Data matching and association based on the arc-segment difference method Postprint

Authors: Jiannan Sun, Zhe Kang, Zhenwei Li, Cunbo Fan, Cunbo Fan

Date: 2025-12-03T00:00:00+00:00

Abstract

In response to the issue of fuzzy matching and association when optical observation data are matched with the orbital elements in a catalog database, this paper proposes a matching and association strategy based on the arcsegment difference method. First, a matching error threshold is set to match the observation data with the known catalog database. Second, the matching results for the same day are sorted on the basis of target identity and observation residuals. Different matching error thresholds and arc-segment dynamic association thresholds are then applied to categorize the observation residuals of the same target across different arc-segments, yielding matching results under various thresholds. Finally, the orbital residual is computed through orbit determination (OD), and the positional error is derived by comparing the OD results with the orbit track from the catalog database. The appropriate matching error threshold is then selected on the basis of these results, leading to the final matching and association of the fuzzy correlation data. Experimental results showed that the correct matching rate for data arc-segments is 92.34% when the matching error threshold is set to 720 , with the arc-segment difference method processing the results of an average matching rate of 97.62% within 8 days. The remaining 5.28% of the fuzzy correlation data are correctly matched and associated, enabling identification of orbital maneuver targets through further processing and analysis. This method substantially enhances the efficiency and accuracy of space target cataloging, offering robust technical support for dynamic maintenance of the space target database.

Full Text

Preamble

Astronomical Techniques and Instruments, Vol. 2, September 2025, 299-309

Article

Open Access

Data Matching and Association Based on the Arc-Segment Difference Method

Jiannan Sun[1,2], Zhe Kang[1], Zhenwei Li[1], Cunbo Fan[3*]

[1]Changchun Observatory, National Astronomical Observatories, Chinese Academy of Sciences, Changchun 130117, China
[2]University of Chinese Academy of Sciences, Beijing 100049, China
[3]Changchun Branch, Chinese Academy of Sciences, Changchun 130022, China

*Correspondence: fancb@ms.ccb.ac.cn

Received: May 6, 2025; Accepted: May 20, 2025; Published Online: May 21, 2025
https://doi.org/10.61977/ati2025029; https://cstr.cn/32083.14.ati2025029

Citation: Sun, J. N., Kang, Z., Li, Z. W., et al. 2025. Data matching and association based on the arc-segment difference method. *Astronomical Techniques and Instruments*, 2(5): 299−309. https://doi.org/10.61977/ati2025029.

Abstract: In response to the challenge of fuzzy matching and association when optical observation data are matched with orbital elements in a catalog database, this paper proposes a matching and association strategy based on the arc-segment difference method. First, a matching error threshold is established to match observation data against a known catalog database. Second, the matching results for the same day are sorted by target identity and observation residuals. Different matching error thresholds and arc-segment dynamic association thresholds are then applied to categorize the observation residuals of the same target across different arc-segments, yielding matching results under various thresholds. Finally, the orbital residual is computed through orbit determination (OD), and the positional error is derived by comparing the OD results with the orbit track from the catalog database. The appropriate matching error threshold is then selected based on these results, leading to the final matching and association of the fuzzy correlation data. Experimental results showed that the correct matching rate for data arc-segments is 92.34% when the matching error threshold is set to 720 , with the arc-segment difference method processing the results of an average matching rate of 97.62% within 8 days. The remaining 5.28% of the fuzzy correlation data are correctly matched and associated, enabling identification of orbital maneuver targets through further processing and analysis.

This method substantially enhances the efficiency and accuracy of space target cataloging, offering robust technical support for dynamic maintenance of the space target database.

Keywords: Optical data processing; Space target identification; Fuzzy correlation; Arc-segment difference method; Orbit determination

1. INTRODUCTION

In recent years, rapid developments in commercial space activities, together with the frequent occurrence of space target collisions and explosive disintegration events, have caused a sharp increase in the number of space targets in orbit. These occurrences have aggravated orbital congestion and are highly prone to triggering cascading collisions, thereby leading to the formation of the Kessler Syndrome [1]. As of December 24, 2024, the number of in-orbit space targets in Space-Track' s publicly released Two-Line Element (TLE) database stands at 30,010. This represents an increase of approximately 1,487 targets compared with the same period in the previous year [2]. The steadily growing number of space targets poses a considerable challenge to the maintenance and expansion of the Space Surveillance and Tracking (SST) database. To address this challenge, continuous observation of space targets is essential to accurately determine their orbital parameters. Consequently, the multi-target capture observation model for sky surveying is used widely in space target cataloging [3]. Nevertheless, an observation mode employing directly multi-target acquire the specific attribute information of a target. To further utilize the observation data, matching and association operations must be performed on the observation arc-segments. identification techniques cannot When the data processing center receives a large volume of observational data, the first step is to match the data against the orbital elements of the catalog database. For observations that do not match successfully, initial orbit determination (IOD) is employed for correlation [4].

