



# Graph Learning for Combinatorial Optimization: A Survey of State-of-the-Art

Yun Peng<sup>1</sup> · Byron Choi<sup>1</sup> · Jianliang Xu<sup>1</sup>

Received: 21 August 2020 / Revised: 16 December 2020 / Accepted: 18 March 2021 / Published online: 28 April 2021  
© The Author(s) 2021

## Abstract

Graphs have been widely used to represent complex data in many applications, such as e-commerce, social networks, and bioinformatics. Efficient and effective analysis of graph data is important for graph-based applications. However, most graph analysis tasks are combinatorial optimization (CO) problems, which are NP-hard. Recent studies have focused a lot on the potential of using machine learning (ML) to solve graph-based CO problems. Most recent methods follow the two-stage framework. The first stage is graph representation learning, which embeds the graphs into low-dimension vectors. The second stage uses machine learning to solve the CO problems using the embeddings of the graphs learned in the first stage. The works for the first stage can be classified into two categories, graph embedding methods and end-to-end learning methods. For graph embedding methods, the learning of the the embeddings of the graphs has its own objective, which may not rely on the CO problems to be solved. The CO problems are solved by independent downstream tasks. For end-to-end learning methods, the learning of the embeddings of the graphs does not have its own objective and is an intermediate step of the learning procedure of solving the CO problems. The works for the second stage can also be classified into two categories, non-autoregressive methods and autoregressive methods. Non-autoregressive methods predict a solution for a CO problem in one shot. A non-autoregressive method predicts a matrix that denotes the probability of each node/edge being a part of a solution of the CO problem. The solution can be computed from the matrix using search heuristics such as beam search. Autoregressive methods iteratively extend a partial solution step by step. At each step, an autoregressive method predicts a node/edge conditioned to current partial solution, which is used to its extension. In this survey, we provide a thorough overview of recent studies of the graph learning-based CO methods. The survey ends with several remarks on future research directions.

**Keywords** Graph representation learning · Graph neural network · Combinational optimization

## Abbreviations

ML	Machine learning	CO	Combinatorial optimization
GNN	Graph neural network	MVC	Minimum vertex cover
DL	Deep learning	MIS	Maximum independent set
RL	Reinforcement learning	TSP	Travelling salesman problem
CNN	Convolutional neural network	GC	Graph coloring
DNN	Deep neural network	MDS	Minimum dominating set
RNN	Recurrent neural network	MM	Maximum matching
MLP	Multi-layer perceptron	MaxCut	Maximum cut
MDP	Markov decision process	MC	Maximum clique
MCTS	Monte Carlo tree search	SI	Subgraph isomorphism
		GSim	Graph similarity
		MF	Matrix factorization
		B & B	Branch and bound
		MILP	Mixed-integer linear programming
		BFS	Breadth-first search
		DFS	Depth-first search

✉ Jianliang Xu  
xujl@comp.hkbu.edu.hk

<sup>1</sup> Department of Computer Science, Hong Kong Baptist University, DLB 626, Level 6, David C. Lam Building, Shaw Campus, Kowloon Tong, Hong Kong

## 1 Introduction

Graphs are ubiquitous and are used in a wide range of domains, from e-commerce [21, 78] to social networking [31, 70] to bioinformatics [20, 76]. Effectively and efficiently analyzing graph data are important for graph-based applications. However, many graph analysis tasks are combinatorial optimization (CO) problems, such as the traveling salesman problem (TSP) [67], maximum independent set (MIS) [14], maximum cut (MaxCut) [23], minimum vertex cover (MVC) [42], maximum clique (MC) [10], graph coloring (GC) [54], subgraph isomorphism (SI) [22], and graph similarity (GSim) [55]. These graph-based CO problems are NP-hard. In the existing literature on this subject, there are three main approaches used to solve a CO problem: exact algorithms, approximation algorithms, and heuristic algorithms. Given a CO problem on a graph  $G$ , exact algorithms aim to compute an optimum solution. Due to the NP-hardness of the problems, the worst-case time complexity of exact algorithms is exponential to the size of  $G$ . To reduce time complexity, approximation algorithms find a suboptimal solution that has a guaranteed approximation ratio to the optimum, with a worst-case polynomial runtime. Nevertheless, many graph-based CO problems, such as general TSP [53], GC [35], and MC [27], are inapproximable with such a bounded ratio. Thus, heuristic algorithms are designed to efficiently find a suboptimal solution with desirable empirical performance. Despite having no theoretical guarantee of optimality, heuristic algorithms often produce good enough solutions in practice.

The practice of applying machine learning (ML) to solve graph-based CO problems has a long history. For example, as far back as the 1980s, researchers were using the Hopfield neural network to solve TSP [28, 60]. Recently, the success of deep learning methods has led to an increasing attention being paid to this subject [5, 14, 43, 67]. Compared to manual algorithm designs, ML-based methods have several advantages in solving graph-based CO problems. First, ML-based methods can automatically identify distinct features from training data. In contrast, human algorithm designers need to study the heuristics with substantial problem-specific research based on intuitions and trial-and-errors. Second, for a graph-based CO problem, ML has the potential to find useful features that it may be hard to specify by human algorithm designers, enabling it to develop a better solution [29]. Third, an ML-based method can adapt to a family of CO problems. For example, S2V-DQN [14] can support TSP, MVC, and MaxCut; GNNTS [42] can support MIS, MVC, and MC. In comparison, it is unlikely for a handcrafted algorithm of one CO problem to be adapted to other CO problems.

Most recent graph learning-based CO methods follow the two-stage framework. The first stage is graph representation learning which embeds the graphs into low-dimension vectors. The second stage uses machine learning to solve the CO problems using the embedding vectors of the graphs learned in the first stage. In this survey, we review the state-of-the-art works of the two stages, respectively.

For the first stage, existing graph representation learning techniques that have been used in ML-based CO methods can be classified into two categories: graph embedding methods and end-to-end learning methods. On one hand, graph embedding methods embed the nodes of a graph into low-dimension vectors. The embedding vectors of the graph learned are inputted to downstream machine learning tasks to solve CO problems. Graph embedding has its learning objective, which may not rely on the CO problems to be solved. The embeddings of the graph are fixed during the solving of the downstream task. On the other hand, in end-to-end learning methods, graph representation learning does not have its own learning objective and is an intermediate step of the learning procedure of solving the CO problem. The embeddings learned are specific to the CO problem being solved.

For the second stage, existing works can be classified into two categories: non-autoregressive methods and autoregressive methods. On one hand, non-autoregressive methods predict a solution for a graph-based CO problem in one shot. For example, for the TSP problem on a graph, a non-autoregressive method predicts a matrix, where each element of the matrix is the probability of an edge belonging to a TSP tour. The TSP tour can be computed from the matrix using beam search. On the other hand, autoregressive methods iteratively extend a partial solution step by step. For the TSP problem, at each step, an autoregressive method predicts an edge conditioned to the current partial TSP tour, which is used to extend the current partial TSP tour.

There have been several surveys of graph representation learning [7, 11, 13, 26, 77]. However, existing surveys mainly focus on the graph representation learning models and their applications in node classification, link prediction or graph classification. In contrast, we focus on using graph learning to solve CO problems. There have also been several previous surveys that have discussed ML-based CO methods [5, 38, 43]. The present survey, however, has different emphases from previous studies. The survey [5] focuses on branch and bound (B&B) search techniques for the mixed-integer linear programming (MILP) problem. Although many graph-based CO problems can be formulated using MILP and solved using the B&B method, most existing ML-based methods for solving graph-based CO problems focus on graph-specific methods. Mazyavkina et al. [43] discuss RL-based CO methods. However, there are ML-based CO methods that do not use RL. This survey is not limited to RL

approaches. Lamb et al. [38] survey the GNN-based neural-symbolic computing methods. Symbolic computing is a broad field and graph-based CO is a topic of it. In contrast, we focus on the graph-based CO problems in this survey.

The rest of this survey is organized as follows. Section 2 presents the notations and preliminaries. Section 3 summarizes graph representation learning techniques. Section 4 discusses the use of ML to solve graph-based CO problems. Section 5 suggests directions for future research. Section 6 concludes this survey.

## 2 Notations and Preliminaries

In this section, we present some of the notations and definitions that are frequently used in this survey.

We denote a graph by  $G = (V, E, \mathbf{X})$ , where  $V$  and  $E$  are the node set and the edge set of  $G$ , respectively,  $\mathbf{X}^{|V| \times d}$  is the matrix of initial features of all nodes, and  $\mathbf{x}_u = \mathbf{X}[u, \cdot]$  denotes the initial features of node  $u$ . We may choose  $u \in G$  or  $u \in V$  to denote a node of the graph, when the choice is more intuitive. Similarly, we may use  $(u, v) \in G$  or  $(u, v) \in E$  to denote an edge of the graph. The adjacency matrix of  $G$  is denoted by  $\mathbf{A}$ . The weight of edge  $(u, v)$  is denoted by  $w_{u,v}$ . We use  $\mathbf{P}$  to denote the transition matrix, where  $\mathbf{P}[u, v] = w_{u,v} / \sum_{v' \in G} w_{u,v'}$ .  $\mathbf{P}^k = \prod_k \mathbf{P}$  is the  $k$ -step transition matrix and  $\mathbf{P}$  is also called the 1-step transition matrix. We use  $g$  to denote a subgraph of  $G$  and  $G \setminus g$  to denote the subgraph of  $G$  after removing all nodes in  $g$ . For a node  $u \in G$ ,  $u \notin g$ , we use  $g \cup \{u\}$  to denote adding the node  $u$  and the edges  $\{(u, v) | v \in g, (u, v) \in G\}$  to  $g$ .  $G$  can be a directed or undirected graph. If  $G$  is directed,  $(u, v)$  and  $(v, u)$  may not present simultaneously in  $E$ .  $N^o(u)$  and  $N^i(u)$  denote the outgoing and incoming neighbors of  $u$ , respectively. If  $G$  is undirected,  $N(u)$  denotes the neighbors of  $u$ . We use a bold uppercase character to denote a matrix (e.g.,  $\mathbf{X}$ ), a bold lowercase character to denote a vector (e.g.,  $\mathbf{x}$ ), and a lowercase character to denote a scalar (e.g.,  $x$ ). The embedding vectors (or embeddings for short) of a node  $u$  and a graph  $G$  are  $d$ -dimensional vectors denoted by  $\mathbf{h}_u$  and  $\mathbf{h}_G$ , respectively. Table 1 summarizes the notations of frequently used symbols.

A graph-based CO problem is formulated in Definition 1.

**Definition 1** Given a graph  $G$  and a cost function  $c$  of the subgraphs of  $G$ , a CO problem is to find the optimum value of  $c$  or the corresponding subgraph that produces that optimum value.

For example, for a graph  $G$ , the maximum clique (MC) problem is to find the largest clique of  $G$ , and the minimum vertex cover (MVC) problem is to find the minimum set of nodes that are adjacent to all edges in  $G$ .

**Table 1** Notations of frequently used symbols and their meaning

Notation	Description
$G = (V, E, \mathbf{X})$	A graph
$\mathbf{A}$	The adjacent matrix of $G$
$u, v$	Node $u$ and node $v$ of $G$
$\mathbf{X}$	A matrix of features of all nodes
$\mathbf{x}_u$	The features of $u$ , i.e., a row of $\mathbf{X}$
$\mathbf{y}$	A signal, i.e., a column of $\mathbf{X}$
$\mathcal{N}_u$	Neighborhood of $u$
$\mathbf{h}_u, \mathbf{h}_G$	Embedding vector of a node $u$ , a graph $G$
$\mathbf{H}$	The matrix of embedding vectors of all nodes
$d$	Dimension of a vector
$f_\theta^l$	The $l$ -th layer of a neural network with parameter $\theta$

### 2.1 Overview of Graph Learning-based CO Methods

Most existing methods that use machine learning to solve the graph-based CO problem follow the two-stage framework, as illustrated in Fig. 1. Given an input graph, the first stage is to learn the representation of the graph in a low-dimension embedding space. The nodes or edges of the graph are represented as embedding vectors (or *embeddings* for short). The techniques for the first stage are discussed in Sect. 3. The second stage uses machine learning to solve the CO problem using the embeddings of the graph learned in the first stage. The techniques for the second stage are discussed in Sect. 4.

There are mainly two ways to learn graph representation in the first stage. In the first way, the embeddings are learned by graph embedding methods. Graph embedding has its own learning objectives that may not rely on the objective of the CO problem to be solved in the second stage. The CO problem is solved as a downstream task of graph embedding and the gradient of the loss of the CO problem in the second stage will not be back-propagated to the first stage. In the second way, the CO problem is solved in the end-to-end manner. The first stage does not have its own learning objective and the gradient of the second stage is back-propagated to the first stage for learning the embeddings of the graph.

There are mainly two different approaches to solve the CO problems in the second stage, namely non-autoregressive methods and autoregressive methods. The non-autoregressive methods predict a solution for a graph-based CO problem in one shot. A non-autoregressive method predicts a matrix that denotes the probability of each node/edge being a part of a solution. The solution of the CO problem can be computed from the matrix by search heuristics, e.g., beam search. The autoregressive methods compute the solution by iteratively extending a partial solution. At each time step, the node/edge that is used to extend the partial solution is predicted conditioned to the current partial solution.

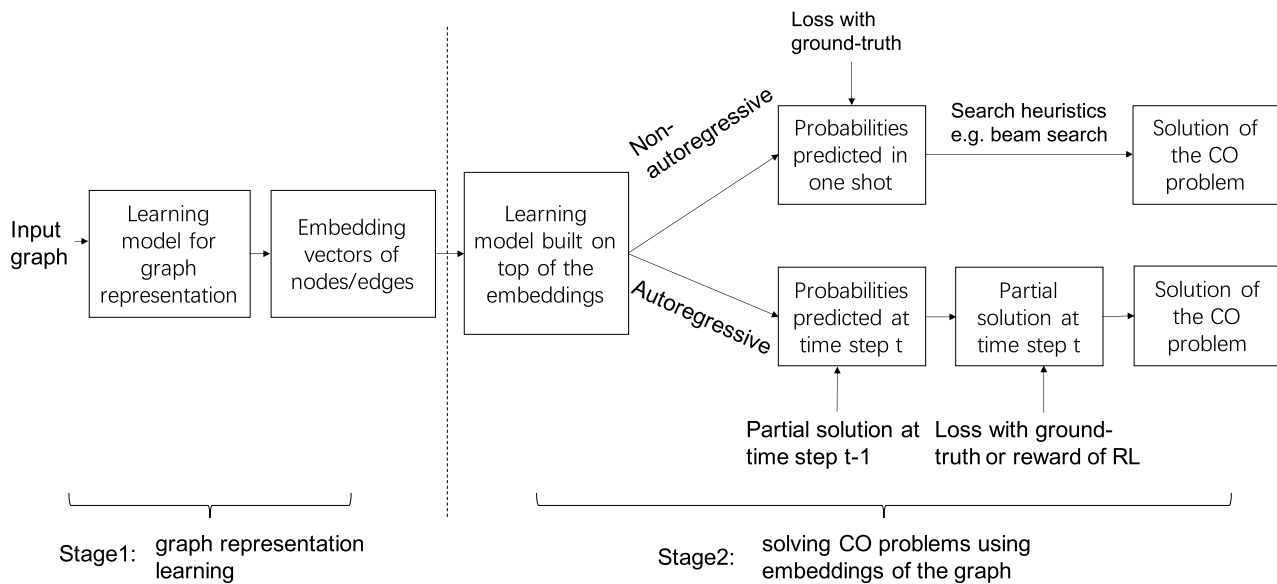


Fig. 1 Overview of the two stages of ML-based CO methods

### 3 Graph Representation Learning Methods

In this section, we survey the graph representation learning methods that have been applied to solve graph-based CO problems. In Sect. 3.1, we review the graph embedding methods, which learn the embeddings of the graph independently to the downstream task of solving the CO problem, and in Sect. 3.2, we review the end-to-end learning methods that learn the embeddings of the graph as an intermediate step of solving the CO problem.

#### 3.1 Graph Embedding Methods

We first review generalized SkipGram and AutoEncoder that are two widely used models in graph embedding.

