# 实验报告

201870258 邢天骋

## 1 Abstract

知识图谱是大型事实数据库的图形表示，它通常存在不完整性。推断实体（节点）之间缺失的关系（链接）是链接预测的任务。最近一种最先进的链接预测方法ConvE，实现了一个卷积神经网络，从连接的主题和关系向量中提取特征。虽然结果令人印象深刻，但该方法是不直观的，且很难理解。而我们提出了一个超网络架构，该架构生成简化的特定关系的卷积滤波，(ⅰ)在标准数据集上优于ConvE和所有以前的方法；(ⅱ)可以作为张量因子化的框架，从而设置在一个成熟的用于链接预测的因子化模型系列中。由此我们可以证明，卷积只是提供了一种引入稀疏性和参数捆绑的方便计算手段，可以在非线性表现力和需要学习的参数数量之间找到一个有效的平衡点。

## 2 Introduction

知识图谱是大型的图结构的事实数据库，例如WordNet、Freebase和Google等。它们包含的信息以三要素（e1, r, e2）的形式存在，其中e1和e2代表主体和客体实体，r是它们之间的关系。它们被认为是重要的信息资源，可用于从问题回答到信息检索和文本总结的各种任务。现有的知识图谱面临的主要挑战之一是它们的不完整性：图谱中的许多实体之间的联系是缺失的。这激发了链接预测领域的大量工作，比如推断知识图谱中缺失链接的任务。

直到最近，许多链接预测的方法都是基于对训练三元组的3个编码的二进制张量表示的不同因子[12,17,23,22]。这种方法是浅层的和线性的，表现力较为有限。然而，试图用额外的全连接层和非线性来增加表现力往往会导致过度拟合[12,17]。出于这个原因，Dettmers等人引入了ConvE，这是一个使用二维卷积在重塑和串联的实体和关系嵌入上的模型[3]。他们鼓励使用卷积，因为它的参数效率高且在GPU上计算速度快，而且有各种计算机视觉的稳健方法来防止过度拟合。尽管ConvE取得的结果令人印象深刻，但很不直观的是：卷积——尤其是二维卷积——应当能有效地从一维实体和嵌入的关系中提取信息。

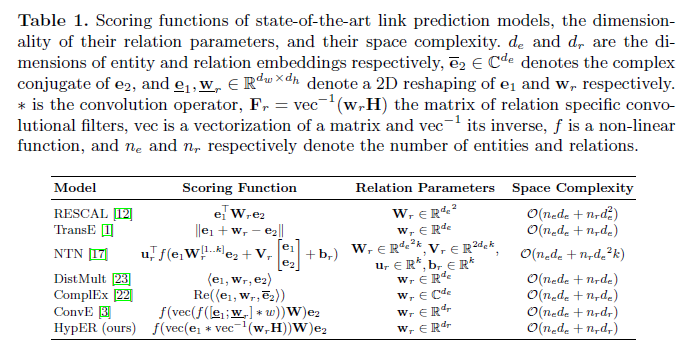
在本文中，我们引入了HypER，一个使用超网络[5]为每个关系生成卷积过滤器权重的模型。超网络是一个网络为另一个网络生成权重的一种方法，可以用来实现跨层的权重共享，并动态地合成输入的权重。在我们的背景下，我们生成特定关系的过滤器权重来处理输入实体，同时也实现了知识图谱中关系的多任务知识共享。我们提出的HypER模型使用超网络来生成一组一维的特定关系过滤器来处理主要实体嵌入。与ConvE相比，这简化了主体实体和关系嵌入之间的互动。在ConvE中，一组全局的二维过滤器在重构和连接过的主体实体和关系嵌入上进行卷积，而这是不直观的，因为它表明词语嵌入中存在二维结构。此外，ConvE中主语和关系之间的交互作用的主语和关系之间的互动取决于对它们如何被重塑和连接的任意选择。相比之下，HypER的超网络产生了特定于关系的过滤器，因此从主语实体嵌入中提取了特定于关系的特征。这不需要二维重塑，并允许实体和关系更完整地互动，而不是只在连接边界附近。我们表明这种简化的方法，除了提高链接预测性能外，还可以从张量因子化的角度来理解，从而将HypER置于一个成熟的因子化模型系列中。因此，在词嵌入中使用卷积的明显晦涩性被解释为：不过是一种引入稀疏性和参数捆绑的方便计算手段。

我们使用标准数据集（FB15k-237, WN18RR, FB15k, WN18, YAGO3-10）对HypER与之前提出的几个链接预测模型进行了评估，在这些数据集中，HypER均表现出了最先进的性能。总而言之，我们的主要贡献是：  
- 提出了一个新的链接预测模型（HypER），该模型在所有标准数据中都达到了最先进的性能。在所有的标准数据集上取得了最先进的性能。  
- 显示了使用卷积层而不是全连接层的好处。层的好处是由于限制了相互作用的维数（即明确的正则化），而不是在嵌入中寻找高维结构（如ConvE所暗示的）；以及  
- 表明尽管使用了卷积，但HypER实际上属于张量分解模型的一个大类，这有助于在表现力和学习参数的数量之间进行良好的权衡。

## 3 问题描述

链接预测：

目的是要训练一个得分函数φ，使得对于每个输入三元组 有 ，其中 为主体和客体的个体，为关系。分数表示对给定的三元组对应于真实事实的预测强度，正数表示真实，负数表示虚假。链接预测模型通常将实体对e1、e2映射到其相应的分布式嵌入表征 中，并使用特定的关系函数分配一个分数。大多数链接预测模型对分数应用了Logistic sigmoid函数，对被查询的事实是否为真给出了一个可解释的概率预测。表1总结了来自整个文献和HypER的模型的评分函数，以及其关系参数的维度和空间复杂性的重要项。



