

面向在线客服系统的服务量预测算法^①



周子馨

(河海大学 计算机与信息学院, 南京 211100)
通讯作者: 周子馨, E-mail: Ada_Chow@189.cn

摘要: 在线客服系统的服务量受到多种因素影响, 为提高系统的服务量预测精度, 本文基于离散变异控制参数的自适应差分进化算法 DMPSADE, 提出了一种改进算法 IDMPSADE, 并将其与长短时记忆神经网络 LSTM 相结合建立了对服务量的预测模型 IDMPSADE-LSTM. 在 IDMPSADE 中, 当子代种群测试函数寻优性能没有父代种群好时, 对父代种群个体进行反向引导, 跳出局部最优, 提升搜索到全局最优能力. 由于 LSTM 的神经元数量、迭代次数、学习率以及训练批次需要通过经验进行设置, 具有较大的随机性, 故利用 IDMPSADE 对这些参数进行寻优. IDMPSADE-LSTM 将分析得到的气温、降水量作为影响因素结合服务量的时间特征对系统的服务量进行预测. 文中实验结果表明, IDMPSADE-LSTM 预测模型比一般的神经网络以及 SARIMA-SVM 混合预测模型的精确度要高.
关键词: 在线客服系统服务量; 时间序列预测; 差分进化算法; 长短时记忆神经网络

引用格式: 周子馨. 面向在线客服系统的服务量预测算法. 计算机系统应用, 2020, 29(4): 137-143. <http://www.c-s-a.org.cn/1003-3254/7337.html>

Service Volume Prediction Algorithm for Online Customer Service System

ZHOU Zi-Xin

(College of Computer and Information, Hohai University, Nanjing 211100, China)

Abstract: The service volume of online customer service system is affected by many factors. In order to improve the prediction accuracy of service volume, an improved algorithm IDMPSADE is proposed on the basis of self-adaptive differential evolution algorithm DMPSADE with discrete mutation control parameters. By combining IDMPSADE with Long-Short Term Memory network (LSTM), an IDMPSADE-LSTM prediction model of service volume is established. IDMPSADE chooses the reverse guidance of the parent population whose child population's performance on test functions is not as good as it, which can escape from the local optimum and improve the capability of searching the optimal solution within defined space. LSTM's parameters, such as number of neurons, epochs, learning rate, and batch-size, are set by experience and have larger randomness, and IDMPSADE could be helpful to optimize these parameters. IDMPSADE-LSTM prediction model uses temperature and precipitation as influencing factors and combines with the temporal characteristics of service volume to predict the service volume. The experimental results show that the proposed IDMPSADE-LSTM prediction model is more accurate compared with general neural networks and SARIMA-SVM hybrid prediction model.

Key words: service volume of online customer service system; time series prediction; differential evolution algorithm; Long-Short Term Memory (LSTM)

① 基金项目: 国网江苏省电力有限公司科技项目 (J2018020)

Foundation item: Science and Technology Project of State Grid Jiangsu Electric Power Co. Ltd. (J2018020)

收稿时间: 2019-08-22; 修改时间: 2019-09-09, 2019-09-19; 采用时间: 2019-09-23; csa 在线出版时间: 2020-04-05

随着客服中心服务量的不断增加,如何保证在高峰时段和突发情况下有足够的客服人员在服务,已成为一个越来越紧迫的问题.实际上,在线客服系统是不能够完全自动化完成客服中心的复杂调度和计划任务.人的交互作用是客服中心必不可少的,因此如何对客服中心的的服务量进行预测,对于保证客服中心提供的服务能够使客户获得较高的满意度是非常重要的^[1].

