

社交网络中基于注意力机制的网络舆情事件演化趋势预测

桑春艳¹ 胥文¹ 贾朝龙¹ 文俊浩²

1 重庆邮电大学软件工程学院 重庆 400065

2 重庆大学大数据与软件学院 重庆 401331

摘要 与传统媒体相比,社交网络在传播新闻、思想、观点等方面发挥着突出的作用,同时也是传播谣言、虚假新闻等负面信息的最佳途径。因此,对网络舆情演化趋势的准确预测和有效控制已成为重要的研究话题。目前,大多数研究从理论建模的角度对网络舆情事件的演化特性和发展趋势进行预测,基于用户行为特征的信息传播演化趋势预测模型的建模及分析有待进一步研究。考虑到信息传播过程中用户之间的相互影响,文中提出一种基于注意力机制的方法,旨在探究社交网络中用户在信息传播过程中的影响来预测信息的传播趋势。首先,利用基于长短时记忆神经网络(Long Shot-Term Memory,LSTM)的网络架构来获取信息传播的轨迹特征。其次,考虑到信息传播和用户行为的复杂性,利用注意力机制挖掘用户之间的依赖性来预测真实的信息传播过程。最后,综合考虑影响信息传播的驱动因素,得到一种基于注意力机制的信息传播演化趋势预测模型(Attention Diffusion Neural Network,ADNN)。在4个对比数据集上的实验结果显示,ADNN模型优于流行的序列模型,该模型能够有效利用驱动因素对信息传播的影响,更准确地预测社交网络中信息的传播趋势。

关键词:信息传播;趋势预测;注意力机制;深度学习

中图法分类号 TP311

Prediction of Evolution Trend of Online Public Opinion Events Based on Attention Mechanism in Social Networks

SANG Chun-yan¹, XU Wen¹, JIA Chao-long¹ and WEN Jun-hao²

1 School of Software Engineering,Chongqing University of Posts and Telecommunications,Chongqing 400065,China

2 School of Big Data & Software Engineering,Chongqing University,Chongqing 401331,China

Abstract Compared with traditional media,social networks play a prominent role in disseminating news,ideas,and opinions,and are also the best way to spread negative information such as rumors and false news. Therefore,accurate prediction and effective control of the evolution trend of online public opinion have become important research topics. At present,most studies predict the evolution characteristics and development trends of online public opinion events from the perspective of theoretical modeling. The modeling and analysis of information dissemination evolution trend prediction models based on user behavior characteristics need to be further studied. Considering the interaction between users in the process of information dissemination,this paper proposes a method based on the attention mechanism,which aims to explore the evolution trend prediction of information dissemination in social networks. Firstly,a network architecture based on long shot-term memory (LSTM) is used to obtain the trajectory characteristics of information propagation. Secondly,considering the complexity of information dissemination and user behavior,the attention mechanism is used to mine the dependence between users to predict the real information dissemination process. Finally,we comprehensively consider the driving factors that affect information dissemination,and obtain an attention diffusion neural network (ADNN) based on the attention mechanism. The experimental results on the four comparative data sets show that the ADNN model is better than the popular sequence model. This model can effectively use the influence of driving factors on information dissemination,and more accurately predict the trend of information dissemination in social networks.

Keywords Information dissemination,Trend prediction,Attention mechanism,Deep learning

1 引言

移动终端设备的广泛应用使其在线社交网络成为用户传

播、交流、共享和收集信息的主要平台^[1-4]。社交网络中关于信息传播的研究已经成为学术界和工业界广泛关注的热点话题。在线社交网络在传播新闻、思想、观点等方面发挥着突出

到稿日期:2020-06-24 返修日期:2020-10-09

基金项目:国家自然科学基金(61672117);重庆市自然科学基金面上项目(cstc2019jcyj-msxmX0588)

This work was supported by the National Natural Science Foundation of China(61672117) and Research project of Chongqing CSTC(cstc2019jcyj-msxmX0588).

