



计算机科学

COMPUTER SCIENCE

基于时序知识图谱嵌入的短期地铁客流量预测

毛慧慧, 赵小乐, 杜圣东, 滕飞, 李天瑞

引用本文

毛慧慧, 赵小乐, 杜圣东, 滕飞, 李天瑞. 基于时序知识图谱嵌入的短期地铁客流量预测[J]. 计算机科学, 2023, 50(7): 213-220.

MAO Huihui, ZHAO Xiaole, DU Shengdong, TENG Fei, LI Tianrui. [Short-term Subway Passenger Flow Forecasting Based on Graphical Embedding of Temporal Knowledge](#) [J]. Computer Science, 2023, 50(7): 213-220.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于遗传算法的恶意软件对抗样本生成方法](#)

Adversarial Malware Generation Method Based on Genetic Algorithm

计算机科学, 2023, 50(7): 325-331. <https://doi.org/10.11896/jsjcx.220800176>

[基于深度学习的活跃IPv6地址预测算法](#)

Deep Learning-based Algorithm for Active IPv6 Address Prediction

计算机科学, 2023, 50(7): 261-269. <https://doi.org/10.11896/jsjcx.220700076>

[面向单一背景的改进RetinaNet目标检测方法研究](#)

Study on Single Background Object Detection Oriented Improved-RetinaNet Model and Its Application

计算机科学, 2023, 50(7): 137-142. <https://doi.org/10.11896/jsjcx.220500066>

[面向自动驾驶的三维目标检测综述](#)

Review of 3D Object Detection for Autonomous Driving

计算机科学, 2023, 50(7): 107-118. <https://doi.org/10.11896/jsjcx.220700090>

[探索站点时空移动模式:长短期交通预测框架](#)

Exploring Station Spatio-Temporal Mobility Pattern:A Short and Long-term Traffic Prediction Framework

计算机科学, 2023, 50(7): 98-106. <https://doi.org/10.11896/jsjcx.220900109>

基于时序知识图谱嵌入的短期地铁客流量预测

毛慧慧 赵小乐 杜圣东 滕飞 李天瑞

西南交通大学计算机与人工智能学院 成都 611756

(hhmslt@163.com)

摘要 短期地铁客流量预测任务是城市智能地铁运营工作的重要组成部分,旨在预测未来短时间内地铁站点的客流量。针对现有方法未能充分利用站点的流入流出客流量信息的问题,提出了一种基于时序知识图谱嵌入(Temporal Knowledge Graph Embedding)结合残差网络(ResNet)和长短期记忆网络(LSTM)的短期地铁客流量预测方法,简称 TKG-ResLSTM。首先,基于地铁客流量数据构建地铁客流量时序知识图谱,并使用时序知识图谱嵌入技术从中获取地铁站点的动态客流量模式。然后,将抽取出的动态客流量模式转换成动态相似矩阵,应用到基于深度学习的地铁客流量预测框架中完成地铁客流量预测任务。最后,利用北京地铁和 A 市地铁客流量数据集,分别在 10 min, 15 min 和 30 min 的时间间隔下进行实验评估。结果表明 TKG-ResLSTM 能够有效地抽取地铁站点的动态客流量模式,在不使用外部信息的情况下,相比 ResLSTM, TKG-ResLSTM 在北京地铁数据集 10min 时间间隔下的预测均方根误差降低了 0.41。

关键词: 深度学习; 时序知识图谱; 时空预测; 动态嵌入; 城市地铁网络

中图法分类号 TP391

Short-term Subway Passenger Flow Forecasting Based on Graphical Embedding of Temporal Knowledge

MAO Huihui, ZHAO Xiaole, DU Shengdong, TENG Fei and LI Tianrui

School of Computing and Artificial Intelligence, Southwest Jiaotong University, Chengdu 611756, China

Abstract Subway short-term passenger flow forecasting is an essential component in urban subway operation, and it aims to forecast the passenger flow of subway stations in a short time in the future. Aiming at the problem that the existing methods fail to make full use of the passenger flow information of stations, a short-term subway passenger flow forecasting method based on temporal knowledge graph embedding combined with residual network and long short-term memory network is proposed, which is called TKG-ResLSTM. First, we use subway passenger flow data to construct a temporal knowledge graph of subway passenger flow, and apply the graphical embedding of temporal knowledge to obtain the dynamic patterns of subway stations passenger flow. Then, the extracted dynamic patterns of passenger flow are converted into dynamic similarity matrices and applied to the forecasting architecture of subway passenger flow based on deep learning to complete the subway passenger flow forecasting task. Finally, experimental evaluations are carried out at time granularities of 10 min, 15 min, and 30 min using the Beijing subway and city A subway passenger flow datasets, respectively. Experimental results show that TKG-ResLSTM can effectively extract the dynamic patterns of subway stations passenger flow. Without using external information, TKG-ResLSTM reduces the root mean square error of forecasting by 0.41 compared with ResLSTM in the time granularity of 10 min of the Beijing subway dataset.

Keywords Deep learning, Temporal knowledge graph, Spatial-Temporal forecasting, Dynamic embedding, Citywide metro network

1 引言

地铁作为城市主要的公共交通工具,在人们的日常出行中起到关键作用。地铁客流量预测是智慧城市建设的重要任务,能够有效地缓解交通拥堵,还能用于按需调整地铁运营

时间,实现城市地铁运营智能化。

早期的地铁客流量预测模型基于数理统计,预测精度低、适应性差、难以捕捉客流量的动态变化^[1],只能预测一个或多个站点的客流量,无法应对大规模的地铁系统。随着机器学习的发展,学者们开始使用 BP 神经网络^[2]、概率树^[3]、支持

到稿日期:2022-06-13 返修日期:2022-11-10

基金项目:国家重点研发计划(2019YFB2101801)

This work was supported by the National Key R&D Program of China(2019YFB2101801).

