Information Technology and Quantitative Management (ITQM 2016)

Forecasting China’s Service Outsourcing Development with an EMD-VAR-SVR Ensemble Method

Siqi Yi¹,²,c,*, Kun Guo¹,², Zhensong Chen¹,²,c

¹School of Economics and Management, Chinese Academy of Science, Beijing, 100190, China
²Research Center on Fictitious Economy & Data Science, Chinese Academy of Science, Beijing 100190, China
³Key Laboratory of Big Data Mining and Knowledge Management, Chinese Academy of Science, Beijing, 100190, China

Abstract

Service outsourcing is playing a more and more important role in the world economy. With the fast development of information technology and service industry, China is witnessing an accelerated growth momentum of service outsourcing. In the first half year of 2015, the total service-outsourcing executive contracts value of China is increased by 9.7% to US $40.8 billion. Looking ahead 10 years, China’s service outsourcing market will receive unprecedented opportunities and challenges, predicting the contract amount of service outsourcing can promote the healthy development of the industry. In this paper, an EMD-VAR-SVR ensemble method is proposed for predicting the executive amount of servicing outsourcing. In order to verify and test the forecasting capability of the proposed EMD-VAR-SVR ensemble method, the single VAR model and single SVR model are used as the benchmark models. The empirical result shows that the proposed ensemble method’s prediction has the best effect.

© 2016 Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Peer-review under responsibility of the Organizing Committee of ITQM 2016

Keywords: Service outsourcing; Empirical mode decomposition (EMD); Vector auto-regression (VAR); Support vector regression (SVR); Forecasting

1. Introduction

Service outsourcing is a business paradigm in which an organization has a part of its business process performed by a service provider [1-2]. It is playing a more and more important role in the world economy. The global information technology outsourcing (ITO) developed rapidly since 1989; the ITO market was only US $10 billion at that time [3]. According to the research of The Gartner Group and Data Monitor, in 2008 alone the top 20 worldwide information technology-outsourcing contracts were worth nearly US $20 billion [4-5]. In 2014, it...
is estimated that the deals exceeded US$700 billion [3]. With the fast development of information technology and service industry, China is witnessing an accelerated growth momentum of service outsourcing. The ministry of commerce found that the total service-outsourcing executive contracts value is increased by 9.7% to US $40.8 billion in the first half year of 2015.

Like most economy activity, service outsourcing is determined by its supply and demand. A number of papers have studied the factors and evolution of service outsourcing. For example, Amiti and Wei [6] find that service outsourcing is positively correlated with labor productivity in the United States. Lacity and Willcocks [7] study what factors can increase the service outsourcing success rate. Loh and Venkatraman [8], using data from fifty-five major U.S. corporations, find that outsourcing is dependent on business governance and financial leverage.

Fei analyzed the driving factors, influence powers and development trends of service outsourcing [9]. JIANG analyze service outsourcing ability of China’s Hainan province through four factors: macro economic factor, industrial foundation factor, human resources number factor and human resources cost factor. Since service outsourcing has been playing an increasingly important role in the world economy, predict the development of service outsourcing and understand the inner mechanism becomes an important and interesting problem [10]. As the service outsourcing contracts series are usually considered a nonlinear and nonstationary time series, which is affected by many factors, predicting service outsourcing contracts is challenging.

In the past decades, many time series models have been proposed to handle the area of forecasting [11-12]. Box and Jenkins proposed ARMA(autoregressive moving average) model to forecast linear stationary series[13]. Bollerslev put forward GARCH(generalized autoregressive conditional heteroskedasticity ) model considering the error variance[14]. However, these traditional time series and econometric models are built on linear assumptions and they cannot capture the nonlinear patterns. In recent years, some nonlinear and artificial intelligent models, like support vector machine (SVM), artificial neural networks (ANN), genetic programming (GP), have been applied to various disciplines, such as business, finance, engineering, management, etc. For example, Roh combined neural network and time-series model to predict the volatility of stock price index [15]. Xie et al. developed a predict system for crude oil price by using support vector regression (SVR) [16]. Artificial intelligent models often had some advantages over traditional models, but they also have shortcomings. ANN often suffers from local minima and over fitting; SVM and GP are sensitive to parameter selection. So some hybrid methods have been proposed. Wang proposed a TEI@I methodology for studying complex systems, using this methodology to predict crude oil price forecasting and obtain good prediction performance [17]. Cheng and Wei proposed a hybrid time-series support vector regression (SVR) model based on empirical mode decomposition (EMD) for forecasting TAIEX (Taiwan stock exchange capitalization weighted stock index) [18].

