IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 28 May 2022, accepted 12 July 2022, date of publication 18 July 2022, date of current version 22 July 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3191784

SURVEY

A Comprehensive Survey of Graph Neural Networks for Knowledge Graphs

ZI YE[1,2], YOGAN JAYA KUMAR[2], GOH ONG SING[2], FENGYAN SONG[3], AND JUNSONG WANG[4]

[1]School of Artificial Intelligence, Wenzhou Polytechnic, Wenzhou 325000, China
[2]Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka, Melaka 76100, Malaysia
[3]Shanghai Gen Cong Information Technology Company Ltd., Shanghai 200000, China
[4]College of Big Data and Internet, Shenzhen Technology University, Shenzhen 518000, China

Corresponding author: Junsong Wang (wjsong2004@126.com)

This work was supported in part by the Basic Research Project of Wenzhou, China, under Grant G2020019; and in part by the National Natural Science Foundation of China under Grant 61473208 and Grant 61876132.

ABSTRACT The Knowledge graph, a multi-relational graph that represents rich factual information among entities of diverse classifications, has gradually become one of the critical tools for knowledge management. However, the existing knowledge graph still has some problems which form hot research topics in recent years. Numerous methods have been proposed based on various representation techniques. Graph Neural Network, a framework that uses deep learning to process graph-structured data directly, has significantly advanced the state-of-the-art in the past few years. This study firstly is aimed at providing a broad, complete as well as comprehensive overview of GNN-based technologies for solving four different KG tasks, including link prediction, knowledge graph alignment, knowledge graph reasoning, and node classification. Further, we also investigated the related artificial intelligence applications of knowledge graphs based on advanced GNN methods, such as recommender systems, question answering, and drug-drug interaction. This review will provide new insights for further study of KG and GNN.

INDEX TERMS Deep learning, distributed embedding, graph neural network, knowledge graph, representation learning.

I. INTRODUCTION

Recently, the fast development of Internet Technology and Web applications has contributed to an explosion of a variety of data on the Internet, which can generate a large amount of valuable knowledge. Therefore, how to organize, represent, and analyze this knowledge has attracted much attention [1]. As a result, the knowledge graph (KG) was created to organize this information semantically and visually appealing [2]. At present, many knowledge graphs have emerged, among which the representative ones are KnowItAll [3], YAGO [4], DBpedia [5], Freebase [6], NELL [7], Probase [8], etc. These knowledge graphs extract, organize, and manage knowledge from enormous data resources to provide users with intelligent services.

The associate editor coordinating the review of this manuscript and approving it for publication was Shadi Alawneh.

However, there are still many technical difficulties during the knowledge graph development. For example, KGs are always incomplete since they have self-defects, which is usually formulated as the link-prediction issue [9]. For the research of the prime technologies such as link prediction [9], knowledge graph alignment [10], knowledge graph reasoning [11], Knowledge Representation Learning (KRL) is the critical foundation, also known as Knowledge Graph Embedding (KGE) [12]. KRL is the process of completing distributed representation of entities and relationships in knowledge graphs. Mapping entities and relationships to low-dimensional vector spaces indirectly captures their semantics [13]. For example, single knowledge graph embedding can be used for link prediction evaluation. Due to the heterogeneity problem between different knowledge graphs, multiple knowledge graphs embedding can be used for entity alignment [14]. Compared with traditional one-hot encoding,

KRL can significantly improve computational efficiency. At the same time, it can ease the issue of data sparseness and achieve the goal of combining heterogeneous information.

Graph Neural Network (GNN) appears as a framework that has used deep learning to learn graph-structured data directly in recent years. The essence of GNN is to gather information from the neighborhood to the target node according to the message passing rules so that entities with similar neighborhoods are close to each other in the embedding space, and it excels in capturing the global or local structural information of the graph [15]. On the other hand, because GNN has a nonlinear solid fitting ability to graph-structure data, it has higher accuracy and better robustness on problems in different fields [16]. A great deal of variants of the GNN algorithm and framework have been proposed in the past few years. Furthermore, these GNN-based KGE models can combine domain data in the knowledge graphs with business scenarios and help domain business upgrade. Typical downstream applications include recommender systems [17], intelligent question answering [18], and drug-drug interactions [19].

Previous survey papers solely focus on general knowledge graphs problems or graph neural networks technology. This study performs a systematic and broad survey of knowledge graph learning in accordance with graph neural network methods. The framework of this article is shown in Fig. 1, and the contributions of this study are presented as follows:

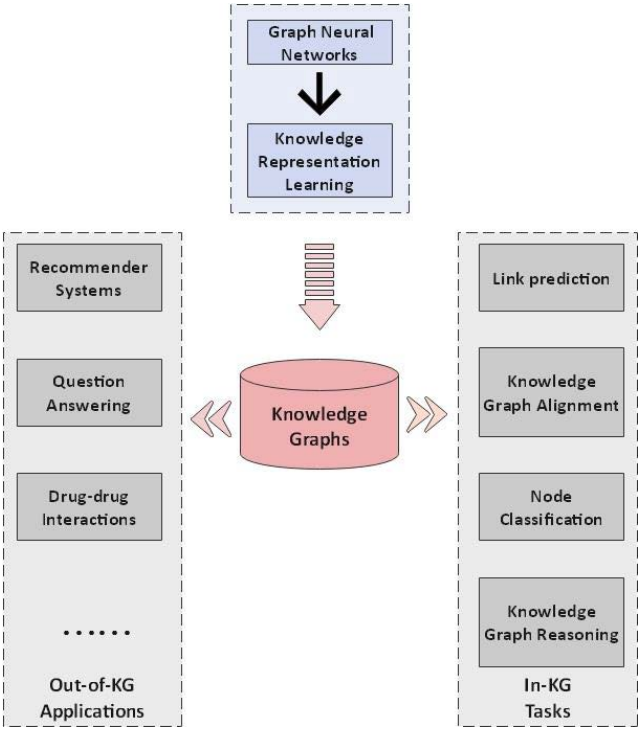

FIGURE 1. The framework of the contents of this study.

–This is the first comprehensive survey paper of graph neural network models for knowledge graph problems to the best of our knowledge. This work covers GNN techniques that solve KG-related tasks. In addition, it offers discussions and comparisons of the introduced methods.

–We explored almost all the state-of-the-art graph learning methods based on graph neural networks for four typical tasks in the knowledge graph. We presented their critical technologies and characteristics from a variety of perspectives.

–Considering the vast application foreground of knowledge graphs, we have further conducted a thorough investigation of downstream tasks, covering recommender systems, question answering, drug-drug interaction, etc.

The following are the main points of the remainder of this study: The definitions of knowledge graphs and graph neural networks are briefly introduced in Section II. Then, Section III shows the approaches proposed for learning knowledge graph representations using GNN technology. The implementation of such information in downstream activities, such as question-answer systems, is next explored in Section IV. Finally, the conclusion of this investigation is shown in Section V.

II. PRELIMINARY KNOWLEDGE

This section first introduces the basic concepts of knowledge graphs and graph neural networks.

A. THE CONCEPT OF KNOWLEDGE GRAPH

A knowledge graph refers to a semantic network graph which is consisted of diverse entities, concepts, and relationships in the real world. It is used to formally describe various things and their associations in the real world [20].

Knowledge graphs are generally represented in triples $G=\{E, R, F\}$. Among them, E represents the entity set $\{e_1, e_2, \cdots, e_E\}$, and the entity e is the most basic element in the knowledge graph, referring to the items that exist objectively and can be distinguished from each other. R represents the relation set $\{r_1, r_2, \cdots, r_R\}$, and the relation r is an edge in the knowledge graph, representing a specific connection between different entities. F represents the fact set $\{f_1, f_2, \cdots, f_F\}$, and each f is defined as a triple $(h, r, t) \in f$, in which h denotes the head entity, r stands for the relationship, and t indicates the tail entity.

B. THE CONCEPT OF GRAPH NEURAL NETWORK

The graph neural network adopts a particular method for describing the nodes. Based on continuous node state updates, it can obtain a state including both neighbor node information and graph topology characteristics [21]. Therefore, the goal is to learn an embedded state $h_v \in R^s$ for each node that encodes its neighbor information, and this state h_v is used to generate the required output o_v, as expressed in

$$H = F(H, X), \tag{1}$$

$$O = G(H, X_N). \tag{2}$$

where H represents the state of all nodes; O represents the result after outputting all nodes; X represents the edge feature; X_N represents the feature of all nodes; $F(\cdot)$, $G(\cdot)$

respectively represent the global transformation function and the global output function.

