

BPSO&1-NN algorithm-based variable selection for power system stability identification

N. N. Au, V. V. Cuong, T. T. Giang, L. T. Nghia, T. V. Hien

Faculty of Electrical and Electronics Engineering, HCMC University of Technology and Education, Ho Chi Minh City, Viet Nam

Abstract— Due to the very high nonlinearity of the power system, traditional analytical methods take a lot of time to solve, causing delay in decision-making. Therefore, quickly detecting power system instability helps the control system to make timely decisions become the key factor to ensure stable operation of the power system. Power system stability identification encounters large data set size problem. The need is to select representative variables as input variables for the identifier. This paper proposes to apply wrapper method to select variables. In which, Binary Particle Swarm Optimization (BPSO) algorithm combines with K-NN ($K=1$) identifier to search for good set of variables. It is named BPSO&1-NN. Test results on IEEE 39-bus diagram show that the proposed method achieves the goal of reducing variables with high accuracy.

Keywords— Identification, Power system, BPSO, Variable selection, K-NN.

I. INTRODUCTION

Modern power systems suffer from operating pressure very close to a stable boundary limit, while power systems are always faced with abnormal agitation that can easily cause instability (Makarov, Reshetov, Stroeve, & Voropai, 2005). Due to the very high nonlinearity of the power system, traditional analytical methods take a lot of time to solve, causing delay in decision-making. Therefore, quickly detecting power system instability helps the control system to make timely decisions become the key factor to ensure stable operation of the power system.

In recent years, the identification method has been applied as an alternative to solving difficult problems that traditional methods of analysis cannot solve in terms of computational (Dong, Rui, & Xu, 2012; Mitra, Benidris, Nguyen, & Deb, 2017; Oliveira, Vieira, Bezerra, Martins, & Rodrigues, 2017; Zhang, Xu, Dong, Meng, & Xu, 2011). By learning the database, the nonlinear input/output relationship between the electrical system operating parameters and stability can be quickly calculated. However, if the identifier works fast, the number of input variables must be characteristic variables. Therefore, the selected variables must be characteristic, eliminate redundant variables and cause noise.

The selection of variables of previously published works mainly applies the filter method that has to go through many stages (Dong et al., 2012; Kalyani & Swarup, 2013; Swarup, 2008). With this method, the first is based on the

statistical criteria to filter out the good variables first, then ask the identifier to evaluate the accuracy to find the appropriate set of variables. This article recommends a wrapper method to select variables. In which, BPSO (Binary Particle Swarm Optimization) algorithm combined with K-NN ($K=1$) identifier to perform the process of searching for good set of variables. It is named BPSO&1-NN. This helps to reduce the step of finding the variable set compared to the filter method and improves the automation of the process of finding the variable set. Test results on IEEE 39-bus diagram show that the proposed method achieves the goal of variable reduction with high accuracy.

II. VARIABLE SELECTION

In the design stages of a stable diagnostic model of the power system using the identification method, the selection of characteristic variables is very important because it directly affects the diagnostic accuracy of the model. This also helps to reduce the number of measurement sensors, reducing computation time for the model. This step defines a specific set of variables that represent the database for training. These initial characteristic variables are the input variable representing the operating parameters of the power system and covering the operating status of the power system. The characteristic variable of the power system in transient mode or dynamic mode is the change in generator

capacity, change of load capacity, change of power on transmission lines, voltage drop at nodes, ... right at the time of the incident. The output variable is labeled binary. '0' is unstable and '1' is stable.

Filter method:

The Filter method applies the search algorithm with the statistical standard function such as Fisher, Divergence, ... to guide the search strategy for the best set of variables to achieve the target value (J) (Dong et al., 2012; Kalyani & Swarup, 2013; Swarup, 2008). The process of searching for the set of variables by the filter method is presented as Figure 1. This method shows that the result is a set of variables meeting statistical standards but does not show the identification accuracy and the number of necessary variables. Therefore, the filter method needs to add an identification accuracy assessment to select the last set of variables.