##### 3.1.1 Generalized SkipGram

The generalized SkipGram model is extended from the well-known SkipGram model [45] for embedding words in natural language processing. The generalized SkipGram model relies on the neighborhood  $\mathcal{N}_u$  of node  $u$  to learn an embedding of  $u$ . The objective is to maximize the likelihood of the nodes in  $\mathcal{N}_u$  conditioned on  $u$ .

$$\max P(v_1, v_2, \dots, v_{|\mathcal{N}_u|} | u), v_i \in \mathcal{N}_u \quad (1)$$

Assuming conditional independence,  $P(v_1, v_2, \dots, v_{|\mathcal{N}_u|} | u) = \prod_{v_i \in \mathcal{N}_u} P(v_i | \mathbf{h}_u)$ .  $P(v_i | \mathbf{h}_u)$  can be

defined as  $\frac{\mathbf{h}_v^T \mathbf{h}_u}{\sum_{v \in G} \mathbf{h}_v^T \mathbf{h}_u}$ . Maximizing  $\prod_{v_i \in \mathcal{N}_u} P(v_i | \mathbf{h}_u)$  is then equivalent to maximizing its logarithm. Hence, (1) becomes

$$\begin{aligned} \max \sum_{v_i \in \mathcal{N}_u} \log P(v_i | \mathbf{h}_u) \\ = \max \sum_{v_i \in \mathcal{N}_u} \log \frac{\mathbf{h}_{v_i}^T \mathbf{h}_u}{\sum_{v \in G} \mathbf{h}_v^T \mathbf{h}_u} \end{aligned} \quad (2)$$

Since computing the denominator of the softmax in (2) is time consuming, many optimization techniques have been proposed. Negative sampling [45] is one of the most well-known techniques. Specifically, the nodes in the neighborhood  $\mathcal{N}_u$  of  $u$  are regarded as positive samples of  $u$ . On the other hand, the nodes not in  $\mathcal{N}_u$  are considered negative samples of  $u$ . Then, maximizing the likelihood in Formula 2 can be achieved as follows:

$$\max \log \sigma(\mathbf{h}_v^T \mathbf{h}_u) + \sum_{i=1}^K \mathbb{E}_{\bar{v} \sim P_n} \log \sigma(-\mathbf{h}_{\bar{v}}^T \mathbf{h}_u), \quad (3)$$

where  $v$  is a positive sample of  $u$ ,  $\bar{v}$  is a negative sample,  $P_n$  is the probability distribution of negative samples,  $\bar{v} \sim P_n$  means sampling a node from the probability distribution  $P_n$ ,  $K$  is the number of negative samples,  $\sigma$  is the sigmoid activation function, and  $\mathbb{E}$  is the expectation.

To conveniently adopt the gradient descent algorithms, maximizing an objective is often rewritten as minimizing its negative. Thus, the objective function of the generalized SkipGram model is to minimize the loss  $\mathcal{L}$  as follows:

$$\begin{aligned} \min \mathcal{L} = & \min -\log \sigma(\mathbf{h}_v^T \mathbf{h}_u) \\ & - \sum_{i=1}^K \mathbb{E}_{v \sim p_n} \log \sigma(-\mathbf{h}_v^T \mathbf{h}_u) \end{aligned} \tag{4}$$

Existing studies on the generalized SkipGram model define the neighborhood in different ways. For example, LINE [63] defines the 1-hop neighbors as the neighborhood in order to preserve the second-order proximity; DeepWalk [56] uses the random walk to define the neighborhood for preserving more global structural information of  $G$ .

### 3.1.2 AutoEncoder

AutoEncoder is composed of an encoder and a decoder. For a graph-based CO problem, the encoder encodes the nodes in the graph into  $d$ -dimensional embedding vectors. The decoder then predicts a solution to the CO problem using the node embeddings (e.g., PointerNet [67]).

Formally, the encoder is a function

$$enc : \mathbb{R}^{d'} \rightarrow \mathbb{R}^d$$

$enc(\mathbf{x}_u)$  embeds node  $u$  into  $\mathbf{h}_u \in \mathbb{R}^d$ .

There are several different types of decoder. For instance, the inner product-based decoder, the reconstruction-based decoder, and the classification-based decoder are three widely-used decoders.

The inner product-based decoder is a function

$$dec : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$$

$dec(\mathbf{h}_u, \mathbf{h}_v)$  returns the similarity of  $\mathbf{h}_u$  and  $\mathbf{h}_v$ . Let  $sim(u, v)$  denotes the proximity of  $u$  and  $v$  in  $G$  (e.g.,  $\mathbf{A}_u[u, v]$  in [68]). The objective function of the inner product decoder is to minimize the loss

$$\mathcal{L} = \sum_{(u,v) \in \mathcal{D}} dist(dec(\mathbf{h}_u, \mathbf{h}_v), sim(u, v)), \tag{5}$$

where  $\mathcal{D}$  is the training dataset and  $dist$  is a user-specified distance function.

The reconstruction-based decoder is a function

$$dec : \mathbb{R}^d \rightarrow \mathbb{R}^{d'}$$

$dec(\mathbf{h}_u)$  outputs  $\hat{\mathbf{x}}_u$  as the reconstruction of  $\mathbf{x}_u$ . The objective function is to minimize the reconstruction loss

$$\mathcal{L} = \sum_{u \in G} ||(dec(\mathbf{h}_u), \mathbf{x}_u)||_2^2$$

The encoder and the decoder can be implemented by different types of neural networks, e.g., the multi-layer perceptron (MLP) [68] or the recurrent neural network (RNN) [65].

### 3.1.3 Generalized SkipGram-Based Graph Embedding Method

This subsection reviews the generalized SkipGram-based graph embedding methods DeepWalk [56], Node2Vec [24], and Struc2Vec [58]; and the subgraph-based graph embedding methods DeepGK [73], Subgraph2Vec [48], RUM [75], Motif2Vec [15], and MotifWalk [50].

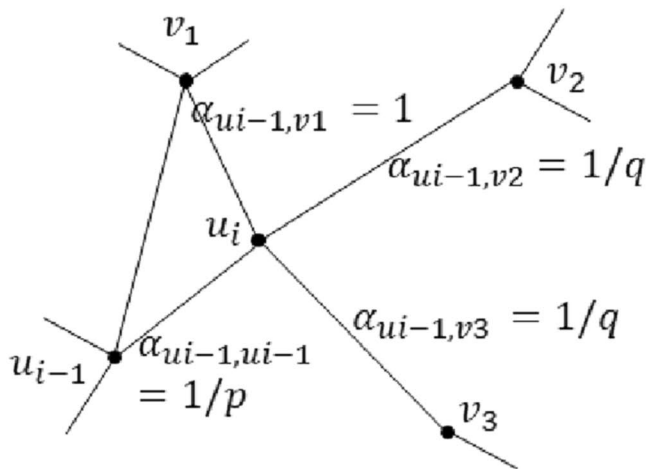
DeepWalk [56] was one of the earliest works to introduce the generalized SkipGram model to graph embedding. The main idea of DeepWalk is to sample a set of truncated random walks of the graph  $G$ , and the nodes in a window of a random walk are regarded as co-occurrence. The neighborhood of a node is the nodes that co-occurred with it. DeepWalk uses the generalized SkipGram model with the negative sampling (refer to Formula 4) to learn the graph embedding.

To incorporate more flexibility into the definition of node neighborhood, Node2Vec [24] introduces breadth-first search (BFS) and depth-first search (DFS) in neighborhood sampling. The nodes found by BFS and DFS can capture different structural properties. Node2Vec uses the second-order random walk to simulate the BFS and DFS. (“second-order” means that when the random walk is at the step  $i$ , the random walk needs to look back to the step  $i - 1$  to decide the step  $i + 1$ .) Two parameters  $p$  and  $q$  are introduced to control the random walk.  $p$  controls the probability of return to an already visited node in the following two steps, and  $q$  controls the probability of visiting a close or a far node in the following two steps. Let  $u_i$  denotes the current node in the walk and  $u_{i-1}$  denotes the previous node. The probability of the random walk to visit the next node  $u_{i+1}$  is defined as below.

$$\begin{aligned} P(u_{i+1}|u_i) = & \begin{cases} \alpha_{u_{i-1}, u_{i+1}} \times w_{u_i, u_{i+1}} & \text{if } (u_i, u_{i+1}) \in G \\ 0 & \text{otherwise} \end{cases} \\ \alpha_{u_{i-1}, u_{i+1}} = & \begin{cases} 1/p & \text{if } dist(u_{i-1}, u_{i+1}) = 0 \\ 1 & \text{if } dist(u_{i-1}, u_{i+1}) = 1 \\ 1/q & \text{if } dist(u_{i-1}, u_{i+1}) = 2 \end{cases} \end{aligned} \tag{6}$$

where  $dist(u_{i-1}, u_{i+1})$  is the shortest distance from  $u_{i-1}$  to  $u_{i+1}$  and  $w_{u_i, u_{i+1}}$  is the weight of the edge  $(u_i, u_{i+1})$ . An example is shown in Fig. 2. The current node of the random walk is  $u_i$ . There are four nodes  $u_{i-1}$ ,  $v_1$ ,  $v_2$  and  $v_3$  that can be the next node of the random walk. The probability of selecting each of them as the next node is shown in Fig. 2.

Struc2Vec [58] argues that the random walks of Node2Vec cannot find nodes that have similar structures but are far away. Struc2Vec builds a multi-layer graph  $G'$  for the input graph  $G$ . The layer  $l$  is a complete graph  $G'_l$ , where each node in  $G$  is a node in  $G'_l$  and each edge  $(u, v) \in G'_l$  is weighted by the structural similarity of the  $l$ -hop



$$p(u_{i-1}|u_i) = w_{u_{i-1}, u_i} / p$$

$$p(v_1|u_i) = w_{u_i, v_1}$$

$$p(v_2|u_i) = w_{u_i, v_2} / q$$

$$p(v_3|u_i) = w_{u_i, v_3} / q$$

**Fig. 2** An example of selecting the next node by the second-order random walk of Node2Vec [24].  $u_i$  is the current node of the random walk and  $u_{i-1}$  is the previous node.  $u_{i-1}$ ,  $v_1$ ,  $v_2$ , and  $v_3$  can be selected as the next node with the corresponding probabilities, respectively

neighborhoods of  $u$  and  $v$  in  $G$ . In this way, two nodes that are far away in  $G$  can reach each other by just one hop in  $G'_i$ . The nodes in  $G'_i$  can have directed edges to the nodes in  $G'_{i-1}$  and  $G'_{i+1}$ . Random walks are sampled on  $G'$ , and the generalized SkipGram model is used to learn the node embedding.

Besides using paths to sample the neighborhood, many works use representative subgraphs of the input graph. The representative subgraphs may be termed motifs, graphlets or kernels in different studies. Yanardag and Vishwanathan [73] propose DeepGK, which is the earliest work embedding the motifs. The neighborhood of a motif  $g$  is defined as the motifs within a small distance from  $g$ . The generalized SkipGram model is used to learn the embeddings for the motifs.

Yu et al. [75] propose a network representation learning method using motifs (RUM). RUM builds a motif graph  $G'$  for the input graph  $G$ , where each node in  $G'$  is a motif of  $G$  and two nodes have an edge in  $G'$  if the corresponding motifs share common nodes. Triangle is used as the graph motif in RUM. RUM uses random walks on the motif graph  $G'$  to define the neighborhood of a motif. Then, the generalized SkipGram model is used to learn the embedding of the motif. An original node  $u$  of  $G$  may occur in multiple motifs of  $G'$ . RUM uses the average of the embeddings of the motifs as the embedding of  $u$ .

Dareddy et al. [15] propose another type of motif graph. Given a graph  $G = (V, E)$ , for each motif  $g$ , Motif2Vec builds a motif graph  $G' = (V, E')$ , where the weight of an edge  $(u, v) \in E'$  is the number of motif instances of  $g$  in  $G$  that contain node  $u$  and  $v$ . Then, Motif2Vec simulates a set of random walks on each motif graph and uses Node2Vec [24] to learn the embeddings of the nodes in  $G$ . A similar idea is also proposed in the MotifWalk method of [50].

Narayanan et al. [48] propose Subgraph2Vec to compute the embeddings of the neighboring subgraphs of the nodes

in the input graph. Let  $g_u$  denotes the neighboring subgraph of a node  $u$ , Subgraph2Vec computes  $\mathbf{h}_{g_u}$  using the generalized SkipGram model. The neighborhood of  $g_u$  is defined as the neighboring subgraphs of the neighbors of  $u$ , i.e.,  $\{g_v | v \in N(u)\}$ .

### 3.1.4 AutoEncoder-based Graph Embedding

AutoEncoder-based graph embedding often preserves the graph structure properties measured by the following proximities.

**Definition 2** Given a graph  $G = (V, E)$ , the *first-order proximity* from  $u$  to  $v$  is the weight of  $(u, v)$ . If  $(u, v) \in E$ ,  $p^{(1)}(u, v) = w_{u,v}$ ; otherwise,  $p^{(1)}(u, v) = 0$ .

The first-order proximity captures the direct relationship between nodes. The second-order proximity captures the similarity of the neighbors of two nodes.

**Definition 3** Given a graph  $G$ , the *second-order proximity* between  $u$  and  $v$  is  $p^{(2)}(u, v) = \text{sim}(\mathbf{p}^{(1)}(u), \mathbf{p}^{(1)}(v))$ , where  $\mathbf{p}^{(1)}(u)$  is the vector of the first-order proximity from  $u$  to all other nodes in  $G$ , i.e.,  $\mathbf{p}^{(1)}(u) = (p^{(1)}(u, v_1), p^{(1)}(u, v_2), \dots, p^{(1)}(u, v_{|V|}))$ , and  $\text{sim}$  is a user-specified similarity function.

The first- and second-order proximities encode the local structures of a graph. Proximities to capture more global structures of a graph have also been proposed in the literature. For example, Cai et al. [7] propose to use  $p^{(k)}(u, v)$  (recursively defined, similar to Definition 3) as the  $k$ -th-order proximity between  $u$  and  $v$ , Cao et al. [8] use the  $k$ -step transition probability  $\mathbf{P}^k[u, v]$  to measure the  $k$ -step relationship from  $u$  to  $v$ , Chen et al. [12] use the node

centrality, Tsitsulin et al. [64] use the Personalized PageRank, and Ou et al. [52] use the Katz Index and Adamic-Adar to measure more global structural properties of  $G$ .

Large-scale information network embedding (LINE) [63] preserves both the first- and second-order proximity in graph embedding using two AutoEncoders, respectively. In order for AutoEncoder to preserve the first-order proximity, the encoder is a simple embedding lookup [9]. The decoder outputs the estimated adjacent matrix using the node embeddings, and the objective is to minimize the loss between the estimated adjacent matrix and the ground truth.

The decoder of LINE is designed as follows. Since adjacent nodes  $u$  and  $v$  in  $G$  have high first-order proximity, they should be close in the embedding space. LINE uses the inner product of  $\mathbf{h}_u$  and  $\mathbf{h}_v$  to measure the distance between  $u$  and  $v$  in the embedding space, as shown below.

$$P_1(u, v) = \frac{1}{1 + \exp(-\mathbf{h}_u^T \mathbf{h}_v)} \tag{7}$$

$P_1(\cdot, \cdot)$  defines the estimated distribution of the first-order proximity (i.e., the estimated adjacent matrix). LINE ensures that the estimated distribution  $P_1(\cdot, \cdot)$  is close to the empirical distribution  $\hat{P}_1(\cdot, \cdot)$  so as to preserve the first-order proximity.

$$\mathcal{L}_1 = \min \text{dist}(\hat{P}_1(\cdot, \cdot), P_1(\cdot, \cdot)) \tag{8}$$

where  $\hat{P}_1(u, v) = \frac{w_{u,v}}{\sum_{(u',v') \in G} w_{u',v'}}$  and  $\text{dist}$  is the distance between two probability distributions. If the KL-divergence is used as  $\text{dist}$ ,  $\mathcal{L}_1$  becomes

$$\mathcal{L}_1 = \min - \sum_{(u,v) \in G} w_{u,v} \log P_1(u, v) \tag{9}$$

In order for AutoEncoder to preserve the second-order proximity, the encoder is a simple embedding lookup [9]. The decoder outputs an estimated distribution between each node and its neighbors. The estimated distribution is reconstructed from the embeddings of the nodes. The objective is to minimize the reconstruction loss between the estimated distribution and the ground truth.

The decoder is designed as follows. Inspired by word embedding [39], the neighbors of  $u$  are regarded as the ‘‘context’’ of  $u$ . LINE uses a conditional probability  $P_2(v|u)$  defined in Formula 10 to model the estimated probability of  $u$  generating a neighbor  $v$ .

$$P_2(v|u) = \frac{\exp(\mathbf{h}'_v{}^T \mathbf{h}_u)}{\sum_{v' \in G} \exp(\mathbf{h}'_{v'}{}^T \mathbf{h}_u)}, \tag{10}$$

where  $\mathbf{h}'$  is the vector of a node when the node is regarded as context.

