

基于在线学习的频谱感知 与接入方法研究

许左宏 张周 王世练 著

清华大学出版社
北京

内 容 简 介

本书针对认知无线网络频谱统计信息未知和时变信道条件下用频设备的高效接入,基于多臂赌博机在线学习理论构建问题模型,提出面向终端能力约束和服务质量要求多样性的频谱感知与接入在线学习方法,实现用频设备对频谱统计信息的在线学习和频谱的高效利用,可为无线通信网络提供理论和方法支撑。通过对本书的学习,读者将掌握基于在线学习的频谱感知与接入技术,为解决未来无线通信网络的关键问题提供有力支持。

本书的研究成果适用于无人机网络、有限电量的无线传感器网络及不同业务类型的物联网等典型通信场景。

本书适合无线通信、认知无线电、物联网等领域的人员、工程师及相关专业的师生阅读。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。举报:010-62782989, beiqinquan@tup.tsinghua.edu.cn。

图书在版编目(CIP)数据

基于在线学习的频谱感知与接入方法研究 / 许左宏, 张周, 王世练著. -- 北京: 清华大学出版社, 2026. 8. -- ISBN 978-7-302-72619-7

I. TN925

中国国家版本馆 CIP 数据核字第 2026445NR2 号

责任编辑: 陈凯仁

封面设计: 刘艳芝

责任校对: 薄军霞

责任印制: 宋 林

出版发行: 清华大学出版社

网 址: <https://www.tup.com.cn>, <https://www.wqxuetang.com>

地 址: 北京清华大学学研大厦 A 座 邮 编: 100084

社 总 机: 010-83470000 邮 购: 010-62786544

投稿与读者服务: 010-62776969, c-service@tup.tsinghua.edu.cn

质量反馈: 010-62772015, zhiliang@tup.tsinghua.edu.cn

印 装 者: 三河市东方印刷有限公司

经 销: 全国新华书店

开 本: 170mm×240mm 印 张: 12.75 插 页: 5 字 数: 243 千字

版 次: 2026 年 8 月第 1 版 印 次: 2026 年 8 月第 1 次印刷

定 价: 69.00 元

产品编号: 111078-01

符号说明



符号	含义
$N!$	$1 \times 2 \times \cdots \times N$
$\binom{N}{M}$	二项式系数组合数 $\frac{N!}{M!(N-M)!}$
$\mathcal{I}_N$	集合 $\{1, 2, \cdots, N\}$, N 为正整数
$\mathbb{R}$	实数集
$\mathbb{R}^d$	d 维实向量集
$E\{\cdot\}$	求数学期望
$\lceil \cdot \rceil$	上取整函数
$\lfloor \cdot \rfloor$	下取整函数
I	指示函数
$D(x, y)$	Kullback-Leibler 距离
mod	取模运算
$\mathbf{X}$	向量或矩阵
$\mathbf{X}^T$	转置
$\ \mathbf{Y}\ _{\mathbf{X}}$	矩阵范数。 $\ \mathbf{Y}\ _{\mathbf{X}} = \sqrt{\mathbf{Y}^T \mathbf{X} \mathbf{Y}}$, 其中 $\mathbf{X}$ 为矩阵
$\mathcal{M}$	集合
$ \mathcal{M} $	集合的势
$\max\{\cdot\}$	求集合中最大值
$\min\{\cdot\}$	求集合中最小值
$\arg \max\{\cdot\}$	求最大值对应的自变量

续表

符号	含义
$\arg \min\{\}$	求最小值对应的自变量
$\cup$	并集
$\emptyset$	空集
$O[g(n)]$	复杂度符号。对于 $f(n)$ ，存在正数 n_0 、 c ，使得当 $n \geq n_0$ 时，始终存在 $0 \leq f(n) \leq c \times g(n)$ ，则记 $f(n) = O[g(n)]$

前言



P R E F A C E

随着下一代无线通信网络技术与协议标准的快速演进，有限的频谱资源和迅速攀升的用户通信需求之间的矛盾加剧，成为当前制约通信系统发展的重要问题。作为缓解频谱资源供需矛盾的重要技术方向，基于在线学习的频谱感知与接入技术赋予了用频设备自主学习和推理能力，其通过实时感知频谱信息和分析无线信道状态的相关数据，自适应地调整接入决策，从而优化频谱资源的利用，在不同无线传输场景中具有重大的应用潜力。

近年来，在无人系统、人工智能等新兴技术驱动下，无线通信网络向复杂动态化、轻量移动化、智能自主化等趋势发展，使得频谱利用面临实时信道信息获取难以支持高效接入、终端有限能量资源要求高能量效率接入、多样化服务质量要求亟须高传输效能接入等难题。针对频谱统计信息未知和时变条件下用频设备高效接入，本书构建基于多臂赌博机（multi-armed bandit, MAB）在线学习理论的问题模型，提出面向终端能力约束和服务质量要求多样性的频谱感知与接入在线学习方法，实现用频设备对频谱统计信息的在线学习和频谱的高效利用，可为无线通信网络提供理论和方法支撑，并且研究成果适用于无人机网络、有限电量的无线传感器网络以及不同业务类型的物联网等典型通信场景。本书的主要研究工作及研究成果具体如下。

（1）开展多信道感知和部分信道接入场景下的高效率频谱利用研究，提出在线学习两步决策方法，从而提高无线传输的平均吞吐量。针对集中式网络传输数据量最大化，构建基于双层 MAB 在线学习的问题模型，提出基于信道空闲概率估计的上置信界限（upper confidence bound, UCB）排序的频谱感知与接入在线学习两步决策方法，并将理论分析与仿真分析相结合。研究表明，累积传输数据量损失与时隙数量呈对数增长关系，同时用频设备可最终识别最优信道并收敛至最优吞吐量。

（2）开展用频设备能量约束下的高能量效率频谱利用研究，提出兼顾能量

消耗和传输性能的在线学习方法,从而提高无线传输的能量效率。针对单用户能量约束下传输数据量最大化,构建基于背包问题和 MAB 在线学习的问题模型,提出基于吞吐量与能量消耗比值估计的 UCB 排序和贪心策略的频谱感知与接入在线学习方法,并将理论分析与仿真分析相结合。研究结果表明,累积传输数据量损失与用户能量呈对数增长关系,用频设备能够在一定能量约束下准确识别最优信道并收敛至最优能量效率。进一步将单用户场景扩展至集中式网络和分布式网络,构建基于多维背包问题、分配问题和 MAB 在线学习理论的问题模型,分别提出基于吞吐量与能量消耗比值估计的 UCB 排序和匈牙利算法、伯特塞卡斯(Bertsekas)拍卖算法的频谱感知与接入在线学习方法。仿真结果表明,所提方法可实现最优信道分配并收敛至最优能量效率。

(3) 开展面向多类型服务质量要求的高传输效能频谱利用研究,提出可根据实时业务调整用频决策的在线学习方法,从而提高频谱利用的灵活性和传输效能。针对单用户多业务传输效能最大化,构建基于多属性决策(multiple attribute decision making, MADM)和上下文 MAB 在线学习的问题模型,提出基于传输效能估计的 UCB 排序的频谱感知与接入在线学习方法,并将理论分析与仿真相结合。研究结果表明,所提方法的累积传输效能损失与时隙数量呈亚线性关系。进一步拓展至集中式网络传输效能最大化,构建基于组合优化理论和上下文 MAB 在线学习的问题模型,提出基于传输效能估计的 UCB 排序和匈牙利算法的频谱感知与接入在线学习方法。进一步构建分布式网络多用户上下文 MAB 在线学习的问题模型,提出基于传输效能估计的 UCB 排序和正交循环策略的频谱感知与接入在线学习方法。仿真结果表明,所提方法能够根据不同传输业务需求调整感知接入策略,从而提高频谱利用的灵活性。

本书的撰写得到研究团队王世练教授、张周副研究员、周壮博士的大力支持。在此,作者表示衷心的感谢!

限于作者水平,本书难免存在不足之处,敬请广大同行、专家和读者批评指正。

许左宏

2025 年 7 月

目录



C O N T E N T S

第 1 章 绪论	1
1.1 研究背景及意义	1
1.2 国内外研究现状	3
1.2.1 基于在线学习的频谱感知与接入方法	5
1.2.2 多臂赌博机在频谱感知与接入中的应用	7
1.2.3 待进一步解决的问题	10
1.3 研究内容与组织结构	13
1.3.1 研究内容	13
1.3.2 组织结构	15
第 2 章 频谱感知与接入及多臂赌博机在线学习理论基础	18
2.1 相关概念与假设	18
2.1.1 相关概念	18
2.1.2 主要假设	19
2.2 频谱感知与接入基础	20
2.2.1 信道感知	23
2.2.2 信道切换	25
2.2.3 信道共享与接入	26
2.2.4 控制信道	27
2.3 面向频谱感知与接入的多臂赌博机在线学习理论	29
2.3.1 多臂赌博机在线学习问题模型	29
2.3.2 多臂赌博机在线学习模型与频谱感知接入映射关系	31
2.3.3 在线学习方法与性能分析	33

2.4 小结	35
第 3 章 基于双层多臂赌博机的多信道感知与部分信道接入方法	36
3.1 集中式网络的频谱感知与接入方法	37
3.1.1 系统模型	37
3.1.2 模型建立与问题描述	38
3.1.3 基于空闲概率估计上置信界限的方法设计	40
3.1.4 性能分析和复杂度分析	42
3.1.5 仿真实验	49
3.2 分布式网络正交频谱感知与接入方法	58
3.2.1 系统模型	58
3.2.2 模型建立与问题描述	58
3.2.3 基于空闲概率估计的上置信界限和正交接入的方法设计	60
3.2.4 性能分析和复杂度分析	64
3.2.5 仿真实验	67
3.3 小结	73
第 4 章 基于背包问题和多臂赌博机的能量约束频谱感知与接入方法	74
4.1 能量约束下单用户频谱感知与接入方法	75
4.1.1 系统模型	75
4.1.2 模型建立与问题描述	77
4.1.3 基于吞吐量与能量消耗比值估计的上置信界限和贪心策略的方法设计	79
4.1.4 性能及复杂度分析	84
4.1.5 仿真实验	94
4.2 能量约束下集中式网络频谱感知与接入方法	100
4.2.1 系统模型	100
4.2.2 模型建立与问题描述	102
4.2.3 基于吞吐量与能量消耗比值估计的上置信界限和匈牙利算法的方法设计	104
4.2.4 复杂度分析	110
4.2.5 仿真实验	111
4.3 能量约束下分布式网络频谱感知与接入方法	116
4.3.1 系统模型	116
4.3.2 模型建立与问题描述	117

4.3.3	基于吞吐量与能量消耗比值估计的上置信界限和拍卖算法的方法设计	119
4.3.4	性能及复杂度分析	124
4.3.5	仿真实验	131
4.4	小结	136
第 5 章 基于上下文多臂赌博机的多类型服务质量要求频谱感知与接入方法		137
5.1	多类型服务质量要求下单用户频谱感知与接入方法	138
5.1.1	多属性决策理论	138
5.1.2	系统模型	139
5.1.3	模型建立与问题描述	142
5.1.4	基于传输效能估计的上置信界限和层次分析法的方法设计	143
5.1.5	性能及复杂度分析	149
5.1.6	仿真实验	154
5.2	多类型服务质量要求下多用户频谱感知与接入方法	160
5.2.1	系统模型	160
5.2.2	基于传输效能估计的上置信界限和匈牙利算法的多用户感知与接入方法设计	163
5.2.3	基于传输效能估计的上置信界限和正交循环策略的多用户感知与接入方法设计	166
5.2.4	仿真实验	170
5.3	小结	176
第 6 章 总结与展望		177
6.1	研究工作总结	177
6.2	展望	179
参考文献		180

绪 论

1.1 研究背景及意义

无线频谱资源是无线通信网络迫切需要拓展的国家战略资源。近年来，随着第五代移动通信技术（5th generation mobile communication technology, 5G）、物联网、移动车联网、无人蜂群、卫星通信、应急通信的迅猛发展，海量的用户数量和日益增长的通信需求对频谱资源的高效利用提出了更高的要求。

根据国际电信联盟（International Telecommunication Union, ITU）关于下一代无线通信网络愿景的报告，未来的无线通信网络呈现以下发展趋势^[1-3]。

（1）复杂动态化。未来，无线通信网络将面临多种不同体制的无线网络共存、网络密集程度增加的局面。一方面，网内及网间用户的用频争夺、共存网络频谱共享、用户动态出入网、不同业务对传输有不同性能需求、非法用户恶意干扰等多种行为，使得用频行为更加复杂；另一方面，物联网、无人设备、卫星通信的兴起极大地丰富了通信网络中的设备种类和数量，不同用频设备对频谱使用提出了不同需求，频谱资源的高效使用更加迫切。

（2）轻量移动化。未来，无线网络正向移动设备的便携化、物联网设备的小型化，以及无人终端的移动化发展。一方面，用频设备受有限电量制约，对高能量效率的需求日益迫切；另一方面，用频设备通常具有有限的数据处理和通信能力。在有限的通信能力和能量资源条件下，如何实现尽可能优的传输性能，已成为未来无线网络技术发展所面临的关键挑战。

（3）智能自主化。人工智能和认知无线电技术的不断发展使得无线网络的智能化水平得到了显著提升。无线网络中的用频设备将具备学习和自主决策的能力，

即能够利用历史数据和环境反馈来优化对频谱的感知与接入方法,从而实现更高效的频谱利用。

为契合无线通信网络的复杂动态化、轻量移动化以及智能自主化的发展趋势,相关研究亟须解决以下 3 个实际问题。

(1) 实时信道信息获取难以支持高效接入:在动态复杂的网络环境中,信道的实时占用状态和信道质量等信息将会影响信息传输的可靠性,这在优化网络接入方面具有关键作用。然而,有限的感知与接入能力、频谱利用的冲突、信道状态的快速变化及网络条件的波动等多个因素使得用频设备对实时信道信息的获取和利用变得越发困难,直接影响用频设备对频谱的高效接入。此外,在多用户网络中,获取频谱信息的准确性、利用频谱资源的竞争性、接入频谱传输的可靠性将共同影响用频设备对频谱资源的利用。因此,需研究如何在有限的频谱感知与接入能力下获取频谱信息和资源共享,以实现频谱资源的高效利用。

(2) 终端有限能量资源要求高能量效率接入:部分用频设备通常具有有限电量,从而制约了频谱接入的长期性能,因此需研究如何设计高效利用能量资源的感知与接入方法,使得用频设备能够实现最大化的频谱接入效率和尽可能长的使用寿命。

(3) 多样化服务质量要求亟须高传输效能接入:复杂多体制共存的网络导致多种不同业务共存,多样化的传输业务对服务质量有不同的需求。用频设备如何根据自身实时业务自适应调整频谱利用决策、在不同业务下公平高效地共享频谱资源是未来无线网络必须面临的挑战。

为应对上述问题,认知无线电(cognitive radio, CR)作为一种智能化且自适应的无线电通信系统,能够实时感知周围环境并动态调整传输参数和接入决策,从而高效利用频谱资源,对提升未来无线网络的频谱利用率具有重大意义。随着人工智能技术的不断发展,基于在线学习理论的频谱感知与接入技术已成为认知无线电领域的研究热点。在该技术中,频谱可被划分为多个不相交的信道。频谱感知与接入通常包含两个过程:首先,用频设备需在较宽的频段内快速有效地进行频谱感知,以获取信道的实时状态和质量信息;其次,根据感知到的信息进行接入,以完成信息传输^[4]。而在线学习理论赋予了用频设备自主学习和推理的能力,可通过实时感知频谱信息和分析无线信道状态的相关数据,自适应调整接入决策,从而优化频谱资源的利用,在不同无线传输场景中具有重大的应用潜力,具体如下:①在实时频谱接入方面,可结合用频设备实际的硬件能力及分析多用户间用频冲突带来的传输影响,设计智能化在线学习感知与接入方法,使得用频设备能够学习频谱信息并与其他设备高效共享频谱资源,从而深入挖掘频谱使用潜

力。② 在优化能量消耗方面,建立频谱感知与接入的能量消耗模型,使得用频设备能够基于自身能量、在线学习的频谱信息及能量消耗信息调整频谱感知与接入方法,从而达到能量消耗与信息传输的平衡。③ 在满足不同业务服务质量要求方面,结合用频设备实时业务类型,建立有效的频谱感知与接入模型和目标函数,使得用频设备可根据自身实时的服务质量要求,自适应选择合适的频谱进行感知与接入。综上所述,基于在线学习的频谱感知与接入技术可为未来无线通信网络亟须解决的问题提供有效的解决方案,具有重要的研究价值。

鉴于上述背景,本书在契合未来无线网络发展趋势的基础上,聚焦实时信道信息获取难以支持高效接入、终端有限能量资源要求高能量效率接入、多样化服务质量要求亟须高传输效能接入这三个问题,基于在线学习理论针对性地设计一系列频谱感知与接入方法,旨在提供智能、灵活的频谱利用方案,助力未来无线网络向更加可靠和高效的方向发展。

1.2 国内外研究现状

认知无线电的概念最早由 Mitola 于 1999 年提出,拓展至多用户网络后,进而构建认知无线网络 (cognitive radio network, CRN)^[5]。在认知无线电技术中,通常存在两类用户:一是使用授权频谱的主要用户,其对频谱享有优先使用权;二是利用频谱机会的次要用户,其通常通过频谱感知获取无线环境信息,基于频谱共享,根据不同策略如次要用户优先级、服务质量 (quality of service, QoS) 要求、公平性等与主要用户和多个次要用户共享频谱资源,以提高频谱利用率和无线传输性能。经过 20 多年的发展,认知无线电技术已得到广泛的研究。

值得说明的是,由于技术和标准的复杂度、与现有系统的兼容性问题、市场和商业模式的不明确性、安全和隐私问题以及部署成本等因素,目前主流的频谱管理政策和法规多是基于静态的频谱分配,认知无线电技术未能够得到广泛应用^[3]。然而,认知无线电中的核心技术,如频谱感知与接入技术,凭借其利于动态优化网络协议并分配资源、深入挖掘频谱利用率等优势,已在一些专用网络,如军事通信、无人机蜂群网络、传感器网络、卫星通信网络中得到了一定应用^[4-7],具体如下:① 在军事通信中,用频设备可快速感知复杂多变干扰的复杂电磁环境,并自适应选择不同网络体制和通信手段遂行作战任务。② 在无人机蜂群网络中,各用频设备可通过频谱感知的手段获取电磁环境态势,再通过频谱共享方法完成对无线资源的利用。③ 在传感器网络中,为实现可靠信息传输,网络管理系统可通过收集多个传感器节点的感知信息建立频谱环境地图,再实时分配合适的信道和

