A combination method using evolutionary algorithms in initial orbit determination for too short arc

Xin-Ran Li [a,b,c,*], Xin Wang [a,c], Yong-qing Xiong [a,c]

[a] *Purple Mountain Observatory, Chinese Academy of Sciences, Nanjing 210008, China*
[b] *University of Science and Technology of China, Hefei 230026, China*
[c] *Key Laboratory for Space Object and Debris Observation, Purple Mountain Observatory, Chinese Academy of Sciences, Nanjing 210008, China*

Abstract

With the combination of two evolutionary algorithms EDA and DE, a new method of initial orbit determination for satellites based on ground-based too-short-arc is established. Compared with other algorithms, the proposed method focuses on the most densely populated region in the solution space rather than the individual with best fitness value. Both the global information and local information are well fused in the search of optimum. In the method (a, e, M) are treated as variables of the optimization, and the optimization procedure is carried out as a two-stage hierarchical optimization problem which has three variables for each stage. Kernel density estimation is applied to build the probability distribution model without any assumptions of the specified distribution, accompanied by handling semi-major axis and eccentricity as a pair of dependent variables in the construction of the probability for the correlation between them in the practice. Numerical experiments with real ground-based observations show that the proposed method is applicable to too-short-arc with even 3 s, and the result of bias in several kilometers can be achieved with $5''$ error added to angular measurements.
© 2018 COSPAR. Published by Elsevier Ltd. All rights reserved.

Keywords: Too short arc; Initial orbit determination; Estimation of distribution algorithm; Differential evolution; Evolutionary algorithm

1. Introduction

With the development of space activities, the number of space objects is increasing rapidly. In order to improve the observation efficiency, sky surveys of space objects have been deployed in recent years. Unlike in the traditional tracking mode, a large number of arcs of only several ten seconds or even several seconds can be obtained in the survey mode. However, with such too-short-arc (TSA), classical methods for initial orbit determination (IOD) are no

more applicable to achieve reasonable solutions. Then how to effectively use TSAs is becoming a research hotspot.

To solve this problem, evolutionary algorithms (EAs) have become a tendency with the development of computing power. Ansalone and Curti (2013) introduced genetic algorithm (GA) with the space-based simulation data. Hinagawa et al. (2014) applied the same algorithm in the IOD of geostationary satellite, but different parameters were chosen.

Li and Wang (2016, 2017) reimplemented the GA with different parameters, objective and operators, and successfully solved the IOD with 10 s arc for near circular orbits. Li and Wang (2015) also investigated the particle swarm optimization (PSO), which has more concise calculation steps and operator selections. These work largely improves the accuracy of the results, however, the choices of meth-

* Corresponding author at: Purple Mountain Observatory, Chinese Academy of Sciences, Nanjing 210008, China.
E-mail addresses: lixr@pmo.ac.cn (X.-R. Li), wangxin@pmo.ac.cn (X. Wang), xyqxcf@pmo.ac.cn (Y.-q. Xiong).

ods, parameters and constraint conditions for specific issues need to be considered correspondingly. Different algorithms have their own advantages and disadvantages, and the case has different characteristics and requirements. For this reason, combining various algorithms has become a common means in the practice of EAs (Cho and Zhang, 2004; Peña et al., 2004; Sun et al., 2005). By mixing the strength of each method, a balance between the various aspects of problems could be got. On the other hand, the inevitable difficulty in mixing different algorithms is that more operators and parameters need to be considered carefully and the optimization result is very sensitive to the selection.

Estimation of distribution algorithm (EDA) is a new class of EA (Larrañaga and Lozano, 2002). It focuses on global statistical information about the search space from the beginning of the search. The distribution of dominant population is what EDA characterizes, rather than the superior individual. Individuals in the dominant population are treated equally. Many experiments show that the superior individual, which has the best fitness, does not stand for the optimal solution in fact, sometimes it is even far away from the optimum (Jia and Wu, 1998). On the contrary, differential evolution (DE) (Storn and Price, 1997) obtains differential information from the current population but is lack of global information. The combination of the two algorithms will create promising solutions by fusing the global information extracted by EDA and the local differential information obtained from DE (Sun et al., 2005).

Parameters of DE are much more concise than other EAs and many experiments show that the different choices only change the computing efficiency while the convergence and results will not be affected at all. Furthermore greedy selection is used in DE, which means that the dominant population will not be worse after the DE process. So inserting DE into EDA will not result in a more complicated problem.

In past researches, the optimization variables were treated independently from each other. However, it's found in the practice of IOD that the semi-major axis a and eccentricity e of the orbit correlate seriously. When semi-major axis increases in the solution, the eccentricity goes with it. For usual EAs, good genes will be broken during searching. Longer semi-major axis can be paired with smaller eccentricity or on the contrary, because both genes are from dominant individuals of the current generation.

This paper is organized as follows. In Section 2, we propose a framework of the combination of EDA and DE after a brief introduction of these two algorithms. In Section 3, details of the implementation of the algorithm are given. Section 4 describes the results of numerical experiments followed by the conclusion in Section 5.

