

支持向量机在边坡稳定分析预测的应用

洪 勇^{1,2}, 邵珠山², 马 力³

(1. 西安建筑科技大学材料与矿资学院, 陕西 西安 710055; 2. 西安建筑科技大学土木工程学院, 陕西 西安 710055; 3. 西安科技大学能源学院, 陕西 西安 710054)

摘 要 目的 提出一种基于改进算法的支持向量机模型(PSO-SVM), 利用边坡的参数分析预测边坡稳定性. 方法 利用支持向量机有效解决小样本、高维数、非线性等问题的优势, 建立粒子群算法(PSO)优化的支持向量机模型, 粒子群算法优化支持向量机参数, 模型中边坡几何参数和强度参数: 边坡角 β 、边坡高度 H 、岩石容重 γ 、黏聚力 c 、内摩擦角 φ 以及孔隙水压力 r_u 作为输入参数, 边坡稳定性系数 FS 和边坡稳定状态 S 作为输出参数. 结果 PSO-SVM模型与网格搜索算法(GS)、遗传算法(GA)优化SVM模型以及人工神经网络ANN模型相比, 具有更高的分类精度和更强的预测能力. 结论 PSO-SVM模型能够准确地获得边坡的稳定性系数, 评价其稳定性, 在边坡稳定分析和预测中具有良好的实际应用价值.

关键词 边坡稳定; 粒子群算法; 支持向量机; 预测

中图分类号 TU457; P642.22

文献标志码 A

Application of a Support Vector Machine for Analysis and Prediction of Slope Stability

HONG Yong^{1,2}, SHAO Zhushan², MA Li³

(1. School of Materials and Mineral Resources, Xi'an University of Architecture and Technology, Xi'an, China, 710055; 2. School of Civil Engineering, Xi'an University of Architecture and Technology, Xi'an, China, 710055; 3. School of Energy Engineering, Xi'an University of Science and Technology, Xi'an, China, 710054)

Abstract: A slope stability analysis and prediction model based on PSO-SVM algorithm, which uses parameters of slopes, is put forward. As support vector machine (SVM) could effectively solve the small sample, high dimension and non-linear problems, the Particle Swarm Optimization and Support Vector Machine (PSO-SVM) coupling model was proposed to analyze the slope stability. The parameters of SVM was optimized by the PSO. The most influential factors on the stability of slope such as slope angle, slope height, unit weight, cohesion, friction angle and pore water pressure coefficient were considered as model inputs, the factor of safety (FS) or the stability status (S) were considered as model outputs. Compared the results of case with optimizing SVM based on grid search (GS), genetic algorithm (GA) and artificial neural networks (ANN) in details, the

收稿日期: 2017-04-09

基金项目: 国家自然科学基金项目(10772143); 陕西省科技统筹项目(2015TZC-G-8-9)

作者简介: 洪勇(1986—), 男, 博士研究生, 主要从事岩土工程、采矿工程方面研究.

results shows that PSO-SVM has the higher classification accuracy and greater prediction ability among the four algorithms, and is therefore considered most suitable for slope stability analysis. it can acquire slope safety factors and evaluate slope stability, which can be well applied to the analysis of slope stability.

Key words: slope stability; support vector machine; particle swarm optimization; prediction

边坡稳定一直是岩土工程界热门研究课题之一,涉及矿山、水利水电、公路铁路等工程领域,关系到企业的经济效益和人民生命财产安全.但是由于边坡工程的复杂性、不确定性、数据不完备性和非线性等特点^[1-3],对其稳定性进行评价和作出准确有效的预测预报往往比较困难,通过对经过工程实践检验的边坡案例进行分析研究具有十分重要的现实意义,从而探讨其边坡稳定性与参数之间的关系.目前,极限平衡法^[4-7]、数值分析法(有限元、边界元、离散元法等)等传统的解析法和数值方法广泛运用于边坡稳定性分析,来减少和降低由边坡失稳所造成的危害.然而,近年来随着计算机技术水平和人工智能的发展,粗糙集^[8-9]、支持向量机^[10](SVM)、人工神经网络(ANN)以及混合算法^[11]等方法应用于边坡稳定性分析,取得了良好的效果.

