

动态数据流分析的在线超限学习算法综述

郭 威¹ 于建江¹ 汤克明¹ 徐 涛²

(盐城师范学院信息工程学院 江苏 盐城 224002)¹

(南京航空航天大学计算机科学与技术学院 南京 210016)²

摘 要 动态数据流分析是一个具有广泛应用价值的研究课题,在线学习方法是其中的一种关键技术。在众多在线学习方法中,在线贯序超限学习机(Online Sequential Extreme Learning Machine, OSELM)是一种新颖且实用的在线学习算法,目前已在动态数据流分析中得到了成功应用。首先,介绍了 OSELM 的理论基础和算法执行过程;然后,以动态数据流分析为应用背景,对各种改进 OSELM 算法进行了分类综述,包括基于滑动窗口的 OSELM、基于遗忘因子的 OSELM、基于样本加权的 OSELM 以及其他方法,重点论述了各类算法的设计思路和实现策略,并对其优缺点进行了比较和分析;最后,探讨了值得进一步研究的问题。

关键词 在线贯序超限学习机,动态数据流分析,滑动窗口,遗忘因子,样本加权

中图分类号 TP181 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2019.04.001

Survey of Online Sequential Extreme Learning Algorithms for Dynamic Data Stream Analysis

GUO Wei¹ YU Jian-jiang¹ TANG Ke-ming¹ XU Tao²

(College of Information Engineering, Yancheng Teachers University, Yancheng, Jiangsu 224002, China)¹

(College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China)²

Abstract Dynamic data stream analysis has become a research focus for its widespread application prospects, and online learning method is key to solve this problem. Among existing online learning methods, online sequential extreme learning machine (OSELM) is a novel and practical online learning algorithm, and it has been successfully applied in the field of dynamic data stream analysis. Firstly, the theoretical foundation and the execution process of OSELM were reviewed. Then, regarding dynamic data flow analysis as the application background, this paper classified and summarized different kinds of improved OSELM algorithms, including the sliding window based OSELM, forgetting factor based OSELM, sample weighting based OSELM and other methods. This paper focused on the design ideas and implementation strategies of different kinds of algorithms, compared and analyzed the advantages and disadvantages of main algorithms. Finally, the possible works for future research were presented.

Keywords Online sequential extreme learning machine, Dynamic data stream analysis, Sliding window, Forgetting factor, Sample weighting

1 引言

动态数据流分析与预测是当前数据挖掘领域的一个重要研究方向,在金融服务、气象测控、交通疏导、工业控制、网络监控等领域被广泛应用,如股价预测、天气预测、交通流量预测、工业系统状态监测、网络入侵检测等^[1-2]。在这些现实应用中,数据样本以数据流的形式贯序到达,而且数据的潜在分布和变化趋势随时间不断发生变化^[3-4],这就要求学习模型能够不断学习新的数据样本以获取最新知识,并保持模型的快速更新,从而对目标系统的实时状态和后续趋势做出及时、准确的分析和预测^[5]。显然,对于此类动态时变环境下的数据

流分析问题,传统的离线学习方法缺乏时变环境下的快速更新能力和动态跟踪能力,在执行效率和预测精度方面都难以满足实际应用的要求。相对地,在线学习方法能以增量的方式实现对流式数据的持续快速学习,已成为当前解决动态数据流分析与预测问题的主流方法^[1,6]。

在众多在线学习方法中,在线贯序超限学习机(OSELM)^[7]是近年来提出的一种新颖而实用的在线学习算法。OSELM以超限学习机(ELM)^[8]的插值理论和逼近定理为基础,将单隐层前馈神经网络(SLFNs)的训练问题转化为线性方程组的求解问题,并采用递归最小二乘(RLS)方法递推计算输出权值,以实现模型的在线更新。借助于ELM强大的

收稿日期:2018-09-13 返修日期:2018-11-21 本文受国家自然科学基金(61603326,61379064,61273106)资助。

郭 威(1983—),男,博士,副教授,主要研究方向为机器学习、数据挖掘, E-mail:weiguo031@163.com(通信作者);于建江(1975—),男,博士,教授,主要研究方向为神经网络、网络控制系统;汤克明(1965—),男,博士,教授,主要研究方向为数据挖掘、智能计算;徐 涛(1962—),男,教授,博士生导师,主要研究方向为数据挖掘、智能信息处理。

学习逼近能力和 RLS 简洁的递推计算方式, OSELM 能够准确且高效地实现对数据样本的在线增量学习。与其他流行的在线学习算法相比, OSELM 具有学习速度快、泛化能力强、实现简单等诸多优点, 非常适合需要在线学习的工程应用。

基于 OSELM 良好的理论基础和优越的在线学习能力, OSELM 自提出便引起了研究者的极大兴趣, 并在理论和应用研究中都得到了快速发展^[9-14], 特别是在动态数据流在线分析与预测研究领域展现出了良好的应用价值和发展潜力, 取得了一批有意义的研究成果^[15-50]。本文旨在对这些研究成果进行归纳、总结和分析: 首先, 简要介绍 OSELM 的理论基础与算法实现; 在此基础上, 重点针对面向动态数据流在线分析的各种改进 OSELM 算法进行分类综述; 然后, 对其中两类主要算法的优缺点进行比较与分析; 最后, 指出该领域进一步的研究方向。

2 OSELM 的理论基础和算法实现

2.1 ELM 理论

超限学习机 ELM^[8]为 OSELM 的理论基础。对于 N 个任意的相异样本 $(x_i, y_i) \in \mathbf{R}^d \times \mathbf{R}^m$, 具有 n 个隐层节点且激活函数为 $g(x)$ 的 SLFNs 的数学模型为:

$$\sum_{j=1}^n \beta_j g_j(x_i) = \sum_{j=1}^n \beta_j G(a_j, b_j, x_i), i=1, 2, \dots, N \quad (1)$$

其中, a_j 和 b_j 是第 j 个隐层节点的学习参数, β_j 是连接第 j 个隐层节点与输出层的输出权值, $g_j(x_i) = G(a_j, b_j, x_i)$ 表示第 j 个隐层节点关于输入向量 x_i 的输出。

该 SLFNs 能以零误差逼近这 N 个样本, 意味着存在 (a_j, b_j) 及 β_j , 使得:

$$\sum_{j=1}^n \beta_j G(a_j, b_j, x_i) = y_i, i=1, 2, \dots, N \quad (2)$$

上面 N 个等式可写成如下矩阵形式:

$$\mathbf{H}\boldsymbol{\beta} = \mathbf{Y} \quad (3)$$

其中:

$$\mathbf{H} = \begin{bmatrix} \mathbf{h}_1 \\ \vdots \\ \mathbf{h}_N \end{bmatrix} = \begin{bmatrix} G(a_1, b_1, x_1) & \cdots & G(a_n, b_n, x_1) \\ \vdots & \cdots & \vdots \\ G(a_1, b_1, x_N) & \cdots & G(a_n, b_n, x_N) \end{bmatrix}_{N \times n} \quad (4)$$