$P_2(\cdot|u)$  defines the estimated distribution of  $u$  over the context. The nodes  $u$  and  $u'$  in  $G$  that have a high second-order proximity should have similar estimated distributions over the context, i.e.,  $P_2(\cdot|u)$  should be similar to  $P_2(\cdot|u')$ . This can be achieved by minimizing the distance between the estimated distribution  $P_2(\cdot|u)$  and the empirical distribution  $\hat{P}_2(\cdot|u)$ , for each node  $u$  in  $G$ . The empirical distribution  $\hat{P}_2(\cdot|u)$  is defined as  $\hat{P}_2(v|u) = w_{u,v} / \sum_{u',v'} w_{u',v'}$ . LINE preserves the second-order proximity as follows.

$$\mathcal{L}_2 = \min \sum_{u \in G} \text{dist}(\hat{P}_2(\cdot|u), P_2(\cdot|u)) \tag{11}$$

Using the KL-divergence for  $\text{dist}$ , Formula 11 produces

$$\mathcal{L}_2 = \min - \sum_{(u,v) \in G} w_{u,v} \log P_2(v|u) \tag{12}$$

LINE trains the two AutoEncoders separately. The node embeddings generated by the two AutoEncoders are concatenated as the embeddings of the nodes. The model of LINE is also adopted by Tang et al. [62] to embed the words in a heterogeneous text graph.

Wang et al. [68] argue that LINE is a shallow model, in the sense that it cannot effectively capture the highly non-linear structure of a graph. Therefore, structural deep network embedding (SDNE) is proposed as a mean of using the deep neural network to embed the nodes. As with LINE, SDNE also preserves the first- and second-order proximity. Both the encoder and decoder of SDNE are MLPs. Given a graph  $G$ , the encoder embeds  $\mathbf{x}_u$  to  $\mathbf{h}_u$ , where  $\mathbf{x}_u$  is the  $u$ -th row in the adjacent matrix  $\mathbf{A}$  of  $G$ , and the decoder reconstructs  $\hat{\mathbf{x}}_u$  from  $\mathbf{h}_u$ .

SDNE preserves the first-order proximity by minimizing the distance in the embeded space for the adjacent nodes in  $G$ .

$$\mathcal{L}_1 = \sum_{(u,v) \in G} \mathbf{A}[u, v] \times \|\mathbf{h}_u - \mathbf{h}_v\|_2^2$$

The second-order proximity is preserved by minimizing the reconstruction loss.

$$\mathcal{L}_2 = \sum_{u \in G} \|\hat{\mathbf{x}}_u - \mathbf{x}_u\|_2^2$$

SDNE combines  $\mathcal{L}_1$ ,  $\mathcal{L}_2$ , and a regularizer term as the objective function and jointly optimizes them by means of a deep neural network. The first- and second-order proximity are preserved and the graph embedding learned is more robust than LINE. As demonstrated in experiments, SDNE outperforms LINE in several downstream tasks (e.g., node classification and link prediction).

Versatile graph embedding method (VERSE) [64] shows that the first- and second-order proximity are not sufficient to

capture the diverse forms of similarity relationships among nodes in a graph. Tsitsulin et al. [64] propose to use a function  $\text{sim}(u, v)$  to measure the similarity between any two nodes  $u$  and  $v$  in  $G$ , where  $\text{sim}(\cdot, \cdot)$  can be any similarity function. The similarity distribution of  $u$  to all other nodes can be defined by  $\text{sim}(u, \cdot)$ . The encoder of VERSE is a simple embedding lookup. The decoder estimates the similarity distribution using the node embeddings, as in Formula 10. The objective is to minimize the reconstruction loss between the estimated similarity distribution and the ground truth.

Dave et al. [16] propose Neural-Brane to capture both node attribute information and graph structural information in the embedding of the graph. Bonner et al. [6] study the interpretability of graph embedding models.

### 3.1.5 Discussion

The generalized SkipGram model is inspired by the word embedding model in natural language processing (NLP). Random walks of the graphs, which are the analog of sentences in texts are widely used by the generalized SkipGram model-based methods for computing the embeddings of the graphs. However, computing random walks are time-consuming. Moreover, the generalized SkipGram model is often regarded as a shallow model when compared to AutoEncoder. AutoEncoder can be deeper by stacking more layers and has more potentials to encode the complex and nonlinear relationships between the nodes of a graph [68]. Recent works of word embedding in NLP also verify the advantage of AutoEncoder [18]. However, designing the architectures of the encoder and decoder and the loss function to encode the structure information of the graph is challenging.

Graph embedding methods can precompute the embedding vectors of graphs. The advantage is that the structure information encoded in the embeddings can be transferred to different downstream tasks. Graph embedding methods learn the embeddings of the graph without considering the downstream CO problems to be solved. The embeddings may not encode the information that are critical for solving the CO problem. There is an opportunity that the performance of the graph embedding-based methods may be inferior to the end-to-end learning methods for solving CO problems. Therefore, there have been recent studies on alternative graph representation learning methods for solving CO problems such as end-to-end learning methods.

## 3.2 End-to-End Method

Graph neural network (GNN) and AutoEncoder are widely used in the end-to-end learning methods of solving CO problems, where computing the embeddings of the graphs are an intermediate step.

### 3.2.1 Graph Neural Network

Graph neural network uses the graph convolution operation to aggregate graph structure and node content information. Graph convolution can be divided into two categories: i) spectral convolutions, defined using the spectra of a graph, which can be computed from the eigendecomposition of the graph's Laplacian matrix, and ii) spatial convolutions, directly defined on a graph by information propagation.

#### A) Graph Spectral Convolution

Given an undirected graph  $G$ ,  $\mathbf{L} = \mathbf{I} - \mathbf{D}^{-1/2}\mathbf{A}\mathbf{D}^{-1/2}$  is the normalized Laplacian matrix of  $G$ .  $\mathbf{L}$  can be decomposed into  $\mathbf{L} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$ , where  $\mathbf{U}$  is the eigenvectors ordered by eigenvalues,  $\mathbf{\Lambda}$  is the diagonal matrix of eigenvalues, and  $\mathbf{\Lambda}[i, i]$  is the  $i$ -th eigenvalue  $\lambda_i$ .

The graph convolution  $*_G$  of an input signal  $\mathbf{s} \in \mathbb{R}^{|V|}$  with a filter  $\mathbf{g}_\theta$  is defined as

$$\mathbf{s} *_G \mathbf{g}_\theta = \mathbf{U}\mathbf{g}_\theta\mathbf{U}^T\mathbf{s} \quad (13)$$

Existing studies on graph spectral convolution all follow Formula (13), and the differences are the choice of the filter  $\mathbf{g}_\theta$  [72]. The  $u$ -th row of the output channel is the embedding  $\mathbf{h}_u$  of a node  $u$ .

#### B) Graph Spatial Convolution

Graph spatial convolution aggregates the information from a node's local neighborhood. Intuitively, each node sends messages based on its current embedding and updates its embedding based on the messages received from its local neighborhood. A graph spatial convolution model often stacks multiple layers, and each layer performs one iteration of message propagation. To illustrate this, we recall the definition given in GraphSAGE [25]. A layer of GraphSAGE is as follows:

$$\mathbf{h}_u^l = \sigma(\mathbf{W}^l[\mathbf{h}_u^{l-1} \parallel \mathbf{h}_{\mathcal{N}_u}^l]) \quad (14)$$

$$\mathbf{h}_{\mathcal{N}_u}^l = \text{AGG}(\{\mathbf{h}_v^{l-1}, v \in \mathcal{N}_u\}), \quad (15)$$

where  $l$  denotes the  $l$ -th layer,  $\parallel$  denotes concatenation,  $\mathcal{N}_u$  is a set of randomly selected neighbors of  $u$ , and  $\text{AGG}$  denotes an order-invariant aggregation function. GraphSAGE suggests three aggregation functions: element-wise mean, LSTM-based aggregator, and max-pooling.

### 3.2.2 GNN-Based Graph Representation Learning

Graph convolutional network (GCN) [36] is a well-known graph spectral convolution model, which is an approximation of the original graph spectral convolution defined in



Formula 13. Given a graph  $G$  and a one-channel input signal  $\mathbf{s} \in \mathbb{R}^{|V|}$ , GCN can output a  $d$ -channel signal  $\mathbf{H}^{|V| \times d}$  as follows:

$$\mathbf{H} = \mathbf{s} *_G \mathbf{g}_\theta = (\tilde{\mathbf{D}}^{-1/2} \tilde{\mathbf{A}} \tilde{\mathbf{D}}^{-1/2}) \mathbf{s} \theta, \tag{16}$$

where  $\theta$  is a  $1 \times d$  trainable parameter vector of the filter,  $\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{I}$  and  $\tilde{\mathbf{D}}$  is a diagonal matrix with  $\tilde{\mathbf{D}}[i, i] = \sum_j \tilde{\mathbf{A}}[i, j]$ . The  $u$ -th row of  $\mathbf{H}$  is the embedding of the node  $u$ ,  $\mathbf{h}_u$ . To allow a  $d'$ -channel input signal  $\mathbf{S}^{|V| \times d'}$  and output a  $d$ -channel signal  $\mathbf{H}^{|V| \times d}$ , the filter needs to take a parameter matrix  $\Theta^{d' \times d}$ . Formula 16 becomes

$$\mathbf{H} = \mathbf{S} *_G \mathbf{g}_\theta = (\tilde{\mathbf{D}}^{-1/2} \tilde{\mathbf{A}} \tilde{\mathbf{D}}^{-1/2}) \mathbf{S} \Theta. \tag{17}$$

Let  $\mathbf{s}_i$  denotes the  $i$ -th channel (i.e., column) of  $\mathbf{S}$ .  $\mathbf{h}_u$  can then be written in the following way.

$$\begin{aligned} \mathbf{h}_u &= \Theta^T \mathbf{y}, \\ \mathbf{y}[i] &= \sum_{v \in N_u \cup \{u\}} \frac{1}{\sqrt{|N_u|} \sqrt{|N_v|}} \mathbf{s}_i[v], 1 \leq i \leq d', \end{aligned} \tag{18}$$

where  $\mathbf{y}$  is a  $d'$ -dimensional column vector.

When multi-layer models are considered, Formulas 17 and 18 are written as Formulas 19 and 20, respectively, where  $l$  denotes the  $l$ -th layer.

$$\mathbf{H}^l = (\tilde{\mathbf{D}}^{-1/2} \tilde{\mathbf{A}} \tilde{\mathbf{D}}^{-1/2}) \mathbf{H}^{l-1} \Theta^l \tag{19}$$

$$\begin{aligned} \mathbf{h}_u^l &= \Theta^{lT} \mathbf{y}^l \\ \mathbf{y}^l[i] &= \sum_{v \in N_u \cup \{u\}} \frac{1}{\sqrt{|N_u|} \sqrt{|N_v|}} \mathbf{H}^{l-1}[v, i] \\ &= \sum_{v \in N_u \cup \{u\}} \frac{1}{\sqrt{|N_u|} \sqrt{|N_v|}} \mathbf{h}_v^{l-1}[i] \end{aligned} \tag{20}$$

From Formula 20, we can observe that GCN aggregates weighted information from a node's neighbors. In particular, for a node  $u$  and a neighbor  $v$  of  $u$ , the information from  $v$  is weighted by their degrees, i.e.,  $1/\sqrt{|N_u| |N_v|}$ . Graph attention network (GAT) [66] argues that the fixed weight approach of GCN may not always be optimal. Therefore, GAT introduces the attention mechanism to graph convolution. A learnable weight function  $\alpha(\cdot, \cdot)$  is proposed, where  $\alpha(u, v)$  denotes the attention weight of  $u$  over its neighbor  $v$ . Specifically, the convolution layer of GAT is as follows.

$$\mathbf{h}_u^l = \sigma \left( \sum_{v \in N(u)} \alpha^l(u, v) \mathbf{W}^l \mathbf{h}_v^{l-1} \right) \tag{21}$$

$$\alpha^l(u, v) = \frac{\exp(\text{LeakyReLU}(\mathbf{a}^{lT} [\mathbf{W}^l \mathbf{h}_u^{l-1} \parallel \mathbf{W}^l \mathbf{h}_v^{l-1}]))}{\sum_{v' \in N(u)} \exp(\text{LeakyReLU}(\mathbf{a}^{lT} [\mathbf{W}^l \mathbf{h}_u^{l-1} \parallel \mathbf{W}^l \mathbf{h}_{v'}^{l-1}]))}, \tag{22}$$

where  $\parallel$  denotes concatenation,  $\mathbf{a}^l$  and  $\mathbf{W}^l$  are the trainable vector and matrix of parameters, respectively.

The attention mechanism enhances models' capacity, and hence, GAT can perform better than GCN in some downstream tasks (e.g., node classification). However, when  $L$  layers are stacked, the  $L$ -hop neighbors of a node are needed to be computed. If the graph  $G$  is dense or a power-law graph, there may exist some nodes that can access almost all nodes in  $G$ , even for a small value of  $L$ . The time cost can be unaffordable.

To optimize efficiency, Hamilton et al. [25] propose a sampling-based method (GraphSAGE). GraphSAGE randomly samples  $k$  neighbors in each layer. Therefore, a model having  $L$  layers only needs to expand  $O(k^L)$  neighbors. Huang et al. [30] further improve the sampling process with an adaptive sampling method. The adaptive sampling in [30] samples

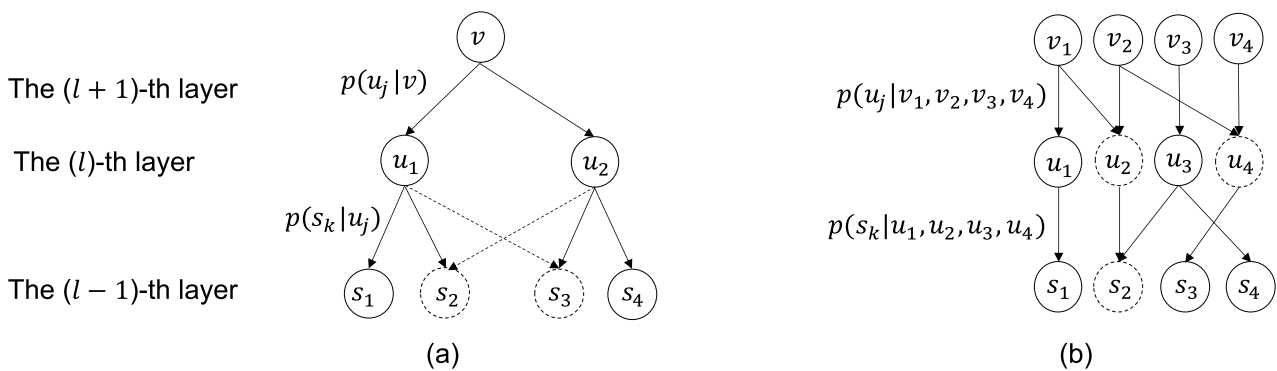


Fig. 3 Adaptive sampling of ASGCN [30]: **a** the node-wise sampling and **b** the layer-wise sampling. In the node-wise sampling, each node in a layer samples its neighbors in the next layer independently. In particular, a node  $v$  in the  $l + 1$ -th layer samples its neighbors in the

$l$ -th layer by  $p(u_j|v)$ . In contrast, all nodes in a layer jointly sample the neighbors in the next layer.  $u_j$  is sampled based on  $p(u_j|v_1, v_2, \dots, v_4)$ . The layer-wise sampling is more efficient than the node-wise sampling

the neighbors based on the embedding of  $u$ , as illustrated in Fig. 3a. The efficiency is further improved by layer-wise sampling, as shown in Fig. 3b. These sampling techniques are experimentally verified effective regarding the classification accuracy.

Yang et al. [74] combine the ideas of attention and sampling and propose the shortest path attention method (SPAGAN). The shortest path attention of SPAGAN has two levels, as shown in Fig. 4. The first level is length-specific, which embeds the shortest paths of the same length  $c$  to a vector  $\mathbf{h}_u^c$ . The second level aggregates  $\mathbf{h}_u^c$  of different values of  $c$  to get the embedding  $\mathbf{h}_u$  of  $u$ .

