Middle-East Journal of Scientific Research 24 (9): 2927-2935, 2016
ISSN 1990-9233

DOI: 10.5829/idosi.mejsr.2016.2927.2935

Informative Vector Machine Based Steady State Security Assessment of Power Systems

R. Thamizhselvan and S. Ganapathy

Department of Electrical Engineering, Annamalai University,
Annamalai Nagar - 608 002, Tamil Nadu, India

Abstract: In this paper, a machine learning based Informative Vector Machine (IVM) algorithm for power system Static Security Assessment (SSA) has been presented. Recently, Machine learning algorithm based classification has become trendy and been successfully applied to many fields including power system and so on. The proposed IVM approach classifies the power system operating scenarios into secure and insecure states that validate accurate measure of system security. This methodology can serve as a decision making tool in power system operation and control to take preventive and remedial actions if system is insecure. The proposed SSA-IVM classifier determines the status of the system security within a less time period and satisfies the demand of low dimensionality of data that supports quick decision in real time. It is a big challenge in reducing the size of input variables and speed up the computation; a correlation based F- value feature selection is used for this purpose in this paper. For that reason, there is a need for finding a smaller set of variables from the larger data set and validating those variables having more discriminatory and useful information to realise about static security. Furthermore, this paper influences a set of selective parameters such as real and reactive power generations that strengthen the performance of the proposed classifier. Experimental results are validated by simulating various scenarios with a standard IEEE-118 Bus system.

Key words: F-value method · Feature selection · Informative Vector Machine · Machine learning · Static Security Assessment

INTRODUCTION

Power system security and reliability have become challenging issues in the present restructured scenario. The problem of monitoring the power flows and bus voltages in a power system is very important in maintaining system security and fast prediction is essential for controlling these quantities. As power systems have become more stressed due to increased loading and large interconnections, there will be an increase in cases of voltage limit violation and line loading limit violation, particularly in contingency conditions like line outage, generator outage etc. Under emergency conditions, power system operator has to take quick decisions without caring much for the optimality of the operating condition [1].

In general, Security assessment (SA) is performed to find out which line is overloaded and up to which level the overloading is present along with the load bus voltage magnitude. It is necessary to ensure that all the power flows and load bus voltage magnitude on the selected network are maintained within the specified limits however, not only in base case condition but also in stressed/line outage conditions. As a result, an accurate as well as fast computation of lines flows, identification of overloaded lines and prediction of line overloading in different overloaded branches are essential for ensuring reliable power system operation. This situation demanding a fast, reliable security assessment is of paramount importance in modern power systems [2].

Security assessment can be classified as Static, Transient and Dynamic security assessment. Static security is one of the main and important aspects of power system security assessment. Static security is defined as the ability of the system to reach a state within the specified secure region following any specified contingency. The standard approach to the security assessment problem is to perform the static security analysis at first, then Transient and dynamic security analyses are planned based on the successive information

Corresponding Author: R. Thamizhselvan, Department of Electrical Engineering, Annamalai University, Annamalai Nagar - 608 002, Tamil Nadu, India.

about SSA. The static security analysis evaluates the post contingent steady state of the system, neglecting the transient behaviour and any other time-dependent variations [3, 4]. On the other hand, the transient and dynamic security analyses evaluate the time-dependent transition from the pre-contingent to the post contingent state. In this paper, only SSA is considered. The key issues in SSA are fast identification of the set of insecure contingencies and their evaluation creating quick impact on the power system operation.

The steady-state load flow model is a well established approximation for understanding a power system in normal operation where load and generation vary slightly around a base case. The assessment of the present and the impact of possible line or generator outages are determined by solving numerically the non-linear load flow equations for all contingencies [2]. In literature, a survey of several power flow methods are available to compute line flows in a power system like Gauss Seidel iterative method, Newton-Raphson method etc, but these are either approximate or too slow for on-line implementation [6, 7].With the expansion of artificial intelligence based techniques such as artificial neural network, fuzzy logic etc. in recent years, there is a growing trend in applying these approaches for the operation and control of power system [5]. Applications of Artificial neural network systems for SSA have gained popularity over the conventional methods as they are efficient in classifying the patterns as secure and insecure.

The use of ANN based pattern recognition (PR) approach, Decision tree based security classifier, Genetic algorithm based neural network , fuzzy logic combined with neural network , query-based learning approach in neural networks [8, 9, 10] for static security evaluation process have also been reported in [11-15]. But these procedures may not have addressed the issue of large number of possible contingencies in power system operation. Furthermore, ANN based SSA is found to be highly time consuming and infeasible for real time applications [12, 13, 14].

