Hongbin ZHANG, Quan CHEN, Weiwen ZHANG, 2022. Improving entity linking with two adaptive features. *Frontiers of Information Technology & Electronic Engineering*, 23(11):1620-1630. https://doi.org/10.1631/FITEE.2100495

Improving entity linking with two adaptive features

Key words: Entity Linking; Local model; Global model; Adaptive features; Entity type

Corresponding author: Weiwen ZHANG
E-mail: zhangww@gdut.edu.cn
ORCID: https://orcid.org/0000-0002-5098-6459

Motivation

1. Entity linking (EL) is essential for efficient information extraction and high-quality knowledge graph construction, which typically shows potential in downstream applications like Question Answering.
2. An EL system is composed of two models. Specifically, the local model can be extended to the global model. However, existing mainstream local models are unable to capture latent semantic information in the context around the mention.
3. Entity type information is usually added to the EL systems. Based on named entity recognition, for all mentions on the AIDA-train dataset, every entity type accounts for a different proportion: 18.40% for PER, 6.58% for GPE, 36.58% for ORG, and 38.44% for UNK. Thus, the training dataset for EL systems contains some uncertain entity type information.

Main idea

1. The first adaptive feature is applied in the local and global models, which can explore latent relationships between individuals from a context.
2. The second adaptive feature is used for entity type modeling, which can describe effective entity type information.
3. These two adaptive features are independent parts of our EL system, and can be linked naturally to work together to handle some uncertain entity type information.

Method

1. The local model is designed to calculate the local context similarity between context and candidates for a mention.

The global model is based on the local model, aiming to calculate dynamic context coherence between previously linked entities and candidates for a mention.

2. The first step explores latent information between the individuals from a context by the embedding matrix ***B*** and “FNN.”

The second step increases the influence of latent information by the embedding matrix ***C*** and “$\otimes$.”

Method (Cont'd)