功率资源,以完成信息传输。④在卫星通信网络中,用频设备具备通过频谱感知获取可用频点完成传输、与地面移动通信网络完成频谱共享、识别被干扰的信道并切换可用频点以及监测网络是否受到其他干扰等功能。此外,频谱感知与接入技术在5G网络中也得到了应用。例如,5G网络的基站可以通过频谱感知技术监测无线环境中频谱的使用状态,并与其他基站共享这些感知信息,采取一定策略调整传输参数,可以避免或降低干扰^[8]。除此之外,相较于目前的无线通信技术,未来6G网络在用户类型多样化、能量效率、智能化、多网络共存等方面得到了进一步的研究,其中频谱感知与接入技术在这些方面能够发挥更显著的应用潜力,具体如下:①在用户类型多样化方面,在增强移动宽带(enhanced mobile broadband, eMBB)场景中,6G网络可基于频谱感知与接入技术和自适应技术,实时监测频谱环境的变化,并根据用户需求动态调整利用更优质的频谱资源,从而提高频谱的使用效率^[9-11]。②在能量效率方面,6G网络可基于频谱感知与接入技术预测频谱的使用状态,并据此优化设备对优质频谱资源的利用和自身能量管理,从而达到更高的能量利用效率,实现绿色通信^[12]。③在智能化方面,6G网络正在研究将人工智能技术融入频谱感知与接入技术,实现网络的自组织、自优化和自学习,从而极大地提升网络的智能化和自适应能力^[13]。④在多网络共存方面,6G网络不仅要与4G网络、5G网络等传统移动通信网络共存,还要处理物联网、无人机通信、卫星通信等多种异构网络的融合问题。6G网络可基于频谱感知接入,使得用户动态利用可用的无线资源,实现更加高效的频谱共享^[14-17]。

频谱感知与接入技术在未来无线网络中具有重大的应用潜力,具有复杂化和轻量移动化的发展趋势的无线网络使得信道的实时可用状态和信道质量快速发生变化,但是用频设备如何迅速准确地获取信道信息将变得越发困难,具体如下:①在无人机集群网络中,无人机频谱感知处理能力的限制、外部无线通信系统频谱活动的时变性和无线信号衰落特性,使得无人机难以实时准确地获得所有信道的状态。②在物联网网络和无线传感器网络中,大量低功耗、弱信号处理能力的节点设备被部署在各种无线环境中,受复杂信道条件(如路径损耗、多径、阴影效应)等影响,节点设备难以精确感知所有信道实时状态和质量信息。③在舰船、低轨卫星等通信系统中,通信设备可能在大范围内移动,电磁传输开放性导致网络可能面临复杂的外部干扰,长距离信息传输的延迟导致实时信道状态获取困难。由于频谱感知与接入是两个耦合的过程,对信道信息获取的困难将会影响后续频谱接入的过程,从而降低用频设备对频谱资源的利用效率。为应对上述挑战,已有研究将在线学习理论用于频谱感知与接入方法的设计中,旨在使得用频设备具有对数据的分析和推理能力。该方法能够显著提高频谱资源的有效利用,并且与

未来无线通信智能化的主流发展方向相契合,因此已经引起了学术界、工业界的广泛关注和极大的研究兴趣。

1.2.1 基于在线学习的频谱感知与接入方法

在线学习是机器学习的一个分支,融合了人工智能、概率统计及控制理论等多个领域,其目标是模拟人类的学习过程,通过顺序地接收数据,实时处理有限的样本,以学习数据中的规律,并且能够对未来的数据点或无法预先观测到的现象进行预测和判断^[18]。强化学习(reinforcement learning, RL)及深度强化学习(deep reinforcement learning, DRL)是在线学习中的常见方法,在频谱感知与接入方法设计中已得到广泛关注^[19]。在强化学习中,用户首先观察当前环境,根据当前状态进行相应的决策,获得瞬时收益并更新状态信息。此后,根据历史收益信息调整下一阶段的决策,不断迭代直至接近最优策略^[20]。文献[21]研究认知无线网络中分布式多用户场景下的频谱感知与接入问题,将频谱决策问题建模为部分可观测的随机博弈问题,采用多智能体强化学习对该问题进行求解,并与启发式方法的性能对比,发现所提方法具有有效性。但该研究假设各次要用户可共享其对信道估计的参数及频谱决策方案,因此不适用于部分分布式网络^①。文献[23]考虑噪声影响下的频谱感知与接入问题,将问题建模为部分可观测马尔可夫决策过程,基于价值迭代方法提出频谱感知与接入的在线学习统计决策方法。但该研究假设各次要用户可共享对信道的观测结果及频谱决策方案,因此不适用于部分分布式网络。考虑到认知无线网络场景的复杂性,强化学习难以针对大规模的网络场景有效地处理高维状态空间,因此相关研究通常将强化学习与深度学习(deep learning, DP)结合得到深度强化学习工具,并基于深度学习拟合高维状态空间及决策对应的Q值完成与无线环境的交互^[24-25]。文献[26]将深度强化学习应用到频谱感知与接入,并通过仿真验证所提方法的有效性。文献[27]将深度强化学习应用到多用户场景下的频谱感知与接入,并通过仿真验证所提方法的有效性。

除上述方法外,多臂赌博机(multi-armed bandit, MAB)作为在线学习领域的重要分支,属于经典统计学问题,最早由Lai和Robbins提出^[28]。其问题示意图如图1.1所示,可描述为一个玩家面对一台具有多个机臂的赌博机,玩家每次可以选择一个机臂,并获得该机臂对应的随机收益。每个机臂的收益分布可能

^① 在本书中,部分分布式网络是指一种无线通信网络,其中次要用户无法获知其他传输对的信道感知接入策略,并且无法与其他传输对进行直接信息交换。在这种网络结构中,每个传输对单独工作,没有中央协调机制或信息共享机制。这种场景特别强调了自主和分布式决策的重要性,在网络管理和资源分配方面呈现出独特的挑战和机遇^[22]。

不同且对次要用户未知。玩家的目的是在不断选择机臂的过程中在线学习机臂的收益信息，尽可能地选取收益最大的机臂，并获得最大的累积收益期望。

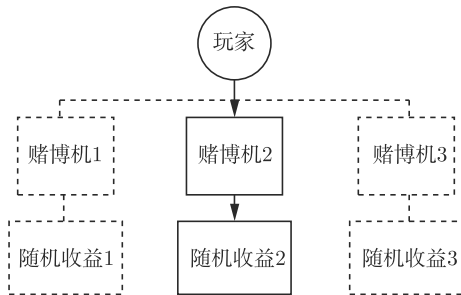


图 1.1 多臂赌博机问题示意图

玩家面对多臂赌博机时，选择机臂的过程可视为一个序列决策过程。玩家为获得最大累积收益期望，需通过不断选取不同机臂获得收益信息，并利用经典统计学原理对各个机臂的收益信息进行估计。基于估计的收益信息，玩家决定下一次选择的机臂，获得收益信息并重复上述过程。在该过程中，玩家面临探索与利用的两难问题：一方面，玩家在选择机臂的过程中，若只选择估计收益信息较大的机臂，则可能错失其他机臂更大收益的机会（因此玩家需尽可能选取不同机臂估计其收益值的统计信息，该过程通常称为探索）；另一方面，玩家若大量选取不同机臂，则可能浪费选择最优机臂（此处最优机臂是指具有最大期望收益值的机臂）以获得最大收益值的机会（因此玩家需尽可能选择最优机臂以获得尽可能大的收益，该过程通常称为利用）。因此，玩家需设计在线学习和决策方法，使尽可能探索不同机臂获得收益信息和尽可能利用最优机臂获得收益值达到平衡。

相较于强化学习和深度强化学习等其他在线学习方法，多臂赌博机具有如下优势：① 轻量化计算与实时性决策。相较于强化学习和深度强化学习，多臂赌博机不需要大量的历史数据来训练模型，而是能够在有限的样本和数据处理中迅速更新信息并做出决策。这种迅速决策的能力能够提供实时有效的解决方案，尤其适用于频谱状态迅速变化的场景。② 便于方法设计。在频谱感知与接入问题中，目的通常是最大化频谱的使用效率，这可以被建模为一个收益最大化问题。多臂赌博机能够提供一种有效的方式使探索不同频谱（感知不同信道）和利用估计的最优频谱（接入信道）达到平衡，也能够良好契合问题目标。此外，多臂赌博机能够在执行决策与无线环境的交互中不断调整策略，便于根据不同无线传输条件针对性展开方法设计。③ 性能保证。多臂赌博机具有可理论分析的特点，通常能够为所设计方法提供理论上的性能保证。这种性能保证对于验证在线学习方法的

有效性至关重要, 为实际应用提供了坚实的理论基础^[29]。

1.2.2 多臂赌博机在频谱感知与接入中的应用

正是凭借上述突出的优势, 多臂赌博机在频谱感知与接入方法设计中得到了广泛的研究。作为该领域的典型代表之一, Lai 等^[30-31] 将独立同分布 (independent and identically distributed, IID) 信道模型下频谱感知与接入问题建模为经典的多臂赌博机问题, 基于文献 [32] 中提出的上置信界限 (upper confidence bound, UCB) 设计在线学习方法, 并证明所提方法的收益损失期望 (即最大累积收益与统计决策方法获得的累积收益的期望间的差值) 与时隙数量 t 呈亚线性关系。此外, 马尔可夫信道模型下的多臂赌博机问题研究同样取得进展。文献 [33] 针对马尔可夫收益模型下的多臂赌博机问题进行研究, 提出统计渐近下的统计决策方法并通过理论分析出所提方法的收益损失的期望值与时隙数量呈 $O(\ln t)$ 关系。基于该研究, Tekin 和 Liu^[34-37] 将马尔可夫信道模型下的频谱感知与接入问题分别建模为休止多臂赌博机 (rested MAB) 和无休止多臂赌博机 (restless MAB) 问题, 基于强马尔可夫特性对信道观测方法和接入方法进行设计, 并通过理论证明了其收益损失期望与时隙数量呈 $O(\ln t)$ 关系。

在上述研究的基础上, 多臂赌博机逐渐拓展到多用户场景下的频谱感知与接入方法设计中。在集中式网络中, 由于存在中心节点 (如基站或网络控制器), 网络的协调和管理变得相对集中化。中心节点可以协调不同用户对信道的感知接入, 从而优化无线资源的分配和使用。在该情况下, 信道感知接入问题可转化为单个用户多信道感知接入问题。在分布式网络中, 根据不同网络场景可将研究分类为静态同构网络、静态异构网络、动态同构网络、动态异构网络下的频谱感知与接入, 其中, 静态、动态是指网络中用户数量是否变化, 同构、异构是指信道占用状态统计信息对各次要用户是否相同, 典型代表如下所述。

(1) 在静态同构网络中, Liu 和 Zhao^[38-39] 将每个时隙划分为多个子时隙, 每个次要用户按照文献 [28] 中的上置信界限方法对信道未知信息进行估计, 以公平性为原则提出时分公平性共享 (time-division fair sharing, TDFS) 策略, 并通过理论分析出所提方法的收益损失期望与时隙数量 t 在统计渐近下为 $O(\ln t)$ 关系。TDFS 算法将每个时隙再次划分时, 对网络同步有较高要求, 因此在大规模网络中实现较为复杂。随后, Anandkumar 等^[40-41] 提出著名的分布式学习和分配 (rho-distributed learning and allocation, ρ^{RAND}) 算法。在该算法中, 各次要用户基于上置信界限算法对信道信息估计值进行排序, 再基于无线电分组交换协议对信道进行争用。当次要用户传输过程中发生冲突时, 其在下一时隙随机选择

信道进行接入。尽管在该过程中次要用户会产生大量冲突,但随着次要用户对信道未知信息估计趋于准确,次要用户最终可实现正交信道分配。通过理论分析可知,该算法的收益损失期望与时隙数量 t 在有限和统计渐近下为 $O(\ln t)$ 关系。基于 ρ^{RAND} 算法,人们对如何降低次要用户的冲突数量展开大量的研究。文献 [42] 利用音乐椅 (musical chair, MC) 方法设计冲突规避机制,提出著名的随机头部 (random top-M, RandTopM) 和音乐椅头部 (musical chair top-M, MCTopM) 算法,并通过理论证明了所提算法与 ρ^{RAND} 相比能够有效降低冲突,并且收益损失期望与时隙数量呈 $O(\ln t)$ 关系。以上算法通过设计机制,可以降低次要用户间的冲突,随着统计信息估计准确最终达到正交状态,而未考虑用户所获得吞吐量或信道接入时间的公平性。针对该问题, Gai 和 Krishnamachari^[44] 假设每个次要用户事先分配索引号,每个时隙中各次要用户根据索引号和上置信界限算法以轮询的方式接入信道,理论分析表明,所提方法的收益损失期望与时隙数量呈 $O(\ln t)$ 关系。在文献 [44] 中,作者设计两步信道接入机制,包括分布式信道协商和轮询式接入过程。通过该机制,每个次要用户能够使用相同的信道接入时间。以上算法皆假设静态同构网络中次要用户数量已知,对于次要用户数量未知的情况,文献 [40] 提出分布式学习和用户估计 (rho-distributed learning and estimation, ρ^{EST}) 算法,在 ρ^{RAND} 算法的基础上添加用户数量估计机制,理论分析表明,该算法的收益损失期望同样可达到与时间呈 $O(\ln t)$ 关系。

(2) 在动态同构网络中,由于次要用户数量发生变化,当次要用户无法得知这一变化且用户间无法进行通信时,上述面临大量冲突的分布式方法难以收敛至正交状态,不再适用于该场景。文献 [45] 基于贪心 (greedy) 算法和无线电分组交换协议设计了基于冲突规避机制的多用户贪心冲突规避 (multi-user ϵ -greedy collision avoiding, MEGA) 算法。理论分析表明,MEGA 算法的收益损失期望与时隙数量呈 $O(t^{1-\beta/2} + \log t + t^\beta)$ 关系,其中 β 为超参数。值得说明的是,尽管该方法未能达到统计渐近意义上的最优性能,但其具有较优的仿真性能。然而,在后续研究中表明,MEGA 算法对于某些动态用户数量的场景性能较差,其原因为冲突数量在某些情况下无法为零,从而产生线性的收益损失期望。为解决该问题,文献 [46] 基于 MC 机制对动态用户数量情况设计动态音乐椅 (dynamic musical chair, DMC) 方法,并通过理论分析出所设计方法的收益损失期望与时隙数量为 $O(\sqrt{xt})$ 关系,其中 x 为 t 个时隙数量内用户变化数量的上界。相较于 MEGA 算法,DMC 算法理论性能更优,但其在信道统计信息估计阶段产生大量的冲突,造成大量性能损失。为解决该问题,文献 [47] 基于徒步 (trekking) 方法提出动态徒步 (dynamic trekking, DT) 算法。该算法包含两部分,即次要用户首先对信道

统计信息进行估计,再基于徒步方法进行信道分配。理论分析表明,所提算法的收益损失期望与时隙数量呈 $O(\sqrt{xt\log t})$ 关系。上述研究中未考虑次要用户间公平性问题,因此文献 [48] 对该问题进行研究,将时隙划分为更小的时隙,基于随机跳跃机制提出动态次要用户协调 (dynamic secondary user coordination with fairness, DSCF) 算法,使得各次要用户能够获得尽可能相同的吞吐量。

(3) 在异构静态网络中,信道的统计信息对不同次要用户未知且不同,因此在方法设计中,次要用户不仅需要学习信道统计信息,还需要考虑次要用户间的信道分配等问题。在该场景下,常采用博弈机制或拍卖算法进行协议设计。作为该类方法的典型代表,最早的在线学习 (learning with linear rewards, LLR) 算法由 Gai 等^[49-50] 提出。该类算法通常将多个次要用户的在线学习和信道分配问题转化为组合多臂赌博机问题,即通过将每一个信道组合方案视为一个机臂,再基于匈牙利算法完成各次要用户的信道分配。经理论分析可知,所提方法的收益损失期望与时隙数量呈 $O(\ln t)$ 关系。但该类算法中假设各次要用户可进行信息交互,因此不适用于部分分布式网络。此外, Kalathil 等^[51] 提出分布式上置信界限 (online index-based distributed learning, dUCB4) 算法,假设各个次要用户可进行信息交互,采用伯特赛卡斯 (Bertsekas) 拍卖算法使得各次要用户经过多次协商最终达到最优正交分配。经理论分析,所提方法的收益损失期望与时隙数量有 $O(\log^2 t)$ 关系。由于 dUCB4 算法的理论性能并未达到最优,文献 [52] 对 dUCB4 算法进行改进,即控制次要用户每次选择信道感知的时间,再结合伯特赛卡斯拍卖机制,提出分布式指数间隔式探索与利用 (decentralized exponentially-spaced exploration and exploitation, dE3) 算法。理论分析表明,所提方法的收益损失期望与时隙数量呈 $O(\log^{1+\delta} t)$ 关系,其中 δ 为超参数。文献 [53] 假设各次要用户间可进行信息交互,提出探索-交互-利用 (explore-signal-exploit repeat, ESER) 算法,并通过理论证明了所提算法的收益损失期望与时隙数量呈 $O(\log^{1+\delta} t)$ 关系。然而,上述方法皆假设次要用户间可进行信息交互,不适用于部分分布式网络。此外,上述文献皆考虑了最大化网络吞吐量,文献 [54] 针对次要用户稳定性分配问题提出可协调稳定分配 (coordinated stable marriage MAB, CSM-MAB) 算法。但该方法假设次要用户具有全频段感知能力,因此无法应用于窄带感知场景。为解决该问题,文献 [55] 提出稳定正交分配 (distributed stable orthogonal configuration for static network, dSOC-SN) 算法,并通过理论分析了静态、动态网络中的冲突数量。

(4) 在异构动态网络中,次要用户数量发生变化且信道统计信息对各次要用户不同。上述 ESER、dE3 算法同样适用于最大化吞吐量场景,dSOC-DN 算法

同样适用于稳定分配场景。但绝大多数算法需次要用户间进行信息交互,因此无法适用于分布式网络场景。考虑到不同网络传输条件和设计目标,目前可应用的成果仍然不足,需进一步展开研究。

1.2.3 待进一步解决的问题

总体来看,基于在线学习的频谱感知与接入方法已经取得了一定的研究成果。然而,这些研究多数聚焦特定场景下的传输效率和可靠性提升,尚未全面涵盖未来不同无线传输条件下的不同挑战。例如,次要用户受天线、处理器等硬件限制,这些限制将影响对频谱状态的感知和信息处理速度。同时,为了满足高能量效率的要求,次要用户在每个时隙中能够接入的信道数量也受到制约。这使得次要用户在每个时隙中面临着需要感知尽可能多数量的信道以获取频谱信息但由于有限硬件能力仅能选择部分信道进行接入的现实矛盾。

