《中文信息学报》稿件排版样式

文章编号：1003-0077（2008）06-0000-00

基于分布式表示和多特征融合的知识库三元组分类*

安波[1]，韩先培[1]，孙乐[1]，吴健[1]

（1. 中国科学院软件研究所 中文信息处理研究室，北京 100190）

摘要：三元组分类是知识库补全及关系抽取的重要技术。当前的 state-of-the-art 三元组分类方法通常基于 TransE 来构建知识库实体和关系的分布式表示。然而， TransE 方法仅仅适用于处理 1 对 1 类型的关系，无法很好的处理 1 对多、多对 1 及多对多类型的关系。针对上述问题，我们在分布式表示的基础上，提出了一种特征融合的方法—TCSF，通过综合利用三元组的距离、关系的先验概率及实体与关系上下文的拟合度来进行分类。在四种公开的数据集（WN11、WN18、FB13、FB15K）上的测试结果显示，TCSF 在三元组分类上的效果超过现有的 state-of-the-art 模型。

关键词：知识库；深度学习；三元组分类

中图分类号：TP391　　**文献标识码：**A

Triple Classification via Synthesized Features based on Knowledge

Bo An[1], Xianpei Han[1] ,Le Sun[1],Jian Wu[1]

(1. Laboratory of Chinese Information Processing, Institute of Software, Chinese Academy of Sciences, Beijing, 100190, China)

Abstract: Triple classification is crucial for knowledge base completion and relation extraction. However, the State-of-the-art methods for triple classification fail to tackle 1-to-n, m-to-1 and m-to-n relations. In this paper, we propose TCSF (Triple Classification based Synthesized Features) method, which can joint exploit the triple distance, the prior probability of relation, and the context compatibility between entity pair and relation for triple classification. Experimental results on four datasets (WN11, WN18, FB13, FB15k) show that TCSF can achieve significant improvement over TransE and other state-of-art triple classification approaches.

Key words: Knowledge Base； Deep Learning； Triple Classification

1 引言

知识库是人工智能和自然语言处理的关键资源，在问答系统、智能检索及情感分析等领域起到重要的作用。目前已经有许多知名的大规模知识库，如 FreeBase、WordNet 等。然而，考虑到知识的规模和更新速度，现有知识库的覆盖度仍然无法达到要求。因此，知识库补全（Knowledge base completion）在近几年成为了热门国际评测任务。三元组分类是知识库补全的关键技术之一，由 Socher[1]等人首先提出。给定三元组（head,relation,tail），head 和 tail 表示头部和尾部实体，relation 表示实体之间的关系。三元组分类的目的是计算三元组的置信度，从而判断该三元组所表示的语义关系是否为真，如果该三元组为真，则该三元组（head,relation,tail）将被加入到知识库。

目前大部分三元组分类系统首先将知识库中的实体和关系编码到低维向量空间，并基于这些连续向量构建置信度模型对三元组进行打分。具体地，对于知识库中的任意实体，使用

* **收稿日期：2016 年 6 月 1 日　　定稿日期：2016 年 8 月 6 日**

基金项目：国家自然科学基金资助项目（61540057，61433015，61272324，61572477）；青海省自然科学基金资助项目（2016-ZJ-Y04，2016-ZJ-740）；

作者简介：安波（1986——），男，博士研究生，自然语言处理和知识表示；韩先培（1984——），男，副研究员，信息抽取、知识库构建以及自然语言处理；孙乐（1971——），男，研究院，信息检索和自然语言处理。

