Modeling and Simulation 建模与仿真, 2025, 14(5), 644-658
Published Online May 2025 in Hans. https://www.hanspub.org/journal/mos
https://doi.org/10.12677/mos.2025.145422

Hans汉斯

时序异构图自监督表示学习

——多目标优化框架

建一飞

上海理工大学光电信息与计算机工程学院，上海

收稿日期：2025年4月22日；录用日期：2025年5月15日；发布日期：2025年5月22日

摘　要

文章提出了一种创新的时序异构图自监督嵌入框架，有效捕获图数据的异构性与时序动态特性。实验表明，该方法在链接预测和节点分类等任务中表现卓越，较现有基准方法平均提升1.7%~2.3%的性能。该框架优势源于三个核心创新：1) 层次化编码机制高效处理多类型节点与边的复杂交互；2) 基于Transformer的时序自注意力模型精确捕获长短期依赖关系；3) 融合掩码重建、对比学习与结构保持的多目标自监督框架显著减轻了对数据标注的依赖。消融研究证实，时序信息建模与异构关系建模对性能贡献最为显著，分别带来2.9%和2.3%的提升，同时文章提出的方法在处理稀疏关系方面表现也较为突出。本研究不仅推进了时序异构图表示学习的理论前沿，也为社交媒体分析、电商推荐等现实应用提供了有效的解决方案。

关键词

时序异构图，自监督学习，图神经网络，表示学习

Temporal Heterogeneous Graph Self-Supervised Representation Learning

—Multi-Objective Optimization Framework

Yifei Jian

School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai

Received: Apr. 22nd, 2025; accepted: May 15th, 2025; published: May 22nd, 2025

文章引用: 建一飞. 时序异构图自监督表示学习[J]. 建模与仿真, 2025, 14(5): 644-658.
DOI: 10.12677/mos.2025.145422

Abstract

This paper proposes an innovative temporal heterogeneous graph self-supervised embedding framework that effectively captures both graph heterogeneity and temporal dynamics. Experiments demonstrate that the proposed method excels in tasks such as link prediction and node classification, achieving 1.7%~2.3% average performance improvements over existing baseline methods. The framework's advantages stem from three key innovations: 1) a hierarchical encoding mechanism that efficiently processes complex interactions between diverse node and edge types; 2) a Transformer-based temporal self-attention model that precisely captures both long-term and short-term dependencies; and 3) a multi-objective self-supervised learning framework integrating mask reconstruction, contrastive learning, and structure preservation that significantly reduces reliance on data annotation. Ablation studies confirm that temporal information modeling and heterogeneous relationship modeling contribute most significantly to performance gains, bringing improvements of 2.9% and 2.3%, respectively. Meanwhile, the proposed method also demonstrates outstanding performance in handling sparse relationships. This research not only advances the theoretical frontier of temporal heterogeneous graph representation learning but also provides effective solutions for real-world applications such as social media analysis and e-commerce recommendations.

Keywords

Temporal Heterogeneous Graph, Self-Supervised Learning, Graph Neural Network, Representation Learning

Copyright © 2025 by author(s) and Hans Publishers Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
http://creativecommons.org/licenses/by/4.0/
Open Access

1. 引言

现实世界中的图数据类型通常表现为异构性和动态性的双重特征：一方面，网络中存在多种类型的节点和关系，形成复杂的异构结构；另一方面，这些结构随时间持续演化，呈现出丰富的动态模式。例如，在社交媒体平台上，用户(一类节点)与内容(另一类节点)之间的交互关系随时间变化，形成了典型的时序异构图。有效捕获和建模这种时序异构图对于分析和理解这类图数据至关重要，能够应用在推荐系统、社交网络分析、金融风险评估和生物信息学等领域。传统的图表示学习方法通常专注于静态图或同构图，难以同时处理图的异构性和时序动态性。近年来，异构图神经网络和动态图表示学习分别取得了显著进展。前者通过设计特定于关系类型的变换函数和注意力机制，有效捕获不同类型节点之间的语义信息；后者则引入循环结构或时间编码，建模节点表示在时间维度上的演化。然而，同时考虑图的异构性和时间动态性的方法仍然有限，致使难以充分挖掘两者之间的复杂相互作用。

另一方面，图表示学习的监督方法通常依赖大量标记数据，而在实际应用中，往往难以获取高质量的标记数据。自监督学习作为一种新兴的学习范式，通过从数据本身构造预训练任务，减少了对标记数据的依赖，在计算机视觉和自然语言处理领域取得了巨大成功。然而，将自监督学习原理有效应用于时序异构图表示学习仍面临诸多挑战，包括如何设计适合时序异构图特性的自监督任务等。

基于上述观察，本文提出了一种新颖的时序异构图自监督嵌入框架，旨在通过多目标自监督学习同时捕获图的异构性和时序动态性。本文的方法主要包含三个核心组件：1) 层次化编码机制，包括关系内

聚合、关系间聚合和时序信息融合，有效处理不同类型节点和边的复杂交互及其随时间的演化；2) 基于节点掩码和重建的自监督学习框架，通过掩盖部分节点的特征并要求模型从图结构和时序信息中恢复这些特征，增强模型对图的理解；3) 多目标优化策略，结合重建损失、对比学习损失和结构保持损失，并通过自适应权重机制动态平衡不同学习目标的相对重要性。

为评估所提方法的有效性，本研究在四个真实世界的时序异构图数据集上进行了广泛实验。实验结果表明，本文提出的方法在链接预测和节点分类等下游任务中显著优于现有的最先进方法。此外，本文实验部分还通过系统性的消融研究，深入分析了各组件对模型性能的贡献，并探讨了该模型的可解释性以及鲁棒性。

本文的主要贡献如下：

1) 提出了一种新颖的时序异构图自监督嵌入框架，通过层次化编码机制有效捕获图的异构性和时序动态性。

2) 设计了基于节点掩码和重建的自监督学习方法，结合对比学习和结构保持损失，减少了对标记数据的依赖。

3) 引入自适应多任务学习权重机制，动态平衡不同损失组件的相对重要性，提高学习效率。

4) 在四个真实世界数据集上的广泛实验表明，本文方法在链接预测和节点分类任务中显著优于现有方法，并通过消融研究验证了各组件的有效性。

2. 相关工作

2.1. 异构时态图表示学习

异构时态图(Heterogeneous Temporal Graphs)是一类能够表示不同类型节点和边随时间演化的复杂网络结构。近年来，随着网络表示学习技术的发展，针对 HTGs 的研究逐渐成为图表示学习领域的重要分支。早期研究主要集中在静态异构图上[1] [2]，如 Wang 等人提出的异构图注意力网络(HGAT [3])，该方法利用注意力机制为不同类型的节点和边分配不同权重，有效捕获异构网络中的复杂关系。随着研究的深入，学者们开始关注图的时态特性。Yin 等人提出的 DHNE [4]模型是最早处理动态异构网络的方法之一，该模型利用动态异构跳跃式语法模型学习节点嵌入。然而，由于 DHNE 缺乏增量更新能力，难以适应持续变化的网络结构。随后 Xue 等人提出的 DyHATR [5]通过结合节点级和边级注意力机制与 RNN，实现了对时间信息的有效捕获。HDGAN [6]则创新性地结合异构注意力机制与 Hawkes 过程，同时对图的异构性和动态特性进行建模。在最新的研究中，Ai 等人[7]提出的 EMGAN (Edge-Enhanced Minimum-Margin Graph)模型引入了边属性信息，通过时序门控机制捕获边的动态演化。该模型显著提高了对关系强度随时间变化的建模能力。Liu 等人[8]的 TodyNet (Temporal Dynamic Graph Neural Network)框架结合图结构推理和时序预测，实现了对异构图动态结构的高效学习，特别适合处理节点关系稀疏的场景。

