对于移动机器人来说，在高度动态和拥挤的场景下实现无碰撞导航仍然是一项艰巨的任务。防撞模块在传统的导航框架通常将运动障碍视为静态的，例如动态窗口方法等(DWA)[5],或者只关注下一步的行动按照一定的交互规则,如交互速度障碍(RVO)[6]和最佳互惠防撞(ORCA)[7]。由于这些方法通过被动反应防止碰撞，并且经常使用一些手动定义的功能来保证安全，它们被发现导致机器人的运动是不自然的、短视的和不安全的。

为了在人群中以社会顺从的方式进行导航，机器人需要有感知、理解和预测周围行人行为的能力。人的运动预测与机器人的运动规划相结合一直是机器人感知导航任务中的一个难题。一个想法是在预测之后进行计划，即在预测其他人未来的轨迹后确定一条安全的路径。一些手工模型(如恒速度模型[8]、离散选择模型[9][10]、社会力量模型及其变体[11][12])和数据驱动方法(如社会LSTM[13]、社会GAN[14])已经被提出来编码行人的意图，以预测他们未来的轨迹。然而，人群行为的高随机性往往会影响行人轨迹预测的计算成本和可靠性。在移动机器人的人工感知导航既需要安全性又需要时间效率的背景下，这种预测后规划方法在实际应用中仍然具有挑战性。

人类感知导航的另一种方法是基于深度强化学习(Deep Reinforcement Learning, DRL)，它将人类运动的预测整合到决策过程中，即机器人从经验中学习以理解拥挤场景，并将人群机器人交互编码到导航策略中。最近在这个领域的工作[15][17][18]已经证明了在模拟中产生群体感知导航策略的优越能力，其中社会注意强化学习(SARL)算法实现了最先进的性能。然而，在现实的实现中存在一些局限性:1)学习的导航策略被限制在与训练过程相关的一定距离内，导致机器人近视，在导航目标较远的情况下会绕行;b)将环境的表示简化为只包含机器人和人群，这将导致机器人无法避开环境中除人类之外的障碍物。这些缺点将限制机器人在现实世界中的导航性能。

基于[18]中指定的SARL模型，我们提出了一种新的算法，即SARL∗，来解决上述问题。我们工作的主要贡献有三个方面:

SARL现实应用问题：

1. 学习到的导航策略被限制在与训练过程相关的一定距离内，导致机器人近视，当导航目标较远时会绕道而行
2. 将环境的表示简化为只包含机器人和人群，这将导致机器人无法避开环境中除人类以外的障碍物。这些缺点将限制机器人在现实世界中的导航性能。

想法：对于移动障碍物的检测避障

将运动的预测整合到决策模块中，

人的避障策略为 ORCA