TELKOMNIKA, Vol. 11, No. 7, July 2013, pp. 3780 ~ 3788
e-ISSN: 2087-278X

Simulation of Radar Track Based on Data Mining Techniques

WU Xiao-chao[1, 2], Wang Lian-dong[1,2], YAN Liao-liao [2], XUE Fang-xia [2]
[1] The State Key Laboratory of Complex Electromagnetic Environment Effects on Electronics and Information System, Luoyang, 471003, China
[2] Electronic Equipment Test Centre of China, Luoyang, 471003, China
*Corresponding author, e-mail: wuxiaochao@126.com

Abstract

In order to make radar target track simulation more realistic in simulation test-bed of C3I system, a new method which utilizes data mining techniques based on historical radar data is proposed. This method includes two stages: The first is "filtering separation - piecewise fitting - feature clustering". In this stage, the radar historical data is divided into the actual true track and noise. Through computing the second-order discrete curvature, the actual true track is decomposed into several segments, such as straight line and arc, which are fitted with multinomial subsequently. On this basis, after analyzing the characteristic vector of radar historical data, the clustering database can be established; the second is "feature association-track recombination". The track in pre-deigned air scenario is segmented by the second-order discrete curvature. After the correlative feature information of the segmented scenario is searched, matched and associated with the information in clustering database, a new track will be restructured by using this output results. This method is very available for its effective and successful application in simulation test-bed of C^3I system.

Keywords: *radar track simulation, data mining techniques, clustering database, simulation test-bed of C^3I system*

1. Introduction

In a simulation test-bed of C^3I system, the radar intelligence simulation is a very important, complex and systematic content by using "black box" modeling and simulation in most cases, and there are two methods: one is to generate the white noise according to the radar measurement accuracy provided by the manufacturer, and then add the white noise to simulation trajectory the radar intelligence simulation. This method is very simple but works badly. The other is to acquire the radar's system error and standard deviation according to the statistical analysis of historical data, and then synthesize radar intelligence by using white Gaussian noise model and simulation trajectory. This method only reflects the statistical regularities of the overall error but fails to show the actual error features, and also debases the fidelity of radar simulation [1]. Since its authenticity, the historical radar data can be used to generate simulation radar intelligence data, the fidelity of which is improved certainly. This paper researches a simulation method to synthesize new radar tracking according to the requirement of C^3I simulation with data mining techniques, which can analyze and cluster the feature vector of the historical radar data. The simulation fidelity can be effectively enhanced by using this method.

2. Process of Radar Intelligence Simulation

Holding a large number of historical radar data is a precondition to generate radar intelligence data. And how to extract the feature vector from historical radar data to flexibly build an air scenario to satisfy the designed requirement is an essential step. This radar intelligence simulation method is divided into two stages, the first stage is through "filtering separation - piecewise fitting - feature clustering" several sectors to establish the cluster database; the second is through "feature association - track recombination" two sectors to recombine a new track according to the requirement of the designed air scenario, as shown in Figure 1.