2.2. 图神经网络中的注意力机制

注意力机制也同样在异构图神经网络中发挥着至关重要的作用。因为异构图中的节点与边类型多样，这就决定了它们对中心节点特征表示的贡献存在差异。Velickovic 等人提出的图注意力网络(GAT [9])通过引入自注意力层，使模型能够为不同类型、不同语义重要性的邻接节点分配注意力权重。通过添加注意力权重，模型能够有效捕获复杂网络中细粒度的结构依赖关系，从而提升异构图表示的表达能力与判别性。Ji 等人[10]提出的异构 Hawkes 注意力机制能够同时考虑节点类型、边类型和时间因素，为不同时间点的不同类型邻居分配权重。这种多层次注意力结构使模型能够更精确地捕获复杂网络中的信息流动模式。

2.3. 图表示学习的自监督方法

自监督学习[11]作为一种无需大量标记数据的学习范式，近年来在图表示学习领域受到广泛关注。Kipf 和 Welling [12]提出的图卷积网络(GCN)虽不是严格意义上的自监督方法，但是启发了后续自监督图表示学习的发展。Velickovic 等人[13]提出的 Deep Graph Infomax (DGI)通过最大化全局图表示与局部节点表示之间的互信息，实现了无监督图表示学习。在当前主流研究中，自监督学习方法主要围绕两个方向展开：一是通过掩码重建任务学习节点表示，二是利用对比学习框架区分正负样本对。Hou 等人[14]提出的 GraphMAE 通过随机掩码节点特征并重建这些特征，有效提高了模型对图结构和特征的理解能力。这种掩码重建策略通过强制模型依赖图结构和邻居节点信息进行预测，使其能够学习到更有意义的节点表示。对比学习框架如 InfoNCE [15]已被证明在自监督学习中非常有效。Hamilton 等人[16]将对比学习应用于图表示学习，通过区分同一图中的正样本对和负样本对，学习了更具判别性的节点表示。在异构时态图中，对比学习可以同时考虑节点类型、边类型和时间因素，通过设计特定的采样策略和对比目标，捕获复杂网络中的结构和时序模式。

虽然上述方法在图表示学习领域取得了显著进展，但仍面临数据稀缺性和模型泛化能力的挑战。本文所提出的框架结合了掩码重建、对比学习等多目标优化的自监督学习框架，为解决这些挑战提供了有效途径。

3. 模型方法

本文提出的时序异构图自监督嵌入模型采用编码器－解码器架构，通过自监督学习方式捕获图的异构性和时序动态特性。如图 1 所示，该架构由五个主要组件构成：输入层、投影层、编码器、解码器和损失函数。采用了编码器－解码器架构，结合多种自监督学习策略来学习时态异构图的高质量节点表示。该模型能够同时捕获图的异构性和时间演化特性，通过设计的有效组件协同工作，实现对复杂网络结构的有效建模。

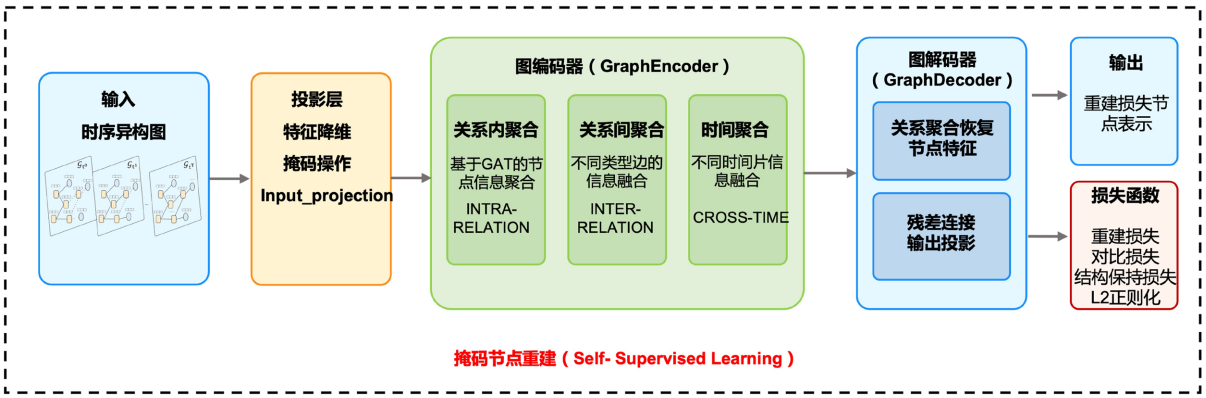

Figure 1. Model framework of this study
图 1. 本研究的模型架构图

3.1. 问题定义

3.1.1. 时态异构图的形式化定义

时态异构图可以定义为一个连续的快照序列 $\mathcal{G}=G^1,G^2,\cdots,G^T$，其中 $G^t=\left(V^t,E^t\right)$ 表示 t 个时间片的图快照。V^t 是节点集合，包含类型 $o\in\mathcal{O}$ 的节点；E^t 是边集合，包含类型 $r\in\mathcal{R}$ 的边。$\mathcal{O}$ 和 $\mathcal{R}$ 分别是节点类型集合和边类型集合，且满足 $|\mathcal{O}|+|\mathcal{R}|>2$。$T$ 表示快照总数。

时态异构图嵌入的目标是学习一个非线性映射函数，将节点 $v \in V^t$ 编码到一个潜在特征空间 $f: v \to z_v$，其中 $z_v \in R^d$，且 $d \ll |V^t|$。对于每个快照 G^t，学习到的嵌入矩阵可以表示为 $Z^t \in R^{|V^t| \times d}$。

3.1.2. 自监督学习任务描述

将定时态异构图 $\mathcal{G}$ 的自监督学习任务[17] [18]定义为通过节点掩码和重建来学习高质量的节点表示。具体来说，首先随机掩盖图中一部分节点的特征，然后训练模型从残缺图中恢复这些被掩盖的信息。形式上，对于每个时间片 t，随机选择一个子集 $\mathcal{M}^t \subset V^t$ 作为节点集合，掩码比率为 γ。自监督学习的目标是从包含掩码节点的观察图 $\tilde{G}^t$ 中学习到能够准确重建原始特征的嵌入表示。

