无线传感器网络

课程报告

**题目：基于多智能体强化学习的云辅助多UAV收费系统的协调调度**

|  |  |
| --- | --- |
| 姓名 | **刘佳鸿** |
| 学号 | **202382060020** |
| 学院 | **计算机科学与工程学院** |
| 班级 | **硕研23-1班** |

**2023年 11 月 30 日**

# 背景

无人机移动基础设施部署的最新进展使网络组织和管理变得灵活可靠。特别是，这种需求正在急剧增加，目前被认为是6G网络场景的主要部分之一。对收集传感数据的大量物联网（IoT）平台进行智能高效的管理，即所谓的协调，有望成为6G网络的主要兴趣之一，连接数量应该是5G要求的10倍左右，以满足6G要求（即每平方公里107个联网设备）。此外，物联网平台预计将具有先进的人工智能（A.I.）和深度学习功能，用于检测和处理各种多模型真实世界传感数据。无人机平台可以被视为一种空中物联网设备，由于其高移动性和灵活的服务覆盖范围的优势，吸引了实时传感、监测和网络应用的兴趣。根据Business Insider Intelligence的报告，预计2021年将有超过2900万架无人机投入使用。

尽管无人机平台有望在6G网络场景中发挥关键作用，但它也面临着一些研究挑战。其中，一个主要问题是无人机平台非常耗电。在现代嵌入式无人机平台中，由于电池的限制，操作时间被限制在几个小时。因此，在无人机移动网络中，节能通信是必不可少的。为了使基于无人机的网络维护在能源效率方面具有鲁棒性和可靠性，最近有几个关于开发无人机平台充电新方法的研究成果。其中，对充电设施的使用进行了积极探讨。本文描述了一种由充电塔组成的充电基础设施。由于充电塔安装在地面上，它们能够不受严格限制地获取电源。此外，我们考虑的充电塔配备了发电和储存装置，如光伏发电机和储能系统，以提供高效和可靠的能源。需要注意的是，由于发射塔可以有多个充电面板用于无线供电/能量传输，每个发射塔可以同时为多架无人机充电。此外，充电塔可以在ESS中共享自己的能源，也可以从其关联的公用事业公司购买能源。这种合作能源共享本质上是为了最大限度地减少公用事业公司的能源购买，以实现经济高效的系统运行。

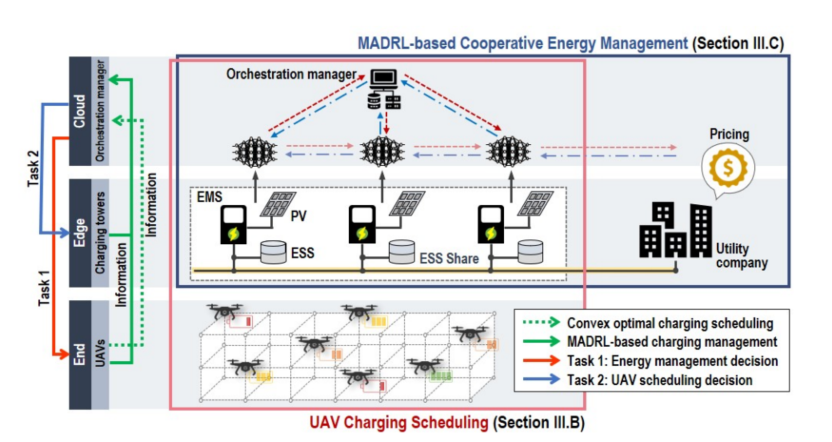


Fig1.编排调度和收费学习框架

# 2. 多智能体强化学习(Multi-agent reinforcement learning)

## 2.1 基本知识

在本文中，马尔可夫决策过程（MDP）被定义为元组（），其中S是所有有效状态的有限集，A是所有有效动作的有限集。函数是转移概率函数，其中是如果智能体在状态中开始执行动作，则转移到状态的概率。函数表示奖励函数，其中。MDP具有有限的时间范围T，求解MDP意味着找到策略，其中用θ参数化（例如，神经层的权重和偏差）；指定哪个动作 必须在每个s中执行，以最大化在有限时间T期间收到的折扣累积奖励。

