A New Machine Learning Framework for Air Combat Intelligent Virtual Opponent

Xiang Lei¹, Anxiang Huang¹, Tao Zhao², Yuqiang Su¹ and Chuan Ren¹

¹ Air Force Command College, Beijing 100097
² China Academy of Railway Sciences, Beijing 100081
Email: 240669315@qq.com

Abstract. In air combat simulation training systems, it is possible to train with “smart” opponents, which can greatly improve the pilot’s combat level. In the new generation of aircraft, an embedded virtual opponent training platform is also designed to enable real-time train with pilots. The core of training is in air combat training. Computers can make decisions based on the current situation. The idea of this paper is to allow pilots to training on specific tactics of warfare in the “human-in-the-loop” combat simulation system, forming strategies under different situation, and then establishing the relationship between machine learning input and output. Through the structured processing of simulation data, the radial neural network model is used for learning and training, and the trained model is used to predict and process the input situation data in real time, and a virtual opponent strategy is generated.

1. Introduction

After AlphaGo defeated Lee Sedol, machine learning as a new field of artificial intelligence has been pushed to new heights. Machine learning plays an important role in areas such as autopilot, speech recognition, and image recognition. The machine learning architecture consists of multiple layers of non-linear arithmetic units. Each lower-level output serves as a high-level input. It can learn effective feature representations from a large number of input data. Many high-level representations of input data are learned. Information is a good way to extract representations from data and can be used for specific problems such as classification, regression, and information retrieval [1]. The field of application of machine learning will be more extensive, especially in the man-machine game, man-to-man coordination, future robot operations and other fields, will play a huge role, with great military value. The air combat intelligent virtual opponent uses a computer to generate an opponent’s aircraft with certain intelligence, capable of possessing the tactical level of an excellent pilot, and becoming an excellent “smart” excellent sparring, in order to improve the pilot’s combat skills. The theory can also be used for UCAV. Machine combat systems, embedded training systems and other aspects. However, because the air-to-air confrontation itself is relatively complex, it involves many dimensions, complicated processes, high costs, and high risks. It has not been well implemented in engineering.

Currently, scholars have conducted a large amount of research. In the literature [2], the intelligent differential game method was introduced into the thinking of modern air combat. The idea of this method is to make use of artificial intelligence methods to predict the decision strategy of enemy aircraft, and to make a global situation assessment for both the enemy and the enemy. In the literature [3]-[4], a decision model for air combat based on manoeuvring movements, namely the expert system model, was established. Twenty-five typical tactical action libraries of three generations of fighter aircrafts were established, and corresponding expert knowledge bases were established. Correspondingly, the goal of corresponding manoeuvring decisions based on real-time air combat
situation information parameters is achieved. In [5]-[6], the air combat situation advantage function and air combat capability advantage function is used as the basis for air combat decision-making. Genetic algorithm is used to solve the air combat decision problem, air combat decision model is constructed according to the established advantage function, and air combat decision-making is completed. In [7]-[8], this paper uses the detection probability of the enemy as the cost function in the current situation to attack the target, and uses the nearest edge of the incoming weapon envelope as the optimal condition for escape. The improved A* is used. The algorithm explores the changes in the next heading and pitching angles of the smart opponent's aircraft, thereby obtaining the position and attitude of the next step of the intelligent opponent's simulation.

In the literature [9] a method for air combat decision-making based on multi-target is proposed by using cooperative priority algorithm. Firstly, the target assignment is performed using cooperative priority algorithm, and then the complexity and dynamics of air battlefield information are addressed by three layers of BP neural network. Uncertainty is used to implement the core air combat decision algorithm, and SOFM network is used to attack the target. This method can effectively solve the air combat decision-making problem under complex dynamic environment. The literature [10] classifies fleets from an air combat environment, and then distributes the tasks based on the results of the air combat situation assessment and the air-to-air combat capability assessment results of both the enemy and myself. After the end of an attack, it is evaluated, the loss is calculated, and the air combat situation is evaluated again. Make a second attack until the end of the air battle.