Based on the hybrid and TEI@I methodologies, this paper attempted to apply the “decomposition-and-ensemble” principle to construct an execution amount of service outsourcing contracts forecasting methodology. In the phase of decomposition, this paper uses EMD to decompose the original data (execution amount of service outsourcing contracts data) into simpler frequency components and highly correlation variables. In the phase of ensemble, the goal is to formulate a consensus forecasting on original data [19]. Based on this principle, this paper proposes an EMD-VAR-SVR ensemble model to predict service outsourcing. In this proposed methodology, the original data series were first decomposed into a finite number of intrinsic mode functions (IMFs) and a residue. An independent SVR or VAR model models the IMFs according to its own characteristics. Finally, prediction results of all IMF components are aggregated, using an SVR to produce an ensemble forecasting result for the original service outsourcing series. Besides, this paper also compared the proposed EMD-VAR-SVR model with single VAR and single SVR.

The rest of the paper is organized as follows: section 2 describes the related work and formulation process of the proposed EMD-VAR-SVR methodology in detail. Section 3 describes the experiments and comparisons, the execution amount of service outsourcing contracts series is used to test the effectiveness of the proposed methodology. Finally, the conclusions and future work are summarized in Section 4.
2. Methodology

2.1. TEI@I methodology

TEI@I methodology is proposed by Wang, which means “text mining + econometrics + intelligence @ integration”. TEI@I embodies the thought of “decomposition-and-ensemble”. For the complex systems, firstly, it uses econometric model to analyze the main trend of complex systems after decomposing. Then, analyzes nonlinear and uncertainty system by the use of artificial intelligence technology. After that, use text-mining technology to deal with the mutability and instability of complex system. Finally, based on the idea of integration, achieve the objective of analyzing complex systems. TEI@I methodology has been used in the area of prediction in recent years. For example, LAI applied TEI@I methodology for crude oil price prediction [20].

2.2. Empirical mode decomposition (EMD)

Since the nonlinear and non-stationary data often involve more than one oscillatory mode, and single Hilbert transform cannot provide the full description of the frequency content, the empirical mode decomposition (EMD) is proposed by Huang et al to solve this problem. EMD has been widely used in many fields, such as geophysics [21-22], Baltic Dry Index [23], medical [24], facial emotion recognition [25], and crude oil price forecasting [26].

The key principle of EMD is to decompose a time series data set into a finite number of intrinsic mode functions (IMFs). The following sifting process can express the decomposition procedure of EMD:

1. Use a cubic spline line to connect all the local maxima and minima as the upper envelope and lower envelope.
2. Compute the mean of the upper envelope and lower envelope as m1.
3. The first component a1 is the different value of data and m1, a1=x1−m1.
4. Check whether a1 is an IMF or not. An IMF should satisfy two requirements: the number of extrema and zero crossing must either equal or differ at most by one; the mean value of local maxima and minima envelope is zero at any point. If a1 is an IMF, the residual r1=x1−a1. If a1 is not an IMF, replace x1 with a1.
5. Repeat step (1)-(4) k times, until a1k is an IMF, that is
   \[ a_1 \left( k-1 \right) - m_{1k} = a_{1k} \]  
   (1)
6. The first IMF c1=a1k, the residual r1=x1−c1 is treated as the new data and subjected to the same sifting process described above.

Repeat the process described above until the last residual r_n becomes a monotonic function from which no more IMFs can be extracted, or r_n is less than the predetermined value of substantial consequence. After the sifting process, we decomposition the data into n-IMFs which are nearly orthogonal to each other, and a residue r_n which represents the main trend of original data x(t). The data series x(t) can be described as:

\[ x(t) = \sum_{i=1}^{n} c_i + r_n \]  
(2)

2.3. Support vector regression (SVR)

Support vector machine (SVM) is a method of machine learning based on statistical learning theory. As for structural risk minimization principle, SVM has a better solution to small sample [27]. Support vector regression is one kind of SVM, which is aimed at finding a function f(x) that matches all input data with an error at the most
\( \varepsilon \) [28]. \( \varepsilon \) can reflect the distance between \( f(x) \) and the true value \( y \), the loss function can be described as the following formula:

\[
L_\varepsilon(f(x), y) = \begin{cases} 
|f(x) - y| - \varepsilon & \text{if } |f(x) - y| \geq \varepsilon \\
0 & \text{otherwise}
\end{cases}
\] (3)

