DOI:10.20033/j.1003-7241.(2025)02-0021-05.

基于改进时空因果卷积网络的PM2.5浓度预测

王媛媛[1,2]，沈 俞[1]，陈秀川[1]，王 超[1]，张兴潮[1]

（1.淮阴工学院 计算机与软件工程学院，江苏 淮安 223001；
2.江苏省物联网移动互联技术工程实验室，江苏 淮安 223001）

摘 要：PM2.5的浓度受到多方面的影响，包括天气、空气中的其他物质成分以及周围地点的空气质量等。为提高深度学习模型对PM2.5浓度预测的准确率与泛化性能，提出了一种改进的时空因果卷积网络(STCN)，使模型在预测时考虑到不同地点对预测地PM2.5浓度的不同影响。以北京市多个地点的空气质量数据进行实验并与原始的STCN网络进行对比，实验表明，改进后的时空因果卷积网络对PM2.5浓度预测的能力高于基线STCN。

关键词：空气质量预测；PM2.5；时空因果卷积网络；空间关联因素
中图分类号：TP18 文献标识码：A 文章编号：1003−7241(2025)02−0021−05

PM2.5 Concentration Prediction Based On Improved Spatiotemporal Causal Convolution Network

WANG Yuan-yuan[1,2], SHEN Yu[1], CHEN Xiu-chuan[1], WANG Chao[1], ZHANG Xing-chao[1]

(1. School of Computer and Software Engineering, Huaian 223001 China;
2. Jiangsu Internet of Things Mobile Interconnection Technology Engineering Laboratory, Huaian 223001 China)

Abstract: The concentration of PM2.5 is affected in many ways, including weather, other air components and the air quality of surrounding places. In order to improve the accuracy and generalization performance of the depth learning model for predicting PM2.5 concentration, an improved spatiotemporal causal convolution network (STCN) is proposed, which allows the model to take into account the different effects of different locations on PM2.5 concentration in the prediction site. The air quality data from multiple locations in Beijing are used for the experiment and compared with the original STCN network. The experiment shows that the improved spatiotemporal causal convolution network has a higher ability to predict PM2.5 concentration than the baseline STCN.

Keywords: air quality prediction; PM2.5; spatiotemporal causal convolution network; spatial correlation factors

0 引言

PM2.5是指空气动力学直径小于或者等于2.5 μm的一种固体颗粒，通过人体呼吸道进入体内，损害肺泡、食道等器官或组织，对人体有严重危害[1]。PM2.5有致癌风险，国际癌症研究中心将PM2.5等空气污染物质列为重点环境致癌物[2]。为缓解空气质量问题对人体和社会的危害，精确预测PM2.5浓度对人们的生理健康和日常出行至关重要[3]。

PM2.5的浓度预测方式主要分为以下三种方式，大气数值模式法、机器学习法以及深度学习法[4−6]。大气数值法是指利用空气中的污染物浓度值以及各类气象数据，基于空气动力学和物理化学等理论，建立模拟污染物形成的数值模型。主要包括多尺度空气质量模型[7]，嵌套空气质量预报模式系统[8]等。机器学习模型是指利用如随机森林(RF)、自回归模型和支持向量机(SVM)等模型对PM2.5相关数据提取特征，并预测未来时间段的浓度值。例如：康俊锋[5]等人对多个机器学习模型进行对比分析，在众多模型中XGBoost模型相对精度更高，效率也较高；贾佳美[9]等人利用改进粒子群算法来提升BP网络的PM2.5预测性能。由于机器学习的模型较小，其特征提取能力有限，因此基于深度学习的PM2.5预测方法成为目前主要的预测方法。其中，序列模型中的循环神经网络(RNN)应用较广。但RNN存在训练困难、不容易收敛等缺点[10]，因此，长短时记忆网络(LSTM)、门控制循环单元(GRU)等被用于PM2.5预测，白盛楠等人[11]应用LSTM结合灰色关联度分析进行PM2.5预测。由深度学习方法衍生出了许多混合模型用来做预测，冀林[12]应用CNN与GRU的混合模型实现预测；杨亚莉等人[13]在RNN神经网络中加入时序注意力和类别注意力构建二向注意力循环

