Adaptive Graph Convolutional Networks with Attention Mechanism for Relation Extraction

Zhixin Li, Yaru Sun, Suqin Tang, Canlong Zhang, Huifang Ma

Guangxi Normal University

CONTENTS

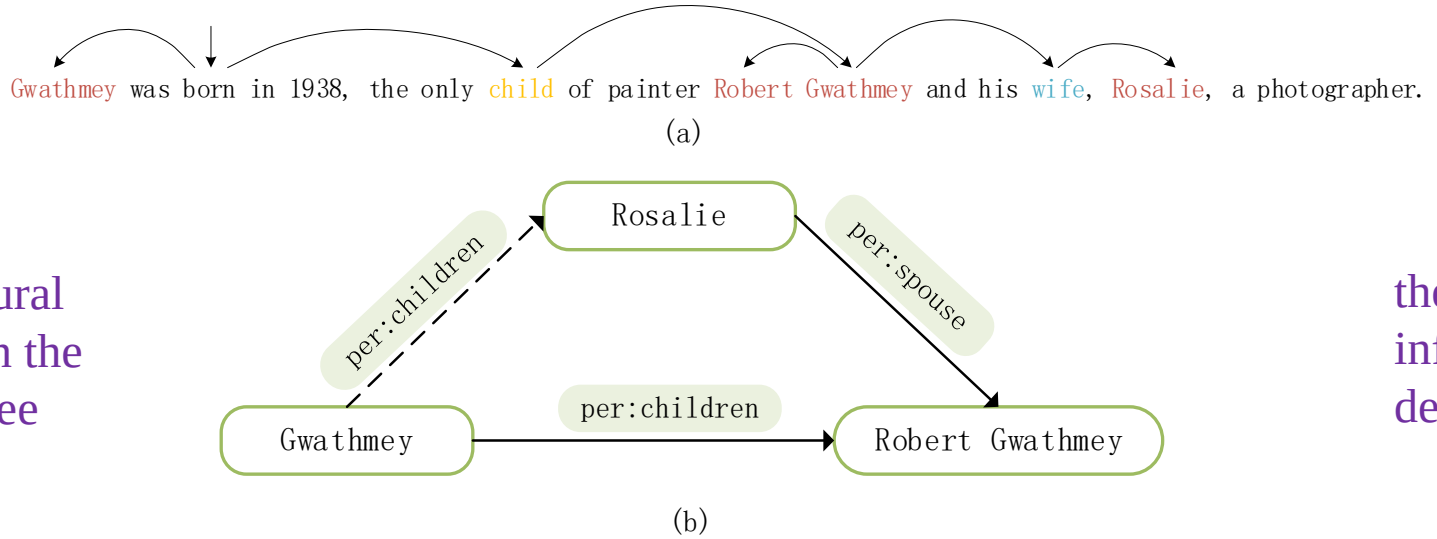
Introduction

The sequence-based model

It operate on word sequences, for example, using cyclic neural networks to encode words to obtain the sentence information.

The dependency-based model

It incorporate the dependency tree of the sentence relationship and effectively use the structural information in the dependency tree to extract features.



the rich structural information on the dependency tree

the multihop influence on the dependent path

Fig.1 Relation extraction of a plain text.

