

# 认知无线网络中提高传输层端到端吞吐率的跨层参数配置\*

江虹<sup>†</sup> 刘从彬 伍春

(西南科技大学信息工程学院, 绵阳 621010)

(2012年4月9日收到; 2012年6月24日收到修改稿)

在认知无线网络中, 传输层端到端 (TCP) 吞吐率是衡量网络性能的重要指标. 前期相关研究大都具有以下两方面缺点: 第一, 大部分研究只考虑了协议底层参数来优化物理链路性能, 对传输层性能有所忽略; 第二, 目前的研究大都基于马尔可夫决策过程建模, 这需要网络具有完全知识, 使得这类模型的应用受到很大限制. 针对以上问题, 本文提出一种新的算法: 网络中每个节点通过联合配置物理层调制方式、发射功率、链路层信道接入和 TCP 拥塞控制因子来找到传输层端到端近似最优吞吐率. 由于无线设备对环境感知存在误差, 本文将网络模型建模为部分可观测马尔可夫决策过程, 并将其转换成信念状态马尔可夫决策过程, 采用  $Q$  值迭代找到近似最优策略. 仿真分析表明, 提出的算法能在动态无线环境下以一定的误差限收敛于最优策略, 能在功率受限条件下, 有效提高传输层端到端吞吐率.

**关键词:** 认知无线网络, 传输层吞吐率, 信念状态马尔可夫决策过程,  $Q$  学习

**PACS:** 88.80.hp

**DOI:** 10.7498/aps.62.038804

## 1 引言

随着无线应用的不断发展, 无线频谱资源变得越来越紧张, 因此, 以参数动态配置方法有效利用无线资源为目的的认知无线电 (CR) 技术<sup>[1-4]</sup>引起了业界的广泛关注. 多节点 CR 网络一般由相互连接的 CR 节点以及授权用户组成, CR 节点通过对环境的感知、学习进行最优参数配置, 以达到提高网络性能的目的. 而传输层端到端吞吐率 (以下称 TCP 吞吐率) 是衡量网络性能的重要指标, 该领域相关研究现状如下.

加州大学洛杉矶分校 Mihaela van der Schaar 等<sup>[5-8]</sup>在 CR 网络参数配置方面进行了较深入的研究. 文献 [5, 6] 分别提出了一种基于优先队列和分解原理的最优路由选择算法. 文献 [7, 8] 结合分

层马尔可夫决策过程 (MDP) 和动态规划, 分别提出了在线最优路由方案和路由与功率联合设计方案. 这类方法的缺点是基于完全知识进行学习, 由于环境多变或感知误差, 使得其应用受到很大限制. 文献 [9, 10] 通过引入部分可观测马尔可夫决策过程 (POMDP), 分别解决 CR 网络中频谱感知和频谱接入的联合设计问题, 提出了对授权用户干扰限制下的认知媒介访问控制 (MAC) 协议. 此类方法将 POMDP 模型引入 CR 网络中, 对求解环境信息部分可知问题具有借鉴意义. 但以上研究只考虑了部分协议层参数, 如物理层 (PHY)、MAC 层或路由层参数, TCP 性能却受到忽视, 还需深入研究.

针对基于无线环境的 TCP 吞吐率改进方面. 文献 [11, 12] 联合优化 PHY 调制编码模式、MAC 层帧长度和信道选择来提高 TCP 性能; 文献 [13] 将信道选择与 TCP 吞吐率联合考虑, 提出一种提高

\* 国家自然科学基金 (批准号: 61072138) 和国防基础科研项目 (批准号: B3120110005) 资助的课题.

<sup>†</sup> 通讯作者. E-mail: jianghong@swust.edu.cn

TCP 吞吐率的最优信道接入方案. 但这类方法具有共同缺点: 采用 TCP Reno 模型, 这种模型将拥塞控制因子固定为 0.5, 使得 TCP 吞吐率模型在无线环境中的适应性受到很大限制; 同时, PHY 信道复用方式效率较低, 造成了较低的频谱利用率.

针对上述问题, 本文提出一种基于信念状态 MDP (BMDP) 的近似最优跨层参数配置方案: 基于 POMDP 对网络建模, 根据观测历史将 POMDP 问题转化为对 BMDP 的求解, 并采用  $Q^{[14]}$  值迭代 (即  $Q$ -BMDP) 求解和优化.

## 2 网络模型描述

设某 CR 网络包括相互通信的授权用户、 $L$  个 CR 节点和  $C$  个可感知信道 (类似于文献 [15, 16]), CR 节点通过探索空闲频段实现数据传输. 本文设决策周期是长度为  $T$  的一个时隙, 每个时隙分为感知决策、包传输和确认三个阶段. 第一阶段启动  $Q$ -BMDP 算法实现跨层参数配置 (发射功率、调制模式、信道选择、拥塞控制因子); 第二阶段将 TCP 数据包在某一组信道上从源节点传输到目的节点; 第三阶段通过公共控制信道确认包到达, 获得回报. 时隙结构如图 1 所示.