The North American Aerospace Defense Command (NORAD) publicly available database is the most commonly used orbital cataloging repository for a wide range of aerospace workers, and TLE data for most space targets are regularly updated on a daily basis through the Space-Track website. The matching rate is 83%−85% when using the TLE database [5,6]. Calculation of TLE data requires the use of an appropriate Simplified General Perturbation Version 4 (SGP4) or Simplified Deep-space Perturbation Version 4 (SDP4) orbital propagation models [7]. Because the accuracy of SGP4/SDP4 models is not published and the TLE data do not contain the corresponding orbital accuracy information, fuzzy matching is possible when using this information, which might affect the accuracy of the space target catalog.

To reduce the probability of fuzzy matching, it is essential to fully understand the accuracy of the propagation model and the orbital accuracy details of the various orbital altitude targets in the TLE database. These can be used as reference values for setting the matching threshold. In 2008, Flohrer analyzed the accuracy of TLE data for 11,286 space targets and indicated that the entire dataset could potentially have a maximum standard deviation of approximately 5 km [8]. Regarding Low Earth Orbit (LEO) satellites, Reising analyzed the accuracy of 634 sets of TLE data for the Flock 1B satellite over the period of 1 month. When using the SGP4 model for 1-error propagation, the positional change was 10−30 km after 1 day of propagation, and 20−70 km after 2 days, mainly in the along-track direction [9]. Typically, the TLE data of LEO satellites are updated every few hours. If not updated promptly, owing to the influence of atmospheric drag, the propagation error of the SGP4 model in the along-track direction will increase markedly. The along-track error can reach several hundred of kilometers in a 7-day prediction [10]. For medium-orbit and high-orbit satellites, Racelis and Joerger' s research analysis showed that during 2013–2015, the TLE along-track errors for most cases ranged from −15 to 15 km, but during a later period (2015–2018), these errors were more constrained, staying between −6 and 6 km [11]. In the case of Elliptical Orbit (EO)/High Elliptical Orbit (HEO) satellites, Wei' s research demonstrated that the larger the eccentricity, the greater the orbital prediction error: The 1-day prediction positional error was within 20 km, and the 3-day prediction positional error was within 100 km [12]. Früh and Schildknecht' s study also revealed a 0.08° deviation between the observation data and the TLE ephemeris of the closest observation moment [13]. Therefore, when choosing the TLE database for target matching, the effects of both the inherent errors in the TLE data and the orbital propagation errors specific to different orbit types must be considered. Furthermore, the handling of orbital parameter anomalies within the TLE data is of vital importance [14]. By setting an appropriate matching error threshold, the correct matching rate of data arc-segments can be improved, thereby reducing the probability of fuzzy correlation.

To address the issue of incorrect associations for GEO targets in dense trajectories, Song proposed a real-time GEO target association algorithm that utilizes two-dimensional judgments of radar ranging and speed. This algorithm improves the association accuracy but is limited by the type of observation data available [15]. Tao put forward a correlation method by analyzing the orbit determination (OD) results of correlated data, achieving the correct correlation of data from different passes of the same target. However, this approach is also constrained by the conditions of OD [16]. Research on data association algorithms has matured remarkably, both domestically and internationally. Notable methods include the nearest-neighbor method, joint probability data association, multi-hypothesis tracking, the IOD method, and the admissible region algorithm [17–19]. All of these algorithms focus on associations between data arc-segments rather than between data arc-segments and known target orbital parameters. Consequently, they do not extract attribute information from data arc-segments, which belong to the category of uncorrelated target processing. From a practical engineering perspective, accurately matching observation data with the orbital elements of the catalog database is essential for enhancing the correct matching rate and improving the ability to analyze anomalous matching data. This improvement is crucial for rapidly advancing space situational awareness and enhancing the cataloging and maintenance capabilities of space targets.

