MINES: Message Intercommunication for Inductive Relation Reasoning Over Neighbor-Enhanced Subgraphs

Ke Liang[1], Lingyuan Meng[1], Sihang Zhou[2*], Wenxuan Tu[1], Siwei Wang[3], Yue Liu[1], Meng Liu[1], Long Zhao[4], Xiangjun Dong[4], Xinwang Liu[1*]

[1]School of Computer, National University of Defense Technology, Changsha, China
[2]School of Intelligence Science and Technology, National University of Defense Technology, Changsha, China
[3]Intelligent Game and Decision Lab, Beijing, China
[4]Qilu University of Technology, Jinan, China

Abstract

GraIL and its variants have shown their promising capacities for inductive relation reasoning on knowledge graphs. However, the uni-directional message-passing mechanism hinders such models from exploiting hidden mutual relations between entities in directed graphs. Besides, the enclosing subgraph extraction in most GraIL-based models restricts the model from extracting enough discriminative information for reasoning. Consequently, the expressive ability of these models is limited. To address the problems, we propose a novel GraIL-based framework, termed MINES, by introducing a Message Intercommunication mechanism on the Neighbor-Enhanced Subgraph. Concretely, the message intercommunication mechanism is designed to capture the omitted hidden mutual information. It introduces bi-directed information interactions between connected entities by inserting an undirected/bi-directed GCN layer between uni-directed RGCN layers. Moreover, inspired by the success of involving more neighbors in other graph-based tasks, we extend the neighborhood area beyond the enclosing subgraph to enhance the information collection for inductive relation reasoning. Extensive experiments prove the promising capacity of the proposed MINES from various aspects, especially for the superiority, effectiveness, and transfer ability.

Introduction

Knowledge graphs (KGs) organize human knowledge in the form of the relational fact triplet. Each triplet consists of a head entity, a tail entity, and a relational edge between them. Recently, many applications have been developed based on KGs, such as information retrieval (Chen et al. 2022e; Li et al. 2023), recommendation systems (Wei et al. 2023; Zhang et al. 2022), integration with LLMs (Pan et al. 2023; Luo et al. 2023b), etc. However, most KGs suffer from incompleteness issues. As an essential way to address the problem, relation reasoning, *i.e.,* relation prediction, can be generally divided into two categories (Chen et al. 2023a), including transductive relation reasoning and inductive relation reasoning (See Fig. 1). In fact, the inductive scenario is more common in the real world, *e.g.,* new users (*i.e.,* entities) are added in e-commerce KGs over continuous time

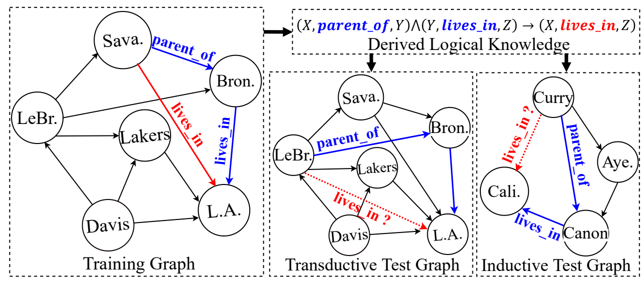
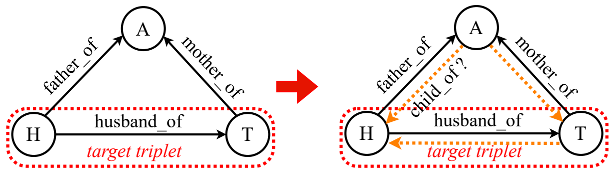


Figure 1: Illustration of transductive and inductive relation reasoning. In the transductive scenario, entities in test graphs are all seen in the model during training. While as for the inductive scenario, unseen entities may exist in test graphs.

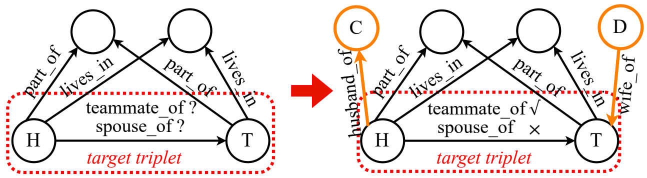
(Chen et al. 2022d, 2023b; Zhang et al. 2023b,a). Therefore, more attention has recently been drawn to inductive models, which can infer missing links between brand-new entities, and our research also falls into this category.

Rule-based and GNN-based methods (Xia et al. 2022) are two typical inductive relation reasoning methods. Rule-based methods, such as NeuralLP (Yang, Yang, and Cohen 2017), RuleN (Meilicke et al. 2018), and DRUM (Sadeghian et al. 2019), induce the entity-independent rules based on observed co-occurrence patterns. Such methods are naturally suitable for inductive scenarios with inherent inductive attributes. However, they suffer from limited expressive ability, and scalability (Teru, Denis, and Hamilton 2020). Inspired by the great achievements of GNN-based methods for other graph-based tasks, several GNN-based inductive relation reasoning models have recently been proposed. Among them, Graph Inductive Learning (Teru, Denis, and Hamilton 2020), *i.e.,* GraIL, is the most influential. It first leverages RGCN (Schlichtkrull et al. 2018) to infer missing triplets based on the enclosing subgraph, which gains great inductive ability. Based on the prototype GraIL, many GraIL-based models (*e.g.,* TACT (Chen et al. 2021), CoMPILE (Mai et al. 2021), Meta-iKG (Zheng et al. 2022), RMPI (Geng et al. 2023), etc.) are proposed. Although proven effective, there are two common limitations in most GraIL-based models, *i.e.,* (1) insufficient message communication, and (2) insufficient neighborhood information collection.

*Corresponding author



(a) Comparison of Message-Passing Passageways w./w.o. the Complete Mutual Message-Passing Passageways



(b) Comparison of Extracted Subgraph Structures w./w.o. the Isolated Neighbors beyond the Enclosing Subgraph.

Figure 2: Limitations of existing GraIL-based models. The differences between ideal scenarios (right figures) and scenarios in previous models (left figures) are colored in orange.

Insufficient message communication. Mutual relationships can always be found between two related entities in the real world. However, due to the incompleteness issue, such mutual relationships will usually not both exist in the given KG, *e.g.,* the edge (*A*,*child_of*,*H*), which represents the mutual relationship corresponding to (*H*,*father_of*,*A*), does not exist in Fig. 2 (a). Meanwhile, existing message-passing mechanisms can only aggregate messages along the given edges. Therefore, without such mutual relational edges, the expressive ability of the model is limited. However, we argue that the absence of edges in graphs does not mean the absence of message-passing passageways. For example, the orange passageways should exist in Fig. 2 (a). Thus, we want a more powerful message communication mechanism by leveraging the omitted mutual information.