通信作者:桑春艳(sangcy@cqupt.edu.cn)

作用。例如,在新冠肺炎疫情防控期间,政府机构利用官方微信、微博、网络新媒体等平台及时发布官方信息,有效阻止了虚假、有害信息的传播。同时,社交平台也为谣言、虚假新闻、有害信息的广泛传播提供了途径。因此,通过对社交网络中的用户行为数据的分析建模,不仅可以挖掘社交网络中信息传播的特点和演化规律,还可以预测社交网络中信息传播的演化趋势,从而对信息传播过程进行有效干预和实时控制。

针对社交网络中用户的个体行为具有很大的随机性和不确定性,Xiong等提出一种改进的病毒传播模型,模型中有些节点可以在随机长度的等待时间段内暂停其传播行为,随后这些节点可能会恢复活动,同时在模型中引用信息的转发机制^[5]。Lagnier等提出了一种将用户的个人资料和用户行为相结合的传播概率模型^[6]。Wang等利用用户关系图的结构特征结合循环神经网络来进行信息传播预测^[7]。Wang等将动态有向无环图和循环神经网络结合,提出了一种拓扑结构的循环神经网络模型来进行信息传播预测^[8]。Bourigault等假设所有用户接收信息都是从初始的信息源头获得,与真实的信息传播过程存在差异,通过计算用户特征向量中的距离来分析用户之间相互影响的可能性大小,从而预测信息的传播过程^[9]。实际上,在线社交网络的信息传播过程中,用户之间错综复杂的关系共同影响着传播过程。

现实中获取详细的图结构信息通常较难甚至是不可能的^[10],而注意力机制可以达到通过一部分图像来推测整体的效果^[11-12]。Luong等^[13]利用注意力机制改善了机器翻译任务,将注意力关注于翻译词语对应的上下文,从而更准确地翻译整个句子。Park等^[14]提出一种利用LSTM预测轨迹序列的模型。鉴于LSTM在序列模型中的良好效果,本文针对社交网络中信息传播的特点,提出融合注意力机制和LSTM的混合模型。为了衡量用户对信息传播过程的影响,该模型利用注意力机制关注用户对整个传播过程的影响,更好地预测信息传播过程。本文的主要贡献如下:

(1)所提模型不需要借助现实中难以获取的图结构,而是直接利用传播序列进行预测,更加符合真实的信息传播过程。

(2)为充分考虑信息传播过程中用户的影响,所提模型利用注意力机制来挖掘用户在信息传播过程中的影响,进而预测真实的信息传播过程。

2 相关工作

在一个社交网络中给定一条消息,描述该消息在社交网络中的传播过程并对其进行预测的研究大致可以分为两类。一类方法借助用户与用户之间的关系,即图结构信息,假设社交网络中的用户只有两种状态:未接收信息的人处于非激活状态和已经接收信息的激活状态,如独立级联模型(Independent Cascade, IC)^[15]、线性阈值模型(Linear Threshold, LT)^[16]等。在IC中,模型抽象了社交网络中人与人之间的行为影响,用边上的权重来表示人与人之间产生影响的概率。与IC模型相比,LT模型的边上不仅有权重,每个节点还存在一个对应的阈值,当节点周围的被激活节点的边权之和达到或超过阈值时,节点也会变为激活状态。考虑到信息在网络中两个相连节点间传播时用连续时间的密度函数表示传播概率,Rodriguez等^[17]对IC模型进行了时间的连续化推广,将

离散的信息传播转化为连续时间上的传播问题,节点的激活可以发生在任何连续时间内,但是,在没有先验图信息指导的情况下,其模型推断方法无法避免构造不现实的传播过程,尤其是当所观察到的传播过程不足或有偏差时。

由于现实中图结构信息难以获取,另一类方法不需要借助用户与用户之间的关系。Rodriguez等^[18]跟踪大量的博客和新闻文章来推断信息如何通过社交网络传播,并提出了一种不需要假设传播机制的方法NETINF,该模型能够最大化训练数据中激活用户的似然值,从而找出信息传播的过程,但其实验的复杂度和结果的准确性还有待提高。Bourigault等^[9]将表示学习应用在信息传播中,为信息传播问题提供了新的视角,将时序信息映射到一个连续的空间中并结合扩散核函数预测信息传播,取得了良好的效果。然而,其假设所有用户接收的信息都是从初始的信息源头获得,这显然与真实的信息传播过程不符。随后Bourigault等^[19]提出了另一个基于IC的表示学习模型,使用用户嵌入来表示信息传播的概率,将传播图模型变成传播空间模型。虽然模型与信息传播过程更加一致,但潜在空间中缺少先验图结构信息仍然导致该模型的泛化能力较弱。

3 基于注意力机制的双向LSTM信息传播模型

信息在社交网络中的传播过程中可以看作一个序列,假设用户集合用 U 来表示,用 q 代表某个信息的传播过程, $q = \{(u_0, t_0), (u_1, t_1), \dots, (u_{|q|-1}, t_{|q|-1})\}$,其中, $u_j \in U$ 表示第 j 个用户在某个传播序列 q 中, t_j 表示用户 u_j 接收事件的时间, $|q|$ 表示传播序列 q 的长度。本文的模型是学习给予的部分信息传播序列 $\{u_0, u_1, \dots, u_j\}$,传播序列在时间 t_j 之前按照接收信息的时间排序,模型的目标是预测下一时间 t_{j+1} 可能接收此信息的用户 u_{j+1} 。

