Software Engineering and Applications 软件工程与应用, 2025, 14(4), 906-918
Published Online August 2025 in Hans. https://www.hanspub.org/journal/sea
https://doi.org/10.12677/sea.2025.144080

Hans汉斯

基于实体关系双向引导的中文少样本关系三元组抽取

李雨辰，张明西*，殷菘泽，曹祥龙，王凌璇

上海理工大学出版学院，上海

收稿日期：2025年7月11日；录用日期：2025年8月15日；发布日期：2025年8月26日

摘　要

关系三元组抽取是知识图谱构建中的关键任务。然而，实际应用中关系类别普遍呈现长尾分布，许多低频关系缺乏充足的标注样本，严重制约了监督学习方法的性能。为此，我们提出了一种基于实体关系双向引导的中文少样本关系三元组抽取方法，通过构建实体识别与关系分类之间的双向信息流，增强子任务间的信息交互与协同建模能力。为提升模型的表示能力和分类判别性能，我们设计了基于token置信度的加权原型构建方法，以动态评估各token的语义贡献并优化原型表示。同时，我们引入原型多样性约束损失，在提升类内一致性的基础上增强类间可分性，从而强化类别判别效果。实验结果表明，所提的方法在多个评价指标上显著优于主流基线模型，验证了其在少样本场景下的有效性与泛化能力。

关键词

关系三元组抽取，知识图谱，少样本学习

Bidirectional Entity-Relation Guided Chinese Few-Shot Relation Triple Extraction

Yuchen Li, Mingxi Zhang*, Songze Yin, Xianglong Cao, Lingxuan Wang

College of Publishing, University of Shanghai for Science and Technology, Shanghai

Received: Jul. 11th, 2025; accepted: Aug. 15th, 2025; published: Aug. 26th, 2025

Abstract

Relation triple extraction is a critical task in the construction of knowledge graphs. However, in real-world applications, the distribution of relation types often follows a long-tail pattern, where

*通讯作者。

文章引用: 李雨辰, 张明西, 殷菘泽, 曹祥龙, 王凌璇. 基于实体关系双向引导的中文少样本关系三元组抽取[J]. 软件工程与应用, 2025, 14(4): 906-918. DOI: 10.12677/sea.2025.144080

many low-frequency relations lack sufficient annotated examples, severely limiting the performance of supervised learning methods. To address this challenge, we propose a Chinese few-shot relation triple extraction method based on a bidirectional entity-relation guidance mechanism. By constructing a bidirectional information flow between entity recognition and relation classification, our method enhances the interaction and joint modeling between the two subtasks. To improve the model's representation capability and classification performance, we further design a token-confidence-based weighted prototype construction method, which dynamically assesses the semantic contribution of each token to optimize prototype representations. In addition, we introduce a prototype diversity constraint loss to improve intra-class consistency while enhancing inter-class separability, thereby strengthening the model's discriminative ability. Experimental results show that our method significantly outperforms mainstream baseline models across multiple evaluation metrics, demonstrating its effectiveness and generalization ability in few-shot scenarios.

Keywords

Relation Triple Extraction, Knowledge Graph, Few-Shot Learning

1. 引言

关系三元组抽取是信息抽取领域的重要任务，旨在从自然语言文本中识别出由主语、谓语和宾语构成的结构化关系，广泛应用于知识图谱构建、问答系统、推荐系统等场景[1]。近年来，预训练语言模型(Pre-trained Language Models, PLMs)的引入显著提升了关系抽取任务的整体性能[2]。然而，在真实应用中，关系类别往往呈现长尾分布特征，尤其在中文语境中，不同关系类别呈现出明显的长尾分布特征，存在大量低频甚至极少样本关系类别，导致基于监督学习的方法在这些长尾关系上的表现较差[3]。如何在标注样本稀缺的条件下准确识别这些关系三元组，已成为当前关系抽取研究中的关键挑战之一。

针对标注样本稀缺的问题，少样本关系三元组抽取(Few-Shot Relation Triple Extraction, FS-RTE)逐渐成为研究热点。现有方法主要分为三种范式：先实体后关系、先关系后实体，以及实体关系双向引导。先实体后关系范式首先识别文本中的实体，再基于实体对进行关系分类[4]。该方式结构清晰，便于实现，但在少样本场景下面临两个主要问题：一是未考虑关系差异性的实体识别难以覆盖所有关系语义；二是可能引入冗余实体，干扰后续关系预测[5]。先关系后实体范式则从关系判别入手，以预测的关系类别作为引导信息辅助实体抽取，从而缓解实体语义差异带来的影响[5]。然而，这两种单向建模范式在建模过程中缺乏实体与关系之间的交互，限制了原型表示的表达能力。为解决上述问题，研究人员提出实体关系双向引导范式，通过构建实体与关系之间的双向交互机制，实现两者表征的协同优化，显著提升了模型在少样本条件下的性能和鲁棒性[6]。尽管实体关系双向引导范式在少样本关系三元组抽取任务中取得了显著进展，但现有方法在原型表示与建模机制上仍存在明显局限。首先，在原型构建方面，大多数方法采用 token 的等权平均方式生成实体或关系原型，忽略了不同 token 对类别语义的重要程度差异。这种构建方式容易受到冗余或无关信息干扰，导致原型表示的语义模糊，影响实体与关系之间的匹配精度。其次，在原型分布建模方面，类内原型往往表现出较大的分散性，而类间原型之间的距离却较为接近，导致整体原型区分度不足。同时，现有方法普遍缺乏专门的结构性约束机制对原型空间进行规范，进一步加剧了匹配边界的模糊性，降低了三元组抽取的稳定性和鲁棒性。