Here, we propose a data matching and association strategy based on the arc-segment difference method to address the fuzzy matching and association issues of observation data during the matching process. Using the matching error threshold and the arc-segment dynamic association threshold, our approach processes the observation residuals of different arc-segments belonging to the same target. Consequently, for every individual arc-segment, one of the following four types of matching and association results will be generated: True match & True association, True match & False association, False match & True association, and False match & False association. The accuracy of these matching correlation results is validated through OD. By processing false matching association data, we can effectively identify elliptical orbit targets, targets with sudden orbital changes (such as maneuvers, collisions, explosions, and fallout), and newly detected targets. Such data can serve as the primary resource for studying these targets, demonstrating substantial application value.

The remainder of this paper is structured as follows. Section 2 details the data matching principle and the matching association strategy of the arc-segment difference method. Section 3 presents the experimental results and validation under different matching error thresholds. Moreover, the application of fuzzy association data is introduced. Section 4 analyzes the causes of fuzzy associations. Finally, Section 5 concludes the paper.

2.1. Data Matching Principle

The positional information derived from the TLE data, computed by SGP4/SDP4 models, is given within the True Equator and Mean Equinox of date (TEME) coordinate system, which is denoted as r_{TEME}. For the Optical Survey Telescope (OST), the coordinate system of observation data obtained using the astronomical positioning method is the J2000 coordinate system. Additionally, the station position represents a state quantity within the Earth-Centered, Earth-Fixed (ECEF) coordinate system, denoted as r_b. When data matching is performed, both the TEME and ECEF coordinate systems need to be converted into the J2000 coordinate system. The relevant transformation relations are expressed as follows [20,21]:

$$r_{\text{brJ2000}} = [M_Z(\Delta \cos \epsilon_0) N A]^T r_{\text{TEME}}; \tag{1}$$

$$r_{\text{J2000}} = [B_2 B_1 N A]^T r_b; \tag{2}$$

where $M_Z(\Delta\cos\epsilon_0)$ represents the matrix for counterclockwise rotation of angle $\Delta\cos\epsilon_0$ around the Z-axis in the original coordinate system, where $\Delta\epsilon_0$ denotes the nutation in longitude; B_2 represents the true obliquity of the ecliptic; B_1 is the polar motion correction matrix; N is the Earth' s rotation matrix; A is the nutation correction matrix; superscript T denotes the transposed transformation matrix; α denotes the right ascension; and δ denotes the declination.

Assuming that the position vector of the space target at time t, which has undergone orbit prediction and been transformed into the J2000 coordinate system, is denoted $R_t(x_t, y_t, z_t)$, and the position vector of the station after coordinate system transformation is $R_{st}(x_{st}, y_{st}, z_{st})$. Then, the position vector of the space target in the J2000 coordinate system with the station as the center, denoted $R_{ct}(x_{ct}, y_{ct}, z_{ct})$, is given by the following:

$$R_{ct}(x_{ct}, y_{ct}, z_{ct}) = R_t(x_t, y_t, z_t) - R_{st}(x_{st}, y_{st}, z_{st}). \tag{3}$$

The theoretical observation value of the space target at this moment is expressed as follows:

$$\begin{cases} \alpha_{ct} = \arctan(y_{ct}/x_{ct}) \\ \delta_{ct} = \arcsin(z_{ct}/\sqrt{x_{ct}^2 + y_{ct}^2 + z_{ct}^2}) \end{cases} \tag{4}$$

The observation residual for the data $(\alpha_{ot}, \delta_{ot})$ is then calculated using Equation (5):

$$\Phi_i = \sqrt{[(\alpha_{ot} - \alpha_{ct})\cos\delta_{ot}]^2 + (\delta_{ot} - \delta_{ct})^2}. \tag{5}$$

After traversing all the targets in the catalog database, the root mean square (RMS) error of the observation residuals is acquired for the entire data arc-segment. The identity information of the target indicated by the minimum RMS error that satisfies the matching error threshold is taken as the initial matching result for that data arc-segment. This initial matching result is marked by a NORAD identification number.

