浙江大学 计算机科学与技术学院
COLLEGE OF COMPUTER SCIENCE AND TECHNOLOGY
ZHEJIANG UNIVERSITY

浙江大学 杭州国际科创中心
ZJU-Hangzhou Global Scientific and Technological Innovation Center

A · Z · F · T

Websoft

Understanding and Improving Knowledge Graph Embedding for Entity Alignment

Lingbing Guo*, Qiang Zhang*, Zequn Sun, Mingyang Chen, Wei Hu, Huajun Chen

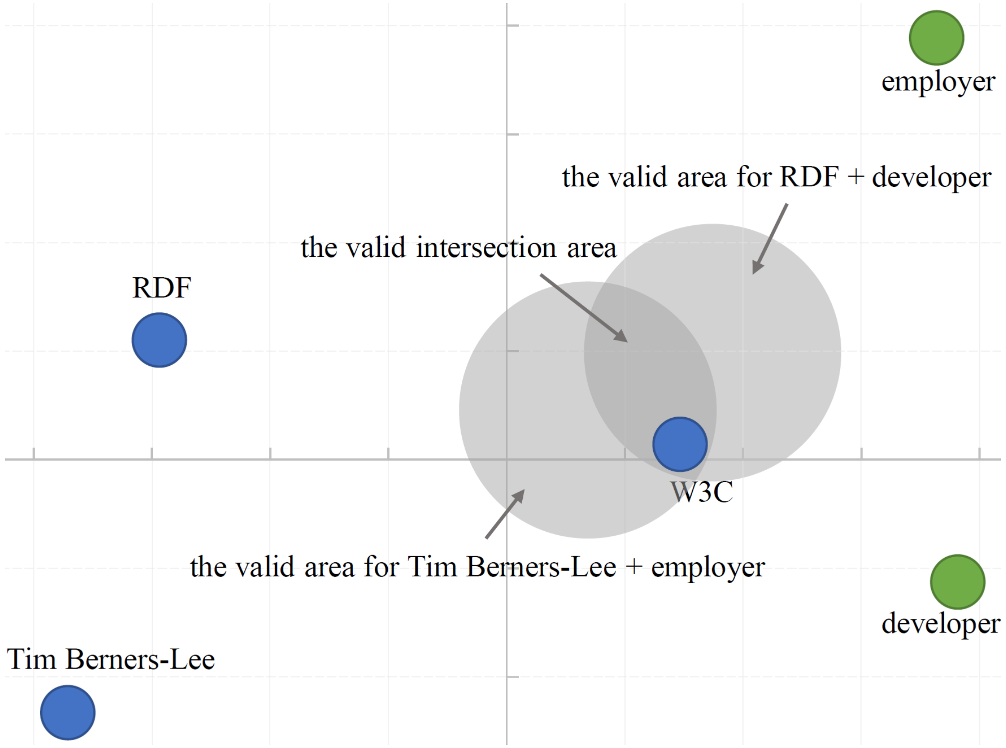
ICML 2022

*: Equal contribution
github.com/guolingbing/NeoEA

Background

The score function of TransE:

$$||\mathbf{e}_i + \mathbf{r} - \mathbf{e}_j|| \leq \lambda,$$



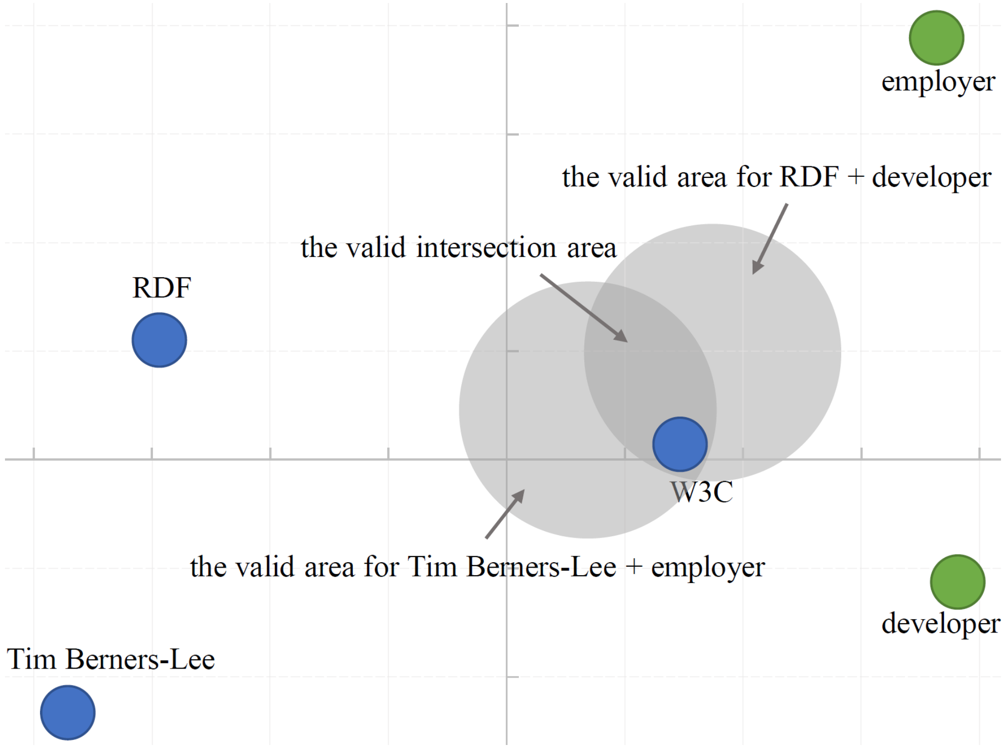
KG Embedding

The score function of TransE:

$$||\mathbf{e}_i + \mathbf{r} - \mathbf{e}_j|| \leq \lambda,$$

The entity **W3C** is constrained by two triples:

(Tim Berners-Lee, employer, **W3C**)
(RDF, developer, **W3C**)



KG Embedding

The score function of TransE:

$$||\mathbf{e}_i + \mathbf{r} - \mathbf{e}_j|| \leq \lambda,$$

The entity **W3C** is constrained by two triples:

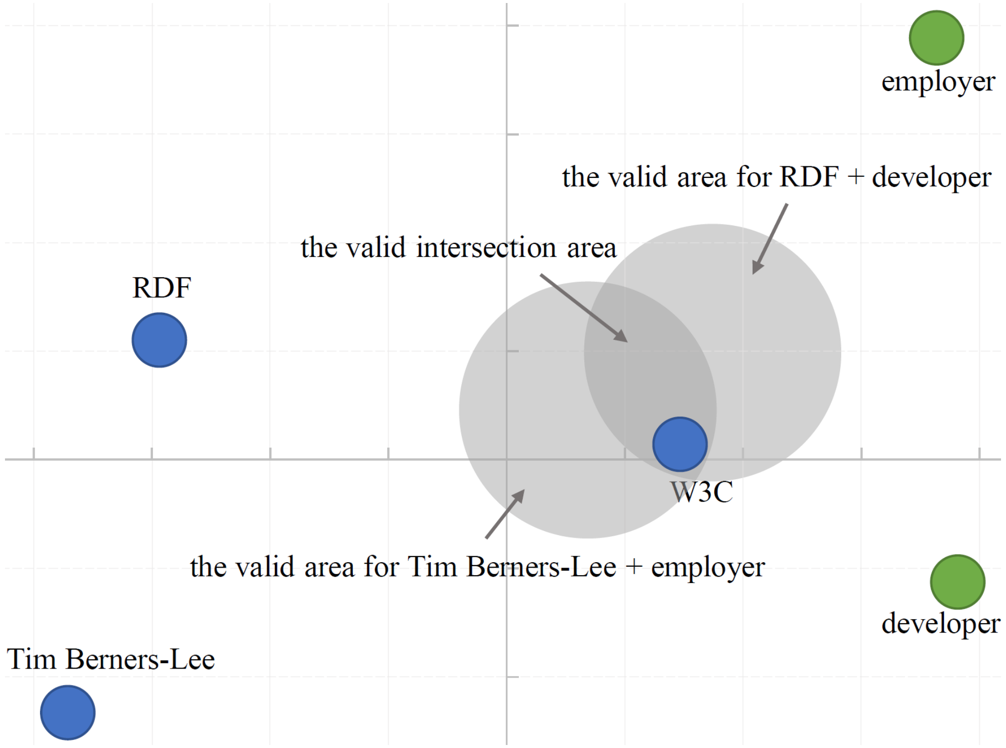
(Tim Berners-Lee, employer, **W3C**)
(RDF, developer, **W3C**)

The embedding of **W3C** is constrained by the intersection area of two circles,
with the **centers**:

Tim Berners-Lee + employer
RDF + developer

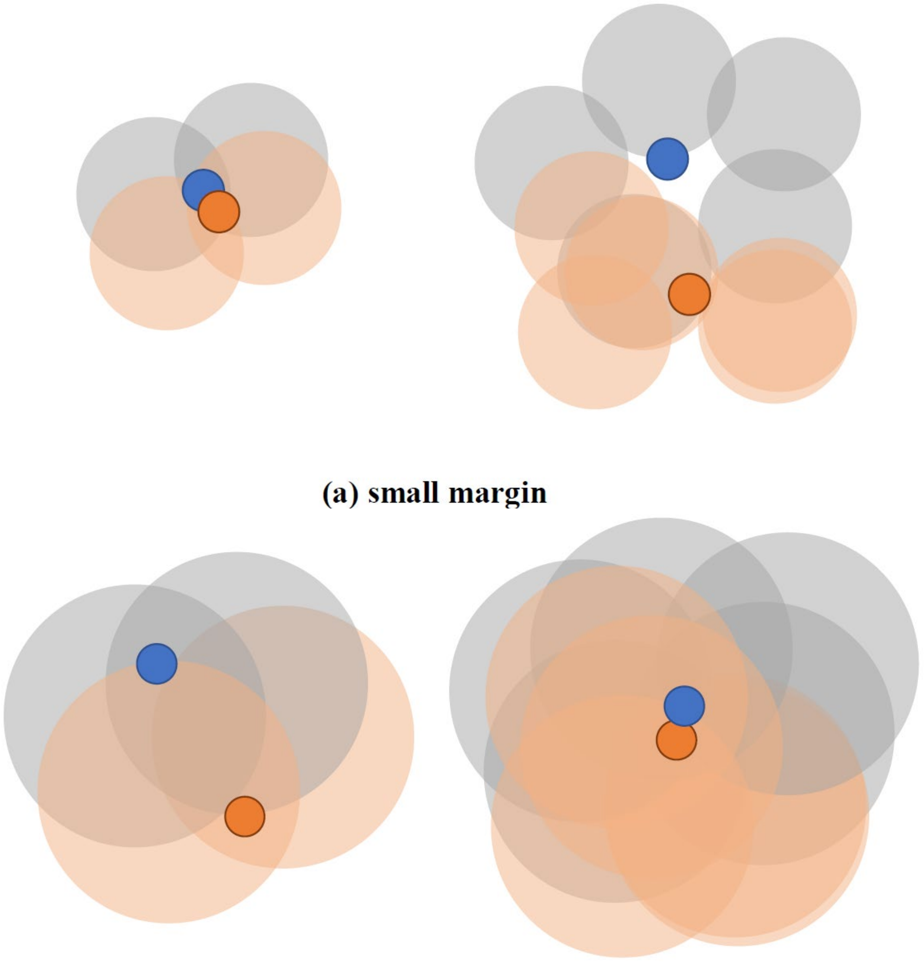
and radii:

$$\lambda$$



Embedding-based Entity Alignment (EEA)

With a small number of **entity pairs as anchors**, we can **learn the entity embeddings** of two KGs in **a unified space**.



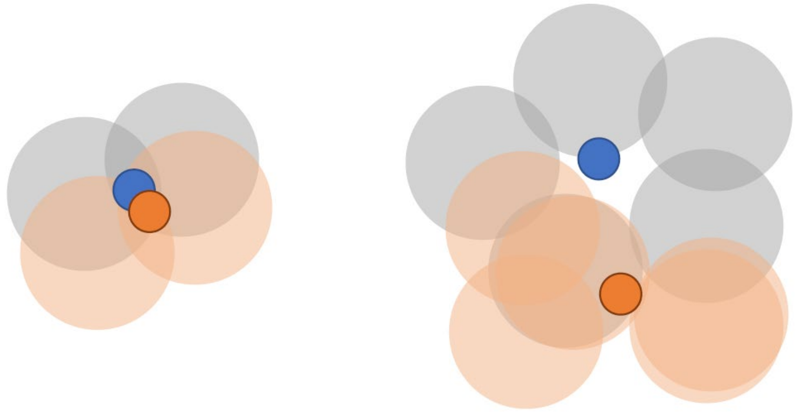
(a) small margin

(b) large margin

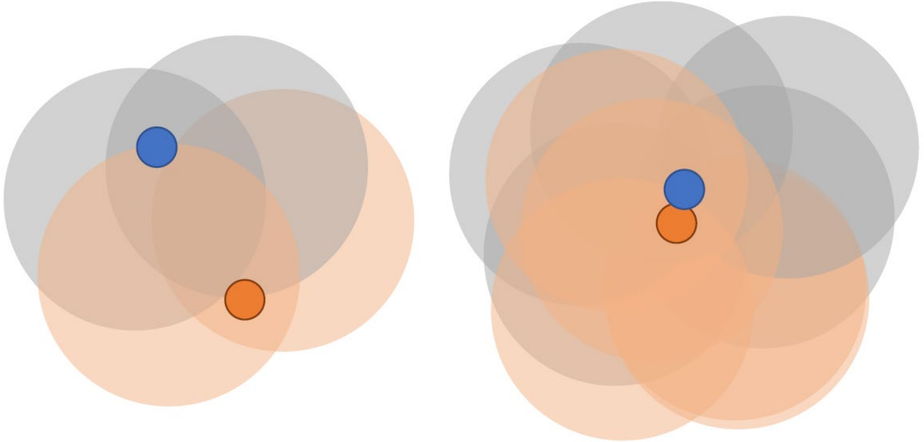
Embedding-based Entity Alignment

With a small number of **entity pairs as anchors**, we can **learn the entity embeddings** of two KGs in **a unified space**.

The blue and orange nodes are entity pairs referred to the same object in real-world.



