

基于离散化动态图的协同过滤推荐算法^①

金佳琪^{②*} 张梦菲^{***} 潘 茂^{**} 褚志海^{****} 方金云^{③*}

(* 中国科学院计算技术研究所 北京 100190)

(** 中国科学院大学 北京 100190)

(*** 国家计算机网络应急技术处理协调中心 北京 100029)

(**** 中国雄安集团有限公司 北京 071700)

摘 要 基于图卷积网络(GCN)模型在学习用户/物品表示方面表现出了强大的性能,给传统的协作过滤(CF)算法带来了新的研究突破。然而,现有的基于 GCN 的 CF 方法仍然都是针对静态图建模,而在实际场景中,用户与物品的交互不是一成不变的,会随着时间的推移而持续演化;GCN 中的过平滑问题会极大地限制现有推荐算法的表示学习建模。为解决上述问题,提出了基于动态图的协同过滤算法(DynGCF),其目的是通过同时捕获图的结构和时态演化信息来学习用户和物品的嵌入表示。DynGCF 首先采用 GCN 学习每个离散快照图上的用户/物品嵌入,然后应用时间卷积网络(TCN)和自注意力机制学习,最终嵌入表示。为缓解过平滑问题,本文改进了传统 GCN 中的关键模块,即邻域聚合,通过在 1 阶交互图和 2 阶共现图建模用户和物品的交互。在 4 个真实数据集上与基于 GCN 的 CF 方法和动态图的基线方法对比,验证了 DynGCF 的性能提升,并分析验证了改进的方法能有效缓解过平滑问题。

关键词 推荐系统; 动态图; 图卷积网络(GCN); 协同过滤(CF)

0 引 言

推荐系统已广泛应用于各种 Web 应用程序中,用于缓解信息过载并帮助用户发现他们潜在的兴趣物品^[1]。许多真实数据可以概括为用户与物品的交互数据,例如,电商或广告中的点击^[2]、新闻^[3]和视频收看^[4]等。作为推荐场景中的常用算法,协同过滤(collaborative filtering, CF)旨在从历史的用户物品交互记录中学习用户/物品的向量嵌入表示(embedding)^[5],并据此预估潜在的交互。

近年来,图卷积网络(graph convolutional network, GCN)蓬勃发展,它将卷积的概念扩展到了图结构数据^[6-8]。GCN 的基本思想是堆叠多个图卷积

层并执行迭代邻域聚合(也称为消息传递)以更新节点表示。由于用户和物品的交互可以自然地建模为一个图,GCN 促进了 CF 中的表示学习建模,将其从单用户项目实例关系建模推广到了基于图结构的多跳邻域聚合,并带来了性能的提升^[9-11]。例如,PinSage^[9]将随机游走算法应用到图卷积中,使得图结构的参数复杂度与输入图大小无关,高效地生成节点嵌入。NGCF(neural graph collaborative filtering)^[10]利用了图中的高阶连通性,从而显式地聚合相关邻居的协同信号。LightGCN^[11]则在 NGCF 的基础上,通过消去不必要的特征变换和非线性激活模块,进而达到更高的效率和更显著的准确性。

尽管有效,但现有的基于 GCN 的 CF 方法仍未

① 北京市科技计划(E031150)和河北省科技计划(E132010)资助项目。

② 男,1994 年生,博士生;研究方向:图学习与推荐系统。

③ 通信作者,E-mail: fangjy@ict.ac.cn。

(收稿日期:2022-02-20)

充分考虑到图的演化。简言之,它们都是基于静态图设计,包含了固定的节点集和边集。然而,许多实际场景中用户和物品的交互图往往是动态的,即会随着时间的推移而不断演变^[12-13]。例如,在电子商务平台中,用户和物品会不断涌现,新的交互也随之出现。在这样一个动态场景中,用户的兴趣是分阶段的,而物品的流行度等特征也会随时间而发生变化。因此,除了图结构信息外,时态信息也有助于学习用户和物品的嵌入表示。另一方面,现有基于 GCN 的 CF 方法^[10-11,14]的研究结果表明,当深度大于 3 时,性能的提升仅实现了微小的改善,甚至开始下降。这是因为单纯的堆叠图卷积层会使得所学的节点嵌入变得难以区分,即存在过平滑^[15]问题。因此,很难通过简单地提高层数来改进推荐结果。

为克服上述局限,本文提出了一种新的推荐框架,即基于动态图的协同过滤算法(dynamic graph-based collaborative filtering, DynGCF),以联合建模用户物品交互中的图结构和时态信息,并缓解过平滑问题。通常,动态图可以分为 2 类:(1)离散动态图(discrete temporal dynamic graph, DTDG)^[16-17],它由一组具有固定时间间隔的有序快照图组成;(2)连续动态图(continuous temporal dynamic graph, CT-DG)^[13],其特点是会增量连续建模每一次的用户-物品交互,交互都会有相应的时间戳。本文基于 DTDG 进行用户和物品的表示建模,根据历史用户物品交互构建快照图,并学习每个快照中的节点嵌入,提出了基于共现图的局部图表示建模方法(co-occurrence graph convolutional network, Co-GCN)来捕获图结构信息。相较于现有基于 GCN 的方法迭代聚合快照图上的节点嵌入,本文提出了简化的图网络,通过利用更多元的拓扑图结构建模更丰富的用户-物品交互、用户-用户相似等特征,进而缓解过平滑问题。为了建模时态演化信息,构建了一个由时域卷积网络(temporal convolutional network, TCN)^[18]模块和自注意层^[19]组成的层次化时态嵌入聚合层,将多阶段的局部节点特征聚合为最终的嵌入表示。实验结果证明, DynGCF 能够在 4 个基准数据集上实现比现有最先进的方法更佳的性能。本文的主要贡献总结为以下 3 点。

(1) 本文强调了利用动态图建模推荐场景下用户和物品表示的重要性,并提出了通过联合建模 1 阶交互和 2 阶共现特征的 Co-GCN。实验结果表明, Co-GCN 还有助于缓解过平滑问题。