2. EDA and DE

2.1. EDA

EDA was first proposed in 1996, and obtained a rapid development. EDA is a mixture of statistical learning and GA, which gets solutions by establishing a probability model for the individuals distribution in the solution space with statistical learning and random sampling to generate a new generation at each iteration. EDA can model the relationship between variables through a probabilistic graph model thus it is an ideal choice for the problem of multivariate correlation optimization.

EDA was aimed at the optimization for binaries in discrete space in the early time. A probability distribution model for a continuous problem is complex, so that it's difficult to design an effective estimation algorithm, and resulting in the slow development of EDA over continuous space. In continuous space, the probability model is generally based on the Gaussian distribution assumption. Since this method is not suitable for engineering problems, for some cases the distribution histograms are often used. However, the histogram distribution is not continuous or smooth, and has many limitations (Wang et al., 2012).

Like all EAs, EDA firstly randomly generates N individuals from the given interval as the initial population, $\mathbf{X} = \{\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_N}\}$, and $\mathbf{x_i}$ stands for a D-dimensional individual, $\mathbf{x_i} = \{x_{i1}, x_{i2}, \ldots, x_{iD}\}$. Every single x_{ij} is an element to be solved. Then on the basis of the fitness, P best individuals from Gth generation are selected as the dominant population. Next, a probability model can be constructed according to the dominant population, and new solutions are subsequently sampled from the constructed model as the $(G + 1)$ th generation. Above operations will be repeated until the stopping criteria is satisfied. One thing needs to be mentioned is that EDA ignores the superior individual and all individuals in the dominant population are equal, while the individual will be selected according to the fitness in GA. Then the searching process will be much more robust.

2.2. DE

DE is a powerful search technique for solving optimization problems over continuous space. Main steps of DE are mutation, crossover and selection, very likely to GA but in reverse order. The initial population $\mathbf{X}^* = \{\mathbf{x_1^*}, \mathbf{x_2^*}, \ldots, \mathbf{x_M^*}\}$ is firstly mutated separately, then the crossover between the initial and mutated populations is taken. Finally the offspring, which has better fitness after crossover, will replace the original one in the new generation. In this schema, neither the population nor individuals will degenerate.

2.3. The combination of EDA and DE

The way of combining EDA and DE is to insert DE into EDA just after obtaining the dominant population. DE process is applied to the dominant population, and the probability model is constructed with the new dominant population. This proposed algorithm is named as EDA/DE and Fig. 1 gives the flowchart:

3. Methodology

3.1. Choice of the variables

High dimensions often bring about difficulties in optimization problems. Ansalone and Curti (2013) used two dimensions by choosing slant distances of the beginning and the end of observation time. Then the optimized results have to be combined with the measurements to achieve the six orbit elements. To prevent that, Li and Wang (2016, 2017) chose (a, e, M) of Kepler orbit elements as variables. With this choice, the orbit elements can be directly obtained after the optimization and are independent of the measurements. However, this choice is only suitable for near circular orbits. For a reasonable earth satellite orbit, semi-major axis a and eccentricity e are correlated and not independent of each other. So when a and e are

Fig. 1. The flowchart of EDA/DE.

taken from reasonable interval separately, $a(1 - e)$ may be less than a_e (a_e is the equatorial radius) which brings about impossible results. For near circular orbits, this problem may be ignored in practice because of the narrow range of eccentricity. But for the IOD problem without any prior information, with the increase of the range of eccentricity to be searched, this situation becomes more serious and must be considered. In this study, new variables $\mathbf{x} = \{x_1, x_2, x_3\}$ are defined in a transformed form:

$$x_1 = a(1 - e), \quad x_2 = ae, \quad x_3 = M_0.$$

x_1 stands for altitude of the perigee. While lower bound of x_1 is determined, any other value in the interval is reasonable.

3.2. Fitness function

For a series of measurements $\{t_i, \alpha_i, \delta_i, i = 1, 2, \ldots, n\}, (\alpha_i, \delta_i)$ are the right ascension and declination at time of t_i. With (a, e, M_0) at t_0, the values of M_i, E_i and f_i at t_i can be got. We can obtain the following relationships (Wu, 2011):

$$r_i = a(1 - e\cos E_i),$$

$$\rho_i = \sqrt{r_i^2 - R_i^2 \sin^2 Z_i} - R_i \cos Z_i,$$

$$\mathbf{r}_i = \rho_i \mathbf{L}_i + \mathbf{R}_i, \tag{1}$$

$$\mathbf{L}_i = (\cos \delta_i \cos \alpha_i, \cos \delta_i \sin \alpha_i, \sin \delta_i)^{\mathrm{T}}.$$

where $\mathbf{r}_i$ is the geocentric position vector of the target, $r_i = |\mathbf{r}_i|$. ρ_i stands for slant distance. And $\mathbf{R}_i$ represents the geocentric position vector of the observation station, $R_i = |\mathbf{R}_i|$. Z_i is the zenith at t_0. For every couple of $(t_k, t_j), t_k > t_j, (\mathbf{r}_k, \mathbf{r}_j)$ and the corresponding (f_k, f_j) can be calculated. The definition of fitness function is as follows:

$$f(X_i) = f((a, e, M_0)_i)$$
$$= \left(\frac{1}{C_n^2} \sum_{j=1}^{n-1} \sum_{k=j+1}^{n} \left(f_k - f_j - \cos^{-1}\left(\frac{\mathbf{r}_k \cdot \mathbf{r}_j}{r_k r_j} \right) \right)^2 \right)^{\frac{1}{2}}.$$

Obviously the smaller fitness value is better.

