

基于 POMDP 的认知无线电自适应频谱感知算法

许瑞琛, 蒋挺

(北京邮电大学 信息与通信工程学院, 北京 100876)

摘要: 针对如何建立适合毫秒级频谱空洞场景的频谱感知机制这一问题, 推导出最佳次级用户单次数据传输时间长度, 解决了数据传输量最大化的问题; 其次提出一种基于部分可观测马尔科夫决策过程 (POMDP, partially observable Markov decision process) 的自适应频谱感知算法, 解决了快速接入毫秒级频谱空洞的问题。该算法根据上一决策时段的信念向量和感知结果自适应确定在当前决策时段内进行频谱感知或数据传输行为。仿真结果表明, 该自适应频谱感知算法能够有效控制次级用户和主用户的冲突概率并增加频谱利用率。

关键词: 部分可观测马尔科夫决策过程; 自适应感知; 数据传输时间; 冲突概率; 频谱利用率

中图分类号: TN911.22

文献标识码: A

文章编号: 1000-436X(2013)06-0049-08

Cognitive radio auto-adaptive sensing algorithm based on POMDP

XU Rui-chen, JIANG Ting

(School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China)

Abstract: In order to design an appropriate spectrum sensing mechanism in millisecond spectrum hole environment, the optimal data transmission time of secondary users was derived to maximize the data throughput. Furthermore, in order to exploit the millisecond spectrum holes, an auto-adaptive spectrum sensing algorithm based on POMDP (partially observable Markov decision process) was proposed. This algorithm executes spectrum sensing or data transmission action in the current decision epoch based on the belief vector and the observation result in the last decision epoch. Simulation results show that the proposed algorithm can constrain the collision probability between primary users and secondary users, and improve systemic spectrum utilization at the same time.

Key words: POMDP; auto-adaptive spectrum sensing; data transmission time; collision probability; spectrum utilization

1 引言

在不影响主用户 (PU, primary user) 使用信道的前提下, 如何建立能够为次级用户 (SU, secondary user) 提供高频谱利用率的频谱感知及数据传输方法已经成为认知无线电 (CR, cognitive radio) 系统^[1] 目前亟待解决的问题之一。

CR 中 SU 需要动态接入频谱空洞实现频谱利用率最大化。目前已有多种解决该问题的频谱感知优化算法^[2,3], 这些算法大多研究如何在最短的感知时间内有效检测 PU 信号以达到最大化频谱利用率的目的。根据实测数据^[4], CR 系统中的频谱空洞时

间长度大多不同, 并且大多长度集中在毫秒级别。而传统 CR 周期性检测机制规定了频谱感知周期, 所以动态变化的 CR 环境往往会导致 SU 和 PU 数据传输的冲突概率上升。例如, 梁宜昌教授提出一种吞吐量最大化数据传输机制^[2], 该机制规定检测周期长为 100 ms, 并且根据不同的 PU 占用信道概率动态选择最优频谱感知时间, 达到最大化频谱利用率的目的。但是该算法并没有考虑 PU 接入信道的频率和单次使用信道的的时间因素。在频谱空洞处于毫秒级别的情况下, 固定的帧长度会导致 SU 和 PU 的冲突概率上升。

部分可观测马尔科夫决策过程 (POMDP, partially

收稿日期: 2012-10-20; 修回日期: 2013-01-04

基金项目: 国家科技重大专项基金资助项目 (2010ZX03006-006); 国家自然科学基金资助项目 (61171176)

Foundation Items: The National Science and Technology Major Project (2010ZX03006-006); The National Natural Science Foundation of China (61171176)

observable Markov decision process) 能够自适应学习 CR 快速变化的外部环境, 并在每个决策时段选择最优的行为。已有多篇文献^[5-7]将 POMDP 应用于 CR 的感知算法当中。然而目前相关的感知算法绝大多数集中在如何选择最佳信道进行感知以缩短频谱感知时间方面, 并且这些算法均使用周期性频谱检测方法。这些算法没有解决如何使 SU 有效接入毫秒级别的频谱空洞的问题。此外, POMDP 模型算法复杂度很大^[2]。在动态变化的 CR 环境中, 算法复杂度过高会造成决策延迟, 这也会成为相关算法在实际应用中的一个瓶颈。

基于以上分析, 本文首先根据 SU 的数据传输需求和 PU 使用信道的特性推导出最佳数据传输时间的解析解。其次, 根据 POMDP 模型设计了自适应非周期频谱感知和数据传输机制。最后, 通过一种简单的单门限决策方式减小所提出算法的复杂度。

2 系统模型

系统由 1 个 SU 和 1 个授权信道组成。系统允许 SU 在授权信道空闲的情况下接入。SU 负责感知信道和数据传输。SU 需要发送 N 个数据分组。在每一个决策时段末尾, SU 基于感知结果做出最佳行为判断: 当不确定 PU 占用信道的情况下, SU 在下一时段进行频谱感知; 当确定 PU 不占用信道的情况下, SU 在下一时段进行数据传输。系统运行时间可以分为多个决策时间段 $\{t_1, t_2, \dots\}$ 。假设 PU 的到达符合均值为 λ 的泊松分布。PU 单次占用信道的的时间符合均值为 $\frac{1}{\mu}$ 的负指数分布。并且 SU 有 λ 和 μ 的先验信息。

