

Security-State Estimation of Electrical Power Systems: A Methodological Proposal

Fredy A. Sanz[1(✉)], C. A. Collazos[1], Iván Ruiz[1], and Farid Meléndez-Pertuz[2]

[1] Vicerrectoría de Investigaciones, Universidad Manuela Beltrán, Bogotá, Colombia
fredy.sanz@umb.edu.co
[2] Departamento de Ciencias de la Computación y Electrónica, Universidad de la Costa, Barranquilla, Colombia

Abstract. Security assessment techniques are a tool that allows the operational personnel of electrical systems to categorize operating conditions in different labels. Nowadays, the large size of power grids and the operating conditions close to their operating limits require efficient, safe and fast procedures to assess the overall grid security, being one of the challenges posed by smart grids conception. In this paper, a supervised learning approach that allows to integrate all the concepts of safety of the electrical network: voltage, angle, frequency, etc., to give an overall assessment of the system in terms of security which may be applied according to the regulations and geographical contexts of each country.

Keywords: Power system security · Power system stability · State estimation · Smart grids

1 Introduction

In the context of Security Analysis (SA), the study of systems behavior depends on a set of contingencies, identifying if the system has necessary security conditions and is capable of surviving after an event or contingency. The power systems behavior and operating restrictions (transformers tabs limits, line power flows limits, etc.), depend on their nature and configuration (power grid topology). Therefore, risks derived from contingencies can affect power system security in different ways, some common faults are: tripping lines, transformers outages, and generators outages, or a combinations of the above, etc. [1, 2].

For on-line SA, different methods have been proposed. In [3] an application using artificial neural network (ANN) models was proposed, considering load changes, and ranking them in the order of severity based on composite security index. Some heuristic and statistical techniques to obtain a power grid diagnosis are used [4–6], although previous works assume that specific contingencies are known, i.e. admittance matrix was assumed known. In [7] the estimation of the equivalent dynamic model through the PMU's (Power Measurement Units) measurements was considered. However, contingencies were not considered, implying an unrealistic case of the power grid due to

© Springer Nature Switzerland AG 2019
S. Misra et al. (Eds.): ICCSA 2019, LNCS 11620, pp. 713–730, 2019.
https://doi.org/10.1007/978-3-030-24296-1_57

topological flexibility and variability that experiment the power grid as a result of failures. On the other hand, in [8] an on-line dynamic security assessment is proposed using the terms of the transient energy function as input features to a machine learning algorithm.

Therefore, in this paper a flexible methodology based on machine learning techniques for on-line security state estimation is proposed, which brings integrally: the concept of security assessment, fast power system security status estimation, flexibility in metrics selection (stability indices), and a reliable estimation of security state under different load scenarios, contingencies and failures. Our proposal consists of the following stages: measurements preparation, stability indices calculation, security state assessment, model building based on machine learning techniques.

This paper is organized as follows. Section 2 presents a brief overview of machine learning approach for power system security-state estimation. The snapshot of methodology proposal is presented in Sect. 3. The IEEE 39-bus test system is used to apply the methodological proposal; implementation and results are shown in Sects. 4 and 5 respectively. Finally discussions and conclusions are presented.

2 Machine Learning in Power System Security-State Estimation

Under the smart grids conception, it is important to note that power systems must integrate new measurement technologies, controls, renewable energies, etc., allowing to have an on-line estimation of the security status. The increasingly inclusion of phasor measurement units (PMUs), permits quickly and accurately measurements of the electrical variables, which are usually stored in a set of instantaneous measurements x_{PMUs}, as shown Eq. 1, including: voltage $V_{i,j}$, currents $I_{i,j}$, active powers $P_{i,j}$, reactive powers $Q_{i,j}$, etc., (where the subscripts i indicate the time when measurement is taken and j involve the bus where a measure is made), serving as input to computational models that support a quick estimation of security status.

$$x_{PMUs} = [V_{i,1} \ldots V_{i,j} \;\; I_{i,1} \ldots I_{i,j} \;\; P_{i,1} \ldots P_{i,j} \;\; Q_{i,1} \ldots Q_{i,j} \ldots] \tag{1}$$

Generally power flows calculation provide data about angles, voltages, active and reactive power, line power flows, switching devices and transformers. In short, a small change in the system load or its configuration involves changes in electrical variables (voltages, currents, power, etc.), i.e. for every change there is a security state of the power grid (normal, alert, emergency, etc.). In this way, as it is shown in Eq. 2, the power grid measurements x_{PMUs} is used to obtain an estimate of the security state y, by means of a function f obtained through machine learning techniques for better time responses. Therefore, the main objective of the methodological proposal is to find a model f, reliable and safe, to predict the security status of the power grid from PMU's measurements.

$$x_{PMUs} \xrightarrow{f} y = \{normal, alert, emergency, \ldots, etc.\} \quad (2)$$

The security state definition depends on particular issues of the power system. In light of smart grids, a flexible methodology that can be made as robust as required will be proposed, which allows the user to define model requirements, according to the on-line measurable variables and limitations of the power system, i.e. it is possible to decide assess: the static security, the dynamic security, or both.

The security state estimation process include two main phases: *(i)* modeling phase, *(ii)* security state estimation phase, as shown Fig. 1. The second phase is the model implementation in the power grid to security status monitoring. The objective of this paper is a methodological proposal for the modeling phase, which includes the measurement preparation, indices calculation, security state assignment, and model building. Details on each components of the methodology are presented in Sect. 3.

