

具有信道切换机制的 EWA 井下信道选择策略

孙 永, 孙立亚, 钱建生

(中国矿业大学 信息与电气工程学院, 江苏 徐州 221008)

摘 要:针对煤矿井下无线安全保障系统的无线信道链路质量问题, 基于经验加权吸引策略, 提出具有信道切换机制的信道选择算法。在时间轴上建立反映信道繁忙和空闲不同状态时间段的模拟无线信道环境, 与 Q 学习在实时信道选择轨迹和不同级别的信道切换次数进行对比, 结果表明, 经验加权吸引策略在优选信道快速捕获的准确性和稳定性方面更具优势。

关键词:井下无线通信; 认知无线电; 强化学习; 信道选择; 经验加权吸引

中图分类号: TN929.5

文献标志码: A

文章编号: 1672-3767(2014)03-0084-09

Channel Selection Strategy Based on EWA Learning with Channel Handoff in Coal Mine Wireless Communications

Sun Yong, Sun Liya, Qian Jiansheng

(School of Information and Electrical Engineering, China University of Mining and Technology, Xuzhou, Jiangsu 221008, China)

Abstract: Reliable wireless channel link quality is the prerequisite of coal mine wireless security system. This paper proposed an experience-weighted attraction (EWA) selection algorithm with channel handoff scheme. Research was conducted through contrastive analysis of Q learning algorithm after the establishment of a wireless simulation environments and time slots of different channel states in timeline. The comparison proves Q learning in real-time channel selection tracks and channel handoff times on different levels, channel selection strategy with EWA learning has higher advantageous performance in term of sensitivity and stability.

Key words: coal mine wireless communications; cognitive radio; reinforcement learning; channel selection; experience-weighted attraction

与传统通信系统不同, 煤矿井下无线通信环境恶劣^[1-2]: 无线信号在煤矿井下传播空间有限且不规则, 电磁波在煤矿井下多为经粗糙巷道壁及各种煤矿开采电气设备的表面折射、散射和漫射形成的非可视距多径传播, 无线电在煤矿井下衰减严重且巷道中电磁干扰严重。虽然理论上可利用的频段较多, 但是由于电磁波在井下传播的隧道效应^[3]以及煤矿巷道特性各不相同, 导致能用于通信传输的频段有限且不尽相同^[4-5], 因此, 煤矿井下真正可利用的频谱资源非常有限。

目前, 矿井无线通信系统主要有矿用超低频透地通信、中频感应通信、矿用漏泄通信、小区制矿井蜂窝状全双工无线通信、小灵通通信以及动力线载波通信等系统^[6]。赵红玉等^[7]针对矿井无线传输信道信号传输损耗大、传输距离短、传输信道窄、噪声干扰大、传输能力有限、语音信号信噪比低的特点, 通过对井下透地传输中语音信号和噪声的分析, 基于语音压缩技术提出了一种用于矿井无线透地通信中语音信号压缩的模型。张国鹏等^[8]根据相邻节点间是否存在可视链路以及巷道多径的密集程度, 对时域超宽带 (ultra wide band,

收稿日期: 2014-02-21

基金项目: 国家高技术研究发展计划 (“863”计划) 项目 (2012AA062103); 江苏省产学研联合创新基金前瞻性联合研究项目 (BY2012081)

作者简介: 孙 永 (1985—), 男, 江苏徐州人, 博士研究生, 主要从事煤矿井下认知无线电和认知无线传感网络方面的研究。

E-mail: SnowandSunShine@msn.com

UWB)井巷传播环境进行综合信道建模和参数化区分,以抗井巷多径传播为目标,提出无码间干扰的 UWB 信号发射策略,实现高效的 UWB 信号收发和改善无线多媒体传感器网络(wireless multi-media sensor network, WMSN)在煤矿井下的误码率。郭银景等^[9]针对井下无线信道干扰严重、节点能量有限的特点,提出了一种基于误码率和能量约束机制的井下无线自组织网络自愈 AODV(Ad hoc on-demand distance vector)路由协议。张婧^[10]根据矿井巷道结构,通过设置不同的网络业务负载、节点发射功率场景,对 DSDR(destination-sequenced distance-vector routing),AODV 和 DSR(dynamic source routing)三种 Ad hoc 无线传感网络路由协议进行了比较分析。陈捷等^[11]针对基于虚拟多进多出技术的无线传感器网络系统,分析了信道衰落情况下的平均误码率,提出了使用各协作节点实现最优传输功率的新方案。王志刚^[12]则提出了一种基于 CFDA (cellular fixed-wireless digital access) 的无线 mesh 网络信息采集方案。然而,不同井下无线通信系统的混合和有限无线资源的激烈争夺,将导致重要(高优先级)通信系统(如煤矿井下无线应急保障系统)无法正常通信和工作,严重危害煤矿井下的安全生产。同时,在采煤工作面,随着采煤机的推进,无线环境的空间几何形状在不断变化,其信道特性也在不断变化,因而传统的单一信道链路质量极为不稳定,易造成通信中断。具有智能感知无线通信环境并实时选择最优通信信道的认知无线电(cognitive radio, CR)在煤矿井下的应用能最大程度地保证重要系统的正常通信。