一个实数向量进行表示，知识库中的任意关系也一个实数向量（维度可以与实体向量的不同）来表示。对知识库的分布式表示任务，TransE[2]是目前效果较好的模型，且时间复杂度较低。但是，就像 TransM[3]中指出的那样，TransE 并不适于处理 1 对多（1-to-N）、多对 1（M-to-1）及多对多（M-to-N）类型的关系。与此同时知识库中的大多数关系都不是 1 对 1（1-to-1）类型的。本文针对知识库补全任务中常用的四个数据集中的关系进行统计分析，分别是 WN11[1]、WN18[2]、FB13[1]和 FB15K[2]，其中 WN11 和 WN18 抽取自 WordNet[1]数据集，FB13 和 FB15K 抽取自 Freebase[2]数据集。其结果如表 1 所示，从表 1 中可以看出，WN11 数据集中的所有 11 种关系均为多对多（M-to-N）类型的关系；WN18 数据集中的所有 18 中关系也全都属于多对多（M-to-N）类型的关系；FB13 数据集中包含 1 中多对 1（M-to-1）的关系和 12 中多对多（M-to-N）的关系；FB15K 数据集中包含 1345 种关系，其中只有 247 种关系为 1 对 1（1-to-1）类型的关系，其余的均为 1 对多、多对 1 及多对多类型的关系。因此，在上述的四种数据集中，绝大多数关系属于非简单的 1 对 1 关系。因此，我们认为，TransE 模型由于不能很好的编码知识库中的 1 对多（1-to-N）、多对 1（M-to-1）、多对多（M-to-N）类型的关系，仅仅依赖于三元组距离|h+r-t|来计算三元组的置信度的现有模型，无法很好的建模多种复杂的关系。

表 1 知识库关系分类信息

数据集	WN11	WN18	FB13	FB15K
关系总数	11	18	13	1345
1-to-1 关系数目	0	0	0	247
1-to-N 关系数目	0	0	0	179
M-to-1 关系数目	0	0	1	225
M-to-N 关系数目	11	18	12	694

针对上述问题，本论文在知识库分布式表示的基础上，提出了一种多特征融合的方法—TCSF，综合利用三元组的距离、关系的先验概率及实体与关系上下文的拟合度来进行分类。具体地，TCSF 通过两个启发式规则来解决上述问题：1）先验规则：一个关系在知识库中出现的次数越多，则给定的实体对（h,t）具有该关系的可能性就越高；2）实体对（h,t）与给定关系 r 上下文拟合程度越高，则该三元组（h,r,t）的置信度越高。本文提出的方法是在知识的分布式表示基础上利用关系的先验信息和关系的上下文特征针对三元组分类任务进行的优化，该方法不针对具体的模型，可以对 TransE、TransH 等基于翻译的关系表示模型进行结果优化。

本文在上述四个数据集上进行了实验。实验结果表明，相比于 TransE[2]及其他 state-of-the-art 模型，本文提出的方法能够大幅提升三元组分类的效果。

2 相关工作

目前大部分三元组分类方法包含两个模块：1）如何将三元组（h,r,t）中的关系和实体映射为低维向量或矩阵表示；2）如何构建损失函数 $f_r(h,t)$ 对三元组的置信度进行打分。以下分别从上述两个模块介绍现有的不同方法。

Unstructured Model (Bordes et al. 2012; 2014)[4] 是一个较为朴素的模型，该模型基于头部实体 h 和尾部实体 t 共现的信息来对知识库中的实体进行编码，但不对关系进行编码，其损失函数如公式（1）所示。

$$f_r(h,t) = \| h - t \|_2^2 \quad (1)$$

Distance Model (Bordes et al. 2011)[5] 与 UM 模型类似，但是增加了对关系的编码。该模型将知识库中的每个关系编码为两个矩阵 (W_{rh}, W_{rt})，其损失函数为公式（2）所示：

[1] http://www.princeton.edu/wordnet

[2] http://www.freebase.com

$$f_r(h,t) = \| W_{rh}h - W_{rt}t \| \quad (2)$$

Single Layer Model (Socher et al. 2013)[1] 通过单层的神经网络来对 DM 模型进行改进，该模型的损失函数如公式（3）所示，其中 g 为 tanh 函数。

$$f_r(h,t) = u_r^T g(\mathrm{M}_{r,1}\mathrm{h} + \mathrm{M}_{r,2}\mathrm{t}) \quad (3)$$

