A Predict Method of Water Pump Operating State Based on Improved Particle Swarm Optimization of Support Vector Machine

Jian Pan*, Yujiang Li and Panfeng Wu
Zhijiang College of Zhejiang University of Technology, 958 Yuezhou Avenue, Keqiao, Shaoxing, China
*Corresponding author email: pj@zjut.edu.cn

Abstract. In order to improve prediction accuracy of water pump operating state, a chaotic prediction model of the pump vibration data based on improved particle swarm optimization of support vector machine is proposed in this paper. Firstly, a grouping optimization strategy particle swarm algorithm based on cosine function is proposed. Then, the training set is obtained on the time series of vibration data by phase space reconstruction. Secondly, The improved particle algorithm is used to optimize the penalty parameters, insensitive loss coefficient and width parameters of support vector machine. Then, a prediction model of vibration data is established by using support vector machine combined with training set and optimal parameters. Finally, the operating state of the pump is predicted according to pump vibration measurement and evaluation method. Compared with the method of linear decreasing weight strategy, the method proposed in this paper is more accurately.

1. Introduction
The vibration is one of important indexes judging water pump operating state, and the vibration intensity reflects the vibration velocity and the amount of energy. At the same time, the vibration intensity is less affected by the vibration frequency. If the trend of vibration intensity could be accurately predicted, the operating state of the pump can be predicted for a period of time in the future so as to make some countermeasures in advance.

In recent years, the methods to monitor and predict pump operating state were studied by scholars. For instance, John E.McInroy[1] proposed an algorithm for condition monitoring and fault diagnosis on motor instantaneous power and current spectrum of oil pump. Hack-Eun Kim[2] proposed a health state probability estimation algorithm to evaluate the remaining life of the device. D.Raheja[3] proposed a general state maintenance architecture to determine operating state of the equipment based on data fusion and data mining technology. Liu Wencai[4] applied the grey theory to the fault diagnosis and prediction of sliding bearing of centrifugal water injection pump and evaluated the operating state of the equipment. Zhao Peng[5] applied a hidden Markov model to fault diagnosis of centrifugal pump and developed a state monitoring and fault diagnosis system. Wu Xingwei[6] used support vector machine(SVM) algorithm to predict the bearing temperature of boiler feed pump.

The time series of vibration intensity show a very strong nonlinear dynamic phenomenon, namely chaos phenomenon. It is difficult to obtain a high predict accuracy because of chaos phenomenon. SVM is widely used because of its good fitting effect and generalization ability in many methods. However, it is necessary to optimize the related parameters to make the SVM fit better. The existing parameter optimization algorithms include grid search algorithm and intelligent search algorithm, and intelligent search algorithms include particle swarm algorithm, genetic algorithm, fish swarm...
algorithm, ant colony algorithm, etc. The grid search method is an exhaustive algorithm, so the running time is too long when the search scope and dimension are large. Particle swarm optimization (PSO) is widely used in intelligent search algorithm because it’s simple and has fewer parameters to set. However, the learning factor and inertia weight in PSO will directly affect the optimization result. Therefore, how to optimize the learning factor and inertia weight becomes particularly important.

In order to improve the prediction accuracy of pump vibration intensity, this paper firstly improves the learning factor and inertia weight selection strategy in PSO, and proposes a grouping optimization PSO algorithm based on cosine function. Then, the improved particle swarm optimization algorithm is used to optimize the parameters of support vector machine, and the improved particle support vector machine chaotic prediction model was established to predict the vibration intensity data. Finally, the operating state of the pump equipment is predicted and evaluated by using the predicted vibration intensity according to pump vibration measurement and evaluation method.

