



DOI:10.13364/j.issn.1672-6510.20220063

基于 GRU 神经网络的可移动性兴趣点的推荐系统

史艳翠, 张弛

(天津科技大学人工智能学院, 天津 300457)

摘要: 目前, 兴趣点推荐算法的研究主要聚焦在静态兴趣点的推荐上, 对可移动性兴趣点的推荐研究比较少. 本文提出了一种包括 3 个网络模块的复合神经网络结构, 可以实现可移动性兴趣点的推荐. 整体算法框架包含门控循环 (gate recurrent unit, GRU) 神经网络专家模块和门网决策模块, 门控循环神经网络专家模块由模式 GRU 和点 GRU 两个子模块组成. 模式 GRU 网络模块融入了迁移学习策略, 负责学习其他数据集中的动态时空模式, 用于模式预测. 点 GRU 网络模块融入了对比学习策略, 以达到扩展目标训练集的样本数量、提高网络泛化能力的目的, 用于点预测; 门网决策模块负责根据输入样本的形式选择相应的 GRU 模块的输出作为预测结果, 实现神经网络专家决策系统. 本文还提出了一种结合 TrAdaBoost + 最大期望 (expectation maximization, EM) 算法的样本过滤算法, 该算法能够从两种学习策略产生的扩容样本中筛选出可信的、可用的样本, 用于训练模式 GRU 以及点 GRU 网络模块. 实验结果表明, 与马尔可夫 (Markov) 模型和高斯混合模型 (Gaussian mixture model, GMM) 方法相比, 本文方法泛化能力强, 可用小训练样本集驱动神经网络, 产生的预测误差以及预测误差分位数更小.

关键词: 移动性兴趣点; 动态时空推荐; 迁移学习; 对比学习; 门控循环神经

中图分类号: TP311 **文献标志码:** A **文章编号:** 1672-6510(2022)06-0054-09

Recommendation System for Mobility Points of Interest Based on GRU Neural Networks

SHI Yancui, ZHANG Chi

(College of Artificial Intelligence, Tianjin University of Science & Technology, Tianjin 300457, China)

Abstract: At present, the research on point of interest recommendation algorithm mainly focuses on the recommendation of static points of interest. However, there is little research on the recommendation of mobile points of interest. In this article, a composite neural network structure including three network modules is proposed to realize the recommendation of mobile points of interest. The overall algorithm framework includes the gate recurrent unit (GRU) neural network expert module and the gate network decision module. The gate recurrent unit neural network expert module is composed of two sub modules: mode GRU and point GRU. The mode GRU module integrates the transfer learning strategy and is responsible for learning the dynamic spatiotemporal modes in other data sets for mode prediction; the point GRU module integrates the contrast learning strategy to expand the number of samples in the target training set and improve the generalization ability of the network for point prediction; the gate network decision-making module is responsible for selecting the output of the corresponding GRU module as the prediction result according to the form of input samples to realize the neural network expert decision-making system. In this article we also propose a sample filtering algorithm combined with TrAdaBoost + expectation maximization (EM). The algorithm can select credible and available samples from the expanded samples generated by the two learning strategies for training mode GRU and point GRU modules. The experimental results show that, compared with Markov model and Gaussian mixture model (GMM) methods, this method has strong generalization ability, can drive the

收稿日期: 2022-03-16; 修回日期: 2022-05-29

基金项目: 国家自然科学基金青年基金项目 (61402331, 61702367, 61807024); 天津市教委理工类基本科研业务费项目 (2018KJ105)

作者简介: 史艳翠 (1982—), 女, 河北人, 副教授, syc@tust.edu.cn

neural network with small training sample set, and produces smaller prediction error and prediction error quantile.

Key words: mobile points of interest; dynamic space-time recommendation; transfer learning; contrast learning; gated recurrent neurons

近年来,随着个人手持终端设备的快速发展以及全球定位系统(global positioning system, GPS)定位技术的日渐成熟,研究人员和商业机构很容易获得大量的个人地理位置信息^[1].大型软件服务商也可以在用户授权的前提下获得丰富的位置和轨迹信息,进而为用户提供基于位置的服务^[2],例如定位、导航、移动中的搜索等功能服务.如何利用时空轨迹数据为用户提供便捷的推荐服务,这类问题成为科研人员关注的重点^[3].对个人活动空间的计算分析,可以挖掘出个体的移动模式、城市的热点区域、兴趣点分布等信息^[4],为城市生活提供有力的数据支撑和高效的服务.在推荐系统研究领域中,基于兴趣点的推荐是应用较为广泛的推荐方法.

目前基于兴趣点的推荐算法很大程度上依赖于以下两点:已有的轨迹中相似的时空运动轨迹^[5]数据以及对个体和群体行为的分析结果.典型的应用场景有行人及车辆轨迹、社交网站签到等.上述应用场景都是以人类社会活动产生的数据为基础,通过对用户行为进行分析得到社会中人群的移动模式、城市的热点区域、兴趣点分布等信息,进而达到让城市生活更加便捷、减少道路拥堵等目的,最终产生了基于兴趣点的推荐系统.这类研究的对象往往都是静态兴趣点+动态用户,并且研究的关注点大都集中在对动态用户轨迹的分析预测问题上,而缺少对兴趣点位置变化的关注.