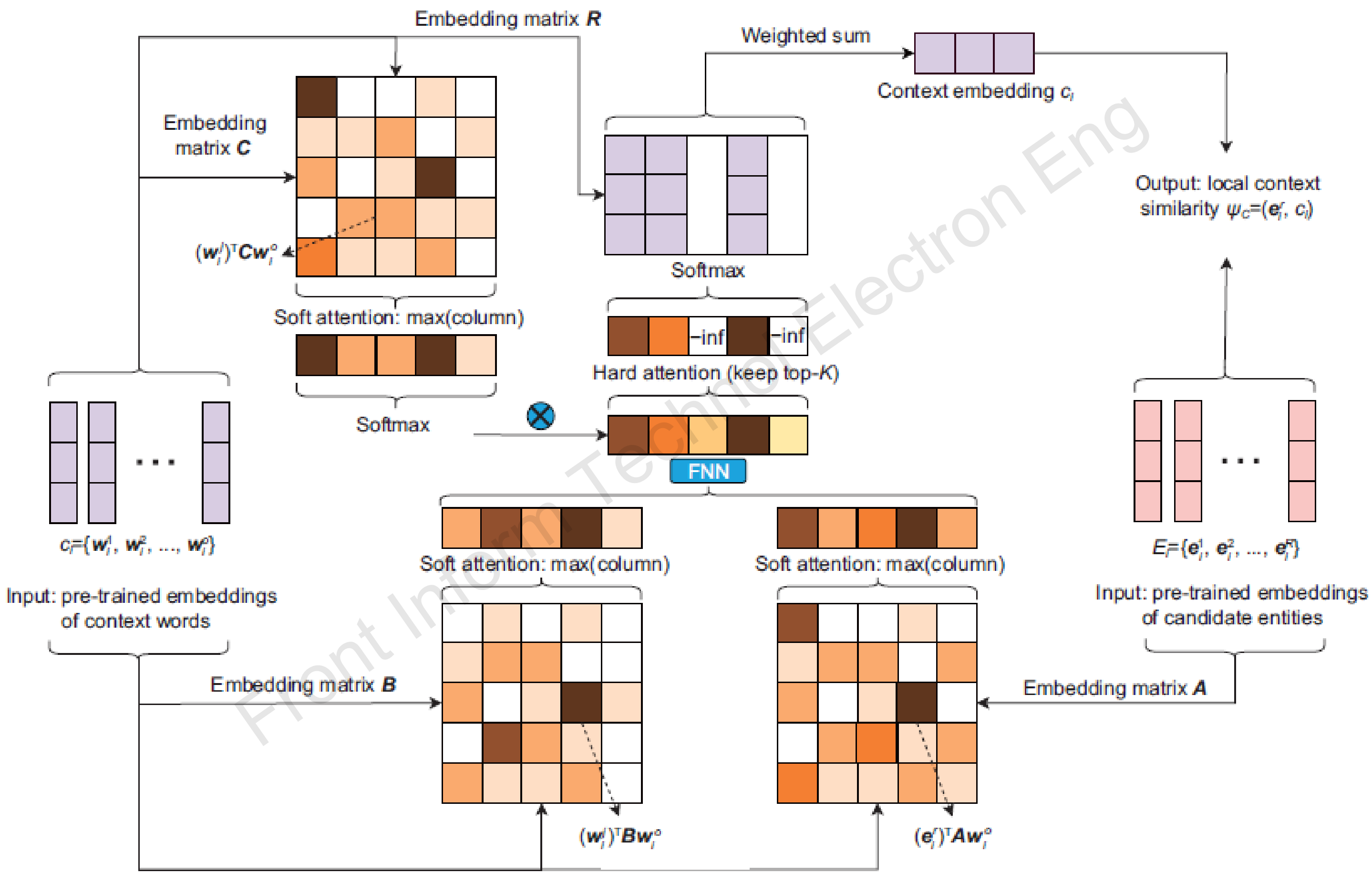
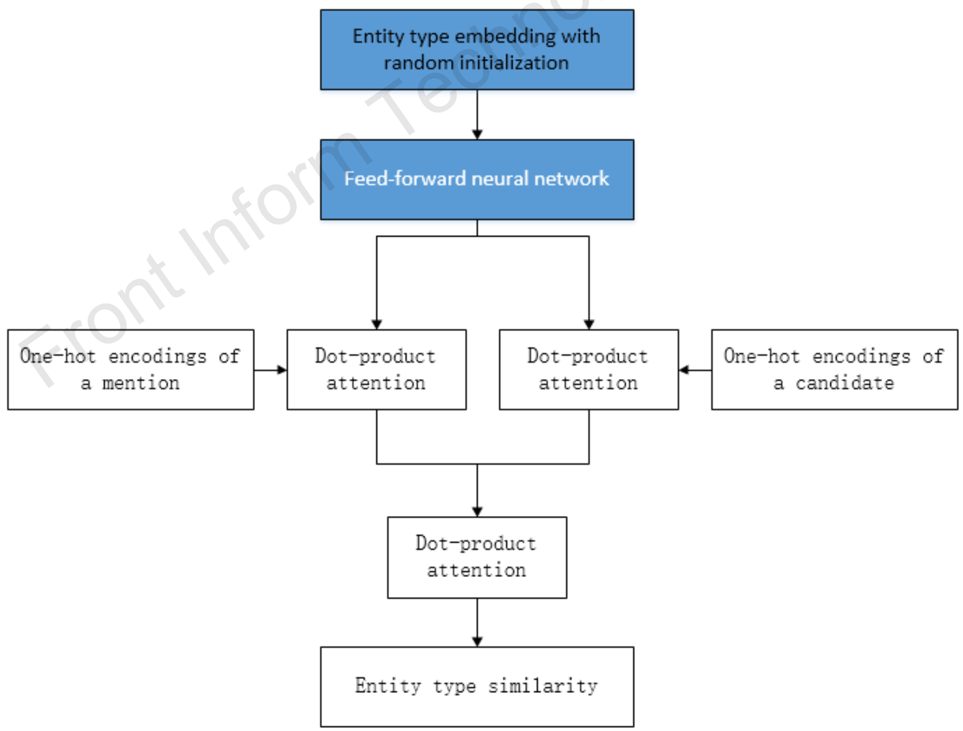


Fig. 1 Computational process of local context similarity, where the first adaptive feature is split into two parts that are respectively associated with the embedding matrices B and C. The first adaptive feature contains two operations, where "FNN" denotes the addition operation which is based on a two-layer feed-forward neural network and "$\otimes$" denotes the multiplication operation

Method (Cont'd)

3. Entity type modeling is applied to calculate entity type similarity between a mention and candidates.

Based on random initialization, four kinds of entity type embeddings are computed with the second adaptive features.



