Multi-source knowledge fusion: a survey

Xiaojuan Zhao [1] · Yan Jia [1] · Aiping Li [1] · Rong Jiang [1] · Yichen Song [1]

Abstract

Multi-source knowledge fusion is one of the important research topics in the fields of artificial intelligence, natural language processing, and so on. The research results of multi-source knowledge fusion can help computer to better understand human intelligence, human language and human thinking, effectively promote the Big Search in Cyberspace, effectively promote the construction of domain knowledge graphs (KGs), and bring enormous social and economic benefits. Due to the uncertainty of knowledge acquisition, the reliability and confidence of KG based on entity recognition and relationship extraction technology need to be evaluated. On the one hand, the process of multi-source knowledge reasoning can detect conflicts and provide help for knowledge evaluation and verification; on the other hand, the new knowledge acquired by knowledge reasoning is also uncertain and needs to be evaluated and verified. Collaborative reasoning of multi-source knowledge includes not only inferring new knowledge from multi-source knowledge, but also conflict detection, i.e. identifying erroneous knowledge or conflicts between knowledges. Starting from several related concepts of multi-source knowledge fusion, this paper comprehensively introduces the latest research progress of open-source knowledge fusion, multi-knowledge graphs fusion, information fusion within KGs, multi-modal knowledge fusion and multi-source knowledge collaborative reasoning. On this basis, the challenges and future research directions of multi-source knowledge fusion in a large-scale knowledge base environment are discussed.

✉ Aiping Li
liaiping@nudt.edu.cn

Extended author information available on the last page of the article

1 Introduction

Knowledge reasoning, intelligent search, intelligent questions and answers(Q&A) and natural language understanding (NLP) need the support of large-scale knowledge base. When building knowledge graph (KG), due to the wide range of knowledge sources, there will be duplication, diversity of semantics and uneven quality among multi-source heterogeneous knowledge. We need to carry out conflict detection, entity disambiguation, entity alignment and other operations, effectively fusion the multi-source knowledge to form a large-scale, high-quality knowledge graph.

From a theoretical point of view, multi-source knowledge fusion is one of the important research topics in the fields of artificial intelligence and natural language processing. The research results of multi-source knowledge fusion can help computers better understand human intelligence, human language and human thinking. From the application point of view, multi-source knowledge fusion can provide effective knowledge support for intelligent search, intelligent recommendation, intelligence analysis, etc. It has great social value and economic benefits.

At present, industry and academia both at home and abroad have carried out extensive research on the key technologies of multi-source knowledge fusion. However, the existing work is aimed at the related technologies in knowledge fusion, such as entity alignment, entity disambiguation, knowledge representation, etc. Some work is to learn from the methods and technologies of multi-source data fusion, and a unified theoretical system has not yet been formed. This paper will introduce the latest research progress of multi-source knowledge fusion technology. Firstly, we introduced several concepts related to multi-source knowledge fusion, such as data fusion, representation learning, entity alignment, and so on. Then, based on the relationship between multi-source knowledge fusion and KG, the research progress of knowledge fusion is introduced in two directions. Then, the related research progress of multi-source knowledge collaborative reasoning is introduced. Finally, the challenges and future research directions of multi-source knowledge fusion in a large-scale knowledge base environment are prospected.

2 Concepts related to multi-source knowledge fusion

2.1 Knowledge fusion and data fusion

Data Fusion, also called multi-sensor data fusion, was first applied in the military field. The goals of data fusion are the most original and untreated records. It mainly discerns the authenticity of the data and the dependability of the information source, solves the numerical conflict between different data sources and seeks out the implied real value, and this processing lays particular emphasis on the data level. Because data has data quality problems (such as data input errors, data loss, etc.), and the data quality affects the algorithm's effect, solving data conflicts and finding the true value of data are considered to be the two basic tasks of data fusion. Multi-source data fusion can obtain more accurate, complete, and reliable estimates and judgments than a single data source.

Knowledge fusion is different from data fusion. The basic problem of knowledge fusion is to study how to combine description information about the same entity or concept from multiple sources. Therefore, knowledge fusion has different names, such as ontology

alignment, ontology matching, record linkage, entity resolution, entity alignment, etc., but their essential work is the same.

Early knowledge fusion was mainly based on traditional data fusion methods. From the traditional data fusion methods, according to the data characteristics of knowledge fusion, some data fusion methods were selected and improved, and they were applied to knowledge fusion [1], or after the knowledge fusion problem is transformed into a data fusion problem, and then the data fusion technology is applied to solve it [2, 3].

2.2 Multi-source knowledge fusion and representation learning

Knowledge representation learning is mainly oriented to entities and relationships in KGs. The entities and vectors are represented in low-dimensional dense vector space by using the method of modeling, and then calculated and reasoned. Knowledge representation learning is of great significance to how computer understands and calculates knowledge. Before the idea of embedding appeared in 2013, people basically used one-hot representation to represent entities. In recent years, the core idea of knowledge representation is how to find an appropriate method to embedding KG into vector space, so as to calculate in vector space. The success of representation learning technology in image, voice, video and NLP has attracted the attention of researchers in the field of KG. Some researchers have begun to study knowledge fusion-oriented representation learning technology.

Vector-based KG representation makes it easier to integrate these data with deep learning models, and the representation of knowledge graph based on vector space has attracted more and more attention. On the one hand, by designing a reasonable knowledge graph representation learning model, the knowledges from different sources can be projected into a unified representation space, which can realize the organic integration of multi-knowledge graphs and adapt to the large-scale application of KGs, It is also of great significance to the research of the integration and completion tasks involved in the construction of KG; on the other hand, the integration of knowledges from different sources can help knowledge graphs capture hidden knowledge more easily, and effectively promote the performance of knowledge representation, which is an iterative process of mutual strengthen.

With the development of KG research and machine learning, the study of network representation learning has attracted extensive attention. Because information networks may contain billions of nodes and edges, it may be very difficult to perform complex reasoning processes throughout the network. Therefore, it has been proposed that one way to solve this problem is network embedding. Network embedding aims at learning the low-dimensional potential representation of nodes in the network. The learned feature representation can be used as the characteristics of various graph-based tasks, such as classification, clustering, link prediction and visualization. The central idea is to find a mapping function that converts every node in the network into a potential representation of low dimensions. Related concepts include graph embedding, graph representation learning and so on. Multi-source knowledge fusion can give full attention to the research results in these areas. In Section 3 representation learning technologies related to multi-source knowledge fusion are proposed.

2.3 Multi-source knowledge fusion and entity alignment

The common abridged general view of knowledge graphs mainly contains three kinds of nodes: entities, concepts and attributes. As the core unit of knowledge graph, entity is also an

important language unit carrying information in text. Different entities have different relationships. Knowledge graphs can be constructed freely by any organization or individual. The data behind them are from a wide range of sources and of uneven quality, resulting in diversity and heterogeneity among them. Knowledge fusion is to integrate different knowledge graphs into a unified form. The commonly used technical methods include ontology alignment (also known as ontology matching) and entity alignment (also known as entity matching). Entity alignment is also called instance alignment, object co-referential resolution. According to the different objects aligned, alignment is generally divided into ontology alignment and instance alignment. Ontology alignment focuses on discovering classes, attributes or relationships that are equivalent or similar (pattern level), while instance alignment focuses on discovering different instances referring to the same object in the real world. According to the alignment method, it can be divided into paired entity alignment and collective entity alignment. Paired entity alignment is also called element-based entity alignment, collective entity alignment is also called structure-based entity alignment, which is divided into global collective entity alignment and local collective entity alignment.

The fundamental problem to be settled in multi-source knowledge fusion is how to reconcile the descriptive information from multiple sources about the same entity or conception. According to the content of the fusion, knowledge fusion can be divided into data schema level fusion and data level fusion. The key mission of multi-source knowledge fusion is data level fusion, consists of entity alignment (EA), attribute alignment, and conflict detection and resolution. Data schema level fusion includes three main aspects: conceptive merging, conceptual hyponymy merging and merging of attribute definitions of concepts. Some researches regard entity alignment and knowledge fusion as two independent segments. They believe that knowledge fusion is based on alignment. After resolving conflicts through conflict detection and truth discovery, knowledge is correlated and merged to form a consistent result. The knowledge fusion mentioned in our research includes the whole process of EA, conflict detection and conflict resolution.

Entity link is to detect new entities in text by the entity recognition technology, to link entities mentioned with corresponding entities in knowledge graph, and to add them to the existing knowledge base, which also belongs to the large scope of knowledge fusion.

3 Multi-source knowledge fusion related technologies

In an open environment, on the one hand, knowledge graphs need to constantly integrate new knowledge from the open Internet, enhance the expansion and coverage of the existing knowledge graph, on the other hand, in order to enhance the application effect of knowledge graphs, we need to integrate multiple knowledge graphs or more semantic information in knowledge graphs.

As shown in Fig. 1, from the perspective of KG construction, multi-source knowledge fusion can be divided into two categories: one is to update the existing KGs, also known as open source knowledge fusion (Section 3.1); this kind of fusion is mainly aimed at large data of the Internet, and studied how to extract useful knowledge from massive fragmented data and integrate it into existing KG, the other is multi-knowledge graph fusion(Section 3.2). It mainly refers to merging multiple knowledge graphs into a large knowledge graph by identifying the equivalent instances, equivalence classes and equivalence attributes of multiple knowledge graphs. Therefore, it is generally considered that the main task of knowledge fusion is entity
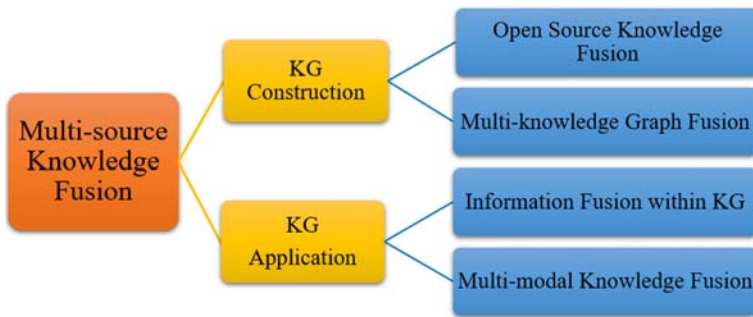
Fig. 1 Classification of Research Progress in Multi-source Knowledge Fusion

alignment. The target of these two kinds of research is to update or construct a new KG. From the perspective of KG application, multi-source knowledge fusion can also be divided into two categories, one is information fusion within knowledge graph (Section 3.3), which mainly refers to taking into account information outsides knowledge graph's structure information in the application process to enhance the application effect. The other is the fusion of multi-modal knowledge (Section 3.4). KG has become very important in the application of intelligent search and recommendation, intelligent Q&A and dialogue system and visual decision support. These two kinds of research are mainly to improve the application quality by better mining the information of multiple knowledge graphs.

