IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received August 24, 2019, accepted September 4, 2019, date of publication September 10, 2019,
date of current version September 23, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2940248

Mining Two-Line Element Data to Detect Orbital Maneuver for Satellite

XUE BAI[1], CHUAN LIAO[2], XIAO PAN[1], AND MING XU[1,3]
[1]School of Astronautics, Beihang University, Beijing 100191, China
[2]Research Institute, China Electronics Technology Group Corporation, Chengdu 610036, China
[3]DFH Satellite Company Ltd., Beijing 100094, China

Corresponding author: Ming Xu (xuming@buaa.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 11772024 and Grant 11432001, and in part by the Shanghai Space Science and Technology Innovation Foundation under Grant SAST2017-033.

ABSTRACT Data clustering analysis is proposed to detect the orbital maneuvers of satellites at different scales. In this study, the unsupervised classification methods of K-means, hierarchical, and fuzzy C-means clustering are used to handle the two-line element (TLE) historical data. The K-means-based contour map method is applied to the characteristic variable selection and cluster number determination. The TLE data of large-, medium-, and small-scale orbital maneuvers are clustered by the aforementioned three methods. Through a series of numerical experiments, it is found that for different scales of orbital maneuvers, the clustering methods have different performances and that they can essentially fulfill the functional requirements of orbital detection. By data mining, the orbital maneuvers of the remote sensing satellites "YAOGAN-9", "TIANHUI-1", and "Envisat" can be easily detected, which will provide useful information for further orbital supervision and prediction.

INDEX TERMS Clustering, data mining, orbital maneuver detection, space situational awareness TLE data.

NOMENCLATURE

a	semi-major axis of ellipse orbit
e	orbit eccentricity
i	orbit inclination
Ω	right ascension of ascending node
ω	argument of perigee
M	mean anomaly
$\boldsymbol{r}$	$[r_x, r_y, r_z]$, position vector from the spacecraft to the Earth's center of mass
$\boldsymbol{v}$	$[v_x, v_y, v_z]$, velocity vector of the spacecraft
μ	gravitational parameter of the Earth

I. INTRODUCTION

Remote sensing satellites rely primarily on either reflected or emitted electromagnetic radiation from the Earth to obtain information on the Earth's surface or the overlying atmosphere [1]. There is a certain orbital maneuver that keeps remote sensing satellites at a prearranged or target orbit so that they can complete their tasks. Detecting this orbital maneuver through effective methods such as data mining has attracted the attention of many scholars.

Recently, data mining has attracted considerable attention across several fields, where it has proved beneficial. There are many people working with the idea that patterns in data can be sought automatically, identified, validated, and used for prediction, which spawned data mining. Data mining involves methods and algorithms that solve problems by analyzing existing data [2]. With the rapid development of computer techniques and the unremitting efforts of practitioners in various industries, data mining technology can automatically accomplish a task and has been greatly developed and applied to many practical uses.

In aerospace engineering, there have been many studies on the use of data mining techniques for exploiting satellite data to solve dynamic problems [3]–[11].

Compared with the physics-based approach, data mining presents a different modeling and prediction capability without an exact model of objects, maneuvers, and space environments. Self points out that data mining is an efficient method for analyzing large volumes of data, such as in the case of clusters of satellites [3]. Sánchez et al. applied the deep neural network to the solution of the

The associate editor coordinating the review of this manuscript and approving it for publication was Zhan Bu.

Hamilton–Jacobi–Belman equation and provided a near-optimal control method for spacecrafts [4]. Du et al. developed an improved ant colony algorithm based on a sensational and consciousness strategy to solve the traveling salesman problem with constraints in the area targets observation mission [5]. Polivka et al. used data mining to process a large amount of historical GPS navigation data to compare the broadcast orbits and clocks to reference products, which provides insight to long-term performance and establishes a baseline for assumptions on the clock and orbit errors and failure rates [6]. Tanner et al. proposed the concept of on-board data mining [7]. Qu et al. applied the Internet of Things, which can be combined with big data, to the low-Earth orbit satellite [8]. Based on the normalization of the large-scale network cooperation of sensor systems, sensor data is processed on the satellite, which can improve network communication capability and reduce costs. To make full use of remote sensing data, Gong et al. designed a multisource remote sensing image data mining system framework based on traditional data mining and knowledge discovery technology [9]. Meanwhile, Peng et al. concentrated on the lack of area-to-mass ratio of a resident space object in most space catalogs, and used Random Forest, a data mining method, to determine the connection between the consistency error and area-to-mass ratio [10]. Peng et al. also used the support vector machine model to learn the errors of orbit prediction, aiming to improve the accuracy of satellite orbit prediction [11]. However, they did not simulate TLE or any other actual on-orbit data.

Orbital maneuver detection is an important component of space situational awareness [12], especially for very little maneuvers that are hard to detect by traditional methods. Through maneuver detection, we can analyze the orbital control purpose of the satellite to obtain the mission objectives. Danial et al. developed an optimal state estimator called the optimal control-based estimator to detect and reconstruct maneuvers with no prior information [13]. The variable dimension filter presented by Bar-Shalom et al. changed the state model for the target by introducing extra state components to realize the track of the maneuvering target instead of relying on the statistical description of the maneuver as a random process [14]. Spingarn and Weidemann proposed a linear regression filter for tracking maneuvering targets [15]. Thorp described the maneuver model as an increase in the driving noise and computed the likelihood for the hypothesis as to whether or not a maneuver occurred [16]. Patera's space detection method was suitable for detecting space events, such as maneuvers, collisions, and explosions, but it focused more on quick events and neglected smaller maneuvers and natural dynamics mismodeling, making its application limited [17]. However, there is currently no systematic method to detect various spatial events based on orbital anomalies. Thus, as anomaly detection and pattern mining are a branch of data mining, it is quite practical to apply data mining technology to orbital maneuver detection.

The data volume of TLE historical data is quite large, and there is very little valuable information. Therefore, a deep analysis is performed on a large amount of data to obtain the information that is beneficial to space mission design and improve model accuracy. So far, some studies on mining TLE data have been conducted. Doornbos et al. converted a large amount of TLE data into satellite resistance data for the problem of the large error in the traditional experience thermal layer model [18]. Lemmens et al. proposed two orbital maneuver detection methods based on TLE historical data. The first method focuses on low-Earth-orbit non-natural anomalous events based on a consistency check between arbitrary element sets. The second method evaluates the unexpected changes by robust statistical and harmonic analyses using time series or derived quantity thereof. Both methods achieve better results in different orbital systems [19]. Kelecy et al. have proposed a maneuver detection algorithm that uses historical TLE data. They examined the data by searching the abnormal differences that exceed the user-specified threshold to realize the detection Their filter requires computing a first-order polynomial fit of all the data within the interval [20].

This paper focuses on the detection of orbital maneuvers through data mining with historical TLE data. The unsupervised classification methods, K-means, hierarchical, and fuzzy C-means clustering, are used to detect the orbital maneuver and maintenance strategy to record the orbital control action. The framework of the data mining approach based on historical TLE data is demonstrated in Figure 1. The TLE historical data is first divided into large-, medium-, and small-scale data according to the magnitudes of orbital maneuvers. Then, data preprocessing is needed to obtain the conforming data to cluster, such as 0-1 standardization, wavelet filtering and Simplified General Perturbations Satellite Orbit Model 4 (SGP4) orbit prediction [21]. After completing the above two steps, the data clustering process is performed First, the K-means-based contour map method is applied to the characteristic variable selection and cluster number determination. Then, according to the TLE data, the three clustering methods are implemented to cluster the data and determine the most suitable method. For the large-, medium-, and small-scale orbital maneuvers, the most efficient clustering methods are different as a result of varying the characteristic variable selection and cluster number. In this paper, the TLE data of the remote sensing satellites, ''YAOGAN-9'' ''TIANHUI-1'' and ''Envisat'' are demonstrated to prove that the clustering method is effective based on traditional engineering experience. Although the three methods have certain missed judgments and return errors in some cases, the convenience of the data mining method is superior to that of traditional methods.

This paper makes three main contributions: **1)** A large number of TLE historical data are clustered in a well-organized manner to present orbital maneuver behaviors without using complicated estimators or filters usually used in classical methods. There is no sensitivity trade-off between

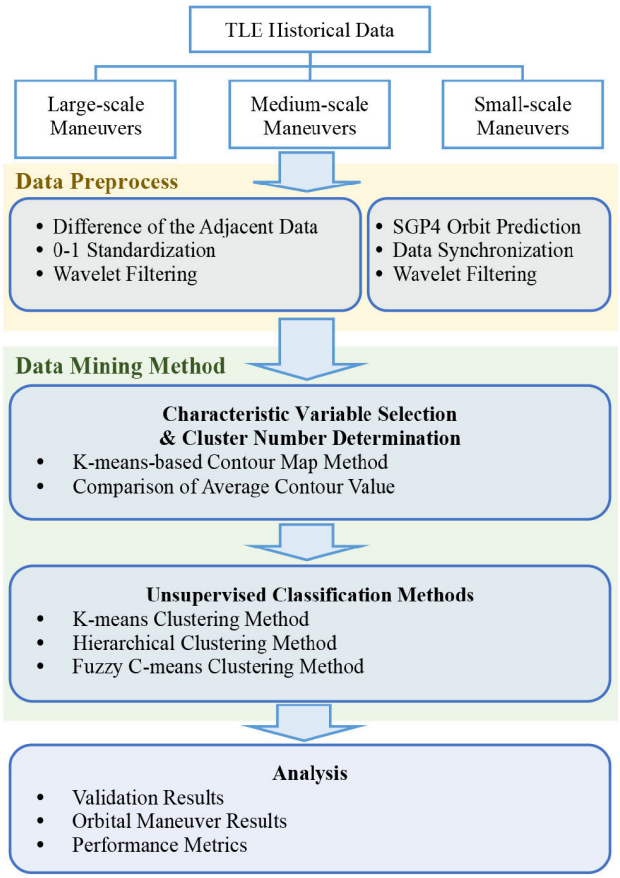


FIGURE 1. **Flowchart of the detection of orbital maneuver by data mining: the three clustering methods are implemented to cluster the data.**

the algorithm parameters and the clustering process is conducted without the artificial threshold derived from the traditional experiences; **2)** The unsupervised classification methods are demonstrated to have good performance with regard to detecting orbital maneuvers. The missing rates and the false rates can be as low as 1%, in comparison to 6%, which occurs for some classical methods; **3)** For different scales of orbital maneuvers, the performances of different clustering methods have been investigated, addressing the shortcoming of only one scale orbital maneuver being examined in classical methods.

