

An algorithm with LightGBM + SVM fusion model for the assessment of dynamic security region

Lulu Liang^{1,*}, Wei Hu¹, Yiwei Zhang¹, Kun Ma¹, Yujia Gu², Bei Tian², Hongqiang Li²

¹ Department of Electrical Engineering, Tsinghua University, Beijing, 100084, China

² State Grid Ningxia Electric Power Co. Ltd., Yinchuan, 750000, China

Abstract. With the development of energy transition, the complexity of power systems' structure, planning and operation is continuously increasing. As to quickly and accurately assess the dynamic security region of power system, there are prominent problems with traditional manual analysis method, i.e. the rules' roughness and a low calculation efficiency while data mining approach could provide a new way to get off such problems. Considering that the performance of SVM algorithm depends on feature selection and the LightGBM, a fast and efficient classification algorithm, can be used for feature selection, this paper proposes a new algorithm based on a fusion model. With the CEPRI-36 bus power system, the results of different algorithms are compared and the proposed algorithm verified.

1 Introduction

With the continuous development of energy transition, the interconnection scale, the access rate of renewable energy, the complexity of planning and operation of power systems are continuously increasing, which makes the power system security analysis much demanding.

In traditional power system security analysis, a few typical operation modes are selected for transient stability calculation, and the calculation results are analyzed manually to obtain the rules about power system security. Power system are typically a complex and high-dimensional nonlinear dynamic system, a simplified analysis with some dimension reduction is almost necessary. However, problems arise with simplification, such as the comprehensiveness of selected operation modes and the accuracy of security rules by manual analysis. In short words, traditional analysis method generally has the roughness of security rules and the low efficiency of analysis process.

The dynamic security region (DSR) of a pre-set fault is the operating region where power system maintains transient stability during the post-fault. DSR is usually defined in injection power space or control variable space^[1]. DSR can help to understand the security margin and then to make corresponding prevention and control decisions. Theoretically, the boundary of DSR is non-knotted, compact; there has no hole in it. Therefore, DSR can be approximately described by hyperplane^[2].

In the traditional security analysis, a few operating points with critical transient stability are found by calculation, and parameters of certain buses are selected to summarizing DSR by manual and empirical fitting. The number of critical points and selected parameter are quite

limited, equivalent to a dimensionality reduction of DSR, resulting in the loss of much system information. In practice, in case of riskiness, the resulted security region may be further zoomed in by a conservative coefficient, resulting in a smaller operational region and a scarification of system operation economy. Actually, it still cannot guarantee the accuracy of operational region. Figure 1 shows an example of two-dimensional DSR, where black dots are transient stability points, the blue dots the transient instability points, and the black dotted line the DSR boundary. If only the X-axis parameter is considered, DSR boundary becomes the light-blue dotted line area, therefore a smaller operational region. In addition, a selection of typical operation mode may make a coverage of operational region incorrect. For example, if operation points in the red circle are not included in typical operation mode, the security region becomes the red dotted line area, where with transient instability operating points.