In literature [11], based on the framework of Markov decision process, the reinforcement learning method of concealed enemies in three-dimensional space is studied, and the advantage regions and exposed regions in the environmental model are defined. Aiming at the dimension disasters faced by high-dimensional state space strategy learning, a Q-learning algorithm based on radial basis function neural network (RBFNN) is presented. The method of hierarchical sampling for training samples is described, and it is aimed at different situations. The manoeuvrre strategy learning of the enemy was simulated. The literature [12] outlines that UCAV elaborated on the key technologies of the UCAV intelligent autonomous air combat from the aspects of air-to-air knowledge acquisition and self-learning, intelligent air combat decision-making, autonomous integrated control and technical verification methods, and proposed flight training and simulation training. Air combat experience and air combat models and tactical knowledge bases were formed. Through knowledge representation and knowledge reasoning, thinking agents and behaviour agents were formed. The literature [13] constructs a fitting network based on Gaussian radial basis function to solve aviation's cognitive behaviour model to provide support for the simulation of air combat decision-making for aviation. Through reinforcement learning, tactical decision-making experience is accumulated, and the demand for resources is greatly reduced and strengthened. The learning cycle, and ultimately the emergence of a reasonable manoeuvring strategy the validity and adaptability of the manoeuvring strategy model was verified through one-to-one air combat simulations, resulting in a warfare trajectory similar to that produced by human pilots.

The literature [14]-[18] used the genetic algorithm to effectively implement the evasive manoeuvring strategy, which is a good escape from the enemy's attack. The effectiveness of air combat decision-making has been improved, making the entire air combat decision more accurate. A completely new method is provided to solve the complex air combat decision problem, so that it searches for the optimal manoeuvring strategy in a large state space. Among the above research results, there are four main methods for solving the smart air combat: the first is based on the method of the expert system, the air combat knowledge base is established, and the action library is triggered under certain conditions to complete the air combat; the second type is based on the game theory method. Strict mathematics is pushed to the calculation, and the calculation is more accurate. The disadvantage is that it solves the multidimensional real-time dynamic air combat problem under the complicated system, the calculation is too cumbersome, and the requirement for the model establishment is relatively high; the third type is based on the neural network algorithm, and the advantage is that Effectively making decisions, the disadvantage is that the sample data acquisition is a bottleneck, which affects the accuracy of the algorithm; the fourth category is the use of genetic algorithms can make the global optimal, the disadvantage is the real-time performance is poor.
This article is based on the third category of neural network algorithm based on the idea of building a "human-in-loop" multi-machine simulation and confrontation system to generate massive sample data, solve the bottleneck problem of air combat against sample data, and use machine learning to learn humans. The tactical tactics of the pilot.

2. Statement of Problem
The intelligent generation of air combat virtual opponents involves the entire air combat process. The core is the organic combination of humans and weapons. It is a high-dimensional Markov process. The difficulty lies in the air combat decision-making. This is because decision-making is not only related to weapon performance, but also depends on the pilot's combat skills and mental state. In order to generate a "smart" air combat virtual opponent with a similar level of tactical tactics as real pilots, conventional optimization theory is not very effective. This is mainly due to:

First, the acquisition of sample data is difficult and costly. Air combat training data requires pilots to use actual equipment for tactical drills to obtain more realistic data. However, actual air combat training is very costly, especially for “multi-vs-multi” flight tactical confrontation training. Its risks and costs are unbearable.

Second, it is difficult to identify and describe the air combat situation. In multi-air combat, it is difficult for computers to determine operational intent through the aircraft's geometric situation. In particular, the initial state of many tactics is inherently deceptive, making computer identification and decision making more difficult. In addition, modern air battles are conducted under systems such as early warning aircraft and jamming machines, presenting disturbances and anti-interference, stealth and anti-stealth, and even the departure of defence-launched missiles. This has added to the description of the air war situation. Dimensions.

Third, it is difficult to make decisions. Air combat is the process of real-time optimization under multiple constraints. It is not only closely related to platform performance, weapon performance, and the state of each target, but also related to pilots' flight skills and psychological quality. Therefore, the establishment of air combat virtual opponents requires consideration of these factors before a decision can be made.

With the idea of machine learning, there are two main problems: one is the sample data problem, and the other is the high-speed solution problem. For the sample data problem, sample data can be generated by a countermeasure simulation system in the “human in the loop”. The problem of high-speed solution is not a big problem at present as the level of computer hardware is improved. Therefore, this method is used to solve the problem of air combat intelligence. Has a certain degree of feasibility.

