文章编号：1003-0077（2011）00-0000-00

基于迁移学习的地理领域概念关系抽取*

熊盛武[1,2]，陈振东[1]，段鹏飞[1,2]*，王娜[1]

（1. 武汉理工大学 计算机科学与技术学院，武汉 430070

2. 交通物联网湖北省重点实验室，武汉 430070）

摘要：在地理等特有领域概念关系抽取过程中，由于其有限的样本标注资源，难以应用深度学习等大规模知识图谱构建技术。迁移学习方法能够利用开放域文本语料资源，帮助解决目标领域训练数据较少的问题。本文针对地理领域文本的时序性特征，利用长短期记忆（Long Short-Term Memory, LSTM）神经网络，构建了基于词特征和句子特征的概念关系抽取模型，针对地理概念关系语料缺乏的问题，提出了基于 LSTM 的迁移学习方法，将开放领域的知识迁移到地理领域，通过权重迁移和重训练调整，显著提升了地理领域概念关系抽取的准确度。

关键词：地理领域；概念关系抽取；迁移学习

中图分类号：TP391　　**文献标识码：**A

Transfer Learning based Concept Relation Extraction in Geographic Domain

Shengwu Xiong[1,2], Zhendong Chen[1], Pengfei Duan[1,2*], Na Wang[1]

(1. Wuhan University of Technology, Wuhan, Hubei 430070, China

2. Hubei Key Laboratory of Transportation Internet of Things, Wuhan University of Technology, Wuhan 430070, China)

Abstract: There is no sufficient corpus in the geographic domain to support the concept relation extraction research, thus it is difficult to apply large-scale knowledge map construction technologies such as deep learning. Alternatively, to solve such problem, several transfer learning based methods have been proposed to transfer the knowledge from the source domain to the target domain. According to the temporal characteristics of geographic text, the LSTM neural network is used in this thesis to construct a concept relation extraction model based on word features and sentence features. According to the lack of geographic concept relation corpus, the LSTM-based transfer learning method is proposed, which transfers the knowledge of the open domain to the geographic domain, through transferring trained network weights and retraining fine-tuned, the accuracy of geographic concept relation extraction can be significantly improved.

Key words: Geographic Domain; Concept Relation Extraction; Transfer Learning

1 引言

概念关系抽取在知识库构建中起着至关重要的作用，而完备且高质量的知识库决定了类人智能系统的智力水平。概念关系抽取主要研究概念对之间是否存在概念关系，但是在地理学科语料资源缺乏的情况下，还无法做到识别出概念对之间具体的概念关系类型。因此，提出地理领域概念关系抽取新方法，对于构建知识库具有重要意义。

知识库构建需要充足的语料作为支撑。例如，由 Google 构建的谷歌知识图谱，关系数目已达到 35 亿条，而且这个数目还在不断扩大；由莱比锡大学、柏林自由大学和 OpenLink

* **收稿日期：** 2017-07-10 **定稿日期：**2017-07-25

基金项目：国家高技术研究发展计划（863 计划）（2015AA015403）；武汉理工大学研究生优秀学位论文培育项目资助（2016-YS-065）

作者简介：熊盛武（1966 年—），男，教授，主要从事机器学习与模式识别的研究；陈振东（1993 年—），男，硕士研究生，主要从事自然语言处理与机器学习研究；段鹏飞（1985 年—），男，讲师，主要从事自然语言处理与机器学习的研究。

Software 联合构建的知识库 DBpedia[1]，截至 2014 年 9 月，关系数已达到 458 万。但是，相较于开放领域文本资源，地理学科的语料资源数据量相差甚远，仅利用这些数据进行地理领域的概念关系抽取研究，其准确率无法得到保证。由此可知，目前地理知识库的构建缺乏充足的地理概念关系语料，不足以支撑概念关系抽取研究。

针对地理概念关系语料不足的问题，传统的解决方法是人工抽取相关地理文本信息，扩充地理概念关系语料库。但这种方法需要人工参与，不仅效率低，人力成本也很高，而且人工处理无法保证语料库的规范性。而使用迁移学习的方法，在目标领域数据量不足的情况下，可以迁移源领域的知识，帮助目标领域更好地完成目标任务。因此，如何利用迁移学习方法解决地理概念关系语料不足的问题，对提升地理概念关系抽取准确率具有重要意义。

