

# 超图学习综述: 算法分类与应用分析\*

胡秉德, 王新根, 王新宇, 宋明黎, 陈纯

(浙江大学 计算机科学与技术学院, 浙江 杭州 310007)

通信作者: 王新根, E-mail: newroot@zju.edu.cn



**摘要:** 随着图结构化数据挖掘的兴起, 超图作为一种特殊的图结构化数据, 在社交网络分析、图像处理、生物反应解析等领域受到广泛关注. 研究者通过解析超图中的拓扑结构与节点属性等信息, 能够有效解决实际应用场景中所遇到的如兴趣推荐、社群划分等问题. 根据超图学习算法的设计特点, 将其划分为谱分析方法和神经网络方法, 根据方法对超图处理的不同手段, 可进一步划分为展开式方法和非展开式方法. 若将展开式方法用于不可分解超图, 则很有可能会造成信息损失. 然而, 现有的超图相关综述文章鲜有就超图学习方法适用于哪类超图这一问题做出相关归纳. 因此, 分别从超图上的谱分析方法和神经网络方法两方面出发, 对展开式方法和非展开式方法展开讨论, 并结合其算法特性和应用场景作进一步细分; 然后, 分析比较各类算法的设计思路, 结合实验结果总结各类算法的优缺点; 最后, 对超图学习未来可能的研究方向进行了展望.

**关键词:** 超图学习; 谱分析; 神经网络; 展开; 非展开

**中图法分类号:** TP181

中文引用格式: 胡秉德, 王新根, 王新宇, 宋明黎, 陈纯. 超图学习综述: 算法分类与应用分析. 软件学报, 2022,33(2): 498–523. <http://www.jos.org.cn/1000-9825/6353.htm>

英文引用格式: Hu BD, Wang XG, Wang XY, Song ML, Chen C. Survey on Hypergraph Learning: Algorithm Classification and Application Analysis. Ruan Jian Xue Bao/Journal of Software, 2022, 33(2): 498–523 (in Chinese). <http://www.jos.org.cn/1000-9825/6353.htm>

## Survey on Hypergraph Learning: Algorithm Classification and Application Analysis

HU Bing-De, WANG Xin-Gen, WANG Xin-Yu, SONG Ming-Li, CHEN Chun

(College of Computer Science and Technology, Zhejiang University, Hangzhou 310007, China)

**Abstract:** With the rise of graph structured data mining, hypergraph, as a special type of graph structured data, is widely concerned in social network analysis, image processing, biological response analysis, and other fields. By analyzing the topological structure and node attributes of hypergraph, many problems can be effectively solved such as recommendation, community detection, and so on. According to the characteristics of hypergraph learning algorithm, it can be divided into spectral analysis method, neural network method, and other method. According to the methods used to process hypergraphs, it can be further divided into expansion method and non-expansion method. If the expansion method is applied to the indecomposable hypergraph, it is likely to cause information loss. However, the existing hypergraph reviews do not discuss that hypergraph learning methods are applicable to which type of hypergraphs. So, this article discusses the expansion method and non-expansion method respectively from the aspects of spectral analysis method and neural network method, and further subdivides them according to their algorithm characteristics and application scenarios. Then, the ideas of different algorithms are analyzed and compared in experiments. The advantages and disadvantages of different algorithms are concluded. Finally, some promising research directions are proposed.

**Key words:** hypergraph learning; spectral analysis; neural network; expansion; non-expansion

图(graph)作为一种高效的关系表达结构, 被广泛地应用于成对关系的建模中, 例如对论文引用关系、私人

\* 基金项目: 广东省重点领域研发计划(2020B0101100005); 浙江省重点研发计划(2021C01014)

收稿时间: 2020-08-07; 修改时间: 2020-09-30; 采用时间: 2021-04-17; jos 在线出版时间: 2021-05-20

社交、蛋白质交互反应等网络的建模. 然而除了成对关系外, 在很多场景中还存在大量一般简单图结构难以表达的非成对关系, 例如社交网络中存在的社区结构、特征关系中的簇结构等. 在这些场景中, 研究者很难甚至无法区分各类结构内部样本与样本之间的交互关系. 而超图具有的一条边内包含任意个数节点的特性, 使其对于这种数据关系的表达有着天然的优势. 具体来说, 超图(hypergraph)是一类一条边可以包含任意节点数量的图结构, 其形式化表达如下: 超图  $H=(X,E)$ , 其中,  $X$  是一组称为节点的元素,  $E$  是一组  $X$  的非空子集, 称为超边或边<sup>[1]</sup>. 图 1 即为一个简单的超图样例.

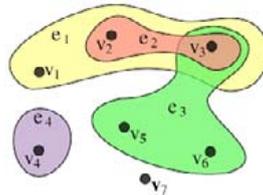


图 1 超图样例<sup>[1]</sup>

事实上, 学术界一直都在进行超图的相关研究, 但早期的研究焦点主要在传统图论问题上<sup>[2-4]</sup>. 随着超图理论的快速发展, 一些更为广泛的应用性问题才在超图上被有针对性地加以研究, 并逐步成为研究热点, 例如图节点分类问题<sup>[5]</sup>、图节点重要性排序问题<sup>[6]</sup>以及一些图像处理任务<sup>[7,8]</sup>. 而近年来, 神经网络研究的突飞猛进, 又使得超图学习有了新的研究方向. 研究者利用神经网络超强的学习能力以及模型本身的灵活性, 对超图的高阶交互关系<sup>[9]</sup>、动态超图<sup>[10]</sup>以及不可分解超图<sup>[11]</sup>等问题的研究正在逐步展开.

从方法层面来看, 谱分析方法作为一直以来图学习的传统主流方法, 在超图学习上也占据着主导地位. 在神经网络被引入到超图学习领域之前, 大部分研究都是建立在谱理论之上的. 但是随着神经网络被引入到超图学习领域, 一系列将传统成对图网络上的学习算法推广到超图上的研究成果被研究者提了出来<sup>[12-14]</sup>. 超图学习首次被作为一种在超图结构上的传播过程是在 Zhou 等人的工作<sup>[12]</sup>中被提出来的, Zhou 等人使用的就是在超图上的谱分析方法. 这类方法通常具有满足严密数理推导的解析过程, 但也正因为如此, 这类方法有很多的局限性.

与此同时, 神经网络的出现为超图学习研究注入了新的活力. 如何将神经网络结构用于超图学习, 成为了一个新的研究热点. 近年来也有很多这方面的优秀工作被先后提出, 其中, 代表性工作 hyperedge based embedding (HEBE)<sup>[15]</sup>方法希望通过超图来为每个事件中涉及的对象学习一个表征的思路为后续的神经网络方法带来了启发. Tu 等人<sup>[11]</sup>证明了 HEBE 方法在稀疏超图下的表现并不理想, 并提出了 deep hyper-network embedding (DHNE)模型. 然而, DHNE 方法直接使用多层感知机对元组关系进行建模, 使其只能应用于均匀超图(所有超边所包含的顶点数相同的图), 将 DHNE 应用于非均匀异构超图会消耗大量的计算资源, 并损失模型的泛化能力. 面对上述挑战, Zhang 等人<sup>[16]</sup>提出了 Hyper-SAGNN, 通过引入自注意力机制达到对非均匀异构超图进行学习挖掘的目的.

然而上述提到的文章在考虑超图各类特性的同时大多忽视了一个潜在的问题, 就是超图的可分解性. 图展开是超图分析的一类经典方法, 即把超边展开成普通边, 具体的又分为团式展开<sup>[5]</sup>、星式展开<sup>[9]</sup>、采样式展开<sup>[17]</sup>等. 但展开式方法需要建立在超图可分解的前提下, 即任意超边的子集能够组成超边. 针对不可分解超图, Tu 等人<sup>[11]</sup>提出了 DHNE 模型, 并从理论上证明: 现有方法中, 任何线性相似性度量在常用的嵌入空间中都无法维持超图网络的不可分解性. 因此, 针对模型算法的分解特性, 本文将对现有的超图学习方法进行梳理.

本文第 1 节介绍超图学习的研究背景并提出超图学习方法的基本分类. 第 2 节对超图学习相关概念及定义进行具体说明. 第 3 节对不同类别的超图学习方法进行归纳和比较. 第 4 节对超图学习的具体应用领域进行说明, 分析超图学习的重要意义. 第 5 节对具有代表性的超图学习算法进行实验比较与分析. 第 6 节与第 7 节对超图学习中未解决的问题和未来展望进行讨论并对本文进行总结.

## 1 背景知识

超图结构是一种比传统成对关系图更加泛化的关系类型. 在过去的几十年间, 超图理论已被证明能够有效解决众多真实世界问题. 超图强大的表达能力, 使其能够很好地对各类网络、数据结构、进程调度以及一些对象关系复杂的系统进行高效建模. 从理论上来看, 超图可以归纳普通图上的某些定理, 甚至可以用一个超图定理代替传统图上的几个定理. 例如, Berge 的弱完全图猜想(一个图是完美的当且仅当它的补图是完美的)就是使用超图概念进行证明的<sup>[18]</sup>. 从实用性角度来看, 超图结构正逐渐变得比传统图结构更受欢迎. 因此, 有关如何从构建的超图结构中学习获取有用的结构信息的研究和实践, 就变得十分迫切且有意义.

超图研究早在 2000 年以前就已被众多学者所关注. 超图理论最早是由 Berge<sup>[19]</sup>在 1987 年提出并发展而来的, 是一种有限组合集理论, 它模拟了许多运筹学和组合优化问题. 那时的主要研究焦点聚集在超图分割<sup>[1,2,20]</sup>、超图最小横截面<sup>[21,22]</sup>以及传统图论理论在超图上的推广<sup>[23]</sup>等传统图论研究方向上. 1993 年, Chung<sup>[23]</sup>提出了  $k$  介同构超图上的拉普拉斯矩阵的定义, 将传统成对图网络上的频谱理论推广到了  $k$  介同构超图上. 在 Chung<sup>[23]</sup>工作的启发下, 很多衍进方法被陆续提出<sup>[9,24]</sup>. 1995 年, Thoms 等人<sup>[4]</sup>定义了超图上的最小截面问题. 1997 年, Gunopulos 等人<sup>[22]</sup>首次在文章中提出超图的截面问题可以转化为在目标数据中查找最大特定语句问题这一观点, 并将其与学习理论即机器学习的概念联系起来. 1999 年, George 等人<sup>[2]</sup>将超图分割算法应用于超大规模集成电路(VLSI circuits), 这一工作受到了学者们的广泛关注.

随着时间的推移, 超图学习的研究方向不再局限于传统图论问题, 超图上的聚类<sup>[25,26]</sup>、标签分类<sup>[5]</sup>以及超图在图像领域的应用<sup>[27,28]</sup>等问题的研究工作开始逐渐增多. 但这些工作大多是在低阶、均匀或是同构超图上开展的, 对于数据的复杂程度有较大限制, 无法适应大数据场景.

随着近年来图结构数据数据量的爆炸式增长以及神经网络模型的快速衍变, 对于异构、高阶、非均匀超图的挖掘学习成为了新的研究热点. 2016 年, Huan 等人<sup>[15]</sup>提出了 HEBE 模型, 将事件与参与对象之间的关系构建成异构超图进行表征学习. 2018 年, Ke 等人<sup>[11]</sup>提出了 DHNE 模型, 这一模型保证了超图结构的完整性, 解决了异构、不可分解超图的表征学习问题, 并证明了常用嵌入空间的线性相似性度量无法维持超图网络中的不可分解性. 2019 年, Zhang 等人<sup>[16]</sup>针对 DHNE 无法处理高阶、非均匀、异构超图的问题提出了 HyperSAGNN 模型. 相较于 DHNE 模型, 该模型泛化能力更强. 此外, 受图卷积神经网络启发, Naganand 等人<sup>[29]</sup>提出了 HyperGCN 模型, 该模型不仅能够挖掘超图的高阶交互关系, 而且模型中还引入了超图节点的特征属性.

超图学习方法种类繁多, 但根据超图学习方法自身特点和所适用的超图结构特点, 超图学习方法可分为:

- 谱分析超图学习方法: 以超图拉普拉斯矩阵为核心的一类矩阵分析方法, 通常具有满足严密数理推导的目标函数最优求解过程. 同时, 相关的理论基础对很多算法模型的设计有着重要的指导意义. 这类方法因为理论限制, 通常泛化性较差, 能够适用的超图结构有限. 并且, 该类超图谱分析方法通过矩阵分解的方式求解, 无法直接应用于大规模超图挖掘场景;
- 神经网络超图学习方法: 以人工神经网络及其衍生结构为模型主体结构的超图学习方法. 这类方法通常学习能力强、结构设计灵活且泛化性高, 很好地弥补了谱分析类方法的缺陷. 将神经网络用于超图学习, 成为近年来一个新的研究热点;
- 其他方法: 在超图学习场景中还有一些既不像谱分析方法那样具有严格的解析解, 同时也不具备神经网络结构的机器学习方法. 这类方法通常在特定的应用场景下结合场景特点定义模型目标, 并通过相应的优化方法求解. 由于相关方法场景特异性强, 本文只给出简要说明.

在谱分析超图学习方法和神经网络超图学习方法中, 又可进一步地划分为:

- 展开式超图学习方法: 展开式超图学习方法针对于可分解超图, 这一类方法将会对超图进行拆解, 从而将超图转化为传统成对图网络. 因此, 这类方法需要超图结构具备可分解性. 这类超图学习方法的普适性较强, 但在对超图进行拆解的同时, 容易造成信息损失;
- 非展开式超图学习方法: 非展开式超图学习方法针对于不可分解超图, 这类方法无需对超图进行拆

解, 而是直接对超图结构进行建模. 在不可分解超图中, 超边中的点通常具有很强的关联性, 而其超边中节点的子集则没有这种强关联性. 例如, 在推荐系统中, 用户、电影、标签三者形成的超边, 用户和标签之间通常就不存在强关联<sup>[11]</sup>. 若对这类超图使用展开式超图学习方法, 则会导致超图信息损失或引入额外噪声.

## 2 超图学习相关概念及定义

- 超图结构

超图  $G$  可表示为  $G(V, E)$ , 其中,  $V = \{v_1, \dots, v_n\}$  表示超图中的  $n$  个节点集合,  $E = \{E_1, \dots, E_m\}$  表示超图中的  $m$  条超边集合.  $E_\alpha$  表示超边  $\alpha$ , 是一个无序节点集合.  $|E_\alpha|$  为超边  $\alpha$  的尺寸, 即集合中点的个数. 当任意超边  $\alpha$  均符合  $|E_\alpha| = 2$  时, 则超图退化为普通图. 超图的关联矩阵  $H$  定义如下:

$$h_{i\alpha} = \begin{cases} 1, & i \in E_\alpha \\ 0, & \text{otherwise} \end{cases}$$

紧接着上述定义, 超图的邻接矩阵  $A$  可定义为  $A = HH^T$ , 其中,  $A_{ij}$  表示节点  $i$  和  $j$  共享超边的数量. 同时, 我们可以定义超边矩阵  $C$  为  $C = H^T H$ , 其中,  $C_{\alpha\beta}$  表示超边  $\alpha$  和超边  $\beta$  共有的节点数量.