*基金项目：江苏省自然科学基金面上项目(BK20211365)；淮阴工学院—淮安经济技术开发区产学研合作项目(Z413H21522)；大学生创新创业训练计划项目(202211049087Y；202211049268XJ)
收稿日期：2023-08-28

神经网络(TDA−RNN)混合模型,实验结果显示混合模型的效果比单一模型更好。

PM2.5浓度受到多方面因素的影响,不仅受气象影响,包括气压、气温、风速等,还受空气中其他颗粒物或者气体的浓度影响,同时,距离较近的城市或地区之间的空气质量也会相互影响,因此预测PM2.5浓度需要考虑多方面因素。一方面,数据必须需要丰富,数据集中需要有多个PM2.5浓度影响因素的数据,且每天的数据尽可能多。另一方面,数据集中需要包括多个不同地点的数据,以此推算不同地区间的空气质量相互影响的程度。宋飞扬等人[4]提出应用KNN算法计算多个地点间的距离求取地点间的关联性,然后把KNN的输出作为输入与原本的空气数据和天气数据一起放入LSTM 进行训练。Zhang[14]等人提出了时空因果卷积网络(STCN),使用1×1的卷积对多个地点的数据进行特征融合,之后把融合后的特征矩阵传给时序因果卷积网络进行训练。

虽然上述部分模型已经有了对不同地点间PM2.5浓度相互影响的研究,但是忽略了仅靠距离判断相关性的局限,还需要参考不同地点间的数据关联性,不同地点对单个地点的颗粒浓度影响程度也是不同的。因此,应用模型预测一个地点的PM2.5浓度时,考虑不同地点对预测地的不同程度的影响因素是必要的,本文改进时空因果卷积网络,使网络在训练时能考虑到上述不同地点的影响。

1 数据与模型

1.1 数据集

本文选取北京的空气数据作为实验样例,数据来源为微软研究院城市计算团队(Urban Computing Team, Microsoft Research)。该数据集[15]拥有较为丰富的输入变量以及较为完整、详细的气象数据和空气质量数据,包含了自2014年5月1日到2015年4月30日的278 023条空气质量数据和116 867条气象数据,数据按每小时进行记录,不同地点的数据源以经纬度进行标明。每个地点都用包含PM2.5、PM10、CO、气压、温度等共12类数据,数据类型说明见表1所示,各类数据均为数值型数据,其中天气以数字表示晴朗、阴天、多云、下雨、下雪等天气;风向也是用数字表示,以0来表示无风,以1到24来表示24个方向。本文选取了北京市内的36个地点的数据进行实验,每个地点以一小时的间隔进行数据采样。

表1 数据集的类型说明

数据属性	变量名	数据类型	单位
空气质量数据	PM2.5	数值型	$\mu g/m^3$
	PM10	数值型	$\mu g/m^3$
	NO_2	数值型	$\mu g/m^3$
	CO	数值型	$\mu g/m^3$
	O_3	数值型	$\mu g/m^3$
	SO_2	数值型	$\mu g/m^3$
气象数据	天气	数值型	-
	温度	数值型	℃
	气压	数值型	hpa
	湿度	数值型	%
	风速	数值型	m/s
	风向	数值型	-

