LGC-CR: Few-shot Knowledge Graph Completion with Local Global Contrastive Learning and LLM Guided Refinement

Yiming Xu, Qi Song, Yihan Wang, Wangqiu Zhou, Junli Liang

Lab for Intelligent Networking and Knowledge Engineering
School of Computer Science, University of Science and Technology of China

Background

In real-world knowledge graphs, many entities/relations show a long-tail distribution (with very few triples), which is defined as **few-shot problem**.

Long-tail distribution(few-shot KG)

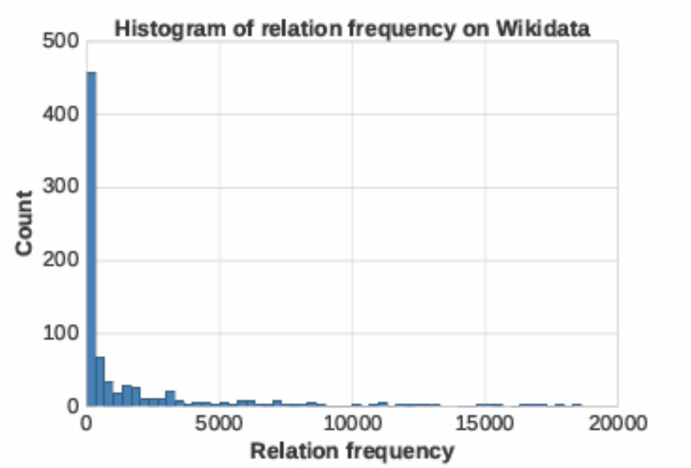
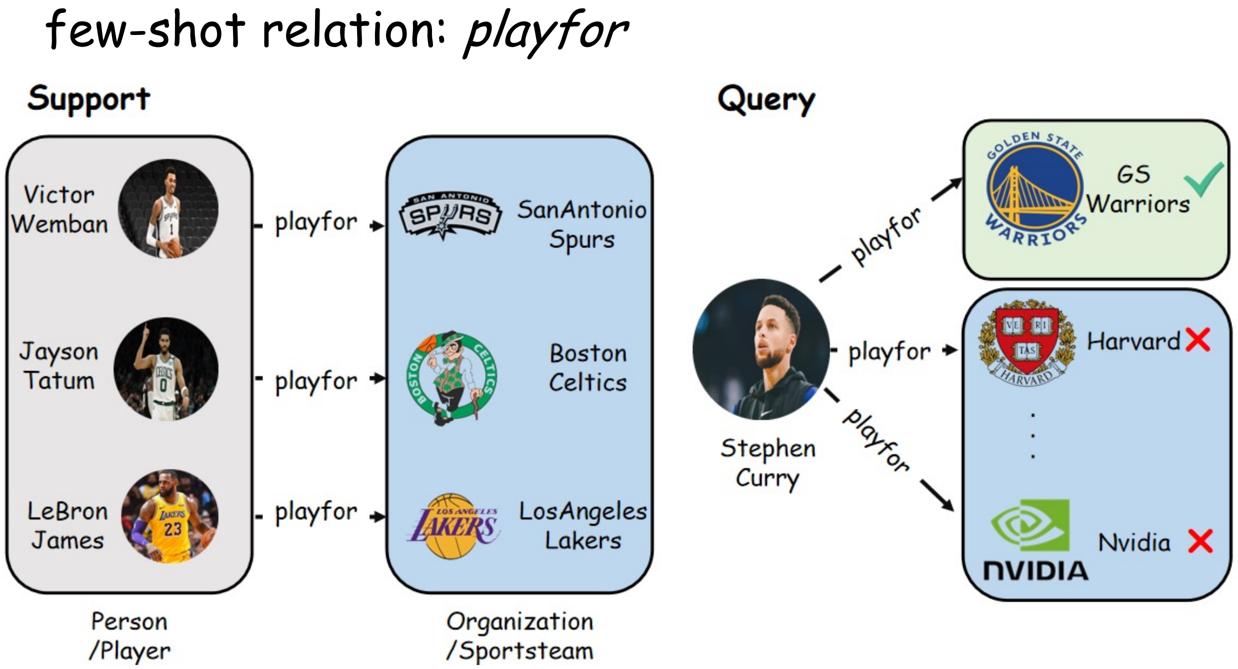


Figure 1: The histogram of relation frequencies in Wikidata. There are a large portion of relations that only have a few triples.

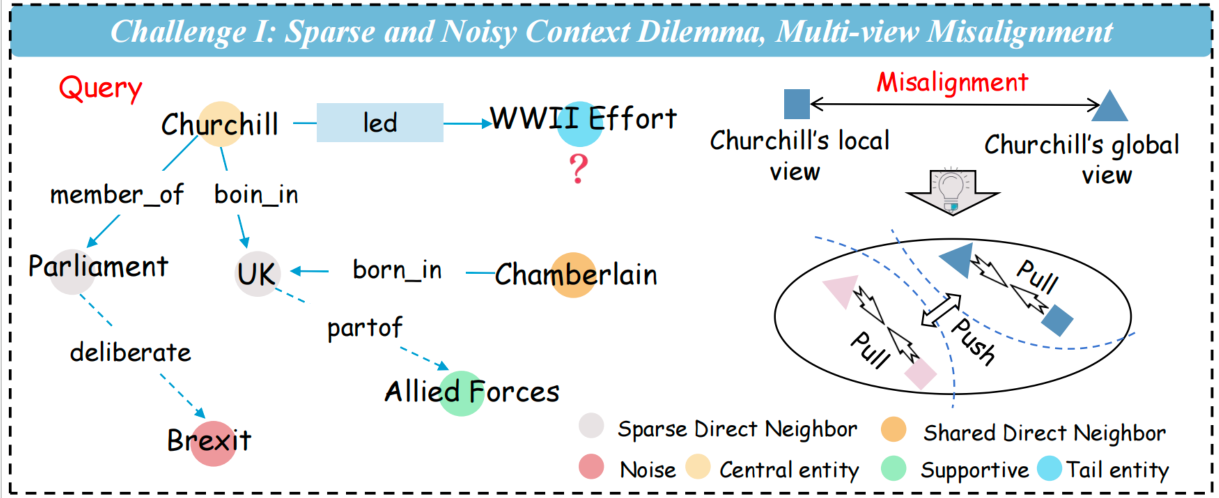
[Xiong, Wenhan and et al. One-shot relational learning for knowledge graphs. EMNLP 2018]

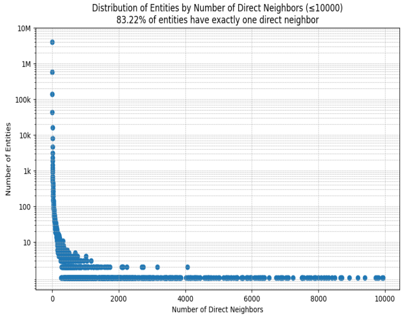
How to effectively achieve knowledge graph completion under few-shot conditions?