- 拉普拉斯矩阵

拉普拉斯矩阵通常用于图的矩阵化表示, 可用于发现许多图的有用性质. 传统成对关系图的拉普拉斯矩阵通常定义为

$$L = D - A,$$

其中,  $D$  为图的度矩阵,  $A$  为图的邻接矩阵. 拉普拉斯矩阵存在对称  $L^{\text{sym}}$  和  $L^{\text{rw}}$  随机游走两类标准化形式:

$$L^{\text{sym}} = D^{-\frac{1}{2}} L D^{-\frac{1}{2}}, \quad L^{\text{rw}} = D^{-1} L.$$

- 展开

将超图转化为传统成对图网络的过程, 团式展开<sup>[5]</sup>和星式展开<sup>[9]</sup>是两种具有代表性的展开思路.

- 异构超图

传统超图结构中, 节点与边的种类只有一种; 而异构超图中, 节点或边的种类在两种以上(目前公开的异构超图数据集中大多为节点异构类型).

- 不可分解超图

不可分解超图通常是异构超图, 这类超图中, 同一超边上的节点有较强的关联关系, 但超边子集中的节点则很可能缺少强关联. 例如, 推荐系统中, 用户、电影、标签形成的三元组能够有效反映用户对电影的评价三者间有强关联关系, 但用户和标签所形成的子集则关联关系较弱<sup>[11]</sup>. 若对这类超图使用可分解超图学习方法, 则会导致超图信息损失或引入额外噪声.

- 结构属性

结构属性是网络结构中最为重要的信息, 通过结构属性, 我们可以构建网络节点之间的拓扑关系. 有效捕捉这种拓扑关系, 是很多图学习方法的研究动机. 结构属性可分为局部结构和全局结构<sup>[11]</sup>: 局部结构是指节点之间的直接邻接关系, 可在超图结构的邻接矩阵中直接观测到, 在一些文献中也被称为一阶关联信息; 全局结构是指除节点一阶关联信息外的高阶结构信息, 在文献中谈及较多的是二阶关联关系. 对于节点  $v_i$  属于超边  $E_i$ ,  $E_i/v_i$  表示节点  $v_i$  的邻居节点. 当  $E_i/v_i$  与  $E_j/v_j$  相似时, 我们认为节点  $i$  和节点  $j$  的二阶关联关系相近.

- 其他属性

除了超图中的拓扑结构信息以外, 超图中还含有其他有价值的信息, 如节点的属性信息、边的属性信息等. 将这些拓扑结构信息以外的信息引入到模型中, 可以让模型从更多的角度去挖掘超图上的潜在相似关系.

### 3 超图学习方法

随着图结构化数据挖掘的兴起,超图作为一种特殊的图结构化数据,也受到研究者广泛研究.本文采用层次化的分类方法对超图学习方法进行分类,首先将超图学习方法分为谱分析方法、神经网络方法以及其他方法,再进一步根据方法建模思路的不同将其分为展开式方法和非展开式方法,具体分类以及代表模型如图 2 所示.本节在讨论具体方法时,将着重于讨论算法的创新点和优劣性.

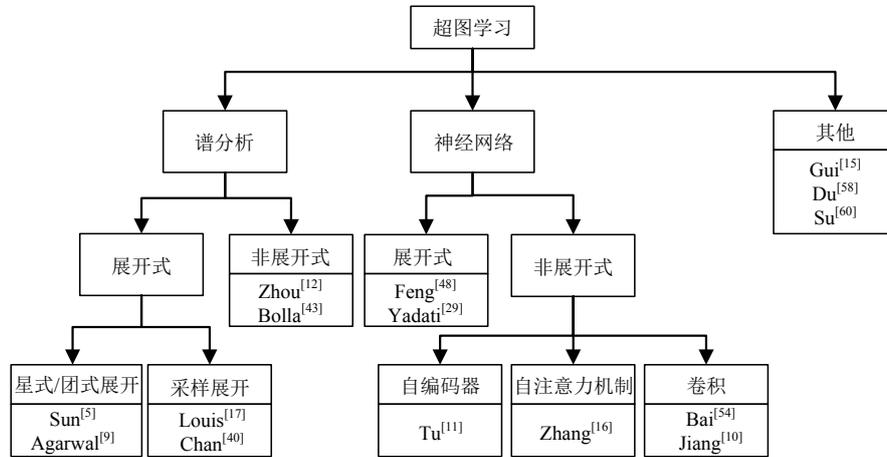


图 2 超图模型分类

#### 3.1 谱分析方法

超图上的谱分析方法是超图学习的传统主流方法,它是一类以谱理论为基础的矩阵分析方法.这类方法通常具有满足严密数理推导的目标函数最优解求解过程,相关的理论基础对很多算法模型的设计有着重要的指导意义.近年来炙手可热的 graph convolutional network (GCN)<sup>[30,31]</sup>就是在谱理论基础上结合深度学习逐步发展而来的.本节将着重讨论超图上的谱分析方法.

##### 3.1.1 展开式谱分析方法

超图上的展开式谱分析方法一般通过将超图转化成传统成对图网络的方式,将超图学习问题简化成传统成对图网络学习问题,再根据拉普拉斯矩阵的谱特性求解.这类方法大多在同构超图上展开研究.

###### (1) 星式展开和团式展开

星式展开(SE)和团式展开(CE)<sup>[9]</sup>是两种比较经典的超图展开式方法,这两种方法被借鉴应用于很多超图算法设计中.最早的标准星式展开和团式展开是在谱理论的基础上提出来的<sup>[5,9,32]</sup>.标准星式展开定义如下<sup>[9]</sup>.

星式展开算法为每个超图  $G(V, E)$  建立一个图  $G^*(V^*, E^*)$ . 图  $G^*(V^*, E^*)$  为超图  $G(V, E)$  上每条超边引入一个节点  $e \in E, V^* = V \cup E$ <sup>[32]</sup>. 每个节点  $e$  连接属于  $e$  所在超边内的所有节点, 即  $E^* = \{(u, e) : u \in e, e \in E\}$ . 值得注意的是: 集合  $E$  中的每条超边在图  $G^*(V^*, E^*)$  中都有一个对应的星状结构, 且  $G^*$  是一个成对图网络. 星式展开将超边权重重新分配给每个对应的普通成对图边:

$$w^*(u, e) = \frac{w(e)}{\delta(e)} \quad (1)$$

其中,  $\delta(e)$  为超边  $e$  中节点个数. 在获得了展开后的成对关系图  $G^*$  后, 我们可以得到  $G^*$  的标准化拉普拉斯矩阵  $L^*$  定义如下<sup>[9]</sup>:

$$L^* = \begin{bmatrix} I & -A \\ -A^T & I \end{bmatrix} \quad (2)$$

其中,  $A$  是一个  $|V| \times |E|$  矩阵:

$$A_{ue} = \frac{h(u,e)w^*(u,e)}{\sqrt{d^*(u)}\sqrt{d^*(e)}} \quad (3)$$

其中,  $d^*(u)$ 和  $d^*(e)$ 分别为超图节点  $u$  和新添超边节点  $e$  在  $G^*$ 上所有连接边权重之和, 其形式化表达对于  $u \in V$  为  $d^*(u) = \sum_{e \in E} h(u,e)w^*(u,e)$ , 对于  $e \in E$  为  $d^*(e) = \sum_{u \in V} h(u,e)w^*(u,e)$ . 星式展开示意图如图 3(a)所示.

团式展开算法<sup>[9]</sup>为原始超图  $G(V,E)$ 构建图  $G^c(V^c,E^c)$ , 原始超图中的每条超边  $e=(u_1,\dots,u_{\delta(e)}) \in E$  在图  $G^c$  中都被边中节点两两相连形成的全连接结构所替代,  $E^c=\{(u,v):u,v \in e, e \in E\}$ <sup>[32]</sup>. 同样需要注意的是, 每条超边在图  $G^c$  中都形成了一个团状结构. 新构建的图  $G^c$  中,  $u$  和  $v$  所形成边的权重  $w^c(u,v)$ 需要满足  $w^c(u,v)$ 与  $u$  和  $v$  所在的原始超边  $e$  的权重差值最小, 即  $w^c(u,v) = \arg \min_{w^c(u,v)} \sum_{e \in E: u,v \in e} (w^c(u,v) - w(e))^2$ . 因此, 团展开后,  $G^c$  中团的每条边的权重与原始超边  $e$  的权重  $w(u)$ 相关. 具体可表示为

$$w^c(u,v) = \mu \sum_{e \in E: u,v \in e} w(e) = \mu \sum_e h(u,e)h(v,e)w(e) \quad (4)$$

其中,  $\mu$ 是一个固定的标量. 同样, 得到图  $G^c$  后, 团式展开的标准化拉普拉斯矩阵  $L^c$  有如下定义<sup>[5]</sup>:

$$L^c = I - D_c^{-\frac{1}{2}} H W H^T D_c^{-\frac{1}{2}} \quad (5)$$

其中,  $D_c$ 是  $G^c$  的度矩阵,  $H$  是原始超图的关联矩阵,  $W$  是由  $w^c$  所组成的权重矩阵. 其示意图如图 3(b)所示.

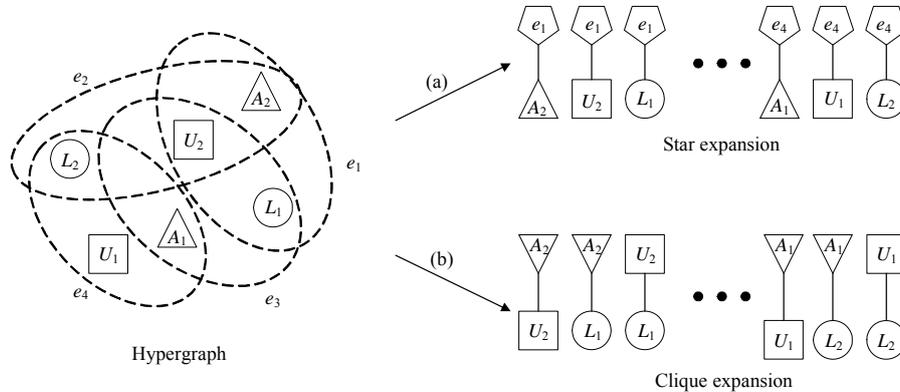


图 3 星式展开和团式展开

有了展开图的标准拉普拉斯矩阵后, 将其带入最小二乘目标函数<sup>[32]</sup>:  $\min tr(XL^T X)$ , subject to  $X^T X = I$ , 问题就转化成了传统图上的切割问题. 根据拉普拉斯矩阵的谱特性, 该目标函数的解即为拉普拉斯矩阵除 0 以外最小特征值对应的特征向量.

两种展开方式的不同点在于: 团式展开中, 同一条超边的节点直接相连具有明确的相似性; 而星式展开中, 同一条超边内节点则是通过超边节点间接相连, 具有隐性相似性. 相同点在于, 两种展开方式都是一个将超图转化为成对图的过程. 在  $k$  阶均匀超图中, 将星式展开图的权重扩大  $(\delta(e)-1)\delta(e)$  倍后获得的成对关系图  $G_{ex}^*(V^*, E^*)$ ,  $G_{ex}^*$  的特征向量恰好是标准化团式展开  $G_c$  的特征向量. 虽然在均匀超图上星式展开和团式展开展开后的成对关系图结构有很大的不同, 但两种展开图的拉普拉斯矩阵的特征分解结果却有着数理相似性. 在非均匀超图中, 星式展开和团式展开有所不同: 团式展开相较于星式展开给予较大的超边更多的权重, 这也将直接导致后续特征分解的差异性<sup>[9]</sup>.

在标准星式展开和团式展开的影响下, 一些以它们为基础的变体展开方式被提了出来.

Yu 等人<sup>[33]</sup>在基于类内散布矩阵的聚类算法中, 提出使用  $A=HW_H(D_e-1)^{-1}H^T$  为超图的展开矩阵, 而后使用拉普拉斯矩阵  $L = I - D_v^{-1/2} A D_v^{-1/2}$  的特征向量作为节点表征进行聚类. 该过程实际上就是一种团式展开的简单变体, 与标准团式展开不同的是, 其权重分配函数如下:

$$w(u, v) = \sum_{e \in E: u, v \in e} \frac{w(e)}{d_e - 1} \quad (6)$$

其中,  $d$  为对应超边的度. 在此基础上提出的 HCIS 算法能够有效分析高阶生物模块的交互关系.

Rodriguez's Laplacian (RL)<sup>[34,35]</sup> 从一个无权超图  $G(V, E)$  构建一个有权图  $G'(V, E'=E^c)$ .  $G'$  为在超图  $G$  上做团式展开后获得的成对关系图, Rodriguez 结合超图结构对  $G'$  的拉普拉斯矩阵给出如下定义:

$$L(G') = D'_v - HH^T \quad (7)$$

其中,  $D'_v$  是图  $G'$  的度矩阵. 不难发现: 将上述拉普拉斯矩阵公式左右乘上  $D_c^{-\frac{1}{2}}$ , 其形式与团式展开的拉普拉斯矩阵极为相似. 实际上, Rodriguez 就是一种将所有超边权重都置为 1 的团式展开. Rodriguez 等人<sup>[34,35]</sup> 在工作中证明了该拉普拉斯矩阵  $L'$  在无权超图上的谱特性以及对于超图的最小代价分割问题的价值.

类似地, Agarwal 等人<sup>[36]</sup> 在团式展开的基础上提出了 clique averaging (CA). 他们认为, 超图的每条超边在展开前后的权值应该都是相等的. 因此, 他们提出以线性函数的变体为基函数对权重分配函数进行建模. 并且, 在使用 Clique Averaging 方法构建图  $G$  后, 使用图  $G$  标准拉普拉斯矩阵的前  $p$  个特征向量对节点进行表征, 然后通过  $k$ -means 聚类算法对节点表征向量进行聚类, 最终验证了该方法在图分割上的实验效果. 通过比较 clique averaging 和 clique expansion 的权重分配函数可以发现: 实际上, clique expansion 的权重分配函数是 clique averaging 的一个低通版本, 两个函数都在解决相同的权重逼近问题.

实际上, 上述几种变体的不同主要集中在展开后超图的权重分配函数上, 权重分配函数不同, 将会直接影响拉普拉斯矩阵的构建结果. 类似的研究还有很多<sup>[37,38]</sup>, Huang 等人<sup>[39]</sup> 更是专门针对超图展开后如何分配权重这一问题对星式展开、团式展开、clique expansion 等算法做出了进一步展开讨论.

## (2) 采样展开

星式展开和团式展开虽然有效, 但其缺点也非常明显: 展开图的边数量远高于原图. 因此, 这些方法在阶超图中难以应用. 针对这一问题, 一些基于采样的展开方式被提了出来<sup>[17,40]</sup>.