目前,对于服务量的预测方法已有多种,如时间序列、排队论、回归预测和神经网络等. Andrews等^[2]和 Antipov等^[3]提出了一种具有解释变量的自回归综合移动平均 ARIMAX 模型,该模型使用如促销周期、广告响应和日历效应等外部变量. Aldor-Noiman等^[4]提出了一种基于混合泊松过程方法的统计模型,该模型考虑事件的影响,例如将账单待出时期作为外部变量. Niu等^[5]提出了一种特殊的回波状态网络,在该网络中每个时间序列具有不同的存储层,并且该网络将外部变量也考虑其中. ARIMAX 模型和基于混合泊松过程方法的统计模型属于参数方法,虽然可以明确各变量之间的数量关系,定量分析变量对因变量的影响大小,但是其方法内部参数的设定受使用者的主观影响较大,不同的使用者很可能会得到相差很大的结果^[6];回波状态网络属于非参数方法,依靠数据来调节模型内部的参数,进而得到变量与因变量之间的关系,但是对于长时间序列,其预测能力有限.因此长短时记忆神经网络(LSTM)被提出,它是一种改进的时间递归神经网络(RNN),可以学习时间序列长短期信息.由于其包含时间记忆单元,因此适合处理和预测时间序列中的间隔和延迟事件.但是 LSTM 仍然有 2 个缺点:一是隐含层数和隐含层神经元数难确定;二是学习率和迭代次数难确定^[6].隐含层数和隐含层神经元数直接决定模型的拟合能力,学习率和迭代次数影响模型的训练过程和效果.在实际应用中,这些参数都是依靠经验来确定,具有较大的随机性,这降低了模型的预测效果.

本文主要贡献和创新如下:

(1) 提出了基于算法 DMPSADE 的一种改进算法 IDMPSADE,即通过反向引导寻优效果不佳的父代种群变异方向,提升搜索全局最优能力.

(2) 建立了 IDMPSADE-LSTM 预测模型,并且分析了在线客服系统的系统历史服务量数据,挖掘出关键影响因素,将其作为 LSTM 的多输入变量,利用 IDMPSADE 对 LSTM 的部分超参数进行寻优,从而提

升了 IDMPSADE-LSTM 预测模型的精确度.

本文内容安排如下:第 1 节主要介绍已有的服务量预测方法.第 2 节针对各预测方法存在的不足,引入构建本文预测模型所需的背景知识.第 3 节介绍改进的算法 IDMPSADE,并与多种其它差分进化算法进行实验对比.第 4 节构建 IDMPSADE-LSTM 预测算法模型.第 5 节通过分析历史数据,选取服务量影响因素作为 LSTM 的多输入变量,利用 IDMPSADE-LSTM 预测模型对服务量进行预测,并且与其他神经网络以及混合预测模型进行实验对比分析.第 6 节对 IDMPSADE-LSTM 预测模型进行总结.

1 相关知识

1.1 长短时记忆神经网络 LSTM 原理

对于标准循环神经网络结构而言,存在一个问题:一般神经网络的信息从输入层到隐含层到输出层单向流动,循环神经网络的信息传递存在定向的循环,很容易出现梯度消失和梯度爆炸^[7-9].LSTM 结构由一组循环连接的子网(称为记忆单元)组成.每个记忆单元包含一个或多个自连接的记忆细胞和三种门(输入门、输出门和遗忘门),这三种门允许记忆细胞存储访问长时间段的信息,因此能够缓解梯度消失的问题^[10].LSTM 中由单个记忆细胞构成的记忆单元如图 1 所示.

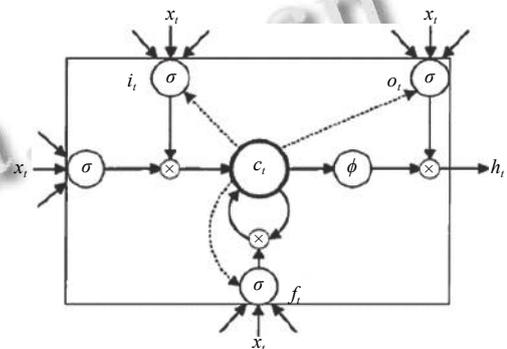


图 1 LSTM 单细胞记忆单元示意图

某一时刻 t 的状态计算过程如下:

$$g_t = \phi(W_{xg} \cdot [x_t, h_{t-1}] + b_g) \quad (1)$$

$$i_t = \sigma(W_{xi} \cdot [x_t, h_{t-1}] + b_i) \quad (2)$$

$$f_t = \sigma(W_{xf} \cdot [x_t, h_{t-1}] + b_f) \quad (3)$$

$$o_t = \sigma(W_{xo} \cdot [x_t, h_{t-1}] + b_o) \quad (4)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot g_t \quad (5)$$

$$h_t = o_t \odot c_t \quad (6)$$

其中, ϕ 和 σ 分别为双曲正切(tanh)和sigmoid函数, W 和 b 为待学习的网络参数. i_t, f_t 和 o_t 分别为 t 时刻的输入门(input gate)、遗忘门(forget gate)和输出门(output gate), c_t 为 t 时刻记忆单元的内部状态. c_t 能够在 t 时刻保持模型对历史输入信息的编码记忆, 输入门控制信息流入, 输出门控制信息流出, 而遗忘门则控制历史信息对当前记忆单元 c_t 的影响大小. h_t 和 h_{t-1} 分别为记忆单元在 $t-1$ 和 t 时刻的输出. 由于能够较好地对比长程和短程依赖关系进行建模, 长短时记忆模型被广泛运用到序列编码中^[11].

