Translating Embedding with Local Connection for Knowledge Graph Completion

Shijun Liu
Shandong University

Translating Embedding with Local Connection for Knowledge Graph Completion

Extended Abstract

Zeyuan Cui
Shandong University
Jinan, China

Shijun Liu*
Shandong University
Jinan, China

Li Pan
Shandong University
Jinan, China

Qiang He
Swinburne University of Technology
Melbourne, Australia

ABSTRACT

Knowledge graphs are extremely useful resources for intelligent applications but suffer from incompleteness. We notice that previous translative embedding models ignore the local connection of the head entity which is important in predicting the tail entity in the triplet. In this paper, we propose a model named TransL, which incorporates local connection into the translative embedding model. We design a generic approach to combine all the entities in the local connection which uses different weights to distinguish the contribution degree of different relations. We evaluate our model on link prediction and triplet classification. The experimental results show that TransL is competitive to existing models.

KEYWORDS

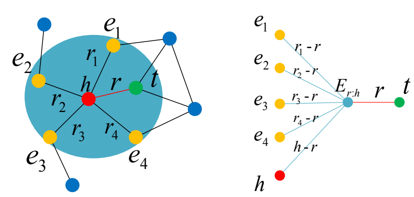
knowledge graph completion, translating embedding, knowledge representation

ACM Reference Format:
Zeyuan Cui, Shijun Liu, Li Pan, and Qiang He. 2020. Translating Embedding with Local Connection for Knowledge Graph Completion. In *Proc. of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020), Auckland, New Zealand, May 9–13, 2020,* IFAAMAS, 3 pages.

1 INTRODUCTION

Many knowledge graphs have been built which are used in intelligent applications. A knowledge graph is a graph composed of entities as nodes and relations as edges. It stores facts in the form of triplets *(head entity, relation, tail entity)*, denoted as *(h, r, t)*. However, existing knowledge graphs are still far from complete. Thus, much work has been devoted to enriching knowledge graphs.

Although existing translation-based models have demonstrated a strong ability to model knowledge graphs, we find that they usually consider each triplet separately. In knowledge graphs, entity nodes are naturally connected to many other entity nodes which can provide important information for predicting the tail entity. For example, if *(Alice, place_of_birth, London)* is true, we can predict that *(Alice, nationality, British)* is likely to be true. This motivates us to explore a model which can make full use of the entities in the local connection of a head entity. As illustrated in Figure 1 (a), the red node is the head entity h and the green node is the tail entity t. The yellow nodes (noted as e_1, e_2, e_3 and e_4) which are connected to the head node should be leveraged to model triplet *(h, r, t)*. We describe our model named TransL in detail below.



(a) (b)

Figure 1: Simple illustration of TransL.

2 METHOD

Given a triplet *(h, r, t)*, TransL first obtains the embedding vector $E_{r:h}$ of the local connections of entity h. The process is shown in Figure 1 (b). The local connections of entity h are $(r_1 : e_1)$, $(r_2 : e_2)$, $(r_3 : e_3)$ and $(r_4 : e_4)$. The embedding of entity e_1 is denoted as V_{e_1} and the weight of entity e_1 is denoted as α_{r_1-r}. The notations are same for other entities. We can obtain the vector $E_{r:h}$ with the following equation: $E_{r:h} = \alpha_{r_1-r}V_{e_1} + \alpha_{r_2-r}V_{e_2} + \alpha_{r_3-r}V_{e_3} + \alpha_{r_4-r}V_{e_4} + \alpha_{h-r}V_h$. In general, we can obtain the embedding of the local connections of the head entity h in relation r with the following equation:

$$E_{r:h} = \sum_{(i:e)\in C_h} \alpha_{i-r} V_e \tag{1}$$

where C_h is the set of the local connections of entity h. For each $(i : e) \in C_h$, i is the relation which connects entity e and head entity h. V_e is the embedding of entity e. And α_{i-r} indicates the importance of relation i to relation r. More specifically, for relation r, each type of relation i corresponds to a weight w_{ir}. We normalize them across all the local connections of h using the softmax function:

$$\alpha_{i-r} = \frac{exp(w_{ir})}{\sum_{(j:e)\in C_h} exp(w_{jr})} \tag{2}$$

*Corresponding author: lsj@sdu.edu.cn, panli@sdu.edu.cn. This work was supported by the National Key Research and Development Program of China (2017YFA0700601).

Proc. of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020), B. An, N. Yorke-Smith, A. El Fallah Seghrouchni, G. Sukthankar (eds.), May 9–13, 2020, Auckland, New Zealand.

