doi:10.3772/j.issn.2095-915x.2023.01.009

基于时间兴趣因子的网络表示学习图书推荐模型研究

开放科学（资源服务）标识码（OSID）

王日花

中国传媒大学图书馆 北京 100024

摘要: [目的/意义]通过分析图书馆的图书流通数据，本文提出一种基于时间兴趣因子融合网络学习的图书推荐模型——TIF_N2V_CF。[方法/过程]评估用户借阅图书的时间间隔并定义兴趣因子权重，根据流通数据构建同质关系网络；网络表示学习将得到的特征矩阵输入融合推荐模型并得到推荐结果。[结果/结论]实验表明，TIF_N2V_CF 模型的召回率在 top z=10 和 z=20 时分别为 0.1302、0.2031，高于未引入时间兴趣因子的 N2V_CF 模型。TIF_N2V_CF 模型将时间兴趣因子引入到网络表示学习，对融合用户和图书的特征矩阵进行相似度计算，解决图书借阅流通数据中同一时间包含多本图书借阅记录造成的难以序列化的问题，缓解数据稀疏和冷启动对模型性能的影响，提高了推荐精度。

关键词：网络表示学习；Node2Vec；协同过滤；时间兴趣因子；Word2Vec

中图分类号：G35；G25

Network Representation Learning Book Recommendation Model Research Based on the Time Interest Factor

WANG Rihua

Library of Communication University of China, Beijing 100024, China

Abstract: [Purpose/significance] By analyzing the circulation data of library book borrowing, this paper proposes a library recommendation model based on time interest factor and network learning-TIF_N2V_CF. [Method/process] Evaluate the time between users to borrow books and define the weights of interest factors, constructing a homogenous relationship network based on circulation data; The network representation learning inputs the obtained feature matrix into the fusion recommendation module and obtains the recommendation results. [Result/conclusion] Experiments show that the Recall of the TIF_N2V_CF model is 0.1302 and 0.2031 at top z=10 and z=20, respectively, which is higher than the N2V_CF model without time interest factor. The TIF_N2V_CF model introduces the time interest factor into the network representation learning, and calculates the similarity of the feature matrix that combines the user and the book, the problem of difficult serialization caused by the

作者简介 王日花（1983-），硕士，副研究馆员，主要研究为信息资源推荐和咨询、知识服务和智慧图书馆建设研究，E-mail：sunshine19626@sina.com。

引用格式 王日花 . 基于时间兴趣因子的网络表示学习图书推荐模型研究 [J]. 情报工程 , 2023, 9(1): 118-127.

book borrowing circulation data containing multiple book borrowing records at the same time is solved, and the impact of data sparseness and cold start on the model performance is alleviated, and the recommendation accuracy is improved.
Keywords: network representation learning; Node2Vec; collaborative filtering; time interest factor; Word2Vec

引言

图书推荐是图书馆运营的重要指标，高效的推荐模型能提升图书馆管理水平，在大数据和智能化背景下成为研究热点。协同过滤模型（Collaborative Filtering，CF）由 Goldberg 等 [1] 提出，是图书推荐的重要算法而被广泛使用；通过分析用户和商品之间的关系构造用户 – 商品的评价矩阵，根据矩阵中各向量与目标向量之间的距离关系输出排序结果，向量排序越靠前，表示用户越感兴趣。但是，模型在实际应用过程中存在下述瓶颈：（1）对于用户 – 商品评价矩阵中未出现过的用户商品关系，该模型无法对其进行距离计算。（2）随着用户和商品数量的增长，用户 – 商品评价矩阵也会变得越来越稀疏；此外，推荐系统会面临冷启动问题 [2]，根据对象分为新用户、新项目和新系统三种冷启动；这将会降低推荐系统精度和效率。

Liang L[3] 提出了基于深度神经网络的图 Embeddings 模型 N2V_CF，通过嵌入技术 Embeddings，表征用户和物品间的关系，压缩评价矩阵维度，解决了没有“用户 – 商品关系向量”而无法计算距离的问题 [4-9]。但是，用户兴趣会随着时间的推移逐步衰减 [10]，如果不考虑时间对用户兴趣的影响，算法则无差别地生成推荐结果，将对模型性能产生一定影响。此外，用户在互联网上每一次交互均会产生日志记录，每条交互日志记录与时间为一对一关系，而对于图书流通数据，每一次借阅记录包含了多本图书的借阅信息，每次借阅记录与时间为多对一的关系，对此类数据进行建模的方法，现存文献中未提及。

基于此，本文考虑到时间对用户兴趣的影响，借助深度神经网络图 Embeddings 模型，通过分析中国传媒大学图书借阅数据，提出基于时间兴趣因子的图书推荐模型 TIF_N2V_CF，模型通过时间兴趣因子加权实现基于用户向量和图书向量的融合推荐。