(2) 提出新的推荐框架 DynGCF,该框架首先基于离散快照图快速提取多阶段节点特征,然后通过层次化聚合多时间步的节点特征学习到了特征更丰富的嵌入表示。

(3) 在 4 个真实数据集上的实验结果表明, DynGCF 模型在召回率(Recall)和归一化折损累计增益(normalized discounted cumulative gain, NDCG)指标上超过了现有方法。

1 相关工作

首先回顾一系列具有代表性的协同过滤算法,然后详细阐述与本文工作相关的基于动态图的协同过滤方法。

1.1 协同过滤

协同过滤(CF)假设相似的用户往往也具有相似的物品偏好,并通过建模用户和物品的交互来探索用户潜在偏好,为用户推荐兴趣物品。早期的 CF 方法,如矩阵分解(matrix factorization, MF)^[20],一般先将用户和物品的编号投影到低维稠密空间,再通过重建历史用户-物品交互矩阵来学习其潜在的表示^[21]。随后,为了进一步推动了多种 CF 方法的发展,基于深度神经网络的模型采用了非线性神经网络^[22-23]。

最近,GCN 的出现为基于图结构的 CF 建模带来了新的思路,即通过迭代执行邻域聚合的方式将图中各节点的局部邻居的特征聚合,并在许多基于图结构的任务中取得了巨大成功^[7-8,24]。得益于图卷积思想指导,有些研究将 GCN 应用到了用户物品交互所构成的二部图中,并达到了良好的性能^[6-8]。Pinsage^[6]主要应用于物品推荐,它通过有效的随机游动和图卷积生成节点的嵌入。NGCF^[7]则利用了图中节点的连通性显式聚合高阶邻居的协作信号,并通过多层 GCN 聚合节点表示。LightGCN^[8]是最近提出的一种方法,它认为许多直接从 GCN 方法中

继承的图操作更适用于节点属性更复杂的场景,而对于主要基于用户/物品的编号建模嵌入表示的协同过滤方法可能并不适用。因此,它通过改进不必要的组件进一步优化了 NGCF 的性能。总的来说,基于 GCN 的 CF 方法性能的提高可归功于表示学习的发展,即通过充分利用图的结构显式地学习用户和物品嵌入。

1.2 基于动态图的方法

基于动态图的研究大致可分为 2 个分支,即基于离散动态图 (DTDG) 的研究和基于连续动态图 (CTDG) 的研究。为了学习图节点的演化模式,早期基于 DTDG 的模型一般采用深度自动编码器^[12]或递归层^[25]进行动态图建模。而在最近的研究中, DySAT (dynamic self-attention network)^[16]是与本文提出的方法最相关的工作,该模型利用图结构化注意力模块和时态注意力模块来联合学习节点的动态表示。与基于 DTDG 的模型相比,基于 CTDG 的方法侧重于在每次交互发生时实时执行嵌入更新。例如, Jodie^[26]使用 2 个循环神经网络 (recurrent neural network, RNN) 组件来建模用户和物品的协同演化。 DynGCF^[27]则采用了 0 阶继承、1 阶传播和 2 阶聚合 3 种更新机制来更新相互作用节点的嵌入。然而,这些模型都没有考虑到图建模中存在的过平滑问题,而本文通过提出结合共现图和 TCN 建模缓解了过平滑问题所带来的性能影响。

2 问题定义

本文研究的是基于动态图协同过滤算法 DynGCF。在给定历史用户-物品交互数据的前提下,首先给出了动态图的基本定义,再阐述具体的研究问题。

2.1 动态图定义

动态图可以形式化定义为一系列具有固定时间间隔的离散快照图: $\mathbb{G} = \{G^1, G^2, \dots, G^T\}$, 其中 $G^t = (V^t, E^t)$ 是第 t 时间步的快照图, V^t 和 E^t 分别是对应时刻下的节点集合和边集合。由于实际场景中用户与物品持续交互,本文以增量的方式定义各时态快照图 G^t 。如图 1 所示,在 $t-1$ 时刻出现的用户和物品节点会被加入 t 时刻的快照图中 ($V^{t-1} \subseteq$

V^t), E^t 则表示时刻 $t-1$ 到 t 之间的增量交互边。

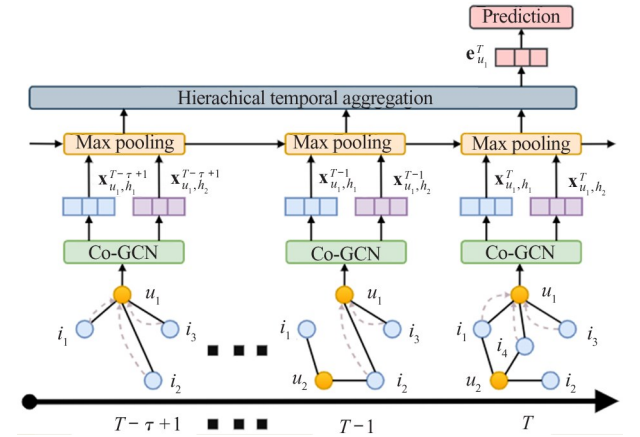


图 1 DynGCF 框架

图 1 所示的框架主要包含 2 个模块。以 u_1 为例,首先利用 Co-GCN 对 u_1 的多时间步局部图结构嵌入进行表示建模,然后经由一个层次化时域聚合器生成最终嵌入 $e_{u_1}^T$ 。图中实线表示关于当前意图具有较高的连通性,虚线则表示连通性较低。

2.2 基于动态图的协同过滤推荐

与基于静态图建模的 CF 方法不同,本文旨在利用多时间步 (如 w 个离散图) 下的历史快照图 ($H = \{G^{T-w+1}, G^{T-w+2}, \dots, G^T\}$) 来建模每个用户/物品在 T 时刻的嵌入 e^T , 并据此预估在 $T+1$ 时刻可能发生的潜在交互。

3 模型

本节首先介绍所提出的模型 DynGCF, 图 1 以用户 u_1 为例说明了节点嵌入的学习过程。本文先设计基于 1 阶交互图和 2 阶共现图的 Co-GCN 算法,以获取每个快照图上的节点局部结构化信息;再应用基于注意力机制的 TCN 模块建模节点的时态演化模式;最后通过节点嵌入内积的方式预估 $T+1$ 时刻的潜在交互。

