The Optimization Select Method of Spacecraft Initial Orbit Based on K-means Clustering Algorithm

Dingxin Yang, Cong Gao, Peng Du

China Satellite Maritime Measurement and Control Department

Jiangyin, China

laoba6065@126.com

Abstract. In the rocket active segment mission, the spacecraft and rocket separation orbit after the initial orbit number is an important basis to determine the success of rocket launchers. At present, there are many data sources that can be used to determine the number of initial orbits. However, it is preferable to rely mainly on manual decisions, resulting in long time-consuming and easily disturbed by the site environment, and the decision-making efficiency and accuracy are not high. A method based on K-means clustering algorithm for spacecraft initial trajectory optimization is proposed, which is automatically classified by machine learning and automatically optimized according to a predetermined strategy to improve decision efficiency and accuracy.

1 Introduction

In the rocket active segment mission, the orbital spacecraft orbital spacecraft orbital separation is the important basis for determining the initial success of a rocket launch [1]. With the continuous enrichment of measurement methods, there are many data sources that can be used to determine the number of initial orbits. However, how to optimize depends mainly on manual decisions, takes longer time, is susceptible to on-site environmental disturbances, and has poor decision-making efficiency and accuracy.

K-means algorithm is a classic data classification algorithm proposed by JB MacQueen in 1967 to deal with the problem of data clustering. The algorithm has the advantages of simple algorithm and high computational efficiency, and put forward earlier, in the field, the algorithm has been studied deeply, but also a more rich and perfect improvement, so in the field of data analysis and application of a wide range of influence [2].

Based on K-means clustering algorithm, a method of spacecraft initial orbit optimization is designed, which can be automatically classified by machine learning and automatically optimized according to predetermined strategies. The theoretical research and experimental data validation show that this method can greatly improve the efficiency and accuracy of decision making.

2 K-means Clustering Algorithm

The K-means clustering algorithm solves the problem of dividing the sample set x={x1, x2, ..., xn} containing n data into k categories C_j, j=1,2, ..., k , The algorithm randomly chooses k data points as the initial centers of k classes. Each data point in the set is divided into the class of the nearest centroid, forming the initial distribution of k clusters. For each class assigned to calculate a new class center, and then continue the data allocation process, so after a predetermined number of iterations, if the class center is no longer a new change, then all the data has been allocated to their own class, The clustering criterion function converges, otherwise continue to iterate until convergence. General clustering criteria function will use the clustering error squared sum criteria function.

2.1 Data Standardization and Similarity Metrics

To avoid the impact of data units of measure on clustering, the data can be normalized as needed to convert the initial measurements to unitless variables. First calculate the mean deviation S, as in (1).

$$S = \frac{1}{n}(|x_1 - m| + |x_2 - m| + ... + |x_n - m|) \tag{1}$$

Where m is the mean of the data set X, i.e., m = $(x_1 + x_2 + ... + x_n) / n$, and then calculate the normalized measured value x′, as in (2).

$$x' = \frac{x_i - m}{S} \tag{2}$$

Using the absolute deviation mean S for standardization, the effect of abnormal (discrete) data is reduced, which helps to improve the classification accuracy. After standardization, the degree of data differences can be determined by the distance between the data, the commonly used Euclidean distance to represent [3], as in (3).

$$d(i, j) = \sqrt{(|x_{i1} - x_{j1}|^2 + |x_{i2} - x_{j2}|^2 + ... + |x_{ip} - x_{jp}|^2)} \tag{3}$$

2.2 K Value Determination Algorithm

K-means clustering algorithm to determine the number of clusters is specific to the data set, the number of clusters in the search within the scope of the operation of the clustering algorithm generates a different number of clustering results, select the appropriate validity index of poly Class results to assess, based on the assessment results to determine the optimal number of clusters [4~7]. Algorithm steps are as follows:

[1] For the sample set X = $\{x_1, x_2, ..., x_n\}$，select the number of clusters [k_{min}, k_{max}], usually make $k_{min}=2$, $k_{max} = Int(\sqrt{n})$.