Lee等^[12]基于高速公路边坡的5个参数,使用ANN方法评价边坡稳定性,取得了良好的效果.Gordan^[13]等使用粒子群优化算法(PSO) PSO-ANN研究了均质边坡的抗震稳定性,并与ANN方法比较,结果表明PSO-ANN方法具有更好的相关性和准确率.Zhou^[14]建立了PSO-SVM耦合模型预测三峡库区八字门滑坡体位移,结果表明其预测结果与滑坡诱发参数有很好的相关性,预测值很好地与实测值相符合.Sakellariou^[1]等使用ANN方法研究了边坡参数对边坡稳定性的影响,并与标准的分析方法进行比较,表明ANN方法具有更好的收敛性.Samui^[2]等使用SVM研究了边坡稳定性问题,结果表明其比传统的ANN方法具有更

高的准确率.Li^[3]等使用GA优化SVM参数,建立了GA-SVM预测滑坡位移,结果表明其比传统的SVM方法具有更高的准确率、更小的均方根误差和更高的相关性系数.

以上所提到智能算法研究边坡的稳定性研究方法,是将边坡的稳定性研究分为两类:一类是判定边坡的稳定性状态稳定或者不稳定,这其实是一个分类问题;另一类是预测边坡的安全系数或者变形边坡的位移,这是一个回归问题.SVM具有扎实的理论基础,以结构风险最小化原理(SRMP)为基础,泛化能力强,能有效解决小样本、高维数、非线性等问题^[15].但是传统的SVM方法惩罚参数 C 和核函数参数 γ_{svm} 的选择确定关系到模型的预测推广泛化能力.PSO算法是Kennedy和Eberhart提出的基于群体智能的一种进化算法,已广泛应用于模型参数的选择和优化^[16-18].因此,笔者使用PSO寻优SVM参数,建立PSO-SVM耦合模型,对边坡稳定性进行分析预测,并与其他算法进行对比分析,评价本文所建立PSO-SVM模型的性能,结果表明,PSO-SVM模型分类准确率更高,预测边坡安全系数精度更高,推广泛化能力更强,研究结果对边坡稳定性分析和预测具有一定的理论和现实意义,能够应用于边坡稳定评价分析.

1 支持向量机原理

支持向量机于1995年由Vapnik和Corinna Cortes首先提出,是基于统计学习理论的机器学习方法,理论基础扎实,目前在计算机、生态学、医学、工程等诸多领域广泛应用.支持向量机分类是将数据映射到高维线

性特征空间,在高维空间寻找最优分类超平面,使得正反例的间隔边缘最大化.对于回归问题,假设有训练样本集 $\{(x_1, y_1), \dots, (x_l, y_l)\} \in (\mathbf{R}^n \times \mathbf{Y})^l$, 其中 $x_i \in \mathbf{R}^n$ 为样本输入变量, $y_i \in \mathbf{Y} = \mathbf{R}, i = 1, \dots, l$ 为对应的样本输出变量, l 为样本个数,则回归问题简化为最优优化问题:

$$\min_{w,b,\xi} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l (\xi_i + \xi_i^*), \quad (1)$$

约束条件为

$$y_i - \sum_{j=1}^n \sum_{i=1}^l (w_j \cdot y_{ji}) - b \leq \varepsilon + \xi_i, i = 1, 2, \dots, l; \quad (2)$$

$$\sum_{j=1}^n \sum_{i=1}^l (w_j \cdot y_{ji}) + b - y_i \leq \varepsilon + \xi_i^*, i = 1, 2, \dots, l; \quad (3)$$