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_1^\top \\ \vdots \\ \beta_n^\top \end{bmatrix}_{n \times m}, \mathbf{Y} = \begin{bmatrix} y_1^\top \\ \vdots \\ y_N^\top \end{bmatrix}_{N \times m} \quad (5)$$

其中, \mathbf{H} 称为网络的隐层输出矩阵, \mathbf{H} 的第 i 行 \mathbf{h}_i 为所有隐层节点关于输入 x_i 的特征映射, \mathbf{H} 的第 j 列为第 j 个隐层节点关于输入 x_1, x_2, \dots, x_N 的输出向量。

根据文献[8]中的插值理论, 给定一个 SLFN 和 N 个任意不同样本, 只要隐层激活函数无限可微, 隐层节点参数 (a_j, b_j) 可随机赋值并保持不变, 此时 \mathbf{H} 为一常数矩阵, 则 SLFNs 的训练问题就转化为求解线性系统 $\mathbf{H}\boldsymbol{\beta} = \mathbf{Y}$ 的最小二乘解 $\hat{\boldsymbol{\beta}}$ 问题, 因此有:

$$\|\hat{\mathbf{H}}\hat{\boldsymbol{\beta}} - \mathbf{Y}\| = \min_{\boldsymbol{\beta}} \|\mathbf{H}\boldsymbol{\beta} - \mathbf{Y}\| \quad (6)$$

根据前馈神经网络的泛化理论, 网络权值的模越小, 网络的泛化能力越强, 故 ELM 采用了最小模最小二乘解作为网络的输出权值, 即:

$$\hat{\boldsymbol{\beta}} = \mathbf{H}^\dagger \mathbf{Y} \quad (7)$$

其中, \mathbf{H}^\dagger 为 \mathbf{H} 的 Moore-Penrose 广义逆。如果 $\mathbf{H}^T \mathbf{H}$ 非奇异, 则式(7)可进一步写为:

$$\hat{\boldsymbol{\beta}} = \mathbf{H}^\dagger \mathbf{Y} = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{Y} \quad (8)$$

2.2 OSELM 算法

为了适应实时在线学习的需要, Liang 等^[7]将贯序学习的思想用于 ELM, 并使用 RLS 方法进行递推计算, 进而提出了一种新的在线学习算法 OSELM, 其学习过程简述如下。

在初始化阶段, 给定初始训练子集 $\Omega_0 = \{(x_i, y_i) \mid i = 1, \dots, N_0\}$, 根据式(8)计算初始输出权值:

$$\boldsymbol{\beta}_0 = \mathbf{P}_0 \mathbf{H}_0^T \mathbf{Y}_0 \quad (9)$$

其中, $\mathbf{P}_0 = (\mathbf{H}_0^T \mathbf{H}_0)^{-1}$, $\mathbf{H}_0 = [\mathbf{h}_1^T \ \mathbf{h}_2^T \ \cdots \ \mathbf{h}_{N_0}^T]^T$, $\mathbf{Y}_0 = [y_1 \ y_2 \ \cdots \ y_{N_0}]^T$ 。

在贯序学习阶段, 每当获取到第 k 个新的训练数据块 $\Omega_k = \{(x_i, y_i) \mid i = (\sum_{j=0}^{k-1} N_j) + 1, \dots, \sum_{j=0}^k N_j\}$, $k = 1, 2, \dots$, 便按照式(10)递推计算输出权值:

$$\mathbf{P}_k = \mathbf{P}_{k-1} - \mathbf{P}_{k-1} \mathbf{H}_k^T (\mathbf{I} + \mathbf{H}_k \mathbf{P}_{k-1} \mathbf{H}_k^T)^{-1} \mathbf{H}_k \mathbf{P}_{k-1} \quad (10)$$

$$\boldsymbol{\beta}_k = \boldsymbol{\beta}_{k-1} + \mathbf{P}_k \mathbf{H}_k^T (\mathbf{Y}_k - \mathbf{Y}_k \boldsymbol{\beta}_{k-1})$$

其中, $\mathbf{H}_k = [\mathbf{h}_{(\sum_{j=0}^{k-1} N_j)+1}^T \ \mathbf{h}_{(\sum_{j=0}^{k-1} N_j)+2}^T \ \cdots \ \mathbf{h}_{\sum_{j=0}^k N_j}^T]^T$, $\mathbf{Y}_k = [y_{(\sum_{j=0}^{k-1} N_j)+1} \ \cdots \ y_{\sum_{j=0}^k N_j}]^T$ 。

特别地, 在贯序学习过程中, 当数据样本总是逐一到达, 即 $N_k \equiv 1$ 时, 则式(10)中输出权值的更新计算公式可简化表示为:

$$\mathbf{P}_k = \mathbf{P}_{k-1} - \frac{\mathbf{P}_{k-1} \mathbf{h}_k^T \mathbf{h}_k \mathbf{P}_{k-1}}{1 + \mathbf{h}_k \mathbf{P}_{k-1} \mathbf{h}_k^T} \quad (11)$$

$$\boldsymbol{\beta}_k = \boldsymbol{\beta}_{k-1} + \mathbf{P}_k \mathbf{h}_k^T (y_k - \mathbf{h}_k \boldsymbol{\beta}_{k-1})$$

根据上述实现过程, OSELM 能以逐一或逐块的方式增量学习训练样本, 即在任意时刻, 只有最新到达的数据(块)需要参与学习, 一旦学习过程完成, 即可立即丢弃该数据(块)并不再使用。此外, 在在线学习过程中, 模型的更新计算完全基于上次迭代计算的中间结果和最新到达的数据(块)递推完成, 不需要保存和重新学习之前的训练样本, 极大地减少了存储开销和计算开销。大量仿真实验和应用实例结果也表明, 相比其他流行的在线学习算法, OSELM 不仅具有更好的泛化能力, 而且在学习速度上具有明显优势。

