

图嵌入技术综述

王天阳

华中科技大学 计算机科学与技术学院, 武汉市 中国 430000

摘 要 图结构广泛存在于现实各种应用场景中, 可以表示现实中许多复杂关系, 现实世界中的复杂系统通常都可以用图的结构来进行建模, 用图中的节点或节点属性表示现实网络系统中的实体或实体标签, 用图中的边来表示真实网络中的实体关系。作为图分析和应用的重要策略, 图嵌入已经成为人工智能中备受关注的课题之一, 通过图嵌入(Graph Embedding, GE)研究, 更好地分析真实网络系统的结构和性质。图嵌入的过程等价于对图中节点或边进行降维表示学习的过程。从图嵌入的目的和意义出发, 分析了图嵌入的原理和优势, 介绍了基于矩阵分解、随机游走和深度学习的图嵌入主要方法。最后对图嵌入做了总结并对未来进行了展望, 考虑了对现有图嵌入方法的改进和优化。

关键词 图模型; 图嵌入; 矩阵分解; 随机游走; 深度学习

中图法分类号 *****

Overview of Graph Embedding Technology

Wang Tianyang

Huazhong University of Science and Technology School of Computer Science and Technology, Wuhan China 430000

Abstract Graph structure widely exists in various application scenarios in reality, and can represent many complex relationships in reality. Complex systems in the real world can usually be modeled by graph structure, and nodes or node attributes in the graph are used to represent real network systems. The entity or entity label in the graph uses the edges in the graph to represent the entity relationship in the real network. As an important strategy for graph analysis and application, graph embedding has become one of the topics that have attracted much attention in artificial intelligence. Through the research of Graph Embedding (GE), the structure and properties of real network systems can be better analyzed. The process of graph embedding is equivalent to the process of dimensionality reduction representation learning of nodes or edges in the graph. Starting from the purpose and significance of graph embedding, the principle and advantages of graph embedding are analyzed, and the main methods of graph embedding based on matrix decomposition, random walk and deep learning are introduced. Finally, I summarized the graph embedding and looked forward to the future, considering the improvement and optimization of the existing graph embedding methods.

Key words Graph model; Graph embedding; Matrix decomposition; Random walk; Deep learning

1 引言

图是复杂系统中常用的信息载体, 图结构(Graph)广泛存在于现实各种应用场景中, 可以表示现实中许多复杂关系, 如社交网络、交通网络、犯

罪网络、人物关系网络等, 因此, 现实世界中的复杂系统通常都可以用图的结构来进行建模, 用图中的节点或节点属性表示现实网络系统中的实体或实体标签, 用图中的边来表示真实网络中的实体关系, 通过图嵌入(Graph Embedding, GE)研究, 可更好地分析真实网络系统的结构和性质。比如社交媒

体,用户之间的关注关系可以构成一个庞大的社交图网络;又比如推荐系统,用户对商品的购买、浏览和评分等行为可以抽象成用户和商品的交互图。然而,现实中的图结构大多都很复杂。一个复杂的图网络中可能包含十亿个节点、千万条边和上万个簇,且不同的边可能代表着节点之间不同的关系。因此,如何有效地对图结构信息进行建模是学术界和工业界持续关注的焦点^[1]。

随着移动互联网的快速发展,我们正处在大数据时代,大数据背景下,图结构和图数据呈现出海量、高维、稀疏、异构、复杂和动态等特点,图结构作为一种非欧几里得数据,很难直接应用卷积神经网络(Convolutional Neural Network, CNN)和循环神经网络(Recurrent Neural Network, RNN)等深度学习方法。近年兴起的图嵌入(Graph Embedding)方法,为图结构的建模提供了很好的解决思路,并在工业界被广泛的应用。图嵌入的过程等价于对图中节点或边进行降维表示学习的过程。图嵌入也面临以下五个方面的挑战:节点海量性、节点属性信息的融合、图的异构性、节点动态增量性和模型的非线性。在图数据分析的研究中,可采用包括图嵌入、模型构建以及问题求解这三个步骤的支撑技术,图嵌入是最关键的一步,是一种图表示学习(Graph Representation Learning)方法^[2]。