被认为是软件无线电(soft radio, SR)的扩展和延伸的认知无线电的概念由 Mitola^[13]首先提出。作为智能无线电(intelligent radio, IR)的革命性发展,认知无线电在软件无线电基础上加入了包含知识库(knowledge base),推理引擎(reasoning engine)和学习引擎(learning engine)在内的认知引擎(cognitive engine, CE),使得软件无线电能够学习和适应周围的无线环境^[14]。作为实现认知无线电的核心部件和关键功能,学习引擎负责通过积累新的环境经验,更新和扩展知识库,这也是区别于传统的预编程软件无线电的标志性特征。

目前已有多种学习算法应用于认知无线电的学习引擎中,包括强化学习(reinforcement learning, RL)算法在内的 Bayesian 统计学习算法、基因遗传算法和神经网络算法^[15-16]。其中,以 Q 学习为代表的强化学习研究最为广泛和深入。

在信道检测和分配中,Oksanen 等^[17]将 Q 学习应用于漏检控制。Chen 等^[18]则将其应用于认知无线网状网的功率控制。

在集中式认知无线网络结构中,Teng 等^[19]将 Q 学习应用于网络信道分配的双拍卖模型中。Zhang 等^[20]将 Q 学习和模糊逻辑判断相结合应用于认知无线网络动态频谱分配(dynamic spectrum allocation, DSA)、频谱效率和公平性的研究中。

在分布式认知无线网络结构中,Galindo-Serrano 等^[21]在无线局域网(wireless regional area networks, WRAN)控制和消除小区间累积干扰问题的研究中应用了多用户 Q 学习算法。而与 Q 学习十分相似的 R 学习算法则被 Zhu 等^[22]应用于长期网络效用最大化和码率最优化的跨层问题研究。

然而,上述算法多对收敛点(如 Q 学习的纳什均衡点)或次优点的存在性进行较多的研究,而对现场应用中的收敛速率等实时性问题缺乏更深入的探讨。基于 EWA(experience-weighted attraction)模型的信道选择算法^[23]能够使系统在线学习井下无线电环境通信信道特性,通过积累自我历史选择信道经验,快速推测并选择和更改当前最优通信信道,实时动态保证通信链路质量,降低系统通信中断概率。前期研究已经通过简单概率方法验证了该算法的有效性^[23],而现有理论模型并不能适应复杂多变的实际应用环境。在前期大量研究工作基础上,引入时间参数 t ,把重点转移至 EWA 智能信道切换机制研究。通过与 Q 学习算法的对比分析,研究所提出算法的时效性和稳定性。

1 EWA 理论

EWA 模型源于多人普通博弈理论。设共有 n 个人参与博弈,每个参与者标记为 $i(i=1,2,\dots,n)$ 。选手 i 的策略空间为 S_i ,包含有 m_i 个离散的选项,即 $S_i = \{s_i^1, s_i^2, \dots, s_i^{m_i}\}$ 。参与者 i 的一个策略 s_i 是 S_i 的一个元素,即 $s_i \in S_i$ 。整个博弈的策略空间 S 是 n 个人策略空间的笛卡尔乘积,即 $S = S_1 \times S_2 \times \dots \times S_n$ 。 s 标记为包

含每个参与者的个人策略组合,即 $s = \{s_1, s_2, \dots, s_n\} \in S$ 。则除去参与者 i 的其他 $n-1$ 个博弈者的个人策略组合表示为 $s_{-i} = \{s_1, s_2, \dots, s_{i-1}, s_{i+1}, \dots, s_{n-1}, s_n\}$ 。记 s_{-i} 的策略组合数为 m_{-i} , 则 $m_{-i} = \prod_{k=1, k \neq i}^n m_k$ 。对于参与者 i 的具有标量值的报酬函数为 $\pi_i(s_i, s_{-i})$ 。考虑时间维度,参与者 i 在 t 时段的策略表示为 $s_i(t)$, 其他参与者的策略集合(向量)则为 $s_{-i}(t)$, 相应地,记参与者 i 的报酬 $\pi_i[s_i(t), s_{-i}(t)]$ 。

EWA 学习算法假定每个策略都有一个吸引值。该模型定义吸引值的初始值、吸引值如何根据经验进行更新以及根据吸引值确定选择该策略的概率。算法的核心是每轮更新的 2 个变量。第一个变量是吸引值 $A_i^j(t)$, 定义为选择策略 j 后参与者 i 的吸引值。吸引值的更新规则为,当前吸引值 $A_i^j(t)$ 是含衰减系数 ϕ 的上一时段吸引值 $A_i^j(t-1)$ 与当前(虚拟)报酬 $\pi_i[s_i(t), s_{-i}(t)]$ 之和,并由更新的经验权重 $N(t)$ 归一化:

$$A_i^j(t) = \frac{\phi \cdot N(t-1) \cdot A_i^j(t-1) + \{\delta + (1-\delta) \cdot I[s_i^j, s_i(t)]\} \cdot \pi_i[s_i^j, s_{-i}(t)]}{N(t)} \quad (1)$$

其中: $I[\cdot]$ 为指示函数,定义如下:

$$I[x, y] = \begin{cases} 1, & x=y \\ 0, & x \neq y \end{cases} \quad (2)$$

该算法在每轮博弈后,用选择该策略后获得的实际报酬和模拟本轮博弈选择其他备选策略后所获得的虚拟报酬乘以权重系数,对吸引值进行更新。参数 $\delta(0 < \delta < 1)$ 在吸引值更新中用以衡量相对于实际获得报酬的虚拟报酬的权重。虚拟报酬可以被理解为选择其他备选策略所获得报酬的预期。

