

Intelligent Substation Virtual Terminal Information Flow Classification Method based on Spectral Clustering Algorithm

Xianzheng Feng, Xiaojun Liao, Yalei Liu and Junnan Liu

State Grid Sichuan Technical Training Center, Chengdu 610072, China

fengxianzheng@qq.com; 15208318375@qq.com; 601508414@qq.com; 349338787@qq.com

Keywords: Spectral clustering; Intelligent substation; Virtual terminal; Information flow

Abstract. In this paper, an intelligent substation virtual terminal information flow method based on spectral clustering algorithm is proposed. By modeling the information flow of the virtual substation virtual loop, the IED equipment relation table is obtained, and an undirected graph is generated, which uses the spectral clustering algorithm. The intelligent substation virtual loop information flow is classified and analyzed, and the correct rate of classification is verified. The simulation results show that the method can classify the information flow between IED devices in the intelligent substation. The algorithm has fast classification and correct classification. The advantages provide a reference for the maintenance personnel to debug.

1. Introduction

There are many virtual circuit information in intelligent substation, which is easy to cause intelligent substation configuration file error, which causes equipment failure. It is necessary to explore a smart substation virtual loop information flow classification method. At present, the methods applied to classification mainly include K-mean, decision tree classification method, naive Bayesian classification method, classification method based on support vector machine, neural network classification method and fuzzy classification method. These algorithms do not effectively classify multidimensional data. At present, the intelligent substation equipment has a complicated structure, flexible interconnection between devices, and virtual circuit information has a large amount of information. The spectral clustering algorithm is based on the spectral theory. Compared with the traditional clustering algorithm, it has the advantage of being able to cluster on the sample space of arbitrary shape and converge to the global optimal solution. The algorithm is based on the given IED connects the relation matrix, and calculates the eigenvalues and eigenvectors of the matrix, and then selects the appropriate feature vector to cluster different data points. On the one hand, the layout results retain the advantages of the spectral clustering algorithm layout. The spectral clustering only needs the similarity matrix between the data, so it is effective for the clustering of sparse data. This traditional clustering algorithm such as K-Means is difficult to achieve. Because of the use of dimensionality reduction, the complexity of processing high-dimensional data clustering is better than the traditional clustering algorithm. The method is beneficial for judging whether the information flow between the IED devices in the intelligent substation is correctly classified, and the algorithm has the advantages of fast classification and correct classification.

2. Spectral Clustering Algorithm

Spectral clustering is a very useful tool for data analysis. It is a typical unsupervised learning method. It divides the data set into a number of clusters, so that the elements in the same cluster have high similarity, while the similarity between different clusters is low. Spectral clustering uses the eigenvectors of the correlation matrix for clustering. It is a clustering method based on graph theory. The essence of spectral clustering algorithm is to transform the clustering problem into the optimal partitioning problem of graphs. It is a point-to-point clustering algorithm and has a good application prospect for data clustering.

Aiming at the shortcomings of existing spectral clustering algorithms, and the disadvantages of some control points being too scattered, an improved spectral clustering algorithm is proposed. The spectral clustering algorithm treats each IED object in the dataset as the vertex V of the graph, and quantizes the connection relationship between the IEDs as the weight of the connected edge E of the corresponding IED device, thus obtaining a similarity-based undirected weighting. Figure $G(V, E)$, where V is the set of vertices of the graph, each vertex in the set represents a data point of the cluster, E is the edge set of the graph, and the weight of each edge is the associated of the two data points. The degree, the higher the weight, indicates that the similarity between the two vertices is higher, so the clustering problem can be transformed into the partitioning problem of the graph. The optimal partitioning criterion based on graph theory is to maximize the internal similarity of subgraphs and minimize the similarity between subgraphs.

3. IED Device Model Construction

Before classifying the intelligent substation information connection diagram, it is first necessary to model the network connection diagram according to the connection relationship between the researched IEDs. The modeling mainly considers the attribute information about the node and the connection information between the nodes. This information has a great influence on the subsequent graph cluster analysis. Taking the virtual terminal connection network between the simple intelligent substation IED as an example, the basic method And the steps are as follows.