In response to these issues, Machine learning algorithms are considered as alternatives for all the type of power system security problems, The application of Support vector machine (SVM) for various studies of security assessment has been reported earlier [16-19]. In this reference, the superior performance of SVM over ANN in terms of accuracy, speed and distribution of high-risk cases has been presented for the steady state security of a large-scale power system. However, SVM learning algorithms suffer from exceeding time and memory requirements, if the training pattern set is very large. To overcome the draw backs of SVM, Core Vector Machine (CVM) and Ball Vector Machine (BVM) have been considered as optional classifiers that reformulate SVM's quadratic programming as a minimum enclosing ball (MEB) problem or enclosing ball (EB) problem and an efficient ($1+\xi$) approximation algorithm to obtain a close- to-optimal SVM solution have been successfully applied to solve many large-scale classification problems [20-22]. However, the online learning issue of the CVM and BVM classifiers is still not addressed. In these two algorithms, data are processed in a batch mode. When a new training sample arrives, the whole training process should be implemented once again to adjust the classifier. On the other hand, Relevance Vector Machine (RVM) is another new machine learning algorithm, which is based on Probabilistic Bayesian framework and has been implemented for better classification performance on SSA and overcomes the drawbacks of existing methods. But it pertains to some inconvenience like previous existing machine learning models particularly misclassification rate is not tolerable with larger power networks [23]. To prevail over the drawbacks of previous machine learning based classifiers, a Gaussian processes based Informative Vector Machine (IVM) [24] has been proposed in this paper. It models more sparsity, good generalization performance, free choice of kernel function and distributive prediction. Because of these advantages, the results obtained from an IVM are superior to existing algorithms. In the proposed IVM approach, the performance is superior as compared to other machine learners due to the possibility of misranking being eliminated [25, 26].

In order to improve the performance of classification, it is necessarily required to elaborate off-line computations for generation of good feature set as the input variables. Therefore, the features which can be considered for static security classifier must be significant variables from the classification point of view. As mentioned in literature [18], under certain justigfied assumptions, the generator currents and load currents can be expressed as a function of generator currents. The generator currents are directly related to the real and reactive powers of the generators. Hence, the real power and reactive power demands can be expressed as a function of real and reactive powers of generators [16]. Hence, in this paper, the proposed initial feature set is having only real and reactive power generations of each generator (i.e., P_{Gi}, Q_{Gi}), capable of providing sufficient discriminating information about the class of system security (secure or insecure). The proposed IVM approach is implemented on a standard IEEE 118- bus system. The simulation results prove that the IVM classifier gives a professional classification, enhancing its suitability for on-line security assessment and even evidences reduction in size of feature space and error rate by exploiting by way of an efficient F -Value feature selection method.

The Proposed Methodology

Significance of Training and Testing Data Set: An off-line process which is based on load flow solution by Newton Raphson (NR) method to represent the dataset for all possible operating condition of the power system has been considered. The data set is generated by varying the load and generation between 80% to 130% of the base case and the voltage magnitude is assumed between 0.90 p.u –1.10 p.u for all test systems and line over load limit in MVA [15] is taken as 130% of base MVA flow. For each operating condition, Single line outage is simulated and steady state variables are obtained from load flow solution that are listed as voltage magnitude V_i, voltage angle δ_i, real power generation P_{Gi}, reactive power generation Q_{Gi}, real power demand P_{Di}, reactive power demand Q_{Di}, active $P_{i\text{-}j}$ the real power flow in line connected between buses i and j, $Q_{i\text{-}j}$ the reactive power flow in line connected between buses i and j, $S_{i\text{-}j}$ line MVA between buses i and j.

For quick evaluation of static security, a number of variables have been considered out of few hundreds of steady state variables of which all of them will not be significant [5]. The main objective handled in this paper is to predict static security status in a faster manner with selection of most effective variables for feature selection and classification. According to the literature, Arora and Surana [17] have derived that if power system is not optimally dispatched, the feature set consisting of pre-disturbance real and reactive power generations and real and reactive power demands at each system bus carry sufficient information about system security. Under certain justified assumptions, the generator currents and load currents can be expressed as a function of generator currents. The generator currents are directly related to the real and reactive powers of the generators. Therefore, the real power and reactive power demands can be expressed as a function of real and reactive powers of generators. Therefore, the proposed initial feature set consists of pre-disturbance real and reactive power generation of each generator, i.e., P_{Gi} and Q_{Gi}, respectively [16, 17]. Having selected the features the next step is to obtain the decision function or classifier for SSA.

Brief Review of Informative Vector Machine: IVM is a novel machine learning algorithm based on Gaussian processes(GP) and introduced by Lawrence *et al*. [24-26] which helps better for classification and regression problems. It is formulated as a sparse Gaussian process based on Bayesian theory, which has high speed in model training, small consumption of memory, strong effectiveness in sparseness and good classifying performance. It is a more fundamental model than the support vector machine for function approximation in learning.

