

基于时间敏感的异构图相似神经网络的新闻推荐

张雯涛^{1,2,3}, 栾方军^{1,2,3*}, 袁帅^{1,2,3}, 刘国奇^{1,2,3}

¹沈阳建筑大学计算机科学与工程学院, 辽宁 沈阳

²辽宁省城市建设大数据管理与分析重点实验室, 辽宁 沈阳

³国家特种计算机工程技术研究中心沈阳分中心, 辽宁 沈阳

收稿日期: 2023年3月12日; 录用日期: 2024年4月12日; 发布日期: 2024年4月18日

摘要

目前网络上新闻信息呈现爆炸式增长趋势, 为方便用户快速找到他们感兴趣的新闻, 个性化新闻推荐变得愈发重要。目前主流的新闻推荐方法大多都是基于静态数据进行研究, 忽略了可以间接反映用户对新闻兴趣程度的动态信息, 如用户阅读新闻的时间等。为了解决此类问题, 文中在GNewsRec模型的基础上提出了一种基于时间敏感的异构图相似神经网络模型(TSHGSN)。采用CNN学习新闻特征, 加入用户阅读新闻时间权重并利用LSTM学习用户点击新闻的序列特征作为用户短期偏好。同时, 构建了一个异构图, 建模用户-新闻-主题关联, 采用新的邻居采样方法聚合节点获取候选新闻特征表示和用户长期偏好。最后, 将用户短期兴趣和长期兴趣与候选新闻分开进行相关性计算, 旨在自适应的调整用户建模中短期兴趣和长期兴趣的重要性。实验数据表明, 与GNewsRec模型相比, 该模型在AUC指标上提高约4%。

关键词

新闻推荐, 图神经网络, 短期偏好, 长期偏好

Time-Sensitive Heterogeneous Graph Similarity Neural Network for News Recommendation

Wentao Zhang^{1,2,3}, Fangjun Luan^{1,2,3*}, Shuai Yuan^{1,2,3}, Guoqi Liu^{1,2,3}

¹School of Computer Science and Engineering, Shenyang Jianzhu University, Shenyang Liaoning

²Liaoning Province Big Data Management and Analysis Laboratory of Urban Construction, Shenyang Liaoning

³Shenyang Branch of National Special Computer Engineering Technology Research Centre, Shenyang Liaoning

Received: Mar. 12th, 2023; accepted: Apr. 12th, 2024; published: Apr. 18th, 2024

*通讯作者。

文章引用: 张雯涛, 栾方军, 袁帅, 刘国奇. 基于时间敏感的异构图相似神经网络的新闻推荐[J]. 计算机科学与应用, 2024, 14(4): 151-162. DOI: 10.12677/csa.2024.144086

Abstract

The explosive growth of news information on the internet has made personalized news recommendation increasingly important for helping users quickly find news articles of interest. Most mainstream news recommendation methods currently rely on static data, overlooking dynamic information that could indirectly reflect user interests in news, such as the time users spend reading news articles. To address this issue, this study proposes a time-sensitive heterogeneous graph similarity neural network model (TSHGSN) based on the GNewsRec model. The TSHGSN model incorporates user reading time weights into the learning of news features using CNN, and utilizes LSTM to learn the sequence features of user-clicked news as short-term user preferences. Additionally, it constructs a heterogeneous graph to model user-news-topic relationships and aggregates node representations of candidate news features and user long-term preferences using a novel neighbor sampling method. Finally, the model separates user short-term and long-term interests from candidate news for relevance calculation, with the aim of adaptively adjusting the importance of short-term and long-term interests in user modeling. Experimental results show that compared to the GNewsRec model, this model achieves approximately a 4% improvement in the AUC metric.

Keywords

News Recommendation, Graph Neural Network, Short-Term Preference, Long-Term Preference

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

随着互联网和大数据等技术持续发展，这些技术也逐步渗透到人们的生活中，带给人们良好的使用体验，其中推荐模型无时无刻不发挥着作用。新闻推荐可以帮助用户从海量新闻中推送其感兴趣的新闻，提高了时间利用的效率。目前，新闻推荐已在改善用户阅读体验方面取得显著的效果，引起学术界的广泛关注[1] [2]。

之前的新闻推荐算法可分为基于协同过滤的推荐算法、基于内容的推荐算法、混合推荐算法[3] [4]。然而，这些方法存在冷启动问题，当用户 - 新闻矩阵稀疏时，推荐结果往往不够准确。近年来，基于深度学习的方法在建模用户 - 新闻交互方面表现出更好的能力[5] [6] [7]，并广泛应用于新闻推荐领域。其中，通过引入图神经网络来建模用户和新闻之间的交互，GNewsRec [8]成为了目前较先进的新闻推荐方法。然而，该方法仍存在两个主要缺点。首先，它们通常只考虑新闻标题和新闻概要作为新闻推荐的特征，忽略了用户在页面上停留的活跃时间，而这个时间能更好地反映用户的兴趣。其次，GNewsRec 模型通过图神经网络聚合获取用户长期偏好和候选新闻表达时只采用了平均的聚合方法，忽略了用户与用户之间的交互。最后，获取用户偏好时只是简单的将短期兴趣和长期兴趣串联在了一起，忽视了长短期兴趣对用户有不同的影响力。

