

基于编解码深度神经网络的交通流预测模型

李山海¹ 吴艳雄¹ 刘玉龙² 张子晔²

(1. 中华全国工商业联合会信息中心 北京市 100035 2. 中国电子科技集团公司第十五研究所 北京市 100083)

摘要: 本文提出了一种基于编码器-解码器结构的深度神经网络模型。在经典的编解码模型基础上, 基于长短时记忆网络模型并扩展时间注意力学习机制, 以自动学习交通流相关序列数据中的隐含表示和非线性相关深度特征, 实现端到端的变长交通流序列预测, 并扩展注意力机制以提升模型的预测性能。基于真实的交通流数据集进行了对比实验分析, 实验结果表明所提出的交通流预测模型相比基准模型具有更优的预测准确率。

关键词: 交通流预测; 编码器-解码器; 时间注意力机制; 深度; 神经网络

1 引言

当前, 新一代人工智能技术在支撑国民经济上的作用更深远、范围更广泛, 应用场景也几乎涵盖智慧城市管理和智能交通各个领域。交通是国民经济发展的主动脉, 当前智能交通产业发展更是各级政府和民营经济的关注热点。随着智慧城市和物联网技术的快速发展, 以及全球定位系统 (GPS) 和地理信息 (GIS) 系统等的普及应用, 交通流数据的采集越来越方便, 交通大数据呈爆炸式增长, 而且交通流相关数据还呈现出复杂的非线性相关特点^[1], 如交通流量、通行速度和通行时间、天气情况等交通流相关多变量数据等, 存在复杂的相互影响。如何有效地分析和建模这些交通流相关多变量序列数据集, 并构建有效的交通流预测模型已经成为智能交通研究的热点和重要发展方向^[2]。研究学者针对这一关键问题提出了很多的研究模型与方法, 包括统计方法和机器学习方法等, 但由于交通序列数据具有的高维性、非线性和相互关联性问题, 如何有效地构建更准确的交通流预测模型仍然是智能交通面临的核心问题, 而深度学习技术的兴起使其成为一种重要的应用参考。

深度学习由于其多层非线性映射原理的深度特征抽取学习能力, 如在图像处理、语音识别和自然语言处理等多个领域取得突破性进展^[3]。由于不同交通时序数据之间通常存在复杂的非线性相关关系, 传统方法模型难以挖掘与交通流量相关的各类序列数据之间的深层关系, 而深度学习对于交通流预测任务来讲是一个很好的选择。

2 相关工作

传统的交通流预测模型可以分为两大类, 一类为参数方法, 另一类为非参数方法。这些传统的方法通常基于浅层模型或核函数进行交通序列数据建模^{[4][5]}, 难以捕获交通流相关序列数据中的相互关联特征。

参数法主要基于经典的时间序列分析技术, 时间序列数据反映了某个 (单变量时间序列) 或者某些随机变量 (多变量时间序列) 随时间不断变化的趋势, 时间序列预测分析就是要获得这种趋势规律以预测未来的序列值, 基于参数方法的交通流预测模型包括自回归滑动平均 (ARMA) 模型, 差分自回归滑动平均 (ARIMA) 模型^[4] 等等。非参数方法是指模型结构及参数需要在训练过程中确定, 也可以用于处理非线性相关的交通流序列数据, 主流的机器学习模型属于此类方法, 比如基于支持向量回归机 (SVR)^[5] 和人工神经网络 (ANN)^[6] 等的交通流预测模型。通过上述两大类传统交通流预测方法分析表明, 传统的交通流预测模型假定交通序列数据中只包含线性自相关结构, 但在真实的交通流预测应用中往往要考虑多种影响因素, 而且这些因素是非线性相关的, 要求模型能够处理非线

性相关交通流数据, 在这方面深度学习模型具有较大的优势, 可以有效获得交通流大数据中的非线性相关特征。比如 Chen 等提出了一种基于深度卷积神经网络的短时交通拥堵预测方法 (PCNN), 有效捕捉了近邻时间段交通拥堵的相似模式与拥堵等级的多尺度性^[7]。Lv 等基于 RNN 和 CNN 网络提出了一种交通速度预测模型 (LC-RNN), 可挖掘周围区域的交通动态变化规律, 实现更准确的交通速度预测^[8]。Gong 等将人流分布信息与历史数据相结合, 提出了三种在线 NMF 模型用于悉尼铁路网的人流预测^[9]。

