d}'$ into (16), respectively, we have the SU's expected utilities of following $\boldsymbol{a}'(j)$ and $\boldsymbol{a}(j)$ as $W(\boldsymbol{a}(j), \boldsymbol{a}^*_{-1})$ and $W(\boldsymbol{a}'(j), \boldsymbol{a}^*_{-1})$, and

$$W(\boldsymbol{a}'(j), \boldsymbol{a}^*_{-1}) - W(\boldsymbol{a}(j), \boldsymbol{a}^*_{-1}) = (1-\xi)g(1-P_F)\cdot\frac{N_s}{M}\frac{E(\hat{\boldsymbol{d}}(\boldsymbol{a}'(j))) - E(\hat{\boldsymbol{d}}(\boldsymbol{a}(j)))}{E(\boldsymbol{\eta}^*) - 1} - c(\boldsymbol{a}'(j) - \boldsymbol{a}(j)). \quad (49)$$

Since $E(\hat{\boldsymbol{d}}(\boldsymbol{a}'(j))) - E(\hat{\boldsymbol{d}}(\boldsymbol{a}(j))) = D_{1,2} - D_{1,1} < 0$ and $\boldsymbol{a}'(j) > \boldsymbol{a}(j)$, we have

$$W(\boldsymbol{a}'(j), \boldsymbol{a}^*_{-1}) - W(\boldsymbol{a}(j), \boldsymbol{a}^*_{-1}) < 0. \quad (50)$$

Similarly, if $j > 1$ and $j' > j$, or $j > 1$ and $1 < j' < j$, we have $E(\hat{\boldsymbol{d}}(\boldsymbol{a}'(j))) - E(\hat{\boldsymbol{d}}(\boldsymbol{a}(j))) = 0$ and $c(\boldsymbol{a}'(j) - 1) = c(\boldsymbol{a}(j) - 1)$. Hence, there is $W(\boldsymbol{a}'(j), \boldsymbol{a}^*_{-1}) - W(\boldsymbol{a}(j), \boldsymbol{a}^*_{-1}) = 0$.

Finally, if $j > 1$ and $j' = 1$, there are $\boldsymbol{a}'(j) = j'$ and $\boldsymbol{a}(j) = j$, then we have $\hat{\boldsymbol{d}}(\boldsymbol{a}'(j)) = (D_{j,2}, D_{j,1})$ and $\hat{\boldsymbol{d}}(\boldsymbol{a}(j)) = (D_{j,1}, D_{j,2})$.

Since $E(\hat{\boldsymbol{d}}(\boldsymbol{a}'(j))) - E(\hat{\boldsymbol{d}}(\boldsymbol{a}(j))) < 0$ and $c(\boldsymbol{a}'(j) - 1) - c(\boldsymbol{a}(j) - 1) < 0$, when $0 < c/g < (1-\xi)(1-P_F)\frac{N_s}{M}\frac{D_{j,2}-D_{j,1}}{(\boldsymbol{a}'(j)-\boldsymbol{a}(j))(E(\boldsymbol{\eta}^*)-1)}$, we have $W(\boldsymbol{a}'(j), \boldsymbol{a}^*_{-1}) - W(\boldsymbol{a}(j), \boldsymbol{a}^*_{-1}) < 0$.

Noting $D_{j,2} = D_{2,2}$ and $D_{j,1} = D_{2,1}$ when $j > 2$, we have

$$c/g < (1-\xi)(1-P_F)\frac{N_s}{M}\frac{D_{j,2} - D_{j,1}}{(\boldsymbol{a}'(j) - \boldsymbol{a}(j))(E(\boldsymbol{\eta}^*) - 1)} < (1-\xi)(1-P_F)\frac{N_s}{M}\frac{D_{2,2} - D_{2,1}}{E(\boldsymbol{\eta}^*) - 1}. \quad (51)$$

In summary, when $0 < c/g < (1-\xi)(1-P_F)\frac{N_s}{M}\frac{D_{2,2}-D_{2,1}}{E(\boldsymbol{\eta}^*)-1}$, there is no such an optimal action rule $\boldsymbol{a}' \neq \boldsymbol{a}^*$ where there is at least one element being not equal to that in $\boldsymbol{a}^*$ for the SU.

