

基于时空注意力的多粒度链路预测算法^{*}

何玉林^{1,2}, 赖俊龙², 崔来中², 尹剑飞², 黄哲学^{1,2}

¹(人工智能与数字经济广东省实验室(深圳), 广东 深圳 518107)

²(深圳大学 计算机与软件学院, 广东 深圳 518060)

通信作者: 何玉林, E-mail: yulinhe@gml.ac.cn



摘 要: 社交网络链路预测有助于揭示网络节点之间的潜在联系, 在好友推荐、合作预测等方面有着重要的实际应用价值. 然而, 现有的链路预测方法忽略了社交网络时间序列的中、长期发展趋势, 且没有从长期的角度考虑网络中节点之间的相互影响关系. 针对以上问题, 提出基于时空注意力的多粒度链路预测算法, 该算法能够融合不同粒度社交网络时间序列的时空特征以提升链路预测的准确性. 首先, 以时间衰减函数构建社交网络快照图的权重, 提出图加权移动平均策略, 生成反映短期、中期和长期趋势的不同粒度社交网络时间序列; 然后, 利用基于多头注意力机制的神经网络提取社交网络序列的全局时间特征; 接着, 结合社交网络序列内节点的历史交互信息, 通过基于掩码注意力机制的神经网络从长期角度自适应地构建网络拓扑结构, 以动态地调整节点之间的相互影响, 并结合图卷积网络建模空间信息; 最后, 提出融合注意力神经网络, 从短期、中期和长期时空特征中提取出有用的短期、中期和长期信息, 并进行特征融合, 准确地预测未来社交网络的链接. 在 4 种社交网络公开数据集上与 7 种现有的链路预测算法的实验对比证实所提方法的有效性和优越性.

关键词: 社交网络; 链路预测; 多粒度; 注意力机制; 图卷积网络

中图法分类号: TP311

中文引用格式: 何玉林, 赖俊龙, 崔来中, 尹剑飞, 黄哲学. 基于时空注意力的多粒度链路预测算法. 软件学报, 2025, 36(9): 4311–4326. <http://www.jos.org.cn/1000-9825/7286.htm>

英文引用格式: He YL, Lai JL, Cui LZ, Yin JF, Huang ZX. Multi-granularity Link Prediction Algorithm Based on Spatiotemporal Attention. Ruan Jian Xue Bao/Journal of Software, 2025, 36(9): 4311–4326 (in Chinese). <http://www.jos.org.cn/1000-9825/7286.htm>

Multi-granularity Link Prediction Algorithm Based on Spatiotemporal Attention

HE Yu-Lin^{1,2}, LAI Jun-Long², CUI Lai-Zhong², YIN Jian-Fei², HUANG Zhe-Xue^{1,2}

¹(Guangdong Laboratory of Artificial Intelligence and Digital Economy (Shenzhen), Shenzhen 518107, China)

²(College of Computer Science and Software Engineering, Shenzhen University, Shenzhen 518060, China)

Abstract: Social network link prediction can help to reveal the potential connections between network nodes, and has important practical application value in friend recommendation and cooperation prediction. However, existing link prediction algorithms ignore the medium and long-term development trend of social network time series, and do not consider the interaction relationship between nodes in the network from a long-term perspective. To address the above-mentioned problems, a spatiotemporal attention-based multi-granularity link prediction algorithm is proposed, which can integrate the spatiotemporal features of social network time series with different granularities to improve the accuracy of link prediction. Firstly, the weights of the social network snapshot graph are constructed with the time decay function, and a graph-weighted moving average strategy is proposed to generate social network time series with different granularities reflecting short-term, medium-term, and long-term trends. Then, a neural network based on the multi-head attention mechanism is used to extract the global temporal features of social network sequences. Next, the historical interaction information of nodes within social network

^{*} 基金项目: 广东省自然科学基金面上项目 (2023A1515011667); 深圳市科技重大专项 (202302D074); 深圳市基础研究项目 (JCYJ20220818100205012, JCYJ20210324093609026)

收稿时间: 2024-04-08; 修改时间: 2024-07-26; 采用时间: 2024-08-26; jos 在线出版时间: 2025-01-15

CNKI 网络首发时间: 2025-01-15

sequences is combined, and the neural network based on the mask attention mechanism is used to adaptively construct the network topology from a long-term perspective to dynamically adjust the interactions between nodes and is combined with graph convolutional network to model spatial information. Finally, the fusion attention neural network is proposed to extract useful short-term, medium-term and long-term information from short-term, medium-term and long-term spatiotemporal features, and perform feature fusion to accurately predict the future links of social networks. Experimental comparisons with seven existing link prediction algorithms on four social network public datasets confirm the effectiveness and superiority of the proposed method.

Key words: social network; link prediction; multi-granularity; attention mechanism; graph convolutional network (GCN)

链路预测是指利用已有的社交网络结构信息预测未来的链接关系, 主要目标是揭示网络中节点之间的潜在联系, 从而帮助理解网络的演化过程、推断未知的关系、发现新的社区结构等. 链路预测在社交网络分析中起着非常重要的作用, 社交网络由节点和边组成, 节点代表个体 (如用户、人物、组织等), 边则代表节点之间的关系 (如友谊、关注、合作等). 在社交网络中, 链路预测的应用场景包括但不限于朋友推荐^[1]、合著关系预测^[2]、社区发现^[3]等. 因此, 研究社交网络链路预测具有重要的实际应用价值, 尤其是在新媒介技术高速发展的今天.

社交网络链路预测任务主要包括对社交网络快照序列的空间结构特征和时间特征进行提取, 以及选择合适的模型预测未来可能出现或断开的链接关系. 不同于一般的网络结构, 如拓扑固定不变的城市交通网络, 社交网络的结构随时间变化而剧烈变化. 现实社交网络中, 用户之间的互动随时发生. 因此, 解决社交网络链路预测任务的关键是建模社交网络时间序列的结构信息和时间信息. 目前, 已有的链路预测方法可以归纳为基于相似性指数的方法、基于矩阵分解的方法以及基于深度学习的方法这 3 类.

现有的链路预测方法通常是对社交网络快照序列中的每个快照提取空间结构特征, 在此基础上再提取该序列的时间特征, 最后利用学习到的时空特征完成对未来链路的预测. 这是一种有效的研究思路, 能够充分建模社交网络演化过程中用户与用户之间的相互影响以及用户好友关系随时间的动态变化. 这些方法取得了很好的预测效果, 但也存在一定的局限性.

● 首先, 在社交网络快照序列中, 对每个快照提取空间特征的做法仅考虑了短期的社交网络拓扑结构, 在学习节点表示时, 将每个邻居的影响程度一致看待会引入冗余信息. 例如, 在图 1 的快照序列内, 用户 a 与用户 c 只互动了一次, 而与用户 b 频繁互动, 这意味着用户 b 对用户 a 的影响程度更大, 但现有的方法无法考虑这种情况, 简单地认为用户 b 和用户 c 对用户 a 的影响程度一样.

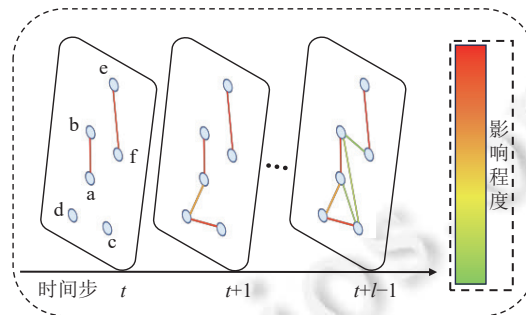


图 1 社交网络时间序列节点相互影响示意图

● 此外, 现有方法通常直接分析原始社交网络的时间序列数据, 忽略了社交网络中期、长期的演化特征. 原始序列数据波动较大, 无法从中洞察网络的中长期演化趋势. 例如, 在图 2 中, 节点 1 与节点 3 之间的互动时间序列波动较大, 反映了节点间短期的变化趋势. 相较之下, 节点互动的 4 步移动平均线和 7 步移动平均线变化趋势较为平缓, 更清晰地呈现了节点间互动的变化趋势. 然而, 现有方法并未将这些序列信息考虑在内.

针对上述问题, 本文提出一种基于时空注意力的多粒度链路预测 (spatiotemporal attention-based multi-granularity link prediction algorithm, SAMG-LP) 算法. 为了充分考虑社交网络中用户之间的相互影响程度和社交网络时间序列的长、短期趋势, 本文从构建自适应邻接矩阵和不同粒度社交网络时间序列两个层面进行社交网络链

路预测. 首先, 扩展用于一维动态数据的加权移动平均^[4], 提出图加权移动平均 (graph weighted moving average, GWMA) 策略来生成不同粒度的社交网络时间序列. 其次, 构建基于多头注意力机制的神经网络提取社交网络快照序列的全局时间特征. 然后, 结合社交网络快照序列内节点之间的交互信息, 通过基于掩码注意力机制的神经网络学习社交网络邻接矩阵, 从长期的角度自适应地学习节点之间的相互影响, 并结合图卷积网络建模空间信息. 最后, 提出融合注意力神经网络来聚合不同粒度时间序列的时空特征, 并通过多层全连接层完成链路预测任务. 本文的主要工作总结如下.

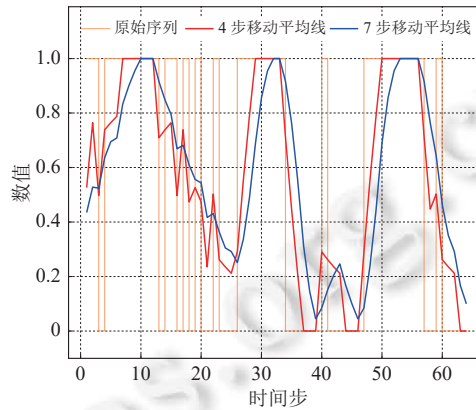


图2 不同粒度社交网络时间序列示意图 (节点1与节点3互动情况)

(1) 结合时间衰减函数提出了图加权移动平均 GWMA 策略, 生成不同粒度的社交网络时间序列. 在此基础上, 提出了用于融合不同粒度时空特征的融合注意力神经网络, 充分考虑了社交网络的短期、中期和长期演化趋势.