Following the success of translative embedding techniques [1, 5, 6, 10], TransL aims to achieve $E_{r:h} + V_r \approx V_t$ when the triplet *(h, r, t)* holds. In other words, V_t should be a near neighbor of $E_{r:h} + V_r$ when *(h, r, t)* is correct, while $E_{r:h} + V_r$ should be far away from V_t otherwise. Given a triplet *(h, r, t)*, we compute a score to indicate how likely that two entities (h, t) are in relation r. The score function is defined as:

$$f(h, r, t) = ||E_{r:h} + V_r - V_t||_2^2 \tag{3}$$

During the training phase, we want the score of a positive triplet to be small while the score of negative triplet to be big. The loss functions for the positive triplet and the negative triplet are given by:

$$\begin{aligned} L_+(h, r, t) &= max(0, f(h, r, t) - \gamma_{pos}) \\ L_-(h, r, t) &= max(0, \gamma_{neg} - f(h, r, t)) \end{aligned} \tag{4}$$

We minimize the loss function below over the training set:

$$L(\theta) = \sum_{(h,r,t) \in T_+} L_+(h, r, t) + \sum_{(h,r,t) \in T_-} L_-(h, r, t) \tag{5}$$

where T_+ denotes the set of positive triplets and T_- denotes the set of negative triplets.

3 EXPERIMENTS

In the experiments, we compare TransL with previous models on the tasks of link prediction and triplet classification.

Link Prediction On the task of link prediction, we compare TransL with three state-of-the-art matrix factorization models, including CP [4], DistMult [11] and ComplEx [9]. We also compare TransL with HolE [7], a holographic embedding model as a baseline method. Another baseline for link prediction is TransE [1] which is a classic model for knowledge graph completion. In addition, we also include R-GCN [8] and ConvE [2] in the comparison. Following previous work, we use FB15k-237 as our evaluation dataset and report both raw and filtered MRR and filtered Hits@1, Hits@3 and Hits@10 for the evaluated models. The results of link prediction are shown in Table 1. Here, “unif” and “bern” are two different strategies of constructing negative triplets in the training phase which are proposed in [10]. As demonstrated, TransL achieves higher results than other models.

Triplet Classification Here, we compare TransL with other translation-based models, including TransE [1], TransH [10], TransR [6] and TransD [5]. And we further compare to KG2E [3] which learns the embeddings of entities and relations in Gaussian distributions. Following previous work, we use WN11 and FB13 as our evaluation datasets. Table 2 shows the accuracy of each model. On WN11, TransL outperforms all the other models by achieving 86.1% for the “unif” setting and 86.6% for the “bern” setting. On FB13, the accuracy achieved by TransL is higher than TransE [1], TransH [10], TransR [6] and KG2E [3]. And it is close to the optimal accuracy achieved by TransD [5].

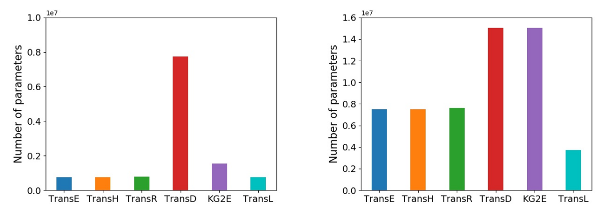
Here, we analyze the actual number of parameters of each model in our experiments of triplet classification. TransL’s parameters to be trained are all the entity embeddings, all the relation embeddings and a relation weight matrix. So, the total number of TransL’s parameters is $O(n_e d_e + n_r d_r + n_r n_r)(d_e = d_r)$, where n_e and n_r represent the numbers of entities and relations respectively. d_e is the dimensionality of the entity embedding space and d_r is the dimensionality of the relation embedding space. The optimal configurations of our model in the experiments are: $d_e = d_r = 20$ on WN11; $d_e = d_r = 50$ on FB13. The actual number of parameters of each model is shown in Figure 2. For WN11, the numbers of parameters of TransE, TransH and TransR are almost the same as TransL. For FB13, the actual number of TransL’s parameters is smaller than TransE, TransH and TransR. We can find that TransD and KG2E require tremendous of parameters to achieve high performance.

Table 1: Evaluation results of link prediction on FB15k-237.

Model	MRR		Hits		
	Raw	Filter	@1	@3	@10
CP [4]	0.080	0.182	0.101	0.197	0.357
DistMult [11]	0.100	0.191	0.106	0.207	0.376
ComplEx [9]	0.109	0.201	0.112	0.213	0.388
HolE [7]	0.124	0.222	0.133	0.253	0.391
TransE [1]	0.144	0.233	0.147	0.263	0.398
R-GCN [8]	0.158	0.248	0.153	0.258	0.414
ConvE [2]	-	0.316	0.239	0.350	0.491
TransL (unif)	0.227	0.342	0.244	0.379	0.535
TransL (bern)	0.248	0.355	0.260	0.389	0.551

Table 2: Evaluation results of triplet classification (%).

