Yuancheng Li, Bei Wang, Yiliang Wang, Xianji Wang

State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources, North China Electric Power University

A Dynamic State Estimation Method Based on Mixed Measurements for Power System

***Abstract**. Dynamic State Estimation (DSE) techniques have the ability to foresee potential contingencies and security risks. Any improvement in its ability to estimate would definitely go a long way in reducing the security risks in the modern power system. One important factor affecting the quality of estimation is the measurement accuracy. Phasor Measurement Unit (PMU) has revolutionized the way state estimation is performed. The unique ability to measure the voltage and current phasors (magnitude and phase angle) with very high accuracy makes PMU extremely useful in modern Energy Management Systems (EMS). Due to the high price, technology level and communication capacity, the PMU can't be equipped in all buses in the system nowadays. Therefore, this paper brings forward an improved method on dynamic state estimation that combines some buses measurements from PMU with measurements from SCADA. As Relevance Vector Machine (RVM) has a better performance on the regression, the state estimation algorithm is based on RVM in this article. Since the input data dimension is too large, pre-processing of data is needed. Autoencoder Network (Autoencoder) can be used for data dimensionality reduction. So this paper uses Autoencoder to reduce the data dimensionality, and then uses RVM to estimate the state of power system.*

***Streszczenie.** W artykule przedstawiono metodę estymacji stanów dynamicznych w sieci elektroenergetycznej, wykorzystujący pomiary fazy i amplitudy napięcia i prądu oraz pomiary SCADA z poszczególnych punktów sieci. W algorytmie wykorzystano maszynę wektorów RVM. Ze względu na zbyt duży wymiar danych wejściowych, zastosowano pre-processing z wykorzystaniem sieci neuronowej auto-encoderowej. (**Estymacja stanów dynamicznych w sieci elektroenergetycznej**).*

Keywords: State Estimation, Power System, Relevance Vector Machine, Autoencoder.
Słowa kluczowe: estymacja stanu, system elektroenergetyczny, RVM, auto-encoder.

1. Introduction

EMS plays an important role in monitoring and controlling the power system and state estimation forms its backbone. With increased operations of power system to its limits and the tendency to utilize the grid to its full potential these days, monitoring and controlling with high level accuracy becomes a trend. Because of the unique ability to predict the state vector one time stamp ahead and the advanced complexity of operating large interconnected networks, electric system uses modern Energy Management System(EMS). The purpose of an EMS is to monitor, control, and optimize the transmission and generation facilities with advanced computer technologies. The aim of the state estimation function is to obtain the best estimate of the current system state [1]. The state estimation is also called Filter. It ensures the secure and economical operation of the power system. The state estimation uses the redundancy of the real-time measurement system to improve the data precision, gets rid of the false information due to the random interferer and estimates the state of power system. The state estimation includes the static state estimation and the dynamic state estimation [2]. The static state estimation is using the measurement data to estimate the power system state at the same time. And the dynamic state estimation is using the measurement data at last time and measurement data now to estimate systems state [3-4].

The dynamic state estimation not only has all the strongpoints of the static state estimation, but also can distribute the power energy economically, forecast the safety of system state, and control the system preventively [5-8]. At present, many state estimation methods in power system are based on the static estimation and the measurement data of these methods come from SCADA/PMU system. Because these data are transmitted to control center once 2 seconds through RTU, the error of RTU and every tache in the transmission make the errors of measurement data bigger, the accuracy of voltage phasor accepted by iterating becomes very imprecise [9-10]. Currently, with the development of GPS and its application in the power system, the time transferring function of GPS is used widely in the power system. The PMU based on GPS starts to be equipped in the power system step by step. Because the measured data from PMU are carried to control center more quickly than the measured data from SCADA, and the PMU can measure the voltage's angle, the precision of measured data from PMU is better than that from SCADA. Moreover, the PMU can measure the branch current phasor and the bus's voltage phasor. If these data are combined in the state estimation method, the precision of state estimation will be improved. The reference [11]-[13] bring forward some methods that combine the mixed measurements in state estimation after some PMU, which are placed in buses and improve the precision in state estimation. However, these methods are all based on measurements at the same time and thus belong to static state estimation.

