

A Method of “Hard-Spot” Threshold Extraction of Pantograph-Catenary Current Collection Parameters of Train-Set

Zhigang Liu, Zhiwei Han, Gang Chen, and Jian Xie

Abstract—The “hard-spot” of catenary is an intrinsic structural defect, and it is also an important index for evaluating the performance of pantograph-catenary current collection. In the traditional method, whether the parameters of “hard-spot” are transfinite only depends on the maximum acceleration in each space, which often results into incorrect judgment. First, the fuzzy c-means (FCM) is adopted to analyze and excavate the data of pantograph-catenary current collection parameters of train-set, and extract the threshold of “hard-spot” in the paper. Then, the reasonable performance index of pantograph-catenary current collection and the evaluation criterion based on three thresholds of “hard-spot” are presented. This criterion fully covers three feature parameters of “hard-spot” in order to avoid the defects with only one index for the transfinite judgment. Experimental result shows that the criterion is reasonable and effective based on the comparison between the sampled data the cluster results.

Index Terms—Train-set, threshold, FCM, hard-spot, pantograph-catenary, current collection.

I. INTRODUCTION

The “hard-spot” means the phenomenon of poor catenary’s flexibility or the concentration mass at catenary joint, section insulator, positioning equipment or mid-point anchor. If there are “hard-spots”, the operation conditions of electric locomotive will become bad, and the mechanical damage and electric arc burn will occur easily, which perhaps induce the faults of pantograph and catenary. At the same time, the local wear of catenary will become bigger and its service life is reduced. With the increasing speed, the influence of “hard-spot” for the pantograph-catenary relation becomes more significant. Hence, the detecting and processing of “hard-spot” is very important for the catenary’s safety in high-speed electric railway traction system [1]-[3]. In order to find the “hard-spot” of catenary, some dynamic detection and diagnosis equipments with the vertical impulsive acceleration of pantograph head are developed [4]-[5]. For the further analysis of detected data, there are several methods such as correlation, classification, sequence and clustering analysis. However, for the correlation analysis, the correlation rules are mainly used for data mining. The

classification analysis is suitable for data analysis with example database, and sequence analysis is suitable for the analysis of pre and post relationship between the sequences [6]. In the paper, the fuzzy c-means (FCM) is adopted to analyze and process the detected data of catenary “hard-spot” from the integrated inspection vehicle (Train-set CRH2-010) during the course of Sixth Raising-Train-Speed in China, and extract the “hard-spot” threshold.

II. FUZZY C-MEANS

Fuzzy c-means is a method of clustering which allows one piece of data to belong to two or more clusters. This method (developed by Dunn in 1973 and improved by Bezdek in 1984 [7]) is frequently used in pattern recognition. It is based on minimization of the following objective function:

$$J(U, V) = \sum_{i=1}^c \sum_{k=1}^n (\mu_{ik})^m (d_{ik})^2 \quad (1)$$

where $X = \{x_1, x_2, \dots, x_n\}$ is data set, x_k is the k th of d -dimensional measured data, $V = \{v_1, v_2, \dots, v_c\}$ is cluster center data set, $2 \leq c \leq n$, v_k is the d -dimension center of the cluster, d_{ik} is the distance between data x_k and center v_k , $m \in (1, \infty)$ is fuzzy factor. When $m \rightarrow \infty$, $J(U, V) \rightarrow \infty$. When $m = 1$, FCM is c-means algorithm.

The algorithm is composed of the following steps:

- 1) Deciding the number of cluster class, fuzzy factor, and stop condition of iteration ε . Initializing partition matrix $U^{(0)}$, $V^{(0)}$, and $r, r = 0, 1, \dots, n$ denotes the step number.
- 2) Computing c centers $\{V_i^{(r)}\}$ in each step.

$$v_i = \frac{\sum_{k=1}^n (u_{ik})^m x_k}{\sum_{k=1}^n (u_{ik})^m}, \quad (i = 1, \dots, c) \quad (2)$$

- 3) Computing the membership matrix $U^{(r)}$ in each step.