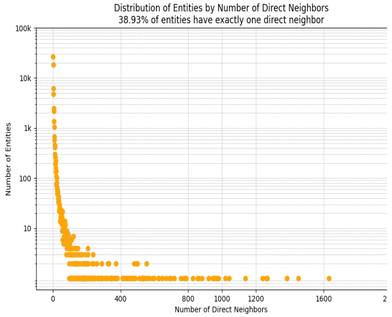
Meta Learning Framework

Motivation

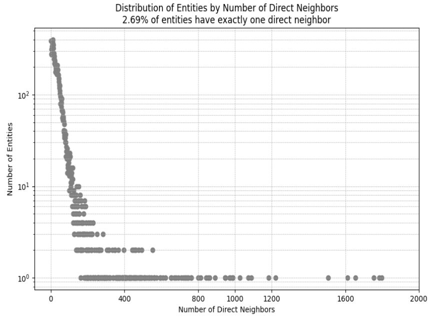




Wikidata



NELL



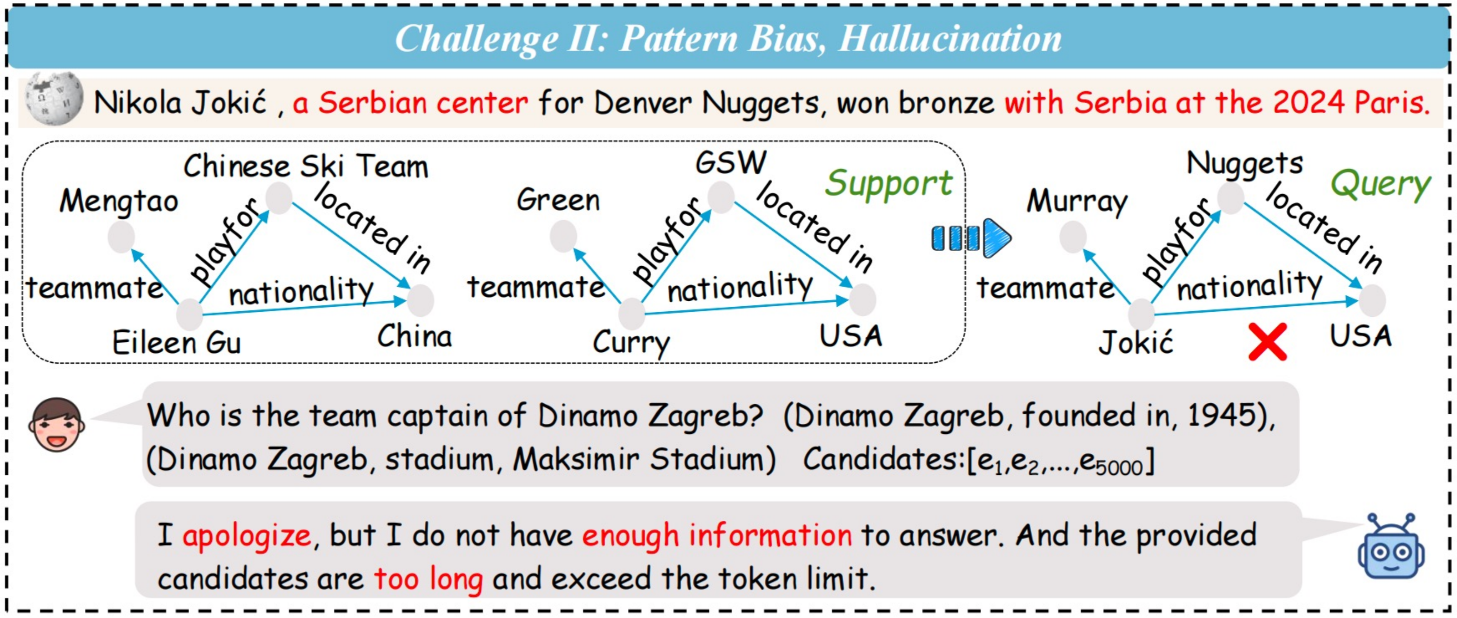
FB15K

Sparse One-hop Neighborhoods!

1. Existing methods over-rely on local neighbors to enhance entity embedding, which struggle to **leverage noisy high-order neighborhoods** and **align multi-view neighbor information**.

Need to Enrich and Align Neighbor Information.

