Knowledge Graph Completion Technology Research

Yongqi Shi*
College of Information and Communication, National University of Defense Technology
Wuhan, China
yongqi_shi@qq.com

Yiwei Lu
College of Information and Communication, National University of Defense Technology
Wuhan, China
yiweilu89@nudt.edu.cn

Ruopeng Yang
College of Information and Communication, National University of Defense Technology
Wuhan, China
yrp_roc@126.com

Yuantao Yang
College of Information and Communication, National University of Defense Technology
Wuhan, China
yangyuantaohit@163.com

***Abstract*—In recent years, knowledge graphs have been widely used in the fields of language cognition, search and recommendation, question answering systems, etc. However, with the continuous expansion of the scale of many knowledge graphs constructed by RDF, how to solve the incompleteness of knowledge graphs has become a research hotspot. knowledge graph completion technology is used to complete the missing content in the knowledge graph, mainly including entity completion and relationship completion. This paper mainly focuses on the completion problem, and sorts out completion methods based on deep representation learning. Through the analysis of the research history and latest progress of completion technology, the practical challenges and future development directions of the technology are put forward.**

Keywords-Knowledge Graph; Entity Completion; Artificial Intelligence; Deep Learning

I. INTRODUCTION

Usually, when people describe a thing or event, it is described according to the structure of subject, predicate and object, which is the common (S, P, O) structure. Or relationship, used to describe the role relationship between subject and object, such as (Michael Jackson, isA, Super star), generally in the form of triples. The database that stores these triples is called a knowledge base. The stored content is mainly the real world. This structure can often be described by a graph structure.

Building a knowledge graph requires acquiring knowledge from external data sources, processing it into triples and storing them in the database. This construction method generally goes through four stages. The early knowledge bases needed to be completed manually, and were basically composed of a small range of experts[1-3]; later, with the emergence of crowdsourcing technology[4], the manual work was carried out by the network The cost is reduced but the quality is not as reliable as before, but the scale has grown rapidly; to better ensure the quality of the data in the knowledge base, automated from semi-structured documents The semi-structured automatic extraction method is produced, but it still relies on a large number of manual rules, learning rules and regular expressions, these rules will have a great impact on the quality of the final data; in order to overcome the above shortcomings, Using artificial intelligence methods, triples can be automatically extracted from unstructured documents[5-9], and the specific implementation details can refer to relevant literature. The current knowledge graph construction process (as shown in Figure 1) can be roughly divided into three steps [10]. (1) Semantic information extraction: mainly refers to the completion of semantic recognition from semi-structured or unstructured data through natural language processing and other technologies, and the completion of entity extraction, relationship extraction, coreference disambiguation and other functions; (2) Multivariate data integration and Verification: mainly refers to verifying the acquired triples, and introducing other knowledge bases if necessary to complete entity matching and relationship category identification across knowledge bases, so as to achieve the purpose of knowledge fusion more effectively; (3) Knowledge graph completion: It mainly refers to predicting the relationship between entities through the acquired knowledge to achieve the completion of the relationship between entities, or the completion of entity type information. This process can use the knowledge inside the knowledge base, or introduce Knowledge of third-party knowledge bases to help complete. In fact, the process of completion can use human cognitive process as a reference, learn existing knowledge, and initially realize the task of knowledge learning through the construction, and then obtain knowledge through the analysis of existing knowledge. Ability to learn, thereby digging out new knowledge.

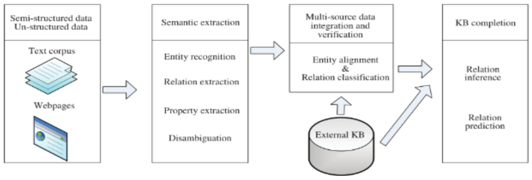
Figure 1. The Construction Process of Knowledge Graph.