通信作者:李天瑞(trli@swjtu.edu.cn)

向量机^[4]、动态贝叶斯网络^[5],以及数理统计和机器学习的混合模型^[6]进行地铁客流量预测。这些模型比单纯的数理统计模型预测精度更高,但无法挖掘出地铁站点之间复杂的时空信息,预测能力有限。深度学习作为机器学习的重要分支,具备从大量数据中抽取复杂特征和数据之间相关性的能力^[7],因此学者们开始尝试使用深度学习来解决地铁客流量预测问题。

使用深度学习进行地铁客流量预测的关键是特征抽取,目前主要考虑时序特征、空间特征以及外部特征。其中,时序特征主要包括邻近性、周期性和趋势性;空间特征主要包括站点之间的邻接性、连通性和功能相似性^[8]。LSTM(Long Short-Term Memory)用于捕获时序特征^[9-10],CNN(Convolution Neural Network)用于捕获空间相关性^[11-12],GCN(Graph Convolutional Network)基于站点的邻接矩阵能同时捕获站点之间的时空信息^[13-14]。为了进一步提高交通流量预测模型的准确率,影响交通流量的外部信息被引入预测模型^[15-16]中,如天气、空气质量、节假日等。但是,外部信息的获取难度通常较大,而且对模型的作用有限。因此,需要研究一种能够充分利用交通流量信息的方法来进一步提高预测模型的准确率。

Yao等^[12]在进行区域出租车需求预测时指出,物理距离远的区域间也存在重要的相关性。他们通过对具有相似时间模式的区域之间的相关性进行建模,捕获了远距离区域之间的交互特征。受此启发,本文希望使用地铁客流量时序知识图谱,充分利用地铁站点的流入流出客流量信息挖掘远距离站点的动态客流量模式,建模远距离站点之间的相关性,从而提高地铁客流量的预测效果。

近年来,知识图谱凭借其强大的知识表示能力受到广泛关注。知识图谱可以有效地组织和表示知识,以便在相关任务中有效地利用知识^[17],知识图谱嵌入(Knowledge Graph Embedding, KGE)是知识图谱能够在下游任务中得以应用的有效途径。加入时间信息的知识图谱被称为时序知识图谱(Temporal Knowledge Graphs, TKG),时序知识图谱嵌入能够有效地抽取知识图谱中事实元组随时间变化的动态信息。

鉴于已有的地铁客流量预测模型未能充分地挖掘出站点流入流出客流量中蕴含的站点相关性,本文提出了一种基于时序知识图谱嵌入的短期地铁客流量预测方法(TKG-ResLSTM),利用时序知识图谱嵌入技术挖掘站点流入流出客流量中蕴含的站点相关性,避免了现有预测模型堆叠外部信息的问题。本文在北京地铁和A市地铁数据集上对TKG-ResLSTM进行验证,实验结果表明TKG-ResLSTM的预测性能与当前的先进方法相比具有优越性。

2 相关工作

本文所提方法涉及地铁客流量预测和时序知识图谱嵌入两个研究领域,下面介绍相关工作。

2.1 地铁客流量预测模型

由于深度学习模型能够从大规模数据中有效地抽取复杂的相关性特征,具有很强的特征提取和表达能力,因此现有的短期地铁客流量预测模型大多基于深度学习方法^[18-20]。

ResLSTM^[16]是一个基于深度学习的短期地铁客流量

预测框架。其使用ResNet, GCN和加入注意力机制的LSTM(Attention LSTM)完成面向整个地铁网络的短期地铁客流量预测,考虑了地铁站点之间的时空相关性、拓扑关系以及天气条件和空气质量。

ResLSTM模型包括4个分支:分支1和分支2分别用地铁站点的流入和流出客流量捕获时空特征,分支3用邻接矩阵提取地铁网络的拓扑信息,分支4代表天气和空气质量对客流量的影响。由于地铁客流量随时间变化且具有明显的周期性,分支1和分支2中考虑了客流量的3种模式,分别为实时模式、天模式、周模式。以分支1为例,流入客流量的时间序列由式(1)给出:

$$\mathbf{X}_{I,p} = \begin{pmatrix} x_{1,p}^{t-n} & x_{1,p}^{t-n+1} & \cdots & x_{1,p}^t \\ \vdots & \vdots & \ddots & \vdots \\ x_{|I|,p}^{t-n} & x_{|I|,p}^{t-n+1} & \cdots & x_{|I|,p}^t \end{pmatrix} \quad (1)$$

其中, $|I|$ 代表整个地铁网络中站点的数量, t 代表每个站点的历史时间步长。按地铁线路排序,每条线路中相邻的站点被放置在相邻的行中。 p 代表3种不同的模式, $\mathbf{X}_{I,r}$, $\mathbf{X}_{I,d}$, $\mathbf{X}_{I,w}$ 分别代表当天、前一天和前一一周在第 $t-n$ 到第 t 个时间间隔的流入客流量时间序列。为了预测第 $t+1$ 个时间间隔的流入量,ResLSTM将流入客流量的3种模式串接成一个输入,如式(2)所示:

$$\mathbf{I}_1 = (\mathbf{X}_{I,r}, \mathbf{X}_{I,d}, \mathbf{X}_{I,w}) \quad (2)$$

由于邻接矩阵只能捕获近距离站点之间的相关性,因此本文将基于站点动态客流量模式的动态相似矩阵引入ResLSTM模型,这样就能够同时捕获近距离和远距离站点之间的相关性,从而进一步提高模型的预测能力。

2.2 时序知识图谱嵌入模型

知识图谱是结构化的事实表示,由实体和关系构成。知识图谱嵌入旨在将知识图谱中的实体和关系嵌入到一个低维的连续语义空间中,已被广泛应用于知识驱动任务^[21]。静态知识图谱用三元组(头实体,关系,尾实体)表示事实,即 (s, r, o) 。时序知识图谱用四元组(头实体,关系,尾实体,时间)表示事实,即 (s, r, o, t) , $s, o \in \mathcal{V}, r \in \mathcal{R}, t \in \mathcal{T}, (s, r, o, t) \in \mathcal{G}$ 。其中 \mathcal{V} 是实体集合, \mathcal{R} 是关系集合, \mathcal{T} 是时间戳集合, \mathcal{G} 是时序知识图谱中包含的所有事实元组的集合。时序知识图谱嵌入在静态知识图谱嵌入的基础上考虑了事实成立的时间信息,把静态知识转换成动态知识,能够更好地表示事实随时间动态变化的过程。

根据嵌入过程中时间信息的融合方式划分,时序知识图谱嵌入方法可以分为3类:关系动态嵌入、实体动态嵌入和时序关系依赖嵌入。关系动态嵌入方法将时间信息与关系结合,然后通过关系进一步影响头实体和尾实体的嵌入^[22]。实体动态嵌入方法以实体嵌入为中心,直接把时间信息融合到头实体和尾实体的嵌入中^[23]。时序关系依赖嵌入方法利用时序关系依赖信息完成时序知识图谱的动态嵌入^[24]。

