Evolutionary Game for Joint Spectrum Sensing and Access in Cognitive Radio Networks

Chunxiao Jiang*[†], Yan Chen*, Yang Gao*, and K. J. Ray Liu*
[illegible]
[†] [illegible]
[illegible]

[illegible]

[illegible]

l l b i p(a) l 1 o U s a l 1 + o (p) l 1 o $U_{\overline{sa}}$ l 1, o

1 s l 1 1 p(a) l 1 1 U s a l 1 + 1 P (a) l 1 1 $U_{\overline{sa}}|_{\mathcal{H}_1}$,

where p(a) l 1 o and p(a) l 1 1 denote the portion of SUs who access the primary channel given channel condition H o and $\mathcal{H}_1$, respectively, which can be calculated by

p(a) l 1 o p(a) (D o p(D o l)) 1 o P $\Big($a 1 - P fM(ps)$\Big)$),

p(a) l 1 1 p(a) (V, d 1 P D(o 1)) 1 1 P $\Big($a 1 - $P_d\Big)$,

where P fM(ps) is the false-alarm probability when M ps SUs cooperatively sense the primary channel.

Thus, we can derive the average utility of SUs who sense the primary channel 1 s, the average utility of SUs who do not sense 1 s, and the average utility of all SUs 1, as follows:

1 s P o 1 l s 1 + o p_1 1 1 s 1, 1

1 s P o 1 s l 1 + o p_1 1 s l 1, 1

1 P s 1 s + (1 - P s) 1 s,

where p_0 is the probability that the PU is absent, i.e., the probability of H o, and p_1 = 1 - p_0 is the probability of H 1. Combining and we can re-write the replicator dynamics of spectrum sensing by below.

C. -6@<;42H?AN 'M>2=;4GN ?7N.@64HAK=N$446GGN

Similar to the analysis of replicator dynamics of spectrum sensing, we should first calculate the average utility of SUs accessing the primary channel 1 a v and the average utility of all SUs given D o, 1 l v. The SUs' utilities with four pure strategies {s a, s a, s a, $\overline{s a}$} given D o can be written by where p(H o l)a(V,o) and p(H o l)(V a) are the probability that the PU is absent when SUs decide access and not access, respectively. U l(vo 1, o) and U l(V 1, d) are SUs' utility functions given {D o, H o} and {D o, H d}, respectively. Since given D o, the access behavior is independent with the status of the PU, we have p(H o l)a(V,o) p(H o l)(V a) p(H o l) v, where p(H o l) v can be calculated by the Bayes' rule given as

p(H o l) v = P o p(D o l) 1 o / P o p(D o l) 1 + o P 1 p(D o l) 1 1

$$= \frac{p_0\,(1 - P_f(Mp_s))}{1 - p_0P_f(Mp_s) - p_1P_d}.$$

Given the action of SUs, utilities are independent with D o, i.e., U l(vo 1, o) U l 1 o and U l(V 1, 1) U l 1 1. In such a case, we can derive the average utility of SUs who access and who do not access the primary channel 1 al v and 1 al v, and the average utility of all SUs 1 l v as follows:

1 al v P s U sal v (1 - P s) U sal v,

1 al v P s U s a v (1 - P s) −

1 l v P a 1 al v (1 - P a) 1 al v

Combining and we can re-write the replicator dynamics of spectrum access by below.

'N $>2<MG;GN ?7N(L?<KH;?>2A;<MN .H23<6N .HBH69MN

At equilibrium, we have Ps 0 and Pa 0. According to and we can get seven possible equilibria: (0, 0), (0, 1), (1, 0), (1, 1), (P s, 1), (1, P a), (P s, P a), where P s satisfies P fM(p s) $\left(\frac{\Theta_s}{R} - p_1P_d\right)/p_0$, P a satisfies J(M P a) − $\frac{(\Theta_a+R)\left(1-p_0P_f(M)-p_1P_d\right)}{p_0\left(1-P_f(M)\right)}$ and (P s, P a) is the solution of the following equations

$$\begin{cases} -\Theta_s+\Big(1-p_a+p_a\big(p_0P_f(Mp_s)+p_1P_d\big)\Big)R=0,\\ \frac{p_0\left(1-P_f(Mp_s)\right)}{1-p_0P_f(Mp_s)-p_1P_d}\mathbb{F}(Mp_a)-\Theta_a-p_sR=0. \end{cases} \tag{28}$$

According to the evolutionary game theory [7], [8], an equilibrium of the replicator dynamic equations is an ESS if and only if it is a locally asymptotically stable point in the dynamic system. In the following +6==2N " and /:6?A6=N " we will check which equilibria are ESSs.