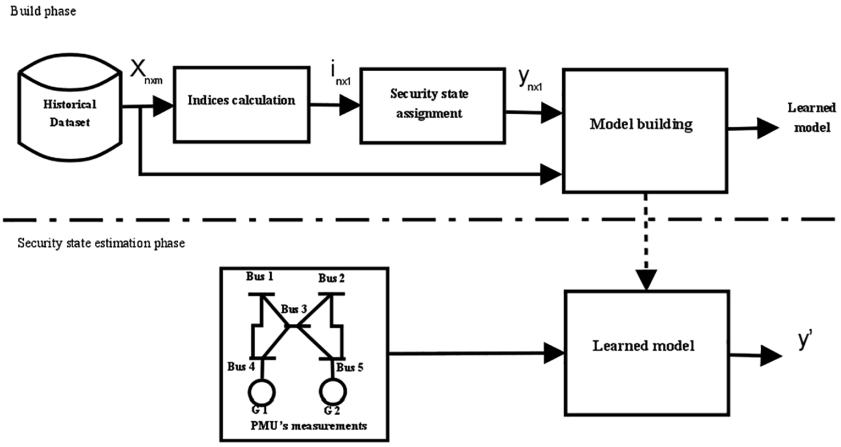

Fig. 1. Security state estimation process for one stability condition (voltage stability, angle stability, frequency stability, etc.)

3 Machine Learning in Power System Security-State Estimation

The methodological proposal is composed of four stages: measurements, indices calculation, security state assignment, and model building. The methodology procedure was schematized previously as modeling phase through Fig. 1, stages are described below. Notice that the procedure described in Fig. 1 includes only one stability condition (e.g. voltage stability, angle stability, frequency stability, etc.). In case that several stability problems are require to analyze, several instances of the procedure must be applied, one for each stability condition.

3.1 Measurements

When a real power system is studied, measurements recorded with PMUs can be considered within a real-time context. However, the power grid contingencies and faults can involve topological changes due to outages in lines, transformers and generators. Which generally cannot be determined immediately, requiring previous analysis for topological changes identification (many authors considers that configuration changes are known after the faults are analyzed and identified [4–7]).

These historical set of measurements X_{PMUs}, is presented in extended form in Eq. 3. where *I, V, P* and *Q* are the currents, voltages, active powers, and reactive power respectively. The subscripts: *i* indicates the measurement of each variable in a moment of time, and *j* involves the bus where a measure is taken. Security assessment is a combination of system monitoring and contingency analysis [9], from the practical point of view these processes are made online and offline respectively, to identify risky situations and to plan remedial actions. Strictly, if a power system requires static and dynamic security evaluation, the following safety criteria may be considered [10]: angle stability, frequency stability, transient stability, voltage stability, small-signal stability, thermal overloading of electrical network equipment, etc.

$$X_{PMUs} = \begin{bmatrix} V_{1,1} \cdots V_{1,j} & I_{1,1} \cdots I_{1,j} & P_{1,1} \cdots P_{1,j} & Q_{1,1} & \cdots & Q_{1,j} \cdots \\ \vdots \ddots \vdots & \vdots \ddots \vdots & \vdots \ddots \vdots & & \ddots & \vdots \\ V_{i,1} \cdots V_{i,j} & I_{i,1} \cdots I_{i,j} & P_{i,1} \cdots P_{i,j} & Q_{i,1} & \cdots & Q_{i,j} \cdots \end{bmatrix} \tag{3}$$

The goal of this stage is to obtain a dataset adapted to the security requirements of a particular power system, this procedure includes: *(i)* taking measurements from the power grid, *(ii)* recording of electrical variables as historical dataset, *(iii)* security assessment according to the anomalies observed in the historical dataset, and *(iv)* dataset refinement taking into account the observed instabilities. The measurements set preparation is illustrated in Fig. 2.

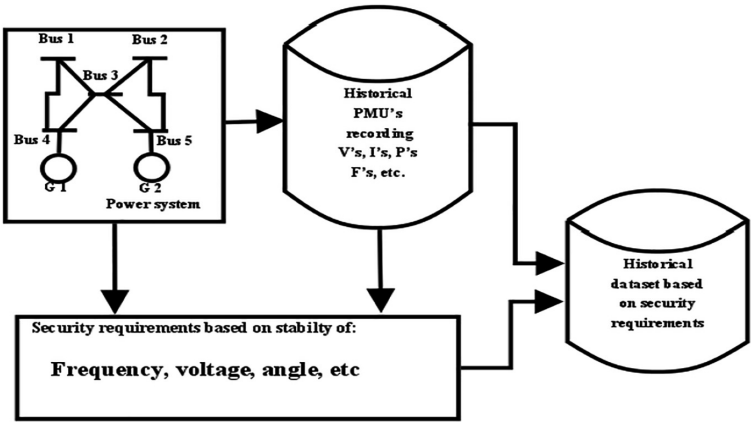

Fig. 2. Measurements set. Using historical dataset, the power grid configuration and their contingencies is possible to determine the existing instabilities and therefore the variables required to estimate the security state.

3.2 Indices Calculation

After deciding which are the stability conditions to consider, appropriate metrics must be selected in order to measure the security degree associated with each stability condition. For example, typical indicators of voltage stability are: Voltage Collapse Proximity Index (VCPI) [11], Fast Voltage Stability index (FVSI) [12], Voltage Stability Index [13], Voltage Margin (Kv) [2], Area of the Voltage Stability Region (AVSR) [14] etc. The stability indicators selected must be subjected to: (i) Knowledge a portion or all of the variables which are used to calculate it, which can be known through PMU's measurements, *(ii)* a familiarity criterion with the index, which regards to operational conditions, efficiency studies, current regulations, etc., providing flexibility to the user's needs.

For stability angle, the following indicators can be used amount others [15]: Expected Stability Margin (EASM), Weighted Probability Index (WPI), and Integral Square Bus Angle (ISBA). In frequency stability analysis some commonly used indices are [16]: Frequency Deviation Assessment (FDSA), Phase Stability Index (PSI), and Constant Frequency Limiter (CFL). On the other hand, considering small-signal stability the following indices may be considered [17]: Single Mode Index (SMI), All Modes Index (AMI), Global Modes Index (GMI), etc. Therefore, the dataset selected must be include the required measurable variables according to the metrics selected for stability evaluation.

According to the types of stability condition selected for the power system, the respective indices are calculated using the measurements contained in Eq. 3. The calculation is performed for each observation (for each instant of time).

3.3 Security State Assignment

Generally, the security states are assigned to measurements groups that fall within a degree of security which are called classes (normal, alert, emergency, …). They must obey to: the security analysis procedures proposed in the literature [2, 4, 18, 19].