在该问题背景下,次要用户在每个时隙中能够接入的信道数量小于能够感知的信道数量,这需要进行两阶段决策:第一阶段需要次要用户基于估计的信道统计信息进行决策选择信道,而第二阶段要求次要用户根据感知到的信道状态做出接入决策。由于接入的决策反过来影响次要用户对信道统计信息进一步的估计及下一阶段的感知决策,两个决策过程之间具有耦合性影响。传统的资源分配方法难以支持次要用户在线学习信道信息,并且不能有效地处理频谱感知与接入之间的复杂相互作用,因此并不适用于解决此类问题,而需建模为两步统计决策过程的设计感知与接入方法。此外,终端轻量化和移动通信严重制约着次要用户的能量供应,这要求未来无线网络不仅要实现高效的频谱利用,还要与能量消耗达成平衡,以获得更优的传输性能。除此之外,随着移动互联网和物联网应用的爆炸式增长,次要用户对多样化业务的需求日益增加,而不同业务对传输质量的要求各不相同,如延迟敏感的应用需要低延迟的网络响应,而数据密集型服务则更侧重于高吞吐量。另外,现有大部分研究未能在信道统计信息未知这一背景下充分考虑多种不同业务类型下的频谱感知与接入方法。综上所述,这些实际的因素不仅影响频谱利用率,还直接关系到次要用户的传输性能、工作寿命和用户满意度等重要指标。因此,本书研究这些特定场景下的频谱感知与接入方法,有助于推动无线通信网络向更加实用和高效的方向发展。

1. 实时信道信息获取难以支持高效接入

目前,已有少量研究结合用频设备有限的硬件能力进一步深入挖掘频谱利用潜力。具体地,考虑次要用户每个时隙中可感知一定数量的信道,但受硬件能力

限制仅能够接入部分信道进行传输,因此带来了多信道感知与部分信道接入问题。在该背景下,次要用户仍然难以实时获取所有信道的状态信息,并且在选择利用频谱资源时需面临两步决策:用频设备需要根据估计的信道信息决定选择哪些信道进行感知,再基于感知结果决定选择哪些信道进行接入。这两步决策过程具有耦合性影响:信道状态的时变性和感知能力的有限性要求用频设备在动态的环境中实时调整感知与接入方法,而信道感知的结果将直接决定接入决策,接入决策又会影响到次要用户对信道状态样本的积累和信道信息的估计,从而影响下一阶段的信道感知决策。因此,多信道感知与部分信道接入并不是一个简单的优化或资源分配问题,而是一个边学习信道信息、边进行决策的问题。

为深入挖掘该问题背景下的频谱利用潜力,文献 [56] 将单个次要用户的频谱感知与接入建模为双层多臂赌博机问题,分别考虑感知和接入决策过程,提出了信道空闲概率的无偏估计的方法,再基于信道状态样本均值的上置信界限设计了频谱感知与接入方法。理论分析表明,该方法的收益损失期望与时隙数量的关系为 $O(\ln t)$ 。但需要注意的是,对于多信道感知与部分接入问题,特别是在多用户集中式网络和分布式网络中,除需要考虑感知与接入两个过程外,还需要考虑用频决策对其他用户的影响。例如,对于集中式多用户网络,需研究如何结合用户实际硬件能力使得多用户协作学习信道信息,以及协调分配信道进行接入;对于分布式多用户网络,需研究如何结合用户实际硬件能力设计能够降低冲突数量的信道接入方法以达到频谱共享。鉴于该问题相关研究较为有限,本书将在既有的单用户研究基础上^[56],针对多用户网络中多信道感知与部分接入在线学习方法展开研究,为充分挖掘频谱资源利用和提高吞吐量提供有效方法。

2. 终端有限能量资源要求高能量效率接入

当次要用户具有有限能量且难以获取信道统计信息时,次要用户需要在时变未知的信道条件和能量约束下寻找最佳的感知与接入方法,以平衡传输性能和工作时长。文献 [57] 对该问题研究的重要性进行了综述,并强调在不确定环境下,能量受限的无线传输技术是未来网络设计的关键研究领域。尽管如此,现有研究主要集中在传统物联网网络和未来智能网络场景。具体地,在物联网网络的背景下,文献 [58] 中描述的问题模型可模拟用户在选择信道过程中获得随机收益并消耗随机能量的情况。该文献提出了基于随机收益和随机能量消耗比值上置信界限的在线学习算法,可通过计算每条信道的平均收益与平均能耗比值的上置信界限进行信道选择。文献 [59] 提出了预算受限的 ε 优先 (budget-limited ε -first) 在线学习算法 (其中, ε 为预先设定的超参数),可模拟用户首先用 $\varepsilon\mathcal{E}$ (其中, $\mathcal{E}$ 表

示用户能量)的能量对所有信道循环式感知接入并计算每条信道的平均收益与平均能耗比值的上置信界限。当 $\varepsilon\mathcal{E}$ 的能量耗尽后,用户用 $(1 - \varepsilon)\mathcal{E}$ 大小的能量持续接入上置信界限最大的信道。此外,文献 [60] 研究了无线网络选择的问题,将用户选择网络接入的过程建模为连续时间多臂赌博机问题,并提出了在线网络选择 (online network selection, ONES) 算法。在该算法中,用户不断估计每个无线网络的平均收益和平均能耗,并通过计算每个网络的指标完成选择的过程。在未来智能网络中,Fouda 等^[61]考虑了可见光频段和无线电频段(包括 2.4GHz、5.25GHz 及 38GHz)下单个用户的信道感知接入统计决策问题。由于各链路衰落特性和空闲概率对次要用户未知,为实现无线链路可靠传输,作者基于上置信界限算法和采样原理,分别设计基于上置信界的能量感知信道选择 (energy-aware band selection with upper confidence bound, EABS-UCB) 算法和基于汤普森采样的能量感知信道选择 (energy-aware band selection with Thompson sampling, EABS-TS) 算法,并通过仿真验证了算法的有效性。但所提出的算法未给出理论性能,并且无法适用于多个用户的场景。

值得说明的是,在认知无线网络的研究领域中,一个值得关注的课题是次要用户在未知信道统计特性且能量有限的条件下,如何设计有效的频谱感知与接入方法。现有研究虽已在此方面取得进展,但鲜有成果是在认知无线网络这一具体背景下展开的。认知无线网络的特殊性在于,次要用户的收益与能量消耗之间存在密切的相关性:次要用户在感知到信道空闲并传输信息时,虽能获得收益,但伴随的是较高的能量消耗;反之,若感知到信道被占用,虽无法传输信息,但感知过程中的能量消耗相对较低。然而,大多数研究假设收益与能量消耗相互独立,忽略了这种相关性,这限制了其在认知无线网络中的应用评估。特别是在多用户参与的集中式和分布式网络环境中,相关研究尤为缺乏。鉴于此,相关研究亟须针对认知无线网络中的能量约束问题进行深入探讨,为提升频谱资源利用率和能量效率提供切实有效的解决方案。

3. 多样化服务质量要求亟须高传输效能接入

在次要用户面临多样性业务和追求高传输效能的接入需求时,需根据实时业务类型和信道统计信息,动态调整频谱感知与接入方法。文献 [62] 针对多频段信道选择问题进行了深入研究,构建了基于不同频段特性和用户服务质量要求的回报函数,并利用 Q-learning 算法设计了多智能体信道选择方法。文献 [63] 则聚焦单信道场景,探讨了如何调整发射参数以满足不同的服务质量要求,通过平均端到端意见分数 (mean opinion score, MOS) 来衡量传输性能,并提出了一种基于

深度强化学习的资源分配技术,有效降低了传输延迟。文献 [64] 研究了如何调整发射参数以获得尽可能大的用户体验,并基于深度确定性策略梯度算法设计了能够提高传输质量和降低失真的方法。尽管目前在满足服务质量要求的频谱感知与接入方法方面已有一定研究,但仍存在不足之处。

首先,传输性能、服务质量要求与信道特征之间内在联系的研究尚需深化。现有研究多采用平均端到端意见分数或特定要求的效能函数来表征传输效能,这在处理复杂决策场景(如平衡传输性能、能量消耗和传输成本等)时难以灵活应用,因此亟须开发一种更通用、可灵活构建的效能函数,以适应这些多样化的需求。

其次,不同的传输业务对传输性能有不同要求,而已有研究大部分是针对某单一服务质量要求展开频谱感知与接入方法设计的。如何将多种服务质量差异纳入统一的频谱决策框架中进行考虑,并设计出灵活的自适应决策方法以最大化传输性能,对提高频谱利用灵活性有着重要意义。鉴于该方面的研究在未知信道统计信息下的多用户网络中较为有限,有必要进一步针对不同传输业务下的频谱感知与接入方法设计开展研究,为提高频谱决策灵活性和传输性能提供有效方法。

1.3 研究内容与组织结构

1.3.1 研究内容

针对频谱统计信息未知和时变条件下用频设备的高效接入,本书围绕无线网络发展所面临的三大挑战,即实时信道信息获取难以支持高效接入、终端有限能量资源要求高能量效率接入、多样化服务质量要求亟须高传输效能接入,基于在线学习理论设计相应的频谱感知与接入方法,可有效提升频谱利用效率、能量效率及传输效能,具体的研究内容和思路概述如下。

1. 基于在线学习的多信道感知与部分信道接入方法

首先,开展次要用户多信道感知与部分信道接入下的方法研究,包括多用户集中式网络和分布式网络两部分。

(1) 针对集中式网络传输数据量最大化,将频谱感知与频谱接入视为两步决策过程,构建了基于双层多臂赌博机在线学习理论的问题模型。同时,提出了基于信道空闲概率估计的上置信界限排序的多信道感知与接入方法,通过理论分析了数据量损失与时隙数量的关系,并通过仿真验证了所提方法在不同用户数量、信道数量、检测概率、虚警概率下的有效性。

(2) 针对分布式网络传输数据量最大化, 综合考虑了次要用户频谱感知与接入两步决策及多用户间频谱利用的冲突, 构建基于多用户双层多臂赌博机在线学习理论的问题模型。同时, 提出了基于信道空闲概率估计的上置信界限排序和随机接入的多信道正交感知与接入方法, 通过理论分析了数据量损失与时隙数量的关系, 并通过仿真验证了所提方法在不同用户数量、信道数量、检测概率、虚警概率下的有效性。

2. 基于在线学习的能量约束频谱感知与接入方法

其次, 开展次要用户在能量约束下基于在线学习的频谱感知与接入的方法研究, 包括单用户、集中式网络和分布式网络三部分。

(1) 针对单用户场景, 研究了如何在未知信道空闲概率信息和信道质量统计信息且受有限能量约束下获得最大数据量。为兼顾感知与接入信道过程中能量消耗、自身剩余能量, 以及对各信道统计信息的在线学习, 构建基于背包问题 (knapsack problem, KP) 和组合多臂赌博机在线学习理论的问题模型。同时, 提出了基于吞吐量与能量消耗比值估计的上置信界限排序和贪心策略的频谱感知与接入方法, 通过理论分析了数据量损失与能量大小间的关系, 并通过仿真验证了所提方法在不同信道感知接入数量、不同能量大小下的有效性。

(2) 针对集中式网络, 研究了中心节点如何在未知信道质量统计信息、各个次要用户受有限能量约束下获得最大数据量。综合考虑各次要用户的剩余能量、感知与接入不同信道带来的能量消耗, 以及各信道统计信息的在线学习, 构建基于多维背包问题和多臂赌博机在线学习的问题模型。同时, 提出了基于吞吐量与能量消耗比值估计的上置信界限排序和匈牙利算法的频谱感知与接入方法, 并通过仿真验证了所提方法在不同用户数量、不同能量大小下的有效性。

(3) 针对分布式网络, 研究了如何在未知信道质量统计信息、各次要用户受有限能量约束、各次要用户间无法信息交互且无法知道其他次要用户感知接入决策、无法得知其他次要用户由于能量耗尽退网情况下获得最大数据量。为兼顾各用户剩余能量、感知接入过程中的能量消耗、对各信道统计信息的在线学习、接入相同信道时的冲突、用户数量的动态变化, 构建基于背包问题的信道分配问题 and 多用户多臂赌博机在线学习理论的问题模型。同时, 提出了基于吞吐量与能量消耗比值估计的上置信界限排序和伯特赛卡斯拍卖算法的频谱感知与接入方法, 通过理论分析了数据量损失与各次要用户能量限制之间的关系, 并通过仿真验证了所提方法在不同用户数量、不同能量大小下的有效性。

3. 基于在线学习的多类型服务质量要求频谱感知与接入方法

最后,开展次要用户在具有多个不同服务质量要求下基于在线学习的频谱感知与接入方法研究,包括单用户、集中式网络和分布式网络三部分。

(1) 针对单用户分类业务模型,研究了如何在未知信道特征统计信息、不同业务类型及服务质量要求下获得最大累积传输效能。为解决该问题,本书研究了如何有效考虑评估业务类型及服务质量要求、信道特征,以及接入信道获得的传输效能之间的关系,构建多属性决策(multiple attribute decision making, MADM)和上下文多臂赌博机(contextual MAB, CMAB)在线学习的问题模型。同时,提出了基于传输效能估计的上置信界限排序的频谱感知与接入方法,分析了累积传输效能损失(即信道特征统计信息已知下的最大累积传输效能期望与所设计的方法获得的累积传输效能期望间的差值)与时隙数量间的关系,并通过仿真验证了所提方法在不同业务类型下的有效性。

(2) 针对集中式网络,研究了中心节点如何在未知信道特征统计信息、各次要用户具有不同业务类型及服务质量要求下获得最大累积传输效能。为解决该问题,本书研究了中心节点如何综合考虑各次要用户的业务类型、估计的信道特征统计信息进行信道分配,构建了基于组合优化问题和上下文多臂赌博机在线学习理论的问题模型,提出了基于传输效能估计的上置信界限排序和匈牙利算法的频谱感知与接入方法。

(3) 针对分布式网络,研究各用户如何尽可能公平地获得传输效能,构建了基于多用户上下文多臂赌博机的在线学习问题模型,提出了基于传输效能估计的上置信界限排序和正交循环策略的频谱感知与接入方法。最终,通过理论分析了累积传输效能损失与时隙数量的关系,并通过仿真验证了所提算法的有效性。

1.3.2 组织结构

本书共分为6个章节,组织结构框图如图1.2所示,具体说明如下。

第1章为绪论,主要介绍频谱感知与接入的研究背景、意义和研究现状,并给出本书的研究内容和组织结构。

第2章对频谱感知与接入及其面向多臂赌博机在线学习理论进行介绍,奠定后续研究的理论基础。

第3章研究了基于双臂赌博机多信道感知与部分信道接入方法,内容包括集中式网络和分布式网络两部分。首先,研究了在集中式网络中如何通过设计频谱感知与接入方法获得尽可能大的数据量。同时,基于双层多臂赌博机在线学习理论构建了问题模型,根据信道空闲概率估计的上置信界限,设计了基于其排序的

频谱感知与接入方法,通过理论分析了算法的数据量损失与时隙数量呈 $O(\ln t)$ 的关系,给出了算法的复杂度,并进一步通过仿真验证了所提算法在不同用户数量、不同信道数量、感知接入数量、不同检测概率、虚警概率条件下的有效性。其次,研究了分布式网络中如何获得尽可能大的数据量。同时,基于多用户双层多臂赌博机在线学习理论构建了问题模型,设计了基于随机接入和信道空闲概率估计的上置信界限排序的频谱感知与接入方法,分析得到了算法的数据量损失与时隙数量为 $O(\ln t)$ 的关系,给出了算法的复杂度,并通过仿真验证了所提算法在不同用户数量、不同信道数量、不同检测概率、虚警概率条件下的有效性。

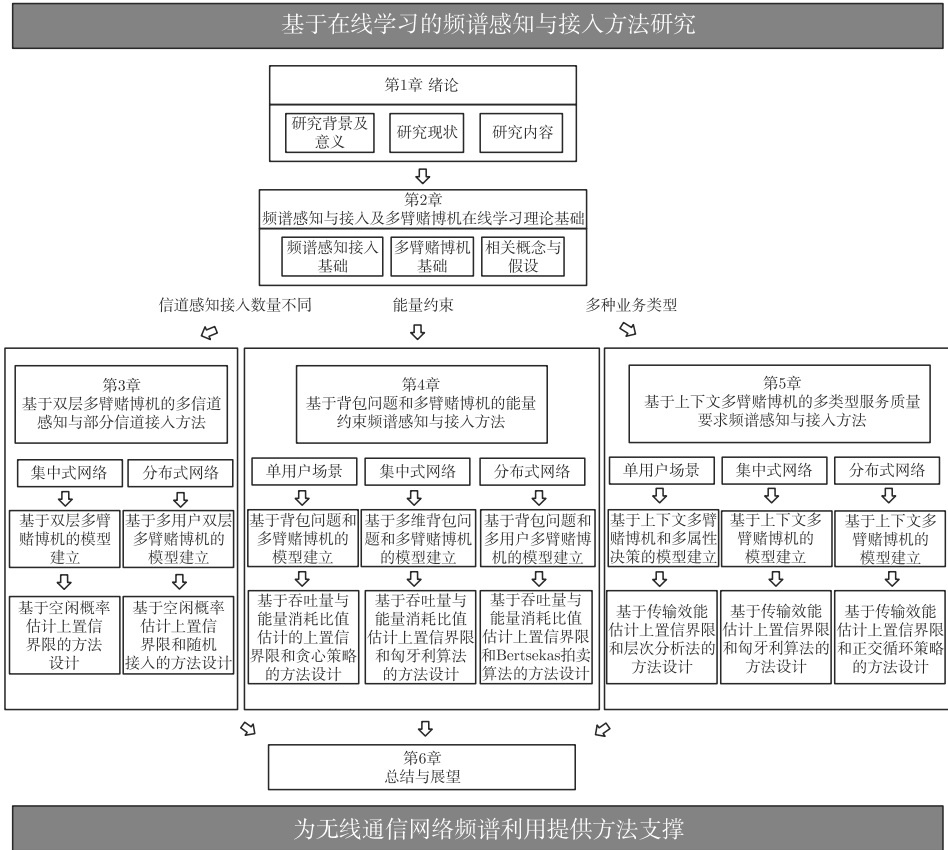


图 1.2 本书组织结构框图

第 4 章研究了基于背包问题和多臂赌博机能量约束条件下的频谱感知与接入方法,内容包括单用户场景、集中式网络和分布式网络三部分。首先,研究了单用户能量受限下如何获得尽可能大的数据量。同时,基于背包问题和多臂赌博机在线