2.2. Matching and Association Strategy Based on Arc-Segment Difference Method

By setting the matching error threshold, we can obtain the identity information of the observation arc-segments corresponding to the known target in the catalog database. If a small matching error threshold is set, the matching result will have high accuracy. However, this will cause the matching success rate of the observation data to be low, and it will affect the data utilization rate. Conversely, if a large matching error threshold is set, the matching success rate of the observation data will be high, but it will result in the formation of fuzzy matching. This will reduce the matching accuracy and affect the cataloging accuracy. Therefore, while ensuring a high data matching rate, it is essential to reprocess the initial matching results. This helps resolve the problem of fuzzy matching and guarantees that each observation arc-segment is correctly matched and associated with other arc-segments.

On the basis of the accuracy of TLE data and the characteristics of orbital error propagation for targets at different altitudes, we adjust the data matching error threshold and the arc-segment dynamic association threshold to classify more precisely the initial matching results. This process ensures correct matching of observation data to known targets and accurate association of data arc-segments.

The data matching error threshold determines whether the observation data are correctly matched to a known target, while the arc-segment dynamic association threshold assesses whether the data arc-segments matched to the same target are correctly associated. After applying these criteria, the initial matching results are categorized into four types.

True match & True association means that the observation data have been matched to the correct target in the database and that the arc-segments have been correctly associated with each other. Therefore, the correct matching rate can be calculated. True match & False association means that the correct matching data arc-segment of the target already exists for the same day, and that according to the dynamic association threshold value, the data arc-segment has a false association; namely, the data do not belong to the marked target. This data type serves as a source for newly detected targets. False match & True association means that the observation residual exceeds the match threshold value, while the arc-segment difference meets the dynamic association threshold between arc-segments. Thus, the arc-segment from the same target data can be accumulated. False match & False association means that the residuals of the data arc-segments exceed the matching association threshold both in matching and in association between arc-segments, but are lower than the set threshold for matching error.

The overall procedure is described in the following.

(1) The matching error threshold of the observation data is set to ΔT, and the dynamic association threshold of the arc-segment is set to ΔD:

$$\Delta D = \delta_i \times (\lfloor \frac{|T_{i+1} - T_i|}{P} \rfloor + 1); \tag{6}$$

where δ_i is the RMS error of the observation residuals of the i-th observation arc-segment, and its unit is arc-seconds; $\lfloor\rfloor$ indicates the downward rounding of the value therein; T_i means the starting observation moment of the i-th observation arc-segment; P represents the orbital period of the matching target, and its unit is minute; and δ_i varies with the sequential order of the observation arc-segments.

(2) The initial matching results are sorted in descending order by target identity and the RMS error values of the observation residuals for each observation arc-segment of the same target. The discrepancy of the arc-segment is calculated as $\Phi' = \Phi_{i+1} - \Phi_i$. The final matching and association result for each observation arc-segment is decided based on the following relationship, and as shown in Fig. 1 [Figure 1: see original paper].

This figure illustrates two relationships: the data matching relationship between observation residuals and the reset matching error threshold (horizontal coordinate), and the arc-segment association relationship between the observation residual difference and the dynamic association threshold (vertical coordinate). If there is only one target-numbered observation arc-segment in the initial matching result, the matching error threshold is used to determine whether the result is a case of True match & True association or False match & False association. The flow chart of the process of data matching and association processing is shown in Fig. 2 [Figure 2: see original paper].

3.1. Evaluation of Experimental Data Accuracy

A data matching and association experiment was performed using observation data collected from an array of photoelectric telescopes at the Changchun Observatory, National Astronomical Observatories, Chinese Academy of Sciences. The parameters of the telescopes are detailed in Fig. 3 [Figure 3: see original paper] and Table 1 .

Observation data were acquired from September 21-28, 2024. During this period, a cooperative target was selected to evaluate the measurement accuracy of the telescope using its Consolidated Prediction Format (CPF) Version 2 ephemeris, which can be accessed via a file transfer protocol site [22]. The CPF ephemeris provides positional information for targets in a geocentric coordinate system, with accuracy typically within several meters for one-day prediction [23], which is sufficient for evaluation of the accuracy of observation data of 9 . Additionally, this assessment is intended to verify the stability of the operating conditions of the telescope, as detailed in Table 2 .

During the observation period, the telescope demonstrated an RMS error of 2.88 in right ascension and 2.84 in declination, resulting in a combined RMS error of 4.03 for the Jason-3 satellite at an altitude of 1,336 km. These values suggest that the measurement accuracy of the telescope remained relatively stable throughout the observation period.