本文的研究工作受到以下观察结果的启发。首先，用户对新闻文章的兴趣可以通过他们在新闻页面上停留的时间来直接反映。如图 2 所示，用户“cx:i8i85z793m9j”在新闻页面“阿根廷爆冷输沙特”停

留了 116 秒(525 个字),而在新闻页面“日本球迷走后的观众席”停留了 15 秒(534 个字)。我们认为用户更偏好前一条新闻。因此,我们在新闻编码器中引入了时间权重。其次,图神经网络中的用户节点之间可能存在相似的兴趣,而取平均值的聚合方法往往忽略了这一信息。为了更好地表示用户的长期兴趣,我们在聚合过程中计算了节点之间的相似度。此外,将用户短期兴趣和长期兴趣分开与候选新闻进行相关性计算,旨在自适应地调整用户建模中短期兴趣和长期兴趣的重要性。

针对上面的想法本文提出了一种时间敏感的异构图相似神经网络(TSHGSN),旨在通过新闻点击行为和新闻阅读行为来建模用户的兴趣。首先在采用 CNN [6]对新闻进行编码提取新闻特征表示,并引入阅读时间权重。接着使用基于注意力的 LSTM 模型[9]获取用户短期兴趣。然后构建一个用户完整点击新闻历史的用户-新闻-主题异质图,其中主题信息可以缓解用户-新闻交互的稀疏性问题[10]。异质图构建完毕后,使用节点相似的采样方法获取用户长期偏好和候选新闻表示。最后,分别采用不同的评分规则计算候选新闻与用户长期兴趣和短期兴趣之间的匹配程度。

2. 方法

2.1. 新闻推荐问题的定义

针对 TSHGSN 模型,本文给出相关定义如下:给定用户集 $U = \{u_1, u_2, \dots, u_k\}$,新闻集 $I = \{d_1, d_2, \dots, d_m\}$, k 和 m 分别表示用户和新闻的数量。根据用户点击历史,用户浏览新闻的记录 $Su = \{d_{u,1}, d_{u,2}, \dots, d_{u,n}\}$,其中 $d_{u,j}$ 表示用户 u 所点击的第 j 条新闻。令 $Y \in R^{K \times M}$ 表示用户 U 和新闻集 I 之间的用户-新闻交互矩阵,其中 $y_{u,d} = 1$ 表示用户 u 点击了新闻 d ,否则 $y_{u,d} = 0$ 。

本文与 GNewsRec 一致,都从新闻的标题和内容(包括新闻实体和实体类型)中提取新闻特征。每个新闻标题 T 都由一系列单词组成的 $T = \{w_1, w_2, \dots, w_f\}$,其中 w_i 表示新闻标题 T 中的第 i 个单词。新闻内容包括新闻实体 $E = \{e_1, e_2, \dots, e_n\}$ 以及实体类型 $C = \{c_1, c_2, \dots, c_n\}$,其中 c_j 是第 j 个实体 e_j 的类型, n 表示实体的个数。