3 面向动态数据流分析的改进 OSELM 算法

作为一种优秀的在线学习算法, OSELM 能够快速且准确地实现对数据样本的在线贯序学习, 为流式数据的在线建模与分析提供了基本的解决方案。但对于时变环境下的动态数据流分析问题, 原始 OSELM 算法仍缺乏对新旧样本的差别学习能力和对失效样本的抛弃能力, 难以取得令人满意的预测效果。为了更好地处理时变数据流中隐含的动态变化特性, 学者们以 OSELM 为基础, 提出了一系列改进的在线超限学习算法, 以进一步提高其在时变环境下的动态跟踪能力和预测精度。总结分析已有的研究成果, 面向动态数据流分析的改进 OSELM 算法可归纳为以下 4 类: 基于滑动窗口的 OSELM 算法、基于遗忘因子的 OSELM 算法、基于样本加权

的 OSELM 算法以及其他方法。

3.1 基于滑动窗口的 OSELM

在动态时变环境下,数据样本往往具有一定的时效性,即数据仅在一段时间内是有效的,过了这段时间之后将会变得无效,如股价数据、天气数据、交通流量数据等,这就要求在线模型在不断学习新样本的同时,还能够及时消除旧的失效样本的影响。滑动窗口策略是驱除过时数据的一种直观方法,鉴于此,学者们将滑动窗口策略与 OSELM 相结合,提出了多种基于滑动窗口的改进 OSELM 算法^[15-22]。这些算法的基本思想如下:在在线学习过程中设定一个固定大小的滑动数据窗,每学习一个新数据样本(块)的同时,删除一个距离当前时刻最远的旧数据样本(块),以逐步摆脱旧样本对当前学习模型的影响。基于滑动窗口的 OSELM 算法的典型实现过程主要有两步实现^[15-19]和一步实现^[20-22]两种方式。

3.1.1 两步实现方式

假定滑动数据窗口的大小为 s ,并且在当前 k 时刻已知的 s 个训练样本为 $(x_i, y_i), i=k-s+1, k-s+2, \dots, k$,则这些样本所对应的隐层输出矩阵为 $\mathbf{H}_k = [\mathbf{h}_{k-s+1}^T \ \mathbf{h}_{k-s+2}^T \ \dots \ \mathbf{h}_k^T]^T$,输出向量为 $\mathbf{Y}_k = [y_{k-s+1} \ y_{k-s+2} \ \dots \ y_k]^T$,根据 OSELM 算法,此时网络的输出权值 β_k 可由式(11)计算得到。然后,在 $k+1$ 时刻,当有新的数据样本 (x_{k+1}, y_{k+1}) 到达时,对应的隐层输出矩阵和输出向量可分别表示为 $\mathbf{H}_{k+1} = [\mathbf{H}_k^T \ \mathbf{h}_{k+1}^T]^T, \mathbf{Y}_{k+1} = [\mathbf{Y}_k^T \ y_{k+1}]^T$ 。根据式(11),当前时刻的输出权值 β_{k+1} 可按下式递推计算:

$$\mathbf{P}_{k+1} = \mathbf{P}_k - \frac{\mathbf{P}_k \mathbf{h}_{k+1}^T \mathbf{h}_{k+1} \mathbf{P}_k}{1 + \mathbf{h}_{k+1}^T \mathbf{P}_k \mathbf{h}_{k+1}} \quad (12)$$

$$\beta_{k+1} = \beta_k + \mathbf{P}_{k+1} \mathbf{h}_{k+1}^T (y_{k+1} - \mathbf{h}_{k+1} \beta_k)$$

根据滑动窗口的思想,在加入新数据样本 (x_{k+1}, y_{k+1}) 的同时,还需要剔除距离当前时刻最远的旧样本 (x_{k-s+1}, y_{k-s+1}) ,此时将对应的隐层输出矩阵和输出向量分别表示为 $\bar{\mathbf{H}}_{k+1}, \bar{\mathbf{Y}}_{k+1}$,并且有 $\mathbf{H}_{k+1} = [\mathbf{h}_{k-s+1}^T \ \bar{\mathbf{H}}_{k+1}^T]^T, \mathbf{Y}_{k+1} = [y_{k-s+1} \ \bar{\mathbf{Y}}_{k+1}^T]^T$ 。经过迭代计算,可最终得到输出权值的计算公式为:

$$\bar{\mathbf{P}}_{k+1} = \mathbf{P}_{k+1} + \frac{\mathbf{P}_{k+1} \mathbf{h}_{k-s+1}^T \mathbf{h}_{k-s+1} \mathbf{P}_{k+1}}{1 - \mathbf{h}_{k-s+1}^T \mathbf{P}_{k+1} \mathbf{h}_{k-s+1}} \quad (13)$$

$$\bar{\beta}_{k+1} = \beta_{k+1} + \bar{\mathbf{P}}_{k+1} \mathbf{h}_{k-s+1}^T (\mathbf{h}_{k-s+1} \beta_{k+1} - y_{k-s+1})$$

3.1.2 一步实现方式

容易看出,上述两步实现算法^[15-19]在在线学习过程中需要一次增样学习(见式(12))和一次减样学习(见式(13))来完成模型的在线更新,因此其计算复杂度较 OSELM 明显增加,并且其只能实现对单个样本的逐一学习。以 OSELM 学习模型为基础,文献^[20-22]给出了滑动窗口策略的另一种实现方式,即仅需要一次迭代计算就可完成模型的在线更新,并且能够以逐一或逐块的方式实现对数据样本的在线学习。算法的在线更新公式如下:

$$\begin{aligned} \mathbf{P}_{k+1} &= \mathbf{P}_k - \mathbf{P}_k \begin{bmatrix} -\mathbf{H}_{k-s+1} \\ \mathbf{H}_{k+1} \end{bmatrix}^T \left(\mathbf{I} + \begin{bmatrix} \mathbf{H}_{k-s+1} \\ \mathbf{H}_{k+1} \end{bmatrix} \mathbf{P}_k \begin{bmatrix} -\mathbf{H}_{k-s+1} \\ \mathbf{H}_{k+1} \end{bmatrix}^T \right)^{-1} \times \begin{bmatrix} \mathbf{H}_{k-s+1} \\ \mathbf{H}_{k+1} \end{bmatrix} \mathbf{P}_k \\ \beta_{k+1} &= \beta_k + \mathbf{P}_{k+1} \begin{bmatrix} -\mathbf{H}_{k-s+1} \\ \mathbf{H}_{k+1} \end{bmatrix}^T \left(\begin{bmatrix} \mathbf{Y}_{k-s+1} \\ \mathbf{Y}_{k+1} \end{bmatrix} - \begin{bmatrix} \mathbf{H}_{k-s+1} \\ \mathbf{H}_{k+1} \end{bmatrix} \beta_k \right) \end{aligned} \quad (14)$$