3.2. Matching and Association Results

The matching error thresholds were set to 60 , 180 , 360 , 720 , 1,080 , 1,800 , and 3,600 , and the average matching success rates of the observation data within 8 days were obtained as 40.56%, 65.43%, 79.98%, 91.77%, 94.97%, 96.75%, and 97.62%, respectively, as shown in Fig. 4 [Figure 4: see original paper].

As depicted in Fig. 4, the matching success rate rises with an increase in the matching error threshold. Evidently, setting a larger matching error threshold enables a higher matching success rate, which is beneficial for maximizing the utilization of observation data. However, a larger threshold also results in fuzzy data matching and ambiguous arc-segment correlations. This will impact the cataloging accuracy and the handling of special target scenarios, such as orbital maneuvers and the detection of new targets.

The matching association strategy based on the arc-segment difference method described in Section 2.2 effectively addresses the aforementioned issues. Given the accuracy of LEO target TLE data and the error propagation characteristics, in addition to the matching results presented in Fig. 4, the matching error thresholds for the observation data were reset to 60 , 180 , 360 , 720 , 1,080 , and 1,800 .

A method involving comparison of the difference between the dynamic association threshold and the arc-segment residual is used to handle the initial matching results. By processing the results, which have an average matching rate of 97.62% within an 8-day period, we obtain four types of results, among which the proportions of the data marked as True match & True association are 44.61%, 69.96%, 84.03%, 94.79%, 97.58%, and 99.15% for the matching error thresholds of 60 , 180 , 360 , 720 , 1,080 , and 1,800 , respectively, as shown in Table 3 .

As shown in Table 3, within the 8-day period, the number of data arc-segments with True match & True association steadily increases as the data matching error threshold rises. Correspondingly, the association success rate also gradually improves, primarily because of the data arc-segments classified as False match & True association. When the matching error threshold exceeds 180 , the results for True match & False association remain unaffected by the magnitude of the threshold. Instead, they are determined by the difference between the residuals of the arc-segments and the dynamic association threshold, with each target typically corresponding to 1–2 passes of data. In the case of False match & True association, a low matching error threshold corresponds to an average of nearly 10 passes of data per target. As the threshold increases, this average is reduced to approximately 3–4 passes per target. For False match & False association, as the matching error threshold rises, the number of False match & False association arc-segments gradually declines, resulting in each space target generally corresponding to 1–2 passes of data.

3.3. Verification of Matching and Association Results

To evaluate the rationality of the settings for the matching error threshold and the dynamic correlation threshold, the OD method is employed. This method is validated through the OD residual. The principle of OD is not the focus of this paper, but relevant details can be found in the book by Montenbruck et al. [24]. Because the data of True match & False association and False match & False association cannot meet the requirements for OD, processing of False match & True association data is conducted to verify the appropriately chosen matching error threshold. Among the False match & True association data, data arc-segments with observation residual errors falling between two adjacent thresholds are selected. Moreover, arc-segments where the same target has observation data for 3 consecutive days are used for OD. The OD result is then compared to the TLE data, with consideration of the value of the OD residual. If the two values change steadily, then the data of this arc-segment belongs to the case of True match & True association; otherwise, they belong to the case of False match & True association.

On the basis of the observation residual data of False match & True association presented in Table 3, data arc-segments whose RMS errors lie within the adjacent matching error threshold intervals are chosen for the purpose of OD. The criterion adopted for OD is that the same target is selected to have at least 3-minute data arc-segments over three consecutive days, with each day comprising 1–2 data passes. Analysis reveals that 165 targets with RMS errors between 60 and 180 meet the OD criteria. In the interval from 180 to 360 , the number of such targets is 143. For the interval from 360 to 720 , there are 88 targets. In the range from 720 to 1,080 , only five targets satisfy the condition. There is merely a single target in the interval from 1,080 to 1,800 . One target was randomly selected from each of these intervals to perform OD, and the corresponding observation data are displayed in Table 4 .

The 1,080 threshold had too few data points for OD to converge; however, for the remaining targets, the OD solutions were successfully completed, and OD residuals were obtained for these targets. Positional errors are calculated by comparing the OD result with the TLE data at the observation times. In Figs. 4 and 5, the number of observation data points is distributed according to the acquired observation time.