目前,图嵌入不仅在节点分类、节点聚类、链接预测、等复杂网络上的机器学习任务中获得成功,还广泛用于社交影响力建模、内容推荐等现实任务^[3],鉴于以上应用的重要价值,近年来图嵌入引起了学界和业界的广泛关注。因此,针对现有的图嵌入方法及应用,归纳图嵌入近期的主要成果,阐述各类方法的核心技术、分析其原理和优缺点、探索面向实际应用的图嵌入解决方案和应用价值具有十分重要的意义。

2 图嵌入的原理与优势

2.1 图嵌入的原理

图嵌入(Graph Embedding)技术的思想是:把图中的节点或者边嵌入到一个低维的向量空间中,将图中的节点以低维稠密向量的形式进行表达,且节点或边在该低维空间的关系能比较完整地保留原图的结构信息,换言之,图嵌入的过程等价于对图中节点或边进行降维表示学习的过程,将属性图转换为向量或向量集。嵌入应该捕获图的拓扑结

构、顶点到顶点的关系以及关于图、子图和顶点的其他相关信息,更多的属性嵌入编码可以在以后的任务中获得更好的结果。

图嵌入一般有两种:节点嵌入和整图嵌入。节点嵌入是将原始图上每个节点通过嵌入得到一个向量表达,嵌入得到表示向量既可以用在简单任务:如判断两个节点的相似性,例如社交网络判断两个用户是否相似从而推荐好友;也可以做为输入表示用在更上游的复杂任务,如将社交网络中节点(用户)表示用在商品的推荐系统中。整图嵌入是将整张图通过嵌入得到一个向量,一般用在比较两个图的相似性上,比如上节提到到判断蛋白质分子的相似性,两个社区是否相似等。图嵌入过程示例如图 2-1 所示。

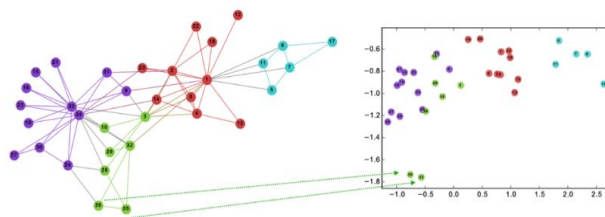


图 2-1 图嵌入过程示例

Word2vec 是近年 Graph Embedding 方法学习图节点嵌入表示的基础,其使用场景均是利用滑动的上下文窗口在序列数据上捕捉节点之间的关系,图嵌入技术需要把 Word2vec 这一高效的语言表示学习模型应用于图的拓扑结构。

(1) Word2vec

2013 年,Google 的 Mikilov 等人提出了 Word2vec 算法^[4],Word2vec 最早应用在自然语言处理领域(NLP)。他们认为:同一语义上下文的单词之间存在天然的语义关系,Word2vec 利用同一上下文的单词来学习各自的词向量表示(Word Embedding)。

具体地,Word2vec 利用语义窗口来捕捉每个句子中的语义上下文,并通过对语义窗口进行滑动,学习每一个句子序列中不同语义上下文窗口中的单词 embedding。在 Word2vec 中,每个词语都关联着两个词向量,分别为中心词向量和背景词向量,取决于当前时刻该词语的角色。为了学习每个单词的中心词向量和背景词向量,Word2vec 提出了两种模型,分别是以中心词预测背景词的 Skip-gram 模型,和以背景词去预测中心词的 C-bow 模型。Skip-gram 模型的输入是特定的一个词的词向量,输出是特定词对应的上下文词向量。而 C-bow 模型

的思路是与 Skip-gram 模型相反, C-bow 模型的训练输入是某一个特征词的上下文相关的词对应的词向量, 而输出就是这特定的一个词的词向量。

(2) 一些经典的图嵌入模型

Graph Embedding 的中心思想就是找到一种映射函数, 该函数将网络中的每个节点转换为低维度的潜在表示。利于计算存储, 不用再手动提特征(自适应性), 图嵌入可根据不同的网络结构和其他信息有许多不同的模型, 几种常见分类如图 2-2 所示。