3.1 Open source knowledge fusion

Massive text, audio and video data on the Internet are important knowledge sources for building KGs. Open source knowledge fusion mainly refers to the real-time fusion of newly added knowledges, which integrates all kinds of information related to KGs contained in the Internet texts.

Integrating various data sources and various forms of knowledge, extracting new entities and new relationships from the knowledge and adding them to the original knowledge graph. This kind of integration can complement and expand the original knowledge graph, so we can regard open source knowledge fusion as a segment in the process of knowledge graph construction, and can also be understood as knowledge graph updating.

Due to the multi-source heterogeneity of the Internet knowledge, knowledge evaluation and verification are indispensable links for open source knowledge fusion.

Due to the uneven quality of Internet knowledge, knowledge must be evaluated and validated in the process of open source knowledge fusion. Knowledge evaluation can judge the authenticity of knowledge, and integrate the validated knowledge with the existing knowledge in knowledge graphs to achieve the fusion of knowledges and improve the reliability and confidence of knowledge. So far, the research work on open source knowledge fusion mainly focuses on the following two aspects: one is knowledge evaluation and verification, the other is entity link.

There are three traditional methods for knowledge evaluation and verification: Bayesian model [4, 5], the D-S evidence theory [6–8], and the fuzzy set theory [9, 10]. With the development of machine learning, knowledge evaluation and verification methods based on graph models [11–14] have been developed in recent years.

Table 1 Comparison of results on entity alignment

Model	DBP-YAGO [22]			DFB [17]			DBP15K (ZH-EN) [20]			DBP-YG(dense) [31]		
	Hits @1	Hits @ 10	MR	Hits @1	Hits @ 10	MR	Hits @1	Hits @ 10	MR	Hits @1	Hits @ 10	MR
TransE [18]	1.22	3.54	24,809									
MTransE [78]	33.46	34.32	7105	13.6	35.1	547.7	30.83	61.41	154	22.8	51.3	0.32
JAPE [20]	33.35	33.37	5296				38.34	68.86	103	26.8	57.3	0.37
IPTransE [17]				71.7	86.5	49				23.6	51.3	0.33
TransE+PS [17]				61.9	79.2	105.2						
JAPE+SE + AE [20]							41.18	74.46	64			
N-gram [22]	89.69	95.83	23									
RSN4EA [31]										82.6	95.8	0.87
KDCoE [79]										56.8	80.4	0.64

Notes on Table 1:

The blanks in the table indicate that there are no comparative experimental results in the related studies.

Database: DBP-YAGO [22] represents data sets DBpedia and YAG; DFB [17] is formed by randomly dividing FB15K triples into two subsets T1 and T2 of similar size); DBP15k(ZH-EN) [20] is built from the multilingual version of DBpedia. DBP-YG (dense) [31] is composed of monolingual datasets DBpedia(English) and YAGO3(English).

Model: TransE+PS representation integrates TransE with Parameter Sharing Model [17]; JAPE represents Joint Attribute-Preserving Embedding [20]; JAPE+SE + AE denotes the combination of JAPE with structure embedding and attribute embedding [20]; N-gram represent *N-gram-based compositional function* to encode the attribute value proposed in [22].

The basic principle of the Bayesian model is: according to the prior probability of the knowledge to be evaluated in advance, and then use the conditional probability observed in the data source to obtain the posterior probability, and select the correct knowledge according to the maximum posterior probability criterion. In fact, the prior probability of knowledge is often very difficult to precognition, so the Bayesian model has boundedness. The D-S evidence theory is a generalization of Bayesian method. This method does not need to know the prior probability, and can well express "uncertainty", and uses "interval estimation" instead of "point estimation" to describe uncertain information. It can be used to solve the conflict problem in multi-source knowledge fusion. Both the D-S evidence theory and the Bayesian model are based on the hypothesis that knowledge from different sources is independent of each other, and when there is a serious conflict among knowledge sources, it often results in contrary conclusions. In addition, the time complexity of the D-S evidence theory has potential exponential explosion, which is not suitable for large-scale knowledge evaluation and verification. The model based on the fuzzy set theory can deal with both inaccurate and uncertain information, but it needs to set up fuzzy rules and membership functions of knowledge based on experience. It is difficult to guarantee the stability and robustness of knowledge evaluation results, and it is not suitable for multi-source heterogeneous knowledge evaluation. Knowledge evaluation based on graph models uses knowledge from the existing knowledge base to fit the prior model, so as to assign a probability to knowledge, and can also be used as a link prediction problem. According to the prediction results, it can guide the quality evaluation of knowledge acquired from data sources. These methods can reduce the wrong

knowledge to a certain extent and improve the reliability and confidence of knowledge. However, the scale of knowledge in the open domain is becoming larger and larger, and it has strong dynamic evolution characteristics. The following research work should consider the time dimension of knowledge and the large-scale knowledge evaluation.

From the point of view of entity links, the research results of open source knowledge fusion are discussed in the next three parts separately, which are not introduced in detail here.

3.2 Multi-knowledge graph fusion

People use different information sources to construct different knowledge graphs. How to fuse and express multi-knowledge graphs is of great significance to establishing a unified large-scale knowledge graph. Because the information sources of different knowledge graphs are different, they may be domain knowledge graphs or general knowledge graphs, and their knowledge description systems are different. The same entities in semantics will have different expressions in different knowledge graphs, and entities with the same name may also represent different things. Multi-knowledge graph fusion is not simply to merge knowledge graphs, but to discover equivalent instances, equivalent attributes or equivalent classes among knowledge graphs, and to determine which entities and relationships from different knowledge graphs will be aligned.

Entity alignment is an important component of multi-source knowledge fusion technology. The aligned entities can be used to transfer knowledge in multi-knowledge graphs, and facilitate the construction of cross-language knowledge graphs and knowledge reasoning. Considering the multi-type relationship in knowledge graphs, [15] proposed a knowledge graph embedding and entity alignment algorithm based on representation learning. They select the alignment-task driven representative relations based on the pre-aligned entity pairs. With the help of the selected relationships, they embed cross-network entities into public space by modeling the head/tail of entities and the corresponding context vectors. For entity alignment tasks, pre-aligned entities are used to facilitate context information transmission across knowledge graphs. In this way, the problem of entity embedding and alignment can be solved simultaneously in a unified framework. A large number of experiments on two multi-lingual knowledge graphs prove the validity of the model. [16] also proposed a multi-source and multi-knowledge base entity alignment algorithm based on network semantic labels. The core of the algorithm is to align the entities between different knowledge graphs by calculating the semantic similarity pairs between two entities. In the alignment process, the description information of entities including unstructured text keywords, semantic tags and category tags is integrated. Firstly, the similarity of three features is calculated separately, and then the similarity is calculated synthetically.

$$SIM(E_1, E_2) = \omega_1 \times SIM(TP_1, TP_2) + \omega_2 \times SIM(C_1, C_2) + \omega_3 \times SIM(S_1, S_2) \quad (1)$$

SIME1,E2 = ω1 × SIMTP1,TP2 + ω2 × SIMC1,C2 + ω3 × SIMS1,S2 Among them, $SIM(TP_1, TP_2)$, $SIM(C_1, C_2)$, $SIM(S_1, S_2)$ respectively represent the semantic similarity based on attribute tags, the semantic similarity based on class tag matching, and the semantic similarity of unstructured text keywords. When the calculated value is greater than a certain threshold, the entity pair with the greatest similarity is taken as the output of the alignment result, which is also considered to have the same semantic orientation. Sun et al. [17] proposed a new method of joint knowledge embedding to achieve entity alignment. The model consists of three parts: knowledge embedding, joint embedding and iterative alignment. Use TransE [18] and PtransE (Path-based TransE) [19] to learn the entities and relationships in different

knowledge graphs separately to obtain knowledge embedding. Because TransE ignores the important multi-step path information in the knowledge graph, the modeling effect on the complex relationship is not ideal, so PTransE is proposed. The joint embedding mapping all individual knowledge embedding into a semantic space. There are three models embedded in the joint: a translation-based model, a linear transformation model, and a parameter sharing model. Iterative alignment is the discovery of more aligned entities by adding "new aligned entities" to the seed set, updating the joint embedding. The objective function consists of three parts:

$$L = K + J + I \tag{2}$$

Where K, J and I denote the score function of knowledge embeddings, joint embeddings, and iterative alignment. Similarly, JAPE [20] uses attribute and text description information to enhance the learning representation of instances, and uses joint representation learning technology to directly embed entities and relationships in different knowledge graphs into a unified vector space.

Zhong et al. [21] proposed CoLink, a general unsupervised framework for the UIL(User Identity Linkage) problem. CoLink employs a co-training algorithm, which manipulates two independent models, the attribute-based model and the relationship-based model, and makes them reinforce each other iteratively in an unsupervised way. The attribute-based model predicts the linked user pairs by only considering the user attributes. It can utilize any classification algorithm. The sequence-to-sequence learning is a very effective implementation of the attribute-based model, which can well handle the challenge of the attribute alignment by treating it as a machine translation problem. The network consists of two parts: the sequence encoder and the sequence decoder. Both the encoder and the decoder use a deep Long Short-Term Memory (LSTM) architecture. Traditional classification algorithms like Support Vector Machines (SVM) can also be employed in the attribute-based model.