More specifically, let  $P_u^c$  be the set of shortest paths starting from  $u$  of the length  $c$  and  $p_{u,v}$  be a shortest path from node  $u$  to node  $v$ .  $\mathbf{h}_u^c$  is computed as follows.

$$\mathbf{h}_u^c = \sum_{p_{u,v} \in P_u^c} \alpha_{u,v} \phi(p_{u,v}),$$

where  $\alpha_{u,v}$  is the attention weight and  $\phi(p_{u,v})$  is a mean pooling that computes the average of the embeddings of the nodes in  $p_{u,v}$ .

$$\alpha_{u,v} = \frac{\exp(\sigma(\mathbf{a}_1[(\mathbf{W}\mathbf{h}_u) \parallel \phi(p_{u,v})]))}{\sum_{p_{u,v'} \in P_u^c} \exp(\sigma(\mathbf{a}_1[(\mathbf{W}\mathbf{h}_u) \parallel \phi(p_{u,v'})]))},$$

where  $\mathbf{a}_1$  and  $\mathbf{W}$  are trainable parameters shared by all nodes, and  $\parallel$  is concatenation. The second level aggregates the paths with different lengths as follows.

$$\mathbf{h}_u = \sigma\left(\sum_{c=2}^C \beta_c \mathbf{h}_u^c\right),$$

where  $C$  is a hyperparameter of the path length limit and  $\beta_c$  is the attention weight.

$$\beta_c = \frac{\exp(\sigma(\mathbf{a}_2[(\mathbf{W}\mathbf{h}_u) \parallel \mathbf{h}_u^c]))}{\sum_{c'=2}^C \exp(\sigma(\mathbf{a}_2[(\mathbf{W}\mathbf{h}_u) \parallel \mathbf{h}_u^{c'}]))},$$

where  $\mathbf{a}_2$  is a trainable parameter vector.

### 3.2.3 AutoEncoder-Based Graph Representation Learning

For the AutoEncoder used in the end-to-end learning, the embeddings of the graph are computed by the encoder. The decoder outputs the probabilities of nodes/edges belonging to the solutions of the CO problems. In recent works, the encoder mainly uses RNN and attention-based model, and the decoder mainly uses MLP, RNN and attention-based model. The encoder corresponds to the first stage of the ML-based CO methods and the decoder corresponds to the second stage (see Fig. 1). In this subsection, we mainly focus on the encoder. The details of the decoders will be discussed in Section 4.

The pointer network (Ptr-Net) proposed by Vinyals et al. [67] is a seminal work of using AutoEncoder to solve the TSP problem. The encoder of Ptr-Net is an RNN taking the nodes of the graph  $G$  as input and outputting an embedding of  $G$ , where the order of the nodes is randomly chosen. Experiments of Ptr-Net observe that the order of input nodes has affects on the quality of the TSP tour found. Therefore, the decoder of Ptr-Net introduces an attention mechanism that can assign weights to the input nodes and ignore the order of them.

Kool et al. [37] use AutoEncoder to sequentially output a TSP tour of a graph  $G$ . The encoder stacks  $L$  self-attention layers. Each layer is defined as follows.

$$\hat{\mathbf{h}}_{v_i} = BN^l(\mathbf{h}_{v_i}^{l-1} + MHA^l(\mathbf{h}_{v_1}^{l-1}, \dots, \mathbf{h}_{v_n}^{l-1}))$$

$$\mathbf{h}_{v_i}^l = BN^l(\hat{\mathbf{h}}_{v_i} + MLP^l(\hat{\mathbf{h}}_{v_i})),$$

where  $\mathbf{h}_i$  denotes the embedding vector of the node  $v_i$ ,  $l$  means the  $l$ -th layer,  $MHA$  denotes the multi-head attention and  $BN$  denotes the batch normalization. The embedding of  $G$   $\mathbf{h}_G = \frac{1}{n} \sum_i \mathbf{h}_{v_i}^L$  and the embedding of each node  $\mathbf{h}_{v_i}^L$  are input to the decoder.

### 3.2.4 Discussions

Most graph neural network-based methods adopt the message propagation framework. Each node iteratively

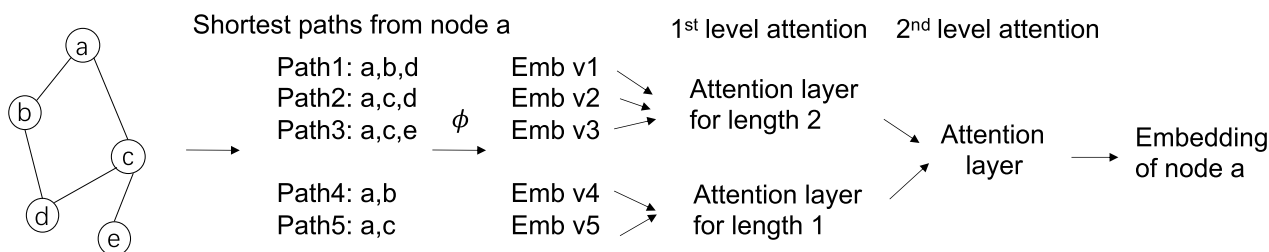


Fig. 4 The two-level convolution of SPAGAN [74]

aggregates the message from neighbors. The structure information of  $k$ -hops of a node can be captured by  $k$  iterations of message aggregation. GNN does not require any node order and can support permutation invariance of CO problems. AutoEncoder-based methods are often used in solving the CO problems having sequential characteristics, *e.g.*, TSP. Sequence model is often used as the encoder to compute the embeddings of the graphs. Attention mechanism is used to support permutation invariance.

End-to-end learning methods learn the embeddings of graph as an intermediate step in solving the CO problem. The embeddings of the graph learned are more specific for the CO problem being solved and are expected to lead to better solutions of the CO problem.

A disadvantage of the GNN-based method is that the GNN is often shallow, due to the over-smooth problem. The attention-based encoder can alleviate this problem, where the encoder with self-attention layers and skip connections can be potentially deeper. However, the time complexity of such encoder on large graphs will be a bottleneck.

For the GNN-based method, the current trend is to use anisotropy GNN (*e.g.* GAT [66]), which can differentiate the information propagated from different neighbors. For AutoEncoder-based method, more recent studies are integrating the attention mechanism with the sequence model to increase the capacity of the model and encode inductive biases.

## 4 Graph Learning-Based Combinatorial Optimization Methods

In this section, we review the works that solve CO problems using graph learning. We review the whole learning procedure in solving a CO problem. For the two stages of the learning procedure, we pay more attention to the second stage, as the first stage has been thoroughly reviewed in the previous section. We will brief the first stage of the ML-based CO methods for the convenience of presentation.

Recent works can be classified into two categories. The first category is the non-autoregressive method which predicts the solution of a CO problem in one shot. The non-autoregressive method predicts a matrix that denotes the probability of each node/edge being a part of a solution. The solution of the CO problem can be found by search heuristics such as beam search. The second category is the autoregressive method, which constructs a solution by iteratively extending a partial solution to obtain a solution of the CO problem. Table 2 lists the selected graph learning-based CO methods.

Section 4.1 summarizes the recent non-autoregressive methods for traveling salesman problem (TSP), graph partition, graph similarity, minimum vertex cover (MVC), graph coloring, maximum independent set, graph matching and graph isomorphism. Section 4.2 presents the recent autoregressive methods for TSP, graph matching, graph alignment, MVC and maximum common subgraph.

**Table 2** Summary of selected CO methods using graph embedding

Method	CO problem	Model
ConvNet [34]	TSP	GNN, non-autoregressive
DTSPGNN [57]	TSP	GNN, non-autoregressive
CPNGNN [59]	MDS, MM, MVC	GNN, non-autoregressive
GAP [49]	Graph partition	GNN, non-autoregressive
GMN [41]	GED	GNN, non-autoregressive
SimGNN [2]	GED	GNN, non-autoregressive
GRAPHSIM [3]	GED	GNN, non-autoregressive
GNNGC [40]	GColor	GNN, non-autoregressive
SiameseGNN [51]	Graph matching, TSP	GNN, non-autoregressive
PCAGM [69]	Graph matching	GNN, non-autoregressive
IsoNN [44]	Graph Iso.	AutoEncoder, non-autoregressive
GNNTS [42]	MIS, MVC, MC	GNN, non-autoregressive
Ptr-Net [67]	TSP	AutoEncoder, autoregressive
LSTMGMatching [46]	Graph matching	AutoEncoder, autoregressive
S2V-DQN [14]	MVC, MaxCut, TSP	GNN, autoregressive
CombOptZero [1]	MVC, MaxCut, MC	GNN, autoregressive
RLMCS [4]	MCS	GNN, autoregressive
CENALP [19]	Graph alignment	SkipGram, autoregressive
TSPImprove [71]	TSP	AutoEncoder, autoregressive
AM [37]	TSP	AutoEncoder, autoregressive

### 4.1 Non-autoregressive CO Methods

Most works in this category use classification techniques to predict the class label of the nodes in the input graph. For a graph  $G$ , the prediction result is a  $|V| \times K$  matrix  $\mathbf{Y}$ , where  $K$  is the number of classes. The  $u$ -th row  $\mathbf{y}_u$  of  $\mathbf{Y}$  is the prediction result for the node  $u$ , where  $\mathbf{y}_u[i]$  is the probability that  $u$  is of the  $i$ -th class, for  $1 \leq i \leq K$ . For example, for the minimum vertex cover (MVC) problem, the classification is binary (i.e.,  $K = 2$ ), and  $\{u | \mathbf{y}_u[1] > \mathbf{y}_u[0]\}$  is the predicted solution. For the graph partition problem,  $K$  is the number of parts, and a node  $u$  is classified to the part with the largest predicted probability. There are some works that predict a score for the input graphs. For example, for the graph similarity problem, the similarity score between two graphs is predicted.

#### A. Travelling Salesman Problem

Joshi et al. [34] propose a GNN-based model (ConvNet) to solve the TSP problem on Euclidean graph. The graph convolution layer of ConvNet is as follows.

$$\mathbf{h}_i^{l+1} = \mathbf{h}_i^l + ReLU\left(BN\left(\mathbf{W}_1^l \mathbf{h}_i^l + \sum_{j \in N_i} \eta_{ij}^l \odot \mathbf{W}_2^l \mathbf{h}_j^l\right)\right)$$

$$\eta_{ij}^l = \frac{\sigma(\mathbf{e}_{ij}^l)}{\sum_{j' \in N_i} \sigma(\mathbf{e}_{ij'}^l) + \epsilon}$$

$$\mathbf{e}_{ij}^l = \mathbf{e}_{ij}^l + ReLU(BN(\mathbf{W}_3^l \mathbf{e}_{ij}^l + \mathbf{W}_4^l \mathbf{h}_i^l + \mathbf{W}_5^l \mathbf{h}_j^l)),$$

where  $BN$  stands for batch normalization,  $\odot$  denotes element-wise product,  $\eta$  is attention weight,  $\epsilon$  is a small value,  $\mathbf{W}_1, \mathbf{W}_2$  and  $\mathbf{W}_3$  are trainable parameters.

The embeddings of the edges outputted by the  $l$ -th layer of ConvNet are fed into a multilayer perceptron (MLP) to predict  $p_{ij}$  the probability of the edge  $e_{ij}$  belongs to the solution of TSP. The cross entropy with the ground-truth TSP tour is used as the loss. The experiments of ConvNet show that ConvNet outperforms recent autoregressive methods but falls short of standard Operations Research solvers.

Prates et al. [57] use GNN to solve the decision version of TSP, which is to decide if a given graph admits a Hamiltonian route with a cost no greater than a given threshold  $C$ . Since the weights of edges are closely related to the cost of a route, Prates et al. compute edge embedding in the graph convolution. Specifically, given a graph  $G = (V, E)$ , an auxiliary

bipartite graph  $G' = (V \cup V', E')$  is constructed, where for each edge  $(u, v)$  in  $G$ ,  $G'$  has a node  $n_{u,v}$  in  $V'$  and edges  $(n_{u,v}, u)$  and  $(n_{u,v}, v)$  are added to  $E'$ . The embeddings of the nodes and edges of  $G$  can be computed by a GNN on the auxiliary graph  $G'$ . Finally, the embeddings of the edges of  $G$  are fed into an MLP to make a binary classification. If the class label of  $G$  is predicted to be 1,  $G$  has a Hamiltonian route with a cost no greater than  $C$ ; otherwise,  $G$  has no such route.

#### B. Graph Partition

Nazi et al. [49] propose GAP as a method for computing a balanced partition of a graph. GAP is composed of a graph embedding module, which uses a GNN model to determine the embedding of the input graph, and a graph partition module, which uses an MLP to predict the partition of nodes. The architecture of GAP is illustrated in Fig. 5. The normalized cut size and the balancedness of the partition are used as the loss. GAP trained on a small graph can be generalized at the inference time on unseen graphs of larger size.

Specifically, suppose  $G = (V, E, \mathbf{X})$  is to be partitioned to  $K$  disjoint parts and  $V_1, V_2, \dots, V_K$  denote the sets of nodes of the parts, respectively. A GNN first computes the embeddings of the nodes in  $G$ . Then, the MLP uses the node embeddings to predict the partition probability  $\mathbf{Y}^{|V| \times K}$  for the nodes, where  $\mathbf{Y}[u, i]$  is the probability that node  $u$  is partitioned to  $V_i$ . Finally, each node can be partitioned to the partition of the largest probability.

The loss of GAP has two components. The first component is to minimize the normalized cut size of the partition:

$$\sum_{i=1}^K \frac{cut(V_i, \bar{V}_i)}{vol(V_i)},$$

where  $\bar{V}_i$  denotes the nodes not in  $V_i$ ,  $cut(V_i, \bar{V}_i)$  denotes the number of edges crossing  $V_i$  and  $\bar{V}_i$ , and  $vol(V_i)$  denotes the total degree of the nodes in  $V_i$ . The second component is to minimize the distance from the balanced partition:

$$\sum_{i=1}^K \sum_{u \in G} \left( \mathbf{Y}[u, i] - \frac{|V|}{K} \right)^2,$$

where  $\frac{|V|}{K}$  is the part size of the balanced partition. The objective function of GAP is as follows.

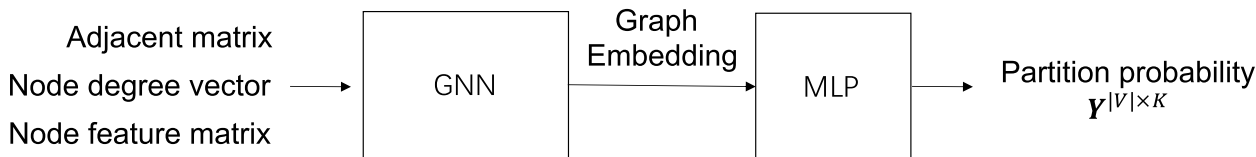


Fig. 5 Overview of GAP [49]

$$\min \sum_{i=1}^K \frac{cut(V_i, \bar{V}_i)}{vol(V_i)} + \sum_{i=1}^K \sum_{u \in G} \left( Y[u, i] - \frac{|V_i|}{K} \right)^2$$

C. Graph Similarity

Bai et al. [2] propose SimGNN as a method for predicting the similarity between two graphs. SimGNN combines two strategies for predicting the similarity between two graphs  $G_1$  and  $G_2$ . The first strategy compares  $G_1$  and  $G_2$  by comparing their global summaries  $\mathbf{h}_{G_1}$  and  $\mathbf{h}_{G_2}$ . The second strategy uses the pair-wise node comparison to provide a fine-grained information as a supplement to the global summaries  $\mathbf{h}_{G_1}$  and  $\mathbf{h}_{G_2}$ . The architecture of SimGNN is shown in Fig. 6.

As shown in Fig. 6, SimGNN first computes the node embeddings of the two input graphs  $G_1$  and  $G_2$  using GCN. For the first strategy, SimGNN computes  $\mathbf{h}_{G_1}$  and  $\mathbf{h}_{G_2}$  from the node embeddings by means of an attention mechanism that can adaptively emphasize the important nodes with respect to a specific similarity metric. Then,  $\mathbf{h}_{G_1}$  and  $\mathbf{h}_{G_2}$  are input to a neural tensor network (NTN) to compute a similarity score vector for  $G_1$  and  $G_2$ .

The attention mechanism to compute  $\mathbf{h}_G$  is defined as follows. For a graph  $G$ , SimGNN introduces a context vector  $\mathbf{c} = \tanh(\mathbf{W} \sum_{u \in G} \mathbf{h}_u)$  to encode the global information of  $G$ .  $\mathbf{c}$  is adaptive to the given similarity metric via  $\mathbf{W}$ . Intuitively, nodes that are close to the global context should receive more attention. Therefore, the attention weight  $\alpha_u$  of a node  $u$  is defined based on the inner product of  $\mathbf{c}$  and  $\mathbf{h}_u$ .  $\alpha_u = \sigma(\mathbf{c}^T \mathbf{h}_u)$ , where  $\sigma$  is the sigmoid function. The embedding of  $G$ ,  $\mathbf{h}_G$ , is computed as  $\mathbf{h}_G = \sum_{u \in G} \alpha_u \mathbf{h}_u$ .