1.2 改进后的时空因果卷积网络

本文提出一种改进的时空因果卷积网络,时空因果卷积网络由空间信息融合和时序因果卷积两部分组成,

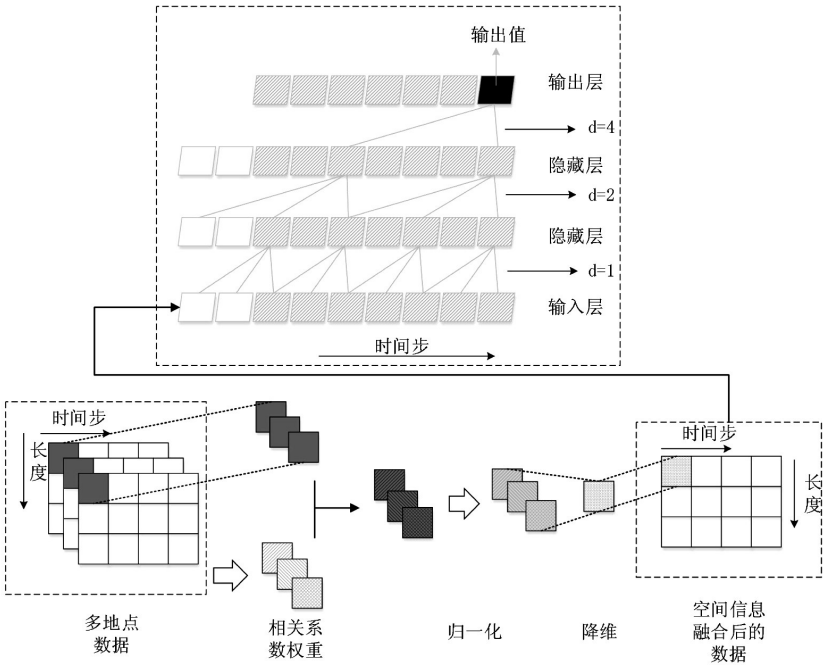


图1 改进后的STCN网络结构

对原始的空间信息融合部分进行改进，调整不同地点对预测结果的影响程度，加深多地点的信息融合。改进后的时空因果卷积网络结构图如图1所示，上半部分为时序因果卷积网络部分，下半部分为空间信息融合部分。

1.2.1 时序因果卷积网络

最近一段时间的研究表明，序列模型都是以循环神经网络为主，如RNN、LSTM、GRU等，Bai[16]等人提出了时序因果卷积网络(TCN)，这是一种由纯卷积进行特征提取的网络，网络主要由因果卷积和空洞卷积组成，并加上残差连接的方式，使得网络性能在序列模型中超过原始RNN以及LSTM与GRU这类门控循环神经网络并且网络结构更加简单，证明了卷积在序列建模上的可行性。

循环网络之所以在序列模型上占有重要的位置，是因为循环网络可以记录输入的时序信息，使得最终的输出结果可以考虑到时序上的先后顺序，但是卷积运算时并不考虑时序信息，于是提出了时序因果卷积，在时序因果卷积中当前时间的输出只依赖于过去与当前的输入，不考虑未来的输入信息。因此时序因果卷积属于单向结构的时间约束模型，模型不会看到未来信息，只关注当前与过去的信息，弥补了卷积在时序信息上的不足。从时序因果卷积的角度看时序问题，利用输入信息$x_1,x_2,\cdots,x_t$预测$y_1,y_2,\cdots,y_t$，那么在x_t处的因果卷积可以用式(1)表示。

$$(F\cdot X)_{x_t}=\sum_{k=1}^{K}f_k x_{t-K+k} \tag{1}$$

式中，F为滤波器，$F=(f_1,f_2,\cdots,f_K)$，X为序列，$X=(x_1,x_2,\cdots,x_T)$。在初次卷积计算时，由于卷积核大小大于1，直接进行卷积操作会获取未来的信息，故在原始的输入前加入补白，使得计算时只关注到当前信息。

卷积操作建立的数据联系会受限于卷积核大小，增加数据联系需要增大感受野范围，通常的做法是增加神经网络的层数，这样会加大训练难度，而通过池化层增加感受野又会造成信息损失，于是TCN采用空洞卷积的方式进行特征提取，在不经过池化层的情况下，以较少的网络层数获得较大的感受野。具体的计算如公式(2)所示，其中F为滤波器，$F=(f_1,f_2,\cdots,f_K)$，X为序列，$X=(x_1,x_2,\cdots,x_T)$，d表示空洞卷积的卷积核间隔数。在初次计算时同样会加入补白以确保不对未来的信息进行计算。并且TCN可以改变空洞卷积的参数d来获取不同的感受野，因此TCN的内存大小是可控的，以此来适应不同的应用场景。

$$(F_{dila=d}\cdot X)_{x_t}=\sum_{k=1}^{K}f_k x_{t-(K-k)d} \tag{2}$$

CNN与RNN相比具有明显的计算优势，卷积在运算中的并行计算能力对于模型的训练效率有着很大的帮助。其次，梯度消失或爆炸是RNN所面临的较大的问题，门控网络对此加以改善，但是门控单元作为网络中存储历史信息的一个组件，在训练时占用较大的内存。TCN由于使用卷积结构，在梯度优化与内存占用方面相比循环网络有较大优势。