为缓解上述现有方法在原型构建和表示方面的不足，本文提出了一种基于实体关系双向引导范式的中文少样本关系三元组抽取方法。在本文方法的设计过程中，主要面临以下两个关键挑战。第一，支持样本稀缺导致类别原型难以准确构建。在少样本条件下，实体与关系的语义表示本就有限，而现有方法通常采用 token 的等权平均方式生成原型，忽略了不同 token 对类别语义的重要程度差异，导致原型表达模糊、易受噪声干扰。为应对这一问题，本文引入了基于 token 置信度的加权原型构建机制，通过动态评估每个 token 的语义贡献，实现更精确的原型聚合，从而提升模型在有限样本下的表示能力。第二，原型空间缺乏结构约束，影响匹配判别稳定性。由于类内原型分布易发散、类间原型区分度不足，模型在三元组抽取过程中容易出现边界模糊和预测不稳定现象。为此，本文进一步设计了原型多样性约束损失函数，在增强类间分离性的同时强化类内一致性，从而提升模型的鲁棒性和泛化性能。总的来说，本研究的主要贡献如下：

- 提出了一种基于 token 置信度的加权原型构建方法，通过引入动态语义评估机制，有效识别并强化对类别语义贡献较大的关键 token，从而缓解传统等权平均策略导致的语义模糊与噪声干扰问题，显著提升了原型表示的准确性与鲁棒性。
- 设计了一种原型多样性约束损失函数，在类间最大化原型差异、类内最小化原型分散，从而规范原型空间的分布结构。该设计有效缓解了少样本条件下类间边界模糊、类内原型发散等问题，显著提升了模型在低资源场景下的判别能力与泛化稳定性。
- 在 FewDuIE1.0 与 FewDuIE2.0 两个中文少样本数据集上进行了系统实验与消融分析，实验结果显示，本文方法在多个评价指标上均显著优于主流基线方法，两个核心模块在不同任务设定下均带来显著性能提升，充分验证了所提方法的有效性、鲁棒性与适应性。

2. 相关工作

关系三元组抽取旨在从自然语言文本中识别由主语、谓语和宾语构成的结构化关系三元组[1]。早期研究主要采用管道式方法[7]，将任务划分为命名实体识别[8]和关系抽取[9]两个阶段。该方法结构简单，但容易受到错误传播的影响，限制了整体性能的提升。为克服流水线方法的不足，部分研究提出了联合抽取模型，通过端到端学习实现实体和关系的同步识别[10]。然而，现有基于监督学习的方法普遍依赖大规模人工标注数据，当面临关系类别分布高度不平衡、特别是长尾关系场景时，其性能仍显著下降，泛化能力受限。

针对上述挑战，近年来研究者将关注重点逐步转向少样本关系三元组抽取任务，旨在利用有限的标注样本，仍能高效完成实体及其语义关系的识别[11]。与传统的监督学习方法类似，该任务通常可分解为少样本命名实体识别[12]和少样本关系抽取[13]两个子任务。早期方法多聚焦于其中之一，分别建模实体或关系信息，未能捕捉三元组内部的语义耦合结构。为此，研究者逐渐转向联合建模策略，提升实体与关系间的交互建模能力与整体抽取性能[4]。目前联合建模方法大致可分为三类主流范式，即先实体后关系、先关系后实体，以及实体关系双向引导。先实体后关系范式首先识别文本中的实体，再基于这些实体预测其间的语义关系。典型方法如 MPE [4]，其采用条件随机场进行实体识别，并结合原型网络完成关系分类。该类方法结构简洁、实现成本低，但在面对新颖或低频关系时易引入冗余实体，干扰关系判别的准确性。先关系后实体范式则以关系识别为起点，随后基于已识别关系引导相关实体的抽取。代表性方法如 RelATE [5]、PTN [14]和 RCTE [15]。此类方法有效缓解了先实体后关系范式中存在的实体语义不一致问题，但其建模过程仍为单向流程，实体与关系之间信息交互不足，且原型表示能力受限。为弥补上述两类方法的信息交互缺陷，近年来研究者提出了实体关系双向引导范式，通过构建实体与关系之间的相互引导机制，实现实体识别与关系分类的协同优化。代表方法如 TGIN [3]，其通过异构图结构与翻

译机制实现信息交互；MG-FTE [6]引入实体感知与关系生成模块进行相互引导；SQGE [16]则融合支持集与查询集的原型信息，结合多级对比学习与实体特征增强策略，在多个少样本数据集上取得了领先的性能表现。尽管上述方法在不同建模范式下均取得了一定进展，但在类别原型的语义表达准确性与原型空间分布的结构化建模方面仍存在显著局限，难以有效应对长尾关系场景下的关系三元组抽取挑战。

3. 方法

3.1. 整体框架

本文提出了一种面向中文文本的少样本关系三元组抽取方法，通过引入实体关系双向引导机制，实现实体识别与关系分类的协同优化。如图 1 所示，整个模型主要包含三个核心模块：(1) 实体关系双向引导模块，构建实体识别与关系分类之间的相互引导机制，通过关系引导实体识别、实体引导关系抽取以及双向注意力机制，实现支持集与查询集的语义对齐和两个子任务的协同学习；(2) 加权原型构建机制，针对传统 token 等权平均方式的局限性，引入基于 token 置信度的动态权重分配策略，通过评估每个 token 的语义贡献实现更精准的原型聚合；(3) 原型多样性约束损失，设计专门的结构化损失函数，解决原型空间中类内发散、类间区分度不足的问题，通过优化类内一致性与类间分离性，提升原型表示的判别能力。模型在支持集上学习原型表示，并在查询集上完成三元组抽取任务。

