

# 一种传输时限下认知无线电网络的动态广播策略



房婷<sup>1</sup> 宫傲宇<sup>1</sup> 张帆<sup>1</sup> 林艳<sup>1</sup> 贾林琼<sup>1,2</sup> 张一晋<sup>1</sup>

<sup>1</sup> 南京理工大学电子工程与光电技术学院 南京 210094

<sup>2</sup> 东南大学移动通信国家重点实验室 南京 210018

(ting.fang@njust.edu.cn)

**摘要** 在具有传输时限要求的认知无线网络中,次用户需要机会式地在给定传输时限内使用无主用户占用的信道广播消息。针对此场景,文中提出一种新的传输时限下认知无线电网络的动态广播策略,允许各等待发送数据的次用户根据每个时隙载波侦听的观测、传输时限剩余时间和主用户占用信道模型实时调整发送概率。首先基于马尔可夫决策过程获得一种载波侦听理想观测假设下的最优策略和最大网络可靠性;然后据此提出一种适用于载波侦听实际观测能力的启发式策略,并通过马尔可夫决策过程建模获得此启发式策略的网络可靠性。仿真结果验证了理论分析的准确性,同时表明所提出的启发式策略的可靠性非常接近理想观测下的最大可靠性,并且明显优于最优静态策略的网络可靠性。

**关键词:** 认知无线电;传输时限;动态广播策略;马尔可夫决策过程;可靠性

**中图分类号** TN911

## Dynamic Broadcasting Strategy in Cognitive Radio Networks Under Delivery Deadline

FANG Ting<sup>1</sup>, GONG Ao-yu<sup>1</sup>, ZHANG Fan<sup>1</sup>, LIN Yan<sup>1</sup>, JIA Lin-qiong<sup>1,2</sup> and ZHANG Yi-jin<sup>1</sup>

<sup>1</sup> School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China

<sup>2</sup> National Mobile Communications Research Laboratory, Southeast University, Nanjing 210018, China

**Abstract** In cognitive radio networks with delivery deadline requirements, secondary users need to opportunistically broadcast messages using the channel unoccupied by primary users (PUs) within a given delivery deadline. For this scenario, this paper proposes a new cognitive radio network dynamic broadcasting strategy under delivery deadline, which allows each secondary user (SU) to adjust the transmission probability according to the observation of carrier sensing in each slot, the remaining time before deadline expiration and the PU traffic model. Based on an ideal assumption on the observation of carrier sensing, this paper obtains an optimal broadcasting strategy and the maximum network reliability using Markov decision process (MDP). Then, according to a practical observation capability of carrier sensing, this paper proposes a heuristic broadcasting strategy and uses another MDP to obtain the network reliability of this heuristic strategy. Simulation results verify the accuracy of the analysis and show that the network reliability of the proposed heuristic strategy is very close to the maximum network reliability under the ideal observation, and is superior to that of the optimal statics strategy.

**Keywords** Cognitive radio, Delivery deadline, Dynamic broadcasting strategy, Markov decision process, Reliability

### 1 概述

近年来,随着无线终端设备的大量普及和无线网络业务的迅猛发展,加上无线频谱管理规则的限制,无线频谱资源变得日益拥挤,成为了制约无线网络发展的瓶颈。认知无线电

网络<sup>[1-2]</sup> (Cognitive Radio Network, CRN) 允许非授权用户(亦称为次用户)机会式地接入未被授权用户(亦称为主用户)占用的频谱,能够有效改善频谱利用率,进而缓解频谱资源分配紧张的局面,因此得到研究人员的广泛关注。

尽管如此,工业控制、远程医疗、环境监测等应用中具有

收稿日期:2020-09-01 返修日期:2021-01-09 本文已加入开放科学计划(OSID),请扫描上方二维码获取补充信息。

基金项目:国家自然科学基金(62071236, 62001225);中央高校基本科研业务经费(30920021127, 30919011227);江苏省自然科学基金(BK20190454);东南大学移动通信国家重点实验室开放研究基金资助课题(2020D19)

This work was supported by the National Natural Science Foundation of China(62071236,62001225), Fundamental Research Funds for the Central Universities of China(30920021127,30919011227), Natural Science Foundation of Jiangsu Province(BK20190454) and Open Research Fund of National Mobile Communications Research Laboratory, Southeast University(2020D19).

通信作者:张一晋(yijin.zhang@gmail.com)

低时延、高可靠性要求的数据业务<sup>[3-4]</sup>仍然为 CRN 次用户接入协议的设计和优化带来巨大的挑战,需要次用户在规定时间内以高可靠性发送数据,否则将会影响服务质量,甚至造成经济损失或人员伤亡。

针对传输时限下基于固定发送概率的次用户静态接入策略已有一系列研究工作。在传输时限同步假设下,文献[5]分析了 ALOHA 机制的渐进性能。在传输时限异步假设下,文献[6]推导了单次传输时 ALOHA 机制的饱和状态下最高可达可靠性以及对应的最优接入概率。基于此工作,文献[7]推导了 ALOHA 机制的饱和状态下最优重传次数,文献[8]推导了排队系统下 ALOHA 机制的最高可达可靠性以及对应的最优接入概率,文献[9]推导了 ALOHA 机制在射频能量搜集场景下的可靠性。文献[10]则将文献[5]的工作拓展至物理层,即能够使用多用户检测技术同时成功接收多个数据的情形。其中仅文献[5,9]考虑了 CRN 对接入概率的影响。然而,文献[5]仅使用信道被主用户占用的稳态概率来刻画此影响,因此无法充分利用信道状态的变化规律来改善可靠性,而文献[9]则仅考虑了单个次用户的场景,因此缺乏对次用户之间竞争的分析。