其中, \mathbf{H}_{k+1} 和 \mathbf{Y}_{k+1} 为当前最新到达的第 $k+1$ 个数据块所对应的隐层输出矩阵和目标输出, \mathbf{H}_{k-s+1} 和 \mathbf{Y}_{k-s+1} 为过时的旧数据块所对应的隐层输出矩阵和目标输出,即模型的在线更新是完全基于当前新数据块和过时旧数据块迭代计算完成的。在此基础上,文献^[23]提出了一种带自适应滑动窗口的元认知 OSELM 算法并将其用于概念漂移学习,该算法中滑动窗口的大小可根据环境的变化进行自适应调整。

与两步实现方式相比,采用一步实现方式的滑动窗口 OSELM 算法由于只需要一次迭代计算就可完成模型的在线更新,因此计算复杂度有所降低。但从式(14)可以看出,该模型的在线更新涉及较多的矩阵转置和矩阵求逆计算,导致算法的计算复杂性仍然较高。

3.2 基于遗忘因子的 OSELM

3.2.1 F-OSELM 算法

对于时变动态系统而言,系统行为经常随时间发生演变;相应地,新的数据样本相比旧的数据样本应当具有更大的价值,并能更有效地反映时变系统的当前状态。遗忘因子方法是自适应滤波理论中描述观测样本在时变环境下动态变化特性的一种有效方法,其核心思想是根据数据样本到达的时间次序赋予每个样本不同的权值,以体现新旧样本的不同贡献。鉴于此,学者们将遗忘因子方法引入到 OSELM 中,提出了基于遗忘因子的改进 OSELM 算法(F-OSELM)^[24-26],算法的实现过程简述如下。

假定在 k 时刻有新的训练样本 (x_k, y_k) 到达,此时 $\mathbf{H}_k = [\mathbf{H}_{k-1}^T \ \mathbf{h}_k^T]^T, \mathbf{Y}_k = [\mathbf{Y}_{k-1}^T \ y_k]^T$,则当前 β_k 表示为:

$$\begin{aligned} \beta_k &= (\mathbf{H}_k^T \mathbf{H}_k)^{-1} \mathbf{H}_k^T \mathbf{Y}_k \\ &= (\mathbf{H}_{k-1}^T \mathbf{H}_{k-1} + \mathbf{h}_k^T \mathbf{h}_k)^{-1} (\mathbf{H}_{k-1}^T \mathbf{Y}_{k-1} + \mathbf{h}_k^T y_k) \end{aligned} \quad (15)$$

考虑到新旧样本的不同时效特性,对旧样本进行遗忘加权,将式(15)重写为:

$$\beta_k = (\lambda \mathbf{H}_{k-1}^T \mathbf{H}_{k-1} + \mathbf{h}_k^T \mathbf{h}_k)^{-1} (\lambda \mathbf{H}_{k-1}^T \mathbf{Y}_{k-1} + \mathbf{h}_k^T y_k) \quad (16)$$

其中, $\lambda \in (0, 1]$ 为遗忘因子,其作用是削弱旧样本的影响,从而间接增强新样本的作用。令 $\mathbf{P}_k = (\lambda \mathbf{H}_{k-1}^T \mathbf{H}_{k-1} + \mathbf{h}_k^T \mathbf{h}_k)^{-1}$,运用 Sherman-Morrison 矩阵求逆公式,可得到 F-OSELM 中输出权值的递推计算公式为:

$$\mathbf{P}_k = \frac{\mathbf{P}_{k-1}}{\lambda} - \frac{\mathbf{P}_{k-1} \mathbf{h}_k^T \mathbf{h}_k \mathbf{P}_{k-1}}{\lambda (\lambda + \mathbf{h}_k^T \mathbf{P}_{k-1} \mathbf{h}_k)} \quad (17)$$

$$\beta_k = \beta_{k-1} + \mathbf{P}_k \mathbf{h}_k^T (y_k - \mathbf{h}_k \beta_{k-1})$$

当 $\lambda=1$ 时, F-OSELM 退化为一般的 OSELM。

3.2.2 F-OSELM 的稳定性分析及改进算法

算法稳定性是评价算法优劣的一个重要指标。由本文第 2 节可知,原始 OSELM 算法的推导实现以隐层输出矩阵的自相关矩阵 $\mathbf{H}^T \mathbf{H}$ 非奇异为假设前提^[7],但由于算法中隐层节点参数的设置具有随机性,该假设在实际应用中未必总是成立,一旦 $\mathbf{H}^T \mathbf{H}$ 为奇异或病态矩阵,对其进行求逆计算的结果将极不可靠,算法的泛化性能将相应地严重下降;另外,我们通过实验验证发现,遗忘因子的引入会进一步加剧算法的不稳定性^[27],即 F-OSELM 算法存在稳定性差的不足。正则化技术是解决病态矩阵求逆问题的常用方法,基于此,学者们将正则化技术与遗忘因子方法同时融合到 OSELM 中,提出了基于正则化与遗忘因子的 OSELM 算法(RF-OS-

ELM)^[28-29]。RF-OSELM 在隐层输出矩阵的自相关矩阵 $\mathbf{H}^T \mathbf{H}$ 中添加了一个额外的正则化项,以增强其良态性以及对其进行矩阵求逆计算的可靠性,即将式(16)重写为:

$$\boldsymbol{\beta}_k = (\lambda \mathbf{H}_{k-1}^T \mathbf{H}_{k-1} + \lambda \delta \mathbf{I} + \mathbf{h}_k^T \mathbf{h}_k)^{-1} (\lambda \mathbf{H}_{k-1}^T \mathbf{Y}_{k-1} + \mathbf{h}_k^T y_k) \quad (18)$$

其中, δ 为正则化参数, \mathbf{I} 为单位矩阵。令 $\mathbf{P}_k = (\lambda \mathbf{H}_{k-1}^T \mathbf{H}_{k-1} + \lambda \delta \mathbf{I} + \mathbf{h}_k^T \mathbf{h}_k)^{-1}$, 进行矩阵求逆计算和递归推导, 可得到 RF-OSELM 中输出权值的递推计算公式与式(17)相同。值得注意的是, 虽然 RF-OSELM 与 F-OSELM 的更新计算公式在形式上是一样的, 但事实上, 在 RF-OSELM 的贯序学习过程中, 其正则化项 $\lambda \delta \mathbf{I}$ 已通过迭代计算隐含嵌入到递推计算公式中。当 $\delta = 0$ 时, RF-OSELM 退化为 F-OSELM; 当 $\delta = 0$ 且 $\lambda = 1$ 时, RF-OSELM 退化为原始的 OSELM。