Bilinear Model (Sutskever et al. 2009,2012)[6] 通过增强三元组中实体对之间的交互关系来改善 DM 模型，该模型将关系编码为一个矩阵，实体编码为向量，其损失函数如公式（4）所示。

$$f_r(h,t) = h^T W_r t \quad (4)$$

Neural Tensor Network (Socher et al.2013)[1] 是当前较为有效的模型，NTN 模型结合 SLM 和 BM 模型的优点，其损失函数如公式（5）所示。该模型的主要缺点是模型参数较多，时间复杂度高。结合公式（3）、（4）、（5）可以看出，NTN 模型是 SLM 模型及 BM 模型的组合优化。

$$f_r(h,t) = u_r^T g(\mathrm{h}^T \mathrm{M}_r\mathrm{t} + \mathrm{M}_{r,1}\mathrm{h} + \mathrm{M}_{r,2}\mathrm{t} + \mathrm{b}_r) \quad (5)$$

TransE (Bordes et al. 2013)[2] 是近期提出的一种简单高效的编码方式，该模型将关系看作是从头部实体到尾部实体的转移，将知识库中的实体和关系编码到相同维度的向量空间，并假设 $h + r - t \approx 0$。其损失函数如公式（6）所示。TransE 以较低的时间复杂度取得了较好的效果，后续的方法大多数是基于 TransE 的改进。

$$f_r(h,t) = \| h + r - t \|_2^2 \quad (6)$$

TransM (Fan et al. 2014)[3] 通过为每个关系增加一个权重因子 w_r 来改进 TransE 模型，其中 w_r 的计算方式如公式（7）所示。该模型的损失函数如公式（8）所示。其中 $\mathrm{hrpt_r}$ 表示，对于每个关系 r，其尾部实体 t 对应的头部实体的 h 的个数。

$$f_r(h,t) = w_r \| h + r - t \|_{\mathrm{L1/.L2}} \quad (7)$$

$$w_r = \frac{1}{\log(\mathrm{h_rpt_r} + \mathrm{t_rph_r})} \quad (8)$$

TransH (Wang et al) [7]）通过将知识库中的实体和关系编码到不同维度向量来处理非 1 对 1 类型的关系。在该模型中，每个关系使用一个转移向量和一个映射向量来表示。映射向量用于将实体向量映射为与关系向量同维度的向量。其映射函数及损失函数如公式(9)、(10）所示。

$$h_\perp = \mathrm{h\text{-}w}_r^T h w_r \quad \text{and} \quad t_\perp = \mathrm{t\text{-}w}_r^T t w_r \quad (9)$$

$$f_r(h,t) = \| h_\perp + r - t_\perp \|_2^2 \quad (10)$$

TransR (Lin et al. 2015)[8] 与 TransH 类似，该模型将知识库中的关系用一个转移向量和一个映射矩阵表示。通过映射矩阵与实体向量的运算，完成从实体空间到关系空间的映射，其损失函数如公式（11）、（12）所示。

$$h_r = hM_r \quad \text{and} \quad t_r = tM_r . \quad (11)$$

$$f_r(h,t) = \| h_r + r - t_r \|_2^2 \quad (12)$$

IIKE (Fan et al. 2015)[9] 将三元组看做实体 h、关系 r 和实体 t 之间的联合概率，该模型的概率计算函数如公式（13）所示。

$$\mathrm{Pr(h,r,t)} = \sqrt[3]{\mathrm{Pr(h|r,t)\,Pr(r|h,t)\,Pr(t|h,r)}}$$

$$\mathrm{Pr(h|r,t)} = \frac{\exp^{D(\mathrm{h,r,t})}}{\sum_{h' \in \mathrm{Eh}} \exp^{D(\mathrm{h',r,t})}} \quad \text{and} \quad D(\mathrm{h,r,t}) = -\| \mathrm{h} + \mathrm{r} - \mathrm{t} \| + \mathrm{b} . \quad (13)$$

综上所述，目前大部分模型通过优化 TransE[2]的损失函数来改善知识库编码，进而改善三元组分类效果。但是上述模型在三元组分类的时候，只使用了实体和关系的分布式表示。分布式表示存在缺乏先验知识和知识库本身不完整的缺点，导致仅依靠知识库的分布式表示