$$\xi_i, \xi_i^* \geq 0, i = 1, 2, \dots, l. \quad (4)$$

求解最后可得回归函数为

$$f(x) = w \cdot x + b, w, x \in \mathbf{R}^n, b \in \mathbf{R}. \quad (5)$$

式中: n 为支持向量数; b 为偏执量; w 为高维空间的权值矢量; 考虑到允许的拟合误差和算法精度 ε , 引进松弛变量 ξ_i, ξ_i^* 和惩罚参数 C . 回归问题的实质是寻找 w 和 b , 使得对于样本以外的 x , 满足 $|f(x) - wx + b| \leq \varepsilon$, 等价于在式(2)、式(3)、式(4)约束下, 求式(1)的最小值优化问题. 对于非线性函数的回归问题, 利用核函数 $K(\cdot)$ 代替相应的内积, 免去高维空间内积计算的复杂性, 于是式(5)改写为

$$f(x) = w \cdot K(x_i \cdot x) + b. \quad (6)$$

核函数的种类有很多, 目前使用最多的有多项式核函数、径向基核函数(Radial Basis Function, RBF)、Sigmoid 核函数等. 研究表明, RBF 能够取得良好的计算效果, 因此笔者选用 RBF.

2 粒子群算法

在 SVM 模型中, 惩罚参数 C 和核函数参数 γ_{SVM} 关系到模型的预测准确程度, 笔者

利用 PSO 算法优化参数 C 和 γ_{SVM} . PSO 算法最早是 Kennedy 和 Eberhart 受到飞鸟集群活动的规律性启发, 于 1995 年提出的一种随机全局优化算法. PSO 算法的一般流程: 首先初始化一群群体规模为 m 的粒子, 其中包括随机位置和速度; 然后根据适应度函数(fitness function), 评价每个粒子的适应度, 也就是判断粒子位置的好坏; 然后对每个粒子, 将其目前的适应值与其个体历史最佳位置对应的适应值作比较, 找出历史最佳位置; 同时对每个粒子, 将其目前的适应值与全局最佳位置(目前所有粒子发现的最佳位置)对应的适应值作比较, 找出全局最佳位置; 最后根据公式更新每个粒子的速度与位置, 达到终止条件则结束搜索. 通常终止条件指的是最大迭代次数或者粒子搜索到的最佳位置的适应值的增量小于预定最小适应阈值.

笔者建立的 PSO - SVM 算法一般流程如下:

- (1) 根据数据集, 对数据进行归一化处理, 建立学习样本和测试样本.
- (2) 定义参数的搜索范围. 进行初始化设置, 设置群体规模、迭代次数, 参数的上下限.
- (3) 确定适应度函数. 建立支持向量机的学习预测模型, 计算个体的适应值. 文中选取均方根误差(MSE)作为适应度函数.
- (4) 更新粒子的位置和速度. 根据前面所述找出粒子的目前个体最好适应值, 并且与个体历史适应值和全局最优值比较, 如果目前的适应值更好, 则保存, 不然继续搜索.
- (5) 停止准则. 判断适应值或迭代次数是否满足要求.
- (6) 利用优化的 SVM 参数建立模型, 预测评估边坡的稳定性.

3 案例研究

影响边坡稳定性的主要因素是边坡的几何参数和强度, 比如边坡角 β 、边坡高度 H 、岩石容重 γ 、黏聚力 c 、内摩擦角 φ 以及孔隙

水压力系数 r_u ,文中研究的数据来源于文献[1],数据由 β 、 H 、 γ 、 c 、 φ 、 r_u 、边坡稳定性系数 FS 和边坡稳定状态 S 构成,其中稳定状态 S 列中0表示失稳,1表示稳定,具体如表1所示. β 、 H 、 γ 、 c 、 φ 、 r_u 作为模型的输入向

量, FS 或者 S 作为输出向量.在边坡稳定状态分析中,前30个样本作为学习样本,后16个样本作为预测样本,建立PSO-SVM边坡稳定性评价模型.