The procedure can be expressed as following [29]:

1. Let the training samples to be \( T = \{(x_1, y_1), \cdots, (x_i, y_i)\} \), where \( x_i \in \mathbb{R}^n, y_i \in \mathbb{R}, i = 1 \cdots n \).
2. Choose an appropriate accuracy \( \varepsilon \) and the penalty parameter \( C > 0 \). The original problem of SVR is:

\[
\min \frac{1}{2} \| w \|^2 + C \sum_{i=1}^{l}(\xi_i + \xi_i^*)
\]

s.t. \( y_i - (\omega \cdot x_i) - b \leq \varepsilon + \xi_i, i = 1, 2 \cdots, l \) (4)

\( (\omega \cdot x_i) + b - y_i \leq \varepsilon + \xi_i^*, i = 1, 2 \cdots, l \) (5)

\( \xi_i, \xi_i^* \geq 0, i = 1, 2 \cdots, l \) (6)

(3) Construct and solve the dual problem:

\[
\min \frac{1}{2} \sum_{i=1}^{l}(a_i - a_i^*) (a_j - a_j^*) K(x_i, x_j) + \varepsilon \sum_{i=1}^{l} (a_i + a_i^*) - \sum_{i=1}^{l} y_i (a_i^* - a_i)
\]

s.t. \( \sum_{i=1}^{l} (a_i - a_i^*) = 0 \) (8)

\( 0 \leq a_i, a_i^* \leq C, i = 1, 2 \cdots, l \) (9)

(4) Compute \( \vec{b} \) and construct the decision function:

\[
f(x) = \sum_{i=1}^{l} (a_i^* - \bar{a}_i) K(x_i, x) + \vec{b}
\] (11)

2.4. Proposed model

Suppose the original data time series is \( x(t), t=1, 2 \cdots N \), we would like to make the l-step ahead prediction, i.e. \( x(t+l) \). Based on the TEI@I methodology, this paper proposed EMD-VAR-SVR model, see figure 1. The overall process of the model is as following:

1. The original data time series \( x(t) \) is decomposed into \( n \) IMF components and one residual component by the method of EMD. The IMFs represent main trend item and nonlinear term of service outsourcing, which means IMFs can reflect the inherent characteristics of the service outsourcing.

2. For the predict of each extracted IMF component and the residual component, we use econometric model to capture the main trend part of the service outsourcing fluctuations, using nonlinear method to fit nonlinear part of volatility. Here VAR and SVR method is adopted. Besides, in order to reflect the economic sense of IMFs and improve the accuracy of the prediction, we add the major influencing factors of service outsourcing into the model.

3. The prediction results of all extracted IMF components and the residue produced in the previous step are combined to generate an aggregated output using an SVR model. Therefore, we can get the final prediction result for the original data time series.

In order to verify the effectiveness of the proposed model, we use execution amount of service outsourcing contracts as predictive index for testing purpose in section 3.
3. Empirical analysis

3.1. Experiment Data and Evaluation criteria

The execution amount of service outsourcing contracts is selected as the dependent variable, standing for the development situation of service outsourcing. The main reasons for choosing this variable are: Firstly, the execution amount of service outsourcing contracts is an important indicator for service outsourcing industry, which can well represent the service outsourcing market. Secondly, it can reflect benchmark fluctuation cycle of this industry. We consider the following three points based on the economic theory to determine the factors affecting the service outsourcing: Firstly, the endogenous factors of service outsourcing, including economies of scale, employee, enterprise, etc. Secondly, support factors, including industry support, infrastructure support, etc. Thirdly, demand factors, including macroeconomic situation, the service industry development, etc. Therefore, we try to find the factors in these three areas as many as possible. Finally, 28 factors that may affect the service outsourcing are collected. The data used in this study is derived from national bureau of statistics of the people’s republic of China and wind database. In this paper, we use monthly data, the time period is from January in 2010 to December in 2014.