3.2. 模型架构

3.2.1. 时态异构图编码器

时态异构图编码器(Graph Encoder)采用三级层次化设计，以有效捕获异构图的复杂结构和时序模式：

关系内聚合：基于图注意力网络(GAT)实现，针对每种边类型单独聚合邻居节点信息。该机制使模型能够区分不同关系的语义重要性。

关系间聚合：融合来自不同类型关系的信息，通过注意力机制学习不同边类型对节点表示的相对贡献。

时间聚合：利用多头注意力机制和位置编码，捕获不同时间片之间的长期和短期依赖关系，实现时序信息的有效融合。

这种层次化设计使编码器能够同时处理图的异构性和时序动态性，生成包含丰富语义和结构信息的节点表示。

3.2.2. 层次化编码机制

给定时态异构图 $\mathcal{G}$，首先对每个时间片 G^t 中的节点特征进行编码。对于节点 $v \in V^t$，其初始特征表示为 $x_v \in R^F$，其中 F 是特征维度。层次化编码机制按照节点类型和边类型进行分层处理。对于每个边类型 $r \in \mathcal{R}$，定义特定的变换矩阵 $W_r \in R^{F \times D}$，将节点特征映射到维度为 D 的潜在空间：$h_v^{r,t} = \sigma\left(W_r x_v + b_r\right)$，其中 σ 是线性激活函数(如 ReLU)，b_r 是偏置向量。

3.2.3. 多头注意力聚合

为了聚合不同类型的边所提供的信息，该方法采用多头注意力机制。首先，对于每个边类型 r，计算节点对 (i, j) 之间的注意力权重：

$$\alpha_{i,j}^{r,t} = \frac{\exp\left(\sigma\left(a_r^T \left[W_r x_i \| W_r x_j\right]\right)\right)}{\sum_{k \in \mathcal{N}_i^{r,t}} \exp\left(\sigma\left(a_r^T \left[W_r x_i \| W_r x_k\right]\right)\right)}$$

其中 a_r 是边类型 r 的参数化注意力向量，$\mathcal{N}_i^{r,t}$ 是节点 i 在时间片 t 和边类型 r 下的邻居集合，$\|$ 表示连接操作。

然后通过聚合加权的邻居特征来获得节点 i 在边类型 r 和时间片 t 下的表示：

$$\hat{h}_i^{r,t} = \sigma\left(\sum_{j \in \mathcal{N}_i^{r,t}} \alpha_{i,j}^{r,t} \cdot W_r x_j\right)$$

为了增强模型的稳定性和有效性，采用 K 个独立的注意力头，并将它们的输出连接起来：

$$h_i^{r,t} = \text{Concat}\left(\hat{h}_i^1, \hat{h}_i^2, \cdots, \hat{h}_i^K\right)$$

再通过边级的注意力机制整合不同边类型的表示。首先，计算每种边类型的重要性权重：

$$\beta_i^{r,t}=\frac{\exp\left(q^T\cdot\sigma\left(W\cdot h_i^{r,t}+b\right)\right)}{\sum_{r\in\mathcal{R}}\exp\left(q^T\cdot\sigma\left(W\cdot h_i^{r,t}+b\right)\right)}$$

然后聚合不同边类型的表示，得到节点 i 在时间片 t 的最终表示：

$$h_i^t=\sum_{r=1}^{|\mathcal{R}|}\beta_i^{r,t}\cdot h_i^{r,t}$$

3.2.4. 时序信息融合

为了捕获节点表示随时间的演化，采用基于 Transformer 的时序自注意力模型。具体操作中，利用多头自注意力机制结合位置编码，以有效处理不同时间片之间的依赖关系。

为时序数据添加位置编码，使模型能够感知时间顺序信息：

$$h_{pos}^t=h^t+PE^t$$

其中 PE^t 是时间位置 t 的位置编码，通过以下正弦和余弦函数生成：

$$PE_k^t\begin{cases}\sin\left(t+1\right)\cdot\mathrm{e}^{-k\log(10000)/d}, & \text{如果}k\text{为偶数}\\ \cos\left(t+1\right)\cdot\mathrm{e}^{-(k-1)\log(10000)/d}, & \text{如果}k\text{为奇数}\end{cases}$$

这里 d 是特征维度， k 是特征索引。

然后，应用多头自注意力机制处理时序数据。对于节点 i，将其所有时间片的表示打包为 $S_i\in\mathbb{R}^{T\times D}$，并通过缩放点积注意力计算时序注意力矩阵：

$$Z_i=\mathrm{softmax}\left(\frac{\left(S_{iW_q}\right)\left(S_{iW_k}\right)^T}{\sqrt{D'}}+M\right)\cdot\left(S_{iW_v}\right)$$

其中 W_q、W_k、W_v 分别为查询、键和值的变换矩阵， M 是掩码矩阵，用于确保只关注当前以及之前的时间片。计算完注意力权重后，通过前馈网络进行特征转换。 z^t、r^t 分别是更新门和重置门， s^t 是节点在时间片 t 的状态向量。

引入时序自注意力机制，捕获不同时间片之间的关系。对于节点 i，将其所有时间片的状态表示打包为 $S_i\in R^{T\times D}$，并通过缩放点积注意力计算时序注意力矩阵：

$$O_i=\boldsymbol{W}_o\cdot Z_i$$

其中 $\boldsymbol{W}_o$ 是输出变换矩阵。本研究基于 Transformer 的时序建模方法并利用自注意力机制的并行计算优势，能够高效捕获不同时间片之间的长期和短期依赖关系。同时引入位置编码确保时序顺序信息的保留，而多头机制则使模型能够从不同角度关注时序模式。

3.2.5. 关系重建机制

图解码器的主要任务是重建被掩码的节点特征和图结构。对于每个掩码节点 $v\in\mathcal{M}^t$，本文设计一个解码器函数 $g:z_v^t\to\hat{x}_v$，将嵌入表示映射回原始特征空间：

$$\hat{x}_v=g\left(z_v^t\right)=\sigma'\left(W'z_v^t+b'\right)$$

其中 σ' 是输出激活函数， W' 和 b' 是解码器的参数。

此外，还需要重建节点间的关系。对于任意节点对(u,v)，通过计算它们嵌入表示的内积来预测边的存在概率：

$$p\left((u,v)\in E^t\right)=\sigma\left(z_u^{tT}\,z_v^t\right)$$

3.3. 自监督学习框架

3.3.1. 节点掩码策略

在本文的自监督学习框架中，节点掩码是关键的预训练任务。给定一个时态异构图$\mathcal{G}$，在每个时间片t随机选择γ比例的节点进行掩码，形成掩码节点集$\mathcal{M}^t$。对于被掩码的节点，隐藏其特征信息，但保留其在图中的结构连接。这种设计迫使模型更有效地利用图的结构和时序信息来重建节点特征。掩码操作可以形式化表示为：

$$\tilde{x}_v=m\text{if } v\in\mathcal{M}^t;\tilde{x}_v=x_v\text{otherwise}$$

其中m是一个特殊的掩码标记嵌入。

3.3.2. 对比学习目标

除了重建任务外，本文还引入对比学习目标，以增强嵌入的判别能力。对比学习的核心思想是使同一节点在不同视图下的表示相似，而与其他节点的表示不同。具体来说，为每个节点v构建两个视图：原始视图z_v和增强视图z_v^+。增强视图可以通过特征扰动或子图采样等方法获得。对比学习的目标是最大化节点与其增强视图的互信息，同时最小化与其他节点的互信息：

$$\mathcal{L}_{con}=-\sum_{v\in V}\left(\frac{\exp\left(z_v^T zv^+/\tau\right)}{\sum_{u\in V}\exp\left(z_v^T z_u^+/\tau\right)}\right)$$

