

基于改进 Real AdaBoost 算法的软件可靠性预测

杜瑞超, 华继学, 翟夕阳, 李志鹏

(空军工程大学防空反导学院, 西安, 710051)

摘要 针对基于单一神经网络的软件可靠性模型预测精度低和可信性差的问题,提出一种基于加权信息熵(WIE)的 Real BP-AdaBoost 算法。首先,用 BP 神经网络个体代替 Real AdaBoost 算法的基分类器,构建 Real BP-AdaBoost 算法。然后,对 Real BP-AdaBoost 算法的加权方式进行改进,以基分类器对训练样本的整体分类权值与基分类器对测试样本的个体分类权值的乘积作为最终的加权系数,得到 WIE Real BP-AdaBoost 算法。最后,通过 2 组软件实际失效数据对 WIE Real BP-AdaBoost 算法的有效性进行验证,并与 SVM、BP 网络、Elman 网络和 Real BP-AdaBoost 算法进行比较研究。实验结果显示,WIE Real BP-AdaBoost 算法对 2 组数据预测的均方误差分别为 0.442 87 和 0.284 71,均低于 4 个对比模型的均方误差,说明了 WIE Real BP-AdaBoost 算法模型具有更高的预测精度和可信性。

关键词 软件可靠性预测;Real AdaBoost 算法;基分类器;加权方式;信息熵

DOI 10.3969/j.issn.1009-3516.2018.01.016

中图分类号 TP302.7 **文献标志码** A **文章编号** 1009-3516(2018)01-0091-06

Research on Software Reliability Prediction Based on Improved Real AdaBoost

DU Ruichao, HUA Jixue, ZHAI Xiyang, LI Zhipeng

(Air and Missile Defense College, Air Force Engineering University, Xi'an 710051, China)

Abstract: Aimed at the problems that the prediction accuracy is low and the dependability of software reliability model is weak based on single neural network, a Real BP-AdaBoost based on weighted information entropy (WIE) is proposed. First, the Real BP-AdaBoost is established by taking BP neural network as the base classifier of Real AdaBoost. Then, the weighted method of base classifier of Real BP-AdaBoost is improved by utilizing the product of the overall weights of the classifier for training samples and individual weight of the classifier to test samples as the final weight, and the WIE Real BP-AdaBoost is produced. Finally, the proposed algorithm is compared with SVM, BP neural network, Elman neural network and Real BP-AdaBoost with respect to two real software failure data. The mean square error of WIE Real BP-AdaBoost of the forecasted two sets data is 0.442 87 and 0.284 71 respectively, both are below the mean square error of the four comparison models. The result shows that WIE Real BP-AdaBoost is higher in prediction accuracy and reliable in dependability.

Key words: software reliability prediction; Real BP-AdaBoost; base classifier; weighted method; information entropy

软件可靠性是软件质量的重要属性之一,对软件失效数据进行分析,建立软件可靠性模型,从而对

收稿日期: 2017-06-12

基金项目: 国家自然科学基金(61402517)

作者简介: 杜瑞超(1991—),男,黑龙江五大连池人,硕士生,主要从事嵌入式系统与软件可靠性研究。E-mail:redu527@163.com

引用格式: 杜瑞超,华继学,翟夕阳,等. 基于改进 Real AdaBoost 算法的软件可靠性预测 [J]. 空军工程大学学报(自然科学版), 2018, 19(1): 91-96. DU Ruichao, HUA Jixue, ZHAI Xiyang, et al. Research on Software Reliability Prediction Based on Improved Real AdaBoost [J]. Journal of Air Force Engineering University (Natural Science Edition), 2018, 19(1): 91-96.

软件可靠性进行预测,是研究人员进行软件可靠性研究的一般思路。一个优秀的软件可靠性模型对软件工程来说具有重要意义,通过软件可靠性模型进行可靠性预测,从而推断软件失效行为,量化软件可靠性行为和状态,有助于软件开发者开发出高质量的软件。

近年来,以预测软件可靠性为目标,研究人员建立的软件可靠性模型数以千计^[1-5]。其中,基于神经网络的软件可靠性模型逐渐成为研究热点^[6-10]。研究表明,神经网络具有很强的非线性运算能力,且可以根据失效过程的动态变化进行调整,适用于对非线性的软件失效数据进行预测。然而,目前单一神经网络对软件可靠性预测存在预测精度低和可信性差的问题,因此,本文考虑使用基于多神经网络集成的软件可靠性模型,在 Real AdaBoost 算法的基础上,利用 BP 神经网络个体对 Real AdaBoost 算法中的基分类器进行替换,通过引入信息熵的概念,改变各神经网络个体的加权方式,对其分配相应的权重,并对权重进行动态调整,最后加权输出预测结果,构建 WIE Real BP-AdaBoost 算法。