## 2.2 DQN

深度Q网络（DQN）是一种无模型强化学习方法，旨在学习具有高维状态空间的最优策略。DQN受到Q-learning的启发，其中使用神经网络来近似Q函数，经验回放D和目标网络是用于稳定优化的两个关键特征。智能体的经验被存储在经验缓冲器D＝(e1,e2,…,et)中，并且被周期性地重新采样以训练Q网络。在第i次训练迭代中，使用小批量重采样经验来更新策略的参数，其中损失函数定义为

其中是目标网络参数，在每个预定义步骤中使用Q网络参数更新，并采用随机梯度下降法优化损失函数。

## 2.3 MADRL

前面提到的方法是用单智能体系统设计的，在这个系统中，智能体只考虑其自身行动的结果。然而，在多智能体系统中，智能体需要同时观察其自身行为的效果以及其他智能体的行为，这一特性会不断重塑环境并导致非平稳性。因此，在多智能体系统中，前人的收敛理论通常无法得到保证。

因此，提出了各种MADRL算法用于多智能体协作和协调。在集中式MADRL算法中，多智能体的状态和动作被表示为单个深度神经网络的输入和输出，因此，只有一种策略决定MADRL的行动。此外，集中式策略收集所有可观察信息（完全可观察信息），并一次确定所有智能体的动作。

# 3. 多无人机充电的协调调度与调度

## 3.1框架描述和设计原理

如图1所示，我们考虑的参考系统包含三个主要组件，即编排管理器(云)，充电塔(能量边缘)和无人机(端)，旨在进行以下两大能量优化智能系统维护和运营。

* 任务1(凸最优充电调度):用于无人机与充电塔之间的凸最优充电调度决策;并对计划对进行相应的能量资源分配数量决策。这一行动需要从无人机和充电塔收集信息;并且需要高性能的计算资源，即相关的业务流程管理器需要利用云计算资源。
* 任务2(基于协同madrl的充电塔EMS学习):针对基于协同madrl的充电塔间能源智能管理。使用基于commnet的合作MADRL算法，因为它可以实现成本效益的能源交易和共享。MADRL需要实时使用庞大的数据集进行数据密集型深度学习计算，本质上需要云辅助的编排管理器。

在业务流程管理器的协调下，两个任务相互关联。任务1优化后的充电计划会对计划充电塔的能量消耗产生影响，最终影响到充电塔之间的协同能量共享。需要注意的是，该充电调度是通过凸规划公式进行优化的，因此可以减少从公用事业公司购买充电塔的电量。另一方面，充电塔之间的协同能量共享也会对充电调度产生影响，因为它决定了每个充电塔的能量状态，这是充电调度优化的主要参数之一。需要注意的是，合作共享的能量可以让充电塔拥有相对足够的能量，这样我们就有更多的机会为预定的无人机提供足够的能量。最后，我们可以确认这两个任务是相互关联和协调的，以实现最佳性能。

## 3.2多无人机充电调度:一个凸优化公式

假设一组无人机需要从充电塔充电，如图2所示。如图2所示，充电塔连接到单个协调管理器，该协调管理器是一个充电塔控制器，用于对充电匹配（即，对充电塔和无人机之间的充电调度进行决策）和调度对的能量分配量进行联合优化（即，为预定充电塔向其相关联的无人机无线传输能量的量的决策）。

