

基于加权聚合图神经网络的音乐推荐模型

肖涵, 孙德志, 游福成

北京印刷学院, 信息工程学院, 北京

收稿日期: 2024年1月18日; 录用日期: 2024年2月21日; 发布日期: 2024年2月29日

摘要

现有基于图神经网络的音乐推荐模型在聚合方式上忽略了节点自身与邻域的权重分配, 多数模型只是简单的相加或拼接, 这种聚合方式没有充分挖掘节点特征的差异性, 针对以上问题, 本文提出了基于加权聚合图神经网络的音乐推荐模型MGWA, 首先根据用户项目交互的行为记录来构建知识图, 然后图神经网络通过逐层聚合邻域信息来丰富项目表示和传播高阶信息, 特别的, 在聚合阶段, 我们对节点自身和其邻域采用门控机制进行权重分配, 从而更好地捕捉节点特征表示。此外, 门控机制可以根据节点自身及其邻域的特征动态地分配权重, 这使得聚合方式更加灵活和可学习, 从而更好地适应不同的图结构和节点特征。最后, 本文评估了所提出模型在真实数据集上的性能, 与常见推荐模型进行对比, 实验结果证明了我们的模型在AUC、F1等指标上均取得了提升。

关键词

音乐推荐, 图神经网络, 加权聚合

Music Recommendation Model Based on Graph Neural Network with Weighted Aggregation

Han Xiao, Dezhi Sun, Fucheng You

School of Information Engineering, Beijing Institute of Graphic Communication, Beijing

Received: Jan. 18th, 2024; accepted: Feb. 21st, 2024; published: Feb. 29th, 2024

Abstract

The existing music recommendation model based on graph neural network ignores the weight distribution between the node itself and the neighborhood, and most models simply add or splice, which does not fully exploit the differences of node characteristics. To solve the above problems, this paper proposes a music recommendation model (MGWA) based on graph neural network with weighted

aggregation. Firstly, a knowledge graph is constructed according to the behavior records of user interaction. Then graph neural network enriches the project representation and spreads high-order information by aggregating neighborhood information layer by layer. Especially, in the aggregation stage, we use gating mechanism to assign weights to nodes themselves and their neighborhoods, so as to better capture the node feature representation. In addition, the gating mechanism can dynamically allocate weights according to the characteristics of nodes themselves and their neighbors, which makes the aggregation method more flexible and learnable, thus better adapting to different graph structures and node characteristics. Finally, this paper evaluates the performance of the proposed model on real data sets, and compares it with the common recommended models. The experimental results prove that our model has achieved improvement in AUC, F1 and other indicators.

Keywords

Music Recommendation, Graph Neural Networks, Weighted Aggregation

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

音乐推荐系统作为电子商务和娱乐产业中的重要应用，一直备受关注。随着互联网和移动设备的普及，用户对个性化、精准的音乐推荐需求日益增加。传统的音乐推荐系统主要基于内容的推荐和协同过滤等方法[1]，这些方法在一定程度上能够为用户提供推荐，但往往忽视了用户和项目之间的复杂关系，难以充分挖掘用户和项目的特征信息。基于内容的方法主要依靠音乐本身的特征，如歌手、流派、风格等来进行推荐，而协同过滤方法则是基于用户行为数据进行推荐，如用户对音乐的播放记录、评分等。然而，这些方法存在一些局限性，例如内容过滤方法容易受到冷启动问题的影响，即新音乐或新用户难以得到准确的推荐；而协同过滤方法则容易出现数据稀疏和推荐热点等问题，影响了推荐的准确性和个性化程度。

为了克服传统音乐推荐系统的局限性，近年来基于深度学习和图神经网络的音乐推荐模型逐渐成为研究热点[2] [3] [4]。这些模型通过建模用户和项目之间的关系图，能够更好地捕捉其特征表示，提高推荐的准确性和个性化程度。图神经网络能够有效地处理图结构数据[5]，将节点和边的信息进行融合，从而更好地挖掘节点之间的关联和特征表示。在音乐推荐领域，图神经网络可以将用户、音乐项目以及它们之间的交互关系表示为一个图，进而进行特征学习和推荐。

然而，当前基于图神经网络的音乐推荐模型在聚合方式上存在一定的局限性，多数模型只是简单地相加或拼接节点特征[6] [7]，忽略了节点自身与邻域的权重分配，导致聚合方式对节点特征的差异性挖掘不足。因此，本文提出了一种基于加权聚合图神经网络的音乐推荐模型 MGWA，通过引入门控机制动态地分配节点自身及其邻域的权重，更好地捕捉节点特征表示。MGWA 模型能够在聚合阶段灵活地适应不同的图结构和节点特征，从而提高推荐的准确性和个性化程度。实验结果表明，我们提出的模型在真实数据集上取得了显著的性能提升，证明了其在音乐推荐领域的潜在应用价值。

