

文章编号: 1000 5013(2010)05-0521-05

一种新型认知无线电信道状态的预测算法

黄川^{1,2}, 郑宝玉¹

(1. 南京邮电大学 信号处理与传输研究院 江苏 南京 210003;

2. 福建师范大学 数学与计算机科学学院 福建 福州 350007)

摘要: 基于部分可测马尔科夫决策过程(POMDP)模型,结合认知无线电频谱侦测技术,提出一种新的多无线电多信道环境下认知无线电检测信道算法.该算法通过对信道状态历史信息分析,推导出信道信念状态的初始分布和转移概率;然后,以此选择出具有最佳回报的信道以供接入,使得次用户能获得最佳带宽回报,从而达到提高信道利用率的目的.仿真结果表明,算法获得相对于传统认知无线电频谱接入方式更高的信道带宽,并接近无漏检和虚警现象的理想情况,有效地提高了信道利用率.

关键词: 认知无线电;多无线电;多信道;马尔科夫模型;频谱侦测

中图分类号: TN 014; TN 914.4

文献标识码: A

新兴的认知无线技术为解决非授权用户有效利用闲置频谱,提高频谱利用率提供了可能^[1].然而,由于网络的未知性,使得次用户很难获得不同信道的精确状态,而必须依据当前的不完全状态信息做出决策.因此,对网络的未知性采用部分可测马尔科夫模型(POMDP)建立信道状态信息模型,并以此做出选择最佳信道的决策具有合理性^[2-3].文[2]提出了一种基于POMDP模型的认知MAC(Media Access Control)协议,解决了异构网络中信道选择决策问题.文[3]在文[2]的基础上,提出更为具体的以POMDP模型为基础的频谱检测策略、信道选择决策,以及接入策略的联合优化方案.上述研究仍存在着两个问题:其一是信道状态间的转移概率是事先给定的,无法根据实际情况变化;另一个是只有一个无线电收发器,使得用户不得不经常中断数据传输过程转而检测授权用户(或称为主用户)是否出现,从而浪费信道带宽.针对上述问题,本文通过建立以POMDP模型为基础的多无线电系统^[4]模型,提出一种与认知无线电频谱侦测技术相结合的在异构网络环境下的新型信道状态预测算法.

1 异构网络环境及多无线电系统模型

假设网络中每个次用户配备两个独立的,具有认知无线电功能的无线电收发设备.当次用户从网络A向不同的网络B移动时,其中一个无线电保持与网络A的连接,称为工作无线电;而另一个无线电处于认知侦测状态,则称为观测无线电.由观测无线电侦测结果组成的信道状态历史信息,由离散观测时间序列 D_T 内的一系列行为 a 、回报 r 、观测状态 z 和应答状态 k 的序列 h 组成.即

$$h = \{(a_1, r_1, z_1, k_1), \dots, (a_t, r_t, z_t, k_t)\}, \quad t \in D_T. \quad (1)$$

假设在网络B中存在 $N(0 < N < \infty)$ 个独立授权信道,每个信道在某一个时隙 t 内处于空闲(1)或忙(0)状态,则网络在某个时隙 t 内的信道状态为

$$S(t) = [S_1(t), \dots, S_M(t)].$$

其中: $M = 2^N$,且 $S_i(t) = s_1(t) \dots s_N(t)$, $s_j(t) \in \{0, 1\}$,每个信道的带宽表示为 $W_{B,j}$, $j = 1, \dots, N$.由于环境的未知性,设次用户可侦测到的信道数目 $n \leq N$.

收稿日期: 2009-11-29

通信作者: 郑宝玉(1945-),男,教授,博士生导师,主要从事无线通信与网络信号处理的研究. E-mail: zby@njupt.edu.cn.