1.2.2 空间信息融合

初始的TCN 网络主要用于NLP和语音领域，并未在环境预测方面进行过实际测试，Zhang[14]等人首次把TCN网络应用在PM2.5浓度预测上，并在原本的TCN上加入空间相关性因素，即时空因果卷积网络(STCN)。在距离相隔并不远的地点之间，由于空气的流动，空气中的污染物浓度也会相互影响，因此，在预测某地PM2.5浓度时，不能只依赖此地过去的数据，也需要参考邻近地区的空气质量。STCN把预测点与周围地点的PM2.5浓度数据，采用皮尔逊相关系数(pearson)进行相关性对比，皮尔逊相关系数的计算见式(3)所示。

$$\rho(x,i)=\frac{\mathrm{cov}(x,i)}{\rho_x\rho_i} \tag{3}$$

式中，ρ_x代表预测点的PM2.5浓度数据方差，ρ_i代表所有地点中第i个PM2.5浓度数据的方差，$cov(x,i)$代表x与i的协方差。再将相关系数较高的地点进行信息融合，随后传入TCN预测结果。

本文认为不同地点的数据对需要预测的地点有着不同的影响，如果直接把不同地区的数据同等地放入网络，那么所有地区间的影响都一样，这不合乎常理。空气质量相关性系数越大的地区之间，对PM2.5浓度的影响理应越大，与相关性系数应成正比。在多地点空间信息融合前，对每个地点给予一个权重，将初始权重矩阵设置为地点间的相关系数，再将权重矩阵进行归一化处理，归一化的目的一方面是降低训练难度、加快收敛，另一方面，本文认为归一化后的权重数据分布更为平滑，这对于下一步的信息融合会有益，可能对实验的精度有影响，具体实验在章节2分析。

STCN选择1×1的卷积作为全连接层对多地点的信息进行融合，融合过程中将多维数据先增维至64维，再降维至1维，最终将多个地点的信息融合在一起。本文认为，升维再降维的目的是让网络更好地学习多地点信息的融合方式，但是在融合前已经加入了权重矩阵，在给定了一个更好的权重比后，不应该再让网络调整信息融合时的权重比，这很可能会导致信息的丢失。

2 实验结果与分析

2.1 实验相关配置

实验环境：计算机语言为python3.7；IDE为pycharm 2021.2.3；深度学习框架为pytorch1.10；操作系统为Window10x64；CPU为Intel i7−10750H，32 GB内存；GPU为NVIDIA GTX1650Ti，4 GB内存。

数据集共315 360条数据，包括北京市内的36个地点，每天24小时，共一整年的数据。数据按照6:2:2的比例划分，每个月的前60%的数据作为训练集，中间20%作为验证集，其余作为测试集。为了获取稳定的实验结果，实验时设置了固定的随机数种子。

2.2 评估指标

实验采用均方差(RMSE)，平均绝对误差(MAE)和决定系数(R^2)三个评估指标。各指标的计算方法如式(4)–式(6)所示。

$$RMSE=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-x_i)^2} \tag{4}$$

$$MAE=\frac{1}{n}\sum_{i=1}^{n}|y_i-x_i| \tag{5}$$

$$R^2=1-\frac{\sum_{i=1}^{n}((y_i-x_i)^2)}{\sum_{i=1}^{n}(|y_i-\bar{y}|+|x_i-\bar{y}|)^2} \tag{6}$$

式中，y_i是实际值，x_i是预测值，$\bar{y}$是n个实际值的均值。RMSE和MAE的值越小、R^2的值越大，模型精度越高。

2.3 模型比较与预测结果

对多个地点进行实验，比较改进后的模型与原始的模型之间性能的差异。随机选取了station ID为1013、1018、1030这三个地点作为实验对象，实验结果如表2所示，改进后模型的性能均优于原来的模型。

表2 模型对比

ID:1013	RMSE	MAE	R2
STCN	17.436	11.746	0.936
改进后	16.869 0	10.927 3	0.940 4
ID:1018			
STCN	22.742	13.979	0.939
改进后	19.869 4	12.491 3	0.947 0
ID:1030			
STCN	18.259 8	11.690 4	0.915 4
改进后	15.605 1	9.536 7	0.938 2