[2] For k = k_{min} to k_{max}

    {

        Randomly select k initial cluster centers Z^k;

        Use K-means clustering algorithm Get c clusters, and to update the membership matrix U^k and the clustering center Z^k;

        Calculate the BWP value [8] for the i-th sample of class j, as in (4);

$$BWP(j,i) = \frac{\min\limits_{1 \le k \le c, k \ne j}(\frac{1}{n_k}\sum\limits_{p=1}^{n_k}\|x_p^{(k)} - x_i^{(j)}\|^2) - \frac{1}{n_j - 1}\sum\limits_{q=1,q \ne i}^{n_j}\|x_q^{(j)} - x_i^{(j)}\|^2}{\min\limits_{1 \le k \le c, k \ne j}(\frac{1}{n_k}\sum\limits_{p=1}^{n_k}\|x_p^{(k)} - x_i^{(j)}\|^2) + \frac{1}{n_j - 1}\sum\limits_{q=1,q \ne i}^{n_j}\|x_q^{(j)} - x_i^{(j)}\|^2} \tag{4}$$

        Calculate the average BWP value, as in (5);

$$avg_{BWP}(k) = \frac{1}{n}\sum_{j=1}^{k}\sum_{i=1}^{n_j}BWP(j,i) \tag{5}$$

    }

[3] When comparing the BWP value, k corresponding to the optimal BWP value is the optimal cluster number k_o，as in (6).

$$k_o = \arg\max_{2 \le k < n}\{avg_{BWP}(k)\} \tag{6}$$

[4] Output clustering results: class center point Z_0, membership matrix U_0, the optimal number of clusters k_0.

2.3 Algorithm Basic Process

The basic process of the algorithm is as follows:

Input: data set $\{x_1, x_2, ..., x_n\}$, specify the number of clusters k;

Output: k categories C_j., j = 1,2, ..., k;

[1] K cluster centers are randomly determined: (m_l, m_2, ..., m_k).

[2] For each sample x_i, find the nearest taxonomic center and assign it to that class.

[3] Recalculate the center of each cluster, as in (7).

$$m_i = \frac{1}{N_i}\sum_{j=1}^{N_i}x_{ij}, i = 1,2,...,k \tag{7}$$

[4] Calculate the deviation value, select the following function as a clustering criterion function, as in (8).

$$J = \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left\|x_{ij} - m_i\right\|^2 \qquad (8)$$

[5] Judge whether J convergence, if convergence, the algorithm terminates, returns K categories C_j, otherwise return to step 2 to continue.

3 Design and Implementation of Initial Orbit Optimization Select Method

3.1 Design and Implementation of The Method

The number of spacecraft initial orbit is generally represented by six elements of the orbital time and the number of orbital elements (t, a, e, i, Ω, ω, m). Since the semi-long axis a can best represent the orbit characteristics, So as a key element of orbital classification. Assuming that the number of orbital types for participation is m, and the number of orbits of a certain type is p_j (where j = (1, 2, 3 ... m)), the preferred flow is as follows:

[1] Determine the number of various types of orbits, each class is denoted by the number of [x_1, x_2, ..., x_m], whichever is the semi-long axis, and data normalization.

[2] According to the measurement accuracy of each measuring device and combined with historical experience to determine the weight of the m-type orbits calculated, defined as [w_1, w_2, ..., w_m], then the weight of a certain orbit in each kind of orbit is w_i / x_i (where i = 1, 2, ..., m).