基金项目: 国家自然科学基金资助项目(60972039)

2 典型的 POMDP 模型

一个典型 POMDP 模型可用六元组表示为 $\langle S, A, T, R, Z, O \rangle^{[5]}$. 其中: S 为系统中有限信道状态集合; A 为次用户采取的有限行为(观测, 接入)的集合, 用 $A = \{a_1, a_2\}$ 表示; T 表示当前信道状态 s 在行为 a 的作用下变为 s' 的转移函数, 记为 $T(s, a, s')$; R 为瞬时回报函数, 记为 $R(s, a)$; Z 为用户对系统状态的有限观测状态集合; O 为观测函数, 记为 $O(s', a, z)$. 此外, $k_i(t) \in \{0, 1\}$ 表示次用户在执行行为 a 后得到的应答. 此处, 设应答是无错的.

由于 S 是未知的, 采用信念状态空间 B 来表示信道状态的概率分布, 有

$$B = \{b: \sum_{s \in S} b(s) = 1, 0, \forall s \in S, b(s) \geq 0\}.$$

其中: $b(s)$ 表示信道处于状态 s 的概率. 根据 Bayes 法则, 可得在 $t+1$ 时隙信念状态更新的表达式为

$$b_{t+1}(s') \propto b_t(a_t, z_{t+1})(s') = \frac{\sum_{s \in S} T(s, a_t, s') O(s', a_t, z_{t+1}) b_t(s)}{\sum_{s \in S} O(s', a_t, z_{t+1}) \sum_{s \in S} T(s, a_t, s') b_t(s)}. \quad (2)$$

对于任何时隙 t , 信念状态是一个充分统计量^[6], 从而表明不会有任何历史信息可以增加次用户的回报. 同时, 定义次用户得到的瞬时回报为空闲信道的带宽为

$$R_B(b, a) = \sum_{s \in S} b(s) R(s, a) = \sum_{s \in S} \sum_{i=0}^n b(s=1) M_{B,i}. \quad (3)$$

策略 π 表示次用户在当前信念状态 b 的情况下, 使用策略 π 从而选择行为 a . 解决一个 POMDP 问题的目标是找寻一个最佳策略 π^* , 以使用户获得的总折扣回报达到最大值. 即次用户所能获得的最大信道带宽, 用值函数表示为

$$V^*(b) = \max_{a \in A} \left[R_B(b, a) + \zeta \sum_{z \in Z} P(z | b, a) V^*(\tau(b, a, z)) \right]. \quad (4)$$

式(4)中: ζ 为折扣因子, $0 < \zeta \leq 1$; P 为条件转移的概率函数.

3 信道预测算法及收敛性证明

以 POMDP 模型为基础的多无线电多信道, 其信道状态预测算法(CSPA)可分为如下两个阶段.

(1) 观测阶段. 通过一段时间的观测, 将获得的系统环境信息记录到 h 中, 期间的执行接入行为仅为向相应频段发送探测包, 并未真正执行信道接入操作. (2) 预测阶段. 通过 h , 对信道的初始状态分布、状态转移概率和观测概率进行估计, 并利用启发式算法找出具有最大折扣回报的策略 π^* , 在接入时隙次用户按其接入. 在预测阶段, 对次用户来说, 信道状态个数 M 是 N 的指数, 要计算出最大折扣回报是很困难的. 然而, 实际网络中的信道一般是独立的, 有如下定理.

定理 1 假设 n 个独立信道, 有 $\Lambda = [\lambda_1, \dots, \lambda_n]$, 其中 λ_i 为信道 i 在某个时隙 t 开始时刻所处的状态, 则 L 是信道状态 $S_i(t)$ 的充分统计量.

证明 参见文[6].

根据定理 1, POMDP 模型中信念状态空间 $B = \{b(S_i(t)), i = 1, \dots, n\}$ 可简化为 $B = \{b(s_k(t)), k = 1, \dots, n\}$, 从而信念空间维度由 2^n 降为 n . 对于每个信道 i 来说, 其最大折扣回报表达式为

$$V_i^*(b_i) = \max_{a_i \in A} \left[R_B(b_i, a_i) + \zeta \sum_{z_i \in Z} P(z_i | b_i, a_i) V_i^*(\tau(b_i, a_i, z_i)) \right]. \quad (5)$$

