Fuzzy Systems and Data Mining VI
A.J. Tallón-Ballesteros (Ed.)
© 2020 The authors and IOS Press.
This article is published online with Open Access by IOS Press and distributed under the terms of the Creative Commons Attribution Non-Commercial License 4.0 (CC BY-NC 4.0).
doi:10.3233/FAIA200692

Application of Grey Prediction in a GA-BP Power Theft Algorithm

Chen Xiaofeng^a, Xu Zhongping^a, Zhang Lipeng^{a,1}, Zhu Feng^a, Qi Xiaoming^a, Chen Fangzhen^b ^aState Grid Information & Telecommunication Accenture Information Technology Co., Ltd, Beijing, 100032, China ^bPingdingshan Power Supply Company of State Grid Henan Electric Power Company, Pingdingshan, 467000, China

Abstract. Statistics show that power theft is one of the main reasons for the dramatic increase in power grid line loss. In this paper, a genetic algorithm is used to optimize a neural network and establish a power theft prediction model. With the grey prediction model, the predicted values of variables are obtained and then applied to the prediction model of a GA-BP neural network to obtain relatively accurate predictions from limited samples, reducing the absolute error. Through the two levels of prediction and analysis, the model is a practical and effective method for power companies to carry out power theft analysis.

Keywords. Genetic Algorithm, Neural Network, Electricity Theft Prediction Model, Grey Model, Electricity Anti-Theft

1. Introduction

Traditional power theft investigation involves the observation of line loss and power consumption data of the area of jurisdiction[1-3]. The staff of the power supply office determine a list of suspects according to their experience and specifications, and determine power theft behavior in combination with on-site investigation.

With the continuous evolution of big, hierarchical data, numerous prediction methods via data analysis have emerged[4-5]. Zheng[6] modeled and analyzed daily electricity consumption data based on the LeNet-5 convolutional neural network, and then used a double-layer deep network to conduct in-depth mining of user information, line loss in the platform area and other data, laying a foundation for accurate anti-power theft. Cai et al.[7] analyzed the correlation of user data and applied a machine learning decision tree method to help staff to determine whether a user has stolen electricity in a time period, and effectively reduce the economic loss. Li et al.[8] studied the potential relationship between electricity consumption behavior and electricity cost risk through big data analysis, so as to accurately locate users who are potentially stealing power. Li et al.[9] constructed a relevant line loss model by using a

¹ Corresponding Author: Zhang Lipeng, State Grid Information & Telecommunication Accenture Information Technology Co., Ltd, Email: Philip_Zhang1@163.com

neural network in a data-driven way, and realized the theoretical calculation of line loss and judgment of potential power theft.

With reference to previous studies, we decided to adopt different methods for prediction and then introduce the grey prediction model. The data obtained using the grey prediction model is applied to the GA-BP algorithm, so as to obtain the proposed method of determining suspicion of electricity theft.

2. Theoretical analysis of the GA-BP algorithm

2.1 BP neural networks

The structure of the backpropagation (BP) neural network topology is shown in Figure 1. A BP neural network input variables through the excitation function and adjust each layer connection weights threshold eventually nonlinear mapping to the output variable. In order to make the network output variable consistent with the expected output, learning generally requires several repetitions; the error value tends to zero and ultimately achieve zero[10-12].



Figure 1: Topological structure of a BP neural network.

The BP neural network algorithm process is as follows:

Suppose the input matrix of the sample is $X = (x_{ij})$, $(i = 1,2,3, \dots, m; j = 1,2,3, \dots p)$, each row of data represents a group of input samples, and each group of input samples corresponds to a group of output samples, so the actual output samples corresponding to all input samples are $XY = [y_1, y_2, y_3, \dots, y_n]^T$. Each of these columns is an index of the sample, so the input sample I_1 of the input layer is equal to X.