For the second strategy, SimGNN constructs a pair-wise node similarity matrix  $M$  by computing the inner product of  $\mathbf{h}_u$  and  $\mathbf{h}_v$  for each  $u \in G_1, v \in G_2$ . SimGNN

uses a histogram of  $M$  to summarize the pair-wise node similarity.

Finally, the similarity score vector outputted by NTN and the histogram are input to a fully connected neural network to predict the similarity between  $G_1$  and  $G_2$ . The mean squared error between the predicted similarity with the ground truth is used as the loss of SimGNN. In the follow-up work GRAPHSIM [3], a CNN-based method is used to replace the histogram of SimGNN.

Li et al. [41] propose the graph matching network (GMN) to solve the graph similarity problem. Instead of embedding each graph independently, GMN embeds two graphs  $G_1$  and  $G_2$  jointly by examining the matching between them. The matching used in GMN is soft matching, which means that a node of  $G_1$  can match to all nodes of  $G_2$  yet with different strengths. The embedding of  $G_1$  can change based on the other graph it is compared against. At inference time, GMN can predict if the distance between two graphs is smaller than a given threshold  $\gamma$ .

Given two graphs  $G_1 = (V(G_1), E(G_1))$  and  $G_2 = (V(G_2), E(G_2))$ , the  $l$ -th convolution layer of GMN is defined as below.

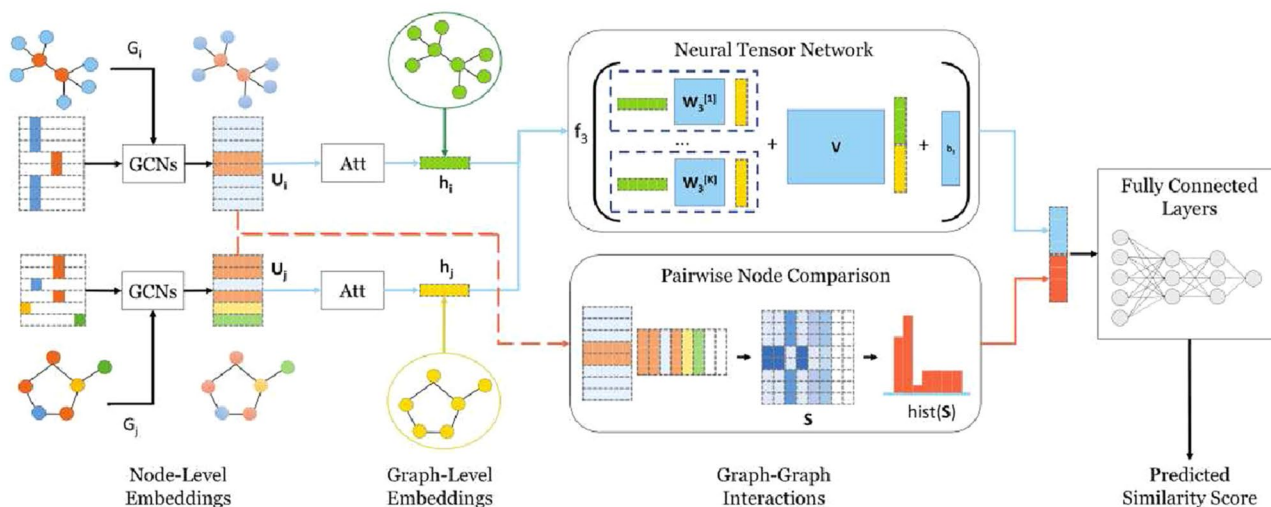


Fig. 6 Overview of SimGNN [2]. The blue solid line illustrates the first strategy of comparing  $G_1$  and  $G_2$  using their global summaries  $\mathbf{h}_{G_1}$  and  $\mathbf{h}_{G_2}$ . The orange dashed line indicates the second strategy of the find-grained pair-wise node comparison

$$\begin{aligned}
 \mathbf{m}_{j \rightarrow i} &= MLP(\mathbf{h}_i^l, \mathbf{h}_j^l), \forall (i, j) \in E(G_1) \\
 \mathbf{m}_{j' \rightarrow i'} &= MLP(\mathbf{h}_{i'}^l, \mathbf{h}_{j'}^l), \forall (i', j') \in E(G_2) \\
 \boldsymbol{\mu}_{j' \rightarrow i} &= f_{match}(\mathbf{h}_i^l, \mathbf{h}_{j'}^l), \forall i \in V(G_1), j' \in V(G_2) \\
 \boldsymbol{\mu}_{i \rightarrow j'} &= f_{match}(\mathbf{h}_i^l, \mathbf{h}_{j'}^l), \forall i \in V(G_1), j' \in V(G_2) \\
 \mathbf{h}_i^{l+1} &= MLP\left(\mathbf{h}_i^l, \sum_{j \in G_1} \mathbf{m}_{j \rightarrow i}, \sum_{j' \in G_2} \boldsymbol{\mu}_{j' \rightarrow i}\right) \\
 \mathbf{h}_{j'}^{l+1} &= MLP\left(\mathbf{h}_{j'}^l, \sum_{i' \in G_2} \mathbf{m}_{i' \rightarrow j'}, \sum_{i \in G_1} \boldsymbol{\mu}_{i \rightarrow j'}\right),
 \end{aligned} \tag{23}$$

where  $\mathbf{m}$  denotes the message aggregation of a node from its neighbors in the same graph,  $\boldsymbol{\mu}$  is the cross-graph matching vector that measures the difference between a node in a graph and all the nodes in the other graph, and  $f_{match}$  can be defined by the following attention-based method.

$$\begin{aligned}
 \alpha_{j' \rightarrow i} &= \alpha_{j' \rightarrow i}(\mathbf{h}_i^l - \mathbf{h}_{j'}^l), \forall i \in V(G_1), j' \in V(G_2) \\
 &= \frac{\exp(\text{dist}(\mathbf{h}_i^l, \mathbf{h}_{j'}^l))}{\sum_{v' \in G_2} \exp(\text{dist}(\mathbf{h}_i^l, \mathbf{h}_{v'}^l))} \\
 \alpha_{i \rightarrow j'} &= \alpha_{i \rightarrow j'}(\mathbf{h}_i^l - \mathbf{h}_{j'}^l), \forall i \in V(G_1), j' \in V(G_2) \\
 &= \frac{\exp(\text{dist}(\mathbf{h}_i^l, \mathbf{h}_{j'}^l))}{\sum_{v \in G_1} \exp(\text{dist}(\mathbf{h}_v^l, \mathbf{h}_{j'}^l))},
 \end{aligned}$$

where  $\text{dist}$  is the Euclidean distance.

Suppose GMN stacks  $L$  layers. The embedding of a graph  $G$  is computed as below.

$$\mathbf{h}_G = MLP(\{\mathbf{h}_i^L\}_{i \in G}), \tag{24}$$

where  $\mathbf{h}_i^L$  is the embedding of node  $i$  outputted by the last convolution layer.

The objective function of GMN is to minimize the margin-based pair-wise loss  $\mathcal{L} = \max\{0, \gamma - t \times (1 - \text{dist}(G_1, G_2))\}$ , where  $\gamma > 0$  is the given margin threshold,  $\text{dist}(G_1, G_2) = \|\mathbf{h}_{G_1} - \mathbf{h}_{G_2}\|_2$  is the Euclidean distance, and  $t$  is the ground truth of the

similarity relationship between  $G_1$  and  $G_2$ , i.e., if  $G_1$  and  $G_2$  are similar,  $t = 1$ ; otherwise,  $t = -1$ .

#### D. Minimum Vertex Cover

Sato et al. [59], from a theoretical perspective, study the power of GNNs in learning approximation algorithms for the minimum vertex cover (MVC) problem. They prove that no existing GNN can compute a  $(2 - \epsilon)$ -approximation for MVC, where  $\epsilon > 0$  is any real number and  $\Delta$  is the maximum node degree. Moreover, Sato et al. propose a more powerful consistent port numbering GNN (CPNGNN), which can return a 2-approximation for MVC. The authors theoretically prove that there exist a set of parameters of CPNGNN that can be used to find an optimal solution for MVC. However, the authors do not propose a method for finding this set of parameters.

CPNGNN is designed based on graph port numbering. Given a graph  $G$ , the ports of a node  $u$  are pairs  $(u, i)$ ,  $1 \leq i \leq |N_u|$ , where  $i$  is the port number. A port numbering is a function  $p$  such that for any edge  $(u_1, u_2) \in G$ , there exists a port  $(u_1, i)$  of  $u_1$  and a port  $(u_2, j)$  of  $u_2$  satisfying  $p(u_1, i) = (u_2, j)$ . Intuitively,  $u_1$  can send messages from the  $i$ th port of  $u_1$  to the  $j$ th port of  $u_2$ . If  $p(u_1, i) = (u_2, j)$ ,  $u_1$  is denoted by  $p_{tail}(u_2, j)$  and  $i$  is denoted by  $p_n(u_2, j)$ . An example of port numbering is shown in Fig. 7.

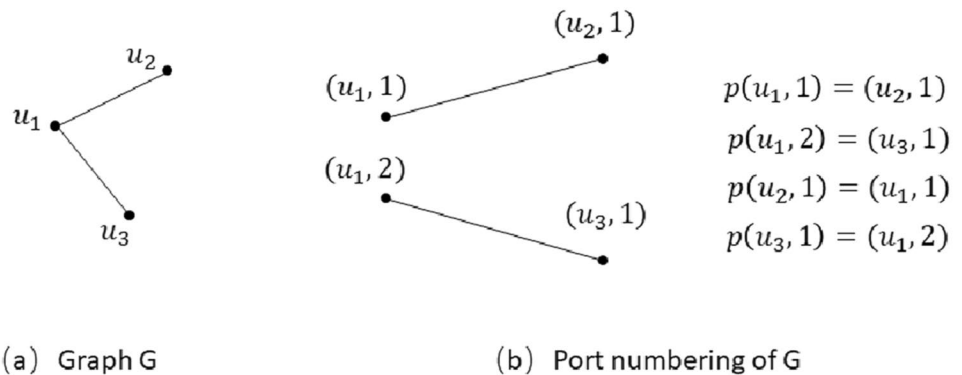
CPNGNN stacks  $L$  convolution layers, and the  $l$ -th layer is defined as follows.

$$\begin{aligned}
 \mathbf{h}_u^l &= ReLU(\mathbf{W}^l[\mathbf{h}_u^{l-1} || \mathbf{x}_{u,1}^{l-1} || \mathbf{x}_{u,2}^{l-1} || \dots || \mathbf{x}_{u,|N_u|}^{l-1}]) \\
 \mathbf{x}_{u,i}^{l-1} &= \mathbf{h}_{p_{tail}(u,i)}^{l-1} || p_n(u,i),
 \end{aligned} \tag{25}$$

where  $\mathbf{W}^l$  is the trainable parameter matrix and  $||$  is concatenation.

Let  $\mathbf{h}_u^L$  denotes the embedding of  $u$  outputted by the last layer of CPNGNN. An MLP takes  $\mathbf{h}_u^L$  as input and outputs the prediction  $\mathbf{y}_u$  for  $u$ , where  $\mathbf{y}_u[1]$  and  $\mathbf{y}_u[0]$  are the probabilities that  $u$  is in an MVC or not, respectively. Then, the nodes  $\{u | \mathbf{y}_u[1] > \mathbf{y}_u[0]\}$  are outputted as an MVC of  $G$ . The approximation ratio of CPNGNN is 2 for MVC. CPNGNN can also solve the minimum dominating set (MDS) problem

Fig. 7 An example of port numbering



and the maximum matching (MM) problem with the approximation ratio  $\frac{\Delta+1}{2}$ .

*E. Graph Coloring*

Lemos et al. [40] propose a graph recurrent neural network to predict if a graph is  $k$ -colorable. Each node  $v$  has an embedding vector  $\mathbf{h}_v$ , and each color  $c$  also has an embedding vector  $\mathbf{h}_c$ . Let  $\mathbf{A}$  denotes the adjacent matrix of the graph  $G$  and  $\mathbf{M}$  denotes the color assignment matrix, where each row of  $\mathbf{M}$  is a node of  $G$  and each column of  $\mathbf{M}$  is a color.  $\mathbf{M}[v, c] = 1$  means the node  $v$  is assigned the color  $c$ . The embeddings of the  $l + 1$ -th iteration  $\mathbf{h}_v^{l+1}$  and  $\mathbf{h}_c^{l+1}$  are computed as follows.

$$\mathbf{h}_v^{l+1}, \mathbf{h}_{nhid}^{l+1} = RNN_1(\mathbf{h}_{nhid}^l, \mathbf{A} \times \mathbf{h}_v^l, \mathbf{M} \times MLP_1(\mathbf{h}_c^l))$$

$$\mathbf{h}_c^{l+1}, \mathbf{h}_{chid}^{l+1} = RNN_2(\mathbf{h}_{chid}^l, \mathbf{M} \times MLP_2(\mathbf{h}_v^l))$$

The embeddings of nodes are fed into an MLP to predict the probability if  $G$  is  $k$ -colorable and the loss is the binary cross entropy between the prediction and the ground-truth. Experiments of [40] show that the proposed techniques outperform the existing heuristic algorithm Tabucol and the greedy algorithm.

*F. Graph Matching*

Nowak et al. [51] study the GNN-based model for the quadratic assignment problem, that can be used to address the graph matching problem. A siamese GNN is constructed to compute the embeddings of two graphs. Let  $\mathbf{Y}$  be the product of the embeddings of the nodes of the two graphs. A stochastic matrix is computed from  $\mathbf{Y}$  by taking the softmax along each row (or column). The cross entropy between the

stochastic matrix and the ground-truth node mapping is the loss. The proposed model can also be used to solve the TSP problem, as TSP can be formulated as a quadratic assignment problem.

Wang et al. [69] propose a GNN-based model to predict the matching of two graphs. Given two graphs  $G_1$  and  $G_2$ , it first uses GNN to compute the embeddings of the nodes of the two graphs. Then, the embeddings are fed to a Sinkhorn layer to obtain a doubly-stochastic matrix. The cross entropy with the ground-truth node mapping is used as the loss. The idea of the Sinkhorn layer is that given a non-negative matrix, iteratively normalize each row and each column of the matrix until the sum of each row and the sum of each column equal to 1, respectively. Experiments of [69] show that the proposed model outperforms the existing learning-based graph matching methods.

*G. Graph Isomorphism*

Meng and Zhang [44] propose an isomorphic neural network (IsoNN) for learning graph embedding. The encoder has three layers: a convolution layer, a min-pooling layer, and a softmax layer. The encoder is shown in Fig. 8. The decoder is an MLP to predict the binary class of  $G$ , and the loss is the cross entropy between the prediction and the ground truth.

