

# 基于表示学习的动态符号社会网络链接预测<sup>①</sup>



刘子豪, 王轶彤

(复旦大学 计算机科学技术学院, 上海 200438)

通信作者: 王轶彤, E-mail: [yitongw@fudan.edu.cn](mailto:yitongw@fudan.edu.cn)

**摘要:** 动态符号网络中的链接预测旨在通过已知的网络拓扑结构和属性特征挖掘节点间的潜在关系. 目前主流的链接预测方法大多基于图表示学习设计, 然而这些工作往往无法同时学习网络中蕴含的符号语义和时间信息. 另外, 现有的图神经网络表示学习模型也难以捕获节点间长距离的信息依赖. 针对上述问题, 本文提出了一种基于表示学习的动态符号社会网络链接预测方法 (RLLP), 该方法设计了一种能够在网络中采样长距离语义路径的时序随机游走策略, 并根据平衡理论和时间感知的长短期记忆神经网络 (T-LSTM) 嵌入每条路径中携带的复杂信息. 最后引入了图注意力机制为节点生成了低维稠密的向量表示, 增强了网络链接的预测能力. 在现实世界的 3 个真实数据集上对本文提出的方法进行验证, 实验结果显示, 相较于其他基线方法, RLLP 在  $F1$  分数和准确率两个指标上均取得了更优的性能表现, 在社会网络链接预测的应用中具有很高的可行性和有效性.

**关键词:** 链接预测; 动态符号网络; 表示学习; 随机游走; 图注意力机制

引用格式: 刘子豪, 王轶彤. 基于表示学习的动态符号社会网络链接预测. 计算机系统应用, 2025, 34(9): 11-21. <http://www.c-s-a.org.cn/1003-3254/9927.html>

## Link Prediction Based on Representation Learning in Dynamic Signed Social Network

LIU Zi-Hao, WANG Yi-Tong

(School of Computer Science and Technology, Fudan University, Shanghai 200438, China)

**Abstract:** Link prediction in dynamic signed networks aims to uncover potential relationships between nodes through known network topologies and attribute features. Most of the current mainstream link prediction methods are designed based on graph representation learning, however, they often fail to simultaneously learn the signed semantic information and temporal information contained in the network. In addition, existing graph neural network representation learning models are also difficult to capture the long-distance information dependence between nodes. To address the above problems, this study proposes a link prediction method based on representation learning (RLLP) in the dynamic signed social network. This method designs a temporal random walk strategy capable of sampling long-distance semantic paths in networks and embeds the complex information carried by each path through balance theory and time-aware long short-term memory neural networks (T-LSTM). Finally, the graph attention mechanism is introduced to generate low-dimensional dense vector representations for nodes, enhancing the predictive capability of network links. The proposed method is verified on three real-world datasets, and the experimental results show that, compared with other baselines, RLLP achieves better performance in both  $F1$ -score and accuracy metrics, demonstrating high feasibility and effectiveness in the applications of social network link prediction.

**Key words:** link prediction; dynamic signed network; representation learning; random walk; graph attention mechanism

① 基金项目: 国家重点研发计划 (2020YFC2008400, 2023YFB2704904)

收稿时间: 2025-01-12; 修改时间: 2025-02-12; 采用时间: 2025-02-24; csa 在线出版时间: 2025-07-25

CNKI 网络首发时间: 2025-07-28

随着社交媒体的发展,越来越多的用户开始在互联网中进行互动并分享内容<sup>[1]</sup>,用户间各式各样的交互构成了复杂多样的社会网络,动态符号社会网络(dynamic signed social network)正是复杂网络中最具代表性的网络之一.如图1所示,现实生活中的许多社会网络都可以表示成动态符号网络,例如交易网络<sup>[2]</sup>、社交网络<sup>[3]</sup>和关系网络<sup>[4]</sup>等.在动态符号社会网络中,节点间形成的交互既带有确切的时间信息,也具备了某一类相反的情感语义,例如信任/不信任、点赞/点踩、朋友/敌人等,可以使用符号{+, -}区分这类相反的符号语义.同时对动态符号社会网络中蕴含的复杂信息进行分析可以帮助我们探索社会网络更深层次的演化规律.

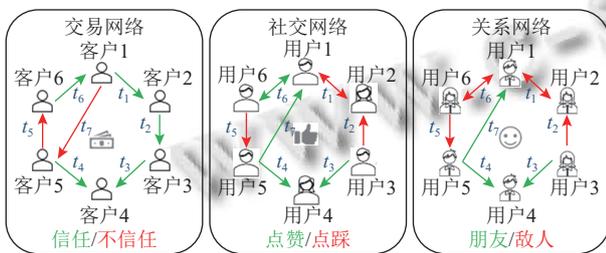


图1 常见的动态符号社会网络

目前,社会网络分析研究中最重要课题之一就是链接预测<sup>[5]</sup>,其目标是根据已知的网络拓扑结构和属性特征挖掘节点间的潜在关系,并计算未来形成新链接的概率,该课题具有广泛的应用场景和重要的现实意义.在社交和电商平台中,链接预测能够基于用户和商品间的潜在关系进行精准的好友推荐<sup>[6]</sup>和商品推荐<sup>[7]</sup>,增加用户对平台的黏性和使用时长.在金融系统中,链接预测可以准确识别网络中的风险交易<sup>[8]</sup>,帮助金融机构及时采取防范措施,降低风险损失.在通信网络中,使用链接预测的方法捕获用户之间的流量增长趋势,可以合理提升网络带宽,提高网络性能<sup>[9]</sup>.在生物网络中,链接预测可以帮助研究者发现蛋白质之间未知的相互作用关系,为生物学研究提供新的线索和方向<sup>[10]</sup>.社会网络中的链接预测涉及计算机科学、社会学、生物学等多个学科领域的知识和技术,其研究和应用过程中促进了跨学科的交流与合作,激发新的研究思路和方法,为解决复杂系统问题提供新的视角和工具.因此,研究复杂社会网络中的链接预测问题至关重要.

早期解决社会网络中的链接预测问题主要使用了相似性度量方法<sup>[11-13]</sup>,这些工作证明了在网络结构或属性特征上相似的两个节点更容易产生新链接.随着

图表示学习(graph representation learning, 也称图嵌入)技术的兴起,基于表示学习的链接预测方法逐渐受到研究者的关注.图表示学习旨在为网络中的节点学习一个低维稠密的向量表示,并且拓扑结构和属性特征相似的节点在向量空间中也是相近的.网络表示学习不仅避免了传统相似性方法的复杂计算,也大幅提升了链接预测任务的性能表现.

然而,现有的基于表示学习的链接预测方法通常只关注网络中静态的符号语义,例如,Wang等人<sup>[14]</sup>首先进行了静态符号网络的表示学习研究,使用随机游走的方法在学习节点嵌入表示的同时捕获了链接的符号语义.Javari等人<sup>[15]</sup>采用网络转换的方法为节点分配了多个网络角色,成功将静态符号图转换为无符号图,在链接预测任务上取得了较好的实验结果.上述方法虽然有效捕获了链接携带的符号信息,但却忽略了节点间交互的动态变化,导致学习得到的节点表示无法体现网络的动态演化过程,最终影响了动态符号社会网络中链接预测任务的准确性.

