DE GRUYTER
OPEN

BULGARIAN ACADEMY OF SCIENCES

CYBERNETICS AND INFORMATION TECHNOLOGIES • Volume **16**, No 6
Special issue with selection of extended papers from 6th International Conference on Logistic, Informatics and Service Science LISS'2016

Sofia • 2016 Print ISSN: 1311-9702; Online ISSN: 1314-4081
DOI: 10.1515/cait-2016-0079

Cloud Computing and Extreme Learning Machine for a Distributed Energy Consumption Forecasting in Equipment-Manufacturing Enterprises

Rui-Dong Wang[1], *Xue-Shan Sun*[2], *Xin Yang*[2], *Haiju Hu*[3]

[1]*Department of Mathematics, Tianjin University of Technology, Tianjin 300384, China*
[2]*ZhongHuan Information College, Tianjin University of Technology, Tianjin 300380, China*
[3]*Economics and Management School, Yanshan University, Qinhuangdao, Hebei 066004, China*
Emails: wangrdtjut@126.com crystal171@live.cn wing.lps@163.com huhaiju@ysu.edu.cn

Abstract**:** *Energy consumption forecasting is a kind of fundamental work of the energy management in equipment-manufacturing enterprises, and an important way to reduce energy consumption. Therefore, this paper proposes an intellectualized, short-term distributed energy consumption forecasting model for equipment-manufacturing enterprises based on cloud computing and extreme learning machine considering the practical enterprise situation of massive and high-dimension data. The analysis of the real energy consumption data provided by LB Enterprise was undertaken and corresponding calculating experiments were completed using a 32-node cloud computing cluster. The experimental results show that the energy consumption forecasting accuracy of the proposed model is higher than the traditional support vector regression and the generalized neural network algorithm. Furthermore, the proposed forecasting algorithm possesses excellent parallel performance, overcomes the shortcoming of a single computer's insufficient computing power when facing massive and high-dimensional data without increasing the cost.*

Keywords**:** *Energy consumption forecasting, cloud computing, online sequential optimization, extreme learning machine, equipment-manufacturing enterprises.*

1. Introduction

Equipment-manufacturing industry is one of the high energy consuming industries in China, so the equipment-manufacturing enterprises are facing tremendous energy-saving pressure and the energy management becomes an effective way to save energy. As a kind of fundamental work of the energy management in equipment-manufacturing enterprises, energy consumption forecasting plays an important role in the enterprise production management [1]. It is the premise for the production scheduling optimization, for the guarantee of stability and economy of the production system, and for the improvement of energy efficiency ultimately.

According to the forecasting duration time, energy consumption forecasting can be classified into short, middle and long-term respectively [2]. Middle and long-term energy consumption forecasting is generally utilized in energy planning and scheduling system [3]. While the short-term energy consumption forecasting plays critical role not only in energy allocation and coordination system [4] but also in energy costs reduction [5, 6]. Therefore, short-term energy consumption forecasting has a major importance in equipment manufacturing enterprises [7, 8]. The larger estimation errors yield, the higher operation costs. It is shown in [9] that a small increase in forecasting accuracy would save millions of dollars in operation costs. In addition, energy purchasing and bidding also require short-term energy consumption forecasting [10, 11].

The modeling problem is a key factor affecting the significance and essence of energy consumption forecasting. So the development of models for energy consumption forecasting has become one of the fastest progressing areas of research. Generally, the quality of the model is evaluated from three aspects: (1) forecasting accuracy, (2) computational complexity and (3) forecasting speed [12]. Current approaches can be classified into three categories: econometric approaches involving the use of time series 0 and regression 0; soft computing approaches, such as gray forecasting 0 and fuzzy logic methods 0; and intelligent optimization approaches, such as genetic algorithms 0, neural networks 0, ant colony and particle swarm optimization 0. In addition, many hybrid forecasting models have also been developed, which can make use of the advantages of each involved technique. Studies show that the energy consumption forecasting is a tricky issue, since large load variations are possible in a period of an hour or a day. These variations depend on many individual parameters such as weather conditions [4, 6], temperature [20], and day of week [20], hour of the day [21], seasonal factors [22, 23], social activities [5] and socio-economic factors [24]. Therefore, those models are based on establishing a relation between the energy consumption and some of the relevant factors to achieve higher accuracies [21] because of the nonlinear characteristic of these relations. Studies regarding the energy consumption forecasting using intelligent optimization approaches, such as Artificial Neural Network (ANN) [4, 20, 25, 26] and Support Vector Regression (SVR) [5, 22, 23] are raised depending on their efficiency, which showed high forecasting accuracy. However, SVR is a type of batch learning algorithm, which cannot replicate real-time online learning. When the new sample data is added to the training set, the batch learning model will be trained using both the original training data and the new data, which will consume a large amount of computing resources and training time. In practical use, especially in short-term energy consumption prediction, the forecasting value is continually added to actual data sets and the next day’s energy consumption value is forecasted; so, as a traditional energy consumption forecasting algorithm, SVR is not suitable for energy consumption forecasting application scenarios.