Trsedya et al. [22] proposed an entity alignment method between knowledge graphs based on attribute embeddings. The framework consists of three components including predicate alignment, embedding learning, and entity alignment. The framework is shown in Fig. 2. In the predicate alignment module, two KGs are merged into one KG by renaming potentially aligned predicates. By calculating the similarity of the name of the predicate (the last part of the URI), the potential aligned pairs of predicates are found and renamed using a unified naming format. For example, its predicate pair, "*dbp: bornIn*" and "*yago: wasBornIn*" will be renamed to "*: bornIn*". An embedding learning module includes structure embedding, attribute character embedding and joint embedding learning. The structural embedding model is built on top of TransE. Unlike TransE, the model wants to pay more attention to aligned triples, that is, triples containing aligned predicates. The model achieves the goal by adding weights. The objective function of structural embedding is:

$$\mathcal{L}_{SE} = \sum_{t_r \epsilon T_r} \sum_{t_{r'} \epsilon T_r'} \max\left(0, \gamma + \alpha\left(f(t_r) - f\left(t_r^{'}\right)\right)\right) \tag{3}$$

$$T_r = \{< h, r, t > \,|\, < h, r, t > \, \epsilon G\} \tag{4}$$
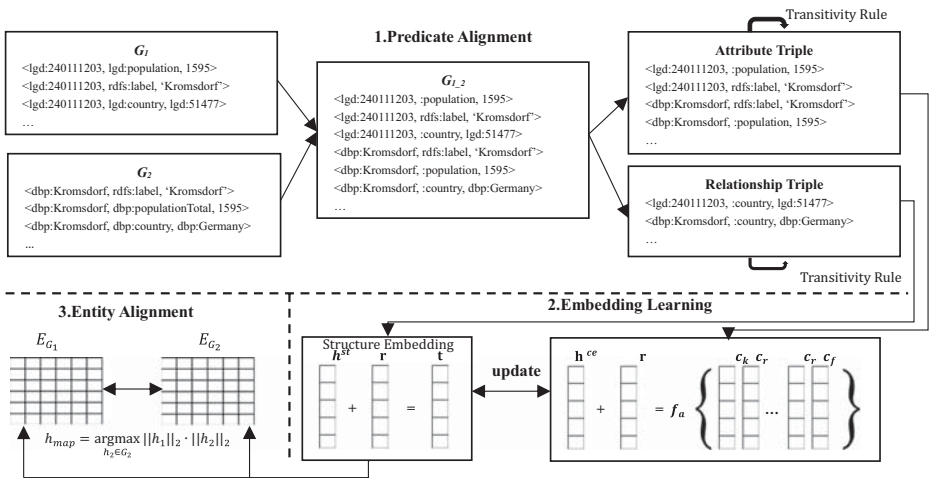
Fig. 2 The Framework of Trsedya et al's Papers

$$T'_r = \left\{ < h', r, t > |h' \in E \right\} \cup \left\{ < h, r, t' > |t' \in E \right\} \tag{5}$$

$$f(t_r) = ||h + r - t|| \tag{6}$$

$$\alpha = \frac{count(r)}{|T|} \tag{7}$$

where count (r) is the number of occurrences relationship r, and $|T|$ is the total number of triples in the merge KG G_{1-2}. Attribute character embedding also follows the idea of TransE. Unlike structure embedding, there are differences in the representation of attributes with the same meaning in different KGs. Hence, Trsedya et al. [22] used a compositional function to encode the attribute value, and the three compositional functions are as follows: the Sum compositional function, the LSTM-based compositional function and the N-gram-based compositional function. The objective function of attribute character embedding is:

$$\mathcal{L}_{CE} = \sum_{t_a \in T_a} \sum_{t_a' \in T'_a} \max \left(0, \gamma_e + \alpha \left(f(t_a) - f\left(t_a'\right) \right) \right) \tag{8}$$

Joint learning uses attribute character embedding to help structure embedding in the same vector space to complete training. The objective function of joint learning is:

$$\mathcal{L}_{SIM} = \sum_{h \in G_1 \cup G_2} \left[1 - ||\mathbf{h}_{se}||_2 \cdot ||\mathbf{h}_{ce}||_2 \right]. \tag{9}$$

The overall objective function of the model is:

$$\mathcal{L} = \mathcal{L}_{SE} + \mathcal{L}_{CE} + \mathcal{L}_{SIM} \tag{10}$$

After the joint learning of structure embedding and attribute character embedding, similar entities from different KGs will have similar embeddings, so potential entity pairs $<h_1, h_{map}>$ can be obtained through computing the following equation:

$$h_{map} = argmax_{h_2 \epsilon G_2} \|h_1\|_2 \cdot \|h_2\|_2 \tag{11}$$

EnAli [23] is an unsupervised method for matching entities in two or more heterogeneous data sources. The research on multi-source heterogeneous data is very important in many fields. For large data sources, aligning all triples of multiple data sources is costly. EnAli employs a generative probabilistic model to incorporate the heterogeneous entity attributes via employing exponential family, handle missing values, and also utilize the locality sensitive hashing schema to reduce the candidate tuples and speed up the aligning process. EnAli is highly accurate and efficient even without any ground-truth tuples. EnAli consists of four components as follows: Candidate tuple generation (employs LSH to block entities from N data sources), Similarity computation, Parameter learning, Decision making. EnAli considers both discrete and continuous similarities as a wider range of probability distributions from the exponential family to model the similarity values of matched and unmatched entity tuples. This is an important extension to handle the heterogenous attribute types, including string, numeric, set, distribution, etc., and these exist in the entity alignment task. Wang et al. [24] proposed a method of enriching entities in ontology by using external definition and context information, and the additional information is used for ontology alignment. Different domains usually have different sentiment expressions, and a general sentiment classifier is not suitable for all domains. Training a domain-specific sentiment classifier for each target domain also faces the problem that the labeled data in the target domain is usually insufficient, and it is costly and time-consuming to annotate enough samples. Multi-source sentiment knowledge fusion can effectively improve the performance of sentiment classification and reduce the dependence on tagged data. Wu et al. [25] constructed a unified fusion framework to train domain-specific sentiment classifier for target domain by fusing sentiment knowledge from multiple sources.

Other studies include: Wang et al. [26] proposed that text data be taken into account in representation learning. Word2vec [27, 28] was used to learn the word representation in Wikipedia text, and TransE [18] was used to learn the knowledge representation in the knowledge base. At the same time, using the link information in the Wikipedia text (the correspondence between anchor text and entity) to make the word representation of entity in text as close as possible to the entity representation in knowledge bases, so as to realize the representation learning of text and knowledge base fusion; Zhong et al. [29] also used similar ideas to fuse entity description information. Sun et al. [30] summarized the current status of entity alignment algorithms in the field of geographical knowledge base research from three aspects of similarity measurement, similarity combination and consistency judgment, summarized the evaluation process of alignment results, and proposed the basic definition and general framework of entity alignment in a geographical knowledge graph. Guo et.al [31] proposed recurrent skipping networks for entity alignment (RSN4EA), which leverages biased RW (Radom Walk) sampling for generating long paths across knowledge graphs and generates the paths with a novel RSN (recurrent skipping network). RSN combines the traditional RNN with residual learning, and only a few parameters can greatly improve the convergence speed and performance.

3.3 Information fusion within knowledge graph

Most of the existing knowledge graph application models only use the triple structure information of knowledge graph, and the information about the entity and the relationship, category information and other information related to the knowledge are not effectively

utilized. There are two main types of research on the internal information fusion of knowledge graphs. One is to consider the entity type, the entity description information and the relationship between the entities in the related research of entity alignment, and the second is to learn the representation of the knowledge graph. Incorporate rich internal information in the knowledge graph to obtain better knowledge representation results.

Zhong et al. [29] performed entity alignment based on entity description information without relying on Wikipedia as anchor text. Inspired by the joint embedding framework in [26], learn the best embedding by minimizing the following loss function:

$$\mathcal{L}\big(\{e_i\}, \{r_j\}, \{w_l\}\big) = \mathcal{L}_K + \mathcal{L}_T + \mathcal{L}_A \tag{12}$$

where $\mathcal{L}_K$, $\mathcal{L}_T$ and $\mathcal{L}_A$ are the loss functions of the knowledge model, the text model and the alignment model respectively. [29] focusing only on the loss function $\mathcal{L}_A$ of the new alignment model, the loss function $\mathcal{L}_K$ of the knowledge model and the loss function in text model $\mathcal{L}_T$ are the same as the counterparts in [26].

Guan et al. [32] proposed a self-learning and embedded entity alignment method (SEEA), which was used to iteratively search for semantic matching entity pairs and make full use of the semantic information contained in entity attributes. See Fig. 3 for an illustration. The knowledge graph is formalized as $G = (E, A, V, R, AT, RT)$, where $E = E_1 \cup E_2$ is the entity set, and E_1 and E_2 are two sets of entities to be aligned. A, V and R represent the set of attributes, the set of attribute values and the set of relationships, respectively. $AT \subseteq E \times A \times V$ is a set of attribute triples, and $RT \subseteq E_1 \times R \times E_2$ is a set of relation triples between entity group E_1 and E_2. The input to the SEEA model is a knowledge graph, which includes two sub-modules: knowledge graph embedding and entity alignment. Knowledge graph embedding includes relation triple learning and attribute triple learning. The self-learning mechanism performs feedback operations from entity alignment to KG embedding. SEEA uses the results of the previous learning iteratively to update the embedding of entities, attributes and attribute values in the next iteration. That is to say, in the self-learning mechanism, the learned relational triples are used to update all embedding in the next iteration.

Yang et al. [33] proposed a Text- Associated Deep Walk (TADW) that incorporates text information. In the framework of matrix decomposition, TADW introduces text features as a supplement to network structure information into network representation learning. Similarly, CANE [34](Context-Aware Network Embedding) is a context-aware embedding method. There are two kinds of embedding for a node V, one is structure-based embedding v_s, the
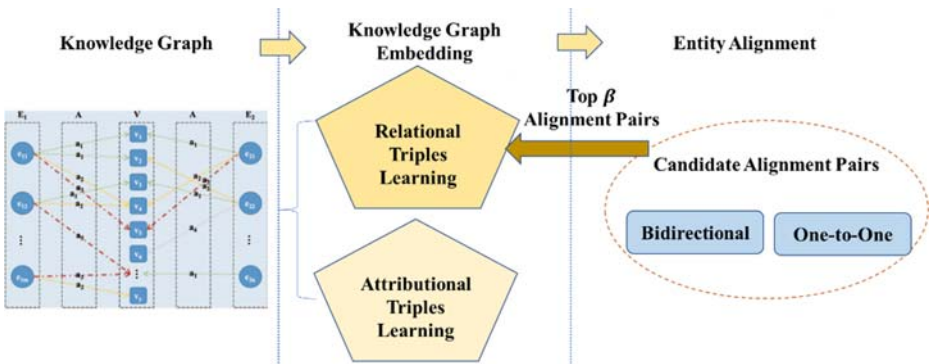

Fig. 3 The framework of the proposed SEEA method

other is text-based embedding v_t (may be context-aware embedding or context-aware embedding), and then they are concatenate to get $v = v_s \oplus v_t$.CANE wants to maximize the objective function of the edge as follows:

$$L = \sum_{e \epsilon E}(L_S(e) + L_t(e)) \tag{13}$$

Where $L_S(e)$ is a structure-based objective function and $L_t(e)$ is a text-based objective function. Context-free Embeddings means that the embedding of a node is fixed and does not change according to its context. Context-aware Embeddings means that CANE learns different embedding based on different context of a node.

Zhang et al. [35] proposed a recommendation system based on Collaborative and knowledge Base Embedding (CKE), as shown in Fig. 4. They introduced structured knowledge, text knowledge, image knowledge and other knowledge graph information to improve the quality of the recommendation system. Among them, structured knowledge uses TransR [36] to get the vector representation of entities. Text knowledge and image knowledge use Stacked Denoising Auto-encoders (SDAE) [37] and Stacked Convolutional Auto-encoders (SCAE) respectively to get vector representation with strong generalization ability.