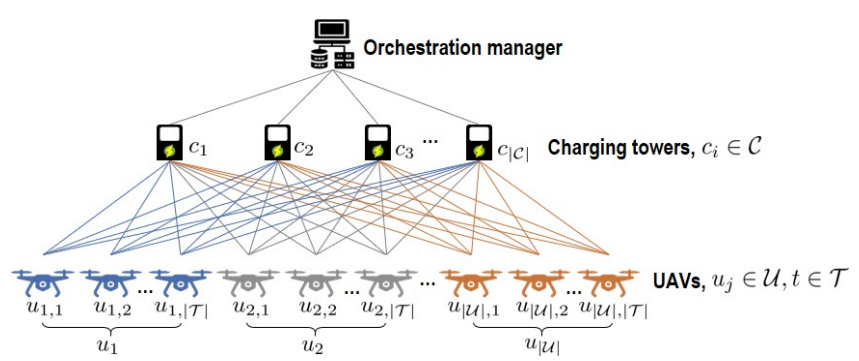


Fig2.多无人机充电调度。

**（1）状态空间(State space)**

假设EMS可以观察到所有给定的充电塔C的过去和现在的信息，而不能观察到它们的未来信息。很明显，它满足马尔可夫性质，MADRL可以用于这个给定的问题设置。状态空间由ESS的能量储备、PV的发电量和负荷组成。ESS能量储备集记为，式中，为i-agent充电塔随时间的ESS能量储备集合，即, 记为。为了安全可靠地使用ESS，考虑了保护比η，即。发电量集合记为，其中是通过i-agent的PV / T产生的能量的集合，记为。通过光伏发电产生的能量受太阳能入射角和充电塔地形条件的影响。因此，这取决于时间和代理的类型。载荷取决于处于充电状态的无人机数量的集合记为 ，式中，为i-agent对T的负载集合，。最后，状态的状态在时刻t记为。 充电塔在t时的状态，用表示，，其中和分别代表ESS充电塔的保障最小和最大容量。注意，充电塔只能观察到过去和当前的信息，而不能观察到未来的信息。此外，其他充电塔的观测值也需要用于MADRL的协同学习和计算。在第i个充电塔的状态信息之上，还包括其他充电塔的状态信息，即 (的PV产生的能量)， (的负荷)， (的ESS中的剩余能量)，其中。因此，第i充电塔的最终状态可表示为

**（2）动作空间(Action space)**

在这里，我们描述了变量和行动空间，通过与所有充电塔(即代理)共享所有ESSs来最大限度地减少公用事业公司的能量。

每个第i个智能体，即，可以选择另一个智能体，然后可以从中获得能量资源，其中能量量为，其中。这被定义为ESS，具体计算过程见算法1。如算法1(3-13行)所示，agent 和进行自我控制的ESS值和为, 。因此，充电塔的作用可表示为，可能的动作数为。为了照顾ESS在动作空间中的荷电状态(SoC)，需要考虑以下两个条件。首先，当一个agent(发送方)的能量状态大于另一个agent(接收方)的能量状态时，可以将自己的能量传递给另一个agent(接收方)。然后，如果agent想要传递自己的能量，那么agent的能量状态应该是emini。供能量q按算法1中的计算过程计算。即使agent 做出决策，如果不能满足上述两个条件，agent也不能做出任何决策。定义第i座充电塔的动作为。最后，动作空间可以表示为。

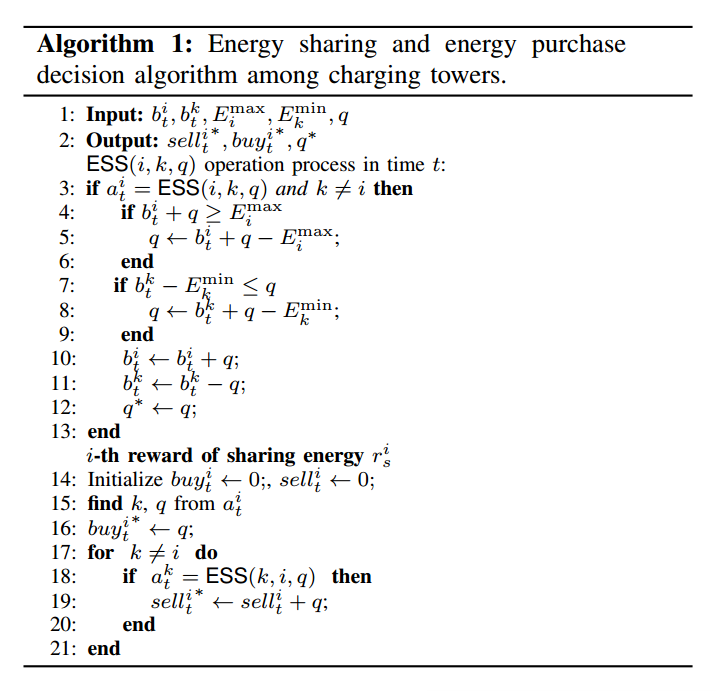


Fig3.算法1。

1. **奖励(Reward)**

我们的业务流程经理操作的EMS的主要目标可以总结如下。

•首先，我们建议的编排经理应该尽量减少从公用事业公司购买的能源数量，因为每个充电塔都有能源发电组件，如用于自充电的光伏发电组件，并且通过在充电塔之间共享ESS(或储能组件)资源，可以进行能源交易/共享。