表1 边坡样本集

Table 1 Training and testing samples of slopes								
编号	$\gamma/(\text{kN}\cdot\text{m}^{-3})$	c/kPa	$\varphi/(^{\circ})$	$\beta/(^{\circ})$	H/m	r_u	S	FS
1	18.68	26.34	15	35	8.23	0	0	1.11
2	16.5	11.49	0	30	3.66	0	0	1
3	18.84	14.36	25	20	30.5	0	1	1.875
4	18.84	57.46	20	20	30.5	0	1	2.045
5	28.44	29.42	35	35	100	0	1	1.78
6	28.44	39.23	38	35	100	0	1	1.99
7	20.6	16.28	26.5	30	40	0	0	1.25
8	14.8	0	17	20	50	0	0	1.13
9	14	11.97	26	30	88	0	0	1.02
10	25	120	45	53	120	0	1	1.3
11	26	150.05	45	50	200	0	1	1.2
12	18.5	25	0	30	6	0	0	1.09
13	18.5	12	0	30	6	0	0	0.78
14	22.4	10	35	30	10	0	1	2
15	21.4	10	30.34	30	20	0	1	1.7
16	22	20	36	45	50	0	0	1.02
17	22	0	36	45	50	0	0	0.89
18	12	0	30	35	4	0	1	1.46
19	12	0	30	45	8	0	0	0.8
20	12	0	30	45	4	0	1	1.44
21	12	0	30	45	8	0	0	0.86
22	23.47	0	32	37	214	0	0	1.08
23	16	70	20	40	115	0	0	1.11
24	20.41	33.52	11	16	10.67	0.35	1	1.4
25	19.63	11.97	20	22	12.19	0.405	0	1.35
26	21.82	8.62	32	28	12.8	0.49	0	1.03
27	20.41	33.52	11	16	45.72	0.2	0	1.28
28	18.84	15.32	30	25	10.67	0.38	1	1.63
29	18.84	0	20	20	7.62	0.45	0	1.05
30	21.43	0	20	20	61	0.5	0	1.03
31	19.06	11.71	28	35	21	0.11	0	1.09
32	18.84	14.36	25	20	30.5	0.45	0	1.11
33	21.51	6.94	30	31	76.81	0.38	0	1.01
34	14	11.97	26	30	88	0.45	0	0.625
35	18	24	30.15	45	20	0.12	0	1.12
36	23	0	20	20	100	0.3	0	1.2
37	22.4	100	45	45	15	0.25	1	1.8
38	22.4	10	35	45	10	0.4	0	0.9
39	20	20	36	45	50	0.25	0	0.96
40	20	20	36	45	50	0.5	0	0.83
41	20	0	36	45	50	0.25	0	0.79
42	20	0	36	45	50	0.5	0	0.67
43	22	0	40	33	8	0.35	1	1.45
44	24	0	40	33	8	0.3	1	1.58
45	20	0	24.5	20	8	0.35	1	1.37
46	18	5	30	20	8	0.3	1	2.05

3.1 边坡稳定状态分析

在样本开始训练之前,为了消除数据尺寸的影响^[14],首先对输入数据按式(7)进行归一化:

$$\bar{x} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (7)$$

式中: \bar{x} 为归一化的结果; x 为原始数据; x_{\max} 为样本数据的最大值; x_{\min} 为样本数据的最小值.

随后运用 PSO 算法对 SVM 模型的惩罚参数 C 和核函数参数 γ_{SVM} 进行寻优,PSO 算法的参数设置:种群数量 20,进化代数 100,加速常量 $c_1 = 1.5, c_2 = 1.7$. 最后搜索的结果: $C_{\text{best}} = 97.172\ 4, \gamma_{\text{SVM best}} = 0.265\ 8$. 参数优化后的 SVM 模型对 16 组测试样本进行分类,在分类的结果中,所有的样本全部被正确分类,准确率 100%,结果如图 1 所示.