Consider a simple hidden variable model and its hidden variable of f , observed variable y and the input data x as given in eqn. (1), (2) and (3) that are independent each other,

$$f = \left\{ f(x_j) \right\} \; (j = 1, ..., N) \tag{1}$$

$$y = \{y_i](j = 1,...,N) \tag{2}$$

$$x = \left\{ x_i \right\} \left(j = 1, ..., N \right) \tag{3}$$

In this model, the hidden function f follows the input data x and the observed value y follows the hidden variable f. By changing the observed value to a continuous value or a discrete value, it can be used for “iteration” or “classification”. The GP assumes a multidimensional normalized distribution with an average of 0 for a joint distribution for any points. As a result, the a priori distribution for the hidden variables is given by the equation.

$$p(f \mid X, \theta) = N(f; 0, K) \tag{4}$$

Note that θ is a parameter for the kernel function; N, the normal distribution; K, the kernel function matrix.

Here, the variance covariance matrix can be replaced by a kernel function matrix. Next, Consider the simultaneous probability distribution for the hidden variable and the observed value is given by the eqn. (5).

$$p\,(y, f \mid X, \theta) = p(f \mid X, \theta)\; p(y \mid f) = p(f \mid X, \theta) \prod_{n=1}^{N} p(y_n \mid f_n) \tag{5}$$

where $p(y_n|f_n)$ is the noise model which gives the function between the hidden variable and the observed value.

By assuming that the noise model is the normal distribution is given by,

$$p(y_n|f_n) = N\ (y;\ 0,\ B^{-1}) \quad (6)$$

where B is the triangular matrix in which the n^{-th} element is given by β_n.

Let us consider the priori distribution $p(y \mid X, \theta)$ for the observed value. This quantity can be obtained by integrating f in Eqn. (5). Also, it can be calculated directly by using normal distribution as given by;

$$p(y \mid X,\theta) = \int N(f;0,K)N(y;0,B^{-1})df = N(y;0,K+B^{-1}) \quad (7)$$

Now concluding the equations, by introducing the distribution for the hidden variable f(x*) in the unknown data set x*. The posteriori distribution $p(f|X,y,\theta)$ can be represented by the following equations using Bayes' theorem with Eqn. (5) and (7),

$$p(f \mid X,y,\theta) = \frac{p(y,f \mid X,\theta)}{y \mid X,\theta} = N(f;\mu,\Sigma) \quad (8)$$

$$\mu = \Sigma By,\ \Sigma = (B+K^{-1}) \quad (9)$$

where $\mu = [\mu_1..........\mu_N]^T$ is the average vector for the posteriori distribution and $\Sigma = [\Sigma_1.....\ \Sigma_N]^T$ is the variance covariance matrix for the posteriori distribution. Moreover, the posteriori distribution for the hidden vector $f(x^*)$ can be readily marginalized by assuming a normal distribution is given by;

$$p(f(x^*) \mid X,y,x^*,\theta) = \int p(f(x^*) \mid f,X,x^*,\theta)p(f \mid X,y,\theta)$$
$$df = N(f(x^*);\mu(x^*),\sigma^2(x^*)) \quad (10)$$

The mean and dispersion for the posteriori distribution for x* in the new data is represented [24] using the following equation based on x* and the input data X, kernel function vector k, the kernel function matrix K for X and eqn. (9) is modified as;

$$\mu(x^*) = \mathrm{kTK}^{-1}.\Sigma\ \mathrm{B\ y} \quad (11)$$

$$\sigma^2(x^*) = k(x^*, x^*) + k^T K^{-1} (\Sigma - k) K^{-1} k \quad (12)$$

where k is the kernel function vector for x* and X; and k(x*, x*) is the kernel function for x* and x*.

Let us select the value of $I = \{n_i\}$ (i = 1, ..., d) as the active set and |I| = d < N. the data selection involves by selection of the data set $J = \{x_j\}/\ I$ (j=1, ..., N) which maximizes the amount of entropy updates for the posteriori distribution after the data is included in the active set.