本文提出的基于注意力机制的信息传播预测模型(ADNN)的框架如图1所示。

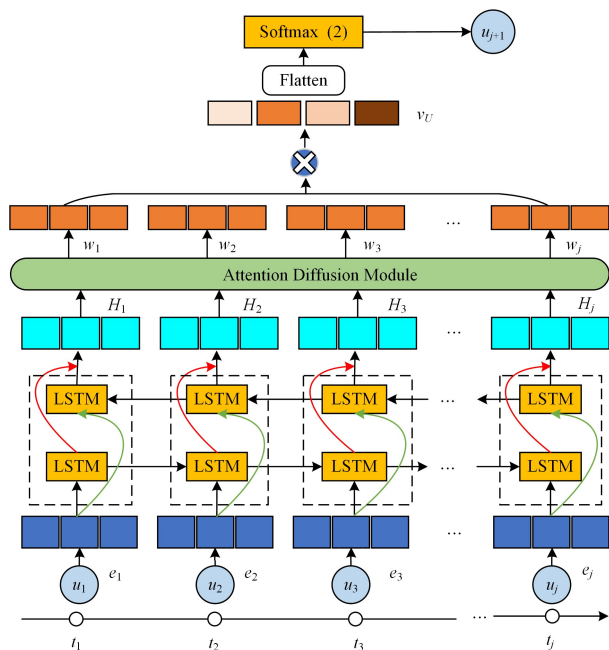


图1 文中提出的模型框架

Fig. 1 Model framework proposed in this paper

首先将用户嵌入向量,通过 LSTM 网络获取信息传播过程的特征。然后通过 Attention Diffusion Module 计算历史用户对信息传播过程的潜在影响,捕获用户在信息传播过程中的内在重要性来预测信息传播过程。在时间 t_n ,模型的输入为信息传播序列 $q_i = \{u_1, u_2, \dots, u_n\}$ 。为了更好地表达用户特征,每个输入用户的原始表示为 $u_i \in \{u_1, u_2, \dots, u_n\}$,将序列中的用户用向量表示为 $e_i = \text{emb}(u_i) \in \mathbb{R}^d$, d 为向量的维度。

3.1 基于双向 LSTM 的信息传播模型

采用长短时记忆(LSTM)方法可以提取信息传播中更复杂的状态,相比经典的概率方法,LSTM 在序列学习任务中表现非常好^[20]。LSTM 中存在 3 种门控机制, c_t 为记忆细胞,可以处理循环神经网络中的梯度衰减问题,并且可以获取时间序列中距离较大的依赖关系。遗忘门 f_t 的目的是决定上一时刻的单元状态 c_{t-1} 有多少保留到当前时刻 c_t 。输入门 i_t 的目的是决定当前时刻网络的输入 x_t 有多少保存到单元状态 c_t 。输出门 o_t 的目的是控制单元状态 c_t 有多少输出到 LSTM 的当前输出值 h_t 。隐含状态向量 h_t 包含输入序列的隐含解释性因素,可以利用 h_t 来预测下一个可能接收信息的用户。具体描述如下:

$$\begin{aligned} i_t &= \sigma(x_t W_{xi} + h_{t-1} W_{hi} + b_i) \\ f_t &= \sigma(x_t W_{xf} + h_{t-1} W_{hf} + b_f) \\ o_t &= \sigma(x_t W_{xo} + h_{t-1} W_{ho} + b_o) \\ \tilde{c}_t &= \tanh(x_t W_{xc} + h_{t-1} W_{hc} + b_c) \\ c_t &= f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \\ h_t &= o_t \odot \tanh(c_t) \end{aligned}$$

其中, x_t 为当前时间的输入; $\tanh$ 和 σ 为激活函数; $W_{xi}, W_{xf}, W_{xc}, W_{xo}$ 和 $W_{hi}, W_{hf}, W_{ho}, W_{hc}$ 为权重参数; b_i, b_f, b_o, b_c 为偏差参数;这些参数都会在神经网络中训练。

如图 2 所示,双向的 LSTM 网络架构可以从正反两个方向处理输入,每个时间步的隐藏状态取决于该时间步之前和之后的信息,隐含状态计算如下:

$$H_t = \vec{h}_{f_t} \oplus \overleftarrow{h}_{b_t} \quad (2)$$

其中, $\vec{h}_{f_t}$ 为前向 LSTM 的隐含状态, $\overleftarrow{h}_{b_t}$ 为后向 LSTM 的隐含状态。这种结构的双向性增加了模型的容量和灵活性。

