FORECASTING TRANSIENT STABILITY OF POWER SYSTEM BY AN ENSEMBLE CLASSIFIER

Nguyen Ngoc Au[1], Le Trong Nghia[1], Quyen Huy Anh[1], Phan Thi Thanh Binh[2]
[1]HCMC University of Technical Education, Ho Chi Minh City, Vietnam
[2]HCMC University of Technology, Ho Chi Minh City, Vietnam

Received 12/11/2018, Peer-reviewed 3/12/2018, Accepted for publication 03/01/2019

ABSTRACT

A large oscillation caused by faults leads power system to instability state. This makes fast forecast a necessity to drive power system into stability state, avoid the risk of blackouts. In recent years, an ensemble classifier has been emerged as a promising approach to enable online transient stable forecast (TSF). The paper proposed an ensemble classifier (EC) that is combined by parallel single classifiers. The single classifiers can compensate for the others by combining in parallel. Then, the EC can improve classification accuracy. The paper proposed the use of Multi-layer Perceptron Networks (MLPN) to build EC. The study is tested on IEEE 39-bus power system network.

Keywords: *Transient stability Forecast; Feature Selection; Power System; Neural Networks; Ensemble classifier.*

1. INTRODUCTION

Nowadays, power systems are forced to operate under highly stressed operating conditions closer to their stability limits due to the rapid growth of electricity demand. The power system operates in the event of unusual problems. Once occurred, it can harm the power system stability and may result in blackouts. To protect a power system from the blackouts, TSF needs to be executed to identify the stable state. Once unstable is identified, emergency control action must be implemented to drive the power system into re-stable state.

In recent years, artificial neural networks (ANN) has been emerged as a promising approach to enable online TSF [1]–[4]. While previous works mainly concentrate on a single classifier model [5], [6], this paper focuses on the model of parallel single classifiers. The single classifiers can compensate for the others by parallel combination. Then, the EC can improve the classification accuracy [7].

In this paper, in data process such as feature selection, K-nearest neighbor classifier (1-NNC, K=1) firstly guides to find feature subset. The 1-NNC is chosen because of its fast calculation and simplicity. With found results, the EC is built by MLPN. The study was implemented on IEEE 39-bus power system using simulation software Power-World 18. The computation was implemented by the Mat-lab R2014a software.

2. METHOD

2.1 Feature and sample

To build EC for TSF, initial features and samples are required. A large number of samples are generated through off-line simulation and the stable status was evaluated for each fault under study. Data for each bus or line fault occurring in the test systems are recorded in which samples of data are kept in a database.

The input features are the vector of system state parameters that characterize the current system state. Fault-on features are variables that characterize at fault-on state of power system occur such as changes in nodal powers, in power flows in a transmission line, and voltage drops in the nodes at the instance of fault.

The output features represent the stable conditions of the power system. The output variables are assigned to label binary variable y {1, 0}. Class 1 {1} is stable class and class 2 {0} is unstable class.

Simulating observation results, if the angle of the relative rotor generators is larger than 180^0 then the system is 'Unstable', and less than 180^0 then the system is 'Stable'. The corresponding data will be put into an unstable and stable class.

2.2 Feature selection criterion

Fisher criterion: This is a criterion that was applied in many works with the ranking feature selection method (RFSM) [8]. By evaluating the criterion of the features as (1), the features are ranked by ordering the best of them and selecting for good features. The bigger feature F is the more important one.

$$F = \frac{|m_1 - m_2|^2}{\sigma_1^2 + \sigma_2^2} \quad (1)$$

Where mi is the mean of class C and σ_i^2 is the variance of class C_i.

Scatter matrices criterion: According to the scatter matrices (SM) theory, the separation of classes of feature subset is based on (2) [9]. The bigger the value of J is the higher the separation of classes of feature subset will be. Where S_m is the dataset covariance matrix; S_w is within-class scatter matrix.

$$J = \mathrm{Tr}\{S_w^{-1} S_m\} \quad (2)$$

In this paper, the Sequential Forward Floating Selection (SFFS) with SM criterion and RFSM with Fisher criterion were applied for feature selection. These were presented in and applied in our previous work [10], [11].

2.3 Proposed Approaches

In our proposed the method we are designing the EC that includes multiple parallel classifiers in order to improve the accuracy. The architecture of the proposed EC is as Figure 1.