式(5)中: $R_B(b_i, a_i) = b(s_i=1) W_{B,i}$. 对于次用户, 则有

$$V^*(b) = \max_{i=1, \dots, n} V_i^*(b_i). \quad (6)$$

对于在时隙 t 内的每一个信道, 可设 θ_i 为在时隙 t 内信道 i 状态为空闲的概率, 即 $b(s_i(t)=1) = \theta_i$; 而状态为忙的概率为 $b(s_i(t)=0) = 1 - \theta_i$. 在 T 个时隙中, 信道 i 状态为空闲出现次数为 y_i , 服从二项式分布 $b(y_i; T, \theta_i)$. 根据 Bayes 法则, θ_i 对 h 的条件概率密度为

$$f(\theta_i | h) = \frac{q(\theta_i) \binom{T}{y_i} \theta_i^{y_i} (1 - \theta_i)^{T - y_i}}{\int_0^1 q(\theta_i) \binom{T}{y_i} \theta_i^{y_i} (1 - \theta_i)^{T - y_i} d\theta_i}. \quad (7)$$

式(7)中: $q(\theta_i)$ 为 θ_i 的先验分布. 对于次用户来说, 信道 i 处于空闲或忙状态是等可能的, 故先验分布 $q(\theta)$ 为 $[0, 1]$ 上的均匀分布, 有

$$f(\theta_i | h) = \frac{\theta_i^{y_i} (1 - \theta_i)^{T - y_i}}{\int_0^1 \theta_i^{y_i} (1 - \theta_i)^{T - y_i} d\theta_i}. \quad (8)$$

用 θ_i 对 h 的条件期望 $E\{\theta_i | h\}$ 估计信道状态为空闲的概率, 有

$$b_i(s_i = 1) = E\{\theta_i | h\} = (y_i + 1) / (T + 2). \quad (9)$$

信道状态转移概率 $T(s, a, s')$ 也是未知的. 设 $p_{s, s'}^a$ 为信道 i 执行行为 a 后, 状态从 s 转移到 s' 的转移概率. 其中: $s, s' \in \{0, 1\}$, 向量 $P_i = (p_{s, s'}^a, a \in A, k = 1, \dots, |S|)$. 在 T 个时隙中信道 i 的状态从 s 到 s' 的转移次数向量 $\phi_i = (\phi_{s, s'}^a, a \in A, k = 1, \dots, |S|)$. 在执行行为 a 条件下, P_i 服从 Dirichlet 分布, $(p_{s, s'}^a, \dots, p_{s, s_1}^a)$ 渐进等于 $\text{Dir}(\phi_{s, s_1}^a, \dots, \phi_{s, s_1}^a)$, 其期望 $E(p_i) = \phi_{s, s'}^a / \sum_{k=1}^{|S|} \phi_{s, s'}^a$.

当信道状态转移后, 向量 $\phi_i = \phi_i + \delta_{s, s'}^a$. 其中: $|\delta_{s, s'}^a| = |S|$, 且 $\delta_{s, s'}^a[s' = j] = 1$, 其余为 0. 用期望值估计转移概率, 有

$$T(s, a, s') = \phi_{s, s'}^a / \sum_{s' \in S} \phi_{s, s'}^a. \quad (10)$$

在频谱侦测中, 由于存在漏检和虚警现象, 因此对于次用户来说, 所观测到的信道状态并不一定与信道真实状态相符. 假设信道为 AWGN, p_d 为检测概率, p_f 为虚警概率, 且采用能量检测器的频谱侦测方法^[2], 则有

$$p_f = P\{Y > \tau | H_0\} = 1 - \Gamma\left(\frac{L}{2}, \frac{\tau}{2\sigma_0^2}\right), \quad (11)$$

$$p_d = P\{Y > \tau | H_1\} = 1 - \Gamma\left(\frac{L}{2}, \frac{\tau}{2\sigma_1^2}\right). \quad (12)$$