另一方面,CRN 次用户动态接入策略<sup>[11]</sup>也已得到广泛研究。代表性工作是文献[12]使用部分观测马尔可夫决策过程(Partially Observable Markov Decision Process, POMDP)对 CRN 的次用户机会式信道接入进行建模,使得次用户根据信道观测做出最优的信道感知和接入选择。受文献[12]的启发,文献[13-15]基于各种约束下的吞吐量、可靠性等一系列优化目标对 CRN 动态接入策略做出了进一步研究,但文献[12-15]均未考虑传输时限的影响。文献[16]探索了次用户发送功率、主次用户干扰、传输时限等约束下的吞吐量最大化问题,但只是假设存在中央调度器根据各次用户对信道的观测进行汇总,并集中式地为各次用户选择信道和设置发送功率,并未考虑各节点仅根据自身观测做出相应决策的分布式情况。另外,文献[17]基于时隙分配、载波分配和功率分配设计了 CRN 射频能量搜集场景下的次用户动态协作接入机制,但仅考虑了传输时限对主用户而非对次用户的影响。

综上所述,目前尚无针对传输时限下认知无线网络的次用户分布式动态广播策略研究。鉴于此,本文提出了一种新的传输时限下的 CRN 动态广播策略,所提策略允许各次用户根据载波侦听的观测、传输时限剩余时间和主用户占用信道模型实时调整发送概率。本文首先基于马尔可夫决策过程(Markov Decision Process, MDP)获得了一种载波侦听理想观测假设下的最优策略和最大网络可靠性,然后据此提出了一种适用于载波侦听实际观测能力的启发式策略,并使用 MDP 获得了此启发式策略的网络可靠性。仿真结果验证了理论分析的准确性,同时表明所提启发式策略的可靠性非常接近于理想观测下的最大可靠性,并且优于其他策略。

## 2 系统模型

如图 1 所示,本节将首先介绍传输时限下的 CRN 模型以及次用户载波侦听观测能力,然后阐述基于此的广播策略设定。

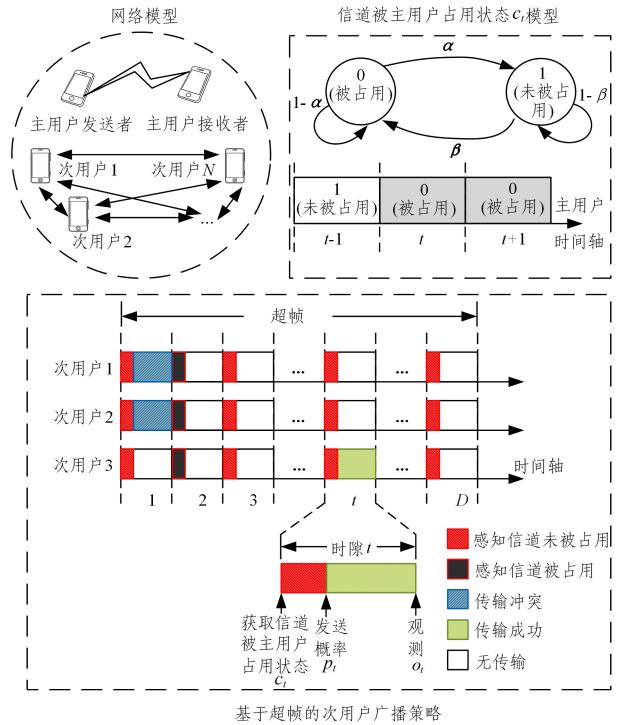


图 1 系统模型图

Fig. 1 System model

### 2.1 传输时限下的 CRN 模型

考虑具有单一时隙化无线信道的 CRN。时隙  $t$  的信道被主用户占用的状态为  $c_t \in \{0, 1\}$ , 0 表示信道在时隙  $t$  被主用户占用, 1 表示信道在时隙  $t$  未被主用户占用。假设此占用状态在时隙边界处的状态转移符合马尔可夫性质, 设  $c_t = 0$  至  $c_{t+1} = 1$  的转移概率为  $\alpha$ ,  $c_t = 1$  至  $c_{t+1} = 0$  的转移概率为  $\beta$ 。

此 CRN 包含  $N \geq 2$  个次用户。各次用户以不断重复的固定长度超帧为传输架构, 在每个超帧开始时刻产生一个数据分组, 需要在每个超帧内将相应产生的数据分组广播至其他次用户, 并且由于广播特性, 不使用应答和重传机制。每个超帧的组成时隙数等于传输时限  $D (D \geq 1)$ 。

假设不存在中央调度器负责调度各次用户传输, 则在任一超帧每个时隙的开始时刻, 每个待发送次用户首先侦听信道, 若侦听到信道未被主用户占用, 则以一定概率在当前时隙内完成发送数据分组, 否则保持静默(即设定发送概率为 0)。

假设一个数据分组当且仅当其独自在一个时隙内进行传输时才会被成功接收。网络可靠性定义为成功接收的数据分组数量占全部产生的数据分组数量的比例。

### 2.2 载波侦听的实际观测能力

根据载波侦听的实际观测能力, 定义任一超帧时隙  $t$  未通信次用户对次用户通信状况的载波侦听观测  $O_t \in \{0$  (未有次用户通信),  $1$  (次用户传输成功),  $*$  (次用户传输冲突)}。这里,  $O_t = 0$  表明时隙  $t+1$  开始时刻的待发送次用户数相比于时隙  $t$  开始时刻未有变化,  $O_t = 1$  表明时隙  $t+1$  开始时刻的待发送次用户数相比于时隙  $t$  开始时刻减少 1, 而  $O_t = *$  表明时隙  $t+1$  开始时刻的待发送次用户数相比于时隙  $t$  开始时刻至少减少 2。