另外,为了嵌入网络的拓扑结构和特征信息,目前基于表示学习设计的链接预测方法大多使用图神经网络(graph neural network, GNN)<sup>[16]</sup>聚合节点的邻居信息以更新节点的嵌入表示.如图2所示,基于图神经网络设计的表示学习模型能够通过堆叠GNN层数的方式聚合目标节点的局部邻域信息,当GNN层数 $l=1$ 时,模型可以捕获节点 $v$ 一阶邻居( $u_1, u_2, u_3$ )的特征.但随着GNN层数的增加,模型的复杂度也会逐渐上升,并且过多的GNN层数也会引起节点嵌入的过度平滑和过度挤压<sup>[17]</sup>的问题,最终导致节点表示准确性的下降,进而影响链接预测任务的性能表现.上述问题限制了图神经网络捕获节点长距离信息依赖的能力,因此仅使用图神经网络学习节点的嵌入表示往往很难聚合动态符号社会网络中高阶邻居的属性特征.

为了同时学习动态符号社会网络携带的符号语义和时间信息,并解决图神经网络模型无法捕获长距离信息依赖的问题,本文提出了一个基于表示学习的动态符号社会网络链接预测方法(representation learning for link prediction, RLLP),该方法在学习节点向量表示的过程中充分保留了网络中的符号语义和时间信息.RLLP方法主要由3个功能模块构成,即时序随机游走模块、路径信息生成模块和注意力嵌入模块.时序随机游走模块提出了一种能够在动态符号网络中采样长

距离语义路径的时序随机游走策略,以获取目标节点和其高阶邻居间的交互关系;路径信息生成模块基于平衡理论将所得路径中的采样链接划分为平衡/不平衡两类,再分别使用2个多层感知机(MLP)和时间感知的长短期记忆网络(time-aware LSTM, T-LSTM)嵌入两类链接中的符号语义和时间信息;注意力嵌入模块引入图注意力机制区分了不同路径对节点嵌入产生的影响,将所有路径信息融合为最终的节点表示。

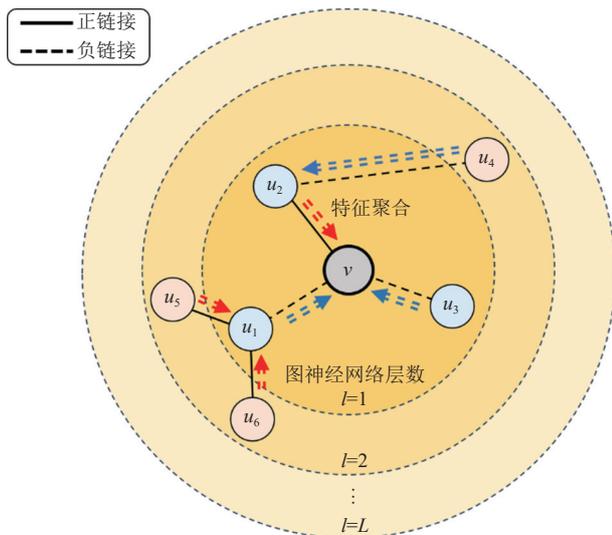


图2 图神经网络表示学习模型结构图

本文的主要贡献有以下几点。

(1) 提出了基于表示学习的动态符号社会网络链接预测方法 RLLP。该方法解决了以往基于表示学习的链接预测方法无法同时捕获网络中符号语义和时间信息的问题。

(2) 设计了3个功能模块用于学习动态符号社会网络中的节点表示。所设计的时序随机游走策略以及融合了平衡理论、T-LSTM和图注意力机制的信息生成和嵌入模块能够在捕获节点长距离信息依赖的同时提高链接预测任务的性能表现。

(3) 在 Epinions、Wikirfa 和 Weibo 这3个公开数据集上进行了实验,并与其他主流的链接预测算法进行了对比,结果表明提出的 RLLP 方法具有很高的可行性和有效性。

## 1 相关工作

### 1.1 基于社会学理论的随机游走方法

随着复杂网络研究的逐渐深入,社会学理论的相关工作成果为符号网络中的链接预测任务提供了新的

指导。目前,符号网络研究中最受人们关注的社会学理论是平衡理论(balance theory)<sup>[18]</sup>。如图3所示,平衡理论将符号网络中任意3个节点间形成的三元关系划分为平衡与不平衡两种类型,其中平衡的三元关系包含偶数条负链接,即满足“朋友的朋友是我的朋友”与“敌人的敌人是我的朋友”这两条假设。平衡理论简明扼要地剖析了符号网络中的链接形成机制,成为研究符号网络链接预测任务的理论基础。



图3 平衡理论示意图

其中,随机游走已经成为建模和研究复杂网络中链接预测问题的重要方法之一。传统的随机游走方法旨在通过探索节点间存在的关联路径帮助我们更好地理解网络的结构特征。然而这些方法大多只能应用于无符号的简单网络中,如何定义符号网络中的随机游走策略,并挖掘符号网络中正负链接形成的拓扑结构信息已成为当下的一个研究热点。

近年来已经有许多基于平衡理论设计的随机游走方法在符号网络链接预测的分析与研究中取得了有效进展。Islam 等人<sup>[19]</sup>设计了一种基于平衡理论计算目标节点与多跳邻居间关系的方法,提出以偶数条负边连接到目标节点的邻居被分配到正关系集合,而以奇数条负边连接到目标节点的邻居被分配到负关系集合,成功地将社会学理论引入符号网络表示学习中,为后续有许多研究提供了指引。Kim 等人<sup>[20]</sup>提出了一种基于截断随机游走的符号网络表示学习方法,通过在似然函数中引入偏差因子来模拟节点间的连通性,并有效分析了网络同质性和平衡理论间的深层关系。Wang 等人<sup>[21]</sup>注意到现实世界的符号网络中平衡三角形出现的频率比不平衡三角形更高,讨论了符号链接和网络结构对随机游走过程的显著影响。Babul 等人<sup>[22]</sup>设计了一个特殊的平衡理论随机游走框架,将随机游走方法的适用范围扩大到更大规模的网络结构之中。

上述基于平衡理论设计的随机游走策略虽然能够有效捕获符号网络的局部和高阶结构特征,但它们并未考虑网络中节点交互携带的时间信息,因此所学习得到的节点嵌入难以准确反映网络的动态演化规律。

### 1.2 基于图神经网络的方法

社会网络中的链路预测任务需要准确判断两个节点间是否可能产生新的链接,该任务依赖于学习图的结构特征,而图神经网络可以有效地从不规则的图数据中挖掘出信息的隐藏模式<sup>[23]</sup>.在平衡理论的指导下,许多研究者开始尝试利用图神经网络捕获社会网络中蕴含的符号语义以提升链接预测任务的性能表现,例如, Derr 等人<sup>[24]</sup>基于平衡理论将图卷积网络 (GCN) 应用于符号图中进行邻域信息的传播和聚合. Huang 等人<sup>[25,26]</sup>在平衡理论的基础上额外考虑了链接方向带来的影响,将图注意力网络 GAT 推广到了符号图上,有效聚合了静态有向符号网络中的符号语义信息.

然而,基于图神经网络设计的表示学习模型往往难以捕获网络中远距离的信息依赖,该问题使得上述方法无法聚合节点的高阶领域信息,进而限制了链接预测任务的性能表现.