本文以地理学科为研究对象，致力于研究能够识别出概念对之间具体概念关系类型的概念关系抽取方法；并研究如何将开放领域的知识有效地迁移到地理领域，解决地理概念关系语料不充足的问题，以提高地理概念抽取的准确率。

2 相关工作

迁移学习的相关技术[2]发展迅速，并且已被广泛应用到计算机视觉与图片处理[3]，自然语言处理[4]，文本分类[5]等众多实际领域。迁移学习方法可以将其他领域已有的知识或数据，迁移到仅有少量甚至没有标签数据的目标领域，协助更好地完成目标任务[6]。

概念关系抽取方面，深度学习技术在标准测试集上得到了比传统机器学习更好的结果。陈宇等使用深度信念网络完成了关系抽取任务，并证明了字特征对于关系抽取研究的有效性[7]。2014 年 Coling 会议上，Zeng 等提出使用 CNN 提取包含实体对的关系语句的句子特征，并将该方法用于关系分类任务[8]，在英文标准关系语料库 SemEval-2010 Task-8[9]上取得了 82.7 的 F 值。在 ACL2015 会议上，Santos 等在 Coling2014 文章的基础上提出了 CR-CNN 模型，该模型主要针对分类器设计进行了改进，将样本类别向量化，并提出了一种新的损失函数[10]，该模型在关系语料库 SemEval-2010 Task-8 上获得了 84.1 的 F 值。

上述方法都只利用了关系语句的部分词语信息进行句子特征提取，没有充分利用整个关系语句的语义信息。为了利用完整的语义信息来提取实体对所在关系语句的句子特征，胡新辰基于 BLSTM 神经网络构建了一个深度学习模型[11]，该模型有效利用了词语间所有的相互信息，并在关系语料库 SemEval-2010 Task-8 上获得了较好的结果。此外，Zhang 等同样认为有助于关系分类的信息可能出现在句子中的任何词语上，故利用双向长短记忆神经网络获取句子中所有词语的信息[12]，在同样的语料库上取得了 84.3 的 F 值，与基于 CNN 的方法相比，其准确率有一定的提升。

以上方法使用的语料库大部分都是 SemEval-2010 Task-8，该语料库由英文编写，不适用于中文地理领域。但对于基础教育地理概念关系抽取的研究，本文可以参照其语料格式，初步构建适量的地理概念关系语料以及可用于迁移学习研究的开放领域概念关系语料库，并将地理概念关系抽取定义为基于给定地理概念对的关系语句分类问题进行研究。

3基于迁移学习的概念关系抽取模型

3.1 构建概念关系语料库

目前公开的语料主要是新闻或博客，难以用作地理概念关系抽取语料；而常用的关系抽取数据集 SemEval-2010 Task-8，由英文编写，不适用于中文关系抽取研究。因此，本文首先选定两类地理概念关系，然后参照 SemEval-2010 Task-8 的数据格式，人工获取包含地理概念对的关系语句，构建适用于地理概念关系抽取研究的语料库，并制定两类标注规则，对语料库进行标注。

根据 MUC7 会议提出的定义，概念关系可分为“分类关系”（Taxonomy）和“非分类关系”（Non-Taxonomy）两大类，本文主要针对地理概念间的两类关系进行研究：上下位关系和同义关系。

例如文本“**衡阳盆地** 中国江南地区具有地域特点的**红层盆地**……龙门山 **沱江**和**岷江**的分水岭，四川省著名地震带”，其中“衡阳盆地”和“红层盆地”两地理概念间存在上下位关系，“沱江”和“岷江”两地理概念间存在同义关系。

地理概念关系语料主要从《中国大百科全书-中国地理》和《百度百科》地理词条两类地理资源获取。其中，《中国大百科全书-中国地理》包含了较完整的地理信息，便于我们从中获取表征地理概念对上下位关系和同义关系的语句，《百度百科》地理词条提供的地理信息，主要用于对《中国大百科全书-中国地理》中无法提取出关系语句的内容进行补充。