Kristiadi et al. [38] considered the semantic information carried by the literal meanings of entity names in knowledge graphs, and proposed a new representation learning mechanism LiteralE (See Fig. 5). The improvement strategy of this mechanism is to integrate the literal information I_j or I_i of entities through transformation function $g(\cdot)$ before scoring the vector representation of entities.

Where $g(\cdot)$can be linear transformations

$$g_{lin}(e_i, I_i) = W^T[e_i, I_i] \tag{14}$$
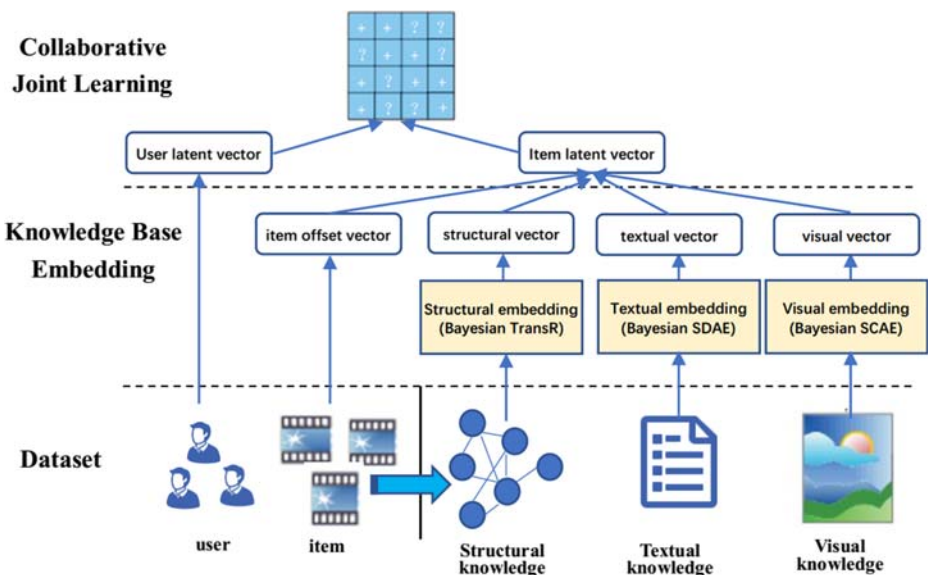
non-linear transformations

Fig. 4 The flowchart of the proposed Collaborative Knowledge Base Embedding (CKE) framework for recommender systems

$$g_{nonlin}(e_i, I_i) = h\big(W^T[e_i, I_i]\big) \tag{15}$$

simple MLPs

$$g_{MLP}(e_i, I_i) = h\big(W_2^T h\big(W_1^T[e_i, I_i]\big)\big) \tag{16}$$

Xie et al. [39] considered that the entity description information provided in Freebase and other knowledge bases can help knowledge representation learning to achieve better results. The representation learning model DKRL(description-embodied knowledge representation learning) proposed in this paper first converts entity description text information into entity representation using CBOW [27, 28] or CNN [40, 41], and then uses the entity representation to learn the objective function of TransE. CBOW extracts keyword sets containing the main concepts of entities from descriptive texts, then selects the first n keywords as input, and then simply adds the coded word vectors as text representations.

$$e_d = x_1 + x_2 + \ldots + x_k \tag{17}$$

Where x_i denotes the embeddings of the first word in the keyword set belonging to entity e. The Convolutional Neural Network (CNN) Encoder consists of five layers. The input is the whole description of a specific entity, and the output is the description-based representation of that entity. CBOW is slightly different from CNN in this model. The former does not consider the word order information of the text, while the latter considers the word order of the text.

TransC [42] is a knowledge graph embedding model which distinguishes concepts from instances. It encodes each concept in knowledge graphs as a sphere and each instance as a vector in the same semantic space. It expresses relations by the spatial inclusion relations between points and spheres and the inclusion relations between spheres. This representation can naturally solve the problem of the transmission of the relations. Concepts and instances, as well as the relative positions between concepts and concepts are described by the relationship between *InstanceOf* and *subClassOf*, respectively. The *InstanceOf* relation is used to indicate whether an instance is in
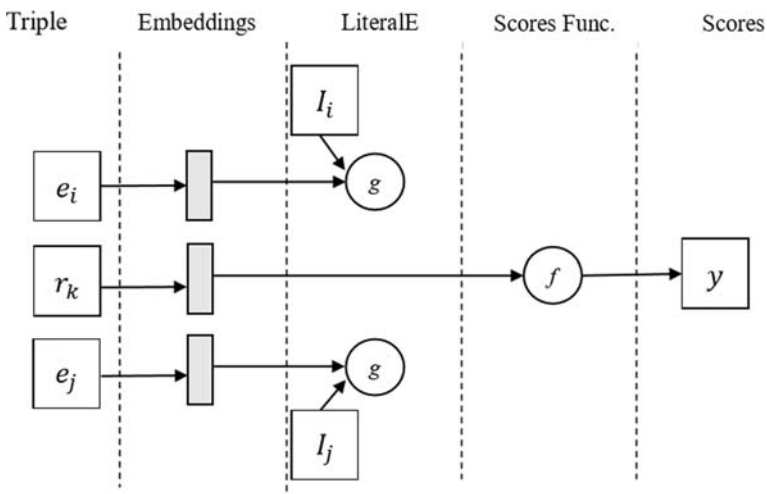
Fig. 5 Overview on how LiteralE is applied to the base scoring function *f*. LiteralE takes the embedding and the corresponding literals as input, and combines them via a learnable function *g*. The output is a joint embedding which is further used in the score function *f*

a sphere represented by a concept, and the *subClassOf* relation is used to indicate the relative position between two concepts. Four possible relative positions are proposed:

As shown in Figure 6, where m is the radius of the sphere, d is the distance between the centers of the two spheres, s_i and s_j represent the spheres represented by concepts i and j, respectively. Figuer6(*a*), 6(*b*), 6(*c*), 6(*d*) respectively represent four kinds of position relations between s_i and s_j. For *InstanceOf* and *subClassOf*, there is a clever design to retain the transitivity of the *isA* relation, that is, the transferability of *instanceOf-subClassOf* is embodied by

$$(i, r_e, c_1) \in S_e \wedge (c_1, r_c, c_2) \in S_c \to (i, r_e, c_2) \in S_e \tag{18}$$

while *subClassOf-subClassOf* is embodied by

$$(c_1, r_c, c_2) \in S_c \wedge (c_2, r_c, c_3) \in S_c \to (c_1, r_c, c_3) \in S_c \tag{19}$$

where (i, r_e, c) means *InstanceOf* triple, (c_i, r_c, c_j) means *SubClassOf* triple. There are three main types of triples: *InstanceOf* Triple, *SubClassOf* Triple, Relational Triple.

the loss function of *instanceOf* triples is defined as:

$$f_e(i, c) = \|i - p\|_2 - m \tag{20}$$

use ζ and ζ' to denote a positive triple and a negative triple, and the margin-based ranking loss for *instanceOf* triples is:

$$\mathcal{L}_e = \sum_{\zeta \in S_e} \sum_{\zeta' \in S_e'} \left[\gamma_e + f_e(\zeta) - f_e\left(\zeta'\right) \right]_+ \tag{21}$$

where $[x]_+ \triangleq \max(0, x)$ and γ_e is the margin separating positive triples and negative triples.

Similarly, we will have the ranking loss for *subClassOf* triples $\mathcal{L}_c$ and relational triples $\mathcal{L}_l$. The overall loss function is the linear combinations of these three functions:

$$\mathcal{L} = \mathcal{L}_e + \mathcal{L}_c + \mathcal{L}_l \tag{22}$$

Other related studies include adding logical rules [31, 43–45], entity types and descriptive text information to knowledge representation learning [46–51], and considering the relationship path in knowledge graph [52–54] Table 2.

3.4 Multi-modal knowledge fusion

Data in different industries come from a wide range of sources and in a variety of forms, each of which can be considered as a modal, such as text, images, video, and audio, different modal have different levels of knowledge representation. Multi-source knowledge focuses on expressing the diversity of data sources. Multi-modal knowledge fusion can make agents
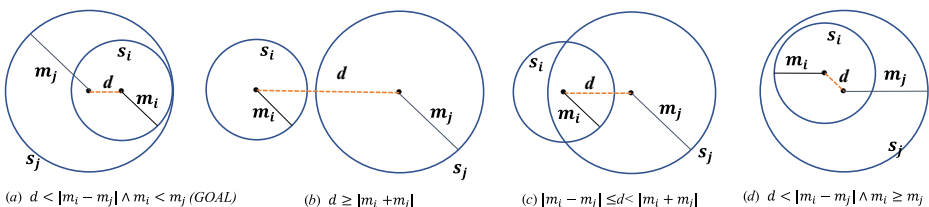

(a) $d < |m_i - m_j| \wedge m_i < m_j$ (GOAL)   (b) $d \geq |m_i + m_j|$ (c) $|m_i - m_j| \leq d < |m_i + m_j|$ (d) $d < |m_i - m_j| \wedge m_i \geq m_j$

Fig. 6 Four relative positions between sphere s_i and s_j

perceive and understand real application scenarios more deeply, and better support industrial applications. Studying the feature representation and learning methods of different modal information can realize the cooperative representation of multi-modal data. In order to overcome the influence of structural differences on multi-modal representation, it is necessary to study the embedded learning method of multi-modal information and its internal and external knowledge, and establish a deep feature learning and association representation model supported by cognitive data, so as to project different modal information, such as language and vision, into a common subspace and realize the multi-modal data co-representation at the knowledge level, and support knowledge acquisition based on multi-modal fusion [55].

Zhang et al. [56] proposed seamless integration of multiple data sources with Bi-GRU (Gated Recurrent Unit) architecture Fig. 7. The model treats four inputs as a sequence $\{s_1, s_2, s_3, s_4\}$ while using a Bi-GRU layer to learn their interdependencies. Subsequently, all hidden units $\{h_1, h_2, h_3, h_4\}$ are concatenated into a new vector representation to preserve their differences and then sent to the final fully connected layer..