(2) 提出了基于掩码注意力机制的图卷积网络, 考虑社交网络快照序列内节点的历史交互信息, 引入掩码注意力机制自适应地学习社交网络的邻接矩阵, 基于该矩阵完成图卷积操作, 有效地学习网络中节点的相互影响关系.

(3) 提出了基于注意力的多粒度链路预测 SAMG-LP 算法. 利用多头注意力神经网络充分地提取社交网络快照序列的时间特征, 以数据驱动的方式构建社交网络的拓扑结构, 利用图卷积网络提取空间特征, 从而得到时空特征, 最后基于注意力神经网络完成不同粒度时空特征的融合与预测.

(4) 在4种真实社交网络数据集上进行了验证性实验, 与7种现有的链路预测算法进行对比验证了 SAMG-LP 算法在社交网络链路预测问题上的优越性. 并对 SAMG-LP 模型进行了系统的消融实验, 证明了模型组件的有效性.

本文第1节介绍链路预测的相关方法和研究现状. 第2节提出基于时空注意力的多粒度链路预测 SAMG-LP 算法. 第3节通过对比实验、消融实验等实验验证所提算法的有效性. 第4节总结全文并提出展望.

1 链路预测相关工作

本节将对3类具有代表性的链路预测方法, 即基于相似性指数的方法、基于矩阵分解的方法和基于深度学习的方法进行介绍.

1.1 基于相似性指数的方法

基于相似性指数的方法通常利用节点之间的相似性来推断它们之间是否存在链接. Sharma 等人^[5]同时考虑网络的局部和全局特性, 提出了一种新的相似性指数. 计算过程分为两个部分, 一个部分是求解节点之间最短路径长度的倒数, 另一个部分是对最短路径中涉及的节点, 求这些节点度数倒数之和, 并求平方根再取倒数, 最后使用平衡因子对这两部分进行加权求和. Jiang 等人^[6]使用互信息量化节点行为同步指标, 并将局部相似性指数与节点行为同步指标相结合进行链路预测, 提出了 SSBS 指数. Charikhi^[7]将 PageRank 算法与局部相似性指数相结合, 以

提高性能并保持局部方法的低复杂性优势。

基于相似性指数方法的另一种思路是以相似性指数作为特征,通过不同的分类或回归模型来推断节点之间是否存在链接。Kumar 等人^[8]以 6 种节点中心性作为特征,并通过 LGBM 分类器进行链路预测,提出了 NC-LGBM 算法。Chaubey 等人^[9]使用 Jaccard 指数、Adar 指数等多种相似性指数作为拓扑特征,使用随机森林和 XG-Boost 算法进行社交网络的链路预测。Choudhury 等人^[10]基于局部聚类系数定义了社区动态,计算每个网络快照的社区动态值得到社区动态时间序列,基于该序列引入了时间相似度、基于相关性的相似度和时间社区感知相似度共 3 种动态相似性度量指数,并将这 3 种相似性指数作为特征输入到多个不同的机器学习分类器中进行预测对比。

基于相似性指数的方法通常比较简单,计算效率较高,预测结果的可解释性较强,但也存在过度依赖节点之间的相似性度量,而忽略了网络结构中的其他信息和特征的缺点,同时需人工提取特征,这些特征难以学习到节点之间复杂的非线性关系,导致预测效果不佳。

1.2 基于矩阵分解的方法

基于矩阵分解的链路预测方法的一般思路是将网络的邻接矩阵分解成多个低秩矩阵的乘积用于构建目标函数,通过最小化目标函数将网络中的节点和边映射到一个低维向量空间,最后利用学习到的潜在表示预测缺失的链接或者新的链接。Ma 等人^[11]将链路预测任务建模为多标签学习任务,基于多标签学习项、特征提取项和时间平滑性项这 3 个部分建立目标函数,并使用交替优化的方式最小化目标函数,提出了 MljFE。Lv 等人^[12]提出了两种新颖的链路预测算法。他们将自编码器结构与深度非负矩阵分解模型相融合,充分挖掘动态网络中隐藏的深层次特征,并通过图正则化和 PageRank 进一步提取网络快照中的局部和全局信息。Yan 等人^[13]从数学推导的角度介绍了用于学习列空间信息的 NMF-C 和用于学习行空间信息的 NMF-R,最后将这些方法结合起来建立了 NMF-CR 模型。Lin 等人^[14]结合原始邻接矩阵和 PageRank 构建了全局连接矩阵,然后在局部邻接矩阵和全局连接矩阵上基于自编码器的结构执行深度矩阵分解,以获取包含丰富结构信息的全局和局部特征表示,此外,引入流形正则化项和稀疏条件来进一步约束深度矩阵分解过程。Mahmoodi 等人^[15]采用对抗性非负矩阵分解,并添加基于 CN 相似性的图形正则化项,交替优化目标函数,提出了 LPANMF 算法。Nasiri 等人^[16]提出了 RGNMF-AN 算法,该算法通过随机游走得到相似性矩阵,然后将相似性矩阵、原始网络结构和节点属性相结合,从而得到目标函数。Chen 等人^[17]提出了 FSSDNMF 算法,它利用 CN 计算网络结构的相似性,再利用深度非负矩阵分解将其映射到多层低维潜在空间,进一步融合稀疏性约束以完成链路预测任务。Yao 等人^[18]将原始网络分成 3 个子网络,应用深度非负矩阵分解模型挖掘子网络内的复杂关系,从而重建网络,然后对重建的网络使用非负矩阵分解进行链路预测。

基于矩阵分解的方法可以有效地建模动态网络的时间特征和空间结构特征,能够考虑更多影响链路预测的因素,例如网络节点的属性,同时这些方法所涉及的模型参数较少,具有较好的计算效率。但这些方法难以学习到数据中的非线性特征,导致预测性能受限,且由于计算过程涉及大量矩阵运算,模型容易出现数值不稳定的问题。

1.3 基于深度学习的方法

基于深度学习的方法目标是将图数据通过合适的深度学习模型转换为节点的特征表示,这些模型可以有效地提取网络的拓扑结构特征和时间特征,从而有效解决链路预测任务。Min 等人^[19]利用图神经网络提取每个网络快照中节点的空间特征,然后设计了一种基于注意力机制的神经网络提取网络快照序列的时间特征,使模型能够有意识地对不同的历史快照进行加权关注,从而提出了 STGSN 模型。Jiao 等人^[20]提出了 TVAE 算法,该方法的思路是利用编码器学习网络快照的低维嵌入,然后使用基于自注意力机制的神经网络学习网络快照的结构特征,接着基于 LSTM 模型建模网络快照序列的时间特征,最后使用两层全连接作为解码器完成链路预测任务。Chen 等人^[21]提出了 GC-LSTM 模型,该模型将图卷积网络嵌入到 LSTM 中,旨在同时学习节点的局部结构特征和动态网络的时间特征。Tan 等人^[22]设计了特征掩码、边扰动和时间步掩码这 3 种策略对图数据进行增强,然后利用增强函数生成两个视图,接着使用图卷积网络学习两个视图的嵌入,并设计了时间感知和频率感知的图对比学习模块,实现更具区分性的节点表示学习,最后利用节点表示进行链路预测。刘林峰等人^[23]利用多层全连接层作为编码器学习

网络的低维表示, 然后使用 GRU 提取时间特征, 最后基于多层全连接层的解码器预测未来链路. 何鹏等人^[24]提出了一种基于图卷积网络的软件系统类之间交互预测方法. 该方法对每个类的内部图进行图卷积以获取类节点的初始特征, 通过对类外部图的图卷积来更新类节点的表示向量, 最后计算类节点之间的评估值以进行交互预测. 赵琳琳等人^[25]提出 TGE-LBSN, 该方法第 1 阶段设计了 IVGS 算法生成初始节点向量, 第 2 阶段对每个社交网络快照提取结构信息, 具体地, 设计了 SAN 策略选取有代表性的邻居节点, 然后设计 SVG 算法学习每个快照中节点的表示. 最后, 结合注意力机制将不同快照融合, 得到节点的最终表示, 进而完成好友推荐任务. Qin 等人^[26]设计了由 GCN、GRU 和对齐单元组成的堆叠结构 GR. 一方面, GCN 和 GRU 的堆叠结构用于提取动态图的时空特征, 另一方面, 注意力节点对齐单元使 RNN 能够解决不同快照中节点数不同的问题. Yin 等人^[27]从图像超分辨率的角度出发, 引入 LR 图和 HR 图的概念, 通过条件归一化流实现 LR 图和 HR 图之间的转换, 从而提出了 SRG 算法. Mei 等人^[28]提出了 GCN_MA 算法. 他们首先提出了 NRNAE 算法, 利用节点度、聚类系数和邻居关系来丰富节点信息表示, 然后以 LSTM 更新 GCN 参数的方式减少模型的参数量, 同时提取时空特征, 最后基于多头注意力进一步学习节点的表示, 增强模型的表征能力以完成链路预测任务.

基于深度学习的方法通过端到端的学习方式直接从原始数据中学习到节点的表示, 无需手工提取特征, 这些方法通过多层非线性变换来学习数据的非线性特征, 可以同时考虑多种影响因素, 具有较强的表达能力和泛化能力. 考虑到现有的基于深度学习的方法仅考虑短期变化趋势的社交网络时间序列, 而忽略了社交网络时间序列的中期和长期变化趋势. 此外, 这些方法没有考虑从长期的角度区分社交网络中节点之间的相互影响来预测未来链接. 针对这些问题, 本文提出图加权移动平均策略来生成不同粒度的时间序列, 以反映社交网络演化的中期、长期趋势, 并基于注意力神经网络有效地提取了社交网络的时间特征和空间特征. 此外, 提出了融合注意力神经网络来融合不同粒度的时空特征, 充分考虑了社交网络时间序列的多种特征.

2 SAMG-LP 算法

2.1 问题定义

定义 1. 社交网络. 社交网络是由节点和节点之间的关系 (通常表示为边或链接) 组成的网络结构, 它是动态变化的, 形式化表示为一个由一系列网络快照构成的集合 $G = \{G_1, G_2, \dots, G_\tau\}$, 其中 $G_t = (V, E_t)$ 代表第 t 个时间步的网络快照, $t \in \{1, 2, \dots, \tau\}$. 所有网络快照共享同一节点集 V 且每个网络快照是无向无权网络. 第 t 个时间步的网络快照 G_t 包含一组节点集 V 和一组边集 E_t , 使用邻接矩阵 $A_t \in R^{N \times N}$ 描述其相应的静态拓扑, 其中 N 是节点集的数量. 特别地, 当第 t 个时间步中的节点 i 和节点 j 之间存在一条边, 即: $(i, j) \in E_t$, 则 $(A_t)_{i,j} = (A_t)_{j,i} = 1$; 否则 $(A_t)_{i,j} = (A_t)_{j,i} = 0$.

