基于马尔科夫决策过程的建模思路：

状态向量s=(时间t，占空比δt，状态s)，时间t可以是一天/一周的各个时间段，占空比是时间间隔Δt内的占空比；

动作a=占空比的变化；

建立传感器能量消耗速度与占空比之间的关系f1：

：传感器消耗能量的速率（J/h）

δ：占空比

：传感器的采样速率（bit/s）

：传感器采样1bit所消耗的能量（J）

：传感器的发送速率（bit/s）

：传感器发送1bit消耗的能量（J）

：传感器放大1bit消耗的能量（J/）

：传感器到汇聚节点的距离（），其中n为路径损失系数

：传感器在休眠状态下的能量消耗速率（J/s）

参考文献使用MICA节点，数据为

|  |  |
| --- | --- |
| n | 2 |
|  | 0.937μJ/bit |
|  | 0.787μJ/bit |
|  | 30μJ/s |
|  | 40 Kbps |

奖赏函数r1为对能量消耗速度的度量，r2为感应延迟时间（h），R=r1和r2的某种组合；

期望效果：可以针对每一天/每一周的发病规律，自动调整占空比，获取人体相关信息，节省能量，但也要考虑及时性，能够应对突发变化。

表征学习效果，可以用Aerage Reward和Frequence of High Delay

生成模拟数据：拟合曲线表达式，纵向关联（均值与季节相关）

延迟考虑通信延迟

后续可考虑两维马尔科夫模型

状态向量把时间序号改为涨幅？（类似大跌、小跌、小涨、大涨的四个状态）

改进qlearning、考虑特征数量的完备性、考虑模式切换的能耗

ε的表达式

奖赏思路：在满足延迟的条件下，能耗降低给与奖励，在不满足延迟的条件下，能耗增加给予奖励，能耗降低给予惩罚