

北京理工大学

研究报告

离散数学在图持续学习领域的应用

The Application of Discrete Mathematics in the Field of Graph
Continual Learning

学 院: 计算机学院

专业名称: 人工智能

班 级: 07182401

姓名学号: 王浩恩 1120242340

庞项祯 1120241349

杨佳车 1120240660

任课教师: 张春霞

2026 年 1 月 16 日

离散数学在图持续学习领域的应用

0 摘 要

本文……。

摘要正文选用模板中的样式所定义的“正文”，每段落首行缩进 2 个字符；或者手动设置成每段落首行缩进 2 个汉字，字体：宋体，字号：小四，行距：固定值 22 磅，间距：段前、段后均为 0 行。阅后删除此段。

摘要是一篇具有独立性和完整性的短文，应概括而扼要地反映出本论文的主要内容。包括研究目的、研究方法、研究结果和结论等，特别要突出研究结果和结论。中文摘要力求语言精炼准确，本科生毕业设计（论文）摘要建议 300-500 字。摘要中不可出现参考文献、图、表、化学结构式、非公知公用的符号和术语。英文摘要与中文摘要的内容应一致。阅后删除此段。

关键词：离散数学；图持续学习；图距离编辑；拓扑正则化

The Application of Discrete Mathematics in the Field of Graph Continual Learning

0 Abstract

In order to study.....

Abstract 正文设置成每段落首行缩进 2 字符，字体：Times New Roman，字号：小四，行距：固定值 22 磅，间距：段前、段后均为 0 行。阅后删除此段。

Key Words: BIT; Undergraduate; Graduation Project (Thesis)

目 录

摘要	I
Abstract	II
第1章 绪论	1
1.1 研究背景与意义	1
1.2 现有挑战	1
第2章 离散数学基础	2
2.1 图的基本定义	2
2.2 子图与同构	2
2.2.1 子图	2
2.2.2 图同构: 结构的等价性	3
2.2.3 子图同构与图匹配问题	3
2.3 图编辑距离 (Graph Edit Distance, GED)	3
2.3.1 图编辑距离的形式化定义	4
2.3.2 计算挑战与算法概览	4
第3章 图持续学习基础	6
3.1 持续学习的核心概念	6
3.2 图数据的持续学习场景	6
第4章 基于 GED 的图动态建模	8
4.1 动态演化的量化模型	8
4.2 动态建模策略	9
4.2.1 策略一: 基于 GED 的经验回放	9
4.2.2 策略二: 拓扑正则化	10
第5章 算法实现与复杂性分析	12
5.1 近似计算优化	12
5.2 算法流程描述	12
第6章 总结与展望	13
6.1 结论	13
6.2 未来工作	13
参考文献	14

北京理工大学本科生论文

附录	15
附录 A	\LaTeX 环境的安装	15
附录 B	BIThesis 使用说明	15
致谢	16

1 絮论

1.1 研究背景与意义

正文……yuFeiJiZongTiDuoXueKeSheJiYouHuaDeXianZhuangYuFaZhanFangXiang2008
.....Hajela2012Application

正文部分：宋体、小四；正文行距：22 磅；间距段前段后均为 0 行。阅后删除此段。

图、表居中，图注标在图下方，表头标在表上方，宋体、五号、居中，1.25 倍行距，间距段前段后均为 0 行，图表与上下文之间各空一行。阅后删除此段。

图-示例：(阅后删除此段)

1.2 现有挑战

2 离散数学基础

2.1 图的基本定义

在离散数学中，图是描述事物间二元关系的一种抽象数学模型，一个图 G 可以定义为一个有序二元组 $G = (V, E)$ ，其中 V 是一个有限的、非空的集合，称为顶点集， E 是顶点集 V 中无序对的集合（对于无向图）或有序对的集合（对于有向图），称为边集。其核心在于忽略事物的具体属性，仅关注对象之间是否存在关联。

相比之下，图持续学习关注的是模型如何从一个按时间顺序到达的图序列中持续地、高效地学习并积累知识，同时避免对已学旧知识的灾难性遗忘。在此语境下，基础的分析单元不再是单个静态图，而是一个动态演化的图序列。

2.2 子图与同构

在对动态图序列的分析中，我们不仅需要描述单个图的结构，更需要精确地刻画不同图之间的结构关系。子图关系描述了图之间的包含与嵌入，而同构关系则定义了图之间结构上的完全等价性。这两个概念是理解图编辑距离与图演化模式的理论基础。