Insufficient neighborhood information collection. Most of the existing GraIL-based models perform reasoning on enclosing subgraphs composed of the paths between the target head and tail entities. Such subgraph extraction fashions abandon many neighbors around the target entities, which have been proven important in other tasks (Chen et al. 2022a; Liu et al. 2022; Niu et al. 2021). In this work, we argue that isolated neighbors around the target entities beyond the enclosing subgraph will benefit the discriminative ability of the models. For example, the subgraph with the isolated neighbors, *i.e.,* *C* and *D*, will be more informative for the model to distinguish the edges representing the *teammate_of* and *spouse_of* relations, compared to the corresponding enclosing subgraph in Fig. 2 (b). Thus, performing reasoning based on a neighbor-enhanced subgraph with more isolated neighbors is worth a try.

To this end, we propose a novel GraIL-based framework (MINES) by introducing a Message Intercommunication mechanism on the Neighbor-Enhanced Subgraph. Concretely, we first extract the neighbor-enhanced subgraph by including isolated neighbors around the target entities beyond the enclosing subgraph. Then, a sequential message intercommunication mechanism is designed to introduce bi-directed information interactions between connected entities. It is achieved by inserting an undirected/bi-directed graph convolutional network (GCN) layer between each of two uni-directed relational graphs convolutional network (RGCN) layers to compensate for the omitted hidden mutual information. Since we only know such missing mutual relational edges are very likely to exist but cannot tell the exact type of relation, the homogeneous subgraph without relations is used for intercommunication. The main contributions are summarized below:

- We propose a novel inductive relation reasoning framework, MINES, which improves the capacity of GraIL-based models by introducing a message intercommunication mechanism on the neighbor-enhanced subgraph.
- We innovatively design the message intercommunication mechanism. It introduces the bi-directed message interactions between connected entities to compensate for the omitted mutual relational information. Besides, we first perform inductive relation reasoning on the neighbor-enhanced subgraph for better discriminative ability.
- Extensive experiments on twelve inductive datasets demonstrate the superiority, effectiveness, and transfer ability of MINES.

Related Work

Transductive Relation Reasoning Methods Transductive relation reasoning methods (Liang et al. 2022) are usually embedding-based methods these years, including TransE (Bordes et al. 2013) and their variants (Trouillon et al. 2017; Sun et al. 2018; Luo et al. 2023a). Besides, such relation reasoning tasks are also extended to temporal (Chen, Liao, and Zhao 2023; Chen et al. 2023d), multimodal scenarios (Zhao et al. 2022; Song, Li, and Li 2023; Chen et al. 2023c), few-shot scenarios (Chen et al. 2022b,c). However, most of these methods inherently assume a fixed entity set (Teru, Denis, and Hamilton 2020; Yang, Cohen, and Salakhudinov 2016), generally referring to transductive scenarios instead of inductive ones. Concretely, transductive settings are the most used and practical scenario, while it is not the best. Under this setting, models are handling tasks with unseen atoms, *e.g.,* relations, entities, by retraining the models again to ensure visibility of the models, which is really resource-consuming.

Inductive Relation Reasoning Methods Rule-based and GNN-based methods (Bi et al. 2023b,a; Yu et al. 2023a,b) are two typical inductive relation reasoning methods. Rule-based methods induce logical rules in KGs according to observed co-occurrences of frequent patterns. RuleN (Meilicke et al. 2018) and AMIE (Galárraga et al. 2013) set empirical thresholds based on the number of statistical results to mine the rules. Moreover, NeuralLP (Yang, Yang, and Cohen 2017) and DRUM (Sadeghian et al. 2019) derive rules in an end-to-end differentiable way. However, they suffer from limited expressive ability and scalability. Inspired by the great achievements of GNN for other graph-based tasks, several GNN-based inductive relation reasoning models have recently been proposed. The GNN models based on Graph Inductive Learning, *i.e.,* GraIL-based models, (Teru, Denis, and Hamilton 2020) are the most influential among them. The prototype GraIL (Teru, Denis, and Hamilton 2020), as the landmark GNN-based model, first leverages RGCN to perform the reasoning based on the local enclosing subgraph. Based on it, many incremental works are developed, including TACT (Chen et al. 2021), CoMPILE (Mai et al. 2021), Meta-iKG (Zheng et al. 2022), RPC-IR (Pan et al. 2021), and etc. These GraIL-based models all achieve promising inductive performances. TACT (Chen et al. 2021) and CoMPILE (Mai et al. 2021) both raise the importance of relation embeddings in the task. Concretely, TACT uses topology-aware correlations between relations to generate representations for triplet scoring. Besides, CoMPILE enhances the message interactions between relations and entities with a novel mechanism. After that, some popular strategies are also integrated, such as contrastive learning models (*e.g.,* RPC-IR (Pan et al. 2021), etc.) and meta-learning models (*e.g.,* Meta-iKG (Zheng et al. 2022), etc.).

Message Passing in GNN-based Models The message-passing schemes aim to achieve information communication between entities. Based on the basic message-passing scheme proposed in vanilla GNN (Scarselli et al. 2008), various strategies are integrated to achieve better expressive ability, such as the GCN (Welling and Kipf 2017) and GAT (Veličković et al. 2018), which shows great capacities on graph-based tasks (Tu et al. 2022, 2023; Yang et al. 2023a,b; Mo et al. 2023a,b; Wen et al. 2023b,a). Then, with these ideas extended from homogeneous to other graph types, more message-passing schemes come out, including HAN (Wang et al. 2019), HGNN (Feng et al. 2019), RGCN (Schlichtkrull et al. 2018), etc. However, the above message-passing mechanisms can only aggregate the message along the given edges. However, the mutual relational edges are usually missing in the given KGs, which restricts the expressive ability of the model. However, we argue that the absence of edges in graphs does not mean the lack of message-passing passageways. For example, the passageway from entity *A* to *T* corresponding to (*T*, *mother_of*, *A*) is supposed to exist in Fig. 2 (a). Some works try to conquer this problem by directly adding "inverse relation" (Vashishth et al. 2019) to provide such bi-directional message communication, which will sometimes lead to incorrect inversing edge construction. It will further hinder the discriminative ability of models. For example, the edge *(A, FatherOf, B)* and its inversing edges *(B, FatherOf, A)* can never both exist.

Subgraph Extraction in GraIL-based Models Recently, GraIL-based models have shown promising potential for the task. The prototype GraIL (Teru, Denis, and Hamilton 2020) first performs inductive relation reasoning on the undirected enclosing subgraph, which only considers the entities within the paths between the target entities. Like GraIL, its variants, *i.e.,* other GraIL-based models, all perform reasoning on the enclosing subgraph, but some make specific implementation modifications on subgraph extraction. For example, CoMPILE (Mai et al. 2021) extracts the directed enclosing subgraph instead of the undirected enclosing subgraph to improve the inference performance on symmetrical triplets. In our work, inspired by the success of exploiting more neighborhood information for GNN-based models for other tasks (Hamilton, Ying, and Leskovec 2017; Chen et al. 2022a; Liu et al. 2022), we notice that a certain number of isolated neighbors outside the enclosing subgraph will also benefit inductive models. However, directly applying their subgraph strategies (Long et al. 2021; Wang et al. 2021) to our task will include more useless nodes in subgraphs, which may hinder the reasoning efficiency and accuracy. Compared to it, MINES is the first model to perform inductive relation reasoning on specific neighbor-enhanced subgraphs. Moreover, the filtering procedure for subgraph extraction (See Sec.3.2) effectively purges the useless nodes and improves the reasoning performance.