To measure the forecasting performance, we use the root mean squared error (RMSE) and Dstat as main criteria evaluation of level prediction. The formula of RMSE is as following:

$$RMSE = \frac{1}{\sqrt{N}} \sum_{t}^{N} (x(t)' - x(t))^2$$  \hspace{1cm} (12)

The directional statistic Dstat can measure predict movement direction, which can be expressed as:

$$Dstat = \frac{1}{N} \sum_{i=1}^{N} a_i$$  \hspace{1cm} (13)

$$a_i = \begin{cases} 0, & [x(t) - x(t - 1)][x(t)' - x(t - 1)] < 0 \\ 1, & [x(t) - x(t - 1)][x(t)' - x(t - 1)] \geq 0 \end{cases}$$  \hspace{1cm} (14)

where N is the number of sample, x(t)' is the predict value, x(t) is actual value.
3.2. Correlation analysis and Granger Causality test

In this section, we select the variables that will enter the model. We select the most correlated variables based on two principles: firstly, the target variables should be correlated with benchmark indicator. Secondly, there should be a cause-and-effect relationship between target variables and the execution amount of service outsourcing contracts. Table 1 shows the selected variables and the correlation coefficient. Table 2 shows the result of Granger Causality test. From table 1, we select seven variables, since they are significantly correlated with the dependent variable. From table 2, we can see that contracts amount of service, the number of enterprises, the number of certification, the number of employers, Gross Domestic Product, and the added value of the tertiary industry can granger cause the dependent variable. The result shows there is no granger causality relationship between the number of contracts and the dependent variable, so we remove this indicator from the model. The final variables that enter the model are CON, ENT, CERT, EMP, GDP and SER.

Table 1 Correlation analysis

<table>
<thead>
<tr>
<th>Indicator</th>
<th>abbreviation</th>
<th>Correlation coefficient</th>
<th>Correlation coefficient(-1)</th>
<th>Correlation coefficient(-2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>contracts amount of service outsourcing</td>
<td>CON</td>
<td>0.9276</td>
<td>0.6522</td>
<td>0.4605</td>
</tr>
<tr>
<td>the number of contracts</td>
<td>NUM</td>
<td>0.9226</td>
<td>0.4989</td>
<td>0.3589</td>
</tr>
<tr>
<td>the number of enterprises</td>
<td>ENT</td>
<td>0.8182</td>
<td>0.7735</td>
<td>0.7363</td>
</tr>
<tr>
<td>the number of certification</td>
<td>CERT</td>
<td>0.8165</td>
<td>0.7717</td>
<td>0.7376</td>
</tr>
<tr>
<td>the number of employers</td>
<td>EMP</td>
<td>0.8214</td>
<td>0.7763</td>
<td>0.7397</td>
</tr>
<tr>
<td>Gross Domestic Product</td>
<td>GDP</td>
<td>0.8230</td>
<td>0.7485</td>
<td>0.5953</td>
</tr>
<tr>
<td>the added value of the tertiary industry</td>
<td>SER</td>
<td>0.5670</td>
<td>0.5522</td>
<td>0.4758</td>
</tr>
<tr>
<td>Dependent variable</td>
<td>EXU</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2 Granger Causality test

<table>
<thead>
<tr>
<th>Null Hypothesis:</th>
<th>Obs</th>
<th>F-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>CON does not Granger Cause Y</td>
<td>58</td>
<td>2.36853</td>
<td>0.1035</td>
</tr>
<tr>
<td>Y does not Granger Cause CON</td>
<td></td>
<td>0.45644</td>
<td>0.6360</td>
</tr>
<tr>
<td>NUM does not Granger Cause Y</td>
<td>58</td>
<td>0.11296</td>
<td>0.8934</td>
</tr>
<tr>
<td>Y does not Granger Cause NUM</td>
<td></td>
<td>1.74385</td>
<td>0.1847</td>
</tr>
<tr>
<td>EMP does not Granger Cause Y</td>
<td>58</td>
<td>18.1748</td>
<td>1.E-06</td>
</tr>
<tr>
<td>Y does not Granger Cause EMP</td>
<td></td>
<td>3.32471</td>
<td>0.0436</td>
</tr>
<tr>
<td>GDP does not Granger Cause Y</td>
<td>58</td>
<td>17.6413</td>
<td>1.E-06</td>
</tr>
<tr>
<td>Y does not Granger Cause GDP</td>
<td></td>
<td>3.74717</td>
<td>0.0301</td>
</tr>
<tr>
<td>SER does not Granger Cause Y</td>
<td>58</td>
<td>3.57293</td>
<td>0.0350</td>
</tr>
<tr>
<td>Y does not Granger Cause SER</td>
<td></td>
<td>1.72792</td>
<td>0.1875</td>
</tr>
<tr>
<td>ENT does not Granger Cause Y</td>
<td>58</td>
<td>16.6330</td>
<td>2.E-06</td>
</tr>
<tr>
<td>Y does not Granger Cause ENT</td>
<td></td>
<td>4.50112</td>
<td>0.0156</td>
</tr>
<tr>
<td>CERT does not Granger Cause Y</td>
<td>58</td>
<td>16.6700</td>
<td>2.E-06</td>
</tr>
<tr>
<td>Y does not Granger Cause CERT</td>
<td></td>
<td>4.37424</td>
<td>0.0174</td>
</tr>
</tbody>
</table>
3.3. The EMD-VAR-SVR model

