系统模型

## **一、MDP = (S, A, P, R, γ)**

### **1. 状态空间 S**：【S = C×L×B】

* 信道状态矩阵：C = [C1, C2, ..., CN] T ∈ RN

采用离散化信道增益等级（如Good/Medium/Bad三级量化），通过CQI反馈构建，*Ci*​ 代表第 *i* 用户的信道状态，总共有 *N* 个用户。每个用户的信道状态通过离散化的量化级别来描述。

缓冲区状态矩阵：

* B = [B1, B2, ..., BN]T ∈ RN

表示各用户解码缓冲区占用量（归一化至[0,1]区间）, 归一化的目的是为了将各用户的缓冲区占用量转换为相同的标准，使得不同用户之间的状态可以直接进行比较。0 表示缓冲区空，1 表示缓冲区满。

*N* 表示系统中存在的用户数量，*Bi*​ 是第 *i* 个用户的缓冲区状态。每个用户的状态都可能因为接收、解码或丢失数据包而变化。

视频层状态：

* 视频层状态L：{l₁,...,lₘ}，m表示当前传输的SVC层数。

L ∈ {Base Layer, EL1, EL2} 表示当前传输的SVC层次。

### **2. 动作空间 A**：【*A*={(*l*,*mcs*) ∣ *l*∈{0,1,2}, *mcs*∈{1,2,...,*M*}}】

传输策略组合：

*A*={(*l*,*mcs*) ∣ *l*∈{0,1,2}, *mcs*∈{1,2,...,*M*}}

*l*=0: 仅传输基础层（Base Layer）

*l*=1: 传输BL+EL1

*l*=2: 传输BL+EL1+EL2

**MCS级别 *mcs***：如MCS1（QPSK）、MCS2（16QAM）、MCS3（64QAM）等。

(1,2)：以MCS2传输BL+EL1；

(2,3)：以MCS3传输所有三层。

### **3. 状态转移概率 P**：【P(s'|s,a)】

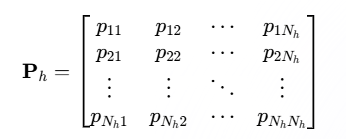
*P*(*s*′∣*s*,*a*) = *PC*​(*c*′∣*c*) ⋅ *PB​*(*b*′∣*b*,*l*,*mcs*) ⋅ *PL*​(*l*′∣*l*,*c*,*mcs*)

**信道转移矩阵***PC*​(*c*′∣*c*)：

* **作用**：描述在当前状态 s 和动作 a 下，系统转移到下一状态 s'​的概率。
* **建模原理**：
* **信道状态转移**：基于马尔可夫链的信道衰落模型（Gilbert-Elliott模型）。
  + **模型原理**：

假设信道状态在离散时间步长内仅依赖前一状态（马尔可夫性）。

定义状态转移概率矩阵 P*h*​，其中元素 *pij*​表示从状态 *i* 转移到状态 *j* 的概率：

* + 

满足每行概率和为1：

* + **时间相关性**：转移概率反映了信道的时间相关性。

若 *pii*​ 较大，说明信道状态相对稳定（如室内静态场景）。

若非对角元素占主导，说明信道快速波动（如高速移动场景）。

* + **对QoS的影响**：

频繁切换低增益状态会增加视频卡顿风险，需动态降低视频层数。

高增益状态的持续概率高，可支持多层传输以提升画质。

**缓存状态转移***PB*​(*b*′∣*b*,*l*,*mcs*)：

* bt+1 = max(0, bt+R(l, mcs) ⋅ Δt ⋅ (1 − BLER(c,mcs)) – μ ⋅ Δt)
* R(l,mcs)：MCS和层数联合决定的传输速率（如MCS3支持更高速率）；
* BLER(c,mcs)：信道状态 c 下MCS的误码率（需查表或经验公式）；

1. **查表法**：

对于不同的 MCS 和信道状态，可以使用 **查表法**，即预先计算并存储对应的 BLER 值。例如：

* 在不同的信道条件（如 Good/Medium/Bad）下，针对每种 MCS（如 MCS1, MCS2, MCS3）有一个对应的 BLER 值。
* 查表方法是最简单且高效的方式，通常采用基于实际网络条件的实验结果或仿真结果。