The vector representation of a user is:

$$v_u = W[h_1 \oplus h_2 \oplus h_3 \oplus h_4] + b_c \tag{23}$$

Table 2 Comparison of various research models

Model	Fused Information Types	Verification task
Zhong et al. (20,150 [29]	text descriptions of entities	Link prediction, Triplet classification, Relational fact extraction, and Analogical reasoning
SEEA [32]	Semantical Information of Attributes of Entities	Entity aligment
TADW [33]	Text features of vertices	Multi-class classification of vertices
CANE [34]	Structure –based information, Text-based context information	Link prediction, Vertex classification
CKE [35]	Structural knowledge, Textual knowledge and visual knowledge, The information of users and items	Movie and book recommendation
DKRL [39]	semantic of entity descriptions	KG completion entity classification (in Zero-shot Scenario)
TransC [43]	Differentiating concepts and instance in entities	Link prediction, Triple classification
KALE [43]	Jointly embedding KGs and logical rules	Link prediction, Triple classification
Rocktaschel et al. (2015) [45]	Logical Background Knowledge	Relation Extraction
Newman-Griffis et al. (2018) [46]	Entities and surfaces forms, Text information	Analogy completion, Entity sense disambiguation
SSE [47]	Additional Semantic information (Semantically Smooth Embedding)	Link prediction, Triple classification
TKRL [48]	Hierarchical entity type information	KG completion, Triple classification
Jointly (A-LSTM) [49]	Both structural and textual information of entities	Link prediction, Triple classification
TEKE [50]	Textual context information (Text-enchanced knowledge embedding)	Link prediction, Triple classification (Capability to handle 1-to-N, N-to 1, N-to-N relations, and KG spareseness)

$$h_i = f_{BiGRU}(s_i) \tag{24}$$

Bi-RNN is used to get the document presentation. The forward hidden layer can get a hidden representation and the backward hidden layer get a representation too. The two hidden layer representations are fused together and then a self-attention mechanism is used to automatically assign weights to different inputs. User nickname, self-introduction, education information, work information and individualized labels are treated as user metadata. After concatenating all the elements of metadata, feeding them into a Bi-RNN layer and an Attention layer to the metadata representation. Network Representation employ LINE [57].

RBMs (Restricted Boltzmann Machines) [58] can be effectively used to model the distribution of binary-valued data. Boltzmann machine models and their extensions to exponential family distributions [59] have been successfully applied in many applications. The Multimodal Deep Boltzmann machine [60] can be used to learn the characteristics of text and pictures separately [61, 62], and then combine these two.

features into a new feature vector as the input feature of the SVM(Support Vector Machines) classifier. The model integrates cross-modal features to set up a fusion representation.

The DCPR (Deep Context-aware Point of view Recommendation) [63] model is a point of view (POI) recommendation model based on deep context-aware. The DCPR model uses LSTM to learn potential user representations and CNN to generate potential representations from comments. An end-to-end depth model is used to consider POI attributes, user preferences, sequential momentum check-ins and so on.When researching the impact of events and investor sentiment on stock price trend, Zhang et al. [64] extracted events from online news, extracted users' emotions from social media, and fused multi-source heterogeneous data by constructing tensors.

Visual appearance score, appearance mixture type and deformation are three important information sources for human pose estimation. [65] proposed to build a multi-source deep model in order to extract non-linear representation from these different.

aspects of information sources. With the deep model, the global, high-order human body articulation patterns in these information sources are extracted for pose estimation. A direct method is to mix information sources with different statistical characteristics in the first hidden layer. As shown in Figure 8 (a), this method has its limitations. Another method, as shown in Figure 8 (b), is to construct the high-level feature representation of each data source with two layers, and then use the other two layers to fuse the high-level representation of different information sources for pose estimation. Auto-encoder and RBM [58] are two common components of unsupervised deep learning algorithms. Similar approaches have been used in the research of representation learning based on a depth model [66–70].

4 Multi-source knowledge cooperative reasoning

The results of multi-source knowledge fusion can be regarded as an important part of multi-source knowledge cooperative reasoning, whether from the perspective of updating and constructing KG or from the perspective of application of KG. Therefore, it is necessary to sort out the research overview of multi-source knowledge cooperative reasoning.

Traditional reasoning refers to the acquisition of new knowledge or conclusions through various methods. Multi-source knowledge collaborative reasoning includes not only inferring
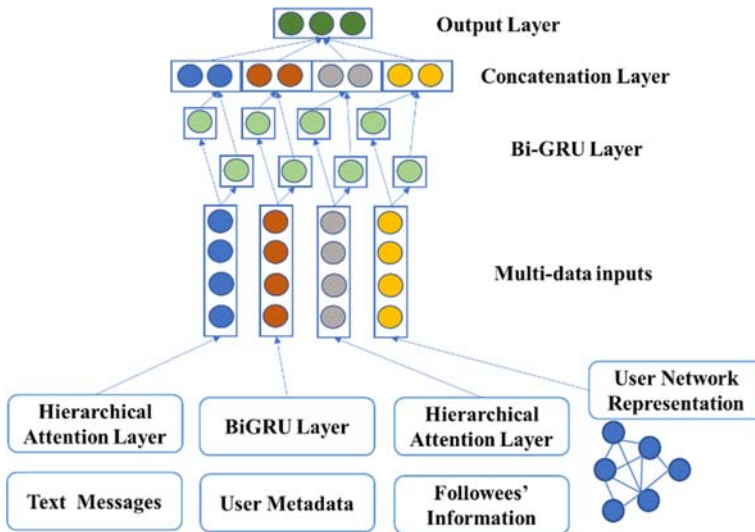
Fig. 7 Illustration of the fusion model. Hierarchical attention layer denotes hierarchical attention network. BiRNN denotes bi-recurrent neural network. Concatenation layer indicates concatenation of all hidden units learned from multi-data inputs

new knowledge from multi-source knowledge, but also conflict detection, that is, identifying wrong knowledge or conflicting knowledge. In multi-source knowledge fusion reasoning, we can regard multi-source knowledge as multi-agent, and multi-agent reasoning related methods are used to study multi-source knowledge fusion reasoning.

Cognitive psychology holds that recursive reasoning, which is to infer what other people think they are thinking, is an inherent thinking mode of human beings and plays an important role in human behavior decision-making in social life. Inspired by human recursive reasoning thinking, Wen et al. [71] introduced the thinking mode of recursive reasoning into the deep reinforcement learning of multi-agent stem for the first time, allowing agents to predict the impact of other agents' reactions on themselves before making decisions. This work improves the depth of AI group thinking, and also provides a new way of thinking for MARL research. Specifically, a recursive probability reasoning framework, Probabilistic Recursive Reasoning (PR2), is proposed, which allows each agent to consider how other agents will respond to their next actions and then make the optimal decision, as shown in Figure 9. Based on the PR2
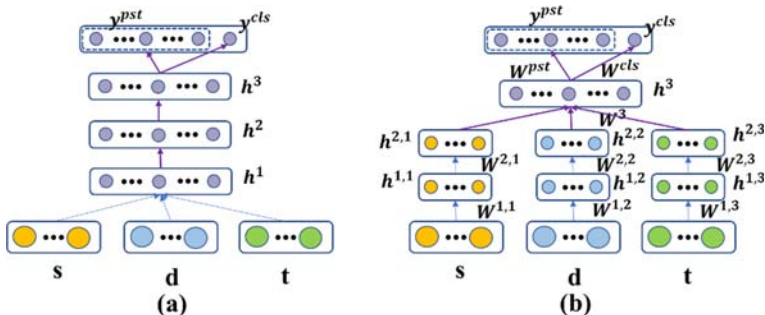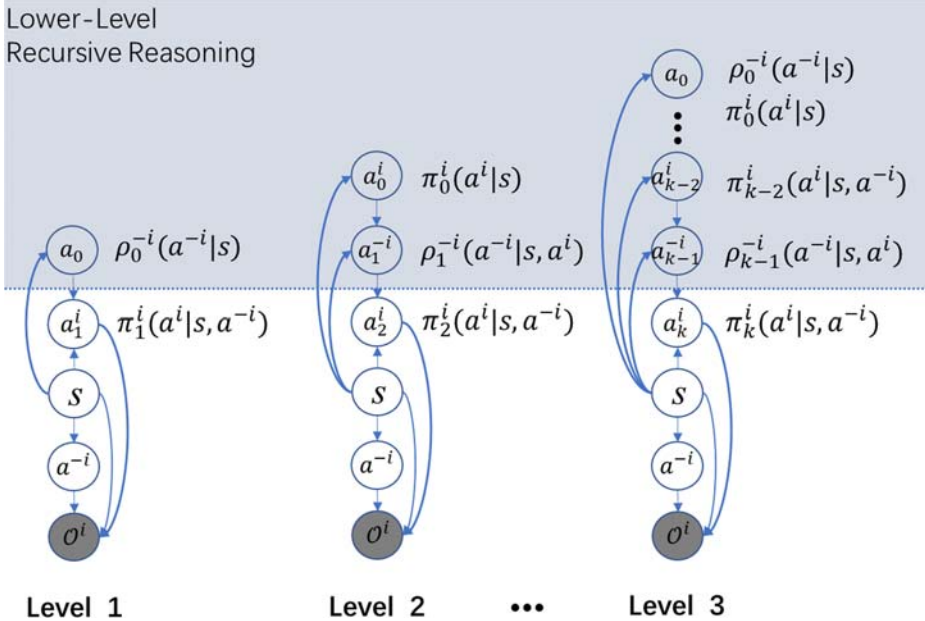

Fig. 8 Direct use of deep model (a) and the deep architecture we propose (b) for part score s, deformation d and mixture type t. Best viewed in color

Fig. 9 Graphical model of the Level-k recursive reasoning. Note that the subfix a^* here stands for the level of thinking not the timestep. The unobservable opponent policies are approximated by ρ^i. The omitted Level-0 model considers opponents fully randomized. Agent i rolls out the recursive reasoning about opponents in its mind (grey area). In the recursion, agents with higher-level beliefs take the best response to the lower-level thinkers' actions. Higher level models would conduct all the computations that the lower-level models have done, e.g. Level-2 contains Level-1

framework, the PR2-Q and the PR2-Actor-Critic algorithms corresponding to continuous and discrete action spaces are proposed. Interestingly, these algorithms are inherently distributed and do not require Centralized Value Function. Many experiments show that PR2 can effectively improve the learning efficiency of a single agent in Multi-Agent Reinforcement learning.

The problem of KG reasoning can be summarized as two steps: path finding and path reasoning. Most of the current methods focus on one step, lacking the interaction between the two steps, which hinders the understanding of diverse inputs and makes the model very sensitive to the impact of noise. In order to increase the robustness of the model and deal with the complex environment, it is necessary to improve the interaction of two steps.

