

基于趋势特征向量的火灾烟气流动与温度分布预测

尹云飞¹ 林跃江¹ 黄发良^{1,2} 白翔宇¹

1 重庆大学计算机学院 重庆 400044

2 南宁师范大学计算机与信息工程学院 南宁 530001

摘要 火灾发生时烟气流动与温度分布预测是建筑和消防领域中的热门技术。针对现有的火灾烟气流动与温度分布预测工作烦琐、预测准确度低的现状,提出基于趋势特征向量的火灾烟气流动与温度分布预测模型,用深度学习方法进行相关数据的训练与预测,对揭示火灾发生及其发展规律有重要意义,可为火灾扑救和人员疏散提供辅助信息。所提模型能够抽取火灾时间序列数据中的趋势特征,并将这些特征作为先验知识来加速和优化深度神经网络的训练过程。文中还设计了 LSTM-TFV (LSTM based on Trend Feature Vector)算法。实验结果表明,所提预测模型提高了火灾烟气流动与温度分布预测的准确度,实现了高效且方便的火灾时间序列数据预测。

关键词: 火灾预测;趋势特征;深度学习;温度分布;烟气流动

中图法分类号 TP391

Prediction of Fire Smoke Flow and Temperature Distribution Based on Trend Feature Vector

YIN Yun-fei¹, LIN Yue-jiang¹, HUANG Fa-liang^{1,2} and BAI Xiang-yu¹

1 College of Computer Science, Chongqing University, Chongqing 400044, China

2 School of Computer and Information Engineering, Nanning Normal University, Nanning 530001, China

Abstract The prediction of smoke movement and temperature distribution when a fire occurs is a popular technology in the field of construction and fire protection. At present, this prediction has not been combined with deep neural network technology. Aiming at the current situation that the prediction of fire smoke movement and temperature distribution is cumbersome and the prediction accuracy is low, a prediction model of fire smoke movement and temperature distribution based on trend feature vector is proposed. The deep learning methods are used to train and predict relevant data, which is of great significance to reveal the law of fire occurrence and development and can provide auxiliary information for fire-fighting and fire evacuation. The proposed model can extract the trend features in the fire time series data, and uses these features as a priori knowledge to accelerate and optimize the training process of the deep neural network. This paper designs LSTM-TFV (LSTM based on Trend Feature Vector) algorithm. Experimental results show that the proposed prediction model improves the accuracy of the prediction of fire smoke movement and temperature distribution, and realizes efficient and convenient fire time series data prediction.

Keywords Fire prediction, Trend characteristics, Deep learning, Temperature distribution, Smoke flow

1 引言

火灾发生时产生的高温和浓烟对人类的生命和财产安全造成了极大的危害^[1-3]。因此,预测火灾的发生和发展过程可以为火灾扑救和人员疏散提供重要的信息^[4]。现有的主流火灾预测模型是 CFAST (Consolidated Model of Fire and Smoke Transport)^[5],它是一种综合的火灾烟气、温度预测模型。目前,CFAST 火灾预测模型已被广泛应用于火灾研究和消防应用领域中。使用 CFAST 火灾预测模型的人员包括火灾调查人员、安全官员、建筑工程师、建筑承包商等,他们使用

CFAST 模型来研究火灾发生时温度和烟气对特定建筑和环境的影响^[6]。在大多数情况下,用户根据经验来设定 CFAST 模型的参数和系统的初始状态,包括起始温度、热释放速率、材料特性、隔间尺寸、机械通风规格、火源特性等^[7],然后根据经验来修正 CFAST 输出的预测结果。因此,CFAST 模型的使用是一个相当复杂且耗时的过程。CFAST 的预测准确度在大多数情况下与实际情况不相符,这就需要工程技术人员根据经验进行修正^[8]。如何修正 CFAST 模型的预测结果是一个十分有趣且值得研究的问题。

当前,学者们使用数据同化的方法将火灾预测模型与测

到稿日期:2020-06-17 返修日期:2020-08-31 本文已加入开放科学计划(OSID),请扫描上方二维码获取补充信息。

基金项目:国家自然科学基金(61962038);广西八桂学者创新团队基金(201979)

This work was supported by the National Natural Science Foundation of China(61962038) and Guangxi Bagui Teams for Innovation and Research (201979).

通信作者:尹云飞(yinyunfei@cqu.edu.cn)

量数据结合起来进行预测,达到了提高预测精度的目的。常用的数据同化方法有顺序同化^[9]、变分同化^[10]、最大条件概率同化^[11]等。近年来,集合卡尔曼滤波方法(Ensemble Kalman Filter, EnKF)被应用于 CFAST 模型预测结果的修正中。EnKF 是一种滤波同化方法,在对 CFAST 预测结果的修正中,其性能超出了其他同化方法。例如, Lin 等^[12]使用 EnKF 方法修正 CFAST 模型对两层建筑进行了实时火灾预测研究。Lin 等^[13]使用 EnKF 方法修正了 CFAST 对多个房间烟雾扩散的预测结果。基于 EnKF 进行反演建模, Ji 等^[14]对隧道火灾进行了预测研究,并通过实验展示了预测效果的提升情况。这些研究虽然都在一定程度上提高了 CFAST 模型预测的准确度,但是存在一些不可避免的缺陷。

首先,用于验证 EnKF 方法的实验场景远比实际场景简单,并且 EnKF 方法无法使用过去已经发生在真实场景下的火灾历史数据进行分析。

其次,EnKF 方法只能对 CFAST 模型预测的下一时刻数据进行修正,而无法对更长时间之后的预测数据进行修正。

最后,在对 CFAST 模型的预测结果进行修正时,EnKF 方法的费用极其昂贵,因为它需要事先部署众多高精度的传感器。对于没有部署传感器的地方,EnKF 方法将失去作用。

现有方法的这些缺点影响了火灾的预测、扑救和疏散。

随着深度学习技术的快速发展,探讨利用深度学习方法来提高 CFAST 模型的预测精度成为了研究的热点。深度学习技术具有全局寻优和万能逼近的特点,它能够从火灾历史数据中学习潜在的知识和规律,并通过神经网络来表达和保存。在对趋势数据的预测中,深度学习方法达到了超越传统方法的效果^[15]。深度学习技术利用深度神经网络的自学习、自组织、自适应特点,可以处理环境信息复杂、知识背景模糊、推理规则不明确的问题。长短期记忆网络(Long Short-Term Memory, LSTM)是一种深度神经网络。LSTM 具有独特的端到端建模方法、容易引入外部变量、能够自动提取特征等优点,因而备受关注^[16-17]。LSTM 已被证明适用于处理与时间序列预测相关的任务,如语音识别^[18]、股票市场预测^[19]、机器翻译^[20]等。