### 2 问题定义

在本文中,动态符号社会网络被表示为带有符号语义和时间信息的无向图  $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{T}, \mathcal{S})$ , 其中  $\mathcal{V}$  是节点集,  $\mathcal{E}$  是边集,  $\mathcal{T}$  和  $\mathcal{S}$  是将每条边分别映射到时间戳

和符号信息的函数,节点  $v$  与节点  $u$  在  $t$  时刻形成的符号为  $s$  的链接可以表示为一个四元组  $(v, u, t, s)$ , 且每条边只拥有一种符号属性,即  $s \in \{+1, -1\}$ . 链接预测问题就是根据已知的网络拓扑结构和属性特征预测将来节点对  $(v, u)$  间是否会形成新链接,即判断链接的类型属于正链接/负链接/无链接. 因此,动态符号社会网络中的链接预测是一个三分类问题.

本文提出的 RLLP 方法旨在学习一个映射函数  $f: v \rightarrow R^d, v \in \mathcal{V}$ , 该函数能够将网络中的每个节点映射到低维向量空间  $R^d$ , 其中  $d \ll |\mathcal{V}|$ , 并且映射得到的节点向量表示应具备挖掘动态符号网络中符号语义 (即正链接相连的节点嵌入应该比负链接相连的节点嵌入更相似) 和时序演化模式 (即节点嵌入能够探索未来网络链接形成机制) 的能力.

### 3 RLLP 方法

RLLP 方法的整体框架结构如图 4 所示,主要包括了 3 个关键部分: 时序随机游走模块、路径信息生成模块以及注意力嵌入模块. 该方法的输入是一张完整的动态符号网络,目标是为动态符号网络中的每一个节点输出对应的低维向量表示.

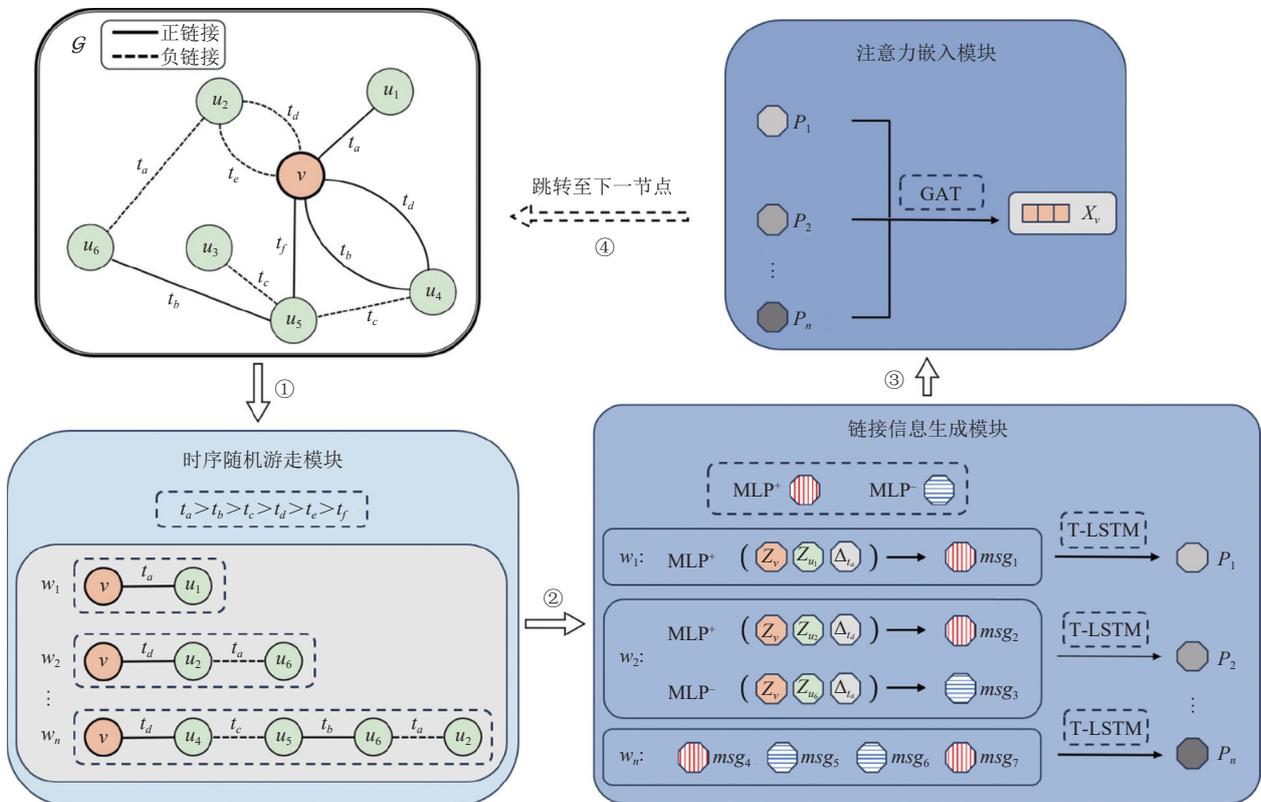


图 4 RLLP 方法整体框架结构图

首先, RLLP 方法会选取网络中任意一个节点作为目标节点 $v$ . 在时序随机游走模块中, RLLP 提出了一种能够在动态符号网络上进行链接采样的时序随机游走策略, 该策略可以根据节点周围的网络结构和链接属性为目标节点 $v$ 生成 $n$ 条带有时间信息和符号语义的时序游走路径 $\{w_1, w_2, \dots, w_n\}$ , 以更好地捕获网络节点间长距离的信息依赖. 接着, RLLP 将上述时序游走路径输入路径信息生成模块, 该模块基于平衡理论设计的多层感知机 (MLP) 以及时序感知的长短期记忆网络 (T-LSTM) 能够为每条游走路径生成一个融合了符号语义和时间信息的向量表示 $\{P_1, P_2, \dots, P_n\}$ . 然后, RLLP 在注意力嵌入模块中引入了图注意力机制为不同游走路径的对应向量表示分配不同权重, 有效聚合了多条路径携带的重要信息, 最终获得目标节点 $v$ 的嵌入表示, 并返回到时序随机游走部分以选取下一个节点作为新的目标节点, 继续进行表示学习的计算过程. 接下来, 将详细介绍 RLLP 方法的各个部分.

### 3.1 时序随机游走

随机游走能够充分捕获节点的局部和高阶领域信息. 在动态符号网络 $\mathcal{G}=(\mathcal{V}, \mathcal{E}, \mathcal{T}, \mathcal{S})$ 中, 受文献[27]启发, 本文定义节点 $v$ 在时刻 $t$ 的时序链接为 $E_t(v)$ , 其含义为节点 $v$ 与任意一阶邻居 $u$ 形成时间 $t' > t$ 的链接集合, 如式 (1) 所示:

$$E_t(v) = \{(v, u, t', s) \mid (v, u, t', s) \in \mathcal{E} \wedge t' > t\} \quad (1)$$

如图 5 所示, 节点 $v$ 在 $t=3$ 的时序链接分别包含了 $(v, u_2, 4, +1)$ 与 $(v, u_3, 5, +1)$ 两条边.

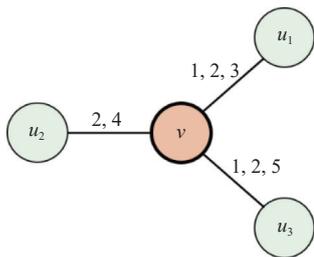


图 5 时序链接示意图

在时序随机游走过程中, 假设 $(u_i, u_j, t, s)$ 为随机游走采样得到的第 $i$ 条链接 $edge_i$ , 则下一条采样链接 $edge_{i+1}$ 服从以下分布:

$$P(edge_{i+1} \mid edge_i) = \begin{cases} \frac{\pi_{\text{ran}}}{M}, & edge_{i+1} \in E_t(u_j) \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

其中,  $\pi_{\text{ran}}$ 是采样链接 $edge_i$ 和 $edge_{i+1}$ 之间的非正则化

转移概率,  $M$ 是正则化常数.

