



---

# VOLTAGE STABILITY ASSESSMENT OF COMPLEX POWER SYSTEM BASED ON GA-SVM

Qiang Li\*, Xiao-feng Liu†

---

**Abstract:** The dynamic stability assessment and prediction of a complex power system is a precondition to take the action of protecting control. This paper presents the four support vector machines (SVMs) with an improved genetic algorithm (GA) to compute their parameters automatically, that one SVM is used to simulate the tangent vector and the others for identifying the instable area. Besides, the GA was initialized by Meta-Learning method to enhance the performance and its optimal solution was selected by last test. Furthermore, a large network simplification was taken for reducing the amount of calculation and responding in real time. Study with the IEEE 118-bus test system indicated that the system status of a complex power system subjected a fault could be predicted based on this technique of the GA-SVM for simulating the tangent vector accurately. Besides, three binary SVM classifiers were trained to locate the instable area, and ranking the levels by the analysis of critical bus is help to management. Based on the test on the networks, the suggested approach can predict accurately with 98.87 % success rate and identify the fault area with 94.91 % success rate.

Key words: *genetic algorithm, power flow, power system simulation, SVM, voltage stability assessment, meta-learning*

Received: April 26, 2018

DOI: 10.14311/NNW.2019.29.027

Revised and accepted: December 30, 2019

## 1. Introduction

The importance of voltage security assessment of power system has been paid much more attention for a long time because a large number of stability accidents occurred in some countries. There is no doubt that a disturbance can have a significant impact on system stability [1].

The measurement of monitoring system is very indispensable before taking action for deciding system status. Phasor measurement units (PMUs) are a key feature tool [2].

---

\*Qiang Li; College of Information, Shan Xi Finance & Taxation College, Taiyuan City, 030024, P. R. China

†Xiao-feng Liu – corresponding author; College of Data Science, Tai Yuan University of Technology, Taiyuan City, 030024, P. R. China, E-mail: [122824940@qq.com](mailto:122824940@qq.com) .v

Many assessment approaches of post-disturbance have been developed in the literature. The nonlinear differential equations based on time-domain simulation are the most straightforward method to indicate the system stability status [3]. However, this method needs a large quantity of network information and consumes a lot of time.

In practice, the simplest way of assessing system status is to monitor performance indices. The one of useful measures to estimate system voltage status is voltage collapse proximity indicator (VCPI) [4]. Where, the power transfer stability index (PTSI) is one better method to determine system stability, which is regarded as dynamic voltage collapse indicator [5]. A new better method used as voltage stability index, called “the components of tangent vector of the power flow jacobian”, has been present in the paper [6]. It is very similar to the  $L$  index for indicating voltage collapse. In addition, it can provide a more effective and more information at points closer to voltage stability limit. To compare with the  $L$  index, it has a less computation and wider range of application.

Recently, many artificial intelligence techniques for solving system stability assessment and safe control problems have well performance in application. Decision trees (DT) technique has been successfully used for security assessment in reference [7]. This method makes decision very quickly, while it utilizes a large quantity of space for constructing trees structure. It was reported that some researchers have utilized fuzzy logic technique for improving robustness [8].

Artificial neural networks (ANNs) have been proposed for the voltage stability assessment, such as BP neural network for voltage assessment [9]. Nevertheless, this requires plenty of training time and space, and easily falls into local minima. Owing to ANNs’ drawback, a better prediction tool is necessary to develop.

Support vector machine (SVM) is a very promising technique for assessing system stability [10, 11]. It has been demonstrated that SVM superior over ANN in terms of accuracy and speed in those references. In the application, it is necessary to set parameters of SVM carefully, which greatly influence the performance of SVM. In the method, consuming time for complex power system is still inevitable.

In order to obtain the important parameters, two classic algorithms were proposed, grid-search strategy and gradient-based methods. The strategy based grid-search is a simple method, but its precision and computational time depends on the length of step in the search space. Although gradient-based algorithm is faster than grid search method, its effect relies on the smooth degree in the search space, and it easily falls into local minima. A method based on the theory of “Quasi-Monte Carlo” was proposed in paper [12] in order to reduce the search scope, called “MOUD”.

The paper [13] demonstrated that the performance of GA and PSO are better than that of grid search strategy and gradient-based method by the experiment. Besides, they adopted an object function to assess the optimization. At same time, it pointed that the initial seed of PSO can influence the performance.

For obtaining the better parameters of SVM, Ana gave the three different initial parameters solutions of GA, and compared them by performances in the experiment [14]. Zhang et al. presented a hybrid method for determining the parameters of SVM that it obtains the parameters by BBDE and set the search interval by the cluster distance in the feature space [15].

Although the GA in the experiment can get a better solution than other algorithms, it is still time-consuming in the large search space of SVM, and often has different results due to different initial values.

Miranda et al. used meta-model to select the parameters for reducing the computation, but this method needs a large number of data training and test beforehand in order to establish a priori model [16, 17]. However, we do not often have a lot of experience in machine learning practically so that the application of this method is limited. Furthermore, this method gives us an idea that as far % as possible the use of prior knowledge is helpful to improve the learning efficiency.

As we known, contradictory usually exists in multi-objective optimization. Chatelain et al. proposed a technique based “Pareto control concept” to achieve the goal of FP and TP balance by EMOO [18]. NSGA II in the solution was used for obtaining Pareto front, but it depends on non-dominated values set as the initial population.