地理领域语料存在一个共性：含有大量无用的描述信息以及少量历史、文化信息等。为了确保语料库的质量，对于例句图 1，本文对地理资源进行了预处理：去除与地理描述无关的历史信息；去除与地理概念关系相关度较低或无关的信息，如英文翻译等；标点符号统一化；语义成分残缺部分补全。最终，我们可以得到表征同义关系的语句：“大相岭是大渡河和青衣江的分水岭，四川省南北部重要的自然地理界线”。

大相岭~~（DaxiangLing）~~大渡河和青衣江的分水岭,四川省南北部重要的自然地理界线。~~又称相公岭和泥巴山。山体蜿蜒于省境西部雅安、荥经和汉源之间,西靠二郎山,东接峨眉山,走向近北西。~~ ~~南北自然景观迥然不同,历史上即有“清风、雅雨、干富林”之说。~~

图 1 语料例句

获取一定量的语料库后，为达到关系抽取训练数据的要求，本文进行了语料分词和标注工作。分词处理过程中，采用 ICTCLAS 工具进行分词，在初步分词结果中，例如“震旦纪变质岩”这种地理概念特定词语，容易出现被分词器切分为多个词的情况。为了提高分词准确度，本文使用了 Sogou 公司提供的通过 10TB 语料训练的词向量表进行了分词修正，词向量表中含有超过 420 万个词向量，分词前利用其中的词表对分词系统的字典进行了扩充，在一定程度上提升了分词结果的准确性。

语料标注过程中，本文针对两种地理概念关系类型指定了特殊的标注模式：在分词基础上，使用特定标识符号分别标识上下位关系和同义关系中的地理概念，而分词结果中剩余的概念无关词语则标识为其他。经过分词和标注处理后，最终构建了一定规模的地理概念关系语料库。

3.2 共享信息迁移

为了完善概念关系抽取模型，继续采用人工方法构建大规模地理概念关系抽取语料的做法不符合实际。由迁移学习方法的思想可知，该方法可以将其他领域已有的知识或数据，迁移到仅有少量标签数据的目标领域，以帮助目标领域更好地完成目标任务，因此本文将迁移学习的思想应用到开放领域和地理领域中，利用开放领域的知识或数据，帮助标注语料较少的地理领域提高概念关系抽取的准确率。

迁移主要是利用开放域文本与地理文本间的共享信息，帮助地理领域更好地完成目标任务，这样的做法可以间接解决地理概念关系语料不足的问题。开放域文本和地理文本间的共享信息描述如图 2 所示。

图 2 中间部分描述的信息“…和(与)…是(属于)…的…”及“…是…之一…”，均为开放域文本和地理文本在语法结构上的相似性，代表了两个领域间的部分共享信息。两个领域间的共享信息越多，利用越充分，迁移的效果就会越理想，因此本文将借助开放域文本进行迁移学习研究。

本文以地理概念关系语料库的构建方式，构建了开放领域的概念关系语料库，两者的区别在于：该语料库充分保留了开放域文本原有的复杂性，关系语句的语义丰富，避免了地理概念关系语料构建时出现的语义表达过于简单等问题。该语料库与地理概念关系语料库相似，同样包含能够表征概念对间上下位关系和同义关系的关系语句。本文构建的开放域语料库共包含了 11082 个概念关系词语，地理领域语料库则包含 2416 个概念关系词语

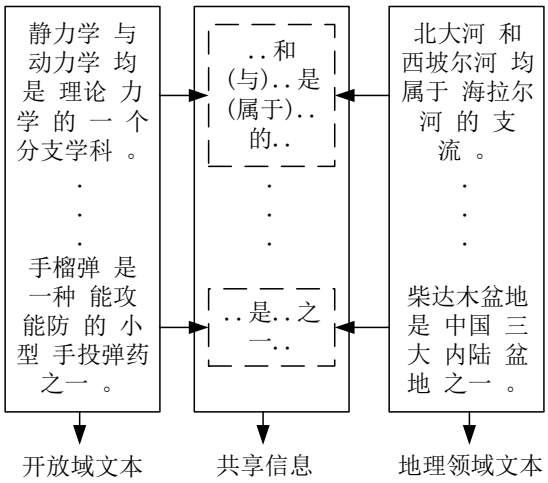

图 2 共享信息描述

3.3 LSTM 网络概念关系抽取模型

本文结合概念对的词特征及概念对所在语句的句子特征，构建了地理概念关系抽取模型，该模型将地理概念关系抽取定义为基于给定地理概念对的关系语句分类问题，详细的概念关系抽取网络结构，如图 3 所示。