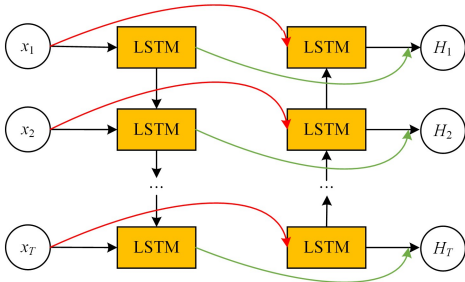


图 2 双向 LSTMNNs 的结构

Fig. 2 Structure of bidirectional LSTMNNs

3.2 基于注意力机制的双向 LSTM 信息传播预测

注意力机制旨在捕获输入用户之间在信息传播中的依赖关系,并提取依赖关系的用户特征。信息传播中的依赖关系指信息传播过程中用户之间激活的概率。

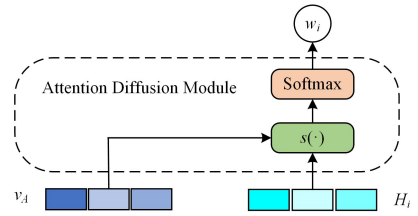


图 3 注意力机制在模型中的应用

Fig. 3 Attention mechanism applied in our model

将信息传播的用户嵌入向量作为输入,通过 LSTMNNs 来提取信息传播的轨迹特征。

$$[H_1; \dots; H_n] = \text{LSTMNNs}([e_1; \dots; e_n]) \quad (3)$$

本文建立注意力机制,利用信息传播的轨迹特征来计算每个用户权重 w ,其中 v_A 为候选用户的嵌入向量表示, v_U 为信息项对于当前候选用户的传播权重表示。

$$v_U = \text{attention}([H_1; \dots; H_n], v_A) = \sum_{j=1}^n w_j H_j \quad (4)$$

权重 w 通过使用轨迹特征 $\{H_1, \dots, H_n\}$ 和 v_A 计算,表示如下:

$$w_i = \text{softmax}(s(H_i, v_A)) = \frac{\exp(s(H_i, v_A))}{\sum_{j=1}^n \exp(s(H_j, v_A))} \quad (5)$$

其中, $s(H_i, v_A)$ 计算第 i 个轨迹特征和 v_A 的依赖程度, D 为输入向量 H_i 的维度。

$$s(H_i, v_A) = \frac{H_i^T v_A}{\sqrt{D}} \quad (6)$$

本文观察一个特定信息传播序列,其传播概率就是用户 u_{j+1} 接收传播轨迹中事件的概率:

$$\hat{p}(u_{j+1}) = \sigma(W_p v_U + b_p) \quad (7)$$

其中, W_p 和 b_p 分别表示权重参数和偏置参数, $\hat{p}(u_{j+1})$ 表示下一个可能接收信息的用户的概率。本文使用交叉熵作为损失函数:

$$L = - \sum_i (p(u_i) \log \hat{p}(u_i) + (1 - p(u_i)) \log (1 - \hat{p}(u_i))) \quad (8)$$

其中, $p(u_i)$ 和 $\hat{p}(u_i)$ 分别表示真实的标签和预测的结果。模型通过利用随时间反向传播算法 (BPTT) 来进行模型训练,在训练过程中最小化损失函数 L ,训练参数通过使用小批量 (mini-batch) 的 Momentum 优化器来更新。

3.3 算法描述

本文中,基于注意力机制的双向 LSTM 信息预测方法通过有效获取用户之间在信息传播中的依赖关系,来对在线社交网络中的舆情事件进行预测。提出的 ADNN 算法的描述如算法 1 所示。

算法 1 ADNN

输入: 用户集合 U , 传播轨迹集合 $\Omega = \{q_1, q_2, \dots, q_n\}$

输出: 模型评价指标 Precision@k, Recall@k, F1-score@k, AUC

1. 模型初始化选择参数
2. Repeat
3. For q in Ω // 迭代传播轨迹集合
4. $e_i = \text{emb}(q)$ // 转为嵌入向量
5. $H_i = \text{LSTMNNs}(e_i)$ // 获取传播的轨迹特征
6. $w_i = \text{softmax}(s(H_i, v_A))$ // 计算轨迹特征和用户之间的权重值
7. $v_U = \sum_{j=1}^n w_j H_j$ // 计算用户在传播轨迹中的权重

