基于 Q-learning 的机会频谱接入信道选择算法*

张 凯，李 鸥，杨白薇
（信息工程大学 通信工程系，郑州 450002）

摘 要：针对未知环境下机会频谱接入的信道选择问题进行研究。将智能控制中的 Q-learning 理论应用于信道选择问题，建立次用户信道选择模型，提出了一种基于 Q-learning 的信道选择算法。该算法通过不断与环境进行交互和学习，引导次用户尽量选择累积回报最大的信道，最大化次用户吞吐量。引入 Boltzmann 学习规则在信道探索与利用之间获得折中。仿真结果表明，与随机选择算法相比，该算法在不需要信道环境先验知识或预测模型下，能够自适应地选择可用性较好的信道，有效提高次用户吞吐量，且收敛速度较快。

关键词：认知无线电；机会频谱接入；Q 学习；信道选择；Boltzmann 规则
中图分类号：TP301.6 **文献标志码**：A **文章编号**：1001-3695(2013)05-1467-04
doi:10.3969/j.issn.1001-3695.2013.05.047

Q-learning based channel selection algorithm for opportunistic spectrum access

ZHANG Kai, LI Ou, YANG Bai-wei
(*Dept. of Communication Engineering, Information Engineering University, Zhengzhou* 450002, *China*)

Abstract: This paper considered the problem of channel selection for opportunistic spectrum access in unknown environment. It adopted the Q-learning of the intelligent control theory in the paper, and proposed a new Q-Learning based algorithm for channel selection considering the channel selection model of secondary user. The algorithm could lead secondary user to select channels with maximum cumulative reward, and maximized secondary user throughput. It adopted a Boltzmann learning rule to achieve well tradeoff between channel exploration and exploitation. From the simulation results, compared with random selection algorithm, the algorithm does not require prior knowledge or prediction models of the channel environment, yet can still select the optimal channel adaptively, improve the secondary user capability and attain to the convergence in short time.

Key words: cognitive radio; opportunistic spectrum access; Q-learning; channel selection; Boltzmann rule

0 引言

基于认知无线电技术[1]（cognitive radio technology）的机会频谱接入[2]（opportunistic spectrum access，OSA）作为一项缓解频谱资源紧张的关键技术，近年来引起了人们的广泛关注。机会频谱接入的核心思想[2]是认知无线电系统的次用户在不对主用户产生很大干扰的前提下，伺机地接入可用频谱，提高频谱利用率，从而可以用来解决频谱稀缺的问题。在机会频谱接入的关键技术中，次用户如何尽可能选择空闲概率高的可用信道，实现频谱的合理有效利用是机会频谱接入的重点和难点问题。信道的选择策略不仅要避免对主用户造成干扰，还要满足用户对频谱资源的需求，增大系统容量，提高频谱利用率。

OSA 研究在学术界受到了很高的重视，相关的研究成果[3~7]也较多。文献[3]考虑物理层频谱侦听和主用户的流量统计特性，提出一种基于部分马尔可夫决策过程（POMDP）的频谱决策算法，文献[4~6]对上述方法进行了更加深入的研究。文献[7]提出一种基于连续时间马尔可夫决策过程（CTMC）的主用户优先的动态频谱分配策略。但是 OSA 中仍然有一些尚未解决的问题。首先，现存的工作都是在假设次用户事先拥有环境的先验知识和动态模型的前提下开展的。其次，算法达到收敛要求环境是静态的。然而，这些假设在现实环境中很难实现[8]，因为：a）用户获取环境的知识需要花费大量的网络资源（如时间、能量和带宽），可能在一些场景中不可行；b）获取其他用户的信息也可能导致巨大的通信开销；c）OSA 系统中的频谱环境总是随时间而变化。因此，在不需要信道环境的先验知识和动态模型的前提下，亟待通过不断与环境进行交互和学习且能够实现优越性能的革新技术出现。