1. **经验公式法**：

在一些情况下，可以使用经验公式来估算 BLER。假设你有 SNR（信噪比）的值，可以通过一些数学公式来估算 BLER，具体公式依赖于调制方式和编码方式。例如，QPSK、16QAM、64QAM 等不同的调制方案会有不同的错误率计算公式。

* 对于 QPSK，可以使用类似于：

BLER =

* 对于高阶调制（如 16QAM, 64QAM），BLER 会随着 SNR 的减小而迅速增大。

1. **物理层仿真**：

基于信道模型（例如，AWGN信道、Rayleigh衰落等），可以进行物理层仿真来计算 BLER。这需要用到具体的信道模型和调制方案，通过仿真得到 BLER 值。

**视频层切换***PL*​(*l*′∣*l*,*c*,*mcs*)：

* 受信道容量约束，若 R(l′,mcs)≤C(c,mcs)（信道容量支持），则允许切换至 *l*′。

在进行视频层切换时，**所选的视频传输速率（R(l′,mcs)）不能超过当前信道的容量（C(c,mcs)）**，否则将无法保证数据的正常传输。

### **4. 奖励函数 R**：【R(s,a)】

多目标加权设计：

R(s,a) = *α Q*(*lt*​) – *β* ( *b*max​−*bt*​ ) −*γp(l,mcs)*

* **视频质量***Q*(*lt*​)：视频质量（如PSNR），随层数增加而提升。
* bmax−bt：缓存饥饿惩罚。
* **功率消耗 *p*(*l*,*mcs*)**：功率消耗惩罚

*p*(*l*,*mcs*) = *p*base​ + *l* ⋅ *p*layer​ + *mcs* ⋅ *p*mcs​

* *p*base​：基础功耗；
* *p*layer​：每增加一层的额外功耗；
* *p*mcs​：高阶MCS的功率增量（如MCS3比MCS1多消耗20%功率）。
* *α*, *β*, *γ*：权重系数，需通过仿真调优。

权重系数通过AHP层次分析法确定

### **5.折扣因子γ：**

为了避免出现状态循环的情况，系统对于将来的预测并不一定都是准确的，所以要打折扣。很显然越靠近1，考虑的利益越长远。

折扣因子 *γ* 反映了系统对**即时奖励**与**未来奖励**的权衡：

**γ→0**：系统更关注**当前时刻的奖励**，倾向于“短视”策略（如为节省资源立即降低视频层数，忽略未来可能的信道改善机会）。

**γ→1**：系统更重视**长期累积奖励**，倾向于“前瞻性”策略（如牺牲当前资源消耗以维持稳定的QoS，即使未来可能获得更高收益）。

## **二、动态优化建模**

**1. 策略优化目标**

求解最优策略π\*:【S→A，最大化长期折扣奖励：】

其中 π : S→A为策略函数，γ∈[0,1) 为折扣因子。

**2. 贝尔曼方程**

采用值迭代算法求解最优值函数V\*(s)：

Vk​(s′)：上一次迭代中，状态 s′ 的价值

**作用**：找到全局最优策略，但计算复杂度高（适用于离线场景）。

**3. 约束条件**

时延约束：Pr{*Bi*​(*t* )< *B*min​} ≤ *ϵ* ∀*i*∈*G*

Bi​(t)：第 i 个用户在时刻 t 的缓冲区占用率（通常归一化到 [0, 1]）

Bmin​：缓冲下限阈值（如 0.2），小于这个值就会出现“播放中断”风险（即卡顿）

ε：允许出现“缓冲区低于下限”的概率上限（例如，设置为 5% 或 10%）

带宽约束：*R*(*l*,*mcs*) ≤ *W*⋅log2​(1 + ​)或者R(l,mcs)≤C(c,mcs)

**信道容量 Ci** 通常根据香农公式计算：

*W*：系统带宽；

SINR：信噪加干扰比，是信道状态的直接反映

*Ci*：用户 *i* 的信道容量

G：组播组成员集合