The motivation of this study is to develop an efficient way for energy consumption forecasting of the equipment-manufacturing enterprises. In recent years, there has been increasing interest in Extreme Learning Machine (ELM), which has the characteristics of self-adaptive ability, independent learning and optimization calculation to a high degree of non-structural and non-precise regularity [27]. Through a large number of experiments, literature [27] shows that: in regression, two-classificatory and multi-classificatory problems, ELM is better than SVR and BP algorithm with regards to computing performance, accuracy and the training and testing time. ELM is introduced into the power load forecasting for the first time in the literature [28], and the better prediction performance is obtained. However, it is not entirely suitable for energy consumption scenarios because of the high dimensional trend of the energy consumption data. Therefore, considering the actual application of the energy consumption forecasting scenarios and the evaluation criterion of the model's quality, a distributed extreme learning machine for energy consumption based on MapReduce is presented. An online sequential extreme learning machine is applied for energy consumption forecasting, and the cloud computing technology and multi-agent technology are introduced to improve the processing ability of the high-dimensional data and the accuracy of the energy consumption forecasting without the increasing of the cost. The parallel performance of the improved algorithm and the accuracy of energy consumption forecasting are tested on a set of clusters gathered in the laboratory, and the actual energy consumption data is used to carry out the case analysis.

The rest of this paper is organized as follows. Section 2 presents the relevant theories and describes the implementation of the proposed methodology. Section 3 provides a case analysis and comparisons with other well-established methods. Section 4 outlines the conclusions.

2. Proposed methodology

2.1. Cloud computing

Hadoop platform developed by Apache is a successful open-access implementation, supported by MapReduce framework and a distributed file system. MapReduce is a programming framework for processing massive datasets in the cluster, originally proposed by Google, which is used to solve the problem of distributed storage and computing. MapReduce framework greatly reduces the difficulty of implementation by masking the underlying specific implementation details. Complex parallel computing on a large scale cluster is abstracted to two functions that can be written by the user, namely "Map" and "Reduce" [29]. MapReduce operating mechanism is shown in Fig.1.

The specific description of MapReduce is as follows:

(1) *Input*. Firstly, the input data is read from the distributed file system, and then is cut into a data sheet. MapReduce framework allocates the data sheet for each Map function.

(2) *Map*. The MapReduce framework considers the input data slice as a (Key, value) pair. According to the Map function which is written by the users, the (Key, value) pair which is distributed by the framework is operated. Finally, the intermediate Key/value pairs are generated.

(3) *Shuffle*. The intermediate key/value pair is transferred from the map nodes to reduce nodes at this stage; it takes longer time than operating the map function and reduces function because of the influence of bandwidth, CPU speed and so on. In this stage all intermediate values associated with the same intermediate key are merged, the key/list of values is formed and the sorting is done.

(4) *Reduce*. The reduce function written by users is executed. All the intermediate values and the corresponding intermediate keys or the intermediate list of values are iterated, the data processing logic set by users is operated, and the new key/value pairs are generated.

(5) *Output*. The output by reduce function is transferred to the specified file system path.

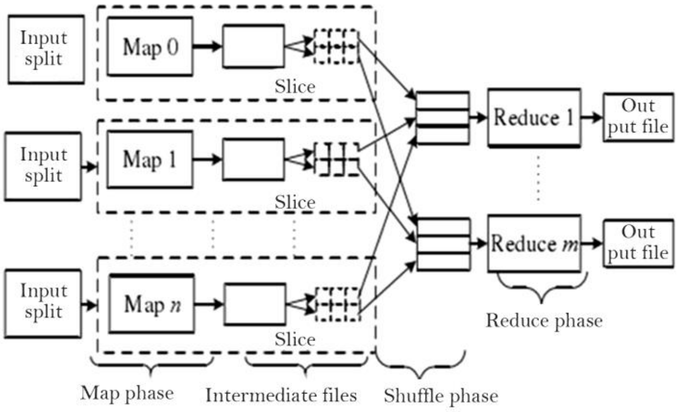

Fig. 1. MapReduce operating mechanism

2.2. Extreme learning machine algorithm

2.2.1. Description of extreme learning machine algorithm

Extreme learning machine algorithm is different from a conventional feedback neural network training learning algorithm because the hidden layer is not iterated; at the same time, the input weights and hidden neurons' threshold is randomly selected. The goal is to minimize the training error and the output weights of the hidden layer are determined by the algorithm. The specific algorithm of extreme learning machine [27] is described as follows:

For N arbitrary distinct samples, (x_i, t_i), where $x_i=[x_{x1}, x_{x2},..., x_{xn}]^{\mathrm{T}} \in R^n$, $t_i=[t_{i1}, t_{i2},..., t_{im}]^{\mathrm{T}} \in R^m$, the limit learning machine model with N hidden neurons and an activation function G can be mathematically modeled as

(1) $$f_{\tilde{N}} = \sum_{i=1}^{\tilde{N}} \beta_i G(a_i, b_i, x_j) = t_j, \quad j = 1, ..., N,$$

where $a_i = [a_1, a_2, ..., a_n]^{\mathrm{T}}$ are the weight vectors of the i-th hidden neuron and the input neurons; b_i is the threshold of the i-th hidden neuron; $\tilde{N}$ is the number of hidden neurons. Formula (1) can be abbreviated as:

(2) $$H\beta = T,$$

(3) $$H(a_1, ..., a_{\tilde{N}}, b_1, ..., b_{\tilde{N}}, x_1, ..., x_N) = \begin{bmatrix} G(a_1, b_1, x_1) & ... & G(a_{\tilde{N}}, b_{\tilde{N}}, x_1) \\ ... & ... & ... \\ G(a_1, b_1, x_N) & ... & G(a_{\tilde{N}}, b_{\tilde{N}}, x_N) \end{bmatrix}_{N \times \tilde{N}},$$

