图神经网络在动态图上的链路预测综述

王天赐
计算机学院
湖北工业大学
武汉, 湖北, 中国

摘要—图形神经网络（GNNs）正迅速成为在图形结构数据上学习的主导方式。链接预测是新的 GNN 模型的一个几乎通用的基准。许多先进的模型，如动态图神经网络（DGNNs）专门针对动态链接预测。然而，这些模型，特别是 DGNNs，很少与其他模型或现有启发式方法进行比较。不同的工作以不同的方式评估他们的模型，因此人们无法直接比较评估指标。受此启发，我们进行了一次全面的比较研究。我们比较了链接预测启发式方法、GNNs、离散 DGNNs 和连续 DGNNs 对动态链接预测的作用。我们发现，简单的链接预测启发式方法往往比 GNN 和 DGNN 表现更好，而在所有被考察的图神经网络中，DGNN 的表现一直优于静态 GNN。

***Index Terms*—动态图神经网络; 启发式; 链接预测;**

I. 引言

A. 研究背景

在本文中主要的研究内容是在动态图上的链接预测, 与静态图的链接预测不同的是, 动态图的链接预测增加了时间维度, 预测难度增加, 但是在现实中的应用更加广泛。

B. 研究方法

在近年来的研究中, 很多链路预测方法取得了发展, 在静态图上研究者们提出了很多高精度的链接预测方法。比如基于启发式算法的共同邻居算法 [1], 以及首次将 GNN 应用在图的链接预测的 SEAL 模型 [1], SEAL 模型

C. 研究问题

近年来，图神经网络（GNNs）作为一个新兴的研究领域，得到了长足的发展，提出并发展了多种体系结构。但是在动态图神经网络（DGNN）领域，这些问题因以下原因而进一步加剧：

1) 数据的动态性质:

2) 缺乏通用术语:

3) 缺乏既定的强大基线（大多数研究不与其他 DGNN 比较性能）:

4) 离散和连续 DGNN 之间的鸿沟:

5) 大量的实验设计选择: 这些选择包括：如何表示动态网络（如快照、时间窗口、连续、边的生存时间等），包括哪些节点特征，如何将数据分成训练-验证-测试集，用哪些指标来评价结果，如何在报告的指标中使用负采样率，以及如何选择/优化神经网络参数（如学习率、早期停止准则、嵌入空间维度等）。所有这些都意味着，通过阅读研究论文来比较方法的性能是不可能的，除非他们清楚地说明他们所有的设计选择，而且这些设计选择在不同的论文中是相同的。这些选择包括：如何表示动态网络（如快照、时间窗口、连续、边的生存时间等），包括哪些节点特征，如何将数据分成训练-验证-测试集，用哪些指标来评价结果，如何在报告的指标中使用负采样率，以及如何选择/优化神经网络参数（如学习率、早期停止准则、嵌入空间维度等）。所有这些都意味着，通过阅读研究论文来比较方法的性能是不可能的，除非他们清楚地说明他们所有的设计选择，而且这些设计选择在不同的论文中是相同的。DGNNs 是建立网络动态模型的一个很有前途的途径，因为它们既能通过 GNNs 编码空间模式，又能通过时间序列组件（如循环神经网络（RNN）或自我注意）编码时间模式。然而，迄今为止提出的 DGNNs 已经在少数数据集上进行了测试，并且很少与其他 DGNNs 进行比较。不同的研究在不同的数据集上进行比较，因为在 DGNN 基准测试中使用哪种数据集的问题上没有共识。

II. 相关研究模型与数据集

A. 相关数据集

我们选择了五个连续的交互网络和一个离散的演化网络（Autonomous）作为数据集。我们选择了互动网络，因为它们允许我们轻松地转换为更粗粒度的时间粒度，如离散网络。更稀疏的快照表明链接和非链接之间有更大的不平衡性，从而使分类问题更难。我们为每个数据集准备了两个版本，一个是有方向的连续交互网络，一个是无方向的离散网络。连续模型对连续网络进行编码。静态和离散模型对离散网络进行编码。在从连续到离散的转换中，互换的边缘被添加到离散网络中，使其成为无定向的。一条边在快照中出现的次数被作为权重加到快照的边上。所有的结果都是对离散网络的预测报告。对于连续模型，这是通过将连续网络的连续部分分割成与离散网络中的快照相对应的快照来实现的。然后，我们让连续模型在目标快照前对连续网络进行编码，然后尝试预测离散网络中的链接发生。在本文中，我们使用了以下数据集：

1) Enron: 这个数据集是一个电子邮件通信网络，其中一个链接是两个人之间发送的电子邮件。Enron 在空间上是一个很小的网络，但在时间上是一个中等规模的网络，有合理数量的连续链接，覆盖的时间跨度超过 3 年。由于节点数量少，边的数量相对较多，它比其他网络要密集得多。

2) UC Irvine Messages: 简称 UC，是加州大学欧文分校的一个在线论坛网络。如果两个学生在同一个论坛帖子上互动，他们就会被连接起来。因此，这原本是一个二方网络，但它被预测为只有一种类型的节点。快照大小的奇特选择来自于 EvolveGCN（Pareja 等人，2020），它观察到较小的快照大小会产生一些没有任何边的快照