Normal Noise Model: A noise model is the necessary parameter for IVM data selection used by GP. Consider a normal distribution is referred to as a normal noise model based on eqn. (3) as given by;

$$p(y_{ni} \mid f_{ni}) = N(y_{ni};0,\beta_{ni}^{-1}) \quad (13)$$

Moreover, the expected value of the noise model, assuming a normal distribution, can be readily calculated and the expected value Z when selecting n_i elements included in the i-th active set is given by the following equation using eqn. (3).

$$Z_{ni}(\mu_{i-1,n_i},\varsigma_{i-1,ni}) = \int p(y_{ni} \mid f_{ni})N(f_{ni};\mu_{i-1,ni},\varsigma_{i-1,ni})df_{ni}$$
$$= N(y_{ni};\mu_{i-1,n_i},\varsigma_{i-1,ni}+\beta_{n_i}^{-1}) \quad (14)$$

where $\mu_{i-1,ni}$ is the n_i element in the posteriori vector μ_{i-1} in the active set; and $\varsigma_{i-1,\ ni,}$ is the n_i triangular element in the posteriori variance covariance matrix Σ_{i-1} in the active set.

Next, the two parameters shown in the following equation are obtained by taking the logarithm of the expected value Z and finding the derived function:

$$g_{ni} = \frac{\partial \log(Z_{ni}(\mu_{i-1,ni},\varsigma_{i-1,ni})}{\partial \mu_{i-1,ni}} = \frac{y_{ni}-\mu_{i-1,ni}}{\beta_{ni}^{-1}+\varsigma_{i-1,ni}} \quad (15)$$

$$G_{ni} = \frac{\partial \log(Z_{ni}(\mu_{i-1,ni},\varsigma_{i-1,ni})}{\partial \varsigma_{i-1,ni}} = -\frac{1}{2(\varsigma_{i-1,ni}+\beta_{ni}^{-1})} + \frac{1}{2}g_{ni}^2 \quad (16)$$

Using Eqns. (15) and (16), we obtain;

$$\upsilon_{ni} = g_{ni}^2 - 2G_{ni} = \frac{1}{\varsigma_{i-1,ni}+\beta_{ni}^{-1}} \quad (17)$$

where v is called the update element and plays a vital role in updating data selection and kernel parameters.

Data Selection: The IVM selects the data in accordance with the difference between the entropy in the posteriori distribution including the new data in active set and the entropy before including the new data. Note that the initial active set is empty. The entropy of the posteriori normal distribution given by the variance covariance matrix Σ and the mean vector μ is defined in the eqn. (18),

$$H(N(.;\mu,\Sigma)) = \frac{N}{2}(1+\log 2\pi) + \frac{1}{2}\log|\Sigma| \tag{18}$$

Therefore, the difference ΔH in the entropy after including the data is given as follows using Eqns. (17) and (18):

$$\Delta H_{i,ni} = \mathrm{H}(\mathrm{N}^{new}) - \mathrm{H}(\mathrm{N}) = -\frac{1}{2}\log(1-v_{i,ni}\varsigma_{i-1,ni}) \tag{19}$$

Note that N^{new} is the normal distribution after data inclusion; I is the N×N unit matrix; and e_{ni} is the n_{ni} element in the unit matrix and ΔH is found for all elements in the data set J= {j}\I (j=1...N), the maximum j is included in the active set I and n_i = J. This is given in the eqn. (20),

$$J = n_i = \arg\max_{n\in J} \Delta H_{i,n} \tag{20}$$

Based on the above, the element n_i included in the active set is selected. However the number of elements in the active set increases as a result of data selection and so μ_I and Σ_I change. μ_I and Σ_I are called the kernel parameters and must be updated every time the data selection is performed for all the d times it is performed.

To perform classification using the active set created through the data selection. The mean and dispersion for a posteriori distribution for x* in unknown data can be represented using the equations below based on eqn. (11) and (12),

$$\mu(x^*) = k_I^T K_1^{-1}\Sigma_1 B_I y_I \tag{21}$$

$$\sigma^2(x^*) = k(x^*,x^*) + k_I^T K_1^{-1}(\Sigma_I - K_I)K_I^{-1}k_I \tag{22}$$

Updating the Kernel Parameters: Updating the kernel parameters is required whenever calculating ΔH_{ini}. The equation for updating the variance covariance matrix Σ_I and the mean vector μ_i in the posteriori distribution can be represented with the following equations using eqn. (13) through (17).

$$\mu_i = \mu_{i-1} + g_{ni}\Sigma_{i-1}e_{ni} \tag{23}$$

$$\Sigma_i = \Sigma_{i-1} - \left(g_{ni}^2 - 2G_{ni}\right)\Sigma_{i-1}e_{ni}e_{ni}^T\Sigma_{i-1} \tag{24}$$

Here, the equations above were derived using Assumed Density Filtering (ADF). Normally, when approximating a posterior distribution, the Kullback-Leibler divergence KL (q||p) between the true posteriori distribution q and the approximated posterior distribution p is minimized based on the variational Bayes method [25]. However, calculating the moment of the true posterior distribution is impossible and so is difficult to work with. In contrast, ADF assumes that the derived function for the inverse Kullback- Leibler divergence KL (q||p) is zero and so can find the parameters for the true posterior distribution [24, 25].