It can be seen that when the state of all nodes is updated from t to $t+1$, it can be expressed as

$$H^{t+1} = F\left(H^t, X\right). \tag{3}$$

III. THE KNOWLEDGE GRAPH LEARNING BASED ON GRAPH NEURAL NETWORK

Since knowledge graphs can represent structured relationships between entities, they have become an important research area in cognition and artificial intelligence. The graph neural network employs the deep neural network to integrate the topology information and attribute feature information in the graph data, and then provides a more refined feature representation of nodes and can be easily combined with downstream tasks through an end-to-end manner. It cleverly meets the requirements of a knowledge graph for learning the attribute features and structural features of entities and relationships. This section mainly introduces the knowledge graph learning method using graph neural networks for four typical tasks.

A. LINK PREDICTION

Link prediction seeks to forecast missing information (links or relations) between the elements in knowledge graphs. The overview of recent GNN-based link prediction models on the knowledge graph is presented in Table 1. In fact, KGs can be represented as multi-relational directed graphs with nodes and edges representing entities and relations, respectively [22]. Therefore, most researchers are motivated to stress the essential effects of various connection patterns between entities, which result in the sufficient capture of the relationships. The models that belong to the Relation-Aware GNN category are shown in Fig. 2.

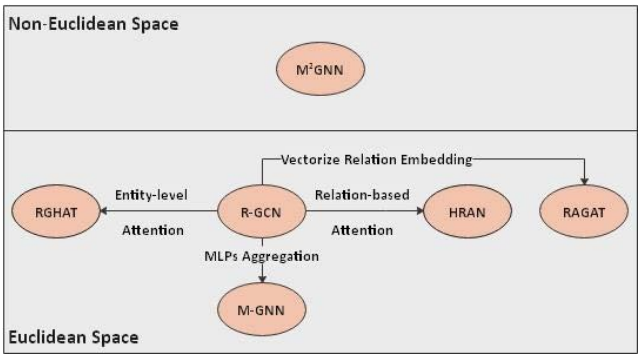

FIGURE 2. The relation-aware GNN models.

To deal with multiple relationships in a GCN base, one of the earliest methods was R-GCN [32], also known as a Relational Graph Convolutional Network. R-GCN and GCN are distinguished by the fact that in R-GCN, edges can represent a variety of different relationships. All edges in GCN can share weights. R-GCN, on the other hand, employs distinct weights for different edge types, and only edges of the same relation type r are associated with the same projection weight. A model like the R-GCN produces exponential parameter expansion, however, especially when dealing with data that is extremely multi-relational. The original study advocates the use of basis decomposition to reduce the amount of model parameters and prevent overfitting. All relationships in the decoder are scored by using the DistMult factorization model, and each relationship is linked to the diagonal matrix. The heterogeneous relation attention network (HRAN) [36] framework also follows a hierarchical structure containing both entity-level and relation-level aggregation for KG embedding. But different from R-GCN,

TABLE 1. Representative models for link prediction.

Category	Models	Relation-Aware GNN	Zero-Shot Learning	Datasets
Link Prediction	RGHAT [23]	√	×	FB15k, WN18, FB15k-237, WN18RR
	Convolutional Transition and Attention-Based Aggregation Graph Neural Network [24]	√	√	WordNet11, NELL-995
	ExpressGNN [25]	×	√	UW-CSE, Cora, synthetic Kinship, FB15K-237
	LSA-GAT [26]	×	×	FB15k-237, WN18RR
	Gravity-Inspired Model [27]	×	×	Cora, Citeseer, Google
	INDIGO [28]	×	√	FB15K-237, NELL-995, WN18RR, WordNet11
	GraIL [29]	√	√	WN18RR, FB15k-237, NELL-995
	SEAL [30]	×	×	USAir, NS, PB, Yeast, C.ele, Power, Router, E.coli
	M^2GNN [31]	√	×	WN18RR, FB15k-237, YAGO3-10
	R-GCNs [32]	√	×	FB15k, WN18
	RAGAT [33]	√	×	FB15k-237, WN18RR
	M-GNN [34]	√	×	WN18, FB15k
	AR-KGAT [35]	×	×	FB15k-237, WN18RR
	HRAN [36]	√	×	WN18, FB15k-237, WN18RR
	LightCAKE [37]	×	×	WN18RR, FB15K-237, NELL995, DDB14

HRAN proposed a novel relation-based attention mechanism with the purpose of obtaining the significance of different relation paths.

Unfortunately, the previous R-GCN model faces three shortcomings, i.e., low representational power, stacking flatly, and poor robustness to noise. A novel multi-level graph neural network (M-GNN) [34] was proposed to deal with the aforementioned challenges. Unlike R-GCN, which utilize mean pooling as their aggregator, the M-GNN introduced multi-layer perception (MLPs) when conducting neighborhood aggregation to improve the representational power of GNN layers. The updating process in M-GNN is as follows: $h_v^{(k)} = MLP^{(k)}((1+\epsilon^{(k)})\cdot h_{r_0,v}^{k-1} + \sum_{\mu \epsilon N_v^r} h_{r,u}^{k-1})$, where N_v^r indicates the set of neighbor indices of the node i under relation $r \epsilon \mathcal{R}$, $h_{r,u}^{k-1}$ denotes the message passing from each neighbor node u under relation r at the $k-1^{th}$ layer, and $h_{r_0,v}^{k}$ denotes the self-connection message. On this basis, the graph coarsening scheme was developed as a strategy to output a series of graphs with different granularities. Then, multiple GNNs on these graphs were stacked to model multi-level structures of the original graph.

Another disadvantage of the R-GCN model is that it has no vectorized relation embedding involved, which can thus limit the model's expansibility. Inspired by the graph attention network, a new graph neural network called Relation Aware Graph Attention network (RAGAT) [33] was proposed. The core idea behind RAGHAT is introducing relation specific network parameters θ_r to study information from neighboring entities under different relations adaptively. The new message function is defined as

$$C_{(u,r,v)}^{r} = \phi_r \left(e_u, e_r, e_v, \theta_r\right). \tag{4}$$

where ϕ_r is combining operator, e_u and e_v denotes entity embedding, e_r is relation embedding.

Three of the link prediction models discussed above are based on relation-level attention. Inspired that not all neighboring entities in a specific relation are equally crucial in indicating the central entity, a novel neighborhood-ware model named Relational Graph neural network with Hierarchical ATtention (RGHAT) [23] was proposed. In addition to computing the weights for different relations, which is the first-level attention, RGHAT is further equipped with entity-level attention. RGHAT highlights the importance of different neighboring entities under the same relation and hierarchically aggregates information in this second level.

Mixed-Curvature Multi-Relational Graph Neural Networks (M^2GNNs) [31] were created by Wang *et al.* to solve the issue that embedding in single-curvature space overlooks the innate heterogeneity structures inherent in rich-structured KGs. To be more exact, a tractable Riemannian product manifold that combines Euclidean, spherical, and hyperbolic spaces can be used to generate the mixed curvature space. Graph Neural Updater is presented to aggregate and update the embedding features of entities and relations to increase the quality of the embedding.

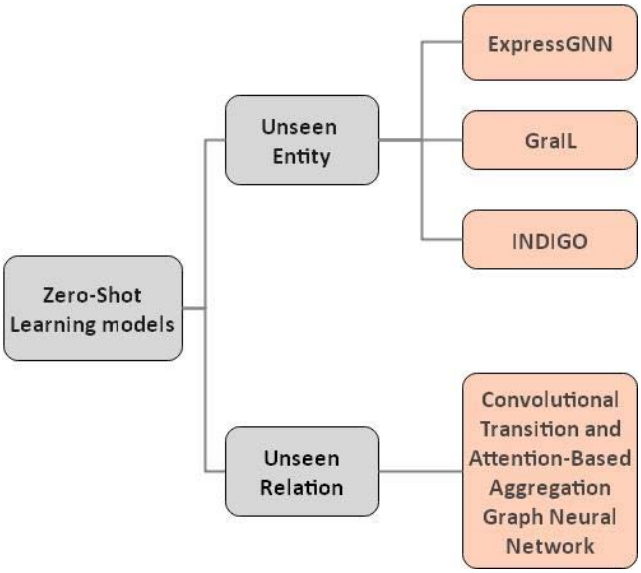

FIGURE 3. The zero-shot learning models.

One of the important issues for link prediction refers to the poor scalability of KGs since numerous real-world KGs are ever-evolving, with new nodes or relations being supplemented over time (also called zero-shot scenarios) [38]. The models that concern Zero-Shot Learning is displayed in Fig. 3. Making predictions on such new entities without expensive re-training is of great necessity for production-ready machine learning models. Markov Logic Networks (MLNs) have emerged as practical tools to deal with the zero-shot learning problem because they don't require many labeled examples for a target task. However, inference in MLN is computationally intensive. In ExpressGNN [25], the advantages of MLNs in logic reasoning and GNN in graph representation learning are both leveraged. It was finally implemented in a constructed zero-shot learning dataset based on FB15K-237 and achieved more capability of performing concerning low amounts of observed data.