Compared with the traditional method, this new method enriches the data-base infinitely, describes the track of system operation more exactly and provides a mass of analyzed data to system for the next analysis of system operation, accident preventing, evaluating and controlling online. This new method predigests the Jacobian matrix greatly, shortens the computing time, improve the precision and astringency on state estimation when the precision of measurements is higher. When the precision of measurements is not high, although the new method that increases the known measurements of system makes the Jacobian matrix more complex, it increases the precision of state estimation value. With the fast development of the power system and the improvement of the computer operation speed, there will be a urgent need for wide and real time measurements in future power system, then the state estimation in millisecond level. Compared with traditional methods, this new method meets the demand of the future power system more developed in computer speed and result precision. As methods to estimate the state in power system, forecasting state data in power system and iterative estimation of state are merged into this state estimator. Hence, in this paper RVM is introduced into state estimation of power system and the experiment results show that this new method can get satisfactory effect. Relevance Vector Machine (RVM) [14] is a new kind of regression method in recent years, which is based on relevance vector algorithms. Compared with the method of

Kalman Filter, Genetic Algorithms, Neural Network Algorithms and SVM, RVM features can make estimation accurate and control the sparseness of the decision-making function in learning machine directly. However, there still remains some problems: 1. the high dimension of the input data required by the state estimation algorithm from SCADA/PMU system; 2. the high time cost of RVM to get satisfactory performance due to RVM using the greedy algorithm to approximate given function by searching a linear combination of basis functions choosing from a redundant basis function dictionary. In this paper, dimensionality reduction is added into state estimator as a preprocessing module to deal with these problems. Autoencoder [15-16] is a kind of network algorithm proposed by G. E. Hinton and R. R. Salakhutdinov, which maps the high dimension features in the input space to new lower dimension coordinates. This pretreatment not only keeps the original topological structure of the original data but also deals with the nonlinear relationship among the input data well. Then RVM is used to train the data after the Autoencoder pretreatment, create training models and forecast the systemic state variables such as voltage magnitude and phase angle value of every node. Autoencoder-RVM is a kind of state estimation method which does not depend on the dimensionality of the problem.

2. Formulation of state estimation

Mathematically, the information model used in power system state estimation uses the equation (1):

(1) $$z = h(x) + e$$

Where: $z:(m\times 1)$ measurement vector, $x:(n\times 1)$ true state vector, $h(x):(m\times 1)$ state equation vector, $e:(m\times 1)$ measurement error vector, m : the number of measurements, n : the number of state variables.

The usual state variables are the voltage magnitude and angle, while the measurements are the real and reactive power flows, node injections and voltage magnitudes. The objective function of the state estimation is the same as that of conventional state estimation as follows:

(2) $$\min J(x) = \sum_{i=1}^{m} w_i (z_i - h_i(x))^2$$

where, w_i : a weighting factor of measurement variable z_i .

Commonly, a criterion that is used in state estimation formulation is to minimize the sum of the differences between the estimated and true values. This approach is called the weighted least squares(WLS) estimation.

3. The proposed state estimation model

One way of handling high-dimensional data is dimensionality reduction. By learning algorithm, we get the status of hidden-variable model. We call this implicit variable essential dimension, and essential dimension is often much smaller than the original data. Autoencoder uses adaptive, multi-layer network coding high-dimensional data into low-nesting. By the middle of training with multiple layers of bi-directional deep neural networks to high-dimensional data into low-dimensional nesting and use similar decoding network from low-dimensional reconstruction of nested high-dimensional data. Since the Autoencoder gives bi-directional mapping between inputted encoding high-dimensional data and low nested data, it overcomes the problem of inverse mapping that most of the non-linear dimensionality reduction methods have.

3.1. The structure of Autoencoder

We should know that $R=\{r_1,\cdots,r_n\}\in\Re^{D\times n}$ is a high dimensional dataset with n vectors where r_i is the i -th data with D dimensions. $M=\{m_1,\cdots,\}\in\Re^{d\times n}$ subject to $d \ll D$, is a kind of low nested data. Through Autoencoder, low nested data M among the high-dimensional data R can be found. The basic procedure is shown in Figure 1. The whole system contains encoding network and decoding network. Encoding network is that dimensionality reduction network is used to transform the high dimensionality data to low nested data. Decoding network is that reconstruction process can be seen as the inverse process of the encoding network and can transform the low nested data to high dimensionality data. Pretraining consists of learning a stack of restricted Boltzmann machines (RBMs), each with only one layer of feature detectors. The learned feature activations of one RBM are used as the "data" for training the next RBM in the stack. After the pretraining, the RBMs are "unrolled" to create a deep autoencoder, which is then fine-tuned using backpropagation of error derivatives.