Louis<sup>[17]</sup> 在算法 Markov clustering (MCL) 中提出了 breaking ties randomly (BTR) 的展开方法, 其核心思想是: 每次生成展开图时, 使用超边中表征向量距离最远的两个点的连接边代替超边本身, 其权重取原超边的权重值. 通过这种采样展开方式, 可以不断优化同一超边内节点的表征, 缩小超边内各节点表征的空间距离. 这种采样展开方法的优点很明显: 该方法采样过程计算简单, 且在投影后大大降低了展开图的复杂度. 但是, BTR 这种展开方式只考虑了节点表征空间距离最大的节点之间的关系, 却忽视了超边中其余节点对超边关系的影响. 这使得该方法在应用于节点差异性较大的超图时, 经常会丢失较多的中间信息.

Chan 等人<sup>[40]</sup> 针对 MCL 算法中每条超边信息流仅从高密度节点流向低密度节点、忽视了具有严格中间密度的中间节点这一问题, 从扩散原理出发提出了将中间节点视为 mediator 的超图展开方法. 对于每条超边  $e$ , 节点  $j \in e$  可以视作一个将节点  $S_e(f)$  中的信息流传递到节点  $I_e(f)$  的 mediator. 节点  $S_e(f)$  和节点  $I_e(f)$  即为 BTR 中, 空间距离相差最大的两节点. Mediator 结构本质上也是一个采样展开过程, 相比于团式展开, 它抛弃了中介节点之间的二元边关系, 仅保留了这些节点与高密度节点和低密度节点之间通信的中间关系. Chan 等人证明: 经过展开后, 最终获得的拉普拉斯矩阵的第二特征值仍满足 Cheeger's 不等式<sup>[41,42]</sup>. 因此, mediator 结构具有广泛的应用价值.

### 3.1.2 非展开式谱分析方法

与展开式方法不同, 非展开式谱分析方法直接对超图进行建模, 即直接构建超图上的拉普拉斯矩阵, 这一建模过程保证了超图信息的完整性.

图 4 所示为一个同构共同作者网络<sup>[12]</sup>, 左侧为团式展开后的投影图, 右侧为超图. 图中作者集为  $E=\{e_1, e_2, e_3\}$ , 文章集为  $V=\{v_1, v_2, v_3, \dots, v_7\}$ . 如图 4 所示: 在展开图中, 我们无法判断一个作者是否同时参与了 3 篇或以上的文章. 很明显, 展开图中丢失了一部分重要的信息.

Carletti 等人<sup>[6]</sup> 在其 random walks on hypergraphs (RWH) 的探索工作中证明, 他们所提出的在超图上直接建模的随机游走方法和在其投影图(团式展开后的成对关系图)上的传统随机游走算法的运算结果有很大的不

同. 通过在一些构造图数据上的算法演绎, Carletti 总结出, 他们提出的超图随机游走算法相较于传统方法对高阶结构更加敏感. Carletti 等人<sup>[6]</sup>的工作也从侧面说明, 在展开图上设计的展开式算法和直接在超图上设计的非展开式方法在性能表现上存在一定的差异.

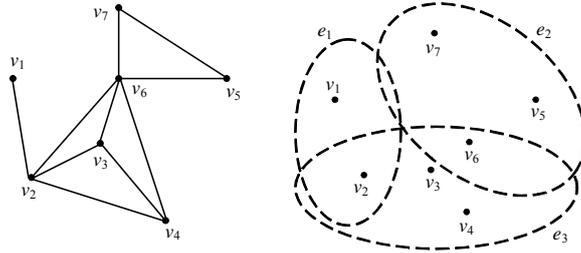


图4 文章共同作者<sup>[12]</sup>

Bolla's Laplacian 方法<sup>[43]</sup>是早期几种非展开式矩阵分解方法之一. Bolla 希望在联通超图上找到节点的  $k$  划分, 使得对应切割集中的边尽可能地少. 以此为目标, Bolla 在无权超图上定义了一个超图拉普拉斯矩阵  $L^o$ , 定义如下:

$$L^o = D_v - HD_e^{-1}H^T \tag{8}$$

其中,  $D_v$  是节点度组成的对角矩阵,  $D_e$  是超边度组成的对角矩阵,  $H$  是超图的关联矩阵.  $L^o$  的特征向量定义了超图节点的最佳欧式距离嵌入. Bolla 证实, 基于  $L^o$  的矩阵分解方法可用于解决超图的最小切割问题.

随后, Zhou 等人<sup>[12]</sup>将传统无向图上的矩阵分解方法推广到了超图上, 提出了 N-cut 算法在超图上的定义, 进一步推动了超图上的非展开式矩阵分解方法的发展. 他们定义了超图划分的形式化表达. 作为一个 NP-complete 问题, Zhou 等人将其放宽到实值优化问题后引入谱理论, 类比于传统图的拉普拉斯矩阵, 定义了超图拉普拉斯矩阵  $\Delta$ :

$$\Delta = I - D_v^{-\frac{1}{2}}HWD_e^{-1}H^TD_v^{-\frac{1}{2}} \tag{9}$$

根据标准线性代数理论, 超图最小切割问题的解即为拉普拉斯矩阵  $\Delta$  的最小非零特征值所对应的特征向量, 因而可以用矩阵分解的方法来解决超图的最小分割问题. 将 Bolla 的拉普拉斯矩阵左右乘上  $D_v^{-\frac{1}{2}}$  后, 与 Zhou 的拉普拉斯矩阵进行比较, 不难发现: Zhou 实际上是将 Bolla 的方法从无权图上推广到了有权超图上, 其矩阵本质并未发生改变, 只是多了一步矩阵标准化计算, 但这一推广使得 Zhou 的拉普拉斯矩阵更具普适性. Zhou 等人<sup>[12]</sup>将相关理论推广到了超图的嵌入、聚类以及传导推理等方面. 此后, 作为经典算法, Zhou 等人<sup>[12]</sup>的工作被众多研究者争相借鉴使用.

Huang 等人<sup>[7]</sup>在处理图像分割问题时, 就在以过分割图像块作为顶点的超图上使用了 Zhou 等人<sup>[12]</sup>的 N-cut 算法对超图进行了分割, 最终解决了图像的过分割问题. Purkait 等人<sup>[44]</sup>在设计大规模超边上的聚类算法时, 在获得了聚类超图后, 同样使用了 N-cut 算法对超图进行分割聚类. Huang 等人<sup>[45]</sup>将 Zhou 等人<sup>[12]</sup>工作中提出的传导推理理论与超图上的概率矩阵相结合, 提出了基于概率的非展开式超图排序矩阵分解算法. Ma 等人<sup>[46]</sup>在 Zhou 等人<sup>[12]</sup>提出的超图标准拉普拉斯算子的基础上提出了其非线性推广 Hypergraph p-Laplacian regularization (HpLapR)来保留数据的几何概率分布.

此外, 在谱分析方法中还有一类超图随机游走算法. 这类算法是以随机游走算法为基础的一类谱分析方法. Huang 等人<sup>[47]</sup>基于超图上的有偏随机游走策略提出了 Hyper2vec 算法, Hyper2vec 算法中, 节点的概率转移公式为

$$p'_2(x|v,u) = \frac{\alpha(x|v,u) \cdot \beta(x) \cdot \sum_{e \in E} w(e) \frac{h(v,e)}{d(v)} \frac{h(x,e)}{\delta(e)}}{Z} \tag{10}$$

其中,  $h(v,e)$ 和  $h(x,e)$ 是超图关联矩阵  $H$  中的元素,  $w(e)$ 为超边权重,  $d(v)$ 为节点  $v$  的度,  $\delta(e)$ 为超边  $e$  的度,  $\alpha(x|v, u)$ 和  $\beta(x)$ 都是分段函数, 对随机游走过程起导向作用. 上述公式的矩阵化表达为  $P = ABD_v^{-1}HWD_e^{-1}H^T/Z$ . 从中我们可以看出, 矩阵化表达  $P$  实际上就是 Zhou 等人<sup>[12]</sup>提出的超图拉普拉斯矩阵的随机游走形式乘上导向函数. Hyper2vec 在 Zhou 等人<sup>[12]</sup>提出的拉普拉斯算子上加上了导向函数, 使得方法本身能够适应不同的网络结构, 从而更好地保留网络的结构和固有属性.

近来, Carletti 等人<sup>[6]</sup>又提出了一种在超图上由广义拉普拉斯算子驱动的超图随机游走算法. 这种算法借鉴模拟了微观物理学模型中, 在同一超边中的多体近邻更容易交换的特性. 该方法构建的拉普拉斯算子等价于超边权重分配函数为  $w_{\bar{a}}=C_{aa}(C_{aa}-1)$  的 Zhou 等人<sup>[12]</sup>提出的超图拉普拉斯算子, 即, Carletti 等人提出的随机游走算法可以说是超图 Zhou 等人工作的一个特例. 相比于传统成对关系图的随机游走, 该方法对高阶超边更加敏感, 即: 相比于传统方法, 在超边规模发生改变时, 该方法计算出的超边内节点重要性变化更大.

谱分析相关算法的对比总结见表 1.

表 1 谱分析方法总结

算法名称	展开方式	权重分配	拉普拉斯矩阵	超图类型	是否展开	优势	劣势
SE	SE	$w^*(u, e) = \frac{w(e)}{\delta(e)}$	$L^* = \begin{bmatrix} I & -A \\ -A^T & I \end{bmatrix}$ $A_{we} = \frac{h(u, e)w^*(u, e)}{\sqrt{d^*(u)}\sqrt{d^*(e)}}$	有权同构	是	能够有效将超图转化成普通图处理	同一超边上的点为隐式关联
CE	CE	$w^c(u, v) = \mu \sum_{e \in E, v \in e} w(e)$ $= \mu \sum_e h(u, e)h(v, e)w(e)$	$L^c = I - S^c$ $S^c = D_c^{-\frac{1}{2}}HWH^T D_c^{-\frac{1}{2}}$	有权同构	是	能够有效将超图转化成普通图处理	展开后会造造成数量的膨胀
HCIS	CE	$w(u, v) = \sum_{e \in E, u, v \in e} \frac{w(e)}{(d_e - 1)}$	$A = HW_H(D_e - I)^{-1}H^T$ $L = I - D_v^{-1/2}AD_v^{-1/2}$	有权同构	是	有效分析高阶模块的交互关系	本质上是团式展开, 因而仍存在展开后边的膨胀问题
RL	CE	$w^r(u, v) =  \{e \in E, u, v \in e\} $	$L'(G^r) = D_v^r - HH^T$	无权同构	是	为无权超图定义了拉普拉斯矩阵并证明其有效性	本质上是团式展开, 因而仍存在展开后边的膨胀问题
CA	CE	$h(z) = \left( \lambda(k) \sum_{v_i, v_j, i < j} g^p(v_i, v_j) \right)^{\frac{1}{p}}$	-	有权同构	是	保证了超图展开前后权重的一致, 使得权重分配函数更加合理	权重分配函数的拟合增加了额外的计算开销
BKR	BKR	$w(i_e, j_e) = w(e)$ or $w(e)/(  S^e   R^e   )$	$L_x = I - A_x$ $A_x = A_x D_x^{-1}$	有权同构	是	大大降低了展开后节点和边的数量, 降低了后续计算的消耗	放弃了过多的中间节点, 丢了较多中间信息
Mediator	Mediator	$w(u, v) = \begin{cases} w_e \cdot \beta_0^e \cdot \max_{i, j \in e} (f_i - f_j) \\ w_e \cdot \beta_j^e \cdot (\max_{s \in e} f_s - f_j) \\ w_e \cdot \beta_j^e \cdot (f_j - \min_{i \in e} f_i) \end{cases}$	$L^m = I - D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$	有权同构	是	降低了展开后边的数量, 且保留了大部分的中间信息	作为一种采样展开方式, 仍会有一定的信息损失
Bolla	-	-	$L^o = D_v - HD_e^{-1}H^T$	无权同构	否	能够有效解决无权图上的最小切割问题	只可应用于无权图, 迁移应用性较差, 且只可应用于连通图

表 1 谱分析方法总结(续)

算法名称	展开方式	权重分配	拉普拉斯矩阵	超图类型	是否展开	优势	劣势
Zhou	-	-	$L^z = I - D_v^{-\frac{1}{2}} H W D_e^{-1} H^T D_v^{-\frac{1}{2}}$	有权同构	否	解决了有权图上的最小切割问题, 具有较强的迁移性, 被众多工作所借鉴	超图结构中 存在孤立 节点时, 拉普拉斯 算子将失效
Hyper2vec	-	-	$L^2 = \frac{A B D_v^{-1} H W D_e^{-1} H^T}{Z}$	有权同构	否	导向函数使得方法能够更好地保留网络的结构和固有属性	导向函数 为分段函数 较为简单
RWH	-	-	$L_{ij} = \delta_{ij} - \frac{k_{ij}^H}{\sum_l k_{il}^H}$	无权同构	否	模拟了微观物理学多体近邻更易交换的特性, 对高阶超边变化敏感	仅可应用于 无权超图

### 3.2 神经网络方法

近年来, 随着神经网络研究的逐渐深入, 研究者将神经网络方法引入到了各个研究领域, 其中就包括了超图学习. 谱分析类方法的优点在于其方法本身拥有较强的数理可解释性, 但也正因如此, 谱分析类方法的灵活性较差, 很多方法适用的超图场景十分有限. 此外, 由于谱分析类方法的求解特性, 很多谱分析方法往往无法直接应用于大规模超图挖掘场景. 而神经网络算法的出现, 则很好地弥补了这些缺陷. 本节将对超图上的神经网络方法进行展开讨论.

#### 3.2.1 展开式神经网络方法

Feng 等人<sup>[48]</sup>受图卷积神经网络的启发, 提出了 hypergraph neural networks framework (HGNN), 将 Zhou 等人<sup>[12]</sup>提出的超图拉普拉斯矩阵  $A=I-\Theta$  带入到传统图卷积神经网络<sup>[31]</sup>中, 定义了超图上的图卷积计算  $g*x = \Phi g(A)\Phi^T x$ . 为了加速计算, Feng 等人借鉴了 Defferrard 等人<sup>[31]</sup>的经验, 使用二阶切比雪夫不等式求得  $g*x$  的近似表达, 最终超图上的图卷积计算可表达为如下形式:

$$Y = D_v^{-\frac{1}{2}} H W D_e^{-1} H^T D_v^{-\frac{1}{2}} X \Theta \tag{11}$$

其中,  $X \in R^{n \times C_1}$  表示节点特征矩阵;  $n$  为节点数;  $C_1$  为节点特征维度;  $W = \text{diag}(w_1, \dots, w_n)$ , 代表超边权重;  $\Theta \in R^{C_1 \times C_2}$  是学习参数;  $Y \in R^{n \times C_2}$  为层输出向量, 可用于分类. 模型整体卷积层设计如图 5 所示.