As is evident from Fig. 5, when the observation residual is less than 720 , the OD for False match & True association data corresponding to different matching error thresholds yields the residual ranging from several arc-seconds to approximately a dozen arc-seconds. For example, the OD residuals for targets with NORAD numbers 21233, 45535, and 58377 are 8.96 , 13.57 , and 12.01 , respectively. Conversely, when the observation residual exceeds 720 , the OD residual surges to hundreds of arc-seconds. An illustrative example is the target NORAD number 60264 in Fig. 5D, which has an OD residual of 228.75 .

Fig. 6 [Figure 6: see original paper] illustrates the positional errors of the OD results compared with the TLE data. The RMS errors corresponding to the positional errors of the targets with NORAD numbers 21233, 45535, 58377, and 60264 are calculated to be 1.82, 1.63, 2.26, and 3.04 km, respectively. Through the consistency of the distribution of the orbital residuals and the position comparison of the OD results with the TLE data error variation range, it is evident that the OD results of the first four targets are consistent with the accuracy of the TLE data. However, for the target with NORAD number 60264, notable disparity exists in the orbital residuals. This suggests that this specific arc-segment is likely to reflect observation data pertaining to other targets. Given that the magnitude of the orbital residual value of this arc-segment is similar to that of its observational residual, it is highly probable that the data of this arc-segment do not belong to the target with NORAD number 60264.

Through OD verification, when the matching error threshold is set at 60 , 180 , and 360 , none of the False match & True association results obtained under these settings can thus accurately indicate that the observation data have been correctly matched and correlated. When the matching error threshold is set to 720 , a certain degree of fuzzy matching correlation exists in the observation data. Hence, it is reasonable to set 720 as the data matching error threshold. Specifically, using the arc-segment difference method, the data arcs with an average matching rate of 97.62% within an 8-day period were processed, and an overall data correct matching rate of approximately 92.34% was achieved. Additionally, the remaining 5.28% of data with fuzzy correlations was classified and processed by allocating the 3,646 data arcs to different false match and association results.

3.4. Application of False Match and Association Data

Combined with the matching results for the same target over multiple days, if the matching residuals for a particular day suddenly increase, it is likely that the target has undergone an orbital maneuver. In the case of True match & False association results, because the existing data arcs are correctly matched to the target information in the catalog database, and because data processing on the same day adopts the set of orbital elements closest to the time of observation, the matching residuals will not suddenly increase in the adjacent orbital period. The occurrence of the target' s orbital maneuver cannot be detected in this scenario. Therefore, most of the True match & False association data should be data relating to the newly detected target. The results in Table 3 list 57 targets with 98 passes of data. In the False match & True association data, it is easier to find the target of an orbital maneuver.

Taking the target with NORAD number 60339 as an example, the results of matching and association for the observation data are listed in Table 5 .

By processing the historical TLE data of the target, the orbital altitude information can be derived. When this information is integrated with the measured data, it is revealed that the target with NORAD number 60339 executed an orbital maneuver on 22 September (Modified Julian Day = 60,575.8020). Specifically, within a span of 2 days, the orbital altitude of the target increased by approximately 28 km, as illustrated in Fig. 7 [Figure 7: see original paper].

In Fig. 7, the TLE data release epoch is 60,577.3333 in Modified Julian Days with a corresponding orbital altitude of 465.520 km. The actual observation moment is 60,577.4441 in Modified Julian Days. When the target makes an orbital maneuver, the altitude of the target orbit is no longer 465.520 km; instead, it should be within the range of 465.520–469.130 km. Because the TLE data evolves unstably during the orbital maneuver, a large matching error value is produced. However, once the target' s orbit stabilizes, the matching errors of the subsequently measured data all fall within the normal matching error threshold for LEO targets.

The same can be found for the orbital maneuver in the False match & False association case. Using the target with NORAD number 60439 as an example, the matching associations for this target are listed in Table 6 .

It is evident that the target started to make an orbital maneuver on 18 September (Modified Julian Day = 60,572.2723). The orbit altitude was increased by approximately 55 km throughout the orbital change, as shown in Fig. 8 [Figure 8: see original paper].

On 22 September (Modified Julian Day = 60,575.8901), a large matching error is evident when the target with NORAD number 60439 underwent an orbital change. However, once the orbit stabilized, the matching error magnitude aligned with the precision of the TLE data.