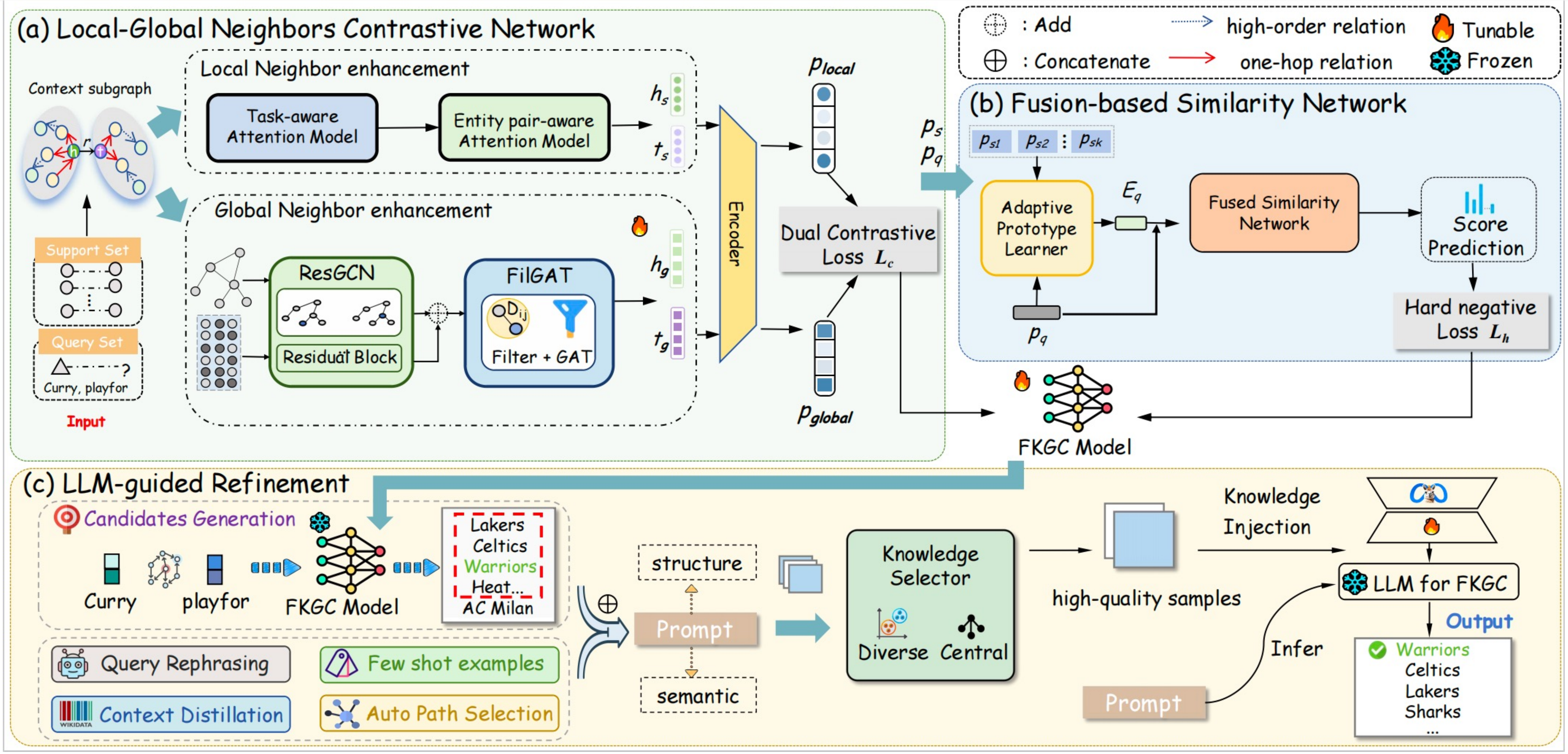
Motivation



2. Embedding-based meta-learning may overfit to **frequent relational patterns** in the support set, while LLMs suffer from **hallucination and input constraints**.

Need more knowledge to Refine the bias.

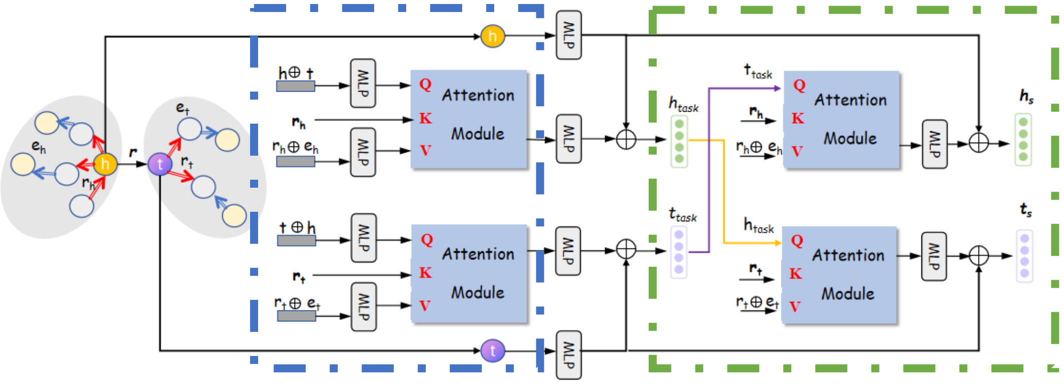
Framework



(a) Local-Global Neighbors Contrastive Network (b) Fusion-based Similarity Network
(c) LLM-guided Refinement

Local-Global Contrastive Network

Cross-Attention Local Enhancement.

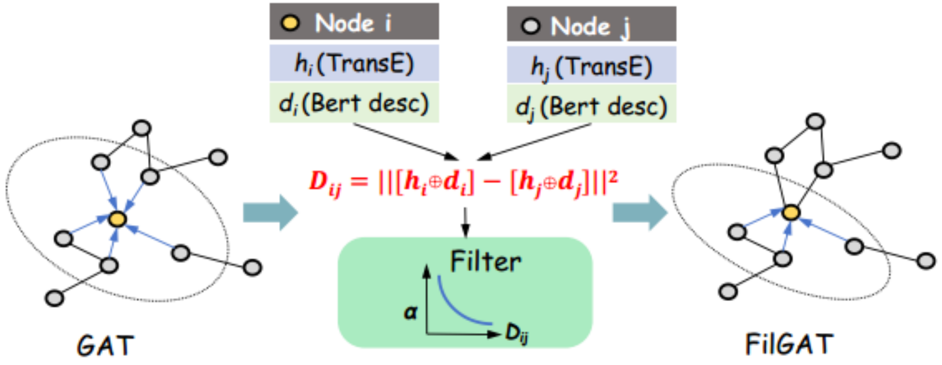


(1)Task-aware (2)Entity pair-aware

(1) Neighbors more task-relevant receive higher weights.

(2) Neighbors more similar to paired entities receive higher weights.

ResGCN-FilGAT Global Enhancement.