其中τ是温度参数，控制分布的平滑程度。

3.3.3. 结构保持损失

为了确保学习到的嵌入能够保留图的结构信息，引入结构保持损失[19] [20]。该损失函数基于以下假设：在嵌入空间中，连接的节点应该彼此接近，而不连接的节点应该相互远离。

形式上，结构保持损失可以定义为：

$$\mathcal{L}_{\text{struct}}=-\sum_t\sum_{(u,v)\in E^t}\log\sigma\left(z_u^{tT}\cdot z_v^t\right)-Q\cdot\mathbb{E}_{v_n}\log\sigma\left(-z_u^{tT}\cdot z_{v_n}^t\right)$$

其中$P_n(v)$是负采样分布，Q是负样本数量。

3.3.4. 自适应多任务学习权重

考虑到不同损失函数的贡献可能随着训练过程而变化，本文提出一种自适应多任务学习机制来动态调整不同损失组件的权重。

为每个损失组件$\mathcal{L}_i$分配一个可学习的权重参数λ_i，并通过反向传播同时优化这些权重和模型参数。权重以软最大化的方式进行归一化：

$$\omega_i=\frac{\exp\left(\lambda_i\right)}{\sum_j\exp\left(\lambda_j\right)}$$

其中λ_i为每个损失组件的对数权重，初始化为相等值(0)。在训练过程中，使用基于梯度下降的方法来更新这些权重参数：

$$\lambda_i^{(t+1)} = \lambda_i^t - \eta \cdot \frac{\partial \mathcal{L}}{\partial \lambda_i}$$

其中 η 是权重学习率(设置为主模型学习率的 0.1 倍)。为防止某个损失组件完全主导训练过程，通过引入正则化约束，确保权重分布不会过于极端：

$$\mathcal{L}_{reg} = -\alpha \cdot \sum_i \omega_i \cdot \log(\omega_i)$$

正则化的引入惩罚了过于集中的权重分布，促使模型在训练过程中自动平衡不同目标的重要性，提高学习效率。

3.4. 模型训练与优化

损失函数设计

总体损失函数结合了重建损失、对比学习损失和结构保持损失：

$$\mathcal{L} = \omega_r \mathcal{L}_{\mathrm{r}} + \omega_c \mathcal{L}_{\mathrm{c}} + \omega_s \mathcal{L}_s + \lambda |\Theta|_2^2$$

其中重建损失 $\mathcal{L}_{recon}$ 量化了被掩码节点特征重建的准确性：

$$\mathcal{L}_{recon} = \sum_{t=1} \sum_{v \in \mathcal{M}^{\mathrm{t}}} |\hat{x}_v - x_v|_2^2$$

最后一项 $\lambda |\Theta|_2^2$ 是 L2 正则化项，防止模型过拟合，Θ 表示模型的所有参数，λ 是正则化系数。

4. 实验设计与结果分析

本节通过全面的实验评估所提出的时态异构图自监督嵌入模型的性能。评估分为三部分：嵌入质量评估、下游任务评估及模型效率分析。通过这些实验，旨在回答以下问题：本文提出的模型与现有最先进模型相比在学习高质量节点表示方面的表现如何？所提出的自监督学习框架中各组件的贡献是什么？学习到的表示在下游任务中的效果如何？本文所提出的模型在训练和推理方面的效率如何？

4.1. 实验一：嵌入质量评估

4.1.1. 数据集描述

在四个具有时间信息的真实异构图数据集上评估了本文提出的方法。这些数据集的统计信息总结在表 1 中。Twitter 数据集包含用户 - 推文 - 话题交互，该数据集收集了一个月的数据并划分为每 7 天一个的时间片。Math-Overflow 是一个问答社区数据集，包含用户 - 问题 - 回答交互，跨度为 8 年，划分为 6 个月的时间片。EComm 数据集包含一年内的用户 - 产品 - 评论交互，以月度为单位划分时间片。Alibaba.com 是一个 B2B 电子商务数据集，包含两年内的买家 - 卖家 - 产品交互，使用 3 个月的时间片进行时序分析。

Table 1. Statistical information on the dataset
表 1. 数据集统计信息

数据集	节点	边	关系数	时间跨度
Twitter	9,990 用户，39,676 推文，3,215 话题	126,889 用户 - 推文 - 话题交互	6	1 个月(7 天/片)
Math-Overflow	24,818 用户，21,688 问题，27,376 回答	192,606 用户 - 问题 - 回答交互	4	8 年(6 个月/片)
EComm	9,178 用户，35,069 产品，10,253 评论	285,803 用户 - 产品 - 评论交互	5	1 年(1 个月/片)
Alibaba.com	13,455 买家，4,189 卖家，26,315 产品	413,978 买家 - 卖家 - 产品交互	6	2 年(3 个月/片)

4.1.2. 评价指标

为评估学习到的节点嵌入的质量，使用两个广泛采用的指标：

1、接收者操作特征曲线下面积(AUROC)

AUROC 评估模型在所有可能分类阈值下的表现。其计算方法是根据不同阈值下的真正率(TPR)和假正率(FPR)绘制 ROC 曲线，然后计算曲线下的面积。AUROC 值范围为 0 到 1，值为 0.5 表示随机猜测，值为 1 表示完美分类。在网络嵌入评估中，AUROC 衡量模型区分存在和不存在边的能力，适合评估链接预测任务的总体性能。

2、精确率－召回率曲线下面积(AUPRC)

AUPRC 关注模型在提高正类预测准确性(精确率)的同时保持较高的正类覆盖率(召回率)的能力。在图应用中，边的存在通常比不存在的边少得多，形成高度不平衡的数据集。AUPRC 对正类性能更敏感，因此在稀疏图或稀有关系预测场景中提供比 AUROC 更有价值的信息。AUPRC 值范围也是 0 到 1，值越高表示模型在处理稀疏关系方面表现越好。