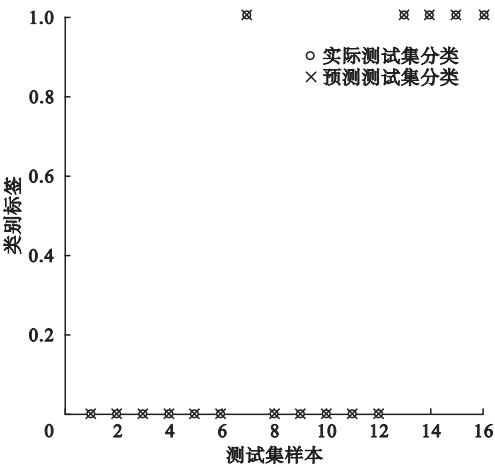


图 1 PSO-SVM 模型边坡稳定状态预测值与实际值比较

Fig. 1 Comparison between predicting results of PSO-SVM with actual values

为了验证 PSO-SVM 模型的优越性,同时与网格搜索算法(GS)、遗传算法(GA)优化的 SVM 模型进行对比分析,GS、GA 算法都是用来优化 SVM 模型的参数 C 和 γ_{SVM} ,最终预测结果如图 2、3 所示. GS、GA 算法优化后的 C 和 γ_{SVM} 值分别为 11.313 7, 38.763 5 和 0.353 6, 0.111 8, 分别有 3 个和 2 个样本被错误分类,对应的准确率分别 81.25% 和 87.50%,准确率分别比 PSO-

SVM 模型低 18.75% 和 12.50%. 从结果对比可以看出,PSO-SVM 模型相比另外两种模型,全局搜索能力更强,预测精度更高,用于边坡稳定性分析是可行的.

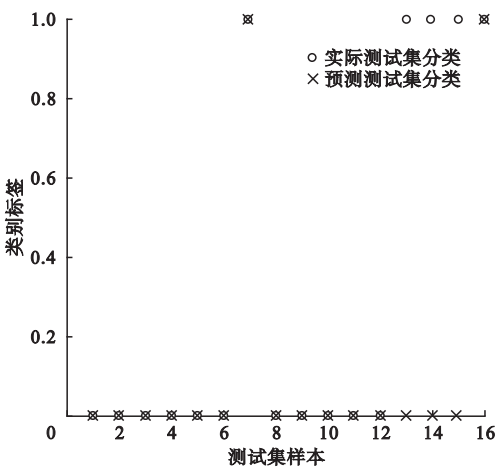


图 2 GS-SVM 模型边坡稳定状态预测值与实际值比较

Fig. 2 Comparison between predicting results of GS-SVM with actual values

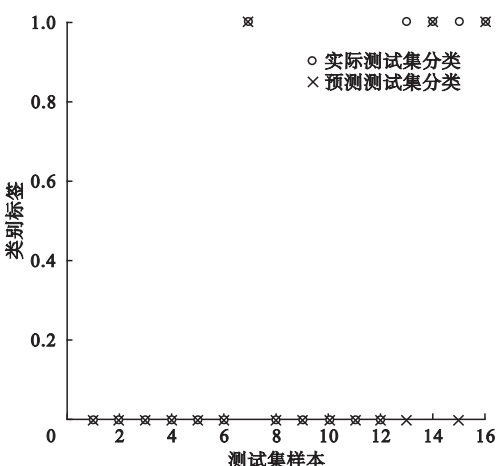


图 3 GA-SVM 模型边坡稳定状态预测值与实际值比较

Fig. 3 Comparison between predicting results of GA-SVM with actual values

3.2 边坡安全系数分析

以上探讨的是边坡稳定状态问题,用建立的 PSO-SVM 模型来判断边坡是否稳定,也就是分类问题.实际工程中,还需要知道边坡的安全系数,继续利用 PSO-SVM 模型来预测边坡的安全系数.在进行稳定性系数预

测前,为了消除输入参数对模型的影响,将原始的46个数据中黏聚力为0的样本剔除,理由是因为当黏聚力为0的样本边坡失稳时,属于平面滑坡,而不是严格的圆弧滑坡,当然

也有部分样本是稳定的,这样过滤后的样本数为31个^[1].笔者以表1中过滤后的31个样本为学习样本,以表2中的21个样本为预测样本^[1],对边坡的安全系数进行预测.