1.2 差分进化算法原理

1.2.1 差分进化算法概述

差分进化算法是一种针对种群演化的计算技术, 初衷是为了解决切比雪夫多项式问题, 在发展过程中, 研究人员发现它也可以用于解决复杂优化问题. 差分进化算法保留了基于种群的全局搜索策略, 采用实数编码, 变异操作采用个体间的差值与变异率相乘, 并且使用一对一的竞争生存策略, 从而降低简化遗传操作. 此外, 因为差分进化算法具备独特的记忆能力, 所以它可以根据跟踪到的动态搜索情况调整搜索策略, 全局收敛能力和健壮性相较其他进化算法优越. 虽然数学方法能够对很多抽象问题具象逻辑求解, 但是自然界中依旧存在大量利用常规的数学归纳方法无法解决的优化问题, 差分进化算法对此类问题无需其特征信息就能求解. 因此, 差分进化算法因其拥有的高效搜索能力对于学术研究和工程类应用而言存在一定的意义与价值^[12].

1.2.2 DMPSADE 算法流程

差分进化算法的搜索能力是由被选择的变异策略和相关的控制参数决定, 所以选择适当的变异策略和控制参数对差分进化算法性能的改善是至关重要的, 基于离散变异控制参数的自适应差分进化算法(DMPSADE)核心是实现控制参数和变异策略的自适应, 同时确保每个变量的控制参数能够进行独立演化^[13].

(1) 初始化

NP 为种群中的个体数量, G_{\max} 为演化过程中的最大迭代数, x 为生成的种群个体, 个体的变化范围为 $[x_{\min}, x_{\max}]$, D 为问题维度. 每一种变异策略的初始个体数量为 $NP/5$, 种群初始化方式如式(7)所示.

$$x = x_{\min} + (x_{\max} - x_{\min}) \times \text{rand}(NP, D) \quad (7)$$

(2) 种群进化

$Setp$ 为决定单个自适应变异策略使用量的常量, G_S 为判断种群个体选用何种变异策略的条件, 令 $G_S = Setp \times G_{\max}$.

变异操作: 对于每个个体 $x_i^G, i = 1, 2, \dots, NP$ 来说, 当 $G < G_S$ 时, 变异策略 $DE/rand/1$ 为第 $G+1$ 代产生候选种群, 如果 $G \geq G_S$, 新生试验向量 v_i^{G+1} 则通过从变异策略池中随机选择出来的变异策略产生.

边界操作: 如果 $v_{ij}^{G+1} < x_j^{\text{low}}$ 或者 $v_{ij}^{G+1} > x_j^{\text{high}}$, 则令 $v_{ij}^{G+1} = x_{r1j}^G$. 其中 $x_j^{\text{low}}, x_j^{\text{high}}$ 分别对应第 j 个个体变量的下界和上界.

交叉操作: 如式(8)所示, 其中 R_j 是 $[0, 1]$ 中的均匀随机数; CR 为交叉控制参数.

$$u_{ij}^{G+1} = \begin{cases} v_{ij}^{G+1}, R_j \leq CR \\ x_{ij}^G, R_j > CR \end{cases}, j = 1, 2, \dots, D \quad (8)$$

选择操作: 如式(9)所示.

$$x_i^{G+1} = \begin{cases} u_i^{G+1}, f(u_i^{G+1}) \leq f(x_i^{G+1}) \\ x_i^G, f(u_i^{G+1}) > f(x_i^{G+1}) \end{cases}, i = 1, 2, \dots, NP \quad (9)$$

(3) 变异策略自适应

如果 $G = G_S$, 5种变异策略 $DE/rand/1$ 、 $DE/rand/2$ 、 $DE/best/2$ 、 $DE/rand-to-best/1$ 和 $DE/rand-to-best/2$ 会被随机分配给种群中所有的个体.

