

认知无线网络中基于随机学习博弈的信道分配与功率控制

汪志勇^{1,2}, 张沪寅¹, 徐 宁¹, 郝 圣¹

(1. 武汉大学计算机学院, 湖北武汉, 430072; 2. 湖北科技学院计算机科学与技术学院, 湖北咸宁, 437100)

摘 要: 传统的认知无线电频谱分配算法往往忽略节点的传输功率对网络干扰的影响, 且存在节点间交互成本高的问题. 为此, 通过量化传输功率等级, 以最大化弹性用户收益为目标, 构建联合频谱分配与功率控制非合作博弈模型, 证明了该博弈为严格潜在博弈且收敛到纳什均衡点. 进一步, 将随机学习理论引入博弈模型, 提出了基于随机学习的策略选择算法, 并给出了该算法收敛到纯策略纳什均衡点的充分条件及严格证明. 仿真结果表明, 所提算法在少量信息交互前提下能获得较高的传输速率, 并提升用户满意度.

关键词: 认知无线网络; 随机学习; 博弈论; 信道分配; 功率控制

中图分类号: TP393 **文献标识码:** A **文章编号:** 0372-2112 (2018)12-2870-08

电子学报 URL: <http://www.ejournal.org.cn> **DOI:** 10.3969/j.issn.0372-2112.2018.12.008

Channel Assignment and Power Control Based on Stochastic Learning Game in Cognitive Radio Networks

WANG Zhi-yong^{1,2}, ZHANG Hu-yin¹, XU Ning¹, HAO Sheng¹

(1. School of Computer, Wuhan University, Wuhan, Hubei 430072, China;

2. College of Computer Science and Technology, Hubei University of Science and Technology, Xianning, Hubei 437100, China)

Abstract: Traditional cognitive radio spectrum allocation algorithms tend to ignore the influence of transmission power on network interference and have the drawback of high interaction cost between nodes. In response to these problems, by quantifying transmission power levels, we formulate the channel assignment and power control problem as a distributed non-cooperative game, in which each second user's purpose is to maximize the elastic traffic rewards. Formally, the formulated game is proved to be an exact potential game and converges to Nash equilibrium (NE) point. Furthermore, introducing the stochastic learning theory into game model, we propose a strategy selection algorithm based on stochastic learning, then the sufficient condition and strict proof for the convergence of this algorithm to pure strategy NE point are given. Finally, Simulation results show that the proposed algorithm can achieve high system throughput and improve users' satisfaction with a small amount of interactions.

Key words: cognitive radio networks; stochastic learning; game theory; channel assignment; power control

1 引言

认知无线网络 (Cognitive Radio Networks, CRN) 允许次用户 (Second Users, SUs) 伺机接入主用户 (Primary Users, PUs) 未使用的空闲频段, 极大地提高了频谱资源利用率^[1]. 然而新的问题也随之产生, 比如: PUs 与 SUs 相互干扰, SUs 对无线频谱开放性特征难以捕捉等问题. 为此, 寻求更高效的频谱资源管理方法已成为该

领域研究的热点^[2].

围绕上述研究点, 文献[3]以信噪比干扰模型建立冲突图, 提出最大独立集启发式图着色信道分配算法. 文献[4]以干扰边及潜在干扰边提出基于差分演化的功率控制与信道分配算法. 文献[5]将链路选择与信道分配建模为混合整数线性规划问题. 文献[6]采用嵌套遗传算法求解频谱分配、功率选择和路由的组合优化问题. 但这些方法通常假设网络中存在一个中心节点

对频谱资源进行统一计算与分配,无法适应无线网络的动态可变性.因此,CRN 中寻求分布式频谱资源分配方法显得尤为重要.

博弈论具有较完备的数学理论及分布式实施特性,比较适合 CRN 频谱分配问题.文献[7]通过主次用户之间协作博弈来进行频谱分配,使系统达到斯坦伯格均衡.文献[8]利用信道特性及距离信息设计目标函数,以协作博弈方式实现用户速率最大化.文献[9]将无悔学习理论引入博弈模型,验证了相关均衡策略博弈进行信道分配的有效性.但是,这些博弈模型中用户策略选择依赖于对手用户所使用的策略,需要用户间大量信息交互,通信成本高.

以减少信息交互量为出发点,文献[10]基于信噪比阈值建立离散反馈的 Q 学习博弈模型.文献[11]综合考虑信道编码与干扰,针对 MIMO 系统提出了随机学习的信道分配算法.文献[12]基于用户需求建立随机博弈模型进行信道选择.几种算法进一步简化了问题求解难度,从不同角度提升网络性能.然而,上述学习模型均是基于用户干扰范围大小固定的假设,忽略了传输功率与用户干扰之间的内在联系,最终限制了频谱的使用效率.

为了尽量避免频谱分配过程中使用过强的假设条件,本文基于物理干扰模型,将传输功率离散化为若干等级,从弹性用户收益角度设计目标函数,提出一种仅需与同信道相邻用户进行信息交互的博弈模型,并从理论上证明其为严格潜在博弈,分析了纳什均衡解的存在及有效性.同时,为了进一步降低了问题求解的复杂度,引入随机学习理论,提出基于随机学习博弈的信道分配与功率控制算法,给出了该博弈收敛到纯策略纳什均衡的形式化证明.最后通过一系列仿真实验验证了所提算法的有效性.