定义 2. 社交网络链路预测. 社交网络链路预测定义为, 给定先前观察到的 l 个网络快照 $\{G_1, G_2, \dots, G_l\}$ 预测第 $l+1$ 个时间步的网络快照 G_{l+1} . 形式化的描述如下:

$$\hat{G}_{l+1} = f(G_1, G_2, \dots, G_l) \quad (1)$$

其中, $\hat{G}_{l+1}$ 为网络快照 G_{l+1} 的预测值, $f(\cdot)$ 为社交网络链路预测算法.

2.2 基于时空注意力的多粒度链路预测算法 SAMG-LP

本文算法整体框架如图 3 所示. 首先, 利用 GWMA 策略从原始社交网络时间序列生成反映中期变化趋势的新的时间序列, 基于中期趋势时间序列再次利用 GWMA 策略生成反映长期趋势的新的时间序列, 以此得到 3 种不同粒度的时间序列, 分别表示社交网络时间序列的短期、中期和长期趋势. 然后, 使用基于多头注意力的神经网络从全局的角度依次提取不同粒度时间序列的时间信息. 接着, 基于掩码注意力神经网络从全局时间特征中自适应地调整图结构中节点之间的连接强度, 以更准确地反映图结构关系, 并基于自适应拓扑结构对时间特征进行图卷积操作, 通过消息传递的方式来更新节点的特征表示. 最后, 进一步利用融合注意力神经网络来融合不同粒度的时空特征, 使得模型能够自适应地聚焦于不同粒度的时空上下文, 实现更准确的社交网络链路预测.

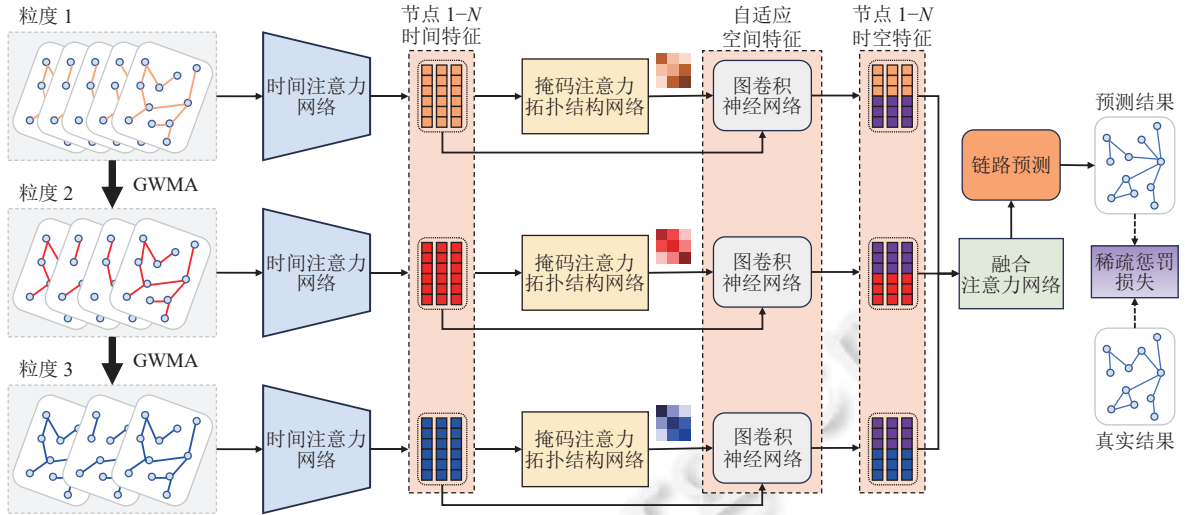


图 3 SAMG-LP 算法框架

2.3 构建不同粒度的时间序列

在社交网络的链路预测任务中, 现有研究都是在原始社交网络快照序列上提取特征, 由于该序列波动性较大, 这会导致节点的演化趋势不够明显, 难以直观地理解网络长期趋势. 此外, 当原始时间序列中存在噪声和波动时, 这些噪声和波动可能会影响预测效果. 为了解决这个问题, 受一维动态数据中移动平均线的启发, 移动平均线可以平滑数据中的噪声和波动, 有助于去除短期波动, 使得数据的长期趋势更加明显, 还可以一定程度地过滤掉数据中的异常值, 提高数据的稳定性和可靠性^[29]. 本文提出了图加权移动平均策略 GWMA, 从原始社交网络快照序列中提取出网络演化的中、长期趋势, 生成新的不同粒度的社交网络时间序列.

给定一个长度为 L 的社交网络快照序列 $\{G_1, G_2, \dots, G_L\}$, 其对应的邻接矩阵序列表示为 $\{A_1, A_2, \dots, A_L\}$, 考虑到历史数据对未来值的影响往往随时间间隔的增长而减弱, 为此, 本文使用时间衰减函数给予不同的网络快照以不同的权重, 使得越接近未来时间步的快照权重越大. 给定图移动平均的项数 n , GWMA 用公式表述如下:

$$\tilde{A}_k = \sum_{t=1}^n d_t \cdot A_{t+k-1} / \sum_{t=1}^n d_t \quad (2)$$

其中, $d_t = (1-\theta)^{n-t}$ 是衰减因子, 表示不同项的权重, θ 是时间衰减函数的参数. $\tilde{A}_k$ 表示 n 项加权移动平均值, $1 \leq k \leq L-n+1$. 为了便于理解, 本文使用如下示例来展示当 $L=4, n=4, \theta=0.1$ 时 $\tilde{A}_k$ 的计算过程, 具体如图 4 所示.

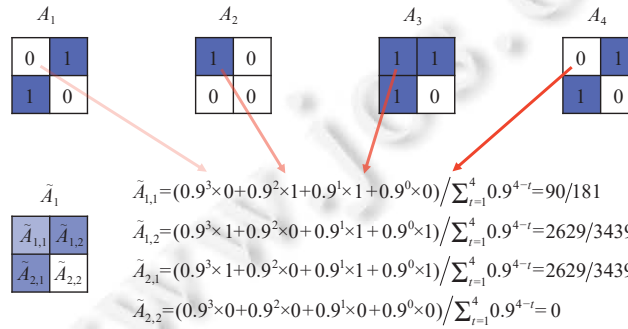


图 4 GWMA 策略示例

对于原始社交网络快照序列, 本文将其表示为 $G^1 = \{G_1^1, G_2^1, \dots, G_L^1\}$, 它表示粒度 1 社交网络时间序列, 其相应的邻接矩阵序列表示为 $A^1 = \{A_1^1, A_2^1, \dots, A_L^1\}$. 将 G^1 利用 GWMA 进行计算, 得到粒度 2 社交网络时间序列 G^2 , 其中

$G^2 = \{G_1^2, G_2^2, \dots, G_{l-n+1}^2\}$, G^2 对应的邻接矩阵序列表示为 $A^2 = \{A_1^2, A_2^2, \dots, A_{l-n+1}^2\}$. 然后, 将 G^2 利用 GWMA 进行计算, 得到粒度 3 社交网络时间序列 $G^3 = \{G_1^3, G_2^3, \dots, G_{l-2n+2}^3\}$, 其对应的邻接矩阵序列为 $A^3 = \{A_1^3, A_2^3, \dots, A_{l-2n+2}^3\}$. 至此生成了 3 种不同粒度的社交网络时间序列, 这 3 个时间序列分别作为短期分量、中期分量和长期分量输入到模型中, 以分别表示社交网络演化过程的短期、中期和长期趋势.

2.4 时间注意力网络

在时间维度方面, 许多工作采用循环神经网络来学习网络序列的时间信息, 这是一种有效的方式, 但其在处理长序列时可能会丢失重要的信息, 导致模型在预测时考虑不到全局信息, 且每个时间步都使用相同的权重来处理输入序列. 为了解决这些问题, 基于注意力机制的神经网络被提出, 并在机器翻译、文本摘要等序列任务中取得了显著的成果. 在本文工作中, 使用基于注意力机制的神经网络提取社交网络序列的全局时间特征, 模型在每个时间步根据输入的不同位置信息来动态地调整权重, 从而更好地学习到关键信息.

在建模社交网络序列的时间信息之前, 本文先使用一个全连接层对社交网络进行约简, 以较低的维度表示网络中的每一个节点. 以短期分量 G^1 为例, 对序列内的每个快照进行约简, 具体如公式 (3) 所示:

$$[H_1^1, H_2^1, \dots, H_l^1] = FC([A_1^1, A_2^1, \dots, A_l^1]) \quad (3)$$

其中, FC 是表示线性变换的全连接层, $H_k^1 \in R^{N \times F}$, $1 \leq k \leq l$ 是约简后的社交网络快照表示, N 是节点数, F 是约简维数. 约简后的短期分量表示为 $G_H^1 = [H_1^1, H_2^1, \dots, H_l^1] \in R^{N \times l \times F}$.

现实中, 社交网络等动态图因多种潜在因素的影响而呈现多元演化的特点^[30]. 例如在工作日和节假日, 社交网络呈现不同的演化趋势. 为了建模社交网络的多元演化, 本文使用多头注意力神经网络从不同角度学习的社交网络的演化. 通过对约简后的社交网络快照表示序列 G_H^1 进行线性变化, 得到 Q_i, K_i, V_i , 计算如公式 (4) 所示:

$$[Q_i, K_i, V_i] = G_H^1 [W_i^Q, W_i^K, W_i^V] \quad (4)$$

其中, $Q_i, K_i, V_i \in R^{N \times d_k}$ 分别是第 i 个头的查询、键和值的表示, $W_i^Q, W_i^K, W_i^V \in R^{F \times d_k}$ 是第 i 个头的权重参数. 然后计算每个社交网络快照的上下文表示, 计算过程如公式 (5) 所示:

$$head_i = \text{Softmax}(Q_i K_i^T / \sqrt{d_k}) V_i \quad (5)$$

其中, $head_i \in R^{N \times d_k}$ 表示第 i 个头的注意力输出. 最后, 将每个头得到的注意力表示进行拼接, 从而将每个头学习到的不同时间特征进行整合, 并对多头合并后的表示进行线性变换得到多头注意力机制的最终输出, 计算方式如下:

$$\mathcal{X}_{\text{tatt}}^1 = \text{concat}(head_1, \dots, head_h) W^O \quad (6)$$

其中, $\mathcal{X}_{\text{tatt}}^1 \in R^{N \times l \times h \times d_k}$ 为时间注意力网络对短期分量社交网络序列学习到的时间特征, $W^O \in R^{h \times d_k \times h \times d_k}$ 是可学习的权重参数, h 是注意力的头数.

