A Integrated IFCM-MPSO-SVM Model for Forecasting Equipment Support Capability

Yong Shuai1*, Tai-liang Song2, Jian-ping Wang1, Hong Shen1, and Wen-bin Zhan3

1 Technical Support Engineering Faculty of Armored Forced Engineering Academy, Beijing 100072, China, CHN
alexshuai@sina.com
dividhand@126.com
riverbird@sohu.com

2 China Defense Science & Technology Information Center, Beijing 100142, China, CHN
songtl@126.com

3 Communication Station of Air Force Equipment Research Academy, Beijing 100085, China, CHN
11416248@qq.com

Received 20 October 2015; Revised 25 January 2016; Accepted 24 May 2016

Abstract. For the sake of improving the accuracy for forecasting equipment support capability, aiming at the problems in support vector machine forecast model, this paper improved fuzzy C-means clustering algorithm about outliers operation and optimization of distance in the clusters and among the clusters firstly. Then this method was used to optimize the input feature sets and reduce the redundancy and excess of the training sample sets. Furthermore, confirmed the Radial Basis Function by comparing the character of the kernel functions. At the same time, modified the particle swarm optimization algorithm about the particle speed, location and the inertia weight value to increase the diversity of particle swarm and avoided the convergence of searching, and this method is used to optimize the SVM parameters and built the forecast model. Finally the example showed the forecast index was objective and the modified forecasting model was accurate.

Keywords: forecasting, FP-growth association rule, fuzzy Bayesian network, fuzzy C-mean clustering, modified particle swarm optimization, support vector machine

1 Introduction

With the comprehensive application of information technology in equipment construction domain, the complexity and technical intensity of modern equipment and weapon are increasing, which cause the equipment support work to be more difficult. This phenomenon exhibits as the complication of weapon, diversification of combat form and frequent maintaining. Due to the low good-rate, the difficulties in maintaining and the inconvenience of combined operations support harmony, the formation of combat capability and support capability is very slow. At the same time, as lots of informatization weapon and equipment are produced and deployed to use, these equipments need higher support capability to exert their performance [1].

Currently, there are many researches about equipment support capability and its design attributes assessment, their aims are making decision by the assessment results. However, the decision-maker not only need to mastery the current situation, but also forecast the future and make decision, which could

* Corresponding Author
find the coming problems and make a long programming for increasing the equipment support capability. Therefore, forecasting the equipment support capability and master the future development level are meaningful to improve equipment support capability. While predicting equipment support capability, we must confirm the prediction data’s characteristics of and the forecast demand, and then select the appropriate forecasting model. According to the definition and related attributes character of equipment support capability, reference [1] concludes the characters of equipment support capability to be nonlinear continuous, mutation, complexity, dynamic openness and incompleteness, and point out the requirement on the basis of the characteristics to be feasible, accurate, objective and complete.

Forecasting theories mainly includes qualitative analysis [2], trend analysis and forecasting [3], exponential smoothing [4], econometric method [5], time series model [6], support vector machine [7], Markov model [8] and so on. Since the support vector machine has simple structure, good performance, high learning speed and good generalization ability, and is suitable for dealing with high dimensional data and avoids the curse of dimensionality, it is used as the main forecast method in this paper [9].

During the process of using SVM model, the objectivity of forecast index, the characteristics of input feature sets and the optimization of parameters have important influence on the forecast accuracy. The forecast index is usually confirmed by experience or mechanism analysis. In the aspect of input feature sets optimization, the common methods include formatting methods and artificial intelligence (such as clustering neural network, wavelet analysis, and fuzzy C-means clustering). The usual method of parameter optimization is empirical optimization method, but there also includes genetic algorithm, particle swarm optimization and so on. However, as the SVM algorithm has difficulty in training samples on large scale, and the current optimization algorithms of input feature sets and parameters have some limitations, for example, the FCM-SVM clustering algorithm [10] can’t deal with outlier data, and it neglects the contribution of the sample to the classification results, thus reduce the reliability of clustering; in the PSO-SVM model [11], the particles speed is only related to the current position of the particles, but not the location information.

Based on the above analysis, this paper is organized as follow. Sect. 2 constructs the forecast index of equipment support capability based on ontology and text mining, Sect. 3 improves the fuzzy C-means clustering algorithm in outlier operation and the distance in and between the clusters to optimize the input feature sets, Sect. 4 modifies the PSO model about the particle speed and the inertia weight value, uses the MPSO to optimizing the parameters and builds the SVM forecasting model. Sect. 5 proves the forecast index is objective and forecast method is feasible and effective by analyzing the example.

2 Equipment Support Capability Forecasting Index System

Since the level of equipment support capability is a qualitative conception, we should assess the equipment support capability level to be a calculable number firstly. Usually, the equipment support capability forecast index is built by the equipment support sort, construction factor and support mode based on subjective experience and mechanism analysis. However, different thinking angle and knowledge range may result in different index and it may affect the reliability and accuracy of the forecast model.