(4) $$\begin{cases} \beta = \left[\beta_1^{\mathrm{T}} ... \beta_{\tilde{N}}^{\mathrm{T}} \right]_{\tilde{N} \times m}, \\ T = \left[t_1^{\mathrm{T}} ... t_N^{\mathrm{T}} \right]_{N \times m}. \end{cases}$$

In Formula (2) H is the hidden layer output matrix of the neural network, the i-th column of H is the i-th hidden neuron's output vector with respect to inputs $x_1, x_2, ..., x_n$. The output weights can be obtained by solving the next linear equations to obtain the least squares solution:

(5) $$\|H\beta - T\| = \|HH^{\mathrm{T}}T - T\| = \min_{\beta} \|H\beta - T\|.$$

The least squares solution is

(6) $$\beta = H^{\mathrm{T}}T,$$

where H^{T} is the Moore-Penrose generalized inverse of the matrix H.

2.2.2. Deficiencies of ELM

In [28], the ELM has utilized for load forecasting showed the ELM outperforms SVM and BP algorithm in computing performance, accuracy, and training and testing time. However, the ELM is not entirely suitable for energy consumption scenarios because it is a batch learning algorithm, at this time, online sequencing is of utmost necessity. When faced with new data added to the training set, the online sequential ELM does not need to re-train. Additionally, in order to cope with the high dimensional trend of the energy consumption data, cloud computing technology and multi-agent distributed technology are used to improve the capability of data processing of massive high-dimensional data and improve the accuracy of energy consumption forecasting. This algorithm is called the MapReduce Weighted Averaged Online Sequential Extreme Learning Machine (MR-OSELM-WA).

2.3. The design of MR-OSELM-WA algorithm based on cloud computing

2.3.1. The proposed online sequential extreme learning machine algorithm

The steps of the sequential extreme learning machine algorithm [30] are as follows:

Step 1. Initial stage. A part of the data set is determined as the initial training set; the number of hidden neurons is manually set as $\tilde{N}$, where $k = 0$, k is a counter. Firstly, the weight vector $w_i = [w_1, w_2, ..., w_n]^{\mathrm{T}}$ of the i-th hidden neuron, the input neurons and the activation function parameters are generated randomly. Then the initial hidden output matrix H_0 and the initial output weight vector β_0 are calculated:

(7) $$P_0 = (H_0^{\mathrm{T}} H_0)^{-1},$$

(8) $$\beta_0 = P_0 H_0 T_0,$$

where $T_0 = [t_1, t_2, ..., t_N]$ is the target value of the output vector.

Step 2. Online sequential learning. When new training data arrives, we assume that it is the $(k + 1)$-th sample data of all the training set data, the hidden layer output matrix is first calculated, i.e.,

(9) $$H_{k+1} = [G(a_1, b_1, x_k), ..., G(a_L, b_L, x_{k+1})],$$

then set $T_{k+1} = \left[t_{(\sum_{j=0}^{k} N_j)+1}, ..., t_{(\sum_{j=0}^{k} N_j)+j} \right]^{\mathrm{T}}$.

Step 3. The output weights β_0^{k+1} can be calculated by the following formula:

(10) $$\begin{cases} P_{k+1} = P_k - P_k H_{k+1}^{\mathrm{T}} (I + H_{k+1} P_k H_{k+1}^{\mathrm{T}})^{-1} H_{k+1} P_k, \\ \beta^{(k+1)} = \beta^{(k)} + K_{k+1}^{-1} H_{k+1}^{\mathrm{T}} (T_{k+1} - H_{k+1} \beta^{(k)}). \end{cases}$$

Step 4. $k = k + 1$, return to Step 1, to train the next data.

2.3.2. MR-OSELM-WA algorithm based on cloud computing

(1) The theory of the MR-OSELM-WA algorithm

With the development of equipment-manufacturing enterprise intelligence, the energy consumption data is increasing in an exponential quantity, it is no longer sufficient to use OSELM algorithm for energy consumption forecasting. With the development of cloud computing models being used in intelligent manufacturing systems, the extreme learning machine is improved with the use of multi-agent thinking and cloud computing technology [31].

The idea of the MR-OSELM-WA algorithm is that the OS-ELM is running at each agent, and the higher the OS-ELM node forecasting accuracy is, the higher weights are obtained by the node. The final forecasting value is calculated by the weighted average of each node's forecasting value. The weight of each node is set as followed:

(11) $$\bar{y} = \sum_{k=1}^{K} \alpha_k y_k \Big/ \sum_{k}^{K} \alpha_k,$$

where $\bar{y}$ is the final forecasting value, y_k is the forecasting value of the k-th node and α_k is the weight of the k-th node.

The weight can be calculated by the standard error function E and the gradient ascent method,

(12) $$E = (1/2)(t - \bar{y})^2.$$

In formula (12), t is the target value of each OS-ELM agent's input training set. According to the gradient ascent strategy, α_k is defined as

(13) $$\Delta\alpha_k = -\eta \partial E / \partial \alpha_k = -\eta(\partial E / \partial \bar{y}) \cdot (\partial \bar{y} / \partial \alpha_k),$$

where η is the learning rate, set at $\eta = 0.1$. It can be written as

(14) $$\Delta\alpha_k = \eta(t - \bar{y}) \left[y_k / \sum_{i=1}^{K} \alpha_i - \sum_{i=1}^{K} \alpha_i y_i / (\sum_{i=1}^{K} \alpha_i)^2 \right],$$

where α_k can be updated by $\alpha_k \leftarrow \alpha_k + \Delta\alpha_k$.