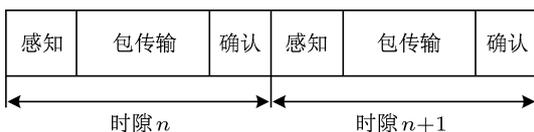


图 1 时隙结构

### 2.1 PHY 模型

在加性高斯白噪声信道下第  $n$  个时隙, 设  $Nos_j \in R, (1 \leq j \leq L)$  为 CR 节点  $j$  的接收背景噪声,  $Pow_i \in R, (1 \leq i \leq L)$  为节点  $i$  的发射功率. 则节点  $j$  的接收信干噪比 (SINR <sub>$j$</sub> ) 为 [8]

$$SINR_j = 10 \lg \frac{G_{ij} Pow_i}{\sum_{k \in L, k \neq i} G_{kj} Pow_k + Nos_j}, \quad (1)$$

其中  $G_{ij}$  为  $i-j$  链路的增益, 本文采用正交频分复用 (OFDM), 子信道采用自适应 MPSK 和 MQAM 调制模式, 由文献 [16, 17] 可推导节点  $j$  接收误比

特率 (BER <sub>$j$</sub> ) 为

$$BER_j = \begin{cases} Q(\sqrt{2\gamma_b}) & \text{QPSK,} \\ \frac{4}{\text{lb}M} \left(1 - \frac{1}{\sqrt{M}}\right) \\ \times \sum_{i=1}^{\sqrt{M}/2} Q\left((2i-1)\sqrt{\frac{3\gamma_b \text{lb}M}{M-1}}\right) \\ \times \text{MQAM: } M = 4^m, \quad m = 2, 3, 4, \dots, \\ \frac{2}{\text{lb}M} Q\left(\sqrt{2\gamma_b \text{lb}M} \sin \frac{\pi}{M}\right) \\ \times \text{MPSK: } M = 2^m, \quad m = 3, 4, 5, \dots, \end{cases} \quad (2)$$

其中  $\gamma_b$  为节点  $j$  的接收比特信干噪比, 由 (1) 式得  $\gamma_b = \frac{W_i}{R_j} \text{SINR}_j$ ,  $W_i$  为节点  $i$  使用的信道带宽,  $R_i$  为数据速率.

### 2.2 MAC 层与 TCP 层模型

在  $C$  个感知信道基础上, 设节点  $i$  可动态选择  $w_i (1 \leq w_i \leq C)$  个信道接入, 总带宽为  $W_i$ . 设链路层帧头长度固定为  $L_{frh}$ , 帧长度为  $L_{fri}$ , 则  $i-j$  链路上的帧错误率可由 BER <sub>$j$</sub>  和  $L_{fri}$  表示为

$$Fe_{ij} = 1 - (1 - BER_j)^{L_{fri}}. \quad (3)$$

设 TCP 包大小为  $L_{tcp}$ , 则每个包需  $N_{fri} = \text{ceil}(L_{tcp}/(L_{fri} + L_{frh}))$  个 MAC 帧组成,  $\text{ceil}(x)$  表示对  $x$  向上取整. 若采用自动重发请求 (ARQ) 协议传输 TCP 包, 最大重传次数为  $R_e$ , 一个 TCP 包在  $i-j$  链路被成功接收的条件是所有帧都被节点  $j$  成功接收, 则包错误率表示为

$$Pe_{ij} = 1 - (1 - Fe_{ij}^{R_e+1})^{N_{fri}}. \quad (4)$$

设 TCP 采用滑动窗口协议进行流量控制, 利用超时重传和捎带确认机制保证传输可靠性. 若 TCP 最大拥塞控制窗口长度为  $cwnd$ , 节点  $j$  在单位时间接收的数据量  $U_{ij}$  表示为 [18]

$$U_{ij} = \left[ \frac{1}{PER_{ij}} + \frac{WD_{ij}^2 Fe_{ij}}{1 + WD_{ij} Fe_{ij}} \right] \div \left[ N_r T_r + PTO(1 + 2PER_{ij} + 4PER_{ij}^2) T_0 + \left(1 + \log \frac{WD_{ij}}{4}\right) T_r \right], \quad (5)$$

$$N_r = \frac{WD_{ij}}{2} + \frac{3}{2} + (1 - PTO)(1 + WD_{ij} Fe_{ij}), \quad (6)$$

其中  $T_0$  表示初始超时,  $T_r$  为 TCP 包往返时间. 在无线网络中, 丢包率  $PER_{ij}$  主要由无线信道衰落引起, 因此在数值上近似等于包错误率  $Pe_{ij}^{[13]}$ , 同时也是 PHY 层  $BER_{ij}$ , ARQ 等的隐函数.  $WD_{ij}$  表示当前拥塞控制窗口长度, 拥塞控制采用  $WD_{ij}$  的平方根反比增加、正比减小的方式, 如 (7) 式,  $\theta_1, \theta_2$  为拥塞控制因子.

$$WD_{ij} \leftarrow \begin{cases} WD_{ij} + \theta_1, & \text{if } N_n > N_{th}, \\ WD_{ij} - \theta_2 WD_{ij}, & \text{if } N_n \leq N_{th}, \end{cases} \quad (7)$$

式中  $N_n = \frac{2WD_{ij}}{RTT} (1 - Fe_{ij})$  表示当前时隙成功传输帧的个数,  $N_{th}$  为每个时隙传输包的期望数, 本文设为 3. PTO 表示由拥塞导致丢包的概率, 计算式如下

$$PTO = 1 - \sum_{m=1}^{WD_{ij}-3} \frac{WD_{ij}!}{m!(WD_{ij}-m)!} \times (1 - PER_{ij})^{\frac{mWD_{ij}}{2}}, \quad (8)$$

考虑到最大窗口长度为  $cwnd$ , TCP 吞吐率可最终表示为

$$TCP_{ij} = \min(cwnd/T_r, U_{ij}). \quad (9)$$

本文旨在部分感知信息前提下, 每个 CR 节点自主配置跨层参数, 最大化 TCP 吞吐率.