3.1 结合共现图的局部图表示建模

对任意快照图 $G^t = (V^t, E^t)$, 定义 $A \in \mathbb{R}^{N \times N}$ 是当前时刻用户-物品交互快照图的邻接矩阵,其中节点数量为 N 。一般来说,GCN 延交互边执行多次迭代邻域聚合,将节点的高阶邻居特征聚合到中心节点,从而平滑各个节点的隐表示,产出用户和物品的嵌入表示 $\{e_v \in \mathbb{R}^d, \forall v \in V^t\}$, 其中 d 表示嵌入的

维度。

尽管 GCN 在表示学习领域取得了巨大的成功,但最近的研究表明,GCN 仍有优化的空间:(1) GCN 的特征变换和非线性激活等组件可能不完全适合于 CF,过多的引入复杂交互特征甚至会对模型训练产生负面影响^[11,28]; (2) 多层的邻域聚合会随着迭代次数增加很快达到边际收益,即存在过平滑问题^[15]。受这些已有工作的启发,本文提出了结合 1 阶交互图和 2 阶共现图的图卷积网络 Co-GCN,它同时利用 2 种图结构来学习节点在不同关联关系中的结构化表示,从而缓解过平滑问题并更好地学习节点嵌入。

3.1.1 图卷积优化

为优化图卷积以适应推荐等基于二部图的图表示建模场景,现有的方法通过消除非线性激活和多层感知机 (multi-layer perceptron, MLP), 以使其具有可解释性和可训练性。结构化图卷积网络 (structured graph convolutional network, SGCN) 通过下述方式形式化了图卷积网络。

$$\begin{aligned} \mathbf{E}^{(K)} &= (\mathbf{A} + \mathbf{I})\mathbf{E}^{(K-1)} = (\mathbf{A} + \mathbf{I})^K \mathbf{E}^{(0)} \\ &= \left(\frac{K}{0}\right)\mathbf{E}^{(0)} + \left(\frac{K}{1}\right)\mathbf{A}\mathbf{E}^{(0)} + \cdots + \left(\frac{K}{K}\right)\mathbf{A}^K \mathbf{E}^{(0)} \end{aligned} \quad (1)$$

其中, $\mathbf{E}^{(K)}$ 是经过 K 次聚合后的嵌入矩阵, $\mathbf{I} \in \mathbb{R}^{N \times N}$ 表示单位矩阵。上式的推导表明,最终嵌入表示等价于对每次卷积得到的各层嵌入作加权和。LightGCN 通过去除二项式系数 $\left(\frac{K}{i}\right)$ 并采用平均化得到最终的嵌入,获得了良好的性能。尽管如此,SGCN 和 LightGCN 仍然存在过平滑问题,这意味着随着深度 K 变大时,节点表示的建模会引入更多来自远距离邻居的“噪声”。另一方面,许多现有的研究^[10-11,14]表明,当图卷积的深度大于 3 时,性能的提升便达到了瓶颈,甚至会有所降低。此外,通过将式(1)中的 K 置为 3,显然可得出 1 阶和 2 阶的邻居对最终嵌入贡献最大的结论。

因此,不再基于单个交互图执行迭代聚合,而是基于 1 阶的交互图和 2 阶的共现图中探索更多潜在的图连接关系。有别于交互图,共现图可以更好地建模具有相同偏好的用户或物品之间的 2 阶相似特

征。更具体地说,共现图主要包括 2 个子图:(1) 用户侧子图,它关联了具有相似偏好的用户;(2) 物品侧子图,它连接具有潜在共性特征的物品。基于交互图的邻接矩阵 $\mathbf{A}$ 得到了式(2)所示的共现图邻接矩阵的形式化表示。

$$\mathbf{A}_{\text{co}} = \text{MASK}(\mathbf{A}^2) \quad (2)$$

其中, $\mathbf{A}_{\text{co}}$ 表示当前快照下共现图的邻接矩阵。此外,考虑到用户对物品的访问是有序的,使用一个 MASK 函数来确保物品侧子图的邻接关系是有向的,以防止特征穿越。最后对节点嵌入的建模可以抽象化为

$$\mathbf{E} = g(\mathbf{E}^{(0)}, \mathbf{A}, \mathbf{A}_{\text{co}}) \quad (3)$$

基于初始化嵌入矩阵 $\mathbf{E}^{(0)} \in \mathbb{R}^{N \times d}$, 1 阶交互邻接矩阵 $\mathbf{A}$ 和 2 阶共现邻接矩阵 $\mathbf{A}_{\text{co}}$, 通过使用合适的网络结构 $g(\cdot)$ 来得到最终嵌入矩阵 $\mathbf{E}$ 。

3.1.2 结构化聚合模块

如前所述,本文的目标是通过减少图聚合操作和利用来自 1 阶交互图和 2 阶共现图中丰富的协同信号来缓解过平滑问题。为此,设计了一个结合 2 阶共现图的图卷积 (Co-GCN) 模块,以形式化地提取并融合从 2 个图中所学得的嵌入特征,如式(4)所示。

$$\mathbf{E}_t = g(\mathbf{E}_t^{(0)}, f_{\text{GCN}}(\mathbf{G}_{\text{inter}}^t, \mathbf{E}_t^{(0)}), f_{\text{GCN}}(\mathbf{G}_{\text{co}}^t, \mathbf{E}_t^{(0)})) \quad (4)$$