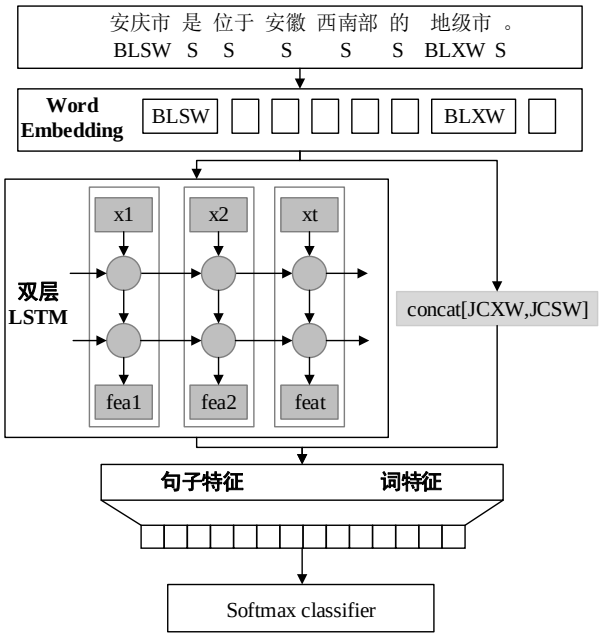

图 3 地理概念关系抽取的网络结构

概念关系抽取网络结构可被分为以下几个主要部分：

（1）Word Embedding 处理层：将所有的输入转换成词向量形式；

（2）LSTM 处理层：使用 LSTM 网络结构抽取概念对所在语句的句子特征，其中 LSTM 网络结构包含两种：单层结构和双层结构。不同 LSTM 网络结构的特征提取能力不同，本文将分别对比单层和双层 LSTM 网络结构提取句子特征的能力，取其优。

（3）特征抽取层：从 Word Embedding 层的输出中提取概念词特征，从 LSTM 层的输

出中提取概念关系语句的句子特征，两种特征的加入，能够提高关系识别的准确率；特征提取层后连接最大池化层，用于提取关系语句重要特征；

（4）特征融合及分类：将以上抽取的两类特征沿着最后一个维度进行拼接实现特征融合，并将融合后的特征送入 Softmax 分类器进行关系分类。

3.4 迁移学习

在 3.2 节中，我们构建了地理概念关系语料库，但语料的规模相对来说并不充足，在这种情况下，为了得到更好的概念关系抽取结果，需要借助其他领域的知识，并迁移到地理领域，然后对地理概念关系抽取进行重新学习。

本文研究的跨领域概念关系抽取问题涉及了开放领域和地理领域。这两个领域的目标任务相同，并且数据都带有标签，不同之处在于：开放领域的概念关系语料充足，而地理领域的概念关系语料较少。所以需要利用迁移学习方法解决的问题是：开放领域完成概念关系抽取后的知识如何有效地迁移到地理领域，辅助地理领域更准确地完成概念关系抽取。

基于深度学习与迁移学习相结合的思想，本文提出基于 LSTM 神经网络的迁移学习方法。本文分别从 LSTM 神经网络的 Embedding 层和网络权重两方面，对开放领域和地理领域间的知识迁移问题进行研究，提出了两类知识迁移方法：（1）基于 Word Embedding 的词向量迁移学习方法，从词向量的角度进行迁移学习尝试；（2）基于网络权重的迁移学习方法，使用开放领域文本训练得到的 LSTM 神经网络权重按层迁移到地理领域，对两个领域间的知识迁移问题进行更深入的研究。