(2) The specific steps of MR-OSELM-WA

The core idea of MapReduce framework is the design of the map and the reduce operations; its parallel implementation whereby the intermediate results of MR-OSELM-WA are saved into the distributed database “HBase” and the distributed cache. MapReduce based distributed MR-OSELM-WA forecasting model is shown on Fig. 2.

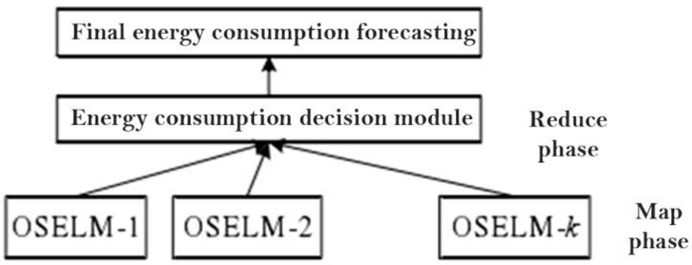

Fig. 2. MapReduce based distributed MR-OSELM-WA forecasting model

The steps of MR-OSELM-WA are as follows:

1) The massive input training set is read from the cloud computing platform of the distributed file system. Through the underlying mechanism of the MapReduce framework and the segmentation of the training set, the k different training subsets can be obtained. The value of k is equal to the number of parallel Maps in the cloud computing cluster.

2) The subset can be trained in accordance with the logic steps of the Map function, that is, the logic of the OS-ELM is to perform parallel training on the set, which is equal to the k different learning machines.

3) The results of the map operation, namely k different machine learning forecasting values, can be transferred from the shuffle phase to the reduce phrase through the MapReduce framework; hence, the weights of forecasting value of map operation output can be determined according to the weight calculation.

4) According to the requirement of the sequential learning mechanism, the energy consumption data of the next day can be predicted along the time axis of the data.

3. Case Analysis

3.1. Preparation phase

3.1.1. Experimental environment

Hadoop platform, which consists of 32 nodes, can be built in the laboratory, with each node configured for Intel (R) Core (TM) i5-2400 4-core CPU @ 2.60 GHz, 4 GBRAM and network bandwidth of 1000 Mbit/s. The Hadoop is Version 0.20.2. The cloud computing platform is shown on Fig. 3.

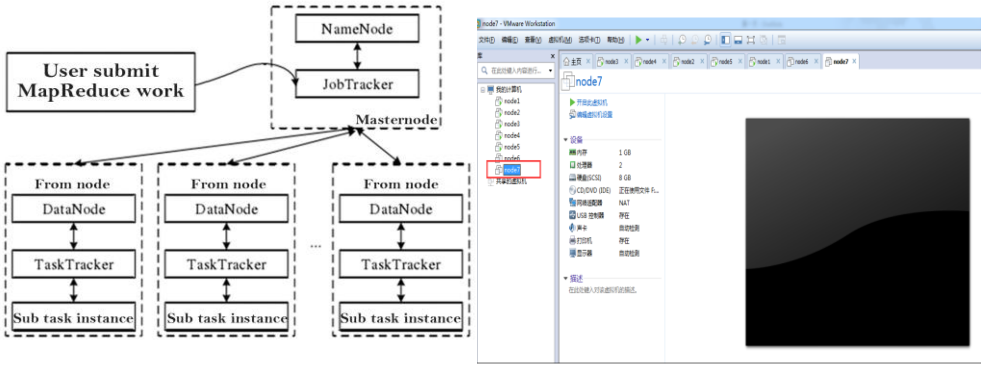

Fig. 3. The cloud computing platform

3.1.2. Background of the LB Enterprise

LB Enterprise is an equipment manufacturing enterprise. Currently, owing to the tight market in China, LB Enterprise is affected by this situation and has experienced order reduction. At the same time, the energy-saving and emission-reduction policies in China also bring great challenges to LB Enterprise. Therefore, it is very urgent to solve the energy consumption forecasting problems in LB Enterprise. It is shown from our investigation of LB Enterprise that electricity is the most highly consumed energy source as it is used in all production processes. Generally, it accounts for 47-67 % of the total consumption, occasionally peaking above 80%. Therefore, electricity consumption is chosen as the forecasting object of the study.

3.1.3. Factors affecting energy consumption in the production process of LB Enterprise

According to the production characteristics of LB Enterprise and the literature reviewed, the factors affecting the energy consumption include not only certain macroscopic elements, such as the environment and economy but also ones related to production processes. The factors from the perspectives of the production, the environment and the economy are summarized in Table 1.

Table 1. The factors affecting energy consumption in the production process

Type	Level 1	Index
Production	Production parameters [32]	Operating time Production time
	Production material parameters [33]	Raw material inputs Raw material
Environment	Weather conditions [34, 35]	Temperature Wind power
Economy	Production outputs [36, 37]	Production
	Market factors	Historical production Historical inventory Historical price

3.1.4. Data sources and pretreatment

The operating time is obtained from statistics of the company and the remaining factors from the company's ERP system. Because there are 32 different types of raw materials used by the forging factory in LB Enterprise, taking them all as input variables of the model will create high-dimensions in the input data and introduce redundant information to limit the performance of the forecasting model. Therefore, we conduct a correlation analysis through the data on historical input, historical energy consumption and historical electricity consumption of all 32 types of raw materials. This analysis identified the types of raw materials that are most relevant to energy consumption as input variables of the model. The environmental data are obtained from the Chinese weather website (**http://www.tianqi.com/**), and the market and economic data are from the China coal industry website **(http://www.coalchina.org.cn/**).

