中国科学院信息工程研究所
INSTITUTE OF INFORMATION ENGINEERING,CAS

Improving Knowledge Graph Embedding Using Simple Constraints

Boyang Ding[1,2], Quan Wang[1,2,3*], Bin Wang[1,2], Li Guo[1,2]

[1]Institute of Information Engineering, Chinese Academy of Sciences
[2]School of Cyber Security, University of Chinese Academy of Sciences
[3]State Key Laboratory of Information Security, Chinese Academy of Sciences

Code and data available at https://github.com/iieir-km/ComplEx-NNE_AER

Outline

- Background
- Our Approach
- Experiments
- Summary

Outline

- Background
- Our Approach
- Experiments
- Summary

Knowledge Graph Embedding

- Knowledge graph
  - Large scale triple set $\{(head\ entity, relation, tail\ entity)\}$
- Key idea
  - Embed entities and relations into continuous vector spaces
  - Simplify the manipulation + preserve the inherent graph structure
- The recent tendency
  - Incorporating external data (lack the universality)
  - More complicated models (decreasing the computation efficiency)

Outline

- Background
- Our Approach
- Experiments
- Summary

Overview

- **Imposing simple constraints on KG embedding**
  - Non-Negativity Constraints for Entities
  - Approximate Entailment Constraints for Relations
- **Benefits**
  - ✓ Low computation complexity
  - ✓ More predictive embedding
  - ✓ More interpretable embedding

Non-Negative Entity Embedding

- Non-Negative Entity Embedding:
  - Require entity embedding lie in a non-negative subspace
  - More sparse representation, batter interpretability
- Intuition
  - Uneconomical to store all negative semantics

Positive semantics for ***"Pairs"***
- The capital of France

Negative semantics for ***"Pairs"***
- Not a capital of China
- Not a capital of US
- Not a capital of Russia

…

Approximate Entailment for Relations

- Approximate relation entailment:
  - $r_p \xrightarrow{\lambda} r_q$: relation r_p entails r_q with confidence level λ
    - e.g $\texttt{bornIn} \xrightarrow{0.8} \texttt{nationality}$
  - Strong prior belief for relation embedding
  - Entailments can be obtained automatically via rule mining software

Overall Model

- Basic embedding model:
  - ComplEx: representing entity e and realtion r as complex vector

$$\mathbf{x} = \mathrm{Re}(\mathbf{x}) + \mathrm{Im}(\mathbf{x})i$$

- Modeling non-negativity:

$$0 \leq \mathrm{Re}(\mathbf{e}), \mathrm{Im}(\mathbf{e}) \leq 1 \quad \forall e \in \mathcal{E}$$

- Modeling approximate entailment:

Using order embedding to model entailment

$$\lambda\boxed{\left(\mathrm{Re}(\mathbf{r}_p) - \mathrm{Re}(\mathbf{r}_q)\right)} \leq \boxed{\xi_1}$$
$$\lambda\boxed{\left|\mathrm{Im}(\mathbf{r}_p) - \mathrm{Im}(\mathbf{r}_q)\right|} \leq \boxed{\xi_2}$$

Using slackness variable to model approximation

- Complexity

$$\mathcal{O}(sd + \bar{n}d + td)$$

  - d dimensionality
  - s # triples in a mini-batch
  - $\bar{n}$ # entities in a mini-batch
  - $t \ll s$ # entailments

Have the complexity of the same order with basic embedding model!