When $d_{ik} = \|x_k - v_i\| \neq 0$,

$$u_{ik}^{(r+1)} = \left[\sum_{j=1}^c \left(\frac{d_{ik}}{d_{jk}} \right)^{2/(m-1)} \right]^{-1} \quad (3)$$

Manuscript received August 10, 2012; revised September 20, 2012. This work was supported in part by National Nature Science Foundation of China (U1134205, 51007074), Railway Ministry Science and Technology Research and Development Program (2011J016-B) and Fundamental Research Funds for Central Universities (SWJTU11CX141) in China.

The authors are with the School of Electrical Engineering, Southwest Jiaotong University, 610031, Chengdu, Sichuan, China (e-mail: liuzg@swjtu.cn, illusioncn@163.com, chengangdadada@163.com).

When $d_{ik} = \|x_k - v_i\| = 0$, let $I_k = \{i | 2 \leq c \leq n; d_{ik}^{(r)}\}$, $\tilde{I}_k = \{1, 2, \dots, c\} - I_k$, then for every $i \in \tilde{I}_k$, satisfies

$$u_{ik}^{(r+1)} = 0, \sum_{i \in \tilde{I}_k} u_{ik}^{(r+1)} = 1 \quad (4)$$

- 4) If $\|U^{(r+1)} - U^{(r)}\| \leq \varepsilon$, stop the iteration. Otherwise, let $r = r + 1_k$ and go back second step.

In fact, ε is the convergence threshold which can influence the clustering precision and speed. In addition, the convergence performance of algorithm is decided by ε [8]-[9].

III. PRE-PROCESSING DATA OF “HARD-SPOT”

The catenary “hard-spot” data include maximum acceleration, peak factor and effective acceleration of “hard-spot”. The data are from the integrated inspection vehicle (Train-set CRH2-010) during the course of Sixth Raising-Train-Speed in China. There are 7068 sets data. If the train’s speed is less 30 km/h, the possibility of “hard-spot” occurrence is very small. Hence, the sampling data in which the train’s speed is more than 30 km/h are chosen.

According to the detected data, if the speed becomes big, the data of “hard-spot” will become big gradually. If the data are divided into two classes, there will be many data in each class, which can reduce the clustering performance and result into the judgment failure of “hard-spot”. Hence, it is very important to pre-process the sampling data. The course of pre-processing includes speed removal and linear regression. The purpose of speed removal is to reduce the speed’s influence on the data of “hard-spot”. The linear rule of “hard-spot” data with speed can be found with linear regression.

The detailed steps are listed below.

- 1) The speed data in data matrix are divided into 10 parts. In each speed-duration, the speed average value can be obtained.
- 2) All data in corresponding speed duration is summed, divided by the number of data, and the average value of parameter in each speed duration can be obtained.
- 3) The linear regression curve of parameter changing rule can be obtained with linear regression method.
- 4) Because the maximum acceleration and effective acceleration of “hard-spot” are absolute values and peak factor is relative value, the linear regression of peak factor is not necessary.

The results are shown in Fig. 1-3.

IV. REALIZATION OF “HARD-SPOT” THRESHOLD

After the pre-processing, the equivalent point of parameter at speed average value can be obtained. The parameter distributions are shown in Fig. 4 (a-c).