Another dominant line to predict the relation in the unseen KGs is to view it as a logical induction problem. One seeks to derive probabilistic logical rules underlying a given KG. GraIL [29] is one of the representative GNN-based frameworks for inductive knowledge graph reasoning. In this approach, the relation between two nodes is predicted from the enclosing sub-graph structure extracted around those two nodes. No node attribute is assumed when labeling the nodes to test the model's ability to generalize solely from the structure. The final step in the current framework aims to adopt a multi-relational GNN for scoring the likelihood of a triplet (u, r_t, v), in which we indicate nodes u and v as target nodes and r_t as the target relation. Besides, the drawback of GraIL is that it relies on heuristics and scoring function so that the inductive capabilities of the GNN are not fully exploited. In contrast to GraIL, INDIGO [28] encodes KGs using a one-to-one correspondence between

triples in the KG and elements of the feature vectors in the innermost and outermost layers of the GNN. Hence, the predicted triples can be read out directly from the last layer of the GNN without the requirement for additional components.

The previously described models can only handle the Out-of-knowledge-graph (OOKG) entity problem. Moreover, they cannot handle the issue of unobserved relations, which is a new zero-shot scenario put forward currently. Convolutional Transition and Attention-Based Aggregation Graph Neural Network [24] was regarded as a novel method that can be used to generate embedding vectors of OOKG relations. This approach novelly sets a convolutional transition function with the purpose of transferring information for OOKG entities and relations in parallel. When computing the embedding of relations, for example, the propagation model for relations in the neighborhood is used. For the purpose of determining the weight value of each information vector, the author also proposed an attention-based aggregation network.

According to the literature review, there are other perspectives on the link prediction problem for knowledge graphs. The SEAL [30] framework develops a new γ-decaying heuristic theory and translates link prediction into a subgraph classification problem in a comprehensive manner. Initially, the SEAL can extract its h-hop enclosing subgraph A for each target link and constructs its node information matrix X, including structural node labels, latent embeddings, and explicit attributes of nodes. Subsequently, SEAL may send (A, X) into a GNN to categorize the existence of the link. By taking unique local structures like cycles and stars into account, a novel graph attention network named LSA-GAT [26] derives a sophisticated representation covering both the semantic and structural information. Lightweight Framework for Context-Aware Knowledge Graph Embedding (LightCAKE) [37] focuses on graph context. The novel aspect of this technique is the construction of a context star network to model the entity/relation context. Following that, each entity/relation node in the newly-framed context star graph combines information from its surrounding context nodes using a scoring algorithm to determine weights. The association rules enhanced knowledge graph attention network (AR-KGAT) [35] aggregates neighborhood information with both association-rules-based and graph-based attention weights. Since a knowledge graph is a directed labeled graph in which the labels have well-defined meanings, the Gravity-Inspired Model [27] is proposed as a new gravity-inspired decoder scheme for the link prediction in directed graphs. Inspired by Newton's theory of universal gravity, this framework learns node embedding from directed graphs, using graph AE and VAE frameworks. Besides, Newton's equations in the resulting embedding are applied, the acceleration $a_{i\to j} = Gm_j/r^2$ of a node i towards a node j is applied because of gravity in the embedding to represent the likelihood that i associated with j in the directed graph.

TABLE 2. Representative models for knowledge graph alignment.

Category	Sub-Categories	Models	Datasets
Knowledge Graph Alignment	Static KG Alignment	CG-MuAlign [40]	Wikipedia, Amazon-Music, Freebase, IMDB
		GrAN [41]	NRW, MUC
		AliNet [42]	DBP15K, DWY100K
		EPEA [43]	DBP15K, DBpedia
		MRAEA [44]	DBP-15K, WK31-60K, SRP_{Normal}^2
		MuGNN [45]	DBP15K, DWY100K
		DANs [46]	DBP15K
	Dynamic KG Alignment	DINGAL [47]	DBP15K

B. KNOWLEDGE GRAPH ALIGNMENT

In the field of knowledge fusion, Entity Alignment, also known as Entity Matching or Entity Resolution, is a critical and foundational technology. Knowledge graphs that depict the same real-world entity can be identified by aligning their entities [39]. In this study, we consider two groups of knowledge graphs, static KGs and dynamic KGs. Table 2 illustrates the representative publications of GNN-based knowledge graph alignment.

We first introduce different embedding models for static KGs. Since the seed alignments are usually insufficient for high-quality entity embedding, most efforts failed to consider structure heterogeneity between different KGs. Alignment-oriented knowledge graph (KG) embeddings can be learned using MuGNN [45], a Multi-channel Graph Neural Network model that robustly encodes two KGs using multiple channels. Each channel encodes KGs using various relation weighting strategies for self-attention toward KG completion and cross-KG attention for trimming exclusive entities independently, both of which are thoroughly integrated using pooling approaches. It can also consistently conclude and transfer rule knowledge for the completion of two KGs. MuGNN may be able to resolve the structural differences between two KGs and better utilize seed alignment data. While MuGNN finds the structure incompleteness of KGs and aims at the rule-based KG completion, the counterpart entities have non-isomorphic neighborhood structures unavoidably. As a result, a new KG alignment network, called AliNet [42], was proposed to mitigate the different neighborhood structures end-to-end. AliNet introduces distant neighbors to enlarge the overlap between their neighborhood structures using an attention mechanism and restricts the equivalent entity pairs' two entities to have the same hidden state in each GAT layer. A relation loss is finally used to refine entity representations. The entity-pair embedding approach (EPEA) [43] for KG alignment introduced the pairwise connectivity graph (PCG) of KGs, whose nodes are entity-pairs and edges are in line with relation-pairs. This approach tries to encode attribute features from entity-pairs

using a CNN and then enhance the features propagation among the neighbors of entity-pairs through GNN with edge-aware attention.

Many attention mechanism-based models for KG representation learning have recently obtained advanced performance in entity alignment tasks. For example, decentralized attention networks (DANs) [46] compute the attention score by purely relying on the target entity's neighbors and ignoring the requirement of its embedding throughout. Such characteristic enables DAN to induce embeddings on unseen entities. Since there are different entity types and relation types in a knowledge graph. Considering the nature of entity and relation types, it makes sense to have a different alignment strategy for the different entity types. Therefore, a multi-type entity alignment algorithm named CG-MuAlign [40] was created to collectively align entities of different types and make predictions on unseen entities using attention mechanisms and neighborhood information. Furthermore, it used relation-aware neighborhood sampling to enhance the computational efficiency of the approach in large-scale data collection.

Many other applicable circumstances can be regarded as problems of graph alignment. For example, Map Fusion (MF) aims to identify nodes from both road networks that match each other [48]. Still, the current GNN approaches show poor performance for the MF task since the information from non-overlapping areas negatively affects the learned node representations. The Graph Alignment Network (GrAN) [41] aggregates information from neighbors by emphasizing nodes that have a good match in the counterpart graph, leading to an inductive bias that neighboring nodes that are likely in the overlapping area are more useful for the target node representation. Cross-lingual entity alignment associates semantically similar entities in knowledge graphs with different languages. However, current approaches fail to model the meta semantics or complex relations such as n-to-n and multi-graphs. A new method, Meta Relation Aware Entity Alignment (MRAEA) [44], operates on cross-lingual KGs through leveraging meta relation-aware embedding and relation-aware self-attention. In addition, this work further adopted an effective iterative training strategy on the basis of the asymmetric nature of alignments.

The prior discussion is based on static KGs, i.e., the knowledge is represented at a specific point in time [49]. However, because of initial incompleteness, almost every KG has some evolution in practice. Moreover, this evolution is smooth with small changes rather than drastic modifications of large subgraphs in most cases. Therefore, to efficiently tackle the issue of updating entity embeddings for the evolving graph topology, a family of algorithms (DINGAL) [47] using graph convolutional networks is proposed. The key idea is to distance the coupling between the parameter matrix in GCN and the underlying graph topology. This work is believed to be the first to study the dynamic knowledge graph alignment problem, and more reaches will likely be made on dynamic knowledge graphs in the future.

TABLE 3. Representative models for knowledge graph reasoning.

