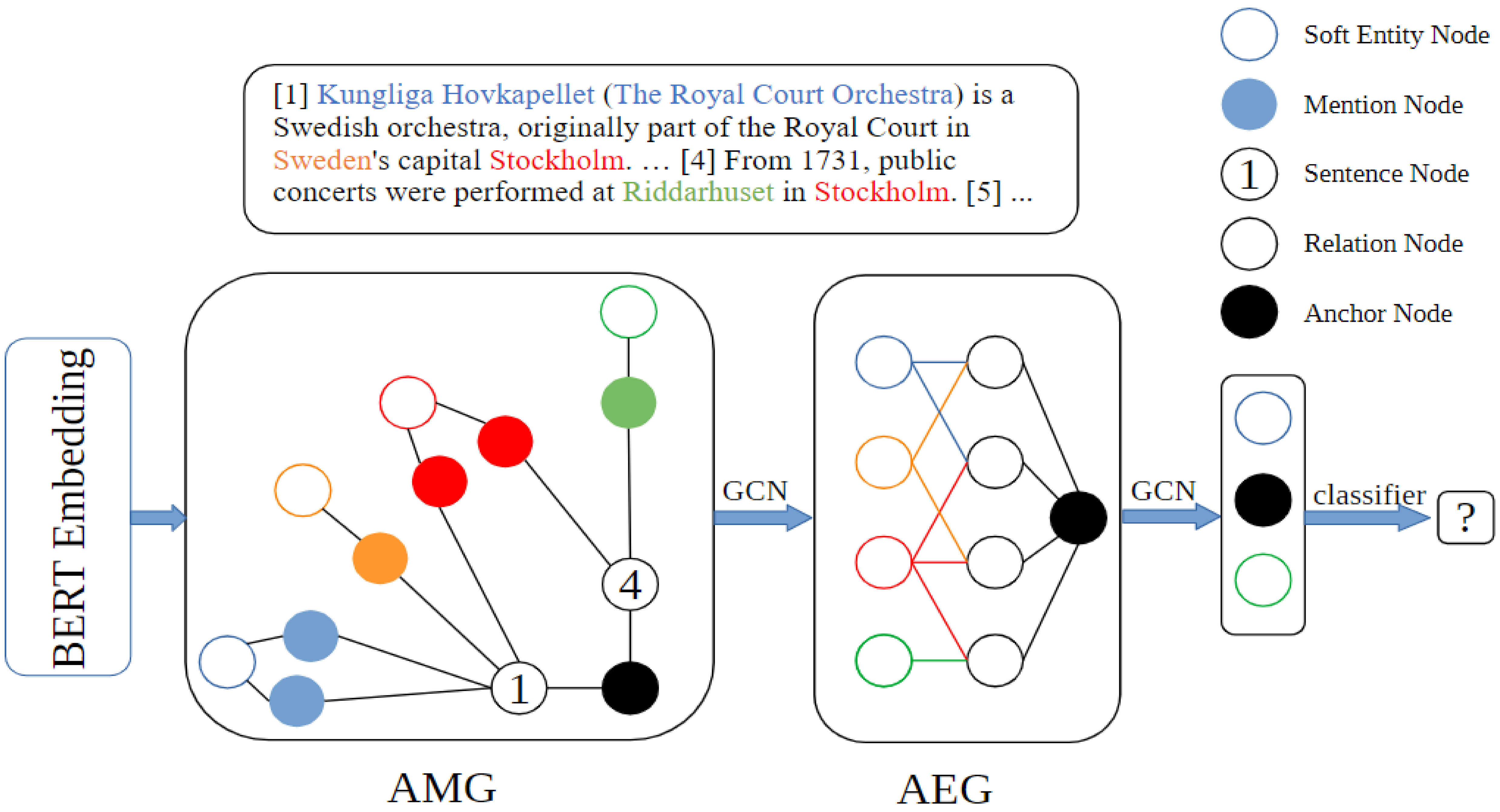
Joint Global and Local Dual–GCN for document–level relation extraction

Wei Zhang, Hong Zhu, Xinjie Sun, Zhangliang Shen, Shuting Gao and Yunhe Pang

Nanjing Power Supply Company,State
Grid Jiangsu Electric Power Co., Ltd.



Background

Document-level relation extraction presents more challenges than sentence-level relation extraction:

- An entity will appear several times in the document, and there may be different representations in different sentences in the document.
- Latent relationship between the two entities which are not in the same or similar context are hard to find out.
- The relationship between entities may need to be determined by the larger context, which requires the aggregation of cross-sentence context information to represent entities

How to effectively and accurately represent the document context and entities?

The approach: Joint Global and Local Dual–GCN

We propose Joint Global and Local Dual-GCN(JDGCN), using double graph neural networks to represent the document and entities in context, including an anchor-based mention-level Graph (AMG) for representation learning and an anchor-based entity-level Graph (AEG) for causal reasoning.

AMG

AMG graph contains four different nodes: mention node , soft entity node , sentence node and a global anchor node. A mention node represents a representation of an entity in a sentence. A special soft entity node is an abstract node of each entity whose initial representation is derived from a mean pooling of all the same entity nodes. The representation of the sentence node comes from the previous section. The anchor node acts as a global node, and the initial state comes from the document embedding in the previous section

AEG

AMG graph contains four different nodes, mention node , soft entity node , sentence node and a global anchor node. Entity node, relationship node and anchor node. The representation of the entity node is from the previous section. If two nodes exist in the same sentence, they are connected to the same relational node, which is connected to a global anchor node that represents the relationship that exists throughout the document.

Representation Learning with GCN

$$h^l = \sigma\left(\sum_{j=1}^{n} A_{ij} W^l h_j^{l-1} + b^l\right)$$

Experiments

Table I shows JDGCN achieves the best results and has a significant improvement over other graph-based models

TABLE I. RESULTS OF JDGCN ON DOCRED DATASET.

Model	DEV				TEST	
	Ign F1	*Ign AUC*	*F1*	*AUC*	*Ign F1*	*F1*
RoBEERTa-RE base	53.85	48.27	56.05	51.35	53.52	55.77
HIN-BERT base	54.29	-	56.31	-	53.70	55.60
LSR-BERT base	52.43	-	59.00	-	56.97	59.05
GAIN-BERT base	59.04	57.76	61.22	60.96	59.00	61.24
JDGCN(OURS)	**60.68**	**58.24**	**61.86**	**61.40**	**60.05**	**61.97**

Ablation study shows the sentence node in AMG and AEG are both important to both graph.

TABLE II. Ablation study

Model	Dev				Test	
	Ign F1	*Ign AUC*	*F1*	*AUC*	*Ign F1*	*F1*
JDGCN	**60.68**	**58.24**	**61.86**	**61.40**	**60.05**	**61.97**
w/o AMG sentence node	59.14	57.86	61.53	60.96	59.32	61.24
w/o AMG anchor node	58.24	56.12	59.84	60.12	58.97	60.56
w/o AEG sentence node	57.68	54.24	59.56	56.17	58.27	59.34

Table III shows that intra-F1 and inter-F1 of the model on the validation set are better than the other baseline, which indicates that the JDGCN model has a high effect on both local and global relation extraction, and its powerful ability is inseparable from the global anchor node, local sentence node and context-based ontology representation.

TABLE III. INTRA AND INTER F1 SCORE COMPARED WITH OTHER MODELS

Model	Intra-F1	Inter-F1
RoBERTa-RE	65.65	50.09
BERT-Two-Step	61.80	47.28
LSR-BERT	65.26	52.05
GAIN-BERT	67.10	53.90
JDGCN	**68.50**	**54.20**