In this study, the EMD model is built by Matlab software package, which is produced by Mathworks Laboratory Corporation. All VAR models are implemented via the Eviews software package, which is produced by Quantitative Micro Software Corporation. The SVR model is built using Libsvm package. The EMD–VAR–SVR model uses VAR, regression and SVR model to predict IMFs extracted by EMD, and applies SVR for combination. In order to compare the forecasting capability of the proposed EMD-VAR-SVR ensemble method with other popular forecasting approaches, the vector auto-regression (VAR) model [30] and the single SVR model are used as the benchmark models.

3.3.1 EMD

We start the prediction experiments in terms of the previous steps shown in Section 2.4. First, we decompose the original data series “execution amount of service outsourcing contracts” into several independent IMFs and one residue by the technique of EMD. Figure 2 illustrate the decomposed result, the execution amount of service outsourcing contracts series is decomposed into three IMFs and one residue. IMFs and residue can always reflect the real economic significance. From the fluctuation of each IMF and residue, we can see that IMF1 is volatile, which may represent the random factor of service outsourcing. IMF2 and IMF3 have certain periodicity; they may have great correlation relationship with influence factors of service outsourcing, such as the economics scale of service outsourcing, macroeconomic development status, etc. The residue reflects the trend of service outsourcing. In order to verify the relationship between the selected variables and IMFs, we will have co-integration and granger causality test in next part.

![Fig2. the decomposition of the execution amount of service outsourcing contracts](image)

3.3.2 VAR and regression model

According to the characteristics of each IMF, we use different model to predict. To start the VAR model, we first do unit root test for all indicators. Under the confidence level of 10%, indicators such as the number of
service outsourcing contracts, the number of enterprises and so on cannot pass the test. After the first order
difference of all indexes, the indicators can pass through the ADF test, except IMF3 and R.

For IMF2, it is correlated with CON, ENT, CERT and EMP. Since the five variables are integrated of order
1, there may be a co-integration relationship among them, so we can carry out the co-integration test. The result
shows that there are at most 2 long-run co-integration relationships among variables. The first co-integration
equation can be present as:

\[ a_0 \text{IMF2}(-1) = a_1 \text{CON}(-1) + a_2 \text{ENT}(-1) + a_3 \text{CERT}(-1) + a_4 \text{EMP}(-1) \]

Where

\[ a_0 = -0.998983, a_1 = 0.100423, a_2 = -0.110722, a_3 = -0.067547, a_4 = -0.000796 \]

For IMF2, we use VAR (p) model to predict, the model can be shown as:

\[ Y_t = C + A_1 \cdot Y_{t-1} + A_2 \cdot Y_{t-2} + \cdots + A_p \cdot Y_{t-p} + \epsilon_t \]  

(15)

where: \( Y_t \) consists of k variables at time t; p is the number of lags; C is an vector modeling the baseline; \( A_i \) is
an matrix with each element reflecting the effect from the lag \( i = 1, \ldots, p \) of endogenous variable; and \( \epsilon_t \) is the
error term. According to the correlation analysis, there is a significant correlation between IMF2 with the
contracts amount of service outsourcing, the number of service outsourcing enterprises, the number of service
outsourcing certification, the number of service outsourcing employers. So we add these four variables into
the VAR model of IMF2. Here, we use VAR(2) model to predict IMF2.

Because IMF3 and residue R cannot pass the ADF test, it is unsuitable to built VAR model. We use regression
model to predict IMF3 and R. For IMF3, there is a significant correlation under the confidence level of 0.01
between IMF3 and Gross Domestic Product of China, which reflects the development status of China’s
macroeconomic, the correlation coefficient is 0.505. So we can build the model as follows:

\[ \text{IMF3} = C_2 + \beta_1 \cdot \text{IMF3}_{t-1} + \beta_2 \cdot \text{IMF3}_{t-2} + \beta_3 \cdot \text{GDP} \]  

(16)