2.5 基于掩码注意力机制的图卷积网络

在空间维度方面, 一种有效且常用的信息表示方法是对社交网络快照序列中的每个快照提取空间结构, 但这也存在一定的局限性, 这种方法无法从长期的角度考虑节点的邻居对节点的影响, 只能从短期的角度简单地认为所有邻居的重要程度是一样的. 为了解决这个问题, 本文使用注意力神经网络从社交网络序列的时间特征中自适应地学习社交网络图结构, 然后基于社交网络快照序列内节点的所有历史邻居信息对图结构进行掩码更新图结构, 避免引入节点的非邻居信息, 并利用学习到的图结构对时间特征进行图卷积, 从而以数据驱动的方式学习社交网络快照序列的空间信息.

首先, 取 $\mathcal{X}_{\text{tatt}}^1$ 中最后一个时间步的特征作为全局时间特征, 记为 $\tilde{\mathcal{X}}_{\text{tatt}}^1 \in R^{N \times h \times d_k}$, 然后使用注意力机制^[31]自适应学习空间维度中节点之间的相互影响关系, 具体计算方式如下:

$$S = V_s \cdot \sigma((\tilde{\mathcal{X}}_{\text{tatt}}^1 W_1) W_2 (\tilde{\mathcal{X}}_{\text{tatt}}^1 W_3)^T + b_s) \quad (7)$$

其中, $W_1 \in R^{h \times d_k \times D_1}$, $W_2 \in R^{D_1 \times D_2}$, $W_3 \in R^{h \times d_k \times D_2}$, $V_s, b_s \in R^{N \times N}$ 是可学习的参数, σ 是 Sigmoid 激活函数. $S \in R^{N \times N}$ 是学习到的注意力矩阵, S 中的元素 $S_{i,j}$ 表示节点 j 对节点 i 的影响程度.

然后, 使用 Softmax 函数对注意力矩阵 S 进行归一化得到注意力矩阵 S' , 计算如公式 (8):

$$S'_{i,j} = \exp(S_{i,j}) \left/ \sum_{r=1}^N \exp(S_{i,r}) \right. \quad (8)$$

接着, 为了除去非邻居节点的影响, 避免引入冗余的信息, 本文基于历史社交网络快照序列内的历史邻居信息对注意力矩阵 S' 进行掩码, 使其只能关注节点的历史邻居. 社交网络快照序列内的历史邻居矩阵计算如下式:

$$M = \sum_{t=1}^l A_t^1 \quad (9)$$

其中, $M \in R^{N \times N}$. 然后使用 M 对注意力矩阵 S' 进行操作如下式:

$$\text{mask}(S') = \begin{cases} S'_{i,j}, & M_{i,j} > 0 \\ 0, & M_{i,j} = 0 \end{cases} \quad (10)$$

其中, $\text{mask}(S')$ 是掩码注意力矩阵, 简单表示为 $\tilde{S}'$.

最后, 由于空间掩码注意力矩阵 $\tilde{S}'$ 能够动态地学习到节点间的相互影响关系, 本文将其与 GCN 结合, 使只能学习静态拓扑结构的 GCN 能够自适应地获得动态的空间结构信息, 从而使模型更适应于真实的社交网络链路预测. 利用掩码注意力矩阵 $\tilde{S}'$ 和时间特征 $\tilde{\mathcal{X}}_{\text{statt}}^1$ 进行图卷积的操作如下所示:

$$\mathcal{X}_{\text{statt}}^1 = \text{ReLU}(\hat{D}^{-1/2} \tilde{S}' \hat{D}^{-1/2} \tilde{\mathcal{X}}_{\text{statt}}^1 W) \quad (11)$$

其中, $\hat{S}' = \tilde{S}' + I$ 是 $\hat{S}'$ 的度矩阵, $W \in R^{h \times d \times D}$ 是可学习的权重参数. $\mathcal{X}_{\text{statt}}^1 \in R^{N \times D}$ 表示学习到的短期分量社交网络序列的时空特征, 类似地, 可以得到中期、长期分量社交网络序列的时空特征 $\mathcal{X}_{\text{statt}}^2 \in R^{N \times D}$ 和 $\mathcal{X}_{\text{statt}}^3 \in R^{N \times D}$.

2.6 融合注意力网络

在得到不同粒度社交网络快照序列的时空特征后, 需要将这些特征进行融合, 并进行链路预测. 为此, 本文提出了融合注意力网络来融合短期、中期和长期分量的时空特征. 具体地, 将短期分量的时空特征作为短期、中期和长期分量注意力神经网络中的查询, 从而从短期、中期和长期分量中提取出有用的短期、中期和长期信息, 然后将短期、中期和长期信息进行相加得到综合的时空特征, 最后将综合时空特征输入到多层全连接神经网络中完成社交网络链路预测任务.

首先, 对粒度 1 的社交网络时空特征 $\mathcal{X}_{\text{statt}}^1$ 进行线性变化, 得到 Q^1, K^1, V^1 , 计算如下所示:

$$[Q^1, K^1, V^1] = \mathcal{X}_{\text{statt}}^1 [W^{Q^1}, W^{K^1}, W^{V^1}] \quad (12)$$

其中, $Q^1, K^1, V^1 \in R^{N \times d}$ 分别是粒度 1 注意力神经网络的查询、键和值, $W^{Q^1}, W^{K^1}, W^{V^1} \in R^{D \times d}$ 是可学习的权重参数. 对粒度 2 的社交网络时空特征 $\mathcal{X}_{\text{statt}}^2$, 本文以粒度 1 的时空特征 $\mathcal{X}_{\text{statt}}^1$ 作为粒度 2 注意力神经网络的查询, 计算 Q^2, K^2, V^2 如下所示:

$$\begin{cases} Q^2 = Q^1 \\ [K^2, V^2] = \mathcal{X}_{\text{statt}}^2 [W^{K^2}, W^{V^2}] \end{cases} \quad (13)$$

其中, $Q^2, K^2, V^2 \in R^{N \times d}$ 分别是粒度 2 注意力神经网络的查询、键和值, $W^{K^2}, W^{V^2} \in R^{D \times d}$ 是可学习的权重参数. 与粒度 2 注意力神经网络类似, 粒度 3 注意力神经网络的查询 Q^3 、键 K^3 和值 V^3 计算如下式:

$$\begin{cases} Q^3 = Q^1 \\ [K^3, V^3] = \mathcal{X}_{\text{statt}}^3 [W^{K^3}, W^{V^3}] \end{cases} \quad (14)$$

其中, $Q^3, K^3, V^3 \in R^{N \times d}$, $W^{K^3}, W^{V^3} \in R^{D \times d}$ 是可学习的权重参数.

接着, 基于 Q^i, K^i, V^i , $i = 1, 2, 3$ 分别提取短期、中期和长期信息, 计算如下式:

$$\text{Attention}(Q^i, K^i, V^i) = \text{Softmax}(Q^i (K^i)^T / \sqrt{d}) V^i, \quad i = 1, 2, 3 \quad (15)$$

其中, $\text{Attention}(Q^i, K^i, V^i) \in R^{N \times d}$ 是第 i 个注意力神经网络的输出, 为方便起见, 简单表示为 Attention_i . Attention_1 、 Attention_2 和 Attention_3 分别表示注意力神经网络提取到社交网络序列的短期、中期和长期信息. 然后将它们进行相加得到综合的社交网络时空特征, 如下所示:

$$\mathcal{X}_{\text{statt}} = \sum_{i=1}^3 \text{Attention}_i \quad (16)$$

其中, $\mathcal{X}_{\text{statt}} \in \mathbb{R}^{N \times d}$.

最后, 通过多层全连接层将综合时空特征 $\mathcal{X}_{\text{statt}}$ 转化为节点之间的链接关系, 具体如下所示:

$$\widehat{A}_{l+1} = \text{MLP}(\mathcal{X}_{\text{statt}}) \quad (17)$$

其中, $\widehat{A}_{l+1} \in \mathbb{R}^{N \times N}$ 表示第 $l+1$ 个时间步社交网络快照 $\widehat{G}_{l+1}$ 对应的邻接矩阵预测结果. MLP 表示多层全连接层, 除最后一层使用 Sigmoid 激活函数外, 其余层均采用 ReLU 激活函数.

2.7 模型的训练优化

现实世界中, 社交网络是稀疏的, 社交网络中节点之间的链接相对于可能的全部链接而言是相对较少的, 这导致网络的链接密度较低. 因此, 设计损失函数时, 需要考虑社交网络的稀疏性. 为了解决社交网络稀疏性的问题, 在反向传播中给予存在的链接更多地关注, 本文使用带稀疏惩罚的损失函数 $\mathcal{L}$ [32], 具体如下所示:

$$\mathcal{L} = \left\| P * (\widehat{A}_{l+1} - A_{l+1}) \right\|_F^2 \quad (18)$$

其中, A_{l+1} 是第 $l+1$ 个时间步真实社交网络快照 G_{l+1} 的邻接矩阵, $*$ 表示逐元素乘法, P 是惩罚矩阵, 具体表示如下式:

$$P = \begin{cases} \beta, & (A_{l+1})_{i,j} > 0 \\ 1, & (A_{l+1})_{i,j} = 0 \end{cases} \quad (19)$$

其中, $\beta > 1$ 为惩罚系数, 对非零元素施加了更多的惩罚来解决社交网络稀疏性问题.

本文提出的 SAMG-LP 算法的训练过程的伪代码如算法 1 所示.

算法 1. 基于注意力的多粒度链路预测 SAMG-LP 算法.

输入: 社交网络快照序列 $G^1 = \{G_1^1, G_2^1, \dots, G_l^1\}$ 和未来社交网络快照 G_{l+1} ;

输出: 优化后的算法及其参数.