Model	WN11	FB13
TransE (unif / bern) [1]	75.9 / 75.9	70.9 / 81.5
TransH (unif / bern) [10]	77.7 / 78.8	76.5 / 83.3
TransR (unif / bern) [6]	85.5 / 85.9	74.7 / 82.5
TransD (unif / bern) [5]	85.6 / 86.4	85.9 / 89.1
KG2E (unif / bern) [3]	83.6 / 85.4	76.4 / 85.3
TransL (unif / bern)	86.1 / 86.6	83.8 / 85.6



(a) WN11 (b) FB13

Figure 2: Number of parameters of each model.

4 CONCLUSION

In this paper, we present a model named TransL for knowledge graph completion. The basic idea of TransL is to explicitly leverage the local connection to learn entity and relation embeddings. The results of experiments conducted on three widely-used benchmark datasets show that TransL is competitive to existing models. Our plans for future work is to further investigate the relationship among entities and to further improve TransL’s performance.

REFERENCES

[1] 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 26*, C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Q. Weinberger (Eds.). Curran Associates, Inc., 2787–2795.

[2] Tim Dettmers, Pasquale Minervini, Pontus Stenetorp, and Sebastian Riedel. 2018. Convolutional 2D Knowledge Graph Embeddings. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, (AAAI-18), the 30th innovative Applications of Artificial Intelligence (IAAI-18), and the 8th AAAI Symposium on Educational Advances in Artificial Intelligence (EAAI-18), New Orleans, Louisiana, USA, February 2-7, 2018*, Sheila A. McIlraith and Kilian Q. Weinberger (Eds.). AAAI Press, 1811–1818.

[3] Shizhu He, Kang Liu, Guoliang Ji, and Jun Zhao. 2015. Learning to Represent Knowledge Graphs with Gaussian Embedding. In *Proceedings of the 24th ACM International Conference on Information and Knowledge Management, CIKM 2015, Melbourne, VIC, Australia, October 19 - 23, 2015*, James Bailey, Alistair Moffat, Charu C. Aggarwal, Maarten de Rijke, Ravi Kumar, Vanessa Murdock, Timos K. Sellis, and Jeffrey Xu Yu (Eds.). ACM, 623–632.

[4] F. L. Hitchcock. 1927. The expression of a tensor or a polyadic as a sum of products. *Studies in Applied Mathematics* 6, 1-4 (1927), 164–189.

[5] Guoliang Ji, Shizhu He, Liheng Xu, Kang Liu, and Jun Zhao. 2015. Knowledge Graph Embedding via Dynamic Mapping Matrix. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing of the Asian Federation of Natural Language Processing, ACL 2015, July 26-31, 2015, Beijing, China, Volume 1: Long Papers*. The Association for Computer Linguistics, 687–696.

[6] Yankai Lin, Zhiyuan Liu, Xuan Zhu, Xuan Zhu, and Xuan Zhu. 2015. Learning entity and relation embeddings for knowledge graph completion. In *Twenty-Ninth AAAI Conference on Artificial Intelligence*. AAAI Press, 2181–2187.

[7] Maximilian Nickel, Lorenzo Rosasco, and Tomaso A. Poggio. 2016. Holographic Embeddings of Knowledge Graphs. In *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, February 12-17, 2016, Phoenix, Arizona, USA*. AAAI Press, 1955–1961.

[8] Michael Sejr Schlichtkrull, Thomas N. Kipf, Peter Bloem, Rianne van den Berg, Ivan Titov, and Max Welling. 2018. Modeling Relational Data with Graph Convolutional Networks. In *The Semantic Web - 15th International Conference, ESWC 2018, Heraklion, Crete, Greece, June 3-7, 2018, Proceedings*. Springer, 593–607.

[9] Théo Trouillon, Johannes Welbl, Sebastian Riedel, Éric Gaussier, and Guillaume Bouchard. 2016. Complex Embeddings for Simple Link Prediction. In *Proceedings of the 33nd International Conference on Machine Learning, ICML 2016, New York City, NY, USA, June 19-24, 2016*. JMLR.org, 2071–2080.

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

[11] Bishan Yang, Wen-tau Yih, Xiaodong He, Jianfeng Gao, and Li Deng. 2015. Embedding Entities and Relations for Learning and Inference in Knowledge Bases. In *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*.