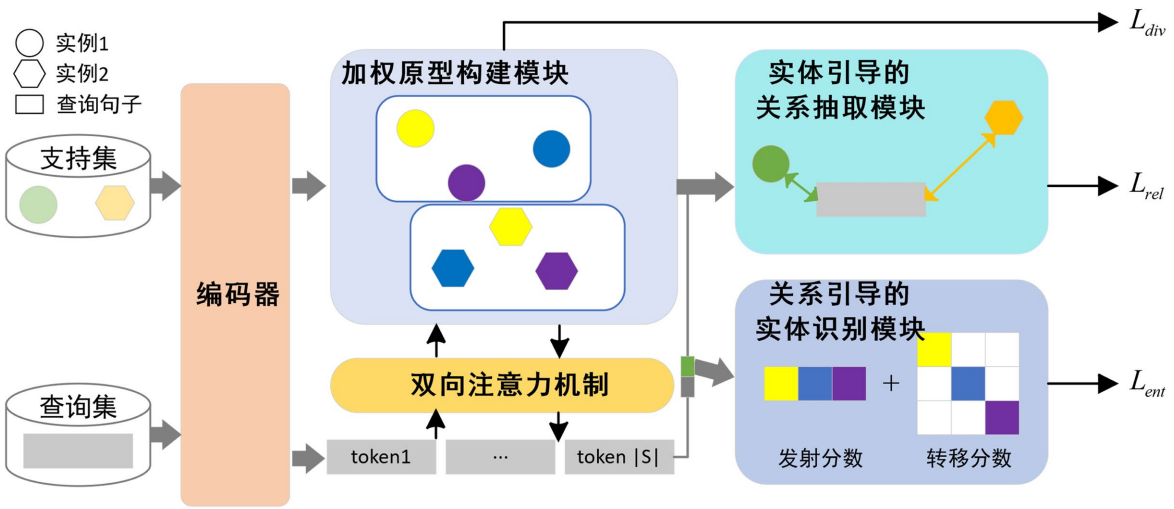

Figure 1. Framework of the proposed model under the 2-way-1-shot setting
图 1. 2-way-1-shot 设置下的模型框架图

3.2. 实体关系双向引导机制

实体关系双向引导机制是本文方法的核心基础，通过构建实体识别与关系分类之间的相互引导，实现两个子任务的协同优化[6]。与传统的单向建模方式不同，该机制建立了实体与关系信息的双向流动，有效缓解了少样本场景下实体与关系表示不充分的问题。

该机制主要包含三个关键组件。关系引导的实体识别利用已有的关系原型引导查询句中的实体抽取过程，通过将关系原型作为语义先验知识，帮助模型更准确地定位与特定关系相关的实体，从而在有限样本条件下提升实体识别的精度。实体引导的关系抽取基于识别出的实体信息反向优化关系原型表示，通过分析实体的语义特征和上下文信息，模型能够动态调整关系原型，使其更好地适应当前的语义场景，提升关系分类的准确性。双向注意力机制[17]实现支持集与查询集之间的语义对齐与原型优化，该机制通

过计算支持样本和查询样本之间的语义相似度，动态分配注意力权重，促进实体与关系信息的双向传递和相互增强。

通过这种双向引导的设计，模型能够在实体识别和关系分类两个子任务间建立密切的信息交互，克服了传统单向建模方式的局限性，为后续的原型构建和优化提供了更加丰富的语义基础。

3.3. 加权原型构建机制

传统的原型构建方法通常采用 token 等权平均的方式生成类别原型，这种方法忽略了不同 token 对类别语义贡献的差异性，容易受到冗余或无关信息的干扰，导致原型表示语义模糊，影响少样本条件下的抽取性能。为解决这一问题，本文提出基于 token 置信度的加权原型构建机制(Weighted Prototype Construction, WPC)，通过动态评估每个 token 的语义重要性，实现更精准的原型聚合。

本文采用基于 L2 范数的置信度计算方法，通过衡量 token 嵌入向量的表示强度来评估其语义贡献。对于属于同一实体标签的 token 集合，首先计算每个 token 嵌入向量的 L2 范数：

$$norms_i = \left\| h_i \right\|_2 \tag{1}$$

其中，h_i 表示第 i 个 token 的嵌入表示，$\left\| \cdot \right\|_2$ 表示 L2 范数计算。

随后，通过温度参数 τ 控制的 softmax 函数进行归一化，得到每个 token 的置信度权重：

$$w_i = \frac{\exp\left(norms_i / \tau\right)}{\sum_{j=1}^{|S|} \exp\left(norms_i / \tau\right)} \tag{2}$$

其中 w_i 表示第 i 个 token 的置信度权重，$\exp(\cdot)$表示指数函数，S 表示对应标签的 token 集合，$|S|$表示集合 S 的大小，j 表示集合中 token 的索引，τ 为温度系数，用于控制权重分布的平滑程度。

基于上述置信度权重 w_i，对该标签下所有 token 的嵌入表示进行加权求和，得到该实体类别在当前支持样本 i 中的原型表示：

$$P_i = \sum_{i=1}^{|S|} w_i \cdot h_i \tag{3}$$

其中 P_i 表示第 i 个支持样本对应的类别原型。

随后，将每个类别在 K 个支持样本中的原型取平均，作为该类别最终的原型表示：

$$P = \frac{1}{K} \sum_{s=1}^{K} P_i \tag{4}$$

通过这种加权聚合策略，模型能够自适应地关注对类别语义贡献更大的 token，从而在少样本条件下构建更加精准和鲁棒的类别原型表示。

3.4. 原型多样性约束损失设计

在少样本关系三元组抽取任务中，由于训练样本稀缺，原型空间容易出现类内原型发散与类间原型区分度不足的问题，导致模型在三元组匹配过程中边界模糊、预测不稳定。为此，本文设计了原型多样性约束损失(Prototype Diversity Loss, PDL)，通过同时优化类间分离度与类内一致性，提升原型表示的判别能力。

该损失函数包含两个关键组件：类间分离损失和类内聚合损失。对于每个批次中的原型张量，首先计算各类别的中心原型表示，然后分别计算类间和类内的约束损失。

类间分离损失旨在最大化不同类别原型之间的距离，增强类间区分度。对于批次中第 i 个样本，首先计算各类别的中心原型：

$$\bar{P}_{i,c}=\frac{1}{T}\sum_{t=1}^{T}P_{i,c,t} \tag{5}$$

其中，$P_{i,c,t}$ 表示第 i 个样本中第 c 个类别的第 t 个原型，T 表示每个类别的原型数量，$\bar{P}_{i,c}$ 表示第 c 个类别的中心原型。