II. FUNDAMENTAL CONCEPTS OF DATA MINING AND ORBITAL MANEUVERS

In this section, the fundamental concepts of the three main clustering methods of data mining are reviewed as well as the orbital maneuvers based on the orbit elements.

A. CLUSTERING METHODS OF DATA MINING

In the face of a huge amount of data, we are interested in the methods for finding and describing structural patterns in the data as an implement to help interpret the data and make predictions from it. Clustering techniques, as one of the main contents of data mining, are usually applied when there is no clear class to be predicted yet instances need to be divided into particular clusters. The data handled in this method are usually vector objects represented as attribute vectors [22] These clusters may reflect mechanisms that act in the domain from which the instances are extracted, which makes certain instances more similar to each other than to the rest [2]. After the clustering is completed, the different clusters are analyzed in detail. Furthermore, to obtain more convincing results, other methods, such as self-learning algorithms and tunable weighting strategies, are used to improve the clustering methods [23], [24].

Based on the principle of the clustering algorithm, the clustering methods can be divided into six groups: partitioning methods, hierarchical methods, density-based methods, model-based methods, grid-based methods and soft-computing methods [25]. Most of traditional clustering methods focus on the density of edges as the attributes and some new methods come up according to the signs of edges to cluster. Li et al. proposed a new efficient clustering algorithm based on the positive and negative update rule to ensure that the network clustering reaches a state of optimal convergence [26].

Presently, the K-means, hierarchical, and fuzzy C-means clustering methods are usually used in practice. The K-means method is a typical hard partitioning clustering method that is simple to implement and widely used. It can be viewed as a degenerate density-based algorithm [27] or a special case of model-based clustering, where all the distributions are assumed to be Gaussians with equal variance [28]. The fuzzy C-means clustering method extends the partitioning notion and suggests a soft clustering schema. The hierarchical method, different from the methods above, construct the clusters by recursively partitioning the instances in either a top-down or bottom-up fashion. The main advantage of the hierarchical method is its applicability to any attribute type. Moreover, the TLE data are typical multi-density data and the effect of grid-based methods used in multi-density data clustering is ''poor'' [29]. The grid-based methods are spatially driven, dividing the embedding space into units independent of the distribution of input objects, different from the other four methods driven by data. To make our analysis targeted and comparable, grid-based methods are not considered in this paper [30].

1) K-MEANS CLUSTERING METHOD

K-means is the most typical distance-based clustering algorithm. It uses distance as a criterion for the similarity between data. Given a predetermined number K, the algorithm divides the data set into K disjoint subsets. Then, the grouping procedure is repeated until it converges. After the number of clusters k and the pieces of data n are given in advance, the processing flow of the K-means algorithm is as follows: **1)** select k data randomly as the initial cluster center from the n data and; **2)** calculate the distance between the remaining data and the cluster center, and re-divide the data according to the minimum distance, which minimizes the following

objective function:

$$E = \sum_{j=1}^{k} \sum_{x_i \in \omega_j} \|x_i - m_j\|^2 \quad (1)$$

where x_i is a data object in a cluster ω_j and m_j is the centroid (mean of objects) of ω_j; **3)** recalculate the mean of each cluster as a new clustering center; **4)** repeat steps 2) and 3) until each cluster no longer changes.

The K-means algorithm is simple and fast [31]. The time complexity of the K-means method is $O(n \cdot k \cdot l)$, where n is the size of the data, k is the number of clusters, and l is the number of iterations taken by the clustering to converge [28]. It indicates that the K-means method has linear time complexity in the size of the data set. The space complexity of the K-means method is $O(n + k)$.

However, as it is a two-stage loop algorithm, it is necessary to continuously re-divide the data, so the computational burden is heavier, and it is sensitive to some abnormal points and a given number of clusters k.

2) HIERARCHICAL CLUSTERING METHOD

The hierarchical clustering method is accomplished by dividing the data into clusters and forming a corresponding tree. The single linkage, average linkage, and complete linkage methods use the minimum, average, and maximum distances between the members of two clusters, respectively [32]. According to the formation process, it can be further divided into agglomerative approaches and divisive approaches. Agglomerative clustering is a bottom-up strategy that first regards each data as a separate cluster before merging different data points into larger clusters until all data is in one cluster or some termination condition is met. Divisive clustering is a top-down strategy that starts with one cluster containing all the data; then, it divides it into different clusters until all the data are merged into a separate cluster or until the termination condition holds.

In this paper, the agglomerative clustering algorithm is used. After the number of clusters k and the pieces of data n are given in advance, the processing flow is as follows: **1)** treat n pieces of data as n clusters; **2)** find the data points that are subordinate to different clusters but closest to each other; **3)** combine the two clusters; and **4)** repeat steps 2) and 3) until the number of current clusters reaches k.

For the hierarchical clustering method, the time and space complexities are $O(n^2 \cdot \log n)$ and $O(n^2)$, respectively because a similarity matrix of size n^2 has to be stored. The agglomeration or divisiveness of hierarchical clustering is difficult to select at a certain stage, which leads to poor quality clustering results, and owing to the need to test and estimate a large amount of data or clusters in the selection process, the flexibility of the method is also modest.

3) FUZZY C-MEANS CLUSTERING METHOD

The fuzzy C-means algorithm extends the hard C-means algorithm to allow data points to partially belong to multiple clusters simultaneously [33]. It can be described as an optimization problem and the goal is to make the objective function minimum:

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

where n is the number of data, c is the number of clusters, and $U = (u_{ik})_{c \times n}$ is the membership matrix. Here, $0 \le u_{ik} \le 1$, $\sum_{i=1}^{c} u_{ik} = 1$, and the total membership of a dataset is 1. $V = (v_1, v_2, \ldots v_c)$ is the center of the cluster. $d_{ik} = \|x_k - v_i\|$ represents the Euclidean distance from the k^{th} data to the i^{th} cluster center. u_{ik}^m is the weighting coefficient, indicating that the k^{th} data belongs to the m^{th} power of the membership of the i^{th} cluster, and m is the parameter that controls the flexibility of the algorithm.

The fuzzy C-means clustering method has a low computation burden and a high efficiency. It has an intermediate $O(n \cdot k)$ complexity [34] It can form a fuzzy similarity matrix according to the relevant data, and directly processes the similarity matrix, avoiding repeated call scans to the database. The parameter m can be dynamically adjusted as needed and has good flexibility. However, the gradient method is used in the optimization process resulting in the local optimum.

B. IMPLEMENTATION OF ORBITAL MANEUVERS

Orbital maneuvers are implemented in the initial orbital phase and orbital maintenance phase to keep the satellite at the prearranged orbit because a launch vehicle usually cannot put the satellite in its final orbit, and there are orbital perturbation factors during its in-orbit phase. Any orbital change is completed by a velocity change applied to the satellite [35]. Therefore, it is comparatively easy to analyze the change in-orbit parameters to obtain the orbital maneuvers and if the observed orbit elements change at different times, the maneuver detection should be considered.

According to the perturbed orbital dynamics [36], an impulse velocity $\Delta \boldsymbol{v}$ exerted at an arbitrary point f during the near-circular orbit will simultaneously change the values of orbit elements as follows:

$$\begin{cases} \Delta a = \dfrac{2}{n} \Delta v_r \\ \Delta e = \dfrac{1}{na} (\sin f \, \Delta v_r + 2 \cos f \, \Delta v_u) \\ \Delta i = \dfrac{1}{na} \cos(\omega + f) \Delta v_h \\ \Delta \Omega = \dfrac{1}{na} \dfrac{\sin(\omega + f)}{\sin i} \Delta v_h \\ \Delta \omega = \dfrac{1}{na} (-\cos f \, \Delta v_r + 2 \sin f \, \Delta v_u - \Delta \Omega \cos i) \\ \Delta M = n + \dfrac{1}{na} (\cos f \, \Delta v_r - 2 \sin f \, \Delta v_u) \end{cases} \quad (3)$$

where Δ is the change in elements caused by the orbital maneuver, $n = \sqrt{\mu / a^3}$, Δv_r, and Δv_u are the impulse velocity components along the radius and along-track directions, respectively, and Δv_h travels along the across-track direction.

Once the differential values of the orbit elements are known, the maneuver parameter Δv can be calculated, and the orbital detection can be realized.

III. DETECTION OF THE ORBITAL MANEUVER AT DIFFERENT SCALES

This section describes how the three methods mentioned above are used to detect the orbital maneuver and maintenance strategy, identify the abnormal orbit behavior, and record the orbital action. With regard to the difficulty in determining the threshold criterion of a traditional method in advance, with the clustering methods, the detection of the orbital maneuver begins from the TLE historical data, which avoids artificial threshold setting and achieves automatic measuring of the distance by algorithm. The number of characteristic variables and clusters is confirmed first by the counter method based on the K-means algorithm; then the TLE historical data can be unsupervised and classified into expected clusters that represent the different orbital behaviors after analyzing the characteristic variables.

According to aerospace dynamics, orbital maneuvers can be divided into three scales based on these events: the establishment and reconfiguration of formation, the maintenance of the low earth orbit, and the maintenance of the high earth orbit. Two of the ''YAOGAN-9'' constellation satellites launched from the Jiuquan Satellite Launch Centre, China, on 5 March 2009, i.e., A and B, are used as the data source for the detection of large-scale orbital maneuvers owing to there being more maneuvers and obvious orbital control in the initial stage of orbit [37]. While ''YAOGAN-9'' is in its late stage of orbit, where moderate/little orbital maintenance is required, its data will be used for the detection of small-scale orbital maneuvers. The ''TIANHUI-1'' satellite launched from the Jiuquan Satellite Launch Centre, China, on 24 August 2010 has more maneuvers and less orbital control during the long-term control period and is used as the data source for the orbit detection of medium-scale orbital maneuvers described in this section [38]. The ''Envisat'' satellite was launched on Mar. 1, 2002, by ESA and operated until Apr. 8, 2012 [39] Its TLE contain a fairly dense maneuver history of small-amplitude orbit-control maneuvers. Furthermore, the variables of data mining depend on the different scales of orbital maneuvers according to the counter method.