本文使用 NTDS 数据集^[11]和 LCCS 数据集^[12]对 WIE Real BP-AdaBoost 算法的有效性进行验证。实验结果表明,WIE Real BP-AdaBoost 算法模型的预测效果较 SVM、BP 网络、Elman 网络、Real BP-Adaboost 算法的预测效果更优。

1 神经网络集成

1990年,神经网络集成(Neural Network Ensembles, NNE)被 Hasen 和 Salamon 正式提出^[13]。1996年, Sollich 和 Krough 提出了相类似的 NNE 的定义^[14]。由于其使用简单且效果明显, NNE 一经提出便大受欢迎,被国际权威 TG Dieteerich 称为当前机器学习四大方向之首^[15]。当前, NNE 广泛应用于图像分类^[16], 汇率预测^[17], 医疗诊断^[18]等领域。 NNE 的实现可以分为以下 2 步: ①使用多个神经网络个体对同一训练样本进行学习, 得到多个神经网络个体; ②将多个神经网络个体对同一问题的预测结果进行集成。

生成神经网络个体的方法主要有 3 种: 变换数据、改变网络特性和神经网络个体优选。结合本文提出的算法, 着重介绍变换训练数据的 AdaBoost 算法^[19]。 AdaBoost 算法的具体步骤如下:

1) 给定训练集 $A = \{(a_1, b_1), \dots, (a_n, b_n)\}$, 根据训练集来确定神经网络结构和最大迭代次数 M ;

2) 迭代 $m=1$, 权重分布 $\omega_m(i) = 1/n$, 平均损失

函数 $\bar{A}_m = 0$;

3) 当平均损失函数 $\bar{A}_m \leq 0.5$ 或 $m \leq M$ 时, 进行迭代。根据权重 $\omega_m(i)$ 从训练样本中抽取样本来训练多个弱预测器 $h_m(a) (i=1, 2, \dots, M)$, 计算训练样本的损失函数 $l_m(i) = |h'_m(a) - b_m|$, $h'_m(a)$ 为网络个体预测值, b_m 为期望值; 计算每个样本的损失函数 $L_m(i) = \frac{l_m^2(i)}{L_i^2}$, 其中, $L_i = \max_{i=1, 2, \dots, n} (l_m(i))$ 。计算平均损失 $\bar{L}_m = \sum_{i=1}^n L_m(i) \omega_m(i)$, 计算 $\alpha_m = \bar{L}_m / (1 - \bar{L}_m)$ 。

4) 权值更新。根据 α_m 调整下一轮训练样本的权值, 公式为:

$$\omega_{m+1}(i) = \frac{\omega_m(i) \alpha_m}{X_m}, (m = 1, 2, \dots, n) \quad (1)$$

式中: X_m 为标准化因子。

5) 经过 M 轮迭代训练后, 得到 M 个弱预测函数 $h_m(a)$, 此时组成一个强预测函数:

$$h_{\text{final}}(a) = \inf [b \in B; \sum_{m, b_m \leq b} \log(1/\alpha_m) \geq \log(1/\alpha_m)] \quad (2)$$

2 WIE Real BP-AdaBoost 算法

2.1 Real AdaBoost 算法

由于 AdaBoost 算法只能输出分类类别而不能输出分类概率, 因此 Schapire 等人于 1999 年提出一种改进的 AdaBoost 算法, 并命名为 Real AdaBoost^[20]。由于划分区间所形成的基分类器具有输出概率的能力, 因而 Real AdaBoost 算法具有连续置信度输出的能力。与 AdaBoost 算法的不同之处也正是基分类器是否概率输出以及测试时基分类器的加权方式。 Real AdaBoost 算法具有连续置信度输出, 能够更好地刻画分类边界, 它的一些改进算法被提出并被应用于图像识别问题^[21]。

现有的关于 Real AdaBoost 算法的研究成果多为理论研究和用于图像识别问题的研究, 应用于软件可靠性预测的成果很少见到, 所以本文对 Real AdaBoost 算法的适用范围以及解决软件可靠性预测问题的有效性进行了研究。