### 3 POMDP 模型

本文将节点  $i$  第  $n$  时隙感知的信道增益作为状态信息, 用  $S_i = G_i = (g_{i1}, \dots, g_{ic}, g_{iC})$  表示, 其中  $g_{ic}$  为第  $c$  ( $1 \leq c \leq C$ ) 个信道增益. 根据第 2 部分模型, 在状态信息基础上, 动作为层间参数的动态配置, 用  $A_i = (Pow_i, Mod_i, X_i, \theta_i)$  表示. 其中  $Pow_i$  为发射功率,  $Mod_i$  为调制模式,  $X_i$  为信道选择,  $\theta_i = \{\theta_1, \theta_2\}$  为拥塞控制因子. 用这些参数在相应信道上传输 TCP 数据包, 此后接收方返回一个传输成功与否的确认消息, 并用  $r(s_i, a_i) = TCP_{ij}$  作为节点  $i$  在状态  $s$  下执行动作  $a$  后的立即回报. 图 2 为每个时隙内链路  $i-j$  上的操作流程图, 其中  $z(s, a, r) \in \bigcup_{k=1}^{n-1} (S_k, A_k, R_k)$  为观测历史.

图 2 说明了在决策  $\pi$  下由状态集  $S$  到动作集  $A$  ( $A = \bigcup A_i$ ) 的映射, 即马尔可夫动态模型, 其目标是使网络在  $N$  个时隙内的平均收益 (传输比特

数) 最大. 但由于节点感知信息存在错误, 导致了 POMDP 的求解模型.

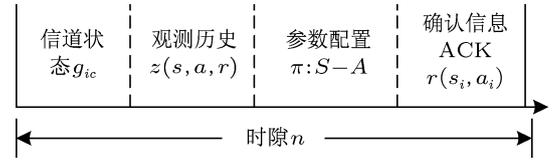


图 2 时隙内的操作

#### 3.1 信念状态

在 POMDP 中, CR 节点须利用现有部分信息、历史动作和立即回报值来进行策略决策, 是非马尔可夫问题, 本文引入信念状态将 POMDP 转化为基于信念状态的 MDP (Belief MDP).

基于观测历史  $Z = \bigcup_z (s, a, r)$ , 定义信念状态  $B(S_n) = \bigcup_{i=1}^L \bigcup_{c=1}^C b_{ic}$ , 其中  $b_{ic}$  为条件概率, 表示在观测历史和转移概率下信道状态为  $g_{ic}$  的概率.  $b(s_{n+1})$  可由贝叶斯公式根据  $b(s_n)$  计算如下<sup>[19]</sup>:

$$b(s_{n+1}) = \eta \cdot o(s_{n+1}, a_n, o_n) \times \sum_{s_n \in S} p(s_n, a_n, s_{n+1}) b(s_n), \quad (10)$$

式中  $\eta$  为正规化常数,  $o(s_{n+1}, a_n, o_n)$  为置信概率, 表示在观测历史基础上, 节点选择动作  $a_n$ , 转移到状态  $s_{n+1}$  和信道状态的分布函数. 由于信念状态在连续空间内取值, 导致 POMDP 求解的维度灾问题<sup>[9]</sup>. 由文献 [9] 知, 本文构建的 POMDP 模型最优值函数可分解为多个分段线性凸函数. 因此本文将信念状态空间分成一个  $K$  段的有限空间, 将  $b(s)$  映射到与其最邻近的值.

由文献 [9] 知, 信念状态为求解最优动作策略  $A^*$  的一个充分统计量, 由此可将 POMDP 问题的求解转化为求解信念状态和策略  $\pi$  的问题, 即求解 BMDP. 如图 3 所示, 模型描述为: 1)  $SE: O \times A \times B(S) \rightarrow B(S)$ , 2)  $\pi: B(S) \rightarrow A$ . 在信念状态  $b$  执行动作  $a$ , 其回报  $r(b, a)$  由  $r(s, a)$  表示为

$$r(b, a) = \sum_{s \in S} b(s) r(s, a). \quad (11)$$

#### 3.2 基于 BMDP 的 Q 学习

Q 学习能有效完成无经验知识的学习, CR 节点的某个动作导致环境正的奖赏, 以后产生这个动作的趋势便会加强, 反之则会减弱. 本文将 Q 学习

映射到 BMDP 的四元组  $\langle B, A, R, SE \rangle$  中, 执行过程如下: 环境状态处于  $s$  的信念状态为  $b(s)$ , 执行动作  $a$  使环境状态转移到  $s'$  的信念值为  $b(s')$ , 同时得到一个奖赏值  $r(b, a)$