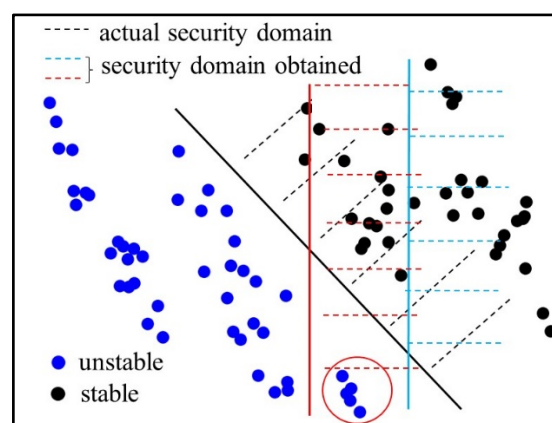


Figure 1. 2D security domain diagram

*Email: liangll19@mails.tsinghua.edu.cn

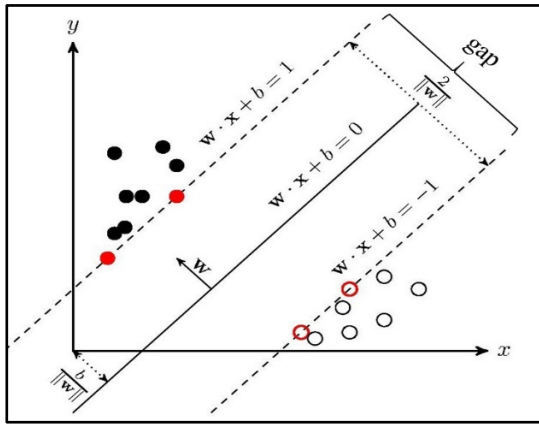


Figure 2. Basic principles of SVM algorithm

For a large-scale power system, with huge transient stability results of typical operating modes, manual analysis is always a formidable task. Data mining algorithm can find related information from massive data, which provides an approach for efficient and accurate search of high-dimensional DSR.

Recently, there are researches of data mining algorithms in power system security and stability analysis, such as artificial neural network (ANN), decision tree (DT) and support vector machine (SVM) [3-5]. However, there is few research of application in DSR. Ref. [6] shows applying SVM algorithm to solve DSR, but the accuracy is unsatisfactory. To this problem, this paper proposes an algorithm based on LightGBM+SVM fusion model, applying LightGBM for feature selection, further improving the efficiency and accuracy of SVM in solving DSR. With the CEPRI-36 bus system, performances of LightGBM, Random Forest (RF), Logical Regression (LR), K-Nearest Neighbor (KNN), XGBoost and SVM algorithms are tested and compared. Furthermore, the performances of LightGBM in feature selection are verified, compared with other relevant algorithms, i.e. principal component analysis (PCA), Relief, Variance Threshold (VT) and Mutual Information (MI).

2 LightGBM + SVM fusion model

2.1 SVM algorithm

SVM algorithm was proposed in 1995 [7], a classification algorithm based on linear discriminant function, by equation (1). To realize sample classification, it uses convex optimization technology to find the optimal discriminant surfaces. Figure 2 shows the basic principles of SVM algorithm.

$$f(x) = w^T x + b \quad (1)$$

where, w and b are respectively the normal vector and displacement term of the hyperplane; x the sample data. Decision rule is: if $f(x) > 0$, it is taken as positive class, i.e. the black dots in Figure 2, namely the stable point; if $f(x) < 0$, as negative class, i.e. the white dots in the Figure 2, namely the instable point; otherwise, it will be any class or rejected.

For the sake of classification accuracy, SVM also maximizes the gap in Figure 2 to minimize the structural risk. The objective function is shown as equation (2).

$$\begin{aligned} \min & \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \varepsilon_i \\ \text{s.t.} & y_i (w^T x_i + b) \geq 1 - \varepsilon_i \\ & \varepsilon_i \geq 0, \quad i = 1, \dots, n \end{aligned} \quad (2)$$

where ε is the relaxation variable, implying that the interval between the allowed samples and the hyperplane is less than the hard threshold 1, that is, a certain number of outliers are allowed; C , the penalty parameter, implying the importance of outliers; n , the number of training samples.

By Lagrange function and its duality, equation (2) can be transformed into equation (3).

$$L(w, b, \alpha) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \varepsilon_i - \sum_{i=1}^n \alpha_i (y_i (w^T x_i + b) - 1 + \varepsilon_i) - \sum_{i=1}^n \gamma_i \varepsilon_i \quad (3)$$

where α, γ_i is the introduced Lagrange multipliers; y_i , the label of sample stability.

According to KKT [8], equation (3) can be reduced to equation (4).

$$\begin{aligned} \max_{\alpha} & \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j x_i x_j \\ \text{s.t.} & 0 \leq \alpha_i \leq C, \quad i = 1, \dots, n \\ & \sum_{i=1}^n \alpha_i y_i = 0 \\ & w = \sum_{i=1}^n \alpha_i y_i x_i \end{aligned} \quad (4)$$

Since data is not linearly separable in most cases, SVM introduces a kernel function $\kappa(x_i, x_j)$ to replace $x_i x_j$ in equation (4) to map the input space to a high-dimensional feature space, and find a hyperplane and realize the classification function. By the introduction of kernel function, SVM omits the high-dimensional transformation of data and directly calculates the inner product in the original input space, which goes around from the dimension explosion of high-dimensional space.