+6==2N "#N The false-alarm probability P f is a decreasing function in terms of P s and the reward from channel access J is a decreasing function in terms of P a, i.e., $\frac{\cdot}{dp_s} < 0$ and $\frac{d\mathbb{F}(Mp_a)}{dp_a} < 0$.

Proof: This can be easily proved by taking derivatives on and . ■

/:6?A6=N"#N For the joint spectrum sensing and access evolutionary game, there are three ESSs (;, p) (1, 0), (1, P a) and (P s, P a) under different conditions of the rewards R s listed as follows:

Ps TP sl (- p) (- 1 s) TP sl (- p)- 8 s + $\Big($ l - p + P $\big($a P O PMfps (+) $p_1P_d\big)\Big)$ $%

Usal v = p(H o l)a(V,o) Usa (V, 1 + o) (1) - p(H o l)a(v,o)) UsalV(o, 1 1)

U sal v = p(H o l)a(V,o) U salV(o, 1 + o) (1) - p(H o l)a(v,o)) U salV(o, 1 1)

Usal v = p(H o l) (Va) UsalV(o, 1 + o) (1) - p(H o l) (a)) Usa (V, 1), 1

$U_{\overline{sa}}|_{\mathcal{D}_0}$ p(H o l)a(,o) V $U_{\overline{sa}}|_{(\mathcal{D}_0,\mathcal{H}_0)}$ + (1 - p(H o l)a(v,o)) $U_{\overline{sa}}|_{(\mathcal{D}_0,\ )}$

P a TP al (P a) (al v 1 - al v) TP al (P a) % po (- lP fM(ps)) / 1 P o P(M fps) p_1P_d J(M P a) - 8 a P R s %

$$(p_s, p_a) = \begin{cases} (1,0), & R > \frac{p_0\left(1-P_f(M)\right)}{1-p_0P_f(M)-p_1P_d} J(1) - 8\ a, \\ (1, P_a)', & R > \frac{\Theta_s}{1-p_{a1}\left(1-p_0P_f(M)-p_1P_d\right)}, \\ (P_{s2}, P_{a2}), & R < R', \end{cases}$$

$$R' = \frac{p_0^2 p_{s_2}\left(1-P_f(Mp_{s_2})\right)\frac{d\mathbb{F}(Mp_{a_2})}{dp_{a_2}} - \frac{p_0p_1(1-P_d)\mathbb{F}(Mp_{a_2})}{1-p_0P_f(Mp_{s_2})-p_1P_d}}{\left(1-p_0P_f(Mp_{s_2})-p_1P_d\right)^2}$$

d Pj(M pS)2

d P2

, E ?8N Due to page limitation, we show the proof in the supplementary information 7 8 ■

(!N $N ';GHA;3KH65N+62D;>9N$<9?A;H:=N7?AN(..N

In the above joint spectrum sensing and access evolutionary game, we have obtained the ESSs for SUs. Thus, a group of SUs can achieve the ESS using the replicator dynamics equations. We can see that solving these equations requires the exchange of utilities among all SUs to find the average utilities such as $\tilde{1}_s$ and $\tilde{1}_a$. However, in a distributed network, it is generally difficult to make each SU reveal such private information. In this section, we will present a distributed learning algorithm that can gradually converge to the ESS without private utility information exchange.

