MATEC Web of Conferences **336**, 09028 (2021) https://doi.org/10.1051/matecconf/202133609028
CSCNS2020

Identification of abuse of market power by power generation enterprises

Fengqing Du[1], *Zhe* Liu[1], *Hongyuan* Tao[2,*], *Huichun* Hua[2], and *Lifeng* Zhang[2]

[1]State Grid Shanghai Electric Power Company, Shanghai, China
[2]State key laboratory of the new energy power system (North China Electricity Power University), Baoding, Hebei province, China

Abstract. At present, the reform of the power market is progressing steadily. To ensure the efficient and healthy operation of the power market, there is an urgent need to strengthen the credit supervision of the electricity market entities. Identifying violations of power generation companies' abuse of market power is a key task in the credit supervision of power market entities. Traditional power generation companies' abuse of market power identification mainly relies on expert decision-making. However, with the increase in market transaction volume, expert decision-making cannot meet the needs of work, and an intelligent identification method suitable for computer analysis must be proposed. This paper first proposes a quantitative definition of abuse of market power, and then takes into account the specific data characteristics of the electricity market, and proposes a method of identifying violations of power generation companies based on improved cost-sensitive support vector machines. Finally, the power market simulation experiment data set labeled by the definition method is used for training and testing. The test results show that the abuse of market power by power generation companies can be quickly and accurately identified, which verifies the effectiveness of the proposed method.

1 Introduction

In the context of the new round of electricity reform, power market transactions are more diversified, and the rapid development is accompanied by corresponding market risks. Because of China's special national conditions, compared with foreign mature power markets, the power generation industry has high barriers to entry, and it is difficult to break the existing oligopoly situation of some power generation enterprises. In this situation, individual power generation companies may abuse their market power in order to obtain higher profits by using their own scale advantages, which will cause market price distortions and reduce market efficiency. At present, the identification method of abuse of market power by domestic power generation enterprises mainly relies on expert decision-making. However, as the scale of transactions in the electricity market increases and the frequency of transactions increases, the efficiency of expert decision-making methods is

* Corresponding author: taohy@163.com

© The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (http://creativecommons.org/licenses/by/4.0/).

relatively low. In order to meet the demand of real-time identification of abuse of market power by power generation enterprises, an intelligent identification method must be found.

Domestic research on the definition of abuse of market power by power generation enterprises has been relatively mature. Literature [1] points out that whether power generation enterprises abuse market power is usually analyzed from three levels. The first level is to analyze whether the power generation enterprises have market power and its size. The second level analyses whether the power generation enterprises make use of market power. The third level analyses the harm degree caused by power generation enterprises using market power. At the same time, literature [1] proposes to use the degree of price abnormality to quantitatively describe the harm caused by abuse of market power. In the identification of abuse of market power by power generation enterprises [2-5], the mainstream methods in China include constructing a supervision index system [6-11], conducting power market simulation [12], and methods based on data processing [13-14]. Literature [6-8] analyses and summarizes the traditional identification indicators of abuse of market power at home and abroad, and literature [9] establishes a set of indicators to measure market power from the perspective of electricity price. The literature [13-14] respectively use AHP and PCA to analyze the abuse of market power by power generation enterprises. But all the above methods require researchers to have a comprehensive understanding of the real situation of power generation enterprises, which is difficult to achieve in the actual power market. In the actual electricity market, the only data that we can easily obtain is quotation data. Therefore, this paper proposes a method for identifying abuse of market power by power generation companies based on quotation data.

The identification of violations of power generation companies is essentially a two-category problem, and the support vector machine (SVM) as an intelligent algorithm has strong applicability in the two-category problem due to its high accuracy and strong generalization ability. Therefore, this paper adopts SVM identify the abuse of market power by power generation companies. However, it is inefficient in calculating large-scale data, so the solving algorithm needs to be improved. Moreover, the data of abuse of market power by power generation enterprises is obviously unbalanced, so it is necessary to improve the algorithm itself in terms of cost sensitivity.