然而,现实世界的需求是动态变化的,兴趣点并非都是静态的.本文针对可移动性兴趣点进行了研究.一个典型的场景是深圳普思英察科技有限公司(PerceptIn公司)在中国深圳工业园区部署的无人智能售货车.从用户的角度看,无人智能售货车是可移动的兴趣点;从无人智能售货车的角度看,用户对其来说也是可移动的兴趣点.这种商业形式的推荐服务需要考虑移动性因素,这与传统静态形式的推荐服务有着本质上的区别.因此,在推荐系统研究中,将静态兴趣点拓展到可移动兴趣点的情景上并实现相应的推荐算法,是一个研究创新点,也是一个研究难点.这种扩展在现实生活中具有广泛的应用环境以及对应的推荐场景^[6].

本文主要解决的问题、实现方法以及创新点如下:

(1)通过赋予兴趣点时空轨迹来引入兴趣点的可

移动性,并解决动态环境下推荐算法的需求问题.

(2)对用户的时空轨迹信息进行提取,然后利用特征层级的迁移学习策略提升网络的预测能力和泛化能力,利用对比学习扩展训练数据的样本规模.为了使扩展后数据集的概率分布不变,本文使用了最大均值差异度量方法(maximum mean discrepancy, MMD)以及 TrAdaBoost + 最大期望(expectation maximization, EM)算法对其进行两步过滤,去除不符合要求的数据.

(3)使用门控循环(gate recurrent unit, GRU)神经网络对用户的未来时空轨迹进行预测,根据用户的聚集方向合理地调整附近的兴趣点位置.

(4)最终结合预测的时空结果,为用户进行合理推荐,使得移动兴趣点与用户能够在一定的时空范围内进行交互.

1 相关工作

时空轨迹预测^[7]的基本实现过程是对时空序列数据进行分析,进而完成建模并在此基础上进行预测^[8].可以将时空序列预测看作是对时间序列的分析在空间中的拓展,其中的时空关系会因为建模方法的不同而有所差异.Cliff等^[9]在1975年提出时空自相关移动平均模型 STARMA,并首次提出了时空序列建模的框架.Martin等^[10]建立了时空自相关移动模型,该模型能够利用时空延迟算子将时间延迟和空间延迟连接起来,从而真正地建立时空一体化模型.对于非平稳、非线性模型,则采用时空序列混合模型,将层次贝叶斯模型、状态空间模型和卡尔曼滤波等方法融合起来.近年来的神经网络^[11-12]、支持向量机^[13]等方法也被拓展并应用到时空序列建模和预测的研究中,比如时空序列神经网络 STANN^[14]、时空序列支持向量机 STSVR等.杨立宁等^[15]提出基于奇异值分解模型和自回归积分滑动平均模型的时空序列分解和预测,该研究结果不仅大幅度降低了训练时间成本,还提高了预测精度.

要解决将某个个体的时空运动预测结果推荐给其他用户的问题,则需要使用适合的时空信息推荐算法.目前有关时空轨迹的兴趣点推荐类型分为两类:一是地点推荐,例如为用户推荐商店、餐馆等;二是

行程推荐,例如为用户推荐旅游路线、为用户导航较小代价的路线。

基于兴趣点的推荐系统的影响因素大致可归纳为4类^[16],即地理因素影响、签到序列影响、社交网络影响和地点属性信息影响。地理因素影响是根据空间轨迹数据进行推荐,如Tobler等^[17]在1970年提出第一地理学法则,该法则强调地理空间的距离是决定事物相似度的关键,距离越近,相似度越高。签到序列影响是基于时空数据进行推荐,如Saleem等^[18]提出的以树形位置轨迹模型为基础的同层聚类方法。该方法虽然可以形成较好的推荐结果,但对算力的需求过高。也有研究人员提出将一天的时间进行等分,将时间分段法与第一地理学法则进行融合,生成时间兴趣点推荐算法^[19]。社交网络影响是根据社交好友信息进行推荐,如Gao等^[20]提出轨迹相似度越高则彼此是好友关系的可能性越大,这样可以根据地理位置和社交关系为用户推荐兴趣点。地点属性信息影响是根据用户对被推荐地点的评价数据进行推荐,这类研究主要应用在影院评分、餐馆评分等方面。Zhao等^[21]利用评价信息作为访问地点的词向量,在进行兴趣点的相似性度量后给出推荐结果。Yin等^[22-23]提出了基于贝叶斯概率模型的推荐方法,计算出用户对新地点的感兴趣程度,结合城市热点和用户轨迹停留点生成兴趣点推荐列表。

本文在已有的研究基础上,提出了一种复合神经网络结构,用于实现可移动性兴趣点的推荐,并且进一步提高推荐结果的准确率。

2 算法结构

本文算法主要包括3个网络模块:模式GRU网络模块、点GRU网络模块以及门网决策模块。

(1)模式GRU网络模块:通过融合迁移学习策略完成模式学习,其目的是学习其他数据集中的动态时空运动模式,进而训练出能够对模式进行预测的子网。在学习过程中利用归一化的方法对数据进行预处理,将不同范围的数据整合到同一范围内。

(2)点GRU网络模块:通过融合对比学习策略完成点学习,其目的是扩展目标训练集的样本数量。在保证扩展后样本概率分布不变的情况下,提高网络的泛化能力,训练出能够根据当前时空点预测下一位置的子网。