Let us consider a diagonal element vector ζ_I for the variance covariance matrix Σ_i.

$$\zeta_I = \mathrm{diag}(\Sigma_i) \tag{25}$$

ζ_I is necessary for calculating the update element v_{ni} and $\Delta H_{i,ni}$. The following equations result when the diag operator is used for Eq. (24):

$$\zeta_I = \zeta_{i-1} - \upsilon_{i,n_i}\, diag\left(s_{i-1,n_i}\, s_{i-1,n_i}^T\right) \tag{26}$$

$$s_{i,n_i} = \sum_i e_{ni} \tag{27}$$

In the same fashion, the mean vector μ_i for the posteriori distribution is necessary to calculate v_{ni} and can be represented as shown in the equations below using Eqn. (23).

$$\mu_i = \mu_{i-1} + g_{ni}S_{i-1,ni} \tag{28}$$

$$\mu_0 = 0 \tag{29}$$

Here, the updated eqn. (26) and (28) are needed to calculate the variance covariance matrix, Σ_i can be represented using the equation below by using the row vector s_{i-1} in Eqn. (24).

$$\Sigma_i = \Sigma_0 - M_i^T M \tag{30}$$

$$\Sigma_0 = K \tag{31}$$

Note that, K is the kernel function matrix for the input data X; and M_i is the i×N matrix where the i-th row is given by $\sqrt{v_{i,ni}S_{i-1,ni}}$, Where S_{i-1}, $n_i = \Sigma_{i-1}\, e_{ni} = K_{ni} - M_{i-1}^T$ m_{i-1} .m_{i-1}, n_i.

The most important evaluator of the IVM performance is the Classification accuracy which is defined as the ratio of successfully classified patterns to the number of patterns in Data set of IEEE 118 test bus system.

$$\text{Classification Accuracy (CA)} = \frac{\text{Number of Samples classified correctly}}{\text{Total number of Samples in the Data set}} \tag{32}$$

Misclassification can be of two distinguished types of error, depending on the actual class of the misclassified patterns,

$$\text{Misclassification (MC)} = \frac{\text{Number of sample classified wrongly}}{\text{Total number of Sample in the Data set}} \tag{33}$$

$$\text{Alert misclassification (AMC)} = \frac{\text{Number of Secure Sample classified as insecure}}{\text{Total number of Secure Sample in the data set}} \tag{34}$$

$$\text{Emergency misclassification (EMC)} = \frac{\text{Number of Insecure Sample Classified secure}}{\text{Total number of Insecure Sample in the dataset}} \tag{35}$$

In power system security evaluation, the alert misclassifications are not of much harm, but emergency misclassification may lead to brutal blackout [2]. It is, therefore, important to ensure that EMC are kept as minimal.

Feature Selection Process: The essential problem in classification using IVM is identifying a representative set of features from which to design a classifier model [18, 19]. This paper addresses the problem of feature selection for machine learning classifier through a correlation based F-Value approach. The central hypothesis is that good feature sets contain features that are highly correlated with the class, yet uncorrelated with each other.

The number of variables characterizing a power system operating state is quite large which all of them will not be significant. This makes the security classifier design complicated and requires large computational resources [2].

In order to reduce the number of features, F-value feature selection which is based on the measure of interest and intraset distances has been used .It can be concluded in the following stages; first the features are selected from pattern vector based on maximization of a criterion function. The F-value defined by eqn. (31) is used as the criterion function for selection of a variable as feature,

$$F = \frac{|\, m_s - m_i \,|}{(\sigma_s + \sigma_i)} \tag{31}$$

where, m_s - Mean of the variable in the secure class, m_i- Mean of the variable in the insecure class
σ_5^2- Variance of the variable in the secure class, σ_i - ^{2} Variance of the variable in the insecure class.

It is clear that, $|m_5 - m_i|$ is the measure of interset distances between the class i and class j and is the measure of dispersion of the variable in the two classes. The F- value can be interpreted as a measure of the classification error that could arise when the variable is used as the only feature. The larger the value of F, the lesser would be the interset distance or/and smaller value would be the dispersion. In other words, larger value of F would result in smaller classification error. Hence variables to be used as features should have high values of F.

The selection of features begins with the computation of F-values for all components (variables) of pattern vector in the training set. The variable with the largest F value is selected as a first feature. Let this variable be z_1. When selecting other features, redundant information is omitted by discarding these variables which are correlated to z_1, i.e. those variables having a correlation coefficient greater than 0.8, [2, 5] say. The proceeding procedure is repeated until the required number of features has been reached, or the F-value of the remaining variables is small. The optimal set of above features serves as an input database for designing the IVM classifier.

