A Survey of TransE-based Representation Learning Methods: Postprint

Authors: Zhang Zhenghang, Qian Yurong, Xing Yanni, Zhao Xin

Date: 2020-09-28T00:00:00+00:00

Abstract

To keep abreast of the latest research progress on TransE-based representation learning methods, this paper categorizes TransE-based representation learning methods into four types through induction and organization: methods based on complex relations, methods based on relation paths, methods based on image information, and methods based on other aspects. A detailed analysis is conducted on the design rationale, advantages, and disadvantages of each method. Simultaneously, a comparison and summary of public datasets and evaluation metrics for TransE-based representation learning methods is provided, along with a comparative analysis of the experimental performance of various TransE-based representation learning algorithms. Finally, the research of the entire paper is summarized, and future research hotspots are prospected. From the research results, the PaSKoGE method, NTransGH method, TCE method, and TransD method demonstrate the best performance on link prediction and triple classification tasks, warranting promotion and further expansion, with potential for further improvement in their path-specific embedding, two-layer neural network, triple context, and dynamically constructed mapping matrices.

Full Text

Preamble

Vol. 38 No. 3
Application Research of Computers
ChinaXiv Cooperative Journal

A Survey of Representation Learning Methods Based on TransE

Zhang Zhenghang[a,b], **Qian Yurong**[a,b†], **Xing Yanni**[a,b], **Zhao Xin**[a,b]

[a]College of Software; [b]Key Laboratory of Signal Detection & Processing in Xinjiang Uygur Autonomous Region, Xinjiang University, Urumqi 830046, China

Abstract: To keep abreast of the latest research progress in TransE-based representation learning methods, this paper systematically classifies these methods into four categories: methods based on complex relationships, methods based on relationship paths, methods based on image information, and methods based on other aspects. For each category, we provide a detailed analysis of the design rationale, advantages, and disadvantages. Additionally, we compare and summarize the common datasets and evaluation metrics used in TransE-based representation learning, and conduct a comparative analysis of the experimental performance of various TransE-based algorithms. Finally, we summarize the research findings and outline future research directions. The results indicate that the PaSKoGE, NTransGH, TCE, and TransD methods demonstrate the best performance on link prediction and triple classification tasks, warranting further promotion and extension. These methods can be further improved through path-specific embedding, two-layer neural networks, triple context, and dynamic mapping matrix construction.

Keywords: knowledge graph; representation learning; TransE model; knowledge graph embedding; translation model

CLC Number: TP
DOI: 10.19734/j.issn.1001-3695.2020.02.0028

0 Introduction

In recent years, inspired by word vector models, representation learning—particularly translation models—has garnered increasing attention in the field of knowledge graphs (KG). Knowledge representation involves mapping symbolic triples into a low-dimensional dense vector space, facilitating computation between entities and relations [?, ?, ?]. In such vector spaces, semantic similarity between entities (or relations) can be inferred by calculating distances, effectively addressing data sparsity challenges and making knowledge acquisition and reasoning more efficient and accurate [?, ?]. Moreover, knowledge representation learning research serves applications such as entity alignment, event extraction, and question answering systems, offering broad prospects [?].

Given the extensive application potential of knowledge representation, scholars have proposed numerous models, including Structured Embedding (SE), Single Layer Model (SLM), Semantic Matching Energy (SME), tensor decomposition models, and translation-based models [?]. Among these classical approaches, translation models represented by TransE [?] have received the most attention and become a current research hotspot. Proposed by Bordes in 2013, TransE demonstrates excellent performance in link prediction with fewer parameters and simple operation. However, TransE suffers from low accuracy when handling complex relationships, such as one-to-many, many-to-one, many-to-many, and reflexive relations, as it cannot precisely infer entities sharing the same relation [?, ?, ?].

Optimizing TransE has remained a popular research topic, with many novel models emerging annually [?, ?, ?, ?]. This survey classifies TransE-based representation learning methods into four categories based on TransE' s limitations: (a) methods for complex relationships, such as STransH, TransD, NTransGH, TransGraph, and TransAH [?, ?, ?]; (b) methods for relationship paths, such as PTransE and PaSKoGE [?, ?]; (c) methods for image information, such as ITMEA and TCE [?, ?]; and (d) other methods [?, ?, ?], such as TransRD [?], TransE-SNS [?], AST_{NZL} [?], and GTrans [?].

The main contributions of this paper are: 1. We introduce the algorithmic principles and pros/cons of TransE-based representation learning methods, providing a comprehensive and reasonable classification and summary [?, ?]. 2. We conduct a detailed analysis and overview of the problems inherent in TransE-based methods. 3. We analyze and summarize commonly used experimental datasets and performance evaluation metrics, compare key metrics across algorithms for each problem category, and identify methods worthy of further promotion and extension [?]. 4. We analyze and summarize current problems, solved issues, and potential future research directions in TransE-based representation learning algorithms [?].

---

1 Overview of TransE Representation Learning

TransE addresses the problem of embedding entities and multi-relational data in low-dimensional vector spaces by interpreting relations as translations operating on entity embeddings [?, ?]. Analyzing its fundamental principles, advantages, disadvantages, and algorithms deepens understanding of its mechanism.