如图 1 所示, SU 在每个决策时段结束时通过决策过程选择下一个决策时段的行为。

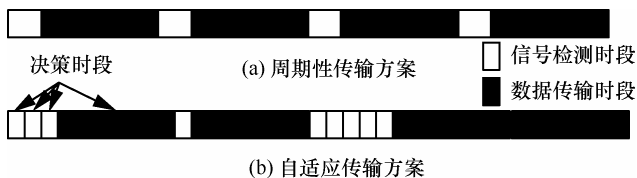


图 1 周期传输方案和自适应传输方案

SU 所选用的行为包括数据传输和频谱感知 2 部分, 现分别对 2 种行为进行说明。

1) 频谱感知

使用 ON-OFF 模型^[5]描述信道的使用情况, H_0

表示 PU 信号不存在, H_1 表示 PU 信号存在。 H_0 和 H_1 情况下, SU 接收的信号可以表示为

$$\begin{cases} H_0 : y(t) = n(t) \\ H_1 : y(t) = hs(t) + n(t) \end{cases} \quad (1)$$

其中, $y(t)$ 表示 SU 接收到的信号, $s(t)$ 为 PU 信号, $n(t)$ 为 AWGN。

当 SU 不能确定信道的真实状态 (H_0, H_1) 时, 选择在频谱感知时间 T_S 内进行能量检测。

如果在第 t 个决策时段进行能量感知, SU 将得到检测结果 E_t ^[8]。

$$E_t = \frac{1}{N_0} \sum_{i=1}^{N_0} |\xi_i|^2 \quad (2)$$

其中, $N_0 = F_S T_S$, F_S 为抽样频率, ξ_i 为 T_S 中信号的第 i 个抽样点。根据文献[8], $n(t)$ 服从高斯分布, 均值和方差分别为 0 和 σ_n^2 。 $s(t)$ 也符合高斯分布, 均值和方差分别为 0 和 σ_s^2 。 H_0 状态下, E_t 服从自由度为 N_0 的卡方分布。当 N_0 足够大时, 根据中心极限定理, E_t 服从高斯分布。在 H_0 和 H_1 状态下, 记 E_t 的均值和方差分别为 μ_0 、 μ_1 、 σ_0^2 、 σ_1^2 ^[8], 满足

$$\begin{cases} H_0 : \mu_0 = \sigma_n^2, \sigma_0^2 = \frac{E|n(t)|^4 - \sigma_n^4}{N_0} \\ H_1 : \mu_1 = (\gamma + 1)\sigma_n^2, \sigma_1^2 = \frac{2(\gamma + 1)^2 \sigma_n^4}{N_0} \end{cases} \quad (3)$$

其中, γ 为信噪比。

为了建立 POMDP 中的状态 (H_0 和 H_1) 和观测值的对应矩阵, 本文对能量检测结果 E_t 进行量化, 得到量化结果 $[\zeta_t] \in [1, 2, \dots, M]$ 。设置量化门限, $\gamma_1 < \gamma_2 < \dots < \gamma_{M+1}$ 。如果 $\gamma_m < E_t < \gamma_{m+1}$, 那么 $\zeta_t = m$ 。为了避免 $E_t < \gamma_1$ 和 $E_t > \gamma_{M+1}$ 的情况出现, 可以设 $\gamma_1 = 0, \gamma_{M+1} \gg \gamma_M$ 。

2) 数据传输

如果 SU 确定授权信道可用, 则开始在时间 T_D 内传输数据。

3 最佳数据传输时间长度的推导过程

由于自适应数据传输模型具备非周期性, 因此无法推导最佳数据传输时间 T_D 的解析解。为简便起见, 这里按照周期数据传输方式求解最佳数据传输时间。和传统的认知无线电固定周期长度不同, 该部分所确定的周期长度是动态自适应的。其中, 频谱感知时间长度是固定的, 并且能够满足 IEEE 802.22 标准^[9]所要求的虚警概率和漏检概率。在固

定频谱感知时间的基础上，根据 PU 使用信道的特性和 SU 的数据传输需求推导得到能够使数据传输量最大化的数据传输时间长度。

如图 2 所示，SU 数据传输周期时间为

$$T = K\tau + T_S \quad (4)$$

其中， T_S 为信道感知耗时， K 为单周期传输的数据分组数， τ 为传输单个数据分组耗时。

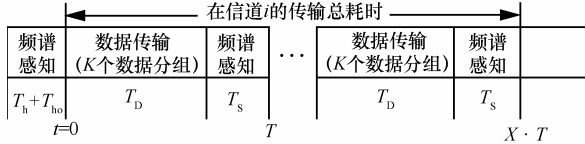


图2 数据传输时间计算模型

设 SU 在信道所传输的数据分组数为

$$N_{AT} = ZK \quad (5)$$

其中， Z 为 SU 在 PU 重新利用该信道前在当前信道连续进行数据传输的周期数， $Z=1,2,3,\dots,X$, s.t. $X \geq N/K$ 。最佳数据传输时间 $T_{D,best}$ 的求解问题就可以描述为

$$T_{D,best} = \left[\arg \max_K E(N_{AT}) \right] \tau \quad (6)$$

其中， $E(N_{AT})$ 为 N_{AT} 的期望。通过上述分析，求解 $T_{D,best}$ 的问题可以转换为求解最佳单次传输数据分组数 K_{best} 的问题。

由于 PU 到达间隔符合负指数分布，可以得到 Z 的分布函数

$$\begin{cases} P\{Z \geq 0\} = 1 \\ P\{Z \geq 1\} = (1 - P_f - P_m) e^{-\lambda T} \\ P\{Z \geq 2\} = (1 - P_f - P_m)^2 e^{-\lambda 2T} \\ \dots \\ P\{Z \geq X-1\} = (1 - P_f - P_m)^{X-1} e^{-\lambda(X-1)T} \\ P\{Z \geq X\} = P\{Z = X\} = (1 - P_f - P_m)^X e^{-\lambda X T} \end{cases} \quad (7)$$