3. General Ideas and Framework
By constructing a "human-in-loop" multi-machine confrontation simulation system, the pilot conducts a number of tactical confrontations against specific tactics through driving tactical simulators to obtain machine learning training samples, and then uses relative positions, postures, and key actions. Describe that the situation at time T is taken as the input of machine learning, and the state of the other side of T+Δt is taken as the output, which is the next strategy. By learning a large number of sample data, an air combat model is established. In the next man-machine confrontation, the trained model was directly used for real-time prediction, and the model with the tactical level of the opponent in the simulation confrontation was obtained. The flow is shown in Figure 1.
Establishing a multi "human vs human" air battle simulation system

Use this model for real-time forecasting and solving in the new "human-machine" training

Machine learning weights and parameter corrections, training prediction models

Figure 1. Air combat virtual intelligent opponent machine learning framework.

The main advantage of this method is that the decision-making part of the air combat is given to the pilot to complete the model, and sample data with decision information. By learning the tactical actions of the pilot to establish a model, it can be more close to the actual, the disadvantage is that the pilot needs to perform a large number of flight drills, but can optimize the processing, Increase the validity of the sample. Mainly in accordance with the five steps in the framework of Figure 1 design and implementation.

3.1 Establishing a Multi-Machine "Human-Human" Air combat Simulation System

The semi-physical simulation system for building human in the loop is mainly to construct a semi-physical simulation platform to build a realistic confrontation environment for pilots, which mainly includes flight control subsystem, fire control subsystem, motion platform subsystem, and vision subsystem. The network management subsystem integrates these subsystems to form a tactical simulator, and is then connected by multiple tactical simulators through a network to form a "man-in-the-loop" multi-machine real-time countermeasure system. A single tactical simulator structure shown in Figure 2.

Figure 2. Tactical simulator composition frame.

A number of tactical simulators are connected through a network control system, data management server, and control server to perform multi-machine real-time simulation. The structure is shown in Figure 3. The system is initialized by a pilot control system and the simulation environment is simulated. The tactical simulator is configured and monitored. The pilot operates the tactical simulator, exercises the specific tactics, and stores the generated data in the database management server according to the established data structure. The frame structure is shown in Figure 3.
3.2 Pilots Conduct Tactics Training
For machine learning, sample data is essential, sample data must be representative, and it must be able to cover problem air combat, but also to reflect the tactical level. If a rule-based strategy is adopted, the strategy to solve the multi-aircraft warfare problem is very obvious difficult. The method of learning the tactical actions of pilots can effectively solve the problem, that is, the organization fights in the "human-human" air-warfare countermeasure simulation system, for specific tactics such as: one-on-one, one-two, two-on-one, etc. You can also add conditions such as warnings and interference. Perform multiple simulations on the same style of tactics. In the simulation process, the pilot’s tactical thinking and operational principles are reflected. The main purpose is to complete the tactical action decision to the pilot. This is also the key difference from other methods. The data sample contains the decision information of the person.

3.3 Comprehensive Processing of Tactical Training Data
After the confrontation, the simulated data is processed. The information of these data mainly includes the aircraft number, model, red and blue attributes, aircraft position (x, y, z), attitude h, p, r, key actions (target locking, real-time interference, launch missiles, time identification (identifies the time stamp of each frame), and its stored data structure is as follows:

```
Structure data_frame{
  Int PlaneID;    //Aircraft flag
  Int PlaneSide;  //Aircraft properties
  Int PlaneType;  //Plane Type
  Float x,y,z;    //Longitude, Latitude, Height
  Float head, pitch, roll;  //Heading, pitch angle, roll angle
  Float v;        //velocity
  .......
  Int KeyAction;  // key action
  Int TimeIndex;  //time tag
}
```

As can be seen from the data structure, the semi-physical simulation system not only needs to record the position and attitude of the aircraft, but also records the key actions of the pilot and records the time stamp. When machine learning is performed, the storage structure needs to be normalized. If absolute flight parameters are directly used as input, the requirements for each pilot to perform drills are higher, and drills need to be conducted in the same space-time and the same situation. This is a relatively high cost. To be able to describe the relative situation data structure, to facilitate the collection of more training samples, you need to do a good correspondence between input and output,
that is, the mapping relationship between them, we T time as an input, the key action parameters of the other side of T + Δt As an output, establish an opponent's learning relationship. In order to reduce the data space, the literature [12] proposed a set of evaluation parameters describing the three-dimensional air combat from the perspective of the two parties, the distance relationship, the speed relationship, and so on. The document b enriched the description of the factors. This article described the air combat description according to actual needs. With key elements such as time stamps and key actions, we describe the data structure of the air combat situation as:

Table 1. Air combat situation description table.

| Name               | sign |
|--------------------|------|
| Relative Height    | A(m,n) |
| Relative Distance  | D(m,n) |
| Relative Heading   | H(m,n) |
| Red Pitch, Blue Pitch | Pr, Pb |
| Red Velocity, Blue Velocity | Vr, Vb |
| Tactical actions   | K    |
| Simulation time identifier | T |

A mapping function is used to represent any two red and blue aircraft r1. The situation of t2 at time T is:

\[ S_{T,t1,t2}(H, D, P, R, V, K, T) \]

For the “M-to-N” multi-aircraft tactical air combat state, the graph is used to represent the undirected graph in the graph theory. The diagonal matrix is used to represent the following 2 pairs of 2, for example, the numbers 0 and 1 are the red planes, and the numbers are 1, 2 For blue planes, then their situation matrix is:

\[
\begin{bmatrix}
S_{0,0} & S_{0,1} & S_{0,2} & S_{0,3} \\
S_{1,0} & S_{1,1} & S_{1,2} & S_{1,3} \\
S_{2,0} & S_{2,1} & S_{2,2} & S_{2,3} \\
S_{3,0} & S_{3,1} & S_{3,2} & S_{3,3}
\end{bmatrix}
\]

\[
\begin{bmatrix}
0 & S_{0,1} & S_{0,2} & S_{0,3} \\
0 & 0 & S_{1,2} & S_{1,3} \\
0 & 0 & 0 & S_{2,3} \\
0 & 0 & 0 & 0
\end{bmatrix}
\]

Because the relative situation between the two aircraft is the same, the upper triangular record can be used. After learning, it is generally converted to the standard list format:

Table 2. Standard format input parameter.

| F1 | F2 | A | D | H | P1 | ... | P2 | V1 | V2 | K | T |
|----|----|---|---|---|----|-----|----|----|----|---|---|
| 0  | 1  | 2000 | 10 | 30 | 10 | 19  | 600 | 700 | 000 | 10 |
| 0  | 2  | -3000 | 30 | 20 | 80 | 11  | 600 | 700 | 000 | 10 |
| 0  | 3  | 3000 | 30 | 20 | 80 | 11  | 600 | 700 | 000 | 10 |
| 1  | 2  | -4000 | 30 | 20 | 80 | 11  | 600 | 700 | 000 | 10 |
| 1  | 3  | 5000 | 30 | 20 | 80 | 11  | 600 | 700 | 000 | 10 |
| 2  | 3  | 6000 | 40 | 20 | 80 | 11  | 600 | 700 | 000 | 10 |

The demonstration data shows the positional parameters between various aircrafts at the simulation time T=10, and the tactical actions used by the red side. A is positive, indicating that the numbered aircraft is higher than the numbered aircraft, and H is positive. Negative means that the numbered aircraft is lower than the numbered aircraft. H denotes the relative heading. Using F1 as a reference,
the positive north is defined as 0, and the positive east is 90. It is subtracted in the clockwise direction, namely: \( H = H_2 - H_1 \).

For the air combat virtual intelligent opponent, the core problem to be solved is: In the current situation, how to make the next decision, this article is based on the pilot's tactical decision-making, in order to be able to parameterize the pilot's actions, we divide the operation into two major categories, one is the operation of the aircraft platform, the main are: the rate of change of speed, is to change the throttle, deceleration and acceleration, the most direct relevant parameter is the rate of change of speed. Pitch angle change rate: mainly by pulling the lever forward and backward direction, change the aircraft's pitch angle, so as to achieve the effect of changing the aircraft height. The third is to change the roll angle and operate the left and right levers to change the course of the aircraft. The other type is the operation of weapons. Operations related to operations mainly include closing radars, opening radars, and locking targets, releasing interferences, and launching missiles. In order to facilitate the recording, use the operation code to record.