1 相关研究

协同过滤算法通常用于商品推荐，利用用户对商品浏览和评价构建矩阵用户 – 商品评价矩阵，矩阵能够表示不同用户和不同物品之间的关系，稠密矩阵能够借助协同过滤算法取得较好的预测结果。但是，对于不停上新的商品和不断增加的用户，构建的矩阵将会变得越来越稀疏，稀疏矩阵会对算法准确性产生较大影响，相似度计算值可信度变低 [11]。

近年来，通过神经网络对自然语言进行处理的方法得到了大范围推广，由 Mikolov 2013 年提出的神经网络语言模型 Word2Vec，对于输入模型的词序列，模型能够输出一组数值向量用于描述词与词之间的关系 [12-13]。Word2Vec 模型有 CBOW(Continuous Bag-of-Words) 和 Skip-Gram 两种实现方式，CBOW 定义的模型输入

doi:10.3772/j.issn.2095-915x.2023.01.009

为给定特征词的上下文，输出为给定特征词向量，Skip-Gram定义的模型输入为给定的特征词，输出为给定特征词上下文词向量。屈冰洋等[14]借鉴了 Word2Vec Skip-Gram 模型输出词向量思想，将文献标题、摘要和关键词等信息输入模型，训练得到一组数值向量用于表示文献之间的关系。这组数值向量可取代用户–物品关系矩阵，解决矩阵稀疏带来的效率和准确性问题。

图模型是一类用图来表示概率分布的技术总称，与传统协同过滤算法不同，DeepWalk 是一种基于图结构的特征提取算法[15]。根据用户–物品之间的关系，可根据用户选择物品序列构建物品关系图，每个物品对应图中一个节点，物品到物品之间连接关系定义为图中的边，DeepWalk 算法则可以根据图中各节点之间的概率关系，在各节点中进行随机游走[16-18]。DeepWalk 算法中关键步骤随机游走，是指从图中某个给定节点开始，根据该节点所有相邻节点边权重，选择一条边作为游走路径，将节点沿着选定边移动到下一节点的过程。游走经过各个节点重新组成新的序列关系，将新的序列带入 Word2Vec 模型中进行训练，即可得到一种表达能力更强的关系向量矩阵。Node2Vec 算法则在此基础上提出了 BFS（宽度优先搜索）和 DFS（深度优先搜索）的随机游走方法，是一种综合考虑 BFS 邻域和 DFS 邻域的特征嵌入方法，权衡了图结构的同质性和结构性，输出向量能够更好表征物品关联信息，并有效缓解了数据稀疏问题[19-20]。

跳转概率的定义是整个序列生成的核心，随着应用场景变化，算法中关键参数边权重决定算法应用效果。本文针对图书馆借阅图书流通记录数据，将各个时刻记录的图书借阅数据进行拆解，通过时间兴趣因子重新定义用户–图书图结构中随机跳转概率权重，实现时间兴趣因子与跳转概率的融合，获得表征用户和图书的向量矩阵，将向量矩阵进行融合计算，根据用户下一次入馆时间和历史借阅图书情况输出推荐结果。

2 TIF_N2V_CF 模型框架

TIF_N2V_CF 模型数据均通过图书流通记录进行变换生成。图书流通记录数据中每一行表示一个用户借阅一本图书的时间信息，由用户 id、图书 id、借书日期、借书时间、还书日期、还书时间和借阅时长这 7 个字段组成。整体框架如图 1 所示，流程如下：

第一步，序列生成。利用图书流通记录数据，按照用户 id、借书日期和借阅时长进行分组，得到用户借阅图书序列；根据图书 id、借书日期和借阅时长进行分组，得到图书被借阅用户序列。

第二步，向量表示。对比用户借阅图书序列和图书被借阅的用户序列，分别生成图书同质网络图和用户同质网络图，网络图中各节点边权重由时间兴趣因子定义，引入时间兴趣因子作为随机游走跳转概率。游走步长值为 1，配置游走次数和游走节点数量来代替截止频率，控制随机游走得到的序列长度，Node2Vec 算法进行随机游走重新生成基于图的图书序列和用户序列。将基于图的图书序列和用户序列输入 Word2Vec 算法，得到表示图书的特征向量和表示用户的特征向量。