3.4.1 词向量迁移

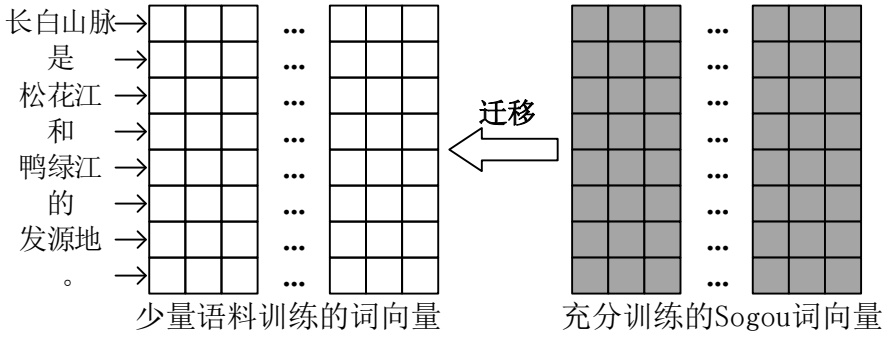

图 4 词向量迁移

Word Embedding 层可以将输入的文本处理成词向量的形式，这是使用 LSTM 神经网络进行模型训练的基本操作。词向量迁移方法的基本思想：将 Sogou 公司提供的词向量迁移到地理领域，然后地理领域使用迁移后的词向量重新完成自身的概念关系抽取。

表 1 开放领域和地理领域的词向量描述

属性	开放领域	地理领域
来源	Sogou 提供	Word2Vec
词向量维度	100	100
词向量数量（个）	4200005（训练语料：10TB）	5071

开放领域的词向量由 Sogou 公司提供，地理领域自身的词向量使用 Word2Vec 处理得到，详细介绍如表 1 所示。Sogou 词向量迁移到地理词向量的描述，如图 4 所示。

3.4.2 权重迁移

与深度学习相结合，是目前基于权重的迁移学习方法的主流方向。目前，基于 CNN 的迁移学习研究较多，如 Hafemann 等先使用数据量充足的开放领域训练一个 CNN 分类模型[13]，然后使用该模型将仅有少量数据的目标领域映射到另一个新的特征空间，并在该特征空间上重新训练一个适用于目标领域的分类器，完成纹理分类任务。Oquab 等基于 CNN，

研究了如何将使用大规模标注数据集 ImageNet 训练得到的图像表征有效地迁移到包含有限训练数据的数据集 PASCAL VOC 上，以更好地完成视觉识别任务[14]；同时还基于 CNN 网络结构，设计了按层迁移的对比实验，实验结果表明：深度神经网络结构的低层网络更适合迁移，因为低层网络抽取的是普遍特征，即不同领域共享的数据特征；而更高层不适合迁移，因为其提取出的特征代表各个领域的特有特征。

由上可知，目前深度学习与迁移学习方法相结合的应用中，常见的是将待迁移的知识转换为神经网络结构中的权重值，因 LSTM 神经网络能有效利用序列数据中长距离的依赖信息，本文基于 LSTM 神经网络针对概念关系语料不足的基础教育地理领域进行迁移学习尝试，该方法主要从应用层面研究如何提升迁移效果。

不同领域的数据经过相同的神经网络训练后，获得的网络权重能够表征每个领域包含的语义信息，对于开放领域和地理领域，两者的目标任务都是完成概念关系抽取，并且概念关系类型相似，因此网络权重存在一定程度的相似性，这些相似性能够用于构建迁移桥梁，也就是使用开放领域与地理领域间相似的知识，解决地理概念关系语料不足的问题。

权重迁移实验中，首先使用开放领域和地理领域文本分别训练 LSTM 概念抽取网络，得到开放领域网络权重后，将按层迁移到地理领域相应的网络部分。迁移后对新组成网络模型的概念关系抽取能力进行评估，为了达到更优的关系抽取效果，本文将在开放领域知识的基础上进行有限次的重训练调整。