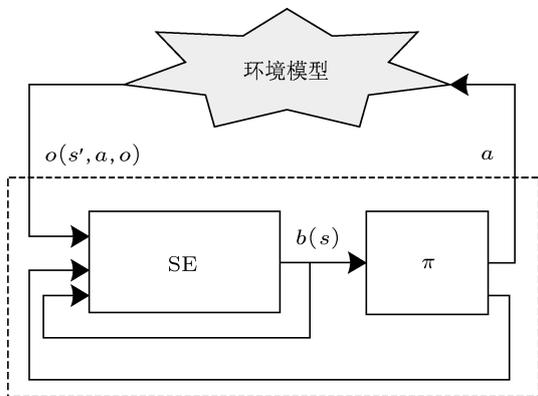


图3 BMDP 模型

在 BMDP 下,  $Q_\pi(b, a)$  定义为在信念状态  $b$  下执行动作  $a$  后, 按照动作策略  $\pi$  选择后续动作, 至最终到达目标期间的累积奖赏期望值.

$$Q_\pi(b, a) = \sum_{s \in S} b(s) E_\pi \times \left[ \sum_{n=0}^{\infty} \gamma^n r^n(s_n, a_n) | s_0 = s, a_0 = a \right], \quad (12)$$

其中  $0 \leq \gamma \leq 1$  为折扣因子, 表示未来回报对当前回报的权重, 将 (11) 式带入 (12) 式可将  $Q$  值改写如下:

$$Q_\pi(b, a) = E \left[ \sum_{n=0}^{\infty} \gamma^n r^n(b_n, a_n) | s_0 = s, a_0 = a \right]. \quad (13)$$

每个动作结束之后, 环境状态转换到下一个状态  $s_{n+1}$ , 对应  $b(s_{n+1})$  由 (10) 式计算, 相应  $Q$  值迭代如下:

$$Q(b, a) \leftarrow Q(b, a) + \alpha(k) \delta(b, a). \quad (14)$$

式中  $\alpha(k)$  为学习速率,  $\delta(b, a)$  为第  $k$  到  $k+1$  步的步长误差, 计算如下:

$$\delta(b, a) = r(b, a) + \gamma \max_{a'} Q(b', a') - Q(b, a), \quad (15)$$

CR 网络中, 对某种特定应用而言, 所有状态-动作对的值函数  $Q_\pi(s, a)$  可由策略  $\pi$  产生的  $Q$  值序列  $\{Q_n(b, a)\}$  表示. 若学习速率  $\alpha_k$  满足条件:  $\sum_{k=1}^{\infty} \alpha_k = \infty, \sum_{k=1}^{\infty} (\alpha_k)^2 < \infty$ , 且  $\{Q_n(b, a)\}$

与最优值  $Q^*(b, a)$  之间最多相差  $\epsilon$ , 则  $\{Q_n(b, a)\}$  能在有限步探索后以  $\frac{1}{2} \gamma \epsilon$  收敛于  $Q^*(b, a)$  (证明见附录 A), 对应的最优策略用  $\pi^*(B)$  表示, 从而得到 BMDP 模型的动作策略

$$\pi^*(b) = \arg \max_a Q^*(b, a). \quad (16)$$

在时隙  $n$ , 为了保证在探索中遍历所有状态-动作对来获得对  $Q$  值的可信估计. 本文采用玻尔兹曼探索策略 [19], 使动作探索与  $Q$  值有效结合. 当前信念状态下访问动作  $a$  的概率为

$$\pi(b, a) = \frac{e^{Q(b, a)/\tau}}{\sum_{a' \in A} e^{Q(b, a')/\tau}}. \quad (17)$$

式中  $\tau$  为温度系数, 该系数控制行为选择的随机程度 [19].

### 3.3 Q-BMDP 结构模型

CR 节点通过实时配置发射功率、调制模式、信道选择、拥塞控制因子来最优化 TCP 吞吐率. 图 4 为  $Q$ -BMDP 结构模型图, 它主要包括无线环境、多个 CR 节点和公共控制信道, 每个节点将 POMDP 转换为 BMDP 来支持  $Q$  学习 ( $Q$ -BMDP). BMDP 通过  $Q$  值迭代找到信念状态对应的动作, 节点执行该动作完成数据包传输, 通过公共控制信道返回一个瞬时回报, 完成  $Q$  值更新, 并参与下一时隙信念状态更新.

## 4 仿真场景设置及结果分析

仿真场景设为  $1000 \text{ m} \times 1000 \text{ m}$ , 如图 5: CR 节点数  $L = 20$ , 组成 10 个源-目的节点传输对. 设源端一直发送数据, 目的端一直接收数据. 每个节点在  $n+1$  时隙的位置随机分布在  $n$  时隙所处位置的  $\delta = 100 \text{ m}$  邻域内. 信道总带宽  $W = 1.25 \text{ MHz}$  ( $W = \bigcup_{i=1}^L W_i$ ), 划分为  $C = 12$  个 OFDM 子信道, CR 节点每次选择  $w_i = 1$  个信道进行传输, 信道模型为 Rayleigh 衰落模型, 噪声功率  $N_0 = 1.1565 \times 10^{-6}$ . 信道间互不相关, 授权用户占用信道的概率为  $p_u = 0.2$ . TCP 包长  $L_{\text{tcp}} = 1500 \text{ bytes}$ , 其最大重传次数  $N_{\text{re}} = 5$ , 最大拥塞窗口  $\text{cwnd} = 6000 \text{ bytes}$ , 拥塞控制因子  $\theta_1 = \frac{3\beta}{2\sqrt{WD_{ij}+1}-\beta}, \theta_2 = \frac{\beta}{\sqrt{WD_{ij}+1}}$ , 本文设  $\beta$  为 1. 初始化超时时长  $T_0 = 2 \text{ s}$ . 数据链

路层为 ARQ 协议, 帧最大重传次数  $N_{fr} = 10$ , 帧头长度  $L_{frh} = 20$  bit, ACK 帧长  $L_{ack} = 24$  bit. PHY 采用 OFDM 信道复用方式, 子带内可选调制方式为 QPSK, 8PSK, 16QAM, 64QAM.  $Q$  学习折扣因子  $\gamma = 0.9$ , 学习速率  $\alpha(k) = 1/(k+1)$ , (17) 式中温度系数在区间 (1, 0.01) 线性递减.