在此基础上, 本文设计了一个动态符号网络中有偏的时序随机游走策略, 该策略倾向于采样交互时间较近的链接, 转移概率 $\pi_{\text{ran}}$ 的计算方式如下:

$$\pi_{\text{ran}} = \frac{1}{\ln(e + \Delta_t)} \quad (3)$$

其中,  $e$ 表示自然常数,  $\Delta_t$ 表示当前链接与上一条采样链接间的时间间隔, 时间间隔越小其对应的转移概率越大.

由于上述随机游走策略有时间顺序的条件限制, 因此本文只规定了时序游走路径的最大长度为 $L=50$ , 以捕获节点的高阶领域信息.

因此, 在动态符号网络中进行时序随机游走采样链接的算法流程如算法 1 所示.

算法 1. 动态符号网络时序随机游走算法

输入: 动态符号网络 $\mathcal{G}=(\mathcal{V}, \mathcal{E}, \mathcal{T}, \mathcal{S})$ , 初始节点 $v$ , 游走路径最大长度 $L$ , 生成路径条数 $n$ .

输出: 节点 $v$ 采样的链接序列 $walks$ .

1. 初始化 $walks$ 为一个空列表 $walks=[]$ ;
2. for  $i=1$  to  $n$  do
3.   初始化序列 $w=[]$ , 时刻 $t=0$ , 节点 $cur=v$ ;
4.   for  $iter=1$  to  $L$  do
5.      $E_t(cur)=GetTemporalEdges(cur, \mathcal{G})$ ;
6.     if  $E_t(cur)$  is NULL do
7.       结束当前循环;
8.     else do
9.        $edge=Sample(E_t(cur), \pi_{\text{ran}})$ ;
10.       将 $edge=(cur, next, t_{\text{new}}, s)$ 添加至 $w$ ;
11.        $cur=next$ ;
12.        $t=t_{\text{new}}$ ;
13.     end if
14.   end for
15.   将 $w$ 添加至 $walks$ ;
16. end for

### 3.2 路径信息生成

以往的工作<sup>[14,27]</sup>在获得随机游走路径后常使用自然语言处理领域的 Skip-gram 模型<sup>[28]</sup>预测中心节点周围的上下文节点, 从而学习目标节点的向量表示. 然而, 这种方法无法嵌入动态符号网络随机游走序列中同时携带的符号语义和时间信息.

为了嵌入动态符号网络中蕴含的复杂数据, 本文在平衡理论的指导下设计了一个路径信息生成模块, 其结构如图 6 所示. 该模块首先会根据相邻链接间的交互时间差计算时间间隔 $\Delta_t$ , 并基于平衡理论为路径

$w_i$ 中的每一条链接构造一个标签  $flag$ . 然后, 根据标签的不同分别使用 2 个多层感知机 (MLP) 以嵌入链接携带的时间信息与符号语义. 最后, 使用时间感知的长短期记忆网络 (T-LSTM) 为随机游走路径生成对应的向量表示  $P_i$ .

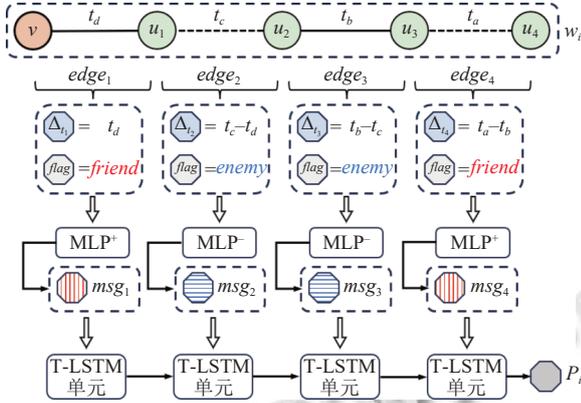


图6 路径信息生成模块结构图

基于平衡理论构造标签  $flag$  的方法如下所示, 如图6, 路径  $w_i$  中共采样得到 4 条链接, 分别记为  $edge_1$ 、 $edge_2$ 、 $edge_3$  和  $edge_4$ . 每条链接都拥有一个标签  $flag$ , 该标签标记了当前链接与源节点  $v$  之间的平衡关系, 满足平衡三元关系假设“该链接到源节点之间的路径包含偶数条负边”的标签记为  $flag=friend$ , 不满足的标签记为  $flag=enemy$ . 以链接  $edge_2$  为例, 由于该条链接到源节点  $v$  之间的路径需要经过奇数条负边, 则由平衡理论可得  $edge_2$  的标签  $flag=enemy$ .

根据路径中各链接标签  $flag$  的不同, 本文分别使用 2 个多层感知机 (MLP) 为任意链接  $(u_i, u_j, t, s)$  生成对应的低维向量表示  $msg_i$ , 如式 (4) 所示:

$$msg_i = \begin{cases} MLP^+(Z_{u_i} \| Z_{u_j} \| \Delta_t), & flag = friend \\ MLP^-(Z_{u_i} \| Z_{u_j} \| \Delta_t), & flag = enemy \end{cases} \quad (4)$$

其中,  $Z_{u_i}$  和  $Z_{u_j}$  表示节点的初始属性特征,  $\Delta_t$  表示当前链接与上一条链接间交互的时间间隔,  $\|$  表示向量的拼接操作. 可以发现, 各链接对应的时间间隔  $\Delta_t$  并不完全相同, 若使用传统的 LSTM 处理该时间间隔不规则的链接向量序列则会影响模型信息聚合的性能表现.

文献[29]最早提出了时序感知的长短期记忆网络 T-LSTM 用于处理医疗患者诊断记录时间间隔不规则的问题. 本文在该方法的基础上构建了一个 T-LSTM 信息聚合器用于计算各条路径对应的向量表示.

T-LSTM 单元的结构如图7所示, 该单元会将上

一时刻输出的节点记忆  $C_{t-1}$  划分为长期记忆  $C_{t-1}^L$  和短期记忆  $C_{t-1}^S$ , 并利用时间衰减函数  $1/\ln(e + \Delta_t)$  对短期记忆  $C_{t-1}^S$  进行遗忘, 如式 (5)–式 (8) 所示:

$$C_{t-1}^S = \tanh(W_s \cdot C_{t-1} + b_s) \quad (5)$$

$$C_{t-1}^L = C_{t-1} - C_{t-1}^S \quad (6)$$

$$\hat{C}_{t-1}^S = C_{t-1}^S \times \frac{1}{\ln(e + \Delta_t)} \quad (7)$$

$$\hat{C}_{t-1} = C_{t-1}^L + \hat{C}_{t-1}^S \quad (8)$$

其中,  $\tanh$  为激活函数,  $W_s$  为可学习的权重矩阵,  $b_s$  为偏置,  $\Delta_t$  表示时间间隔. 如式 (7) 所示, 时间间隔越小, T-LSTM 对短期记忆的衰减就越小, 表明在链接向量序列中近期交互携带的信息对当前神经网络的输出具有更高的贡献.