3.3. Generation of the initial population

Choose a pair of (t_k, t_j). From Eq. (1)

$$\Delta f_{kj} = \cos^{-1}\left(\frac{\mathbf{r}_k \cdot \mathbf{r}_j}{r_k r_j} \right),$$

define $r_j = r_k = r$. Regarding

$$\left| \sqrt{\frac{\mu}{r^3}} - \frac{\Delta f_{kj}}{t_k - t_j} \right|$$

as the objective, r can be optimized (Wu, 2011). Obviously x_1 must be less than r. According to the altitude of space objects, $[1.03a_e, r]$ is sufficient for the interval of x_1, and x_2 ranges in $[0, 4a_e]$, $x_3 \in [0, 2\pi]$. Then the initial population is sampled uniformly from the above interval separately.
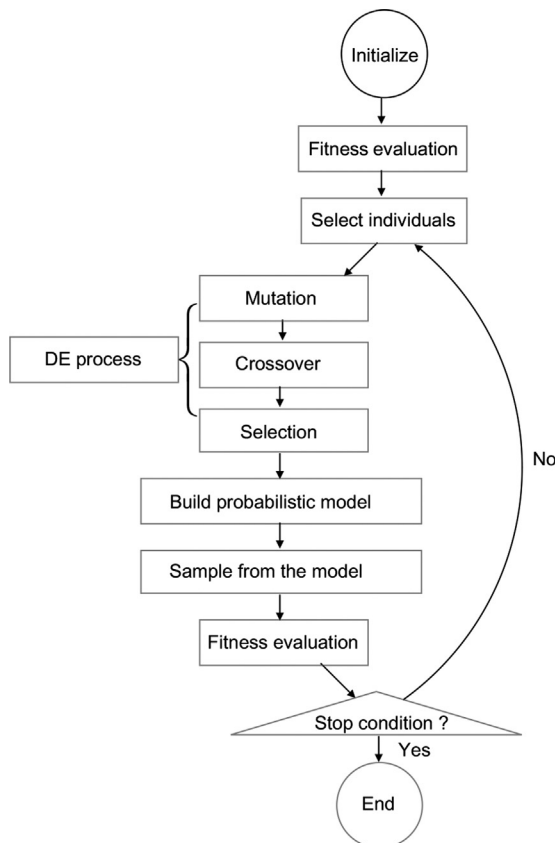
3.4. DE process for the dominant population

Truncation selection is used to pick up the P best individuals from Gth generation as the dominant population $\mathbf{X}_E^G$, and $\mathbf{X}_E^G$ is regarded as the initial population $\mathbf{X}$ in DE. The mutation is carried out as follows:

$$\mathbf{V}_i = \mathbf{X}_{r_0} + F(\mathbf{X}_{r_1} - \mathbf{X}_{r_2}).$$

where $r_i (i = 0, 1, 2)$ are exclusive integers randomly generated within the range $[1, P]$, and $r_i \neq i$. F is a positive control parameter and usually $F \in [0.4, 1]$.

After the mutation, crossover operation is applied to each pair of $\mathbf{X}_i$ and the mutant $\mathbf{V}_i$ to generate trial vector $\mathbf{U}_i$:

$$U_{ij} = \begin{cases} V_{ij}, r < CR \quad or \quad j = rand \\ X_{ij}, others \end{cases}.$$

U_{ij}, V_{ij}, X_{ij} are the jth element of the corresponding ith vector respectively. r is a uniform random number in $[0, 1]$, and j is a randomly chosen integer in $[1, P]$. The crossover rate CR is a constant within the range $[0, 1)$. This operation ensures that $\mathbf{U}_i$ will be different from $\mathbf{X}_i$ with at least one element. Finally the greedy selection is used:

$$\mathbf{X}_i^* = \begin{cases} \mathbf{U}_i, f(\mathbf{U}_i) < f(\mathbf{X}_i) \\ \mathbf{X}_i, others \end{cases}.$$

Obviously the new generation will not be worse than before. After the DE process, the dominant population $\mathbf{X}_E^G$ evolves to a better population $\mathbf{X}_E^{G*}$.

3.5. Construction of the probability model

The probability model is constructed with the new dominant population $\mathbf{X}_E^{G*}$. (a, e) are considered as a pair of dependent variables, and M remains independent.

