Table 3: Entity Type Prediction Accuracy on FB15kET

Method	MRR		Hits @		
	Raw	Filt	1	3	10
Rescal	0.07	0.19	9.71	19.58	37.58
Rescal-ET	0.10	0.24	12.17	27.92	50.72
TransE	0.15	0.45	31.51	51.45	**73.93**
TransE-ET	**0.17**	0.46	33.56	52.96	71.16
HolE	0.08	0.22	13.29	23.35	38.16
HolE-ET	0.15	0.42	29.40	48.04	66.73
ETE	**0.17**	**0.50**	**38.51**	**55.33**	71.93

Table 4: Entity Type Prediction Accuracy on YAGO43kET

Method	MRR		Hits @		
	Raw	Filt	1	3	10
Rescal	0.05	0.08	4.24	8.31	15.31
Rescal-ET	0.04	0.09	4.32	9.62	19.40
TransE	0.09	0.21	12.63	23.24	38.93
TransE-ET	**0.10**	0.18	9.19	19.41	35.58
HolE	0.05	0.16	9.02	17.28	29.25
HolE-ET	0.08	0.18	10.28	20.13	34.90
ETE	**0.10**	**0.23**	**13.73**	**26.28**	**42.18**

4.3 Experimental Results

For our experiments, we used the following evaluation protocol: for each triple (ϵ, p, τ) in the test set, first τ is replaced by τ', and we compute the score of (ϵ, p, τ') for $\forall \tau' \in T$, and then we rank all of these *corrupted* triples by the scores. It is possible that multiple corrupted versions of a triple exist in a dataset because an entity could have multiple entity types. In this case, only one corrupted triple is considered as the correct one for each test triple. To avoid this issue, we remove all of the corrupted triples from the ranking, except for the correct one. In Tables 3 and 4, Raw indicates the cases where we do not remove the additional corrupted triples; those where we only include the correct triple are shown as Filt (filtered).

For our experiments, the prediction accuracy of ETE was compared to three state-of-the-art KG embedding methodsbaselines: Rescal, TransE and HolE on the FB15kET and YAGO43kET datasets. We also extended Rescal, TransE and HolE in the same way as ETE that they are trained with the FB15k and YAGO43k training sets first and then used to train embeddings of entity types on the FB15kET and YAGO43kET training sets. The trained vectors of entity types are used for the test with FB15kET and YAGO43kET validation and test sets. We call the extended baselines Rescal-ET, TransE-ET and HolE-ET, respectively.

From Tables 3 and 4, the effectiveness of our approach can clearly be seen. The extended baselines Rescal-ET, TransE-ET and HolE-ET show better accuracy performance than the original methods in all cases, except for TransE-ET on YAGO43kET. Furthermore, these experimental results show that our method ETE consistently outperforms all the baseline methods.

5 CONCLUSION

We proposed an embedding method ETE for entity type prediction. The main benefits of our approach are: (1) higher prediction accuracy compared to state-of-the-art baseline algorithms, and (2) higher accuracy both for inferring missing entity types as well as for inferring missing entities and relation types. We achieve these benefits while preserving linear scalability with the number of entity types. From the results of our experiments, we show that our method consistently gives higher prediction accuracy than baseline methods on two kinds of real-world knowledge graph.

6 ACKNOWLEDGEMENTS

This material is based upon work supported in whole or in part with funding from the Laboratory for Analytic Sciences (LAS). Any opinions, findings, conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the LAS and/or any agency or entity of the United States Government.

REFERENCES

[1] Antoine Bordes, Sumit Chopra, and Jason Weston. 2014. Question answering with subgraph embeddings. In *Proceedings of Conference on Empirical Methods in Natural Language Processing (EMNLP).*

[2] Antoine Bordes, Nicolas Usunier, Alberto Garcia-Duran, Jason Weston, and Oksana Yakhnenko. 2013. Translating embeddings for modeling multi-relational data. In *Advances in Neural Information Processing Systems.* 2787–2795.

[3] Antoine Bordes, Jason Weston, Ronan Collobert, and Yoshua Bengio. 2011. Learning structured embeddings of knowledge bases. In *Conference on Artificial Intelligence.*

[4] Kai-Wei Chang, Scott Wen-tau Yih, Bishan Yang, and Chris Meek. 2014. Typed tensor decomposition of knowledge bases for relation extraction. In *Proceedings of Conference on Empirical Methods in Natural Language Processing (EMNLP).*

[5] John Duchi, Elad Hazan, and Yoram Singer. 2011. Adaptive subgradient methods for online learning and stochastic optimization. *Journal of Machine Learning Research* 12, Jul (2011), 2121–2159.

[6] Yuan Fang and Ming-Wei Chang. 2014. Entity linking on microblogs with spatial and temporal signals. In *Transactions of the Association for Computational Linguistics*, Vol. 2. 259–272.

[7] Yankai Lin, Zhiyuan Liu, Maosong Sun, Yang Liu, and Xuan Zhu. 2015. Learning Entity and Relation Embeddings for Knowledge Graph Completion. In *AAAI Conference on Artificial Intelligence.* 2181–2187.

[8] Laurens van der Maaten and Geoffrey Hinton. 2008. Visualizing data using t-SNE. *Journal of Machine Learning Research* 9, Nov (2008), 2579–2605.

[9] Changsung Moon, Steve Harenberg, John Slankas, and Nagiza Samatova. 2017. Learning Contextual Embeddings for Knowledge Graph Completion. In *Pacific Asia Conference on Information Systems (PACIS).*

[10] Arvind Neelakantan and Ming-Wei Chang. 2015. Inferring missing entity type instances for knowledge base completion: New dataset and methods. In *the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (HLT-NAACL).* 515–525.

[11] Maximilian Nickel, Lorenzo Rosasco, and Tomaso Poggio. 2016. Holographic Embeddings of Knowledge Graphs. In *AAAI Conference on Artificial Intelligence.* 1955–1961.

[12] Maximilian Nickel, Volker Tresp, and Hans-Peter Kriegel. 2011. A three-way model for collective learning on multi-relational data. In *Proceedings of the 28th international conference on machine learning (ICML-11).* 809–816.

[13] Richard Socher, Danqi Chen, Christopher D Manning, and Andrew Ng. 2013. Reasoning with neural tensor networks for knowledge base completion. In *Advances in Neural Information Processing Systems.* 926–934.

[14] Zhen Wang, Jianwen Zhang, Jianlin Feng, and Zheng Chen. 2014. Knowledge Graph Embedding by Translating on Hyperplanes. In *AAAI Conference on Artificial Intelligence.* 1112–1119.

[15] Ruobing Xie, Zhiyuan Liu, and Maosong Sun. 2016. Representation Learning of Knowledge Graphs with Hierarchical Types. In *Proceedings of International Joint Conference on Artificial Intelligence (IJCAI-16).* 2965–2971.

[16] Xuchen Yao and Benjamin Van Durme. 2014. Information Extraction over Structured Data: Question Answering with Freebase. In *the Association for Computational Linguistics.* Citeseer, 956–966.