第二个变量是 $N(t)$, 定义为过去经验折合成当前经验的观测等价值。 $N(t)$ 值越大,过去经验对当前吸引值的影响越大。经验权值 $N(t)$ 的更新规则是乘以衰减系数 $\rho(0 < \rho < 1)$ 上一时段的经验权值 $N(t-1)$ 加上当前经验增量值 1:

$$N(t) = \rho \cdot N(t-1) + 1 \quad (3)$$

定义 $A_i^j(0)$ 和 $N(0)$ 为变量 $A_i^j(t)$ 和 $N(t)$ 的初始值。其中, $N(0)$ 可理解为博弈前思考对吸引值影响的等价评估值。

吸引值确定选择相应策略的概率。即概率 $P_i^j(t)$ 随吸引值 $A_i^j(t)$ 单调递增,而随 $A_i^l(t)$ (其中 $l \neq j$) 单调递减。采用指数形式的选择概率表达式如下:

$$P_i^j(t+1) = e^{\lambda \cdot A_i^j(t)} / \sum_{k=1}^{m_i} e^{\lambda \cdot A_i^k(t)} \quad (4)$$

参数 λ 用于衡量参与者对吸引值的敏感程度。参与者对吸引值越敏感, λ 也越大。

2 信道选择模型

在无线通信信道选择问题上,不同的无线信道应具有不同的信道可用性,即认知无线电所感知的不同信道的可用概率 α 应有区别。

设该无线电传输环境可划分为 n 个信道,则信道 $i(1 \leq i \leq n)$ 的可用概率用 α_i 表示,写成向量的形式 $\mathbf{A} = [\alpha_1, \alpha_2, \dots, \alpha_n]$ 。在时间轴 t 上,信道的可用概率体现在信号传输时段和空闲时段的信道占空比上。为此需引入 2 个新的参数,即信道数据传输完成速率 μ 和数据传输请求速率 γ ,则信道 $i(1 \leq i \leq n)$ 的信道数据传输完成速率 μ_i 和信道数据传输请求速率 γ_i 的向量形式分别为 $\mathbf{M} = [\mu_1, \mu_2, \dots, \mu_n]$ 和 $\mathbf{\Gamma} = [\gamma_1, \gamma_2, \dots, \gamma_n]$ 。信道可用概率和数据传输请求完成速率满足关系式 $\alpha_i = \gamma_i / (\mu_i + \gamma_i)$ 。考虑无线信道特性随时间的变化,不同时刻的信道可用概率以及 μ_i 和 γ_i 也不相同。上述参数引入时间参数 t 后的形式为

$$\alpha_i(t) = \gamma_i(t) / [\mu_i(t) + \gamma_i(t)]$$

根据 EWA 经验权值 $N(t)$ 的更新为式(3)。而吸引值 $A_i^j(t)$ 的更新表达式则应做相应修改,以适应认知无线信道选择和传输特点。改进后的数学表达式如下:

$$A_i^j(t) = \frac{\phi \cdot N(t-1) \cdot A_i^j(t-1) + \{\delta + (1-\delta) \cdot I[s_i^j, s_i(t)]\} \cdot \pi_i[s_i^j, s_{-i}(t)] \cdot \alpha_i(t)}{N(t)}$$

$$\frac{\eta + (1-\eta) \cdot I[1, x(j)] \cdot I[s_i^j, s_{-i}(t)] + (1-\eta) \cdot \{1 - I[s_i^j, s_{-i}(t)]\}}{N(t)} \quad (5)$$

其中:

$$x(j) = \begin{cases} 0, & \text{策略选择信道 } j \text{ 作为优先信道通信传输失败;} \\ 1, & \text{策略选择信道 } j \text{ 作为优先信道通信传输成功;} \end{cases} \quad (6)$$

$$\pi_i[s_i^j, s_{-i}(t)] = \begin{cases} 0, & \text{信道 } j \text{ 感知繁忙不可用} \\ 1, & \text{信道 } j \text{ 感知空闲可利用} \end{cases} \quad (7)$$

则信道选择概率 $P_i^j(t+1)$ 由式(4)决定。

由式(7)可知,在认知无线电进行无线环境传输信道感知搜索阶段,当感知到信道 j 当前状态繁忙(干扰严重以致无法选择该信道进行传输)时,将信道 j 的可用状态标记为 0,即不可用,则选择信道 j 策略应得不到任何报酬,即选择信道 j 策略相应的报酬函数 $\pi_i[s_i^j, s_{-i}(t)]$ 取值为 0;感知到信道 j 当前状态空闲(干扰低于传输门限可选择该信道进行传输)时,将信道 j 的可用状态标记为 1,即可用,则选择信道 j 策略理应得到相应的报酬,设此时的报酬函数 $\pi_i[s_i^j, s_{-i}(t)]$ 值为 1。

当认知无线电进行无线环境传输信道感知搜索结束后,状态繁忙的信道所得报酬为 0,而状态空闲的信道将标记为可用候选信道,当前状态可用但未被策略选择的信道可获得 $\delta \cdot \pi_i[s_i^j, s_{-i}(t)]$ 的(虚拟)报酬。本研究在式(5)引入了信道感知反馈机制,将当前状态可用且实际被策略选中信道的吸引值分两个阶段进行更新。第一阶段,当前状态可用且被策略选中的信道在传输开始前即获得 $\eta \cdot \pi_i[s_i^j, s_{-i}(t)]$ 的报酬,其中 $0 < \eta < 1$;第二阶段,当选择的信道 j 成功完成传输后将获得 $(1-\eta) \cdot \pi_i[s_i^j, s_{-i}(t)]$ 的报酬,如果选择的信道 j 在传输过程中通信中断失败,则得不到相应的反馈报酬。