A. THE NUMBER OF CLUSTERS

In Section II, the TLE historical data is deconstructed into a standardized in-orbit state comprising 13 combined parameters, including orbital elements, and the position and velocity vectors. To facilitate the application of the subsequent clustering methods, clustering characteristic variable selection and cluster number determination must be solved first. Generally speaking, to judge whether orbital maneuvering occurs, manually determining the characteristic variable as a criterion and observing whether there is a significant change in the characteristic variables before and after the orbital control is necessary. This method is based on the empirical or statistical theory for the selection of the characteristic variables and the determination of the threshold, mostly requiring human intervention. This method combines the characteristic variable selection with the contour method, calculates the average contour value by using a single in-orbit state, sorts according to the calculation result, selects the combination of the criterion variables, and solves the problems of characteristic variable selection and number determination.

The contour value of the i^{th} data in the dataset is defined as:

$$S(i) = \frac{\min(\boldsymbol{b}) - a}{\max[a, \min(\boldsymbol{b})]}, \quad i = 1, \ldots, n \tag{4}$$

where *min* and *max* are the minimum and maximum functions, respectively, a is the average distance between the i^{th} data and other identical cluster data, and $\boldsymbol{b}$ is a vector where the elements represent the average distance between the i^{th} data and other different cluster data. Obviously, the value range of $S(i)$ is $[-1, 1]$. The larger the value is, the more reasonable the classification result of the data is. There are two conclusions that can be drawn from Equation (4): **1)** For the calculation of the contour value of a certain data, it is necessary to complete clustering; and **2)** the average contour value of a dataset can be used as the basis for judging the current clustering result.

For different scales of orbital control maneuvers, cluster numbers and characteristic variables correspond to different average contour values.

The TLE historical data of satellite A of ''YAOGAN-9'' are used as the basis of the simulation. The data should be processed first, and the characteristic variables before and after the orbital maneuver will change greatly. Therefore, the adjacent two data are differentiated, which means the data form $[a, e, i, \Omega, \omega, M, r_x, r_y, r_z, v_x, v_y, v_z]$ is converted into the differentiated form $[\delta a, \delta e, \delta i, \delta\Omega, \delta\omega, \delta M, \delta r_x \delta r_y \delta r_z, |\delta \boldsymbol{r}|, \delta v_x, \delta v_y, \delta v_z, |\delta \boldsymbol{v}|]$. It is worth mentioning that in the natural evolution stage of the orbit, some difference results may be quite close to 0, which is in conflict with the clustering algorithm and exceeds the computer floating point precision. Therefore, 0-1 standardization is performed again to avoid this problem.

The K-means method is then used to cluster the data and draw the contour map for the following reasons: **1)** the K-means algorithm is relatively scalable and efficient with a certain universality for the following analysis **2)** the algorithm is simple and fast empirically, reducing the time of execution required to find an appropriate number of clusters for the huge dataset; **3)** the algorithm has less computation complexity compared with other methods. **4)** it is the most typical distance-based clustering algorithm using distance as a criterion for the similarity between datasets. Consequently, it is used for characteristic variable selection and cluster number determination. Figure 2 shows the contour map when Δa was chosen as the characteristic variable and the number of clusters is 2, 3, 4, and 5. The cluster with label ''1'' has the largest proportion and it is most likely the data of the

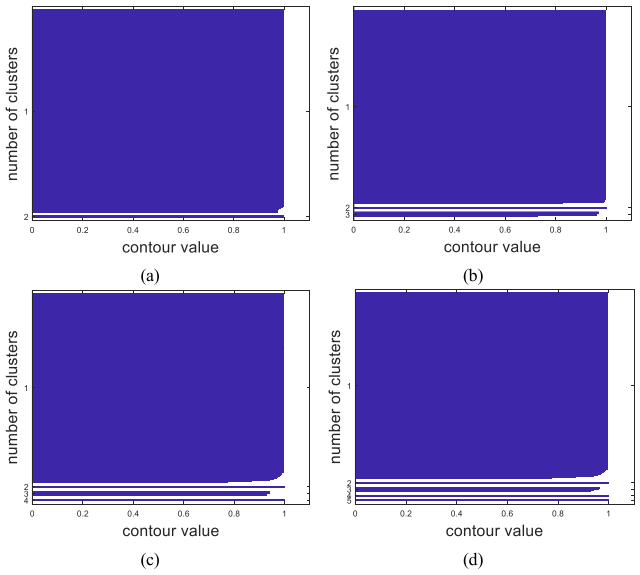


FIGURE 2. Contour map: δ a was chosen as the characteristic variable; a the number of clusters is 2; b the number of clusters is 3; c the number of clusters is 4; and d the number of clusters is 5.

natural evolution stage of the orbit. The clusters labeled "2", "3", and "4" are relatively small and should correspond to the different categories, such as ascending points, descending points, and abnormal points, which are not frequent during satellite operation and account for a small proportion. However, it is difficult to directly determine the number of clusters.

Figure 3 shows the variation in the average contour value in the case of different cluster numbers and different characteristic variables. As can be seen from the foregoing, the clustering results are effective for the average contour value closer to 1. Figure 3a demonstrates the trend of the average contour value corresponding to 14 characteristic variables as the number of clusters changes from 2 to 15. It can be seen that when the number of clusters is greater than 3, the average contour value has a certain decline and the overall trend starts to diverge. Therefore, Figure 3b focuses on the case of the average contour value corresponding to each characteristic variable when the number of clusters is 2 and 3. Except for those of $\delta\Omega$, δM, and $\delta\,|\boldsymbol{r}|$, the average contour value corresponding to most of the characteristic variables has a tendency to converge toward 1 as the number of clusters changes from 2 to 3. Therefore, setting the number of clusters to 3 is a reasonable choice.

The determination of the characteristic variables and the number of clusters are based on the TLE historical data of satellite A of "YAOGAN-9", which mainly comprises large-scale maneuvers in the deployment phase. For medium- and small-scale orbital maneuvers, the number of the clusters is basically consistent except that the selection of the characteristic variables requires a separate analysis. The specific discussion can be found in **Sections** III.C and III.D.

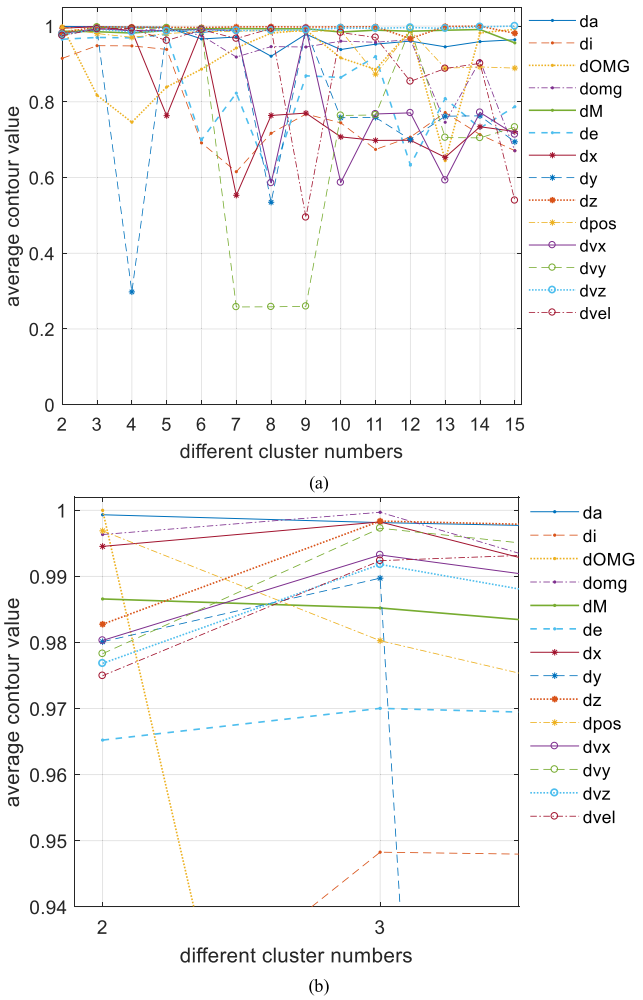


FIGURE 3. Average contour value in the case of different cluster numbers and different characteristic variables: TLE data from YAOGAN-9A; a trend of the average contour value corresponding to 14 characteristic variables as the number of clusters changes from 2 to 15; and b average contour value corresponding to each characteristic variable when the number of clusters is 2 and 3.

B. DETECTION OF LARGE-SCALE ORBITAL MANEUVER

When the satellite constellation is in the deployment phase, it will produce larger and more obvious orbital maneuvers, so the TLE historical data of satellite A of "YAOGAN-9" was chosen as the basis for the simulation on a large scale. The results of averaging the average contour values of all the characteristic variables with different numbers of clusters are shown in Table 1. It can be quantitatively seen that when the number of clusters is greater than 3, the mean average contour value presents a downward trend despite some fluctuations.

However, the average values of the numerical results of the three clusters, shown in Figure 3, are smaller than those of the two clusters, which can be seen in Table 2. This is because the average contour value corresponding to $\delta\Omega$ decreases by 18.27% when the number of clusters increases from 2 to 3, lowering the average contour value of cluster number 3. In fact, if $\delta\Omega$ is not taken into account, the average contour value of Cluster 3 is 0.9879, which is greater than that of Cluster 2, which is 0.9790. In addition,

TABLE 1. Average contour value of different numbers of clusters.

Number of clusters	2	3	4	5	6
Average contour value	0.9804	0.9757	0.9172	0.9569	0.9378
Number of clusters	7	8	9	10	11
Average contour value	0.8568	0.8044	0.8588	0.8706	0.8717
Number of clusters	12	13	14	15	
Average contour value	0.8719	0.8135	0.8607	0.8081	

TABLE 2. Average contour value of different characteristic variables when the number of clusters is 3.

Characteristic variable	δa	δi	$\delta\Omega$	$\delta\omega$	δM
Average contour value	0.9981	0.9483	0.8173	0.9996	0.9852
Characteristic variable	δe	δr_x	δr_y	δr_z	$\delta\lvert\boldsymbol{r}\rvert$
Average contour value	0.9700	0.9982	0.9897	0.9984	0.9803
Characteristic variable	δv_x	δv_y	δv_z	$\delta\lvert\boldsymbol{v}\rvert$	
Average contour value	0.9932	0.9973	0.9918	0.9924	

as can be seen from Table 2, δa is not the optimum characteristic variable when the number of clusters is 3 because the average contour values corresponding to the characteristic variables such as $\delta\omega$, δr_x, and δr_z are higher than δa, which is different from the traditional method where the semi-major axis variation is regarded as the first criterion [20].