### 2.3 广播策略设定

显而易见,每个待发送次用户需要在不同时刻根据传输的紧迫性和网络的竞争程度调整发送概率。鉴于此,在任一超帧每个时隙  $t$  的开始时刻,每个待发送次用户根据超帧已进行时隙数  $t$ ,时隙  $t$  之前信道被主用户占用状态的历史  $c_1, c_2, \dots, c_t$ ,时隙  $t$  之前采用的发送概率历史以及时隙  $t$  之前对次用户通信状况的观测历史  $O_1, O_2, \dots, O_{t-1}$ ,使用相同的发送概率。因此,可以将时隙  $t$  的发送概率表示为:

$$p_t(c_1, p_1(c_1), O_1, \dots, c_{t-1}, p_{t-1}(c_1, p_1(c_1, O_0), \dots, O_{t-2}, c_{t-1}), O_{t-1}, c_t) \quad (1)$$

本文以最大化网络可靠性为目标,对任一超帧内每个时隙  $t$  的发送概率进行优化。

传输时限  $D$  与次用户数  $N$  的关系需要符合应用场景要求,因此不同场景具有不同的约束关系。鉴于此,本文的策略设计、分析和优化将适用于各种  $D$  和  $N$  的约束关系,并将仿真验证一系列  $D$  和  $N$  取值下的策略性能。

### 3 理想观测下的最优广播策略

如第 2.2 节所述,当次用户传输发生冲突时,此事件尽管可以被各待发送次用户观测,但其对于待发送次用户数的影响无法被确定性获知。因此,需要应用 POMDP 对信道未被主用户占用时的一系列各待发送次用户的发送概率进行优化,但 POMDP 的求解也要求较高的计算复杂度。为了简化分析,本节理想化地假设各待发送次用户能够根据历史载波侦听观测信息,准确获知任一时刻开始时刻的待发送次用户数量,即理想化地假设各待发送次用户不仅能够准确获知之前每次成功传输均仅涉及一个次用户,而且能够准确获知之前每次冲突传输涉及的次用户数量,从而准确获知待发送次用户数量的变化。因此,将使用 MDP 替代 POMDP 进行协议建模,获得理想观测下的最优广播策略以及最大网络可靠性。

基于以上理想观测,可以设定待发送次用户在时隙  $t$  的发送概率仅由信道在时隙  $t$  被主用户占用状态和时隙  $t$  开始时刻待发送次用户数量  $n_t$  决定,因此可以将发送概率(见式(1))简化地表示为:

$$p_t(c_t, n_t) \quad (2)$$

显然, $c_t=0$  或  $n_t=0$  时, $p_t(c_t, n_t)=0$ ;  $c_t=1$  且  $n_t>1$  时, $0<p_t(c_t, n_t)<1$ 。另外,为了使得数据尽早被成功发送, $c_t=n_t=1$  时,设定  $p_t(c_t, n_t)=1$ 。

#### 3.1 理想观测下的 MDP 建模

给定次用户总数  $N$ 、传输时限  $D$  和信道被主用户占用状态的马尔可夫链参数  $\alpha, \beta$ ,并且根据各待发送次用户能够准确获知待发送次用户数量这一理想观测假设对第 2 节所述广播策略进行 MDP 建模,并对其状态、动作、决策规则、状态转移概率、期望奖励以及期望总奖励进行定义。

状态:定义状态空间  $S$  为:

$$S \triangleq \{(c, n), c \in \{0, 1\}, n \in \{0, 1, \dots, N\}\} \quad (3)$$

定义时隙  $t$  的状态  $s_t$  为:

$$s_t \triangleq (c_t, n_t) \in S \quad (4)$$

其中, $c_t$  为时隙  $t$  的信道被主用户占用状态, $n_t$  为时隙  $t$  开始

时刻待发次用户数。

动作:定义动作空间  $A$  为:

$$A \triangleq \{0, 1, \dots, N\} \quad (5)$$

其中, $a \in A$  表示  $a$  个次用户在时隙  $t$  发送数据。

决策规则:在时隙  $t$ ,当  $c_t=0$  或  $n_t=0$  时, $p_t(c_t, n_t)=0$ ,因此动作  $a=0$  在时隙  $t$  被确定性执行;在时隙  $t$ ,当  $c_t=n_t=1$  时,信道未被主用户占用,唯一待发送次用户以概率  $p_t(c_t, n_t)=1$  发送数据,因此动作  $a=1$  在时隙  $t$  被确定性执行;在时隙  $t$ ,当  $c_t=1$  以及  $n_t>1$  时,信道未被主用户占用, $n_t$  个待发送次用户均以概率  $0<p_t(1, n_t)<1$  发送数据,因此任一动作  $a \in \{1, 2, \dots, n_t\}$  按概率:

$$\binom{n_t}{a} p_t(1, n_t)^a (1 - p_t(1, n_t))^{n_t - a} \quad (6)$$

被执行,而使得其他  $a \in \{0, n_t + 1, n_t + 2, \dots, N\}$  不会被执行。

状态转移概率:将状态  $s_t$  转移至状态  $s_{t+1}$  的概率表示为  $\Pr(s_{t+1} | s_t)$ 。在时隙  $t$ ,当  $c_t=0$  时,对于任一  $n_t \in \{0, 1, \dots, N\}$ ,容易得到:

$$\Pr((c_{t+1}=0, n_{t+1}=n_t) | (c_t=0, n_t)) = 1 - \alpha \quad (7)$$

$$\Pr((c_{t+1}=1, n_{t+1}=n_t) | (c_t=0, n_t)) = \alpha \quad (8)$$

在时隙  $t$ ,当  $c_t=1$  且  $n_t \in \{0, 1\}$  时,容易得到:

$$\Pr((c_{t+1}=0, n_{t+1}=0) | (c_t=1, n_t)) = \beta \quad (9)$$

$$\Pr((c_{t+1}=1, n_{t+1}=0) | (c_t=1, n_t)) = 1 - \beta \quad (10)$$

在时隙  $t$ ,当  $c_t=1$  时,对于任一  $n_t \in \{2, 3, \dots, N\}$  和任一  $a \in \{1, 2, \dots, n_t\}$ ,可以得到:

$$\Pr((c_{t+1}=0, n_{t+1}=n_t - a) | (c_t=1, n_t)) = \binom{n_t}{a} p_t(1, n_t)^a (1 - p_t(1, n_t))^{n_t - a} \beta \quad (11)$$

$$\Pr((c_{t+1}=1, n_{t+1}=n_t - a) | (c_t=1, n_t)) = \binom{n_t}{a} p_t(1, n_t)^a (1 - p_t(1, n_t))^{n_t - a} (1 - \beta) \quad (12)$$

除以上情况, $\Pr(s_{t+1} | s_t) = 0$ 。

期望奖励:在时隙  $t$ ,定义期望奖励  $r_t(s_t)$  为成功传输的数据分组数量的期望值,即:

$$r_t(s_t) \triangleq \begin{cases} n_t p_t(1, n_t) (1 - p_t(1, n_t))^{n_t - 1}, & \text{if } c_t=1, n_t > 1 \\ 1, & \text{if } c_t=n_t=1 \\ 0, & \text{otherwise} \end{cases} \quad (13)$$

期望总奖励:定义期望总奖励  $R(s_1)$  为初始状态  $s_1$  条件下,从时隙 1 至时隙  $D$  期间成功传输的数据分组总数的期望,即:

$$R(s_1) \triangleq \mathbb{E}_{s_1} \left\{ \sum_{t=1}^D r_t(s_t) \right\} \quad (14)$$

#### 3.2 理想观测下的策略优化

最大化 3.1 节所构建 MDP 的期望总奖励等价于理想观测下最大化网络可靠性。因此本节将探索最大期望总奖励以及相应的最优策略。

定义值函数  $u_t(s_t)$  为初始状态  $s_t$  条件下,从时隙  $t$  至时隙  $D$  期间成功传输的数据分组数量的期望,即:

$$u_t(s_t) \triangleq \mathbb{E}_{s_t} \left\{ \sum_{\tau=t}^D r_\tau(s_\tau) \right\} \quad (15)$$

可见, $R(s_1) = u_1(s_1)$ 。进一步,根据贝尔曼方程可以将时隙  $t$

的值函数写为时隙  $t$  的即时奖励和时隙  $t+1$  的值函数期望值之和,即:

$$u_t(s_t) = r_t(s_t) + \sum_{s_{t+1} \in S} \Pr(s_{t+1} | s_t) u_{t+1}(s_{t+1}) \quad (16)$$

因此,获得各时隙  $t$  值函数的具体步骤如算法 1 所示。

#### 算法 1 值函数 $u_t(s_t)$ 求解过程

1. 设置  $t=D$ , 对于任一  $s_D \in S$ , 获得:

$$u_D(s_D) = r_D(s_D) \quad (17)$$

2. 如果  $t=1$ , 停止迭代; 否则进入步骤 3。

3. 令  $t=t-1$ , 对于任一  $s_t \in S$ , 获得:

$$u_t(s_t) = r_t(s_t) + \sum_{s_{t+1} \in S} \Pr(s_{t+1} | s_t) u_{t+1}(s_{t+1}) \quad (18)$$

4. 返回步骤 2。

进一步,可以运用动态规划方法来获得最优广播策略的最大化期望总奖励,即寻求所有  $1 \leq t \leq D, 1 < n_t \leq N$  情况下  $p_t(1, n_t)$  的最优值,  $p_t^*(1, n_t)$  以按时隙  $D$  至 1 的顺序最大化上述每一个值函数,从而最大化期望总奖励。具体步骤如算法 2 所示。

#### 算法 2 策略优化

1. 设置  $t=D$ , 对于任一  $s_D \in S$ , 如果  $c_D=0$  或  $n_D=0$ , 获得  $u_D(s_D)$  的最大值  $u_D^{\max}(s_D)=0$ ; 如果  $c_D=n_D=1$ , 获得  $u_D^{\max}(s_D)=1$ ; 否则计算最优发送概率:

$$p_D^*(c_D, n_D) = \arg \max_{p_D(1, n_D) \in (0, 1)} u_D(s_D) \quad (19)$$

并将其代入算法 1 获得  $u_D^{\max}(s_D)$ 。

2. 如果  $t=1$ , 则停止迭代; 否则进入步骤 3。

3. 令  $t=t-1$ , 对于任一  $s_t \in S$ , 如果  $c_t=0$  或  $n_t=0$ , 获得  $u_t(s_t)$  的最大值  $u_t^{\max}(s_t)$  为:

$$u_t^{\max}(s_t) = \sum_{s_{t+1} \in S} \Pr(s_{t+1} | s_t) u_{t+1}^{\max}(s_{t+1}) \quad (20)$$