A technique for solving big data computing is resolved by network simplification and distribution computation on a cluster system.

In this paper, a new method utilizes SVM to simulate the tangent vector of power flow jacobian for predicting the system status and takes network simplification for reducing the amount of calculation.

An improved genetic algorithm is suggested for obtaining the optimized combination parameters of SVM. While the Meta-Learning is used to accelerate the convergence of the GA and the configuration of the GA-SVM is decided by last test.

The structure of this paper is as such that, Section 2 gives a brief description of tangent vector of PFJ; Mathematical model of SVM and the genetic algorithm are presented in Section 3; Experiment method in Section 4 includes model, tools and data so on; Results and analysis are described in Section 5; Finally, conclusions are summarized in Section 6.

## 2. Tangent vector of the power flow jacobian

The  $L$  index is one most popular voltage stability indicator, which is an easy method to compute system status in practice.

$L$  index method is one of voltage collapse proximity indicators, and it is used to estimate the status of a system. When an exception occurs to a node, the value of  $L$  could be more than 1.

$$L_i = \left| 1 + \frac{\bar{V}_{0i}}{\bar{V}_{2i}} \right| = \left| \frac{\bar{S}_i}{V_i^2 Y_{ii}} \right| = \left| \frac{S_i}{V_i^2 Y_{ii}} \right|, \quad (1)$$

where  $\bar{S}_i$ ,  $Y_{ii}$  and  $V_i$  are the transformed power transformed admittance and node voltage respectively.

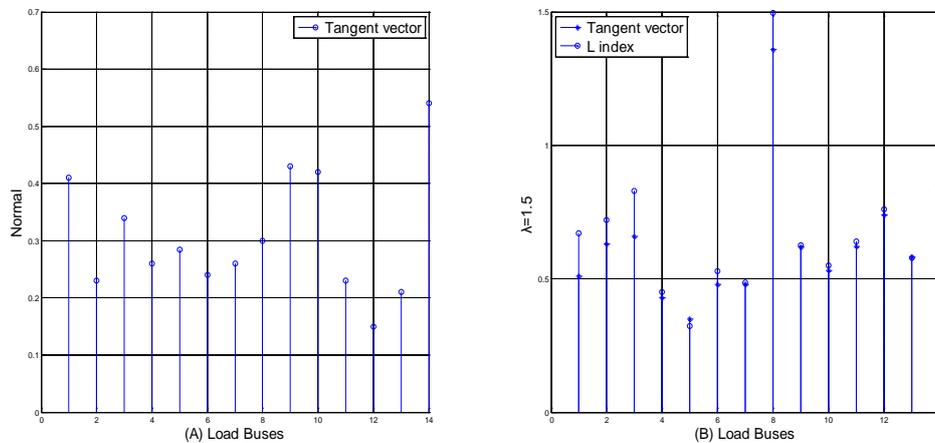
It has been described in reference [6] that  $L$  index cannot express system instability in the absolute way actually and be used for assessing radial distribution systems.

The following equation of power flow computation describes the typical variations of the  $i$ -th node.

$$\begin{aligned}
 P_{Gi} - P_{Li} - P_{Ti} &= 0, & P_{Ti} &= \sum_{j=1}^n V_i V_j [G_{ij} \cos \theta_{ij} + B_{ij} \sin \theta_{ij}], \\
 Q_{Gi} - Q_{Li} - Q_{Ti} &= 0, & Q_{Ti} &= \sum_{j=1}^n V_i V_j [G_{ij} \sin \theta_{ij} - B_{ij} \cos \theta_{ij}],
 \end{aligned} \tag{2}$$

where  $P$  is an active power and  $Q$  is a reactive power,  $\cos \phi$  is power factor.

The power flow jacobian matrix ( $PFJ$ ) at any operating point is defined by the following formula.



**Fig. 1** TV and L Index performance on standard IEEE 14-bus system.

The component  $\Delta|V|$  of  $TV$  represents the change of voltage magnitudes at one load bus. This value is below 1 and above 0 for stable state, otherwise for unstable state. Hence, its component  $\Delta|V|$  is regarded as voltage stability index (VSI) [6].

Practically, it is very difficult for obtaining the  $TVs$  of all operating points in time because the amount of computation behind  $TV$  is also very large. Therefore, this paper presents the GA-SVM method for reducing the amount of calculating.

The subfigure (A) in Fig. 1 shows that the performances of  $L$  index and Tangent Vector are same basically in the normal state. However,  $L$  index fails to note that it is the critical bus prone to system collapse, when bus 14 increases load with rate of 150 % ( $\lambda = 1.5$ ) alone in the subfigure (B). Thus, Tangent Vector is more reliable and accurate than  $L$  index.

$$J_{\text{powerflow}} = \begin{pmatrix} \frac{\partial P_{Ti}}{\partial \theta_j} & \frac{\partial P_{Ti}}{\partial V_j} \\ \frac{\partial Q_{Ti}}{\partial \theta_j} & \frac{\partial Q_{Ti}}{\partial V_j} \end{pmatrix} \tag{3}$$