\( \text{R} \) is correlated with the added value of the tertiary industry, that means, the development of service
outsourcing are greatly influenced by the service industry; the increase of service industry development level will
drive China's service outsourcing contract revenue growth. The correlation coefficient between \( R \) and the added
value of the tertiary industry is 0.539, and it is significant correlation under the confidence level of 0.01. The
regression model of \( R \) is as follows:

\[ R = C_3 + \gamma_1 \cdot R_{t-1} + \gamma_2 \cdot R_{t-2} + \gamma_3 \cdot \text{SER} \]  

(17)

3.3.3 SVR and result

Since IMF1 is volatile, we use SVR to model it. The result of IMF1’s autocorrelation analysis shows that
IMF1 is correlated with its lag phase 6 data. So we can set up the prediction step length to be 6. That is to say,
for the SVR model, the output variable is \( \text{IMF1}_{t} \), the corresponding input variables are
\{IMF1\_{t-6}, IMF1\_{t-5}, IMF1\_{t-4}, IMF1\_{t-3}, IMF1\_{t-2}, IMF1\_{t-1}\} . The formula is as following:

\[ \text{IMF1}_t = \sum_{i=1}^{6} a_i \cdot \text{IMF1}_{t-i} \]  

(18)

For the ensemble model, the SVR model is used to gain the final predict result. The predict value of IMF1,
IMF2, IMF3 and R produced by the above model are treated as the input variables, the original execution amount
of service outsourcing contracts as the output variable. We put the data from January 2010 to September 2014 as
the training sample data, the data from October 2014 to December 2014 as the test sample data. Based on the
same data set, this paper uses single SVR and single VAR model for comparison, the prediction results are shown
in table 5. From this table, we can generally see that the prediction result of the proposed EMD-VAR-SVR
method is promising for the execution amount of service outsourcing contracts series. The measurement of
forecasting performance is goodness-of-fit, such as RMSE and Dstat .The experimental results show that the
EMD-VAR-SVR integrated forecasting method is superior to the single SVR model and single VAR model.

<table>
<thead>
<tr>
<th>time</th>
<th>Real value</th>
<th>EMD-VAR-SVR</th>
<th>SVR</th>
<th>VAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>2014-10</td>
<td>66.6</td>
<td>66.9</td>
<td>67.9</td>
<td>59.1</td>
</tr>
<tr>
<td>2014-11</td>
<td>77.7</td>
<td>83.0</td>
<td>75.9</td>
<td>65.9</td>
</tr>
<tr>
<td>2014-12</td>
<td>123.9</td>
<td>122.1</td>
<td>90.8</td>
<td>75.0</td>
</tr>
<tr>
<td>RMSE</td>
<td>3.24</td>
<td>19.1</td>
<td>29.3</td>
<td></td>
</tr>
<tr>
<td>Dstat</td>
<td>1</td>
<td>0.667</td>
<td>0.667</td>
<td></td>
</tr>
</tbody>
</table>

4. Conclusion

This paper proposes an EMD-VAR-SVR integrated method to predict the execution amount of service outsourcing contracts. The experimental results prove that the proposed EMD-VAR-SVR integrated model has better prediction precision than the single SVR and single VAR method, which has the minimum RMSE and largest Dstat in all models. At present, quantitative analysis of the research on the influencing factors of service outsourcing is always based on simple econometric models, this article first proposed EMD-VAR-SVR integrated approach based on TEI @ I methodology. By decomposing original service outsourcing time series data to several composition, we are able to describe the economic meaning, influence factors and characteristics of service outsourcing industry more accurately. Our research shows that, the execution amount of service outsourcing contracts is mainly influenced by economic scale of service outsourcing, service industry development and operation situation of macroeconomic. In order to promote the healthy development of service outsourcing industry and maintain the stability growth of the service outsourcing revenue, we think we can take measures from several aspects. Firstly, we should actively cooperate with the well-known international software enterprises, thus, accelerating the globalization process of service outsourcing enterprises. Secondly, adjust the structure of service outsourcing industry, improve the quality of the service outsourcing enterprises, and cultivate leading enterprises of industry. Thirdly, increase the education investment in service outsourcing industry.

Acknowledgements

We would like to thank China’s service outsourcing research center for providing the data set. This work has been partially supported by grants from National Natural Science Foundation of China (NO. 71331005, NO71110107026, NO. 71501175), the Generalized Virtual Economy Research Fund (NO. GX2011-1001(Z)), and the Open Project of Key Laboratory of Big Data Mining and Knowledge Management, Chinese Academy of Sciences.

References