很难取得很好的效果。针对分布式表示的这种问题，本文提出，综合三元组的距离、关系的先验概率及实体与关系上下文的拟合度来提高分类精确度。

3 基于分布式表示和多特征融合的三元组分类方法

为解决前文中提到的问题，我们首先对 TransE 模型进行分析以得出解决方案。首先，根据 TransE 的损失函数公式（6）可知，当两个实体 h1 和 t1 之间存在两种不同关系 r_1 和 r_2 时，有$(h_1,r_1,t_1) \in \Delta^+$ 且 $(h_1,r_2,t_1) \in \Delta^+$，从而可以得到$h_1+r_1-t_1\approx 0$和$h_1+r_2-t_1\approx 0$。依次可以推出 $r_1\approx r_2$。此时，当知识库中存在三元组$(h_2,r_1,t_1) \in \Delta^+$，则可以推出$h_2+r_1-t_1\approx 0$，结合上面推出的 $r_1\approx r_2$，可以推出 $h_2+r_2-t_1\approx 0$，推出$(h_2,r_2,t_1) \in \Delta^+$。但是很多情况下 h_2 和 t_1 并没有 r_2 关系。具体的例子如下：

(Obama, born_in, USA) $\in \Delta^+$、 (Obama, president_of, USA) $\in \Delta^+$

⇨ Emb(born_in) ≈Emb(president_of)

(Justin_Biber, born_in, USA) $\in \Delta^+$

⇨ Emb(Justin_Biber) + Emb(born_in)-Emb(USA) ≈0

⇨ Emb(Justin_Biber) + Emb(president_of)-Emb(USA) ≈0

⇨ (Justin_Biber, president_of, USA) $\in \Delta^+$

本文对四个常用的数据集进行统计分析，其结果如表 2 所示。从表中可以看出，数据集中实体对具有不同关系的情况普遍存在，且在 FB13 和 FB15K 特别明显。由上述分析可推出，当一个实体对具有多种不同关系的时候，TransE[2]会将其编码到相似的空间，容易引起错误的分类。

表 2 数据集统计信息

数据集	WN11	WN18	FB13	FB15K
#实体	38,696	40,943	75,043	14,951
#训练集	112,581	141,442	316,232	483,141
#验证集	2,609	5,000	5,908	100,000
#测试集	10,544	5,000	23,733	118,142
M-to-N 关系数目	11	18	12	694
#实体相同&关系不同三元组	148	277	20,665	125,203
#r 和 t 相同&h 不同三元组	≈0.9m	≈1.1m	≈1.6m	≈44m
#h 和 r 相同&t 不同三元组	≈1m	≈1.2m	≈4m	≈26m

注： “#”表示数量；“&”表示“与”关系；

在进行三元组分类时，通过公式（6）计算给定（h,r,t）的距离。但当实体对具有关系 r’，且 r’与 r 的编码相似时，就会错误的将三元组（h,r,t）判定为真。因此，解决上述问题的关键是如何正确的区分编码相近的关系。针对该问题，本文通过启发式的方法来对相似空间的不同关系进行区分，以提高三元组分类的准确率。本文使用的启发式规则包括关系的先验概率和关系的上下文。在分布式表示模型中，实体或者关系之间语义的相似度通常使用计算向量之间的相似度表示。具体地，TCSF 使用欧几里得距离公式（14）来计算两个关系的相似度：

$$dis(r,r') = \sqrt{\sum_{i=1}^{d}(r_i - r_i^{'})} \qquad (14)$$

3．1 关系的先验概率

本文使用的第一个启发式规则是关系的先验概率，即一个关系出现的次数越多，则实体对（h,t）具有该关系的概率就越大。给定候选三元组（h,r,t），本文通过对比给定关系 r 最相似的关系 r’来确定关系 r 的先验概率，如公式（15）所示：

$$prob(r) = \frac{N_r}{N_r + N_{r'}} \quad (15)$$

其中 N_r 是关系 r 在训练集中出现的次数，$N_{r'}$ 是与关系 r 最相似的关系。本文只考虑最相似关系的原因是相似的关系得到损失值较为类似，容易引起错误分类。当关系的表示相似度差别较大的时候，根据公式（6）不容易产生误分类。

Algorithm 1 The triple classification of TCSF

Input：

Training set T={(h,r,t)}, valid set V={(h,r,t)}, test set S={(h,r,t)}, entities and relations set E and R, margin γ, embedding dim. d, norm L1/L2, step size st, epoche ep, threshold ϵ.