According to the above analysis, for the detection of large-scale orbital maneuvers, the number of clusters is determined to be 3 and $[\delta a, \delta\omega, \delta r_z, \delta v_x]$ are chosen as the characteristic variables. Focusing on the first 200 TLE historical data points after Satellite A is in orbit, we obtain the semi-major axis variation curve and the K-means clustering result shown in Figure 4, in which the clustering result is plotted as a dot-dash line to highlight the contrast between them. The semi-major axis corresponding to Cluster 1 has no obvious fluctuations, indicating that Cluster 1 represents the natural evolution points of the orbit; while Cluster 2 always appears behind Cluster 3, demonstrating that Cluster 3 represents the starting point of orbital maneuver, while Cluster 2 represents the end point.

By the K-means clustering method, the orbital maneuver implemented in the deployment phase of Satellite A is illustrated in Figure 5. The clustering results show that the entire maneuvering process comprised two descending and four ascending nodes. However, from the semi-major axis curve, the first descending is not obvious, and it is in the early stage of the orbit, indicating that it is not a highly probable active maneuver of satellite A. After the second ascending, it seems there is still an ascending maneuver according to the semi-major axis curve, which is excluded from the clustering results.

Figure 6 shows the distance matrix between the 200 data points, which is the degree of similarity of each cluster. It can be seen from the color bar on the right side that the lighter the color of the dots in the figure, the larger their distance from other data points, and the lower the similarity between

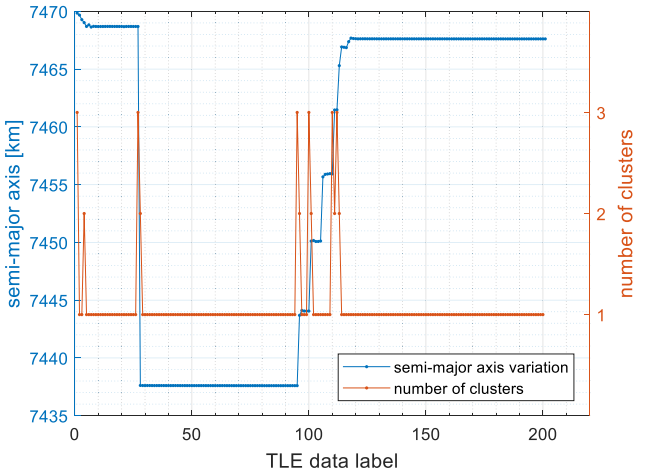


FIGURE 4. Semi-major axis variation curve and the K-means clustering result: the blue line indicates the semi-major axis variation with time and the red line indicates the clustering result.

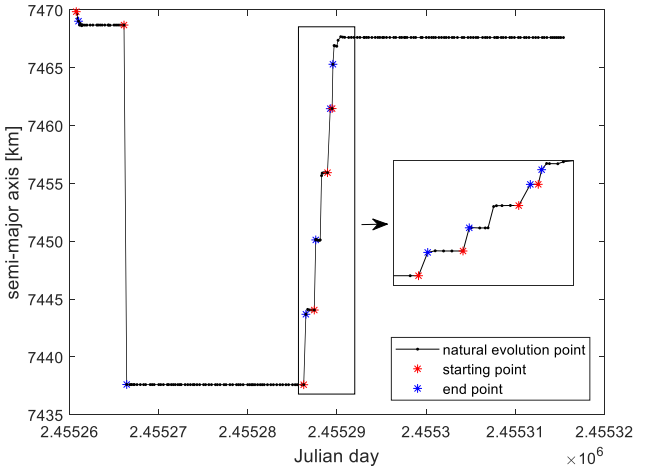


FIGURE 5. Semi-major axis curve marked with orbit control points by the K-means clustering method: the black line indicates the natural evolution process; the red asterisk indicates the starting point of the orbital maneuver; and the blue asterisk indicates the end point of the orbital maneuver.

the clusters; hence, it is reasonable to prove that the number of clusters is 3.

Figure 7 shows the semi-major axis variation curve and hierarchical clustering results of satellite A. The two semi-major axes of Cluster 2 have no obvious fluctuations, indicating that Cluster 2 represents the natural evolution points of the orbit; only one point is classified into Cluster 1, and only one point is placed into Cluster 3, which cannot show the ascending or descending maneuver clearly.

According to the detection through the hierarchical clustering method, the orbital maneuver implemented in the deployment phase of satellite A is given in Figure 8. The results show that the entire process had one descending and one ascending maneuver only, which is quite different from the actual maneuvers indicating that the hierarchical clustering missed several maneuvers.

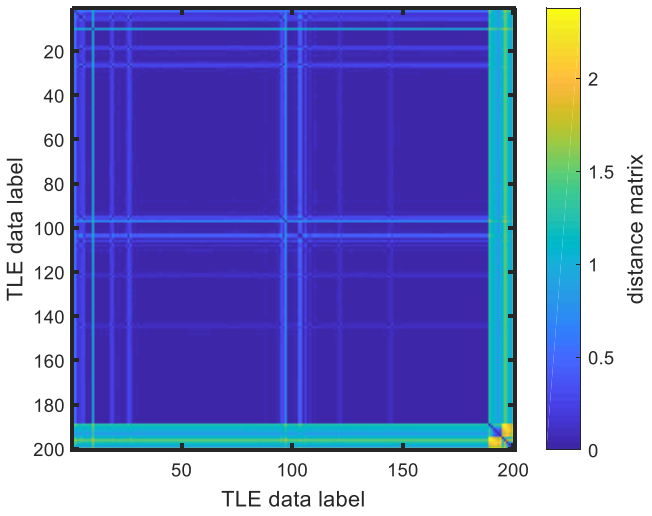


FIGURE 6. Distance matrix of the K-means clustering result: the color represents the degree of similarity of each cluster.

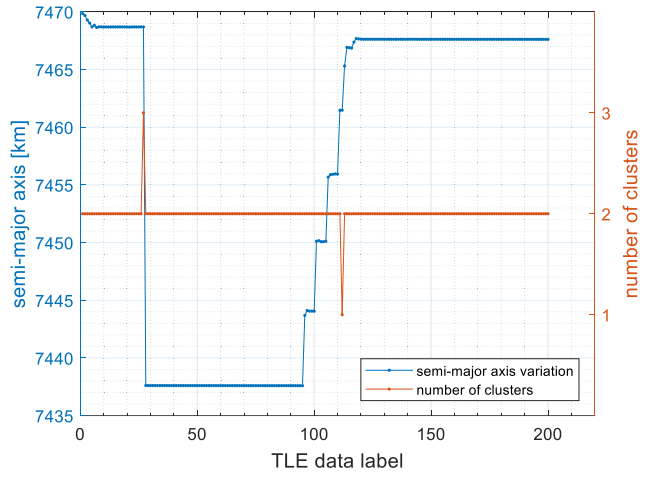


FIGURE 7. Semi-major axis variation curve and the hierarchical clustering result: the blue line indicates the semi-major axis variation with time and the red line indicates the clustering result.

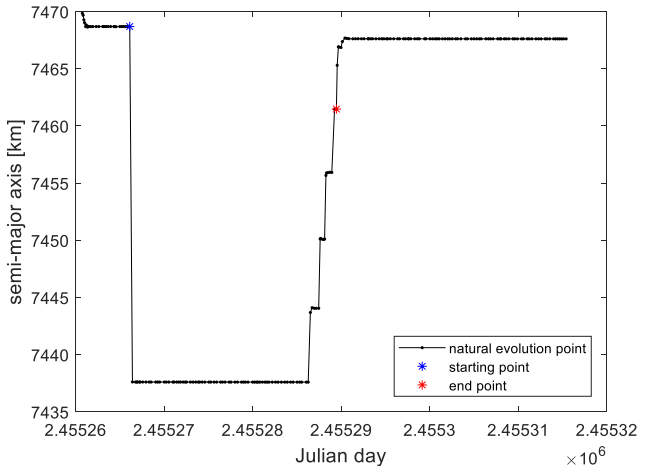


FIGURE 8. Semi-major axis curve marked with orbit control points by the hierarchical clustering method: the black line indicates the natural evolution process; the blue asterisk indicates the starting point of the orbital maneuver; and the red asterisk indicates the end point of the orbital maneuver.

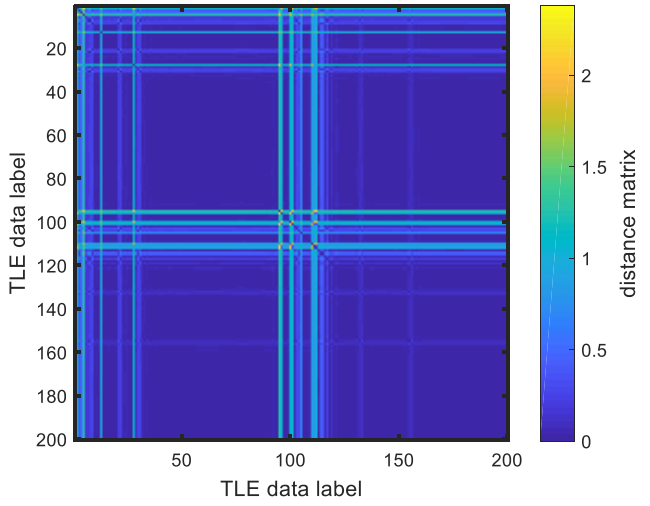


FIGURE 9. Distance matrix of the hierarchical clustering result: the color represents the degree of similarity of each cluster.

Figure 9 shows the distance matrix between the 200 data points. As there are fewer data points divided into Cluster 1 and Cluster 3, it can be seen that the hierarchical clustering method is not effective.

Figure 10 demonstrates the semi-major axis variation curve and the fuzzy C-means clustering results of satellite A. It can be seen that except for the descending and ascending points, there are some unclear points, which is similar to the K-means result This indicates that the fuzzy C-means clustering method misjudged some points. The orbital maneuver implemented in the deployment phase of satellite A is given in Figure 11. The semi-major axis corresponding to Cluster 1 has no obvious fluctuation, indicating that it represents the orbit natural evolution points, and Cluster 3 always appears after Cluster 2, indicating that Cluster 3 represents the starting point of the orbital maneuver, and Cluster 2 represents the end point of the orbital maneuver. Figure 12 illustrates the distance matrix between the 200 data points.