Category	Models	Datasets
Knowledge Graph Reasoning	Deep-IDA* [51]	NELL-995
	DPMPN [52]	FB15K, FB15K-237, WN18, WN18RR, NELL995, YAGO3-10
	CS-GNN-N [53]	DBP15K-num, WN18RR-num, FB15k-237-num, Mil-assistant
	TRAR [54]	FB15K-237, WN18RR, Kinship
	T-GraphR [55]	LegalLPP

C. KNOWLEDGE GRAPH REASONING

The knowledge reasoning for knowledge graphs studied in this paper refers to using specific methods to infer new conclusions or identify wrong information based on existing data. For example, a fact like triple $(X, BirthPlace, Y)$ is given in DPpedia, the missing triple $(X, Nationality, Y)$ can be obtained through reasoning [50]. To some extent, knowledge reasoning is similar to link prediction, while the relationships acquired by knowledge reasoning mostly need a multi-hop reasoning process in the knowledge graph. Table 3 illustrates the representative publications of GNN-based knowledge graph reasoning.

Graph neural networks learn the feature representation automatically and provide structured explanations suitable for knowledge reasoning. The target relational attention-oriented reasoning (TRAR) [54] model proposes a novel embedding-based approach to aggregate the information by designing node-level and relational subgraph-level attention mechanisms. In addition, the multiple target relational attention-oriented layers concentrate more on the relations that match the target relation, whereas each subgraph uses a hierarchical attention mechanism for obtaining node-level information. The Dynamic Pruned Message Passing Networks (DPMPN) [52] also model the reasoning process for large-scale knowledge graphs by constructing local subgraphs dynamically. This approach developed a two-GNN framework to learn the reasoning by explaining through some understandable form. The cascaded attention mechanism makes explanation efficient by selecting relevant nodes to construct subgraphs regardless of how large the underlying graph is. Even though the above-mentioned embedding methods have successfully acquired promising findings in specific KG reasoning tasks, they fall short for modeling multi-hop relational paths in more complicated reasoning tasks. The Deep-IDA* [51] framework empowers embedding-based methods by combining path-based algorithms. It is the first to integrate the traditional path searching algorithms and deep neural networks for KG reasoning.

Furthermore, the knowledge graphs reasoning ability can be widely applied in many downstream tasks, and the illustration is shown in Fig. 4. Developing a reasoning method is one of the specific directions of knowledge graph in the military field. It is a critical technology that can stimulate the intellectual development of military combat command. A knowledge reasoning method was proposed for military decisions by mixing rule learning, rule injection, and graph

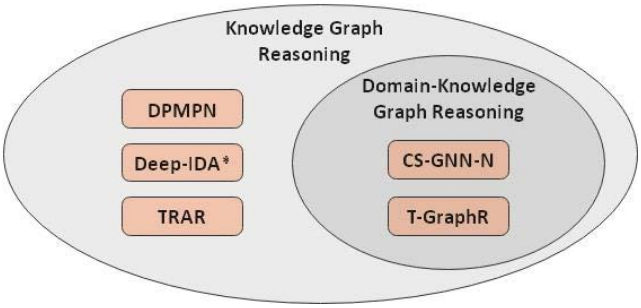

FIGURE 4. The models on knowledge graph reasoning.

TABLE 4. Representative models for node classification.

Category	Models	Datasets
Node Classification	EvolveGCN [56]	Elliptic
	HDGI [57]	DBLP, ACM, IMDB
	H-GCN [58]	NELL

neural network learning together, named context-surrounding graph neural networks with numbers (CS-GNN-N) [53]. With the advancement of technology in recent years, legal artificial intelligence has flourished. The Text-guided Graph Reasoning (T-GraphR) [55] refers to a novel approach for the Legal provision prediction (LPP) task that seeks to predict the associated legal provisions of affairs. This approach briefly consists of two main components: text representation learning and legal graph reasoning. We believe that more knowledge reasoning methods will be explored and attempted on more domain KGs in the future.

D. NODE CLASSIFICATION

Node classification, in which an attribute of each node in a graph is predicted, is another of the most popular and commonly adopted tasks on graph data—for example, assigning a categorical class to each node (binary or multiclass classification) or forecasting a continuous number (regression). Table 4 illustrates the representative publications of GNN-based knowledge graph node classification.

To solve the problem of the limited receptive field caused by the lack of the ''graph pooling'' mechanism, a novel deep Hierarchical Graph Convolutional Network (H-GCN) [58] is presented for semi-supervised node classification. Coarsening layers and symmetric refining layers make up the H-GCN model. This approach can achieve a more significant receptive field and good information propagation by clustering structurally related nodes into hyper-nodes.

Still, labeled data is not always available, and the training sample scarcity problem has aroused extensive research interest. The Heterogeneous Deep Graph Infomax (HDGI) [57] attempted to learn high-level representations containing graph-level structural information without any supervised label by maximizing local-global mutual information. HDGI also combines the meta-path technique to represent the composite relations with distinct semantics in heterogeneous graph studies.

In real-world circumstances, we are frequently confronted with continually developing graphs, so the question can be formulated as to how the graph neural network handles such a dynamism. A new method called evolving graph convolutional network (EvolveGCN) [56] was developed to capture the dynamism of graphs, based on the recent success of graph convolutional network (GCN) for static networks. Instead, the RNN governs the GCN model at each step in this method. As a result, it effectively executes model adaptation, which focuses on the model rather than the node embeddings, leading to the non-restriction of the node appearance. It's worth noting that the GCN parameters aren't trained in this manner. Only the RNN parameters are trained because they are computed from the recurrent network.

IV. THE KNOWLEDGE GRAPH APPLICATION BASED ON GRAPH NEURAL NETWORK

At present, the higher-quality knowledge graph obtained after the graph neural network representation learning can be used for multiple applicable downstream tasks combined with specific scenarios. Recommender Systems, Question Answering, and Drug-drug Interaction are explained in detail.

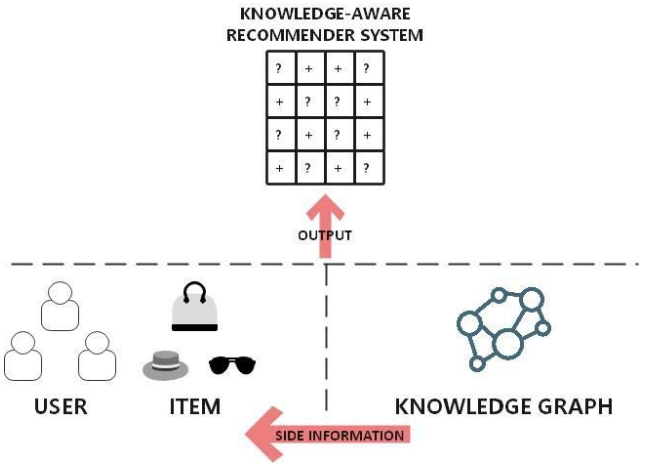

FIGURE 5. Illustration of user-item interactions and the knowledge graph.

A. RECOMMENDER SYSTEM

Recommender Systems (RS), as one of the most famous and significant uses of Artificial Intelligence (AI), have been widely adopted to assist consumers in making appropriate choices among the enormous amount of products and services available. However, when the data has cold-start problems, depending merely on user-item interactions spoils recommendation performance. As a result, existing research suggests using knowledge graphs (KGs) as side information to investigate implicit or high-order connectivity relations between users or items to improve their representations and thus improve recommendation effectiveness, as shown in Fig. 5. The GNN technique addresses two critical downstream tasks: explicit feedback and implicit feedback. For example, Ranking prediction models employ explicit

feedback to deliver a customized ranked list of recommended products to the user. On the other hand, Click-through rate (CTR) prediction leverages implicit feedback to forecast the likelihood of people clicking adverts or objects. As a result, the knowledge-aware recommendation task can be expressed as

$$\hat{y}_{uv} = \mathcal{F}(u, v|\ominus, \mathcal{G}_Y, \mathcal{G}_K). \tag{5}$$

where $\hat{y}_{uv}$ refers to the prediction of user's interest in item v, $\mathcal{F}$ indicates the learned prediction function with weights $\ominus$, $\mathcal{G}_Y$ is a user-item bipartite graph, and $\mathcal{G}_K$ is the knowledge graph.

Generally, the workflow of a GNN-based Knowledge Aware Deep Recommender System initially studies to yield an embedding by graph embedding module for every graph node, containing the user and item nodes, encoding the information distilled from the input knowledge. Finally, the model parameters can be updated by dealing with the objective function [68]. There are two ways for applying graph feature learning to recommender systems: One-by-one Learning (OL) and Joint Learning (JL). In terms of one-by-one learning, the knowledge graph characteristics are initially adopted for obtaining entity vectors and relationship vectors. Subsequently, these low-dimensional vectors are presented into the recommender system in order to study the target user vectors and item vectors. Additionally, joint learning can combine the knowledge graph feature learning with the objective role of the recommender system, which is an end-to-end method. Table 5 illustrates the representative publications of the GNN-based knowledge-aware recommender system.