In order to verify the effectiveness of the intelligent identification method for the abuse of market power by power generation companies, this paper first gives a quantitative definition of the abuse of market power by power generation companies. Secondly, according to the characteristics of the power market data and the shortcomings of the SVM itself, corresponding improvements have been made to the SVM. Finally, the power market simulation experiment data set labeled by the definition method is used to train and test the improved algorithm.

2 A quantitative definition of abuse of market power

In order to judge whether power generation enterprises abuse market power more accurately, it is necessary to put forward its quantitative definition first. In the market economy, enterprises aim at making profits, so to judge whether power generation enterprises have the intention of abusing market power should be analyzed based on the power generation enterprises' strategy of quoting high prices. What is certain is that the higher the proportion of a power generation enterprise reporting high price or the higher the declared electricity price, the more obvious the intention of the enterprise abusing market power.

Secondly, to determine the quantitative definition, it is necessary to give the calculation formulas of two values: the quoted cost deviation value D (the degree that the declared price deviates from the production cost) and the quoted high price ratio R (the proportion of electricity whose declared price is high). The calculation formulas are as follows:

$$D = \frac{P - M}{M} \tag{1}$$

$$R = \frac{q}{Q} \times 100\% \tag{2}$$

where P represents the declared electricity price, M represents the marginal power generation cost enterprise, q represents the declared price is high-priced declared electricity, Q represents the total declared electricity.

The Delphi method, a classical algorithm for determining the critical value in engineering, can be used to determine the critical value of the quoted cost deviation and the quoted high price ratio. In view of the current situation of insufficient competition in the domestic electricity market, in order to prohibit violations to the greatest extent, it is necessary to set a more stringent definition of violations to avoid abnormal prices in the electricity market. By Delphi method, the critical values of the quoted cost deviation and the quoted high price ratio are determined to be 0.3 and 30%, respectively.

Because the marginal power generation cost in the above quantitative definition is difficult to obtain in the actual power market, it is impossible to use the definition method to identify the abuse of market power by power generation companies. Therefore, this paper proposes an intelligent identification method based on the data declared by power generation companies.

3 Intelligent recognition algorithm

Violation identification is essentially a classification problem, that is, the transaction data is divided into two categories: violation and non-violation. Support vector machine (SVM) is a high-performance classification algorithm, which can accurately identify illegal data in transactions. According to the unbalanced data of power market, this paper adopts cost-sensitive support vector machine (CSVM). Since support vector machines have low solution efficiency when the data scale is too high, this paper makes improvements to the support vector machine solution algorithm.

2.1 Cost-sensitive support vector machine (CSVM)

Set a given training set $\{(x_1,y_1),(x_2,y_2),\ldots,(x_n,y_n)\}$, where n is the number of power generating enterprises; $y_i \in \{1,-1\}$ is the definition identification result, 1 means non-violation, -1 means violation ; x_i is power generation enterprise quotation data, for example, in the three-stage quotation rules, $\mathbf{x}_i = (x_{i1}^p, x_{i1}^q, x_{i2}^p, x_{i2}^q, x_{i3}^p, x_{i3}^q)$, where x_{ij}^p , $,x_{ij}^q$ respectively represent the declared electricity price and declared electricity quantity of the i-th power generation company in segment j. Set the optimal classification interface as follows:

$$w^T x + b = 0 \tag{3}$$

The mathematical model of linear separable cost-sensitive support vector machine is:

$$\begin{aligned}&\min f(w,\xi)=\frac{1}{2}w^T w+C_+\sum_{i=1}^{m}\xi_i+C_-\sum_{i=1}^{l}\xi_i\\&\textit{subject to}:\\&y_i(w^T+b)\geq 1-\xi_i\\&\xi_i\geq 0\end{aligned}\tag{4}$$

In which ξ_i is a slack variable, C_+ is a positive sample punishment parameter, and C_- is a negative sample punishment parameter ($C_+=l/m\cdot C_-$).