第三步，图书推荐。结合图书和用户的特征向量，设计融合推荐算法模型，根据用户新入馆用户 id 和入馆时间，算法可以输出推荐图书序列，用于引导用户借阅图书。

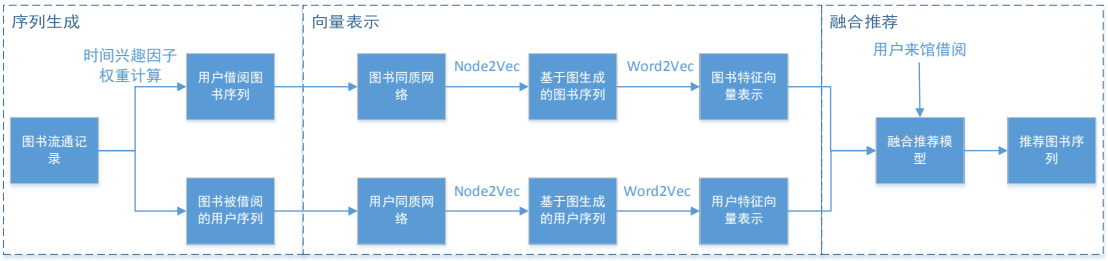


图 1 TIF_N2V_CF 模型框架

3 TIF_N2V_CF 模型流程

3.1 序列生成

用户借阅图书行为与平时在网页上点击物品行为不同，用户每次在网页上点击的物品，和点击时间结合形成一条唯一的兴趣记录，根据时间维度生成物品兴趣序列。图书馆记录的用户借书时间发生在用户完成图书选择后，到前台登记的时间；该时间记录用户选择多本图书的兴趣记录，不能通过该时间推断出多本图书之间兴趣影响的先后关系；因此，数值向量化关键在于物品兴趣序列的生成。

提出一种针对用户借阅图书行为的兴趣序列生成方法。在同一时刻，各图书相互影响权重相等，在不同时刻，历史借阅图书对该时刻借阅图书影响随时间间隔的长短变化，权重由时间兴趣因子 TIF(Time Interest Factor) 决定。以用户二次借阅同一本书的行为作为时间兴趣因子评估依据，对二次借阅之间间隔和产生二次借阅行为的用户分布进行分析，结果如图 2 所示。

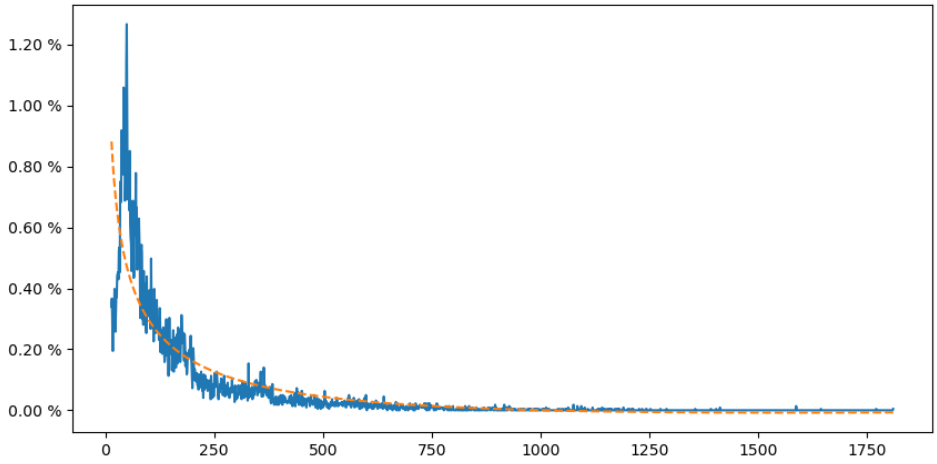


图 2 二次借阅用户占比随时间的变化及其拟合曲线

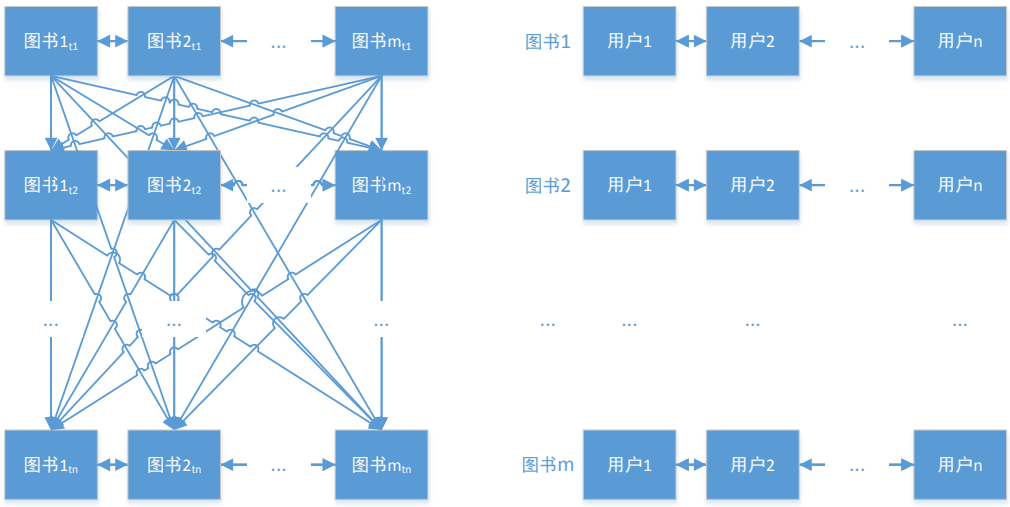
图 2 中横坐标表示用户两次借阅行为的时间间隔，纵坐标表示在该时间间隔产生二次借阅的用户数占比。图 2 中实线表示真实产生二次借阅行为的用户数，表明随着时间间隔的拉长，用户借阅同一本书行为会逐渐减少。通过以下方程拟合得到了虚线，用于表示用户的兴

doi:10.3772/j.issn.2095-915x.2023.01.009

趣会随着时间推移逐渐减弱，因此，时间兴趣因子 TIF(Δt) 可定义为：

$$\begin{cases} 1, \Delta t = 0 \\ \alpha \log\left(\frac{1}{\Delta t}\right)^2 + \beta \log\left(\frac{1}{\Delta t}\right) + b, \Delta t > 0 \end{cases} \quad (1)$$