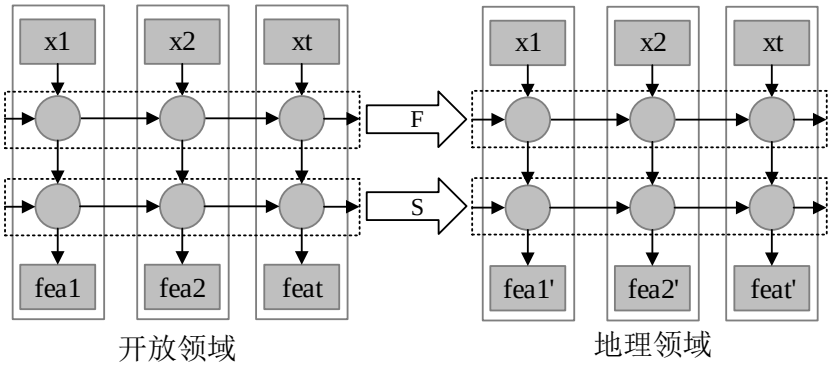

图 5 网络权重按层迁移

在确保了概念关系抽取网络结构相同的情况下，我们可以按层迁移网络权重，并且在这种对等的情况下，单独研究网络权重迁移对结果的影响。基于双层 LSTM 网络结构的网络权重按层迁移过程，如图 5 所示，单层 LSTM 迁移方式同理可得。

由图 5 可知，基于双层 LSTM 网络结构按层迁移网络权重，有三种迁移方式：迁移第一层、迁移第二层及两层全迁移，再依据是否对迁移的网络权重进行重训练调整，本文最终将地理领域的重训练方式分为六种，相关实验及结果分析将在第 4 节详细描述。

4 实验及结果分析

依据 MUC 会议的标准，关系抽取性能的好坏主要根据准确率（P）和召回率（R），为了综合评价性能的好坏，本文采用 F 值作为评价指标，F 值即召回率和准确率的加权几何平均值，其定义如公式（3）所示。

$$P=\frac{\text{某关系类型被正确分类的实例数}}{\text{模型预测某关系类型的实例数}} \tag{1}$$

$$R=\frac{\text{某关系类型被正确分类的实例数}}{\text{测试集中某关系类型的实例数}} \tag{2}$$

$$F=\frac{((w^2+1.0)*P*R)}{((w^2*P)+R)} \tag{3}$$

本文将 w 的值设置为 1，认为准确率和召回率同等重要。最终采用以上三类评价指标的加权平均值作为性能评测标准，分别记为 P_{Avg} 、 R_{Avg} 和 F_{Avg} 。

4.1 LSTM 网络概念关系抽取

概念关系抽取研究中，深度学习技术的应用已经非常广泛，如 Zeng 等人[8]结合句子特征构建了 CNN 进行关系分类研究，Zhang 等人[12]利用 Bi-LSTM 更充分地利用词语间的相互信息，因此本文结合概念对的词特征和概念对所在关系语句的句子特征，共做了四组实验。第一组实验使用单层 LSTM 网络结构抽取句子特征，第二组则使用双层 LSTM 网络结构，三四组实验分别使用 CNN 和 Bi-LSTM 网络替换本文网络的句子特征抽取部分，以进行对比试验。实验过程中，将地理概念关系语料库 80%的数据作为训练集，20%作为测试集，结果详见表 2。

表 2 地理概念关系抽取结果对比

网络结构	P_{Avg}	R_{Avg}	F_{Avg}
单层 LSTM	94.01	93.94	93.95
双层 LSTM	**94.81**	**95.15**	**94.94**
CNN[8]	93.28	93.34	93.31
Bi-LSTM[12]	94.31	**95.15**	94.64

从实验结果可以看出，使用双层 LSTM 网络结构抽取句子特征，在平均 F 值等各方面相比单层 LSTM 网络都有一定的提升， CNN、Bi-LSTM 网络的对比实验结果中，在标识了语句中概念对关系的情况下，两种网络抽取结果没有突破，可以看出，对于本文的地理领域概念关系抽取网络结构， 双层 LSTM 网络结构能更充分地将词语间的相互关系信息运用到每个词的训练中，因此信息提取能力更强。

4.2 词向量迁移实验

基于 4.1 节实验结果，本文的概念关系抽取网络最终选用了双层 LSTM 网络结构。接下来使用迁移后的地理词向量完成了迁移实验，实验结果如表 3 所示。

表 3 词向量迁移实验结果对比

词向量类型	P_{Avg}	R_{Avg}	F_{Avg}
地理词向量	94.81	95.15	94.94
迁移词向量	96.87	93.33	95.03

结果表明，在平均准确率及平均 F 值两个指标上有一定的提升，但是在平均召回率上有所下降，该方法对地理概念关系抽取的准确率提升不明显。分析原因有二：

1） 地理概念关系语料库数据量少

本文训练词向量过程依赖的地理概念关系语料库规模较小，而 Sogou 词向量是由大量开放领域文本库训练得到，相对来说 Sogou 词向量特征表达效果比本文地理词向量更加精准。

