各位老师上午好，我开题报告的题目是基于强化学习的无人集群抗干扰研究

内容分为以上六个方面，其中重点是该研究方向的发展综述和研究内容。

首先是该课题研究的背景和意义：

无人机自组网可以借助多跳通信，搭载不同任务载荷来提供可靠的多样化实时服务，广泛应用于各种军事以及民用领域场景，但同时无线网络的开放特性使它容易受到各种物理层的干扰攻击，从而对无人机网络通信安全造成严重威胁。

按照干扰信号和干扰产生方式将目前常见的对物理层产生干扰的方式进行如上的分类：干扰信号分为压制式干扰和欺骗式干扰两种，同时按产生干扰的方式，干扰又分为持续干扰、随机干扰、扫频干扰、反应式干扰和智能干扰，其中智能干扰能够自适应地调整干扰产生的时段和频段，抗干扰难度更大。

现有的无人集群抗干扰技术主要从节点移动、网络层、MAC层和物理层进行抗干扰实现：

节点层面和网络层是从空间角度避开干扰，MAC层的跳频扩频技术是研究最主要的抗干扰研究手段。

由于现有场景下存在对智能干扰应对能力不足的主要问题，本课题研究的意义在于利用强化学习联合多种抗干扰手段实现智能干扰场景下的抗干扰能力的提升。

接下来将分别对抗干扰技术、强化学习技术、基于强化学习的抗干扰技术的发展综述进行介绍；

首先是抗干扰技术的综述：

传统的MAC层抗干扰技术应用于无人机网络，针对干扰机产生干扰的信道切换，能够切换不同的信道集进行干扰规避，但是由于信道的有限，无法应对多信道智能干扰

第二类抗干扰方式利用了无人机网络特有的移动特性，将路径规划建模为一个优化问题，从而实现最小干扰路径和发射功率的求解，但是其需要对信道、任务和干扰进行建模，复杂度高，适应性差。

第三类是基于监督式机器学习的抗干扰技术，用于优化传统抗干扰的问题，缺点是需要大量先验知识和学习数据。

然后对强化学习的发展进行介绍

首先是强化学习原理的介绍：

强化学习算法利用基于马尔可过程的前提和贝尔曼方程，实现对每一个状态的价值或者对应的Q值进行预测，从而输出每个状态下的最优决策问题。

基于价值的强化学习的训练过程就是通过输入状态，动作，回报和下一状态的四元组，更新价值估计，对应的数学形式就是上述贝尔曼方程。

强化学习现有的发展历程如上：基于价值的方法从Q-learning到DQN，再到DQN的各种变体，基于策略的方法主要是AC架构，即演员评论家架构，变体为DDPG等

上述不同的强化学习算法主要的应用区别就是: dqn引入神经网络，解决了输入状态空间连续的问题，基于策略的方法相比与基于价值的方法解决了输出的决策空间无限的问题。

于是根据状态空间和决策空间维数的区别，需要使用不同的网络实现‘

然后是利用强化学习进行抗干扰的研究; 这几篇文献分别利用强化学习算法实现如信道接入，轨迹规划，中继转发等不同抗干扰方式的优化，

其利用强化学习技解决抗干扰通信过程中因无人机移动性和视距信道造成的信道状态波动大，以及干扰模型未知等问题。

但现有研究存在以下两个问题，

一是仅仅基于无人机之间的通信抗干扰的场景研究较少

二是仍旧针对单一的抗干扰方式进行决策或者参数优化，在智能干扰，强对抗环境下下抗干扰能力有限。

然后是研究内容，本课题的研究内容主要分为两点：  
第一点是基于深度强化学习的集中式联合抗干扰技术：

该研究方向利用全局信息：比如任务信息，网络通信指标信息，节点物理信息包括所有无人机节点收集的动态空间谱信息、移动信息和功率信，输入到设计好的DQN网络中，从而充分提取当前全局电磁场景信息和任务信息下的特征，输出网络的决策信息包括：跳频的子信道切换，轨迹规划的下一时帧位置，功率以及是否参与中继发等组合决策，用于指导训练的回报函数是节点的链路质量指标，实现优化综合运用多种抗干扰技术决策的智能组网

中间的网络模型目前打算构建基于混合注意力机制的卷积神经网络，提高抗干扰能力模型的预测精度与模型训练的收敛速度。

该研究方向采用集中式训练决策，所有决策都由基站或选中的中心节点输出，并且输出的决策是多种抗干扰技术的组合。

第二个研究内容是基于多智能体强化学习的分布式联合抗干扰技术，

这个研究方向主要是为了解决研究方向一中集中式抗干扰在实际中难以部署的问题，通过引入多智能体强化学习思想，建模各节点行为，将输入信息限制为本地节点所能获取到的信息，输出决策限制为本地节点的决策，

每架无人机都能够利用本地信息独立学习而不受无人机群规模大小影响，避免额外通信开销，适应高动态性。

然后是研究方案设计和可行性分析

本研究采用计算机仿真的方式研究，利用exata网络仿真软件和强化学习库LIBtorch完成相应模块的设计和开发

通过场景建模、实现对应的多种抗干扰算法，设计对应的网络模型，训练集中式和多智能体模型，验证对应算法效果。

由于单一抗干扰研究方案和深度强化学习方案成熟，该方案的可行性较高，模型训练效果还需要进一步验证。

然后是研究挑战和创新点

本研究的创新点和研究内容一致：

1. 首先是基于基于全局信息的联合抗干扰模型：与以往单一抗干扰手段不同，联合多种抗干扰手段，优化输出决策
2. 然后是基于多智能体的分布式联合抗干扰模型：具有分布式和可扩展性，不依赖全局信息，解决了全局信息交互的额外开销问题。

本研究的面临的主要挑战是模型的收敛性问题，模型参数和超参数的调整将对模型的训练效果产生大的影响，由于输出多种抗干扰决策，较高的决策动作空间也将增大训练难度。

最后是研究预期成果和计划

利用深度强化学习训练模型，模型训练阶段输入的数据为各节点接收机收到的信息，包括到达率、功率以及链路质量等信息，通过设计回报函数指导对应的抗干扰参数优化以及对应的抗干扰方式选择，能够处理在压制式智能干扰场景下实现针对单一传统抗干扰手段抗干扰能力的提升。

初步研究计划安排如上。

最后是研究预期成果和计划

利用深度强化学习训练模型，模型训练阶段输入的数据为各节点接收机收到的信息，包括到达率、功率以及链路质量等信息，通过设计回报函数指导对应的抗干扰参数优化以及对应的抗干扰方式选择，能够处理在压制式智能干扰场景下实现针对单一传统抗干扰手段抗干扰能力的提升。

初步研究计划安排如上。