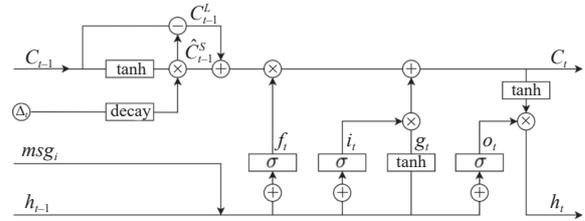


图7 T-LSTM 单元结构图

经过衰减后的节点记忆  $\hat{C}_{t-1}$  会参与标准的 LSTM 计算, 如式 (9)–式 (14) 所示:

$$f_t = \sigma(W_f \cdot msg_i + U_f \cdot h_{t-1} + b_f) \quad (9)$$

$$i_t = \sigma(W_i \cdot msg_i + U_i \cdot h_{t-1} + b_i) \quad (10)$$

$$o_t = \sigma(W_o \cdot msg_i + U_o \cdot h_{t-1} + b_o) \quad (11)$$

$$\tilde{C}_t = \tanh(W_c \cdot msg_i + U_c \cdot h_{t-1} + b_c) \quad (12)$$

$$C_t = f_t \times \hat{C}_{t-1} + i_t \times \tilde{C}_t \quad (13)$$

$$h_t = o_t \times \tanh(C_t) \quad (14)$$

其中,  $\sigma$  为激活函数 Sigmoid,  $W$  和  $U$  是可学习的权重矩阵,  $b$  为偏置.

对于向量输入序列  $M = (msg_1, msg_2, \dots, msg_i, \dots)$ , 路径信息生成模块会将经过非线性变换后 T-LSTM 最终输出的隐层状态  $h_i$  作为当前游走路径  $w_i$  对应的向量表示  $P_i$ , 如式 (15) 和式 (16) 所示:

$$h_i = T-LSTM(M) \quad (15)$$

$$P_i = \tanh(W_i h_i + b_i) \quad (16)$$

### 3.3 注意力嵌入

第3.2节得到的路径向量表示成功嵌入了动态符号网络中目标节点 $v$ 与其高阶邻居间长距离交互所携带的符号语义和时序信息。然而,仅使用T-LSTM聚合器很难区分相似度高的路径信息。为了提高表示学习的效果,并为不同路径分配不同的权重系数以减少噪声对信息聚合的影响,本文引入了图注意力机制(graph attention mechanism)将路径信息生成模块计算得到的 $n$ 个路径向量聚合为目标节点的嵌入表示。

文献[30]最早提出了图注意力网络(GAT),利用改良后的自注意力层来解决以往图卷积网络无法区分节点重要性的问题。文献[25]进一步将图注意力网络推广到了符号图中,用于计算不同符号类型的链接对节点嵌入产生的影响。在上述工作的基础上,本文将图注意力机制应用于长距离路径的信息融合过程,设计了该注意力嵌入模块以更全面地捕捉时序符号路径中携带的关键信息。

注意力嵌入模块的输入为目标节点 $v$ 对应的 $n$ 条游走路径向量表示组成的集合 $path = \{P_1, P_2, \dots, P_n\}$ ,则各路径对应注意力系数的计算过程如式(7)所示:

$$\alpha_i = \frac{\exp\left(\text{LeakyReLU}\left(\vec{a}^T \left[ \mathbf{W}_\alpha Z_v \parallel \mathbf{W}_\beta P_i \right]\right)\right)}{\sum_{P_i \in path} \exp\left(\text{LeakyReLU}\left(\vec{a}^T \left[ \mathbf{W}_\alpha Z_v \parallel \mathbf{W}_\beta P_i \right]\right)\right)} \quad (17)$$

其中, $\mathbf{T}$ 表示向量转置, $\parallel$ 表示向量拼接操作, $\mathbf{W}_\alpha$ 和 $\mathbf{W}_\beta$ 是可学习的权重矩阵, $\vec{a}$ 为注意力权重向量, $\text{LeakyReLU}$ 是非线性激活函数。

在得到归一化的注意力系数后,注意力嵌入模块使用该权重系数对 $n$ 条游走路径对应的向量表示进行加权求和,其结果作为每个节点的最终嵌入表示,如式(18)所示:

$$X_v = \sigma \left( \sum_{P_i \in path} \alpha_i \mathbf{W}_x P_i \right) \quad (18)$$

其中, $X_v \in \mathbb{R}^d$ , $d$ 表示节点嵌入的维度, $\mathbf{W}_x$ 是可学习的权重矩阵。

### 3.4 损失函数

动态符号网络中的链接预测问题是一个三分类任务,即需要预测两个节点间形成的链接类型是正链接/负链接/无链接。因此本文构造了一个带有Softmax激活函数的MLP预测器,通过拼接两个节点间的嵌入表示

判断节点对 $(v, u)$ 间可能形成链接类型的概率,如式(19):

$$\hat{y} = \text{Softmax}(\text{MLP}(X_v \parallel X_u)) \quad (19)$$

本文使用标签 $\{1, 2, 3\}$ 分别代表正链接/负链接/无链接这3种链接类型,并最小化如下交叉熵损失函数以训练RLLP方法。

$$\text{Loss} = - \sum_{(v, u, t, s) \in \mathcal{M}} \sum_{i=1}^3 y_i \log \hat{y}_i \quad (20)$$

其中, $\mathcal{M}$ 表示动态符号网络的训练边集合 $\mathcal{E}_{\text{train}}$ 与负采样得到的不相连节点对构成的无链接集合 $\mathcal{E}_{\text{neg}}$ 的并集,即 $\mathcal{M} = \mathcal{E}_{\text{train}} \cup \mathcal{E}_{\text{neg}}$ , $y_i$ 是节点对 $(v, u)$ 间形成的真实链接标签,使用one-hot编码表示, $\hat{y}_i$ 是节点对 $(v, u)$ 间形成链接类型预测为第 $i$ 类的概率,满足 $\sum \hat{y}_i = 1$ 。

该损失函数旨在评估模型预测概率分布与真实标签概率分布之间的差异程度。通过最小化该损失函数,RLLP方法能够找到最接近真实标签出现概率的参数配置。

## 4 实验分析

### 4.1 实验数据集

本文的实验部分共选取了3个真实世界的动态符号网络数据集以评估RLLP方法在链接预测任务上的性能表现,分别为Wikirfa、Epinions和Weibo(微博),各数据集的汇总统计信息如表1所示。

表1 动态符号社会网络数据集汇总统计

数据集	节点	正链接数	负链接数	时间跨度(天)
Wikirfa	1956	35560	10500	365
Epinions	2328	70035	15833	425
Weibo	14021	146782	167757	460

(1) Wikirfa<sup>[31]</sup>: 网络中的节点代表维基百科成员,带时间戳的符号链接代表“请求管理员(request for adminship)”的投票结果,社区成员可对维基百科管理员的选举投支持、中立或反对的一票,本文只选取了其中表示支持和反对的投票数据构建动态符号网络。

(2) Epinions<sup>[32]</sup>: 是一个在线商品评价网络,网络中的节点表示网站用户,用户之间存在信任/不信任关系,这也构成了网络中带时间信息的符号链接。

(3) Weibo<sup>[33]</sup>: 微博是一个基于用户关系的社交媒体平台,网络中的节点表示平台中的注册用户,用户可以在转发其他用户文章或评论的同时发表带有正面或负面情感类型的言论。

## 4.2 实验评价指标

本文选用  $F1$  分数 ( $F1$ -score) 和准确率 ( $Accuracy$ ) 作为动态符号网络中链接预测任务的评价指标.  $F1$  分数为精确率 ( $precision$ ) 和召回率 ( $recall$ ) 的调和平均数, 计算方法如式 (21)–式 (23) 所示:

$$precision = \frac{1}{3} \sum_{i=1}^3 \frac{TP_i}{TP_i + FP_i} \quad (21)$$

$$recall = \frac{1}{3} \sum_{i=1}^3 \frac{TP_i}{TP_i + FN_i} \quad (22)$$

$$F1\text{-score} = 2 \cdot \frac{precision \cdot recall}{precision + recall} \quad (23)$$

准确率表示预测正确的样本数占总样本数的比例, 计算方法如式 (24) 所示:

$$Accuracy = \frac{\sum_{i=1}^3 (TP_i + TN_i)}{\sum_{i=1}^3 (TP_i + TN_i + FP_i + FN_i)} \quad (24)$$

其中,  $TP_i$  表示模型把第  $i$  类正样本正确预测为正类的数量;  $TN_i$  表示模型把第  $i$  类负样本正确预测为负类的数量;  $FP_i$  表示模型把第  $i$  类负样本错误预测为正类的数量;  $FN_i$  表示模型把第  $i$  类正样本错误预测为负类的数量.