但是从实验结果来看，迁移了特征表达能力较强的 Sogou 词向量后，特征抽取能力并没有显著提高，由此可知，迁移词向量方法没有真正解决地理概念关系语料库数据量少的问题，实验结果不会有太显著的改善。

2） 迁移后地理词向量的特征空间不一致

迁移 Sogou 词向量过程中，存在 212 个词向量迁移失败的情况，本文对迁移失败的处理方式是保持原来向量值，这可能导致迁移失败的词向量与 3695 个迁移成功词向量的特征空间不一致[15]，进而导致信息表达不准确甚至断层，引发错误，使得迁移词向量方法的实

验结果不佳。

综上可知，基于 Word Embedding 的词向量迁移学习方法，虽然有一定的迁移效果，但它没有真正解决地理概念关系语料不足的问题，同时还产生了特征空间不一致的负作用，无法有效提升地理概念关系抽取的准确率。

4.3 网络权重迁移实验

由于 LSTM 网络结构分层的结构特点，网络权重迁移中有三种迁移方式：迁移第一层、迁移第二层及两层全迁移。网络模型组合新的权重即得到了新的迁移模型，在评估新模型的概念关系抽取能力后，本文对网络模型进行了有限次的重训练调整，因此地理领域的重训练方式分为六种，详见表 4。

表 4 权重迁移训练描述

重训练调整策略	迁移 LSTM 网络结构层次	简称
保持不变（R）	双层 LSTM 网络结构	R+D
保持不变（R）	双层 LSTM 第一层	R+F
保持不变（R）	双层 LSTM 第二层	R+S
重新训练（T）	双层 LSTM 网络结构	T+D
重新训练（T）	双层 LSTM 第一层	T+F
重新训练（T）	双层 LSTM 第二层	T+S

本文首先使用“保持不变”（R）的迁移方式进行了网络权重迁移实验，以“R+D”训练方式为例，实验结果详见表 5。

表 5 权重迁移“R+D”实验结果

P_{Avg}	R_{Avg}	F_{Avg}
7.57	1.21	2.08

由表 5 可知，将开放领域的网络权重迁移到地理领域后，若权重保持不变，直接用于地理领域的概念关系抽取，最终获得的准确率、召回率都非常低，意味着迁移失败。基于“R”迁移方式其他两类训练方式同样迁移失败，具体实验结果不再赘述。

跨领域研究的前提是不同领域间存在一定的相似性，也就是指跨领域间的共享信息，而除了共享信息，领域内还有各不相同的特有信息。在知识迁移过程中，领域间共享的信息越多，利用越充分，迁移效果就会越好；反之，若源领域的特有信息过多迁移到目标领域，会导致迁移效果不理想。本组“R”迁移方式实验直接将开放领域权重整体迁移到了地理领域，也就代表着共享信息迁移的同时，特有信息也全部迁移，这种迁移方式没有考虑到地理领域自身的特有信息，因此导致了迁移实验的失败。

其次，本文使用“重新训练”（T）迁移方式的三类训练方式进行了网络权重迁移实验，实验结果如表 6。

表 6 权重迁移“T+ D/F/ S”实验结果

迁移方式	P_{Avg}	R_{Avg}	F_{Avg}
双层 LSTM（不迁移）	94.81	95.15	94.94
T+D	95.86	**96.97**	**96.41**
T+F	94.81	95.76	95.25
T+S	**96.25**	93.94	95.05

