

# Document-level Relation Extraction via Subgraph Reasoning

Xingyu Peng\*, Chong Zhang\* and Ke Xu†

State Key Lab of Software Development Environment, Beihang University, Beijing, 100191, China  
{xypeng, chongzh, kexu}@buaa.edu.cn

## Abstract

Document-level relation extraction aims to extract relations between entities in a document. In contrast to sentence-level relation extraction, it deals with longer texts and more complex entity interactions, which requires reasoning over multiple sentences with rich reasoning skills. Most current researches construct a document-level graph first, and then focus on the overall graph structure or the paths between the target entity pair in the graph. In this paper, we propose a novel **subgraph reasoning (SGR) framework** for document-level relation extraction. SGR combines the advantages of both graph-based models and path-based models, integrating various paths between the target entity pair into a much simpler subgraph structure to perform relational reasoning. Moreover, **the paths generated by our designed heuristic strategy** explicitly model the requisite reasoning skills and roughly cover the **supporting sentences** for each relation instance. Experimental results on DocRED show that SGR outperforms existing models, and further analyses demonstrate that our method is both effective and explainable. Our code is available at <https://github.com/CrystalOvo/SGR>.

## 1 Introduction

Relation extraction (RE), which aims to identify semantic relations between entities from plain text, is one of the most fundamental tasks in information extraction and natural language processing. Previous research efforts on RE mainly concentrate on sentence-level, attempting to predict the relations between two entities within a sentence [Zeng *et al.*, 2014; Zhou *et al.*, 2016; Zhang *et al.*, 2018; Wei *et al.*, 2020]. However, a large number of relations are expressed by more than one sentence and can only be extracted by reading and reasoning over multiple sentences. For this reason, many researchers start to explore document-level RE [Quirk and Poon, 2017; Gupta *et al.*, 2019; Christopoulou *et al.*, 2019].

\*Equal contribution.

†Corresponding author.

<b>Espoo Cathedral</b> [1] The <i>Espoo Cathedral</i> is a medieval stone church in <i>Espoo</i> , <i>Finland</i> and the seat of the <i>Diocese of Espoo</i> of the <i>Evangelical Lutheran Church of Finland</i> . [2] The cathedral is located in the district of <i>Espoon keskus</i> , near the <i>Espoonjoki</i> river. [3] The oldest parts of the church were completed in <i>the 1480s</i> and it is thus the oldest preserved building in the city. ... [6] In addition to being the seat of the <i>Diocese of Espoo</i> , the <i>Espoo Cathedral</i> serves as the church for the <i>Espoo Cathedral Parish</i> and ...	
Subject: <i>Espoo</i> Object: <i>Finland</i> Relation: <b>country</b>	Supporting Sentences: [1]
Subject: <i>Espoo Cathedral Parish</i> Object: <i>Finland</i> Relation: <b>country</b>	Supporting Sentences: [1], [6]
Subject: <i>Espoo Cathedral</i> Object: <i>Espoon keskus</i> Relation: <b>location</b>	Supporting Sentences: [1], [2]

Figure 1: An example of document-level RE from DocRED. Three annotated relational facts are presented, with entity mentions involved in these facts colored and other entity mentions underlined for clarity.

Recently, A large-scale, manually-annotated, and general-purpose dataset for document-level RE, DocRED, is proposed in [Yao *et al.*, 2019]. DocRED attracts much attention and further propels the document-level RE forward.

In document-level RE, there are **several major challenges**. First of all, the target entities involved in a relational fact may reside in different sentences, **making it impossible to determine their relations solely based on a single sentence**. Additionally, an entity may be mentioned in many sentences, so the information from corresponding mentions must be properly aggregated to better represent the entity. Moreover, **many relations need to be identified with rich reasoning skills** (i.e., pattern recognition, logical reasoning, co-reference reasoning, and common-sense reasoning) [Yao *et al.*, 2019], which often requires reasoning over global and complex information in the document.

As Figure 1 shows, it is **easy to identify the intra-sentence relational fact** (*Espoo*, country, *Finland*) that the subject and object appear in the same sentence. On the contrary, identifying the inter-sentence relational facts (*Espoo Cathedral Parish*, country, *Finland*) and (*Espoo Cathedral*, location, *Espoon keskus*) that the subject and object do not appear in the

same sentence and thus have long-distance dependencies is not trivial. Therefore, the identification of these two relational facts requires reasoning over more than one sentence. For example, the relation between *Espoo Cathedral Parish* and *Finland* is indirectly established by a reference *The Cathedral* at the beginning of sentence 2, which requires performing co-reference reasoning over sentence 1 and sentence 2. While the relation between *Espoo Cathedral* and *Espoon keskus* is indirectly established by a bridge entity *Espoo Cathedral*, requiring performing logical reasoning over sentence 1 and sentence 6.