(a) small margin



(b) large margin

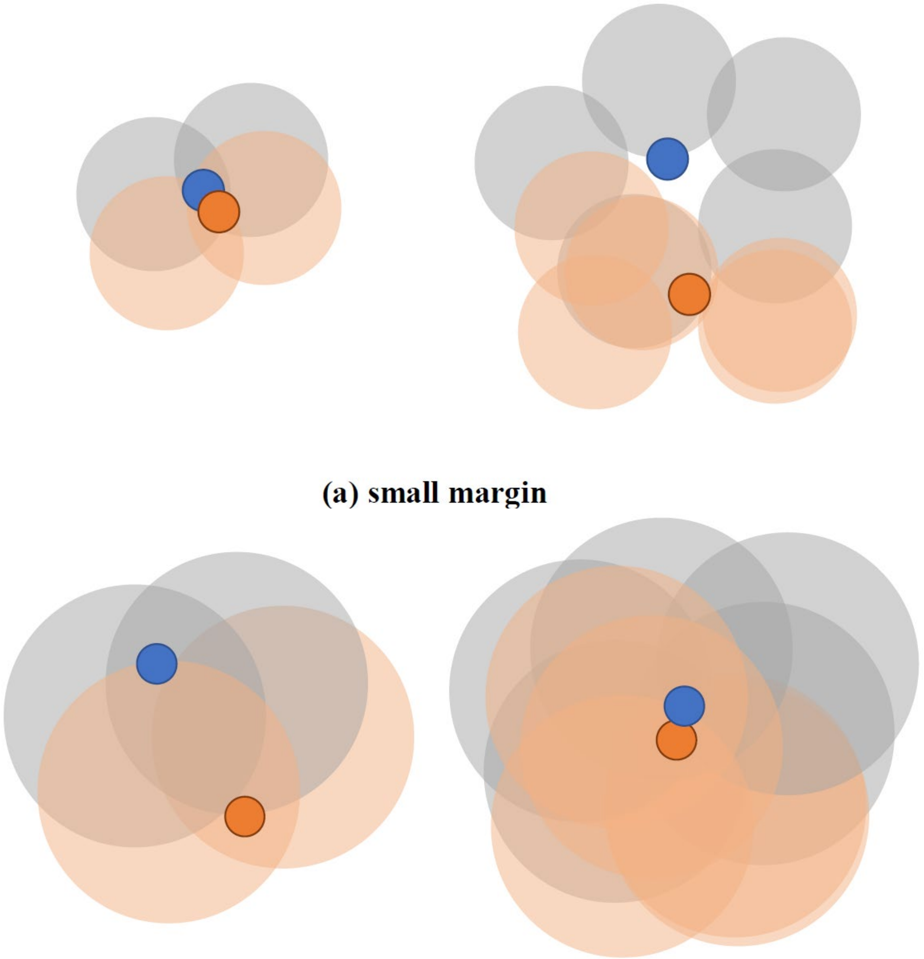
Embedding-based Entity Alignment

With a small number of **entity pairs as anchors**, we can **learn the entity embeddings** of two KGs in **a unified space**.

The blue and orange nodes are entity pairs referred to the same object in real-world.

A small margin (figure a) gives tight constraints, but fails to model entities with more triples.

A large margin can model the entities with more triples, but the bound is too loose for the long-tail entities.



(a) small margin

(b) large margin

Proposition 2.2 (Discrepancy Bound). *The embedding difference of two potentially aligned entities* (e_x^1, e_y^2) *is bound by* ϵ*, which is proportional to the hyper-parameter* λ*:*

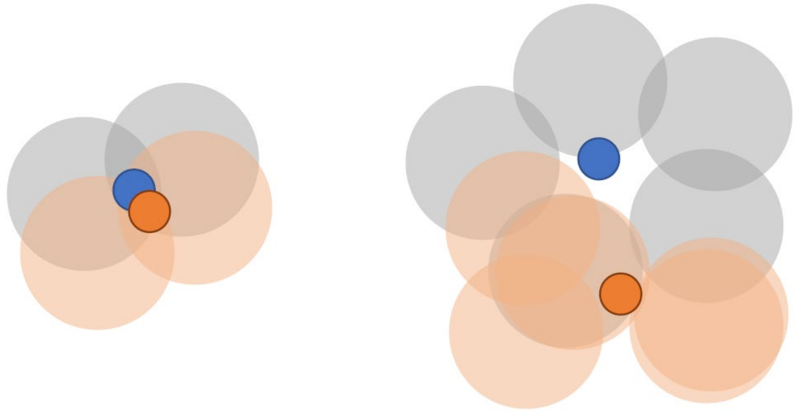
$$\exists \epsilon \propto \lambda, ||\mathbf{e}_x^1 - \mathbf{e}_y^2|| \leq \epsilon. \quad (4)$$