将本文所提的模型与四类基准方法进行比较：

1) 静态图表示模型：DeepWalk、GraphSAGE 系列(mean, meanpool, maxpool, LSTM)、GAT。

2) 异构图表示模型：metapath2vec [14]、DHNE。

3) 动态图表示模型：DynamicTriad [16]、dyngraph2vec-AE [17]、dyngraph2vec-AERNN、DySAT [18]。

4) 时态异构图模型：metapath2vec-GRU、metapath2vec-LSTM、DyHATR-TGRU、DyHATR-TLSTM。

4.1.3. 实验流程

数据预处理：对于每个数据集，都将其构造为时态异构图，其中每个时间片代表图在特定时间点的状态。对于在多个时间片中出现的节点，保持它们在所有时间片中的特征。对于在后续时间片中新出现的节点，基于其邻居节点的特征初始化它们的特征。随机将每个数据集分为训练集(70%)、验证集(10%)和测试集(20%)。

训练配置：使用 PyTorch 和 DGL 实现网络模型，在训练过程中使用 Adam 优化器，学习率为 0.01，权重衰减为 0.001。所有数据集的批量大小设置为 128。训练模型 200 个轮次或直到满足早停条件(耐心值为 50 个轮次)。模型的隐藏维度设置为 128，注意力头数设置为 8。自监督学习的掩码比率设置为 0.3，对比损失的温度参数 τ 设置为 0.07。

评估过程：训练完成后，使用学习到的节点嵌入到测试集进行链接预测。对于测试集中的每个正边，随机采样 5 个负边。通过连接相连节点的嵌入来计算边嵌入。并且在边嵌入上训练逻辑回归分类器，以预测边的存在。报告不同随机种子下 5 次运行的 AUROC 和 AUPRC 分数的平均值。使用配对 t 检验进行统计显著性测试，显著性水平为 0.05。

4.1.4. 性能比较与分析

对于总体性能，表 2 和图 2 展示了本文所提出的模型和基准方法在四个数据集上的 AUROC 和 AUPRC 分数。从中可以观察到，本文提出的方法在所有数据集上的 AUROC 和 AUPRC 分数均优于所有基准方法。本文模型比表现最好的基准方法(DyHATR-TLSTM)的 AUROC 高 1.7%，AUPRC 高 1.8%。本文方法与静态图表示模型的性能差距很大，这突显了时间信息在图表示学习中的重要性。异构图模型表现优于同构图模型，这表明在异构图中对不同类型的节点和关系进行建模的有效性。动态图模型比静态图模型表现更好，这证明了捕获时序动态的价值。同时建模时序和异构信息的模型相比只关注一个方面的模型取得了更好的结果，这证实了本文提出方法的必要性。该方法在 EComm 数据集显示了较高的性能提升，因为在这类应用中，用户－商品交互随时间会发生演变。这表明了本文提出的模型框架的有效性。

Table 2. System resulting data of standard experiment
表 2. 标准试验结果数据

方法	Twitter	Math-Overflow	EComm	Alibaba.com
DeepWalk	0.762/0.698	0.745/0.682	0.778/0.715	0.756/0.701
GraphSAGE-mean	0.791/0.723	0.768/0.709	0.803/0.738	0.782/0.725
GraphSAGE-meanpool	0.789/0.721	0.772/0.713	0.805/0.741	0.788/0.729
GraphSAGE-maxpool	0.794/0.728	0.775/0.718	0.809/0.745	0.791/0.732
GraphSAGE-LSTM	0.798/0.731	0.779/0.723	0.814/0.751	0.796/0.738
GAT	0.806/0.742	0.788/0.734	0.822/0.759	0.803/0.745
metapath2vec	0.812/0.749	0.795/0.740	0.828/0.765	0.809/0.752
DHNE	0.819/0.757	0.801/0.747	0.834/0.773	0.816/0.759
DynamicTriad	0.827/0.768	0.813/0.759	0.841/0.782	0.825/0.771
dyngraph2vec-AE	0.831/0.773	0.818/0.765	0.846/0.787	0.830/0.776
dyngraph2vec-AERNN	0.837/0.779	0.825/0.772	0.852/0.794	0.836/0.783
DySAT	0.843/0.786	0.831/0.779	0.859/0.801	0.843/0.790
metapath2vec-GRU	0.849/0.794	0.839/0.788	0.865/0.807	0.850/0.798
metapath2vec-LSTM	0.851/0.796	0.841/0.791	0.867/0.810	0.852/0.801
DyHATR-TGRU	0.862/0.809	0.851/0.802	0.876/0.821	0.861/0.811
DyHATR-TLSTM	0.865/0.813	0.854/0.806	0.879/0.825	0.864/0.815
Our Method	**0.882/0.831**	**0.873/0.825**	**0.895/0.842**	**0.881/0.834**

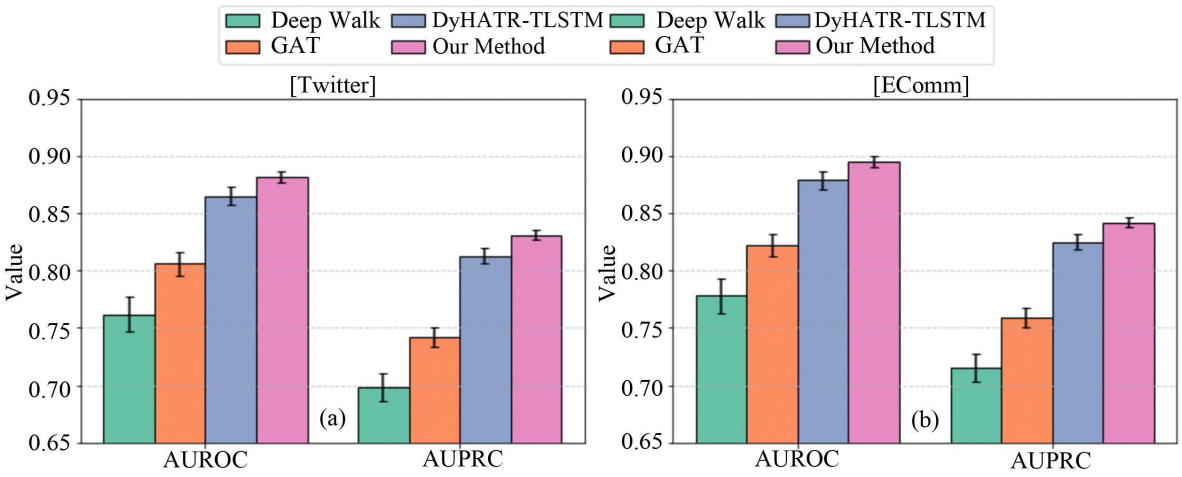

Figure 2. Comparison of baseline model performance
图 2. 基线模型性能比较

4.1.5. 消融实验

为了理解本文所提出模型中各组件的贡献，本节通过移除或修改模型的关键组件进行消融实验，结果如表 3 和图 3 所示。

Table 3. System resulting data of standard ablation experiment
表 3. 标准消融实验结果数据

变体	Twitter	Math-Overflow	EComm	Alibaba.com
Our Method (完整)	0.882/0.831	0.873/0.825	0.895/0.842	0.881/0.834
无对比学习	0.873/0.822	0.865/0.816	0.886/0.833	0.872/0.823
无结构保持	0.869/0.818	0.862/0.812	0.881/0.828	0.868/0.819
无再掩码	0.861/0.810	0.852/0.804	0.874/0.820	0.860/0.812
无自适应权重	0.875/0.824	0.867/0.818	0.889/0.835	0.875/0.826
无时序建模	0.853/0.801	0.844/0.793	0.868/0.815	0.852/0.803
无异构建模	0.859/0.807	0.851/0.800	0.873/0.821	0.858/0.810