如果 $G > G_S$, 计算出每一种变异策略的累积概率, 并采用轮盘法决定第 $G+1$ 代种群个体的变异策略.

(4) 控制参数自适应

由于迭代数 G 、上一代变异控制参数值与交叉控制参数值对下一代变异控制参数和交叉控制参数的更新有一定影响, 故以此为判断条件, 通过计算每个个体的权值, 获取每个个体的变异控制参数加权平均值与种群的交叉控制参数加权平均值, 利用正态分布函数对控制参数进行更新.

(5) 迭代

种群迭代一次, 即 $G = G + 1$.

(6) 评估

重复步骤(2)~(5)直到迭代数值达到最大值或评估函数收敛精度达到要求.

2 IDMPSADE 算法

DMPSADE 算法主要思想是种群在经过初步的初

始化、种群进化之后,进行变异策略和控制参数的自适应行为,然后选择评估值小的种群作为最佳种群.IDMPSADE以DMPSADE算法思想为基础,初始化种群时均采用正向引导变异,当正向引导的效果不佳即子代种群适应度函数大于父代种群适应度函数时,此处的适应度函数为算法的寻优函数,对父代种群个体进行反向引导,即更改引导方向,选择反向引导更为直观,从而提高跳出局部最优和搜索到全局最优的可能性.反向引导体现在改进策略公式的 F 参数符号中,具体如式(10),式(11)至式(14)所示.

DE/rand/1 策略:

$$v_i^{G+1} = x_{r_1}^G - F * (x_{r_2}^G - x_{r_3}^G) \quad (10)$$

DE/rand/2 策略:

$$v_i^{G+1} = x_{r_1}^G - F * (x_{r_2}^G - x_{r_3}^G) - F * (x_{r_4}^G - x_{r_5}^G) \quad (11)$$

DE/best/2 策略:

$$v_i^{G+1} = x_{best}^G - F * (x_{r_1}^G - x_{r_2}^G) - F * (x_{r_3}^G - x_{r_4}^G) \quad (12)$$

DE/current-to-best/1 策略:

$$v_i^{G+1} = x_i^G - F * (x_{best}^G - x_i^G) - F * (x_{r_1}^G - x_{r_2}^G) \quad (13)$$

DE/current-to-best/2 策略:

$$v_i^{G+1} = x_i^G - F * (x_{best}^G - x_i^G) - F * (x_{r_1}^G - x_{r_2}^G + x_{r_3}^G - x_{r_4}^G) \quad (14)$$

IDMPSADE算法的控制参数包括变异控制参数与交叉控制参数,参数的定义与DMPSADE算法的一致,公式如式(15)至式(17)所示.

$$w_i^G = \frac{|f(u_i^{G+1}) - f_{max}|}{\sum_{i=1}^{NP} |f(u_i^{G+1}) - f_{max}|} \quad (15)$$

$$F_{w_j}^{G+1} = \sum_{i=1}^{NP} (w_i^G \times F_{ij}^G) \quad (16)$$

$$CR_w^{G+1} = \sum_{i=1}^{NP} (w_i^G \times CR_i^G) \quad (17)$$

当更新条件为 $G < 0.5 \times G_{max}$ 并且个体对应的变异控制参数 F_{ij}^{G+1} 大于1,或者 F_{ij}^{G+1} 小于 $0.1 \times (0.1 - (G/G_{max})^2)$ 时,变异控制参数需要按照 $N(F_{w_j}^{G+1}, \sigma_1)$ 进行更新;若更新条件为 $G > 0.5 \times G_{max}$ 并且 F_{ij}^{G+1} 大于1,则将变异控制参数 F_{ij}^{G+1} 置为1;若更新条件为 $F_{ij}^{G+1} < 0$,则将变异控制参数 F_{ij}^{G+1} 按照 $|N(0, 0.1 \times (1 - (G/G_{max})^2))|$ 进行更新.