如果  $c_t=n_t=1$ , 获得  $u_t^{\max}(s_t)$  为:

$$u_t^{\max}(s_t) = 1 + \sum_{s_{t+1} \in S} \Pr(s_{t+1} | s_t) u_{t+1}^{\max}(s_{t+1}) \quad (21)$$

否则计算最优发送概率:

$$p_t^*(1, n_t) = \arg \max_{p_t(1, n_t) \in (0, 1)} \{r_t(s_t) + \sum_{s_{t+1} \in S} \Pr(s_{t+1} | s_t) u_{t+1}^{\max}(s_{t+1})\} \quad (22)$$

并将其代入算法 1 获得  $u_t^{\max}(s_t)$ 。

4. 返回步骤 2。

图 2 为  $N=3, D=10, c_1=0, \alpha=0.7, \beta=0.3$  或  $\beta=0.1$  时, 算法 1 求解所有  $1 \leq t \leq D, 1 < n_t \leq N$  情况下  $p_t(1, n_t)$  最优值的过程图。可得到理想观测假设下  $\beta=0.3$  时的最大期望总奖励为  $u_1^{\max}((0, 3))=2.098$ , 理想观测下最大网络可靠性为  $u_1^{\max}((0, 3))/N=0.699$ ; 而  $\beta=0.1$  时的最大期望总奖励为  $u_1^{\max}((0, 3))=2.259$ , 理想观测下最大网络可靠性为  $u_1^{\max}((0, 3))/N=0.753$ 。

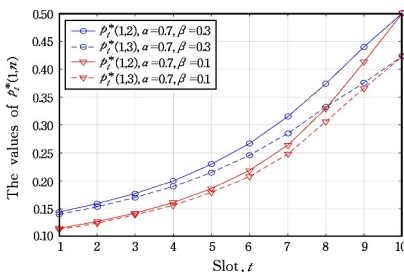


图 2  $N=3, D=10, c_1=0$  条件下最优发送概率的求解过程  
Fig. 2 Procedure to obtain optimal transmission probabilities under  $N=3, D=10, c_1=0$

## 4 适用于实际观测的启发式广播策略

第 3 节的 MDP 建模是基于各待发送次用户能够准确获知任一时隙开始时刻的待发送次用户数量这一理想观测假设, 因此其获得最优广播策略并不适用于第 2.2 节所述的实际观测能力。鉴于此, 本节受第 3 节工作的启发, 提出了一种适用于实际观测能力的启发式广播策略, 并通过 MDP 建模获得了其网络可靠性。

### 4.1 启发式策略设计

多于两个次用户在时隙  $t$  同时发送数据分组, 称为时隙  $t$  发生了严重冲突。表 1 为使用理想观测下最优广播策略造成严重冲突的仿真概率, 每项数据均由 10 次独立仿真且每次仿真持续  $10^7$  个时隙的实验平均求取。如表 1 所列, 在不同参数组合下发生严重冲突的最大概率仅为 0.0560, 表明理想观测下使用其对应的最优广播策略极少造成严重冲突。

表 1 发生严重冲突的仿真概率

Table 1 Simulative probabilities of occurrence of severe collisions

		$N=10$	$N=20$	$N=30$	$N=40$
$\alpha=0.7,$ $\beta=0.3$	$D=10$	0.0462	0.0538	0.0553	0.0560
	$D=20$	0.0162	0.0531	0.0543	0.0547
	$D=30$	0.0063	0.0351	0.0541	0.0546
$\alpha=0.4,$ $\beta=0.6$	$D=10$	0.0286	0.0307	0.0311	0.0312
	$D=20$	0.0215	0.0306	0.0309	0.0311
	$D=30$	0.0121	0.0302	0.0308	0.0311

基于以上观察, 本节的设计启发是: 使用实际观测能力下的最优广播策略极少造成严重冲突, 因此可以假设次用户冲突均仅由两个次用户参与。引入参数  $v_t$  表示时隙  $t$  开始时刻基于上述启发来估计得出的待发送次用户数, 并且规定  $c_t=0$  或  $v_t=0$  时, 各待发送次用户在时隙  $t$  确定性不发送数据;  $c_t=v_t=1$  时, 各待发送次用户在时隙  $t$  确定性发送数据, 其他情况时各待发送次用户在时隙  $t$  以 3.2 节所推导的概率  $p_t^*(c_t, v_t)$  发送数据。

### 4.2 启发式策略的可靠性分析

给定次用户总数  $N$ 、传输时限  $D$  和信道被主用户占用状态的马尔可夫链参数  $\alpha, \beta$ , 本节通过 MDP 建模对上述启发式广播策略进行可靠性分析。其状态、动作、决策规则、状态转移概率、期望奖励以及期望总奖励的定义如下。

状态: 定义状态空间  $H$  为:

$$H \triangleq \{(c, v, n), c \in \{0, 1\}, v, n \in \{0, 1, \dots, N\}\} \quad (23)$$

定义时隙  $t$  的状态  $h_t$  为:

$$h_t \triangleq (c_t, v_t, n_t) \in H \quad (24)$$

其中,  $c_t$  为时隙  $t$  的信道被主用户占用状态,  $v_t$  表示时隙  $t$  开始时刻根据实际观测估计得到的待发送次用户数,  $n_t$  为时隙  $t$  开始时刻真实的待发送次用户数。显然, 由于实际观测估计保守地计算参与冲突的次用户数, 可以得到  $v_t \geq n_t$ 。