ResGCN $$H^{(1)} = \mathrm{ReLU}\left(\tilde{D}^{-\frac{1}{2}}\tilde{A}\tilde{D}^{-\frac{1}{2}}XW^{(0)}\right) + W^r X,$$

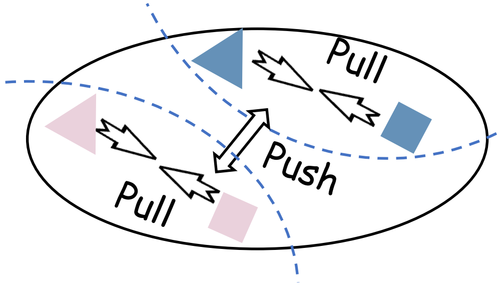
one-hop: stable and comprehensive

FilGAT $$\alpha_{ij} = \frac{o_{ij} \cdot \mathrm{sigmoid}(-\beta D_{ij})}{\sum_{k \in N_i} o_{ik} \cdot \mathrm{sigmoid}(-\beta D_{ik})},$$

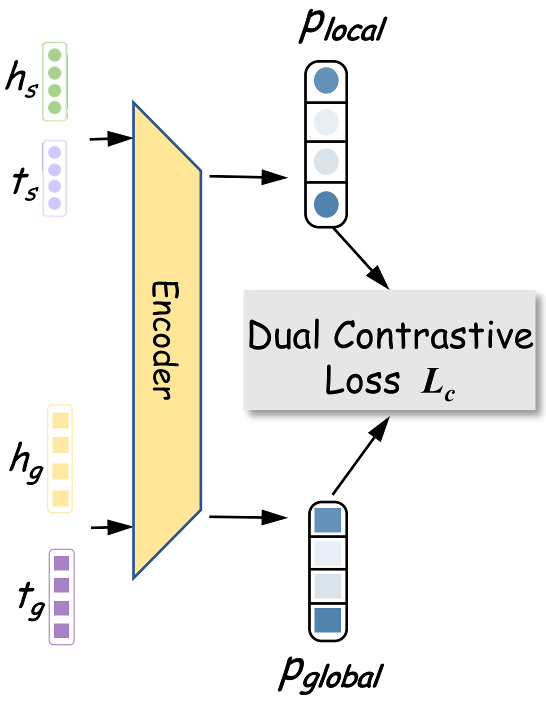
$$H_i^{(2)} = \sum_{j \in N(i)} \alpha_{ij} W^l H_j^{(1)},$$

high-order: Not all neighbors matter: adaptive filtering based on structure-semantic dissimilarity.

Local-Global Contrastive Network



- Local and global neighbors from different views may generate embeddings with **certain gaps**.
- Optimizing from a single perspective (local or global only) may **overfits to noise in the current view**.
- **Maximize local–global neighborhood consistency** for the same entity pair.
- **Interactive alignment**.



□ Project to the same semantic space

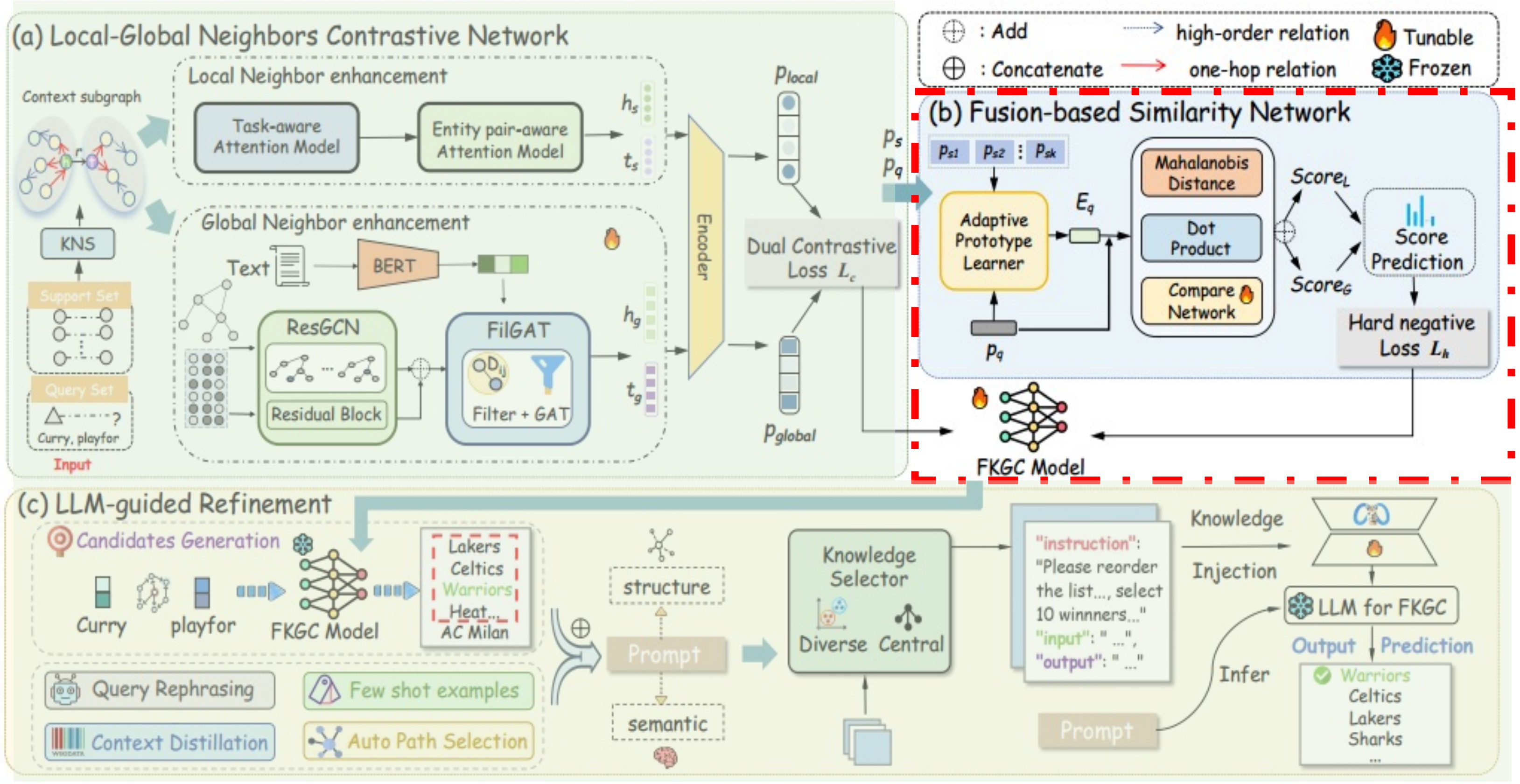
□ Dual Neighbor Contrastive Loss

$$L_{local} = \frac{1}{|Q_r|} \sum_{q \in Q_r} -\log \frac{\exp(sim(p_{local}^{q}, p_{global}^{q})/\tau)}{\sum_i \exp(sim(p_{local}^{q}, p_{global}^{q_i^-})/\tau)}$$