3 模型设计

3.1 编码器-解码器

编码器-解码器模型 (Encoder-decoder) 是由 Cho 于 2014 年提出^[10]。该模型最早应用于自然语言处理, 后来被引入时间序列预测领域。

编解码器结构包括编码器 (encoder) 和解码器 (decoder)。其中编码器对输入数据特征进行编码, 解码器对应预测输出目标。Encoder-decoder 用于交通流预测计算过程如下:

在编码阶段读入交通流数据序列 $X=(x_1, x_2, \dots, x_{T_x})$ 。通过编码器中的神经网络层计算后, 得到隐层状态 h 由公式 (1) 所示。

$$h_t=f(x_t, h_{t-1}) \quad (1)$$

编码后获得了各个时间段的隐藏层, 将隐藏层的信息汇总, 生成编码好的隐含表示向量 C :

$$C=q(h_1, h_2, \dots, h_{T_x}) \quad (2)$$

经典的 Encoder-decoder 模型通常选取最后一个时刻的隐层作为解码器的输入数据, 即公式 (3)。

$$C=q(h_1, h_2, \dots, h_{T_x})=h_{T_x} \quad (3)$$

上述公式中 $h_t \in \mathbb{R}^n$ 是在时间步 t 的隐藏状态, q 是编码采用的神经网络层, 如采用 LSTM 神经网络可以表示为 $q(\{h_1, h_2, \dots, h_T\})=h_T$ 。

$$y_t = \operatorname{argmax}_P(y) = \prod_{t=1}^T p(y_t | y_1, y_2, \dots, y_{t-1}, C) \quad (4)$$

Decoder 负责对输入数据进行解码, 利用向量 C 和所用先前的预测输出 $\{y_1, y_2, \dots, y_{t-1}\}$ 计算预测输出 y_t , 一般通过联合条件概率计算预测输出, 如公式 (4) 所示。

3.2 模型结构设计

经典的神经网络模型输入序列和输出序列一般是等长的, 但在交通流预测应用中, 模型输入序列和输出序列并不一致, 这就需要设计新的模型解决这一问题。由于 Encoder-decoder 模型中输入序列和输出序列可以不等长, 通过参数设置, 可以适应不同的交通流