By using Kernel density estimation, the model is built, and based on the sample y_{ij}, the probabilistic density (Scott, 2015) at $\mathbf{y}^*$ can be calculated as:

$$D(\mathbf{y}^*) = \frac{1}{P h_1 \cdots h_d} \sum_{i=1}^{P} \left\{ \prod_{j=1}^{d} K\left(\frac{y_j^* - y_{ij}}{h_j}\right) \right\}. \tag{2}$$

In each dimension the same kernel K is used but the smoothing parameter h_i is different. For M, the equation is one-dimensional, $d = 1$, and for (a, e), it's a two-dimensional problem, $d = 2$. Epanechnikov kernel function (Epanechnikov, 1969) is chosen as:

$$K(x) = \begin{cases} \frac{3}{4}(1 - x^2), & |x| < 1 \\ 0, & |x| > 1 \end{cases}.$$

The probability model which is established with the dominant group of Gth generation is expressed as D_G^E, then the probability model of $(G+1)$th generation is

$$D_{G+1}^E = (1 - \alpha)D_{G-1}^E + \alpha D_G^E.$$

α indicates learning rate. The concept of Monte Carlo simulation is used in the paper. However, the way solutions obtained from this evolution makes sure that during the construction of the probability density function the key point is the shape of the probability distribution or the location of the peak, other than the accuracy of the probability density. And this property is retained with the learning step.

It is noteworthy that h must be changing during the convergence process of the probability density function. According to rule of thumb (Silverman, 1986, 43), h is chosen adaptively as:

$$h = \left(\frac{4}{3P}\right)^{1/5} \sigma. \tag{3}$$

where σ is the standard deviation of the samples in each dimension.

3.6. Sampling method

With EDA, the new generation is sampled from the built probabilistic model. The implement of sampling is on the basis of the density calculated from Eq. (2). For kernel density estimated model, a very efficient sampling method was given by Silverman (1986, 142–144). The process is as follows, for the dataset $\{y_1, y_2, \ldots y_P\}$,

Step 1 Randomly select an integer μ from the range $[1, P]$.
Step 2 Randomly sample ε from $K(\cdot)$.
Step 3 $z_i = y_\mu + \varepsilon h$.

After repeating, the final set of z_i will have the same distribution as y_μ.

For our problem, the adapted sampling process is:

Step 1 Randomly select 2 integers μ_1 and μ_2 from $[1, P]$.
Step 2 Randomly sample $\varepsilon_j (j = 1, 2, 3)$ from $K(\cdot)$.
Step 3 $a_i = a_{\mu_1} + \varepsilon_1 h_a, e_i = e_{\mu_1} + \varepsilon_2 h_e$.
Step 4 $M_i = M_{\mu_2} + \varepsilon_3 h_M$.
Step 5 $z_i = (a_i, e_i, M_i)$.

The new generation will be fulfilled by repeating this process.

For the Epanechnikov kernel we selected, a more convenient way for sampling is available (Silverman, 1986, 42–48). v_1, v_2, v_3 are randomly generated within the range $[0, 1)$ respectively, then ε is expressed as:

$$\varepsilon = \begin{cases} v_3, & v_3 \geqslant v_2 \ and \ v_3 \geqslant v_1 \\ v_2, & others \end{cases}.$$

3.7. Stopping criteria

In common practice, the evolution is stopped when the iterations are performed over the predefined parameter

G_{max} of maximum generation or when the number of iterations of the fitness value's stagnation exceeds C generations. These usual techniques are adopted in our algorithm.

Besides, convergence of the probability density is considered as another termination condition in our study. From Eq. (3), the convergent degree of the density can be represented by σ. When σ becomes smaller, h tends to be 0, and the population will not update anymore. Therefore, another stopping criteria is also taken into account, which is the sum of standard deviations in three dimensions $\sigma_1 + \sigma_2 + \sigma_3 < T_\sigma$.

Just like GA, the evolutionary process will have a destructive effect on the parent generation, and the best individual of the parent generation is not guaranteed to appear in the child generation in EDA/DE. In GA, the elitist scheme is used to keep the best individual in the next generation. While in EDA/DE, the solution with best fitness value is recorded as $\mathbf{x}_{best}$. Considering the final status of σ at the end of evolution, when σ goes to 0, all the individuals in the population converged to a single point, and this solution is recorded as $\mathbf{x}_{prob}$. Both $\mathbf{x}_{best}$ and $\mathbf{x}_{prob}$ will be given in the later experiments.

3.8. Solution of (i, ω, Ω)

From Eq. (1), $\mathbf{r}_i$ can be deduced with the solution of (a, e, M_0), then the full set of six orbit elements can be obtained without any difficulties. In consideration of the accuracy requirement, an easy method is employed. Calculating a set of (i, ω, Ω) from every couple of $(\mathbf{r}_i, \mathbf{r}_j)$, then the mean or median value of these results can be obtained as the final solution.

4. Numerical experiment

The algorithm is implemented in MATLAB. In two-body model, real measurements from the optical observation network of Chinese Academy of Sciences are applied to verify the performance of the proposed algorithm. The accuracy of measurement is $5''$ and the sampling rate is 1 Hz. First 10 points are taken for the experiment. The station coordinates in terrestrial reference frame are $(-0.20a_e, 0.88a_e, 0.42a_e)$, and the measurements in topocentric equatorial coordinate system (J2000.0) $\{t_i, \alpha_i, \delta_i\}$ are listed in Table 1.