Method

Preliminary

The KG is the directed relational graph, denoted as *KG* $= (\mathcal{E}, \mathcal{R}, \mathcal{G})$, where the entity (*i.e.,* node) set and the relation (i.e., edge label) set are represented as $\mathcal{E}$ and $\mathcal{R}$, respectively, and $\mathcal{G} = \{(e_u, r_{u,v}, e_v) \mid e_u, e_v \in \mathcal{E}, r_{u,v} \in \mathcal{R}\}$ is the set of fact triplets (*i.e.,* edges) in the given *KG*. The main goal of the inductive relation reasoning is to predict the likelihood of the target relation r_t between the target head e_h and target tail e_t by scoring the target triplet (e_h, r_t, e_t) in the given *KG*. MINES is implemented based on the prototype GraIL (Teru, Denis, and Hamilton 2020), and the main ideas in our paper have good scalability, which can be easily applied to other GraIL-based inductive models. For a fair comparison, we follow the settings in previous GraIL-based models to perform reasoning solely based on the structural semantics derived from the subgraph. The proposed MINES have four steps (See Fig. 3): (1) neighbor-enhanced subgraph extraction, (2) entity labeling and embedding initialization, (3) message intercommunication, and (4) triplet scoring.

Neighbor-Enhanced Subgraph Extraction

The neighbor-enhanced (N-E) subgraph in MINES is extracted by including more isolated neighbors beyond the paths between the target entities based on the enclosing subgraph extracted. Similar to (Teru, Denis, and Hamilton 2020), we first generate the *k-hop* neighbors around the target head and tail for both incoming and outgoing edges, denoted as $N_k(e_h)$ and $N_k(e_t)$. Then, we take the intersection of the neighbor sets and get the enclosing subgraph by filtering out the entities which are isolated beyond the paths between target entities. However, different from the enclosing subgraph in previous GraIL-based models, the neighbor-enhanced subgraph further contains the *k-hop* isolated entities around the target head and target tail entities, together with the corresponding edges, which constitutes the *k-hop* path between each isolate entity and the target entity (See **orange lines** in Fig. 3). In this way, MINES enlarges the original enclosing subgraph to the more informative neighbor-enhanced open subgraph for reasoning.

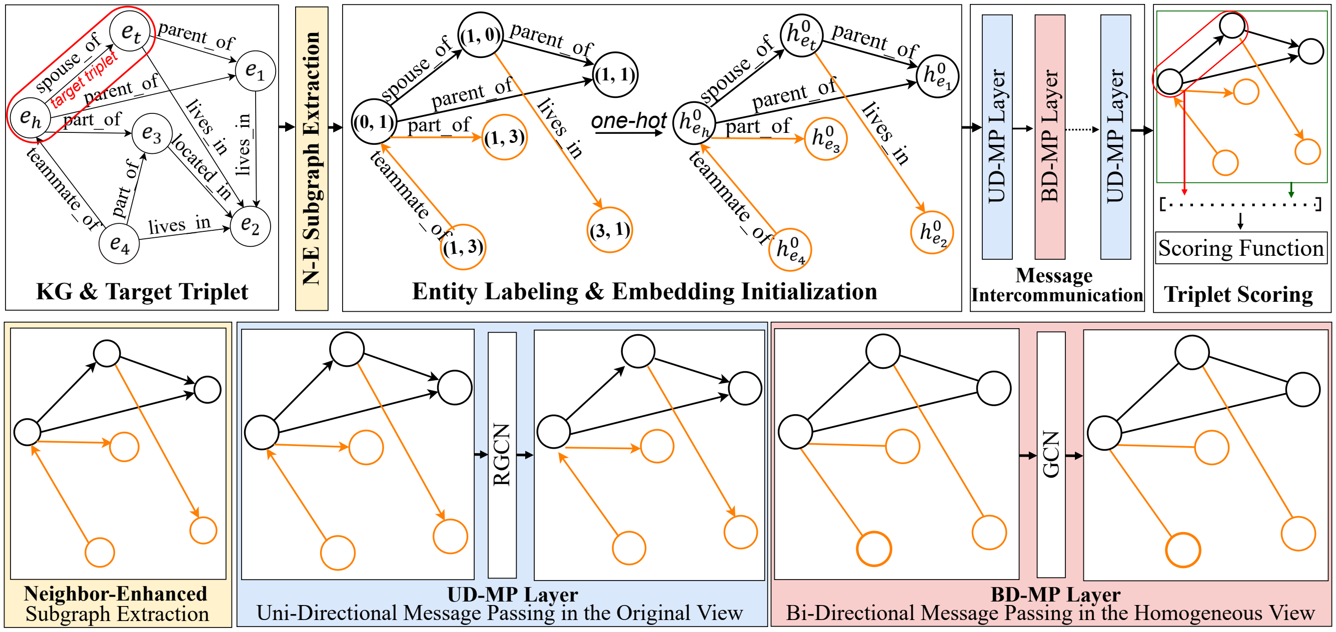


Figure 3: The framework of the proposed MINES, which includes four main steps: neighbor-enhanced subgraph extraction, entity labeling & embedding initialization, message intercommunication, and triplet scoring. Our main contribution lies in the first and the third step. Precisely, in the first step, we extract the neighbor-enhanced subgraph by adding neighboring entities and edges (colored in orange) to the enclosing subgraph (colored in black). In the previous work, only the black entities and relations are included in the subgraph for reasoning; In the third step, we feed subgraphs into the novel message intercommunication module to learn representations. In this step, GCN is integrated with RGCN to achieve better information interactions.

Entity Labeling and Embedding Initialization

Entity features are required for the message communication mechanism in GNN models. Since no prior entity attributes are used in our framework, we leverage the structural semantics to initialize entity features, including two steps, *i.e.,* entity labeling, and embedding initialization. Concretely, each entity e_i in the subgraph for the target triplet (e_h, r_t, e_t) is labeled with the distance tuple $(d(e_i, e_h), d(e_i, e_t))$, where $d(e_i, e_h)$ represents the shortest distance of the undirected path between e_i and e_h without counting any path through e_h (likewise for $d(e_i, e_t)$). An example is present in Fig. 3. Afterward, we initialize the embedding by the one-hot operation, denoted as $\mathbf{h}_{e_i}^0 = [\text{one-hot}(d(e_i, e_h)) \oplus \text{one-hot}(d(e_i, e_t))]$. Note that two target entities, e_h and e_t are uniquely labeled as (0, 1) and (1, 0).