2.2. 方法总体框架

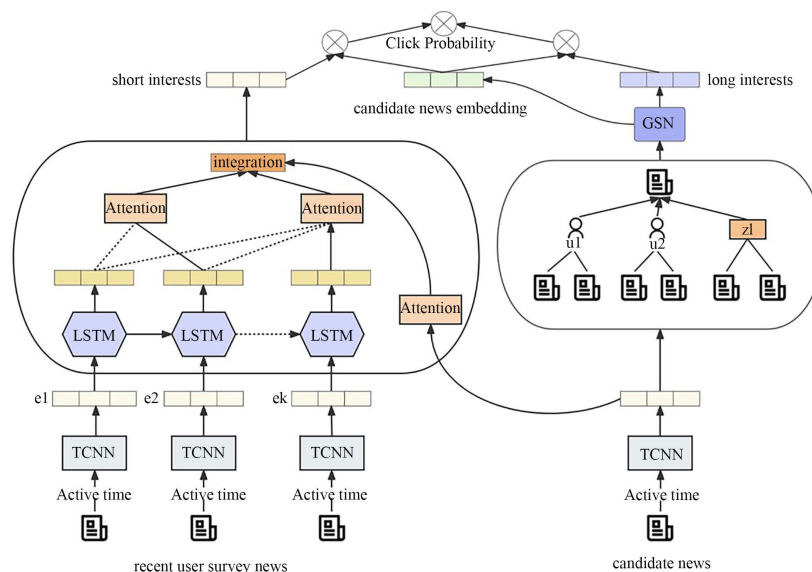


Figure 1. The overall framework of the proposed model

图 1. 所提出方法总体框架

本文基于 GNewsRec 模型，提出 TSHGSN 模型，模型框架如图 1 所示。在图 1 中，TSHGSN 包括新闻编码器，用户短期兴趣建模，用户长期兴趣建模和感知预测四个模块。新闻编码模块使用卷积神经网络从新闻标题、新闻实体、实体类型和用户阅读新闻时间提取新闻特征。用户短期兴趣模块使用基于注意力机制的 LSTM 根据最近点击新闻历史获取用户短期兴趣。用户长期兴趣模块在构建用户 - 新闻 - 主题异质图后，利用 GNN 网络采用节点相似的聚合方法以获取用户长期兴趣和候选新闻表示。感知预测模块将用户短期兴趣和长期兴趣分开与候选新闻进行相关性计算，旨在自适应地调整用户建模中短期兴趣和长期兴趣的重要性。

2.3. 新闻编码器

在新闻特征提取部分，在 GNewsRec 新闻编码器的基础上加入了用户阅读新闻时间权重。

将新闻标题的单词放入词矩阵中来获取新闻标题的嵌入表示

$$\bar{T} = [w_1, w_2, \dots, w_f] \in R^{f \times k_1} \quad (1)$$

其中， w_i 表示新闻标题 T 中第 i 个单词的嵌入向量， k_1 是词矩阵的维度。新闻实体的嵌入表示

$$\bar{E} = [e_1, e_2, \dots, e_n] \in R^{n \times k_2} \quad (2)$$

新闻实体类型的嵌入表示

$$\bar{C} = [c_1, c_2, \dots, c_n] \in R^{n \times k_3} \quad (3)$$

其中， e_j 表示第 j 实体的嵌入向量， c_j 表示与其对应实体类型的嵌入向量， k_2 ， k_3 分别表示新闻实体及实体类型的词矩阵嵌入维度。因为新闻实体及类型的对应关系，将两者拼接起来，得到的新闻内容特征表示

$$\bar{P} = [e_1, f(c_1), e_2, f(c_2), \dots, e_n, f(c_n)] \in R^{2n \times k_2} \quad (4)$$

其中， $f(c) = W_c c$ 表示转换函数， $W_c \in R^{k_2 \times k_3}$ 表示可以训练的矩阵。

此外，新闻页面上的用户的活动时间 $L = [a_1, a_2, \dots, a_m]$ ，隐含地反映了用户对新闻的兴趣。因此，我们将用户在新闻上的活动时间作为输入的额外权重，标题嵌入可以表示为

$$T^\alpha = [w_1 * a_\alpha, w_2 * a_\alpha, \dots, w_f * a_\alpha] \in R^{f \times k_1} \quad (5)$$

其中， $*$ 表示元素间乘积。新闻内容的特征表示

$$P^\alpha = [e_1 * a_\alpha, f(c_1) * a_\alpha, e_2 * a_\alpha, f(c_2) * a_\alpha, \dots, e_n * a_\alpha, f(c_n) * a_\alpha] \in R^{2n \times k_2} \quad (6)$$

分别将加入阅读时间的新闻标题词矩阵 T^α 和新闻内容词矩阵 P^α 送入两个具有独立权重参数的 CNN 中，并行卷积后得到它们的深层特征表示 \tilde{T} 和 \tilde{P} 并连接作为最终的新闻文本特征表示

$$d = f_c \left(\left[\begin{array}{c} \tilde{T} \\ \tilde{P} \end{array} \right] \right) \quad (7)$$

其中， $d \in R^D$ ， D 表示新闻特征表示向量 d 的维度， f_c 表示一个紧密连接层。

2.4. 用户短期兴趣建模

用户在阅读新闻的时，可能会被某些新闻暂时吸引。图 2 是某用户在一段连续时间内的点击新闻序列，用户在时刻 t_{j-5} 阅读一篇关于“卡塔尔世界杯开幕”的新闻，后续就会持续关注世界杯的赛事，显然该用户暂时被世界杯事件吸引，即用户的短期兴趣。

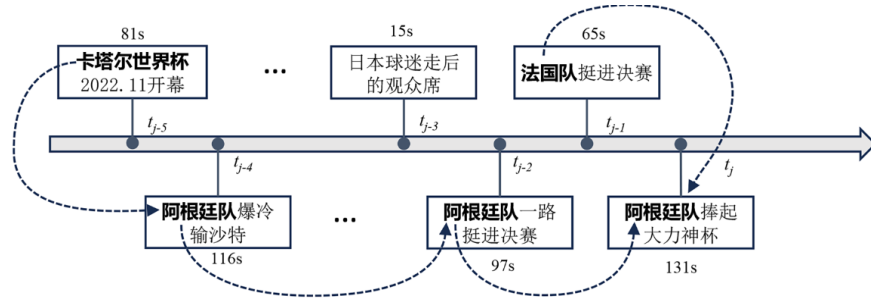


Figure 2. Title sequence of news read by user
图 2. 用户阅读新闻标题序列

用户对于不同新闻的兴趣是不同的(如图 2 所示点击了阿根廷相关的新闻就有三次), 给定一个用户 u , 在 l 最近时间点击新闻 $\{d_1, d_2, \dots, d_l\}$, 使用一个注意机制来赋予不同用户最近点击的新闻对候选新闻 d 的权重:

$$u_i = \tanh(W'd_i + b') \quad (8)$$

$$u = \tanh(Wd + b) \quad (9)$$

$$\alpha_i = \frac{\exp(v^T(u + u_i))}{\sum_i \exp(v^T(u + u_i))} \quad (10)$$

$$u_c = \sum_i \alpha_i d_i \quad (11)$$

其中 u_c 是用户整体兴趣嵌入, α_i 为点击新闻 $d_i (i=1, \dots, l)$ 对候选新闻 d 的影响权重。 $W', W \in R^{D \times D}, d_i^m, b_w, b_i, v^T \in RD, D$ 是新闻的嵌入维度。

应用注意机制对用户当前整体偏好进行建模后, 仍关注用户点击新闻的顺序对用户偏好的影响, 基于此通过基于注意的 LSTM [11] 来捕获用户点击新闻的顺序特征来捕获顺序特征。

在图 1 中, 整个左侧模块是用户的短期兴趣, 本文将用户最近所点击的新闻表示按顺序输入到 LSTM 中, 以获取用户点击新闻的顺序特征。由于每个用户的当前点击行为, 可能会受到先前点击新闻的影响, 将上面提到的新闻内容注意力机制应用于每个隐藏状态 h_i 及其以前的隐藏状态 $\{h_1, h_2, \dots, h_{i-1}\}$ ($h_j = LSTM(h_{j-1}, d_j)$) 的 LSTM, 获得在不同点击时间更丰富的顺序特性表示 $s_i (i=1, \dots, l)$ 。这些内容 (s_1, s_2, \dots, s_l) 由 CNN 集成所获得的, 由此得到用户点击新闻的顺序特征表示 \tilde{s} 。将用户整体兴趣嵌入和顺序级嵌入的串联, 得到最终用户的短期兴趣嵌入:

$$u_s = W_s [u_c; \tilde{s}] \quad (12)$$

2.5. 用户长期兴趣建模

在 GNewsRec 的基础上, 本文亦利用异质图捕捉用户和新闻之间的高阶关系[10]。同样采用 LDA [12] 算法将新闻进行主题分类, 根据用户完整点击新闻历史完成用户 - 新闻 - 主题异质图的构建, 最后利用图相似神经网络(GSN)学习用户和新闻节点的高阶嵌入表示。

通过 GSN 计算以用户为中心节点的一般步骤如下。在聚合前, 新闻已加入阅读时长权重, 由于用户和新闻节点之间都是有一定的相似性的, 因此如图 3 所示, 此处采用了节点相似的邻居采样方法:

$$\alpha_j = \frac{u * d_j}{|u || d_j |} \quad (13)$$

$$E_N = \sum_{d_j \in N(u)} \alpha_j d_j \tag{14}$$

其中用户 u 是通过公式(11)初始化的, $N(u)$ 表示与用户直接相连的新闻, α_j 表示邻居节点和中心节点的余弦相似度。

然后通过邻居节点嵌入对用户长期兴趣嵌入进行更新:

$$u_i = \text{ReLU}(W_i E_N + b_i) \tag{15}$$

其中 $W_i \in R^{D \times D}$ 和 $b_i \in R^{D \times D}$ 分别表示 GSN 第一层的转换权重和偏置。

为了表示新闻 d , 首先计算邻居采样的相似度权重:

$$\partial_k = \frac{e_k * d}{|e_k \| d |} \tag{16}$$

$$d_N = \sum_{e \in S(d)} \partial_k e_k \tag{17}$$

其中 $e \in R^D$ 新闻相邻用户的向量表示, 包括用户 u 和主题 z , 其中用户 u 是通过公式(11)初始化的, 新闻的主题 z 是随机初始化的, 而 d 是用新闻编码器生成的, ∂_k 表示邻居节点和中心节点的余弦相似度。为了保持每批的计算模式固定和更有效, 对每个新闻 d 统一采样一组大小固定的邻居 $S(d)$ 。