Based on real measurements, as calculated in Section 3.3, the intervals of (x_1, x_2, x_3) are $x_1 \in [1.03a_e, 1.15a_e]$, $x_2 \in [0, 4.0a_e], x_3 \in [0, 2\pi]$. No more information is used in the determination of the intervals. In EDA/DE, value of parameters are as follows: for EDA part, $N = 30, P = 9$, $\alpha = 0.1, T_\sigma = 10^{-6}, G_{max} = 200, C = 140$, and for DE part, $F = 1.0, CR = 0.9$. Compared to Monte Carlo simulation, the number of initial population is only 30 individuals, which is far less than what the Monte Carlo method needs. In fact, Monte Carlo simulation builds the distribution in the whole solution space, while EDA/DE focuses on the

Table 1
Measurements of 10 s arc.

year	mon	day	h	min	s	$\alpha/(°)$	$\delta_i/(°)$
2006	2	2	22	4	29.108499	317.13694	58.49152
2006	2	2	22	4	30.100500	316.93530	58.47358
2006	2	2	22	4	31.101499	316.73286	58.45491
2006	2	2	22	4	32.105500	316.52872	58.43533
2006	2	2	22	4	33.103499	316.32486	58.41669
2006	2	2	22	4	34.100500	316.12155	58.39658
2006	2	2	22	4	36.098499	315.71230	58.35538
2006	2	2	22	4	37.098499	315.50594	58.33330
2006	2	2	22	4	38.101499	315.29733	58.31063
2006	2	2	22	4	39.101500	315.09116	58.28825

distribution of sub-space of the dominant populations, which is the most intensive place in the solution space or position of the peak.

Fig. 2 shows the convergence process of the semi-major axis a and the fitness in a whole evolution. It clearly shows that the efficiency is pretty good, and since the 40th generation the solution is very close to the final optimum. In IOD, the precision of the track surface is slightly better than the precision of the track, and it is the key to get a more accurate a in the calculation of IOD. Then values of (a, e) are the most important part (Wu, 2011), and it's easy to turn (x_1, x_2) to (a, e). The probability distributions of (a, e) at generation 0 and 195 with different coordinate scales are given in Fig. 3. Points in Fig. 3 represent the sampled individuals location of the current generation, and the contour lines show the probability density distribution of (a, e) obtained from individuals distribution. The density increases as the color becomes shallower. Clearly the distribution eventually gathers around the correct solution, which means a small number of population are enough to find the sub-space of the dominant population, and throughout the process the property that regarding the location of the peak as the focus of search is retained. Besides, as shown in Fig. 3, randomly generated populations have no certain distribution laws, and the distribution is spread in a wide range at the beginning. Yet it doesn't
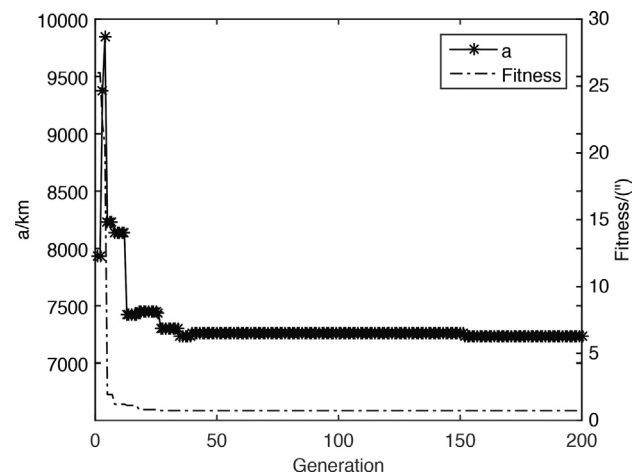

Fig. 2. The convergence progress of the semi-major axis a and the fitness.
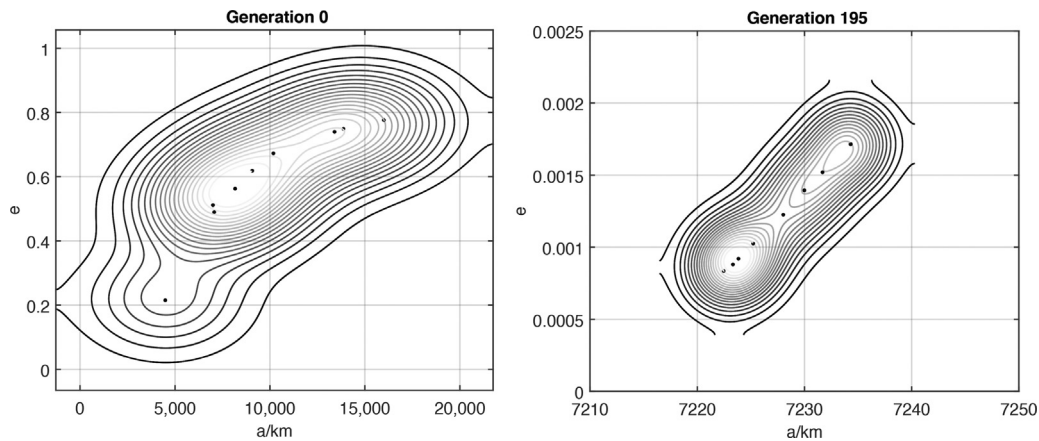
Fig. 3. Probability densities of the initial and the 195th generation.