数据同化方法依赖数据的实时同步和矫正,而神经网络方法依赖历史数据的训练和学习。

为了深入挖掘时间序列数据预测的准确度,本文研究了利用时间序列数据的内在特征来辅助深度学习的方法。具体做法是:从时间序列数据中抽取趋势特征信息,并将这些信息作为深度学习的外部变量,最终构建出具有先验知识的深度学习模型。常见的趋势特征计算方法有 Cochran-Armitage 方法^[21]、Mann-Kendall 方法^[22]、Kendall's tau-b 方法^[23]。这些方法能够从火灾时间序列数据中抽取出变化趋势,并且能进一步检测出时间序列数据是否发生了突变。趋势特征分析被广泛应用于气候分析^[24]、水文分析^[25]、植被分析^[26]等方面。

主流的深度学习模型有 CNN, RNN, LSTM 等。由于火灾烟气流动与温度分布预测是一个与时间相关的预测任务,因此本文选用基于时间序列数据处理的 LSTM 进行建模。相比其他深度学习模型, LSTM 模型具有记忆时间长、在大数

据集上表现效果好、易与趋势特征相结合等优点。

由于测量误差、环境因素、人为因素等的存在,某一特征的真实值并不存在^[27]。在研究工作中,一般使用观察测量值(即观测值)来近似替代真实值。

由于目前学术界、产业界的研究目标是研发提高 CFAST 预测准确度的技术,因此本文的动机是利用深度神经网络 LSTM 模型来对火灾发生时烟气流动和温度分布数据进行深度学习,并且在学习过程中利用趋势特征信息来加速和优化学习过程,最后通过学习到的神经网络结构和参数对火灾烟气流动和温度分布进行预测。本文还通过实验验证了所提方法的有效性,探讨了各种因素对该方法的影响。

本文的贡献如下:

(1)提出了基于趋势特征向量的火灾烟气流动与温度分布预测模型。

(2)首次将深度神经网络 RNN, LSTM 等用于对火灾发生时的烟气时间序列数据、温度时间序列数据的训练和预测。

(3)探讨了趋势特征、深度学习参数对时间序列数据预测准确性的影响。

本文第 2 节首先简要介绍学术界和产业界普遍使用的 CFAST 火灾预测模型,然后提出基于趋势特征向量的火灾烟气流动与温度分布预测模型,其中包括模型的结构、深度神经网络设计、趋势特征向量计算、算法设计等;第 3 节通过实验对比所提预测模型与其他模型的性能,并且在趋势特征和 LSTM 参数的影响等方面对所提模型进行了实验、分析和探讨;最后总结全文。

2 模型与方法

2.1 火灾预测模型 CFAST

CFAST 是一个双区域火灾预测模型,由美国国家标准与技术研究院(National Institute of Standards and Technology, NIST)提供。它可以预测多室区域内由火灾引起的热环境。在 CFAST 中,每个房间被分为上气体层和下气体层,火焰通过羽流将燃烧产物从下层驱动到上层。在每个房间的上气体层和下气体层中,假设其温度是均匀的。

根据质量守恒方程、动量守恒方程、能量守恒方程、理想气体定律、羽流模型、排气流动方程、辐射和燃烧模型,构建火灾烟气运动和温度分布方程组^[28]。通过小增量时间内的焓和质量流特性求解状态变量,如烟气状态、温度变化等。在使用 CFAST 火灾预测模型时,通常需要设定材料性能、隔间尺寸、机械通风规格、火源特性、喷头和探测器规格等参数,这样就可以预测状态变量的变化,如烟层高度(Smoke Layer Height, SLH)的变化、下层温度(Low Temperature, LT)的变化、上层温度(Upper Temperature, UT)的变化等。

2.2 基于趋势特征向量的火灾烟气流动与温度分布预测模型

目前,对于火灾发生时烟气流动和温度分布的预测,大多数机构和研究人员是通过 CFAST 来进行的。本文提出了一种基于趋势特征向量的火灾烟气流动与温度分布预测模型,它改进了 LSTM 深度神经网络的预测准确度,并且与 CFAST 深度融合。

针对目前学术界、产业界努力提高 CFAST 预测准确度

的需求,本文定位于研究 CFAST 预测值与其对应的真实值(可以用观测值来替代)之间的偏差。

定义 1 CFAST 的校准偏差(CFAST Observation Error, COE),即 CFAST 的预测值与观测值(真实值)之间的偏差,如式(1)所示:

$$COE = y_{\text{obs}} - y_{\text{CFAST}} \quad (1)$$

其中, y_{obs} 是火灾发生时烟气流动和温度分布的观测值, y_{CFAST} 是 CFAST 对火灾发生时烟气流动和温度分布的预测。通过 COE 可以校准 CFAST 的预测结果。

图 1 是基于趋势特性向量的火灾烟气流动与温度分布预测模型结构图。

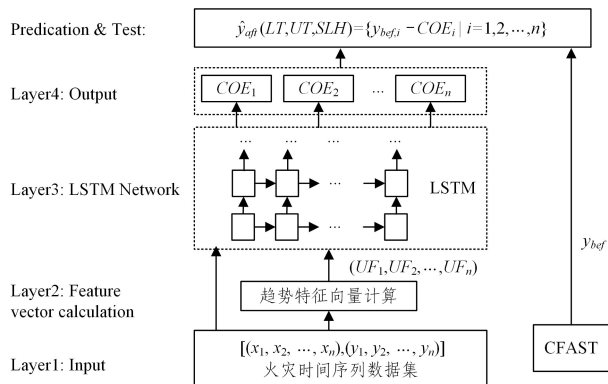


图 1 基于趋势特征向量的火灾烟气流动与温度分布预测模型

Fig. 1 Prediction model of fire smoke flow and temperature distribution based on trend feature vector

在图 1 中,输入层是火灾序列数据 $[(x_1, x_2, \dots, x_n), (y_1, y_2, \dots, y_n)]$, 它包括 LT, UT 和 SLH 3 个维度, 其中 (x_1, x_2, \dots, x_n) 表示火灾时间序列数据, (y_1, y_2, \dots, y_n) 表示 (x_1, x_2, \dots, x_n) 对应的观测值(或者真实值)。输入层的火灾序列数据有很多, 因为只有达到一定数量的输入数据才能保证深度神经网络的训练结果具有实用性。特征提取层的作用是从火灾序列数据中提取趋势特征向量。网络层是 LSTM 深度神经网络。输出层是火灾时间序列数据的校准偏差预测, 即 COE , 它包括 LT 序列的校准偏差预测、 UT 序列的校准偏差预测、 SLH 序列的校准偏差预测。 $COE_1, COE_2, \dots, COE_n$ 是 LSTM 网络的输出结果和优化结果。在预测和测试层, y_{bef} 是测试数据, 它来源于 CFAST 的一次预测; 通过 $(COE_1, COE_2, \dots, COE_n)$ 对 y_{bef} 进行校准, 就可以得到基于深度神经网的火灾烟气流动与温度分布预测结果。