[3] Using statistical methods to dynamically adjust the weights [9~11]. Calculate the standard deviation of the semi-major axes of all kinds of orbits, denoted as [σ_1, σ_2, ..., σ_m], weights obtained from the standard deviation, denoted as [ω_1, ω_2, ..., ω_m], as in (9).

$$\omega_i = \frac{\frac{1}{\sigma_i}}{\sum_{j=1}^{m}\frac{1}{\sigma_i}}(\sigma_i \neq 0) \qquad (9)$$

[4] Use the mean-square error of the semi-major axis of all kinds of orbits to correct the weights, denoted as [w_1, w_2, ..., w_m]; as in (10), then the weight of a certain orbit in each type of track is w_i / x_i.

$$w_i' = \frac{w_i + \omega_i}{2}(w_i \neq 0) \qquad (10)$$

[5] The revised weight is used to calculate the sum of the weights of each category. The highest category is the final preferred category, and the group of tracks closest to the category center is taken as the output, as shown in Fig. 1.
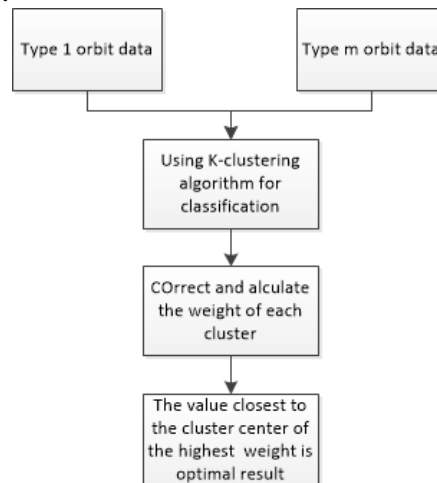

Figure 1. Flow chart of initial rail optimization method

3.2 Method validation

Using a certain spacecraft orbit into the orbit measurement of the first measurement of the number of root to verify, a total of five categories, according to the order of theoretical accuracy were marked as ABCDE, the weight according to the measurement accuracy is set to [0.3,0.2,0.2,0.2 , 0.1]. Calculated according to the number of various types of orbits a single weight of [0.05,0.02,0.0105,0.0167,0.1].The semimajor axis is shown as Table I, and the mean square error of all kinds of orbits' semi-major axis is calculated. According to the mean square error, the weighted values are [0.6144, 0.1194, 0.1078, 0.1084, 0.05] ,The number of single weights calculated as [0.1024,0.012,0.0057,0.009,0.05].

Table I. Semi -long axis data of various initial orbits of a spacecraft

Name (Quantity)	Semi-long axis (km)	Mean square deviation
A(6)	*4436.313,*4436.033,*4436.439,*4435.921 *4436.169,*4436.001	11.942
B(10)	*4444.332,*4442.480,*4440.683,*4438.425 *4436.203,*4434.217,*4432.604,*4431.728 *4432.039,*4431.639	4.773
C(19)	*4432.774,*4433.945,*4430.761,*4432.505 *4431.331,*4426.087,*4423.781*4423.440 *4477.969,*4427.876,*4422.327,*4423.532 *4428.721,*4426.920*4428.662,*4430.593 *4430.535,*4433.589,*4424.822	0.199
D(12)	*4477.770,*4477.593,*4477.732,*4477.719 *4477.692,*4477.849,*4456.721,*4456.665 *4456.780,*4456.728,*4456.679,*4456.818	10.964
E(1)	*4474.701	0
		* represents the same value

For Type E, since there is only a small number of measurements and only one set of orbits, the mean squared deviation is zero. Therefore, the mean squared error corresponds to a weight of zero and the weight adjustment uses formula 10. In terms of the mean square deviation of other types of orbits, under normal circumstances, the data with high theoretical measurement accuracy tend to have smaller mean square error. From the dynamically adjusted result of the new weight, it can be seen that the data with high measurement accuracy can obtain a larger total weight.