RESULTS AND DISCUSSION

It has been observed that in many security assessments, the ability of any good classifier can be realised by its performance level when incorporated on a larger power system rather than a smaller and simple test system. Hence in this paper, a larger IEEE 118 bus system has been considered to prove the efficiency of the proposed IVM classifier.

The proposed IVM classification approach to static security evaluation is implemented an IEEE 118 bus power system [27]. The effectiveness of the proposed classifier has been demonstrated by comparing with existing machine learners such as SVM, CVM and RVM. The required data set for training and testing phases are obtained by off-line simulation by Newton Raphson load flow performed using MATPOWER Toolbox with MATLAB 7.1 [28]. This data set is obtained by random fluctuation of generation and load from 80% to 130% of their base case value with generation variation limited to their minimum and maximum limits. The security limit for bus voltage magnitude is taken as 0.90 to 1.10 p.u. The MVA limit of system branches is assumed as 130% of base case.

The IEEE-118 bus sample system details [29] are shown in Table 1 and 2, one at a time, outage studies are performed and form the set of disturbances to be utilized for steady state security in the power system. The patterns or variables are generated through the load flow results. The generated variable set consists of 118 numbers of voltage magnitude variables (V_i), 118 numbers of voltage angle (δ_i), 19 numbers of real power generation variables (P_{Gi}), 54 numbers of reactive power generation variables (Q_{Gi}), 99 numbers of real power demand variables (P_{Di}), 99 numbers of reactive power demand variables (Q_{Di}), 186 numbers of active real power flow variables (Pi-j) , 186 numbers of reactive power flow variables (Qi-j), 186 numbers of reversal of real power flow (Pj-i), 186 numbers of reversal of rective power flow (Qj-i), 186 numbers of real power losses(Ploss), 186 numbers of reactive power losses(Qloss) and 186 numbers of line MVA variables (Si-j). As a result, the total system variables are counted to 1809 variables initially.

Irrelevant variables at certain buses such as zero loads, zero generation and constant values are neglected. Likewise, the system variable that represents real and reactive power losses has been rejected that amounting to 1001 operating states as input variables for classification. All feasible 1001 variables are subjected to static security check with voltage limit and line flow limit.

Table 1 and 2 show the complete details about an IEEE 118 bus system. It can be seen that, for a possible 1063 cases as operating scenarios, 744 operating scenarios are found to be secure and the remaining 319 cases are found to be insecure. The training and testing samples are split in random by the ratio of 80 % (850 cases) for training phase and 20% (213 cases) for testing phase as given in Table 3.

Hence, the IVM classifier is obtained by use of the splitted training and test data sets to provide better classification results with important features from input attributes. Table 4 shows an optimal set of patterns selected using F-Value feature selection process. Among all these variables, effective input features have been selected by using correlation based feature ranking analysis. By having a threshold value of 0.8, 412 and 11 featured inputs (uncorrelated components) were obtained.

Table 1: Details of IEEE 118 bus system components

IEEE 118 Bus System Components	
No of Buses	118
No of Generators	54
No of Committed Gens	54
No of Loads	99
No of Fixed loads	99
No of Shunts components	14
Total no of Branches	186
No of Transformers	9

Table 2: Details of IEEE 118 bus system capacity

IEEE 118 Bus System	P(Mw)	Q(Mvar)
Total gen capacity	9966.2	--
Generation (actual)	4374.9	795.7
Load	4242	1438
Fixed	4242	1438
Shunt (inj)	--	84.4
Losses (I^2 *Z)	132.86	783.79
Branch Charging (inj)	--	1341.7

Table 3: Data set for Training and Testing Phases

Scenarios	Overall	Training	Testing
Total No of Cases	1063	850	213
Secure Cases	744	597	147
Insecure Cases	319	253	66

The effectiveness of the dimensionality reduction has been determined with a threshold value of 0.8 and the highly correlated variables are discarded from the total pattern variables with all feasible variables including P_G, Q_G as a scenario and another case study considering P_G and Q_G variables alone.

It is clear from Table 5, that the training accuracy of IVM classifier gives better performance by choosing the optimized kernel parameters. In order to justify this training result, any good classifier desires to ensure higher value of accuracy and less error rate forever. From Table 6, it is evident that the performance of IVM classifier is improved with selection of good feature set and elimination of redundant data.

As seen from Tables 6 and 7, the proposed IVM is very superior in classification in terms of higher classification accuracy and less misclassifications. The blend of IVM classifier with F-Value algorithm yields zero error or 100% accuracy for classifying the data patterns for both the case studies under consideration.