Real AdaBoost 算法假设训练样本集为 $A = \{(a_1, b_1), \dots, (a_n, b_n)\}$, n 为样本总数, $b_i \in \{-1, +1\}$ 。弱分类器空间记为 H , h_t 为第 t 个分类器。算法的具体步骤如下:

步骤 1 初始化权值, $\omega_i = 1/n, i=1, 2, \dots, n$ 。

步骤 2 For $t=1, 2, \dots, T$ 。

1) 对 A 进行划分, $A = A_1 \cup A_2 \cup \dots \cup A_n$, 且 $i \neq j$ 时, $A_i \cap A_j = \emptyset$ 。

2)统计 A_i 中 -1 和 $+1$ 的累积样本权, $W^j =$

$$\sum_{i:(a_i \in A_j) \cap (b_i = -1)} w_i^j, W_+^j = \sum_{i:(a_i \in A_j) \cap (b_i = +1)} w_i^j, j = 1, 2, \dots, m。$$

3)定义基分类器 $h_i(a): \forall a \in A$, 令:

$$h_i(a) = \frac{1}{2} \ln(W_+^j + \gamma) / (W_-^j + \gamma), j = 1, 2, \dots, m \quad (3)$$

式中: γ 为平滑因子, $\gamma \in (0, 1)$ 。

4)选取 $h_i(a)$, 最小化 $Z_i = 2 \sum_{j=1}^m \sqrt{W_+^j W_-^j}$, 即:

$$h_i(a) = \arg \min_{h \in H} Z_i \quad (4)$$

5)权值调整:

$$w_i^{t+1} = w_i^t \times \frac{\exp(-b_i h_i(a_i))}{Z_i} \quad (5)$$

步骤 3 强分类器:

$$H(a) = \text{sign}(f(a)) \quad (6)$$

式中: $f(a) = \sum_{i=1}^T h_i(a)$ 。

实验结果表明: Real AdaBoost 算法并不能有效解决软件可靠性预测精度低和可信性差的问题, 因此, 本文首先利用 BP 神经网络个体对 AdaBoost 算法中构成集成的个体进行替换, 构建 Real BP-AdaBoost 算法。

2.2 WIE Real BP-AdaBoost 算法

BP 神经网络是一种能够进行概率输出的弱学习算法, BP 神经网络的输出记为 $O = [x_1, x_2, \dots, x_K]$, K 为训练样本的类数, x_i 为样本属于第 l 类的可能性, $l = 1, 2, \dots, K$ 。BP 神经网络中有多种传递函数, 为了与对比模型 SVM 保持一致, 使用“logsig”函数, 因此 $0 \leq x_i \leq 1$ 。令 $X = \sum_{l=1}^K x_l$, BP 神经网络的概率输出可以记为 $Y = [y_1 \ y_2 \ \dots \ y_K]$, $y_l = \frac{x_l}{X}$ 。以概率输出的 BP 神经网络为基分类器的 AdaBoost 算法, 本文将其命名为 Real BP-AdaBoost 算法。

由于同一分类器对不同的样本分类能力是不同的, 在 Real BP-AdaBoost 算法中对测试样本进行分类时并没有考虑此问题, 基分类器的加权系数只与分类错误率有关, 对所有的测试样本均赋予相同的加权值。为改进 Real BP-AdaBoost 算法中基分类器的加权方式, 本文引入信息熵的概念, 设计一种动态加权方法。

信息熵是为了解决对信息的量化度量问题而由香农于 1948 年提出的概念^[22]。信息熵越大表示信号分到各个类别的不确定性越强, 信号所包含的信息越多, 信息熵的计算公式为:

$$H(U) = - \sum_{i=1}^n p_i \log p_i \quad (7)$$

式中: n 表示信号总的类别数; p_i 表示信号分到第 i 类的概率。

BP 神经网络的概率输出文中已经给出, 当分到各类的概率差别较大时, 信息熵较小, 信息的确定性较高, 分类器此时应该被赋予较大的权值, 反之亦然。基于此, 本文设计一种分类器对不同样本分类的权值计算公式, 假设共有 m 个基分类器, 第 i 个基分类器的加权系数 ωh_i 为:

$$\omega h_i = \frac{\exp(-H_i(x))}{\sum_{j=1}^m \exp(-H_j(x))} \quad (8)$$

信息熵的准确性依赖于分类边界的准确性, 当一个基分类器分类较差, 分类边界与最优边界差别较大时, 单一的信息熵并不能准确反应基分类器对样本的分类能力。在此, 定义以分类器正确率为基准的另一加权系数, 第 i 个基分类器的加权系数 ωe_i 为:

$$\omega e_i = \frac{1 - \text{err}_i}{\text{err}_i} \quad (9)$$

式中: err_i 为第 i 个基分类器对训练样本的分类错误率。 ωe 表示分类器整体的分类能力, ωh 表示分类器对于某一具体样本分类结果的确定性, $\omega e \times \omega h$ 更能准确地表示基分类器对具体样本的分类能力, 本文将其称之为加权信息熵 (Weighted Information Entropy, WIE), 将基于 WIE 的 Real BP-AdaBoost 算法命名为 WIE Real BP-AdaBoost 算法。

不同于 Real BP-AdaBoost 算法的基分类器的加权方式, WIE Real BP-AdaBoost 算法的基分类器加权步骤如下:

步骤 1 根据公式 (9) 求出基分类器的整体加权值 ωe_i ;

步骤 2 求解基分类器对样本的个体加权值:

For $i = 1, 2, \dots, m$

For $j = 1, 2, \dots, n$

$$H_{ij} = - \sum_{j=1}^n p_{ij} \log p_{ij} \quad (10)$$

$$\omega h_{ij} = \frac{\exp(-H_{ij})}{\sum_{i=1}^m \exp(-H_{ij})} \quad (11)$$

则加权信息熵为: $\omega_{ij} = \omega e_i \times \omega h_{ij}$ 。

综上, WIE Real BP-AdaBoost 算法的具体步骤如下:

步骤 1 初始化权值: $w_i^1 = 1/m (i = 1, 2, \dots, m)$;

步骤 2 DO FOR $t = 1, 2, \dots, T$; 其中 T 为训练轮数, 即得到的弱分类器个数;

1)用弱学习算法 P , 基于样本权值 w_i^t 训练得到弱分类器 $h_t(x): x \rightarrow \{-1, +1\}$;

2)计算 $h_t(x)$ 的错误率:

$$\epsilon_t = \sum_{i=1}^m w_i^t (h_t(x_i) \neq y_i) \quad (12)$$

$$\text{令 } \beta_t = \epsilon_t / (1 - \epsilon_t), \alpha_t = \frac{1}{2} \log(1/\beta_t);$$

3)调整权值:

$$w_i^{t+1} = \frac{w_i^t}{z_t} \begin{cases} e^{-\alpha_t}, & h_t(x) = y \\ e^{\alpha_t}, & h_t(x) \neq y \end{cases} \quad (13)$$

式中: z_t 是权值 w_i^t 的归一化因子;

步骤 3 循环结束后, 最后的强分类器为:

$$H(x) = \text{sign}(f(x)) \quad (14)$$

其中:

$$f(x) = \sum_{i=1}^T \omega_{ij} h_t(x) \quad (15)$$

至此, WIE Real BP-AdaBoost 算法构建完成。

基于 WIE Real BP-AdaBoost 算法的软件可靠性预测见图 1。

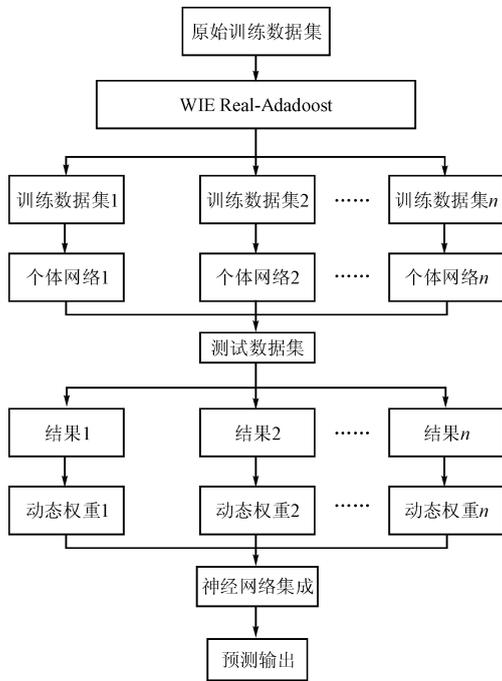


图 1 基于 WIE Real BP-AdaBoost 算法的软件可靠性预测示意图

Fig. 1 The sketch map of software reliability prediction based on WIE Real BP-AdaBoost