Considering Additional Constraints

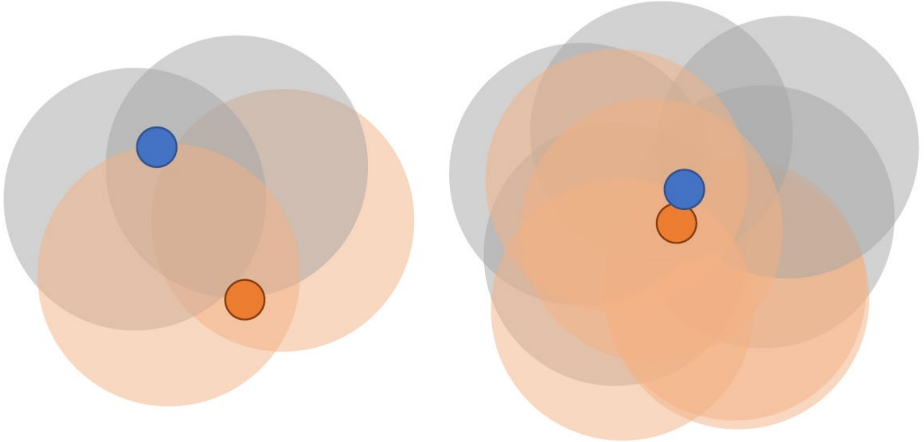
Align the embedding distributions of two KGs:

- A discriminator to discriminate the entities from different KG.
- The embeddings from different KGs try to fool the discriminator.
- Use the method in Domain Adaptation area, an empirical Wasserstein distance based loss is:

$$\mathcal{L}_{\mathbb{A}_E} = \mathbb{E}_{\mathbb{A}_E^1}[f_w(\mathbf{e})] - \mathbb{E}_{\mathbb{A}_E^2}[f_w(\mathbf{e})], \qquad (12)$$



(a) small margin

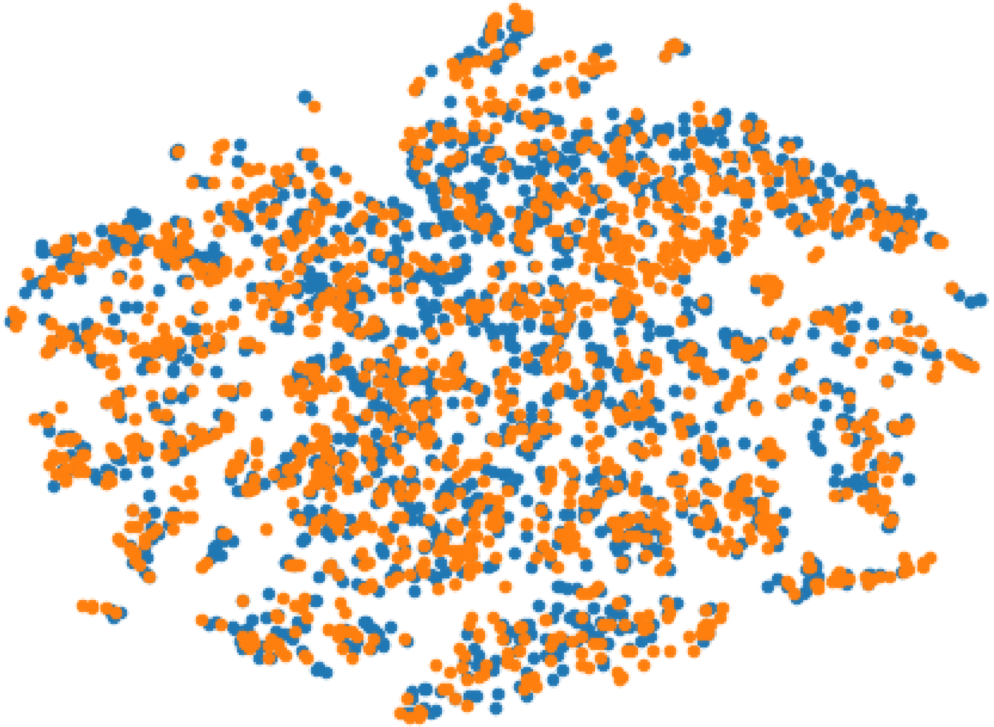


(b) large margin

Conditional Distributions

The direct solution may be less helpful for KG embedding.

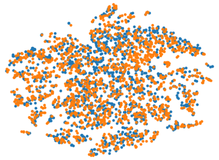
a. The embeddings usually tend to uniformly distribute over the space, leading to that the two distributions are already ``aligned''.



Conditional Distributions

The direct solution may be less helpful for KG embedding.

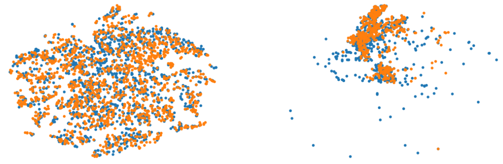
a. The embeddings usually tend to uniformly distribute over the space, leading to that the two distributions are already ``aligned''.
b. The entities conditioning on some relations, their distributions are actually different.



Conditional Distributions

The direct solution may be less helpful for KG embedding.

