

# 无线网络中的信息时效性优化

(申请清华大学工学博士学位论文)

培养单位：电子工程系

学 科：信息与通信工程

研 究 生：汤皓玥

指 导 教 师：王 劲 涛 教 授

二〇二一年十二月



# **Data Freshness Optimization in Wireless Networks**

Dissertation Submitted to  
**Tsinghua University**  
in partial fulfillment of the requirement  
for the degree of  
**Doctor of Philosophy**  
in  
**Information and Communication Engineering**

by

**Haoyue Tang**

Dissertation Supervisor: Professor Wang Jintao

**December, 2021**



## 学位论文公开评阅人和答辩委员会名单

### 公开评阅人名单

牛志升	教授	清华大学
宋健	教授	清华大学
艾渤	教授	北京交通大学

### 答辩委员会名单

主席	周世东	教授	清华大学
委员	牛志升	教授	清华大学
	王劲涛	教授	清华大学
	费泽松	教授	北京理工大学
	杨晨阳	教授	北京航空航天大学
	周盛	副教授	清华大学
秘书	张超	副研究员	清华大学



# 关于学位论文使用授权的说明

本人完全了解清华大学有关保留、使用学位论文的规定，即：

清华大学拥有在著作权法规定范围内学位论文的使用权，其中包括：（1）已获学位的研究生必须按学校规定提交学位论文，学校可以采用影印、缩印或其他复制手段保存研究生上交的学位论文；（2）为教学和科研目的，学校可以将公开的学位论文作为资料在图书馆、资料室等场所供校内师生阅读，或在校园网上供校内师生浏览部分内容；（3）根据《中华人民共和国学位条例暂行实施办法》及上级教育主管部门具体要求，向国家图书馆报送相应的学位论文。

本人保证遵守上述规定。

作者签名: 汤皓明

导师签名: 孙海涛

日期: 2021.10.28

日期: 2021.10.28



## 摘要

在未来自动驾驶系统中，车辆有效和安全的控制依赖于车速、路况等状态信息。这些随时间不断变化的状态信息由部署在车辆和环境中的传感器节点观测和采集。状态信息的时变性要求传感器源源不断进行状态信息采集，通过通信网络直接发送或由基站转发给用户。然而，网络中的随机传输时延、通信系统有限的带宽和传输速率、以及车辆移动所导致的信道时变性为采集和推送时效性强的状态信息带来了挑战。另一方面，好的状态信息时效性与传统通信质量要求，例如高吞吐量与低传输时延有本质不同。因此，我们需要重新设计传输和网络调度算法，优化网络中用户和决策中心的状态信息时效性，从而保证系统的有效和安全性。

本文针对车联网中，基站通过传感器采集、用户从基站主动索取和被动态送状态信息的场景，研究优化状态信息时效性的信息采集和传输调度策略。论文依据状态信息更新特点，选用信息年龄和同步年龄段量信息时效性。主要贡献如下：

1. 点对点链路信息年龄最小化的自适应状态信息采样。考虑状态信息传输存在随机时延，以最小化接收端平均信息年龄为目标，针对传输时延统计信息未知的挑战，设计了自适应在线采样策略。理论分析证明，随着时间的采样数目趋于无穷，算法得到的平均信息年龄趋于最优。
2. 多传感器网络信息年龄最小化的采样、功率控制和调度联合设计。考虑基站通过时变信道从多个传感器采集状态信息，以信息年龄最小化为目标，研究满足带宽约束的传输策略。面对多传感器调度问题中的“维度诅咒”，使用带宽松弛和拉格朗日乘子法进行多传感器解耦，设计了松弛约束下的最优策略，并基于此提出了满足带宽约束的调度策略。理论分析证明，所提策略在大规模网络用户数量趋于无穷时有渐近最优性。
3. 最小化用户索取状态信息副本的信息年龄的基站端缓存更新。考虑基站端缓存与远端服务器之间的链路带宽受限，针对状态信息流行度时变和下载所需时长时变的特点，通过线性规划和分支定界法设计了基站端缓存更新策略，有效降低了用户从基站端缓存得到的状态信息副本的信息年龄。
4. 同步年龄最小化的广播与推送。针对状态信息变化与传输存在双重随机性，以最小化全网用户平均同步年龄为目标，基于怀特指数提出了一种广播推送方案，仿真结果发现所提算法得到的平均同步年龄接近理论下界。

**关键词：**信息年龄；同步年龄；随机网络优化；调度；状态更新系统

## Abstract

The safety and efficient control of autonomous driving systems relies on status information such as driving speed and highway conditions. These status information are time-sensitive, and are observed by sensors deployed on cars and the surrounding environment. The timeliness of status information requires the communication networks collect status updates from sensors, and send them out to interested users in a timely manner. However, the delivery of fresh status to users is limited by the random transmission delay, the limited communication rate and the time-varying channel conditions due to vehicular mobility. Moreover, data freshness requirement is different from traditional quality of service (QoS) optimization goals such as high throughput and low delay. Thus, in order to guarantee the safety and efficient control of real time applications, it is important to redesign transmission strategies and network scheduling algorithms to obtain a good data freshness performance.

This thesis focuses on two scenarios in vehicular networks: the base station collecting status updates from sensors and sending out status updates to interested users. The goal is to propose sampling, transmission and scheduling algorithms that optimize data freshness performance. Depending on the evolution characteristics of status information, two recently proposed metrics, the Age of Information (AoI) and the Age of Synchronization (AoS) are used to measure the freshness of information at the receiver. The main contributions are as follows:

1. Adaptive sampling and transmission strategies in point-to-point communication systems for AoI minimization. Consider a sensor samples status update and transmit them to the destination through a channel with random delay. To overcome the challenge of unknown delay statistics, we propose an online sampling strategy based on Monro-Robbins algorithm. Theoretic analysis proves that the proposed algorithm adaptively learns the optimal sampling policy. The average AoI optimality gap diminishes when the number of samples goes to infinity.
2. AoI optimal sampling, cross-layer transmission and scheduling for multi-sensor networks with time-varying channels. Consider that multiple sensors are connected to the base station via time-varying communication links using a shared bandwidth, the goal is to design scheduling policy that meets the bandwidth constraint of the network and average

---

## Abstract

---

power constraint of the sensors. To overcome “the curse of dimension” in this decision making problem, a sampling and cross-layer scheduling algorithm is proposed through bandwidth relaxation and sensor decoupling. Theoretic analysis shows the asymptotic performance of the algorithm in large scale networks.

3. Caching update strategies for AoI minimization. Consider copies of status information are stored at the base station. When users request for status updates, they obtain the possible outdated copies at the local cache. The goal is to design cache updating strategies for the base station so that the AoI of user requested files can be minimized. AoI minimum cache updating strategies under time-varying files popularity, time-varying update durations are proposed by using the Branch-Reduce-Bound (BRB) and Linear Programming (LP). Simulation results validate the effectiveness of the proposed algorithm in terms of AoI reduction.

4. Data freshness oriented broadcasting in multi-user networks. Consider both status updates generation and transmission encounters randomness. The goal is to minimize the expected average Age of Synchronization (AoS) performance over the entire network by designing efficient broadcasting strategy for the base station. The AoS minimization broadcasting problem is reformulated into a restless multi-arm bandit problem and broadcasting strategy based on the Whittle’s index is proposed. Simulation results show the average AoS performance of the proposed algorithm is close to the theoretic lower bound.

**Keywords:** Age of Information; Age of Synchronization; Stochastic Network Optimization; Scheduling; Status Update

## 目 录

摘要.....	I
Abstract.....	II
目录.....	IV
插图和附表清单.....	VIII
符号和缩略语说明.....	X
第1章 引言 .....	1
1.1 车联网中的状态信息 .....	1
1.2 随机网络中的状态信息时效性优化：约束条件与技术难点 .....	2
1.2.1 状态信息度量.....	2
1.2.2 状态信息时效性优化：难点与挑战.....	4
1.3 随机网络中的状态信息时效性优化研究综述 .....	4
1.3.1 时效性最优的状态信息采集.....	5
1.3.2 时效性最优的状态信息推送.....	7
1.4 研究思路与结构安排 .....	8
1.4.1 研究思路.....	8
1.4.2 结构安排.....	10
第2章 单传感器信息年龄最小化的在线状态信息采集研究 .....	12
2.1 引言 .....	12
2.2 问题建模 .....	12
2.2.1 系统模型.....	12
2.2.2 连续时间信息年龄.....	13
2.2.3 优化问题建模.....	13
2.3 问题求解 .....	15
2.3.1 更新过程优化问题.....	15
2.3.2 一种传输时延分布 $\mathbb{P}_D$ 未知时的在线优化算法 .....	16
2.3.3 算法理论分析.....	18
2.4 仿真结果 .....	19
2.4.1 无采样频率约束下的仿真结果.....	19

2.4.2 有采样频率限制下的仿真结果.....	20
2.5 本章小结 .....	21
2.6 附录 .....	22
2.6.1 问题 2.2 推导.....	22
2.6.2 推论 2.1 证明.....	23
2.6.3 定理 2.2 证明.....	23
2.6.4 引理 2.1 证明.....	26
2.6.5 引理 2.2 证明.....	27
2.6.6 引理 2.3 证明.....	27
2.6.7 定理 2.3 证明.....	30
2.6.8 引理 2.4 证明.....	32
<b>第 3 章 多传感器信息年龄最小化的状态信息采集和跨层传输算法研究 .....</b>	<b>34</b>
3.1 引言 .....	34
3.2 系统模型和问题建模 .....	34
3.2.1 网络模型.....	34
3.2.2 信息年龄.....	35
3.2.3 优化问题建模.....	36
3.3 单传感器解耦及其求解.....	37
3.3.1 单传感器解耦.....	37
3.3.2 受限马尔可夫决策过程及其阈值结构.....	39
3.3.3 受限马尔可夫决策过程的线性规划求解.....	40
3.4 多传感器机会调度 .....	43
3.4.1 时间平均带宽约束下最优策略 $\pi_R^*$ 求解.....	44
3.4.2 一种满足每时隙带宽严格约束的调度方案及其理论特性.....	47
3.5 仿真结果 .....	47
3.5.1 全网平均信息年龄表现.....	48
3.5.2 信息年龄与功率折中关系.....	49
3.6 本章小结 .....	51
3.7 附录 .....	52
3.7.1 引理 3.1 证明.....	52
3.7.2 引理 3.2 证明.....	53
3.7.3 引理 3.3 证明.....	55
3.7.4 定理 3.2 证明.....	56

3.7.5 引理 3.4 证明.....	58
3.7.6 引理 3.5 证明.....	60
<b>第 4 章 信息年龄最小化的基站端缓存更新算法研究 .....</b>	<b>62</b>
4.1 引言 .....	62
4.2 针对状态信息流行度时变的缓存更新方案研究 .....	62
4.2.1 系统模型.....	62
4.2.2 问题求解.....	64
4.2.3 数值仿真.....	69
4.2.4 本节小结.....	70
4.3 针对更新时长随信息年龄变化的基站端缓存更新策略研究 .....	71
4.3.1 系统模型和问题描述.....	71
4.3.2 优化问题建模.....	73
4.3.3 一个可行的松弛问题次优解.....	74
4.3.4 一种可以满足带宽约束的更新策略.....	79
4.3.5 数值仿真.....	80
4.3.6 本节小结.....	83
4.4 本章小结 .....	83
<b>第 5 章 状态更新随机时同步年龄最小化的推送算法研究 .....</b>	<b>84</b>
5.1 引言 .....	84
5.2 系统模型 .....	84
5.2.1 网络模型.....	84
5.2.2 信息年龄与同步年龄.....	85
5.2.3 问题建模.....	87
5.2.4 全网平均同步年龄的理论下界.....	88
5.3 调度策略 1：基于马尔可夫决策过程的调度方案设计.....	88
5.4 调度策略 2：基于怀特指数的调度算法设计.....	91
5.4.1 解耦后的单用户调度策略设计.....	91
5.4.2 可索引性质.....	92
5.4.3 怀特指数 (Whittle's Index) 的推导.....	94
5.4.4 基于怀特指数的调度算法.....	95
5.5 数值仿真 .....	95
5.6 本章小结 .....	97

## 目 录

---

5.7 附录 .....	98
5.7.1 定理 5.1 证明.....	98
5.7.2 推论 5.1 的推导过程.....	102
5.7.3 推论 5.2 的推导过程.....	104
<b>第 6 章 总结与展望 .....</b>	<b>107</b>
6.1 研究内容总结 .....	107
6.2 进一步研究方向 .....	108
<b>参考文献.....</b>	<b>110</b>
<b>致 谢.....</b>	<b>117</b>
<b>声 明.....</b>	<b>119</b>
<b>个人简历、在学期间完成的相关学术成果.....</b>	<b>120</b>
<b>指导教师学术评语.....</b>	<b>123</b>
<b>答辩委员会决议书.....</b>	<b>124</b>

## 插图和附表清单

图 1.1 车联网状态信息采集与推送示意图 .....	1
图 1.2 信息年龄的定义 .....	3
图 1.3 研究内容、挑战及其相互关系 .....	9
图 2.1 一个点对点状态更新系统 .....	12
图 2.2 连续时间信息年龄 $x(t)$ 的演进曲线 .....	14
图 2.3 信息年龄比值 $\hat{X}_k$ 随着采样数目 $k$ 的演进关系图 .....	20
图 2.4 平均信息年龄 $\bar{X}_{\pi,t}$ 随着时间 $t$ 的演进关系图。 .....	21
图 2.5 有采样频率约束下, 针对某一条采样路径的时间平均信息年龄演进 .....	21
图 2.6 使用不同 $V$ 时在线采样策略得到的平均采样间隔 .....	22
图 3.1 多传感器网络状态信息采集系统模型 .....	35
图 3.2 离散时间随意采样 (generate-at-will) 模型中信息年龄演进曲线 .....	36
图 3.3 两状态信道 $Q = 2$ 的 $(x, q)$ 状态跳转示意图 .....	41
图 3.4 全网平均信息年龄随网络传感器数目 $N$ 变化关系图 .....	48
图 3.5 带宽 $M/N$ 保持不变时, 全网平均信息年龄随传感器数目 $N$ 的变化关系图 .....	49
图 3.6 全网平均信息年龄与功率限制因子关系图 .....	50
图 3.7 同一网络中不同功率传感器的平均信息年龄 .....	50
图 3.8 网络中不同功率约束节点在不同状态 $(x, q)$ 下被调度概率 .....	52
图 4.1 远程服务器、基站端缓存、用户系统图 .....	62
图 4.2 被索取状态信息的平均信息年龄与流行度转移概率 $q$ 关系图 .....	70
图 4.3 状态信息 $n$ 的信息年龄演进关系图 .....	72
图 4.4 每份状态信息更新时长为不同的常数时, 用户索取得到状态信息的平均信息年龄随状态信息数量 $N$ 演进曲线图 .....	81
图 4.5 每份状态信息更新时长时变时, 被索取状态信息平均信息年龄随状态信息数量 $N$ 的演进曲线图 .....	82
图 4.6 不同状态信息下载和更新所占用的带宽 .....	82
图 4.7 不同状态信息的更新时间占比 (上图) 和平均信息年龄 .....	83
图 5.1 基站向用户推送状态信息系统图 .....	85
图 5.2 信息年龄和同步年龄演进曲线对比图 .....	86
图 5.3 三用户网络的全网平均同步年龄与全网所有信源更新概率总和的关系 ..	96

图 5.4 全网平均同步年龄与网络用户数 $N$ 的关系 .....	97
图 5.5 用户对应的信源更新概率不同时，他们处于不同同步年龄被调度的概率 .....	97
图 5.6 基站到用户传输成功概率不同时，他们处于不同同步年龄被调度的概率 .....	98
图 5.7 用户的同步年龄状态转移图。 .....	102
图 6.1 研究内容总结 .....	109

## 符号和缩略语说明

AoI	信息年龄 (Age of Information)
AoS	同步年龄 (Age of Synchronization)
KKT	KKT 条件 (Karush-Kuhn-Tucker Condition)
MDP	马尔可夫决策过程 (Markov Decision Process)
CMDP	受限马尔可夫决策过程 (Constrained Markov Decision Process)
SMDP	半马尔可夫决策过程 (Semi-Markov Decision Process)
LP	线性规划 (Linear Programming)
RMAB	无休止多臂赌博机 (Restless Multi-Armed Bandit)
i.i.d	独立同分布 (independent and identically distribution)
BRB	分支定界法 (Branch-Reduce-Bound)
RR	更新过程 (Renewal-Reward-Process)
QoS	服务质量 (Quality of Service)
$\mathbb{E}$	期望运算符
$\mathbb{M}$	均值运算符
Pr	概率测度
$\mathcal{A}$	集合 $\mathcal{A}$ 的基数
$\mathbb{R}$	实数集合
$x(t)$	时刻 $t$ 的信息年龄
$\mathcal{O}$	渐近符号
$\mathbf{a}$	矢量 $\mathbf{a}$
$\mathbf{A}$	矩阵 $\mathbf{A}$
$(\cdot)^T$	矩阵或矢量的转置

# 第1章 引言

## 1.1 车联网中的状态信息

随着通信技术和硬件设备的不断进步，自动驾驶系统蓬勃发展并受到了广泛关注。根据 Navigant Research 的预测，自动驾驶汽车将在 2035 年达到全球年销量 9,540 万辆，占据整体车辆市场的 75%<sup>[1]</sup>。自动驾驶和车联网也被列入我国顶层发展规划<sup>[2-3]</sup>。

自动驾驶中的安全性和有效控制依赖于准确及时的路况和车速等状态信息。这些状态信息由部署在车辆和环境中的传感器采集，比如车辆摄像头的图像可以为车辆提供路况和环境信息、车胎上部署的传感器为自动驾驶提供胎压和车速信息等。由于车况等信息随着时间不断变化，如图1.1所示，为了保证用户和决策中心的状态信息时效性，车联网中基站通过网络从传感器不断采集这些带有时间戳的信息，用户和网络中的决策中心从基站端缓存主动索取或者被动广播推送状态信息。

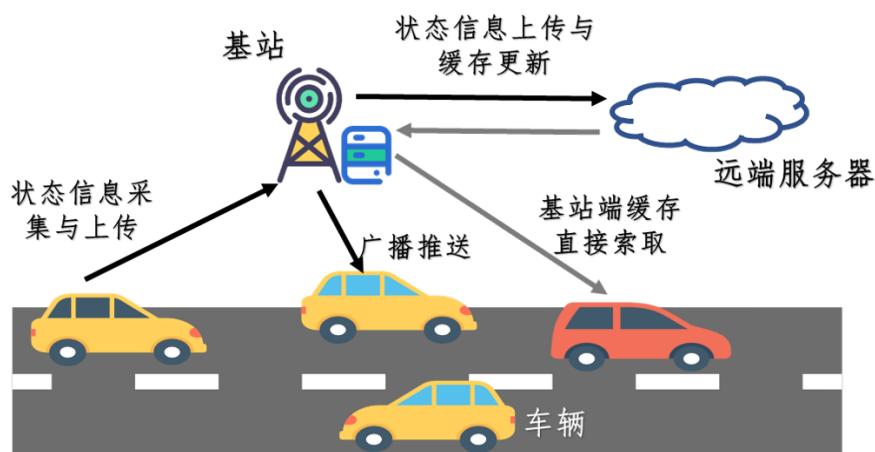


图 1.1 车联网状态信息采集与推送示意图

综上所述，状态信息时效性对于车联网和自动驾驶的有效安全控制尤为重要。而状态信息的采集和分发过程均存在着通信时延和传输随机性，针对这些传输随机性，我们需要设计合理的调度传输算法，确保车联网中用户和决策中心的状态信息时效性。

## 1.2 随机网络中的状态信息时效性优化：约束条件与技术难点

然而，信息接收端并不总是拥有最新的状态信息。回顾图1.1中车联网系统的状态信息采集与推送过程，具体原因有以下三方面：一是状态信息由接入网络的传感器进行采集，如果传感器因为功率等原因并不总是能连续不断地生成和发送状态信息数据包，那么接收端的信息就一定不是最新的；二是来源于通信过程中的随机性。由于信道衰落的随机性，状态信息的采集和推送过程均存在着通信随机性，那些被成功接收的数据包在传输过程中也不可避免地会经历排队和传输时延；三是为了降低通信时延，用户索取状态信息时可以直接从基站端缓存获取状态信息副本，但由于基站端缓存与云端服务器之间的通信链路受限，基站端缓存所保存的状态信息副本并不一定是最新的。因此，我们需要设计优化调度算法，对抗通信网络传输的随机性，从而获得保证接收端的信息时效性，确保系统的安全有效控制。

### 1.2.1 状态信息度量

在建立研究的数学模型之前，我们首先需要一个定量的指标，度量状态信息时效性。注意到接收端所拥有状态信息的产生时刻相比于当前时刻总是有时间差，也就是说，接收端的状态信息总是过时的。进一步地，我们所选取的度量需要可以刻画收端信息过时的程度。我们将首先说明，那些常被用于度量网络中服务质量 (Quality of Service, QoS) 的度量，例如传输时延和吞吐量，并不能刻画接收端信息时效性的过时程度：

- **传输时延：**传输时延可以刻画每个状态更新数据包被发送给接收端所需要的时间。用传输时延度量接收端信息时效性有两个问题：一方面那些在传输过程中因为信道衰落和编解码错误而传输失败的数据包没有时延；另一方面，数据包的传输时延短并不意味着接收端的信息时效性好，例如在路况更新中，如果每过十分钟才进行一次路况信息采样，用户成功收到下一次路况更新前会有至少十分钟未收到新的路况，这十分钟内他所拥有的路况信息就不再新了。在这样的情形中，虽然状态时效性较差，但通信时延无法刻画这一较差的时效性。
- **吞吐量：**信道的吞吐量表示单位时间内传输的比特数，但这一针对通信信道的刻画并不能反映信息时效性。比如在先到先服务系统中，当系统吞吐量逼近稳定区域极限，传输时延会急剧增大<sup>[4]</sup>。显然，一个较大的传输时延将会导致状态信息的传输需要经过很长时间才能被接收到，这一较长的传输时延会导致用户所拥有的状态信息变得非常过时。因此，系统吞吐量大并不能等

同于状态信息时效性强。

- **估计误差：**实际上，我们可以利用当前接收端对于状态信息的估计与状态信息当前的差异作为信息时效性的度量。使用这一度量要求我们拥有状态信息变化的先验知识。注意到在车联网系统中，状态信息包括路况、车速、车距等，信息类别多样演进过程复杂，拥有这些状态信息变化的先验知识并不可行。

因此，我们需要一种较为普适地可以刻画信息时效性的度量。一种新的度量，信息年龄，近年来受到了广泛的关注。根据定义，接收端当前的信息年龄为当前时刻与接收端所拥有最新信息的产生时刻之间的时间差<sup>[5]</sup>，它演进关系如图1.2所示。由图可见，接收端在每时每刻都有信息年龄，同时，一旦新的数据包被接收到，信息年龄将会降低为数据包的传输时延，而如果没有新的状态信息被收到，信息年龄将会持续增长。因此，信息年龄可以刻画状态信息过时的程度，我们希望通信系统中，接收端的信息年龄越小越好。注意到在信息年龄的每次增长和下降演进中，最大值依赖于两次数据包接收之间的时间间隔，最小值为数据包的传输时延。因此为了保证接收端信息时效性好，系统需要兼具高吞吐量和低传输时延。

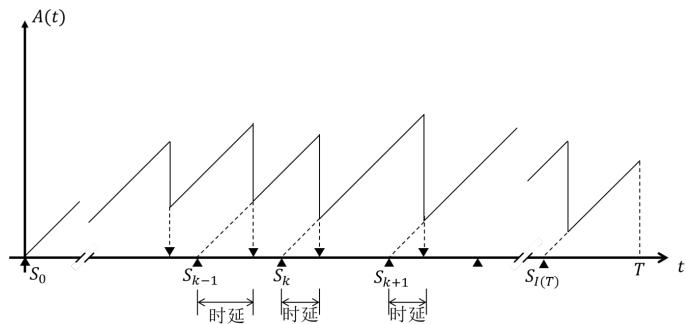


图 1.2 信息年龄的定义

如果状态信息在相邻两次更新数据包接收的区间内没有变化，那么接收端所拥有的信息时效性将不发生变化。然而依据定义，接收端的信息年龄将不断增长。在这样的场景中，使用信息年龄度量时效性并不合理。作为信息年龄的补充，同步年龄<sup>[6]</sup>可以用于度量当前时刻与接收端不同步时刻之间的时间差。

本论文中，我们将依据状态信息更新特点，分别使用信息年龄和同步年龄度量状态信息时效性。接下来，我们将讨论在信息年龄和同步年龄最小化方面的挑战和难点。

### 1.2.2 状态信息时效性优化：难点与挑战

我们将依据图1.1中状态信息传输过程，分析在状态信息采集、用户主动索取和广播推送环节中的难点与挑战：

- **传感器功率限制：**状态信息由网络接入传感器进行采集。然而，网络所能提供给传感器的电量是有限的，这意味着这些传感器并不总是能源源不断地采集最新的状态信息进行发送，传感器需要依据自身信道状态进行传输功率和采样策略的设计，在满足功率约束的条件下尽可能地为接收端提供最新的状态信息。
- **传感器与基站通信链路随机性：**状态信息由传感器推送给基站的过程中需要经过无线信道。由于车联网中车辆的移动性，连接传感器与通信链路之间的信道是时变的。另外网关处存在着数据包排队。以上现象导致了状态信息在采集的过程中存在着丢包、时延等随机性。
- **基站端缓存处状态信息副本的过时性：**部分状态信息（如地图等信息）保存在云端服务器上。为了降低用户在索取这部分信息时的通信时延，基站提前将这些状态信息副本保存在基站端缓存中供用户索取。基站端缓存与云端服务器带宽受限的链路将导致：基站端缓存不能随时随地获取云端服务器上最新的状态信息，这一约束将导致基站端缓存所拥有的副本信息过时。
- **状态信息推送中的带宽限制：**通信网络需要通过有限的带宽将信息推送给用户。由于网络中接入用户数量多，因此用户并不是每时每刻都可以被基站服务并获得自己感兴趣的状态信息更新。在没有被服务的时刻，用户未能收到更新信息，因此信息时效性变差，信息年龄增大。
- **不同信源之间差异：**每一个状态信息的变化有一定差异，在刻画和优化信息时效性时，我们需要将状态信息之间的差异纳入考量。

总的来说，信息时效性优化的核心挑战在于，如何依据动态变化的信道、用户需求和状态信息特点，适配有限的传输带宽、功率等通信资源，对抗通信中的随机性，从而保证发送给用户的状态信息时效性。

## 1.3 随机网络中的状态信息时效性优化研究综述

本节将对随机网络中传输资源受限情况时，时效性最优的信息采集与推送的相关研究进行综述。

### 1.3.1 时效性最优的状态信息采集

#### 1.3.1.1 单链路系统

针对单传感器状态信息采集系统，已有研究分别针对状态信息更新数据包由外部信源控制、由传感器自行控制两种场景，以最小化平均信息年龄为目标，设计了相应的采集与跨层优化算法。

当状态信息更新数据包由外部信源产生时，文献<sup>[5,7-8]</sup>分析了使用不同传输顺序所得到的时间平均信息年龄。这些分析发现高吞吐和低传输时延并不能保证接收端拥有较小的平均信息年龄，我们需要根据通信系统的约束条件，重新设计信息年龄最小、时效性最优的调度和传输策略。其中，离散时间系统中的平均信息年龄最小化问题可以被建模为一个马尔可夫决策过程<sup>[9-16]</sup>。文献<sup>[11]</sup>考虑传输功率受限时，最小化接收端信息年龄的数据包选择与发送联合设计策略；针对传输信道时变的场景，文献<sup>[9]</sup>提出了一种数据包选择和跨层优化联合设计，可以在一定功率限制下最小化接收端的信息年龄；进一步地，考虑传输存在丢包，文献<sup>[12]</sup>结合混合重传控制协议 (Hybrid Automatic Repeat Request, HARQ)，设计了一种满足功率约束的信息年龄最小化重传与数据包更新机制；文献<sup>[10]</sup>考虑了传输带宽受限时，信息年龄最小的服务中断策略；

当传输节点可以随时采样状态信息，生成并发送新的数据包时，采样、服务顺序和跨层传输的联合优化可以有效降低接收端的平均信息年龄<sup>[13,15-18]</sup>。在连续时间场景中，最优化时效性的采样策略设计可以被转化为一个半马尔可夫决策过程 (Semi-Markov Decision Process, SMDP)，文献<sup>[15-16]</sup>证明了该 SMDP 问题的最优解存在阈值结构，即当接收端的信息年龄超过一定阈值后，最优策略将立刻采样新的状态信息并发送给接收端。

我们注意到以上信息年龄最优的采样策略设计需要通信信道的统计信息已知。当信道的统计信息（如传输时延分布、丢包概率等）未知时，信息年龄最小的传输策略设计可以被建模为一个在线序列决策问题，在线学习和强化学习算法可以帮助我们自适应地设计传输策略，使得接收端信息年龄最小。当状态更新数据包生成由外部控制时，文献<sup>[19-20]</sup>基于多臂赌博机理论 (Multi-Armed Bandit, MAB) 设计了两种时效性最优的传输链路选择算法；当接收端的信息时效性被建模为一个有关信息年龄的时变惩罚函数时，文献<sup>[21]</sup>基于敌对赌博机理论 (Adversarial Bandits) 设计了一种鲁棒的在线时效性优化策略。以上算法理论上都被证明可以达到次线性遗憾界，即随着系统运行时间变长，所提算法逐渐收敛到最优。当传感器可以控制采样过程和状态信息数据包生成时，文献<sup>[22]</sup>将系统的时效性需求建模为平均信息年龄最大值的约束条件，并提出了一种可以满足该约束条件的网络效用最大化

调度算法。这一研究的主要缺陷在于优化目标仍然为优化网络最大效用，并非接收端状态信息时效性，即信息年龄。文献<sup>[13,23-26]</sup>使用了不同的强化学习方法（如Q-Learning、SARSA、Actor-Critic）等进行了最小化信息年龄的采样、调度联合设计，并通过仿真说明了这些算法在网络统计信息未知时的有效性。然而，这些算法的收敛性和收敛速度并没有严格数学分析。同时，使用深度强化学习有较大的计算开销和计算复杂度，这意味着这些算法难以应用到动态变化的网络中。

综上所述，在单传感器状态信息采集问题上，现有工作多是针对信道统计信息已知的场景设计采样和传输策略。虽有部分工作使用在线学习和强化学习等算法对信道统计信息未知信息年龄最小的调度问题进行了研究，但这些算法缺少收敛性和遗憾界分析。

### 1.3.1.2 多传感器网络

在单传感器最优传输方案的基础上，研究者们进一步考虑带宽受限的多传感器网络，设计了最小化中心控制节点信息年龄的采样和调度策略<sup>[27-39]</sup>。当信源产生过程由外部随机过程控制时，文献<sup>[32]</sup>考虑每次传输理想不会丢包的场景，设计了一种信息年龄最小的传输调度的方案；文献<sup>[33,38]</sup>提出了一种分布式的节点状态信息上传算法，证明了所提算法在大规模网络中的渐近最优化；针对传输过程可能存在丢包的场景，文献<sup>[37]</sup>通过无休止多臂赌博机 (Restless Multi-Armed Bandit) 框架，近似计算了用户再不同信息年龄下的怀特指数 (Whittle's index)，提出了一种可以有效减小全网平均信息年龄的调度方案；为了降低近似怀特指数的计算复杂度，文献<sup>[36]</sup>基于李雅普诺夫 (Lyapunov) 优化框架设计了两种最小化全网平均信息年龄的调度算法，并分析了他们的理论最优化。当数据包的生成过程可以由传感器控制时，为了优化信息年龄，我们还需要将采样过程纳入联合设计中。其中，文献<sup>[35]</sup>针对传感器与中央控制节点信道时不变的场景，提出了四种信息年龄最小化的传输策略；特别地，文献<sup>[40-41]</sup>证明了文献<sup>[27-28]</sup>所提出的基于怀特指数的调度算法在大规模网络中的渐近最优化；进一步考虑对于节点收集到的信息还有总吞吐量、单位时间有效吞吐量的约束，文献<sup>[35,39]</sup>基于李雅普诺夫漂流加惩罚 (Lyapunov Drift Plus Penalty, DPP) 框架设计了几种有理论最优化保证的调度算法；我们注意到以上研究都假设传输信道时不变，而在车联网等时效性要求较高的网络中，移动性导致了传感器与中央控制节点的信道经历快衰落时变。以上这些算法无法直接推广到信道时变的网络中使用。

文献<sup>[42-49]</sup>对多用户网络中信道时变的场景进行了建模，并提出了优化吞吐量、传输时延的调度、传输方案。其中，论文<sup>[42-43,49]</sup>将信道状态随时间的变化建模成马尔可夫链，通过值迭代<sup>[43,49]</sup>和怀特指数近似<sup>[42]</sup>的方法设计了传输时延最小的

功率选择和调度策略。注意到值迭代方法复杂度较高，而基于怀特指数近似的调度方案缺乏理论保证。同时，正如我们在节1.2所指出的，这些针对吞吐量和时延的优化并不能保证网络中的用户有一个好的时效性表现。

目前已有部分工作针对多用户网络中信道状态时变时，信息年龄最小的调度方案展开了研究。其中，研究<sup>[30-31]</sup>考虑了中心控制器和用户间信道时变的调度策略设计，但该论文假设信道只有“好”和“坏”两种状态，这一假设较为简单，无法刻画通信网络中信道复杂多变的衰落场景；文献<sup>[50]</sup>提出了一种网络中信道时变时，信息年龄渐近最优的传输调度方案，但该方案最优性成立的条件之一是网络中的带宽多于网络中用户数，这一条件在实际系统中较难实现；文献<sup>[51]</sup>针对一类特定的有记忆的衰落信道 Gilbert-Elliott 信道设计了接入调度准则，仿真结果说明了该方案在降低全网信息年龄方面的有效性。但该方案缺少理论最优性保证。

综上所述，多用户网络中信息采集与调度问题，尤其是用于上传状态信息的信道时变的场景，优化全网信息年龄时效性的信息采集与调度策略研究仍处于起步阶段。需要进一步设计可以适应传输信道时变特点、有理论保证的多用户信息采集与调度策略。

### 1.3.2 时效性最优的状态信息推送

#### 1.3.2.1 缓存更新策略

如节1.2所述，好的信息时效性要求通信内容有较小的通信时延。为了减小状态信息分发过程中的时延，缓存<sup>[52-53]</sup>技术将用户可能索取文件的副本提前存储在靠近用户的基站中，当用户索取文件时，直接从基站端缓存中获取副本。由于用户不再需要通过基站从服务器中索取信息，因此，基站端缓存技术可以有效减小传输时延并成为近年来的研究热点<sup>[54-55]</sup>。

当远程服务器上的状态信息数量较多且随着时间变化时，基站端缓存需要通过带宽受限的信道不断更新基站端缓存中的文件。已有针对缓存更新的研究主要优化系统命中率<sup>[53]</sup>从而减小实时通信速率<sup>[56]</sup>。而这些指标并不能保证用户从基站端缓存中得到的文件时效性尽可能好。

实际上，用户兴趣也应当是网络缓存策略更新设计的重要依据<sup>[57-58]</sup>。目前通信系统中针对信息年龄的优化<sup>[15,27-28,32]</sup>大多假设用户的兴趣是完全不变的，而考虑用户兴趣随时间变化的研究中<sup>[59-61]</sup>，他们缓存策略设计的出发点仍然是优化用户索取命中率。

综上所述，当用户从基站端缓存索取得到的文件是动态变化的状态信息时。针对基站端缓存与远程服务器连接带宽受限的场景，需要依据用户动态变化的需求

和下载时间变化，设计缓存更新和推送策略，保证用户得到的信息具有较好的信息时效性。

### 1.3.2.2 多用户网络广播推送

当新的状态信息生成由外部随机过程控制时，如何通过带宽受限的随机网络将他们分发给感兴趣的用户引起了广泛的关注<sup>[27-29,32-35,62-64]</sup>。其中，文献<sup>[32-34]</sup>以最小化网络中中央控制器的信息年龄为目标，分别基于无休止多臂赌博机和李雅普诺夫框架设计了两种低复杂度的调度算法。实际上，当状态信息代表控制系统中的警报<sup>[65]</sup>或是新的网页<sup>[66]</sup>，状态信息产生的间隔可能比较长，在没有新的状态信息产生时，接收端所拥有的过去的信息完全可以刻画当前的状态，但信息年龄反而很大。这一现象说明了信息年龄在刻画状态时效性方面的局限性。文献<sup>[67]</sup>将状态信息由外部控制时的时效性表现建模为服务间隔，而文献<sup>[68]</sup>利用错误状态惩罚函数度量状态信息时效性。我们注意到，这两种度量只能刻画状态的不同步，无法体现不同步时间所带来的影响。

为了刻画状态信息的生成由外部控制时接收端的信息时效性，一种新的度量—“同步年龄”（Age of Synchronization, AoS）被提出<sup>[6]</sup>。根据定义，同步年龄表示当前时刻与接收端的信息最近一次不同步的时间差。相比于信息年龄，同步年龄可以更好地反映出接收端所保存的信息在内容上的不同步状态。文献<sup>[69-70]</sup>研究了数据库中针对同步状态的文件下载策略，文献<sup>[71]</sup>利用李雅普诺夫框架提出了一种在重负载区域内同步年龄渐近最优的调度算法，进一步地，文献<sup>[34]</sup>使用李雅普诺夫漂流加惩罚框架设计了一种可以满足网络中用户异质需求的调度方案。但这些研究都假设传输过程是理想的，即每次文件更新和下载都将成功。与有线网络不同，由于无线信道的随机性、编解码过程可能出现的错误，传输过程非理想有丢包是无线通信的重要特点。

综上所述，在时效性强的状态信息推送方面，现有研究对于网络传输条件的假设过于理想。

## 1.4 研究思路与结构安排

### 1.4.1 研究思路

未来车联网、工业物联网中接入的海量节点将为网络中的决策中心提供海量状态信息。网络中信道的随机性、状态信息的重要程度变化以及状态信息的动态变化特点为采集和推送时效性强的状态信息带来了挑战。本文分别针对状态信息的采集和推送两种场景，以信息年龄和同步年龄为时效性度量指标，研究如何适

应信道变化随机性、传输时间和用户兴趣的动态变化，在有限的通信资源下通过采样、调度、跨层优化等联合设计，最小化用户信息年龄和同步年龄表现。研究内容、挑战及相互关系如图1.3所示。

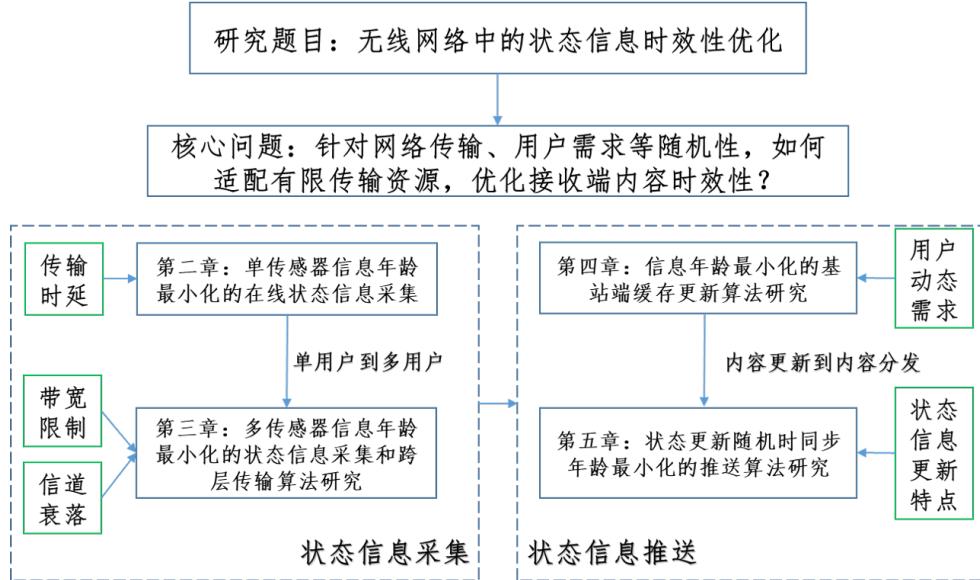


图 1.3 研究内容、挑战及其相互关系

具体研究思路如下：

- (1) 单传感器状态信息采集系统中，信息年龄最小的自适应传输优化策略研究。考虑发送端通过有随机传输时延的信道，采样最新的状态信息并发送给接收端。过去的研究<sup>[16]</sup>表明，信息年龄最小的采样策略具有阈值结构，即接收端的信息年龄在某一阈值以下时，发送端将进行等待，而接收端的信息年龄一旦超过某一阈值，接收端将立刻采样并发送新的数据包，这一最优阈值需要依据传输时延分布的闭式表达式进行计算。本论文研究解决的核心问题是：在信道传输时延统计信息未知时，如何进行在线采样算法设计，逐步学习出时效性最优的等待策略，从而最小化接收端信息年龄？为此，本文提出了“边等待，边学习”的解决思路，基于随机优化中的 Monro-Robbins 算法<sup>[72]</sup>，设计了在线等待策略。Monro-Robbins 算法是随机优化中的一个重要技术，适用于随机梯度下降（Stochastic Gradient Descent, SGD）<sup>[73]</sup>、更新过程优化（Renewal-Reward Process）<sup>[74]</sup>等多种场景。信息年龄的演进并不是一个天然的更新过程，而信息年龄演进中的潜在的无界性也为 Monro-Robbins 算法的直接使用带来了困难。针对以上难点，我们对于信息年龄演进关系进行了分析，通过合理的切割将信息年龄最小化问题转化为更新过程优化问题，根据信息年龄演进的性质修改了 Monro-Robbins 算法。进一步地，我们分析

了使用该算法时学习代价（learning regret）的上界，即对比传输时延已知的理想策略的最大性能损失。

- (2) 多传感器网络中，传输信道时变、带宽受限时，全网平均信息年龄最小的状态信息采样策略研究。本论文研究解决的核心问题是：如何机会地利用信道的起伏衰落，设计合理的采样、跨层优化算法从传感器中采样，使得全网平均信息年龄较小？该问题可以被转化为一个多用户离散时间决策问题，解决这一问题的核心挑战是决策空间和状态空间的维度随着用户数指数增长。为此，我们通过带宽松弛和拉格朗日乘子法将多用户决策问题转化为单用户代价最小化问题，并利用最优解的阈值结构将其进一步转化为线性规划问题求解。这些单用户代价最小化问题的最优决策帮助我们之后通过截断调度，最终设计了一个可以满足每时刻网络中带宽严格约束的调度方案。进一步地，我们证明了所提算法在大规模网络中的理论最优性。
- (3) 为了减小用户索取状态信息这一过程中的时延，我们考虑基站将状态信息副本保存在缓存中，用户直接从基站端缓存获取状态信息。由于带宽和通信资源的限制，基站端缓存中所保存的状态信息可能不是最新的。而为了能够向用户持续不断提供时效性强的状态信息，基站端缓存需要通过有限的带宽不断更新状态信息副本。本论文研究解决的核心问题是：如何根据用户的需求，即不同状态信息的流行度，设计满足通信资源限制的基站端缓存更新方案？我们首先考虑状态信息更新时长随时间变化这一场景，将信息年龄最小的状态更新问题转化为一个有关更新时间占用比例的单调优化问题，提出了一种满足带宽约束的基站端缓存更新策略；进一步地，我们考虑状态信息的流行度随时间变化的场景，在传输带宽限制下，通过带宽松弛和拉格朗日乘子法，将多状态信息更新问题转化为单状态信息代价最小化问题，并设计了在状态信息数量足够多时，信息年龄渐近最优的缓存更新方案。
- (4) 进一步地，考虑状态信息产生过程中的随机性。我们选用同步年龄<sup>[6]</sup>这一新度量刻画用户端所保存的状态信息与信源。本论文研究解决的核心问题时：如何在状态信息更新与传输信道由双重随机性下，根据基站-用户信道特点和状态信息更新特点，设计推送方案，使得网络中所有用户的同步年龄最小？本文将同步年龄最小的调度问题转化为多臂赌博机问题，利用怀特指数设计了一种低复杂度、接近全网同步年龄理论下界的调度算法。

#### 1.4.2 结构安排

基于1.4.1的研究思路，本文后续章节的内容安排如下：

第2章针对单传感器状态信息采集系统中传输时延统计信息未知的场景，以最

小化接收端的平均信息年龄为目标,提出了“边等待,边学习”的采样算法。我们将信息年龄最小化问题建模为更新过程优化问题,提出了基于在线学习的自适应采样发送策略,并通过理论分析,得到了使用该策略平均信息年龄与信道统计信息已知时最小信息年龄差距的上界,证明了随着采样数目增加,所提算法可以学习得到统计信息已知时信息年龄最小的采样策略。本章相关工作已被 IEEE Infocom2022 收录<sup>[75]</sup>,另有一篇论文拟投稿 IEEE Transactions on Information Theory。

第 3 章在第 2 章单传感器信息年龄最小化采样策略设计的基础上,针对多传感器网络中带宽受限、传输信道时变这一场景,设计了最小化全网平均信息年龄的采样和跨层优化方案。我们通过将多传感器信息年龄最小化的采样和跨层优化问题解耦为单传感器代价最小化问题,利用单用户问题的阈值策略,设计了一种满足系统带宽约束的调度方案,并证明了所提算法在大规模网络中的最优性。本章相关工作已经发表在 2019 57th Annual Allerton Conference on Communication, Control, and Computing (Allerton)<sup>[76]</sup> 国际会议论文集以及 IEEE Journal on Selected Areas in Communications 期刊<sup>[77]</sup>。

第 4 章考虑状态信息推送中基站端缓存的内容更新策略。为了减轻通信时延,基站端缓存将远程服务器中的文件副本提前保存在基站中,当用户索取文件中,他们可以直接从基站缓存索取文件副本。本章工作考虑远程服务器中的文件保存动态变化的状态信息(如实时路况、新闻等),缓存系统通过一个带宽受限的网络,从远程服务器更新文件副本这一场景,针对文件更新时间和文件流行度随时间变化两种情况,在带宽限制条件下,优化用户获取到文件信息年龄的缓存更新算法。本章相关工作已经发表在 2020 Information Theory Workshop<sup>[78]</sup> 和 IEEE 18th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOPT 2020 Workshops)<sup>[79]</sup> 会议论文集中。

第 5 章考虑状态信息由外部随机过程产生,基站将状态信息广播推送给用户。本章使用同步年龄度量用户信息时效性。在推送带宽约束下,全网同步年龄最小化的调度问题可以被转化为无休止多臂赌博机。论文推导了全网平均同步年龄的理论下界,基于马尔可夫决策过程和怀特指数提出了两种调度方案,所提算法可以得到达到接近于理论下界的全网平均同步年龄。本章相关工作已经发表在 2019 International Symposium on Information Theory<sup>[80]</sup>, 2020 International Conference on Computer Communications Workshops<sup>[81]</sup> 以及 IEEE Transactions on Wireless Communications 期刊<sup>[82]</sup> 上。

第 6 章总结了本论文的主要研究内容和相应创新点,并对未来研究方向进行了展望。

## 第2章 单传感器信息年龄最小化的在线状态信息采集研究

### 2.1 引言

节1.2已经指出，时效性最优并不等同于吞吐量最大或是传输时延最短。我们需要依据不同网络中的通信资源限制、传输过程中的随机性，重新设计传输调度算法，优化接收端信息时效性。

在车联网中，车辆和环境所搭载的传感器对环境进行采样并通过网络将状态信息发送给接收端。这一过程可以被建模为一个点对点状态更新系统。本章以最小化点对点系统接收端平均信息年龄为目标，在信道传输统计信息未知时设计在线采样传输策略。研究将信息年龄最小的采样策略设计问题转化为一个更新过程优化问题，基于随机优化 Monro-Robbins<sup>[72,74]</sup>算法，在线设计了一种自适应采样策略，所得到的平均信息年龄在采样数目增多时收敛到最优。本章进一步分析了自适应采样策略的学习代价，即对比传输时延统计信息已知时信息年龄最小策略的性能损失，并在数值仿真中验证了所提算法的理论性能。

### 2.2 问题建模

#### 2.2.1 系统模型

与文献类似<sup>[16,83]</sup>，我们考虑如图2.1所示的点对点传输系统。其中传感器观测一个时变物理过程，采样他的状态信息形成数据包，并将这些数据包通过信道发送给接收端。数据包在信道中的发送顺序为先到先服务（First-Come-First-Serve, FCFS），每个数据包在传输过程中的传输时延是随机的。受到传输时延和服务顺序的影响，在上一个数据包的传输完成前，其他的数据包将在队列中等待。每当一个数据包完成了传输并被成功接收，一个传输反馈 ACK 将被立刻无差错地发送给传感器。

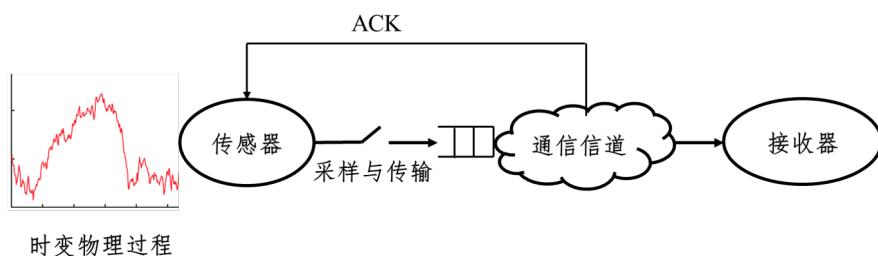


图 2.1 一个点对点状态更新系统

与文献<sup>[15]</sup>相同，我们考虑连续时间系统，传感器可以在任意时间  $t \in \mathbb{R}^+$  采样并发送数据包。记第  $k$  个状态更新数据包的产生时刻为  $S_k$ ，在信道中的传输时延为  $D_k$ 。我们假设每次传输时延独立同分布，其概率测度  $\mathbb{P}_D$  满足如下假设：

**假设 2.1：** 概率测度  $\mathbb{P}_D$  在  $[0, \infty)$  上绝对连续，它的期望和二阶矩上下有界，即：

$$0 < \bar{D}_{lb} \leq \bar{D} \triangleq \mathbb{E}_{\mathbb{P}_D}[D] \leq \bar{D}_{ub} < \infty, \quad (2.1a)$$

$$0 < M_{lb} \leq \mathbb{E}_{\mathbb{P}_D}[D^2] \leq M_{ub} < \infty. \quad (2.1b)$$

记  $R_k$  为第  $k$  个状态更新数据包被成功接收的时刻，它可以通过如下公式迭代计算：

$$R_k = \max\{R_{k-1}, S_k\} + D_k.$$

如果第  $k$  个数据包的采样在第  $k-1$  个数据包被成功接收之前，即  $R_{k-1} > S_k$ ，那么第  $k$  个数据包就将在队列中等待直到时刻  $R_{k-1}$  后才可以进入信道被发送给接收端，他的时效性也将因此变差。所以说，为了保证接收器的时效性最新，一个好的采样策略应“等待”第  $(k-1)$  个状态更新数据包被成功接收，即时间  $R_k$  之后，再进行下一次采样，也就是说  $S_k \geq R_{k-1}$ 。使用这样的“等待”策略，第  $k$  个状态更新数据包被成功接收的时刻  $R_k$  可以被简化为  $R_k = S_k + D_k$ 。为了方便之后的分析，我们记  $W_k := S_{k+1} - R_k$  为收到第  $k$  个数据包后，在采样和传输第  $(k+1)$  个状态更新数据包之前的等待时间。