(3)门网决策模块:在工作时起到切换开关的作用,负责根据输入的数据形式选择是模式GRU的结果作为输出还是点GRU的结果作为输出。

果作为输出还是点GRU的结果作为输出。

图1(a)为本文方法网络的计算流程。输入数据由点GRU网络模块、模式GRU网络模块和门网决策模块共同接收,由3个网络模块协同工作,输出最终结果。图1(b)为本文方法网络的数据和样本处理流程。其中有两条主线,即对比学习部分和迁移学习部分。对比学习部分:将主数据集分成测试集和训练集,然后利用对比学习策略将训练集数据样本进行扩展,利用本文提出的两步样本过滤法对扩展后的数据集进行过滤处理,以确保扩展后的数据集的样本概率分布不变,最终达到训练出有效的点GRU网络的目的。迁移学习部分:与主数据集相似的外部数据集能够提供有用的时空信息,利用迁移学习可以学到不同数据集在时空运动上呈现出的共同规律,这些共同规律也可看作是能有效提升网络预测性能的训练样本。这两部分的结果会在决策门网模块的控制下进行选择性输出,最终得到推荐结果。引入迁移学习和对比学习策略的目的以及优势在于:在已有数据集规模较小的情况下,仍然能够训练出预测准确度高、泛化能力强的模型。

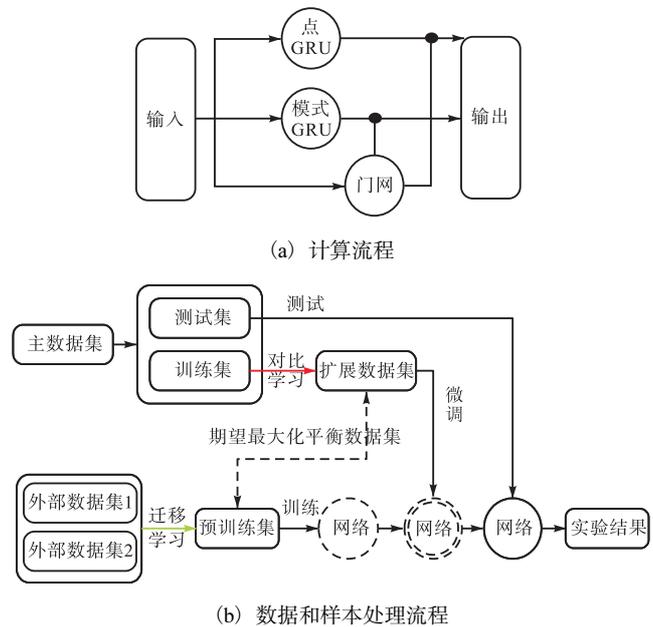


图1 网络的计算流程以及数据和样本处理流程

Fig. 1 Calculation process, data and sample processing flow of the network

2.1 GRU网络

在本文方法中,GRU网络^[24]负责分别与迁移学习策略和对比学习策略进行融合,使之能完成对点以及模式的预测。本文中 x 代表数据集中一个用户产生的连续时空位置信息, $x = \{\dots, x_{t-1}, x_t, x_{t+1}, \dots\}$ 。由于

内部数据点存在时间顺序, 因此 \mathbf{x} 本质为向量, 但在后面处理过程中需要对 \mathbf{x} 内部的数据点进行逐个处理, 为了方便, 将其表示成集合的形式, 其中 x_t 为用户在 t 时刻产生的数据(用户当前的时空位置), 作为本文 GRU 网络的输入部分. h_t^j 表示由用户当前的时空位置状态所得的隐状态.

$$h_t^j = (1 - z_t^j)h_{t-1}^j + z_t^j\tilde{h}_t^j \quad (1)$$

式中: t 为当前时刻, 在 t 时刻激活线性插值 h_t^j 和 h_{t-1}^j , 其中 j 是本程序中的计数项, 用于计数循环, 无实际计算意义. 更新门 z_t^j 为 t 时刻激活更新时的逻辑门, 决定了单元激活其中的更新信息.

$$z_t^j = \sigma(W_z x_t + U_z h_{t-1}^j) \quad (2)$$

目标激活 h_t^j 为

$$\tilde{h}_t^j = \tanh(W_r x_t + r_t \odot U_r h_{t-1}^j) \quad (3)$$

式中: σ 为 sigmoid 函数; W_z 、 W_r 为权重矩阵, 用于线性变换; U_z 和 U_r 分别为权重矩阵 W_z 、 W_r 变换后的表达形式; \odot 为哈达玛积 (Hadamard product); r_t 为一组重置单元, 若关闭重置门 r_t^j , 重置单元将使单元忘记先前的计算状态.

重置门 r_t^j 与更新门的计算类似, 为

$$r_t^j = \sigma(W_r x_t + U_r h_{t-1}^j) \quad (4)$$

最终得到用户下一时刻的预测点 y_t 作为本文 GRU 模块的输出

$$\text{GRU}(x_t) = y_t = x_t \oplus \tilde{h}_t^j \quad (5)$$

式中 \oplus 表示进行矩阵加法操作.

需要注意的是, GRU 网络处理过程中涉及连续时空点中时间 t 的概念, 因此输入和输出的结果下标要注明 t , 但在后面的决策门网模块中不涉及时间概念, 因此下标省略了时间 t .

2.2 基于迁移学习的模式 GRU 训练

利用迁移学习策略^[25]训练模式 GRU. 利用迁移学习扩展用于训练神经网络结构的数据集, 本文方法需要将已有的用户时空轨迹数据、特征等高层知识进行转化, 用于解决新的轨迹预测问题. 即使是不同的时空轨迹数据集, 其特征空间也具有一定的共同性, 这样就可以采用样本级迁移方法解决本质上相同的问题. 本文方法应用到主数据集和外部数据集, 其中主数据集分为训练集和测试集, 外部数据集采用的是相同领域的时空点数据集. 利用从外部数据集提取到的时空轨迹对训练数据进行扩容, 以实现运动模式这种高级知识的迁移学习.

如图 2 所示, 从时间和空间的数据点角度来看,

在迁移学习过程中, 应对用户的运动模式进行迁移学习, 而非对用户具体的时空点进行迁移.