动作: 定义与 3.1 节相同。

决策规则: 在时隙  $t$ , 当  $c_t=0$  或  $v_t=0$  时,  $p_t^*(c_t, v_t)=0$ , 因此动作  $a=0$  在时隙  $t$  被确定性执行。在时隙  $t$ , 当  $c_t=v_t=1$  时, 待发送次用户以概率  $p_t^*(c_t, v_t)=1$  发送数据, 因此动作

$a=1$  在时隙  $t$  被确定性执行。在时隙  $t$ , 当  $c_t=1$  以及  $v_t>1$  时,  $n_t$  个待发送次用户均以概率  $0 < p_t^*(c_t, v_t) < 1$  发送数据, 因此任一动作  $a \in \{1, 2, \dots, n_t\}$  按概率:

$$\left(\frac{n_t}{a}\right) p_t^*(1, v_t)^a (1 - p_t^*(1, v_t))^{n_t - a} \quad (25)$$

被执行, 而使得其他  $a \in \{0, n_t+1, n_t+2, \dots, N\}$  不会被执行。

状态转移概率: 将状态  $h_t$  转移至状态  $h_{t+1}$  的概率表示为  $\text{Pr}'(h_{t+1} | h_t)$ 。在时隙  $t$ , 当  $c_t=0$  时, 对于任一  $n_t \in \{0, 1, \dots, N\}$  以及  $v_t \in \{n_t, n_t+1, \dots, N\}$ , 容易得到:

$$\text{Pr}'((c_{t+1}=0, v_{t+1}=v_t, n_{t+1}=n_t) | (c_t=0, v_t, n_t)) = 1 - \alpha \quad (26)$$

$$\text{Pr}'((c_{t+1}=1, v_{t+1}=v_t, n_{t+1}=n_t) | (c_t=0, v_t, n_t)) = \alpha \quad (27)$$

在时隙  $t$ , 当  $c_t=1, v_t \in \{0, 1\}, n_t \in \{0, v_t\}$  时, 容易得到:

$$\text{Pr}'((c_{t+1}=0, v_{t+1}=0, n_{t+1}=0) | (c_t=1, v_t=0, n_t=0)) = \beta \quad (28)$$

$$\text{Pr}'((c_{t+1}=1, v_{t+1}=0, n_{t+1}=0) | (c_t=1, v_t=0, n_t=0)) = 1 - \beta \quad (29)$$

$$\text{Pr}'((c_{t+1}=0, v_{t+1}=0, n_{t+1}=0) | (c_t=1, v_t=1, n_t=1)) = \beta \quad (30)$$

$$\text{Pr}'((c_{t+1}=1, v_{t+1}=0, n_{t+1}=0) | (c_t=1, v_t=1, n_t=1)) = 1 - \beta \quad (31)$$

在时隙  $t$ , 当  $c_t=1$  时, 对于任一  $v_t \in \{2, 3, \dots, N\}$ , 任一  $n_t \in \{0, 1, \dots, v_t\}$  和任一  $a \in \{0, 1, \dots, n_t\}$ , 可得:

$$\text{Pr}'((c_{t+1}=0, v_{t+1}=v_t - \min(a, 2), n_{t+1}=n_t - a) | (c_t=1, v_t, n_t)) = \left(\frac{n_t}{a}\right) p_t^*(1, v_t)^a (1 - p_t^*(1, v_t))^{n_t - a} \beta \quad (32)$$

$$\text{Pr}'((c_{t+1}=1, v_{t+1}=v_t - \min(a, 2), n_{t+1}=n_t - a) | (c_t=1, v_t, n_t)) =$$

$$\left(\frac{n_t}{a}\right) p_t^*(1, v_t)^a (1 - p_t^*(1, v_t))^{n_t - a} (1 - \beta) \quad (33)$$

除以上情况,  $\text{Pr}'(h_{t+1} | h_t) = 0$ 。

期望奖励: 在时隙  $t$ , 定义期望奖励  $w_t(h_t)$  为成功传输的数据分组数量的期望值, 即:

$$w_t(h_t) \triangleq \begin{cases} n_t p_t^*(1, v_t) (1 - p_t^*(1, v_t))^{n_t - 1}, & \text{if } c_t=1, v_t > 1, n_t \geq 1 \\ 1, & \text{if } c_t=v_t=n_t=1 \\ 0, & \text{otherwise} \end{cases} \quad (34)$$

期望总奖励: 定义期望总奖励  $W(h_1)$  为初始状态  $h_1$  条件下从时隙 1 至时隙  $D$  期间成功传输的数据分组总数的期望, 即:

$$W(h_1) \triangleq \mathbb{E}_{h_1} \left\{ \sum_{t=1}^D w_t(h_t) \right\} \quad (35)$$

类似于第 3.2 节, 可以运用逆向归纳法获得期望总奖励  $W(h_1)$  及其对应的启发式策略的网络可靠性。

定义值函数  $m_t(h_t)$  如下:

$$m_t(h_t) \triangleq \mathbb{E}_{h_t} \left\{ \sum_{t'=t}^D w_{t'}(h_{t'}) \right\} \quad (36)$$