### Table 3. Air combat output parameter table.

| Aircraft operation platform | Weapon operation                      |
|-----------------------------|---------------------------------------|
| 1, velocity Change value \( \Delta v \) | 4, Close radar, Operation code 000    |
| 2, pitch Change value \( \Delta p \)       | 5, Open radar, Operation code 001     |
| 3, roll Change value \( \Delta r \)        | 6, Lock target, Operation code 111    |
|                             | 7, Release interference, Operation code 100 |
|                             | 8, fire the missile, Operation code 110 |

The output represented by the map is:

\[
f_{T+\Delta t,i,\Delta v_i,\Delta p_i,\Delta r_i,code_i}(S_{T,m,n})
\]

Its meaning is: In the \( T \) situation, the blue side pilot \( i \) makes a decision at \( T+\Delta t \), which includes the pitch angle, roll angle, pitch angle and weapon key operation, based on the time \( T \), plus these The amount of change can lead to the next step.

After experiments, in the sample data, the stored data is determined according to the simulation refresh cycle, such as tactical level of confrontation, the refresh rate is greater than 40 frames per second, will make the visual scene has a continuous change trend, no jitter and delay feeling. There is no need to set the storage cycle so low, because people's decision cycle is not the millisecond level, the data interval is too thin, and it is not meaningful. The specific setting depends on the specific performance of the simulation system. The standard of measurement is that the current frame and the next frame must have obvious changes.

### 3.4 Machine Learning

As an important branch of artificial intelligence, machine learning is also a very big concept. Its core is to learn the laws from a large amount of historical data, so as to intelligently identify new samples or make predictions for the future. 3.1 And 3.2 are mainly for constructing and producing samples. The platform and method of data in 3.3, the main purpose is to normalize the data format, process the sample data of the flight drill, and establish the correspondence between the input and output, which is to establish a sample of machine learning. There are many types of air combat, which are generally based on combat units with “4 vs 4” or less. Different styles, different numbers of inputs and outputs, such as “2 vs 2” air battles, at a certain moment, the situation of 4 aircrafts at the time of input, And output is the next strategy for the two aircraft, different confrontation styles, different training models.

Because the strategy learning in the discrete state space requires a look-up table to store the values of state and action pairs, the scale of the state space is too large after being discretized, and the resources required for storing the query table are huge. In order to overcome the “dimensional disaster” problem of continuous state space strategy learning, this paper draws on the existing research results [19-20], and uses the radial basis function neural network (RBF-NN) excellent function approximation capability. Fitting continuous function of state and action pairs can greatly reduce the need for sample
size. RBFNN is a typical local approximation neural network. It has the advantages of parallelism, fault tolerance and best approximation capability\cite{21}. If the new input is the same as the sample input, then the output of the sample is followed. If the new input does not exist in the sample, it is the fitting value with the smallest error, thus solving the discrete sample correspondence problem in the multidimensional space. The basis function often used is a Gaussian function, that is:

\[
\sigma_j(x) = \varphi_j\|x - c_j\|_2/\sigma_j = \exp\left(-\frac{\|x - c_j\|^2}{\sigma_j^2}\right)
\]

In RBFNN, there are generally many hidden layers that are needed for input. This is clearly not suitable for air combat against such a large sample space. Therefore, before training, clustering methods should be used. Sample data is clustered to reduce the position of the centre point. We use the "2 to 2" air combat style as an example to establish a neural network learning model as shown in Figure 4. The input of the model is the situation at time T, which is the position and attitude relationship between the aircraft. The state matrix in 2.3 is used as the input, and the output of the model is the mapping function in 2.3. Samples can be derived from "human-human" simulation simulations.

![Diagram of "2 vs 2" air combat against data machine learning.](image)

The training process can be divided into two phases: the first phase is to determine the centre value \(c_j\) of the Gaussian kernel function at each node of the hidden layer based on all input samples, and the normalization constant \(\sigma_j^2\) to obtain the input layer and the hidden layer Radial basis functions. The second stage: After the parameters of the hidden layer are determined, the weight of the output layer is determined according to the sample using the principle of least squares. Through training, the model can be determined and a stable forecasting network can be obtained.