学习理论构建了问题模型,分析了信道统计信息已知下的最优感知与接入方法和数据量。根据平均吞吐量与平均能量消耗比值的上置信界限,设计了基于其排序的频谱感知与接入方法,通过理论分析了算法的数据量损失与能量大小呈亚线性关系,给出了算法复杂度,并结合仿真验证了所提算法的有效性。其次,研究了集中式网络能量受限下如何获得尽可能大的数据量。同时,基于多维背包问题和多臂赌博机在线学习理论构建了问题模型,设计了基于贪心策略求解多维背包问题、匈牙利算法求解组合优化问题、平均吞吐量与平均能量消耗比值的上置信界限排序的频谱感知与接入方法,给出了算法的复杂度,并仿真验证了所提算法在不同用户数量、不同信道衰落模型下的有效性。最后,研究了分布式网络能量受限下如何获得尽可能大的数据量。同时基于信道分配和多用户多臂赌博机在线学习理论构建了问题模型,设计了基于伯特赛卡斯拍卖算法求解信道分配问题、平均吞吐量与平均能量消耗比值的上置信界限排序的频谱感知与接入方法,通过理论分析了所提算法数据量损失与各次要用户能量大小呈亚线性关系,给出了算法的复杂度,并通过仿真验证了所提算法在不同用户数量、不同信道衰落模型下的有效性。

第5章研究了基于上下文多臂赌博机的多类型服务质量要求的频谱感知与接入方法,内容包括单用户场景、集中式网络和分布式网络三部分。首先,研究了单用户分类业务场景下如何获得尽可能大的累积传输效能。同时基于MADM理论和上下文多臂赌博机在线学习理论构建了问题模型,设计了基于层次分析法(analytic hierarchy process, AHP)量化信道特征对服务质量影响、岭回归分析多未知参数估计、平均传输效能上置信界限排序的频谱感知与接入方法,通过理论分析了所提算法的累积传输效能损失与时隙数量呈 $O(\ln t)$ 的关系,给出了算法的复杂度,并仿真验证了所提算法的有效性。其次,研究了集中式网络分类业务场景下如何获得尽可能大的累积传输效能。同时,基于组合优化和上下文多臂赌博机在线学习理论构建了问题模型,设计了基于匈牙利算法求解组合优化问题、平均传输效能的上置信界限排序的频谱感知与接入方法,给出了算法的复杂度,并通过仿真验证了所提算法在不同传输要求下的有效性。最后,研究了分布式网络分类业务场景下如何获得尽可能公平的累积传输效能。同时基于多用户上下文多臂赌博机在线学习理论构建了问题模型,设计了基于公平性原则、正交循环接入和平均传输效能的上置信界限排序的频谱感知与接入方法,通过理论分析了算法的累积传输效能损失与时隙数量呈 $O(\ln t)$ 的关系,并通过仿真验证了所提算法在不同传输要求下的有效性。

第6章对本书研究进行总结,概括本书的主要创新点和结论,并对后续的研究方向提出展望。

第2章

频谱感知与接入及多臂赌博机 在线学习理论基础

本章分别对频谱感知与接入中的关键技术及多臂赌博机理论基础进行介绍。本节结构如下：2.1 节对本书研究内容的一些相关概念和基本假设进行说明；2.2 节介绍频谱感知接入的概念、主要用户信道占用时隙结构模型、频谱感知接入框架以及频谱感知接入关键技术；基于上述基础，2.3 节介绍面向频谱感知接入的多臂赌博机相关理论模型、统计决策以及性能分析方法；2.4 节总结本章内容。

2.1 相关概念与假设

2.1.1 相关概念

1. 中心节点

中心节点是集中式认知无线网络中具有控制或决策功能的重要节点，其能够通过节点交换控制信息收集网络中各节点的状态信息、发送数据请求，根据传输目标采用一定的优化算法（如图论、分配算法、匈牙利算法、博弈论等）制定频谱分配策略，协调整个网络中的频谱分配和访问控制，并根据需要实时调整和优化传输性能。中心节点可有效地降低信道接入冲突、提高频谱利用率、提供更加可靠和高质量的无线传输服务^[65]。

2. 信道统计信息

统计信息是指关于无线通信信道特性的各种数据和信息，这些信息通常包括信道空闲概率、信道容量和信道质量等指标。由于未来无线网络趋于复杂动态化

发展,次要用户通常难以实时获取所有信道状态信息。此外,在初始入网阶段,次要用户难以获取信道的统计信息。这些假设符合实际场景中的无线通信系统。为最大化传输性能,需设计频谱感知与接入在线学习方法,以使次要用户在每个时隙中能够通过选择信道感知其实时信息、基于感知结果选择信道进行接入、根据感知或传输结果更新对信道统计信息的估计,从而在与无线环境不断交互的过程中获得更优的传输性能。

3. 后悔/数据量损失/累积传输效能损失

在多臂赌博机理论中,后悔 (regret) 是一个关键概念,衡量的是如果始终选择最佳臂能获得的理论最大收益与实际通过特定策略获得的收益之间的差距,即用于衡量某种策略在一定时间内相对于最优策略的性能损失。理想情况下,一个有效策略的后悔值会随着时间的推移而逐渐趋于平稳,即越来越接近始终选择最佳臂的理论最优性能。本书同样采用该工具来衡量所提算法的有效性。为更加直观地表现研究内容,在非对称感知接入和能量受限条件下的频谱感知与接入策略中,本书采用数据量损失这一概念衡量统计决策方法在一定时间内获得的数据量与理论上最优数据量间的差值。在分类业务驱动下的频谱感知与接入策略中,本书采用累积传输效能损失这一概念衡量统计决策方法在一定时间内获得的累积传输效能与理论上最优累积传输效能间的差值。

2.1.2 主要假设

1. 关于信道实时质量探测的假设

本书假设次要用户发射端可通过向接收端发送数据包,并接收反馈信息获取实时信道质量,例如,在文献 [66] 中,当次要用户感知信道为空闲后,以最大功率向接收端发送请求发送 (request-to-send, RTS) 数据包。接收端根据接收到的 RTS 数据包估计信道实时质量,然后向发射端发送包含信道质量信息的数据包。通过这种方式,次要用户的发射端可以实时了解并适应当前的信道状况,从而优化其通信策略。这种方法强调了实时信息交换在动态无线通信环境中的重要性,为有效利用可用资源提供了一种有效途径。

2. 关于发射机——接收机信道感知接入决策同步的假设

考虑分布式发射机—接收机传输对和多条信道组成的无线通信网络。在该网络中进行信息传输前,发射机和接收机需通过控制信息调整至相同频段。该内容并非本书研究的重点内容,因此,本书假设发射机在选择信道进行感知接入时,

可与接收机通过某一技术手段（如公共控制信道）传递控制信息，并调整至相同频段^[67]。

3. 关于信道状态感知的假设

本书研究的重点是频谱感知与接入策略，属于媒体接入控制层的研究内容。因此，本书假设次要用户发射端可通过某种信道感知技术（如能量检测法）获取实时信道状态，而不关注某一具体技术。

关于本书的研究思路，做以下说明：在多臂赌博机问题研究中，为了评估所设计的统计决策方法的有效性，通常需要将设计方法的性能与统计信息已知下的最优方法性能进行对比。因此，本书在对某一问题展开研究时，首先研究统计信息已知下的最优方法，并以此作为性能基准；其次研究并设计未知信息下的统计决策方法^[68]。

2.2 频谱感知与接入基础

作为提升频谱效率的重要使能技术，认知无线网络中与主要用户共存的次要用户可通过实时感知主要信道占用状态，寻找可用的时频资源，与其他次要用户协作或者竞争，伺机接入信道完成信息传输。频谱是一种有限的资源，需要考虑各次要用户间公平合理的分配和利用，因此，在认知无线网络中需要对频谱资源进行有效的管理和调度，通常将频段划分为多个不相交的信道（如 IEEE 802.22 协议标准^[69]中将频段划分为多个没有交集的信道，每条信道带宽为 6MHz、7MHz、8MHz）。为实现最佳传输性能，直观的方式是次要用户感知所有信道，再优选信道进行接入。然而，实际上认知无线网络可工作在由多个主要用户网络的频域资源组成的较宽频段范围中，如电视频段（5MHz~1GHz）、工业的、科学的、医学的（industrial scientific medical, ISM）部分频段（2.40~2.48GHz）等^[70]，次要用户通常难以实时感知所有信道：从硬件限制的角度来看，较大的带宽对次要用户的数据处理速率要求较高，使得较大的带宽需要次要用户具有信号处理功能强大的硬件设备，这将会增加硬件成本^[71]；从感知时间限制的角度来看，较宽频段的感知将导致检测时间的增加，从而增加了时延和降低了网络性能；从能量消耗因素的角度来看，较宽频段的感知将会消耗大量的能量，这对于依赖电池供电的移动设备来说难以实现。上述实际因素产生了频谱感知中最为典型的多信道频谱感知的问题，大量的研究集中于次要用户每个时隙中能够感知接入部分信道的场景，研究如何选择信道进行感知和接入的策略，以获得尽可能优的传输性能。

用户信道占用结构模型按照时间可分类为时隙结构模型和非时隙结构模型,如图 2.1 所示。在时隙结构模型中(图 2.1 (a)),主要用户信道占用状态在每个时隙中保持不变,并在每个时隙中以独立或相关(如马尔可夫信道模型)的方式变化。次要用户与主要用户保持同步且以同样时隙的方式进行感知接入。在每个时隙初始时,次要用户对信道进行感知并基于感知结果决定是否利用信道资源。在非时隙结构模型中(图 2.1 (b)),主要用户可以随时使用频谱。在该模型下,可通过对次要用户施加冲突约束将非时隙结构模型转换为时隙结构模型^[72]。考虑到时隙结构模型便于工程实现,目前大多数标准化工作及研究均围绕该模型展开,本书专注于研究时隙结构模型下的信道感知与接入。

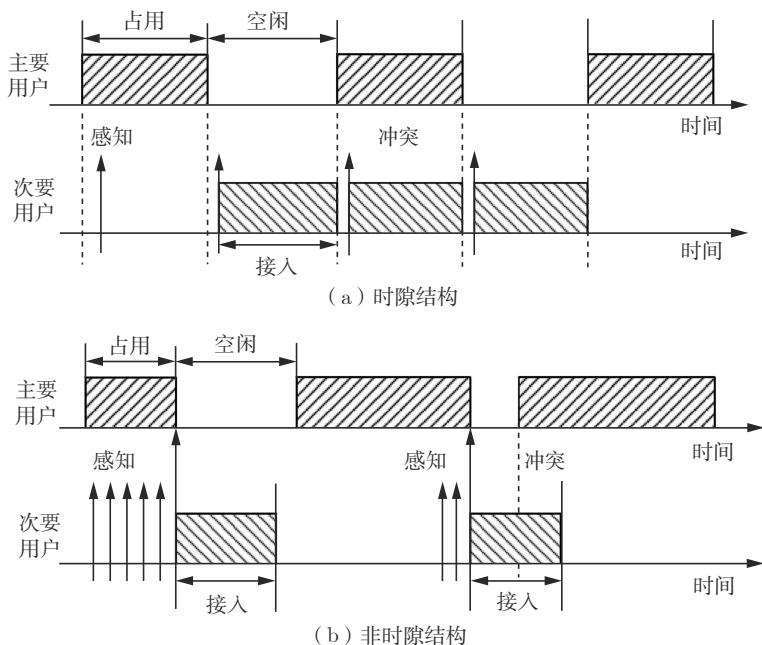


图 2.1 用户信道占用结构模型

在主要用户网络时隙结构模型下,认知无线网络中频谱感知与接入框架可分为两类,分别为并行感知框架和序贯感知框架,如图 2.2 所示。在并行感知框架中(图 2.2 (a)),次要用户在每个时隙初始时仅选择部分信道进行感知,当信道感知为空闲时,次要用户接入信道并进行信息传输;若信道感知为占用,次要用户放弃该时隙中的传输机会。在序贯感知框架中(图 2.2 (b)),次要用户按照某种标准将信道排序,依次序对信道进行感知,并在满足某种规则的条件下停止感知并选择接入。考虑到并行感知框架易于工程实现,目前大多数标准化工作围绕

该模型展开，本书专注于研究并行感知与框架下的信道感知与接入问题。

在上述功能实现过程中，次要用户与无线环境的交互可视为一个认知系统。文献 [73] 对频谱感知与接入中的关键技术进行如下分类：信道感知、信道切换、共享与接入，以及控制信道等，如图 2.3 所示。其中，信道感知负责获取可用的无线环境信息；信道接入是次要用户获取信道信息后对无线资源的利用；信道切换负责主要用户占用信道或者信道质量无法承载传输任务时，令次要用户停止对现有信道的接入而伺机接入其他信道；信道共享是次要用户与主要用户及其他次要用户共享可用的频谱资源；控制信道是确定合适的时间和频段使得各次要用户进行信息交互，以达到感知信息共享、接入策略共享、建立通信链路等目的。下面分别对频谱感知与接入中的关键技术进行介绍。

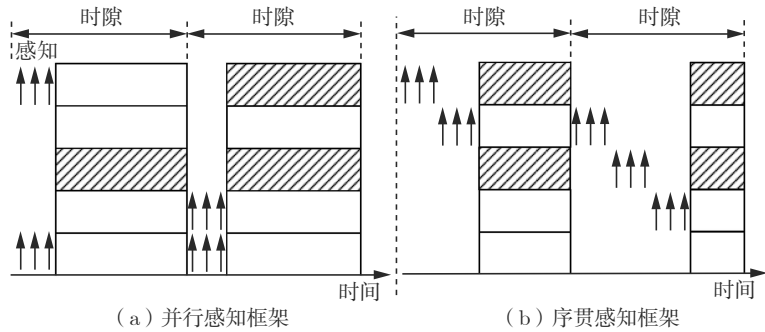


图 2.2 感知接入框架

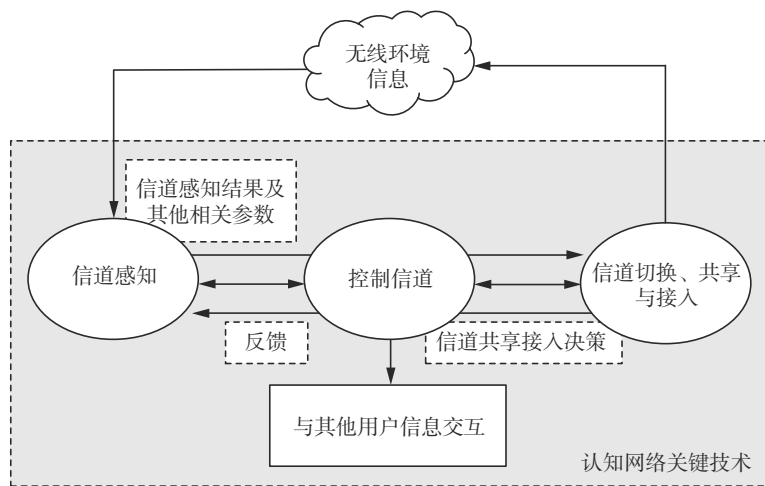


图 2.3 频谱感知与接入关键技术图

2.2.1 信道感知

信道感知是频谱感知与接入的前提和基础,其主要功能是检测某频带内是否存在主要用户活动,即时域、频域、空域上的空闲频谱,分析频谱可用性并评估频谱的统计学特征,为后续频谱共享和接入提供必要依据^[74]。

信道感知是指在通信系统物理层中,通过对电磁频谱的监测和分析来识别当前环境中的无线信号和频谱使用情况。其本质可视为一个二元假设问题,即次要用户采用一定技术手段和方法对特定频段进行监测,然后基于某一标准进行判决以判断检测信道的可用状态。一个感知系统通常由无线电前端、信号处理单元、控制单元及用户接口组成,其中,无线电前端包括天线、射频接收器和模拟前端电路,用于接收无线电信号并将其转换为可处理的电信号;信号处理单元负责对接收到的信号进行数字化处理和分析;控制单元可设置感知策略,用于控制和协调感知系统的操作;用户接口可用于调整感知参数、显示感知结果等。目前,信道感知技术已得到广泛的研究,按照信道检测的自主性,感知方法可分为辅助信道感知和独立信道感知两大类^[75]。

(1) 在辅助信道感知技术中,信道空闲状态信息并非来自于认知无线网络内部,而是来源于主要网络或者其他实体辅助,其具体实现需要设计额外的频谱管理协议,如增加通信系统的控制信息开销使得次要用户从主要网络或实体辅助中获取信道空闲信息,其具体技术包括基于频谱租赁、基于频谱政策、基于频谱资源数据库等。这些方法有助于提高信道感知的可靠性和精度,较好地保护主要用户网络不受干扰,从而支持更有效的频谱管理和资源分配策略,对于实现动态频谱共享具有重要意义。然而,该方法需要次要网络连接实体辅助,并增加了通信系统的控制信息成本^[76]。

(2) 在独立信道感知技术中,次要用户通过数字信号处理方法感知信道空闲信息,按照参与的用户数量可分为协作信道感知和单用户信道感知。① 协作信道感知是利用空间分集和用户多样性,将多个次要用户本地检测结果按照一定的规则融合后判决信道可用状态。其主要技术包括集中式协作检测、分布式协作检测和基于中继的协作检测^[77-79]。② 在单用户信道感知中,按照感知能力可分为窄带感知和宽带感知。其中在窄带感知中,次要用户每次仅能感知一个窄带信道,主要技术包括能量检测法、匹配滤波器检测法和循环平稳特征检测法。具体地,在能量检测法中,次要用户通过接收信号并将其转换为电信号,然后测量接收到信号的能量,即感知的信号在一个时间段内的功率总和^[80]。判断信道是否被占用可建模为一个二元假设问题,即

$$\begin{cases} H_0 : r(t) = n(t) \\ H_1 : r(t) = s(t) + n(t) \end{cases} \quad (2.1)$$

其中, H_0 表示主要用户不存在; H_1 表示主要用户存在; $r(t)$ 表示次要用户检测的信号; $s(t)$ 表示接收到的主要用户的信号; $n(t)$ 表示信道的高斯白噪声。次要用户进行信道检测时, 首先将接收的信号输入相匹配的带通滤波器, 对采样的信号求平方和并与检测门限 ε 比较以判断信道状态。对于时隙 t 下的信道 i , 其感知结果可表示为

$$Z_i(t) = \begin{cases} \sum_{j=1}^I n_j^2, & \text{空闲} \\ \sum_{j=1}^I (s_j + n_j)^2, & \text{占用} \end{cases} \quad (2.2)$$

其中, $Z_i(t)$ 表示次要用户在时隙 t 中感知阶段对信道 i 采样的平方和; n_j 和 s_j , $j \in [1, I]$ 分别表示噪声信号和主要用户信号的采样值; I 表示感知阶段的总采样数。考虑主要用户信号为复数相移键控信号, 噪声为 IID 的循环对称复高斯噪声, 则次要用户在信道占用时能够检测出信道实际占用状态的概率, 即检测概率可表示为

$$P_d(\varepsilon, I) = P(Z_i(t) \geq \varepsilon | \text{信道占用}) = \frac{1}{2} \operatorname{erfc} \left(\frac{\varepsilon - I - I\gamma_p}{2\sqrt{2}\sqrt{\frac{I}{2} + I\gamma_p}} \right) \quad (2.3)$$