PFJ can derive the tangent vector ( $TV$ ) of the  $i$ -th node as bellow:

$$TV = \begin{bmatrix} \Delta\theta \\ \Delta|V| \end{bmatrix} = J_{\text{powerflow}}^{-1} \begin{bmatrix} P_{Gi} - P_{Li} \\ Q_{Gi} - Q_{Li} \end{bmatrix} \quad (4)$$

### 3. GA-SVM model

#### 3.1 SVM regression and classification

Given training sample set  $\{x_i, y_i\} i = 1, 2 \dots n$ , with  $x_i \in R^d$  as input and  $y_i \in R$  as output. In SVM, Nonlinear function  $\varphi(x)$  is used for mapping input vector into a high-dimensional feature space. In the space, the linear regression function  $f(x)$  is defined in expression (5), and  $\omega$  represents weight as well as  $b$  is a deviation [19].

$$f(x) = (\omega \cdot \varphi(x)) + b \quad (5)$$

The slack variables  $\xi_i, \xi_i^*$  are introduced to measure the deviation degree  $\varepsilon$  of training sample out of band. The regression function is optimized as bellow:

$$\begin{aligned} \min & \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^m (\xi_i + \xi_i^*), \quad C > 0, \\ \text{s.t.} & \begin{cases} y_i - \omega \cdot \varphi(x) - b \leq \varepsilon + \xi_i, \\ \omega \cdot \varphi(x) + b - y_i \leq \varepsilon + \xi_i^* \end{cases}, \quad \xi_i, \xi_i^* \geq 0, \quad i = 1, 2, \dots, n, \end{aligned} \quad (6)$$

where the parameter  $\varepsilon$  is a nonsensitive loss, and the parameter  $C$  is a penalty factor, which controls on the punishment degree of the sample beyond error. For changing the optimized problem into the dual problem, the Lagrange multipliers  $\alpha_i, \alpha_i^*$  are introduced in the Eq. 7.

$$\begin{aligned} \max & \sum_{i=1}^n y_i (\alpha_i - \alpha_i^*) - \varepsilon \sum_{i=1}^n (\alpha_i + \alpha_i^*) \\ & - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (\alpha_i - \alpha_i^*) (\alpha_j - \alpha_j^*) K(x_i, x_j) \\ \text{s.t.} & \sum_{i=1}^n (\alpha_i - \alpha_i^*) = 0, \quad C \geq \alpha_i, \alpha_i^* \geq 0 \\ & K(x_i, x_j) = \varphi(x_i) \cdot \varphi(x_j), \end{aligned} \quad (7)$$

where  $K(x_i, x_j)$  is called kernel function, which replaces the dot product operation in high dimensional space and realizes the function in the input space in order to avoid dimensional disaster. Furthermore, a kernel function that could satisfy the Mercer's theorem computes a dot product in feature space. Main kernel functions include linear, polynomial and Gaussian RBF. In this work, Gaussian kernel function in Eq. 8 is selected for nonlinear mapping with  $\sigma$  as kernel parameter.

$$\text{Gaussian RBF: } K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\delta^2}\right) \quad (8)$$

Finally, the SVM regression function is expressed as the Eq. 9.

$$f(x) = \sum_{i=1}^m (\alpha_i - \alpha_j^*) K(x_i, x) + b \quad (9)$$

**SVM classification for stability assessment is as shown in reference [10].** To validate the model, a  $k$ -fold cross-validation will be carried out in the experiment, which  $k-1$  folds are used for training and the remaining for testing. Besides, training and testing is repeated until all  $k$  parts have been used as a test set [10].

### 3.2 Basic genetic algorithm

The penalty factors  $C$ , kernel parameter  $\sigma$ , nonsensitive loss  $\varepsilon$  greatly influence the performance of SVM, thus they need be set by combinatorial optimization. Usually, most researchers set them by experience so that the error of calculation is bigger. Therefore, in this paper, genetic algorithm (GA) is utilized to search for better combinations of the parameters instead of manual setting for obtaining the higher accuracy [20]. The optimization process is shown in Fig. 2.

Due to the genetic algorithm has the implicit parallelism and very strong global searching ability, it can search the global optimal point in a very short period. Therefore, in this paper, genetic algorithm is used to optimize these parameters, the specific steps are:

- (i) The SVM parameters encoding and GA parameters settings.
  - (a) Encode the SVM training parameters  $(\varepsilon, \sigma, C)$  in accordance with chromosome gene.
  - (b) Set GA parameters such as crossover rate and mutation rate and so on.
  - (c) Generate a group of seeds randomly as the initial chromosome gene of the SVM parameters.
- (ii) Calculate the fitness of the each individual of population. The average relative error as a fitness function is used to measure the pros and cons of selected parameters.

$$p_{si} = \frac{f_i}{\sum f_i}, \quad (10)$$

when the fitness of sample  $x_i$  is assumed as  $f_i$ , the possibility  $p_{si}$  of sample  $x_i$  is calculated by above Eq. 10. That strategy keeps next generation population fitness and elitist.

- (iii) Genetic selection, crossover and mutation. Replicate the chromosome that has the higher fitness according to the individual fitness through the roulette method; Exchange the two genes in the chromosome randomly through a single point of crossover; Inverse the information bit “0” or “1” of gene encoding in chromosome.