## 4.3 实验环境及参数设置

实验运行在 Ubuntu 18.04 操作系统上, 处理器为 Intel(R) Xeon(R) Silver 4310 CPU @ 2.10 GHz, 显卡为 NVIDIA GeForce RTX 4090, 模型代码在 Python 3.9.16 中使用 PyTorch 1.13.1 开发实现.

在 RLLP 的训练过程中, 本文使用了网格搜索的

方法寻找最优参数, 并使用 Adam 优化器进行自适应学习率优化, 设置模型的初始学习率为 0.001, batch 大小为 512, 节点表示的嵌入维度为 128, 同时为了防止过拟合, 设置 dropout 为 0.5.

## 4.4 方法对比及结果分析

为了验证 RLLP 方法在链接预测任务上的有效性, 本文将其与相似性度量方法和基于表示学习的方法进行比较, 共选取了以下 7 个基线方法.

(1) CN<sup>[12]</sup>: 该方法基于两个节点的共同邻居数量设计了相似性度量函数, 即  $S = |N_v \cap N_u|$ .

(2) AA<sup>[34]</sup>: 该方法基于共同邻居的节点度设计了相似性度量函数, 即  $S = \sum 1/\log(k_w), w \in |N_v \cap N_u|$ .

(3) node2vec<sup>[35]</sup>: 该方法是一种在静态无符号网络中学习可扩展节点嵌入的半监督算法, 它结合了 BFS 和 DFS 策略来指引随机游走过程.

(4) SIGNet<sup>[19]</sup>: 该方法根据平衡理论提出了静态符号图上的随机游走策略.

(5) SGCN<sup>[24]</sup>: 该方法根据平衡理论提出符号图上的卷积图神经网络, 有效聚合了节点的局部邻域信息.

(6) SiGAT<sup>[25]</sup>: 该方法考虑了静态符号网络边的有向性, 使用了图注意力网络 GAT 来学习节点表示.

(7) SDGNN<sup>[26]</sup>: 该模型在应用图注意力网络的基础上, 将社会学理论引入了损失函数的计算中.

表 2 显示了本文提出的 RLLP 方法与其他基线方法在 Wikirfa、Epinions 和 Weibo 数据集上的实验结果. 表中数据为 10 次实验的平均值, 其中粗体表示最佳结果, 下划线表示次优结果. 接下来对表 2 中的结果进行分析.

表 2 各对比方法在 3 个数据集上的链接预测实验结果

类别	模型	Wikirfa		Epinions		Weibo	
		$F1$ -score	$Accuracy$	$F1$ -score	$Accuracy$	$F1$ -score	$Accuracy$
相似性度量	CN	0.308	0.347	0.311	0.370	0.264	0.289
	AA	0.312	0.352	0.318	0.369	0.271	0.307
表示学习	node2vec	0.472	0.547	0.460	0.588	0.389	0.454
	SIGNet	0.700	0.783	0.698	0.804	0.581	<u>0.742</u>
	SGCN	0.684	0.749	0.709	0.823	0.570	0.716
	SiGAT	0.701	0.816	0.716	<u>0.862</u>	0.556	0.734
	SDGNN	<u>0.715</u>	<u>0.830</u>	<u>0.729</u>	0.860	<u>0.597</u>	0.739
	RLLP	<b>0.734</b>	<b>0.852</b>	<b>0.750</b>	<b>0.881</b>	<b>0.615</b>	<b>0.769</b>

首先, 在 3 个数据集的实验中, 相似性度量方法 CN 和 AA 的性能表现较差, 其对应的  $F1$  分数和准确率都较低, 原因在于这两种方法仅考虑节点在网络中的拓扑结构, 而忽略了动态符号网络中出现的符号语

义和时间信息, 并且根据两个节点间的相似性分数很难精确地将动态符号网络中的链接区分为正链接/负链接/无链接这 3 种类型.

其次, 在 Wikirfa 和 Epinions 数据集中, 基于表示

学习和图注意力网络的 SiGAT 和 SDGNN 方法取得了较优的性能表现,说明图注意力机制能够有效地为不同类型的局部领域信息分配不同权重,生成更具代表性的网络节点嵌入表示。而经典的 node2vec 方法由于无法学习网络中同时出现的符号和时间信息只取得了略优于相似性度量方法的实验结果。

另外,在 Weibo 数据集中大多数模型的性能表现都不如在 Wikirfa 和 Epinions 数据集上得到的结果,这是因为 Weibo 数据集中节点间形成的负链接关系占比达到 53%,并且 Weibo 网络拥有较大的时间跨度,频繁出现的负向符号语义和复杂的网络时序演化会影响表示学习方法的信息聚合过程,最终干扰链接预测任务的性能表现。其中,基于平衡理论指导的随机游走 SIGNet 方法在 Weibo 数据集上取得了较优的准确率表现,说明在带符号的社会网络中应用平衡理论和随机游走策略可以较好地捕获长距离路径中频繁出现的负向符号语义信息,生成更准确的网络表示。

最后,本文提出的 RLLP 方法在所有对比方法中表现最佳,说明同时嵌入动态符号网络中出现的符号语义和时间信息能够提升链接预测任务的性能。RLLP 方法主要包含 3 个关键部分,时序随机游走模块内设计的链接采样策略捕获了目标节点与其高阶邻居间长距离的路径依赖;基于平衡理论和 T-LSTM 设计的路径信息生成模块有效融合了长路径中同时携带的符号和时间信息;而注意力嵌入模块能够通过图注意力机制为每个节点生成更富表达能力的嵌入表示。与现有的相似性度量方法及基于表示学习的链接预测方法相比,RLLP 方法在  $F1$  分数和准确率上展现出一定优势。

通过分析在 3 个数据集中的链接预测实验结果,RLLP 方法的优点主要体现在以下 3 个方面。首先,与传统的基于相似性度量的方法相比,本方法提出的动态符号网络时序随机游走算法能够充分捕获目标节点周围的拓扑结构信息,避免了相似性方法的复杂计算;其次,与仅能挖掘网络静态结构特征的表示学习方法相比,本方法基于平衡理论和 T-LSTM 设计的路径信息生成模块能够有效嵌入长距离路径中出现的符号语义和时间信息,在时间跨度较大、负链接占比较高的 Weibo 数据集中也能准确捕获网络的复杂演化过程。最后,本方法在计算节点最终嵌入时使用的图注意力机制能够考虑不同路径向量的聚合权重,以降低数据噪声对表示学习准确性产生的影响。因此,RLLP 方法