其中, ε 表示检测门限; γ_p 表示次要用户接收到的主要用户信号的信噪比; $\operatorname{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2} dt$ 表示标准高斯分布的错误概率补函数。主要用户未占用信道但次要用户判定信道为占用的概率, 即虚警概率可表示为

$$P_f(\varepsilon, I) = P(Z_i(t) \geq \varepsilon | \text{信道空闲}) = \frac{1}{2} \operatorname{erfc} \left(\frac{\varepsilon - I}{2\sqrt{2}\sqrt{\frac{I}{2}}} \right) \quad (2.4)$$

能量检测法具有简单且易于实现、计算复杂度低、无须知道主要用户信号先验信息等优点, 在工程和学术研究中得到了广泛应用; 缺点为无线环境中噪声方差的不确定性将直接影响检测性能, 该方法的判决门限受噪声影响而难以确定, 该方

法无法对噪声、干扰和信号进行区分（从而难以准确判断被检测信号）。匹配滤波器检测法中次要用户事先具有目标信号的特征，将监测的信号与匹配滤波器进行卷积运算，根据滤波器输出的幅值或能量判断信道状态。该方法的优点是次要用户能够利用信号的特征精确地感知信道，在干扰较强的无线环境中能够有效地区分目标信号和干扰信号^[81]；缺点是用户需要已知一定的光验信息，或对信号特征具有较强的感知能力。循环平稳特征检测法的主要思想是基于信号的循环平稳特性提取信号特征，通过分析信号谱相关函数中循环频率的特性来确定主要用户信号的存在^[82]。

在宽带感知中，次要用户对整个监测频段信号进行采样，然后通过相关重构法重建原信号，再根据功率或能量比较对各子信道空闲情况进行判决。其主要技术包括压缩感知和小波变换等。相较于窄带感知，宽带感知需要次要用户的射频前端有更强的信号处理能力，其硬件实现复杂（如高分辨率高速 A/D 转换器），价格昂贵，并且需要更加可靠有效的信号处理算法提高频谱检测的准确性^[83]。

2.2.2 信道切换

信道切换又称频谱移动，是在次要用户工作的信道质量无法满足传输需求或主要用户在该信道上出现时，次要用户停止信息传输，而寻找其他的信道进行接入的过程。次要用户信息传输被打断时，将会造成较大的时延，因此良好的信道切换方式能够保证用户在较小延迟情况下传输信息。文献 [73] 将信道切换分为静态切换、被动式切换、主动式切换和杂交式切换四种方式（图 2.4）。① 静态切换方式假设次要用户在固定信道上进行传输，当主要用户出现时停止传输，直到信道空闲时再重新传输。该类方法具有延迟性高、缺乏灵活性等缺点。② 被动式切换方式是指当次要用户检测到主要用户出现时进行信道切换。该类方法可能造成较高切换延迟，不适用于实时类型应用。③ 主动式切换方式是指次要用户利用在线学习方法预测信道空闲状态，主动选择信道进行传输。该类方法具有较低延迟、较少切换次数等优点，但对在线学习方法有较高要求。④ 杂交式切换方式是指次要用户利用主动式信道感知和被动式信道切换。在该方法中，次要用户通常存储备用信道列表并进行信道感知，当次要用户需进行信道切换时，其直接从信道列表中选择信道进行切换。该类方法能够达到高延迟与高复杂度间的折中。

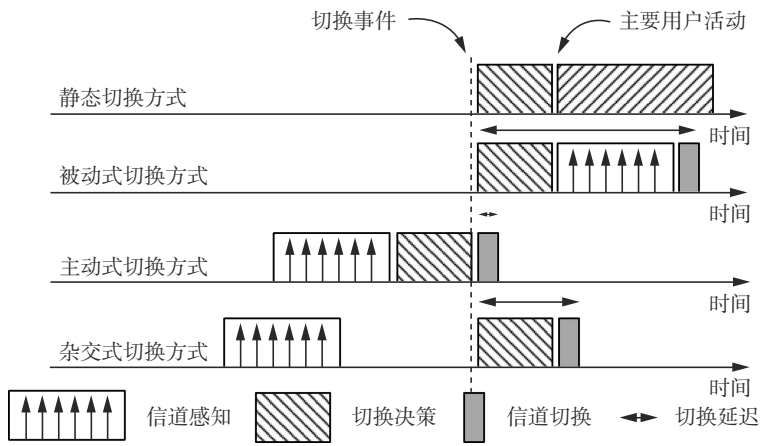


图 2.4 不同信道切换过程

2.2.3 信道共享与接入

信道共享与接入是指次要用户基于感知到的无线环境信息，与主要用户和其他次要用户共用无线资源的过程，是提高频谱利用率的关键技术。信道共享的分类如图 2.5 所示。

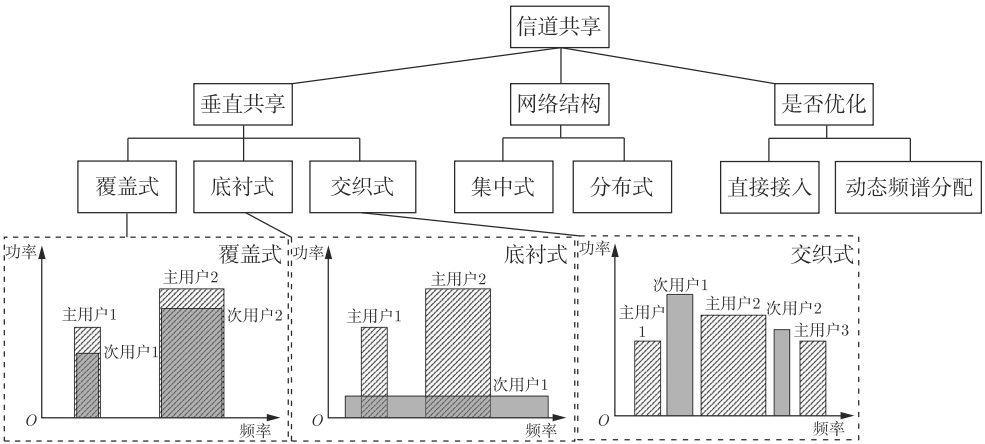


图 2.5 信道共享的分类

(1) 按照主要用户网络频谱利用方式进行划分,信道共享为垂直共享方式,它可分为覆盖式 (overlay)、底衬式 (underlay) 和交织式 (interweave) 三种类型。
① 覆盖式又称为干扰减轻,允许次要用户与主要用户在相同信道上进行传输,但次要用户需事先知道主要用户的码本,采用干扰抵消和消除技术 (如编码技术和

中继技术)对主要用户信息进行译码或转发^[84]。② 底衬式又称为干扰控制。在该方式下,次要用户在不干扰主要用户的前提下与主要用户使用同一时频资源。为达到该目的,次要用户通常采取控制发射功率、扩频或超带宽技术,使得对主要用户的干扰低于干扰门限^[85-86]。该种方式的难点在于如何准确衡量次要用户是否对主要用户产生较大干扰,实际应用中将面临较多问题。③ 在交织式中,次要用户通过对信道进行周期性感知,获取空闲时频资源并进行信息传输。当主要用户在该信道上进行传输时,次要用户需及时退避以避免对主要用户的干扰^[87]。由于交织式无须事先知道主要用户码本信息以及主要用户干扰门限,技术实现简单,可用于频谱感知接入的频分多址、时分多址、正交频分多址系统中,是目前频谱感知接入中最为广泛应用与研究的模型。

(2) 按照网络结构进行划分,信道共享与接入可分为集中式网络共享和分布式网络共享。① 在集中式网络共享中,存在一个中心节点获取系统中各个次要用户的信息,对各个次要用户的频谱感知和接入进行统一决策,可获得全局最优的网络性能,但方法复杂度高、信令开销大;② 在分布式网络共享中,无中心节点对各次要用户进行调度,各次要用户可通过公共控制信道与邻居节点进行局部信息交互或各自独自进行频谱决策,与集中式网络相比,其减少了信息交互,但次要用户间可能会产生干扰,严重影响网络性能。

(3) 按照网络中是否考虑优化进行划分,信道共享与接入可划分为直接接入和动态频谱分配^[88]。① 直接接入是指次要用户不考虑系统是否具有全局最优的频谱分配,而仅考虑自身最优的接入方案,其方式可分为两类,分别为基于竞争的方式和基于协调的方式,其中基于竞争的方式是指多个次要用户对信道进行争用,赢得信道机会的次要用户进行信息传输;基于协调的方式是指次要用户与其相邻节点共享信道决策信息,以提高系统整体性能。② 动态频谱分配是指次要用户采用优化、博弈论、图论、启发式等方法,自适应调整发射参数复用无线资源,以满足不同的网络设计目标。

2.2.4 控制信道

相较于传统无线网络,频谱感知与接入中需进行大量信息交互,如时钟同步、信道感知决策信息、信道接入决策信息、信道感知融合、分布式用户间信道协商等。目前实现控制信息交互大致可分为基于专用控制信道、基于时隙拆分、基于跳变序列、基于多汇聚的控制信息管理四种方法^[73],具体描述如下。

(1) 在基于专用控制信道的方法中,通常设置一个专用的控制信道用于信令交互,次要用户需配备额外的收发器工作于控制信道所在频段,在传输过程中同

时侦听控制信息^[89-90]。其协调机制如图 2.6 所示。在该类方法中，次要用户发射端将可用信道通过 RTS 数据包发送给接收端，接收端获得双方皆可用的信道，再通过清除发送（clear-to-send, CTS）数据包发送给发射端，从而确认传输信道。该类方法中，各次要用户无须保持严格同步，但用户数量较大时，将会使得控制信道产生拥塞降低信道利用率。



图 2.6 基于专用控制信道的协调机制

（2）在基于时隙拆分的方法中，一个时隙被拆分为信令控制阶段和数据传输阶段^[91]。其协调机制如图 2.7 所示。在信令控制阶段，所有次要用户对公共信道进行侦听，需发送数据的次要用户争用信道，完成 RTS/CTS 信令交互后在数据传输阶段完成数据传输。该类方法的主要缺点为在控制阶段仅在控制信道上进行信令交互，其他信道处于空闲状态，造成频谱资源的浪费。

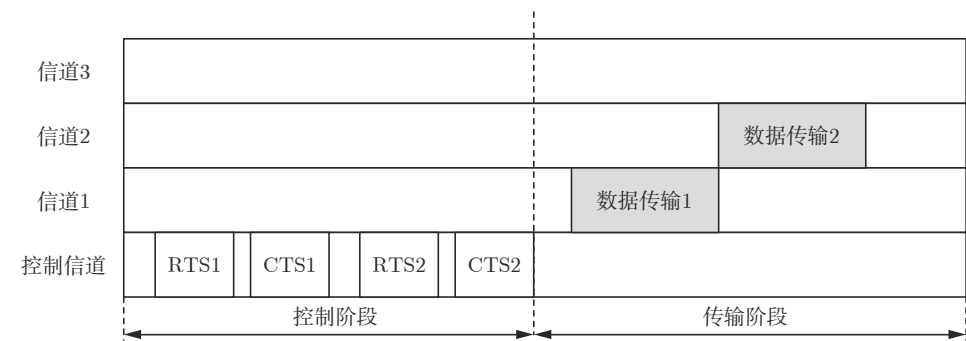


图 2.7 基于时隙拆分的协调机制

（3）在基于跳变序列的方法中，各次要用户保持严格同步且共有一相同跳频序列^[92]。其协调机制如图 2.8 所示。在每个时隙中，各次要用户按照跳频序列跳转到信道进行侦听。当发射端向接收端发送数据时，首先发送 RTS，若接收端接收到报文，则在当前信道上回复 CTS 给发射端开始数据传输过程。在数据传输时，收发两端停止跳变，而其余节点继续跳变过程。当信息传输结束后，收发两端

加入跳频序列继续跳转过程。该方法不需要专用控制信道进行控制信息发送,所有次要用户可随时进行信道协商。但该方法信道接入时延较大,过高的频谱切换会导致能量浪费等情况。



图 2.8 基于跳变序列的协调机制

(4) 在基于多汇聚的控制信息管理方法中,控制信息交互不需要所有次要用户使用相同的跳频序列^[93]。其协调机制如图 2.9 所示。在每个时隙中,各次要用户发射端、接收端按照自身的跳频序列跳转信道,当双方跳转到相同信道并建立通信链路时,完成汇聚过程。当发射端向接收端发送数据时,首先发送 RTS,若接收端接收到报文,则在当前信道上回复 CTS 给发射端开始数据传输过程。当信息传输结束后,各次要用户返回到各自原来的跳频序列继续跳转信道。

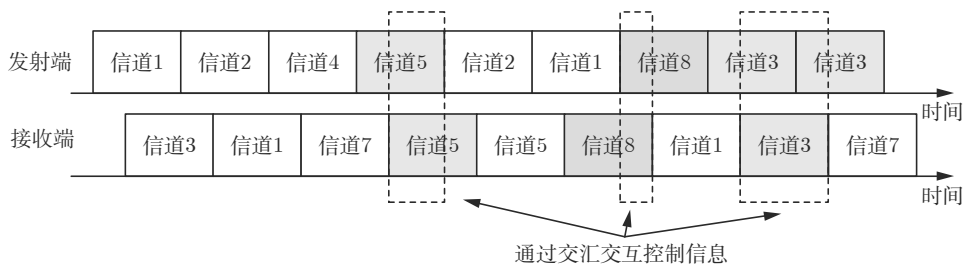


图 2.9 基于多汇聚控制信息管理的协调机制

2.3 面向频谱感知与接入的多臂赌博机在线学习理论

本节介绍多臂赌博机模型,以阐明其在频谱感知与接入策略设计中的一般方法和研究思路。

2.3.1 多臂赌博机在线学习问题模型

在经典随机多臂赌博机问题模型中,假设存在一个玩家面对 N ($N \geq 2$) 个相互独立的机臂。在每个时隙 $t \in \mathcal{I}_T = \{1, 2, \dots, T\}$ 中,玩家根据某个策略选择

一个机臂, 得到一个随机收益值 $X_i(t)$, $i \in \mathcal{I}_N = \{1, 2, \dots, N\}$ 。同一机臂不同时隙中的收益值为 IID, 并且机臂 $i \in \mathcal{I}_N$ 的收益值服从均值为 μ_i 的分布 $\mathcal{D}_i$, 其中 $\mu_i = E\{\mathcal{D}_i\} = E\{X_i(t)\}$ 。在该问题模型中, 机臂的收益值分布和均值对玩家都是未知的, 玩家的目标是设计某种策略, 使得玩家在一定数量的时隙内根据积累的收益观测值选择机臂, 以获得尽可能大的累积收益期望。

对于策略 π , 令 $U^\pi(t)$ 表示 t 个时隙中的累积收益值, 其期望可表示为

$$E\{U^\pi(t)\} = E\left\{\sum_{j=1}^t X_{\pi_j}(j)\right\} = \sum_{i=1}^N \sum_{j=1}^t E\{X_{\pi_j}(j)I[\pi_j = i]\} = \sum_{i=1}^N \mu_i E\{T_i(t)\} \quad (2.5)$$

其中, π_j 表示时隙 j 中选择的机臂; $T_i(t)$ 表示 t 个时隙中选择机臂 i 的次数; 在多臂赌博机问题模型中, 在不对期望符号进行特别说明的情况下, $E\{\cdot\} = E_\pi\{\cdot\}$ 可理解为在给定策略 π 下的条件期望, 其中 $\{\cdot\}$ 中可能存在一个或多个随机变量。该条件期望可用于描述在多个随机变量组成的联合分布中, 在给定策略 π 下玩家可获得的收益值均值^[68]。

在所有策略中, 当各机臂收益值分布 $\mathcal{D}_i$, $\forall i \in \mathcal{I}_N$ 已知时, 存在能够获得最大累积收益期望的最优策略, 即在每个时隙中选择收益均值最大的机臂 μ^* (简称为最优机臂), 其中 $\mu^* = \max_{i \in \mathcal{I}_N} \{\mu_i\}$ 。在实际问题中, 机臂收益值分布对玩家未知, 因此, 玩家不可避免地选择非最优机臂, 该过程将会让任意策略 π 带来的累积收益期望与理论最大累积收益期望之间存在一个差值, 定义为后悔, 有如下形式:

$$\begin{aligned} R(t) &= E\{U^{\pi^*}(t)\} - E\{U^\pi(t)\} = \mu^* t - E\{U^\pi(t)\} \\ &= \sum_{i: \mu_i < \mu^*} (\mu^* - \mu_i) E\{T_i(t)\} \end{aligned} \quad (2.6)$$

值得说明的是, 多臂赌博机问题中所设计策略的后悔值通常采用期望的形式。这种方式能够有效地解决两个核心问题: 收益随机性带来的不确定性和策略性能的数学分析的便利性。首先, 由于玩家获得的收益具有随机性, 不同时间点的累积收益会有显著差异。采用期望形式的后悔值作为度量, 能够在平均意义上衡量策略的性能, 这样就可以在一定程度上消除随机性引起的波动。这种方法提供了一个长期平均的性能指标, 使得评估更加稳定和可靠。其次, 期望形式的后悔值在数学分析上更加方便处理与分析。这种方法提供了一个清晰的框架, 便于分析和比较不同策略的效果, 尤其是在处理复杂的概率模型和优化问题时, 使得理论分析和实证研究都更为可行, 有助于深入理解和改进策略。

在上述模型基础上,对多臂赌博机问题的研究衍生出多种变体形式,如多个玩家下的多用户多臂赌博机问题^[94]、机臂收益值在不同时隙中呈相关的休止多臂赌博机 (rested MAB) 和无休止多臂赌博机 (restless MAB)^[36]、玩家受有限资源限制下的背包多臂赌博机 (bandits with knapsack, BwK) 问题^[95]、多个机臂对应不同玩家组成一个机臂的组合多臂赌博机 (combinatorial MAB) 问题^[96]、考虑玩家实时状态的上下文多臂赌博机 (contextual MAB)^[97]、机臂收益值的期望随时间变化的非静态多臂赌博机 (non-stationary MAB)^[98]、存在恶意干扰使得机臂收益值动态变化的对抗多臂赌博机 (adversarial MAB)^[99]、机臂收益值为连续值的连续多臂赌博机 (lipschitz MAB)^[100]。这些模型可针对不同无线网络场景进行建模,具有指导作用和重大应用潜力。

2.3.2 多臂赌博机在线学习模型与频谱感知接入映射关系

由上述模型可知,认知无线网络中信道统计信息未知下的频谱感知接入问题可基于多臂赌博机进行有效建模^[101-102]。其主要思路是基于认知无线网络问题场景和设计目标,将频谱感知接入过程与多臂赌博机问题建立映射关系,如图 2.10 所示。具体地,认知无线网络中次要用户可视为多臂赌博机模型中的玩家,信道可视为机臂,次要用户在信道上的吞吐量、接入时间或效能收益可视为拉动机臂获得的收益,次要用户进行信道感知接入的过程可视为玩家连续选取机臂获得收益的过程。当次要用户感知的信道状态为空闲时,其可获得相应收益;当感知信道为占用时,其可获得的收益为零。