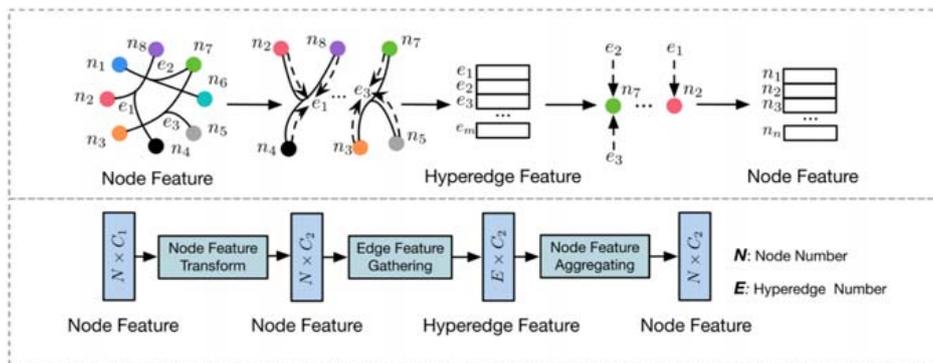


图 5 超边卷积层图示<sup>[48]</sup>

实际上, 上述图卷积计算过程就是在模拟前文提到的星式展开过程的一个神经网络推广,  $D_e^{-1} H^T D_v^{-\frac{1}{2}} X \Theta$  部分计算就是在将原始节点信息聚合到边节点上,  $D_v^{-\frac{1}{2}} H W$  部分的计算则是重新将获得边节点信息放回

始节点上. 相比于原始的星式展开算法, HGNN 模型引入了可拟合参数, 使得模型的学习能力更强, 能够对超图的结构属性信息和节点特征信息进行更好的融合. 但是由于该卷积计算的拉普拉斯算子需要求取  $D_v^{-\frac{1}{2}}$  与  $D_e^{-1}$ , 因此在超图结构存在孤立点时, 该算子将会失效. 而这一情况在社交网络数据中较为常见.

受 HGNN 模型的启发, Ji 等人<sup>[49]</sup>提出了 DHCF 双通道协同过滤算法. 该算法通过其定义的 JHConv 结构, 分别对用户和项目的结构信息进行了提取, 并将提取后的两个表征矩阵通过一个共享权重矩阵关联起来, 最终获得项目和用户的表征矩阵. 其中, JHConv 结构的定义如下:

$$X^{(l+1)} = \sigma \left( D_v^{-\frac{1}{2}} H D_e^{-1} H^T D_v^{-\frac{1}{2}} X^{(l)} \Theta^{(l)} + X^{(l)} \right) \quad (12)$$

JHConv 的计算过程通过在卷积过程中额外引入上一层的表征结果  $X^{(l)}$ , 加速了模型的收敛速度, 并且这种双通道结构在达到协同过滤目的的同时, 也保持了用户和项目各自的表征特点. 但模型中仅通过共享权重的方式来构建不同对象之间的关联关系, 使得不同对象之间的联系完全依赖于数据质量, 这将难以应对噪声数据.

Yi 等人<sup>[50]</sup>更是将 HGNN 与 RNN 相结合, 提出了 HGC-RNN. 该方法用图卷积神经网络在时序超图上的时间切片上提取特征后, 将其送入 RNN 中做最后的时序预测任务. 与现有的基于 GNN 的结构化时间序列模型相比, HGC-RNN 所涉及的参数更少, 对于复杂网络具有更好的鲁棒性. 但该方法只适用于无权时序超图.

Yadati 等人<sup>[29]</sup>认为: HGNN 模型在信息融合的过程中引入了过多的噪声, 这在一些基于超图的半监督学习场景下将会有不良影响. 因此, Yadati 等人<sup>[29]</sup>借鉴前文提到的 BTR<sup>[17]</sup>和 Mediator<sup>[40]</sup>两种采样式展开思路, 提出了 HyperGCN、FastHyperGCN、1-HyperGCN 模型. 其展开示意图如图 6 所示. 该方法通过采样的方式对超图展开后的二元边进行了筛选, 从而一定程度上过滤了可能存在的数据噪声. 但同样, 该方法可能会忽略掉一些有效信息, 且采样后的网络结构虽然降低了每一次训练的复杂度, 但可能会需要更多的迭代次数才能收敛到理想结果.

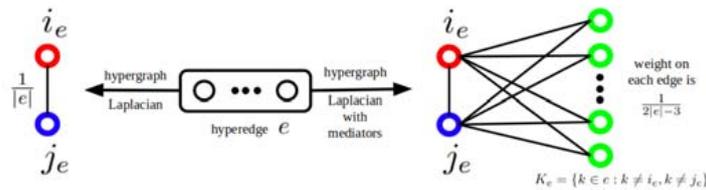


图 6 超图拉普拉斯算子和带中介的超图拉普拉斯算子<sup>[29]</sup>

Yang 等人<sup>[51]</sup>认为: Feng 等人<sup>[48]</sup>和 Yadati 等人<sup>[29]</sup>的工作只是构建了简单的加权图, 然后应用成熟的图学习算法, 这一过程对于高阶学习来说是远远不够的. 因此, Yang 等人<sup>[51]</sup>提出了一种新的超图展开方法 line expansion (LE). 超图  $G_H(V, E)$  经过 LE 展开后获得  $G_l(V_l, E_l)$ ,  $G_l$  中的节点由超图中的节点-超边对  $(v, e)$  组成. 图 7 为 LE 展开示意图.

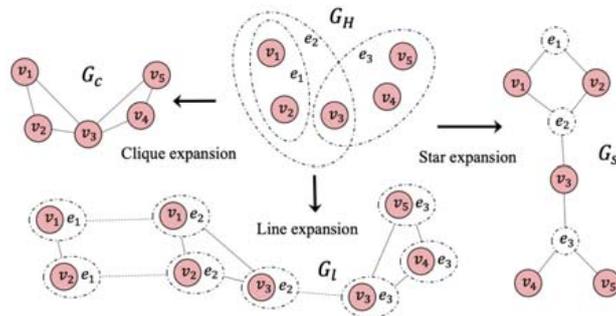


图 7 线式展开<sup>[51]</sup>

在图  $G_l$  中, Yang 等人首先将节点相似邻居和超边相似邻居的表征通过卷积的方式聚合起来, 其公式如下:

$$h_{(v,e)}^{(k+1)} = \sigma \left( \sum_{e'} w_e h_{(v,e')}^{(k)} \Theta^{(k)} + \sum_{v'} w_v h_{(v',e)}^{(k)} \Theta^{(k)} \right) \quad (13)$$

其中,  $h_{(v,e)}^{(k)}$  和  $h_{(v',e)}^{(k)}$  为  $G_l$  中节点  $(v,e')$  和  $(v',e)$  在卷积网络  $k$  层的表征,  $\Theta^{(k)}$  为  $k$  层滤波器参数,  $w_e$  和  $w_v$  是用于参数化节点相似度和边相似度的超参数. LE 展开方式保留了原始超图数据的共现性, 作者通过这种展开图特性对在相同边上的邻居节点和具有相似邻居节点的节点信息进行了提取, 很大程度上保留了原始超图的结构信息.

### 3.2.2 非展开式神经网络方法

#### (1) 基于自编码器的方法

Tu 等人<sup>[11]</sup>认为: 在一些超图的超边中, 同一超边节点集合的子集不能够构成超边关系, 这类超边被称为不可分解超边. 在这类超边上, 一些展开式方法将不再适用, 因而提出了专门针对不可展开超边的 DHNE 模型. Tu 等人定义了超图的一阶相似, 即同一超边内节点是邻近相似的; 与二阶相似, 即具有相似邻居结构的节点之间的相似性. 为了保留超图的二阶相似性, 他们使用 autoencoder 结构来对包含了邻居结构信息的拉普拉斯矩阵  $A=HH^T-D_v$  进行重建. 将重建获得的节点特征向量作为输入进行超图的一阶相似计算, 其计算公式如下所示:

$$\left. \begin{aligned} L_{ijk} &= \sigma(W_a^{(2)} * X_i^a + W_b^{(2)} * X_j^b + W_c^{(2)} * X_k^c + b^{(2)}), \\ S_{ijk} &\equiv S(X_i^a, X_j^b, X_k^c) = \sigma(W^{(3)} * L_{ijk} + b^3) \end{aligned} \right\} \quad (14)$$

$X_i^a$ 、 $X_j^b$ 、 $X_k^c$  即为重建获得的特征向量, 其中,  $\sigma$  是 sigmoid 函数. DHNE 模型在保留二阶相似性时使用的矩阵  $A$ , 其实就是前文提到的 Bolla's Laplacian<sup>[45]</sup>. DHNE 虽然是针对异构超图设计的模型且保留了节点的一阶、二阶信息, 但是由于其模型结构对输入元组的数目有严格要求, 因此只适用均匀超图, 很难扩展到任意超图.

#### (2) 基于自注意力机制的方法

针对 DHNE 方法仅限于处理固定类型和固定大小的异构超边这一问题, Zhang 等人<sup>[16]</sup>提出了 Hyper-SAGNN 模型. 他们在模型中使用了经典自注意力机制<sup>[52,53]</sup>来对超图信息进行聚合, 构造了节点与节点之间的成对注意力系数作为节点的动态特征, 同时, 结合节点原始的静态特征来对节点进行描述. 其最终计算公式如下:

$$o_i = W_o^T ((\bar{d}_i - \bar{s}_i)^2) + b \quad (15)$$

其中,  $\bar{d}_i$  为动态特征,  $\bar{s}_i$  为静态特征. 由于 Hyper-SAGNN 对输入元组的节点类型、数目没有要求, 因此, Hyper-SAGNN 与 DHNE 相比具有更好的泛化性. 但是由于 Hyper-SAGNN 构造了很多中间特征, 使其模型计算复杂度较高.

#### (3) 基于卷积的方法

类似地, Bai 等人<sup>[54]</sup>也把注意力机制运用到了超图上. 不同的是, Bai 等人将图注意力机制与图卷积两种结构进行了结合. 他们定义的卷积计算与前文提到的 Feng 等人<sup>[48]</sup>定义的卷积计算相同, 但 Bai 等人认为: 在定义了关联矩阵  $H$  之后, 图卷积内在的注意力机制是不可学习和训练的. 将注意力机制引入图卷积恰好可以解决这一问题, 因此, Bai 等人利用注意力机制重新计算了关联矩阵  $H$ , 并将其作为卷积神经网络的输入. 图 8 所示为模型注意力机制示意图. 在节点与超边属于同一同构域的假设前提下, 对于节点  $x_i$  与其关联超边  $x_j$  之间的关联关系  $H_{ij}$  定义如下:

$$H_{ij} = \frac{\exp(\sigma(\text{sim}(x_i P, x_j P)))}{\sum_{k \in \mathcal{N}_i} \exp(\sigma(\text{sim}(x_i P, x_k P)))} \quad (16)$$

其中,  $\sigma(\cdot)$  是一个非线性激活函数,  $\text{sim}(\cdot)$  是计算两两节点相似度的相似度函数,  $\mathcal{N}_i$  是节点  $i$  的邻接超边集合. 经过上述计算获得的关联矩阵, 受注意力机制的影响, 在神经网络反向传播的训练过程中变得可学习, 因此, 相

较于传统的超图卷积神经网络, 该模型更加灵活. 但需要注意的是, 上述注意力机制的建立过程需要保证节点与超边的相似度之间存在可比性.

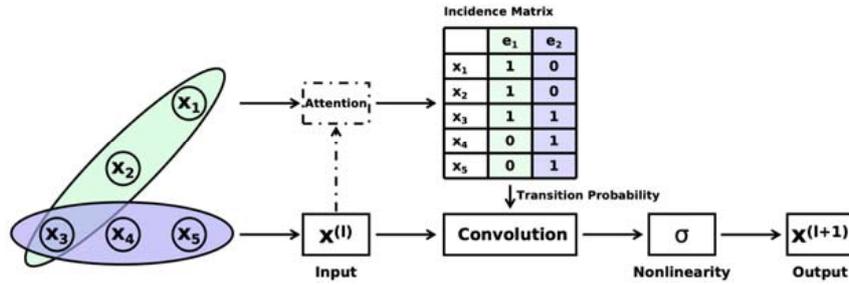


图 8 注意力机制<sup>[54]</sup>

非展开式超图神经网络还被引入到了流形学习中, Jin 等人<sup>[55]</sup>认为: 超图结构能够诱导卷积流形网络, 显著提高模型对复杂数据的表示能力. 因而提出了 hypergraph induced convolutional manifold network (H-CMN) 模型, 该模型的目标函数定义如下:

$$J_{\gamma, \rho}(\theta, Q) = \frac{1}{N} \sum_{i=1}^N L(R_i, f_Q(x_i)) + \frac{\gamma}{2} \|\theta\|^2 + \frac{\rho}{N^2} \sum_{i=1}^N \sum_{j=1}^N A_{ij} \|f_Q(x_i) - f_Q(x_j)\|^2 \quad (17)$$

其中, 前项是 softmax 损失函数,  $L(R_i, f_Q(x_i))$  表示  $x_i$  被预测为  $R_i$  类的概率; 后项则为流形损失函数,  $f_Q(x_i)$  是滤波器为  $Q$  的卷积神经网络对样本  $i$  的表征,  $A_{ij}$  是  $x_i$  和  $x_j$  两个样本在超图上的相似度. 虽然该目标函数与大多数流形损失函数<sup>[56,57]</sup>类似, 目标是使邻近的两个节点表征的空间距离接近, 但在该目标函数之上, Jin 等人提出了一种基于小批量数据的超图数据重构训练方法. 每次将前一次训练的训练集节点特征按照标签放入特征缓存, 在当前训练集中选取样本, 根据其标签, 使用  $k$  近邻方法与在缓存中选取的节点构建超边. 通过这种不断迭代的方式实现了 H-CMN 的小批量数据训练. 图 9 为其超图构建过程示例. 这一训练过程使得 H-CMN 能够实现大规模数据集的训练, 且使模型对于数据中的噪声有更好的鲁棒性.

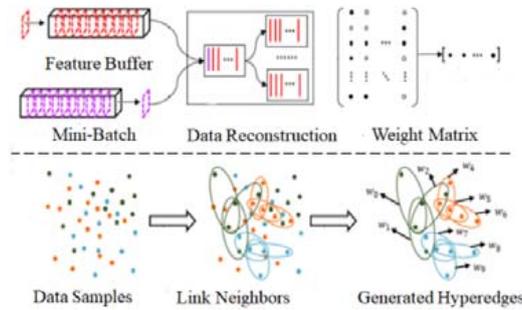


图 9 超图构建过程示例<sup>[55]</sup>

与上述工作不同的是, Jiang 等人<sup>[10]</sup>着眼于将卷积神经网络用于动态超图的挖掘研究, 提出了 dynamic hypergraph construction (DHG) 动态超图构建方法和 hypergraph convolution (HGC) 超图卷积模型. 他们从  $n$  个数据的表征  $X=[x_1; x_2; \dots; x_n]$  出发, 通过  $k$ -means 聚类和  $k$  近邻算法对每个样本构建超图. 因为超图关系是通过数据表征构建的, 所以在模型优化过程中, DHG 方法构建的超图会随着表征的不断优化而发生动态变化. 实际上, 上文提到的基于小批量数据的数据重构方法<sup>[55]</sup>也是一种动态超图构建方法. Jiang 等人结合 DHG 和 HGC 提出的 DHGNN 模型能够通过不断地重构超图结构来应对原始数据中隐藏的节点关系, 但由于该方法每次迭代都需要对节点进行聚类重构超图, 因此该方法的效率较低无法应用于大规模超图网络.