1.1 TransE Method Overview

TransE is a classic method among knowledge representation translation models, with the Trans series extending from it. In TransE, entities and relations are mapped to a vector space, where their representations become vector operations [?, ?]. The core idea treats a triple (h, r, t) as a translation operation from head entity vector to tail entity vector [?, ?]. TransE makes several assumptions about mapping triples (h, r, t) to vector space (Figure [Figure 1: see original paper]), positing that each triple can be represented as:

$$\mathbf{h} + \mathbf{r} \approx \mathbf{t}$$

where $\mathbf{h}$ denotes the head entity vector, $\mathbf{r}$ the relation vector, and $\mathbf{t}$ the tail entity vector. Equation (1) signifies that in vector space, the head entity vector plus the relation should equal the tail entity vector. If this holds, the triple (h, r, t) is valid in the KG. TransE training requires:

For positive triples: $\mathbf{h} + \mathbf{r} \approx \mathbf{t}$
For negative triples: $\mathbf{h}' + \mathbf{r} \neq \mathbf{t}'$

The relationship between (h, r, t) and (h', r, t') approximates vector similarity. TransE employs Euclidean distance as the scoring function:

$$f(h, r, t) = \|\mathbf{h} + \mathbf{r} - \mathbf{t}\|_{L_1/L_2}$$

Lower scores benefit positive triples, while higher scores benefit negative triples. TransE then designs a loss function to evaluate representation learning effectiveness:

$$L = \sum_{(h,r,t)\in S} \sum_{(h',r,t')\in S'} [\gamma + f(h, r, t) - f(h', r, t')]_+$$

where S is the positive sample set, S' contains negative samples generated by randomly replacing head or tail entities, $\gamma > 0$ is the margin parameter, and $[x]_+ = \max(0, x)$. Through continuous training, TransE minimizes this loss function.

Algorithm 1: Learning TransE
Input: Training set S, entity set E, relation set L, margin γ, embedding dimension k
Initialize: $\mathbf{r} \sim \text{uniform}(-\frac{6}{\sqrt{k}}, \frac{6}{\sqrt{k}})$ for each $r \in L$, $\mathbf{e} \sim \text{uniform}(-\frac{6}{\sqrt{k}}, \frac{6}{\sqrt{k}})$ for each entity $e \in E$
Loop:
Sample a minibatch of size b
Initialize $T_{batch} \leftarrow \emptyset$
For $(h, r, t) \in S_{batch}$:
Sample $(h', r, t') \leftarrow \text{sample}(S'_{(h,r,t)})$
$T_{batch} \leftarrow T_{batch} \cup \{((h, r, t), (h', r, t'))\}$
Update embeddings w.r.t. $\sum_{((h,r,t),(h',r,t'))\in T_{batch}} \nabla[\gamma + f(h, r, t) - f(h', r, t')]_+$
End loop

1.2 Advantages and Disadvantages of TransE

TransE was proposed to handle multi-relational data, offering a simple and efficient KG representation learning method capable of various link prediction tasks [**?**, **?**]. Its simplicity demonstrates that KG representation learning can automatically capture reasoning features without manual design, making it suitable for large-scale complex KGs. However, TransE suffers from insufficient expressiveness. KG relations can be categorized as one-to-one, one-to-many, many-to-one, and many-to-many based on head-to-tail entity ratios, but TransE cannot effectively handle one-to-many, many-to-one, many-to-many, or reflexive relations. Additionally, TransE minimizes a margin loss function shared across all relation paths to determine entities, relations, and multi-step paths, which cannot consider differences between relation paths [**?**]. TransE also exhibits poor performance with image information, low-quality negative triples, slow convergence, poor generalization, and weak margin identification.

2 TransE-Based Representation Learning Methods

To address TransE's limitations, researchers have developed various optimization models, primarily focusing on complex relationships, relationship paths, and image information, with fewer works on other aspects.

2.1 Methods Based on Complex Relationships

Complex relationship issues refer to TransE' s inability to handle one-to-many, many-to-one, many-to-many, and reflexive relations. Classic methods include TransR, TransH, and TransA [**?**]. The conventional approach projects entities and relations into vector space before computing loss functions to infer semantic connections [**?**], but this suffers from inefficiency in link prediction and triple classification.

Five main solutions excel at link prediction and triple classification:

STransH (Single-layer Neural Network Embedding) embeds entities and relations in separate spaces, using single-layer neural network nonlinear operations to strengthen semantic connections. Inspired by TransH, it introduces relation-specific hyperplane projection, allowing entities to play different roles in different relations. *Advantages:* Relation-specific hyperplane projection. *Disadvantages:* Does not consider relation paths.

TransD (Dynamic Mapping Matrix Construction) represents each symbol (entity/relation) with two vectors: one for meaning and another for dynamically constructing mapping matrices, considering both relation diversity and entity characteristics (Figure [Figure 2: see original paper]). *Advantages:* Dynamic matrix construction provides flexible projection. *Disadvantages:* Ignores intrinsic relation correlations, has many parameters, and hinders knowledge sharing.