An ensemble of vectors can be modeled using a two-layer network called a "Restricted Boltzmann machine"(RBM)[14, 15]. The structure of RBM contains a visual layer , a hidden layer and the connection between a visual layer and a hidden layer. A joint configuration (v, h)of the visible and hidden units has an energy [16] given by

(3) $$E(\mathrm{v},\mathrm{h}) = -\sum_{i\in visible} b_i v_i - -\sum_{i\in hidden} b_i h_i - -\sum_{i,j} v_i h_j w_{ij}$$

where v_i and h_i are the binary states of visual unit i and hidden unit j , b_i and b_j are their biases, and w_{ij} is the weight between them.

The states of the hidden units are then updated once more so that they represent features of the confabulation. The change in a weight is given by

(4) $$\Delta w_{ij} = \varepsilon(\langle v_i h_j\rangle_{data} - \langle v_i h_j\rangle_{recon})$$

where ε is a learning rate, $\langle v_i h_j\rangle_{data}$ is the mean of the original w training data and $\langle v_i h_j\rangle_{recon}$ is the mean of the reconstruction data.

3.2. The structure of RVM

Relevance vector machine is a sparse probability model based on support vector machine proposed by Michael E Tipping in 2001. Its training is carried on under Bayesian framework, so we can get the distribution of predicted values by regression estimate with RVM. Compared with SVM, RVM has the following advantages:

(1) With RVM, we can get probability forecasts;

(2) In the inference process, setting error parameter subjectively can be avoided;

(3) The relevant vectors used for training are less than the SVM;

(4) Kernel function has a greater range of choices for it need not satisfy the Mercer conditions.

The biggest difference between RVM and SVM is that RVM turns subjective division into objective division under probability, which makes classification function reach likelihood function maximum for the training set.

The output of RVM model is as follows:

(5) $$y(x) = \sum_{j=1}^{m} \omega_j \phi_j(x) + \omega_0$$

where $\phi_j(x)$ is a non-linear kernel function, ω_j is model weights. The kernel function used in SVM must satisfy Mercer theorem but RVM doesn’t have the limitation.

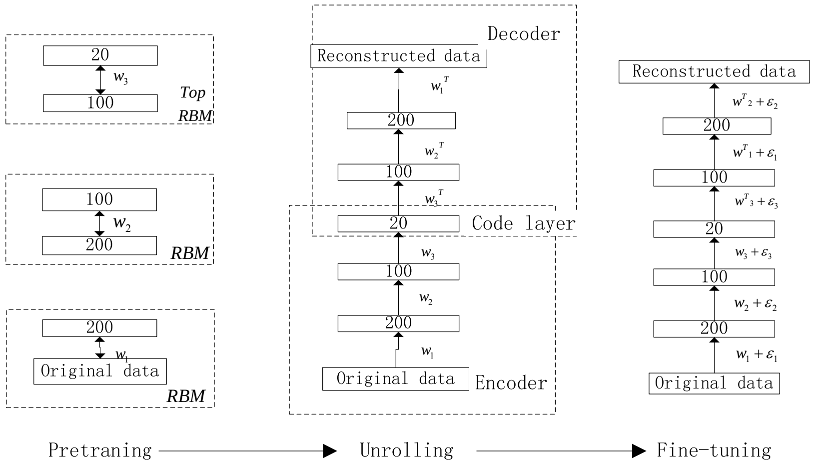

Figure 1. The structure of the Autoencoder network.

After defining the model (5) basis functions, we can train the model weights ω_j with maximum likelihood method under Bayesian framework, which may avoid learning problems and improve model generalization ability. Therefore, RVM defines the priori probability distribution for each model weight:

(6) $$p(\omega_j \mid \alpha_j) = \left[\frac{\alpha_j}{2\pi}\right]^{1/2} \exp\left[-\frac{1}{2}\alpha_j\omega_j^2\right]$$

Where ω_j is hyper-parameter of the priori distribution of model weight α_j.

