

# 基于可拓展自注意力时空图卷积神经网络的用户轨迹识别模型

雷天亮, 吉立新, 王庚润, 刘树新, 巫 岚

(信息工程大学信息技术研究所, 河南郑州 450002)

**摘 要:** 用户轨迹识别作为一项重要的时空数据挖掘任务, 广泛应用于基于位置的个性化服务推荐、行程规划、犯罪行为检测和目标跟踪等领域, 但依然面临预测精度不高的问题, 主要原因是轨迹数据低采样且稀疏、轨迹类别数量巨大等. 针对上述问题提出了基于可拓展自注意力时空图卷积神经网络的用户轨迹识别模型(Expandable Self-Attention Spatio-Temporal Graph Convolutional Neural Networks, ESAST-GCNN), 该模型采用时空图卷积神经网络方式, 深度挖掘时序特征与空间特征关系, 并进行预测与拓展, 结合自注意力机制获取用户轨迹特征向量内部相关性, 最终根据该特征向量进行用户轨迹身份识别. 在两个真实数据集上进行测试后发现, ESAST-GCNN 相较于 TULER-GRU (TUL via Embedding and RNN) 在 Geolife 与 Gowalla 中准确率分别提高了 13.95%、10.63%, 实验结果表明 ESAST-GCNN 优于其他模型, 识别效果更好, 适用范围更广.

**关键词:** 用户轨迹识别; 时空图卷积神经网络; 自注意力机制; 深度学习; 时空序列

**基金项目:** 河南省重大科技专项项目(No.221100210700)

**中图分类号:** TP391

**文献标识码:** A

**文章编号:** 0372-2112(2024)11-3741-10

**电子学报 URL:** <http://www.ejournal.org.cn>

**DOI:** 10.12263/DZXB.20221225

## User Trajectory Identification Based on Expandable Self-Attention Spatio-Temporal Graph Convolutional Neural Networks

LEI Tian-liang, JI Li-xin, WANG Geng-run, LIU Shu-xin, WU Lan

(Institute of Information Technology, Information Engineering University, Zhengzhou, Henan 450002, China)

**Abstract:** As an important spatio-temporal data mining task, user trajectory identification is widely used in the fields of location-based personalized service recommendation, itinerary planning, crime behavior detection, and target tracking. However, it still has low prediction accuracy, mainly due to low sampling and sparse trajectory data, and a huge number of trajectory categories. To fill the research gaps, a user trajectory identification model based on an expandable self-attention spatio-temporal graph convolutional neural network (ESAST-GCNN) is proposed, which adopts the spatio-temporal graph convolutional neural network to deeply mine the relationship between time sequence features and spatial features to predict and expand the sequence. This model combines the self-attention mechanism to obtain the internal correlation of user trajectory feature vectors and identify user trajectories. After testing on two real datasets, the results show that the accuracy of ESAST-GCNN is improved by 13.95% and 10.63% in Geolife and Gowalla compared with TUL via Embedding and RNN (TULER-GRU), respectively. The experimental results illustrate that ESAST-GCNN is superior to other comparative models, with better identification effect and wider applicability.

**Key words:** user trajectory identification; spatio-temporal graph convolutional neural network; self-attention mechanism; deep learning; space-time sequence

**Foundation Item(s):** Major Science and Technology Project of Henan Province (No.221100210700-2)

## 1 引言

随着互联网的快速发展,基于位置的社交网络得到大规模使用,如微信、QQ、Facebook等,大量的人类移动轨迹数据被记录,为用户轨迹识别方向的发展提供了丰富的数据支持.近年来,用户轨迹识别技术已经引起学术界广泛关注,该技术应用前景广阔,包括地点推荐<sup>[1]</sup>、异常轨迹检测<sup>[2]</sup>、城市交通规划<sup>[3]</sup>等.这在改善用户体验、提高政府社会服务效能等方面具有重要意义.

用户轨迹识别是最近引入的轨迹挖掘任务<sup>[4]</sup>,着重学习轨迹与用户之间的相关性,最终将目标轨迹准确地关联到对应已知用户.用户轨迹识别技术在多个领域都发挥着十分积极的作用,如通过跨平台的轨迹数据识别用户身份,将有助于身份验证和隐私保护<sup>[5]</sup>;同时,平台能够从已知用户轨迹数据中高效地挖掘出恶意用户,如从轨迹数据(如瞬态电话信号和签到)<sup>[4]</sup>中识别潜在的罪犯或恐怖分子,维护社会治安.此外,该技术可以关联不同网络中同一用户,能够更完整地刻画用户短期偏好<sup>[6]</sup>和长期偏好<sup>[7]</sup>.

为更好地完成用户轨迹识别任务,前期许多研究者尝试了多种用户轨迹识别的方法<sup>[5,8,9]</sup>,为后期的研究奠定了坚实的基础.早期的用户轨迹识别模型是基于序列统计方法,如主题模型<sup>[10,11]</sup>、矩阵分解模型<sup>[12]</sup>、马尔可夫模型<sup>[13]</sup>等,虽然能捕捉到用户位置访问偏好以及移动规律,但是无法有效学习到用户轨迹内的序列依赖关系.然而循环神经网络(Recurrent Neural Networks, RNN)模型及其变种<sup>[14]</sup>不仅在长序列处理任务上取得巨大成功<sup>[15]</sup>,同时在用户轨迹识别问题上也表现出很好的效果.

目前主流的用户轨迹识别是以RNN为基础,将用户轨迹识别任务看作多分类任务<sup>[14-19]</sup>进行研究. Gao等人<sup>[4]</sup>首次引入了用户轨迹识别问题的定义,将目标轨迹与对应生成轨迹的用户联系起来,提出了一种基于RNN的用户轨迹识别模型TULER-GRU. TULVAE<sup>[9]</sup>是建立在TULER-GRU基础之上的变体,主要采取半监督方式,利用BiLSTM提取轨迹复杂的依赖关系. Wang等人<sup>[16]</sup>提出了一个多任务模型,通过用RNN使得轨迹向量与对应的生成用户向量在某个特征空间中更加相似,从而实现用户轨迹识别的目的. 同样RNN的其他变体在用户轨迹识别任务中也得到了广泛应用,如长短时记忆(Long Short Term Memory, LSTM)网络<sup>[14]</sup>和门控循环单元(Gated Recurrent Unit, GRU)<sup>[17]</sup>网络,在用户轨迹识别问题上展现了很好的效果. Miao等人<sup>[18]</sup>将注意力机制引入用户轨迹识别中,提出一个注意力循环神经网络模型DeepTUL,该模型结合了多种特征,并从标记的历史轨迹中学习用户移动性的多周期性性质,提高模型的准确性. Zhou等人<sup>[19]</sup>首次提出将图神经

络模型GNNTUL应用到用户轨迹识别. 该模型由一个图神经网络(Graph Neural Networks, GNN)模块和一个分类器组成,将用户轨迹以图的方式表示,能有效地学习人类的流动性,挖掘用户轨迹背后的内隐转换模式,同时提取用户独特的运动特征,并识别用户身份. 部分学者将基于位置点预测的方法用于解决用户轨迹识别问题,如DeepMove<sup>[20]</sup>、CATHI<sup>[21]</sup>和VaNext<sup>[22]</sup>等.