(1) 优化目标

本文中,模型的优化目标是 minimized 损失函数。损失函数的定义为:

$$L = \sum_{i=1}^m (COE_i^{\text{exp}} - COE_i^{\text{pred}})^2 \quad (2)$$

其中, L 是损失函数, COE_i^{exp} 是 COE 的期望值, COE_i^{pred} 是深度神经网络 LSTM 对 COE 的预测结果。 COE_i 是第 i 个 LSTM 单元的处理对象, 它包括 COE 的期望值和预测值。

(2) 数据集的融合

由于火灾事故的数据集很少, 单个火灾场景的数据量难以满足训练要求, 因此, 我们融合多个场景的火灾数据集进行实验, 并将最小化损失函数作为 LSTM 优化的目标。

2.2.1 深度神经网络设计

LSTM 是长短期记忆深度神经网络, 它解决了传统的递归神经网络在处理长时间序列数据时存在梯度消失和梯度爆炸的问题。 LSTM 通过“门”单元使得误差以常数的形式在网络中传递。 LSTM 的内部单元结构如图 2 所示。

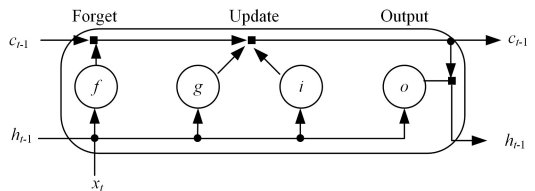


图 2 深度神经网络 LSTM 的网络单元

Fig. 2 Network unit of deep neural network LSTM

在图 2 中, x_t 表示本单元在 t 时刻的输入数据; C_{t-1} 表示上一单元在 $t-1$ 时刻的状态, C_t 表示本单元在 t 时刻的状态; h_{t-1} 表示上一单元在 $t-1$ 时刻其记忆单元的输出, h_t 表示本单元在 t 时刻其记忆单元的输出; f 是“遗忘门”, i 是“输入门”, o 是“输出门”; g 是神经元状态更新操作。

$$f_t = \sigma(W_f x_t + R_f h_{t-1} + b_f) \quad (3)$$

其中, f_t 表示 t 时刻本单元的“遗忘门”, 它通过 σ 函数选择一部分 x_t 和 h_{t-1} 的数据, 忘记一部分 x_t 和 h_{t-1} 的数据。 W_f 和 R_f 是权重, b_f 是偏置。

$$g_t = \tanh(W_g x_t + R_g h_{t-1} + b_g) \quad (4)$$

其中, g_t 是 t 时刻本单元神经元的更新操作, 它是 x_t 和 h_{t-1} 的 \tanh 函数。 W_g 和 R_g 是权重, b_g 是偏置。

$$i_t = \sigma(W_i x_t + R_i h_{t-1} + b_i) \quad (5)$$

其中, i_t 是 t 时刻本单元的“输入门”, 它通过 σ 函数选择 x_t 和 h_{t-1} 的一部分信息到神经元的记忆单元中。 W_i 和 R_i 是权重, b_i 是偏置。

$$o_t = \sigma(W_o x_t + R_o h_{t-1} + b_o) \quad (6)$$

其中, o_t 是 t 时刻本单元的“输出门”, 它通过 σ 函数选择 x_t 和 h_{t-1} 的一部分信息输出到下一单元。 W_o 和 R_o 是权重, b_o 是偏置。

使用式(7)和式(8)表示本单元在 t 时刻的状态 c_t 和输出 h_t , 其中“ \otimes ”表示点积。

$$c_t = f_t \otimes c_{t-1} + i_t \otimes g_t \quad (7)$$

其中, f_t 是 t 时刻的“遗忘门”, c_{t-1} 是 $t-1$ 时刻的状态, i_t 和 g_t 分别是 t 时刻的“输入门”和神经元状态更新操作。

$$h_t = o_t \otimes \tanh(c_t) \quad (8)$$

其中, o_t 是 t 时刻的“输出门”, c_t 是 t 时刻的状态。对 c_t 进行 \tanh 函数后与 o_t 进行点乘即可得到 t 时刻的输出 h_t 。

得益于 3 个控制门和记忆单元, 深度神经网络 LSTM 可以容易地保存、读取、复位和更新长时间序列数据的信息。值得注意的是, 由于深度神经网络 LSTM 内部参数的共享机制, 输出的维度可以通过设置权重矩阵的维度来控制。

2.2.2 趋势特征向量计算

对于具有 n 个数据的时间序列 $X = \{x_1, x_2, \dots, x_n\}$, 构造一个秩序列:

$$S_k = \sum_{i=1}^k \sum_{j=1}^i r_{i,j}, k=1, 2, \dots, n \quad (9)$$

其中, S_k 表示 X 中第 k 个数据的秩, $r_{i,j}$ 的取值如式(10)所示:

$$r_{i,j} = \begin{cases} 1, & x_i > x_j \\ 0, & \text{otherwise} \end{cases}, 1 \leq j \leq i \quad (10)$$

其中, $r_{i,j}$ 表示 x_i 和 x_j 的比较结果(取 0 或 1)。式(10)即, 当 x_i 大于 x_j 时, $r_{i,j} = 1$; 否则, $r_{i,j} = 0$ 。 x_i 和 x_j 分别表示 X 中的第 i 个数据和第 j 个数据。

S_k 是秩序列中 k 时刻的元素, 表达了 k 时刻之前任意两个时刻对应数据出现增长的概率(即后面的数据大于前面数据的统计)。

定义统计变量 UF_k :

$$UF_k = \frac{[S_k - E(S_k)]}{\sqrt{Var(S_k)}}, k=1, 2, \dots, n \quad (11)$$

其中, UF_k 是 k 时刻的统计变量, S_k 是 k 时刻的秩元素, $E(S_k)$ 是 S_k 的均值, $Var(S_k)$ 是 S_k 的方差。称 $UF = \{UF_1, UF_2, \dots, UF_n\}$ 为趋势特征向量。

在火灾烟气流动和温度分布预测中, 取 $UF_1 = 0, E(S_k) = (k(k+1))/4, Var(S_k) = (k(k-1)(2k+5))/72$ 。

令 $X' = \{x_n, \dots, x_2, x_1\}$, 重复上述步骤, 计算趋势特征向量 UB 。显然, $UB = \{UB_1, UB_2, \dots, UB_n\}$ 是另一个趋势特征向量。

UF 与 UB 满足式(12):

$$UB_k = -UF_k, k=(n, n-1, \dots, 1), UB_1 = 0 \quad (12)$$

其中, UB_k 是第 k 个数据的统计变量。

由于在火灾发生过程中存在较多的外界干扰, 严重影响了预测的准确度, 因此我们设计了容差比较机制:

$$r_i = \begin{cases} 1, & f_{\text{toler}}(x_i, x_j, w) \\ 0, & \text{otherwise} \end{cases}, j=1, 2, \dots, i \quad (13)$$