- (iv) Repeat steps from (ii) to (iv), so that a set of initial parameters of support vector machine evolves continuously until the termination condition is satisfied.
- (v) Decode the optimal values to obtain optimized parameters  $(\varepsilon, \sigma, C)$ .

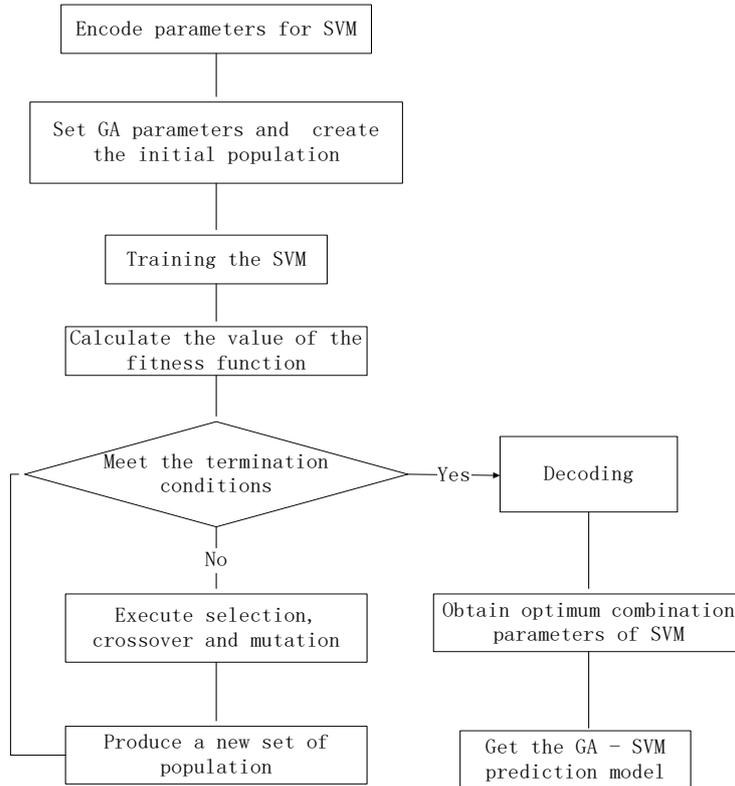


Fig. 2 Optimization process of GA-SVM model.

### 3.3 Genetic algorithm for real-coding problems

The hyperparameters are usually real-valued and continuous in practice, so binary-coded could lead to reduce the efficiency [14]. Furthermore, we take a new strategy based on real-coded method for improving GA. Firstly, the value needs to be changed into a string representation according to preset the number of bits. The crossover operator still takes the single point cross method. Given the two string parents and the length  $l$  of them, an integer is generated from  $0$  to  $l - 1$ , which will be taken as cross point. Then, the latter parts of them are permuted from the cut point to produce two new children strings.

The operator of mutation utilizes the way of non-uniform alteration. The variation probability of element  $x_{i+1}$  is the  $1/10$  of element  $x_i$  in numeric string. Each

element  $x_i$  takes the method of random variation in the integer range of 0 to 9, as  $(x_i = \rho : \rho \neq x_i; \rho = \text{rand}(); \rho \in [0 - 9])$ .

The above method still takes the process steps of basic GA, whose modifications are only the real-coding representation and the two operators of crossover and mutation.

### 3.4 Model selection and initial method

Meta-learning give us a good thought that utilizing the prior knowledge and experience as possible can help to improve learning [16,17]. A number of the same and similar experiments on relatively small-scale system had been implemented before we used this model for assessing complex power system. Thus, we has created a Meta-Database based the Meta-learning model to select the initial parameters and kernel type. The algorithm of Meta-learning is shown as bellow: Given a problem, its input can be represented by a vector  $x = (x^1, \dots, x^n)$ . A function was defined to calculate the distance between the meta-attributes of the input and meta-examples in Meta-base. The Meta-learner searches the  $k$  most similar meta-examples by the function.

$$\text{dis}(x, u) = \frac{1}{n} \sum_{i=1}^n \sqrt{(x_i - u_i)^2} \quad (11)$$

In our Meta-Examples, 11 datasets had been collected including standard model system and non-standard model system, such as standard 3 machine-9 bus system, IEEE 14 bus system and so on [11]. Each of them includes the three parameters and the success rate of different kernels type (linear kernel, polynomial kernel and RBF kernel). Where, the RBF kernel performed relatively better than others did, so we remained the RBF kernel for the SVM. The parameters in database were combined as the initial 1331 configurations of this experiment. The paper [18] proposed a new technology based on “ROC front” concept for selecting the trade-off point of multi-object, but the determined condition is very difficult to be acquired at runtime, so the only left way to decide it is by testing. According to the above conclusion, we firstly recombine the parameters of Meta-Examples to set the initial configuration of GA. Then final goal will be decided by testing the produced population in GA.

## 4. Experiment method

### 4.1 Test system

The proposed method is tested for the IEEE 118-bus systems, which has 118 buses and 50 generators. Many researchers use this standard test system to validate their algorithm [22].

### 4.2 Network simplification

There are many factors impact voltage such as generator level, system status, and load change and so on. In the PFJ matrix, the change of voltage magnitude is a function of active power, reactive power and system load so that input sample should include them. However, the amount of computation is too large for a large

power system. There are 118 nodes with 354 input parameters at least in Fig. 3, which will lead to reducing the convergence speed of the algorithm greatly. Therefore simplifying network is a very necessary work. An important simplified principle is the unchanged boundary conditions. A large power system can be divided into several regions according to the characteristics of geography or electrical, that its power flow computation and stability in each area are calculated respectively, and the link-line situations between areas reflect the mutual influence [21].