其中，α 和 β 均为拟合系数，b 为偏置项。Δt 为两次借阅之间的间隔天数。公式（1）表明，当在同一日期借阅多本书时，多本书之间的借阅间隔天数为 0，因此，影响用户借阅行为的兴趣因子为 1；而在不同日期借阅多本书之间，兴趣因子将受到借阅间隔日期影响，假设时刻 tn>t1，则仅有用户在 t1 时刻借阅的图书会对 tn 时刻的借阅行为产生影响，而不会有 tn 时刻借阅的图书对 t1 时刻借阅的行为产生影响。根据兴趣因子与借阅行为的关系，用户在 t1–tn 时刻借阅图书关联关系由图 3(a) 所示。



(a)用户借阅图书序列　　(b)图书被借阅的用户序列

图 3 用户借阅图书和图书被用户借阅关系

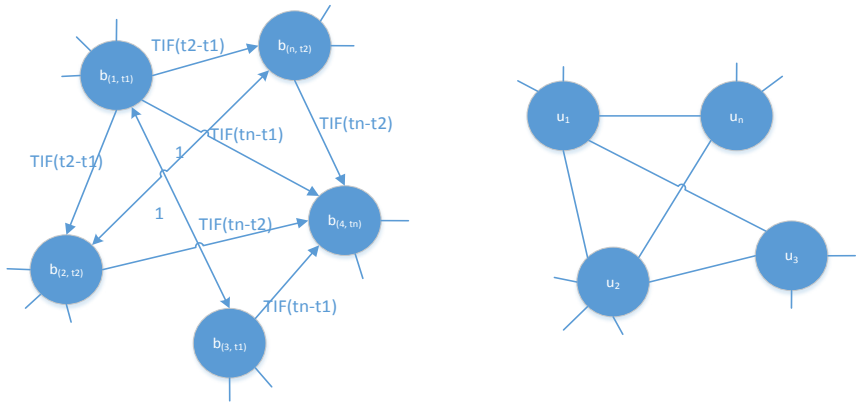
图 3 中每条线连接的两本图书可以形成图书序列，箭头方向表示前一时刻借阅图书对后一时刻借阅图书的影响，例如，用户在 t1 时刻生成的图书兴趣序列为：$(b_{(1,t1)}, b_{(2,t1)}) \cdots (b_{(1,t1)}, b_{(m,t1)}), (b_{(2,t1)}, b_{(1,t1)}) \cdots (b_{(2,t1)}, b_{(m,t1)}), \cdots (b_{(m,t1)}, b_{(1,t1)}) \cdots (b_{(m,t1)}, b_{(m-1,t1)})$。$(b_{(x,ti)}, b_{(y,tj)})$ 表示在 ti 时刻用户借阅的图书 b_x 与 tj 时刻用户借阅图书 b_y 之间的关系，其中，时刻 ti < tj, $b_x, b_y \in B$，B 为图书的集合。增加了兴趣因子后，序列每一项将增加一个权重值得到新序列：$(b_{(1,t1)}, b_{(m,t1)}, 1) \cdots (b_{(1,t1)}, b_{(m,t1)}, 1), (b_{(2,t1)}, b_{(1,t1)}, 1) \cdots (b_{(2,t1)}, b_{(m,t1)}, 1) \cdots (b_{(m,t1)}, b_{(1,t1)}, 1) \cdots (b_{(m,t1)}, b_{(m-1,t1)}, 1)$。在同一时刻借阅多本书时，多本书之间的借阅间隔为 0，因此，用户借阅行为的兴趣因子权重为 1。ti 时刻借阅图书 b_x 与 tj 时刻借阅图书 b_y 之间的关系，可用三元组的形式表示为：$\{(b_x, b_y, TIF(tj - ti))\}$。

与生成图书序列的方法相同，根据图书 id 进行分组，可得到图书被借阅的用户序列。借阅同一本图书用户之间的关系与时间无关，用户序列没有权重项，如图 3(b) 表示。借阅图书 1 至图书 m 的用户序列均可表示为 $(u_1, u_2) \cdots (u_1, u_n), (u_2, u_1) \cdots (u_2, u_n) \cdots (u_n, u_1) \cdots (u_n, u_{n-1})$，可得到图书 1 至图书 m 被借阅的用户序列二元组为：