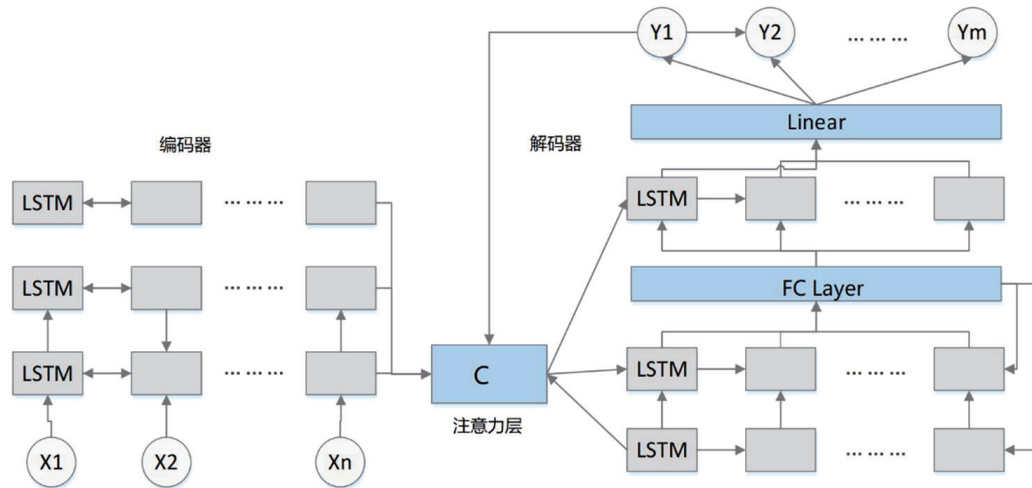


图 1: 交通流预测模型结构设计图

序列预测问题。同时，本模型设计中扩展了 attention 机制，可以进一步提高模型预测的准确率。模型结构设计如图 1。

整个模型结构设计分为三层，编码器层、注意力层和解码器层，编解码器中都采用多层 LSTM 神经网络进行构造。首先交通流数据集通过数据归一化得到输入向量 $\{X_1, X_2, \dots, X_n\}$ ，编码器对交通流输入进行编码，得到编码向量 c ，再流入解码器进行解码。在这个过程中，模型扩展设计了 attention 机制，在编码器每一步的解码中，都有预测输出反馈回解码器输入，以对输入交通流所处理的 LSTM 编码层进行权值更新，实现交通流预测输出和输入序列的某些变量相关，以提升模型的预测准确率。

3.3 注意力机制

注意力机制设计不要求编码器将所有输入交通流序列都编码为一个固定长度的向量，而是在解码的时候，每一步都会选择性的从向量序列中挑选一个子集进行处理。这样，在做每一步交通流预测输出的时候，可以尽量利用输入序列信息特征，下面给出注意力机制核心计算过程。交通流预测输出的条件概率计算如下公式所示：

$$p(y_i | y_1, y_2, \dots, y_{i-1}, X) = g(y_{i-1}, s_i, c_i) \quad (5)$$

上式 s_i 表示解码器 i 时刻的隐藏状态：

$$s_i = f(s_{i-1}, y_{i-1}, c_i) \quad (6)$$

条件概率与每个目标预测输出 y_i 相对应的编码向量 c_i 有关。其中是 c_i 由编码时的隐藏向量序列 (h_1, h_2, \dots, h_n) 按权重相乘并求和，如下所示：

$$c_i = \sum_{j=1}^{T_x} a_{ij} h_j \quad (7)$$

上述计算过程由于 h_i 中包含了交通流输入序列中第 i 个时间步的信息。将隐藏向量序列按权重相加，表示在生成第 j 个时间步预测输出时的注意力权重是不同的。 a_{ij} 的值越高，表示第 i 个时间步输出在第 j 个时间步输入上相关性越强。 a_{ij} 计算公式如下：

$$a_{ij} = \frac{\exp(e_{ij})}{\sum_{k=1}^{T_x} \exp(e_{ik})} \quad (8)$$

$$e_{ij} = a(s_{i-1}, h_j) \quad (9)$$

公式 (9) 中 s_{i-1} 先跟每个 h 分别计算得到一个分数，然后使用 softmax 得到 i 时刻的输出在 T_x 个隐藏状态中的注意力分配向量，从而得到计算 c_i 的权重。

4 实验对比与分析

为对所提出模型的有效性进行验证，基于 Keras 深度学习框架

设计实现了本文提出模型，采用 Scikit-learn 机器学习库用于构建基准对比相关浅层学习模型。服务器硬件配置为 12G 显存的 NVIDIA GeForce RTX 2080Ti GPU, Intel(R) Core(TM) i9-7920X CPU, 内存 64G。

并将 RMSE 和 MAE 作为模型误差分析指标，用于评估各种方法的交通流预测性能，误差指标计算公式如下：

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (10)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (11)$$

其中 \hat{y}_i 代表预测的交通流量值， y_i 代表真实的交通流量值， n 是测试数据集样本数量。

4.1 实验数据集

实验数据集来源于英国政府开发数据平台，简称为 Highway England Dataset^[11]，该平台发布有关英格兰的两类主要公路网的交通流数据信息。Highway England 交通流数据集是一个典型的多变量时间序列，该数据集采集了每间隔 15 分钟的平均通行时间，通行速度和交通流量等信息，实验所用的数据集时间跨度为 2013 年 1 月 1 日至 2014 年 2 月 31 日。