Dual Attention-Guided GCN

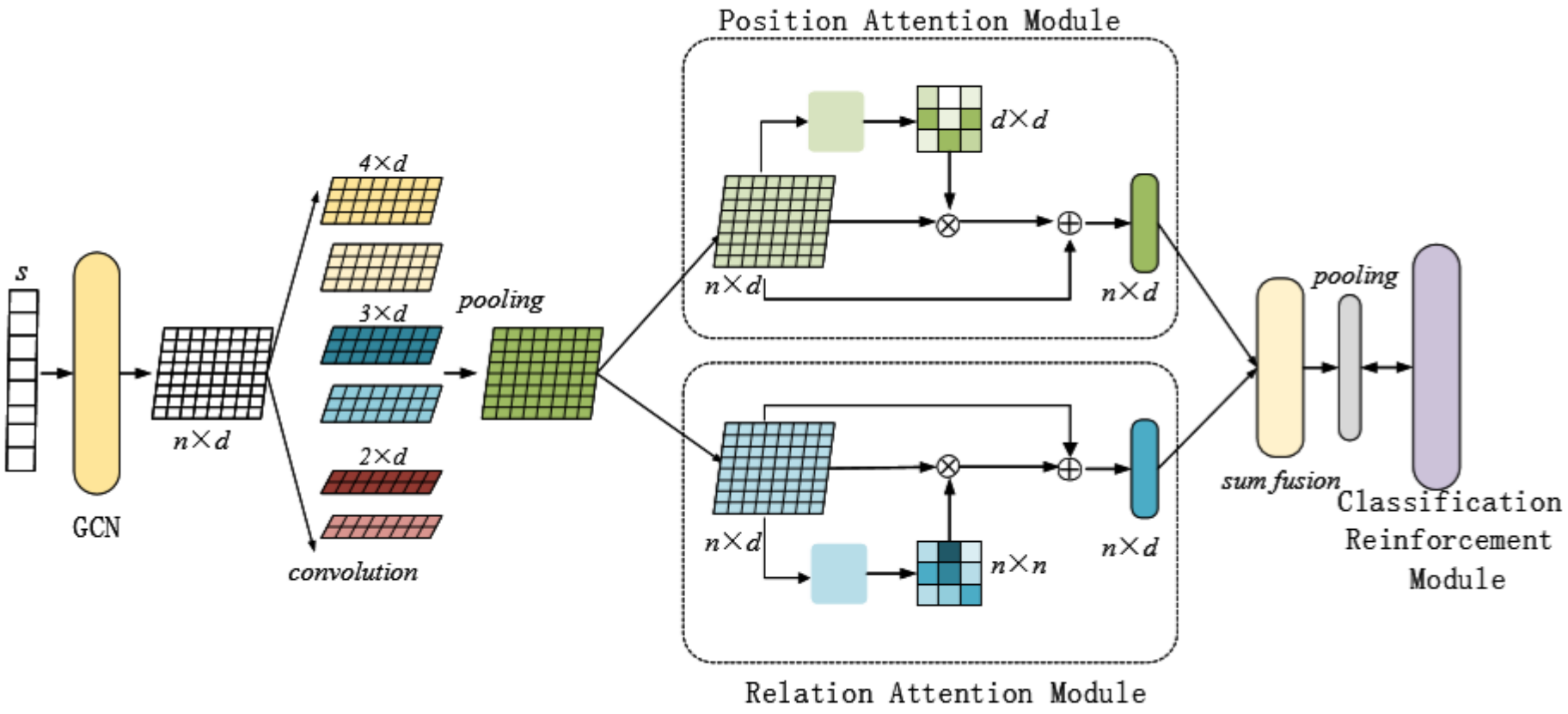


Fig. 2. The DAGCN model.

Dual Attention-Guided GCN

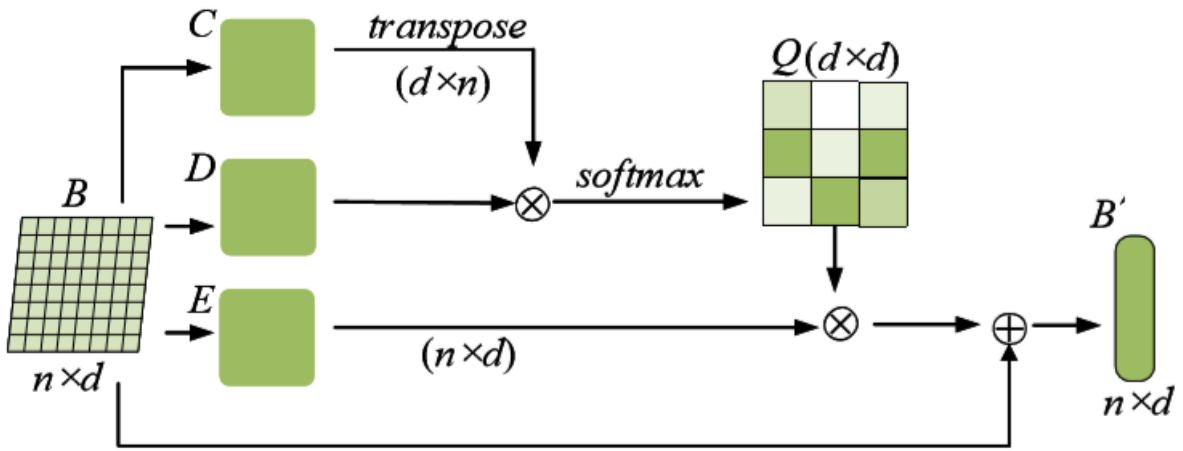


Fig.3 The position attention module.

$$q_{ij} = \frac{\exp(D_j \cdot C_i)}{\sum_{j=1}^{d} \exp(D_j \cdot C_i)}$$

$$B'_j = \alpha \sum_{i=1}^{d} (q_{ij} E_i) + B_j$$

Dual Attention-Guided GCN

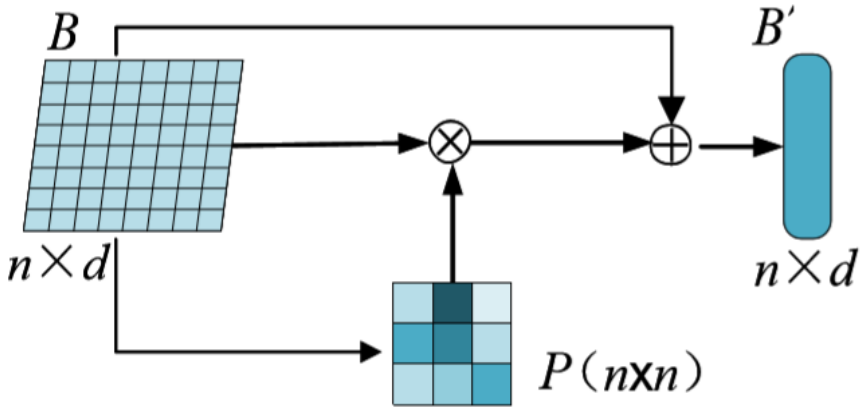


Fig.4 The relation attention module.