•其次，每个ESS中的能量状态应介于和之间。因此，如果PV产生的能量大于ESS的剩余容量，则该能量将被丢弃(剩余能量)。

•最后，业务流程管理器应该能够监控所有给定充电塔的状态。在此基础上，管理者可以在考虑剩余能量最小化和从公用事业公司购买能源最小化的情况下定义奖励。对于这种工作负载密集的MADRL学习计算，基本上需要基于云的计算资源。

根据目标，总奖励可以根据以下概念来表示。

•应尽量减少从公用事业公司购买的能源量。因此，可以将其表示为负奖励，记为，其中,表示购买的能量(负奖励)。

•将剩余能量最小化，表示为负奖励，记为，其中

表示剩余能量(负奖励)。

•ESSs之间的能源共享应最大化，以减少购买和浪费的能源。因此，应将能量共享最大化表示为正奖励，记为，其中代表能量共享(正奖励)

## 3.3 基于MADRL的协同EMS学习

系统模型:edge - cloud的概念系统架构，即业务流程管理器与收费塔之间基于MADRL的协作功能关系，如图4所示。我们的编排经理控制充电塔，每个充电塔可以是一个产消者，产生和消耗能源的可再生能源，如光伏太阳能。公用事业公司可以为多个充电塔提供能源。该公司根据充电塔的要求，以自己确定的价格实时交易自己的能源。此外，我们可以假设公用事业价格是由公司以使用时间(ToU)的形式确定的;因此价格是预定不变的。为了实时调度进出公用事业公司、无人机充电和ESS充放电的能量流，每个充电塔共享它们的EMS。充电塔的EMSs旨在通过减少从公用事业公司购买的能源来降低不确定情况下的运营成本。本文提出的系统假设EMS只能管理直连ESS的充放电。此外，假设业务流程管理器管理的充电塔可以共享存储在ESS中的能量。如前所述，存在一组光伏可再生能源系统，可记为， ，其中每个与充电塔相关联;。每个PV系统通过PV板将太阳光转化为能量。光伏发电用于帮助满足每个时隙开始的需求(例如无人机需求和负载需求)。此外，还有一套配备了充电塔的ESS。每个ESS都有自己的最大容量;;它能够根据每个请求存储和释放自己的能量。如上所述，存储在每个ESS中的能量可以与其他ESS共享，并注意共享的能量不能出售。