These examples indicate that a model for document-level RE needs to be able to effectively capture complex interactions between long-distance entities across multiple sentences in the document. Most recent studies construct a document-level graph based on syntactic dependency, heuristics, and other factors, then use graph neural networks to perform multi-hop reasoning to obtain meaningful entity representations [Zhang *et al.*, 2018; Nan *et al.*, 2020; Zeng *et al.*, 2020]. However, these methods only take into account the overall graph structure, so they may ignore local contextual information surrounding the target entity pair. Moreover, as a result of the over-smoothing problem, these methods may fail to model the interactions between long-distance entities. In addition to graph-based models, some path-based models attempt to extract paths between the target entity pair that retain sufficient information to predict the relations [Xu *et al.*, 2021; Huang *et al.*, 2021]. These methods explicitly consider the above-mentioned reasoning skills and can alleviate the problem of modeling long-distance entity interactions, but they treat each path separately and not all entity pairs can be connected by a path.

In this paper, we propose a novel SubGraph Reasoning (SGR) framework for document-level RE. SGR combines the benefits of both graph-based and path-based models, integrating various paths into a much simpler subgraph structure to perform various reasonings at once. More specifically, we start by building a heterogeneous graph with three different types of nodes: entity nodes, mention nodes, and sentence nodes. Then, considering the reasoning skills required for document-level RE, we heuristically design a simple but extremely effective strategy to generate reasoning paths. These paths not only intuitively model all potential reasoning skills, but they also roughly cover the annotated supporting sentences in practice. Furthermore, the path generation strategy ensures that all entity pairs could be connected by a path. After the reasoning paths are generated, inspired by the success of incorporating subgraph structure in knowledge graph completion [Teru *et al.*, 2020], question answering [Yasunaga *et al.*, 2021], link prediction [Cai and Ji, 2020], etc., we extract a subgraph around the target entity pair based on the previously generated paths and finally apply an R-GCN [Schlichtkrull *et al.*, 2018] on the subgraph. In this way, the model could focus on the most crucial entities, mentions, and sentences and perform joint reasoning over various reasoning paths.

We evaluate SGR on DocRED, experimental results show that SGR achieves better performance than existing models, and further analyses demonstrate that our method is both effective and explainable.

## 2 Methodology

In this section, we first briefly introduce the document-level RE task formulation, and then describe the proposed model in detail. Figure 2 gives an illustration of our approach.

### 2.1 Task Formulation

Given a document  $d$  containing sentences  $\{s_i\}_{i=1}^{n_s}$  and entities  $\{e_i\}_{i=1}^{n_e}$ , the task of document-level RE is to predict the relations  $\mathcal{R}' \subset \mathcal{R}$  between each entity pair  $(e_h, e_t)$ , where  $\mathcal{R}$  is a pre-defined relation set,  $e_h$  and  $e_t$  refer to head and tail entities, respectively. An entity  $e_i$  may appear multiple times in the document via its corresponding mentions  $\{m_j^i\}_{j=1}^{n_{e_i}}$ . A relation  $r \in \mathcal{R}$  exists between entities  $e_h$  and  $e_t$  if it is expressed by any pair of their mentions.

### 2.2 Document Encoding

For a document  $d = \{w_i\}_{i=1}^{|d|}$ , we encode it into a sequence of vectors  $\{g_i\}_{i=1}^{|d|}$ :

$$\{g_1, g_2, \dots, g_{|d|}\} = \text{Encoder}(\{x_1, x_2, \dots, x_{|d|}\}) \quad (1)$$

where  $x_i$  is the concatenation of  $w_i$ 's word embedding and entity type embedding.

$$x_i = [E_w(w_i); E_t(t_i)] \quad (2)$$

where  $E_w(\cdot)$  and  $E_t(\cdot)$  denote the word embedding layer and entity type embedding layer respectively.  $t_i$  is the corresponding entity type id of  $w_i$ .

### 2.3 Subgraph Reasoning

#### Document Graph Construction

Since an entity may spread across multiple sentences in the document via its corresponding mentions, we build a heterogeneous graph  $\mathcal{G}$  comprised of entity nodes, mention nodes, and sentence nodes. The node embeddings are initialized as follows:

$$\mathbf{h}_{s_i} = \text{Max}(\{g_j\}_{w_j \in s_i}) \quad (3)$$

$$\mathbf{h}_{m_i} = \text{Max}(\{g_j\}_{w_j \in m_i}) \quad (4)$$

$$\mathbf{h}_{e_i} = \text{Max}(\{\mathbf{h}_{m_j}\}_{m_j \in e_i}) \quad (5)$$

Then, to explicitly model the interactions among entities, mentions, and sentences, we introduce three types of edges.

- **Mention-Entity Edge:** We connect a mention node to an entity node if the mention refers to the entity in order to model the co-reference of mentions.
- **Mention-Sentence Edge:** The co-occurrence of mentions in a sentence may indicate a relation. Therefore, we add an edge between a mention node and a sentence node if the mention resides in the sentence.
- **Sentence-Sentence Edge:** Unlike previous methods that connect all sentence nodes with each other [Christopoulou *et al.*, 2019; Wang *et al.*, 2020], we only add edges between two sentence nodes that corresponding sentences are adjacent in the document to keep the sequential information.