2. Improved PSO Algorithm

A Group Optimization Strategy (GOS) based on cosine function is proposed in this paper. In each iteration of PSO, $c_1$, $c_2$ and $w$ are grouped into a Group and optimized simultaneously. The method can be described as:

\[
\begin{align*}
    &c_1 = (c_{1,\text{start}} - c_{1,\text{end}}) \cos \frac{\pi t}{2T} + c_{1,\text{end}} \\
    &c_2 = (c_{2,\text{end}} - c_{2,\text{start}}) \cos \pi \left(1 + \frac{t}{2T}\right) + c_{2,\text{end}} \\
    &w = (w_{\text{start}} - w_{\text{end}}) \cos \frac{\pi t}{2T} + w_{\text{end}}
\end{align*}
\]

Where, $c_{1,\text{start}}$ and $c_{1,\text{end}}$ are the value range of $c_1$, and $c_{2,\text{start}}$ and $c_{2,\text{end}}$ are the value range of $c_2$. It can be seen from the above equation that when $t$ is small, cosine function is in a slow decline stage. At this time, $c_1$ and $w$ are larger, $c_2$ is smaller, and PSO can maintain longer global search ability. As $t$ increases, cosine function transits to fast decline stage. At this time, $c_1$ and $w$ are smaller and $c_2$ is larger. So the PSO can enter the local search stage more quickly, and the probability of finding the global optimal value is improved.

In order to prove the effectiveness of the improved PSO algorithm proposed, in this section, PSO with fixed learning factors and weights (standard PSO), linear decreasing weight PSO (LDW-PSO) and grouping optimization strategy based on cosine function proposed in this paper (GOS-PSO) are respectively used to optimize the two classical test functions, and then the optimization effect and convergence ability of the three algorithms are compared. Inertia weight calculation formula in LDW-PSO is as follows:

\[
w = \frac{(w_{\text{start}} - w_{\text{end}})(T-t)}{T} + w_{\text{end}}
\]

Where, $w_{\text{start}}$ is the upper limit of $w$ and $w_{\text{end}}$ is the lower limit of $w$. The two test functions are as follows:

Sphere function: \[f(x) = \sum_{i=1}^{d} x_i^2\] \hspace{1cm} (3)

Griewank function: \[f(x) = \sum_{i=1}^{d} \frac{x_i^2}{4000} - \prod_{i=1}^{d} \cos \left(\frac{x_i}{\sqrt{i}}\right) + 1\] \hspace{1cm} (4)

Where, Sphere function is single-peak function, Griewank function is multi-peak function. They both achieve the global minimum 0 when $x_i = 0$.

For the three PSO algorithms to be tested, the common parameters are uniformly set as: the number of iterations $T = 1000$, the number of particles $M = 40$. For Sphere function, other parameters are set as: dimension of Sphere function $d = 20$, particle position range is $[-100, 100]$, particle velocity range is $[-30, 30]$. For Griewank function, other parameters are set as follows: dimension of Griewank function $d = 20$, particle position range is $[-600, 600]$, particle velocity range is $[-180, 180]$. The value of $c_1$ [9] is that $c_1 = 2.05$ in standard PSO and LDW-PSO, $c_{1,\text{start}} = \ldots$
2.75 and $c_{1,\text{end}} = 1.25$ in GOS-PSO. The value of $c_2$ [9] is that $c_2 = 2.05$ in standard PSO and LDW-PSO, $c_{2,\text{start}} = 0.5$ and $c_{2,\text{end}} = 2.25$ in GOS-PSO. The value of $w$ is that $w = 0.75$ in standard PSO, $w_{\text{start}} = 0.9$ and $w_{\text{end}} = 0.4$ in LDW-PSO and GOS-PSO. For each test function, the three PSO algorithms were executed 500 times respectively, and then the time of obtaining the minimum value of each PSO algorithm in 500 times of running and the average number of iterations converging to the minimum value are counted. The optimization results are shown in Table 1. The minimum times of convergence to the three gradients within 5, 1 and 0.2 of the three PSO algorithms as well as the average iteration times of convergence to the minimum are calculated in Table 1.