由于信道真实状态的未知性, 次用户可根据执行行为 a 后得到的应答信息 k , 来验证观测状态的正确与否, 有

$$O(s', a, z) = P(z | s', a) = P(z | k, a) = \begin{cases} p_d, & k = z = 0, \\ 1 - p_f, & k = z = 1, \\ p_f, & k = 1, z = 0, \\ 1 - p_d, & k = 0, z = 1, \end{cases} \quad (13)$$

从而可求出次用户可获得的最大折扣回报. 即

$$\begin{aligned} V^*(b) = \max_{i=1, \dots, n} \max_{a_i \in A} [& b(s_i = 1) W_{B, i} + \zeta [p_d p_{0,0}^a b(s_i = 0) + \\ & p_d p_{1,0}^a b(s_i = 1) + p_f p_{0,1}^a b(s_i = 0) + p_f p_{1,1}^a b(s_i = 1)] V_i^*(\tau(b_i, a_i, 0)) + \\ & \zeta [(1 - p_d) p_{0,0}^a b(s_i = 0) + (1 - p_d) p_{1,0}^a b(s_i = 1) + \\ & (1 - p_f) p_{0,1}^a b(s_i = 0) + (1 - p_f) p_{1,1}^a b(s_i = 1)] V_i^*(\tau(b_i, a_i, 1))]. \end{aligned} \quad (14)$$

(1) 当 $k = z = 0$ 时, 信念状态更新为

$$\tau(b, a, z) = \frac{[p_{0,0}^a b(s = 0) + p_{1,0}^a b(s = 1)]}{p_d [p_{0,0}^a b(s = 0) + p_{1,0}^a b(s = 1)] + p_f [p_{0,1}^a b(s = 0) + p_{1,1}^a b(s = 1)]}.$$

(2) 当 $k = z = 1$ 时, 信念状态更新为

$$\tau(b, a, z) = \frac{[p_{0,1}^a b(s = 0) + p_{1,1}^a b(s = 1)]}{(1 - p_d) [p_{0,0}^a b(s = 0) + p_{1,0}^a b(s = 1)] + (1 - p_f) [p_{0,1}^a b(s = 0) + p_{1,1}^a b(s = 1)]}.$$

(3) 当 $k = 1, z = 0$ 时, 信念状态更新为

$$\tau(b, a, z) = \frac{p_f [p_{0,1}^a b(s = 0) + p_{1,1}^a b(s = 1)]}{p_d [p_{0,0}^a b(s = 0) + p_{1,0}^a b(s = 1)] + p_f [p_{0,1}^a b(s = 0) + p_{1,1}^a b(s = 1)]}.$$

(4) 当 $k = 0, z = 1$ 时, 信念状态更新为

$$\tau(b, a, z) = \frac{(1 - p_d) [p_{0,0}^a b(s = 0) + p_{1,0}^a b(s = 1)]}{(1 - p_d) [p_{0,0}^a b(s = 0) + p_{1,0}^a b(s = 1)] + (1 - p_f) [p_{0,1}^a b(s = 0) + p_{1,1}^a b(s = 1)]}.$$

4 仿真结果及分析

为了测试 CSPA 算法的性能,引入随机接入算法(Random Access Algorithm, RAA)^[2]与之比较.在 RAA 算法中,次用户在刚进入未知新网络时,不使用信道状态预测方法,而是通过每一接入时隙开始时刻的侦测来获知可以采用的若干信道,并随机选择其中一个接入.为了体现公平性,设新网络中每一信道的带宽均为 1 个单位,且每个时隙为 1 个单位时间.设折扣因子 $\zeta=1$.同时比较 CSPA 算法与理想情况下所获得的信道带宽回报,即与不存在漏检和虚警现象条件下的对比.

信噪比(R_{SN})和检测样本数目(L)的不同时,虚警率 P_f 值对检测率 P_d 的影响,如图 1 所示.从图 1 中可知,当 $R_{SN}=10, L=5$ 时,虚警率 P_f 的变化对检测率 P_d 的影响最小,且 P_d 值在 0.9~1.0 之间变动.因此,设 $P_d=0.95$.仿真中,算法对 10 000 个随机信道进行运算,然后取回报的平均值(单位每时隙).