---

4. ANALYSIS AND DISCUSSION

In the process of matching and associating space target data with a known catalog database, multiple factors can give rise to fuzzy matching and association. These factors include the orbital accuracy of the cataloged targets, orbital maneuvers executed by the targets, and the detection of new targets. The data matching and association strategy proposed in this paper, based on the arc-segment difference method, not only enables effective identification of fuzzy association data but also allows for its refinement and classification. By setting the data matching threshold and the arc-segment dynamic association threshold, this strategy showcases the specific application of such data, thereby enhancing the data utilization rate.

According to Table 3, when the matching error threshold is set at 720 , statistical analysis reveals that all the True match & False association data pertain to LEO targets. Among the False match & True association targets, 98.07% consist of LEO targets, while the remaining 1.93% are composed of 11 HEO targets and 2 MEO targets. In the case of False match & False association targets, 89.10%

are LEO targets, and the remaining 40 targets include 23 HEO targets and 17 MEO targets. It is evident that when handling data of LEO and HEO targets, fuzzy correlations are likely to occur owing to issues related to orbital accuracy. Given that the slant distance of LEO targets from the station is assumed to be 1,000 km, the line-of-sight error corresponding to 720 is approximately 3.49 km, which is in line with the accuracy of LEO targets in the TLE database.

As shown in Fig. 5D, the OD residual for the case of False match & True association is 228.75 . When this is combined with the positional error of 3.04 km depicted in Fig. 6D, it indicates that the target' s OD scheme is correct and that the matching result corresponds to a target within the known catalog database. However, the observation data include data from other targets, which results in an excessively large OD residual. After eliminating the interfering data, the OD residual is found to be 18.57 , which is of similar magnitude to that of the OD residuals of the other three targets. From Tables 5 and 6, through processing the matching results of the same target over multiple days (as illustrated in Figs. 7 and 8), the false matching correlation data have found specific applications. On this basis, it is deemed that setting the matching threshold at 720 is appropriate. However, if a satellite performs a continuous low-thrust maneuver and the TLE is the latest before that maneuver (i.e., the TLE prediction is relatively accurate), false matching association data might fail to detect the maneuver. To enhance the accuracy of determining target maneuvers, further research could be conducted, such as incorporating multi-source features (for example, the stability of the photometric curve and the mutation rate of orbital parameters). Meanwhile, experiments with MEO and HEO observation data should be increased to verify the universality of the proposed method.

---

5. CONCLUSIONS

This paper introduces a data matching and arc-segment association strategy based on the arc-segment difference method, designed to address challenges in fuzzy matching and arc-segment association during the processing of observation data. By optimizing the matching error threshold and the dynamic arc-segment association threshold, the proposed approach substantially enhances the accuracy of data matching and mitigates the fuzzy association issues caused by orbital maneuvers or propagation errors.

Experimental results demonstrated that setting the matching error threshold to 720 enables the arc-segment difference method to effectively handle fuzzy matching correlation data, thereby improving data utilization while maintaining a high matching rate. Furthermore, by classifying and processing fuzzy associated data, the method not only ensures accurate association of data from the same target but also facilitates the identification of potential orbital maneuver targets. For new target validation, the method leverages observation arc-segments from the same target to extend the arc length and enhance the

OD accuracy. In the case of maneuvering targets, the measured data provide robust evidence of orbital changes, underscoring the method' s effectiveness in processing fuzzy correlation data.

The paper also validates the appropriateness of the selected matching error threshold through OD analysis, demonstrating that the recommended threshold effectively screens fuzzy correlation data. This approach improves the efficiency of the space target catalog and enhances the precision of target orbits in the database.

Future research could further explore the practical applications and value of fuzzy associated data, building on the foundation established in this study.

ACKNOWLEDGEMENTS

This research was supported by National Natural Science Foundation of China (12273080).

AI DISCLOSURE STATEMENT

Deepseek was employed for language and grammar checks within the article. The authors carefully reviewed, edited, and revised the Deepseek-generated texts to their own preferences, assuming ultimate responsibility for the content of the publication.

AUTHOR CONTRIBUTIONS

Jiannan Sun conceived the ideas and wrote the manuscript. Jiannan Sun and Cunbo Fan developed the methodology. Jiannan Sun, Zhe Kang, and Zhenwei Li verified the research data. Zhe Kang managed and maintained the research data and provided study resources. Zhenwei Li provided the funding support. Cunbo Fan supervised and revised the paper. All authors read and approved the final manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

Baccelli, F., Candel, S., Perrin, G., et al. 2024. Large satellite constellations: Challenges and impact. *Académie des sciences*, hal−04626677.