From the view of ontology and text mining, this paper points out a method to confirm the index and weight. This method uses the thought of ontology to mine the text in a high level and produces high level or multi level rules, which have semantic meaning and are composed by high level abstract concept. For example, “support personnel” is used to generalize “human resource” and “personal factor”, and “facility and equipment support” is used to generalize “material support” and “maintenance equipment”. The final forecast index and weight are confirmed by text mining, which show the objectivity of the index. The steps to confirm the index by ontology and data mining are as follow:

(1) Preprocessing method of text data

The paper takes the reverse maximum method to preprocess the text data. This method chooses a character string which contains 6 to 8 Chinese characters as the maximum character string and matches the maximum character string with the word entries in the dictionary. If they can’t match, then cut off a Chinese character and continue matching until finding corresponding word location in the dictionary, the matching direction is from left to right.

(2) Feature representation

Text feature representation is the textual metadata and it is divided into descriptive features (such as the text name, date, size, type) and its semantic features (such as the author, institutions, title, content of
text, etc.). The feature representation represents documents by some feature terms, and it only needs to process these feature items when mining the text so as to process unstructured text.

(3) Feature extraction

The feature extraction algorithm evaluates the feature by constructing an evaluation function, and then it arranges according to the score, the highest predetermined score is selected. In text processing, the most commonly used evaluation functions are information gain, expected cross entropy, mutual information, weight of text evidence and word frequency. This paper takes the weight of text evidence to extract feature, the evaluation function is used to measure the difference between the probability of class and the conditional probability of given characteristic class, and its effect in the experiment is superior to the expected cross entropy.

(4) Curtailment of feature set

Feature sets are curtailed by the latent semantic indexing method and the singular value decomposition in matrix theory. The curtailment of feature set transforms the word frequency matrix to a $K \times K$ singular matrix. The basic steps are as follows:

① Establish the word frequent matrix.

② Singular value analysis of word frequent matrix, decompose the frequency matrix to 3 matrixes $U \cdot S \cdot V$. The matrix $U$ and $V$ are orthogonal matrices, and $S$ is the singular values of diagonal matrix ($K \times K$).

③ To every document $d$, use the new vectors gained by eliminating the word in SVD to replace the original vector.

④ Save all the vector sets, use advanced multidimensional indexing technology to create an index.

⑤ Calculate the similarity by the converted document vectors.

Step 5 Text special processing

During the process of text mining, combine or replace some strongly correlative searching words, for example, we use support personnel to replace all the personnel factor, human resource and support officer, and we combine the establishment support, facility support, material support and spare part support to be facility equipment support. When processing the frequency of each words, assume the number of forecast index is 5 (this number may be changed as other number, and it is just a hypothesis), select 5 indexes whose frequencies are the biggest five indexes, and regard them to be able to obtain the equipment support capability evaluation value. The threshold of the words frequency is the frequency of the smallest index in the above five index. To the weight value, this paper defines it as the ration of the summary of this index’s frequency to the summary of the total frequency.

$$ p_i = \frac{\sum_{j=1}^{n} q_{ij}}{\sum_{j} q_j} \tag{1} $$

Where $q_{ij}$ is the weight value of each keyword in the $j$-th text, $n_j$ is the frequency of this keyword in all the texts, $q_i$ is the summary weight value of the total frequency.

Based on the normalized parameter values and their weight values, equipment support capability evaluation algorithm is:

$$ P = \sum_{i=1}^{5} p_i D_i \tag{2} $$

Where $p_i$ is the weight of the $i$-th index, $D_i$ is the normalized parameter value of the $i$-th index.

Based on the above method, this paper collects and analyzes 26 publicly published Chinese and English papers about equipment support capability index, and finally support personnel, support technology, equipment support command, equipment support management and facility equipment support are extracted to be the forecast index. Each index and its weight value are shown in Table 1.
Table 1. Equipment support capability index and frequency based on text mining

<table>
<thead>
<tr>
<th>Index</th>
<th>Support personnel</th>
<th>Support technology</th>
<th>Equipment support command</th>
<th>Equipment support management</th>
<th>Facility equipment support</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relevant parameters</td>
<td>Employable testing score</td>
<td>Support technology testing score</td>
<td>Support command testing score</td>
<td>Support management testing score</td>
<td>Availability and matching rate</td>
</tr>
<tr>
<td>Frequency</td>
<td>13</td>
<td>12</td>
<td>11</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>Weight ratio (%)</td>
<td>23.6</td>
<td>21.8</td>
<td>20</td>
<td>18.2</td>
<td>16.4</td>
</tr>
</tbody>
</table>

In Table 1, the way to calculate the value of facility equipment assessment is:

\[ S_p = \frac{S_n + S_t}{2} \]  \hspace{1cm} (3)

\( S_p \) represents the value of facility equipment assessment, \( S_n \) represents the availability facility equipment, \( S_t \) represents the matching rate of facility equipment.