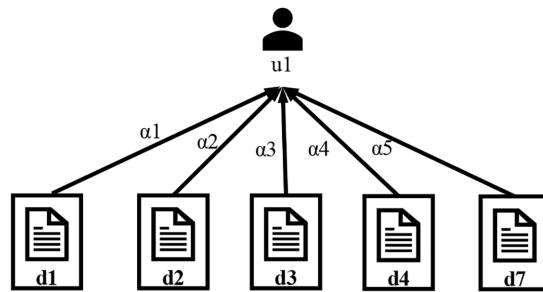


Figure 3. User long-term preference
图 3. 用户长期偏好

然后用 d_N 更新候选新闻表示:

$$\tilde{d} = \text{ReLU}(W_i d_N + b_i) \tag{18}$$

其中 $W_i \in R^{D \times D}$ 和 $b_i \in R^{D \times D}$ 分别表示 GSN 第一层的转换权重和偏置。

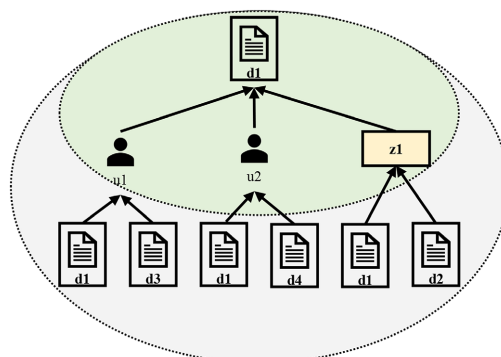


Figure 4. Two-layers GSN
图 4. 二层 GSN

如图 4 所示, 绿色部分是单层 GSN 的聚合过程, 候选新闻仅聚合最近节点。为了捕捉用户和新闻之间的高阶关系, GSN 可以从一层扩展到二层, 更加丰富节点信息。如图 4 所示, 结合灰色部分就可以得到二层 GSN 后的新闻表示。首先, 通过聚合相邻的新闻嵌入来得到与它直接连接的用户嵌入 u_l 和主题嵌入 z , 然后将聚合得到的嵌入 u_l 和 z 聚合以获得两层 GSN 的新闻表示。一般来说, 第一层 GSN 的新闻表示是初始化用户节点和主题节点的混合产物, 直到第二层 GSN 才更为准确。

2.6. 感知预测

GNewsRec 模型中是将长期和短期兴趣串联起来和候选新闻进行相关性计算的, 但是用户长期和短期兴趣对候选新闻的影响可能并不相同。因此, 本文提出了将用户的长期和短期兴趣分开和候选新闻进行相关性的计算, 如此可以更准确地预测用户对新闻的点击。首先计算候选新闻嵌入 \tilde{d} 和用户的长期兴趣嵌入 u_l 之间的相关性:

$$\hat{y}_l = u_l \cdot \tilde{d} \quad (19)$$

再计算候选新闻 \tilde{d} 和用户短期兴趣 u_s 之间的相关性:

$$\hat{y}_s = u_s \cdot \tilde{d} \quad (20)$$

统一的点击分数 \hat{y} 为两个分数的线性组合

$$\hat{y} = \alpha \hat{y}_l + (1 - \alpha) \hat{y}_s \quad (21)$$

其中 α 为一个可学习的参数。

为了训练模型 TSHGSN, 正样本为用户点击过的新闻, 负样本在用户未点击的新闻中随机选择。一个训练样本被表示为 $X = (x, u, y)$, 其中 x 是预测是否点击的候选新闻。对于每个正输入样本, 则为 $y = 1$, 否则为 $y = 0$ 。在训练模型时, 每个负样本的阅读时间权重为 -1 。在测试模型时, 正负样本的阅读时间权重均为训练的模型的平均数。训练模型时每个输入样本都有一个各自的估计概率 $\hat{y} \in [0, 1]$, 表示用户是否会点击候选新闻 x , 使用交叉熵损失作为的损失函数:

$$L = - \left\{ \sum_{X \in \Delta^+} y \log \hat{y} + \sum_{X \in \Delta^-} (1 - y) \log (1 - \hat{y}) \right\} + \lambda \|W\|_2 \quad (22)$$

其中, $\|W\|_2$ 是对所有可训练参数的 L2 正则化, λ 为惩罚权重, Δ^+ 是正样本集, Δ^- 是负样本集。此外, 我们还采用了退出和早期停止, 以避免过拟合。

3. 实验及分析

3.1. 数据集

本文在一个真实的在线新闻数据集 Adressa [13] 上进行实验, 这是一个点击日志数据集, 大约有 2000 万次页面访问, 以及一个 270 万次点击的子样本。

数据集的情况如表 1 所示, Adressa 分为 2 个数据子集: Adressa-1week 和 Adressa-10weeks。分别包含 Adressa 一周和十周的数据。Adressa 是一个基于事件的数据集, 然而并不是所有事件包含新闻阅读时间, 本文使用如下 7 个属性: 会话开始、会话结束、用户 ID、新闻 ID、阅读时间、标题和概要文件, 用于生成数据集。