步骤 1. 构建多粒度特征: 对粒度 1 快照序列 G^1 使用 GWMA 生成粒度 2 快照序列 G^2 , 类似地, 利用粒度 2 快照序列 G^2 生成粒度 3 快照序列 G^3 ;

步骤 2. 建模时间特征: 对 3 种不同粒度的社交网络快照序列分别使用时间注意力网络提取时间信息, 由公式 (6) 得到时间特征;

步骤 3. 建模空间特征: 由公式 (7)–公式 (10) 自适应构建社交网络的空间拓扑结构, 并由公式 (11) 得到 3 种不同粒度的社交网络快照序列的时空特征;

步骤 4. 由公式 (15) 从 3 种不同粒度的时空特征分别提取短期、中期和长期信息, 由公式 (16) 完成特征融合, 并由公式 (17) 将综合时空特征转化为社交网络链路预测结果;

步骤 5. 由公式 (18) 得到损失函数, 并通过梯度下降更新算法的参数.

本文提出的 SAMG-LP 算法时间复杂度主要包括对步骤 2、步骤 3 和步骤 4 共 3 部分的计算. 在步骤 2 中主要计算量是公式 (6), 其相应的计算复杂度是 $O(Nlh^2d_k^2)$, 其中 N 是社交网络节点数, l 是社交网络快照序列的长度, h 是注意力的头数, d_k 是每个注意力头的输出维度. 步骤 3 中主要计算量是公式 (7), 计算复杂度是 $O(N^2D_2)$, 其中 D_2 是权重参数 W_2 的输出维度. 步骤 4 的主要计算量在公式 (15), 计算复杂度 $O(N^2d)$, 其中 d 是注意力神经网络的输出维度. 由于真实社交网络中的节点数往往是巨大的, 因此, 算法时间复杂度可简约地表示为 $O(N^2)$.

3 实验分析

本节主要在真实数据集上验证本文提出的 SAMG-LP 算法的链路预测效果. 实验主要分成 4 部分: 第 1 部分是在 4 个真实社交网络数据集上, 通过与基准模型比较, 验证 SAMG-LP 算法在链接预测任务上的优越性; 第 2 部分是通过消融实验, 分析 SAMG-LP 各个组件对算法性能的影响; 第 3 部分是分析不同参数设置对 SAMG-LP 算

法预测性能的影响;第 4 部分是分析 SAMG-LP 算法训练时间消耗.

3.1 实验数据集

本文在公开数据集 Huggle、LH10、LyonSchool 和 workplace^[33,34]上进行实验.将数据集按照 7:1:2 的比例划分成训练集、验证集和测试集.这些数据集都是社交网络数据集,其统计信息如表 1 所示.这些数据集简要描述如下.

表 1 数据集的统计信息

数据集	节点数	链接总数	快照数量
Huggle	274	28 244	70
LH10	687	150 126	71
LyonSchool	497	1 048 576	473
workplace	973	9 827	89

- (1) Huggle 通过无线设备记录人与人的联系,节点表示人,节点间的链接表示他们之间有过接触.
- (2) LH10、LyonSchool 都是使用 RFID 技术来识别人与人之间的接触,其中 LH10 数据收集地点是医院, LyonSchool 数据收集地点是小学.
- (3) workplace 表示工作场所中人与人之间交流的动态网络.

3.2 评价指标及基准模型

本文使用 3 种常用的性能指标来比较不同的链路预测算法,分别是 AUC、MR^[35]和 Recall. AUC 反映了算法的预测准确率, Recall 反映了算法对正样本的识别能力,二者均为指标值越大效果越好.而 MR 反映了算法预测错误的比例,指标值越小效果越好.

为了评估所提出算法的预测性能,本文将 SAMG-LP 算法与 7 种具有代表性的链路预测算法进行比较,分为基于相似性指数的方法、基于矩阵分解的方法和基于深度学习的方法.基于相似性指数的方法分为两类:第 1 类是直接构造相似性指标的 PR-CN 算法^[7];第 2 类是以相似性指数作为特征,通过分类器进行链路预测的 NC-LGBM 算法^[8].基于矩阵分解的方法包括:基于静态图链路预测的 LPANMF^[15]非负矩阵分解算法和基于动态图链路预测的 MijFE^[11]非负矩阵分解算法.基于深度学习的链路预测方法分为两类:第 1 类是提取时空特征进行链路预测的 GC-LSTM^[21]和 GCN_MA^[28];第 2 类是从图像生成角度完成链路预测的 SRG 算法^[27].

所有实验均在 Ubuntu 服务器上进行 (CPU: Intel(R) Xeon(R) Gold 5118 CPU @ 2.30 GHz, GPU: NVIDIA GeForce RTX 2080Ti).实验均采用 10 个历史时间步窗口作为社交网络快照序列的长度,取 5 次随机实验的平均值作为实验结果.对于静态图链路预测算法,使用最后一个网络快照作为算法的输入.所有基准算法都遵循他们论文中的最佳参数.而对于 SAMG-LP 算法而言,图移动平均的项数设置为 4,时间衰减函数的参数设置为 0.1,时间注意力网络和基于掩码注意力机制的图卷积网络中的所有隐藏层维度均设置为 32,融合注意力网络除 MLP 最后一层维度设置为节点数外,其余隐藏层维度均设置为 64,损失函数惩罚系数设置为 2, Adam 优化器的学习率为 0.0005,训练迭代次数设置为 300,实验基于 PyTorch 2.0.0.

3.3 对比实验结果分析

本节分别在 4 个社交网络数据集上对比了本文提出的 SAMG-LP 和基准算法的 AUC、MR 和 Recall.实验结果如表 2 所示,其中加粗部分表示最优结果.由表 2 可知,相比于 7 种基准算法,本文算法几乎在所有指标上都取得了最优结果.其中 SAMG-LP 在 LH10 和 workplace 数据集上的预测性能有较为明显的提升.对于 LH10 数据集, SAMG-LP 在 AUC、MR 和 Recall 指标上分别取得了 3.4%–16.9%、13.3%–72.4% 和 2.5%–30.0% 的性能提升;对于 workplace 数据集, SAMG-LP 在 AUC、MR 和 Recall 指标上分别取得了 3.7%–55.9%、16.9%–94.6% 和 4.6%–90.3% 的性能提升.

- 1) 与 2 种基于相似性指数的方法相比, SAMG-LP 在 AUC、MR 和 Recall 指标上至少取得了 4.5%、14.7% 和 2.5% 的性能提升.这是由于社交网络往往具有复杂的时空特征,不能简单地通过一种或几种相似性指数进行描述.这使得混合型相似性指数 PR-CN 在 4 种数据集上的表现较差.虽然基于机器学习分类器的相似性指数算法

NC-LGBM 比 PR-CN 预测性能有所提升,但由于输入的多种相似性指数特征难以表示社交网络的复杂特征而使得其预测能力受限.与这类方法不同,本文没有构建具体的网络相似性指数,而是通过建立深度学习模型来建模社交网络的时空特征,从社交网络数据中自适应地提取出深层的抽象特征.

表 2 SAMG-LP 与基准算法性能比较

算法	Haggle			LH10			LyonSchool			workplace		
	AUC	MR	Recall	AUC	MR	Recall	AUC	MR	Recall	AUC	MR	Recall
PR-CN	0.864	0.107	0.791	0.806	0.091	0.691	0.714	0.101	0.494	0.814	0.059	0.680
NC-LGBM	0.851	0.124	0.814	0.810	0.261	0.876	0.884	0.095	0.807	0.655	0.545	0.853
MljFE	0.804	0.283	0.734	0.896	0.112	0.821	0.923	0.083	0.894	0.619	0.903	0.897
LPANMF	0.867	0.105	0.824	0.853	0.092	0.770	0.878	0.136	0.838	0.702	0.091	0.495
GC-LSTM	0.902	0.043	0.826	0.896	0.107	0.824	0.862	0.099	0.771	0.827	0.062	0.701
SRG	0.887	0.062	0.802	0.831	0.083	0.719	0.893	0.315	0.867	0.931	0.078	0.901
GCN_MA	0.887	0.072	0.824	0.911	0.091	0.854	0.889	0.167	0.933	0.908	0.088	0.872
SAMG-LP	0.915	0.039	0.853	0.942	0.072	0.898	0.924	0.081	0.900	0.965	0.049	0.942

2) 与 2 种基于矩阵分解的方法相比, SAMG-LP 在 AUC、MR 和 Recall 指标上比 MljFE 平均取得了 18.7%、54.7% 和 7.8% 的性能提升; 比 LPANMF 平均取得了 14.7%、42.8% 和 29.5% 的性能提升. MljFE 虽然考虑了网络的动态性和拓扑结构, 并通过非负矩阵分解的方式建立目标函数, 然后使用优化后的特征矩阵和参数矩阵完成链路预测, 但这种简单的矩阵分解难以学习复杂的非线性关系, 它的特征表示能力有限, 导致预测性能一般. LPANMF 在邻接矩阵中加入了攻击者矩阵, 并通过对抗训练学习节点的低维表示, 这一定程度提高了模型的特征提取能力, 但该算法只考虑了网络的稀疏性和空间结构, 忽略了网络的时间动态性, 使得学习到的节点表示不够全面, 链路预测能力有限. 而本文的算法首先利用多头注意力神经网络从时间维度提取时间特征, 然后利用基于数据驱动学习到的拓扑结构从空间维度提取空间特征, 并考虑网络的稀疏性以提高节点的表示能力, 在理论层面上更为合理, 实验结果也证明了本文算法相比于基于矩阵分解的方法预测性能更好.