II. THE ORIGIN OF THE PROBLEM

In the process of research and application of knowledge graph and its related technologie. Moreover, since the construction of most knowledge graphs is implemented in a manual and semi-automatic way, such as extracting and filling from text or existing knowledge bases, it is unrealistic to cover all facts, and the internal associations between information are complex and diverse. There are a lot of hidden entities and

2022 International Conference on Computing, Robotics and System Sciences (ICRSS) | 978-1-6654-6403-1/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICRSS57469.2022.00024

978-1-6654-6403-1/22/$31.00 ©2022 IEEE
DOI 10.1109/ICRSS57469.2022.00024



relationships waiting to be mined, especially in large knowledge graphs. Moreover, the coverage of the commonsense knowledge base on which it is built is limited, and there are still a lot of facts missing. Therefore, the construction of large-scale knowledge graphs must be implemented in a long-term iterative loop.

Knowledge graph completion is an important research direction. Based on the evolutionary process of completion research, this paper discusses two categories: first, traditional, including rule-based reasoning and probabilistic graph modeling, represented primarily by Markov logic networks. The other type is deep learning-based knowledge graph complementation methods, including translation model-based methods, semantic matching model-based methods, network feature learning-based methods, and other neural network-based methods [11]. Model method. This paper summarizes the typical research work in the various completion methods described above, and compares and analyzes the strengths and weaknesses of the various methods.

III. Traditional Knowledge Graph Completion Method

A. Rule-based Knowledge Graph Reasoning Completion Method

The most classic is the reasoning component based on the NELLs language learning system proposed by Mitchel et al. [12]. The system first learns probabilistic rules, and then instantiates the rules through manual screening, builds a structured knowledge base, and finally uses relational learning algorithms to infer. Create a new instance. In addition, Wei et al. proposed a knowledge base embedding framework KGRL [13], which is based on semantic rules to represent OWL2RL for reasoning, which can effectively derive more implicit information. In addition, KGRL can also eliminate redundant data through reasoning and judgment.

The above single-step inference model implemented by rules is based on a number of rules and features. It is difficult to get rules and constraints that are less computable and have high coverage. Therefore, related research is beginning to focus on multi-hop rules. Development of inference direction.

The probabilistic language reasoning method ProPPR [14] is proposed by Wang et al. ProPPR's reasoning is based on a personalized PageRank process rather than proofs constructed by the SLD parsing theorem prover. The triples are concatenated into clauses, and the clauses are used as inference targets, take the relationship as the inference step, and fuse the relationship weight. This method expands and refines the acquired rules, and provides a more precise basis for the reasoning process.

Therefore, the rule-based reasoning method can combine artificially defined logic rules with various probabilistic graphical models to conduct knowledge reasoning on the basis of the constructed logic network, thereby obtaining new facts.

B. Knowledge Graph Completion Method Based on Probabilistic Graph Model

A probabilistic graphical model is a model that computes probabilistic relationships on graph data. It integrates the expressive and computational capabilities of probability theory and graph theory, and provides directions for predicting new facts based on probability distribution inference.

Methods based on probabilistic graph models are mostly based on Markov logic networks [15] and Bayesian networks [16]. The NELL cleaning system based on Markov logic developed by Jiang et al [17], which enables the knowledge base to use joint probabilistic reasoning, use the logistic regression model to extract feature patterns as the basis for interpreting the reliability of extracted facts, and calculate the establishment of candidate entities and relationships. The probability of predicting the correct triplet completes the knowledge graph. This method defines the latent characteristics of the graph based on the logistic regression model and performs link prediction. The problem is that the high relationship complexity restricts the feature learning and parameter estimation capabilities of the model.

Han et al. using the Bayesian network's ability to model uncertain information [18], making statistics on the nodes of the commodity knowledge graph, and the probability of the establishment of the link between the product and the user entity is predicted.The model integrates the semantic information contained in the knowledge graph through the external data set, and achieves better prediction accuracy and model completion effect, but there is still room for improvement in computational efficiency.

This method can integrate multiple information such as relational semantics, has greater advantages in the prediction of new knowledge, is easy to explain and understand, and has a more flexible graph structure, which is often used to solve uncertain problems. This kind of method has improved computing power, but the algorithm complexity limits its application in large-scale multi-relational graphs.