$$p_{ij} = \frac{\exp(P_i \cdot P_j)}{\sum_{j=1}^{n} \exp(P_i \cdot P_j)}$$

$$B_i' = \beta \sum_{j=1}^{n} (p_{ij} B_j) + B_i$$

Classification Reinforcement Module

We use distributional reinforcement learning to optimize the representation of relationship features.

$$r_{ij} = FFNN(h_{ei}, h_{ej}, h_{sent})$$

$$h_{sent} = f(B')$$

$$\mathbb{P}(r_{ij} \mid h_{ei}, h_{ej}, h_{sent}) = soft\max(MLP(r_{ij}))$$

$$Q(h,r) = \mathbb{E}R(h,r) + \gamma \mathbb{E}Q\left(h', r'\right)$$

$$Z(h,r) = R(h,r) + \gamma Z\left(h', r'\right)$$

$$\mathcal{L} = \sum_{s \in \mathcal{S}} \sum_{i \neq j} \log Z\left(h, r_{ij} \mid i, j, s\right)$$

Experiments

TABLE I
RESULTS ON TACRED DATASETS.

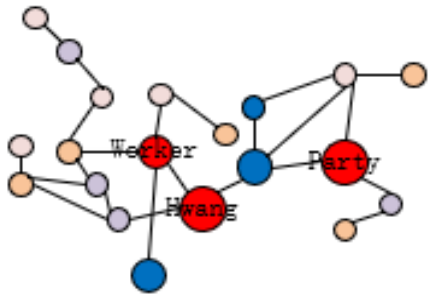
Model	P	R	F1
LR [2]	**73.5**	49.9	59.4
SDP-LSTM [20]	66.3	52.7	58.7
Tree-LSTM [21]	66.0	59.2	62.4
PA-LSTM [2]	65.7	64.5	65.1
GCN [6]	69.8	59.0	64.0
C-GCN [6]	69.9	63.3	66.4
AGGCN [9]	69.9	60.9	65.1
C-AGGCN [9]	71.8	66.4	69.0
DAGCN(ours)	70.1	63.5	66.8
C-DAGCN(ours)	72.6	**68.7**	**70.6**

TABLE II
RESULTS ON SEMEVAL DATASETS.

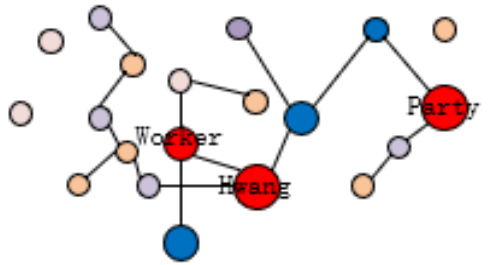
Model	F1
SVM [22]	82.2
SDP-LSTM [20]	83.7
SPTree [8]	84.4
PA-LSTM [2]	82.7
C-GCN [6]	84.8
C-AGGCN [9]	85.7
C-DAGCN(ours)	**86.9**

Experiments

Hwang, architect of the Pyongyang regime' s ideology of juche, was onec secretary of the ruling workers' Party and a tutor to current leader.

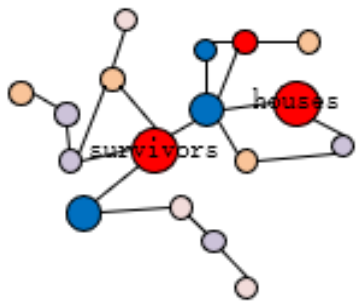


(a)

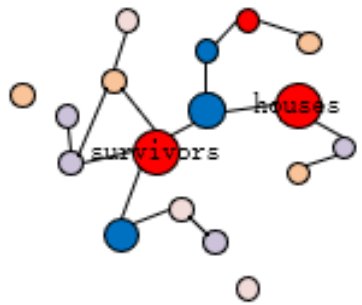


(b)

The government against the disaster and set up makeshift houses, ten million quake survivors moved into makeshift houses.



(c)



(d)

Fig. 5. Visualization of node association.

Experiments

(a) (b)

Fig. 6. Visualization of entity relationship classification.

Conclusion

Our contributions are summarized as follows.

We propose a dual attention graph neural network. By capturing a long-range of contextual dependency information and improves the discriminate ability of feature representations in relation extraction.

We propose a position attention module to learn the spatial correlation of features and also propose a relation attention module to capture the dependency information between nodes.

We propose a classification reinforcement module to optimize the representation of relationship features and improve the accuracy of relationship classification.

Our model achieved the new state-of-the-art performance on the TACRED and SemEval datasets.

Thank you.