其中, $f_{\text{toler}}(x_i, x_j, w)$ 是容差函数, 它用来减少干扰对 x_i 与 x_j 之间差值的影响。例如, $x_i - x_j > w$, 则 $r_i = 1$, 否则 $r_i = 0$ 。

为了分析序列 X 的变化趋势, 需要考查趋势特征向量 UF 和 UB 的取值情况。当 $UF_k > 0$ 时, 表明序列 X 呈上升趋势; 当 $UF_k < 0$ 时, 表示序列 X 呈下降趋势; 具有相似变化趋势的序列具有相近的 UF_k 和 UB_k 。

针对火灾发生时上层温度、下层温度和烟气层高度预测的工程和研究需求, 总结火灾发生时烟气流动和温度分布的特性, 具体如下:

(1) 当火灾烟气流动和温度分布稳定时, 即时间序列趋于稳定时, 式(9)和式(10)的计算容易受到外界扰动的影响, 这使得趋势特征向量 UF 和 UB 的计算会出现较大偏差。

(2) 在火灾发生的前期, LT 和 UT 变化幅度较大, 而在火灾发生的后期, 它们的变化幅度较小。

(3) 因为趋势特征向量 UB 是时间逆序计算得到的, 它更容易受到后期外界扰动的影响。

(4) 当序列数据某一维变量的取值范围较大时, 外界干扰对预测性能的影响小; 当序列数据某一维变量的取值范围较小时, 外界干扰对预测性能的影响大。举例说明, 序列数据中的 UT 和 LT 对应的取值范围大, 即 UT 和 LT 的变化范围为 $20 \sim 1300^\circ\text{C}$, 那么外界干扰对它的影响是次要的; SLH 对应的序列数据变化范围小, 即 SLH 的变化范围为 $0 \sim 2\text{m}$, 那么外界出现的很小的干扰也能很明显地表现出来。

根据火灾发生时烟气流动和温度分布的特性, 本文设计

了基于趋势特征向量的时间序列预测机制:

(1) 使用式(9)~式(11)进行 LT 、 UT 的趋势特征的刻画, 即计算它们的趋势特征向量 UF 的值;

(2) 使用式(9)、式(13)、式(11)进行 SLH 的趋势特征的刻画, 即计算它的趋势特征向量 UF 的值。

2.2.3 算法设计

为了实现基于趋势特征向量的火灾烟气流动与温度分布预测程序, 需要设计对应的算法, 即 LSTM-TFV (LSTM based on Trend Feature Vector) 算法。该算法的思想是, 利用火灾烟气流动与温度分布的标准数据作为模型的输入数据, 利用趋势特征向量计算方法计算下层温度 LT 的趋势特征向量、上层温度 UT 的趋势特征向量和烟气层高度 SLH 的趋势特征向量。在每一时刻 t , 向深度神经网络 LSTM 的神经单元输入 x_t, y_t 和 UF_t 。随后, 进入 LSTM 神经网络的迭代训练过程。在每一次迭代中, 计算对应的校准偏差 $COE = (COE_1, COE_2, \dots, COE_n)$, 再经过损失函数处理后, 进行反向求梯度和迭代优化。

最后, 根据式(14), 用训练好的模型参数预测火灾发生时的烟气流动和温度分布。

$$\hat{y}(t) = x_t - COE_t, t=1, 2, \dots, n \quad (14)$$

其中, $\hat{y}(t)$ 是 t 时刻 LT, UT, SLH 的预测值; x_t 是 t 时刻 CFAST 的预测值; COE_t 是 t 时刻 CFAST 预测数据存在的误差, 即把这个误差作为矫正值来矫正 CFAST 的预测。

在对模型进行充分训练后, 不需要观测值, 只需要 CFAST 的输出值, 即可准确地基于 CFAST 的输出预测出烟气流动和温度分布。改进后的模型提高了 CFAST 的实用性。

算法 1 基于趋势特征向量的 LSTM 预测算法 (LSTM-TFV)

输入: 时间序列数据 $[(x_1, x_2, \dots, x_n), (y_1, y_2, \dots, y_n)]$; 火灾特性参数 MP, CS, MVS, FSC, ST, Hrr; 学习速率 a ; 隐藏层单元的数量 h ; 训练过程中的批处理大小 b

输出: COE ; 预测值 \hat{y}

1. $x_train \leftarrow [(x_1, x_2, \dots, x_n), (y_1, y_2, \dots, y_n)]$;
2. for each LT in x_train : // LT 是下层温度
3. $UF \leftarrow TFV_LT(LT)$; // TFV_LT 是式(9)~式(11), 它可以计算 UF
4. for each UT in x_train : // UT 是上层温度
5. $UF \leftarrow TFV_UT(UT)$; // TFV_UT 是式(9)~式(11), 它可以计算 UF
6. for each SLH in x_train : // SLH 是烟气层高度
7. $UF \leftarrow TFV_SLH(SLH)$; // TFV_SLH 是式(9)、式(11)、式(13), 它可以计算 UF
8. $COE \leftarrow LSTM(x_train, UF, a, h, b, otherParas)$; // 使用 LSTM 进行模型训练
9. $x_test = CFAST(MP, CS, MVS, FSC, ST, Hrr)$; // 获得 CFAST 的一组输出作为测试数据
10. $\hat{y} \leftarrow x_test - COE$; // 根据式(14)计算预测值

在算法 1 中, 第 1 步被用来获得输入数据, 即进行输入层操作。第 2~7 步用于进行趋势特征向量的计算。第 8 步用于进行深度神经网络 LSTM 的训练, 对应的参数包括训练数据集、趋势特征向量、学习速率以及其他常用的神经网络训练

参数。第9步用于获取测试数据,即通过 CFAST 计算出一组火灾烟气流动和温度分布数据,CFAST 使用的参数包括材料特性(Material Performance,MP)、隔间尺寸(Compartment Size,CS)、机械通风规格(Mechanical Ventilation Specifications,MVS)、火源特性(Fire Source Characteristics,FSC)、起始温度(Start Temperature,ST)、热释放速率(Heat Release Rate,Hrr)。第10步用于计算预测值。

3 实验

3.1 数据集

为了验证基于趋势特征向量火灾烟气流动与温度分布预测模型的有效性,本文采用美国国家标准与技术研究院 NIST 提供的火灾烟气运动标准数据集。实验数据集的收集、下载和使用许可参见文献[5]。