Received January 27, 2013; Revised April 17, 2013; Accepted April 28, 2013

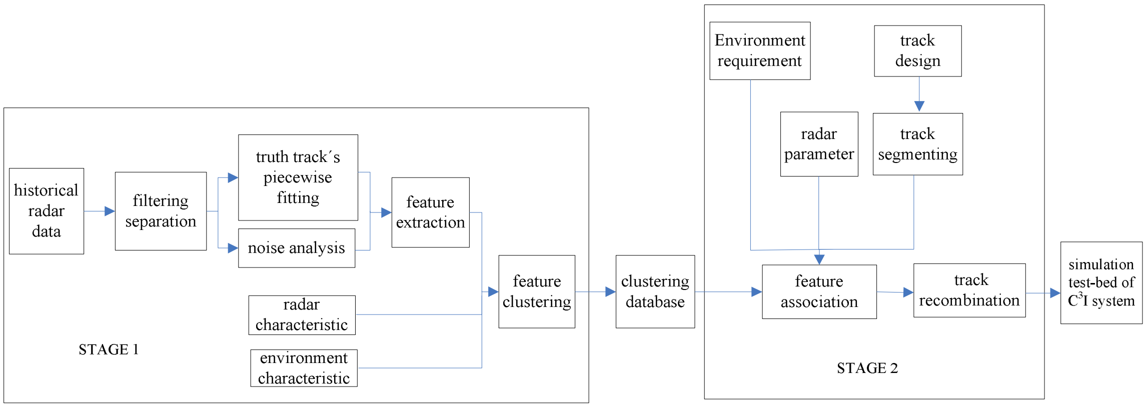

Figure 1. The process of radar track simulation based on data mining techniques

Stage 1. In this stage, with the data mining techniques, a large number of historical radar intelligence data will be sorted by the same kinds of sample group of feature vector to construct clustering database of track feature vector. The main steps are as follows:

a) By using wavelet decomposition and filtering in the frequency domain, the track data in relative coordinate system will be separated into two parts: noise signal and actual true track.

b)This step has two sectors. The first is to divide the actual true track into several segments, such as straight lines and arcs, using the curvature method to improving the matching probability at the time of track generating, and then build an approximate multinomial model for each track segment. The second is to get the statistical information of noise signal by analyzing the frequency spectrum performance of noise in the each track segment.

c)The computing result of above track data, radar parameters (including radar detection accuracy, radar coverage and scan cycle, etc), and environment features (including weather, landform, etc) are all gathered into database server, and then classify similar feature information into respective groups by clustering method to establish a data repository, through which data classifying and hierarchical management can be achieved, so as to ease executing data query during data generating.

Stage 2. In the second stage, new track model will be recombined in the light of track model in clustering database, data association and rule according to the requirement in designed air scenario. There are two steps:

a) According to simulation requirements, knowledge of the simulation track is transformed into parameters, such as the radar parameter, track velocity, environment configuration, and designed air scenario (e.g. straight line track, annular track, and figure-eight track, etc.). After the track in designed air scenario is divided into several segments, the same or similar segment track data resource can be found in the clustering database by seeking and matching these parameters.

b) This segment track data resource can be restructured into a new track model to fit for simulation requirements. At last, this new track is sent to the simulation test-bed of C^3I system to complete the automatic generation process of the radar simulation track.

3. Main Algorithm of Radar Intelligence Generation

3.1. Track Data Separation Based on Wavelet Transform

The radar target track data can be separated by filtering method. But the traditional digital filter method leads to phase shift and signal distortion, and the filter parameters can not accommodate the diversification of the track noise frequency, hence it will induce an unfavorable filtering result. It is widely accepted that the wavelet transform is a very practical tool to deal with non-stationary signal for its multi-resolution and time-frequency location property. A wavelet transform decomposes a signal into basis functions which are known as wavelets. Wavelet transform is calculated separately for different segments of the time-domain signal at different frequencies resulting in Multi-resolution analysis or MRA. Here, the signal-decomposition purpose is to extract actual true track and noise signal.

A wavelet transform in which the wavelets are discretely sampled are known as *discrete wavelet transform* (DWT). The DWT gives a multi-resolution description of a signal which is very useful in analyzing "real-world" signals. Essentially, a discrete multi-resolution description of a continuous-time signal is obtained by a DWT. It converts a series $x(i)$ (i=1,2,3,...,n) into one low pass coefficient series known as “approximation” and one high pass coefficient series known as “detail”. Length of each series is n/2.In real life situations, such transformation is applied recursively on the low-pass series until the desired number of iterations is reached. As an air target track, the noise added to the signal data is higher in frequency as compared to the actual signal. Hence if we can remove the high frequency components of the signal, we would be able to separate the noise. Reader can get more information about wavelet algorithm from other literatures. Thus the purpose can be completed after acquiring these two part signals. Here, the wavelet DB4 is chosen to effectively reduce track signal noise and to guarantee signals minimum losses. The main steps to execute this process are as follows:

a) Wavelet decomposition. After determining the wavelet basis, the wavelet coefficient series of flight track data would be obtained by computing and iterating a given number of times according to MRA.