In the evolutionary biology, the Wright-Fisher model has been widely used to study the population reproduction dynamics under natural selection 7 8. The model is based on the assumption that the probability of an individual adopting a certain strategy is proportional to the expected utility of the population using this strategy. Let P_s and P_a be the probabilities of a SU sensing and accessing the primary channel, respectively. According to the Wright-Fisher model, P_s is proportional to the total utility of SUs sensing the channel. In such a case, SUs' strategy of spectrum sensing at time $t+1$, $P_s(t+1)$, can be calculated by

$$P_s(t+1) = \frac{P_s(t)U_s(t)}{P_s(t)\tilde{1}_s(t) + \left(1 - P_s(t)\right)\tilde{V}_s(t)},$$

where $U_s(t)$ and $U_s(t)$ are the average utilities of SUs who have sensed and not sensed the channel at the t-th time slot, respectively, the denominator is the average utility of all SUs, which is the normalization term that ensures $P_s + \tilde{p}_{\bar{s}} = 1$. If SUs observe that the PU is present after cooperative spectrum sensing at the beginning of the time slot, they will always not access the primary channel within this slot. Otherwise, each SU will access the channel with probability P_a. With the Wright-Fisher model, P_a is proportional to the total utility of SUs who choose to access the channel. In such a case, one SU's strategy of channel access, P_a, can be computed by

$$P_a(t'+1) = \frac{P_a(t')U_a(t')}{P_a(t')\tilde{1}_a(t') + \left(1 - P_a(t')\right)\tilde{1}_a(t')},$$

where t' and $t'+1$ represent the time slots when SUs observe that the PU is absent, $U_a(t')$ and $U_a(t')$ are the average utilities of SUs who have accessed and not accessed the channel at the t'-th time slot, respectively.

Based on the assumption that the number of SUs M is sufficiently large, the proportion of SUs who sense the primary channel is equal to the probability of one individual SU choosing to sense the channel, i.e., $p_s = P_s$. Similarly, $p_a = P_a$ if SUs observe the PU is absent. In such a case, we have

From we can see that when $\tilde{1}_s = \tilde{1}_{\bar{s}}$ or $P_s = 0$ or 1, $P_s = 0$, i.e., the equilibrium is achieved. According to the replicator dynamics equation of spectrum sensing, $U_s = U_{\bar{s}}$ and $P_s = 0$ or 1 are the solutions to the replicator dynamics. The same argument can be applied to P_a. Therefore, the Wright-Fisher model is equivalent to the replicator dynamics equation when M is sufficiently large. Note that although the total number of SUs M may be large, the ESS, e.g., (P_{s2}, P_{a2}), can ensure that only a portion of SUs will sense and access the channel. In such a case, the actual number of SUs that need to exchange sensing data or share the channel simultaneously is small, which can be seen from the simulation results.

From we can see that in order to update $P_s(t+1)$, each SU needs to learn about the average utilities $U_s(t)$ and $U_s(t)$. Here, we assume that each slot can be further divided into L subslots and each SU uses the same strategy of channel sensing and access during all L subslots, e.g., same P_s and P_a. During each time slot t or t', the SU computes the approximated average utilities $U_s(t)$, $U_s(t)$, $U_a(t')$ and $U_a(t')$ by

$$\tilde{1}_s(t) = \frac{\sum_{g=1}^{n} U_{sa}(t, g) + \sum_{h=1}^{m} U_{sa}(t, h)}{n+m},$$

$$\tilde{1}_s(t) = \frac{\sum_{g=1}^{n} U_{sa}(t, g) + \sum_{h=1}^{m} U_{sa}(t, h)}{n+m},$$

$$\tilde{1}_a(t') = \frac{\sum_{g'=1}^{n} U_{sa}(t', g') + \sum_{h'=1}^{m} U_{sa}(t', h')}{n+m}$$

$$\tilde{1}_a(t') = \frac{\sum_{g'=1}^{n} U_{sa}(t', g') + \sum_{h'=1}^{m} -}{n+m}$$