尽管 RF-OSELM 通过使用正则化技术使得算法的稳定性有了一定的提高, 但文献[27]从推演计算的角度证明了其稳定性是不持续的。理论上, RF-OSELM 算法等价于最小化求解如下带遗忘因子和正则化项的最小二乘代价函数:

$$J_{\text{RF}}(\boldsymbol{\beta}_k) = \sum_{i=1}^k \lambda^{k-i} |y_i - \mathbf{h}_i \boldsymbol{\beta}_k|^2 + \delta \lambda^k \|\boldsymbol{\beta}_k\|^2 \quad (19)$$

分析式(19)可以发现, RF-OSELM 等价使用了一种特殊的指数遗忘正则化项 $\delta \lambda^k \|\boldsymbol{\beta}_k\|^2$, 这种特殊的选择使得 RF-OSELM 所得到的模型更新计算公式具有与原始 OSELM 同样的简洁性, 但这同时也对模型的稳定性造成了不利影响。在时变环境下, 遗忘因子 λ 通常取略小于 1 的常数。在此情形下, λ^k 将随着时间 k 的不断增加而单调递减并趋近于 0; 相应地, $\delta \lambda^k \|\boldsymbol{\beta}_k\|^2$ 的正则化功能将逐渐弱减直至完全消失, 从而最终导致 RF-OSELM 出现与原始 OSELM 同样的由于病态矩阵求逆而引起的算法发散问题。基于以上分析, 文献[27]提出了一种具有广义正则化与遗忘机制的在线贯序超限学习机 GRF-OSELM。与 RF-OSELM 不同, GRF-OSELM 在其代价函数中引入一种新的常数项 $\delta \|\boldsymbol{\beta}_k\|^2$ 来替代式(19)中的指数遗忘正则化项 $\delta \lambda^k \|\boldsymbol{\beta}_k\|^2$, 使得其正则化功能不会因遗忘因子的影响而被逐渐削弱, 从而保证了算法在整个在线学习过程中都具有恒定的正则化效果和持续的稳定性。GRF-OSELM 的代价函数可表示为:

$$J_{\text{GRF}}(\boldsymbol{\beta}_k) = \sum_{i=1}^k \lambda^{k-i} |y_i - \mathbf{h}_i \boldsymbol{\beta}_k|^2 + \delta \|\boldsymbol{\beta}_k\|^2 \quad (20)$$

对式(20)进行递推求解和近似计算, 可得到 GRF-OSELM 算法中输出权值 $\boldsymbol{\beta}_k$ 的递推计算公式为:

$$\begin{aligned} \mathbf{P}_k^* &= \frac{1}{\lambda} \mathbf{P}_{k-1} - \frac{\delta(1-\lambda)}{\lambda^2} \mathbf{P}_{k-1} (\mathbf{I} - \frac{\delta(1-\lambda)}{\lambda} \mathbf{P}_{k-1}) \mathbf{P}_{k-1} \\ \mathbf{P}_k &= \mathbf{P}_k^* - \frac{\mathbf{P}_k^* \mathbf{h}_k^T \mathbf{h}_k \mathbf{P}_k^*}{1 + \mathbf{h}_k \mathbf{P}_k^* \mathbf{h}_k^T} \end{aligned} \quad (21)$$

$$\boldsymbol{\beta}_k = \boldsymbol{\beta}_{k-1} + \mathbf{P}_k \mathbf{h}_k^T (y_k - \mathbf{h}_k \boldsymbol{\beta}_{k-1}) - \delta(1-\lambda) \mathbf{P}_k \boldsymbol{\beta}_{k-1}$$

3.2.3 VFF-OSELM 算法

上述文献^[24-29]通过理论分析和实验证明了遗忘因子方法对于动态时变环境下的数据流分析与预测问题是积极有效的, 但实际应用中, 在没有先验知识的情况下要提前设定一个准确的遗忘因子是非常困难甚至不可能的^[51]。此外, 在复杂时变环境下, 数据流的变化速率可能是不规律的, 即有时变化快, 有时变化慢, 在此情形下预设的固定遗忘因子并不能保证

对时变系统的动态变化有全局的自适应性。通常, 可变遗忘因子(VFF)方法是一种更好的选择。为此, 学者们又提出了多种基于 VFF 的 OSELM 算法^[30-38], 以进一步提高对复杂时变系统的动态跟踪能力和预测精度。预测误差是表征时变系统动态变化最直观、有效的指标; 当系统变化较快时, 预测误差将增大, 遗忘因子应相应减小, 以加速遗忘旧的失效状态并及时跟踪时变系统的最新状态; 而当系统变化较慢或趋于平稳时, 随着预测误差的降低, 遗忘因子应相应增大, 以提高系统在稳态下的预测精度。基于以上分析, 不少学者直接以在线预测误差作为核心观测变量, 并通过设计满足一定区间范围的修正函数来动态调整遗忘因子的大小^[30-34]。这些 VFF 方法最突出的优点就是简洁、高效, 几乎不产生额外的计算开销; 但对于不同的应用问题, 需要设置不同的控制参数, 这无疑增加了算法应用的难度。文献[35]提出了一种方向遗忘因子 OSELM 算法(DFF-OSELM), 该算法认为数据样本具有方向性, 并按照一定的方向来遗忘旧的数据样本, 从而提高对动态系统的预测性能, 但该算法同样存在较多的辅助参数需要设置。相对地, 文献[36]提出了一种更为通用的自适应遗忘因子 OSELM 算法(AFF-OSELM), 该算法采用误差梯度下降方法进行遗忘因子的自适应迭代调整, 无须进行繁杂的参数设置, 具有较好的普适性, 并已在机械故障预测^[28]、非平稳系统识别^[36]、混沌时间序列预测^[37]等问题中展现出了良好的应用效果。然而, 上述 AFF-OSELM 算法并未考虑迭代学习过程中可能出现的病态矩阵求逆问题或采用传统的指数遗忘正则化方法来解决该问题, 因此它们均存在潜在的不稳定性。在文献[36]的基础上, 文献[38]以广义正则化为前提条件, 推导出一种新的自适应遗忘因子方法并整合到 OSELM 中, 使得算法同时兼具良好的动态跟踪能力和持续的稳定性, 但该算法的计算复杂度较高。

最近, 文献[39]将滑动窗口方法与遗忘因子方法相结合, 提出一种记忆退化 OSELM 算法, 该算法使用滑动窗口策略抛弃旧的失效样本, 并使用自适应遗忘因子方法对滑动窗口内的有效样本进行遗忘加权以体现其时效性, 从而进一步提升了算法的预测精度和泛化性能。