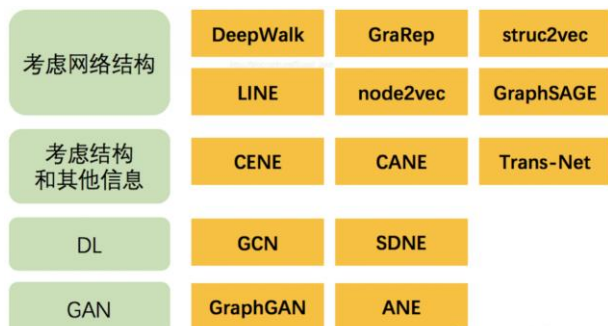


图 2-2 图嵌入的几种常见分类

2.2 图嵌入的优势

图嵌入技术有以下优势:

(1) 解决了图结构作为一种非欧几里得数据, 很难直接应用卷积神经网络和循环神经网络等机器学习和深度学习方法的局限性。图由节点和边构成, 这些向量关系一般只能用数学、统计或特定的子集进行表示, 但是嵌入之后的向量控件具有更加灵活和丰富的计算方式。

(2) 图嵌入能够压缩数据, 能够起到降维的作用。图中节点间的连接一般使用邻接矩阵来描述, 邻接矩阵的维度是 $|V| \times |V|$, 其中 $|V|$ 是图中节点的个数, 矩阵中的每列和每行都代表一个节点, 矩阵中的非零值表示两个节点相邻, 将邻接矩阵用大型图的特征空间几乎是不可能的。而嵌入可以看做是一种压缩技术, 使用图嵌入技术能够起到降维的作用, 可以有效解决维度过大的问题。

(3) 图嵌入后得到向量表示, 向量计算比直接在图上操作更加的简单、快捷。

(4) 图嵌入可用于快速比较相似数据结构的数据结构。

3 近期主要成果

早期的图嵌入算法主要用于数据降维, 通过邻域关系构建相似度图, 将节点嵌入低维向量空间, 并保持相连节点向量的相似性, 典型的算法是

LE(Laplacian Eigenmap)^[5]、Isomap(Isometric Feature Mapping)^[6]和 LLE(Locally Linear Embedding)^[7]。但是这几种算法的时间复杂度是 $O(|V|^2)$, 这类方法通常时间复杂度高, 很难扩展到大型图上。

从 2010 年起, 面对大规模且稀疏的真实网络, 学者们开始研究应用更广泛的图嵌入技术, 如 Deepwalk 模型、Text-Associated DeepWalk (TADW) 模型、Discriminative Matrix Factorization (DMF) 模型, 这些技术除了实现图的降维以外, 还可实现节点分类与聚类、链接预测、图的重构和可视化等。

最近, 随着图嵌入技术越来越受到关注, 人们针对不同应用场景提出了不同的扩展性强的图嵌入算法和模型, 主要包括基于矩阵分解的图嵌入方法^[8], 使用邻接矩阵的近似分解作为嵌入、基于随机游走的图嵌入方法^[9]将游走序列输入到 Skip-Gram 或 C-bow 生成嵌入向量^[10]、基于深度学习的图嵌入方法, 这些方法利用图的稀疏性降低了时间复杂度。

目前具有代表性的方法由表 1 概括。下面分别阐述几种典型的图嵌入方法。

表 1 图嵌入方法概括

Matrix Factorization	Random Walk	Deep Learning	Others
LE	DeepWalk	SDNE	LINE
CGE	Node2vec	DNGR	NEU
Isomap	Metapath2vec	GraphSAGE	CDK
LLE	—	—	DeepCas
GraRep	—	—	TransE
HOPE	—	—	DPMQ

3.1 基于矩阵分解的图嵌入方法

基于矩阵分解的图嵌入方法的思想是用矩阵表示图中节点之间的关联信息, 通过矩阵分解方法对图的关联矩阵进行降维, 为每个节点生成一个低维向量。其中, 用矩阵表示图的方法包括拉普拉斯矩阵和邻接矩阵等。具体研究中, 根据矩阵性质选择合适的矩阵分解方法, 目前可分为两类, 一类是基于图拉普拉斯特征映射的分解(Graph Laplacian Eigenmap Factorization), 另一类是相似矩阵的分解(Node Proximity Matrix Factorization)^[11]。