The TransD mapping function is:

$$\mathbf{M}_{rh} = \mathbf{r}_p \mathbf{h}_p^\top + \mathbf{I}, \quad \mathbf{M}_{rt} = \mathbf{r}_p \mathbf{t}_p^\top + \mathbf{I}$$

where $\mathbf{I}$ is the identity matrix. Head and tail entities use different mapping matrices $\mathbf{M}_{rh}$ and $\mathbf{M}_{rt}$, determined by relation vector $\mathbf{r}_p$ and entity mapping vectors $\mathbf{h}_p$ or $\mathbf{t}_p$. The projected vectors are:

$$\mathbf{h}_\perp = \mathbf{M}_{rh}\mathbf{h}, \quad \mathbf{t}_\perp = \mathbf{M}_{rt}\mathbf{t}$$

The TransD loss function becomes:

$$f(h, r, t) = \|\mathbf{h}_\perp + \mathbf{r} - \mathbf{t}_\perp\|_{L_1/L_2}$$

NTransGH (Generalized Hyperplane Translation Mechanism) combines a translation mechanism modeling relations as generalized hyperplanes with neural networks capturing complex entity-relation interactions. *Advantages:* Two-layer neural network for scoring functions enables low-dimensional embeddings. *Disadvantages:* Requires many parameters.

NTransGH uses basis vectors (instead of TransH's single normal vector) to define generalized hyperplanes (Figure Figure 3: see original paper). It introduces a two-layer neural network scoring function to find low-dimensional embeddings for complex relations. For triple (h, r, t), NTransGH projects entities onto a generalized hyperplane using basis vectors, but this increases parameter count.

TransGraph (Cross-Training Mechanism Embedding) uses stochastic gradient descent to optimize objective function L, updating vectors and parameters through gradient computation. It introduces a vector-sharing cross-training mechanism on top of TransE to deeply fuse network structure and triple information. *Advantages:* Integrates KG network structure features. *Disadvantages:* Ignores entity description text and web text.

TransAH (Hyperplane Model Embedding) first builds relation-specific hyperplane models, then embeds head and tail entities into these hyperplanes for relation reasoning. *Advantages:* Relation-specific hyperplane models effectively represent complex relations. *Disadvantages:* Does not consider text relation extraction or entity clustering tasks.

For complex relationships, STransH, TransD, NTransGH, TransGraph, and TransAH all show good performance. STransH uses single-layer neural networks but remains weak in predicting complex relations; adding hyperplane models significantly improves performance. TransD dynamically constructs relation projection matrices through outer products, reducing parameters in KGs with many relations but few entities while enhancing global feature capture. NTransGH' s core is its two-layer neural network scoring function—a trend in improving TransE' s performance. TransGraph' s vector-sharing cross-training mechanism is novel and worth referencing. TransAH needs strengthening in entity clustering tasks.

2.2 Methods Based on Relationship Paths

Relationship path issues arise because TransE only considers direct relations in representation learning, while multi-step paths contain rich inference patterns and ignore differences between paths [**?**]. Path-based optimization is a recent hotspot, though few methods excel at link prediction and triple classification.

Two main solutions perform well:

PTransE (Path Representation Learning) treats relation paths as translations between entities, representing paths through semantic composition of relation embeddings. *Advantages:* Builds triples from entity pairs connected by relation paths, facilitating reliable path selection. *Disadvantages:* Only considers direct relations and simple inference patterns.

PaSKoGE (Path-Specific Embedding) adaptively determines margin-based loss functions for each path by encoding correlations between relations and multi-step paths for any entity pair. *Advantages:* Path-specific embedding with adaptive margins per path. *Disadvantages:* Requires additional time to compute path-specific margins.

The PaSKoGE margin-varying objective function is:

$$\Delta_{opt}(h, r, t) = \mathbb{E}_{p \in P(h,t)}[\|h + r - t\| - \gamma_p + \|h' + r' - t'\|]_+$$

where $P(h, t)$ is the path set between h and t, and γ_p is the optimal margin separating positive and negative triples, defined similarly to TransA. Training uses SGD, with negative triples constructed by randomly replacing components. Modeling correlations between relations and multi-step paths requires extra time.

Both PTransE and PaSKoGE effectively address TransE' s path issues, with different emphases: PTransE focuses on path-as-translation and reliability measurement, while PaSKoGE minimizes path-specific margin losses. Both remain promising research directions.

2.3 Methods Based on Image Information

TransE treats KGs as independent triples during learning, failing to utilize entity features and numerous inter-triple connections. It also focuses on text while ignoring image information, leaving visual features underutilized. Typical solutions include ITMEA and TCE.

ITMEA (Multi-modal Entity Alignment) jointly embeds multi-modal (image, text) data using a combination of TransE and TransD, enabling multi-modal data embedding into low-dimensional semantic space. It iteratively learns alignment mappings between aligned multi-modal entities and applies them to unaligned entities. *Limitations:* No confidence evaluation for newly aligned entities during iteration, resulting in low iteration effectiveness.