3 Input Feature Set Optimization Based on IFCM

The input feature set has influence on the forecast accuracy of SVM. After optimizing the input feature set by artificial intelligence, SVM can be used to establish the forecasting model according to different clusters, which makes each forecasting model only retains the sample points near the optimal level, gets rid of the distant points and reduces the redundancy and excess of the training sample set, finally affects the forecast accuracy.

FCM clustering ignores the difference about the contribution of the sample characteristics to the classification results, so it can’t distinguish the outlier samples. Its improved algorithm weighted FCM (WFCM) only considers minimizing the distance in the samples but ignores maximizing the distance between the samples. This paper proposes an improved fuzzy C-means clustering algorithm (IFCM). This clustering algorithm can not only effectively distinguish the outlier samples in data, but also minimize the distance in the samples and maximize the distance between samples. Therefore, it can reduce training samples and improve the training speed without affecting the SVM performance, and it will improve the forecast accuracy of the model ultimately.

3.1 Outliers Acquisition and Correction

The outliers in the data sample are usually defined as some class value whose appearing time is small and rare in a categorical variable, or some value which is too large in an interval variable. In general, processing the outliers will affect the accuracy and stability of the forecasting model. [7] The processing methods of the outliers generally include direct deletion, residual analysis, substituting by mean or statistic, considering them as missing value and filling them with statistical model, using robust model modification, using sampling technique or simulation technology to accept more reasonable standard error, etc. In this paper, the method of random linear method is used to correct the outliers, and the outliers will be replaced by the modified value. The steps of outliers correction are shown as follows:

**Outliers acquisition.** The algorithm of outlier detection is as follow:

\[ P = \sqrt[\frac{1}{2}]{{\sum_{j=2}^{j} T^2}} \]  \hspace{1cm} (4)

Where, \( P \) is the sample number of the outliers set, \( j \) is the sample sequence number, and \( T \) is the difference characteristic formula of the abnormal factor.

The algorithm for the threshold of outliers is:

\[ r(w) = \sum_{j=2}^{p} Y_j / P \]  \hspace{1cm} (5)
Compare the observed sample data, if it measures up the condition of outliers data, extract them, otherwise, continue to calculate and judge. The judging method is comparing the size of $r(w)$ and $\lambda$, if the left number is large, define it as outlier, otherwise it is not a outlier. Where the size of $\lambda$ is defined by the operator, and it could be determined by the revised accuracy.

**Outlier correction.** The methods of outlier correction include SVM, machine learning, small world model and information entropy, etc. Because of the complexity of IFCM clustering model, for the sake of easy calculation, this paper uses random linear method to correct the outliers, produces different correction values and clustering them.

To the $j$-th outlier $x_j$, set $x_j = k_1x_{j-1} + k_2x_{j+1}$, where $x_{j-1}$ and $x_{j+1}$ is the previous one normal values and posterior one normal value of $x_j$ and $k_1, k_2 \in [0,1]$, $k_1 + k_2 = 1$. Because $k_1$ and $k_2$ are random values, the modified number is generated by cyclic iteration, which can guarantee the maximum membership degree.

### 3.2 Objective Function and Constraint Condition

The objective function of IFCM clustering algorithm is:

$$J(U, P, c) = \frac{\sum_{i=1}^{n} \sum_{c=1}^{c} (u_{ik})^m w_i (d_{ik})^2}{\sum_{j=1}^{c} \sum_{j=1}^{n} (p_i - p_j)}$$  \hspace{1cm} (6)

Subject to:

$$0 \leq u_{ik} \leq 1, 1 \leq i \leq c, 1 \leq k \leq n$$

$$\sum_{i=1}^{n} u_{ik} = 1, 1 \leq k \leq n$$

$$0 < \sum_{k=1}^{n} u_{ik} < n, 1 \leq i \leq c$$

$$1 \leq m \leq \infty, p_i \neq p_j$$  \hspace{1cm} (7)

Among them, the distance between the classes is $p_i - p_j = d_{ik} - d_{jk}$, the weight value is $w_i$, the weighted index is $m$, the distortion is $d_i$, the prototype vector is $p_i (i=1,2,\cdots,c)$.

The clustering criterion satisfies the following constraints:

$$0 \leq u_{ik} \leq 1, 1 \leq i \leq c, 1 \leq k \leq n$$

$$\sum_{i=1}^{n} u_{ik} = 1, 1 \leq k \leq n$$

$$0 < \sum_{k=1}^{n} u_{ik} < n, 1 \leq i \leq c$$

$$1 \leq m \leq \infty, d_{ik} \neq d_{jk}$$  \hspace{1cm} (8)

Under this condition, the error square $J(U, P, c)$ between the sample and the typical sample reaches the minimum value. When the objective function $J(U, P, c)$ reaches the maximum, the distance among the classes is the largest and the distance in the classes reaches the minimum.