其中, $\mathbf{G}_{\text{inter}}^t$ 和 $\mathbf{G}_{\text{co}}^t$ 分别代表了交互图和共现图在时间步 t 下的图结构信息 (节点集合与边集合)。以交互图为例,通过 $f_{\text{GCN}}(\mathbf{G}_{\text{inter}}^t, \mathbf{E}_t^{(0)}) = \tilde{\mathbf{A}}^t \mathbf{E}_t^{(0)}$ 执行单步图卷积操作聚合每个图中的节点特征,生成各节点所对应的 d 维嵌入,其中 $\tilde{\mathbf{A}}^t = \mathbf{D}_t^{-\frac{1}{2}} \mathbf{A}^t \mathbf{D}_t^{-\frac{1}{2}}$, $\mathbf{D}_t \in \mathbb{R}^{N \times N}$ 是 $\mathbf{A}^{(t)}$ 的对角阵。 $g(\cdot)$ 表示基于不同子图的特征聚合网络。接下来,通过执行如图 2 所示聚合操作产出子图在时间步 t 下的节点嵌入矩阵 $\mathbf{E}_t$ 具体操作可形式化为

$$\mathbf{x}_u^t = (\mathbf{x}_{u,h_0}^t \parallel \text{MaxPooling}(\mathbf{x}_{u,h_1}^t, \mathbf{x}_{u,h_2}^t)) \mathbf{W} + \mathbf{b} \quad (5)$$

其中, $\mathbf{x}_{u,h_0}^t$ 为用户 u 卷积前的初始化嵌入向量, $\mathbf{x}_{u,h_1}^t$ 和 $\mathbf{x}_{u,h_2}^t$ 分别为从交互图和共现图中学得的节点嵌入, $\mathbf{W} \in \mathbb{R}^{2d \times d}$ 表示特征变换矩阵, $\mathbf{b}$ 为偏置向量,

而 $\parallel$ 则指代向量拼接操作。首先通过一个 MaxPooling 模块融合 2 个子图所产出的嵌入,然后对拼接后的表示进行线性变换,生成最终的快照图结构化嵌入。

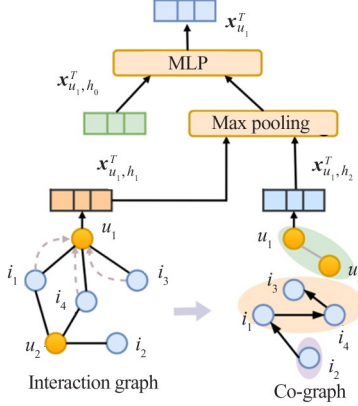


图2 Co-GCN 图结构化聚合框架

3.2 层次化时态嵌入聚合层

本文设计了一个层次化时态嵌入传播层来捕捉用户-物品动态交互图中的特征演化。对每个节点 v , 将其在结构化聚合模块中得到的多时间步下的嵌入定义如下 $\{z_v^{t-w+1}, \dots, z_v^t\}, z_v^k \in \mathbb{R}^d$, 其中, w 是窗口大小, z_v^k 是节点 v 从第 k 个快照图中所学到的嵌入表示。先应用 TCN 来捕获用户(物品)的阶段兴趣(受欢迎度);再叠加一个注意力层对每个用户或物品的多阶段嵌入表示执行加权融合,产出最终的嵌入表示。

3.2.1 时域卷积网络

TCN 是一维 CNN 的一种特殊变体,常用于编码序列信息。相对于原始的 CNN,TCN 主要做了 2 个重要的改进,即因果卷积和空洞卷积。因果卷积通过将时刻 t 之后的历史嵌入掩藏,以保证在卷积过程中不会出现特征穿越。此外,考虑到一维 CNN 的序列建模性能受其有限的感受野影响,TCN 采用多层的空洞卷积核聚合节点的多阶段历史嵌入。具体而言,第 r 层 TCN 的空洞卷积可以写成:

$$z_v^t = F(x_v^t) = (X * l_r f)(t) = \sum_{i=0}^{k-1} f(i) \cdot x_v^{t-l_r i} \quad (6)$$

其中, f 是滤波器, X 表示输入序列的嵌入矩阵, $*$ 表示卷积运算, $l_r = (k-1)^{r-1}$ 为滤波器的空洞卷积系数。不同于基于循环神经网络和自注意力机制,

多层的 TCN 模块使得模型能依指数增长地聚合历史的嵌入特征,且在不丢失长期历史信息的时候还能更多地捕获近期历史特征。因此,它更适用于需要同时兼顾短期时效性和长期记忆性的动态图建模场景。

3.2.2 多时域自注意力聚合层

为了进一步捕获时态演化模式,在 TCN 之上堆叠了一个自注意层(self-attention layer)。该层的输入是经 TCN 层输出的一系列时序嵌入: $\{z_v^{t-w+1}, \dots, z_v^t\}, z_v^k \in \mathbb{R}^d$ 。为建模直到时间步 t 为止的节点演化模式,依注意力权重对多个时序嵌入进行聚合:

$$e_v^t = \sum_{i=0}^{w-i} \pi(i) z_v^{t-i} \quad (7)$$

其中 $\pi(i)$ 是用于控制加权聚合的权重系数,通过式(8)计算注意力权重。

$$\pi(i) = \tanh(z_v^{t-iT} z_v^t) \quad (8)$$

其中,使用 $\tanh$ 作为非线性激活函数。以用户为例,上述的等式根据用户的历史兴趣表示 z_v^{t-i} 和当前阶段表示 z_v^t 的关系计算重要性权重参数。为了表述简洁,仅使用内积的方式计算权重,而不将其作为本文讨论的重点。

此后,通过应用 softmax 函数对上述注意力得分进行归一化,可得归一后注意力权重为

$$\pi(i) = \frac{\exp(\pi(i))}{\sum_{j=0}^{w-1} \exp(\pi(j))} \quad (9)$$

最终,通过从历史的快照图中依注意力权重分层聚合各阶段的节点结构化特征,可以获得用户和物品的最终表示。算法 1 给出了基于离散动态图的协同过滤算法伪代码。

算法 1 基于离散动态图的协同过滤推荐算法

Input: $H = \{G^{T-w+1}, \dots, G^T\}, \{E_{T-w+1}^{(0)}, \dots, E_T^{(0)}\}$

Output: E^{T+1}

1: Initialize $\{E_{T-w+1}^{(0)}, \dots, E_T^{(0)}\}, E_{\text{list}} = \{\};$

2: **for** $t = T - w + 1$ to T **do**

3: Initialize $G_{\text{inter}}^t, G_{\text{co}}^t;$

4: $E_t = \text{Co-GCN}(E_t^{(0)}, G_{\text{inter}}^t, G_{\text{co}}^t);$

5: $E_{\text{list}} \leftarrow \text{append}(E_t)$

6: **end for**

7: $Z = \text{TCN}(E_{\text{list}}); \quad // Z = \{Z_{T-w+1}, \dots, Z_T\}, Z_t \in \mathbb{R}^{N \times d}$