$$L_{global} = \frac{1}{|Q_r|} \sum_{q \in Q_r} -\log \frac{\exp(sim(p_{global}^{q}, p_{local}^{q})/\tau)}{\sum_i \exp(sim(p_{global}^{q}, p_{local}^{q_i^-})/\tau)}$$

$$L_{cl}$$

Fusion based Similiarity Network

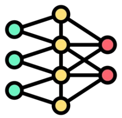


- A fusion similarity network integrates three **complementary similarity metrics**, which constructs a **learnable non-linear** metric space.

- Hard Negative Loss
Instead of treat all negative samples equally, we **focus more on hard negative samples** that are difficult to distinguish during training.

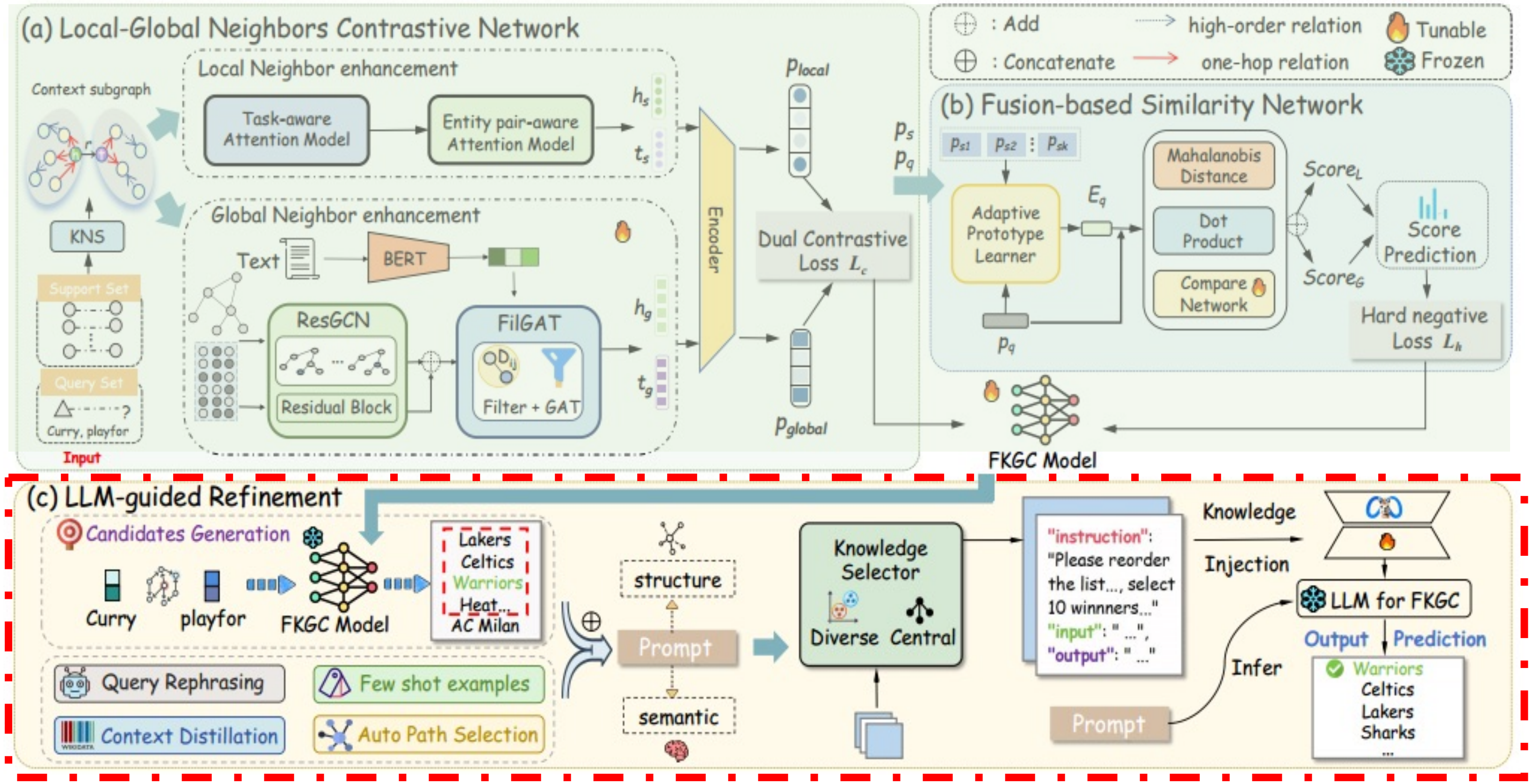
Dual Contrastive Loss

Hard Negative Loss



Optimize FKGC Model

LLM-guided Refinement



- **Instruction Retrieval**

Retrieve query-relevant **structural and semantic knowledge**.

- **Knowledge Selector**

Diversity-aware Sampling: selects **representative relations** covering diverse KG information.

Centrality-aware Sampling: focusing on **high-centrality relations** while avoiding noise.

> Prompt Template

You are an expert in refining initial meta-learning outputs. Please refine the list of tail entities based on the given information, select 10 winnners and output as ent1 | ent2...