火灾烟气流动和温度分布预测实验涉及4种不同的场景,即 ATF corridors^[29],FM/SNL^[30-31],NIST/NRC^[32],WTC^[33]。其中,ATF corridors 是不同火源下两层建筑连接走廊的温度和烟气传播场景,它由火灾研究实验室 ATF 提供。FM/SNL 是核电站中各种典型隔间火灾烟气运动的场景,它由 Factory Mutual 和 Sandia National 实验室提供。Factory Mutual 和 Sandia National 实验室进行了一系列实验,其中涉及的参数包括火灾强度、通风速率和火灾位置等,这些实验验证了核电站中各种典型隔间类型的计算机模型。NIST/NRC 也是核电站中的火灾场景,但是与 FM/SNL 关注的场景不同,NIST/NRC 的实验场景中还包括电力、电缆,以及不同热特性的墙壁和天花板。WTC 是商业大厦中的火灾场景,它来源于世贸中心火灾调查事件,因此它对类似的商业大厦消防工作具有重要的参考价值。表1是实验中用到的数据集。

表1 数据集
Table 1 Datasets

Data sets	Adapted scenes	Number of samples	Sample length
D1	ATF,FM/SNL,NIST/NRC,WTC	43	90~200
D2	ATF,FM/SNL,NIST/NRC,WTC	43	90~200
D3	ATF,FM/SNL,NIST/NRC,WTC	43	90~200

火灾烟气流动和温度分布预测实验涉及3个数据集,每个数据集有43个样本序列(时间序列数据),这些样本序列的长度大部分在90~200之间。从43个样本序列中随机选出4个分别作为4种场景的测试集,剩余的39个序列作为训练集。对实验结果的验证采用的是 Holdout 方法。

3.2 实验平台及参数选择

实验平台选用 DELL Optiplex(Intel i5-2400 CPU @ 3.10GHz Dual core,10G Memory)+Windos 7(sp1) 64 位。深度学习框架选用 Matlab2018b 提供的深度学习工具箱。

火灾数据参数是根据火灾实验场景的物理特性进行设置的。神经网络参数的选择涉及隐藏层单元的数量(hidden units)、训练过程中的批处理大小(batch size)和学习速率(learning rate)。这些参数的选择方法是通过实验尝试,最终确定一个合适的值。具体做法是:先选定其中一个参数,同时固定其他参数,进行多次训练,记录每次实验的均方根误差(Root Mean Squared Error, RMSE)。重复上述实验,最后选

取表现最佳的参数。预防过拟合的方法是 Dropout。

3.3 评估指标

为了评估算法的预测性能,我们使用 RMSE 和决定系数(Coefficient of Determination, R Squared)作为评价指标。RMSE 值越小,预测值与观测值的偏差就越小;R Squared 越接近1,拟合效果就越好。

RMSE 的计算方法如式(15)所示:

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^N (y(t) - \hat{y}(t))^2} \quad (15)$$

其中,N 是样本长度,y(t)是观测值, $\hat{y}(t)$ 是预测值。

R Squared 的计算方法如式(16)所示:

$$R^2(y, \bar{y}) = 1 - \frac{\sum_{t=1}^N (y(t) - \hat{y}(t))^2}{\sum_{t=1}^N (y(t) - \bar{y}(t))^2} \quad (16)$$

其中,N 是样本长度;y(t)是观测值; $\bar{y}(t)$ 是 y(t)的均值,即 $\bar{y}(t) = \frac{1}{N} \sum_{t=1}^N y(t)$; $\hat{y}(t)$ 是预测值。

3.4 性能比较

为了评估算法的性能,本文对以下算法进行代码实现和实验。

(1)CFAST-RNN 算法:深度神经网络 RNN 被用来对火灾数据进行训练,并且基于训练的模型对 CFAST 的预测值进行校准;

(2)CFAST-LSTM 算法:深度神经网络 LSTM 被用来对火灾数据进行训练,并且基于训练的模型对 CFAST 的预测值进行校准;

(3)CFAST-GRU 算法:深度神经网络 GRU 被用来对火灾数据进行训练,并且基于训练的模型对 CFAST 的预测值进行校准;

(4)LSTM-TFV 算法:根据本文提出的基于趋势特征向量的预测方法,利用火灾数据来对深度神经网络 LSTM 进行训练,并且基于训练的模型对 CFAST 的预测值进行校准。

实验集中对比 LSTM-TFV, CFAST-LSTM, CFAST-RNN, CFAST-GRU 和 CFAST 5 种方法在 4 种场景(即 ATF corridors 场景、FM/SNL 场景、NIST/NRC 场景、WTC 场景)下,对 SLH,LT 及 UT 的预测精度。其中,CFAST-RNN 和 CFAST-LSTM 的算法流程与 LSTM-TFV 的算法流程类似。CFAST-RNN 和 CFAST-LSTM 的训练集中不包含趋势统计序列数据,CFAST-RNN 使用 RNN 网络代替 LSTM 网络。

表2~表4列出了5种方法在4个场景下对烟气层高度 SLH、下层温度 LT 和上层温度 UT 预测的性能。表2中的数据表示5种方法在不同场景下的评估数据,其中 RMSE 代表均方根误差、R Squared 代表决定系数。

从表2可以看出,LSTM-TFV 算法在 ATF,FM/SNL, WTC 3 个场景下取得了最好的效果。在 NIST/NRC 场景下,LSTM-TFV 算法的预测效果与 CFAST-LSTM 算法接近。这是因为 NIST/NRC 场景在运行一段时间后突然受到较大的外界干扰并且表现出随机性。因此,该场景的时间趋势特征相对不太明显并且有混沌现象,这使得 LSTM-TFV 方法的趋势特征预测能力没有得到最大的发挥。

表3列出了 CFAST, CFAST-RNN, CFAST-LSTM, CFAST-GRU, LSTM-TFV 5 种预测方法对上层温度 LT 的

预测性能。从 R Square 指标来看,本文提出的 LSTM-TFV 算法在 4 个场景下的预测性能均达到了最优;从 RMSE 指标来看,LSTM-TFV 算法表现出的性能远超其他算法。

表 4 列出了 CFAST, CFAST-RNN, CFAST-LSTM, CFAST-GRU, LSTM-TFV 5 种预测方法对下层温度 UT 的

预测性能。其中,本文提出的 LSTM-TFV 算法在 ATF, FM/SNL, WTC 3 个场景下表现出最优的性能。在 NIST/NRC 场景下, LSTM-TFV 算法与 CFAST-LSTM 算法的预测性能相近。这是因为 NIST/NRC 场景在运行一段时间后受到了较大的外界干扰,这影响了趋势特征向量的计算结果。