本文的其余部分结构安排如下：第 2 节介绍了相关工作；第 3 节详细阐述了 MGWA 模型的设计与实现；第 4 节给出了实验结果和性能评估；最后，第 5 节对全文进行了总结和展望。

2. 相关工作

图神经网络在推荐系统中的应用是一个备受关注的研究领域，先前的研究已经证明了图神经网络在

推荐系统中具有很大的潜力[4]。图神经网络能够更好地捕捉用户和项目之间的复杂关系，从而提高推荐的准确性和个性化程度。图神经网络通过利用图结构来建模用户和项目之间的关系，其中图的节点表示用户或项目，边表示它们之间的交互或关联。通过这种方式，图神经网络可以更好地捕捉用户和项目之间的复杂关系，包括用户之间的相似性、项目之间的相似性以及用户和项目之间的交互模式。

在图神经网络中，节点表示学习是一个核心问题，它涉及如何将用户和项目表示为低维向量，以便进行推荐。先前的研究表明，图神经网络可以通过学习节点表示来实现更准确的推荐。这通常涉及将图神经网络应用于表示学习任务，通过在图上进行信息传播和聚合来丰富节点的表示，从而捕捉用户和项目之间的复杂关系。利用图神经网络来学习用户和项目的表示，从而实现个性化的推荐[8]。通过在图上进行信息传播和聚合，图神经网络可以更好地捕捉用户和项目之间的关系，实现更准确的推荐。此外，研究人员将知识图谱和图神经网络结合起来用于推荐[9]。Wang 等人提出了推荐系统的知识图卷积网络模型[7]，该模型通过选择性地和有偏地聚合邻域信息，将图神经网络方法扩展到知识图谱，能够学习知识图谱的结构信息和语义信息以及用户的个性化和潜在兴趣。Chang 等人提出了一种基于图的解决方法[10]，将松散的项目序列重新构造为紧密的项目兴趣图，该模型利用图神经网络强大的能力，从嘈杂的用户行为序列中动态融合和提取用户激活的核心兴趣。

3. 模型设计

在本节中，我们将详细讨论提出的模型，并详细介绍了相关方法。

3.1. 符号定义

本文所使用的符号定义如下，在音乐推荐场景中，我们定义集合 $U = \{u_1, u_2, u_3, \dots, u_m\}$ ，和 $V = \{v_1, v_2, v_3, \dots, v_n\}$ ，分别代表用户和项目的集合，用户、项目交互矩阵 $Y = \{y_{uv} | u \in U, v \in V\}$ ，其中， $y_{uv} = 1$ 代表用户 u 和项目 v 之间存在交互行为，否则没有交互行为。根据交互矩阵 Y ，我们可以构建可用的知识图 \mathcal{G} ，它由海量的实体、关系组成实体三元组 (h, r, t) 。这里 $h \in \mathcal{E}$ ， $r \in \mathcal{R}$ 和 $t \in \mathcal{E}$ 分别表示知识三元组的头部、关系和尾部， \mathcal{E} 和 \mathcal{R} 表示 KG 中的实体和关系集。给定交互矩阵 Y 和知识图 \mathcal{G} ，我们的目标是预测用户 u 是否对他之前没有交互的项目 v 具有潜在兴趣。