By deductive analysis, we conclude that there is no such an optimal action rule where there are more than one element being not equal to those in $\boldsymbol{a}^*$. Consequently, the optimal action rule for SUs in the soft fusion scenario is $\boldsymbol{a}^* = (1, 2, \ldots, J)$ when the condition in (51) satisfied. ■

REFERENCES

[1] D. Lahat, T. Adali, and C. Jutten, "Multimodal data fusion: An overview of methods, challenges, and prospects," *Proc. IEEE*, vol. 103, no. 9, pp. 1449–1477, Sep. 2015.
[2] L. Zhang, G. Ding, Q. Wu, Y. Zou, Z. Han, and J. Wang, "Byzantine attack and defense in cognitive radio networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 3, pp. 1342–1363, 3rd Quart., 2013.
[3] Y. N. Wang, J. F. Ye, G. J. Xu, Q. M. Chen, H. Y. Li, and X. R. Liu, "Novel hierarchical fault diagnosis approach for smart power grid with information fusion of multi-data resources based on fuzzy Petri net," in *Proc. IEEE FUZZ*, Beijing, China, Jul. 2014, pp. 1183–1189.
[4] D. Macii, A. Boni, M. D. Cecco, and D. Petri, "Tutorial 14: Multisensor data fusion," *IEEE Instrum. Meas. Mag.*, vol. 11, no. 3, pp. 24–33, Jun. 2008.
[5] G. Tan and S. A. Jarvis, "A payment-based incentive and service differentiation scheme for peer-to-peer streaming broadcast," *IEEE Trans. Parallel Distrib. Syst.*, vol. 19, no. 7, pp. 940–953, Jul. 2008.
[6] N. H. Tran, L. B. Le, S. Ren, Z. Han, and C. S. Hong, "Joint pricing and load balancing for cognitive spectrum access: Non-cooperation versus cooperation," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 5, pp. 972–985, May 2015.
[7] Z. Song, E. Ngai, J. Ma, X. Gong, Y. Liu, and W. Wang, "Incentive mechanism for participatory sensing under budget constraints," in *Proc. IEEE WCNC*, Istanbul, Turkey, Apr. 2014, pp. 3361–3366.
[8] S. Buchegger and J.-Y. Le Boudec, "Self-policing mobile ad hoc networks by reputation systems," *IEEE Commun. Mag.*, vol. 43, no. 7, pp. 101–107, Jul. 2005.
[9] A. Satsiou and L. Tassiulas, "Reputation-based resource allocation in P2P systems of rational users," *IEEE Trans. Parallel Distrib. Syst.*, vol. 21, no. 4, pp. 466–479, Apr. 2010.
[10] M. E. Mahmoud and X. Shen, "Credit-based mechanism protecting multi-hop wireless networks from rational and irrational packet drop," in *Proc. IEEE GLOBECOM*, Miami, FL, USA, Dec. 2010, pp. 1–5.
[11] H. Marzi and A. Marzi, "A security model for wireless sensor networks," in *Proc. IEEE CIVEMSA*, Ottawa, ON, Canada, May 2014, pp. 64–69.
[12] F. Gao, W. Yuan, W. Liu, W. Cheng, and S. Wang, "A robust and efficient cooperative spectrum sensing scheme in cognitive radio networks," in *Proc. IEEE ICC Workshops*, Cape Town, South Africa, May 2010, pp. 1–5.
[13] M. Pan and Y. Fang, "Bargaining based pairwise cooperative spectrum sensing for cognitive radio networks," in *Proc. IEEE MILCOM*, San Diego, CA, USA, Nov. 2008, pp. 1–7.
[14] K. Cao and Z. Yang, "A novel cooperative spectrum sensing method based on cooperative game theory," *J. Electron. (China)*, vol. 27, no. 2, pp. 183–189, Mar. 2010.
[15] V. Balaji and C. Hota, "Efficient cooperative spectrum sensing in cognitive radio using coalitional game model," in *Proc. IEEE iC3I*, Mysore, India, Nov. 2014, pp. 899–907.
[16] X. Hao, M. H. Cheung, V. W. S. Wong, and V. C. M. Leung, "A coalition formation game for energy-efficient cooperative spectrum sensing in cognitive radio networks with multiple channels," in *Proc. IEEE GLOBECOM*, Houston, TX, USA, Dec. 2011, pp. 1–6.
[17] W. Wang, B. Kasiri, J. Cai, and A. S. Alfa, "Distributed cooperative multi-channel spectrum sensing based on dynamic coalitional game," in *Proc. IEEE GLOBECOM*, Miami, FL, USA, Dec. 2010, pp. 1–5.
[18] H. Ohtsuki, Y. Iwasa, and M. A. Nowak, "Indirect reciprocity provides only a narrow margin of efficiency for costly punishment," *Nature*, vol. 457, pp. 79–82, Jan. 2009.
[19] M. A. Nowak and K. Sigmund, "Evolution of indirect reciprocity," *Nature*, vol. 437, pp. 1291–1298, Oct. 2005.
[20] Y. Chen and K. J. R. Liu, "Indirect reciprocity game modelling for cooperation stimulation in cognitive networks," *IEEE Trans. Commun.*, vol. 59, no. 1, pp. 159–168, Jan. 2011.