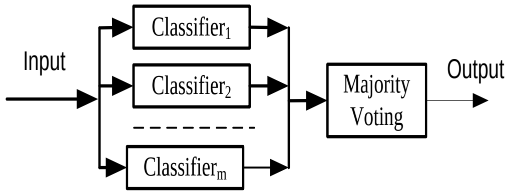

Figure 1. *The architecture of proposed EC*

By learning from a dynamic stability database, the nonlinear relation between input and output can be captured and memorized in the EC. In the online mode, the system's stability status can be identified in real time as soon as the EC receives the input signal. Necessity is to select a compact data set that is the representative of all data set with the aim of reducing memory, saving cost.

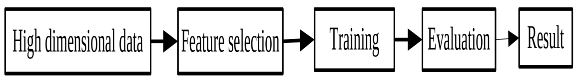

Figure 2. *The design EC process*

To build EC, feature selection is implemented. Feature selection is actually dimensional space reduction. The design EC process is proposed as Figure 2. This process consists of the following steps: Feature selection, Training, and Evaluation.

2.3.1 Feature selection

The EC-based TSF can be formulated as a mapping $y_i = f(x_i)$ after learning from a dynamic stability database Feature selection The EC-based TSF can be formulated as a mapping $y_i = f(x_i)$ after learning from a dynamic stability database $D = \{x_i, y_i\}_{i=1}^{n}$. Where: x_i is feature; it is n-dimensional input vector that characterizes the system operating state; and y_i is output vector. The feature subset selection consists of selecting a d dimensional feature vector z, d < n; The d selected features to represent the original one in a new knowledge base Feature selection.

2.3.2 Training

The classifiers learn the input and output relationship to build knowledge. Classification is that data are classified into labeled classes basically. The classes have labeled as ' 1: Stable ', and '0: Unstable'. Each classifier assigns each data point in the entire space of data into one of two classes. The EC has multiple classifiers that are a combination of parallel by training them and fusing their decisions to produce the final decision.

Each time a neural network is trained can result in a different solution due to different initial weight and bias values and different divisions of data into training, validation, and test sets [8], [12]. As a result, different neural networks trained on the same problem can give different outputs for the same input. Larger numbers of neurons in the hidden layer give the network more flexibility because the network has more parameters it can optimize. Basing on those, the paper proposed a new combine model of MLPNs for TSF problem. As figure 1, the model has *m* single MLPN classifiers (SMLPNC) that are connected in parallel, all of that have the same input features. The training steps are proposed as follow:

Step 1. Data generating. A large number of samples are generated through off-line simulation and the stable status is evaluated for each fault under study. Data for each bus or line fault occurring in the test systems are recorded in which samples of data are kept in a database. The dataset is normalized before training. The data are split into training set (Dtrain) and testing set (Dtest) randomly.

Step 2. Training. The selected range of hidden neural numbers is [N1,N2]. Each of the value in the range, the single MLPN is trained by 10 times. The accuracy rate (AR) is calculated as Eq. (3). The ARs are ranked in descending order.

Step 3. EC construction. Combine SMLPNCs in parallel from the top to the bottom of ARs sequentially. The paper used majority voting to calculate the final out of EC as Eq.(6).

Step 4. EC Selection. A selected number of SMLPNC is that EC yields the highest accuracy rate in the observed range.

2.3.3 Evaluation

The overall accuracy rate (AR) is calculated by Eq.(3).

$$AR(\%) = \frac{n}{N}.100 \quad (3)$$

Where n is a number of the correct test samples; N is a total number of the test samples.

Due to the computing errors, a threshold needs to be defined as the decision borderline for the neural networks (NN) to classify. So, a typical rule for a neural network classifier to classify a new sample can be as (4).

$$\begin{cases} If \quad y \le 0.5 \quad then \quad y = 0, \text{'Unstable'} \\ If \quad y > 0.5 \quad then \quad y = 1, \text{'Stable'} \end{cases} \quad (4)$$

In the proposed EC as Fig. 1, a straightforward voting technique is majority voting [13]. It considers only the most likely class provided by each classifier and chooses the most frequent class label among this crisp output set [13]. Experimentally, the number of classifiers used for voting is usually odd.