TCE (Triple Context Exploitation) defines triple context comprising neighbor context and path context, evaluating correlations between triples and their contexts rather than using triples individually. It incorporates context into a scoring function assessing triple confidence. For each triple, two structural information types are considered as context: outgoing relations of entities and relation paths between entity pairs (Figure [Figure 4: see original paper]). *Limitations:* Underperforms compared to classical translation models.

Triple context (h, r, t) consists of head entity neighbor context and entity pair (h, t) path context, formalized as:

$$C(h, r, t) = C_h \cup C_{(h,t)}$$

The scoring function introduces triple context as conditional probability given context and all embeddings:

$$f(h,r,t) = P((h,r,t)|C(h,r,t);\Theta)$$

The objective maximizes joint probability of all triples in KG K:

$$\Theta = \arg\max_{\Theta} \prod_{(h,r,t)\in K} P((h,r,t)|C(h,r,t);\Theta)$$

Neighbor and path contexts are merged via:

$$P((h,r,t)|C(h,r,t)) \approx P(h|C_h) \cdot P(t|C_t) \cdot P(r|C_{(h,t)})$$

Both ITMEA and TCE have advantages: ITMEA strengthens multi-modal embedding, while TCE effectively utilizes KG structure, particularly local triple context. Addressing their limitations remains a key research focus.

2.4 Methods Based on Other Aspects

TransE also suffers from poor generalization, low-quality negative triples, slow convergence, and weak margin identification. Few scholars have addressed these issues, leaving substantial room for development.

Poor Generalization: More triples during training yield better embedding, but few triples result in insufficient updates and poor performance [**?**]. **TransRD (Unequal Transfer Matrix Embedding)** addresses this by projecting head/tail entities with unequal transfer matrices, using ADADELTA for learning rate adjustment and grouping relations to share projection matrices. *Advantages:* Uses shared projection matrices to capture relations, improving generalization by modeling intrinsic correlations. *Disadvantages:* Does not utilize KG relation path information.

Low-Quality Negative Triples: KGE training generates negatives by randomly replacing entities, producing many low-quality samples. **TransE-SNS (K-means Clustering)** clusters KG entities via K-means, then replaces head or tail entities with similar entities from the same cluster. *Advantages:* Similarity-based negative sampling improves entity similarity. *Disadvantages:* Difficult for sparse large-scale KGs.

Slow Convergence: Random training data may be hard to train, causing delays and poor model quality. **AST_{NZL} (Adaptive Selection)** selects relation categories probabilistically, then randomly samples instances from selected groups, adaptively adjusting selection probabilities based on training effectiveness. *Advantages:* "Loss non-zero" mechanism improves probability estimation accuracy. *Disadvantages:* Does not explore relation grouping.

Weak Margin Identification: Poor relation representation with obvious noise from other relations. **GTrans (Generic Translation Framework)** explains each entity through intrinsic state (inherent features) and mimic state (relation-influenced features), building dynamic relation spaces by assigning different weights to entities. *Advantages:* Dynamic relation spaces provide flexibility and reduce noise. *Disadvantages:* Only considers single triple extraction.

TransRD, TransE-SNS, AST_{NZL}, and GTrans each address specific TransE limitations effectively. However, these remain challenging problems with limited breakthroughs. Future research may combine these methods with attention mechanisms and convolutional neural networks.

---

3 Experiments on TransE-Based Representation Learning Algorithms

This section introduces common datasets, evaluation metrics, and comparative analyses of the aforementioned algorithms [?, ?, ?].

3.1 Common Datasets

Standard entity-relation datasets are required for scientific evaluation [?]. Commonly used datasets include:

- **WN18:** WordNet subset with 18 relations and 40,943 entities
- **FB15K:** Dense FreeBase subset with 1,345 relations and 14,951 entities
- **WN11:** WordNet subset with 11 relations and 38,696 entities
- **FB13:** Dense FreeBase subset with 13 relations and 75,043 entities
- **FB40K:** Dense FreeBase subset with 1,336 relations and 39,528 entities
- **MPBC_{20}:** 20 relations and 175,624 entities
- **FB15K-237:** FreeBase subset with 237 relations and 14,541 entities

Table summarizes these datasets, comparing relations, entities, training/validation/test sets.

3.2 Evaluation Metrics

Metrics are categorized into Mean Rank, Hits@10, and Accuracy [?, ?, ?]. Mean Rank and Hits@10 measure link prediction, while Accuracy measures triple classification [?, ?].

a) Mean Rank: Average rank of correct entities (lower is better) [?].
b) Hits@10: Probability of correct entities ranking in top 10 (higher is better).
c) Accuracy: For triple classification:

$$ACC = \frac{T_p + T_n}{N_{pos} + N_{neg}}$$

where T_p and T_n are correctly predicted positive/negative triples, and N_{pos}, N_{neg} are total positive/negative triples. Higher ACC indicates better classification [?, ?, ?].
d) Running Time: Compares method efficiency (lower is better) [?, ?].

3.3 Algorithm Analysis and Comparison

Table compares representative TransE-based algorithms, including classification, publication year, datasets, metrics, principles, and pros/cons [**?**, **?**].