If the hidden layer of the network contains m neurons, weight matrix $W = (W_{ij}), (i = 1, 2, 3, \dots, m; j = 1, 2, 3, \dots p)$, threshold matrix $B = [b_2, b_2, \dots b_m]^T$, the input of the hidden layer is:

$$I_2 = W_{m \times p} X_{p \times n} + Bones_{1 \times n}$$
(1)

ones_{$1\times n$} is a matrix where all of the elements of the $1 \times n$ vector are 1. The selection of the hidden layer excitation function is a unipolar S function, whose expression is:

$$f(x) = \frac{1}{1 + e^{-x}}$$
(2)

The output of the hidden layer can be obtained as follows: $O_2 = f(I_2)$. The input

to the output layer is:

$$I_3 = X_{ik}O_2 + B_{ik} + ones_{1 \times n}$$
(3)

Since the transfer function is a linear function, the output of the output layer can be considered to be $O_3 = I_3$. In order to obtain connection weight threshold adjustment between layers, the chain rule is used:

$$\Delta W_{jk} = -\eta \frac{\partial E}{\partial W_{jk}} = -\eta (Y - O_3) O'_2 \tag{4}$$

$$\Delta B_{jk} = -\eta \frac{\partial E}{\partial B_{jk}} = -\eta (Y - O_3) \text{ones}_{n \times 1}$$
(5)

Again, $ones_{1 \times n}$ is a matrix where all of the elements of $1 \times n$ are 1. By formula (2), f'(x) = f(x)[1 - f(x)]. The weight threshold adjustment between hidden layers is further calculated as:

$$\Delta W_{jk} = -\eta W'_{jk} (Y - O_3) O_2 (1 - O_2)$$
(6)

$$\Delta B_{jk} = -\eta \frac{\partial E}{\partial B_{jk}} = -\eta W'_{jk} \tag{7}$$

According to formulae $(4) \sim (7)$ it can be seen that the correct initial threshold can be determined from the second to the first cumulative power threshold; thus, the accuracy of the training is decided by the selection of the initial weight threshold. If the initial threshold selection is not appropriate, it is likely to result in slow convergence and ease of falling into local optimal solutions, so the right choice of initial threshold is important[12].

2.2 BP neural network optimazation with GA

A genetic algorithm (GA) is a global heuristic search algorithm inspired by biological evolution[13]. As a GA is a global search algorithm that can find the global optimal solution, this can address the deficiency of BP neural networks that have a tendency to converge to sub-optimal solutions. Hence, this paper uses a GA to optimize the BP neural network to ensure the accuracy of the model prediction results[14].

2.2.1 Coding and population initialization

The premise of using a genetic algorithm is to encode the solution of a problem into the form of a string. This is achieved using a binary encoding, the simplest coding method [15]. N initial string structures are randomly generated at first, each representing an individual; the N individuals constitute a population. The algorithm begins evolution with this population, initially.

2.2.2 Fitness assessment

Fitness indicates the quality of an individual solution, and the mean squared error between the predicted result and the actual result is taken as the evaluation of individual fitness [16], The calculation is as follows:

$$F = mse(Y - 0) = \frac{1}{n} \sum_{i=1}^{n} (y_i - o_i)^2$$
(8)

where n is the number of inputs to the samples; y_i is the prediction of the network; o_i is the actual result, **mse** is the mean squared error function.

2.2.3 Model building

Based on the optimization of weights and thresholds of the neural network by the GA, the structure of the neural network is designed, including the selection of the input, output, data processing, the number of hidden layer nodes, and the selection of the activation function.

Step 1: Vector selection Inputting.

The input vector of the mathematical model in this paper is the evaluation of a user's electricity consumption status, including the month, energy efficiency level of the user, monthly electricity consumption, service meter type, line loss rate, three-phase voltage imbalance rate, three-phase current imbalance rate, power factor, single consumption and event record. The internal characteristics of each index intersect, and the index data are correlated and influenced by each other.

Step 2: Selection of output vector.

In order to evaluate the power consumption status of users, the suspected power stealing factor and mode are considered as the output of the network. Here, this is divided into either not electricity-stealing (represented by 0) or suspected electricity-stealing (represented by 1).