Table 3 provides a comparison of the three clustering methods in which the number of each cluster and the number of false and missed judgments are regarded as the comparative evaluation. It can be seen from Table 3 that the K-means clustering method can clearly distinguish different maneuver points despite two missed points and two fault points. The hierarchical clustering method can only detect the two points of maneuvering and the fuzzy C-means method detects several misjudged points. The missing rates of the three methods are 1%, 5%, and 1%, respectively, and the false rates are 1%, 0%, and 11%, respectively, meaning that both the hierarchical and fuzzy C-means methods cannot detect the orbit control points efficiently. Thus, the K-means method is chosen as the orbital detection method and the subsequent content on the detection of large-scale orbital maneuvers will be based on K-means.

It is confirmed that satellite A underwent one descending and four ascending maneuvers after entering the orbit.

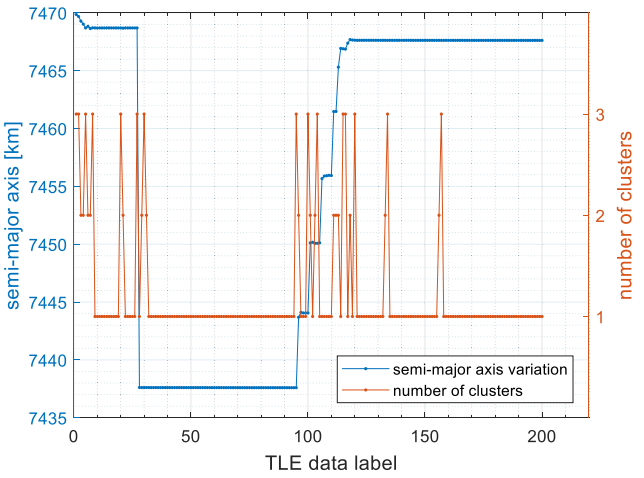


FIGURE 10. Semi-major axis variation curve and the fuzzy C-means clustering result: the blue line indicates the semi-major axis variation with time and the red line indicates the clustering result.

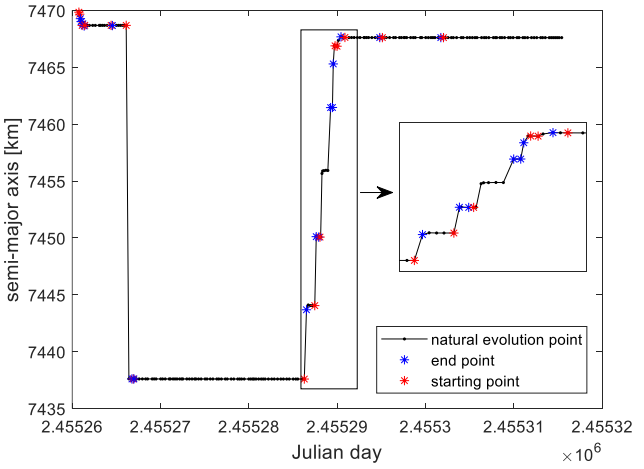


FIGURE 11. Semi-major axis curve marked with orbit control points by the fuzzy C-means clustering method: the black line indicates the natural evolution process; the blue asterisk indicates the starting point of the orbital maneuver; and the red asterisk indicates the end point of the orbital maneuver.

TABLE 3. Comparison of the three clustering methods.

Clustering method	Assessment	Number/Times
K-means clustering	1	188
	2	6
	3	6
	false	2
	missed	2
Hierarchical clustering	1	1
	2	198
	3	1
	false	0
	missed	10
Fuzzy C-means clustering	1	169
	2	16
	3	15
	false	22
	missed	2

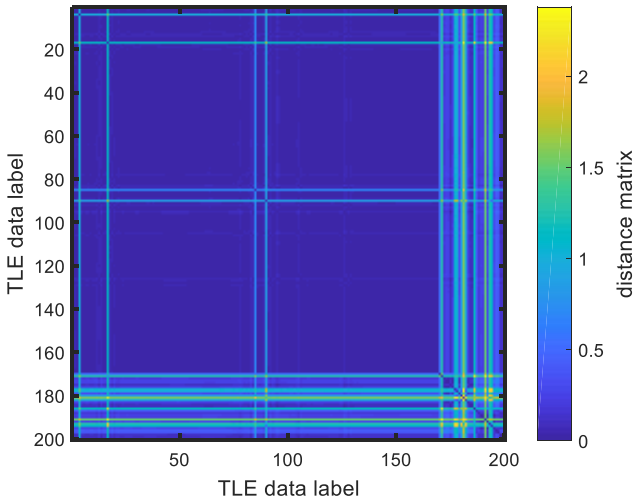


FIGURE 12. Distance matrix of the fuzzy C-means clustering result: the color represents the degree of similarity of each cluster.

However, based on engineering experience, it is difficult for the spaceborne orbital control engine to achieve a descending maneuver up to 30 km at a time, and the true descending process should have the same multiple maneuvers as the ascending process. In fact, there are another three additional TLE historical datasets of the spatial catalogs (the final stage of the rocket, the bracket, etc.), and through analysis of these data, it is found that the actual early descending maneuver of satellite A is recorded in the datalogs of the last stage of the rocket's orbit.

Figures 13 and 14 show the detection results through the K-means method for the early descending maneuver of satellite A. Consistent with the previous analysis, the descending process is performed three times. From the semi-major axis variation curve, there is a missed judgment at the second descending, and the actual descending process should contain four maneuvers. Furthermore, it should be noted that the in-orbit state of almost all the spatial targets fluctuates near a few points after orbit-injection, so the clustering results near the orbit-injection points are not discussed.

C. DETECTION OF MEDIUM-SCALE ORBITAL MANEUVER

Section III.B shows that the clustering method is effective in detecting large-scale orbital maneuvers with magnitudes of approximately 5 km at one time. While for medium-scale orbital maneuvers, the TLE historical data of "TIANHUI-1", of which the maneuver is approximately 50–200 m during the long-term control period, is used as the basis for the simulation. The characteristic variables and the number of clusters are reselected. Figure 15a illustrates the variation in the average contour value when different clusters and different characteristic variables are taken. When the number of clusters is greater than 6, the average contour value converges to around 0.7. Figure 15b shows an enlarged view of the

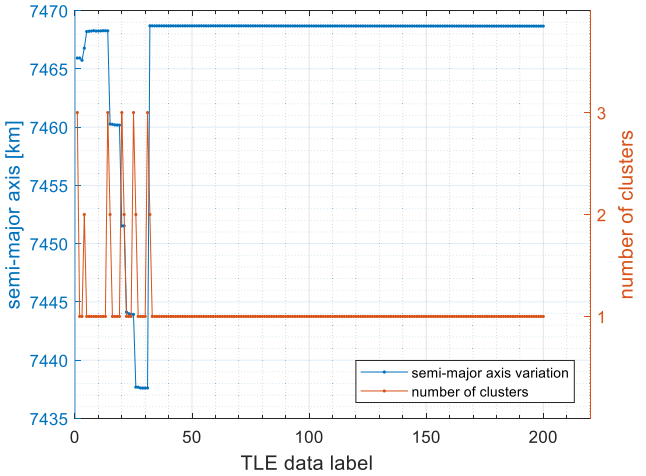


FIGURE 13. Early semi-major axis variation curve and the clustering result: the blue line indicates the early semi-major axis variation with time and the red line indicates the clustering result.

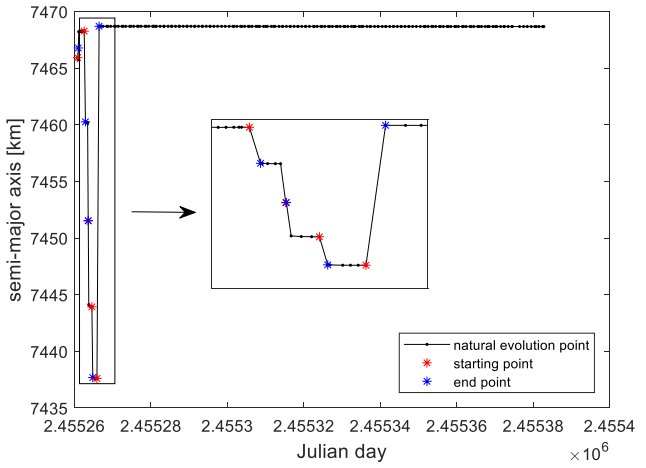


FIGURE 14. Early semi-major axis curve marked with orbital control points: the black line indicates the natural evolution process; the red asterisk indicates the starting point of the orbital maneuver; and the blue asterisk indicates the end point of the orbital maneuver.

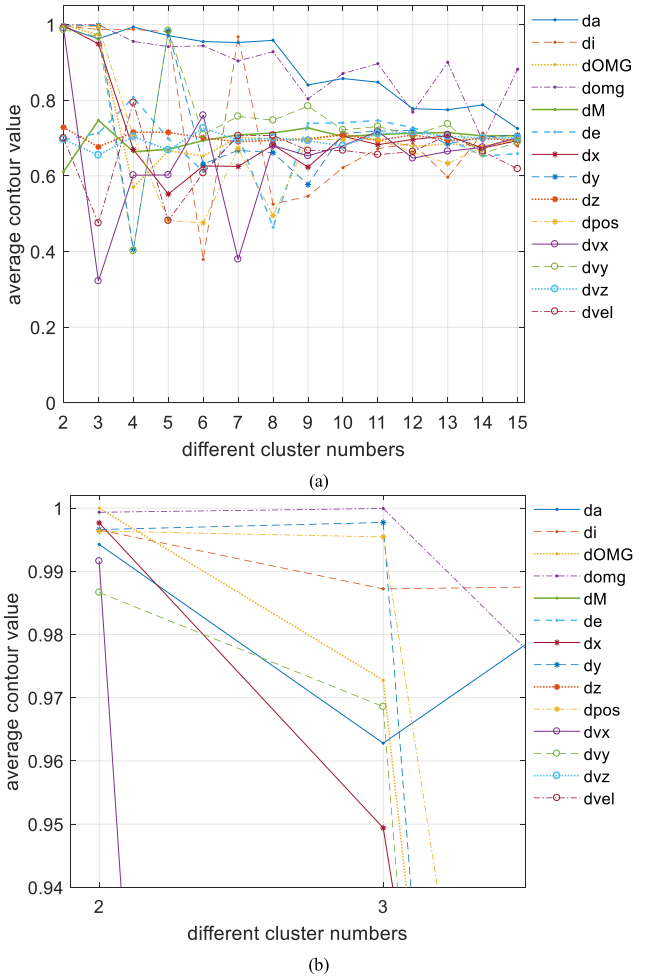


FIGURE 15. Average contour value in the case of different cluster numbers and different characteristic variables: TLE data from TIANHUI-1; a trend of average contour value corresponding to 14 characteristic variables as the number of cluster changed from 2 to 15; and b average contour value corresponding to each characteristic variable when the number of clusters was 2 and 3.

average contour values for 2 clusters and 3 clusters. Based on the order of the average contour values, $[\delta a, \delta i, \delta\Omega, \delta\omega]$ are selected as the feature variables.