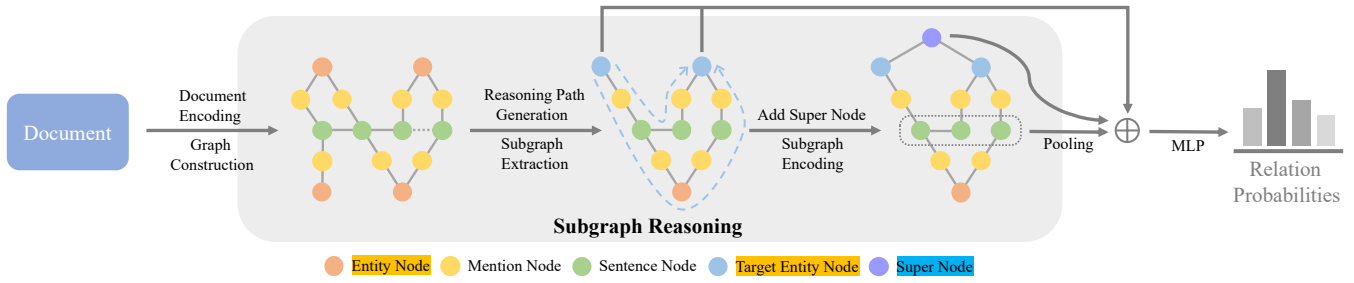


Figure 2: Overview of our approach. The generated reasoning paths are indicated by dashed lines.

### Reasoning Path Generation

Our approach aims to predict the relations by focusing on relevant entities, mentions, and sentences rather than the entire document. To this end, we assume that the paths between two target entities, containing a subset of all nodes in the graph, provide sufficient information to determine the relations. Therefore, inspired by [Xu *et al.*, 2021], we introduce the following reasoning paths, which connect two target entities and explicitly model the four types of reasoning skills we discussed earlier. The meta-path schemes of these paths are illustrated in Figure 3.

- **Intra-sentence Reasoning Path:** The relation between two entities co-occurred in the same sentence (e.g., 1st relational fact in Figure 1) can be formulated as a path in the form of **E-M-S-M-E**. “E”, “M”, and “S” are entity node, mention node, and sentence node, respectively, “-” denotes edge. The intra-sentence reasoning path models two types of reasoning: pattern recognition and common-sense reasoning, as they generally perform reasoning inside the sentence.
- **Inter-sentence Reasoning Path:** For two entities that have no mentions in the same sentence, we use the inter-sentence reasoning path to model the relations between them. It can be thought of as adding additional bridges to the intra-sentence reasoning path, as shown in Figure 3. According to the different kinds of bridges, we introduce logical reasoning path and co-reference reasoning path.

**Logical Reasoning Path:** The relation indirectly established by a bridge entity (e.g., 2nd relational fact in Figure 1) can be formulated as a path in the form of **E-M-S-M-E-M-S-M-E**, which models logical reasoning by performing reasoning over sentences containing the head, bridge, and tail entities, respectively.

**Co-reference Reasoning Path:** The relation between two entities in adjacent sentences is mostly established by a reference (e.g., 3rd relational fact in Figure 1). It can be formulated as a path in the form of **E-M-S-S-M-E**, which models co-reference reasoning by performing reasoning over two adjacent sentences containing the head and tail entities respectively.

Naturally, these two types of reasoning paths are just special cases of the inter-sentence reasoning path, which contain only a single bridge. In fact, with repetition, there could be several different kinds of bridges in an inter-sentence reasoning path.

As a result of reasoning path generation, all entity pairs could be connected by at least one path and these paths are able to capture both the semantic and structural correlations between the entity pair. The problem is that even if we generate paths under the meta-path scheme, there would still be a large number of paths for an entity pair, but only few of them are necessary.

Hence, we simply constrain the scale of paths by limiting the number of bridges, thus generating fewer but better paths to remove the noise from irrelevant paths. Specifically, given an entity pair, we start with searching for intra-sentence reasoning paths (0 bridges) between them, if none are found, turn to search for inter-sentence reasoning paths without repetitions (1 bridge), gradually relaxing the limit on the number of bridges by allowing more and more repetitions until at least one path is found.

To verify the effectiveness of our heuristic path generation strategy, we compare the generated paths with the annotated supporting sentences on the development set of DocRED. The results show that the supporting sentences for 75.2% of relation instances are covered by the sentences contained in the generated paths and 53.5% are exactly the same. Moreover, we find that the paths between each entity pair contain an average of 1.8 sentences, implying that up to 80% of the sentences in the document could be filtered out. Consequently, the reasoning paths generated by our strategy are both sufficient and non-redundant for relational reasoning.