Specifically, the encoder of IsoNN is designed as follows. Given a set of motifs, the convolution layer of the encoder extracts a set of isomorphism features from  $G$  for each motif. Suppose  $\mathbf{K}_i$  is the adjacent matrix of the  $i$ -th motif that has  $k$  nodes. The L2-norm between  $\mathbf{K}_i$  and a  $k$  by  $k$  submatrix  $\mathbf{A}_{x,y,k}$  of the adjacent matrix  $\mathbf{A}$  of  $G$  is an isomorphism feature extracted by  $\mathbf{K}_i$  with respect to  $\mathbf{A}_{x,y,k}$ , where  $x$  and  $y$  denote

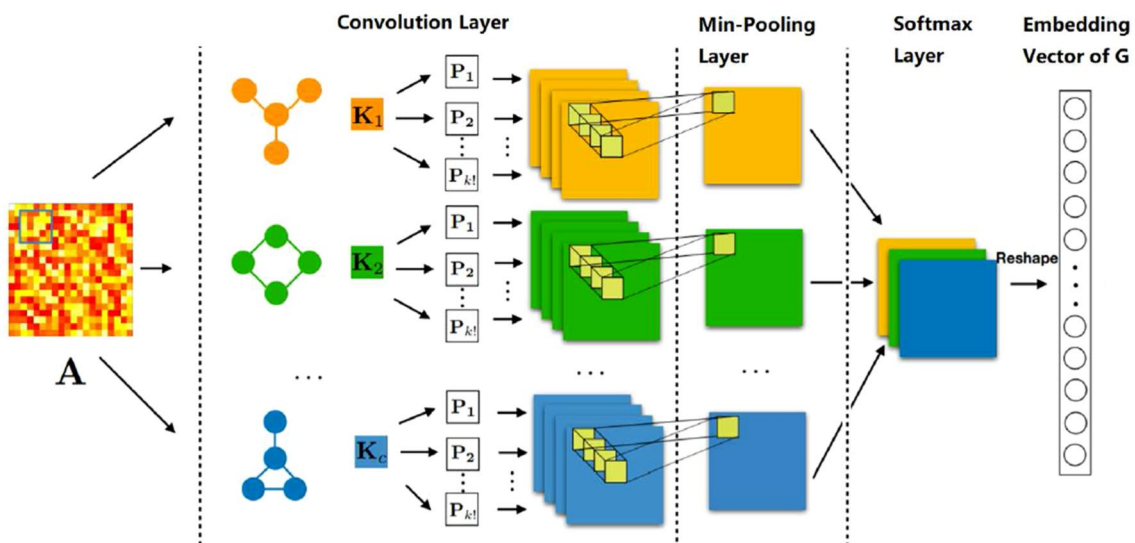


Fig. 8 Overview of IsoNN [44]

the top-left corner of the submatrix in  $\mathbf{A}$ . IsoNN examines  $k!$  permutations of  $\mathbf{K}_i$  and extracts  $k!$  isomorphism features for  $\mathbf{A}_{x,y,k}$ . The smallest one is regarded as the optimal isomorphism feature extracted by  $\mathbf{K}_i$  for  $\mathbf{A}_{x,y,k}$ , which is computed by the min-pooling layer. Since the optimal isomorphism features for  $\mathbf{A}_{x,y,k}$  extracted by different motifs can have different scales, the softmax layer is used to normalize them. Finally, the normalized isomorphism features extracted by all motifs for all values of  $x$  and  $y$  are concatenated as the embedding of  $G$ .

#### H. Maximum Independent Set

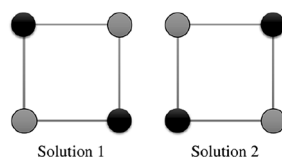
Li et al. [42] propose a GNNTS model that combines GNN and heuristic search to compute the maximum independent set (MIS) of a graph. GNNTS trains a GCN  $f$  using a set of training graphs, where the MISs of a graph can be used as the ground truth labels of the graph. For a graph  $G = (V, E)$ , the prediction result of  $f$  is a  $|V| \times 2$  matrix  $\mathbf{Y}$ , where  $\mathbf{Y}[\cdot, 1]$  and  $\mathbf{Y}[\cdot, 0]$  are the probabilities of the nodes being in or not in an MIS of  $G$ , respectively.

The basic idea of GNNTS is to use  $f$  as the heuristic function within a greedy search procedure. Specifically, in each iteration, the nodes of  $G$  are sorted by  $\mathbf{Y}[\cdot, 1]$ . The greedy algorithm picks the node  $u$  with the largest value in  $\mathbf{Y}[\cdot, 1]$ , marks  $u$  as 1, and adds  $u$  to a partial solution  $U$ . All neighbors of  $u$  are marked as 0.  $u$  and its neighbors are removed from  $G$ , and the remaining graph is input to  $f$  for the next iteration. Once all nodes in  $G$  are marked,  $U$  is returned as the MIS of  $G$ .

The basic method described above has the disadvantage that it cannot support the case in which  $G$  has multiple solutions. For the example shown in Fig. 9, the square graph of four nodes has two MISs and the basic method predicts that each node has a probability 0.5 of belonging to an MIS, which is not useful.

To address this disadvantage, the GNN  $f$  is extended to output multiple prediction results, i.e.,  $f(G) = \{f^1(G), f^2(G), \dots, f^m(G)\}$ , where  $f^i(G)$  is a  $|V| \times 2$  matrix  $\mathbf{Y}^i$ ,  $1 \leq i \leq m$ , and  $m$  is a hyperparameter. Then, the GNN  $f$  is used in a tree search procedure. Specifically, GNNTS maintains a tree of partial solutions, where each leaf is a partial solution to be extended. At each step, GNNTS randomly picks a leaf  $n_{leaf}$  from the search tree and uses  $f$  to output  $m$  prediction results  $\mathbf{Y}^1, \mathbf{Y}^2, \dots, \mathbf{Y}^m$ . Then, for each  $\mathbf{Y}^i$ , GNNTS uses the basic method to compute an extension of  $n_{leaf}$ . The  $m$  newly obtained partial solutions are inserted to the search tree as the children of  $n_{leaf}$ . If a leaf of the search tree cannot be extended anymore, the leaf is a maximal independent set.

**Fig. 9** Illustration of the two MISs of the square graph [42]



The largest computed maximal independent set is outputted. GNNTS can also solve the minimum vertex cover (MVC) and maximal clique (MC) problems by reducing to MIS.

#### 4.1.1 Discussions

Non-autoregressive methods output a solution in one shot. The advantage is that the inference of non-autoregressive methods is faster than autoregressive methods [33]. However, the probability of a node/edge being a part of a solution does not depend on that of other nodes/edges. There is an opportunity that non-autoregressive methods are not able to outperform autoregressive methods for solving the CO problems having sequential characteristics, such as TSP. Therefore, there are many recent works studying autoregressive methods.

#### 4.2 Autoregressive CO Methods

Autoregressive methods iteratively extend a partial solution. In each iteration, a node/edge is added to the partial solution. Most existing works use sequence model-based methods or reinforcement learning-based methods to iteratively extend the partial solution.

##### A. Sequence Model-Based Methods

The pointer network (Ptr-Net) proposed by Vinyals et al. [67] is a seminal work in this category. It uses an RNN-based AutoEncoder to solve the travelling salesman problem (TSP) on a Euclidian graph. The encoder of Ptr-Net is an RNN taking the nodes of the graph  $G$  as input and outputting an embedding of  $G$ , where the order of the nodes is randomly chosen. The decoder of Ptr-Net is also an RNN. In each time step, the decoder computes an attention over the input nodes, and selects the input node that has the largest attention weight as output.

Specifically, given a graph  $G$ , suppose the nodes of  $G$  are sequentially input as  $v_1, v_2, \dots, v_{|V|}$  to the encoder, and the decoder sequentially outputs  $v_{j_1}, v_{j_2}, \dots, v_{j_{|V|}}$ . Let  $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_{|V|}$  and  $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_{|V|}$  denote the sequences of the hidden states of the encoder and the decoder, respectively. For the  $k$ -th time step of the decoder, the decoder selects one node in  $v_1, v_2, \dots, v_{|V|}$  as  $v_{j_k}$  by an attention weight vector  $\alpha^k$  over  $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_{|V|}$ .  $\alpha^k$  is defined as:

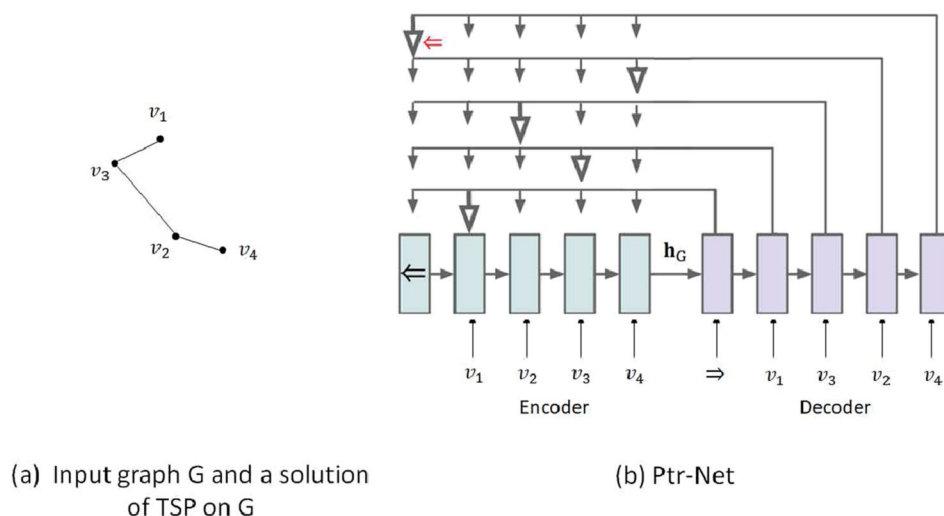
$$\alpha^k[j] = \mathbf{c}^T [\tanh(\mathbf{W}_1 \mathbf{a}_j + \mathbf{W}_2 \mathbf{b}_k)], 1 \leq j \leq |V|$$

where  $\mathbf{c}$ ,  $\mathbf{W}_1$ , and  $\mathbf{W}_2$  are trainable parameters. Then, the decoder outputs  $v_{j_k} = v_i$ , where  $i = \operatorname{argmax} \alpha^k$ .

For example, Fig. 10a shows a Euclidean graph  $G$  with four nodes and a solution  $v_1, v_3, v_2, v_4$ . Fig. 10b shows the procedure of Ptr-Net for computing the solution. The hollow



**Fig. 10** An example of using Ptr-Net [67]. **a** shows a Euclidean graph  $G$  on a 2D plane, and the solution is marked by the edges. **b** shows the encoder and the decoder of Ptr-Net for finding the solution on  $G$



arrow marks the node that has the largest attention weight at each time step of the decoder.

Milan et al. [46] propose a LSTM-based method to solve the graph matching problem. Given two graphs  $G_1$  and  $G_2$  of  $n$  nodes, from the features of nodes and edges of  $G_1$  and  $G_2$ , a  $n^2$  by  $n^2$  similarity matrix  $\mathbf{M}$  can be computed, where  $\mathbf{M}_{ij,kl}$  is the similarity of the edge  $(v_i, v_j) \in G_1$  and  $(v_k, v_l) \in G_2$ , and  $\mathbf{M}_{ii,ll}$  is the similarity of the node  $v_i$  of  $G_1$  and  $v_l$  in  $G_2$ .  $\mathbf{M}$  is input to the LSTM as the input feature. At each step, the LSTM will predict a node pair of matching. The cross entropy with the ground-truth matching is used as the loss. However,  $\mathbf{M}$  is of  $O(n^4)$  size, which is too large for matching large graphs.

Du et al. [19] observe that link prediction and graph alignment are inherently related and the joint learning of them can benefit each other. Given two graphs  $G_1$  and  $G_2$ , crossing edges between all nodes of  $G_1$  and  $G_2$  are added. The network alignment model predicts the probability of accepting a crossing edge, i.e., the end nodes of the crossing edge are aligned. The link prediction model predicts the probability of inserting an edge  $(u, v)$  to  $G_1$  based on if  $(u', v')$  is in  $G_2$ , where  $u$  and  $v$  are aligned to  $u'$  and  $v'$ , respectively. Both the network alignment model and the link prediction model need the embeddings of the nodes of  $G_1$  and  $G_2$ , which are computed by the generalized Skip-Gram model using the random walks crossing the two graphs. Suppose the random walk is on  $G_1$ , it will switch to  $G_2$  at the next step with probability  $p$ . If the random walk switches, the probability of walking from a node  $v$  in  $G_1$  to a node  $u$  in  $G_2$  is  $p'(v, u)$ . If the crossing edge between  $v$  and  $u$  is an accepted crossing edge,  $p'(v, u) = 1$ ; otherwise,  $p'(v, u) = \frac{w(v,u)}{Z}$ , where  $w(v, u)$  is the structure similarity between  $v$  and  $u$  and  $Z = \sum_{u' \in G_2} w(v, u')$ .  $w(v, u)$  is measured by the degree distributions of the neighbors of  $v$  and  $u$  in  $G_1$  and  $G_2$ , respectively. In each iteration, the pair of

nodes of the two graphs having the largest predicted probability by the graph alignment model is aligned and the edges of  $G_1$  and  $G_2$  whose probabilities predicted by the link prediction model exceed the threshold are added to  $G_1$  and  $G_2$ , respectively. Node embeddings are recomputed in each iteration, as the alignment between  $G_1$  and  $G_2$  and edges in  $G_1$  and  $G_2$  are updated. Experiments of [19] show that link prediction and graph alignment can benefit each other and the proposed techniques are suitable for aligning graphs whose distribution of the degree of aligned nodes is close to linear or the graphs having no node attribute information.

**B. Reinforcement Learning-Based Searching**

When iteratively extending a partial solution, each iteration selects the node in order to optimize the final solution. Such a sequential decision process can be modeled as a Markov decision process (MDP) and solved by reinforcement learning (RL). Therefore, we first presents a brief review of RL.

**B.1 Review of Reinforcement Learning**

In RL, an agent acts in an environment, collecting rewards and updating its policy to select future actions. It can be formulated as an MDP  $(S, A, T, R, \gamma)$ , where

- $S$  is the set of states, and some states in  $S$  are end states;
- $A$  is the set of actions;
- $T : S \times A \times S \rightarrow [0, 1]$  is the transition function,  $T(s, a, s')$  is the transition probability to state  $s'$  after taking action  $a$  in state  $s$ ;
- $R : S \times A \rightarrow \mathbb{R}$  is the reward of taking action  $a$  in state  $s$ ; and
- $\gamma$  is a discount factor.

The agent uses a policy  $\pi : S \rightarrow A$  to select an action for a state. RL is to learn an optimal policy  $\pi^*$  that can return the optimal action for each state in terms of the overall reward. RL relies on the state-value function and the action-value function to optimize the policy. The state-value function  $V^\pi(s)$  denotes the overall reward starting from the state  $s$  following the policy  $\pi$ . The action-value function  $Q^\pi(s, a)$  denotes the overall reward starting from the state  $s$  and the action  $a$  following the policy  $\pi$ . Formally,

$$V^\pi(s) = \mathbb{E}_\pi \left[ \sum_{t=0}^T \gamma^t R(s_t, a_t) | s_0 = s \right],$$

$$Q^\pi(s, a) = \mathbb{E}_\pi \left[ \sum_{t=0}^T \gamma^t R(s_t, a_t) | s_0 = s, a_0 = a \right],$$

where  $\mathbb{E}_\pi$  denotes the expected value given that the agent follows the policy  $\pi$ ,  $t$  is the time step and  $T$  is the time step of reaching an ending state. The state-value function and the action-value function of the optimal policy  $\pi^*$  are denoted by  $V^*$  and  $Q^*$ , respectively.

RL can learn  $\pi^*$  by iteratively optimizing the value functions, which is called as the value-based method. The value-based methods compute  $Q^*$  and output the optimal policy  $\pi^*(s) = \max_a Q^*(s, a)$ . Q-learning is a well-known value-based RL method. Suppose  $Q$  is the current action-value function. At each state  $s_t$ , Q-learning selects the action  $a_t$  by the  $\epsilon$ -greedy policy, which is selecting  $\max_a Q(s, a)$  with a probability  $1 - \epsilon$  and selecting a random action with a probability  $\epsilon$ , and updates  $Q$  as Formula 26.

$$Q(s_t, a_t) = Q(s_t, a_t) + \alpha_t [R(s_t, a_t) + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t)], \tag{26}$$

where  $\alpha_t$  is the learning rate at the time step  $t$ . Q-learning converges to  $Q^*$  with probability 1, if each state-action pair is performed infinitely often and  $\alpha_t$  satisfies  $\sum_{n=1}^\infty \alpha_t = \infty$  and  $\sum_{n=1}^\infty \alpha_t^2 < \infty$ .

Q-learning needs a table, namely Q-table, to store the action values. The size of the Q-table is  $|S| \times |A|$ , which can be too large to support the applications having a large number of states and actions. Therefore, many methods have been proposed to approximate the Q-table by parameterized functions. For example, deep Q-learning network (DQN) uses a deep neural network as the function approximation of the Q-table [47].

The value-based methods first optimize the value functions and then improve the policy based on the optimized value functions. There are also many methods that directly optimize the policy based on policy gradient. We refer the reader to [61] for more details of RL.

### B.2 Reinforcement Learning-Based CO Methods

Since iteratively extending a partial solution of a CO problem is inherently a sequential decision process, several works use reinforcement learning (RL) to extend the partial solution. The partial solution and the input graph together determine the state of RL, whereas the node that can be added to the partial solution is the action. RL can learn an optimal policy to find the optimal node for a partial solution.

Dai et al. propose S2V-DQN [14] that combines GNN and deep Q-learning to tackle the MVC problem. Given a graph  $G$ , let  $U$  denotes the current partial solution and  $\bar{U} = V \setminus U$ . The RL task for MVC can be formulated as follows.