$\{(u_m, u_n)\}$。

3.2 向量表示

根据序列生成步骤得到的图书三元组 $\{(b_x, b_y, TIF(tj-ti))\}$，以 b_x 和 b_y 作为图中的节点，将 $TIF(tj-ti)$ 作为连接 b_x 和 b_y 节点边的权重，根据节点和边的关系，把图书三元组绘制成图，该图即为图书同质网络 $G_b=\{V_b, E_b, W_b\}$，其中 V_b 表示图书节点集合，E_b 表示节点之间边的集合，W_b 表示边的权重集合。与生成图书同质网络的方法相同，把用户二元组 $\{(u_m, u_n)\}$ 绘制成图，得到用户同质网络 $G_u=\{V_u, E_u\}$，其中 V_u 表示用户节点集合，E_u 表示节点之间边的集合，图书同质网络和用户同质网络如图 4 所示。



(a)图书同质网络　　(b)用户同质网络

图 4 图书同质网络和用户同质网络

对 G_b 和 G_u 使用 Node2Vec 在同质网络中各节点执行随机游走，分别得到基于图生成的图书序列和用户序列。Node2Vec 定义随机游走节点跳转概率的公式为：

$$p(v_i \mid v_j) = \alpha_{ij} \frac{M_{ij}}{\sum_{j \in N+(v_i)} M_{ij}} \tag{2}$$

其中，对于有向有权图的图书同质网络 G_b，$N_{+(vi)}$ 是节点 v_i 所有出边的集合，M_{ij} 是节点 V_i 到节点 V_j 边的权重，$M_{ij} = TIF(tj-ti)$，对于无向无权图 G_u，$N_{+(vi)}$ 为节点 v_i 所有边的集合，$M_{ij} = 1$。α_{ij} 为 Node2Vec 定义的跳转权重，公式定义为：

$$\begin{cases} \alpha_{ij} = \dfrac{1}{p}, d_{ij} = 0 \\ \alpha_{ij} = 1, d_{ij} = 1 \\ \alpha_{ij} = \dfrac{1}{q}, d_{ij} = 0 \end{cases} \tag{3}$$

其中 d_{ij} 是节点 v_i 到节点 v_j 的距离。p 和 q 是用来控制近邻搜索策略的两个超参数，其中参数 p 控制返回至已游走节点的概率，q 控制游走的方向。若 q>1，偏向 BFS 游走策略。若 q<1，则倾向 DFS 游走策略。对 Node2Vec 随机游走节点跳转概率的两个超参数 p 和 q 进行调整，在同质网络 G_b 和 G_u 中分别进行随机游走，得到基于图生成的图书序列 T_b 和用户序列 T_u。

基于随机游走得到图书序列 T_b 和用户序列 T_u，每条节点序列视为句子，节点序列中的节点视为词语，通过 Word2Vec 的 Skip-Gram 模型对 T_b 和 T_u 进行训练，分别得到表示图书的特征向量和表示用户的特征向量。节点的向量表示，其任务为最大化如下目标函数：

doi:10.3772/j.issn.2095-915x.2023.01.009

$$p(v_o \mid v_i) = \frac{\exp(v_o^T v_i)}{\sum_{c \in C} \exp(v_c^T v_i)} \quad (4)$$

其中，v_i 为当前节点的向量表示，v_o 为上下文节点的向量表示，v_c 为上下文节点之外的任意节点向量表示，C 为所有词向量的集合。Skip-Gram 把当前节点作为输入来预测其上下文节点，通过最大化上下文节点与真实节点之间的概率，得到一组低维向量来表示当前节点，模型结构如图 5 所示。

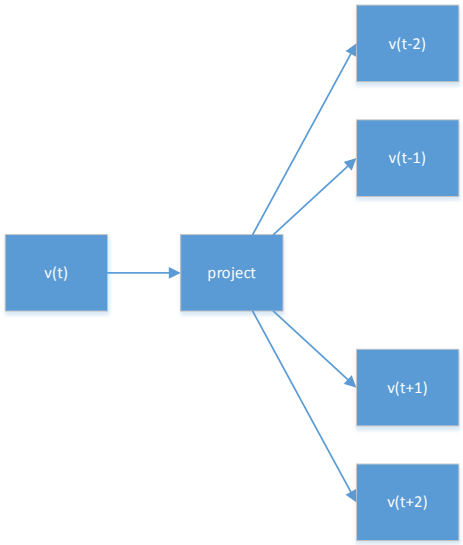


图 5　Word2Vec 的 Skip-Gram 模型结构

定义训练输出图书特征向量维度 d_b 和用户特征向量维度 d_u，通过最大化上述目标函数值，分别对图书同质网络 G_b 和用户同质网络 G_u 中各节点进行训练，得到每个图书节点和用户节点的向量表示。所有图书向量表示组成图书特征矩阵 M，所有用户的向量表示组成用户特征矩阵 N。图书特征矩阵 M 和用户特征矩阵 N，均是由距离意义的向量组成，避免了新用户未借阅图书而产生全 0 的图书向量；同时，避免了新增图书未被借阅而产生全 0 的用户向量，这些均缓解了图书和用户的模型冷启动问题。