为建立认知无线网络中频谱感知接入模型,首先需考虑主要用户信道占用模型。目前常用的主要用户信道占用模型为 IID 信道模型和马尔可夫信道模型。在 IID 信道模型中,每条信道在不同时隙中的主要用户占用状态呈 IID,信道实际状态服从伯努利分布;在马尔可夫信道模型中,每条信道的主要用户占用活动服从马尔可夫随机过程的统计学模型,信道状态在不同时隙中并非统计独立,而是与上一时刻的信道状态有关^①。值得说明的是, IID 信道模型能够较好地描述信道状态并为频谱感知接入算法提供决策基础,因此,本书主要针对该信道模型展开研究。

① 马尔可夫过程: 给定随机过程 $\{X(t) : t \in \mathcal{I}_T\}$, $E = \{x : X(t) = x, t \in \mathcal{I}_T\}$ 为其状态空间。如果对于 $\forall n \geq 2, t_1 < t_2 < \dots < t_n \in \mathcal{I}_T$, 有 $P\{X(t_n) < x_n | X(t_{n-1}) = x_{n-1}, \dots, X(t_1) = x_1\} = P\{X(t_n) < x_n | X(t_{n-1}) = x_{n-1}\}$, 则称 $\{X(t) : t \in \mathcal{I}_T\}$ 为马尔可夫过程^[103]。

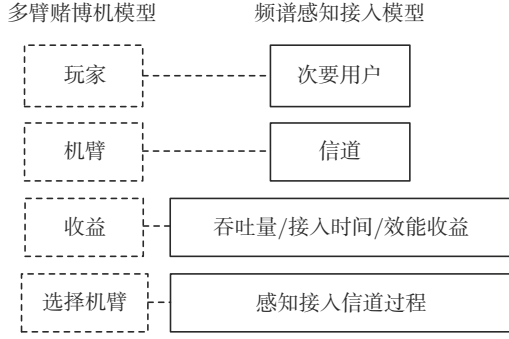


图 2.10 多臂赌博机与频谱感知接入映射关系

1. IID 信道模型

IID 信道模型下的频谱感知接入问题模型建立如下。假设网络中有 N 条信道，信道 $i \in \mathcal{I}_N$ 对应的收益服从未知分布，期望收益记为 μ_i 。在时隙 j 中，次要用户进行频谱决策选择一条信道 $\pi_j \in \mathcal{I}_N$ 进行感知，并获得实时收益 $X_{\pi_j}(j)$ 。对于任意频谱感知与接入策略，其后悔可根据式 (2.6) 等价表示为

$$R(t) = \mu^* t - E \left\{ \sum_{j=1}^t X_{\pi_j}(j) \right\} \quad (2.7)$$

其中， $\mu^* = \max_{i \in \mathcal{I}_N} \{\mu_i\}$ 表示所有信道中次要用户能够获得的最大收益的期望， $\mu^* t$ 表示次要用户已知信道收益期望信息前提下的最优传输性能。在实际网络中，次要用户可基于在线学习等方式估计各信道的收益统计信息，在信息估计准确后，持续性选择收益期望最大的信道进行感知接入，可最终逼近统计信息已知下最优方法的统计性能。IID 信道模型下的频谱感知接入问题在于设计有效的在线学习统计决策方法，使得后悔值尽可能小。

2. 多信道频谱感知接入模型

在认知无线网络中，当次要用户采取并行感知框架时，其可同时感知多条信道，该场景同样适用于多臂赌博机问题模型的建立，即玩家可在每个时隙同时选取多个机臂^[33]。假设次要用户每个时隙中可同时感知 M 条信道，则后悔可表示为

$$R(t) = t \sum_{m=1}^M \mu_m^* - E \left\{ \sum_{j=1}^t \sum_{m=1}^M X_{\pi_j^m}(j) \right\} \quad (2.8)$$

其中, $\sum_{m=1}^M \mu_m^*$ 表示 M 个最优信道的收益期望之和, π_j^m 表示时隙 j 下选择的第 m 个信道。值得说明的是, 对于集中式认知无线网络, 由于中心节点可对各次要用户的感知接入进行调度, 集中式网络的频谱感知接入模型可转化为单用户多信道频谱感知接入模型。

3. 多用户频谱感知接入模型

认知无线网络中存在多个次要用户时, 频谱感知接入问题可建模为多用户多臂赌博机问题。假设网络中存在 U 个次要用户, 对于任意统计决策方法, 多用户频谱感知接入的后悔可建模为

$$R(t) = t \sum_{u=1}^U \mu_u^* - E \left\{ \sum_{j=1}^t \sum_{u=1}^U X_{\pi_{u,j}}(j) \right\} \quad (2.9)$$

其中, μ_u^* 表示次要用户 u 的最优信道, $\pi_{u,j} \in \mathcal{I}_N$ 表示次要用户 u 时隙 j 中选择的信道。值得说明的是, 式 (2.9) 中建立的问题模型未考虑分布式用户间信息传输产生冲突等情况, 在实际多用户场景中, 应根据不同条件和问题目标, 针对性地进行问题建模和方法设计。

2.3.3 在线学习方法与性能分析

关于多臂赌博机问题的分析已得到了广泛研究, 并构建了完备的性能分析框架, 可用来分析频谱感知与接入策略的有效性。作为多臂赌博机经典问题的研究先驱, Lai 和 Robbins^[28] 通过分析证明: 在 IID 收益模型下 N 个机臂的多臂赌博机问题中, 任意自适应统计决策方法^①的收益损失 $R(t)$ 在统计渐近意义上存在下界, 即

$$\liminf_{t \rightarrow \infty} \frac{R(t)}{\ln t} \geq \sum_{i=1, i \neq i^*}^N \frac{\mu^* - \mu_i}{D(\mu_i, \mu^*)} \quad (2.10)$$

其中, $D(\mu_i, \mu^*)$ 表示期望分别为 μ_i 和 μ^* 的伯努利分布之间的库尔贝克-莱布勒 (Kullback-Leibler) 距离。式 (2.10) 表明, 任意自适应统计决策方法在时间 $t \rightarrow \infty$ 时, $R(t)$ 的下界阶数为 $O(\ln t)$, 这表明对于任意问题和任意统计决策方法, 所能达到的收益损失在时隙趋于无穷大时的阶数最小为 $O(\ln t)$ 。此外, Lai

① 根据文献 [104] 的定义, 自适应统计决策方法是指任意时刻 t 的决策只依赖之前样本值的决策方法。

和 Robbins^[28] 提出了基于概率计算的在线学习方法, 并通过理论分析证明其可达到统计渐近收益损失的下界, 即 $R(t)$ 在 $t \rightarrow \infty$ 时阶数为 $O(\ln t)$ 。

在上述研究基础上, Anantharam 等^[33,105] 考虑次要用户每个时隙中可选取多个机臂的场景, 基于概率计算设计在线学习方法, 并证明收益损失渐近统计意义上的下界为

$$\liminf_{t \rightarrow \infty} \frac{R(t)}{\ln t} \geq \sum_{i \notin M\text{-best}}^N \frac{\mu_M^* - \mu_i}{D(\mu_i, \mu_M^*)} \quad (2.11)$$

其中, $M\text{-best}$ 表示收益期望最优的 M 个机臂集合, μ_M^* 表示第 M 个最优机臂的收益期望。

考虑上述研究所设计方法中, 存在需存储所有样本信息、计算复杂度高等问题, Rajeer^[106] 和 Auer 等^[30] 相继提出了索引系数算法, 即基于测度论等工具为每个机臂构建一个索引标号, 通过比较不同机臂的索引标号对机臂的期望收益大小进行排序, 相比于之前的算法, 该方法运算复杂度低、利于工程实现。值得说明的是, Rajeer^[106] 提出方法的收益损失在统计渐近情况下阶数为 $O(\ln t)$, 但无法保证在有限时间内的性能。Auer 等^[30] 利用当前样本均值的上置信界限作为机臂实际收益信息的索引标号, 提出了两种统计决策方法, 分别记为 UCB1 和 UCB2。对于 UCB1 算法, 机臂 i 在时隙 t 中的索引标号表示为

$$g_i(t) = \bar{\mu}_i(t) + \sqrt{\frac{2 \ln t}{T_i(t)}} \quad (2.12)$$

其中, $T_i(t)$ 表示机臂 i 在 t 个时隙中被选取的次数, $\bar{\mu}_i(t)$ 表示 t 个时隙中机臂 i 的平均收益。机臂的索引标号实际上是平均收益的上置信界限, 表示的是机臂可能具有的收益上限。该算法的核心思想是在选取机臂的过程中对每个机臂的收益估计一个置信区间, 每一次选取上置信界限最大的机臂, 再根据观测到的结果更新各机臂的上置信界限。该方法在统计渐近和有限时间下收益损失期望的阶数可证明为 $O(\ln t)$ 。收益损失期望的上界满足如下关系:

$$R(t) \leq 8 \sum_{i \neq i^*}^N \frac{\ln t}{\mu^* - \mu_i} + \left(1 + \frac{\pi^2}{3}\right) \sum_{i=1}^N (\mu^* - \mu_i) \quad (2.13)$$

其中, i^* 表示期望收益最大的机臂。对于 UCB2 算法, 机臂 i 每次被选取时被拉动 $\lceil (1 + \alpha)^{T_i(t)+1} \rceil - \lceil (1 + \alpha)^{T_i(t)} \rceil$ 次, 在时隙 t 中的索引标号表示为

$$g_i(t) = \bar{\mu}_i(t) + \sqrt{\frac{(1 + \alpha) \ln\{e \cdot i / \lceil (1 + \alpha)^{T_i(t)} \rceil\}}{2 \lceil (1 + \alpha)^{T_i(t)} \rceil}} \quad (2.14)$$

该方法在统计渐近和有限时间条件下的收益损失期望阶数可证明为 $O(\ln t)$ 。收益损失期望的上界满足如下关系：

$$R(t) \leq \sum_{i \neq i^*}^N \left[\frac{(1 + \alpha)(1 + 4\alpha) \ln 2e(\mu^* - \mu_i)^2 i}{2(\mu^* - \mu_i)} + \frac{c_\alpha}{\mu^* - \mu_i} \right] \quad (2.15)$$

其中, c_α 是一个与 α 负相关的参数。

除上述方法外, 设计多臂赌博机问题的统计决策方法还有概率匹配算法^[107]、 ϵ -贪心算法^[108]、softmax 算法^[109]、自适应追踪算法^[110]等。其中, 在概率匹配算法中, 信道 i 在时隙 t 中被玩家选取的概率 $p_i(t) = \bar{\mu}_i(t) / \sum_{i=1}^K \bar{\mu}_i(t)$; 在 ϵ -贪心算法中, 最优机臂 i 在时隙 t 中被选取的概率为 $1 - \epsilon + \frac{\epsilon}{K}$, 非最优机臂被选取的概率为 ϵ/K ; 在 softmax 算法中, 信道 i 在时隙 t 中被玩家选取的概率 $p_i(t) = e^{\bar{\mu}_i(t)/T} / \sum_{i=1}^K e^{\bar{\mu}_i(t)/T}$, 其中 T 表示一预设参数; 在自适应追踪算法中, 机臂 i 在时隙 t 中被选取的概率为

$$p_i(t+1) = \begin{cases} p_i(t) + \beta[1 - p_i(t)], & i = \arg \max_{i \in \mathcal{I}_K} \{\bar{\mu}_i(t)\} \\ p_i(t) + \beta[0 - p_i(t)], & i \neq \arg \max_{i \in \mathcal{I}_K} \{\bar{\mu}_i(t)\} \end{cases} \quad (2.16)$$

其中, β 为一预设参数。在上述方法中, 用户皆基于收益的估计以一定概率选择机臂, 因此, 该类方法的理论性能通常难以得到保证。相较于上述方法, 尽管上置信界限算法具有良好的理论性能, 但在不同多臂赌博机问题场景中, 上置信界限算法的性能并不一定为最优。因此, 在实际问题中, 应结合具体的场景特点和需求设计统计决策方法。

2.4 小结

本章对频谱感知接入中的关键技术进行了介绍, 阐明了本书在认知无线网络众多信道占用模型、感知框架中聚焦时隙结构模型、并行感知框架、IID 信道状态模型下的频谱感知与接入策略设计, 为后续研究奠定技术基础。同时, 对面向频谱感知与接入策略设计的多臂赌博机相关理论进行了介绍, 描述了该工具的问题模型、与频谱感知接入问题的对应关系、基于多臂赌博机建立的频谱感知接入问题模型、方法设计的一般思路和相关算法, 为后续方法设计提供了理论基础。

第3章

基于双层多臂赌博机的多信道感知与部分信道接入方法

在大多数频谱感知与接入方法中，通常假设次要用户感知的信道数量和能够接入的信道数量相同。这类研究着重探讨次要用户如何选择信道进行感知，以及如何与其他用户共享频谱资源。然而，在部分实际通信系统中，受硬件和功率等限制，次要用户在每个时隙中虽能感知多条信道，但仅能选择其中部分信道进行接入，这使得频谱感知与接入过程需划分为两个步骤：次要用户需首先决定选择哪些信道进行感知，其次根据感知结果进行第二次决策，即决定选择哪些信道进行接入，以完成与其他用户的频谱资源共享及信息传输。这种两步决策的过程能够充分考虑次要用户的硬件限制，在动态变化的无线环境中更有效地分配和管理频谱资源，从而更深入地提高频谱利用效率。然而，在信道统计信息未知的情况下，相较于大多数频谱感知与接入方法，多信道感知与部分接入在方法设计上遇到了新的问题。这主要是因为次要用户的信道感知和接入两个阶段之间存在耦合影响：感知结果会影响接入决策，而接入决策又会影响次要用户对信道空闲概率的估计，进而影响下一时隙中的信道感知决策。这种相互依赖关系使得感知与接入并非简单的相互接续的两个独立过程，而是前后接续又相互约束的两个过程，传统的资源分配和优化理论难以对这两个阶段进行精确地描述以及设计相应的方法。因此，需要构建有效的模型并设计新的方法来揭示这两个阶段的相互作用。目前，尽管已有少数研究尝试针对单用户场景下的频谱感知接入进行探讨^[56]，但在多用户网络环境下，这方面的深入研究仍然不足，如对于集中式网络，需研究中心节点如何调度各次要用户进行信道感知、如何根据累积的观测样本更新信道信息，以及如何根据感知结果调度各次要用户进行信道接入；对于分布式网络，多个次要用

户接入相同的信道可能会导致冲突,降低网络性能。因此,需要进一步研究如何设计冲突避免机制,以有效管理多个用户对频谱资源的利用。值得说明的是,在认知无线网络中,次要用户发现空闲信道并传输成功时,表明次要用户成功利用了频谱机会,因此越大的吞吐量越表明,次要用户更高效地利用了频谱资源。

本章首先考虑次要用户硬件和功率等限制,基于双层多臂赌博机在线学习理论构建多信道感知与接入问题模型;其次,设计算法使得次要用户通过信道感知获取信道的实时状态,并在线更新信道统计信息的估计值;再根据感知到的信道状态结果优选信道进行接入;最后,通过定义的数据量损失和平均吞吐量验证所提算法的有效性。

本章结构如下:3.1节对集中式网络的频谱感知与接入方法展开研究。首先将中心节点分配信道给各次要用户进行感知、接入的问题建模为双层多臂赌博机模型;其次分别设计了信道空闲概率信息已知和未知两种情况下的频谱感知与接入方法,并通过理论分析与仿真,验证了所设计的方法在不同信道数量、不同用户数量、不同检测概率下的有效性。3.2节对分布式网络的频谱感知与接入方法展开研究。首先将分布式用户感知接入决策建模为多用户双层多臂赌博机模型,研究了信道统计信息已知情况下的最优感知与接入方法;其次,以此为性能基准,研究了基于空闲概率估计的上置信界限和随机接入的频谱感知与接入方法;最后,通过理论分析与仿真,验证了所设计的方法在不同信道数量、不同用户数量、不同检测概率下的有效性。3.3节对研究进行总结。

3.1 集中式网络的频谱感知与接入方法

3.1.1 系统模型

本节基于文献 [111] ~ 文献 [113] 构建集中式网络问题模型。集中式网络中有 N 条信道,记 $\mathcal{I}_N = \{1, 2, \dots, N\}$ 为信道集合。网络中存在 U 个用户,记 $\mathcal{I}_U = \{1, 2, \dots, U\}$ 为用户集合。网络采用等长时隙结构,每个时隙长度为 T 。考虑上行传输场景,中心节点在每个时隙中首先收集各个次要用户的发送数据请求,进行信道感知决策,然后将信道分配方案广播给各次要用户,各次要用户根据分配方案对信道进行感知,再将感知结果上报给中心节点。基于感知结果,中心节点再进行信道接入决策,并将接入结果广播给各次要用户进行接入。根据接入决策,各次要用户每个时隙在每条信道上传输 B 比特数据,传输成功后中心节点发送肯定应答 (acknowledgement, ACK) 信息给各次要用户。

本节考虑的是 IID 信道模型。假设信道状态在不同时隙中相互独立且服从块衰落模型,即信道状态在一个时隙内保持不变。令 θ_i 表示信道 i 的空闲概率,则每个时隙中信道 i 以概率 θ_i 为空闲状态、概率 $1 - \theta_i$ 为占用状态,其状态与时隙无关。另外,假设信道的空闲概率有 $\theta_1 > \theta_2 > \dots > \theta_N$ 。

考虑次要用户受硬件和功率限制无法进行全频谱感知,每个时隙中仅能同时感知 M 条信道且最多接入 K 条感知为空闲的信道。具体地,当次要用户在进行频谱感知时,发射端采用能量检测法检测信道实时状态。令检测概率为 P_d ,虚警概率为 P_f ,并且各次要用户检测能力相同。若信道感知为空闲,则次要用户采用固定传输速率在该时隙剩余时间内发送数据;若信道感知为占用,则次要用户放弃该时隙的传输机会。

3.1.2 模型建立与问题描述

本节研究的问题中考虑了信道空闲概率对认知无线网络未知、各次要用户对信道状态的感知可能存在错误(此时次要用户感知到的信道状态与实际信道状态未必完全一致,即感知为空闲的信道的实际状态可能并非为实际空闲,次要用户接入该信道时无法成功传输),以及次要用户每个时隙中仅能感知 M 条信道并接入最多 K 条信道等情况。为获得尽可能大的数据量,需要设计频谱感知与接入方法,使得中心节点能够基于存在错误的样本观测值对信道空闲概率进行估计,同时分别进行感知决策和接入决策,调度各次要用户对信道的感知接入。