## 2.2.2 连续时间信息年龄

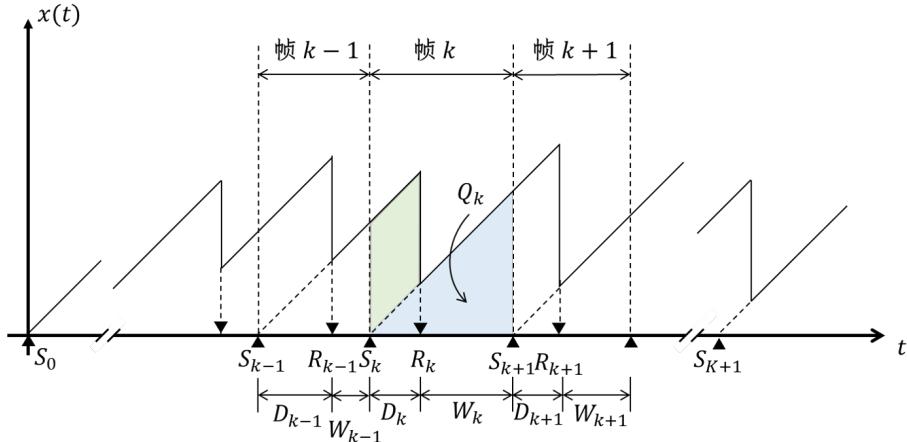
接下来我们介绍章节中考虑的连续时间信息年龄的定义和演进。记  $i(t) := \arg \max_{k \in \mathbb{N}} \{k | R_k \leq t\}$  为时间  $t$  接收端所拥有的最新信息编号。根据定义<sup>[5]</sup>，接收端在时间  $t$  的信息年龄  $x(t)$  可以通过如下公式计算：

$$x(t) \triangleq t - S_{i(t)}. \quad (2.2)$$

为了方便之后的分析和计算，我们在图2.2中展示了一条采样路径（sample path）中信息年龄  $x(t)$  的演进曲线。

## 2.2.3 优化问题建模

我们希望通过设计一个采样和更新策略  $\pi$ ，也就是一系列采样等待时间  $\pi \triangleq \{W_1, W_2, \dots\}$ ，使得接收器的平均信息年龄最小。特别地，我们仅考虑“因果”采样策略，即第  $(k+1)$  次采样前的等待时间  $W_k$  仅由过去的采样和传输时延  $\{(S_i, D_i)\}_{i=1}^k$  决定，未来的采样传输时延  $\{D_i\}_{i>k}$  不可预测也不可用于策略设计。为了方便之后的分析，我们假设每次等待时间都有一个上界  $W_{ub}$ ，并记  $\Pi$  为满足采样等待时间

图 2.2 连续时间信息年龄  $x(t)$  的演进曲线

$W_k \in [0, W_{ub}]$ ,  $\forall k$  的所有因果策略集合<sup>①</sup>. 记  $K$  为观察时间内传感器进行采样的总次数, 使用策略  $\pi$  得到的平均信息年龄  $J(\pi)$  定义为<sup>②</sup>:

$$J(\pi) \triangleq \limsup_{K \rightarrow \infty} \frac{\mathbb{E} \left[ \int_{t=0}^{S_{K+1}} x(t) dt \right]}{\mathbb{E}[S_{K+1}]}, \quad (2.3)$$

其中的数学期望是针对传输时延  $D_k$  和等待时间  $W_k$  求取的。

为了方便之后的计算和分析, 我们定义“帧”  $k$  为第  $k$  次和第  $(k+1)$  次采样之间的时间间隔。由于传输时延  $D_k$  在每一帧中均以概率  $\mathbb{P}_D$  独立同分布, 第  $K+1$  次采样时间  $S_{K+1}$  的期望可以写成:  $\mathbb{E}[S_{K+1}] = \mathbb{E} \left[ \sum_{k=1}^K (D_k + W_k) \right]$ 。类似地, 记  $A_k := \int_{t=S_k}^{S_{k+1}} x(t) dt$  为帧  $k$  内的累计信息年龄, 根据图2.2, 他是一个平行四边形和一个三角形的面积之和, 即:

$$A_k = (D_{k-1} + W_{k-1})D_k + \frac{1}{2}(D_k + W_k)^2.$$

时间  $[0, S_{K+1})$  内的累计信息年龄可以被写成是序列  $A_k$  的和, 即

$$\mathbb{E} \left[ \int_{t=0}^{S_{K+1}} x(t) dt \right] = \mathbb{E} \left[ \sum_{k=1}^K A_k \right] = \mathbb{E} \left[ \sum_{k=1}^K q(D_{k-1}, W_{k-1}, D_k, W_k) \right], \quad (2.4)$$

其中函数

$$q(d', w', d, w) := (d' + w')d + \frac{1}{2}(d + w)^2.$$

在采样频率限制下, 设计最小化平均信息年龄年龄的采样策略  $\pi$  可以被建模为如

① 这一限制在  $W_{ub}$  取值较大时不会影响所设计策略的最优性。这是因为等待无穷长时间再进行下一次采样将会导致一个无穷大的信息年龄, 对于时效性优化不利。

② 另一种时间平均信息年龄的定义为: 时间窗  $[0, T)$  内的累计信息年龄除以时间窗长度  $T$ , 即  $\limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[ \int_0^T X(t) dt \right]$ 。这两种定义均合理。特别地, 当策略  $\pi$  为时不变策略, 使用策略  $\pi$  所得到的马尔科夫链  $\{(D_k, W_k)\}$  仅有一个遍历类时, 文献<sup>[84]</sup>证明了这两种定义等价。

下数学问题:

**问题 2.1:**

$$\pi^* \triangleq \arg \min_{\pi \in \Pi} \limsup_{K \rightarrow \infty} \frac{\mathbb{E} \left[ \sum_{k=1}^K q(D_{k-1}, W_{k-1}, D_k, W_k) \right]}{\mathbb{E} \left[ \sum_{k=1}^K (D_k + W_k) \right]}, \quad (2.5a)$$

$$\text{s.t. } \liminf_{K \rightarrow \infty} \frac{1}{K} \mathbb{E} \left[ \sum_{k=1}^K (D_k + W_k) \right] \geq \frac{1}{f_{\max}}, \quad (2.5b)$$

其中  $f_{\max}$  为系统所允许的最大采样频率。我们假设  $f_{\max}$  满足如下条件:

**假设 2.2:** 问题2.1有严格可行解, 即存在  $\epsilon > 0$  和时不变策略  $\pi_\epsilon \in \Pi$ , 使得采样频率约束可以严格满足, 即:

$$\mathbb{E} [D + \pi_\epsilon(D)] \geq \frac{1}{f_{\max}} + \epsilon. \quad (2.6)$$

我们注意到最优策略  $\pi^*$  被证明具有一个阈值结构<sup>[16]</sup>, 当传输时延分布  $\mathbb{P}_D$  的闭式表达式已知时, 最优阈值可以通过二分搜索得到。本问题中, 我们假设传输时延分布  $\mathbb{P}_D$  未知。采样策略设计仅能依据平均传输时延和二阶矩的上下界  $\bar{D}_{lb}, \bar{D}_{ub}, M_{lb}, M_{ub}$  以及历史传输时延  $\{D_k\}$ 。

## 2.3 问题求解

本节中, 我们首先将回顾研究传输时延分布  $\mathbb{P}_D$  已知时, 信息年龄最小采样策略  $\pi^*$  的结构, 从而将信息年龄最小化问题2.1转化为更新过程优化问题。节2.3.2将提出一种适用于时延分布函数  $\mathbb{P}_D$  未知的场景, 在线学习和设计最优策略  $\pi^*$  的采样算法, 节2.3.3将对所提算法进行理论分析。

### 2.3.1 更新过程优化问题

**定义 2.1:** 策略  $\pi$  为时不变的确定性策略 (stationary deterministic), 如果它所选取的等待时间  $W_k$  可以写成传输时延  $D_k$  的一个确定性函数  $w : [0, \infty) \mapsto [0, W_{ub}]$ , 即第  $(k+1)$  次采样前的等待时间  $W_k = w(D_k)$ 。

记  $\Pi_{SD} \subseteq \Pi$  为所有时不变确定性策略集合:

$$\Pi_{SD} \triangleq \{\pi \in \Pi : W_k = w(D_k), \forall k\}.$$

当传输时延分布  $\mathbb{P}_D$  已知且绝对连续时, 根据文献<sup>[16]</sup>, 信息年龄最小的采样策略  $\pi^*$  具有如下性质:

**定理 2.1:** <sup>[16]</sup>定理2 存在一个时不变确定性策略  $\pi^* \in \Pi_{SD}$ , 它可以最小化问题2.1中的平均信息年龄。

接下来我们记  $\pi(d)$  为时不变确定性  $\pi$  观测到传输时延  $d$  后选择的采样等待时间。根据定理2.1，存在一个确定性策略  $\pi^*$  可以最小化问题2.1中的平均信息年龄。接下来，记  $L^2$  为 Lebesgue 空间有界函数族，我们将最优时不变确定性策略  $\pi^*$  的搜索问题转化为如下问题2.2中的更新过程优化：

**问题 2.2 (更新过程优化):**

$$\pi^* = \arg \min_{\pi \in L_2} \frac{\mathbb{E}\left[\frac{1}{2}(D + \pi(D))^2\right]}{\mathbb{E}[D + \pi(D)]} + \bar{D}, \quad (2.7a)$$

$$\text{s.t. } \mathbb{E}[D + \pi(D)] \geq \frac{1}{f_{\max}}. \quad (2.7b)$$

具体推导详见附录2.6.1。问题2.2可以被看作是更新过程的优化，这是因为：

- 在每一帧内观测到的传输时延  $D_k$  以概率  $\mathbb{P}_D$  独立同分布。
- 记帧  $k$  的长度  $L_k := D_k + \pi(D_k)$ 。由于每次传输时延  $D_k$  独立同分布， $\pi(\cdot)$  为一个确定性函数，每一帧的长度  $L_k$  也为独立同分布的随机变量。
- 我们将  $Q_k := \frac{1}{2}(D_k + \pi(D_k))^2$  认为是帧  $k$  内所得到的“奖励”。同样地，根据传输时延  $D_k$  独立同分布、 $\pi$  为确定性函数这两条性质，帧  $k$  内得到的奖励  $Q_k$  也是独立同分布的随机变量。

以上分析说明每一帧  $k$  的帧长和奖励  $(L_k, Q_k)$  都是独立同分布的，同时，由于不等式  $\mathbb{E}[L_k] \leq \mathbb{E}[D + W_{\text{ub}}] < \infty$  和  $\mathbb{E}[Q_k] \leq \mathbb{E}\left[\frac{1}{2}(D + W_{\text{ub}})^2\right] < \infty$ ，帧  $k$  的长度和奖励的期望都有界。最优采样策略  $\pi^*$  的设计可以方便地通过更新过程求解。

### 2.3.2 一种传输时延分布 $\mathbb{P}_D$ 未知时的在线优化算法

本节中我们将首先讨论传输时延分布  $\mathbb{P}_D$  已知时，如何用更新过程优化框架设计最优的时不变确定性采样策略  $\pi^*$ 。接下来基于  $\pi^*$  的设计，针对  $\mathbb{P}_D$  未知的场景，设计一种在线采样算法  $\pi_{\text{online}}$ 。为了简化后续分析，记策略族  $\Pi_{\text{cons}}$  为满足采样频率限制的  $f_{\max}$  所有时不变确定性策略集合，即：

$$\Pi_{\text{cons}} \triangleq \{\pi | \pi \in \Pi_{\text{SD}}, \mathbb{E}_{\pi}[D + \pi(D)] \geq \frac{1}{f_{\max}}\}$$

#### 2.3.2.1 传输时延分布 $\mathbb{P}_D$ 已知时最优策略 $\pi^*$ 的设计

记  $\bar{X}^* := J(\pi^*)$  为使用最优策略  $\pi^*$  得到的平均信息年龄，那么对于满足采样频率约束的任意策略  $\pi \in \Pi_{\text{cons}}$ ，我们有以下不等式：

$$\frac{\mathbb{E}\left[\frac{1}{2}(D + \pi(D))^2\right]}{\mathbb{E}[D + \pi(D)]} + \bar{D} \geq \bar{X}^*, \forall \pi \in \Pi_{\text{cons}}. \quad (2.8)$$

在不等式 (2.8) 的两边同时减去  $\bar{D}$  并乘以  $\mathbb{E}[D + \pi(D)]$ ，记  $\gamma^* := \bar{X}^* - \bar{D}$ ，

即可得到如下不等式：

$$\frac{1}{2}\mathbb{E}[(D + \pi(D))^2] - \gamma^* \mathbb{E}[D + \pi(D)] \geq 0, \forall \pi \in \Pi_{\text{cons}}. \quad (2.9)$$

注意到式(2.9)中的等号成立当且仅当策略  $\pi$  就是信息年龄最优策略  $\pi^*$ 。因此，如果参数  $\gamma^*$  已知，可以通过求解以下函数优化问题得到信息年龄最优的采样策略  $\pi^*$ ：

**问题 2.3 (函数优化问题):**

$$\pi^* = \arg \min_{\pi \in L^2} \mathbb{E} \left[ \frac{1}{2}(D + \pi(D))^2 - \gamma^*(D + \pi(D)) \right], \quad (2.10a)$$

$$\text{s.t. } \mathbb{E}[D + \pi(D)] \geq \frac{1}{f_{\max}}. \quad (2.10b)$$

接下来，我们通过拉格朗日乘子法，将约束条件式(2.10b)通过拉格朗日乘子  $v \geq 0$  加入到目标函数式(2.10a)中，并构建如下拉格朗日函数：

$$\mathcal{L}(\gamma, v, \pi) := \mathbb{E} \left[ \frac{1}{2}(D + \pi(D))^2 - (\gamma + v)(D + \pi(D)) \right] + v \frac{1}{f_{\max}}. \quad (2.11)$$

记  $v^* := \arg \max_{v \geq 0} \min_{\pi \in \Pi_{\text{SD}}} \mathcal{L}(\gamma^*, v, \pi)$  为优化问题在  $\gamma = \gamma^*$  的对偶乘子。根据<sup>[16]</sup>定理<sup>4</sup>，在观测到传输时延  $d$  后，最优策略  $\pi^*$  选取等待时间的准则为：

$$\pi^*(d) = (\gamma^* + v^* - d)^+. \quad (2.12)$$

记  $\tau^* := \gamma^* + v^*$  为最优阈值。当传输时延分布  $\mathbb{P}_D$  已知，最优阈值  $\tau^*$  可以通过二分法计算。得到最优策略  $\pi^*$  后，即可进一步计算最优参数  $\gamma^*$ ，他的取值满足以下推论：

**推论 2.1：** 最优参数  $\gamma^*$  的取值范围满足：

$$\gamma_{lb} \leq \gamma^* \leq \gamma_{ub},$$

其中

$$\begin{aligned} \gamma_{lb} &= \frac{\frac{1}{2}M_{lb}}{\bar{D}_{ub} + W_{ub}}, \\ \gamma_{ub} &= \frac{\frac{1}{2}M_{ub} + \bar{D}_{ub} \frac{1}{f_{\max}} + \frac{1}{2} \frac{1}{f_{\max}^2}}{\bar{D}_{lb} + \frac{1}{f_{\max}}}. \end{aligned} \quad (2.13)$$

### 2.3.2.2 基于 Monro-Robbins 算法的自适应在线采样算法 $\pi_{\text{online}}$

当传输时延统计信息  $\mathbb{P}_D$  未知时，我们无法通过二分法求解最优阈值  $\tau^*$  并得到  $\gamma^*$ 。在这样的场景中，我们通过 Monro-Robbins 算法<sup>[72]</sup>维护一个序列  $\{\gamma_k\}$ ，每个元素  $\gamma_k$  代表着在第  $k$  帧我们对于最优参数  $\gamma^*$  的估计。同时为了满足系统采样

频率限制，我们用序列  $U_k := k \frac{1}{f_{\max}} - S_k$  记录截至时间  $S_k$  的采样间隔违背。初始点  $\gamma_0$  随机从区间  $[\gamma_{lb}, \gamma_{ub}]$  中抽取，即  $\gamma_1 \sim \text{Uni}([\gamma_{lb}, \gamma_{ub}])$ 。接下来的每一帧  $k$  中，等待时间  $W_k$  的选取和参数  $\gamma_k$  的更新如下：

- 在帧  $k$  开始后，我们将收到第  $k$  个数据包成功传输后返回的 ACK 并计算传输时延  $D_k$ 。根据当前的估计  $\gamma_k$  和采样频率  $U_k$ ，我们选取等待时间

$$W_k = \left( \gamma_k + \frac{1}{V} U_k^+ - D_k \right)^+. \quad (2.14a)$$

我们将在等待  $W_k$  后，在时刻  $S_{k+1} = S_k + D_k + W_k$  采样并发送第  $k+1$  个数据包，帧  $k$  也将在时刻  $S_{k+1}$  结束。可以计算出帧  $k$  的时长  $L_k = D_k + W_k$  和奖励  $Q_k = \frac{1}{2}(D_k + W_k)^2$  用于之后计算和更新参数  $\gamma_k$ 。

- 接下来我们依据 Monro-Robbins 算法<sup>[72]</sup> 更新我们对于系数  $\gamma^*$  的猜测：

$$\gamma_{k+1} = [\gamma_k + \eta_k (Q_k - \gamma_k L_k)]_{\gamma_{lb}}^{\gamma_{ub}}, \quad (2.14b)$$

其中  $[\gamma]_a^b = \min\{b, \max\{\gamma, a\}\}$ ， $\{\eta_k\}$  为一组逐渐缩小的步长，他的取值形式为：

$$\eta_k = \begin{cases} \frac{1}{2D_{lb}}, & k = 1; \\ \frac{1}{(k+2)D_{lb}}, & k \geq 2. \end{cases} \quad (2.14c)$$

- 接下来我们更新采样频率违背序列  $U_k$ ，他的更新函数如下：

$$U_{k+1} = U_k + \left( \frac{1}{f_{\max}} - L_k \right). \quad (2.14d)$$

### 2.3.3 算法理论分析

我们注意到使用所提在线采样策略  $\pi_{\text{online}}$  得到的平均信息年龄与  $\bar{X}^*$  之间的差距与时间  $t$  演进较难分析。作为替代，与文献<sup>[74]</sup>相同，我们定义截止到第  $K$  次采样时累计信息年龄的期望与时间  $S_{K+1}$  期望的比值为：

$$\tilde{X}_K \triangleq \frac{\mathbb{E} \left[ \int_{t=0}^{S_{K+1}} x(t) dt \right]}{\mathbb{E}[S_{K+1}]}.$$

同时，记策略  $\pi_K$  为帧  $K$  内的等待时间选取策略，即函数  $\pi_K(d) = (\gamma_K + \frac{1}{V} U_K^+ - d)^+$ 。我们将从在线算法比值  $\tilde{X}_K$  与最优信息年龄  $\bar{X}^*$  之间的差距  $\tilde{X}_K - \bar{X}^*$ ，以及策略  $\pi_K$  与最优策略  $\pi^*$  得到的平均信息年龄之差两个角度分析所提在线算法的收敛速度。算法主要性能如下：

**定理 2.2：** 当问题2.2没有采样频率的限制，即  $f_{\max} = \infty$ ，如果传输时延有上界  $B$ ，

即  $D < B < \infty$ , 所提在线算法得到的平均信息年龄和采样阈值满足以下性质:

$$\lim_{K \rightarrow \infty} \frac{\int_{t=0}^{S_{K+1}} x(t) dt}{S_{K+1}} \stackrel{\text{a.s.}}{=} \bar{A}_{\pi^*}, \quad \lim_{K \rightarrow \infty} \gamma_K \stackrel{\text{a.s.}}{=} \gamma^*. \quad (2.15a)$$

在线算法  $\pi_{\text{online}}$  得到的信息年龄比值  $\tilde{X}_K$  满足如下条件:

$$\tilde{X}_K - \bar{X}^* \leq \frac{L_{\text{ub}}^4}{DD_{\text{lb}}^2} \frac{(1 + \ln(K+1))}{K} = \mathcal{O}\left(\frac{\ln K}{K}\right), \quad (2.15b)$$

其中  $L_{\text{ub}} = B + \gamma_{\text{ub}}$ 。记  $J(\pi)$  为使用策略  $\pi$  得到的平均信息年龄, 那么使用策略  $\pi_K$  与使用最优策略  $\pi^*$  得到的平均信息年龄之间的差有如下上界:

$$J(\pi_K) - J(\pi^*) \leq \frac{L_{\text{ub}}^4}{DD_{\text{lb}}^2} \frac{1}{K} = \mathcal{O}\left(\frac{1}{K}\right). \quad (2.15c)$$

几乎处处收敛性证明与<sup>[74]</sup>[定理 7、定理 8] 类似, 在此不再展开。收敛速度证明详见附录2.6.3。

**注 2.1:** 以上结果说明, 当系统没有采样频率约束时, 在线算法  $\pi_{\text{online}}$  将最终学习到使得平均信息年龄最小的采样策略  $\pi^*$ 。

**定理 2.3:** 所提算法  $\pi_{\text{online}}$  可以满足系统采样频率要求, 即

$$\liminf_{K \rightarrow \infty} \mathbb{E} \left[ \frac{1}{K} \sum_{k=1}^K (W_k + D_k) \right] \geq \frac{1}{f_{\max}}. \quad (2.16)$$

证明详见附录2.6.7。

## 2.4 仿真结果

本节通过仿真评估所提出的自适应采样策略的平均信息年龄和学习代价。我们考虑传输时延服从参数为  $(\mu, \sigma)$  的对数正态分布 (log-normal distribution), 即概率测度  $\mathbb{P}_D$  的密度函数  $p(d)$  服从如下表达式:

$$p(d) := \frac{\mathbb{P}_D(dd)}{dd} = \frac{1}{d\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln d - \mu)^2}{2\sigma^2}\right).$$

### 2.4.1 无采样频率约束下的仿真结果

我们首先研究在点对点链路没有采样频率约束 ( $f_{\max} = \infty$ ) 所提算法的自适应。我们仿真并比较了以下三种策略: (1) 零等待策略  $\pi_{\text{zw}}$ , 即一旦收到上一个数据包被成功接收的 ACK 后, 立刻采样并发送一个新的数据包, 该策略的等待时间确定规则为:  $\pi_{\text{zw}}(d) = 0, \forall d$ ; (2) 当传输时延统计信息已知时, 文献<sup>[16]</sup>所提出信息年龄最小的采样策略  $\pi^*$ ; (3) 节2.3.2.2所提的在线采样算法  $\pi_{\text{online}}$ 。仿真过程中,

我们考虑传输时延服从参数  $\mu = 1, \sigma = 1.3$  的对数正态分布。图2.3画出了信息年龄比值  $\tilde{X}_k = \frac{\mathbb{E}\left[\int_0^{S_{k+1}} x(t)dt\right]}{\mathbb{E}[S_{k+1}]}$  随着帧  $k$  的演进关系曲线。图2.4画出了截至  $t$  时，时间平均信息年龄  $\bar{X}_{\pi,t} = \frac{1}{t} \int_{t'=1}^t x(t')dt'$  的平均值以及置信区间。我们通过对 100 次仿真求平均完成了期望的计算。如图2.3所示，通过在线采样算法得到的信息年龄比值  $\tilde{X}_k$  随着采样次数  $k$  增多逐渐收敛到统计信息已知时的最小信息年龄  $\bar{X}^* = J(\pi^*)$ ，这与我们在定理2.2中的理论分析相符。同时，我们发现使用在线策略  $\pi_{\text{online}}$  时，时间平均信息年龄  $\bar{X}_{\pi,t}$  与统计信息已知时最小信息年龄  $\bar{X}^*$  之间的差距随着  $t$  的增大逐渐减小并消失。

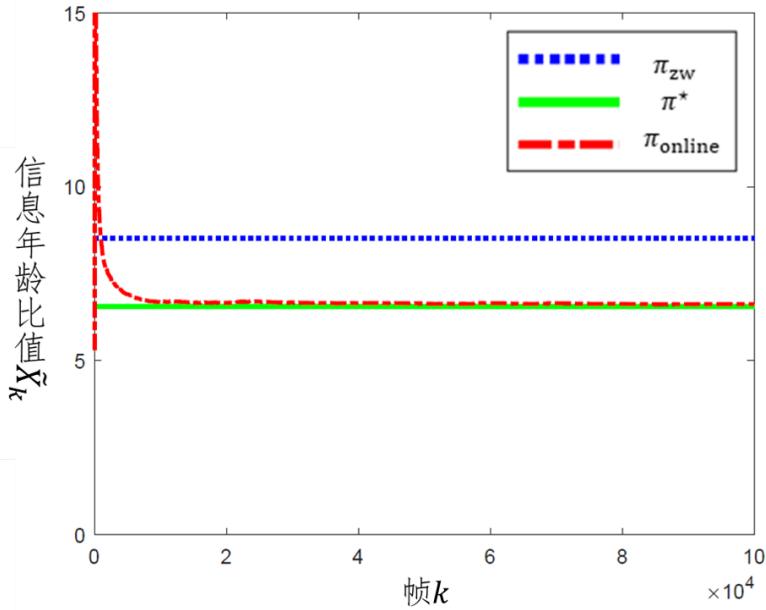
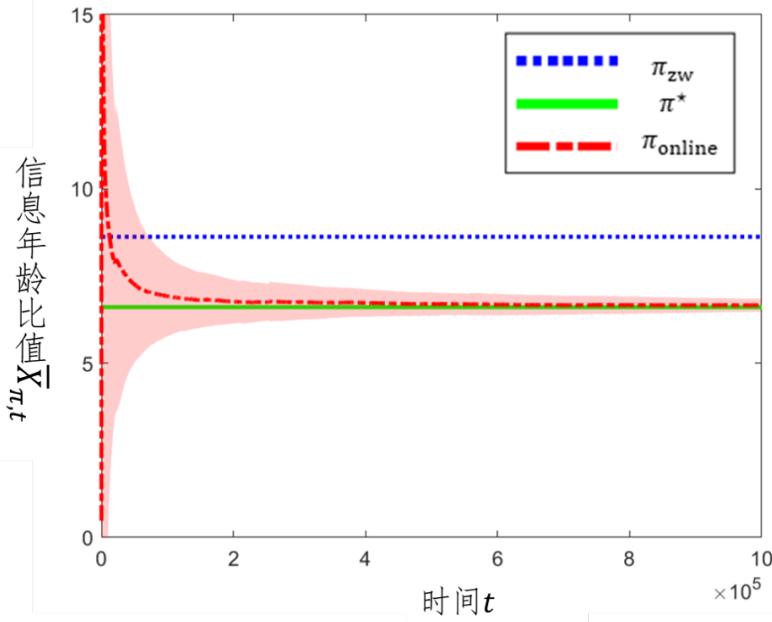


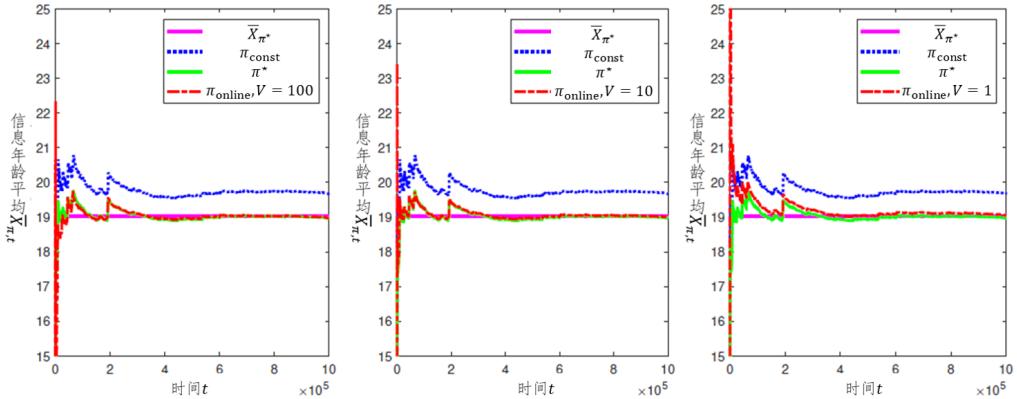
图 2.3 信息年龄比值  $\tilde{X}_k$  随着采样数目  $k$  的演进关系图

#### 2.4.2 有采样频率限制下的仿真结果

接下来我们研究所提算法在有采样频率约束下的理论性能。考虑到采样频率的限制，我们仿真了并比较以下三种算法：(1) 满足采样频率限制的常数等待策略  $\pi_{\text{const}}$ ，即每一次收到 ACK 后的等待时间为常数  $\pi_{\text{const}}(d) \equiv \frac{1}{f_{\max}} - \bar{D}, \forall d$ ；(2) 当传输时延统计信息已知时，文献<sup>[16]</sup>所提出的信息年龄最小采样策略  $\pi^*$ ；(3) 节2.3.2.2所提在线采样策略。仿真中传输时延服从参数  $\mu = 1, \sigma = 1.5$  的对数正态分布，其中采样频率的限制为  $f_{\max} = \frac{1}{10\bar{D}}$ 。对于某一条传输时延固定的采样路径，使用以上三种策略得到的时间平均信息年龄、平均采样间隔如图2.5和图2.6所示。图2.5说明随着时间  $t$  增长，所提在线学习策略得到的时间平均信息年龄逐渐收敛于最优策略  $\pi^*$  得到的平均信息年龄。同时随着  $t$  增加，使用不同数值  $V$  得到的

图 2.4 平均信息年龄  $\bar{X}_{\pi,t}$  随着时间  $t$  的演进关系图。

平均采样间隔都接近于  $\frac{1}{f_{\max}}$ ，这验证了定理2.3中所提算法满足采样频率限制这一结论。通过对比三张子图，我们发现当  $V$  越小，对采样频率限制的服从条件越好，平均信息年龄的演进也更加平稳。



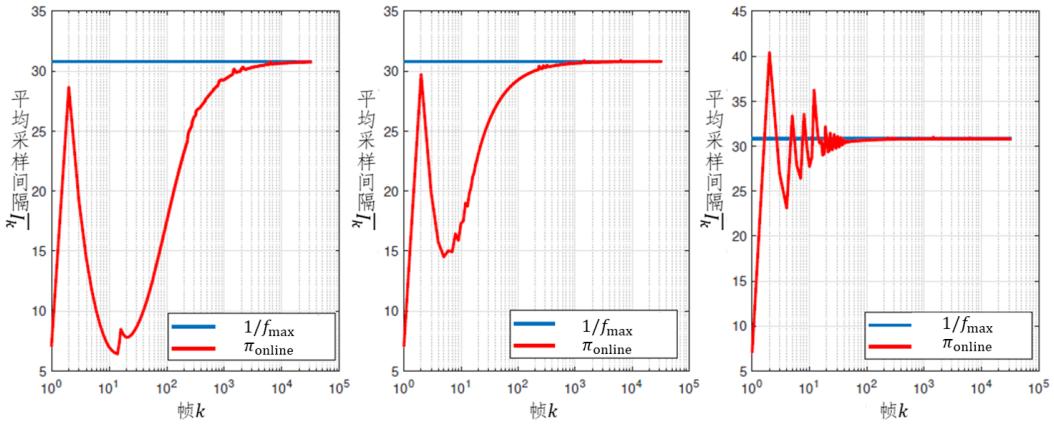
参数  $V$  的取值从左到右分别为 100, 10, 1。

图 2.5 有采样频率约束下，针对某一条采样路径的时间平均信息年龄演进

## 2.5 本章小结

本章研究了点对点状态更新系统中，传输时延统计信息未知时，最小化接收端平均信息年龄的采样策略。

- 我们将信息年龄最小化问题转化为更新过程优化问题，通过 Monro-Robbins



参数  $V$  的取值从左到右分别为 100, 10, 1。

图 2.6 使用不同  $V$  时在线采样策略得到的平均采样间隔

算法设计了一种“边等待，边学习”的在线采样策略。

- 理论分析证明，随着时间  $t$  和采样次数  $k$  趋于无穷，所提在线采样策略可以学习到统计信息已知时信息年龄最小的采样策略。
- 仿真结果验证了所提算法的渐近最优性。通过讨论仿真中关键参数对于信息年龄和采样间隔违背的影响，为算法在实际系统中的部署提供了指导。

## 2.6 附录

### 2.6.1 问题2.2推导

由于概率测度  $\mathbb{P}_D$  绝对连续，等待策略  $\pi(d) \in [0, W_{ub}]$  所选取的等待时间有上界  $W_{ub}$ ，根据<sup>[85]</sup>，我们得到  $\int_0^\infty |\pi(d)|^2 \mathbb{P}_D(dd) \leq \int_0^\infty W_{ub}^2 \mathbb{P}_D(dd) \leq W_{ub}^2$ ，因此函数  $\pi(d)$  属于 Lebesgue 空间的函数族  $L^2$ 。

接下来我们计算使用任意时不变确定性策略  $\pi$  得到的平均信息年龄  $J(\pi)$ 。根据式 (2.3)， $J(\pi)$  计算如下：

$$\begin{aligned} J(\pi) &= \limsup_{K \rightarrow \infty} \frac{\mathbb{E} \left[ \sum_{k=1}^K q(D_{k-1}, \pi(D_{k-1}), D_k, \pi(D_k)) \right]}{\mathbb{E} \left[ \sum_{k=1}^K (D_k + \pi(D_k)) \right]} \\ &= \limsup_{K \rightarrow \infty} \frac{\sum_{k=1}^K \mathbb{E}[(D_{k-1} + \pi(D_{k-1}))D_k]}{\sum_{k=1}^K \mathbb{E}[D_k + \pi(D_k)]} + \limsup_{K \rightarrow \infty} \frac{\sum_{k=1}^K \mathbb{E} \left[ \frac{1}{2}(D_k + \pi(D_k))^2 \right]}{\sum_{k=1}^K \mathbb{E}[D_k + \pi(D_k)]} \\ &\stackrel{(a)}{=} \limsup_{K \rightarrow \infty} \frac{\sum_{k=1}^K \mathbb{E}[(D_{k-1} + \pi(D_{k-1}))] \overline{D}}{\sum_{k=1}^K \mathbb{E}[D_k + \pi(D_k)]} + \limsup_{K \rightarrow \infty} \frac{\sum_{k=1}^K \mathbb{E} \left[ \frac{1}{2}(D_k + \pi(D_k))^2 \right]}{\sum_{k=1}^K \mathbb{E}[D_k + \pi(D_k)]} \end{aligned}$$

$$\stackrel{(b)}{=} \limsup_{K \rightarrow \infty} \frac{\sum_{k=1}^K \mathbb{E} \left[ \frac{1}{2} (D_k + \pi(D_k))^2 \right]}{\sum_{k=1}^K \mathbb{E}[D_k + \pi(D_k)]} + \bar{D}, \quad (2.17)$$

其中等式(a)是因为：传输时延  $D_k$  独立同分布、策略  $\pi$  为时不变策略，因此第  $k$  个数据包的传输时延  $D_k$  和第  $(k-1)$  个数据包的传输时延  $D_{k-1}$ 、等待时间  $\pi(D_{k-1})$  均独立，我们得到  $\mathbb{E}[D_{k-1} + \pi(D_{k-1})D_k] = \mathbb{E}[D_{k-1}] + \mathbb{E}[\pi(D_{k-1})]\bar{D}$ ；等式(b)成立是因为传输时延  $D_k$  独立同分布，因此  $\limsup_{K \rightarrow \infty} \frac{\sum_{k=1}^K \mathbb{E}[(D_{k-1} + \pi(D_{k-1}))]}{\sum_{k=1}^K \mathbb{E}[D_k + \pi(D_k)]} = 1$ 。

## 2.6.2 推论2.1证明

证明：最优系数  $\gamma^*$  的下界由如下式子得到：

$$\gamma^* = \frac{\mathbb{E} \left[ \frac{1}{2} (D + \pi^*(D))^2 \right]}{\mathbb{E}[D + \pi^*(D)]} \geq \frac{\mathbb{E} \left[ \frac{1}{2} (D + \pi^*(D))^2 \right]}{\mathbb{E}[D + W_{ub}]} = \frac{\frac{1}{2} M_{lb}}{\bar{D}_{ub} + W_{ub}}. \quad (2.18)$$

接下来我们推导最优系数  $\gamma^*$  的上界。我们考虑常数等待策略  $\pi_{const}$ ，即帧  $k$  中的等待时间  $W_k \equiv \frac{1}{f_{max}}$  为一个与时延  $D_k$  无关的常数。根据式(2.7a)，使用策略  $\pi_{const}$  得到的平均信息年龄上界如下：

$$\begin{aligned} J(\pi_{const}) &= \frac{\mathbb{E} \left[ \frac{1}{2} (D + \pi_{const}(D))^2 \right]}{\mathbb{E}[D + \pi_{const}(D)]} + \bar{D} \\ &\leq \frac{\frac{1}{2} M_{ub} + \bar{D} \frac{1}{f_{max}} + \frac{1}{2} \frac{1}{f_{max}^2}}{\bar{D}_{lb} + \frac{1}{f_{max}}} + \bar{D}. \end{aligned} \quad (2.19)$$

注意到策略  $\pi_{const}$  得到的平均信息年龄可能不是最优的，也就是说  $J(\pi^*) \leq J(\pi_{const})$ ，而最优系数  $\gamma^*$  满足  $\gamma^* = J(\pi^*) - \bar{D}$ 。根据不等式(2.19)，即得到最优系数  $\gamma^*$  的上界：

$$\gamma^* \leq J(\pi_{const}) - \bar{D} \leq \frac{\frac{1}{2} M_{ub} + \bar{D}_{ub} \frac{1}{f_{max}} + \frac{1}{2} \frac{1}{f_{max}^2}}{\bar{D}_{lb} + \frac{1}{f_{max}}} =: \gamma_{ub}. \quad (2.20)$$

■

## 2.6.3 定理2.2证明

证明：我们注意到任意帧  $k$  中使用的系数  $\gamma_k$  均有上界，即  $\gamma_k \leq \gamma_{ub}$ ，同时传输时延也有上界  $D \leq B$ 。因此帧长和每一帧内的奖励均有上界，即：

$$\begin{aligned} L &\leq D + (\gamma - D)^+ \leq B + \gamma_{lb} =: L_{ub}, \\ Q &= \frac{1}{2} L^2 \leq L_{ub}^2. \end{aligned} \quad (2.21)$$

记使用信息年龄最优采样策略  $\pi^*$  时的平均帧长  $\bar{L}^* := \mathbb{E}[D + \pi^*(D)]$ , 每一帧平均奖励  $\bar{Q}^* := \mathbb{E}[\frac{1}{2}(D + \pi^*(D))^2]$ 。接下来我们将分析帧  $k$  的平均长度  $L_k$ 、奖励  $Q_k$  与使用最优策略得到的  $\bar{L}^*$ 、 $\bar{Q}^*$  之间的关系。我们将首先给出如下引理:

**引理 2.1:** 帧  $k$  长度的条件期望  $\mathbb{E}[L_k | \mathcal{F}_{k-1}]$  和奖励  $\mathbb{E}[Q_k | \mathcal{F}_{k-1}]$  满足以下不等式:

$$\mathbb{E}[Q_k - \gamma_k L_k | \mathcal{F}_{k-1}] \leq (\gamma^* - \gamma_k) \bar{L}^*, \quad (2.22a)$$

$$\mathbb{E}[Q_k - \gamma^* L_k | \mathcal{F}_{k-1}] \leq -(\gamma^* - \gamma_k) (\mathbb{E}[L_k | \mathcal{F}_{k-1}] - \bar{L}^*). \quad (2.22b)$$

**引理 2.2:** 注意到前  $K$  帧累计信息年龄可以通过式(2.4)计算, 即  $\mathbb{E}\left[\int_0^{S_{K+1}} x(t) dt\right] = \mathbb{E}\left[\sum_{k=1}^K A_k\right]$ , 其中第  $k$  帧内累计信息年龄  $A_k = Q_k + D_k L_{k-1}$ 。截至第  $K$  帧结束, 前  $K$  帧累计信息年龄之间期望满足如下不等式:

$$\mathbb{E}\left[\sum_{k=1}^K ((Q_k + D_k L_{k-1}) - (\gamma^* + \bar{D}) L_k)\right] \leq \mathbb{E}\left[\sum_{k=1}^K (\gamma^* - \gamma_k)^2\right]. \quad (2.23)$$

引理2.1和2.2在附录2.6.4和2.6.5中给出。以上引理即给出截至到第  $K$  帧, 比值  $\tilde{X}_K$  与最优信息年龄  $\bar{X}^*$  之间差别的上界:

$$\begin{aligned} & \tilde{X}_K - \bar{X}^* \\ &= \frac{\mathbb{E}\left[\sum_{k=1}^K (Q_k + L_{k-1} D_k)\right]}{\mathbb{E}\left[\sum_{k=1}^K L_k\right]} - (\gamma^* + \bar{D}) \\ &= \frac{\mathbb{E}\left[\sum_{k=1}^K (Q_k + L_{k-1} D_k - (\gamma^* + \bar{D}) L_k)\right]}{\mathbb{E}\left[\sum_{k=1}^K L_k\right]} \\ &\stackrel{(a)}{\leq} \frac{\mathbb{E}\left[\sum_{k=1}^K (\gamma^* - \gamma_k)^2\right]}{\mathbb{E}\left[\sum_{k=1}^K L_k\right]}. \end{aligned} \quad (2.24)$$

其中, 不等式 (a) 可以由引理2.2直接得到。由于帧长  $L_k$  有下界, 为了得到  $\tilde{X}_K$  与  $\bar{X}^*$  之间的差, 我们只需要进一步确定系数  $\gamma_k$  与  $\gamma^*$  累计误差的期望  $\mathbb{E}\left[\sum_{k=1}^K (\gamma^* - \gamma_k)^2\right]$ 。下一个引理将帮助我们确定帧  $k$  内系数  $\gamma_k$  与最优系数  $\gamma^*$  之间的均方误差期望的上界, 它的证明将在附录2.6.6中给出:

**引理 2.3:** 帧  $k$  内系数  $\gamma_k$  与最优系数  $\gamma^*$  之间的误差期望有如下上界:

$$\mathbb{E}[(\gamma_k - \gamma^*)^2] \leq \frac{1}{k} \frac{L_{\text{ub}}^4}{D_{\text{lb}}^2}. \quad (2.25)$$

引理2.1-2.3将帮助我们证明定理2.2中的两个结论:

### 2.6.3.1 式(2.15b)的证明

将引理2.3中的不等式(2.3)从帧  $k=1$  加到  $K$  可得:

$$\begin{aligned}
 & \mathbb{E} \left[ \sum_{k=1}^K (\gamma^* - \gamma_k)^2 \right] \\
 & \leq \frac{L_{\text{ub}}^4}{D_{\text{lb}}^2} \left( \sum_{k=1}^K \frac{1}{k} \right) \\
 & \stackrel{(b)}{\leq} \frac{L_{\text{ub}}^4}{D_{\text{lb}}^2} \left( 1 + \int_{k=1}^{K+1} \frac{1}{k} dk \right) \\
 & \stackrel{(c)}{=} \frac{L_{\text{ub}}^4}{D_{\text{lb}}^2} (1 + \ln(K+1)), \tag{2.26}
 \end{aligned}$$

其中不等式(b)是因为  $\frac{1}{k} \leq \int_{k'=k}^{k+1} \frac{1}{k'} dk'$ , 等式(c)是因为  $\int_a^b \frac{1}{x} dx = \ln b - \ln a$ 。将不等式(2.26)带入到式(2.24)中, 我们有:

$$\begin{aligned}
 \tilde{X}_K - \bar{X}^* &= \frac{\mathbb{E} \left[ \sum_{k=1}^K (\gamma^* - \gamma_k)^2 \right]}{\mathbb{E} \left[ \sum_{k=1}^K L_k \right]} \\
 &\leq \frac{L_{\text{ub}}^4}{DD_{\text{lb}}^2} \frac{(1 + \ln(K+1))}{K}. \tag{2.27}
 \end{aligned}$$

### 2.6.3.2 式(2.15c)的证明

注意到使用系数为  $\gamma_K$  的时不变确定性策略  $\pi_K$  时, 平均信息计算如下:

$$J(\pi_K) = \frac{\mathbb{E} \left[ \frac{1}{2}((\gamma_K - D)^+ + D)^2 \right]}{\mathbb{E}[(\gamma_K - D)^+ + D]} + \bar{D}.$$

注意到  $\pi^*$  为最小化平均信息年龄的最佳策略, 因此其他任意时不变确定性策略  $\pi_k$  得到的平均信息年龄都满足:  $J(\pi_K) \geq J(\pi^*)$ 。策略  $\pi_K$  得到的平均信息年龄距离最优信息年龄  $J(\pi^*)$  之间的差满足以下不等式:

$$\begin{aligned}
 & J(\pi_K) - J(\pi^*) \\
 &= \frac{\mathbb{E} \left[ \frac{1}{2}((\gamma_K - D)^+ + D)^2 \right]}{\mathbb{E}[(\gamma_K - D)^+ + D]} - \gamma^* \\
 &= \frac{\mathbb{E} \left[ \frac{1}{2}((\gamma_K - D)^+ + D)^2 - \gamma_K((\gamma_K - D)^+ + D) \right]}{\mathbb{E}[(\gamma_K - D)^+ + D]} + (\gamma_K - \gamma^*)
 \end{aligned}$$

$$\begin{aligned}
&\stackrel{(d)}{=} \frac{\mathbb{E}[Q_K - \gamma_K L_K]}{\mathbb{E}[L_K]} + (\gamma_K - \gamma^*) \\
&\stackrel{(e)}{\leq} \frac{(\gamma^* - \gamma_K) \bar{L}^*}{\mathbb{E}[(\gamma_K - D)^+ + D]} + (\gamma_K - \gamma^*) \\
&= (\gamma_K - \gamma^*) \left( \frac{\mathbb{E}[(\gamma_K - D)^+ + D] - \mathbb{E}[(\gamma^* - D)^+ + D]}{\mathbb{E}[(\gamma_K - D)^+ + D]} \right) \\
&= \frac{(\gamma_K - \gamma^*)}{\mathbb{E}[(\gamma_K - D)^+ + D]} \mathbb{E}[(\gamma_K - D)^+ - (\gamma^* - D)^+] \\
&\leq \frac{1}{D} (\gamma_K - \gamma^*)^2,
\end{aligned} \tag{2.28}$$

其中不等式 (d) 因为根据定义  $Q_k = \frac{1}{2} ((\gamma_k - D_k)^+ + D_k)^2$ ,  $L_k = (\gamma_k - D_k)^+ + D_k$ , 另外每一帧内的传输时延  $D_k$  独立同分布; 不等式 (e) 可以通过对引理2.1中的第一个结论  $\mathbb{E}[Q_K - \gamma_K L_K | \mathcal{F}_{K-1}] \leq (\bar{Q}^* - \gamma_K \bar{L}^*)$  两边取期望得到。

将系数  $\gamma_k$  的收敛情况不等式 (2.25) 带入到不等式 (2.28), 即可完成定理2.2中第二个结论的证明:

$$\mathbb{E}[J(\pi_K) - J(\pi^*)] \leq \frac{L_{ub}^4}{DD_{lb}^2} \frac{1}{K}.$$

■

#### 2.6.4 引理2.1证明

**证明:** 注意到在帧  $k$  中, 等待时间  $W_k$  的选取是用于最小化目标函数式 (2.10a), 因此我们有:

$$\begin{aligned}
\mathbb{E}[Q_k - \gamma_k L_k | \mathcal{F}_{k-1}] &\stackrel{(a)}{\leq} (\bar{Q}^* - \gamma_k \bar{L}^*) \\
&= (\bar{Q}^* - \gamma^* \bar{L}^*) + (\gamma^* - \gamma_k) \bar{L}^* \stackrel{(b)}{=} (\gamma^* - \gamma_k) \bar{L}^*,
\end{aligned} \tag{2.29}$$

其中不等式 (a) 是因为在  $\gamma = \gamma_k$  时, 帧  $k$  中所使用的策略  $\pi_k$  可以最小化目标函数 (2.10a), 但策略  $\pi^*$  并不一定可以优化拉格朗日函数 (2.10a); 等式 (b) 是因为最优系数  $\gamma^*$  满足  $\bar{Q}^* = \gamma^* \bar{L}^*$ 。以上推导可以证明引理2.1中的第一个结论。

接下来, 在不等式 (2.29) 的两边同时加上  $(\gamma_k - \gamma^*) \mathbb{E}[L_k | \mathcal{F}_{k-1}]$ , 我们将得到:

$$\mathbb{E}[Q_k - \gamma^* L_k | \mathcal{F}_{k-1}] \leq (\gamma_k - \gamma^*) \mathbb{E}[L_k - \bar{L}^* | \mathcal{F}_{k-1}]. \tag{2.30}$$

■

### 2.6.5 引理2.2证明

**证明：**为了得到累计误差  $\mathbb{E} \left[ \sum_{k=1}^K ((Q_k + L_{k-1} D_k) - (\gamma^* + \bar{D}) L_k) \right]$  的上界，首先我们在不等式 (2.22b) 的两端同时加上  $\mathbb{E}[L_{k-1} D_k | \mathcal{F}_{k-1}]$  并得到：

$$\mathbb{E}[(Q_k + L_{k-1} D_k) - \gamma^* L_k | \mathcal{F}_{k-1}] \leq -(\gamma^* - \gamma_k) \left( \mathbb{E}[L_k | \mathcal{F}_{k-1}] - \bar{L}^* \right) + \mathbb{E}[L_{k-1} D_k | \mathcal{F}_{k-1}].$$

进一步地，将以上不等式化简得到：

$$\begin{aligned} & \mathbb{E}[(Q_k + L_{k-1} D_k) - \gamma^* L_k | \mathcal{F}_{k-1}] \\ & \stackrel{(a)}{\leq} -(\gamma^* - \gamma_k) \left( \mathbb{E}[L_k | \mathcal{F}_{k-1}] - \bar{L}^* \right) + L_{k-1} \bar{D} \\ & \stackrel{(b)}{\leq} (\gamma^* - \gamma_k)^2 + L_{k-1} \bar{D}, \end{aligned} \quad (2.31)$$

其中不等式 (a) 是因为传输时延  $D_k$  与  $L_{k-1}$  独立所以  $\mathbb{E}[L_{k-1} D_k | \mathcal{F}_{k-1}] = L_{k-1} \bar{D}$ ；不等式 (b) 是因为：

$$\mathbb{E}[L_k - \bar{L}^* | \mathcal{F}_{k-1}] = \mathbb{E}[(\gamma_k - D)^+ - (\gamma^* - D)^+] \leq |\gamma_k - \gamma^*|. \quad (2.32)$$

将不等式 (2.31) 从帧  $k = 1$  加到帧  $K$  并对于  $\mathcal{F}_{K-1}$  取期望，我们得到：

$$\mathbb{E} \left[ \sum_{k=1}^K ((Q_k + L_{k-1} D_k) - \gamma^* L_{k-1}) \right] \leq \mathbb{E} \left[ \sum_{k=1}^K (\gamma^* - \gamma_k)^2 \right] - \mathbb{E} \left[ \sum_{k=1}^K L_k \right] \bar{D}. \quad (2.33)$$

对以上不等式两边同时减去  $\mathbb{E} \left[ \sum_{k=1}^K L_k \right] \bar{D} + \mathbb{E}[L_K] \gamma^*$ ，我们将得到：

$$\begin{aligned} \mathbb{E} \left[ \sum_{k=1}^K ((Q_k + L_{k-1} D_k) - (\gamma^* + \bar{D}) L_k) \right] & \leq \mathbb{E} \left[ \sum_{k=1}^K (\gamma^* - \gamma_k)^2 \right] - \gamma^* \mathbb{E}[L_K] \\ & \leq \mathbb{E} \left[ \sum_{k=1}^K (\gamma^* - \gamma_k)^2 \right]. \end{aligned} \quad (2.34)$$

■

### 2.6.6 引理2.3证明

**证明：**为了方便后续分析，我们记

$$z_{k+1} := \gamma_k + \eta_k (Q_k - \gamma_k L_k). \quad (2.35)$$

由于系数  $\gamma_{k+1} = [z_{k+1}]_{\gamma_{lb}}^{\gamma_{ub}}$ ，即可通过  $(z_{k+1} - \gamma^*)^2$  确定第  $k$  帧内系数  $\gamma_k$  的误差  $(\gamma_{k+1} - \gamma^*)^2$  上界：

$$(\gamma_{k+1} - \gamma^*)^2 = ([z_{k+1}]_{\gamma_{lb}}^{\gamma_{ub}} - [\gamma^*]_{\gamma_{lb}}^{\gamma_{ub}})^2 \leq (z_{k+1} - \gamma^*)^2. \quad (2.36)$$