表2 边坡安全系数测试样本集
Table 2 Testing samples of slopes safety factor

编号	$\gamma/(\text{kN}\cdot\text{m}^{-3})$	C/kPa	$\varphi/(^{\circ})$	$\beta/(^{\circ})$	H/m	r_u	FS
1	21.00	20.00	40.00	40.00	12.00	0.00	1.84
2	21.00	45.00	25.00	49.00	12.00	0.30	1.53
3	21.00	30.00	35.00	40.00	12.00	0.40	1.49
4	21.00	35.00	28.00	40.00	12.00	0.50	1.43
5	20.00	10.00	29.00	34.00	6.00	0.30	1.34
6	20.00	40.00	30.00	30.00	15.00	0.30	1.84
7	18.00	45.00	25.00	25.00	14.00	0.30	2.09
8	19.00	30.00	35.00	35.00	11.00	0.20	2.00
9	20.00	40.00	40.00	40.00	10.00	0.20	2.31
10	18.85	24.80	21.30	29.20	37.00	0.50	1.07
11	18.85	10.34	21.30	34.00	37.00	0.30	1.29
12	18.80	30.00	10.00	25.00	50.00	0.10	1.40
13	18.80	25.00	10.00	25.00	50.00	0.20	1.18
14	18.80	20.00	10.00	25.00	50.00	0.30	0.97
15	19.10	10.00	10.00	25.00	50.00	0.40	0.65
16	18.80	30.00	20.00	30.00	50.00	0.10	1.46
17	18.80	25.00	20.00	30.00	50.00	0.20	1.21
18	18.80	20.00	20.00	30.00	50.00	0.30	1.00
19	19.10	10.00	20.00	30.00	50.00	0.40	0.65
20	22.00	20.00	22.00	20.00	180.00	0.00	1.12
21	22.00	20.00	22.00	20.00	180.00	0.10	0.99

样本训练前同样根据式(7)对数据进行归一化处理,然后利用 PSO 算法对 SVM 参数优化,PSO 算法初始设定的参数同前文一致,最后搜索的结果: $C_{\text{best}}=19.718\ 6$, $\gamma_{\text{SVM best}}=0.320\ 2$.同样,将 PSO-SVM 模型与网格搜索算法(GS)、遗传算法(GA)优化的 SVM 模型以及文献[1]的结果进行对比分析,结果见表3.

表3 各模型预测性能
Table 3 Prediction performance of various models

数据类型	MSE			
	PSO-SVM	ANN	GA-SVM	GS-SVM
31个学习样本	0.006 5	0.000 25	0.007 2	0.007 8
21个测试样本	0.054 4	0.110 0	0.057 4	0.054 5

从最终的结果可以看出,笔者建立的 PSO-SVM 模型预测结果要好于其他三种模型,预测精度较高,均方根误差较少.虽然在学习样本上 ANN 模型的均方根误差较 PSO-SVM 的小,但是在测试样本中 ANN 模型的均方根误差反而更大,说明 ANN 模型的推广泛化能力弱,预测效果较差;同时与 GS、GA 优化的 SVM 相比,PSO-SVM 学习样本和测试样本的均方根误差都要小于两者,说明 PSO 算法优化效果较好,推广泛化能力强,能够用于边坡安全系数的分析预测中.

4 结 语

笔者建立了粒子群优化的支持向量机耦

合模型,利用粒子群算法优化支持向量参数,相比于其他的支持向量机模型,全局搜索能力更强,建立的模型推广泛化能力更好。模型运用于边坡稳定性状态判断和安全系数预测中,相比于GS-SVM、GA-SVM以及ANN模型,PSO-SVM模型分类准确率更高,预测边坡安全系数精度更高,推广泛化能力更强,能够应用于边坡稳定评价分析。