然后计算类间距离矩阵，并通过负距离的形式构建类间分离损失：

$$L_{inter}=-\frac{1}{B}\sum_{i=1}^{B}\frac{1}{|M|}\sum_{(c_1,c_2)\in M}\left\|\bar{P}_{i,c_1}-\bar{P}_{i,c_2}\right\|_2 \tag{6}$$

其中，$\left\|\bar{P}_{i,c_1}-\bar{P}_{i,c_2}\right\|_2$ 表示类别 c_1 和 c_2 之间的欧氏距离，M 表示上三角掩码集合用于避免重复计算，$|M|$ 表示有效距离对的数量，B 表示批次大小。

类内聚合损失旨在最小化同一类别内不同原型之间的距离，增强类内一致性：

$$L_{intra}=\frac{1}{B\times N}\sum_{i=1}^{B}\sum_{c=1}^{N}\frac{1}{T}\sum_{t=1}^{T}\left\|P_{i,c,t}-\bar{P}_{i,c}\right\|_2 \tag{7}$$

其中，N 表示类别数量，$\left\|P_{i,c,t}-\bar{P}_{i,c}\right\|_2$ 表示第 t 个原型与其类别中心的欧氏距离。

最终，原型多样性约束损失函数为：

$$L_{div}=L_{inter}+L_{intra} \tag{8}$$

通过联合优化类间分离性与类内一致性，所提损失函数能够引导模型学习到更加紧凑且具有判别性的原型表示。类间分离损失确保不同类别的原型在特征空间中保持足够距离，而类内聚合损失则促使同类原型聚集在相近区域，从而显著提升了模型在少样本条件下的泛化能力与判别稳定性。

3.5. 目标函数

在完成各个模块的设计与损失函数的构建之后，本文进一步定义了模型的整体损失函数，以统一优化实体识别、关系分类以及原型表示学习三个核心任务。总体损失由三部分组成：关系分类损失 L_{rel}、实体识别损失 L_{ent} 以及原型多样性约束损失 L_{div}。

在关系分类方面，模型需要基于双向引导机制增强的特征表示，准确判断查询样本所属的关系类别。为此，采用交叉熵损失函数度量预测关系分布与真实标签之间的差异：

$$L_{rel}=-\frac{1}{B\times N\times Q}\sum_{i=1}^{B\times N\times Q}\sum_{j=1}^{\mathrm{N}}y_{i,j}^{rel}\log\left(\hat{p}_{i,j}^{rel}\right) \tag{9}$$

其中，$y_{i,j}^{rel}$ 表示第 i 个样本对第 j 个关系类别的真实标签，$\hat{p}_{i,j}^{rel}$ 表示模型预测第 i 个样本属于第 j 个关系类别的概率，Q 表示每个类别的查询样本数。该损失鼓励查询样本的表示与对应类别的原型具有更高的内积相似度，从而实现有效的关系预测。

在实体识别方面，由于实体边界判定具有序列依赖性，简单的分类损失难以捕捉标签间的转移约束。因此，采用条件随机场建模标签序列的全局最优路径，通过负对数似然损失优化：

$$L_{ent}=-\frac{1}{B\times N\times Q}\sum_{i=1}^{B\times N\times Q}\log\left(P\left(y_i^{ent}\mid x_i,\theta\right)\right) \tag{10}$$

其中，y_i^{ent} 表示第 i 个样本的真实实体标签序列，x_i 表示输入序列，θ 表示 CRF 参数，$P(y_i^{ent}\mid x_i,\theta)$ 表示给定输入序列下真实标签序列的条件概率。通过该损失函数，模型能够有效学习实体边界的识别能力。

最后，原型多样性约束损失 L_{div} 已在 3.4 节中详细阐述，通过联合优化类间分离性和类内一致性来增强原型表示的判别能力。

为实现实体识别、关系分类与原型表示优化的协同训练，本文设计了多任务联合优化的总损失函数：

$$L_{total} = L_{rel} + L_{ent} + \lambda_{div} \cdot L_{div} \tag{11}$$

其中，λ_{div} 是控制原型多样性损失权重的超参数，用于平衡不同任务之间的优化目标。该损失函数在整个训练过程中被最小化，使模型能够同时学习到具有强泛化能力的实体与关系表示，并确保支持集中的原型具备良好的判别性，从而在查询集中实现更准确的关系三元组抽取。

Table 1. Statistics of two datasets
表 1. 两个数据集的统计信息

类别	FewDuIE1.0	FewDuIE2.0
训练关系数	15	15
验证关系数	14	14
测试关系数	14	13
每个关系的句子数	400/50	500/50

4. 实验

4.1. 实验设置

4.1.1. 数据集

为了适配少样本关系三元组抽取任务，我们重新构建了两个数据集。表 1 展示了这些数据集的统计数据。需特别指出的是，训练集、验证集和测试集之间不存在类别重叠的情况。

FewDuIE1.0 数据集是基于公开数据集 DuIE [18]构建的一个子集，其原始语料来源于百度百科与百度新闻等中文文本资源。DuIE 已被广泛用于中文实体关系抽取任务的评估[19] [20]。由于原始数据中部分句子包含多个关系三元组，为避免模型混淆并确保任务设定的一致性，我们仅保留包含单一三元组的句子，最终获得覆盖 43 个关系类别的子集。考虑到各关系类别样本数量的不均衡性，我们选取其中句子数量少于 400 的 14 个关系类别作为测试集；再从其余样本数较多的 29 个关系类别中，随机划分出 15 个作为训练集，14 个作为验证集。在完成类别划分后，我们从各类别中随机抽取指定数量的样本，用于构建最终的数据集。其中，每类关系在训练集、验证集和测试集中分别包含 400、400 和 50 个实例。