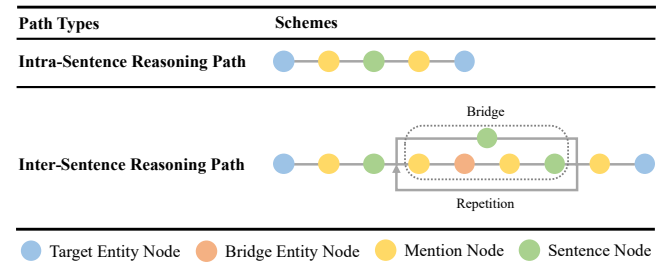


Figure 3: Reasoning path types and their corresponding meta-path schemes.

### Subgraph Extraction

In document-level RE, there could be multiple relations between an entity pair, corresponding to multiple reasoning paths. In addition, determining a single implicit relation may also require multiple reasoning paths.

Previous works [Huang *et al.*, 2021; Xu *et al.*, 2021] tend to use each path to predict the relations independently and then aggregate the results. In contrast, we extract an enclosing subgraph  $\mathcal{G}'$  around the entity pair in the graph  $\mathcal{G}$  to integrate different paths. Specifically, the subgraph  $\mathcal{G}'$  is an induced subgraph of  $\mathcal{G}$ , formed from the nodes that appear at least once on the paths and all of the edges in  $\mathcal{G}$  among those nodes.

As an example shown in Figure 2, for two target entities that do not appear in the same sentence but can be associated by a bridge entity or a bridge sentence, we generate two inter-sentence reasoning paths without repetitions (a logical reasoning path and a co-reference reasoning path) between them, and then use these paths to extract a subgraph.

### Subgraph Encoding

In this stage, we first introduce a **super node**  $z$ , and connect  $z$  to the target entity nodes in the subgraph. The super node embedding  $\mathbf{h}_z$  is initialized as the max-pooling of the target entity node embeddings.

Then we employ an  **$L$ -layer stacked R-GCN** [Schlichtkrull *et al.*, 2018] on the subgraph, which applies message passing for different types of edges separately in each layer. Given node  $u$  at layer  $l$ , the aggregation of  $u$ 's immediate neighbors is defined as:

$$\mathbf{h}_u^{(l+1)} = \sigma \left( \sum_{t \in \mathcal{T}} \sum_{v \in \mathcal{N}_u^t} \frac{1}{c_{u,t}} W_t^{(l)} \mathbf{h}_v^{(l)} + W_0^{(l)} \mathbf{h}_u^{(l)} \right) \quad (6)$$

where  $\sigma(\cdot)$  is an activation function,  $\mathcal{N}_u^t$  is the neighbors of node  $u$  under edge type  $t \in \mathcal{T}$ ,  $c_{u,t} = |\mathcal{N}_u^t|$  is a normalization constant,  $W_t^{(l)}, W_0^{(l)} \in \mathbb{R}^{d_h \times d_h}$  are trainable parameters.

In particular, we set the  **$L$  to 2 to enable the super node to aggregate information from the target entity nodes and their subordinate mention nodes**, as well as the sentence nodes to aggregate information from all entity and mention nodes in the subgraph.

### 2.4 Classification

We formulate the document-level RE task as a multi-label classification task. To calculate the probabilities of relations between an entity pair  $(e_h, e_t)$ , we use information from both the global document encoder and the local subgraph. Specifically, we concatenate the following representations: (1) the linear combination of the initial target entity node embeddings  $\mathbf{h}_c = W_c([\mathbf{h}_{e_h}; \mathbf{h}_{e_t}])$ , which provides global entity-aware information of the target entity pair; (2) the learned **super node embedding**  $\mathbf{h}_z^{(L)}$ , which provides local entity-aware information of the target entity pair; (3) the **max-pooling of all learned sentence node embeddings**  $\mathbf{h}_{\mathcal{G}'}^{(L)} = \text{Max} \left( \left\{ \mathbf{h}_{s_i}^{(L)} \right\}_{s_i \in \mathcal{G}'} \right)$ , which provides local contextual information of the target entity pair; (4) the entity distance embedding  $E_d(d_{ht})$ , where  $d_{ht}$  is the relative distance between the first mentions of the target entity pair in the document,  $E_d(\cdot)$  is the relative distance embedding layer. The concatenated representation is then passed through an MLP:

$$P(r | e_h, e_t) = \text{MLP} \left( \left[ \mathbf{h}_c; \mathbf{h}_z^{(L)}; \mathbf{h}_{\mathcal{G}'}^{(L)}; E_d(d_{ht}) \right] \right) \quad (7)$$

We use **binary cross entropy loss** to train our model:

$$\mathcal{L} = - \sum_{\mathcal{D} \in \mathcal{S}} \sum_{h \neq t} \sum_{r_i \in \mathcal{R}} \mathbb{I}(r_i = 1) \log P(r_i | e_h, e_t) + \mathbb{I}(r_i = 0) \log (1 - P(r_i | e_h, e_t)) \quad (8)$$

where  $\mathcal{S}$  denotes the whole corpus,  $\mathbb{I}(\cdot)$  refers to the indication function.