Based on the principle of structural risk minimization and the introduction of kernel function technology, SVM has good classification effect, strong interpretability and generalization ability. When a linear kernel function is used, SVM can give the expression of classification hyperplane, namely, the boundary of DSR. However, when the number of features is large, the cost of SVM training is very high and likely to overfit. As a result, a process of feature selection is needed and the results' quality has a great impact on the accuracy of SVM.

2.2 LightGBM algorithm

LightGBM was proposed in 2017 [9], an improved version of Gradient Boosting Decision Tree (GBDT). GBDT, an integrated algorithm, consists of a series of linear combinations of submodels. Based on the principle of iteration, it takes the regression tree as the submodel and adds the submodels one by one to decrease the loss

function of learner. GBDT can be expressed as equation (5).

$$F_k(x) = \sum_{i=1}^k f_i(x | \theta_i) \quad (5)$$

where $f_i(x|\theta_i)$ is the regression tree submodel newly added in the i -th iteration; θ_i , parameter of the submodel; k , the number of submodels; x , the sample data. θ_i is obtained by minimizing the loss function, by equation (6).

$$\theta_i = \arg \min L[F_{i-1}(x) + f_i(x | \theta_i)] \quad (6)$$

where L is the loss function that the learner uses for prediction.

GBDT uses pre-sorting algorithm for feature selection and splitting, thus is very time- and memory- consuming, and not suitable for massive processing. On the contrary, LightGBM uses histogram algorithm and leaf growth strategy to replace GBDT's pre-sorting algorithm and layer growth strategy respectively, greatly improving the speed and efficiency of the algorithm.

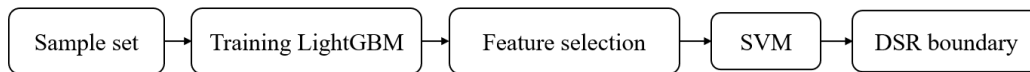


Figure 3. Scheme of DSR algorithm with LightGBM + SVM fusion model

2.3 LightGBM +SVM fusion model

For classification task, LightGBM is fast, efficient and difficult to over fit, especially for high-dimensional data, but it can only give label classification. SVM algorithm with linear kernel function can give hyperplane representing DSR, but its effect depends on the quality of feature selection. Taking the advantages of the two algorithms, this paper proposes a DSR algorithm based on LightGBM + SVM fusion model, shown in Figure 3. Firstly, the training set is sent to LightGBM for training. In the training process, features importance used for feature selection are determined by two parts, i.e. times being used and their gain to final classification results. Finally, selected features are sent to SVM for training to solve the DSR boundary.

3 Examples and analysis

3.1 CEPRI-36 bus system and samples

To verify the proposed method, the CEPRI-36 bus system with 8 generators, 32 lines and 10 loads, as in [6] is taken as test system. The pre-set fault is: 3-phase short circuit occurs on the sending end of line bus16-bus20 at 0s, and the line is cut off at 0.2s. Transient process of the system within 5s after the fault appearance is simulated. By changing system's operation point, i.e. the active power of each generator and the power of each load varies in a range of 80%~120%, a total of 8000 samples are generated. Taking the injected active and reactive power, the voltage amplitude and phase angle of buses as the initial features, and the stability of the system transient as the label, there are 228 initial features, 5407 stable samples and 2593 unstable samples.

3.2 Calculation and analysis

3.2.1 Algorithm performance test

70% of the samples are randomly selected into the training set, and the remaining 30% of the samples into the test set. Optimal parameters are found through Grid Search and 5-fold Cross Validation, and test results of each data mining algorithm under parameter tuning are shown in Table 1. LightGBM mainly optimizes six parameters: tree model depth (P_{max_dep}), number of leaf nodes (P_{num_leaves}), learning rate ($P_{learning_rate}$), minimum number of data on a leaf ($P_{min_data_in_leaf}$), random selection of features and data proportion in each iteration (P_{fea_fra} , P_{bag_fra}); SVM uses linear kernel function.

Table 1. Efficiency and accuracy of various data mining algorithms

Algorithm	Training time(s)	Accuracy(%)	
		Training set	Test set
SVM	18.745	96.32	95.38
RF	8.5978	100	96.45
KNN	3.6848	96.17	94.08
LR	2.875	96.48	95.21
XGBoost	13.400	99.98	96.92
LightGBM	0.8891	100	98.08

The results of Table 1 shows that LightGBM algorithm has the fastest training speed and highest accuracy. SVM takes a long time to train although with the advantage of domain visualization.