表 2 SLH 的预测结果

Table 2 Predictive performance of SLH

Evaluation Scene	RMSE				R Squared			
	ATF	FM/SNL	NIST/NRC	WTC	ATF	FM/SNL	NIST/NRC	WTC
CFAST	0.715	1.811	0.426	1.292	-0.034	0.215	0.488	0.160
CFAST-RNN	1.285	0.268	0.162	1.638	-0.058	0.786	0.887	0.142
CFAST-LSTM	0.099	0.205	0.075	0.133	0.659	0.841	0.960	0.817
CFAST-GRU	0.0671	0.2622	0.0755	0.1389	0.8715	0.8299	0.9600	0.7508
LSTM-TFV	0.056	0.173	0.112	0.124	0.947	0.918	0.934	0.849

表 3 LT 的预测结果

Table 3 Predictive performance of LT

Evaluation Scene	RMSE				R Squared			
	ATF	FM/SNL	NIST/NRC	WTC	ATF	FM/SNL	NIST/NRC	WTC
CFAST	2.176	9.815	19.400	33.207	0.808	-0.354	0.608	0.324
CFAST-RNN	3.871	5.028	18.084	14.035	0.641	0.614	0.670	0.203
CFAST-LSTM	1.429	3.773	8.272	7.791	0.826	0.823	0.844	0.762
CFAST-GRU	2.6479	3.6741	7.4199	4.8080	0.3809	0.9042	0.8614	0.8878
LSTM-TFV	1.389	2.374	4.656	7.058	0.828	0.946	0.930	0.838

表 4 UT 的预测性能

Table 4 Predictive performance of UT

Evaluation Scene	RMSE				R Squared			
	ATF	FM/SNL	NIST/NRC	WTC	ATF	FM/SNL	NIST/NRC	WTC
CFAST	12.071	6.663	23.051	62.842	0.487	0.528	0.806	0.787
CFAST-RNN	11.958	6.778	24.257	41.487	0.632	0.821	0.743	0.769
CFAST-LSTM	9.769	4.140	8.286	35.365	0.723	0.870	0.975	0.898
CFAST-GRU	4.9784	4.3006	12.3717	31.427	0.8981	0.864	0.9365	0.9131
LSTM-TFV	4.843	3.482	10.918	28.021	0.929	0.917	0.956	0.921

为了对火灾发生时烟气流动和温度分布进行准确的预测,需要选择合适的预测方法。与 CFAST 相比, CFAST-RNN 方法的预测性能已经有所提高,但是在某些情况下 CFAST-RNN 方法的预测性能略低于 CFAST。这说明了 CFAST-RNN 的预测性能在总体上有所提高,但是并不是适合所有的场景,表现出了不稳定性。CFAST-LSTM 很大程度上提高了预测的准确性,并且也表现出了很好的稳定性(普遍适用性)。因此,它的各项指标都远优于 CFAST,并且比 CFAST-RNN 好。LSTM-TFV 的预测性能在所有方法中是最好的,它的适应性也最强。从实验结果可以看出,在两个评价指标上,大多数的最优表现都由 LSTM-TFV 方法获得。

相比对比模型,本文提出的 LSTM-TFV 模型表现出了较好的性能。其内在原因是,在 LSTM 神经网络的训练过程中,所提方法利用了从时间序列数据中获得的趋势特征。利用趋势特征向量来计算校准偏差 COE ,可使训练得到的 LSTM 网络的性能达到相对较优。

基于以上分析,可以得出以下结论:

(1) 基于深度神经网络的火灾烟气流动和温度分布预测能显著提高预测的准确度。

(2) 对于长期的时间序列任务,即时间序列数据较长时, LSTM-TFV 方法和 CFAST-LSTM 方法优于 CFAST-RNN 方法。

(3) 不同的火灾场景,虽时间序列数据长度不同,但都可以采用 LSTM-TFV 方法和 CFAST-LSTM 方法。通过深度神经网络来预测 COE ,进而通过 COE 来校准预测结果的方法是可行的。这种方法提高了 CFAST 的预测准确度。

(4) LSTM-TFV 方法充分利用了时间序列数据的趋势特征性质,因此在所有趋势特征较明显的场景中, LSTM-TFV 方法的性能优于 CFAST-LSTM 方法。

3.5 趋势特征的影响

为了进一步验证趋势特征向量在时间序列数据预测中的作用,本文进行了趋势特征对时间序列数据预测准确性的研究,对火灾时间序列数据进行了多次训练和预测,并在每次训练和预测中记录其对应的均方根误差 $RMSE$ 。

图 3 给出了趋势特征对火灾时间序列数据预测准确性的影响。

箱型图箱体的上边缘表示 25% 分位数,中间红线表示 50% 分位数(中位数),箱体的下边缘表示 75% 分位数,箱型图上面的短横线表示最小值,箱型图下面的短横线表示最大值。

从图 3 的实验结果来看,对于火灾 SLH, LT, UT 的预测, LSTM-TFV 方法对应的箱型图的各个指标(25% 分位数、50% 分位数、75% 分位数、最小值、最大值)均明显低于 CFAST-LSTM 方法。这说明趋势特征的加入提高了模型的预测准确度和适应性。

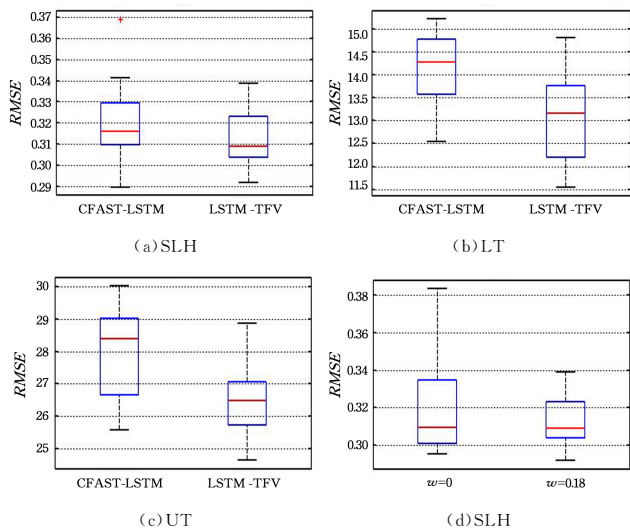


图3 趋势特征对火灾时间序列数据预测准确性的影响

(电子版为彩色)

Fig. 3 Influence of trend features on prediction accuracy of fire time series data

在 LSTM-TFV 方法对烟气层高度的预测中,设置式(13)的参数 $\omega=0, 0.18$ 。与 $\omega=0$ 的情况相比, $\omega=0.18$ 时均方根误差 RMSE 更集中、变化范围更小。因此,动态调整 ω 既可以保证时间序列数据预测的准确性,又能提高方法的适用性。