RNN对时序数据的特征提取能力,使其在用户轨迹识别任务上得到发展,但对于具有时间与空间属性特征的轨迹数据来说,对于空间的提取能力也十分重要,于是许多学者开展卷积神经网络(Convolutional Neural Networks, CNN)在有关时空属性数据上的尝试. 1998年LeCun等人<sup>[23]</sup>提出了经典的CNN模型LeNet-5,在原有LeNet模型基础上添加了池化层,具有良好的学习和识别能力. Uddin等人<sup>[24]</sup>提出将从运动视频图像中提取的时空特征应用于CNN进行活动训练和识别,证实基于深度学习的活动识别方法优于其他传统方法. 为了更好地权衡CNN的准确性和复杂性,Agymen等人<sup>[25]</sup>提出增强池化操作,池化内核同时对特征图进行子采样,使得3D CNN性能显著提高. Nikhil等人<sup>[26]</sup>提出一种基于CNN的人类轨迹预测方法,该方法支持并行与有效的时间表示,能够更快地预测出结果. Yin等人<sup>[27]</sup>将注意力机制与CNN相结合,提出采用一种多头注意力的CNN模型来预测步态轨迹,与广泛使用的RNN模型相比预测精度更高、计算时间更短. Yan等人<sup>[28]</sup>提出用于人类动作识别的时空图卷积神经网络(Spatio-Temporal Graph Convolutional Neural Networks, ST-GCNN). 该模型克服了传统方法依赖于遍历规则的局限性,充分学习数据的时空特性,在人类动作识别方面取得了优异的成绩.

自注意力机制对时序数据具有强大建模能力,因此可以将其应用到时序数据预测上. 2017年,以自注意力机制为基本单元的Transformer模型<sup>[29]</sup>使得注意力机制得到广泛的关注,被应用于目标检测、语义分割、图像处理和视频理解等领域. 随后Google AI团队发布的以Transformer框架为基础、采用双向预训练方式的Bert模型<sup>[30]</sup>,极大地提升模型的运行效率,很大程度上推动了学术界和工业界在自然语言处理(Natural Language Processing, NLP)领域的进一步发展. Dosovitskiy等人<sup>[31]</sup>提出以Transformer为基础的ViT模型,其模型简单,效果更佳且可拓展性强,成为了Transformer在计算机视觉领域应用的里程碑著作,但ViT模型的自注意力模块计算复杂度非常高. 针对该问题, Liu等人<sup>[32]</sup>提出Swin-transformer模型,该模型不仅采用了金字塔式的分层结构,同时还提出了一种线性复杂度的Attention计算,在下游任务上表现十分强劲. 基于Transformer变体的

TimeFormer 模型<sup>[33]</sup>中的自注意力机制能够捕捉到数据的时空依赖性,并且训练速度较快,能够减少计算量. Ying 等人<sup>[34]</sup>将 GNN 与 Transformer 结合提出了 Graphormer 模型,该模型通过中心性编码、空间编码和边编码 3 种编码方式将图的结构信息融入到模型中,在图结构数据任务上具有较好表现.

尽管以上模型取得非常重大的成果,但仍有 3 个问题需要解决. (1) 数据稀疏:现实场景下,受用户个人的生活习惯与周围环境的影响,用户不会一直上传其位置信息到相关平台,使得位置数据具有局部性,造成大量位置数据缺失,同时为了保护用户隐私,完备的轨迹数据信息很难获取,极大地提升了研究的难度; (2) 分类类别数量庞大:用户轨迹识别任务可视作为一种多分类任务,与一般多分类任务不同,在该分类任务中每个用户就是一个类别,在真实场景下需要识别的用户数量往往十分庞大,意味着分类时用户类别数也十分巨大,使得更加难以准确地刻画用户与目标轨迹之间的关联性; (3) 未充分利用轨迹数据特征以及特征之间的关联性:以往的方法只利用空间特征,而没有利用轨迹数据中存在的其他特征,如时间特征等,且未考虑空间特征与时间特征的关联性.

为了解决上述挑战,提出了基于可拓展自注意力时空图卷积神经网络(Expandable Self-Attention Spatio-Temporal Graph Convolutional Neural Networks, ESAST-GCNN)的用户轨迹识别模型来解决用户轨迹识别任务. 该模型将轨迹数据以图的方式表示,然后通过时空图卷积神经网络提取用户移动轨迹的时空关联特征,并根据得到的特征进一步预测与扩展用户轨迹时空变化特征向量,缓解数据稀疏和低数据质量等问题,自注意力机制可以捕捉轨迹特征向量内部的相关性等重要的上下文信息,解决长序列依赖问题,提高用户轨迹识别的准确率. 综上所述,本文主要贡献如下:

(1) 采取了图表示方法,不再只依赖于用户轨迹访问序列,可以更好地提取用户轨迹数据的移动规律,该方法相较于其他方法更具通用性和有效性.

(2) 提出了一种基于可拓展自注意力时空图卷积神经网络的用户轨迹识别模型,充分利用用户移动轨迹数据的时空特性,有效地缓解数据稀疏和低数据质量的问题,并且利用自注意力机制计算轨迹内各位置点的互相影响程度,解决长序列依赖问题,并增强网络对用户移动模式独特性的学习.

(3) 在数据密度相差较大的两类真实数据集上进行了实验,结果表明,ESAST-GCNN 的性能比现有模型具有显著的提升.

## 2 问题定义

用户轨迹数据包含多个位置点数据,每个位置点的数据主要包含 3 个要素:用户名、时间戳和空间坐标信息,同一用户的多个轨迹位置点按照时间顺序排列便形成了用户的轨迹,其具体定义如下.