神经网络相关算法的对比总结见表 2.

表 2 超图神经网络方法总结

算法名称	主要结构	展开方式	输入	超图类型	是否展开	优势	劣势
HGNN	基于谱的图卷积网络	Star expansion	节点特征、邻接矩阵	同构有权	是	能够对超图的结构属性信息和节点特征信息进行更好的融合	该卷积计算的拉普拉斯算子在超图结构存在孤立点时会失效
DHCF	基于谱的图卷积网络	Star expansion	节点特征、邻接矩阵	异构无权	是	双通道超图协同过滤在对不同对象关系进行学习的同时,保证了不同对象自身的表征特性	模型仅通过共享权重的方式来构建不同对象之间的关联关系,使得不同对象之间的联系完全依赖于数据质量,这将难以应对噪声数据
HGC-RNN	基于谱的图卷积网络、循环神经网络	Star expansion	节点特征、邻接矩阵	同构无权	是	与现有基于 GNN 的结构化时间序列模型相比, HGC-RNN 所涉及的参数更少,对于复杂网络具有更好的鲁棒性	该方法只适用于无权时序超图
Hyper-GCN	基于谱的图卷积网络	Mediator	节点特征、邻接矩阵	同构有权	是	采样过程一定程度上过滤了可能存在的数据噪声	可能会忽略掉一些有效信息,且需要更多的迭代次数才能够收敛
LE	基于谱的图卷积网络	LE	节点特征、邻接矩阵	同构有权	是	展开方式保留了原始超图数据的共现性	展开后的结构为非常见结构,需要特殊处理
DHNE	自编码器、注意力机制	-	超边元组	异构无权	否	适用于异构超图且能够保留节点的一阶、二阶信息	只适用均匀超图,很难扩展到任意规模超图上
Hyper-SAGNN	注意力机制	-	超边元组	异构无权	否	模型输入的元组节点可为任意类型,与 DHNE 相比具有更好的泛化性	模型计算复杂度较高
AM	注意力机制、基于谱的卷积网络	-	节点特征、邻接矩阵	同构有权	否	使图卷积神经网络获得了可学习的注意力机制	需要保证节点与超边的相似度之间存在可比性的前提下
H-CMN	基于谱的卷积网络	-	节点特征、邻接矩阵	同构无权	否	能够实现大规模数据集的训练,且模型对于数据中的噪声有更好的鲁棒性	重建过程存在随机性,可能丢失信息
DHGNN	基于谱的卷积网络	-	节点特征、邻接矩阵	同构无权	否	能够不断地重构超图结构来适应原始数据中隐藏的节点关系	训练效率较低,无法应用于大规模超图网络

### 3.3 其他方法

除了上述两类超图学习方法外,在超图学习场景中,还有一些既不像谱分析方法那样具有严格的解析解,同时也不具备神经网络结构的机器学习方法.这类方法通常在特定的应用场景下,结合场景特点定义模型目标,并通过相应的优化方法求解.

HEBE<sup>[15]</sup>就是其中具有代表性的工作之一.作为早期的异构超图学习算法,HEBE 开创性地将超边内除核心节点外其余节点当作上下文信息,通过条件概率对超边内节点的共现概率  $P(u|C)$  进行计算,然后将计算出的概率分布  $P(u|C)$  和经验分布  $\hat{P}(\cdot|C_t)$  通过 KL 散度进行拟合,最终获得核心节点的表征.其目标函数如下:

$$L = -\sum_{t=1}^T \sum_{C_t \in P_t} \lambda_{C_t} KL(\hat{P}(\cdot|C_t), P(\cdot|C_t)) \quad (18)$$

其中,  $\lambda_{C_t}$  是衡量上下文重要性的一个参数.由于 HEBE 模型在对核心节点进行表征时结合了与核心节点相关的一些上下文信息,因此,HEBE 模型对于稀疏数据具有一定的鲁棒性.此外,HEBE 模型的核心概率计算过程具有良好的扩展性,因此,HEBE 模型可以应用于大规模真实超图网络数据.这一工作对于后续 Zhang 等人<sup>[16]</sup>和 Tu 等人<sup>[11]</sup>的工作具有很大的启发意义.

Du 等人<sup>[58]</sup>在视觉跟踪领域提出了基于几何超图学习的跟踪方法. Du 等人认为: 基于传统成对图的跟踪器仅考虑了局部零件之间的成对几何关系, 它们没有充分利用目标的内在结构, 在大变形和大遮挡的情况下, 容易受到成对关系误差的干扰. 因此, 他们提出一种基于置信度的采样方法, 来选择具有代表性的顶点和超边来构造几何超图, 将视觉跟踪问题归结为高阶几何关系模式的搜索问题. 该方法能够有效应对大尺度噪声问题.

类似地, 近来, Wen 等人<sup>[59]</sup>在多目标跟踪(MOT)场景下, 希望通过非均匀超图  $G(V, \epsilon, A)$  来编码同一物体在不同帧之间的位置关系. 他们以物体在每一帧中的位置作为节点, 通过在这些节点的邻域中搜索相应的稠密结构来构建非均匀超图. 对于节点  $v_s$ , 稠密结构搜索问题可以被表达为如下组合优化问题:

$$\begin{aligned} & \arg \max_y \sum_{d=1}^D \lambda_d \sum_{v_{1:d} \in N(v_s)} A(v_{1:d}) \overbrace{y_1 \cdots y_d}^d \\ \text{s.t. } & \sum_{i=1}^{|N(v_s)|} y_i = 1, y_s = \frac{1}{\alpha}, \forall i, y_i \in \left\{0, \frac{1}{\alpha}\right\} \end{aligned} \quad (19)$$

其中,  $N(v_s)$  是  $v_s$  的邻居节点.  $y$  是指示向量, 非零位表示对应节点被包括在稠密结构搜索中. 最终构建的非均匀超图能够有效提取各类对象之间的不同依赖关系.

Su 等人<sup>[60]</sup>提出, 在 3D 目标分类任务中引入有权超图. 在由待分类 3D 目标相关性建立的超图上, 对超图的权重进行组合优化, 使得不同的超边具有不同的权重, 且该权重能够代表超边的影响力. 在这个有权超图上, 重新对对象的相关性进行评估, 最终获得节点之间的关联关系. 该方法通过超图结构, 有效提取了 3D 目标多视图之间的复杂关系, 在实验评估指标上超过了经典深度神经网络方法 MVCNN.

Lin 等人<sup>[61]</sup>针对超图在处理视觉几何模型时节点数量过多、结构关系过于复杂的问题, 提出了 HOMF 算法. 其核心过程分为自适应线性估计(AIE)和迭代超边优化(IHO)两部分: 在 AIE 中, Lin 等人基于 Epanechnikov 核函数, 结合几何节点残差对节点权重进行估算, 并基于每个节点的权重对节点的信息熵进行计算, 从而选取具有代表性的重要数据点; 在 IHO 中, 则根据 AIE 产生的新节点权重对保留的节点重新构建超图关系, 达到优化超图结构的目的. 经过上述的采样和结构优化过程, 使得 HOMF 算法能够在保证采样效率的同时, 有效处理具有严重异常值的几何数据.

Lee 等人<sup>[62]</sup>则针对超图上的模体展开了研究. 他们定义了超图上的 h-motif 结构, 并提出了 MoCHy 并行算法用于计算 h-motif 的结构. 在实验中, 通过 5 个域上 11 个真实世界超图中 h-motif 结构数量与随机超图的比较实验, 证明了 h-motif 结构在真实超图中的意义, 并通过对 h-motif 结构变化趋势的分析, 得出了真实世界协作网络随时间变化的规律.

对于上述方法的总结介绍, 再次说明了超图结构普遍存在于真实世界中, 对于超图学习方法的研究具有普适的现实意义.

## 4 应用

超图学习作为一种能够有效挖掘超图信息的手段, 被广泛应用于图像处理任务<sup>[7,8]</sup>、超大规模集成电路<sup>[2]</sup>、推荐系统<sup>[63]</sup>等领域. 大量实验结果表明, 超图结构在表示数据的高阶交互关系方面要优于传统图结构. 接下来, 本节将对超图学习方法的应用方向进行展开讨论.

### 4.1 超图分割

超图分割<sup>[12,64]</sup>是超图学习方法的一个重要研究方向, 也是早期几个应用方向之一. 最大流最小割问题一直都是图论研究的一个热门问题, 在超图中也不例外. 如何快速求得该问题的可行解, 是很多领域所面临的问题.

George 等人<sup>[2]</sup>提出一种基于多级范式的超图划分算法, 将多级超图分割应用于超大规模集成电路(VLSI circuits)的工作就曾引起了各界的广泛关注. 此外, 很多早期超图学习算法都是以超图分割为目标而提出的, 如 Hadley<sup>[37]</sup>、Ihler<sup>[38]</sup>和 Lawler<sup>[65]</sup>就是早期针对均匀超图切割提出的, 为切割超边赋予固定权重标量作为惩罚

项的切割算法. 而对于非均匀超图, 也有一些基于边权重函数而非固定标量的超图分割算法<sup>[66,67]</sup>. 随着超图理论不断发展以及切割目标的改变, 越来越多的超图切割算法被提出. 目前, 超图分割是超图学习方法的一个重要应用方向之一.

## 4.2 聚类

网络聚类是指将网络顶点划分为一组簇的任务, 使得同一簇中的顶点彼此紧密连接, 而不同簇之间连接较弱. 这样的集群结构广泛存在于生物信息学、计算机科学、物理学、社会学等众多网络系统中, 并具有重要意义.

实际上, 超图分割也能实现聚类的目的, 在很多超图聚类的研究工作中, 将超图分割算法也归类于聚类算法中<sup>[67,68]</sup>. 因此, 对于自定义边权重函数的研究也是超图聚类中的一个重要研究方法. 但是通过一些超图的表征方法<sup>[69,70]</sup>将超图的结构信息与节点属性信息投影到向量空间, 然后借由其空间距离对超图进行聚类, 也是超图聚类的一条有效技术路线. 最终的聚类结果在保留了图结构特性的同时, 还综合了节点属性信息.

## 4.3 节点分类

除了上面提到的超图分割、聚类分析外, 节点分类也是超图学习方法的一大应用研究方向. 节点分类任务是建立在相似节点具有相似标签的假设前提下的一类任务. 大多数节点分类算法都是将有标签节点的数据分析结论推广到其他无标签数据上的一个过程, 这在超图中也不例外.

标签传播算法是一类非常经典的半监督图节点分类算法. 超图上的标签传播算法最早是由 Zhou 等人<sup>[12]</sup>提出的, 方法的目的是最小化共享同一条超边的顶点标签之间的差异. 与标签传播不同的是, 近来一些将超图结构信息映射到向量空间后再进行节点分类的模型<sup>[29,47]</sup>被提出, 这类模型希望能够挖掘除邻接信息外更高阶的结构信息, 从而提高节点分类效果. 而这类方法的主要缺点在于仅使用了初始超图结构, 忽略了结构信息映射到向量空间后对超图结构本身的影响. 因此, 一些动态超图模型被提出<sup>[10,71]</sup>, 此类模型在训练过程中会对图结构进行优化, 从而获得更好的分类结果.

## 4.4 连接预测

连接预测<sup>[72]</sup>是一类在现实世界中被广泛使用的应用, 尤其是在社交网络和推荐系统中. 真实图数据中通常存在部分缺失关系, 如何在有限的节点关系(即边)中抽取出现图的数据模式, 是这类问题的一大难点. 因为超图的结构特性, 超图连接预测输入相对复杂. 连接预测在超图中的输入通常是一个节点集合, 算法模型需要判断在节点集合包含的节点之间是否存在超边.

超图上的连接预测模型主要分为两类: 一类是以 DHNE<sup>[11]</sup>为代表的针对均匀超图的超边预测模型, 另一类则是以 Hyper-SAGNN<sup>[16]</sup>为代表的以任意非均匀超图为目标超边预测模型. 此外, 很多方法在设计过程中为了满足场景特性还会引入一些除传统超图结构外的额外信息, HSimple、HypE 算法<sup>[73]</sup>即在连接预测时引入了超边内包含节点的相对位置信息.

## 4.5 重要性排序

随着数据量爆炸式的增长, 节点重要性排序一直以来都是国内外研究的焦点, PageRank 等算法也被各界广泛应用. 在大规模超图上如何定量分析节点的重要程度, 也是超图网络研究所需要解决的问题.

随机游走是一类高效的重要性排序算法, 超图上的随机游走可以追溯到 Zhou 等人<sup>[12]</sup>提出的工作. 在所有超边都具有相同数量的节点这一假设下, 一些针对均匀超图的随机游走算法被提出<sup>[74,75]</sup>. 近来, Carletti 等人<sup>[6]</sup>提出了针对任意超图的随机游走算法, 且在其工作中证明: 在一些情况下, 超图随机游走产生的节点重要性排序与传统随机游走产生的节点重要性排序相比存在倒置的情况.