3.3 基于样本加权的 OSELM

为了处理数据分布随时间动态变化的数据集, 文献[40]提出了一种基于样本加权的 Timeliness OSELM 算法(T-OSELM)。与 F-OSELM 算法^[24-26]对旧样本施以小于 1 的遗忘因子作为权值不同, T-OSELM 直接对新样本添加一个大于 1 的权值系数以突出新样本更加重要的作用, 从而更好地跟踪目标系统的实时状态。同时, T-OSELM 在在线学习过程中采用一种自适应迭代策略, 有效保证了算法的稳定性和收敛性。在股价数据、天气数据、手势数据等真实数据集上的测试结果表明, 对于时变性较强的数据集, T-OSELM 具有较对比算法更高的学习精度、更快的收敛速度和更好的稳定性。在 T-OSELM 算法的基础上, 文献[41]从统计学习理论中经验风险和结构风险的平衡性分析出发, 通过引入正则化因子来平衡这两种风险, 从而进一步提高了算法的泛化性能, 并有效避免了算法中潜在的病态问题。此外, 基于不同的加权策略, 学者们还提出了多种类似的加权 OSELM 算法, 并在非平

稳时间序列在线预测^[42]、时变非线性系统建模^[43]、电能质量识别^[44]等问题中取得成功应用。

3.4 其他方法

除了以上3类方法,学者们还提出了其他一些基于OSELM的改进算法,以实现动态流式数据的在线分析与预测。为有效表征时变系统的实时动态特征,文献[45-46]提出了模型结构可自适应调整的OSELM算法,其基本思想是:当系统状态发生改变时,算法可自动调节隐层节点个数,以更准确地描绘出系统当前的实时状态,从而提高对时变系统的预测精度。针对非平稳环境下的在线学习问题,文献[47]将在线超限学习算法扩展到时变神经网络的训练中,提出了一种时变OSELM算法并将其用于时变系统的在线识别。针对具有不确定性的时变系统预测问题,文献[48]将灰色理论与OSELM模型相结合,提出了一种基于OSELM的灰色预测模型,并将其应用于船舶横摇运动的在线预测。此外,文献[49]将集成学习的思想用于OSELM,提出了一种集成子集OSELM算法,该算法采用集成学习中经典的动态加权多数投票方法来处理概念漂移学习。文献[50]通过融合使用协同训练策略和OSELM算法,实现了动态环境下移动物体阴影的检测。

上述改进OSELM算法通过扩展网络模型^[45-47]或融合使用其他方法^[48-50],有效地提高了对动态时变系统的在线建模与分析能力,但这也使得算法的计算复杂性有所增加。此外,这些算法大多用于解决某一特定类型的应用问题,其实用性和可扩展性有待进一步验证。

4 算法的比较与分析

在上述4类面向动态数据流分析的改进OSELM算法中,基于滑动窗口的OSELM和基于遗忘因子的OSELM是学者们研究最为集中的两类典型算法,本节将主要对这两类算法进行比较和分析。

1)在实现方式上,基于滑动窗口的OSELM既可实现单个样本的逐一在线学习^[15-19],也可实现数据样本块的逐块在线学习^[20-22]。基于遗忘因子的OSELM大多采用逐一的方式进行样本的在线学习^[24-29]。

2)在计算复杂性方面,基于滑动窗口的OSELM在在线学习过程中需要增样学习和减样学习这两个学习步(式(12)和式(13))来完成模型的在线更新,或通过一次较为复杂的迭代计算(式(14))来完成,其计算时间较OSELM增加了近一倍。相对而言,基于遗忘因子的OSELM的模型更新公式(式(17))具有与原始OSELM算法(式(11))同样的简洁性,二者的计算时间基本相当。显然,与基于滑动窗口的OSELM相比,基于遗忘因子的OSELM的计算复杂性更低、学习效率更高。

3)在自适应性方面,已有基于滑动窗口的OSELM算法^[15-22]绝大多数采用固定大小的滑动数据窗,滑动窗口的自适应调节的实现技术上较为复杂^[23]。相对地,面对不同的应用问题,多种基于可变遗忘因子的OSELM算法^[28,30-38]被广泛提出并取得了良好的应用效果,展现出更好的自适应性。

4)在使用灵活性方面,基于滑动窗口的OSELM可被看

作基于遗忘因子的OSELM的一种特例,即在基于遗忘因子的OSELM中,将一定数量的最近样本的遗忘因子均设置为1(对应滑动窗口内的有效样本),并将其他旧样本的遗忘因子均设置为0(对应滑动窗口外的无效样本),即可等价得到基于滑动窗口的OSELM。从这个意义上讲,基于遗忘因子的OSELM在新旧样本的加权使用上无疑更为灵活。

5)在稳定性方面,文献[15]以基于滑动窗口的OSELM算法为基础,通过融合使用正则化技术进一步增强了算法的稳定性,但笔者通过实验验证发现该方法并不具有普适性,在某些情形下仍然存在失效的风险。而在基于遗忘因子的OSELM中,遗忘因子的引入进一步加剧了算法中潜在的病态问题,从而导致F-OSELM算法极不稳定^[27]。虽然基于正则化与遗忘因子的RF-OSELM算法^[28-29]通过引入正则化技术在一定程度上提升了算法的稳定性,但这种稳定性仍然是不持续的。在此基础上,笔者进一步提出了基于广义正则化与遗忘因子的GRF-OSELM算法^[27],从根本上保证了算法具有恒定的正则化效果和持续的稳定性,但这也一定程度上增加了算法的计算复杂性。因此,在稳定性方面,这两类算法均存在一定的缺陷。

综上,基于滑动窗口策略的OSELM算法的优点是简单直观、实现方式多样,对处理具有分段突变特征的时变应用具有理想的效果。而对于其他一些具有复杂时变特征的应用场景,基于遗忘因子的OSELM算法在计算效率、自适应性和应用灵活性方面更有优势,实用性更强。

结束语 动态数据流在线分析与预测是一个具有广泛应用价值的研究课题,研究建立具有快速更新能力和自适应跟踪能力的在线学习算法是解决这一问题的关键。本文以OSELM这一新兴在线学习算法作为基础理论模型,对面向动态数据流分析的各种改进OSELM算法按不同类别进行了归纳和总结,并对其中两类主要算法的优缺点进行了比较与分析。

作为一种优秀的在线学习算法,OSELM为解决动态数据流在线分析与预测问题提供了良好的技术支持,并在理论分析和实际应用中取得了许多令人满意的研究成果。总结分析已有的研究成果,本文认为该领域仍有以下问题值得进一步研究。