地铁网络是一个动态的稳定网络,其动态性主要体现在地铁网络中的客流量会随时间动态变化,其稳定性由地铁的固有属性决定,地铁建成后不会频繁增加或修改站点,这与常见的事件知识图谱构成的事件网络有很大不同。此外,本文

认为无论采用何种方式构建关于地铁的时序知识图谱,地铁站点一定是其中最重要的实体,并且该实体的特征相对稳定。因此,本文认为属于实体动态嵌入方法的 DEEMB(Diachronic Entity Embedding)模型^[23]更适合与地铁相关的时序知识图谱的嵌入。

DEEMB 模型通过 DEEMB 函数捕获实体在任意时间的动态特征,在实体嵌入中添加时间信息,将已有的静态知识图谱嵌入模型转换为动态的时序知识图谱嵌入模型。DEEMB 模型中给出了一种表现优异的 DEEMB 函数。令 $\mathbf{z}'_v \in \mathbb{R}^d$ 是 $DEEMB(v, t)$ 中的一个向量,代表实体 v 在时间 t 的动态嵌入, $DEEMB(v, t) = (\dots, \mathbf{z}'_v, \dots)$ 。 \mathbf{z}'_v 的定义如式(3)所示:

$$\mathbf{z}'_v[n] = \begin{cases} \mathbf{a}_v[n] \cdot \sigma(\mathbf{w}_v[n] \cdot t + \mathbf{b}_v[n]), & 1 \leq n \leq \gamma d \\ \mathbf{a}_v[n], & \gamma d < n \leq d \end{cases} \quad (3)$$

其中, d 是实体向量的维度, $\mathbf{a}_v \in \mathbb{R}^d$, $\mathbf{w}_v \in \mathbb{R}^{\gamma d}$, $\mathbf{b}_v \in \mathbb{R}^{\gamma d}$ 是具有可学习参数的特定于实体的向量, \mathbf{a}_v 可以看作实体的静态嵌入, $\mathbf{a}_v[n]$ 代表向量 \mathbf{a}_v 中的第 n 个元素, σ 是一个激活函数。DEEMB 模型认为实体的部分特征随时间改变,其余特征固定不变。在式(3)中, \mathbf{a}_v 的前 γd 个元素用于捕获实体随时间变化的特征,剩余的 $(1-\gamma)d$ 个元素用于捕获实体的静态特征。 $0 \leq \gamma \leq 1$ 是一个控制随时间变化的特征在实体嵌入中所占百分比的超参数。

本文使用 DEEMB 模型完成地铁客流量时序知识图谱嵌入,为构建地铁网络的动态相似矩阵奠定了基础。

3 基于时序知识图谱嵌入的短期地铁客流量预测方法

本文提出的 TKG-ResLSTM 方法主要包括 4 个部分:1) 利用地铁站点的流入流出客流量数据构建地铁客流量时序知识图谱;2) 面向地铁客流量时序知识图谱生成实体(地铁站点)的动态嵌入,即挖掘地铁站点的动态客流量模式;3) 利用地铁站点的动态嵌入生成动态相似矩阵,构建站点之间的近距离相关性和远距离相关性;4) 将动态相似矩阵融入用于短期地铁客流预测的深度神经网络,进行短期地铁客流量预测。本节将分别对以上 4 个部分进行详细介绍。

3.1 问题定义

本文的目标是用地铁网络中所有站点的历史客流值和地铁网络的动态相似矩阵,预测未来一段时间内所有地铁站点的进站客流量。令 $x_i^{a,t} \in \mathbb{R}$ 代表第 i 个站点在第 t 个时间间隔的第 a 个特征的值, $A = \{a_1, a_2, \dots, a_{|A|}\}$ 代表特征集合,本文中 A 包括进站客流量、出站客流量,以及动态相似矩阵; $\mathbf{X}_i^t = \{x_i^{a_1}, x_i^{a_2}, \dots, x_i^{a_{|A|}}\} \in \mathbb{R}^{|A| \times |A|}$ 代表所有站点在第 t 个时间间隔的所有特征; $\mathbf{Y}_i^t = \{y_i^1, y_i^2, \dots, y_i^{|I|}\} \in \mathbb{R}^{|I|}$ 代表所有站点在第 t 个时间间隔的进站客流量。本文的预测任务可以定义为 $\mathbf{Y}_i^{t+1} = f(\mathbf{X}_i^{t-n}, \mathbf{X}_i^{t-n+1}, \dots, \mathbf{X}_i^t)$ 。

3.2 构建地铁客流量时序知识图谱

本文使用的北京地铁和 A 市地铁客流量数据集中记录了所有地铁站点在每个时间间隔的流入和流出客流量,以 10 min 时间间隔为例,数据的具体形式如表 1 所列。表中 Inflow 和 Outflow 分别代表站点在一个时间间隔内的进站客流量和出站客流量。

表 1 北京地铁客流量数据

Table 1 Beijing subway passenger flow data

Station	Inflow/Person	Outflow/Person	Time
Xidan	22	3	2016-02-19 05:00-05:10
Xidan	41	11	2016-02-19 05:10-05:20

与基于文本的知识图谱相比,基于地铁客流量数据构建的地铁客流量时序知识图谱在构建过程中不涉及实体和关系的抽取工作。本文构建的地铁客流量时序知识图谱由地铁客流量四元组构成,基本形式为 (s, r, o, t) , 其中 s, r, o, t 分别代表头实体、关系、尾实体以及时间信息。

本文认为地铁站点是地铁客流量时序知识图谱中最重要的实体,并希望通过构建的地铁客流量时序知识图谱挖掘地铁站点的动态客流量模式。因此,地铁客流量时序知识图谱必须能够充分体现地铁站点和客流量的关系。图 1 给出了真实地铁网络中客流的传播方式,图中的圆圈代表地铁站点,箭头代表客流的流动方向,据此可以得到事实元组(站点 1, 流入, 站点 2)。但是,此事实元组无法体现具体的客流量,并且无法为其标注合适的时间信息。