**定义 1** 在时间间隔  $\Delta t$  内,轨迹表示为  $T^i = \{p_1^i, p_2^i, \dots, p_m^i\}$ ,  $p^i$  表示第  $i$  个用户的位置点,  $m$  为轨迹中位置点的个数. 位置点  $p_m^i = (l_m^i, t_m^i)$ ,  $l_m^i$  表示用户  $i$  的第  $m$  个位置点坐标(即经纬度),  $t_m^i$  表示用户  $i$  的第  $m$  个位置点上对应的时间(即时间戳形式).

**定义 2**  $\Gamma = \{T_1, T_2, \dots, T_M\}$  表示未识别的目标轨迹集合,  $U = \{u_1, u_2, \dots, u_N\}$  ( $M \gg N$ ) 表示用户集合,用户轨迹识别旨在构建一个映射关系,可以正确识别用户轨迹身份:  $\Gamma \rightarrow U$ .

## 3 ESAST-GCNN 模型

ESAST-GCNN 模型包括 3 个部分:特征表示层、时空图卷积层和用户轨迹识别层. 第一层的主要目的是获得用户轨迹的特征,将这些特征以图的方式表示;然后将这些被表示的特征输入到时空图卷积层,根据已有的轨迹数据深度预测与拓展轨迹数据,缓解数据稀疏情况;最后通过用户轨迹识别层将未识别的目标轨迹进行用户身份识别. 具体模型框架如图 1 所示.

### 3.1 特征表示层

为获得用户轨迹的图表示形式,特征表示层构造了空间图  $G_t$ , 将  $G_t$  定义为  $G_t = (V_t, E_t)$ , 其中  $V_t = \{v_i^t | \forall i \in \{1, 2, \dots, N\}\}$  是图  $G_t$  的顶点集,位置点  $v_i^t$  表示用户  $i$  在  $t$  时刻的位置嵌入向量. 传统的一维热向量嵌入表示会随着位置点的数量增加导致维度诅咒,因此采用 NLP 任务中常用的词嵌入方法创建一个可学习的词字典,每一个位置点对应一个多维词向量,避免了维度诅咒,并提高了模型学习效率.

$E_t$  是图  $G_t$  中的边,  $E_t = \{e_{ij}^t | \forall i, j \in \{1, 2, \dots, N\}\}$ . 如果  $v_i^t$  和  $v_j^t$  连接,则  $e_{ij}^t = 1$ , 否则  $e_{ij}^t = 0$ . 为了模拟两个节点相互影响的强度,本文使用用户位置点的 L2 范数距离的倒数  $a_{L,t}^{ij}$  来表示相互影响强度,引入  $a_{L,t}^{ij}$  作为  $e_{ij}^t$  的核函数,其定义为

$$a_{L,t}^{ij} = \begin{cases} 1 / \left( \|v_i^t - v_j^t\|_2, \|v_i^t - v_j^t\|_2 \neq 0 \right) \\ 0, \text{ otherwise} \end{cases} \quad (1)$$

### 3.2 时空图卷积层

ESAST-GCNN 模型的时空图卷积部分主要由 ST-GCNN 与时空拓展卷积神经网络(Spatio-Temporal Extended Convolutional Neural Networks, STE-CNN)组成. ST-GCNN 对用户轨迹以图表示的方式进行时空卷积运算来提取用户轨迹特征. STE-CNN 将这些特征作为输

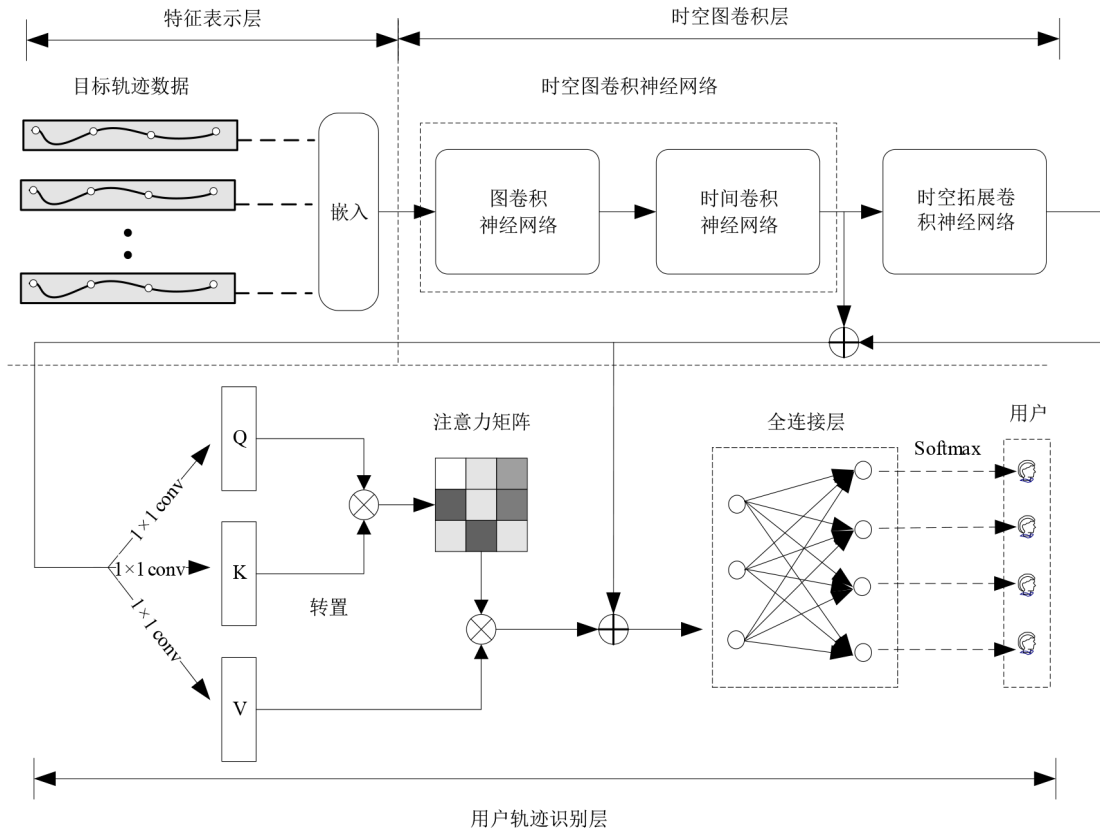


图1 基于可拓展自注意力时空图卷积神经网络的用户轨迹识别模型框架

入,并作为一个整体来预测与拓展所有用户的轨迹向量.该模型总体框架如图2所示.