According to Lagrange multiplier method, the optimal iteration formula of $u_{ik}$ and $p_j$ can be deduced as follow:
\[
U_{ik} = \frac{\left( \frac{1}{n-1} \sum_{j=1}^{c} \left( \frac{(d_{ik})^2}{\sum_{j=1}^{c}(d_{ik} - d_{jk})^2} \right) \right)}{\sum_{j=1}^{c} \left( \frac{(d_{ik})^2}{\sum_{j=1}^{c}(d_{ik} - d_{jk})^2} \right)}
\]

\[p_i = \frac{\sum_{i=1}^{n} w_{im} x_i}{\sum_{i=1}^{n} w_{im}}\]

3.3 Stopping Condition and Output

(1) Set dividing matrix as \(U^0 = [u_{ik}]_{c \times n}\), \(u_{ik}\) represents the fuzzy membership degree of the sample \(x_k\) belongs to the \(i\)-th cluster, and guarantee \(\sum_{i=1}^{c} u_{ik} = 1\); Set stopping threshold \(\varepsilon\) and iteration counter \(b = 0\).

(2) Calculate the fuzzy centroid of every fuzzy subset as follow:

\[p_i^{(b)} = \frac{\sum_{i=1}^{n} u_{im} x_i}{\sum_{i=1}^{n} u_{im}}\]

(3) Calculate the weighted matrix \(w_i^{(b)}\) of every fuzzy cluster as follow:

\[(w_i^{(b)})_{kl} = \begin{cases} 
\left( \frac{1}{\sum_{i=1}^{c} u_{im}} \right) & k = l \\
Tr((\sum_{i=1}^{c} u_{im}))^{(b)} & k \neq l \\
0 & \text{otherwise}
\end{cases}\]

Where \(Tr(\bullet)\) is the operation of calculating the matrix trace, \(\sum_{i=1}^{c} u_{im}\) is the \(i\)-th fuzzy covariance matrix, the calculating process is as follow:

\[\left( \sum_{i=1}^{c} u_{im} \right)^{(b)} = \frac{\sum_{i=1}^{n} (\mu_{ik}^{(b)})^w (d_{ik})^2}{\sum_{i=1}^{n} (\mu_{ik}^{(b)})^w}, i = 1, 2, \ldots, c\]

(4) Calculate the distance \(d_{ik}\) of sample \(x_k\) to the \(i\)-th class centroid as follow:

\[d_{ik}^{(b)} = \|x_k - p_i^{(b)}\|^2 = (x_k - p_i^{(b)})^\top w_i (x_k - p_i^{(b)})\]

(5) Use the following formula to update the fuzzy partition matrix \(U^{(b+1)} = [\mu_k]_{c \times n}\).

To \(\forall i, k\), if \(\forall d_{ik}(t) > 0\), then
\[
\mu_{ik}^{(b+1)} = \frac{1}{m+1} \sum_{j=1}^{m} \frac{(d_{ik}^{(b+1)})^2}{(d_{ik}^{(b+1)} - d_{jk}^{(b+1)})^2},
\]

If \( \exists i, r \), and \( d_{ir}^{(b)} = 0 \), so
\[
\mu_{ir}^{(b)} = 1, \text{ and to } j \neq r \ , \ \mu_{ij}^{(b)} = 0
\]

(6) When \( \|U^{(b+1)} - U^{(b)}\| < \epsilon \), then the calculating process is stopped and outputting partition matrix \( U \) and cluster antitype \( P \), otherwise set \( b = b + 1 \), then go to Step (1). In the formula \( \|\cdot\| \) is a suitable matrix norm.

4 MPSO-SVM Forecast Modeling

Kernel function and its parameters have important influence on the forecast accuracy of SVM. Based on the advantages of RBF kernel function, this paper uses it to be the kernel function of the forecasting model. In the aspect of parameter selection, this paper uses the modified particle swarm optimization algorithm to increase the diversity of particle swarm and avoid the convergence of searching, which can effectively jump out of the local minimum and get more objective parameters, and ultimately affects the forecast accuracy.

4.1 Kernel Function Selection

Kernel function plays an important role in the SVM forecast model, and the effect of different kernel functions on SVM is also different. Common SVM kernel functions include radial basis function (RBF), linear kernel function, Sigmoid kernel function and polynomial kernel function. \([8]\) When the number of features is smaller than the number of samples, the linear kernel function is not adaptive, so this paper does not choose linear kernel function. The characteristics of the polynomial kernel function, RBF kernel function and Sigmoid kernel function are shown in Table 2. This paper chooses it to be the kernel function of the forecast model based on the advantages of the RBF kernel function.