3.6 LSTM 参数的影响

为了验证深度神经网络 LSTM 结构的设计对时间序列数

据预测的影响,本文进行了 CFAS 观测值误差 COE 的研究。

LSTM 有 3 个重要参数,即隐藏层单元的数量(h)、训练过程中的批处理大小(b)和学习速率(a)。

对于隐藏层单元的数量,如果隐藏层单元数量太少,样本中不同场景之间的 COE 相关性就不能充分展现出来;如果隐藏层单元数量太多,网络结构会表现出很大的随机性。

对于批处理大小,由于数据集中各个时间序列数据的长度不相同,因此在训练时需要填充较短的序列数据,以保障所有序列数据具有相同的长度。但是,过多的填充会影响神经网络的训练性能。为了避免这种情况的发生,我们先根据序列的长度对训练数据进行排序;然后将排序后的序列数据划分成不同的数据段,数据段的大小等于批处理的大小;最后,在训练时,对于进行批处理的数据段,将该数据段中的每个序列数据填充成长度相同的序列,它们的长度与它们中最长的序列的长度相同。

对于学习率,较大的学习率会加快训练过程但是会产生不收敛的现象,而较小的学习率会使训练过程变慢。因此,学习率会影响模型训练时的收敛效率和预测的准确度。

为了比较各个参数对 SLH, LT, UT 预测准确性的影响,本文进行了 LSTM 参数影响实验。令 $h \in \{10, 30, 50, 70, 90\}$, $b \in \{2, 4, 6, 8\}$, $a \in \{0.001, 0.005, 0.01, 0.015, 0.02, 0.025, 0.03, 0.035, 0.04, 0.045, 0.05\}$, 进行多次训练,记录每次实验的均方根误差 RMSE。

图 4 给出了 LSTM 参数对 SLH, LT, UT 预测准确性的影响。

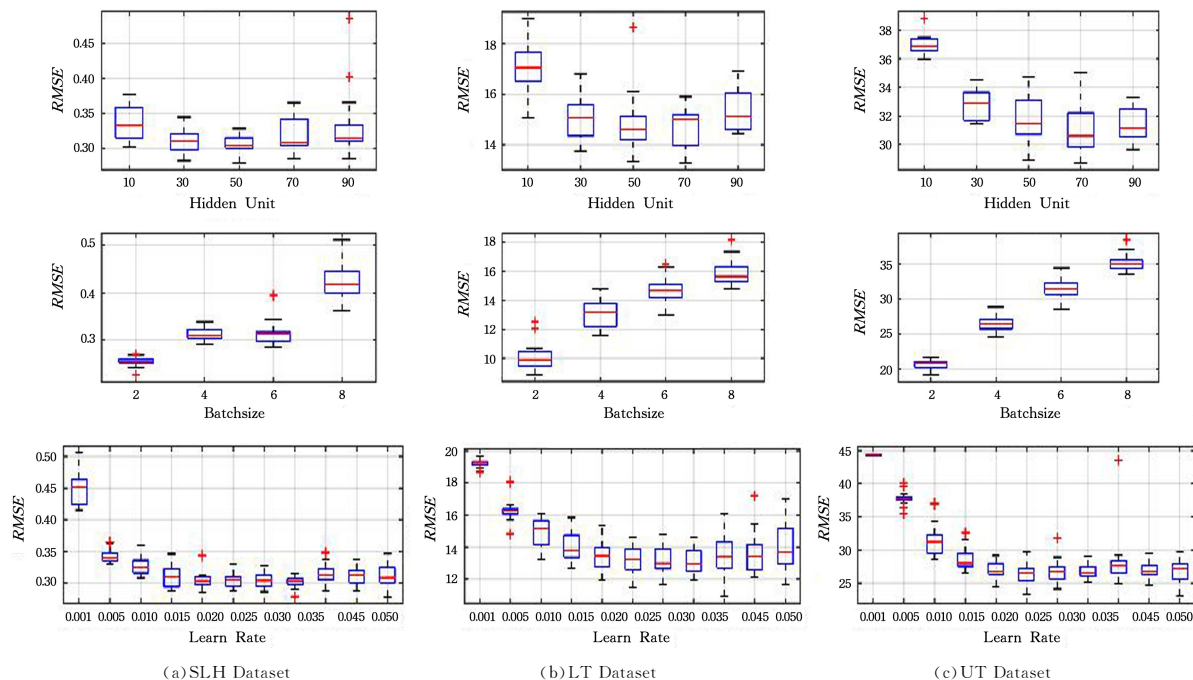


图4 LSTM 参数对火灾时间序列数据预测准确性的影响

Fig. 4 Influence of LSTM parameters on prediction accuracy of fire time series data

图 4 中,从上到下分别是隐藏层单元的数量、批处理大小、学习速率对火灾时间序列数据预测准确性的影响;从左到右分别是 SLH 的变化、 LT 的变化、 UT 的变化。

从图 4 的实验结果来看,随着隐藏层单元数量的增多, SLH, LT, UT 呈现先下降后上升的趋势。其中, $SLH, LT,$

UT 分别在 $h=50, 50, 70$ 时表现最好。随着批处理的增大, SLH, LT, UT 逐渐表现出上升趋势,即 $b=2$ 时, SLH, LT, UT 预测的准确度最高。这说明了批处理的大小(batch size)应该选择 2。随着学习率的提高, SLH, LT, UT 也呈现先下降后上升的趋势。 SLH, LT, UT 分别在 $a=0.02, 0.03,$

0.025 时取得最好的预测性能。

结束语 火灾数据分析与预测对于火灾的及时扑救和人员的及时疏散具有重要意义。本文分析了火灾发生时烟气流动与温度分布预测的需求,对现有的火灾预测模型及方法进行了介绍。在分析现有模型和方法存在不足的基础上,提出了一种基于趋势特征向量的火灾烟气流动与温度分布预测模型。本文提出的模型是一种改进的长短期记忆神经网络模型,它的改进在于利用趋势特征加速和优化神经网络的训练过程。实验结果表明,所提方法具有良好的预测性能,提高了火灾发生时烟气流动和温度分布预测的准确度和适用性。

在后续工作中,我们将主要研究以下问题:在不同场景的火灾时间序列数据中优化趋势特征向量的计算,使用动态损失函数改进 LSTM 神经网络的训练性能,使用 GAN 网络 (Generative Adversarial Network) 对火灾序列数据和 CFAST 的预测结果进行训练和预测。