1)稳定性分析。如前文所述,在时变环境下,一般的正则化技术对于提升OSELM的稳定性具有局限性,虽然广义正则化技术^[27]能一劳永逸地解决该问题,但这是以增加算法的计算复杂度为代价得到的。因此,如何在显著增加算法计算复杂度的前提下保证算法具有良好的稳定性仍有待进一步研究。

2)鲁棒性分析。在现实应用中,实际采集到的数据流样本中往往包含离群值,离群值的存在不仅会严重影响系统建模的准确性,也会对比时变环境下突变拐点的判断造成严重干扰^[34]。虽然目前已有学者对增强OSELM在离群值环境下的鲁棒性进行了研究并提出了相应的算法^[52],但如何将其拓展到对比时变离群值环境是一个值得继续研究的问题。

3)遗忘因子的自适应调整方法。在已有研究中,基于可变遗忘因子的OSELM算法^[28,30-38]是处理复杂时变应用最直接有效的方法,然而这些算法仍然存在辅助参数设置复杂、计

算复杂度高等不足。因此,如何设计出更加简洁高效、普适性更强的可变遗忘因子方法仍有较大的研究空间。

参 考 文 献

- [1] OLORUNNIMBE M K, VIKTOR H L, PAQUET E. Dynamic adaptation of online ensembles for drifting data streams[J]. *Journal of Intelligent Information Systems*, 2018, 50(2): 291-313.
- [2] MEJRI D, LIMAM M, WEIHS C. A new dynamic weighted majority control chart for data streams[J]. *Soft Computing*, 2018, 22(2): 511-522.
- [3] MENCAGLI G, TORQUATI M, DANELUTTO M. Elastic-PPQ: A two-level autonomic system for spatial preference query processing over dynamic data streams[J]. *Future Generation Computer Systems*, 2018, 79(3): 862-877.
- [4] MOHAMAD S, BOUCHACHIA A, SAYEDMOUCHAWEH M. A Bi-Criteria Active Learning Algorithm for Dynamic Data Streams[J]. *IEEE Transactions on Neural Networks & Learning Systems*, 2018, 29(1): 74-86.
- [5] KHAMASSI I, SAYED-MOUCHAWEH M, HAMMAMI M, et al. Discussion and review on evolving data streams and concept drift adapting[J]. *Evolving Systems*, 2018, 9(1): 1-23.
- [6] LI C, WEI F, DONG W, et al. Dynamic Structure Embedded Online Multiple-Output Regression for Streaming Data[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2018, doi:10.1109/TPAMI.2018.2794446.
- [7] LIANG N Y, HUANG G B, SARATCHANDRAN P, et al. A fast and accurate online sequential learning algorithm for feedforward networks[J]. *IEEE Transactions on Neural Networks*, 2006, 17(6): 1411-23.
- [8] HUANG G B, ZHU Q Y, SIEW C K. Extreme learning machine: Theory and applications [J]. *Neurocomputing*, 2006, 70(1): 489-501.
- [9] HUYNH H T, WON Y. Regularized online sequential learning algorithm for single-hidden layer feedforward neural networks [J]. *Pattern Recognition Letters*, 2011, 32(14): 1930-1935.
- [10] HUANG S, WANG B, QIU J, et al. Parallel ensemble of online sequential extreme learning machine based on MapReduce[J]. *Neurocomputing*, 2016, 174: 352-367.
- [11] SCARDAPANE S, COMMINELO D, SCARPINITI M, et al. Online Sequential Extreme Learning Machine With Kernels[J]. *IEEE Transactions on Neural Networks & Learning Systems*, 2015, 26(9): 2214-2220.
- [12] JIA X, WANG R, LIU J, et al. A semi-supervised online sequential extreme learning machine method [J]. *Neurocomputing*, 2016, 174: 168-178.
- [13] WONG P K, XIANG H G, WONG K I, et al. Online extreme learning machine based modeling and optimization for point-by-point engine calibration [J]. *Neurocomputing*, 2018, 277: 187-197.
- [14] OUYANG J Q, ZHOU Y, TANG H R. A Meteorological Prediction Model Based on Storm and Online Sequential Extreme Learning Machine[J]. *Journal of Computer Research & Development*, 2017, 54(8): 1736-1743. (in Chinese)
- 欧阳建权, 周勇, 唐欢容. 基于 Storm 的在线序列极限学习机的气象预测模型[J]. *计算机研究与发展*, 2017, 54(8): 1736-1743.
- [15] ZHOU X, LIU Z, ZHU C. Online Regularized and Kernelized Extreme Learning Machines with Forgetting Mechanism [J]. *Mathematical Problems in Engineering*, 2014, 2014: 1-11.
- [16] ZHOU X, WANG C. Cholesky factorization based online regularized and kernelized extreme learning machines with forgetting mechanism [J]. *Neurocomputing*, 2016, 174: 1147-1155.
- [17] ZHANG H, ZHANG S, YIN Y. Kernel online sequential ELM algorithm with sliding window subject to time-varying environments [J]. *Memetic Computing*, 2016, 10(1): 43-52.
- [18] WANG X, HAN M. Online sequential extreme learning machine with kernels for nonstationary time series prediction [J]. *Neurocomputing*, 2014, 145: 90-97.
- [19] ZHANG X, WANG H L. Fixed-memory extreme learning machine and its applications [J]. *Control and Decision*, 2012, 27(8): 1206-1210. (in Chinese)
- 张弦, 王宏力. 限定记忆极端学习机及其应用 [J]. *控制与决策*, 2012, 27(8): 1206-1210.
- [20] ZHAO J, WANG Z, PARK D S. Online sequential extreme learning machine with forgetting mechanism [J]. *Neurocomputing*, 2012, 87: 79-89.
- [21] WANG J, RAN R, ZHOU Y. A Short-Term Photovoltaic Power Prediction Model Based on an FOS-ELM Algorithm [J]. *Applied Sciences*, 2017, 7(4): 423.
- [22] XIAO D, LI B, ZHANG S. An online sequential multiple hidden layers extreme learning machine method with forgetting mechanism [J]. *Chemometrics and Intelligent Laboratory Systems*, 2018, 176: 126-133.
- [23] MIRZA B, LIN Z. Meta-cognitive online sequential extreme learning machine for imbalanced and concept-drifting data classification [J]. *Neural Networks*, 2016, 80: 79-94.
- [24] ZHANG X, WANG H L. Selective forgetting extreme learning machine and its application to time series prediction [J]. *Acta Physica Sinica*, 2011, 60(8): 68-74. (in Chinese)
- 张弦, 王宏力. 具有选择与遗忘机制的极端学习机在时间序列预测中的应用 [J]. *物理学报*, 2011, 60(8): 68-74.
- [25] LU J, HUANG J, LU F. Sensor Fault Diagnosis for Aero Engine Based on Online Sequential Extreme Learning Machine with Memory Principle [J]. *Energies*, 2017, 10(1): 1-15.
- [26] TIAN Z, WANG G, REN Y, et al. An Adaptive Online Sequential Extreme Learning Machine for Short-Term Wind Speed Prediction Based on Improved Artificial Bee Colony Algorithm [J]. *Neural Network World*, 2018, 28(3): 191-212.
- [27] GUO W, XU T, TANG K M, et al. Online sequential extreme learning machine with generalized regularization and forgetting mechanism [J]. *Control & Decision*, 2017, 32(2): 247-254. (in Chinese)
- 郭威, 徐涛, 汤克明, 等. 具有广义正则化与遗忘机制的在线贯序超限学习机 [J]. *控制与决策*, 2017, 32(2): 247-254.
- [28] DU Z L, LI X M, ZHENG Z G, et al. Extreme learning machine based on regularization and forgetting factor and its application