通过对式(5)更深入的分析,可以得出参数 δ 和 η 的关系:

$$\delta = \eta + \tau \quad (8)$$

式(8)表明,(虚拟)报酬 $\delta \cdot \pi_i[s_i^j, s_{-i}(t)]$ 包含两个部分:信道感知时可用的实际报酬 $\eta \cdot \pi_i[s_i^j, s_{-i}(t)]$ 和假设以信道 j 作为通信传输信道,以一定概率通信传输成功所获得的虚拟报酬 $\tau \cdot \pi_i[s_i^j, s_{-i}(t)]$,其中, $0 < \tau < 1$ 。而相对应的实际通信传输成功所获得的报酬实为 $(1-\eta) \cdot \pi_i[s_i^j, s_{-i}(t)]$ 。

式(5)和式(8)的实际意义在于引入了多级报酬反馈体系:第一层次为信道感知反馈机制,在优选信道通信传输前,标记感知空闲可用信道为通信传输备选信道,给予 $\eta \cdot \pi_i[s_i^j, s_{-i}(t)]$ 可用于通信传输的实际报酬反馈;第二层次为信道传输反馈机制,已被标记为备选信道但实际未被策略选中进行通信传输的信道,给予 $\tau \cdot \pi_i[s_i^j, s_{-i}(t)]$ 可能通信传输成功的虚拟报酬反馈,被策略选中的优选信道且实际成功完成通信传输后给予 $(1-\eta) \cdot \pi_i[s_i^j, s_{-i}(t)]$ 的实际报酬反馈。本研究的意义在于建立了多级别多层次的报酬反馈体制,细化报酬成分要素,优化报酬反馈结构,使不同层次和级别的报酬反馈机制更加细致和合理。

根据智能信道切换机制,认知无线电在当前信道发生数据通信中断后,具有中中断续传功能,即记录通信数据中断点,自动切换至当前中断时间其他空闲信道,接入新的空闲信道后续传未完成的通信数据,这是信道切换的主要过程。如在当前中断点检出多个空闲可用信道,则根据算法选择概率最大的空闲信道作为信道切换接入的目标信道;若在当前中断点未检出空闲信道,则定义为发生全信道阻塞,等待并周期循环检测信道环境,直至检出并接入最先出现空闲的信道,然后续传剩余未完成的通信数据。

3 仿真分析

假设模拟仿真环境下的无线电传输信道数为 5,即 $n=5$ 。初始状态下各信道的选择概率相同 $P_i^j(0) = 1/n = 1/5 = 0.2$ 。经验权值归一化的初值 $N(0) = 1$,EWA 吸引值归一化的初值 $A_i^j(0) = 1$ 。衰减系数和敏感系数根据经验选取 $\phi = \rho = \lambda = 0.9$ 。因传输成功所获得的归一化报酬为 1,传输失败所获得的报酬为 0,假定在无任何先验概率的条件下,传输成功和失败的概率相同,则虚拟报酬权重系数 $\delta = 1 \times 0.5 + 0 \times 0.5 = 0.5$ 。由上节分析可知,反馈报酬权重系数应比虚拟报酬权重系数小,本研究取 $\eta = \delta/2 = 0.25$ 。为体现信道可用概率的一般性,选取 $0 \sim 1$ 均匀分布的概率向量作为各信道可用概率,即初始信道可用概率向量 $\mathbf{A}_1 = [0.1, 0.3, 0.5, 0.7, 0.9]$,则相应的信道数据传输完成速率向量 $\mathbf{M}_1 = [1/0.1, 1/0.3, 1/0.5, 1/0.7, 1/0.9]$ 和信道

数据传输请求速率 $\Gamma_0 = [1/0.9, 1/0.7, 1/0.5, 1/0.3, 1/0.1]$ 。为验证信道在可用概率发生改变的情况下,该智能算法能否在线实时准确决策和引导认知无线电切换到新的可用概率最高的传输信道,在模拟仿真过程中(25 回合后),初始的信道空闲概率向量 \mathbf{A}_0 将改变为 $\mathbf{A}_1 = \{0.9, 0.7, 0.5, 0.3, 0.1\}$,则相应的信道数据传输完成速率向量 $\mathbf{M}_1 = \{1/0.9, 1/0.7, 1/0.5, 1/0.3, 1/0.1\}$ 和信道数据传输请求速率 $\Gamma_0 = \{1/0.1, 1/0.3, 1/0.5, 1/0.7, 1/0.9\}$ 。认知无线电的数据传输完成和请求速率在整个过程中维持不变,分别为 $\mu_s = 1/0.1$ 和 $\gamma_s = 1/0.9$ 。考虑到实际无线环境下上述两种参数具有一定的突发性和随机性,在模拟仿真中生成服从各相应参数呈指数分布的随机数。