3) Bitcoin-OTC: 这个数据集是一个在比特币 OTC 平台上交易的谁信任谁的网络。一个链接是一个用户对另一个用户的评价。比特币网络在节点方面是中等规模的，然而，它的大多数边是唯一的边，这表明很少有边是重复出现的。缺少重复出现的边缘导致每个快照比其他大多数数据集要稀疏得多。

4) Autonomous-systems: 这个数据集是一个互联网路由器通信网络。一个链接是一个路由器与一个对等体交换流量。这个网络已经被汇总为一个离散的网络并选择前 99 天，并将其作为我们的数据集。这是迄今为止拥有最多边缘的数据集。

5) Wikipedia: 这个数据集是一个双联的维基百科页面编辑网络。节点是一个维基百科用户或一个维基百科页面。一个链接是一个编辑维基百科页面的用户。维基百科也有很少的重现边缘，与比特币类似，然后有比较稀疏的时间快照。

6) Reddit: 这个数据集是一个双联的 Reddit 发布网络。节点是一个 Reddit 用户或一个 subreddit。一个链接是一个用户在一个子 reddit 上的发帖。Reddit 是空间上最大的网络，因为它有最大数量的节点和独特的边。

B. 相关研究模型

1) GCN: GCN 借用了卷积神经网络（CNN）的卷积概念，直接根据图的连通结构对图进行卷积，作为滤波器进行邻域混合

2) GAT: 将注意机制与 GNN 相结合，旨在更有效地学习邻里特征。图形注意层作为 GAT 的组成部分，对 GNN 起着聚合作用。它首先对每个节点应用一个共享的线性变换，由权重矩阵 W 参数化　射频 0×F

3) GC-LSTM: 这个新的深度模型中的 GCN 能够为每个时间段滑动节点结构学习网络快照，而 LSTM 负责网络快照的时间特征学习。此外，当前的动态链路预测方法只能处理删除的链接，GC-LSTM 可以同时预测添加或删除的链接。

4) TGAT: 时间图注意力（TGAT）层来有效地聚合时间拓扑邻域特征以及学习时间特征交互通过堆叠 TGAT 层，网络将节点嵌入识别为时间的函数，并能够随着图的发展归纳推断新节点和观察到的节点的嵌入。TGAT 可以同时处理节点分类和链接预测任务，并且可以自然地扩展到包括时间边缘特征。

III. 相关方法

A. Graph Convolutional Networks (GCN)

GCN[9] 借用了卷积神经网络（CNN）的卷积概念，直接根据图的连通结构对图进行卷积，作为滤波器进行邻域混合。架构可以简单地概括为

$$H^{l+1} = \sigma\left(\tilde{D}^{-1/2}\tilde{A}^{-1/2}\tilde{D}^{-1/2}H^{(l)}W^{(l)}\right) \qquad (1)$$

在这里　表示 sigmoid 函数，u 和 v 是两个邻居，而 vn 是负样本，Q 是负样本数。第一项的目标是最大化 u 和

v 的嵌入之间的相似性，而第二项则试图区分负样本的嵌入

1) Graph Attention Networks (GAT): GAT[17] 将注意机制与 GNN 相结合，旨在更有效地学习邻里特征。图形注意层作为 GAT 的组成部分，对 GNN 起着聚合作用。它首先对每个节点应用一个共享的线性变换，由权重矩阵 W 参数化　射频 0×F。然后在节点上执行自我注意，其中使用共享注意机制来计算捕获节点 j 的特征对节点 i 的重要性的注意系数，即：

$$e_{ij} = a\left(\mathbf{W}h_i, \mathbf{W}h_j\right) \tag{2}$$

GAT 使用 softmax 函数对 j 上的系数进行归一化。因此，注意机制是一个由权重向量 a 参数化的单层网络 R2F 0，然后进行非线性激活（例如 LeakyReLU），并获得归一化系数，如下所示：

$$\alpha_{ij} = \frac{\exp\left(\text{LeakyReLU}\left(a\left[\mathbf{W}h_i, \mathbf{W}h_j\right]\right)\right)}{\sum_{k\in N_i}\exp\left(\text{LeakyReLU}\left(a\left[\mathbf{W}h_i, \mathbf{W}h_k\right]\right)\right)} \tag{3}$$

这些注意系数用于计算邻居特征的线性组合以获得每个节点的聚集特征，即：

$$h_i' = \sum_{k\in N_i}\alpha_{ij}\mathbf{W}h_k \tag{4}$$

最后，多头注意（即一套独立的注意机制）被用来稳定自我注意的学习过程。而多头注意力需要用到的算力可能有些相对较大了，不一定说一定适合个人训练，可以使用相关领域的预处理模型来迁移学习。