3.1 IED Node Attribute Information

There are many attributes of the IED node. In the SCD file, the basic name, attribute, manufacturer, version, etc. are included. In order to better classify the IED nodes, other attribute information needs to be considered. First, the IED device type can be set according to the node IED name prefix: protection, measurement and control, merging unit, intelligent terminal and other types. Secondly, the attributes can be set according to the voltage level: 110kV, 220kV, 10kV, main transformer node, etc. In addition to this, the relevant calculation data can also be used as node attributes, such as the number of IED nodes (number of IED connections), the degree of access (indicating the number of IED subscriptions), and so on.

3.2 IED Connection Attribute Information

The virtual connection between the IED constitutes the edge of the network diagram. Since the virtual connection of the IED includes input and output, some are one-way connections, such as merging unit to protection, and some are bidirectional connections, such as protection and connection between intelligent terminals. So this diagram is a directed graph, and you need to consider the connection direction property. Secondly, the basic types of connections can be considered according to the type of signal transmission, such as Goose signal, SV signal, MMS signal. Of course, other connection properties, such as control signals and interlocking signals, can be considered for more advanced analysis applications.

3.3 IED Node Connection Table

After obtaining the data according to the IED node attribute information and the connection attribute information as described above, the node-connection table of the network connection diagram can be constructed. A typical table is as follows:

Table 1 Node Table

No.	IED Name	IED Description	IED Type	Voltage Level
0	pzb1a	Main transformer protection	protection	220kV
1	pmx220a		Merging unit	110kV

Table 2 Branches

Source node ID	destination node ID	direction	edge number	branch name	weight	Information Type	signal Types
0	1	Directed	0	ds1-ds2-5	2	M1	Trip
1	2	Directed	1	m1-m2-5	1	G1	Voltage
2	0	Directed	2	input-ds1-5	1	2	Current

4 Algorithm Flow

The specific process of the algorithm is as follows:

The first step is to construct a similarity matrix S between the various IED devices based on the connection mode between the various IED devices.

In the second step, the adjacency matrix W is constructed according to the similarity matrix S , and the degree matrix D is constructed. The method of adjacency matrix is K-proximity method. The KNN algorithm is used to traverse all sample points, and the nearest k points of each sample are taken as neighbors, only between the k points closest to the sample distance. A degree matrix, that is, a diagonal matrix obtained by summing each row (or each column) of the similarity matrix.

The third step calculates the Laplacian matrix L and constructs a normalized Laplacian matrix. Laplacian matrix $L = D - W$, calculate the normalized Laplacian matrix:

The fourth step calculates the feature vector corresponding to each of the smallest eigenvalues, normalizes the feature vector, and finally forms the feature matrix F of the dimension.

In the fifth step, each row in the feature matrix F is used as a sample of a dimension, and a total of n samples are clustered by K-Means clustering method.

The sixth step outputs the visualization according to the clustering result. Use the plot function to draw the optimized node coordinates.

5 Simulations

The algorithm proposed in this paper is programmed by Matlab. The information flow of SCD file in a smart substation is visually classified. The substation is a typical 500kV intelligent station. The main wiring is three-thirds of the wiring. The number of IEDs is 358. IED There are 604 information flow relationships. Figure 2 shows the layout classification using the K-mean algorithm, Figure 3 shows the layout classification results using the spectral clustering algorithm, and Figure 4 shows the process data, including the similarity matrix, Laplacian matrix, eigenvectors, and eigenvalue normalization. The results show that the spectral clustering algorithm classification is correcter than the K-mean algorithm. This algorithm is convenient for viewing the classification between nodes. The program runs as figure 2, figure 3, figure 4.