Message Intercommunication Mechanism

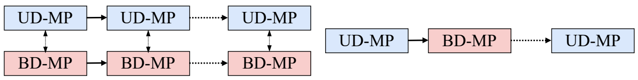
The message communication mechanisms in previous GraIL-based models for inductive relation reasoning mainly rely on the R-GCN (Schlichtkrull et al. 2018). With such a uni-directional message-passing scheme, the model will leave out the hidden mutual relational information underlying each edge in the given KG. Thus, the expressive ability of models is restricted without such information.

To achieve sufficient message communication, we propose a message intercommunication mechanism composed of the Uni-Directional Message Passing (*i.e., UD-MP*) and Bi-Directional Message Passing (*i.e., BD-MP*) layers, which are two different network layers with different message passing schemes. The message passing scheme in *UD-MP* layers is based on the original subgraph (*i.e.,* original view), and the R-GCN is selected to update the embeddings as same as GraIL-based models. In comparison, the message passing scheme in the *BD-MP* layer is selected as the simple GCN based on the homogeneous view of the subgraph. Besides, we use a sequential framework by inserting a *UD-MP* layer (*i.e.,* undirected GCN layer) between each of two directed *BD-MP* layers (*i.e.,* directed RGCN layers) in the proposed message intercommunication mechanism.

Uni-Directional Message Passing in the Original View We adopt the RGCN (Schlichtkrull et al. 2018), denoted as $\mathbf{g}_{ud}(\cdot)$, for embedding updating in the original view of the subgraph, which is the initially extracted subgraph without any modifications. The *UD-MP* layers persevere the similar message communication schemes in previous GraIL-based models. They focus on the uni-directional message passing along the given edges in the KGs.

$$\mathbf{h}_{e_i}^{l+1} = \mathbf{g}_{ud}(\mathbf{h}_{e_i}^l) = \sigma\left(\sum_{r\in\mathcal{R}}\sum_{e_j\in\mathcal{N}_{e_i}^r}\frac{1}{c_{e_i,r}}\mathbf{h}_{e_j}^l\mathbf{W}_r^l + \mathbf{h}_{e_i}^l\mathbf{W}_0^l\right), \tag{1}$$

where the feature vector of entity e_i at the l^{th} layer is present as $\mathbf{h}_{e_i}^l$. Besides, the set of neighbour indices for specific entity e_i with the relation $r \in \mathcal{R}$ is marked as $\mathcal{N}_i^r$. $c_{e_i,r} = |\mathcal{N}_{e_i}^r|$ is a normalization constant. $\mathbf{W}_r^l$, $\mathbf{W}_0^l$ are two weight parameters. Moreover, $\sigma(\cdot)$ is an activation function.



(a) The Parallel Framework (b) The Sequential Framework

Figure 4: Illustration of parallel and sequential intercommunication frameworks.

Bi-Directional Message Passing in the Homogeneous View The basic GCN (Welling and Kipf 2017), denoted as $\mathbf{g}_{bd}(\cdot)$, is adopted for embedding updating in the homogeneous view of the subgraph, which is generated by replacing the uni-directed labelled edges with the bi-directed/undirected unlabeled edges in the extracted subgraph (See Fig. 3). Since we only know such missing mutual edges are very likely to exist but cannot tell the exact relation (*i.e.,* edge label) without any language models, *e.g.,* Bert (Devlin et al. 2019), we just build up the missing message-passing passageways with unlabeled edges in this work to bridge the message intercommunication between entities.

$$\mathbf{h}_{e_i}^{l+1} = \mathbf{g}_{bd}(\mathbf{h}_{e_i}^{l}) = \sigma\left(\sum_{e_j \in \mathcal{N}_{e_i}} \frac{1}{c_{e_i,e_j}} \mathbf{h}_{e_j}^{l} \mathbf{W}^{l}\right), \tag{2}$$

where we leverage $\mathbf{h}_{e_i}^{l}$ to denote as the entity feature vector of e_i at the l^{th} layer. Besides, the set of neighbor indices of node e_i is present as $\mathcal{N}_i$. Morever, a normalization constant for edge (e_i, e_j) is calculated as $c_{e_i,e_j} = \sqrt{|\mathcal{N}_{e_i}| \cdot |\mathcal{N}_{e_j}|}$. Meanwhile, $\mathbf{W}^{\mathbf{l}}$ and $\sigma(\cdot)$ denote the weight parameter and activation function separately.

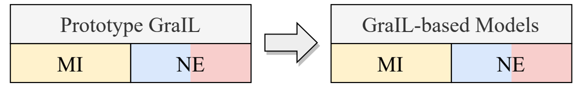
Framework of the Intercommunication Mechanism The sequential framework integrates the *UD-MP* and *BD-MP* layers in our message intercommunication mechanism. Concretely, we insert a *UD-MP* layer (*i.e.,* undirected GCN layer) between each of two directed *BD-MP* layers (*i.e.,* directed RGCN layers) to compensate for the hidden mutual information omitted by each RGCN layer. Besides, we reassign the entity embedding generated in one view to the corresponding entity in the other view for the cross-view communication of the entity embeddings. Moreover, our model fixes the first and last network layer as the *UD-MP* layer. The main reason for selecting a sequential framework instead of a parallel one is to reduce the complexity of the model. As shown in Fig. 4, parallel frameworks will generally take more GNN layers than sequential ones. Meanwhile, the substitution of the simpler GCN layer for the RGCN layer will lead to a reduction in the number of parameters.

Triplet Scoring

The scoring function $f(e_h, r_t, e_t)$, which aims to measure whether the inferring is of high possibility or not, is calculated as follows in MINES.

$$f(e_h, r_t, e_t) = \mathbf{W}[\mathbf{h}_{KG_{NE_{(e_h,r_t,e_t)}}} \oplus \mathbf{h}_{e_h} \oplus \mathbf{h}_{e_t} \oplus \mathbf{h}_{r_t}], \tag{3}$$

where $\mathbf{W}$ denotes as the weight parameter, $\mathbf{h}_{e_h}$ and $\mathbf{h}_{e_t}$ represent the hidden embeddings of entities for head and tail respectively, the learned embedding of the target relation is



(a) MINES (b) MINES+

Figure 5: Transferring MINES from prototype GraIL to GraIL-based models (from MINES to MINES+).

marked as $\mathbf{h}_{r_t}$. $\mathbf{h}_{KG_{NE(e_h,r_t,e_t)}}$, as the subgraph representation, is calculated as follows:

$$\mathbf{h}_{\mathrm{KG}_{NE_{(e_h,r_t,e_t)}}} = \frac{1}{|\mathcal{E}_{NS}|} \sum_{e_i \in \mathcal{E}_{NE}} \mathbf{h}_{e_i}^{L}, \tag{4}$$

where $|\mathcal{E}_{NE}|$ is the quantity of the entity elements in the set, and L represents the number of network layers used in the model, *i.e.,* layer quantity of *UD-MP* and *BD-MP* layers.