For a given set of training samples $\{x_i, t_i\}_{i=1}^N$, we can assume that the target value t_i is independent, and the noise of inputted data obey Gaussian distribution of which the variance is σ^2. In this way, likelihood function of the given training sample set is as follows:

(7) $$p(t \mid \omega_j, \sigma^2) = (2\pi\sigma^2)^{-N/2} \exp\left[-\frac{1}{2\sigma^2}\alpha_j \left\|t - \Phi\omega\right\|^2\right]$$

Where $t = (t_1, t_2, \cdots, t_n)^T$, $\omega = (\omega_1, \omega_2, \cdots, \omega_n)^T$, Φ is matrix of which the rows include the response of all kernel functions to input x_i:

(8) $$(\Phi)_i = \left[1, \phi_1(x_i), \phi_2(x_i), \cdots, \phi_n(x_i),\right]$$

Based on priori probability distribution and likelihood distribution, we calculate the posterior probability distribution of model weights with Bayesian method. The formula can be written as:

(9) $$p(\omega \mid t, \alpha, \sigma^2) = \frac{p(t \mid \omega, \sigma^2)\, p(\omega \mid \alpha)}{p(t \mid \alpha, \sigma^2)}$$

The posterior distribution of model weight is multivariate Gaussian distribution, that is:

(10) $$p(\omega \mid t, \alpha, \sigma^2) = N(\mu, \Sigma)$$

Where $\Sigma = (\sigma^{-2}\Phi^T\Phi + A)^{-1}$ is covariance, A is diagonal matrix of $(\alpha_0, \alpha_1, \cdots, \alpha_n)$ and $\mu = \sigma^{-2}\sum \Phi^T t$ is mean value. The likelihood distribution of training target value can realize marginalization by integration.

(11) $$p(\omega \mid t, \alpha, \sigma^2) = p\int (t \mid \omega, \sigma^2)\, p(\omega \mid \alpha)\, d\omega$$

In this way, we can get marginal likelihood distribution of the hyper-parameters:

(12) $$p(t \mid \alpha, \sigma^2) = N(0, C)$$

Here covariance $C = \sigma^2 I + \Phi A^{-1} \Phi^T$.

Finally, the estimated value of model weights in RVM method is given by the mean value of posterior distribution and it is a maximum posteriori (MAP) estimation. The MAP estimation of model weight depends on hyper- parameters a and noise variance σ^2 and its estimated value $\overline{\alpha}$ and $\overline{\sigma}^2$ can be obtained by maximizing the marginal likelihood distribution. The uncertainty of model weight optimal value reflected by posterior distribution may show the uncertainty of model predictions. For the given input value x^*, the corresponding probability distribution of the output is as follows:

(13) $$p(t^* \mid t^*, \overline{\alpha}, \overline{\sigma}^2) = p\int (t^* \mid x^*, \omega, \overline{\sigma}^2)\, p(\omega \mid t, \overline{\alpha}, \overline{\sigma}^2)$$

The formula (13) obeys the form of Gaussian distribution, that is:

(14) $$p(t^* \mid t^*, \overline{\alpha}, \overline{\sigma}^2) = N(y^*, \sigma^{*2})$$

Where the predicted mean value $y^* = \mu^T \Phi(x^*)$ and the variance (uncertainty) $\sigma^{*2} = \overline{\sigma}^2 + \Phi^T(x^*) \sum \Phi(x^*)$.

RVM solves the problem of parameters selection with significance under Bayesian framework which has wide applicability. Using RVM for regression prediction, we can obtain better predicted value and variance range.