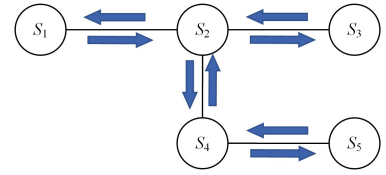


图 1 地铁客流示意图

Fig. 1 Schematic diagram of subway passenger flow

为了使地铁客流量时序知识图谱能够体现出站点的动态客流量模式,本文用客流量实体为站点实体创建连接,将地铁站点作为头实体,将流入客流量和流出客流量作为尾实体,对应的关系分别为流入(inflow)和流出(outflow)。此外,四元组中的时间信息包含年、月、日,以及时间间隔。由于地铁客流量预测工作通常基于一个月左右的客流数据,因此所有客流量的时间戳中年份相同。为了加强地铁客流量四元组中时间信息的相异度,本文仅保留月、日、时间间隔这 3 个时间元素。表 1 中完整的时间间隔表示方法过于复杂,不适用于知识图谱,因此本文按照时间间隔在一天中的顺序对其进行编号,在时间戳中使用时间间隔的编号代替原始的时间间隔。由于实际情况下地铁客流量会不断变化,而且不同站点之间的客流量存在较大差异,因此直接使用真实客流量不仅会导致知识图谱中尾实体数量过多,还会使其失去原本的构建意义,无法体现出站点的客流量模式。本文采用流量等级解决了此问题,用流量等级值代替真实流量值能够将客流量进行等级归类,便于体现站点的客流量模式。以 20 人为一个等级,地铁站点的客流量在 0~19 人之间对应的流量等级值为 0,客流量在 20~39 人之间对应的流量等级值为 1。最终构建的地铁客流量时序知识图谱如表 2 所列。

表 2 北京地铁客流量时序知识图谱

Table 2 TKG for passenger flow of Beijing subway

s	r	o	Real flow value	t
Xidan	inflow	1	22	2-19-1
Xidan	outflow	0	3	2-19-1
Xidan	inflow	2	41	2-19-2
Xidan	outflow	0	11	2-19-2

在本文构建的地铁客流量时序知识图谱中,客流量实体是地铁站点实体之间交互的桥梁,在相同时间拥有相同流入客流量或流出客流量的站点客流量模式的相似度更高。图2中蓝色圆圈代表地铁站点实体;橙色圆圈代表客流量实体;蓝色箭头代表流入关系,即当前时间间隔进站的客流量;橙色箭头代表流出关系,即当前时间间隔出站的客流量。图2中站点的物理连接情况和图1相同,由图2可知,虽然站点 S_1 和 S_5 的物理距离较远,但是本文构建的地铁客流量时序知识图谱依旧能够捕获二者的相关性。

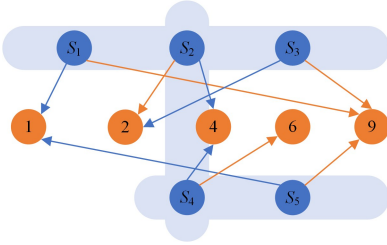


图2 地铁客流量时序知识图谱示意图(电子版为彩图)

Fig. 2 Schematic diagram of temporal knowledge graph of subway passenger flow

3.3 基于地铁客流量时序知识图谱嵌入构建动态相似矩阵

本文使用 DEEMB 模型^[23] 从上文构建的地铁客流量时序知识图谱中抽取地铁站点的动态嵌入,挖掘站点的动态客流量模式。为了使嵌入简单高效,本文选择使用 DE-DistMult^[23],即在 DistMult^[25]的基础上使用 DEEMB 函数进行改进,以完成地铁客流量时序知识图谱的嵌入任务。

DE-DistMult 首先通过 DistMult 模型得到实体和关系的静态嵌入,然后在静态嵌入中应用式(3)得到头实体和尾实体的动态嵌入 z_r 和 z_o ,最后通过式(4)计算融合时间信息后的动态三元组的得分函数。其中 z_r 是关系的静态嵌入, \circ 代表哈达玛乘积(Hadamard product)。DE-DistMult 在本文构建的地铁客流量时序知识图谱中得到的 z_r 是地铁站点实体的嵌入, z_o 是客流量实体的嵌入, z_r 和 z_o 中蕴含了地铁站点和客流量之间的相关性。以客流量实体为桥梁,可以构建地铁站点之间的相关性。

$$score = z_r \circ z_o \circ z_t \quad (4)$$

使用 DE-DistMult 完成地铁客流量时序知识图谱的嵌入后,为了进一步利用嵌入信息中蕴含的地铁站点动态客流量模式挖掘站点之间的相关性,本文首先将第 t 个时间间隔下所有地铁站点实体的嵌入按照站点的物理邻接顺序取出,构成第 t 个时间间隔的地铁站点嵌入矩阵 M' ,如式(5)所示:

$$M' = \begin{bmatrix} m_{11} & m_{12} & \cdots & m_{1d} \\ \vdots & \vdots & \ddots & \vdots \\ m_{|I|1} & m_{|I|2} & \cdots & m_{|I|d} \end{bmatrix} \in \mathbb{R}^{|I| \times d} \quad (5)$$

其中, $|I|$ 代表地铁网络中站点的数量, d 代表每个站点的嵌入维度。然后,使用式(6)计算站点之间的相似度:

$$similarity = \cos(m_i, m_j) \quad (6)$$

其中, m_i 和 m_j 分别代表第 i 个站点和第 j 个站点的嵌入向量,通过式(6)可以得到两个站点的相似度,将其应用到 M' 的每一行数据,即可得到由所有站点相似度构成的相似矩阵 S' 。如式(7)所示, s_{ij} 代表第 i 个站点和第 j 个站点的相似度,

所有时间间隔下的相似矩阵共同构成动态相似矩阵。为了便于理解动态相似矩阵的含义,可以将其与交通流量预测模型中的邻接矩阵做类比。邻接矩阵包含地铁网络的静态拓扑信息,能够为邻接的站点建立联系。动态相似矩阵刻画不同时间间隔下站点之间客流量模式的相似度,能够同时建立远近距离站点之间的动态相关性。

$$S' = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1|I|} \\ \vdots & \vdots & \ddots & \vdots \\ s_{|I|1} & s_{|I|2} & \cdots & s_{|I||I|} \end{bmatrix} \in \mathbb{R}^{|I| \times |I|} \quad (7)$$