其中， P_f 和 P_m 分别为所设定的虚警概率和漏警概率上限，令 $P_f + P_m = P$ ， $a = (1 - P) e^{-\lambda T}$ ，可得

$$\begin{cases} P\{Z = 0\} = P\{Z \geq 0\} - P\{Z \geq 1\} = 1 - a \\ P\{Z = 1\} = P\{Z \geq 1\} - P\{Z \geq 2\} = a - a^2 \\ P\{Z = 2\} = P\{Z \geq 2\} - P\{Z \geq 3\} = a^2 - a^3 \\ \dots \\ P\{Z = X-1\} = P\{Z \geq X-1\} - P\{Z \geq X\} \\ \quad = a^{X-1} - a^X \\ P\{Z = X\} = a^X \end{cases} \quad (8)$$

那么 N_{AT} 的期望为

$$\begin{aligned} E(N_{AT}) &= \sum_{n=0}^X nKP(Z = n) \\ &= K\{(a - a^2) + 2(a^2 - a^3) + 3(a^3 - a^4) + \dots + (X-1)(a^{X-1} - a^X) + Xa^X\} \\ &= K(a + a^2 + \dots + a^X) \\ &= Ka(1 - a^X)/(1 - a) \end{aligned} \quad (9)$$

化简 $E(N_{AT})$ ，求 $\frac{d[E(N_{AT})]}{d(K)}$ ，可得

$$\begin{aligned} \frac{d[E(N_{AT})]}{d(K)} &= \frac{-C^2 EK^2 - 2BCEK + BE - B^2E - CD}{(B + CK)^2} \end{aligned} \quad (10)$$

其中， $B = (P + \lambda T_S - P\lambda T_S)$ 、 $C = (1 - P)\lambda\tau$ 、 $D = NP(1 - E)$ 和 $E = 1 - e^{-\lambda\tau N}$ 。式 (10) 的推导过程见附录。

令 $\frac{d[E(N_{AT})]}{d(K)} = 0$ ，得

$$K_{best} = \frac{-BE + \sqrt{BE(BE - CD)}}{CE} \quad (11)$$

$$T_{D,best} = K_{best} \tau \quad (12)$$

图 3 描述了 $E(N_{AT})$ 和 T_D 之间的关系，并给出了由式 (12) 所得到的最佳单周期数据传输时间 $T_{D,best}$ 和使用 MATLAB 得到的最优解 $T_{MD,best}$ 。其中 $N = 1000$ 、 $P_f = 0.01$ 、 $P_m = 0.01$ 、 $\tau = 0.85$ ms、 $T_S = 1$ ms、 $\lambda = \{10, 20, 30, 40, 50, 60, 70, 90\}$ 。由图 3 可知，在不同的 λ 取值下， $T_{D,best}$ 和 $T_{MD,best}$ 接近，二者得到的 $E(N_{AT})$ 之差小于 2。

4 基于 POMDP 的决策过程

PU 的行为可以描述为一个两状态马尔科夫过程^[5-7]。其中，一个状态表示 PU 占用信道，另外一个状态表示 PU 没有占用信道。CR 系统并不知道 PU 的真实状态，只能从所得到的频谱感知结果推断 PU 的真实状态，所以使用隐马尔科夫模型 (HMM, hidden Markov model) 架构可以通过观测值预测隐藏的信道状态。但是 CR 系统环境随着时间动态变化，SU 需要根据环境变化情况自适应选择最佳行为。基于 POMDP 的 SU 对环境的自适应学习过程如图 4 所示。POMDP 模型^[5]可以自适应

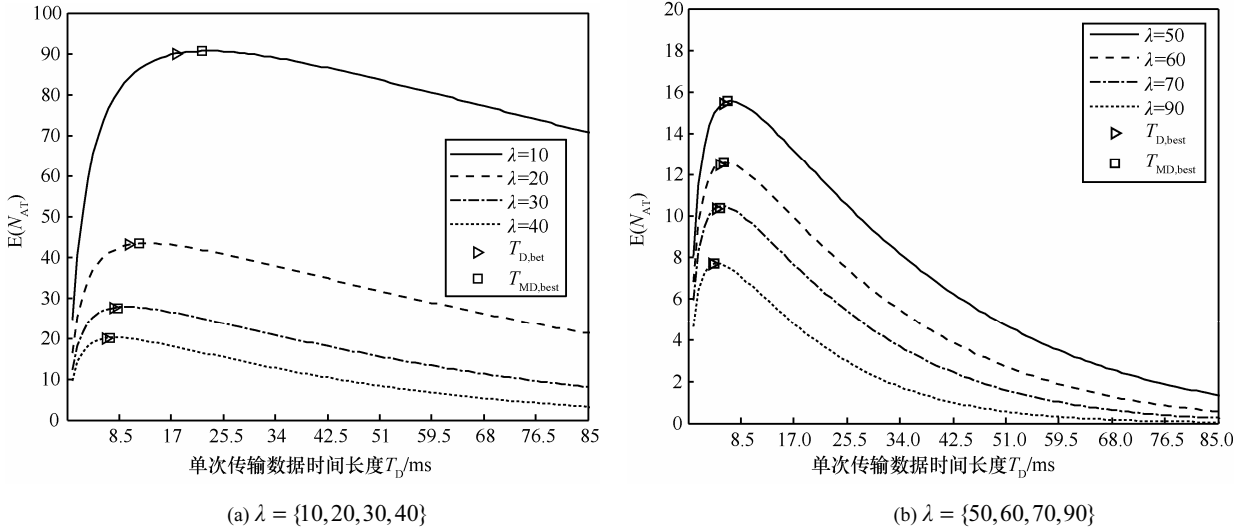


图 3 K_t 和 $E(N_{AT})$ 的关系

学习 CR 外部快速变化的环境, 并做出最佳决策(频谱感知或数据传输)。第二部分的 PU 到达符合泊松分布的假设也符合马尔科夫性, 所以这里使用 POMDP 模型描述认知无线电中信道的状态转移和 SU 的学习过程。POMDP 模型包括状态、行为、观测值和反馈 4 个部分, 可以用 $\{S, A, Y, R\}$ 表示, 具体参见图 4。下面将分别介绍自适应频谱感知算法中的状态、行为、状态转移矩阵、观测值矩阵、目标函数、反馈函数和决策过程, 最后提出一种低复杂度的单门限决策过程替代基于 POMDP 的动态规划算法。