Q-learning[9]（Q 学习）作为一种无模型、无监督的在线强化学习算法（model-free, teacher-free, on-line reinforcement learning），是解决上述问题的有效途径之一。文献[10]利用强化学习理论解决认知无线网络的功率控制问题，然而仅限于频谱重叠共享[11]（spectrum underlay）环境中。文献[12]基于 Q 学习研究了认知无线电系统的感知管理算法。文献[8]将纳什均衡（NE）的思想引入未知环境下的 OSA 系统，实现了频谱分配的公平性。文献[13]尝试采用 Q 学习方法对 OSA 系统进行建模分析，但是备用信道的设置会增加频谱资源的开销，导致频带资源的浪费。

本文在上述研究的启发下，将 Q-learning 的思想应用于 OSA 系统的信道选择策略中，运用 Q-learning 算法选择合理的信道，提高 OSA 系统中次用户的吞吐量。本文考虑的 OSA 系统有如下特点：a）频谱空洞（spectrum holes）随时间而变化；b）无须知道信道空闲的统计信息；c）信道的设置符合实际情景，不预留备用信道，充分利用频带资源。不仅如此，次用户通过

收稿日期：2012-09-21；**修回日期**：2012-10-30 **基金项目**：国家科技部重大专项课题（2008ZX03006）

作者简介：张凯（1988-），男，江苏徐州人，硕士研究生，主要研究方向为认知网络和机会频谱接入（meetkai@126.com）；李鸥（1961-），男，教授，博士，主要研究方向为计算机网络和认知网络；杨白薇（1975-），女，副教授，博士，主要研究方向为无线传感器网络.

不断地与环境进行交互和学习,采用 Boltzmann 方法[14],借鉴模拟退火的思想优化动作的选择,使得基于 Q-learning 的信道选择算法能够尽可能地选择空闲概率高的信道,提高次用户的吞吐量,并且收敛速度较快。

本文给出了系统的模型并作出相应的假设,详细讨论本文给出的基于 Q-learning 的信道选择算法,并通过仿真详细分析算法性能。

1 系统模型与假设

假设认知无线电系统中的授权频谱可以划分为 m 个互不干扰的正交频段,即信道,每个信道采用 OFDM 的接入方式,一个信道可能对应于若干个用户,且每个信道的传输频率和覆盖范围各不相同。主用户不需要知道次用户的存在,可以随时接入信道,次用户必须及时退让。

在业务流方面,本文假设每个信道的主用户业务流分别以速率 λ_i 和 μ_i 的泊松过程到达和离开信道 i,即信道 i 上 2 个相继到达的主业务之间的时间间隔服从均值为 λ_i^{-1} 的指数分布,而每个主业务在信道 i 上的驻留时间服从均值为 μ_i^{-1} 的指数分布。设 η_i 为主用户在信道 i 上的业务负载,则可以表示为 $\mu_i^{-1}/(\lambda_i^{-1}+\mu_i^{-1})$。

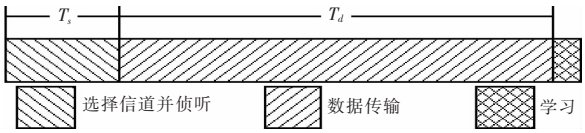

图 1 时隙结构图

次用户的传输结构如图 1 所示。假设次用户由于受到硬件限制在一个时隙内仅能选择一个信道用于传输[3]。在每个时隙的开始阶段,次用户根据当前的信道选择策略选择一个信道进行侦听。简单起见,假设信道侦听结果完全正确。如果选择的信道检测为空闲,次用户接入信道并使用,否则,次用户不得不中断传输。在时隙结束后,次用户根据当前获得的吞吐量收到一个回报,然后应用至学习算法更新信道选择策略。