<table>
<thead>
<tr>
<th>Kernel function</th>
<th>Nonlinear mapping capability</th>
<th>Parameter number</th>
<th>Numerical limiting conditions</th>
<th>Global superiority</th>
<th>Positive definite</th>
</tr>
</thead>
<tbody>
<tr>
<td>Radial basis function</td>
<td>Yes</td>
<td>Less</td>
<td>Have constraint</td>
<td>Locality is strong</td>
<td>Positive definite</td>
</tr>
<tr>
<td>Sigmoid kernel function</td>
<td>When define specific parameters</td>
<td>More</td>
<td>Maybe ineffective</td>
<td>Locality is strong</td>
<td>Non positive definite</td>
</tr>
<tr>
<td>Polynomial kernel function</td>
<td>No</td>
<td>Uncertain</td>
<td>Uncertain</td>
<td>Global superiority is strong</td>
<td>Positive definite</td>
</tr>
</tbody>
</table>

RBF kernel function algorithm is as follows, where \( \delta \) means kernel parameter.

\[
k(x, y) = \exp \left( -\frac{\|x - y\|^2}{\delta^2} \right), \delta \in R
\]
4.2 MPSO Parameters Optimization Modeling

Particle swarm optimization (PSO) is a kind of parallel swarm intelligence optimization algorithm, which is characterized by less experience parameters, fast calculation speed and strong searching ability. The traditional PSO algorithm only considers the speed of the particles being related to the current position and ignores the location information, and the PSO algorithm is easy to fall into local optimum, which leads to the low convergence accuracy. This paper proposes a Modified PSO (MPSO) algorithm, which considers both the diversity and the convergence problem, and prevents the generation of local minima.

MPSO hypotheses are as follows:

1. The velocity of the particle is not only related to the current position of the particle, but also related to front position information and back position information of the particle.
2. The velocity of the particle is not entirely dependent on the speed of the previous moment.
3. The position function of the particle is independent to the front position information and back position information of the particle.
4. The inertia weight is related to the relative position of the particle.

The general particle swarm optimization algorithm step could be seen in reference [10], and this paper mainly modifies the PSO-SVM algorithm as follow:

**Modification of the particle’s velocity and position.** The particle velocity is in accordance with the following formula:

\[
v_{i,j}(t+1) = \omega v_{i,j}(t) + c_1 r_1 p(t) + c_2 r_2 g(t)
\]

Where

\[
p(t) = p_{i,j} - s_{11} x_{i,j} (t-1) - s_{12} x_{i,j} (t) - s_{13} x_{i,j} (t + 1)
\]

\[
g(t) = g_{i,j} - s_{21} x_{i,j} (t-1) - s_{22} x_{i,j} (t) - s_{23} x_{i,j} (t + 1)
\]

\[
s_{11}, s_{12}, s_{13} \in [0,1], s_{11} + s_{12} + s_{13} = 1, c_1 \text{ and } c_2 \text{ are learning factors, } c_1 \text{ influences the cognitive ability and } c_2 \text{ influences the information-sharing ability of particle, } c_1 + c_2 \in (2-\varepsilon, 2+\varepsilon), \varepsilon \text{ is any small positive value.}
\]

\[
x_{i,j}(t) = k_1 x_{i,j} (t+1) + k_2 x_{i,j} (t+1) - v_{i,j} (t+1)
\]

And it satisfies \(k_1, k_2 \in (0,1), k_1 + k_2 = 1\). During the programming process, in order to prevent infinite loop, the definition of the accuracy of the random value is 0.001.

**Inertia weight value modification.** The inertia weight is proposed to avoid the premature of the optimization process and the local search being trapped too early, it widens the searching space of the PSO algorithm and could balance the searching ability in the local and global. The value of inertia weight is usually a constant value, and it is also defined as a linear function [11] or linear function based on relative distance [12]. Due to the complexity of the optimization process and the difference between particles, the above methods have some limitations to the problem. This paper considers that the inertia weight of the particle is changed according to the inertia weight of the particle, the relationship between them is nonlinear, and the purpose is to make the particles with better adaptation value converge accurately, and the particles with worse adaptation value converge fast. Therefore, inertia weight should be in the form of exponential decline.

The reference [13] and [14] have discussed in the form of exponential decline, but the trends are different. Concluding the two methods, this paper considers that the inertia weight satisfies the following conditions.

\[
w = w_{\min} + (w_{\max} - w_{\min})e^{-f(t^2)}
\]

Where \(f(t^2) < 0\). In order to facilitate the computation of the computer cycle, set \(f(t^2) = -k_1 t^{2k_2}\), \(k_1 \in (0,1), k_2 \text{ is positive integer and } k_2 \leq 10\), the calculation accuracy of \(k_1\) is 0.01, and the optimal value of \(k_1\) and \(k_2\) is calculated by the iterative operation.

240
4.3 Forecast modeling process

The forecast model is constructed in the following process, and the modeling process is shown in Fig. 1.