Table II. Clustering and comprehensive weight calculation result data

Cluster center (km) (genus / quantity)	Clustering data (km) genus	Comprehensive weight
*4456.732 (D/6)	*4456.721 D,*4456.665 D,*4456.780 D *4456.728 D,*4456.679 D,*4456.818 D	0.055
*4437.196 (A/6,B/6,C/2)	*4433.945 C,*4433.589 C,*4444.332 B *4442.480 B,*4440.683 B,*4438.425 B *4436.203 B,*4434.217 B *4436.313 A *4436.033 A,*4436.439 A,*4435.921 A *4436.169 A,*4436.001 A	0.697
*4428.634 (B/4,C/16)	*4432.774 C,*4430.761 C,*4432.505 C *4431.331 C,*4426.087 C,*4423.781 C *4423.440 C,*4427.876 C *4422.327 C *4423.532 C,*4428.721 C,*4426.920 C *4428.662 C,*4430.593 C,*4430.535 C *4424.822 C *4432.604 B,*4431.728 B *4432.039 B,*4431.639 B	0.139
*4477.378 (D/6,C/1,E/1)	*4477.770 D,*4477.593 D,*4477.732 D *4477.719 D,*4477.692 D,*4477.849 D *4474.701 E,*4477.969 C	0.109
		* represents the same value

The optimal clustering number is 4, and each clustering and clustering center is calculated by K-means clustering algorithm. Shown as Table II, according to the weights of all kinds of orbits, the sum of each cluster can be calculated Weights, based on the center of the cluster with the highest combined weight, are selected as the final set of the best output for the final result. Shown as Table III, it can be seen that the accurate result of the preferred result and the post-processing Very close to completely meet the accuracy of the choice of the initial track requirements, verify the feasibility and accuracy of the method.

Table III. Comparison of optimal results with post processing accurate results

Final selection results (genus)	Accurate results of post processing
*4436.313(A)	*4436.471
* represents the same value	

4 Conclusions

This paper presents a K-means clustering algorithm based on multi-data source spacecraft initial orbit optimization method, introduced the algorithm flow and several key steps, and the use of programs to achieve the algorithm, the use of a spacecraft measured initial orbit The method verifies the feasibility of the method and the correctness of the method. The method can greatly improve the efficiency of the initial-track selection under the premise of ensuring the accuracy.

References

[1] JS Li, Spacecraft orbit determination, Beijing: National defense industry, 2003, pp.84-103.

[2] JG Sun , J Liu , LY Zhao , Clustering algorithms research, Journal of Software, vol. 19, issue 1, pp. 48-61, 2008.

[3] S Jiang, Research on K- means clustering algorithm, Xian: Shaanxi Normal University, 2010.

[4] SB Zhou, Research and application of optimal cluster number determination method in clustering analysis, Wuxi: Jiangnan University, 2009.

[5] SB Zhou, ZY Xu, XQ Tang, New method for determining optimal number of clusters in k-means clustering algorithm, Computer Engineering and Applications, vol. 46, issue 16, pp. 27-31, 2010.

[6] Y Wang, J Tang, QF Rao, CY Yuan, High efficient K-means algorithm for determining optimal number of clusters, Journal of Computer Applications, vol. 34, issue 5, pp. 1331-1335, 2014.

[7] RY Jia, JL Song, K-means Optimal Clustering Number Determination Method Based on Clustering Center Optimization, Microelectronics & Computer, vol. 33, issue 5, pp. 62-66, 2016.

[8] SB Zhou, ZY Xu, XQ Tang, Method for determining optimal number of clusters in K-means clustering algorithm, Journal of Computer Applications, vol. 30, issue 8, pp. 1995-1998, 2010.

[9] YL Hao, ZP Liu, Fusion algorithm of angular velocity weighted averages for GFSINS based on dynamically allocating weights, Journal of Chinese Inertial Technology, vol. 18, issue 1, pp. 16-21, 2010.

[10] J Jiang, C Gao, LN Yu, Application of the Weighted Average Method Based on Dynamically Allocating Weights in Attitude Solution, Aeronautical Manufacturing Technology, vol. 506, issue 11, pp. 106-109, 2016.

[11] QX Wang, FX Wang, Application of the Weighted Average Method Based on Dynamically Allocating Weights in Attitude Solution, Bulletin of Soil and Water Conservation, vol. 36, issue 2, pp. 151-155, 2016.