When sample points are linearly indivisible, it is necessary to introduce kernel functions and transform them into linearly separable data for processing. The kernel function is defined as:

$$K\left(\boldsymbol{x}_i,\boldsymbol{x}_j\right)=\langle\phi\left(\boldsymbol{x}_i\right),\phi\left(\boldsymbol{x}_j\right)\rangle=\phi^T\left(\boldsymbol{x}_i\right)\cdot\phi\left(\boldsymbol{x}_j\right)\tag{5}$$

According to the Lagrange duality principle, it can be transformed into:

$$\begin{aligned}&\max_{\alpha}\left(\sum_{i=1}^{n}\alpha_i-\frac{1}{2}\sum_{j=1}^{n}\sum_{i=1}^{n}\alpha_i\alpha_j y_i y_j \boldsymbol{x}_i^T\boldsymbol{x}_j\right)\\&\text{subject to:}\\&\sum_{i=1}^{n}\alpha_i y_i=0\\&C_i\geq\alpha_i\geq 0\end{aligned}\tag{6}$$

where $C_i=C_+(y_i=1),C(i)=C_-(y_i=-1)$.

Suppose $H(i,j)=y_i y_j K(x_i,x_j)$, convert it to matrix form:

$$\begin{aligned}&\min -e^T\alpha+1/2\alpha^T H\alpha\\&s.t.\ \ 0\leq\alpha\leq Ce\\&\qquad y^T\alpha=0\end{aligned}\tag{7}$$

where $C(i)=C_+(y_i=1),C(i)=C_-(y_i=-1)$, e is a column vector with all ones.

2.2 improve the cost-sensitive support vector machine (CPPA-CSVM)

Because the core of support vector machine is to solve quadratic programming problems, when the data scale is too large, the operation speed will be reduced. The scale of transaction data in power market is related to the number of segments of quotations. Each segment of the quotation involves two data: declared price and declared electricity quantity. When there are many segments, the data scale will be too large. Therefore, the multi-dimensional analysis (MDS) algorithm is used to simplify the data before solving, so as to improve the operation efficiency. For solving the quadratic programming problem, this paper first transforms it into a convex optimization problem with linear inequality constraints, then transforms it into a variational inequality problem, and finally uses the Customized Proximal Point Algorithm (CPPA) [15] to solve the variational inequality problem. The specific steps are as follows:

Step 1: Assume that the zero-space matrix of the data label vector $\boldsymbol{y}$ is Z, then α in the original problem can be expressed as $\alpha=Z\upsilon$. Then the original problem can be transformed into a linear inequality constrained convex optimization problem:

$$\begin{aligned} &\min \theta(\upsilon)=\frac{1}{2}\upsilon Z^T HZ\upsilon - e^T Z\upsilon \\ &s.t.\, A\upsilon \geq b \left(A=\begin{bmatrix} Z \\ -Z \end{bmatrix}, b=\begin{bmatrix} 0 \\ -Ce \end{bmatrix} \right) \end{aligned} \tag{8}$$

Step 2: Firstly, the convex optimization problem is written as Lagrange function.

$$L(\upsilon,\lambda)=\theta(\upsilon)-\lambda^T(A\upsilon-b) \tag{9}$$

Then find the optimal point (υ^*,λ^*) of Lagrange function and transform it into a variational inequality problem:

$$\theta(\upsilon)-\theta(\upsilon^*)+\left(\begin{pmatrix} \upsilon \\ \lambda \end{pmatrix}-\begin{pmatrix} \upsilon^* \\ \lambda^* \end{pmatrix}\right)^T \begin{pmatrix} -A\lambda^* \\ A\upsilon^*-b \end{pmatrix} \geq 0 \tag{10}$$