Table 2
The results of IOD with different random seeds.

	a/km	e	$i/(°)$	$\Omega/(°)$	$(\omega + M)/(°)$	Fitness/($''$)
$\mathbf{x}_{best}$	7241.90203	0.00420	98.58661	31.59582	138.41453	0.72564
$\mathbf{x}_{prob}$	7241.91894	0.00420	98.58657	31.59589	138.41436	0.72564
Δ	0.01691	0.00000	0.00004	0.00007	0.00017	0.00000
$\mathbf{x}_{best}$	7227.81559	0.00302	98.60158	31.56393	138.48108	0.72561
$\mathbf{x}_{prob}$	7214.70170	0.00046	98.69119	31.42707	138.83872	0.72825
Δ	13.11389	0.00256	0.08961	0.13686	0.35764	0.00264
$\mathbf{x}_{best}$	7222.61632	0.00324	98.58730	31.58038	138.42886	0.72637
$\mathbf{x}_{prob}$	7241.00546	0.00649	98.47779	31.75113	137.98890	0.74570
Δ	18.38914	0.00325	0.10951	0.17075	0.43996	0.01933
$\mathbf{x}_{best}$	7237.86677	0.00395	98.58372	31.59679	138.40562	0.72552
$\mathbf{x}_{prob}$	7235.26952	0.00345	98.60215	31.56868	138.47937	0.72564
Δ	2.59725	0.00050	0.01843	0.02811	0.07375	0.00012
$\mathbf{x}_{best}$	7228.25230	0.00283	98.61123	31.55049	138.51884	0.72562
$\mathbf{x}_{prob}$	7232.81442	0.00267	98.64448	31.50697	138.64612	0.72658
Δ	4.56212	0.00016	0.03325	0.04352	0.12728	0.00096
$\mathbf{x}_{best}$	7254.92853	0.00553	98.57100	31.62813	138.34546	0.72606
$\mathbf{x}_{prob}$	7238.13588	0.00324	98.64051	31.51659	138.62687	0.72662
Δ	16.79265	0.00229	0.06951	0.11154	0.28141	0.00056
$\mathbf{x}_{best}$	7219.52052	0.00228	98.62142	31.52948	138.56277	0.72580
$\mathbf{x}_{prob}$	7215.67336	0.00068	98.68122	31.44223	138.79980	0.72837
Δ	3.84716	0.00160	0.05980	0.08725	0.23703	0.00257
$\mathbf{x}_{best}$	7237.62763	0.00391	98.58545	31.59416	138.41256	0.72552
$\mathbf{x}_{prob}$	7239.78874	0.00407	98.58496	31.59652	138.40936	0.72557
Δ	2.16111	0.00016	0.00049	0.00236	0.00320	0.00005
$\mathbf{x}_{best}$	7241.40252	0.00416	98.58658	31.59547	138.41469	0.72562
$\mathbf{x}_{prob}$	7241.46929	0.00417	98.58642	31.59575	138.41407	0.72562
Δ	0.06677	0.00001	0.00016	0.00028	0.00062	0.00000
$\mathbf{x}_{best}$	7211.12643	0.00001	98.70248	31.40843	138.88185	0.72883
$\mathbf{x}_{prob}$	7211.14623	0.00001	98.70249	31.40843	138.88197	0.72883
Δ	0.01980	0.00000	0.00001	0.00000	0.00012	0.00000

affect the evolution of the latter. At generation 195, the region becomes concentrated and finds the solution space.

Table 2 lists 20 results from 10 runs with different random seeds, where both $\mathbf{x}_{best}$ and $\mathbf{x}_{prob}$ are given, and the Δ represents difference between each element of $\mathbf{x}_{best}$ and $\mathbf{x}_{prob}$. Statistical result is summarized in Table 3, where POD (precise orbit determination) is the reference solution by the precise orbit determination with multi-day observations.

Table 3
Statistical result of IOD.

	a/km	e	$i/(°)$	$\Omega/(°)$	$(\omega + M)/(°)$
POD	7229.64165	0.00173	98.63644	31.51627	138.72688
Med	7236.44857	0.00335	98.59444	31.57453	138.45412
Mean	7231.74911	0.00313	98.61176	31.55262	138.51829
Med(best)	7232.93997	0.00357	98.58695	31.58727	138.42177
Mean(best)	7232.30586	0.00331	98.60374	31.56431	138.48663
Med(prob)	7236.70270	0.00335	98.62133	31.54263	138.55312
Mean(prob)	7231.19235	0.00294	98.61978	31.54093	138.54995
σ	12.28508	0.00173	0.05367	0.08358	0.21476
σ(best)	12.85984	0.00147	0.03769	0.06141	0.15229
σ(prob)	12.35118	0.00203	0.06721	0.10332	0.26824