可见  $W(h_1) = m_1(h_1)$ 。基于贝尔曼方程, 值函数求解过程如算法 3 所示。

### 算法 3 值函数 $m_t(h_t)$ 求解过程

1. 设置  $t=D$ , 对于任一  $h_D \in H$ , 获得:

$$m_D(h_D) = w_D(h_D) \quad (37)$$

2. 如果  $t=1$ , 停止迭代; 否则进入步骤 3。

3. 令  $t=t-1$ , 对于任一  $h_t \in H$ , 获得:

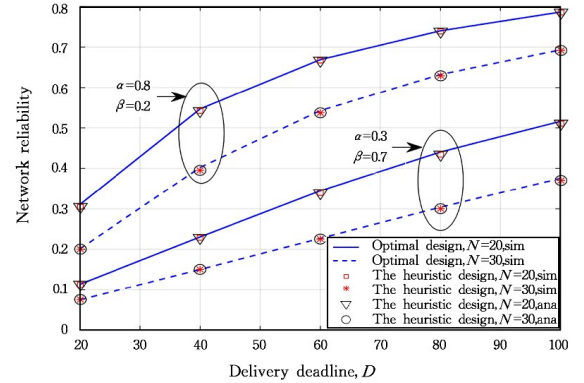
$$m_t(h_t) = w_t(h_t) + \sum_{h_{t+1} \in H} \text{Pr}'(h_{t+1} | h_t) w_{t+1}(h_{t+1}) \quad (38)$$

4. 返回步骤 2。

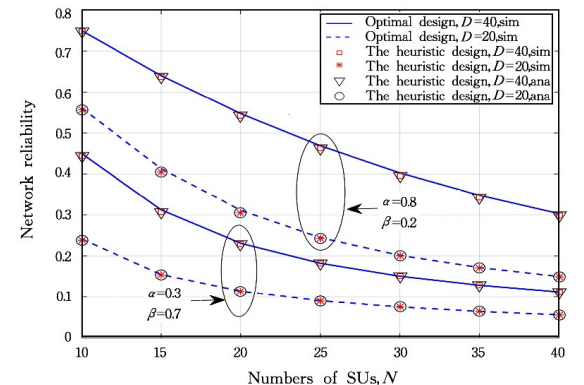
## 5 仿真结果

本节使用仿真实验对比理想观测下的最优策略、符合实际观测的启发式策略和最优静态广播策略的可靠性能, 并验证 4.2 节启发式策略理论分析的准确性。考虑的最优静态策略为: 给定网络参数, 各待发送次用户在信道未被占用时始终以相同且固定的最优静态发送概率发送数据, 此最优静态发送概率通过遍历搜索获得。

仿真工具使用 MATLAB 软件, 仿真场景设置数据传输速率为  $8 \times 10^6$  bit/s, 时隙时间为  $30 \mu\text{s}$ , 侦听主用户状态和侦听次用户传输状态需要时间各为  $5 \mu\text{s}$ , 次用户传播延迟不大于  $1 \mu\text{s}$ , 次用户传输的数据分组长度为 24 字节。每项仿真数据均由 10 次独立仿真且每次仿真持续  $10^6$  个时隙的实验平均求取。在仿真实验的每个超帧开始时刻, 根据  $c_1=1$  的稳态概率  $\alpha/(\alpha+\beta)$  随机产生信道被占用状态。



(a) 网络可靠性随着传输时限的变化



(b) 网络可靠性随着次用户数目的变化

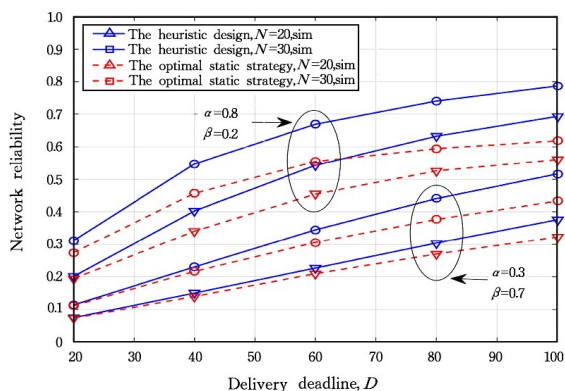
图 3 启发式策略和理想观测下最优策略的可靠性对比

Fig. 3 Comparison of reliability between heuristic strategy and optimal strategy under ideal observation

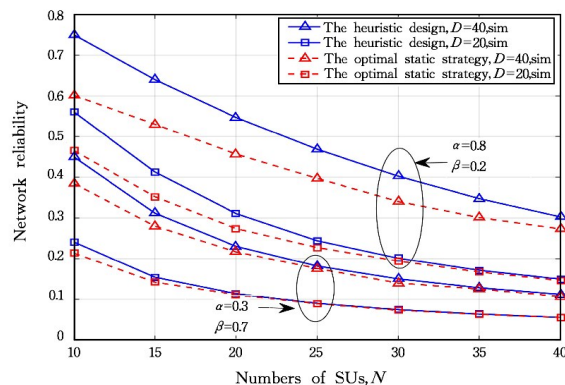
图 3 在一系列网络参数下对比了符合实际观测的启发式

策略网络可靠性的仿真值和理论值,以及理想观测下最优策略的网络可靠性仿真值。可以发现,三者取值在所有情况下均非常接近,表明了启发式策略设计的合理性以及其理论分析的准确性。另外,给定  $N, \alpha, \beta$ , 三者的取值均随着  $D$  的增加而增加并逐渐趋于稳定。此现象的原因是传输时限的增加使得传输机会增加并且冲突概率减小,从而增加了网络可靠性。进一步可以发现,给定  $N$  和  $D, \alpha/(\alpha+\beta)$  增大使得信道不被占用的时间增加,从而可以改善网络可靠性。此外,  $N$  增大会加剧次用户的竞争,使得网络可靠性能恶化。