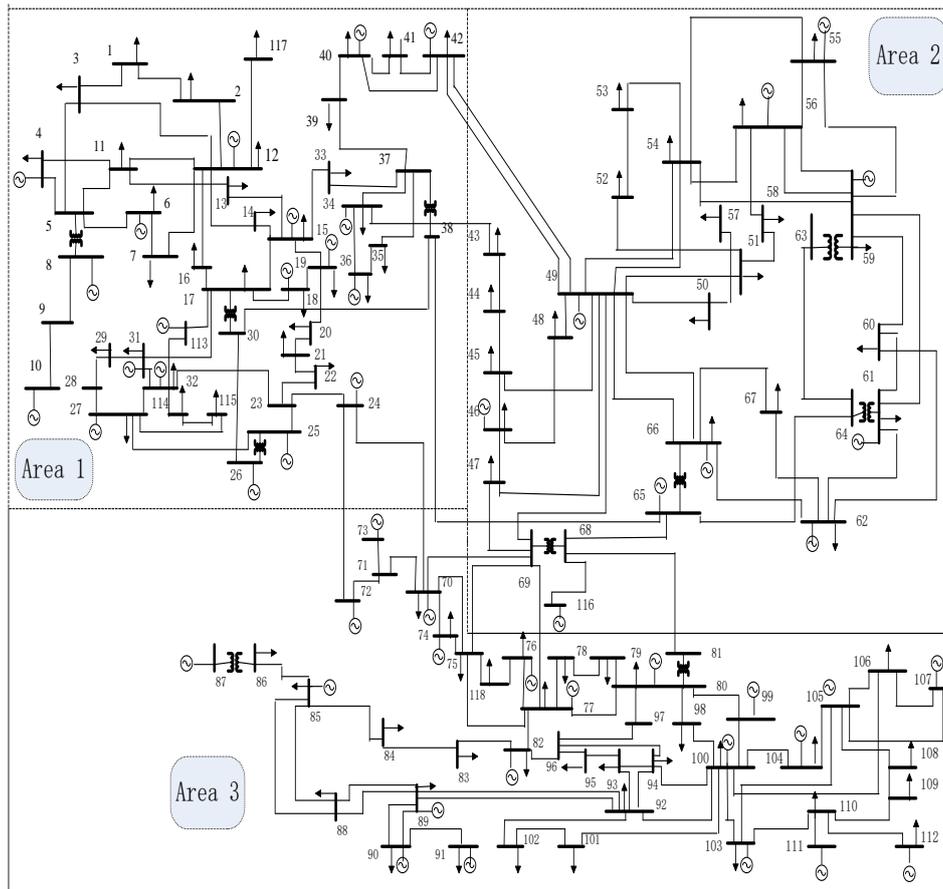


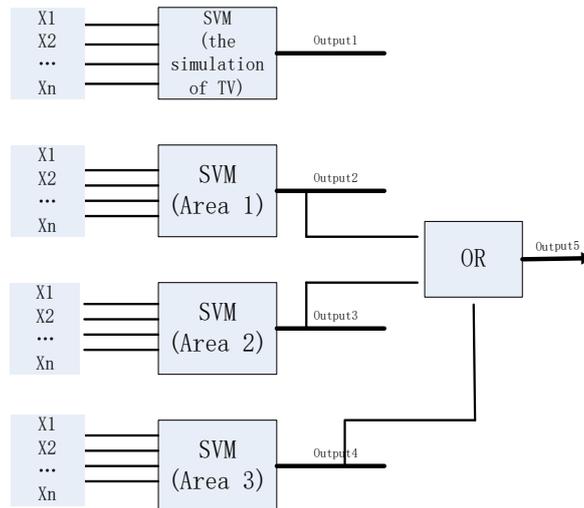
Fig. 3 Network of IEEE 118-bus.

The partition principle is:

- The area of independent calculation has less contact with other area as possible.
- Because the electrical stability has spatial locality, the topology structure remains the same in the area that the study node and its surrounding nodes are in closer distance geographically and electrically.

The IEEE 118-bus test network in Fig. 3 is divided in the three areas of area1, area2 and area3 so that each area has fewer nodes and remains the same topology (the dotted line divides the three areas).

### 4.3 Experimental model



**Fig. 4** System model.

In this paper, four support vector machines are employed for the system status prediction shown in Fig. 4. The three of them are used as the status classifiers for identifying the unstable area and the other is used to simulate the component of TV in real-time. Thus, there are five outputs, that output1 is the simulation value; output2, output3 and output4 represent the stability status of each area respectively; output5 is on behalf of the system state.

### 4.4 Experimental tools and generation of data

The implementation of this algorithm is an improvement for the libsvm and the GA algorithm of MATLAB. All data in the experiment is generated by power system analysis toolbox (PSAT) [22]. 900 loading scenarios were generated by changing the load at each bus and randomly in four levels ( $\pm 50\%$  of base case,  $\pm 25\%$  of base case). Thus this system generated 424800 data ( $900 \times 472$ ).