交叉控制参数 CR_i^{G+1} 的更新条件与变异控制参数的一致,在各条件下对应的更新式分别为 $N(CR_w^{G+1}, \sigma_2)$, $1, |N(0, 0.1 \times (1 - (G/G_{max})^2))|$.其中, $f(u_i^{G+1})$ 为个体适应度函数,即目标函数, f_{max} 为种群最优适应度函数.由于 CR 的值越大,种群个体信息量交换的程度越大.反之,如果 CR 的值偏小,将会使种群的多样性快速减小,不利于全局寻优; F 主要影响算法的全局寻优能力, F 越小,算法对局部的搜索能力更好, F 越大算法越能跳出局部极小点,但是收敛速度会变慢.此外, F 还影响种群的多样性^[14],故令 $\sigma_1 = \sigma_2 = 0.8 - 0.5 \times (1 - (G/G_{max})^2)$,图2为 σ 的变化曲线,在寻优过程中以保证更新后的 CR 与 F 均在0.5左右分布,从而平衡算法的寻优能力与收敛速度.

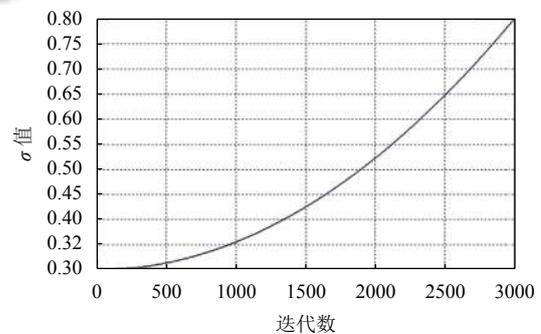


图2 σ 变化曲线

由于IDMPSADE算法的控制参数更新方式与DMPSADE算法相同,故根据已有的实验分析^[13]可知,控制参数具有极大的随机性,每个种群个体有其变异控制参数与交叉控制参数,适应性在不同的测试函数上表现也不同.

为了测试IDMPSADE的寻优性能,除选择DMPSADE算法进行实验对比,还选择了jDE^[15]和SaDE^[16]在4种常用测试函数上进行测试,这4种测试函数具有不同的特性,可抽象地表示实际生活的不同问题.测试函数分别为F1-Shifted Schwefel's Problem 1.2 with Noise in Fitness, F2-Shifted Rosenbrock's Function, F3-Non-Continuous Rotated Hybrid Composition Function, F4-Rotated Hybrid Composition Function without bound,其中F1为单峰函数,F2为多峰函数且极值点个数较多,F3、F4为组合测试函数^[17].

设置参数 $NP=100, G_{max}=3000, D=30, Setp=0.175, Msp=0.02$.评估性能的指标为测试函数均值与标准差,均值与标准差越小,说明算法的收敛精度越高,寻优性

能越好. 由表 1 知, IDMPsADE 的平均值和标准差均比 jDE, SaDE 和 DMPSADE 的小, 说明 IDMPsADE 算法具有较好的性能.

表 1 各差分进化算法标结果对比

函数	算法	平均值	标准差
F1	jDE	2.45E-02	1.69E-01
	SaDE	9.56E+01	1.25E+01
	DMPSADE	2.87E+02	1.80E+02
	IDMPsADE	3.02E-03	0.02E-03
F2	jDE	2.23E+01	2.36E+01
	SaDE	4.34E+01	3.23E+01
	DMPSADE	6.23E-01	1.60E+00
	IDMPsADE	1.20E-04	5.74E-04
F3	jDE	8.27E+02	2.09E+02
	SaDE	6.90E+02	2.87E+02
	DMPSADE	5.39E+02	4.17E-05
	IDMPsADE	5.34E+02	1.72E-05
F4	jDE	2.17E+02	1.46E+00
	SaDE	2.23E+02	3.77E+00
	DMPSADE	2.20E+02	1.96E+00
	IDMPsADE	2.11E+02	0.84E+00

3 IDMPsADE-LSTM 预测模型

参数的选取对 LSTM 的预测效果存在较大影响. 采用 IDMPsADE 算法对 LSTM 参数进行优化, 需要进行寻优操作的主要参数有第一层隐藏层的神经元数 h_1 , 第二层隐藏层的神经元数 h_2 , 学习率 r , 训练迭代次数 e , 训练批次 b , 激活函数设置为 tanh 函数, 由于是预测回归问题, 故全连接层输出维度设置为 1.

Step 1. 初始化. 将 LSTM 模型中的控制参数数目作为 IDMPsADE 算法的问题维度, 即 D 的取值; 将在线客服系统的服务量划分为训练 LSTM 的训练数据与预测数据. 选取 LSTM 的拟合值与真实值之间的 RMSE 作为目标函数. 初始化 IDMPsADE 算法的基本参数包括种群规模、最大迭代次数、变异参数和交叉参数. IDMPsADE 模型的种群按照式 (7) 产生, 种群中的每个个体均是 LSTM 模型的一个参数组合.