4.2 基准对比模型及参数设置

本文所提出模型与以下基准方法进行比较：

(1) VAR，向量自回归模型，可以对多变量时间序列数据之间的隐含关系进行建模。

(2) ARIMA，自回归综合移动平均值法是一种广泛用于时间序列分析的模型，它结合了移动平均值和自回归方法。

(3) SVR，支持向量回归 (SVR) 是支持向量机器学习模型的一种变体方法，经常用于时间序列预测。

(4) RNN，这是用于处理序列任务的最为传统的深度学习方法，LSTM 和 GRU (门控循环单位)，是两种最流行的基于 RNN 变体的深度学习模型。

(5) Seq2Seq，一种可用于交通流预测的经典序列到序列深度学习模型，该模型选择 LSTM 作为编码器和解码器组件。

本文对比所用的神经网络基准模型，其权重初始化均使用 Keras 框架默认值，使用 Adam 作为模型优化器，学习率设定为 0.001，使用的 batch size 设为 64。编码器 - 解码器的 LSTM 神经网络层设置为 3 层，在每两层 LSTM 间设置一层全连接层，每个隐藏层神经

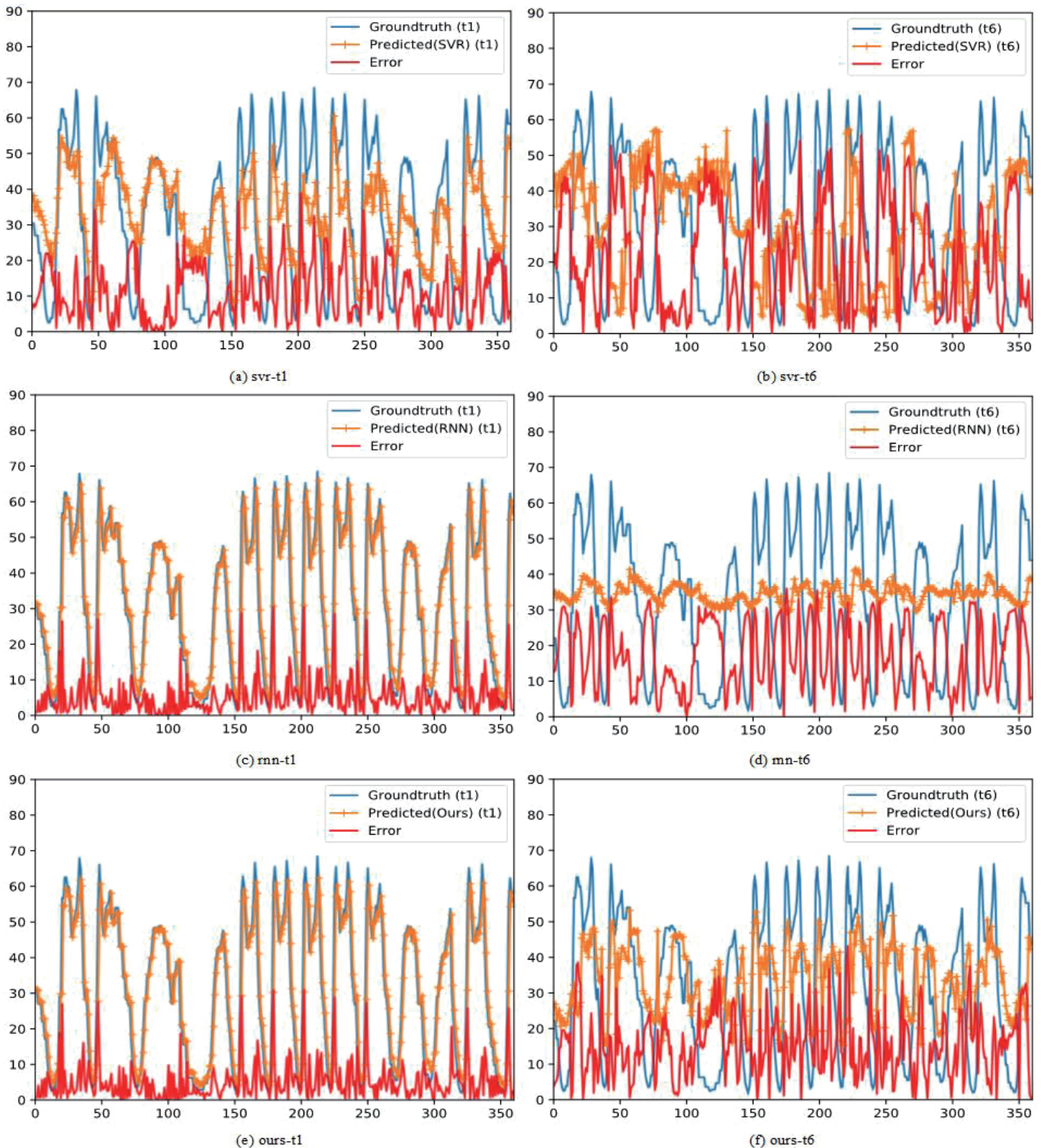


图 2: 不同模型 (SVR, RNN, 本文所提出模型) 在两周内的交通流预测值与真实值曲线对比图

元个数为 128。

4.3 实验结果对比分析

不同模型的交通流预测误差指标对比如表 1 所示, 给出了本文所提出模型 (Ours) 与基准模型的多步 (从 $t+1$ 到 $t+12$ 的未来 12 小时内) 交通流预测误差 (RMSE 和 MAE) 结果。从表中数据可以看到, 无论预测时间步长是多少, 本文模型的预测性能都比其他基准模型要好, 即使对于未来 12 个小时后的长时间步预测, 所提

模型预测性能也能保持最优, 预测误差值保持最低。

SVR 模型, RNN 模型与本文所提出模型在两周内的交通流预测值与真实值曲线对比图如图 2, 图中的误差为预测值与真实值相减后的绝对误差。图 2(a) 为 SVR 单步预测; 图 2(b) 为 SVR 未来第 6 个时间步预测; 图 2(c) 为 RNN 单步预测; 图 2(d) 为 RNN 未来第 6 个时间步预测; 图 2(e) 为本文模型的单步预测; 图 2(f) 为本文模型在未来第 6 个时间步预测。其中水平轴 x 表示观测的时间步