Table 2. Input variables of the energy consumption forecasting model

Serial number	Public input variables	Raw materials
1	operating time	RM6
2	production time	RM8
3	Input of raw material	RM17
4	average high temperature	RM20
5	average low temperature	RM25
6	the average wind power	
7	production	
8	coal production of China	
9	coal stockpiles of China	
10	CCPI	
11	BSPI	

The coal price was measured using the price index from the China National Coal Association, including China's Coal Prices Index (CCPI) and the Bohai Sea Power Coal Prices Index (BSPI). Therefore, the final input variables of the energy consumption forecasting model are shown in Table 2, where RM stands for certain

Raw Materials. The sample includes every day's input variables for energy consumption of the forging factory from June 2014 to July 2015.

We standardize data using the Z-Score standardization method instead of the minimum-maximum standardization to avoid the effects of isolated points. The formula of the Z-Score is typically used as follows:

$$\overline{x_i} = \frac{1}{n}\sum_{j=1}^{n} X_i^j,\ \ \sigma_i^2 = \frac{1}{n-1}\sum_{j=1}^{n}(X_i^j - \bar{X}_i)^2,\ \ \bar{X}_i^j - \frac{x_i^j - \bar{X}_i}{\sigma_i}, \tag{15}$$

where $\bar{X}_i$ and σ_i are the mean and standard deviation of the attribute, x_i^j and $\bar{X}_i^j$ are the original data and standardized data, $j = 1,\ldots\ldots, n,$ and n is the dimension. $E = \left[E_{i-7}, E_{i-6}, E_{i-5}, E_{i-4}, E_{i-3}, E_{i-2}, E_{i-1}\right],$ indicates the energy consumption value of 7 days before the forecasting day. The experimental objective is to forecast every day's energy consumption in July 2015. The output of the training set is $y_i = E_i,$ the prediction of the energy consumption value.

3.1.5. The testing index

(1) The measure of accuracy of the energy consumption forecasting is called M_{APE}, which is used as the testing index,

$$M_{\mathrm{APE}} = 100\left[\sum_{i=1}^{n}\left|\left[\left(E_i - \hat{E}_i\right)/E_i\right]/n\right|\right],\ n = 31, \tag{16}$$

where E_i and $\hat{E}_i$ indicate the actual energy consumption value and the forecasting value of the i-th day respectively; n is the number of days of the forecasting month. Regarding the energy consumption, the smaller the value of M_{APE}, the greater the accuracy of the energy consumption forecast.

(2) The parallel performance of the MR-OSELM-WA is tested by the s_{peedup} and the s_{caleup} values. s_{peedup} is a measure of the performance and effectiveness of the parallel system or program, as shown:

$$s_{\mathrm{peedup}} = \frac{\text{Execution time on a single machine}}{\text{execution time on cloud}}; \tag{17}$$

s_{caleup} is used to compare the execution time of an m-times increase of the cluster with the increase of m-times of the task data with the original data set,

$$s_{\mathrm{caleup}} = \frac{m \cdot \text{Execution time of the data set}}{\text{Execution time of the orginal data set}}. \tag{18}$$

3.2. The results analysis

3.2.1. The forecasting accuracy of MR-OSELM-WA

The MR-OSELM-WA, SVR [38] and functional networks [39] are compared in this experiment; SVR and functional networks show extremely good ability in energy consumption forecasting. By comparing with these two algorithms, the performance of MR-OSELM-WA algorithm proposed in this paper is tested.

In order to calculate the M_{APE} value of the formula (16) as the objective function, the optimal parameters of the three algorithm models are obtained by cross-validating ten times. To obtain the optimal values of the parameters for each algorithm, the historical data of June 2014 to June 2015 are used as the training sets, and the MR-OSELM-WA, functional network algorithm and SVR are trained separately to forecast the energy consumption value in July 2015.

To ensure the objectivity of the experimental results, the experiment is performed 50 times, and the average value is used as the final result of the experiment. The M_{APE} values in energy consumption forecasting of the three algorithms are shown in Table 3. It is obviously that the M_{APE} value of MR-OSELM-WA is smallest which means the accuracy of the MR-OSELM-WA in energy consumption forecasting is highest. In addition, the SVR and functional networks use batch learning mode, so when the training set is large, the operation of the batch learning algorithm requires more memory space. Once the memory space exceeds a certain limit, the efficiency of the implementation of the algorithm will be substantially reduced; however, the training set's bulk of the MR-OSELM-WA is smaller than the ELM and the above phenomenon is not readily apparent.

Figs 4 and 5 are the comparisons of the actual energy consumption value and the forecasting value by MR-OSELM-WA, SVR and functional network algorithm in July 2015.