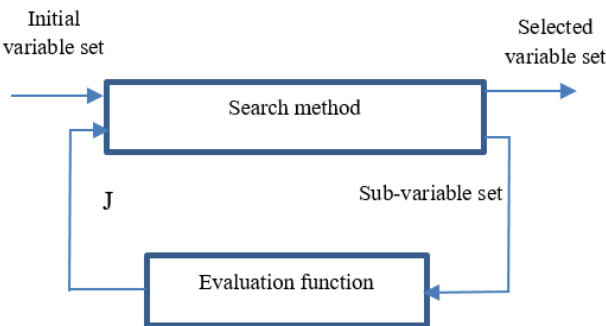


Fig. 1: Diagram of Filter method

Wrapper method:

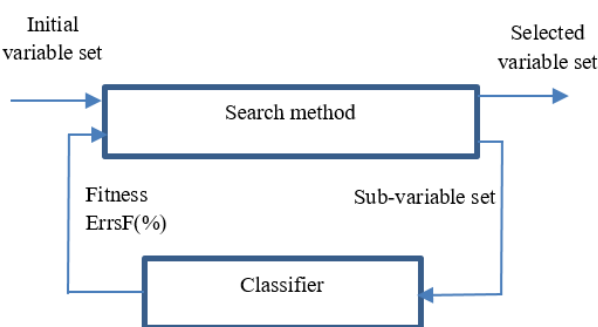


Fig. 2: Diagram of Wrapper method

The Wrapper method is presented as in Figure 2. However, this method differs from the Filter method in that the target function considering the identity accuracy. Thus, this method results in a set of variables to be selected with a specific identification accuracy.

The fitness function for selecting variables is presented in formulas (1), (2), and (3).

$$Fitness = \frac{ErrsF}{TotalF - sF} \tag{1}$$

$$AccsF(\%) = \frac{AccSample}{TotalSample}(\%) \tag{2}$$

$$ErrsF(\%) = 1 - AccsF \tag{3}$$

where:

Fitness is the objective function.

AccsF is the identification accuracy of the selected set of variables.

AccSample is the number of the correct identifier. TotalSample is the total number of samples.

TotalF: the total number of variables in the data set.

sF is the number of selected variables.

ErrsF is the identification error of the selected variable.

III. BPSO ALGORITHM IN VARIABLE SELECTION

PSO (Particle Swarm Optimization) is the optimal search algorithm proposed by Kennedy and Eberhart (Kennedy & Eberhart, 1995). In the PSO, each candidate answer of the problem is encoded as an instance that moves through the search space. The whole herd seeks the optimal solution by updating the position of each individual based on their own experience and on neighboring individuals.

Generally, the vector $x_i = (x_{i1}, x_{i2}, \dots, x_{iD})$ used in the PSO represents the position of the i th instance. The vector $v_i = (v_{i1}, v_{i2}, \dots, v_{iD})$ is used in the PSO to represent the velocity of the i th instance. D is the size of the search space. During the search, the best position for each previous individual was recorded as pbest. The best location of the herd is the gbest. The herd was randomly generated from the population. Finding the best solution by updating the velocity and position of each individual according to equations (4) and (5).

$$x_{id}^{t+1} = x_{id}^t + v_{id}^{t+1} \tag{4}$$

$$v_{id}^{t+1} = w * v_{id}^t + c_1 r_{1i} * (p_{id} - x_{id}^t) + c_2 r_{2i} * (p_{gd} - x_{id}^t) \tag{5}$$

Where: t is the t th iteration of the search process. d is the size in the search space, $d \in D$. c_1, c_2 are acceleration constants. r_{1i}, r_{2i} are random values, valid in the range $[0,1]$. p_{id} and p_{gd} represent pbest and gbest particles of size

d. w is the inertial weight. v_{id}^t is the velocity, limited to the maximum velocity v_{max} , $v_{id}^t \in [-v_{max}, v_{max}]$.

The original PSO algorithm applied to the problem of continuity. Kennedy and Eberhart developed the BPSO algorithm for the discrete problem (Kennedy & Eberhart, 1997). The velocity in BPSO represents the element that can take the value 1. Equation (4) is still used to update the velocity while x_{id} , p_{id} get the value 0 or 1. The function sigmoid $s(v_{id})$ is used to convert the value of v_{id} into a range of values (0,1). The BPSO updates each instance's position using equations (6) and (7).

$$x_{id} = \begin{cases} 1, & \text{if } rand() < s(v_{id}) \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

$$s(v_{id}) = \frac{1}{1 + e^{-v_{id}}} \quad (7)$$

The function $rand()$ is a random function whose value is distributed in (0,1).