- A state  $s$  is determined by  $G$  and  $U$ ,  $s = f_{state}(G, U)$ . If  $U$  is a vertex cover of  $G$ , the state is an end state;
- An action  $a_v$  is adding a node  $v \in \bar{U}$  to  $U$ ;
- The transition  $T(f_{state}(G, U), a_v) = f_{state}(G, U \cup \{v\})$ ; and
- The reward of an action  $R(s, a_v) = -1$  so as to minimize the vertex cover.

The representation of state  $s$  can be computed by embedding  $G$  and  $U$  using a GNN as follows.

$$f_{state}(G, U) = \sum_v \mathbf{h}_v^L$$

$$\mathbf{h}_u^l = ReLU \left( \theta_1 x_u + \theta_2 \sum_{v \in N_u} \mathbf{h}_v^{l-1} + \theta_3 \sum_{v \in N_u} ReLU(\theta_4 w_{u,v}) \right), \tag{27}$$

where  $L$  is the total number of layers of the GNN,  $x_u = 1$  if  $u \in U$  and otherwise,  $x_u = 0$ ,  $w_{u,v}$  is the weight of the edge  $(u, v)$ , and  $\theta_1, \theta_2, \theta_3$  and  $\theta_4$  are trainable parameters.

We can use the embedding of  $v$ ,  $\mathbf{h}_v$  to represent the action  $a_v$ . The representations of the state  $s$  and the action  $a_v$  are fed into an MLP to compute  $Q(s, a_v)$  as below.

$$Q(s, a_v) = \theta_5 ReLU \left( Concat \left( \theta_6 \sum_{u \in V} \mathbf{h}_u^L, \theta_7 \mathbf{h}_v^L \right) \right), \tag{28}$$

where  $\theta_5, \theta_6$ , and  $\theta_7$  are trainable parameters.

Deep Q-learning is used to optimize the parameters. After the MLP and the GNN are trained, they can be generalized to compute MVC for unseen graphs. S2V-DQN can also solve the MaxCut and TSP problems.

Bai et al. [4] propose to compute the maximum common subgraph (MCS) of two graphs using GNN and Q-learning. Given two graphs  $G_1$  and  $G_2$ , the partial solution is a subgraph  $g_1$  of  $G_1$  and a subgraph  $g_2$  of  $G_2$  satisfying  $g_1$  and  $g_2$  are isomorphic. The RL task for MCS is formulated as follows.

- A state  $s$  is determined by  $G_1, G_2, g_1$  and  $g_2, s = f_{state}(G_1, G_2, g_1, g_2)$ . If  $g_1$  and  $g_2$  cannot be extended, the state is an end state;
- An action  $a_{u,v}$  is to select a node  $u$  from  $G_1 \setminus g_1$  and a node  $v$  from  $G_2 \setminus g_2$  and add them to  $g_1$  and  $g_2$ , respectively;
- The transaction  $T(f_{state}(G_1, G_2, g_1, g_2), a_{u,v}) = f_{state}(G_1, G_2, g_1 \cup \{u\}, g_2 \cup \{v\})$ . The isomorphism between  $g_1 \cup \{u\}$  and  $g_2 \cup \{v\}$  needs to be assured; and
- The reward  $R(s, a_{u,v}) = 1$ .

The representation of the state  $s$  can be computed by a GNN on an auxiliary graph  $G'$ .  $G'$  is constructed by adding a pseudo node  $n_s$  connecting to the nodes in  $g_1$  and the nodes in  $g_2$ . Then, a GNN is used to compute the node embeddings for  $G'$ . Note that the node embeddings change with the extension of the partial solution  $g_1$  and  $g_2$ .  $\mathbf{h}_{G_1}$  and  $\mathbf{h}_{G_2}$  can be computed by the summation of the embeddings of the nodes in  $G_1$  and  $G_2$ , respectively. The concatenation of  $\mathbf{h}_{n_s}, \mathbf{h}_{G_1}$  and  $\mathbf{h}_{G_2}$  is the representation of the state  $s$ . The action  $a_{u,v}$  is represented by the concatenation of  $\mathbf{h}_u$  and  $\mathbf{h}_v$ . The representations of the states and the actions are fed into an MLP to predict  $Q$ . Fig. 11a, b show an example.

Rather than just selecting one node with the largest Q-value as in [14], Bai et al. [4] propose to select  $k$  nodes utilizing the beam search. At each time step, the agent of RL is allowed to transit to at most  $k$  best next states. The beam search builds an exploration tree, where each node of the tree is a state and each edge of the tree is an action. Figure 11c shows an example of  $k = 3$ . The partial solution is returned as a maximal independent set if it cannot be extended. The largest one among the computed maximal independent sets is outputted.

Inspired by AlphaGo Zero, which has surpassed human in the game Go, Abe et al. [1] propose CombOptZero, combining GNN and Monte Carlo tree search (MCTS)-based RL to solve the MVC problem. The formulation of the RL task is as S2V-DQN [14]. The key difference is that CombOptZero uses the MCTS-based searching for the next action. For a state  $s$ , suppose  $U$  is the partial solution, a GNN embeds  $G$  and  $U$  and outputs two vectors  $\mathbf{p}$  and  $\mathbf{v}$ , where  $\mathbf{p}[a]$  is the probability of taking the action  $a$  for the state, and  $\mathbf{v}[a]$  is the estimated overall reward from the state  $s$  with action  $a$ .  $\mathbf{p}$  and  $\mathbf{v}$  are input to a MCTS, which can produce a better

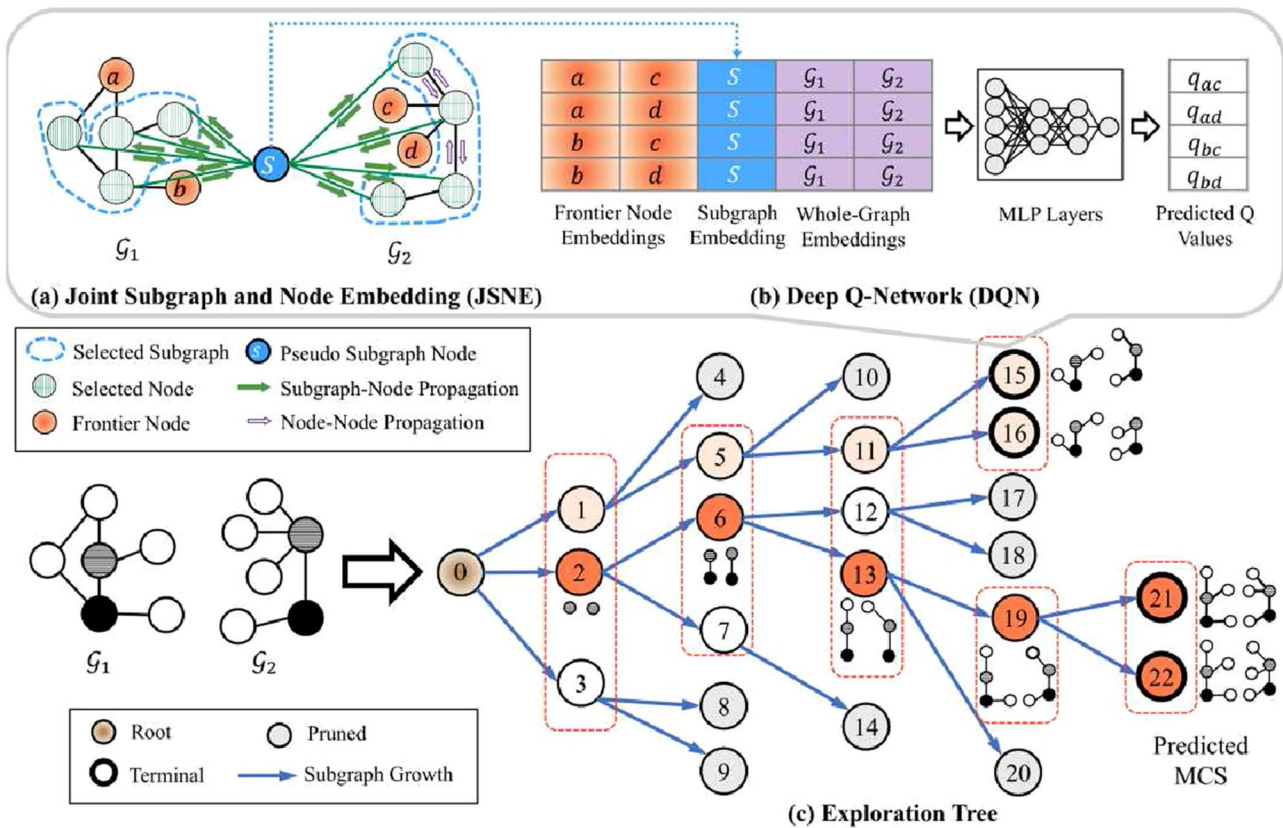


Fig. 11 Overview of RLMCS [4]

action prediction  $\mathbf{p}'$  than  $\mathbf{p}$ .  $\text{argmax}_a \mathbf{p}'[a]$  is outputted as the optimal action selected for  $s$ . CombOptZero can also solve the MaxCut problem.

Kool et al. [37] use AutoEncoder to sequentially output a TSP tour of a graph  $G$ . The encoder stacks  $L$  self-attention layers. The details of the encoder are presented in Sect. 3.2.3.

The decoder of [37] sequentially predicts the next node to be added to the partial solution  $seq$ , i.e., a partial TSP tour. At the  $t$ -th step of decoding,  $seq$  has  $t - 1$  nodes. A special context vector  $\mathbf{h}_c$  is introduced to represent the decoding context. At the  $t$ -th step of decoding,  $\mathbf{h}_c = \mathbf{h}_G || \mathbf{h}_{seq_{t-1}} || \mathbf{h}_{seq_0}$ , where  $||$  denotes concatenation,  $seq_0$  denotes the 0-th node in  $seq$  and  $seq_{t-1}$  denotes the  $t - 1$ -th node in  $seq$ . The embedding of a node  $v_i$  is computed as  $\mathbf{h}_{v_i} = \sum_{j \in N_i} \alpha_j \mathbf{W}_1 \mathbf{h}_j$ , where the attention weight  $\alpha_j = \frac{e^{u_j}}{\sum_{v_j \in N_i} e^{u_j}}$  and  $u_j = (\mathbf{W}_2 \mathbf{h}_c)^T (\mathbf{W}_3 \mathbf{h}_{v_j})$ , if  $v_j \notin seq$ ; otherwise,  $u_j = -\infty$ . The probability of choosing  $v_i$  to add to  $seq$  at the  $t$ -th step is  $p_{v_i} = \frac{e^{u_j}}{\sum_{v_j \in G} e^{u_j}}$ .  $\mathbf{W}_1$ ,  $\mathbf{W}_2$  and  $\mathbf{W}_3$  are trainable parameters. The REINFORCE algorithm is used to train the model.

The experiments presented in [37] show that the proposed method can support several related problems of TSP, including vehicle routing problem (VRP), orienteering problem (OP), prize collecting TSP (PCTSP) and stochastic PCTSP (SPCTSP) with the same set of hyperparameters. However, the proposed method does not outperform the specialized algorithm for TSP (e.g., Concorde).

There are works not iteratively extending a partial solution to a solution of a CO problem but iteratively improving a suboptimal solution to a better solution. For example, Wu et al. [71] propose to improve the solution of TSP (i.e., a TSP tour) on  $G$  using RL. The MDP is defined as follows. A TSP tour of  $G$  is a state  $s = (v_1, v_2, \dots, v_n)$ ,  $n$  is the number of nodes in  $G$  and  $v_i \neq v_j$  for  $i \neq j$ . A 2-opt operator is an action. Given two nodes  $v_i, v_j$  in  $s$ , the 2-opt operator selects a pair of nodes  $v_i$  and  $v_j$  and reverses the order of nodes between  $v_i$  and  $v_j$  in  $s$ . The transition of an action is deterministic. The reward of an action is the reduction of the TSP tour with respect to the current best TSP tour so far. The architecture of Transformer is adopted to compute node embeddings. The compatibility of a pair of nodes  $v_i$  and  $v_j$  is computed as  $(\mathbf{W}_1 \mathbf{h}_i)^T (\mathbf{W}_2 \mathbf{h}_j)$ , where  $\mathbf{W}_1$  and  $\mathbf{W}_2$  are trainable parameters. The compatibilities of all pairs of nodes are stored in a matrix  $\mathbf{Y}$ .  $\mathbf{Y}$  is fed into a masked softmax layer as follows.

$$\mathbf{Y}'_{i,j} = \begin{cases} C \cdot \tanh(Y_{i,j}), & \text{if } i \neq j \\ -\infty, & \text{if } i = j \end{cases}$$

$$\mathbf{P} = \text{softmax}(\mathbf{Y}')$$

where  $\mathbf{P}_{i,j}$  is the probability of selecting the pair of nodes  $v_i$  and  $v_j$  in the 2-opt operator. REINFORCE is used to train the model. The experiments reported in [71] show that the

proposed techniques outperform the heuristic algorithms for improving TSP tours.

### 4.3 Discussions

Non-autoregressive methods predict the probabilities that each node/edge being a part of a solution in one shot. The cross entropy between the predicted probabilities and the ground-truth solution of the CO problem is used as the loss function. Autoregressive methods predict the node/edge to add to the partial solution step by step. The inference of non-autoregressive methods can be faster than autoregressive methods, as when performing inference non-autoregressive methods predict a solution in one shot. Fast inference is desired for some real-time decision-making tasks, e.g., the vehicle routing problem. However, non-autoregressive methods inherently ignore some sequential characteristics of some CO problems, e.g., the TSP problem. Autoregressive methods can explicitly model this sequential inductive bias by attention mechanism or recurrent neural networks. Experimental comparison in [33] shows that the autoregressive methods can outperform the non-autoregressive methods in terms of the quality of the tour found for the TSP problem but takes much longer time. However, for the problem without sequential characteristic, non-autoregressive methods can produce better solution, e.g., molecule generation task [32].

The non-autoregressive methods need the ground-truth solution for supervised training. It is a drawback of the non-autoregressive methods as it is hard to compute the ground-truth solution for the CO problems on large graphs, considering the NP-hardness of the CO problems. The autoregressive methods with reinforcement learning-based searching do not need the ground-truth, which has the potential to support larger graphs. Moreover, the supervised learning of non-autoregressive models that having a large number of parameters can make the models remember the training instances and the generalization is limited on unseen instances. Although reinforcement learning can overcome this problem, the sample efficiency needs to be improved.

Regarding the comparison with traditional heuristic algorithms for the CO problems, current learning-based CO methods can have competitive performance with the traditional problem-specific heuristic algorithms on small graphs, but current learning-based CO methods do not scale well to large graphs. As large graphs have been emerging in many applications, there have been a trend of studying learning-based methods on large graphs.

The techniques and ideas of traditional heuristics for the CO problems can benefit the learning-based CO methods. For example, Dai et al. [14] present that incorporating the idea of adding the farthest nodes first and the 2-opt operation of traditional heuristics can improve the performance

of the learning-based method for TSP. Exploring the chance of integrating the ideas and operations of traditional heuristics into the learning-based methods is attracting increasing research attention [14, 17, 37].

## 5 Future Work

Although there are recent significant advances of using graph learning models in solving several different CO problems, graph learning-based methods for CO problems are still at the early stage and there are many open problems for further studies. Some possible directions are listed as follows.

*Encoding Global Information of Graphs* In many graph-based CO problems, the global information of the graph is needed for solving the CO problem (e.g., graph edit distance, TSP). However, existing graph learning models, especially graph convolution, is aggregating local information from neighbors. Although more global information can be obtained by adding more graph convolution layers, there may be a non-trivial over-smooth problem. Therefore, how to effectively encode more global information of graphs is an important direction.

*Designing Task-Dependent Model* A GNN architecture is used to support diverse types of CO problems. However, each problem has its own characteristics. How to encode inductive bias into GNN architectures in order to better capture the characteristics of the CO problems is an important direction.

The loss function that is generally used in classification or regression (e.g., cross entropy) is widely used in the learning-based methods for solving CO problems. However, the general loss function may not have a strong relationship with the objective of the CO problems. For example, switching two nodes in a TSP tour will produce a TSP tour of very different score with respect to the objective of TSP. However, the two TSP tours can have the same loss in terms of cross entropy [46]. Therefore, designing the problem-specific loss function needs to be studied.

*Generalization* Most existing learning-based methods for a CO problem cannot outperform traditional heuristic algorithms specifically designed for the CO problem on a larger graph or the graphs unseen in training, although the learning-based methods can be on par with or better than the traditional heuristic algorithm on small graphs. Therefore, an important direction is to rethink the learning pipeline for CO problem in order to generalize to larger graphs and unseen graphs [33].