2 网络模型

类似文献[12]中的 CRN 模型,网络中包含若干认知用户,每个认知用户代表一个分簇子网,簇首节点负责本子网的信道接入及传输功率控制^[13].图 1 右边部分为三个分簇用户的细节展示,该网络模型在实际的生活中也较为常见,比如分簇无线局域网、蜂窝无线网络等^[14].设定网络中随机分布 N 个 SUs,记为 $V = \{1, 2, \dots, N\}$,频谱资源被划分为 M 个相互正交的无线信道,记为 $C = \{1, 2, \dots, M\}$,定义 SUs 信道分配策略 $c_i = \{c_i^1, c_i^2, \dots, c_i^M\}$,当 $c_i^m = 1$ 表示用户 i 选择信道 m ,否则 $c_i^m = 0$.将通信功率范围离散化为 L 个数量等级,定义 SUs 功率使用策略 $q_i = \{q_i^1, q_i^2, \dots, q_i^L\}$,当 $q_i^l = 1$ 表示认

知用户 i 使用功率等级 l ,否则 $q_i^l = 0$.任意用户在某时刻最多只能使用一条信道传输,则有 $\sum_{m \in C} c_i^m \leq 1$ 和 $\sum_{l \in L} q_i^l \leq 1$.定义簇首与最远成员节点之间的距离为认知用户的通信半径,记为: d_i .

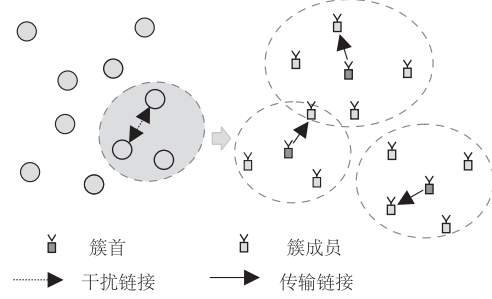


图 1 网络模型

2.1 信道干扰模型与用户传输速率

用户 i 对 j 产生干扰,则有 $c_i = c_j$,且 $d_{ij} < R_i + d_j$, d_{ij} 表示用户 i 与用户 j 簇首节点之间的距离, $R_i = (\tilde{q}_i^l / (q_N \cdot \alpha))^{1/\gamma}$ 为用户 i 在当前功率下的干扰范围,其中 \tilde{q}_i^l 表示功率等级为 l 的实际功率值, q_N 为环境噪声功率, α 为信噪比干扰阈值, γ 为路径损失系数.当认知用户 i 所选择的信道与功率等级分别为 c_i 和 q_i ,记为 $a_i = (c_i, q_i)$,定义用户 i 的相邻用户集为 $B_i = \{j \in V: d_{ij} < R_i + d_j\}$,定义与 i 同信道的干扰用户集为 $I_i = \{j: c_i = c_j, j \in B_i\}$,则有 $I_i \subseteq B_i$.同信道 SUs 实行等概率竞争交替接入^[15],将位于通信半径处的成员节点的可达传输速率作为用户的信道收益,则根据香农公式,用户 i 的传输速率为:

$$v_i(a) = \frac{c_i^m w_m}{1 + |I_i|} \log_2 \left(1 + \frac{\tilde{q}_i^l \times d_i^{-1/\gamma}}{q_N} \right) \quad (1)$$

其中 $a = \{a_1, a_2, \dots, a_N\}$ 为所有用户当前策略选择, w_m 为信道带宽, $|I_i|$ 表示集合 I_i 中元素的个数.

2.2 用户满意度与问题描述

本文构建弹性用户满意度信道容量模型,当用户传输速率高于设定阈值 v_{th} 时,用户满意度与 $v_i(a)$ 成正比,反之随 $v_i(a)$ 的递减而指数下降,其定义为:

$$r_i(a) = \begin{cases} v_i(a), & v_i(a) \geq v_{th} \\ v_{th} \exp\left(-\beta \frac{v_{th} - v_i(a)}{v_{th}}\right), & \text{其它} \end{cases} \quad (2)$$

其中 β 为衰减系数,取值越大函数曲率越大, $r_i(a)$ 随 $v_i(a)$ 减少越明显.整个网络用户满意度表示为:

$$U_r(a) = \sum_{i \in V} r_i(a) \quad (3)$$

形式上,联合分配问题可以表述为:

$$\begin{aligned}
a_{\text{opt}} &= \operatorname{argmax} U_r(a) \\
\text{s. t.} \\
\sum_{i \in V} c_i^m &\leq 1, i \in V, m \in C \\
\sum_{i \in V} q_i^l &\leq 1, i \in V, l \in L \\
\sum_{i \in V} q_i^l &\leq \sum_{i \in V} c_i^m, i \in V, m \in C, l \in L
\end{aligned} \quad (4)$$

式(4)前两个约束条件限定用户只能单信道及功率操作,第三个约束条件确保当 $c_i = 0$ 时 $q_i = 0$. 上述问题为典型的 NP 问题^[16],其解空间规模为 $(M \times L)^N$,当 $M = 3, L = 5, N = 20$ 时,即使速度百亿亿次/秒的超级计算机,搜索最优解耗时约十年之久. 因此,寻找问题的分布式解决方案尤为重要.

3 信道分配与功率控制博弈模型及纳什均衡

3.1 博弈模型

将 SUs 作为博弈参与者,构建 CRN 联合信道分配与功率控制博弈模型 $\Gamma = \{V, A, B_n, u_n\}$,其中 V 为 SUs 集, $A = \{a_n: \forall n \in N\}$ 为用户信道及功率策略集, B_n 为相邻用户集, u_n 为用户目标函数,定义如下:

$$\begin{aligned}
u_n(a_n, a_{B_n}) &= r_n(a_n, a_{B_n}) + \sum_{k \in B_n} r_k(a_k, a_{B_i}) \\
&\quad - [r_n(a_{n_0}, a_{B_n}) + \sum_{k \in B_n} r_k(a_k, a_{B_{i/n}})] \quad (5)
\end{aligned}$$

其中 a_{B_n} 为集合 B_n 中元素的策略状态, a_{n_0} 表示用户 n 退出信道竞争时的策略状态, $a_{B_{i/n}}$ 为用户 n 的相邻用户在 n 退出信道竞争时的策略状态. 然而博弈过程中, n 只对集合 I_n 中的用户产生干扰,则有:

$$r_n(a_n, a_{B_n}) = r_n(a_n, a_{I_n}) \quad (6)$$

且在用户 n 退出信道竞争前后,集合 $B_n - I_n$ 中用户目标函数值保持不变,即:

$$r_k(a_k, a_{B_i}) = r_k(a_k, a_{B_{i/n}}), k \in B_n - I_n \quad (7)$$

当 n 退出信道竞争时令 $r_n(a_{n_0}, a_{B_n}) = 0$,则有:

$$\begin{aligned}
u_n(a_n, a_{I_n}) &= r_n(a_n, a_{I_n}) \\
&\quad + \sum_{k \in I_n} (r_k(a_k, a_{I_i}) - r_k(a_k, a_{I_{i/n}})) \quad (8)
\end{aligned}$$