Table 1. Convergence of three PSO algorithms

| Convergence Condition | standard PSO | LDW-PSO | GOS-PSO |
|-----------------------|-------------|---------|---------|
|                       | Sphere      | Sphere  | Sphere  |
|                       | Griewank    | Griewank| Griewank|
| Convergence times($\leq$5) | 12          | 14      | 212     |
| Average number of iterations ($\leq$5) | 641.41      | 215.01  | 480.05  |
| Convergence times($\leq$1) | 1           | 1       | 47      |
| Average number of iterations ($\leq$1) | 742         | 496     | 524.15  |
| Convergence times($\leq$0.2) | 0           | 0       | 4       |
| Average number of iterations ($\leq$0.2) | -           | -       | 553.75  |

The convergence of LSW-PSO and GOS-PSO is significantly better than that of standard PSO, no matter for Sphere function or Griewank function optimization. Comparing the optimization of LSW-PSO and GOS-PSO, it is found that although the partial average iteration times of GOS-PSO are slightly higher than that of LSW-PSO, but the convergence is obviously better than that of LSW-PSO, especially the convergence times within 5 and 1, which is almost 10 to 20 times of the convergence times of LSW-PSO. It is obviously worthwhile to increase the probability of finding the optimal solution by several times by slightly increasing the number of iterations. Therefore, a grouping optimization strategy based on cosine function proposed in this paper is effective.

3. A Prediction Model Based on Improved PSO Optimization of SVM

Suppose that the time series of vibration intensity on a certain position of water pump is $X_k = \{x_1, x_2, ..., x_l \}$, where $K$ represents the position number ($k>0$). The time series of vibration intensity is smoothed to $Y_k = \{y_1, y_2, ..., y_l \}$. Assuming that the delay time of $Y_k$ is $\tau$ and the embedding dimension is $m$, the phase space reconstruction of $Y_k$ is carried out to obtain the training set $S_k = \{D_1, D_2, ..., D_l, ..., D_{n-(m-1)\tau}\}$, where $D_l = \{y_1, y_{1+\tau}, y_{1+2\tau}, ..., y_{l+(m-1)\tau}\}$ is phase points.

The predict steps of the improved PSO optimization of SVM are as follows:

1. $X_k$ is collected and then de-trended and smoothed to obtain $Y_k$.
2. The delay time and embedding dimension of $Y_k$ are calculated, and the phase space of $Y_k$ is reconstructed to obtain $S_k$.
3. The improved PSO parameters are set including the total number of particles $n$, the maximum number of iterations $T$, the initial position of particles, the initial velocity of particles, the local optimal value of particles and the global optimal value of particles. The range of penalty parameter $C$ in SVM, insensitive loss coefficient $\epsilon$ and width parameter $g$ in Gaussian kernel function are set.
4. The particle fitness is calculated and the local and global optimal values are updated.
5. Learning factors $c_1, c_2$ and inertia weight $w$ are calculated and updated according to Equation 1.
(6) Calculate and update particle position and velocity.
(7) If the fitness value meets the requirements or the maximum number of iterations has been reached, the optimization is finished and $C$, $\varepsilon$ and $g$ are obtained from the global optimal position; otherwise, step (4) is performed.
(8) The $C$, $\varepsilon$ and $g$ obtained by the optimization are brought into SVM, and the final SVM prediction model is obtained by combining the training set. Finally, the performance of SVM prediction model is analyzed by using the test set.

4. Experiment and Result Analysis

4.1. Data Preprocessing
We collected 840 sets of vibration intensity data and stored in $X_1$, $X_2$ and $X_3$ which respectively represents the data of drive end bearing (DEB), pump shell (PS) and inlet flange (IF) on water pump for experiment. Five-point cubic smoothing method [10] was used to preprocess $X_1$, $X_2$ and $X_3$ respectively, and the results are $Y_1$, $Y_2$ and $Y_3$. The first 805 vibration intensity data of $Y_1$, $Y_2$ and $Y_3$ were used as the training set, and the last 35 vibration intensity data were used as the test set. The vibration intensity data in $Y_1$, $Y_2$ and $Y_3$ sets are normalized to the range of $[-1,1]$, which can make the data comparable. The calculation formula is as follows:

$$y_i^* = 2 \frac{y_i - y_{\text{min}}}{y_{\text{max}} - y_{\text{min}}} - 1$$

Where $y_{\text{max}}$ and $y_{\text{min}}$ are the maximum and minimum values respectively.