For responding in time, the MATLAB tools running on a distribution system, that there were four node workers for three areas job, which took the parallel configuration on a Linux cluster system.

### 4.5 The configuration of system

The system parameters need to be set before executing including cross rate, mutation rate and so on, and they are described in Tab. I. Where, 1331 population

Name	Value	Role
Rl	5	Real-coding length
PN	1500	Initial population number
Pt	0.72	Tournament selection rate
Pc	0.8	Cross rate
Pm	0.02	Mutation rates:
Loop	$10^5$	Max iterations number
MSE	0.0005	

**Tab. I** *System parameters.*

initial values were from the recombining records ( $\varepsilon, \sigma, C$ ) of Meta-Database, the remaining was produced randomly in the range of 0 to  $2^{16}$ .

The step size and range of three hyper-parameters is set in Tab. II.

Parameter	Begin	End	Step size
$C$	0	10	0.01
$\sigma$	-1	1	0.00001
$\varepsilon$	-1	1	0.00001

**Tab. II** *Hyper-parameters configuration.*

## 4.6 Training and testing data

In this experiment, a 5-fold cross validation was decided to train and test data set. There were 424800 generated data in the IEEE 118-bus system totally. Where, 339840 data (80% of the total data) were used for training this algorithm and remaining 84960 data (20% of the total data) used for testing.

## 4.7 Experimental steps

The following steps describe the implementation of GA-SVM algorithm.

- (i) Build the IEEE 118-bus test model on PSAT.
- (ii) Carry out simulations to generate experimental data considering.
  - (a) Change the three parameters (active power, reactive power, load) at all buses.
  - (b) Change the three parameters at individual buses.
  - (c) Change each parameter alone.
  - (d) Reflect power factors through link-line between areas and other area regarded as an equivalent system.

- (iii) Make the input vector and output vector. Active power, reactive power and system load as the components of input vector, and the output vector is a value that simulates the component  $\Delta|V|$  of TV.
- (iv) Select 80% of the total data as training set.
- (v) Select Gaussian Kernel as kernel type, and calculate the penalty parameters  $C$  and kernel parameters by the GA.
- (vi) Begin to train the SVM.
- (vii) Compare the output values with target value by means of mean square error (MSE) to adjust the internal weight of SVM.

$$\text{MSE} = \sqrt{\frac{\sum_{i=1}^n \left(\frac{y-y^*}{y}\right)^2}{n}} \quad (12)$$

where  $y$  is the output value and  $y^*$  is target value.

- (viii) Meet the terminal condition, the algorithm produce the candidates.
- (ix) Decide the best parameters by testing the candidates.

## 5. Result and analysis

### 5.1 SVM parameters

The kernel type considered for this model was RBF. The MSE was set less than 0.0005 for reaching the satisfactory accuracy. The GA-SVM algorithm produced the 101 candidates in this experiment. The best parameters were determined by their performance in the last test and their performance was shown in Tab. III.

Algorithm	The average number of iterations	The average convergence time [sec]	The global optimal probability [%]
The GA	231	27000	87.53
Grid-search	23000	833000	73.27
Gradient-search	11450	532000	32.45

**Tab. III** Results of genetic algorithm.

It can be seen from Tab. III that the GA has an obvious superiority, compared with grid-search and gradient-search, in terms of efficiency and accuracy. Therefore, we think the solution of the GA as a better method to set the parameters of SVM. The best settings are chosen in Tab. IV including the GA, grid-search, gradient-search and manual setting. The GA method has the lowest MSE by comparing with the others in test, so it is easy concluded that the proposed method has optimal accuracy.

Parameters	the GA	grid-search	gradient-search	manual
$C$	9.34000	22.17000	11.22000	15
$\sigma$	0.47300	0.01425	0.57352	0.5
$\varepsilon$	0.00330	0.02475	0.003344	0.01
MSE	0.00623e-005	0.03022e-005	0.02514e-005	0.83371e-005

**Tab. IV** Selected parameters and MSE.

## 5.2 Algorithm performance

In order to make a comparison between the proposed method and other methods, their performances are shown in Tab. V. In the experiment, all methods took the  $K$ -fold cross validation for training and testing. In Tab. V, the learning time and testing time are the average of  $k$  times validation. The mean learning time includes only the time required to train the model after performing the hyper-parameters selection. Besides, the spending time of the proposed method during the evolutionary task was 270000 s. From Tab. V, it is found that the training time for GA-SVM, SVM and BPNN are 662 ms, 1041 ms and 2400 ms respectively, and the testing time for the three methods are 96 ms, 216 ms and 339 ms respectively. This notes that the GA-SVM is better than SVM and BPNN in performance. At same time, it is also noted that the GA-SVM has higher accuracy than the others do. It can be concluded from the above analysis that the GA-SVM is much more suitable for the dynamic assessment of a large power system.

Model	Mean learning time [ms]	Mean testing time [ms]	Mean squared error [MSE]
GA-SVM	662	96	0.00623e-005
SVM	1041	216	0.01422e-005
BPNN	2400	339	0.09613e-005

**Tab. V** Different algorithms for system value.

## 5.3 Test result

There are 28320 data for unstable case and 56640 data for stable case in the total of testing data.