Although results in Table 2 are different, even for the same pair of $\mathbf{x}_{best}$ and $\mathbf{x}_{prob}$, the difference in fitness is very small and far beyond the accuracy of measurements. From the view of optimization, these values are equal. Thereupon all the results are valid. Moreover, the median and mean of each set of data are pretty close, and have smaller bias and deviation compared to previous studies. This proofs that the proposed algorithm is effective, and in practice median
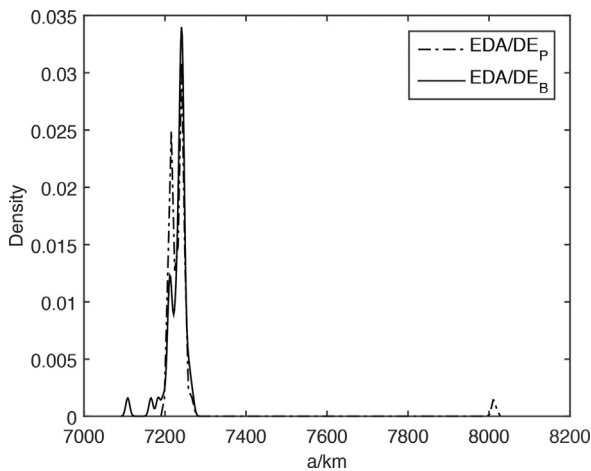

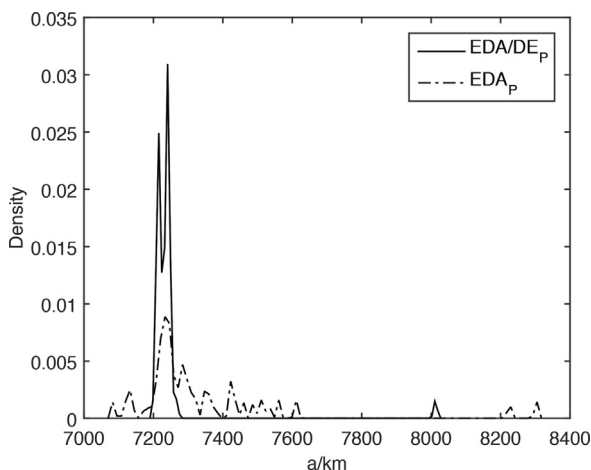Fig. 4. Distributions of semi-major axis a_{best} and a_{prob}.

Fig. 5. Distributions of semi-major axis a by EDA/DE and EDA.

or mean value of several runs can be recommended as the final result.

Fig. 4 shows the distribution of semi-major axis a after 50 runs with the same dataset. Solid line stands for the distribution for a_{best}, while dashed line is for a_{prob}. The results of two groups are very close and a_{prob} spreads in a little bit narrower range.

Fig. 5 compares the distribution of a_{prob} obtained with and without DE process in the EDA framework. It clearly shows that with DE process the peak is much more concentrated. In other words, the result has much higher confidence level.

Moreover, an attempt for shorter arcs with higher frequency is carried out with the proposed algorithm. A real observation of 10 points in 3 s is used in the experiment. The corresponding station coordinates in terrestrial reference frame are $(-0.47a_e, 0.49a_e, 0.73a_e)$, and Table 4 lists the measurements $\{t_i, \alpha_i, \delta_i\}$ in topocentric equatorial coordinate system (J2000.0).

Ansalone and Curti (2013) mentioned that the $2'' - 3''$ error in measurements strongly impacts the result of IOD for TSAs. The shorter the arc is, the more serious the influence is. Parametric bootstrap method (Efron and Tibshirani, 1993), which is used by Li and Wang (2016, 2017) in previous study, is taken to check the influence. $5''$ error is added to the original data. Table 5 summarizes the results of 50 runs.

With the comparison between mean and median of $\mathbf{x}_{best}$ and $\mathbf{x}_{prob}$, the solution of $\mathbf{x}_{prob}$ is actually better than $\mathbf{x}_{best}$. That means the solution where probability density converges is more appropriate to be the result rather than the solution with the smallest fitness. Fig. 6 gives the distribution of a_{prob} in dashed line and a_{best} in solid line. Though

Table 4
Measurements of 3 s arc.

year	mon	day	h	min	s	$\alpha/(°)$	$\delta_i/(°)$
2012	07	15	12	09	01.889783	287.45947	−20.51230
2012	07	15	12	09	02.244244	287.46754	−20.46708
2012	07	15	12	09	02.589936	287.47581	−20.42063
2012	07	15	12	09	02.805785	287.48059	−20.39216
2012	07	15	12	09	03.022276	287.48590	−20.36291
2012	07	15	12	09	03.210433	287.49015	−20.33852
2012	07	15	12	09	03.529566	287.49789	−20.29588
2012	07	15	12	09	03.848292	287.50534	−20.25316
2012	07	15	12	09	04.155389	287.51282	−20.21258
2012	07	15	12	09	04.494840	287.52063	−20.16691

Table 5
The statistical result of IOD with 3 s arc length.