Outline

- Background
- Our Approach
- Experiments
- Summary

Link Prediction

- To complete a triple (e_i, r_k, e_j) with e_i or e_j missing

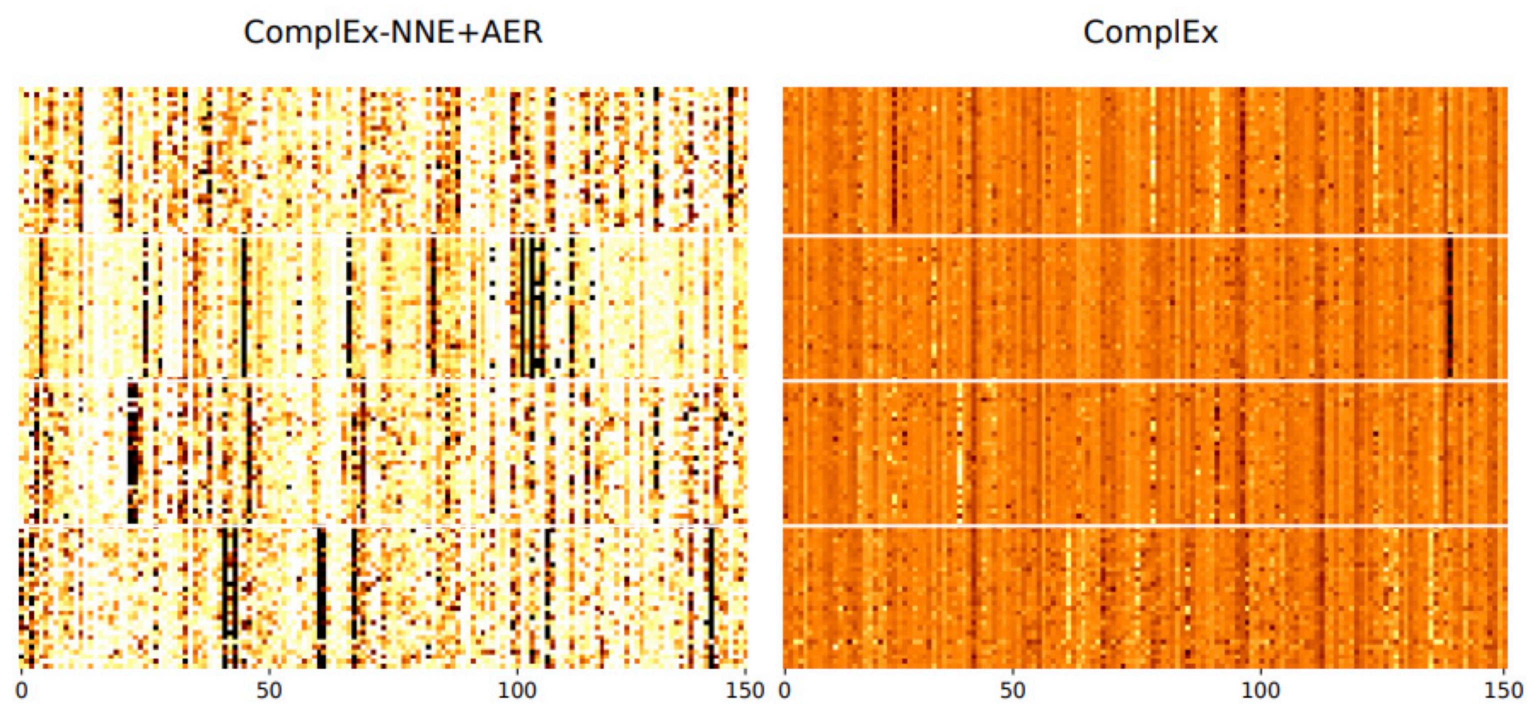
	WN18			FB15K			DB100K		
	MRR	HITS@1	HITS@3	MRR	HITS@1	HITS@3	MRR	HITS@1	HITS@3
TransE(2013)	0.454	0.089	0.823	0.380	0.231	0.472	0.111	0.016	0.164
DistMult(2015)	0.822	0.728	0.914	0.654	0.546	0.733	0.233	0.115	0.301
HolE(2016)	0.938	0.930	0.945	0.524	0.402	0.613	0.260	0.182	0.309
ComplEx(2016)	0.941	0.936	0.945	0.692	0.599	0.759	0.242	0.126	0.312
ANALOGY(2017)	0.942	0.939	0.944	0.725	0.646	0.785	0.252	0.143	0.323
RUGE(2018)	—	—	—	0.768	0.703	0.815	0.246	0.129	0.325
ComplEx^{R}(2017)	0.940	—	0.943	—	—	—	0.253	0.167	0.294
R-GCN(2017)	0.814	0.686	0.928	0.651	0.541	0.736	—	—	—
R-GCN+(2017)	0.819	0.697	0.929	0.696	0.601	0.760	—	—	—
ConvE(2018)	0.942	0.935	**0.947**	0.745	0.670	0.801	—	—	—
Single DistMult(2017)	0.797	—	—	0.798	—	—	—	—	—
ComplEx-NNE	0.941	0.937	0.944	0.727	0.659	0.772	0.298	0.229	0.330
ComplEx-NNE+AER	**0.943**	**0.940**	0.945	**0.803**	**0.761**	**0.831**	**0.306**	**0.244**	**0.334**

Basic embedding model | Incorporating logical rules | Complicated models

ComplEx-NNE(ComplEx with Non-Negative Entities) and ComplEx-NNE-AER(ComplEx with Approximate Entailment for Relations) outperform compared methods