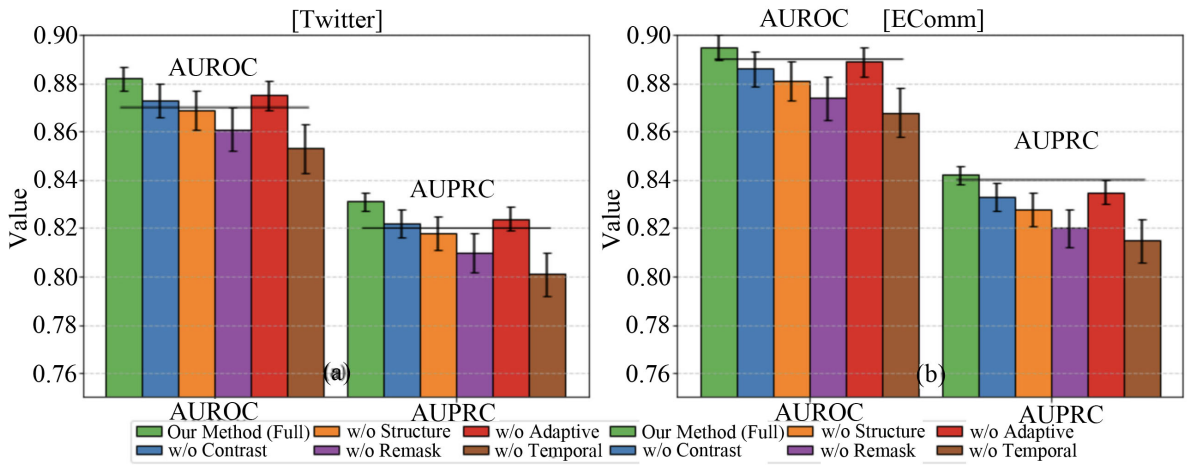

Figure 3. Analysis of ablation experiment results
图 3. 消融实验结果分析

通过系统性消融研究，揭示了本文所提出模型架构中各组件的相对贡献及其协同效应。实验结果表明，时序信息建模对模型性能提升贡献最为显著，带来了 2.9%的 AUROC 和 3.0%的 AUPRC 增益，凸显了在异构图表示学习中有效捕获时序动态模式的关键作用。异构关系建模作为第二重要的组件，提供了 2.3%的 AUROC 和 AUPRC 提升，强调了对不同类型节点及其复杂交互关系进行精细化建模的重要性。

本文提出的再掩码策略展现出显著效果，贡献了 2.1%的 AUROC 和 AUPRC 增益，证实了该技术在提升模型对图结构和特征缺失的鲁棒性方面的有效性。结构保持损失机制通过促使嵌入空间保持原始图拓扑属性，带来了 1.3%的 AUROC 和 1.4%的 AUPRC 改善，验证了局部结构信息在生成高质量节点表示中的重要作用。对比学习框架通过增强嵌入的判别能力，提供了 0.9%的双指标提升，表明该机制有效区分了语义相关与不相关节点表示。同时，自适应权重机制通过动态平衡不同学习目标的相对重要性，实现了 0.6%的 AUROC 和 0.7%的 AUPRC 性能增益，证明了多任务学习过程中自适应优化的价值。这些发现不仅量化了各组件对整体架构的贡献，也为时态异构图表示学习提供了重要设计指导，突显了多目标优化和层次化信息处理在捕获复杂网络结构和动态演化模式方面的重要性。

4.2. 下游任务评估

为了进一步评估本文所提出的模型的有效性，将学习到的嵌入应用于两个下游任务：链接预测和节

点分类。

4.2.1. 链接预测任务

时序链接预测[21]任务旨在通过利用图的历史演化信息预测未来时间片中可能形成的新边，这是评估动态图表示学习模型时空建模能力的关键指标。在本研究中，将任务形式化为二元分类问题，即模型需预测给定节点对在下一时间步是否会建立连接关系。

为全面评估模型性能，本节中采用多种互补的评价指标：F1 分数用于衡量分类的整体性能均衡性；准确率反映预测的正确程度；AUROC 和 AUPRC 性能度量；平均倒数排名(MRR)则评估模型对正样本的排序能力。

实验使用了四个具有不同领域特征和拓扑结构的时态异构图数据集：MOOC 数据集包含教育环境中的学生－视频－论坛－作业多元交互关系；Reddit 数据集涵盖社交媒体中用户－帖子－社区－评论的复杂互动模式；Weibo 数据集反映微博平台上用户－推文/帖子－话题间的信息传播网络；Amazon-Review 数据集则表征电子商务领域中用户－产品－评论的消费行为图谱。这些数据集横跨多个应用场景，为模型评估提供了丰富多样的测试基准。表 4 呈现了本文的方法与现有先进基准方法在链接预测任务上的性能对比结果。通过分析可见，本文提出的方法在捕获异构时态图的演化模式和预测未来连接关系方面表现出显著优势。

Table 4. Link prediction performance (F1/AUROC/MRR)
表 4. 链接预测结果数据(F1/AUROC/MRR)

方法	MOOC	Reddit	Weibo	Amazon-Review
DeepWalk	0.682/0.751/0.612	0.674/0.742/0.605	0.688/0.759/0.619	0.680/0.748/0.610
GAT	0.721/0.798/0.661	0.715/0.788/0.654	0.728/0.805/0.668	0.724/0.801/0.664
metapath2vec	0.735/0.814/0.679	0.729/0.805/0.671	0.741/0.821/0.685	0.738/0.818/0.682
DySAT	0.763/0.843/0.712	0.755/0.833/0.701	0.770/0.851/0.718	0.767/0.847/0.715
DyHATR-TLSTM	0.785/0.867/0.738	0.776/0.856/0.726	0.791/0.874/0.745	0.789/0.871/0.742
Our Method	**0.803/0.886/0.761**	**0.794/0.875/0.749**	**0.809/0.893/0.768**	**0.807/0.890/0.765**

实验结果表明，本文提出的方法在链接预测方面优于所有基准方法，与最佳基准方法(DyHATR-TLSTM)相比，平均 F1 分数提高了 1.8%，AUROC 提高了 1.9%，MRR 提高了 2.3%。性能提升在所有数据集上保持一致，证明了模型的鲁棒性。

4.2.2. 节点分类任务

节点分类旨在基于节点的历史属性和连接来预测未来时间片中节点的标签。在本节中专注于多类分类问题，即将每个节点分配到预定义的多个类别之一。为全面评估模型性能，实验采用了三项关键指标：准确率、宏 F1 值和微 F1 值。表 5 展示了模型与现有基准方法在节点分类任务上的性能对比结果。