where g denotes the subslot when the SU chooses sensing and accessing, h denotes the subslot when the SU chooses sensing but not accessing, while g' and h' denote the subslot when SUs observe that the PU is absent. At the end of each slot t, the SU adjusts its strategy according to and. Thus, each SU can gradually learn the ESS. Based on each SU can gradually learn the ESS. In Algorithm we summarize the detailed procedure of the proposed distributed learning algorithm. In the following, we will examine the effectiveness of the learning algorithm through simulations.

```
Algorithm A Distributed Learning Algorithm for ESS.
• Given the time slot index t = 0, the SU initializes
  strategies Ps and Pa with Ps(0) and Pa(0).
for each time slot t do
  for n = 1 : L do
    • Sense the primary channel with probability Ps(t).
    • Exchange the sensing data with others through the
      signalling channel.
    if The SU observes that the PU is absent then
      • Access the sensed channel with probability Pa(t).
      • Estimate the average utilities of sensing and not
        sensing using  and
      • Estimate the average utilities of accessing and
        not accessing using  and
    else
      • Do not access the sensed channel.
      • Estimate the average utilities of sensing and not
        sensing using  and
    end if
  end for
  • Update the probability of sensing and accessing,
    Ps(t + 1) and Pa(t + 1), using (g) and
end for
```

IV. SIMULATION RESULTS

In this section, we conduct simulations to verify the effectiveness of our analysis. All the parameters used in the simulation are listed in Table I. We simulate the proposed learning algorithm with 20 SUs, and adjust the value of the reward R to see which ESS the system will converge.

TABLE I
PARAMETERS USED IN THE SIMULATION.

Parameter	Value	Parameter	Value
p_0	0.9		0.9
	-15dB	T_s	10ms
T_a	100ms		1MHz
	8MHz	SNR	-10dB
INR	-20dB	E_1	0.03mw/bit
E_2	0.5mw/s	E_3	2mw/s

$!N &?>L6A96>46N?7N(..N

In Fig. we show the convergence of the population states Ps and Pa, where the reward to SUs who only contribute sensing but do not access the channel R is set as 40, 100 and 400, respectively. In the simulations, Ps = 0.1 and Pa = 0.9 at the beginning, which means a large portion of SUs access the primary channel without contributing to channel sensing. From Fig. it can be seen that with our scheme, SUs will quickly give up such an undesired strategy, and the system finally converges to different ESSs under different R.

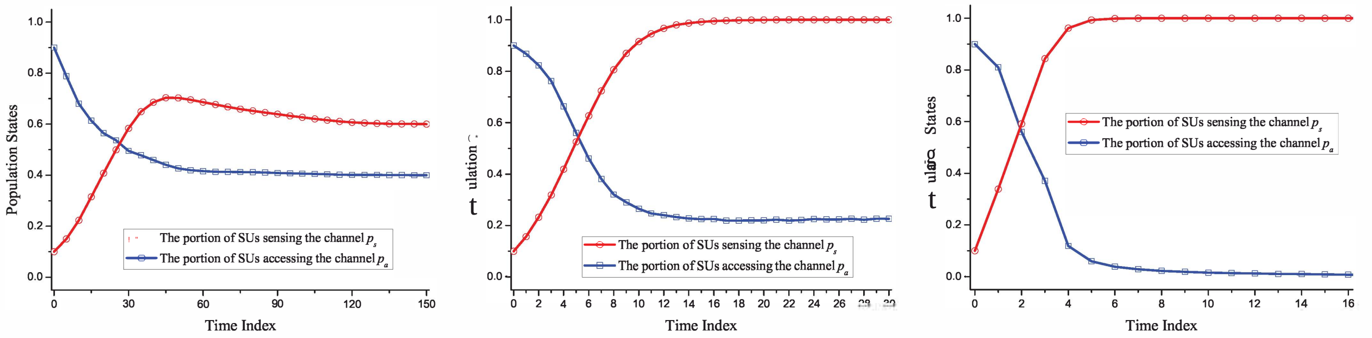
In Fig. 1-(a), when the reward R = 40, the ESS is (p; ,) = (0.6, 0.4) which corresponds to the case that part of SUs sense and access the primary channel. In such a case, Ps converging to 0.6 is because 60 percent of SUs cooperatively sensing the channel can already achieve relatively low false-alarm probability; and Pa converging to 0.4 is because more than 40 percent of SUs simultaneously accessing the channel will severely impair the throughput of each other.