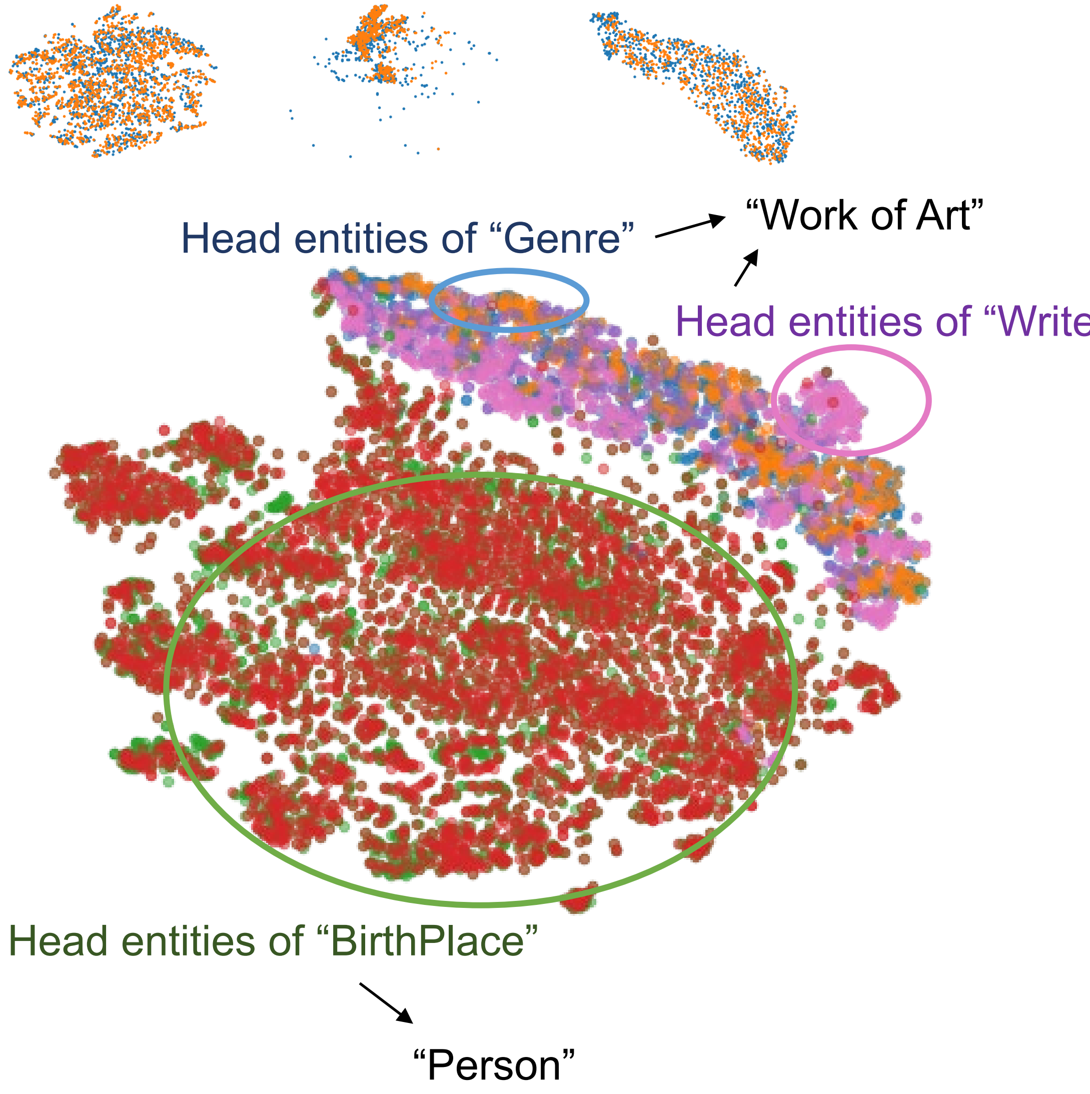
a. The embeddings usually tend to uniformly distribute over the space, leading to that the two distributions are already ``aligned''.
b. The entities conditioning on some relations, their distributions are discriminative.
c. Learn to align the conditional distributions to reduce the discrepancy.



Conditional Distributions

The direct solution may be less helpful for KG embedding.