The label outputs of classifiers are given as c-dimensional binary vectors $[d_{i,1}, d_{i,2},\ldots,d_{i,c}]^T \in \{0,1\}^c$, i=1,2,…,L. Where d_{ij}='1' if D_i labels x in ω_j, and '0' otherwise. The plurality vote will result in an ensemble decision for class ω_k if as (5).

$$\sum_{t=1}^{L} d_{i,k}(x) = \max_{j=1,..c} \sum_{i=1}^{L} d_{i,j} \quad (5)$$

The number of classifiers, L, is odd. According to (5), the majority vote will give an accurate class label as rule (6).

$$(\lfloor L_T/2 \rfloor + 1) > 0.5*L \quad (6)$$

Where $\lfloor a \rfloor$ denotes the floor function, that is, the nearest integer smaller than a. L_T is the numbers of correct answers.

3. RESULTS – DISCUSSION

3.1 Features and samples

The study was tested on the IEEE 39-bus scheme. It includes 39 buses, 19 loads, 10 generators. The diagram IEEE 39-bus scheme is shown as Fig. 5. This is a well-known single-line diagram. It was used in many published works [5], [6], [8]. The off-line simulation was implemented to collect data for training. Load levels are (20,30,…,120)% normal load. The setting fault clearing time (FCT) is 50ms. In this paper, all kinds of faults such as single phase to ground, double phase to ground, three phases to ground and phase-to-phase short-circuit are considered. Faults are tested in any buses and in each of 5% distances of long transmission lines of the test systems. For each of the considered load samples, the generator samples have been got accordingly by running optimal power flow (OPF) tool of Power-World software.

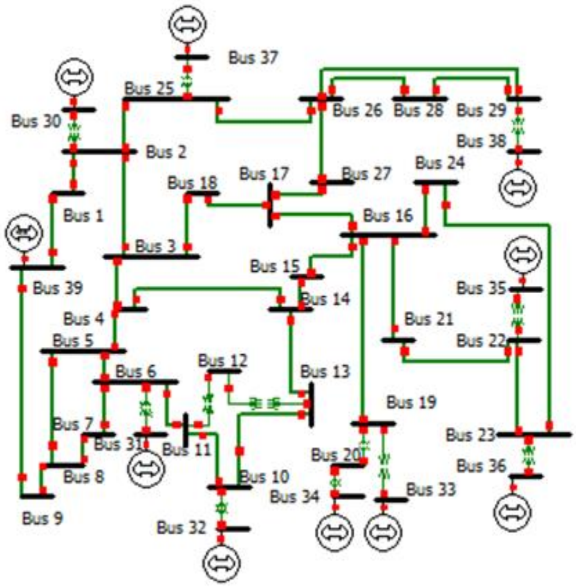

Figure 3. *The IEEE 39-bus diagram*

The input and output feature are $x\{delV_{bus}, delP_{Load}, delP_{flow}\}$ and $y\{1,0\}$. Total of input features is 104, $x\{104(39+19+46)\}$. The number of output feature is one, $y\{1,0\}$. From simulating results, there are 3551 samples that include 2649 S samples and 902 U samples, $D(S,U)=D(2649,902)$.

3.2 Feature selection results

The 1-NNC firstly used as a classifier for evaluating accurate classification because of its simplicity. The dataset was randomly divided into ten sub-datasets. Each training sub-dataset had 2384 stable samples, 812 unstable samples. Each testing sub-dataset had 265 stable samples, 90 unstable samples. It is a 10-fold cross-validation. Average evaluation result was executed by ten times on 10-fold. Results of the selected feature with SFFS and RFSM are shown in Table 1. In that, testing accurate rate of SFFS is higher than that of RFSM. So, there are the feature set of 15 selected features with SFFS.