将新闻按发布时间进行排序, 训练 Adressa-1week 数据集时, 将前 5 天的历史数据用于构建异质图, 训练集第 6 天的历史数据, 验证集为第 7 天的 20%, 测试集为剩余的 80% 的数据。训练 Adressa-10 周的数据集时, 将前 50 天的数据构建异质图, 训练集为的 51~60 天的数据, 验证集为 61~62 的数据, 测试集

为最后 8 天。

Table 1. Description of experiment datasets

表 1. 实验数据详细信息

类别	Adressa-1week	Adressa-10weeks
用户数	267,718	1,013,972
新闻数	9426	48,007
事件数	1,321,557	11,655,367
实体类型	11	11
标题平均单词数	4.87	4.99
新闻平均实体数	21.79	20.98

为了减少文本数据噪声，在放入词矩阵前：通过挪威语停顿词文档从新闻标题中删除无用的停顿词，并去除新闻概要文件中的无意义实体和实体类型。

3.2. 实验环境和结果对比

与 GNewsRec 模型一致，本文亦采用 AUC 和 F1 作为评价指标。模型中新闻标题、实体及实体类型的词矩阵维度 k_1 、 k_2 和 k_3 均设置为 50，新闻、用户和主题的词向量嵌入维度均设为 128，用户短期兴趣模块里点击的新闻数 m 设为 10。在异质图中 LDA 中主题数与 GNewsRec 一致设为 20，聚合用户节点时 L_u 设置为 10，聚合新闻节点时 L_n 设置为 30。本文高斯分布进行随机初始化参数，其中高斯分布均值设为 0、标准差设 0.1，并采用学习率为 0.0003，失活率为 0.5 的 Adam [14] 对参数进行优化。模型在训练过程中，采用验证集进行验证。上述参数是根据模型在验证集上性能最优时确定的，保留在验证集上最优的模型，并在测试集上进行测试。

为了验证模型的性能，选择如下经典的推荐算法与近期效果突出的新闻推荐算法进行对比实验。

(1) Wide & Deep (Wide & Deep Learning for Recommend Systems) [15] 经典推荐模型，同时兼顾记忆和泛化能力，既可以利用广义线性模型来记忆历史数据的模式，又可以利用深度神经网络来学习数据的抽象特征表示，从而提高推荐系统的预测准确性和推荐效果。

(2) DeepFM (Factorization-Machine Based Neural Network) [16] 经典推荐模型，由线性模型和深度学习两个模块组成，结合广度和深度将 FM 和神经网络相结合用于推荐模型。

(3) DKN (Deep Knowledge-Aware Network for News Recommendation) [5] 经典新闻推荐模型，提出用并行 CNN 提取新闻特征和实体，并使用 AM 提取用户的整体偏好。

(4) DAN (Deep Attention Neural Network for News Recommendation) [6] 基于深度注意力神经网络，将新闻标题和新闻实体及类型通过 CNN 并行提取，并根据用户点击历史使用注意力机制获取用户整体偏好，同时使用基于注意的 LSTM 获取用户时序偏好。

(5) GNewsRec (Graph News Recommendation with Long-term and Short-term Interesting Model) [8] 结合用户长期短期兴趣，基于图神经网络构建用户长期兴趣，采用 LSTM 结合注意力机制建模用户短期兴趣，串联用户长期兴趣和短期兴趣形成用户的最终表示。

(6) GNUD (Graph Neural News Recommendation with Unsupervised Preference Disentangle) [17] 在新闻-用户图结构的基础上，加入偏好解缠正则化以动态识别用户点击新闻的偏好因素，进一步提高模型性。

各模型在两个数据集上的指标值对比结果如表 2 所示，可得出如下结论：(1) 经典推荐模型在小数据

集上的表现普遍不如新闻推荐模型，可是新闻推荐模型都加入了注意力机制的缘故，可以更加准确的表达用户的偏好。表中，新闻推荐模型在大数据集上的性能均低于小数据的性能，可能是大数据记录的用户 10 周阅读的新闻，由于新闻的时效性，时间较久的新闻可能会对用户兴趣建模造成干扰。(2) 相比于其他模型 TSHGSH 各项指标值均最高。这主要是由于本文同时考虑了用户点击行为和用户阅读行为来提取深度新闻表示，利用 GNN 在聚合时采用邻居相似的聚合方法的结果。

Table 2. Performance comparison of different models

表 2. 个模型对比

模型	Adressa-1week		Adressa-10week	
	AUC	F1	AUC	F1
Wide & Deep	68.35	69.33	73.28	69.53
DeepFM	69.23	62.01	74.04	65.82
DKN	75.57	76.95	74.32	72.29
DAN	75.93	74.01	76.76	71.64
GNewsRec	81.26	82.85	78.62	81.01
GNUD	84.01	83.90	83.21	81.09
TSHGSN	85.37	84.93	84.16	82.02

3.3. 参数影响分析

本节进行多次实验，以验证参数对实验结果的影响，在 Adressa-1week、Adressa-10weeks 数据集上分析模型中训练集负样本阅读时长参数对实验结果的影响和 GSN 层数对实验结果的影响。