8. $\hat{p}(u_{j+1}) = \sigma(\mathbf{W}_p \mathbf{v}_U + \mathbf{b}_p)$ // 计算用户的传播概率
9. $L = -\sum_i (p(u_i) \log \hat{p}(u_i) + (1 - p(u_i)) \log (1 - \hat{p}(u_i)))$ // 计算误差
10. 根据 L 更新参数
11. Until convergence // 收敛或最大迭代步
12. 输出模型评价指标

4 实验

本实验的操作系统为 Windows 10,硬件环境为 Intel® Core™ i7-9700(3.0 GHz),RAM 为 32 GB,实验使用 Tensorflow 来实现,模型运行在显卡 Nvidia GTX 1060(6 GB)上,基于 CUDA 9.0、Tensorflow 1.13.1。为了评估所提方法在信息传播过程中的效果,本文采用不同的评价方法和不同的真实数据集,通过与相关方法进行对比实验,来验证本文所提模型的有效性。

4.1 数据集描述

级联(cascade)是一个信息项通过一组用户传播的过程,每个级联均包含(user,timestamp)的列表对,列表对表示用户在时间戳上被激活。文中所采用的 4 个数据集的统计描述如表 1 所列。

表 1 数据集统计
Table 1 Statistics of datasets

Datasets	User	Link	Cascade
Memetracker	498	158 194	8 304
Last. FM	1 892	12 523	186 479
Weibo	8 190	43 365	925 009
Irvine	540	62 605	471

Last. FM:一个流行的音乐网站^[21]。该数据集包含来自 Last. fm 在线音乐网站的大约 2 000 名用户收听的音乐艺术家信息。本文将每个艺术家视为信息传播项,并删除收听记录不超过 5 次的用户。

Weibo:此数据集包含从新浪微博抓取的信息发布记录^[22],每个信息的发布记录代表一个信息传播过程,其中用户按照信息发布的时间进行排序,并过滤不活跃的用户,以减轻数据噪声对模型的影响。

Memetracker:该数据集收集了一百万条新闻故事和博客文章,并跟踪最常用的引语和短语^[23](也称作模因),来研究模因在人群中的传播。每个模因都被视为信息传播项,网站或博客的每个 URL 被视为用户。本文仅保留了最活跃的 URL,以减轻数据噪声对模型的影响。

Irvine:该数据集收集于加州大学欧文分校学生在线社

区^[24]。学生可以在不同的论坛上参与和撰写帖子。本文将每个论坛视为一个信息项,并删除参加不超过 5 个论坛的用户。

4.2 对比算法和评价方法描述

RNN:基本的循环神经网络模型^[25]。

LSTM:长短期记忆神经网络,一种广泛用于序列数据建模的神经网络框架。先前已经有人使用 LSTM 进行信息传播的预测^[26]。

GRU:门控神经网络^[27],是 LSTM 的一种简化变形。

Self-Attention^[28]:在序列数据内部做注意力计算,寻找序列内部的联系,不需要额外的外部信息。

将本文模型与 4 种对比方法进行对比,以此来评估本文模型在信息传播任务上的表现。在模型中,本文定义 $S_u(k)$ 为信息项的前 k 个预测传播用户列表, G_u 为真实的下一时间的信息传播用户列表。在测试集中共有 M 个信息项。具体评价指标如下。

precision@k:计算预测用户列表的精度。

$$precision@k = \frac{1}{M} \sum_{u=1}^M \frac{|S_u(k) \cap G_u|}{k}$$

(9)

recall@k:计算预测用户列表的召回率。

$$recall@k = \frac{1}{M} \sum_{u=1}^M \frac{|S_u(k) \cap G_u|}{|G_u|}$$

(10)

F1-score@k:计算准确率(precision)和召回率(recall)的加权平均,最大值为 1,最小值为 0,该值越大意味着模型越好。

$$F1_score@k = \frac{2 \cdot precision@k \cdot recall@k}{precision@k + recall@k}$$

(11)

AUC:ROC 曲线下的面积,横坐标为 FPR,代表将负例分类为正例的概率;纵坐标为 TPR,代表能将正例正确分类的概率。曲线下面积越大,模型的学习效果越好。

$$AUC = \frac{1}{|Q^{Test}|} \sum_{q \in Q^{Test}} \frac{1}{|U_q^+| |U_q^-|} \sum_{i \in U_q^+} \sum_{j \in U_q^-} \delta(\hat{p}_{q,i} > \hat{p}_{q,j})$$

(12)