[21] B. Zhang, Y. Chen, and K. J. R. Liu, "An indirect-reciprocity reputation game for cooperation in dynamic spectrum access networks," *IEEE Trans. Wireless Commun.*, vol. 11, no. 12, pp. 4328–4341, Dec. 2012.

[22] Y. Gao, Y. Chen, and K. J. R. Liu, "Cooperation stimulation for multiuser cooperative communications using indirect reciprocity game," *IEEE Trans. Commun.*, vol. 60, no. 12, pp. 3650–3661, Dec. 2012.

[23] L. Xiao, Y. Chen, and K. J. R. Liu, "Anti-cheating prosumer energy exchange based on indirect reciprocity," in *Proc. IEEE Int. Conf. Commun.*, Sydney, NSW, Australia, Jun. 2014, pp. 599–604.

[24] C. Jiang, Y. Chen, and K. J. R. Liu, "Multi-channel sensing and access game: Bayesian social learning with negative network externality," *IEEE Trans. Wireless Commun.*, vol. 13, no. 4, pp. 2176–2188, Apr. 2014.

[25] Y. Zhu, W. Wu, D. Li, and L. Ding, "A double-auction-based mechanism to stimulate secondary users for cooperative sensing in cognitive radio networks," *IEEE Trans. Veh. Technol.*, vol. 64, no. 8, pp. 3770–3781, Aug. 2015.

[26] T. Wang, L. Song, Z. Han, and W. Saad, "Distributed cooperative sensing in cognitive radio networks: An overlapping coalition formation approach," *IEEE Trans. Commun.*, vol. 62, no. 9, pp. 3144–3160, Sep. 2014.

[27] B. Zhang, Y. Chen, C.-Y. Wang, and K. J. R. Liu, "A Chinese restaurant game for learning and decision making in cognitive radio networks," *Comput. Netw.*, vol. 91, pp. 117–134, Nov. 2015.

[28] L. Shen, H. Wang, W. Zhang, and Z. Zhao, "Blind spectrum sensing for cognitive radio channels with noise uncertainty," *IEEE Trans. Wireless Commun.*, vol. 10, no. 6, pp. 1721–1724, Jun. 2011.

[29] Y. Chen, B. Wang, W. S. Lin, Y. Wu, and K. J. R. Liu, "Cooperative peer-to-peer streaming: An evolutionary game-theoretic approach," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 10, pp. 1346–1357, Oct. 2010.

[30] Y.-C. Liang, Y. Zeng, E. C. Y. Peh, and A. T. Hoang, "Sensing-throughput tradeoff for cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 4, pp. 1326–1337, Apr. 2008.