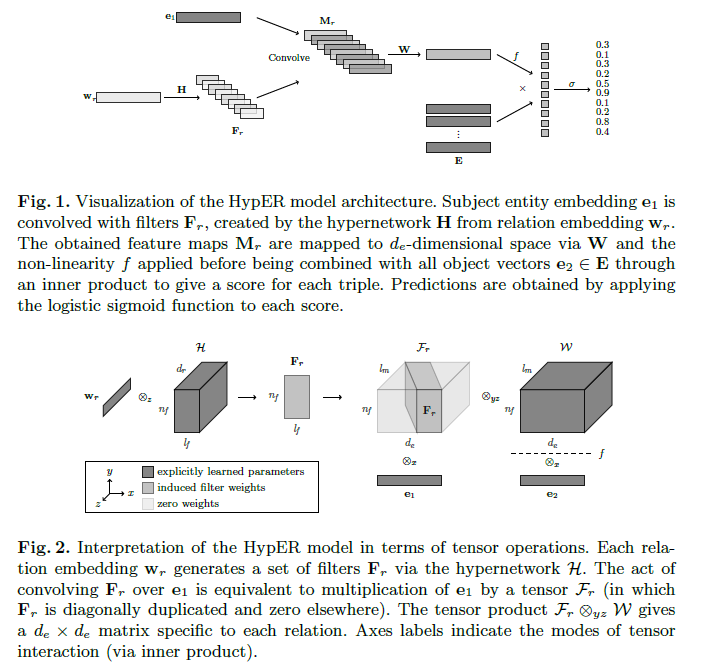
## 4 输入、输出、模型算法描述

文章中，作者提出了一个新颖的超网络模型，用于知识图谱中的链接预测。该网络通过全连接层投射每个关系的向量嵌入，重塑其结果，来给每个关系提供一组卷积滤波器权重向量，以下为详细过程。

## 4.1 Score Function与模型结构

HypER模型中关系特定的Score Function如下：

其中 将向量重塑为矩阵，以及非线性的 被选为校正后的线性单元(ReLU)



框架图

在Feedforward过程中，模型从实体中获取输入三元组的嵌入以及关系嵌入矩阵。超网络是应用在嵌入关系上的完全连接层。其结果被重塑以生成一个卷积滤波的矩阵 。虽然过滤集总体维数为 ，秩被限制为 ，来使得参数在关系之间可以共享。

## 4.2 张量分解

在了解了 HypER 的结构之后，我们能够通过把超网络H和权重矩阵W认为是张量，从而将其视为一系列的张量操作，其中 。根据主要实体嵌入的卷积行为，等同于与分散张量相乘的结果。线性性质允许将乘积 单独视为为每个关系生成的一个矩阵。更进一步，与其复制的入口给，我们能够将概括为无关关系的四维分散张量 ，通过将入口替换为维度的。因此 HypER 模型可以被清晰表示为使用核心张量的关于的乘法。这个视角使得 HypER 可以用类似链接预测中因子分解方法家族的方式来看待。

## 5 评价指标及计算公式

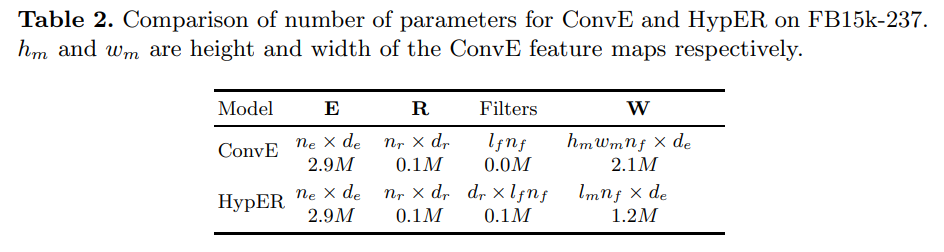
## 5.1 训练过程

根据上述训练的过程，使用根据Adam优化的 来最小化二进制交叉熵损失：

是一个标签向量，向量中1代表真实三元组，反之则为0，其受制于Label Smoothing。Label Smoothing是一种增强概括性的普遍使用的技术。Label Smoothing 通过增加一个统一的优先级来改变真实有效标签值的分布，从而使得模型减少置信度，达成正则化的效果。是指同时评分，与相比，提供了显著的加速作用：训练速度为3倍，测试则为300倍。对上述HypER模型的一个潜在扩展是将卷积滤波同时应用在主体和客体的实体嵌入上。

## 6 对比方法及引用出处

文章对比了ConvE 和 HypER 使用的参数数量。可以发现，总的来说HypER使用了更少的参数，由于其直接将关系转变为了卷积滤波。



## 7 结果

文章的主要发现是：

* 尽管比最近的ConvE参数要少，HypER在所有数据集上均优于其他所有模型，在链接预测任务上实现了最先进的效果
* 滤波维度研究表明，与一位实体嵌入向量相比，对重塑后的二维实体嵌入进行卷积没有任何好处，大多数信息可以用非常小的卷积滤波器进行提取

总而言之，除了WN18上的平均倒数秩和WN18RR、FB15k-237、WN18和YAGO3-10上的平均秩以外，HypER在所有指标上均优于其他所有其他的模型。考虑到平均排名对异常值高度敏感，这表明HypER正确地对许多真实三元组进行了秩计算，并能名列前十，但在其他地方产生了更大的错误。