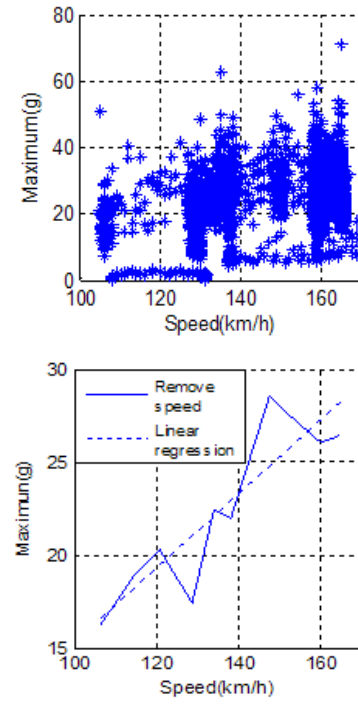


Fig. 1. (a) Maximum acceleration distribution of “hard-spot” (b) Speed changing trend of maximum acceleration.

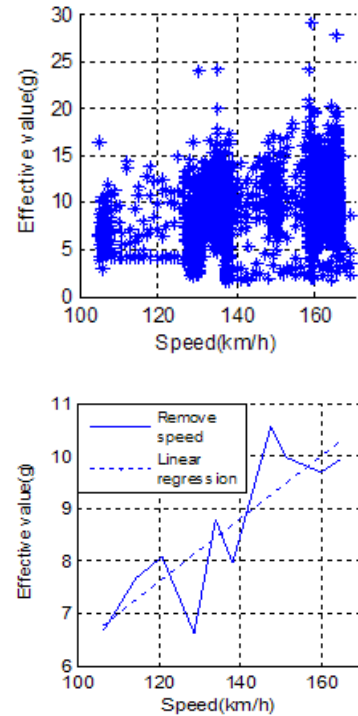
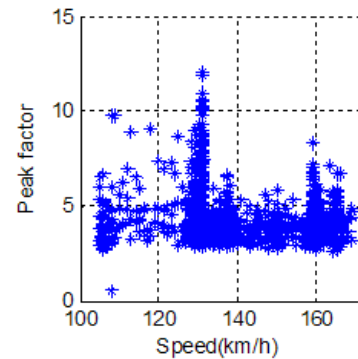


Fig. 2. (a) Effective acceleration distribution of “hard-spot” (b) Speed changing trend of effective acceleration.



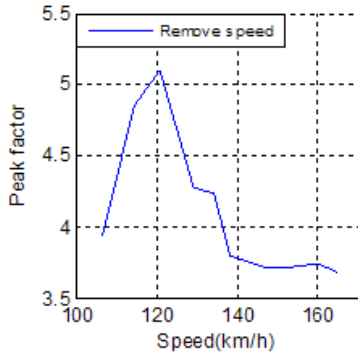


Fig. 3. (a) Peak factor distribution of "hard-spot" (b) Speed changing trend of peak factor.

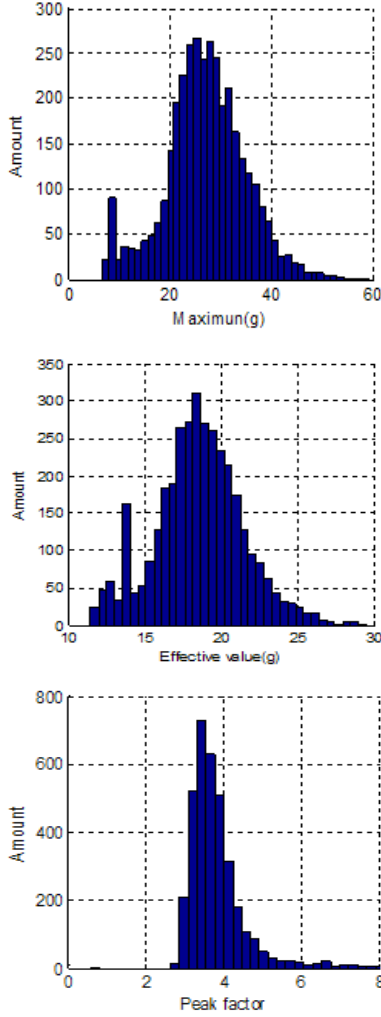


Fig. 4. (a) Amount distribution of maximum acceleration (b) Amount distribution of effective acceleration (c) Amount distribution of peak factor

From Figure 4(a-c), the parameter distributions after speed removal are basically normal distributions. We draw about 10% data in the sampling data after speed removal, and then a matrix W with 370 data is formed. The data in matrix W can be divided into two classes with FCM, and the "hard-spot" threshold can be confirmed based on clustering centers.