本文采用三种仿真方案作为对比: 1) CR 节点能感知网络的完全知识, 即基于知识的  $Q$ -MDP 算法; 2) CR 节点只能感知部分知识, 即  $Q$ -BMDP 算法; 3) 采用传统固定参数配置方案.

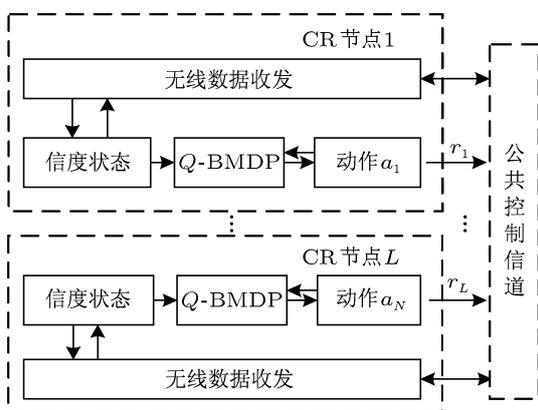


图4 Q-BMDP 结构模型图

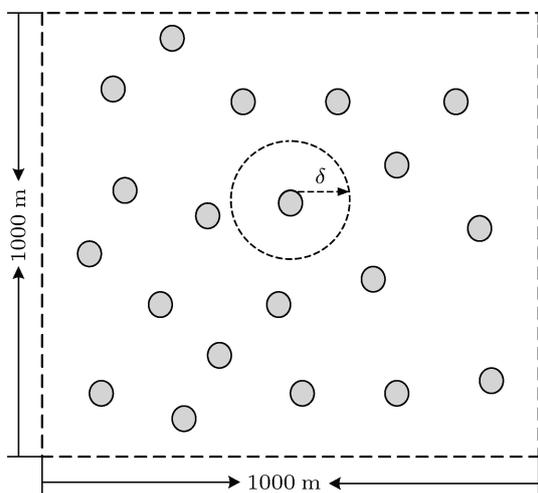


图5 节点分布情况

#### 4.1 功率非受限条件下性能分析

设最大发射功率为 1 mW, SINR 门限为 5.231814 dB, BER 需求是  $10^{-5}$ . 分别采用上述三种方案反复实验, 记录每次实验达到 QoS 的运算次数, 图 6 为其中连续 50 次实验过程中的运算次数比较. 从图知, 固定参数配置方法由于缺乏与

环境的交互, 收敛性最差, 由此可能使网络性能不稳定; 当环境知识完全可知时, 节点可以进行最优决策, 从图中可以看出  $Q$ -MDP 具有最好的收敛性, 能在 20 次运算内找到最优参数配置; 本文提出的  $Q$ -BMDP 算法非常接近  $Q$ -MDP, 也能在 20 次运算内收敛到近似最优状态, 与  $Q$ -MDP 比较只有 1.5%

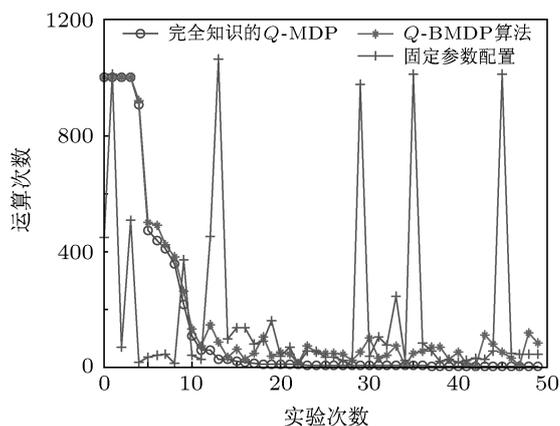


图6 运算次数比较

左右的误差. 图 7 是在相同条件下, 网络总 TCP 吞吐率变化情况. 由图知,  $Q$ -MDP 算法的 TCP 平均吞吐率最高, 而  $Q$ -BMDP 算法的 TCP 吞吐率非常接近  $Q$ -MDP, 二者性能只相差约 2% 左右; 由于固定参数配置方案没有考虑环境的变化, 其吞吐率最低, 仅为  $Q$ -BMDP 的 59% 左右.