### 3.2.1 图卷积神经网络

图卷积神经网络从特征表示层中获得图邻接矩阵 $A_l$ 来描述轨迹变量之间的关联关系.卷积运算如下所示:

$$z^{(l+1)} = \sigma \left( \sum_{h=1}^k \sum_{w=1}^k (p(z^{(l)}, h, w)) \cdot W^{(l)}(h, w) \right) \quad (2)$$

其中, $k$ 是卷积核大小, $p(\cdot)$ 是聚合以 $z$ 为中心的邻居信息的采样函数, $\sigma(\cdot)$ 是一个激活函数, $W^{(l)}$ 为 $l$ 层的可训练参数矩阵, $l$ 表示第 $l$ 层.

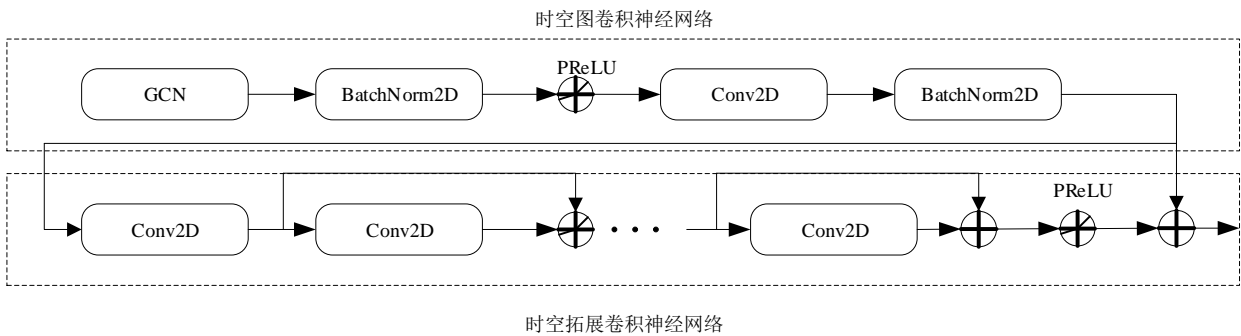


图2 时空图卷积模型框架

结合式(2)可以推导出图的卷积运算式为

$$v^{i(l+1)} = \sigma \left( \frac{1}{\Omega} \sum_{v^{j(l)} \in B(v^{i(l)})} p(v^{i(l)}, v^{j(l)}) \cdot W(v^{i(l)}, v^{j(l)}) \right) \quad (3)$$

其中, $B(v^i) = \{v^j | d(v^i, v^j) \leq D\}$ 是顶点 $v^i$ 的邻居集, $d(v^i, v^j)$

表示连接 $v^i$ 和 $v^j$ 的最短路径; $\frac{1}{\Omega}$ 是一个归一化项, $\Omega$ 是邻居集数.

### 3.2.2 时空图卷积神经网络

ST-GCNN将空间图卷积扩展到时空图卷积,该图G

的属性集是  $G_t, G_t$  包含了用户轨迹的时空信息. 图  $G$  的网络拓扑结构是根据用户间的位置关系构建, 所有用户之间都存在空间位置关系, 只是关联程度不同, 所以  $G_1, G_2, \dots, G_T$  的拓扑结构是相同的, 而当  $t$  变化时  $v_t$  会分配到不同的属性, 用户间的边也会分配到不同的属性. 与  $G$  对应的加权邻接矩阵  $A$  是  $\{A_1, \dots, A_t, \dots, A_T\}$  的集合.

为了便于学习, 首先将加权邻接矩阵使用归一化公式<sup>[35]</sup>进行对称归一化:

$$A_t = A_t^{-\frac{1}{2}} \hat{A}_t A_t^{-\frac{1}{2}} \quad (4)$$

其中,  $\hat{A}_t = A_t + I, A_t$  是  $\hat{A}_t$  的对角节点度矩阵.

将时间点  $t$  和网络层  $l$  处的顶点集嵌入表示为  $V_t^{(l)}$ . 图卷积网络层如图 3 所示, 根据上述定义, 可以实现 ST-GCNN 层:

$$f(V_t^{(l)}, A) = \sigma \left( A_t^{-\frac{1}{2}} \hat{A}_t A_t^{-\frac{1}{2}} V_t^{(l)} W^{(l)} \right). \quad (5)$$

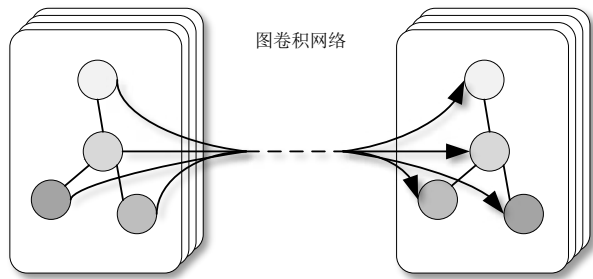


图 3 图卷积网络

在经过 ST-GCNN 处理之后, 可以得到图的特征  $\bar{V}$ , 然后将特征  $\bar{V}$  传递给由一系列残差连接的 CNN 组成的 STE-CNN, 该模块将时间维度视为特征通道, 并根据 ST-GCNN 得到的图表征特征  $\bar{V}$  中的时空特征变化规律预测用户未来时刻轨迹的特征向量, 与原轨迹特征向量拼接进一步拓展原数据集, 丰富训练数据, 有利于后续用户轨迹识别层对数据特征的认识, 缓解数据稀疏的问题. 计算式如下:

$$\bar{V}^{(l+1)} = \bar{V}^{(l)} + \sigma[\text{Conv2D}(\bar{V}^{(l)})] \quad (6)$$

其中,  $\text{Conv2D}(\cdot)$  表示图卷积函数.

总的来说, ESAST-GCNN 和 ST-GCNN 之间主要的区别在于: ESAST-GCNN 除了用户轨迹识别层之外, 还增加了使用 STE-CNN 来增强时间维度的灵活性, 进一步挖掘时空关联特征信息.