3.4 引入动态相似矩阵的短期地铁客流量预测方法

随着深度学习的发展,学者们提出了一系列基于深度学习的地铁客流量预测框架^[18-20]。本文在已有的地铁客流量预测框架 ResLSTM^[16]中引入动态相似矩阵,在不使用邻接矩阵和天气数据等外部信息的前提下取得了更好的预测结果。

本文所提模型 TKG-ResLSTM 的整体架构如图3所示,主要由流入客流量、流出客流量,以及动态相似矩阵3个分支构成。

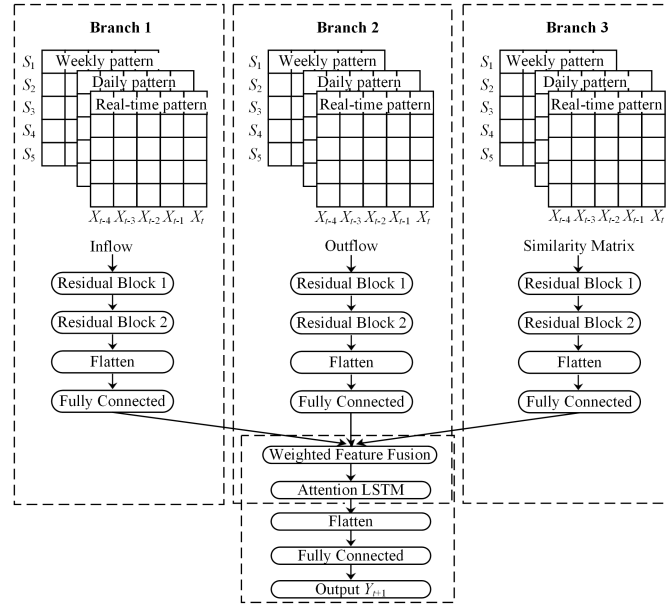


图3 TKG-ResLSTM 整体架构

Fig. 3 Overall architecture of TKG-ResLSTM

分支1使用地铁站点的流入客流量捕获时空特征,分支2使用地铁站点的流出客流量捕获时空特征,分支3使用动态相似矩阵捕获远近距离站点之间的动态相关性。其中前两个分支以及分支的特征融合部分与 ResLSTM 相同。第三个分支使用基于动态客流量模式的动态相似矩阵捕获远近距离站点之间的动态相关性,动态客流量模式指站点的进站客流量和出站客流量随时间变化的规律,蕴含在站点实体的嵌入中。由于不同时间间隔地铁网络的相似矩阵不同,且会随着时间的变化呈现明显的周期性,因此本文考虑动态相似矩阵的实时模式、天模式、周模式,三者分别对应3个时间序列 $S'_{t,r}$, $S'_{t,d}$, $S'_{t,w}$,分别代表当天第 t 个时间间隔地铁网络的相似矩阵以及前一天和前一周期相同时间间隔地铁网络的相似矩阵。本文通过动态相似矩阵调节站点的流入客流量,从而达到提高模型预测能力的目标。流入客流量也分为实时模式、天模式、

周模式,如式(2)所示。最终,分支3的输入如式(8)所示:

$$I_3 = (S_{l,r}^t X_{l,r}^t, S_{l,d}^t X_{l,d}^t, S_{l,w}^t X_{l,w}^t) \quad (8)$$

4 实验结果与分析

4.1 实验数据

本文在北京地铁数据集(公开数据集)^[16]和A市地铁数据集(某市地铁刷卡数据集)上对TKG-ResLSTM模型进行实验评估。北京地铁数据集中包含北京地铁2016年2月29日到2016年4月3日05:00-23:00的地铁客流

量数据,共17条线路,276个站点,25个工作日。A市地铁数据集中包含A市地铁2020年11月2日到2020年11月30日06:00-24:00的地铁客流量数据,共8条线路,168个站点,21个工作日。地铁数据集中每条客流量数据的关键信息是站点、流量种类(流入量或流出量)、流量值和时间戳。本文在实验中使用的时间间隔是10 min,15 min和30 min。

本文使用北京地铁和A市地铁客流量数据构建的地铁客流量时序知识图谱的统计信息如表3所列。

表3 地铁客流量时序知识图谱统计信息

Table 3 Statistics for TKG of subway passenger flow

Dataset	Number of entities	Number of relationship types	Number of quadruples in training set	Number of quadruples in validation set	Number of quadruples in test set
Beijing 10 min	466	2	119 232(75%)	19 872(12.5%)	19 872(12.5%)
Beijing 15 min	549	2	119 232(75%)	19 872(12.5%)	19 872(12.5%)
Beijing 30 min	761	2	119 232(75%)	19 872(12.5%)	19 872(12.5%)
City A 10min	281	2	72 576(75%)	12 096(12.5%)	12 096(12.5%)
City A 15 min	317	2	72 576(75%)	12 096(12.5%)	12 096(12.5%)
City A 30 min	431	2	72 576(75%)	12 096(12.5%)	12 096(12.5%)

4.2 模型配置和评价指标

本文使用TensorFlow和Keras在GPU中训练模型。以北京地铁数据集为例,使用前4周数据训练模型,最后1周数据测试模型,验证集用来校准模型,训练时使用前5个时间间隔的数据预测下一个时间间隔站点的进站客流量,使用3个指标评估模型性能,分别为均方根误差(Root Mean Square Error, RMSE)、平均绝对误差(Mean Absolute Error, MAE)和加权平均绝对百分比误差(Weighted Mean Absolute Percentage Error, WMAPE),计算式如式(9)~式(11)所示:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (9)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (10)$$

$$WMAPE = \frac{\sum_{i=1}^n \left(\frac{y_i}{\sum_{j=1}^n y_j} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \right)}{\sum_{j=1}^n y_j} \quad (11)$$