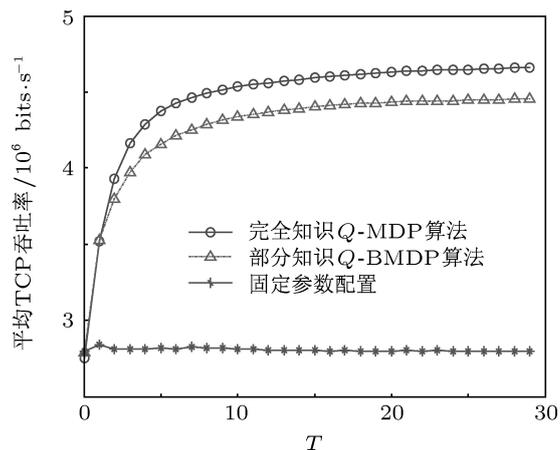


图7 网络总 TCP 吞吐率比较

#### 4.2 功率受限条件下性能分析

设定网络总功率  $P$ , 参数配置中增加功率受限条件 1)  $\sum_{i=1}^L Pow_i \leq P, Pow_i \geq 0, 1 \leq i \leq L$ , 2)  $Pow_i \leq P_r, P_r$  为节点允许的最大发射功率, 其他

仿真参数与 4.1 相同. 图 8 为将发射功率归一化后三种算法在不同 SINR 情况下的 TCP 吞吐量对比关系.

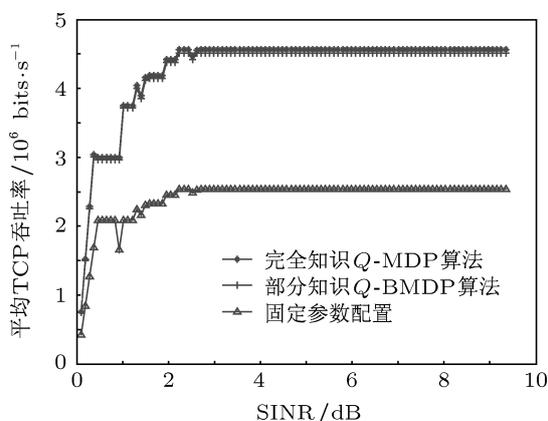


图 8 功率受限下的吞吐量比较

由图 8 知, 当 SINR 小于 0.7 时, 由于受环境限制, 三者性能相差不大, TCP 吞吐量较低; 当 SINR 从 0.7 增加到 2.0 左右时, 固定参数配置方法与另外两种算法的性能差异越来越明显; 当 SINR 大于 2.0 以后, 吞吐量差异最大, 且趋于稳定. 由于 Q-MDP 对环境完全可知, 能够在功率受限情况下进行最优决策, 其吞吐量最大; Q-BMDP 能够在部分感知情况下进行近似最优决策, 其 TCP 性能与 Q-MDP 性能非常接近, 二者相差约 1% 左右; 固定参数配置方案的吞吐量最差, 仅为 Q-BMDP 的 56% 左右.

### 4.3 模型适应性分析

设无线信道干扰噪声  $N_0$  分别为 (a)  $1.1565 \times 10^{-6}$ , (b)  $1.1565 \times 10^{-5}$ , (c)  $1.1565 \times 10^{-4}$  和 (d)  $1.1565 \times 10^{-3}$ , 在与文献 [11] 相同参数条件下, 对 Q-BMDP 所采用模型与文献 [11] 所采用模型进行仿真对比分析. 图 9 为不同噪声条件下两模型对应 TCP 吞吐量随时隙的变化情况, 图中横坐标代表时隙  $T$  ( $\times 10$  slots), 纵坐标代表 TCP 吞吐量 (bit/s), 子图 (a), (b), (c), (d) 分别为四种噪声条件对应的 TCP 吞吐量. 当  $N_0 = 1.1565 \times 10^{-6}$  时, 文献 [11] 中模型和 Q-BMDP 所采用模型分别在第 30 和第 17 个时隙获得较稳定的吞吐量; 当  $N_0 = 1.1565 \times 10^{-5}$  时, 文献 [11] 中模型和 Q-BMDP 所采用模型分别在第 40 和第 22 个时隙获得较稳定的吞吐量. 从 (a), (b) 可知, 噪声功率较小时, Q-BMDP 所采用模型比文献 [11] 所采用模型能更快达到稳定状态, 但在稳定状态下 TCP 吞吐量相差不大. 当  $N_0 = 1.1565 \times 10^{-4}$  时, Q-MDP 所采用模型大概在第 30 个时隙后获得较稳定的吞吐量, 而文献 [11] 对应模型在这段时隙内只有 Q-BMDP 对应模型吞吐率的 75% 左右, 并且要在第 53 个时隙后才能获得与 Q-BMDP 对应模型相近的吞吐量; 当  $N_0$  为  $1.1565 \times 10^{-3}$  时, 由于噪声功率较大, TCP 吞吐量在启动阶段波动较大, 但在第 50 个时隙后 Q-BMDP 对应模型吞吐量比文献 [11] 对应模型吞吐量高出 30% 左右. 由此可知, 在无线信道噪声功率较低时,