3.1.1 基于图拉普拉斯特征映射的矩阵分解

基于图拉普拉斯特征映射的矩阵分解的基本思想是用节点对的相似性来保持图的性质, 在图嵌入过程中, 如果相似的节点对被映射到向量空间中的

距离更远,该方法就给予这对节点更大的罚值。基于这一思想,该方法使用如下目标函数来实现节点对的低维映射:

$$Y = \sum_{i \neq j} (Y_i - Y_j)^2 W_{ij} = 2Y^T L Y$$

其中, L 是 G 的拉普拉斯矩阵, $L = D - W$, D 是对角矩阵, $D = \sum_j W_{ji}$, D_{ii} 值越大、 Y_i 的嵌入越重要。目标函数的最小化问题可描述为:

$$Y = \min_{Y^T D Y = I} Y^T L Y = \min \frac{Y^T L Y}{Y^T D Y}$$

式(2)中的约束条件是 $Y^T D Y = I$, 可消除嵌入中的平凡解。以下几种学习方法都可归结为拉普拉斯特征映射方法, 如经典的 LE[5]、CGE(Cauchy Graph Embedding)和 Isomap[6]等。这些方法的不同之处在于对邻接矩阵 W_{ij} 采用的计算方法不同, 或者是在目标函数中是否引入线性函数 $Y = X T a$ 等。

3.1.2 基于节点相似性的矩阵分解

基于节点相似性的矩阵分解用矩阵分解方法在低维空间中渐进地模拟图中相似的节点, 从而保存节点的相似性, 在给定图的原嵌入向量矩阵 Y 情况下, 最小化以下损失函数获得最终的 d 维嵌入向量:

$$Y^* = \min \| Y - Y_s Y_t^T \|$$

其中, Y_s 和 $Y_t \in \mathbb{R}^{|V| \times d}$, Y_s 和 Y_t 分别是低维空间中源节点和目标节点的嵌入向量。通常, 使用 SVD(Singular Value Decomposition)对原嵌入向量矩阵 Y 分解, 方法如下:

$$Y = \sum_{i=1}^{|V|} \sigma_i u_i u_i^T \approx \sum_{i=1}^d \sigma_i u_i u_i^T$$

$\{\sigma_1, \sigma_2, \dots, \sigma_{|V|}\}$ 是以降序排序的奇异值, u_i 和 u_i^T 分别是奇异值 σ_i 对应的奇异向量, 嵌入的目的是获得最大的 d 个奇异值和奇异向量, 具体表示形式为:

$$Y_s = [\sqrt{\sigma_1} u_1, \dots, \sqrt{\sigma_d} u_d]$$

$$Y_t = [\sqrt{\sigma_1} u_1, \dots, \sqrt{\sigma_d} u_d]$$

节点 i 的嵌入形式为 Y_i 或 $[Y_i, Y_t]$, 采用何种形式主要取决于是否保持图的非对称结构。目前, 除了用 SVD 对基于节点相似性矩阵进行分解之外, 还可以用其他矩阵分解法对其进行分解。

3.2 基于随机游走的图嵌入方法

基于随机游走的图嵌入方法由自然语言处理中的词-向量(Word2vec) 模型演化而来, 通常能很好

地表示图中节点之间的相互关系。尤其针对大规模图中无法观察到图的整体、而图的局部关系较容易发现的情形, 基于随机游走的图嵌入方法表现优异。该方法采用不同的游走策略获取图的局部结构和全局结构, 并且生成节点序列, 再利用 Skip-Gram 算法完成图中节点的降维嵌入。对于同构和异构图, 由于其具有不同结构, 可采用不同的基于随机游走的嵌入方法。

3.2.1 同构网络中的随机游走

(1) DeepWalk 模型

Perozzi 等人在 KDD 2014 提出了 DeepWalk 模型^[12], 是图嵌入方法的经典模型之一, 是语言模型和无监督学习从单词序列到图上的扩展, 它成功在 Word2vec 和 Graph Embedding 之间架起了连接的桥梁。为了学习图中每个节点的嵌入表示, DeepWalk 提出了一种二阶段的图嵌入学习框架: (1) 采用截断式随机游走(Truncated Random Walks)的方式把图中每个节点的局部拓扑结构转换成序列信息;