3.4 Experimental Results

Experiments compare methods on link prediction (Table) and triple classification (Table) using public datasets. Parameters: learning rate $\alpha \in \{0.002, 0.005, 0.01\}$, margin $\gamma \in \{0.25, 0.5, 1, 2\}$, dimension $k \in \{50, 75, 100\}$, weight $\eta \in \{0.05, 0.0625, 0.25, 1.0\}$, batch size $B \in \{20, 75, 200, 1200, 4800\}$.

Link Prediction: On WN18, TransD, NTransGH, TransGraph, PaSKoGE, TransRD, TransE-SNS, and GTrans outperform TransE in MeanRank, with PaSKoGE best; for Hits@10, STransH, TransD, NTransGH, TransGraph, PaSKoGE, TransRD, TransE-SNS, AST_{NZL}, and GTrans outperform TransE, with NTransGH best [**?**, **?**, **?**]. On FB15K, STransH, TransD, NTransGH, TransGraph, TransAH, PTransE, PaSKoGE, TCE, TransRD, TransE-SNS, and GTrans outperform TransE in MeanRank, with TCE best; all methods improve Hits@10 over TransE, with PaSKoGE best [**?**, **?**]. **Conclusion:** PaSKoGE, TCE, and NTransGH are worth promoting.

Triple Classification: On WN11, STransH, TransD, NTransGH, TransGraph, TransAH, and TransE-SNS outperform TransE, with NTransGH best. On FB13, STransH, TransD, NTransGH, TransGraph, TransAH, and TransE-SNS outperform TransE, with TransD best. On FB15K, STransH, TransD, NTransGH, TransAH, and TransE-SNS outperform TransE, with NTransGH best [**?**, **?**]. **Conclusion:** NTransGH and TransD are worth promoting.

---

4 Summary and Outlook

TransE-based representation learning has become a KG research hotspot. This paper provides a scientific analysis and summary, introducing methods for complex relationships, relationship paths, image information, and other aspects. PaSKoGE, NTransGH, TCE, and TransD show the best performance on link prediction and triple classification, warranting further extension through path-specific embedding, two-layer neural networks, triple context, and dynamic mapping matrices.

Future research trends include:

1) Gaussian Mixture Embedding Methods
TransE struggles with multiple relation semantics where one relation may have multiple meanings across entity pairs. We propose **TransGM**, a novel Gaussian mixture model that leverages mixture of relation component vectors for embedding factual triples. Experiments show this model requires further improvement.

2) Fewer-Parameter Single-Layer Neural Network Methods
While deep neural networks can enhance entity description embedding for TransE, they require additional parameter storage and hyperparameter tuning. We propose a single-layer, fewer-parameter model that maximizes log-likelihood of observed knowledge by measuring triple probabilities with entity descriptions, simultaneously learning embeddings for entities, relations, and description words. Experiments show research significance.

3) Separate Entity and Relation Space Embedding Methods
Models like TransE treat relations as translations in a shared semantic space. However, entities have multiple aspects, and relations focus on different aspects, making shared spaces insufficient. We propose **TransR** to build separate entity and relation spaces, projecting entities into relation spaces before learning translations. Experiments show this model needs further refinement.

---

5 Conclusion

TransE cannot effectively handle one-to-many, many-to-one, many-to-many, or reflexive relations, and suffers from poor image information processing, low-quality negative triples, slow convergence, poor generalization, and weak margin identification. This survey classifies TransE-based methods into four categories based on these problems and validates their effectiveness through link prediction and triple classification experiments. Future work will test effectiveness on entity alignment.

References

[1] Sabina K, Leonids N, Natalya P, Vladimir U, Colleen Heinemann. Intelligent Collaborative Educational Systems and Knowledge Representation [J]. *Procedia Computer Science*, 2017 (104): 166–173.

[2] Kargin A, Petrenko T. Knowledge Representation in Smart Rules Engine [C]// 2019 3rd International Conference on Advanced Information and Communications Technologies (AICT), Lviv: IEEE Press, 2019: 231-238.

[3] Ren Lijuan, Lu Jun, Guo Wei. Multi-source Knowledge Embedding Research of Knowledge Graph [C]// 2019 IEEE 3rd International Conference on Circuits, Systems and Devices (ICCSD), Ukraine, 2019: 234-239.

[4] Cesar S, Zhang Haoxi, Imran S, Md M W, Caterine Silva de Oliveira, Edward Szczerbicki. Experience based knowledge representation for Internet of Things and Cyber Physical Systems with case studies [J]. *Future Generation Computer Systems*, 2019 (92): 604–616.

[5] Miao Fan, Qiang Zhou, Thomas F Z, Ralph G. Distributed representation

learning for knowledge graphs with entity descriptions [J]. *Pattern Recognition Letters*, 2017 (93): 31–37.

[6] Duc-Hong P, Anh-Cuong L. Learning multiple layers of knowledge representation for aspect based sentiment analysis [J]. *Data & Knowledge Engineering*, 2018 (114): 26–39.

[7] Chen Xiaojun, Xiang Yang. STransH: An improved knowledge representation model based on translation model [J]. *Computer Science*, 2019, 46 (09): 184-189.