回顾式 (2.14b) 中  $\gamma_k$  的更新准则，当没有采样频率约束时，采样频率违背

$U_k^+ \equiv 0$ , 我们有如下不等式:

$$\begin{aligned}
 & \frac{1}{2}(z_{k+1} - \gamma^*)^2 \\
 & \stackrel{(a)}{=} \frac{1}{2}(\gamma_k - \gamma^* + \eta_k(Q_k - \gamma_k L_k))^2 \\
 & = \frac{1}{2}(\gamma_k - \gamma^*)^2 + \frac{1}{2}\eta_k^2(Q_k - \gamma_k L_k)^2 + \eta_k(\gamma_k - \gamma^*)(Q_k - \gamma_k L_k) \\
 & \stackrel{(b)}{\leq} \frac{1}{2}(\gamma_k - \gamma^*)^2 + \frac{1}{2}\eta_k^2 L_{\text{ub}}^4 + \eta_k(\gamma_k - \gamma^*)(Q_k - \gamma_k L_k), 
 \end{aligned} \tag{2.37}$$

其中等式(a)由式(2.35)中对于 $z_k$ 的定义直接得到; 不等式(b)是因为第 $k$ 帧的奖励满足 $Q_k = \frac{1}{2}L_k^2 \leq L_{\text{ub}}^2$ 以及 $\gamma_k L_k \leq L_{\text{ub}}^2$ 。接下来, 对于不等式(2.37)的两边求取条件概率 $\mathcal{F}_{k-1}$ , 我们将得到以下不等式:

$$\begin{aligned}
 & \frac{1}{2}\mathbb{E}[(z_{k+1} - \gamma^*)^2 | \mathcal{F}_{k-1}] \\
 & = \frac{1}{2}(\gamma_k - \gamma^*)^2 + \frac{1}{2}\eta_k^2 L_{\text{ub}}^4 + \eta_k(\gamma_k - \gamma^*)\mathbb{E}[Q_k - \gamma_k L_k | \mathcal{F}_{k-1}]. 
 \end{aligned} \tag{2.38}$$

接下来我们将分情况讨论不等式(2.38)中最后一项的上界

$$(\gamma_k - \gamma^*)\mathbb{E}[Q_k - \gamma_k L_k | \mathcal{F}_{k-1}] \tag{2.39}$$

- 如果当前系数满足 $\gamma_k - \gamma^* \geq 0$ , 将不等式(2.22a)直接带入式(2.39)我们有:

$$\begin{aligned}
 & (\gamma_k - \gamma^*)\mathbb{E}[Q_k - \gamma_k L_k | \mathcal{F}_{k-1}] \\
 & \leq -(\gamma_k - \gamma^*)^2 \bar{L}^* \stackrel{(c)}{\leq} -(\gamma_k - \gamma^*)^2 \bar{D}, 
 \end{aligned} \tag{2.40}$$

其中不等式(c)是因为平均帧长满足 $\bar{L}^* \geq \bar{D}$ 。

- 如果当前系数满足 $\gamma_k - \gamma^* \leq 0$ , 我们用以下方法获取式(2.39)的上界:

$$\begin{aligned}
 & (\gamma_k - \gamma^*)\mathbb{E}[Q_k - \gamma_k L_k | \mathcal{F}_{k-1}] \\
 & = (\gamma_k - \gamma^*)\mathbb{E}[Q_k - \gamma^* L_k | \mathcal{F}_{k-1}] \\
 & \quad - (\gamma_k - \gamma^*)^2 \mathbb{E}[L_k | \mathcal{F}_{k-1}] \\
 & \stackrel{(d)}{\leq} (\gamma_k - \gamma^*)(\bar{Q}^* - \gamma^* \bar{L}^*) - (\gamma_k - \gamma^*)^2 \mathbb{E}[L_k | \mathcal{F}_{k-1}] \\
 & = -(\gamma_k - \gamma^*)^2 \mathbb{E}[L_k | \mathcal{F}_{k-1}] \\
 & \stackrel{(e)}{\leq} -(\gamma_k - \gamma^*)^2 \bar{D}, 
 \end{aligned} \tag{2.41}$$

其中不等式(d)是因为 $\mathbb{E}[Q_k - \gamma^* L_k | \gamma_k] \geq \bar{R}^* - \gamma^* \bar{L}^* = 0$ , 不等式(e)是因为 $\mathbb{E}[L_k | \mathcal{F}_{k-1}] \geq \bar{D}$ 。

将不等式(2.40)和(2.41)带入到式(2.38)中, 我们有:

$$\begin{aligned} & \frac{1}{2} \mathbb{E}[(z_{k+1} - \gamma^*)^2 | \mathcal{F}_{k-1}] \\ &= \left(\frac{1}{2} - \eta_k \bar{D}\right) (\gamma_k - \gamma^*)^2 + \frac{1}{2} \eta_k^2 L_{\text{ub}}^4 \\ &= \left(\frac{1}{2} - \eta_k \bar{D}_{\text{lb}}\right) (\gamma_k - \gamma^*)^2 + \frac{1}{2} \eta_k^2 L_{\text{ub}}^4. \end{aligned} \quad (2.42)$$

针对滤流  $\mathcal{F}_{k-1}$  求取期望, 即可确定  $\gamma_{k+1}$  与最优系数  $\gamma^*$  之间的误差:

$$\begin{aligned} & \frac{1}{2} \mathbb{E}[(\gamma_{k+1} - \gamma^*)^2] \stackrel{(f)}{\leq} \frac{1}{2} \mathbb{E}[(z_k - \gamma^*)^2] \\ & \leq \frac{1}{2} \left(1 - 2\eta_k \bar{D}_{\text{lb}}\right) \mathbb{E}[(\gamma_k - \gamma^*)^2] + \frac{1}{2} \eta_k^2 L_{\text{ub}}^4. \end{aligned} \quad (2.43)$$

其中不等式(f)由不等式(2.36)得到。

接下来, 我们通过数学归纳法证明, 当选取的步长满足  $\eta_1 = \frac{1}{2\bar{D}_{\text{lb}}}$  和  $\eta_k = \frac{1}{(k+2)\bar{D}_{\text{lb}}}, \forall k > 1$  时, 第  $k$  帧的系数  $\gamma_k$  与最优系数  $\gamma^*$  之间的差距满足如下不等式:

$$\frac{1}{2} \mathbb{E}[(\gamma_k - \gamma^*)^2] \leq \frac{1}{2k} \frac{L_{\text{ub}}^4}{\bar{D}_{\text{lb}}^2}. \quad (2.44)$$

证明过程如下:

- 当  $k = 2$  时, 将选取的步长  $\eta_1 = \frac{1}{2\bar{D}_{\text{lb}}}$  直接带入到式(2.42)我们得到:

$$\frac{1}{2} \mathbb{E}[(\gamma_2 - \gamma^*)^2] \leq \frac{1}{8} \frac{L_{\text{ub}}^4}{\bar{D}_{\text{lb}}^2} \leq \frac{1}{4} \frac{L_{\text{ub}}^4}{\bar{D}_{\text{lb}}^2}.$$

- 当  $k > 2$  时, 假定不等式  $\frac{1}{2} \mathbb{E}[(\gamma_k - \gamma^*)^2] \leq \frac{1}{2k} \frac{L_{\text{ub}}^4}{\bar{D}_{\text{lb}}^2}$  成立, 注意到步长  $\eta_k$  满足  $\eta_k = \frac{1}{(k+2)\bar{D}_{\text{lb}}}$ , 即可得到以下不等式:

$$\begin{aligned} & \frac{1}{2} \mathbb{E}[(\gamma_{k+1} - \gamma^*)^2] \\ & \leq \left(\frac{1}{2} - \eta_k \bar{D}_{\text{lb}}\right) (\gamma_k - \gamma^*)^2 + \frac{1}{2} \eta_k^2 L_{\text{ub}}^4 \\ & \leq \left(1 - \frac{2}{k+2}\right) \frac{1}{2k} \frac{L_{\text{ub}}^4}{\bar{D}_{\text{lb}}^2} + \frac{1}{2} \frac{1}{(k+2)^2} \frac{L_{\text{ub}}^4}{\bar{D}_{\text{lb}}^2} \\ & = \frac{1}{2} \left(\frac{1}{k+2} + \frac{1}{(k+2)^2}\right) \frac{L_{\text{ub}}^4}{\bar{D}_{\text{lb}}^2} \\ & = \frac{1}{2} \frac{k+3}{(k+2)^2} \frac{L_{\text{ub}}^4}{\bar{D}_{\text{lb}}^2} \end{aligned}$$

$$\stackrel{(g)}{\leq} \frac{1}{2(k+1)} \frac{L_{ub}^4}{D_{lb}^2}, \quad (2.45)$$

其中不等式 (g) 是因为  $(k+1)(k+3) \leq (k+2)^2$ 。

■

### 2.6.7 定理2.3证明

**证明：**注意到平均采样频率是否违背  $f_{max}$  的约束可以通过采样间隔违背序列  $U_k$  进行刻画，其中  $U_k$  的计算公式如下：

$$U_k = \sum_{k'=1}^k \left( \frac{1}{f_{max}} - (D_{k'} + W_{k'}) \right).$$

序列  $U_k$  可以被看作是一个排队系统。根据文献[86]定理2.8，为了证明所提策略满足采样频率约束，我们只需要证明以下不等式成立：

$$\lim_{K \rightarrow \infty} \frac{1}{K} \sum_{k=1}^K \mathbb{E}[U_k^+] < \infty. \quad (2.46)$$

我们用 Lyapunov 分析进行证明，定义采样违背为  $U_k$  时的 Lyapunov 函数为：

$$\mathcal{J}(U_k) := \frac{1}{2} [U_k^+]^2, \quad (2.47)$$

以及相应的 Lyapunov Drift 为：

$$\Delta(U_k) := \mathbb{E} [\mathcal{J}(U_{k+1}) - \mathcal{J}(U_k) | \mathcal{F}_{k-1}]. \quad (2.48)$$

首先我们要分析 Lyapunov Drift 的上界。注意到第  $k$  帧中的采样频率违背可以被写成如下所示的递归表达式：

$$U_{k+1} = U_k + \left( \frac{1}{f_{max}} - L_k \right), \forall k, \quad (2.49)$$

其中  $L_k = W_k + D_k$  为第  $k$  帧的长度。对于  $U_{k+1}^+$  取平方，我们有：

$$\begin{aligned} [U_{k+1}^+]^2 &= \left[ \max\{U_k - L_k + \frac{1}{f_{max}}, 0\} \right]^2 \\ &\leq \left[ \max\{U_k^+ - L_k + \frac{1}{f_{max}}, 0\} \right]^2 \\ &\leq \left[ U_k^+ - L_k + \frac{1}{f_{max}} \right]^2. \end{aligned} \quad (2.50)$$

考虑到任意采样等待时间均有上确界，即  $W_k \leq W_{ub}$ ，帧长  $L_k \leq B + W_{ub}$ ，即

可得到以下上界:

$$[U_{k+1}^+]^2 - [U_k^+]^2 \leq -2U_k^+ \left( L_k - \frac{1}{f_{\max}} \right) + \left( (B + W_{ub})^2 + \frac{1}{f_{\max}^2} \right). \quad (2.51)$$

对于式(2.51)两边关于传输时延  $D_k$  取期望, 即可得到 Lyapunov Drift  $\Delta(U_k) = \mathbb{E}[J(U_{k+1}) - J(U_k) | \mathcal{F}_{k-1}]$  的上界:

$$\Delta(U_k) \leq -U_k^+ \mathbb{E} \left[ L_k - \frac{1}{f_{\max}} | \mathcal{F}_{k-1} \right] + \frac{1}{2} \left( (B + W_{ub})^2 + \frac{1}{f_{\max}^2} \right). \quad (2.52)$$

接下来, 我们需要确定  $\mathbb{E} \left[ L_k - \frac{1}{f_{\max}} | \mathcal{F}_{k-1} \right]$  的一个下界。我们有如下引理:

**引理 2.4:** 根据问题2.1由严格可行解的假设2.2, 我们有如下不等式:

$$-U_k^+ \mathbb{E} \left[ L_k - \frac{1}{f_{\max}} | \mathcal{F}_{k-1} \right] \leq -\epsilon U_k^+ + V \left( \frac{1}{2}(B + W_{ub} + \epsilon)^2 + \gamma_{ub}(B + W_{ub}) \right). \quad (2.53)$$

它的证明将在附录2.6.8给出。将引理2.4中的不等式(2.53)带入式(2.52)即可得到:

$$\begin{aligned} & \mathbb{E} \left[ \frac{1}{2} [U_{k+1}^+]^2 - \frac{1}{2} [U_k^+]^2 | \mathcal{F}_{k-1} \right] \\ & \leq -\epsilon U_k^+ + \frac{1}{2} \left( (B + W_{ub})^2 + \frac{1}{f_{\max}^2} \right) + V \left( \frac{1}{2}(B + W_{ub})^2 + \gamma_{ub}(B + W_{ub}) \right). \end{aligned} \quad (2.54)$$

记

$$C := \frac{1}{2} \left( (B + W_{ub})^2 + \frac{1}{f_{\max}^2} \right) + V \left( \frac{1}{2}(B + W_{ub})^2 + \gamma_{ub}(B + W_{ub}) \right) < \infty. \quad (2.55)$$

将不等式(2.54)从帧  $k = 1$  加到帧  $K$  并对  $\mathcal{F}_K$  取期望, 注意到初始帧的采样违背被设置为  $U_1^+ = 0$ , 我们有:

$$\epsilon \sum_{k=1}^K \mathbb{E} [U_k^+] \leq \sum_{k=1}^K \mathbb{E} [\Delta(U_k)] + KC. \quad (2.56)$$

我们注意到 Lyapunov 函数式(2.47)是非负的, 初始点  $U_0 = 0$ , 对总帧数  $K$  取极限  $K \rightarrow \infty$  即可得到:

$$\limsup_{K \rightarrow \infty} \mathbb{E} \left[ \sum_{k=1}^K U_k^+ \right] < \frac{C}{\epsilon} < \infty, \quad (2.57)$$

以上不等式证明条件式(2.46)成立, 因此所提算法满足平均采样率约束条件。 ■

### 2.6.8 引理2.4证明

证明： 我们首先定义函数：

$$f(u, w, d) := -u^+(w + d) + V \left( \frac{1}{2}(d + w)^2 - \gamma(d + w) \right).$$

函数  $f(u, w, d)$  对于  $w$  的偏导计算如下：

$$\frac{\partial f(u, w, d)}{\partial w} = V \left( w + d - \left( \gamma + \frac{1}{V} u^+ \right) \right).$$

因此，对于给定的  $u$  和  $d$ ，最小化函数  $f(u, w, d)$  的参数  $w \geq 0$  为：

$$\arg \min_{w \geq 0} f(u, w, d) = \left( \gamma + \frac{1}{V} u^+ - d \right)^+. \quad (2.58)$$

我们注意到等待时间的选取规则与上式表达形式相同

$$W_k = \left( \gamma_k + \frac{1}{V} U_k^+ - D_k \right)^+.$$

也就是说，对于给定的传输时延  $D_k$  和采样频率违背  $U_k$ ，根据式 (2.58)，采样策略  $W_k$  的选取最小化函数  $f(u = U_k, w, d = D_k)$ ，因此对于其他等待时间选取策略  $W = \pi(D)$ ，我们有：

$$\begin{aligned} & -U_k^+(W_k + D_k) + V \left( \frac{1}{2}(D_k + W_k)^2 - \gamma_k(D_k + W_k) \right) \\ & \leq -U_k^+(\pi(D_k) + D_k) + V \left( \frac{1}{2}(D_k + \pi(D_k))^2 - \gamma_k(D_k + \pi(D_k)) \right). \end{aligned} \quad (2.59)$$

对式 (2.59) 两边关于传输时延  $D_k$  取期望后再在两边添加项  $U_k^+ \frac{1}{f_{\max}}$ ，即可得到：

$$\begin{aligned} & -U_k^+ \mathbb{E} \left[ D_k + W_k - \frac{1}{f_{\max}} | \mathcal{F}_{k-1} \right] + V \mathbb{E} \left[ \frac{1}{2}(D_k + W_k)^2 - \gamma_k(D_k + W_k) | \mathcal{F}_{k-1} \right] \\ & \leq -U_k^+ \mathbb{E} \left[ \pi(D_k) + D_k - \frac{1}{f_{\max}} | \mathcal{F}_{k-1} \right] + V \mathbb{E} \left[ \frac{1}{2}(D_k + \pi(D_k))^2 - \gamma_k(D_k + \pi(D_k)) | \mathcal{F}_{k-1} \right]. \end{aligned} \quad (2.60)$$

根据假设2.2，采样频率限制式 (2.5b) 可以被严格满足，我们定义策略  $\pi_\epsilon$ ，观测到传输时延  $D$  后，策略  $\pi_\epsilon$  选择等待  $(\gamma^\star + v^\star + \epsilon - D)^+$  并进行下一次采样。显然，策略  $\pi_\epsilon$  满足如下不等式：

$$\mathbb{E} [\pi_\epsilon(D) + D] \geq \frac{1}{f_{\max}} + \epsilon. \quad (2.61)$$

注意到传输时延  $D_k$  独立同分布，将式 (2.61) 带入到式 (2.60) 即可得到：

$$\begin{aligned} & -U_k^+ \mathbb{E} \left[ L_k - \frac{1}{f_{\max}} | \mathcal{F}_{k-1} \right] + V \mathbb{E} \left[ \frac{1}{2}(D_k + W_k)^2 - \gamma_k(D_k + W_k) | \mathcal{F}_{k-1} \right] \\ & \leq -U_k^+ \mathbb{E} \left[ D_k + \pi_\epsilon(D_k) - \frac{1}{f_{\max}} \right] + V \mathbb{E} \left[ \frac{1}{2}(D_k + \pi_\epsilon(D_k))^2 - \gamma_k(D_k + \pi_\epsilon(D_k)) | \mathcal{F}_{k-1} \right] \end{aligned}$$

$$\stackrel{(a)}{\leq} -\epsilon U_k^+ + V \mathbb{E} \left[ \frac{1}{2} (D_k + \pi_\epsilon(D_k))^2 - \gamma_k (D_k + \pi_\epsilon(D_k)) | \mathcal{F}_{k-1} \right], \quad (2.62)$$

其中不等式(a)可以由式(2.61)直接得到。进一步化简以上不等式即可证明:

$$\begin{aligned} & -U_k^+ \mathbb{E} \left[ L_k - \frac{1}{f_{\max}} | \mathcal{F}_{k-1} \right] \\ & \leq -\epsilon U_k^+ + V \left( \frac{1}{2} (B + W_{ub} + \epsilon)^2 + \gamma_{ub} (B + W_{ub}) \right). \end{aligned} \quad (2.63)$$

■

## 第3章 多传感器信息年龄最小化的状态信息采集和跨层传输算法研究

### 3.1 引言

在第二章的基础上，本章进一步考虑车联网中的中央控制器通过有限带宽从多个传感器进行状态信息采集的场景，旨在设计优化接收端状态信息时效性的采样和调度算法。如节1.3.1.2所述，目前多传感器网络时效性最优的采样方案主要考虑传输信道时不变的场景，在传输信道时变的场景中，如何设计信息年龄最小的状态信息采样和传输方案以及这些方案的理论性能有待进一步研究。

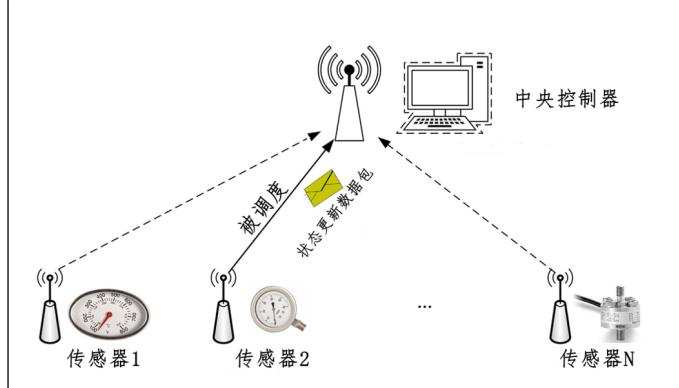
本章研究如图3.1所示的多传感器无线网络中，信道衰落有记忆、传输功率受限时，最小化网络中的平均信息年龄的采样、传输功率和调度跨层设计问题。该问题可以被建模为一个离散时间马尔可夫决策过程 (Markov Decision Process, MDP)，解决该 MDP 的主要困难在于状态数随着传感器数量指数增长面临“维度爆炸”。为了应对这一挑战，我们首先松弛每时刻带宽限制为时间平均的带宽限制，通过拉格朗日乘子法转化为单传感器马尔可夫决策过程，并利用最优解的阈值结构将其进一步转化为线性规划问题。接下来通过对偶法搜索可以满足时间平均带宽约束的最优拉格朗日乘子，最后通过截断调度法设计了一个可以满足原每时刻带宽约束的调度方案并证明了所提算法在大规模网络中的理论最优性。

### 3.2 系统模型和问题建模

#### 3.2.1 网络模型

考虑如图3.1所示的网络，中心控制节点通过无线信道从  $N$  个传感器采集状态信息。考虑离散时间系统，我们用  $t \in \{1, \dots, T\}$  表示当前时刻，并记变量  $u_n(t) \in \{0, 1\}$  为时刻  $t$  中心控制节点对于传感器  $n$  的调度策略。其中  $u_n(t) = 1$  表示中心节点决定在时刻  $t$  从传感器  $n$  采集信息，在时隙  $t$  开始时传感器  $n$  就将它当前的观测发送给中心控制节点。考虑传输是无丢包的，传输时延为 1 时隙。由于带宽限制，每时隙至多有  $M$  ( $M < N$ ) 个传感器可以一起被调度发送状态信息，这一限制对于调度策略  $\{u_n(t)\}$  的要求是：

$$\sum_{n=1}^N u_n(t) \leq M, \forall t. \quad (3.1)$$



$N$  个传感器向中心控制器发送信息，由于带宽限制，每时刻只有不多于  $M$  个传感器可以发送信息。

图 3.1 多传感器网络状态信息采集系统模型

接下来我们介绍对于时变信道的建模。我们将信道分为  $Q$  个状态，记  $q_n(t)$  为传感器  $n$  在第  $t$  个时隙的信道状态。为了刻画信道传输的时变和有记忆性，我们假设每个传感器的信道状态序列  $\{q_n(t)\}_{t=1}^{\infty}$  是一个遍历马尔可夫链，记  $p_{i,j}^{(n)}$  为传感器  $n$  的马尔可夫链从状态  $i$  到状态  $j$  的一步转移概率  $p_{i,j}^{(n)} \triangleq \Pr(q_n(t+1) = j | q_n(t) = i)$ 。

若传感器  $n$  在信道状态为  $q$  时被中心节点选择调度，为了保证可靠传输，即信道容量大于状态信息数据包大小，它将使用  $\omega(q)$  的能量进行发送。那么对于一组时隙  $1 \sim T$  的传输决策  $\mathbf{u}_n(\pi) := [u_n(1), \dots, u_n(T)]$ ，这  $T$  个时隙内传感器  $n$  所消耗的平均功率  $E_n$  为：

$$E_n(\mathbf{u}_n(\pi)) := \frac{1}{T} \sum_{t=1}^T u_n(t) \omega(q_n(t)). \quad (3.2)$$

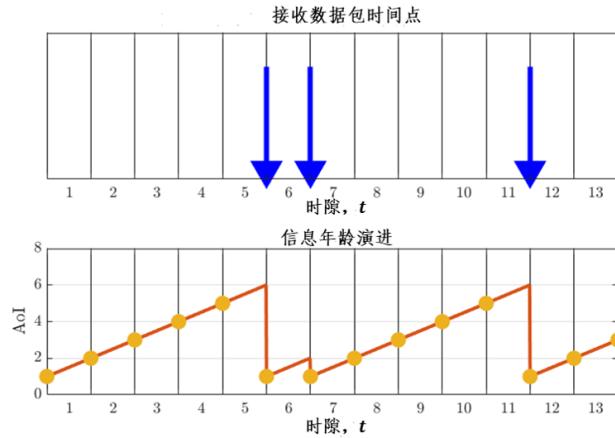
### 3.2.2 信息年龄

我们用信息年龄<sup>[5]</sup>度量中央控制器所拥有的全网传感器的状态信息时效性。根据定义，信息年龄为当前时刻  $t$  与接收端所拥有的最新信息产生时刻之间的时间差。我们把传感器  $n$  在第  $t$  时隙开始时的信息年龄记作  $x_n(t)$ 。与研究<sup>[15,28]</sup>中的假设相同，我们考虑随意采样模型 (generate-at-will)，传感器每时刻都可以采样状态信息并产生新的数据包用于发送。也就是说，如果传感器  $n$  在第  $t$  时隙被调度，一个包含有第  $t$  时隙最新信息的数据包将在第  $t$  时隙末尾被中央控制器收到，那么根据定义，在第  $t+1$  时隙开头，中央控制器将拥有传感器  $n$  在第  $t$  时隙开头最新的状态信息，因此  $x_n(t+1) = 1$ ；若传感器  $n$  没有在第  $t$  时隙被调度，接收端所拥有的最新信息产生时刻不变，因此  $t+1$  时隙刚开始的信息年龄  $x_n(t+1) = x_n(t) + 1$ 。

综上所述，本章中，信息年龄  $x_n(t)$  的演进公式如下：

$$x_n(t+1) = \begin{cases} 1, & u_n(t) = 1; \\ x_n(t) + 1, & u_n(t) = 0. \end{cases} \quad (3.3)$$

图3.2展示了信息年龄随着调度决策的演进曲线。



上图：蓝色向下的箭头表示中央控制器接收到传输数据包的时刻；下图：红色曲线表示信息年龄的演进，其中每时隙开始的信息年龄使用黄色圆点进行标注。由图可知，每当一个新的数据包被收到，接收端的信息年龄就降为1。

图 3.2 离散时间随意采样 (generate-at-will) 模型中信息年龄演进曲线

### 3.2.3 优化问题建模

我们考虑有  $N$  个传感器接入的网络，每个传感器和中央控制器之间的信道被建模为转移概率为  $\{p_{i,j}^{(n)}\}$  的马尔可夫链。我们将  $J(\pi)$  记为使用策略  $\pi$  时，全网所有传感器在每时隙开始的平均信息年龄：

$$J(\pi) \triangleq \lim_{T \rightarrow \infty} \frac{1}{NT} \mathbb{E}_\pi \left[ \sum_{t=1}^T \sum_{n=1}^N x_n(t) | \mathbf{x}(0) \right], \quad (3.4)$$

其中向量  $\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_N(t)] \in \mathbb{N}^N$  表示全网所有传感器在时隙  $t$  开始的信息年龄。本研究中，我们假设时隙  $t = 0$  时，中央控制器处所有传感器的信息年龄都是 1，即  $\mathbf{x}(0) = \mathbf{1}$ 。

我们用集合  $\Pi$  表示因果策略，即每时刻的调度策略  $\{u_n(t)\}_{n=1}^N$  仅依赖于当前时刻所有传感器的信息年龄  $\{x_n(t)\}_{n=1}^N$ 、信道状态  $\{q_n(t)\}_{n=1}^N$  以及信道状态的转移概率  $\{p_{i,j}^{(n)}\}$ 。未来的信道状态将不会被用于当前时刻的决策。考虑各节点的功率约束在中央传感器处已知，我们寻找一个因果策略  $\pi := \{u_n(t)\} \in \Pi$ ，在满足网络带宽和传感器功率约束下，优化全网信息年龄。这一带宽、功率双约束优化问题如

下：

**问题3.1(带宽功率双约束优化问题):**

$$\pi^* = \arg \min_{\pi \in \Pi} \lim_{T \rightarrow \infty} \left\{ \frac{1}{NT} \mathbb{E}_\pi \left[ \sum_{t=1}^T \sum_{n=1}^N x_n(t) \right] \right\}, \quad (3.5a)$$

$$\text{s.t. } \mathbb{E}_\pi \left[ \sum_{n=1}^N u_n(t) \right] \leq M, \forall t, \quad (3.5b)$$

$$\lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E}_\pi \left[ \sum_{t=1}^T u_n(t) \omega(q_n(t)) \right] \leq \mathcal{E}_n, \forall n. \quad (3.5c)$$

由于带宽约束式(3.5b)，每时隙总计有 $\binom{N}{1} + \dots + \binom{N}{M}$ 种可能的调度策略。这一维度爆炸现象使得值迭代和策略迭代不能直接用于问题求解。为了解决以上这一困难，我们提出如下方案：

- 受到研究<sup>[43,49,62]</sup>的启发，在节3.3，我们首先通过带宽松弛和拉格朗日乘子法进行传感器的解耦(3.5b)。解耦后，每个传感器可以独立去设计策略，最小化它们自己的信息年龄和调度代价之和。
- 节3.4，我们将在单个传感器最优传输策略设计的基础上，设计一种随机调度方案满足原带宽、功率双约束优化问题中每时隙的严格带宽约束(3.5b)。

### 3.3 单传感器解耦及其求解

本节中我们将问题3.1解耦为单传感器代价最小化问题。进一步地，我们利用单传感器代价最小化问题最优采样和传输方案的阈值结构，将最优策略求解问题转化为一个线性规划。

#### 3.3.1 单传感器解耦

首先我们将每时刻调度的传感器数量小于等于 $M$ 这一严格约束(3.5b)松弛为时间平均上调度的传感器总数小于等于 $M$ ，松弛后的优化问题如下：

**问题3.2(时间平均带宽约束、功率约束优化问题):**

$$\pi_R^* = \arg \min_{\pi \in \Pi} \lim_{T \rightarrow \infty} \left\{ \frac{1}{NT} \mathbb{E}_\pi \left[ \sum_{t=1}^T \sum_{n=1}^N x_n(t) \right] \right\}, \quad (3.6a)$$

$$\text{s.t. } \lim_{T \rightarrow \infty} \mathbb{E}_\pi \left[ \frac{1}{T} \sum_{t=1}^T \sum_{n=1}^N u_n(t) \right] \leq M, \quad (3.6b)$$

$$\lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E}_\pi \left[ \sum_{t=1}^T u_n(t) \omega(q_n(t)) \right] \leq \mathcal{E}_n, \forall n. \quad (3.6c)$$

注意到任意一个满足原问题3.1中每时隙带宽严格约束的调度策略  $\pi$  一定可以满足问题3.2中的平均带宽约束，因此，问题3.1中的可行策略是松弛问题3.2可行策略的子集，采用问题3.2中的最优策略  $\pi_R^\star$  得到的全网平均信息年龄  $J(\pi_R^\star)$  是问题3.1中最优策略  $J(\pi^\star)$  的下界，即  $J(\pi_R^\star) \leq J(\pi^\star)$ 。我们注意到带宽松弛后的问题仍然是一个决策空间为  $2^N$  的受限马尔可夫决策过程，仍然无法直接求解。为了解决问题3.2，我们首先介绍以下引理：

**引理 3.1：** 问题3.2有一个可以写成如下形式的最优调度方案  $\pi_R^\star$ :

$$\pi_R^\star = \otimes_{n=1}^N \pi_{R,n}^\star, \quad (3.7)$$

策略  $\pi_R^\star$  每时刻是否调度传感器  $n$  由策略  $\pi_{R,n}^\star$  根据传感器  $n$  当前的信道状态  $q_n(t)$  和信息年龄  $x_n(t)$  决定，与其他传感器的信息年龄和信道状态无关。

接下来我们寻找如引理3.1所示的最优策略  $\pi_R^\star = \otimes_{n=1}^N \pi_{R,n}^\star$ 。我们将松弛后的带宽约束通过拉格朗日乘子  $W \geq 0$  写出拉格朗日函数：

$$\mathcal{L}(\pi, W) = \lim_{T \rightarrow \infty} \left\{ \frac{1}{NT} \mathbb{E}_\pi \left[ \sum_{n=1}^N \sum_{t=1}^T \left( x_n(t) + W u_n(t) - \frac{WM}{N} \right) \right] \right\}. \quad (3.8)$$

定义  $\pi(W) \triangleq \arg \min_{\pi \in \Pi} \mathcal{L}(\pi, W)$  为满足功率约束式 (3.5c) 条件下最小化式 (3.8) 的最优策略。根据文献<sup>[87]</sup>中第四部分的推导，对于给定的拉格朗日乘子  $W$ ，策略  $\pi(W)$  可以解耦写成  $\pi(W) = \otimes_{n=1}^N \pi_{d,n}(W)$  的形式，每个解耦后的策略  $\pi_{d,n}(W)$  为如下问题最优解：

**问题 3.3 (解耦后单节点功率限制下信息年龄和调度代价总和最小化问题):**

$$\pi_{d,n}^\star(W) = \arg \min_{\pi \in \Pi} \mathcal{L}(\pi_n, W), \text{ 其中} \quad (3.9a)$$

$$\mathcal{L}_n(\pi_n, W) = \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E}_{\pi_n} \left[ \sum_{t=1}^T x_n(t) + W u_n(t) \right], \quad (3.9b)$$

$$\text{s.t. } \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E}_{\pi_n} \left[ \sum_{t=1}^T u_n(t) \omega(q_n(t)) \right] \leq \mathcal{E}_n. \quad (3.9c)$$

对于给定的拉格朗日乘子，求解最优的调度策略  $\pi(W)$  可以转化为求解  $N$  个问题3.3。接下来我们将把问题3.3建模为受限马尔可夫决策过程，分析它的最优阈值结构，这一阈值结构将允许我们使用线性规划的方法分别求解每个传感器  $n$  对应的问题3.3。

### 3.3.2 受限马尔可夫决策过程及其阈值结构

由于这  $N$  个解耦后的问题3.3结构相同，本节中我们讨论求解问题3.3的一般方法并忽略下标  $n$ 。问题3.3是一个受限马尔可夫决策过程，它的四要素  $(\mathbb{S}, \mathbb{A}, \text{Pr}(\cdot|\cdot), C(\cdot, \cdot))$  如下：

- **状态空间  $\mathcal{S}$** : 受限马尔可夫决策过程在时间  $t$  的状态为当前时刻的信息年龄和信道状态  $(x(t), q(t))$ 。由于信道状态  $q(t)$  从有限可数集  $Q$  中选取，信息年龄  $x(t) \in \mathbb{N}^+$ ，状态空间  $\mathbb{S} = \{\mathbb{N}^+ \times Q\}$  可数但有无限多的元素。
- **决策空间  $\mathcal{A}$** : 每时刻的决策  $u(t) \in \mathbb{A} = \{0, 1\}$ ，其中  $u(t) = 1$  表示传感器在时刻  $t$  开始被中心控制器调度并发送状态信息，而  $u(t) = 0$  表示传感器在时刻  $t$  没有被调度。
- **状态转移函数  $\text{Pr}(\cdot \rightarrow \cdot | \cdot)$** : 如果传感器在  $t$  时刻没有被调度，即  $u(t) = 0$ ，中心节点处关于传感器的信息年龄线性增长  $x(t+1) = x(t) + 1$ ；如果传感器在时刻  $t$  被调度了，那么在时隙  $t+1$  刚开始，他的信息年龄  $x(t+1) = 1$ 。根据信道状态的演进特性， $q(t+1)$  仅和前一时刻的信道状态  $q(t)$  有关，与当前时刻的信息年龄  $x(t)$  无关。因此，马尔可夫链从状态  $(x, q)$  转移到状态  $(x', q')$  的概率为：

$$\text{Pr}((x, q) \rightarrow (x', q')) = \begin{cases} p_{q, q'}, & \{u = 0, x' = x + 1\} \text{ or } \{u = 1, x' = 1\}; \\ 0, & \text{其他情况.} \end{cases} \quad (3.10)$$

- **单步代价函数  $C(\cdot, \cdot)$** : 问题3.3需要优化的目标函数包括信息年龄和调度代价之和。在状态  $(x, q)$  下选取策略  $u$  时，受限马尔可夫问题的单步代价为：

$$C_X(x, q, u) = x + W u, \quad (3.11a)$$

每一步的功率消耗由信道状态  $q$  和决策  $u$  共同决定：

$$C_Q(x, q, u) = \omega(q)u. \quad (3.11b)$$

这一受限马尔可夫决策过程的目标是设计策略  $\pi$ ，在满足平均功率约束下：

$$\lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E}_\pi \left[ \sum_{t=1}^T C_Q(x(t), q(t), s(t)) \right] \leq \mathcal{E},$$

最小化平均代价（即信息年龄和传输代价的总和）

$$\lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E}_\pi \left[ \sum_{t=1}^T C_X(x(t), q(t), s(t)) \right].$$

在刻画上述受限马尔可夫决策过程的最优解结构之前，我们首先给出本章中

静态随机策略和静态确定性策略的定义:

**定义 3.1:** 我们用  $\Pi_{\text{SR}}$  和  $\Pi_{\text{SD}}$  分别表示静态随机策略和静态确定性策略。给定状态  $(x(t) = x, q(t) = q)$ , 一个静态随机策略  $\pi_{\text{SR}} \in \Pi_{\text{SR}}$  以一个不随时间  $t$  变化的概率  $\xi_{x,q} \in [0, 1]$  选取策略  $u = 1$ 。对于静态确定性策略  $\pi_{\text{SD}} \in \Pi_{\text{SD}}$ , 每时刻的策略可以写成函数  $u(t) = a(x, q)$ , 其中函数  $a(\cdot) : (x, q) \mapsto \{0, 1\}$  是从状态空间  $\mathbb{S} \rightarrow \mathbb{A}$  到决策空间的确定性映射。

根据文献<sup>[88]</sup>中定理 4.4 和定理 6.3, 上述受限马尔可夫决策过程的最优策略具有如下性质:

**推论 3.1:** 存在一个静态随机策略  $\pi_d^* \in \Pi_{\text{SR}}$  为问题3.3的最优解。同时, 存在一个拉格朗日乘子  $\lambda \geq 0$ , 最优策略  $\pi_d^*$  同时也是最小化如下拉格朗日函数的最佳策略 (我们将拉格朗日函数中的常数项  $-\lambda \mathcal{E}$  省略)

$$\pi_d^*(\lambda) \triangleq \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E}_{\pi} \left[ \sum_{t=1}^T [C_X(x(t), q(t), s(t)) + \lambda C_Q(x(t), q(t), s(t))] \right]. \quad (3.12)$$

接下来我们证明所有最小化拉格朗日函数式 (3.12) 的策略都有阈值结构。记  $\gamma$  为使用最优策略得到的平均代价, 那么存在一个值函数  $V(x, q)$  满足如下贝尔曼方程:

$$V(x, q) + \gamma = \min\{C_X(x, q, 0) + \sum_{q'=1}^Q p_{q,q'} V(x+1, q'), \quad (3.13)$$

$$C_X(x, q, 1) + \sum_{q'=1}^Q p_{q,q'} V(1, q') + \lambda C_Q(x, q, 1)\}, \quad (3.14)$$

贝尔曼方程将决定最小化拉格朗日函数 (3.12) 的阈值结构, 我们有如下引理:

**引理 3.2:** 对于任意给定的拉格朗日乘子  $\lambda$ , 策略  $\pi_d^*(\lambda)$  均有阈值结构。也就是说, 对于每一个信道状态  $q$ , 存在阈值  $\tau_q$ , 如果当前信息年龄  $x > \tau_q$ , 问题3.3的最优策略  $\pi_d^*(\lambda)$  选择调度传感器; 当  $x < \tau_q$  时, 策略  $\pi_d^*(\lambda)$  选择空闲。

**证明:** 阈值结构的证明附在本章附录节3.7.2中。这里我们提供一个直观的解释。由于传感器与中央控制器的通信受到功率的限制, 为了达到一个较小的信息年龄, 我们只在信息不够新 (即信息年龄较大) 和信道条件好的时候进行调度。 ■

### 3.3.3 受限马尔可夫决策过程的线性规划求解

接下来我们求解受限马尔可夫决策过程3.3的最优静态随机策略。对于任意一个静态随机策略  $\pi \in \Pi_{\text{SR}}$ , 记  $\xi_{x,q}$  为传感器处于状态  $(x, q)$  时被调度的概率。我们的目标是寻找一组最优的调度概率  $\{\xi_{x,q}^*\}$  最小化问题3.3的目标函数。根据上一节得到的阈值结构, 存有一组  $\{\tau_q\}$ , 使得最优策略  $\pi_d^*$  对应的调度概率满足

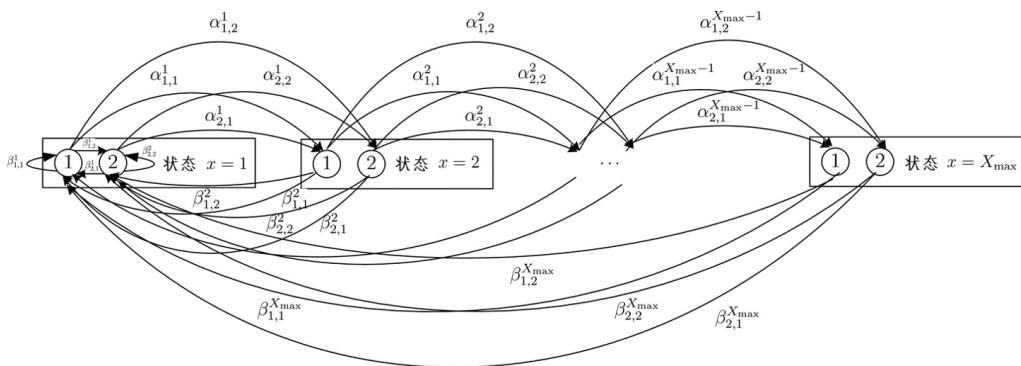
$\xi_{x,q}^* = 1, \forall x \geq \max_q \tau_q$ 。为了寻找这样的最优策略，在接下来的讨论中我们选取一个足够大的  $X_{\max} \geq \max_q \tau_q$ ，在搜索过程中只关注那些满足  $\xi_{x,q} = 1, \forall x \geq X_{\max}$  的策略，其余策略因为不满足阈值结构将被排除在搜索范围以外。

我们记  $\mu_{x,q}$  为选用策略  $\{\xi_{x,q}\}$  时，传感器处于信息年龄为  $x$ 、信道状态为  $q$  的稳态概率。当  $X_{\max} \geq \max_q \tau_q$  时，传感器的信息年龄在  $x = X_{\max}$  时都将被调度并将在下一时刻回到 1，因此最优策略可以保证  $\mu_{X_{\max}+1,q} = 0, \forall q$ ，我们也不再关注那些  $x \geq X_{\max}$  的状态。

为了方便之后计算静态随机策略带来的信息年龄和调度代价，图3.3展示了一个两信道状态（即  $Q = 2$ ）下，信息年龄-信道状态转移图。记  $\alpha_{q,q'}^x$  为使用策略  $\{\xi_{x,q}\}$  时，从状态  $(x, q)$  到状态  $(x + 1, q')$  一步转移概率，记  $\beta_{q,q'}^x$  为状态  $(x, q)$  到状态  $(1, q')$  的转移概率，这些参数可以通过如下公式计算：

$$\alpha_{q,q'}^x = \Pr((x, q) \rightarrow (x + 1, q')) = (1 - \xi_{x,q})p_{q,q'}, \quad (3.15a)$$

$$\beta_{q,q'}^x = \Pr((x, q) \rightarrow (1, q')) = \xi_{x,q}p_{q,q'}. \quad (3.15b)$$



圆圈表示信道状态  $q$ ，长方形表示信息年龄  $x$ 。从状态  $(x, q)$  到  $(x + 1, q')$  的转移概率为  $\alpha_{q,q'}^x$ ，从状态  $(x, q)$  到状态  $(1, q')$  的转移概率为  $\beta_{q,q'}^x$ 。

图 3.3 两状态信道  $Q = 2$  的  $(x, q)$  状态跳转示意图

记  $\boldsymbol{\mu} = [\mu_{1,1}, \dots, \mu_{1,Q}, \dots, \mu_{X_{\max},1}, \dots, \mu_{X_{\max},Q}]^T$  为马尔可夫链的稳态概率分布。接下来我们根据图3.3写出状态转移矩阵  $\boldsymbol{\Phi}$ ：

$$\boldsymbol{\Phi} = \begin{bmatrix} \boldsymbol{\beta}^1 & \boldsymbol{\beta}^2 & \cdots & \boldsymbol{\beta}^{X_{\max}-1} & \boldsymbol{\beta}^{X_{\max}} \\ \boldsymbol{\alpha}^1 & \mathbf{0}_Q & \cdots & \mathbf{0}_Q & \mathbf{0}_Q \\ \mathbf{0}_Q & \boldsymbol{\alpha}^2 & \cdots & \mathbf{0}_Q & \mathbf{0}_Q \\ \cdots & \cdots & \cdots & \cdots \cdots & \cdots \\ \mathbf{0}_Q & \mathbf{0}_Q & \cdots & \boldsymbol{\alpha}^{X_{\max}-1} & \mathbf{0}_Q \end{bmatrix}, \quad (3.16)$$

其中  $\mathbf{0}_Q$  为一个  $Q$  维全零向量。矩阵  $\boldsymbol{\alpha}^x$  和  $\boldsymbol{\beta}^x$  为状态  $\{x\} \times Q$  出发向状态  $\{(x+1)\} \times Q$

和状态  $\{1\} \times Q$  转移的概率，可以通过如下表达式计算：

$$\boldsymbol{\alpha}^x = \begin{bmatrix} \alpha_{1,1}^x & \alpha_{2,1}^x & \cdots & \alpha_{Q,1}^x \\ \alpha_{1,2}^x & \alpha_{2,2}^x & \cdots & \alpha_{Q,2}^x \\ \cdots & \cdots & \cdots & \cdots \\ \alpha_{1,Q}^x & \alpha_{2,Q}^x & \cdots & \alpha_{Q,Q}^x \end{bmatrix}, \quad (3.17a)$$

$$\boldsymbol{\beta}^x = \begin{bmatrix} \beta_{1,1}^x & \beta_{2,1}^x & \cdots & \beta_{Q,1}^x \\ \beta_{1,2}^x & \beta_{2,2}^x & \cdots & \beta_{Q,2}^x \\ \cdots & \cdots & \cdots & \cdots \\ \beta_{1,Q}^x & \beta_{2,Q}^x & \cdots & \beta_{Q,Q}^x \end{bmatrix}. \quad (3.17b)$$

马尔可夫链的稳态分布满足  $\boldsymbol{\Phi}\boldsymbol{\mu} = \boldsymbol{\mu}$ 。对于信息年龄  $\forall x \geq X_{\max} + 1$  的状态，稳态分布概率  $\mu_{x,q} = 0, \forall q$ 。因此，我们得到  $\sum_{x=1}^{X_{\max}} \sum_{q=1}^Q \mu_{x,q} = 1$  这样一个限制条件。使用策略  $\{\xi_{x,q}\}$  时，马尔可夫链的稳态分布  $\boldsymbol{\mu}$  为以下线性方程组的解：

$$\begin{bmatrix} \mathbf{Q} - \mathbf{I}_{QX_{\max}} \\ \mathbf{1}_{QX_{\max}}^T \end{bmatrix} \boldsymbol{\mu} = \begin{bmatrix} \mathbf{0}_{QX_{\max}} \\ 1 \end{bmatrix}, \quad (3.18)$$

其中  $\mathbf{1}_{QX_{\max}}$  为  $(Q \times X_{\max})$  维全 1 向量， $\mathbf{I}_{QX_{\max}}$  为  $(Q \times X_{\max})$  维单位阵。

我们引入新变量  $y_{x,q} = \mu_{x,q}\xi_{x,q}$ ，他表示传感器处于状态  $(x, q)$  且被调度的概率。我们有如下定理：

**定理 3.1：** 求解单传感器受限马尔可夫决策过程 3.3 最优策略可以转化为求解如下线性规划问题：

**问题 3.4 (单传感器受限马尔可夫决策过程的等价线性规划问题)：**

$$\{\mu_{x,q}^*, y_{x,q}^*\} = \arg \min_{\{\mu_{x,q}, y_{x,q}\}} \sum_{x=1}^{X_{\max}} \sum_{q=1}^Q (W y_{x,q} + x \mu_{x,q}), \quad (3.19a)$$

$$\text{s.t. } \mu_{1,q} = \sum_{x=1}^X \sum_{q'=1}^Q y_{x,q'} p_{q',q}, \quad (3.19b)$$

$$\mu_{x,q} = \sum_{q'=1}^Q (\mu_{x-1,q'} - y_{x-1,q'}) p_{q',q}, \quad (3.19c)$$

$$\sum_{x=1}^{X_{\max}} \sum_{q=1}^Q \mu_{x,q} = 1, \quad (3.19d)$$

$$y_{x,q} \leq \mu_{x,q}, \quad (3.19e)$$

$$\sum_{x=1}^{X_{\max}} \sum_{q=1}^Q y_{x,q} \omega(q) \leq \mathcal{E} \quad (3.19f)$$

$$0 \leq \mu_{x,q} \leq 1, 0 \leq y_{x,q} \leq 1, \forall x, q. \quad (3.19g)$$

**证明：** 我们首先把问题3.3的优化目标式(3.9b)写成关于 $\{\mu_{x,q}\}$ 和 $\{y_{x,q}\}$ 的函数。注意到传感器处于状态 $(x, q)$ 的概率为 $\mu_{x,q}$ 。处于状态 $(x, q)$ 时，传感器将以概率 $\xi_{x,q}$ 被调度并产生代价 $C_X(x, q, 1) = x + W$ ，以概率 $1 - \xi_{x,q}$ 传感器将处于空闲并产生代价 $C_X(x, q, 0) = x$ 。因此采用策略 $\{\xi_{x,q}\}$ 得到的时间平均代价可以通过以下公式计算：

$$\sum_{x=1}^{X_{\max}} \sum_{q=1}^Q \mu_{x,q} (\xi_{x,q}(x + W) + (1 - \xi_{x,q})x) = \sum_{x=1}^{X_{\max}} \sum_{q=1}^Q (W y_{x,q} + x \mu_{x,q}). \quad (3.20)$$

若传感器在 $(x, q)$ 被调度，他将使用 $\omega(q)$ 功率发送信息。因此，使用策略 $\{\xi_{x,q}\}$ 带来的平均能量开销为：

$$\sum_{x=1}^{X_{\max}} \sum_{q=1}^Q \mu_{x,q} \xi_{x,q} \omega(q) = \sum_{x=1}^{X_{\max}} \sum_{q=1}^Q y_{x,q} \omega(q). \quad (3.21)$$

因此，受限马尔可夫决策过程的能量约束式(3.5c)可以转化为线性规划中的约束式(3.19f)。约束式(3.19b) - (3.19d)来自于 $\xi_{x,q}$ 与稳态分布 $y_{x,q}, \mu_{x,q}$ 之间的关系式(3.18)。注意到每个状态 $(x, q)$ 的调度概率都要满足 $\xi_{x,q} \leq 1$ ，我们可以得到约束式(3.19e)。 ■

到目前为止，我们通过求解线性规划获得了在给定拉格朗日乘子 $W$ 下，问题3.3的最优静态策略对应的稳态分布 $\mu_{x,q}$ 和 $y_{x,q}$ ，优化拉格朗日函数式(3.9b)的最优策略 $\xi_{x,q}$ 可以通过 $\xi_{x,q} = \frac{y_{x,q}}{\mu_{x,q}}$ 获得。根据阈值结构，我们还有如下推论可以帮助我们简化后续计算：

**推论 3.2：** 对于任意的信道状态 $q$ ，优化问题3.3的调度概率 $\xi_{\cdot,q}^*$ 是一个非递减序列，即

$$\xi_{x_1,q}^* \leq \xi_{x_2,q}^*, \forall 1 \leq x_1 < x_2. \quad (3.22)$$

### 3.4 多传感器机会调度

上一节中我们讨论了给定拉格朗日乘子 $W$ ，如何通过线性规划的方法求解受限马尔可夫决策过程(问题3.3)的最优策略。本节中，我们将研究如何选取 $W$ 得到满足时间平均带宽约束的最优策略 $\pi_R^*$ 。最后，我们将基于策略 $\pi_R^*$ 提出一种可以满足原每时隙带宽约束式(3.5b)的调度策略并分析其理论特性。