3 仿真实验及分析

3.1 实验数据

本次实验选取 2 组公开发表的软件失效数据 NTDS(Naval Tactical Data System, NTDS)数据和 LCCS(Large-scale Command and Control System,

LCCS)数据。NTDS 数据(见表 1)是美国海军战术数据系统在开发、测试和用户 3 个阶段发现的错误数和故障间隔时间,在这 3 个阶段分别发现了 26、5 和 3 共计 34 个数据,并以失效间隔时间的形式给出。LCCS 数据(见表 2)是美国空军大型指挥控制软件系统在 15 周测试时间内,观察到的每周失效数据,并以时间失效数的形式给出。

表 1 NTDS 失效间隔时间数据

Tab. 1 Time between failures of NTDS

失效	间隔	失效	间隔	失效	间隔	失效	间隔
1	9	10	7	19	6	28	47
2	12	11	1	20	1	29	12
3	11	12	6	21	11	30	9
4	4	13	1	22	33	31	135
5	7	14	9	23	7	32	258
6	2	15	4	24	91	33	16
7	5	16	1	25	2	34	35
8	8	17	3	26	1		
9	5	18	3	27	87		

表 2 LCCS 软件失效数据

Tab. 2 Software failure data of LCCS

周	失效数	周	失效数	周	失效数
1	203	6	71	11	27
2	186	7	54	12	42
3	183	8	57	13	55
4	47	9	80	14	62
5	46	10	64	15	11

3.2 实验设计

为验证 WIE Real BP-AdaBoost 算法的有效性,选取 SVM、BP 网络、Elman 网络、Real BP-AdaBoost 算法,同 WIE Real BP-AdaBoost 算法进行比较,设定最高迭代次数为 100 次。在获得算法取得最优解的迭代次数后,设置迭代次数为最优迭代次数。在进行实验时,分别画出各算法模型在 NTDS 数据数据下的累计失效时间图、在 LCCS 数据下的累计失效数图,分别画出各算法模型预测 NTDS 数据和 LCCS 数据的误差图,计算各模型的均方误差 MSE(Mean Square Error, MSE),并以表格的形式给出。

均方误差,用来评测算法预测软件可靠性的精度与可信性,其计算公式如下:

$$MSE = \sum_{i=1}^m \frac{(x_i - n(t_i))^2}{m} \quad (21)$$

式中: x_i 表示软件程序实际失效数; $n(t_i)$ 表示模型预测失效数, $i = 1, 2, \dots, m$, m 表示累计失效数。

MSE 的值越小,意味着模型的预测精度与可信性越高,模型的拟合性越好。

3.3 实验结果分析

首先分析各模型对 NTDS 数据的实验结果(见图 2、图 3 和表 3),从图 2 可以看出,各模型均拟合出了 NTDS 数据的累计失效时间的整体趋势,但 WIE Real BP-AdaBoost 算法模型的拟合结果最为精确;通过图 3 可以更为直观地观察到,WIE Real BP-AdaBoost 算法模型的误差在零基准线上下浮动最小,而其余对比模型表现出了不稳定性 and 预测误差较大的状态;在表 3(精确到小数点后 5 位)中 WIE Real BP-AdaBoost 算法模型的 MSE 最小,同样说明了 WIE Real BP-AdaBoost 算法的有效性。

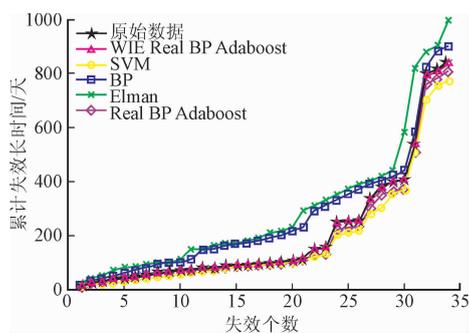


图 2 各模型预测 NTDS 数据的累计失效时间对比
Fig. 2 Comparison of NTDS's accumulative failure time of five forecast models