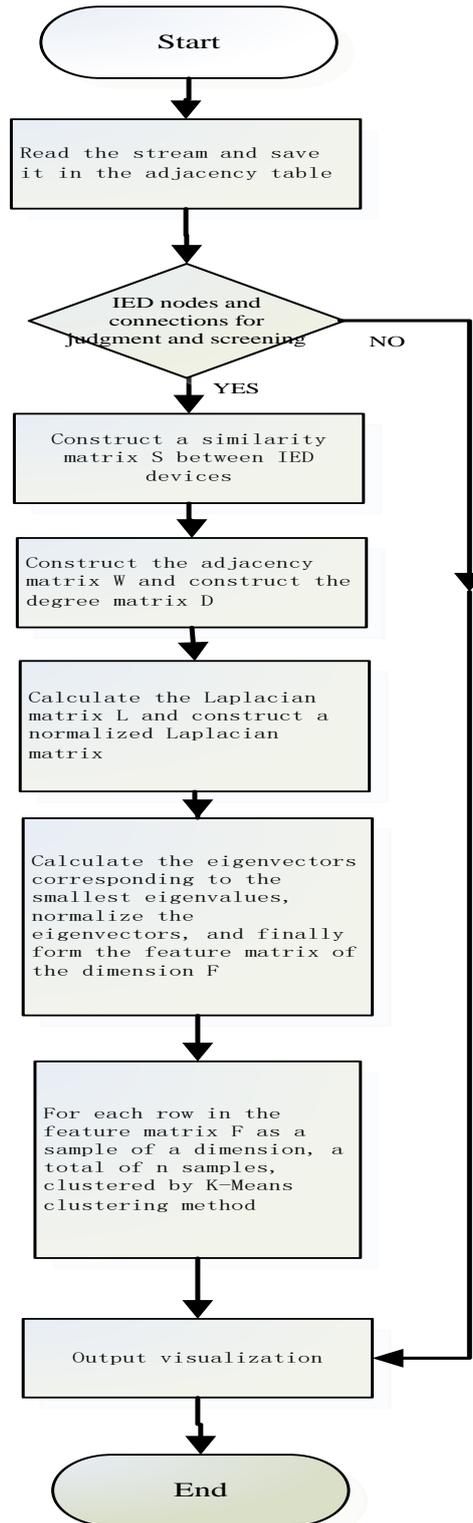


Figure 1. Spectrum clustering algorithm flow chart

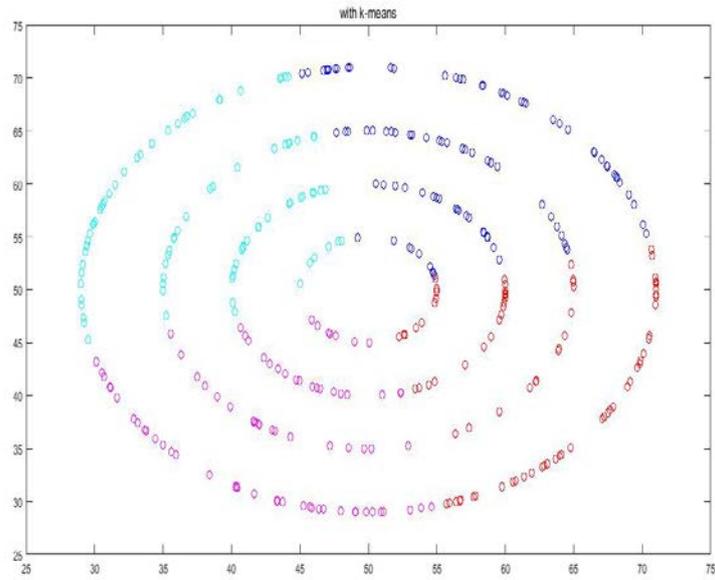


Figure 2. K-mean algorithm layout classification

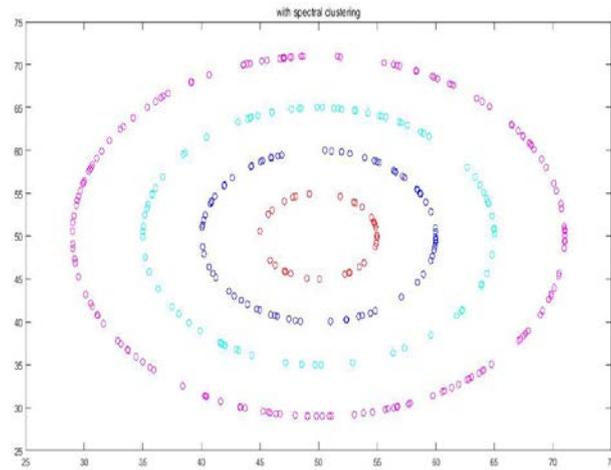


Figure 3. Layout classification of spectral clustering algorithm

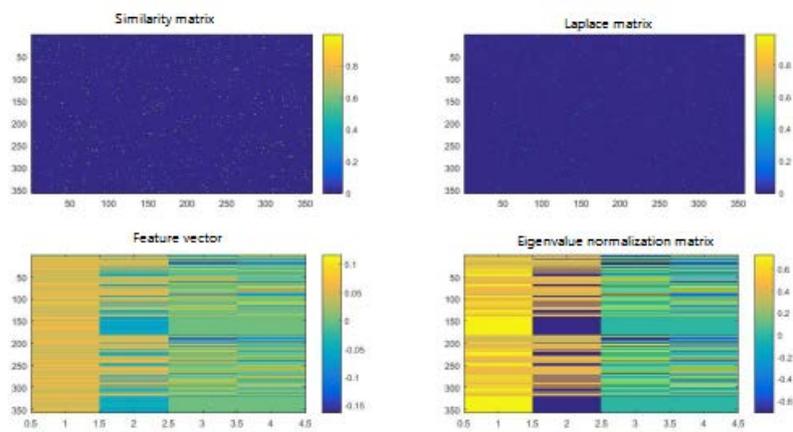


Figure 4. process data

6 Conclusions

By modeling the IED equipment relationship model of the intelligent substation, the spectral clustering algorithm is used to classify the relationship diagram between the information flows, which proves the advantages of the spectral clustering algorithm on the classification of the intelligent substation IED equipment relationship: spectrum aggregation Classes only require a similarity matrix between the data, so they are very efficient for clustering of sparse data. This traditional clustering algorithm such as K-Means is difficult to achieve. Because of the use of dimensionality reduction, the complexity of processing high-dimensional data clustering is better than the traditional clustering algorithm. At the same time, the clustering effect depends on the similarity matrix, and the final clustering effect obtained by different similar matrices may be different. The example results demonstrate the feasibility of the method and improve the accuracy of the classification results.

References