C. Knowledge Graph Completion Method Based on Graph Computing

Method based on graph computing focuses on abstracting the structure of the knowledge graph into graph data, that is, a collection of nodes and edges, and realizes the completion by predicting the structure of the graph network.

Lao et al. proposed the Path Ranking Algorithm (PRA) [19]. This algorithm is the most typical practice in this type of method. For a certain relationship, a random walk algorithm is used to search for a specific path in the whole graph as a training feature of the model, and then the model is used to predict the links between nodes, and then complete Knowledge Graph.

Wang et al. proposed the CPRA model [20] on the basis of the above method. Unlike PRA, which models a single relationship, the method for measuring the similarity between entity nodes is based on the common paths of nodes, that is, a set of relationships. The common path accounts for the proportion of all paths, and combined with multi-task learning, the features of the common path and a single relationship are saved at one time. This method outperforms the PRA algorithm in knowledge prediction performance on a large-scale knowledge graph through the coupling relationship.

The advantage of the completion method based on graph computing is that it introduces a new perspective[21], taking graph structure as another predictive feature of knowledge reasoning, including information such as the in-and-out degree of entity nodes, adjacency graph structure, and path. However, when applied to large graphs, due to the large number of features such as paths, it faces problems such as high memory usage and feature explosion. At the same time, the model has high computational complexity and weak portability. This is also the common problem of traditional completion methods. Right question. In the face of the semantic nature of knowledge graphs, traditional methods have poor performance in encoding semantic information.

IV. Knowledge Graph Completion Method Based on Deep Representation Learning

The knowledge graph can be showing a network structure. However, graphs will encounter the following problems when applied. First, it is computationally inefficient. Because the complexity of graph algorithms is difficult to control, and it needs to be specially designed for different graphs, the portability is poor, and it is difficult to carry out calculations for knowledge graphs with a wide coverage; secondly, such knowledge graphs usually contain long-tailed distributions[22]. This results in sparse representation of entities and relationships. The main purpose is to use machine learning methods to represent entities, relationships and other semantic information in a continuous vector space in the form of dense low-dimensional real-valued vectors. It simplifies the operation while simplifying the inherent graph structure of the knowledge graph. A typical knowledge representation learning technique generally includes the following parts: determining the continuous embedded representation of relations and entities, judging the true probability of triples by scoring, learning the representation of entity relations, and solving the problem of maximizing the probability of being established for visible facts.

A. Knowledge Graph Completion Method Based on Translation Model

Based on the above research, Bordes et al. proposed the classic translation model TransE [23], which projected the head and tail entities and relations into the embedding space, and then designed a score function to calculate the relationship after translation and transformation of the relation vector. Euclidean distance, according to which to determine the probability that the triplet is true, and obtain new facts to complete the knowledge map. This method is efficient and concise, and the prediction and completion effect is good. The limitation is that each feature weight is not discriminative, and it lacks the ability to calculate complex relationships. After that, many researchers started a series of studies on the TransE variant model inspired by this.

Wang et al. proposed TransH [24] for scene modeling with various complex relationships between entities, which generates embeddings based on hyperplanes of different relationships, and uses the connection relationship between triple element vectors on the hyperplane to determine new Whether the fact is true or not expands the translation model's ability to deal with complex relationships, and at the same time uses algorithms to effectively control the model complexity.

The TransD [25] model proposed by Ji et al. takes into account the different attributes contained in entities, and will obtain different embeddings under the transformation of various relationships. The elements in the knowledge graph are represented in different vector spaces by matrix dynamic mapping, which simplifies the calculation of the projection operation, which reduces the parameters in the learning process, is a further extension of the translation model.

Considering the semantic properties of the encoded knowledge graph, the TransT [26] model proposed by Ma et al. combines triple structured information with prior knowledge of type information to generate embeddings for its representation in different contexts, and Using mixture Gaussian distribution description, the practice of embedding model in semantic similarity calculation is realized. Defects. This method expresses entities and relationships in a more intuitive way, and also has a certain breakthrough in scalability.