Let $m=1.5, c=2, \varepsilon=0.001$, and suppose v_1, v_2 be initial different clustering centers. The orders of each point in matrix are corresponding to the horizontal ordinates. We define the distance between the two clustering centers.

$$d = |v_1 - v_2| \quad (5)$$

The distances from first point $w(1)$ in matrix w to clustering centers v_1, v_2 are defined.

$$d_{11} = |v_1 - w(1)| \quad d_{12} = |v_2 - w(1)| \quad (6)$$

The memberships of $w(1)$ relative to v_1, v_2 are computed respectively.

$$u_{11} = d_{11} / (d_{11} + d_{12}) \quad u_{12} = d_{12} / (d_{11} + d_{12}) \quad (7)$$

And so on, the membership matrix $U = [u(k, i)]$ with 2×370 can be obtained. The objective function is as follow.

$$J = \sum_{k=1}^{370} \sum_{i=1}^2 u(k, i)^{3/2} d(k, i)^2 \quad (8)$$

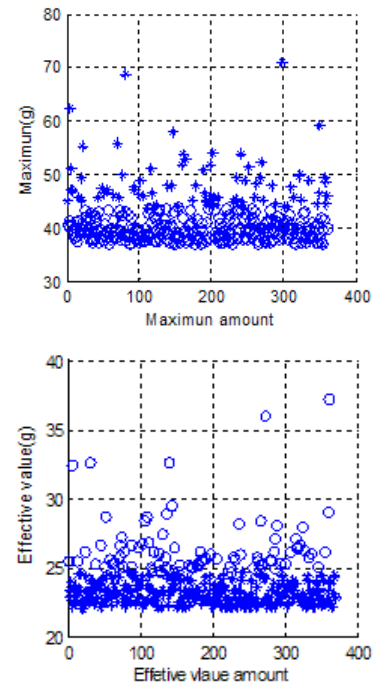
Modifying the clustering centers, let

$$v'_1 = \frac{\sum_{k=1}^{370} u(k, 1)^{3/2} w(k)}{\sum_{k=1}^n u(k, 1)^{3/2}} \quad (9)$$

$$v'_2 = \frac{\sum_{k=1}^{370} u(k, 2)^{3/2} w(k)}{\sum_{k=1}^n u(k, 2)^{3/2}} \quad (10)$$

v'_1, v'_2 are new clustering centers. The distances from all points in the matrix to v'_1, v'_2 are recomputed, as well as the new membership matrix $U = [u(k, i)]$. After n th modifications, the objective function J_n can be obtained. The iteration will stop until $J_n - J_{n-1} < \varepsilon$.

The clustering distributions of parameters are shown in Figure 5(a-c).



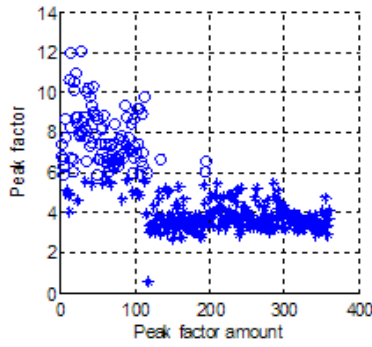


Fig. 5. (a) Clustering distribution of maximum acceleration (b) Clustering distribution of effective acceleration (c) Clustering distribution of peak factor

The clustering centers of “hard-spot” parameters are listed in Table 1.