2 基于 Q-learning 的信道选择算法

2.1 Q-learning 理论

Q-learning 基于的一个关键假设是智能体和环境的交互可看做是一个 Markov 决策过程(MDP),即由有限、离散的环境状态的集合 $S=\{s_1,s_2,\cdots,s_n\}$,有限、离散的学习者动作的集合 $A=\{a_1,a_2,\cdots,a_n\}$,标量的强化信号 r 和学习者的策略 $\pi:S\rightarrow A$ 等基本要素组成。在每一轮的迭代中,学习者感知环境状态 $s\in S$,并根据当前策略 π 选择动作 $a\in A$ 作用于环境;环境状态由此变化为 $s'\in S$,同时产生一个强化信号(称为"回报")$r(s,a)$ 反馈给学习者;学习者根据此更新其策略,并进入下一轮迭代。通过不断的反射学习优化一个可以迭代计算的 Q 函数提高学习能力,学习的最终目标是找到每个状态的最佳策略 $\pi^*(s)\in A$ 以最大化期望的长期累积回报(即状态的"值"):

$$V^{\pi}(s)=E(\sum_{t=0}^{\infty}\gamma^t r(s_t,\pi(s_t))\mid s_0=s) \tag{1}$$

其中:$\gamma\in(0,1)$ 为常数时间折现因子,它体现了未来回报相对当前回报的重要性。根据 Bellman 最优准则,式(1)的最大值为

$$V^*(s)=V^{\pi^*}(s)=\max_{\pi}V^{\pi}(s)=\max_{a\in A}[R(s,a)+\gamma\sum_{s'\in S}P_{s,s'}(a)V^*(s')] \tag{2}$$

其中:$R(s,a)$ 为 $r(s_t,a_t)$ 的数学期望;$P_{s,s'}(a)$ 为状态 s 在动作 a 的作用下达到状态 s' 的转移概率。

Q 学习的思想是不去估计环境模型,能够在未知 $R(s,a)$ 和 $P_{s,s'}(a)$ 的情况下,通过简单的 Q 值迭代找到最优的策略 π^* 满足式(2)。将策略 π 下的每一对状态和动作 (s,a) 与一个 Q 值相关联:

$$Q^{\pi}(s,a)=R(s,a)+\gamma\sum_{s'\in S}P_{s,s'}(a)V^{\pi}(s') \tag{3}$$

由式(2)和(3)可以得到

$$V^*(s)=\max_{a\in A}Q^*(s,a) \tag{4}$$

$$\pi^*(s)=\arg\max_{a}Q^*(s,a) \tag{5}$$

Q 学习通过以下迭代规则[9]来获得 $Q^*(s,a)$:

$$Q_{t+1}(s,a)=(1-\alpha)Q_t(s,a)+\alpha(r_t+\gamma\max_{a'}Q_t(s',a')) \tag{6}$$

其中:$\alpha=1/(1+\mathrm{visit}(s,a))$,$\alpha\in(0,1]$ 为学习率;$\mathrm{visit}(s,a)$ 为状态—行动对 (s,a) 被访问的总次数。随着 $t\rightarrow\infty$,若每对 (s,a) 的 Q 值能够经历无穷多次更新,且 α 递减至 0,则 $Q_t(s,a)$ 将以概率 1 收敛到最优值 $Q^*(s,a)$。此时,最优策略 π^* 可以由式(5)得到。

2.2 基于 Q-learning 的信道选择算法

本文将 Q-learning 应用于机会频谱接入系统中的信道选择策略中,信道选择和接入体系框架如图 2 所示。次用户根据信道选择和接入策略选择信道进行频谱侦听,观察无线频谱环境获得信道的忙/闲信息,作出信道接入决策,从而获得回报,实现与环境的交互。笔者把机会频谱接入建模成一个有限马尔可夫决策过程,它的状态空间 S、动作空间 A、状态转移函数 δ 和回报函数 r 的划分和选取分别如下:

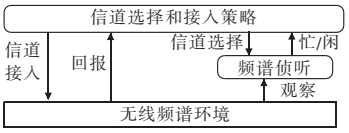

图 2 信道选择和接入体系框架

a)状态空间 S。S 由 m 个状态 $\{s_1,s_2,\cdots,s_m\}$ 构成。当次用户在当前时隙使用频带 b_i 且同时没有主用户在此频带上,则选取 s_i 作为当前状态。

b)动作空间 A。A 由 m 个动作组成,$A=\{a_1,a_2,\cdots,a_m\}$。其中,执行动作 a_i 时总是能够进入状态 s_i。若系统中的频带个数很大时,就会产生大量的状态—动作对组合,消耗过多的存储空间,这个问题可以通过神经网络[15]得到解决。