### 3.3 用户轨迹识别层

用户轨迹识别层由上下文信息生成模块与全连接网络组成. 上下文信息生成模块通过自注意力机制从已拓展的时空特征向量  $\bar{V}$  中获取上下文信息, 并将其与特征  $\bar{V}$  拼接获得更高级的特征向量表示  $x$ , 然后将向

量  $x$  输入全连接层并利用 Softmax 函数映射成  $N$  维的向量  $y$ , 得到分类结果.

$$x = \text{Softmax} \left( \frac{Q \times K^T}{\sqrt{d_k}} \right) V \quad (7)$$

$$y = \text{Softmax}(W_n x + b_n) \quad (8)$$

其中,  $Q = \bar{V}' \times W_Q$  为查询向量,  $W_Q$  为查询权重矩阵;  $K = \bar{V}' \times W_K$ , 为键向量,  $W_K$  为键权重矩阵;  $V = \bar{V}' \times W_V$ , 为值向量,  $W_V$  为值权重矩阵;  $d_k$  为向量的维数;  $W_n$  为全连接层权重矩阵,  $b_n$  为偏差,  $y$  是经过归一化后的概率分布向量, 该向量中每一维  $y_i$  表示属于用户  $i$  的概率.

## 4 实验结果与分析

### 4.1 实验数据集

实验采用了两种真实的用户移动数据集: Gowalla 数据集和 Geolife 数据集.

(1) Gowalla 数据集: 在 Gowalla 签到数据集中, 每个用户的每一次签到都作为一条记录, 签到的具体属性分列表示, 具体包括用户 ID、签到时间、签到地点经纬度坐标, 以及每个地点的经纬度所唯一对应的地点 ID. 处理后用于实验的 Gowalla 数据集稀疏度为 66.41%, 数据较为稀疏, 因此数据质量较低.

(2) Geolife 数据集: 该数据集是 GPS 类型轨迹数据集, 它是在微软亚洲研究院的格里夫项目中收集的. 该数据集的 GPS 轨迹由一系列时间戳点表示, 每个点都包含纬度、经度和高度的信息. 记录了大量用户广泛的户外活动, 不仅包括回家上班等生活习惯, 还有一些娱乐和体育活动. 处理后用于实验的 Geolife 数据集稀疏度为 0%, 数据密集, 数据质量较高.

为便于学习, 本文将两个数据集的数据构建成轨迹形式, 选取一些轨迹数量较多的用户, 能够更好地让模型学习用户的移动模式, 将用户轨迹  $T$  切分为子轨迹集  $\Gamma = \{T_1, T_2, \dots, T_M\}$ . 该实验将每位用户的轨迹数量的 70% 作为训练集、15% 作为验证集、15% 作为测试集. 实验采用的 Gowalla 与 Geolife 数据集信息如表 1 所示.

表 1 实验数据集统计信息

数据集	用户数	训练集	验证集	测试集
Gowalla	156	79 872	17 004	17 160
Geolife	85	32 725	6 970	7 055

### 4.2 对比模型

(1) LCSS<sup>[36]</sup> (Longest Common Sub-Sequence): 一种通过寻找两个序列最大公共子序列的方法比较两个序列之间相似性.

(2) LDA<sup>[37]</sup> (Linear Discriminant Analysis): 一种被

广泛应用的线性判别方法. 该模型通过将轨迹嵌入到一维热向量中, 并利用奇异值分解优化求解类内散点矩阵的方法来解决用户轨迹识别问题.

(3) SVM<sup>[38]</sup> (Support Vector Machine): SVM 是轨迹数据挖掘任务中常见的模型之一, 采用了线性核进行分类学习.

(4) TULER-GRU<sup>[4]</sup> (TUL via Embedding and RNN): 将 RNN 用于用户轨迹识别任务的前沿工作, 具有很好的识别效果, 如带有门控循环单元的神经网络.

(5) TimeSformer<sup>[33]</sup>: 其使用自注意力机制结构充分学习数据潜在的时空特征, 计算复杂度小, 模型训练速度快.

(6) Graphormer<sup>[34]</sup>: 将 Transformer 的思想应用在 Graph 任务上, 通过中心性编码、空间编码和边编码 3 种编码方式将图的结构信息融入到模型中, 在图预测任务上取得较好的效果.

根据本文的模型, 分别提出了变体模型: 基于时空图卷积神经网络模型 ST-GCNN、基于时空图卷积神经网络与自注意力机制 (Self-Attention Spatio-Temporal Graph Convolutional Neural Networks, SA-STGCNN) 模型. 虽然 SA-STGCNN 相较于 ESAST-GCNN 缺少了一个 STE-CNN 模块, 但相较于 ST-GCNN 保留了自注意力机制, 以便进行模型效果对比, 突出各个模块的作用.

### 4.3 评价指标

为了更好地与其他实验方法进行对比, 本文的评价指标采用了 ACC@K (Top K 准确率) 和 Macro-P、Macro-R 与 Macro-F1 等指标. 其中, ACC@K 是指若将轨迹的多个预测结果按照概率大小进行排序, 前  $k$  个预测结果中包含了轨迹真实的用户身份, 则认为预测结果正确的概率, 计算式如下:

$$ACC@K = \frac{\#Correctly\ Identified\ Trajectories@K}{\#Trajectories} \quad (9)$$

Macro-F1 作为衡量多分类模型准确性的一项重要指标, 其计算式为

$$Macro-F1 = \frac{2 \times Macro-P \times Macro-R}{Macro-P + Macro-R} \quad (10)$$

其中, Macro-P 和 Macro-R 分别为多分类任务中的平均准确率和平均召回率.

### 4.4 实验参数

本文实验的模型训练时采用 Adam 优化器优化网络, ST-GCNN 层数设置为 1, STE-CNN 层数设置为 2, 初始学习率设为 0.001, Dropout 设为 0.2, Epoch 设为 100, 批次大小设置为 64, 实验均在 Pytorch 框架上完成.

### 4.5 实验结果分析

本文提出的 ESAST-GCNN 模型在签到类型数据集 Gowalla 与 GPS 类型数据集 Geolife 的实验结果如表 2 所示, 其中最佳结果用粗体表示.