Step 2. 种群进化. 首先判断是否达到迭代结束条件, 即全局极小值是否达到了设定的精度或者是否达到了设定的最大迭代次数. 如果是, 则停止迭代输出最优个体; 否则, 种群继续进行变异、边界和交叉操作得到下一代种群个体 x_i^{G+1} , 其中种群的迭代次数以及种群个体目标函数值的分布差异性决定了种群对五种变异策略的选择概率, 即变异策略的自适应, 转下一步.

Step 3. 控制参数自适应. 由于迭代数 G 、上一代

变异控制参数值与交叉控制参数值对下一代变异控制参数和交叉控制参数的更新有一定影响, 根据种群个体目标函数值的分布差异性以及当前的迭代次数, 更新变异控制参数 F 与交叉控制参数 CR , 转下一步.

Step 4. 更改变异方向. 如果下一代种群个体 x_i^{G+1} 的目标函数值比上一代种群个体 x_i^G 的大, 则保留上一代种群个体 x_i^G , 并更改其在 Step 2 中的变异方向, 然后继续 Step 2 与 Step 3 的步骤. 否则, 进行 Step 5 操作.

Step 5. 迭代. 种群迭代一次, 即 $G = G + 1$, 转 Step 2.

Step 6. 得到 LSTM 最优参数组合以及在线客服系统服务量的预测结果.

4 实验结果分析

4.1 服务量影响因素选取

服务量数据来自于某电力客服中心从 2016 年 1 月至 2018 年 12 月的数据中心记录, 包括每天各时间点的服务记录. 影响服务量的因素有很多. 由于在冬夏两季受温度影响, 居民用电量较多, 可能出现跳闸断电、电费相较过去有所增加等情况, 导致系统服务量变化较大. 雨雪天气影响居民的外出, 居民用电量较多, 可能出现电力设备故障、电费相较过去有所增加等情况, 导致系统服务量变化较大. 所以选取温度和降水量为主要影响因素.

在将温度和降水量作为 LSTM 的输入前, 需要对其进行相关性分析. 分析方法利用 Pearson 相关性分析和 Spearman 秩相关性分析^[18]. 首先将近一年内的在线服务量按照工作日与非工作日划分, 同时对异常在线服务量进行数据清洗, 分别将温度、降水量作为第一变量, 将近一年的工作日在线服务量作为第二变量, 对第一、第二变量进行相关系数计算, 当 Pearson 相关度和 Spearman 相关度均大于 0.5 时, 说明影响因素与服务量的相关性强. 影响因素相关性分析结果如表 2 所示.

由计算结果得出, 温度和降水量这两类气象影响因素与在线服务量的相关性强, 并且温度和服务量的相关性高于降水量和服务量的相关性, 温度和降水量可作为 LSTM 预测系统服务量的输入.

表 2 影响因素相关性分析

影响因素	Pearson 相关度	Spearman 相关度
温度	0.661 169	0.611 445
降水量	0.618 556	0.601 73

4.2 服务量预测结果分析

(1) IDMPADE 参数与 LSTM 参数设置

种群规模 NP 为 100, 最大迭代次数 G_{max} 为 3000, 目标函数维度 D 为 5, 决定单个自适应变异策略使用量的参数 $Setp$ 为 0.175, 5 种变异策略的可选概率最大变化量 Msp 为 0.02, 由于是自适应变异控制参数与交叉控制参数, 因此根据种群个体目标函数值的分布差异性以及迭代次数, 种群个体的更新变异控制参数 F 与交叉控制参数 CR 的加权平均值如图 2 和图 3 所示. 经过 IDMPADE 算法的优化, LSTM 的第一层隐藏层神经元数为 40, 第二层隐藏层神经元数为 30, 学习率为 0.01, 训练迭代次数为 10, 训练批次为 72.