(2) 把 Word2vec 模型应用于阶段一产生的序列数据, 学习序列中每个节点的 embedding 表示。

DeepWalk 算法的整体流程如图 3-1 所示。

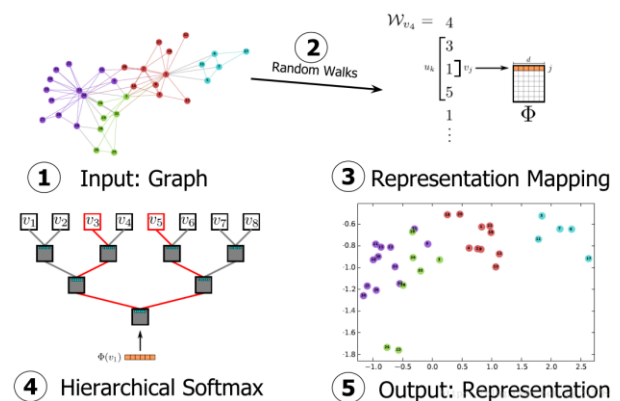


图 3-1 DeepWalk 算法整体流程

Deepwalk 不仅在数据量较少时有较好的表现, 还可以扩展到大型图的表示学习。随机游走不仅可以完成图结构到序列信息的转换, 还可以并行地为每个节点生成序列信息, 这为 DeepWalk 模型应用在大规模图结构上提供了可行性。DeepWalk 这种二阶段的图嵌入学习框架, 也被后续很多 Graph Embedding 方法所采用, 因此 DeepWalk 在学术界和工业界, 都是一个很常见的 Graph Embedding baseline。但是, 该由于优化过程中未使用明确的目标函数, 使得模型保持网络结构的能力受到限制。而且不适合带权图, 只能保持图的二阶相似性; 有限的步长会影响上下文信息的完整性; 面对大规模

的图, 调整超参数较复杂, 且游走步数超过 25 后嵌入效果不够显著。

(2) Node2vec 模型

节点-向量模型 Node2vec 也是通过改进 Word2vec 模型而建立的一种随机游走方法^[13], 通过分别调整广度优先搜索(Breadth-First Search, BFS) 和深度优先搜索(Depth-First Search, DFS) 策略的参数, 来获取图的全局结构和局部结构。具体步骤包括: 计算转移概率, 结合 BFS 和 DFS 生成随机游走序列, 用 Skip-Gram 模型对生成的游走序列进行嵌入。其优点是每一步都能可并行处理, 能保持语义信息和结构信息, 但对于含有特定属性的节点嵌入效果仍有待提高。

3.2.2 异构网络中的随机游走

(1) Metapath2vec 模型

元路径-向量模型 Metapath2vec 也是一种由 Word2vec 模型改进并应用到异构网络上的图嵌入方法^[14], 采用基于元路径的随机游走策略得到的嵌入效果, 明显优于基于同构网络的随机游走方法在异构网络上的嵌入效果。具体步骤包括: 对网络中的节点建立元路径, 根据节点在元路径中的距离和节点之间的语义关系生成游走序列, 对节点序列采用 Skip-Gram 模型实现节点嵌入。但是, 该方法在大规模图上的嵌入效果不佳, 且不能自动学习有意义的元路径, 也不能实现节点嵌入的增量更新。

(2) TriDNR 模型

TriDNR^[3]是首个利用标签信息进行表示学习的深层神经网络模型, 能够同时利用网络结构、节点特征和节点标签学习节点嵌入表示。对于网络结构, 通过随机游走序列最大化共现概率来保持图中节点间的邻近关系; 对于节点特征, 通过最大化给定节点的单词序列的共现概率捕获节点与单词的相关性; 对于节点标签, 通过最大化给定类别标签的单词序列的概率建模标签与单词的对应关系。最后, 使用耦合神经网络的算法将三部分信息合开为节点嵌入。

3.3 基于深度学习的图嵌入方法

随着深度学习研究的快速推进, 越来越多的深度神经网络技术应用于图分析中, 旨在改进基于浅层模型实现图嵌入的方法, 因为浅层模型无法捕获高度非线性结构, 进而导致非最优解的生成。目前, 典型的深度学习模型有 SDNE^[15]、DNNGR^[16]和 GraphSAGE^[17]等都是基于编码器-译码器框架的模型。其中编码阶段是将图中的节点映射为一个低维