[31] S. H. A. Ahmad, M. Liu, T. Javidi, Q. Zhao, and B. Krishnamachari, "Optimality of myopic sensing in multichannel opportunistic access," *IEEE Trans. Inf. Theory*, vol. 55, no. 9, pp. 4040–4050, Sep. 2009.

[32] R. A. Fisher, *The Genetical Theory of Natural Selection*. Oxford, U.K.: Clarendon, 1930.

[33] J. Rajasekharan and V. Koivunen, "Cooperative game-theoretic approach to spectrum sharing in cognitive radios," *Signal Process.*, vol. 106, pp. 15–29, Jan. 2015.

[34] 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, Mar. 2010.

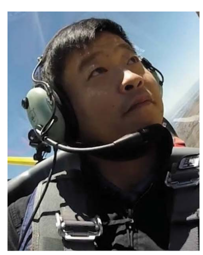
[35] T. Wang, L. Song, Z. Han, and W. Saad, "Overlapping coalition formation games for emerging communication networks," *IEEE Netw.*, vol. 30, no. 5, pp. 46–53, Sep./Oct. 2016.

Biling Zhang (M'12) received the B.S., M.S., and Ph.D. degrees from the Beijing University of Posts and Telecommunications (BUPT), Beijing, China, in 2000, 2003, and 2013, respectively. From 2011 to 2012, she was a Visiting Scholar with the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD, USA. She is currently an Associate Professor with BUPT. Her research interests include wireless resource allocation, wireless communications and networking, game theory, and smart grid.

Yan Chen (SM'14) received the B.S. degree from the University of Science and Technology of China in 2004, the M.S. degree from The Hong Kong University of Science and Technology in 2007, and the Ph.D. degree from the University of Maryland, College Park, MD, USA, in 2011. Since 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 a recipient of multiple honors and awards, including the 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, the Finalist of the Dean's Doctoral Research Award from the A. James Clark School of Engineering, the University of Maryland in 2011, and the Chinese Government Award for outstanding students abroad in 2010.

Jung-lang Yu (M'04) received the B.S. degree from National Cheng-Kung University, Tainan, Taiwan, in 1986, and the M.S. and Ph.D. degrees from National Taiwan University, Taipei, Taiwan, in 1988 and 1997, respectively, all in electrical engineering. From 1998 to 2000, he was an Associate Professor with the Department of Electronic Engineering, Tungnan University, Taipei. He is currently a Professor with the Department of Electrical Engineering and the Associate Dean of the College of Science and Engineering with Fu Jen Catholic University. His current research interests include massive MIMO, MIMO-OFDM systems, cooperative network communications, and space–time signal processing.

Bin Chen received the B.S. degree from the Beijing Institute of Technology in 2014, and the M.S. degree from the Beijing University of Posts and Telecommunications in 2017. He is currently an Engineer with Huawei Technologies Company Ltd. His current research interests include wireless communications and game theory.



Zhu Han (S'01–M'04–SM'09–F'14) received the B.S. degree in electronic engineering from Tsinghua University in 1997, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Maryland, College Park, MD, USA, in 1999 and 2003, respectively.

From 2000 to 2002, he was a Research and Development Engineer with JDSU, Germantown, MD, USA. From 2003 to 2006, he was a Research Associate with the University of Maryland. From 2006 to 2008, he was an Assistant Professor with Boise State University, Boise, ID, USA. He is currently a Professor with the Electrical and Computer Engineering Department and with the Computer Science Department, University of Houston, Houston, TX, USA. His current research interests include wireless resource allocation and management, wireless communications and networking, game theory, big data analysis, security, and smart grid. He received the NSF Career Award in 2010, the Fred W. Ellersick Prize of the IEEE Communication Society in 2011, the EURASIP Best Paper Award for the *Journal on Advances in Signal Processing* in 2015, IEEE Leonard G. Abraham Prize in the field of Communications Systems (best paper award in IEEE JSAC) in 2016, and several best paper awards in IEEE conferences. He is currently an IEEE Communications Society Distinguished Lecturer.