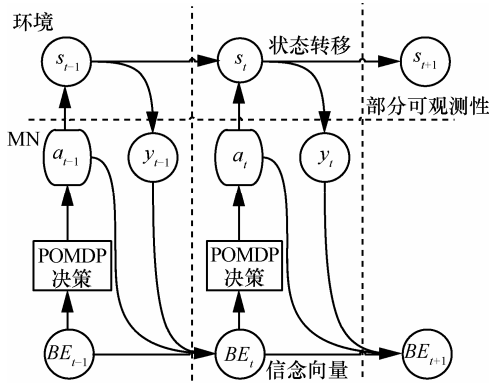


图 4 基于 POMDP 的 SU 和环境的自适应学习过程

4.1 状态、行为、状态转移矩阵和观测值矩阵

s_t 表示第 t 个决策时段的状态, $s_t = 0, 1$ 分别表示信道可用和不可用。 a_t 表示第 t 个决策时段所选用的行为。 $a_t = S, D$ 分别表示在第 t 个决策时段执行频谱感知和数据传输行为。

$p_{i,j}^a = P(s_{t+1} = j | s_t = i, a_t = a)$ 表示第 t 个决策

时段处于状态 i 并执行行为 a , 状态在第 $t+1$ 个决策时段转移到 j 的概率。这里假设 PU 的状态在 T_D 或 T_S 内只能变化一次。基于该假设, 4 种状态转移概率分别为 $p_{0,0}^a = e^{-\lambda T_a}$ 、 $p_{0,1}^a = 1 - e^{-\lambda T_a}$ 、 $p_{1,0}^a = 1 - e^{-\mu T_a}$ 、 $p_{1,1}^a = e^{-\mu T_a}$ 。

记 y_t 为第 t 个决策时段得到的观测值。如果 SU 选择能量感知行为, 根据式 (2) 得到感知结果 E_t , 通过量化得到和 $y_t = \zeta_t$ 。如果 SU 选择数据传输行为, 不会得到观测值, 所以 $y_t = \text{Null}$ 。记 $b_{i,y}^a = P\{y_t = y | s_{t+1} = i, a_t = a\}$ 为在状态 i 情况下, 采用行为 a 获得观测值 y 的概率。因为 $a_t = D$ 时不产生观测值, 所以 $b_{i,\text{Null}}^D = 1$ 。当 $a_t = S$ 时, $b_{i,y}^S = P\{y_y < E_t < \gamma_{y+1} | s_t = i, a_t = S\}$, $y \in [1, \dots, M]$ 。根据 PU 在不同状态下的 E_t 的概率计算 $b_{0,y}^S$ 和 $b_{1,y}^S$ 。

4.2 目标函数和反馈函数

在 POMDP 中, 目标函数是加权总反馈值的期望。

$$E \left\{ \sum_{t=1}^{\infty} \alpha^t R(s_t, a_t) \right\} \quad (13)$$

其中, α 为折扣因子, $0 < \alpha < 1$, 在认知无线电相关文献^[10,11]中 α 的取值大多在 0.95~0.999 之间, 这里设 $\alpha = 0.99$ 。 $R(s, a)$ 是在状态 s 执行行为 a 得到的反馈。按照最大化频谱利用率并且最小化 SU 和 PU 的冲突概率的原则设计 $R(s, a)$ 。如果在状态 $s_t = 0$ 时传输数据, 可以提高频谱利用率并且没有造成和 PU 的冲突, 所以 $R(0, D)$ 设为正数。同理 $R(1, D)$ 设为负数。由于 $R(s_t, S)$ 进行频谱感知, 并没有造成和 PU 的冲突, 但是在没有数据传输的前提下耗费了时间,

降低了频谱利用率。所以 $R(s_t, S)$ 设为负数并且 $|R(s_t, S)| \ll |R(1, D)|$ 。基于以上定义，反馈函数能够寻找频谱利用率和 PU 冲突概率之间的折中方案。例如可以通过增加 $|R(1, D)|$ 的方式，以频谱利用率为代价，减小 SU 和 PU 的冲突概率。

4.3 决策过程

基于 POMDP 的决策过程通过在每个决策时段计算信念向量的方式选择最佳行为。信念向量^[6]包含选择最佳行为的所有信息。在 POMDP 中，决策函数用于将每个决策时段的信念向量映射到下一个时段所选用的行为。在所有可能的策略中，POMDP 旨在寻找能够最大化目标函数的最佳策略。下面将分别介绍信念向量和最佳策略的计算过程。

记 $BE_t = (BE_t^0, BE_t^1)$, s.t. $BE_t^0 + BE_t^1 = 1$ 为第 t 个决策时段的信念向量。其中， BE_t^0 、 BE_t^1 分别表示由第 $t-1$ 决策时段的行为和观测值推断得到的 PU 在 t 决策时段处于状态 0 和状态 1 的概率。在第 t 个决策时段结束时，决策过程根据 a_t 和 y_t 更新 BE_t 至 BE_{t+1} 。如图 5 所示，PU 驻留时间均值为 $\frac{1}{\mu}$ ，并且 PU 到达间隔均值为 $\frac{1}{\lambda}$ 。所以状态的初始分布为 $BE_0 = \left(\frac{\mu}{(\lambda + \mu)}, \frac{\lambda}{(\lambda + \mu)} \right)$ 。