首先,根据各信道的数据传输完成速率和信道数据传输请求速率模拟生成信道在时间轴上繁忙和空闲时段,如图 1 所示。

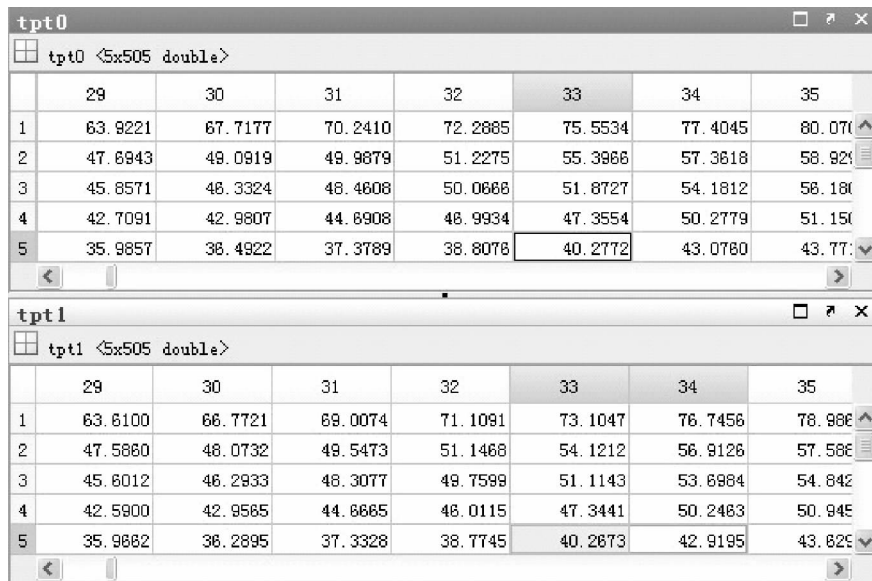


图 1 各信道状态转移时间仿真界面图

Fig. 1 The simulation interface of each channel state shift timeline

图 1 中,行标表示信道序号,列标表示时间轴标记序号。表 tpt1 记录各信道由空闲状态变为繁忙状态的时间点,表 tpt0 记录各信道由繁忙状态变为空闲状态的时间点。记录下各信道的状态转移时间点也就确定了各信道不同状态下的时段。如图中标记表示信道 5 在 40.267 3 至 40.277 2 的时间段为繁忙不可用状态,而在 40.277 2 至 42.919 5 的时间段为信道 5 的空闲可用状态。

无线信道环境建立之后,具有信道切换机制的基于 EWA 学习的信道选择算法即可实现实时在线模拟仿真,如图 2 所示。

图 2 中,表 sc 表示认知无线电选择的优选信道,表 tst1 表示该信道数据传输开始的时刻,tst0 表示该信道数据传输结束的时刻。图中标记记录了信道切换的过程:认知无线电在信道 5 从时间点 42.919 1 开始传输数据,认知无线电在时间点 42.919 5 检测到当前信道 5 由空闲可用状态变为繁忙状态,被动撤出当前传输信道,选择和接入空闲可用信道 4,继续传输剩余未完成的数据。需要说明的是,当算法选择信道 0 时,表明认知无线电遭遇全信道阻塞情况。在第 43 回合,基于 EWA 学习的信道选择算法检测到了全信道阻塞。在等待 0.012 5 时间单位 $[tst1(44)-tst0(43)]$ 后,自动接入最先转入空闲状态的信道 1,继续传输剩余数据。

经历 100 回合在线模拟仿真后,基于 EWA 学习的信道选择概率轨迹记录如图 3 所示。

由图 3 可以看出,基于 EWA 学习的信道选择算法,在初始各信道等概率的情况下随机地选择信道 3 作为接入信道,在初始化的过渡过程后,成功地判断跟踪和锁定信道 1 作为其优选信道。在第 25 回合后,由于可用概率的变化,信道 1 的选择概率急剧下降,而信道 5 的选择概率迅速攀升,并在第 38 回合后稳定超过信道 1 的选择概率,信道 5 最终取代信道 1 成为新的信道可用概率状态下的优选信道。