2.2.1 子图

子图关系反映了图结构之间的部分与整体，是分析局部模式、识别功能模块以及量化结构变化的核心工具。

在离散数学的定义中，令 $G = (V, E)$ 和 $G' = (V', E')$ 为两个图，若 $V' \subseteq V$ 且 $E' \subseteq E$ ，并且 E' 中的每条边所关联的顶点均属于 V' ，称 G' 是 G 的子图。而子图在图持续学习中有着重要意义，其核心动态图序列的演化可以自然地看成一个子图扩展过程，即 G_t 可视为由子图 $G_{(t-1)}$ 演化而来，新增部分则代表了在时间间隔内发生的变化。这种视角将全局演化问题转化为局部结构增量的识别问题。

2.2.2 图同构：结构的等价性

当我们需要判断两个图在结构上是否“完全相同”时，就需要用到离散数学中图同构的概念。同构是图论中最深刻的等价关系，也是许多图算法（如图匹配、图数据库查询）的核心。两个图 $G = (V, E)$ 和 $H = (U, F)$ 是同构的，即 $G \cong H$ ，当且仅当存在一个双射函数 $\phi: V \rightarrow U$ ，使得对 V 中任意两个顶点 u, v ，

$$[u, v] \in E \iff [\varphi(u), \varphi(v)] \in F. \quad (2-1)$$

这样的函数 φ 称为从 G 到 H 的一个同构映射。对于有向图何标记图，则条件要分别加强到有向边相同和标签相同。

同构关系表明，两个图虽然在表面表示上可能不同（顶点编号、空间布局），但其连接拓扑结构完全一致。子图同构判定是 NP 完全问题，它在多个领域有根本性应用：在化学信息学中，用于查找分子数据库中包含特定功能基团的分子；在社交网络分析中，用于检测网络中的特定社群结构（如小团体、星型结构）；在计算机视觉中，用于从场景图中识别已知的物体关系模式。

2.2.3 子图同构与图匹配问题

一个更普遍且在实践中至关重要的难题是子图同构问题，它询问一个图是否包含另一个图的同构副本。

在图持续学习的动态建模背景下，子图同构的概念为我们提供了一种定性地描述结构稳定性的工具。然而，当 H 在演化过程中发生微小畸变（如一名成员暂时断开连接）时，严格的子图同构关系将不再成立，但结构相似性依然很高。此时，我们需要一种定量的、能够容忍微小差异的度量，来刻画这种近似包含或近似等价的关系——这正是接下来的图编辑距离所要解决的问题。

2.3 图编辑距离 (Graph Edit Distance, GED)

在动态图序列的分析中，我们不仅需要识别结构的定性与定量变化，更需要一种能够精确量化这种变化、并能揭示变化细节的数学工具。图编辑距离 (Graph Edit Distance, GED) 正是这样一种工具，它将图之间的差异转化为一系列可解释的原子

操作，并通过最小化操作总代价来定义距离。

2.3.1 图编辑距离的形式化定义

对于两个标记图 $G_1 = (V_1, E_1, l(v_1), l(E_1))$ 和 $G_2 = (V_2, E_2, l(V_2), l(E_2))$ ，基本的图编辑操作包括：

1. 顶点替换：将图 G_1 中的一个顶点 $u \in V_1$ 映射为图 G_2 中的一个顶点 $U \in V_2$ ，记为 $(u \rightarrow U)$ 。若 $l_V(u) \neq l_V(U)$ ，则此操作实质为修改顶点标签。
2. 顶点插入：将图 G_2 中的一个顶点 $U \in V_2$ 映射到一个特殊的空顶点 ε ，记为 $(\varepsilon \rightarrow U)$ ，表示在 G_1 中插入该顶点。
3. 顶点删除：将图 G_1 中的一个顶点 $u \in V_1$ 映射到空顶点 ε ，记为 $(u \rightarrow \varepsilon)$ ，表示从 G_1 中删除该顶点。
4. 边替换、边插入、边删除：定义与顶点操作类似，作用于边集。