### 3.4.1 时间平均带宽约束下最优策略 $\pi_R^*$ 求解

对于任意拉格朗日乘子  $W \geq 0$ , 记  $g(W)$  为对偶函数, 即

$$g(W) \triangleq \min_{\pi \in \Pi} \mathcal{L}(\pi, W). \quad (3.23)$$

由于最小化拉格朗日函数式 (3.8) 可以被解耦为  $N$  个受限马尔可夫决策过程 (问题3.3), 因此对偶函数  $g(W)$  还可以写作:

$$\begin{aligned} g(W) &= \frac{1}{N} \sum_{n=1}^N g_n(W) - WM, \text{ 其中} \\ g_n(W) &\triangleq \min_{\pi_n \in \Pi} (\mathcal{L}_n(\pi_n, W)), \text{ s.t. } \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \omega(q_n(t)) u_n(t) \leq \mathcal{E}_n. \end{aligned} \quad (3.24)$$

根据定理3.1, 在约束条件式 (3.9b) 下, 寻找最小化函数  $\mathcal{L}_n(\pi, W)$  的策略  $\pi_{d,n}^*(W)$  可以被转化为一个线性规划, 我们记  $\mu_{x,q}^{n,W}, y_{x,q}^{n,W}$  为第  $n$  个传感器对应线性规划的最优解,  $g_n(W)$  即为该线性规划问题的最小值。我们用  $\bar{X}_n(W)$  和  $\bar{B}_n(W)$  分别表示拉格朗日乘子为  $W$  时, 使用最优策略  $\pi_{d,n}^*(W)$  后, 传感器  $n$  的平均信息年龄和调度概率, 他们可以由线性规划的解  $\{y_{x,q}^{n,W}\}$  直接计算得到:

$$\bar{X}_n(W) = \sum_{x=1}^{X_{\max}} \sum_{q=1}^Q x \mu_{x,q}^{n,W}, \quad (3.25a)$$

$$\bar{B}_n(W) = \sum_{x=1}^{X_{\max}} \sum_{q=1}^Q y_{x,q}^{n,W}. \quad (3.25b)$$

通过  $\bar{X}_n(W)$  和  $\bar{B}_n(W)$ , 我们可以直接写出对偶函数  $g_n(W)$  的值:

$$g_n(W) = \bar{X}_n(W) + W \bar{B}_n(W). \quad (3.25c)$$

根据文献<sup>[88]</sup>, 满足松弛带宽的最优策略  $\pi_R^*$  所对应的稳态概率分布  $\{\mu_{x,q}^{n,\star}, y_{x,q}^{n,\star}\}$  为定理3.1中拉格朗日乘子  $W = W^\star$  处的最优解, 其中  $W^\star$  的定义如下:

$$W^\star := \sup \{W \mid \sum_{n=1}^N \bar{B}_n(W) \leq M\}.$$

同时, 最优解  $\pi_R^*$  还要满足以下条件:

$$W^\star \left( \sum_{n=1}^N \bar{B}_n(W^\star) \right) = 0$$

我们用对偶乘子法寻找这样的  $W^\star$  和相应的解  $\{\mu_{x,q}^{n,\star}, y_{x,q}^{n,\star}\}$ 。记  $W^{(k)}$  为第  $k$  次迭代时使用的拉格朗日乘子, 根据文献<sup>[89]</sup>[式 6.1.1], 对偶函数  $g(W^{(k)})$  在  $W^{(k)}$  处

的一个次梯度为：

$$d_W g(W^{(k)}) = \sum_{n=1}^N \bar{B}_n(W^{(k)}) - M. \quad (3.26)$$

我们从  $W^{(0)} = 0$  开始，求解每个传感器  $n$  的线性规划问题3.4的最优解  $\{\mu_{x,q}^{n,W^{(0)}}, y_{x,q}^{n,W^{(0)}}\}$ ，并计算时间平均调度概率  $\bar{B}_n(W^{(0)})$ 。如果总网络占用带宽  $\sum_{n=1}^N \bar{B}_n(W^{(0)}) - M \leq 0$ ，说明这组稳态分布所对应的调度策略已经可以满足带宽约束，我们记  $\{\mu_{x,q}^{n,\star}, y_{x,q}^{n,\star}\} = \{\mu_{x,q}^{n,W^{(0)}}, y_{x,q}^{n,W^{(0)}}\}$  并终止迭代；否则我们用次梯度法寻找对偶乘子  $W^\star$ 。与文献<sup>[12]</sup>相似，我们选取  $\gamma_k = \frac{1}{k}$ ，并按照如下准则更新第  $k$  次迭代使用的拉格朗日乘子  $W^{(k)}$ ，其中次梯度  $d_W g(W^{(k-1)})$  的计算由式 (3.26) 给出：

$$W^{(k)} = W^{(k-1)} + \gamma_k d_W g(W^{(k-1)}). \quad (3.27)$$

迭代过程的终止条件为  $|W^{(k)} - W^{(k-1)}| < \varepsilon$  且  $\sum_{n=1}^N \bar{B}_n(W^{(k)}) \leq M$ 。如果停止时， $\sum_{n=1}^N \bar{B}_n(W^{(K)}) = M$ ，我们记  $\{\mu_{x,q}^{n,\star}, y_{x,q}^{n,\star}\} = \{\mu_{x,q}^{n,W^{(K)}}, y_{x,q}^{n,W^{(K)}}\}$ 。否则，根据求解线性规划问题的对偶单纯形法，我们首先从序列  $W^{(k)}$  寻找拉格朗日乘子  $W_l$  和  $W_u$ ，他们的定义如下：

$$W_l := \arg \max_{W^{(k)}} \sum_{n=1}^N \bar{B}_n(W^{(k)}), \text{ s.t. } \sum_{n=1}^N \bar{B}_n(W^{(k)}) \leq M, \quad (3.28a)$$

$$W_u := \arg \min_{W^{(k)}} \sum_{n=1}^N \bar{B}_n(W^{(k)}), \text{ s.t. } \sum_{n=1}^N \bar{B}_n(W^{(k)}) \geq M. \quad (3.28b)$$

我们将拉格朗日乘子  $W_l$  和  $W_u$  对应的线性规划最优解分别记为  $\{\mu^{n,W_l}, y^{n,W_l}\}$  和  $\{\mu^{n,W_u}, y^{n,W_u}\}$ ，这两个策略所使用的带宽分别记为  $M_l := \sum_{n=1}^N \bar{B}_n(W_l)$  和  $M_u := \sum_{n=1}^N \bar{B}_n(W_u)$ 。根据对偶单纯形法，最优策略  $\pi_R^\star$  所得到的稳态概率分布  $\{\mu^{n,\star}, y^{n,\star}\}$  为  $\{\mu^{n,W_l}, y^{n,W_l}\}$  和  $\{\mu^{n,W_u}, y^{n,W_u}\}$  的加权线性组合，可以通过如下式子计算：

$$\{\mu^{n,\star}, y^{n,\star}\} = \nu \{\mu^{n,W_l}, y^{n,W_l}\} + (1 - \nu) \{\mu^{n,W_u}, y^{n,W_u}\}, \quad (3.29)$$

其中系数  $\nu$  为。

$$\nu = \frac{M_u - M}{M_u - M_l}.$$

在得到  $\{\mu^{n,\star}, y^{n,\star}\}$  之后，依据解耦单传感器（问题3.3）的最优决策的阈值结构，我们可以计算出松弛问题3.2的最优决策  $\pi_R^\star$  在传感器  $n$  处于状态  $(x, q)$  时的调

度概率:

$$\xi_{x,q}^{n,\star} = \begin{cases} 1, & \xi_{x-1,q}^{n,\star} = 1 \text{ or } \mu_{x,q}^{n,\star} = 0 \text{ or } x \geq X_{\max}; \\ \frac{y_{x,q}^{n,\star}}{\mu_{x,q}^{n,\star}}, & \text{其他情况.} \end{cases} \quad (3.30)$$

搜索调度概率  $\xi_{x,q}^{n,\star}$  的全过程流程图如算法图3.1所示。

---

### 算法 3.1 时间平均带宽约束下的最优策略 $\pi_R^*$ 调度概率 $\xi_{x,q}^{n,*}$ 求解

---

**输入:** 节点数  $N$ , 带宽  $M$ , 信道转移矩阵  $\{p_{ij}^{(n)}\}$ , 功率约束  $\mathcal{E}_n$ 。  
**输出:** 策略  $\pi_R^*$  在传感器  $n$  处于状态  $(x, q)$  时的最优调度概率  $\xi_{x,q}^{n,\star}$ 。  
**初始化:**  $W^{(0)} \leftarrow 0, M_l \leftarrow 0, M_u \leftarrow 2M$ , 对于每个传感器  $n$ , 求解受限马尔可夫决策过程问题3.3对应的线性规划问题3.4的最优解  $\{\boldsymbol{\mu}^{n,W^{(0)}}, \mathbf{y}^{n,W^{(0)}}\}$  和占用的带宽  $\overline{B}_n(W^{(0)})$ 。  
**if**  $\sum_{n=1}^N \overline{B}_n(W^{(0)}) \leq M$  **then**  
     $\{\boldsymbol{\mu}^{n,\star}, \mathbf{y}^{n,\star}\} \leftarrow \{\boldsymbol{\mu}^{n,W^{(0)}}, \mathbf{y}^{n,W^{(0)}}\}$   
**else** {通过次梯度法迭代求解}  
    **repeat**  
         $k \leftarrow k + 1$   
        计算次梯度  $d_W g(W^{(k-1)}) = \sum_{n=1}^N \overline{B}_n(W^{(k-1)}) - M$   
        更新拉格朗日乘子  $W^{(k)} \leftarrow W^{(k-1)} + \gamma_k d_W g(W^{(k-1)})$   
        当拉格朗日乘子  $W = W^{(k)}$  时, 针对每个传感器  $n$  分别求解线性规划问题3.4并将其最优解记为  $\{\boldsymbol{\mu}^{n,W^{(k)}}, \mathbf{y}^{n,W^{(k)}}\}$ 。计算每个传感器  $n$  在对应策略下所需使用的调度带宽  $\overline{B}_n(W^{(k)})$   
        **if**  $M_l < \sum_{n=1}^N \overline{B}_n(W^{(k)}) \leq M$  **then**  
             $M_l \leftarrow \sum_{n=1}^N \overline{B}_n(W^{(k)}), W_l \leftarrow W^{(k)}$   
        **else if**  $M < \sum_{n=1}^N \overline{B}_n(W^{(k)}) \leq M_u$  **then**  
             $M_u \leftarrow \sum_{n=1}^N \overline{B}_n(W^{(k)}), W_u \leftarrow W^{(k)}$   
        **end if**  
        **until**  $|W^{(k)} - W^{(k-1)}| < \epsilon$  和  $\sum_{n=1}^N \overline{B}_n(W^{(k)}) \leq M$   
         $v \leftarrow \frac{M_u - M}{M_u - M_r}$   
         $\{\boldsymbol{\mu}^{n,\star}, \mathbf{y}^{n,\star}\} \leftarrow \lambda \{\boldsymbol{\mu}^{n,W_l}, \mathbf{y}^{n,W_l}\} + (1 - \lambda) \{\boldsymbol{\mu}^{n,W_u}, \mathbf{y}^{n,W_u}\}$   
**end if**  
    依据式 (3.30) 计算  $\{\xi_{x,q}^{n,\star}\}$ .

---

策略  $\pi_R^*$  即为: 在每时隙  $t$ , 中央控制器观察传感器  $n$  当前的信息年龄  $x_n(t)$  和信道状态  $q_n(t)$ , 以概率  $\xi_{x_n(t),q_n(t)}^{n,\star}$  调度传感器  $n$ 。

使用时间平均带宽约束下最优策略  $\pi_R^*$  得到的全网平均信息年龄  $J(\pi_R^*)$  可以通过稳态分布概率  $\{\boldsymbol{\mu}^{n,\star}, \mathbf{y}^{n,\star}\}$  直接计算。根据节3.3.1中的讨论, 策略  $\pi_R^*$  得到的全网平均信息年龄是原带宽、功率双约束问题3.1中最优策略平均信息年龄的下界, 也就是说, 原问题3.1的一个全网平均信息年龄下界  $\text{AoI}_{LB}$  可以通过如下式子计算:

$$\text{AoI}_{LB} = J(\pi_R^*) = \frac{1}{N} \sum_{n=1}^N \sum_{x=1}^{X_{\max}} \sum_{q=1}^Q x \mu_{x,q}^{n,\star}. \quad (3.31)$$

### 3.4.2 一种满足每时隙带宽严格约束的调度方案及其理论特性

本部分中，我们将基于时间平均带宽约束下信息年龄最小的调度策略  $\pi_R^*$ ，构造一个满足每时隙调度传感器数目都小于等于  $M$  的策略  $\tilde{\pi}$ 。我们把  $\pi_R^*$  选取的调度方案记作  $u_n^R(t)$ ，并记集合  $\Omega(t) := \{n | u_n^R(t) = 1\}$  为策略  $\pi_R^*$  在时隙  $t$  内选择调度的所有传感器。为了满足每时隙调度传感器数量小于  $M$  这一约束，我们按照如下准则进行调度方案  $u_n(t)$  的分配与设计：

- 如果策略  $\pi_R^*$  选取调度的传感器数目满足  $|\Omega(t)| \leq M$ ，策略  $\tilde{\pi}$  就将集合  $\Omega(t)$  中的传感器全部调度，即  $u_n(t) = 1, \forall u_n^R(t) = 1$ 。
- 如果策略  $\pi_R^*$  拟选取调度的传感器数目多于带宽  $M$ ，即  $|\Omega(t)| > M$ ，策略  $\tilde{\pi}$  随机选取集合  $\Omega(t)$  中一个大小不超过  $M$  的子集  $\mathcal{M}(t) \subset \Omega(t), |\mathcal{M}(t)| = M$ ，集合  $\Omega(t)$  中每个传感器被选取的概率是相同的。由于带宽的限制，未被选入集合  $\mathcal{M}(t)$  中的传感器，即集合  $\Omega(t)/\mathcal{M}(t)$  中的传感器将不被调度。

**定理 3.2：** 考虑网络中传感器  $N$  有相同的信道转移概率  $p_{i,j}^n$  和相同的能量限制  $\mathcal{E}_n$ <sup>①</sup>。如果策略  $\pi_R^*$  导出的马氏链  $(x, q)$  是遍历的<sup>②</sup>，当带宽  $M$  和传感器数量  $N$  的比值  $\frac{M}{N} = \theta$  保持不变时，随着  $N$  不断增大，使用策略  $\tilde{\pi}$  得到的全网平均信息年龄与式 (3.31) 中的理论下界之间的差距以  $\mathcal{O}(\frac{1}{\sqrt{N}})$  的速度衰减。也就是说，在大规模传感器网络  $N \rightarrow \infty$  中，当  $\frac{M}{N} = \theta$  保持不变时，调度方案  $\tilde{\pi}$  是问题3.1的渐近最优解。

**证明：** 定理证明详见附录节3.7.4。 ■

## 3.5 仿真结果

本节利用数值仿真验证所提调度策略的理论特性。在仿真中我们考虑一个  $Q = 4$  状态的信道，信道转移特性如矩阵  $\mathbf{P}$  所示，其中第  $i$  行第  $j$  个元素  $p_{ij}$  表示信道状态从  $i$  演进到  $j$  的概率：

$$\mathbf{P} = \begin{bmatrix} 0.4 & 0.3 & 0.2 & 0.1 \\ 0.25 & 0.3 & 0.25 & 0.2 \\ 0.2 & 0.25 & 0.3 & 0.25 \\ 0.1 & 0.2 & 0.3 & 0.4 \end{bmatrix}.$$

我们假设所有传感器的信道状态都相同，他们在每个信道状态下的稳态分布概率为  $\boldsymbol{\eta} = [0.2368, 0.2632, 0.2632, 0.2368]$ 。使用不同策略的全网平均信息年龄期望由  $T = 10^6$  连续时隙取平均得到。

① 这一结论可以被拓展到网络中的传感器被分为几组，每组中所有的传感器信道转移状态和能量限制相同。

② 这一遍历性要求较为容易满足，由信道转移的随机性很容易得到  $(X, Q)$  的非周期和常返性。

注意到文献<sup>[28]</sup>发现，当全网所有传感器信道状态相同时，每次选取信息年龄最大的传感器进行调度这一贪心方案是优化全网信息年龄的最优方案。实现这一贪心的调度方案需要每个传感器  $n$  的平均功率约束  $\mathcal{E}_n \geq \mathcal{E}^{\text{RR}} = \frac{M}{N} \sum_{q=1}^Q \eta_q \omega(q)$ 。在接下来的仿真和分析中，我们用  $\rho_n = \mathcal{E}_n / \mathcal{E}^{\text{RR}}$  度量传感器  $n$  的功率约束是否严苛，一个较小的  $\rho$  表明对应的传感器  $n$  所拥有的平均传输功率较小。

### 3.5.1 全网平均信息年龄表现

图3.4研究了固定带宽  $M \in \{2, 5\}$  两种情况下，传感器数目  $N$  不断增长时，使用所提调度方案和贪心法的全网平均信息年龄表现。其中第  $n$  个传感器的功率约束比值为  $\rho_n = 0.2 + \frac{1.4}{N-1}(n-1)$ 。我们对于文献<sup>[28]</sup>中的贪心法做一些改进使得它可以满足本问题中的平均传输功率约束。记  $C_n(t)$  为到时间  $t$  为止传感器  $n$  所消耗的发送功率，我们用集合  $\mathcal{R}(t) = \{n | \mathcal{E}_n t - C_n(t) \geq 0\}$  记录那些到时隙  $t$  为止没有违背平均功率约束的传感器。改进后的贪心法从集合  $\mathcal{R}(t)$  选取  $M$  个当前信息年龄最大的节点进行调度，如果  $\mathcal{R}(t)$  中的传感器数目小于  $M$ ，我们就将  $\mathcal{R}(t)$  中的传感器全部调度。通过仿真图可以发现，相比与贪心法，我们所提出的调度策略  $\tilde{\pi}$  得到的全网平均信息年龄更接近于全网平均信息年龄的理论下界  $\text{AoI}_{\text{LB}}$ 。随着网络中传感器数目  $N$  逐渐增多，当  $N = 50$  时，所提调度方案  $\tilde{\pi}$  相比于贪心法可以降低全网平均信息年龄约 40%。

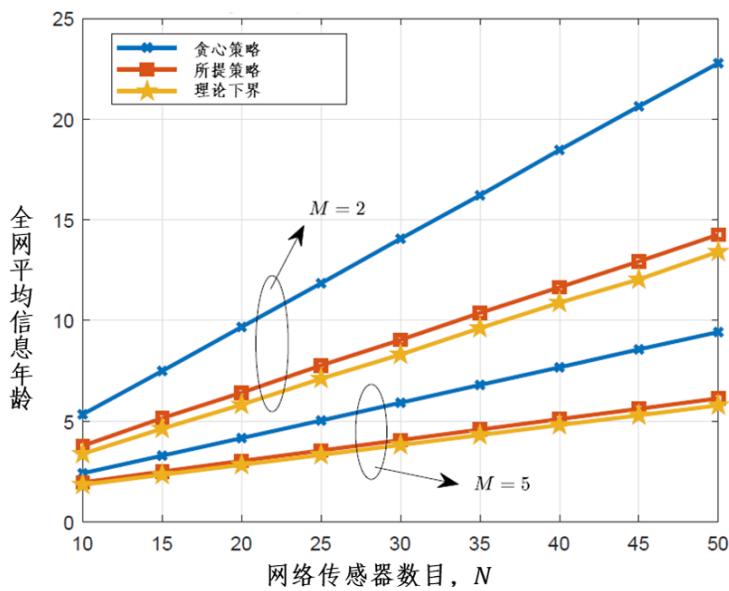


图 3.4 全网平均信息年龄随网络传感器数目  $N$  变化关系图

图3.5展示所提算法  $\tilde{\pi}$  的渐近最优性。我们保持  $\frac{M}{N} \in \{\frac{1}{5}, \frac{1}{8}\}$  固定为常数。每个传感器的功率约束为  $\rho_n = 0.2 + \frac{1.4}{N-1}(n-1)$ 。从图3.5中可以发现，随着网络中的传

感器数目  $N$  逐渐变多, 所提算法得到的全网平均信息年龄与理论下界  $\text{AoI}_{\text{LB}}$  之间的差距不断减小。

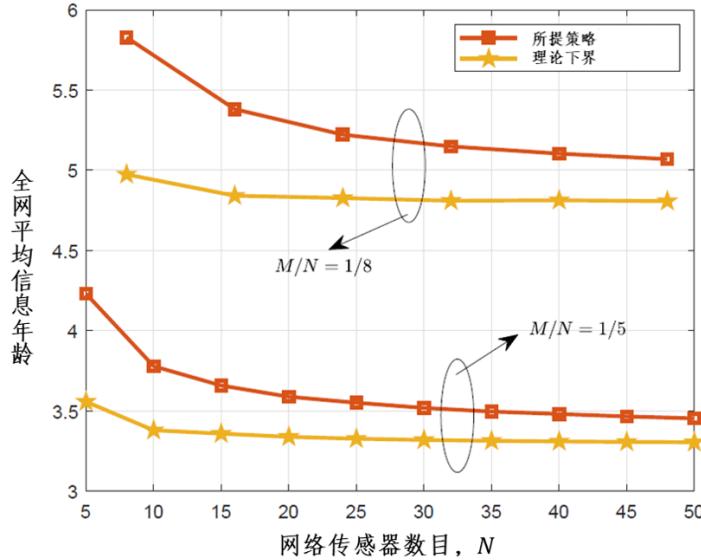


图 3.5 带宽  $M/N$  保持不变时, 全网平均信息年龄随传感器数目  $N$  的变化关系图

### 3.5.2 信息年龄与功率折中关系

图3.6研究了网络中传感器数目为  $N \in \{4, 8, 16\}$  时, 功率约束对全网平均信息年龄的影响。我们假设网络中的传感器有相同的信道衰落特性和功率约束因子  $\rho$ 。带宽约束为  $M = 1$ , 也就是说每时隙只有一个传感器可以给中央控制器发送信息。我们在图3.6中画出没有功率约束时的全网信息年龄理论下界 (绿色虚线)、不同功率约束  $\rho$  下全网信息年龄理论下界 (黄色实线) 和使用所提策略  $\tilde{\pi}$  得到的全网平均信息年龄 (红色方框)。仿真图再次验证了所提调度算法  $\tilde{\pi}$  得到的全网平均信息年龄与理论下界的值非常接近。随着功率约束不断放松 (即  $\rho$  变大), 全网信息年龄单调下降。当  $\rho \rightarrow 1$  时, 每个传感器都有足够的功率支持中央控制器进行轮询的调度。所提算法得到的全网平均信息年龄接近于理论下界; 当  $\rho \rightarrow 0$ , 即传感器所拥有的功率越来越少, 全网信息年龄剧烈上升。网络中的信息年龄与可用能量  $\{\mathcal{E}_n\}$  存在折中关系。

图3.8中, 我们研究了功率限制不同的传感器使用所提策略  $\tilde{\pi}$  的调度方案得到的平均信息年龄  $\bar{X}_n$ 。我们考虑一个  $N = 8$  的传感器网络, 每时隙有  $M = 2$  个传感器可以被调度, 每个传感器  $n$  的功率约束因子为  $\rho_n = 0.2n$ 。图3.7中, 对于功率限制比较强严苛的第一和第二个传感器 ( $\rho_1 = 0.2, \rho_2 = 0.4$ ), 所提调度方案  $\tilde{\pi}$  相比于贪心法可以帮助他们降低平均信息年龄约 40%; 对于那些传输功率较为充足的

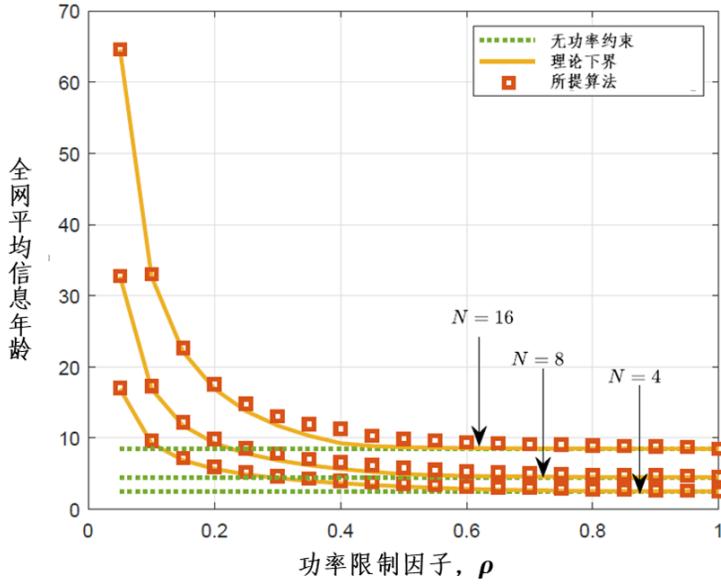
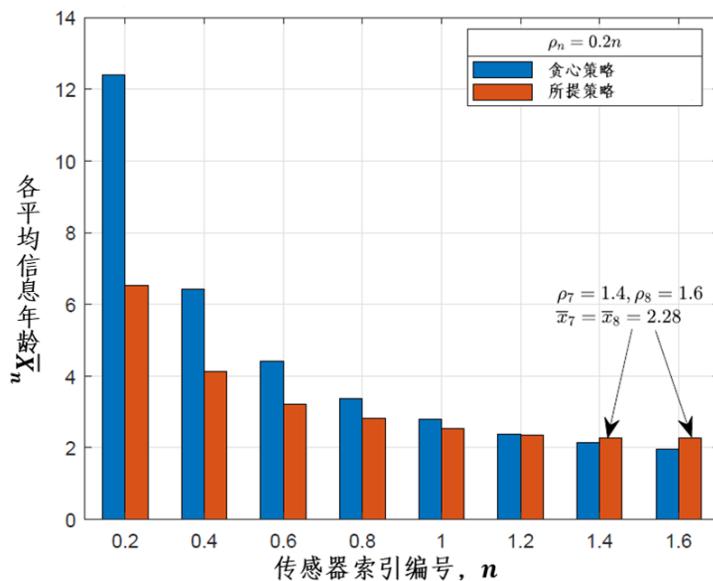


图 3.6 全网平均信息年龄与功率限制因子关系图

传感器（例如  $\rho_7 = 1.4, \rho_8 = 1.6$ ），所提算法  $\tilde{\pi}$  将保证将以一定频率规律地从这些传感器中采集状态信息，他们的平均信息年龄  $\bar{X}_n$  表现与贪心算法是相似的。总体来说，所提算法  $\tilde{\pi}$  之所以可以拥有较低的信息年龄表现，是因为他根据传感器的信道变化，对带宽和传输功率这两个资源都重新做了合理的分配。那些功率受限传感器的信息年龄下降尤其比较多。



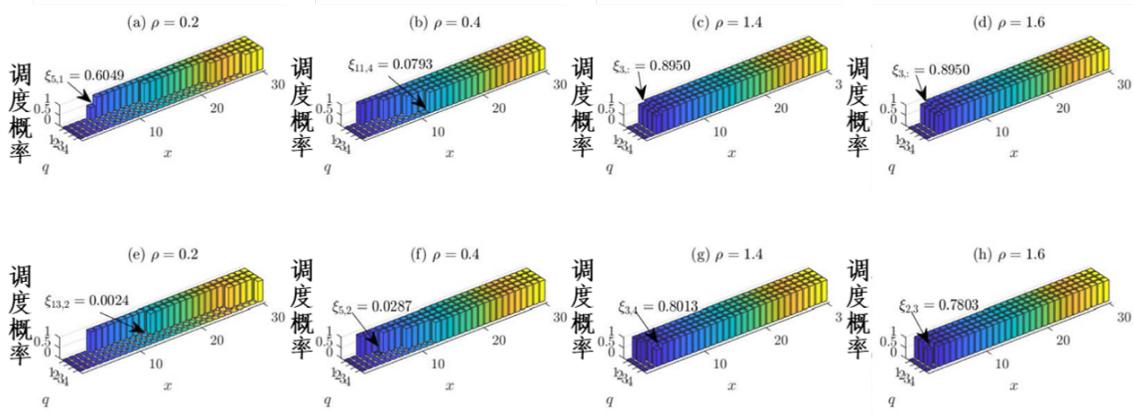
$N = 8, M = 2$  的多传感器网络中，每个传感器功率约束因子  $\rho_n = 0.2n$ 。

图 3.7 同一网络中不同功率传感器的平均信息年龄

注意到调度算法  $\tilde{\pi}$  是在策略  $\pi_R^*$  的基础上进行构建的,  $u_n^R(t)$  代表着传感器  $n$  在当前状态  $(x_n(t), q_n(t))$  向网络中心节点提供时效性强的状态信息的紧迫程度。为了更好地研究调度算法  $\tilde{\pi}$  的设计直觉, 我们在图3.8画出了不同编号的传感器处于不同状态  $(x, q)$  时, 策略  $\pi_R^*$  的调度概率。其中, 子图 (a)-(d) 表示的是传感器  $\{1, 2, 7, 8\}$  (功率限制因子为  $\rho = \{0.2, 0.4, 1.4, 1.6\}$ ) 使用  $\pi_R^*$  的调度概率  $\xi_{x,q}^{n,*}$ ; 我们在子图 (e)-(h) 中画出不考虑带宽约束  $M$  时, 传感器  $\{1, 2, 7, 8\}$  在不同状态下的调度概率。观察子图 (a) 和 (b) 可以发现, 当传输功率非常少, 越大的  $q$  代表较为糟糕的信道状态, 需要较大的功率  $\omega(q)$  来确保可靠传输, 因此传输功率的激活阈值  $\tau_q$  是信道状态  $q$  的一个上升序列。注意图3.8(b) 中, 每个信道状态的激活阈值都比图3.8(a) 的激活阈值小, 这意味着在有充足的功率下, 每个传感器将在信息年龄更小的时候被调度。在子图 (a) 和 (b) 中, 不同信道激活阈值  $\tau_q$  之间的差别比子图 (e) 和 (f) 中激活阈值之间的差别小, 这一阈值调整的来源是所提算法在满足功率的同时, 还需要满足全网传输带宽的限制, 所以算法设计需要尽可能地少占用带宽。总的来说, 对于功率受限的传感器, 调度策略的设计要充分利用信道条件较好的时刻进行机会性调度, 同时尽量少地利用带宽, 对于这些传感器来说, 如果当前信道状态较差, 中心控制器会选择等待, 直到传感器的信道变好或者信息年龄过大才会调度他们。图 (c)、(g), 图 (d)、(h) 中, 传感器都有比较多的传输功率 ( $\rho = \{1.4, 1.6\}$ ), 由图可见, 我们所提出的调度方案并没有把这些传输功率用尽, 同时所有的信道有相同的激活阈值。这一阈值的设定主要是为了满足网络中的带宽约束。相比于贪心法来说, 所提算法在功率和带宽分配方面更加合理, 对于那些功率约束比较严格的节点可以取得大幅度的信息年龄降低。总的来说, 对于网络中传感器的功率约束不同时, 所提调度策略有着不同的调度方案。对于那些功率限制比较严格的节点, 为了尽可能节约能量, 中央控制器寻找一些信道比较好的时刻进行调度; 对于平均传输功率比较多的传感器, 中央调度器将以一定频率更新和调度他们, 保证每个传感器都能规律地向中央控制器发送状态信息。

### 3.6 本章小结

本章考虑多传感器带宽、传输功率双受限网络中信息年龄最小的采样和调度问题。我们将每个传感器与中央控制器之间时变信道建模为遍历马尔可夫链, 不同信道状态下传感器将使用不同的发送功率确保可靠传输。我们将这个多传感器调度问题通过时间平均带宽约束松弛和拉格朗日乘子法, 解耦为单个传感器代价最小化问题。我们证明了单传感器代价最小化问题的最优解结构, 将其建模为受限马尔可夫决策过程, 并通过线性规划求解。通过选取合适的拉格朗日乘子, 我



(a)-(d):  $N = 8, M = 2$  网络中, 不同功率限制  $\rho_n$  的传感器调度策略; (e)-(h): 仅有功率限制, 无带宽限制网络中, 不同功率限制  $\rho_n$  的最优调度策略。

图 3.8 网络中不同功率约束节点在不同状态  $(x, q)$  下被调度概率

们首先得到了在时间平均带宽约束下的最优策略, 在此基础上我们提出了一个满足原每时隙带宽约束的随机调度策略。进一步地, 我们证明了所提策略在大规模网络中的渐近最优化。我们的研究发现, 当传感器用于发送数据的能量比较少时, 应充分挖掘信道的随机起伏, 在信道条件比较好和传感器信息年龄较大时进行调度; 那些功率供给充足的传感器在不同信道下将拥有固定的激活阈值, 以一定频率周期性地被调度。

## 3.7 附录

### 3.7.1 引理3.1证明

我们需要证明问题3.2存在一个最优策略  $\pi_R^*$ , 它可以写成  $N$  个独立的单传感器调度策略。为了证明这一引理, 我们引入变量  $\mathbf{b} = [b_1, \dots, b_N]^N$ , 其中每个元素  $b_n$  表示对于传感器  $n$  占用带宽的限制, 问题3.2可以写成如下等价形式:

**问题 3.5 (松弛问题3.2的等价问题):**

$$\pi_R^* = \arg \min_{0 \leq \mathbf{b} \leq 1, \mathbf{1}^T \mathbf{b} = M} J(\pi_E^*(\mathbf{b})),$$

$$\text{其中 } \pi_E^*(\mathbf{b}) \triangleq \min_{\pi \in \Pi_{NA}} \lim_{T \rightarrow \infty} \left\{ \frac{1}{NT} \mathbb{E}_{\pi} \left[ \sum_{t=1}^T \sum_{n=1}^N x_n(t) \right] \right\}, \quad (3.32a)$$

$$\text{s.t. } \lim_{T \rightarrow \infty} \mathbb{E}_{\pi} \left[ \frac{1}{T} \sum_{t=1}^T u_n(t) \right] \leq b_n, \quad (3.32b)$$

$$\lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E}_{\pi} \left[ \sum_{t=1}^T u_n(t) \omega(q_n(t)) \right] \leq \mathcal{E}_n, \forall n. \quad (3.32c)$$

根据文献<sup>[87]</sup>中公式(15), 针对每一组  $\{b_n\}_{n=1}^N$ , 策略  $\pi_E^\star(\mathbf{b})$  都可以被分解成  $N$  个独立的策略  $\pi_{E,n}^\star(b_n)$ , 其中  $\pi_{E,n}^\star(b_n)$  为如下优化问题最优解:

问题 3.6 (松弛问题3.2的等价问题):

$$\pi_{E,n}^\star(b_n) \triangleq \min_{\pi \in \Pi_{NA}} \lim_{T \rightarrow \infty} \left\{ \frac{1}{T} \mathbb{E}_\pi \left[ \sum_{t=1}^T x_n(t) \right] \right\}, \quad (3.33a)$$

$$\text{s.t. } \lim_{T \rightarrow \infty} \mathbb{E}_\pi \left[ \frac{1}{T} \sum_{t=1}^T u_n(t) \right] \leq b_n, \quad (3.33b)$$

$$\lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E}_\pi \left[ \sum_{t=1}^T u_n(t) \omega(q_n(t)) \right] \leq \mathcal{E}_n. \quad (3.33c)$$

记  $\mathbf{b}^\star = \arg \min_{0 \leq \mathbf{b} \leq 1, \mathbf{1}^T \mathbf{b} = M} J(\pi_E^\star(\mathbf{b}))$ 。那么所要寻找的最优策略即为  $\pi_R^\star = \pi_E^\star(\mathbf{b}^\star) = \otimes_{n=1}^N \pi_{E,n}^\star(b_n^\star)$ , 也就是说每时隙  $t$ , 对于传感器  $n$  的调度策略只由策略  $\pi_{E,n}^\star(b_n^\star)$  依据传感器  $n$  当前的信道状态  $q_n(t)$  和信息年龄  $x_n(t)$  决定。

### 3.7.2 引理3.2证明

**证明:** 我们考虑一个固定的拉格朗日乘子  $\lambda \geq 0$ 。我们首先研究无穷时间  $\alpha$  折扣代价最小化的阈值结构, 从而证明最小化时间平均代价式(3.12), 即策略  $\pi_d^\star(\lambda)$  的阈值结构。记折扣因子为  $0 < \alpha < 1$  时, 从状态  $(x, q)$  出发, 使用静态策略  $\pi$  得到的无穷长时间内  $\alpha$  折扣代价  $J_{\alpha,\pi}(x, q)$  为:

$$J_{\alpha,\pi}(x, q) \triangleq \lim_{T \rightarrow \infty} \mathbb{E}_\pi \left[ \sum_{t=0}^T \alpha^t (C_X(x(t), q(t), s(t)) + \lambda C_Q(x(t), q(t), s(t))) \mid x(0) = x, q(0) = q \right]. \quad (3.34)$$

记  $V_\alpha(x, q) := \min_{\pi \in \Pi} J_{\alpha,\pi}(x, q)$  为从状态  $(x, q)$  出发最小的  $\alpha$  累计折扣代价。由马尔可夫决策过程的性质, 函数  $V_\alpha(x, q)$  满足以下贝尔曼方程:

$$V_\alpha(x, q) = \min \{ C_X(x, q, 0) + \alpha \sum_{q'=1}^Q p_{q,q'} V_\alpha(x+1, q'), \\ C_X(x, q, 1) + \lambda C_Q(x, q, 1) + \alpha \sum_{q'=1}^Q p_{q,q'} V_\alpha(1, q') \}. \quad (3.35)$$

$\alpha$  折扣代价最小化策略的阈值结构将通过以上贝尔曼方程证明, 我们首先在如下引理中给出  $V_\alpha(x, q)$  的性质:

**引理 3.3:** 对于给定折扣因子  $\alpha$  和信道状态  $q$ , 值函数  $V_\alpha(\cdot, q)$  随着  $x$  递增。

以上引理证明详见节3.7.3。接下来我们利用这个引理证明最优策略的阈值结

构。记  $\Delta(x, q)$  为使用不同策略  $u = \{0, 1\}$  时值函数的差别, 即:

$$\begin{aligned}\Delta(x, q) &:= \left( C_X(x, q, 0) + \alpha \sum_{q'=1}^Q p_{q,q'} V_\alpha(x+1, q') \right) \\ &\quad - \left( C_X(x, q, 1) - \lambda C_Q(x, q, 1) - \alpha \sum_{q'=1}^Q p_{q,q'} V_\alpha(1, q') \right) \\ &= \alpha \sum_{q'=1}^Q p_{q,q'} (V_\alpha(x+1, q') - V_\alpha(1, q')) - (\lambda \omega(q) + W).\end{aligned}\quad (3.36)$$

我们注意到最优的策略  $u_\alpha^\star(x, q)$  在状态  $(x, q)$  时选取的策略由值函数  $V_\alpha(x, q)$  和以上贝尔曼方程决定。也就是说, 如果最优策略满足  $u_\alpha^\star(x, q) = 1$ , 即在状态  $(x, q)$  时调度传感器最优, 那么将式 (3.11a) 中的单步代价函数带入  $\Delta(x, q) \geq 0$ , 我们可以得到以下不等式:

$$\alpha \sum_{q'=1}^Q p_{q,q'} (V_\alpha(x+1, q') - V_\alpha(1, q')) - (\lambda \omega(q) + W) \geq 0.\quad (3.37)$$

根据引理3.3, 给定信道状态  $q$ , 值函数  $V_\alpha(\cdot, q)$  单调递增。因此, 对于任意状态  $x' > x$ ,  $\Delta(x', q)$  的下界为:

$$\begin{aligned}\Delta(x', q) &= \alpha \sum_{q'=1}^Q p_{q,q'} (V_\alpha(x'+1, q') - V_\alpha(1, q')) - (\lambda \omega(q) + W) \\ &\stackrel{(a)}{>} \alpha \sum_{q'=1}^Q p_{q,q'} (V_\alpha(x+1, q') - V_\alpha(1, q')) - (\lambda \omega(q) + W) \geq 0,\end{aligned}\quad (3.38)$$

其中不等式 (a) 由  $V(\cdot, q)$  的递增性得到。由于对于满足  $x' > x$  的状态  $x'$ , 差值  $\Delta(x', q)$  总是正的, 因此最优策略  $u_\alpha^\star(x, q) = 1$ , 即选择调度传感器;

同样地, 如果在状态  $(x, q)$  处最优决策为不调度, 即  $u_\alpha^\star(x, q) = 0$ , 那么对于状态  $x' < x$ , 最优的策略  $u_\alpha^\star(x', q) = 0$  选择空闲;

如果存在状态  $\Delta(x, q) = 0$ , 则最优策略可以以任意概率  $p \in [0, 1]$  调度传感器。根据以上的分析, 对于任意  $x' > x$ , 根据式 (3.38), 则可以证明  $\Delta(x', q) > 0$ ; 同样对于  $x' < x$ , 可以证明  $\Delta(x', q) < 0$ 。因此, 对于信道状态  $q$ , 这样的调度和不调度都为最优的信息年龄  $x$  只有一个。

接下来, 我们证明对于任意的  $\lambda$ , 激活阈值小于无穷。注意到对于任意的信道状态  $q$ , 依据贝尔曼方程, 在状态为  $(x, q)$  时, 采用不调度  $u = 0$  和调度  $u = 1$  带来

的无穷长时间  $\alpha$  折扣代价之差计算如下：

$$\begin{aligned}
 & (C_X(x, q, 0) + \alpha \mathbb{E}[V_\alpha(x+1, q')]) - (C_X(x, q, 1) + \alpha \mathbb{E}[V_\alpha(1, q')]) \\
 & \geq x + \alpha(x+1) - (x + \alpha \mathbb{E}[V_\alpha(1, q')] + W + \lambda \omega(q)) \\
 & = \alpha x + \alpha - \alpha \mathbb{E}[V_\alpha(1, q')] - W - \lambda \omega(q),
 \end{aligned} \tag{3.39}$$

可见以上差距随着  $x$  增大线性增长。因此，存在一个小于无穷的阈值使得不等式 (3.37) 可以被满足，也就是说最优的调度方案不会一直不调度传感器。这说明，对于信道状态  $q$ ，存在一系列阈值  $\tau_q$ ，对于信息年龄大于阈值  $x > \tau_q$  的状态，最优策略选择调度  $u_\alpha^\star(x, q) = 1$ ；对于小于阈值的状态  $x < \tau_q$ ，最优策略选择空闲  $u_\alpha^\star(x, q) = 0$ 。

最后，我们从无穷时间  $\alpha$  折扣代价的最优策略推导最小化时间平均代价的最优策略。我们取一系列最终收敛到 1 的折扣因子  $\lim_{k \rightarrow \infty} \alpha_k = 1$ ，根据文献<sup>[90]</sup>，最小化无穷时间  $\alpha_k$  折扣代价的策略  $u_{\alpha_k}^\star$  收敛到最小化时间平均代价的策略，引理3.2中的阈值得证。 ■

### 3.7.3 引理3.3证明

**证明：** 本部分我们将证明  $\alpha$  折扣代价最小化问题的值函数  $V_\alpha(x, q)$  单调上升。值函数可以通过贝尔曼方程式 (3.35) 迭代计算。记  $V_\alpha^{(k)}(x, q)$  为第  $k$  次迭代后得到的值函数，我们通过归纳法证明值函数单调性。

假设第  $k$  次迭代得到的函数  $V_\alpha^{(k)}(\cdot, q)$  是单调递增的。不失一般的，考虑两状态  $x_1, x_2$  满足  $x_1 < x_2$ 。根据受限马尔可夫决策过程一步代价函数，对于任意的决策  $u \in \mathcal{A}$ ，我们有：

$$C_X(x_1, q, u) < C_X(x_2, q, u), C_Q(x_1, q, u) = C_Q(x_2, q, u). \tag{3.40}$$

记  $J_{\alpha, u}^{(k)}(x, q) := C_X(x, q, u) + \alpha \mathbb{E}_{x', q' | x, q, u} [V_\alpha^{(k)}(x', q')]$ 。我们有如下不等式：

$$\begin{aligned}
 & J_{\alpha, 0}^{(k)}(x_1, q) \\
 & = C_X(x_1, q, 0) + \alpha \sum_{q'=1}^Q p_{q, q'} V_\alpha^{(k)}(x_1 + 1, q) \\
 & \stackrel{(a)}{<} C_X(x_2, q, 0) + \alpha \sum_{q'=1}^Q p_{q, q'} V_\alpha^{(k)}(x_2 + 1, q) \\
 & = J_{\alpha, 0}^{(k)}(x_2, q),
 \end{aligned} \tag{3.41}$$

其中不等式 (a) 来自于函数  $V_\alpha^{(k)}(\cdot, q)$  的单调性。同样地，我们可以得到  $J_{\alpha, 1}^{(k)}(x_1, q) < J_{\alpha, 1}^{(k)}(x_2, q)$ ，即函数  $J_{\alpha, 1}^{(k)}(\cdot, q)$  同样具有单调递增性质。对于所有的决策  $u \in \mathcal{A}$  取最

小值，可以得到  $J_{\alpha,s}^{(k)}(x_1, q) < J_{\alpha,s}^{(k)}(x_2, q)$ 。

注意到第  $(k+1)$  次迭代得到的值函数可以通过如下公式写出：

$$V_{\alpha}^{(k+1)}(x, q) = \min_s J_{\alpha,s}^{(k)}(x, q).$$

因此，我们得到  $V_{\alpha}^{(k+1)}(x_1, q) < V_{\alpha}^{(k+1)}(x_2, q)$ 。令迭代次数  $k \rightarrow \infty$ ，根据值迭代特性，我们有  $V_{\alpha}^{(k)}(x, q) \rightarrow V_{\alpha}(x, q)$ 。因此，值函数  $V_{\alpha}(x, q)$  是单调增的。 ■

### 3.7.4 定理3.2证明

**证明：**注意到使用策略  $\pi_R^*$  得到的全网平均信息年龄  $J(\pi_R^*)$  是原问题3.1的理论下界。我们通过比较所提策略全网平均信息年龄  $J(\tilde{\pi})$  与  $J(\pi_R^*)$  之间的差证明定理3.2。

首先，注意到策略  $\pi_R^*$  满足时间平均带宽约束，记  $\bar{\Omega}(t) := \mathbb{E}[|\Omega(t)|]$ ，采用策略  $\pi_R^*$  调度的平均传感器数目满足以下不等式：

$$\bar{\Omega} := \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \bar{\Omega}(t) = \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[ \sum_{t=1}^T |\Omega(t)| \right] \leq M. \quad (3.42)$$

根据引理3.2，每一个解耦后的单传感器代价最小化策略都有阈值结构。记  $\Gamma_n = \max_q \tau_{n,q} - \min_q \tau_{n,q}$  为传感器  $n$  的所有调度阈值中最大的差距。假设在时隙  $t$  中，策略  $\pi_R^*$  选择策略  $u_n^R(t) = 1$ ，单传感器  $n$  最终没有被策略  $\tilde{\pi}$  调度。在下一时隙  $t+1$  传感器  $n$  仍未被调度有以下两种可能：（1）传感器  $n$  在下一时刻的信道状态  $q'$  对应的阈值  $\tau_{n,q'}$  大于时刻  $t$  的阈值  $\tau_{n,q}$ ，我们假设这一事件发生的概率是  $p$ ；（2）在下一时隙，仍然有多余  $M$  个传感器被策略  $\pi_R^*$  选择进行调度，而传感器  $n$  再次在随机选择中被策略  $\tilde{\pi}$  丢弃，这一概率小于  $\frac{N-M}{N}$ 。因此，如果传感器  $n$  满足  $u_n^R(t) = 1$  但是最终没有被策略  $\tilde{\pi}$  调度，他在下一时刻仍然未被调度的概率小于：

$$p + (1-p) \frac{N-M}{N} = \frac{N-M}{N} + \frac{M}{N} p. \quad (3.43)$$

事件  $\tau_{n,q'} > \tau_{n,q}$  发生的概率  $p$  小于信道离开当前状态  $q$  的概率，即  $p \leq 1 - \min_{n,q} p_{q,q}^{(n)}$ 。因此，对于任意一个传感器  $n$ ，在第  $t$  时隙满足  $u_n^R(t) = 1, u_n(t) = 0$  的情况下，它在时隙  $t+1$  时仍然未被策略  $\tilde{\pi}$  调度的概率小于一个常数  $z$ ：

$$z := \frac{N-M}{N} + \frac{M}{N} \left( 1 - \min_{n,q} P_{q,q}^{(n)} \right). \quad (3.44)$$

同理，该传感器  $n$  在接下来的  $t'$  个时隙中都未被调度的概率一定小于  $z^{(t'-\Gamma_n)^+}$ ，其中  $(\cdot)^+ = \max\{\cdot, 0\}$ 。如果时隙  $t$  传感器已经被调度，在时隙  $t+t'$  他的信息年龄将为  $t'$ ，而直到时隙  $t+t'$  仍未调度传感器  $n$  将导致它的信息年龄变成  $x_n(t) + t'$ ，因此，连续  $t'$  个时隙都没调度传感器  $n$  这一事件将为全网带来信息年龄  $x_n(t)$  的增

长。

接下来, 我们研究每次随机选择  $M$  个传感器进行调度这一策略相比于  $\pi_R^*$  带来的全网平均信息年龄增长。我们通过提出一种新的调度策略  $\hat{\pi}_R^*$  进行信息年龄增量的研究, 策略  $\hat{\pi}_R^*$  如下: 如果策略  $\pi_R^*$  选择调度的传感器数目多于带宽, 即  $|\Omega(t)| > M$ , 新策略  $\hat{\pi}_R^*$  将集合  $\Omega(t)$  中的传感器全部调度, 但在计算总的信息年龄时, 对于每个  $n \in \Omega(t)/\mathcal{M}(t)$  的传感器, 额外加上代价  $\sum_{t'=0}^{\infty} z^{(t'-\Gamma_n)^+} x_n(t) = (\Gamma_n + \frac{1}{1-z})x_n(t)$ 。由于  $M$  个传感器是随机被抽取的, 这一额外代价的期望可以通过如下公式计算:

$$\begin{aligned} & \mathbb{E}_{|\Omega(t)| > M} \sum_{n=1}^N (\Gamma_n + \frac{1}{1-z})x_n(t) \frac{|\Omega(t)| - M}{|\Omega(t)|} \\ & \leq \mathbb{E}_{|\Omega(t)| > M} \sum_{n=1}^N (\Gamma_n + \frac{1}{1-z})x_n(t) \frac{|\Omega(t)| - M}{M} \end{aligned} \quad (3.45)$$

我们注意到, 新策略  $\hat{\pi}_R^*$  带来的全网平均信息年龄一定大于  $\tilde{\pi}$ 。因此, 策略  $J(\tilde{\pi})$  与策略  $\pi_R^*$  之间信息年龄的差距有如下上界:

$$\begin{aligned} & (J(\tilde{\pi}) - J(\pi_R^*)) \\ & \leq (J(\hat{\pi}_R^*) - J(\pi_R^*)) \\ & = \lim_{T \rightarrow \infty} \frac{1}{NT} \mathbb{E}_{\pi_R^*} \left[ \sum_{t=1}^T \mathbb{E}_{|\Omega(t)| > M} \left( \sum_{n=1}^N (\Gamma_n + \frac{1}{1-z})x_n(t) \times \frac{|\Omega(t)| - M}{M} \right) \right] \\ & = \lim_{T \rightarrow \infty} \frac{1}{NT} \mathbb{E}_{\pi_R^*} \left[ \sum_{t=1}^T \left( \sum_{n=1}^N (\Gamma_n + \frac{1}{1-z})x_n(t) \frac{(|\Omega(t)| - \bar{\Omega}(t) + \bar{\Omega}(t) - M)^+}{M} \right) \right] \\ & \leq \lim_{T \rightarrow \infty} \frac{\max_n \Gamma_n + \frac{1}{1-z}}{MNT} \mathbb{E}_{\pi_R^*} \left[ \sum_{t=1}^T \sum_{n=1}^N x_n(t) \left( (|\Omega(t)| - \bar{\Omega}(t))^+ + (\bar{\Omega}(t) - M)^+ \right) \right] \\ & \stackrel{(a)}{\leq} \lim_{T \rightarrow \infty} \frac{\max_n \Gamma_n + \frac{1}{1-z}}{MNT} \mathbb{E}_{\pi_R^*} \left[ \sum_{t=1}^T \sum_{n=1}^N x_n(t) \left( |\Omega(t)| - \bar{\Omega}(t) + (\bar{\Omega}(t) - M)^+ \right) \right] \\ & \stackrel{(b)}{\leq} \lim_{T \rightarrow \infty} \frac{\max_n \Gamma_n + \frac{1}{1-z}}{MNT} \mathbb{E}_{\pi_R^*} \left[ \sum_{t=1}^T \sum_{n=1}^N \max_q \tau_{n,q} \left( |\Omega(t)| - \bar{\Omega}(t) + (\bar{\Omega}(t) - M)^+ \right) \right] \\ & \stackrel{(c)}{=} \lim_{T \rightarrow \infty} \frac{(\max_n \Gamma_n + \frac{1}{1-z}) \sum_{n=1}^N \max_q \tau_{n,q}}{\theta N^2} \times \\ & \quad \left( \lim_{T \rightarrow \infty} \mathbb{E}_{\pi_R^*} \left[ \frac{1}{T} \sum_{t=1}^T |\Omega(t)| - \bar{\Omega}(t) \right] + \lim_{T \rightarrow \infty} \mathbb{E}_{\pi_R^*} \left[ \frac{1}{T} \sum_{t=1}^T |(\bar{\Omega}(t) - M)^+| \right] \right), \end{aligned} \quad (3.46)$$