在此基础上，我们进一步构建了更具挑战性的 FewDuIE2.0 数据集。该数据集基于 DuIE2.0 [21]构建，原始语料同样来源于百度百科、百度新闻等中文开放领域文本。相较于 DuIE，DuIE2.0 的语言风格更口语化，且在关系结构上更为复杂，进一步提升了任务的语言多样性与建模难度。我们采用与 FewDuIE1.0 相同的筛选策略，仅保留包含单一三元组的样本，最终得到覆盖 42 个关系类别的数据子集。根据同样的划分策略，我们将样本数量少于 500 的 13 个关系类别划为测试集，从其余 29 个类别中随机选取 15 个作为训练集，14 个作为验证集。随后从各类别中随机抽取指定数量的样本，构建最终数据集，其中每类关系在训练集、验证集和测试集中分别包含 500、500 和 50 个实例。

4.1.2. 基线

为评估模型效果，我们选取了几种典型的基线方法进行对比实验。

RelATE [5]：属于先关系后实体的抽取范式。该方法通过双重注意力机制生成关系原型，再以该原型为引导，分别识别对应的头实体与尾实体，从而缓解标签爆炸、实体差异等问题，适用于少样本场景下的关系三元组抽取。

PTN [14]：属于先关系后实体的抽取范式。该方法提出了一种基于视角迁移的框架，依次通过关系视角、实体视角和三元组视角进行推理，并在视角间循环迁移信息，最终实现联合三元组抽取。

MGFTE [6]：属于实体关系相互引导的抽取范式。该方法提出双原型解码器协同推理，并通过原型级融合模块增强实体与关系间的关联性，从而提升少样本场景下的抽取鲁棒性。

SQGE [16]：属于实体关系相互引导的抽取范式。该方法通过融合支持集与查询集的原型信息，结合多级对比学习和实体特征增强策略，有效缓解类内差异与类间混淆问题，实现少样本关系三元组抽取。

4.1.3. 实现细节

我们采用 Adam 优化器[22]对模型参数进行优化，其中预训练语言模型部分的学习率设置为 1×10^{-5}，其余模块的学习率为 1×10^{-3}，权重衰减系数为 1×10^{-4}。训练过程中使用 StepLR 学习率调度策略，每训练 1000 步将学习率衰减一次。训练最多执行 5000 步，若在每 500 步一次的连续三次验证评估中验证集性能无提升，则提前终止训练，以避免过拟合。

在任务构造方面，我们遵循 N-way K-shot 的少样本学习设定，分别在 FewDuIE1.0 和 FewDuIE2.0 数据集上构建 5-way 5-shot 与 3-way 3-shot 两类实验任务。具体而言，每个训练任务包含 $N \times K$ 个标注样本的支持集和 1 个样本的查询集。所有实验均在训练集上进行训练，通过验证集性能确定最优模型参数，并在测试集上报告最终性能。此外，在多任务损失函数中，温度系数 τ 设置为 1.2，权重平衡系数 λ_{div} 设置为 0.05。

本研究采用表现优异且应用广泛的中文预训练编码器 chinese-bert-wwm-ext [23]，该模型在本任务中经过对比实验验证具有良好的性能优势。

模型的评估方法沿用了以往研究的做法，使用 F1 分数、精确率和召回率来评估模型的性能。

所有实验在一台配备 Intel Xeon Silver 4310 CPU @ 2.10GHz、NVIDIA RTX 3090 GPU 及 256GB 内存的服务器上完成，操作系统为 Windows Server 2022 Standard。实验环境基于 PyTorch 1.10.0 深度学习框架，CUDA 版本为 11.1。

Table 2. Main results on FewDuIE1.0
表 2. FewDuIE1.0 的主要结果

模型	3-way 3shot			5-way 5shot		
	精确率	召回率	F1 分数	精确率	召回率	F1 分数
PTN [14]	43.40	43.40	43.40	40.80	40.80	40.80
RelATE [5]	48.30	52.73	50.42	47.46	50.32	48.85
MG-FTE [6]	64.05	61.46	62.43	63.57	61.51	62.43
SQGE [16]	31.47	19.06	23.74	34.69	17.68	23.42
Ours	66.46	65.21	65.65	66.18	65.29	65.64

4.2. 实验主要结果

表 2 总结了不同方法在 FewDuIE1.0 数据集上的性能表现。可以看出，在 FewDuIE1.0 数据集上，所提模型在两个任务设定下均显著优于现有主流方法，F1 分数分别达到 65.65% (3-way 3-shot)和 65.64% (5-way 5-shot)，分别比性能最优的对比方法 MG-FTE [6]提高了约 3.2%。此外，与结构较为简洁的单向建模方法 PTN [14]和 RelATE [5]相比，本文方法在 F1 指标上分别提高了超过 20%和 15%，进一步验证了实体关系双向引导机制在低资源条件下对信息交互与联合建模的增强效果。为更直观展示实验结果，图 2