**Fig. 1.** Forecast modeling process

**Step1:** Collect relevant information, use ontology and text mining method to confirm the forecast index and weight value. To forecast equipment support capability, the first step is to build an assessment index system and obtain the evaluation value of equipment support capability. Based on ontology and unstructured data, this paper uses text mining to confirm the indexes and weights of equipment support capability. By this method, the evaluation value of equipment support capability will be more objective and the input sample data of forecasting model could be more credible.

**Step2:** Use IFCM model to correct the outliers of the input feature sets. Since the input feature set influences the forecast accuracy of SVM, by analyzing the defeat in FCM and WFCM, this paper build an IFCM model to distinguish the outlier samples, minimize the distance in the samples and maximize the distance between samples. By IFCM algorithm, the training samples are reduced and the training speed is accelerated, and it will improve the prediction accuracy ultimately.

**Step3:** Select the appropriate kernel function, and use MPSO to optimize the parameters. By comparing the characteristics of nonlinear mapping capability, parameter number, numerical limiting conditions, global superiority and positive definite about kernel functions, this paper selects the RBF kernel function. Then modify the particle swarm optimization algorithm about the particle speed, location and the inertia weight value, increase the diversity of particle swarm and avoid the convergence of searching and finally find the best parameters of SVM.

**Step4:** Build forecast model and analyze the forecast results. As the input data is operated by IFCM and the best parameters of SVM are obtained, use SVM regression algorithm to build forecasting model and predict the evaluation value of equipment support capability. Use RMSE and MAPE to validate the accuracy of this model by comparing with other forecasting methods.

5 Case analysis

5.1 Case Introduction

To validate the accuracy and applicability of the integrated forecasting model, this paper takes three datasets from equipment support comprehensive management information system.

Every dataset has different characteristics. Dataset A is obtained from a department which has no emergent missions, personnel alteration or support equipment variation. Therefore, its equipment support capability has increased slowly. Dataset B is obtained from a unit which has many emergent missions to execute. At the same time, it has changed the main support equipment during the period of data collection, the leaders of this unit have changed for 4 times and its station has changed twice, so its equipment support capability assessment value may change largely. Dataset C has the situation of both Dataset A and Dataset B, in the first 15 equipment support capability assessment value, the change of the value is small, however, the last 15 equipment support capability assessment value change largely. The equipment support capability value is calculated by the normalized index value and the weight value.
5.2 Case Validation

Based on the forecast modeling process, this paper uses IFCM and MPSO-SVM hybrid forecast model to forecast the equipment support capability and validate the forecast accuracy.

In order to verify the validity of the IFCM clustering algorithm, this paper uses FCM and WFCM algorithms for clustering at the same time. Set the number of clusters is 3 (this value can be defined as any other positive integer), use clustering membership degree to evaluate the accuracy of the clustering, the comparison of membership degree in each clustering algorithm is shown in Table 3. Clustering membership algorithm is as follows:

For any kind of center i and element r, if there exists \( d_{ir}^{(k)} > 0 \), the membership algorithm is:

\[
u_{ij}^{(k)} = \left[ \sum_{i=1}^{c} \left( \frac{d_{ij}^{(k)}}{d_{ir}^{(k)}} \right)^{2} \right]^{-1}
\]

(23)

If \( i \) and \( r \) make \( d_{ir}^{(k)} = 0 \), then \( \nu_{ij}^{(k)} = 1(j \neq r, \nu_{ij}^{(k)} = 0) \).

