Journal of computer science and software applications
https://www.mfacademia.org/
ISSN:2377-0430
Vol. 2, No. 4, 2022

Enhancing Power Communication Network Efficiency: A Hybrid Approach Using DBSCAN Clustering and Sliding Window Techniques for Alarm Data Reduction

Sisi Wang
St. Catherine University
Wangsisi123@gmail.com

Abstract:

The original alarm data in power communication networks exhibit issues such as discreteness, redundancy, and temporal asynchrony. To address these challenges, this paper proposes a method that integrates DBSCAN clustering with a sliding window approach. Initially, the DBSCAN clustering algorithm is utilized to effectively tackle the problem of discreteness. Subsequently, the sliding window technique is employed to mitigate alarm redundancy and temporal asynchrony. Simulations demonstrate that the proposed method significantly reduces the number of alarms, achieving an average reduction of up to 47.68%.

Keywords:

Power communication network, alarm data, DBSCAN, sliding window.

1. Introduction

The power communication network plays a major role in ensuring the safe and stable operation of the electric power system. The large-scale and complicated power communication network generates a large amount of operation and maintenance data in the process of daily network maintenance [1] [2]. In the power communication network, equipment and line faults are manifested in the form of alarm messages. The power communication network generates a large number of alarm messages on a daily basis, and it is a pressing need to determine how to mine valuable information from a massive amount of alarm information. It appears to be difficult for the conventional network-based management system to cope with an enormous amount of alarm data. Meanwhile, due to the heterogeneity of the power communication network and the application of power communication facilities supplied by different manufacturers, it is inevitable for the alarm information generated by the power communication network to show discreteness, redundancy, and time asynchronism. In this case, not all of the alarms in the huge alarm data are valuable for operation and maintenance. Besides, some noise will case disruption to the maintenance work. In reference[3],a method is proposed to process the alarm data with reliance on search trees. By aggregating the alarm data falling into the same category and replacing this category with less information, the number of alarms can be reduced effectively. In reference[4],a warning mechanism is introduced into the integrated management system for the electric power communication network, which provides a theoretical basis for future fault handling. However, neither of them have given consideration to the processing of the original alarm data. Therefore, it is necessary to develop a sort of technology that can reduce the size of alarm data, while improving the availability of alarm data effectively.

2. Alarm Data Processing of Power Communication Network based on DBSCAN and Sliding Window

2.1. Alarm Data Processing of Power Communication Network based on DBSCAN

Based on density clustering,Ester proposed DBSCAN algorithm[5][6],which can automatically determine the number of clusters, and has two parameters, including the clustering radius Eps and the minimum density threshold MinPts, The clustering radius Eps determines the size of the cluster, and the MinPts determines whether the randomly selected object P is the core object. When the DBSCAN clustering algorithm is applied to process the original alarm data of the power communication network, it is sufficient to input the original alarm data set D, the clustering radius Eps, and the minimum density threshold MinPts, as the cluster analysis can be completed automatically. The clustering approach involves two parts: One is the cluster discovery and the other is neighborhood expansion. The process of DBSCAN is illustrated in Figure 1:

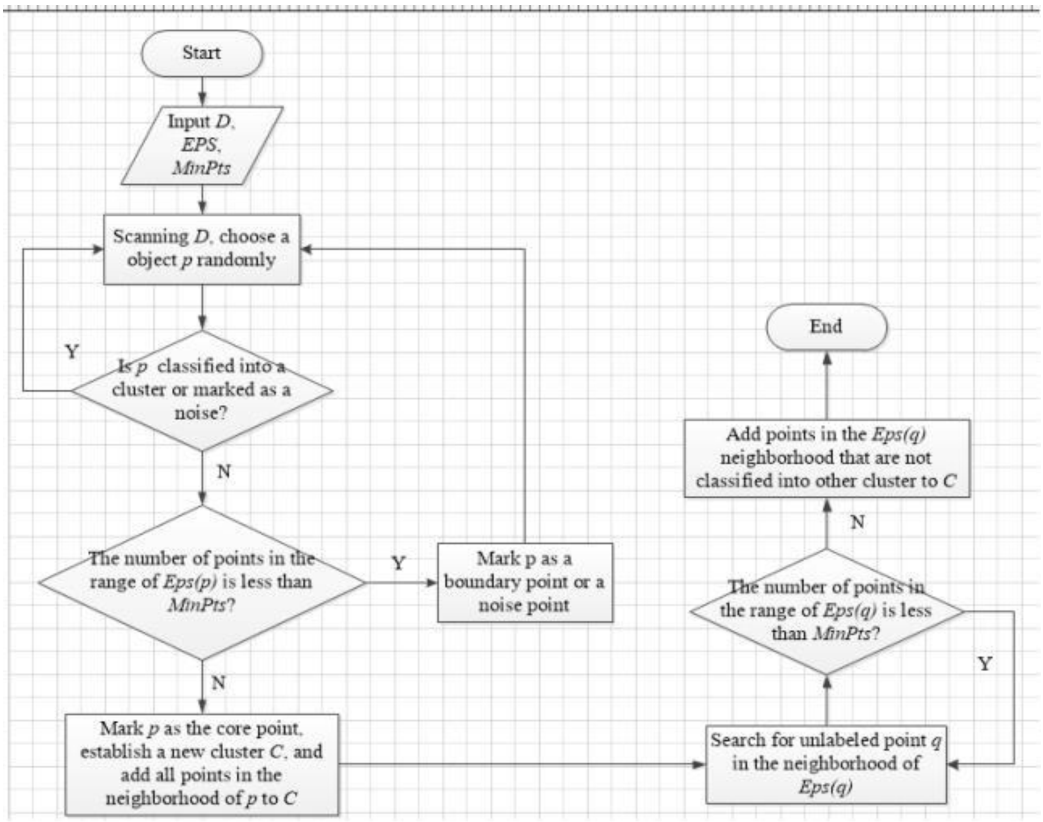

Figure 1. DBSCAN algorithm execution process

2.2. Alarm Data Processing of Power Communication Network based on SlidingWindow

Sliding windows have a widespread applications in the processing of time-based streams[7][8], while power communications network alarms data are typically time-series data. The sliding window contains two parameters, which are the sliding window size W and sliding step S. As shown in Figure 2,A to F are different alarm transactions, the sliding window is 5 and sliding step is 2.After sliding widow process we can obtain the following data {(ABC) (CBD) (BCDE) (DBEF) (EF) }. Compared with the original data, there are no duplicate alarm transaction itemsin each window, as a result of which alarm transactions redundancy is eliminated.

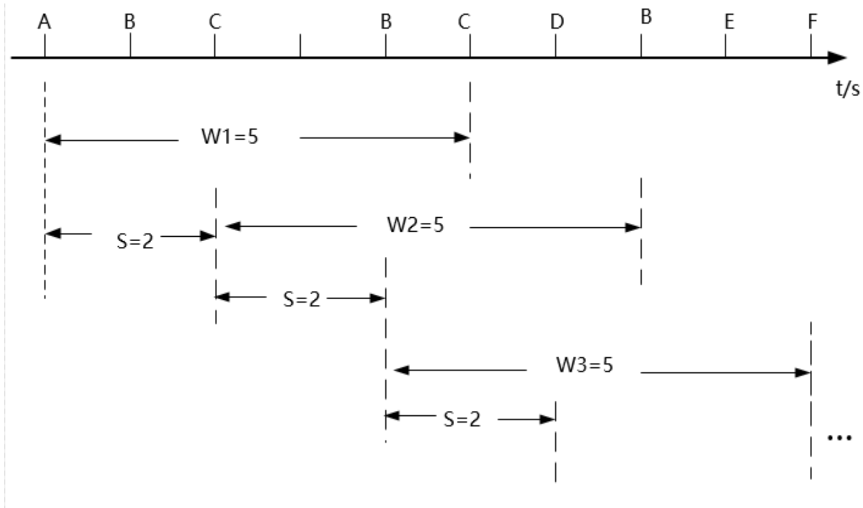

Figure 2. Example of sliding window

2.3. Alarm Data Processing of Power Communication Network based on DBSCAN and Sliding Window

As for such problems as discreteness and redundancy encounter by the original power alarm data, we choose to combine clustering and sliding window, and proposed the processing of power alarm data processing based on clustering and sliding window. The main steps as follows: Step1: In order to address the problem of discrete of raw alarm information, we choose the DBSCAN clustering algorithm to cluster the time period of the alarm information. Which is purposed to find the alarm-intensive time periods, and to remove the scattered alarm information.