一个编辑路径 $P = (o_1, o_2, \dots, o_k)$ 是一个有限的编辑操作序列，这些操作将图 G_1 逐步转换为一个与 G_2 同构的图。

基于此，我们可以给定一个代价函数 c ，为每一个编辑操作 o 赋予一个非负代价 $c(o) \geq 0$ 。图 G_1 和 G_2 间的图编辑距离定义为所有可能编辑路径中的最小总代价：

$$\text{GED}(G_1, G_2) = \min_{P \in \mathcal{P}(G_1, G_2)} \sum_{o \in P} c(o), \quad (2-2)$$

其中 $\mathcal{P}(G_1, G_2)$ 是所有能将 G_1 转换为 G_2 的编辑路径的集合。

2.3.2 计算挑战与算法概览

图编辑距离的计算是 NP-hard 问题，且对于一般的图，其判定问题（给定两个图和阈值 K ，判断 GED 是否小于等于 K ）是 NP-完全的。

这一复杂性源于编辑路径数量的组合爆炸。给定两个具有 n 和 m 个顶点的图，可能的顶点映射（即顶点替换的集合）数量高达 $O(n^m)$ 量级。因此，在实际应用中，尤其是大规模图或动态图序列中需要频繁计算时，必须依赖近似算法或启发式方法。以下是三类主要方法：

1. **基于搜索的精确算法**: 例如 A* 算法, 它利用启发式函数剪枝搜索空间, 在中小规模图上可求得精确解, 但最坏情况下仍是指数复杂度。
2. **整数线性规划 (ILP) 模型**: 将 GED 计算建模为一个整数线性规划问题, 利用优化求解器 (如 Gurobi、CPLEX) 求解。该方法能获得精确解, 但同样受限于问题规模。
3. **近似与启发式算法**: 这是处理大规模图的主要途径。
 - **二分图匹配法**: 将顶点映射问题转化为二分图最小权匹配问题, 通过忽略部分边结构获得多项式时间的近似解。著名的 Hungarian 算法和 Volgenant-Jonker 算法常被用于此。
 - **局部搜索与迭代改进**: 如 Beam Search (束搜索), 在搜索过程中只保留有限数量的最有希望的路径。
 - **基于深度学习的 GED 预测**: 近年来, 利用图神经网络 (GNN) 直接学习图的表示并预测 GED 值的方法展现出潜力, 能够实现快速但近似的距离估计。

3 图持续学习基础

我们在此处通过系统性地阐述图持续学习的基本范式，以明晰其核心概念、独特挑战以及与其它学习范式的区别，并分类阐述其典型应用场景，从而为后续“基于图编辑距离的动态建模”提供明确的问题定义和需求背景。

3.1 持续学习的核心概念

持续学习旨在模拟智能体在动态环境中持续积累并整合知识的能力，其核心是处理数据分布的非平稳性。

作为一种机器学习范式，持续学习要求模型从一系列按时间或顺序到达的任务或数据分布 D_1, D_2, \dots, D_T 中依次进行学习。其核心目标是，在学习新任务 D_t 的同时，能够最大限度地保持对先前所有已学任务 $\{D_1, \dots, D_{t-1}\}$ 的性能。

灾难性遗忘是持续学习中最根本的挑战。当神经网络等模型使用基于梯度的优化方法学习新任务时，其参数会为了最小化新任务的损失而更新，这一过程往往会上覆盖或破坏存储在过去参数中关于旧任务的知识，导致模型在旧任务上的性能急剧下降，仿佛“遗忘”了之前所学。

为了缓解灾难性遗忘，图持续学习主要围绕三大策略展开研究，分别是：基于回放的方法，基于正则化的方法，基于动态架构的方法。而当持续学习的对象从独立的图像或文本样本，转变为内部存在复杂结构关联的图数据时，上述挑战和策略被置于一个全新的、更复杂的维度中。

3.2 图数据的持续学习场景

图持续学习并非简单地将持续学习技术应用于图神经网络。图数据自身的结构性、关联性和演化特性，使得其持续学习问题呈现出独特的场景和更高的复杂度。

图数据因为其自身的结构关联，演化特性，在持续学习问题上有更独特更复杂的场景，无论是在节点粒度预测时的分类、回归任务，还是在动态演化的图数据中进行适应、学习，都对算法的计算效率提出了极高的要求。