- in fault prediction[J]. Chinese Journal of Scientific Instrument, 2015, 36(7):1546-1553. (in Chinese)
- 杜占龙,李小民,郑宗贵,等.基于正则化与遗忘因子的极限学习机及其在故障预测中的应用[J].仪器仪表学报,2015,36(7):1546-1553.
- [29] WU Z, TANG H, HE S, et al. Fast dynamic hysteresis modeling using a regularized online sequential extreme learning machine with forgetting property[J]. The International Journal of Advanced Manufacturing Technology, 2018, 94:3473-3484.
- [30] ZHANG H, ZHANG S, YIN Y. Online sequential ELM algorithm with forgetting factor for real applications[J]. Neurocomputing, 2017, 261:144-152.
- [31] LIU D, WU Y, JIANG H. FP-ELM: An online sequential learning algorithm for dealing with concept drift[J]. Neurocomputing, 2016, 207:322-334.
- [32] LI Y, ZHANG S, YIN Y, et al. A Novel Online Sequential Extreme Learning Machine for Gas Utilization Ratio Prediction in Blast Furnaces[J]. Sensors, 2017, 17(8):1847.
- [33] LU J, HUANG J, LU F. Time Series Prediction Based on Adaptive Weight Online Sequential Extreme Learning Machine[J]. Applied Sciences, 2017, 7(3):217.
- [34] GUO W, XU T, YU J J, et al. Online Sequential Extreme Learning Machine Based on M-estimator and Variable Forgetting Factor[J]. Journal of Electronics & Information Technology, 2018, 40(6):1360-1367. (in Chinese)
- 郭威,徐涛,于建江,等.基于 M-estimator 与可变遗忘因子的在线贯序超限学习机[J].电子与信息学报,2018,40(6):1360-1367.
- [35] SOARES S G, ARA JO R. An adaptive ensemble of online Extreme Learning Machines with variable forgetting factor for dynamic system prediction[J]. Neurocomputing, 2016, 171:693-707.
- [36] LIMJS, LEES K, PANGH S, et al. Low complexity adaptive forgetting factor for online sequential extreme learning machine (OS-ELM) for application to nonstationary system estimations [J]. Neural Computing & Applications, 2013, 22(3-4):569-576.
- [37] LIU J. Adaptive forgetting factor OS-ELM and bootstrap for time series prediction[J]. International Journal of Modeling Simulation & Scientific Computing, 2017, 8(3):1-19.
- [38] GUO W, XU T, TANG K, et al. Online Sequential Extreme Learning Machine with Generalized Regularization and Adaptive Forgetting Factor for Time-Varying System Prediction [J]. Mathematical Problems in Engineering, 2018, 2018:1-22.
- [39] ZOU Q Y, WANG X J, ZHOU C J, et al. The memory degradation based online sequential extreme learning machine[J]. Neurocomputing, 2018, 275:2864-2879.
- [40] GU Y, LIU J, CHEN Y, et al. TOSELM: Timeliness Online Sequential Extreme Learning Machine[J]. Neurocomputing, 2014, 128:119-127.
- [41] LUO X, YANG X, JIANG C, et al. Timeliness online regularized extreme learning machine[J]. International Journal of Machine Learning and Cybernetics, 2016, 9(3):465-476.
- [42] MAO W, WANG J, WANG L, et al. Online Sequential Prediction for Nonstationary Time Series with New Weight-Setting Strategy Using Extreme Learning Machine[J]. Mathematical Problems in Engineering, 2015, 2015:1-13.
- [43] LU X, ZHOU C, HUANG M, et al. Regularized online sequential extreme learning machine with adaptive regulation factor for time-varying nonlinear system[J]. Neurocomputing, 2016, 174:617-626.
- [44] SAHANI M, DASH P K. Variational mode decomposition and weighted online sequential extreme learning machine for power quality event patterns recognition[J]. Neurocomputing, 2018, 310:10-27.
- [45] YIN J C. A variable-structure online sequential extreme learning machine for time-varying system prediction[J]. Neurocomputing, 2017, 261:115-125.
- [46] XU S, WANG J. Dynamic extreme learning machine for data stream classification[J]. Neurocomputing, 2017, 238:433-449.
- [47] YE Y, SQUARTINI S, PIAZZA F. Online sequential extreme learning machine in nonstationary environments[J]. Neurocomputing, 2013, 116:94-101.
- [48] YIN J C, ZOU Z J, XU F, et al. Online ship roll motion prediction based on grey sequential extreme learning machine[J]. Neurocomputing, 2014, 129:168-174.
- [49] MIRZA B, LIN Z, LIU N. Ensemble of subset online sequential extreme learning machine for class imbalance and concept drift [J]. Neurocomputing, 2015, 149:316-329.
- [50] GHIMIRE D, LEE J. Online sequential extreme learning machine-based co-training for dynamic moving cast shadow detection[J]. Multimedia Tools and Applications, 2015, 75(18):11181-11197.
- [51] CAI Y, LAMARE R C D, ZHAO M, et al. Low-Complexity Variable Forgetting Factor Mechanism for Blind Adaptive Constrained Constant Modulus Algorithms [J]. IEEE Transactions on Signal Processing, 2012, 60(8):3988-4002.
- [52] GUO W, XU T, TANG K. M-estimator-based online sequential extreme learning machine for predicting chaotic time series with outliers[J]. Neural Computing & Applications, 2017, 28(12):4093-4110.