Step2: After clustering, we take the sliding window approach to process the clustered alarm data. Base on the principle that" alarm data appears in the same window are regarded as occurring at the same time, and the same alarm information recorded only once" alarm data redundancy is eliminated and time asynchronism is resolved. The realization process of the method is illustrated in Figure 3:

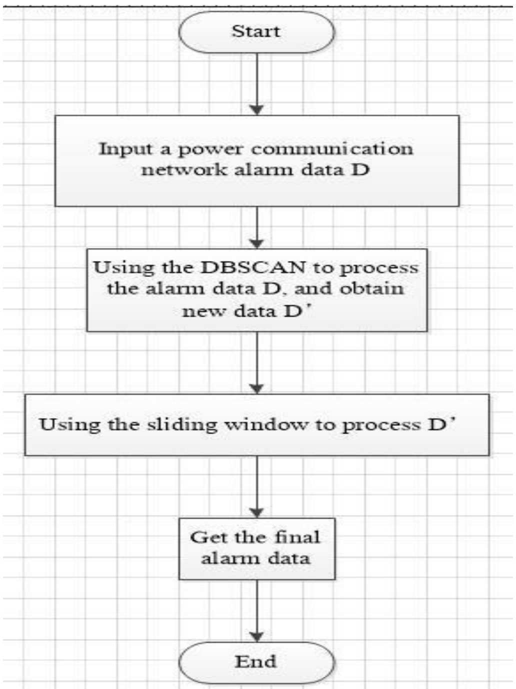

Figure 3. The operation process of proposed method

3. Experimental and Simulation

In order to demonstrate the advantage of the method proposed in this paper in the processing of power communication network alarm data, we select three groups of power communication network alarm data on a random basis, and apply the method proposed in this paper, sliding window and clustering to deal with the alarm data. Finally, we compare the number of alarms produced using the three methods. Table 1 lists the alarm numbers of the selected alarm data.

Table 1. Selected alarm data

Number of groups	Number of alarm transactions
1	165
2	105
3	126

The third-party module in machine learning Scikit-learn (sklearn) can easily and efficiently achieve cluster analysis[9][10]. Our technical route is to achieve cluster analysis by calling Sklean.cluster.DBSCAN. Before calling this function, we must initialize parameters of the DBSCAN clustering algorithm. The parameters are initialized as follows:

DBSCAN(eps=8,min_samples=10,metric='euclidean',metric_params=None,algorithm= 'auto',leaf_size=30,p=None,n_jobs=None).

In python, sliding window analysis can be achieved by calling the rolling function in the pandas module. At the same time, the two parameters of the sliding window are 8 and 2. The experimental results are shown below. Among them, Figure 4 is the result of processing the original power communication network alarm data using the clustering and sliding window (DBSW)method proposed in this paper, Figure 5 is the result of using DBSCAN clustering algorithm (DB), and Figure 6 is the result of using the sliding window (SW), and Figure 7 is the comparison of the results of the three approaches. From Figure 7, we can see that the method proposed in this paper has distinct advantages in processing the raw power alarm data.