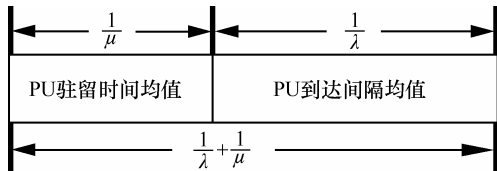


图 5 PU 驻留时间和到达间隔均值

如果 $a_t = D$ ，则决策算法得到观测值 $y_t = \text{Null}$ 。此时决策算法并没有得到任何指示 PU 真实状态的有用信息。在这种情况下根据 $BE_{t+1} = f(BE_t)$ 更新信念向量。

$$f(BE_{t+1}) = \left(\sum_{i=0}^1 p_{i,0}^D BE_t^i, \sum_{i=0}^1 p_{i,1}^D BE_t^i \right) \quad (14)$$

如果 $a_t = S$ ，则 SU 进行能量感知，并将量化后的信号能量作为观测值。利用状态转移概率和贝叶斯理论，根据 $BE_{t+1} = g(BE_t)$ 更新信念向量。

$$g(y, BE_t) = \left(\frac{b_{0,y}^S \sum_{i=0}^1 p_{i,0}^S BE_t^i}{\bar{\omega}(y,t)}, \frac{b_{1,y}^S \sum_{i=0}^1 p_{i,1}^S BE_t^i}{\bar{\omega}(y,t)} \right) \quad (15)$$

其中， $\bar{\omega}(y,t)$ 为归一化因子。

$$\bar{\omega}(y,t) = b_{0,y}^S \sum_{i=0}^1 p_{i,0}^S BE_t^i + b_{1,y}^S \sum_{i=0}^1 p_{i,1}^S BE_t^i \quad (16)$$

记 $\rho: BE_t \rightarrow \{S, D\}$ 为将信念向量映射到行为的策略。在每个决策时段中，决策算法选择 $\rho(BE_t)$ 作为当前决策时段的行为。在所有可能的策略中，基于 POMDP 的决策算法旨在寻找能够使目标函数最大化的最佳策略 ρ^* 。根据文献[12]，最佳代价函数为 $V^* = BE \rightarrow \mathbb{R}$ 。 V^* 能够将当前决策时段的信念向量映射为最佳策略所对应的总反馈值。

$$V^*(BE_t) = \max_{a \in \{D, S\}} U_a^*(BE_t) \quad (17)$$

$$U_D^*(BE_t) = \sum_{i=0}^1 BE_t^i R(i, D) + \alpha V^*(f(BE_{t-1})) \quad (18)$$

$$U_S^*(BE_t) = \sum_{i=0}^1 BE_t^i R(i, S) + \alpha \sum_{y=1}^M \bar{\omega}(y,t) V^*(g(y, BE_{t-1})) \quad (19)$$

通过 $V^*(BE)$ 可以得到最佳策略可表示为

$$\rho^*(BE_t) = \arg \max_{a \in \{D, S\}} U_a^*(BE_t) \quad (20)$$

可以使用多种动态规划算法^[13]得到最佳策略 ρ^* ，但是动态规划算法存在计算复杂度大的缺点^[13]，这可能会导致预测和判断出现延迟。

这里通过一个例子来说明使用网格算法^[13]得到的 BE_t^1 、最佳策略 ρ^* 和不同行为所对应的最佳代价函数值 $U_a^*(BE_t)$ 之间的关系，并说明采用单门限决策方法代替动态规划算法的原因。该例子中，所使用的参数如下： $\lambda = 50$ ， $\mu = 50$ ， $R(1, D) = -1$ ， $R(0, D) = 1$ ， $R(1, S) = R(0, S) = -0.1$ ， $\alpha = 0.99$ ， $T_S = 1 \text{ ms}$ ， $T_D = 5 \text{ ms}$ ， $F_S = 6 \text{ MHz}$ ， $SNR = -10 \text{ dB}$ 。由图 6 可知， $U_a^*(BE_t)$ 可以表示为 BE_t^1 的函数。当 $BE_t^1 \leq 0.15$ 时， $U_S^*(BE_t) \leq U_D^*(BE_t)$ ，选择数据传输行为能够获得更高的代价，所以最佳策略为 $a = D$ ；当 $BE_t^1 > 0.15$ 时， $U_S^*(BE_t) \geq U_D^*(BE_t)$ ，选择频谱感知行为能够获得更高的代价，所以最佳策略为 $a = S$ 。根据以上分析，可以通过一种简单的单门限方式确定下一决策时段的最佳策略 ρ^* 。此例中，判决门限 η 为 $U_S^*(BE_t)$ 和 $U_D^*(BE_t)$ 相交点所对应 BE_t^1 的值。此外，使用单门限决策方法仅需要在每个决策时

段使用式(14)~式(16)计算 BE_t^1 ，并不需要计算动态规划算法中的式(13)、式(17)、式(18)、式(19)、式(20)。单门限决策方法计算复杂度小于动态规划算法，所以这里采用单门限决策方法代替动态规划算法。