BPSO

Begin

Data set, $kfold$; D : dimensionality of search space
 N : the population size; T : maximum iterations;

C_1, C_2, v_{max}, W

Randomly initialise the position and velocity of each particle;

while $t \leq T$

 evaluate fitness of each particle according to Equation (1);

for $i=1$ to N

 update the p_{best} of particle i ;
 update the g_{best} of particle i ;

end

for $i=1$ to N

for $d=1$ to D

 update the velocity of particle i according to Equation (4);
 update the position of particle i according to Equations (6) and (7);

end

end

end

calculate the classification error of the selected feature subset;

return the position of g_{best} (the selected feature subset);

return the classification error of the selected feature subset;

end

IV. RESULTS

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. 2. It was used in many published works. 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 (Glover, Sarma, & Overbye, 2012). 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.

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 1167 samples that include 834 S samples and 783 U samples. The K-NN (K=1) is used as a classifier for evaluating accurate classification because of its simplicity. Assessment of classification error is a cross-assessment method. It is a 10-fold cross-validation.

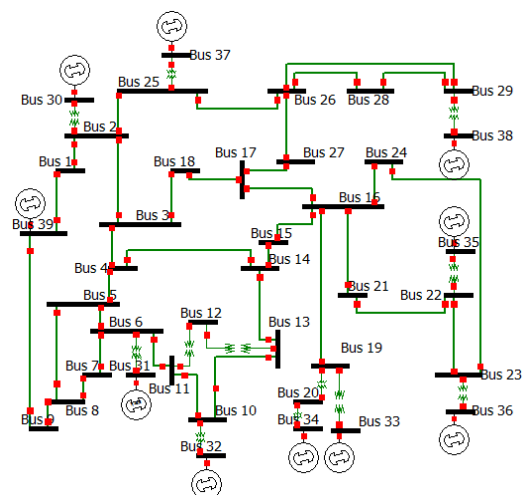


Fig. 2: The IEEE 39-bus diagram

Feature selection results:

The BPSO&1-NN algorithm works with different N values, namely 10, 20, 30, 40, and 50. The values of w are

0.4, 0.5, 0.6, 0.7, 0.8, and 0.9. The number of iterations is 100 for the program executions, $T=100$. The values of c_1 and c_2 are selected unchanged during program execution, $c_1=2$, $c_2=2$. The program is also executed on Matlab 2018a software. Identification tool, 1-NN, is supported from this software.

The convergence characteristics of BPSO&1-NN variable selection algorithm are shown as Fig. 3 to Fig. 8. The results of the selected number of variables, the value of the best fitness function and the test error are presented in Table 1.

Table 1. Best results during the implementation of BPSO program with $w = 0.9$, $N = 50$, the number of variables in the set of variables is 23, the identification error is 5.07%. Thus, the algorithm BPSO&1-1NN has reduced the number of variables from 104 to 23 variables. The number of variables is reduced by 78% while test accuracy reaches 94.93%. This is the accepted result in previously published (Abdelaziz & El-Dessouki, 2013; Amjady & Majedi, 2007; Dong et al., 2012; Haidar, Mustafa, Ibrahim, & Ahmed, 2011; Kalyani & Swarup, 2013; Karami & Esmaili, 2013).

Table 1. Results of BPSO algorithm

w	N	f	Best Fitness	ErrsF(%)
0.4	10	32	7.3868e-04	5,07
	20	32	7.2150e-04	5,07
	30	33	7.3166e-04	5,01
	40	30	6.7693e-04	5,13
	50	37	7.1073e-04	5,01
0.5	10	33	7.1424e-04	5,19
	20	32	6.9573e-04	5,26
	30	33	7.0553e-04	5,07
	40	34	7.2445e-04	5,07
	50	31	7.3703e-04	5,26
0.6	10	33	7.2295e-04	4,95
	20	32	7.2150e-04	5,19
	30	36	7.3666e-04	5,01
	40	33	6.5327e-04	4,89
	50	36	6.9119e-04	5,13
0.7	10	31	7.0315e-04	5,13
	20	35	7.1702e-04	5,26
	30	30	6.9364e-04	5,26
	40	29	7.0089e-04	5,13

	50	26	6.6600e-04	4,95
0.8	10	32	6.6138e-04	5,07
	20	30	6.6021e-04	4,89
	30	33	6.5327e-04	5,26
	40	30	6.6857e-04	5,07
	50	27	6.1843e-04	5,26
0.9	10	32	6.6996e-04	5,44
	20	27	6.2646e-04	5,38
	30	27	6.1040e-04	5,13
	40	23	5.9552e-04	5,13
	50	23	6.1079e-04	5,07