为解决上述问题,本节将集中式网络的频谱感知接入过程建模为双层多臂赌博机模型,如图 3.1 所示。具体地,多臂赌博机中的玩家对应无线网络中的中心节点及多个次要用户,多臂赌博机的机臂对应感知、接入的信道,多臂赌博机的收益对应次要用户接入信道获得的数据量,多臂赌博机中选择机臂的过程对应中心节点为各次要用户分配信道进行感知与接入两步决策过程。

对于集中式网络的任意频谱的感知与接入方法 π^C , t 个时隙中所能获得数据量的期望可表示为

$$E\{U^{\pi^C}(t)\} = B \sum_{j=1}^t \sum_{u=1}^U \sum_{i=1}^N E\{E\{S_i(j)|X_{u,i}(j)=1\} I[i \in a_u(j)]\} \quad (3.1)$$

其中, $X_{u,i}(j)$ 表示次要用户 u 在时隙 j 下对信道 i 的感知结果; $X_{u,i}(j) = 1$ 、 $X_{u,i}(j) = 0$ 分别表示信道感知为空闲和占用; $S_i(j)$ 表示时隙 j 下信道 i 实际占用状态,当 $S_i(j) = 1$ 、 $S_i(j) = 0$ 时,其分别表示信道实际状态为空闲和占用; $a_u(j)$ 表示次要用户 u 时隙 j 下接入的信道集合; $I[\cdot]$ 表示指示函数, $I[i \in a_u(j)] = 1$

表示时隙 j 中次要用户 u 感知信道 i , $I[i \in a_u(j)] = 0$ 表示时隙 j 中次要用户 u 未感知信道 i ; $E\{S_i(j)|X_{u,i}(j) = 1\}$ 表示次要用户 u 在时隙 j 中感知信道 i 为空闲并接入时能够成功传输的条件期望; $E\{\cdot\}$ 为外层期望, 表示次要用户在算法 π^C 下接入信道集合 $a_u(j)$, $\forall u \in \mathcal{I}_U$ 的期望。

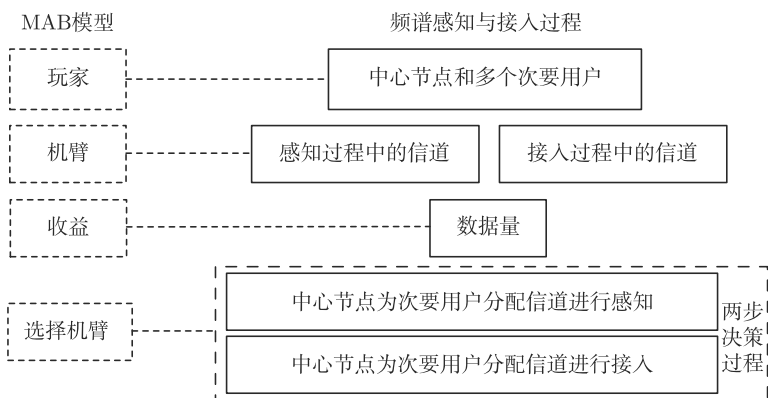


图 3.1 集中式网络基于双层多臂赌博机的模型建立

在所有频谱感知与接入方法中, 存在一个能够获得最大数据量期望的最优方法 π^{*C} (即可理解为中心节点具有信道空闲概率先验信息时, 能够在每个时隙中实时感知接入空闲概率最大的信道, 因此, 可获得最大数据量期望), 该方法获得的最大数据量期望可作为未知信道空闲概率下所设计的感知与接入方法的性能基准。本节研究的问题中的信道空闲概率未知, 次要用户无法实时感知接入数据量期望最大的信道集合, 因此任意方法所获得的数据量期望与最大数据量期望存在一个差值, 可将其定义为数据量损失期望, 表示为

$$R(t; \pi^C) = E\{U^{\pi^{*C}}(t)\} - E\{U^{\pi^C}(t)\} \quad (3.2)$$

由式 (3.2) 可知, 为获得尽可能大的数据量期望, 应使设计的频谱感知与接入方法获得的数据量期望 $E\{U^{\pi^C}(t)\}$ 尽可能趋于最大数据量期望 $E\{U^{\pi^{*C}}(t)\}$, 使得 $R(t; \pi^C)$ 尽可能小, 因此, 该问题可转化为以下两个问题。

(1) 问题 (P1): 求解 $E\{U^{\pi^{*C}}(t)\}$, 即首先考虑信道空闲概率信息对认知无线网络已知条件下获得最大数据量期望的方法设计, 求得的数据量期望可作为未知信息下方法设计的性能基准。

(2) 问题 (P2): 求解 $E\{U^{\pi^C}(t)\}$, 即考虑信道空闲概率信息未知下的方法设计, 使得 $R(t; \pi^C)$ 尽可能小。

接下来, 将首先对问题 (P1) 进行求解, 即首先研究信道空闲概率已知下的最优频谱感知与接入方法。以该方法为性能基准, 再对问题 (P2) 进行求解, 即研究信道空闲概率信息未知下的方法设计。

3.1.3 基于空闲概率估计上置信界限的方法设计

1. 统计信息已知时的性能基准

基于上述模型, 次要用户感知信道为空闲时存在两种情况, 即信道实际状态为空闲而被感知为空闲, 该事件概率为 $\theta_i(1 - P_f)$, 其中 P_f 为虚警概率; 信道实际状态为占用而被感知为空闲, 该事件概率为 $(1 - \theta_i)(1 - P_d)$, 其中 P_d 为检测概率。因此, 信道 i 被感知为空闲的概率可表示为

$$f(\theta_i) = \theta_i(1 - P_f) + (1 - \theta_i)(1 - P_d) = \theta_i(P_d - P_f) + 1 - P_d \quad (3.3)$$

次要用户 u 感知信道 i 为空闲并接入该信道 (每个时隙在每条信道上传输 B 比特数据) 可获得的数据量期望可表示为

$$\mu_i = BE\{S_i(t)|X_{u,i}(t) = 1\} = \frac{B\theta_i(1 - P_f)}{f(\theta_i)} \quad (3.4)$$

为获得最大数据量, 中心节点需考虑如何基于信道空闲概率先验信息为 U 个次要用户分配信道进行感知, 再基于感知到的结果分配信道给各用户进行接入。由于在该问题背景下信道空闲概率对各次要用户相同, 各次要用户的信道感知结果汇报给中心节点, 并且各次要用户具有相同的检测概率、虚警概率, 该场景下集中式网络感知接入问题可视为单用户多信道感知接入问题, 即每个时隙中中心节点分配 UM 条 (UM 表示 $U \times M$) 信道给 U 个次要用户进行感知的过程, 可视为单个次要用户每个时隙中选择 UM 条信道进行感知; 中心节点根据感知结果分配最多 UK 条 (UK 表示 $U \times K$) 信道给各次要用户接入的过程, 可视为单个次要用户每个时隙中基于感知结果选择最多 UK 条信道进行接入的过程。

对于单个用户多信道感知问题, 由文献 [112] 可知, 信道空闲概率信息已知时的最优感知与接入方法为用户每个时隙中感知 UM 条具有最大数据量期望的信道, 并接入最多 UK 条数据量期望最大、感知为空闲的信道。因此, 集中式网络的最优感知与接入方法 (记为算法 π^{*C} -IID) 为中心节点每个时隙中选择数据量期望 (即 $\mu_i, \forall i \in \mathcal{I}_N$) 最大的 UM 条信道分配给各次要用户进行感知, 获得感知为空闲的信道集合 $\mathcal{I}(t)$; 再选择最多 UK 条感知为空闲且数据量期望最大的信道分配给次要用户进行传输。

2. 统计信息未知下的方法设计

本节对信道统计信息未知下的频谱感知与接入方法进行设计，其核心思想为中心节点在每个时隙中，首先通过信道感知获取信道状态样本，实时更新估计各信道的概率，并为每条信道计算概率估计值的上置信界限。此后，中心节点根据感知到的信道状态及信道的上置信界限，进行接入决策，即分配信道给各次要用户进行信息传输。本节设计的算法命名为 π^C -IID，详见算法 3.1。

算法 3.1 集中式的网络频谱感知与接入算法 π^C -IID

```

1: 初始化: 对时隙  $l = 1 : \left\lceil \frac{N}{UM} \right\rceil$ , 次要用户  $u$  感知  $M$  条信道, 即信道  $\{(u-1)M+1+(l-1)UM : uM+(l-1)uM \bmod N\}$ 。感知结束后, 次要用户  $u$  将感知结果汇报给中心节点;
2: 中心节点更新  $\mathbf{T}(l)$  和  $\mathbf{Y}(l)$  统计量, 并随机选择  $\max\{|\mathcal{I}(l)|, UK\}$  条空闲信道分配给次要用户进行接入;
3: for 时隙  $t$  do
4:   中心节点根据式 (3.7) 计算所有信道索引标号  $g_i(t)$ ,  $\forall i \in \mathcal{I}_N$ , 进行频谱感知决策, 选择  $UM$  个索引标号最大的信道集合  $\mathcal{M}(t)$  并将信息发送给所有次要用户;
5:   次要用户  $u$  感知  $\mathcal{M}(t)$  中第  $(u-1)M+1+(l-1)UM : (u-1)M+1$  条信道, 并将感知结果上报给中心节点;
6:   中心节点获得信道状态感知为空闲的集合  $\mathcal{I}(t)$ ;
7:   if  $|\mathcal{I}(t)| \geq UK$  then
8:     中心节点在集合  $\mathcal{I}(t)$  里选择前  $UK$  个信道索引标号最大的信道, 并广播给次要用户; 次要用户  $u$  对  $\mathcal{I}(t)$  中的第  $(u-1)K+1+(l-1)UK : (u-1)K+1$  条信道以固定传输速率进行接入;
9:   else if  $0 < |\mathcal{I}(t)| < UK$  then
10:     中心节点随机调度次要用户接入所有空闲信道;
11:   else
12:     不接入信道;
13:   end if
14:   中心节点更新  $T_i(t)$ ,  $\forall i \in \mathcal{M}(t)$  和  $Y_i(t)$ ,  $\forall i \in \mathcal{I}(t)$ ;
15: end for

```

在算法初始化阶段, 如时隙 $l = 1 : \left\lceil \frac{N}{UM} \right\rceil$, 中心节点选择信道集合 $\mathcal{M}(l)$, 将其中第 $(u-1)M+1+(l-1)UM : uM+(l-1)uM \bmod N$ 条信道分配给次要用户 u 进行感知, 其中 $\bmod$ 表示取模运算。感知结束后, 各次要用户将感知结果汇报给中心节点, 中心节点将 $\max\{|\mathcal{I}(l)|, UK\}$ 条信道分配给 U 个次要用户进行接入, 其中 $\mathcal{I}(l)$ 表示时隙 l 中感知为空闲的信道集合, $|\mathcal{I}(l)|$ 表示空闲信

道集合中元素数量。中心节点更新信道统计量 $\mathbf{T}(l) = [T_1(l), T_2(l), \dots, T_N(l)]^T$ 和 $\mathbf{Y}(l) = [Y_1(l), Y_2(l), \dots, Y_N(l)]^T$, 其中 $T(l)$ 表示时隙 l 中对 N 条信道感知的次数, $Y(l)$ 表示时隙 l 中各信道感知为空闲的次数。

在后续感知接入过程中, 如时隙 t , 中心节点对每条信道空闲概率进行估计, 即对于信道 i , 有

$$E \left\{ \frac{Y_i(t-1)}{T_i(t-1)} \right\} = f(\theta_i) \quad (3.5)$$

将式 (3.5) 代入式 (3.3), 中心节点对信道 i 空闲概率 θ_i 的估计可表示为

$$\hat{\theta}_i(t) = \frac{\frac{Y_i(t-1)}{T_i(t-1)} + P_d - 1}{P_d - P_f} \quad (3.6)$$

中心节点根据 $\hat{\theta}_i(t)$ 的上置信界限为各信道计算索引标号为

$$g_i(t) = \frac{\frac{Y_i(t-1)}{T_i(t-1)} + P_d - 1}{P_d - P_f} + \frac{1}{P_d - P_f} \sqrt{\frac{2 \ln(t-1)}{T_i(t-1)}} \quad (3.7)$$

基于上述索引标号, 中心节点进行信道感知决策, 即在时隙 t 中选择 UM 个索引标号最大的信道集合 $\mathcal{M}(t)$ 并将信息广播给所有次要用户。次要用户 u 对集合中第 $(u-1)M+1+(l-1)UM:(u-1)M+1$ 条信道进行感知, 并将结果上报至中心节点。

基于各次要用户上报的感知结果, 中心节点可得到时隙 t 中的空闲信道集合 $\mathcal{I}(t)$ 。若 $|\mathcal{I}(t)| \geq UK$, 则中心节点在所有空闲信道中选择索引标号最大的 UK 条信道分配给各次要用户, 次要用户 u 选择第 $(u-1)K+1+(l-1)UK:(u-1)K+1$ 条信道接入; 若 $|\mathcal{I}(t)| < UK$, 则中心节点调度所有空闲信道随机分配给次要用户进行接入; 若无空闲信道, 则中心节点不进行信道分配。传输完成后, 中心节点更新信道统计量, 即

$$T_i(t) = T_i(t-1) + 1, \forall i \in \mathcal{M}(t) \quad (3.8)$$

$$Y_i(t) = Y_i(t-1) + 1, \forall i \in \mathcal{I}(t) \quad (3.9)$$

3.1.4 性能分析和复杂度分析

1. 性能分析

本节对所提的算法 π^C -IID 进行性能分析。为方便后续表述, 称信道集合 $\mathcal{M}^* = \{1, 2, \dots, UM\}$ 为最优信道 [在系统模型中假设信道空闲概率有 $\theta_1 >$

$\theta_2 > \dots > \theta_N$, 根据式 (3.4) 不难得出, 各信道数据量期望有 $\frac{B\theta_1(1-P_f)}{f(\theta_1)} > \frac{B\theta_2(1-P_f)}{f(\theta_2)} > \dots > \frac{B\theta_N(1-P_f)}{f(\theta_N)}$, 因此, 此处 UM 条最优信道是指 UM 条数据量期望最大的信道, 其他信道为非最优信道。按照算法 3.1 可知, 数据量损失来源于次要用户未接入最优信道。该情况可拆分为两个事件: 中心节点未分配最优信道给次要用户感知, 因此, 次要用户难以接入最优信道; 中心节点分配最优信道给次要用户感知, 但未分配数据量期望最大的信道给次要用户接入。上述事件对应的数据量损失期望上界可表示为

$$R(t; \pi^C) \leq \Delta \quad (3.10)$$

$$UK \underbrace{\sum_{j=1}^t E\{I[\mathcal{M}(j) \neq \mathcal{M}^*]\}}_{(A)} + \underbrace{\sum_{j=1}^t E \left\{ I[\mathcal{M}(j) = \mathcal{M}^*] \cdot \sum_{i < k, i, k \in \mathcal{I}_{\mathcal{M}^*}(j)} I[g_i(j) \leq g_k(j)] \right\}}_{(B)}$$

其中

$$\Delta = \max_{i, k \in \mathcal{I}_N} \left\{ \frac{(1-P_f)B\theta_i}{f(\theta_i)} - \frac{(1-P_f)B\theta_k}{f(\theta_k)} \right\} \quad (3.11)$$

表示次要用户感知接入任意两条信道对应的数据量期望的最大差值; (A) 表示各次要用户在 t 个时隙中, 未感知最优信道集合的时隙数量的期望, 系数 UK 来源于集中式网络每个时隙中信道接入的数量; (B) 表示各次要用户感知最优信道时, 在感知为空闲信道集合中, 未接入数据量期望最大的信道这一事件的期望, 其中 $I[\cdot]$ 为指示函数, $I[\mathcal{M}(j) = \mathcal{M}^*]$ 用于刻画中心节点分配集合 $\mathcal{M}^*$ 给各次要用户进行感知这一事件, 求和符号的下标 $\mathcal{I}_{\mathcal{M}^*}(j)$ 表示集合 $\mathcal{M}^*$ 在时隙 j 中被感知为空闲的信道集合, $i < k$ 的含义为 $\theta_i > \theta_k$, 由式 (3.4) 不难得出 $\mu_i > \mu_k$, 表示信道 i 的数据量期望大于信道 k 的数据量期望, 此时中心节点应优先选择信道 i 给次要用户进行接入, $I[g_i(j) \leq g_k(j)]$ 表示时隙 j 中信道 k 的索引标号大于 i 的索引标号, 此时中心节点按照算法 π^C 选择信道 k 给次要用户进行接入, $\sum_{i < k, i, k \in \mathcal{I}_{\mathcal{M}^*}(j)} I[g_i(j) \leq g_k(j)]$ 表示时隙 j 中信道集合 $\mathcal{I}_{\mathcal{M}^*}(j)$ 中任意两条信

道 $i < k$ 但 $g_i(j) \leq g_k(j)$ 发生的次数 (此时中心节点应选择信道 i 进行接入但错误地接入了信道 k , 因此获得了较小的数据量, 造成了数据量损失); $E\{\cdot\}$ 表示对 $\mathcal{I}_{\mathcal{M}^*}(j)$ 和错误接入信道的时隙数量求期望。下面分别对事件 (A) 和事件 (B) 进行分析。

对于事件 (A), 即中心节点未分配最优信道集合给各次要用户进行感知, 则至少存在一条信道 $i \notin \mathcal{M}^*$ 在某时隙中被某次要用户感知, 因此有

$$\sum_{j=1}^t E\{I[\mathcal{M}(j) \neq \mathcal{M}^*]\} \leq \sum_{i=UM+1}^N E\{T_i(t)\} \quad (3.12)$$

其中, $E\{T_i(t)\}$, $\forall i \notin \mathcal{M}^*$ 有如下结论。

引理 3.1 在 t 个时隙中, $E\{T_i(t)\}$, $\forall i \notin \mathcal{M}^*$ 有如下结论:

$$E\{T_i(t)\} \leq \frac{8 \ln t}{(\theta_{UM} - \theta_i)^2 (P_d - P_f)^2} + 1 + \frac{MU\pi^2}{3} \quad (3.13)$$

证明 对于任意正整数 l , t 个时隙内信道 i 被感知的次数可表示为

$$T_i(t) = 1 + \sum_{j=\lceil \frac{N}{UM} \rceil + 1}^t I[i \in \mathcal{M}(j)] \quad (3.14)$$

若将 $T_i(j-1)$ 拆分为 $T_i(j-1) \geq l$ 和 $T_i(j-1) < l$, 则有

$$\begin{aligned} T_i(t) = 1 + & \sum_{j=\lceil \frac{N}{UM} \rceil + 1}^t I[i \in \mathcal{M}(j), T_i(j-1) \geq l] + \\ & \sum_{j=\lceil \frac{N}{UM} \rceil + 1}^t I[i \in \mathcal{M}(j), T_i(j-1) < l] \end{aligned} \quad (3.15)$$

