随着移动通信的发展，大量基站的铺设满足了海量不同类型设备的接入需求。As a consequence，通信设施能源消耗大幅提升，造成了大量碳排放，这阻碍了移动通信的进一步发展。实际上，基站占据了总消耗的60%-80%，并且基站流量负载大多时候处于较低值（ITJ18）。因此，在低负载时控制基站进入低功耗状态的动态睡眠控制受到了大量兴趣。

基站休眠控制分为整体和组件两种模式。ITJ18，TN21中采用数据驱动的方法建模流量模型，然后用强化学习方法控制基站的分钟级休眠。TMC20，TWC22研究了异构接入网（HetNet）中微基站群的流量卸载和睡眠。这些工作都强调基站整体而不是部分组件的睡眠，无法充分挖掘能效潜力。PIMRC13，INFOCOM11中根据流量负载变化调整基站多组模块的休眠与唤醒。Ali El Amine等在TNSM22中研究了4中固定的睡眠等级，做到了微秒级的精细睡眠，类似的工作还有WCNC19，NCA19。然而，这些工作没有考虑睡眠和唤醒切换的高频操作导致的能耗。

5G支持大量不同的用户案例，这意味着基站同一时间面对着多样化的流量模式和QoS需求。网络切片和网络功能虚拟化技术把通用物理设施隔离并虚拟化为可分配资源，交给切片调用（WC19）。在这样的场景下，现有的休眠控制方法难以根据切片的具体流量模式和QoS需求进行自适应调整，以取得睡眠和QoS之间的trade-off。

我们提出了一个QoS and Traffic Aware的多智能体框架，用来控制基站的切片组件的毫秒级睡眠，从而达到全局的能效与服务质量的平衡。通过设计全局增益奖励，我们把切片组件单独的睡眠控制优化转化为全局优化，并利用多智能体强化学习方法分别控制切片组件的睡眠，以避免维度灾难。据我们所知，这是第一个根据不同的切片流量模式和QoS需求来控制基站组件睡眠的框架。

具体来说，我们的贡献如下：

1. 提出一种切片QoS和Traffic Aware的基于多智能体强化学习的基站休眠控制框架，在保证QoS的前提下，充分挖掘睡眠潜力，并避免了维度灾难问题。
2. 所提框架支持不同RL算法。我们部署了目前SOTA的RL算法，包括： DDPG、TD3、PPO、SAC（to do，来个RL综述），代码is available：
3. 通过系统仿真，对比了所提框架下不同RL算法的训练和表现效果
4. 与Baseline方法进行对比，取得了显著提升（具体是xxx）

**待定：**

5. 基站休眠和激活本质是一种任务调度行为，但我们同时在资源分配和天线功率上与其进行匹配；基站活跃期间的提高天线功率、采用激进的资源分配策略，进一步提升休眠潜力和资源利用效率（具体方法待定）

6. 针对实际环境中流量波动影响RL离线训练的问题（临界值不稳，难收敛），提出xxx解决方案（针对离线学习RL算法的经验回放池，包括：DDPG、TD3、SAC，注意PPO不是）