3.2 纳什均衡分析

若当前博弈策略为 $a_{\text{NE}} = \{a_1^*, a_2^*, \dots, a_N^*\}$,任意用户单方面改变策略都不能使其目标函数值增大,即:

$$u_n(a_n^*, a_{I_n}^*) \geq u_n(a_n, a_{I_n}^*), \forall n \in V, a_n^* \neq a_n \quad (9)$$

则称 a_{NE} 为纳什均衡点. 通常纳什均衡点的存在与否决定博弈能否收敛到最优解或近似最优解.

如果博弈模型存在一个潜在函数 $V(a)$,使得

$$V(a_n', a_{-n}) - V(a_n, a_{-n}) =$$

$$u_n(a_n', a_{-n}) - u_n(a_n, a_{-n}), \forall n \in V, a_n' \neq a_n \quad (10)$$

则该博弈为严格潜在博弈,且具备两个重要的性质:其一,潜在博弈至少存在一个纳什均衡解. 其二,潜

在函数的一个极值点代表一个纳什均衡解.

定理 1 联合信道分配与功率控制博弈 Γ 为潜在博弈,至少存在一个纯策略纳什均衡解.

证明:定义潜在函数为所有认知用户满意度函数值之和,表示为:

$$V(a_1, a_2, \dots, a_N) = V(a_n, a_{-n}) = \sum_{n \in V} r_n(a_n, a_{I_n}) \quad (11)$$

其中 $a_{-n} = (a_1, \dots, a_{n-1}, a_{n+1}, \dots, a_N)$,当用户 n 将策略 a_n 改变为 a_n' ,且 $a_n \neq a_n'$,则潜在函数改变量 ΔV 为:

$$\begin{aligned}
\Delta V &= V(a_n, a_{-n}) - V(a_n', a_{-n}) \\
&= (r_n(a_n, a_{-n}) - r_n(a_n', a_{-n})) \\
&\quad + \sum_{k \in I_n} (r_k(a_k, a_{I_i}) - r_k(a_k, a_{I_{i/n}})) \\
&\quad + \sum_{k \in I_n'} (r_k(a_k, a_{I_{i/n}'}) - r_k(a_k, a_{I_i}')) \\
&\quad + \sum_{k \in V - I_n - I_n' - n} (r_k(a_k, a_{I_i}) - r_k(a_k, a_{I_{i/n}})) \quad (12)
\end{aligned}$$

其中 I_n' 为 n 采用 a_n' 策略的干扰用户集. 为了简化公式推导,用 $\Delta X = \sum_{k \in X} (r_k(a_k, a_{I_i}) - r_k(a_k, a_{I_{i/n}}))$ 表示集合 X 中用户满意度值改变量,则(12)式可简写为:

$$\begin{aligned}
\Delta V &= (\Delta n - \Delta n') + \Delta I_n - \Delta I_n' \\
&\quad + \Delta(V - I_n - I_n' - n) \quad (13)
\end{aligned}$$

从式(13)中可知 ΔV 由四部分组成,第一部分为用户 n 自身满意度函数值的改变量,第二部分为用户 n 放弃 a_n 策略后干扰用户集 I_n 满意度函数值改变量,第三部分为用户 n 选择 a_n' 策略后干扰用户集 I_n' 满意度函数值改变量,第四部分为剩余用户满意度函数值改变量. 由于集合 $V - I_n - I_n' - n$ 中用户或处于 n 的干扰范围之外,或与 n 使用不同信道,故有

$$\sum_{k \in V - I_n - I_n' - n} (r_k(a_k, a_{I_i}) - r_k(a_k, a_{I_{i/n}})) = 0 \quad (14)$$

因此联合(12)~(14)有

$$\Delta V = (\Delta n - \Delta n') + \Delta I_n - \Delta I_n' \quad (15)$$

根据式(8),用户 n 在不同策略下目标函数改变量 Δu_n 为:

$$\begin{aligned}
\Delta u_n &= u_n(a_n, a_{-n}) - u_n(a_n', a_{-n}) \\
&= r_n(a_n, a_{I_n}) + \sum_{k \in I_n} (r_k(a_k, a_{I_i}) - r_k(a_k, a_{I_{i/n}})) \\
&\quad - r_n(a_n', a_{I_n}') - \sum_{k \in I_n'} (r_k(a_k, a_{I_i}') - r_k(a_k, a_{I_{i/n}'})) \\
&= (\Delta n - \Delta n') + \Delta I_n - \Delta I_n' \quad (16)
\end{aligned}$$

由式(10)、(15)、(16)可知联合信道与功率分配博弈 Γ 为潜在博弈,至少存在一个纯策略纳什均衡点,证毕.

为了达到收敛,SUs 通常采用最优响应方式,需知道对手用户策略而进行大量交互,并遍历所有可用策略,估算用户收益,计算量大. 为突破上述限制,本文将随机学习理论引入博弈模型,进一步降低问题求解难

度,同时保证算法收敛及用户收益.