The metrics were previously selected considering measurable parameters with PMUs. After this, security conditions of the power system should be selected and classified in ranges for each index, in detail, the security state ranking process these stages *(i)* electrical variables are measured on-line, *(ii)* a parallel estimator for each stability type is calculated separately, *(iii)* a concept of system security is issued, considering the kind of stability with higher risk. This procedure is shown in Fig. 3, considering different security states and a ranking processor which must select the more severe stability condition and thus emit a security state concept.

The security states assignment described in Fig. 1 includes only one stability condition, however if power system have multiple stability risks, the security state assignment will be obtained through the ranking process described about and shown in Fig. 3.

3.4 Model Building

The goal in this phase is to learn a model f that maps the measurements from PMUs to a security state. The model building includes two stages: *(i)* selection of the most discriminative variables, and *(ii)* learning a classifier that discriminates (separate) observed variables into different classes. The basis of these stages are explained bellow.

Variables Selection. The goal in the variables selection stage is to obtain a lower dimension representation of an original dataset with many samples while maintaining the discriminability between different classes, i.e., security states in the context of this work.

Given the dataset $\{\boldsymbol{X}, \boldsymbol{y}\}$, where $\boldsymbol{X} \in R^{N,M}$ is the matrix containing measured variables from the power system, $\boldsymbol{y} \in R^{N}$ is a vector of discrete values containing the security state $y \in \{1, 2, \ldots, K\}$ assigned to each sample, M is the total number of variables recorded from the PMUs, K is the total number of security states and N is the total number of observed samples; the variable selection aims to obtain a new dataset $\{\boldsymbol{X'}, \boldsymbol{y}\}$, where $\boldsymbol{X'} \in R^{N,P}$ is the matrix containing P selected variables and $P < M$. If the least significant buses to estimate the security status are identified, then the buses that govern the power grid stability are also known. [20]. Generally, variables selection techniques can be chosen through the modeler experience or by benchmarking, some useful techniques are: Square Correlation coefficient (R^2) [21], Fisher's score (f_{score}) [22], Correlation-based Feature Selection (CFS) [23].

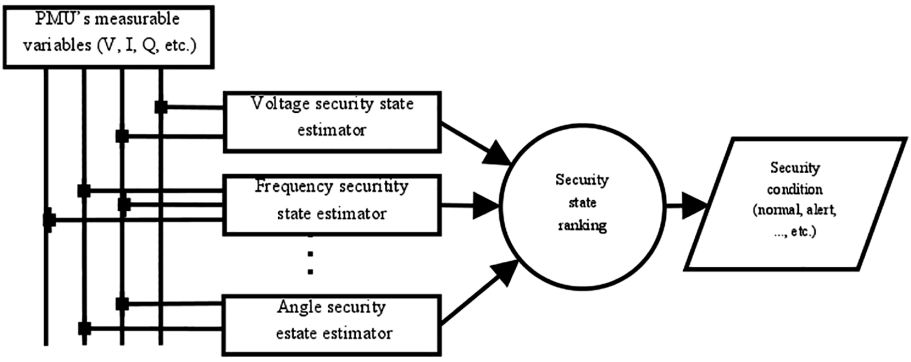

Fig. 3. Security state ranking for two or more stability condition

Classifier. The classifier is responsible for estimating the security state, i.e. it is the model that can take as input the selected PMU's measurements and infers the security state of the power system. The classifier takes an input data $\mathbf{x}_{\text{PMUs}}$ representing the observations from the PMUs in a given time and assigns it to one of K discrete classes.

$C = \{C_1, C_2, \ldots, C_K\}$, representing the security states of the power system [24]. The classification model has to be learned from a training dataset $\{\boldsymbol{X}, \boldsymbol{y}\}$, where $\boldsymbol{X} \in R^{N,M}$ is matrix with N samples of the M variables measured from the PMUs, $\boldsymbol{y} \in R^{N}$ is a vector of the security state labels assigned to each sample, $y \in \{1, 2, \ldots, K\}$ and K is the number of security states or classes. To do so, the classifier

divides the space of the input vector into decision regions separated by decision boundaries or hyperplanes [24]. Different techniques have been documented as Naive Bayes (NB) [25], Support Vector Machines (SVM) [26], Feed Forward Neural Networks (NN) [27] and Linear Discriminant Analysis (LDA) [28]. Considering the flexibility and future effectiveness of the proposed methodology, we recommend to review the most prominent existing learning algorithms techniques.

4 Methodology Implementation

In order to demonstrate and validate the usability of the proposed methodology, the IEEE 39-bus test system (New England) was used [29]. This test power systems is shown in Fig. 4. This power system was assumed to be vulnerable to voltage stability, leaving aside other stability types. Therefore, the power grid security is modeled and evaluated to be robust respect to voltage stability.

Measurements. As the selected power system is vulnerable respect to voltage stability, voltage measurements and the configuration of the power grid defined in the admittance matrix $\boldsymbol{Y}$ are needed to compute the security state [30].

The New England system was simulated under different conditions to obtain a dataset. The system was subjected to outage-line contingencies on each line, and also to generator-outage contingency in each generator. Therefore the dataset includes the New England voltages obtained by simulation using power flows under different load profiles, lines and generators outages in each devices.

As a result, the measurement matrix is $\boldsymbol{X}_{VS} \in R^{N,M}$, where $M = 39$ is the number of voltage buses and $N = 18764$ is the total number of observations in the different scenarios.