## 3 Experiments

### 3.1 Dataset

We evaluate our model on DocRED, a large-scale human-annotated dataset for document-level RE constructed from Wikipedia and Wikidata. DocRED contains 3,053 documents for training, 1,000 for development, and 1,000 for testing, involving 96 relation types, 132,275 entities, and 56,354 relational facts. Moreover, more than 40.7% of relational facts can only be extracted from multiple sentences, and 61.1% of relation instances require multiple reasoning skills. It also provides supporting sentences for each relation instance as part of the annotation.

### 3.2 Experimental Settings

Following the settings of [Yao *et al.*, 2019], We use GloVe (100d) [Pennington *et al.*, 2014] and BiLSTM (256d) [Hochreiter and Schmidhuber, 1997] as **word embedding and encoder**, respectively. With setting the batch size to 4, we train our model using AdamW [Loshchilov and Hutter, 2019] optimizer, a linear learning rate scheduler with 6% warmup, and a maximum learning rate of 0.01. All hyperparameters are tuned based on the development set.

### 3.3 Baselines and Evaluation Metric

We compare our model with existing **sequence-based models**, including CNN/LSTM/BiLSTM based models evaluated in [Yao *et al.*, 2019] and HIN [Tang *et al.*, 2020]; **graph-based models**, including GAT [Velickovic *et al.*, 2018], GCNN [Sahu *et al.*, 2019], EoG [Christopoulou *et al.*, 2019], AGGCN [Guo *et al.*, 2019], LSR [Nan *et al.*, 2020], GAIN [Zeng *et al.*, 2020], and SIRE [Zeng *et al.*, 2021]; and **path-based models**, including DRN [Xu *et al.*, 2021] and Paths+BiLSTM [Huang *et al.*, 2021].

Following [Yao *et al.*, 2019], we use **F1 and Ign F1** as the evaluation metrics. Ign F1 denotes F1 scores excluding relational facts that exist in both training and development/testing sets. Moreover, **F1 scores for intra- and inter-sentence entity pairs are also evaluated on the development set**. The results on the testing set is evaluated through CodaLab<sup>1</sup>.

### 3.4 Main Results

As shown in Table 1, our proposed model SGR achieves 57.49% F1 and 55.82% Ign F1 on the development set, **57.15% F1 and 55.12% Ign F1 on the testing set**, significantly outperforming all strong baselines. Specifically, SGR obtains 3.85 F1 and 3.97 Ign F1 points higher than the sequence-based model HIN on the testing set. Compared with the recent path-based models DRN and Paths+LSTM, SGR also

<sup>1</sup><https://competitions.codalab.org/competitions/20717>



Model		Dev				Test	
		Intra-F1	Inter-F1	F1	Ign F1	F1	Ign F1
Sequence-based	CNN [Yao <i>et al.</i> , 2019]	51.87*	37.58*	43.45	41.58	42.26	40.33
	LSTM [Yao <i>et al.</i> , 2019]	56.57*	41.47*	50.68	48.44	50.07	47.71
	BiLSTM [Yao <i>et al.</i> , 2019]	57.05*	43.49*	50.94	48.87	51.06	48.78
	<b>HIN-Glove</b> [Tang <i>et al.</i> , 2020]	-	-	52.95	51.06	53.30	51.15
Graph-based	GAT* [Velickovic <i>et al.</i> , 2018]	58.14	43.94	51.44	45.17	49.51	47.36
	GCNN* [Sahu <i>et al.</i> , 2019]	57.78	44.11	51.52	46.22	51.62	49.59
	EoG* [Christopoulou <i>et al.</i> , 2019]	58.90	44.60	52.15	45.94	51.82	49.48
	AGGCN* [Guo <i>et al.</i> , 2019]	58.76	45.45	52.47	46.29	51.45	48.89
	LSR-Glove [Nan <i>et al.</i> , 2020]	60.83	48.35	55.17	48.82	54.18	52.15
	GAIN-Glove [Zeng <i>et al.</i> , 2020]	61.67	48.77	55.29	53.05	55.08	52.66
	<b>SIRE-Glove</b> [Zeng <i>et al.</i> , 2021]	62.94	48.97	55.91	54.10	55.96	54.04
Path-based	DRN-Glove [Xu <i>et al.</i> , 2021]	-	-	56.49	54.61	56.33	54.35
	Paths+BiLSTM [Huang <i>et al.</i> , 2021]	62.73	49.11	56.54	-	56.23	-
Ours	SGR	<b>63.58</b>	<b>49.94</b>	<b>57.49</b>	<b>55.82</b>	<b>57.15</b>	<b>55.12</b>

 Table 1: Performance on DocRED. Results with \* are reported in [Nan *et al.*, 2020].

achieves a better performance. Moreover, SGR outperforms the state-of-the-art graph-based model SIRE in the same experimental setting by 1.19% and 1.08% in terms of F1 and Ign F1 on the testing set. These results indicate that our proposed model combines the benefits of both path-based models and graph-based models by leveraging both well-designed reasoning paths and subgraph structure to learn more expressive features.