c)状态转移函数 δ。$S\times A\rightarrow S$ 代表给定当前状态 s_j 执行动作 a_k 系统进入至下一状态,对于任意状态—动作对 (s_j,a_k),$\delta(s_j,a_k)=s_k$。

d)回报函数 r。其设计基于系统性能的指标,系统的目的是执行动作后次用户能够使用可用频谱,获取最大系统吞吐量,那么规定这时的回报为正回报,其余的情况下回报为零。当前状态下的次用户在第 j 个时隙在频带 b_i 上获得的回报为

$$r(j)=\frac{T_d}{T_s+T_d}I_i\Phi_{T_d,j}(t)B_i \tag{7}$$

其中:I_i 表示频带 b_i 是否空闲;$\Phi_{T_d,j}(t)$ 为第 j 个时隙内频带 b_i 在传输时隙 T_d 内的空闲时间;B_i 为频带 b_i 的带宽。

2.3 Boltzmann 学习规则

对于在状态 s 的次用户选择动作的一个明显策略是选择使 $Q_t(s,a)$ 最大化的动作,从而"利用"(exploit)其当前近似的

Q。然而,使用此策略存在风险,次用户可能过度束缚到在早期训练中有高 Q 值的动作,而不能够“探索”(explore)到其他可能有更高值的动作。实际上,Q 学习的收敛性定理[9]要求每个状态—动作转换无限频繁地发生。显然,如果次用户总选择使当前 $Q_t(s,a)$ 最大的动作,将不能保证无限频繁性。“探索”使系统尝试未做过的动作,期望有得到更多回报的机会;而在“利用”过程中,系统更倾向于采取先前得到更多回报的动作。因此,为了在“探索”和“利用”之间取得折中,本文利用 Boltzmann 学习规则,使用概率的途径来选择动作。有较高 Q 值的动作被赋予较高的概率,但所有动作的概率都非 0。赋予这种概率的方法[14]是

$$P(a_i/s)=\exp\left[\frac{\hat{Q}(s,a_i^*)}{T}\right]/\sum_j \exp\left[\frac{\hat{Q}(s,a_j)}{T}\right] \tag{8}$$

其中:$P(a_i/s)$ 为次用户在状态 s 时选择动作 a_i 的概率;$T>0$ 为退火过程中的温度参数,它确定此选择优先考虑高 Q 值的程度。较小的 T 值会将较高的概率赋予超出平均 Q 的动作,致使次用户利用它所学习的知识来选择它认为会是回报最大的动作。相反,较大的 T 值会使其他动作有较高的概率,随机性较强,导致次用户探索当前 Q 值还不高的动作。本文的策略是以一个非常大的 T 值开始并逐渐减小至最终值 T_{final},从而在时间上从“探索”平滑地过渡到“利用”。

2.4　算法流程

通过以上状态、动作和回报等要素的选取可以看出,本文提出的基于 Q-learning 的机会频谱接入信道选择算法实际上是通过回报值的设计指导次用户选择信道,使其朝着最大化用户吞吐量的方向选择信道。并且,Q-learning 中的探索策略使得算法不会总为次用户选择一个回报值最大的动作,而是以一定概率选择其他动作,那么当信道环境发生变化时,这种机制保证了次用户能够在不同的信道上积累经验,信道忙/闲信息发生改变时,自适应地优化用户选择行为,选取最优信道。

基于 Q-learning 信道选择算法流程如下:

a)初始化。初始化 Q 值矩阵 $\boldsymbol{Q}=[Q(s,a)]_{M\times M}$ 为零矩阵,设置式(1)中的折现因子 γ,动作选择算法中的初始参数 T_0 和 T_{final}。从状态空间 S 中选取一个状态作为当前的一个激活状态 s,该状态下的当前信道作为激活信道。

b)动作的选择和执行。根据当前的 Q 值矩阵,采用 Boltzmann 学习规则,从动作空间 $A=\{a_1,a_2,\cdots,a_m\}$ 中以概率 $P(a/s)$ 选择一个动作 a,则当前状态 s 执行动作 a 系统进入至下一状态。