Run TransE and generate the embeddings of entities and relations

foreach r ∈ R *//对于每个关系，计算出其头部实体和尾部实体的均值*

sim(r)=min{dis(r’,r)}, r’ ∈ R and r’≠r

$$avgHead(r) = \frac{1}{n}\sum_{i=1}^{n} h$$, h ∈ H_r and n =| H_r |

$$avgTail(t) = \frac{1}{n}\sum_{i=1}^{n} t$$, t ∈ H_t and n =| H_t |

foreach (h,r,t) ∈ S do *//结合翻译距离及先验概率计算三元组的可信性*

if $f_r(h,t) < th_r$

if $f_{r'}(h,t) < th_{r'}$,r’= sim(r)

if $plausible((h,t),r) > plausible((h,t),r')$

(h,r,t) is positive

else

(h,r,t) is positive

end if

end if

end if

end foreach

3．2 关系的上下文

TransE[2]进行训练时，训练集中的每个三元组（h,r,t）作为正例，负例通过随机替换实体 h 或 t 构造。该方法类似于 Word Embedding 中的训练方法。基于上述讨论，我们认为所有具有关系 r 的实体对集合{(h_1,t_1), (h_2, t_2), …}就构成了关系 r 表示的上下文。在计算（h,r, t）三元组打分的时候，我们可以不仅仅使用关系 r 的表示，同时也可以计算实体对（h,t）与 r 的上下文之间的匹配程度。

具体地，本文将关系 r 的上下文表示为具有关系 r 的三元组的头部实体编码向量的均值

表示及尾部实体的均值表示。向量的均值反应了该关系对应的头部实体和尾部实体的主要特征，如关系 born_in 对应的头部实体通常为人名，尾部实体通常为地名。根据 TransE[2]得到的结果分析可知，相同类型的实体会编码到比较相近的区域，如人名会较为集中的分布在一个区域，地名也会被集中的分布式另外一个区域。因此，给定一个三元组（h,r,t），实体对（h,t）与关系 r 的合理程度可以通过判断头部实体 h 和尾部实体 t 是否与关系 r 对应的头部实体和尾部实体的均值（关系的上下文）的拟合度来进行判断，当实体与关系的上下文不属于同一个领域的时候，则该三元组的置信度较低。上下文的拟合度可以通过公式（16）计算得到。

$$align((h,t),r) = e^{-\sqrt{(h-avgHead(r))^2+(t-avgTail(r))^2}} \quad (16)$$

3．3 算法

TCSF 利用关系的先验概率和实体对与关系上下文的拟合度的对相似的关系进行区分，计算方法如公式（17）所示。Algorithm 1 是模型的简化算法。

$$plausible((h,t),r) = prob(r) * align((h,t),r) \quad (17)$$

4 实验

4．1 实验数据

本文使用四个常用的数据集来测试模型的效果，分别是 WN11[1]、WN18[2]、FB13[1]和 FB15K[2]，这些数据摘自 WordNet 和 FreeBase 知识库。在 TransE[2]的实验设计中，测试集中的负例是通过随机的替换三元组的头部实体或者尾部实体进行构造。为了构建质量更高的测试数据，本文采用 Socher[1]等人提出的方法。在进行实体替换时，替换的实体应该在相应的位置出现过，且不能是已经存在知识库中的三元组。例如，当替换头部实体时，该实体在训练集中作为头部实体出现过。WN11、FB13 及 FB15K 数据集，我们直接使用 Fan[3]等人的构造的数据。本文根据 Fan[3]中的方法构建了 WN18 的验证集和测试集。四种数据集的统计信息如表 2 所示。