- [1]Gao Xu, Yu Qingguang, Ma Yingxin, Tang Xiaobing. Research on the corresponding method of optical substation optical real circuit based on graph theory deep traversal algorithm [J/OL]. Electrical measurement and instrument: 1-6
- [2]Ching Yee Yong, Kim Mey Chew, Rubita Sudirman. Human wrist motion classification using decision tree and principle component analysis[J]. Materials Today: Proceedings, 2019, 16.
- [3]Wu Chenyang, Wang Xiaohua, Wang Jinbo, Song Jiawei. Research on Knowledge Classification Based on KNN and Naive Bayesian Algorithms[J]. Journal of Physics: Conference Series, 2019, 1213(3).
- [4]Utpal Barman, Ridip Dev Choudhury. Soil texture classification using multi class support vector machine[J]. Information Processing in Agriculture, 2019.
- [5]Ciritzis Alexander, Rossi Cristina, Eberhard Matthias, Marcon Magda, Becker Anton S, Boss Andreas. Automatic classification of ultrasound breast lesions using a deep convolutional neural network mimicking human decision-making.[J]. European radiology, 2019, 29(10).
- [6]S. Shahnawazuddin, Nagaraj Adiga, B Tarun Sai, Waquar Ahmad, Hemant K. Kathania. Developing speaker independent ASR system using limited data through prosody modification based on fuzzy classification of spectral bins[J]. Digital Signal Processing, 2019, 93.
- [7]Daryl R. DeFord, Scott D. Pauls. Spectral clustering methods for multiplex networks[J]. Physica A: Statistical Mechanics and its Applications, 2019, 533.