其中, Q^{Test} 为测试信息序列集; $\hat{p}_{q,i}$ 表示在某个信息传播序列 q 中用户 i 接收信息的概率; $|U_q^+|$ 和 $|U_q^-|$ 分别表示传播序列 q 的接收信息的用户集和未接收信息的用户集;当 $\hat{p}_{q,i} > \hat{p}_{q,j}$ 时函数 $\delta(\hat{p}_{q,i} > \hat{p}_{q,j})$ 返回 1,否则返回 0。

4.3 实验结果

本实验在 Memetracker, Irvine, Last. FM, Weibo 数据集上进行了验证,实验结果如表 2 和图 4 所示。

表 2 各模型在真实数据集上的表现

Table 2 Performance of each model on real data

Model	Memetracker			Irvine			Last. FM			Weibo		
	P@3	R@3	F1@3	P@3	R@3	F1@3	P@3	R@3	F1@3	P@3	R@3	F1@3
RNN	0.4842	0.7263	0.5810	0.3768	0.5652	0.4522	0.5891	0.8837	0.7069	0.3723	0.5585	0.4468
GRU	0.4836	0.7254	0.5803	0.4545	0.6817	0.5453	0.5889	0.8834	0.7067	0.3723	0.5584	0.4467
LSTM	0.4848	0.7271	0.5817	0.3540	0.5311	0.4248	0.5892	0.8838	0.7071	0.3719	0.5579	0.4463
Self-Attention	0.4830	0.7245	0.5796	0.3509	0.5264	0.4211	0.5693	0.8540	0.6832	0.3707	0.5561	0.4449
ADNN	0.4900	0.7350	0.5880	0.4638	0.6957	0.5565	0.5910	0.8864	0.7091	0.3777	0.5666	0.4533

由表 2 可以看出,GRU 模型在 Irvine 数据集上取得了较好的结果,但是在 Memetracker 数据集上表现不佳。而本文

模型在 4 个数据集上都取得了较好的结果,说明本文模型泛化能力较强,能够利用注意力机制获取序列的局部结构特征

和全局节点的关联性,动态调整特征权重,更好地预测真实的信息传播过程。如图 4 所示,本文模型在 Memetracker, Irvine, Last, FM, Weibo 数据集上的 AUC 结果相比对比模型是最优的。

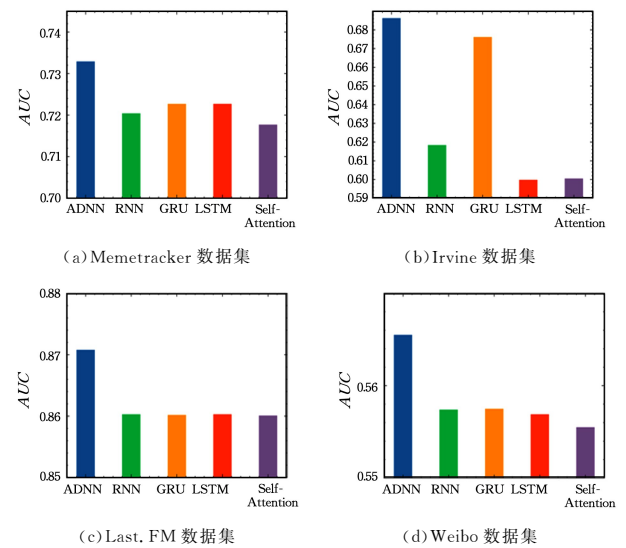


图 4 4 种不同数据集上的 AUC

Fig. 4 AUC on four different datasets

4.4 性能评价

表 2 和图 4 给出了 RNN,GRU,LSTM 和 Self-Attention 4 种模型在 Memetracker, Irvine、Last, FM、Weibo 数据集上性能表现,可以得到以下结论:

(1)仅利用注意力机制的模型(Self-Attention)在不同数据集上的表现均差于循环神经网络模型。Self-Attention 虽然利用注意力机制来提取用户之间的依赖性,在历史信息传播过程中选择对信息传播过程有影响的用户,但这是根据历史信息传播过程得出的,并没有利用信息传播的轨迹特征来提取用户之间的依赖性,导致模型表现较差。实验结果表明,仅利用注意力机制对于预测信息传播过程效果是有限的。

(2)基于循环神经网络的模型(RNN,GRU,LSTM)可以提取信息传播过程的轨迹特征,对信息传播过程进行建模。总体而言,该模型在 4 个数据集上均有较好的性能表现,但是由于无法利用用户之间的依赖性,其考虑影响信息传播的驱动因素较少,因此循环神经网络模型在数据集上的表现仍然不及本文模型。