c)获得回报。在侦听时隙阶段(T_s),次用户侦听此状态下的信道,判断当前信道是否空闲;若信道空闲,则次用户进入数据传输时隙阶段(T_d),进行传输业务,依据式(7)获得回报值;若信道繁忙,则次用户不进行传输业务,获得回报为零。

d)Q 值更新。在时隙结束后,结合记录的状态—动作对 (s,a) 以及相应的回报 r,依据式(6)更新 Q 值矩阵。

e)参数更新。每个时隙迭代结束后,需要更新退火过程中的温度参数 T。本文设置它以负指数规律随着学习的过程逐渐减小,以满足 Q 学习中动作的选择,从探索平滑地过渡到利用。

f)重复步骤 b)~e),直至时间结束。

3　仿真结果与分析

仿真评估基于 Q-learning 的机会频谱接入信道选择算法的性能。作为参考,本文还仿真了随机选择策略下的性能。在随机选择策略中,次用户在每个时隙中以相同的概率任意选择信道进行侦听。本文采用蒙特卡洛实验策略,图中的每一点都是经过1 000次相互独立且在相同参数设置下的实验计算而得。信道选择算法的参数设置如下:信道个数 $m=7$,各信道带宽 $B=200$ kHz,每个时隙的总长度 $T_l=100\times10^{-3}$ s,并且为了满足侦听需求,设置侦听时间 $T_s=5\times10^{-3}$ s,则传输时间 $T_d=95\times10^{-3}$ s。设置时间折现因子 $\gamma=0.9$,突显未来回报的重要性。为了充分探索所有信道的忙/闲信息,退火过程的初始温度 $T=1\times10^{50}$,且以参数 $a=0.9$ 的负指数规律递减至最终温度 T_{final},T_{final} 设为 0.1。

不同信道上承载着不同的主用户业务流,如 VoIP、WLAN 和 Webmail 等,为了充分反映不同信道主业务分布和负载的不同,信道 i 上主业务驻留时间参数 μ_i^{-1} 的取值为[1,100]之间的随机数,η_i 的取值为(0,1)之间的随机数,λ_i^{-1} 则由业务负载 η_i 和 μ_i^{-1} 决定,单位为 ms。系统中主用户业务的平均负载为 $\bar{\eta}=(1/m)\sum_{i=1}^{m}\eta_i$(在图 4 中用 yTar 表示),设信道间主用户业务负载的平均变化系数 $\text{Cov}=\sigma/\bar{\eta}$,其中 σ 为业务负载的标准差。Q 学习的优势是在不需要主用户行为的先验知识的前提下,仍能依靠其自身能力收敛至理想的行为,而主用户的行为服从指数分布的假设是为了产生信道环境样例,用于通过不断的交互验证 Q 学习算法。

本文首先讨论次用户在主用户平均业务负载 $\bar{\eta}$ 所能实现的系统吞吐量。为了能够评价系统的总体平均性能,将 50 个时隙组成 1 个长时隙(即 5 s),以每个长时隙所能实现的吞吐量占理想情况下所能实现的吞吐量的比例作为评价指标。在此仿真场景中,设置 $\bar{\eta}=0.5$,Cov = 0.5。如图 3 所示,随着时间的增加,基于 Q-learning 的信道选择算法能够通过与环境的不断交互和学习获得显著优于随机接入策略的性能,并且在 20 个长时隙之后获得的吞吐量的比例逐渐稳定,收敛速度较快。

图 4 为基于 Q-learning 信道选择算法在不同平均业务负载 $\bar{\eta}$ 环境中的性能比较。此场景中 Cov = 0.5,平均业务负载 $\bar{\eta}$ 分别为 0.5 和 0.8。如图 4 所示,与预期一致,信道的平均业务负载 $\bar{\eta}$ 越小,系统获得的吞吐量越大。

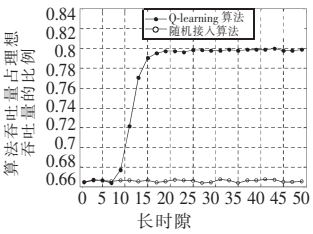

图 3　不同算法吞吐量占理想吞吐量的比例($\bar{\eta}$=0.5,Cov=0.5)