Meanwhile, we can see that SGR outperforms all baseline models on both intra- and inter-sentence settings, which demonstrates that subgraph reasoning improves performance on inter-sentence relation instances, while intra-sentence relation instances benefit from it as well.

### 3.5 Performance Analysis

To further assess the effectiveness of SGR, we divide the development set into several subsets based on different analytical strategies, and evaluate models trained with or without subgraph reasoning on each subset.

**Entity Distance.** In this part, we explore our model’s capacity for capturing the relations between entity pair in long-distance. For this purpose, we examine the model performance in terms of entity distance, which is defined as the relative distance between the first mentions of the target entity pair in the document, and report the F1 score in Figure 4a. The first thing to note is that as the entity distance increases, the F1 score gradually decreases, which is consistent with human intuition that capturing long-distance relations is still a challenging problem for document-level RE. Also, the model with subgraph reasoning consistently outperforms the model without subgraph reasoning, especially when the entity distance  $\geq 64$ . This is because explicitly incorporating the graph structure breaks the limitations of sequence modeling. In addition, the finer subgraph structure further makes model focus more on the target entity pair and related entities, thus reducing the influence of unrelated entities in long-distance.

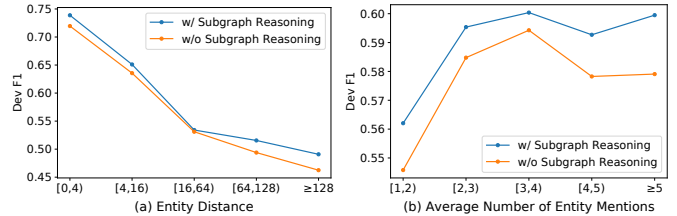


Figure 4: Performance analysis results in terms of (a) entity distance and (b) average number of entity mentions.

**Average Number of Entity Mentions.** In order to investigate our model’s capacity for aggregating information from multiple mentions, we examine the model performance in terms of average mention number of the target entity pair. From Figure 4b we can first observe that models do not perform well when the average mention number is small, which reveals that the information carried by a single mention is quite limited, making relations harder to be predicted. When the average mention number is large, however, not all mentions are necessary to predict the relations, and aggregating information indiscriminately may introduce noise from irrelevant mentions. Nevertheless, the subgraph structure makes model focus more on mentions that are relevant to the target entity pair, so the model with subgraph reasoning consistently maintains a relatively high performance with the increase of average mention number, especially when the average mention number  $\geq 4$ .

### 3.6 Case Study

In Figure 5, we present two case studies to better illustrate the reasoning process and the effectiveness of our proposed model. In the first case, we can first infer that the *Skai TV* in sentence 1 is part of the *Skai Group* in sentence 2 by co-reference reasoning. Then, using *Skai TV* as a bridge entity, we can further infer that *Greece* is the country of *Skai Group*.

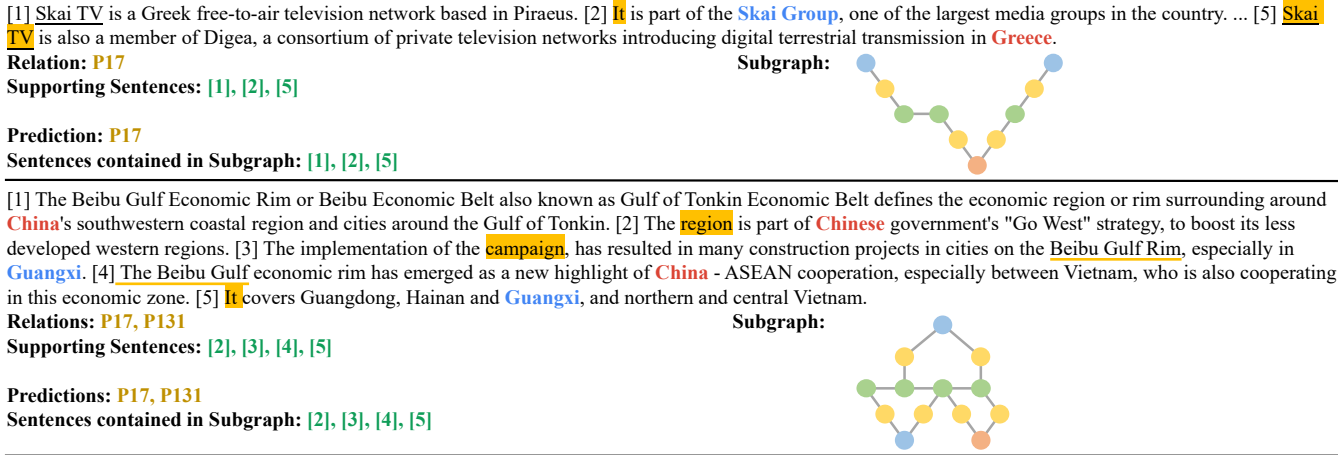


Figure 5: Case studies of our approach. “P17” refers to “country”. “P131” refers to “located in the administrative territorial entity”.