其中不等式 (a) 是因为  $(\cdot)^+ \leq |\cdot|$ ; 不等式 (b) 是因为使用策略  $\pi_R^*$  时, 每个解耦的

传感器都有一系列调度阈值，因此任意传感器的信息年龄  $x_n(t)$  都不会超过最大的调度阈值  $\max_q \tau_{n,q}$ ；等式(c)是因为  $M = N\theta$ 。

接下来我们通过下面两个引理分别计算式(3.46)中最后两项的上界。这两个引理的证明分别在节3.7.5和节3.7.6中。

**引理3.4：**

$$\lim_{T \rightarrow \infty} \mathbb{E}_{\pi_R^*} \left[ \frac{1}{T} \sum_{t=1}^T (\bar{\Omega}(t) - M)^+ \right] = 0. \quad (3.47)$$

**引理3.5：**

$$\lim_{T \rightarrow \infty} \mathbb{E}_{\pi_R^*} \left[ \frac{1}{T} \sum_{t=1}^T |\Omega(t) - \bar{\Omega}(t)| \right] \leq \sqrt{N}. \quad (3.48)$$

注意到对于给定的功率限制条件  $\mathcal{E}_n$ ，调度阈值之差  $\Gamma_n$  与网络中的传感器数目无关  $N$ 。同时，根据式(3.44)，常数  $z$  的取值与网络中节点数目也无关。当网络中  $\frac{N}{M} = \theta$  保持为常数，可用的带宽资源  $M$  随传感器数目  $N$  线性增长，调度阈值  $\max_q \tau_{n,q}$  也不会随着  $N$  增长，因此  $\Gamma_n$  保持不变。

结合式(3.46)-(3.48)，我们可以得到  $J(\tilde{\pi})$  与  $J(\pi_R^*)$  的上界：

$$J(\tilde{\pi}) - J(\pi_R^*) = \mathcal{O}\left(\frac{1}{\sqrt{N}}\right). \quad (3.49)$$

我们注意到策略  $\pi^*$  得到的平均信息年龄  $J(\pi_R^*)$  一定大于无功率约束时轮询策略  $\pi^{RR}$  得到的全网平均信息年龄，即  $J(\pi^{RR}) = \frac{1}{2}(\frac{N}{M} + 1)$ 。因此，当我们保持  $N/M = 1/\theta$  为常数，我们可以得到  $J(\pi_R^*)$  的下界：

$$J(\pi_R^*) \geq J(\pi^{RR}) = \frac{1}{2}(\frac{1}{\theta} + 1). \quad (3.50)$$

结合以上不等式，我们最终可以证明策略  $\tilde{\pi}$  在大规模网络中的渐近最优性：

$$\frac{J(\tilde{\pi}) - J(\pi_R^*)}{J(\pi_R^*)} = \mathcal{O}\left(\frac{1}{\sqrt{N}}\right). \quad (3.51)$$

即

$$\lim_{N \rightarrow \infty} \frac{J(\tilde{\pi}) - J(\pi_R^*)}{J(\pi_R^*)} = 0. \quad (3.52)$$

■

### 3.7.5 引理3.4证明

**证明：**注意到使用策略  $\pi_R^*$  时，每个传感器的信息年龄和信道状态  $(x_n(t), q_n(t))$  可以用马尔可夫链描述。如果这些马尔可夫链都是遍历的，传感器  $n$  在状态  $(x, q)$  的稳态分布为  $\mu_{x,r}^{n,*}$ 。由马尔可夫链稳态分布的性质，对于任意的  $\epsilon > 0$ ，都有阈值  $T_\epsilon$ ，

使得对任意  $t > T_\epsilon$  我们有:

$$\left| \Pr(x_n(t) = x, q_n(t) = q) - \mu_{x,r}^{n,\star} \right| \leq \epsilon, \forall 1 \leq x \leq \max_{n,q} \tau_{n,q}, \quad (3.53)$$

而对于状态  $x > \max_{n,q} \tau_{n,q}$ , 由阈值结构, 他们的稳态分布为 0, 即

$$\Pr(x_n(t) = x, q_n(t) = q) = 0. \quad (3.54)$$

将传感器  $n$  在时刻  $t$  被策略  $\pi_R^\star$  调度的概率记为  $\Pr(u_n^R(t) = 1)$ 。对于所有满足  $\forall t > T_\epsilon$  的时隙, 我们有:

$$\begin{aligned} & \Pr(u_n^R(t) = 1) \\ &= \sum_{q=1}^Q \sum_{x=1}^{\infty} \Pr(x_n(t) = x, q_n(t) = q) \Pr(u_n^R(t) = 1 | x_n(t) = x, q_n(t) = q) \\ &\stackrel{(a)}{=} \sum_{r=1}^R \sum_{x=1}^{\max_{n,q} \tau_{n,q}} \Pr(x_n(t) = x, q_n(t) = q) \xi_{x,q}^{n,\star} \\ &= \sum_{q=1}^Q \sum_{x=1}^{\max_{n,q} \tau_{n,q}} \mu_{x,q}^{n,\star} \xi_{x,q}^{n,\star} + \sum_{q=1}^Q \sum_{x=1}^{\max_{n,q} \tau_{n,q}} \left( \Pr(x_n(t) = x, q_n(t) = q) - \mu_{x,q}^{n,\star} \right) \xi_{x,q}^{n,\star}. \end{aligned} \quad (3.55)$$

其中等式 (a) 由式 (3.54) 得到, 而  $\xi_{x,q}^{n,\star}$  为使用策略  $\pi_R^\star$  时调度概率, 在状态  $(x, q)$  中调度传感器  $n$  的概率。

式 (3.55) 说明, 对于任意  $t \geq T_\epsilon$ , 我们由如下不等式

$$\begin{aligned} & \left| \Pr(u_n^R(t) = 1) - \left( \sum_{q=1}^Q \sum_{x=1}^{\max_{n,q} \tau_{n,q}} \mu_{x,r}^{n,\star} \xi_{x,r}^{n,\star} \right) \right| \\ &= \left| \sum_{q=1}^Q \sum_{x=1}^{\max_{n,q} \tau_{n,q}} \left( \Pr(x_n(t) = x, q_n(t) = q) - \mu_{x,q}^{n,\star} \right) \xi_{x,q}^{n,\star} \right| \\ &\stackrel{(a)}{\leq} \epsilon \max_{n,q} \tau_{n,q}, \end{aligned} \quad (3.56)$$

其中不等式 (a) 来源于式 (3.53)。

注意到每时隙调度传感器数目的平均值为:

$$\overline{\Omega}(t) = \sum_{n=1}^N \mathbb{E}_{\pi_R^\star}[u_n^R(t)] = \sum_{n=1}^N \Pr(u_n^R(t)).$$

同时, 策略  $\pi_R^\star$  可以满足时间平均带宽约束, 我们有如下不等式:

$$\sum_{n=1}^N \sum_{q=1}^Q \sum_{x=1}^{\max_{n,q} \tau_{n,q}} \mu_{x,r}^{n,\star} \xi_{x,q}^{n,\star} = \overline{\Omega} \leq M. \quad (3.57)$$

因此, 对于  $t > T_\varepsilon$ , 我们可以得到  $\bar{\Omega}(t) - M$  的一个上界:

$$\begin{aligned}
& \left( \bar{\Omega}(t) - M \right)^+ \leq \left| \bar{\Omega}(t) - \bar{\Omega} \right| \\
& \leq \left| \sum_{n=1}^N \left( \Pr(u_n^R(t) = 1) - \sum_{q=1}^Q \sum_{x=1}^{\max_{n,q} \tau_{n,q}} \mu_{x,r}^{n,*} \xi_{x,r}^{n,*} \right) \right| \\
& \leq \sum_{n=1}^N \left| \Pr(u_n^R = 1) - \sum_{q=1}^Q \sum_{x=1}^{\max_{n,q} \tau_{n,q}} \mu_{x,r}^{n,*} \xi_{x,r}^{n,*} \right| \\
& \stackrel{(b)}{\leq} N \varepsilon Q \max_{n,q} \tau_{n,q}. \tag{3.58}
\end{aligned}$$

其中不等式 (b) 由式 (3.56) 直接得到。将式 (3.58) 从时间  $1 \sim T$  加和, 我们可以得到以下不等式:

$$\begin{aligned}
& \lim_{T \rightarrow \infty} \mathbb{E}_{\pi_R^\star} \left[ \frac{1}{T} \sum_{t=1}^T (\bar{\Omega}(t) - M)^+ \right] \\
& = \lim_{T \rightarrow \infty} \mathbb{E}_{\pi_R^\star} \left[ \frac{1}{T} \left( \sum_{t=1}^{T_\varepsilon} (\bar{\Omega}(t) - M)^+ + \sum_{t=T_\varepsilon+1}^T (\bar{\Omega}(t) - M)^+ \right) \right] \\
& \leq \lim_{T \rightarrow \infty} \left( \frac{T_\varepsilon N}{T} + \frac{T - T_\varepsilon}{T} \varepsilon N Q \max_{n,q} \tau_{n,q} \right) \\
& = \varepsilon N Q \max_{n,q} \tau_{n,q}. \tag{3.59}
\end{aligned}$$

其中最后一步由 Cesaro 求和公式得到。

选取足够小的  $\varepsilon$  即可完成引理3.4的证明。 ■

### 3.7.6 引理3.5证明

**证明:** 首先我们有以下不等式:

$$\mathbb{E}_{\pi_R^\star} \left[ \left| |\Omega(t)| - \bar{\Omega}(t) \right| \right] \stackrel{(a)}{\leq} \sqrt{\mathbb{E}_{\bar{\pi}_R^\star} \left[ \left| |\Omega(t)| - \bar{\Omega}(t) \right|^2 \right]} \stackrel{(b)}{=} \sqrt{\text{var}[|\Omega(t)|]}. \tag{3.60}$$

其中不等式 (a) 由 Jensen 不等式得到。等式 (b) 式源于定义  $\bar{\Omega}(t) = \mathbb{E}_{\pi_R^\star}[|\Omega(t)|]$ 。

我们注意到  $|\Omega(t)| = \sum_{n=1}^N u_n^R(t)$ 。根据引理3.1中策略  $\pi_R^\star$  的解耦特性,  $u_n^R(t)$  的取值仅依赖于传感器  $n$  当前状态  $(x_n(t), q_n(t))$ 。也就是说, 调度决策随机变量  $\{u_n^R(t)\}_n$  彼此独立。因此

$$\text{var}[|\Omega(t)|] = \sum_{n=1}^N \text{var}[u_n^R(t)] \stackrel{(c)}{\leq} N. \tag{3.61}$$

其中不等式 (a) 是因为对于所有传感器  $n$ , 决策变量  $u_n^R(t) \in \{0, 1\}$ , 因此我们有

$$u_n^R(t) \leq 1.$$

将式(3.61)带回到式(3.60)可以得到引理中的结论:

$$\mathbb{E}_{\pi_R^\star} \left[ \left| |\Omega(t)| - \overline{\Omega}(t) \right| \right] \leq \sqrt{N}. \quad (3.62)$$

■

## 第4章 信息年龄最小化的基站端缓存更新算法研究

### 4.1 引言

车联网中的用户通常会从基站端缓存中索取不断变化的路况信息、地图等状态信息副本。本章研究车联网中基站端缓存更新策略。如节1.2所述，过去基站端缓存系统中，文件副本更新策略设计旨在优化用户索取副本时的命中率，针对用户索取状态信息副本的时效性优化尚处于起步阶段。

本章首先针对状态信息流行度随时间变化的场景，设计了基站端缓存中状态信息副本的更新策略，优化用户索取得到副本的信息年龄；进一步地，论文工作考虑了状态信息更新时长随时间变化的场景，利用下载信息时长随信息年龄变化的特点，设计了满足带宽约束的基站端缓存更新策略。

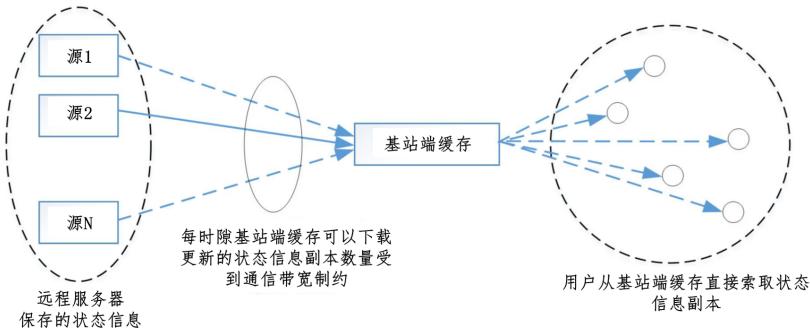


图 4.1 远程服务器、基站端缓存、用户系统图

### 4.2 针对状态信息流行度时变的缓存更新方案研究

#### 4.2.1 系统模型

本节针对基站端缓存所保存的状态信息流行度随时间变化的场景，设计满足带宽要求的基站端缓存更新策略，旨在优化用户索取得到副本的平均信息年龄。我们在文献<sup>[62]</sup>的基础上考虑状态信息流行度时变这一特性。考虑远程服务器存有  $N$  个随着时间变化的状态信息，基站端缓存拥有这些状态信息的副本（但可能不是最新的）。考虑离散时间系统，每时隙  $t \in \{1, \dots, T\}$  中，基站端缓存通过带宽受限的网络从远程服务器下载并更新状态信息副本，用户直接从基站端缓存获取状态信息副本。对于每份状态信息  $n \in \{1, \dots, N\}$ ，我们记  $\{u_n(t)\} \in \{0, 1\}$  为时隙  $t$  的下载决策。也就是说，如果  $u_n(t) = 1$ ，基站端缓存将在时隙  $t$  开始时从远程服务器

下载最新的状态信息  $n$ ; 如果  $u_n(t) = 0$ , 基站端缓存就不下载也不更新相应的状态信息。受限于带宽限制, 基站端缓存每时隙只能从远程服务器下载更新不多于  $M$  份状态信息, 这一限制条件对于  $u_n(t)$  有如下要求:

$$\sum_{n=1}^N u_n(t) \leq M, \quad \forall t. \quad (4.1)$$

我们用  $x_n(t)$  表示基站端缓存中, 状态信息  $n$  的副本在时隙  $t$  的信息年龄。我们假设在时间  $t = 0$  时, 基站端缓存中所有状态信息都与远程服务器进行了一次同步, 因此所有状态信息在时隙  $t = 1$  时的信息年龄  $x_n(t) = 1, \forall n$ 。在此之后, 与第二章相同, 基站端缓存状态信息  $n$  的信息年龄演进关系如下:

$$x_n(t+1) = \begin{cases} 1, & u_n(t) = 1; \\ x_n(t) + 1, & u_n(t) = 0. \end{cases} \quad (4.2)$$

在每时隙  $t$ , 索取状态信息  $n$  的用户数量依赖于状态信息  $n$  在此时的流行度模式  $r_n(t) \in \mathcal{R} := \{1, \dots, R\}$ , 其中  $R > 0$ 。每时隙索取状态信息  $n$  的用户数期望可以用函数  $\omega_n : \mathcal{R} \mapsto \mathbb{R}^+$  表示, 即当状态信息  $n$  的流行度模式为  $r_n(t)$  时, 下载它的总用户数期望为  $\omega_n(r_n(t)) > 0$ 。我们假设对状态信息  $n$  来说, 它的流行度模式  $\{r_n(t)\}_{t=1}^T$  为一个  $R$  状态的马尔可夫链, 对于任意两个状态  $(r, r') \in \mathcal{R}^2$ , 马氏链的转移概率可以写作  $P_{r,r'}^n := \Pr(r_n(t+1) = r' | r_n(t) = r)$ 。

考虑因果策略集合  $\Pi$ , 即  $t$  时隙的下载策略  $\{u_n(t)\}$  仅由当前和过去的流行度模式  $\{r_n(\tau)\}_{\tau \leq t}$ 、信息年龄  $\{x_n(\tau)\}_{\tau \leq t}$  以及流行度模式的转移概率  $\{P_{r,r'}^n\}$  决定。未来的流行度  $\{r_n(\tau)\}_{\tau > t}$  未知, 因此不可以用来设计下载和更新策略。研究的目标是设计基站端缓存更新策略  $\pi \in \Pi$ , 使得用户索取得到状态信息副本的信息年龄期望可以被最小化。这一问题从数学上建模如下:

**问题 4.1 (原问题):**

$$\pi^\star = \arg \min_{\pi \in \Pi} \lim_{T \rightarrow \infty} \mathbb{E}_\pi \left[ \frac{1}{T} \sum_{t=1}^T \sum_{n=1}^N \omega_n(r_n(t)) x_n(t) \right], \quad (4.3a)$$

$$\text{s.t. } \sum_{n=1}^N u_n(t) \leq M, \quad \forall t. \quad (4.3b)$$

我们注意到问题4.1可以被建模为一个马尔可夫决策过程, 每一时刻的状态包括当前时刻基站端缓存中副本的信息年龄  $\mathbf{x}(t) = [x_1(t), \dots, x_N(t)]^T$  和流行度模式  $\mathbf{r}(t) = [r_1(t), \dots, r_N(t)]^T$ ; 决策空间需要满足带宽限制条件式 (4.3b), 即可行决策集合  $\mathcal{A} := \{\mathbf{u}(t) | \sum_{n=1}^N u_n(t) \leq M\}$ 。我们注意到决策集合  $\mathcal{A}$  的大小随着带宽约束  $M$  和状态信息总数  $N$  指数增长, 因此对于稍大的  $M$  和  $N$  我们仍然很难通过值迭代

代<sup>[91]</sup>进行最优策略求解。

接下来，我们首先放松带宽限制式（4.3b），并在节4.2.2中讨论求解这一新约束下的最优更新策略  $\pi_R^*$ 。受这一更新策略启示，在节4.3.4，我们将提出一种可以满足问题4.1中带宽约束式（4.3b）的策略  $\hat{\pi}$  并分析其理论性能。节4.3.5提供了数值仿真结果，用以说明所提策略在降低用户获取副本信息年龄方面的表现，节4.3.6进行本工作的小结。

## 4.2.2 问题求解

### 4.2.2.1 松弛优化问题

与文献<sup>[28,62,76,92-93]</sup>相同，我们首先将约束（4.3b）松弛为时间平均的带宽限制条件。松弛后我们将得到如下优化问题：

**问题 4.2 (松弛优化问题):**

$$\pi_R^* = \arg \min_{\pi \in \Pi} \lim_{T \rightarrow \infty} \mathbb{E}_\pi \left[ \frac{1}{T} \sum_{t=1}^T \sum_{n=1}^N \omega_n(r_n(t)) x_n(t) \right], \quad (4.4a)$$

$$\text{s.t. } \lim_{T \rightarrow \infty} \mathbb{E}_\pi \left[ \frac{1}{T} \sum_{t=1}^T \sum_{n=1}^N u_n(t) \right] \leq M. \quad (4.4b)$$

注意到问题4.2是一个受限马尔可夫决策过程<sup>[88]</sup>，它的决策空间相比问题4.1甚至更大。实际上，通过对问题4.2中每份状态信息进行解耦，我们可以解决缓存更新问题中的“维度爆炸”。

文献<sup>[94]</sup>证明，问题4.2的最优策略  $\pi_R^*$  可以通过拉格朗日乘子法得到。我们将限制条件式（4.4b）通过拉格朗日乘子  $W \geq 0$  加入优化目标式（4.4a）中，得到如下拉格朗日函数：

$$\mathcal{L}(\pi, W) = \lim_{T \rightarrow \infty} \mathbb{E}_\pi \left[ \frac{1}{T} \sum_{t=1}^T \sum_{n=1}^N (\omega_n(r_n(t)) X_n(t) + W U_n(t)) - WM \right]. \quad (4.5)$$

接下来，我们将首先针对一个固定的  $W$ ，讨论最小化式（4.5）的基站端缓存更新策略。节4.2.2.3中，我们将确定松弛问题4.2最优策略对应的拉格朗日乘子  $W$ 。

每个固定的  $W$  可以被视作状态信息下载更新代价，式（4.5）即为所有状态信息下载更新代价与信息年龄的总和。通过交换式（4.5）时间  $t$  和状态信息  $n$  的求和顺序（因为对于  $n$  求和得到的数值有限），根据 [95, 第四章]，最小化拉格朗日函数式（4.5）的最优策略  $\pi_F^*(W)$  可以被分解为：

$$\pi_F^*(W) = \bigotimes_{n=1}^N \pi_F^{n,*}(W), \quad (4.6)$$

其中策略  $\pi_F^{n,\star}(W)$  定义如下：

**问题 4.3 (单状态信息松弛解耦优化问题):**

$$\pi_F^{n,\star}(W) \triangleq \arg \min_{\pi \in \Pi} \lim_{T \rightarrow \infty} \mathbb{E}_\pi \left[ \frac{1}{T} \sum_{t=1}^T (\omega_n(R_n(t))X_n(t) + WU_n(t)) \right]. \quad (4.7)$$

#### 4.2.2.2 问题4.3的求解算法

接下来我们对于给定的拉格朗日乘子  $W$  求解优化问题4.3。注意到每份状态信息  $n$  对应的优化问题均有相同的形式，在接下来的分析中，我们略去状态信息下标  $n$ 。

问题4.3为一个马尔可夫决策过程，它的状态为  $(x(t), r(t))$ ，每时刻决策  $u(t) \in \{0, 1\}$ ，在状态  $(x(t), r(t)) = (x, r)$  时采取行动  $u(t) = u$  的一步代价为：

$$C(x, r, u) = \omega(r)x + Wu. \quad (4.8)$$

根据式 (4.2)，若基站端缓存在时隙  $t$  决定更新状态信息，即  $u(t) = 1$ ，下一时刻状态信息副本的信息年龄将降为 1；否则信息年龄将增大 1。因此，该马尔可夫决策过程的状态转移函数为：

$$\Pr((x(t), r(t)) \rightarrow (1, r')) = P_{r(t), r'}, \text{ 如果 } u(t) = 1; \quad (4.9a)$$

$$\Pr((x(t), r(t)) \rightarrow (x(t) + 1, r')) = P_{r(t), r'}, \text{ 如果 } u(t) = 0. \quad (4.9b)$$

**定义 4.1：**一个策略  $\pi$  被称为是静态的，如果对于任意时间  $t$ ，假设状态信息和流行度模式满足  $(x(t) = x, r(t) = r)$ ，策略  $\pi$  以概率  $\xi_{x,r}$  选择下载  $u(t) = 1$  该状态信息，其中概率  $\xi_{x,r}$  仅和信息年龄、流行度  $(x, r)$  有关，和时间  $t$  无关。

接下来的定理将说明最优策略  $\pi_F^\star$  是一个有阈值结构的静态策略：

**定理 4.1：**问题4.3有一个最优策略  $\pi_F^\star$ ，即我们可以找到一系列的阈值  $\{\tau_r\}_{r \in \mathcal{R}}$ ，当状态信息的流行度模式为  $r$  时，如果它的信息年龄  $x > \tau_r$ ，策略  $\pi_F^\star$  以概率 1 更新状态信息；否则当信息年龄  $x \leq \tau_r$ ，策略  $\pi_F^\star$  不更新对应状态信息。

该定理的证明与本论文中引理3.2的证明相似，在此省略。

根据定理4.1，由于状态信息流行度在有限状态集合  $\mathcal{R}$  中选取，记  $X_{\max} = \max_{r \in \mathcal{R}} \tau_r$  为最大调度阈值。根据优化目标式 (4.7)，一定存在一个最优策略  $\pi_F^\star$ ，它的调度阈值最大值满足：

$$X_{\max} \leq \max_{r \in \mathcal{R}} \frac{W + \omega(r)}{\omega(r)} =: X_{\text{ub}}. \quad (4.10)$$

不等式 (4.10) 成立是因为：对于任意不满足式 (4.10) 的策略  $\pi$ ，我们可以找到一个  $\pi$  的改进策略，每当信息年龄  $X > \frac{W + \omega(r)}{\omega(r)}$  时就去更新状态信息，这样的策略得到的信息年龄更小。进一步地，对于任意的流行度模式，索取该状态信息的用户

数总是满足  $\omega(r) > 0$ , 因此阈值  $X_{ub}$  是有限的。需要强调的是这个最大的阈值依赖于状态信息  $n$  的流行度演进特性和拉格朗日乘子  $W$ , 在讨论拉格朗日乘子  $W$  选取的部分中, 我们将使用  $X_{ub}^n(W)$  体现这一最大阈值与  $W$  之间的相关性。

接下来我们可以将马尔可夫决策过程问题4.3转化为状态空间有限的决策问题。本节中我们同样通过文献<sup>[88]</sup>提供线性规划方案, 将问题4.3转化为关于概率占用状态  $\{\mu_{x,r}\}$  和  $\{v_{x,r}\}$  的线性规划。其中  $\mu_{x,r}$  表示使用静态策略时, 状态信息处于状态  $(x,r)$  的概率, 而占用测度 (occupation measure)  $v_{x,r}$  表示在稳态分布下, 状态信息处于状态  $(x,r)$  被下载和更新的概率。即, 当状态信息的信息年龄为  $x$ 、流行度模式处于  $r$  时, 静态策略以概率  $\xi_{x,r} = v_{x,r}/\mu_{x,r}$  下载和更新该状态信息。我们记  $0/0 = 1$ , 因为稳态概率为 0 说明相应状态信息的信息年龄已经超过了最大下载更新阈值  $X_{max}$ 。通过以上的定义, 依据<sup>[88]</sup>定理 4.3, 我们可以通过如下的定理求解问题4.3:

**定理 4.2 (定理 4.3<sup>[88]</sup>的等价形式):** 记  $X_{ub}$  为式 (4.10) 中更新阈值的上界, 问题4.3的最优策略  $\{\xi_{x,r}^\star\}_{x,r}$  为:

$$\xi_{x,r}^\star = \frac{v_{x,r}^\star}{\mu_{x,r}^\star}, \quad (4.11)$$

其中  $v_{x,r}^\star$  和  $\mu_{x,r}^\star$  可以由以下线性规划问题得到:

$$\{\mu_{x,r}^\star, v_{x,r}^\star\} = \arg \min_{\{\mu_{x,r}, v_{x,r}\}} \sum_{x=1}^{X_{ub}} \sum_{r=1}^R (\omega(r)x\mu_{x,r} + Wv_{x,r}), \quad (4.12a)$$

$$\text{s.t. } \mu_{1,r} = \sum_{x=1}^{X_{ub}} \sum_{r'=1}^R v_{x,r'} P_{r',r}, \quad (4.12b)$$

$$\mu_{x,r} = \sum_{r'=1}^R (\mu_{x-1,r'} - v_{x-1,r'}) P_{r',r}, \forall x > 1, \quad (4.12c)$$

$$\sum_{x=1}^{X_{ub}} \sum_{r=1}^R \mu_{x,r} = 1, \quad (4.12d)$$

$$v_{x,r} \leq \mu_{x,r}, \quad (4.12e)$$

$$0 \leq \mu_{x,r}, 0 \leq v_{x,r}, \forall x, r. \quad (4.12f)$$

**注 4.1:** 根据<sup>[88]</sup>定理 4.3, 定理4.2中的线性规划问题对应的每一个可行解可以通过式 (4.11) 转化成一个静态策略。同样的, 问题4.3中任意一个静态策略对应着定理4.2线性规划问题的一个可行解。

**注 4.2:** 上述定理4.2中, 线性规划问题的约束条件除了式 (4.12e) 以外都可以通

过[88]定理4.3直接得到或转化。约束(4.12e)是因为状态 $(x, r)$ 的稳态分布为所有决策 $u$ 占位测度的和，而 $\nu_{x,r}$ 表示决策 $u = 1$ 的占位测度，因此 $\nu_{x,r} \leq \mu_{x,r}$ 。

#### 4.2.2.3 松弛问题4.2的求解

节4.2.2.2中，我们提出了一种针对固定拉格朗日乘子 $W$ ，求解问题4.3线性规划方案。根据式(4.6)，这些策略合起来可以优化拉格朗日函数式(4.5)。也就是说，在每时刻 $t$ ，当状态信息的信息年龄 $(x_1(t), \dots, x_N(t)) = [x_1, \dots, x_N]$ ，流行度状态 $(r_1(t), \dots, r_N(t)) = [r_1, \dots, r_N]$ 时，最小化拉格朗日函数式(4.5)的最优策略 $\pi_F^*(W)$ 以概率 $\xi_{x_n, r_n}^{n, \star}(W) = \frac{\nu_{x_n, r_n}^{n, \star}(W)}{\mu_{x_n, r_n}^{n, \star}(W)}$ 更新状态信息 $n$ ，这一决策与缓存中的其他状态信息 $n'$ 没有关系。使用策略 $\pi_F^*(W)$ 时，平均每时隙下载的状态信息数量为：

$$d^*(W) := \sum_{n=1}^N \sum_{x=1}^{X_{ub}^n(W)} \sum_{r=1}^R \nu_{x,r}^{n, \star}(W), \quad (4.13)$$

使用策略 $\pi_F^*(W)$ ，用户从基站端缓存索取获得状态信息副本的平均信息年龄为：

$$\bar{X}^*(W) \triangleq \sum_{n=1}^N \sum_{x=1}^{X_{ub}^n(W)} \sum_{r=1}^R \omega(r)x \mu_{x,r}^{n, \star}(W). \quad (4.14)$$

因此，对于给定的拉格朗日乘子 $W$ ，式(4.5)的最小代价，即使用策略 $\pi_F^*(W)$ 时的代价函数可以通过如下公式计算：

$$\mathcal{L}(\pi_F^*(W), W) = \bar{X}^*(W) + W d^*(W) - M. \quad (4.15)$$

接下来，我们将讨论如何选取拉格朗日乘子 $W$ 。根据文献[96]，对于任意的拉格朗日乘子 $W$ 和下载策略 $\pi$ ，我们都：

$$\bar{X}^*(W) + W d^*(W) - MW \leq J(\pi) + W d(\pi) - MW, \quad (4.16)$$

其中 $J(\pi)$ 和 $d(\pi)$ 分别表示使用策略 $\pi$ 时，用户索取得到副本的平均信息年龄和平均每时刻下载的状态信息数量。根据式(4.16)，如果存在拉格朗日乘子 $\tilde{W}$ 使得 $d(\pi_F^*(\tilde{W})) = M$ ，则该策略对应的平均信息年龄 $J(\pi_F^*(\tilde{W}))$ 为所有满足带宽约束的策略 $\pi$ (即 $d(\pi) \leq M$ )中最小的。根据[96]引理3.3，函数 $d(\pi_F^*(W))$ 关于 $W$ 非递增，如果这样的 $\tilde{W}$ 存在，我们定义 $\pi^*$ ：

$$W^* \triangleq \inf\{W | d^*(W) \leq M\}. \quad (4.17)$$

在这样的情况下，问题4.2的最优策略就是 $\pi_R^* = \pi_F^*(W^*)$ 。

当函数 $W \mapsto d(\pi_F^*(W))$ 在 $W^*$ 处不连续时，我们通过定两个中间策略寻找 $\pi_R^*$ 。我们注意到策略 $\pi_F^*(W)$ 为所有策略 $\pi_F^{n, \star}(W)$ 的连乘，其中每个策略

$\{\xi_{x,r}^{n,\star}(W)\}_{1 \leq x \leq X_{ub}^n(W), r \in \mathcal{R}}$  可以通过求解定理4.2中的线性规划问题得到。首先考虑拉格朗日乘子  $W \in [0, W^\star]$  的情况，根据式(4.10)，我们定义调度阈值上界  $X_{ub,ls}^n = \max_{W \in [0, W^\star]} X_{ub}^n(W)$ 。考虑这些拉格朗日乘子对应的最优调度策略  $\{\xi_{x,r}^{n,\star}(W)\}_{1 \leq x \leq X_{ub,ls}, r \in \mathcal{R}}$ ，对于任意状态信息  $n$  的任意状态  $(x, r)$ ，最优解  $\xi_{x,r}^{n,\star}(W)$  的上确界小于等于 1，因此存在一个子列  $\{W_{ls,\ell}\}_{\ell \geq 1}$  左收敛于  $W^\star$ ，即序列左极限  $\xi_{x,r,ls}^{n,\star} := \lim_{\ell \rightarrow \infty} \xi_{x,r}^{n,\star}(W_{ls,\ell})$  存在。同样地，考虑拉格朗日乘子  $W \geq W^\star$ ，取  $X_{ub,rs}^n = \max_{W \in [W^\star, \bar{W}]} X_{ub}^n(W)$ ，同样存在一个收敛子列  $\{W_{rs,\ell}\}_{\ell \geq 1}$  从右边逼近于  $W^\star$ ，即序列  $\xi_{x,r,rs}^{n,\star} := \lim_{\ell \rightarrow \infty} \xi_{x,r}^{n,\star}(W_{rs,\ell})$  的右极限也存在。因此，我们可以分别定义

$$\pi_{F,ls}^\star := \lim_{W \uparrow W^\star} \pi_F^\star(W) \mathbb{I} \quad (4.18)$$

和

$$\pi_{F,rs}^\star := \lim_{W \downarrow W^\star} \pi_F^\star(W). \quad (4.19)$$

注意到策略  $\pi_{F,ls}^\star$  和  $\pi_{F,rs}^\star$  可能会因为收敛子列的选取有所不同，但这一不同将不会影响本节的最终结果。

接下来，我们构造一个最优静态策略  $\pi_R^\star$  满足松弛后的带宽条件并最小化索取副本信息年龄。我们通过定义以下的稳态分布和占用测度进行构造：

$$\bar{\mu}_{x,r}^{n,\star} \triangleq \lambda \mu_{x,r,ls}^{n,\star} + (1 - \lambda) \mu_{x,r,rs}^{n,\star}, \quad (4.20)$$

$$\bar{v}_{x,r}^{n,\star} \triangleq \lambda v_{x,r,ls}^{n,\star} + (1 - \lambda) v_{x,r,rs}^{n,\star}, \quad (4.21)$$

其中  $\{\mu_{x,r,ls}^{n,\star}, v_{x,r,ls}^{n,\star}\}$  和  $\{\mu_{x,r,rs}^{n,\star}, v_{x,r,rs}^{n,\star}\}$  分别表示策略  $\pi_{F,ls}^\star$  和  $\pi_{F,rs}^\star$  的稳态分布概率和占用测度。根据  $\{\bar{\mu}_{x,r}^{n,\star}, \bar{v}_{x,r}^{n,\star}\}$ ，我们可以构造策略  $\pi_R^\star$ ：在时间  $t$ ，基站端缓存中状态信息副本  $n$  的信息年龄  $x_n(t) = x$ ，流行度  $r_n(t) = r$  时，策略  $\pi_R^\star$  以概率  $\bar{\xi}_{x,r}^{n,\star} \triangleq \frac{\bar{v}_{x,r}^{n,\star}}{\bar{\mu}_{x,r}^{n,\star}}$  更新状态信息  $n$ 。定义  $\bar{X}_{ub}^n = \max(X_{ub,ls}^n, X_{ub,rs}^n)$ ，使用新策略  $\pi_R^\star$  时，用户索取得到状态信息副本的平均信息年龄为：

$$J(\pi_R^\star) = \sum_{n=1}^N \sum_{x=1}^{\bar{X}_{ub}^n} \sum_{r=1}^R \omega(r) x [\lambda \mu_{x,r,ls}^{n,\star} + (1 - \lambda) \mu_{x,r,rs}^{n,\star}] \quad (4.22)$$

平均每时隙下载的状态信息总数满足

$$\sum_{n=1}^N \sum_{x=1}^{\bar{X}_{ub}^n} \sum_{r=1}^R [\lambda v_{x,r,ls}^{n,\star} + (1 - \lambda) v_{x,r,rs}^{n,\star}] = M. \quad (4.23)$$

#### 4.2.2.4 一种满足原每时隙带宽约束的基站端缓存更新方案

接下来我们回到问题4.1，我们在策略  $\pi_R^*$  的基础上提出一种满足问题4.1中每时隙带宽约束的基站端缓存更新策略  $\hat{\pi}$ : 若时隙  $t$  时，策略  $\pi_R^*$  选择下载更新多于  $M$  份状态信息，策略  $\hat{\pi}$  从  $\pi_R^*$  选取的状态信息中随机选取  $M$  份进行更新；否则，策略  $\hat{\pi}$  就将策略  $\pi_R^*$  选取的状态信息全部下载更新。接下来，我们将证明保持  $N/M$  固定不变，当状态信息总数  $N \rightarrow \infty$  时，节4.2.2.4所提出的调度策略  $\hat{\pi}$  得到的平均信息年龄趋于问题4.1中最优策略  $\pi^*$  得到的平均信息年龄。

**定理 4.3:** 考虑  $N$  份状态信息有相同的流行度转移概率  $P_{r,r'}$  和相同的被索取用户数期望  $\omega(r)$ <sup>①</sup>。固定比值  $\theta := N/M$  保持不变并与  $N$  无关，如果策略  $\pi_R^*$  导出的马氏链  $(x, r)$  是遍历的<sup>②</sup>，节4.2.2.4所提策略  $\hat{\pi}$  有渐近最优性，即：

$$\lim_{\substack{N \rightarrow \infty \\ \theta = N/M}} \frac{J(\hat{\pi}) - J(\pi^*)}{J(\pi^*)} = 0, \quad (4.24)$$

其中策略  $\pi^*$  为问题4.1中的最优策略。

**证明:** 本部分证明与定理3.2的证明类似。在此省略。 ■

#### 4.2.3 数值仿真

在数值仿真中，我们考虑每份状态信息有两个流行度状态  $\mathcal{R} = \{1, 2\}$ ，它的流行度转移矩阵如下，其中  $q \in (0, 1)$ :

$$\mathbf{P}^n = \begin{bmatrix} q & 1-q \\ 1-q & q \end{bmatrix}, \quad \forall n \in \{1, \dots, N\}. \quad (4.25)$$

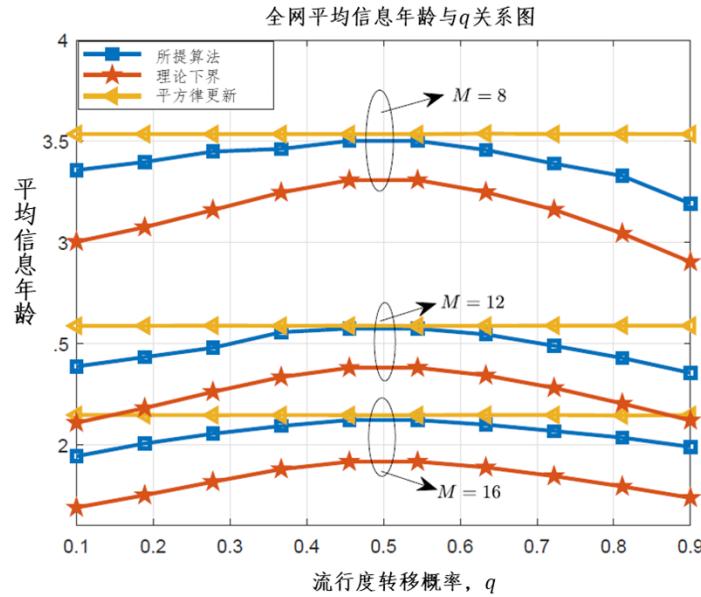
状态信息  $n$  在不同流行度模式下期望的被索取数目分别为  $\omega_n(1) = 0.2\bar{\omega}_n$ ,  $\omega_n(2) = 1.8\bar{\omega}_n$ ，其中  $\bar{\omega}_n \propto 1/n^\alpha$  服从参数  $\alpha = 1.5$  的 Zipf 分布。根据式 (4.25)，每份状态信息处于不同流行度模式的稳态分布概率都是一样的。因此  $\bar{\omega}_n$  是状态信息  $n$  被索取的用户数期望。

图4.2画出了基站端缓存中存有  $N = 64$  份状态信息时，采用策略  $\hat{\pi}$ ,  $\pi_R^*$  和平方律更新策略，用户获取状态信息副本的平均信息年龄在不同带宽  $M$  下，随着转移概率  $q$  的变化曲线。注意到策略  $\pi_R^*$  得到的平均信息年龄为问题4.1的理论下界；平方律更新策略<sup>[62]</sup>没有考虑状态信息流行度模式  $r_n(t)$  的时变特性，仅依赖于  $\bar{\omega}_n$  进行基站端缓存更新策略设计。

首先我们注意到所提策略在  $q \neq 0.5$  时得到的平均信息年龄都比使用平方律更新策略得到的平均信息年龄小。当  $q = 0.5$  时，下一时刻的流行度与当前时刻状态

① 这一结论可以被拓展到状态信息被分为几组，每组中状态信息有相同的流行度转移概率和被索取数期望。

② 这一遍历条件很容易满足，这是因为状态信息流行度转移的随机性很容易导出状态演进的非周期和正常返特性。

图 4.2 被索取状态信息的平均信息年龄与流行度转移概率  $q$  关系图

信息的流行度完全无关，因此当前状态信息流行度不能为下一时刻的状态信息流行度提供任何信息，所提策略在降低用户获取状态信息副本的平均信息年龄方面效果有限；当  $q$  离 0.5 相差越远，使用所提策略  $\hat{\pi}$  在降低平均信息年龄方面效果越显著，这是因为下一时刻的状态信息流行度与当前时刻更为相关，利用当前时刻状态信息流行度可以较好地对状态信息下一时刻的流行度进行预测和估计。当网络中可用带宽  $M$  变大，被索取状态信息的信息年龄减小，因为每时隙可以更新的状态信息数目变多。

#### 4.2.4 本节小结

本节针对基站端缓存的状态信息流行度时变的场景，在下载更新带宽受限时，针对用户所获取状态信息副本的信息年龄最小化问题，提出了一种基站端缓存更新方案。我们通过松弛每时刻更新状态信息数量的严格带宽约束，将原问题转化为受限马尔可夫决策过程，使用拉格朗日乘子法对状态信息进行解耦，通过线性规划求解了受限马尔可夫决策过程的最优策略。在带宽松弛问题最优策略基础上，我们提出了一种可以满足原每时隙严格带宽约束的基站端缓存更新策略，该策略在状态信息数量趋于无穷时有渐近最优性。我们通过数值仿真，说明了所提算法相比于平方律更新策略，在状态信息流行度时变时可以显著降低用户从基站端缓存获取状态信息副本的平均信息年龄。

### 4.3 针对更新时长随信息年龄变化的基站端缓存更新策略研究

在本节中，我们进一步考虑状态信息的更新时长随着它的信息年龄变化，在带宽限制下设计基站端缓存更新策略，最小化用户获取到状态信息副本的信息年龄。实际上，如果一份状态信息副本当前信息年龄较小，那么当前副本与远程服务器上状态信息差别不会很大，也就是说下载和更新所用时长很可能较短；相反，如果一份状态信息副本信息年龄较大，那么基站端缓存中大部分的内容都需要被更换，也就是说更新的时间会相对较长。

类似于文献[62,80,97]，本节考虑基站端缓存可以在任意时间选择从远程服务器上更新和下载状态信息。我们假设每份状态信息被索取的概率随着时间固定不变，而每次更新所需时长与它的信息年龄有关。我们设计了基站端缓存更新方案，优化用户索取得到状态信息副本的平均信息年龄。

#### 4.3.1 系统模型和问题描述

本节考虑的系统如图4.1所示，远程服务器存有  $N$  个随时间变化的状态信息。假设时间  $t = 0$  时，基站端缓存存有每份状态信息最新的副本。在接下来的事件中，每份状态信息随时间不断变化，而为了保证这些存有的副本内容尽可能新，基站端缓存将会不断从远程服务器下载并更新本地状态信息副本。受限于传输带宽限制，每时刻基站端缓存只能更新一份状态信息。同时，我们假设在之前的下载结束后，新状态信息的更新和下载可以在任意时刻开始。

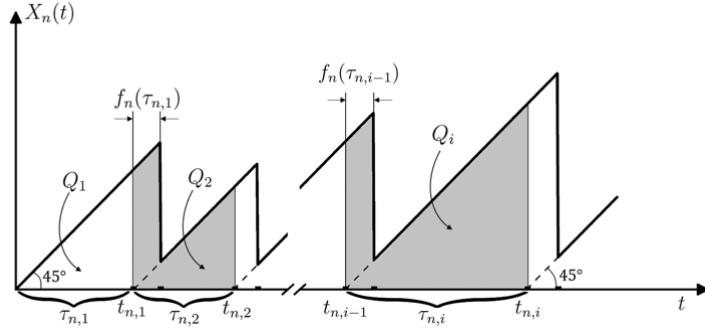
我们考虑从 0 到  $T(T \geq 0)$  的一段连续时间。我们用  $K_n$  表示状态信息  $n \in \{1, \dots, N\}$  在这段时间内被下载更新的总次数，并将  $0 < t_{n,1} < \dots < t_{n,K_n}$  记为这些更新的开始时刻。我们将  $\tau_{n,i}$  记为状态信息  $n$  第  $i$  次与第  $(i+1)$  次更新的时间间隔， $\tau_{n,K_n+1}$  的计算公式如下：

$$\tau_{n,i} := \begin{cases} t_{n,1}, & i = 1; \\ t_{n,i} - t_{n,i-1}, & i = 2, \dots, K_n; \\ T - t_{n,K_n}, & i = K_n + 1. \end{cases} \quad (4.26)$$

接下来我们介绍本问题中每次下载状态信息所需要的时间以及信息年龄的演进。在此之前首先我们给出  $N$  个函数  $f_1(\cdot), \dots, f_N(\cdot)$ ，他们具有如下性质：

**假设 4.1(更新时长函数):** 我们假设对于任意  $n \in \{1, \dots, N\}$ ，函数  $f_n(\cdot) : \mathbb{R}^+ \mapsto \mathbb{R}^+$  在定义域  $\mathbb{R}^+$  上严格大于 0，是有界非递减的。同时，这些函数都是可微、下凹的。

接下来我们介绍本问题中每份状态信息  $n$  的信息年龄  $\{x_n(t)\}_{t \geq 0}$  的演进。假设时刻  $t = 0$ ，基站端缓存有最新的状态信息，因此对于任意状态信息  $n$  来说，它的

图 4.3 状态信息  $n$  的信息年龄演进关系图

信息年龄在此刻都是  $x_n(0) = 0$ 。在基站端缓存成功接收到状态信息  $n$  下一次的更新前，它的信息年龄以  $x_n(t) = t$  的形式线性增长。注意到  $t_{n,1}$  为基站端缓存第一次开始更新状态信息  $n$  的时刻，这一次的下载和更新将在时刻  $t_{n,1} + d_{n,1}$  停止，其中  $d_{n,1}$  表示状态信息  $n$  第一次下载所用去的时长。本章考虑状态信息下载时长依赖于更新开始时（即时刻  $t_{n,1}$ ）状态信息的信息年龄，即下载更新时长由函数  $f_n$  决定：

$$d_{n,1} := f_n(x_n(t_{n,1})) = f_n(t_{n,1}) = f_n(\tau_{n,1}), \quad (4.27)$$

其中等号成立的原因是在  $t_{n,1}$  时刻，根据式(4.26)，状态信息  $n$  的信息年龄  $x_n(t_{n,1}) = t_{n,1}$ 。

在状态信息  $n$  第一次更新结束（时刻  $t = t_{n,1} + d_{n,1}$ ）时，根据信息年龄的定义，基站端缓存中，状态信息  $n$  副本的信息年龄将下降为当前时间与最新版本的创建时间之差，即  $x_n(t_{n,1} + d_{n,1}) = d_{n,1}$ ，接下来在状态信息  $n$  的第二次更新结束之前，他的信息年龄线性增长，即  $x_n(t) = t - t_{n,1}$ ，当第二次更新结束（时刻  $t = t_{n,2} + d_{n,2}$ ），基站端缓存中，状态信息  $n$  副本的信息年龄下降为  $d_{n,2}$ ，接着再次以  $x_n(t) = t - t_{n,2}$  线性增长……信息年龄的演进曲线如图4.3所示：

总结来说，当状态信息下载时间随信息年龄变化时，任意一份状态信息  $n \in \{1, \dots, N\}$  的信息年龄计算如下：

$$x_n(t) = \begin{cases} t, & t \in [0, t_{n,1} + d_{n,1}); \\ (t - t_{n,k}), & t \in [t_{n,k} + d_{n,k}, t_{n,k+1} + d_{n,k+1}), \end{cases} \quad (4.28)$$

其中  $d_{n,k}$  为状态信息  $n$  第  $k$  次下载所用时间：

$$d_{n,k} := f_n(\tau_{n,k}), \quad (4.29)$$

状态信息  $n$  在  $(0, T]$  区间内的平均信息年龄计算如下：

$$\bar{x}_n = \frac{1}{T} \int_0^T x_n(t) dt. \quad (4.30)$$

考虑用户从基站端缓存中索取状态信息时，每份状态信息  $n \in \{1, \dots, N\}$  被索取的概率为  $p_n$ 。记  $J(\pi)$  为使用下载更新策略  $\pi$  时，用户索取得到状态信息副本的平均信息年龄：

$$J(\pi) := \sum_{n=1}^N p_n \bar{x}_n. \quad (4.31)$$

其中每次更新的开始时刻  $\{t_{n,1}, \dots, t_{n,N}\}_{n=1}^N$  需要满足带宽限制，即任意时间只有一份状态信息正在被下载更新。这一约束要求任意状态信息  $n \in \{1, \dots, N\}$  和更新  $i \in \{1, \dots, K_n\}$  的所使用时间  $[t_{n,i}, t_{n,i} + d_{n,i})$  彼此不相交。

### 4.3.2 优化问题建模

本节的研究目标是通过设计满足带宽要求的更新策略  $\{t_{n,1}, \dots, t_{n,K_n}\}_{n=1}^N$ ，最小化索取状态信息平均信息年龄  $J(\pi)$  (式 (4.31))。为了解决该问题，我们首先将平均信息年龄  $J(\pi)$  写成基站端缓存更新间隔与传输时间的函数。注意到时间  $[0, T)$  内状态信息  $n$  的累计信息年龄可以写成  $K_n + 1$  个时间间隔中的信息年龄之和，即：

$$\bar{x}_n = \frac{1}{T} \sum_{i=1}^{K_n+1} \int_{t=t_{n,i-1}}^{t_{n,i}} x_n(t) dt = \frac{1}{T} \sum_{i=1}^{K_n+1} Q_{n,i}, \quad (4.32)$$

其中时间间隔内  $[t_{n,i-1}, t_{n,i})$  的累计信息年龄  $Q_{n,i}$  如图4.3所示，可以写成一个平行四边形和三角形的面积和：

$$Q_{n,i} := \int_{t=t_{n,i-1}}^{t=t_{n,i}} X_n(t) dt = \tau_{n,i-1} \cdot f_n(\tau_{n,i-1}) + \frac{1}{2} \tau_{n,i}^2. \quad (4.33)$$

经过化简后，本节的优化问题建模如下：

**问题 4.4：**

$$\min_{\{K_n, \tau_{n,i}\}} \sum_{n=1}^N p_n \left( \sum_{i=1}^{K_n+1} \frac{1}{2} \tau_{n,i}^2 + \sum_{i=1}^{K_n} \tau_{n,i} \cdot f_n(\tau_{n,i}) \right) \quad (4.34a)$$