CSPA 算法与 RAA 算法的对比,如图 2 所示.图 2 中:观测时隙 $T_o=30$,接入时隙 $T_A=30$,信道数目 $n=2$.从图 2 的平均回报值(δ)曲线可以看到,CSPA 算法在刚开始的时隙获得约 0.55 的平均回报值,在第 14 个接入时隙上升至 0.73,而后增长平稳,逐渐趋近于 0.74;RAA 算法每个时隙的平均回报值稳定在 0.5 左右,在理想情况下,其平均回报值稳定在 0.755 左右.

从图 2 的平均回报值百分比(φ)曲线可以看到,CSPA 算法能获得的平均回报值比 RAA 算法平均多出约 43%.表明次用户采用 CSPA 算法,在每个时隙都能够取得最佳的接入策略,因此获得的信道带宽平均回报优于采用传统认知无线电随机频谱接入方式.但是,由于频谱侦测中存在漏检和虚警现象,使得算法与无漏检和虚警现象存在的理想值有一定的偏差.

采用不同的观测时隙值来比较 CSPA 算法和 RAA 算法所获得的平均回报值,结果如图 3 所示.从图 3 中可以看到,观测时隙分别为 30, 100 时,CSPA 算法所获得的平均回报值略有差异.如果在观测阶段观测越充分(观测时隙越长),CSPA 算法所获得的初始信念状态和状态转移概率越准确,其获得的回报越精确.然而,观测时隙具体的取值应根据实际情况而定,这也是下一步研究工作的重点.从图中 3 可以看到,RAA 算法由于没有观测时隙,故其在不同观测时隙条件下,其所获得的平均回报值变化不大,基本稳定在 0.5 左右.

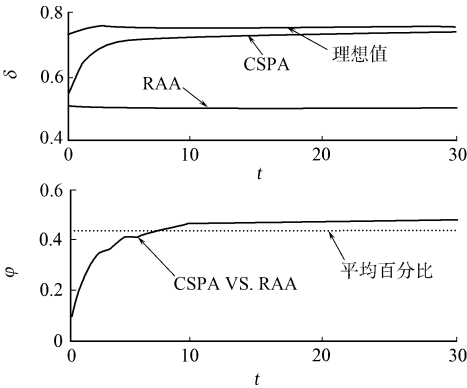


图 2 CSPA 算法与 RAA 算法的对比
Fig. 2 Comparison of CSPA and RAA

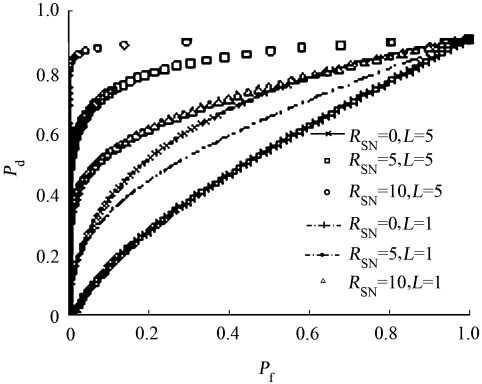


图 1 虚警率和检测率的关系
Fig. 1 Relationship of probability
of false alarm and detection

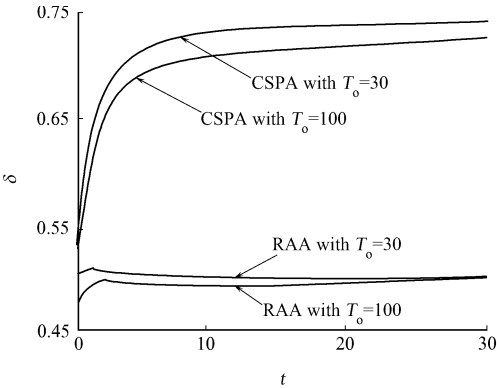


图 3 不同观测时隙值对算法的影响
Fig. 3 Effect of different values of
observation time slots on the algorithm

信道数目对算法的影响,如图 4 所示.从图 4 中可以看出,当可被侦测到的信道数目由 2 增加到 12 时,CSPA 算法获得的平均回报值有显著的提升. $n=2$ 时的平均回报值趋近于 0.74,而当 $n=12$ 时,平均回报值趋近于 0.98.这是因为可用的信道数目越多,CSPA 算法在每个时隙可能获得的最大回报机

会越大, 从而获得的平均回报越多. 由于 RAA 算法信道选择的随机性, 信道数目的增多对其影响不大, 基本维持在 0.5 左右. 如果不考虑漏检和虚警现象对系统的影响, 当信道数目足够大时, 其获得的平均回报值能达到 1. 由此说明, CSPA 算法更适用于多信道环境.