3) 与 3 种基于深度学习的方法相比, SAMG-LP 除了在 LyonSchool 数据集上, 与 GCN_MA 模型相比在 Recall 指标上取得了次优的结果外, 其余情况都取得了最优的结果. 这 3 种基于深度学习的方法都考虑了社交网络的时间信息和空间结构信息, 预测性能较好, 这也证明了基于深度学习的方法在链路预测问题上的优势, 但它们也存在一些局限性. GC-LSTM 对 LSTM 模型进行改进, 将 GCN 嵌入到 LSTM 的内部结构中来同时提取网络快照序列的时空特征, 但其预测性能不如 SAMG-LP, 这是因为相比于本文算法, GC-LSTM 只能学习到短期的空间特征和局部的时间特征. SRG 深度生成模型设计了基于条件归一化流的 DGFlow 模块, 实现 LR 图与 HR 图的转换, 利用 LR 图对未来 HR 图进行预测, 进而通过 DGFlow 将未来 HR 图生成未来网络. SRG 简单地考虑网络的拓扑结构, 忽略了节点间的影响程度不同, 且在训练过程没有考虑到复杂网络的稀疏特性. GCN_MA 利用 LSTM 更新 GCN 参数来同时建模局部时间特征和空间特征, 然后基于多头注意力机制从全局角度建模时间特征. 虽然 GCN_MA 在 Recall 指标上具有良好的竞争力, 但其整体表现不如 SAMG-LP, 这是由于 GCN_MA 只考虑了单一粒度的社交网络时间序列, 无法学习到它的中、长期趋势, 且在考虑空间特征时只从预先定义的社交网络结构出发, 无法自适应地从数据中学习社交网络的拓扑结构, 进而更有效地学习到社交网络的空间结构信息. 与 GCN_MA 相比, SAMG-LP 可以从多个不同粒度的社交网络时间序列中学习信息更丰富的时空特征, 并从稀疏惩罚损失函数考虑到了社交网络的稀疏性, 这使得 SAMG-LP 的综合预测性能更好.

3.4 消融实验

为了进一步验证 SAMG-LP 算法中各组件的有效性, 本文在 4 种真实的社交网络数据集上进行消融实验并进行分析. 本文设计了不同网络结构和损失函数的 SAMG-LP 算法变体, 具体地, model0 表示仅使用原始社交网络快照序列的单粒度模型, 探究中、长期粒度社交网络快照序列对 SAMG-LP 算法预测性能的影响; model1 表示仅使用单头注意力机制提取时间特征的 SAMG-LP 算法, 探究社交网络多元演化对预测性能的影响程度; model2 表示仅使用历史邻居矩阵进行图卷积的 SAMG-LP 算法, 探究自适应提取拓扑结构的组件对 SAMG-LP 算法的预测性

能是否有促进作用; model3 表示去除基于掩码注意力机制的图卷积网络组件的 SAMG-LP 算法, 探究空间特征对链路预测性能的重要程度; model4 表示将融合注意力网络换为直接对不同粒度时空特征相加融合的 SAMG-LP 算法, 探究融合不同粒度的注意力网络组件的有效性; model5 表示将稀疏惩罚损失函数替换为 MSE 损失函数的 SAMG-LP 算法, 探究网络稀疏特性对 SAMG-LP 算法预测性能的影响. 消融实验的结果如表 3 所示, 其中加粗部分表示最佳结果. 可以观察到, 所有变体模型的综合表现都不如 SAMG-LP 算法, 实验表明 SAMG-LP 算法各个组件的有效性.

表 3 消融实验结果

算法	Haggle			LH10			LyonSchool			workplace		
	AUC	MR	Recall	AUC	MR	Recall	AUC	MR	Recall	AUC	MR	Recall
model0	0.907	0.041	0.838	0.921	0.101	0.887	0.818	0.084	0.696	0.962	0.052	0.934
model1	0.870	0.072	0.772	0.918	0.081	0.862	0.874	0.091	0.802	0.880	0.081	0.814
model2	0.865	0.078	0.764	0.922	0.087	0.869	0.915	0.089	0.889	0.914	0.096	0.885
model3	0.894	0.068	0.809	0.584	0.403	0.504	0.846	0.509	0.891	0.635	0.443	0.708
model4	0.892	0.047	0.809	0.936	0.086	0.890	0.819	0.082	0.684	0.954	0.058	0.922
model5	0.711	0.016	0.433	0.897	0.101	0.837	0.716	0.071	0.475	0.585	0.770	0.900
SAMG-LP	0.915	0.039	0.853	0.942	0.072	0.898	0.924	0.081	0.900	0.965	0.049	0.942

在 model0 中, 该变体相比于 SAMG-LP 算法在 AUC、MR 和 Recall 指标上的预测性能平均下降了 3.7%、13.8% 和 6.6%, 这表明了多粒度社交网络快照序列对 SAMG-LP 的促进作用, 原始的社交网络快照序列仅提供短期趋势的信息, 而 GWMA 生成的快照序列能够有效地补充中、长期变化趋势的信息, 丰富了节点的嵌入表示; 在 model1 中, 变体模型的表现都不如 SAMG-LP 算法, 这表明社交网络多元演化在链路预测中的重要性; 在 model2 中, 仅使用历史邻居矩阵提取空间特征无法区别邻居节点的重要性, 使得预测性能不如 SAMG-LP 算法; 在 model3 中, 其相比于 SAMG-LP 算法在 AUC、MR 和 Recall 指标上的预测性能至少下降了 2.3%、74.4% 和 1.0%, 特别是在 LH10 和 workplace 数据集上, model3 的预测表现相比于 SAMG-LP 算法急剧下滑, 这突出了社交网络空间结构特征对链路预测的重要性和必要性; 在 model4 中, 相比于 SAMG-LP, 其在各指标上的表现都有所下降了, 这证明了本文提出的融合注意力网络能够提升 SAMG-LP 算法预测效果; 在 model5 中, 虽然 SAMG-LP 算法相比于 model5 在 Haggle 和 LyonSchool 数据集的 MR 指标上取得了次优结果, 但其余情况下 SAMG-LP 算法预测性能都比 model5 更好. 这是因为 MSE 损失函数不区分真实标签中的零与非零元素, 社交网络的稀疏性使得真实标签中的零元素远多于非零元素, MSE 损失函数使得模型更倾向于将链接存在与否预测为 0, 进而导致模型的 MR 指标表现较好, 但这也导致模型在 AUC 和 Recall 指标上表现较差. 当模型倾向于将预测结果预测为 0, 其应用价值会受到影响, 特别是在需要准确识别链接存在的情况下. 因此, 实验有效证明了稀疏损失函数能够考虑到了社交网络的稀疏特性, 使得模型的预测结果更加可靠.

3.5 参数影响分析

为了探究超参数对 SAMG-LP 算法预测性能的影响, 本节在 4 个真实社交网络数据集上分别对注意力头数 h 、时空特征嵌入维度 D 、损失函数的惩罚系数 β 和时间衰减函数参数 θ 共 4 种模型核心超参数进行实验. 为分析不同超参数的影响, h 的取值范围为 {1, 2, 4, 8, 16}, D 的取值范围为 {16, 32, 64, 128}, β 的取值范围为 {1, 1.5, 2, 2.5, 3}, θ 的取值范围为 {0, 0.1, 0.2, 0.4, 0.8}. 当研究其中一个超参数时, 其余超参数的设定与第 3.2 节实验的设置保持一致. 超参数的实验结果如图 5 所示.

从图 5(a) 可以看到, 在 Haggle 数据集上, 随着注意力头数的增加, 模型在 AUC 和 Recall 指标上的预测效果先增加后减少, 然后趋于稳定; 而在 MR 指标上呈现波动状态. 特别地, 当 $h = 2$ 时, AUC 和 Recall 达到最大值, 而 MR 达到最小值. 因此, 在 Haggle 数据集上, 设定注意力头数为 2, 此时 SAMG-LP 算法取得最优预测结果. 从图 5(b) 和图 5(d) 中可以看到, 在 LH10 和 workplace 数据集上, 当注意力头数设置为 2 时, SAMG-LP 算法在 AUC、MR 和 Recall 指标上都取得较好的表现. 从图 5(c) 中观察到, 在 LyonSchool 数据集中, 随着 h 的增加, SAMG-LP 算法

在 AUC 和 Recall 指标上总体呈现增长趋势, 在 MR 指标上大致呈现下降趋势, 当 $h = 16$ 时, SAMG-LP 算法在 3 个指标上取得最优效果。