开放领域的网络权重按层迁移到地理领域网络中，无论是迁移第一层、第二层或者两层全迁移，使用地理文本对迁移的网络权重再度训练后，最终的平均 F 值都有显著提升，迁

移学习成功运用在了开放领域和地理领域之间。

5 结语

本文以地理学科为研究对象，利用 LSTM 神经网络，对地理概念关系抽取进行了研究。首先，本文转换了中文地理概念关系抽取问题，并定义为基于给定概念对关系的语句分类问题，参照 SemEval-2010 Task-8 数据集构建了适用于地理概念关系抽取的标注语料库；其次，针对地理概念关系语料不足的问题，本文从深度学习的角度出发，提出了基于 LSTM 神经网络的迁移学习方法，分别从词向量迁移和网络权重迁移两个方向进行了研究，实验结果表明，基于 Word Embedding 的词向量迁移学习方法不能显著提升地理概念关系抽取的准确率；网络权重迁移后若不进行地理领域文本的重训练，将会导致迁移失败。最终经过重新训练的权重迁移实验结果表明，开放域文本和地理文本间存在共享信息，可用于构建迁移桥梁。同时，也说明各个领域含有领域特有信息，迁移后的网络权重若保持不变，即将开放领域的特有信息全部迁移到地理领域，会导致迁移失败；若使用地理文本对迁移后的网络权重进行重训练调整，即考虑地理领域自身的特有信息，能较显著地提升地理概念关系抽取的准确率。

参考文献

[1] Bizer C, Lehmann J, Kobilarov G, et al. DBpedia-A crystallization point for the Web of Data [J]. Web Semantics Science Services & Agents on the World Wide Web, 2009, 7(3): 154-165.

[2] Zhuang F, Luo P, Xiong H, et al. Cross-Domain Learning from Multiple Sources: A Consensus Regularization Perspective [J]. IEEE Transactions on Knowledge & Data Engineering, 2010, 22(12):1664-1678.

[3] Kandaswamy C, Silva L M, Alexandre L A, et al. Improving Deep Neural Network Performance by Reusing Features Trained with Transductive Transference[C]. Artificial Neural Networks and Machine Learning. 2014:265-272.

[4] Huang J T, Li J, Yu D, et al. Cross-language knowledge transfer using multilingual deep neural network with shared hidden layers[C]. IEEE International Conference on Acoustics, Speech and Signal Processing, 2013:7304-7308.

[5] Behbood V, Lu J, Zhang G. Text categorization by fuzzy domain adaptation[C]. IEEE International Conference on Fuzzy Systems. IEEE, 2013:1-7.

[6] Pan S J, Yang Q. A survey on Transfer Learning [J]. IEEE Transactions on Knowledge & Data Engineering, 2010, 22(10):1345-1359.

[7] 陈宇，郑德权，赵铁军. 基于 Deep Belief Nets 的中文名实体关系抽取[J]. 软件学报，2012, 23(10):2572-2585.

[8] Zeng D, Liu K, Lai S, et al. Relation classification via convolutional deep neural network[C]. Proceedings of COLING. 2014: 2335-2344.

[9] Hendrickx I, Su N K, Kozareva Z, et al. SemEval-2010 task 8: multi-way classification of semantic relations between pairs of nominals[C]. The Work-shop on Semantic Evaluations: Recent Achievements and Future Directions. Association for Computational Linguistics, 2009:94-99.

[10] Santos C N D, Xiang B, Zhou B. Classifying Relations by Ranking with Convolutional Neural Networks [J]. Computer Science, 2015.

[11] 胡新辰. 基于 LSTM 的语义关系分类研究[D]. 哈尔滨工业大学, 2015.

[12] Zhang S, Zheng D, Hu X, et al. Bidirectional Long Short-Term Memory Networks for Relation Classification [J]. 2015.

[13] Hafemann L G, Oliveira L S, Cavalin P R, et al. Transfer learning between texture classification tasks using Convolutional Neural Networks[C]. International Joint Conference on Neural Networks. 2015: 1-7.

[14] Oquab M, Bottou L, Laptev I, et al. Learning and Transferring Mid-level Image Representations Using Convolutional Neural Networks[C]. IEEE Conference on Computer Vision and Pattern Recognition. 2014: 1717-1724.

[15] Pan J, Hu X, Li P, et al. Domain adaptation via Multi-Layer Transfer Learning [J]. Neurocomputing, 2016, 190:10-24.

作者联系方式：段鹏飞，湖北省武汉市珞狮路 122 号武汉理工大学鉴湖校区鉴主 1106，430070，15972100809，duanpf@whut.edu.cn