Major results

Our local model considers only the first adaptive features, but our EL system consists of two adaptive features.
The local model and the EL system were both trained on the AIDA-train dataset, and we validate them on the AIDA-A dataset.

Table 2 Micro F1 score for performance evaluation of local models and EL systems with adaptive features on the AIDA-B test set (adapted from Zhang et al. (2022), Copyright 2022, with permission from ACM)

Local model	Micro F1 score (%)
Prior (Medelyan et al., 2009)	71.90
Plato+star (Globerson et al., 2016)	87.90
Skip-gram (Yamada et al., 2016)	87.20
ETHZ-Attn (Yang et al., 2019)	90.88
BERT-Entity-Sim (Chen et al., 2020)	90.06
Knowledge-aware (Deng et al., 2020)	90.38
Our local model	**90.99**
EL system	**Micro F1 score (%)**
MulFocal-Att (Globerson et al., 2016)	91.00
Two-step (Yamada et al., 2016)	91.50
Deep-ED (Ganea and Hofmann, 2017)	92.22
Ment-norm (Le and Titov, 2018)	93.07
DCA (Yang et al., 2019)	93.73
BERT-Entity-Sim (Chen et al., 2020)	93.60
Knowledge-aware (Deng et al., 2020)	93.60
Our system	**94.20**

The best results are in bold

Major results (Cont'd)

Our EL system has been trained on the AIDA-train dataset, and we test it on out-domain datasets.

Table 3 Micro F1 score for performance evaluation of EL systems on out-domain datasets (adapted from Zhang et al. (2022), Copyright 2022, with permission from ACM)

EL system	Micro F1 score (%)					Average score (%)
	MSNBC	AQUAINT	ACE2004	CWEB	WIKI	
WNED (Guo and Barbosa, 2018)	92.00	87.00	88.00	77.00	**84.50**	85.70
Deep-ED (Ganea and Hofmann, 2017)	93.70	88.50	88.50	**77.90**	77.50	85.22
Ment-norm (Le and Titov, 2018)	93.90	88.30	89.90	77.50	78.00	85.52
DCA (Yang et al., 2019)	93.80	88.25	90.14	75.59	78.84	85.32
CoSimTC (Xin et al., 2019)	94.16	**90.90**	**92.92**	76.96	75.02	85.99
FGS2EE (Hou et al., 2020)	94.26	88.47	90.70	77.41	77.66	85.70
Graph-based-EL (Deng et al., 2020)	**94.41**	89.23	90.54	76.64	78.20	85.80
Our system	**94.41**	90.21	90.54	76.97	78.16	**86.06**

The best results are in bold. If a result has a confidence interval, we determine the median from it

Major results (Cont'd)

Ablation analysis, where the first adaptive features have two steps

Table 4 Micro F1 score for EL systems of ablation analysis on out-domain datasets

EL system	Micro F1 score (%)					Average score (%)
	MSNBC	AQUAINT	ACE2004	CWEB	WIKI	
only_AP_first_model	**94.72**	88.39	90.14	75.64	77.46	85.27
only_AP_model	94.41	88.25	**91.35**	75.70	**78.72**	**85.69**
only_AP_type	93.96	**89.51**	89.34	**76.23**	78.41	85.49

The best results are in bold

Conclusions

1. Two adaptive features can enable the EL system to capture latent information, properly adjust four kinds of entity type embedding, and handle some uncertain entity type information.

2. Two adaptive features can be integrated in the local model, global model, and entity type modeling of EL systems.

3. Our system achieved the best performance on the AIDA-B and MSNBC datasets. It also achieved the best average performance on out-domain datasets.

Hongbin ZHANG received his M.E. degree in software engineering from Guangdong University of Technology in 2022. He is currently pursuing the Ph.D. degree in computer science and technology at School of Computer Science and Technology, Guangdong University of Technology. His research interests include natural language processing and deep learning.