Table 1. *The highest average classification rate of feature selection methods.*

Algorithm	d(feature)	AR(%)
RFSM	17	95.32
SFFS	15	95.63
-	104	95.60

3.3 Ensemble Classifier

Multilayer Perceptron Network Classifier: The MLPNC is a general model used in problems where the relationship between predictor and response variable is high non-linear. The MLPNC is a feed forward neural network model mapping sets of input data onto a set of appropriate output. The basic structure of a MLPN consists of three layers that are input, hidden and output layer. In this work, there are 15 input features and one output feature. The activation function log-sig was used in this study for the hidden layer. The output layer consists of a single neuron with activation function pureline. Levenberg-Marquardt (trainlm) optimization based for weight and bias updating algorithm is selected. The Levenberg-Marquardt algorithm was used and recommended in many works by fast computation and high accuracy [12]. The parameters setting was used during the training process for all experiment: performance goal =1e-5, the maximum number of epochs to train =1e3, and the others were fixed with their default values. Numbers of neurons for the hidden (N) layer were selected by trial and error experiments [14]. The results of the training steps:

Step 1. Data generating. Data have 15 features. The original data, D(2649,902), is randomly split by 75% for training, $D_{train}(1987,676)$, and 25% for testing,

$D_{test}(662,226)$. There are 2663 samples as the training set and 888 samples as the testing set.

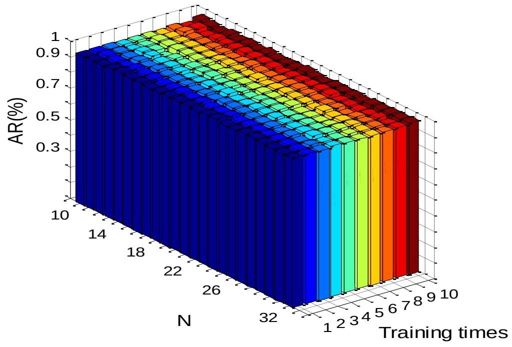

Figure 4. *The testing accuracy rate for all training times*

Step 2. Training. The selected range of hidden neural numbers is [10,32]. Each of the value in the range, the single MLPN classifier (MLPNC) is trained by 10 times. The accuracy rates (AR) are shown as Fig.4. The ARs are ranked in descending order as Fig 5.

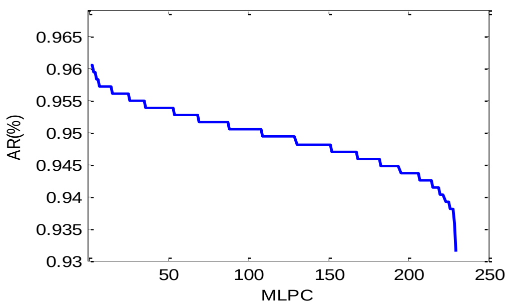

Figure 5. *The ranked ARs in descending order.*

Step 3. EC model building. Combine SMLPNCs in parallel from the top to the bottom of ARs sequentially. Applying the majority voting, the output results of the model are shown by Fig. 6.

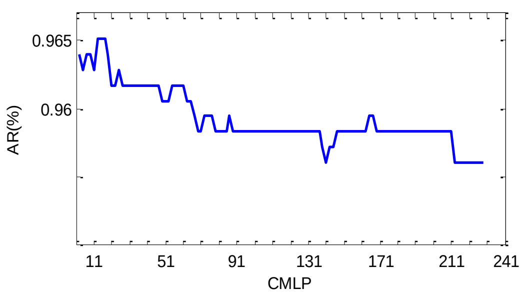

Figure 6. *The accuracy rate for EC model building*

Step 4. EC model Selection. In according to Fig. 6, the highest AR is 96.51% at 13 connected SMLPNCs in parallel. So, the proposed EC has 13 SMLPNCs and the ARs as Fig. 7, in which the 14th is the AR of the proposed EC and the others are the ARs of SMLPNCs.

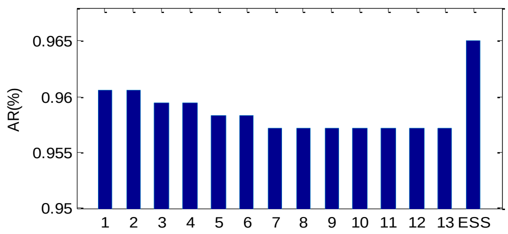

Figure 7. *The AR of the proposed EC*

3.4 Discussion

Table 1, with 1-NNC, the AR was 95.63% with the 15 selected features of the SFFS, 95.32% with the 17 selected features of the RFSM. The SFFS yielded higher AR than RFSM. Comparing with all features set, the average classification rate of the 15 selected feature subset of the SFFS increased by 0.03%, while the number of features decreased by 85.57%. The SFFS removed the abundant features or weak and noisy features, while recognition accuracy is still not decreased.