B. Knowledge Graph Completion Method Based on Semantic Matching Model

To obtain the probability of new facts being true by matching the representations of elements in the knowledge graph in the embedding space to predict new knowledge, in the inference completion process, KGs are usually represented as tensors, and then tensor decomposition is used to infer unknown facts. The entry is set to zero for unknown and unseen relations. Then, the triplet scores are calculated according to the vectors obtained by factoring, and the candidates with higher scores are selected as the inference results.

Nickel et al. proposed the bilinear model RESCAL [27], which uses knowledge graph as the correlation between interconnected nodes of relational data, learns relational features, using relational matrices to model implicit interactions between knowledge graph elements, and finally using bilinear functions to score triples[28]. It can perform bilinear transformation on relations to describe the association, but the model complexity is high.

Trouillon et al. further proposed the extended model ComplEx [29], for the asymmetric relationship, the element is represented in the complex space. Use a tensor slice to represent the adjacency matrix of a specific relationship, perform low-rank decomposition on it, and then correspond to the elements of the triplet, and finally score the given triplet based on the multilinear product between vectors, and judge the triplet.

Bordes et al. proposed that for multi-relational data, linear relational embeddings model relations in the feature space with additional constraints, thereby learning projections from entities [30], and based on this, the research on the Semantic Matching Energy function [31] was carried out. Modeling based on vector representations, learn distributed forms of multi-relational data and apply to scenarios with a large number of relational types by sharing the states and parameters of entities.

Neural Tensor Networks proposed by Socher et al [32] is a typical application of this idea, modeling triples based on neural networks and descriptive information, and using bilinear tensors in neural networks Do linear transformation to realize the connection of entities embedded in different dimensions. NTN can accurately predict the hidden relationship between entities.

C. Knowledge Graph Completion Method Based on Network Representation Learning

Perozzi et al. proposed the Deepwalk algorithm[33], which constructs a random walk path on the network, learns similar embeddings, and obtains node sequence information in multiple local network search paths at the same time, saving computing time and storage space, and when the graph structure changes,it can be iteratively updated without secondary computation for the whole graph. The disadvantage is that the transition probability between nodes is ignored, which is not in line with practical application scenarios.

Based on the shortcomings of the above models, Tang et al. proposed the LINE model[34]. By calculating the first-order similarity and second-order similarity of network nodes, the local and global information of the network structure is saved, and weights are assigned to different nodes and edges. Then, the embedding representation of each vertex is obtained by minimizing the distance function.

D. Knowledge Graph Completion Methods Based on Other Neural Network Models

Deep neural networks methods are effective to alleviate data sparseness in natural language learning tasks. The method based on neural networks uses expressive capabilities of neural networks to model knowledge graphs, which can achieve good results.

In order to form a new and effective vector structure feature space, it is necessary to construct new spatial basic elements in the knowledge graph. In the data structure of the knowledge graph, there is a lot of textual information, but because the key data information is not well extracted from the knowledge graph, in addition, there are many entities and entity relationships in various text structures. Therefore, in order to filter out meaningless text content, you need to set the text neighborhood between entities.

In the knowledge graph missing connection completion algorithm considering data sparseness, it is necessary to label the basic entities in the corpus, and define each entity as a text neighborhood in the knowledge graph. At the same time, in the text neighborhood of the knowledge graph, data vocabulary and entities appear at the same time. Therefore, the expression for a text neighborhood that fuses entity relationships looks like this:

$$N_h=N_T(h)\cup N_S(h) \quad (1)$$

$$N_t=N_T(t)\cup N_S(t) \quad (2)$$

In the automatic completion algorithm considering the knowledge graph missing connection of data sparseness, the knowledge graph embedding representation model is constructed by using the latent feature vector of the entity, and its triples are explained, and the knowledge graph embedding representation model based on data sparse is constructed to extract unknown relationships of data or reason about the relationships of missing data, and represent them in low-dimensional vector space.