Based on the scoring function, we train the model to score positive triplets higher than the negative triplets via the noise-contrastive hinge loss (Bordes et al. 2013):

$$\mathcal{L} = \sum_{(e_h,r_t,e_t) \in \mathcal{G}} \max(0, f(e_h^n, r_t^n, e_t^n) - f(e_h, r_t, e_t) + \gamma), \tag{5}$$

where γ is the margin hyper-parameter, and we use (e_h, r_t, e_t) and (e_h^n, r_t^n, e_t^n) to represent the positive and negative triplets separately. In particular, we generate a negative triplet by replacing the head (or tail).

From MINES to MINES+

Although MINES is developed based on prototype GraIL in previous sections, however, as most of the GraIL-based models are also developed based on prototype GraIL, the two most important modules in MINES can be easily extended to other GraIL-based models, as shown in Fig. 5. Such plug-and-play attributes are promising and proven by transfer analysis in experiment sections.

Experiment

Experiment Setting

Most KG datasets are originally created for transductive settings. To evaluate the inductive ability, 12 datasets based on FB15K-237, NELL-995, and WN18RR, which contain v1, v2, v3, v4 subsets (Teru, Denis, and Hamilton 2020). We implement MINES based on the prototype GraIL model (Teru, Denis, and Hamilton 2020), and experiments are conducted based on a single NVIDIA TITAN XP. We select the 3-layer model (*i.e., UD-MP+BD-MP+UD-MP*) and *3-hop* extracted subgraphs as same as the prototype GraIL to compare with SOTA models fairly. Besides, the dimension of the feature representation and dropout rate is set to 32 and 0.5 separately. Moreover, the batchsize and the margin parameter γ are set to 16 and 10 separately.

Main Results

Tab. 1, Tab. 2 and Tab. 3 show that MINES significantly outperforms other compared baselines on inductive datasets for

Methods		WN18RR				FB15K-237				NELL-995			
		v1	v2	v3	v4	v1	v2	v3	v4	v1	v2	v3	v4
Rule-Based	Neural-LP	86.02	83.78	62.90	82.06	69.64	76.55	73.95	75.74	64.66	83.61	87.58	85.69
	DRUM	86.02	84.05	63.20	82.06	69.71	76.44	74.03	76.20	59.86	83.99	87.71	85.94
	RuleN	90.26	89.01	76.46	71.75	75.24	88.70	91.24	91.79	84.99	88.40	87.20	80.52
GraIL-based	GraIL	94.32	94.18	85.80	92.72	84.69	90.57	91.68	94.46	86.05	92.62	93.34	87.50
	TACT	94.64	97.45	86.33	97.97	83.82	92.98	91.28	94.42	88.72	94.80	94.79	85.76
	CoMPILE	98.23	**99.56**	93.60	**99.80**	85.50	91.68	93.12	94.90	80.16	95.88	96.08	85.48
	RPC-IR	99.41	93.76	98.75	87.24	92.75	93.93	95.26	84.23	88.12	94.12	96.10	87.81
Ours	MINES	**99.69**	99.48	**99.27**	99.58	**99.01**	**99.41**	**99.56**	**99.48**	**99.55**	**99.59**	**99.70**	**97.52**

Table 1: AUC-PR on the inductive benchmark datasets. Best results are boldfaced, and the second best ones are underlined.

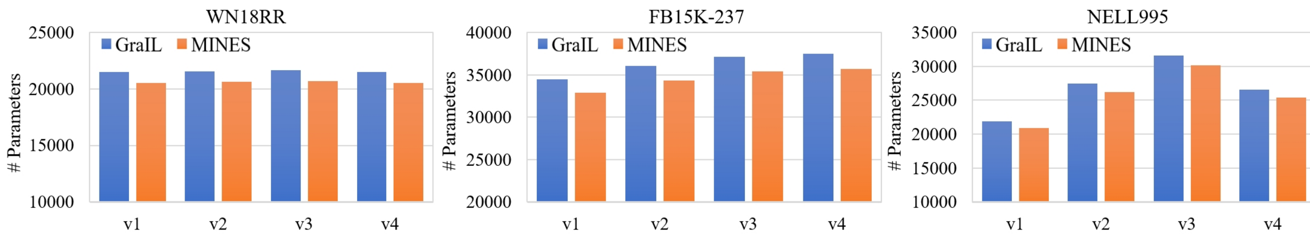


Figure 6: Comparison of the parameter numbers of 3-layer MINES and prototype GraIL for training on benchmark datasets.

Methods		WN18RR			
		v1	v2	v3	v4
Rule-Based	Neural-LP	74.37	68.93	46.18	67.13
	DRUM	74.37	68.93	46.18	67.13
	RuleN	80.85	78.23	53.39	71.59
GraIL-based	GraIL	82.45	78.68	58.43	73.41
	TACT	83.24	81.63	62.73	76.27
	CoMPILE	83.60	79.82	60.69	75.49
	RPC-IR	85.11	81.63	62.40	76.35
Ours	MINES	**87.23**	**83.87**	**69.42**	**79.04**

Table 2: Hit@10 results on the WN18RR.

Methods		FB15K-237		NELL-995	
		v1	v2	v3	v4
Rule-Based	Neural-LP	52.92	58.94	82.71	80.58
	DRUM	52.92	58.73	82.71	80.58
	RuleN	49.76	77.82	77.26	61.35
GraIL-based	GraIL	64.15	81.80	91.41	73.19
	TACT	65.61	83.05	91.35	74.69
	CoMPILE	67.64	82.98	92.77	75.19
	RPC-IR	67.56	82.53	94.01	71.82
Ours	MINES	**67.67**	**83.18**	**95.92**	**81.61**

Table 3: Hit@10 results on the FB15K-237 and NELL-995.

both the Hits@10 and AUC-PR evaluation metrics. On average, our method makes 4.01% on AUC-PR and 2.78% on Hit@10 boosts on each dataset compared to the previous best performances. Specifically, MINES improves the best AUC-PR performance by an average of 6.06% on NELL-995 and FB15K-237 datasets. It further highlights the better discriminative and expressive ability of MINES with the novel message intercommunication mechanism and novel strategy of reasoning on the neighbor-enhanced subgraph. Besides, we observe that the improvement of our Hit@10 performance on FB15K-237 is not apparent, which indicates that our model may be more effective for the sparser datasets with fewer relations. Besides, Fig. 6 shows that MINES is a lightweight model compared to the prototype GraIL model. For each dataset, MINES reduces about 1500 parameters for training. Such parameter reduction is caused by replacing an RGCN layer with a simpler GCN layer. Thus, the above results demonstrate the superiority of MINES from both the evaluation metrics and model complexity aspects.