2.2.4 Slection Practicing

The selection operator in GAs aims to select individuals with strong adaptability to reproduce in the next generation. This paper adopts the proportional selection strategy; assuming that the total number of individuals in the population is N, then the fitness of individuals n is:

$$p_i = \frac{k/F_i}{\sum_{i=1}^N k/F_i} \quad (k \text{ is the coefficient})$$
(9)

2.2.5 Crossover and mutation

In this paper, some genes were exchanged among chromosomes by single point crossover. For mutation, a simple bit flip approach is used.

3. Application of grey prediction in GA-BP power theft prediction

This section details the GA-BP neural network model for power theft analysis. The evaluation indexes of the power sample data is used as input to the grey model. The output is the suspected behavior: 1 represents unusual behavior, whereas 0 means normal behavior [17].

3.1 Data preprocessing

The test data in this paper are from a power supply company, which is low-voltage non-residential user data. The data of 560 users used. Of these, 10 had stolen electricity,

while the remaining 550 had not.

First of all, it is necessary to determine the evaluation indexes regarding electrical power theft from the power metering data. According to the result of data analysis and measurement of electrical parameter can affect the watt-hour meter and watt-hour meter measurement error to some changes in the electrical parameters, selected the monthly electricity consumption and the area line loss, maximum line loss, power factor, the three-phase imbalance rate and contract constitute the power capacity than the six electrical parameters evaluation index system, and then in order to eliminate the index data and the dimensional difference between the order of magnitude. An index is normalized as follows [18]:

$$x' = b + (a - b) \frac{x - x_{\min}}{x_{\max} - x_{\min}}$$
 (10)

The data in the electric theft evaluation index system are normalized to the interval [a, b]. Here, x' is the normalized data and x is the original data for each index. The maximum value for each index is x_{max} , and the minimum value is x_{min} . In this paper, [0,1] normalization is adopted [19].

3.2 GM(1,1) prediction

The grey prediction model GM(1,1) requires less modeling information, is easy to operate, and has high modeling accuracy. Therefore, it has been widely used in various prediction fields, and is an effective tool to deal with small-sample prediction problems[20]. The following carries out correlation analysis on the data of selected power users. The variation law of the original data system is generated by gray, and a data sequence with strong regularity is generated, so the corresponding differential equation model is established. By the above methods, we obtain the time point of power user data or the predicted value of [21] in the future and its development trend.

3.2.1 Parameter definition

The original data sequence is:

$$x^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots x^{(0)}(n))$$
(11)

The new sequence $x^{(1)}$ is generated from the growing sequence of original data:

$$\mathbf{x}^{(1)} = (\mathbf{x}^{(1)}(1), \mathbf{x}^{(1)}(2), \dots \mathbf{x}^{(1)}(n))$$
(12)

The grey derivative is defined as follows:

$$dx^{(1)}(k) = x^{(0)}(k) = x^{(1)}(k) - x^{(1)}(k-1)$$
(13)

Let $z^{(1)}$ approximate the sequence $x^{(1)}$:

$$z^{(1)}(k) = 0.5x^{(1)}(k) + 0.5x^{(1)}(k-1), (k = 1, 2, \dots, n)$$
(14)

$$z^{(1)} = \left(z^{(1)}(2), z^{(1)}(3), \cdots, z^{(1)}(n) \right)$$
(15)

Therefore, the grey differential equation, GM(1,1), is:

$$dx^{(1)}(k) + az^{(1)}(k) = b (16)$$

Let $x^{(0)}(k) + az^{(1)}(k) = b$, $x^{(0)}(k)$ is grey differentiable, then a is uniformly continuous, parameter $z^{(1)}(k)$ is white, b for the grey parameters. Entering the parameter k = 1,2,3,...,n into

$$x^{(0)}(k) + az^{(1)}(k) = b$$
 (17)

we obtain:

130

$$\begin{cases} x^{(0)}(2) + az^{(1)}(2) = b \\ x^{(0)}(3) + az^{(1)}(3) = b \\ \dots \\ x^{(0)}(n) + az^{(1)}(n) = b \end{cases}$$
(18)
Let $Y = (x^{(0)}(2), x^{(0)}(3), \dots, x^{(0)}(n))^{T}, u = (a, b)^{T},$
$$B = \begin{bmatrix} -z^{(1)}(2), 1 \\ -z^{(1)}(2), 1 \\ \vdots \\ -z^{(1)}(n), 1 \end{bmatrix}$$
(19)

Y is a data vector, B is a data matrix, where $z^{(1)}(i)$ is the row vector formed by the prediction data obtained from (14) and (15), and u is a parameter vector. Therefore, the GM(1,1) model can be expressed as a matrix equation Y = Bu. This can be obtained by the least squares method:

$$\hat{u} = (\hat{a}, \hat{b})^T = (B^T B)^{-1} B^T Y$$
(20)

3.2.2 Prediction calculation

By calculating the grey differential equation, we can get the prediction model:

$$\begin{cases} x^{(1)}(k+1) = (x^{(0)}(1) - \frac{u}{a})e^{-ak} + \frac{u}{a} \\ \hat{x}^{(0)}(k+1) = \exp\left(\hat{x}^{(k)}(k+1) - \hat{x}^{(1)}(k)\right), \\ k = 1, 2, 3 \cdots, n \end{cases}$$
(21)

At the same time, the original data sequence is obtained through a dimension reduction calculation:

$$\hat{\mathbf{x}}^{(0)}(\mathbf{i}) = \mathbf{x}^{(1)}(\mathbf{i}) - \mathbf{x}^{(1)}(\mathbf{i} - 1)$$
(22)

When using the GM(1,1) grey prediction model to carry out analysis of power theft, the sequence is as follows:

$$\mathbf{x}^{(i)} = (\mathbf{x}^{(i)}(1), \mathbf{x}^{(i)}(2), \cdots \cdot \mathbf{x}^{(i)}(n))$$
(23)

 $x^{(i)}$ is defined as the metering value of the electrical energy, so $(x^{(i)}(1), x^{(i)}(2), \dots, x^{(i)}(n))$ is the sequence formed by the influencing factors of electric energy measurement, and i is a time-related variable, so as to achieve time series prediction.

The next part will perform data processing on the influencing factor variables that need to be considered for prediction (as shown in Table 1). The predicted values of these variables are used in the prediction model of the GA-BP neural network.

Number	Electricity	Maximum	Line-loss in the	Three phase	Power	capacity	Power
	consumption	line-loss	platform area	unbalance rate	factor	ratio	stealing
1001	0.42412	0.98963	0.99522	0.75942	0.12934	0.82645	0
1002	0.39484	0.98412	0.98373	0.28438	0.23643	0.75244	0
1003	0.21948	0.99762	0.97633	0.26474	0.28233	0.82545	0
1004	0.47283	0.93266	0.99274	0.26154	0.25344	0.71265	0
1005	0.52934	0.97883	0.93744	0.57584	0.17464	0.82514	0
1006	0.54654	0.87354	0.91983	0.43495	0.27544	0.82645	0
1007	0.40285	0.92364	0.91954	0.74387	0.03634	0.87555	0
1008	0.21847	0.89736	0.87625	0.67352	0.12846	0.86456	0
1009	0.08475	0.93764	0.89736	0.83575	0.28746	0.68553	0
1010	0.22847	0.98964	0.91747	0.03754	0.27645	0.67329	0
1011	0.47383	0.97253	0.93865	0.53856	0.12764	0.67454	0
1012	0.68873	0.58373	0.92745	0.83263	0.24675	0.73585	0
1013	0.41847	0.96386	0.92764	0.54925	0.12644	0.97353	0
1014	0.28465	1.00000	0.95274	0.85636	0.37563	0.85347	0
1015	0.58736	0.98363	0.92424	0.74495	0.47563	0.59463	0
1016	0.56285	0.96342	1.00000	0.53956	0.42545	0.73554	0
1017	0.60745	0.91735	0.93856	0.35753	0.24633	0.67833	0
1018	0.57957	0.89622	0.95854	0.66833	0.26496	0.83685	0
1019	0.34756	0.87654	0.93246	0.75554	0.00000	0.68364	0
1020	0.05833	0.96252	0.93654	0.32755	0.27614	0.56533	0
1021	0.00000	0.87226	0.82558	0.02154	0.12755	1.00000	0
1022	0.27445	0.95252	0.85553	0.28564	0.28468	0.38653	0
1023	0.54123	0.93846	0.98365	0.44954	0.28645	0.68365	0
1024	0.72389	0.58323	0.95853	0.22385	0.25385	0.74638	0
1025	0.82745	0.00000	0.85853	0.13587	0.12865	0.65385	1
1026	1.00000	0.07116	0.85255	0.00000	1.00000	0.00000	1
1027	0.97847	0.04001	0.00000	0.04734	0.49766	0.49275	1