3.2. 模型总体结构

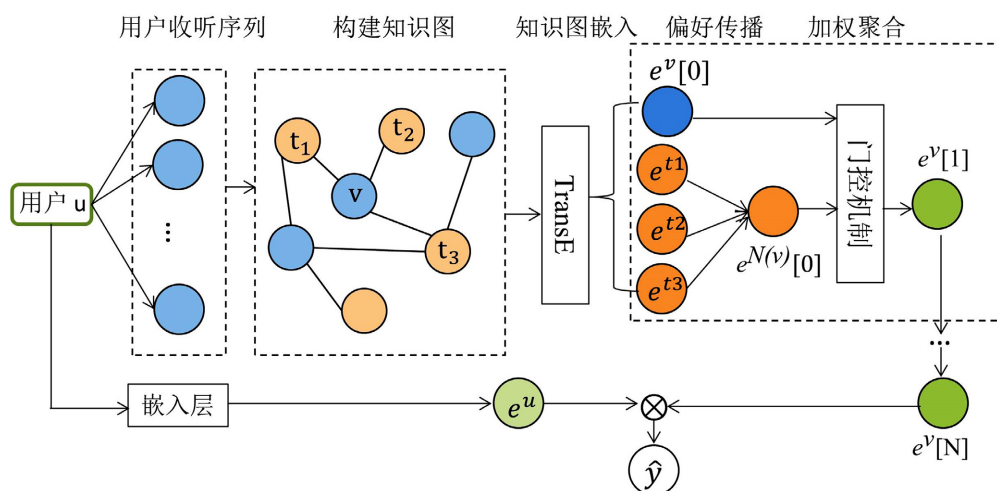


Figure 1. Overall structure of the model
图 1. 模型总体结构

模型总体结构如图 1 所示，首先根据用户项目交互的行为记录来构建知识图，然后图神经网络通过逐层聚合邻域信息来丰富项目表示和传播高阶信息，特别的，在聚合阶段，我们对节点自身和其邻域采用门控机制进行权重分配，从而更好地捕捉节点特征表示。

3.3. 知识图嵌入

TransE 是一种常用的知识图谱嵌入方法[11]，用于将知识图中的实体和关系映射到低维连续向量空间中。它的核心思想是通过学习实体和关系之间的转移关系来进行知识图谱的建模。具体来说，对于知识图中的每个三元组 (h, r, t) ，其中 h 和 t 分别表示头实体和尾实体， r 表示它们之间的关系。TransE 的目标是学习到一个映射函数，使得对于每个三元组 (h, r, t) ，满足 $h + r \approx t$ ，其中 $+$ 表示向量的加法。换句话说，TransE 希望通过学习实体和关系的向量表示，使得头实体、关系和尾实体之间的关系可以在向量空间中用简单的平移来表示。

具体的训练过程是通过最大化正向三元组得分和最小化负向三元组得分来实现的。对于每个正向三元组 (h, r, t) ，通过计算 $h + r$ 和 t 的距离，得到一个得分，使得正向三元组的得分尽可能小；同时，对于每个负向三元组 (h, r, t') ，其中 t' 表示一个错误的尾实体，通过计算 $h + r$ 和 t' 的距离，得到一个得分，使得负向三元组的得分尽可能大。通过这种方式，可以学习到实体和关系的向量表示，从而实现知识图谱的嵌入。公式如下：

$$L = \sum_{(e^h, e^r, e^t) \in T} \sum_{(e^{h'}, e^{r'}, e^{t'}) \in T} \max \left(\left[\gamma + \mathcal{E}(e^h, e^r, e^t) - \mathcal{E}(e^{h'}, e^{r'}, e^{t'}) \right], 0 \right) \quad (1)$$

其中， e^h, e^r, e^t 分别代表头部实体、关系和尾部实体的向量表示。 $\mathcal{E}(e^h, e^r, e^t)$ 为评分函数，定义如下：

$$\mathcal{E}(e^h, e^r, e^t) = \|e^h + e^r - e^t\| \quad (2)$$

通过训练目标函数，TransE 可以将实体和关系映射到一个低维向量空间，在这个空间中，相似的实体和关系更加接近。

3.4. 偏好传播

在获取实体和关系的嵌入后，我们需要计算它的邻域嵌入。实体和关系的嵌入可以通过知识图嵌入得到，邻域嵌入是实体的邻域嵌入的线性组合。为了区分不同邻居的重要性，在计算邻居的线性组合时，我们使用归一化的重要性分数作为邻居嵌入的权重。因此，邻域嵌入的计算公式如下：

$$e^{N(v)} = \sum_{a \in N(v)} \tilde{s}_{r,v,a}^u e^a \quad (3)$$

其中 $\tilde{s}_{r,v,a}^u$ 是归一化的重要性分数， $r_{v,a}$ 表示实体 v 与其邻居 a 的关系， $N(v)$ 表示实体 v 的邻居集合，归一化公式如下：

$$\tilde{s}_{r,v,a}^u = \frac{\exp(s_{r,v,a}^u)}{\sum_{a \in N(v)} \exp(s_{r,v,a}^u)} \quad (4)$$