4．2 实验结果

本文与主流的几种方法进行比较，包括 TransE[2]、TransM[3]、TransH[7]、TransR[8]。因为实现和参数调整的问题，我们没有得到相应论文中的最好结果，因此本文直接使用各个模型在相应论文中的最高准确率作为对比依据。为了减少因为参数随机初始化对结果造成的影响，本文对每一组参数都进行 10 次实验，并取其平均值和最好值作为最终的结果。通过多次实验验证，得到每种数据集的参数，(d=200, γ=2.0, st=0.02, ep =300 for WN11; d=200, γ=1.0, st=0.003, ep =10 for WN18; d=100, γ=2.0, st=0.03, ep =20 for FB13; d=400, γ=3.0, st=0.002, ep =30 for FB15k)。d 代表实体和关系的维度、γ 是 margin 值、st 为初始的学习率、ep 是迭代的次数。

三元组分类任务使用准确率作为评价指标，计算方法如公式（18）所示。式中 T_p 和 T_n 表示预测正确的正例数和负例数，N_{pos} 和 N_{neg} 代表训练集中的正例数和负例数。ACC 越高表示系统预测的准确率越好。其

$$ACC = \frac{T_p + T_n}{N_{pos} + N_{neg}} \quad (18)$$

实验结果如表 3 所示，TCSF 在四种数据集上的最优结果明显高于其他系统。但平均值在 FB15K 上略低于 IIKE[9]和 TransR[8]。本文提出的方法是在 TransE[2]上的改进，实验结果

显示（表 4 中第一行 TransE 的结果来来自论文[8]，TransE-avg 和 TransE-best 的结果来自于本文的实现），在四种数据集上的三元组分类的准确率的均值和最高值能提升 10 个以上的百分点。

表 3 模型在不同数据集上的准确率

	WN11	WN18	FB13	FB15K
TransE	75.9%	--	70.9%	79.6%
TransH	77.7%	--	76.5%	80.2%
TransM	77.8%	--	72.1%	89.9%
TransR	85.5%	--	74.7%	81.7%
IIKE	--	--	--	91.1%
TransE-avg	66.0%	68.2%	63.3%	76.6%
TransE-bes	71.2%	80.6%	75.6%	80.0%
TCSF -avg	78.0%	79.3%	75.6%	90.2%
TCSF -best	**86.8%**	**94.2%**	**80.4%**	**91.4%**

注： “avg”表示 10 次实验的平均值；“best”表示 10 次实验的最好值；

为了验证我们的方法在不同类型上的关系上的效果，我们在 FB15K 上针对不同类型的关系进行了三元组分类任务，因为该数据集包含的关系数量较多，且如表 1 所示，包含全部四种类型的关系。得到的结果如表 4 所示，从表 4 中可以看出，我们的方法相对于 TransE 在 1-N、M-1 和 M-N 类型的关系上取得了明显的提升，在 1-1 类型的关系上的准确率没有变化。因此我们的方法可以有效的提高非 1-1 类型的关系的分类准确率。

表 4 模型在 FB15K 中不同关系类型上的结果

	#三元组	TransE	TCSF
1-1	454	98.2%	98.2%
1-N	2,078	77.9%	92.7%
M-1	6,084	86.2%	95.5%
M-N	109,526	**79.6%**	**91.2%**

在本文中，我们主要利用了关系的先验概率和关系的上下文两种特征来增强三元组分类的准确性，为了更好的分析不同的特征在不同类型的关系类型中的效果，本文通过在 TransE 训练的分布式表示上分别增加不同别的特征在三元组分类任务上进行测试，结果如表 5 所示：

表 5 不同特征在 FB15K 中不同关系类型上的结果

	#三元组	TransE	TransE+Prop	TransE+Context	TCSF
1-1	454	98.2%	98.2%	98.2%	98.2%
1-N	2,078	77.9%	84.9%	89.2%	92.7%
M-1	6,084	86.2%	88.2%	92.3%	95.5%
M-N	109,526	**79.6%**	**83.6%**	**90.1%**	**91.2%**

注： TransE+Prop 表示在 TransE 的基础上增加先验特征；TransE+Context 表示在 TransE 的基础上增加关系上下文特征；

通过表 5 的结果可以看出，关系的先验特征能够在 1 对多（1-to-N）、多对 1（M-to-1）、多对多（M-to-N）类型的关系上改进三元组分类的效果；关系的上下文特征也可以在 1 对多（1-to-N）、多对 1（M-to-1）、多对多（M-to-N）类型的关系上改进三元组分类的效果。