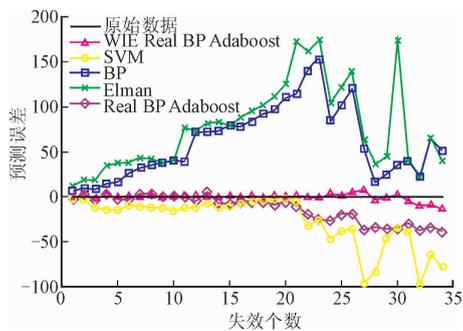


图 3 各模型预测 NTDS 数据的误差对比
Fig. 3 Comparison of five models' error during forecast NTDS data

表 3 NTDS 数据:各模型 MSE 对比

Tab. 3 NTDS data: comparison of five forecast models' MSE

模型	WIE Real BP-AdaBoost	SVM	BP	Elman	Real BP-AdaBoost
MSE	0.442 87	10.711 43	18.798 33	105.051 83	1.296 82

接着分析各模型对 LCCS 数据的实验结果见图 4、图 5 和表 4,虽然 LCCS 数据相对少于 NTDS 数据,这在某种程度上影响了 WIE Real BP-AdaBoost 算法模型的预测精度,但其预测结果仍然是所有模

型中最好的。图 4 直观反映出其拟合结果与原始数据最为接近;观察图 5, WIE Real BP-AdaBoost 算法模型的误差波动幅度较预测 NTDS 数据时要大,甚至在预测 LCCS 数据第 7 周时的误差比 BP 网络和 Elman 网络还要大,但这只是预测过程中的个别现象,对模型的整体预测效果并无影响,其误差波动在所有模型中依然最小;表 4 的 MSE 值表明,WIE Real BP-AdaBoost 算法在所有对比模型中表现最优。

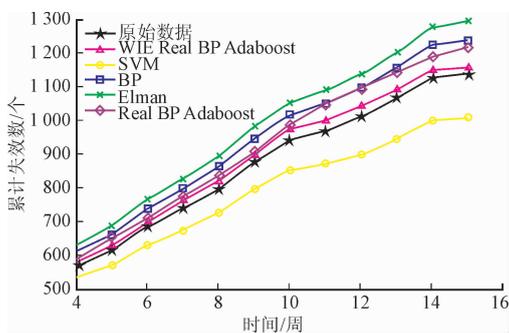


图 4 各模型预测 LCCS 数据的累计失效数对比
Fig. 4 Comparison of LCCS's accumulative failure of five forecast models

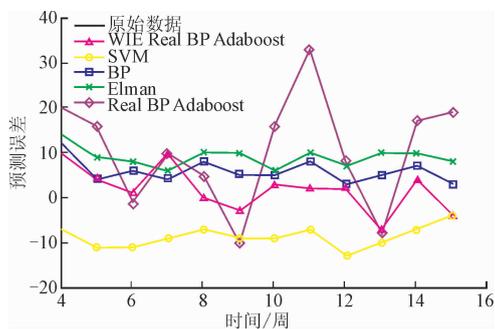


图 5 各模型预测 LCCS 数据的误差对比
Fig. 5 Comparison of five models' error during forecast LCCS data

表 4 LCCS 数据:各模型 MSE 对比

Tab. 4 LCCS data: comparison of five forecast models' MSE

模型	WIE Real BP-AdaBoost	SVM	BP	Elman	Real BP-AdaBoost
MSE	0.284 71	1.111 60	0.709 14	1.658 17	2.460 60

综上,通过 2 组数据对 WIE Real BP-AdaBoost 算法的有效性进行了验证,基本上解决了基于单一神经网络的软件可靠性模型预测精度低和可信性差的问题。

4 结语

本文通过对 Real AdaBoost 算法的分类器和加权方式进行改进,提出了 WIE Real BP-AdaBoost

算法,并与相关算法(Real BP-AdaBoost 算法)和其他算法模型进行比较,实验结果验证了新算法的有效性,达到了预期效果。

值得注意的是,通过对实验结果的分析,发现本文提出的算法对不同类型数据的预测效果存在差异,这说明本文所提出的模型的适用性不强。另外,本文实验所采取的数据量较少,不能突显出算法的优势。下一步工作应围绕探寻更具普适性的软件可靠性模型和将软件可靠性模型应用于大型软件工程 2 个方面来展开。

参考文献(References):