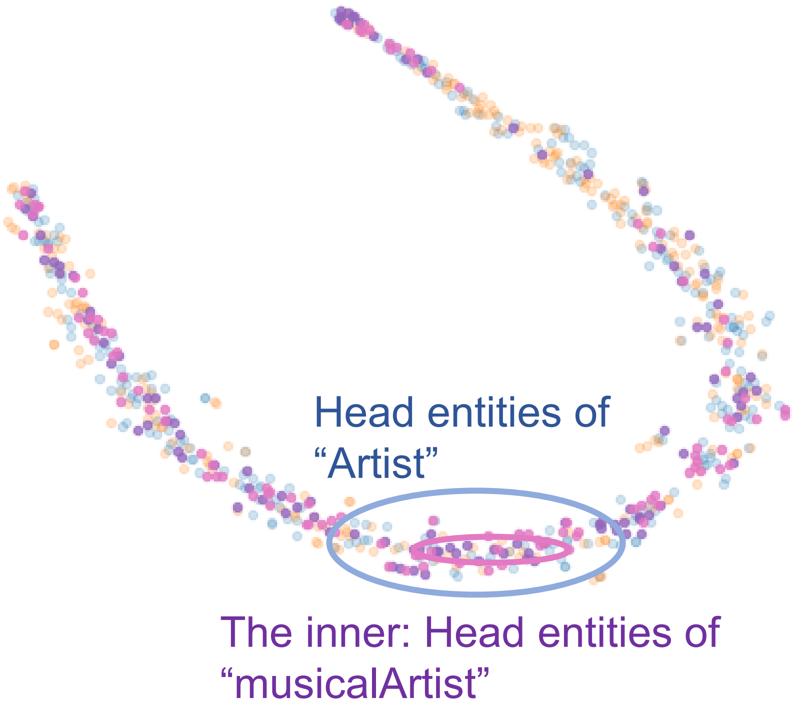
a. The embeddings usually tend to uniformly distribute over the space, leading to that the two distributions are already ``aligned''.
b. The entities conditioning on some relations, their distributions are discriminative.
c. Learns to align the conditional distributions to reduce the discrepancy.
d. Some implicit ontology-level information also be captured



Conditional Distributions

The direct solution may be less helpful for KG embedding.

a. The embeddings usually tend to uniformly distribute over the space, leading to that the two distributions are already ``aligned''.
b. If we condition the entities on some relations, their distributions are discriminative.
c. Learns to align the conditioned distributions to reduce the discrepancy.
d. Some implicit ontology-level information also be captured
e. The triple embeddings conditioning on relation embeddings.



Conditional Distributions

The direct solution may be less helpful for KG embedding.

a. The embeddings usually tend to uniformly distribute over the space, leading to that the two distributions are already ``aligned''.
b. If we condition the entities on some relations, their distributions are discriminative.
c. Learns to align the conditioned distributions to reduce the discrepancy.
d. Some implicit ontology-level information also be captured
e. Condition the triple embeddings on relation embeddings.

Proposition 3.2 (Expressiveness)**.** *Aligning the conditional neural axioms minimizes the embedding discrepancy of two KGs at the ontology level, without the need of type/class information.*

Experiments

- All baseline methods gained significant improvement.

- The training time did not increase too much, even compared with the fast baseline method.

Table 1: Results on V1 datasets (5-fold cross-validation).

Models	EN-FR			EN-DE			D-W			D-Y		
	H@1	H@5	MRR	H@1	H@5	MRR	H@1	H@5	MRR	H@1	H@5	MRR
BootEA (Sun et al., 2018)	.507	.718	.603	.675	**.820**	**.740**	.572	.744	.649	.739	.849	.788
BootEA + NeoEA	**.521**	**.733**	**.617**	**.676**	**.820**	**.740**	**.579**	**.753**	**.658**	**.756**	**.859**	**.797**
SEA (Pei et al., 2019a)	.280	.530	.397	.530	.718	.617	.360	.572	.458	.500	.706	.591
SEA + NeoEA	**.320**	**.584**	**.443**	**.586**	**.766**	**.668**	**.389**	**.608**	**.490**	**.549**	**.752**	**.638**
RSN (Guo et al., 2019)	.393	.595	.487	.587	.752	.662	.441	.615	.521	.514	.655	.580
RSN + NeoEA	**.399**	**.597**	**.490**	**.600**	**.759**	**.673**	**.450**	**.624**	**.530**	**.522**	**.663**	**.588**
RDGCN (Wu et al., 2019)	.755	.854	.800	.830	.895	.859	.515	.669	.584	.931	.969	.949
RDGCN + NeoEA	**.775**	**.868**	**.817**	**.846**	**.908**	**.874**	**.527**	**.671**	**.592**	**.941**	**.972**	**.955**