DIVA [72] modeled the link-missing reasoning problem in Q&A tasks based on KGs as a potential variable graph model. The path was regarded as a potential variable, and relationship as a variable that can be observed after a given entity pair. Therefore, the Path-Finding Module is used as a prior distribution to infer potential links and paths. As a likelihood distribution, the inference module divides potential links into several categories. Based on the above assumptions, an approximate posterior module is introduced and a variational auto-encoder (VAE) [73] is designed. The model (DIVA) consists of three parts: a posterior approximator, a prior (path finder), and a likelihood (path reasoner). The framework of variational reasoning is introduced, which combines path finder and path reasoner closely to conduct joint reasoning. In the path reasoning module, the convolution neural network and the feed-forward neural network are used. The input is path sequence, and the output is probability distribution of

relationships. The problem of path finding is modeled as the Markov decision process. Recursive predictive actions are based on history. The hidden state is calculated using an LSTM neural network.

DeepPath [74] and MINERVA [75] (Meandering In Networks of Entities to Reach Veri-similar Answers) can be considered as the optimization of path search procedures. Compound reasoning [52] and reasoning chains [76] can be understood as the optimization of path inferring. For a more complex question, because of the incompleteness of knowledge graph, it is necessary to combine multiple knowledge graphs for inferring in order to seek out the proper answer. DeepPath modeled the process of searching answers to complex questions as MDP (Markov decision process) < S, A, P, R > and resolved it by reinforcement learning. The environment system in the Reinforcement learning system is responsible for the dynamic interaction between knowledge graphs and Agent. However, DeepPath needs to know the target entity in advance and use the target entity to guide the process of finding the reasoning path. MINERVA is to find the correct answer in all the entities in the knowledge graph, it required neither advance knowledge of the target entity nor any pre-training, nor a special designed reward function, but only used LSTM to express the historical state.

5 Prospects for future research

Language knowledge bases are becoming an important source of knowledge for human and artificial intelligence-related applications. Researching cross-lingual knowledge graph fusion technology will undoubtedly provide a general method to help extract and apply this knowledge. The aforementioned knowledge fusion technologies have achieved good results in the application of monolingual knowledge graphs, but the application of these technologies to cross-language knowledge fusion remains to be explored. In addition, academia and industry are beginning to focus on large-scale knowledge graphs. The scale of the knowledge graph is getting larger and larger. The original knowledge graph fusion technology needs to reconsider the accuracy and execution efficiency of the algorithm.

5.1 Cross-lingual knowledge graph fusion

Cross-lingual knowledge graph fusion promotes the tasks of knowledge-driven cross-lingual NLP and promotes the tasks of cross-lingual reasoning. With the development of representation learning, scholars begin to use the information of relational and the text description of entities in multilingual knowledge graphs for cross-lingual representation learning. [77] jointly trained the embedding model of cross-lingual knowledge graphs and the embedding model of cross-lingual description. MTransE [78] solved the problem of representation learning and matching of cross-lingual knowledge graphs through a transfer-based approach. It first uses TransE to learn the representation of a single knowledge graph, and then learns the linear transformation of different knowledge representation spaces for instance matching. MTransE includes three different transfer methods: axis calibration method, transfer vector method and linear transformation method. By using different loss functions, MTransE has designed five different varieties.

The accuracy of cross-lingual reasoning is often unsatisfactory due to the low degree of entity alignment among the multilingual knowledge graph. Cross-lingual knowledge graph alignment based on embedding strategy can effectively improve the accuracy of reasoning if the text description of entities in knowledge graph is taken into account. Chen et al. [79]

proposed a semi-supervised learning method, KDCoE, to deal with cross-lingual knowledge graph alignment. Based on the embedding strategy, this paper collaboratively trained the mutilingual knowledge graph embedding (KGEM) model and the mutilingual entity description embedding model (DEM). Multilingual knowledge graph embedding is composed of knowledge model and alignment model. The traditional TransE method is used to construct the knowledge model, which can preserve the entities and relationships in the embedding space, while the alignment model refers to the linear transformation strategy in MTransE. However, this paper only calculates the embedding of cross-lingual entities rather than the whole embedding of triples. The embedding process of multilingual entity description includes two parts: encoding and cross-lingual embedding. This paper uses Attentive Gated Recurrent Unit encoder (AGRU) to encoding the multi-linguistic entity description. The cross-lingual embedding part uses word embedding method to measure and find similar words between different languages. In order to better reflect the lexical level semantic information described by multilingual entities, cross-lingual Bilbowa [80] word embeddings are pre-trained using cross-lingual parallel corpus Europarl V7 and monolingual corpus in Wikipedia. Then, the entity description text is converted into vector sequence using the embeddings mentioned above, and then input into the encoder.

Xu et al. [81] defined the task of entity alignment in cross-lingual knowledge graph as the task of finding new alignment data based on the existing set of aligned entities. Given two knowledge graphs G_1 and G_2, a set of pre-aligned entities $S = \{(e_{i1}, e_{i2})\}_{i=1}^m$, GCN (Graph Convolutional Networks) is used to embed entities from different languages into a unified vector space, and the aligned entities are expected to be closer. The input of GCN is the eigenvector of the node and the structure of the graph, and the output is the entity embedding at the node level. GCN encodes the neighborhood information of a node into a real vector. In the problem of entity alignment, it is assumed that: (1) equivalent entities tend to have similar relationships, and (2) equivalent entities tend to have equivalent neighbors. GCN can combine attribute information and structure information. Entity alignment is based on the distance between entities. For $e_i \in G_1$ and $e_j \in G_2$ the distance between them is calculated as follows:

$$D\left(e_i, e_j\right) = \beta \frac{f\left(\left(h_s(e_i), h_s\left(e_j\right)\right)\right)}{d_s} + (1-\beta) \frac{f\left(\left(h_a(e_i), h_a\left(e_j\right)\right)\right)}{d_a} \tag{25}$$

Where $(x, y) = \|x - y\|_1$, and β is the parameters that balance the importance of two kinds of embedding.

Wu et al. [82] used the bilingual topic model to solve the cross-lingual ontology matching problem, proposed the modeling of disordered word pairs (called BiBTM) in bilingual documents. On this basis, the word co-occurrence relationship and hierarchical structure relationship between classes are further added, that is, the subsequent C-BiBTM [83]. In order to solve the problem of cross-lingual attribute matching, Zhang et al. [84] proposed the EAFG model, which not only considers the characteristics of attributes to itself, but also considers the correlation between attributes.

5.2 Large-scale knowledge graph fusion

In the big data environment, the existence of multiple links makes the KG more and more complex and larger. The construction of a KG and multi-source knowledge fusion need to reconsider the accuracy and efficiency of the algorithm.

Parallel processing technology mainly takes algorithm as the core, parallel languages as description, software and hardware as implementation tools, which provides some new directions for solving large-scale knowledge graph fusion. Parallel technology mainly includes two aspects: one is multi-core and multi-processor technology in single-machine environment, such as multi-threading and the GPU technology; the other is the distributed technology based on network communications in a multi-machine environment, such as MapReduce computing framework, the Peer-To-Peer network framework, etc. For the languages with low expressive abilities such as RDFS and OWL, parallel processing under a single-machine environment can effectively improve real-time processing efficiency. With the maturity of the distributed technology, more and more researchers begin trying to use a distributed framework in data reasoning. Many works have proposed reasoning methods for large-scale ontology based on MapReduce's open source implementation design. The experimental results show that it can accomplish the reasoning of tens of billions of RDF triples on large clusters, and can accomplish many large data volume reasoning tasks that cannot be accomplished in a single computer environment.

Li et al. [85] proposed a new RDFS reasoning method based on the Spark context. Mcbrien et al. [86] used Spark to reasoning large ontologies in OWL. Similar studies include [82, 87, 88]. Common sense reasoning simulates the human cognitive ability, and multi-source knowledge fusion also includes integrating common sense into existing knowledge graphs. However, this type of knowledge is quite extensive, and the integration with common sense knowledge bases will make query based on knowledge graphs more and more difficult and slow. Tran et al. [89] proposed a new fast subgraph matching method GPsense, which takes advantage of the large-scale parallel processing capability of modern GPUs. It is designed for a scalable large-scale parallel architecture and can support the next generation of large data sentiment analysis and natural language processing applications [90]. used a common sense knowledge base to solve real-time multimodal analysis problems. In particular, the problem of multimodal sentiment analysis includes simultaneous analysis of different emotional and polarity detection methods, such as voice and video. Graph traversal based on GPUs can quickly extract important features from multi-modal sources. The experimental results on YouTube dataset show that the accuracy of this method is better than that of previous systems. In terms of processing speed, compared with the corresponding method based on CPU, the feature extraction method has several orders of magnitude improvements.

Liu et al. [91] proposed a flow reasoning method based on a large number of RDF data, which simplifies the flow reasoning problem into a time reasoning problem and uses graphics processing units (GPUs) to improve performance. Donkal et al. [92] proposed a multi-modal fusion framework based on Spark to ensure fast processing of large data in a parallel computing environment. The experimental results in intrusion detection systems show that compared with the existing technologies, the accuracy of data and test time have obvious advantages. Ju et al. [93] designed and implemented the RDFS reasoning and the RETE algorithm in parallel with Apache Spark.

Large-scale knowledge graphs have been widely used in intelligent search, intelligent QA and other fields. In order to compute large-scale knowledge graphs with millions of entities and facts, knowledge graphs need to be partitioned. However, the existing partitioning algorithms are difficult to meet the requirements of both partition efficiency and partition quality. Based on the power-law-distribution of social networks in the real world, Zhong et al. [94] proposed a graph-partitioning algorithm based on message cluster and stream partitioning (MCS). Compared with the traditional algorithm, the partition quality of MCS is closer to or

even more than that of the Metis package. In terms of partitioning efficiency, the PageRank algorithm in the Spark cluster system is used to calculate twitter graphics data. The total time of MCS is lower than that of hash partitioning. With the increase of iteration times, the effect is more obvious, which proves the effectiveness of MCS. For Qualitative Spatial Temporal Reasoning (QSTR), most of the work is focused on a relatively small constrained network composed of hundreds or mostly thousands of relationships. With the emergence of qualitative spatial temporal knowledge graphs with hundreds of thousands or even millions of relationships, traditional QSTR cannot carry out such large-scale reasoning. Mantle et al. [95] put forward the a parallel and distributed QSTR technology, PARQR, and implemented it using Apache Spark framework. The effectiveness of this method is proved in large-scale synthetic data sets and real KGs.

Through the incremental reasoning algorithm KGRL Incre, [96, 97] effectively updated the previous reasoning results incrementally, avoiding the complete re-reasoning of the extended KG. This method filters irrelevant triples, reduces the size of data to be processed, and a delayed reasoning strategy, which limits the number of iterations and keeps the relative integrity of the final results. Through a large number of experiments and comprehensive evaluation, the experimental results show that KGRL increment can significantly reduce the time consumption compared with the extended reasoning method in the target scenario.