绘制了各模型在 FewDuIE1.0 数据集上的 F1 分数柱状图，进一步凸显了本文方法的性能优势。

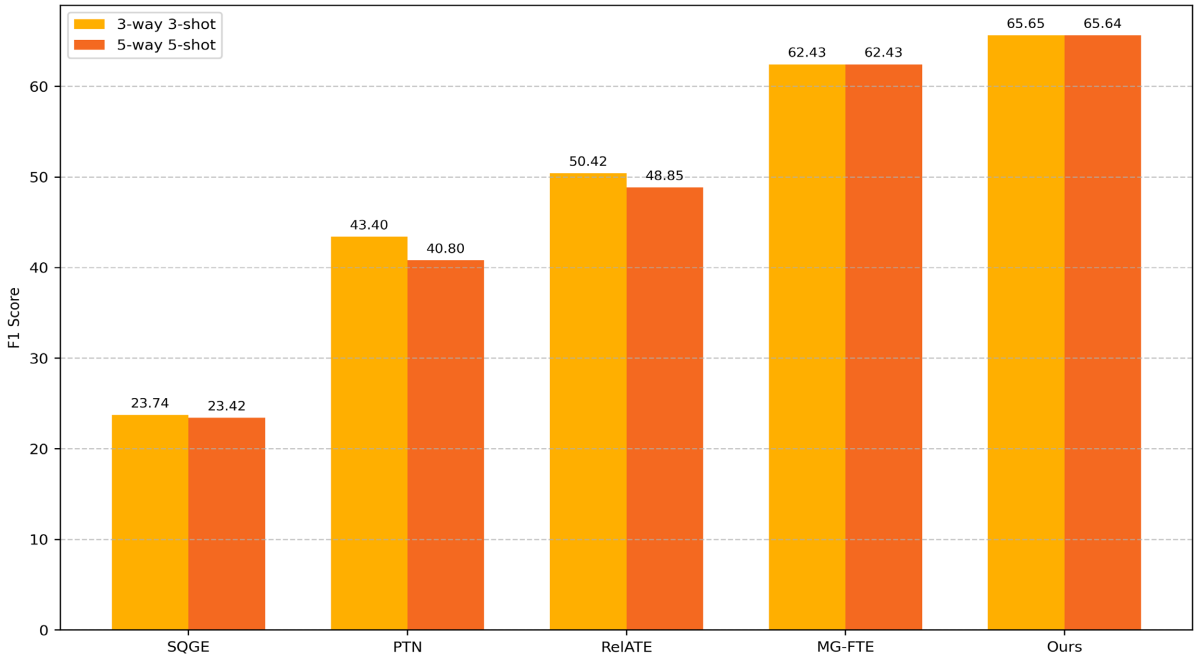

Figure 2. F1 score comparison of models on FewDuIE1.0
图 2. FewDuIE1.0 上模型 F1 分数对比

Table 3. Main results on FewDuIE2.0
表 3. FewDuIE2.0 的主要结果

模型	3-way 3shot			5-way 5shot		
	精确率	召回率	F1 分数	精确率	召回率	F1 分数
PTN [14]	28.60	28.60	28.60	29.40	29.40	29.40
RelATE [5]	42.43	45.43	43.89	40.84	43.81	42.27
MG-FTE [6]	59.60	60.47	59.82	59.82	60.46	60.03
SQGE [16]	17.56	14.09	15.63	15.90	11.27	13.19
Ours	62.63	62.15	62.18	62.76	62.39	62.46

表 3 展示了不同方法在 FewDuIE2.0 数据集上的实验结果。由于 FewDuIE2.0 的数据结构更为复杂且包含更多样化的语言现象，因此对模型的建模能力提出了更高的要求。尽管如此，本文提出的方法仍然在 3-way-3-shot 和 5-way-5-shot 任务中均表现出色，在两个任务中分别达到 62.18%和 62.46%的 F1 分数，同样超越所有对比方法。其中，与 MG-FTE [6]相比，F1 分数分别提升了 2.36%和 2.43%，显示出所提方法在面对更复杂语义结构时依然具备良好的泛化能力和稳定性。从图 3 的 F1 性能对比柱状图中可以看出，本文方法在 FewDuIE2.0 数据集上依然保持领先，充分说明其在复杂任务中的适应性与鲁棒性。

从精确率与召回率的角度来看，本文方法在两个数据集的不同设定下均取得了较为平衡的性能，表明其不仅能够较为准确地识别出正确的三元组，同时具备较强的覆盖能力。这种平衡性得益于本文所引入的加权原型构建机制与原型多样性约束损失，在提升原型表达判别力的同时，有效缓解了原型空间中类别间混淆与类别内发散的问题。

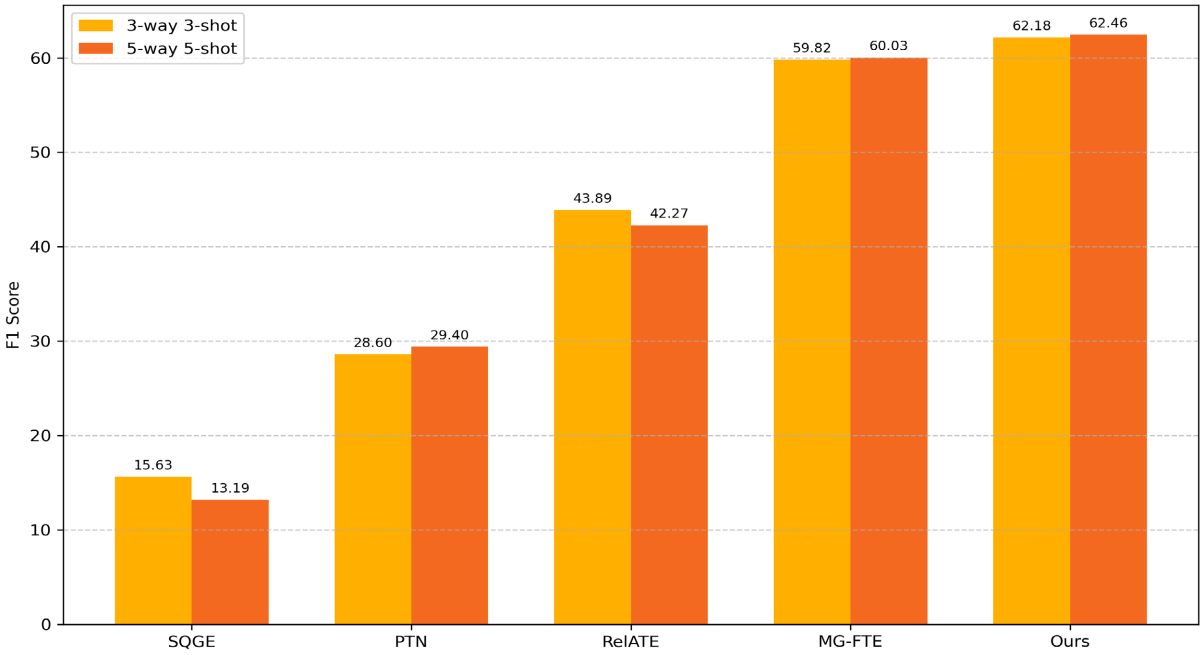

Figure 3. F1 score comparison of models on FewDuIE2.0
图 3. FewDuIE2.0 上模型 F1 分数对比