The whole modeling could be concluded in next steps:

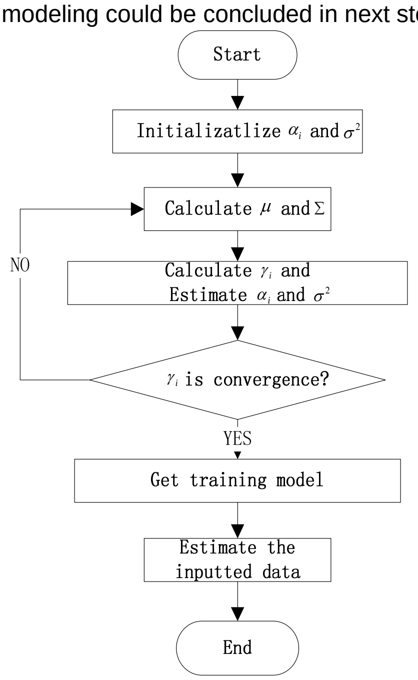

Figure 2. The structure of the RVM

The reason of combining forecasting systemic state variables in power systems with iterative estimation of state estimation is a main kind of method to solve the problem of state estimation. As a result, the problem of state estimation in power system can be regarded as a regression problem and RVM gets the approximation functions to estimate these systemic state variables.

The predictive process of RVM can be shown simply as follows.

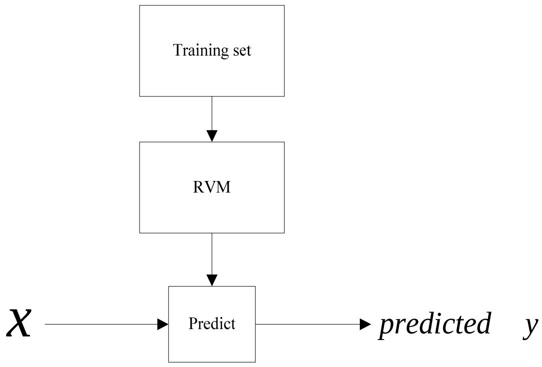

Figure 3. Simple predictive process of RVM

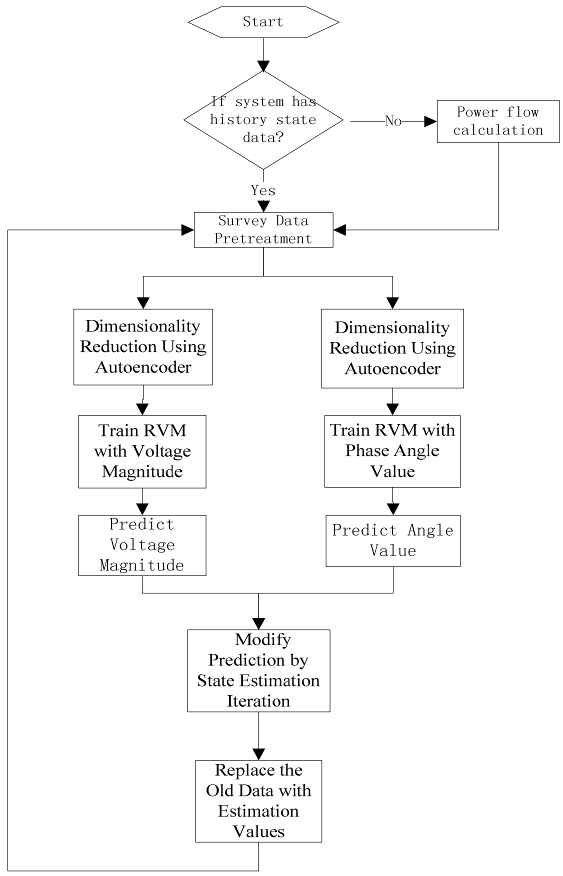

Figure 4. Framework based on Autoencoder-RVM

3.3. The proposed method using Autoencoder-RVM

Some surveyed data which always consist of power, voltage and current of every node and branch in the power network from SCADA/PMU system need to be sampled and standardized into state estimator. And the scale of the data become more and more complex, scale larger and dimensions higher. There are two problems remaining. Firstly, the time cost of training RVM to build the predictive model will become uncontrollable. Secondly, the precision and generalization ability will be influenced by the nonlinear character of the surveyed data at the same time.

For above reasons, a pretreatment for the input surveyed data set has to be merged into the process of state estimation based on RVM. Some traditional methods such as PCA is a good treatment for dimensionality reduction, but relations of input data from SCADA/PMU system are nonlinear and the original topological structure of the original data cannot be kept. Autoencoder, which is a non-linear dimensionality reduction method base on network, which can keep the original topological structure of the original data and deal with the nonlinear relationship among the input data well, maping the high dimension features in the input space to a new lower dimension space. This new method of state estimation based on Autoencoder and RVM can perform better in the control of sparseness and the predictive speed.