Indices Calculation. There are several alternatives to select a metric to compute the voltage stability [2, 11–14]. In this work the Voltage Collapse Proximity Index (VCPI) was adopted as suggested in [11]. The VCPI for each bus is computed as:

$$VCPI_J = \left| 1 - \frac{\sum_{m=1 m \neq j}^{M} V_m^{'}}{V_j} \right| \tag{4}$$

Where $V_m^{'}$ is defined as:

$$V_m^{'} = \frac{Y_{jm}}{\sum_{l=1 l \neq j}^{M} Y_{jl}} \tag{5}$$

Where, V_j is the voltage phasor at bus j, V_m is the voltage phasor at bus m, Y_{jm} is the admittance between bus j and bus m, Y_{jl} is the admittance between bus i and bus l, j is any monitored bus, m is other bus connected to bus j, M is the number of buses. The term $V_m^{'}$ indicates the current flow participation of bus m. The $VCPI_j$ value ranges between 0 and 1. If the value of this metric is close to 0, it is assumed that the node is stable; however, if the value of the index is close to 1, the node is unstable and a

voltage collapse in the power systems is expected. Using historical data to build a model is advantageous, since post-contingency analysis can be performed, to learn more about the power grid as configuration and fault types, i.e. the admittance matrix Y is known, and the voltage stability index may be calculated. Note that if the historical dataset includes voltages measurements, each voltage pattern corresponds to an stability state.

Security State Assignment. Four security states were considered *(K = 4): (i)* normal, all system variables are within the nominal range, the system operates in a secure manner; *(ii)* alert, if security levels fall below a certain limit of adequacy; *(iii)* emergency, when the system has been weakened to a level where a contingency may cause an overloading of equipment that places the system in a hazardous state; and (iv) in extremis, occurs if the disturbance is very severe, this state may result after the alert state occurs.. In this work, the VCPI ranges associated with each security level are presented in Table 1. These ranges were applied to the security state of each sample recorded from the selected power system.

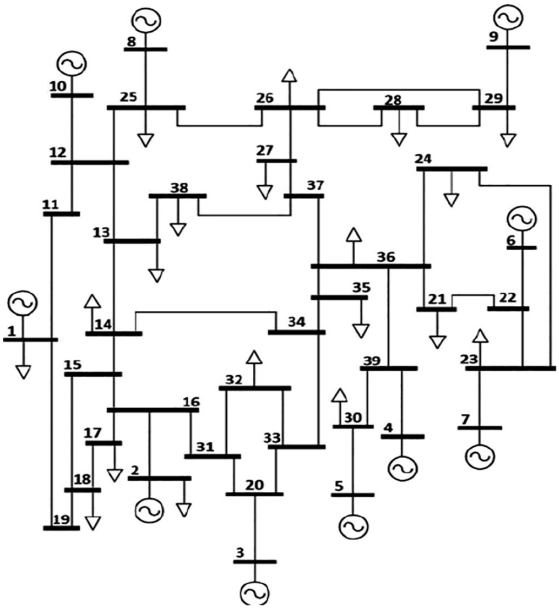

Fig. 4. IEEE 39-bus test system

Table 1. Security-state ranges for New England system

Normal	Alert	Emergency	In extremis
$0 \leq VCPI < 0.25$	$0.25 \leq VCPI < 0.5$	$0.5 \leq VCPI < 0.75$	$VCPI \geq 0.75$

This criterion is applied to each previously calculated VCPI. As a result, each measurement contained in the dataset is associated to a security state or equivalently to one discrete value of {1, 2, 3, 4} for normal, alert, emergency and in extremis, respectively.

Model Building. Given the measurement matrix and the security state, we can construct the dataset $\{\boldsymbol{X}, \boldsymbol{y}\}$, where $\boldsymbol{X} \in R^{N,M}$ is the matrix containing voltages measured from the power system, $\boldsymbol{y}$ is a vector of discrete values containing the security state $y \in \{1, 2, 3, 4\}$ assigned to each sample, N is the number of samples and M is the number of bus voltages or variables recorded from the PMUs. This dataset is:

$$X_{PMUs} = \begin{bmatrix} V_{1,1} & \cdots & V_{1,M} \\ \vdots & \ddots & \vdots \\ V_{N,1} & \cdots & V_{N,M} \end{bmatrix}, y = \begin{matrix} y_1 \\ \vdots \\ y_N \end{matrix} \tag{6}$$

As a result of the simulation process $N = 18764$ measurements were obtained, of which *3858* belongs to normal, *1286* to alertness, *1336* to emergency state, and *12284* to in extremis. In the case of in extreme state a greater number of observations are included because the system was subjected to multiple contingencies, forcing it to severe conditions that affect security, and in the case of normal state there is a smaller number of observations because this state is only one of the studied cases.

Variables Selection. The variables selection techniques used in this implementation are: square correlation coefficient, Fisher's score, and Correlation-based features selection. These techniques are described below.

Squared correlation coefficient *(R^2)*: This technique computes the squared correlation coefficient for every variable measured from the power system and the vector of security state values or class labels [21], and then selects the P variables with higher R^2 values. The square value of the correlation coefficient R^2 is computed independently for each variable as:

$$R_j^2 = \frac{\sum_{i=1}^{N}(x_{i,j} - \bar{x}_j)(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{N}(x_{i,j} - \bar{x}_2)^2 \sum_{i=1}^{N}(y_i - \bar{y})^2}} \tag{7}$$

Where $x_{j,i}$ is the *i-th* sample of the *j-th* variable measured from the PMUs, y_i is the security state or class label associated with the *i-th* sample (measurement at time *i*) and the bar notation represents the average value across all observed samples. After the R^2 values of all variables are computed, the $P < M$ variables presenting the highest R^2 are selected, but, no specific criterion exists to choose P. Although the R^2 method is simple and easy to understand, it presents some limitations: *(i)* the R^2 value is computed independently for each variable, thus, correlated or redundant variables cannot be identified, *(ii)* the method could discharge variables with low individual R^2 value, but with very high R^2 value when combined with other variables and *(iii)* the number of P selected variable has to be manually defined. These limitations may occur in the security state estimation in power systems using voltage measurements, since neighbor buses may have similar behavior, i.e. they are correlated [20].