During testing phase, an overall efficiency of 99.53% has been achieved with all possible steady state variables for the IEEE 118 bus system, whereas maximum efficiency of 100 % has been achieved for the same118 bus system with F-value feature selection. With the implementation of F-value feature selection, the size of the input variables gets reduced from 1001 variables to 412 variables.

Table 4: Feature Selection Process

Case study	IEEE118 bus system (All feasible variables)	IEEE118 bus system Only P_G and Q_G
No. of pattern variables	1001	73
No. of features selected	412	11

Table 5: Choice of kernel parameters for better accuracy

Model (kernel)	Alpha max	epsilon	Noise model	Training accuracy of IVM (%)
Gaussian	e^{-12}	e^{4}	7	100 (850/850)
Linard, rbf	e^{-12}	e^{5}	9	99.76(848/850)

Table 6: Results of IVM-SSA classifier before and after feature selection

Performance Evaluation of IVM	Training with 1001 Variables considering 850 patterns	Testing with 1001 Variables considering 213 patterns	Training with 412 Variables considering 850 patterns	Testing with 412 Variables considering 213 patterns
CA (%)	100 (850/850)	99.53(212/213)	100 (850/850)	100 (213/213)
MC (%)	0 (0/850)	7.526(1/213)	0 (0/850)	0 (0/213)
AMC (%)	0(0/597)	0 (0/147)	0 (0/597)	0 (0/147)
EMC (%)	0 (0/253)	1.515 (1/66)	0 (0/253)	0 (0/66)

Table 7: Results of IVM-SSA classifier with only P_G and Q_G as selected

Performance Evaluation of IVM	Training with 73 Attributes and 850 patterns	Testing With 73 Attributes and 213 patterns	Training With 11 Attributes and 850 patterns	Testing With 11 Attributes and 213 patterns
CA(%)	100 (850/850)	98.12 (209/213)	100 (850/850)	100 (213/213)
MC (%)	0 (0/850)	1.87 (4/213)	0 (0/850)	0 (0/213)
AMC (%)	0 (0/597)	0 (0/147)	0 (0/597)	0 (0/147)
EMC (%)	0 (0/253)	6.06 (4/66)	0 (0/253)	0 (0/66)

Similarly, the steady state security assessment has been carried out by limiting the variables corresponding to P_G and Q_G alone. The number of input variables considering only P_G and Q_G without feature selection is 73 and its corresponding classification accuracy is 98.12 %, whereas, the number of input variables considering only P_G and Q_G with feature selection is 11 and its corresponding classification accuracy is 100 %. Results prove that, the performance of IVM model considering P_G, Q_G as input variables is similar to that of considering all the steady state variables. Hence, it can be concluded that the steady state security assessment considering only P_G and Q_G is sufficient for obtaining efficient classification with maximum accuracy.

CONCLUSION

IVM-based SSA technique for an IEEE 118 bus power system has been proposed. The proposed correlation based F-value feature selection algorithm is more relevant and suitable to deal with the problem of high dimensionality as it will restrict the number of variables that are needed to be obtained accurately. Simulation results prove that, such technique is feasible and provides a deeper insight into the system performance, since it presents fast, accurate and relevant information about the system state. The proposed model holds the promise as fast classifier for static security of large scale power systems with the presence of real and reactive power generation alone in the IEEE 118 Bus system. A Gaussian based kernel has been established to segregate the data space which is relevant to current condition of the system into several classes, thus improving accuracy and to minimize error rate. The selected Gaussian kernel can identify the resemblance of distinctive input and outputs of the dataset quickly. Future work will focus on the application of IVM classifier with different kernels and optimizing its parameters for real time transient and dynamic security assessment of power systems.

REFERENCES

1. Balu, N., T. Bertram, A. Bose, V. Brandwajn, G. Cauley, D. Curtice, A. Fouad, L. Fink, M.G. Lauby, B.F. Wollenberg and J.N. Wrubel, 1992. On-line power system security analysis. Proceedings of the IEEE, 80(2): 262-282.
2. Pang, C.K., A.J. Koivo and A.H. El-Abiad, 1973. Application of Pattern Recognition to Steady-State Security Evaluation in a Power System. IEEE Transactions on Systems, Man and Cybernetics, 3(6): 622-631.