4.2. Analysis of Chaotic Characteristics of Data
Mutual information method [11], Cao method [12] and Wolf method [13] are used to calculate the delay time $\tau$, embedding dimension $m$ and maximum Lyapunov exponent of normalized vibration intensity data in $Y_1$, $Y_2$ and $Y_3$ respectively. The results are shown in Table 2 and the vibration intensity of different parts of the pump has chaotic characteristics since Maximum Lyapunov exponent > 0.

| Data item | $Y_1$ | $Y_2$ | $Y_3$ |
|-----------|-------|-------|-------|
| $\tau$    | 2     | 2     | 2     |
| $m$       | 8     | 9     | 8     |
| Maximum Lyapunov exponent | 0.059592 | 0.048477 | 0.075810 |

The training set $S_k$ (k=1,2,3) can be obtained by using $\tau$ and $m$ to reconstruct the phase space of $Y_k$ (k=1,2,3) where the matrix form of $S_k$ is shown in Equation 6. The first $m - 1$ pieces of data of each phase point in $S_k$ are used as the input and the $m - th$ data as the output of the model.

$$S_k^T = \begin{pmatrix} D_1 \\ D_2 \\ \vdots \\ D_{824} \end{pmatrix}^T = \begin{pmatrix} y_1 & y_3 & \cdots & y_{17} \\ y_2 & y_4 & \cdots & y_{18} \\ \vdots & \vdots & \ddots & \vdots \\ y_{824} & y_{826} & \cdots & y_{940} \end{pmatrix}^T$$

4.3. Model Evaluation Index
The mean square error (MSE), mean absolute error (MAE) and mean relative error (MRE) are used to evaluate the prediction results. The calculation formula of each evaluation index is as follows:

$$MSE = \frac{1}{L} \sum_{i=1}^{L} (y_i - y'_i)^2$$

$$MAE = \frac{1}{L} \sum_{i=1}^{L} |y_i - y'_i|$$

$$MRE = \frac{1}{L} \sum_{i=1}^{L} \left| \frac{y_i - y'_i}{y_i} \right|$$
Where, \( y_i \) is the actual vibration data, \( y'_i \) is the predicted data and \( L \) is the number of data.

### 4.4. Prediction Results and Analysis

GOS-PSO and LDW-PSO were used to optimize the \( C, \varepsilon \) and \( g \) of SVM. The number of iterations of PSO is set to 100, and the number of particles is set to 25. In GOS-PSO, \( c_{1,\text{start}} = 2.75, c_{1,\text{end}} = 1.25, c_{2,\text{start}} = 0.5, c_{2,\text{end}} = 2.25, w_{\text{start}} = 0.9, w_{\text{end}} = 0.4 \). In LDW-PSO, \( c_1 = c_2 = 2.05, w_{\text{start}} \) and \( w_{\text{end}} \) are the same as GOS-PSO. The results are shown in Table 3.

**Table 3. Optimization results of parameters**

| Data item | GOS-PSO | LDW-PSO | GOS-PSO | LDW-PSO | GOS-PSO | LDW-PSO |
|-----------|---------|---------|---------|---------|---------|---------|
| on DEB    | 1       | 1       | 0.1578  | 0.2323  | 0.4350  | 0.4489  |
| on PS     | 4.1213  | 1.9289  | 0.1536  | 0.3099  | 0.3461  | 0.1552  |
| on IF     | 173.9420| 1.2456  | 0.3240  | 0.4541  | 0.0500  | 0.4386  |

The optimal parameters in Table 3 are put into SVM, and the SVM is trained separately with the training set. The GOS-PSO-SVM and LSW-PSO-SVM prediction models were used to predict the data in the test set respectively, and the prediction results were shown in Figure 1 and Figure 2. The evaluation indexes of the prediction model are shown in Table 4.