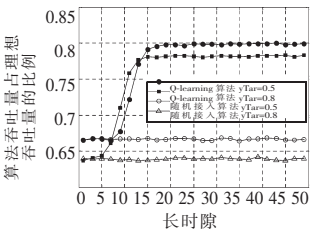

图 4　不同平均业务负载对吞吐量的影响

图 5 为基于 Q-learning 信道选择算法在不同的业务负载平均变化系数 Cov 环境中的性能:Cov = 0 和 Cov = 0.6(Cov 反映了不同信道间业务负载的差异,Cov 越大,信道之间的差异就越大),平均业务负载 $\bar{\eta}=0.8$。如图 5 所示,在 Cov = 0.6 时,基于 Q-learning 的信道选择算法所能实现的吞吐量占理想状态下的 80%,并且 Cov 越大,获得的吞吐量越多,这说明不同信道间较大的变化系数能够增大发现高可用信道的概率,另一方面,较大的变化系数能够使得次用户以较小时间开销发现可用信道;在 Cov = 0 时,Q-learning 信道选择算法随机接入算

法实现的性能较为接近,分别为 55% 和 54%。这是由于当信道间差异较大时,Q 学习策略在“探索”阶段能够发现空闲概率高的信道,而在“利用”阶段则以较高的概率使用它们,从而能够获得较大的吞吐量。

为了进一步评估文中所提信道选择算法的有效性,图 6 通过观察所提算法的信道选择过程,分析其性能。此场景中,时隙的个数设置为 4 000 个,蒙特卡洛实验次数为 1 000 次,设置 $\bar{\eta}=0.5$, Cov = 0.5,信道 1 ~ 7 的业务负载分别为[0.90, 0.88, 0.45, 0.44, 0.23, 0.43, 0.21]。如图 6 所示,在开始阶段,次用户处于“探索”阶段,等概率地尝试接入不同的信道,期望有获得更多回报的机会,随着时间的进行,次用户选择信道 5 和 7(平均业务负载较小)的次数 z 占当前时隙内总选择次数的比例逐渐增大,信道选择算法逐渐从“探索”阶段平滑地过渡到“利用”阶段。次用户选择理想信道的动作占动作选择总次数的 80%。这说明 Q-learning 算法通过在线学习,能够有效地将已有经验用于后续的信道选择中,从而获得性能的提升。

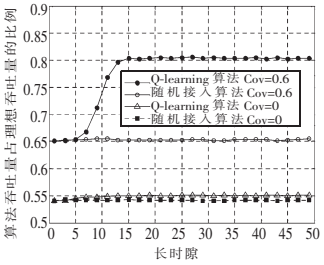

图 5 不同 Cov 对吞吐量的影响

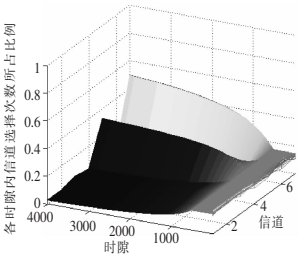

图 6 信道选择次数比例随时隙变化图

4 结束语

本文提出了一种基于 Q-learning 的信道选择算法,以解决机会频谱接入系统中的信道选择问题。同时,采用 Boltzmann 方法优化动作的选择,在信道“探索”和“利用”之间获得了很好的折中。仿真结果表明,基于 Q-learning 的信道选择算法在不需要信道环境先验知识和动态模型的前提下,通过与环境的不断交互和学习尽可能选择空闲概率大的信道,获得较好的吞吐量,并且收敛速度较快。

参考文献:

[1] HAYKIN S. Cognitive radio: brain-empowered wireless communications[J]. IEEE Journal on Selected Areas in Communications, 2005, 23(2): 201-220.

[2] ZHAO Qing, SADLER B M. A survey of dynamic spectrum access [J]. IEEE Signal Processing Magazine, 2007, 24(3): 79-89.

[3] ZHAO Qing, TONG Lung, 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.

[4] CHEN Yun-xia, ZHAO Qing, 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.