北京理工大学本科生论文

为得到最优的算法，我们必须借助图编辑距离，以此量化一个图结构的变化程度，据此解析出具体的演化过程，帮助机器识别出所需关注的子图。

4 基于 GED 的图动态建模

基于离散数学和图编辑距离，我们提出一个分层框架：首先，利用 GED 构建动态演化的量化模型，将宏观序列差异映射为微观结构单元的演化属性；其次，基于此量化模型，设计两种新颖的、具有明确拓扑语义的持续学习策略——基于 GED 的经验回放与基于 GED 的拓扑正则化。该框架使得模型不仅能感知变化的大小，更能理解变化的性质，从而实现更精准的知识巩固与适应。

4.1 动态演化的量化模型

给定动态图序列，我们构造一个加权的有向演化图 $\mathcal{G} = (\mathcal{V}, \mathcal{A}, w)$ ，其中：

- 顶点集 \mathcal{V} ：每个顶点 V_t 对应原序列中的一个图快照 G_t 。
- 弧集 \mathcal{A} ：对于任意 $t \geq 2$ ，存在一条从 V_{t-1} 指向 V_t 的弧 $a_{t-1,t}$ 。
- 权重函数 w ：弧 $a_{t-1,t}$ 的权重 $w_{t-1,t} = \text{GED}(G_{t-1}, G_t)$ 。

演化图 \mathcal{G} 是动态序列的一阶宏观抽象。通过分析其权重序列 $\{w_{t-1,t}\}$ ，我们可以识别系统演化的模式：

- 平稳演化期：连续出现较小且波动不大的 GED 值。
- 剧变期：GED 值出现显著峰值，指示重大结构重组或外部事件冲击。
- 趋势分析：计算 GED 序列的移动平均与自相关系数，可判断演化是趋向稳定还是持续动荡。

为了将宏观演化与微观结构变化联系起来，我们引入顶点演化轨迹与结构单元活跃度的概念。

对于一个在时间窗 $[t_s, t_e]$ 内存在的顶点 u ，其演化轨迹 $\text{Traj}(u)$ 是一个序列，记录其在每个快照 G_t ($t_s \leq t \leq t_e$) 中的邻接向量（或高阶表征）及标签（如有）。顶点的活跃度 $\alpha(u)$ 定义为在其生命周期内，其邻接向量（或表征）的逐时间步变化量的范数之和，该变化可通过对比相邻快照中的局部子图的 GED 来近似计算。高 $\alpha(u)$ 表明该顶点连接关系不稳定，是网络中的活跃点或边缘点。

对于一条边 $e = (u, v)$ ，其持久性 $p(e)$ 定义为在连续快照中同时包含顶点 u 和 v 且边 e 存在的时间步比例。对于一个连通子图 H （如 k -clique、 k -core），其持久性

$p(H)$ 定义为在连续快照中能找到与 H 同构（或近似同构，GED 小于阈值 τ ）的子图的时间步比例。

高持久性的边或子图是网络的骨干结构或功能模块，如长期稳定的合作团队、频繁交互的蛋白复合物。它们是需要被模型长期牢固记忆的知识实体。

基于 GED 的量化流程：

1. 输入：动态图序列 $\mathbb{G} = \{G_1, G_2, \dots, G_T\}$ 。
2. 解析编辑路径：从 $P_{t-1,t}$ 中提取涉及的顶点和边操作，更新对应顶点和边的活跃度与持久性计数器。
3. 输出量化指标：每个顶点的活跃度 $\alpha(u)$ 、每条边的持久性 $p(e)$ 、每个显著子图模式 H 的持久性 $p(H)$ ，以及演化图。

此量化模型将 GED 的数值输出转化为一系列附着于图结构元素上的、具有明确拓扑语义的属性标签，为后续的持续学习策略提供了精准的方向标。

4.2 动态建模策略

基于上述量化模型，我们提出两种核心的持续学习策略。它们共享一个核心理念：依据结构单元在演化中的角色，实施差异化的学习处理。

4.2.1 策略一：基于 GED 的经验回放

传统经验回放通常随机或基于多样性选择节点/边。我们提出 GED-Guided Replay (GGR)，其核心是利用量化模型输出的持久性指标，智能选择回放样本，确保模型持续复习网络中最稳定、最核心的知识。GGR 的核心步骤为：

1. 识别候选记忆单元：在历史快照中，识别所有满足 $p(e) > \theta_p$ (高持久性边) 的边，以及由这些边诱导出的连通子图作为核心子图。同时，将活跃度极低 $\alpha(u) < \theta_\alpha$ 的顶点标记为稳定顶点。
2. 构建记忆缓冲区：记忆缓冲区 \mathcal{M} 存储两种类型的记忆项：