3.3　融合推荐

推荐规则：假设用户 u_{target}，在 t_{target} 时刻选择了图书 $b_{target1}$、$b_{target2}$、$b_{targetn}$，则根据用户特征矩阵 N，进行余弦相似度计算 Similar，找出与 u_{target} 相似度最高的 top x 用户，标记为 $u_1 \cdots u_n$，对于每个相似的用户 u_i，得出其在 t_{i1} 时刻至 t_{im} 时刻借阅图书的情况，其中 $t_{im} < t_{target}$，所借阅图书标记为 $b_{i1} \cdots b_{im}$。则定义推荐 b_j 的推荐度为 $R(b_j)$，可用公式表示为：

$$R(b_j) = \sum_i^x TIF(t_{target} - t_i) Similar(u_i, u_{target}) \quad (5)$$

因此，得到基于用户的推荐度集合 $R_u = \{R(b_1) \cdots R(b_m)\}$。

根据 u_{target} 在 t_1 至 t_n 时刻历史的图书借阅情况，其中 $t_n < t_{target}$，对 u_{target} 在 t_i 时刻所借阅的图书标记为 $b_{i1} \cdots b_{im}$，根据图书特征矩阵 M 进行余弦相似度计算，每本书得到相似度最高的 top y 本书，与 b_{ij} 相似的 y 本书可以表示为 $b_{ij1} \cdots b_{ijy}$，则 y 本书中第 k 本的推荐度定义为：

$$R(b_k) = \sum_j^m \sum_j^n TIF(t_{target} - t_i) Similar(b_{ij}, b_{ijk}) \quad (6)$$

因此，得到基于图书的推荐度集合 $R_b = \{R(b_1) \cdots R(b_{(y*m)})\}$。

R_u 和 R_b 两集合对应图书的推荐度进行相加后，进行倒序排序，排序结果取前 z 个推荐的图书形成集合，该集合即为最终 top z 的推荐结果。

4　实验与结果

4.1　数据集

采用中国传媒大学图书馆流通记录数据进行实验，数据集包含 25 416 个用户、借阅 192 251 本图书、共计 474 705 条借阅记录。实验使用数据集的用户 id、图书 id、借书日期、借书

时间、还书日期、还书时间和借阅时长等 7 个字段，将 2016 年 1 月 1 日—2021 年 4 月 1 日以前的数据作为训练集，用于生成图书特征矩阵 M 和用户特征矩阵 N，以 2021 年 4 月 1 日—2021 年 5 月 17 日的数据作为测试集，根据图书特征矩阵 M 和用户特征矩阵 N，仅通过用户 id 和借书时间，使用融合推荐模型输出 top z 本用户可能选择的图书。如果该用户在该时刻确实选择了 top z 本推荐图书中的任意一本或多本，则本条借阅数据记为 1，如果 top z 本推荐图书中没有一本图书被选择，则记为 0。

4.2 评价指标

召回率（Recall）表示样本中正例被正确预测的比例，采用召回率作为提出模型的衡量指标，其计算公式为：

$$\text{Recall}=\frac{TP}{TP+FP} \tag{7}$$

其中 TP 为预测为 1 且实际为 1 的样本数量，FP 为预测为 1 但实际为 0 的样本数量。Recall 指标结果越高，说明模型的预测精度越高。

4.3 结果分析

在序列生成步骤中增加兴趣因子作为序列权重的模型定义为 TIF_N2V_CF，与未增加兴趣因子权重的模型 N2V_CF 进行对比 [3]。通过设置不同的 top z 值，得到两种模型的召回率情况。

根据公式 (5) 和 (6) 分别得到推荐图书集合 R_u 和 R_b，在不进行集合融合的情况下，直接使用 R_b 进行图书推荐，则 TIF_N2V_CF(R_b) 和 N2V_CF(R_b) 的推荐实验结果如图 6 所示。