b) Threshold processing of “detail”. Threshold is chosen to compare with the magnitude of absolute value of “detail”. If one absolute value in the “detail” is less than or equal to the threshold, its value of number would be set to zero. Here, heuristic threshold principle is used. Suppose that the length of signal $x(i)$ is n, let:

$$\sigma = \frac{\|x\|^2 - n}{n}$$
$$\zeta = \frac{(\sqrt{\log(n)/\log 2})^3}{\sqrt{n}} \tag{1}$$

then threshold θ is:

$$\theta = \begin{cases} \theta_2, & \sigma < \zeta \\ \min(\theta_1, \theta_2), & \text{others} \end{cases} \tag{2}$$

where θ_1 is Stein unbiased risk estimation threshold, and θ_2 is fixed threshold. Let $y(i)$ be ascending sequence of $|x(k)|$, and let $y_1(k) = y(k)^2$, then:

$$y_2(k) = \sum_{i=1}^{k} y_1(k)$$
$$r(k) = \frac{n - 2k + y_2(k) + (n-k)y_1(k)}{n} \tag{3}$$
$$\theta_1 = \sqrt{\min(r)}$$

and:

$$\theta_2 = \sqrt{2\log(n)} \tag{4}$$

Thus, we can make use of threshold θ to remove the noise from the “detail” of radar track data.

c) Signal reconstruction and separation. After the reconstruction of wavelet, the filtered radar track data is actual true track $\overline{x}(i)$ which very approaches realistic track was achieved. The data of noise signal can also be achieved by the function $s(i) = x(i) - \overline{x}(i)$, where $x(i)$ is raw data sequence of radar track. Thus, actual true track $\overline{x}(i)$ and the noise signal $s(i)$ are achieved.

3.2. Track Segmenting and Modeling

Track data segmenting, i.e. to divide the filtered track data into several sections, has two purposes: the first one is to build a more accurate target trajectory model; the second is to facilitate the implementation of data classification as well as searching & matching track. Multinomial modeling can be obtained through curve fitting to resemble the actual true track in multinomial modeling to ensure a successful matching between actual true tracks and designed.

3.2.1. Method of Track Segmenting

Radar target track is a space geometric curve. According to the differential geometry theory, the curvature of a random point on the space curve doses not change in the pace with curve location and direction for its inherent characteristics. Hence, an inherent feature for curve is curvature and its distribution. Through analyzing this feature, the radar target trajectory can be divided into several segments according to respective curvature that thereby reflects the different state of the track. Since the track data is discrete, the first-order curvature and the second-order curvature can be used to analyze the track curvature.

If the abscissa and ordinate of one track respectively is $x(k)$ and $y(k)$, where $k = 1, 2, \cdots, n$, the first-order discrete curvature is [2]:

$$\theta(k) = \tan^{-1}(\Delta y(k) / \Delta x(k)) \tag{5}$$

Where

$$\begin{aligned} \Delta x(k) &= \frac{x(k+1) - x(k-1)}{\sqrt{(x(k+1) - x(k-1))^2 + (y(k+1) - y(k-1))^2}} \\ \Delta y(k) &= \frac{y(k+1) - y(k-1)}{\sqrt{(x(k+1) - x(k-1))^2 + (y(k+1) - y(k-1))^2}} \end{aligned} \tag{6}$$

The first-order discrete curvature can reflect some changes of the trajectory curvature, for instance, a straight line is a fixed value, and the arc is an oblique line, but these results are still not intuitive. So it is necessary to use the second-order discrete curvature to get more information. The second-order discrete curvature is:

$$\theta_s(k) = \Delta x(k)\Delta^2 y(k) - \Delta^2 x(k)\Delta y(k) \tag{7}$$

where Δ^2 is the second-order difference operator. Then the discrete second-order difference formulas of $\Delta^2 x(k)$ and $\Delta^2 y(k)$ are:

$$\begin{aligned} \Delta^2 x(k) &= 2\frac{\dfrac{x(k+1) - x(k)}{\sqrt{(x(k+1) - x(k))^2 + (y(k+1) - y(k))^2}} - \dfrac{x(k) - x(k-1)}{\sqrt{(x(k) - x(k-1))^2 + (y(k) - y(k-1))^2}}}{\sqrt{(x(k+1) - x(k-1))^2 + (y(k+1) - y(k-1))^2}} \\ \Delta^2 y(k) &= 2\frac{\dfrac{y(k+1) - y(k)}{\sqrt{(x(k+1) - x(k))^2 + (y(k+1) - y(k))^2}} - \dfrac{y(k) - y(k-1)}{\sqrt{(x(k) - x(k-1))^2 + (y(k) - y(k-1))^2}}}{\sqrt{(x(k+1) - x(k-1))^2 + (y(k+1) - y(k-1))^2}} \end{aligned} \tag{8}$$