After comparing the performance of the three clustering methods, the K-means method finally is used to perform orbital maneuver detection on the first 2000 TLE historical data points after "TIANHUI-1" was in orbit. The hierarchical and the fuzzy C-means clustering methods miss a lot of maneuvers.

Figure 16 shows the detection results when the number of clusters is 3, and 16 orbital maneuvers are detected. However, there are 14 missed points and 5 fault points, and the correspondence between the start and the end points of the maneuver is not accurate.

In the case where the number of clusters is 3, the time and the orbital control of 16 maneuvers are as shown in Table 4. It can be seen that the orbital control fluctuates between 100 and 1000 m, but lies mostly around 450 m, while the interval of control timing fluctuates between 15 and 35 d, but lies mostly around 25 d. The orbital control and maneuver are proportional to the orbital control intervals; this result is compatible with traditional orbital dynamics.

Through the above-mentioned orbital maneuver detection of the "TIANHUI-1" during the long-term control period, the applicability of the clustering method is verified; furthermore, the method shows good performance in the case of medium-scale orbital maneuvers.

D. DETECTION OF SMALL-SCALE ORBITAL MANEUVER

In this section, two conditions of small-scale orbital maneuvers are discussed: maneuvers by absolute orbit elements and maneuvers by relative orbit elements.

1) MANEUVERS BY ABSOLUTE ORBIT ELEMENTS

"Envisat", as an Earth observation satellite, provides observational parameters to improve environmental studies. A 2-year segment of the "Envisat" dataset running

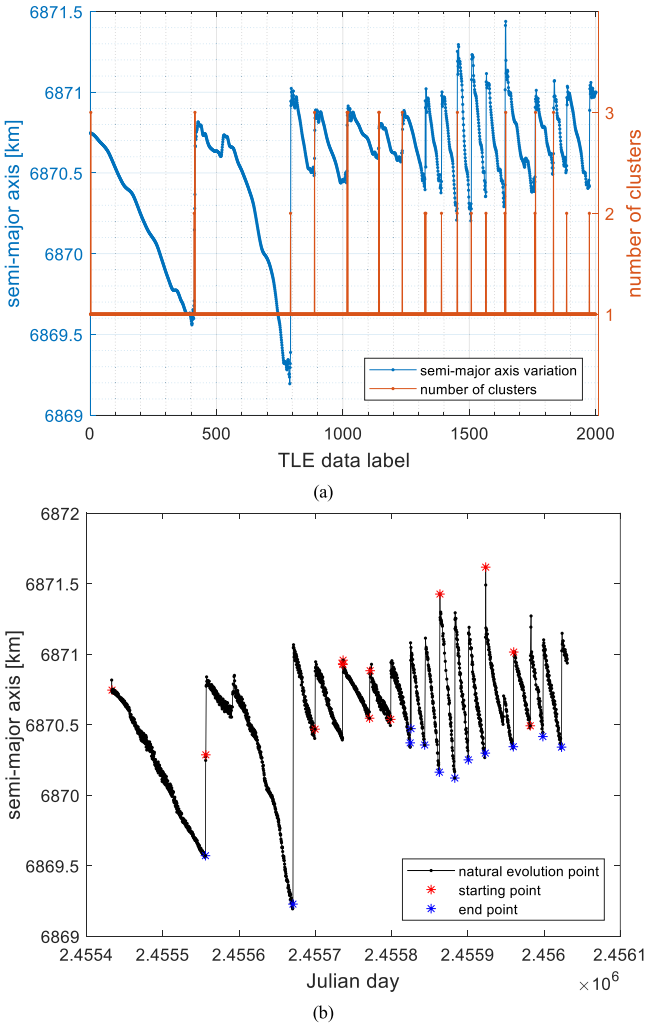


FIGURE 16. Semi-major axis variation curve and the clustering result of "TIANHUI-1": a the blue line indicates the semi-major axis variation with time; the red line indicates the clustering result; b the red asterisk indicates the start point of the orbital maneuver; the blue asterisk indicates the end point of the orbital maneuver.

TABLE 4. Time and orbital control of 16 maneuvers of "TIANHUI-1".

Number	Time	Orbital control (m)	Interval (day)
1	2010/12/25 10:25:6	440	/
2	2011/4/19 7:25:29	1556	114.8753
3	2011/5/18 3:32:53	174	28.83847
4	2011/6/23 7:56:6	284	36.18279
5	2011/7/28 19:38:28	122	35.48775
6	2011/8/25 12:30:50	133	27.70303
7	2011/9/20 1:54:57	550	25.55841
8	2011/10/9 1:31:19	532	18.98359
9	2011/10/28 5:38:42	974	19.17179
10	2011/11/17 11:12:40	800	20.23192
11	2011/12/5 3:2:24	682	17.65954
12	2011/12/27 9:46:57	590	22.28094
13	2012/2/2 7:43:51	431	36.91451
14	2012/2/24 12:46:12	283	22.20997
15	2012/3/12 2:59:27	422	16.59253
16	2012/4/5 10:36:33	472	24.31743

from 2003–2005 is examined by the clustering method. It can be easily seen that the orbit maneuver applied in Envisat during 2003 to 2005 is approximately 50 m and below, which

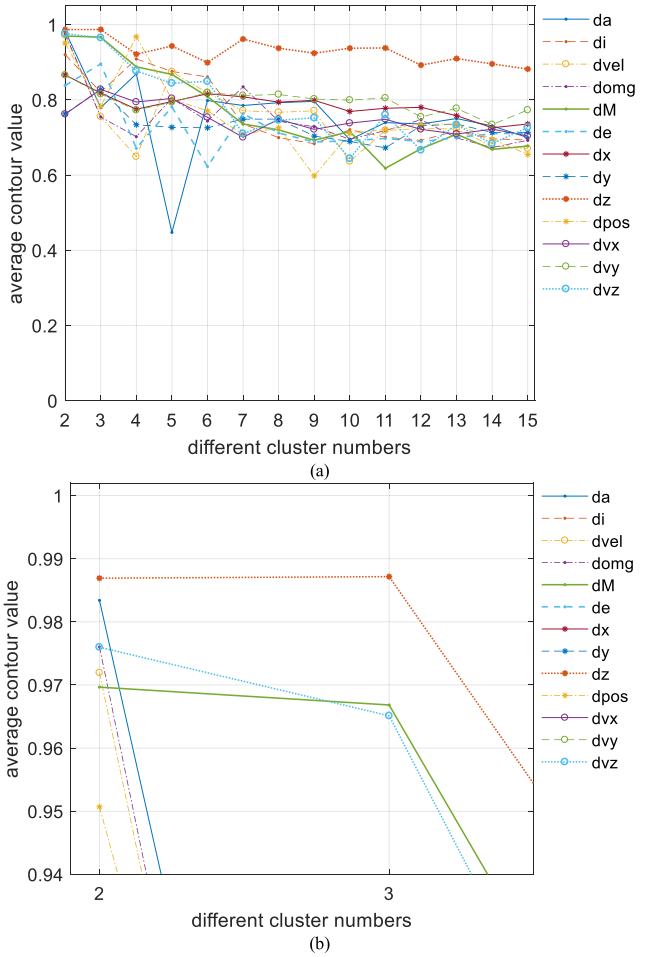


FIGURE 17. Average contour value in the case of different cluster numbers and different characteristic variables: a TLE data from "Envisat"; a trend of average contour value corresponding to 13 characteristic variables as the number of cluster changes from 2 to 15; b average contour value corresponding to each characteristic variable when the number of clusters is 2 and 3.

can be devoted to the detection of small-scale orbital maneuvers. The characteristic variables and the number of clusters are reselected. Figure 17 shows the variation in the average contour value in the case of different cluster numbers and different characteristic variables.

Figure 17a illustrates the variation in the average contour value when different clusters and different characteristic variables are taken. Figure 17b shows an enlarged view of the average contour values for 2 clusters and 3 clusters. In order of the average contour values, setting the number of clusters to 2 is a reasonable choice and $[\delta z, \delta a, \delta\omega]$ are selected as the feature variables. According to the results of the semi-major variation curves of the three methods with the orbital control points, the fuzzy C-means clustering method is useful in detecting the orbital maneuvers of "Envisat". The K-means and hierarchical clustering methods cannot effectively classify the TLE data.

As shown in Figure 18, the fuzzy C-means clustering method can clearly distinguish the semi-major axis increases

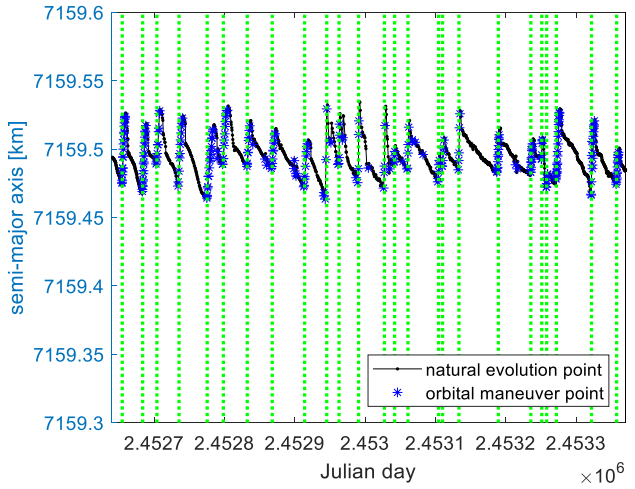


FIGURE 18. Semi-major axis variation curves marked with orbital control points: the black line indicates the natural evolution process; the blue asterisk indicates the orbital maneuver points; and the green dotted lines represent the detected orbital maneuvers.