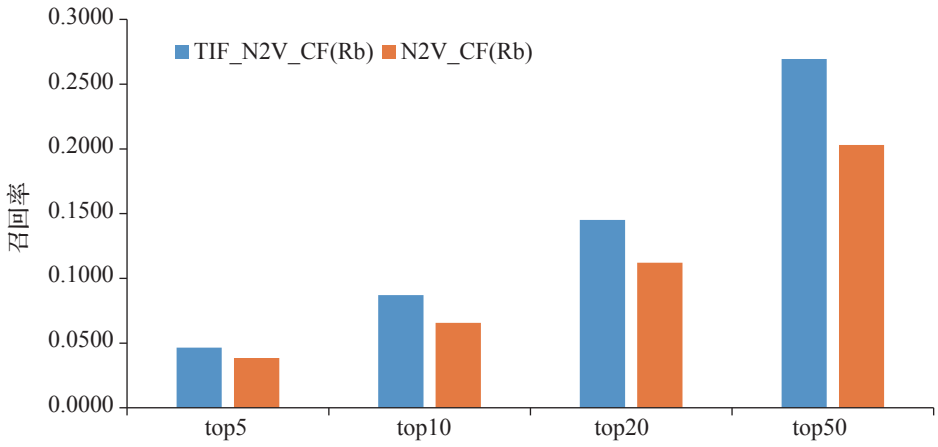


图 6 TIF_N2V_CF(R_b) 和 N2V_CF(R_b) 召回率对比

对比图 6 生成 TIF_N2V_CF(R_b) 和 N2V_CF(R_b) 召回率数据，如表 1 所示：

表 1 TIF_N2V_CF(R_b) 和 N2V_CF(R_b) 召回率对比表

	TIF_N2V_CF(R_b)	N2V_CF(R_b)
top5	0.0464	0.0384
top10	0.0871	0.0657
top20	0.1451	0.1121
top50	0.2695	0.2031

由表 1 实验结果得出，未使用 R_u 序列进行融合的 TIF_N2V_CF(R_b)，在 z 为任何值时，召回率均优于 N2V_CF(R_b)。

对 R_u 和 R_b 集合进行融合使用完整的 TIF_N2V_CF 和 N2V_CF 模型进行实验，实验结果如图 7 所示。

doi:10.3772/j.issn.2095-915x.2023.01.009

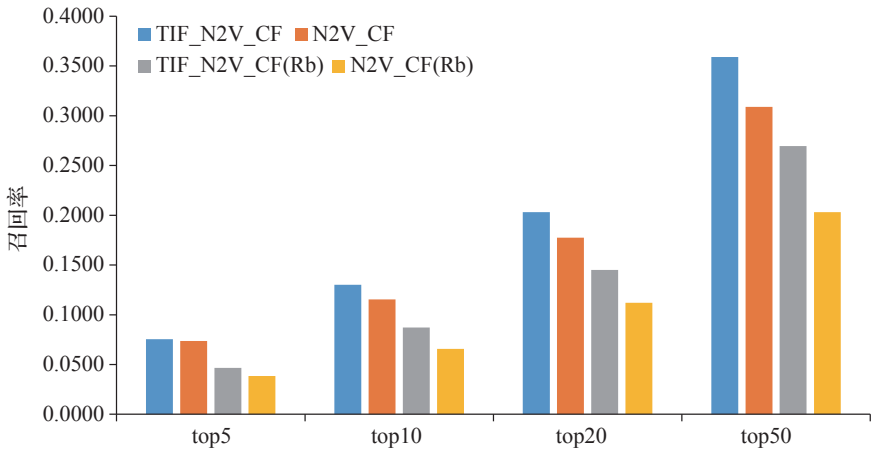


图 7 TIF_N2V_CF、N2V_CF、TIF_N2V_CF(R_b) 和 N2V_CF(R_b) 召回率对比

对比图 7 生成的 TIF_N2V_CF、N2V_CF、TIF_N2V_CF(R_b) 和 N2V_CF(R_b) 召回率数据，如表 2 所示：

表 2 TIF_N2V_CF、N2V_CF、TIF_N2V_CF(R_b) 和 N2V_CF(R_b) 召回率对比表

	TIF_N2V_CF	N2V_CF[3]	TIF_N2V_CF(R_b)	N2V_CF(R_b)
top5	0.0754	0.0736	0.0464	0.0384
top10	0.1302	0.1153	0.0871	0.0657
top20	0.2031	0.1774	0.1451	0.1121
top50	0.3591	0.3090	0.2695	0.2031

由表 2 得出，同时使用 R_u 和 R_b 集合融合后进行图书推荐的 N2V_CF 模型，要优于未进行集合融合的 TIF_N2V_CF(R_b) 和 N2V_CF(R_b) 模型；增加了兴趣因子权重的 TIF_N2V_CF 在 $z < 10$ 时召回率与未增加兴趣因子权重的 N2V_CF 保持一致；对于 $z \geq 10$ 时，TIF_N2V_CF 的召回率要优于 N2V_CF。实验表明，TIF_N2V_CF 模型相较于未增加兴趣因子权重的 N2V_CF 模型，具有更优的预测精度和表现。

5 结语

本文考虑到时间对用户兴趣的影响，通过自然语音处理中深度神经网络训练词向量模型的方法，结合图计算的宽度优先搜索和深度优先搜索，给出一种基于时间兴趣因子的图书推荐模型 TIF_N2V_CF。主要优势是对用户二次借阅图书的情况进行分析，定义图结构中各节点间的时间权重兴趣因子，有效解决同一时间借阅记录中多本图书无法序列化的问题；利用随机游走的方式生成用户和图书序列，使用词向量训练方法得到用户和图书的关系向量矩阵，相对于传统协同过滤的评价矩阵，神经网络训练得到的关系向量矩阵能够更好的表征物品关联关系，有效缓解数据稀疏和冷启动的问题。实验证明，引入时间兴趣因子权重计算相似度的融合推荐模型性能更优。

本文是在测试环境中对算法进行了基础研究，未考虑生产环境中的硬件和网络现状；后续，为保证算法在使用过程中的效能，考虑引入用户和图书信息等元数据，可对算法参数进行优化和调整。

参 考 文 献

[1] Goldberg D, Nichols D, Oki B M, et al. Using collaborative filtering to weave an information

tapestry[J]. Communications of the Acm, 1992, 35(12): 61-70.