The new procedure in state estimation proposed in this paper is shown as Figure 4.

4. Case Studies

In this paper, a proposed system is applied to IEEE14 bus system to validate the new method's performance. In the simulation case, system power flow calculation is used as the true result. As the measurement data in the system are susceptible to interference with white Gaussian noise, the data is added with the corresponding normally distributed random measurement error using as the simulated data. In most power system, PMUs are installed in the important power plants and important transformer substation. In this simulation case, PMUs are fixed in the transformer substation bus 4, 5 and power plant bus 1, 2, 3, 6, 8, and other nodes bus are installed with SCADA. The state estimation program fits for the power system state variables by minimizing these errors. Ideally state estimation should run at the scanning rate of the telemetry system (at every two seconds). Due to computational limitations, most practical state estimators run every few minutes or when major changes occur.

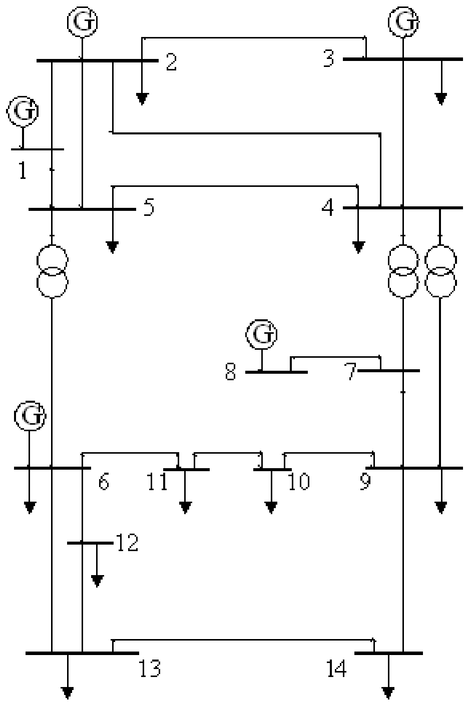

Figure 5. IEEE14 bus system

4.1 Analysis of experimental data

In this study, the number of analysed node is 6. In order to improve the prediction accuracy, there are dimensions that are selected. Through Autoencoder, the dimensions are reduced to 60 dimensions. In the Autoencoder reduction process, 60 samples are used for training and 40 samples are for testing. The Figure 6 shows the errors of the reconstruction.

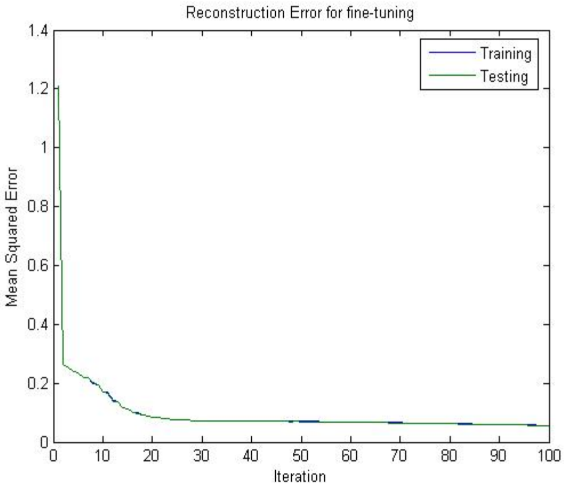

Figure 6. the errors of reducing the data dimensions.

After 20 iterations, the average mean square error has become less than 0.1, and dimensionality reduction effect is ideal. After the data's dimension is reduced, RVM is used to estimate the system state. The Figure 7 is the contrast of estimated value and true value. The voltage in power system is transferred to p.u. voltage. Then it is normalized to [0,1] in our algorithm which is the Y axis in the figure. The Time axis represents the every two second the dynamic state estimation runs.

Through figure 6, the Autoencoder fast reduce the dimensionality of data and RVM is applied to predict state value of power system. The output value of Autoencoder is between [0,1] and then is used as input of Rvm. The Autoencoder is advanced and non-linear method for reducting dimensionality and principle information relating data is preserved by the method.