In Fig. 1-(b), when the reward R = 100, the ESS is (p; ,) = (1, 0.25) which corresponds to the case that all SUs sense but part of them access the primary channel. Although the false-alarm probability is already low enough when Ps = 0.6, the increase of the reward R enhances the utility of sensing but not accessing, which attracts more SUs to sense but fewer SUs to access the primary channel.

In Fig. 1-(c), when the reward R = 400, the ESS is (p; , p) = (1, 0) which corresponds to the case that all SUs sense but no one accesses the primary channel. In such case, the reward R is too high that each SU can already obtain relatively high utility from just sensing without accessing the channel. Therefore, we can see that R should be properly set so that the network converges to a desired ESS.

%!N .H23;<;JN?7N(..N

In order to verify the stability of the ESS, let SUs deviate from the equilibrium when the system is at ESS. As shown in Fig. we first let SUs deviate from cooperative sensing by setting Ps = 0.1 at t = 200. It can be seen that both Ps and Pa return back to the ESS quickly after the perturbation. We can also see that Pa increases a little when Ps falls to 0.1. This is because a huge reduction of Ps leads to the decrease of both 1a and 1a. If the reduction 1a < that of 1a, Pa will be larger



(a) %* R40. (b) %* R100. (c) %* R400.

Fig. 1. Population states of the joint spectrum sensing and access evolutionary game.

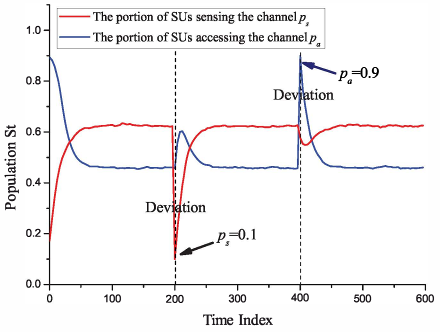


Fig. 2. Stability of ESS with R #* 50.

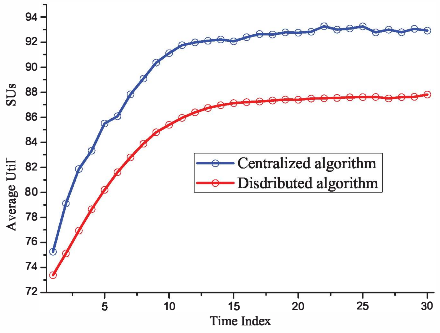


Fig. 3. Utility comparison.

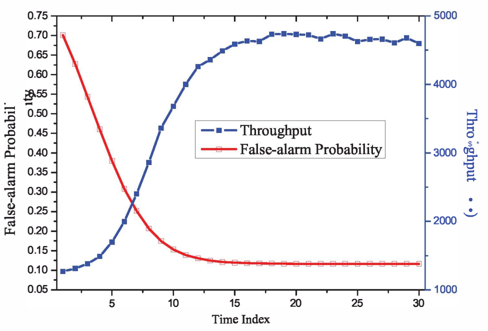


Fig. 4. Sensing and access performances.

than 0 according to (2), which results in the increasing of p_a. When $t = 400$, we let SUs deviate from the equilibrium again by setting $p_a = 0.9$. In such a case, the utility from channel access is extremely low and SUs will not sense and access the channel. That is why p_s begins to drop down when p_a is set to be 0.9.

&!N ,6C?A=2>46N (L2<K2H;?>N

We first compare the performance of our distributed learning algorithm with that of the centralized algorithm. In the centralized model, there is a data center in charge of collecting all SU's utility information in each slot and globally adjusting SUs's strategies p_s and p_a in the next time slot. Fig. i shows the comparison results in terms of the average utility of all SUs, from which we can see that the gap between our distributed algorithm and the centralized one is about i. Nevertheless, the centralized algorithm requires SUs to truthfully report their private utility information, while our distributed algorithm does not.