嵌入向量, 译码阶段是将学习的到的低维嵌入向量重构为原图。

(1) SDNE 模型

结构化深度网络嵌入 (Structrual Deep Network Embedding, SDNE)^[15]是一种深度自动编码器, 它保持网络的一阶和二阶相似性, 通过高度的非线性函数和优化目标函数来获得最终的嵌入向量, 首先求 K 层的隐嵌入向量 $y(k)_i$ 和输出数据 x_i , 再求损失函数, 最后用随机梯度下降法优化整个深度学习模型。该方法能保持图的一阶和二阶相似性, 但面对百万级的节点数量时运行效率较低, 且该模型无法实现图的增量更新。

(2) DNNGR 模型

基于深度神经网络的图表示学习 (Deep Neural Networks for Learning Graph Representation, DNNGR)^[16]是结合随机漫游和深度自动编码器的方法, 主要解决传统线性降维方法无法保持图非线性结构的问题, 具体步骤包括: 用随机漫游算法生成概率共现矩阵 (Probabilistic Co-Occurrence Matrix, PCO Matrix), 由概率共现矩阵计算正点互信息矩阵 (PPMI Matrix), 用栈式降噪自动编码器生成具有高阶相似性的图嵌入。该方法能获得图的复杂特征和非线性结构, 采用 PPMI Matrix, 避免对大量无效节点的嵌入。然而, 深度学习模型调参较复杂, 无法实现图的增量更新。

(3) GraphSAGE 模型

基于图采样和聚合的图嵌入 (Graph Sample and aggreGatE, GraphSAGE)^[17]是利用节点的特征信息, 通过改进的图卷积网络生成新增节点或新增子图的嵌入向量, 最终能实现节点嵌入的增量更新和保持图的特征信息和结构信息, 具体步骤包括: 类似广度优先遍历的方式对图中每个节点的 k 深度邻近节点进行采样, 用聚类函数对每个节点的邻近节点特征信息进行聚类, 用非线性激活函数生成新的嵌入向量。但是, 当采样深度 k 大于 2 时, 性能明显下降, 能否采用非均值采样的方法, 且能否在有向图上实现优异的性能, 目前仍不明确。

3.4 其他图嵌入方法

其他图嵌入方法, 主要包括前述图嵌入方法扩展、以及基于其他理论或技术的图嵌入方法。基于矩阵分解、随机游走和深度学习的图嵌入扩展方法, 具有代表性的是 LINE (Largescale Information Network Embedding)^[18]和 NEU (Network Embedding Update)^[19]等。LINE 通过 KL 散度 (Kullback-Leibler

Divergence), 最小化由邻接矩阵生成的节点对概率分布和由嵌入内积产生的概率分布之间的距离, 最终实现图嵌入的最优化。NEU 概括了构造相似矩阵和降维的网络嵌入步骤, 并得出了以下结论: 构造更高阶、更准确的相似矩阵可以提高网络嵌入的质量; 准确地计算高阶相似矩阵对于大规模网络并不可行。

基于信息传递的图嵌入方法, 主要包括 CDK(Content Diffusion Kernel)^[20]和 DeepCas(End-to-end Deep Learning Architecture)^[21]。CDK 是运用热传递核理论(Heat Diffusion Kernel)将图中不同信息级联(Information Cascade)路径中的节点嵌入到欧式空间中, 用随机梯度下降法优化目标函数, 使得信息级联路径中较近的节点对在嵌入后对应向量的距离也较近。DeepCas 是将一个信息级联图, 用深度学习的门控循环单元 GRU(Gated Recurrent Unite)模型嵌入到一个向量空间, 用于预测图中信息级联的增量。

基于知识表示的图嵌入方法, 主要以翻译模型(Relationships as Translations in The Embedding Space, TransE)^[22]为代表, 受到 Word2vec 中词向量空间平移不变现象的启发, Bordes 等^[22]提出了 TransE 模型, 将知识库中词汇的语义关系看作向量空间中实体向量间的某种平移向量, 优点是简单有效。目前有大量研究工作对 TransE 进行扩展, 如对其改进的模型有 TransH^[23]、TransR^[24]、TransD^[25]、TranSparse^[26]和 KG2E^[27]等。基于图数据库的图嵌入方法中, DPMQ(DistanceBased Pattern Match Queries)^[28]将图数据库生成的图节点用 LLR(Linial, London and Robinovich)^[29]方法转化成空间中的向量, 将图中模式匹配查询问题转化成向量空间中基于距离的多向联结(Multi-way Joint)问题, 提高了查询的效率, 降低了图中节点之间最短距离的计算复杂度;但是, 基于图嵌入向量之间的距离实现模式匹配的查询相当于近似查询, 查询结果未必正确。