Fisher's score (f_{score}): This technique also selects a lower dimension dataset where each variable is chosen according to a score computed under the Fisher criterion [31]. The score is computed independently for each variable considering the distance between samples for the same class (expected to be small) and the distance between samples in different classes (expected to be large). Thus, the Fisher score for each variable is given by [32]:

$$f_{score} = \frac{\sum_{k=1}^{K} n^k (\mu_j^k - \mu_j)^2}{(\sigma_j)^2} \tag{8}$$

Where $(\sigma_j)^2 = \sum_{k=1}^{K} n^k (\sigma_j^k)^2$, μ_j and σ_j is the mean and the standard deviation of the *j-th* variable, μ_j^k and σ_j^k is the mean and the standard deviation of the *j-th* variable only for samples in the *k-th* security state, and n^k is the number of samples in the *k-th* security state.

Then, the P variables with larger scores are selected, but no specific criterion exits to choose P. Similarly to R^2, the f_{score} technique cannot detect and discharge redundant variables, it could discharge variables with low individual scores but with high scores when combined with other variables, and the number of selected variables P is not automatically determined.

Correlation-based filter selection (CFS): This technique simultaneously assesses correlations between variables with the class labels, as well as correlations between variables, and automatically selects the P variables with the highest discriminative power. CFS is based on symmetrical uncertainty (SU), which is defined as [23]:

$$SU(X, Y) = 2\left[\frac{H(X) - H(X|Y)}{H(X)H(Y)}\right] \tag{9}$$

Where SU(X, Y) is the SU between the random variables X and Y, $H(x) = -\sum_i P(x_i) log_2(P(x_i))$ is the entropy of X and $H(X|Y) = -\sum_i P(y_i) \sum_i P(x_i|y_i) log_2(P(x_i|y_i))$ is the conditional entropy of variable X given variable Y [33]. However, SU assumes that the random variables X and Y are categorical (i.e., qualitative variables or attribute variables), but the variables that represent the measurements from PMU are continuous.

This problem can be solved by applying the multi-interval discretization method [34]. Then, the automatic selection of variables is performed in two steps. First, a subset of relevant features (i.e., features that are correlated to the class) is selected by choosing all features for which SU(X, Y) > δ, where δ is determined the minimum required relevance to be analyzed for redundancy and is heuristically chosen (in this study, δ = 0 was chosen). Second, redundant variables are detected and discarded if they form a Markov blanket with any other variable, i.e., variable x_i is an approximate Markov blanket for variable x_j, with respect to class label y, if SU(x_i, y) ≥ SU(x_j, y) and SU(x_i, x_j) ≥ SU(x_j, y) [35]. The advantage of CFS over R^2 and f_{score} is that it automatically selects the P relevant and non-redundant variables.

Classifier. In the present methodology implementation was evaluated using two classification methods with different characteristics. Naive Bayes (NB): Naive Bayes is a probabilistic classifier based on the Bayes theorem which selects the class with the highest probability, assuming that the variables in the dataset are independent [25, 36], Hence the name *naive*. Given the vector of variables $x = [x_1, x_2, \ldots, x_M]$ and the complete set of classes or security states $C = \{C_1, C_2, \ldots, C_K\}$, the posterior probability of each class given the observed variables $p(C_k|x), k = 1, 2, \ldots, K$ according to the Bayes theorem is [24]:

$$p(C_k|\boldsymbol{x}) = \frac{p(\boldsymbol{x}|C_k)p(C_k)}{p(\boldsymbol{x})} \tag{10}$$

Notice that to compute the posterior probability $p(C_k|\boldsymbol{x})$, the probability of the variables given the class $p(\boldsymbol{x}|C_k)$ is needed. However, if the dimension of the variable vector $\boldsymbol{x}$ is high, it is necessary to have an extremely large amount of data samples. In consequence, the computation of $p(\boldsymbol{x}|C_k)$ is highly expensive. In order to simplify the computation of this join probability, NB assumes that all variables in $\boldsymbol{x}$ are independent. Thus, the joint probability reduces to:

$$p(x_1, x_2, \ldots, x_M|C_k) = \prod_{j=1}^{M} p(x_j|C_k) \tag{11}$$

Thus, the NB output is simply computed as:

$$y = \underset{C_k \in C}{argmax}\, p(C_k) \prod_{j=1}^{M} p(x_j|C_k) \tag{12}$$

Where y is the security state or class label that presents the maximum probability among all classes. The advantage provided by the NB classifier is its simplicity. Given a training dataset $(X; y); i = 1, \ldots, N$ we only need to estimate the parameters of the probability distribution for each variable in each class (i.e., the mean and standard deviation of a normal distribution), then, when a new sample x is available, the posterior probabilities for each class is computed and subsequently the class with the highest posterior probability is selected.

Support Vector Machine (SVM): The goal in this classifier is to obtain a separation hyperplane, possibly in a higher dimensional and non-linear representation of the variables, which discriminates data into different classes. To do so, the method maximizes the separating margins between the hyperplane and the so-called support vectors i.e., the nearest data points in each class [37]. Thus, the hyperplane is computed through the solution of the following optimization problem:

$$\min_{W,\xi} \frac{1}{2} \left\| \boldsymbol{w}^{\boldsymbol{T}} \boldsymbol{w} \right\|^2 + C \sum_{i=1}^{N} \xi_i \tag{13}$$

Subject to,

$$y_i\left(\boldsymbol{w}^{\boldsymbol{T}}\boldsymbol{X}_{\boldsymbol{i}}+b\right)\geq 1-\xi_i,\xi_i\geq 0,\forall i=1,\ldots,N \tag{14}$$

Where $\boldsymbol{w}$ and b are the weight vector and the bias term of the hyperplane, respectively, ξ_i are slack variables introduced because an optimal separating hyperplane might not exist when there is overlap between classes, N is the number of training samples and C is a penalty term for misclassification that determines a trade o between margin width and training error [38].

In this case, the separating hyperplane is linear. However, non-linear decision boundaries can be created through a kernel function applied to the observed data $K(x, x_i)$ [39]. Here, we used as kernel the radial basis function (RBF) which is given by:

$$K(x,x_i)=e^{-\frac{\|x-x_i\|^2}{2\sigma^2}} \tag{15}$$

Where σ determines width of the kernel. In the SVM with RBF as kernel, the parameters C and σ need to be defined. In this work, they were optimized through 5-fold cross validation using exclusively training data [39].