4 随机学习策略选择算法及收敛分析

4.1 随机学习博弈

构建随机学习博弈 (Stochastic Learning Game, SLG), 扩展纯策略博弈 $\Gamma = \{V, A, I_n, u_n\}$ 为混合策略博弈 $\Gamma' = \{V, P, I_n, u_n\}$, 其中 $P = (p_1, p_2, \dots, p_N)$ 为系统混合策略选择概率规划, $p_n = \{p_n^{c,q} : 1 \leq c \leq M, 1 \leq q \leq L\}$ 包含 $M \times L$ 元素, 其中 $p_n^{c,q}$ 表示用户 n 选择策略行为 $a_n = (c, q)$ 的概率, 且满足 $\sum_c \sum_q p_n^{c,q} = 1, p_n^c = \sum_q p_n^{c,q}$ 及 $p_n^q = \sum_c p_n^{c,q}$, p_n^c 和 p_n^q 分别为用户 n 选择信道 c 和使用功率等级 q 的概率, 博弈过程中用户依概率进行策略选择, 直至算法收敛, 伪代码如算法 1 所示.

算法 1 随机学习信道与功率选择算法

- 1: 初始化. 令 $t=0, V_1 = V$, 初始化联合信道与功率策略空间选择概率为 $p_n^{c,q}(t) = 1/(M * L)$.
- 2: 策略选择. 在集合 V_1 随机选择用户 n 进行博弈, 根据概率 $p_n^{c,q}$ 确定用户行为策略 $a_n(t) = (c, q)$.
- 3: 目标函数值计算. 每个时隙周期内, 所有用户进行信道感知、竞争及数据传输. 时隙周期尾, 用户根据式 (8) 计算本次迭代目标函数值 $u_n(a)$.
- 4: 策略概率更新. 所有用户依规则 (17) 更新 $t+1$ 时刻策略选择概率:

$$p_n^{c,q}(t+1) = p_n^{c,q}(t) + b\tilde{u}_n(t)(1 - p_n^{c,q}(t)), a_n(t) = (c, q)$$

$$p_n^{c,q}(t+1) = p_n^{c,q}(t) - b\tilde{u}_n(t)p_n^{c,q}(t), a_n(t) \neq (c, q) \quad (17)$$
 其中, $0 < b < 1$ 为步长因子, $\tilde{u}_n(t)$ 为归一化目标函数值, 计算方式为:

$$\tilde{u}_n(t) = u_n(a_n, a_n) / \max(r_n(a_n', a_n')) \quad (18)$$
 $\max(r_n(a_n', a_n'))$ 可由用户 n 独占信道最大功率传输计算得到. 为了确 $\tilde{u}_n(t) \in [0, 1]$, 引入归一化调整因子 ε 对其做适当调整

$$\tilde{u}_n(t) = \max(0, \min(1, \tilde{u}_n(t) + \varepsilon)) \quad (19)$$
- 5: $V_1 = V_1 - n$, 若 $V_1 \neq \emptyset$, 返回步骤 2, 否则令 $V_1 = V$.
- 6: $t = t + 1$, 当算法满足停止条件时迭代结束, 否则返回步骤 2 继续执行.

4.2 收敛性分析

由式 (17) 可知, 策略概率更新规则 $\{P(t), t \geq 0\}$ 实际上是一个马尔柯夫更新过程, 当前状态由 N 个 $M \times L$ 概率矩阵构成, 策略空间为 $\Omega = \{P \in [0, 1]^{M * L}\}$, 用户基于随机学习博弈的目标函数为:

$$g_n(P) = \sum_{a_1, \dots, a_N} u_n(a_1, \dots, a_N) \prod_k p_k^{a_k} \quad (20)$$

当用户 n 采用纯策略 $a_n = (c, q)$, 其它用户采用混合策略时, 其目标函数 $h_{na_n}(P)$ 为:

$$h_{na_n}(P) = \sum_{a_{-n}} u_n((c, q), a_{-n}) \prod_{k \neq n} p_k^{a_k} \quad (21)$$

当步长因子足够小时, 更新过程 $\{P(t), t \geq 0\}$ 收敛到特定常微分方程的解, 并且方程的稳定点等价于过程 $\{P(t), t \geq 0\}$ 的纳什均衡点^[17], 常微分方程表示为:

$$\frac{dP}{dt} = f(P), P(0) = P_0 \quad (22)$$

其中 P_0 为初始策略概率, $f: \Omega \rightarrow [0, 1]^{M * L}$ 为条件期望概率, 代表当前时刻概率改变量的期望值, 表示为:

$$f(P) = E[G(P(t), a(t), u(t)) | P(t) = P] \quad (23)$$

基于此, 定理 2 给出了所提随机学习博弈收敛到纯策略纳什均衡点充分条件.

定理 2 假定存在有界可微函数 $H(P): P \rightarrow R$ 对于常量 $\kappa > 0$ 满足

$$\frac{\partial H(P)}{\partial p_n^{a_n}} = \kappa h_{na_n}(P), \forall n, a_n \quad (24)$$

则所提随机学习信道分配与功率控制算法收敛到纯策略纳什均衡.

证明: 记 $p_n^{a_n}$ 表示用户 n 选择纯策略 $a_n = (c, q)$ 的概率, 联合式 (17)、(21) ~ (23), 则有:

$$\begin{aligned} \frac{dp_n^{a_n}}{dt} &= f(P) \\ &= (p_n^{a_n}) (1 - p_n^{a_n}) E(\tilde{u}_n | (a_n, P_{-n})) \\ &\quad + \sum_{a_n', a_n' \neq a_n} p_n^{a_n'} (-p_n^{a_n}) E(\tilde{u}_n | (a_n', P_{-n})) \\ &= p_n^{a_n} \sum_{a_n'} p_n^{a_n'} [h_{na_n}(P) - h_{na_n'}(P)] \end{aligned} \quad (25)$$

由式 (24)、(25) 可得:

$$\begin{aligned} \frac{dH(P)}{dt} &= \sum_{n, a_n} \frac{\partial H(P)}{\partial p_n^{a_n}} \frac{dp_n^{a_n}}{dt} \\ &= \kappa \sum_n \sum_{a_n} \sum_{a_n'} p_n^{a_n} p_n^{a_n'} [(h_{na_n}(P))^2 - h_{na_n}(P) h_{na_n'}(P)] \\ &= \kappa \sum_n \sum_{a_n} \sum_{a_n' > a_n} p_n^{a_n} p_n^{a_n'} [(h_{na_n}(P)) - (h_{na_n'}(P))]^2 \geq 0 \end{aligned} \quad (26)$$