4 总结与展望

本文首先介绍了图嵌入技术的研究目的和意义, 介绍了图嵌入的原理和优势, 分析了目前主流的图嵌入方法, 从基本思想、实现步骤和优缺点等角度, 主要总结分析了基于矩阵分解、基于随机游走和基于深度学习的图嵌入方法。

总的来看, 基于矩阵分解的图嵌入方法主要包括 Isomap、LE 和 LLE 等, 容易受邻接矩阵边的约束, 在保持图的二阶甚至高阶相似性方面仍存在不足。基于随机游走的图嵌入方法主要包括 DeepWalk、Node2vec、Metapath2vec 和 TriDNR 等, 主要通过 Skip-Gram 生成嵌入向量, 但是 SkipGram 属于浅层模型, 对图中的复杂语义信息获取不够、且扩展性不强。基于深度学习的图嵌入方法主要包括 SDNE、DNDR 和 GraphSAGE 等, 但其可解释性较差。

将来可考虑对现有图嵌入方法进行改进优化:

①将现有的深度学习和矩阵分解方法结合, 采用深度学习技术, 用矩阵分解对获取的特征向量降维, 既增强了模型的可解释性又提高了算法的计算效率。

②将随机游走和矩阵分解方法结合, 通过随机游走获取节点间的二阶甚至高阶相似性, 使用矩阵分解方法对获取的特征向量进行低维嵌入, 弥补基于随机游走和矩阵分解的图嵌入方法的不足。

③研究基于并行算法的图嵌入方法, 基于 MapReduce 或 Spark 等分布式计算来扩展, 是实现图嵌入增量更新的有效解决方案。

致 谢 感谢施展老师和童薇老师的指导。

参 考 文 献

- [1] 祁志卫,王箭辉,岳昆,乔少杰,李劲.图嵌入方法与应用:研究综述[J].电子学报,2020,48(04):808-818.
- [2] 王啸,崔鹏,朱文武.网络表征学习中的基本问题初探[J].中国计算机学会通讯,2018,14(3):11-15. WANG Xiao, CUI Peng, ZHU Wenwu. A preliminary research on the fundamental problems of network representation learning[J]. Communications of the CCF,2018,14(3):11-15. (in Chinese): 922-929.
- [3] 袁立宁,李欣,王晓冬,刘钊.图嵌入模型综述[J/OL].计算机科学与探索:1-30[2022-01-05].<http://kns.cnki.net/kcms/detail/11.5602.TP.20210913.1016.002.html>.
- [4] Mikolov T, Chen K, Corrado G, et al. Efficient Estimation of Word Representations in Vector Space[J]. Computer Science, 2013.
- [5] BELKIN M, NIYOGI P. Laplacian eigenmaps and spectral techniques for embedding and clustering [A]. Proceedings of Advances in Neural Information Processing Systems [C]. Vancouver, British Columbia, Canada: MIT Press, 2001. 585-591.
- [6] TENENBAUM J, DE SILVA V, LANGFORD J. A global geometric framework for nonlinear dimensionality reduction [J]. Science, 2000, 290