by **logical reasoning**. The preceding reasoning process can be expressed as an **inter-sentence reasoning path with both a bridge sentence and a bridge entity**. In the other more complicated case that two target entities, *China* and *Guangxi*, are mentioned in each of the five sentences, we can generate four reasoning paths between them and use these paths to extract a subgraph to make predictions. It should be noticed that, in these two cases, not only does our proposed model correctly predict the relations, but the sentences contained in the subgraph extracted by our proposed model also accurately cover the supporting sentences.

## 4 Related Work

Various approaches have been shown effective in document-level RE, which can be roughly divided into sequence-based models, graph-based models, and path-based models.

**Sequence-based Models.** These models mostly rely on LSTM- or Transformer-based architectures, which can model complex interactions among entities by implicitly capturing long-distance dependencies. [Yao *et al.*, 2019] use CNN/LSTM/BiLSTM to encode the document into a sequence of hidden state vectors, and then compute entity representations to predict the relations for each entity pair. The hierarchical inference network (HIN) is proposed by [Tang *et al.*, 2020], which uses multi-granularity inference information including entity-level, sentence-level, and document-level to better model complex interactions between entities. [Ye *et al.*, 2020] propose a language representation model to better capture and represent the co-reference information. [Zhou *et al.*, 2021] use attentions learned from transformer to locate relevant context and adopt an adaptive threshold for each entity pair.

**Graph-based Models.** These models generally construct a graph with words, mentions, entities, or sentences as nodes, and predict the relations by reasoning on the graph. [Christopoulou *et al.*, 2019] propose an edge-oriented graph neural model that focuses on encoding information into edge representations rather than node representations. [Nan *et al.*,

2020] propose the latent structure refinement (LSR) model, which dynamically induces the latent document-level graph to empower the relational reasoning across sentences. [Zeng *et al.*, 2020] propose the graph aggregation and inference network (GAIN), which constructs two graphs of different granularity: **heterogeneous mention-level graph** and **entity-level graph** to respectively model the interactions among different mentions and conduct a path reasoning mechanism to predict the relations. [Zeng *et al.*, 2021] propose the separate intra- and inter-sentential reasoning (SIRE) model to represent intra- and inter-sentential relations in different ways, as well as a new and straightforward form of logical reasoning module to cover all cases of logical reasoning chains.

**Path-based Models.** These models typically extract reasoning paths between the entity pair, and perform relational reasoning with them instead of reading the entire document. [Xu *et al.*, 2021] model the paths of reasoning skills explicitly, propose the discriminative reasoning network (DRN) to encode the reasoning paths, and then estimate the relation probability distribution of different reasoning paths. [Huang *et al.*, 2021] select supporting sentences from the document using several well-designed handcrafted rules and combining them with a BiLSTM to predict the relations.

## 5 Conclusion

In this paper, we propose a novel subgraph reasoning (SGR) framework for document-level RE that combines the advantages of both graph-based models and path-based models. In particular, we extract a subgraph around the target entity pair based on corresponding reasoning paths and perform relational reasoning by applying an R-GCN on the subgraph, which considers all potential reasoning skills. Besides, the reasoning paths generated by our designed heuristic strategy explicitly model the document-level RE required reasoning skills and roughly cover the supporting sentences for each relation instance. Experimental results show that SGR outperforms existing models on both intra- and inter-sentence settings. Further analyses demonstrate that our method is both effective and explainable.

## Acknowledgments

This research was supported by the National Natural Science Foundation of China under grant 61932002.