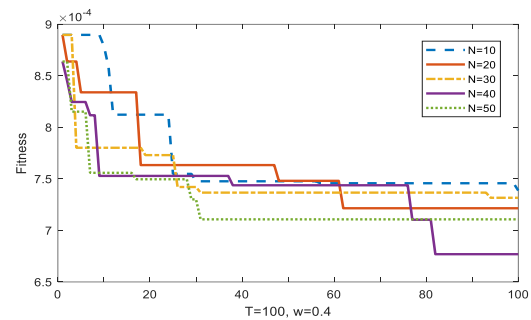


Fig. 3: Convergence characteristics of BPSO&1-NN, $w=0,4$

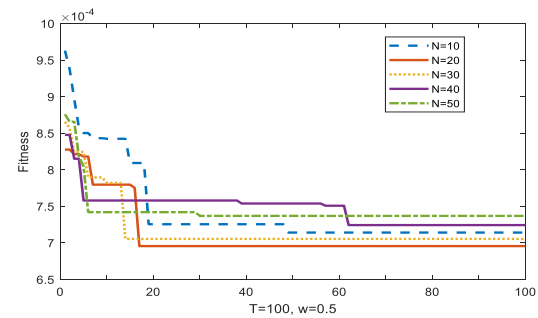


Fig. 4: Convergence characteristics of BPSO&1-NN, $w=0,5$

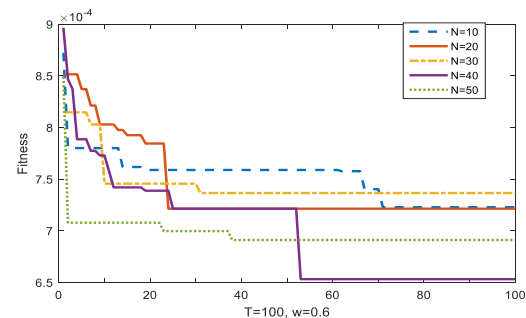


Fig. 5: Convergence characteristics of BPSO&1-NN, $w=0,6$

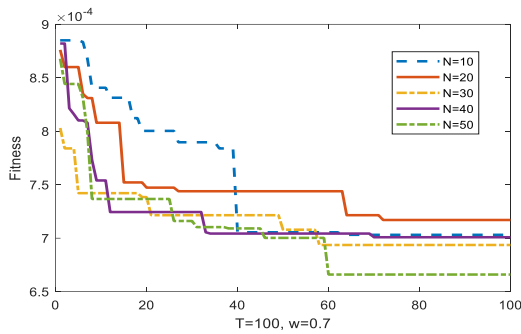


Fig. 6: Convergence characteristics of BPSO&1-NN,
 $w=0,7$

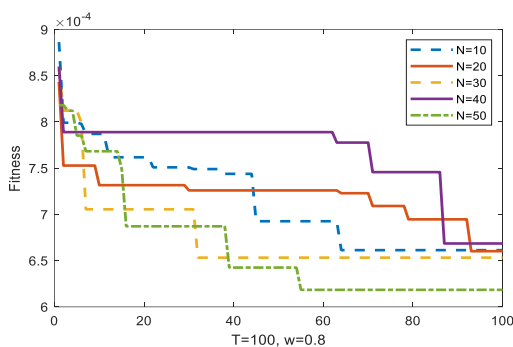


Fig. 7: Convergence characteristics of BPSO&1-NN,
 $w=0,8$

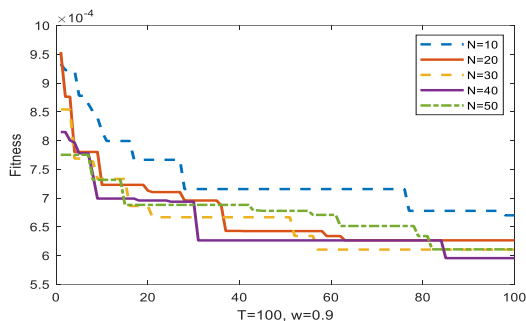


Fig. 8: Convergence characteristics of BPSO&1-NN,
 $w=0,9$

V. CONCLUSION

The paper successfully applied BPSO algorithm combined with 1-NN identifier in the variable selection approach to the problem of identifying stability of power system. Test results on IEEE-39 bus diagram showed that the number of variables decreased significantly, but the test accuracy still met expectations. This achieved result will be the driving force for the research and application of evolutionary algorithms to select variables for the problem of identifying stability of power systems.