(3)本文模型(ADNN)利用双向的 LSTM 网络架构增加了模型的容量和灵活性,能够全面地提取信息传播过程的轨迹特征。将轨迹特征通过模型的 Attention Diffusion Module 捕获信息传播之间的依赖关系,挖掘用户在信息传播过程中的驱动因素。与其他模型进行对比可以发现,本文模型能够更好地预测信息传播过程,这进一步说明本文模型获取用户之间的依赖性对于预测信息传播过程是更加有效的。

相较于其他模型,本文模型有以下优势:

- (1)从预测结果上来看本文模型表现更好。
- (2)本文模型避免了复杂的特征工程,且该模型不依赖于图结构信息,更加符合真实的信息传播过程。
- (3)该模型更具通用性,在多种数据集上模型性能表现稳

定,适用于多种数据集。

4.5 参数选择

(1)Embedding 维度的影响。本文研究了 Embedding 维度对模型性能的影响,让 Embedding 维度在 5~25 之间变化,步长为 5。实验结果如表 3 所列。

表 3 不同 Embedding 维度下本文模型在 4 个数据集上的 AUC
Table 3 AUC results of proposed model with different dimension of Embedding on four datasets

Embedding dimension	25	20	15	10	5
Last, FM	0.8681	0.8708	0.8705	0.8701	0.8694
Irvine	0.6871	0.6863	0.6879	0.6863	0.6863
Memetracker	0.7308	0.7329	0.7336	0.7350	0.7345
Weibo	0.5654	0.5655	0.5689	0.5688	0.5688

从表 3 可以看出,AUC 值的趋势为先升高然后降低。可能的原因是小尺寸的 Embedding 会使模型失去足够的表现力,而大尺寸的 Embedding 会使模型出现过度拟合,导致模型性能下降。

(2)Regularization 系数的影响。本文中 Regularization 系数从 0.001 变化到 10,实验结果如表 4 所列。

表 4 不同 Regularization 系数下本文模型在 4 个数据集上的 AUC
Table 4 AUC results of proposed model with different Regularization on four datasets

Regularization	0.001	0.01	0.1	1	10
Last, FM	0.7590	0.8708	0.8756	0.8742	0.8705
Irvine	0.6673	0.6863	0.6832	0.6793	0.6483
Memetracker	0.7335	0.7329	0.7406	0.7384	0.7351
Weibo	0.5550	0.5655	0.5839	0.5839	0.5824

从表 4 可以看出,除 irvine 数据集外,其余数据集上系数为 0.1 时模型的表现最佳。实际上 Regularization 系数可以限制模型其他参数的值和避免模型的过度拟合。因此我们需要在 Regularization 系数和精度优化之间进行适当的权衡。

结束语 针对社交网络中网络舆情事件的特点和国内外研究现状,本文提出一种融合 LSTM 和注意力机制的信息传播预测模型 ADNN。该模型结合注意力机制可以更好地地关注用户之间的依赖关系,采用基于 LSTM 的网络架构来捕获信息传播过程的轨迹特征。为验证该模型的有效性和可行性,本文通过不同的数据集对其进行了实验验证和对比分析。4 个真实数据集上的实验结果验证了本文模型的有效性。

社交网络中,用户个体行为与一条信息的最终的传播过程之间的影响是极为复杂的,真实的传播过程并不严格符合序列假设,利用注意力机制仅从用户层面提取依赖关系存在局限性。为了能够更准确地发现舆情事件的动态演化趋势,在今后的研究中将关注用户对传播项的兴趣,从非序列的角度考虑信息传播过程。

参 考 文 献

[1] STAI E,MILAIUO E,KARYOTIS V,et al. Temporal dynamics of information diffusion in twitter; Modeling and experimentation[J]. IEEE Transactions on Computational Social Systems, 2018,5(1):256-264.