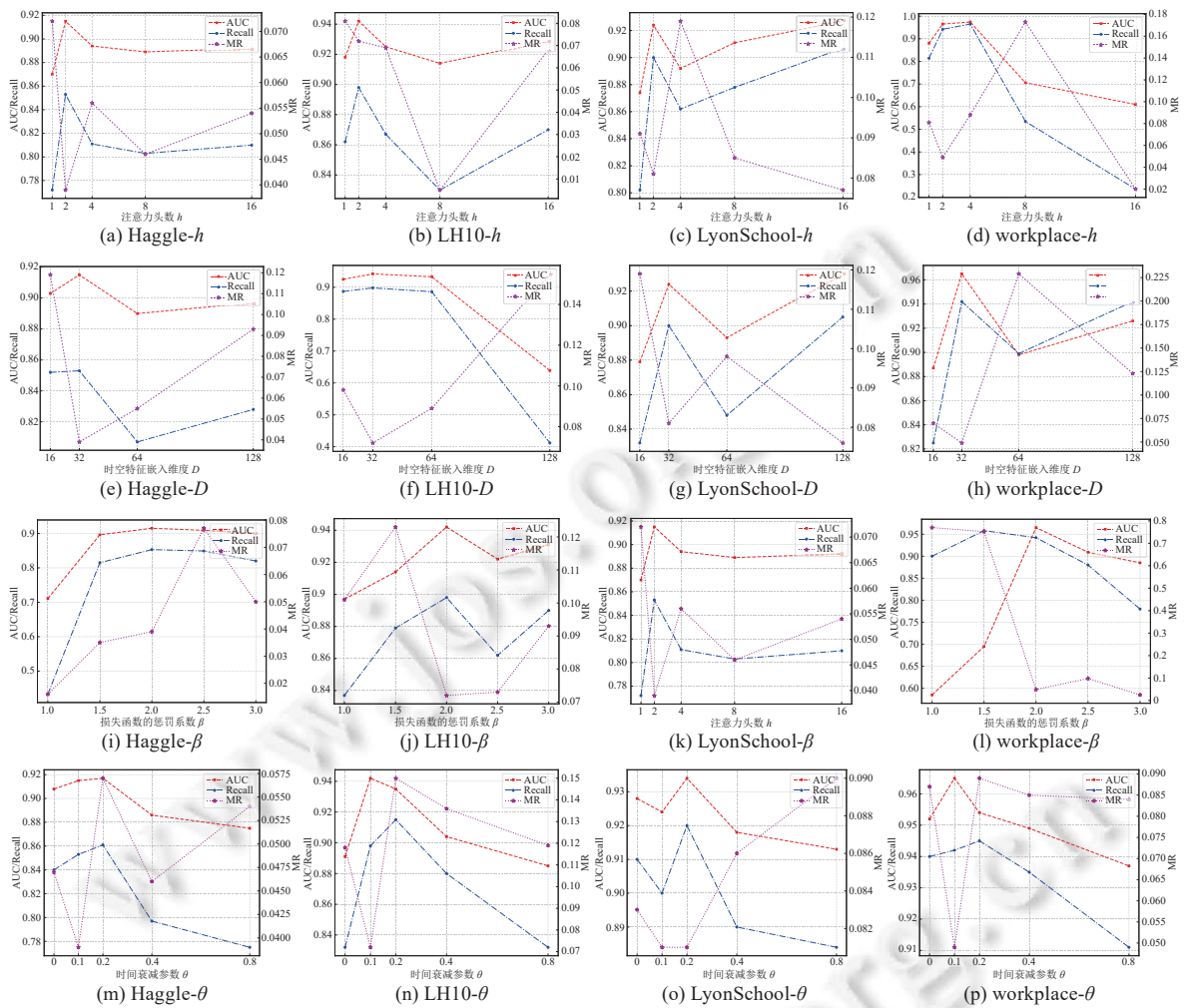


图5 参数对 SAMG-LP 算法性能的影响

由图 5(e)–图 5(h) 分析在不同数据集中, 时空特征嵌入维度对模型预测性能的影响, 可以得出, 在 Haggle、LH10 和 workplace 数据集中, 时空特征嵌入维度 D 取值为 32 时, SAMG-LP 算法的预测性能最佳; 而在 LyonSchool 数据集中, D 取值 128 时, SAMG-LP 在 3 种评估指标中均取得最优表现。类似地, 从图 5(i)–图 5(l) 中分析在不同数据集中, 损失函数的惩罚系数 β 对 SAMG-LP 算法的影响, 在 Haggle 和 LH10 数据集中, 选定 $\beta = 2$ 时, SAMG-LP 算法的 AUC 和 Recall 取得最大值, 且 MR 指标表现较好。在 LyonSchool 数据集上, 随着惩罚系数的增大, AUC 和 Recall 也在增大, 当 $\beta = 3$ 时, AUC 和 Recall 均取得最大值, 此时 MR 取得次小值。在 workplace 数据集上, 当惩罚系数设置为 2 时, 模型的 AUC 取得最大值, 而 Recall 和 MR 取得次优的效果。

图 5(m)–图 5(p) 展示了时间衰减参数 θ 在不同数据集上对模型预测性能的影响。可以看出, 在 4 个真实数据集中, 随着时间衰减参数 θ 的增大, SAMG-LP 算法的 AUC 和 Recall 指标先上升后下降, 而 MR 指标则大体表现为先下降后上升。具体而言, 当时间衰减参数 $\theta = 0$ 时, 衰减因子 $d_t = 1$, 即: 所有社交网络快照的权重相同, 这导致对未来值影响较小的历史数据也被同等对待, 造成冗余信息的引入, 所以预测性能较差。随着时间衰减参数 θ 的增加, 衰减因子 d_t 逐渐减小, 即: 历史社交网络快照的权重不断降低。当时间衰减参数 θ 取值在 0.1–0.2 时, 模型预测

性能达到较好的效果; 当时间衰减参数 θ 超过 0.2 以后, 预测性能逐步下降. 这是因为适度地减小历史网络快照的权重有助于提高特征的区分度, 从而提升预测性能; 而时间衰减参数 θ 过大时, 历史网络快照信息会被过度丢弃, 导致模型预测性能下降. 实验结果表明, 时间衰减参数 θ 取值在 0.1–0.2 之间较为合适.

3.6 时间分析

为探讨 SAMG-LP 算法与所有基准算法的训练时间消耗, 我们对所有算法在相同训练样本上的训练时间进行了对比, 记录 5 次随机实验结果的平均值如表 4 所示, 其中加粗部分表示一类方法中的最佳结果. 从理论分析, 所有算法的时间复杂度几乎都是 $O(N^2)$, 但它们具体训练时长却有所不同. 这是因为理论时间复杂度忽略了常数因子和低阶项的影响, 而这些因素在实际运行时会对性能产生影响. 从实验结果分析, 深度学习链路预测方法所需的训练成本比矩阵分解方法、相似性指数方法更高. 这是因为在相似性指数方法中, 它们通过梯度提升等算法训练, 甚至不需要训练. 例如, PR-CN 直接根据输入数据预测未来结果; 矩阵分解方法则以梯度下降法来优化矩阵参数, 因其涉及的矩阵参数较少而具有较好的训练速度. 然而, 这两类方法的学习能力有限, 导致其在实际应用中的预测表现不佳.

表 4 SAMG-LP 和基线方法在 4 个数据集上的训练时间成本 (s)

方法类型	算法	Haggle	LH10	LyonSchool	workplace	理论时间复杂度
相似性指数方法	PR-CN	—	—	—	—	$O(N^2)$
	NC-LGBM	44.62	281.10	1096.78	673.88	$O(N^2)$
矩阵分解方法	MljFE	132.76	413.88	1369.31	852.23	$O(N^2)$
	LPANMF	191.99	323.20	1708.04	904.50	$O(N^2)$
深度学习方法	GC-LSTM	246.43	640.56	3640.91	1320.94	$O(N^2)$
	SRG	291.85	862.43	5180.83	2263.77	$O(N^3)$
	GCN_MA	259.68	249.19	2637.50	712.59	$O(N^2)$
	SAMG-LP	235.47	371.39	2362.52	695.61	$O(N^2)$

注: “—”表示方法不适用

在深度学习方法中, SRG 立方级的时间复杂度使其所需训练时长比其他方法更多. 相比之下, GCN_MA 和 GC-LSTM 在训练时间上有所不同. GC-LSTM 将 GCN 嵌入 LSTM 结构中, 由于嵌入后模块的高计算消耗以及需要逐步处理每一个时间步, 导致训练成本较高. GCN_MA 则通过解耦 GCN、LSTM 和注意力神经网络, 分别学习社交网络不同层次的时空特征, 从而在训练时间上具有优势. 相比上述模型, SAMG-LP 摒弃了逐步从每个时间步学习空间特征的思路, 通过结合掩码注意力机制在整个图快照序列中学习网络节点间的连接强度, 进而只需一次图卷积操作即可完成空间特征的提取. 此外, SAMG-LP 模块之间采用解耦的方式, 确保了单个模块计算的低消耗性. 综合而言, SAMG-LP 在速度和预测性能方面能够达到更好的平衡.

4 总 结

针对现有链路预测算法仅考虑单一粒度的社交网络快照序列和无法从长期角度考虑节点邻居不同重要程度等问题, 本文提出一种基于注意力的多粒度链路预测 SAMG-LP 算法. 通过扩展加权移动平均提出图加权移动平均策略, 用于构建反映短、中、长期变化趋势的不同粒度社交网络时间序列, 并基于多头注意力神经网络从全局角度依次建模不同粒度社交网络时间序列的时间信息, 然后引入掩码注意力机制自适应地学习社交网络中节点之间的相互影响关系, 结合图卷积网络构建网络的空间特征, 进一步提出了融合注意力神经网络有效地融合来自不同粒度的时空上下文, 最后使用稀疏惩罚损失实现更准确可靠的链路预测. 在 4 种真实社交网络数据集上的大量对比实验和消融实验验证了 SAMG-LP 算法的优越性和有效性. 在未来工作中, 将进一步探索链路预测中的多分量特征联合建模, 如对网络序列中的高、低频分量联合建模, 或对谱域、空间域分量的联合建模等, 充分利用网络序列中的多种信息进一步提高模型链路预测的能力.

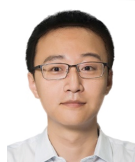
References:

- [1] Zhu HR, Xiong F, Chen HS, Xiong X, Wang L. Incorporating a triple graph neural network with multiple implicit feedback for social recommendation. *ACM Trans. on the Web*, 2024, 18(2): 23. [doi: [10.1145/3580517](https://doi.org/10.1145/3580517)]
- [2] Nunes I, Heddes M, Vergés P, Abraham D, Veidenbaum A, Nicolau A, Givargis T. DotHash: Estimating set similarity metrics for link prediction and document deduplication. In: *Proc. of the 29th ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining*. Long Beach: ACM, 2023. 1758–1769. [doi: [10.1145/3580305.3599314](https://doi.org/10.1145/3580305.3599314)]
- [3] Ma ZM, Nandy S. Community detection with contextual multilayer networks. *IEEE Trans. on Information Theory*, 2023, 69(5): 3203–3239. [doi: [10.1109/TIT.2023.3238352](https://doi.org/10.1109/TIT.2023.3238352)]
- [4] Billah MM, Sultana A, Bhuiyan F, Kaosar MG. Stock price prediction: Comparison of different moving average techniques using deep learning model. *Neural Computing and Applications*, 2024, 36(11): 5861–5871. [doi: [10.1007/s00521-023-09369-0](https://doi.org/10.1007/s00521-023-09369-0)]
- [5] Sharma A, Yadav AK, Rai AK. A novel and precise approach for similarity-based link prediction in diverse networks. *Social Network Analysis and Mining*, 2023, 14(1): 11. [doi: [10.1007/s13278-023-01160-2](https://doi.org/10.1007/s13278-023-01160-2)]
- [6] Jiang SY, Xu XK, Xiao J. Link prediction by combining local structure similarity with node behavior synchronization. *IEEE Trans. on Computational Social Systems*, 2023, 11(3): 3816–3825. [doi: [10.1109/TCSS.2023.3335295](https://doi.org/10.1109/TCSS.2023.3335295)]
- [7] Charikhi M. Association of the PageRank algorithm with similarity-based methods for link prediction in complex networks. *Physica A: Statistical Mechanics and Its Applications*, 2024, 637: 129552. [doi: [10.1016/j.physa.2024.129552](https://doi.org/10.1016/j.physa.2024.129552)]
- [8] Kumar S, Mallik A, Panda BS. Link prediction in complex networks using node centrality and light gradient boosting machine. *World Wide Web*, 2022, 25(6): 2487–2513. [doi: [10.1007/s11280-021-01000-3](https://doi.org/10.1007/s11280-021-01000-3)]
- [9] Chaubey VP, Sharma A, Sharma T, Sharma S, Kumar A. Link prediction for social network analysis using random forest and XG-boost algorithm. In: *Proc. of the 2nd Int'l Conf. on Informatics*. Noida: IEEE, 2023. 1–6. [doi: [10.1109/ICI60088.2023.10421655](https://doi.org/10.1109/ICI60088.2023.10421655)]
- [10] Choudhury N. Community-aware evolution similarity for link prediction in dynamic social networks. *Mathematics*, 2024, 12(2): 285. [doi: [10.3390/math12020285](https://doi.org/10.3390/math12020285)]
- [11] Ma XK, Tan SY, Xie XH, Zhong XX, Deng JJ. Joint multi-label learning and feature extraction for temporal link prediction. *Pattern Recognition*, 2022, 121: 108216. [doi: [10.1016/j.patcog.2021.108216](https://doi.org/10.1016/j.patcog.2021.108216)]
- [12] Lv LS, Bardou D, Liu YQ, Hu P. Deep Autoencoder-like non-negative matrix factorization with graph regularized for link prediction in dynamic networks. *Applied Soft Computing*, 2023, 148: 110832. [doi: [10.1016/j.asoc.2023.110832](https://doi.org/10.1016/j.asoc.2023.110832)]
- [13] Yan LP, Yu WR. Non-negative matrix factorization for link prediction preserving row and column spaces. In: *Proc. of the 2023 IEEE Int'l Conf. on Data Mining*. Shanghai: IEEE, 2023. 1451–1456. [doi: [10.1109/ICDM58522.2023.00190](https://doi.org/10.1109/ICDM58522.2023.00190)]
- [14] Lin XY, Chen XY, Zheng ZW. Deep manifold matrix factorization autoencoder using global connectivity for link prediction. *Applied Intelligence*, 2023, 53(21): 25816–25835. [doi: [10.1007/s10489-023-04887-9](https://doi.org/10.1007/s10489-023-04887-9)]
- [15] Mahmoodi R, Seyedi SA, Tab FA, Abdollahpour A. Link prediction by adversarial nonnegative matrix factorization. *Knowledge-based Systems*, 2023, 280: 110998. [doi: [10.1016/j.knsys.2023.110998](https://doi.org/10.1016/j.knsys.2023.110998)]
- [16] Nasiri E, Berahmand K, Li YF. Robust graph regularization nonnegative matrix factorization for link prediction in attributed networks. *Multimedia Tools and Applications*, 2023, 82(3): 3745–3768. [doi: [10.1007/s11042-022-12943-8](https://doi.org/10.1007/s11042-022-12943-8)]
- [17] Chen GF, Wang HB, Fang YL, Jiang L. Link prediction by deep non-negative matrix factorization. *Expert Systems with Applications*, 2022, 188: 115991. [doi: [10.1016/j.eswa.2021.115991](https://doi.org/10.1016/j.eswa.2021.115991)]
- [18] Yao YB, He YY, Huang ZT, Xu ZP, Yang F, Tang JX, Gao K. Deep non-negative matrix factorization with edge generator for link prediction in complex networks. *Applied Intelligence*, 2024, 54(1): 592–613. [doi: [10.1007/s10489-023-05211-1](https://doi.org/10.1007/s10489-023-05211-1)]
- [19] Min SJ, Gao Z, Peng J, Wang L, Qin K, Fang B. STGSN—A spatial-temporal graph neural network framework for time-evolving social networks. *Knowledge-based Systems*, 2021, 214: 106746. [doi: [10.1016/j.knsys.2021.106746](https://doi.org/10.1016/j.knsys.2021.106746)]
- [20] Jiao PF, Guo X, Jing X, He DX, Wu HM, Pan SR, Gong MG, Wang WJ. Temporal network embedding for link prediction via VAE joint attention mechanism. *IEEE Trans. on Neural Networks and Learning Systems*, 2022, 33(12): 7400–7413. [doi: [10.1109/TNNLS.2021.3084957](https://doi.org/10.1109/TNNLS.2021.3084957)]
- [21] Chen JY, Wang XK, Xu XH. GC-LSTM: Graph convolution embedded LSTM for dynamic network link prediction. *Applied Intelligence*, 2022, 52(7): 7513–7528. [doi: [10.1007/s10489-021-02518-9](https://doi.org/10.1007/s10489-021-02518-9)]
- [22] Tan SY, You JY, Li DY. Temporality- and frequency-aware graph contrastive learning for temporal network. In: *Proc. of the 31st ACM Int'l Conf. on Information & Knowledge Management*. Atlanta: ACM, 2022. 1878–1888. [doi: [10.1145/3511808.3557469](https://doi.org/10.1145/3511808.3557469)]
- [23] Liu LF, Yu ZX, Zhu H. A link prediction method based on gated recurrent units for mobile social network. *Journal of Computer Research and Development*, 2023, 60(3): 705–716 (in Chinese with English abstract). [doi: [10.7544/issn1000-1239.202110432](https://doi.org/10.7544/issn1000-1239.202110432)]
- [24] He P, Wei C, Lü SK, Zeng C, Li B. GoGCN for interaction prediction between classes in software system. *Ruan Jian Xue Bao/Journal of Software*, 2023, 34(11): 5029–5041 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/6678.htm> [doi: [10.13328/j.cnki.jos.006678](https://doi.org/10.13328/j.cnki.jos.006678)]

- [25] Zhao LL, Wu AB, Yuan Y, Li Y, Wang GR. Graph representation learning on location-based social networks. Chinese Journal of Computers, 2022, 45(4): 838–857 (in Chinese with English abstract). [doi: 10.11897/SP.J.1016.2022.00838]
- [26] Qin M, Zhang CR, Bai B, Zhang G, Yeung DY. High-quality temporal link prediction for weighted dynamic graphs via inductive embedding aggregation. IEEE Trans. on Knowledge and Data Engineering, 2023, 35(9): 9378–9393. [doi: 10.1109/TKDE.2023.3238360]
- [27] Yin YT, Wu YJ, Yang XB, Zhang WS, Yuan XJ. Super resolution graph with conditional normalizing flows for temporal link prediction. IEEE Trans. on Knowledge and Data Engineering, 2024, 36(3): 1311–1327. [doi: 10.1109/TKDE.2023.3295367]
- [28] Mei P, Zhao YH. Dynamic network link prediction with node representation learning from graph convolutional networks. Scientific Reports, 2024, 14(1): 538. [doi: 10.1038/s41598-023-50977-6]
- [29] Chen XQ, Wu SB, Shi CJ, Huang YG, Yang YS, Ke RM, Zhao JS. Sensing data supported traffic flow prediction via denoising schemes and ANN: A comparison. IEEE Sensors Journal, 2020, 20(23): 14317–14328. [doi: 10.1109/JSEN.2020.3007809]
- [30] Sankar A, Wu YH, Gou L, Zhang W, Yang H. DySAT: Deep neural representation learning on dynamic graphs via self-attention networks. In: Proc. of the 13th Int'l Conf. on Web Search and Data Mining. Houston: ACM, 2020. 519–527. [doi: 10.1145/3336191.3371845]
- [31] Guo SN, Lin YF, Feng N, Song C, Wan HY. Attention based spatial-temporal graph convolutional networks for traffic flow forecasting. In: Proc. of the 33rd AAAI Conf. on Artificial Intelligence. Honolulu: AAAI, 2019. 922–929. [doi: 10.1609/aaai.v33i01.3301922]
- [32] Chen JY, Zhang J, Xu XH, Fu CB, Zhang D, Zhang QP, Xuan Q. E-LSTM-D: A deep learning framework for dynamic network link prediction. IEEE Trans. on Systems, Man, and Cybernetics: Systems, 2021, 51(6): 3699–3712. [doi: 10.1109/TSMC.2019.2932913]
- [33] Rossi R, Ahmed N. The network data repository with interactive graph analytics and visualization. In: Proc. of the 29th AAAI Conf. on Artificial Intelligence. Austin: AAAI, 2015. 4292–4293. [doi: 10.1609/aaai.v29i1.9277]
- [34] Kunegis J. Konec: The Koblenz network collection. In: Proc. of the 22nd Int'l Conf. World Wide Web. Rio de Janeiro: ACM, 2013. 1343–1350. [doi: 10.1145/2487788.2488173]
- [35] Lei K, Qin M, Bai B, Zhang G, Yang M. GCN-GAN: A non-linear temporal link prediction model for weighted dynamic networks. In: Proc. of the 2019 IEEE Conf. on Computer Communications. Paris: IEEE, 2019. 388–396. [doi: 10.1109/INFOCOM.2019.8737631]

附中文参考文献:

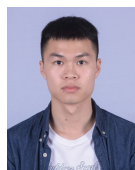
- [23] 刘林峰, 于子兴, 祝贺. 基于门控循环单元的移动社会网络链路预测方法. 计算机研究与发展, 2023, 60(3): 705–716. [doi: 10.7544/issn1000-1239.202110432]
- [24] 何鹏, 卫操, 吕晟凯, 曾诚, 李兵. 基于 GoGCN 的软件系统类交互关系预测. 软件学报, 2023, 34(11): 5029–5041. <http://www.jos.org.cn/1000-9825/6678.htm> [doi: 10.13328/j.cnki.jos.006678]
- [25] 赵琳琳, 吴安彪, 袁野, 李扬, 王国仁. 位置社交网络上的图表示学习. 计算机学报, 2022, 45(4): 838–857. [doi: 10.11897/SP.J.1016.2022.00838]



何玉林(1982—), 男, 博士, 研究员, CCF 专业会员, 主要研究领域为大数据系统计算技术, 多样本统计分析方法, 机器学习算法及应用.



尹剑飞(1974—), 男, 博士, 副教授, 主要研究领域为大数据, 机器学习, 统计和数值优化.



赖俊龙(2000—), 男, 硕士生, 主要研究领域为图数据挖掘, 时空序列分析, 链路预测.



黄哲学(1959—), 男, 博士, 教授, 博士生导师, CCF 专业会员, 主要研究领域为大数据系统计算技术, 机器学习算法及应用.



崔来中(1984—), 男, 博士, 教授, 博士生导师, 主要研究领域为互联网体系结构, 边缘计算, AI 驱动的新型网络优化设计.