表 2 Gowalla 和 Geolife 数据集上各模型的性能结果对比

单位: %

数据集	模型	ACC@1	ACC@3	ACC@5	Macro-P	Macro-R	Macro-F1
Gowalla	LCSS	46.71	55.47	58.39	25.69	23.36	24.47
	LDA	55.31	66.92	71.50	56.15	63.71	59.69
	SVM	56.98	68.43	72.93	57.95	66.65	62.00
	TULER-GRU	71.89	81.06	83.72	78.85	71.89	75.21
	TimeSformer	77.76	84.36	89.02	78.58	79.38	78.98
	Graphormer	79.39	85.91	89.24	80.27	80.08	80.17
	ST-GCNN	41.95	51.04	54.80	47.66	41.95	44.62
	SA-STGCNN	74.75	82.82	85.20	79.76	74.75	77.17
	ESAST-GCNN	<b>82.52</b>	<b>88.68</b>	<b>91.23</b>	<b>84.48</b>	<b>86.50</b>	<b>85.48</b>
Geolife	LCSS	55.69	71.23	77.82	37.03	44.10	40.26
	LDA	62.65	76.61	81.64	63.58	57.68	60.49
	SVM	71.01	88.94	91.21	72.46	58.27	64.59
	TULER-GRU	75.18	89.61	92.67	76.04	75.18	75.61
	TimeSformer	86.41	94.65	96.03	86.06	86.27	86.16
	Graphormer	88.40	<b>96.57</b>	97.68	<b>89.45</b>	88.40	<b>88.92</b>
	ST-GCNN	85.38	93.87	95.51	86.71	85.38	86.04
	SA-STGCNN	85.85	94.91	96.83	86.26	86.68	86.47
	ESAST-GCNN	<b>89.13</b>	95.81	<b>97.94</b>	88.49	<b>88.43</b>	88.46

从表 2 可以看出, ESAST-GCNN 在大多数指标上表现最好, 尤其在数据较为稀疏的 Gowalla 数据集中, 这

得益于 ESAST-GCNN 的预测与拓展轨迹向量的能力, 在一定程度上缓解数据稀疏问题. 与 Graphormer 模型

相比较, ESAST-GCNN 模型仅采用自注意力模块而未使用了 Transformer 完整结构, 一方面是为了防止神经网络层数过多导致过拟合现象, 另一方面可以适当降低模型运算成本. 尽管 Graphormer 模型具有强大的图结构信息学习能力, 但其忽略了轨迹图之间的时间依赖关系, 而 ESAST-GCNN 出色的用户轨迹时空特征信息的挖掘能力与对用户移动模式独特性的学习能力, 使其在数据密集地 Geolife 数据集中与 Graphormer 的实验结果差距不大, 复杂的网络结构并不能带来明显的性能增益.

从图 4 与图 5 的实验结果可以看出, 各模型在 Geolife 数据集中的表现均优于 Gowalla 数据集. 这是由于 Gowalla 数据集的数据稀缺性导致本身的序列关系不明显, 而 Geolife 数据集中的序列关系非常完整, 所以模型能更准确地学习到每个用户的移动特性.

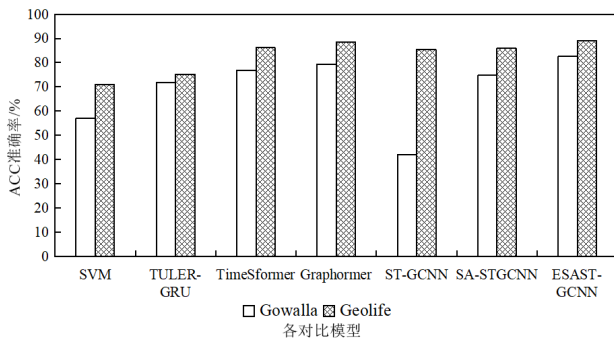


图 4 各模型的 ACC@1 指标结果对比

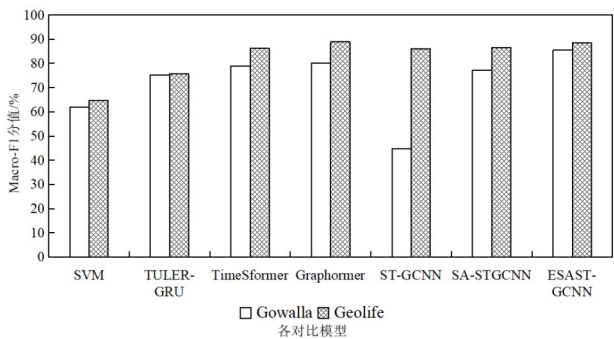


图 5 各模型的 Macro-F1 指标结果对比

本文在 Gowalla 与 Geolife 两个数据集上对提出的 ESAST-GCNN 中的各个模块的重要性进行了一系列的消融研究, 消融实验结果见表 3. 这里以 ST-GCNN 模型作为基准实验模型.

表 3 Gowalla 和 Geolife 数据集上消融实验结果

Model	ST-GCNN	Self-attention	STE-CNN	Macro-F1	
				Gowalla	Geolife
ST-GCNN	√	—	—	44.62%	86.04%
SA-STGCNN	√	√	—	77.17%	86.47%
ESAST-GCNN	√	√	√	85.48%	88.46%

该实验将包含自注意力机制的 SA-STGCNN 模型与无自注意力机制的 ST-GCNN 模型进行比较. 从图 6 与图 7 的对比结果中可以看出, 在数据密集的条件下两个模型之间的性能差距较小, 但在数据较为稀疏的环境下 SA-STGCNN 模型的性能明显更好、更稳定, 这表明由于自注意力机制模块计算出轨迹向量内隐向量的上下关系, 增强了模型对用户移动模式特异性的提取能力与模型的鲁棒性. 同理, 将 ESAST-GCNN 模型与 SA-STGCNN 模型进行比较, 可以看出 ESAST-GCNN 模型在两个数据集中性能表现得更好, 这是由于 STE-CNN 具有根据用户轨迹内的时空特征变化规律进行预测与拓展轨迹向量的能力, 增强模型鲁棒性以及为用户移动模式特异性的提取能力.