5 Simulations and Results

The dataset generated from the IEEE 39-bus test system shown in Fig. 4 that is presented in Eq. 6 has $N = 18764$ samples, $M = 39$ voltage variables and $K = 4$ security states. For this dataset, *13858* samples belong to normal, 1286 to alertness, *1336* to emergency state, and *12284* to in extremis. In this case problems may arise when a classification algorithm is trained. To examine the effect of the balanced samples, a balanced dataset was randomly selected from the original one, where the number of samples for each class is *1286* or total number of samples is $N = 5144$.

Different experiments are performed, combining the use of each variables selection technique with each classifier. This procedure was applied to both dataset, balanced and unbalanced. In order to measure the classification ability of each classifier accompanied by each variables selection technique a 10-fold cross validation process was carried out [40]. Considering the variable analysis techniques used, different variables were selected in each case, Table 2(A) and (B) show the selected variables for unbalanced and balanced dataset respectively. Classification accuracy results are presented in Table 3(A) and (B) for balanced and unbalanced dataset respectively, where the accuracy is defined as the percentage of correct classification.

In these results, the number of variables selected by the R^2 and F_{scores} methods is $P = 10$. Note that this number was imposed prior to simply accept of 25% of the $M = 39$ variables. However, the number of variables selected by the CFS methods is $P = 6$, which was automatically and is less than 10 selected for other methods.

Comparing box plots in Fig. 5, which explain different classifiers accuracy, the median in the box plot shows better predictability when CFS and all features are used independently of the classifier. However, if all features are used in classification

process prediction the time is longer, and comparatively the accuracy between CFS and all features have the same order.

In these results, the number of variables selected by the R^2 and F_{scores} methods is $P = 10$. Note that this number was imposed prior to simply accept of 25% of the $M = 39$ variables. However, the number of variables selected by the CFS methods is $P = 6$, which was automatically and is less than 10 selected for other methods. Comparing box plots in Fig. 5, which explain different classifiers accuracy, the median in the box plot shows better predictability when CFS and all features are used independently of the classifier. Observing Table 3 and in contrast with Fig. 5, it can be concluded that the classifier with better accuracy is SVM. The confusion matrices shown in Tables 4 and 5 explain the percentage of classification achieved by each class respect to all classes. From the confusion matrices, it is evident that balanced datasets allow better prediction between classes than unbalanced datasets. Moreover, if the CFS

Table 2. Variables selection results using 10-fold cross validation

(a) Unbalanced data		
Feature selection algorithm	# of selected nodes	Selected nodes
R^2	10	5, 11, 12, 19, 20, 22, 23, 25, 29, 13
Fisher score	10	1, 2, 3, 6, 9, 16, 21, 24, 26, 28
CFE	6	6, 7, 9, 10, 20, 25
(b) Balanced data		
Feature selection algorithm	# of selected nodes	Selected nodes
R^2	10	4, 17, 19, 20, 22, 23, 25, 29, 2, 12
Fisher score	10	1, 3, 6, 7, 8, 9, 13, 26, 28, 24
CFE	6	5, 6, 7, 9, 14, 25

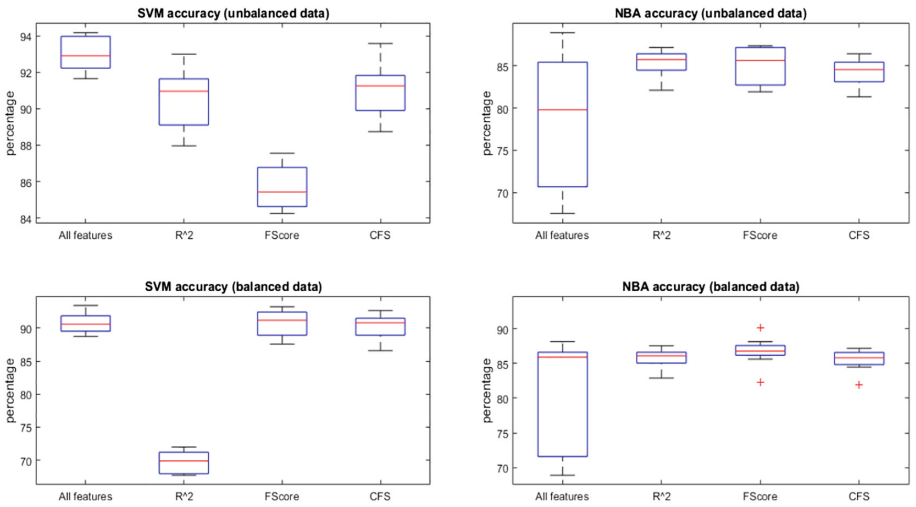

Fig. 5. Classifiers accuracy box plot

Table 3. Classification results using 10-fold cross validation

(a) Unbalanced data			
Feature selection algorithm	# of selected features	Naïve Bayes accuracy	SVM accuracy
None	39	78.91%	92.96%
R^2	10	85.30%	90.85%
Fisher score	10	85.09%	85.61%
CFS	6	84.29%	91.10%
(b) Balanced data			
Feature selection algorithm	# of selected features	Naïve Bayes accuracy	SVM accuracy
None	39	81.76%	90.79%
R^2	10	85.65%	69.83%
Fisher score	10	86.72%	90.73%
CFS	6	85.46%	91.16%

Table 4. Cconfusion matrix of svm classification (unbalanced data). the entry at the i-th row and the j-th column represents the probability of classifying an state from ci as an state in cj. one of the benefits of the confusion matrices is that allow to find if the model is confusing two classes.