假设时隙 j 前信道 i 已被感知 l 次, 则有

$$T_i(t) \leq l + \sum_{j=\lceil \frac{N}{UM} \rceil + 1}^t I[i \in \mathcal{M}(j), T_i(j-1) \geq l] \quad (3.16)$$

当非最优信道 i 在时隙 j 中被感知时, 存在至少一条信道 $k \in \mathcal{M}^*$, 使得时隙 j 中信道 k 的索引标号小于信道 i 的索引标号, 即 $g_k(j) \leq g_i(j)$ (其含义为中心节

点应分配信道 k 给次要用户感知接入, 但信道标号排序 $g_k(j) \leq g_i(j)$, 所以中心节点分配信道 i 给次要用户), 因此有

$$T_i(t) \leq l + \sum_{j=\lceil \frac{N}{UM} \rceil}^{t-1} I \left[\min_{k \in \mathcal{M}^*} \left\{ \hat{\theta}_k^T[T_k(j)] + \frac{1}{P_d - P_f} \sqrt{\frac{2 \ln(j)}{T_k(j)}} \right\} \leq \right. \quad (3.17)$$

$$\left. \hat{\theta}_i^T[T_i(j)] + \frac{1}{P_d - P_f} \sqrt{\frac{2 \ln(j)}{T_i(j)}}, T_i(j) \geq l \right]$$

即得

$$T_i(t) \leq l + \sum_{j=\lceil \frac{N}{UM} \rceil}^{t-1} \sum_{k=1}^{UM} I \left[\hat{\theta}_k^T[T_k(j)] + \frac{1}{P_d - P_f} \sqrt{\frac{2 \ln j}{T_k(j)}} \leq \right. \quad (3.18)$$

$$\left. \hat{\theta}_i^T[T_i(j)] + \frac{1}{P_d - P_f} \sqrt{\frac{2 \ln j}{T_i(j)}}, T_i(j) \geq l \right]$$

由于无法得知哪一具体时隙中信道 k 的索引标号小于信道 i 的索引标号, 考虑式 (3.18) 中各信道在所有时隙中的索引标号大小, 有如下关系:

$$T_i(t) \leq l + \sum_{j=\lceil \frac{N}{UM} \rceil}^{t-1} \sum_{k=1}^{UM} I \left[\min_{0 < t_1 \leq j} \left\{ \hat{\theta}_k^T(t_1) + \frac{1}{P_d - P_f} \sqrt{\frac{2 \ln j}{t_1}} \right\} \leq \right. \quad (3.19)$$

$$\left. \max_{l \leq t_2 \leq j} \left\{ \hat{\theta}_i^T(t_2) + \frac{1}{P_d - P_f} \sqrt{\frac{2 \ln j}{t_2}} \right\} \right]$$

放缩可得

$$T_i(t) \leq l +$$

$$\sum_{k=1}^{UM} \sum_{j=1}^t \sum_{t_1=1}^j \sum_{t_2=l}^j I \left[\underbrace{\hat{\theta}_k^T(t_1) + \frac{1}{P_d - P_f} \sqrt{\frac{2 \ln j}{t_1}} \leq \hat{\theta}_i^T(t_2) + \frac{1}{P_d - P_f} \sqrt{\frac{2 \ln j}{t_2}}}_{(C)} \right] \quad (3.20)$$

式 (3.19) 用于描述索引标号 $g_k(t_1)$ 在前 j 个时隙内小于索引标号 $g_i(t_2)$ 在时隙 $l \sim j$ 中的时隙数量, 最终放缩可得式 (3.20)。在式 (3.20) 中, 事件 (C) 表明以下三个事件至少一件发生^[30] (为得到该结论, 可假设下述三个事件皆不成立, 经传递关系推理可发现, 它与事件 (C) 中不等式矛盾, 因此至少存在一个事件成立): ① $\hat{\theta}_k^T(t_1) \leq \theta_k - \frac{1}{P_d - P_f} \sqrt{\frac{2 \ln j}{t_1}}$; ② $\hat{\theta}_i^T(t_2) \geq \theta_i + \frac{1}{P_d - P_f} \sqrt{\frac{2 \ln j}{t_2}}$;

③ $\theta_k < \theta_i + \frac{2}{P_d - P_f} \sqrt{\frac{2 \ln j}{t_2}}$ 。

因此, 对事件 (C) 取期望, 可得如下不等式:

$$\begin{aligned} & E \left\{ I \left[\hat{\theta}_k^T(t_1) + \frac{1}{P_d - P_f} \sqrt{\frac{2 \ln j}{t_1}} \leq \hat{\theta}_i^T(t_2) + \frac{1}{P_d - P_f} \sqrt{\frac{2 \ln j}{t_2}} \right] \right\} \leq \\ & E \left\{ I \left[\hat{\theta}_k^T(t_1) \leq \theta_k - \frac{1}{P_d - P_f} \sqrt{\frac{2 \ln j}{t_1}} \right] \right\} + \\ & E \left\{ I \left[\hat{\theta}_i^T(t_2) \geq \theta_i + \frac{1}{P_d - P_f} \sqrt{\frac{2 \ln j}{t_2}} \right] \right\} + \\ & E \left\{ I \left[\theta_k < \theta_i + \frac{2}{P_d - P_f} \sqrt{\frac{2 \ln j}{t_2}} \right] \right\} \end{aligned} \quad (3.21)$$

根据切诺夫-霍夫丁 (Chernoff-Hoeffding, CH) 不等式^[30], 式 (3.21) “ $\leq$ ” 的右侧前两项的上界为 $E \left\{ I \left[\hat{\theta}_k^T(t_1) \leq \theta_k - \frac{1}{P_d - P_f} \sqrt{\frac{2 \ln j}{t_1}} \right] \right\} \leq j^{-4}$, $E \left\{ I \left[\hat{\theta}_i^T(t_2) \geq \theta_i + \frac{1}{P_d - P_f} \sqrt{\frac{2 \ln j}{t_2}} \right] \right\} \leq j^{-4}$ 。此外, 当 $t_2 \geq \frac{8 \ln t}{(\theta_{UM} - \theta_i)^2 (P_d - P_f)^2}$ 时, 式 (3.21) “ $\leq$ ” 的右侧第三项 $E \left\{ I \left[\theta_k < \theta_i + \frac{2}{P_d - P_f} \sqrt{\frac{2 \ln j}{t_2}} \right] \right\} = 0$ 。

因此, 令 $l = \left\lceil \frac{8 \ln t}{(\theta_{UM} - \theta_i)^2 (P_d - P_f)^2} \right\rceil$, 此时根据式 (3.20), 可得

$$E\{T_i(t)\} \leq \left\lceil \frac{8 \ln t}{(\theta_{UM} - \theta_i)^2 (P_d - P_f)^2} \right\rceil + \sum_{k=1}^{UM} \sum_{j=1}^{\infty} \sum_{t_1=1}^j \sum_{t_2=\left\lceil \frac{8 \ln t}{(\theta_{UM} - \theta_i)^2 (P_d - P_f)^2} \right\rceil}^j 2j^{-4} \quad (3.22)$$

对上式进行放缩, 可表示为

$$E\{T_i(t)\} \leq \frac{8 \ln t}{(\theta_{UM} - \theta_i)^2 (P_d - P_f)^2} + 1 + \sum_{k=1}^{UM} \sum_{j=1}^{\infty} 2j^{-2} \quad (3.23)$$

对上式进行放缩, 可表示为

$$E\{T_i(t)\} \leq \frac{8 \ln t}{(\theta_{UM} - \theta_i)^2 (P_d - P_f)^2} + \frac{UM\pi^2}{3} + 1 \quad (3.24)$$

证毕。

当事件 $I[\mathcal{M}(j) \neq \mathcal{M}^*]$ 发生时, 一定存在信道 $i \notin \mathcal{M}^*$ 被感知接入, 因此, 事件 $I[\mathcal{M}(j) \neq \mathcal{M}^*]$ 发生的时隙数量上界为所有非最优信道被感知接入的时隙数量。根据式 (3.12) 和引理 3.1, 可得

$$\begin{aligned} \sum_{j=1}^t E\{I[\mathcal{M}(j) \neq \mathcal{M}^*]\} &\leq \\ \sum_{i=UM+1}^N \frac{8 \ln t}{(\theta_{UM} - \theta_i)^2 (P_d - P_f)^2} + (N - UM) \left(\frac{UM\pi^2}{3} + 1 \right) \end{aligned} \quad (3.25)$$

接下来对事件 (B) 进行分析。令 $T'(t)$ 表示中心节点分配最优信道进行感知, 在感知为空闲的信道集合中, 未分配数据量期望最大的信道进行接入的时隙数量, 有如下结论。

引理 3.2 $T'(t)$ 的期望上界有如下关系:

$$E\{T'(t)\} \leq \sum_{i=1}^{UM-1} \sum_{k=i+1}^{UM} \left[\frac{8 \ln t}{(\theta_i - \theta_k)^2 (P_d - P_f)^2} + 1 + \frac{\pi^2}{3} \right] \quad (3.26)$$

证明 中心节点在时隙 j 中分配最优信道进行感知、在感知为空闲集合中未分配数据量期望最大的信道进行接入, 其原因为中心节点在该时隙中对最优信道索引标号排序存在错误, 因此 t 个时隙内该事件的上界 [类似于式 (3.10) 中式 (B) 的表达形式] 可表示为

$$\begin{aligned} T'(t) &\leq \sum_{j=1}^t I[\mathcal{M}(j) = \mathcal{M}^*] \cdot \sum_{i < k, i, k \in \mathcal{I}_{\mathcal{M}^*}(j)} I \left[\hat{\theta}_i^T(T_i(j-1)) + \frac{1}{P_d - P_f} \sqrt{\frac{2 \ln(j-1)}{T_i(j-1)}} \leq \right. \\ &\quad \left. \hat{\theta}_k^T(T_k(j-1)) + \frac{1}{P_d - P_f} \sqrt{\frac{2 \ln(j-1)}{T_k(j-1)}} \right] \end{aligned} \quad (3.27)$$

类似于式 (3.16) 的分析, 对于任意正整数 $l_{i,k}$, 有

$$T'(t) \leq \sum_{i < k, i, k \in \mathcal{I}_{\mathcal{M}^*}(j)} \left\{ l_{i,k} + \sum_{j=\lceil \frac{N}{UM} \rceil + 1}^t I[\mathcal{M}(j) = \mathcal{M}^*, T_k(j-1) \geq l_{i,k}] \cdot \right. \\ \left. I \left[\hat{\theta}_i^T(T_i(j-1)) + \frac{1}{P_d - P_f} \sqrt{\frac{2 \ln(j-1)}{T_i(j-1)}} \leq \right. \right. \\ \left. \left. \hat{\theta}_k^T(T_k(j-1)) + \frac{1}{P_d - P_f} \sqrt{\frac{2 \ln(j-1)}{T_k(j-1)}} \right] \right\} \quad (3.28)$$

类似于式 (3.18) ~ 式 (3.20) 的分析, 可得

$$T'(t) \leq \sum_{i < k, i, k \in \mathcal{I}_{\mathcal{M}^*}(j)} \left\{ l_{i,k} + \right. \\ \left. \sum_{j=1}^t \sum_{t_1=1}^j \sum_{t_2=l_{i,k}}^j I \left[\hat{\theta}_i^T(t_1) + \frac{1}{P_d - P_f} \sqrt{\frac{2 \ln j}{t_1}} \leq \hat{\theta}_k^T(t_2) + \frac{1}{P_d - P_f} \sqrt{\frac{2 \ln j}{t_2}} \right] \right\} \quad (3.29)$$

根据式 (3.20) ~ 式 (3.24) 的推导, 关于式 (3.29) 可得如下结论, 即

$$E\{T'(t)\} \leq \sum_{i=1}^{UM-1} \sum_{k=i+1}^{UM} \left[\frac{8 \ln t}{(\theta_i - \theta_k)^2 (P_d - P_f)^2} + 1 + \frac{\pi^2}{3} \right] \quad (3.30)$$

证毕。

根据引理 3.1、引理 3.2 及式 (3.10), 所提算法有如下结论。

定理 3.1 算法 π^C -IID 的数据量损失期望上界为

$$R(t; \pi^C) \leq \Delta \cdot \sum_{i=UM+1}^N \frac{8UK \ln t}{(\theta_{UM} - \theta_i)^2 (P_d - P_f)^2} + \\ \Delta \cdot \sum_{i=1}^{UM-1} \sum_{k=i+1}^{UM} \frac{8 \ln t}{(\theta_i - \theta_k)^2 (P_d - P_f)^2} + \\ \Delta \cdot UK(N - UM) \left(\frac{UM\pi^2}{3} + 1 \right) + \Delta \cdot \binom{UM}{2} \left(\frac{\pi^2}{3} + 1 \right) \quad (3.31)$$

其中, $\binom{UM}{2} = \frac{UM \cdot (UM - 1)}{2}$ 。

由上述分析可知, 算法 π^C -IID 的数据量损失期望上界与时隙数量的关系为 $O(\ln t)$ 。该结论的直观含义为, 信道统计信息未知下次要用户需用 $O(\ln t)$ 的时间对非最优信道进行感知, 使用 $t - O(\ln t)$ 的时间感知接入最优信道。该结论的潜在含义如下: 当时隙数量较小时, 中心节点将尽可能将不同信道分配给次要用户进行探索; 当时隙数量增大时, 中心节点趋于将最优信道分配给各次要用户进行感知接入。

2. 复杂度分析

本小节分别对算法的空间复杂度和时间复杂度进行分析, 其中, 空间复杂度用来度量算法在运行过程中临时占用存储空间的大小, 时间复杂度表示执行算法所需的计算工作量^[114]。

1) 空间复杂度

在信道感知接入过程中, $\mathbf{T}(t)$ 、 $\mathbf{Y}(t)$ 和 $g_i(t)$, $\forall i \in \mathcal{I}_N$ 的数据长度为 N , 因此中心节点在时隙 t 中以空间复杂度 $O(N)$ 对上述数据进行记录; U 个次要用户信道感知结果的数据长度为 UM , 因此中心节点以空间复杂度 $O(UM)$ 记录上述数据; 中心节点再以空间复杂度 $O(UM)$ 进行频谱决策选择 UM 、 UK 条信道, 因此, 每个时隙中算法 π^C -IID 的空间复杂度为 $O(N)$ 。

2) 时间复杂度

在信道感知接入过程中, $\mathbf{T}(t)$ 、 $\mathbf{Y}(t)$ 和 $g_i(t)$, $\forall i \in \mathcal{I}_N$ 的数据长度为 N , 中心节点在每个时隙中花费 $O(N)$ 的时间对上述数据进行更新。此后, 中心节点花费 $O(N^2)$ 的时间对信道进行排序^①, 花费时间 $O(1)$ 选择 UM 、 UK 条信道进行频谱决策并更新数据量, 因此, 算法 π^C -IID 的时间复杂度为 $O(N^2)$ 。

3.1.5 仿真实验

本节对集中式网络中算法 π^C -IID 在 MATLAB R2018b 软件平台上进行验证。本节采用数据量损失和平均吞吐量对算法有效性进行度量。当平均吞吐量趋于最优吞吐量时, 表明次要用户可有效地识别并接入最优的信道, 从而验证了算法可高效地利用频谱资源。

值得说明的是, 本节仿真中数据量损失采用多臂赌博机研究中常采用的归一化形式^[114], 即 $R(t; \pi^C)/\ln t$ 的形式。其潜在含义如下: 若 $R(t; \pi^C)/\ln t$ 随着时间的增加趋于平缓且小于某一常数 C , 即可验证 $R(t; \pi^C)$ 的上界阶数为 $O(\ln t)$; 若平均吞吐量定义为所获得的数据量与时隙数量的比值, 当平均吞吐量趋于最优平均吞

① 假设本算法中采用直接插入法对 N 条信道进行排序^[115], 最差情况的时间复杂度为 $O(N^2)$ 。

吐量（此时表示次要用户每个时隙中感知接入最优信道所获得的平均吞吐量）时，表明次要用户随时间的增加趋于感知接入最优信道，与减少数据量损失等价。

本节考虑以下场景：① 不同用户数量、信道感知接入数量下仿真验证；② 不同检测概率 P_d 、虚警概率 P_f 下仿真验证；③ 不同信道数量下仿真验证；④ 不同时隙下中心节点对各信道空闲概率估计；⑤ 不同时隙下对各信道感知接入次数；⑥ 不同算法性能比较。设次要用户有连续的数据业务，每空闲时隙中次要用户传输速率为 1 兆比特每秒（megabits per second, Mbit/s）。

1. 不同用户数量、信道感知接入数量

考虑网络中信道数量 $N = 9$ ，信道空闲概率为 $[0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3, 0.2, 0.1]^T$ ，用户数量 $U = 2, 3, 4$ ，信道感知接入数量为 $M = 1, 2, 3$ ，接入数量为 $K = 1, 2$ ，蒙特卡罗次数设置为 65。

在算法 π^C -IID 的仿真中，中心节点按照算法 3.1 中描述的过程，首先令各次要用户感知接入信道一次并更新相关信息。在后续每个时隙中，中心节点首先计算所有信道索引标号，将标号最大的信道分配给各次要用户感知接入，并更新信道相关信息。

为验证算法 π^C -IID 的有效性，在仿真中考虑算法 π^{*C} -IID 作为性能基准，即假设中心节点已知信道空闲概率信息，按照 3.1.3 节描述的步骤和上述场景进行仿真。通过计算 π^{*C} -IID 得到的数据量均值（即多次蒙特卡罗仿真中获得的总数据量除以蒙特卡罗次数）与算法 π^C -IID 得到的数据量均值的差值，可求得数据量损失；通过计算 π^{*C} -IID 得到的数据量均值与时隙数量的比值，可得到最优平均吞吐量。

图 3.2 表示数据量损失与时隙数量的关系，其中横轴表示时隙数量，纵轴表示数据量损失。由结果可知，数据量损失与时隙数量呈亚线性关系，即随着时隙数量的增加，数据量损失的增长趋于平缓，表明中心节点趋于选择最优信道（否则，数据量损失与时隙数量呈线性关系）。此外，相同用户数量 U 、相同信道感知数量 M 下， K 值越大，则数据量损失越大。其原因为 K 越大，集中式网络接入非最优信道次数越多，产生的数据量损失越大；相同 M 、 K 下，用户数量 U 越大，数据量损失越小，其原因为集中式网络每个时隙中感知的信道数量 UM 变大，则对信道空闲概率信息的估计速度增加，更易于在有限时隙数量下选取最优信道进行感知；相同 U 、 K 下， M 越大，则数据量损失越小，其原因为集中式网络每个时隙下感知的信道数量 UM 变大，对未知参数的估计速度增加，更易于选择最优信道进行感知。