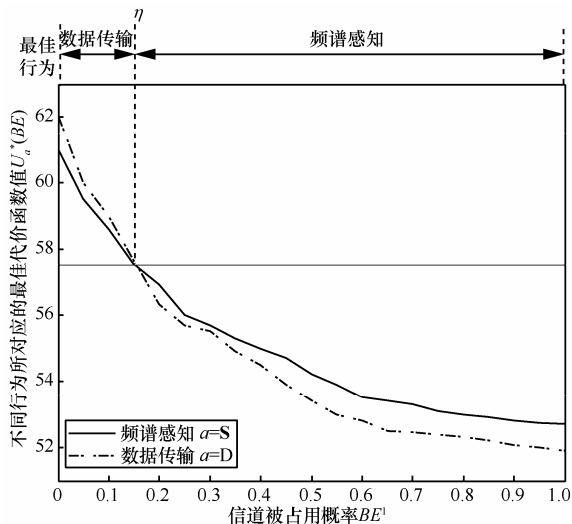


图 6 不同行为所对应的最佳代价函数值 $U_a^*(BE)$ 和 BE^1 的关系

4.4 基于单门限的低复杂度决策过程

获得 ρ^* 的流程如下。

1) 初始化

$$t = 0, \quad BE_0 = \left(\frac{\mu}{(\lambda + \mu)}, \frac{\lambda}{(\lambda + \mu)} \right)$$

2) 迭代 $t = 1, 2, 3, \dots$

如果 $BE_t^1 < \eta$, $\rho^*(BE_t) = D$ 。SU 在 T_D 内进行数据传输，并且按照 $f(BE_t)$ 更新 BE_t 至 BE_{t+1} 。

如果 $BE_t^1 > \eta$, $\rho^*(BE_t) = S$ 。SU 在 T_S 内进行频谱感知，根据式 (2) 得到能量感知结果 E_t ，量化后得到观测值 y_t 。按照 $g(y_t, \pi_t)$ 更新 BE_t 至 BE_{t+1} 。

3) 结束

数据传输完毕，结束迭代过程。

不同的门限 η 不仅会影响频谱利用率和碰撞概率之间的平衡，同样也会影响 CR 系统的性能。可以通过设定不同的 η 动态适应不同的应用环境。例如可以通过减小 η 确保只有在 SU 确信 PU 信号不存在的情况下才进行数据传输，以降低频谱利用率为代价换取低 SU 和 PU 的冲突概率。

5 仿真结果分析

该部分首先分析门限 η 对所提出的自适应频谱感知算法 (ASA, auto-adaptive sensing algorithm) 的

影响，其次在不同的 PU 单次驻留时间参数 μ 和不同的 SNR 的情况下验证 ASA 在不同 CR 环境下的顽奋赴性，并比较 ASA 和 STT^[2] (sensing-throughput tradeoff) 的性能。最后比较 SASA (same epoch ASA) 和 ASA 的性能。其中，SASA 和 ASA 的算法结构相同，二者的区别在于 SASA 的数据传输决策时段长度是固定的，并且和频谱感知决策时段长度相同， $T_S = T_D = 1 \text{ ms}$ 。而 ASA 的数据传输决策时段长度是由式 (12) 得到的最佳数据传输时间长度，频谱感知决策时段长度为 $T_S = 1 \text{ ms}$ 。此外，STT 算法在长度为 100 ms 的帧中，根据 PU 占用信道的概率动态调节感知时间以最大化系统吞吐量。

仿真参数如下。

根据 IEEE 802.22 标准^[9]，使用 QPSK 调制方式，单数据分组传输耗时为 $\tau = 0.85 \text{ ms}$ ， $T_S = T_H = 1 \text{ ms}$ 。抽样频率为 6 MHz，PU 信号的 SNR 取值范围为 $[-10 \text{ dB}, 0 \text{ dB}]$ ，PU 到达强度为 $\lambda = 50$ ，PU 驻留时间均值参数取值范围为 $\mu = [50, 60, \dots, 240]$ ，门限 η 的取值范围为 $[0.01, 0.91]$ ，数据传输量为 $N = 1000$ 。

相关名词解释如下。

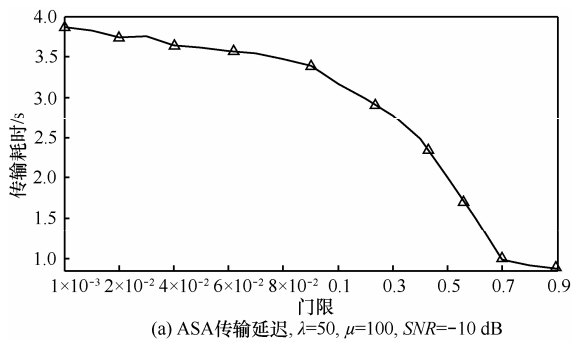
碰撞概率：PU 和 SU 同时传输数据的时间和 SU 数据传输时间的比值。

频谱利用率：在 SU 和 PU 传输数据不冲突的前提下，SU 传输数据的时间和信道空闲时间的比值。

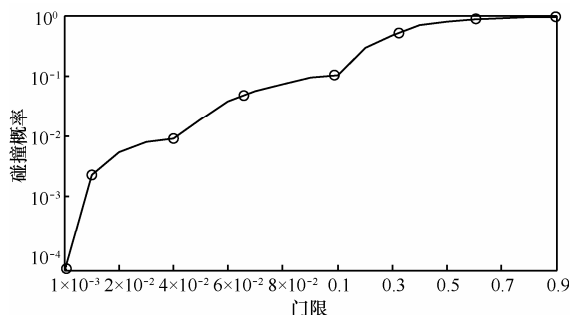
不同门限 η 下，ASA 的传输延迟和碰撞概率如图 7 所示。由图 7 可知，门限 η 是限制碰撞概率和传输延迟的关键因素，可以通过增加 η 以提高冲突概率为代价减小数据传输延迟。另外， η 和碰撞概率并不都是对等关系。例如 $\eta = 0.1$ 时，碰撞概率为 0.102，而 $\eta = 0.04$ 时，碰撞概率为 0.009 3。根据仿真结果，可以通过设置不同的 η 达到碰撞概率和传输延迟的折中。例如设 $\eta = 0.04$ 可以使碰撞概率小于 0.01 和传输延迟小于 3.7 s。