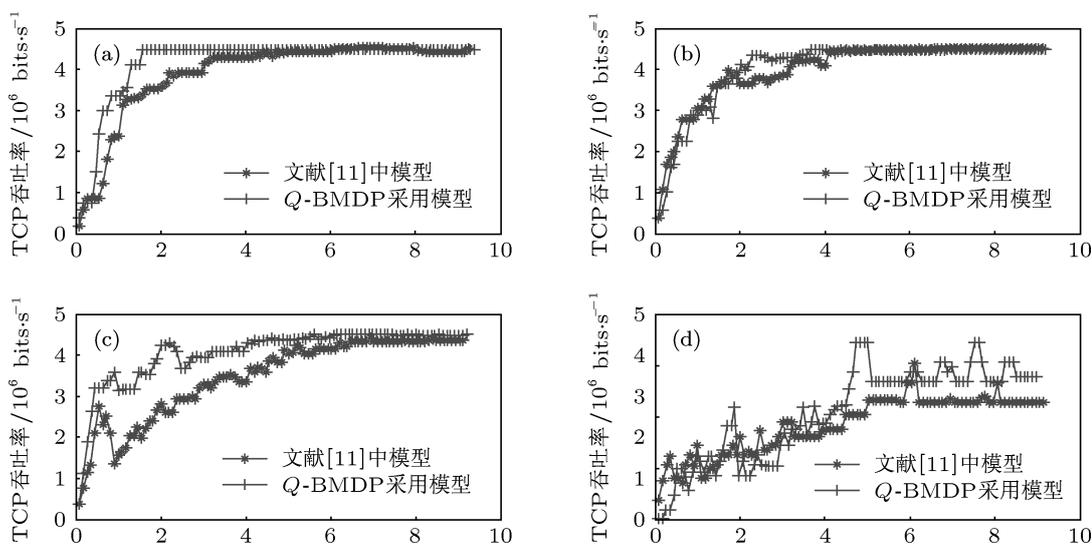


图 9 不同噪声下两模型 TCP 吞吐量比较 (a)  $N_0 = 1.1565 \times 10^{-6}$ ; (b)  $N_0 = 1.1565 \times 10^{-5}$ ; (c)  $N_0 = 1.1565 \times 10^{-4}$ ; (d)  $N_0 = 1.1565 \times 10^{-3}$

两个模型稳定状态对应的 TCP 吞吐率相差不大,但随着噪声功率的增加,  $Q$ -BMDP 所采用模型的 TCP 性能明显优于文献 [11] 中的模型. 仿真说明本文  $Q$ -BMDP 对应模型更适用于无线环境.

### 5 结论

当前 TCP 吞吐率研究大都基于完全知识对网络建模, 没有考虑到感知误差, 使其应用受到很大限制. 针对认知无线电环境不完全可知情况下的 TCP 吞吐率优化问题, 本文联合考虑物理层、链路层与 TCP 层协议, 将网络模型建模为部分可观测马尔可夫决策过程, 提出了基于  $Q$  值迭代的  $Q$ -BMDP 认知无线网络参数配置方案. 仿真分析表明, 与固定参数配置方案相比, 本文的  $Q$ -BMDP 算法能有效提高 TCP 吞吐率; 与完全知识的  $Q$ -MDP 算法相比, 两者性能非常接近.

### 附录 A

**命题** 若对于  $\forall b$ ,  $Q$ -BMDP 近似算法的  $Q(b, a)$  与实际最优值  $Q^*(b, a)$  最多相差  $\varepsilon$ , 则近似最优值  $\tilde{Q}(b, a)$  与  $Q^*(b, a)$  最多相差  $\frac{1}{2}\gamma\varepsilon$ .

**证明** 由条件知,  $Q^*(b, a) - \varepsilon \leq Q(b, a) \leq Q^*(b, a) + \varepsilon$ . 用  $a^* \in A^*$  表示实际最优  $Q$  值对应的动作策略,  $\tilde{a} \in \tilde{A}$  表示  $Q$ -BMDP 近似算法最优  $Q$  值对应的动作策略. 根据 (14), (15) 式得:

$$\begin{aligned} \tilde{Q}(b, a) &= [1 - \alpha(k)]Q(b, \tilde{a}) \\ &\quad + \alpha(k) \left[ r(b, \tilde{a}) + \gamma \max_{a'} Q(b', a') \right] \\ &\geq [1 - \alpha(k)]Q(b, \tilde{a}) \end{aligned} \quad (18)$$

$$+ \alpha(k) \left\{ r(b, \tilde{a}) + \gamma \max_{a'} [Q(b', a') - \varepsilon] \right\}. \quad (19)$$

由于  $\tilde{a} \in \tilde{A}$  为近似算法最优  $Q$  值对应的动作, 则在  $Q$ -BMDP 中, 对于  $a^* \in A^*$ ,  $Q(b, \tilde{a}) \geq Q(b, a^*)$ , (19) 式可改写为

$$\begin{aligned} \tilde{Q}(b, a) &\geq [1 - \alpha(k)]Q(b, a^*) \\ &\quad + \alpha(k) \left[ r(b, a^*) + \gamma \max_{a'} Q(b', a') \right] \\ &\quad - \alpha(k)\gamma\varepsilon. \end{aligned} \quad (20)$$

因为  $\alpha(k) = \frac{1}{k+1} \leq \frac{1}{2}$ , 所以

$$\begin{aligned} \tilde{Q}(b, a) &\geq Q^*(b, a) - \alpha(k)\gamma\varepsilon \\ &\geq Q^*(b, a) - \frac{1}{2}\gamma\varepsilon \end{aligned} \quad (21)$$