展现出很高的可行性和有效性。

#### 4.5 消融实验

本文旨在通过消融实验来评估 RLLP 方法中不同功能模块的有效性。以下是 RLLP 方法的 3 个变体。

(1) -RW: 将随机游走路径最大长度  $L$  设置为 1,此时 RLLP 方法仅能够聚合目标节点的一阶邻居信息。

(2) -T-LSTM: 用均值聚合器替代原始 RLLP 方法路径信息生成模块中的 T-LSTM 聚合器,即将链接向量序列的平均值作为当前游走路径对应的向量表示。

(3) -GAT: 移除注意力嵌入模块使用的图注意力机制,即将各路径向量的注意力权重系数设置为相同的值。

本文在 3 个数据集上进行了消融实验,并选取  $F1$  分数作为性能衡量指标,实验结果如图 8 所示。可以发现单独移除 3 个功能模块都会使 RLLP 方法性能产生下降,下面将深入分析消融实验结果。

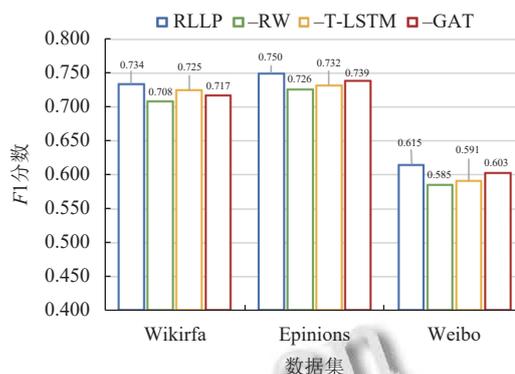


图 8 消融实验结果

移除 RLLP 方法中的时序随机游走模块,该方法变体 -RW 在 Wikirfa、Epinions 和 Weibo 这 3 个数据集上的  $F1$  分数分别下降了 3.5%、3.2%、4.8%,这表明通过在动态符号网络中引入有偏的时序随机游走采样方法能够更有效地捕获网络中节点间长距离的信息依赖,从而提高了方法在链接预测任务上的性能表现。

将使用 T-LSTM 生成路径信息的聚合方法替换为均值聚合后,变体 -T-LSTM 在 Weibo 数据集上的  $F1$  分数下降了 3.9%,证明了 T-LSTM 可以有效嵌入路径中出现的时序信息,并充分考虑不均匀交互时间间隔带来的影响。

当图注意力机制被移除后,各条语义路径的注意力系数被设置成相同的值,变体 -GAT 在 Wikirfa 数据集上的  $F1$  分数下降了 2.3%,这表明图注意力机制能够赋予关键语义路径更高的权重系数,削弱噪声带来的影响,进而增强节点嵌入的表达力。

#### 4.6 参数敏感性分析

RLLP 方法包含 2 个重要的参数: 时序随机游走生成的路径条数  $n$  以及节点嵌入的最终维度  $d$ . 本文在 3 个数据集上进行了参数的敏感性分析实验, 并选取 F1 分数作为实验的衡量指标.

(1) 时序随机游走生成的路径条数  $n$ . 本文设置生成的路径条数  $n = [10, 20, 30, 40, 50, 60]$ , 实验结果如图 9 所示. 可以看出, 随着路径条数  $n$  的增加, RLLP 方法在链接预测任务上的性能表现也逐渐上升, 这说明更多的路径条数能够帮助 RLLP 方法更充分地捕获网络中蕴含的符号语义和时序信息. 然而, 当  $n$  足够大时, 时序随机游走的时间开销也会进一步增加, 并且对于网络中一些交互较少的节点而言, 设置过大的路径条数会生成许多重复的路径信息, 从而影响 RLLP 方法学习节点嵌入的准确性.

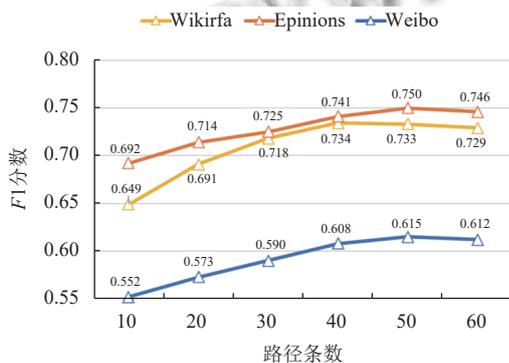


图 9 路径条数  $n$  敏感性分析结果

(2) 节点嵌入维度  $d$ . 本文设置节点嵌入维度  $d = [16, 32, 64, 128, 256]$  维, 实验结果如图 10 所示. 可以看出, 随着  $d$  从 16 维增加至 128 维, RLLP 方法性能有着明显的上升趋势, 这表示更高的维度会使节点嵌入拥有更强大的信息表示能力, 但在  $d$  超过 128 维达到 256 维后, 模型性能会开始下降, 这是因为嵌入维度过大导致方法中参数增加, 使得表示学习的训练更加困难, 并且可能导致过拟合, 损害链接预测的性能.

#### 5 结束语

本文提出了一种基于表示学习的动态符号社会网络链接预测方法 RLLP. 针对当前大多数基于节点嵌入的链接预测算法无法同时学习网络中的符号语义和时间信息的问题, 本文所提出的基于平衡理论设计的 T-LSTM 聚合器能够有效捕获网络中蕴含的复杂信息. 其次, 针对图神经网络模型难以捕获目标节点与高阶

邻居间长距离信息依赖的问题, 本文在平衡理论的指导下设计了一种能够在动态符号网络中进行时序随机游走的链接采样策略用于生成可供学习的时序符号语义路径, 并且融合了图注意力机制的节点嵌入模块可以提取路径中的关键信息, 生成更具表达能力的节点低维向量表示. 此外, 在 3 个真实数据集上的对比实验和进一步的消融实验和参数分析实验结果都验证了 RLLP 方法在链接预测任务中的优越性和有效性.

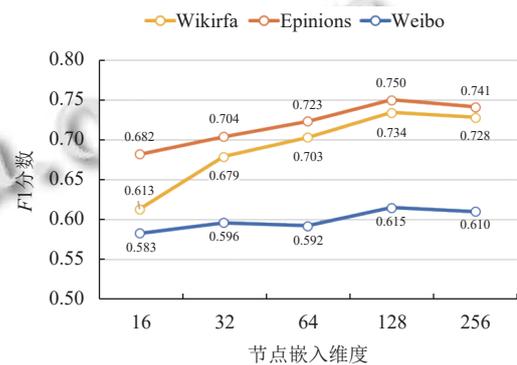


图 10 节点嵌入维度  $d$  敏感性分析结果

#### 参考文献

- 1 Tarigan IM, Harahap MAK, Sari DM, *et al.* Understanding social media: Benefits of social media for individuals. *Jurnal Pendidikan Tambusai*, 2023, 7(1): 2317–2322.
- 2 Banerjee T, Bernstein A, Feinstein Z. Dynamic clearing and contagion in financial networks. *European Journal of Operational Research*, 2025, 321(2): 664–675. [doi: 10.1016/j.ejor.2024.09.046]
- 3 Leskovec J, Huttenlocher D, Kleinberg J. Signed networks in social media. arXiv:1003.2424, 2010.
- 4 Leskovec J, Huttenlocher D, Kleinberg J. Predicting positive and negative links in online social networks. arXiv: 1003.2429, 2010.
- 5 Kumar A, Singh SS, Singh K, *et al.* Link prediction techniques, applications, and performance: A survey. *Physica A: Statistical Mechanics and its Applications*, 2020, 553: 124289. [doi: 10.1016/j.physa.2020.124289]
- 6 Chen L, Xie YZ, Zheng ZB, *et al.* Friend recommendation based on multi-social graph convolutional network. *IEEE Access*, 2020, 8: 43618–43629. [doi: 10.1109/ACCESS.2020.2977407]
- 7 Liu GG. An ecommerce recommendation algorithm based on link prediction. *Alexandria Engineering Journal*, 2022, 61(1): 905–910. [doi: 10.1016/j.aej.2021.04.081]
- 8 Hao W, Zhan HL, Bao XJ, *et al.* Bond transaction link prediction based on dynamic network embedding and time