Reliability of GA-SVM classification is determined in Tab. VI. Stable cases were classified as stable all the time. However, only less than 3.4% of the unstable cases (955 out of 28320) were classified as stable by mistake.

Cases	Classified as stable [%]	Classified as unstable [%]
Stable case	100 (56640/56640)	0 (0/56640)
Unstable case	3.372 (955/28320)	96.628 (27365/28320)

**Tab. VI** Prediction Accuracy on testing data.

#### 5.4 Unstable area prediction

The output of the status classifier in each area is 0 or 1 with 0 for stability and 1 for instability. In order to find which bus could produce the error of locating area in algorithm, each bus will be to increase load (+50%) alone with the other buses remaining constant. It can be seen from Tab. VII that there are 6 buses classified wrongly for GA-SVM and 12 buses for BPNN.

Bus No.	Area No.	GA-SVM	BPNN	Bus No.	Area No.	GA-SVM	BPNN
24	1	3	3	69	2	2	3
34	1	1	2	70	3	3	3
38	1	1	2	72	3	1	1
42	1	2	2	5	3	2	2
43	2	1	1	77	3	2	2
49	2	2	1	81	3	3	3
68	2	2	3				

**Tab. VII** Prediction of border buses.

Tab. VIII shows that the GA-SVM method can locate the area with the correct rate of 94.91% and BPNN with rate of 89.83. Thus, the GA-SVM compared with BPNN in classification with higher success rate.

Method	Error No.	Correct rate [%]
GA-SVM	6	94.91
BPNN	12	89.83

**Tab. VIII** Prediction accuracies.

#### 5.5 Identifying critical buses

Critical bus is the one vulnerability of system after it falls into a fault. It will reduce the complex work to find the critical bus and take more attention on it. For the system of IEEE 118-bus, the load is raised with rate of 150% ( $\lambda = 1.5$ ) at each bus alone and the other buses keep constant respectively. From Tab. IX, it was observed that there are 13 critical buses, the value of which is greater than

1. It is also seen that there are some buses less than 1 and near to 1 in Fig. 5. Besides, the rest of buses are not sensitive. Thus, they can be ranked three levels for management that the 13 critical buses are the first level that need to pay more attention, the buses in the range of 0.6 and 1 for second level with less attention, and the rest buses for the third level. Therefore, this work is very significant that it can simplify management, reduce costs and improve efficiency.

Bus No.	Area	Value of TV
12	1	1.78
15	1	1.34
17	1	1.92
37	1	1.34
49	2	1.59
56	2	1.27
68	2	1.30
69	2	1.34
70	3	1.33
77	3	1.77
80	3	1.65
100	3	1.44
105	3	1.33

Tab. IX Critical buses.

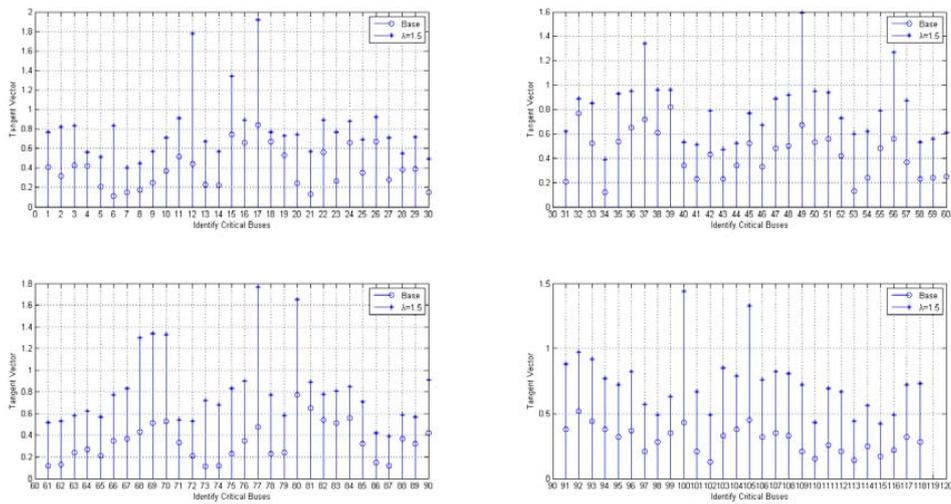


Fig. 5 Identify critical buses.

## 6. Conclusion

This paper presented an accurate, fast and reliable model to analyze and predict the transient process of a large power system after it is subjected to a fault. In this technique, an improved genetic algorithm is used to obtain the better combinational parameters for SVM instead of manual setting. In order to enhance the performance of the GA, the parameters in Meta-Database were used to set the initial population. The method utilized SVM as learning machine to calculate the component of  $TV$  quickly and as a classifier to predict the state of each area and system.

The study carried out for IEEE 118-bus test system indicated that the prediction reliability could achieve success rate of 98.87%. It was demonstrated that this approach is able to predict system status and trigger protection control application in real time. The analysis of critical buses based on GA-SVM is helpful to find the weakness of system, and it will be beneficial for maintaining of the complex system. Besides, there are still a lot of research chances. For examples, this technique is applied in radial distribution system and topology changing.

Additionally, when the system is identified unstable using this proposed method, the locating of the fault component is further research direction.

## Acknowledgement

This work was supported by the National Natural Science Foundation of China (61502330) and Science and Technology Innovation Project of Shanxi Province (20171119) and Project of Soft Science Project of Shanxi Province (2016041008-5).