表 1: 本文模型与基准模型的交通流预测性能对比

Models	Model Error							
	RMSE				MAE			
	t+1	t+3	t+6	t+12	t+1	t+3	t+6	t+12
VAR	7.12	16.54	19.21	20.44	5.23	13.21	17.11	17.89
ARIMA	7.24	17.32	20.24	21.52	5.21	13.46	18.24	17.56
SVR-LINEAR	7.25	18.24	22.08	22.35	5.22	14.29	19.20	18.28
RNN	7.20	15.33	18.24	17.44	5.15	12.54	15.89	14.43
GRU	7.21	15.39	18.67	17.40	5.19	12.63	16.42	14.42
LSTM	7.22	15.42	19.49	17.85	5.19	12.57	16.90	14.56
SEQ2SEQ	7.20	14.43	16.82	17.19	5.01	11.55	14.21	13.97
Ours	7.14	13.22	15.30	16.24	4.87	11.04	12.15	12.82

长, 每个间隔为 1 小时, 垂直轴 y 代表交通流量, 分别比较了不同模型在单步预测 (t+1) 和多步预测条件下 (t+6) 的交通流预测误差情况。从图中我们可以看到, 本文所提模型在不同时间步长情况下都能保持最优的预测性能, 预测值曲线与真实值曲线能够很好的匹配。模型预测性能不管是在平时工作日情况还是周末情况下, 都要优于以 SVR 为代表的浅层模型和 RNN 为代表的所有基准对比模型, 从图 2 中的交通流波峰和波谷时间段范围内, 各模型的预测对比情况更为明显。

综上, 通过在真实交通流数据集上进行的实验对比分析, 本文所提出模型相比基准方法具有更好的预测性能, RMSE 与 MAE 预测误差最小, 不管是在单步预测还是在长时多步预测条件下, 本模型的交通流预测值都可以与真实值保持较好的匹配。