Table 3. Forecasting M_{APE} based on MR-OSELM-WA, functional networks and SVR in test samples

Forecasting algorithm	M_{APE}
MR-OSELM-WA	1.872
Functional networks	3.250
SVR	2.463

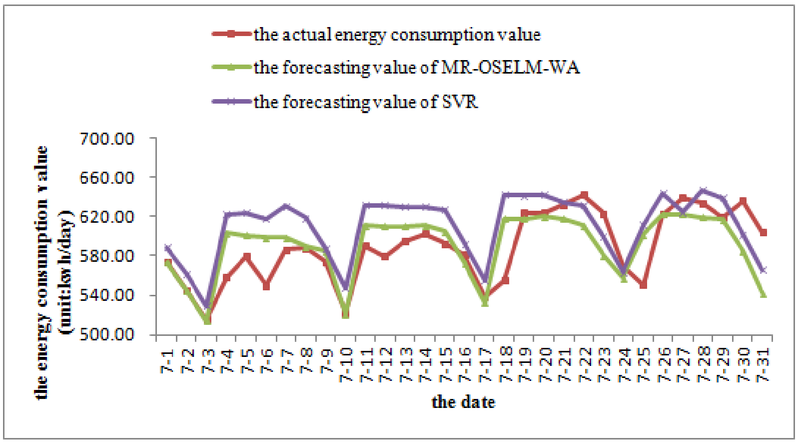

Fig. 4. Comparisons of the actual energy consumption value and the forecasting value by MR-OSELM-WA and SVR in July 2015

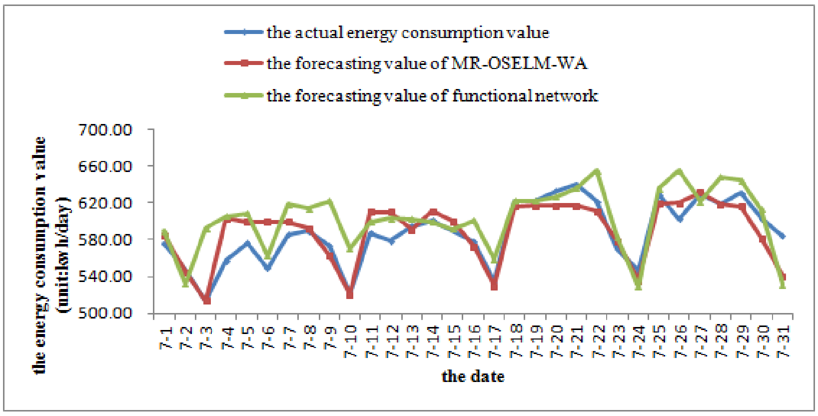

Fig. 5. Comparisons of the actual energy consumption value and the forecasting value by MR-OSELM-WA and the functional network in July 2015

3.2.2. The parallel performance of MR-OSELM-WA

In order to reflect the parallel performance of the MR-OSELM-WA, the energy consumption sample data provided by the example can be expanded to 1000 times, 2000 times, 4000 times and 8000 times of the original data set artificially. The four different-sized data sets can be formed and run on the cloud platform of the number of 4, 8, 16, 32 cluster nodes respectively, in order to calculate the s_{peedup} and the S_{caleup}.

The s_{peedup} will be close to one in the perfect parallel system. But in practical application, with the increase of nodes, the consumption of transmission among nodes in the network will increase continually, as shown in Fig. 6, so the linear accelerated ratio is very difficult to achieve. From Fig. 6, it is apparent that as the size of the data set increases, that rate of increase of the MR-OSELM-WA s_{peedup} value increases in a linear fashion, especially for large data sets. In practical application, the greater the number of the data, the higher the value of MR-OSELM-WA s_{peedup}, that is, the MR-OSELM-WA meets the needs of the energy consumption forecasting of massive, high dimensional data.

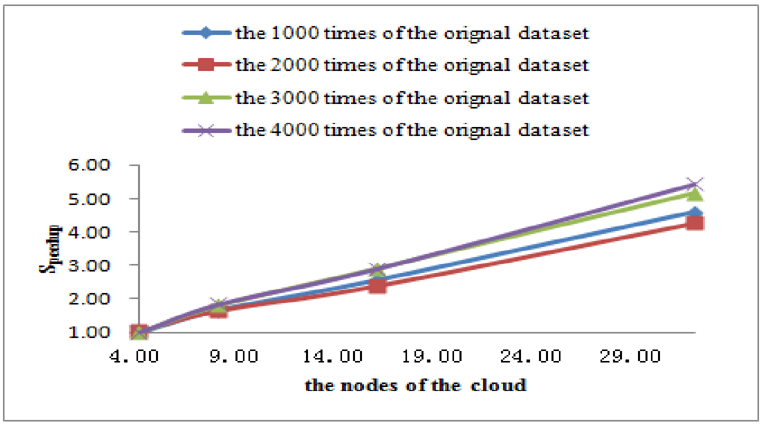

Fig. 6. The s_{peedup} of MR-OSELM-WA

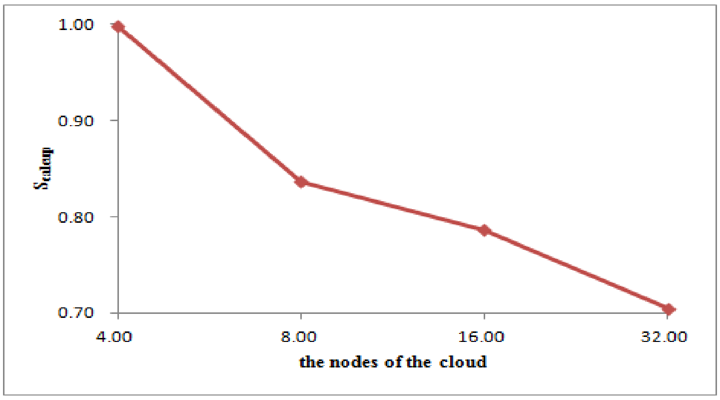

Fig. 7. The s_{caleup} of MR-OSELM-WA

In the perfect parallel system, the s_{caleup} is equal to one constant, but it is impossible to achieve this in practical application. As the size of the dataset increases, the s_{caleup} of the parallel system is gradually reduced. Experimental results are shown on Fig. 7. When the data set is larger, the contraction rate of the MR-OSELM-WA becomes smaller, so the performance of s_{caleup} of the MR-OSELM-WA is better.