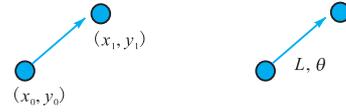


图 2 直角坐标转换为模式 GRU 的训练数据格式
Fig. 2 Convert rectangular coordinates to training data format of the mode GRU

譬如, A 用户在甲地, B 用户在乙地, 考虑到甲、乙两地可能相隔甚远, 且 A、B 用户所用的交通工具不同, 所以 A、B 两者的坐标无法直接进行匹配学习. 但是, 在加速时, 无论 A、B 用户是步行还是驾驶机动车, A、B 两者都会出现坐标逐渐稀疏, 在减速过程中坐标则会逐渐密集, 对于转弯等也有相似的运动模式. 因此, 应该对不同用户采集的数据使用密度聚类 (density-based spatial clustering of applications with noise, DBSCAN) 方法进行预先聚类, 以解决轨迹中出现的因长时间驻留某地, 而产生坐标重复采样的问题; 然后, 利用归一化方法, 消除用户产生数据的度量差异, 利用外部数据集构建多个此种模式序列供 GRU 学习, 进而完成时空轨迹模式的迁移, 并将模式 GRU 输出的预测结果重新映射到实际空间中, 达到利用模式 GRU 提供可借鉴的动态轨迹模式的目的. 从数据层级上实现动态运动模式的迁移学习是一个比较容易实现的操作方法, 但在本文中, 该方法需要着重构建时空点链信息, 供模式 GRU 进行学习.

2.3 基于对比学习的点 GRU 训练

使用对比学习策略^[26]对点 GRU 进行训练. 对时空数据 x 训练得到一个编码器 f , 使得 $H(f(x), f(x^+)) \gg H(f(x), f(x^-))$, 其中: x^+ 是和 x 相似的正样本, x^- 是与 x 不相似的负样本, H 是度量函数, 用于度量样本的相似度. 对比学习的损失函数 L_N 可表示为

$$L_N = -E_x \left[\log \frac{\exp(f(x)^T f(x^+))}{\exp(f(x)^T f(x^+)) + \sum_{j=1}^{N-1} \exp(f(x)^T f(x_j^-))} \right] \quad (6)$$

其中 E 为期望. 通常正负样本都将被应用, 且正负样本区分度越大越好, 而在本文, 仅使用生成的正样本对 GRU 进行训练, 以达到扩大 GRU 的训练样本集和增强网络泛化性能的目的.

如图 3 所示: 已知用户 A 的运动轨迹, 该轨迹所生成的正样本集用绿色标记, 代表该用户可能走到的

点(小范围区域);而负样本集用红色标记,代表该用户不可能走到的点(远距离区域).期望通过上述编码器找到最大区分度的正样本,将正样本集合进行筛选后,供 GRU 训练使用.

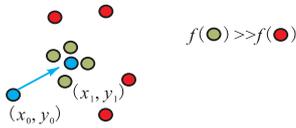


图3 对比学习策略(生成区分度高的正负样本)

Fig.3 Contrast learning strategies (generating positive and negative samples with high discrimination)

本文采用对比学习策略,对训练集进行扩容的同时,也要保证生成样本的可用性与合理性,因而设计两步样本过滤法对生成的数据进行筛选.

2.4 两步样本过滤法

对迁移学习算法和对比学习算法生成的样本进行两步过滤:第一步使用 MMD 方法^[27]确保样本的总体概率统计特征一致;第二步使用本文提出的 TrAdaBoost + EM 的算法,对样本计分后进行进一步筛选,最终保证训练样本数据扩展的有效性.

本文在 MMD 方法中设定两个分布: s 和 r 分别代表用户时空位置源数据集(source)和用户时空位置生成数据集(resulting).由源数据集与生成数据集组成了扩展数据集,通过函数 $\varphi(\cdot)$ 将数据点映射到再生核希尔伯特空间(reproducing kernel Hilbert space, RKHS),则 s 和 r 之间的 MMD 度量可定义为

$$\text{MMD}^2(s, r) = \sup_{\|\varphi\|_k \leq 1} E_{X^s \sim s} [\varphi(X^s)] - E_{X^r \sim r} [\varphi(X^r)] \quad (7)$$

式中: E 为期望, $\|\varphi\|_k \leq 1$ 为符合 RKHS 条件的一组函数.如果 s 和 r 相同,则 $\text{MMD}(s, r) = 0$;如果从 s 和 r 中分别选取样本,则 MMD 的实际估计为

$$\text{MMD}^2(D_s, D_r) = \left\| \frac{1}{M} \sum_{i=1}^M \varphi(x_i^s) - \frac{1}{N} \sum_{j=1}^N \varphi(x_j^r) \right\|_k^2 \quad (8)$$

为了筛选数据,并得到符合要求的数据集,在利用 MMD 方法完成第一步筛选后,得到 MMD 值较小、较符合要求的待过滤数据集 D_{Fs} . 下一步,利用 TrAdaBoost 结合 EM 的思想,对源数据集和生成数据集的样本进行判别并筛选,算法步骤如下:

输入:首先输入待过滤数据集 D_{Fs} , 设置标签为 0;生成数据集的训练集 D_t , 设置标签为 1;最大迭代次数设置为 N . 然后进行初始化操作,初始集合权重向量 $D_w = (D_{w_1}, \dots, D_{w_{n+m}})$; D_{Fs} 样本累计向量设为

$D_{ws} = (D_{ws_1}, \dots, D_{ws_n})$, D_t 样本累计向量设为 $D_{wt} = (D_{wt_1}, \dots, D_{wt_m})$.