Multi-source knowledge fusion is a challenging task. Although parallel processing technology has been applied to knowledge graph related research, the existing technology pays more attention to knowledge reasoning, and there are still many problems to be studied and solved on how to establish a large-scale knowledge fusion framework.

6 Concluding remarks

Knowledge graphs is essentially a large-scale semantic network, which is the basis of machine cognitive intelligence. The main goal of knowledge graphs is to describe various entities and concepts existing in the real world, as well as their relationship, and to express knowledge in a form closer to the human cognitive world. It is widely used in intelligent search, personalized recommendation, intelligent question answering and other fields. Multi-source knowledge fusion can effectively promote the study and development of KGs in the related domains such as Big Search in Cyberspace, NLP and so forth, effectively promote the construction of domain knowledge graphs, and bring gigantic social effect and huge economic benefits.

References

1. Dong, X. L. , & Srivastava, D.: Knowledge Curation and Knowledge Fusion: Challenges, Models and Applications[J], (2015)
2. Wang, H. , Fang, Z. , Zhang, L. , Pan, J. Z. , & Ruan, T.: Effective online knowledge graph fusion. In: Proceedings of ISWC, pp. 286–302. (2015)
3. Dong, X.L., Gabrilovich, E., Heitz, G., Horn, W., Murphy, K., Sun, S., et al.: From data fusion to knowledge fusion[J]. Proceedings of the VLDB Endowment. 7(10), 881–892 (2014)
4. Dong, X. , & Naumann, F.: Data Fusion - Resolving Data Conflicts for Integration[J]. Proceedings of the Vldb Endowment, 2(2),1654–1655(2009)
5. Zhou, F., Wang, P.B. , &Han, L.Y .:Multi-source knowledge fusion algorithm[J]. Journal of Beijing University of Aeronautics & Astronautics, (2013). (In Chinese)
6. Dempster A P .: Upper and Lower Probabilities Induced By A Multivalued Mapping[J]. Annals of Mathematical Statistics, 38 (1967)
7. Rota G C.: A mathematical theory of evidence: G. Shafer, Princeton University Press, pp.297(1976). [J]. Advances in Mathematics, 24(3),341–341 (1977)
8. Couso, I., Sánchez, L.: Upper and lower probabilities induced by a fuzzy random variable[J]. Fuzzy Sets & Systems. 165(1), 1–23 (2011)
9. Keller, J. M. , Liu, D. , & Fogel, D. B.: Fuzzy Measures and Fuzzy Integrals[C]// Wiley-IEEE Press, (2000)
10. Tahani, H., Keller, J.M.: Information fusion in computer vision using the fuzzy integral[J]. IEEE Transactions on Systems, Man and Cybernetics. 20(3), 733–741 (1990)
11. Lao, N. , Mitchell, T. M. , & Cohen, W. W.: Random Walk Inference and Learning in a Large Scale Knowledge Base[C]// Conference on Empirical Methods in Natural Language Processing. (2011)
12. Zhao, B., Han, J.: A Probabilistic Model for Estimating Real-Valued Truth from Conflicting Sources[J]. Proc. of QDB, (2012)
13. Dong, X. , Gabrilovich, E. , Heitz, G. , Horn, W. , Lao, N. , & Murphy, K. , et al.: Knowledge Vault: a Web-Scale Approach to Probabilistic Knowledge Fusion[J]. (2014)
14. Dong, X.L., Gabrilovich, E., Murphy, K., Dang, V., Horn, W., Lugaresi, C., et al.: Knowledge-based trust: estimating the trustworthiness of web sources[J]. Proceedings of the VLDB Endowment. 8(9), 938–949 (2015)
15. Zhang, Y., Liu, L., & Fu, S., et al.: Entity Alignment Across Knowledge Graphs Based on Representative Relations Selection[C]// 5th ICSAI. IEEE: 1056–1061. (2018)
16. Wang, X.P., Liu, K., He, S.Z., Liu, S.L., Zhang, Y.Z., & Zhao, J.: Multi-Source Knowledge Bases Entity Alignment by Leveraging Semantic Tags[J]. Chinese Journal of Computers, (2017).(In Chinese)
17. Sun, M. , Zhu, H. , Xie, R. , & Liu, Z.: Iterative Entity Alignment Via Joint Knowledge Embeddings[C]// International Joint Conference on Artificial Intelligence. AAAI Press, (2017)
18. Bordes, A., Usunier. N., Garcia-Duran, A., et al.: Translating embeddings for modeling multi-relational data[C]//Advances in neural information processing systems, pp.2787–2795(2013)
19. Lin, Y. , Liu, Z. , Luan, H. , Sun, M. , Rao, S. , & Liu, S. .: Modeling Relation Paths for Representation Learning of Knowledge Bases[C], in Proceedings of EMNLP. (2015)
20. Sun, Z. , Hu, W. , & Li, C.: Cross-lingual entity alignment via joint attribute-preserving embedding. In: Proceedings of ISWC, pp.628–644 (2017)
21. Zhong, Z.X., Cao, Y., Guo, M., & Nie, Z.Q.: CoLink: an Unsupervised Framework for User Identity Linkage[C]//Thirty-Second AAAI Conference on Artificial Intelligence. (2018)
22. Trsedya, B., Qi, J.Z., &Zhang, R.: Entity Alignment between Knowledge Graphs Using Attribute Embeddings ,AAAI. (2019)
23. Kong, C. , Gao, M. , Chen, X. U. , Yunbin, F. U. , Qian, W. , & Zhou, A.: EnAli: entity alignment across multiple heterogeneous data sources[J]. Frontiers of Computer Science, 13(1). (2019)
24. Wang, L. L. , Bhagavatula, C. , Neumann, M. , Lo, K. , Wilhelm, C. , & Ammar, W.: Ontology Alignment in the Biomedical Domain Using Entity Definitions and Context[J]. arXiv preprint arXiv:1806.07976, (2018)
25. Wu, F.Z., Huang, Y.F., Yuan, Z.G.: Domain-specific sentiment classification via fusing sentiment knowledge from multiple sources[J]. Information Fusion. 35, 26–37 (2017)
26. Wang, Z., Zhang, J., Feng, J., Chen, Z.: Knowledge Graph Embedding by Translating on Hyperplanes [C]// Twenty-Eighth Aaai Conference on Artificial Intelligence. AAAI Press (2014)
27. Mikolov, T., Sutskever, I., Chen, K., Corrado, G., Dean, J.: Distributed representations of words and phrases and their compositionality[J]. Adv. Neural Inf. Proces. Syst. 26, 3111–3119 (2013)

28. Mikolov, T., Chen, K., Corrado, G., et al.: Efficient Estimation of Word Representations in Vector Space[J]. Computer Science (2013)

29. Zhong, H., Zhang, J., &Wang, Z., et al.: Aligning Knowledge and Text Embeddings by Entity Descriptions[C], in Proceedings of EMNLP, Pp 267–272. (2015)

30. Sun, K., Zhu, Y., Song, J.: Progress and challenges on entity alignment of geographic knowledge bases[J]. ISPRS Int. J. Geo Inf. **8**(2), 77 (2019)

31. Guo, L., Sun, Z., Cao, E., & Hu, W.: Recurrent Skipping Networks for Entity Alignment[J]. (2018)

32. Guan, S., Jin, X., Wang, Y., Jia, Y., Cheng, X.: Self-learning and embedding based entity alignment[J]. Knowl. Inf. Syst. **24**, 1–26 (2018)

33. Yang, C., Liu, Z., Zhao, D., Sun, M., & Chang, E.: Network Representation Learning with Rich Text Information[C]// International Conference on Artificial Intelligence. AAAI Press, (2015)

34. Tu, C., Liu, H., & Liu, Z., et al.: CANE: Context-Aware Network Embedding for Relation Modeling[C]// Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). (2017)

35. Zhang, F., Yuan, N.J., Lian, D., Xie, X., Ma, W.Y.: Collaborative Knowledge Base Embedding for Recommender Systems[C]// the 22nd ACM SIGKDD International Conference. ACM (2016)

36. Lin, Y., Liu, Z., Sun, M., Liu, Y., & Zhu, X.: Learning Entity and Relation Embeddings for Knowledge Graph Completion[C]// Twenty-Ninth AAAI Conference on Artificial Intelligence. (2015)

37. Vincent, P., Larochelle, H., Lajoie, I., Bengio, Y., Manzagol, P.A.: Stacked Denoising Autoencoders: Learning Useful Representations in a Deep Network with a Local Denoising Criterion[J]. Journal of Machine Learning Research. **11**(12), 3371–3408 (2010)

38. Kristiadi, A., Khan, M. A., Lukovnikov, D., Lehmann, J., & Fischer, A.: Incorporating Literals into Knowledge Graph Embeddings[J], (2018)

39. Xie, R.B., Liu, Z.Y., Jia, J., Luan, H.B., &Sun, M.S.: Representation learning of knowledge graphs with entity descriptions[C], in Proceedings of AAAI, (2016)

40. Collobert, R., &Weston, J.: A unified architecture for natural language processing: Deep neural networks with multitask learning[C]//Proceedings of the 25th international conference on Machine learning. ACM, 160–167. (2008)

41. Collobert, R., Weston, J., Bottou, L., Karlen, M., Kavukcuoglu, K., Kuksa, P.: Natural language processing (almost) from scratch[J]. Journal of Machine Learning Research. **12**(1), 2493–2537 (2011)

42. Lv, X., Hou, L., Li, J., & Liu, Z.: Differentiating Concepts and Instances for Knowledge Graph Embedding[J]. (2018)

43. Guo, S., Wang, Q., &Wang, L., et al.: Jointly Embedding Knowledge Graphs and Logical Rules[C]// Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. (2016)

44. Demeester, T., Rocktäschel, Tim, & Riedel, S.: Lifted rule injection for relation embeddings. In Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing (EMNLP), pp.1389–1399 (2016)

45. Rocktäschel, T., Singh, S., & Riedel, S.: Injecting Logical Background Knowledge into Embeddings for Relation Extraction[C]// North American Chapter of the Association for Computational Linguistics. (2015)

46. Newman-Griffis, D., Lai, A. M., &Fosler-Lussier, E.: Jointly Embedding Entities and Text with Distant Supervision[J]. (2018)

47. Guo, S., Wang, Q., Wang, B., Wang, L., Guo, L.: SSE: semantically smooth embedding for knowledge graphs. IEEE Transactions on Knowledge & Data Engineering, Journal. **29**(4), 884–897 (2017)

48. Xie, R.B., Liu, Z.Y., Sun, M.S.: Representation Learning of Knowledge Graphs with Hierarchical Types[C]// International Joint Conference on Artificial Intelligence. AAAI Press (2016)