3. Kirschen, D.S., 2002. Power system security. Power Engineering Journal, 16(5): 241-248.
4. Echavarren, F.M., E. Lobato, L. Rouco and T. Gomez, 2011. Formulation, computation and improvement of steady state security margins in power systems. Part I: Theoretical framework, International Journal of Electrical Power & Energy Systems, 33(2): 340-346.
5. Sinha, A.K. and I.J. Nagrath, 1984. Pattern Recognition Method for Power System Steady state Security Assessment. Journal of Institution of Engineers (India), 64: 269-271.
6. John Kaye, R. and Felix F. Wu, 1984. Analysis of Linearized Decoupled Power Flow Approximations for Steady-State Security Assessment. IEEE Transactions on Circuits and Systems, 31(7): 623-636.
7. Mitra, P., V. Vittal, B. Keel and J. Mistry, 2016. A Systematic Approach to-1-1 Analysis for Power System Security Assessment. IEEE Power and Energy Technology Systems Journal, 3(2): 72-80.
8. Matos, M.A., N.D. Hatziargriou and J.P. Lopes, 2000. Multicontingency steady state security evaluation using fuzzy clustering techniques, IEEE Transactions on Power Systems, 15(1): 177-183.
9. Srivani, J. and K.S. Swarup, 2008. Power system static security assessment and evaluation using external system equivalents. International Journal of Electrical Power & Energy Systems, 30(2): 83-92.
10. Kim, H. and C. Singh, 2005. Power system probabilistic security assessment using Bayes classifier. Electric Power Systems Res., 74(1): 157-165.
11. Reddy, G., T. Ananthapadmanabha, N. Mohan and L. Kumar, 2016. Optimal DG placement based on static security assessment, International Journal of Electrical Engineering & Technology, 7(2): 37-49.
12. Sekhar, P. and S. Mohanty, 2016. An online power system static security assessment module using multi-layer perceptron and radial basis function network, International Journal of Electrical Power & Energy Systems, 76: 165-173.
13. Lekshmi, M., A. EEE and M.S. Nagaraj, 2014. Development of an Online Static Power System Security Assessment Module Using Artificial Neural Networks in 118-Bus Test System. International Journal of Modern Computer Science, 2(6): 74-79.
14. Saeh, I., W. Mustafa and N. Al-geelani, 2016. New Classifier Design for Static Security Evaluation Using Artificial Intelligence Techniques, International J. Electrical and Computer Engineering, 6(2): 870-876.
15. Kalyani, S. and K.S. Swarup, 2009. Power System Security Assessment using Binary SVM Based Pattern Recognition, International Journal of Electrical and Electronics Engineering, 3(11): 710-716.
16. Niazi, K.R., C.M. Arora and S.L. Surana, 2004. Power system security evaluation using ANN: feature selection using divergence, Electric Power Systems Research, 69(2): 161-167.
17. Arora, C.M. and S.L. Surana, 1996. Transient security evaluation and preventive control of power systems using PR techniques, Journal of Institution of Engineers India, 76: 199-203.
18. Ruba, S. and M.S. Thanabal, 2016.Optimization Algorithms for Feature Selection in Classification: A Survey, International Journal of Innovative Research in Computer and Communication Engineering, 4(2): 1124-1127.
19. Saeys, Y., I. Inza and P. Larranaga, 2007. A review of feature selection techniques in bioinformatics. Bioinformatics, 23(19): 2507-2517.
20. Tsang, I.W., J.T. Kwok and P.M. Cheung, 2005. Core vector machines: Fast SVM training on very large data sets, Journal of Machine Learning Research, 6: 363-392.
21. Mohammadi, M. and G.B. Gharehpetian, 2009. Application of core vector machines for on-line voltage security assessment using a decision tree-based feature selection algorithm. IET Generation, Transmission & Distribution, 3(8): 701-712.
22. Mohammadi, M. and G.B. Gharehpetian, 2010. On-line transient stability assessment of large-scale power systems by using ball vector machines, Energy Conversion and Management, 51: 640-647.
23. Tipping, M.E., 2001. Sparse Bayesian learning and the relevance vector machine. Journal of Machine Learning Research, 1: 211-244.
24. Lawrence, N.D., J.C. Platt and M.I. Jordan, 2005. Extensions of the informative vector machine. In Deterministic and Statistical Methods in Machine Learning, 1: 56-87.
25. Kurata, E. and H. Mori, 2009. Short-term load forecasting using informative vector machine, Electrical Engineering in Japan, 166(2): 23-31.
26. Elattar, E.E., 2013. Day-ahead price forecasting of electricity markets based on local informative vector machine, IET Generation, Transmission & Distribution, 7(10): 1063-1071.
27. Lawrence, N.D., and C.P. John, 2004. IVM SOFTWARE: Learning to learn with the informative vector machine. Proceedings of the twenty-first international conference on Machine learning. ACM, NY, USA.
28. Zimmerman, R. and C. Murillo-Sanchez, 2007. MATPOWER: a MATLAB Power System Simulation Package, Power Systems Engineering Research Centre (PSERC).
29. Test Case Archive, 1996. Electrical Engineering, University of Washington, Seattle, WA.