[8] Bordes A, Usunier N, Garcia-Duran A, Weston J, Yakhnenko O. Translating Embeddings for Modeling Multi-relational Data [C]// NIPS, 2013: 2787-2795.

[9] Wang Zhen, Zhang Jianwen, Feng Jianlin, Chen Zheng. Knowledge Graph Embedding by Translating on Hyperplanes [C]// AAAI, 2014: 1112-1119.

[10] Lin Yankai, Liu Zhiyuan, Sun Maosong, Liu Yang, Zhu Xuan. Learning Entity and Relation Embeddings for Knowledge Graph Completion [C]// AAAI, 2015: 2181-2187.

[11] Ji Guoliang, He Shizhu, Xu Liheng, Liu Kang, Zhao Jun. Knowledge Graph Embedding via Dynamic Mapping Matrix [C]// ACL, 2015: 687-696.

[12] Xiao Han, Huang Minlie, Zhu Xiaoyan. TransG: A Generative Model for Knowledge Graph Embedding [J]. *Computer Science*, 2016 (1): 2316-2325.

[13] Zhang Yi, Cao Wanhua, Liu Juntao. A Novel Negative Sample Generating Method for Knowledge Graph Embedding [C]// Proceedings of the 2019 International Conference on Embedded Wireless Systems and Networks, 2019: 401-406.

[14] Kohei I, Yukihisa K, Shinsuke K, Atsushi S. Structuring Engineers' Implicit Knowledge of Forming Process Design by Using a Graph Model [J]. *Procedia CIRP*, 2018 (67): 563-568.

[15] Wu Zhenyong, Liao Jihua, Song Wenyan, Mao Hanling, Huang Zhenfeng, Li Xinxin, Mao Hanying. Semantic Hyper-graph-based Knowledge Representation Architecture for Complex Product Development [J]. *Computers in Industry*, 2018 (100): 43-56.

[16] Ji Guoliang, Liu Kang, He Shizhu, Zhao Jun. Knowledge Graph Completion with Adaptive Sparse Transfer Matrix [C]// AAAI, 2016: 985-991.

[17] Zhu Qiannan, Zhou Xiaofei, Zhang Peng, Shi Yong. A Neural Translating General Hyperplane for Knowledge Graph Embedding [J]. *Journal of Computational Science*, 2019 (30): 108–117.

[18] Fang Yang, Zhao Xiang, Tan Zhen, Yang Shiyu, Xiao Weidong. An Improved Representation Method of Knowledge Map Based on Translation [J]. *Computer Research and Development*, 2018 (55): 139-150.

[19] Lin Yankai, Liu Zhiyuan, Luan Huanbo, Sun Maosong, Rao Siwei, Liu Song. Modeling Relation Paths for Representation Learning of Knowledge Bases [C]//

EMNLP, 2015: 705-714.

[20] Jia Yantao, Wang Yuanzhuo, Jin Xiaolong, Cheng Xueqi. Path-specific Knowledge Graph Embedding [J]. *Knowledge-Based Systems*, 2018 (151): 1-12.

[21] Wang Huiyong, Lun Bing, Zhang Xiaoming, Sun Xiaoling. Multimodal Entity Alignment Based on Joint Knowledge Representation Learning [J]. *Control and Decision*, 2019: 1-10.

[22] Gao Shiun, Qi Huan, Zhou Guilin, Zhang Quan. Knowledge Graph Embedding with Triple Context [C]// CIKM, 2017: 2299-2302.

[23] Liu Huchen, Xue Luan, Li Zhiwu, Wu Jianing. Linguistic Petri Nets Based on Cloud Model Theory for Knowledge Representation and Reasoning [J]. *IEEE Transactions on Knowledge and Data Engineering*, 2018, 30 (4): 717-728.

[24] Pan Yunhe. Multiple Knowledge Representation of Artificial Intelligence [J]. *Engineering*, 2019 (3): 2095-8099.

[25] Zhu Yanli, Yang Xiaoping, Wang Liang, Zhang Zhiyu. TransRD: A Knowledge Graph Embedding Representation Model with Unequal Features [J]. *Chinese Journal of Information Technology*, 2019, 33 (11): 73-82.

[26] Rao Guanjun, Gu Tianlong, Chang Liang, Bin Chenzhong, Qin Saige, Xuan Wen. Knowledge Graph Embedding Based on Similarity Negative Sampling [J]. *Journal of Intelligent Systems*, 2019, 23 (10): 1-9.

[27] Ouyang Dantong, Ma Xiang, Lei Jingpei, Feng Shasha. Adaptive Selection in Knowledge Graph Embedding [J]. *Journal of Jilin University (Engineering Edition)*, 2019: 1-8.

[28] Zhen Tan, Xiang Zhao, Yang Fang, Xiao Weidong. GTrans: Generic Knowledge Graph Embedding via Multi-State Entities and Dynamic Relation Spaces [J]. *IEEE Access*, 2018 (6): 2169-3536.

[29] Zhang Wen. Knowledge Graph Embedding with Diversity of Structures [C]// WWW' 17 Companion, 2017: 747–753.