其中，计算邻居重要性得分公式如下：

$$s_r^u = \sigma(e^{u^T} e^r) \quad (5)$$

在知识图谱中，用户的偏好信息会通过实体和关系之间的连接传播到其他相关实体上，在偏好传播的每一层中，通过邻居实体的迭代聚合来更新领域嵌入表示，公式如下：

$$e^{N(v)}[h-1] = \sum_{a \in N(v)} \tilde{s}_{r,v,a}^u e^a [h-1] \quad (6)$$

3.5. 加权聚合

我们通过图神经网络将实体表示 e^v 及其邻域表示 $e^{N(v)}$ 聚合到单个向量中。这里我们使用的聚合方式是通过门控机制来控制实体本身的嵌入与其相邻的嵌入的权重，从而根据具体的节点和邻居节点的特征动态地分配权重。门控卷积网络是由 Dauphin 等人提出的一个模型[12]，用于控制应该传播哪些信息以预测语言建模任务中的下一个单词，在这里，我们采用类似的策略来选择哪些特征用于预测与未来项目相关的特征。原论文中的公式如下：

$$z = (X * W + b) \otimes \sigma(X * V + c) \quad (7)$$

其中， X 代表输入， W 、 V 、 b 、 c 是可学习的参数， σ 是 sigmoid 函数， $*$ 代表卷积运算， \otimes 代表矩阵之间的逐元素乘积。

为了减少可学习参数的数量，我们在原始的门控卷积网络中应用内积而不是卷积操作，我们的输入是节点自身嵌入及其邻域嵌入表示，修改后公式如下：

$$z = (\{e^v, e^{N(v)}\} \cdot W + b) \otimes \sigma(\{e^v, e^{N(v)}\} \cdot V + c) \quad (8)$$

因此，门控机制单层聚合的公式如下：

$$e^v[1] = (\{e^v[0], e^{N(v)}[0]\} \cdot W + b) \otimes \sigma(\{e^v[0], e^{N(v)}[0]\} \cdot V + c) \quad (9)$$

以上是一层聚合的情况。此时， $e^v[0] = e^v$ ， $e^{N(v)}[0] = e^{N(v)}$ 。 $e^v[1]$ ，是一次迭代的结果。假设迭代层数为 n ，实体 v 的最终聚合表示为 $e^v[N]$ ，公式如下：

$$e^v[N] = (\{e^v[N-1], e^{N(v)}[N-1]\} \cdot W + b) \otimes \sigma(\{e^v[N-1], e^{N(v)}[N-1]\} \cdot V + c) \quad (10)$$

3.6. 评分预测

最终的 n 阶实体表示表示为 $e^v[N]$ ，它与用户表示 e^u 一起被送入函数 $f: R^d \times R^d \rightarrow R$ ，用于预测点击概率：

$$\hat{y} = f(e^v[N]^T e^u) \quad (11)$$

\hat{y} 表示用户收听音乐的预测概率。

3.7. 模型训练

我们使用时间反向传播算法来训练 MGWA 模型，训练集、评估集和测试集的比例为 6:2:2。所有参数都可以通过反向传播的梯度下降来计算。

4. 实验结果及分析

这部分是详细的实验设计，包括数据集、实验基准模型、评价指标和实验结果对比分析。

4.1. 数据集

Table 1. Basic information statistics of the dataset

表 1. 数据集的基本信息统计

		Last.FM			
用户	项目	交互	实体	关系	三元组
1872	3846	21,173	9366	60	15,518

Last.FM 包含来自 Last.fm 在线音乐系统的 1872 个用户的音乐家收听信息。数据集的知识图谱是由微软 Satori 构建的，并且只保留置信度大于 0.9 的整个 KG 三元组。实验数据集的统计数据总结在表 1 中。

4.2. 实验基准模型

RippleNet [6]: 这种方法利用知识图谱信息，以用户的历史项目为种子，在知识图中的路径上迭代的传播用户对实体集的偏好，最后通过相加每一层项目响应来聚合用户偏好。

KGCN [7]: 提出了一种结合知识图和图卷积神经网络特点的模型，该模型基于项目传播信息的多跳邻域性质，获得增强的项目表示，最后通过简单拼接或相加来聚合项目自身与其邻域的嵌入表示。

4.3. 评价指标

为了评估我们模型的性能，我们使用 AUC、ACC 和 F1 作为度量，AUC 是指 ROC 曲线下的面积。ROC 曲线是由 TPR 作为纵坐标和 FPR 作为横坐标形成的曲线。ACC 表示正确预测的样本数与样本总数的比率。F1 评分是对查准率和查全率的综合评价，公式如下：

$$F1 = \frac{2 * recall * precision}{recall + precision} \quad (12)$$

4.4. 实验结果对比分析

为了验证我们构建的模型的有效性，我们将提出的模型 MGWA 与经典基线进行了比较。推荐任务的实验结果见表 2。

Table 2. Comparison between model MGWA and baselines

表 2. 模型 MGWA 和基线的对比

Methods	AUC	ACC	F1
RippleNet	0.724	0.677	0.668
KGCN	0.856	0.778	0.776
MGWA (ours)	0.899	0.808	0.803