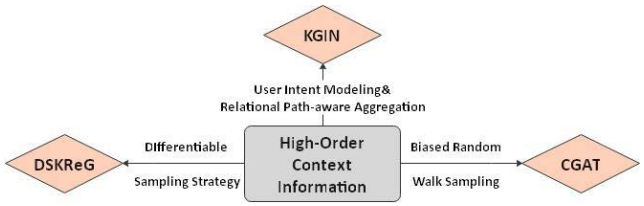

FIGURE 6. The models focusing on capturing high-order context information.

In practice, one group of approaches automatically captures both structure and semantic information of KGs, such as the most related high-order neighbors. The overall view of the models concerned is displayed in Fig. 6. Contextualized Graph Attention Network (CGAT) [59] was proposed by Yang *et al.* to take advantage of KG entities' local and non-local graph context information. A user-specific graph attention technique is used to aggregate the relation-ware neighborhood information of an item in order to consider local context information (one-hop). Based on a biased random walk, essential entities for the target entity are extracted from the entire KG to incorporate the non-local context in KG. Using a GRU module, this embedding can be explicitly aggregated.

TABLE 5. Representative models of GNN-based knowledge-aware recommender system.

Category	Tasks	Models	Datasets	Methods
Recommender System	Explicit Feedback	CGAT [59]	Last-FM, Movielens-1M, Book-Crossing	JL
		DSKReG [60]	Last-FM, BookCrossing, MovieLens-Sub	OL
		DKMN [61]	Yelp open dataset	OL
		JNSKR [62]	Amazon-book, Yelp2018	JL
		KGIN [63]	Amazon-Book, Last-FM, Alibaba-iFashion	JL
		SARC [64]	Movielens-1M, Book-Crossing, Last.fm-2K	JL
	Implicit Feedback	Fi-GNN [65]	Criteo, Avazu	OL
		GraphSW [66]	Last.FM2011, Book-Crossing, movie, LFM-1b 2015, Amazon-book, Yelp 2018	JL
		KGNN-LS [67]	MovieLens-20M, Book-Crossing, Last.FM, Dianping-Food	JL

Note that OL stands for One-by-One learning, and JL stands for Joint Learning.

However, the exponential increase of a node's receptive field places a severe constraint on high-order aggregation. Differentiable Sampling on Knowledge Graph for Recommendation with Relational GNN (DSKReG) [60] proposed learning the relevance distribution of related items from knowledge graphs and sampling relevant items in accordance with this distribution. With this model's addition of a differentiable sampling method, the selection of appropriate objects can be optimized as the model is being trained.

The Knowledge Graph-based Intent Network, or KGIN [63], is primarily concerned with analyzing the user intents that lie beneath user-item interactions. It does this by utilizing item KG to improve the performance of recommendations and the interpretability of those recommendations. This approach represents each intent as an attentive combination of KG relations and introduces a new relational path-aware aggregation scheme to aggregate long-range connectivity relation sequences and refine multi-hop path representations.

The heterogeneous knowledge graph is a new hot pot because real-world applications, such as social networks display beneficial and essential information on node connections. Some existing models proposed attempt to handle

relations in heterogeneous graphs with different semantics. The Knowledge-aware Graph Neural Networks with Label Smoothness regularization (KGNN-LS) [67] calculates user-specific item embeddings by initially employing a trainable function to identify meaningful knowledge graph relationships for a given user. For example, a given user is more concerned with the ''director'' relationship between movies and people than the ''lead actor'' relationship. Furthermore, the proposed label smoothness constraint and leave-one-out loss offer strong regularization for studying the edge weights in KGs. The Dual Knowledge Multimodal Network (DKMN) [61] is an extension work of KGNN-LS. This work aims to work on knowledge graphs and other multimodal data, e.g., textual reviews and pictures.

Unlike previous studies, which have primarily focused on exploring novel neural networks, some researchers consider reducing the massive computational cost while maintaining the pattern of extracting features. The GraphSW [66] technique is based on a stage-wise training framework that only examines a subset of KG entities at each stage. In the succeeding steps, the network receives the learned embedding from the previous stages, and the model can gradually learn the information from the KG. It has been discovered that the existing non-sampling strategy computes the gradient over the entire data set, resulting in high computational costs. A novel Jointly Non-Sampling learning model for Knowledge graph enhanced Recommendation (JNSKR) [62] first designed a new efficient non-sampling loss for knowledge graph embedding learning, significantly reducing complexity. The surrounding entities of an item are then aggregated with attention mechanisms to help learn accurate user preferences over items.

Also, there are some other methods for studying GNN-based recommender systems. The Feature interaction Graph Neural Networks (Fi-GNN) [65] identified a limitation in modeling sophisticated interactions using simple unstructured combinations. Hence, it intends to consider the structure of multi-field features. The Fi-GNN adopted a graph structure representing multi-field features called a feature graph. Each node in the graph intuitively is consistent with a feature field, and diverse fields can make the interaction through edges. As a result, modeling sophisticated interactions between feature fields can be decreased to modeling node interactions on the feature graph. The novel Split-And-ReCombine strategy (SARC) [64] separates the user-item-entity interactions into three two-way interactions: the ''user-item'', ''user-entity'', and ''item-entity'' interactions. Besides, the two-way interactions can be represented as a graph, thus be modeled using Graph Neural Networks (GNN) and knowledge graph embeddings. In the second stage, SARC uses the representation of users and items learned in the first step to make a suggestion.

B. KNOWLEDGE BASE QUESTION ANSWERING

Knowledge base question answering (KBQA) aims to respond to a question using information from a knowledge base (KB). In recent years, academics have focused chiefly on GNN-based solutions for addressing the difficulties of answering complicated questions, as shown in Table 6 for relevant models.

TABLE 6. Representative models for knowledge base question answering.

Category	Models	Datasets
Knowledge Base Question Answering	Relational GNN for Open-domain Question Answering [69]	WebQuestionsSP
	QA-GNN [70]	CommonsenseQA, OpenBookQA, MedQA-USMLE
	DELFT [71]	QbLink, QANTA, TriviaQA

TABLE 7. Representative models for drug-drug interactions.

Categories	Models	Datasets
Drug-Drug Interactions	GENN [72]	DeepDDI, BIOSNAP-sub
	KGNN [73]	DrugBank, KEGG-drug

Subgraph reasoning is a common method for obtaining the answer. QA-GNN [70] is an end-to-end question answering model in which the QA context as an additional node connects the topic entities to form the subgraph of KG. It is then proposed to use pre-trained language models to score KG nodes' relevance to the QA context based on their importance. Finally, the joint graph representation is updated through graph-based message passing. Deciphering Entity Links from Free Text (DELFT) [71] produces a dense and high-coverage semantic subgraph by linking question entity nodes to candidate entity nodes using text sentences from Wikipedia. This innovative graph neural network performs better on entity-rich questions due to the extensive coverage of its free-text evidence.

Open-domain question answering is a difficult task in KBQA, which aims to answer a question in natural language in accordance with large-scale unstructured documents. The Relational GNN for Open-domain Question Answering [69] is an OpenQA architecture. This proposed model can update embeddings from a knowledge graph and a collection of linked texts together to learn contextual knowledge graph embeddings. Contextualized relations are utilized in knowledge graphs to enrich them. The bi-directional attention mechanism and hierarchical representation learning are also used for open-domain question answering tasks in this approach as well.

C. DRUG-DRUG INTERACTIONS

Over the years, graph neural network has been an emerging tool for Drug-drug interaction (DDI) prediction but has not been widely applied, and the concerned models are shown in Table 7. The graph energy neural network (GENN) [72] is the first proposed model explicitly for drug link type correlations. Motivated by the intuition that an ''energy'' can be derived over the graph, a new energy function defined by the graph neural networks is formulated and used to incorporate the dependency structures. However, this

TABLE 8. KGs-based GNN models in other fields.

Models	Objectives	Datasets	Methods	Conclusions
SA GNN [74]	To improve SA with collective AI over KGs.	Robotics KG, Generic KG, Research KG, Academic KG	Four ideas with collective AI; Object representations from agents	Better than individual AI
Power System Network Topology Identification [75]	To identify the topology of a power network.	IEEE 118-bus	Conflicting and Contradictory information	Accurate in conflicting and missing information
DEAP-FAKED [76]	A KG framework for the detection of fake news is utilized.	Kaggle Fake News, CoAID	NLP-based and GNN-based technique	F1-score of 88% and 78%
Optical Network Fault Localization [77]	To identify fault location in optical networks.	Alarm data from OTN	Alarm KG; GGNN-based method	Reasonable accuracy
Graph-Based Knowledge Tracing [78]	To predict the process of coursework proficiency with the latent knowledge structure.	ASSISTments 2009-2010, Algebra 2006-2007	Time-series node-level classification	More interpretable predictions
KG-Enhanced Review Generation Model [79]	To propose a knowledge-enhanced PRG model based on capsule GNN.	Amazon Electronic & Book IMDb Movie	Caps-GNN; Sequence and sentence generation	Better at both aspect and word levels
Few-Shot KG-To-Text Generation Model [80]	To generate a natural language text in KG with the few-shot setting.	AGENDA, WebNLG, GenWiki Fine	Representation alignment; Relation-biased KG; Multi-task learning	Better at both fully supervised and few-shot settings
NKD-GNN [81]	An analysis of the indirect relationships between identified entities is used to generate a caption for a news image.	TopNews, GoodNews	TopNews KG; NKD-GNN	Effective, but unable to extract news images from metaphors

approach only considers the correlations between drugs, and a neglected deficiency is that the relations between drugs and other entities such as targes and genes are not considered.