## 4.6 推荐系统

电子商务在世界范围内规模不断扩大, 商品类目随之快速增长. 如何为顾客推荐想要的商品, 成为各大电商的难题, 为了解决这一问题, 推荐系统应运而生. 作为学术界与工业界联系最为紧密的应用方向之一, 研

究者们对于推荐系统的研究层出不穷。

面对推荐的复杂场景,很多研究工作提出使用超图对多元信息进行建模<sup>[63,76,77]</sup>,在保证信息完整性的同时,能够有效地对高阶关系建模,从而更好地实现信息融合。基于超图的推荐算法已被证明在音乐推荐<sup>[63,77]</sup>、文本推荐<sup>[76]</sup>、电影推荐<sup>[49]</sup>场景下有显著效果。

#### 4.7 视觉任务

超图学习算法在视觉任务上也有广泛的应用。在计算机视觉中,超图用于描述视觉特征之间的关系,例如视觉分类<sup>[78]</sup>、图像检索<sup>[45]</sup>和视频对象分割<sup>[7]</sup>。也有使用超图结构在 3D 模型分类工作中进行标签传播的<sup>[71]</sup>,以及在社交媒体上处理多模态数据的<sup>[79,80]</sup>。超图学习算法在处理视觉任务中相对多变,算法需要根据具体目标任务的特点而设计。在视觉任务中引入超图结构的主要原因在于:超图能够很好地表达主体之间的复杂关系,发掘主体时间的高阶关系。

#### 4.8 化学分析

分子结构的图表示在计算化学中有着广泛的应用。而传统图结构不能充分描述非经典结构的化合物,不准确的分子图使得分子结构分析无法顺利进行。为了解决这一问题,Konstantinova 等人<sup>[81]</sup>在其工作中提出用超图来作为非经典分子结构的数学模型,并将一系列超图理论应用于分子结构分析。

近来,Kajino<sup>[82]</sup>进一步提出在分子优化过程中引入超图。分子优化的目的是发现具有理想性质的新分子,而在生成新分子的过程中常常会出现解码错误。而 Kajino 提出的分子超图语法(MHG)则能有效避免这一问题。

#### 4.9 生物网络

在生物网络中,节点通常描述蛋白质、代谢物、基因或其他生物实体,而边缘代表节点之间的功能关系或相互作用,如“结合”“催化”或“转化为”。传统图的一个重要性质是每条边都连接两个节点,然而许多生物过程的特点是有两个以上的参与伙伴。用传统图常常无法恰当表达这些生物过程,而超图提供了一个框架能够有效解决这一问题。

近来,Tsuyuzaki<sup>[83]</sup>就在其工作中针对以往的数据分析方法只关注两种细胞类型之间的一对一的交互这一问题,提出了 scTensor 用于提取代表性交互关系超图的新方法,从而发现一些新的细胞交互模式。类似地,Yu 等人<sup>[33]</sup>也基于类内散布矩阵提出了 HCIS,有效分析了高阶生物模块的交互关系。Wu 等人<sup>[84]</sup>更是直接将生物信号通路建模为超图,从而更好地研究信号通路的相互作用。在 Wu 等人<sup>[84]</sup>工作的启发下,Schwob 等人<sup>[85]</sup>认为,考虑时间依赖性有助于进一步改善细胞信号通路的表现。他们定义了时间依赖超图,更好地模拟了生物信号通路。

## 5 实验与分析

本节将针对超图上的节点分类任务和连接预测任务分别展开对比实验。在节点分类任务中,我们将对谱分析方法中提到的 3 种展开式方法 star expansion<sup>[9]</sup>、clique expansion<sup>[9]</sup>、Rodriguez<sup>[34]</sup>和两种非展开式方法 Zhou<sup>[12]</sup>和 Bolla<sup>[43]</sup>进行对比实验。此外,实验中还加入了 HyperGCN<sup>[29]</sup>、HGNN<sup>[48]</sup>、Hyper-SAGNN<sup>[16]</sup>、DHGNN<sup>[10]</sup>这 4 种神经网络方法。在连接预测任务中,本文对 DHNE<sup>[11]</sup>和 Hyper-SAGNN<sup>[16]</sup>这两种方法展开了对比实验。

### 5.1 数据集介绍

为了对比上述方法,实验数据除了经典网络引文数据 Cora 外,实验使用  $k$  近邻算法构建了 4 个超图数据集,数据集具体信息见表 3。

- ORL: 人脸数据库 ORL 包含了 40 个不同的主题,每个主题有 10 幅不同的图像。这些图像是在不同时间拍摄的,因此在光线、表情、面部细节等方面存在不同。这里,我们使用  $k=3$  的  $k$  近邻算法对 ORL 数据进行超图构建,每张图片为一个节点,每条超边由每个节点在向量空间上欧氏距离最近的 3 个邻

居所构成. 因此, ORL 超图数据有 400 个节点和 400 条边, 40 类标签;

- COIL20<sup>[86]</sup>: 灰度图片数据集 COIL20 包含对 20 个物体从不同角度的拍摄, 每隔 5°拍摄一张图像, 因此每个物体有 72 张图像. 同样, 我们使用  $k=3$  的  $k$  近邻算法对 COIL20 数据进行超图构建, 每张图片为一个节点, 每条超边由每个节点在向量空间上欧氏距离最近的 3 个邻居节点所构成;
- Yeast<sup>[87]</sup>: 酵母菌分类数据集 Yeast 包含 1 484 个酵母菌实例. 这里, 我们去除酵母菌的命名特征, 以最后一维 nuc 特征为标签, 酵母菌实例为节点, 使用  $k=3$  的  $k$  近邻算法对 Yeast 数据进行超图构建;
- Cora<sup>[88]</sup>: 引文数据集 Cora 包含 2 708 个科学出版物, 每个出版物属于 7 个类别中的一个. 这里使用的是 HyperGCN 一文中提供与处理过的超图数据集. 但由于该数据集中存在孤立点, 为了避免拉普拉斯矩阵变成奇异矩阵, 本文所使用的 Cora 数据集去除了这些孤立点;
- MNIST<sup>[86]</sup>: 手写数字数据集 MNIST 包含 60 000 个训练样本和 10 000 个预测样本, 这里我们仅使用 10 000 个预测样本来构建我们的超图数据集.

表 3 用于评估节点分类的数据集

数据集	节点数量	超边数量	平均超边尺寸	特征维度	类别数
ORL	400	400	3	1 024	40
COIL20	1 440	1 440	3	1 024	20
Yeast	1 484	1 484	3	7	10
Cora	2 388	1 072	4.3	1 433	7
MINIST	10 000	10 000	3	784	10

对于连接预测实验, 我们使用的数据集信息如下(见表 4).

- GPS<sup>[89]</sup>: 数据集描述用户在某个位置加入活动. (用户、位置、活动)关系用于构建超网络;
- MovieLens<sup>[90]</sup>: 该数据集描述了对于电影的个人标记活动. 每部电影至少有一种类型. (用户、电影、标签)关系构成了超图的超边;
- Drug<sup>[91]</sup>: 该数据集来自 FDA 不良事件报告系统(FAERS), 它包含了提交给 FDA 的不良事件和药物错误报告的信息. 通过(用户、药物、反应)关系来构建超图.

表 4 用于评估连接预测的数据集

数据集	节点数量	边数量	平均超边尺寸
GPS	221	1 436	3
MovieLens	17 100	47 957	3
drug	7 486	171 756	3

## 5.2 实验设置

在节点分类任务对比实验中, 我们将 ORL、COIL20、Yeast、Cora、MINIST 这 5 个数据集进行了随机划分, 50%的数据用于训练, 另外 50%的数据用于测试, 以对比不同模型的节点分类准确率, 随机数种子设置为 0. 在传统谱分析方法中, Zhou、Star expansion、Clique expansion 这 3 种有权算法使用了 LLRE<sup>[39]</sup>、CENTROID<sup>[39]</sup>、TRACE<sup>[39]</sup>、VOLUME<sup>[39]</sup>这 4 种超边权重计算方法. 此外, 本文在节点分类对比中还分别使用了 10%、20%、30%、40%、50%这 5 个不同的训练数据比例进行训练, 其余数据用于测试, 以对比不同算法的鲁棒性. 最后, 本文还将 Star expansion、Clique expansion、Rodriguez、Zhou、Bolla、HyperGCN、HGNN、Hyper-SAGNN、DHGNN 这 9 种节点分类算法在 Cora 数据集上获得的节点表征, 经过 tsne 降维后进行了可视化.

在连接预测实验中, 本文将 DHNE、Hyper-SAGNN 这两种算法在 GPS、MovieLens、Drug 数据集上进行了对比实验, 分别使用了 10%、20%、30%、40%、50%的超边信息进行训练, 其余用于测试.

## 5.3 结果与分析

本节对实验中的算法性能进行分析和对比. 以下图表中的结果均为错误率.

由表 5 可以看出, Zhou、Star expansion、Clique expansion 这 3 种谱方法在不同的超边权重计算方式下的

结果有明显差异. 其中, LLRE 和 CENTROID 这两种超边权重计算方式在 3 种算法、4 个数据集中表现相对稳定; 5 种谱方法在 5 个数据集中的性能表现有所差异. 在谱方法中: Zhou 在 COIL20 和 Cora 数据集上的错误率是在 5 种方法中最低的; Clique expansion 方法在 COIL20 和 ORL 数据集上, 在 5 种方法中表现最佳. 但总体上来说, 5 种方法在 5 个数据集上的性能表现相对平稳, 错误率比较接近.

表 5 谱方法节点分类结果 (%)

Database	Zhou		Star-expansion		Clique-expansion		Bolla	Rodriguez
COIL20 ( $k=3$ )	LLRE	5.42±1.81	LLRE	5.62±2.01	LLRE	5.42±1.81	5.42±1.81	5.49±1.87
	CENTROID	<b>5.35±1.87</b>	CENTROID	5.62±2.01	CENTROID	<b>5.35±1.87</b>		
	TRACE	5.69±1.81	TRACE	5.83±1.94	TRACE	5.69±1.81		
	VOLUME	5.62±1.6	VOLUME	5.56±1.81	VOLUME	5.62±1.6		
ORL ( $k=3$ )	LLRE	10.75±1.75	LLRE	13.75±1.75	LLRE	<b>10.25±2.25</b>	10.75±1.75	11.25±0.75
	CENTROID	10.75±2.25	CENTROID	11.0±2.0	CENTROID	10.5±2.5		
	TRACE	14.25±1.75	TRACE	14.5±1.5	TRACE	14.25±1.75		
	VOLUME	13.25±1.25	VOLUME	14.75±1.75	VOLUME	13.5±1.5		
MINIST ( $k=3$ )	LLRE	15.34±0.16	LLRE	15.75±0.01	LLRE	14.94±0.16	14.82±0.04	<b>14.7±0</b>
	CENTROID	15.26±0.18	CENTROID	15.65±0.5	CENTROID	14.94±0.22		
	TRACE	18.44±0.18	TRACE	18.66±0.12	TRACE	18.49±0.19		
	VOLUME	15.78±0.04	VOLUME	16.4±0.06	VOLUME	15.49±0.01		
Yeast ( $k=3$ )	LLRE	49.66±0.2	LLRE	<b>46.36±1.35</b>	LLRE	49.12±0.2	49.26±0.47	46.9±2.02
	CENTROID	49.73±0.27	CENTROID	46.43±1.28	CENTROID	49.12±0.2		
	TRACE	49.66±0.2	TRACE	<b>46.36±1.35</b>	TRACE	49.12±0.2		
	VOLUME	49.8±0.2	VOLUME	47.24±2.09	VOLUME	49.73±0.4		
Cora	LLRE	<b>36.26±1.26</b>	LLRE	40.45±0.75	LLRE	36.73±1.21	36.35±1.34	37.86±1.59
	CENTROID	36.31±1.3	CENTROID	40.49±0.8	CENTROID	36.64±1.21		
	TRACE	37.9±1.55	TRACE	41.83±1.13	TRACE	38.19±1.68		
	VOLUME	38.23±1.13	VOLUME	39.24±1.38	VOLUME	38.07±1.21		

神经网络的节点分类结果见表 6, 其中, HGNN 模型在 COIL20、ORL、MINIST、Cora 这 4 个数据集中错误率最低, 在 Yeast 数据集中稍逊于 hyperGCN. DHGNN 模型在 5 个数据集中的错误率略高于 HGNN, 但整体表现较为平稳. hyperGCN 模型在 ORL 数据集中错误率要显著高于 HGNN 和 DHGNN, 这可能是因为 HyperGCN 模型的建模思路不适合小规模数据集. 而 Hyper-SAGNN 模型的节点分类实验, 我们依照论文中提到的将模型倒数第 2 层的输出取出作为节点的表征, 再用该表征结合逻辑回归模型进行分类实验. 从最终结果中我们可以看到: Hyper-SAGNN 模型在所有数据集中均表现不佳, 甚至低于一些谱方法. 由此可见, Hyper-SAGNN 这种以节点间是否存在超边为目标所获得的节点嵌入对于节点分类任务是不适用的. 从整体上来看, 在节点分类任务中, 神经网络方法要优于传统谱方法.

表 6 图神经网络方法节点分类结果 (%)

Database	HyperGCN	HGNN	Hyper-SAGNN	DHGNN
COIL20( $k=3$ )	1.04±0.07	0.28±0	51.18±0.76	0.56±0
ORL( $k=3$ )	13±2.5	6.75±0.25	27.75±0.75	8.5±0.5
MINIST( $k=3$ )	5.37±0.17	4.63±0.07	40.6±0.71	5.54±0.02
Yeast( $k=3$ )	43.87±0.47	44.21±0.27	59.43±1.89	51.82±0.07
Cora	17.47±0.8	16.88±0.12	45.06±0.75	19.35±0.17

上述 9 种方法的节点表征经过 tsne 算法降维后的可视化结果如图 10 所示, 子图括号中为节点表征对应节点分类结果的错误率. 从图中我们可以看到: 传统谱方法的节点表征中大部分在空间上有所分离, 但仍有一些节点存在较为严重的混杂. 而 HyperGCN、DHGNN、HGNN 这 3 种模型产出的节点表征在降维后的二维空间中有一个较为清晰的分布. Hyper-SAGNN 模型则是可视化结果最差的, 大部分节点都混杂在一起. 这也在一定程度上说明了 Hyper-SAGNN 在节点分类任务中表现不佳的原因, 其同类节点表征在嵌入空间上没有明显的区分度.

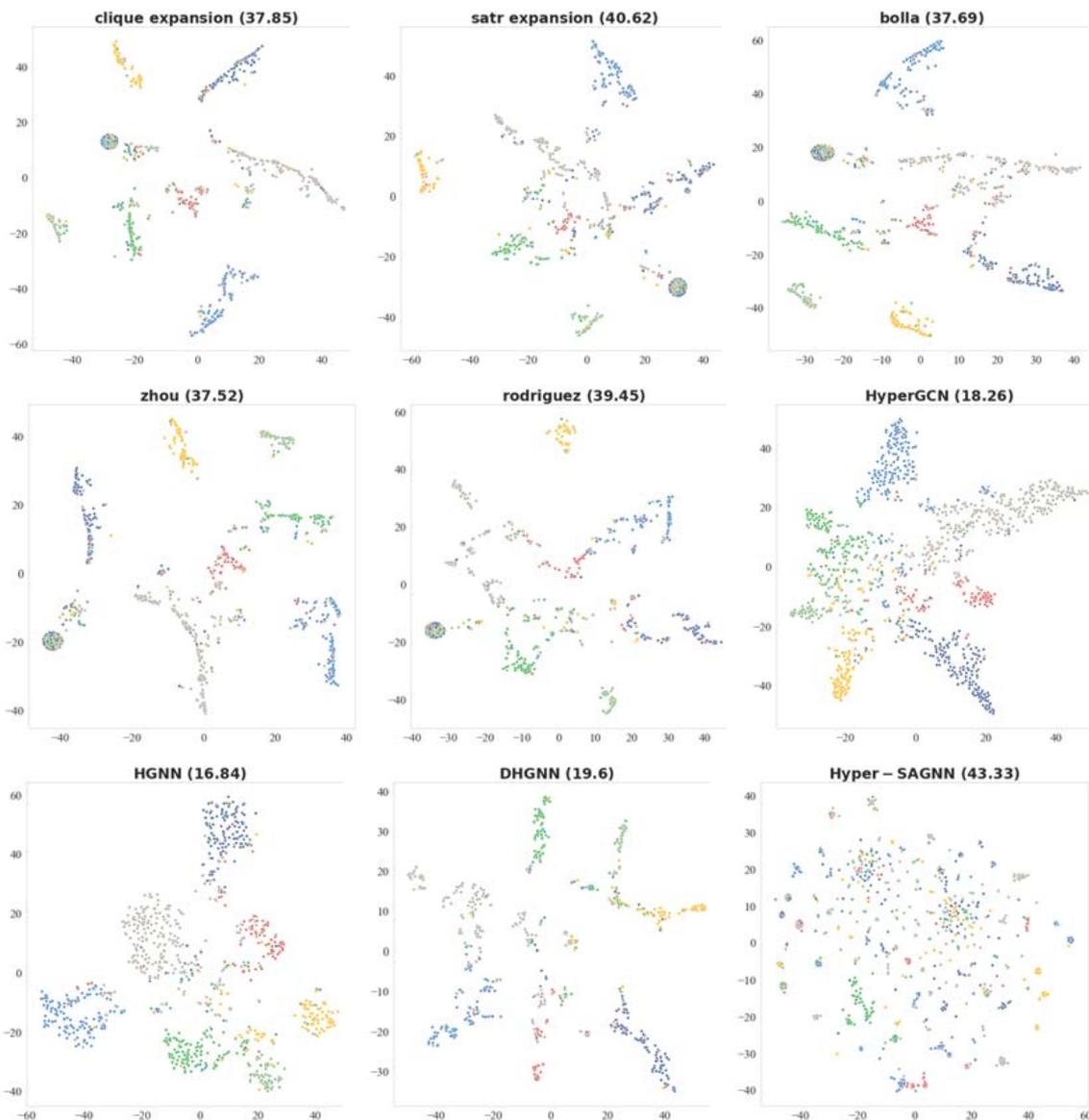


图 10 节点分类任务的节点嵌入结果

此外, 模型鲁棒性实验的结果如图 11 所示(由于 Hyper-SAGNN 模型的节点分类性能表现不佳, 我们将其从鲁棒性实验中剔除)。从图中我们可以看到: 在 ORL 两个数据集中, 传统谱方法和神经网络方法的结果较为接近。但从总体的实验结果中可以看出: 随着训练数据百分比的变化, 神经网络方法的错误率变化情况要小于传统谱方法, 神经网络方法的模型稳定性优于传统谱方法。

