Hui-fang WANG, Chen-yu ZHANG, Dong-yang LIN, Ben-teng HE, 2019. An artificial intelligence based method for evaluating power grid node importance using network embedding and support vector regression. *Frontiers of Information Technology & Electronic Engineering*, 20(6):816-828. https://doi.org/10.1631/FITEE.1800146

An artificial intelligence based method for evaluating power grid node importance using network embedding and support vector regression

Key words: Power grid; Artificial intelligence; Node importance; Text-associated DeepWalk; Network embedding; Support vector regression

Corresponding author: Hui-fang WANG
E-mail: huifangwang@zju.edu.cn
ORCID: http://orcid.org/0000-0002-1483-364X

Motivation

1. Some large blackouts caused by the collapse of certain nodes aroused wide awareness of the need for protection of the key nodes in a power grid.

2. Evaluation methods using an index system can reflect the importance of a power grid node from many critical perspectives. However, power networks vary in scale, structure, and even tolerance to different consequences caused by the node fault.

3. Since the power grid is an important system in the format of the network along with the Internet, social networks, and so on, NE methods have the potential to be applied to deal with power grid problems.

Main idea

1. An NE method is used to vectorize features of power grid nodes with regard to their structural and electrical information. Based on this, a machine learning framework is established to evaluate node importance.

2. Then, a sample set can be established based on a set of steady-state and node fault transient simulations on different operational modes of each specific power network. It can reflect the relationship between features of nodes and their importance

3. Then, a machine learning method called support vector regression (SVR) is used to automatically mine the valuable information contained in these samples.

Method

1. Introducing feature extraction for power network nodes using the network embedding method, text-associated DeepWalk (TADW) algorithm, including details when applied in power network problems.

2. Introducing node fault consequence evaluation and fault sample set establishment methods.

3. First, we use feature selection methods to improve the performance of the whole framework. Then, we apply the machine learning SVR algorithm to the establishment of the model. Finally, we carry out various experiments based on the proposed framework.

Major results

1. The proposed model can extract reasonable feature expressions for the power grid nodes.

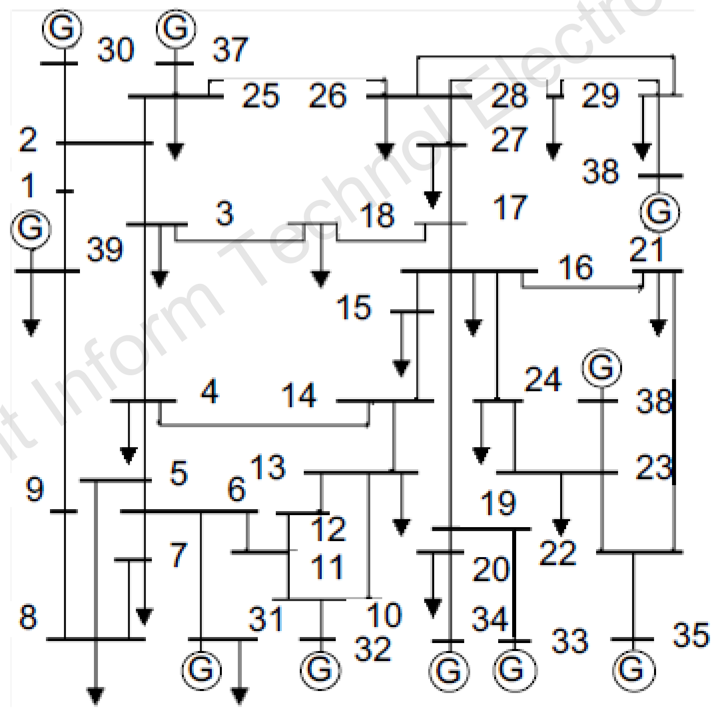

Fig. 3 IEEE-39 power system

Major results

2. Features calculated by the TADW algorithm can be further selected by some feature selection methods to improve the performance and efficiency of the machine learning framework.