TABLE I: THE CLUSTERING CENTERS OF “HARD-SPOT” PARAMETERS

Index	First clustering center	Second clustering center	Average
Maximum Acceleration (g)	48.3072	39.3654	43.8635
Effective Acceleration (g)	26.6887	23.0383	24.8635
Peak Factor	7.8051	3.7830	5.79405

The average value of two clustering centers means the threshold of in the paper. For example, in a parameter set (speed 135.070 km/h, maximum acceleration 49.882 g, effective acceleration 49.882 g and peak factor 4.850), the maximum is bigger than threshold, but the effective acceleration and peak factor are smaller than threshold. We cannot judge that the “hard-spot” of the sampling data is transfinite. Hence, whether a “hard-spot” is transfinite needs depend on the integrated judgment of three thresholds of feature parameters.

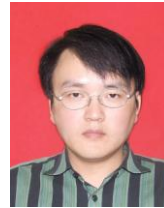
V. CONCLUSION

At present, in pantograph-catenary relationship detection system of China, whether a “hard-spot” is transfinite only depends on the maximum acceleration in each catenary space, which results into the failure of “hard-spot” judgment. In the paper, an idea on extracting threshold of pantograph-catenary current collection parameters of train-set is proposed. Firstly, three current collection feature parameters are pre-processed. Then, with the FCM method, the clustering centers of maximum acceleration, effective acceleration and peak factor

are computed, and the corresponding thresholds of pantograph-catenary current collection parameters can be obtained. In the end, the judgment criterion of “hard-spot” based on the three feature thresholds can be obtained.

REFERENCES

- [1] H. Li, “Analysis on catenary hard-spot,” *Electric Railway*, vol. 3, pp. 31-32, June, 2002.
- [2] B. Q. Zhang and M. H. Jia, “On the catenary hard-spot and its preventive measure,” *Railway Locomotive and Car*, vol. 5, pp. 39-41, October, 2005.
- [3] Z. Q. Zhang, “The reason of catenary hard-spot and its thinking and practice of regulation,” *Electric Railway*, vol. 4, pp. 36-38, August, 2007.
- [4] N. Fumio, I. Matsutaro, and N. Hiroki, “Measurement of contact loss by detecting spark,” *Quarterly Reports*, vol. 25, no. 3, pp. 95-98, March, 1984.
- [5] T. X. Han, “Reason analysis for continuous spark in current collection of catenary-pantograph system,” *Railway Locomotive and Car*, vol. 23, no. 3, pp. 58-61, March, 2003.
- [6] G. Chen, “The current collecting performance analysis of motor train-set’s pantograph and catenary based on fuzzy clustering algorithm,” *Master Degree Thesis, Dept. Elect. Eng.*, Southwest Jiaotong University, Chengdu, Sichuan, 2007.
- [7] J. C. Bezdeck, R. Ehrlich, and W. Full, “FCM: Fuzzy C-means algorithm,” *Computers and Geosciences*, vol. 23, pp. 16-20, 1984.
- [8] G. Tao and B. Dubuisson, “A loose-pattern process approach to clustering fuzzy data sets,” *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 7, no. 3, pp. 366-372, March, 1985.
- [9] L. D. Mantilras and L. Valverde, “New results in fuzzy clustering based on the concept of in distinguish ability relation,” *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 10, no. 5, pp. 754-757, May, 1988.



Zhigang Liu received the PhD degree in Power System and its Automation from Southwest Jiaotong University of China in 2003. He is currently a full Professor of School of Electrical Engineering, Southwest Jiaotong University in China. His research interests are signal processing and its application in power system, electrification railway technology, and computational intelligence.

Zhiwei Han received the Master degree in Power System and its Automation from Southwest Jiaotong University of China in 2009. He is pursuing PhD. His research interests are signal processing and its application in electrification railway.

Gang Chen received the Master degree in Power System and its Automation from Southwest Jiaotong University of China in 2012. Her research interests are signal processing and its application in power system.

Jian Xie received the Master degree in Power System and its Automation from Southwest Jiaotong University of China in 2010. His research interests are signal processing and its application in electrification railway.