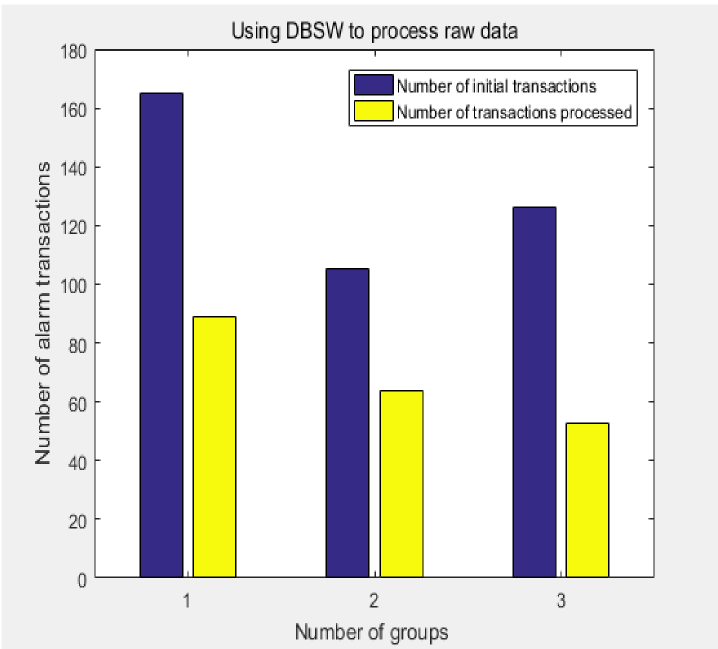

Figure 4. Using DBSW to cope with alarm data

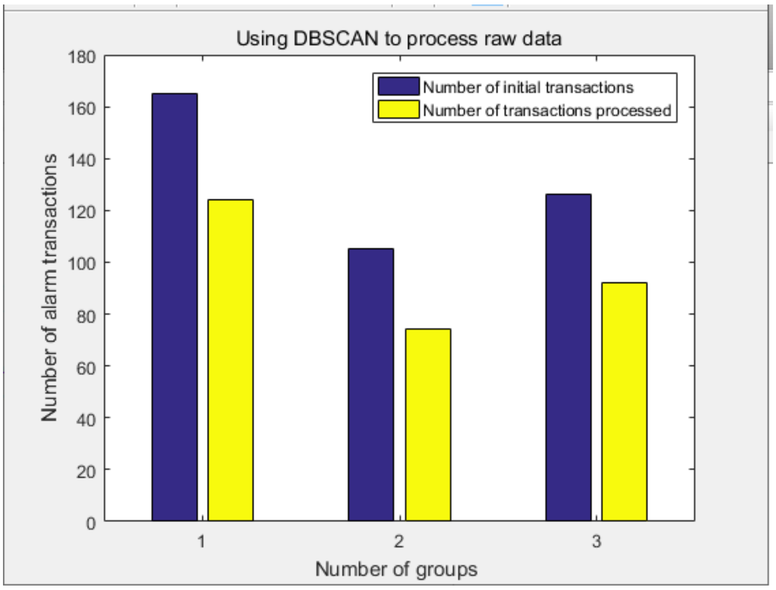

Figure 5. Using DBSCAN only to cope with alarm data

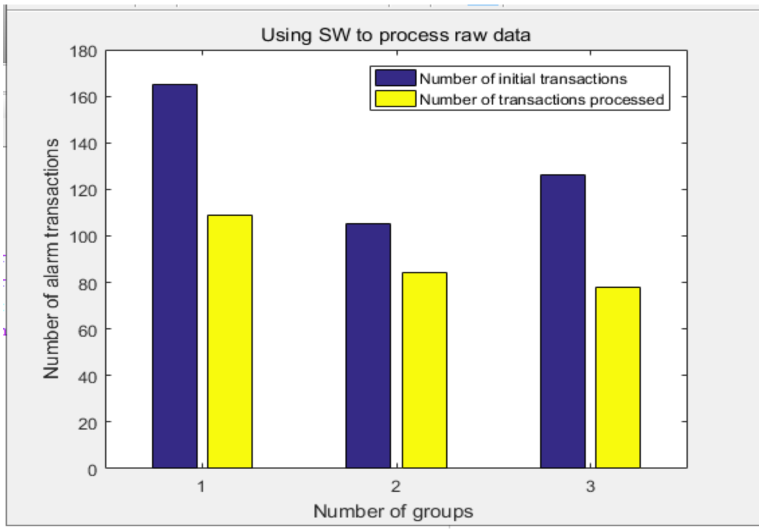

Figure 6. Using SW only to cope with alarm data

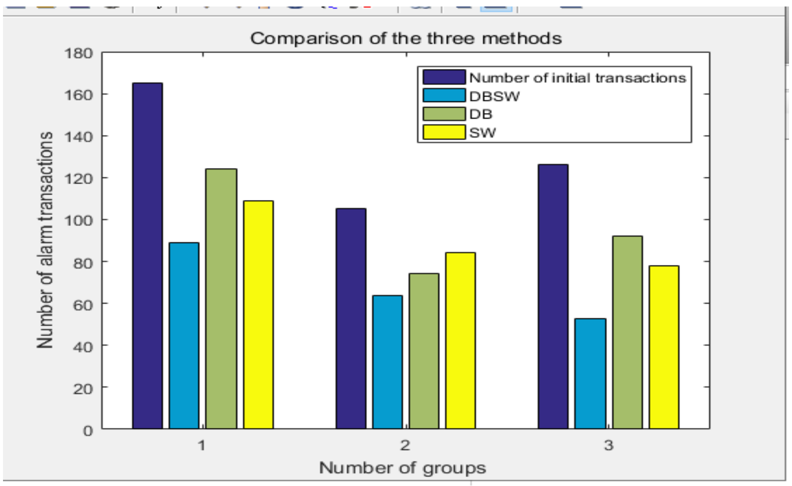

Figure 7. Comparison of the three methods

In order to further demonstrate that the method proposed in this paper is effective in reducing the number of alarms. Table 2 indicates the number of remaining alarms in each group of raw alarm data after processing by three methods.

Table 2. Comparison of three methods

Three kinds of methods	Group one		Group two		Group three		Average decrease (%)
	Before	After	Before	After	Before	After	
DBSW	165	89	105	64	126	53	47.68
DB	165	124	105	72	126	92	27.10
SW	165	109	105	89	126	78	29.81

From Table 2, it can be seen that after a combination of clustering and sliding windows to process the alarm data of power communication network, the comparison made with the other two methods reveals that the average number of alarms is reduced by as much as 47.68%, which evidences the effectiveness of the proposed method in reducing the number of alarms in the power communication network and in improving the availability of alarm data for the power communication network.

4. Conclusion

In order to resolve various problems with the alarm data in the power communication network, This paper proposes a method that combines clustering and sliding windows. Firstly, the problem of discrete alarm in power communication network is solved by clustering, and then sliding window is used to solve the problems of redundancy and time asynchronous in the power communication network alarm data. Finally, the superiority of the method proposed in this paper when processing the alarm data of electric power communication network is verified by experiments.