Table 1: Normalized sample data sheet

3.3 Genetic algorithm optimization

In order to improve the data obtained from the grey prediction model, this article uses the genetic algorithm optimization toolbox[22] from the University of North Carolina. In the genetic algorithm used in this paper, the specific parameter settings are: 100 iterations maximum, the population size is 500, the probability of selection is 0.9, the mutation probability is 0.09, and the crossover probability is 0.5. The changes in the optimal fitness and the average fitness of population individuals are shown in Figure 2.

After this process, the global optimal initial weight threshold can be obtained and used for BP network training. The BP network training parameters are set as follows: six input layer neurons, nine hidden layer neurons, 1 output layer neuron, a unipolar sigmoid function is used as the transfer function, the gradient descent trained function is used as the corresponding training function, use the largest number of training for 2000 times, the learning rate is 0.1, the target minimum mean variance is 0.003[23].



Figure 2. Fitness change over 100 iterations.

3.4 Model testing

The data obtained from the grey prediction model are used by the optimized BP neural network model, and the suspected power theft in the training data could be predicted more accurately. However, the predicted results exhibit overfitting after network optimization, so it is still necessary to use the data in the test samples and further verify the accuracy of the model. The test set data is imported to verify the model, and the predicted results are shown in Figure 3.



Figure 3. BP and GA-BP model test results.

As can be seen from Figure 3, the predictions of the BP neural network optimized by the GA are closer to the expected values than the unoptimized BP neural network.

4. Conclusion

Using the implementation in this paper, via the grey prediction model, we forecasted data related to power consumption on the time dimension. The prediction after data processing is input to the GA-BP neural network model in order to predict power theft. To a certain extent, the grey prediction model reduced the absolute error of the test data, and provides a long-term and short-term forecast analysis. By combining the grey

prediction model with the neural network analysis, the prediction and analysis on two levels make the model have better universality for the prediction of electric theft behavior, and provide a practical and effective method for related industries to carry out power theft detection. Future work will focus on accessing the large data collected by the information acquisition system and stored in a database, completing the assessment of the power consumption status of a large range of power users, and providing technical support for realizing the intelligent control of the power industry.