8: $E^{T+1} = \text{Self-Att}(Z);$

3.3 模型预测

为了预测潜在的交互,使用内积来估计用户对目标物品的偏好:

$$\hat{y}(u, i) = \mathbf{e}_u^t \mathbf{e}_i^t \quad (10)$$

其中 $\mathbf{e}_u^t$ 和 $\mathbf{e}_i^t$ 分别为用户 u 和物品 i 在时间步 t 的嵌入。为了学习可训练的参数,使用贝叶斯个性化排序(Bayesian personalized ranking, BRP)目标函数^[21]对所提出的模型进行优化。目标函数为

$$L_{\text{BPR}} = - \sum_{(u, i, j) \in D} \ln \sigma(\hat{y}_{ui} - \hat{y}_{uj}) + \lambda (\|\mathbf{E}\|^2 + \|\mathbf{W}\|^2 + \|\Theta\|^2) \quad (11)$$

其中, $D = \{(u, i, j) | (u, i) \in D^+, (u, j) \in D^-\}$ 表示训练所用的数据集, D^+ 和 D^- 分别是已观测和未观测到的样本。 $\mathbf{E}$ 和 $\mathbf{W}$ 分别为嵌入矩阵和网络的权重参数, Θ 表示所有可训练参数, λ 为惩罚系数。采用 Adam^[29] 优化器在学习过程中自动调整学习率。

4 实验与结果分析

本节首先阐述实验设置,然后在 4 个基准数据集上对 DynGCF 进行详细的评估实验,并给出相关分析。

4.1 实验数据集

在 4 个不同规模、领域和密度的真实数据集上评估本文提出的模型: Amazon-Beauty^[30]、ML-25M^[31]、Gowalla、Yelp^[32]。Amazon-Beauty 来自亚马逊评论数据集,包含大量用户-物品交互数据,例如用户评论和评分。Yelp 则是一个用户关于地方企业(如餐馆和酒吧)偏好的评分数据集。Gowalla 是一个基于位置的社交网站,本文使用用户的签住信息构造训练数据。MovieLens 是一个广泛应用于评估协同过滤算法的基准数据集。本文使用了 ML-25M 中的标签数据集,它包含用户对电影的标签信息。

由于每个数据集的数据分布不同,将它们按不同的时间间隔分割成一系列的离散快照图;选用最近的 w 个快照图进行训练,并将最后一个快照上的节点交互样本作为测试集;此外,应用负采样策略,

将观测到的用户物品对作为正例,将未观测到的交互作为负例进行训练。经预处理后的数据集的统计信息如表 1 所示。

表 1 4 个数据集的统计情况

数据集	用户数量	物品数量	交互数量	稠密度
Amazon-Beauty	6079	11 448	35 755	0.051%
Yelp	10 371	14 821	404 422	0.263%
Gowalla	16 278	6150	538 113	0.538%
ML-25M	2255	8540	243 740	1.266%

4.2 评价指标

采用了 2 个经典的 Recall@K(召回)和 NDCG@K(排序)指标来评估模型的 top-K 推荐性能。对于每个用户,Recall@K 推荐列表的前 K 个物品中实际被点击的数量占有所有标签数量的比例。NDCG@K 是归一化折损累计增益(NDCG)。默认情况下,设置 $K=20$,并汇报测试集中所有用户 top-K 指标的平均值。

4.3 对比模型

将 DynGCF 与现有最先进的方法进行了比较,包括基于 CF 的方法(BPRMF)、基于 GCN 的方法(PinSage、NGCF、LightGCN)和基于动态图的方法(DynAERNN、DySAT、Jodie、DGCF)。

(1) BPRMF^[21]:这是一种经典的矩阵分解方法,通过贝叶斯个性化排名(BPR)优化,以模拟用户物品交互。

(2) PinSage^[9]:该模型采用多参数共享图卷积模块聚合特征信息,使得参数复杂度与输入图大小无关。

(3) NGCF^[10]:该方法在用户物品交互图上采用 3 个 GCN 层来显式聚合来自多跳邻居的协作信号。

(4) LightGCN^[11]:该模型在 CF 任务中应用了 GCN 建模图结构,但去除了非必须的特征变换和非线性激活模块,从而获得更好的性能。

(5) DynAERNN^[25]:一种由密集自动编码器与递归层组成的深度神经网络,用于捕获时间图演化。

(6) DySAT^[16]:该模型应用多头自注意机制来学习离散动态图上的结构和时间特征。

(7) Jodie^[26]:它是基于协同演化模型的代表方法,使用用户 RNN、物品 RNN 以及投影操作来建模用户和物品的嵌入。

(8) DGCF^[27]:这是一种目前效果最佳的基于连续动态建模的方法,它采用 3 种更新机制,即 0 阶继承、1 阶传播和 2 阶聚合,对 1 跳用户项交互进行建模,并同时更新用户和物品节点。本文将 GAT 聚合器实现为 2 阶聚合函数。

4.4 实验设置

本节所述方法及基线模型的嵌入维度统一设置为 64。对于应用图形卷积的所有基线方法,将卷积层深度设置为 3。此外,训练批大小设置为 1024,并

使用 Adam 算法优化所有基线方法。在超参数方面,应用网格搜索,在 $\{0.0001, 0.0005, 0.001, 0.01\}$ 中搜索最佳的 L_2 系数,并在 $\{0.0001, 0.0005, 0.001\}$ 中调整学习率,并应用 Xavier 初始化器初始化所有模型的模型参数。特别地,本文在 $\{0, 0.1, \dots, 0.7, 0.8\}$ 中搜索 NGCF 的丢弃(dropout)比率。

4.5 性能比较

为了展示 DynGCF 的整体性能,评估了其与基线方法在 Recall(R)@20 和 NDCG(N)@20 上的指标评分。令 $w = 8$, 总体结果如表 2 所示。其中,最佳的结果以黑体标识,次优的结果则用下划线标识。通过分析比较,得出了以下观测结果。