当 $\eta = 0.1$ 时，ASA 算法在不同的 PU 驻留时间参数 μ 和 SNR 下的传输延迟和碰撞概率如图 8 所示。由图 8 可知，当 μ 和 SNR 增大时，ASA 的传输延迟逐渐减小，碰撞概率始终稳定在 0.1 附近，并且散布范围较小。这也验证了 ASA 在不同 CR 外部环境中的顽健性。传输延迟逐渐减小的原因分析如下。1) 当 λ 固定时， μ 的增大导致频谱空洞时间增长，从而 ASA 能够获得更长的数据传输时间并减小数据传输延迟；2) 从宏观分析，SNR 的增大使得 ASA 能够在更少的感知时间内判断 PU 是否使用信道。从微观分

析, SNR 越大, 量化得到的 $b_{0,y}^S$ 和 $b_{1,y}^S$ 之间的欧氏距离越大。利用式(15)能够得到区别更大的 BE_i^0 和 BE_i^1 , ASA 算法能够较快地确定 PU 是否占用信道的结果。



(a) ASA 传输延迟, $\lambda=50, \mu=100, SNR=-10$ dB



(b) ASA 碰撞概率, $\lambda=50, \mu=100, SNR=-10$ dB

图 7 不同门限 η 下的传输延迟和碰撞概率

STT 在不同 μ 的情况下的传输延迟和碰撞概率如图 8 所示, STT 的数据传输时间小于 ASA, 但碰撞概率高于 ASA。原因分析: STT 的帧长固定为 100 ms, 所以传输 1 000 个数据分组只需要 9 个周期, 一旦检测到频谱空洞就在 91~98 ms 内进行数据传输, 能有效减小数据传输延迟。在频谱空洞为毫秒级的情况下, STT 的较长的数据传输时间会造成很大的碰撞概率。所以非自适应的固定帧长的数据传输方法并不适用于毫秒级频谱空洞的情况。图 9 为不同的 SNR 下 ASA 算法的传输延迟和碰撞概率。

ASA、SASA 在不同的碰撞概率下的频谱利用率和每秒钟的迭代次数如图 10 所示 (SASA 每秒的迭代次数为 1 000)。由图 10 可知, 二者的频谱利用率随着碰撞概率的增大而增大, SASA 的频谱利用率和 ASA 基本相同, 但是 SASA 的迭代次数要大于 ASA。并且, ASA 的迭代次数随着碰撞概率的增大而减小。原因分析: SASA 的决策时段均相等, 并没有考虑最佳单次数据传输时间。ASA 所使用的决策时段不等, 但是考虑了最佳单次数据传输时间 ($T_D > T_S$), 能够实现单信道最大数据分组传输量

$E(N_{AT})$ 。所以 ASA 能够在迭代次数小于 SASA 的情况下, 达到和 SASA 相近的频谱利用率。

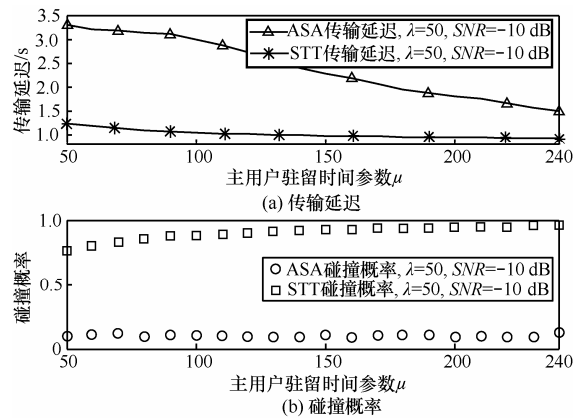


图 8 不同主用户驻留参数 μ 情况下的传输延迟和碰撞概率

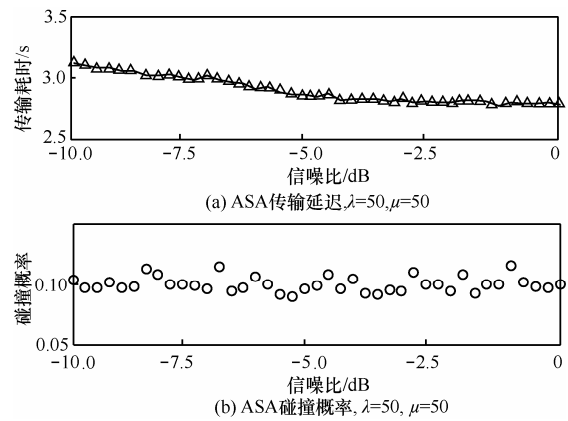
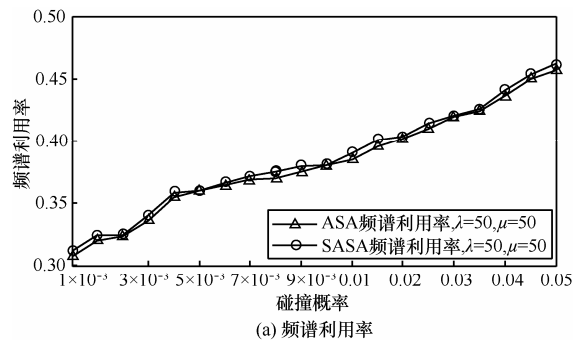
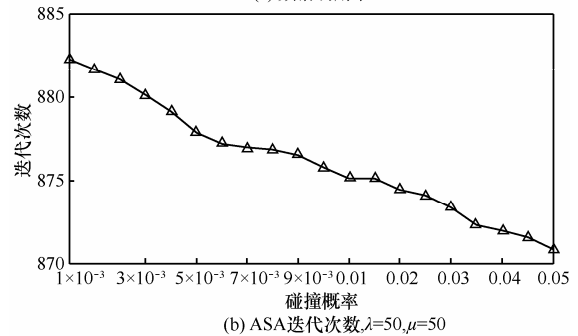