4. Conclusions

With the development of intelligence in equipment-manufacturing enterprises, the trend towards the use of massive and high dimensional data is inevitable; therefore, the complexity of the algorithm which is widely used in energy consumption forecasting will continue to increase, which means that a stand-alone computer cannot handle such large computational requirements. In recent years, the prevalence of big data processing technology has provided an effective way to solve this problem, and the problem of parallelism of the algorithm has become an important area of research within the field of energy consumption forecasting. The algorithm proposed in this paper can not only shorten the training time and reduce the consumption of computing resources, but also significantly improve the accuracy of energy consumption forecasting.

***Acknowledgments*:** The authors are supported by the Natural Science Foundation of China (Grant No 11301384).

References

1. D u, Y.-W. The Design and Implementation of Energy Management and Data Correction System in Iron and Steel Enterprise. Central South University, 2010.
2. R e, A.-A. Improving Electric Load Forecasts Using Network Committees. – Electric Power Syst. Res., Vol. **74**, 2005, pp. 83-94.
3. H a h n, H., S. M e y e r-Ni e b e r g, S. P i c k l. Electric Load Forecasting Methods: Tools for Decision Making. – Eur. J. Oper. Res., Vol. **199**, 2009, No 3, pp. 902-907.

4. C h o w, T. W. S., C. T. L e u n g. Neural Network Based Short-Term Load Forecasting Using Weather Compensation. – IEEE Trans. Power Syst., Vol. **11**, 1996, No 4, pp. 1736-1742.
5. H o n g, W. C. Electric Load Forecasting by Support Vector Model. – Appl. Math. Model, Vol. **33**, 2009, No 5, pp. 2444-2454.
6. M e t a x i o t i s, K., A. K a g i a n n a s, D. A s k o u n i s, J. P s a r r a s. Artificial Intelligence in Short Term Electric Load Forecasting: A State-of-the-Art Survey for the Researcher. – Energy Convers Manage, Vol. **44**, 2003, No 9, pp. 1525-1534.
7. A u n g, Z., M. T o u k h y, J. W i l l i a m s, A. S a n c h e z, S. H e r r e r o. Towards Accurate Electricity Load Forecasting in Smart Grids. – In: 4th International Conference on Advances in Databases, Knowledge, and Data Applications, DBKDA, 2012. p. 51-57.
8. H e r n a n d e z, L., C. B a l a d r ó n, J. M. A g u i a r, B. C a r r o, A. J. S a n c h e z-E s g u e v i l l a s, J. L l o r e t. Short-Term Load Forecasting for Microgrids Based on Artificial Neural Networks. – Energies, Vol. **6**, 2013, No 3, pp. 1385-1440.
9. B u n n, D. W., E. D. F a r m e r. Comparative Models for Electrical Load Forecasting. Chichester, NY, Wiley, 1985.
10. J a v e d, F., N. A r s h a d, F. W a l l i n, I. V a s s i l e v a, E. D a h l q u i s t. Forecasting for Demand Response in Smart Grids: An Analysis on Use of Anthropologic and Structural Data and Short Term Multiple Loads Forecasting. – Appl Energy, 2012, No 96, pp. 150-160.
11. S o l i m a n, S. A., A. M. A l k a n d a r i. Electrical Load Forecasting: Modeling and Model Construction. Burlington, MA, Butterworth-Heinemann, 2010.
12. T i a n, Z.-Q., X.-B. H u a n g. Overview of Data Mining Technologies for Industrial Processes. – Control Engineering of China, 2009, No 16, pp. 120-123 (in Chinese).
13. K u m a r, U., V. K. J a i n. Time Series Models (Grey-Markov). Grey Model with Rolling Mechanism and Singular Spectrum Analysis to Forecast Energy Consumption India. – Energy, Vol. **35**, 2010, No 4, pp. 1709-1716.
14. B r a u n, M. R., H. A l t a n, S. B. M. B e c k. Using Regression Analysis to Predict the Future Energy Consumption of a Supermarket in the UK Applied. – Energy, 2014, No 130, pp. 305-313.
15. Z h o u, P., B. W. A n g, K. L. P o h. A Trigonometric Grey Prediction Approach to Forecasting Electricity Demand. – Energy, Vol. **31**, 2006, No 14, pp. 2839-2847.
16. W u, Z., J. X u. Predicting and Optimization of Energy Consumption Using System Dynamics-Fuzzy Multiple Objective Programming in World Areas. – Energy, 2013, No 49, pp. 19-31.
17. Y u, F., Z. X i a o. A Short-Term Load Forecasting Model of Natural Gas Based on Optimized Genetic Algorithm and Improved BP Neural Network. – Applied Energy, 2014, No 134, pp. 102-113.
18. E k o n o m o u, L. Greek Long-Term Energy Consumption Prediction Using Artificial Neural Networks. – Energy, 2010, No 35, pp. 512-517.
19. A r a s h, G., S. M. R. K a z e m i, F. M e h m a n p a z i r, M. M. N a k h o s t i n. A Cooperative Ant Colony Optimization-Genetic Algorithm Approach for Construction of Energy Demand Forecasting Knowledge-Based Expert Systems. – Knowledge-Based Systems, 2013, No 39, pp. 194-206.
20. L u, C. N., H. T. W u, S. V e m u r i. Neural Network Based Short Term Load Forecasting. – IEEE Trans. Power Syst., Vol. **8**, 1993, No 1, pp. 336-342.
21. H i p p e r t, H. S., C. E. P e d r e i r a, R. C. S o u z a. Neural Networks for Short-Term Load Forecasting: A Review and Evaluation. – IEEE Trans. Power Syst., Vol. **16**, 2001, No 1, pp. 44-55.
22. P a i, P.-F., W.-C. H o n g. Forecasting Regional Electricity Load Based on Recurrent Support Vector Machines with Genetic Algorithms. – Electric Power Syst. Res., 2005, No 74, pp. 417-425.
23. P a i, P. F., W. C. H o n g. Support Vector Machines with Simulated Annealing Algorithms in Electricity Load Forecasting. – Energy Convers Manage, Vol. **46**, 2005, No 17, pp. 2669-2688.
24. A l f a r e s, H. K., M. N a z e e r u d d i n. Electric Load Forecasting: Literature Survey and Classification of Methods. – Int. J. Syst. Sci., Vol. **33**, 2002, No 1, pp. 23-34.
25. M o g h r a m, I., S. R a h m a n. Analysis and Evaluation of Five Short-Term Load Forecasting Techniques. – IEEE Trans. Power Syst., Vol. **4**, 1989, No 4, pp. 1484-1491.