Since the second-order discrete curvature can obtain the change in the equal curvature locus, Eq. 9 can be used to locate the feature point on this locus:

$$C(k) = (\theta_s(k+1) - \theta_s(k)) / T_s \tag{9}$$

where T_s is the sampling interval. The location criterion is 3σ. If $C(k) > 3\sigma$, the corresponding coordinate values $(x(k), y(k))$ is the feature point on the track. The calculation equation for σ is as follows:

$$\sigma = \left(\frac{1}{n-1}\sum_{i=1}^{n}(C(i) - \bar{C})^2\right)^{\frac{1}{2}} \tag{10}$$

where $\bar{C} = \frac{1}{n}\sum_{i=1}^{n} C(i)$. This track's feature point is just the position of segmenting.

To display its effectiveness, Figure 2 is the result of MATLAB simulation for a radar race track.

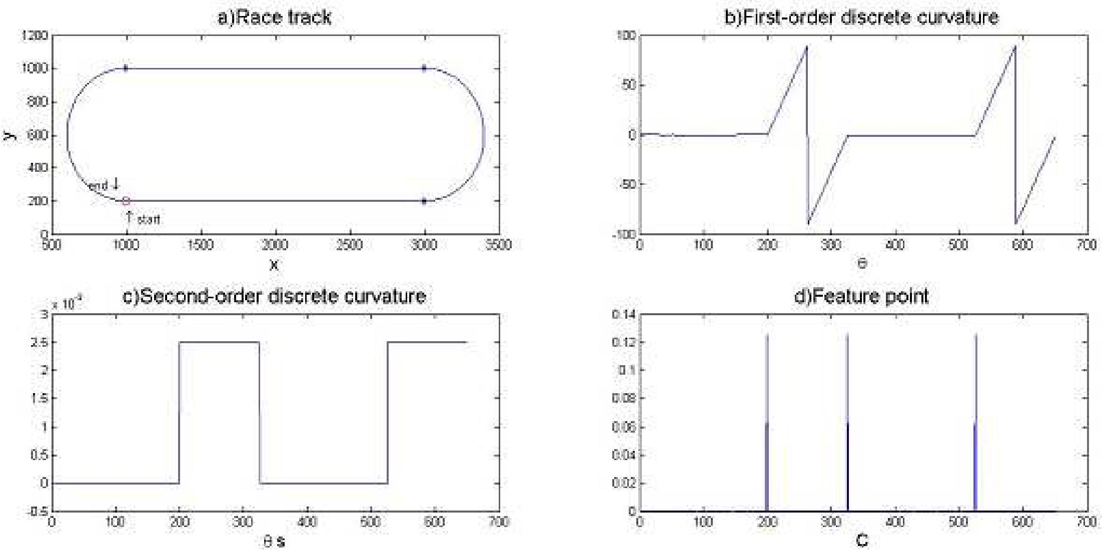

Figure 2. The simulation result of track segmenting method

As Figure 2 shows, this method has a good result. The radar track is divided into four segments by computing and analyzing. Here, the locations of three pulses in figure d), just corresponding to the locations of the three flags "*" in figure a), are inflection points of this radar track, so the result keeps consistent in each feature point with race track. From figure c), the value of second-order discrete curvature for a straight line is zero, and others are non-zero. Thereby, this method has solved the problem to divide radar track. But, in the process of practical application, the filtered data still involve some disturbance to deteriorate right result. The threshold must be chosen to evaluate the curve type. For instance, if θ_s locates in $[-5^\circ, 5^\circ]$, this curve is straight line, otherwise, it is arc.