For loop = 1, ..., N (循环步骤 1—8):

1. 令 $D = (D_{Fs} \cup D_t)$, 并将 D 随机划分为 $n + m$ 个集合,记为 D_1, \dots, D_{n+m} .

2. 两个集合的总子集数为 $n + m$ 个. 设定每个集合的权重为

$$p^t = w^t / \sum_{i=1}^{n+m} w_i^t \quad (9)$$

式中: p^t 为第 t 个集合的权重, w_i^t 为 t 集合中第 i 个元素的权重.

3. D 集合的统计参数 E, M 以及标签构成输入数据,进行 EM 分类.

4. E-step:

$$Q_i(z_i) = p(z_i | x_i, \theta_j) \quad (10)$$

式中: x_i 表示第 i 个样本; z_i 表示其对应的隐含数据 (x_i, z_i) , 为完全数据; θ_j 为模型参数; $p(\cdot)$ 表示概率.

5. M-step:

$$L(\theta, \theta_j) = \sum_{i=1}^n \sum_{z_i} Q_i(z_i) \log \frac{p(x_i, z_i; \theta)}{Q_i(z_i)} \quad (11)$$

式中: $L(\cdot)$ 为似然函数.

6. 得到 EM 的模型参数 θ 后,对 D 内的 $n + m$ 个集合进行判别,并计算错误率

$$\varepsilon_t = \sum_{i=n+1}^{n+m} \frac{w_i^t |h_t(x_i) - c(x_i)|}{\sum_{i=n+1}^{n+m} w_i^t} \quad (12)$$

式中: ε_t 表示错误率, $h_t(x_i)$ 表示二分类预测输出(值为 0 或 1), $c(x_i)$ 表示二分类实际输出.

7. 令

$$\beta_t = \varepsilon_t / (1 - \varepsilon_t) \quad (13)$$

并更新权重 w

$$w_i^{t+1} = \begin{cases} w_i^t \beta^{h_t(x_i) - c(x_i)}, & i = 1, \dots, n \\ w_i^t \beta^{-h_t(x_i) - c(x_i)}, & i = 1, \dots, n+m \end{cases} \quad (14)$$

8. 令每个集合的样本共享权重为 w_i^{t+1} , 并分别累计到 D_s 和 D_t 上,且 w_s 和 w_t 一同更新:

$$D'_s = D_s + w_i^{t+1} \quad (15)$$

$$D'_t = D_t + w_i^{t+1} \quad (16)$$

$$w'_s = w_s + D'_s \quad (17)$$

$$w'_t = w_t + D'_t \quad (18)$$

式中: D_s 为滤掉数据的数据集; D_t 为剩余数据的数据集; w_s 和 w_t 分别为 D_s 和 D_t 集合权重.

将数据集 D_t 中的数据进行重新整理得到集合 D_{trans} .

输出: $D_{\text{trans}} = \{x_1, \dots, x_k\}$, 且 $x_i \in D_i$; D_{trans} 样本对应的权重为 $W_{\text{trans}} = \{w_{\text{trans},1}, \dots, w_{\text{trans},k}\}$, $w_{\text{trans},i} \in \{D_s \cup D_i\}$ 且 $\forall w_i < \min(W_{\text{trans}})$.

经过两步样本过滤法的处理, 得到合格的训练样本集后, 再根据源数据集内的轨迹形式(根据用户 ID、时间顺序)将该集合中的数据还原成各个空间轨迹 \mathbf{x} , $\mathbf{x} = \{\dots, x_{t-1}, x_t, x_{t+1}, \dots\}$, \mathbf{x} 作为训练 GRU 网络的输入数据. 上述过滤方法输出中的权重用于判断样本是否合格, 因此仅在过滤方法内部有效, 不作为 GRU 网络的输入数据.

2.5 门网

门网^[28]是混合专家网络的一部分, 根据功能差异分为竞争型门网和协作型门网. 本文选择并构建竞争型门网. 在决策门网模块中, 将用户产生的连续时空位置信息 \mathbf{x} 视为输入, 其中 $\mathbf{x} = \{\dots, x_{t-1}, x_t, x_{t+1}, \dots\}$, 根据输入数据的不同形式, 由门网的决策结果决定选择点 GRU 或模式 GRU 的计算结果作为最终输出.

由于系统的输入为连续序列, 序列里的数据项包括时间、空间位置, 用户的移动状态(例如停留、徘徊)可能会导致序列里的数据在位置项上发生重复,

Example of check-in information

[user]	[check-in time]	[latitude]	[longitude]	[location ID]
196514	2010-07-24T13:45:06Z	53.364 811 9	-2.272 346 583 3	145064
196514	2010-07-24T13:44:58Z	53.360 511 233	-2.276 369 017	1275991
196514	2010-07-24T13:44:46Z	53.365 389 594 5	-2.275 408 704 6	376497
196514	2010-07-24T13:44:38Z	53.366 370 583 3	-2.270 076 433 3	98503
196514	2010-07-24T13:44:26Z	53.367 408 752 4	-2.278 381 347 7	1043431
196514	2010-07-24T13:44:08Z	53.367 566 337 7	-2.278 631 763	881734
196514	2010-07-24T13:43:18Z	53.367 964 062 6	-2.279 294 368 9	207763
196514	2010-07-24T13:41:10Z	53.364 905	-2.277 082 4	1042822

图 4 用户签到信息示例

Fig. 4 Example of user check-in information

(2) 外部数据集: 迁移学习策略部分涉及的数据集有 Facebook V: Predicting Check Ins 数据集和 Yelp Dataset Check Ins 数据集. 这两个外部数据集的数据格式与主数据集核心一致, 都能够提供用户的时空轨迹信息, 包括对比实验中涉及的最关键的经度、纬度、时间 3 个信息.