由于基站端缓存从服务器下载更新状态信息时有带宽约束，更新次数  $K_n$  和更新时间间隔  $\tau_{n,i}$  的选取需要满足以下条件：

$$\sum_{i=1}^{K_n+1} \tau_{n,i} = T, \quad \forall n \in \{1, \dots, N\}, \quad (4.34b)$$

$$\left( \sum_{k=1}^i \tau_{n,k} - \sum_{k=1}^j \tau_{n',k} \right) \notin \left( -f_n(\tau_{n,i}), f_{n'}(\tau_{n',j}) \right], \forall (n, i) \neq (n', j). \quad (4.34c)$$

其中约束条件式(4.34b)是每份状态信息每次更新间隔加和为 $T$ , 约束条件式(4.34c)来源于带宽限制下, 任意两段文件下载时长 $\left[\sum_{k=1}^i \tau_{n,k}, \sum_{k=1}^i \tau_{n,k} + f_n(\tau_{n,i})\right]$ 、 $\left[\sum_{k=1}^j \tau_{n',k}, \sum_{k=1}^j \tau_{n',k} + f_{n'}(\tau_{n',j})\right]$ 彼此不相交。

### 4.3.3 一个可行的松弛问题次优解

问题4.4即使在渐近区域 $T \rightarrow \infty$ 中仍然为一个困难的整数规划问题, 因此我们将约束条件(4.34c)放松为:

$$\sum_{n=1}^N \sum_{i=1}^{K_n} f_n(\tau_{n,i}) \leq T, \quad (4.35)$$

即每时刻多份状态信息可以被同时下载和更新, 只要所有状态信息总的下载更新时长小于时间 $T$ 即可。当每份状态信息下载时间为常数时, 文献<sup>[62]</sup>证明了当状态信息更新间隔也为常数时, 用户索取得到状态信息的信息年龄最小。在接下来的研究中, 我们考虑对于任意一份状态信息每次更新时间间隔都是一样的:

$$\tau_{n,i} = \bar{\tau}_n := \frac{T}{K_n + 1}. \quad (4.36)$$

通过这样的简化, 问题4.4就转化为寻找一组最优的状态信息更新间隔 $\{\bar{\tau}_n\}$ 最小化索取副本的平均信息年龄, 我们将在节4.3.3.1讨论这一问题的求解。节4.3.4中, 我们将提出一种满足每时刻只下载一份状态信息的基站端缓存更新策略, 并通过仿真说明该算法与采用等间隔更新策略得到的平均信息年龄相近。

#### 4.3.3.1 带宽松弛优化问题

在本节中, 我们考虑观测时间 $T \rightarrow \infty$ 。根据式(4.34a), 当时间 $T \rightarrow \infty$ 时, 只有对于所有状态信息 $n$ , 它们下载和更新总次数都满足 $K_n \rightarrow \infty$ , 用户索取状态信息的平均信息年龄才可能有界从而达到最小。在接下来的研究中, 我们只考虑 $\lim_{T \rightarrow \infty} K_n \rightarrow \infty$ 的情况, 也就是说:

$$\lim_{T \rightarrow \infty} \frac{K_n}{K_n + 1} = 1. \quad (4.37)$$

将式(4.36)和式(4.37)带入式(4.34a)和式(4.35), 我们可以得到如下新的优化问题, 在时间 $T \rightarrow \infty$ 时它可以看成是原问题4.4的一个近似:

**问题 4.5 (带宽松弛优化问题-1):**

$$\min_{\{\bar{\tau}_n\}_{n \geq 0}} \sum_{n=1}^N p_n \left( \frac{1}{2} \bar{\tau}_n + f_n(\bar{\tau}_n) \right) \quad (4.38a)$$

s.t.

$$\sum_{n=1}^N \frac{f_n(\bar{\tau}_n)}{\bar{\tau}_n} \leq 1. \quad (4.38b)$$

接下来我们将优化问题4.5进一步转化为一个单调优化问题进行求解。首先，我们记  $\lambda_n$  为使用下载间隔  $\bar{\tau}_n$  时，状态信息  $n$  的下载时长与下载时间间隔的比值：

$$\lambda_n := \frac{f_n(\bar{\tau}_n)}{\bar{\tau}_n}. \quad (4.39)$$

对于有限时间  $T$ ，状态信息  $n$  的下载时间占比为  $(1/T) \sum_{i \in \{1, \dots, K_n\}} f_n(\tau_{n,i}) = K_n f_n(\bar{\tau}_n) / ((K_n + 1) \bar{\tau}_n)$ ，根据式 (4.37)，当时间  $T \rightarrow \infty$  时，下载时间占用比率趋近于  $\lambda_n = \frac{f_n(\bar{\tau}_n)}{\bar{\tau}_n}$ 。根据式 (4.38b)，我们得到  $\lambda_n$  的约束条件： $\sum_{n=1}^N \lambda_n \leq 1$ 。

为了进一步将问题4.5转化为单调优化问题，我们定义函数  $g_n(t)$  为：

$$g_n(t) := \frac{f_n(t)}{t}. \quad (4.40)$$

我们可以得到函数  $g_n(t)$  的一些性质：

**引理 4.1：** 在假设4.1下，函数  $t \mapsto g_n(t)$  在  $t \in \mathbb{R}_+$  上严格递减。他的像为  $(0, \infty)$ 。同时函数  $g_n$  的逆函数  $g_n^{(-1)}$  也是严格递减的。

**证明：** 函数  $g_n(t)$  的导数的上界为：

$$g'_n(t) = \frac{tf'_n(t) - f_n(t)}{t^2} \stackrel{(a)}{\leq} \frac{-f_n(0)}{t^2} \stackrel{(b)}{<} 0, \quad (4.41)$$

其中不等式 (a) 来源于函数  $f_n$  的凹性，不等式 (b) 是因为函数  $f_n$  在定义域上总是正的。因此，函数  $g_n$  是严格递减的。由于  $\lim_{t \rightarrow 0} f_n(t) = f_n(0) > 0$ ，我们有  $\lim_{t \rightarrow 0} \frac{f_n(t)}{t} = \infty$ 。又因为函数  $f_n$  有上界，我们有  $\lim_{t \rightarrow \infty} \frac{f_n(t)}{t} = 0$ 。最后，根据函数  $f_n$  连续可微、严格递减，它的像为  $(0, \infty)$ 。函数  $g_n^{-1}$  严格递减性质的证明是相似的。 ■

因此，记  $\bar{\tau}_n = g_n^{(-1)}(\lambda_n)$ ，问题4.5可以转化为有关比值  $\{\lambda_n\}_{n=1}^N$  的优化问题：

**问题 4.6 (带宽松弛优化问题-2)：** 记

$$h_n(\lambda) := g_n^{(-1)}(\lambda) \left( \frac{1}{2} + \lambda \right). \quad (4.42)$$

问题4.5与以下问题等价：

$$\min_{\{\lambda_n\}_{n=1}^N \geq 0} \sum_{n=1}^N p_n \cdot h_n(\lambda_n) \quad (4.43a)$$

s.t.

$$\sum_{n=1}^N \lambda_n \leq 1. \quad (4.43b)$$

接下来，我们将首先证明问题4.6的最优解  $\{\lambda_n\}_{n=1}^N$  落在约束条件的边界上，这一性质将帮助我们设计问题4.6的求解方案。我们将提出两种求解方案，一种利用函数  $h_n$  的单调特性求解优化问题，这种方案可以用于任意单调、凹的更新函数  $f_n$  但是求解复杂度较高；当函数  $f_n$  的选取使函数  $h_n(\cdot)$  是凸的，我们可以利用 KKT 条件进行问题求解。

我们后续求解依赖于以下引理：

**引理 4.2：** 函数  $h_n(\cdot)$  在定义域  $\mathbb{R}_+$  上严格递减。

**证明：** 根据引理4.1，函数  $g_n^{(-1)}(\lambda)$  是严格递减的。同时，根据式 (4.40)，我们有  $f_n(g_n^{(-1)}(\lambda)) = g_n^{(-1)}(\lambda) \cdot g_n(g_n^{(-1)}(\lambda)) = g_n^{(-1)}(\lambda)\lambda$ 。由于函数  $f_n$  非递减，函数  $g_n^{(-1)}$  严格递减，复合函数  $f_n \circ g_n^{(-1)}$  是非递增的。 ■

**命题 4.1：** 记  $\{\lambda_n^*\}_n$  为问题4.6的最优解。那么可以证明它满足：

$$\sum_{n=1}^N \lambda_n^* = 1, \quad (4.44)$$

也就是说，最优解  $\{\lambda_n^*\}_{n=1}^N$  可行集边界上。

**证明：** 我们通过反证法进行证明，假设  $\sum_{n=1}^N \lambda_n^* < 1$ ，那么对于任意的  $n_0 \in \{1, \dots, N\}$ ，我们将下载更新时长占用率  $\lambda_{n_0}^*$  用  $\lambda_{n_0}^* + \delta_{n_0}$  替代。由于函数  $h_n$  是严格递减的，我们得到  $h_n(\lambda_{n_0}^* + \delta_{n_0}) < h_n(\lambda_{n_0}^*)$ 。因此对于新的满足条件的解  $\lambda_{n_0}^\dagger = \lambda_{n_0}^* + \delta_{n_0}$ ,  $\lambda_n^\dagger = \lambda_n^*, \forall n \neq n_0$ ，这组解得到的平均信息年龄将比原  $\{\lambda_n^*\}$  更小，证明结束。 ■

#### 4.3.3.2 一种基于单调优化的求解方案

根据引理4.2，问题4.6中的约束条件均为线性（也是凸的），代价函数是单调递减的，因此该问题为一个单调优化问题<sup>[98-100]</sup>，接下来我们通过分支定界法 (Branch-Reduce-Bound, BRB)<sup>[100]</sup> 求解。注意到函数  $h_n(\lambda)$  在  $\lambda \rightarrow 0$  时会变得无界，而根据式 (4.53)，当  $\lambda \rightarrow 0$  时，状态信息将几乎不会被更新，它的信息年龄将会无限增长。因此，在初始时刻，我们令分支定界法需要从原点 0 处一个小的邻域开始搜索。

首先，我们给出搜索区域“盒子”(box) 的定义：

**定义 4.2 (盒子)：** 记  $\lambda = [\lambda_1, \dots, \lambda_N] \in \mathbb{R}^N$  为这  $N$  份状态信息的下载时间占比，记  $\mathbf{c} = [c_1, \dots, c_N] \in \mathbb{R}_+^N$  和  $\mathbf{d} = [d_1, \dots, d_N] \in \mathbb{R}_+^N$  为盒子的上下界，它们每个元素满足  $c_n \leq d_n, \forall n$ 。由向量  $\mathbf{c}, \mathbf{d}$  导出的  $N$  维搜索盒子  $\mathcal{B}$  为：

$$\mathcal{B} := \{\lambda | c_n \leq \lambda_n \leq d_n, \forall n.\}$$

根据引理4.2，代价函数式 (4.42) 单调递减，因此对于任意盒子  $\mathcal{B}$  中的向量  $\lambda \in \mathcal{B} = [\mathbf{c}^\mathcal{B}, \mathbf{d}^\mathcal{B}]$ ，我们可以给出函数  $h(\lambda) := \sum_{n=1}^N h_n(\lambda_n)$  的上下界，即  $h(\mathbf{d}^\mathcal{B}) =$

$$h_{\min}^{\mathcal{B}} \leq h(\lambda) \leq h_{\max}^{\mathcal{B}} = h(\mathbf{c}^{\mathcal{B}}).$$

接下来，我们给出分支定界法的具体步骤：记  $\mathcal{S}$  为所有搜索的盒子集合，初始化时， $\mathcal{S} \leftarrow \{[\boldsymbol{\delta}, \mathbf{1}]\}$ ，其中  $\boldsymbol{\delta} \in \mathbb{R}^{+N}$  为一个接近于 0 的  $N$  维矢量， $\mathbf{1}$  为  $N$  维全 1 矢量。

**分支 (Branch)** 这一步的目的是确定本轮搜索“盒子”的范围。首先我们从已有的盒子集合  $\mathcal{S}$  中选取信息年龄下界最小的盒子  $\mathcal{B}_{\min} = [\mathbf{c}, \mathbf{d}]$ ：

$$\mathcal{B}_{\min} = \arg_{\mathcal{B} \in \mathcal{M}} h_{\min}^{\mathcal{B}}. \quad (4.45)$$

接下来我们沿着盒子  $\mathcal{B}_{\min}$  最长的边  $i = \arg \max_n (d_n - c_n)$  将盒子  $\mathcal{B}_{\min}$  分为两半，我们将得到如下两个新盒子  $\mathcal{B}_1$  和  $\mathcal{B}_2$ ：

$$\mathcal{B}_1 = [\mathbf{c}, \mathbf{d} - \frac{1}{2}(d_i - c_i)\mathbf{e}_i], \mathcal{B}_2 = [\mathbf{c} + \frac{1}{2}(d_i - c_i)\mathbf{e}_i, \mathbf{d}], \quad (4.46)$$

其中  $\mathbf{e}_i$  为一个单位向量，其中第  $i$  个元素为 1 其余元素为 0。

同样的，我们接下来从集合  $\mathcal{S}$  选取信息年龄上界最大的盒子  $\mathcal{B}_{\max} = \arg_{\mathcal{B} \in \mathcal{S}} h_{\max}^{\mathcal{B}}$ ，用同样的方法将其沿着最长的边分为两个盒子  $\mathcal{B}_3$  和  $\mathcal{B}_4$ 。

**减支 (Reduce)** 接下来，我们将删除新盒子  $\{\mathcal{B}_1 \dots \mathcal{B}_4\}$  中那些一定不包含最优解的盒子。根据推论4.1，如果  $\sum_{n=1}^N (d_n - \frac{1}{2}(d_n - c_n)) < 1$ ，那么盒子  $\mathcal{B}_1$  不包含问题4.5约束区域的任意边界点，因此也不包含最优解  $\lambda^*$ ，我们将盒子  $\mathcal{B}_1$  从备选集合  $\mathcal{S}$  中移除；如果盒子的下界  $\sum_{n=1}^N (c_n + \frac{1}{2}(d_n - c_n)) > 1$ ，那么盒子  $\mathcal{B}_2$  在问题4.5的可行区域以外，它将同样被移除。我们对新盒子  $\mathcal{B}_3$  和  $\mathcal{B}_4$  采用同样的操作，并将那些  $\mathcal{B}_1 \sim \mathcal{B}_4$  中没有被移除的盒子放入备选集合  $\mathcal{S}$  中。

**定界 (Bound)** 如果备选盒子  $\mathcal{B}_2$  没有被移除，那么盒子  $\mathcal{B}_2$  中至少包含了问题4.5中的一个可行解。对于任意  $\lambda \in \mathcal{B}_2$ ，代价函数  $h(\lambda)$  的上界为  $h_{\max}^{\mathcal{B}_2} = h(\mathbf{c} + \frac{1}{2}(d_i - c_i)\mathbf{e}_i)$ ，因此，问题4.5的最优解一定满足  $h(\lambda^*) \leq h_{\max}^{\mathcal{B}_2}$ 。接下来，我们从备选集合  $\mathcal{S}$  中删除那些一定不包含最优解的盒子，即对于任意盒子  $\mathcal{A} = [\mathbf{a}, \mathbf{b}] \in \mathcal{S}$ ，如果  $h_{\min}^{\mathcal{A}} = h(\mathbf{b}) \geq h_{\max}^{\mathcal{B}_2}$ ，盒子  $\mathcal{A}$  一定不包含最优解  $\lambda^*$ ，我们可以把它从待搜索集合  $\mathcal{S}$  中移除。

在每一轮分支定界法结束之后，我们检查当前盒子中，最小和最大的差值是否满足  $\max_{\mathcal{B} \in \mathcal{S}} h_{\max}^{\mathcal{B}} - \min_{\mathcal{B} \in \mathcal{S}} h_{\min}^{\mathcal{B}} \leq \epsilon$ ，其中  $\epsilon$  为代表分支定界法的停止条件（即搜索精度）。迭代结束后，集合  $\mathcal{S}$  中剩余的盒子将位于最优解  $\lambda^*$  附近。接下来，求取这些盒子的平均值， $\lambda = \frac{1}{|\mathcal{S}|} \sum_{\mathcal{A}=[\mathbf{a},\mathbf{b}] \in \mathcal{S}} \frac{1}{2}(\mathbf{a} + \mathbf{b})$ ，我们可以得到最优下载时间占

用率  $\lambda^*$  的一组近似。接下来，每份状态信息  $n$  的平均更新间隔  $\bar{\tau}^* = [\bar{\tau}_1^*, \dots, \bar{\tau}_N^*]$  可以根据式  $g_n(\bar{\tau}_n^*) = \frac{f_n(\bar{\tau}_n^*)}{\bar{\tau}_n^*} = \lambda_n^*$  求解计算。

#### 4.3.3.3 基于 KKT 条件的求解方案

我们注意到  $h_n$  是由函数  $f_n$  决定的，如果  $f_n$  对应的函数  $h_n$  是凸的（例如节4.3.3.4 中考虑的函数），我们可以通过 KKT 条件找到  $\{\lambda_n^*\}$ ，这一算法相比于分支定界法求解复杂度较低。在本节中，我们主要讨论对于这些凸函数  $h_n$ ，使用 KKT 条件求最优  $\{\lambda_n^*\}$ 。

定义拉格朗日函数：

$$\mathcal{L}(\lambda_1, \dots, \lambda_N, v) = \sum_{n=1}^N h_n(\lambda_n) + v \left( \sum_{n=1}^N \lambda_n - 1 \right),$$

其中  $v \geq 0$  表示拉格朗日乘子。KKT 条件要求最优解  $\{\lambda_n^*\}$  和对偶乘子  $v^*$  必须满足：

$$p_n h'_n(\lambda_n^*) + v^* = 0, \quad \forall n \quad (4.47)$$

$$v^* \left( \sum_{n=1}^N \lambda_n^* - 1 \right) = 0, \quad (4.48)$$

其中  $h'_n$  为函数  $h_n$  的一阶导数：

$$h'_n(\lambda) := g_n^{(-1)'}(\lambda) \left( \frac{1}{2} + \lambda \right) + g_n^{-1}(\lambda). \quad (4.49)$$

由于函数  $h'_n$  可逆，原函数  $h_n$  连续单调递减，逆函数  $h_n'^{(-1)}$  的像为非正实数  $(-\infty, 0]$ 。我们可以将 KKT 条件式 (4.47) 重写为：

$$\lambda_n^* = h_n'^{(-1)} \left( -\frac{v^*}{p_n} \right), \quad \forall n. \quad (4.50)$$

根据推论4.1，KKT 条件中式 (4.48) 被转化为对最优解的约束  $\sum_{n=1}^N \lambda_n^* = 1$ 。当最优对偶乘子  $v^*$  已知时，原问题最优解  $\lambda_1^*, \dots, \lambda_N^*$  可以由式 (4.50) 直接计算得到，而对偶乘子  $v^* \geq 0$  的选取还需要满足条件  $\sum_{n=1}^N \lambda_n^* = 1$ 。对于一些给定的函数  $f_n$  来说，我们可以找到其一阶导的逆函数  $h_n'^{(-1)}$  的闭式表达形式；对于另一些没有闭式表达式的函数  $h_n$ ，我们可以通过列写函数表的方法计算逆函数  $h_n'^{(-1)}$  在不同点的取值。

通过变量替换，我们可以将 KKT 条件用最优更新间隔写出  $\bar{\tau}_n^* = g_n^{(-1)}(\lambda_n^*)$ 。其中，根据  $g_n^{(-1)'}(\lambda_n^*) = \frac{1}{g_n'(\bar{\tau}_n^*)} = \frac{(\bar{\tau}_n^*)^2}{f_n'(\bar{\tau}_n^*)\bar{\tau}_n^* - f_n(\bar{\tau}_n^*)}$ ，式 (4.50) 等价为：

$$\bar{\tau}_n^* + \frac{(\bar{\tau}_n^*)^2}{f_n'(\bar{\tau}_n^*)\bar{\tau}_n^* - f_n(\bar{\tau}_n^*)} \left( \frac{1}{2} + \frac{f_n(\bar{\tau}_n^*)}{\bar{\tau}_n^*} \right) = -\frac{v^*}{p_n}, \quad \forall n, \quad (4.51)$$

其中  $f'_n$  为  $f_n$  的一阶导数，以上方程可以根据  $f_n$  的表达式直接求解。

#### 4.3.3.4 一个基于函数 $f_n$ 求解的典型例子

本部分中我们针对一种特定形式的  $f_n$  给出基于 KKT 条件的求解方案，这一求解方案将广泛应用于之后的仿真中。我们假设状态信息  $n$  的一部分内容由于时间流逝过时了，基站端缓存和远程服务器都知道这些过时的部分。这些过时的内容可以被描述为擦除掉的比特，我们用一个参数为  $\Delta_n$  的二元擦除信道（Binary Erasure Channel, BEC）描述状态信息  $n$  的变化和演进。假设状态信息  $n$  最初有  $B_n$  比特，那么在  $t$  个时间单位之后，他的每个比特都经历了参数为  $\Delta_n$  的  $t$  次二元擦除信道变化。 $t$  次变化之后的比特转移可以被写成是  $1 - (1 - \Delta_n)^t$ ， $t$  个单位时间之后，被擦除掉的平均比特数为  $B_n(1 - (1 - \Delta_n)^t)$ 。

注意到当  $\lim_{t \rightarrow 0} f_n(t) = 0$  时，解  $\bar{\tau}_n^* = 0$  显然最优，但是这并不现实因为实际更新总有因为信令和信道估计等带来的固定时延。考虑到这一特点，我们假设状态信息更新时长的信息年龄的关系为：

$$f_n(t) = B_n - (B_n - \epsilon_n)e^{-\beta_n t} \quad (4.52)$$

其中  $\beta_n = -\log(1 - \Delta_n)$ ,  $\epsilon_n$  为因为信令和信道估计带来的固定时延。注意到当  $\beta_n > 0$  时，函数  $f_n$  满足假设4.1。

对于式 (4.52) 所示的下载时长函数，更新时长占比  $g_n(\cdot)$  的逆函数为：

$$g_n^{(-1)}(\lambda) = \frac{B_n}{\lambda} + \frac{1}{\beta_n} W\left(-\frac{\beta_n(B_n - \epsilon_n)}{\lambda} e^{-\frac{\beta_n B_n}{\lambda}}\right), \quad (4.53)$$

其中  $W(\cdot)$  为 Lambert 函数。我们注意到对于一些特定的  $\epsilon_n, B_n, \beta_n$  取值来说（例如节4.3.5中的仿真条件），函数  $h_n = g_n^{(-1)}(\lambda)(\frac{1}{2} + \lambda)$  是凸的。在这样的情况下，我们可以直接使用 KKT 条件求取问题4.6最优解。对于其他的  $\epsilon_n, B_n, \beta_n$  取值 ( $B_n = 1, \epsilon_n = 0.02$  以及  $\beta_n = 10$ )，函数  $h_n$  非凸，我们用4.3.3.2求取松弛问题4.6最优解。

#### 4.3.4 一种可以满足带宽约束的更新策略

接下来的问题是，我们如何设计一种可以满足原问题4.4中带宽约束的基站端缓存更新方案？即每一次更新下载必须在上一次更新下载结束后才能开始。

为了方便我们对于更新策略的描述，我们记  $\{\lambda_n^*\}$  为问题4.6的最优解，它可以通过分支定界法或者 KKT 条件求解得到。我们定义状态信息  $n$  的下载时间占比所对应的更新时间间隔为  $\bar{\tau}_n^* := g_n^{(-1)}(\lambda_n^*)$ 。如果在时间  $t$  时，基站端缓存刚结束了上一份状态信息的更新，它将立刻选取当前最亟需被更新的状态信息并立刻开始更

新和下载，即选取状态

$$n_0(t) = \arg \min_{n \in \{1, \dots, N\}} \bar{\tau}_n^* - x_n(t).$$

### 4.3.5 数值仿真

注意到当状态信息更新时长为常数时，平方律更新策略<sup>[62]</sup>可以最小化用户得到状态信息副本的平均信息年龄。本部分中，我们将通过数值仿真，比较所提更新策略、松弛问题的策略以及平方律策略得到的平均信息年龄<sup>①</sup>。当更新时长为常数时，最优的更新时间占比  $\{\lambda_n^*\}$  可以由平方律关系得到：

$$\lambda_n^* = \lambda_n^{\text{sqrt}} = \frac{\sqrt{p_n}}{\sum_{i=1}^N \sqrt{p_i}}. \quad (4.54)$$

我们考虑以下两种更新函数进行仿真：(1)  $f_n(t) = B_n$ ；(2) 式 (4.52) 给出的更新时长与信息年龄相关函数  $f_n(t)$ 。在接下来的仿真图中，蓝线和红线表示使用平方律策略得到的仿真结果，黄线和紫线表示通过求解问题4.6本章所提基站端缓存更新策略的仿真结果。其中实线表示求解松弛问题得到的平均信息年龄，虚线表示使用节4.3.4所提出满足带宽约束的下载更新算法得到的平均信息年龄。

#### 4.3.5.1 更新时长 $f_n(t)$ 与状态信息 $n$ 有关但是与信息年龄无关

我们首先考虑函数  $f_n(t) = B_n$ 。在这样的情况下，函数  $g_n(t) = B_n/t$ ,  $g_n^{(-1)}(\lambda) = B_n/\lambda$ 。函数  $h_n(t)$  可以写作：

$$h_n(\lambda) = \frac{B_n}{2\lambda} + B_n.$$

我们注意到函数  $h_n(t)$  是凸的，因此可以通过 KKT 条件 (4.50) 直接得到松弛问题4.6对应的最优下载时间占比  $\lambda_n^*$ ：

$$\lambda_n^* = \frac{\sqrt{p_n B_n}}{\sum_{i=1}^N \sqrt{p_i B_i}}. \quad (4.55)$$

式 (4.55) 与平方律更新策略式 (4.54) 的不同点在于：更新时间占比的计算中加入了更新时长的权重  $B_n$ 。根据式 (4.55)，对于任意两份更新时长相同的状态信息，被索取概率更高（即  $p_n$  较大）的状态信息将有更大的更新时间占比；对于两个索取概率相同的状态信息，每次更新时长较长的信息将有更大的更新时间占比。注意到状态信息的更新频率（即在一定时间  $T$  内被更新的次数）等于  $\lambda_n/B_n$ ，实际所需更新时长更长的状态信息对应的更新频率将会更低一些。

我们考虑  $N$  份状态信息可以被分成两类的情况，对于第一类状态信息  $n \in$

<sup>①</sup> 文献<sup>[62]</sup>中的优化变量为更新频率，在每次更新所用时间都是常数这种情况下，更新频率与本研究中的更新时间占比两概念等价。

$\{1, 2, \dots, N/2\}$ , 它们的更新间隔  $B_n = 1$ , 被索取概率  $p_n \propto 1/n^\alpha$ ; 对于第二类状态信息  $n \in \{N/2 + 1, 2, \dots, N\}$ , 它们的更新间隔  $B_n = 5$ , 被索取概率为  $p_n = p_{n-N/2} \propto 1/(n - N/2)^\alpha$ 。也就是说, 每一类状态信息被索取的概率服从一个参数为  $\alpha$  的 Zipf 分布。仿真中我们选取参数  $\alpha = 1.8$ 。

图4.4我们画出使用不同策略时, 用户所得到的状态信息平均信息年龄  $J(\pi)$  随着状态信息总数  $N$  的变化曲线。我们注意到, 所提算法得到的平均信息年龄比平方律更新策略得到的信息年龄小。比如在  $N = 50$  这个点, 我们的算法可以降低平均信息年龄约 10%。同时, 所提更新策略得到信息年龄与松弛问题4.6最优解之间的差别很小。

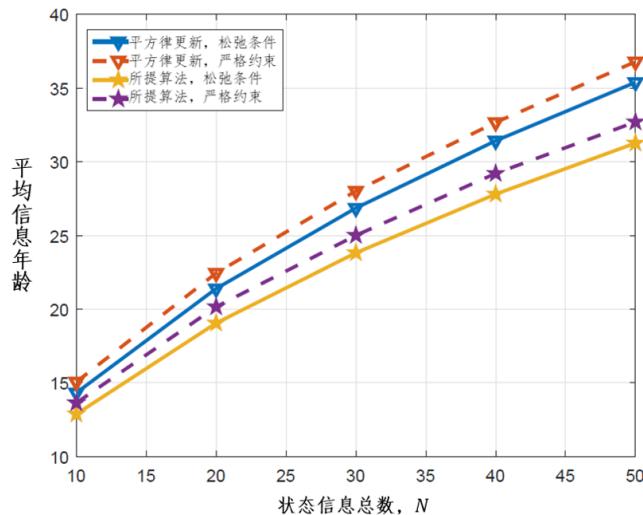


图 4.4 每份状态信息更新时长为不同的常数时, 用户索取得到状态信息的平均信息年龄随状态信息数量  $N$  演进曲线图

#### 4.3.5.2 更新时长 $f_n(t)$ 与状态信息和信息年龄均有关

在本部分的仿真中, 我们考虑如式 (4.52) 所示的更新函数  $f_n$ 。对于所有的状态信息  $n$ , 我们选取参数  $\varepsilon_n = 0.02, B_n = 1$ 。其中, 我们考虑 (1) 参数  $\beta_n$  均一样但是被索取概率  $p_n$  不同; (2) 每份状态信息对应的  $\beta_n$  不同这两种场景。本部分中我们同样考虑状态信息被索取的概率服从参数  $\alpha = 1.8$  的 Zipf 分布, 并通过节4.3.3.2中的分支定界法求解松弛问题4.6的最优更新时间占比。

图4.5考虑的场景中, 参数  $\beta_n = 0.015, \forall n$ , 也就是说所有状态信息更新时长随信息年龄变化的函数都是一样的。我们画出了被索取状态信息的信息年龄随着基站端缓存中状态信息总数  $N$  的变化曲线。当  $N = 5$  时, 所提出算法相比于平方律更新策略可以降低被用户索取得到状态信息的信息年龄约 50%。

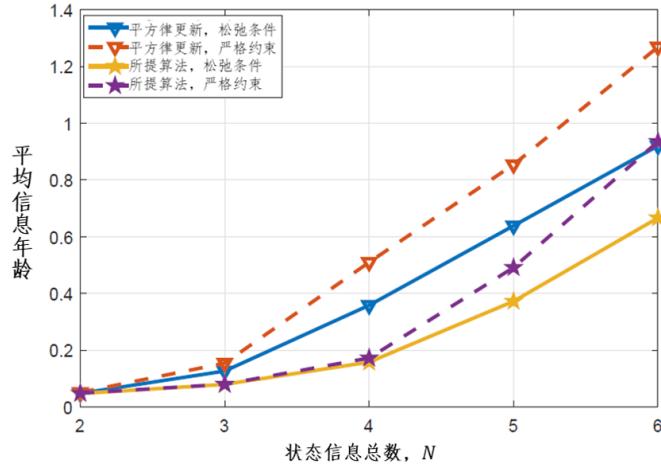


图4.5 每份状态信息更新时长时变时，被索取状态信息平均信息年龄随状态信息数量  $N$  的演进曲线图

图4.6画出了  $N = 5$  时，每份状态信息的更新时间占比。同样我们选取参数  $\beta_n = 0.015, \forall n$ 。我们注意到问题4.6所得到的更新时间占比相比于平方律更新有显著不同。事实上，那些被索取概率较大的状态信息由于较为频繁地被更新，他们的平均更新时长也更短，由图可见，它们的更新时间占比其实相对较小。

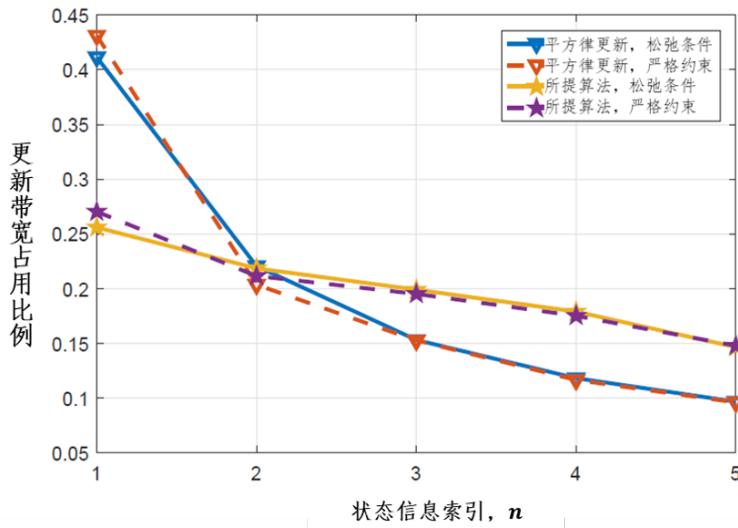


图4.6 不同状态信息下载和更新所占用的带宽

图4.7研究了状态信息被索取概率相同即  $(p_n = 1/N, \forall n)$ ，但是更新时长随信息年龄变化关系不同时，所提算法分配给每份状态信息的更新时间占比。我们考虑  $N = 5$ ，对于每份状态信息  $n$ ，更新函数变化参数  $\beta_n = 0.1 \times n, \forall n$ 。注意到对于编号较大的状态信息  $n$ ，参数  $\beta_n$  相对较大，根据式 (4.52)，更新函数  $f_n$  几乎

为一个常数。对于这些状态信息来说，所提算法的更新时间占比与平方律算法几乎相同（其中平方律算法在更新时间为常数的情况下就是最优的）。而对于编号较小的状态信息  $n$ ，参数  $\beta_n$  较小，函数  $f_n$  在信息年龄较小的情况下严格递增。对于这些状态信息来说，当它们的更新间隔较短时，每次更新所需要的时长也比较短，因此我们的算法所分配的更新时间占比显著地小于平方律分配策略；对于那些编号处于中间地带的状态信息  $n$ ，所提算法给它们分配的更新时间占比较大。这是因为一方面，这些状态信息相对于那些  $\beta$  较大的状态信息来说更新频率更高，另一方面相对于那些  $\beta$  较小的状态信息来说，更新它们每次所需要的时长也比较大。

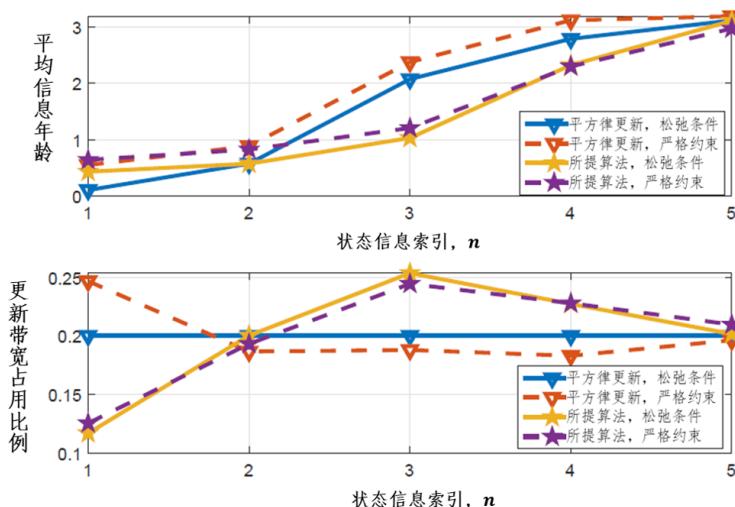


图 4.7 不同状态信息的更新时间占比（上图）和平均信息年龄

#### 4.3.6 本节小结

本节中我们针对基站端缓存更新状态信息所需时长随信息年龄变化这一场景，提出了在带宽约束下，减小用户获得状态信息副本平均信息年龄的有效方案。

### 4.4 本章小结

本章面向基站端缓存系统，在多个假设下探究了最小化用户获得状态信息副本平均信息年龄的缓存更新方案。首先我们考虑状态信息的下载传输时间为常数，针对状态信息流行度随时间变化这一场景，我们提出了一种状态信息趋于无穷多时渐近最优的基站端缓存更新方案；接下来，我们考虑状态信息的下载时长随时间变化，将传输时间建模为当前信息年龄的函数，我们给出了一种在数值仿真中接近最优的缓存更新方案。

## 第 5 章 状态更新随机时同步年龄最小化的推送算法研究

### 5.1 引言

本章在第四章的基础上，考虑基站将保存的路况、故障等随机变化的状态信息通过广播推送给用户。如节1.2所述，当状态信息演进由外部过程控制时，使用信息年龄度量接收端信息时效性存在一定局限性，这要求我们寻找更加合理的度量，并设计调度算法，在状态更新和传输双重随机性下，保证用户时效性。

本章选用同步年龄<sup>[6]</sup>度量状态信息变化随机时，用户所保存状态信息的信息时效性。考虑传输带宽限制，依据用户信道状态和信源更新特点，设计调度策略，针对最小化全网用户同步年龄这一问题展开研究。本章首先推导了全网同步年龄的理论下界，并基于马尔可夫决策过程和无休止多臂赌博机提出了两种调度推送方案，仿真中所提算法得到的全网平均同步年龄接近理论下界。

### 5.2 系统模型

#### 5.2.1 网络模型

如图，基站保存有  $N$  份状态信息，并将这些信源的状态更新信息推送给  $N$  个用户。我们假设用户  $n$  只对状态信息  $n$  感兴趣。考虑离散时间系统，并使用  $t \in \{1, \dots, T\}$  表示当前时隙的索引。在每时隙  $t$ ，信源  $n$  以相同的概率  $\lambda_n \in (0, 1]$  产生新的状态更新<sup>①</sup>。我们用示性函数  $\Lambda_n(t) \in \{0, 1\}$  记录信源  $n$  在时隙  $t$  是否有更新产生。如果  $\Lambda_n(t) = 1$ ，时隙  $t$  中信源产生了一次更新，基站将可以在时隙  $(t + 1)$  开头将更新推送给用户。

在每时隙开始，基站通过无线信道将信源的状态更新发送给用户。考虑用户只对信源最新的状态感兴趣，基站仅需保存每个信源最新状态更新的数据包。即对于每个信源来说，陈旧的状态信息将被新的状态信息替换掉。我们用  $u_n(t) \in \{0, 1\}$  表示时隙  $t$  是否推送信源  $n$ 。如果时隙  $t$ ，基站没有发送用户  $n$  所感兴趣的信源，那么  $u_n(t) = 0$ ；否则当  $u_n(t) = 1$  时，信源  $n$  的最新更新将被基站传输给用户，如果传输成功，用户  $n$  将在时隙  $t$  末尾接收到相应的数据包。我们假设从基站到用户  $n$  的传输服从概率  $p_n$  的伯努利分布。时隙  $t$  的末尾，用户  $n$  将传输结果无差错地反馈给基站。

---

<sup>①</sup> 当  $\lambda_n = 1$  时，信源  $n$  每时每刻都将产生更新。此时同步年龄和信息年龄等价。在之后的推导中，我们也将比较在  $\lambda_n = 1$  场景下同步年龄和信息年龄的推导结果，用于阐述两种度量之间联系。

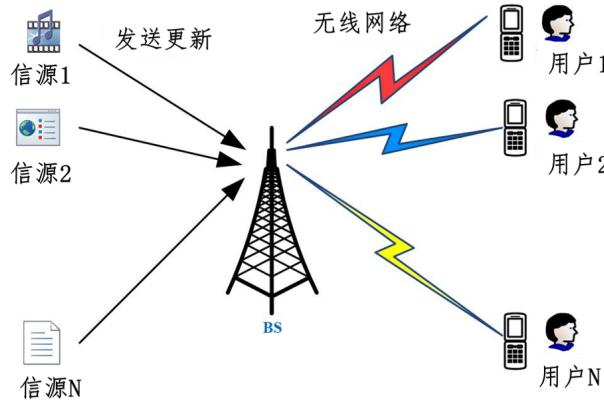


图 5.1 基站向用户推送状态信息系统图

与文献[28,33,62]类似，本节中我们考虑基站和用户之间的通信带宽受限，每时刻基站只能推送一个信源的更新给用户，也就是说调度策略  $\{u_n(t)\}$  的设计需满足：

$$\sum_{n=1}^N u_n(t) \leq 1, \forall t. \quad (5.1)$$

### 5.2.2 信息年龄与同步年龄

为了介绍和比较信息年龄与同步年龄之间的联系和差异，我们将首先考虑一个点对点信源-接收端通信链路。我们将首先回顾同步年龄和信息年龄的定义，并介绍本问题中同步年龄随着调度决策  $\{u_n(t)\}$  和外部更新过程  $\{\Lambda_n(t)\}$  的演进。假设第  $i$  个状态更新数据包产生于时隙  $g_i$ ，记  $r_i$  为该数据包被成功接收的时刻，如果因为丢包等传输过程中的随机性，数据包  $i$  未能被接收端成功接受，我们记  $r_i = +\infty$ 。

记  $q(t) = \max_{i \in \mathbb{N}^+} \{i | r_i < t\}$  为接收端在时隙  $t$  所拥有的最新更新的索引，他的产生时间为  $g_{q(t)}$ 。在时隙  $t$  开头，信息年龄  $x(t)$  的计算公式为：

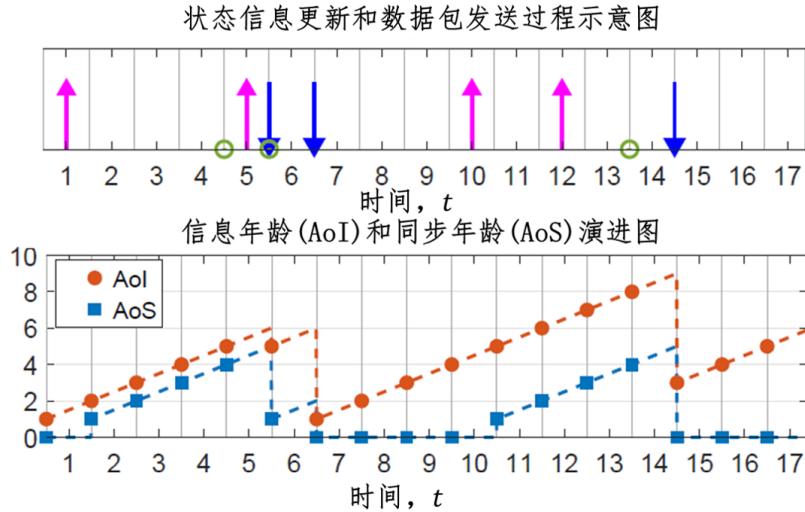
$$x(t) = t - g_{q(t)}. \quad (5.2)$$

同步年龄刻画了接收端与信源不同步的时长[6]。接收端所拥有的最新信息索引为  $q(t)$ ，那么  $q(t) + 1$  即为时隙  $g_{q(t)}$  之后状态信息首次更新。根据定义，在时隙  $t$  开头接收端的同步年龄  $s(t)$  为：

$$s(t) = (t - g_{q(t)+1})^+, \quad (5.3)$$

其中函数  $(\cdot)^+ = \max\{0, \cdot\}$ 。根据式 (5.3)，如果时刻  $g_{q(t)}$  到当前时刻  $t$ ，信源  $n$  没有产生新的状态更新，即  $g_{q(t)+1} \geq t$ ，那么在时隙  $t$ ，接收端的最新信息与当前信源同步，即  $s(t) = 0$ 。为了方便比较信息年龄和同步年龄的不同，我们将它们的演进曲线画在图5.2中。如图5.2所示，同步年龄在下一次更新到达之前将保持为 0，一

一旦有新的更新产生，同步年龄将线性增长，直到接收到最新的状态更新；而信息年龄在没有接收到新的数据包更新前将持续线性增长。



上图：更新产生时刻（粉色向上箭头）、更新决策（绿色圆圈）和信源更新实际被收到时刻的序列（蓝色向下箭头）；下图：信息年龄和同步年龄的演进曲线。

图 5.2 信息年龄和同步年龄演进曲线对比图

接下来我们回到多用户的场景，并介绍本文所考虑的同步年龄的演进关系。记  $s_n(t)$  为用户  $n$  在时隙  $t$  开始时的同步年龄。我们将依据信源同步状态分类讨论：

- 我们首先考虑时隙  $t$  开始时，用户  $n$  保存的信息与信源  $n$  状态同步，也就是说  $s_n(t) = 0$ 。下一时刻的同步年龄  $s_n(t+1)$  将由时隙  $t$  中是否有更新产生决定：
  - \* 如果信源在时隙  $t$  没有更新产生，即  $\Lambda_n(t) = 0$ ，那么下一时刻信源  $n$  和用户  $n$  仍然同步，下一时刻的同步年龄  $s_n(t+1) = 0$ 。
  - \* 如果信源在时隙  $t$  内产生了更新，即  $\Lambda_n(t) = 1$ ，在时隙  $t+1$  的一开始，用户  $n$  处关于信源  $n$  的信息将立刻不同步，即  $s_n(t+1) = 1$ 。
- 如果用户  $n$  和信源  $n$  在时隙  $t$  开始已经不同步，即  $s_n(t) \neq 0$ ：
  - + 如果时隙  $t$ ，基站决定广播用户  $n$  所感兴趣的更新，即  $u_n(t) = 1$  并且传输成功，那么信源  $n$  在时隙  $t-1$  末尾的最新状态将在时隙  $t$  末尾被用户  $n$  收到。在这种情况下：
    - \* 如果  $\Lambda_n(t) = 0$ ，时隙  $t$  内信源  $n$  没有产生更新，因此用户  $n$  接收到的信息即为信源的最新状态，在时隙  $(t+1)$  开始的同步年龄  $s_n(t+1) = 0$ 。
    - \* 如果  $\Lambda_n(t) = 1$ ，由于信源  $n$  产生了新的更新，用户在时隙  $t$  末尾接收到的信息将立刻变得不同步，时隙  $t+1$  开始时，用户的同步年龄

$$s_n(t+1) = 1.$$

- + 如果时隙  $t$  基站决定不推送用户  $n$  所感兴趣的更新（即  $u_n(t) = 0$ ）或者传输失败，用户  $n$  将无法获得最新的数据包，根据定义他的同步年龄将线性增长，在时隙  $t+1$  开头用户  $n$  的同步年龄即为： $s_n(t+1) = s_n(t) + 1$ 。

根据以上分析，本研究中用户  $n$  的同步年龄演进关系如下所示：

$$s_n(t+1) = \begin{cases} 0, & s_n(t) = 0, \Lambda_n(t) = 0; \\ 1, & s_n(t) = 0, \Lambda_n(t) = 1; \\ 0, & \Lambda_n(t) = 0, u_n(t) = 1, \text{传输成功}; \\ 1, & \Lambda_n(t) = 1, u_n(t) = 1, \text{传输成功}; \\ s_n(t) + 1, & \text{其他。} \end{cases} \quad (5.4)$$

### 5.2.3 问题建模

我们用  $J_T(\pi)$  表示使用策略  $\pi$  时，在连续  $T$  时隙内全网用户的平均同步年龄，即

$$J_T(\pi) \triangleq \mathbb{E}_\pi \left[ \frac{1}{NT} \sum_{t=1}^T \sum_{n=1}^N s_n(t) | \mathbf{s}(0) \right],$$

其中向量  $\mathbf{s}(t) = [s_1(t), s_2(t), \dots, s_N(t)]^T \in \mathbb{N}^N$  表示时隙  $t$  开始时全网所有用户的同步年龄。我们假设在  $t = 0$  时刻，所有用户已经和信源同步，即  $\mathbf{s}(0) = 0$ 。在接下来的讨论中，我们将忽略  $\mathbf{s}(0)$  的影响。

记  $\Pi$  为所有因果策略的集合（因果策略的定义详见定义2.1），即时隙  $t$  的调度策略仅依赖于信道传输概率的统计信息  $\{p_n\}$ 、过去和现在用户的同步年龄  $\{s_n(\tau)\}_{\tau < t}$ 。未来的丢包情况、信源更新情况和同步年龄不可以用于决策。我们的目标是设计一个非因果策略  $\pi \in \Pi$  优化全网平均同步年龄，优化问题如下：

**问题 5.1 (同步年龄最小调度问题):**

$$\pi^\star \triangleq \arg \min_{\pi \in \Pi} \limsup_{T \rightarrow \infty} J_T(\pi), \quad (5.5a)$$

$$\text{其中 } J_T(\pi) = \mathbb{E}_\pi \left[ \frac{1}{NT} \sum_{t=1}^T \sum_{n=1}^N s_n(t) \right], \quad (5.5b)$$

$$\text{s.t. } \mathbb{E}_\pi \left[ \sum_{n=1}^N u_n(t) \right] \leq 1, \forall t. \quad (5.5c)$$

### 5.2.4 全网平均同步年龄的理论下界

本节中我们将计算同步年龄的理论下界以方便之后的算法评估。

**定理 5.1:** 考虑网络将更新频率为  $\{\lambda_n\}$  的状态信息推送给信道统计信息为  $\{p_n\}$  多用户，任意因果调度算法的平均同步年龄下界为：

$$\text{AoS}_{\text{LB}} = \frac{1}{N} \sum_{n=1}^N \gamma_n^* \left[ \frac{1}{2} \left( \frac{1}{\gamma_n^*} - \frac{1-\lambda_n}{\lambda_n} \right)^2 + \frac{1}{2} \left( \frac{1}{\gamma_n^*} - \frac{1-\lambda_n}{\lambda_n} \right) \right], \quad (5.6)$$

其中  $\gamma_n^* = \max\{1/\sqrt{\left(\frac{1-\lambda_n}{\lambda_n}\right)^2 - \left(\frac{1-\lambda_n}{\lambda_n}\right) + \frac{2\mu^* N}{p_n}}, \lambda_n\}$ , 参数  $\mu^*$  的选取需要满足  $\sum_{n=1}^N \frac{\gamma_n^*}{p_n} = 1$ 。

定理的详细证明见附录5.7.1。

### 5.3 调度策略 1：基于马尔可夫决策过程的调度方案设计

本节中我们将依据马尔可夫决策过程设计一个优化全网同步年龄的调度策略。问题5.1可以被建模为一个马尔可夫决策过程，它的要素如下所示：

- **状态空间  $S$ :** 为时隙  $t$  开始时全网所有用户的同步年龄  $\mathbf{s}(t)$ 。本问题中状态空间可数，但由于传输可能失败所以状态空间无限大。
- **决策空间  $A$ :** 我们记时隙  $t$  的决策  $a(t)$  为被选择的用户索引，决策  $a(t)$  与调度策略的转换关系为  $u_n(t) = 1_{(n=a(t))}, \forall n$ 。如果基站选择不推送任何信源的更新，我们记  $a(t) = 0$ 。决策空间  $A = \{0, 1, 2, \dots, N\}$  有限可数。
- **状态转移概率:** 记  $\Pr(\mathbf{s}'|\mathbf{s}, a)$  为选取决策  $a$  时状态  $\mathbf{s}(t) = \mathbf{s} = [s_1, s_2, \dots, s_N]^T$  转移到状态  $\mathbf{s}(t+1) = \mathbf{s}' = [s'_1, s'_2, \dots, s'_N]^T$  的概率。由于每时刻每个信源  $n$  的更新以及基站发送信息给用户  $n$  是否成功彼此独立，这一解耦后的转移概率为：

$$\Pr(\mathbf{s}'|\mathbf{s}, a) = \prod_{n=1}^N \Pr(s'_n|s_n, a), \quad (5.7)$$

其中  $\Pr(s'_n|s_n, a)$  表示决策为  $a$  时用户  $n$  同步年龄的转移，依据式 (5.4) 我们

有：

$$\Pr(s'_n | s_n, a) = \begin{cases} 1, & s'_n = s_n + 1, s_n \neq 0, a \neq n; \\ 1 - p_n, & s'_n = s_n + 1, s_n \neq 0, a = n; \\ \lambda_n p_n, & s'_n = 1, s_n \neq 0, a = n; \\ (1 - \lambda_n)p_n, & s'_n = 0, s_n \neq 0, a = n; \\ \lambda_n, & s'_n = 1, s_n = 0, \forall a; \\ 1 - \lambda_n, & s'_n = 0, s_n = 0, \forall a; \\ 0, & \text{其他.} \end{cases} \quad (5.8)$$