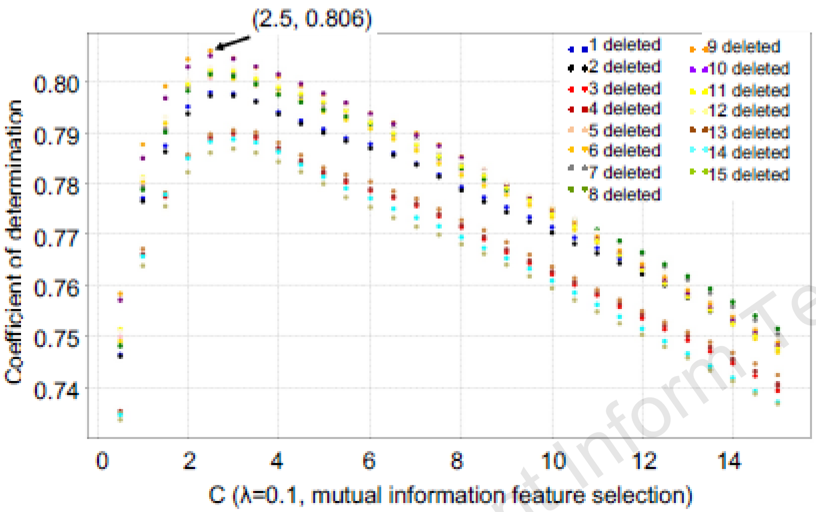

Fig. 4 CD-C curves when λ=0.1 and the mutual information method is used as the feature selection method

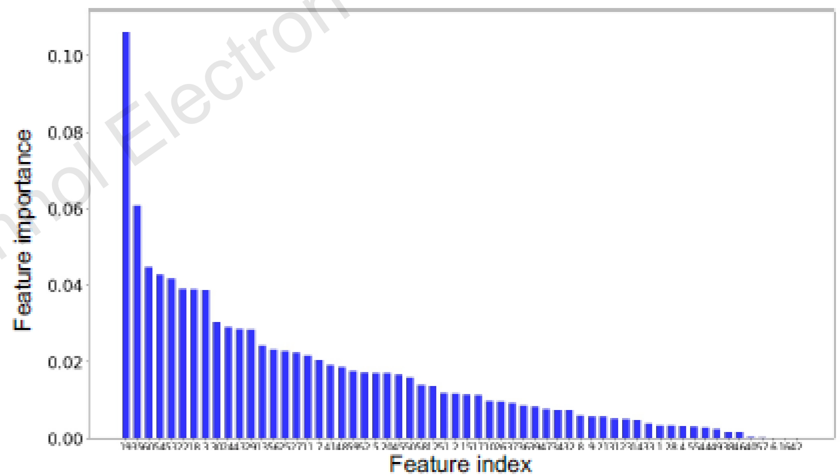

Fig. 5 Feature importance calculated by the GBDT model when λ=0.1

3. Based on the extracted features, machine learning methods can be applied to predict the fault consequence of the power grid nodes more precisely and effectively.

Major results

4. The proposed method can provide more reasonable ranking results of power grid nodes than the index-based method.

Table 4 The nodes with severe fault consequences according to PSAT simulation

Index of node that may cause a large power angle difference of generators	Maximum power angle difference (°)
19	6328
16	5945
34	5710
17	5213
22	4820
24	4602
29	4563
20	3584
38	2960
33	2840
25	1562
28	960

Table 3 Evaluation results of node importance ranking by three methods

Ranking	Node index		
	Method A	Method B	Method C
1	**19**	**16**	4
2	**16**	**17**	8
3	**22**	3	**16**
4	**34**	**34**	27
5	**17**	26	15
6	**20**	**24**	3
7	**24**	**25**	6
8	**38**	15	26
9	**29**	39	**24**
10	**33**	14	**17**
11	23	**28**	11
12	**25**	6	9
13	21	**19**	5
14	**28**	18	**20**
15	26	8	23
30	8	28	22
31	36	38	25
32	7	31	29
33	31	32	31
34	39	35	14
35	30	37	21
36	12	36	28
37	1	7	13
38	9	30	1
39	32	12	12

Method A: proposed method; Method B: electrical betweeness method (Xu et al., 2010); Method C: PageRank method (Li et al., 2014)

Conclusions

1. First, in terms of universality, under the framework in this paper, a particular sample set can be established for each specific power network. This sample set can reflect specific physical characteristics of the power network. The flexibility in the data labeling process taking account of the specific preference to different node fault consequences also guarantees universality. Thus, the sample set can avoid the possible bias of methods based on an index system given the diversity of power networks.

2. Second, in terms of calculation speed, the proposed method can perform rapid and efficient online calculation according to the operational information of the power grid. The time-consuming steady-state and transient simulation process, which aims at accumulating enough samples for the SVR model, is performed offline.