(a) All Features

	C_1	C_2	C_3	C_4
C_1	0.997	0	0.003	0
C_2	1	0	0	0
C_3	0	0	1	0
C_4	0	0	0	1

(b) R2

	C_1	C_2	C_3	C_4
C_1	1	0	0	0
C_2	1	0	0	0
C_3	0.313	0	0.687	0
C_4	0	0	0	1

(c) Fisher Score

	C_1	C_2	C_3	C_4
C_1	1	0	0	0
C_2	1	0	0	0
C_3	1	0	0	0
C_4	0	0	0	1

(d) CFS

	C_1	C_2	C_3	C_4
C_1	1	0	0	0
C_2	1	0	0	0
C_3	0.246	0	0.754	0
C_4	0	0	0	1

Table 5. Confusion matrix of SVM classification (balanced data)

(a) All Features

	C_1	C_2	C_3	C_4
C_1	0.668	0.325	0.006	0
C_2	0.037	0.963	0	0
C_3	0	0	1	0
C_4	0	0	0	1

(b) R2

	C_1	C_2	C_3	C_4
C_1	0.496	0.496	0.008	0
C_2	0.500	0.500	0	0
C_3	0.052	0.013	0.935	0
C_4	0.020	0	0.066	0.914

(c) Fisher Score

	C_1	C_2	C_3	C_4
C_1	0.693	0.289	0.018	0
C_2	0.059	0.941	0	0
C_3	0	0	0	0
C_4	0	0	0	1

(d) CFS

	C_1	C_2	C_3	C_4
C_1	1	0	0	0
C_2	1	0	0	0
C_3	0.246	0	0.754	0
C_4	0	0	0	1

technique is selected then (Figs. 4(d) and 5(d)) the 64.7% and 96.3% of normal and alert states are successfully predicted, while 100% of states of emergency and in extremis can be successfully predicted. Even if the first two classes were not accurately classified, the most risk situations were classified with the 100% of accuracy, which is an important advance in security state estimation.

5.1 Conclusions

An alternative approach to achieve dimension reduction in the original set of variables is to perform feature extraction based on linear or non-linear transformations [41]. Based on this theory the complete m-dimensional set of independent variables is transformed to a q-dimensional reduced set of independent variables [41]. However, with this new data it is very difficult to find a physical meaning of the new variables. This is undesirable in power systems because some voltage buses in a power grid are more vulnerable or unstable due to important loads or generators attached to it. Therefore, if bus voltage data is combined or transformed, the distinctive patterns that show abnormal behaviors may disappear.

In this paper a supervised learning framework for power systems security-state estimation was proposed. The contribution suggests an open methodology to build security status models of the power grid, considering the stability problems, where other stability issues can be included to increase the estimation robustness. The methodology is proposed to be applied using historical measures of power systems, which must be structured in a dataset that allows the metrics calculation for each stability condition, and then building a model that operates with online measurable signals through the PMUs, to estimate security status of the power grid. To explain the methodology application, the IEEE 39-bus test system was used, detailing each steps required to find a reliable model for security state estimation. The dataset was obtained through power flows simulation taking into account multiple load conditions and contingencies. Two datasets were taken into account, balanced and unbalanced, for which models were constructed, tested and validated, showing that balanced dataset is more promise for each assigned class (normal, alert, emergency, in extremis). The obtained models conformed great predictability especially in emergency or in extreme conditions, which is priority for power system monitoring. The obtained results show that machine learning techniques may be successful for power system security on-line monitoring. Further research is required to improve the robustness of the security models to be applied in a real power grid. The methodology used in this work is applicable to real processes such as those shown in [42] and may be used for state estimation.

References

1. Avalos, J.: Analysis and application of optimization techniques to power system security and electricity markets, Ph.D. thesis, University of Waterloo (2008)
2. Jian-Hua, Z., Guohua, Z., Manyin, D., Jingyan, Y.: Indices system and methods for urban power grid security assessment. In: Proceedings of APPEEC 2009, pp. 1–5 (2009)

3. Sunitha, R., Kumar, S.K., Mathew, A.T.: Online static security assessment module using artificial neural networks. IEEE Trans. Power Syst. **28**(4), 4328–4335 (2013). https://doi.org/10.1109/tpwrs.2013.2267557
4. Zhao, F., Guo, Q., Sun, H., Wang, M., Wang, B.: Power system online security operational trend analysis and simulation results. In: 2013 IEEE Power and Energy Society General Meeting (PES), pp. 1–15 (2013). https://doi.org/10.1109/pesmg.2013.6672773
5. Sun, H., et al.: Automatic learning of fine operating rules for online power system security control. IEEE Trans. Neural Netw. Learn. Syst. **PP**(99) (2015). https://doi.org/10.1109/tnnls.2015.2390621
6. Wang, J., et al.: Online probabilistic security risk assessment implementation at china southern power grid towards smart control center. In: 2015 Seventh Annual IEEE Green Technologies Conference (GreenTech), pp. 74–81. https://doi.org/10.1109/greentech.2015.18
7. Beiraghi, M., Ranjbar, A.M.: Online voltage security assessment based on wide-area measurements. IEEE Trans. Power Deliv. **28**(2), 989–997 (2013). https://doi.org/10.1109/TPWRD.2013.2247426
8. Geeganage, J., Annakkage, U.D., Weekes, T., Archer, B.A.: Application of energy-based power system features for dynamic security assessment. IEEE Trans. Power Syst. **30**(4), 1957–1965 (2015). https://doi.org/10.1109/TPWRS.2014.2353048
9. Chakrabarti, A., Kothari, D., de Abhinandan, Mukhopadhyay, A.: An Introduction to Reactive Power Control and Voltage Stability in Power Transmission Systems. PHI Learning Private Limited (2010)
10. Chychykina, I., Styczynski, Z.A., Heyde, C.O., Krebs, R.: Power system instability prevention and remedial measures with online dynamic security assessment. In: 2015 IEEE Eindhoven PowerTech, pp. 1–5 (2015). https://doi.org/10.1109/ptc.2015.7232303
11. Perez Londono, S., Rodriguez Garcia, L., Lopez, Y.U.: Effects of doubly fed wind generators on voltage stability of power systems. In: Proceedings of the Sixth IEEE/PES Transmission and Distribution: Latin America Conference and Exposition (T&D-LA), pp. 1–6 (2012)
12. Taylor, C.: Power System Voltage Stability. McGraw-Hill, New York (1993)
13. Haque, M.: On-line monitoring of maximum permissible loading of a power system within voltage stability limits. IEE Proc. Gener. Transm. Distrib. **150**(1), 107–112 (2003)
14. Lee, C.Y., Tsai, S.-H., Wu, Y.-K.: A new approach to the assessment of steady-state voltage stability margins using the p-q-v curve. Int. J. Electr. Power Energy Syst. **32**(10), 1091–1098 (2010)
15. Momoh, J.A., Makarov, Y.V., Elfayoumy, M.: Expected angle stability margin. In: 1998 IEEE International Conference on Systems, Man, and Cybernetics, vol. 4, pp. 3753–3758 (1998). https://doi.org/10.1109/icsmc.1998.726671
16. Zhang, H., Liu, Y.: New index for frequency deviation security assessment. In: IPEC, 2010 Conference Proceedings, pp. 1031–1034 (2010). https://doi.org/10.1109/ipecon.2010.5696966
17. Sevilla, F.R.S., Vanfretti, L.: A small-signal stability index for power system dynamic impact assessment using time-domain simulations. In: 2014 IEEE PES General Meeting | Conference Exposition, pp. 1–5 (2014). https://doi.org/10.1109/pesgm.2014.6938842
18. Nguyen, T.B.: Dynamic security assessment of power systems using trajectory sensitivity approach, Ph.D. thesis, University of Illinois at Urbana-Champaign (2002)
19. Li, W.: Risk Assessment of power systems. In: IEEE Press Series on Power Engineering (2005)