To address this limitation, a novel end-to-end framework, the Knowledge graph neural network (KGNN) [73], is introduced to explore the topological information of each entity in the knowledge graph, which is beneficial for DDI prediction. In addition, KGNN aggregates rich neighborhood information with a bias to learn both high-order structures and semantic relations of the KG.

D. RESEARCH IN OTHER AREAS

The GNN-based knowledge graphs also have applications in many other fields. Table 8 lists the recently proposed models, including scenarios of Fake News Detection (FND), Fault Localization (FL), Image-Text Matching (ITM), Personalized Review Generation (PRG), Situational Awareness (SA), Knowledge Tracing (KT), and Power System Network Topology Identification (PSNTI).

The DEAP-FAKED [76] encoded news content using an NLP-based approach and then identified, extracted, and mapped the named entities to a KG. Finally, the entities in the KG were encoded using a GNN-based technique. This approach can achieve better results by utilizing only the news articles' titles and handling the bias. The Optical Network Fault Localization [77] designed an alarm KG to assist network administrators to explore and visualize the relationship between alarms. Then, a GGNN-based method was put forward to reason the association between alarms and detect the root alarm. It is experimentally verified that the knowledge graphs help to construct an easy-to-understand alarm knowledge system, which leads to good accuracy. There are three methods used in Few-Shot KG-To-Text Generation Model [80]: representation alignment to help bridge the semantic gap between KG codings and PLMs, relation-biased KG linearization to derive input representations, as well as multi-task learning to learn how KG and text correspond to each other. All of these assists in producing effective semantic representations for both few-shot and fully supervised settings. According to the News Knowledge Driven Graph Neural Network (NKD-GNN) [81], the KG was built using named entities extracted from the TopNews dataset, and the relationships between those entities were explored across the KG. Therefore, by examining all relationships and implicitly assuming relationships between named entities in the news knowledge graph, this model is able to choose the best candidate for each placeholder in the news image template caption. The KG-Enhanced Review Generation Model [79] adopted the Caps-GNN to learn graph capsules to encode underlying characteristics from the HKG. The generation process contains aspect sequence generation and sentence generation. This model performs better than all the baselines due to the significant difference that KG information was included in the multi-stage generation process. The SA GNN [74] discussed four ideas of making predictions with collective AI and proposed a GNN framework that jointly learns object representations from multiple agents. Since each AI has a unique, incomplete view of the knowledge graph with noise, multiple AI agents are required to produce a forecast collectively.

The Graph-Based Knowledge Tracing [78] reformulated the knowledge tracing task as a time-series node-level classification issue in the GNN. This approach exhibited better interpretable predictions because it directly models the knowledge state for each concept and further models the edge weights using K separate neural networks for K edge types. The graph neural network can serve as a technological tool among entities in the knowledge graph, which works for missing information. Therefore, both can be merged to be adopted in topological recognition. The Power System Network Topology Identification [75] contained an additional process for inferring conflicting information. Meanwhile, the knowledge inference on contradictory information was conducted on the basis of GNN.

V. CONCLUSION

Knowledge graph, a form of data representation that uses graph structure to model the connections between things, has attracted much attention and faces many challenges. This paper depicted the importance and necessity of knowledge graph embedding and how KGE is used to solve KG problems. We present a thorough review of existing GNN-based approaches that mainly focus on four types of KG tasks, i.e., link prediction, knowledge graph alignment, knowledge graph reasoning, and node classification. We went over the specifics of the model as well as the benefits and contributions of such strategies. After that, this work focuses on how the KG method based on GNN can be applied to practical application areas such as recommender systems, question answering, and drug-drug interaction. This is the first complete graph neural network technology survey for knowledge graphs. We believe that the investigation of KGs on using GNN will receive increasing attention in the near future.

ACKNOWLEDGMENT

The authors would like to thank Universiti Teknikal Malaysia Melaka for supporting this research.

REFERENCES

[1] S. Sagiroglu and D. Sinanc, "Big data: A review," in *Proc. CTS*, San Diego, CA, USA, 2013, pp. 42–47.

[2] D. Fensel, U. Şimşek, K. Angele, E. Huaman, E. Kärle, O. Panasiuk, I. Toma, J. Umbrich, and A. Wahler, *Knowledge Graphs*. Cham, Switzerland: Springer, 2020, pp. 1–10.

[3] B. Jiang, L. Tan, Y. Ren, and F. Li, "Intelligent interaction with virtual geographical environments based on geographic knowledge graph," *ISPRS Int. J. Geo-Inf.*, vol. 8, no. 10, p. 428, Sep. 2019.

[4] D. Ringler and H. Paulheim, "One knowledge graph to rule them all? Analyzing the differences between DBpedia, YAGO, Wikidata & co.," in *Proc. Joint German/Austrian Conf. Artif. Intell.*, Dortmund, Germany, 2017, pp. 366–372.

[5] S. Auer, C. Bizer, G. Kobilarov, J. Lehmann, R. Cyganiak, and Z. Ives, "DBpedia: A nucleus for a web of open data," in *The Semantic Web*. Berlin, Germany: Springer, 2007, pp. 722–735.

[6] H. Paulheim, "Knowledge graph refinement: A survey of approaches and evaluation methods," *Semantic Web*, vol. 8, no. 3, pp. 489–508, Dec. 2016.

[7] J. Pujara, H. Miao, L. Getoor, and W. W. Cohen, "Large-scale knowledge graph identification using PSL," in *Proc. AAAI-FSS*, Arlington, TX, USA, 2013, pp. 1–4.

[8] C. Wang, M. Gao, X. He, and R. Zhang, "Challenges in Chinese knowledge graph construction," in *Proc. 31st IEEE Int. Conf. Data Eng. Workshops*, Apr. 2015, pp. 59–61.

[9] A. Rossi, D. Barbosa, D. Firmani, A. Matinata, and P. Merialdo, "Knowledge graph embedding for link prediction: A comparative analysis," *ACM Trans. Knowl. Discovery Data*, vol. 15, no. 2, pp. 1–49, Apr. 2021.

[10] Z. Sun, W. Hu, Q. Zhang, and Y. Qu, "Bootstrapping entity alignment with knowledge graph embedding," in *Proc. 27th Int. Joint Conf. Artif. Intell.*, Stockholm, Sweden, Jul. 2018, pp. 4396–4402.

[11] X. Chen, S. Jia, and Y. Xiang, "A review: Knowledge reasoning over knowledge graph," *Expert Syst. Appl.*, vol. 141, Mar. 2020, Art. no. 112948.

[12] N. Guan, D. Song, and L. Liao, "Knowledge graph embedding with concepts," *Knowl.-Based Syst.*, vol. 164, pp. 38–44, Jan. 2019.

[13] Z. Liu, M. Sun, and Y. Lin, "Knowledge representation learning: A review," *J. Comput. Res. Develop.*, vol. 53, no. 2, pp. 247–261, 2016.

[14] X. Zhao, W. Zeng, J. Tang, W. Wang, and F. M. Suchanek, "An experimental study of state-of-the-art entity alignment approaches," *IEEE Trans. Knowl. Data Eng.*, vol. 34, no. 6, pp. 2610–2625, Jun. 2022.

[15] F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner, and G. Monfardini, "The graph neural network model," *IEEE Trans. Neural Netw.*, vol. 20, no. 1, pp. 61–80, Jan. 2008.

[16] J. Zhou, G. Cui, S. Hu, Z. Zhang, C. Yang, Z. Liu, L. Wang, C. Li, and M. Sun, "Graph neural networks: A review of methods and applications," *AI Open*, vol. 1, pp. 57–81, Apr. 2021.

[17] I. Portugal, P. Alencar, and D. Cowan, "The use of machine learning algorithms in recommender systems: A systematic review," *Expert Syst. Appl.*, vol. 97, pp. 205–227, May 2018.

[18] W. Ahmed, A. Dasan, and A. P. Babu, "Developing an intelligent question answering system," *Int. J. Educ. Manage. Eng.*, vol. 7, no. 6, pp. 50–61, Nov. 2017.