Ablation Study

The ablation studies are performed on multiple benchmarks to investigate the effectiveness and robustness of the strategy of reasoning on the neighbor-enhanced (N.E.) subgraph and message intercommunication (M.I.) mechanism in MINES. Two compared models (*i.e.,* GraIL w. M.I. and GraIL w. N.E.) are generated. In GraIL w. M.I., only the message intercommunication mechanism is integrated with the GraIL. In GraIL w. N.E., the original uni-directional message communication mechanism is kept, but the neighbor-enhanced subgraph substitutes the enclosing subgraph. Tab. 4 shows that the average AUC-PR values on WN18RR, FB15K-237, and NELL-995 increased by 7.66%, 8.77%, and 7.64% with the M.I. mechanism, which is higher compared to improvements of the neighbor-enhanced subgraph (*i.e.,* 2.01%, 0.33%, 1.67%) on these datasets. It suggests that the message intercommunication mechanism is more effective for classification performance. Besides, Tab. 5 and Tab. 6 show that the average Hit@10 values are increased by 1.98% with the M.I. mechanism, while lower than the average improvement of 3.62% brought by leveraging the N.E. procedure. It indicates that the ranking performance benefits more from the N.E. subgraph.

Methods	WN18RR				FB15K-237				NELL-995			
	v1	v2	v3	v4	v1	v2	v3	v4	v1	v2	v3	v4
GraIL	94.32	94.18	85.80	92.72	84.69	90.57	91.68	94.46	86.05	92.62	93.34	87.50
GraIL w. M.I.	99.50↑	99.43↑	99.17↑	99.56↑	98.46↑	99.08↑	99.52↑	99.41↑	97.16↑	99.50↑	99.56↑	93.83↑
GraIL w. N.E.	95.68↑	95.66↑	88.80↑	95.22↑	85.38↑	90.94↑	91.71↑	94.69↑	87.80↑	94.38↑	94.21↑	89.81↑
MINES	99.69↑	99.48↑	99.27↑	99.58↑	99.01↑	99.41↑	99.56↑	99.48↑	99.55↑	99.59↑	99.70↑	97.52↑

Table 4: Ablation study of AUC-PR on the benchmark datasets w./w.o. M.I. and N.E.. ↑ denotes "increasing".

Methods	WN18RR			
	v1	v2	v3	v4
GraIL	82.45	78.68	58.43	73.41
GraIL w. M.I.	84.04↑	80.27↑	60.99↑	75.79↑
GraIL w. N.E.	87.03↑	82.35↑	68.21↑	78.90↑
MINES	87.23↑	83.87↑	69.42↑	79.04↑

Table 5: Ablation study of Hit@10 on WN18RR w./w.o. M.I. and N.E.. ↑ denotes "increasing".

Methods	FB15K-237		NELL-995	
	v1	v2	v3	v4
GraIL	64.15	81.80	91.41	73.19
GraIL w. M.I.	64.90 ↑	82.43 ↑	93.33 ↑	77.62 ↑
GraIL w. N.E.	66.29 ↑	82.53 ↑	91.41	75.81 ↑
MINES	67.67 ↑	83.18 ↑	95.92 ↑	81.61 ↑

Table 6: Ablation study of Hit@10 on the FB15K-237 and NELL-995 w./w.o. M.I. and N.E.. ↑ denotes "increasing".

Transfer Analysis on TACT and CoMPILE

The results in previous sections have shown that our strategies can benefit the prototype GraIL model. In this section, we further extend our idea to TACT and CoMPILE, two typical GraIL-based models, to evaluate the scalability and generalizability of our approach. The new models (*i.e.,* TACT w. M.I.N.E. and CoMPLIE w. M.I.N.E.) are implemented by replacing the subgraph extraction and message communication modules with the neighbor-enhanced subgraph extraction and message intercommunication mechanism in MINES. Tab. 7 shows the significant improvements in both AUC-PR and Hit@10 metrics on two benchmark datasets (*i.e.,* WN18RR v1 and NELL-995 v1) for both of the new models. It demonstrates that the ideas in MINES can be well-scaled to other GraIL-based models.

Intercommunication Framework Analysis

To demonstrate the suitability of the proposed sequential intercommunication framework, we investigate various combinations of the *BD-MP* and *UD-MP* layers in MINES on three benchmark datasets. Concretely, the compared frameworks include (1) RRR, (2) Bi-RRR, (3) GGG, (4) GRR, (5) RRG, (6) CCC and (7) RGR (*i.e.,* the baseline without any inter-communication frameworks), where R represents one *UD-MP* layer, *i.e.,* one RGCN layer in the original view of the subgraph, and G represents one *BD-MP* layer, *i.e.,* one GCN layer in the homogeneous view of the subgraph. In particular, Bi-RRR is a 3-RGCN-layer model in the bi-directional subgraph that adds all of the inversing edges to the original subgraph. CCC represents the model

Methods	WN18RR v1		NELL-955 v1	
	AUC-PR	Hit@10	AUC-PR	Hit@10
GraIL	94.32	82.45	86.05	59.50
GraIL w. M.I.N.E.	**99.69**	**87.23**	**99.55**	**63.50**
TACT	94.64	83.24	85.58	55.00
TACT w. M.I.N.E.	**95.55**	**85.11**	**88.72**	**62.00**
CoMPILE	98.23	83.60	80.16	58.38
CoMPILE w. M.I.N.E.	**100.00**	**87.50**	**100.00**	**60.00**

Table 7: Evaluation on MINES+: transfer experiments on TACT and CoMPILE. The best results are boldfaced.

Methods	WN18RR v1		FB15K-237 v1		NELL-955 v1	
	AUC-PR	Hit@10	AUC-PR	Hit@10	AUC-PR	Hit@10
RRR	95.68	87.03	84.69	64.15	86.05	59.50
Bi-RRR	95.96	82.45	83.10	62.44	70.64	58.50
GGG	97.00	84.04	82.99	61.22	67.11	47.50
CCC	89.74	82.18	82.21	62.20	67.16	49.00
GRR	96.02	84.04	85.61	64.15	86.04	60.50
RRG	99.50	84.04	98.78	54.63	99.51	54.50
RGR	**99.69**	**87.23**	**99.01**	**67.67**	**99.55**	**63.50**

Table 8: Performance comparison between different intercommunication frameworks. The best results are boldfaced.

that the RGCNs in UN-MP layers in GraIL are replaced by COMPGCNs, where new weights and embeddings for reversed relation^{-1} are trained. Tab. 8 shows that the RGR used in MINES outperforms Bi-RRR and GGG on average by 16.63% on AUC-PR and 6.78% on Hit@10. It indicates that the negative impacts of the redundant relation information in the Bi-RRR and the relationship information loss in GGG influence the performance more than the positive impacts of the intercommunication. Besides, compared to GRR and RRG, the average AUC-PR and Hit@10 are higher by 5.17% and 5.82%. It suggests that fixing the first and last RGCN layers benefits the discriminative ability of the models for the relation reasoning tasks in KGs. Thus, our RGR is most proper.

Conclusion

In this paper, we propose a novel GraIL-based framework, termed MINES, by introducing Message Intercommunication mechanism on the Neighbor-Enhanced Subgraph. As a result, our model is of better discriminative and expressive ability due to sufficient information communication. Extensive experiments prove the promising capacity of the proposed MINES from various aspects, especially for the superiority, effectiveness, and transfer ability. In the future, we aim to investigate neighbor-enhanced subgraphs in a more fine-grained and efficient manner.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (no. 62325604, 62276271).