- [1] LYU M. Handbook of Software Reliability Engineering [M]. New York: IEEE Computer Society Press, McGraw Hill, 1996.
- [2] GUO P, LYU M. A Case Study on Stacked Generalization with Software Reliability Growth Modeling Data [C]// Proceedings of the 8th International Conference on Neural Information Processing. Shanghai, China, 2001: 3763-3769.
- [3] GUO P, LYU M. A New Approach to Optical Multilayer Learning Neural Network [C]// Proceedings of the 2001 International Conference on Artificial Intelligence. Las Vegas, Nevada, SUA, 2001: 426-432.
- [4] XING F, GUO P, LYU M. A Novel Method for Early Software Quality Prediction Based on Support Vector Machine [C]// Proceedings of the 16th International Symposium on Software Reliability Engineering. Chicago, Illinois, USA, 2005: 213-222.
- [5] LI S, YIN Q, GUO P. A Hierarchical Mixture of Software Reliability Model for Prediction [J]. Applied Mathematics and Computation, 2007, 185(2): 1120-1130.
- [6] KHOSHGOFTAAR T, SZABO R. Using Neural Networks to Predict Software Faults during Testing [J]. IEEE Transaction on Reliability, 1996, 45(3): 456-462.
- [7] HU Q, XIE M, NG S. Robust Recurrent Neural Network Modeling for Software Fault Detection and Correct Prediction [J]. Reliability Engineering and System Safety, 2007, 92(3): 332-340.
- [8] SU Y, HUANG C. Neural-Network-Based Approaches for Software Reliability Estimation Using Dynamic Weighted Combinational Models [J]. Journal of Systems and Software, 2007, 80(4): 606-615.
- [9] ALJAHDALI S H, SHETA A, RINE D. Prediction of Software Reliability: A Comparison between Regression and Neural Network Non-Parametric Models [J]. ACS/IEEE International Conference on Computer Systems and Applications, 2001, 22(10): 470-473.
- [10] ZHENG J. Predicting Software Reliability with Neural Network Ensembles [J]. Expert Systems with Applications, 2009, 36(2): 2116-2122.
- [11] 徐仁佐, 谢曼, 郑人杰. 软件可靠性模型及应用 [M]. 北京: 清华大学出版社, 1994.
XU R Z, XIE M, ZHENG R J. Software Reliability Model and Its Application [M]. Beijing: Tsinghua University Press, 1994. (in Chinese)
- [12] 黄锡滋. 软件可靠性、安全性与质量保证 [M]. 北京: 电子工业出版社, 2002: 126.
HUANG X Z. Software Reliability, Safety and Quality Assurance [M]. Beijing: Electronic Industry Press, 2002: 126. (in Chinese)
- [13] HANSEN L K, SALAMON P. Neural Network Ensembles [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1990, 12(10): 993-1001.
- [14] SOLLICH P, KROUGH A. Learning with Ensembles: How Over-Fitting Can be Useful [J]. Advances in Neural Information Processing Systems, 1996, 8: 190-196.
- [15] DIETTERICH T G. Machine-Learning Research [J]. AI Magazine, 1997, 18(4): 97-136.
- [16] HAN M, ZHU X, YAO W. Remote Sensing Image Classification Based on Neural Network Ensemble Algorithm [J]. Neurocomputing, 2012, 78(1): 133-138.
- [17] YU L, LAI K K, WANG S. Multistage RBF Neural Network Ensemble Learning for Exchange Rates Forecasting [J]. Neurocomputing, 2008, 71(16): 3295-3302.
- [18] ZHOU Z H, JIANG Y, YANG Y B, et al. Lung cancer Cell Identification Based on Artificial Neural Network Ensembles [J]. Artificial Intelligence in Medicine, 2002, 24(1): 25-36.
- [19] FREUND Y, SCHAPIRE R E. Experiments with a New Boosting Algorithm [J]. Proceedings of the 13th International Conference on Machine Learning, 1996, 13: 148-156.
- [20] SCHAPIRE R E, SINGER Y. Improved Boosting Algorithms Using Confidence-Rated Predictions [J]. Machine Learning, 1999, 37(3): 297-336.
- [21] 付忠良. 关于 Real AdaBoost 算法的分析与改进 [J]. 电子科技大学学报, 2012, 41(4): 545-551.
FU Z L. Analysis and Improvement on Real AdaBoost Algorithm [J]. Journal of University of Electronic Science and Technology of China, 2012, 41(4): 545-551. (in Chinese)
- [22] HILLBORN R C. Chaos and Nonlinear Dynamics [M]. Oxford: Oxford University Press, 1994.

(编辑: 徐楠楠)