Table 4. Ablation studies on FewDuIE1.0
表 4. 在 FewDuIE1.0 上的消融研究

类别	3-way 3shot			5-way 5shot		
	精确率	召回率	F1 分数	精确率	召回率	F1 分数
ours	66.46	65.21	65.65	66.18	65.29	65.64
w/o PDL	65.83	64.22	64.83	65.27	64.54	64.81
w/o WPC	65.27	64.09	64.49	65.06	64.54	64.71

4.3. 消融实验

为进一步验证所提方法中关键模块的有效性，本文设计了消融实验与编码器对比实验，分析不同模块与编码器对模型性能的具体影响。

Table 5. Ablation studies on FewDuIE2.0
表 5. 在 FewDuIE2.0 上的消融研究

类别	3-way 3shot			5-way 5shot		
	精确率	召回率	F1 分数	精确率	召回率	F1 分数
ours	62.63	62.15	62.18	62.76	62.39	62.46
w/o PDL	61.32	61.49	61.19	61.78	61.61	61.59
w/o WPC	61.24	60.59	60.76	61.25	60.71	60.87

在消融实验中，我们分别去除模型中的加权原型构建机制和原型多样性约束损失，并在 FewDuIE1.0 与 FewDuIE2.0 两个数据集上进行对比实验，结果如表 4 与表 5 所示。结果表明，无论是去除加权原型构

建机制还是原型多样性约束损失，都会导致模型性能下降，说明两者对最终效果均具有积极贡献。其中，加权原型机制对 F1 分数提升更为显著，有助于缓解原型表达模糊的问题；而多样性约束则在优化原型空间结构、提升匹配边界清晰度方面发挥关键作用，两者共同促进了模型在少样本场景下的稳定性与泛化能力。

在编码器对比实验中，我们将所提模型分别搭配六种主流中文预训练编码器进行测试。这些编码器中，bert-base-chinese 由 Google 发布[2]，chinese-roberta-wwm-ext、chinese-macbert-base、chinese-bert-wwm 和 chinese-bert-wwm-ext 均由哈工大讯飞联合实验室发布[23]，ernie-3.0-base-chinese 则由百度发布[24]。我们在 FewDuIE1.0 数据集上开展了 3-way 3-shot 与 5-way 5-shot 两类实验任务，结果如表 6 所示。结果显示，chinese-bert-wwm-ext 编码器在两个任务设定中均获得最高的 F1 分数，显著优于其他编码器。该模型在原始 Whole Word Masking (WWM)策略基础上，采用了更大规模的语料和更丰富的训练细节优化，进一步提升了对中文词语级语义的建模能力，尤其在命名实体识别和上下文关系抽取任务中表现出更强的适应性和稳定性。相比之下，RoBERTa 和 ERNIE 等模型虽在部分指标上具备优势，但整体表现略有波动，泛化能力不足。因此，选用适配性更强的预训练编码器不仅有助于提升基础表示能力，也为少样本三元组抽取任务提供了更可靠的语义支撑。

Table 6. Performance comparison of different encoders on the FewDuIE1.0
表 6. 不同编码器在 FewDuIE1.0 上的性能比较

编码器	3-way 3shot			5-way 5shot		
	精确率	召回率	F1 分数	精确率	召回率	F1 分数
bert-base-chinese	63.90	58.58	60.86	63.43	58.74	60.87
chinese-roberta-wwm-ext	61.48	60.01	60.42	60.31	59.48	59.79
ernie-3.0-base-chinese	60.34	59.00	59.36	59.22	58.12	58.57
chinese-macbert-base	62.32	61.02	61.42	61.86	61.37	61.53
chinese-bert-wwm	62.61	58.86	60.35	61.95	58.65	60.14
chinese-bert-wwm-ext	64.05	61.46	62.43	63.57	61.51	62.43

5. 结论

本文针对中文少样本关系三元组抽取任务中类别原型表示模糊、实体与关系交互建模不充分等问题，提出了一种基于实体关系双向引导的新型抽取方法。该方法通过构建实体识别与关系分类之间的双向信息流，实现两个子任务的协同优化。进一步地，本文引入加权原型构建机制与原型多样性约束损失函数，分别提升原型语义表达的准确性与结构区分能力。实验结果显示，所提方法在 FewDuIE1.0 与 FewDuIE2.0 数据集上均优于主流基线，验证了其在少样本条件下的有效性与泛化能力。此外，消融实验证实了两个关键模块对性能提升具有显著贡献，编码器对比实验也进一步体现了中文预训练语言模型在该任务中的重要性。

基金项目

国家自然科学基金项目(62002225)；上海市自然科学基金项目(21ZR1445400)。

参考文献

[1] Yan, Y., Sun, H. and Liu, J. (2021) A Review and Outlook for Relation Extraction. *Proceedings of the 5th International Conference on Computer Science and Application Engineering*, Sanya, 19-21 October 2021, 1-5. https://doi.org/10.1145/3487075.3487103

[2] Devlin, J., Chang, M.-W., Lee, K. and Toutanova, K. (2019) BERT: Pre-Training of Deep Bidirectional Transformers

for Language Understanding. *Proceedings of the* 2019 *Conference of the North American Chapter of the Association for Computational Linguistics*, Minneapolis, June 2019, 4171-4186.

[3] Wang, J., Zhang, L., Liu, J., *et al.* (2022) TGIN: Translation-Based Graph Inference Network for Few-Shot Relational Triplet Extraction. *IEEE Transactions on Neural Networks and Learning Systems*, **35**, 9147-9161.