表 2 DynGCF 方法在 4 个数据集上与基准方法的对比实验结果

方法	Amazon-Beauty		Yelp		Gowalla		ML-25M	
	R@ 20	N@ 20	R@ 20	N@ 20	R@ 20	N@ 20	R@ 20	N@ 20
BPRMF	0.0316	0.0232	0.0298	0.0352	0.1232	0.1461	0.0323	0.0885
PinSage	0.0376	0.0251	0.0353	0.0408	0.1280	0.1533	0.0348	0.0737
NGCF	0.0448	0.0276	0.0379	0.0419	0.1606	0.1538	0.0372	0.0863
LightGCN	<u>0.0483</u>	<u>0.0287</u>	<u>0.0417</u>	<u>0.0468</u>	<u>0.1762</u>	0.1587	0.0458	0.0939
DynAERNN	0.0366	0.0206	0.0274	0.0352	0.1351	0.1517	0.0234	0.0711
DySAT	0.0454	0.0273	0.0371	0.0423	0.1689	0.1637	<u>0.0572</u>	<u>0.1058</u>
Jodie	0.0408	0.0229	0.0332	0.0397	0.1538	0.1465	0.0365	0.0874
DGCF	0.0452	0.0265	0.0392	0.0447	0.1717	<u>0.1642</u>	0.0543	0.0982
DynGCF	0.0529	0.0308	0.0455	0.0513	0.1968	0.1844	0.0615	0.1123
涨幅	9.52%	7.32%	9.11%	9.62%	11.69%	12.30%	7.52%	6.14%

(1) 本文提出的模型 DynGCF 在 4 个数据集上取得了显著的改进,这说明了本文提出模型的优越性。

(2) BPRMF 在 4 个数据集上表现不佳,这表明仅使用内积来建模并不足以建模用户和物品之间的关系。

(3) 基于物品的连续动态图方法的方法(Jodie, DGCF)与基于离散动态图的方法在 Amazon-Beauty 和 Yelp 上性能接近,但在 Gowalla 和 ML-25M 上表现不佳。一个可能的原因是用户的偏好在频繁的交互中会更快地发生迁移;而 DGCF 中的 2 阶聚合操作会反复引入历史特征,因此难以及时捕获时态变化信息。

(4) DynGCF 始终优于基于 GCN 的静态方法。

这进一步证明了利用动态图对演化模式建模的关键重要性。

(5) 相比于 DySAT, DynGCF 表现出更优异的性能。这主要是由于过平滑问题导致 DySAT 在建模用户物品交互的过程中会引入噪声。与现有基于动态图的模型不同, DynGCF 具有以下 2 个优点:其一是 DynGCF 同时利用交互图中的 1 跳协作信号和共现图中的 2 跳相似性进行局部图结构特征学习,更适用于建模推荐场景中普遍存在的用户-物品二部交互图;其二是 DynGCF 应用了 TCN 能更好地建模用户和物品的阶段性特征。

4.6 消融实验

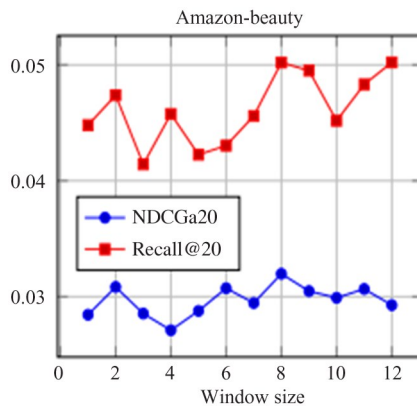
接下来,进行了一组消融实验,在 Gowalla 和 ML-25M 2 个数据集上验证 Co-GCN 和时域聚合层

的有效性。表 3 列举出了消融实验在 Recall@20 和 NDCG@20 上的结果,以体现每个模块对 DynGCF 性能的贡献。方法(1)禁用了 TCN 模块,仅使用自注意层来聚合多阶段嵌入;而方法(2)与(1)相反,禁用了自注意层,仅使用 TCN 模块;方法(3)、(4)、(5)和(6)使用原始的图卷积网络代替 Co-GCN,并将深度分别设置为 1、2、3、4;方法(7)则是用于比较的完整 DynGCF 模型。根据表 3,可以得出以下结论。

首先,通过比较方法(1)、(2)和(6)的结果,可

表 3 消融对比实验结果

方法	Gowalla		ML-25M	
	R@ 20	N@ 20	R@ 20	N@ 20
(1) DynGCF-TCN	0.1833	0.154	0.0597	0.1207
(2) DynGCF-ATT	0.1522	0.141	0.0684	0.1305
(3) DynGCF-1	0.1404	0.1306	0.0642	0.1058
(4) DynGCF-2	0.1493	0.1366	0.0649	0.1190
(5) DynGCF-3	0.1482	0.1407	0.0715	0.1248
(6) DynGCF-4	0.1476	0.1405	0.0653	0.1222
(7) DynGCF	0.1919	0.169	0.0697	0.1313



以看出 TCN 模块和自注意层共同提高了模型建模时态演化的能力。其次,方法(1)和(2)的结果说明了,虽然 TCN 效果显著,但仅使用 TCN 还不足以充分捕获时态信息,自注意层可以在聚合过程中自适应地控制节点在各时间步上历史嵌入的聚合权重。然后,方法(3)~(6)的结果说明了仅靠增加卷积的层数并不能提升推荐效果,即过平滑问题。最后,通过比较方法(5)和(7)的结果,验证了 DynGCF 的有效性,它相比直接基于 GCN 的方法能取得更佳的性能。

4.7 离散快照图数量对性能的影响

作为一个重要超参,是否能恰当地选取离散快照图窗口大小 w 决定了 DynGCF 的性能上限。本节在 Amazon-Beauty 和 ML-25M 数据集上验证了快照图数量的影响,结果如图 3 所示。当窗口大小较短时,推荐的指标波动相对较剧烈;而随着窗口大小的增加,模型的性能趋于稳定。一个可能的原因是,较长的历史交互可以帮助 DynGCF 捕获更多来自用户和物品的多阶段特征。