References

Bi, W.; Cheng, X.; Xu, B.; Sun, X.; Xu, L.; and Shen, H. 2023a. Bridged-GNN: Knowledge Bridge Learning for Effective Knowledge Transfer. In *CIKM*, 99–109.

Bi, W.; Xu, B.; Sun, X.; Xu, L.; Shen, H.; and Cheng, X. 2023b. Predicting the silent majority on graphs: Knowledge transferable graph neural network. In *Proceedings of the ACM Web Conference 2023*, 274–285.

Bordes, A.; Usunier, N.; Garcia-Duran, A.; Weston, J.; and Yakhnenko, O. 2013. Translating embeddings for modeling multi-relational data. *NeurIPs*, 26: 2787–2795.

Chen, H.; Huang, Z.; Xu, Y.; Deng, Z.; Huang, F.; He, P.; and Li, Z. 2022a. Neighbor enhanced graph convolutional networks for node classification and recommendation. *KBS*, 246: 108594.

Chen, J.; He, H.; Wu, F.; and Wang, J. 2021. Topology-Aware Correlations Between Relations for Inductive Link Prediction in Knowledge Graphs. *AAAI*, 35: 6271–6278.

Chen, M.; Zhang, W.; Geng, Y.; Xu, Z.; Pan, J. Z.; and Chen, H. 2023a. Generalizing to Unseen Elements: A Survey on Knowledge Extrapolation for Knowledge Graphs. *arXiv preprint arXiv:2302.01859*.

Chen, M.; Zhang, W.; Yao, Z.; Chen, X.; Ding, M.; Huang, F.; and Chen, H. 2022b. Meta-learning based knowledge extrapolation for knowledge graphs in the federated setting. *arXiv preprint arXiv:2205.04692*.

Chen, M.; Zhang, W.; Zhu, Y.; Zhou, H.; Yuan, Z.; Xu, C.; and Chen, H. 2022c. Meta-Knowledge Transfer for Inductive Knowledge Graph Embedding. In *SIGIR*.

Chen, W.; Wan, H.; Guo, S.; Huang, H.; Zheng, S.; Li, J.; Lin, S.; and Lin, Y. 2022d. Building and exploiting spatial–temporal knowledge graph for next POI recommendation. *KBS*.

Chen, W.; Wan, H.; Wu, Y.; Zhao, S.; Cheng, J.; Li, Y.; and Lin, Y. 2023b. Local-Global History-aware Contrastive Learning for Temporal Knowledge Graph Reasoning. *arXiv preprint arXiv:2312.01601*.

Chen, X.; Zhang, J.; Wang, X.; Wu, T.; Deng, S.; Wang, Y.; Si, L.; Chen, H.; and Zhang, N. 2023c. Continual Multimodal Knowledge Graph Construction. *arXiv preprint arXiv:2305.08698*.

Chen, X.; Zhang, N.; Xie, X.; Deng, S.; Yao, Y.; Tan, C.; Huang, F.; Si, L.; and Chen, H. 2022e. Knowprompt: Knowledge-aware prompt-tuning with synergistic optimization for relation extraction. In *Proceedings of the ACM Web conference 2022*, 2778–2788.

Chen, Z.; Li, D.; Zhao, X.; Hu, B.; and Zhang, M. 2023d. Temporal Knowledge Question Answering via Abstract Reasoning Induction. *arXiv preprint arXiv:2311.09149*.

Chen, Z.; Liao, J.; and Zhao, X. 2023. Multi-granularity Temporal Question Answering over Knowledge Graphs. In *ACL*, 11378–11392. Toronto, Canada.

Devlin, J.; Chang, M.-W.; Lee, K.; and Toutanova, K. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *NACCL*.

Feng, Y.; You, H.; Zhang, Z.; Ji, R.; and Gao, Y. 2019. Hypergraph neural networks. In *AAAI*.

Galárraga, L. A.; Teflioudi, C.; Hose, K.; and Suchanek, F. 2013. AMIE: Association Rule Mining under Incomplete Evidence in Ontological Knowledge Bases. In *WWW*.

Geng, Y.; Chen, J.; Pan, J. Z.; Chen, M.; Jiang, S.; Zhang, W.; and Chen, H. 2023. Relational message passing for fully inductive knowledge graph completion. In *ICDE*, 1221–1233. IEEE.

Hamilton, W. L.; Ying, R.; and Leskovec, J. 2017. Inductive representation learning on large graphs. In *NeurIPs*.

Li, Q.; Guo, S.; Ji, C.; Peng, X.; Cui, S.; and Li, J. 2023. Dual-Gated Fusion with Prefix-Tuning for Multi-Modal Relation Extraction. *arXiv preprint arXiv:2306.11020*.

Liang, K.; Meng, L.; Liu, M.; Liu, Y.; Tu, W.; Wang, S.; Zhou, S.; Liu, X.; and Sun, F. 2022. Reasoning over different types of knowledge graphs: Static, temporal and multi-modal. *arXiv preprint arXiv:2212.05767*.

Liu, S.; Ying, R.; Dong, H.; Li, L.; Xu, T.; Rong, Y.; Zhao, P.; Huang, J.; and Wu, D. 2022. Local augmentation for graph neural networks. In *ICML*, 14054–14072. Baltimore, Maryland USA: PMLR.

Long, X.; Huang, C.; Xu, Y.; Xu, H.; Dai, P.; Xia, L.; and Bo, L. 2021. Social recommendation with self-supervised metagraph informax network. In *CIKM*.

Luo, L.; Li, Y.-F.; Haffari, G.; and Pan, S. 2023a. Normalizing flow-based neural process for few-shot knowledge graph completion. In *SIGIR*.

Luo, L.; Li, Y.-F.; Haffari, G.; and Pan, S. 2023b. Reasoning on Graphs: Faithful and Interpretable Large Language Model Reasoning. *arXiv preprint arxiv:2310.01061*.

Mai, S.; Zheng, S.; Yang, Y.; and Hu, H. 2021. Communicative Message Passing for Inductive Relation Reasoning. In *AAAI*.

Meilicke, C.; Fink, M.; Wang, Y.; Ruffinelli, D.; Gemulla, R.; and Stuckenschmidt, H. 2018. Fine-grained evaluation of rule-and embedding-based systems for knowledge graph completion. In *ISWC*, 3–20. Monterey: Springer.

Mo, Y.; Chen, Y.; Lei, Y.; Peng, L.; Shi, X.; Yuan, C.; and Zhu, X. 2023a. Multiplex Graph Representation Learning Via Dual Correlation Reduction.

Mo, Y.; Lei, Y.; Shen, J.; Shi, X.; Shen, H. T.; and Zhu, X. 2023b. Disentangled Multiplex Graph Representation Learning. In *ICML*.

Niu, G.; Li, Y.; Tang, C.; Geng, R.; Dai, J.; Liu, Q.; Wang, H.; Sun, J.; Huang, F.; and Si, L. 2021. Relational learning with gated and attentive neighbor aggregator for few-shot knowledge graph completion. In *SIGIR*, 213–222. Montréal, Quebec: ACM.