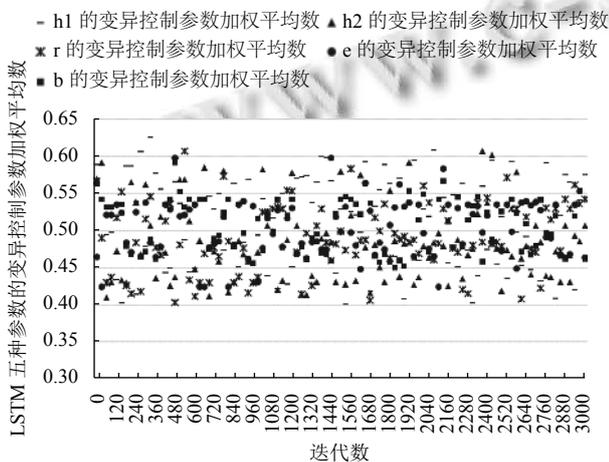


图3 LSTM 五种参数的变异控制参数 F 加权平均值

如图 3 与图 4 所示, 变异控制参数 F 的加权平均值在 0.5 上下浮动, 范围在 0.4 到 0.65 之间; 交叉控制参数 CR 的加权平均值在 0.5 上下浮动, 范围在 0.4 到 0.6 之间, 有效地平衡了算法的收敛精度与收敛速度.

(2) 服务量预测精确度分析

基于统一测试数据集, 利用经过 IDMPADE 算法优化后的 LSTM 神经网络对在线客服系统的服务量进行预测, 预测出代表 336 个小时内服务量的 336 个样本点数与原实际数据的对比如图 4 所示. 算法的预测性能指标选取均方误差 MSE 和平均绝对误差 MAE , 如式 (18) 和式 (19) 所示.

$$MSE = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n} \quad (18)$$

$$MAE = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n} \quad (19)$$

其中, y_i 表示真实值, \hat{y}_i 表示预测值, n 表示测试样本数, MSE 与 MAE 的值越小, 表示误差越小, 预测性能越好.

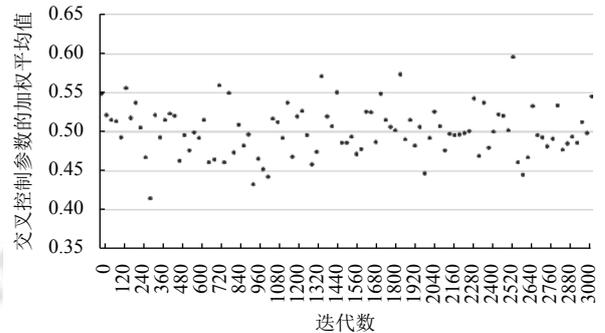


图4 交叉控制参数 CR 加权平均值

采用传统的 BP 神经网络、RBF 神经网络、SARIMA-SVM 模型预测效果与 IDMPADE-LSTM 模型预测效果对比如表 3 所示.

表3 不同模型预测结果对比

模型	MSE	MAE
BP	0.056	0.076
RBF	0.048	0.069
SARIMA-SVM	0.041	0.055
IDMPADE-LSTM	0.021	0.035

由图 5 可以看出, 在线客服系统的服务量具备一定的周期性, 预测结果相较实际数据的极值范围更小, 能够较好地预测出服务量的总体变化趋势, 有助于客服中心对客服提前排班以应对服务量变化情况.

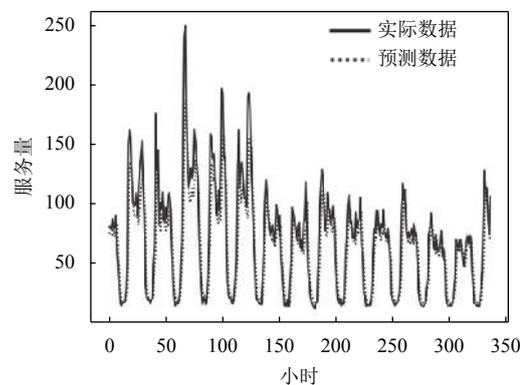


图5 IDMPADE-LSTM 服务量预测结果

由表 3 可以看出, IDMPADE-LSTM 混合预测模型的预测误差小于 BP 神经网络、RBF 神经网络以及

SARIMA-SVM 混合预测模型, 说明该服务量预测模型具有一定优越性, 预测精度有所提升.

5 总结

系统服务量预测的准确性、有效性、实时性是公司为客服人员安排技能培训、进行工作规划的重要依据, 应用面向在线客服系统的服务量预测算法可为人力资源调度优化等工作提供科学的依据.