并且通过结果的对比可知，关系的上下文比关系的先验概率对结果的提高更明显，并且两种特征的叠加可以取得更号的结果。

通过表 3、4、5 的结果可知，本文提出的方法基于 TCSF 能够很大程度上提高三元组分类的效果，我们提出方法有两个关键贡献：

（1）我们的方法能够很大程度的提高三元组分类的结果，产生新的 state-of-art 的结果；

（2）我们通过对比不同的特征对不同类型的关系上的分类效果，更直观的反映了不同的特征在三元组分类中的作用；

（3）本文提出的方法是在知识表示的基础上针对三元组分类任务的扩展，不仅可以在 TransE 上使用，也可以在 TransM、TransH 等模型上使用，具有较好的通用性。

5 结束语

本文在知识库分布式表示的基础上提出了一种特征融合的方法，从而有效提高三元组分类的准确率。实验结果显示，TransE[2]等系统无法很好的建模非 1 对 1 类型的关系。本文通过加入关系的先验知识和关系的上下文特征，能够有效提高三元组分类上准确率，验证了不同的特征对不同类型关系的分类效果。

参考文献：

[1] R. Socher, D. Chen, C. D. Manning, and A. Ng. 2013. Reasoning with neural tensor networks for knowledge base completion. In Advances in Neural Information Processing Systems.

[2] A. Bordes, N. Usunier, A. Garcia-Duran, J. Weston, and O. Yakhnenko. 2013. Translating embeddings for modeling multi-relational data. In Advances in Neural Information Processing Systems 26.

[3] Miao Fan, Qiang Zhou, Emily Chang, Thomas Fang Zheng. 2014. Transition-based Knowledge graph embedding with relation mapping properties. In PACLIC, pages 328–337

[4] Bordes A,Glorot X, Weston J, and Bengio Y. 2012. A semantic matching energy function for learning with multirelational data. Machine Learning 1–27.

[5] A. Bordes, J. Weston, R. Collobert, Y. Bengio et al. 2011. Learning structured embeddings of knowledge bases. In AAAI.

[6] I. Sutskever, R. Salakhutdinov, and J. B. Tenenbaum. 2009. Modelling relational data using bayesian clustered tensor factorization. In NIPS, pages 1821-1828.

[7] Wang Z, Zhang, J, Feng J, and Chen Z. 2014. Knowledge graph embedding by translating on hyperplanes. In Proceedings of AAAI, pages 1112-1119.

[8] Yankai Lin, Zhiyuan Liu, Maosong Sun and Yang Liu, Xuan Zhu. 2015. Learning entity and relation embeddings for knowledge graph completion. In AAAI, pages 2181-2187.

[9] Miao Fan, Kai Cao, Yifan He. 2015. Jointly Embedding Relations and Mentions for Knowledge Population. arXiv preprint arXiv:1504.01683.

[10] Socher, R.; Chen, D.; Manning, C. D.; and Ng, A. 2013. Reasoning with neural tensor networks for knowledge base completion. In Proceedings of NIPS, 926–934.

[11] Bengio, Y.; Ducharme, R.; Vincent, P.; and Jauvin, C. 2003. A neural probabilistic language model. JMLR 3:1137–1155.

[12] Jenatton, R.; Roux, N. L.; Bordes, A.; and Obozinski, G. R. 2012. A latent factor model for highly multi-relational data. In Proceedings of NIPS, 3167–3175.

作者联系方式：
安波　北京市海淀区中关村南四街 4 号中科院软件园 5 号楼 1201 100190　电话 15010199396 电子邮箱 anbo@nfs.iscas.ac.cn

韩先培 北京市海淀区中关村南四街 4 号中科院软件园 5 号楼 1201 100190　电话 13581558099 电子邮箱 hanxianpei@qq.com

孙乐　北京市海淀区中关村南四街 4 号中科院软件园 5 号楼 1201 100190　电话 18600598168 电子邮箱 lesunle@163.com

吴健　北京市海淀区中关村南四街 4 号中科院软件园 5 号楼 1201 100190　电话 13601332546 电子邮箱 wujian@nfs.iscas.ac.cn