另外, 同样由 (14), (15) 得:

$$\begin{aligned} \tilde{Q}(b, a) &= [1 - \alpha(k)]\tilde{Q}(b, a) \\ &\quad + \alpha(k) \left[ \tilde{r}(b, a) + \gamma \max_{a'} \tilde{Q}(b', a') \right] \end{aligned} \quad (22)$$

$$\begin{aligned} &\leq [1 - \alpha(k)]\tilde{Q}(b, a) \\ &\quad + \alpha(k) \left\{ \tilde{r}(b, a) + \gamma \max_{a'} [Q^*(b', a') + \varepsilon] \right\} \end{aligned} \quad (23)$$

$$\begin{aligned} &\leq [1 - \alpha(k)]\tilde{Q}(b, a) \\ &\quad + \alpha(k) \left[ \tilde{r}(b, a) + \gamma \max_{a'} Q(b', a') \right] \\ &\quad + \alpha(k)\gamma\varepsilon \end{aligned} \quad (24)$$

$$\leq Q^*(b, a) + \alpha(k)\gamma\varepsilon \leq Q^*(b, a) + \frac{1}{2}\gamma\varepsilon \quad (25)$$

命题得证, 因此, 对于  $\forall b$ ,  $Q$ -BMDP 的近似最优值  $\tilde{Q}(b, a)$  与实际最优值  $Q^*(b, a)$  最多相差  $\frac{1}{2}\gamma\varepsilon$ .

[1] Wang B B, Liu K, Ray J 2011 *IEEE J. Sel. Top. Sign. Process.* **5** 5  
 [2] Zu Y X, Zhou J 2011 *Acta Phys. Sin.* **60** 079501 (in Chinese) [祖云霄, 周杰 2011 物理学报 **60** 079501]  
 [3] Zhao Z J, Xu S Y, Zheng S L, Yang X N 2009 *Acta Phys. Sin.* **58** 5118 (in Chinese) [赵知劲, 徐世宇, 郑世链, 杨小牛 2009 物理学报 **58** 5118]  
 [4] Mwangoka J W, Letaief K B, Cao Z G 2009 *PHYCOM* **2** 103  
 [5] Shiang H P, van der Schaar M 2007 *IEEE J. Sel. Areas Commun.* **25** 770  
 [6] Fu F W, van der Schaar M 2010 *IEEE Trans. Signal Process.* **58** 1401  
 [7] Zhang Y, Fu F W, van der Schaar M 2010 *IEEE Trans. Signal Process.* **58** 3108  
 [8] Lin Z C, van der Schaar M 2011 *IEEE Trans. Wireless Commun.* **1** 102  
 [9] Zhao Q, Lang T, Ananthram, Chen Y X 2007 *IEEE J. Sel. Areas Commun.* **25** 589  
 [10] Wang Y C, Ren P Y, Su Z 2011 *IEICE Trans. Commun.* **E94-B** 1621  
 [11] Lou C Q, Yu F R, Ji H, Leung V C M 2010 *IEEE Trans. Veh. Technol.* **59** 2485  
 [12] Wei Y F, Yu F R, Song M, Zhang Y 2011 *IET Commun.* **5** 2257  
 [13] Luo C Q, Yu F R, Ji H, Leung V C M 2011 *Wirel. Netw.* **17** 479

- [14] Kwon W Y, Suh I H, Lee S 2011 *Int. J. Control, Automat. Syst.* **9** 328–47 181
- [15] Maskery M, Krishnamurthy V, Zhao Q 2009 *IEEE Trans. Commun.* **57** 459
- [16] Krishnaswamy D 2002 *Proceedings of the 56th IEEE Conference on Vehicular Technology* Vancouver, BC, Canada, September 24–28, 2002 p1312
- [17] Lu J, Letaief K B, Chuang J C I, Liou M L 1999 *IEEE Trans. Commun.* **47** 181
- [18] Parvez N, Mahanti A, Williamson C 2010 *IEEE/ACM Trans. Networking* **18** 448
- [19] Wiering M A, Kooi T 2010 *Proceedings of the IEEE International Joint Conference on Neural Networks* Barcelona, Spain, July 18–23, 2010 p1

# Crosslayer parameter configuration for TCP throughput improvement in cognitive radio networks\*

Jiang Hong<sup>†</sup> Liu Cong-Bin Wu Chun

(School of Information Engineering, Southwest University of Science and Technology, Mianyang 621010, China)

(Received 9 April 2012; revised manuscript received 24 June 2012)

## Abstract

In cognitive radio network (CRN), TCP end to end throughput is one of the key issues to measure its performance. However, most of existing research efforts devoted to TCP performance improvement have two weaknesses as follows. First, most of them only consider the underlying parameters to optimize the physical performance, but the TCP performance is neglected. Second, they are largely formulated as a Markov decision process (MDP), which requires a complete knowledge of network and cannot be directly applied to CRNs. To solve the above problems, a  $Q$ -BMDP algorithm is proposed in this paper. Each user in CRN combines modulation type and transmitting power at the physical layer, access channels at the media access control layer and TCP congestion control factor to maximize the TCP throughput. Due to the existence of perception error of environment, this issue is formulated as a partial observable Markov decision process (POMDP) which is then converted to belief state MDP, with  $Q$ -value iteration to find the approximately optimal strategy. Simulation and analysis results show that the proposed algorithm can be approximately converged to optimal strategy under a maximum error limit, and can effectively improve TCP throughput in a dynamic wireless network under the premise of the limited power consumption.

**Keywords:** cognitive radio networks, TCP throughput, belief Markov decision process,  $Q$ -learning

**PACS:** 88.80.hp

**DOI:** 10.7498/aps.62.038804

\* Project supported by the National Natural Science Foundation of China (Grant No. 61072138) and the National Defense Fundamental Scientific Research Program, China (Grant No. B3120110005).

<sup>†</sup> Corresponding author. E-mail: jianghong@swust.edu.cn