本文在对系统服务量数据调查分析的基础上, 提出了一种 IDMPSADE-LSTM 服务量预测模型, 与 BP 神经网络、RBF 神经网络以及 SARIMA-SVM 混合预测模型相比, 本文提出的预测模型将关键影响因素与服务量自身的时间序列特征进行了有效的结合, 并且预测的精确度更高. 此外, 本文提出的 IDMPSADE-LSTM 预测模型解决了神经网络部分超参数的赋值随机性问题, 该模型除了可应用于系统服务量预测, 对于交通、金融等领域的相关数据预测也具备一定的参考价值.

参考文献

- Zhang J, Tian WF, Cui CY. Traffic forecasting of mobile communication in call center. Proceedings of the 2010 3rd International Conference on Computer Science and Information Technology. Chengdu, China. 2010. 160–164.
- Andrews BH, Cunningham SM. L. L. Bean improves call-center forecasting. Interfaces, 1995, 25(6): 1–13. [doi: 10.1287/inte.25.6.1]
- Antipov A, Meade N. Forecasting call frequency at a financial services call centre. Journal of the Operational Research Society, 2002, 53(9): 953–960. [doi: 10.1057/palgrave.jors.2601415]
- Aldor-Noiman S, Feigin PD, Mandelbaum A. Workload forecasting for a call center: Methodology and a case study. The Annals of Applied Statistics, 2009, 3(4): 1403–1447. [doi: 10.1214/09-AOAS255]
- Niu DX, Ji L, Xing M, et al. Multi-variable echo state network optimized by Bayesian regulation for daily peak load forecasting. Journal of Networks, 2012, 7(11): 1790–1795.
- 李万, 冯芬玲, 蒋琦玮. 改进粒子群算法优化 LSTM 神经网络的铁路客运量预测. 铁道科学与工程学报, 2018, 15(12): 3274–3280.
- Hochreiter S. Untersuchungen zu dynamischen neuronalen netzen[Master's thesis]. Munchen: Technische Universität München, 1991.
- Kolen JF, Kremer SC. Gradient flow in recurrent nets: The difficulty of learning longterm dependencies. In: Kolen JF, Kremer SC, eds. A Field Guide to Dynamical Recurrent Networks. New York: Wiley-IEEE Press, 2001. 237–243.
- Bengio Y, Simard P, Frasconi P. Learning long-term dependencies with gradient descent is difficult. IEEE Transactions on Neural Networks, 1994, 5(2): 157–166. [doi: 10.1109/72.279181]
- Hochreiter S, Schmidhuber J. Long short-term memory. Neural Computation, 1997, 9(8): 1735–1780. [doi: 10.1162/neco.1997.9.8.1735]
- 孔德江. 面向序列数据的深度学习算法研究[博士学位论文]. 杭州: 浙江大学, 2019.
- 刘波, 王凌, 金以慧. 差分进化算法研究进展. 控制与决策, 2007, 22(7): 721–729. [doi: 10.3321/j.issn:1001-0920.2007.07.001]
- Fan QQ, Yan XF. Self-adaptive differential evolution algorithm with discrete mutation control parameters. Expert Systems with Applications, 2015, 42(3): 1551–1572. [doi: 10.1016/j.eswa.2014.09.046]
- 呼忠权. 差分进化算法的优化及其应用研究[硕士学位论文]. 秦皇岛: 燕山大学, 2013.
- Brest J, Greiner S, Boskovic B, et al. Self-adapting control parameters in differential evolution: A comparative study on numerical benchmark problems. IEEE Transactions on Evolutionary Computation, 2006, 10(6): 646–657. [doi: 10.1109/TEVC.2006.872133]
- Qin AK, Huang VL, Suganthan PN. Differential evolution algorithm with strategy adaptation for global numerical optimization. IEEE Transactions on Evolutionary Computation, 2009, 13(2): 398–417. [doi: 10.1109/TEVC.2008.927706]
- Suganthan PN, Hansen N, Liang JJ, et al. Problem definitions and evaluation criteria for the CEC 2005 special session on real-parameter optimization. Singapore: Nanyang Technological University, 2005.
- Xiao CW, Ye JQ, Esteves RM, et al. Using spearman's correlation coefficients for exploratory data analysis on big dataset. Concurrency and Computation: Practice and Experience, 2016, 28(14): 3866–3878. [doi: 10.1002/cpe.3745]