Various feature selection algorithms are used to be combined with SVM for training, and the test results are listed in Table 2.

Relief algorithm in Table 2 is the feature selection method in [6]. In the test, SVM uses a linear kernel function, and the penalty parameter C is 100. 30 features selected by each selection algorithm are sent to SVM training. The parameters of LightGBM algorithm are set to: $P_{max_dep} = 8$, $P_{num_leaves} = 80$, $P_{learning_rate} = 0.01$, $P_{fea_fra} = 0.5$, $P_{bag_fra} = 0.5$, $P_{min_data_in_leaf} = 20$. Results shows that the model proposed in this paper outperforms all other existing models in accuracy.

Table 2. Accuracy of various feature selection algorithms with SVM

Model	Accuracy(%)	
	Training set	Test set
PCA + SVM	86.33	85.83
MI + SVM	91.61	91.43

VT + SVM	93.93	93.66
Relief + SVM	90.80	90.75
LightGBM + SVM	96.28	96.24

3.2.2 Sensitivity analysis

Further, LightGBM is used as feature engineering to build features automatically. In LightGBM, the setting of parameter P_{num_leaves} is generally as: $classes < P_{num_leaves} < 2^{P_{max_dep}}$, this paper takes $2^{P_{max_dep}}$; the smaller the parameter $P_{learning_rate}$, the higher the precision. Sensitivity analysis focuses on mainly the influence of the parameters P_{max_dep} , P_{fea_fra} , P_{bag_fra} , $P_{min_data_in_leaf}$ on the final result. Parameter setting is as in Section 3.2.1, the control variable method is employed for experiment, and the results are shown in Figure 4.

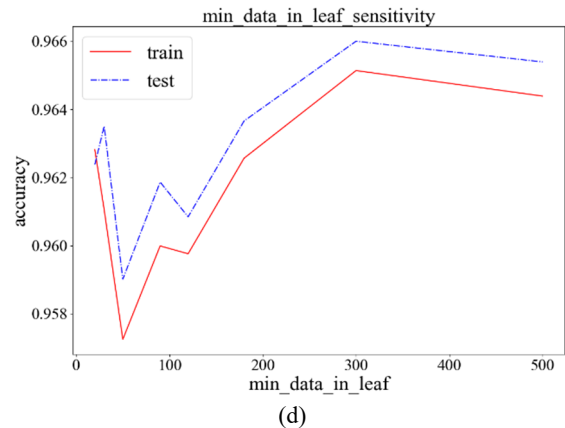
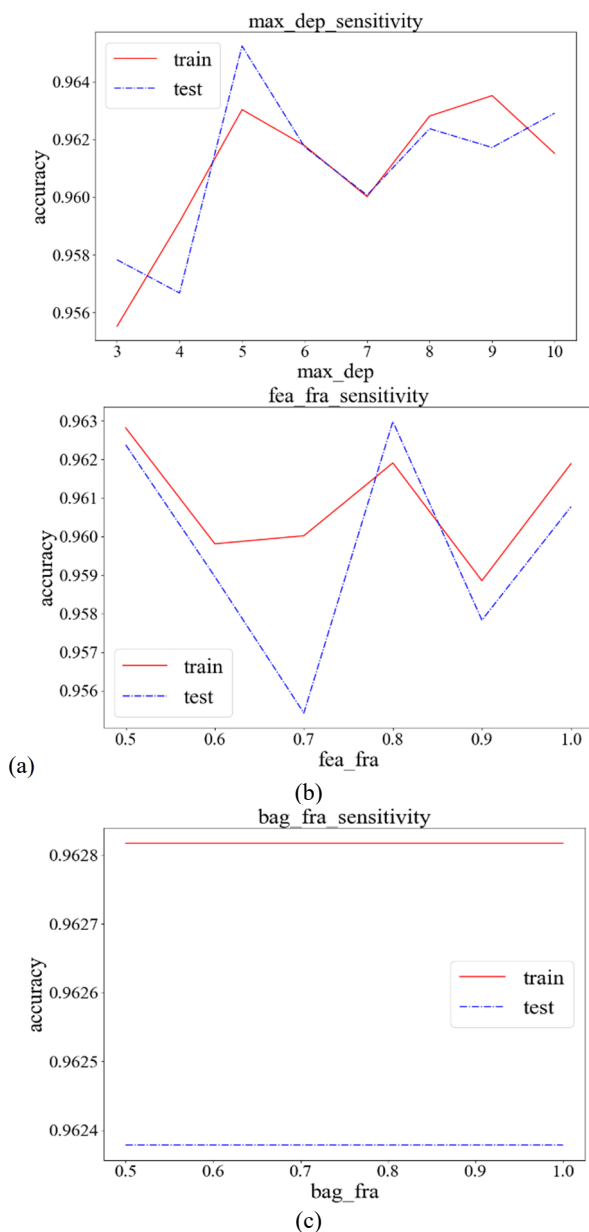


