**一、DeepWalk**

DeepWalk是一种基于随机游走的网络表示学习算法，其目的是将网络中的节点表示为低维向量，这些向量可以用于后续的机器学习任务。DeepWalk的数学原理主要包括两个部分：随机游走和Skip-Gram模型。

**1. 随机游走**

随机游走是一种在图中随机选择路径的方法，其目的是生成图中节点的局部邻域信息。对于一个给定的图G(V, E)，其中V表示节点集合，E表示边集合，随机游走从一个初始节点v开始，沿着图的边随机地选择相邻节点，并按照一定的概率转移到下一个节点。这个过程重复进行若干步，直到达到预定的游走长度T。

在DeepWalk中，每个节点都可以作为初始节点，并且可以执行多次随机游走。通过这种方式，我们可以生成大量的节点序列，这些序列可以捕捉到图中的局部结构信息。

**2. Skip-Gram模型**

Skip-Gram模型是一种基于词嵌入的方法，用于将离散的词汇表示为连续的向量。在DeepWalk中，将随机游走生成的节点序列视为“句子”，将节点视为“词汇”。Skip-Gram模型的目标是学习节点的向量表示，使得给定一个节点，能够最大化预测其邻居节点的条件概率。具体来说，给定一个节点序列，Skip-Gram模型的目标函数可以表示为：

其中，是序列中的节点，表示的邻居节点集合，表示节点的向量表示。条件概率可以通过softmax函数计算：

其中，是节点的一个邻居节点，是图中的任意节点，和分别表示节点和的向量表示。通过最大化目标函数，我们可以学习到每个节点的向量表示。

总结一下，DeepWalk的数学原理主要包括两个步骤：首先使用随机游走生成图中节点的局部邻域信息，然后利用Skip-Gram模型学习节点的低维向量表示。这些向量表示可以捕捉到图中的结构信息，并可以用于后续的机器学习任务。

**二、在PGRP的基础上进行改进**

首先，

**1、PGRP的State.** 第t步的状态用三元组定义，这里是起始user，是第 t 步 agent到达的entity，是第 t 步之前的历史，是过去 k 步 entities 和 relations 的集合。初始态，给一个确定的T，终态为。注意，这里的都是transE训练的向量。

**改进后的State.** 除了transE训练的，还有deepwalk训练的。即

然后，

**2、PGRP的Action**.：对于状态的全状态空间定义成相连的所有edges和entities，但是不包括里已经有的relations 和entities。正式来说。因为出度的分布有明显的长尾效应，某些结点的出度很大，因此存他们的全状态空间浪费资源。因此，在给定起始用户 u 的条件下通过一个scoring function计算每一边的score，有效保留重要 edges 。这个 scoring function在给定 user 的情况下把边映射到一个score实数。最后根据 score 排序，选取top 个action：

是预定义参数，用以控制action数量的上界，函数细节在下节讨论。

**改进后的Action：改进了其中的**scoring function *。*在下面reward部分详细说明。

**3、PGRP的Reward：**Reward即由上面提到的scoring function计算得到。这个模型中的agent 被鼓励去探索尽可能多的 "good paths"。直觉上看一个 "good path" 当然要把 user 引导到他可能会接触的 item 上。因此我们只针对终态（T时刻），计算 scoring function 定义为：

这里的**即是用transE模型产生的u向量和向量内积。**

**改进后的Reward：**

**4、PGRP的网络：**利用了一个策略网络 (policy network) 和价值网络 (value network) 它们共享特征层。策略网络用状态向量和二值化向量作为输入，计算出每个action的概率，对于不在的action 这个概率取0。价值网络将状态向量映射到一个实数，这两种网络的结构定义：

其中s是，即TransE模型的向量的拼接向量。

**改进后的网络**：除了s中TransE模型的向量，还利用s中DeepWalk的向量，即：

简言之，这篇论文只用transE算法学习了用户和商品之间的关系信息，我们额外使用了deepwalk算法学习了用户和商品本身的路径信息（语义信息，它周边的节点信息），能更精准推荐。

**三、实验结果**

**Amaon\_Beauty**

**NDCG=6.135 | Recall=7.355 | HR=15.476 | Precision=1.738 | Invalid users=0**

**Amazon\_Cellphones**

**NDCG=5.626 | Recall=6.433 | HR=12.702 | Precision=1.348 | Invalid users=1**

**比PGRP在指标上好一些。**

