Supplementary Materials: LoginMEA

Taoyu Su[a,b], Xinghua Zhang[a,b], Jiawei Sheng[a,*], Zhenyu Zhang[c] and Tingwen Liu[a,b]

[a]Institute of Information Engineering, Chinese Academy of Sciences, Beijing, China
[b]School of Cyber Security, University of Chinese Academy of Sciences, Beijing, China
[c]Baidu Inc.
{sutaoyu, zhangxinghua, shengjiawei, liutingwen}@iie.ac.cn, zhangzhenyu07@baidu.com

1 Supplementary Details

1.1 Datasets details

In our experiments, we use two types of multi-modal EA datasets. (1) Cross-KG datasets: we select FB15K-DB15K and FB15K-YAGO15K public datasets, which are deemed as the most typical datasets in multi-modal entity alignment tasks built-in [3]. FB15K is a representative subset extracted from the Freebase knowledge base. Aiming to maintain an approximate entity number of FB15K, DB15K from DBpedia, and YAGO15K from YAGO are mainly selected based on the entities aligned with FB15K. (1) Bilingual datasets: DBP15k is a widely used cross-lingual EA benchmark. It consists of four language-specific knowledge graphs from DBpedia and includes three bilingual entity alignment settings: French-English (FR-EN), Japanese-English (JA-EN), and Chinese-English (ZH-EN). Additionally, DBpedia has released images for the English, French, and Japanese versions. Since Chinese images are not released in DBpedia, EVA [2] extracted them from the raw Chinese Wikipedia dump with the same process as described by DBpedia[1] The details of all multi-modal EA datasets are listed in 1.

Table 1. Statistics of the Datasets (Ent.→Entity, Rel.→Relation, Rel tr.→Relation triple, Attr.→Attribute, Attr tr.→Attribute triple).

Dataset	KG	#Ent.	#Rel.	#Rel tr.	#Attr.	#Attr tr.	#Image	#EA pairs
FB15K-DB15K	FB15K	14,951	1,345	592,213	116	29,395	13,444	12,846
	DB15K	12,842	279	89,197	225	48,080	12,837	
FB15K-YAGO15K	FB15K	14,951	1,345	592,213	116	29,395	13,444	11,199
	YAGO15K	15,404	32	122,886	7	23,532	11,194	
DBP15K$_{\text{ZH-EN}}$	ZH (Chinese)	19,388	1,701	70,414	8,111	248,035	15,912	15,000
	EN (English)	19,572	1,323	95,142	7,173	343,218	14,125	
DBP15K$_{\text{JA-EN}}$	JA (Japanese)	19,814	1,299	77,214	5,882	248,991	12,739	15,000
	EN (English)	19,780	1,153	93,484	6,066	320,616	13,741	
DBP15K$_{\text{FR-EN}}$	FR (French)	19,661	903	105,998	4,547	273,825	14,174	15,000
	EN (English)	19,993	1,208	115,722	6,422	351,094	13,858	

1.2 Metric Details

To evaluate our IBMEA approach, we adopt the classical rank-based evaluation protocol of knowledge graph entity alignment. The following metrics are used:

Hits@N: Hits@N is the proportion of true aligned entities that appear in the first N entities of the sorted rank list. Hits@N can be

* Corresponding Author.

Table 2. Best hyper-parameter settings of model and the search space for hyper-parameters used.

Hyper-parameters	Best setting	Search space
Batch size	3500	1000, 1500, 3500, 7500, 10000
Train epoch	600	500, 600, 700, 800, 900, 1000, 1500, 2000
Learning rate	5e-3	3e-4, 6e-4, 3e-3, 6e-3, 3e-2
Weight decay	1e-2	1e-3, 5e-3, 1e-2, 5e-2
Decomposition factors R	8	2, 4, 8, 16, 32, 64, 128,256
Graph input hidden dimension	300	200, 300, 400, 500
Graph layer	3	1, 2, 3, 4, 5
Graph feature size	300	100, 200, 300, 400, 500
Visual feature size	100	100, 200, 300, 400, 500
Attribute feature size	100	100, 200, 300, 400, 500
Relation feature size	100	100, 200, 300, 400, 500
temperature factor τ	0.1	0.05, 0.1, 0.15, 0.2, 0.25, 0.3

Table 3. Hardware specifications of the used machine.

hardware	specification
RAM	251 GB
CPU	Intel(R) Xeon(R) Silver 4110 CPU @ 2.10GHz
GPU	NVIDIA(R) A100(80GB) x 4

defined as

$$\text{Hits@N} = \frac{1}{|\mathcal{S}|}\sum_{q\in\mathcal{S}} \mathbb{I}[\text{rank}(i) \leq N], \tag{1}$$

where $\mathcal{S}$ is the number of all testing alignment sets, rank_i refers to the rank position of the first correct mapping for the i-th query entities, and $\mathbb{I}[\text{rank}(i) \leq N]$ yields 1 if i is ranked between 1 and $\mathcal{S}$, 0 otherwise. This metric is bounded in the [0, 1] range and its values increase with $\mathcal{S}$, where the higher the better. Note that, Hits@1 should be preferable, and it is equivalent to precision widely-used in conventional entity alignment.

MRR: Mean reciprocal rank (MRR) measures the number of aligned entity pairs predicted correctly. MRR is the average of the reciprocal ranks of results for a sample of candidate alignment entities:

$$\text{Hits@N} = \frac{1}{|\mathcal{S}|}\sum_{q\in\mathcal{S}} \frac{1}{\text{rank}(i)}, \tag{2}$$

MRR is a useful metric because it not only considers if the EA algorithm correctly aligns entities, but also the rank of the first correctly aligned entity. This means that MRR penalizes lower ranks more severely than higher ones, which is often more reflective of real-world performance. Higher MRR values indicate better performance, with 1 being the maximum achievable value.

1.3 Implementation details

We report our best hyper-parameter settings across two MMKGs datasets and hyper-parameter search space in Table 2. It's noteworthy that all hyperparameter configurations were carefully tuned using a 10-trial grid search technique. Instead of always choosing the best-performing model, we balance the memory limit and model performance. We train and evaluate all our models on a machine with the specifications listed in Table 3.

References

[1] J. Lehmann, R. Isele, M. Jakob, A. Jentzsch, D. Kontokostas, P. N. Mendes, S. Hellmann, M. Morsey, P. van Kleef, S. Auer, and C. Bizer. Dbpedia - A large-scale, multilingual knowledge base extracted from wikipedia. Semantic Web, 6(2):167–195, 2015.

[2] F. Liu, M. Chen, D. Roth, and N. Collier. Visual pivoting for (unsupervised) entity alignment. In Proceedings of AAAI, pages 4257–4266, 2021.

[3] Y. Liu, H. Li, A. García-Durán, M. Niepert, D. Oñoro-Rubio, and D. S. Rosenblum. MMKG: multi-modal knowledge graphs. In Proceedings of ESWC, volume 11503, pages 459–474, 2019.