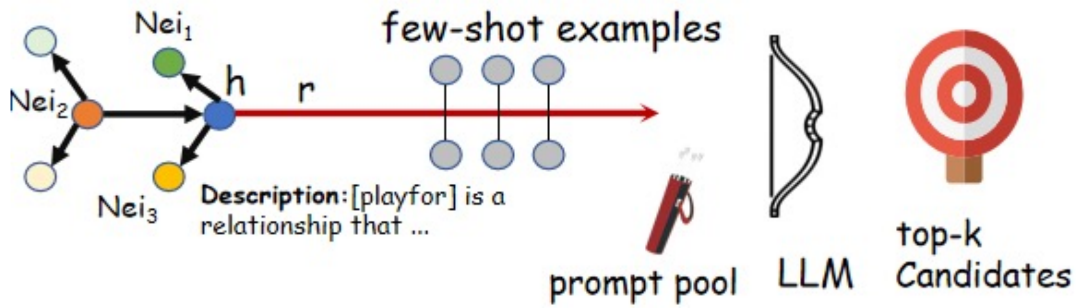
Query: Which team does Curry play for?
What team is Curry currently on?

few-shot examples:
Lebron James, playfor, Lakers...

Entity Des: American basketball....
Relation Des: [playfor] is a relation...

Path triples:
Draymond Green, teammates, Curry
Curry, team_location, San Francisco, venue_of, Chase

top-m entities:
Lakers, Celtics, Warriors, Heat, ...



High quality Facts

Train LLM to Learn Refining efficiently!

Datasets and Settings

Datasets

Datasets	*Entities(E)*	*Triples*	*Relations(R)*	*Task(Train/val/Test)*	*Avg.Description Length (E/R)*
NELL-One	68,545	181,109	358	67(51/5/11)	30.5/10.2
Wiki-One	4,838,244	5,859,240	822	183(133/16/34)	24.7/8.4
FB15K-One	14,478	309,621	237	45(32/8/5)	19.6/7.1

Implementation Details

platform	4 NVIDIA RTX A800 GPUs
base model	Deepseek-R1-Distill-Qwen-14B
fine-tuning approach	LoRA

Experiments

Model	Nell-One MRR 3-shot	Nell-One MRR 5-shot	Nell-One Hit@5 3-shot	Nell-One Hit@5 5-shot	Nell-One Hit@1 3-shot	Nell-One Hit@1 5-shot	Wiki-One MRR 3-shot	Wiki-One MRR 5-shot	Wiki-One Hit@5 3-shot	Wiki-One Hit@5 5-shot	Wiki-One Hit@1 3-shot	Wiki-One Hit@1 5-shot	FB15K-One MRR 5-shot	FB15K-One Hit@5 5-shot	FB15K-One Hit@1 5-shot
GMatching(2018)[37]	-	0.176	-	0.233	-	0.113	-	0.263	-	0.337	-	0.197	0.189	0.274	0.101
MetaR (2019)[4]	0.245	0.261	0.360	0.350	0.144	0.168	0.317	0.323	0.379	0.385	0.261	0.270	0.203	0.291	0.107
FSRL(2020)[40]	0.219	0.195	0.296	0.279	0.139	0.108	0.102	0.113	0.131	0.135	0.050	0.056	0.223	0.364	0.102
FAAN(2020)[30]	0.247	0.279	0.309	0.364	0.183	0.200	0.298	0.341	0.368	0.395	0.228	0.281	0.259	0.424	0.178
GANA(2021)[26]	0.322	0.344	0.432*	0.437	0.225	0.246	0.331	0.351	0.389	0.407	0.283	0.299	0.209	0.334	0.107
CIAN(2022)[19]	0.344*	0.373*	0.417	0.453	0.266*	0.294*	0.358	0.383	0.438	0.453	0.284	0.318	0.248	0.426	0.144
TransAM(2023)[23]	0.235	0.263	0.361	0.311	0.175	0.205	0.315	0.330	0.345	0.405	0.273	0.258	-	-	-
APINet(2023)[20]	0.305	0.318	0.405	0.412	0.208	0.225	0.342	0.347	0.419	0.428	0.283	0.297	-	-	-
SuperRL(2024)[12]	0.312	0.330	0.404	0.441	0.223	0.234	0.359*	0.388*	0.446*	0.458*	0.297*	0.320*	-	-	-
MVSE(2024)[25]	0.302	0.332	0.405	0.465*	0.178	0.211	0.349	0.351	0.382	0.422	0.295	0.296	0.285*	0.458*	0.195*
HNII(2024)[21]	-	0.365	-	0.453	-	0.283	-	0.352	-	0.457	-	0.315	-	-	-
NFAA(2025)[8]	0.303	0.332	0.368	0.409	0.233	0.263	0.332	0.341	0.384	0.416	0.285	0.282	-	-	-
LGC-FKGC	0.382	0.390	0.467	0.488	0.304	0.311	0.362	0.377	0.458	0.473	0.301	0.304	0.310	0.469	0.208
LGC+DeepSeek-R1	-	0.392	-	0.473	-	0.322	-	0.486	-	0.525	-	0.459	0.396	0.518	0.300
LGC-CR	**0.428**	**0.440**	**0.515**	**0.523**	**0.342**	**0.375**	**0.544**	**0.567**	**0.573**	**0.606**	**0.511**	**0.537**	**0.474**	**0.566**	**0.401**

Best results are **bolded**, runner-up results are underlined, and * indicates the SOTA baseline metrics. Missing metrics are due to unavailability in the original papers.

- ✓ LGC-CR outperforms the state-of-the-art methods on three benchmark datasets.
  NELL: **8.1%** Hit@1 Wiki: **21.7%** Hit@1 FB15K: **20.6%** Hit@1
- ✓ LGC-FKGC is also competitive, particularly under 3-shot setting.
  Stable and denoised global neighbor information compensate for lack of support set.
- ✓ With the addition of LLM refinement, our method significantly outperforms embedding-based meta-learning.

Experiments-Ablation Study

Table 3: Ablation study results on Nell-One and Wiki-One.