Therefore, the knowledge graph embedding representation model based on data sparseness mainly includes: the vector of entity relationships, matrices, tensors and other embedding representations and the scoring function for judging triples. In addition, through the knowledge graph attention network model, the neighborhood structure between nodes is fully considered, and the knowledge graph embedding representation model is optimized according to the data semantic structure in the knowledge graph.

For large-scale knowledge graphs, a two-dimensional convolutional neural network model is used for the representation learning for knowledge graphs[35]. Connection layers and convolutional layers to model the associations between entities and relations, using 2D convolution to predict new knowledge. At the same time, a shared weight mechanism is set up. The advantages of the model are that the parameter utilization is high and scalable, but the convolutional neural network can only operate on regular data.

For problems in the non-Euclidean graph data field, combining the models, Shang et al. proposed an end-to-end Structure-Aware Convolutional Network [36] for completion. The model uses a graph convolutional network coding with additional weights to aggregate the node structure, and has learnable weights. The decoder combines the ConvE and TransE models. Translation between and relations. SACN improves the performance of the model by directly convolving entities and relationships in the same dimension, and uses graph convolutional networks to effectively create node embedding representations that aggregate messages from their neighborhoods for nodes.

The above neural network model cannot handle the sequential nature of triples connected as clauses. Guo et al. proposed a serialized completion method DSKG [37] based on the recurrent neural network model, and used a knowledge graph-specific sampling method to train the model. This study uses a two-layer model to process entities and relationships using different RNN units, feeding them into the model as different types of sequence elements and then performing cyclic transformations to predict the next hidden state, solving triples into sequences. The problem of insufficient contextual information required for post-inference.

Combining the advantages of multi-class algorithms, Lei et al. proposed a path inference method that integrates attention mechanism and long short-term memory neural network model for completion task [38]. The model learns explicit representations between entities and semantic information, while distinguishing multiple representations of entities under different semantic relations.

V. SUMMARY AND OUTLOOK

From the above description, we can see that the development of knowledge base completion technology is closely linked with people's cognition of knowledge and the progress of knowledge representation methods. With the advent of the era of big data, the representation, storage and Relevant computing will inevitably face huge challenges. For completion technology, the challenges and main development directions in the future should be reflected in the following three aspects:

(1) The sparsity of entities and relationships is more prominent. In statistics, the long-tail phenomenon of data itself is very common, so there are many methods in statistics to solve this problem. In the construction of large-scale knowledge graphs, such a long tail phenomenon will also occur, which is mainly reflected in the repeated occurrence of many entities and relationships with high frequency, such as US presidents or sports stars, news about such entities or Articles are very rich, so there are many related relationship instances; but for other entities and relationships, there are very few instances, such as ordinary people, although the frequency is not high, but the number is large, resulting in related relationship instances. It is also very sparse, and this situation will become more obvious as the amount of data continues to increase. It can be seen from the aforementioned completion problem that there are few instances of relationships and attributes of long-tail entities, which will lead to serious problems of missing relationships and attributes, and will have a serious impact on the completion process. Therefore, in the problem of completion, it will become more and more urgent to solve the sparsity of long-tail entities and their relationships, which may need to be comprehensively solved by the combination of various technologies such as information retrieval, relationship discovery, and entity linking.

(2) The one-to-many, many-to-one and many-to-many problems of entity relationships become more serious. As can be seen from the introduction of the previous methods, the early knowledge graph construction and completion technologies were designed based on the one-to-one relationship. With the continuous increase and enrichment of knowledge, one-to-many, many-to-one and many-to-many There will be more and more many-to-many relationships, which will be very prominent in the construction of domain knowledge graphs. For example, in the field of life sciences, a certain gene will be related to hundreds or even thousands of proteins, a certain reaction pathway will appear repeatedly in thousands of reaction sequences, and certain attributes will appear in most genomes. Compared with the problems in the open field, the one-to-many, many-to-one and many-to-many relationships at this time are not as simple as one-to-ten or dozens of orders of magnitude, but hundreds of thousands of orders of magnitude. Traditional solutions, the solution cannot be effective or even impossible to solve the relational learning problem of this magnitude, which requires adding new control variables and constraints on the basis of the existing solutions, or even proposing completely different solutions or representation models.