3.1.2 数据集分割

在主数据集中, 对每个用户的记录信息进行分割: 20% 训练集, 80% 测试集. 与通常实验数据分割不同, 本实验将 20% 作为训练数据, 80% 作为测试数据, 目的在于体现出本文所提出的方法能够以少量的

因此单从序列里数据的个数来看, 无法判断应该将哪个专家的结果作为输出.

本方法需要训练出一个门网, 用于精准判断输入数据的形式, 进而选择合适的子网结果作为输出. 决策门网模块将输入数据进行处理, 将距离较近甚至相同的点合并. 若得到的数据串较短(以长度 1 为参考), 会激活点 GRU 的输出; 若得到的数据串较长, 将激活模式 GRU 的输出.

3 实验与结果分析

3.1 实验数据集以及数据预处理

3.1.1 数据集

本文使用了 1 个主数据集和 2 个外部数据集. 主数据集用于训练并测试点 GRU 网络以及测试模式 GRU 网络, 贯穿整个结构的训练以及测试; 2 个外部数据集都用于训练模式 GRU 网络, 因此也称为迁移学习模式中的外部数据集.

(1) 主数据集: Gowalla check-in dataset, 该数据集是公开数据集, 其中数据项以及数据格式如图 4 所示.

数据驱动神经网络系统的优势. 去除主数据集中签到记录过少的用户内容, 比如数据集中 ID 为 196585 的用户只有两个时空点记录.

3.2 对比实验

3.2.1 评价指标

与其他实验中考察的准确率、召回率等指标不同, 对本文可移动兴趣点的推荐效果评价较为直接有效的方法是计算具体的空间距离误差, 所以多次实验后, 对算法的预测误差进行统计分析是关键且切实可行的手段. 在 t 时刻预测 $t+1$ 时刻的兴趣点位置, 需记录并计算预测点 $x' = (ptx'_{t+1}, pty'_{t+1})$ 与真实点

$x = (ptx_{t+1}, pty_{t+1})$ 的欧氏距离 D , 则

$$D(x', x) = \sqrt{(ptx'_{t+1} - ptx_{t+1})^2 + (pty'_{t+1} - pty_{t+1})^2} \quad (19)$$

式中: x' 表示预测坐标点, x 表示实际坐标点, ptx'_{t+1} 表示预测坐标点的横坐标, pty'_{t+1} 表示预测坐标点的纵坐标, ptx_{t+1} 表示实际坐标点的横坐标, pty_{t+1} 表示实际坐标点的纵坐标。

分别整合不同算法的误差结果并按照从小到大进行排列, 得到误差的有序集合 Q , 进而得到预测结果的误差分布, 再统计误差结果后算出四分位数^[29], 即可得到不同算法的预测误差对比结果。四分位数 Q_1 、 Q_2 、 Q_3 分别为 Q 中所有数值由小到大排列后第 25% 的数字、第 50% 的数字、第 75% 的数字。

3.2.2 基准对照实验的设计

采用的基准方法为马尔可夫模型 + DBSCAN 方法和高斯混合模型 + DBSCAN 方法。马尔可夫模型和高斯混合模型都需要结合 DBSCAN 方法进行协同工作。其步骤如下: 首先将轨迹点簇根据不同用户整合成各个用户的路线; 然后分别利用马尔可夫模型和高斯混合模型对用户的路径进行预测; 最后利用 DBSCAN 方法将预测结果进行整合, 得到用户聚集较多的时空位置, 并将该位置推荐给用户。

3.3 结果分析

3.3.1 短期可移动性兴趣点推荐对比

短期可移动性兴趣点的预测结果要求预测未来的 1 个时空点并进行 3 轮对比, 每一轮对用户集合随机抽取 50% 的测试样本进行测试, 统计出各个用户预测误差的四分位数, 并据此绘制四分位数箱式图。

图 5 展示了 3 种对比算法的短期对比实验结果。由此对比实验结果可见, 本文方法 (A) 和马尔可夫模型 + DBSCAN 方法 (B) 比高斯混合模型 + DBSCAN 方法 (C) 在性能方面表现更优。在统计预测距离误差的中位数方面, 本文的方法比高斯混合模型 + DBSCAN 方法的中位数要低 32% ~ 65%, 比马尔可夫模型 + DBSCAN 方法的中位数要低 15% ~ 47%; 可见本文方法所提供的预测结果具有更好的稳定性。同时在几轮预测实验的距离误差中位数方面, 可以发现马尔可夫模型 + DBSCAN 方法略优于高斯混合模型 + DBSCAN 方法。

统计结果的下四分位数能够反映可移动性兴趣点预测结果的精准度。从表 1 可见, 在 3 轮对比实验中, 本文方法有两轮相比于其他两种方法有着最优的下四分位数结果, 特别是在第 3 轮实验中, 有 25% 的

预测结果的绝对距离小于 0.53 km (箱式图的下边界)。这个预测结果表明本文方法在现实世界中也能够较好地预测可移动兴趣点的短期位置。

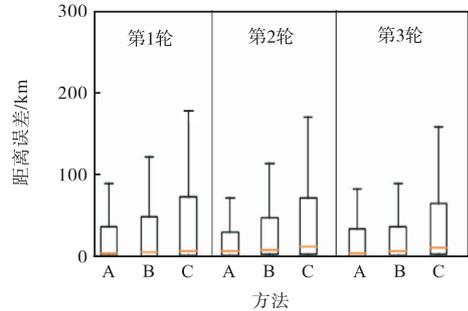


图 5 短期可移动性兴趣点推荐对比结果

Fig. 5 Comparison results of short-term mobile points of interest recommendation

在对比上四分位数的统计结果方面, 表 1 中本文方法也体现了较好的性能, 具有更小的误差。除了马尔可夫模型 + DBSCAN 方法在第 3 轮对比实验中表现出较为理想的预测误差上四分位数, 本文方法比另外两种方法的预测误差上四分位数低 25% ~ 30%。