[2] KUHNLE A,ALIM M A,LI X,et al. Multiplex influence maxi-

- mization in online social networks with heterogeneous diffusion models[J]. *IEEE Transactions on Computational Social Systems*, 2018, 5(2): 418-429.
- [3] MATSUBARA Y, SAKURAI Y, PRAKASH B A, et al. Nonlinear dynamics of information diffusion in social networks[J]. *ACM Transactions on the Web (TWEB)*, 2017, 11(2): 1-40.
 - [4] JIANG Y, JIANG J C. Understanding social networks from a multiagent perspective[J]. *IEEE Transactions on Parallel and Distributed Systems*, 2013, 25(10): 2743-2759.
 - [5] XIONG F, WANG X M, CHENG J J. Subtle role of latency for information diffusion in online social networks [J]. *Chinese Physics B*, 2016, 25(10): 108904.
 - [6] LAGNIER C, DENOYER L, GAUSSIER E, et al. Predicting information diffusion in social networks using content and user's profiles[C] // *European Conference on Information Retrieval*. Berlin, Heidelberg: Springer, 2013: 74-85.
 - [7] WANG Z, CHEN C, LI W. A Sequential Neural Information Diffusion Model with Structure Attention[C] // *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, 2018: 1795-1798.
 - [8] WANG J, ZHENG V W, LIU Z, et al. Topological recurrent neural network for diffusion prediction[C] // *2017 IEEE International Conference on Data Mining (ICDM)*. IEEE, 2017: 475-484.
 - [9] BOURIGAULT S, LAGNIER C, LAMPRIER S, et al. Learning social network embeddings for predicting information diffusion [C] // *Proceedings of the 7th ACM International Conference on Web Search and Data Mining*, 2014: 393-402.
 - [10] SUBBIAN K, PRAKASH B A, ADAMIC L. Detecting large re-share cascades in social networks[C] // *Proceedings of the 26th International Conference on World Wide Web*, 2017: 597-605.
 - [11] ANDERSON P, HE X, BUEHLER C, et al. Bottom-up and top-down attention for image captioning and visual question answering[C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018: 6077-6086.
 - [12] WANG F, JIANG M, QIAN C, et al. Residual attention network for image classification[C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017: 3156-3164.
 - [13] LUONG M T, PHAM H, MANNING C D. Effective approaches to attention-based neural machine translation[C] // *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, 2015: 1412-1421.
 - [14] PARK S H, KIM B D, KANG C M, et al. Sequence-to-sequence prediction of vehicle trajectory via LSTM encoder-decoder architecture[C] // *2018 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2018: 1672-1678.
 - [15] GASKÓN, KÉPES T, SUCIU M A, et al. Considerations about using the Shapley Value for Influence Maximization in the case of the Weighted Cascade Model[C] // *2020 IEEE 18th World Symposium on Applied Machine Intelligence and Informatics (SAMI)*. IEEE, 2020: 257-262.
 - [16] BOZORGI A, SAMET S, KWISTHOUT J, et al. Community-based influence maximization in social networks under a competitive linear threshold model[J]. *Knowledge-Based Systems*, 2017, 134: 149-158.
 - [17] GOMEZ-RODRIGUEZ M, BALDUZZI D, SCHÖLKOPF B. Uncovering the temporal dynamics of diffusion networks[C] // *Proceedings of the 28th International Conference on International Conference on Machine Learning*, 2011: 561-568.
 - [18] GOMEZ-RODRIGUEZ M, LESKOVEC J, KRAUSE A. Inferring networks of diffusion and influence[J]. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 2012, 5(4): 1-37.
 - [19] BOURIGAULT S, LAMPRIER S, GALLINARI P. Representation learning for information diffusion through social networks: an embedded cascade model[C] // *Proceedings of the Ninth ACM international conference on Web Search and Data Mining*, 2016: 573-582.
 - [20] SAGHEER A, KOTB M. Time series forecasting of petroleum production using deep LSTM recurrent networks[J]. *Neurocomputing*, 2019, 323: 203-213.
 - [21] WANG H, ZHANG F, ZHANG M, et al. Knowledge-aware graph neural networks with label smoothness regularization for recommender systems[C] // *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2019: 968-977.
 - [22] WANG Z, LI W. Hierarchical Diffusion Attention Network [C] // *IJCAI*, 2019: 3828-3834.
 - [23] YANG C, SUN M, LIU H, et al. Neural Diffusion Model for Microscopic Cascade Study[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2021, 33(3): 1128-1139.
 - [24] OPSAHL T, PANZARASA P. Clustering in weighted networks [J]. *Social networks*, 2009, 31(2): 155-163.
 - [25] DU N, DAI H, TRIVEDI R, et al. Recurrent marked temporal point processes: Embedding event history to vector[C] // *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016: 1555-1564.
 - [26] HU W, SINGH K K, XIAO F, et al. Who will share my image? Predicting the content diffusion path in online social networks [C] // *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*, 2018: 252-260.
 - [27] DEY R, SALEMT F M. Gate-variants of gated recurrent unit (GRU) neural networks[C] // *2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS)*. IEEE, 2017: 1597-1600.
 - [28] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C] // *Advances in Neural Information Processing Systems*, 2017: 5998-6008.



SANG Chun-yan, born in 1983, Ph.D., associate professor. Her main research interests include service recommendation and data mining.