### Table 3. Sum of membership degree in each clustering algorithm

<table>
<thead>
<tr>
<th></th>
<th>Dataset A</th>
<th></th>
<th></th>
<th>Dataset B</th>
<th></th>
<th></th>
<th>Dataset C</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number</td>
<td>Membership(10^-4)</td>
<td>FCM</td>
<td>WFCM</td>
<td>IFCM</td>
<td>Number</td>
<td>Membership(10^-4)</td>
<td>FCM</td>
</tr>
<tr>
<td>1</td>
<td>5443</td>
<td>5733</td>
<td>5798</td>
<td>1</td>
<td>8652</td>
<td>8564</td>
<td>8762</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>5709</td>
<td>5707</td>
<td>5857</td>
<td>2</td>
<td>5042</td>
<td>5141</td>
<td>6248</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>5474</td>
<td>5884</td>
<td>6665</td>
<td>3</td>
<td>7371</td>
<td>6971</td>
<td>7692</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>6516</td>
<td>7310</td>
<td>7349</td>
<td>4</td>
<td>8406</td>
<td>8313</td>
<td>9175</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>5574</td>
<td>5765</td>
<td>6848</td>
<td>5</td>
<td>6646</td>
<td>5964</td>
<td>6681</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>6941</td>
<td>7188</td>
<td>7724</td>
<td>6</td>
<td>7361</td>
<td>8058</td>
<td>8682</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>7850</td>
<td>7681</td>
<td>8456</td>
<td>7</td>
<td>5243</td>
<td>5181</td>
<td>6374</td>
<td>7</td>
</tr>
<tr>
<td>8</td>
<td>8407</td>
<td>8233</td>
<td>8952</td>
<td>8</td>
<td>7073</td>
<td>7205</td>
<td>7005</td>
<td>8</td>
</tr>
<tr>
<td>9</td>
<td>7477</td>
<td>7717</td>
<td>7786</td>
<td>9</td>
<td>7319</td>
<td>7771</td>
<td>7806</td>
<td>9</td>
</tr>
<tr>
<td>10</td>
<td>5399</td>
<td>5798</td>
<td>5526</td>
<td>10</td>
<td>5042</td>
<td>5749</td>
<td>5811</td>
<td>10</td>
</tr>
<tr>
<td>11</td>
<td>8170</td>
<td>8501</td>
<td>9434</td>
<td>11</td>
<td>5318</td>
<td>6149</td>
<td>6061</td>
<td>11</td>
</tr>
<tr>
<td>12</td>
<td>4591</td>
<td>5496</td>
<td>5551</td>
<td>12</td>
<td>8492</td>
<td>8450</td>
<td>8716</td>
<td>12</td>
</tr>
<tr>
<td>13</td>
<td>7605</td>
<td>7159</td>
<td>7662</td>
<td>13</td>
<td>7746</td>
<td>8361</td>
<td>8604</td>
<td>13</td>
</tr>
<tr>
<td>14</td>
<td>7765</td>
<td>8230</td>
<td>8166</td>
<td>14</td>
<td>7319</td>
<td>7506</td>
<td>7270</td>
<td>14</td>
</tr>
<tr>
<td>15</td>
<td>8755</td>
<td>9078</td>
<td>9022</td>
<td>15</td>
<td>5337</td>
<td>5813</td>
<td>5823</td>
<td>15</td>
</tr>
<tr>
<td>16</td>
<td>6317</td>
<td>6097</td>
<td>6810</td>
<td>16</td>
<td>6818</td>
<td>7390</td>
<td>7779</td>
<td>16</td>
</tr>
<tr>
<td>17</td>
<td>8406</td>
<td>8821</td>
<td>9261</td>
<td>17</td>
<td>7553</td>
<td>7700</td>
<td>8286</td>
<td>17</td>
</tr>
<tr>
<td>18</td>
<td>7599</td>
<td>7747</td>
<td>8281</td>
<td>18</td>
<td>8832</td>
<td>8732</td>
<td>9378</td>
<td>18</td>
</tr>
<tr>
<td>19</td>
<td>8375</td>
<td>7875</td>
<td>8687</td>
<td>19</td>
<td>6846</td>
<td>6930</td>
<td>7416</td>
<td>19</td>
</tr>
<tr>
<td>20</td>
<td>5600</td>
<td>5779</td>
<td>6203</td>
<td>20</td>
<td>5564</td>
<td>5499</td>
<td>6155</td>
<td>20</td>
</tr>
<tr>
<td>21</td>
<td>6971</td>
<td>7115</td>
<td>7788</td>
<td>21</td>
<td>6399</td>
<td>6442</td>
<td>6588</td>
<td>21</td>
</tr>
<tr>
<td>22</td>
<td>7524</td>
<td>7554</td>
<td>8738</td>
<td>22</td>
<td>5810</td>
<td>6066</td>
<td>6575</td>
<td>22</td>
</tr>
<tr>
<td>23</td>
<td>6822</td>
<td>7065</td>
<td>6931</td>
<td>23</td>
<td>6255</td>
<td>6815</td>
<td>7776</td>
<td>23</td>
</tr>
<tr>
<td>24</td>
<td>5222</td>
<td>5082</td>
<td>6324</td>
<td>24</td>
<td>8517</td>
<td>8980</td>
<td>9298</td>
<td>24</td>
</tr>
<tr>
<td>25</td>
<td>7407</td>
<td>7076</td>
<td>8137</td>
<td>25</td>
<td>4990</td>
<td>5241</td>
<td>5688</td>
<td>25</td>
</tr>
<tr>
<td>26</td>
<td>8659</td>
<td>8817</td>
<td>9605</td>
<td>26</td>
<td>5385</td>
<td>5566</td>
<td>5330</td>
<td>26</td>
</tr>
<tr>
<td>27</td>
<td>7770</td>
<td>8529</td>
<td>8922</td>
<td>27</td>
<td>8870</td>
<td>8281</td>
<td>9680</td>
<td>27</td>
</tr>
<tr>
<td>28</td>
<td>5544</td>
<td>5512</td>
<td>5678</td>
<td>28</td>
<td>6942</td>
<td>7441</td>
<td>7214</td>
<td>28</td>
</tr>
<tr>
<td>29</td>
<td>8755</td>
<td>8842</td>
<td>9524</td>
<td>29</td>
<td>6547</td>
<td>7499</td>
<td>7833</td>
<td>29</td>
</tr>
<tr>
<td>30</td>
<td>6901</td>
<td>7806</td>
<td>7822</td>
<td>30</td>
<td>8358</td>
<td>8357</td>
<td>9640</td>
<td>30</td>
</tr>
</tbody>
</table>