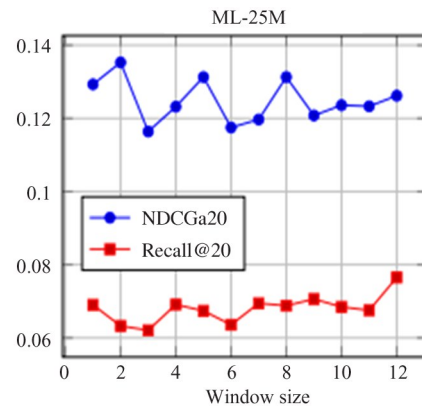


图 3 离散快照数量对模型性能的影响

4.8 对过平滑问题的缓解

为了评估对过平滑问题的缓解效果,采用 MAD (mean average distance) 作为评价指标^[15],它通过计算任意节点对之间的全局平均距离来评估图节点嵌入的平滑度。直观地说,较小的 MAD 值意味着存在更严重的过平滑问题。将 DynGCF 的 MAD 值与以下 4 种基于静态图的模型 (Co-GCN、Pinsage、NGCF 及 LightGCN) 进行比较,以验证其性能。统一

将图卷积深度固定为 3,结果如图 4 所示。从图 4 可以得到以下结论。

(1) Co-GCN 比基于 GCN 的 CF 方法获得更高的 MAD 值,这证明了 Co-GCN 在缓解过度平滑问题方面的有效性。

(2) 在基于 GCN 的模型中,LightGCN 的 MAD 值最低。一个主要原因是,LightGCN 为了简化图卷积操作没有使用非线性激活函数,而这种完全线性

的聚合方式在提高性能的同时也间接使得各节点更加平滑,节点的嵌入更加趋同。

(3) DynGCF 的 MAD 值较之 Co-GCN 更高,这说明 DynGCF 本身的动态建模架构也有助于缓解过

度平滑问题。这主要是由于 DynGCF 的时域聚合模块是从多阶段的快照图中分别学习的嵌入表示,这相较于单个静态图的建模方式更加稳定。

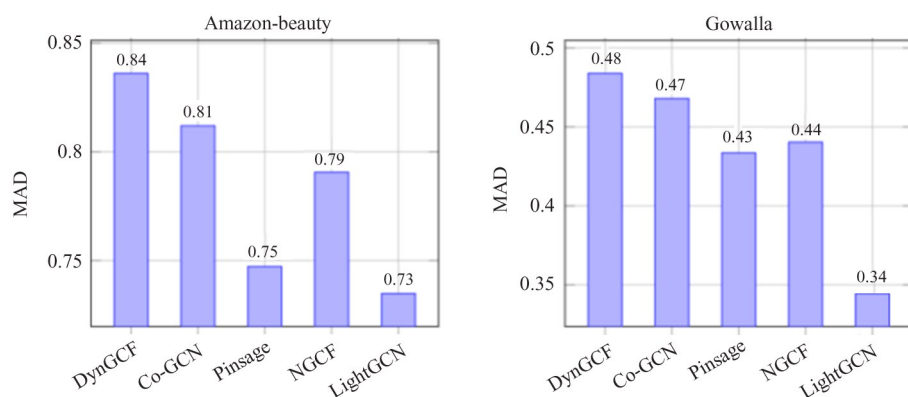


图4 DynGCF 与基线模型在 Amazon-Beauty 和 Gowalla 上的 MAD 值

5 结 论

本文研究了推荐场景中基于动态图的协同过滤方法并应用于用户和物品的交互建模中;设计了一个新的模型 DynGCF,它可以有效地从历史快照图中捕获节点的局部图结构和时态演化信息。具体地,本文首先提出了 Co-GCN 来捕获图结构特征,然后进一步利用 TCN 和自注意机制来构建一个层次化时态嵌入聚合层,以模拟节点的演化模式。综合实验结果表明,该模型的性能明显优于现有基线方法,且可以缓解过度平滑问题。

参考文献

- [1] 蔺丰奇,刘益. 信息过载问题研究述评[J]. 情报理论与实践, 2007, 305:710-714.
- [2] 张玉洁,董政,孟祥武. 个性化广告推荐系统及其应用研究[J]. 计算机学报, 2021, 3(44): 531-563.
- [3] 梁仕威,张晨蕊,曹雷,等. 基于协同表示学习的个性化新闻推荐[J]. 中文信息学报, 2018, 32(11): 72-78.
- [4] 王娜,何晓明,刘志强,等. 一种基于用户播放行为序列的个性化视频推荐策略[J]. 计算机学报, 2020, 43(1):123-135
- [5] SARWAR B, KARYPIS G, KONSTAN J, et al. Item-

based collaborative filtering recommendation algorithms [C] // Proceedings of the 10th International Conference on World Wide Web. Hong Kong: Association for Computing Machinery, 2001:285-295.

- [6] 徐冰冰,岑科廷,黄俊杰,等. 图卷积神经网络综述[J]. 计算机学报, 2020, 43(5):26.
- [7] HAMILTON W, YING Z T, LESKOVEC J. Inductive representation learning on large graphs [C] // Proceedings of the 31st Conference on Neural Information Processing Systems. Long Beach: Curran Associates Inc., 2017:1024-1034.
- [8] KIPF T N, WELING M. Semi-supervised classification with graph convolutional networks [C] // The 5th International Conference on Learning Representations. Toulon: ICLR, 2017:3294-3302.
- [9] YING R, HE R N, CHEN K F, et al. Graph convolutional neural networks for web-scale recommender systems [C] // Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. London: Association for Computing Machinery, 2018: 974-983.
- [10] WAN X, HE X N, WANG M, et al. Neural graph collaborative filtering [C] // Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval. Paris: Association for Computing Machinery, 2019:165-174.