[30] Katrin J, André B. Semi-automated Site Equipment Selection and Configuration Through Formal Knowledge Representation and Inference [J]. *Advanced Engineering Informatics*, 2018 (38): 488–500.

[31] Loredana C, Giuseppe P, Genoveffa T, Daniele I. A Knowledge Representation Framework to Enhance Automatic Video Surveillance [J]. *Expert Systems With Applications*, 2019 (131): 190–207.

[32] Tang Xing, Chen Ling, Cui Jun, Wei Baogang. Knowledge Representation Learning with Entity Descriptions, Hierarchical Types, and Textual Relations [J]. *Information Processing and Management*, 2019 (56): 809–822.

[33] Zhao Chao, Jiang Jingchi, Guan Yi, Guo Xitong, He Bin. EMR-based Medical Knowledge Representation and Inference via Markov Random Fields

and Distributed Representation Learning [J]. *Artificial Intelligence in Medicine*, 2018 (87): 49–59.

[34] Xaro B, Angel C, Esther V, Ana G S, Juan C. FCA-based Knowledge Representation and Local Generalized Linear Models to Address Relevance and Diversity in Diverse Social Images [J]. *Future Generation Computer Systems*, 2019 (100): 250–265.

[35] Pablo A Q, Jessica C M, Concepcion R, Juan C M. A Body of Knowledge Representation Model of Ecotourism Products in Southeastern Ecuador [J]. *Heliyon*, 2018 (4): 1063.

[36] Nie Binling, Sun Shouqian. Knowledge Graph Embedding via Reasoning over Entities, Relations, and Text [J]. *Future Generation Computer Systems*, 2019 (91): 426-433.

[37] Zhang Wen. Knowledge Graph Embedding with Diversity of Structures [C]// WWW' 17 Companion, 2017: 747–753.

[38] Xia Xiaoqiang, Zhang Dehai, Liu Qing, Wang Yanpeng, Cui Menglong. Synergistic Union of Word Embedding and Knowledge Graph for Words Semantic Similarity Measure [C]// ICCC, 2018: 2349-2353.

[39] Guan Niannian, Song Dandan, Liao Lejian. Knowledge Graph Embedding with Concepts [J]. *Knowledge-Based Systems*, 2019 (164): 38-44.

[40] He Ming, Du Xiangkun, Wang Bo. Representation Learning of Knowledge Graphs via Fine-Grained Relation Description Combinations [J]. *IEEE Access*, 2019 (7): 26466-26473.

[41] Arpad G, Adrian F, Ugo F, Paolo Z, Lucian V. Performance and Energy Optimisation in CPUs Through Fuzzy Knowledge Representation [J]. *Information Sciences*, 2019 (476): 375–391.

[42] Liu Wenqiang, Liu Jun, Wu Mengmeng, Samar A, Hu Wei, Wei Bifan, Zheng Qinghua. Representation Learning over Multiple Knowledge Graphs for Knowledge Graphs Alignment [J]. *Neurocomputing*, 2018 (320): 12–24.

[43] David P, Sun Yu. A Survey of Knowledge Representation in Service Robotics [J]. *Robotics and Autonomous Systems*, 2019 (118): 13–30.

[44] Debarpita S, Swapan K B, Jyotsna K M, Subrata G. Rough Set Based Lattice Structure for Knowledge Representation in Medical Expert Systems: Low Back Pain Management Case Study [J]. *Expert Systems With Applications*, 2020 (145): 113084.

[45] Kaushalya K, Nikola K, Denise T. Deep Learning and Deep Knowledge Representation in Spiking Neural Networks for Brain-Computer Interfaces [J]. *Neural Networks*, 2020 (121): 169–185.

[46] Ji Guoliang, Liu Kang, He Shizhu, Zhao Jun. Knowledge Graph Completion with Adaptive Sparse Transfer Matrix [C]// AAAI, 2016: 985-991.

[47] Jiang Feng, Gu Qing, Hao Huizhen, Li Na, Guo Yanwen, Chen Daochu. Overview of Content-based Image Segmentation Methods [J]. *Journal of Software*, 2017, 28 (01): 160-183.

[48] Wang Fan, Li Heng, Dong Chao, Ding Lieyun. Knowledge Representation Using Non-parametric Bayesian Networks for Tunneling Risk Analysis [J]. *Reliability Engineering and System Safety*, 2019 (191): 106-115.

[49] Didem G, Aneta V F, Jad E, Swarup K M. Knowledge Representation of Cyber-physical Systems for Monitoring Purpose [J]. *Procedia CIRP*, 2018 (72): 468–473.

[50] A E Suleimankadieva, V I Pilipenko, J Sági. Knowledge Company: Approaches to Assessing New Knowledge and Representation it to Society [J]. *Procedia Computer Science*, 2019 (150): 730–736.

[51] Han Xiao, Zhang Chunhong, Sun Ting, Ji Yang, Hu Zhen. A Triple-Branch Neural Network for Knowledge Graph Embedding [J]. *IEEE Access*, 2018 (6): 76606-76615.

[52] Qiu Jiangnan, Zuo Min, Yan Shunin, Shi Huayan. A Qualitative Knowledge Representation Model and Application for Crisis Events [J]. *Procedia Computer Science*, 2018 (126): 1828–1836.