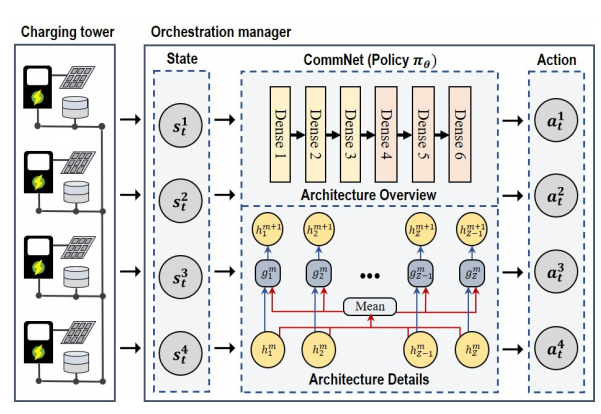


Fig4.edge - cloud的结构。

# 4. 性能评估

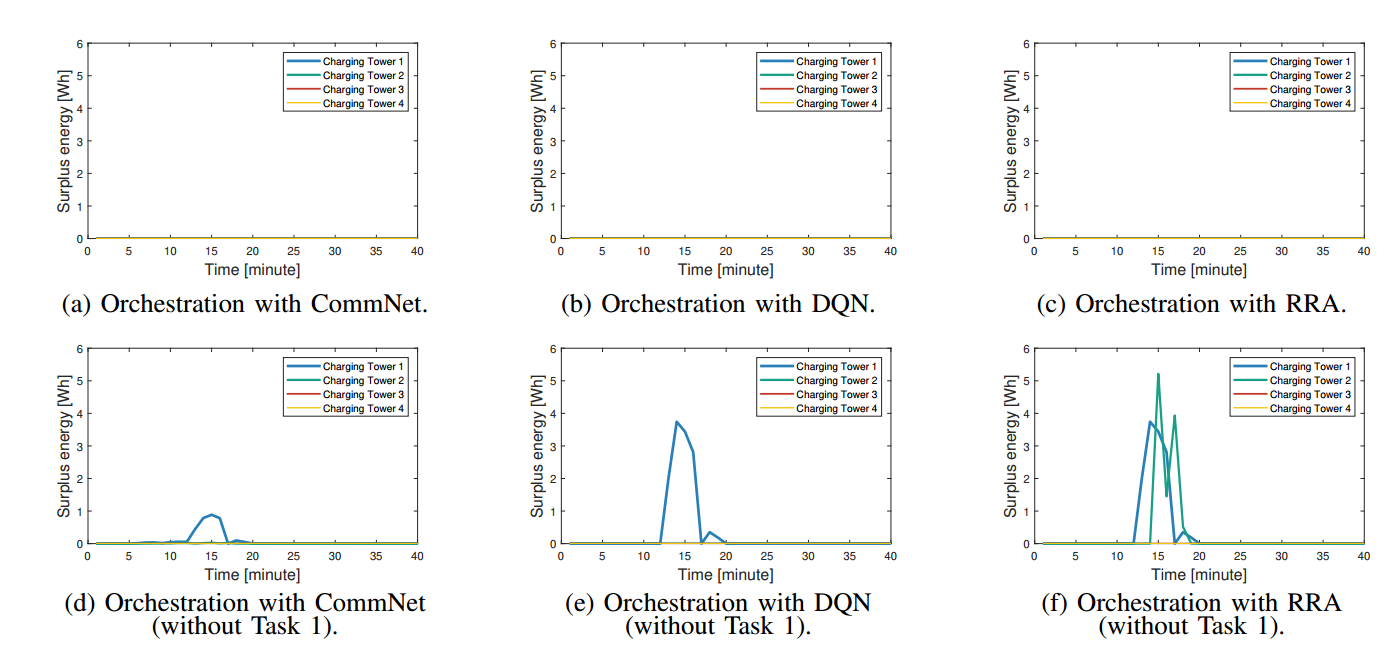
## 4.1性能指标

本节给出了我们提出的两个任务的性能评估结果，即凸最优充电调度(任务1)和基于MADRL/ commnet的能源共享EMS学习(任务2)。

## 4.2评估结果

# 

分别在(a-c)中的任务1和(d-f)中不包含任务1的拟议的编排方法中，从公用事业公司购买能源量。



我们建议的包含(a-c)任务1的编排方法和不包含(d-f)任务1的编排方法中的剩余能量。

## 4.3仿真结果评估

仿真结果验证了所提出的凸最优充电调度在充电塔的能量购买和剩余最小化方面是有效的。

如果充电调度工作不正常，则边缘云(即云和充电塔之间的系统)的能量量会变低，因此无法有效运行。这意味着我们参考系统中的网络组件，即无人机(端)，充电塔(边缘)和编排管理器(云)，是相关的，因此，有效的编排(或协同设计)本质上是必需的，正如我们可以进一步证实，在我们提出的“与CommNet的编排”中，购买的能量和剩余的能量被最小化，共享的能量被最大化，因为购买的能量和剩余的能量被设置为负奖励，而共享的能量被设置为正奖励。

# 总结

本文提出了一种云辅助的无人机网络联合充电调度与能量管理框架。为了给那些极度耗电的无人机充电，充电塔被考虑在运行时进行即插即用充电。充电塔应该具有成本效益，因此它配备了光伏发电和储能系统功能。此外，塔楼应该通过智能能源共享来提高成本效益。根据需求和设置，提出了1)无人机与塔间的充电调度和2)塔间的协同能量管理。在充电调度中，对无人机和塔进行充电能量最大化的调度，由调度对确定充电能量分配量。这里，两个决策是相关的，即，它是一个非凸问题。我们重新构造了保证最优解的非凸到凸问题。最后，利用多智能体深度强化学习，设计并实现了塔间的协同能源共享，实现了塔间的智能能源共享。我们可以看到，这两种方法是相互关联的，应该由一个集中的编排管理者在考虑公平性、能源效率和成本效益的情况下进行管理、协调和协调。