(3) Dynamic knowledge graph completion and how to automatically expand the scale of knowledge graphs have become the focus of research. Because in the real world, new entities and new relationships may be continuously generated, and the existing translation models cannot meet the needs of automatically adding new entities and new relationships. Compared with static knowledge map completion, dynamic knowledge map completion can It can effectively establish the relationship, and it can also update the data in the knowledge graph in time, so it has better practical value. How to design an efficient learning algorithm for dynamic completion is a good research point at present.

REFERENCES

[1] Lenat D B, "CYC: a large-scale investment in knowledge infrastructure," Commun ACM, vol. 38, pp. 33–38.

[2] Miller G A, "WordNet: a lexical database for English," Commun ACM, vol. 38, pp. 39–41.

[3] Bodenreider O, "The unified medical language system (UMLS): integrating biomedical terminology," Nucleic AcidsRes, vol. 32, 2004, pp. 267–270

[4] Bollacker K D, Evans C, Paritosh P, "Freebase: a collaboratively created graph database for structuring humanknowledge," In: Proceedings of International Conference on Management of Data, 2008, pp. 1247–1250

[5] Fan J, Ferrucci D A, Gondek D, "PRISMATIC: inducing knowledge from a large scale lexicalized relationresource," In: Proceedings of North American Chapter of the Association for Computational Linguistics, 2010, pp. 122–127

[6] Fader A, Soderland S, Etzioni O, "Identifying relations for open information extraction," In: Proceedings of EmpiricalMethods in NaturalLanguage Processing, 2011, pp. 1535–1545

[7] Nakashole N, Weikum G, Suchanek F M, "PATTY: a taxonomy of relational patterns with semantic types," In: Proceedings of Empirical Methods in NaturalLanguage Processing, 2012, pp. 1135–1145

[8] Niu F, Zhang C, ReC, "Elementary: large-scale knowledge-base construction via machine learning and statistical inference," Int J Semantic Web Inf Syst, vol. 8, 2012, pp. 42–73

[9] Nickel M, Murphy K, Tresp V, "A review of relational machine learning for knowledge graphs," Proc IEEE, vol. 104, 2016, pp. 11–33

[10] Schmitz M, Soderland S, Bart R, "Open language learning for information extraction," In: Proceedings of Empirical Methods in NaturalLanguage Processing, 2012, pp. 523–534

[11] Guan S P, Jin X L, Jia Y T, "Research progress of knowledge reasoning for knowledge graph," Journal of Software, vol. 29, 10. 2018, pp. 2966-2994.

[12] Carlson A, Betteridge J, Kisiel B, "Toward an Architecture for Never-ending Language Learning,," Proceedings of the AAAI Conference on Artificial Intelligence, vol. 24, 1. 2010, pp. 1306–1313.

[13] Wei Y, Luo J, Xie H, "KGRL: An OWL2 RL Reasoning System for Large ScaleKnowledge Graph," 2016 12th InternationalConference on Semantics, Knowledge and Grids (SKG) IEEE, 2016, pp. 83-89.

[14] Wang W Y, Mazaitis K, Cohen W W, "Programming with Personalized Pagerank: A Locally Groundable First-order Probabilistic logic," Proceedings of the 22nd ACM international conference on Information & Knowledge Management, 2013, pp. 2129-2138.

[15] Richardson M, Domingos P, "Markov logic networks," Machine Learning, vol. 63, 2. 2006, pp. 207-207.

[16] Heckerman D, Geiger D, Chickering D M, "Learning bayesian networks: The combination of knowledge and statistical data," Machine Learning, vol. 20, 3. 1995, pp. 197-243.

[17] Jiang S, Lowd D, Dou D, "Learning to Refine an Automatically Extracted Knowledge Base Using Markov Logic," 2012 IEEE 12th InternationalConference on Data Mining IEEE, 2012, pp. 912-917.

[18] Lu H, Zhidu Y, Yujie W, "Link prediction of knowledge graph based on Bayesian network," Journal of Frontiers of Computer Science and Technology, vol. 11, 5. 2017, pp. 742-751.