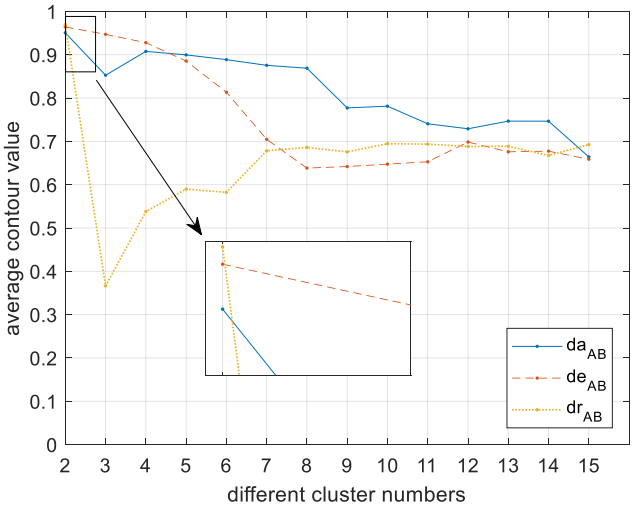


FIGURE 19. Average contour value in the case of different cluster numbers and different characteristic variables: the trend of the average contour value corresponding to 3 characteristic variables as the number of clusters changes from 2 to 15.

and decreases: the semi-major axis corresponding to Cluster 1 exhibits an ascending trend, and the semi-major axis corresponding to Cluster 2 exhibits a descending trend, indicating that Cluster 1 represents the orbital control process. To make it more apparent, the green dotted lines represent the detected orbital maneuvers by the clustering method. The false detection rate is a little over 1.4%.

To prove the efficiency of the proposed detection method, the comparison with the classical algorithm presented by Kelecy et al is made [20]. Their study focused on Envisat for which high-quality TLE data and known maneuver histories are available. The main differences and advantages are as follows: **1)** There are no complicated estimators to establish in our proposed method. While the detection proposed by Kelecy et al needs a linear polynomial filter and computes the filtered differences between each nearby data segment to find anomalously large differences; **2)** because the clustering methods are based on the similarity of data, the maneuver detection proposed in this paper is insensitive to the artificial threshold derived from the traditional experiences. However, a sensitivity trade-off between the algorithm parameters exists in the classical method. The filtering window length, the order of the filtering polynomial and the n-σ detection level can influence the performance metrics; **3)** and the false detection level of the clustering method is about 1.4%, which is lower than the 6% of the classical method used by Kelecy et al. 4) For computational complexity, the clustering methods are at least linearly proportional to the size of the data set, which is higher than that of the classic algorithm with linear complexity, i.e. $O(n)$, where n is the size of data.

2) MANEUVERS BY RELATIVE ORBIT ELEMENTS

When the satellite formation of ''YAOGAN-9'' was deployed in its late stage of orbit, there was moderately little orbital maintenance required at about 5–20 m devoted to the detection of small-scale orbital maneuvers. This section mainly focuses on satellite A and satellite B.

In fact, the detection of small-scale orbital maneuvers is different from that of large- or medium-scale orbital maneuvers. As the formation configuration changes periodically owing to the different orbit elements of each satellite, the relative orbit elements at each moment also change. Therefore, an orbital maneuver can be detected by the relative orbit elements of a fixed moment within the satellite orbit period; this requires the orbital data of each satellite in the formation to be synchronized. Because the sampling time interval is large and the sampling time of each satellite is different, the TLE data downloaded directly from the website must be processed to be used as a clustering database. First, the SGP4 model is used to predict the orbit data when satellite A runs to the orbit phase of 0, the orbital data of satellite B is also predicted at this time to obtain the relative orbit elements between the two satellites at the same time. Owing to certain noise in the TLE data, there are some errors in the prediction data, so wavelet filtering is needed to preprocess the data. The adjacent two datasets are differentiated and the relative orbit elements are converted into differential form. δa_{AB}, δe_{AB}, and δr_{AB} are chosen as the characteristic variables owing to the insignificance of the other relative orbit elements.

Figure 19 shows the variation in the average contour value in the case of different cluster numbers and different characteristic variables. It demonstrates the trend of the average contour value corresponding to 3 characteristic variables as the number of clusters changes from 2 to 15. It can be seen that when the number of clusters is greater than 2, the average contour value has a certain decline and the overall trend starts to diverge. Therefore, the number of clusters is determined to be 2 for the detection of small-scale orbital maneuvers. When the number of clusters is 2, it can be seen from Figure 19 that the average contour value of δr_{AB} is higher than that of δa_{AB} and δe_{AB}, so δr_{AB} is chosen as the characteristic variable.

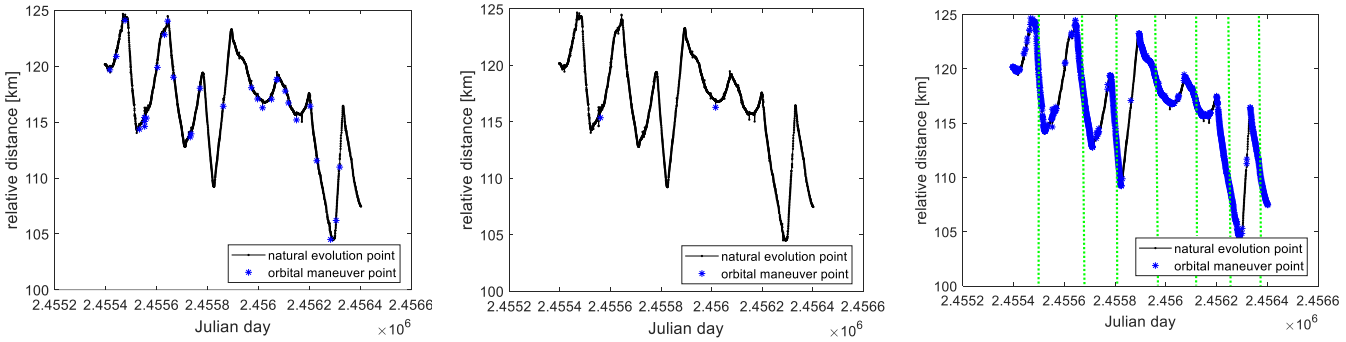


FIGURE 20. Relative distance variation curves marked with orbital control points: a by the K-means clustering method; b by the hierarchical clustering method; and c by the fuzzy C-means clustering method; the green dotted lines represent the detected orbital maneuvers.

Different from those of the large orbital control and the medium orbital control, the characteristic variable of the small-scale orbital maneuver detection, δr_{AB}, is a relative-position quantity, so it is necessary to compare the clustering methods again. Figure 20 shows the relative distance variation curves of the three methods with the orbital control points. In Figure 20a, only 7 points indicating the start or end maneuver points are detected by the K-means clustering method and there are some meaningless points. In Figure 20b, only 2 maneuver points are detected, which does not agree with the actual maneuvers, demonstrating that the hierarchical clustering method is not accurate. Both of these methods can not indicate the accurate orbital maneuvers. As shown in Figure 20c, it can be seen that the fuzzy C-means clustering method can clearly distinguish the relative distance increases and decreases: the relative distance corresponding to Cluster 1 exhibits an ascending trend, and the relative distance corresponding to Cluster 2 exhibits a descending trend, indicating that Cluster 2 represents the orbital control process. The green dotted lines represent the detected orbital maneuvers by the clustering method. The missing rate is approximately 4%. Therefore, the K-means and hierarchical clustering methods cannot effectively classify the relative distance data, while the fuzzy C-means clustering method is useful in detecting the orbital maneuvers.

IV. CONCLUSION

In this paper, the capabilities of data mining in detecting orbital maneuvers at different scales are explored, as well as its application to the remote sensing satellites "YAOGAN-9" "TIANHUI-1" and "Envisat". The unsupervised classification methods, K-means, hierarchical, and fuzzy C-means clustering, are used to process TLE historical data to obtain the maneuver points to record the orbital control action. Different scales of orbital maneuvers of the TLE data are simulated to evaluate the performance of the clustering methods. According to the results, the clustering methods can capture the orbital maneuver points although there are certain missed judgments and errors in the clustering results. The characteristic variables selection and cluster number determination are crucial to the accuracy of clustering methods. Through a series of numerical experiments, for the detection of orbital maneuvers at different scales, the three clustering methods have the following performances: The K-means clustering method is suitable for detecting large- and medium-scale orbital maneuvers, while the fuzzy C-means clustering method can detect small-scale orbital maneuvers more efficiently. The current work illustrates how to mine the law based on the existing orbital data of the Chinese remote sensing satellites. Future work can be conducted with the purpose of predicting the orbital maneuvers based on the historical data, combined with global or local coverage on the ground, eclipses, geometrical relationship with other satellites, and so on, which can be extended to satellites of other categories, except remotesensing satellites.

REFERENCES

[1] J. B. Campbell and H. W. Randolph, "History and scope of remote sensing," in *Introduction to Remote Sensing*, 5th ed. New York, NY, USA: Guilford, 2011, pp. 3–30.

[2] I. H. Witten, E. Frank, and M. A. Hall, "What's it all about," in *Data Mining: Practical Machine Learning Tools and Techniques*, 3rd ed. Cambridge, MA, USA: Morgan Kaufmann, 2011, p. 338.

[3] L. Self, "Use of data mining on satellite data bases for knowledge extraction," presented at the 13th Int. Florida Artif. Intell. Res. Soc. Conf., Orlando, Florida, May 2000.

[4] C. Sánchez-Sánchez, D. Izzo, and D. Hennes, "Learning the optimal state-feedback using deep networks," in *Proc. IEEE Symp. Series Comput. Intell. (SSCI)*, Athens, Greece, Dec. 2016, pp. 1–8. doi: 10.1109/SSCI.2016.7850105.

[5] B. Du, S. Li, Y. She, W. Li, H. Liao, and H. Wang, "Area targets observation mission planning of agile satellite considering the drift angle constraint," *Proc. SPIE*, vol. 4, Oct. 2018, Art. no. 047002.

[6] A. Polívka, M. Houdek, P. B. Ober, and M. Tossaint, "Satellite navigation data mining (SENDAI)," in *Proc. Int. Assoc. Inst. Navigat. World Congr. (IAIN)*, Prague, Czech Republic, Oct. 2015, pp. 1–6. doi: 10.1109/IAIN.2015.7352225.