We further conduct simulation to evaluate the performance of our joint channel sensing and access algorithm. Fig. i shows the performance of SUs's false-alarm probability and throughput during the ESS converging process. We can see that along with the system converging to the ESS, the false-alarm probability gradually converges to the lowest limit, while the throughput gradually achieves the highest limit. Moreover, from Fig. i and Fig. i, we can see that the false-alarm probability is a decreasing function in terms of p_s and SUs' throughput is a decreasing function in terms of p_a, which is consistent with the results proved in +6==2N "!N

1!N CONCLUSION

In this paper, we analyzed how SUs should cooperate with each other in the joint spectrum sensing and access problem using evolutionary game theory. Through solving the joint replicator dynamic equations of channel sensing and accessing, we derived different ESSs under different conditions. Based on the nature selection theory, we proposed a distributed learning algorithm that enables SUs to achieve the ESS purely based on their own utility history. From simulation results, we can see that by adjusting the reward to the contributor, the population states of the network will converge to the desired ESS.

REFERENCES

[1] B. Wang and K. J. R. Liu. "Advances in cognitive radios: A survey," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 1, pp. 5-23, 2011.

[2] S. Shankar, C. Cordeiro, and K. Challapali, "Spectrum agile radios: utilization and sensing architectures," in *Proc. IEEE DySPAN*, 2005, pp. 160-169.

[3] H. Tang, "Some physical layer issues of wide-band cognitive radio systems," in *Proc. IEEE DySPAN*, 2005, pp. 151-159.

[4] A. Ghasemi and E. S. Sousa, "Collaborative spectrum sensing for opportunistic access in fading environments," in *Proc. IEEE DySPAN*, 2005, pp. 131-136.

[5] E. Visotsky, S. Kufner, and R. Peterson, "On collaborative detection of TV transmissions in support of dynamic spectrum sharing," in *Proc. IEEE DySPAN*, 2005, pp. 338-345.

[6] Q. Zhao, L. Tong, A. Swami, and Y. Chen, "Decentralized cognitive MAC for opportunistic spectrum access in Ad Hoc networks: a POMDP framework," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 3, pp. 589-600, 2007.

[7] C. Jiang, Y. Chen, K. J. R. Liu, and Y. Ren, "Renewal-theoretical dynamic spectrum access in cognitive radio networks with unknown primary behavior," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 3, pp. 406-416, 2013.

[8] Z. Ji and K. J. R. Liu, "Dynamic spectrum sharing: a game theoretical overview," *IEEE Commun. Mag.*, vol. 45, no. 5, pp. 88-94, 2010.

[9] A. A. El-Sherif and K. J. R. Liu, "Joint design of spectrum sensing and channel access in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 6, pp. 1743-1753, 2011.

[10] W. Saad, Z. Han, R. Zheng, A. Hjorungnes, T. Basar, and H. V. Poor, "Coalitional games in partition form for joint spectrum sensing and access in cognitive radio networks;' *IEEE J. Sel. Topics Signal Process.*, vol. 6, no. 2, pp. 195-209, 2012.

[11] R. Cressman, *Evolutionary Dynamics and Extensive Form Games*. Cambridge, MA: MIT Press, 2003.

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

[13] D. Niyato and E. Hossain, "Dynamics of network selection in heterogeneous wireless networks: An evolutionary game approach," *IEEE Trans. Veh. Tech.*, vol. 58, no. 4, pp. 2008-2017, 2009.

[14] X. Chen and J. Huang, "Evolutionarily stable open spectrum access in a many-users regime," in *Proc. IEEE GLOBECOM*, 2011, pp. 1-5.

[15] 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. Circuit Syst. Video Technol.*, vol. 20, no. 10, pp. 1346-1357, 2010.

[16] P. Baronti and etc, "Wireless sensor networks: A survey on the state of the art and the 802.15.4 and zigbee standards," *Computer Communication*, vol. 30, no. 7, pp. 1655-1695, 2007.

[17] I. F. Akyildiz, B. F. Lo, and R. Balakrishnan, "Cooperative spectrum sensing in cognitive radio networks: A survey," *Physical Communication*, vol. 4, no. 3, pp. 40-62, 2011.

[18] C. Jiang, Y. Chen, Y. Gao, and K. J. R. Liu, "Supplementary information for evolutionary game for joint spectrum sensing and access in cognitive radio networks," [Online]. Available: http://www.sig.umd.edu/jcx/GLOBECOM2013SuppInfo.pdf.

[19] R. Fisher, *The Genetical Theory of Natural Selection*. Clarendon Press, 1930.