References

[1] X. Miao, J. Li and X. Chen, "Research on reliability assessment method of electric power communication network and application," 2012 China International Conference on Electricity Distribution, Shanghai, 2012, pp. 1-3.

[2] X. Zhang et al., "Architecture Design of Electric Power Emergency Communication Based on 4G LTE Network," 2019 IEEE 4th International Conference on Cloud Computing and Big Data Analysis (ICCCBDA), Chengdu, China, 2019, pp. 600-603.

[3] Wang Baoyi, Guo Yawei, Shi Zhancheng, Zhang Shaomin. Research of alarm correlation method based on dependency search tree in electric power communication network[J]. Realy,2008(06):59-64.

[4] Li LiDa. Design and Realization of fault alarm mechanism in network management system of power communications network[J]. Telecommunications for electric power system,2006(12):13-16.

[5] A. Smiti and Z. Elouedi, "DBSCAN-GM: An improved clustering method based on Gaussian Means and DBSCAN techniques," 2012 IEEE 16th International Conference on Intelligent Engineering Systems (INES), Lisbon, 2012, pp. 573-578.

[6] Huan Yu and Wenhui Zhang, "DBSCAN data clustering algorithm for video stabilizing system," Proceedings 2013 International Conference on Mechatronic Sciences, Electric Engineering and Computer (MEC), Shengyang, 2013, pp. 1297-1301.

[7] Wang Dewen, Yang Liping. Streaming processing method and condition monitoring anomaly detection of smart grid big data [J]. Automation of Electric Power Systems, 2016, 40 (14): 122-128.

[8] Li Guohui, Chen Hui. Mining frequent patterns in an arbitrary sliding time window of a data stream[J]. Journal of Software, 2008 (10): 2585-2596.

[9] Information on https://scikit-learn.org/stable/modules/clustering.html#dbscan.

[10] Huang YongChang. Machine learning by scikit-learn: Algorithms and Practices[M]. Beijing: China Machine Press, 2018.