参考文献

- [1] SAKELLARIOU M G, FERENTINOU M D. A study of slope stability prediction using neural networks[J]. Geotechnical & geological engineering, 2005, 23(4): 419.
- [2] SAMUI P. Slope stability analysis: a support vector machine approach [J]. Environmental geology, 2008, 56(2): 255 - 267.
- [3] LI X Z, KONG J M. Application of GA-SVM method with parameter optimization for landslide development prediction [J]. Natural hazards & earth system sciences, 2013, 1(5): 5295 - 5322.
- [4] FELLENIUS W. Calculation of stability of earth dams [J]. Transactions second congress on large dams, 1936, 4: 445.
- [5] BISHOP A W. The use of slip circle in the stability of slopes [J]. Geotechnique, 1955, 5(1): 7 - 17.
- [6] BISHOP A W, MORGENSTERN N R. Stability coefficients for earth slopes [J]. Geotechnique, 1960, 10(4): 129 - 150.
- [7] MORGENSTERN N R, PRICE V E. The analysis of the stability of general slip surfaces [J]. Geotechnique, 1965, 15(1): 79 - 93.
- [8] 闫长斌. 边坡稳定性预测的粗糙集-距离判别模型及其应用 [J]. 工程地质学报, 2016, 24(2): 204 - 210.
(YAN Changbin. Rough set-distance discriminant analysis model of slope stability prediction and its application [J]. Journal of engineering geology, 2016, 24(2): 204 - 210.)
- [9] 李云, 刘霁. 基于 RS-CPM 模型的边坡失稳灾害预测及应用 [J]. 中南大学学报(自然科学版), 2013, 44(7): 2971 - 2976.
(LI Yun, LIU Ji. Slope instability disaster forecast and its application based on RS-CPM model [J]. Journal of central south university (science and technology), 2013, 44(7): 2971 - 2976.)
- [10] 俞俊平, 陈志坚, 武立军, 等. 基于蚁群算法优化支持向量机的边坡位移预测 [J]. 长江科学院院报, 2015, 32(4): 22 - 27.
(YU Junping, CHEN Zhijian, WU Lijun, et al. Forecasting slope displacement based on support vector machine optimized by colony algorithm [J]. Journal of Yangtze river scientific research institute, 2015, 32(4): 22 - 27.)
- [11] 李胜, 韩永亮. 基于 MFOA-SVR 露天矿边坡变形量预测研究 [J]. 中国安全生产科学技术, 2015, 11(1): 11 - 16.
(LI Sheng, HAN Yongliang. Research on forecasting of slope deformation in open-pit mine based on MFOA-SVR [J]. Journal of safety science and technology, 2015, 11(1): 11 - 16.)
- [12] LEE T L, LIN H M, YUH-PIN L U. Assessment of highway slope failure using neural networks [J]. Journal of Zhejiang university, 2009, 10(1): 101 - 108.
- [13] GORDAN B, ARMAGHANI D J, HAJIHASANI M, et al. Prediction of seismic slope stability through combination of particle swarm optimization and neural network [J]. Engineering with computers, 2015, 32(1): 1 - 13.
- [14] ZHOU C, YIN K, CAO Y, et al. Application of time series analysis and PSO-SVM model in predicting the Bazimen landslide in the three gorges reservoir, China [J]. Engineering geology, 2016, 204: 108 - 120.
- [15] VAPNIK V N. The nature of statistical learning theory [J]. IEEE transactions on neural networks, 1995, 10(5): 988 - 999.
- [16] KENNEDY J, EBERHART R. Particle swarm optimization [C]. IEEE international conference on neural networks, Germany: Springer. 1995.
- [17] OLSSON A E. Particle swarm optimization: theory, techniques, and applications [J]. Engineering tools techniques & tables, 2011, 34: 21 - 28.
- [18] NIETO P J G, GARCÍA-GONZALO E, FERNÁNDEZ J R A, et al. A hybrid PSO optimized SVM-based model for predicting a successful growth cycle of the *Spirulina platensis*, from raceway experiments data [J]. Journal of computational & applied mathematics, 2016, 291(1): 293 - 303.