[19] B. Percha and R. B. Altman, "Informatics confronts drug–drug interactions," *Trends Pharmacolog. Sci.*, vol. 34, no. 3, pp. 178–184, Mar. 2013.

[20] S. Auer, V. Kovtun, M. Prinz, A. Kasprzik, M. Stocker, and M. E. Vidal, "Towards a knowledge graph for science," in *Proc. 8th Int. Conf. Web Intell., Mining Semantics*, Novi Sad, Serbia, Jun. 2018, pp. 1–6.

[21] Z. Liu and J. Zhou, "Introduction to graph neural networks," *Synth. Lectures Artif. Intell. Mach. Learn.*, vol. 14, no. 2, pp. 1–127, Mar. 2020.

[22] Y. Zhang, Q. Fang, S. Qian, and C. Xu, "Multi-modal multi-relational feature aggregation network for medical knowledge representation learning," in *Proc. 28th ACM Int. Conf. Multimedia*, Seattle, WA, USA, Oct. 2020, pp. 3956–3965.

[23] Z. Zhang, F. Zhuang, H. Zhu, Z. Shi, H. Xiong, and Q. He, "Relational graph neural network with hierarchical attention for knowledge graph completion," in *Proc. AAAI-AI*, New York, NY, USA, 2020, pp. 9612–9619.

[24] M. Zhao, W. Jia, and Y. Huang, "Attention-based aggregation graph networks for knowledge graph information transfer," in *Proc. PAKDD*, Singapore, 2020, pp. 542–554.

[25] Y. Zhang, X. Chen, Y. Yang, A. Ramamurthy, B. Li, Y. Qi, and L. Song, "Efficient probabilistic logic reasoning with graph neural networks," 2020, *arXiv:2001.11850*.

[26] K. Ji, B. Hui, and G. Luo, "Graph attention networks with local structure awareness for knowledge graph completion," *IEEE Access*, vol. 8, pp. 224860–224870, 2020.

[27] G. Salha, S. Limnios, R. Hennequin, V.-A. Tran, and M. Vazirgiannis, "Gravity-inspired graph autoencoders for directed link prediction," in *Proc. 28th ACM Int. Conf. Inf. Knowl. Manage.*, Beijing, China, Nov. 2019, pp. 589–598.

[28] S. Liu, B. Grau, I. Horrocks, and E. Kostylev, "INDIGO: GNN-based inductive knowledge graph completion using pair-wise encoding," in *Proc. Adv. Neural Inf. Process. Syst.*, 2021, pp. 2034–2045.

[29] K. Teru, E. Denis, and W. Hamilton, "Inductive relation prediction by subgraph reasoning," in *Proc. 37th Int. Conf. Mach. Learn.*, 2020, pp. 9448–9457.

[30] M. Zhang and Y. Chen, "Link prediction based on graph neural networks," in *Proc. 32nd NeurIPS*, Montreal, QC, Canada, 2018, pp. 5171–5181.

[31] S. Wang, X. Wei, C. N. N. D. Santos, Z. Wang, R. Nallapati, A. Arnold, B. Xiang, P. S. Yu, and I. F. Cruz, "Mixed-curvature multi-relational graph neural network for knowledge graph completion," in *Proc. Web Conf.*, Ljubljana, Slovenia, Apr. 2021, pp. 1761–1771.

[32] M. Schlichtkrull, T. N. Kipf, P. Bloem, R. van den Berg, I. Titov, and M. Welling, "Modeling relational data with graph convolutional networks," in *Proc. ESWC*, Crete, Greece, 2018, pp. 593–607.