- **单步代价函数：**记  $C(\mathbf{s}(t), a(t))$  为状态  $\mathbf{s}(t)$ 、决策  $a(t)$  时的单步代价函数，即全网用户在时隙  $t$  的平均同步年龄。

$$C(\mathbf{s}(t), a(t)) = \frac{1}{N} \sum_{n=1}^N s_n(t).$$

问题5.1中的最优策略  $\pi^*$  可以通过求解以上马尔可夫决策过程得到。为了使用值迭代求解以上马尔可夫决策过程，我们记  $J_\alpha(\mathbf{s}, \pi)$  从状态  $\mathbf{s}$  出发，使用策略  $\pi$  得到的  $\alpha$ -累计折扣代价期望：

$$J_\alpha(\mathbf{s}, \pi) := \lim_{T \rightarrow \infty} \mathbb{E}_\pi \left[ \sum_{t=1}^T \alpha^{t-1} C(\mathbf{s}(t), a(t)) \right], \quad (5.9)$$

接下来，我们将在  $\alpha \rightarrow 1$  时，通过最小化  $J_\alpha(\mathbf{s}, \pi)$  设计一个接近  $\pi^*$  的策略。记  $\pi_\alpha^*$  最小化  $\alpha$  折扣代价的调度策略，即：

$$\pi_\alpha^* := \arg \min_{\pi \in \Pi} J_\alpha(\mathbf{s}, \pi), \forall \mathbf{s}, \quad (5.10a)$$

$$\text{s.t. } \mathbb{E}_\pi \left[ \sum_{n=1}^N u_n(t) \right] \leq 1, \forall t. \quad (5.10b)$$

可以证明，存在一个静态确定性策略最小化  $\alpha$ -累计折扣代价。接下来，我们将利用最优解的阈值结构，通过值迭代求解  $\pi_\alpha^*$ 。记  $\pi_\alpha^*(\mathbf{s})$  为策略  $\pi_\alpha^*$  在时隙  $t$ ，状态为  $\mathbf{s}$  时所选取的决策。记  $V_\alpha(\mathbf{s})$  为从状态  $\mathbf{s}$  出发，使用策略  $\pi_\alpha^*$  得到的  $\alpha$ -累计折扣代价，即：

$$V_\alpha(\mathbf{s}) \triangleq \lim_{T \rightarrow \infty} \mathbb{E}_{\pi_\alpha^*} \left[ \sum_{t=1}^T \alpha^{t-1} C(\mathbf{s}(t), a(t)) \right] = \min_{\pi} J_\alpha(\mathbf{s}, \pi). \quad (5.11)$$

值函数  $V_\alpha(\cdot)$  满足以下贝尔曼方程:

$$V_\alpha(\mathbf{s}) = \min_{a \in \mathbb{A}} \{C(\mathbf{s}, a) + \alpha \sum_{\mathbf{s}'} V_\alpha(\mathbf{s}') \Pr(\mathbf{s}' | \mathbf{s}, a)\}. \quad (5.12)$$

对于状态空间有限可数的情况，我们可以直接通过策略迭代求解得到最优策略  $\pi_\alpha^\star$ 。对于本问题中状态空间无限的情况，我们通过给每个用户的同步年龄设置最大值  $S_n^{\max}$ ，对原马尔可夫决策过程进行近似求解。这一近似的合理性在于连续传输失败的概率随着传输尝试的次数的增加指数衰减，因此当每个用户的截断上界  $S_n^{\max}$  逐渐增大，截断后马尔可夫决策过程的最优策略将趋向于  $\pi_\alpha^\star$ 。

简单起见，我们为全网所有用户选取相同的截断阈值  $S_n^{\max} = m$ ，记  $\tilde{s}_n(t)$  为截断之后用户  $n$  的同步年龄，即  $\tilde{s}_n(t) = \min\{s_n(t), m\}$ 。接下来，我们依据截断后的同步年龄  $\tilde{s}_n$  构建新的马尔可夫决策过程

- **状态空间:** 我们将原状态  $\mathbf{s}(t)$  用截断后的同步年龄  $\tilde{\mathbf{s}}(t) = [\tilde{s}_1(t), \tilde{s}_2(t), \dots, \tilde{s}_N(t)]^T$  进行替换。
- **概率转移函数:** 由于同步年龄被截断，状态转移函数相应地发生变化，记  $\Pr(\tilde{\mathbf{s}}' | \tilde{\mathbf{s}}, a)$  为状态  $\tilde{\mathbf{s}}(t) = \tilde{\mathbf{s}}$  向状态  $\tilde{\mathbf{s}}(t+1) = \tilde{\mathbf{s}}'$  转移的概率。同样地，这一概率可以被解耦写成:

$$\Pr(\tilde{\mathbf{s}}' | \tilde{\mathbf{s}}, a) = \prod_{n=1}^N \Pr(\tilde{s}'_n | \tilde{s}_n, a). \quad (5.13)$$

其中，转移概率  $\Pr(\tilde{s}'_n | \tilde{s}_n, a)$  与式 (5.8) 的  $\Pr(s'_n | s_n, a)$  除了以下情况外完全相同:

$$\Pr(\tilde{s}'_n | \tilde{s}_n, a) = \begin{cases} 1, & \tilde{s}'_n = \tilde{s}_n = m, a \neq n; \\ 1 - p_n, & \tilde{s}'_n = \tilde{s}_n = m, a = n. \end{cases} \quad (5.14)$$

对于给定的截断阈值  $m$ ，我们可以通过值迭代得到一个最优的确定性调度策略。记策略迭代的初始策略为  $\pi^{(0)}(\tilde{\mathbf{s}}) = \arg \max_{\tilde{s}_n} \tilde{s}_n$ ，即总是调度截断后同步年龄最大的用户这样的贪心策略。给定第  $k$  轮迭代后得到的策略  $\pi^{(k)}(\tilde{\mathbf{s}})$  和值函数  $V_\alpha^{(k)}(\tilde{\mathbf{s}})$ ，第  $(k+1)$  轮的策略  $\pi^{(k+1)}(\tilde{\mathbf{s}})$  与值函数  $V_\alpha^{(k+1)}(\tilde{\mathbf{s}})$  可以通过策略迭代计算。随着迭代轮次  $k$  增加，策略  $\pi^{(k)}(\mathbf{x})$  和值函数  $V_\alpha^{(k)}(\mathbf{x})$  将最终收敛。算法流程图见流程图5.1所示:

通过以上策略迭代算法我们可以得到一个静态确定性策略  $\pi$ 。基于这一策略的调度算法如下：在每时隙  $t$  基站计算所有用户的同步年龄  $\mathbf{s}(t)$  以及截断之后的  $\tilde{\mathbf{s}}(t)$ ，根据策略  $\pi$  调度用户  $a(t) = \pi(\tilde{\mathbf{s}}(t))$ 。

这一算法的主要问题是可能的状态总计有  $m^N$  个，因此计算复杂度随着用户数  $N$  增加指数爆炸。下一节我们将针对大规模网络场景设计一种低复杂度调度方案。

**算法 5.1 相对值迭代**


---

```

1: 初始化: 对于每个状态  $\tilde{\mathbf{s}}$ , 初始化策略  $\pi^{(0)}(\tilde{\mathbf{s}}) = \arg \max_n \tilde{s}_n$ , 记策略初始值  $V_\alpha^{(0)}(\tilde{\mathbf{s}}) = \sum_{n=1}^N \tilde{s}_n$ .
2: repeat
3:    $\pi^{(k+1)}(\tilde{\mathbf{s}}) \leftarrow 0, \forall \tilde{\mathbf{s}}, \Delta = 0$ .
4:   for  $\tilde{\mathbf{s}} \in \mathcal{S}$  do
5:     值函数更新  $V_\alpha^{(k+1)}(\mathbf{x}) \leftarrow \min_{a \in \mathcal{A}} \left( C(\mathbf{x}, a) + \alpha \mathbb{E}_{\mathbf{x}'} [V_\alpha^{(k)}(\mathbf{x}') | \mathbf{x}, a] \right)$ 
6:   end for
7:    $V_\alpha^{(k+1)}(\tilde{\mathbf{s}}) \leftarrow V_\alpha^{(k+1)}(\tilde{\mathbf{s}}) - V_\alpha^{(k+1)}(0)$ , for all  $\tilde{\mathbf{s}}$ .
8:    $\Delta = \max_{\tilde{\mathbf{s}} \in \mathcal{S}} \{ \Delta, |V_\alpha^{(k+1)}(\tilde{\mathbf{s}}) - V_\alpha^{(k)}(\tilde{\mathbf{s}})| \}$ 
9:    $k \leftarrow k + 1$ 
10:  until  $\Delta < \delta$ .
    输出:
11:  策略  $\pi(\tilde{\mathbf{s}}) = \arg \min_{\tilde{\mathbf{s}} \in \mathcal{S}} (C(\tilde{\mathbf{s}}, a) + \alpha \mathbb{E}[V_\alpha(\tilde{\mathbf{s}}') | \tilde{\mathbf{s}}, a])$ 

```

---

## 5.4 调度策略 2：基于怀特指数的调度算法设计

为了降低计算复杂度, 本节中我们将调度问题转化为无休止多臂赌博机<sup>[101]</sup>, 通过怀特指数 (Whittle's index) 设计调度方案。

注意到网络中的  $N$  的状态更新信息源可以被看作  $N$  个赌博机, 每个赌博机  $n$  在时隙  $t$  一开始的状态为同步年龄  $s_n(t)$ , 用户  $n$  下一时隙  $t+1$  同步年龄仅依赖于调度策略  $u_n(t)$  和当前同步年龄  $s_n(t)$ 。因此, 每个用户的同步年龄类似于无休止赌博机演进。向用户推送状态信息可以被看作是激活赌博机, 而保持空闲不向用户推送最新的状态信息可以被看作是不激活赌博机。怀特指数可以解决无休止赌博机的优化问题<sup>[102]</sup>, 为了使用这一技术, 本章中, 我们将首先证明问题的可检索性 (indexability), 接下来计算怀特指数的闭式表达式并提出相应的调度策略。

### 5.4.1 解耦后的单用户调度策略设计

为了使用怀特指数, 我们首先松弛每时刻只能调度一个用户这一约束, 得到如下所示的时间平均带宽约束:

$$\mathbb{E}_\pi \left[ \frac{1}{T} \sum_{t=1}^T \sum_{n=1}^N u_n(t) \right] \leq 1.$$

记  $W \geq 0$  为拉格朗日乘子, 我们将松弛后的带宽约束放入目标函数中, 并得到如下代价最小化问题:

$$\text{minimize } \lim_{T \rightarrow \infty} \mathbb{E}_\pi \left[ \frac{1}{NT} \sum_{t=1}^T \sum_{n=1}^N \left( s_n(t) + W u_n(t) - \frac{W}{N} \right) \right]. \quad (5.15)$$

对于固定的  $W$ , 以上问题可以被解耦成  $N$  个子问题分别求解。在接下来的讨论中, 我们略去下标  $n$ 。接下来, 我们对每一个解耦后的子问题, 求取最小化同步年

龄与调度代价之和的策略  $\mu^*(W)$ , 即

$$\mu^*(W) \triangleq \arg_{\pi \in \Pi_{NA}} \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E}_{\mu} \left[ \sum_{t=1}^T (s(t) + W u(t)) \right]. \quad (5.16)$$

这一解耦后的策略优化问题同样可以被看成是一个马尔可夫决策过程, 这一马尔可夫决策过程的要素如下:

- **状态空间:** 时隙  $t$  的状态为用户当前的同步年龄  $s(t) \in \mathbb{N}$ , 同样地, 状态空间无限大但是可数。
- **决策空间:** 每时隙基站都有两种策略可以选取,  $a(t) = 1$  即调度相应用户,  $a(t) = 0$  表示不调度。注意到我们目前研究的是一个解耦后的调度问题,  $a(t)$  不同于  $u(t)$ , 本身没有带宽限制。
- **状态转移概率:** 根据式 (5.4), 解耦后的状态转移概率  $\Pr(s'|s, a)$  可以写作:

$$\Pr(s'|s, a) = \begin{cases} \lambda, & s' = 1, s = 0, a = 0, 1; \\ 1 - \lambda, & s' = 0, s = 0, a = 0, 1; \\ p\lambda, & s' = 1, s \neq 0, a = 1; \\ p(1 - \lambda), & s' = 0, s \neq 0, a = 1; \\ 1 - p, & s' = s + 1, s \neq 0, a = 1; \\ 1, & s' = s + 1, s \neq 0, a = 0; \\ 0, & \text{其他情况.} \end{cases} \quad (5.17)$$

- **单步代价函数:** 对于固定的拉格朗日乘子  $W$ , 赌博机处于状态  $s(t)$  时采用策略  $a(t)$  的一步代价为同步年龄和调度代价之和, 即:

$$C(s(t), a(t)) = s(t) + W a(t). \quad (5.18)$$

根据文献<sup>[103]</sup>, 存在一个静态随机策略可以最小化以上马尔可夫决策过程在无穷时间内的平均代价。接下来, 我们将证明最优策略  $\mu^*(W)$  具有阈值结构, 并依据最优阈值与  $W$  之间的关系证明同步年龄最小化问题的可索引 (indexable) 性质。

#### 5.4.2 可索引性质

根据文献<sup>[101]</sup>, 无休止多臂赌博机的可索引性定义如下:

**定义 5.1:** 记集合  $\Omega(W)$  为策略  $\mu^*(W)$  选择不调度的状态集合。我们称无休止多臂赌博机是可检索的, 如果集合  $\Omega(W)$  中的元素随着拉格朗日乘子的增大逐渐增加, 即  $\Omega(W) \subseteq \Omega(W'), \forall W \geq W'$ 。

接下来我们将首先证明以上马尔可夫决策过程的最优策略有阈值结构，证明过程将通过研究  $\alpha$  折扣代价最小化策略的阈值结构得到。对于给定的拉格朗日乘子  $W$ ，记  $J_{\alpha,W}(s, \mu)$  为从状态  $s(1) = s$  开始的  $\alpha$  累计折扣代价，即

$$J_{\alpha,W}(s, \mu) := \limsup_{T \rightarrow \infty} \mathbb{E}_{\mu} \left[ \sum_{t=1}^T \alpha^{t-1} C(s(t), a(t)) \right].$$

记  $\mu_{\alpha}^*(W) = \arg \min_{\mu \in \Pi_{NA}} J_{\alpha,W}(s, \mu)$  为最小化  $\alpha$  累计折扣代价的最优策略，记  $V_{\alpha,W}(s) = \min_{\mu} J_{\alpha}(s, \mu) = J_{\alpha,W}(s, \mu_{\alpha}^*(W))$  为使用这一策略时，从状态  $s$  出发得到的  $\alpha$  累计折扣代价。函数  $V_{\alpha(s),W}$  满足以下贝尔曼方程：

$$V_{\alpha,W}(s) = \min_{a \in \mathbb{A}} \{C(s, a) + \alpha \sum_{s'} V_{\alpha,W}(s') \Pr(s' | s, a)\}. \quad (5.19)$$

**引理 5.1：** 对于拉格朗日乘子  $W > 0$ ，值函数  $V_{\alpha,W}(\cdot)$  单调递增。

引理 5.1 与引理 3.3 的证明相似，在此我们省略具体证明过程。

**定理 5.2：** 最小化时间平均代价式 (5.16) 的策略  $\mu^*(W)$  具有阈值结构。也就是说，如果策略  $\mu^*(W)$  在同步年龄处于  $s$  时不发送对应状态信息，那么对于所有  $s' < s$  的状态，策略  $\mu^*(W)$  都将选择不发送；否则，如果最优策略在状态  $s$  时选择发送用户对应的状态信息，那么对于所有大于  $s$  的状态  $s+1, s+2, \dots$ ，最优策略  $\mu^*(W)$  都将选择向用户推送最新的状态信息。

**证明：** 我们将首先研究使得  $\alpha$  累计折扣代价最小化的策略  $\mu_{\alpha}^*(W)$  的阈值结构。策略  $\mu_{\alpha}^*$  选取的动作由贝尔曼方程式 (5.19) 决定，因此，如果处于状态  $s$  时，最优策略  $\mu_{\alpha}^*$  选择调度用户，值函数需要满足不等式：

$$W + \alpha p (\lambda V_{\alpha,W}(1) + (1 - \lambda) V_{\alpha,W}(0)) \leq \alpha p V_{\alpha,W}(s+1). \quad (5.20)$$

根据  $V_{\alpha,W}(\cdot)$  单调递增的性质，对于满足  $s' > s$  的状态  $s'$ ，我们均有：

$$W + \alpha p (\lambda V_{\alpha,W}(1) + (1 - \lambda) V_{\alpha,W}(0)) \alpha p V_{\alpha,W}(s) \leq \alpha p V_{\alpha,W}(s'). \quad (5.21)$$

以上不等式说明，对于满足不等式  $s' \geq s$  的状态  $s'$ ，策略  $\mu_{\alpha}^*(W)$  将在状态  $s'$  时选择调度这些用户。同样地，如果最优策略  $\mu_{\alpha}^*(W)$  在状态  $s$  时选择不调度用户，那么对于所有  $s' \leq s$  的状态，策略  $\mu_{\alpha}^*(W)$  都将选择不调度。注意到这一阈值结构对于  $\alpha \in (0, 1)$  均成立。取  $\alpha \rightarrow 1$ ，策略  $\mu_{\alpha}^*$  将趋于策略  $\mu^*(W)$ ，因此  $\mu^*(W)$  也有这样的阈值结构。 ■

接下来，我们通过证明激活阈值随着  $W$  增加而增大来证明本问题的可索引性。我们首先计算激活阈值为  $\tau$  时，马尔可夫决策过程的时间平均代价，详细的推导在附录 5.7.2 中；

**推论 5.1：** 考虑一个阈值策略，在同步年龄  $s \geq \tau$  进行状态信息推送，当  $s < \tau$  时

不推送状态信息。记  $F(\tau, W)$  为给定拉格朗日乘子  $W$ , 使用这一阈值  $\tau$  得到的时间平均代价。 $F(\tau, W)$  可以通过如下公式进行计算:

$$F(\tau, W) = \frac{\tau(\tau - 1)}{2} \xi_1^{(\tau)} + \frac{\xi_1^{(\tau)}}{p} \left( \frac{1}{p} - 1 \right) + \frac{\xi_1^{(\tau)}}{p} (\tau + W), \quad (5.22)$$

其中  $\xi_s^{(\tau)}$  表示使用激活阈值  $\tau$  时无休止赌博机处于状态  $s$  的稳态分布:

$$\xi_1^{(\tau)} = 1 / \left( \frac{1 - \lambda}{\lambda} + \tau + \frac{1}{p} - 1 \right).$$

接下来,我们将依据推论5.1推导拉格朗日乘子为  $W$  时,最优激活阈值  $\tau_{\text{opt}}(W)$ 。最优激活阈值需要满足  $F(\tau_{\text{opt}} + 1, W) \geq F(\tau_{\text{opt}}, W)$  以及  $F(\tau_{\text{opt}} - 1, W) \geq F(\tau_{\text{opt}}, W)$ , 求解这两个不等式得到的  $\tau_{\text{opt}}$  如推论5.2所述,详细的推导见附录5.7.3:

**推论 5.2:** 对于给定拉格朗日乘子  $W$ , 最优激活阈值可以通过如下式子计算:

$$\tau_{\text{opt}} = \lfloor \left( \frac{5}{2} - \frac{1}{p} - \frac{1}{\lambda} \right) + \sqrt{\left( \frac{5}{2} - \frac{1}{p} - \frac{1}{\lambda} \right)^2 + 2\left( \frac{W}{p} - \frac{1 - \lambda}{\lambda} \frac{1 - p}{p} \right) + 2\frac{1 - p}{p}} \rfloor. \quad (5.23)$$

我们注意到最优激活阈值  $\tau_{\text{opt}}$  随着拉格朗日乘子  $W$  单调非减,也就是说选择不推送状态信息这一动作的集合  $\Omega(W)$  随着  $W$  单调递增。特别地,当拉格朗日乘子  $W = 0$  时,最优阈值  $\tau_{\text{opt}} = 0$ ,也就是说  $\Omega_0 = \emptyset$  为空集。因此赌博机的可检索性得到了证明。

### 5.4.3 怀特指数 (Whittle's Index) 的推导

我们将赌博机处于状态  $s$  时的怀特指数记为  $I(s)$ , 它度量了在状态  $s$  下激活赌博机的收益。根据定义,  $I(s)$  是使得决策  $a = 1$  和  $a = 0$  在状态  $s$  下可以取得同样的收益的拉格朗日乘子  $W$  增量<sup>[102]</sup>。记  $\phi(\tau)$  为激活阈值为  $\tau$  时, 赌博机被激活的总时间占比, 根据阈值结构, 这一时间占比为赌博机处于状态  $s \geq \tau$  的稳态分布概率。注意到  $\xi_s^{(\tau)}$  表示赌博机使用阈值策略  $\tau$  时, 处于状态  $s$  的稳态分布,  $\phi(\tau)$  可以通过如下式子计算:

$$\phi(\tau) = \sum_{s=\tau}^{\infty} \xi_s^{(\tau)} = \sum_{s=\tau}^{\infty} \xi_1^{(\tau)} (1 - p)^{s-\tau} = \frac{\xi_1^{(\tau)}}{p}. \quad (5.24)$$

根据文献<sup>[101]</sup>式 6.11, 怀特指数为:

$$I(s) = \frac{F(s+1, 0) - F(s, 0)}{\phi(s) - \phi(s+1)} = \frac{p(F(s+1, 0) - F(s, 0))}{\xi_1^{(s)} - \xi_1^{(s+1)}}. \quad (5.25)$$

<sup>②</sup> 当  $\lambda = 1$  时, 同步年龄等价于信息年龄, 我们有  $\xi_1^{(s)} = \frac{1}{s + \frac{1}{p} - 1}$  以及  $F(s+1, 0) - F(s, 0) = \left( \frac{s(s-1)}{2} + \frac{1}{p^2} - \frac{1}{p} + \frac{s}{p} \right) (\xi_1^{(s+1)} - \xi_1^{(s)}) + (s + \frac{1}{p}) \xi_1^{(s+1)}$ 。在这样的情况下, 根据我们的推导, 怀特指数  $I(s) =$

#### 5.4.4 基于怀特指数的调度算法

最后我们给出一种基于怀特指数的调度算法。在每时隙  $t$  开始，基站计算每个用户的同步年龄  $s_n(t)$  和相应的怀特指数  $I_n(s_n(t))$ 。接下来，选择怀特指数最大的用户所对应的状态信息进行广播推送（如果出现怀特指数相同的情况，则从怀特指数最大的用户中随机选取一个）。我们注意到计算怀特指数的过程是针对每个用户单独进行的，因此计算复杂度为  $\mathcal{O}(N)$ ，相比于节5.3中基于马尔可夫决策过程的调度算法，计算复杂度几乎可以忽略不计。

我们注意到这一怀特指数的调度方案同样适用于带宽限制为  $M \geq 1$  的场景中。在这样的场景中，在时隙  $t$  基站计算出所有用户的怀特指数  $I_n(s_n(t))$  后，为了使得全网同步年龄最小，基站将选取怀特指数最大的  $M$  个用户并广播这些用户所感兴趣的状态信息。

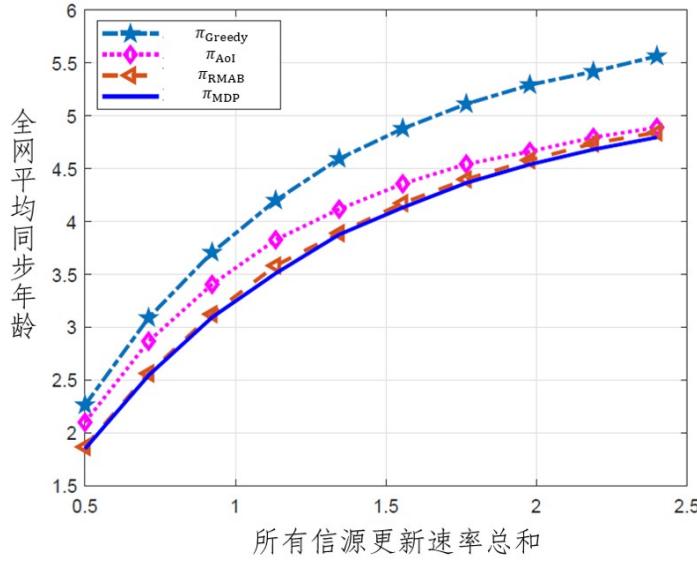
### 5.5 数值仿真

本部分我们将通过数值仿真说明所提算法在全网平均同步年龄的表现以及调度策略的设计启示。我们仿真并比较了以下四种调度算法所得到的平均同步年龄：1) 贪心法：每次都推送同步年龄最大的用户对应的信源  $\pi_{\text{Greedy}}$ ；2) 当信源更新由外部控制时，文献<sup>[37]</sup>提出的信息年龄最小的广播策略  $\pi_{\text{Aol}}$ ；3) 节5.3中基于马尔可夫决策过程设计的广播策略  $\pi_{\text{MDP}}$ 。4) 节5.4根据怀特指数的广播策略，记作策略  $\pi_{\text{RMAB}}$ 。在接下来的仿真中，我们记所有信源的更新速率总和为  $\lambda_{\text{tot}} := \sum_{n=1}^N \lambda_n$ 。全网平均同步年龄为  $T = 10^5$  个时隙所有用户同步年龄的均值。

图5.3中，我们考虑了三用户广播网络，对应信源的更新概率  $\lambda = [0.3, 0.4, 0.3]\lambda_{\text{tot}}$ ，传输成功概率  $p = [0.2, 0.55, 0.9]$ 。基于马尔可夫决策过程的调度决策使用的截断阈值  $m = 20$ 。图5.4画出网络中存在更多用户时的平均同步年龄，其中所有信源更新频率的总和  $\lambda_{\text{tot}} = 2$ ，每个信源  $n$  的更新频率  $\lambda_n = \frac{2n}{N(N+1)}\lambda_{\text{tot}}$ ，基站到用户  $n$  的传输成功概率  $p_n = \frac{n}{N}$ 。由于多用户网络中使用马尔可夫决策过程设计推送策略有较高的计算复杂度，我们用推导得到的全网平均同步年龄理论下界代替  $\pi_{\text{MDP}}$  得到的平均同步年龄。在图5.3中，基于无休止多臂赌博机的调度策略得到的全网平均同步年龄与基于马尔可夫决策过程得到的全网平均同步年龄相接近；图5.4中基于无休止多臂赌博机的调度策略得到的平均同步年龄与全网同步年龄的理论下界接近。尤其是当网络中用户数  $N$  逐渐增加， $\pi_{\text{Greedy}}$  和  $\pi_{\text{Aol}}$  得到的平均同步年龄与  $\pi_{\text{RMAB}}$  之间的差距不断变大，显示出所提策略在最小化全网平

---

$-p \left( \frac{s(s-1)}{2} + \frac{1}{p^2} - \frac{1}{p} + \frac{s}{p} \right) + p(s + \frac{1}{p})(s + \frac{1}{p} - 1) = \frac{ps}{2}(\tau + \frac{2-p}{p})$ 。这与文献<sup>[28]</sup>式<sup>(56)</sup>中  $T = 1$ ,  $\alpha = 1$  时得到的怀特指数相同。



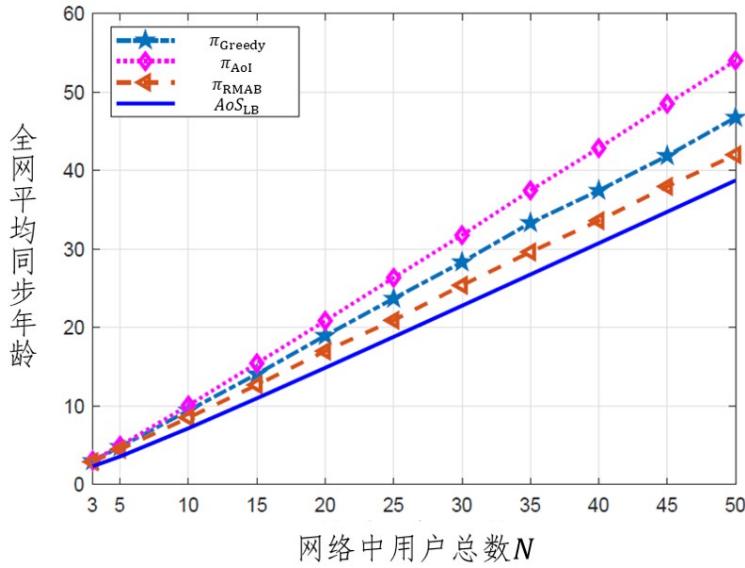
信源更新概率  $\lambda = [0.3, 0.4, 0.3]\lambda_{total}$ 、传输成功概率  $p = [0.2, 0.55, 0.9]$ 。

图 5.3 三用户网络的全网平均同步年龄与全网所有信源更新概率总和的关系

均同步年龄方面的良好性能。

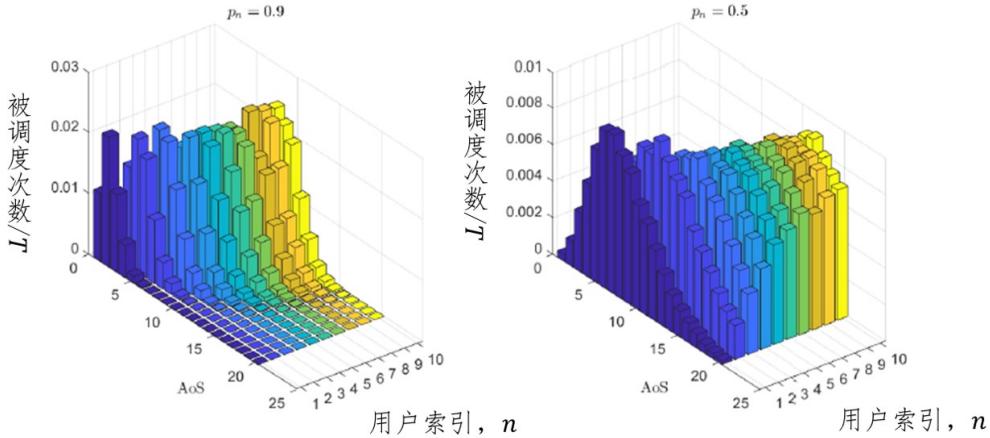
根据我们在节5.2.2的分析，当状态信息更新概率  $\lambda = 1$  时，信息年龄和同步年龄等价。图5.3中，当所有信源总的更新频率  $\lambda_{tot} \rightarrow 2.4$  时，每个信源的更新概率  $\lambda_n$  趋近于 1，也就是说几乎每个时隙，信源都有新的状态信息。在这样的场景中，信息年龄最小化策略  $\pi_{\text{AoI}}$  得到的平均同步年龄接近于  $\pi_{\text{RMAB}}$ 。当信源的外部更新频率  $\lambda_n$  远小于 1 或者网络中用户数量增加时，信息年龄最小化的调度策略  $\pi_{\text{AoI}}$  得到的平均同步年龄将显著高于策略  $\pi_{\text{RMAB}}$ ，甚至比贪心策略  $\pi_{\text{Greedy}}$  得到的平均同步年龄还差。这一现象进一步说明了，信息年龄和同步年龄是两类完全不同的度量，信息年龄最小化的调度策略不能保证一个好的同步年龄表现。

为了揭示同步年龄最小化调度策略的设计启示，我们计算了用户处于不同同步年龄时被推送相应状态信息的概率，并画出5.5所示的柱状图。图5.5考虑了  $N = 10$  的多用户网络，每个用户  $n$  对应信源  $n$  的随机更新概率  $\lambda_n = 0.5n/N$ ，左图中基站到用户传输成功的均为  $p_n = 0.9$ ，右图中基站到用户成功传输的概率均为  $p_n = 0.5$ 。图5.6中，每个用户成功传输概率  $p_n = n/N$ ，左图中信源更新频率  $\lambda_n = 0.2$ ，右图中信源更新频率  $\lambda_n = 0.5$ 。图5.5和5.6揭示了，那些更新概率  $\lambda_n$  较小信源、对应传输成功概率  $p_n$  较大的用户将优先获得状态信息推送。



全网所有信源的更新到达概率总和  $\lambda_{\text{tot}} = 2$ , 每个信源  $n$  的更新到达概率  $\lambda_n = \frac{2n}{N(N+1)} \lambda_{\text{total}}$ , 基站到用户  $n$  的传输成功概率  $p_n = n/N$ 。

图 5.4 全网平均同步年龄与网络用户数  $N$  的关系

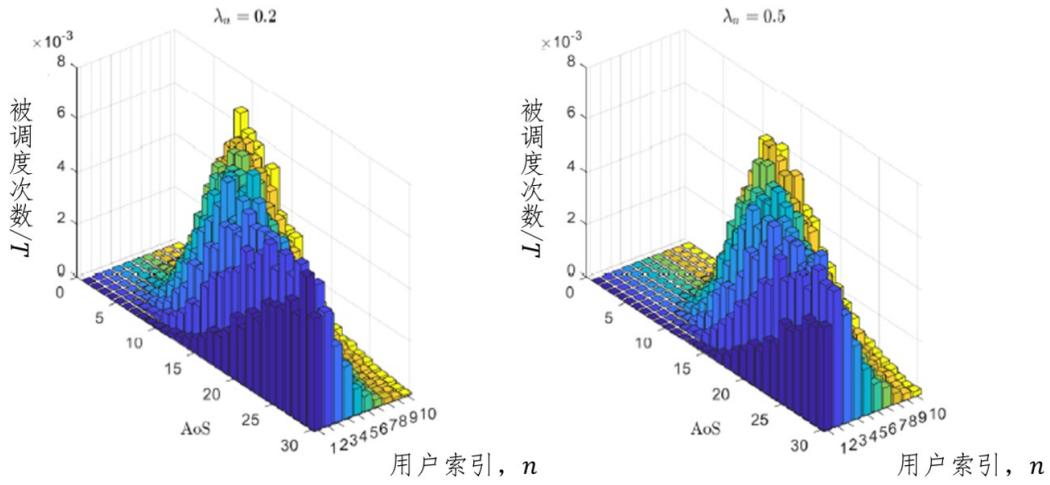


$N = 10$  多用户网络, 每个用户在不同同步年龄时收到状态信息推送的概率, 信源  $n$  产生更新的概率  $\lambda_n = 0.5n/N$ 。左图中基站给所有用户推送成功概率均为  $p_n = 0.9$ , 右图中基站所有用户推送成功概率均为  $p_n = 0.5$ 。

图 5.5 用户对应的信源更新概率不同时, 他们处于不同同步年龄被调度的概率

## 5.6 本章小结

本章中, 我们考虑一个基站将随机的状态更新信息通过无线信道广播和推送给用户。我们通过同步年龄度量接收端用户的时效性。基于马尔可夫决策过程和无休止多臂赌博机, 论文工作提出了两种最小化全网平均同步年龄的调度策略。仿真结果说明基于无休止多臂赌博机的调度策略得到的平均同步年龄与马尔可夫决



$N = 10$  多用户网络每个用户在不同同步年龄时的调度概率，基站推送内容给用户  $n$  的成功概率  $p_n = n/N$ 。左图中所有信源产生更新的概率  $\lambda_n = 0.2, \forall n$ ，右图中所有信源产生更新的概率  $\lambda_n = 0.5$ 。

图 5.6 基站到用户传输成功概率不同时，他们处于不同同步年龄被调度的概率

策过程得到的平均同步年龄相似。由于信息年龄和同步年龄有本质不同，使用信息年龄最小化的调度策略不能保证一个好的同步年龄。优化同步年龄的调度策略需要同时考虑状态信息更新频率和信道状态，那些更新频率较慢的信源、传输成功概率较大的用户将优先被广播和推送。

## 5.7 附录

### 5.7.1 定理5.1证明

**证明：** 考虑一种符合  $\pi \in \Pi$  带宽约束条件的因果调度策略  $\pi$ 。当用户  $n$  的同步年龄为 0 时，基站推送源  $n$  的状态无法降低用户  $n$  的同步年龄，因此只考虑那些同步年龄  $s_n(t) > 0$  时才会广播和推送源  $n$  的策略  $\pi$ 。考虑使用策略  $\pi$  得到的一条采样路径  $\omega$ ，截止时隙  $T$ ，假设基站推送信源  $n$  的更新  $L_n^T$  次，其中  $K_n^T$  次推送被用户  $n$  成功收到。与文献<sup>[28]</sup>的分析相同，当  $T \rightarrow \infty$  时，任意推送信源  $n$  次数小于或等于某个常数的策略都会导致用户  $n$  产生无限大的同步年龄。我们称这样的策略为“饥饿”策略，它们绝非最小化同步年龄的调度推送方案。在接下来的分析中我们仅考虑“非饥饿策略”  $\pi$ ，也就是说，当时间  $T$  趋于无穷时，广播和推送源  $n$  的次数以概率 1 趋于无穷：

$$\lim_{T \rightarrow \infty} L_n^T = \infty, \text{w.p.1.} \quad (5.26)$$

由于每次推送传输成功事件独立并服从参数为  $p_n$  的伯努利分布，我们有：

$$\lim_{T \rightarrow \infty} \frac{K_n^T}{L_n^T} = p_n, \text{w.p.1.} \quad (5.27)$$

以上等式说明用户  $n$  成功接收到信源  $n$  推送的次数以概率 1 趋于无穷，即：

$$\lim_{T \rightarrow \infty} K_n^T = \infty, \text{w.p.1.} \quad (5.28)$$

假设在时隙  $t_{n,i}$  时隙末尾，用户  $n$  第  $i$  次收到了信源  $n$  的更新。记  $\tau_{n,i}$  为用户第  $(i-1)$  次收到更新与第  $i$  次收到更新之间的时间间隔，即：

$$\tau_{n,i} := t_{n,i} - t_{n,i-1}, \forall i \in \{1, 2, \dots, K_n^T\}. \quad (5.29)$$

由于所有的用户一开始就被假设完全同步，我们记所有用户收到第 0 次更新的时刻  $t_{n,0} = 0, \forall n$ 。为了方便分析和计算同步年龄在时间  $[t_{n,K_n^T}, T]$  的变化，我们规定第  $K_n^T$  次更新间隔  $\tau_{n,K_n^T+1}$  如下：

$$\tau_{n,K_n^T+1} = T - t_{n,K_n^T}. \quad (5.30)$$

这样的定义使得  $\{\tau_{n,i}\}$  加和满足：

$$\sum_{i=1}^{K_n^T+1} \tau_{n,i} = T. \quad (5.31)$$

根据同步年龄演进，如果在用户  $n$  接收到信源  $n$  最新的状态信息后，信源  $n$  一直都没有新的状态信息产生，他的同步年龄  $s_n(t)$  将持续保持为 0。记  $v_{n,i}$  为用户  $n$  第  $i$  次接收到状态信息后，同步年龄持续为 0 的时隙数量，即对于  $j \in [1, v_{n,i}]$ ，同步年龄  $s_n(t_{n,i} + j) = 0$ 。在时隙  $t_{n,i} + v_{n,i} + 1$  开头，用户的同步年龄  $s_n(t_{n,i} + v_{n,i} + 1) = 1$  并将持续线性增长，直到用户在时隙  $t_{n,i+1}$  收到更新。为了简化后续分析，我们记  $w_{n,i} := s_n(t_{n,i+1}) = t_{n,i+1} - t_{n,i} - v_{n,i}$  为用户第  $i$  和第  $i+1$  次收到更新区间内，同步年龄不为 0 的总时隙数。那么，用户  $n$  在时间  $[t_{n,i} + 1, t_{n,i+1}]$  每时隙开头的同步年龄总和为：

$$\sum_{j=t_{n,i}+1}^{t_{n,i+1}} s_n(j) = \frac{s_n(t_{n,i+1})(s_n(t_{n,i+1}) + 1)}{2} = \frac{w_{n,i}^2}{2} + \frac{w_{n,i}}{2}. \quad (5.32)$$

记  $S_T(\omega)$  为采样路径在时间  $[1, T]$  内的平均同步年龄，它可以写成序列  $\{K_n^T\}$  与  $\{w_{n,i}\}$  的函数：

$$S_T(\omega) := \frac{1}{NT} \sum_{n=1}^N \sum_{i=1}^{K_n^T+1} \left( \sum_{j=t_{n,i-1}+1}^{t_{n,i}} s_n(j) \right)$$

$$= \frac{1}{N} \sum_{n=1}^N \left( \frac{K_n^T + 1}{T} \frac{\sum_{i=1}^{K_n^T+1} \frac{w_{n,i}^2}{2} + \frac{w_{n,i}}{2}}{K_n^T + 1} \right). \quad (5.33)$$

记  $\mathbb{M}[\cdot]$  为一系列随机变量的平均值, 记  $\gamma_n$  为推送信源  $n$  的状态信息的时间占比, 即  $\gamma_n := \frac{K_n^T+1}{T}$ 。我们可以得到同步年龄  $S_T(\omega)$  的一个下界:

$$\begin{aligned} S_T(\omega) &= \frac{1}{N} \sum_{n=1}^N \gamma_n \left( \frac{1}{2} \mathbb{M}[w_n^2] + \frac{1}{2} \mathbb{M}[w_n] \right) \\ &\stackrel{(a)}{\geq} \frac{1}{N} \sum_{n=1}^N \gamma_n \left( \frac{1}{2} \mathbb{M}[w_n]^2 + \frac{1}{2} \mathbb{M}[w_n] \right), \end{aligned} \quad (5.34)$$

其中不等式 (a) 由不等式  $\mathbb{M}[w_n^2] \geq \mathbb{M}[w_n]^2$  得到。为了进一步计算得到  $S_T(\omega)$  的下界, 接下来我们首先推导  $\mathbb{M}[w_n]$  与  $\gamma_n$  之间关系, 进而通过  $\{\gamma_n\}$  的内在约束, 把求解同步年龄  $\lim_{T \rightarrow \infty} \mathbb{E}[S_T(\omega)]$  的理论下界这一问题转化为序列  $\{\gamma_n\}$  的优化问题。

给定  $\gamma_n$ , 序列  $\{w_{n,i}\}_{i=1}^{K_n^T}$  的平均值  $\mathbb{M}[w_n]$  计算将分为两步: 首先我们推导  $\mathbb{M}[w_n] + \mathbb{M}[v_n]$  与  $\gamma_n$  的关系, 接下来我们计算  $\mathbb{M}[v_n]$ 。序列  $\{w_{n,i}\}_{i=1}^{K_n^T}$  与序列  $\{v_{n,i}\}_{i=1}^{K_n^T}$  平均值之和  $\mathbb{M}[w_n] + \mathbb{M}[v_n]$  计算如下:

$$\mathbb{M}[w_n] + \mathbb{M}[v_n] \stackrel{(a)}{=} \frac{\sum_{i=1}^{K_n^T+1} w_{n,i}}{K_n^T + 1} + \frac{\sum_{i=1}^{K_n^T+1} v_{n,i}}{K_n^T + 1} \stackrel{(b)}{=} \frac{T}{K_n^T + 1} = \frac{1}{\gamma_n}, \quad (5.35)$$

其中等式 (a) 来源于  $\mathbb{M}[\cdot]$  的定义, 等式 (b) 来源于定义  $\tau_{n,i+1} = w_{n,i} + v_{n,i}$  和式 (5.31)。

由于每时隙状态信息  $n$  是否有更新服从参数为  $\lambda_n$  的伯努利分布, 因此用户  $n$  在第  $i$  次收到更新后同步年龄保持为 0 的时长  $v_{n,i}$  服从参数为  $\lambda_n$  的几何分布。根据大数定律, 当时间  $T$  趋于无穷时, 我们有:

$$\mathbb{M}[v_n] = \mathbb{E}[v_{n,i}] = \frac{1 - \lambda_n}{\lambda_n}, \text{ w.p.1.} \quad (5.36)$$

将式 (5.36) 带入到式 (5.35) 中, 我们可以将  $\mathbb{M}[w_n]$  写成关于  $\gamma_n$  的一个函数:

$$\mathbb{M}[w_n] = \frac{1}{\gamma_n} - \frac{1}{\lambda_n} + 1, \text{ w.p.1.} \quad (5.37)$$

接下来我们讨论  $\gamma_n$  取值的限制。第一个约束条件由  $\mathbb{M}[w_n]$  的下界得到。由于策略  $\pi$  仅在用户的同步年龄不为 0 时才会推送源  $n$  的更新, 因此在每次推送更新的时隙  $t_{n,i}$ , 用户  $n$  的同步年龄均  $s_n(t_{n,i}) \geq 1$ , 也就是说  $\mathbb{M}[w_n] \geq 1$ 。利用式 (5.37) 中  $\mathbb{M}[w_n]$  与  $\gamma_n$  之间的关系, 我们得到如下有关  $\gamma_n$  的约束条件:

$$\frac{1}{\gamma_n} - \frac{1}{\lambda_n} + 1 \geq 1 \Rightarrow \gamma_n \leq \lambda_n, \text{ w.p.1.} \quad (5.38)$$

序列  $\{\gamma_n\}$  的第二个约束条件来源于网络带宽限制。由于每时刻基站只能广播

和推送一个信源的更新，在区间  $[1, T]$ ，基站推送给状态信息最多  $T$  次，因此：

$$\sum_{n=1}^N L_n^T \leq T. \quad (5.39)$$

当时间  $T \rightarrow \infty$ ，我们有  $\gamma_n = \lim_{T \rightarrow \infty} \frac{K_n^T + 1}{T} = \lim_{T \rightarrow \infty} \frac{K_n^T}{T}$ 。将式 (5.27) 带入式 (5.39)，我们可以得到序列  $\{\gamma_n\}$  需满足的第二个约束条件：

$$\sum_{n=1}^N \frac{\gamma_n}{p_n} \leq 1. \quad (5.40)$$

因此，对于任意非饥饿策略  $\pi \in \Pi$ ，它的平均同步年龄下界都可以通过计算该策略所对应的系数  $\{\gamma_n\}$ ，之后利用式 (5.37) 计算  $\mathbb{M}[w_n]$  并带入式 (5.34) 得到。因此，记  $\text{AoS}_{LB}$  为全网用户平均同步年龄的理论下界，它可以通过求解以下关于  $\gamma = \{\gamma_n\}_{n=1}^N$  的优化问题得到：

问题 5.2 (同步年龄理论下界问题)：

$$\text{AoS}_{LB} := \min_{\gamma \geq 0} \frac{1}{N} \sum_{n=1}^N \gamma_n \left[ \frac{1}{2} \left( \frac{1}{\gamma_n} - \frac{1}{\lambda_n} + 1 \right)^2 + \frac{1}{2} \left( \frac{1}{\gamma_n} - \frac{1}{\lambda_n} + 1 \right) \right], \quad (5.41a)$$

$$\text{s.t., } \sum_{n=1}^N \frac{\gamma_n}{p_n} \leq 1, \quad (5.41b)$$

$$\gamma_n \leq \lambda_n, \forall n. \quad (5.41c)$$

采用任意策略  $\pi$  所得到的全网平均同步年龄将以概率 1 大于  $\text{AoS}_{LB}$ 。接下来我们讨论优化问题 5.2 的求解。目标函数为凸，优化变量  $\gamma$  在闭的凸集中选取，我们可以使用拉格朗日函数求解优化问题最优解：

$$\begin{aligned} \mathcal{L}(\gamma, \mu, \nu) &= \frac{1}{N} \sum_{n=1}^N \gamma_n \left( \frac{1}{2} \left( \frac{1}{\gamma_n} - \frac{1-\lambda_n}{\lambda_n} \right)^2 + \frac{1}{2} \left( \frac{1}{\gamma_n} - \frac{1-\lambda_n}{\lambda_n} \right) \right) \\ &\quad + \mu \left( \sum_{n=1}^N \frac{\gamma_n}{p_n} - 1 \right) + \sum_{n=1}^N \nu_n (\gamma_n - \lambda_n), \end{aligned} \quad (5.42)$$

其中  $\mu$  和  $\nu := [\nu_1, \dots] \geq 0$  为拉格朗日乘子。根据 Karush-Kuhn-Tucker (KKT) 条件，参数  $\gamma$  在最优点的取值需满足： $\nabla_\gamma \mathcal{L}(\gamma, \mu, \nu) = 0, \forall n$ ，也就是说：

$$\frac{1}{2N} \left[ \left( \frac{1-\lambda_n}{\lambda_n} \right)^2 - \left( \frac{1-\lambda_n}{\lambda_n} \right) \right] - \frac{1}{2N} \frac{1}{\gamma_n^2} + \frac{\mu}{p_n} + \nu_n = 0. \quad (5.43)$$

因此，最优点处  $\gamma_n^*$  可以被写成对偶乘子  $\mu^*$  和  $v_n^*$  的函数：

$$\gamma_n^* = 1/\sqrt{\left(\frac{1-\lambda_n}{\lambda_n}\right)^2 - \frac{1-\lambda_n}{\lambda_n} + N\left(\frac{2\mu^*}{p_n} + 2v_n^*\right)}. \quad (5.44)$$

接下来我们考虑 Complete Slackness (CS) 条件：

$$\mu^* \left( \sum_{n=1}^N \frac{\gamma_n^*}{p_n} - 1 \right) = 0, \quad (5.45a)$$

$$v_n^*(\gamma_n^* - \lambda_n) = 0, \forall n. \quad (5.45b)$$

结合约束条件  $\gamma_n \leq \lambda_n$  暗示了，当  $\gamma_n^* < \lambda_n^*$  时，对偶乘子  $v_n^* = 0$ 。因此最优  $\gamma_n^*$  可以通过如下公式计算：

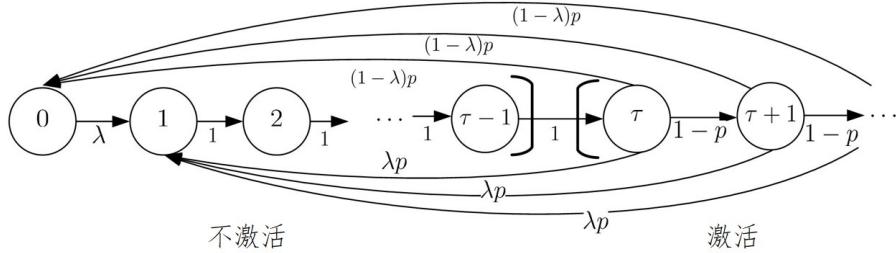
$$\gamma_n^* = \max\{1/\sqrt{\left(\frac{1-\lambda_n}{\lambda_n}\right)^2 - \left(\frac{1-\lambda_n}{\lambda_n}\right) + \frac{2\mu^* N}{p_n}}, \lambda_n\}, \quad (5.46)$$

其中对偶乘子  $\mu^*$  的选取使得  $\sum_{n=1}^N \frac{\gamma_n^*}{p_n} = 1$ 。

■

### 5.7.2 推论5.1的推导过程

根据用户  $n$  的同步年龄随调度策略状态转移概率，我们可以画出使用激活阈值  $\tau$  时的状态转移图：



考虑激活阈值为  $\tau$ ，对于同步年龄小于  $\tau$  的状态，无休止赌博机保持空闲（即不调度）；对于同步年龄大于和等于  $\tau$  的状态，无休止赌博机将被激活和调度。状态之间的转移概率标注在箭头下方。

图 5.7 用户的同步年龄状态转移图。

记  $\xi_s^{(\tau)}$  为使用激活阈值  $\tau$  时，无休止赌博机的同步年龄为  $s$  的稳态分布概率。根据式 (5.17) 和图5.7所示的状态转移关系，稳态分布概率  $\{\xi_s^{(\tau)}\}$  需要满足

1) 对于状态  $s < \tau$ ，无休止赌博机保持空闲不被激活，因此：

$$\xi_1^{(\tau)} = \dots = \xi_\tau^{(\tau)}. \quad (5.47a)$$

2) 对于状态  $s \geq \tau$ ，赌博机将被激活，用户将获得相应信息的推送。以概率

$1 - p$  广播推送失败，用户的同步年龄线性增长  $s + 1$ ，因此：

$$\xi_{s+1}^{(\tau)} = (1 - p)\xi_s^{(\tau)}, \forall s \geq \tau. \quad (5.47b)$$