我们可以观察到，模型 MGWA 在所有评估设置中具有最佳性能。这表明了我们的模型采用加权聚合的方法的有效性。我们的模型优于 RippleNet 模型，是因为 RippleNet 模型未考虑知识图谱中的实体权重，从而未将推荐重点放在重要程度较高的实体上。我们的模型表现优于 KGCN 模型，是因为 KGCN 模型在聚合时采用的方法是简单的相加和拼接，忽略了节点自身的重要性，从而影响推荐准确率。

5. 总结和展望

本文提出了一种基于加权聚合图神经网络的音乐推荐模型 MGWA，通过引入门控机制动态地分配节点自身及其邻域的权重，从而更好地捕捉节点特征表示。实验结果表明，所提出的模型在真实数据集上取得了显著的性能提升，这证明了加权聚合的有效性和门控机制的灵活性和可学习性。未来，我们可以进一步探索模型的扩展性和泛化能力，尝试在更大规模的数据集上进行验证，同时也可以考虑将该模型应用到其他领域的推荐系统中，以期获得更广泛的应用和影响。

参考文献

- [1] Adomavicius, G. and Tuzhilin, A. (2005) Toward the Next Generation of Recommender Systems: A Survey of the

-
- State-of-the-Art and Possible Extensions. *IEEE Transactions on Knowledge and Data Engineering*, **17**, 734-749. <https://doi.org/10.1109/TKDE.2005.99>
- [2] Wang, X., He, X., Wang, M., *et al.* (2019) Neural Graph Collaborative Filtering. *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, New York, 21-25 July 2019, 165-174. <https://doi.org/10.1145/3331184.3331267>
- [3] Wu, J., He, X., Wang, X., *et al.* (2022) Graph Convolution Machine for Context-Aware Recommender System. *Frontiers of Computer Science*, **16**, Article 166614. <https://doi.org/10.1007/s11704-021-0261-8>
- [4] Ying, R., He, R., Chen, K., *et al.* (2018) Graph Convolutional Neural Networks for Web-Scale Recommender Systems. *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, New York, 19-23 August 2018, 974-983. <https://doi.org/10.1145/3219819.3219890>
- [5] Gori, M., Monfardini, G. and Scarselli, F. (2005) A New Model for Learning in Graph Domains. *Proceedings 2005 IEEE International Joint Conference on Neural Networks*, Montreal, 31 July-5 August 2005, 729-734.
- [6] Wang, H., Zhang, F., Wang, J., *et al.* (2018) RippleNet: Propagating User Preferences on the Knowledge Graph for Recommender Systems. *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, New York, 22-26 October 2018, 417-426. <https://doi.org/10.1145/3269206.3271739>
- [7] Wang, H., Zhao, M., Xie, X., *et al.* (2019) Knowledge Graph Convolutional Networks for Recommender Systems. *The World Wide Web Conference*, San Francisco, 13-17 May 2019, 3307-3313. <https://doi.org/10.1145/3308558.3313417>
- [8] Tai, C.-Y., Wu, M.-Y., Chu, Y.-W., *et al.* (2019) GraphSW: A Training Protocol Based on Stage-Wise Training for GNN-Based Recommender Model.
- [9] Wang, X., He, X., Cao, Y., *et al.* (2019) Kgat: Knowledge Graph Attention Network for Recommendation. *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, New York, 4-8 August 2019, 950-958. <https://doi.org/10.1145/3292500.3330989>
- [10] Chang, J., Gao, C., Zheng, Y., *et al.* (2021) Sequential Recommendation with Graph Neural Networks. *The 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*, New York, 11-15 July 2021, 378-387. <https://doi.org/10.1145/3404835.3462968>
- [11] Bordes, A., Usunier, N., Garcia-Duran, A., *et al.* (2013) Translating Embeddings for Modeling Multi-Relational Data. *Advances in Neural Information Processing Systems 26 (NIPS 2013)*, Harrahs and Harveys, 5-11 December 2013, 2787-2795.
- [12] Dauphin, Y.N., Fan, A., Auli, M., *et al.* (2016) Language Modeling with Gated Convolutional Networks. *Computation and Language*.