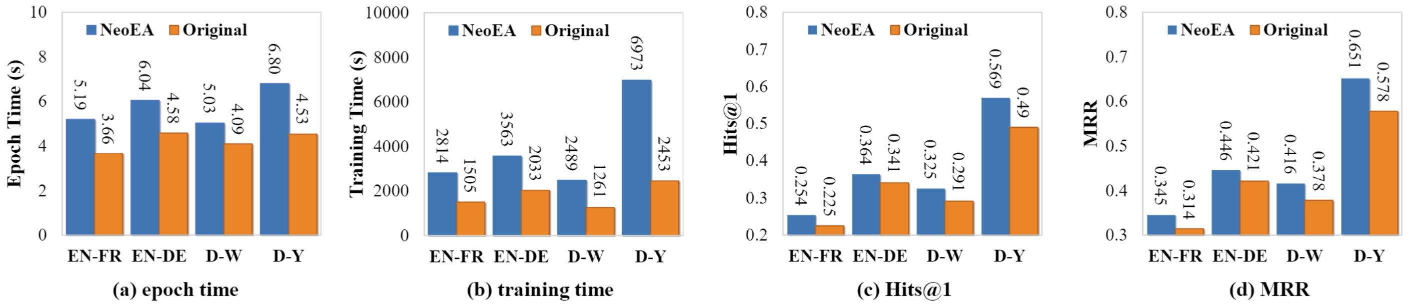
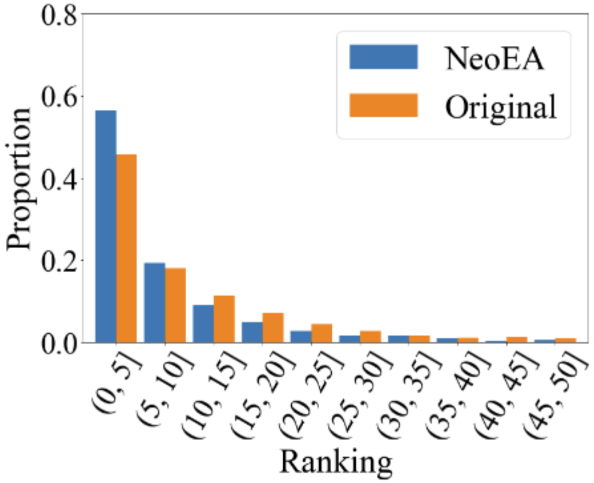
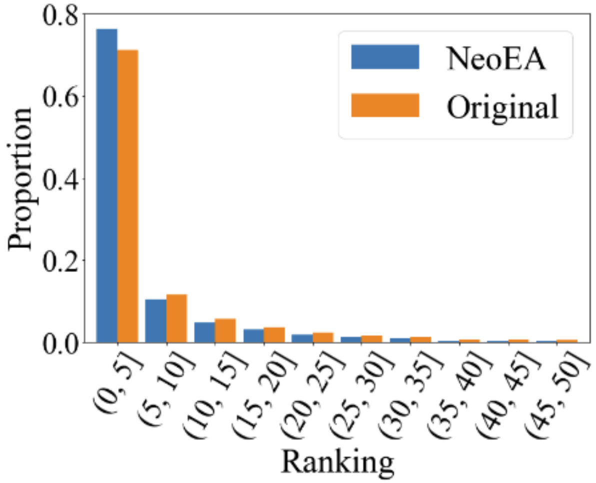


Figure 5: Results on OpenEA 100K datasets, with the fastest EEA model SEA as baseline.

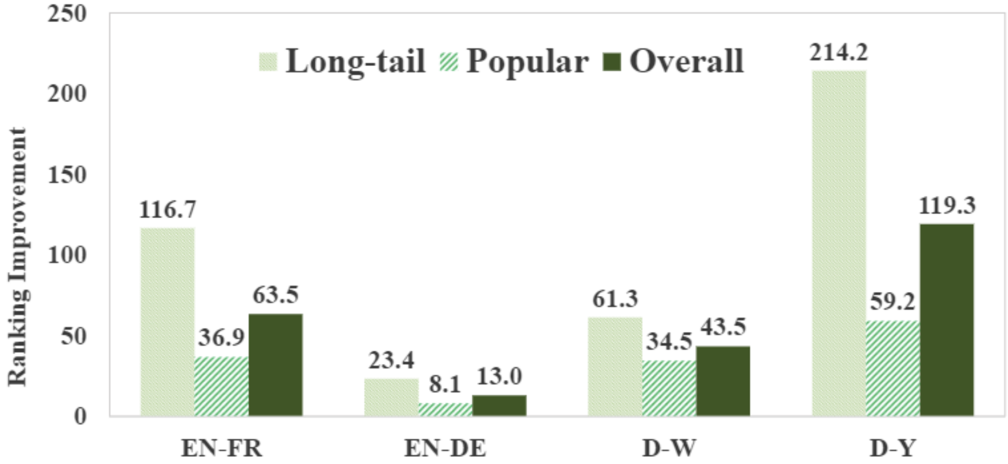
Experiments



(a) long-tail entities (EN-FR)



(b) popular entities (EN-FR)



(c) average ranking improvement on four datasets

- The long-tail entities (less constrained ones) gained significantly more improvement.
- On most datasets, around three times than popular entities.

Conclusion

- Analyzing the current EEA methods theoretically.
- A new approach to learn KG embeddings for entity alignment, capturing the implicit ontology-level information.
- Extensive experiments demonstrate consistent and significant improvement overall all baseline methods.

Thanks for your attention!

- Code and datasets are available at https://github.com/guolingbing/NeoEA

- This work is funded by NSFCU19B2027/NSFC91846204,
- National Key R&D Program of China (Funding No.SQ2018YFC000004),
- Zhejiang Provincial Natural Science Foundation of China (No.LGG22F030011).