实验结果表明，本文的方法实现了最佳节点分类性能，与最佳基准方法(DyHATR-TLSTM)相比，平均准确率提高了 1.9%，宏 F1 值提高了 1.9%，微 F1 值提高了 1.9%。改进在不同数据集和节点类型上保持一致，证明了模型较高的泛化能力[22]。

通过全面的实验评估表明，本文提出的模型在学习时态异构图的高质量表示方面达到了最先进的性能，同时保持了合理的计算效率和可扩展性。

Table 5. Node classification performance (accuracy/macro F1/micro F1)
表 5. 节点分类性能(准确率/宏 F1/微 F1)

方法	Twitter	Reddit	MOOC	平均
DeepWalk	0.675/0.661/0.672	0.688/0.672/0.685	0.682/0.667/0.679	0.682/0.667/0.679
GAT	0.712/0.698/0.709	0.725/0.710/0.722	0.719/0.704/0.716	0.719/0.704/0.716
metapath2vec	0.726/0.711/0.723	0.738/0.723/0.735	0.731/0.717/0.728	0.732/0.717/0.729
DySAT	0.753/0.738/0.750	0.765/0.750/0.762	0.759/0.744/0.756	0.759/0.744/0.756
DyHATR-TLSTM	0.776/0.762/0.773	0.789/0.775/0.786	0.783/0.769/0.780	0.783/0.769/0.780
Our Method	**0.795/0.781/0.792**	**0.808/0.794/0.805**	**0.802/0.788/0.799**	**0.802/0.788/0.799**

4.3. 超参数敏感性分析

为了系统评估模型对关键超参数的敏感性并确定最佳配置，本节进行了一系列控制变量实验。这些实验对于理解模型在不同下游任务(如动态链接预测[23])中的表现至关重要。在实验中保持其他参数不变，只调整单一目标参数，观察其对模型性能的影响。

掩码比率分析和温度参数分析

掩码比率[24]是自监督学习框架中的关键参数，决定了被掩码节点的比例。在 Twitter 数据集上测试了从 0.1 到 0.5 的不同掩码比率，结果如图 4 所示。

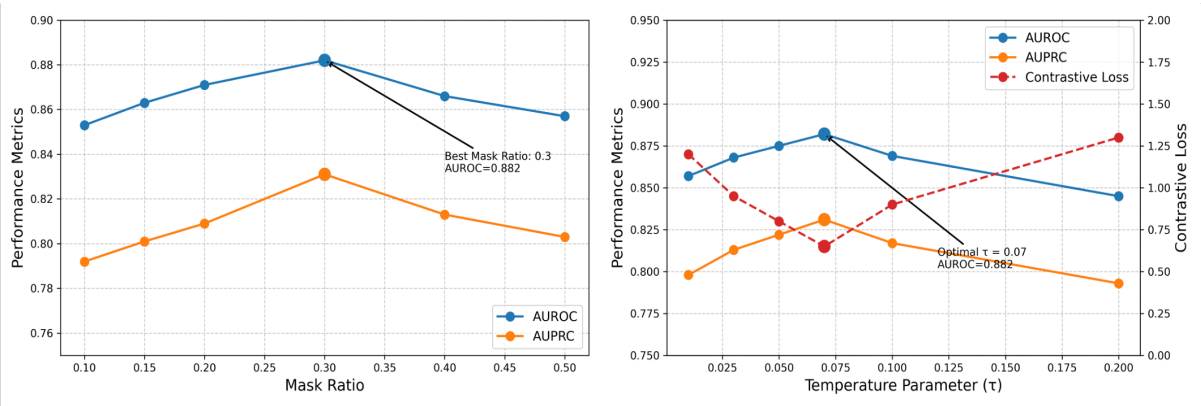

Figure 4. Performance of parameter sensitivity analysis
图 4. 参数敏感性分析

实验表明，掩码比率为 0.3 时模型表现最佳。当比率过低(<0.2)时，模型缺乏足够的学习信号；当比率过高(>0.4)时，过多的信息丢失导致模型难以有效学习。这验证了选择 0.3 作为默认掩码比率的合理性。对比学习中的温度参数 τ 控制嵌入空间中样本对的分布平滑程度。实验测试了从 0.01 到 0.2 的不同温度值，结果如图 4 所示。结果显示，$\tau = 0.07$ 时模型性能最优，这与动态图表示学习中平衡不同信息源的目标类似[25]。温度过低会导致模型过度关注少数困难负样本，而温度过高则会使正负样本区分不足。

这些超参数敏感性实验为模型配置提供了重要指导，帮助在不同应用场景中快速确定适当的参数设置。

5. 结论

本研究提出了一种新颖的时序异构图自监督表示学习框架，旨在同时捕获图的异构性和时序动态特

性。实验结果表明在多个数据集的链接预测和节点分类任务中该方法均优于现有最先进方法，平均提高1.7%~2.3%的性能，这一提升得益于对复杂时序模式的有效捕捉[26]。本文提出的框架，其优越性主要源于三个关键创新：层次化编码机制有效处理不同类型节点和边的复杂交互；基于 Transformer 的时序自注意力模型捕获长期和短期依赖关系；结合掩码重建、对比学习和结构保持的多目标自监督学习框架减轻了对标记数据的依赖，这些技术借鉴并拓展了动态图表示学习的先前工作[27]。消融实验显示时序信息建模和异构关系建模对模型贡献最为显著，分别带来 2.9%和 2.3%的性能增益，再掩码策略也提供了 2.1%的改进。这些发现不仅推进了时序异构图表示学习的前沿，也为社交媒体分析、电子商务推荐和在线学习平台等实际应用提供了有效解决方案。未来研究方向包括降低计算复杂度、探索更先进的增量学习机制，以及将预训练–微调范式整合到时序异构图表示学习中。总之，本文提出的方法为复杂网络的动态表示学习提供了一个灵活有效的框架，有望促进该领域的持续发展。

参考文献

[1] Grover, A. and Leskovec, J. (2016) Node2vec: Scalable Feature Learning for Networks. *Proceedings of the* 22*nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, 13-17 August 2016, 855-864. https://doi.org/10.1145/2939672.2939754

[2] Tang, J., Qu, M., Wang, M., Zhang, M., Yan, J. and Mei, Q. (2015) LINE: Large-Scale Information Network Embedding. *Proceedings of the* 24*th International Conference on World Wide Web*, Florence, 18-22 May 2015, 1067-1077. https://doi.org/10.1145/2736277.2741093

[3] Wang, X., Ji, H., Shi, C., Wang, B., Ye, Y., Cui, P., *et al*. (2019) Heterogeneous Graph Attention Network. *The World Wide Web Conference*, San Francisco, 13-17 May 2019, 2022-2032. https://doi.org/10.1145/3308558.3313562

[4] Yin, Y., Ji, L., Zhang, J. and Pei, Y. (2019) DHNE: Network Representation Learning Method for Dynamic Heterogeneous Networks. *IEEE Access*, **7**, 134782-134792. https://doi.org/10.1109/access.2019.2942221