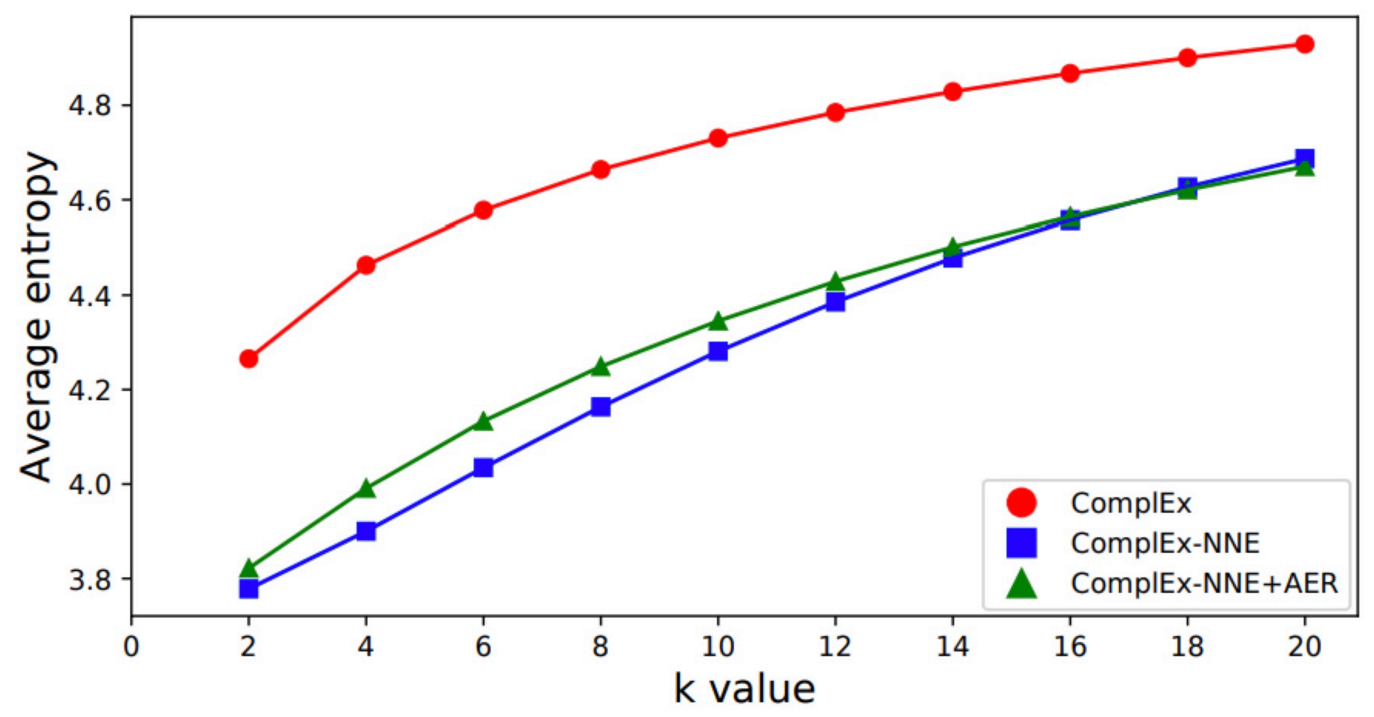
Visualization Analysis of Entity

❖ Semantic distribution



Each row is a type of randomly selected entities' embedding. The darker the color the greater the value