其中, y_i 代表真实值, \hat{y}_i 代表预测值。

本文使用DE-DistMult完成地铁客流量时序知识图谱的嵌入工作。我们为每一个事实元组 $f = (s, r, o, t) \in test$ 创建两个查询,分别为 $(s, r, ?, t)$ 和 $(?, r, o, t)$, $test$ 代表测试集。对于第一个查询,DE-DistMult将地铁客流量时序知识图谱中的所有实体进行排序,其中实体属于 $o \cup \bar{C}_{f,o}, \bar{C}_{f,o} = \{o' : o' \in \mathcal{V}, (s, r, o', t) \notin \mathcal{G}\}$ 。然后,对第二个查询做同样的操作。式(12)中的 $k_{f,o}$ 和 $k_{f,s}$ 分别代表两个查询中实体 o 和 s 的排名, $|test|$ 代表测试集中事实元组的数量。式(13)中,当 $cond$ 成立时 1_{cond} 为1,否则 1_{cond} 为0。我们使用平均倒数排名(Mean Reciprocal Rank, MRR)和Hit@n评估嵌入结果,计算式如式(12)、式(13)所示:

$$MRR = \frac{1}{2 \times |test|} \sum_{f=(s,r,o,t) \in test} \left(\frac{1}{k_{f,o}} + \frac{1}{k_{f,s}} \right) \quad (12)$$

$$Hit@n = \frac{1}{2 \times |test|} \sum_{f=(s,r,o,t) \in test} (1_{k_{f,o} \leq n} + 1_{k_{f,s} \leq n}) \quad (13)$$

4.3 地铁客流量时序知识图谱嵌入结果

本文分别构建北京地铁和A市地铁在10 min, 15 min和

30 min时间间隔下的地铁客流量时序知识图谱,嵌入结果如表4所列。由表4可知,时间间隔越短的地铁客流量时序知识图谱的嵌入结果越好。

表4 时序知识图谱嵌入结果

Table 4 Results of TKG embedding

Datasets/ Indicators	MRR	Hit@1/%	Hit@3/%	Hit@10/%
Beijing 10 min	0.4114	27.56	48.27	66.45
Beijing 15 min	0.3207	18.70	37.23	59.16
Beijing 30 min	0.2773	15.51	30.75	53.81
City A 10 min	0.4422	29.39	52.11	72.64
City A 15 min	0.4423	29.53	51.60	73.39
City A 30 min	0.3675	22.51	41.82	66.84

为了证明地铁客流量时序知识图谱嵌入模型能够捕捉地铁站点的动态客流量模式,建立站点之间的动态相关性,本文以A市地铁数据为例,对站点嵌入进行可视化分析。

考虑到不同站点在人流量较少的时间段内客流量相差不大,本文选择在A市地铁的早高峰和晚高峰时间段对时间间隔为10 min的站点的动态嵌入进行可视化分析。由原始数据推测出08:00-8:30属于A市地铁早高峰时间段,17:30-18:00属于A市地铁晚高峰时间段。图4(a)、图4(b)、图4(c)分别代表A市地铁2020年11月2日08:00-08:10、08:10-08:20、08:20-08:30所有站点的动态嵌入,图4(d)、图4(e)、图4(f)分别代表A市地铁2020年11月2日17:30-17:40、17:40-17:50、17:50-18:00所有站点的动态嵌入。图4中的一个点代表一个站点的嵌入,其中红色点和绿色点是同一条线路上的换乘站,地铁客流量模式相似,蓝色点是另一条线路上的起始站,与前两个站点的客流量模式差异较大。图4的可视化结果与理论分析结果相同,在早高峰和晚高峰期间的6个时间段内红色点和绿色点都非常接近,与此同时二者与蓝色点都相距很远,这说明基于本文构建的A市地铁客流量时序知识图谱生成的站点嵌入能够有效地捕捉站点的动态客流量模式。

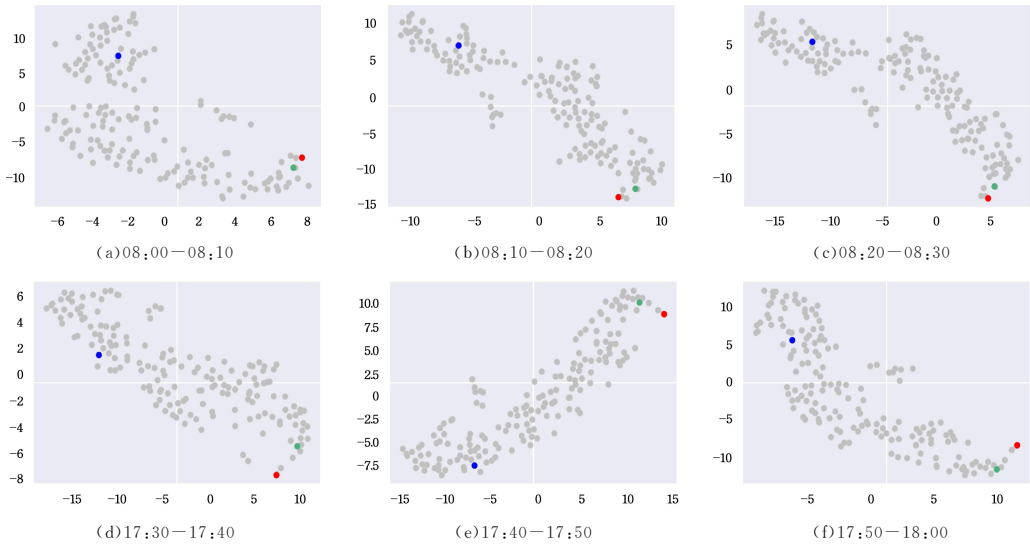


图4 A市地铁早晚高峰时段内的地铁站点嵌入(电子版为彩图)

Fig. 4 Subway station embedding in the morning and evening rush hours of subway in city A

为了能够更直观地说明站点间客流量模式的相似度,本文在图5中给出了站点B和与其客流量模式相似度最高的站点C的真实流量值,在图6中给出了站点B和与其客流量模式相似度最低的站点D的真实流量值。图中横坐标代表一周包含的所有时间间隔, $108 \times 5 - 5 = 535$, 纵坐标代表在一个时间间隔内站点的进站客流量。由图5和图6可知,本文构建的地铁客流量时序知识图谱能够有效地捕捉地铁站点的动态客流量模式。

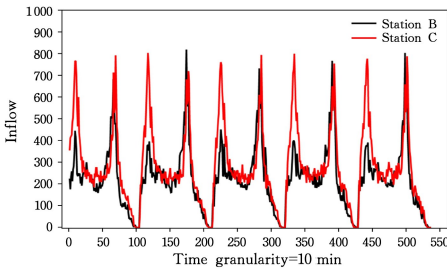


图5 站点B和站点C真实客流量对比图

Fig. 5 Comparison of real passenger flow between station B and station C

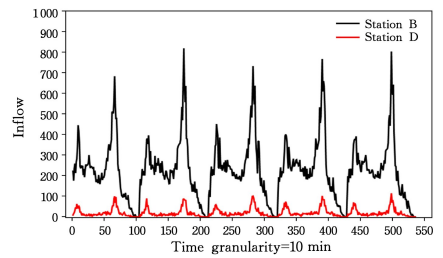


图6 站点B和站点D真实客流量对比图

Fig. 6 Comparison of real passenger flow between station B and station D

4.4 短期地铁客流量预测结果

短期地铁客流量的预测结果如表5所列,其中ARIMA基于数理统计,RBF-SVR和BPNN基于机器学习,其余模型基于深度学习。由表5可知,在大多数情况下,基于深度学习的模型明显优于基于数理统计和基于机器学习的模型。