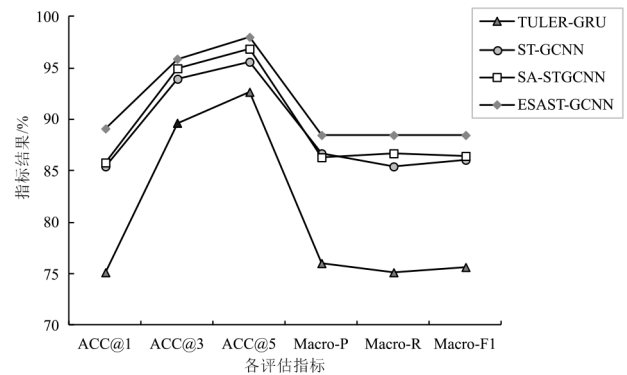


图 6 Geolife 数据集实验结果

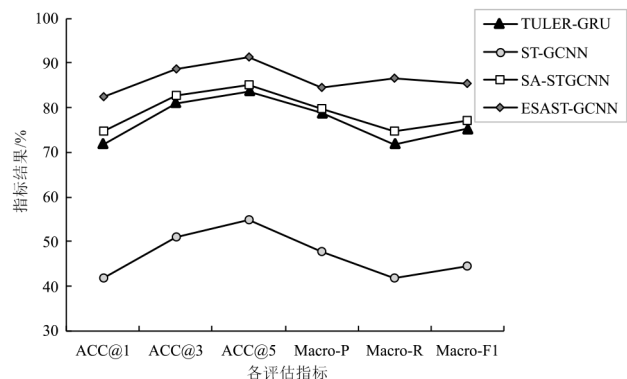


图 7 Gowalla 数据集实验结果

如图 8 与图 9 所示, ESAST-GCNN 及其变体在 Geolife 这种高密度数据集中的收敛速度均优于 TULER-GRU, 但在 Gowalla 这种低密度数据集中, ST-GCNN 与 SA-STGCNN 收敛速度明显落后于 TULER-GRU, 然而 ESAST-GCNN 的收敛速度依然保持领先, 这也说明 STE-CNN 模块使得 ESAST-GCNN 受到数据集稀疏性的影响更小, 提高了该模型的准确性与效率.

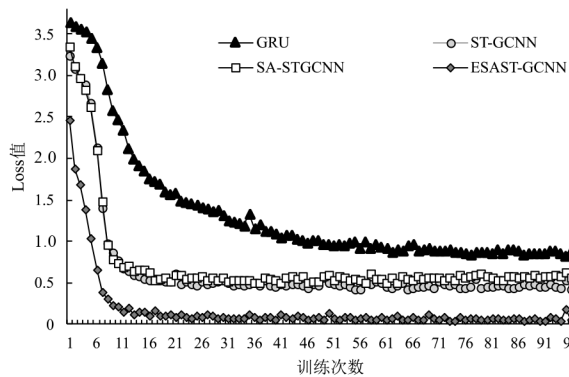


图8 Geolife数据集上的模型收敛趋势

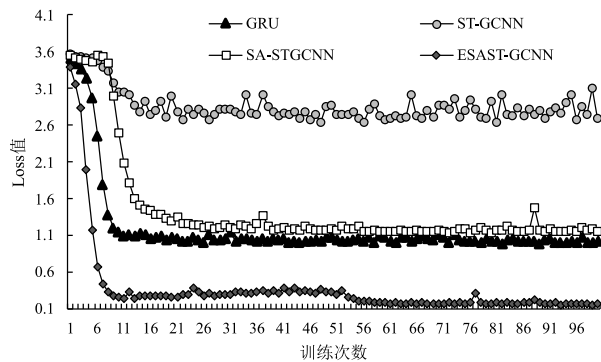


图9 Gowalla数据集上的模型收敛趋势

## 5 总结

本文针对用户轨迹识别任务中数据稀疏、时空特征关联程度低等问题,提出了基于可拓展自注意力时空图卷积神经网络的用户轨迹识别模型 ESAST-GCNN. 该模型由特征表示层、时空图卷积层以及用户轨迹识别层组成,将用户轨迹以图的方式呈现,并获得用户轨迹图特征向量,经过时空图卷积层获取用户移动的时空关联特征,并进一步拓展与预测该时空关联特征,充分挖掘用户轨迹时空特征信息,然后利用自注意力机制来增强网络对用户移动模式独特性的学习,提高模型性能. 通过在两个真实移动性数据集上进行广泛实验,证明了该模型的有效性,并给出了直观的解释. 未来工作重点将集中在提高用户轨迹识别模型鲁棒性以及动态用户轨迹时空关系的挖掘上,进一步提升用户轨迹识别的准确性与适用性.

## 参考文献

- [1] YIN H, CUI B. Spatio-Temporal Recommendation in Social Media[M]. Singapore: Springer Singapore, 2016.
- [2] 毛嘉莉, 金澈清, 章志刚, 等. 轨迹大数据异常检测: 研究进展及系统框架[J]. 软件学报, 2017, 28(1): 17-34.
- MAO J L, JIN C Q, ZHANG Z G, et al. Trajectory big data anomaly detection: Research progress and system framework[J]. Journal of Software, 2017, 28(1): 17-34. (in Chinese)
- [3] ZHOU X, NOULAS A, MASCOLO C, et al. Discovering latent patterns of urban cultural interactions in WeChat for modern city planning[C]//Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM, 2018: 1069-1078.
- [4] GAO Q, ZHOU F, ZHANG K, et al. Identifying human mobility via trajectory embeddings[C]//Proceedings of the 26th International Joint Conference on Artificial Intelligence. Palo Alto: AAAI, 2017: 1689-1695.
- [5] ZHOU F, LIU L, ZHANG K, et al. Deep Link: A deep learning approach for user identity linkage[C]//IEEE INFOCOM. Piscataway: IEEE, 2018: 1313-1321.
- [6] HIDASI B, KARATZOGLOU A, BALTRUNAS L, et al. Session-based recommendations with recurrent neural networks[C]//Proceedings of the 4th International Conference on Learning Representations. Piscataway: IEEE, 2016: 1-12.
- [7] QUADRANA M, KARATZOGLOU A, HIDASI B, et al. Personalizing session-based recommendations with hierarchical recurrent neural networks[C]//Proceedings of the 11th ACM Conference on Recommender Systems. New York: ACM, 2017: 130-137.
- [8] 陈鸿昶, 徐乾, 黄瑞阳, 等. 一种基于用户轨迹的跨社交网络用户身份识别算法[J]. 电子与信息学报, 2018, 40(11): 2758-2764.
- CHEN H C, XU Q, HUANG R Y, et al. A cross social network user identification algorithm based on user trajectory[J]. Journal of Electronics and Information Technology, 2018, 40(11): 2758-2764. (in Chinese)
- [9] ZHOU F, GAO Q, TRAJCEVSKI G, et al. Trajectory-user linking via variational autoencoder[C]//Proceedings of the 27th International Joint Conference on Artificial Intelligence. Palo Alto: AAAI, 2018: 3212-3218.
- [10] KURASHIMA T, IWATA T, HOSHIDE T, et al. Geo topic model: Joint modeling of user's activity area and interests for location recommendation[C]//Proceedings of the 6th ACM International Conference on Web Search and Data Mining. New York: ACM, 2013: 375-384.
- [11] YIN H, SUN Y, CUI B, et al. LCARS: A location-content-aware recommender system[C]//Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM, 2013: 221-229.
- [12] ZHENG W, ZHENG Y, XIE X, et al. Collaborative lo-