表 1 短期可移动性兴趣点推荐误差统计

Tab. 1 Error statistics of short-term mobile points of interest recommendation km

轮次	四分位数位置	A	B	C
第 1 轮	上四分位	37.32	49.61	73.58
	中位数	4.93	5.80	7.34
	下四分位	2.03	0.93	2.22
第 2 轮	上四分位	30.12	47.68	71.52
	中位数	6.68	8.41	12.62
	下四分位	1.62	2.34	3.51
第 3 轮	上四分位	33.75	36.63	65.18
	中位数	3.73	7.07	10.77
	下四分位	0.53	0.98	2.93

在短期可移动性兴趣点推荐对比实验中, 本文方法具有最优的中位数和上四分位数, 在预测距离误差方面展示了稳定性和准确性。但是, 其下四分位数在其中一轮实验中没有取得优势, 可能原因在于引入对比学习策略扩展训练集的样本, 使短期预测的误差变大。但总体而言, 引入对比学习策略和迁移学习策略能够很好地降低预测误差, 加强预测效果的稳定性。

3.3.2 中期可移动性兴趣点推荐对比

中期可移动性兴趣点的预测结果要求预测未来的第 3 个时空点, 与短期可移动性兴趣点的预测对比实验类似, 进行 3 轮对比, 并采用同样的统计对比方式进行评估。图 6 展示了 3 种算法的中期对比实验结果。从此对比实验结果可见, 高斯混合模型 +

DBSCAN 方法已无法保证预测的准确性和稳定性, 其原因在于高斯混合模型的内在机制不如马尔可夫模型和 GRU 网络具有较好的短期、中期的记忆和预测能力. 在统计预测距离误差的中位数方面, 本文方法的误差比另外两种对比算法的误差低 35% ~ 65%, 在 3 轮实验中均未超过 10 km.

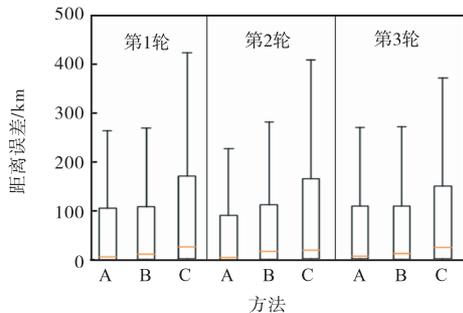


图 6 中期可移动性兴趣点推荐对比结果

Fig. 6 Comparison results of mid-term mobile points of interest recommendations

表 2 中列出了详细的中期可移动性兴趣点推荐对比统计结果. 马尔可夫模型 + DBSCAN 方法具有较好的表现, 其学习能力和记忆能力都保证了预测结果的准确性. 但是, 由于该方法缺少了对比学习策略和迁移学习策略的辅助, 在预测过程中泛化能力不足, 预测的误差中位数较高, 仅在第 1 轮实验中有最低值, 为 13.36 km. 而本文方法在对比实验中表现出更好的泛化能力, 并且因为将其他数据集中的时空模式信息进行了迁移, 所以表现出更好的预测准确性.

表 2 中期可移动性兴趣点推荐误差统计

Tab. 2 Error statistics of mid-term mobile points of interest recommendation km

轮次	四分位数位置	A	B	C
第 1 轮	上四分位	106.82	109.36	171.75
	中位数	8.65	13.36	28.10
	下四分位	1.80	2.39	4.74
第 2 轮	上四分位	92.23	114.50	166.63
	中位数	7.20	18.73	22.09
	下四分位	1.64	3.16	4.53
第 3 轮	上四分位	111.06	111.09	151.44
	中位数	9.67	14.72	27.53
	下四分位	3.69	3.02	4.42

对比实验的结果表明, 在解决可移动性兴趣点推荐的问题时, 本文提出的方法能够同时发挥对比学习策略和迁移学习策略的优势, 在样本过滤后能保证预测结果的分布更稳定且降低预测误差, 算法性能优于马尔可夫模型 + DBSCAN 方法以及高斯混合模型 + DBSCAN 方法.

4 结 语

本文从可移动兴趣点的推荐及预测问题入手, 提出了一种复合神经网络结构. 为了提高训练网络的效果和增强网络预测问题时的泛化能力, 将基于样本的迁移学习策略和对比策略进行了融合, 并使用 MMD 方法和本文提出的 TrAdaBoost + EM 算法对扩容样本进行过滤, 保证了样本的有效性、可用性、分布一致性, 最终使得复合神经网络结构能够很好地激活内部专家决策结构进行推荐.

本文提出的复合神经网络和学习策略架构比较灵活, 在后续算法改进研究中, 可以考虑变换学习策略或融入更多的学习策略、扩展专家网络结构等.