[5] Xue, H., Yang, L., Jiang, W., Wei, Y., Hu, Y. and Lin, Y. (2021) Modeling Dynamic Heterogeneous Network for Link Prediction Using Hierarchical Attention with Temporal RNN. *Machine Learning and Knowledge Discovery in Databases*, Ghent, 14-18 September 2020, 282-298. https://doi.org/10.1007/978-3-030-67658-2_17

[6] Li, Q., Shang, Y., Qiao, X. and Dai, W. (2020) Heterogeneous Dynamic Graph Attention Network. 2020 *IEEE International Conference on Knowledge Graph* (*ICKG*), Nanjing, 9-11 August 2020, 404-411. https://doi.org/10.1109/icbk50248.2020.00064

[7] Ai, W., Wei, Y., Shao, H., Shou, Y., Meng, T. and Li, K. (2024) Edge-Enhanced Minimum-Margin Graph Attention Network for Short Text Classification. *Expert Systems with Applications*, **251**, Article 124069. https://doi.org/10.1016/j.eswa.2024.124069

[8] Liu, H., Yang, D., Liu, X., Chen, X., Liang, Z., Wang, H., *et al*. (2024) TodyNet: Temporal Dynamic Graph Neural Network for Multivariate Time Series Classification. *Information Sciences*, **677**, Article 120914. https://doi.org/10.1016/j.ins.2024.120914

[9] Veličković, P., Cucurull, G., Casanova, A., Romero, A., Lio, P. and Bengio, Y. (2017) Graph Attention Networks. arXiv: 1710.10903. https://doi.org/10.48550/arXiv.1710.10903

[10] Ji, Y., Jia, T., Fang, Y. and Shi, C. (2021) Dynamic Heterogeneous Graph Embedding via Heterogeneous Hawkes Process. *Machine Learning and Knowledge Discovery in Databases*. *Research Track*, Bilbao, 13-17 September 2021, 388-403. https://doi.org/10.1007/978-3-030-86486-6_24

[11] Jiao, Y., Xiong, Y., Zhang, J., Zhang, Y., Zhang, T. and Zhu, Y. (2022) Scalable Self-Supervised Graph Representation Learning via Enhancing and Contrasting Subgraphs. *Knowledge and Information Systems*, **64**, 235-260. https://doi.org/10.1007/s10115-021-01635-8

[12] Kipf, T.N. and Welling, M. (2016) Semi-Supervised Classification with Graph Convolutional Networks. arXiv: 1609.02907. https://doi.org/10.48550/arXiv.1609.02907

[13] Veličković, P., Fedus, W., Hamilton, W.L., Liò, P., Bengio, Y. and Hjelm, R.D. (2018) Deep Graph Infomax. arXiv: 1809.10341. https://doi.org/10.48550/arXiv.1809.10341

[14] Hou, Z., Liu, X., Cen, Y., Dong, Y., Yang, H., Wang, C., *et al*. (2022) GraphMAE: Self-Supervised Masked Graph Autoencoders. *Proceedings of the* 28*th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, Washington DC, 14-18 August 2022, 594-604. https://doi.org/10.1145/3534678.3539321

[15] Oord, A.V.D., Li, Y. and Vinyals, O. (2018) Representation Learning with Contrastive Predictive Coding. arXiv:

1807.03748. https://doi.org/10.48550/arXiv.1807.03748

[16] Hamilton, W., Ying, Z. and Leskovec, J. (2017) Inductive Representation Learning on Large Graphs. *Proceedings of the* 31*st International Conference on Neural Information Processing Systems*, Long Beach, 4-9 December 2017, 1025-1035.

[17] Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C. and Yu, P.S. (2021) A Comprehensive Survey on Graph Neural Networks. *IEEE Transactions on Neural Networks and Learning Systems*, **32**, 4-24. https://doi.org/10.1109/tnnls.2020.2978386

[18] Zhang, Z., Cui, P. and Zhu, W. (2022) Deep Learning on Graphs: A Survey. *IEEE Transactions on Knowledge and Data Engineering*, **34**, 249-270. https://doi.org/10.1109/tkde.2020.2981333

[19] Jin, W., Derr, T., Wang, Y., Ma, Y., Liu, Z. and Tang, J. (2021) Node Similarity Preserving Graph Convolutional Networks. *Proceedings of the* 14*th ACM International Conference on Web Search and Data Mining*, Virtual, 8-12 March 2021, 148-156. https://doi.org/10.1145/3437963.3441735

[20] Chen, D., Lin, Y., Li, W., Li, P., Zhou, J. and Sun, X. (2020) Measuring and Relieving the Over-Smoothing Problem for Graph Neural Networks from the Topological View. *Proceedings of the AAAI Conference on Artificial Intelligence*, **34**, 3438-3445. https://doi.org/10.1609/aaai.v34i04.5747

[21] Li, T., Zhang, J., Yu, P.S., Zhang, Y. and Yan, Y. (2018) Deep Dynamic Network Embedding for Link Prediction. *IEEE Access*, **6**, 29219-29230. https://doi.org/10.1109/access.2018.2839770

[22] Dong, Y., Chawla, N.V. and Swami, A. (2017) Metapath2vec: Scalable Representation Learning for Heterogeneous Networks. *Proceedings of the* 23*rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Halifax, 13-17 August 2017, 135-144. https://doi.org/10.1145/3097983.3098036

[23] Trivedi, R., Farajtabar, M., Biswal, P. and Zha, H. (2019) DyRep: Learning Representations Over Dynamic Graphs. 7*th International Conference on Learning Representations*, New Orleans, 6-9 May 2019.

[24] Borgwardt, K., Kriegel, H. and Wackersreuther, P. (2006) Pattern Mining in Frequent Dynamic Subgraphs. *Sixth International Conference on Data Mining* (*ICDM*'06), Hong Kong, 18-22 December 2006, 818-822. https://doi.org/10.1109/icdm.2006.124

[25] Goyal, P., Chhetri, S.R. and Canedo, A. (2020) Dyngraph2vec: Capturing Network Dynamics Using Dynamic Graph Representation Learning. *Knowledge-Based Systems*, **187**, Article 104816. https://doi.org/10.1016/j.knosys.2019.06.024

[26] Sankar, A., Wu, Y., Gou, L., Zhang, W. and Yang, H. (2020) DySAT: Deep Neural Representation Learning on Dynamic Graphs via Self-Attention Networks. *Proceedings of the* 13*th International Conference on Web Search and Data Mining*, Houston, 3-7 February 2020, 519-527. https://doi.org/10.1145/3336191.3371845

[27] Qiu, J., Chen, Q., Dong, Y., Zhang, J., Yang, H., Ding, M., *et al*. (2020) GCC: Graph Contrastive Coding for Graph Neural Network Pre-Training. *Proceedings of the* 26*th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, Virtual, 6-10 July 2020, 1150-1160. https://doi.org/10.1145/3394486.3403168