- 子图记忆项：对于每个核心子图 H ，存储其拓扑结构 A_H 、顶点特征 X_H 以及在上次被观测到时模型对其内部节点的预测输出（软标签）。
 - 节点上下文记忆项：对于每个稳定顶点 u ，存储其当前特征 c_u 、其 k -hop 邻域的子图结构、以及模型对其的预测输出。
3. 自适应回放调度：当在新快照 G_t 上训练时，设计回放概率。对于记忆项 m_i ，其被抽中的概率

$$P_i \propto p(\text{单元}i) \cdot \exp(-\beta \cdot \text{recall}_i),$$

其中 p 是其持久性， recall_i 是近期被回放的次数， β 是控制多样性的超参数。这使高持久性且近期较少被复习的记忆项有更高优先级。

4. 基于结构的回放增强：在回放一个子图记忆项时，并非孤立地输入该子图，而是将其“植入”到当前快照 G_t 的一个相似上下文中（例如，找到当前图中一个与记忆子图表征相似的节点簇，将记忆子图作为附加结构注入计算图），以模拟该核心结构在新环境下的作用，增强知识的可迁移性。

4.2.2 策略二：拓扑正则化

正则化方法通过约束参数更新来保护旧知识。我们提出 Topology-Aware Elastic Weight Consolidation (TA-EWC)，将 GED 分析得出的拓扑稳定性信息融入参数重要性估计中，实现正则化的“拓扑感知”。

1. 计算拓扑加权的 Fisher 信息矩阵：在标准的 EWC 中，参数 θ_i 的重要性正比于 Fisher 信息矩阵对角线元素 F_i ，它度量了该参数对预测结果的敏感度。我们引入拓扑权重进行修正。
 - 首先，在任务 A （对应历史图数据）上，对于每个训练样本（节点），计算其损失对参数的梯度 $\nabla_{\theta} \mathcal{L}_A(f_{\theta}(v))$ 。
 - 然后，计算拓扑加权 Fisher 信息：

$$\tilde{F}_i^{(A)} = \frac{1}{|\mathcal{V}_A|} \sum_{v \in \mathcal{V}_A} \underbrace{(\alpha(v) + \varepsilon)^{-1}}_{\text{稳定性权重}} [\nabla_{\theta} \mathcal{L}_A(f_{\theta}(v))]_i^2,$$

这里 $\alpha(v)$ 是顶点活跃度。关键洞察：对稳定顶点（低 $\alpha(v)$ ）预测正确的参数更为重要，因为这些顶点表征了网络的静态核心。因此，我们用活跃度的倒数进行加权，赋予稳定点对应的梯度平方更高权重。

2. 定义正则化损失：当学习新任务 B （对应新快照）时，总损失函数为

$$\mathcal{L}_{\text{total}}(\theta) = \mathcal{L}_B(\theta) + \frac{\lambda}{2} \sum_i \tilde{F}_i^{(A)} (\theta_i - \theta_i^*)^2,$$

其中 θ^* 是任务 A 训练结束后的最优参数， λ 控制正则化强度。

相比于其他的算法，TA-EWC 的主要优势是正则化更具有解释性，与图结构紧密耦合，而且可与回放策略自然融合。

5 算法实现与复杂性分析

5.1 近似计算优化

5.2 算法流程描述

6 总结与展望

6.1 结论

6.2 未来工作

如果您遇到生僻字无法显示、列表项间距过大、三线表需要跨页等问题，请参考[在线文档的「疑难杂症」部分](#)。

6 参考文献

附 录

附录相关内容…

A.1 L^AT_EX 环境的安装

L^AT_EX 环境的安装。

B.2 BIThesis 使用说明

BIThesis 使用说明。

附录是毕业设计（论文）主体的补充项目，为了体现整篇文章的完整性，写入正文又可能有损于论文的条理性、逻辑性和精炼性，这些材料可以写入附录段，但对于每一篇文章并不是必须的。附录依次用大写正体英文字母 A、B、C……编序号，如附录 A、附录 B。阅后删除此段。

附录正文样式与文章正文相同：宋体、小四；行距：22 磅；间距段前段后均为 0 行。阅后删除此段。

致 谢

值此论文完成之际，首先向我的导师……

致谢正文样式与文章正文相同：宋体、小四；行距：22 磅；间距段前段后均为 0 行。阅后删除此段。