对比北京地铁数据的ResLSTM-No W&A和TKG-ResLSTM可以发现,本文提出的动态相似矩阵比目前流量预测模型中广泛使用的邻接矩阵效果更好。

表5 地铁客流量预测结果

Table 5 Prediction results of subway passenger flow

Model	Beijing subway passenger flow datasets								
	10 min			15 min			30 min		
	RMSE	MAE	WMAPE/%	RMSE	MAE	WMAPE/%	RMSE	MAE	WMAPE/%
ARIMA	50.5436	27.3968	15.530	79.9580	42.2139	15.950	189.3329	100.3590	18.950
RBF-SVR	72.3753	61.2418	34.700	118.5395	100.2861	37.630	256.0734	206.0826	39.330
BPNN	45.6946	24.5322	14.240	73.3083	40.8457	15.450	163.9720	87.1259	16.810
Vanilla RNN	49.6506	26.2275	14.750	74.6021	41.4159	15.430	127.9259	71.8328	13.840
LSTM	37.1903	21.9925	12.710	53.9216	29.5340	11.290	96.3534	55.8265	10.760
GRU	44.9115	24.7353	14.330	56.6550	33.0309	12.560	86.6984	52.0717	10.030
CNN	29.8125	18.5460	10.413	40.2673	25.1231	9.375	64.0458	39.6867	7.472
ConvLSTM	28.7943	17.4780	9.814	37.0923	22.4236	8.380	61.4978	36.9768	6.962
ResLSTM-GCN	30.0706	17.4331	9.775	36.7801	21.6615	8.099	59.7289	33.6304	6.334
TKG-ResLSTM-S	33.0726	19.3005	10.856	42.9536	24.5391	9.165	68.3067	39.7635	7.487
ResLSTM-No W&A	29.6219	17.6344	9.911	38.7386	23.0008	8.589	60.1340	34.1360	6.428
ResLSTM	<u>28.3661</u>	<u>16.6318</u>	<u>9.352</u>	<u>36.0444</u>	<u>20.8783</u>	<u>7.805</u>	<u>56.9649</u>	<u>32.5819</u>	<u>6.134</u>
TKG-ResLSTM	27.9609	16.4257	9.238	35.7266	20.4850	7.660	<u>57.5938</u>	32.3309	6.090

(续表)

Model	City A subway passenger flow datasets								
	10 min			15 min			30 min		
	RMSE	MAE	WMAPE/%	RMSE	MAE	WMAPE/%	RMSE	MAE	WMAPE/%
ResLSTM-GCN	20.9699	12.1141	13.760	<u>26.8179</u>	15.2489	11.508	45.4617	24.2477	9.466
TKG-ResLSTM-S	22.7666	12.5296	14.242	31.5256	17.3283	13.199	50.9626	26.4578	10.297
ResLSTM-No W&A	<u>20.9656</u>	<u>11.6319</u>	13.257	26.5445	<u>14.8427</u>	<u>11.230</u>	44.1183	22.8548	8.916
TKG-ResLSTM	19.9086	11.1759	12.749	27.1942	14.7396	11.179	<u>45.0700</u>	<u>23.6160</u>	9.188

注:ResLSTM-GCN表示只有分支3(邻接矩阵分支)的ResLSTM;TKG-ResLSTM-S表示只有分支3(相似度分支)的TKG-ResLSTM;ResLSTM-No W&A表示去除分支4(天气数据分支)的ResLSTM;粗体代表最优结果,下划线代表次优结果。

对比北京地铁数据的ResLSTM和TKG-ResLSTM可以发现,当时间间隔为10 min和15 min时,TKG-ResLSTM模型在3个指标下均优于添加了天气和空气质量信息的ResLSTM模型。在未使用外部信息和邻接矩阵的情况下,TKG-ResLSTM模型取得了更好的预测结果,说明本文提出的动态相似矩阵中蕴含地铁站点间的近距离相关性和远距离相关性。

对比北京地铁数据和A市地铁数据的TKG-ResLSTM-S和TKG-ResLSTM以及ResLSTM-GCN和ResLSTM可以发现,地铁站点的流入和流出客流量对动态相似矩阵的影响比对邻接矩阵的影响更大,这说明本文通过地铁客流量时序知识图谱嵌入捕获的地铁站点动态客流量模式与原始的流入流出客流量相匹配,三者相辅相成,共同提高模型的预测效果。

4.5 地铁客流量时序知识图谱中的流量等级对TKG-ResLSTM预测结果的影响

由表6可知,构建地铁客流量时序知识图谱时选择的流量等级会直接影响TKG-ResLSTM的预测结果。地铁客流量时序知识图谱中的实体包括站点实体和客流量实体,站点实体的数量随地铁网络固定,客流量实体的数量受流量等级影响。当Class=1时,地铁客流量时序知识图谱中的客流量实体数量过多,嵌入过程中很难捕获站点的客流量模式,导致TKG-ResLSTM的预测结果最差。当Class=1000时,地铁客流量时序知识图谱中的客流量实体数量过少,造成地铁站点客流量模式信息的丢失,预测结果也随之降低。不同数据集的最佳流量等级不同,需要根据实验结果确定。

表6 流量等级对预测结果的影响,以北京地铁为例

Table 6 Impact of traffic class on prediction results, taking Beijing subway as an example

Beijing 15min				
Class\ Indicators	Number of entities	RMSE	MAE	WMAPE/%
1	4048	38.6796	22.5123	8.410
20	549	35.7266	20.4850	7.660
40	422	36.4540	20.8894	7.803
60	377	36.0087	20.5210	7.671
80	355	36.2602	20.9121	7.820
1000	283	36.9822	21.3716	7.989