[7] S. Tanner, C. Stein, and S. J. Graves, "On-board data mining," in *Scientific Data Mining and Knowledge Discovery*. Berlin, Germany: Springer, 2009, pp. 345–376. doi: 10.1007/978-3-642-02788-8_13.

[8] Z. Qu, G. Zhang, H. Cao, and J. Xie, "Leo satellite constellation for Internet of Things," *IEEE Access*, vol. 5, pp. 18391–18401, 2017. doi: 10.1109/ACCESS.2017.2735988.

[9] H. Gong, W. Zhao, and J. Li, "The technological framework of data mining from the polygenetic remotely sensed data," *J. Image Graph.*, vol. 10, no. 5, pp. 620–623, 2005.

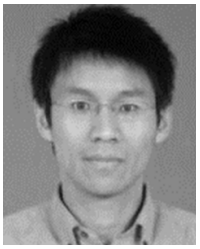
[10] H. Peng and X. Bai, "Recovering area-to-mass ratio of resident space objects through data mining," *Acta Astronautica*, vol. 142, pp. 75–86, Jan. 2018. doi: 10.1016/j.actaastro.2017.09.030.

[11] H. Peng and X. Bai, "Exploring capability of support vector machine for improving satellite orbit prediction accuracy," in *J. Aerosp. Inf. Syst.*, vol. 15, no. 6, pp. 366–381, Mar. 2018. doi: 10.2514/1.i010616.
[12] Z. J. Folcik, P. J. Cefola, and R. I. Abbot, "GEO maneuver detection for space situational awareness (AAS 07-285)," *Adv. Astron. Sci.*, vol. 129, no. 1, p. 523, 2008.
[13] D. P. Lubey, D. J. Scheeres, and R. S. Erwin, "Maneuver detection and reconstruction of stationkeeping spacecraft at GEO using the optimal control-based estimator," *IFAC-PapersOnLine*, vol. 48, no. 9, pp. 216–221, 2015. doi: 10.1016/j.ifacol.2015.08.086.
[14] Y. Bar-Shalom and K. Birmiwal, "Variable dimension filter for maneuvering target tracking," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-18, no. 5, pp. 621–629, Sep. 1982. doi: 10.1109/TAES.1982.309274.
[15] K. Spingarn and H. L. Weidemann, "Linear regression filtering and prediction for tracking maneuvering aircraft targets," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-8, no. 6, pp. 800–810, Nov. 1972. doi: 10.1109/TAES.1972.309612.
[16] J. S. Thorp, "Optimal tracking of maneuvering targets," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-9, no. 4, pp. 512–519, Jul. 1973. doi: 10.1109/TAES.1973.309633.
[17] R. P. Patera, "Space event detection method," *J. Spacecraft Rockets*, vol. 45, no. 3, pp. 554–559, May 2008. doi: 10.2514/1.30348.
[18] E. Doornbos, H. Klinkrad, and P. Visser, "Use of two-line element data for thermosphere neutral density model calibration," *Adv. Space Res.*, vol. 41, no. 7, pp. 1115–1122, 2008. doi: 10.1016/j.asr.2006.12.025.
[19] S. Lemmens and H. Krag, "Two-line-elements-based maneuver detection methods for satellites in low earth orbit," *J. Guid., Control, Dyn.*, vol. 37, no. 3, pp. 860–868, Mar. 2014. doi: 10.2514/1.61300.
[20] T. Kelecy, D. Hall, K. Hamada, and D. Stocker, "Satellite maneuver detection using two-line elements data," presented at the Adv. Maui Opt. Space Surveill. Technol. Conf., Maui, Hawaii, Sep. 2007.
[21] F. R. Hoots and R. L. Roehrich, "Spacetrack report no. 3, models for propagation of NORAD element sets," U.S. Air Force Aerosp. Defense Command, Colorado Springs, CO, USA, Tech. Rep. IEICE 103(531), Dec. 1980.
[22] Z. Bu, J. Cao, H.-J. Li, G. Gao, and H. Tao, "GLEAM: A graph clustering framework based on potential game optimization for large-scale social networks," *Knowl. Inf. Syst.*, vol. 55, no. 3, pp. 741–770, Sep. 2017. doi: 10.1007/s10115-017-1105-6.
[23] Z. Bu, H.-J. Li, J. Cao, Z. Wang, and G. Gao, "Dynamic cluster formation game for attributed graph clustering," *IEEE Trans. Cybern.*, vol. 49, no. 1, pp. 328–341, Jan. 2019. doi: 10.1109/tcyb.2017.2772880.
[24] H.-J. Li, Z. Bu, Z. Wang, J. Cao, and Y. Shi, "Enhance the performance of network computation by a tunable weighting strategy," *IEEE Trans. Emerg. Topics Comput.*, vol. 2, no. 3, pp. 214–223, Jun. 2018. doi: 10.1109/tetci.2018.2829906.
[25] V. Estivill-Castro and J. Yang, "Fast and robust general purpose clustering algorithms," presented at the Pacific Rim Int. Conf. Artif. Intell., Berlin, Germany, Aug. 2000. doi: 10.1007/3-540-44533-1_24.
[26] H.-J. Li, Z. Bu, Y. Li, Z. Zhang, Y. Chu, G. Li, and J. Cao, "Evolving the attribute flow for dynamical clustering in signed networks," *Chaos, Solitons Fractals*, vol. 110, pp. 20–27, May 2018. doi: 10.1016/j.chaos.2018.02.009.
[27] B. Tjaden and J. Cohen, "A survey of computational methods used in microarray data interpretation," *Appl. Mycol. Biotechnol.*, vol. 6, pp. 161–178, 2006. doi: 10.1016/s1874-5334(06)80010-9.
[28] O. Maimon and L. Rokach, *DataMining And Knowledge Discovery Handbook*. New York, NY, USA: Springer, 2005. doi: 10.1007/b107408.
[29] S. Saket and S. Pandya, "An overview of partitioning algorithms in clustering techniques," *Int. J. Adv. Res. Comput. Eng. Technol.*, vol. 5, no. 6, pp. 1943–1946, Jun. 2016.
[30] S. Agarwal, "Data mining: Data mining concepts and techniques," in *Proc. Int. Conf. Mach. Intell. Res. Advancement*, Katra, India, Dec. 2013, pp. 203–207. doi: 10.1109/ICMIRA.2013.45.
[31] B. Anukampa and C. Sujata, "A comparative study on hierarchical, K-means and fuzzy C-means clustering algorithms and application to microarray gene expression data," *Int. J. Adv. Res. Eng. Technol.*, vol. 3, no. 1, pp. 36–41, Jan. 2015.
[32] M. Kaur and U. Kaur, "Comparison between K-mean and hierarchical algorithm using query redirection," *Int. J. Adv. Res. Comput. Sci. Softw. Eng.*, vol. 3, no. 7, pp. 1454–1459, Jul. 2013.
[33] A. K. Jain, M. N. Murty, and P. J. Flynn, "Data clustering: A review," *ACM Comput. Surv.*, vol. 31, no. 3, pp. 264–323, Sep. 1999.
[34] P. Berkhin, "A survey of clustering data mining techniques," in *Grouping Multidimensional Data*, vol. 43, no. 1. Berlin, Germany: Springer, 2006, pp. 25–71.
[35] M. J. Sidi, "Orbit dynamics," in *Spacecraft Dynamics and Control: A Practical Engineering Approach*, vol. 7. New York, NY, USA: Cambridge Univ. Press, 2000.
[36] M. J. Sidi, "Spacecraft dynamics and control: A practical engineering approach," *J. Spacecraft Rockets*, vol. 34, no. 6, pp. 851–852, 1997.
[37] K. Zhang, F.-Q. Zhou, and J. Xiong, "Research of multi-satellite synchronous orbit determination based on serial integration," in *Proc. Int. Conf. Multimedia Technol.*, Hangzhou, China, Jul. 2011, pp. 2536–2540. doi: 10.1109/ICMT.2011.6002407.
[38] T. Wu, X. Hu, Y. Zhang, L. Zhang, P. Tao, and L. Lu, "Automatic cloud detection for high resolution satellite stereo images and its application in terrain extraction," *ISPRS J. Photogram. Remote Sens.*, vol. 121, pp. 143–156, Nov. 2016. doi: 10.1016/j.isprsjprs.2016.09.006.
[39] B. Barrett, C. Pratola, A. Gruber, and E. Dwyer, "Intercomparison of soil moisture retrievals from *in situ*, ASAR, and ECV SM data sets over different European sites," in *Satellite Soil Moisture Retrieval Techniques and Applications*. Amsterdam, The Netherlands: Elsevier, 2016, pp. 209–228. doi: 10.1016/b978-0-12-803388-3.00011-5.

XUE BAI is currently pursuing the B.S. degree in spacecraft design &engineering with the School of Astronautics, Beihang University, Beijing, China. Her research interests include formation flying, circular restricted three-body problem, and orbital big data mining.

CHUAN LIAO received the B.S. degree from the School of Aerospace Engineering, Beijing Institute of Technology, Beijing, China, in 2016, and the M.S. degree in aerospace engineering from Beihang University, Beijing, in 2019. He is currently with the Research Institute, China Electronics Technology Group Corporation, Chengdu, China. His research interests include spacecraft orbital dynamics and control, orbital big data mining, and inversion.

XIAO PAN received the B.S. degree from the School of Astronautics, Northwestern Polytechnical University, Shaanxi, China, in 2015. She is currently pursuing the Ph.D. degree in spacecraft design & engineering with Beihang University, Beijing, China. Her research interests include circular restricted three-body problem, trajectory design, navigation technology, and formation flying.



MING XU received the B.S. and Ph.D. degrees in aerospace engineering from Beihang University, Beijing, China, in 2003 and 2008, respectively. He was an Engineer of orbital design and operation with DFH Satellite Company Ltd., China Academy of Space Technology, Beijing, until 2010. He is currently an Associate Professor with the School of Astronautics, Beihang University. He has 50 publications in journals, books, and proceedings. His current research interests include the applications of dynamical systems theory into astrodynamics and orbital control. He received National Top 100 Excellent Doctoral Dissertation Award nomination, in 2010, and Third Class Prizes of the National Defense Technology Invention Award, in 2016. He serves as associate editors for the journals of astrodynamics and *Advances in Aircraft and Spacecraft Science*.