Models	Nell-One		Wiki-One	
	MRR	Hit@1	MRR	Hit@1
Full Model	**0.440**	**0.375**	**0.567**	**0.537**
w/o. KS	0.420	0.356	0.538	0.508
w/o. KS&sft	0.392	0.322	0.485	0.459
w/o. CR	0.390	0.311	0.377	0.304
w/o. CR&FS	0.384	0.302	0.364	0.295
w/o. CR&DCL	0.376	0.286	0.359	0.288
w/o. CR&LGC	0.353	0.269	0.342	0.280
LLM-only(shuffle)	0.267	0.282	0.336	0.269
LLM-only(generate)	0.110	0.135	0.110	0.226

1. Knowledge selector **selects high diversity, high-centrality facts** helps reduce noise from marginal relations.

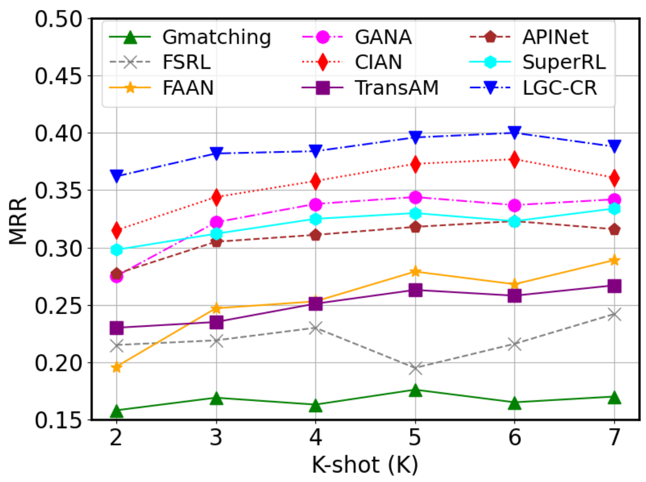
2. **Knowledge injection** through fine-tuning enhances LLM performance and contributes to better refinement results.

3. Dual contrastive learning show its effectiveness in **aligning local-global neighbors** and enhancing model robustness.

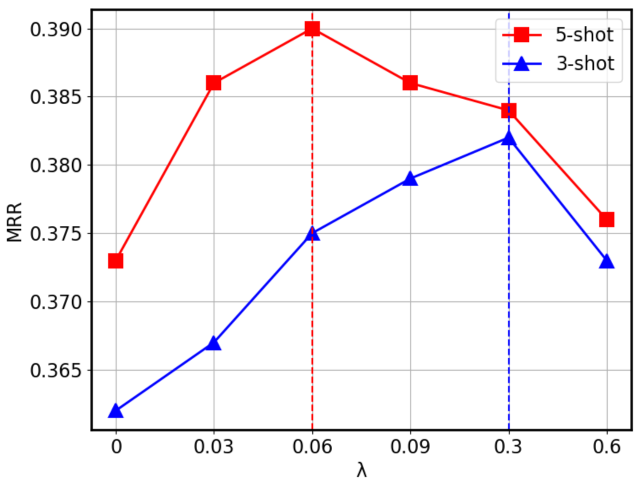
4. Relying solely on a single metric can **introduce similarity bias.**

5. Rich **high-order neighbor information and noise filtering** are crucial to model effectiveness.

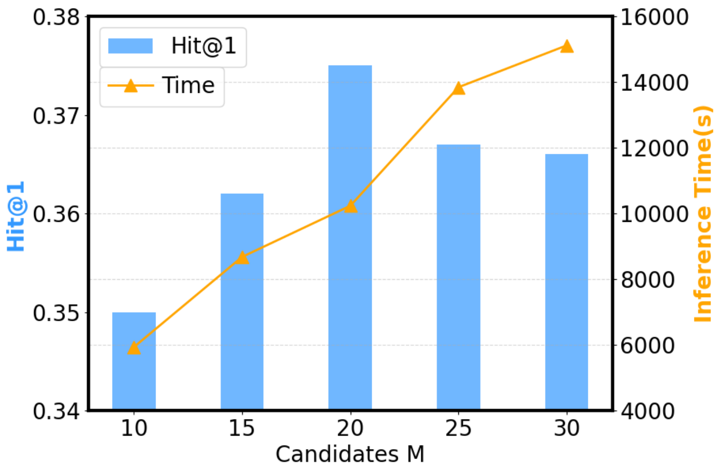
Experiments-Parameter Analysis



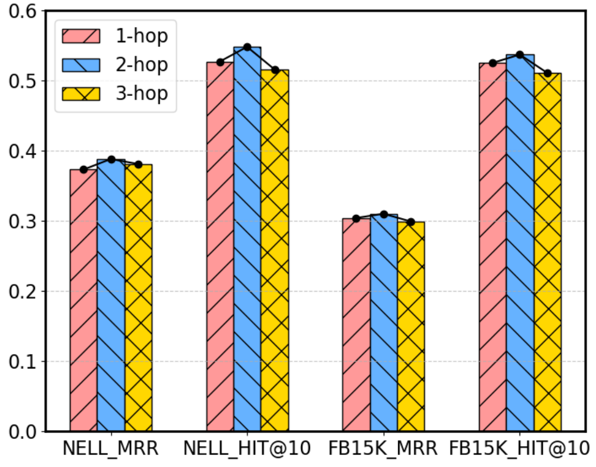
Impact of few-shot size K on NELL dataset