Reference

- SHI Yuliang, RONG Yiping, ZHU Weiyi. Stealing Behavior Recognition Method Based on Electricity Characteristics Analysis[J], Journal of Computer Research and Development, 2018,55(08):1599-1608.
- [2] LAI Zhe. Analysis of Anti-Stealing Electricity with Big Data Technology[J]. Electronic Test, 2019(16):50-51+38.
- [3] ZHANG Yumei. Intelligent Anti-Stealing Data Acquisition Technology and Application Analysis[J]. Technology Innovation and Application, 2017(20):41+43.
- [4] ZHANG Dongxia, MIAO Xin, LIU Liping. Research on Development Strategy for Smart Grid Big Data[J]. Proceedings of the CSEE, 2015,35(01):2-12.
- [5] ZHAI Haini, PANG Xuwei. YOU Mingyu. Analysis and Design of Application Value and Sharing Platform of Power Big Data[J]. Management and Administration, 2017(07):104-108.
- [6] ZHENG Jianning. Detection Method of Electricity Stealing Behavior Based on Deep Learning[J]. Information Technology, 2019(02):156-159.
- [7] CAI Jiarong, WANG Shunyi, WU Guangcai. The User's Electric Power Predicition and Electricity Inspection Plan Based on Machine Learning Research on Auxiliary Arrangement[J], Electronic Test, 2018(02):108-109.
- [8] LI Jianbin, WU Binbin, ZHU Yakui, et al. Risk Prediction and Prevention and Control of Customer Electricity Charge Based on Big Data Analysis[J]. Power Systems and Big Data, 2019,22(02):1-6. 2018(02):108-109.
- [9] LI Zhipeng, HOU Huiyong, JIANG Sifan, et al. Line Loss Calculation and Electricity Theft Analysis Based on Artificial Neural Network[J]. Southern Power System Technology, 2019,13(02):7-12+50.
- [10] Tianhua Liu, Shoulin Yin. An Improved Particle Swarm Optimization Algorithm Used for BP Neural Network and Multimedia Course-Ware Evaluation[J]. Multimedia Tools and Applications, 2017, 76(9).
- [11] LIN Jingdong, WU Xinyi, CHAI Yi, et al. Structure Optimization of Convolutional Neural Networks: A survey[J/OL]. Acta Automatica Sinica: 1-14[2020-02-21].https://doi.org/10.16383/j.aas.c180275.
- [12] Yilun Shang. Deffuant Model of Opinion Formation in One-Dimensional Multiplex Networks[J]. Journal of Physics A: Mathematical and Theoretical, 2015, 48(39).
- [13] XU Junxing, LI Chuan, LI Yingna. Power Stealing Analysis Based on Genetic Algorithm Optimized BP Neural Network[J]. Computer Engineering & Software, 2017,38(11):18-23.
- [14] Xiaohong Lu, Yongquan Wang, Jie Li, Yang Zhou, Zongjin Ren, Steven Y Liang. Three-dimensional coordinate measurement algorithm by optimizing BP neural network based on GA[J]. Engineering Computations, 2019, 36(6).
- [15] WANG Zhibin, KONG Yanan, LIU Yongcheng et al. Predicition of Junction Temperature for High Power LED by Optimizing BP Neural Based on Genetic Algorithm[J]. Journal of optoelectronics: laser, 2014,25(07):1303-1309.
- [16] MIN Jiangtao. Back Analysis Model of BP Neural Network Based on Genetic Algorithm Optimization[J]. Construction Materials & Decoration, 2018(45):296.
- [17] WANG Jingli, LI Liang, YU Lei, et al. Automated Fugl-Meyer Assessment Based on Genetic Algorithm and Extreme Learning Machine[J]. Journal of Computer Applications, 2014,34(03):907-910+914.
- [18] XU Zhi, LI Hongjiao, CHEN Jingjing. Predicition of User Stealing Behavior Based on Machine Learning[J]. Journal of Shanghai University of Electric Power, 2017,33(04):389-393.
- [19] YE Zhensai, XIAO Hui. Research on Transmission Line Visualization Model Based on Economic Big Data[J]. Power Systems and Big Data, 2019,22(09):28-34.
- [20] Hu Bin, Li Fang, Zhou Hou-Shun. Robustness of Complex Networks under Attack and Repair. 2009, 26(12):128901 (4pp).

- [21] DING Leiming, YANG Xiaolei, HUANG Jinbo, et al. Application of Improved Grey model Based on BP Neural Network Optimization in Electric Quantity Prediction[J]. Mechanical and Electrical Information, 2019(36):36-39.
- [22] SONG Xiuying. Application of GM (1,1) Model based on MATLAB in Economic Analysis[J]. Mathematics Learning and Research, 2011(11):93-95.
- [23] QIAN Guachao, MA Yutang, MA Yi, et al. Application of Genetic Algorithm in Fault Diagnosis of Grounding Grid Corrosion[J]. Yunnan Electric Power, 2016,44(02):113-117+129.