Table 3. Experimental results of different negative sample reading time values

表 3. 不同负样本阅读时间取值

模型	Adressa-1week		Adressa-10week	
	AUC	F1	AUC	F1
TSHGSN _{-0.00001}	79.94	79.29	78.84	77.98
TSHGSN ₋₀	82.37	81.63	81.21	80.28
TSHGSN ₋₁	85.37	84.93	84.16	82.02

Table 4. Experiment results of different GSN layers

表 4. 不同 GSN 层数实验结果

模型	Adressa-1week		Adressa-10week	
	AUC	F1	AUC	F1
TSHGSN ₋₁	80.09	78.73	80.57	77.61
TSHGSN ₋₂	85.37	84.93	84.16	82.02
TSHGSN ₋₃	81.92	80.76	80.23	79.37

训练集负样本的不同阅读时长对实验结果的影响如表 3 所示，由表 3 可看出，本文测试了三个常用参数分别是 0.00001, 0, -1。当参数 0.00001 和 0 时，卷积后由于常量参数 b 的影响，对最后模型训练

结果还是有一定的干扰。当参数为-1时，卷积后加上常量 b 时，负样本的结果接近于 0，干扰最小并且 AUC 和 F1 表现最优。不同 GSN 层数对实验结果的影响如表 4 所示，由表 4 可看出，本文将 GSN 层数分别设置为 1、2、3，当 GSN 层数为 2 时，两个评价指标效果最优。这是因为一层 GSN 只是初始化节点聚合的混合物，而三层 GSN 可能会给模型带来更多的噪声影响模型的性能。

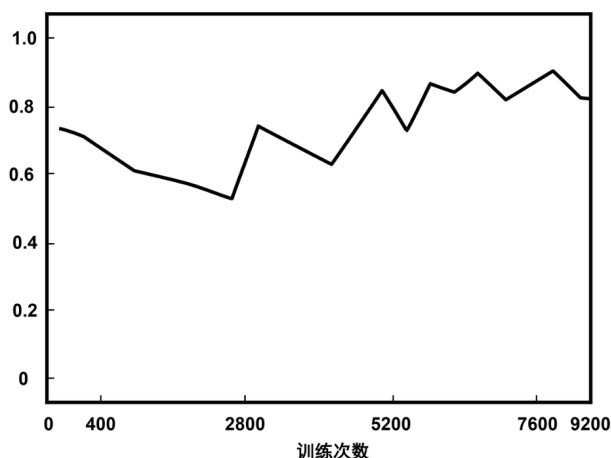


Figure 5. Variation curve of learnable parameter α
图 5. 可学习参数 α 变化曲线

学习参数 α 的变化曲线如图 5 所示。由图可知，当 $\alpha = 0.85$ 在验证集上效果最优。这说明用户长期兴趣在用户兴趣建模中占主导作用，用户短期兴趣在用户兴趣中起到补充作用。

3.4. 消融实验分析

本节为了验证各模块的有效性进行消融实验，定义如下三个子模型。

- 1) TSHGSN_{-time} 表示从 TSHGSN 模型里去掉新闻阅读时间
- 2) TSHGSN_{-short} 表示从 TSHGSN 模型里去掉短期偏好
- 3) TSHGSN_{-GSN} 表示用普通的平均邻居采样方法

各模型在 Adressa-1week 和 Adressa-10week 实验的指标值如表 5 所示，表中黑色字体为最优值。由表 5 可看出，当 TSHGSN 删除阅读时长的权重后，AUC 和 F1 出现大幅下降，这说明阅读新闻时长可以有效表达用户的阅读满意度，能够提高推荐的准确性。当 TSHGSN 删除用户短期兴趣建模模块后，AUC 和 F1 下降约 2%，这表明用户短期兴趣建模模块的有效性，应同时考虑用户的长期兴趣和短期兴趣。当 TSHGSN 删除 GSN 后，使用平均的聚合方法时，AUC 和 F1 下降了 1%，这验证 GSN 对 TSHGSN 的提升效果。综上所述，本文提出的各模块均是有效的。

Table 5. Result of ablation experiment

表 5. 消融实验结果

模型	Adressa-1week		Adressa-10week	
	AUC	F1	AUC	F1
TSHGSN _{-time}	83.16	82.29	80.94	79.87
TSHGSN _{-short}	83.45	82.55	81.76	80.82
TSHGSN _{-GSN}	84.58	83.95	83.32	81.29
TSHGSN	85.37	84.93	84.16	82.02