Step 3: Use the Customized Proximal Point Algorithm (CPPA) to solve the variational inequality problem. Customized Proximal Point Algorithm is expressed as follows:

$$\begin{aligned} &\tilde{\omega}^k=(\tilde{\upsilon}^k,\tilde{\lambda}^k),\omega^k=(\upsilon^k,\lambda^k) \\ &\begin{cases} \tilde{\upsilon}^k = \arg\min\{\theta(\upsilon)+\frac{r}{2}\left\|\upsilon-[\upsilon^k+\frac{1}{\tau}A^T\lambda^k]\right\|^2\} \\ \tilde{\lambda}^k=\{\lambda^k-\frac{1}{s}[A(2\tilde{\upsilon}^k-\upsilon^k)-b]\}_+ \\ \omega^{k+1}=\omega^k-\tau(\omega^k-\tilde{\omega}^k)\,\tau\in(0,2) \\ r\cdot s\geq \left\|A^T A\right\| \end{cases} \end{aligned} \tag{11}$$

4 Case analysis

In this paper, the power market simulation data set labeled by the definition method is used to analyze the examples, and 39 sample points are selected to construct the data set. By calculating the deviation value of quotation cost and the ratio of quoted price in each quotation, 5 samples were found to be in violation and 34 samples were found to be in non-violation. Then multidimensio-nal scale analysis method is used to reduce the dimension of training data. In this case, the first three elements are taken, and the cumulative contribution rate reaches 99.03%. Part of the original data is shown in Table 1, and some data after dimensionality reduction are shown in Table 2.

Table 1. Part of the original data.

Power generation enterprises	1	10	19	29	39
First bidding capacity	48617.0	48721.0	75142.0	271848.2	98324.0
First bidding price	299.0	299.0	300.3	298.0	299.0
First deal capacity	48617.0	48721.0	75142.0	271848.2	95768.0
Second bidding capacity	13210.0	24360.5	37571.0	61231.5	47884.0
Second bidding price	299.1	299.9	302.4	301.4	301.7
Second deal capacity	13210.0	24360.5	37571.0	61231.5	47884.0
Third bidding capacity	3173.0	16918.5	7287.0	66920.3	3792.0
Third bidding price	305.1	304.6	304.1	306.7	309.9

Third deal capacity	0.0	0.0	7287.0	0.0	0.0
Marginal generation cost	232.0	240.0	231.5	231.5	241.0

Table 2. Partial data after dimensional reduction through multidimensional scaling analysis.

Power generation enterprises	x_1	x_2	x_3
1	0.915	-0.549	-0.2914
10	0.7297	0.1823	2.9268
19	0.9408	0.6183	2.5026
29	-1.7628	1.8628	-0.2217
39	0.3407	0.863	-0.2093

3.1 Improve cost-sensitive support vector machine training

In order to verify the performance of the Improve cost-sensitive support vector machine, this paper uses 39 sample points of the data set to conduct an experimental test on a PC. The experimental environment is as follows: 2.3GHz CPU, 8G RAM, software MATLAB R2014b. For the convenience of comparison, this paper uses ordinary support vector machine (SVM) and improved cost-sensitive support vector machine (CPPA-CSVM) for training and testing. Draw a three-dimensional scatter diagram of the data set. It can be seen from Figure 1 that the data set is linearly inseparable and needs to be mapped by kernel function. This paper chooses Gaussian kernel function for mapping.

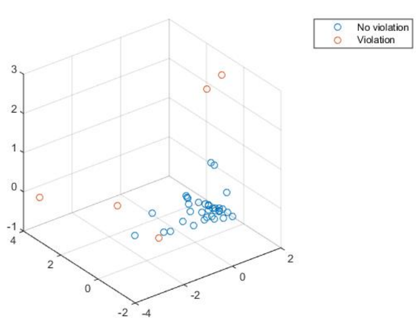

Fig. 1. 3D scatter diagram of the data set.