## References

- [Cai and Ji, 2020] Lei Cai and Shuiwang Ji. A multi-scale approach for graph link prediction. In *AAAI*, 2020.
- [Christopoulou *et al.*, 2019] Fenia Christopoulou, Makoto Miwa, and Sophia Ananiadou. Connecting the dots: Document-level neural relation extraction with edge-oriented graphs. In *EMNLP-IJCNLP*, 2019.
- [Guo *et al.*, 2019] Zhijiang Guo, Yan Zhang, and Wei Lu. Attention guided graph convolutional networks for relation extraction. In *ACL*, 2019.
- [Gupta *et al.*, 2019] Pankaj Gupta, Subburam Rajaram, Hinrich Schütze, and Thomas A. Runkler. Neural relation extraction within and across sentence boundaries. In *AAAI*, 2019.
- [Hochreiter and Schmidhuber, 1997] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 1997.
- [Huang *et al.*, 2021] Quzhe Huang, Shengqi Zhu, Yansong Feng, Yuan Ye, Yuxuan Lai, and Dongyan Zhao. Three sentences are all you need: Local path enhanced document relation extraction. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, 2021.
- [Loshchilov and Hutter, 2019] Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization. In *ICLR*, 2019.
- [Nan *et al.*, 2020] Guoshun Nan, Zhijiang Guo, Ivan Sekulic, and Wei Lu. Reasoning with latent structure refinement for document-level relation extraction. In *ACL*, 2020.
- [Pennington *et al.*, 2014] Jeffrey Pennington, Richard Socher, and Christopher Manning. GloVe: Global vectors for word representation. In *EMNLP*, 2014.
- [Quirk and Poon, 2017] Chris Quirk and Hoifung Poon. Distant supervision for relation extraction beyond the sentence boundary. In *EACL*, 2017.
- [Sahu *et al.*, 2019] Sunil Kumar Sahu, Fenia Christopoulou, Makoto Miwa, and Sophia Ananiadou. Inter-sentence relation extraction with document-level graph convolutional neural network. In *ACL*, 2019.
- [Schlichtkrull *et al.*, 2018] Michael Schlichtkrull, Thomas N Kipf, Peter Bloem, Rianne Van Den Berg, Ivan Titov, and Max Welling. Modeling relational data with graph convolutional networks. In *ESWC*, 2018.
- [Tang *et al.*, 2020] Hengzhu Tang, Yanan Cao, Zhenyu Zhang, Jiangxia Cao, Fang Fang, Shi Wang, and Pengfei Yin. Hin: Hierarchical inference network for document-level relation extraction. In *PAKDD*, 2020.
- [Teru *et al.*, 2020] Komal Teru, Etienne Denis, and Will Hamilton. Inductive relation prediction by subgraph reasoning. In *ICML*, 2020.
- [Velickovic *et al.*, 2018] Petar Velickovic, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Liò, and Yoshua Bengio. Graph attention networks. In *ICLR*, 2018.
- [Wang *et al.*, 2020] Difeng Wang, Wei Hu, Ermei Cao, and Weijian Sun. Global-to-local neural networks for document-level relation extraction. In *EMNLP*, 2020.
- [Wei *et al.*, 2020] Zhepei Wei, Jianlin Su, Yue Wang, Yuan Tian, and Yi Chang. A novel cascade binary tagging framework for relational triple extraction. In *ACL*, 2020.
- [Xu *et al.*, 2021] Wang Xu, Kehai Chen, and Tiejun Zhao. Discriminative reasoning for document-level relation extraction. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP*, 2021.
- [Yao *et al.*, 2019] Yuan Yao, Deming Ye, Peng Li, Xu Han, Yankai Lin, Zhenghao Liu, Zhiyuan Liu, Lixin Huang, Jie Zhou, and Maosong Sun. DocRED: A large-scale document-level relation extraction dataset. In *ACL*, 2019.
- [Yasunaga *et al.*, 2021] Michihiro Yasunaga, Hongyu Ren, Antoine Bosselut, Percy Liang, and Jure Leskovec. QA-GNN: Reasoning with language models and knowledge graphs for question answering. In *NAACL*, 2021.
- [Ye *et al.*, 2020] Deming Ye, Yankai Lin, Jiaju Du, Zhenghao Liu, Peng Li, Maosong Sun, and Zhiyuan Liu. Coreferential Reasoning Learning for Language Representation. In *EMNLP*, 2020.
- [Zeng *et al.*, 2014] Daojian Zeng, Kang Liu, Siwei Lai, Guangyou Zhou, and Jun Zhao. Relation classification via convolutional deep neural network. In *COLING*, 2014.
- [Zeng *et al.*, 2020] Shuang Zeng, Runxin Xu, Baobao Chang, and Lei Li. Double graph based reasoning for document-level relation extraction. In *EMNLP*, 2020.
- [Zeng *et al.*, 2021] Shuang Zeng, Yuting Wu, and Baobao Chang. SIRE: Separate intra- and inter-sentential reasoning for document-level relation extraction. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP*, 2021.
- [Zhang *et al.*, 2018] Yuhao Zhang, Peng Qi, and Christopher D. Manning. Graph convolution over pruned dependency trees improves relation extraction. In *EMNLP*, 2018.
- [Zhou *et al.*, 2016] Peng Zhou, Wei Shi, Jun Tian, Zhenyu Qi, Bingchen Li, Hongwei Hao, and Bo Xu. Attention-based bidirectional long short-term memory networks for relation classification. In *Proceedings of the 54th annual meeting of the association for computational linguistics (Volume 2: Short Papers)*, 2016.
- [Zhou *et al.*, 2021] Wenxuan Zhou, Kevin Huang, Tengyu Ma, and Jing Huang. Document-level relation extraction with adaptive thresholding and localized context pooling. In *AAAI*, 2021.