20. Sanz, F.A., Ramirez, J.M., Correa, R.E.: Experimental design for a large power system vulnerability estimation. Electr. Power Syst. Res. **121**, 20–27 (2015). https://doi.org/10.1016/j.epsr.2014.11.026
21. Guyon, I.: An introduction to variable and feature selection. J. Mach. Learn. Res. **3**, 1157–1182 (2003)
22. Gu, W., Wan, Q.: Linearized voltage stability index for wide-area voltage monitoring and control. Int. J. Electr. Power Energy Syst. **32**(4), 333–336 (2010)
23. Hall, M.A.: Correlation-based feature selection for machine learning, Ph.D. thesis, University of Waikato (1999)
24. Bishop, C.M.: Pattern Recognition and Machine Learning. Information Science and Statistics. Springer, New York (2006)
25. Morash, V., Bai, O., Furlani, S., Lin, P., Hallett, M.: Classifying fEEGg signals preceding right hand, left hand, tongue, and right foot movements and motor imageries. Clin. Neurophysiol. **119**(11), 2570–2578 (2008). https://doi.org/10.1016/j.clinph.2008.08.013
26. Guyon, I., Gunn, S., Nikravesh, M., Zadeh, L.A.: Feature Extraction: Foundations and Applications. Studies in Fuzziness and Soft Computing. Springer, New York (2006). https://doi.org/10.1007/978-3-540-35488-8
27. Jensen, C.A., et al.: Inversion of feedforward neural networks: algorithms and applications. Proc. IEEE **87**(9), 1536–1549 (1999). https://doi.org/10.1109/5.784232
28. Innah, H., Asfani, D.A., Hiyama, T.: Voltage stability assessment based on discrimination principle. In: 2012 22nd Australasian Universities Power Engineering Conference (AUPEC), pp. 1–6 (2012)
29. Lakkireddy, J., Rastgoufard, R., Leevongwat, I., Rastgoufard, P.: Steady state voltage stability enhancement using shunt and series facts devices. In: 2015 Clemson University Power Systems Conference (PSC), pp. 1–5 (2015). https://doi.org/10.1109/psc.2015.7101706
30. Kundur, P., Balu, N., Lauby, M.: Power System Stability and Control, EPRI Power System Engineering Series. McGraw-Hill, New York (1994)
31. Shi, Z., Huang, Y., Zhang, S.: Fisher score based naive Bayesian classifier. In: 2005 International Conference on Neural Networks and Brain, vol. 3, pp. 1616–1621 (2005)
32. Singh, B., Sankhwar, J.S., Vyas, O.P.: Optimization of feature selection method for high dimensional data using Fisher score and minimum spanning tree. In: 2014 Annual IEEE India Conference (INDICON), pp. 1–6 (2014)
33. Yu, L., Liu, H.: Efficient feature selection via analysis of relevance and redundancy. J. Mach. Learn. Res. **5**, 1205–1224 (2004)
34. Fayyad, U.M., Irani, K.B.: Multi-interval discretization of continuous-valued attributes for classification learning. In: IJCAI, pp. 1022–1029 (1993)
35. Chidlovskii, B., Lecerf, L.: Scalable feature selection for multi-class problems. In: Daelemans, W., Goethals, B., Morik, K. (eds.) ECML PKDD 2008. LNCS, vol. 5211, pp. 227–240. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-87479-9_33
36. Fukunaga, K.: Introduction to Statistical Pattern Recognition. Computer Science and Scientific Computing. Elsevier Science (2013)
37. Vapnik, V.N.: The Nature of Statistical Learning Theory. Springer, New York (1995). https://doi.org/10.1007/978-1-4757-2440-0
38. Scholkopf, B., Smola, A.J.: Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond, Adaptive Computation and Machine Learning. MIT Press, Cambridge (2002)
39. Burges, C.J.: A tutorial on support vector machines for pattern recognition. Data Min. Knowl. Discov. **2**, 121–167 (1998)

40. Bengio, Y., Grandvalet, Y.: Bias in estimating the variance of *K*-fold cross-validation. In: Duchesne, P., RÉMillard, B. (eds.) Statistical Modeling and Analysis for Complex Data Problems, pp. 75–95. Springer, Boston (2005). https://doi.org/10.1007/0-387-24555-3_5
41. Li, J., et al.: Feature selection: a data perspective. CoRR abs/1601
42. Duarte, J., Garcia, J., Jimenez, J., Sanjuan, M.E., Bula, A., Gonzalez, J.: Autoignition control in spark-ignition engines using internal model control structure. J. Energy Resour. Technol. **139**(2) (2016). http://energyresources.asmedigitalcollection.asme.org/article.aspx?articleid=2532521