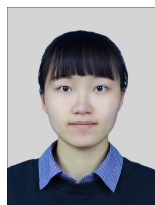
结束语 本文提出了一种基于时序知识图谱嵌入的短期地铁客流量预测方法(TKG-ResLSTM),该方法利用时序知识图谱嵌入技术获取地铁站点的动态客流量模式,构建动态相似矩阵,在不使用邻接矩阵和外部信息的情况下提高了地铁站点客流量的预测效果。实验结果表明,本文提出的TKG-ResLSTM模型在北京地铁公开数据集上的性能优于

相关工作,此外,在本文构建的A市地铁数据集上也取得了较为优异的结果。TKG-ResLSTM具有良好的可扩展性,其基于地铁客流量时序知识图谱抽取站点客流量模式的方法可以方便地扩展到其他方法中。在后续工作中,我们将尝试对地铁客流量时序知识图谱嵌入模型和客流量预测模型进行联合训练,以进一步提高预测准确率,提升模型的稳定性。

参考文献

- [1] LIU J, GUAN W. A summary of traffic flow forecasting methods[J]. Journal of Highway and Transportation Research and Development, 2004, 21(3): 82-85.
- [2] WANG P, LIU Y. Network traffic prediction based on improved BP wavelet neural network[C]// 2008 4th International Conference on Wireless Communications, Networking and Mobile Computing. IEEE, 2008: 1-5.
- [3] LENG B, ZENG J, XIONG Z, et al. Probability tree based passenger flow prediction and its application to the Beijing subway system[J]. Frontiers of Computer Science, 2013, 7(2): 195-203.
- [4] SU H, YU S. Hybrid GA based online support vector machine model for short-term traffic flow forecasting[C]// International Workshop on Advanced Parallel Processing Technologies. Springer, 2007: 743-752.
- [5] ROOS J, BONNEVAY S, GAVIN G. Short-term urban rail passenger flow forecasting: A dynamic bayesian network approach [C]// 2016 15th IEEE International Conference on Machine Learning and Applications(ICMLA). IEEE, 2016: 1034-1039.
- [6] SUN Y, LENG B, GUAN W. A novel wavelet-SVM short-time passenger flow prediction in Beijing subway system[J]. Neurocomputing, 2015, 166: 109-121.
- [7] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521(7553): 436-444.
- [8] ZHANG J, CHEN F, GUO Y, et al. Multi-graph convolutional network for short-term passenger flow forecasting in urban rail transit[J]. IET Intelligent Transport Systems, 2020, 14(10): 1210-1217.
- [9] YANG D, CHEN K, YANG M, et al. Urban rail transit passenger flow forecast based on LSTM with enhanced long-term features[J]. IET Intelligent Transport Systems, 2019, 13(10): 1475-1482.
- [10] TANG T T, ZHOU W. Research on commodity sales forecast oriented on deep learning[J]. Journal of Chongqing University of Technology(Natural Science), 2022, 36(7): 310-317.
- [11] ZHANG J, ZHENG Y, QI D. Deep spatio-temporal residual networks for citywide crowd flows prediction[C]// Proceedings of

- the AAAI Conference on Artificial Intelligence, 2017; 1655-1661.
- [12] YAO H, WU F, KE J, et al. Deep multi-view spatial-temporal network for taxi demand prediction[C] // Proceedings of the AAAI Conference on Artificial Intelligence, 2018; 2588-2595.
- [13] HAN Y, WANG S, REN Y, et al. Predicting station-level short-term passenger flow in a citywide metro network using spatio-temporal graph convolutional neural networks[J]. ISPRS International Journal of Geo-Information, 2019, 8(6): 243.
- [14] YU B, YIN H, ZHU Z. Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting[J]. arXiv:1709.04875, 2017.
- [15] TANG L, ZHAO Y, CABRERA J, et al. Forecasting short-term passenger flow: An empirical study on shenzhen metro[J]. IEEE Transactions on Intelligent Transportation Systems, 2018, 20(10): 3613-3622.
- [16] ZHANG J, CHEN F, CUI Z, et al. Deep learning architecture for short-term passenger flow forecasting in urban rail transit[J]. IEEE Transactions on Intelligent Transportation Systems, 2020, 22(11): 7004-7014.
- [17] JI T, HE Y, ZHU Y Y, et al. Power user clustering method based on a combination of knowledge graph and modified Gaussian mixture model[J]. Journal of Chongqing University of Technology(Natural Science), 2022, 36(12): 92-101.
- [18] LIU Y, LIU Z, JIA R. DeepPF: A deep learning based architecture for metro passenger flow prediction[J]. Transportation Research Part C: Emerging Technologies, 2019, 101: 18-34.
- [19] HAO S, LEE D H, ZHAO D. Sequence to sequence learning with attention mechanism for short-term passenger flow prediction in large-scale metro system[J]. Transportation Research Part C: Emerging Technologies, 2019, 107: 287-300.
- [20] YE J, ZHAO J, YE K, et al. Multi-STGCnet: A graph convolution based spatial-temporal framework for subway passenger flow forecasting[C] // 2020 International Joint Conference on Neural Networks(IJCNN). IEEE, 2020: 1-8.
- [21] LIN Y, HAN X, XIE R, et al. Knowledge representation learning: A quantitative review[J]. arXiv:1812.10901, 2018.
- [22] GARCÍA-DURÁN A, DUMANČIĆ S, NIEPERT M. Learning sequence encoders for temporal knowledge graph completion[J]. arXiv:1809.03202, 2018.
- [23] GOEL R, KAZEMI S M, BRUBAKER M, et al. Diachronic embedding for temporal knowledge graph completion[C] // Proceedings of the AAAI Conference on Artificial Intelligence, 2020: 3988-3995.
- [24] JIANG T, LIU T, GE T, et al. Towards time-aware knowledge graph completion[C] // Proceedings of the 26th International Conference on Computational Linguistics: Technical Papers, 2016: 1715-1724.
- [25] YANG B, YIH W, HE X, et al. Embedding entities and relations for learning and inference in knowledge bases[J]. arXiv:1412.6575, 2014.



MAO Huihui, born in 1998, postgraduate. Her main research interests include knowledge graph and traffic flow prediction.



LI Tianrui, born in 1969, Ph.D, professor, Ph.D supervisor, is a distinguished member of China Computer Federation. His main research interests include big data intelligence, rough sets and granular computing.

(责任编辑:李亚辉)