26. A m j a d y, N. Short-Term Bus Load Forecasting of Power Systems by a New Hybrid Method. – IEEE Trans. Power Syst., Vol. **22**, 2007, No 1, pp. 333-341.
27. H u a n g, G. B., Q. Y. Z h u, C. K. S i e w. Extreme Learning Machine: A New Learning Scheme of Feedforward Neural Networks. – 2004 IEEE International Joint Conference on Neural Networks, Budapest, Hungary, IEEE, 2004, pp. 985-990.
28. M a o, L i., Y. W a n g, X. L i u et al. Short-Term Power Load Forecasting Method Based on Improved Extreme Learning Machine. – Power System Protection and Control, Vol. **40**, 2012, No 20, pp. 140-144 (in Chinese).
29. Z h a n g, S., J. L i u, B. Z h a o et al. Cloud Computing-Based Analysis on Residential Electricity Consumption Behavior. – Power System Technology, Vol. **37**, 2013, No 6, pp. 1542-1546 (in Chinese).
30. S o n g, Y., G. Z h o u, Y. Z h u. Present Status and Challenges of Big Data Processing in Smart Grid. – Power System Technology, Vol. **37**, 2013, No 4, pp. 927-935 (in Chinese).
31. L i, J., H. W e i. A Multi-Agent System to Draw up Energy-Saving Power Generation Dispatching Schedule. – Power System Technology, Vol. **35**, 2011, No 5, pp. 90-96 (in Chinese).
32. W e n, H., H. Z h a n g. Management System for Energy-Requirements Standard Understanding and Application Analyze. – Information Technology&Standardization, 2011, No 37, pp. 65-67 (in Chinese).
33. L u, H., B. Z h u, D. L i. More Energy Consumption Characteristics Analysis and Modeling of Forging Production System Based on Rough Sets Theory. – Machine Design and Manufacturing Engineering, 2013, No 12, pp. 20-23 (in Chinese).
34. T r i b b l e, A. N. The Relationship between Weather Variables and Electricity Demand to Improve Short-Term Load Forecasting. Norman, Oklahoma, University of Oklahoma, 2003.
35. H e r r m a n n, C., S. T h i e d e. Process Chain Simulation to Foster Energy Efficiency in Manufacturing. – CIRP Journal of Manufacturing Science and Technology, Vol. **1**, 2009, No 4, pp. 221-229.
36. T a o, L., L. H u, R. C h e n, A. C h e n. Establishment and Simulation of the RBF Model of the Unit Process Energy Consumption of an Automobile Enterprise. – In: International Workshop on Intelligent Systems and Applications, 2009, ISA 2009, 2009, No 1, pp. 23-24.
37. A z a d e h, M. S. A., S. F. G h a d e r i, S. M. A s a d z a d e h. An Integrated DEA PCA Numerical Taxonomy Approach for Energy Efficiency Assessment and Consumption Optimization in Energy Intensive Manufacturing Sectors. – Energy Policy, Vol. **35**, 2007, No 7, pp. 3792-3806.
38. Z e n g, F., S. M i a o, Y. S h a n g et al. Parallel Implementation of Node-Impedance Matrix Generation Algorithm Based on Decomposition and Coordination. – Power System Technology, Vol. **35**, 2011, No 9, pp. 99-104 (in Chinese).
39. C h e n, B., M. W. C h a n g. Load Forecasting Using Support Vector Machines: A Study on EUNITE Competition 2001. – IEEE Transactions on Power Systems, Vol. **19**, 2004, No 4, pp. 1821-1830.