参考文献:

- [1] DJUKNIC G M, RICHTON R E. Geolocation and assisted GPS[J]. IEEE Computer, 2001, 34(2): 123-125.
- [2] DEY A, HIGHTOWER J, DE LARA E, et al. Location-based services[J]. IEEE Pervasive computing, 2010, 9(1): 396-402.
- [3] ISINKAYE F O, FOLAJIMI Y O, OJOKOH B A. Recommendation systems: principles, methods and evaluation[J]. Egyptian informatics journal, 2015, 16(3): 261-273.
- [4] 杨丽, 王时绘, 朱博. 基于动态和静态偏好的兴趣点推荐算法[J]. 计算机应用, 2021, 41(2): 398-406.
- [5] 史覃覃, 韩文军, 陶宁, 等. 利用时空轨迹挖掘移动对象聚集移动模式[J]. 小型微型计算机系统, 2019, 40(5): 1099-1106.
- [6] 段淳林, 宋成. 用户需求、算法推荐与场景匹配: 智能广告的理论逻辑与实践思考[J]. 现代传播(中国传媒大学学报), 2020, 42(8): 119-128.
- [7] 林志威. 行人移动轨迹中基于意图增强的下一个地点预测模型的研究[D]. 杭州: 浙江大学, 2020.
- [8] 周明, 卓为, 郭志鹏. 基于时空信息的交通流预测[J]. 工业控制计算机, 2020, 33(4): 42-45.
- [9] CLIFF A D, HAGGETT P, ORD J K, et al. Elements of spatial structure[M]. Cambridge: Cambridge University Press, 1975.
- [10] MARTIN R L, OEPPEN J E. The identification of regional forecasting models using space-time correlation functions[J]. Transactions of the institute of British geographers, 1975, 66: 95-118.
- [11] CHERKASSKY V, MA Y. Practical selection of SVM parameters and noise estimation for SVM regression[J].

- Neural networks, 2004, 17(1): 113–126.
- [12] ZHANG G P. Time series forecasting using a hybrid ARIMA and neural network model[J]. Neurocomputing, 2003, 50: 159–175.
- [13] KRIZHEVSKY A, SUTSKEVER I, HINTON G E, et al. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84–90.
- [14] HE Z, CHOW C Y, ZHANG J D. STANN: a spatio-temporal attentive neural network for traffic prediction[J]. Institute of electrical and electronics engineering access, 2018, 7: 4795–4806.
- [15] 杨立宁, 李艳婷. 基于 SVD 和 ARIMA 的时空序列分解和预测[J]. 计算机工程, 2021, 47(3): 53–61.
- [16] 于正一. 兴趣点推荐系统研究综述[J]. 德州学院学报, 2020, 36(4): 52–59.
- [17] TOBLER W R. A computer movie simulating urban growth in the Detroit region[J]. Economic geography, 1970, 46: 234–240.
- [18] SALEEM M A, LEE Y K, LEE S. Dynamicity in social trends towards trajectory based location recommendation[C]//ICOST. International Conference on Smart Homes and Health Telematics. Berlin: Springer, 2013: 86–93.
- [19] LI X, JIANG M, HONG H, et al. A time-aware personalized point-of-interest recommendation via high-order tensor factorization[J]. ACM Transactions on information systems, 2017, 35(4): 1–23.
- [20] GAO H, TANG J, LIU H. Exploring social-historical ties on location-based social networks[EB/OL]. [2021–11–23]. https://www.researchgate.net/publication/268348632_Exploring_Social-Historical_Ties_on_Location-Based_Social_Networks.pdf.
- [21] ZHAO K, CONG G, YUAN Q, et al. SAR: a sentiment-aspect-region model for user preference analysis in geo-tagged reviews[C]//IEEE. 2015 IEEE 31st International Conference on Data Engineering. New York: IEEE, 2015: 7113324.
- [22] YIN H, SUN Y, CUI B, et al. LCARS: a location-content-aware recommender system[C]//ACM. Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Chicago: ACM, 2013: 221–229.
- [23] YIN H, CUI B, SUN Y, et al. LCARS: a spatial item recommender system[J]. ACM Transactions on information systems, 2014, 32(3): 1–37.
- [24] DEY R, SALEMT F M. Gate-variants of gated recurrent unit (GRU) neural networks[C]//IEEE. 2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS). New York: IEEE, 2017: 8053243.
- [25] 葛轶洲, 刘恒, 王言, 等. 小样本困境下的深度学习图像识别综述[J]. 软件学报, 2022, 33(1): 193–210.
- [26] 张玲玲, 陈一苇, 吴文俊, 等. 基于对比约束的可解释小样本学习[J]. 计算机研究与发展, 2021, 58(12): 2573–2584.
- [27] SUTHERLAND D J, TUNG H Y, STATHMANN H, et al. Generative models and model criticism via optimized maximum mean discrepancy[EB/OL]. [2021–11–23]. <https://arxiv.org/pdf/1611.04488.pdf>.
- [28] 郭亮, 杨兴耀, 于炯, 等. 基于注意力机制和门控网络相结合的混合推荐系统[J/OL]. 计算机科学: 1–10 [2021–11–23]. <http://kns.cnki.net/kcms/detail/50.1075.TP.20220223.1411.002.html>.
- [29] 张云华. 统计学中四分位数的计算[J]. 中国高新技术企业, 2009(20): 173–174.

责任编辑: 郎婧

(上接第 53 页)

- [25] 陈亚瑞, 杨巨成, 史艳翠, 等. 概率生成模型变分推理方法综述[J]. 计算机研究与发展, 2022, 59(3): 617–632.
- [26] 陈亚瑞, 蒋硕然, 杨巨成, 等. 混合变分自编码[J]. 计算机研究与发展, 2020, 57(1): 136–144.
- [27] 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.
- [28] NETZER Y, WANG T, COATES A, et al. Reading digits in natural images with unsupervised feature learning[EB/OL]. [2021–11–01]. http://ufldl.stanford.edu/housenumbers/nips_2011_housenumbers.pdf.
- [29] KINGMA D P, BA J. Adam: a method for stochastic optimization[EB/OL]. [2021–11–01]. <https://arxiv.org/abs/1412.6980>.

责任编辑: 郎婧