为验证不同的修改对实验准确率的影响，本文进行消融实验，以station ID为1013的地点为实验对象，如表3所示，其中α表示在空间关系上加入了相关系数权重，β表示在α上的基础上增加了归一化操作，γ表示取消原来的升维操作，对空间信息直接融合。结果显示，直接取消信息融合时的升维操作对实验结果有益，但是直接增加权重系数并不会带来提升的效果，在两者相结合后，即添加归一化的相关系数权重后再进行信息直接融合的模型性能最好，证明了增加归一化后的权重系数与取消升维操作两者相结合对模型改进的有效性。

使用测试数据集在改进前后的模型上分别预测未来一小时的PM2.5浓度，测试集使用station ID为1013地点的数据，预测日期为每个月的最后6天，图2、图3分别展示了两个模型改进前后的7月份真实数据与预测数据的对比，图4、图5分别为图2与图3数据的散点图。根据图4、图5中的R2显示，改进前的数据相关系数为0.918 3，改进后为0.919，表明在改进后真实数据与预测数据拟合度高于改进前，预测值更精准。

表3 消融实验

模型	RMSE	MAE	R2
STCN	17.436 0	11.746 0	0.936 0
STCN+α	19.426 4	12.353 3	0.920 9
STCN+β	18.465 4	11.734 6	0.928 5
STCN+γ	17.317 6	11.039 0	0.937 1
STCN+α+γ	17.402 2	11.303 0	0.936 5
STCN+β+γ(改进后)	16.869 0	10.927 3	0.940 4

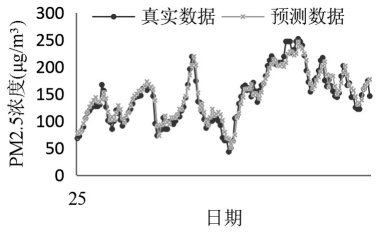


图2 改进前的预测图

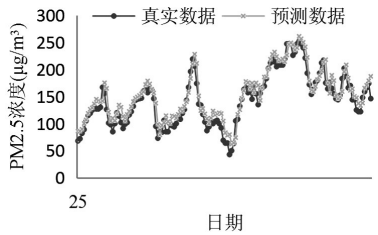


图3 改进后的预测图

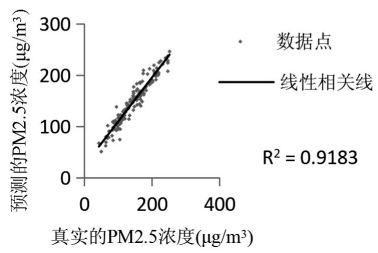


图4 改进前的散点图

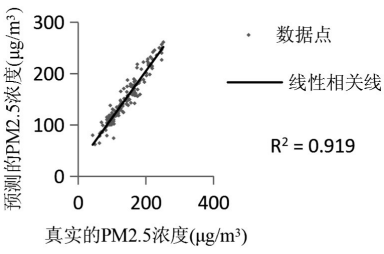


图5 改进后的散点图

3 结束语

为提高模型的PM2.5浓度预测能力，本文提出了一种改进的时空因果卷积网络，改进了不同地点对预测结

果的不同影响，改进后的模型对PM2.5浓度预测更精准。气溶胶光学厚度(AOD)与PM2.5浓度之间有着密切的联系，在接下来的实验中需要考虑AOD因素的影响。另外，不同气体之间会发生化学反应产生新物质，新物质还会受到温度和渗透压的影响，如果将化学影响加以考虑，实验结果很可能还会有所提升。

参考文献：

[1] Sorenden M，Daneshvar B，Hansen M，et al.Personal PM2.5 exposure and markers of oxidative stress in blood[J].Envir onmental Health Perspectives，2003，111(2)：161−166.

[2] Loomis D，Grosse Y，Lauby−Secretan B，et al，2013.The carcinogenicity of outdoor air pollution[J].The Lancet Oncol−ogy，2013，14(13)：1262−1263.

[3] 裘阅，汪家权，胡淑恒.2015−2016年安徽省PM2.5和PM10～2.5时空分布特征[J].合肥工业大学学报(自然科学版)，2020，38(1)：113−118.

[4] 宋飞扬，铁治欣，黄泽华，等.基于KNN−LSTM的PM2.5的浓度预测模型[J].计算机系统应用，2020，29(7)：193−198.

[5] 康俊锋，黄烈星，张春艳，等.多机器学习模型下逐小时PM2.5预测及对比分析[J].中国环境科学，2020，40(5)：1895−1905.