Figure 4. Sensitivity analysis of parameters P_{max_dep} , P_{fea_fra} , P_{bag_fra} , $P_{min_data_in_leaf}$

It can be seen that: (1) the model accuracy with all parameters is more than 95.5%, indicating that the model is robust to parameters and thus work well without complex parameter adjustment; (2) in various cases, the model accuracy for training set and test set is almost the same, implying that the model does not incline to overfit and has good generalization ability.

4 Conclusion

Power system is a typical high-dimensional nonlinear dynamic system, which makes it almost impossible to solve the accurate dynamic security region manually. Data mining algorithm is a key technology of mining information from massive data, so this paper applies it to solve the problem of dynamic security region. Main work and conclusions of this paper include: (1) A dynamic security region solving algorithm based on LightGBM+SVM fusion model is proposed, where LightGBM algorithm is used for feature selection before feeding data into SVM training. Test results of SVM, RF, XGBoost, LR, KNN and LightGBM show that LightGBM classification algorithm is fast and accurate; comparing the accuracy of SVM with LightGBM, PCA, Relief, MI and VT respectively used as feature selection algorithms, results show that LightGBM has great advantages as a feature selection algorithm. (2) Sensitivity analysis shows that the algorithm proposed in this paper has high accuracy in both training set and test set under different parameters, implying that it is insensitive to parameters, less difficult to adjust parameters, not inclines to over-fit and has strong generalization ability.

Acknowledgments

This work was supported by the Science and Technology Project of the State Grid Corporation of China (No. 5400-201935258A-0-0-00)

References

1. Yu, Y. (2002) Security region of bulk power system. In: Proceedings of the 2002 International

- Conference on Power System Technology. 2002, Oct 13–17. Kunming, China. pp. 13–7.
2. Dong, C., Yu, Y. (2005) PDSR in Phase Angle Space and SPM Based Security Monitoring of Power Systems. In: TENCN 2005 - 2005 IEEE Region 10 Conference. 2005, Nov 21-24. Melbourne, VIC, Australia. pp. 1-5.
 3. Sobajic, D. J., Pao, Y. H. (1989) Artificial neural-net based dynamic security assessment for electric power systems. *IEEE Power Engineering Review*, 9(2), pp. 55-55.
 4. R. Diao, V. Vittal and N. Logic. (2010) Design of a Real-Time Security Assessment Tool for Situational Awareness Enhancement in Modern Power Systems. *IEEE Transactions on Power Systems*, 25(2), pp. 957-965.
 5. Moulin, L. S., Silva, A. P. A. D., El-Sharkawi, M. A., Li, R. J. M. (2004) Support vector machines for transient stability analysis of large-scale power systems. *IEEE Transactions on Power Systems*, 19(2), pp. 818-825.
 6. Zhang, P., Hu, W., Liu, X., Xu, X., Shao, G. (2017) Study on practical dynamic security region of power system based on big data. In: 2017 12th IEEE Conference on Industrial Electronics and Applications (ICIEA). 2017 June 18-20. Siem Reap. pp. 1996-1999.
 7. Cortes, C., Vapnik, V. (1995) Support-vector networks. *Machine learning*, 20(3), pp. 273-297.
 8. Kuhn, H. W., Tucker, A. W. (1951) Nonlinear programming. *Berkeley Symposium on Mathematical Statistics & Probability*. University of California Press.
 9. Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., et al. (2017) LightGBM: a highly efficient gradient boosting decision tree. In: *Advances in Neural Information Processing Systems*. Cambridge. pp. 3146-3154.