[2] 史海燕, 倪云瑞.推荐系统冷启动问题研究进展[J]. 图书馆学研究 , 2021(12):2-10.DOI:10.15941/j.cnki.issn1001-0424.2021.12.001.

[3] Liang L, Tang R. An Improved Collaborative Filtering Algorithm Based on Node2vec[C]. 2nd International Conference on Computer Science and Artificial Intelligence. ShenZhen: Shenzhen Univ, Coll Comp Sci & Software Engn; Kalbis, Inst 2018: 218-222.

[4] Vasile F, Smirnova E, Conneau A. Meta-Prod2Vec - Product Embeddings Using Side-Information for Recommendation[C]. Proceedings of the 10th ACM Conference on Recommender Systems, 2016: 225-232.

[5] He X, Zhang H, Kan M Y, et al. Fast Matrix Factorization for Online Recommendation with Implicit Feedback[C]. Proceedings of the 39th International ACM SIGIR conference, 2016: 549-558.

[6] He X, Liao L, Zhang H, et al. Neural Collaborative Filtering[C]. Proceedings of the 26th International World Wide Web Conferences Steering Committee, 2017: 173-182.

[7] Wang X, He X, Wang M, et al. Neural Graph Collaborative Filtering[C]. Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval. 2019: 165-174.

[8] Yang Z, Cohen W, Salakhutdinov R. Revisiting Semi-Supervised Learning with Graph Embeddings[C]. Proceedings of the 33rd International Conference on Machine Learning, 2016: 40–48.

[9] 崔岩 , 祁伟 , 庞海龙 , 赵辉 . 融合协同过滤和 XGBoost 的推荐算法 [J]. 计算机应用研究 , 2020, 37(1): 62-65.
Cui Yan, Qi Wei, Pang Hailong, Zhao Hui. Extreme gradient boosting recommendation algorithm with collaborative filtering[J]. Application Research of Computer, 2020, 37(1): 62-65.

[10] 刘乔 , 刘彬 . 基于时间加权的协同过滤推荐算法的改进 [J]. 计算机工程与设计 , 2016, 37(7): 1827-1830, 1872.
Liu Qiao, Liu, Bin. Improvement of collaborative filtering recommendation algorithm based on time weighting[J]. Computer Engineering and Design, 2016, 37(7): 1827-1830, 1872.

[11] 冷亚军 , 陆青 , 梁昌勇 . 协同过滤推荐技术综述 [J]. 模式识别与人工智能 , 2014, 27(8): 720-734.
Leng Yajun, Lu Qing, Liang Changyong. Survey of Recommendation Based on Collaborative Filtering[J]. Pattern Recognition and Artificial Intelligence, 2014, 27(8): 720-734.

[12] Mikolov T, Chen K, Corrado G, et al. Efficient Estimation of Word Representations in Vector Space[J]. Computer Science, 2013.

[13] Mikolov T, Sutskever I, Chen K, et al. Distributed Representations of Words and Phrases and their Compositionality[J]. Advances in Neural Information Processing Systems, 2013, 26: 3111-3119.

[14] 屈冰洋 , 王亚民 . 基于深度学习的科技信息文献推荐模型研究 [J]. 情报理论与实践 , {3}, {4}{5}:1-10.

[15] Perozzi B, Al-Rfou R, Skiena S. Deepwalk: Online learning of social representations[C]. Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2014: 701-710.

[16] Tang J, Qu M, Wang M, et al. LINE: Large-scale Information Network Embedding[C]. International World Wide Web Conferences Steering Committee, 2015: 1067-1077.

[17] Yao W L, He J, Huang G Y, et al. A Graph-based model for context-aware recommendation using implicit feedback data[J]. World Wide Web-internet & Web Information Systems, 2015, 18(5): 1351-1371.

[18] Grbovic M, Radosavljevic V, Djuric N, et al. E-commerce in Your Inbox: Product Recommendations at Scale[C]. Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2015: 1809-1818.

[19] Grover A, Leskovec J. node2vec: Scalable feature learning for networks[C]. Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016: 855-864.

[20] 何瑾琳 , 刘学军 , 徐新艳 , 毛宇佳 . 融合 Node2Vec 和深度神经网络的隐式反馈推荐模型 [J]. 计算机科学 , 2019, 46(6): 41-48.
He Jinlin, Liu Xuejun, Xu Xinyan, Mao Yujia. Implicit Feed Recommendation Model Combining Node2Vec and Deep Neural Networks[J]. Computer Science, 2019, 46(6): 41-48.