- (5500):2319-2323.
- [7] ROWEIS S, SAUL L. Nonlinear dimensionality reduction by locally linear embedding [J]. Science, 2000, 290 (5500):2323-2326.
- [8] AHMED A, SHERVASHIDZE N, NARAYANAMURTHY S, et al. Distributed large-scale natural graph factorization [C]//Proceedings of the 22nd international conference on World Wide Web. 2013: 37-48.
- [9] LOVÁSZ L. Random walks on graphs: A survey[J]. Combinatorics, Paul erdos is eighty, 1993, 2(1): 1-46.
- [10] GOLDBERG Y, LEVY O. word2vec Explained: deriving Mikolov et al.'s negative-sampling word-embedding method[J]. arXiv preprint arXiv:1402.3722, 2014.
- [11] HAMILTON W, YING R, LESKOVEC J. Representation learning on graphs: methods and applications[J]. Data Engineering Bulletin, 017, 40(1):52-54.
- [12] PEROZZI B, AL-FOU R, SKIENA S. Deepwalk: Online learning of social representations[C]//Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining. 2014: 701-710.
- [13] GROVER A, LESKOVEC J. Node2vec: Scalable feature learning for networks[A]. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining[C]. San Francisco, CA, USA: ACM, 2016. 855-864.
- [14] DONG Y, CHAWLA N, SWAMI A. metapath2vec: Scalable representation learning for heterogeneous networks [A]. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining [C]. Halifax, NS, Canada: ACM, 2017. 35-144.
- [15] WANG D, CUI P, ZHU W. Structural deep network embedding[A]. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining [C]. San Francisco, CA, USA:ACM, 2016. 1225-1234.
- [16] CAO S, LU W, XU Q. Deep neural networks for learning graph representations[A]. Proceedings of the 33rd AAAI Conference on Artificial Intelligence[C]. Phoenix, Arizona, USA: AAAI, 2016. 1145-1152.
- [17] HAMILTON W, YING R, LESKOVEC J. Inductive representation learning on large graphs [A]. Proceedings of Advances in Neural Information Processing Systems [C]. Long Beach, CA, USA:MIT Press, 2017. 1024-1034.
- [18] TANG J, QU M, WANG M, et al. Line: Large scale information network embedding[A]. Proceedings of the 24th International World Wide Web Conference[C]. Florence, Italy: ACM, 2015. 1067-1077.
- [19] YANG C, SUN M, LIU Z, et al. Fast network embedding enhancement via high order proximity approximation[A]. Proceedings of the 26th International Joint Conference on Artificial Intelligence[C]. Melbourne, Australia: Morgan Kaufmann, 2017. 3894-3900.
- [20] BOURIGAULT S, LAGNIER C, LAMPRIER S, et al. Learning social network embeddings for predicting information diffusion[A]. Proceedings of the 7th ACM International Conference on Web Search and Data Mining[C]. New York, NY, USA: ACM, 2014. 393-402.[30] LI C, MA J, GUO X, et al. DeepCas: An end-to-end predictor of information cascades[A]. Proceedings of the 26th International Conference on World Wide Web[C]. Perth, Australia: ACM, 2017. 577-586.
- [22] BORDES A, USUNIER A, et al. Translating embeddings for modeling multirelational data[A]. Proceedings of the 27th Annual Conference on Neural Information Processing Systems[C]. Lake Tahoe, Nevada, United States: MIT Press, 2013. 2787-2795.
- [23] WANG Z, ZHANG J, FENG J, et al. Knowledge graph embedding by translating on hyperplanes[A]. Proceedings of the 28th AAAI Conference on Artificial Intelligence[C]. Québec City, Québec, Canada:AAAI, 2014. 1112-1119.
- [24] LIN Y, LIU Z, SUN M, et al. Learning entity and relation embeddings for knowledge graph completion[A]. Proceedings of the 29th AAAI Conference on Artificial Intelligence[C]. Austin, Texas, USA:AAAI, 2015. 2181-2187.
- [25] JI G, HE S, XU L, et al. Knowledge graph embedding via dynamic mapping matrix[A]. Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics[C]. Stroudsburg, PA: ACL, 2015. 687-696.
- [26] JI G, LIU K, HE S, et al. Knowledge graph completion with adaptive sparse transfer matrix[A]. Proceedings of the 30th AAAI Conference on Artificial Intelligence[C]. Phoenix, Arizona, USA: AAAI, 2016. 985-991.
- [27] HE S, LIU K, JI G, et al. Learning to represent knowledge graphs with Gaussian embedding[A]. Proceedings of the 24th ACM International Conference on Information and Knowledge Management[C]. Melbourne, VIC, Australia: CIKM, 2015. 623-632.
- [28] ZOU L, CHEN L, ZHAO D, et al. Answering pattern match queries in large graph databases via graph embedding[J]. VLDB Journal, 2012, 21(1):97-120.
- [29] RIBA P, LLAD S J, FORNS A, et al. Largescale graph indexing using binary embeddings of node contexts for information spotting in document image databases [j]. Pattern Recognition Letters, 2017, 87:203-211.