3.2.2. Multinomial Modeling

If a normal flight target trajectory is continuous and smooth, namely n-order and differentiable, it can be arbitrary approximated by a multinomial. If there is a continuous and smooth curves $X(t)$, its multinomial fitting model is $\hat{X}(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n$, where n is fitting order, $A = [a_0, a_1, \cdots, a_n]$ is coefficients of the equation, $X = [x_1, x_2, \cdots, x_n]$ is output vectors, and $T = [t_1^n, t_1^{n-1}, \cdots, 1; t_2^n, t_2^{n-1}, \cdots, 1; \cdots; t_n^n, t_n^{n-1}, \cdots, 1]$ is time power vector. So the formula to solve the A through least squares fitting is:

$$A = (T^T T)^{-1} T^T X^T \tag{11}$$

Because radar target track is segmented, which not only reduces the fitting of the order, but also can improve the fitting precision, generally, a fifth-order multinomial can get a satisfactory effect.

3.3. Feature Clustering

Clustering is a process that the sets of physical or abstract objects are grouped into many classes according to the similar feature by dividing the object set based on the similarity between these objects and the expression of the mutual relationship in the space. The purpose in radar track data's feature vector is to establish a clustering parameter vector, to construct a clustering. The clustering parameter is a data structure or array that describes the radar track information [3~6]. What the contents of the track clustering parameter vector include are as follows:

a) Trajectory type: There are seven trajectory types which are distinguished by the shape of the flight path, like linear type, arc type, racetrack type, circular type, square type, figure-eight type, and others.

b) Fight attitude: Flight attitude reflects the aircraft flight state which includes level flight, climb flight, descending flight and others.

c) Average speed: Average speed is the whole track's average speed per hour.

d) Aircraft type: Aircraft type is the size of aircraft which includes large aircraft, medium-size aircraft, and small aircraft.

e) Environment feature: Environment feature includes the geography and the weather. The contents of the geography are plain, hill, forest, mountain and ocean. The contents of the weather are sunny, cloudy, rainy, and windy. In general, the windy day has more influence on the track data.

f) Radar parameter: Radar parameter includes radar detection accuracy, radar detection range, data sampling cycle, noise characteristics, and outlier characteristics. Noise characteristics consist of noise variance, noise rms value, power spectral density and envelope diagram of noise distribution. Outlier characteristics contain the occurring probability and the range of outlier peak value.

A successful and rational clustering depends on an appropriate threshold, so the threshold selection is an important step. Generally, the threshold is set from experience whose main principle is to ensure that the different characteristic of radar data can be distinguished, and to be much good to discriminate the data for its generation.

3.4 Track Matching and Restructuring

The purpose of matching is to pick out the historical radar data from the clustering database which has the minimum distance, the maximal similarity or the same feature value compared with designed air scenario. Feature matching has the following processes:

a) Some feature value, such as data sampling cycle, flight attitude, and environment feature and so on, must be same. This match will be not successful as long as one of these values can not be matched. Therefore, these values should be a priority for matching. If matched successful, the next step for matching could be continued.

b) Some matching items are fuzzy variables, which must be know how closely the variables of the historical radar data fit for the designed air scenario, so the similarity matching method can be used to resolve this problem. It is necessary to determine a threshold for every matching item. The closest track segment data, whose matching items belong to the range of its threshold, can be picked out from the clustering database.

c) The historical track segment data associates with the designed track segment data through searching the correlated data segments in clustering database. The air scenario created by the simulation test-bed of C^3I system is divided into straight lines and arcs through the way mentioned above. Then these segments are processed through matching spatial distance with the track data generated by multinomial fitting equation in time domain. The equation of Euclidean spatial distance is used to compute the spatial distance:

$$d(x, y) = \|x - y\| = [\sum_{i=1}^{n} (x_i - y_i)^2]^{1/2} \quad (12)$$

where x is the historical track segment data , and y is the track segment data of the designed air scenario. The best matching is the historical track segment data whose $d(x,y)$ is the smallest and below threshold.