[6] 韦惠红，李剑，张文言，等.基于深度学习和支持向量机集成学习的PM2.5浓度24h预测[J].华中师范大学学报(自然科学版)，2022，56(2)：262−269.

[7] Dennis R L，Byun D W，Novak J H.The next gener−ation of integrated air quality modeling：EPA's models−3[J].Atmosp heric Environment，1996，30(12)：1925−1938.

[8] 王自发，谢付莹，王喜全，等.嵌套网格空气质量预报模式系统的发展与应用[J].大气科学，2006，31(5)：778−790.

[9] 贾佳美，池凯凯，吴哲翔.改进粒子群优化BP神经网络的PM2.5预测[J].计算机工程与设计，2021，42(12)：3495−3501.

[10] 王乃钰，叶育鑫，刘露，等.基于深度学习的语言模型研究进展[J].软件学报，2021，32(4)：1082−1115.

[11] 白盛楠，申晓留.基于LSTM循环神经网络的PM2.5预测[J].计算机应用与软件，2019，36(1)：67−70，104.

[12] 冀林.基于CNNs−GRU深度学习的PM2.5预测研究与实现[D].重庆：重庆邮电大学，2019.

[13] 杨亚莉，李智伟，钟卫军.基于二向注意力循环神经网络的PM2.5浓度预测[J].空军工程大学学报(自然科学版)，2020，21(6)：101−106.

[14] Zhang L，Na J，Zhu J，et al.Spatiotem poral causal convolutional network for forecasting hourly PM2.5 concen−trations in Beijing，China[J].Computers&Geosci ences，2021，155(104869)：1−10.

[15] Zheng Y，Capra L，Wolfson O，et al.Urban Comput−ing：concepts，methodo logies，and applications[J].ACM Transac tion on Intelligent Systems and Techno logy，2014，5(3)：1−55.

[16] Bai S，Kolter J，Koltun V.An Empiri cal E−valua−tion of Generic Convolutio nal and Rec−urrent Networks for Sequence Modeling[DB/OL].2018，https：//arxiv.org/abs/1803.01271.

作者简介：王媛媛(1981−)，女，博士，副教授，研究方向：机器学习、计算机视觉。

(上接第16页)

率的增长，两个巡检机器人的覆盖率均开始下降，巡检机器人1的最低覆盖率为84.68%，巡检机器人2的最低覆盖率为85.31%，均超过设置的最低覆盖率。实验证明：在不同障碍物比率时，本文方法巡检机器人协同路径规划的覆盖率均较高，即路径规划效果较优。

3 结束语

针对复杂环境下，输电线路巡检存在的效率低、碰撞危险高、能耗高等问题，研究复杂环境输电线路巡检机器人协同路径规划方法，结合单目标路径规划与基于交通规则的路径规划，建立复杂环境输电线路巡检机器人协同路径规划模型，缩短协同规划路径距离，降低巡检机器人总能耗，有效避免巡检机器人出现碰撞危险。

参考文献：

[1] 崔宪伟，杨翎，赵勤坤，等.基于计算机视觉的车载轨道缺陷智能巡检系统设计[J].自动化技术与应用，2024，43(7)：35−38.

[2] 董翔宇，季坤，朱俊，等.对特高压变电站巡检机器人路径规划改进蚁群算法的研究[J].电力系统保护与控制，2021，49(18)：154−160.

[3] 刘胜，张豪，晏齐忠，等.基于ACO−SA算法的变电站巡检机器人路径规划[J].南方电网技术，2022，16(9)：75−82.

[4] 邱硕涵，谭章禄，蔡晓梅.基于D__Star算法的巡检机器人路径规划优化研究[J].中国矿业大学学报，2021，50(5)：1011−1018.

[5] 路浩，陈洋，吴怀宇，等.受路网和测量约束的变电站巡检机器人路径规划[J].中国机械工程，2021，32(16)：1972−1982.

[6] 王秀丽，周鹏，侯静楠，等.面向变电站机器人巡检路径规划中的算法研究[J].计算机工程与应用，2021，57(14)：245−250.

[7] 童枭军，陈洋，路浩，等.面向表计读数的变电站巡检机器人路径规划[J].电子测量与仪器学报，2022，36(8)：167−177.

作者简介：郑武略(1987−)，男，本科，高级工程师，研究方向：输电线路运行维护、数字化智能化。