49. Xu, J., Chen, K., Qiu, X., & Huang, X.: Knowledge Graph Representation with Jointly Structural and Textual Encoding[J]. (2016)

50. Wang, Z., & Li, J.: Text-Enhanced Representation Learning for Knowledge Graph[C]// International Joint Conference on Artificial Intelligence. AAAI Press, (2016)

51. Wang, Z., Zhang, J., Feng, J., & Chen, Z.: Knowledge graph and text jointly embedding. EMNLP2014, pp 1591–1601 (2014)

52. Neelakantan, A., Roth, B., & Mccallum, A.: Compositional Vector Space Models for Knowledge Base Completion[J]. Computer Science, 1–16. (2015)

53. Guu, K., Miller, J., Liang, P.: Traversing Knowledge Graphs in Vector Space[J]. Computer Science (2015)

54. Lin, Y., Liu, Z., Luan, H., Sun, M., Rao, S., Liu, S.: Modeling Relation Paths for Representation Learning of Knowledge Bases[J]. Computer Science (2015)

55. Knowledge acquisition, https://en.wikipedia.org/wiki/Knowledge_acquisition, last accessed, 2019/04/10

56. Zhang, L., Fu, S., Jiang, S., Bao, R., & Zeng, Y.: A Fusion Model of Multi-Data Sources for User Profiling in Social Media[J]. (2018)

57. Tang, J. , Qu, M. , Wang, M. , Zhang, M. , Yan, J. , & Mei, Q.: LINE: Large-scale information network embedding[J]. 24th International Conference on World Wide Web, WWW 2015, (2015)

58. Smolensky, P.: Information Processing in Dynamical Systems: Foundations of Harmony Theory[R]. Colorado Univ at Boulder Dept of Computer Science (1986)

59. Welling, M. , Rosen-Zvi, M. , & Hinton, G.: Exponential family harmoniums with an application to information retrieval[C]//Advances in neural information processing systems. pp.1481–1488. (2005)

60. Salakhutdinov, R., & Hinton, G.: Deep boltzmann machines[C]//Artificial intelligence and statistics. pp.448–455. (2009)

61. Srivastava, N., &Salakhutdinov, R.: Multimodal learning with Deep Boltzmann Machines[C]// International Conference on Neural Information Processing Systems. Curran Associates Inc. 2012:2222–2230. (2012)

62. Srivastava ,N., & Salakhutdinov, R.: Learning representations for multimodal data with deep belief nets[A].//International Conference on Machine Learning Representation Learning Workshop[C],(2012)

63. Wang, F. , Qu, Y. , Zheng, L. , Lu, C. T. , & Yu, P. S.: Deep and broad learning on content-aware POI recommendation[C]//2017 IEEE 3rd International Conference on Collaboration and Internet Computing (CIC). IEEE, 369–378. (2017)

64. Zhang, Y., Wang, S., Yao, Y., Fang, B., Yu, P.S.: Improving stock market prediction via heterogeneous information fusion[J]. Knowl.-Based Syst. **143**, 236–247 (2018)

65. Ouyang, W. , Chu, X. , & Wang, X.: Multi-source Deep Learning for Human Pose Estimation[C]// IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, pp. 2337–2344. (2014)

66. Ngiam, J., Khosla, A., Kim, M., Nam, J., Lee, H. & Ng, A. Y.: Multimodal Deep Learning.. In L. Getoor & T. Scheffer (eds.), ICML ,pp. 689–696, : Omnipress. (2011)

67. Deng, L., Jia, Y., Zhou, B., Huang, J., Han, Y.: User interest mining via tags and bidirectional interactions on Sina Weibo[J]. World Wide Web. **21**(1), 1–22 (2017)

68. Zhu, J. , Zhang, J. , Zhang, C. , Wu, Q. , Jia, Y. , & Zhou, B. , et al.: CHRS: Cold Start Recommendation Across Multiple Heterogeneous Information Networks[J]. IEEE Access, PP(99):1–1. (2017)

69. Neverova, N., Wolf, C., Taylor, G.W., Nebout, F.: Moddrop: adaptive multi-modal gesture recognition[J]. IEEE Trans. Pattern Anal. Mach. Intell. **38**(8), 1692–1706 (2016)

70. Liu, Z. , Zhang, W. , Quek, T. Q. S. , & Lin, S.: Deep fusion of heterogeneous sensor data[C]// IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE, pp.5965–5969. (2017)

71. Wen, Y. , Yang, Y. , Lu, R. , & Wang, J.: Multi-Agent Generalized Recursive Reasoning[J]. (2019)

72. Chen, W., Xiong, W., Yan, X., & Wang, W.: Variational Knowledge Graph Reasoning[J]. (2018)

73. Kingma D P, Welling M.: Auto-Encoding Variational Bayes[J]. (2013)

74. Xiong, W. , Hoang, T. , & Wang, W. Y.: DeepPath: a Reinforcement Learning Method for Knowledge Graph Reasoning[J]. (2017)

75. Das, R. , Dhuliawala, S. , Zaheer, M. , Vilnis, L. , Durugkar, I. , & Krishnamurthy, A. , et al.: Go for a Walk and Arrive at the Answer: Reasoning over Paths in Knowledge Bases Using Reinforcement Learning[J]. (2017)

76. Das, R. , Neelakantan, A. , Belanger, D. , & Mccallum, A.: Chains of Reasoning over Entities, Relations, and Text Using Recurrent Neural Networks[J]. (2016)

77. Costa G A, de Oliveira J M P.: Linguistic Frames as Support for Entity Alignment in Knowledge Graphs[C]//Proceedings of the 20th International Conference on Information Integration and Web-based Applications & Services. ACM, pp.226–229. (2018)

78. Chen, M. , Tian, Y. , Yang, M. , & Zaniolo, C.: MTransE: Multilingual knowledge graph embeddings for cross-lingual knowledge alignment. In:Proceedings of IJCAI,pp. 1511–1517, (2017)

79. Chen, M. , Tian, Y. , Chang, K. W. , Skiena, S. , & Zaniolo, C.: Co-Training Embeddings of Knowledge Graphs and Entity Descriptions for Cross-Lingual Entity Alignment[J]. (2018)

80. Gouws ,S. , Bengio, Y. , & Corrado, G .: BilBOWA: Fast Bilingual Distributed Representations without Word Alignments[J]. Eprint Arxiv, pp.748–756. (2014)

81. Xu, K. , Wang, L. , Yu, M. , Feng, Y. , Song, Y. , & Wang, Z. , et al.: Cross-Lingual Knowledge Graph Alignment Via Graph Matching Neural Network[J]. (2019)

82. Wu, T. , Qi, G. , Wang, H. , Xu, K. , & Cui, X.: Cross-lingual taxonomy alignment with bilingual biterm topic model. In: Proceedings of AAAI, pp.287–293. (2016)

83. Wu, T. , Zhang, L. , Qi, G. , Cui, X. , & Xu, K.: Encoding category correlations into bilingual topic modeling for cross-lingual taxonomy alignment. In: Proceedings of ISWC, pp.728–744. (2017)

84. Zhang, Y. , Paradis, T. , Hou, L. , Li, J. , Zhang, J. , & Zheng, H.: Cross-Lingual Infobox Alignment in Wikipedia Using Entity-Attribute Factor Graph[J]. (2017)

85. Li, R. , Zhang, Q. , Wang, H. , & Wang, G.: Distributed RDFS Rules Reasoning for Large-Scaled RDF Graphs Using Spark[C]// International Conference on Service Science. IEEE Computer Society, (2016)

86. Mcbrien, P., & Liu, Y.: SPOWL: Spark-Based OWL 2 Reasoning Materialisation[C]// Acm Sigmod Workshop on Algorithms & Systems for Mapreduce & beyond. ACM, (2017)

87. Liu, Z. , Feng, Z. , Zhang, X. , Wang, X. , & Rao, G.: RORS: Enhanced Rule-Based OWL Reasoning on Spark.[C]// Asia-pacific Web Conference. Springer International Publishing, (2016)
88. Zhou, Z. , Qi, G. , Liu, C. , Mutharaju, R. , & Hitzler, P.: Reasoning with Large Scale OWL 2 EL Ontologies Based on MapReduce.[J]. (2016)
89. Tran, H.N., Cambria, E., Hussain, A.: Towards GPU-based common-sense reasoning: using fast subgraph matching[J]. Cogn. Comput. **8**(6), 1074–1086 (2016)
90. Tran, N. H. , & Cambria, E.: GPU-Based Commonsense Paradigms Reasoning for Real-Time Query Answering and Multimodal Analysis[J]. (2018)
91. Liu, C., Urbani, J., Qi, G.: Efficient RDF Stream Reasoning with Graphics Processing Units (GPUs). ACM, International Conference on World Wide Web (2014)
92. Donkal, G., Verma, G.K.: A multimodal fusion based framework to reinforce IDS for securing Big Data environment using Spark[J]. Journal of Information Security and Applications. **43**, 1–11 (2018)
93. Ju, H. , & Oh, S.: Enabling RETE Algorithm for RDFS Reasoning on Apache Spark[C]// 2018 IEEE 8th International Symposium on Cloud and Service Computing (SC2). IEEE Computer Society, (2018)
94. Zhong, J. , Wang, C. , Li, Q. , & Li, Q.: A New Graph-Partitioning Algorithm for Large-Scale Knowledge Graph[C]//: 14th International Conference, ADMA 2018, Nanjing, China, November 16–18, 2018, Proceedings. Advanced Data Mining and Applications. (2018)
95. Mantle, M., Batsakis, S., Antoniou, G.: Large scale distributed spatio-temporal reasoning using real-world knowledge graphs[J]. Knowl.-Based Syst. **163**, 214–226 (2019)
96. Wang, Y.F., Luo, J.: An incremental reasoning algorithm for large scale knowledge graph, in knowledge science, engineering and management. Cham. Switzerland: Springer. **2018**, 503–513 (2018)
97. Luo, J., Wang, Y.F., and Xu, Y.: Incremental Theory Closure Reasoning for Large Scale Knowledge Graphs[J].IEEE Access.99,1–1 (2019)

Publisher's note   Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Affiliations

Xiaojuan Zhao[1] · **Yan Jia**[1] · **Aiping Li**[1] · **Rong Jiang**[1] · **Yichen Song**[1]

Xiaojuan Zhao
zhaoxiaojuan18@nudt.edu.cn

Yan Jia
jiayanjy@vip.sina.com

Rong Jiang
jiangrong@nudt.edu.cn

Yichen Song
songyichen14@nudt.edu.cn

[1]   College of Computer, National University of Defense Technology, Changsha, China