![Figure 1. GOS-PSO-SVM prediction results and prediction error diagram](image1)

![Figure 2. Prediction results and prediction error diagram of LDW-PSO-SVM](image2)
It can be seen that MSE, MAE and MRE of GOS-PSO-SVM are the smallest among the three prediction models and GOS-PSO-SVM has the best predict effect. In conclusion, the improved PSO-SVM data prediction model has high prediction accuracy and generalization ability, and can accurately predict the vibration intensity of different parts of the water pump.

### Table 4. Comparison of evaluation indexes of vibration intensity

| Prediction Model | MSE       | MAE       | MRE       |
|------------------|-----------|-----------|-----------|
|                  | DEB | PS | IF | DEB | PS | IF | DEB | PS | IF |
| GOS-PSO-SVM      | 0.0062% | 0.035% | 0.01% | 0.64% | 1.51% | 0.87% | 4.8% | 5.74% | 9.61% |
| LDW-PSO-SVM      | 0.0066% | 0.04%  | 0.012% | 0.67% | 1.60% | 0.94% | 5.02% | 6.01% | 10.43% |

### 5. Prediction of Pump Operating State

In this section, pump vibration measurement and evaluation method (GB/T 29531-2013)[14] and the evaluation experience of water supply enterprises will be used to carry out a predictive evaluation of the operating state of water pump based on the predicted vibration intensity data. The process of judging the pump operating state can be described as follows: First of all, the category of the pump unit needs to be determined. Secondly, the vibration intensity level table of water pump is compared, and the points of different level are counted from vibration intensity data. Thirdly, multiply the points of different levels by the corresponding weight of the vibration intensity to get the points allocated according to the weight. Next, add up the points of the same level, and calculate their ratios. If the ratio of the maximum level is more than 20%, this level is the maximum effective intensity level. If the ratio of the maximum level is no more than 20%, this level is the remaining intensity level of the maximum intensity level, the cycle continues to judge, until found. Finally, the operating state of the pump equipment can be obtained according to the intensity level.

We use the weight values and the predicted data in Section 3, Section 4 to complete the prediction of the pump operating state. Because the rated speed of the pump unit is 1750rpm and pump center height is 243mm, according to the category of pump table, we can judge that the pump unit belongs to the second class. The levels of the predicted and actual data are mainly distributed on the three levels of 0.11, 0.18 and 0.28. For each intensity level, the predicted data points on this level are basically the same as the actual data points. Multiply the points by the corresponding weight values and add them up, then calculate the proportion of each level. The results are shown in Table 5.

### Table 5. The number of points belonging to different intensity classes

| Intensity level | 0.11 Predicted | 0.18 Predicted | 0.28 Predicted | 0.11 Actual | 0.18 Actual | 0.28 Actual |
|-----------------|----------------|----------------|----------------|-------------|-------------|-------------|
| DEB             | 0.75           | 1.5            | 0              | 25.5        | 24.75       | 0           |
| PS              | 0              | 0              | 3.75           | 3.9         | 1.35        | 1.35        |
| IF              | 2.8            | 2.8            | 0.7            | 0.7         | 0           | 0           |
| Sum points      | 3.55           | 4.3            | 29.95          | 29.35       | 1.5         | 1.35        |
| Ratios          | 10.14          | 12.29          | 85.57          | 83.86       | 4.28        | 3.86        |