In the process of optimizing the parameters, As MPSO method focuses on the interrelationship between the particles, based on the meaning and area of \( c_1 \) and \( c_2 \), setting \( c_2 \) to be a little bigger than \( c_1 \) may help get the best parameters faster. Therefore, set \( c_1=0.995 \) and \( c_2=1.005 \). The iterations number is confirmed by the sample number, forecasting accuracy requirement and forecasting time. Since this paper
has used random number to select the inertia weight value, the iterations number needs to be bigger. At the same time, because the MPSO just focuses on the accuracy and not consider the forecasting time, so increase the iterations number may add the prediction accuracy, thus set the maximum number of iterations in the algorithm to be $10^9$. Other parameters are as follow: set $\varepsilon = 10^{-4}$, the range of $\delta^2$ is $(10^{-4}, 10^4)$, the random parameters are iterated by the designed calculating accuracy, and the remaining parameters are taken as the default value.

For the sake of validating the validity of the forecasting model, this paper uses SVM, FCM-SVM and PSO-SVM model to forecast the result. In order to test the accuracy of the forecast results, RMSE and MAPE are chosen as the indexes to analyze the forecasting accuracy.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (d_i - \hat{d}_i)^2}$$

$$MAPE = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{d_i - \hat{d}_i}{d_i} \right|$$

Where $N$ is the number of data sets, $d_i$ is the real value, $\hat{d}_i$ is the forecast value.

Table 4 shows the comparison of RMSE and MAPE in different models, and the comparisons of forecasting value and evaluation value in dataset A to dataset C are shown in Fig. 2 to Fig. 4.

![Fig. 2. Comparison of forecasting value and calculating value in Dataset A](image)

![Fig. 3. Comparison of forecasting value and calculating value in Dataset B](image)
5.3 Case Analysis

To equipment support capability forecasting index, since the confirming method is based on ontology and text mining, it overcomes the subjectivity and randomness. Form Table 3 we can find that the membership of IFCM algorithm is almost higher than WFCM and FCM model, which proves that the clustering accuracy of IFCM algorithm is higher than WFCM and FCM model.

From Table 4 and Fig. 2 to Fig. 4, we can find that to different characteristic of example datasets, when data in the example vary significantly, RMSE and MAPE are large; when data in the example vary slightly, RMSE and MAPE are little. However, no matter changes in the example are large or small, RMSE and MAPE in this paper is smaller than SVM, FCM-SVM and PSO-SVM model, and the forecasting values of IFCM-MPSO-SVM model are more closed to the evaluation values than SVM, FCM-SVM and PSO-SVM model. Therefore, we can conclude that the accuracy in this paper is higher than SVM, FCM-SVM and PSO-SVM model and the IFCM-MPSO-SVM model has favorable adaptability.

Table 4. RMSE and MAPE comparison of each forecast method

<table>
<thead>
<tr>
<th>Sample</th>
<th>Dataset A</th>
<th>Dataset B</th>
<th>Dataset C</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RMSE</td>
<td>MAPE</td>
<td>RMSE</td>
</tr>
<tr>
<td>This paper</td>
<td>0.0090</td>
<td>0.0042</td>
<td>0.0184</td>
</tr>
<tr>
<td>SVM</td>
<td>0.0445</td>
<td>0.0188</td>
<td>0.0486</td>
</tr>
<tr>
<td>FCM-SVM</td>
<td>0.0292</td>
<td>0.0121</td>
<td>0.0494</td>
</tr>
<tr>
<td>PSO-SVM</td>
<td>0.0189</td>
<td>0.0082</td>
<td>0.0395</td>
</tr>
</tbody>
</table>

6 Conclusion

This paper points out the issue of forecasting equipment capability, builds forecast index based on ontology and text mining, uses IFCM model and MPSO model to improve the SVM forecast model and builds hybrid forecast model. The case analysis shows that the index and weight value are objective, the clustering effect is better and the forecast model is more accurate, and all of the above shows the method in this paper is effective.

At the same time, this paper has the following aspects to be improved.

1. Since the index confirming method is rooted from the data sets, the selected sample data should be general and accurate.

2. As the random numbers has been used for many times in the paper, the time of program training has been increased. Especially, precision control may increase the difficulty to find the optimum points, and improve the accuracy may result in geometric multiples increase of the training time. Therefore, it is needed to find a more suitable way to reduce the training time about the optimization of the random number.
References