- cation and activity recommendations with GPS history data[C]//Proceedings of the 19th International Conference on World Wide Web. New York: ACM, 2010: 1029-1038.
- [13] HE R, JULIAN M. Fusing similarity models with Markov chains for sparse sequential recommendation[C]//Proceedings of the 16th IEEE International Conference on Data Mining. Piscataway: IEEE, 2016: 191-200.
- [14] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. *Neural Computation*, 1997, 9(8): 1735-1780.
- [15] KALCHBRENNER N, BLUNSOM P. Recurrent continuous translation models[C]//Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing. Piscataway: IEEE, 2013: 1700-1709.
- [16] WANG G, LIAO D, LI J. Complete user mobility via user and trajectory embeddings[J]. *IEEE Access*, 2018, 6: 72125-72136.
- [17] CHUNG J, GULCEHRE C, CHO K, et al. Empirical evaluation of gated recurrent neural networks on sequence modeling[C]//NIPS 2014 Workshop on Deep Learning. Piscataway: NIPS, 2014: 1-14.
- [18] MIAO C, WANG J, YU H, et al. Trajectory-user linking with attentive recurrent network[C]//Proceedings of the 19th International Conference on Autonomous Agents and Multi Agent Systems. Richland: IFAAMS, 2020: 878-886.
- [19] ZHOU FAN, CHEN SHUPEI, WU JIN, et al. Trajectory-user linking via graph neural network[C]//Proceedings of IEEE International Conference on Communications. Piscataway: IEEE, 2021: 1-6.
- [20] FENG J, LI Y, ZHANG C, et al. Deep Move: Predicting human mobility with attentional recurrent networks[C]//Proceedings of the 2018 World Wide Web Conference. New York: ACM, 2018: 1459-1468.
- [21] ZHOU F, YUE X, TRAJCEVSKI G, et al. Context aware variational trajectory encoding and human mobility inference[C]//Proceedings of the 2019 World Wide Web Conference. New York: ACM, 2019: 3469-3475.
- [22] GAO Q, ZHOU F, TRAJCEVSKI G, et al. Predicting human mobility via variational attention[C]//Proceedings of the 2019 World Wide Web Conference. New York: ACM, 2019: 2750-2756.
- [23] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. *Proceedings of the IEEE*, 1998, 86(11): 2278-2324.
- [24] UDDIN M Z, KHAKSAR W, TORRESEN J. Human activity recognition using robust spatiotemporal features and convolutional neural network[C]//2017 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI). Piscataway: IEEE, 2017: 144-149.
- [25] AGYEMAN R, RAFIQ M, SHIN H K, et al. Optimizing spatiotemporal feature learning in 3d convolutional neural networks with pooling blocks[J]. *IEEE Access*, 2021(9): 70797-70805.
- [26] NIKHIL N, MORRIS B T. Convolutional neural network for trajectory prediction[M]//Lecture Notes in Computer Science. Cham: Springer International Publishing, 2019: 186-196.
- [27] YIN J, XUE T, ZHANG T. Real-time gait trajectory prediction based on convolutional neural network with multi-head attention[C]//2022 27th International Conference on Automation and Computing (ICAC). Piscataway: IEEE, 2022: 1-6.
- [28] YAN S, XIONG Y, LIN D. Spatial temporal graph convolutional networks for skeleton-based action recognition[C]//Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence. Palo Alto: AAAI, 2018: 7444-7452.
- [29] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. New York: ACM, 2017: 6000-6010.
- [30] DEVLIN J, CHANG M-W, LEE K, et al. BERT: Pre-training of deep bidirectional transformers for language understanding[C]//Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Minneapolis: ACL, 2019: 4171-4186.
- [31] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16x16 words: Transformers for image recognition at scale[C]//International Conference on Learning Representations. New York: ACM, 2021: 1-12.
- [32] LIU Z, LIN Y, CAO Y, et al. Swin transformer: Hierarchical vision transformer using shifted windows[C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway: IEEE, 2021: 9992-10002.
- [33] BERTASIUS G, WANG H, TORRESANI L. Is space-time attention all you need for video understanding?[C]//Proceedings of the 38th International Conference on Machine Learning. Piscataway: IEEE, 2021: 813-824.
- [34] YING C, CAI T, LUO S, et al. Do transformers really per-

form bad for graph representation?[C]//Proceedings of the 35th Conference on Neural Information Processing Systems (NeurIPS 2021). Sydney: NIPS, 2021: 28877-28888.

- [35] KIPF T N, WELING M. Semi-supervised classification with graph convolutional networks[C]//Proceedings of the 5th International Conference on Learning Representations. Piscataway: IEEE, 2017: 1-13.
- [36] DAS G, GUNOPULOS D, MANNULA H. Finding similar time series[C]//Proceedings of the 1st European Symposium on Principles of Data Mining and Knowledge Discovery. Berlin: Springer, 1997: 88-100.
- [37] MIKOLOV T, CHEN K, CORRADO G, et al. Efficient estimation of word representations in vector space[C]//Proceedings of the 1st International Conference on Learning Representations. Piscataway: IEEE, 2013: 1-15.
- [38] CORTES C, VLADIMIR V. Support-vector network[J]. Machine Learning, 1995, 20(3): 273-297.



刘树新 男,1987年出生,山东潍坊人. 现为国家数字交换系统工程技术研究中心助理研究员. 主要研究方向为链路预测、通信网络安全.  
E-mail: liushuxin11@126.com



巫 岚 女,1994年出生,四川遂宁人. 现为国家数字交换系统工程技术研究中心研究实习员. 主要研究方向为行为分析和网络空间安全.  
E-mail: lanwundsc@163.com

#### 作者简介



雷天亮 男,1998年出生,安徽宣城人. 现为信息工程大学硕士研究生. 主要研究方向为用户身份识别.  
E-mail: leitianliang1@163.com



吉立新 男,1970年出生,江苏淮安人. 现为国家数字交换系统工程技术研究中心研究员,博士生导师. 主要研究方向为数据挖掘、电信网安全.  
E-mail: jlxdsc@139.com



王庚润 男,1987年出生,安徽蒙城人. 现为国家数字交换系统工程技术研究中心副研究员. 主要研究方向为电信网安全、数据处理.  
E-mail: wanggenrun@gmail.com