ACKNOWLEDGEMENTS

This work belongs to the project grant No: T2020-26TĐ funded by Ho Chi Minh City University of Technology and Education, Vietnam.

REFERENCES

- [1] Abdelaziz, A. Y., & El-Dessouki, M. A. (2013). Transient Stability Assessment using Decision Trees and Fuzzy Logic Techniques. *International Journal of Intelligent Systems and Applications*, 5(10), 1–10. <https://doi.org/10.5815/ijisa.2013.10.01>
- [2] Amjady, N., & Majedi, S. F. (2007). Transient stability prediction by a hybrid intelligent system. *IEEE Transactions on Power Systems*, 22(3), 1275–1283. <https://doi.org/10.1109/TPWRS.2007.901667>
- [3] Dong, Z. Y., Rui, Z., & Xu, Y. (2012). Feature selection for intelligent stability assessment of power systems. *2012 IEEE Power and Energy Society General Meeting*, 1–7. <https://doi.org/10.1109/PESGM.2012.6344780>
- [4] Glover, J. D., Sarma, M. S., & Overbye, T. (2012). Power System Analysis and Design. In *Global Engineering: Christopher M.Shortt* (Fifth Edit). <https://doi.org/10.1145/633615.810652>
- [5] Haidar, A. M. A., Mustafa, M. W., Ibrahim, F. A. F., & Ahmed, I. A. (2011). Transient stability evaluation of electrical power system using generalized regression neural networks. *Applied Soft Computing Journal*, 11(4), 3558–3570. <https://doi.org/10.1016/j.asoc.2011.01.028>
- [6] Kalyani, S., & Swarup, K. S. (2013). Pattern analysis and classification for security evaluation in power networks. *International Journal of Electrical Power and Energy Systems*, 44(1), 547–560. <https://doi.org/10.1016/j.ijepes.2012.07.065>
- [7] Karami, A., & Esmaili, S. Z. (2013). Transient stability assessment of power systems described with detailed models using neural networks. *International Journal of Electrical Power and Energy Systems*, 45(1), 279–292. <https://doi.org/10.1016/j.ijepes.2012.08.071>
- [8] Kennedy, J., & Eberhart, R. (1995). Particle Swarm Optimization. *IEEE International Conference , Perth, WA, Australia*, (ISBN: 0-7803-2768-3), 1942–1948.
- [9] Kennedy, J., & Eberhart, R. C. (1997). Discrete binary version of the particle swarm algorithm. *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, 5, 4104–4108. <https://doi.org/10.1109/icsmc.1997.637339>
- [10] Makarov, Y. V., Reshetov, V. I., Stroev, V. A., & Voropai, N. I. (2005). Blackout prevention in the United States, Europe, and Russia. *Proceedings of the IEEE*, 93(11), 1942–1954. <https://doi.org/10.1109/JPROC.2005.857486>
- [11] Mitra, J., Benidris, M., Nguyen, N., & Deb, S. (2017). A Visualization Tool for Real-Time Dynamic Contingency Screening and Remedial Actions. *IEEE Transactions on Industry Applications*, 53(4), 3268–3278. <https://doi.org/10.1109/TIA.2017.2686353>

- [12] Oliveira, W. D., Vieira, J. P. A., Bezerra, U. H., Martins, D. A., & Rodrigues, B. das G. (2017). Power system security assessment for multiple contingencies using multiway decision tree. *Electric Power Systems Research*, 148, 264–272. <https://doi.org/10.1016/j.epsr.2017.03.029>
- [13] Swarup, K. S. (2008). Artificial neural network using pattern recognition for security assessment and analysis. *Neurocomputing*, 71(4–6), 983–998. <https://doi.org/10.1016/j.neucom.2007.02.017>
- [14] Zhang, R., Xu, Y., Dong, Z. Y., Meng, K., & Xu, Z. (2011). Intelligent systems for power system dynamic security assessment: Review and classification. *2011 4th International Conference on Electric Utility Deregulation and Restructuring and Power Technologies (DRPT)*, 134–139. <https://doi.org/10.1109/DRPT.2011.5993876>