knowledge graph completion with adaptive sparse transfer matrix

1. 本文的贡献/动机：

问题：由于之前的翻译模型并没有考虑异构性和不平衡性，导致了复杂关系欠拟合、简单关系过拟合等。

1）提出了新的方法，考虑到了异构性（各个关系连接的总的实体数量不同，即关系的复杂度不同）和不平衡性（一个关系连接的头实体和尾实体的数量不同）。

2）本文的方法是很高效并且有很少的参数，更易于知识图谱的扩展。

3）提出了两种转移矩阵的稀疏模式（非零元素集中在对角线和非零元素均匀分布在矩阵当中）并比较了优缺点。

4）在三元组分类和连接预测的任务上，都达到了最好的性能。

1. 相关工作

TransE：只适合一对一的关系。

TransH：和TransE相比，它将关系映射到另一个空间，解决了TransE对一对多，多对一，多对多关系建模的难题。

TransR/CTransR：也用到了转移矩阵，然而由于每一个关系都有相同的参数，易造成简单关系过拟合，复杂关系欠拟合。

TransD：考虑到了实体和关系的多种类型，取代了转移矩阵，使用了两个投影向量。

KG2E：使用Gaussian Distribution 来表示实体和关系，提出了用Gaussian Distribution的协方差来表示实体和关系的不确定度的新思想。

1. 本文的工作

针对不同难度的实体间关系，使用不同稀疏程度的矩阵（不同数量的参数）来进行表征，从而防止对复杂关系欠拟合或者对简单关系过拟合。

本文针对两类问题，提出了两种稀疏矩阵的初始化方法：TranSparse(share)、TranSparse(separate)。

1）TranSparse(share)：模型的参数（稀疏度）由如下公式确定：

IMG_256

其中，Nr是关系r所连接的三元组的个数，Nr\*是数据集中最大的关系三元组的个数。

有了这个转移矩阵就可以获得头实体和尾实体的投影向量：

IMG_256

1. TranSparse(separate)：在头实体和尾实体使用不同稀疏度的转移矩阵

IMG_256

和上面的方法类似，只不过将Nr和Nr\*替换成了实体的个数。

有了这个转移矩阵同样可以获得头实体和尾实体的投影向量：

IMG_256

上述的两种方法使用相同的得分函数：

IMG_256

这个打分希望在目标三元组更低，在错误的三元组更高。

模型的训练：与之前的方法相同，为了训练，我们人为的增加一些错误的三元组。

损失函数为：

IMG_256

1. 总结

这些方法的核心思想就是如何找到合适的方法来将知识图谱emmbedding到向量空间，从而在向量空间中进行计算。

尽管都获得了不错的性能，但知识表示学习仍然面临着挑战，主要包括以下几个方面：

1）如何进行更加复杂的知识推理。

2）对于知识图谱无法表达的信息，应该进行如何表示和推理。