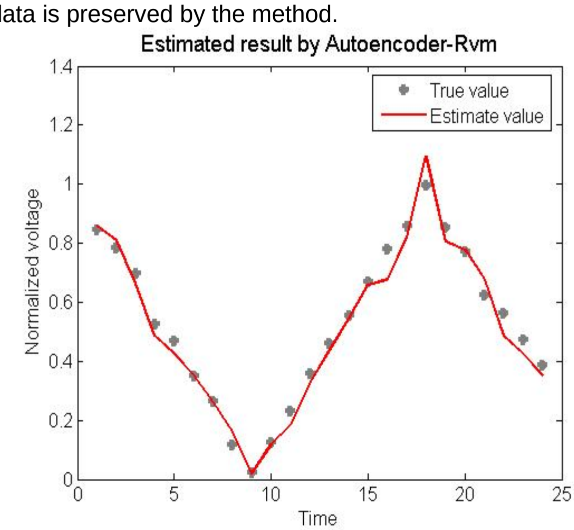

Figure 7. Estimated result of Autoencoder-Rvm for voltage

From the Figure 7, we can conclude that estimated value can better fit for the true value. In order to make a comparison, RVM is used as the state estimator. Estimated result is described in Figure 8.

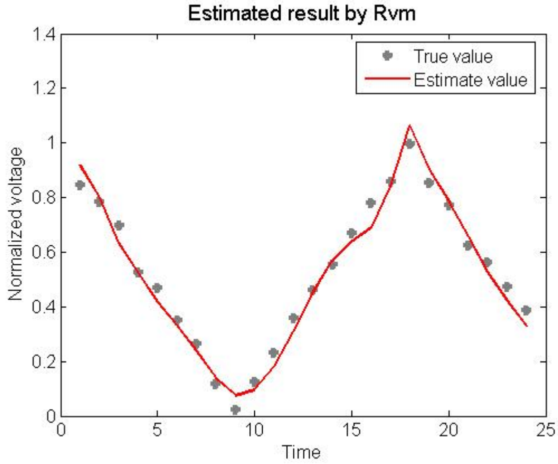

Figure 8. Estimated result of Rvm for voltage

4.2 Contrast of results

The results of the experiment show that the predictive error of Autoencoder-RVM is almost the same compared with RVM. The prediction performs better at smooth curve than the curve in conditions when abrupt changes occur. At the same time, prediction of angle doesn't perform as well as prediction of voltage since angle always changes more greatly than voltage in power system.

Tab.1 shows the results of prediction performance and computing time for training and testing data by SVM using all features.

Table 1. Comparison of experiment results

State estimator	Error (RMS)	Trained Time(ms)	Tested time(ms)	Used time(ms)
Rvm	0.046	860	82	942
Autoencoder-Rvm	0.051	660	76	736

Two conclusions are gained as follows:

1. The training time and testing time are decreased by using Autoencoder-Rvm because Rvm algorithm can implement simply without iterations.
2. The Autoencoder-Rvm has better performance in sparse solutions and does not decrease the real time performance.

Sample set is rolled by time and dimensionality is reduced during every prediction and total time cost will decrease to some extent. But the precision performance of prediction is almost the same. Considering the real-time performance of dynamic state estimation and computing capability of modern computers, the time precision performance of prediction is acceptable.

5. Conclude

Autoencoder for dimensionality reduction and RVM for regression are introduced successfully into the state estimator and a satisfactory performance is obtained. From the theoretical analysis and the experimental results, we can see that the pre-treatment of Autoencoder could efficiently deal with the nonlinear feature of initial data and reduce their dimensionality and the following training time cost of RVM. Autoencoder - RVM has powerful learning ability, good generalization ability, sparse solution and low dependency on sample data. The state changes of a power system can be tracked quickly and precisely. This indicate that predictive control based on Autoencoder - RVM has potential applications in realizing nonlinear control. With more and more new Machine Learning algorithms being introduced into the field of state estimation, the dynamic state estimation in power system will have a new complexion.