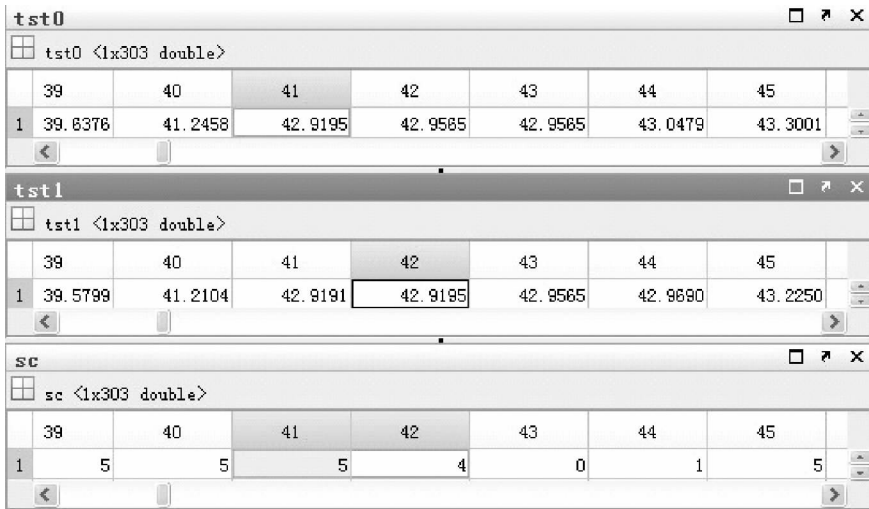


图 2 信道切换和信道阻塞仿真界面图

Fig. 2 The simulation interface of channel handoff and channel block

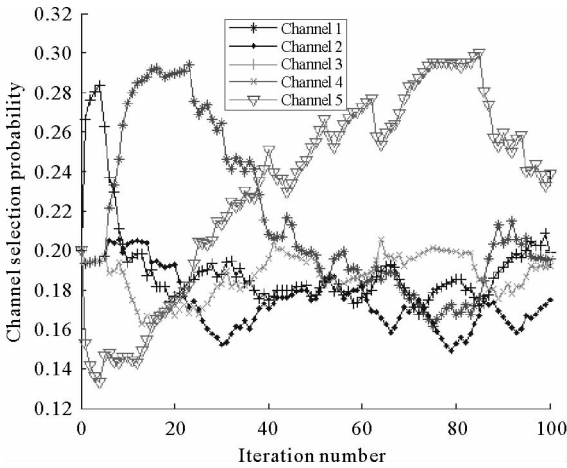


图 3 基于 EWA 学习的信道选择概率轨迹图

Fig. 3 Channel selection probability based on EWA learning

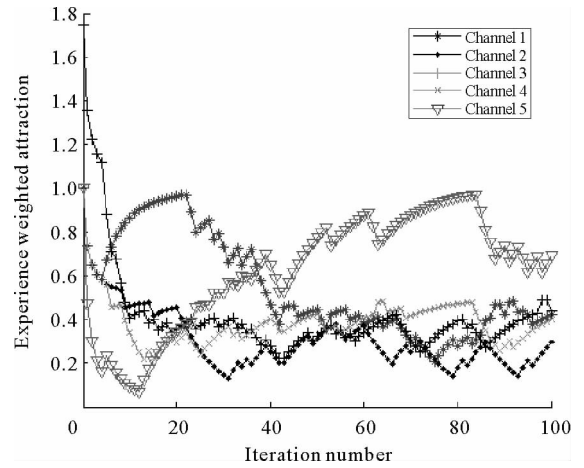


图 4 基于 EWA 学习的信道吸引值轨迹图

Fig. 4 Channel experience weighted attraction based on EWA learning

相应地,基于 EWA 学习的信道吸引值曲线如图 4 所示。对比图 3 可以发现,吸引值的变化趋势与信道选择概率的变化趋势相对一致。

为便于凸显 EWA 算法的优越性,在相同信道环境和初始条件下将其与 Q 学习的信道选择策略进行对比,如图 5 所示。

EWA 策略和 Q 学习信道选择算法的第一次差异出现在第 8 回合,Q 学习的信道选择在第 6、第 7 回合之后都选择第 1 信道作为优选信道的情况下,在第 8 回合却选择第 3 信道作为其传输信道,为了分析这一现象产生的原因,导出记录每回合后各信道选择概率表,如图 6 所示。

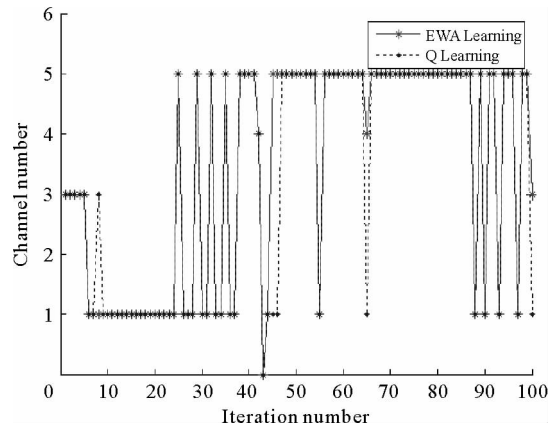


图 5 基于 EWA 学习和 Q 学习的信道选择结果比较

Fig. 5 The results of channel selection based on EWA learning and Q learning

表 prom 记录基于 EWA 学习算法每回合计算

得出的信道选择概率值;表 promr 记录基于 EWA 学习算法每回合计算得出的信道选择概率值。由图 6 可以看出,第 8 回合表 prom 中信道 1 的选择概率最大,其值为 0.233 2;而表 promr 中,信道 3 的选择概率虽在第 6 和第 7 回合连续呈现下降趋势,但在第 8 回合依然处于最大值 0.277 1,而此时信道 3 在经历了第 6 和第 7 回合的繁忙不可用状态后,在第 8 回合恢复了空闲可用状态。所以,第 8 回合重新选择信道 3 作为其优选信道是 Q 学习的信道选择按照其规则必然的判断和选择结果。这也直接体现了两种基于 EWA 学习的信道选择算法具有稳定的更快跟踪锁定切换到最优信道的性能。而在第 43 回合全信道阻塞后二者信道选择差异的原因同样可以用上述分析进行解释。

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	0.2000	0.1936	0.1942	0.1944	0.1945	0.1971	0.2217	0.2332	0.2461	0.2636	0.2744	0.2800	0.2847
2	0.2000	0.1936	0.1942	0.1944	0.1945	0.1971	0.2051	0.2037	0.2058	0.1986	0.2032	0.2043	0.2052
3	0.2000	0.2665	0.2759	0.2805	0.2833	0.2624	0.2371	0.2296	0.2113	0.2031	0.1940	0.1963	0.1982
4	0.2000	0.1936	0.1942	0.1944	0.1945	0.1971	0.1881	0.1897	0.1939	0.1889	0.1823	0.1748	0.1687
5	0.2000	0.1528	0.1416	0.1363	0.1333	0.1464	0.1480	0.1438	0.1429	0.1458	0.1462	0.1445	0.1431