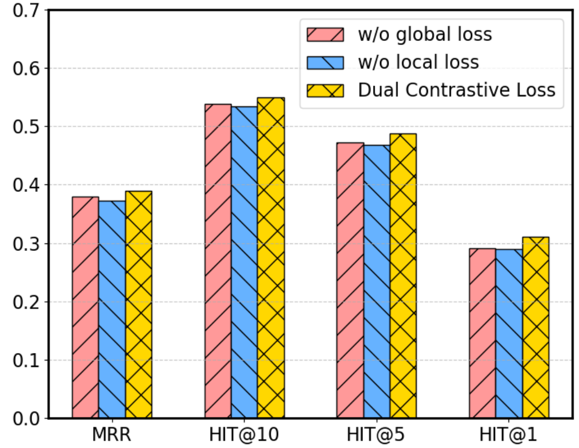
Impact of hyperparameter λ



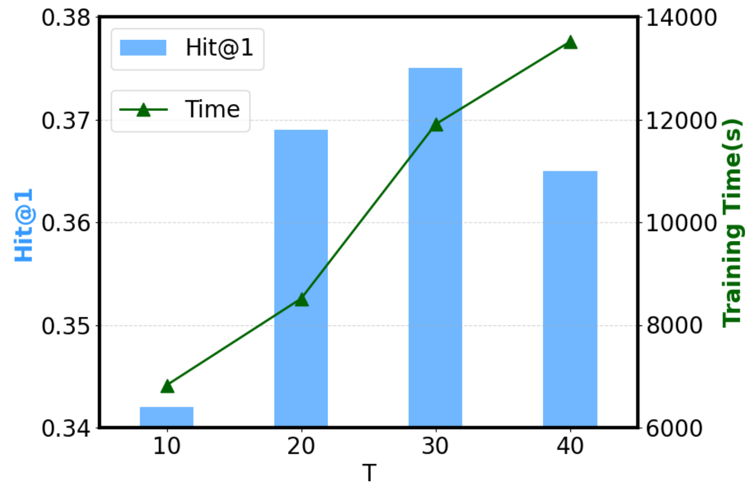
Impact of number of candidates



Impact of neighbor hop



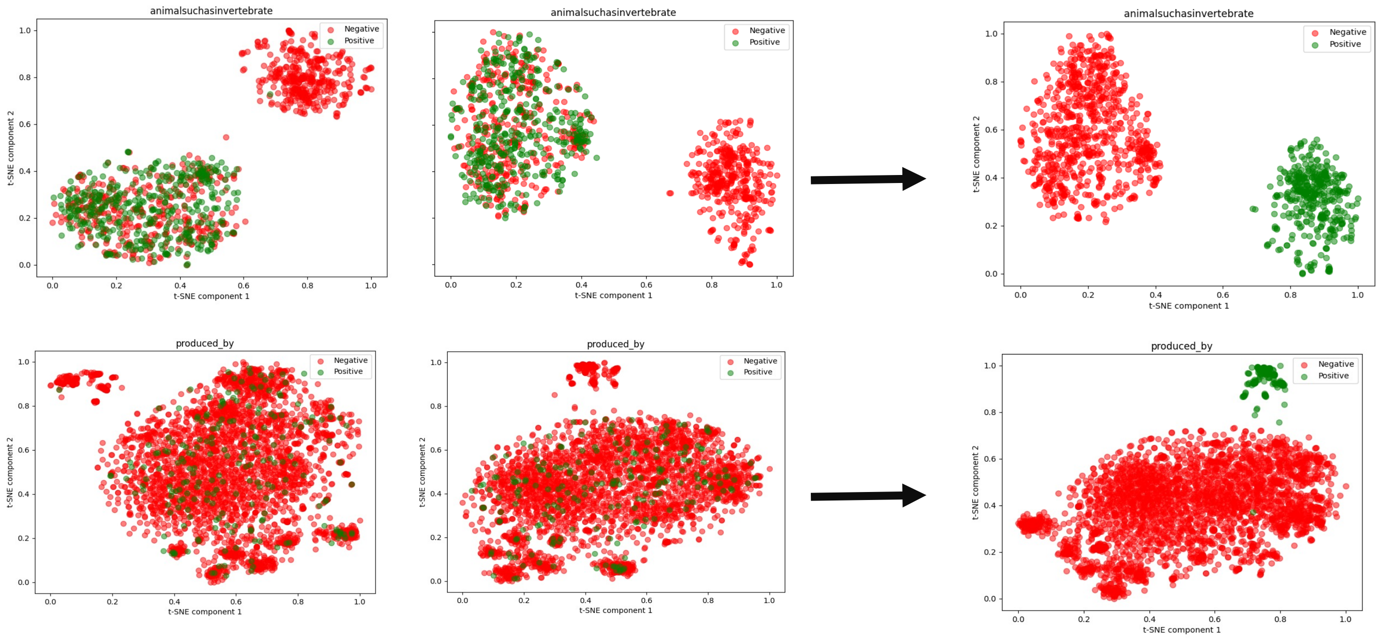
Impact of Contrastive Learning Strategies



Impact of Knowledge Selector ratio

Experiment-Case Study

- Our LLM-Guided Refinement integrates **textual evidence and factual knowledge** to promote the correct answer to the top, effectively **mitigating the pattern bias** of meta-learning methods.



CIAN(2022) MVSE(2024) Ours

Table 5: Case study of Top-5 candidate rankings for two relation types. Correct entities are in bold.

Query Task	LGC-FKGC	LGC-CR
(dallas_mavericks, teamcoach, ?) *True:* coach:jason_kidd	phil_jackson gregg_popovich scott_skiles rick_carlisle **jason_kidd**	**jason_kidd** rick_carlisle gregg_popovich scott_skiles marc_iavaroni
(possible_agrippina_major, location_of_discovery, ?) *True:* béziers	siselen müllheim **béziers** labastide_marnhac isarrnig	**béziers** siselen müllheim labastide_marnhac river_ness

- Our model excels at producing highly **distinctive and well-separated** represen- -tations for different types of entity pairs.

Conclusion

01 02 03 04

Conclusion

- We introduced LGC-CR, a novel few-shot knowledge graph completion framework that integrates an *enhanced meta-learning model with LLM knowledge.*
- At the data level-to address Challenge 1, we design a *local-global contrastive network* to enrich and align neighborhood information.
- At the model level-to address Challenge 2, we introduce *LLM refinement*, which retrieves relevant contexts and selects diverse, knowledge-centric facts to fine-tune LLMs for optimizing meta learning results.
- Experiments on three benchmark datasets demonstrate that LGC-CR outperforms the state-of-the-art baselines by **8.1-20.6%** in terms of Hit@1 and shows robust performance in few-shot scenarios.

Future Work : - Inject long-tail knowledge into LLMs in a more effective manner.
- More modalities knowledge (such as image) could be considered.

Thanks

xym30 @ mail.ustc.edu.cn

Yiming Xu would like to **thank** all the collaborators and the team members.