

未来预测的一个关键挑战是高度的不确定性，这在很大程度上是由于不知道其他代理的意图和潜在特征。例如，车辆通常具有多模态的未来分布：它可以转弯、直行、减速、加速等。根据其他场景元素，它可以通过、让行、改变车道或驶入车道。在过去的几年里，这一挑战引起了很多人的兴趣。对高度多模态进行建模的一种方法是采用可以从中抽取样本的灵活隐式分布——条件变分自动编码器(CVAEs)[1]、生成对抗网络(GANs)[2]、和单步策略推出方法[3]。尽管它们的表现具有竞争力，但使用隐变量来建模意图会阻止它们被解释，并且通常需要测试时间抽样来评估概率查询（例如，“代理向左转的可能性有多大？”）。此外，在大型机器学习社区[4]中，特别是自动驾驶汽车中，已经有相当大的努力来解决这类模型中的模式崩溃问题[5,6]。

为了解决这些限制，我们观察到，对于我们的任务（例如车辆和行人轨迹预测），在较长的未来中的不确定性主要可以通过对代理可能目标的预测来捕获。这些目标不仅基于可解释的物理实体（例如位置），而且与意图密切相关（例如变道或右转）。我们推测目标空间可以在场景中被离散化——允许确定性模型并行生成不同的目标——并且后来改进得更精确了。

这些观察导致我们提出了目标驱动轨迹预测框架，名为 TNT。我们首先将未来预测问题转化为预测在离散目标上的分布状态，然后制定一个概率模型，其中轨迹估计和可能性是基于这样的目标。生成的框架有三个端到端训练的阶段：（1）目标预测 估计给定场景背景下的候选目标上的分布；（2）以目标为条件的移动估计 预测每个目标的轨迹状态序列；（3）评分和选择 估计每个预测轨迹的可能性，并考虑所有其他预测轨迹的背景。我们通过排序可能性和抑制冗余的轨迹最终得到了一系列紧凑的多样预测。我们三个阶段模型的说明应用于车辆轨迹预测如图1所示。尽管我们的模型是端到端训练的，但是它的三个阶段的布局，在每个阶段都有可解释的输出，紧密遵循了传统机器人运动预测和规划系统的典型处理步骤 [7, 8]，因此，很容易在部署期间合并领域知识。