4. 总结

本文提出基于时间敏感的异质图相似神经网络的新闻推荐模型。TSHGSN 算法使用卷积神经网络(CNN)学习新闻特征,并结合用户阅读新闻的时间权重,利用长短期记忆网络(LSTM)学习用户点击新闻的序列特征,以捕捉用户的短期兴趣。同时,构建了一个异构图,并使用节点相似性的方法来聚合节点,获取候选新闻特征表示和用户长期兴趣。最后,用户短期兴趣和长期兴趣与候选新闻分开比较,旨在自适应地调整用户建模中短期兴趣和长期兴趣的重要性。经过对比试验,证明了其在推荐任务上的优异表现。此外,还进行了多个消融实验以检验模型中某些组件的有效性,并分析了不同参数值的设置对算法效果的影响。今后将考虑引入图注意力网络,提高新闻推荐的准确性。

基金项目

国家自然科学基金青年基金项目(61602323)。

参考文献

- [1] Ferrell, M., Ferrell, S., Ottwell, R., *et al.* (2021) Superlative Use within News Articles Relating to Therapies for Multiple Sclerosis. *Multiple Sclerosis and Related Disorders*, **49**, 102736. <https://doi.org/10.1016/j.msard.2021.102736>
- [2] Wu, F., Qiao, Y., Chen, J.-H., *et al.* (2020) Mind: A Large-Scale Dataset for News Recommendation. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 5-10 July 2020, 3597-3606. <https://doi.org/10.18653/v1/2020.acl-main.331>
- [3] 田萱, 丁琪, 廖子慧. 基于深度学习的新闻推荐算法研究综述[J]. *计算机科学与探索*, 2021, 15(6): 971-998.
- [4] Wu, C., Wu, F., Huang, Y., *et al.* (2023) Personalized News Recommendation: Methods and Challenges. *ACM Transactions on Information Systems*, **41**, Article No. 24. <https://doi.org/10.1145/3530257>
- [5] Wang, H., Zhang, F., Xie, X., *et al.* (2018) DKN: Deep Knowledge-Aware Network for News Recommendation. *Proceedings of the World Wide Web Conference*, 23-27 April 2018, 1835-1844. <https://doi.org/10.1145/3178876.3186175>
- [6] Zhu, Q., Zhou, X., Song, Z., *et al.* (2019) DAN: Deep Attention Neural Network for News Recommendation. *Proceedings of the AAAI Conference on Artificial Intelligence*, **33**, 5973-5980. <https://doi.org/10.1609/aaai.v33i01.33015973>
- [7] Wu, C., Wu, F., An, M., *et al.* (2019) NPA: Neural News Recommendation with Personalized Attention. *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, Barcelona, 4-8 August 2019, 2576-2584. <https://doi.org/10.1145/3292500.3330665>
- [8] Hu, L., Li, C., Shi, C., *et al.* (2020) Graph Neural News Recommendation with Long-Term and Short-Term Interest Modeling. *Information Processing & Management*, **57**, Article ID: 102142. <https://doi.org/10.1016/j.ipm.2019.102142>
- [9] An, M., Wu, F., Wu, C., *et al.* (2019) Neural News Recommendation with Long-and Short-Term User Representations. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, Florence, 28 July-2 August 2019, 336-345. <https://doi.org/10.18653/v1/P19-1033>
- [10] Lee, D., Oh, B., Seo, S., *et al.* (2020) News Recommendation with Topic-Enriched Knowledge Graphs. *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*, 19-23 October 2020, 695-704. <https://doi.org/10.1145/3340531.3411932>
- [11] Hochreiter, S., Schmidhuber, J. (1997) Long Short-Term Memory. *Neural Computation*, **9**, 1735-1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- [12] Blei, D.M., Ng, A.Y. and Jordan, M.I. (2003) Latent Dirichlet Allocation. *Journal of Machine Learning Research*, **3**, 993-1022.
- [13] Gulla, J.A., Zhang, L., Liu, P., *et al.* (2017) The Adressa Dataset for News Recommendation. *Proceedings of the International Conference on Web Intelligence*, Leipzig, 23-26 August 2017, 1042-1048. <https://doi.org/10.1145/3106426.3109436>
- [14] Kingma, D.P. and Ba, J. (2014) Adam: A Method for Stochastic Optimization. arXiv preprint arXiv:1412.6980.
- [15] Cheng, H.-T., Koc, L., Harmsen, J., *et al.* (2016) Wide & Deep Learning for Recommender Systems. *Proceedings of the 1st Workshop on Deep Learning for Recommender Systems*, Boston, 15 September 2016, 7-10.

<https://doi.org/10.1145/2988450.2988454>

- [16] Guo, H., Tang, R., Ye, Y., *et al.* (2017) DeepFM: A Factorization-Machine Based Neural Network for CTR Prediction. arXiv preprint arXiv:1703.04247,. <https://doi.org/10.24963/ijcai.2017/239>
- [17] Hu, L., Xu, S., Li, C., *et al.* (2020) Graph Neural News Recommendation with Unsupervised Preference Disentanglement. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 5-10 July 2020, 4255-4264. <https://doi.org/10.18653/v1/2020.acl-main.392>