5 结束语

文中基于 POMDP 模型, 提出了一种在多无线电多信道环境下带有认知无线电频谱侦测功能的信道状态预测算法(CSPA), 以实现用户在多信道切换时能得到最佳信道带宽回报. 仿真结果表明, CSPA 算法获得相对于传统认知无线电频谱接入方式更高的信道带宽, 并接近无漏检和虚警现象的理想情况, 从而有效地提高了信道利用率.

参考文献:

[1] AKYILDIZ I, LEE W Y, VURAN M C, et al. A survey on spectrum management in cognitive radio networks[J]. IEEE Communications Magazine, 2008, 46(4): 40-48.

[2] ZHAO Q, TONG L, SWAMI A, et al. Decentralized cognitive MAC for opportunistic spectrum access in ad hoc networks: A POMDP framework[J]. IEEE Journal on Selected Areas in Communications, 2007, 25(3): 589-600

[3] CHEN Y X, ZHAO Q, SWAMI A. Joint design and separation principle for opportunistic spectrum access in the presence of sensing errors[J]. IEEE Trans on Information Theory, 2008, 54(5): 2053-2071.

[4] PIAO G, DAVID K. Multi-standard radio resource management for integrated voice and data services[C] // IEEE 65th Vehicular Technology Conference Dublin: IEEE, 2007: 990-995.

[5] KAELBLING L P, LITTMAN M L, CASSANDRA A R. Planning and acting in partially observable stochastic domains[J]. Artificial Intelligence, 1998, 101(1): 99-134.

[6] SMALLWOOD R D, SONDIK E J. The optimal control of partially observable Markov processes over a finite horizon[J]. Operations Research, 1973, 21(5): 1071-1088.

A Novel Channel State Prediction Algorithm of Cognitive Radio

HUANG Chuan^{1,2}, ZHENG Bao-yu¹

(1. Institute of Signal Processing and Transmission, Nanjing University of Posts and Telecommunications, Nanjing 210003, China;
2. School of Mathematics and Computer Science, Fujian Normal University, Fuzhou 350007, China)

Abstract: Based on the theory of partially observable Markov decision process (POMDP) model, a novel cognitive radio channel sensing algorithm integrated with spectrum sensing technique for cognitive radio under multi-radio multi-channel environment. By the analysis of the channel state historical information, the initial distribution of the belief state and transition probability is derived and the channel with optimal reward is selected for unlicensed user to improve the spectrum utilization. The simulation results demonstrate that the proposed algorithm has better performance than classical algorithms.

Keywords: cognitive radio; multi-radio; multi-channel; Markov model; spectrum sensing

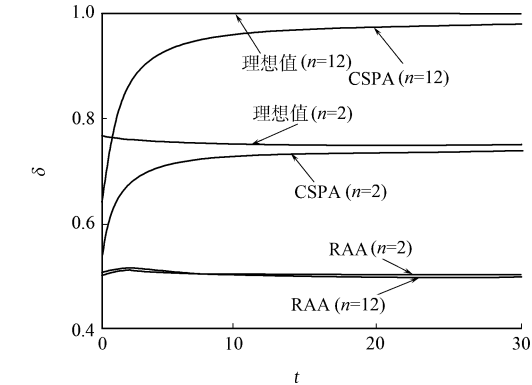


图 4 信道数目对算法的影响
Fig. 4 Effect of different numbers
of channel on the algorithm

(责任编辑: 黄仲一 英文审校: 吴逢铁)