5 结论

新一代人工智能技术在支撑国民经济上的作用越来越凸显, 特别是以深度学习为代表的新兴技术, 在智能交通领域的应用方兴未艾。交通作为国民经济发展的主动脉, 使得智能交通产业发展备受各级政府和民营经济的关注, 比如阿里巴巴的城市大脑平台, 智能交通就是其应用的核心领域。而交通流预测是一个经典的智能交通研究课题, 随着交通大数据时代的到来, 传统的统计建模方法面临技术瓶颈。本文引入当前主流的人工智能技术, 从经典的编码器-解码器深度神经网络模型入手, 设计了一种新的基于编码器-解码器结构和融合注意力机制的深度神经网络模型。通过编解码过程中长短时记忆网络模型的应用并扩展时间注意力学习机制, 能自动学习交通流相关序列数据中的隐含表示和非线性相关深度特征, 实现端到端的变长交通流序列预测。最后在真实的交通流数据集上与多个基准模型进行对比实验, 验证了所提出模型的有效性和准确性, 为智能交通应用提供了一定参考。在未来工作中, 将进一步研究该模型的结构设计优化, 并结合参数设置搜索, 以提升模型的预测性能。

参考文献

- [1]Lippi M, Bertini M, Frasconi P, "Short-term traffic flow forecasting: An experimental comparison of time-series analysis and supervised learning," IEEE Trans. Intelligent Transportation Systems, vol. 14, no. 2, pp. 871-882, 2013.
- [2]樊娜, 赵祥模, 戴明, 安毅生. 短时交通流预测模型 [J]. 交通运输工程学报, 2012, 12 (04): 114-119.
- [3]Schmidhuber J, "Deep learning in neural networks: an overview," Neural Networks, vol. 61, pp. 85-117, 2015.
- [4]韩超, 宋苏, 王成红. 基于 ARIMA 模型的短时交通流实时自适应预测 [J]. 系统仿真学报, 2004 (07): 1530-1535.

- [5]Castro-Neto M, Jeong Y S, Jeong M K, et al, "Online-SVR for short-term traffic flow prediction under typical and atypical traffic conditions," Expert Systems with Applications, vol. 36, no. 3, pp. 6164-6173, 2009.
- [6]Chan K Y, Dillon T S, Singh J, et al, "Neural-network-based models for short-term traffic flow forecasting using a hybrid exponential smoothing and Levenberg-Marquardt algorithm," IEEE Trans. Intelligent Transportation Systems, vol. 13, no. 2, pp. 644-654, 2012.
- [7]Chen M, Yu X, Liu Y. PCNN: Deep Convolutional Networks for Short-Term Traffic Congestion Prediction[J]. IEEE Transactions on Intelligent Transportation System, 2018, 19 (11): 3550-3559.
- [8]Lv Z, Xu J, Zheng K, et al. LC-RNN: A Deep Learning Model for Traffic Speed Prediction[C]. Proceedings of the 27th International Joint Conference on Artificial Intelligence, 2018: 3470-3476.
- [9]Gong Y, Li Z, Zhang J, et al. Network-wide Crowd Flow Prediction of Sydney Trains via Customized Online Non-negative Matrix Factorization[C]. Proceedings of the 27th ACM International Conference on Information and Knowledge Management, 2018: 1243-1252.
- [10]Cho K, Van Merriënboer B, Gulcehre C, et al, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," arXiv preprint arXiv:1406.1078, 2014.
- [11][Online]. Available: <https://data.gov.uk/dataset/dft-eng-srn-routes-journey-times/>

作者简介

李山海 (1963-), 男, 学士学位, 现为中华全国工商业联合会信息中心主任、助理研究员。研究方向为信息系统顶层设计。
吴艳雄 (1975-), 男, 学士学位, 现为中华全国工商业联合会信息中心技术处处长、工程师。研究方向为信息网络与信息系统。
刘玉龙 (1981-), 男, 硕士学位, 现为中国电子科技集团公司第十五研究所系统八部主任、研究员级高工。研究方向为信息系统顶层设计、数据资产管理与共享技术、数据挖掘与分析预测技术等。
张子晔 (1995-), 女, 硕士学位, 现为中国电子科技集团公司第十五研究所助理工程师。研究方向为数据信息系统建设、数据资产管理与共享技术。