If the track matching is successful, the track segments picked out are recombined according to the designed air scenario. If no track is found, Gaussian noise would be produced according to the given radar detection accuracy and added to the designed air scenario. Finally, a new track can be restructured to meet the simulation requirement.

4. Software Implementation

This software is called *radar intelligence automatic generating system*, which includes two parts: data mining and data generating. The software exploits the technology of mixed-language programming of VC and MATLAB. VC is for the interface problem, and MATLAB is for the matrix computing and data processing. All the historical radar data and its correlative information are deposited in Oracle database. The software can dispose multiple objects automatically in real time according to the setting. Figure 3 is the software's startup interface. There are three buttons in this interface: "Build Clustering Database", "Generate Radar Intellgence" and "Exit". Clicking "Build Clustering Database", the software will turn to the data mining interface to execute the process of stage 1. And clicking "Generate Radar Intellgence", the data generating interface will display to go on the process of stage 2. Clicking "Exit" will end the software. The software is installed in scenario server, which belongs to simulation test-bed of C^3I system.

Figure 3. The software's startup interface

Figure 4 is the data mining interface. This interface can import historical radar data to construct the clustering database. It is necessary to firstly set up related parameters about the information of radar track data, which are helpful for getting correct feature value. This software can automatically collect the radar data if only a data communication link is built between the software and actual radar based on wired or wireless networks.

Figure 5 is the data generating interface. This interface can match the designed air scenario to generate a new track that meets the setting conditions. The resut of new radar track can show in a chart, as Figure 5 shows, so that we can affirm its veracity. At last, the correct simaltion of radar track is sent to scenario generator of simulation test-bed of C^3I system in according with XML document.

The simlation of radar track has these advantages:a) get essentially the same as real radar track data; b) meet the data under many conditions, such as weather, geography and status; c) create and configure a flexible combination for tracks of various shapes ; d) construct many kinds of flying targets; e) have a good extensibility; f) achieve simplicity. Its software has already been successful applied to simulation test-bed of C[3]I system, and gotten a very fine effect.