[19] Lao N, Mitchell T, Cohen W, "Random Walk Inference and Learning in a Large Scale Knowledge Base," Proceedings of the 2011 conference on empirical methods in natural language processing, 2011, pp. 529-539.

[20] Wang Q, Liu J, Luo Y, "Knowledge Base Completion via Coupled Path Ranking," Meeting of theAssociation for Computational Linguistics, 2016, pp. 1308-1318.

[21] Wang Q, Mao Z, Wang B, "Knowledge graph embedding: A survey of approaches and applications," IEEE Transactions on Knowledge and Data Engineering, vol. 29, 12. 2017, pp. 2724-2743.

[22] Mikolov T, Chen K, Corrado G, "Efficient estimation of word representations in vector space," Computer Science, 2013.

[23] Bordes A, Usunier N, Garcia-Duran A, "Translating Embeddings for Modeling Multi-relational Data," Neural Information Processing Systems (NIPS), 2013, pp. 1-9.

[24] Wang Z, Zhang J, Feng J, "Knowledge Graph Embedding by Translating on Hyperplanes," Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence, 2014, pp. 1112-1119.

[25] Ji G, He S, Xu L, "Knowledge Graph Embedding via Dynamic Mapping Matrix," Proceedings of the 53rd annual meeting of the association for computational linguistics and the 7th international joint conference on natural language processing, 2015, pp. 687-696.

[26] Ma S, Ding J, Jia W, "Transt: Type-based Multiple Embedding Representations for Knowledge Graph Completion," Joint European Conference on Machine Learning and Knowledge Discovery in Databases Springer, Cham, 2017, pp. 717-733.

[27] Nickel M, Tresp V, Kriegel H P, "A Three-way Model for Collective Learning on Multi-relational Data," Proceedings of the 28th International Conference on International Conference on Machine Learning, 2011, pp. 809-816.

[28] Yang B, Yih W, He X, "Embedding entities and relations for learning and inference in knowledge bases," arXiv preprint arXiv:1412.6575, 2014.

[29] Trouillon T, Welbl J, Riedel S, "Complex Embeddings for Simple Link Prediction," International Conference on Machine Learning. PMLR, 2016, pp. 2071-2080.

[30] Bordes A, Weston J, Collobert R, "Learning Structured Embeddings of Knowledge Bases," 25th Conference on Artificial Intelligence (AAAI). 2011, pp. 301-306.

[31] Bordes A, Glorot X, Weston J, "A semantic matching energy function for learning with multi-relational data," Machine Learning, vol. 94, 2. 2014, pp. 233-259.

[32] Socher R, Chen D, Manning C D, "Reasoning with Neural Tensor Networks for Knowledge Base Completion," Advances in neural information processing systems. 2013, pp. 926-934.

[33] Perozzi B, Al-Rfou R, Skiena S, "Deepwalk: Online Learning of Social Representations," Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining. 2014, pp. 701-710.

[34] Tang J, Qu M, Wang M, "Line: Large-scale Information Network Embedding," Proceedings of the 24th international conference on world wide web, 2015, pp. 1067-1077.

[35] Dettmers T, Minervini P, Stenetorp P, "Convolutional2D Knowledge Graph Embeddings," 32nd AAAI Conference on Artificial Intelligence, AAAI 2018, AAI Publications, vol. 32, 2018, 1811-1818.

[36] Shang C, Tang Y, Huang J, "End-to-end Structure-aware Convolutional Networks for Knowledge Base Completion," Proceedings of the AAAI Conference on Artificial Intelligence, vol. 33, 1. 2019, pp. 3060-3067.

[37] Guo L, Zhang Q, Ge W, "DSKG: A Deep Sequential Model for Knowledge Graph Completion," China Conference on Knowledge Graph and Semantic Computing Springer, Singapore, 2018, pp. 65-77.

[38] Lei K, Zhang J, Xie Y, "Path-based reasoning with constrained type attention for knowledge graph completion," Neural Computing and Applications, 2019, pp. 1-10.