连接预测实验, 我们同样选取了 10%、20%、30%、40%、50% 的超边信息进行训练, 其余用于测试。在图 12 中: 随着用于训练的超边数据比例不同, DHNE 和 Hyper-SAGNN 两类模型的错误率各有高低。但从错误率变化曲线上来看, Hyper-SAGNN 模型的稳定性要高于 DHNE 模型; 且在 50% 训练数据的情况下, Hyper-SAGNN 模型准确率高于 DHNE 模型。但 Hyper-SAGNN 模型的训练速度要远远慢于 DHNE 模型。

表 7 是在 drug 数据集上 50% 数据的训练结果, 我们可以看到: 虽然 Hyper-SAGNN 错误率低于 DHNE, 但在训练时间上, Hyper-SAGNN 是 DHNE 的 10 倍。

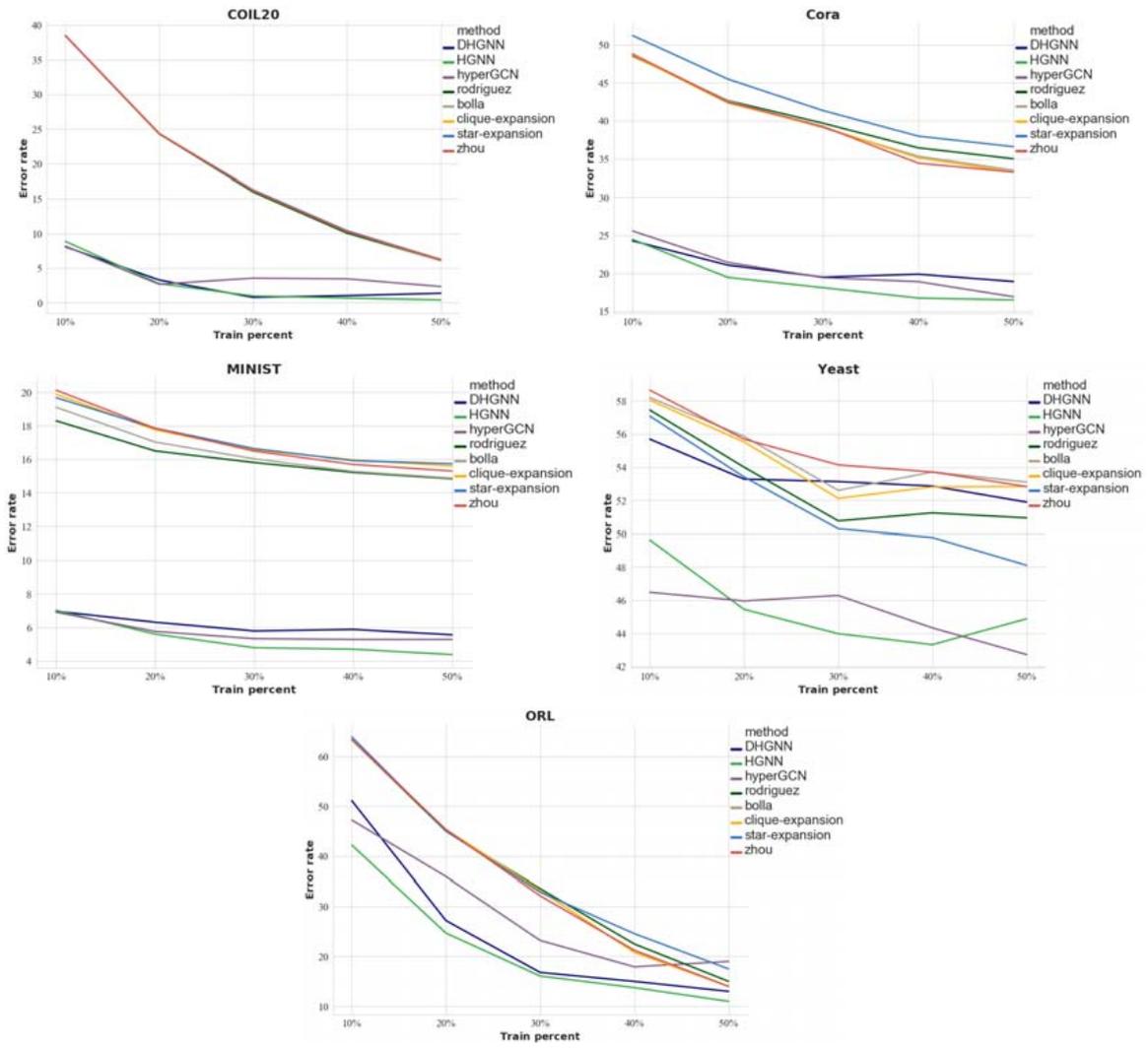


图 11 不同百分比数据训练模型的结果

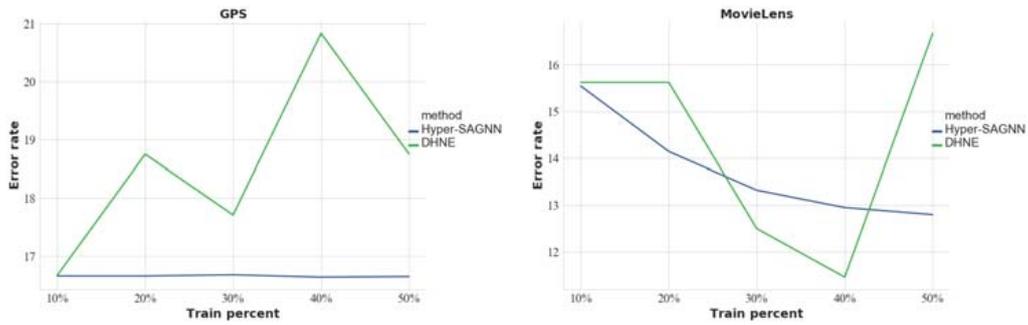


图 12 链接预测结果

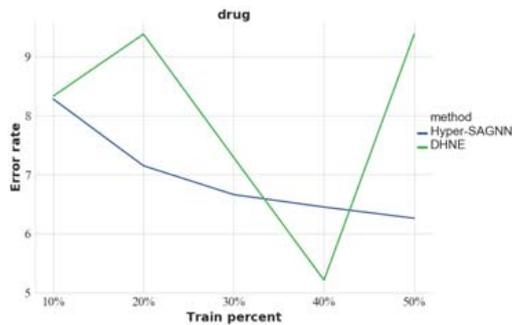


图 12 链接预测结果(续)

表 7 durg 连接预测结果

Method	Time (ms)	Erro rate (%)
Hyper-SAGNN	195	6.36
DHNE	18	9.38

## 6 未来展望

随着图结构研究的兴起,作为一种可以对高阶关系进行建模的灵活数据结构,超图结构近几年受到的关注也逐渐增多.由于图卷积神经网络、图注意力机制、自编码器神经网络结构的发展和成熟,研究者对于超图的研究手段变得更加灵活、多变.虽然超图学习算法应用于解决聚类、视觉等任务由来已久,但大规模真实网络数据场景下超图学习算法的研究近来才被研究者所重视.在大规模真实网络数据场景下,超图学习算法面临着异构节点、非均匀超边以及模型鲁棒性等诸多问题的挑战.针对目前超图学习算法的不足,本节提出超图学习算法研究的几个潜在的研究方向.

- (1) 有很大一部分超图学习算法只能处理均匀超图或同构超图,这类算法显然无法在大规模真实网络数据场景下使用.因为这类场景下通常有很多异构节点以及复杂的节点关系,受超图本身复杂结构的影响(一条边包含多个节点),上述问题将会变得更加棘手.一些算法虽然可以处理异构超图,但因模型复杂度过高因而不具备应用价值.如何在保证模型性能的前提下处理超图中的异构节点和各类超边,是超图学习算法的一个重要研究方向;
- (2) 高阶关系广泛存在于真实场景中,但这种高阶关系往往容易被忽略.很多真实场景下的数据仅有节点信息和节点的成对关系而没有超边信息.超图结构想要得到进一步的应用推广,首先要解决如何在这类数据上根据目标任务建立超边关系,从而利用超图学习算法来处理这类任务获得更好的性能表现.因而,如何根据节点信息和成对关系获得超边关系,也是一个超图学习的重要研究方向;
- (3) 超图结构数据在真实场景中由于数据观测的局限性,一些隐含的重要关系并没有直接体现在固有结构中,且一些局部超图结构还有可能会误导模型.因而有人提出动态超图模型,即:在模型优化的过程中对超图结构进行动态修改,从而使得超图结构可以更加贴合目标任务,以获得更好的性能表现.动态超图能够在一定程度上解决超图上的不一致性问题,是一个极具潜力的研究方向;
- (4) 像很多传统图应用场景一样,超图应用场景也存在标签稀疏的问题.因而,通过部分有标签节点来预测其余节点标签的超图半监督学习,也是超图学习的一大热点问题.与传统图不同的是,超图上的每条超边涉及两个以上的节点,如何处理同一超边上的节点之间的关系就变得尤为重要.一些超图学习算法通过将超图展开成普通的方式来解决这一问题.这一类方法虽然直观、灵活,但超边展开过程容易丢失超边信息;
- (5) 超图结构信息除了节点信息、节点邻居信息外,还包括一些高阶全局信息.一些方法通过注意力机制、图卷积等结构来抽取超图的节点信息和节点邻居信息,但鲜有方法可将高阶全局信息引入超图模型.如何有效抽取超图中的高阶结构信息,将其引入到超图学习模型中,是一个仍待解决的问题;
- (6) 超图分割一直以来都是超图的研究热点之一,尽管已有很多算法能够实现对超图结构在不同的目标条件下进行切割,但研究者对于降低算法的时间和空间复杂度这一问题的研究一直未停止.所以,如何快速获得符合目标条件的超图切割结果,也是一个重要研究方向;

- (7) 除此之外, 在超图上还存在一些特殊应用场景的研究, 如利用超图进行多元关系推理、解决图像处理问题等. 研究这类问题时需要将超图学习方法和研究场景特点进行结合, 对超图学习方法进行一些特异性设计, 从而使超图机制能够贴合任务场景.

## 7 总 结

本文对超图学习方法进行了梳理分析: 首先, 将超图学习方法分为谱分析方法、神经网络方法以及其他方法, 再进一步根据方法建模思路的不同将其分为展开式方法和非展开式方法; 在讨论具体方法时, 根据方法的特点进行了进一步细分, 并对同类方法进行了分析比较; 通过实验对比了主流超图学习方法的算法性能; 最后, 对当前超图学习的研究方向进行了展望.

### References:

- [1] <https://en.wikipedia.org/wiki/Hypergraph#Definitions>
- [2] Karypis G, Aggarwal R, Kumar V, *et al.* Multilevel hypergraph partitioning: Applications in VLSI domain. *IEEE Trans. on Very Large Scale Integration (VLSI) Systems*, 1999, 7(1): 69–79.
- [3] Karypis G, Kumar V. Multilevel  $k$ -way hypergraph partitioning. *VLSI Design*, 2000, 11(3): 285–300.
- [4] Eiter T, Gottlob G. Identifying the minimal transversals of a hypergraph and related problems. *SIAM Journal on Computing*, 1995, 24(6): 1278–1304.
- [5] Sun L, Ji S, Ye J. Hypergraph spectral learning for multi-label classification. In: *Proc. of the 14th ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining*. Las Vegas: Association for Computing Machinery, 2008. 668–676. [doi: 10.1145/1401890.1401971]
- [6] Carletti T, Battiston F, Cencetti G, *et al.* Random walks on hypergraphs. *Physical Review E*, 2020, 101(2-1): 022308.
- [7] Huang Y, Liu Q, Metaxas D. Video object segmentation by hypergraph cut. In: *Proc. of the 2009 IEEE Conf. on Computer Vision and Pattern Recognition*. IEEE, 2009. 1738–1745.
- [8] Gao Y, Wang M, Tao D, *et al.* 3-D object retrieval and recognition with hypergraph analysis. *IEEE Trans. on Image Processing*, 2012, 21(9): 4290–4303.
- [9] Agarwal S, Branson K, Belongie S. Higher order learning with graphs. In: *Proc. of the 23rd Int'l Conf. on Machine Learning*. Pittsburgh: Association for Computing Machinery, 2006. 17–24. [doi: 10.1145/1143844.1143847]
- [10] Jiang J, Wei Y, Feng Y, *et al.* Dynamic hypergraph neural networks. In: *Proc. of the IJCAI*. 2019. 2635–2641.
- [11] Tu K, Cui P, Wang X, *et al.* Structural deep embedding for hyper-networks. In: *Proc. of the AAAI Conf. on Artificial Intelligence*. 2018.
- [12] Zhou D, Huang J, Schölkopf B. Learning with hypergraphs: Clustering, classification, and embedding. In: *Advances in Neural Information Processing Systems*, 2006, 19: 1601–1608.
- [13] Liu Y, Shao J, Xiao J, *et al.* Hypergraph spectral hashing for image retrieval with heterogeneous social contexts. *Neurocomputing*, 2013, 119: 49–58.
- [14] Wu F, Han YH, Zhuang YT. Multiple hypergraph clustering of Web images by mining word2image correlations. *Journal of Computer Science and Technology*, 2010, 25(4): 750–760.
- [15] Gui H, Liu J, Tao F, *et al.* Large-scale embedding learning in heterogeneous event data. In: *Proc. of the 16th IEEE Int'l Conf. on Data Mining (ICDM)*. IEEE, 2016. 907–912.
- [16] Zhang R, Zou Y, Ma J. Hyper-SAGNN: A self-attention based graph neural network for hypergraphs. In: *Proc. of the Int'l Conf. on Learning Representations (ICLR)*. 2020.
- [17] Louis A. Hypergraph Markov operators, eigenvalues and approximation algorithms. In: *Proc. of the 47th Annual ACM Symp. on Theory of Computing*. 2015. 713–722.
- [18] Bretto A. *Hypergraph Theory — An Introduction*. Cham: Springer, 2013.
- [19] Berge C. *Hypergraphs: Combinatorics of Finite Sets*. Elsevier, 1984.
- [20] Kahng AB. Fast hypergraph partition. In: *Proc. of the 26th ACM/IEEE Design Automation Conf.* 1989. 762–766.
- [21] Alon N. Transversal numbers of uniform hypergraphs. *Graphs and Combinatorics*, 1990, 6(1): 1–4.