REFERENCES

[1] Yu Erkeng, Power system state estimation, Beijing: Hydraulic and Electric Power Press, 1985
[2] Tripathy S. C., Chohan S., On line tracking state estimation in power system, Tournal of the Institution of Engineers India: Electrical Engineering Division, 74(1994), 162-164
[3] Tong-Koo Park, The Concept and Design of Dynamic State Estimation, in Proceedings of the American Control Conference San Diego, Cali-forma, Tune, 1999
[4] Al-Othman A. K., Irving M.R., Uncertainty modelling in power system state estimation, IEE Proceedings: Generation, Transmission and Distribution, 152(2005), 233-239
[5] Sinha A.K., Mondal T.K., Dynamic State Estimation Using ANN Based Bus load Pre-diction, IEEE Transactions on Power Systems, 14(1999), 1219-1225
[6] Amjady, Nima, Short-term bus load forecasting of power systems by a new hybrid method, IEEE Transactions on Power Systems, 22(2007), 333-334
[7] Holbert, Keith E., Mishra, Intrusion detection through SCADA systems using fuzzy logic-based state estimation methods, International Tournal of Critical Infrastructures, 3(2007), 58-87
[8] Knmar, Ashwani, Robust dynamic state estimation of power harmonics, International Energy Systems, 28(2006), 65-74
[9] Watson, N.R., An adaptive Kahnan filter for dynamic harmonic state estimation and harmonic injection tracking, IEEE Transactions on Power Delivery, 20(2005), 1577-1584
[10] Thakur S.S., Sinha A, K., A robust dynamic state estimator for electric power systems, Toumal of the Institution of Engineers (India): Electrical Engineering Division, 84(2003), 42-46,.
[11] Huang S.T., Shih K.R., Dynamic state estimation scheme including nonlinear measurement-function considerations, IEE Proceedings: Generation, Transmission and Distribution, 149(2002), 673-678
[12] Lee Tae-Won, Lee S., New data fusion method and its application to state estimation of nonlinear.namic systems, Proceedings-IEEE International Conference on Robotics and Automation, 4(2000), 3525-3530
[13] Sinha A.K., Mandal T.K., Hierarchical dynamic state estimator using ANN-based dynamic load prediction IEE Proceedings: Generation, Transmission and Distribution, 146(1999), 541-549
[14] Smolensky P., Parallel Distributed Processing:Foundations, D. E. Rumelhart J., McClelland L., MITPress, Cambridge, 1(1986), 194–281
[15] Hinton G. E., Neural Comput. 14(2002)
[16] Hopfield J. J., Proc. Natl. Acad. Sci. U.S.A. 79(1982), 2554
[17] Nga-Viet Nguyen, Shin V., Shevlyakov G., Power system state estimation with fusion method Computer and Automation Engineering (ICCAE), 2010 The 2nd International Conference on Volume: 5 (2010) 71 – 76
[18] Bian X., Li X.R., Chen H., Gan D., Qiu J., Joint Estimation of State and Parameter With Synchrophasors—Part I: State Tracking Power Systems, (2011), 1 – 13
[19] Okon T., Wilkosz K., Consideration of different operation modes of UPFC in power system state estimation Environment and Electrical Engineering (EEEIC), (2011) 1 – 4
[20] Gao W., Wang S., On-line dynamic state estimation of power systems, North American Power Symposium (NAPS), 2010, 1 – 6
[21] Tang P., Jiang R., Zhao M.m, Feature Selection and Design of Intrusion Detection System Based on k-Means and Triangle Area Support Vector Machine,Future Networks, 2010. ICFN '10. Second International Conference on (2010) 144 – 148
[22] Rodriguez-Martinez E., Goulermas J.Y., Mu T., Ralph J.F., Automatic Induction of Projection Pursuit Indices, Neural Networks, IEEE Transactions on (2010) 1281 – 1295

Authors: prof. Yuancheng Li, Institute of Smart Grid and Information Security, North China Electric Power University, 2 Beinong road, Huilongguan Town, Changping Distict, Beijing, China, E-mail: dflyc@163.com; Ms. Bei Wang, Institute of Smart Grid and Information Security, North China Electric Power University, E-mail: tinybaby007@163.com; Mr. Yiliang Wang, Institute of Smart Grid and Information Security, North China Electric Power University, E-mail: wangyiliang206@163.com; Mr. Xianji Wang, Institute of Smart Grid and Information Security, North China Electric Power University, E-mail: book987@163.com;

Corresponding author:
Mr. Yiliang Wang, Institute of Smart Grid and Information Security, North China Electric Power University, E-mail: wangyiliang206@163.com;