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	0.2000	0.1808	0.1735	0.1697	0.1674	0.1658	0.1966	0.2216	0.2108	0.2311	0.2482	0.2626	0.2749
2	0.2000	0.1808	0.1735	0.1697	0.1674	0.1658	0.1683	0.1691	0.1681	0.1685	0.1684	0.1680	0.1675
3	0.2000	0.2769	0.3061	0.3213	0.3305	0.3366	0.2983	0.2711	0.2848	0.2633	0.2465	0.2333	0.2225
4	0.2000	0.1808	0.1735	0.1697	0.1674	0.1658	0.1683	0.1691	0.1681	0.1685	0.1684	0.1680	0.1675
5	0.2000	0.1808	0.1735	0.1697	0.1674	0.1658	0.1683	0.1691	0.1681	0.1685	0.1684	0.1680	0.1675

图 6 信道选择概率表界面图

Fig. 6 The probability table of channel selection

引入一定时间内信道切换次数,能够更加直观地体现 EWA 算法对比 Q 学习的信道选择策略上的优势。二者对比结果如图 7 所示。

图 7 中,Handoff 1 表示发生 1 次信道切换后完成数据传输的次数;Handoff 2 则表示连续 2 次信道切换后完成数据传输的次数;Handoff 3 表示连续 3 次及以上信道切换后完成数据传输的次数。可以看出,在 1 次信道切换后 Q 学习策略达到了 23 次,而 EWA 学习策略仅为 18 次;在连续 2 次信道切换后 EWA 学习策略则为 3 次,而 Q 学习策略的次数是 6 次,虽然连续 3 次及以上信道切换后二者相同(均为 2 次),但在总信道切换次数上,EWA 学习策略对比 Q 学习策略体现出了明显的优势。因而,基于 EWA 的信道选择算法比 Q 学习的信道选择策略具有更强的实时应变性,具有快速发现和适应信道环境变化的能力。

4 结束语

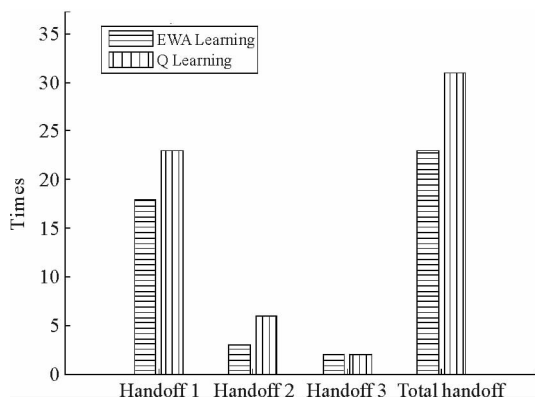


图 7 EWA 学习策略与 Q 学习策略的信道切换次数对比图

Fig. 7 Channel handoff times comparison between EWA learning and Q learning

提出了基于 EWA 学习的信道选择算法,引入了信道感知反馈和信道切换机制,使得系统能够在线学习

井下无线通信信道特性,通过经验积累,选择或更改当前最优通信信道,实时动态保证通信链路质量,降低系统通信阻塞中断概率。仿真结果表明,该算法能够在线实时准确决策并引导认知无线电切换到新的可用概率最高的传输信道,与当前强化学习中主流研究的Q学习算法策略相比,基于EWA学习算法的信道选择策略具有更加稳定的性能和更加快速准确捕获优势信道的能力。

参考文献:

- [1] 盛惠兴,方志远,霍冠英. OFDM技术及在认知无线电中的应用[J]. 信息安全与通信保密,2007(5):71-73.
Sheng Huixing, Fang Zhiyuan, Huo Guanying. OFDM technologies and applications in cognitive radio[J]. Information Security and Communications Privacy,2007(5):71-73.
- [2] 王文清. 煤矿井下无线电噪声的研究[J]. 煤炭学报,2011,36(10):1773-1776.
Wang Wenqing. Study on the radio noise in coal mine[J]. Journal of China Coal Society,2011,36(10):1773-1776.
- [3] 张守祥,王汝琳,刘芳. 基于OFDM的巷道和工作面无线移动通信研究[J]. 煤矿安全,2006,37(3):13-16.
Zhang Shouxiang, Wang Rulin, Liu Fang. Study of wireless mobile communication for laneway and face based on OFDM[J]. Safety in Coal Mines,2006,37(3):13-16.
- [4] 孙继平,成凌飞,张长森. 截面尺寸对矩形巷道中电磁波传播的影响[J]. 中国矿业大学学报,2005,34(9):596-599.
Sun Jiping, Cheng Lingfei, Zhang Changsen. Influence of transverse dimensions on electromagnetic waves propagation in rectangular tunnels[J]. Journal of China University of Mining & Technology,2005,34(9):596-599.
- [5] 孙继平,成凌飞,张长森. 电导率对巷道中电波传播的影响[J]. 辽宁工程技术大学学报:自然科学版,2007,26(2):96-98.
Sun Jiping, Cheng Lingfei, Zhang Changsen. Influence of conductivity on radio waves propagation in tunnels[J]. Journal of Liaoning Technical University: Natural Science Edition,2007,26(2):96-98.
- [6] 胡穗延. 煤矿自动化和通信技术现状与发展趋势[J]. 煤炭科学技术,2007(8):1-4.
Hu Suiyan. Present status and development tendency of mine automation and communication technology[J]. Coal Science and Technology,2007(8):1-4.
- [7] 赵红玉,陶晋宜,王园媛,等. 矿井无线透地通信中语音信号压缩的研究[J]. 煤炭工程,2014(3):20-22,25.
Zhao Hongyu, Tao Jinyi, Wang Yuanyuan, et al. Study on compression of voice signal in mine wireless through-the-earth communication[J]. Coal Engineering,2014(3):20-22,25.
- [8] 张国鹏,王艳芬,丁恩杰. 矿井无线多媒体传感器网络UWB信号收发策略研究[J]. 煤炭科学技术,2013(12):71-75.
Zhang Guopeng, Wang Yanfen, Ding Enjie. Study on UWB signal transmitting and receiving strategy of mine wireless multimedia sensor network[J]. Coal Science and Technology,2013(12):71-75.
- [9] 郭银景,杨洪鸾,杨涛,等. 一种井下无线自组织网络自愈路由算法[J]. 山东科技大学学报:自然科学版,2012,31(1):84-88.
Guo Yinjing, Yang Hongqian, Yang Tao, et al. A self-healing routing algorithm for the underground wireless self-organizing network[J]. Journal of Shandong University of Science and Technology: Natural Science,2012,31(1):84-88.
- [10] 张婧. 基于Ad hoc的矿井巷道无线传感器网络路由协议分析[J]. 山西大学学报:自然科学版,2013,36(4):586-590.
Zhang Jing. Analysis on routing protocols for WSNs based on Ad hoc in the mine tunnel[J]. Journal of Shanxi University: Natural Science Edition,2013,36(4):586-590.
- [11] 陈捷,阚保强. 衰落信道下的无线传感器网络虚拟MIMO传输优化设计[J]. 山东科技大学学报:自然科学版,2013,32(5):96-101.
Chen Jie, Kan Baoqiang. Optimal design of WSN based on virtual MIMO under fading channels[J]. Journal of Shandong University of Science and Technology: Natural Science,2013,32(5):96-101.
- [12] 王志刚. 一种无线Mesh网络煤矿监控系统实现[J]. 煤炭技术,2014(1):161-163.
Wang Zhigang. A wireless mesh network in coal mine monitoring system[J]. Coal Technology,2014(1):161-163.
- [13] Mitola J. Cognitive radio for flexible mobile multimedia communications[J]. Mobile Networks & Applications,2001,6(5):435-441.
- [14] Bantouna A, Stavroulaki V, Kritikou Y, et al. An overview of learning mechanisms for cognitive systems[J]. Eurasip Journal on Wireless Communications and Networking,2012(22):1-6.
- [15] Tsagkaris K, Katidiotis A, Demestichas P. Neural network-based learning schemes for cognitive radio systems[J]. Computer Communications,2008,31(14):3394-3404.

- [16] Katidiotis A, Tsagkaris K, Demestichas P. Performance evaluation of artificial neural network-based learning schemes for cognitive radio systems[J]. Computers & Electrical Engineering, 2010, 36(3): 518-535.
- [17] Oksanen J, Lunden J, Koivunen V. Reinforcement learning based sensing policy optimization for energy efficient cognitive radio networks[J]. Neurocomputing, 2012, 80: 102-110.
- [18] Chen X F, Zhao Z F, Zhang H G, et al. Reinforcement learning enhanced iterative power allocation in stochastic cognitive wireless mesh networks[J]. Wireless Personal Communications, 2011, 57(1): 89-104.
- [19] Teng Y L, Yu F R, Han K, et al. Reinforcement-learning-based double auction design for dynamic spectrum access in cognitive radio networks[J]. Wireless Personal Communications, 2013, 69(2): 771-791.
- [20] Zhang W Z, Liu X C. Centralized dynamic spectrum allocation in cognitive radio networks based on fuzzy logic and Q-learning[J]. China Communications, 2011, 8(7): 46-54.
- [21] Galindo-Serrano A, Giupponi L. Distributed Q-learning for aggregated interference control in cognitive radio networks[J]. IEEE Transactions on Vehicular Technology, 2010, 59(4): 1823-1834.
- [22] Zhu J, Wang J, Luo T, et al. Adaptive transmission scheduling over fading channels for energy-efficient cognitive radio networks by reinforcement learning[J]. Telecommunication Systems, 2009, 42(12): 123-138.
- [23] Sun Y, Qian J S. Cognitive radio channel selection strategy based on experience-weighted attraction learning [J]. TELKOMNIKA Indonesian Journal of Electrical Engineering, 2014, 12(1): 149-156.

(责任编辑: 吕文红)

“矿山物联网技术”研究专栏征稿

征稿范围:

- | | |
|--------------|-------------|
| ◇ 矿山物联网架构 | ◇ 矿山通信网络 |
| ◇ 矿用传感器技术及网络 | ◇ 矿山数据仓库 |
| ◇ 井下人员定位技术 | ◇ 矿山应急指挥系统 |
| ◇ 矿山 GIS | ◇ 数字矿山理论与技术 |
| ◇ 矿山监测监控技术 | ◇ 矿山虚拟现实技术 |

欢迎相关领域专家学者和工程技术人员踊跃投稿, 来稿请注明“矿山物联网技术”研究专栏。稿件通过专家评审后优先发表。

投稿平台: http://xuebao.sdust.edu.cn/index_z.asp

电子邮箱: xbgjcl@126.com

联系电话: 0532-86057826

山东科技大学学报(自然科学版)编辑部