[4] Yu, H., Zhang, N., Deng, S., Ye, H., Zhang, W. and Chen, H. (2020) Bridging Text and Knowledge with Multi-Prototype Embedding for Few-Shot Relational Triple Extraction. *Proceedings of the* 28*th International Conference on Computational Linguistics*, Barcelona, 8-13 December 2020, 6399-6410. https://doi.org/10.18653/v1/2020.coling-main.563

[5] Cong, X., Sheng, J., Cui, S., Yu, B., Liu, T. and Wang, B. (2022) Relation-Guided Few-Shot Relational Triple Extraction. *Proceedings of the* 45*th International ACM SIGIR Conference on Research and Development in Information Retrieval*, Madrid, 11-15 July 2022, 2206-2213. https://doi.org/10.1145/3477495.3531831

[6] Yang, C., Jiang, S., He, B., Ma, C. and He, L. (2023) Mutually Guided Few-Shot Learning for Relational Triple Extraction. 2023-2023 *IEEE International Conference on Acoustics*, *Speech and Signal Processing (ICASSP*), Rhodes Island, 4-10 June 2023, 1-5. https://doi.org/10.1109/icassp49357.2023.10096232

[7] Chan, Y.S. and Roth, D. (2011) Exploiting Syntactico-Semantic Structures for Relation Extraction. *Proceedings of the* 49*th Annual Meeting of the Association for Computational Linguistics*: *Human LANGUAGE Technologies*, Stroudsburg, 19-24 June 2011, 551-560.

[8] Zhao, S., Hu, M., Cai, Z., Chen, H. and Liu, F. (2021) Dynamic Modeling Cross- and Self-Lattice Attention Network for Chinese Ner. *Proceedings of the AAAI Conference on Artificial Intelligence*, **35**, 14515-14523. https://doi.org/10.1609/aaai.v35i16.17706

[9] Lin, Y., Liu, Z. and Sun, M. (2017) Neural Relation Extraction with Multi-Lingual Attention. *Proceedings of the* 55*th Annual Meeting of the Association for Computational Linguistics*, Vancouver, July 30-August 4, 2017, 34-43. https://doi.org/10.18653/v1/p17-1004

[10] Zeng, X., Zeng, D., He, S., Liu, K. and Zhao, J. (2018) Extracting Relational Facts by an End-to-End Neural Model with Copy Mechanism. *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*, Melbourne, 15-20 July 2018, 506-514. https://doi.org/10.18653/v1/p18-1047

[11] Wang, Y., Yao, Q., Kwok, J.T. and Ni, L.M. (2020) Generalizing from a Few Examples: A Survey on Few-Shot Learning. *ACM Computing Surveys*, **53**, 1-34. https://doi.org/10.1145/3386252

[12] Feng, J., Xu, G., Wang, Q., Yang, Y. and Huang, L. (2024) Note the Hierarchy: Taxonomy-Guided Prototype for Few-Shot Named Entity Recognition. *Information Processing & Management*, **61**, Article 103557. https://doi.org/10.1016/j.ipm.2023.103557

[13] Han, J., Cheng, B., Wan, Z. and Lu, W. (2023) Towards Hard Few-Shot Relation Classification. *IEEE Transactions on Knowledge and Data Engineering*, **35**, 9476-9489. https://doi.org/10.1109/tkde.2023.3240851

[14] Fei, J., Zeng, W., Zhao, X., Li, X. and Xiao, W. (2022) Few-Shot Relational Triple Extraction with Perspective Transfer Network. *Proceedings of the* 31*st ACM International Conference on Information & Knowledge Management*, Atlanta, 17-21 October 2022, 488-498. https://doi.org/10.1145/3511808.3557323

[15] Liao, B., Lu, Z. and Guo, Y. (2024) RCTE: Relation Candidate-Guided Few-Shot Relational Triple Extraction. *Proceedings of the* 2024 3*rd International Conference on Cyber Security*, *Artificial Intelligence and Digital Economy*, Nanjing, 1-3 March 2024, 486-491. https://doi.org/10.1145/3672919.3673006

[16] Gao, C., Zhang, X., Jin, Z., Cai, W., Wang, D., Du, K., *et al.* (2025) SQGE: Support-Query Prototype Guidance and Enhancement for Few-Shot Relational Triple Extraction. *Neural Networks*, **185**, Article 107172. https://doi.org/10.1016/j.neunet.2025.107172

[17] Seo, M., Kembhavi, A., Farhadi, A. and Hajishirzi, H. (2017) Bidirectional Attention Flow for Machine Comprehension. 5*th International Conference on Learning Representations*, Toulon, 24-26 April 2017. https://openreview.net/forum?id=HJ0UKP9ge

[18] Li, S., He, W., Shi, Y., Jiang, W., Liang, H., Jiang, Y., *et al.* (2019) Duie: A Large-Scale Chinese Dataset for Information Extraction. In: Tang, J., Kan, M.Y., Zhao, D., Li, S. and Zan, H., Eds., *Lecture Notes in Computer Science*, Springer International Publishing, 791-800. https://doi.org/10.1007/978-3-030-32236-6_72

[19] 张龙辉, 尹淑娟, 任飞亮, 等. BSLRel: 基于二元序列标注的级联关系三元组抽取模型[J]. 中文信息学报, 2021, 35(6): 74-84.

[20] He, Y., Li, Z. and Dou, Z. (2025) Chinese Relation Extraction for Constructing Satellite Frequency and Orbit Knowledge Graph: A Survey. *Digital Communications and Networks*. In Press. https://doi.org/10.1016/j.dcan.2025.05.002

[21] 百度. DuIE2.0 中文关系抽取数据集[EB/OL]. https://aistudio.baidu.com/datasetdetail/180082, 2025-07-01.

[22] Kingma, D.P. (2014) Adam: A Method for Stochastic Optimization. arXiv:1412.6980.

[23] Cui, Y., Che, W., Liu, T., Qin, B. and Yang, Z. (2021) Pre-Training with Whole Word Masking for Chinese Bert. *IEEE*/*ACM*

Transactions on Audio, *Speech*, *and Language Processing*, **29**, 3504-3514. https://doi.org/10.1109/taslp.2021.3124365

[24] Sun, Y., Wang, S., Feng, S., *et al*. (2021) Ernie 3.0: Large-Scale Knowledge Enhanced Pre-Training for Language Understanding and Generation. arXiv:2107.02137.