[53] Pridi S, Nathalie J, Peter H, George M. A General Framework Dedicated to Computational Morphogenesis Part II–Knowledge Representation and Architecture [J]. *BioSystems*, 2018 (173): 314–334.

[54] Nathaniel R, Philip M, Fernbach, Steven A. Sloman. Individual Representation in a Community of Knowledge [J]. *Trends in Cognitive Sciences*, 2019 (23): 10.

[55] Manuel C, Juan C, Darren P, John A. A New Paradigm for Uncertain Knowledge Representation by Plausible Petri Nets [J]. *Information Sciences*, 2018 (453): 323–345.

[56] Masita, Masila A J, Chia P L, Noor M M, Fatihah M. Knowledge Representation Model for Crime Analysis [J]. *Procedia Computer Science*, 2017 (116): 484–491.

[57] Amelec V, Omar B P. An Intelligent Approach for the Design and Development of a Personalized System of Knowledge Representation [J]. *Procedia Computer Science*, 2019 (151): 1225–1230.

[58] Honoré-Chedozeau C, Lelièvre-Desmas M, Ballester J, Chollet S, Valentin D. Knowledge Representation Among Assessors Through Free Hierarchical Sorting and a Semi-directed Interview: Exploring Beaujolais Wines [J]. *Food Quality and Preference*, 2017 (57): 17–31.

[59] Qi Jinda, Ding Lan, Samsung L. Ontology-based Knowledge Representation of Urban Heat Island Mitigation Strategies [J]. *Sustainable Cities and Society*, 2020 (52): 101875.

[60] Liu Huchen, You Jianxin, Li Zhiwu, Tian Guangdong. Fuzzy Petri Nets for Knowledge Representation and Reasoning: A Literature Review [J]. *Engineering Applications of Artificial Intelligence*, 2017 (60): 45–56.

[61] Giovannini A, Aubry A, Panetto H, Haouzi H, Canciglieri Junior O, Pierrel L. Knowledge Representation, Retrieval and Reuse for Product Family Design: An Anti-logicist Approach [J]. *Computers & Industrial Engineering*, 2016 (101): 391–402.

[62] J Montero, H Bustince, C Franco, J T Rodríguez, D Gómez, M Pagola, J Fernández, E Barrenechea. Paired Structures in Knowledge Representation [J]. *Knowledge-Based Systems*, 2016 (100): 50–58.

[63] Zhao Chao, Jiang Jingchi, Guan Yi, Guo Xitong, He Bin. EMR-based Medical Knowledge Representation and Inference via Markov Random Fields and Distributed Representation Learning [J]. *Artificial Intelligence in Medicine*, 2018 (87): 49–59.

[64] Yang Li, Geng Xinyu, Cao Xiedong. A Novel Knowledge Representation Model Based on Factor State Space [J]. *Optik*, 2016 (127): 5141–5147.

[65] Paul W, Paul M, Trevor C, Enrico M. Improving Comprehension of Knowledge Representation Languages: A Case Study with Description Logics [J]. *International Journal of Human-Computer Studies*, 2019 (122): 145–167.

[66] Han Xiao, Zhang Chunhong, Guo Chenchen, Ji Yang, Zheng Hu. Distributed Representation of Knowledge Graphs with Subgraph-aware Proximity [J]. *Theoretical Computer Science*, 2020 (803): 48–56.

[67] Xiao Han, Huang Minlie, Zhu Xiaoyan. TransG: A Generative Model for Knowledge Graph Embedding [J]. *Computer Science*, 2016 (1): 2316-2325.

[68] Zhang Yi, Cao Wanhua, Liu Juntao. A Novel Negative Sample Generating Method for Knowledge Graph Embedding [C]// 2019 International Conference on Embedded Wireless Systems and Networks, 2019: 401–406.

[69] Kohei I, Yukihisa K, Shinsuke K, Atsushi S. Structuring Engineers' Implicit Knowledge of Forming Process Design by Using a Graph Model [J]. *Procedia CIRP*, 2018 (67): 563-568.

[70] Wang Huiyong, Lun Bing, Zhang Xiaoming, Sun Xiaoling. Multimodal Entity Alignment Based on Joint Knowledge Representation Learning [J]. *Control and Decision*, 2019: 1-10.

[71] Gayathri R, V Uma. Ontology Based Knowledge Representation Technique, Domain Modeling Languages and Planners for Robotic Path Planning: A Survey [J]. *ICT Express*, 2018 (4): 69–74.

[72] Wu Zhenyong, Liao Jihua, Song Wenyan, Mao Hanling, Huang Zhenfeng, Li Xinxin, Mao Hanying. Semantic Hyper-graph-based Knowledge Representation Architecture for Complex Product Development [J]. *Computers in Industry*, 2018 (100): 43–56.

[73] Guo Shu, Wang Quan, Wang Bin, Wang Lihong. SSE: Semantically Smooth Embedding for Knowledge Graphs [J]. *IEEE Transactions on Knowledge and Data Engineering*, 2017 (29): 884-897.

Note: Figure translations are in progress. See original paper for figures.