图 4 仿真比较了启发式策略与最优静态策略的可靠性性能。可以发现,启发式策略在所有情况下均明显表现更佳,但当  $\alpha/(\alpha+\beta)$  较小,  $D$  较小或  $N$  较大时,启发式策略的性能优势将变小。这是由于在这两种情况下,信道未被主用户占用的时间变少,从而造成发送概率动态调整的机会变少,影响了其对改善网络可靠性的作用。而  $N$  较大的情况下,次用户之间的竞争变得更加激烈,从而限制了启发式策略的作用。



(a) 网络可靠性随着传输时限的变化



(b) 网络可靠性随着次用户数目的变化

图 4 启发式策略和最优静态策略的可靠性对比

Fig. 4 Comparison of reliability between heuristic strategy and optimal static strategy

**结束语** 本文设计了一种传输时限下的 CRN 动态广播策略,基于 MDP 提出了一种理想观测下的最优策略,并受此启发提出了一种适用于实际观测的启发式策略以及此策略可靠性的理论分析方法。仿真结果表明,该启发式策略获得的可靠性非常接近于理想观测下最优策略的可靠性,并且明显优于最优静态策略。将来的工作将进一步探索如何结合非正

交多址接入技术<sup>[18]</sup>以及网络编码技术<sup>[19]</sup>设计传输时限下的 CRN 动态广播协议。

## 参考文献

- [1] KHAN A A, REHMANI M H, RACHEDI A. Cognitive-radio-based Internet of things: Applications, architectures, spectrum related functionalities, and future research directions[J]. IEEE Wireless Communications, 2017, 24(3): 17-25.
- [2] RYSZARD S. Cognitive radio, spectrum, and evolutionary heuristics[J]. IEEE Communications Magazine, 2018, 56(6): 166-171.
- [3] POPOVSKI P, NIELSEN J J, STWFANOVIC C, et al. Wireless access for ultra-reliable low-latency communication: principles and building blocks[J]. IEEE Network, 2018, 32(2): 16-23.
- [4] BENNIS M, DEBBAH M, POOR H V. Ultra-reliable and low-latency wireless communication: Tail, risk and scale[J]. Proceeding of IEEE, 2018, 106(10): 1834-1853.
- [5] DENG L, DENG J, CHEN P N. On the asymptotic performance of delay-constrained slotted ALOHA[C] // 27th International Conference on Computer Communication and Networks (ICCCN). 2018: 1-8.
- [6] BAE Y H. Analysis of optimal random access for broadcasting with deadline in cognitive radio networks[J]. IEEE Communications Letters, 2013, 17(3): 573-575.
- [7] BAE Y H. Optimal retransmission-based broadcasting under delivery deadline constraint[J]. IEEE Communications Letters, 2015, 19(6): 1041-1044.
- [8] BAE Y H. Queueing analysis of deadline - constrained broadcasting in wireless networks[J]. IEEE Communications Letters, 2015, 19(10): 1782-1785.
- [9] BAE Y H. Modeling timely-delivery ratio of slotted aloha with energy harvesting[J]. IEEE Communications Letters, 2017, 21(8): 1823-1826.
- [10] ZHANG Y, LO Y H, SHU F, et al. Achieving maximum reliability in deadline-constrained random access with multiple-packet reception[J]. IEEE Transactions on Vehicular Technology, 2019, 68(6): 5997-6008.
- [11] SULTANA A, FERNANDO X, ZHAO L. An overview of medium access control strategies for opportunistic spectrum access in cognitive radio networks[J]. Peer-to-Peer Networking and Applications, 2017, 10: 1113-1141.
- [12] 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.
- [13] JODA R, ZORZI M. Access policy design for cognitive secondary users under a primary type-I HARQ process[J]. IEEE Transactions on Communications, 2015, 63(11): 4037-4049.
- [14] URGONKAR R, NEELY M J. Opportunistic scheduling with reliability guarantees in cognitive radio networks[J]. IEEE Transactions on Mobile Computing, 2019, 8(6): 766-777.

- [15] NAPARSTEK O, COHEN K. Deep multi-user reinforcement learning for distributed dynamic spectrum access [J]. *IEEE Transactions on Wireless Communications*, 2019, 18(1): 310-323.
- [16] EWAISHA A E, TEPEDELENLIOGLU C. Throughput optimization in multichannel cognitive radios with hard-deadline constraints[J]. *IEEE Transactions on Vehicular Technology*, 2016, 65(4): 2355-2368.
- [17] XU D, LI Q. Cooperative resource allocation in cognitive wireless powered communication networks with energy accumulation and deadline requirements[J]. *SCIENCE CHINA Information Sciences* 2019, 62(8): 141-154.
- [18] THAKUR P, KUMAR A, PANDIT S, et al. Frameworks of non-orthogonal multiple access techniques in cognitive radio communication systems[J]. *China Communications*, 2019, 16(6): 129-149.
- [19] LI Y, ZHOU J, WANG J, et al. On data dissemination enhanced by network coded device-to-device communications [J]. *IEEE Transactions on Wireless Communications*, 2020, 19(6): 3963-3976.



**FANG Ting**, born in 1998, postgraduate, is a member of China Computer Federation. Her main research interests include design and optimization for protocols in wireless networks.



**ZHANG Yi-jin**, born in 1982, Ph.D, professor. His main research interests include sequence design, wireless networks, and artificial intelligence.