图 9 不同 SNR 下 ASA 算法的传输延迟和碰撞概率



(a) 频谱利用率



(b) ASA 迭代次数, $\lambda=50, \mu=50$

图 10 ASA 和 SASA 在不同碰撞概率下的频谱利用率和每秒的迭代次数

6 结束语

在认知无线电的实际应用中, 存在多种毫秒级频谱空洞场景。在设计次级用户频谱感知和数据传输机制时, 一方面要保证对主用户的保护, 另一方面要最大化认知无线电系统的频谱利用率。为实现上述目的, 本文首先推导出最佳单次数据传输时间, 并根据部分可观测马尔科夫过程建立了次级用户的频谱感知和数据传输调度算法。另外, 为了减小该算法的复杂度, 提出一种简单的基于单门限的迭代方式代替该算法。仿真结果表明, 和周期传输机制相比, 所提出的算法更适合实际的认知无线电应用环境。该算法能够通过设定不同的门限限制次级用户和主用户的碰撞概率, 并最大化频谱利用率。

附录: 式(10)推导过程

$$E(N_{AT}) = \frac{Ka(1-a^x)}{(1-a)}$$

$$\frac{a}{1-a} = \frac{(1-P)e^{-\lambda(K\tau+T_S)}}{1-(1-P)e^{-\lambda(K\tau+T_S)}}$$

当 x 比较小的情况下, $e^x \approx 1+x$, 所以

$$\frac{a}{1-a} \approx \frac{(1-P)(1-\lambda(K\tau+T_S))}{1-(1-\lambda(K\tau+T_S))}$$

$$= \frac{1-(P+\lambda T_S - P\lambda T_S) - (1-P)\lambda\tau K}{P+\lambda T_S - P\lambda T_S + (1-P)\lambda\tau K}$$

$$= \frac{1-B-CK}{B+CK}$$

$$K(1-a^x) = K - K((1-P)e^{-\lambda(K\tau+T_S)})^x$$

$$= K - K(1-P)^x e^{-\lambda(K\tau+T_S)x}$$

$$= K - K(1-P)^x e^{-\lambda T_S x} e^{-\lambda\tau Kx}$$

$$\approx K - K[(1-P)(1-\lambda T_S)]^x e^{-\lambda\tau Kx}$$

由于 $X = \frac{\delta}{(N/K)}$, $\delta(x)$ 表示采用进一法对 x 取整,

$X \approx \frac{N}{K}$, 所以

$$K(1-a^X) \approx K - K\left(1 - \frac{NP}{K_i}\right)e^{-\lambda\tau N}$$

$$= K_i(1 - e^{-\lambda\tau N}) + NP e^{-\lambda\tau N}$$

$$= EK + D$$

所以

$$E(N_{AT}) = \frac{(1-B-CK)(EK+D)}{B+CK}$$

对 $E(N_{AT})$ 求导便可得到式 (10)。

参考文献:

- [1] STOTAS S, NALLANATHAN A. On the throughput and spectrum sensing enhancement of opportunistic spectrum access cognitive radio networks[J]. IEEE Transactions on Wireless Communication, 2012, 11(1): 97-107.
- [2] LIANG Y C. Sensing-throughput tradeoff for cognitive radio networks[J]. IEEE Transactions on Wireless Communication, 2008, 7(4): 1326-1337.
- [3] WANG C L, HAN W. Throughput maximization for cognitive radio networks with wideband spectrum sensing[A]. IEEE WCNC 2012[C]. Paris, France, 2012. 1293-1298.
- [4] WILLKONMM D, MACHIRAJU S. Primary user behavior in cellular networks and implications for dynamic spectrum access[J]. IEEE Communications Magazine, 2009, 47(3):88-95.
- [5] ZHAO Q, YE J. Quickest detection in multiple on-off processes[J]. IEEE Transactions on Signal Processing, 2010, 58(12): 5994-6006.
- [6] ZHANG T Y, WU Y, LANG K. Optimal scheduling of cooperative spectrum sensing in cognitive radio networks[J]. IEEE System Journal, 2010, 4(4):535-549.
- [7] AHMAD S, LIU M, JAVIDI T. Optimality of myopic sensing in multichannel opportunistic access[J]. IEEE Transactions on Information Theory, 2009, 55(9):4040-4050.
- [8] SONG J, FENG Z, ZHANG P. Spectrum sensing in cognitive radios based on enhanced energy detector[J]. IET Communications, 2012, 6(8):805-809.
- [9] IEEE Std 802.22. Cognitive Wireless Ran Medium Access Control and Physical Layer Specifications[S]. 2011.
- [10] VIKRAM K. Algorithms for optimal scheduling and management of hidden Markov model sensors[J]. IEEE Transactions on Signal Processing, 2002, 50(6):1382-1398.
- [11] JAYAKRISHNAN U, VENUGOPAL V. Algorithms for dynamic spectrum access with learning for Cognitive radio[J]. IEEE Transactions on Signal Processing, 2010, 58(2):750-760.
- [12] MONAHAN G E. A survey of partially observable Markov decision processes: theory, models, and algorithms[J]. Manage Sci, 1982, 28(1): 1-16.
- [13] Tony's POMDP page[EB/OL]. <http://www.cs.brown.edu/research/ai/pomdp/index.html>, 2012.

作者简介:



许瑞琛 (1984-), 男, 河南焦作人, 北京邮电大学博士生, 主要研究方向为认知无线电频谱接入机制和隐马尔科夫理论。



蒋挺 (1962-), 男, 四川威远人, 博士, 北京邮电大学教授, 主要研究方向为通信技术、信息理论研究和应用。