3.2 Result analysis

This paper selects 75% of the data set as the training set and 25% of the data set as the test set. The test results and evaluation indexes of the two algorithms are shown in Table 3. It can be seen from Table 3 that the training accuracy and recall rate of the two algorithms are both 100%, and both of them have reached 100% accuracy and recall rate in the test set.

The training time of the improved support vector machine is only 0.034 seconds while ensuring the correct rate, which is obviously faster than the ordinary support vector machine. The results show that the proposed method can quickly and effectively identify the irregularities of power generation enterprises abusing market power.

Table 3. Test results.

	SVM	CPPA-CSVM
Training time/s	0.205	0.034
Training set Correct rate	100%	100%
Training set recall rate	100%	100%
Test set correct rate	100%	100%
Test set recall rate	100%	100%

5 Conclusion

In this paper, the quantitative definition of abuse of market pow er is given, and an identification algorithm based on improved cost-sensitive support vector machine is designed. Combined with the quoted data of power generation companies, this algorithm can realize the real-time identification of the abuse of market power by power generation companies and promote the development of credit evaluation of market entities. The main conclusions of this paper are as follows:

(1) Based on the data declared by power generation companies in the transaction, a quantitative definition of abuse of market power is proposed. Aiming at the imbalance of illegal data in power generation companies, this paper chooses cost-sensitive support vector machines (CSVM).

(2) The improved cost-sensitive support vector machine (CPPA-CSVM) overcomes the problem of low efficiency of SVM in solving large-scale data while ensuring the high accuracy of the original algorithm, which lays a solid foundation for real-time identification of abuse of market power by power generation enterprises.

This paper is supported by the Science and Technology Projects of State Grid Shanghai Electric Power Research Institute (Grant No. SGSHDK00HZJS2000254).

References

1. Q. Xia, C.B. Li, J.J. Jiang, et al, Power System Technology, **27**, 1-4 (2003)
2. M.L. Bao, Y. Ding, C.Z. Shao, et al, Proc Chin Soc Elect Eng, **17,** 4-15+330 (2017)

3. M. Bai, A.M. He, Price: Theory & Practice, **04**,17-21 (2017)
4. L.Z. Zhang, China Energy News, **004** (2017)
5. B.B. Chakrabarti, D.G. Goodwin, POWERCON (2008)
6. J. Zhang, L.Z. Zhang, L. Yu, Proceedings of the Chinese Society for Electrical Engineering, **06**, 123-128 (2006)
7. W.M. Liu, J.J. Wu, K. Yang, Power System Technology, **S2**, 211-214 (2007)
8. S.H. Shi, D.N. Liu, H.W. Hu, et al, Price : Theory & Practice, **08**, 51-54 (2018)
9. J.W. Ding, Y. Shen, C.Q. Kang, et al, Automation of Electric Power Systems, **13**,24-29+67 (2003)
10. L.J. Yang, Y. Zhao, Z.F. Tan. Power System Technology, **07**, 26-31 (2006)
11. M. Shafie-Khah, M.P. Moghaddam, M.K. Sheikh-El-Eslami, IET GENER TRANSM DIS, **10**, 1842-1852 (2016)
12. S. Soleymani, A.M. Ranjbar, A. Jafari, et al, IEEE Power India Conference, IEEE (2006)
13. Y.G. Yong, Z.G. Zhang, X.Y. Wang, Journal of North China Electric Power University, **04**, 76-79+84 (2007) L. Hongze, W. Bao, G. Sen. Modern Electric Power, **28**, 85-89 (2011)
14. H.Z. Li, B. Wang, S. Guo, Modern Electric Power, **28**, 85-89 (2011)
15. Y. Peng, Research on Support Vector Machine Algorithm Based on Variational Inequality[D]. Nanjing University (2016)
16. J.K. Lin, Y.X. Ni, F.L. Wu, Power System Technology, **11**, 70-76 (2002)
17. S. Prabhakar Karthikeyan, I. Jacob Raglend, D.P. Kothari, INT J ELEC POWER, **48**, 139-147 (2013)