- [11] HE X N, DENG K, WANG X, et al. LightGCN:simplifying and powering graph convolution network for recommendation[C] // Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval. Virtual; Association for Computing Machinery, 2020:639-648.
- [12] GOYAL P, KAMRA N, HE X, et al. DynGEM: deep embedding method for dynamic graphs[EB/OL]. (2018-05-29) [2022-03-20]. <https://arxiv.org/pdf/1805.11273.pdf>.
- [13] MA Y, GUO Z, REN Z, et al. Streaming graph neural networks[C]//Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval. Virtual; Association for Computing Machinery,2020:719-728.
- [14] WANG X, HE X, CAO Y, et al. KGAT: knowledge graph attention network for recommendation[C] // Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Diego; Association for Computing Machinery, 2019: 950-958.
- [15] CHEN D, LIN Y, LI W, et al. Measuring and relieving the over-smoothing problem for graph neural networks from the topological view[C] // Proceedings of the 34th AAAI Conference on Artificial Intelligence. New York; AAAI Press, 2020: 3438-3445.
- [16] SANKAR A, WU Y H, GOU L, et al. DySAT:deep neural representation learning on dynamic graphs via self-attention networks[C] //The 13th ACM International Conference on Web Search and Data Mining. Houston; Association for Computing Machinery, 2020: 519-527.
- [17] PAREJA A, DOMENICONI G, CHEN J, et al. Evolve GCN: evolving graph convolutional networks for dynamic graphs[C] //The 34th AAAI Conference on Artificial Intelligence. New York; AAAI Press, 2020: 5363-5370.
- [18] LEA C N, FLYNN M D, VIDAL R, et al. Hager temporal convolutional networks for action segmentation and detection[C] //2017 IEEE Conference on Computer Vision and Pattern Recognition. Hawaii; IEEE, 2017: 1003-1012.
- [19] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C] //The 31st Annual Conference on Neural Information Processing Systems. Long Beach; Curran Associates Inc. , 2017: 5998-6008.
- [20] KOREN Y, BELLR. Matrix factorization techniques for recommender systems[J]. Computer, 2009, 42(8):30-37.
- [21] RENDLE S, FREUDENTHALER C, GANTNER Z, et al. BPR:Bayesian personalized ranking from implicit feedback [C]// Proceedings of the 25th Conference on Uncertainty in Artificial Intelligence. Montreal; AUAI Press, 2009: 452-461.
- [22] HE X N, LIAO L, ZHANG H, et al. Neural collaborative filtering[C] //Proceedings of the 26th International Conference on World Wide Web. Perth; IEEE, 2017: 173-182.
- [23] DENG Z, HUANG L, WANG C, et al. DeepCF: a unified framework of representation learning and matching function learning in recommender system[C] // Proceedings of the 33th AAAI Conference on Artificial Intelligence. Hawaii; AAAI Press, 2019: 61-68.
- [24] VELICKOVIC P, CUCURULL G, CASANOVA A, et al. Graph attention networks[C] //The 6th International Conference on Learning Representations. Vancouver; ICLR, 2018(Poster).
- [25] GOYAL P, CHHETRI S R. Dyngraph2vec: capturing network dynamics using dynamic graph representation learning[J]. Knowledge Based Systems, 2020, 187: 104816.
- [26] KUMAR S, ZHANG X, LESKOVEC J. Predicting dynamic embedding trajectory in temporal interaction networks[C] //Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Diego; Association for Computing Machinery, 2019: 1269-1278.
- [27] LI X, ZHANG M, WU S, et al. Dynamic graph collaborative filtering[C] //2020 IEEE International Conference on Data Mining. Sorrento; IEEE, 2020: 322-331.
- [28] WU FELIX, AMAURI H, JR S, et al. Simplifying graph convolutional networks[C] //Proceedings of the 36th International Conferenceon Machine Learning. Long Beach; ICML,2019:6861-6871.
- [29] COVINGTON P, ADAMS J, SARGIN E. Deep neural networks for you tube recommendations[C] // Proceedings of the 10th ACM Conference on Recommender Systems. Boston; Association for Computing Machinery,

- 2016; 191-198.
- [30] HE R N, MCAULEY J. Ups and downs: modeling the visual evolution of fashion trends with one-class collaborative filtering[C] // Proceedings of the 25th International Conference on World Wide Web. Montreal: Association for Computing Machinery, 2016;507-517.
- [31] MAXWELL H F, JOSEPH A K. The MovieLens datasets: history and context[J]. ACM Transactions on Interactive Intelligent Systems, 2015, 19;1-19.
- [32] LUCA M. Reviews, reputation, and revenue: the case of yelp.com[R]. Harvard: Harvard Business School NOM Unit Working Paper, 2016.

Discrete temporal dynamic graph based collaborative filtering

JIN Jiaqi^{* **}, ZHANG Mengfei^{***}, PAN Mao^{* **}, ZHU Zhihai^{****}, FANG Jinyun^{*}

(^{*} Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190)

(^{**} University of Chinese Academy of Sciences, Beijing 100190)

(^{***} The National Computer Network Emergency Response Technical Team Coordination Center of China, Beijing 100029)

(^{****} China Xiongan Group, Beijing 071700)

Abstract

The graph convolution network(GCN) based models have shown powerful performance in learning the users' / items' representations and achieved new state-of-the-art for collaborative filtering(CF). Nevertheless, existing GCN based CF methods still have following limitations: They are all target static graphs while many real-life graphs evolve over time. Existing work that adapts GCN to recommender systems suffers from performance limitations due to the over-smoothing issue. To tackle the aforementioned problems, a dynamic graph based collaborative filtering (DynGCF) is proposed, which aims to learn the representations of users and items by capturing both graph structural and temporal information. Specifically, DynGCF adapts GCN to learn discrete user/item embeddings on each graph snapshot at first, then employs temporal convolutional networks(TCN) and self-attention mechanism to learn the final embeddings. To alleviate the over-smoothing issue, we analyze and simplify the neighborhood aggregation, which is a pivot component in GCN, by jointly using only 1-hop interaction and 2-hop co-occurrence graph to model the user-item interactions. Extensive experiments are conducted on four real-world datasets to demonstrate the significant performance gains for DynGCF over state-of-the-art GCN based CF methods and dynamic graph based methods. Further analysis proves that the alleviation of the over-smoothing benefits from the hop graphs.

Key words: recommender system, dynamic graph, graph convolution network (GCN), collaborative filtering (CF)