[33] X. Liu, H. Tan, Q. Chen, and G. Lin, "RAGAT: Relation aware graph attention network for knowledge graph completion," *IEEE Access*, vol. 9, pp. 20840–20849, 2021.
[34] Z. Wang, Z. Ren, C. He, P. Zhang, and Y. Hu, "Robust embedding with multi-level structures for link prediction," in *Proc. 28th Int. Joint Conf. Artif. Intell.*, Macao, China, Aug. 2019, pp. 5240–5246.
[35] Z. Zhang, J. Huang, and Q. Tan, "Association rules enhanced knowledge graph attention network," *Knowl.-Based Syst.*, vol. 239, Mar. 2022, Art. no. 108038.
[36] Z. Li, H. Liu, Z. Zhang, T. Liu, and N. N. Xiong, "Learning knowledge graph embedding with heterogeneous relation attention networks," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Feb. 19, 2021, doi: 10.1109/TNNLS.2021.3055147.
[37] Z. Ning, Z. Qiao, H. Dong, Y. Du, and Y. Zhou, "LightCAKE: A lightweight framework for context-aware knowledge graph embedding," in *Proc. PAKDD*, 2021, pp. 181–193.
[38] A. Hogan, E. Blomqvist, M. Cochez, C. d'Amato, G. D. Melo, C. Gutierrez, S. Kirrane, J. E. L. Gayo, R. Navigli, and S. Neumaier, "Knowledge graphs," *Synth. Lect. Comput. Archit.*, vol. 12, no. 2, pp. 1–257, Nov. 2021.
[39] B. D. Trisedya, J. Qi, and R. Zhang, "Entity alignment between knowledge graphs using attribute embeddings," in *Proc. AAAI-AI*, Honolulu, HI, USA, 2019, pp. 297–304.
[40] Q. Zhu, H. Wei, B. Sisman, D. Zheng, C. Faloutsos, X. L. Dong, and J. Han, "Collective multi-type entity alignment between knowledge graphs," in *Proc. Web Conf.*, Taipei, Taiwan, Apr. 2020, pp. 2241–2252.
[41] E. Faerman, O. Voggenreiter, F. Borutta, T. Emrich, M. Berrendorf, and M. Schubert, "Graph alignment networks with node matching scores," in *Proc. NeurIPS*, Vancouver, BC, Canada, 2019, pp. 1–7.
[42] Z. Sun, C. Wang, W. Hu, M. Chen, J. Dai, W. Zhang, and Y. Qu, "Knowledge graph alignment network with gated multi-hop neighborhood aggregation," in *Proc. AAA-AI*, New York, NY, USA, 2020, pp. 222–229.
[43] Z. Wang, J. Yang, and X. Ye, "Knowledge graph alignment with entity-pair embedding," in *Proc. Conf. Empirical Methods Natural Lang. Process. (EMNLP)*, Punta Cana, Dominican Republic, 2020, pp. 1672–1680.
[44] X. Mao, W. Wang, H. Xu, M. Lan, and Y. Wu, "MRAEA: An efficient and robust entity alignment approach for cross-lingual knowledge graph," in *Proc. 13th Int. Conf. Web Search Data Mining*, Houston, TX, USA, Jan. 2020, pp. 420–428.
[45] Y. Cao, Z. Liu, C. Li, Z. Liu, J. Li, and T.-S. Chua, "Multi-channel graph neural network for entity alignment," 2019, *arXiv:1908.09898*.
[46] L. Guo, W. Wang, Z. Sun, C. Liu, and W. Hu, "Decentralized knowledge graph representation learning," 2020, *arXiv:2010.08114*.
[47] Y. Yan, L. Liu, Y. Ban, B. Jing, and H. Tong, "Dynamic knowledge graph alignment," in *Proc. AAAI-AI*, 2021, pp. 4564–4572.
[48] Y. Yue, C. Yang, Y. Wang, P. G. C. N. Senarathne, J. Zhang, M. Wen, and D. Wang, "A multilevel fusion system for multirobot 3-D mapping using heterogeneous sensors," *IEEE Syst. J.*, vol. 14, no. 1, pp. 1341–1352, Mar. 2020.
[49] Y. Tao, Y. Li, and Z. Wu, "Temporal link prediction via reinforcement learning," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Toronto, ON, Canada, Jun. 2021, pp. 3470–3474.
[50] G. Wan, S. Pan, C. Gong, C. Zhou, and G. Haffari, "Reasoning like human: Hierarchical reinforcement learning for knowledge graph reasoning," in *Proc. 29th Int. Joint Conf. Artif. Intell.*, Yokohama, Japan, Jul. 2020, pp. 1926–1932.
[51] Q. Wang, Y. Hao, and F. Chen, "Deepening the IDA* algorithm for knowledge graph reasoning through neural network architecture," *Neurocomputing*, vol. 429, pp. 101–109, Mar. 2021.
[52] X. Xu, W. Feng, Y. Jiang, X. Xie, Z. Sun, and Z.-H. Deng, "Dynamically pruned message passing networks for large-scale knowledge graph reasoning," 2019, *arXiv:1909.11334*.
[53] K. Nie, K. Zeng, and Q. Meng, "Knowledge reasoning method for military decision support knowledge graph mixing rule and graph neural networks learning together," in *Proc. Chin. Autom. Congr. (CAC)*, Shanghai, China, Nov. 2020, pp. 4013–4018.
[54] X. Zhao, Y. Jia, A. Li, R. Jiang, K. Chen, and Y. Wang, "Target relational attention-oriented knowledge graph reasoning," *Neurocomputing*, vol. 461, pp. 577–586, Oct. 2021.
[55] L. Li, Z. Bi, H. Ye, S. Deng, H. Chen, and H. Tou, "Text-guided legal knowledge graph reasoning," in *Proc. CCKS*, Guangzhou, China, 2021, pp. 27–39.
[56] A. Pareja, G. Domeniconi, J. Chen, T. Ma, T. Suzumura, H. Kanezashi, T. Kaler, T. Schardl, and C. Leiserson, "EvolveGCN: Evolving graph convolutional networks for dynamic graphs," in *Proc. AAAI Conf. Artif. Intell.*, New York, NY, USA, 2020, pp. 5363–5370.
[57] Y. Ren, B. Liu, C. Huang, P. Dai, L. Bo, and J. Zhang, "Heterogeneous deep graph infomax," 2019, *arXiv:1911.08538*.
[58] F. Hu, Y. Zhu, S. Wu, L. Wang, and T. Tan, "Hierarchical graph convolutional networks for semi-supervised node classification," 2019, *arXiv:1902.06667*.
[59] Y. Liu, S. Yang, Y. Xu, C. Miao, M. Wu, and J. Zhang, "Contextualized graph attention network for recommendation with item knowledge graph," *IEEE Trans. Knowl. Data Eng.*, early access, May 24, 2021, doi: 10.1109/TKDE.2021.3082948.
[60] Y. Wang, Z. Liu, Z. Fan, L. Sun, and P. S. Yu, "DSKReG: Differentiable sampling on knowledge graph for recommendation with relational GNN," in *Proc. 30th ACM Int. Conf. Inf. Knowl. Manage.*, Oct. 2021, pp. 3513–3517.
[61] J. Ouyang, Y. Jin, and X. Zhou, "Dual knowledge multimodal network for recommender system," to be published.
[62] C. Chen, M. Zhang, W. Ma, Y. Liu, and S. Ma, "Jointly non-sampling learning for knowledge graph enhanced recommendation," in *Proc. 43rd Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, Xi'an, China, Jul. 2020, pp. 189–198.
[63] X. Wang, T. Huang, D. Wang, Y. Yuan, Z. Liu, X. He, and T.-S. Chua, "Learning intents behind interactions with knowledge graph for recommendation," in *Proc. Web Conf.*, Ljubljana, Slovenia, Apr. 2021, pp. 878–887.
[64] W. Zhang, Y. Cao, and C. Xu, "SARC: Split-and-recombine networks for knowledge-based recommendation," in *Proc. IEEE 31st Int. Conf. Tools Artif. Intell. (ICTAI)*, Portland, OR, USA, Nov. 2019, pp. 652–659.
[65] Z. Li, Z. Cui, S. Wu, X. Zhang, and L. Wang, "Fi-GNN: Modeling feature interactions via graph neural networks for CTR prediction," in *Proc. 28th ACM Int. Conf. Inf. Knowl. Manage.*, Beijing, China, Nov. 2019, pp. 539–548.
[66] C.-Y. Tai, M.-R. Wu, Y.-W. Chu, and S.-Y. Chu, "GraphSW: A training protocol based on stage-wise training for GNN-based recommender model," 2019, *arXiv:1908.05611*.
[67] H. Wang, F. Zhang, M. Zhang, J. Leskovec, M. Zhao, W. Li, and Z. Wang, "Knowledge-aware graph neural networks with label smoothness regularization for recommender systems," in *Proc. 25th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Anchorage, AK, USA, Jul. 2019, pp. 968–977.
[68] C. Gao, X. Wang, X. He, and Y. Li, "Graph neural networks for recommender system," in *Proc. 15th ACM Int. Conf. Web Search Data Mining*, Phoenix, AZ, USA, Feb. 2022, pp. 1623–1625.
[69] S. Vivona and K. Hassani, "Relational graph representation learning for open-domain question answering," 2019, *arXiv:1910.08249*.
[70] M. Yasunaga, H. Ren, A. Bosselut, P. Liang, and J. Leskovec, "QA-GNN: Reasoning with language models and knowledge graphs for question answering," 2021, *arXiv:2104.06378*.
[71] C. Zhao, C. Xiong, X. Qian, and J. Boyd-Graber, "Complex factoid question answering with a free-text knowledge graph," in *Proc. Web Conf.*, Taipei, Taiwan, Apr. 2020, pp. 1205–1216.
[72] T. Ma, J. Shang, C. Xiao, and J. Sun, "GENN: Predicting correlated drug-drug interactions with graph energy neural networks," 2019, *arXiv:1910.02107*.
[73] X. Lin, Z. Quan, Z.-J. Wang, T. Ma, and X. Zeng, "KGNN: Knowledge graph neural network for drug-drug interaction prediction," in *Proc. 29th Int. Joint Conf. Artif. Intell.*, Yokohama, Japan, Jul. 2020, pp. 2739–2745.
[74] M. Jiang, "Improving situational awareness with collective artificial intelligence over knowledge graphs," *Proc. SPIE*, vol. 11413, Apr. 2020, Art. no. 114130.
[75] C. Wang, J. An, and G. Mu, "Power system network topology identification based on knowledge graph and graph neural network," *Frontiers Energy Res.*, vol. 8, p. 403, Feb. 2021.
[76] M. Mayank, S. Sharma, and R. Sharma, "DEAP-FAKED: Knowledge graph based approach for fake news detection," 2021, *arXiv:2107.10648*.
[77] Z. Li, Y. Zhao, Y. Li, S. Rahman, F. Wang, X. Xin, and J. Zhang, "Fault localization based on knowledge graph in software-defined optical networks," *J. Lightw. Technol.*, vol. 39, no. 13, pp. 4236–4246, Apr. 8, 2021.

[78] H. Nakagawa, Y. Iwasawa, and Y. Matsuo, "Graph-based knowledge tracing: Modeling student proficiency using graph neural network," in *Proc. IEEE/WIC/ACM Int. Conf. Web Intell.*, Thessaloniki, Greece, Oct. 2019, pp. 156–163.
[79] J. Li, S. Li, W. X. Zhao, G. He, Z. Wei, N. J. Yuan, and J.-R. Wen, "Knowledge-enhanced personalized review generation with capsule graph neural network," in *Proc. 29th ACM Int. Conf. Inf. Knowl. Manage.*, Ireland, U.K., Oct. 2020, pp. 735–744.
[80] J. Li, T. Tang, W. X. Zhao, Z. Wei, N. J. Yuan, and J.-R. Wen, "Few-shot knowledge graph-to-text generation with pretrained language models," 2021, *arXiv:2106.01623*.
[81] Z. Yumeng, Y. Jing, G. Shuo, and L. Limin, "News image-text matching with news knowledge graph," *IEEE Access*, vol. 9, pp. 108017–108027, 2021.

ZI YE received the bachelor's degree in mathematics and statistical science from University College London, U.K., in 2009, and the master's degree in applied statistics from the University of Oxford, U.K., in 2010. She is currently pursuing the Ph.D. degree with Universiti Teknikal Malaysia Melaka. Her research interests include artificial intelligence and machine learning.

YOGAN JAYA KUMAR received the bachelor's and master's degrees from Universiti Sains Malaysia, and the Ph.D. degree in the field of computer science, in 2014. He is currently a Senior Lecturer with Universiti Teknikal Malaysia Melaka. His research interests include the field of text mining, information extraction, and AI applications.

GOH ONG SING is currently an Assistant Vice Chancellor at Office of Industry and Community Network. His research interests include the development of intelligent agent, machine learning and speech technology, conversational robot, and mobile services. He has led research grants funded by Malaysian Government's Intensified Research.

FENGYAN SONG received the M.S. degree from the Department of Physics, Peking University. He is proficient in quantum field theory. He currently switches to the field of artificial intelligence.

JUNSONG WANG received the B.Eng. degree in electrical engineering from the Harbin Institute of Technology, Harbin, China, in 1994, and the Ph.D. degree in automatic control from Tianjin University, China, in 2008. From 2008 to 2010, he was a Postdoctoral Fellow with the Department of Automation, Tsinghua University, Beijing, China. He is currently a Full Professor with the College of Big Data and Internet, Shenzhen Technology University, Shenzhen, China. He has published over 40 scientific articles in international journals. His current research interests include brain-inspired artificial intelligence, computer vision, and machine learning.