参 考 文 献

- [1] SPYROU K J, KOROMILA J A. A risk model of passenger ship fire safety and its application[J]. *Reliability Engineering & System Safety*, 2020, 200: 1-14.
- [2] WANG W H, ZHU D H, PENG C H, et al. Uncertainty analysis for parameters of CFAST based on Electrical cabinet fire scenario in main control room[J]. *Nuclear Power Engineering*, 2018, 39(2): 153-156.
- [3] DING C X, PAN M Z, CHEN H, et al. An anionic polyelectrolyte hybrid for wood-polyethylene composites with high strength and fire safety via self-assembly[J]. *Construction and Building Materials*, 2020, 249: 1-12.
- [4] COWLARD A, JAHN W, ABECASSIS-EMPIS C, et al. Sensor assisted firefighting[J]. *Fire Technology*, 2010, 46(3): 719-741.
- [5] PEACOCK R D, RENEKE P A, FORNEY G P. CFAST—Consolidated Fire and Smoke Transport (Version 7)—Volume 3: Verification and Validation Guide [R]. National Institute of Standards and Technology, MD, USA, 2015.
- [6] PEACOCK R D, RENEKE P A, FORNEY G P. CFAST-Consolidated Model of Fire Growth and Smoke Transport (Version 7) Volume 2: User's Guide [R]. National Institute of Standards and Technology, MD, USA, 2017.
- [7] WANG W G, GUO Y, PENG C H. Uncertainty Analysis for Parameters of CFAST in the Main Control Room Fire Scenario [J]. *Atw-International Journal for Nuclear Power*, 2017, 62(7): 461-465.
- [8] HAN S, XIAO L, LI H Y, et al. Simulation research on safety evacuation in the underground commercial building fire [J]. *Fire Science and Technology*, 2018, 37(7): 910-914.
- [9] MAHJOUR S K, SANTOS A A S, Correia M G. Developing a workflow to select representative reservoir models combining distance-based clustering and data assimilation for decision making process [J]. *Journal of Petroleum Science and Engineering*, 2020, 190: 1-20.
- [10] CHANDRAMOULI P, MEMIN E, HEITZ D. 4D large scale variational data assimilation of a turbulent flow with a dynamics error model[J]. *Journal of Computational Physics*, 2020, 412: 1-29.
- [11] BURMAN E, FEIZMOHAMMADI A, OKSANEN L. A Finite Element Data Assimilation Method For The Wave Equation [J]. *Mathematics of Computation*, 2020, 89(324): 1681-1709.
- [12] LIN C C, WANG L. Forecasting smoke transport during compartment fires using a data assimilation model[J]. *Journal of Fire Science*, 2015, 33(1): 3-21.
- [13] LIN C C, ZHAO G C, WANG L Z L. Using real-time sensing data for predicting future state of building fires [C]// *Proceedings of the 2015 IEEE International Conference on Automation Science and Engineering (CASE)*. Gothenburg, Sweden, 2015.
- [14] JI J, TONG Q, WANG L L, et al. Application of the EnKF method for real-time forecasting of smoke movement during tunnel fires [J]. *Adv. Eng. Softw.* 2018, 115: 398-412.
- [15] LI X, ZHANG W, XU N X. Deep Learning-Based Machinery Fault Diagnostics With Domain Adaptation Across Sensors at Different Places [J]. *IEEE Transactions on Industrial Electronics*, 2020, 67(8): 6785-6794.
- [16] ASSAAD M, BONÉ R, CARDOT H. A new boosting algorithm for improved time-series forecasting with recurrent neural networks [J]. *Information Fusion*, 2008, 9(1): 41-55.
- [17] YU C Y, QI X, Ma H. LLR: Learning learning rates by LSTM for training neural networks[J]. *Neurocomputing*, 2020, 394: 41-50.
- [18] GRAVES A, NAVDEEP J. Towards end-to-end speech recognition with recurrent neural networks [C]// *Proceedings of the 31st International Conference on Machine Learning*. Beijing, China, 2014: 1764-1772.
- [19] CAO J, LI Z, LI J. Financial time series forecasting model based on CEEMDAN and LSTM [J]. *Physica A*, 2019, 519: 127-139.
- [20] VAN HOUUDT G, MOSQUERA C, NAPOLES G. A review on the long short-term memory model [J]. *Artificial Intelligence Review*, 2020, 53: 5929-5955.
- [21] HSIEH P-H, WU O, GEUE C. Economic burden of rheumatoid arthritis: a systematic review of literature in biologic era [J]. *Annals of the Rheumatic Diseases*, 2020, 79(6): 771-777.
- [22] SHIKWAMBANA L, KGANYAGO M. Trends in atmospheric pollutants from oil refinery processes: a case study over the United Arab Emirates [J]. *Remote Sensing Letters*, 2020, 11(6): 590-597.
- [23] HU C, BAI Y, LI J, et al. Prognostic value of systemic inflammatory factors NLR, LMR, PLR and LDH in penile cancer [J]. *BMC Urology*, 2020, 20(1): 1-9.
- [24] FATHIAN F, DEHGHAN Z, BAZR KAR M H, et al. Trends in hydrological and climatic variables affected by four variations of the Mann-Kendall approach in Urmia Lake basin [J]. *Iran. Hydrolog. Sci. J.* 2016, 61(5): 892-904.
- [25] NOURANI V, MEHR A D, AZAD N. Trend analysis of hydro-climatological variables in Urmia lake basin using hybrid wavelet Mann-Kendall and Şen tests [J]. *Environ Earth Science*, 2018, 77: 1-18.

- [26] ARAGHI A, MOUSAVI-BAYGI M, ADAMOWSKI J. Detecting soil temperature trends in Northeast Iran from 1993 to 2016 [J]. *Soil. Till. Res.* 2017, 174:177-192.
- [27] NI X L, XU M, CAO C X, et al. Forest height estimation and change monitoring based on artificial neural network using Geoscience Laser Altimeter System and Landsat data [J]. *Journal of Applied Remote Sensing*, 2020, 14(2):1-9.
- [28] JAIN S J. Numerical simulation of fire in a tunnel: Comparative study of CFAST and CFX predictions [J]. *Tunnelling Underground Space Technology*, 2018, 23(2):160-170.
- [29] SHEPPARD D T, KLEIN B W. Burn tests in two story structure with hallways [R]. Ammdale, Maryland: ATF Laboratories, 2009.
- [30] NOWLEN S P. Enclosure Environment Characterization Testing for the Base Line Validation of Computer Fire Simulation Codes [R]. Albuquerque, NM: Sandia National Laboratories, 1987.
- [31] CHAVEZ J M, NOWLEN S P. An Experimental Investigation of Internally Ignited Fires in Nuclear Power Plant Control Cabinets: Part 2, Room Effects Tests [R]. Albuquerque, NM: Sandia National Laboratories, 1988.
- [32] HAMINS A P, MARANGHIDES A, JOHNSON E L, et al. Report of experimental results for the international fire model benchmarking and validation exercise 3 [R]. MD, USA: National Institute of Standards and Technology, 2003.
- [33] HAMINS A P, MARANGHIDES A, MCGRATTAN K B, et al. Experiments and Modeling of Structural Steel Elements Exposed to Fire. Federal Building and Fire Safety Investigation of the World Trade Center Disaster (NIST NCSTAR 1-5B) [R]. National Institute of Standards and Technology, MD, USA, 2005.



YIN Yun-fei, born in 1976, Ph.D, associate professor, is a member of China Computer Federation. His main research interests include machine learning and IoT engineering.