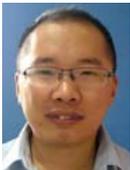
- [22] Gunopulos D, Mannila H, Khardon R, *et al.* Data mining, hypergraph transversals, and machine learning. In: Proc. of the 16th ACM SIGACT-SIGMOD-SIGART Symp. on Principles of Database Systems. 1997. 209–216.
- [23] Chung F. The Laplacian of a hypergraph. In: Expanding Graphs (DIMACS Series). 1993. 21–36. [doi: 10.1090/dimacs/010/03]
- [24] Cooper J, Dutle A. Spectra of uniform hypergraphs. *Linear Algebra and Its Applications*, 2012, 436(9): 3268–3292.
- [25] Buló SR, Pelillo M. A game-theoretic approach to hypergraph clustering. In: Proc. of the NIPS. Citeseer, 2009. 1571–1579.
- [26] Papa DA, Markov IL. Hypergraph partitioning and clustering. In: *Handbook of Approximation Algorithms and Metaheuristics*. 2007. 61–71.
- [27] Zass R, Shashua A. Probabilistic graph and hypergraph matching. In: Proc. of the 2008 IEEE Conf. on Computer Vision and Pattern Recognition. IEEE, 2008. 1–8.
- [28] Bretto A, Gillibert L. Hypergraph-Based image representation. In: Proc. of the Int'l Workshop on Graph-based Representations in Pattern Recognition. Springer-Verlag, 2005. 1–11.
- [29] Yadati N, Nimishakavi M, Yadav P, *et al.* HyperGCN: A new method for training graph convolutional networks on hypergraphs. In: *Advances in Neural Information Processing Systems*, 2019, 32: 1511–1522.
- [30] Kipf TN, Welling M. Semi-supervised classification with graph convolutional networks. arXiv Preprint arXiv: 160902907, 2016.
- [31] Defferrard M, Bresson X, Vandergheynst P. Convolutional neural networks on graphs with fast localized spectral filtering. In: Proc. of the 30th Int'l Conf. on Neural Information Processing Systems. 2016. 3844–3852.
- [32] Zien JY, Schlag MD, Chan PK. Multilevel spectral hypergraph partitioning with arbitrary vertex sizes. *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, 1999, 18(9): 1389–1399.
- [33] Yu L, Shen X, Jiang X, *et al.* Hypergraph clustering based on intra-class scatter matrix for mining higher-order microbial module. In: Proc. of the 2019 IEEE Int'l Conf. on Bioinformatics and Biomedicine (BIBM). IEEE, 2019. 240–243.
- [34] Rodríguez JA. On the Laplacian eigenvalues and metric parameters of hypergraphs. *Linear and Multilinear Algebra*, 2002, 50(1): 1–14.
- [35] Rodríguez JA. On the Laplacian spectrum and walk-regular hypergraphs. *Linear and Multilinear Algebra*, 2003, 51(3): 285–297.
- [36] Agarwal S, Lim J, Zelnik-Manor L, *et al.* Beyond pairwise clustering. In: Proc. of the 2005 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR 2005). IEEE, 2005. 838–845.
- [37] Hadley SW. Approximation techniques for hypergraph partitioning problems. *Discrete Applied Mathematics*, 1995, 59(2): 115–127.
- [38] Ihler E, Wagner D, Wagner F. Modeling hypergraphs by graphs with the same mincut properties. *Information Processing Letters*, 1993, 45(4): 171–175.
- [39] Huang S, Elgammal A, Yang D. On the effect of hyperedge weights on hypergraph learning. *Image and Vision Computing*, 2017, 57: 89–101.
- [40] Chan THH, Liang Z. Generalizing the hypergraph Laplacian via a diffusion process with mediators. *Theoretical Computer Science*, 2020, 806: 416–428.
- [41] Alon N, Milman VD.  $\lambda_1$ , isoperimetric inequalities for graphs, and superconcentrators. *Journal of Combinatorial Theory, Series B*, 1985, 38(1): 73–88.
- [42] Alon N. Eigenvalues and expanders. *Combinatorica*, 1986, 6(2): 83–96.
- [43] Bolla M. Spectra, euclidean representations and clusterings of hypergraphs. *Discrete Mathematics*, 1993, 117(1-3): 19–39.
- [44] Purkait P, Chin TJ, Sadri A, *et al.* Clustering with hypergraphs: The case for large hyperedges. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2016, 39(9): 1697–1711.
- [45] Huang Y, Liu Q, Zhang S, *et al.* Image retrieval via probabilistic hypergraph ranking. In: Proc. of the 2010 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition. IEEE, 2010. 3376–3383.
- [46] Ma X, Liu W, Li S, *et al.* Hypergraph  $\mathcal{L}$ -Laplacian regularization for remotely sensed image recognition. *IEEE Trans. on Geoscience and Remote Sensing*, 2018, 57(3): 1585–1595.
- [47] Huang J, Chen C, Ye F, *et al.* Hyper2vec: Biased random walk for hyper-network embedding. In: Proc. of the Int'l Conf. on Database Systems for Advanced Applications. Springer, 2019. 273–277.
- [48] Feng Y, You H, Zhang Z, *et al.* Hypergraph neural networks. In: Proc. of the AAAI Conf. on Artificial Intelligence. 2019. 3558–3565.
- [49] Ji S, Feng Y, Ji R, *et al.* Dual channel hypergraph collaborative filtering. In: Proc. of the 26th ACM SIGKDD Int'l Conf. on Knowledge Discovery & Data Mining. 2020. 2020–2029.

- [50] Yi J, Park J. Hypergraph convolutional recurrent neural network. In: Proc. of the 26th ACM SIGKDD Int'l Conf. on Knowledge Discovery & Data Mining. 2020. 3366–3376.
- [51] Yang C, Wang R, Yao S, *et al.* Hypergraph learning with line expansion. arXiv: 200504843, 2020.
- [52] Vaswani A, Shazeer N, Parmar N, *et al.* Attention is all you need. In: Proc. of the 31st Int'l Conf. on Neural Information Processing Systems. 2017. 6000–6010.
- [53] Veličković P, Cucurull G, Casanova A, *et al.* Graph attention networks. arXiv: 171010903, 2017.
- [54] Bai S, Zhang F, Torr PH. Hypergraph convolution and hypergraph attention. 2020. <https://arxiv.org/pdf/1901.08150.pdf>
- [55] Jin T, Cao L, Zhang B, *et al.* Hypergraph induced convolutional manifold networks. In: Proc. of the IJCAI. 2019. 2670–2676.
- [56] Belkin M, Niyogi P, Sindhvani V. Manifold regularization: A geometric framework for learning from labeled and unlabeled examples. *Journal of Machine Learning Research*, 2006, 7(11).
- [57] He X, Niyogi P. Locality preserving projections. *Advances in Neural Information Processing Systems*, 2004, 16(16): 153–160.
- [58] Du D, Qi H, Wen L, *et al.* Geometric hypergraph learning for visual tracking. *IEEE Trans. on Cybernetics*, 2017, 47(12): 4182–4195.
- [59] Wen L, Du D, Li S, *et al.* Learning non-uniform hypergraph for multi-object tracking. In: Proc. of the AAAI Conf. on Artificial Intelligence. 2019. 8981–8988.
- [60] Su L, Gao Y, Zhao X, *et al.* Vertex-weighted hypergraph learning for multi-view object classification. In: Proc. of the 26th Int'l Joint Conf. on Artificial Intelligence. 2017. 2779–2785.
- [61] Lin S, Xiao G, Yan Y, *et al.* Hypergraph optimization for multi-structural geometric model fitting. In: Proc. of the AAAI Conf. on Artificial Intelligence. 2019. 8730–8737.
- [62] Lee G, Ko J, Shin K. Hypergraph motifs: Concepts, algorithms, and discoveries. *Proc. of the VLDB Endowment*, 2020, 13(12): 2256–2269.
- [63] Tan S, Bu J, Chen C, *et al.* Using rich social media information for music recommendation via hypergraph model. *ACM Trans. on Multimedia Computing, Communications, and Applications (TOMM)*, 2011, 7(1): 1–22.
- [64] Hein M, Setzer S, Jost L, *et al.* The total variation on hypergraphs-learning on hypergraphs revisited. In: Proc. of the 26th Int'l Conf. on Neural Information Processing Systems, Vol. 2. 2013. 2427–2435.
- [65] Lawler EL. Cutssets and partitions of hypergraphs. *Networks*, 1973, 3(3): 275–285.
- [66] Li P, Milenkovic O. Inhomogeneous hypergraph clustering with applications. *Advances in Neural Information Processing Systems*, 2017, 2017: 2309–2319.
- [67] Veldt N, Benson AR, Kleinberg J. Minimizing localized ratio cut objectives in hypergraphs. In: Proc. of the 26th ACM SIGKDD Int'l Conf. on Knowledge Discovery & Data Mining. 2020. 1708–1718.
- [68] Whang JJ, Du R, Jung S, *et al.* MEGA: Multi-view semi-supervised clustering of hypergraphs. *Proc. of the VLDB Endowment*, 2020, 13(5): 698–711.
- [69] Yang D, Qu B, Yang J, *et al.* Revisiting user mobility and social relationships in LBSNs: A hypergraph embedding approach. In: Proc. of the World Wide Web Conf. 2019. 2147–2157.
- [70] Martino A, Giuliani A, Rizzi A. (Hyper) Graph embedding and classification via simplicial complexes. *Algorithms*, 2019, 12(11): 223.
- [71] Zhang Z, Lin H, Gao Y. Dynamic hypergraph structure learning. In: Proc. of the 27th Int'l Joint Conf. on Artificial Intelligence. 2018. 3162–3169.
- [72] Li D, Xu Z, Li S, *et al.* Link prediction in social networks based on hypergraph. In: Proc. of the 22nd Int'l Conf. on World Wide Web. 2013. 41–42.
- [73] Fatemi B, Taslakian P, Vazquez D, *et al.* Knowledge hypergraphs: Prediction beyond binary relations. arXiv: 190600137, 2019.
- [74] Helali A, Löwe M. Hitting times, commute times, and cover times for random walks on random hypergraphs. *Statistics & Probability Letters*, 2019, 154: 108535.
- [75] Lu L, Peng X. High-ordered random walks and generalized Laplacians on hypergraphs. In: Proc. of the Int'l Workshop on Algorithms and Models for the Web-graph. Springer, 2011. 14–25.
- [76] Zhu Y, Guan Z, Tan S, *et al.* Heterogeneous hypergraph embedding for document recommendation. *Neurocomputing*, 2016, 216: 150–162.
- [77] Bu J, Tan S, Chen C, *et al.* Music recommendation by unified hypergraph: Combining social media information and music content. In: Proc. of the 18th ACM Int'l Conf. on Multimedia. 2010. 391–400.

- [78] Wang M, Liu X, Wu X. Visual classification by  $\ell_1$ -hypergraph modeling. *IEEE Trans. on Knowledge and Data Engineering*, 2015, 27(9): 2564–2574.
- [79] Chen F, Gao Y, Cao D, *et al.* Multimodal hypergraph learning for microblog sentiment prediction. In: *Proc. of the 2015 IEEE Int'l Conf. on Multimedia and Expo (ICME)*. IEEE, 2015. 1–6.
- [80] Ji R, Chen F, Cao L, *et al.* Cross-modality microblog sentiment prediction via bi-layer multimodal hypergraph learning. *IEEE Trans. on Multimedia*, 2018, 21(4): 1062–1075.
- [81] Konstantinova EV, Skorobogatov VA. Application of hypergraph theory in chemistry. *Discrete Mathematics*, 2001, 235(1–3): 365–383.
- [82] Kajino H. Molecular hypergraph grammar with its application to molecular optimization. In: *Proc. of the Int'l Conf. on Machine Learning*. PMLR, 2019. 3183–3191.
- [83] Tsuyuzaki K, Ishii M, Nikaido I. Uncovering hypergraphs of cell-cell interaction from single cell RNA-sequencing data. *bioRxiv*, 2019. 566182.
- [84] Wu MY, Hill CS. TGF- $\beta$  superfamily signaling in embryonic development and homeostasis. *Developmental Cell*, 2009, 16(3): 329–343.
- [85] Schwob M, Zhan J, Dempsey A. Modeling cell communication with time-dependent signaling hypergraphs. *IEEE/ACM Trans. on Computational Biology and Bioinformatics*, 2019.
- [86] <http://www.cad.zju.edu.cn/home/dengcai/Data/MLData.html>
- [87] <https://archive.ics.uci.edu/ml/datasets/Yeast>
- [88] <https://github.com/malllabisc/HyperGCN/tree/master/data/coauthorship/cora>
- [89] Zheng V, Cao B, Zheng Y, *et al.* Collaborative filtering meets mobile recommendation: A user-centered approach. In: *Proc. of the AAAI Conf. on Artificial Intelligence*. 2010.
- [90] Harper FM, Konstan JA. The movielens datasets: History and context. *ACM Trans. on Interactive Intelligent Systems (TIIS)*, 2015, 5(4): 1–19.
- [91] <http://www.fda.gov/Drugs/>



胡秉德(1994—), 男, 博士生, 主要研究领域为数据挖掘, 图神经网络.



王新根(1982—), 男, 博士, 讲师, CCF 专业会员, 主要研究领域为分布式计算, 软件性能.



王新宇(1979—), 男, 博士, 教授, 博士生导师, CCF 专业会员, 主要研究领域为数据实时智能, 人工智能, 软件工程.



宋明黎(1976—), 男, 博士, 教授, 博士生导师, CCF 专业会员, 主要研究领域为计算机视觉, 模式识别与人工智能, 机器学习.



陈纯(1955—), 男, 博士, 教授, 博士生导师, 主要研究领域为计算机软件与理论, 大数据实时智能处理技术.