将所有策略线性排列, $a_n > a_n'$ 表示 a_n' 策略位于 a_n 之前. 由式 (25)、(26) 可知 $H(P)$ 是关于时间 t 的非减函数, $H(P)$ 在 $f(P) = 0$ 处取得极值点, 即常微分方程式 (22) 的稳定点, 更新过程 $\{P(t), t \geq 0\}$ 的纳什均衡点. 令 $H(P) = g_n(P)$, 为用户使用混合策略的目标函数值之和, 则有:

$$\begin{aligned} H(P) &= \sum_{a_1, \dots, a_N} u_n(a_1, \dots, a_N) \prod_k p_{ka_k} \\ &= \sum_{a_n} p_n^{a_n} h_{na_n}(P) \end{aligned} \quad (27)$$

$$\begin{aligned} \frac{\partial H(P)}{\partial p_n^{a_n}} &= \sum_{a_{-n}} u_n((c, q), a_{-n}) \prod_{k \neq n} p_k^{a_k} \\ &= h_{na_n}(P) \end{aligned} \quad (28)$$

满足条件式 (24), 由此可知本文所提随机学习信道分配与功率控制博弈收敛到纯策略纳什均衡, 证毕.

5 实验及分析

实验场景设定为 $100 * 100m^2$ 区域随机部署 N 个用户, 用户通信半径 d_n 在范围 $[5, 20]$ 随机分布, 路径损

失系数 $\gamma = 4$, 收益衰减系数 $\beta = 4$, 信道带宽 $w_c = 1$ MHz. 最小传输功率 $p^{\min} = -30$ dBm, 最大传输功率 $p^{\max} = 0$ dBm, 噪声功率为 $p_N = -80$ dBm, 信噪比干扰阈值为 $\alpha = 10$ dB, 默认速率阈值为 $V_{th} = 1.5$ Mbps/s, 策略概率更新步长 $b = 0.15$.

实验从收敛性、平均传输速率及归一化网络满意度来衡量算法的性能. 对比算法主要为几种常见的分配算法. (1) 本文所提随机学习博弈算法 (Stochastic Learning Game, SLG). (2) 完全信息动态博弈 (Complete Information of Dynamic Game, CIDG), 利用公式(2)为目标函数的一种最佳响应博弈模型. (3) 本地博弈 (Local Game), 利用式(2)为目标函数的随机学习算法. (4) 随机选择算法 (Random Selection).

5.1 收敛性

预设 $N = 20, M = 3, L = 6$, 图2和图3分别为随机选取单个SU的信道和功率策略概率演化过程. 从图中可以看出经过200次迭代后, 用户信道选择概率向量由初始值 $\{1/3, 1/3, 1/3\}$ 最终收敛到 $\{0, 0, 1\}$, 并维持恒定不变. 同时用户在6个功率等级上的选择概率也同样的表现, 结论与定理1和定理2相符.

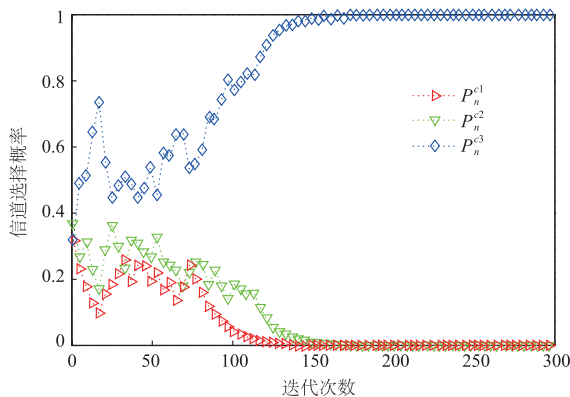


图2 不同信道选择概率演化过程

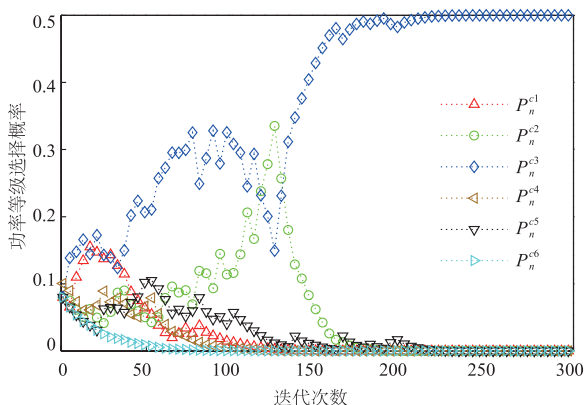


图3 不同功率等级选择概率演化过程

图4展示了算法收敛后网络用户的拓扑、信道和功

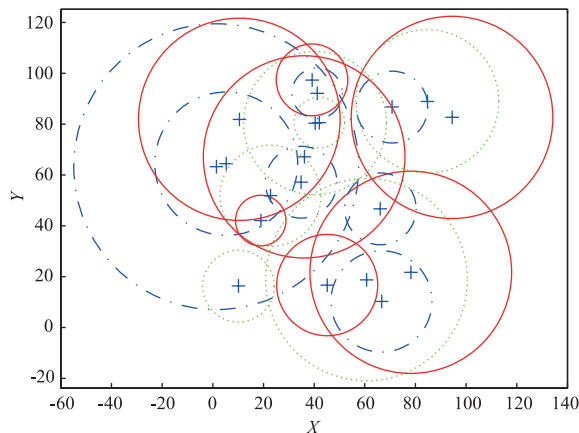


图4 信道与功率选择及干扰范围

率选择及干扰范围信息. 符号“+”为簇首所在位置, 不同颜色线条代表不同的信道选择, 不同半径代表不同功率等级形成的干扰范围. 从图中可以看出用户策略分布于不同的信道与功率等级, 重复实验还发现处于拓扑外围的SUs功率等级较高, 而处于拓扑内部的SUs则选择相对低的功率等级, 这正是SLG算法为减少干扰的一种自适应调节行为.