According to Figure 4 and 5, the highest accuracy classification gets 96.01% within 230 training times. In Figure 6 and 7, the proposed EC obtains 96.51%. It is shown that the proposed EC has improved accuracy classification by 0.5%.

4. CONCLUSIONS

Feature selection is aimed at reducing the number of features, reducing memory, and saving cost. This paper proposed the use of the SFFS and RFMS for feature selection in power system transient stability classification. The results showed that the SFFS achieved effective reduction features and recognition accuracy of SFFS higher than that of RFMS.

The paper is proposed the design process and the procedure SMLPNCs training for the

EC. It can be implemented in offline training, and online application. The proposed EC is tested on the IEEE 39-bus system and compared with single classifier. The simulation results show that the EC can effectively increase the classification accuracy over a single classifier.

ACKNOWLEDGEMENT

The authors acknowledge the support from the Electrical Power System and Renewable Lab of Ho Chi Minh City University of Technology and Education, Vietnam.

REFERENCES

[1] S. Zarrabian, R. Belkacemi, and A. A. Babalola, “Intelligent mitigation of blackout in real-time microgrids: Neural network approach,” *Power Energy Conf. Illinois (PECI), 2016 IEEE*, 2016.

[2] Y. Zhou, J. Wu, L. Hao, L. Ji, and Z. Yu, “Transient Stability Prediction of Power Systems Using Post-disturbance Rotor Angle Trajectory Cluster Features,” *Electr. Power Components Syst.*, vol. 44, no. 17, pp. 1879–1891, 2016.

[3] Y. Zhang, T. Li, G. Na, G. Li, and Y. Li, “Optimized extreme learning machine for power system transient stability prediction using synchrophasors,” *Math. Probl. Eng.*, vol. 2015, p. 8, 2015.

[4] A. Hoballah and I. Erlich, “Transient stability assessment using ANN considering power system topology changes,” *2009 15th Int. Conf. Intell. Syst. Appl. to Power Syst. ISAP ’09*, 2009.

[5] A. M. A. Haidar, M. W. Mustafa, F. A. F. Ibrahim, and I. A. Ahmed, “Transient stability evaluation of electrical power system using generalized regression neural networks,” *Appl. Soft Comput. J.*, vol. 11, no. 4, pp. 3558–3570, 2011.

[6] S. Kalyani and K. S. Swarup, “Pattern analysis and classification for security evaluation in power networks,” *Int. J. Electr. Power Energy Syst.*, vol. 44, no. 1, pp. 547–560, 2013.

[7] R. Ebrahimpour, “Transient Stability Assessment of a Power System by Mixture of Experts,” vol. 2, no. 4, pp. 93–104, 2010.

[8] Z. Y. Dong, Z. Rui, and Y. Xu, “Feature selection for intelligent stability assessment of power systems,” in *2012 IEEE Power and Energy Society General Meeting*, 2012, pp. 1–7.

[9] A. R. Webb and K. D. Copsey, *Statistical Pattern Recognition*. 2011.

[10] N.N.Au, Q.H.Anh, and P.T.T.Binh, “Feature Subset Selection in Dynamic Stability Assessment Power System Using Artificial Neural Networks,” *Sci. Technol. Dev. Vol.18, No.K3*, 2015.

[11] N. A. Nguyen, T. N. Le, H. A. Quyen, B. P. T. Thanh, and T. B. Nguyen, “Hybrid Classifier Model for Dynamic Stability Prediction in Power System,” *Proc. - 2017 Int. Conf. Syst. Sci. Eng. ICSSE 2017*, vol. 2017, no. Icsse, pp. 144–147, 2017.

[12] M. H. Beale, M. T. Hagan, and H. B. Demuth, “Neural Network Toolbox ™ User ’ s Guide R 2014 a,” 2014.

[13] R. Polikar, “Ensemble Based Systems,” *IEEE Circuits Syst. Mag.*, vol. 6, no. 3, pp. 21–45, 2006.

[14] K. G. Sheela and S. N. Deepa, “Review on methods to fix number of hidden neurons in neural networks,” *Math. Probl. Eng. Hindawi Publ. Corp.*, p. 11 p, 2013.

Corresponding author:
Nguyen Ngoc Au
HCMC University of Technical Education, Ho Chi Minh City
Email: ngocau@hcmute.edu.vn