2) Graph Convolution Embedded LSTM: LSTM 被作为主要框架来学习动态网络的所有快照的时间特征。而对于每个快照，GCN 被用来捕捉节点的局部结构特性以及它们之间的关系。一个好处是，GC-LSTM 可以预测添加和删除的链接。GCN 最初是为无定向网络设计的，这意味着对称网络的拉普拉斯矩阵。为了将 GCN 嵌入提议的 GC-LSTM 中，我们首先需要重新定义拉普拉斯矩阵。按照 Ma 等人的想法，有向拉普拉斯矩阵被定义为:

$$L^{sym} = I - \frac{1}{2}\left(\Phi^{1/2}P\Phi^{-1/2} + \Phi^{-1/2}P^T\right. \tag{5}$$

GC-LSTM 模型主要依靠两个状态值，一个是用于提取最后一次输入信息的隐藏状态 h，一个是用于保存长期信息的单元状态 c。GC-LSTM 的本质是它在前进过程中有一个单元状态 c，导致信息在单元状态 c 上长期传输。单元状态可以通过遗忘门和输入门进行更新，定义如下:

$$\begin{aligned}
\bar{c}_t &= \tanh\left(A_tW_c + GCN_o^K\left(\tilde{A}_{t-1}, h_{t-1}\right) + b_c\right),\\
i_t &= \sigma\left(A_tW_i + GCN_c^K\left(\tilde{A}_{t-1}, h_{t-1}\right) + b_i\right)\\
c_t &= f_t \odot GCN_c^K c_{t-1} + i_t \cdot \bar{c}_t.
\end{aligned} \tag{6}$$

在动态网络链接预测任务中，我们需要考虑邻居的隐藏状态对节点隐藏状态的影响，以及邻居的细胞状态的影响。

3) TGAT: TGAT 可以有效地聚集时间-拓扑邻域特征，并学习时间特征的相互作用。通过使用自我注意机制作为构建模块，并在经典的 Bochner 定理的基础上开发了一种新的功能性时间编码技术，即谐波分析法。通过堆叠 TGAT 层，网络将节点嵌入识别为时间函数，并且能够在图的演变过程中归纳推断新节点和观察到的节点的嵌入。根据原始的自我注意机制，我们首先得到实体-时间特征矩阵为:

$$\mathbf{Z}(t) = \left[\tilde{\mathbf{h}}_0^{(l-1)}(t)\left\|\Phi_{d_T}(0), \tilde{\mathbf{h}}_1^{(l-1)}(t_1)\right\|\Phi_{d_T}(t-t_1), \ldots, \tilde{\mathbf{h}}_N^{(l-1)}(t_N)\right\| \tag{7}$$

并将其转发给三个不同的线性投影，以获得'query','key'和'value'。

$$\mathbf{q}(t) = [\mathbf{Z}(t)]_0\mathbf{W}_Q, \mathbf{K}(t) = [\mathbf{Z}(t)]_{1:N}\mathbf{W}_K, \mathbf{V}(t) = [\mathbf{Z}(t)]_{1:N}\mathbf{W}_V \tag{8}$$

然后，我们将其传递给一个前馈神经网络，以捕捉特征之间的非线性相互作用，如:

$$\begin{aligned}
&\tilde{\mathbf{h}}_0^{(l)}(t) = \text{FFN}\left(\mathbf{h}(t)\|\mathbf{x}_0\right) \equiv \text{ReLU}\left(\left[\mathbf{h}(t)\|\mathbf{x}_0\right]\mathbf{W}_0^{(l)} + \mathbf{b}_0^{(l)}\right)\mathbf{W}_1^{(l)} + \mathbf{b}_1^{(l)}\\
&\mathbf{W}_0^{(l)} \in \mathbb{R}^{(d_h+d_0)\times d_f}, \mathbf{W}_1^{(l)} \in \mathbb{R}^{d_f\times d}, \mathbf{b}_0^{(l)} \in \mathbb{R}^{d_f}, \mathbf{b}_1^{(l)} \in \mathbb{R}^d
\end{aligned} \tag{9}$$

因此，TGAT 层可以使用半监督学习框架实现节点分类任务，也可以使用编码器-解码器框架实现链接预测任务。

IV. 结论

在本文中我们介绍并实现几种现有的离散和连续 DGNNs 的框架。我们对不同的 GNN 在动态链接预测任务上进行了综合比较。我们的实验表明，这些 GNN 架构在其他链路预测任务的基准测试中表现相似。本文的未来有几个有趣的方向。首先，我们的基准数据集仍然相对较小，将来我们可以在更大的图上评估模型，特

别是在现实世界中的应用。本文的第二个有趣的方向是实现最近开发的 GNN 模型。此外，我们可以尝试设计和开发我们自己的 DGNN 架构和链路预测任务的基准测试。

参考文献

[1] J. Zhang, J. Tang, B. Liang, Z. Yang, S. Wang, J. Zuo, and J. Li, "Recommendation over a heterogeneous social network," in *Web-Age Information Management, 2008. WAIM'08. The Ninth International Conference on*. IEEE, 2008, pp. 309–316. 1

[2] O. Ronneberger, P. Fischer, and T. Brox, *U-Net: Convolutional Networks for Biomedical Image Segmentation*. Springer International Publishing, 2015.