**Statement** We declare that there is no conflict of interest regarding the publication of this paper.

## References

- [1] KARLSSON D., LINDAHL S. Wide area protection and emergency control, In: *Proc. IEEE Power Eng. Soc. General Meeting*, 2004, 1, p. 5.
- [2] PHADKE A.G. Synchronized phasor measurements in power systems, *IEEE Comput. Appl. Power*, 6(2), Apr. 1993, pp. 10–15.
- [3] ANDERSON P.M., FOUAD A.A. *Power System Control and Stability*, 2nd ed. Piscataway, NJ: IEEE, 2003.
- [4] KESSEL P. GLAVITCH H. Estimating the voltage stability of power systems, *IEEE Trans. on Power Delivery*, 1986, 1(3), pp. 346–354.
- [5] NIZAM M., MOHAMED A., HUSSAIN A. Performance evaluation of voltage stability indices for dynamic voltage collapse prediction, *Journal of Applied Science*, 6(5), 2006, pp. 1104–1113.
- [6] ACHARYA N.V., RAO P.S.N. A new voltage stability index based on the tangent vector of the power flow jacobian, *ISGT Asia (978-1-4799-1346-6)*, 2013 IEEE, pp. 1–6.
- [7] KAI S., LIKHATE S., VITTAL V., KOLLURI V.S., MANDAL S. An online dynamic security assessment scheme using phasor measurements and decision trees, *IEEE Trans. Power Syst.*, 22(4), Nov. 2007, pp. 1935–1943.
- [8] GAVOYIANNIS A.E., VOGIATZIS D.G., GEORGIA D.R., HATZIARGYRIOU N.D. Combined support vector classifiers using fuzzy clustering for dynamic security assessment, In: *Proc. IEEE Power Eng. Soc. Summer Meeting*, 2001, 2, pp. 1281–1286.

- [9] HAN X., ZHENG Z., TIAN N., HOU Y. Voltage stability assessment based on BP neural network, Asia-Pacific Power & Energy Engineering Conference (978-1-4244-2486-3), March 2009, doi: [10.1109/APPEEC.2009.4918962](https://doi.org/10.1109/APPEEC.2009.4918962).
- [10] GOMEZ F.R., RAJAPAKSE A.D., ANNAKKAGE U.D. Support Vector Machine-Based Algorithm for Post-Fault Transient Stability Status Prediction Using Synchronized Measurements, IEEE Transactions on Power Systems, August 2011, 26(3), pp. 1417–1483.
- [11] LI Q. Analysis and assessment of voltage stability based on GA-SVM, Journal of system simulation, 2014, 26(10), pp. 2369-2373,, +2406.ISSN:1004-731X.
- [12] LI W., LIU L., GONG W. Multi-objective uniform design as a svm model selection tool for face recognition, Expert Syst. Appl., 2011, 38(6) pp. 6689–6695.
- [13] YUAN R., GUANGCHEN B. Determination of Optimal SVM Parameters by Using GAPSO, journal of computers, 5(8), 2010, pp. 1160–1168.
- [14] LORENA A.C., DE CARVALHO A.C. Evolutionary tuning of SVM parameter values in multiclass problems, Neurocomputing, 2008, 71, pp. 16–18.
- [15] ZHANG X., QIU D., CHEN F. Support vector machine with parameter optimization by a novel hybrid method and its application to fault diagnosis, Neurocomputing, 2015, 149, pp. 641–651.
- [16] MIRANDA B.C., PRUDENCIO R.B.C., CARVALHO A.C.P.L.F., SOARES C. Combining meta-learning with multi-objective particle swarm algorithms for svm parameter selection: an experimental analysis, in: Proceedings of the 2012 Brazilian Symposium on Neural Networks, SBRN '12, IEEE Computer Society, Washington, DC, USA, 2012, pp. 1–6.
- [17] MIRANDA B.C., PRUDENCIO R.B.C., CARVALHO A.C.P.L.F., SOARES C. Multi-objective optimization and meta-learning for svm parameter selection, In: The 2012 International Joint Conference on Neural Networks (IJCNN), 2012, pp. 1–8.
- [18] CHATELAIN C., ADAM S., LECOURTIER Y., HEUTTE L., PAQUET T. A multi-model selection framework for unknown and/or evaluative misclassification cost problems, Pattern Recognit, 2010, 43(3), pp. 815–823.
- [19] JIN-LONG AN, QING-XIN YANG, ZHEN-PING MA Study on SVM on-line function regression method for mass data, Machine Learning and Cybernetics, 2007 International Conference (978-1-4244-0973-0), 2007, 5, pp. 2773–2777.
- [20] ZHOU J., BAI T., SUO C. The SVM optimized by culture genetic algorithm and its application in forecasting share price, Granular Computing (978-1-4244-2512-9), 2008, pp. 838–843.
- [21] XINYUAN L., YONGHUA S., JUNYONG L., YING F., MIN Z. Post-fault voltage security assessment of complex power system based on ANN, Proceedings of the Chinese Society of Electrical Engineering, 1996, 16(3), pp. 147–150.
- [22] Power System Test Archive-UWEE (University of Washington) Available-online, <http://www.ee.washington.edu/research/pstc>