Pan, S.; Luo, L.; Wang, Y.; Chen, C.; Wang, J.; and Wu, X. 2023. Unifying Large Language Models and Knowledge Graphs: A Roadmap. *arXiv preprint arxiv:2306.08302*.

Pan, Y.; Liu, J.; Zhang, L.; Hu, X.; Zhao, T.; and Lin, Q. 2021. Learning First-Order Rules with Relational Path Contrast for Inductive Relation Reasoning. In *arXiv preprint arXiv:2110.08810*.

Sadeghian, A.; Armandpour, M.; Ding, P.; and Wang, D. Z. 2019. Drum: End-to-end differentiable rule mining on knowledge graphs. *NeurIPs*, 32.

Scarselli, F.; Gori, M.; Tsoi, A. C.; Hagenbuchner, M.; and Monfardini, G. 2008. The graph neural network model. *T-NN*.

Schlichtkrull, M.; Kipf, T. N.; Bloem, P.; Berg, R. v. d.; Titov, I.; and Welling, M. 2018. Modeling relational data with graph convolutional networks. In *ESWC*.

Song, S.; Li, X.; and Li, S. 2023. How to Bridge the Gap between Modalities: A Comprehensive Survey on Multimodal Large Language Model. *arXiv preprint arXiv:2311.07594*.

Sun, Z.; Deng, Z.-H.; Nie, J.-Y.; and Tang, J. 2018. RotatE: Knowledge Graph Embedding by Relational Rotation in Complex Space. In *ICLR*.

Teru, K. K.; Denis, E.; and Hamilton, W. L. 2020. Inductive Relation Prediction by Subgraph Reasoning. *ICML*.

Trouillon, T.; Dance, C. R.; Gaussier, É.; Welbl, J.; Riedel, S.; and Bouchard, G. 2017. Knowledge Graph Completion via Complex Tensor Factorization. *JMLR*, 18: 1–38.

Tu, W.; Zhou, S.; Liu, X.; Ge, C.; Cai, Z.; and Liu, Y. 2023. Hierarchically Contrastive Hard Sample Mining for Graph Self-Supervised Pretraining. *IEEE Transactions on Neural Networks and Learning Systems (TNNLS)*, 1–14.

Tu, W.; Zhou, S.; Liu, X.; Liu, Y.; Cai, Z.; Zhu, E.; Zhang, C.; and Cheng, J. 2022. Initializing Then Refining: A Simple Graph Attribute Imputation Network. In *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence (IJCAI)*, 3494–3500.

Vashishth, S.; Sanyal, S.; Nitin, V.; and Talukdar, P. 2019. Composition-based Multi-Relational Graph Convolutional Networks. In *ICLR*.

Veličković, P.; Cucurull, G.; Casanova, A.; Romero, A.; Liò, P.; and Bengio, Y. 2018. Graph Attention Networks. In *ICLR*.

Wang, P.; Agarwal, K.; Ham, C.; Choudhury, S.; and Reddy, C. K. 2021. Self-supervised learning of contextual embeddings for link prediction in heterogeneous networks. In *WWW*.

Wang, X.; Ji, H.; Shi, C.; Wang, B.; Ye, Y.; Cui, P.; and Yu, P. S. 2019. Heterogeneous graph attention network. In *WWW*.

Wei, W.; Ren, X.; Tang, J.; Wang, Q.; Su, L.; Cheng, S.; Wang, J.; Yin, D.; and Huang, C. 2023. LLMRec: Large Language Models with Graph Augmentation for Recommendation. *arXiv preprint arXiv:2311.00423*.

Welling, M.; and Kipf, T. N. 2017. Semi-supervised classification with graph convolutional networks. In *ICLR*.

Wen, Y.; Liu, S.; Wan, X.; Wang, S.; Liang, K.; Liu, X.; Yang, X.; and Zhang, P. 2023a. Efficient Multi-View Graph Clustering with Local and Global Structure Preservation. In *MM*.

Wen, Y.; Wang, S.; Liao, Q.; Liang, W.; Liang, K.; Wan, X.; and Liu, X. 2023b. Unpaired Multi-View Graph Clustering with Cross-View Structure Matching. *T-NNLS*.

Xia, J.; Zhu, Y.; Du, Y.; and Li, S. Z. 2022. A survey of pretraining on graphs: Taxonomy, methods, and applications. *arXiv preprint arXiv:2202.07893*.

Yang, F.; Yang, Z.; and Cohen, W. W. 2017. Differentiable learning of logical rules for knowledge base reasoning. *NeurIPs*, 30.

Yang, X.; Jin, J.; Wang, S.; Liang, K.; Liu, Y.; Wen, Y.; Liu, S.; Zhou, S.; Liu, X.; and Zhu, E. 2023a. DealMVC: Dual Contrastive Calibration for Multi-view Clustering. In *ACM MM*.

Yang, X.; Tan, C.; Liu, Y.; Liang, K.; Wang, S.; Zhou, S.; Xia, J.; Li, S. Z.; Liu, X.; and Zhu, E. 2023b. CONVERT: Contrastive Graph Clustering with Reliable Augmentation. In *ACM MM*.

Yang, Z.; Cohen, W.; and Salakhudinov, R. 2016. Revisiting semi-supervised learning with graph embeddings. In *ICLR*, 40–48. New York, USA: PMLR.

Yu, S.; Liu, S.; Wang, S.; Tang, C.; Luo, Z.; Liu, X.; and Zhu, E. 2023a. Sparse Low-Rank Multi-View Subspace Clustering With Consensus Anchors and Unified Bipartite Graph. *T-NNLS*.

Yu, S.; Wang, S.; Wen, Y.; Wang, Z.; Luo, Z.; Zhu, E.; and Liu, X. 2023b. How to Construct Corresponding Anchors for Incomplete Multiview Clustering. *T-CSVT*.

Zhang, M.; Wu, S.; Yu, X.; Liu, Q.; and Wang, L. 2022. Dynamic graph neural networks for sequential recommendation. *IEEE T-KDE*, 35(5): 4741–4753.

Zhang, M.; Xia, Y.; Liu, Q.; Wu, S.; and Wang, L. 2023a. Learning Latent Relations for Temporal Knowledge Graph Reasoning. In *ACL*, 12617–12631.

Zhang, M.; Xia, Y.; Liu, Q.; Wu, S.; and Wang, L. 2023b. Learning Long-and Short-term Representations for Temporal Knowledge Graph Reasoning. In *Proceedings of the ACM Web Conference 2023*, 2412–2422.

Zhao, Y.; Cai, X.; Wu, Y.; Zhang, H.; Zhang, Y.; Zhao, G.; and Jiang, N. 2022. MoSE: Modality Split and Ensemble for Multimodal Knowledge Graph Completion. In *EMNLP*, 10527–10536.

Zheng, S.; Mai, S.; Sun, Y.; Hu, H.; and Yang, Y. 2022. Subgraph-aware Few-Shot Inductive Link Prediction via Meta-Learning. *T-KDE*.