[5] AHMAD S, LIU Ming-yan, JAVIDI T, *et al*. Optimality of myopic sensing in multichannel opportunistic access[J]. IEEE Trans on Information Theory, 2009, 55(9): 4040-4050.

[6] CHEN Yun-xia, ZHAO Qing, SWAMI A. Distributed spectrum sensing and access in cognitive radio networks with energy constraint[J]. IEEE Trans on Signal Processing, 2009, 57(2): 783-797.

[7] WANG Bei-bei, JI Zhu, LIU K, *et al*. Primary-prioritized Markov approach for dynamic spectrum allocation[J]. IEEE Trans on Wireless Communications, 2009, 8(4): 1854-1865.

[8] XU Yu-hua, WANG Jin-long, WU Qi-hui, *et al*. Opportunistic spectrum access in unknown dynamic environment: a game-theoretic stochastic learning solution[J]. IEEE Trans on Wireless Communications, 2012, 11(4): 1380-1391.

[9] GOSAVI A. Reinforcement learning: a tutorial survey and recent advances[J]. Informs Journal on Computing, 2009, 21(2): 178-192.

[10] ZHOU Pan, CHANG Yu-sun, COPELAN J A. Reinforcement learning for repeated power control game in cognitive radio networks[J]. IEEE Journal on Selected Areas in Communications, 2012, 30(1): 54-69.

[11] AKYILDIZ I F, WON-YEOL L, VURAN M C, *et al*. A survey on spectrum management in cognitive radio networks[J]. IEEE Communications Magazine, 2008, 46(4): 40-48.

[12] 李默,徐友云,蔡跃明. 基于 Q-learning 的认知无线电系统感知管理算法[J]. 电子与信息学报, 2010, 32(3): 623-628.

[13] VENKATRAMAN P, HAMDAOUI B, GUIZANI M. Opportunistic bandwidth sharing through reinforcement learning[J]. IEEE Trans on Vehicular Technology, 2010, 59(6): 3148-3153.

[14] GAO Mao-zu, LIU Yang, MALEC J. A new Q-learning algorithm based on the metropolis criterion[J]. IEEE Trans on Systems Man, and Cybernetics, Part B: Cybernetics, 2004, 34(5): 2140-2143.

[15] WEISSENSTEINER A. A Q-learning approach to derive optimal consumption and investment strategies[J]. IEEE Trans on Neural Networks, 2009, 20(8): 1234-1243.

(上接第 1457 页)

[5] QI Y, KOBAYASHI H, SUDA H. Analysis of wireless relocation in a non-line-of-sight environment[J]. IEEE Trans on Wireless Communications, 2006, 5(3): 672-681.

[6] LAITINEN H, NORDSTROM T, LAHTEENMAKI J. Database correlation method for GSM location[C]//Proc of IEEE Vehicular Technology Conference. 2001.

[7] ZIMMERMANN D, BAUMANN J, LAYH M, *et al*. Database correlation for positioning of mobile terminals in cellular networks using wave propagation models [C]//Proc of IEEE Vehicular Technology Conference. 2004: 4682-4686.

[8] LI B, WANG Y, LEE H K, *et al*. Method for yielding a database of location fingerprints in WLAN [C]//Proc of IEEE Communications Conference. 2005: 580-586.

[9] LI B, RIZOS C, LEE H K. Utilizing Kriging to generate a NLOS error correction map for network based mobile positioning[C]//Proc of International Symposium on GPS/GNSS. 2004: 179.

[10] LEE H K, LI B, RIZOS C. Implementation procedure of wireless signal map matching for location-based services[C]//Proc of IEEE International Symposium on Signal Processing and Information Technology. 2005: 429-434.

[11] 陆大絟. 随机过程及其应用[M]. 北京: 清华大学出版社, 2005.

[12] 唐文胜,李珊,匡旺秋. RF 室内定位指纹库空间相关生成算法[J]. 计算机工程与应用, 2008, 44(23): 226-229.

[13] 王赛伟. 基于位置指纹的 WLAN 室内定位方法研究[D]. 哈尔滨: 哈尔滨工业大学, 2009.