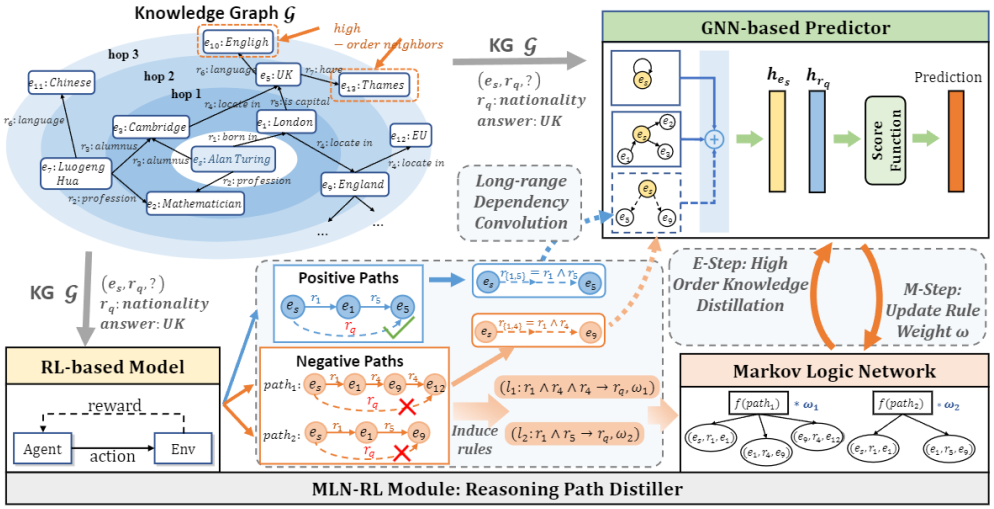
Exploring & exploiting high-order graph structure for sparse knowledge graph completion

Tao HE, Ming LIU, Yixin CAO, Zekun WANG, Zihao ZHENG, Bing QIN

Frontiers of Computer Science, DOI: 10.1007/s11704-023-3521-y

Problems & Ideas

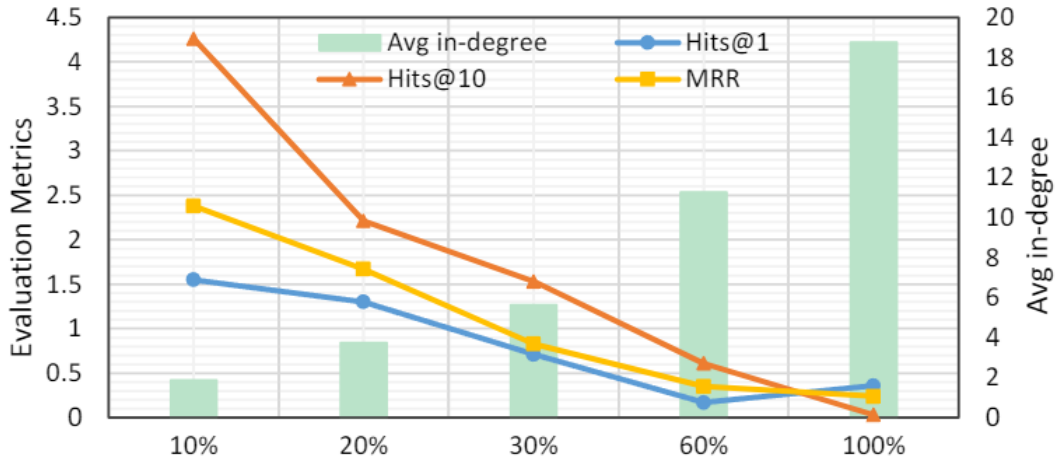
- Challenges of sparse Knowledge Graph Completion:
  - KGs are too sparse to produce sufficient neighborhoods to depict entities features.
  - Conventional GNN-based KGC methods can not filter noisy and irrelevant neighborhoods during aggregating messages.
- Ideas: Apply Reinforcement Learning to search valuable reasoning paths and Markov Logic Network to distill logical reasoning knowledge into GNN-based methods.

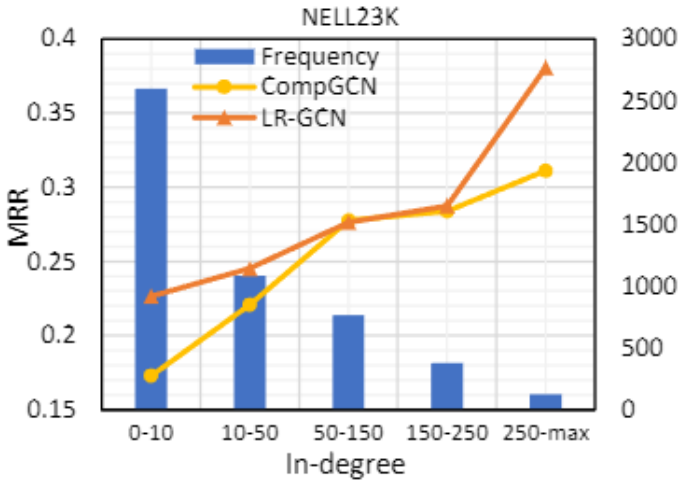


Our framework consists of two modules: GNN-based model and MLN-RL model. Given a query (es, rq, ?), MLN-RL first reasons paths and classify the paths into two parts according to whether the path is correct. The positive paths and negative path segments are applied to construct new edges to capture long range dependency explicitly. While the negative paths are fed into MLN to distill knowledge for the GNN-based model by variational EM

Main Contributions

- Contributions:
  - A novel GNN-based framework, LR-GCN, explores and exploits high-order graph structures to relieve the challenge of sparse KGC;
  - A novel path-based method, MLN-RL, which generates reasoning paths with delicately calibrated rule weights. This approach effectively filters out noisy paths and explores more instructive high-order graph structures;
  - Two strategies exploit the mined high-order graph structure information explicitly and implicitly.





Analysis results on FB15K-237 and NELL23K. Left: Improvements of LR-GCN on FB15K-237 and 4 sparse datasets against to CompGCN; Right: MRR results and entity frequency grouped by entity in-degree on NELL23K.