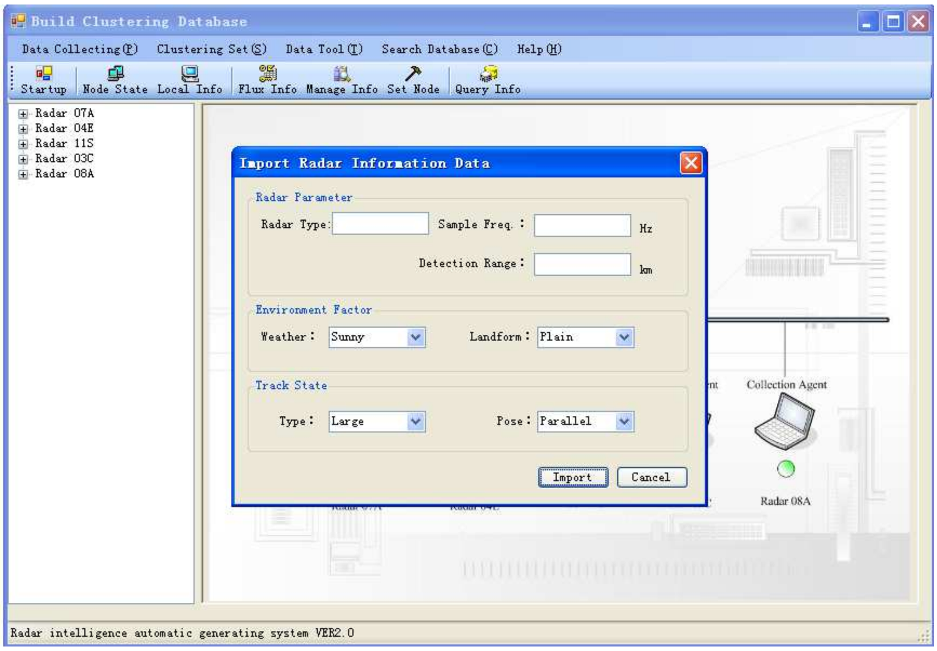

Figure 4. The data mining interface

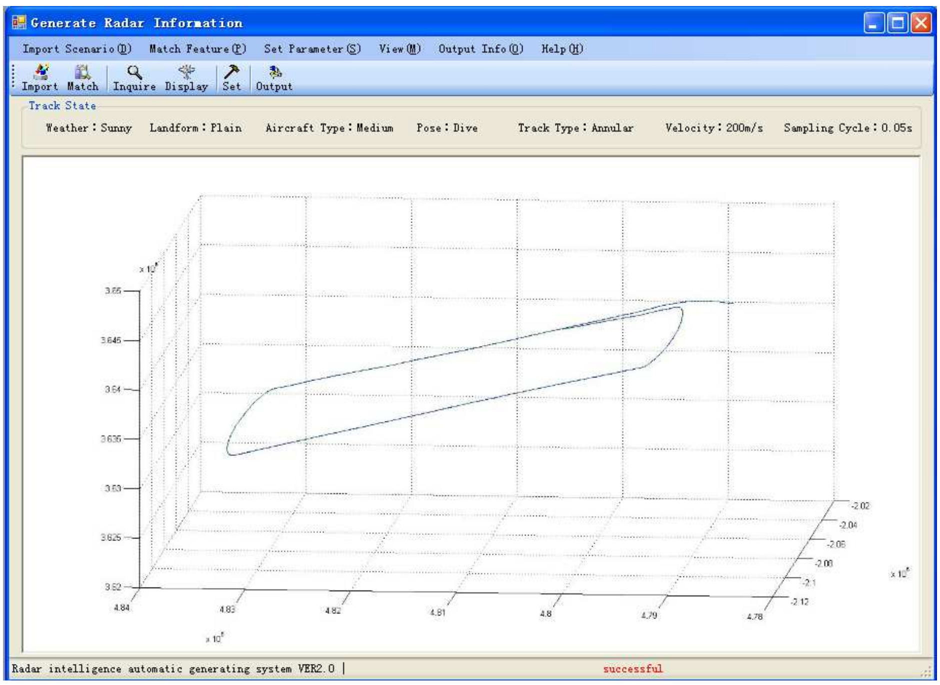

Figure 5. The data generating interface

5. Conclusion

How realistic the simulating radar air intelligent information directly influences the testing ability and effect of simulation test-bed of C^3I system. This paper proposes a new method to compose simulating radar air intelligent information by data mining technology through making use of historical radar data. Based on data mining techniques, the feature vector clustering database of historical radar data is built by implementing the steps that include filtering separation, piecewise fitting and feature clustering. Then the air intelligent data, which meets the scenario requirement of simulation test-bed of C^3I system, is generated by implementing the steps including feature association and track recombination based on the data in clustering database. This method has the advantages of good configuration flexibility, strong ability for generating many and varied track, and high fidelity. From its application effect in the C^3I simulated test, this new method can commendably satisfy the requirement of radar simulation, and has a certain value of application.

References

[1] GAO Jing, ZHOU Yu-fang, CHEN Li-feng. Research and application of radar detection error model based on data mining. *Command Control & Simulation*. 2009; 2(1): 94-97.
[2] MENG Na, ZHOU Yi-qi, QU Juan, CHEN Xin-li. Direct feature extraction of sectional curve based on a discrete curvature arithmetic. *Journal of China University of Mining & Technology*. 2009; 38(4): 571-575.
[3] CAI Jingying, XIE Fuding, ZHANG Yong. New fuzzy clustering algorithm based on feature weighted Mahalanobis distances. *Computer Engineering and Applications*. 2012; 48(5): 198-200.
[4] LI Bin, ZHANG Yin-he, CHEN Lu-hua, ZHANG Hong-tao. Radar characteristic of target with spatial precession. *Modern Defence Technology*. 2011; 39(4): 113-117.
[5] ZHU Ping, TIAN Chengjuan. Analysis on Doppler Radar Characteristic of a Short–time Severe Convective Weather in the Eastern Part of Qinghai Province. *Journal of Arid Meteorology*. 2011; 29(3): 336-342.
[6] GONG Xi, PEI Tao, SUN Jia, LUO Ming, Review of the research progresses in trajectory Clustering methods. *Progress in Geography*. 2011; 30(5): 522-534.