5.2 用户传输速率

本节实验考查不同环境参数下, 用户的信道收益情况对比, 实验结果取多次平均值. 图5为四种算法在不同用户数量下平均用户传输速率的变化曲线. 总体上用户传输速率随网络规模增大而减少, 因为信道资源一定时, 位于同一信道干扰范围内竞争用户数增多导致. 其次SLG算法性能与CIDG最为接近, 尤其是网络密度较大的时候, 证明了SLG算法所构建潜在博弈的有效性. 本地博弈及随机选择算法缺少信息观察存在盲目性导致效率较低.

图6中平均用户传输速率随信道资源的增多而增大, 同时四种算法间差异增大, CIDG因完全掌握环境信息, 理论上当迭代次数趋向无穷大时, 能找到系统最优解. 相比之下, SLG模型平均减少了近 $(M-1) * N/M$ 的交互量, 传输速率仍能逼近CIDG, 进一步说明了SLG算法的有效性.

将用户功率范围分别进行2、4、11、16四种不同的等级划分, 不同数量用户采用SLG算法所获得的传输速率对比情况如图7所示. 从图中可知功率等级划分越精细, SLG算法能对功率实施更有效性的控制, 获得高的传输速率.

5.3 归一化网络满意度

定义传输速率超过阈值的用户数量与总用户数量的比值为归一化网络满意度, 记为 $U_{net} = |\{n: n \in N, v_n > v_{th}\}| / N$.

设定 $M = 3, L = 6$ 时, 从图8可以看出, SLG在 $N \leq$

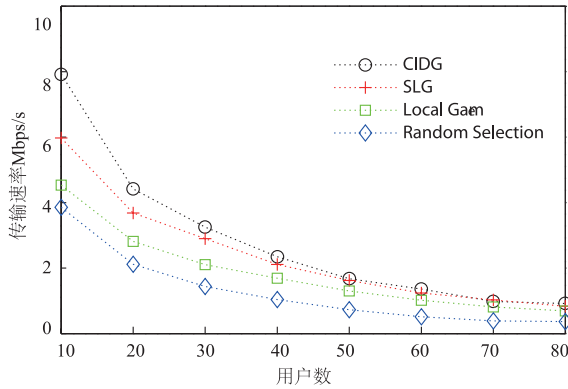


图 5 不同用户数量传输速率对比 ($M=3, L=6$)

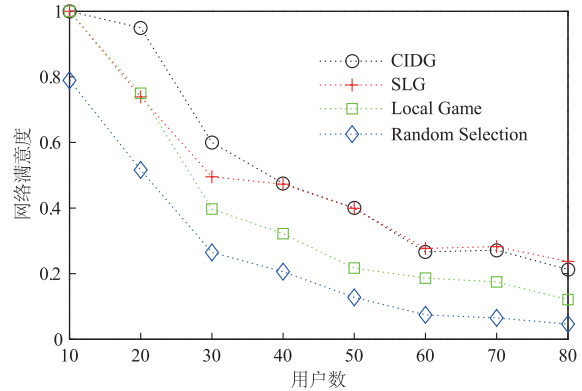


图 8 不同用户数量网络满意度对比

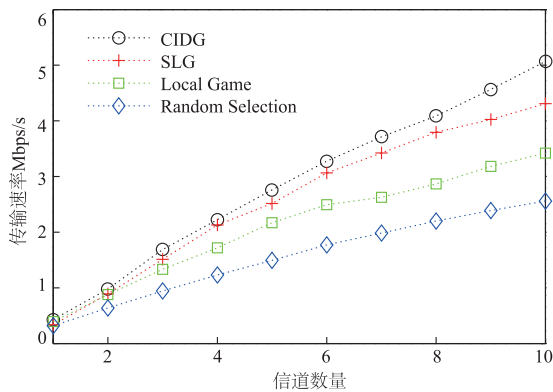


图 6 不同信道数量传输速率对比 ($N=80, L=6$)

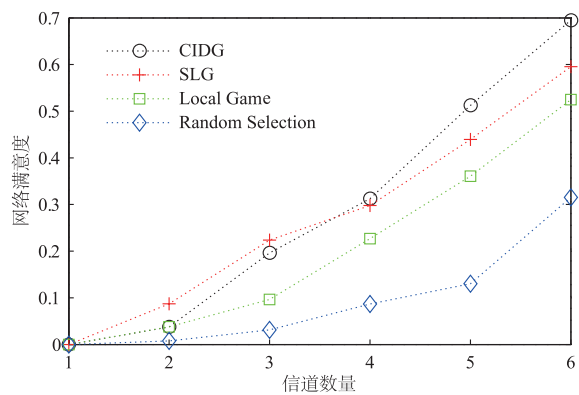


图 9 不同信道数量网络满意度对比

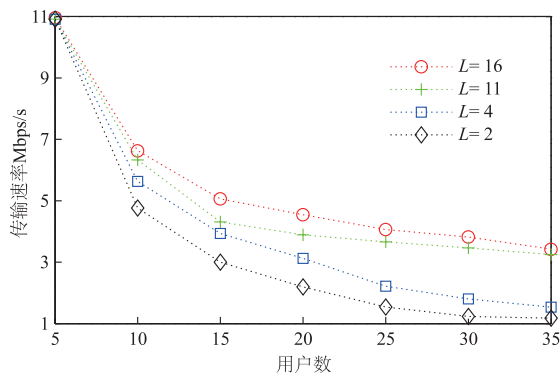


图 7 不同功率等级传输速率对比 ($M=3$)

40 时满意度低于 CIDG,而在 $N > 40$ 时满意度逼近 CIDG,甚至超过 CIDG 算法,这主要是因为当用户密度变大时, SUs 间干扰增大, CIDG 采取的是一种利己的最佳响应策略,易陷入局部最优解; SLG 算法为潜在博弈,兼顾个体与全局收益,故能在少量信息观察前提下保持良好的网络满意度。