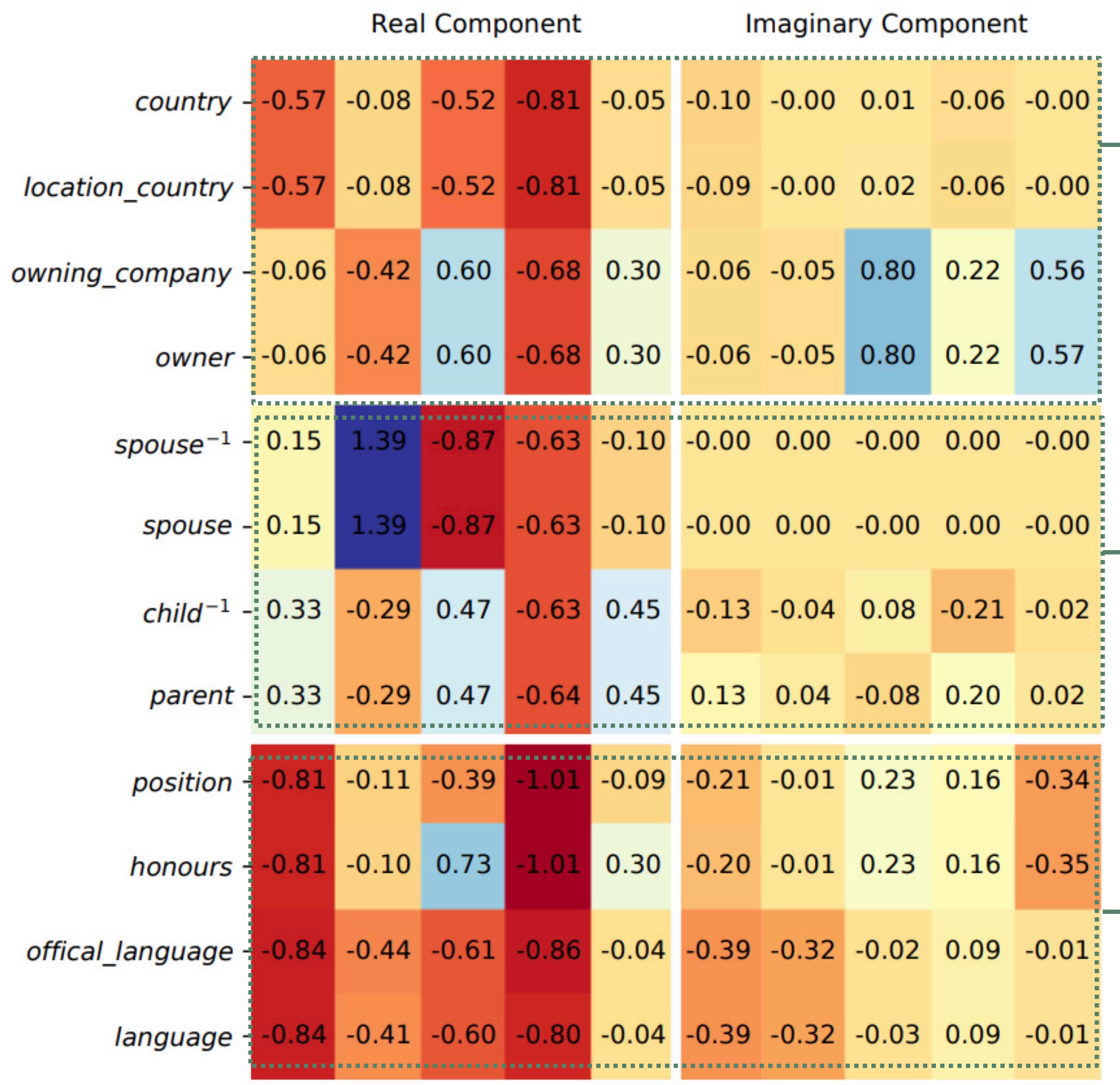
❖ Semantic purity



Measuring the semantic purity of each dimension via calculating the entropy of top ranked entities' type, the lower the entropy the higher the purity

Visualization Analysis of Relation

- Relation embedding entailment visualization



Equivalence should satisfy:

$$\mathrm{Re}(\mathbf{r}_p) = \mathrm{Re}(\mathbf{r}_q)$$
$$\mathrm{Im}(\mathbf{r}_p) = \mathrm{Im}(\mathbf{r}_q)$$

Inversion should satisfy:

$$\mathrm{Re}(\mathbf{r}_p) = \mathrm{Re}(\mathbf{r}_q)$$
$$\mathrm{Im}(\mathbf{r}_p) = -\mathrm{Im}(\mathbf{r}_q)$$

Other entailment should satisfy:

$$\mathrm{Re}(\mathbf{r}_p) \leq \mathrm{Re}(\mathbf{r}_q)$$
$$\mathrm{Im}(\mathbf{r}_p) = \mathrm{Im}(\mathbf{r}_q)$$

Outline

- Background
- Our Approach
- Experiments
- Summary

Takeaway Message

- Simple constraints:
  - Non-negativity constraints for entity embedding
  - Approximate entailment constraints for relation embedding
- Experiment results:
  - Efficient
  - Effective
  - Interpretable

Code and data available at https://github.com/iieir-km/ComplEx-NNE_AER

Thank you!

Q&A

dingboyang@iie.ac.cn