Space-Track.org. Public catalog object count. Available from https://www.space-track.org/#/spaceOpsTempo. [Accessed on 2024−12−24].

Wang, X. 2015. Optimization of strategy of space object of survey with electro-optical telescopes. *Journal of Spacecraft TT&C Technology*, 34(2): 201−206. (in Chinese)

Wu, L. D. 2011. *Orbit and detection of satellite and space debris*. Beijing: China Science and Technology Press. (in Chinese)

Ding, et al. 2019. Observational data analysis of space objects with space debris photo-electric telescope array. *Science Technology and Engineering*, 19(18): 25−31. (in Chinese)

Lei, X. X., Sang, J. Z., Li, Z. W. 2019. Preliminary analysis of observations obtained from ground-based electro-optical sensor array at Changchun. *Journal of Geomatics*, 44(1): 41−44. (in Chinese)

Vallado, D., Crawford, P. 2008. SGP4 orbit determination. In *Proceeding of AIAA/AAS Astrodynamics Specialist Conference and Exhibit*.

Flohrer, T., Krag, H., Klinkrad, H. 2008. Assessment and categorization of TLE orbit errors for the US SSN catalogue. In *Proceedings of the Advanced Maui Optical and Space Surveillance Technologies Conference*.

Riesing, K. 2015. Two line element sets of cubesats in LEO: Accuracy assessment and estimation techniques for improvement. In *Proceeding of 29th Annual AIAA/USU Conference on Small Satellites*.

Xu, X. L., Xiong, Y. Q. 2019. Study on orbit prediction error of space objects based on historical TLE. *Acta Astronomica sinica*, 60(4): 28−40. (in Chinese)

Racelis, D., Joerger, M. 2018. High-integrity TLE error models for MEO and GEO satellites. In *Proceeding of AIAA SPACE and Astronautics Forum and Exposition*.

Wei, D., Zhao, C. Y. 2010. An accuracy analysis of the SGP4/SDP4 model. *Chinese Astronomy and Astrophysics*, 34(1): 69−76.

Früh, C., Schildknecht, T. 2012. Accuracy of two-line-element data for geostationary and high-eccentricity orbits. *Journal of Guidance, Control, and Dynamics*, 35(5): 1483−1491.

Wang, R. L., Liu, W., Liu, S. Q., et al. 2014. An orbital anomaly analysis method based on TLE data. *Chinese Journal of Space Science*, 34(2): 208−213. (in Chinese)

Song, L. P., Chen, D. F., Tian, T., et al. 2023. A real-time correlation algorithm for GEO targets based on radar ranging and velocity measurement. *Journal of Beijing University of Aeronautics and Astronautics*, 49(8): 2167−2175. (in Chinese)

Tao, Y., Hu, W. D., Xu, J. 2006. A method of space object identification based on orbit matching and improving. *Radar Science and Technology*, 4(3): 134−139. (in Chinese)

Gong, X., Le, Z. C., Wang, H., et al. 2020. Survey of data association technology in multi-target tracking. *Computer Science*, 47(10): 136−144. (in Chinese)

Vallado, D. 2013. *Fundamentals of Astrodynamics and Applications*, 4th edition. Hawthrone: Microcosm Press.

Milani, A., Gronchi, G. F., Vitturi, M. D. M., et al. 2004. Orbit determination with very short arcs. I admissible regions. *Celestial Mechanics and Dynamical Astronomy*, 90: 57—85.

Li, J. S. 1995. *Satellite Precision Orbit Determination*, 1st edition. Beijing: People' s Liberation Army Press. (in Chinese)

Chen, L., Han, L., Bai, X. Z., et al. 2010. *Orbital dynamics and error analysis of space object*. Beijing: National Defense Industry Press. (in Chinese)

CPF ephemeris download. Available from ftp://edc.dgfi.tum.de/ [Accessed 2024—12—22].

Zhang, H. F., Zhang, Z. P. 2006. The new SLR ephemeris and its application to predicting satellites. *Annals of Shanghai Observatory Academia Sinica*, 27: 8—16. (in Chinese)

Montenbruck, O., Gill, E., Lutze, F. H. 2012. *Satellite orbits: Models, methods, and applications*. Beijing: National Defense Industry Press. (in Chinese)

Note: Figure translations are in progress. See original paper for figures.