当固定 $N = 80, L = 6$ 时,不同信道数量下几种算法网络满意度对比情况如图 9 所示,从中可以看出当信道资源紧张时,基于 SLG 算法的网络满意度最高,当信道资源宽松时,SLG 算法网络满意度略低于 CIDG,其原因与前述基本一致。

设定 $N = 80, M = 3, L = 6$ 时,不同速率阈值变化下几种算法网络满意度对比情况如图 10 所示. CIDG 和 Random Selection 由于其算法特性,用户收益趋向平均主义,当阈值较低时能获得较好的网络满意度. 然而阈值增大时, SLG 算法则能充分发挥弹性目标函数及潜在博弈的优越性,网络满意度能高于另外几种算法,进一步说明了其有效性。

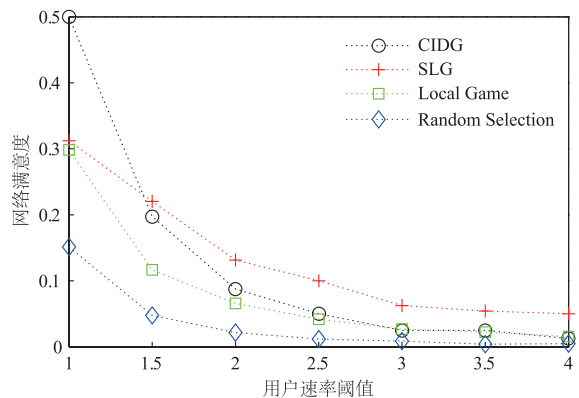


图 10 不同速率阈值网络满意度对比

5.4 PUs 存在时算法性能评估

本节将验证 PUs 与 SUs 同时存在于 CRN 时,几种频谱资源分配算法性能对比. 在网络中部署 4 个 PUs,

分别位于(30,30) (30,70) (70,30)和(75,30)处,干扰半径均为30m,以概率 θ 分别对信道1、2、3和1独占使用,则信道空闲概率为 $\mu = 1 - \theta$.

验证 PUs 存在时算法的收敛性. 设定 $N = 20, M = 3, L = 6$, 图 11 为随机选取某 SU 的信道选择概率的演化过程, 从图中可知经 150 次迭代后, 该用户以概率 1 收敛到某一固定信道, 其功率策略亦是如此, 最终三条信道 SUs 数目分别为 5、7 和 8, 且保持不变, 证明了 SLG 算法在 PUs 存下的前提下最终也能收敛到纯策略纳什均衡.

验证 PUs 存在时算法的有效性. 为了避免拓扑不均匀, 加大 SUs 的部署数量. 图 12 为 $N = 80, L = 6, M = 3$ 时, 不同算法平均传输速率对比. 从图中可知, CIDG 算法有全局信息的优势, 与 PUs 是零冲突, 信道收益最高, 但选择信道始终随 PUs 不断变化, 信道切换成本高. SLG 算法借助少量交互, 博弈过程中不断减少 PUs 使用相对频繁信道的选择概率, 尽可能避免冲突, 仍然能够获接近 CIDG 算法的传输速率, 验证了其在 PUs 存在时的有效性.

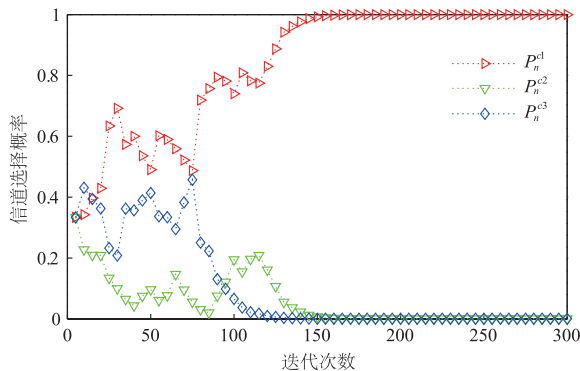


图 11 PUs 存在时信道选择概率演化过程

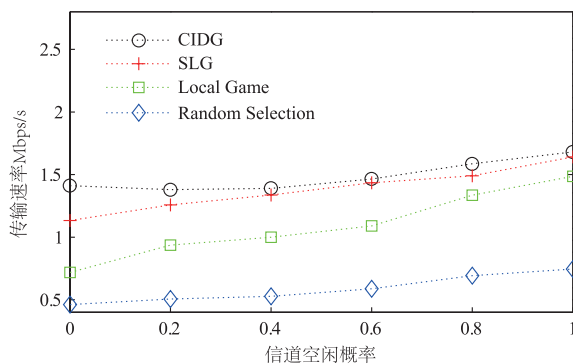


图 12 不同信道空闲概率传输速率对比

图 13 为不同信道空闲概率时网络满意度对比情况. 当 PUs 存在时, SLG 算法能够充分发挥弹性目标函数及潜在博弈的优势, SUs 逐渐减少所在区域 PUs 所使用信道的概率, 获得最好网络满意度. CIDG 算法作为

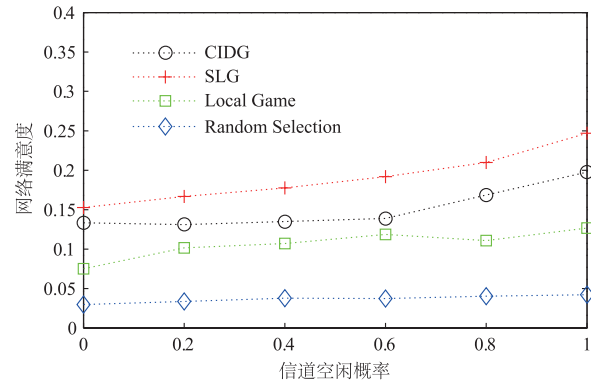


图 13 不同信道空闲概率网络满意度对比

一种理想态, 虽然能有效避免与 PUs 产生冲突, 却加重了 SUs 间的竞争, 导致较差的网络满意度. Local Game 算法因参考信息较少, 且非潜在博弈, 性能不及 SLG 与 CIDG. 随机算法对所有策略机会均等, SUs 与 PUs 及 SUs 间存在较高的冲突, 故网络满意度表现最差.