*Integration of Traditional Heuristics* Integrating traditional heuristics can improve the performance of learning-based CO methods. For example, Dai et al. [14] present

that incorporating the idea of adding the farthest nodes first and the 2-opt operation of traditional heuristics can improve the performance of the learning-based method for TSP. Therefore, identifying the operations of traditional heuristics of a CO problem that can benefit the learning-based methods for the CO problem and integrating the operations appropriately into the learning procedure need to be studied.

*Supporting Many Graphs* Most existing graph learning-based CO methods focus on a graph or two graphs. Another possible future direction is to study the problems that involve a large number of graphs, for example, by optimizing the query evaluation on a large graph database such as graph similarity search, graph pattern matching query and sub-graph isomorphism search.

## 6 Conclusion

In this survey, we provided a thorough overview of the recent graph learning methods for solving CO problems. Existing works fall into two main categories. First, non-autoregressive methods predict the solution of a CO problem in one shot. Second, autoregressive methods iterative extend a partial solution step by step. Heuristic search and reinforcement learning are widely used in the autoregressive methods to extend the partial solution. In these graph learning-based CO methods, a graph is represented in numerical vectors. Then, we also survey the recent graph representation learning methods, including the generalized SkipGram-based methods, the AutoEncoder-based methods and the GNN-based methods. Several possible directions for future research are discussed as well.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Abe K, Xu Z, Sato I, Sugiyama M (2019) Solving np-hard problems on graphs with extended alpha zero. [arXiv:1905.11623](https://arxiv.org/abs/1905.11623)
2. Bai Y, Ding H, Bian S, Chen T, Sun Y, Wang W (2019) SimGNN: a neural network approach to fast graph similarity computation.

- In: Proceedings of the ACM international conference on web search and data mining (WSDM'19), pp 384–392
3. Bai Y, Ding H, Gu K, Sun Y, Wang W (2020) Learning-based efficient graph similarity computation via multi-scale convolutional set matching. In: Proceedings of the AAAI conference on artificial intelligence (AAAI'20), pp 3219–3226
  4. Bai Y, Xu D, Wang A, Gu K, Wu X, Marinovic A, Ro C, Sun Y, Wang W (2020) Fast detection of maximum common subgraph via deep Q-learning. [arXiv:2002.03129](https://arxiv.org/abs/2002.03129)
  5. Bengio Y, Lodi A, Prouvost A (2018) Machine learning for combinatorial optimization: A methodological tour d'horizon. [arXiv:1811.06128](https://arxiv.org/abs/1811.06128)
  6. Bonner S, Kureshi I, Brennan J, Theodoropoulos G, McGough AS, Obara B (2019) Exploring the semantic content of unsupervised graph embeddings: an empirical study. *Data Sci Eng* 4(3):269–289
  7. Cai H, Zheng VW, Chang K (2018) A comprehensive survey of graph embedding: problems, techniques, and applications. *IEEE Trans Knowl Data Eng* 30(09):1616–1637
  8. Cao S, Lu W, Xu Q (2015) GraRep: learning graph representations with global structural information. In: Proceedings of ACM international on conference on information and knowledge management (CIKM'15), pp 891–900
  9. Chami I, Abu-El-Hajja S, Perozzi B, Ré C, Murphy K (2020) Machine learning on graphs: a model and comprehensive taxonomy. [arXiv:2005.03675](https://arxiv.org/abs/2005.03675)
  10. Chang L (2019) Efficient maximum clique computation over large sparse graphs. In: Proceedings of ACM SIGKDD international conference on knowledge discovery and data mining (KDD'19), pp 529–538
  11. Chen F, Wang YC, Wang B, Kuo CCJ (2020) Graph representation learning: a survey. *APSIPA Trans Signal Inf Process* 9:e15
  12. Chen H, Yin H, Chen T, Nguyen QVH, Peng W, Li X (2019) Exploiting centrality information with graph convolutions for network representation learning. In: IEEE international conference on data engineering (ICDE'19), pp 590–601
  13. Cui P, Wang X, Pei J, Zhu W (2019) A survey on network embedding. *IEEE Trans Knowl Data Eng* 31(5):833–852
  14. Dai H, Khalil EB, Zhang Y, Dilkina B, Song L (2017) Learning combinatorial optimization algorithms over graphs. In: Proceedings of international conference on neural information processing systems (NIPS'17), pp 6351–6361
  15. Daredy MR, Das M, Yang H (2019) Motif2Vec: Motif aware node representation learning for heterogeneous networks. In: 2019 IEEE international conference on big data (Big Data'19), pp 1052–1059
  16. Dave VS, Zhang B, Chen PY, Al Hasan M (2019) Neural-brane: neural bayesian personalized ranking for attributed network embedding. *Data Sci Eng* 4(2):119–131
  17. Deudon M, Cournut P, Lacoste A, Adulyasak Y, Rousseau LM (2018) Learning heuristics for the TSP by policy gradient. In: International conference on the integration of constraint programming, artificial intelligence, and operations research, pp 170–181
  18. Devlin J, Chang MW, Lee K, Toutanova K (2019) BERT: Pre-training of deep bidirectional transformers for language understanding. *Proc Conf N Am Chapter Assoc Comput Linguist Hum Lang Technol* 1:4171–4186
  19. Du X, Yan J, Zha H (2019) Joint link prediction and network alignment via cross-graph embedding. In: Proceedings of international joint conference on artificial intelligence, IJCAI'19, pp 2251–2257
  20. Duvenaud DK, Maclaurin D, Iparraguirre J, Bombarell R, Hirzel T, Aspuru-Guzik A, Adams RP (2015) Convolutional networks on graphs for learning molecular fingerprints. *Adv Neural Inf Process Syst* 28:2224–2232
  21. Fan W, Ma Y, Li Q, He Y, Zhao E, Tang J, Yin D (2019) Graph neural networks for social recommendation. In: The world wide web conference, pp 417–426
  22. Fan Z, Peng Y, Choi B, Xu J, Bhowmick SS (2014) Towards efficient authenticated subgraph query service in outsourced graph databases. *IEEE Trans Serv Comput* 7(4):696–713
  23. Goemans MX, Williamson DP (1995) Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *J ACM* 42(6):1115–1145
  24. Grover A, Leskovec J (2016) Node2Vec: Scalable feature learning for networks. In: Proceedings of ACM SIGKDD international conference on knowledge discovery and data mining (KDD'16), pp 855–864
  25. Hamilton WL, Ying R, Leskovec J (2017) Inductive representation learning on large graphs. In: Proceedings of international conference on neural information processing systems (NIPS'17), pp 1025–1035
  26. Hamilton WL, Ying R, Leskovec J (2017) Representation learning on graphs: methods and applications. *IEEE Data Eng Bull*
  27. Håstad J (1999) Clique is hard to approximate within  $n^{1-\epsilon}$ . *Acta Math* 182(1):105–142
  28. Hopfield J, Tank D (1985) Neural computation of decisions in optimisation problems. *Biol Cybern* 52:141–152
  29. Huang J, Patwary MMA, Diamos GF (2019) Coloring big graphs with alphagozero. *CoRR* [arXiv:1902.10162](https://arxiv.org/abs/1902.10162)
  30. Huang W, Zhang T, Rong Y, Huang J (2018) Adaptive sampling towards fast graph representation learning. In: Proceedings of international conference on neural information processing systems (NIPS'18), pp 4563–4572
  31. Huang X, Lakshmanan LV, Xu J (2019) Community search over big graphs. Morgan & Claypool Publishers, New York
  32. Jin W, Barzilay R, Jaakkola T (2018) Junction tree variational autoencoder for molecular graph generation. *Proc Int Conf Mach Learn* 80:2323–2332
  33. Joshi CK, Cappart Q, Rousseau LM, Laurent T, Bresson X (2020) Learning TSP requires rethinking generalization. [arXiv:2006.07054](https://arxiv.org/abs/2006.07054)
  34. Joshi CK, Laurent T, Bresson X (2019) An efficient graph convolutional network technique for the travelling salesman problem. *CoRR* [arXiv:1906.01227](https://arxiv.org/abs/1906.01227)
  35. Khot S (2001) Improved inapproximability results for maxclique, chromatic number and approximate graph coloring. In: Proceedings IEEE symposium on foundations of computer science, pp 600–609
  36. Kipf TN, Welling M (2017) Semi-supervised classification with graph convolutional networks. In: International conference on learning representations (ICLR'17)
  37. Kool W, van Hoof H, Welling M (2019) Attention, learn to solve routing problems! In: International conference on learning representations
  38. Lamb LC, Garcez AD, Gori M, Prates MO, Avelar PH, Vardi MY (2020) Graph neural networks meet neural-symbolic computing: a survey and perspective. In: Proceedings of international joint conference on artificial intelligence (IJCAI'20), pp 4877–4884
  39. Le Q, Mikolov T (2014) Distributed representations of sentences and documents. In: Proceedings of international conference on machine learning, vol 32, ICML'14, pp II–1188–II–1196
  40. Lemos H, Prates MOR, Avelar PHC, Lamb LC (2019) Graph colouring meets deep learning: Effective graph neural network models for combinatorial problems. *CoRR* [arXiv:1903.04598](https://arxiv.org/abs/1903.04598)
  41. Li Y, Gu C, Dullien T, Vinyals O, Kohli P (2019) Graph matching networks for learning the similarity of graph structured objects. In: Proceedings of international conference on machine learning (ICML'19), pp 3835–3845

42. Li Z, Chen Q, Koltun V (2018) Combinatorial optimization with graph convolutional networks and guided tree search. In: Proceedings of international conference on neural information processing systems (NIPS'18), pp 537–546
43. Mazyavkina N, Sviridov S, Ivanov S, Burnaev E (2020) Reinforcement learning for combinatorial optimization: a survey. [arXiv:2003.03600](https://arxiv.org/abs/2003.03600)
44. Meng L, Zhang J (2019) IsoNN: Isomorphic neural network for graph representation learning and classification. CoRR [arXiv:1907.09495](https://arxiv.org/abs/1907.09495)
45. Mikolov T, Sutskever I, Chen K, Corrado G, Dean J (2013) Distributed representations of words and phrases and their compositionality. In: Proceedings of the 26th international conference on neural information processing systems (NIPS'13), pp 3111–3119
46. Milan A, Rezatofighi SH, Garg R, Dick A, Reid I (2017) Data-driven approximations to np-hard problems. In: Thirty-first AAAI conference on artificial intelligence
47. Mnih V, Kavukcuoglu K, Silver D, Rusu AA, Veness J, Bellemare MG, Graves A, Riedmiller M, Fidjeland AK, Ostrovski G et al (2015) Human-level control through deep reinforcement learning. *Nature* 518(7540):529–533
48. Narayanan A, Chandramohan M, Chen L, Liu Y, Saminathan S, Subgraph2Vec: learning distributed representations of rooted subgraphs from large graphs
49. Nazi A, Hang W, Goldie A, Ravi S, Mirhoseini A (2019) GAP : Generalizable approximate graph partitioning framework. URL ICLR workshop
50. Nguyen H, Murata T (2017) Motif-aware graph embeddings. In: Proceedings of international joint conference on artificial intelligence (IJCAI'17), pp 1–1
51. Nowak A, Villar S, Bandeira AS, Bruna J (2017) A note on learning algorithms for quadratic assignment with graph neural networks. In: Proceeding of international conference on machine learning (ICML'17), pp 22–22
52. Ou M, Cui P, Pei J, Zhang Z, Zhu W (2016) Asymmetric transitivity preserving graph embedding. In: Proceedings of ACM SIGKDD international conference on knowledge discovery and data mining (KDD'16), pp 1105–1114
53. Papadimitriou CH, Vempala S (2006) On the approximability of the traveling salesman problem. *Combinatorica* 26(1):101–120
54. Peng Y, Choi B, He B, Zhou S, Xu R, Yu X (2016) VColor: A practical vertex-cut based approach for coloring large graphs. In: IEEE international conference on data engineering (ICDE'16), pp 97–108
55. Peng Y, Fan Z, Choi B, Xu J, Bhowmick SS (2015) Authenticated subgraph similarity search in outsourced graph databases. *IEEE Trans Knowl Data Eng* 27(7):1838–1860
56. Perozzi B, Al-Rfou R, Skiena S (2014) DeepWalk: Online learning of social representations. In: Proceedings of ACM SIGKDD international conference on knowledge discovery and data mining (KDD'14), pp 701–710
57. Prates M, Avelar P, Lemos H, Lamb L, Vardi M (2019) Learning to solve NP-Complete problems - a graph neural network for decision TSP. In: Proceedings of AAAI conference on artificial intelligence (AAAI'19), pp 4731–4738
58. Ribeiro LF, Saverese PH, Figueiredo DR (2017) Struc2Vec: Learning node representations from structural identity. In: Proceedings of ACM SIGKDD international conference on knowledge discovery and data mining (KDD'17), pp 385–394
59. Sato R, Yamada M, Kashima H (2019) Approximation ratios of graph neural networks for combinatorial problems. In: Proceedings of the neural information processing systems (NIPS'19)
60. Smith-Miles K (1999) Neural networks for combinatorial optimization: a review of more than a decade of research. *INFORMS J Comput* 11:15–34
61. Sutton RS, Barto AG (2018) Reinforcement learning: an introduction. MIT Press, Cambridge
62. Tang J, Qu M, Mei Q (2015) PTE: Predictive text embedding through large-scale heterogeneous text networks. In: Proceedings of ACM SIGKDD international conference on knowledge discovery and data mining (KDD'15), pp 1165–1174
63. Tang J, Qu M, Wang M, Zhang M, Yan J, Mei Q (2015) LINE: Large-scale information network embedding. In: Proceedings of international conference on world wide web (WWW'15), pp 1067–1077
64. Tsitsulin A, Mottin D, Karras P, Müller E (2018) VERSE: Versatile graph embeddings from similarity measures. In: Proceedings of world wide web conference (WWW'18), pp 539–548
65. Tu K, Cui P, Wang X, Yu PS, Zhu W (2018) Deep recursive network embedding with regular equivalence. In: Proceedings of ACM SIGKDD international conference on knowledge discovery & data mining (KDD'18), pp 2357–2366
66. Veličković P, Cucurull G, Casanova A, Romero A, Liò P, Bengio Y (2018) Graph Attention Networks. In: International conference on learning representations (ICLR)
67. Vinyals O, Fortunato M, Jaitly N (2015) Pointer networks. *Adv Neural Inf Process Syst* 28:2692–2700
68. Wang D, Cui P, Zhu W (2016) Structural deep network embedding. In: Proceedings of ACM SIGKDD International conference on knowledge discovery and data mining (KDD'16), pp 1225–1234
69. Wang R, Yan J, Yang X (2019) Learning combinatorial embedding networks for deep graph matching. In: Proceedings of the IEEE international conference on computer vision (ICCV'19), pp 3056–3065
70. Wasserman S, Faust K et al (1994) Social network analysis: methods and applications, vol 8. Cambridge University Press, Cambridge
71. Wu Y, Song W, Cao Z, Zhang J, Lim A (2019) Learning improvement heuristics for solving the travelling salesman problem. CoRR [arXiv:1912.05784](https://arxiv.org/abs/1912.05784)
72. Wu Z, Pan S, Chen F, Long G, Zhang C, Yu PS (2019) A comprehensive survey on graph neural networks. CoRR [arXiv:1901.00596](https://arxiv.org/abs/1901.00596)
73. Yanardag P, Vishwanathan S (2015) Deep graph kernels. In: Proceedings of ACM sigkdd international conference on knowledge discovery and data mining (KDD'15), pp 1365–1374
74. Yang Y, Wang X, Song M, Yuan J, Tao D (2019) SPAGAN: Shortest path graph attention network. In: Proceedings of international joint conference on artificial intelligence (IJCAI'19), pp 4099–4105
75. Yu Y, Lu Z, Liu J, Zhao G, Wen J (2019) RUM: Network representation learning using motifs. In: IEEE international conference on data engineering (ICDE'19), pp 1382–1393
76. Yue X, Wang Z, Huang J, Parthasarathy S, Moosavinasab S, Huang Y, Lin SM, Zhang W, Zhang P, Sun H (2019) Graph embedding on biomedical networks: methods, applications and evaluations. *Bioinformatics* 36(4):1241–1251
77. Zhang D, Yin J, Zhu X, Zhang C (2020) Network representation learning: a survey. *IEEE Trans Big Data* 6(1):3–28
78. Zhang F, Yuan NJ, Lian D, Xie X, Ma WY (2016) Collaborative knowledge base embedding for recommender systems. In: Proceedings of ACM SIGKDD international conference on knowledge discovery and data mining (KDD'16), pp 353–362