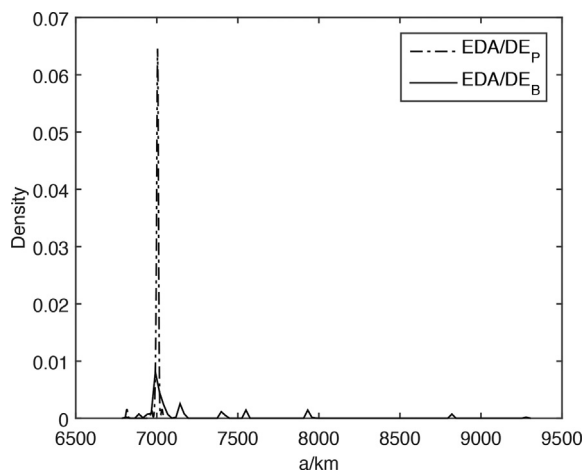| Method	a/km	e	$i/(°)$	$\Omega/(°)$	$(\omega + M)/(°)$
POD	7011.48506	0.00371	97.8273	265.2511	36.32832
Med(best)	7005.45024	0.00744	98.17428	265.4725	36.63901
Mean(best)	7153.67456	0.00322	98.73802	265.7831	38.53654
Med(prob)	7005.33650	0.00579	97.96349	265.3824	35.96765
Mean(prob)	6999.96399	0.00216	98.03313	265.4130	36.18842

Fig. 6. Distributions of semi-major axis a with 3 s arc.

the result spreads over a much larger region compared with Fig. 4, it is not surprising that the locations of the peaks of both groups are quite close and the peak of a_{prob} is much more higher this time. Although high accuracy observation is very beneficial to the IOD of TSAs, with proposed algorithm, the impact is reduced significantly. Even when the arc is shortened from 10 s to 3 s, with EDA/DE, the result is still accuracy, and the peak is concentrated obviously.

5. Conclusion

In this work, the combination of EDA and DE is successfully applied to the IOD of TSAs. Compared to GA, PSO and other EAs, the use of EDA eliminates the constraint of eccentricity, and the DE process concentrates the result. And during the building of probabilistic model, (a, e) are considered correlated, which is more accordance with the practice.

Differing from the previous studies, the proposed algorithm abandons the traditional view of optimum which has the best fitness, and focuses on the most intensive place in the solution space. The searching does not try to find the best individual any more, instead, it just finds the region of the solution and shrinks it until the region is densely populated enough.

With these improvements, the accuracy and the confidence level are increased greatly. As the numerical experiments show that the peak of EDA/DE is much more concentrated than EDA, and compared with $\mathbf{x}_{best}, \mathbf{x}_{prob}$

spreads in a little bit narrower range, and for even shorter arc of 3 s, the result is with a high accuracy.

References

Ansalone, L., Curti, F., 2013. A genetic algorithm for Initial Orbit Determination from a too short arc optical observation. Adv. Space Res. 52, 477–489.

Cho, D.Y., Zhang, B.T., 2004. Evolutionary continuous optimization by distribution estimation with variational bayesian independent component analyzers mixture model. Lect. Notes Comput. Sci. 3242, 212–221.

Efron, B., Tibshirani, R., 1993. An Introduction to the Bootstrap, first ed. Chapman & Hall, New York.

Epanechnikov, V.A., 1969. Nonparametric estimation of a multidimensional probability density. Theory Probability Appl. 14, 156–162.

Hinagawa, H., Yamaoka, H., Hanada, T., 2014. Orbit determination by genetic algorithm and application to GEO observation. Adv. Space Res. 53, 532–542.

Jia, P.Z., Wu, L.D., 1998. On the optimal accuracy and double solutions of initial orbit determination. Acta Astron. Sin. 39, 337–343.

Larrañaga, P., Lozano, J.A., 2002. Estimation of Distribution Algorithms: A New Tool for Evolutionary Computation, first ed. Springer, US.

Li, X.R., Wang, X., 2015. Particle swarm optimization for initial Orbit Determination with too Short Arc. J. Spacecraft TT & C Technol. 34, 545–551.

Li, X.R., Wang, X., 2016. Genetic algorithm for initial Orbit Determination with too Short Arc. Acta Astron. Sin. 57, 66–77.

Li, X.R., Wang, X., 2017. Genetic algorithm for Initial Orbit Determination with too Short Arc. Chin. Astron. Astrophys. 41, 254–262.

Peña, J.M., Robles, V., Larrañaga, P., Herves, V., Rosales, F., Pérez, M. S., 2004. GA-EDA: hybrid evolutionary algorithm using genetic and estimation of distribution algorithms. Lect. Notes Comput. Sci. 3029, 361–371.

Scott, D.W., 2015. Multivariate Density Estimation: Theory, Practice, and Visualization, first ed. John Wiley & Sons, New Jersey.

Silverman, B.W., 1986. Density Estimation for Statistics and Dara Analysis, first ed. Chapman and Hall, London.

Storn, R., Price, K.V., 1997. Differential evolution-a simple and efficient heuristic for global optimization over continuous spaces. J. Global Optim. 11, 341–359.

Sun, J.Y., Zhang, Q.F., Tsang, E.P.K., 2005. DE/EDA: a new evolutionary algorithm for global optimization. Inf. Sci. 169, 249–262.

Wang, S.Y., Wang, L., Fang, C., Xu, Y., 2012. Advances in estimation of distribution algorithms. Control Decision 27, 961–966.

Wu, D.L., 2011. The Orbit and Observation of Artificial Satellites and Space Debris, first ed. China Science and Technology Press, Beijing.