6 结语

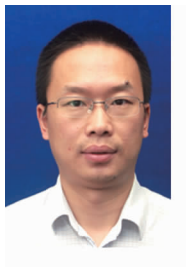
本文研究了 CRN 动态信道分配及功率控制问题. 根据次用户策略收益定义弹性满意度目标函数, 构建联合频谱分配与功率控制的分布式非合作博弈模型, 形式化证明了该博弈为潜在博弈, 从理论上分析了其纳什均衡点的存在性及最终收敛解的有效性. 为了降低交互成本和算法求解复杂度, 将纯策略博弈扩展为混合策略博弈, 提出了仅需与当前信道相邻用户交互的随机学习策略选择算法, 并给出了随机学习算法收敛到纯策略纳什均衡的严格证明. 仿真结果表明网络中无论 PUs 是否存在, 该算法能够在少量信息交互前提下求解纯策略纳什均衡解, 获得较高的传输速率, 并提升用户满意度, 尤其是在用户密度较大时, 其性能接近甚至超过完全信息动态博弈.

参考文献

- [1] Lai J, Dutkiewicz E, Liu R P, et al. Opportunistic spectrum access with two channel sensing in cognitive radio networks [J]. IEEE Transactions on Mobile Computing, 2015, 14 (1): 126 - 138.
- [2] Audhya G K, Sinha K, Ghosh S C, et al. A survey on the channel assignment problem in wireless networks [J]. Wireless Communications & Mobile Computing, 2011, 11 (5): 583 - 609.
- [3] Chaudhry A U, Hafez R H M, Chinneck J W. On the impact of interference models on channel assignment in multi-radio multi-channel wireless mesh networks [J]. Ad Hoc Networks, 2015, 27 (C): 68 - 80.
- [4] 贾杰, 李燕燕, 等. 认知无线网状网中基于差分演化的功

- 率控制与信道分配[J]. 电子学报, 2013, 41(1): 62–67.
- Jia J, Li Y, et al. Channel allocation and power control based on differential evolution algorithm in cognitive radio mesh network [J]. Acta Electronica Sinica, 2013, 41(1): 62–67. (in Chinese)
- [5] Wang X, Huang L, Leng B, et al. Joint channel and sink assignment for data collection in cognitive wireless sensor networks[J]. International Journal of Communication Systems, 2017, 30(5): 1–15.
- [6] Jia J, Wang X, Chen J. A genetic approach on cross-layer optimization for cognitive radio wireless mesh network under SINR model[J]. Ad Hoc Networks, 2015, 27(C): 57–67.
- [7] Haddad M, Hayel Y, Habachi O. Spectrum coordination in energy efficient cognitive radio networks[J]. IEEE Transactions on Vehicular Technology, 2012, 64(5): 2112–2122.
- [8] Duarte P B F, Fadlullah Z M, Vasilakos A V, et al. On the partially overlapped channel assignment on wireless mesh network backbone: a game theoretic approach [J]. IEEE Journal on Selected Areas in Communications, 2012, 30(1): 119–127.
- [9] Zheng J, Cai Y, Xu Y, et al. Distributed channel selection for interference mitigation in dynamic environment: a game-theoretic stochastic learning solution[J]. IEEE Transactions on Vehicular Technology, 2014, 63(9): 4757–4762.
- [10] Gao Z, Chen J, Xu Y. Opportunistic spectrum access with discrete feedback in unknown and dynamic environment [J]. KSII Transactions on Internet & Information Systems, 2015, 9(10): 3867–3886.
- [11] Tseng L C, Chien F T, Chang R Y, et al. Distributed channel assignment for network MIMO: game-theoretic formulation and stochastic learning [J]. Wireless Networks, 2015, 21(4): 1–16.
- [12] Wu Q, Wu D, Xu Y, Wang J. Demand-aware multichannel opportunistic spectrum access: a local interaction game approach with reduced information exchange [J]. IEEE Transactions on Vehicular Technology, 2015, 64(10): 4899–4904.
- [13] 徐宁, 张沪寅, 等. 认知 Ad Hoc 网络中基于信道相似度的分簇算法研究 [J]. 电子学报, 2016, 44(10): 2323–2329.
- Xu N, Zhang H, et al. Channel similarity based clustering algorithm in cognitive Ad Hoc network [J]. Acta Electronica Sinica, 2016, 44(10): 2323–2329. (in Chinese)
- [14] Babadi B, Tarokh V. GADIA: A greedy asynchronous distributed interference avoidance algorithm [J]. Transactions on Information Theory IEEE, 2011, 56(12): 6228–6252.
- [15] Hu S, Yao Y D, Yang Z. MAC protocol identification using support vector machines for cognitive radio networks [J]. IEEE Wireless Communications, 2014, 21(1): 52–60.
- [16] Noroozoliaee M J, Hamdaoui B, Tumer K. Efficient objective functions for coordinated learning in large-scale distributed OSA systems [J]. IEEE Transactions on Mobile Computing, 2013, 12(5): 931–944.
- [17] Sastry P S, Phansalkar V V, Thathachar M. Decentralized learning of Nash equilibria in multi-person stochastic games with incomplete information [J]. IEEE Transactions on Systems Man & Cybernetics, 1994, 24(5): 769–777.

作者简介



汪志勇 男, 1978 年出生于湖北武汉, 博士研究生, 主要研究方向为认知无线网络、上下文认知计算。
E-mail: zywang_whu@whu.edu.cn



张沪寅 (通信作者) 男, 1962 年生于江苏苏州, 教授、博士、博士生导师, 主要研究方向为网络 QoS、计算机网络、新一代互联网体系结构。
Email: zhy2536@whu.edu.cn