It can be seen that the sum points and Ratios of the predicted data are very close to those of actual data, which proves that the prediction model has high accuracy. The intensity level that the maximum intensity grade is 0.28, but the sum points and ratios of predicted data and actual data are 4.28% and 3.86% respectively, both of which are less than 20%. Therefore, 0.28 is not the maximum effective intensity level, and 0.18 is the maximum effective intensity grade. Compared with the vibration level evaluation table of the pump, it can be determined that the vibration level of the pump unit is A which is excellent and suggest the sustainable and stable operation.
6. Conclusion
In order to improve the accuracy of pump operating state prediction, the GOS-PSO-SVM is proposed for parameter optimization of support vector machine in this paper. Firstly, the learning factor and inertia weight selection strategy in standard PSO algorithm is improved, and a grouping optimization strategy (GOS) based on cosine function is proposed to optimize the learning factor and inertia weight selection. Meanwhile, the GOS-PSO proposed in this paper is compared with the standard PSO and LSW-PSO. Then, the actual vibration intensity data of the pump are pre-processed, the delay time and embedding dimension are calculated, and the chaotic characteristics are judged. The phase space of the data is reconstructed by using the delay time and embedding dimension, and the training set is obtained. Next GOS-PSO is used to optimize the penalty parameters, insensitive loss coefficient and width parameters in SVM. We compared the GOS-PSO-SVM with LSW-PSO-SVM model on the train set, and the results shows that the GOS-PSO-SVM has better accuracy. Finally, the predicted data are used to predict the operation state of the water pump, and the results shows that the prediction model proposed in this paper has a certain practicality.

7. Acknowledgment
This work was supported by the Basic Public Welfare Research Project of Zhejiang Province, China (Grant No. LGF20F020015).

8. References
[1] McInroy J and Legowski S 2001 Using power measurements to diagnose degradations in motor drive power systems: a case study of oilfield pump jacks IEEE transactions on industry applications 37(06) pp 1574-81
[2] Kim H, Hwang S, Tan A, Mathew J and Choi B 2012 Integrated approach for diagnostics and prognostics of HP LNG pump based on health state probability estimation Journal of mechanical science and technology 26(11) pp 3571-85
[3] Raheja D, Llinas J, Nagi R and Romanowski C 2006 Data fusion/data mining-based architecture for condition-based maintenance International journal of production research 44(14) pp 2869-87
[4] Wencai L 2010 Research on fault diagnosis and prediction of sliding bearing of centrifugal pump based on Grey Theory. China University of Petroleum, Beijing
[5] Peng Z 2011 Research on vibration fault diagnosis method and system realization of centrifugal pump North China Electric Power University
[6] Xingwei W and Daocai C 2009 Dynamic Forecast of Temperature State for Water Feeding Pump Bearings 2 pp 53-57
[7] Cristianini N and Shawetaylor J 2001 An Introduction to Support Vector Machines and Other Kernel-based Learning Methods Kybernetes 30(1) pp 103-115
[8] Kennedy J and Eberhart R 1995 Particle swarm optimization Proceedings of ICNN’95 - International Conference on Neural Networks 4 pp 1942-48
[9] Xiang F, Guolong C and Wenzhong G 2006 Settings and Experimental Analysis of Acceleration Coefficients in Particle Swarm Optimization Algorithm Journal of Jimei University(Natural Science) 11(02) pp 164-151
[10] Wuwei S, Biao C and Jianfeng W 2013 Study on reverse deduction of reservoir-inflow based on cubical smoothing algorithm with five-point approximation Water Resources and Hydropower Engineering 44(12) pp 100-102
[11] Fraser A and Swinney H 1986 Independent coordinates for strange attractors from mutual information Physical Review A 33(02) pp 1134-40
[12] Cao L 1997 Practical method for determining the minimum embedding dimension of a scalar time series Physica D 110(01-02) pp 302-10
[13] Wolf A, Swift B and Swinney J 1985 Determining Lyapunov exponents from a time series Physica D Nonlinear Phenomena 16(09) pp 302-310
[14] Yuyan Z and Weiwei L et al 2013 GB/T 29531-2013 Methods of measuring and evaluating vibration of pumps (Beijing: Standards Press of China)