以概率  $p$  传输和推送将成功被用户  $n$  接收到。以概率  $\lambda$ ，在时隙  $t$  一开始信源  $n$  会产生一次新的更新，无休止赌博机将回到状态  $s = 1$ ：

$$\xi_1^{(\tau)} = \sum_{s=\tau}^{\infty} \lambda p \xi_s^{(\tau)} + \lambda \xi_0^{(\tau)}. \quad (5.47c)$$

以概率  $1 - \lambda$ ，状态信息  $n$  将不产生更新，那么用户  $n$  在下一时隙开始的同步年龄将回到  $s = 0$ ：

$$\xi_0^{(\tau)} = \sum_{s=\tau}^{\infty} (1 - \lambda) p \xi_s^{(\tau)} + (1 - \lambda) \xi_0^{(\tau)}. \quad (5.47d)$$

所有状态的稳态分布概率加和为 1，即：

$$\sum_{s=0}^{\infty} \xi_s^{(\tau)} = 1. \quad (5.47e)$$

求解式 (5.47b) - (5.47e) 对应方程可以得到无休止赌博机的稳态分布概率  $\xi_s^{(\tau)}$ ：

$$\xi_s^{(\tau)} = \begin{cases} \frac{1-\lambda}{\lambda} / \left( \frac{1-\lambda}{\lambda} + \tau + \frac{1}{p} - 1 \right), & s = 0; \\ 1 / \left( \frac{1-\lambda}{\lambda} + \tau + \frac{1}{p} - 1 \right), & 1 \leq s \leq \tau; \\ \frac{(1-p)^{s-\tau}}{\left( \frac{1-\lambda}{\lambda} + \tau + \frac{1}{p} - 1 \right)}, & s > \tau. \end{cases} \quad (5.48)$$

因此，使用激活阈值  $\tau$ ，对于给定的拉格朗日乘子  $W$ ，无休止赌博机的总代价  $F(\tau, W)$  为：

$$\begin{aligned} F(\tau, W) &= \sum_{s=0}^{\tau-1} s \xi_s^{(\tau)} + \sum_{s=\tau}^{\infty} (s + W) \xi_s^{(\tau)} \\ &= \sum_{s=1}^{\tau-1} s \xi_1^{(\tau)} + \sum_{s=\tau}^{\infty} (s + W) \xi_1^{(\tau)} (1 - p)^{s-\tau} \\ &= \frac{\tau(\tau-1)}{2} \xi_1^{(\tau)} + \frac{\xi_1^{(\tau)}}{p} \left( \frac{1}{p} - 1 \right) + \frac{\xi_1^{(\tau)}}{p} (\tau + W). \end{aligned} \quad (5.49)$$

### 5.7.3 推论5.2的推导过程

记  $V(s)$  为状态  $s$  下的值函数，取  $\beta$  为无休止赌博机使用最优激活策略得到的平均代价。贝尔曼方程如下：

$$V(s) + \beta = \min\{(W + s) + (1 - p)V(s + 1) + p(\lambda V(1) + (1 - \lambda)V(0)), s + V(s + 1)\}. \quad (5.50)$$

根据定理5.2，解耦后无休止赌博机的最优决策具有阈值结构。假设  $\tau_{\text{opt}}$  记为最优激活阈值，即当同步年龄满足  $0 \leq s < \tau_{\text{opt}}$  时，最优策略不调度用户；而当同步年龄满足  $s \geq \tau_{\text{opt}}$ ，无休止赌博机被激活，用户被调度。以上贝尔曼方程说明，当无休止赌博机处于状态  $\tau_{\text{opt}}$  时，值函数满足：

$$W + p((1 - \lambda)V(0) + \lambda V(1)) \leq pV(\tau_{\text{opt}} + 1), \quad (5.51a)$$

当赌博机处于状态  $\tau_{\text{opt}} - 1$  时，最优策略选择不调度用户，因此：

$$W + p((1 - \lambda)V(0) + \lambda V(1)) \geq pV(\tau_{\text{opt}}). \quad (5.51b)$$

计算最优阈值  $\tau_{\text{opt}}$  将分为两步：首先我们将建立值函数  $V(\tau_{\text{opt}})$ 、平均代价  $\beta$  与最优激活阈值  $\tau_{\text{opt}}$  之间的关系。接下来，我们将通过两种不同的方式计算  $V(\tau_{\text{opt}})$ ，通过求解方程计算最优阈值  $\tau_{\text{opt}}$ 。不失一般地，我们假设  $\lambda V(1) + (1 - \lambda)V(0) = 0$ 。

为了计算最优阈值  $\tau_{\text{opt}}$  处的值函数  $V(\tau_{\text{opt}})$ ，我们首先考虑满足  $s \geq \tau_{\text{opt}}$  的状态。根据  $\tau_{\text{opt}}$  的最优性，由于调度用户是最优的，根据贝尔曼方程式 (5.50)，我们可以得到如下  $V(s)$  与  $V(s + 1)$  之间的关系：

$$V(s) = (-\beta + W + s) + (1 - p)V(s + 1). \quad (5.52)$$

将  $V(s + 1) = (-\beta + W + s) + (1 - p)V(s + 2)$  带入到以上等式，我们可以得到：

$$V(s) = (-\beta + W + s) + (1 - p)(-\beta + W + s + 1) + (1 - p)^2V(s + 2). \quad (5.53)$$

将以上过程重复  $K$  次我们得到：

$$V(s) = \sum_{k=0}^{K-1} (1 - p)^k(-\beta + W + s + k) + (1 - p)^K V(s + K). \quad (5.54)$$

考虑重复次数  $K \rightarrow \infty$ ，由于  $\lim_{K \rightarrow \infty} (1 - p)^K V(s + K) = 0$ ，状态  $s \geq \tau_{\text{opt}}$  处的值函数  $V(s)$  可以通过以下式子计算：

$$V(s) = \frac{1}{p}(-\beta + W + s) + \frac{1 - p}{p^2}. \quad (5.55)$$

结合式 (5.51a)、式 (5.51b) 和假设  $(1 - \lambda)V(0) + \lambda V(1) = 0$ ，考虑到  $\tau_{\text{opt}}$  为最优

激活阈值，我们有以下不等式：

$$V(\tau_{\text{opt}}) \leq \frac{W}{p} < V(\tau_{\text{opt}} + 1). \quad (5.56)$$

注意到值函数  $V(s)$  单调递增，以上不等式说明存在  $\gamma \in [0, 1)$  使得  $V(\tau_{\text{opt}} + \gamma) = \frac{W}{p}$ 。将式 (5.55) 带入到  $V(\tau_{\text{opt}} + \gamma)$ ，我们可以通过如下表达式计算使用最优策略的平均代价  $\beta$ ：

$$\beta = \tau_{\text{opt}} + \gamma + \frac{1-p}{p}. \quad (5.57)$$

接下来我们考虑满足  $s < \tau_{\text{opt}}$  的状态，并将  $(1 - \lambda)V(0) + \lambda V(1)$  写成关于  $\tau_{\text{opt}}$  的函数。由于最优策略在状态  $s < \tau_{\text{opt}}$  保持空闲，根据贝尔曼方程，我们有：

$$V(s) = (-\beta + W + s) + V(s + 1). \quad (5.58)$$

将  $V(s - 1) = (-\beta + s - 1) + V(s)$  带入到以上不等式，并将这一过程重复  $s - 1$  次，我们可以得到  $V(1)$  与  $V(\tau_{\text{opt}})$  之间的关系：

$$V(1) = \frac{(\tau_{\text{opt}} - 1)(-2\beta + \tau_{\text{opt}})}{2} + V(\tau_{\text{opt}}). \quad (5.59)$$

考虑同步年龄处于 0 时的贝尔曼方程，值函数  $V(0)$  满足以下条件：

$$V(0) = -\beta + (1 - \lambda)V(0) + \lambda V(1),$$

也就是说：

$$V(0) = -\frac{\beta}{\lambda} + V(1). \quad (5.60)$$

将以上等式带入到假设  $(1 - \lambda)V(0) + \lambda V(1) = 0$ ，我们可以得到同步年龄为 1 时的值函数  $V(1)$ ：

$$V(1) = \frac{1 - \lambda}{\lambda} \beta. \quad (5.61)$$

将式 (5.60) 和式 (5.61) 带入到式 (5.59) 中，我们可以得到  $V(\tau_{\text{opt}})$  的表达式：

$$V(\tau_{\text{opt}}) = -\frac{(\tau_{\text{opt}} - 1)(-2\beta + \tau_{\text{opt}})}{2} + \frac{1 - \lambda}{\lambda} \beta. \quad (5.62)$$

接下来我们通过另一种方式计算状态  $\tau_{\text{opt}}$  的值函数。通过式 (5.55)，当赌博机状态处于  $s = \tau_{\text{opt}}$  时，我们有：

$$V(\tau_{\text{opt}}) = \frac{1}{p}(-\beta + W + \tau_{\text{opt}}) + \frac{1-p}{p^2}. \quad (5.63)$$

由于通过以上两个不等式计算出的  $V(\tau_{\text{opt}})$  数值应相同, 即我们得到以下方程:

$$-\frac{1-\lambda}{\lambda}\beta + \frac{(\tau_{\text{opt}}-1)(-2\beta+\tau_{\text{opt}})}{2} + \frac{1}{p}(-\beta+W+\tau_{\text{opt}}) + \frac{1-p}{p^2} = 0. \quad (5.64)$$

将  $\beta = \tau_{\text{opt}} + \gamma + \frac{1-p}{p}$  带入到以上等式, 我们可以得到:

$$-\frac{1-\lambda}{\lambda}(\tau_{\text{opt}} + \gamma + \frac{1-p}{p}) + \frac{(\tau_{\text{opt}}-1)(-\tau_{\text{opt}} - 2\gamma - 2\frac{1-p}{p})}{2} + \frac{1}{p}(W - \gamma) = 0. \quad (5.65)$$

从以上等式我们可以得到最优阈值  $\tau_{\text{opt}}$  的二次方程:

$$\frac{1}{2}\tau_{\text{opt}}^2 + (\gamma + \frac{1}{p} + \frac{1}{\lambda} - \frac{5}{2})\tau_{\text{opt}} - \frac{1}{p}(W - \gamma) + \frac{1-\lambda}{\lambda}(\gamma + \frac{1-p}{p}) - 2\frac{1-p}{p} - 2\gamma = 0.$$

由于激活阈值  $\tau_{\text{opt}}$  为整数,  $\gamma \in [0, 1)$ , 对于给定拉格朗日乘子  $W$ , 我们最终可以得到最优激活阈值为:

$$\tau_{\text{opt}} = \lfloor (\frac{5}{2} - \frac{1}{p} - \frac{1}{\lambda}) + \sqrt{\left(\frac{5}{2} - \frac{1}{p} - \frac{1}{\lambda}\right)^2 + 2\left(\frac{W}{p} - \frac{1-\lambda}{\lambda}\frac{1-p}{p}\right) + 2\frac{1-p}{p}} \rfloor. \quad (5.66)$$

## 第6章 总结与展望

### 6.1 研究内容总结

为了实现自动驾驶的愿景，海量传感器接入车联网，监测车速、位置等状态信息，并通过无线网络将状态信息发送给车联网中的用户和控制单元。用户和控制单元的状态信息时效性将对自动驾驶和车联网系统的安全性有重大影响。本文面向车联网中状态信息采集、用户通过基站主动索取和被动推送状态信息两种场景，针对通信有时延、信道变化随机、用户需求随机、通信资源受限的挑战，基于在线学习、马尔可夫决策过程等方法，使用信息年龄和同步年龄等指标度量接收端状态信息时效性，设计了时效性最优采样、传输和调度策略。

本文的主要贡献和创新点总结如下：

- 1 针对点对点通信链路中传输时延统计信息未知的场景，以最小化接收端平均信息年龄为目标，提出了一种自适应的采样传输算法。论文工作将信息年龄最小化问题转化为更新过程优化，并基于李雅普诺夫理论，设计了更新间隔违背的虚拟队列。通过随机优化算法和李雅普诺夫漂流加惩罚框架，设计了满足传输资源限制下，最小化接收端平均信息年龄的自适应采样和传输算法。论文工作证明了所提算法得到的平均信息年龄与统计信息已知时最小信息年龄之间的差值随着采样次数  $K$  的衰减速率为  $\mathcal{O}\left(\frac{\log K}{K}\right)$ 。仿真结果表明，对比贪心法等采样传输策略，所提算法得到的平均信息年龄较小；相比于传输时延统计信息已知时的最优采样策略，所提算法有很好地自适应性能，随着运行时间的增加，所提算法得到的平均信息年龄趋于最优。
- 2 针对基站通过时变无线信道从多传感器采样状态信息的场景，提出了一种最小化全网信息年龄的采样、跨层优化、调度联合设计。针对多用户采样、调度决策问题中“维度爆炸”的挑战，论文工作通过约束条件松弛、拉格朗日乘子法将多传感器决策问题解耦单传感器代价最小化问题。利用单传感器代价最小化问题最优解的阈值结构，使用线性规划求解了单传感器代价最小化问题中最优采样和发送策略。进一步地，论文工作提出了一种满足网络传输带宽约束的多传感器采样、跨层优化和调度算法，证明了其在大规模网络中的理论最优化。所提策略针对信道状态时变的特点，通过机会调度传感器采样和发送状态信息，保证了基站端的信息时效性。典型参数下，所提算法可以降低全网平均信息年龄约 30%。
- 3 针对车联网中用户从基站端缓存直接索取状态信息副本的场景，考虑状态信

息流行度和传输时长随时间变化的特点，以最小化用户索取得到的状态信息副本的信息年龄为目标，提出了相应的基站端缓存更新策略。为了解决有限带宽下状态信息副本更新这一整数规划问题，我们通过带宽松弛和拉格朗日乘子法将多状态信息更新问题解耦为单状态信息代价最小化问题。利用状态信息更新时长随信息年龄单调递增、递增速度逐渐减小的特点，将调度策略设计转化为带宽分配问题，利用这一问题单调优化的特性，使用分支定界法求解最优带宽分配策略，基于此提出了满足网络带宽约束的基站端缓存更新算法。所提算法根据用户需求，结合不同状态信息下载时长的变化特性，将有限更新带宽优先分配给流行度较高、更新所需时长较短的状态信息副本。

4 针对状态信息更新随机的特点，选用同步年龄作为时效性度量，提出了带宽限制下同步年龄最小的内容推送策略。论文工作通过采样路径定理推导了全网同步年龄的理论下界，将带宽约束下同步年龄最小化这一调度问题转化为无休止多臂赌博机代价最小化问题，基于马尔可夫决策过程和怀特指数设计了优化全网同步年龄的调度方案。所提方案将优先向信道条件较好的用户推送那些更新频率较低的状态信息。

各章节核心成果及其相互关系如图6.1所示。本文针对车联网中状态信息优化问题，以最小化接收端信息年龄、同步年龄为目标，在状态信息采集和推送环节设计了优化信息时效性的采样和传输调度策略。研究内容1考虑点对点链路信息年龄最小化问题，得到了信息年龄最小化策略的设计启示，即信息年龄最小化的采样和传输策略具有阈值结构，一旦接收端的信息年龄超过某一阈值应立刻采样和发送新的状态信息，这一点对点链路中的设计启示是时效性最优的采样、传输策略设计核心。在此基础上，研究内容2考虑多传感器网络中传输信道时变的特点，提出了大规模网络中信息年龄最优的采样和调度算法。前两项研究内容侧重于状态信息采集。后两项研究内容在此基础上考虑了基站将状态信息推送和分发给用户。其中，研究内容3考虑用户从基站端缓存主动索取状态信息副本的场景，对基站端缓存中状态信息副本的更新给出了理论指导；研究内容4进一步考虑了状态信息变化特点和单播推送过程中的随机性，选取同步年龄作为度量指标，提出了带宽限制下优化全网同步年龄的广播推送方案。

## 6.2 进一步研究方向

提供时效性强的状态信息是未来车联网、工业物联网中安全有效控制的关键。在本文研究调度算法的基础上，仍有一些开放问题值得进一步研究：

- 块衰落（block fading）信道下时效性最优的自适应调度算法。本文在研究内

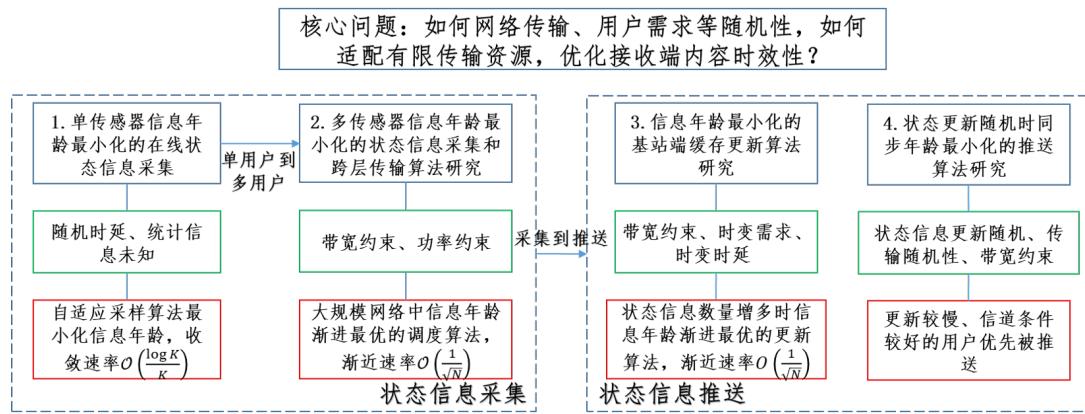


图 6.1 研究内容总结

容 1 中考虑了信道状态不变场景中时效性最优的自适应采样传输设计，在研究内容 2 中考虑了信道状态快衰落场景下信息年龄最小的采样、传输联合优化。但在实际块衰落场景中，如何设计自适应的调度算法，既能尽快检测信道状态和传输时延的改变，又能快速适应新的信道状态从而优化接收端状态信息时效性有待进一步研究。

- 利用状态信息先验和状态信息内在相关性的编码、传输联合设计。状态信息的一个重要特点是不同时刻的状态信息彼此相关，而过去对于信源压缩编码的研究主要是针对信源信息彼此独立进行的。当我们对于状态信息的演进有一定先验知识（例如它是一个维纳过程，或是一个线性时不变系统），充分利用先验知识将有助于设计更加高效的编码和传输方案。未来我们将探索利用先验知识和节点间协作的编码、传输策略的联合设计，将有限的传输资源充分利用。
- 基于已有状态信息有效、安全的控制算法。时效性强的状态信息最终将用于决策和控制。由于通信系统的资源限制，即使我们进行了编码压缩、传输和调度优化，接收端所拥有的状态信息仍然不是最新的。在这样的情况下，如何依据这些部分过时的状态信息，设计有效安全的决策算法，是保证车联网、工业物联网有效控制的关键。

## 参考文献

- [1] Autonomous vehicles will surpass 95 million in annual sales by 2035.[J]. *navigantresearch.com*, 2013.
- [2] 工业和信息化部. 汽车驾驶自动化分级[J]. 2019.
- [3] 中国信息通信研究院. 车联网白皮书[Z]. 2020.
- [4] Eryilmaz A, Srikant R. Asymptotically tight steady-state queue length bounds implied by drift conditions[J]. *Queueing Syst. Theory Appl.*, 2012, 72(3-4): 311-359.
- [5] Kaul S, Yates R, Gruteser M. Real-time status: How often should one update?[C]//2012 Proceedings IEEE INFOCOM. 2012: 2731-2735.
- [6] Zhong J, Yates R D, Soljanin E. Two Freshness Metrics for Local Cache Refresh[C]//2018 IEEE International Symposium on Information Theory (ISIT). 2018: 1924-1928.
- [7] Soysal A, Ulukus S. Age of information in g/g/1/1 systems[C]//2019 53rd Asilomar Conference on Signals, Systems, and Computers. 2019: 2022-2027.
- [8] Najm E, Nasser R, Telatar E. Content based status updates[J]. *IEEE Transactions on Information Theory*, 2020, 66(6): 3846-3863.
- [9] Wang Y, Chen W. Adaptive power and rate control for real-time status updating over fading channels[J]. *IEEE Transactions on Wireless Communications*, 2021, 20(5): 3095-3106.
- [10] Wang B, Feng S, Yang J. When to preempt? age of information minimization under link capacity constraint[J]. *Journal of Communications and Networks*, 2019, 21(3): 220-232.
- [11] Zhou B, Saad W. Joint status sampling and updating for minimizing age of information in the internet of things[J]. *IEEE Transactions on Communications*, 2019, 67(11): 7468-7482.
- [12] Ceran E T, Gündüz D, György A. Average age of information with hybrid arq under a resource constraint[C]//2018 IEEE Wireless Communications and Networking Conference (WCNC). 2018: 1-6.
- [13] Abd-Elmagid M A, Dhillon H S, Pappas N. A reinforcement learning framework for optimizing age of information in rf-powered communication systems[J]. *IEEE Transactions on Communications*, 2020, 68(8): 4747-4760.
- [14] Arafa A, Yang J, Ulukus S, et al. Age-minimal transmission for energy harvesting sensors with finite batteries: Online policies[J]. *IEEE Transactions on Information Theory*, 2020, 66(1): 534-556.
- [15] Yates R D. Lazy is timely: Status updates by an energy harvesting source[C]//2015 IEEE International Symposium on Information Theory (ISIT). 2015: 3008-3012.
- [16] Sun Y, Uysal-Biyikoglu E, Yates R D, et al. Update or wait: How to keep your data fresh[J]. *IEEE Transactions on Information Theory*, 2017, 63(11): 7492-7508.
- [17] Arafa A, Yang J, Ulukus S, et al. Online timely status updates with erasures for energy harvesting sensors[C]//2018 56th Annual Allerton Conference on Communication, Control, and Computing (Allerton). 2018: 966-972.

- 
- [18] Baknina A, Ulukus S, Oze O, et al. Sening information through status updates[C]//2018 IEEE International Symposium on Information Theory (ISIT). 2018: 2271-2275.
  - [19] Bhandari K, Fatale S, Narula U, et al. Age-of-information bandits[C]//2020 18th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOPT). 2020: 1-8.
  - [20] Atay E U, Kadota I, Modiano E. Aging bandits: Regret analysis and order-optimal learning algorithm for wireless networks with stochastic arrivals[Z]. 2020.
  - [21] Tripathi V, Modiano E. An online learning approach to optimizing time-varying costs of aoi [Z]. 2021.
  - [22] Li B. Efficient learning-based scheduling for information freshness in wireless networks[C]// IEEE INFOCOM 2021 - IEEE Conference on Computer Communications. 2021.
  - [23] Ceran E T, Gündüz D, György A. Reinforcement learning to minimize age of information with an energy harvesting sensor with harq and sensing cost[C]//IEEE INFOCOM 2019 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS). 2019: 656-661.
  - [24] Ceran E T, Gündüz D, György A. A reinforcement learning approach to age of information in multi-user networks with harq[J]. IEEE Journal on Selected Areas in Communications, 2021, 39(5): 1412-1426.
  - [25] Kam C, Kompella S, Ephremides A. Learning to sample a signal through an unknown system for minimum aoi[C]//IEEE INFOCOM 2019 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS). 2019: 177-182.
  - [26] Leng S, Yener A. Age of information minimization for wireless ad hoc networks: A deep reinforcement learning approach[C]//2019 IEEE Global Communications Conference (GLOBECOM). 2019: 1-6.
  - [27] Kadota I, Uysal-Biyikoglu E, Singh R, et al. Minimizing the age of information in broadcast wireless networks[C]//2016 54th Annual Allerton Conference on Communication, Control, and Computing (Allerton). 2016: 844-851.
  - [28] Kadota I, Sinha A, Uysal-Biyikoglu E, et al. Scheduling policies for minimizing age of information in broadcast wireless networks[J]. IEEE/ACM Transactions on Networking, 2018, 26 (6): 2637-2650.
  - [29] Talak R, Karaman S, Modiano E. Minimizing age-of-information in multi-hop wireless networks[C]//2017 55th Annual Allerton Conference on Communication, Control, and Computing (Allerton). 2017: 486-493.
  - [30] Talak R, Kadota I, Karaman S, et al. Scheduling policies for age minimization in wireless networks with unknown channel state[C]//2018 IEEE International Symposium on Information Theory (ISIT). 2018: 2564-2568.
  - [31] Talak R, Karaman S, Modiano E. Optimizing age of information in wireless networks with perfect channel state information[C]//2018 16th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt). 2018: 1-8.

- 
- [32] Hsu Y, Modiano E, Duan L. Age of information: Design and analysis of optimal scheduling algorithms[C]//2017 IEEE International Symposium on Information Theory (ISIT). 2017: 561-565.
  - [33] Jiang Z, Krishnamachari B, Zheng X, et al. Decentralized status update for age-of-information optimization in wireless multiaccess channels[C]//2018 IEEE International Symposium on Information Theory (ISIT). 2018: 2276-2280.
  - [34] Lu N, Ji B, Li B. Age-based scheduling: Improving data freshness for wireless real-time traffic [C]//MobiHoc '18: Proceedings of the Eighteenth ACM International Symposium on Mobile Ad Hoc Networking and Computing. New York, NY, USA: ACM, 2018: 191-200.
  - [35] Kadota I, Sinha A, Modiano E. Optimizing age of information in wireless networks with throughput constraints[C]//IEEE INFOCOM 2018 - IEEE Conference on Computer Communications. 2018: 1-9.
  - [36] Kadota I, Modiano E. Age of information in random access networks with stochastic arrivals [C]//IEEE INFOCOM 2021 - IEEE Conference on Computer Communications. 2021: 1-10.
  - [37] Sun J, Jiang Z, Zhou S, et al. Optimizing information freshness in broadcast network with unreliable links and random arrivals: An approximate index policy[C]//IEEE INFOCOM 2019 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS). 2019.
  - [38] Jiang Z, Krishnamachari B, Zheng X, et al. Timely status update in wireless uplinks: Analytical solutions with asymptotic optimality[J]. IEEE Internet of Things Journal, 2019, 6(2): 3885-3898.
  - [39] Sun J, Wang L, Jiang Z, et al. Age-optimal scheduling for heterogeneous traffic with timely throughput constraints[J]. IEEE Journal on Selected Areas in Communications, 2021, 39(5): 1485-1498.
  - [40] Jiang Z. Analyzing age of information in multiaccess networks by fluid limits[C]//IEEE INFOCOM 2021 - IEEE Conference on Computer Communications. 2021: 1-10.
  - [41] Maatouk A, Kriouile S, Assad M, et al. On the optimality of the whittle's index policy for minimizing the age of information[J]. IEEE Transactions on Wireless Communications, 2021, 20(2): 1263-1277.
  - [42] Borkar V S, Kasbekar G S, Pattathil S, et al. Opportunistic scheduling as restless bandits[J]. IEEE Transactions on Control of Network Systems, 2018, 5(4): 1952-1961.
  - [43] Chen K, Huang L. Timely-throughput optimal scheduling with prediction[J]. IEEE/ACM Transactions on Networking, 2018, 26(6): 2457-2470.
  - [44] Wang M, Liu J, Chen W, et al. On delay-power tradeoff of rate adaptive wireless communications with random arrivals[C]//GLOBECOM 2017 - 2017 IEEE Global Communications Conference. 2017: 1-6.
  - [45] Wang M, Liu J, Chen W, et al. Joint queue-aware and channel-aware delay optimal scheduling of arbitrarily bursty traffic over multi-state time-varying channels[J]. IEEE Transactions on Communications, 2019, 67(1): 503-517.
  - [46] Yang J, Ulukus S. Delay-minimal transmission for average power constrained multi-access communications[J]. IEEE Transactions on Wireless Communications, 2010, 9(9): 2754-2767.

- 
- [47] Uysal-Biyikoglu E, Prabhakar B, El Gamal A. Energy-efficient packet transmission over a wireless link[J]. IEEE/ACM Transactions on Networking, 2002, 10(4): 487-499.
  - [48] Berry R A, Gallager R G. Communication over fading channels with delay constraints[J]. IEEE Transactions on Information Theory, 2002, 48(5): 1135-1149.
  - [49] Singh R, Kumar P R. Throughput optimal decentralized scheduling of multihop networks with end-to-end deadline constraints: Unreliable links[J]. IEEE Transactions on Automatic Control, 2019, 64(1): 127-142.
  - [50] Qian Z, Wu F, Pan J, et al. Minimizing age of information in multi-channel time-sensitive information update systems[C]//IEEE INFOCOM 2020 - IEEE Conference on Computer Communications. 2020: 446-455.
  - [51] Yao G, Bedewy A M, Shroff N B. Age-optimal low-power status update over time-correlated fading channel[C]//2021 IEEE International Symposium on Information Theory (ISIT). 2021: 2972-2977.
  - [52] Przybylski S A. Cache and memory hierarchy design: a performance-directed approach[C]// 1990.
  - [53] Cohen E, Shenker S. Replication strategies in unstructured peer-to-peer networks[J]. SIGCOMM Comput. Commun. Rev., 2002, 32(4): 177-190.
  - [54] Hoang D T, Niyato D, Nguyen D N, et al. A dynamic edge caching framework for mobile 5g networks[J]. IEEE Wireless Communications, 2018, 25(5): 95-103.
  - [55] Chen W, Poor H V. Wireless caching: Making radio access networks more than bit-pipelines [J]. Network, 2021, 1(2): 146-164.
  - [56] Chen X, He L, Xu S, et al. Hit ratio driven mobile edge caching scheme for video on demand services[C]//2019 IEEE International Conference on Multimedia and Expo (ICME). 2019: 1702-1707.
  - [57] Ahmed M, Li Y, Waqas M, et al. A survey on socially aware device-to-device communications [J]. IEEE Communications Surveys Tutorials, 2018, 20(3): 2169-2197.
  - [58] Baştug E, Bennis M, Debbah M. Social and spatial proactive caching for mobile data offloading [C]//2014 IEEE International Conference on Communications Workshops (ICC). 2014: 581-586.
  - [59] Bharath B, Nagananda K, Gunduz D, et al. Caching with time-varying popularity profiles: a learning-theoretic perspective[J]. IEEE Transactions on Communications, 2018, 66(9): 3837-3847.
  - [60] Gao J, Zhang S, Zhao L, et al. The design of dynamic probabilistic caching with time-varying content popularity[J]. IEEE Transactions on Mobile Computing, 2020.
  - [61] Kam C, Kompella S, Nguyen G D, et al. Information freshness and popularity in mobile caching [C]//IEEE International Symposium on Information Theory (ISIT). 2017: 136-140.
  - [62] Yates R D, Ciblat P, Yener A, et al. Age-optimal constrained cache updating[C]//2017 IEEE International Symposium on Information Theory (ISIT). 2017: 141-145.

- 
- [63] Talak R, Karaman S, Modiano E. Distributed scheduling algorithms for optimizing information freshness in wireless networks[C]//2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC). 2018: 1-5.
  - [64] Hsu Y. Age of information: Whittle index for scheduling stochastic arrivals[C]//2018 IEEE International Symposium on Information Theory (ISIT). 2018: 2634-2638.
  - [65] Sia K C, Cho J, Cho H K. Efficient monitoring algorithm for fast news alerts[J]. IEEE Transactions on Knowledge and Data Engineering, 2007, 19(7): 950-961.
  - [66] Cho J, Garcia-Molina H. Effective page refresh policies for web crawlers[J]. ACM Trans. Database Syst., 2003, 28(4): 390-426.
  - [67] Li B, Li R, Eryilmaz A. Throughput-optimal scheduling design with regular service guarantees in wireless networks[J]. IEEE/ACM Transactions on Networking, 2015, 23(5): 1542-1552.
  - [68] Maatouk A, Kriouile S, Assaad M, et al. The age of incorrect information: A new performance metric for status updates[J]. IEEE/ACM Transactions on Networking, 2020, 28(5): 2215-2228.
  - [69] Li X, Cline D B H, Loguinov D. On sample-path staleness in lazy data replication[C]//2015 IEEE Conference on Computer Communications (INFOCOM). 2015: 1104-1112.
  - [70] Cho J, Garcia-Molina H. Synchronizing a database to improve freshness[C]//SIGMOD '00: Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data. New York, NY, USA: Association for Computing Machinery, 2000: 117-128.
  - [71] Joo C, Eryilmaz A. Wireless scheduling for information freshness and synchrony: Drift-based design and heavy-traffic analysis[J]. IEEE/ACM Transactions on Networking, 2018, 26(6): 2556-2568.
  - [72] Robbins H, Monro S. A Stochastic Approximation Method[J]. The Annals of Mathematical Statistics, 1951, 22(3): 400 - 407.
  - [73] Saad D. Online algorithms and stochastic approximations[J]. Online Learning, 1998, 5: 6-3.
  - [74] Neely M J. Fast learning for renewal optimization in online task scheduling[Z]. 2021.
  - [75] Tang H, Chen Y, Wang J, et al. Sending timely status updates through channel with random delay via online learning[C]//IEEE INFOCOM 2022 - IEEE Conference on Computer Communications (INFOCOM 2022). London, United Kingdom (Great Britain), 2022.
  - [76] Tang H, Wang J, Song L, et al. Scheduling to minimize age of information in multi-state time-varying networks with power constraints[C]//2019 57th Annual Allerton Conference on Communication, Control, and Computing (Allerton). 2019: 1198-1205.
  - [77] Tang H, Wang J, Song L, et al. Minimizing age of information with power constraints: Multi-user opportunistic scheduling in multi-state time-varying channels[J]. IEEE Journal on Selected Areas in Communications, 2020, 38(5): 854-868.
  - [78] Tang H, Ciblat P, Wang J, et al. Cache updating strategy minimizing the age of information with time-varying files' popularities[C]//IEEE Information Theory Workshop, ITW 2020, Riva del Garda, Italy, April 11-15, 2021. IEEE, 2020: 1-5.
  - [79] Tang H, Ciblat P, Wang J, et al. Age of information aware cache updating with file- and age-dependent update durations[C]//18th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks, WiOPT 2020, Volos, Greece, June 15-19, 2020. IEEE, 2020: 1-6.

- [80] Tang H, Wang J, Tang Z, et al. Scheduling to Minimize Age of Synchronization in Wireless Broadcast Networks with Random Updates[C]//IEEE International Symposium on Information Theory (ISIT). 2019.
- [81] Zhang Q, Tang H, Wang J. Minimizing the age of synchronization in power-constrained wireless networks with unreliable time-varying channels[C]//IEEE INFOCOM 2020 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS). 2020: 936-941.
- [82] Tang H, Wang J, Tang Z, et al. Scheduling to minimize age of synchronization in wireless broadcast networks with random updates[J]. IEEE Trans. Wirel. Commun., 2020, 19(6): 4023-4037.
- [83] Arafa A, Yates R D, Poor H V. Timely cloud computing: Preemption and waiting[C]//2019 57th Annual Allerton Conference on Communication, Control, and Computing (Allerton). 2019: 528-535.
- [84] Ross S M. Applied probability models with optimization applications[M]. Courier Corporation, 2013.
- [85] Rudin W. Real and complex analysis p. 2[M]. McGraw-Hill, 1970.
- [86] Neely M J. Stochastic network optimization with application to communication and queueing systems[J]. Synthesis Lectures on Communication Networks, 2010, 3(1): 1-211.
- [87] Tajan R, Ciblat P. Information-theoretic multi-user power adaptation in retransmission schemes [C]//2016 IEEE 17th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC). 2016: 1-5.
- [88] Altman E. Constrained markov decision processes: volume 7[M]. CRC Press, 1999.
- [89] Bertsekas D P, Scientific A. Convex optimization algorithms[M]. Athena Scientific Belmont, 2015.
- [90] Sennott L I. Average cost optimal stationary policies in infinite state markov decision processes with unbounded costs[J]. Operations Research, 1989, 37(4): 626-633.
- [91] Puterman M. Markov decision processes[M]. Wiley, 1994.
- [92] Weber R R, Weiss G. On an index policy for restless bandits[J]. Journal of Applied Probability, 1990, 27(3): 637-648.
- [93] Hsu Y P, Modiano E, Duan L. Scheduling algorithms for minimizing age of information in wireless broadcast networks with random arrivals[J]. IEEE Transactions on Mobile Computing, 2020, 19(12): 2903-2915.
- [94] Sennott L I. Constrained discounted markov decision chains[J]. Probability in the Engineering and Informational Sciences, 1991, 5(4): 463-475.
- [95] Stefanov S. Applied optimization: volume 53 separable programming: Theory and methods [M]. SpringerScience, 2001.
- [96] Beutler F, Ross K. Optimal policies for controlled markov chains with a constraint[J]. Journal of Mathematical Analysis and Applications, 1985, 112(1): 236 - 252.
- [97] Bastopcu M, Ulukus S. Age of Information for Updates with Distortion[C]//IEEE Information Theory Workshop (ITW). 2019.

- [98] Jorswieck E, Larsson E. Monotonic optimization framework for the two-user MISO interference channel[J]. IEEE Transactions on Communications, 2010, 58(7): 2159-2169.
- [99] Björnson E, Zheng G, Bengtsson M, et al. Robust monotonic optimization framework for multicell MISO systems[J]. IEEE Transactions on Signal Processing, 2012, 60(5): 1-16.
- [100] Björnson E, Jorswieck E. Optimal Resource Allocation in Coordinated Multi-Cell Systems: volume 9[M]. Now Publishers, 2013: 113-381.
- [101] Gittins J, Glazebrook K, Weber R. Multi-armed bandit allocation indices[M]. John Wiley & Sons, 2011.
- [102] Weber R R, Weiss G. On an index policy for restless bandits[J]. Journal of Applied Probability, 1990, 27(3): 637-648.
- [103] Bertsekas D P. Dynamic programming and optimal control[M]. 2nd ed. Athena Scientific, 2000.

## 致 谢

感谢我的导师王劲涛教授，给我提供了广阔的天空自由探索，鼓励我去探索和寻找自己真正感兴趣的课题，小心翼翼保护着我们的好奇心和科研兴趣。特别感谢导师对我生涯规划的指导，在我为各种选择纠结和犹豫时，老师总能帮我理清头绪带着我倾听自己内心的声音。谢谢宋健教授百忙中对大方向的把控和对每一篇论文的修改和指导。

特别感谢本科在慕尼黑交换时讲授 *<Information Theory>* 课程的 Gerhard Kramer 教授，这是我第一次感受到信息论和通信理论研究的简洁优美和大气磅礴！也感谢 Kramer 创办的信息论暑期学校给我提供了接触到“信息年龄”的机会和成为 Michèle 学生的可能性。感谢巴黎高等电信学院的 Michèle Wigger 教授和 Philippe Ciblat 教授对我科研和生活的关爱。非常想念农历春节前后 Michèle 亲自下厨开 party 炸的春卷，还有在我自己跟自己急得掉眼泪时她充满安全感的拥抱。那些午后在办公室落地窗的阳光下一起改 paper、在学校走廊上喝咖啡讨论卢浮宫什么时候人最少、在她的两个宝贝 monsters 敲桌子的背景音乐中开香槟闲聊瑞士最佳旅游季节的时光让我慢慢理解了“慢慢走，欣赏啊”的科研和生活态度。Michèle 是做信息论的，但当发现我更喜欢网络优化的方法切入课题时，仍然尽己所能帮助我，努力创造了许许多多的机会让我认识她的好朋友们。我喜欢每次讨论时 Michèle 看向我们那双清澈明亮的大眼睛，谢谢那漂亮蓝眼睛中闪烁的光所传递的自信、鼓励和无穷无尽的力量，感谢她在我博士后申请时无私的关心和帮助。她是两个孩子的妈妈，也是我们温柔善良偶尔严厉的 research mom，还是可以一起品尝咖啡聊人生的朋友。不管是做妈妈还是做老师或者 Mentor，我都希望自己能和 Michèle 一样陪伴着孩子们快乐成长，也祝 Martis 和 Céline 两个宝宝成长快乐；感谢罗格斯大学的 Roy Yates 教授、奥本大学的 Yin Sun 老师以及耶鲁大学的 Leandros Tassiulas 教授，和老师们合作的过程中受益良多也特别开心。接触到信息年龄是因为孙老师的讲座，博士最后一个工作也有幸得到了孙老师很多帮助，实现了下一个工作一定要比上一个做的好的目标，是一个美好的结束。感谢周世东教授、周盛副教授对论文工作的宝贵建议和长期以来的鼓励，感谢牛志升教授的严格要求和在答辩现场的提问。感谢沈渊副教授在博一《统计推断》课程中对我大作业的宝贵建议，鼓励我跳出舒适圈寻找自己更感兴趣的课题，感谢老师在我博后申请时的推荐和帮助。感谢高飞飞副教授逻辑丝滑严谨的研究生课程，科研方法论和论文修改指导。感谢王禹皓老师教授了我学生阶段最后一门课，规律地

读 paper 和接受新知识帮助我调整了博四糟糕的状态，还要感谢老师对于最后一个工作中不可达界证明的帮助和各方面的鼓励。感谢杨朋昆老师帮忙收拾我捅出的信息论本子巨锅并带团队飞疯狂通宵写本子，卓越的 ddl 型选手广阔的知识面、准确的直觉以及对于逻辑、表达和条理的严格要求让我受益良多，逼一逼我才知道一周可以学这么多东西，让我认识到自己之前是有多划水：（感谢白铂老师和吕晖老师本科的培养，谢谢 Jing Yang, Lele Wang 和 Zhiying Wang 老师们一起开会一起吃喝玩乐留下的美好回忆和分享鼓励，感谢 Tara Javidi 教授在 WITHITS workshop 对于 research pitch 的指导，还要感谢 Tara 和汪玉教授对于生涯规划的宝贵建议。

感谢我在华为实习期间的 mentor 李宗鹏老师以及我的同事刘少腾、缪丽华、陈昊、陈祥玉、张子君，以及顾问 Frank Kelly 老爷爷，实习的四个月让我充分感受了优美的网络优化理论是如何在系统中发挥作用的；感谢我在微软亚洲研究院实习期间的 mentor 刘畅老师带我开拓视野，一起发现机器学习之美。

谢谢姜之源学长和孙宇璇学姐在我需要时总能分享他们宝贵的经验，谢谢何龙桩学长和唐子涵学长在我低年级孤立无援的时候耐心听我讲问题帮我想办法。谢谢一起读博、开会、玩耍和发泄焦虑的小伙伴们（然而现在已经喊不少人老师了）：陈斌老师、闫起发老师、李龙光老师、杨恒杰、刘一锐、刘柏，尤其要谢谢杨教授各类 Proposal，他和刘一锐在开会溜出去玩的路上还在讨论学术问题，我不仅啥也没听懂还在试图理解两位编码大佬时丢了钱包，最终在陈斌老师广播信道的帮助下扬名 ISIT2019；特别感谢孙径舟同学陪我讨论问题一起修改论文，谢谢组里读博读硕、做毕设和 SRT 的小朋友们：邓程昊、陈宇超、张启宁、占文皓、侯欣语、何雨桥、谭亦钦、陈国智，谢谢你们给我尝试带小朋友们飞的机会、让我体验和大家一起做科研的快乐，和你们的讨论让我收获超多，不断在提高我的表达能力和严谨性，我为你们五彩斑斓的未来感到自豪！感谢实验室的孙跃学长、张肖晖学长、陆智麟、史旭、杨素芳、王薛涵、曾睿和王浩宇小朋友，谢谢你们让实验室充满欢乐，祝浩宇中考顺利。谢谢雨婷姐姐和惠铄迪一起共度博一快乐的辅导员生活。谢谢周小菁老师、晓夏姐和艺体队的队友们，艺术体操和舞蹈点亮了读博期间最无助的日子，那些耗软度、拉素质的酸爽和队友的欢笑声是我抑郁焦虑时每周生活的盼头，老师教练们对热爱的事业纯粹的追求和理想主义深深感染了我，我会向两位纯粹的理想主义者学习，把追求美作为一辈子的目标。

最后感谢我的爸爸妈妈，感谢爸妈的养育之恩，家的港湾给我精神上的鼓励和无限支持。谢谢爸爸妈妈给我发现美好的双眼和感受美好的心脏，让我充分享受探索新知识的快乐。他们总是用自己的乐观、开明和积极向上带着我感受生活的美好。我无从报答他们对我无条件的爱和宽容，唯有加倍努力追求幸福。

声 明

---

声 明

本人郑重声明：所呈交的学位论文，是本人在导师指导下，独立进行研究工作所取得的成果。尽我所知，除文中已经注明引用的内容外，本学位论文的研究成果不包含任何他人享有著作权的内容。对本论文所涉及的研究工作做出贡献的其他个人和集体，均已在文中以明确方式标明。

签 名：汤皓娟 日 期：2021.12.14

## 个人简历、在学期间完成的相关学术成果

### 个人简历

1995年1月11日出生于江苏省南京市。

2013年9月保送清华大学电子工程系电子信息科学与技术专业，2017年7月本科毕业并获得工学学士学位。

2017年9月免试进入清华大学电子工程系攻读信息与通信工程专业博士至今，期间获得国家奖学金、海英之星奖学金、腾讯犀牛鸟精英人才培养计划、电子之星奖学金、IEEE Globecom2021 Student Travel Grant、微软亚洲研究院“明日之星”优秀实习生等荣誉。

### 在学期间完成的相关学术成果

#### 学术论文：

- [1] **Tang H**, Wang J, Song L., Song J. Minimizing Age of Information With Power Constraints: Multi-User Opportunistic Scheduling in Multi-State Time-Varying Channels. *IEEE Journal on Selected Areas in Communications*, 2020, 38(5): 854-868. (SCI 收录, 检索号: LR4QY, 影响因子: 11.42)
- [2] **Tang H**, Wang J, Tang Z, Song J. Scheduling to Minimize Age of Synchronization in Wireless Broadcast Networks With Random Updates. *IEEE Transactions on Wireless Communications*, 2020, 19(6): 4023-4037. (SCI 收录, 检索号: MC2VB, 影响因子: 6.779)
- [3] **Tang H**, Wang J, He L. Off-Grid Sparse Bayesian Learning-Based Channel Estimation for MmWave Massive MIMO Uplink. *IEEE Wireless Communications Letters*, 2019, 8(1): 45-48. (SCI 论文, 检索号: HM5IU, 影响因子: 4.66)
- [4] **Tang H**, Ciblat P, Wang J, Wigger M, Yates R. Cache Updating Strategy Minimizing the Age of Information with Time-Varying Files' Popularities. *Information Theory Workshop (ITW)*, Venice, Italy, 2021, pp. 1-6. (EI 收录, 检索号: 20213410816917)
- [5] **Tang H**, Ciblat P, Wang J, Wigger M, Yates R. Age of Information Aware Cache Updating with File- and Age-Dependent Update Durations,” *International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks Workshop (WiOPT)*, Volos, Greece, 2020, pp. 1-6. (EI 收录, 检索号: 20204009297976)
- [6] **Tang H**, Wang J, Song L, Song J. Scheduling to Minimize Age of Information in Multi-State Time-Varying Networks with Power Constraints. *Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Monticello, IL,

- USA, 2019. (EI 收录, 检索号: 20200308042337)
- [7] **Tang H**, Wang J, Tang Z, Song J. Scheduling to Minimize Age of Synchronization in Wireless Broadcast Networks with Random Updates. IEEE International Symposium on Information Theory (ISIT), Paris, France, 2019, pp. 1027-1031. (EI 收录, 检索号: 20194207537797)
- [8] **Tang H**, Wang J, Zheng Y. Covert communications with extremely low power under finite block length over slow fading. IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), Honolulu, HI, 2018, pp. 657-661. (EI 收录, 检索号: 20183105634684)
- [9] **Tang H**, Wang J, He L. Mutual information maximization for optimal spatial modulation MIMO system. IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB), Cagliari, 2017, pp. 1-4. (EI 收录, 检索号: 20173304047231)
- [10] **Tang H**, Hou X, Wang J, Song J. Joint Link Rate Selection and Channel State Change Detection in Block-Fading Channels. Global Communications Conference (已录用)
- [11] **Tang H**, Chen Y, Wang J, Sun J, Song J. Sending Timely Status Updates through Channel with Random Delay via Online Learning. IEEE Conference on Computer Communications 2022 (已录用)
- [12] Chen Y, **Tang H**, Wang J, Song J. Optimizing Age Penalty in Time-Varying Networks with Markovian and Error-Prone Channel State. Entropy 23(1): 91 (2021) (SCI 收录, 检索号: PV7BN, 影响因子: 2.587)
- [13] Chen Y, **Tang H**, Wang J. Optimizing Age of Information in Multicast Unilateral Networks”, International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB), Paris, 2020, pp. 1-4. (EI 收录, 检索号: 20211410172302)
- [14] Zhan W, **Tang H**, Wang J. Delay Optimal Cross-Layer Scheduling Over Markov Channels with Power Constraint. International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB), Paris, 2020, pp. 1-6. (EI 收录, 检索号: 20211410172352)
- [15] Zhang Q, **Tang H**, Wang J. Minimizing the Age of Synchronization in Power-Constrained Wireless Networks with Unreliable Time-Varying Channels. IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), Toronto, ON, Canada, 2020, pp. 936-941. (EI 收录, 检索号: 20203909247862)
- [16] Tan Y, **Tang H**, Wang J, Song J. Early Drop: A Packet-Dropping Incentive Rate Control Mechanism to Keep Data Fresh under Heterogeneous QoS Requirements. International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB), Chengdu, 2021, pp. 1-6. (EI 收录, 检索号: 20214110998320)
- [17] C. Liu, X. Sun, J. Wang, **H. Tang**, T. Li, T. Qin, W. Chen, and T. Liu, “Learn-

- ing Causal Semantic Representation for Out-of-Distribution Prediction” , accepted NeurIPS2021
- [18] C. Liu, **H. Tang**, T. Qin, J. Wang and Tie-Yan Liu, “On the Generative Utility of Cyclic Conditionals” , accepted NeurIPS2021.

## 指导教师学术评语

论文采用随机优化、李雅普诺夫分析、马尔可夫决策过程等工具优化了随机网络中的信息时效性，并分析了所提算法理论性能。选题具有重要的理论意义和应用价值。论文工作主要创新点如下：

1、针对链路传输时延统计信息闭式解未知，提出了优化接收端平均信息年龄的在线采样算法。证明所提算法平均信息年龄相比与统计信息已知时的增量为  $O(\log K/K)$ ，即所提算法得到信息年龄随着采样次数增加收敛到传输时延统计信息已知时的最小值，这一增量与任意因果算法的极小极大下界阶数匹配，即所提算法满足极小极大率最优。在数值仿真中所提算法可以有效降低链路接收端信息年龄。

2、考虑多传感器网络中的带宽、功率双约束，提出了满足网络带宽、节点传输功率限制的采样、功率控制和调度联合设计算法。论文工作证明了所提算法与全网平均年龄下界之间差距随着网络中节点数目  $N$  增多，以速率  $O(1/\sqrt{N})$  衰减，即所提算法在大规模网络中的渐近最优性。

3、考虑用户通过基站端缓存从远程服务器索取状态信息副本，针对用户需求时变、基站端缓存副本更新时长时变两种场景，设计了满足带宽约束的基站端缓存更新方案。证明了所提算法在状态信息数量足够多时的渐进最优性。

4. 考虑状态信息产生与发送具有双重随机性，基于无休止多臂赌博机提出了一种满足带宽约束的调度方案，仿真中所提算法得到的全网平均同步年龄接近理论下界。

## 答辩委员会决议书

论文研究无线网络中的信息时效性优化，选题具有重要的理论意义和应用前景。

论文的创新性成果为：

- 1、针对链路传输时延独立同分布且未知的场景，提出了一种最小化平均信息年龄的在线采样算法，证明了所提算法的渐近最优性；
- 2、针对信道时变且带宽受限的场景，提出了一种最小化平均信息年龄的多传感器采样、传输联合优化算法，证明了所提算法在大规模网络中的渐近最优性；
- 3、针对基站端缓存更新周期时变、用户对内容需求时变的场景，提出了一种最小化用户索取信息副本信息年龄的缓存更新算法，及一种带宽限制下同步年龄最小的推送策略。

论文工作表明作者在信息与通信工程学科具有坚实宽广的基础理论和系统深入的专门知识，具有独立从事科学研究工作的能力。论文写作条理清楚、逻辑严密、图表格式规范，答辩叙述清楚、回答问题正确。经答辩委员会 6 人无记名投票表决，一致同意通过论文答辩，并建议授予汤皓玥工学博士学位。