- series analysis. Proceedings of the 6th International Conference on Systems and Informatics. Shanghai: IEEE, 2019. 1471–1477.
- 9 Zhang YJ, Li JB, Li Y, *et al.* Cellular traffic offloading via link prediction in opportunistic networks. *IEEE Access*, 2019, 7: 39244–39252. [doi: [10.1109/ACCESS.2019.2891642](https://doi.org/10.1109/ACCESS.2019.2891642)]
- 10 Abbas K, Abbasi A, Dong S, *et al.* Application of network link prediction in drug discovery. *BMC Bioinformatics*, 2021, 22(1): 187. [doi: [10.1186/s12859-021-04082-y](https://doi.org/10.1186/s12859-021-04082-y)]
- 11 Daud NN, Ab Hamid S H, Saadoon M, *et al.* Applications of link prediction in social networks: A review. *Journal of Network and Computer Applications*, 2020, 166: 102716. [doi: [10.1016/j.jnca.2020.102716](https://doi.org/10.1016/j.jnca.2020.102716)]
- 12 Zeng S. Link prediction based on local information considering preferential attachment. *Physica A: Statistical Mechanics and Its Applications*, 2016, 443: 537–542. [doi: [10.1016/j.physa.2015.10.016](https://doi.org/10.1016/j.physa.2015.10.016)]
- 13 Zhao H, Du L, Buntine W. Leveraging node attributes for incomplete relational data. Proceedings of the 34th International Conference on Machine Learning. Sydney: JMLR.org, 2017. 4072–4081.
- 14 Wang SH, Tang JL, Aggarwal C, *et al.* Signed network embedding in social media. Proceedings of the 2017 SIAM International Conference on Data Mining. Houston: Society for Industrial and Applied Mathematics, 2017. 327–335.
- 15 Javari A, Derr T, Esmailian P, *et al.* ROSE: Role-based signed network embedding. Proceedings of the 2020 Web Conference. Taipei: ACM, 2020. 2782–2788.
- 16 Wu ZH, Pan SR, Chen FW, *et al.* A comprehensive survey on graph neural networks. *IEEE Transactions on Neural Networks and Learning Systems*, 2021, 32(1): 4–24. [doi: [10.1109/TNNLS.2020.2978386](https://doi.org/10.1109/TNNLS.2020.2978386)]
- 17 Giraldo JH, Skianis K, Bouwmans T, *et al.* On the trade-off between over-smoothing and over-squashing in deep graph neural networks. arXiv:2212.02374v2, 2023.
- 18 Heider F. Attitudes and cognitive organization. *The Journal of Psychology*, 1946, 21(1): 107–112. [doi: [10.1080/00223980.1946.9917275](https://doi.org/10.1080/00223980.1946.9917275)]
- 19 Islam MR, Aditya Prakash B, Ramakrishnan N. SIGNet: Scalable embeddings for signed networks. Proceedings of the 22nd Pacific-Asia Conference on Advances in Knowledge Discovery and Data Mining. Melbourne: Springer, 2018. 157–169.
- 20 Kim J, Park H, Lee JE, *et al.* SIDE: Representation learning in signed directed networks. Proceedings of the 2018 World Wide Web Conference. Lyon: International World Wide Web Conferences Steering Committee, 2018. 509–518.
- 21 Wang HJ, Qu CQ, Jiao CZ, *et al.* Self-avoiding pruning random walk on signed network. *New Journal of Physics*, 2019, 21(3): 035001. [doi: [10.1088/1367-2630/ab060f](https://doi.org/10.1088/1367-2630/ab060f)]
- 22 Babul SA, Tian Y, Lambiotte R. Strong and weak random walks on signed networks. *npj Complexity*, 2025, 2(1): 4. [doi: [10.1038/s44260-025-00027-1](https://doi.org/10.1038/s44260-025-00027-1)]
- 23 Zhou J, Cui GQ, Hu SD, *et al.* Graph neural networks: A review of methods and applications. *AI Open*, 2020, 1: 57–81. [doi: [10.1016/j.aiopen.2021.01.001](https://doi.org/10.1016/j.aiopen.2021.01.001)]
- 24 Derr T, Ma Y, Tang JL. Signed graph convolutional networks. arXiv:1808.06354v1, 2018.
- 25 Huang JJ, Shen HW, Hou L, *et al.* Signed graph attention networks. arXiv:1906.10958v3, 2019.
- 26 Huang JJ, Shen HW, Hou L, *et al.* SDGNN: Learning node representation for signed directed networks. Proceedings of the 35th AAAI Conference on Artificial Intelligence. Palo Alto: AAAI Press, 2021. 196–203.
- 27 王岩, 任浩, 王喆. 动态连续时间网络表示学习. *计算机工程与应用*, 2022, 58(12): 163–169. [doi: [10.3778/j.issn.1002-8331.2012-0285](https://doi.org/10.3778/j.issn.1002-8331.2012-0285)]
- 28 Mikolov T, Sutskever I, Chen K, *et al.* Distributed representations of words and phrases and their compositionality. Proceedings of the 27th International Conference on Neural Information Processing Systems. Lake Tahoe: Curran Associates Inc., 2013. 3111–3119.
- 29 Baytas IM, Xiao C, Zhang X, *et al.* Patient subtyping via time-aware LSTM networks. Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Halifax: ACM, 2017. 65–74.
- 30 Veličković P, Cucurull G, Casanova A, *et al.* Graph attention networks. arXiv:1710.10903, 2017.
- 31 West R, Paskov HS, Leskovec J, *et al.* Exploiting social network structure for person-to-person sentiment analysis. *Transactions of the Association for Computational Linguistics*, 2014, 2: 297–310. [doi: [10.1162/tacl\\_a\\_00184](https://doi.org/10.1162/tacl_a_00184)]
- 32 Richardson M, Agrawal R, Domingos P. Trust management for the Semantic Web. Proceedings of the 2nd International Semantic Web Conference. Sanibel Island: Springer, 2003. 351–368.
- 33 Zhang J, Liu B, Tang J, *et al.* Social influence locality for modeling retweeting behaviors. Proceedings of the 23rd International Joint Conference on Artificial Intelligence. Beijing: AAAI Press, 2013. 2761–2767.
- 34 Adamic LA, Adar E. Friends and neighbors on the Web. *Social Networks*, 2003, 25(3): 211–230. [doi: [10.1016/S0378-8733\(03\)00009-1](https://doi.org/10.1016/S0378-8733(03)00009-1)]
- 35 Grover A, Leskovec J. node2vec: Scalable feature learning for networks. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco: ACM, 2016. 855–864.

(校对责编:王欣欣)

Special Issue 专论·综述 21