### 3.5 Using the Model to Predict

The predicted steps using the trained model are: First, set the initial position of the virtual intelligent opponent, which can also be randomly generated by the system; second, calculate the situation between the red and blue parties, and use the method in 3.3 to perform normalization; It is through the trained network to predict the position and tactical action parameters of the next frame of blue squares. Fourth, through these parameters, and then based on the current situation, through these changes, the position parameters of the next frame are solved. The fifth is to use the situation of the red and blue sides of the next frame as the next input. This cycle until the end of air combat.
4. Platform Construction and Experiments

The platform is mainly based on the comparatively mature aviation combat simulation system, which normalizes the flight data, and then conducts learning in the radial neural network to determine the weight of each parameter, and then carries out real-time forecast output. Through the conversion of the flight data generated by the simulation platform, a relatively stable model is trained, and then the data of each entity in the simulation system is acquired as the input of the model in real time, and then the predicted data is forwarded to generate the air combat virtual intelligent opponent. The experimental combat space is 100*100Km. The three-generation aircraft is used against the three-generation aircraft. The scope of the missile is 40Km. The pilots are organized to perform "2 vs 2" countermeasures 20 times. The simulation flight data is stored and processed in the aviation combat simulation system, added machine learning module. A total of 20 (intercepted from the mid-range interception phase data) minutes of air combat confrontation, 40 frames per second data, a total of 48,000 data, a total of 20 times, a total of 960,000 sample data, through machine learning, got a "2 vs 2" air battle confrontation model.

Fig. 6a shows the situation of a "2 vs.2" confrontation at time T, and Fig. 6b shows the situation of "man-machine" versus time T+100. The left side of the figure shows the red side track and the right side shows the blue side track. The blue square (right) in FIG. 6c represents the trajectory of a certain T time predicted by machine learning, and the blue square (right) in FIG. 6d represents the predicted T+1000 time trajectory after machine learning. As can be seen in the figure, the air-to-air fictitious opponent’s aircraft generates a relatively continuous warfare trajectory.

5. Conclusion

In this paper, the “human-in-loop” air combat simulation system is used to generate machine learning data samples, and the simulation data is normalized for input and output. A multi-aircraft air combat situation description method is established and radial neural network learning is used. The framework learns simulation data, trains a relatively stable network model, and then makes real-time predictions.
It can generate similar human trajectories and key actions in real time. When the actual project is used, there are still some issues that need further study:

The first is the construction of the aviation combat simulation system, involving various weapon models. The accuracy of these models determines the credibility of the sample data. The performance and operational mechanism of different aircraft types and weapons are different. Therefore, many models need to be constructed. The accuracy of the sample data itself cannot be guaranteed.

The second is that when pilots conduct virtual combat in air combat, the use of weapons and tactics are all random. Some of them use different tactical decisions under the same conditions, which leads to the diversity of our data samples, which may lead to insufficient learning convergence.

Third, air combat is a high-dimensional Markov process. When a pilot conducts an air-to-air battle, the sample space needs multiple air-to-air battles. It can try to cover the boundary value of the sample space and reduce the fitting error. It is necessary to use experimental design before simulation. Theoretical methods: Scientifically design the factor levels of each factor dimension, such as the height difference, and perform multiple experiments at various height differences.

Fourth, the learning framework needs to be further optimized. This paper only adds a prototype system of the machine learning framework to the original simulation system. It also needs to further optimize and improve the framework and weight parameters to improve the learning efficiency and the convergence rate.

6. References

[1] Jianwei Liu, Yuan Liu and Xionglin Luo 2014 Research and development on deep learning 2012 Application Research of Computers 31 1921-29
[2] Yining Wang and Yuxian Jiang 2003 An Intelligent Differential Game on Air Combat Decision Flight Dynamics 21 66 -70
[3] Guierez L, Vanhtsevanos G and Henk B 2003 to the adaptive mode transition control of unmanned aerial vehicle Proceedings of 2003 American Control Conferemme Denver Colorado
[4] Qiang Wei and Deyun Zhou 2007 Research on UCAV's Intelligent Decision-Making System based on Expert System Fire Control & Command Control 21 66-70
[5] Tao Zhang, Lei Yu and Zhongliang Zhou 2012 Decision-making for Air Combat Maneuvering Based on Variable Weight Pseudo-parallel Genetic Algorithm Flight Dynamics 30 470-474
[6] James S, Jonathan P, Brian W, et al 2010 Air combat strategy using approximate dynamic programming Journal of Guidance, Control and Dynamic 33 1641-1654
[7] Xiang Lei, Anxiang Huang, Xiaowen Feng, Jinsong Li and Yunpeng Gao 2014 Approach for Intelligent Rival Airplane Real-time Route Planning Journal of System Simulation 25 1881-85
[8] Xiang Lei, Anxiang Huang, Yuqiang Su, etc 2016 Approach for Intelligent Rival-Air-Plane Threats Evading16th Asia Simulation Conference and SCS Autumn Simulation Multi-Conference AsiaSim/SCS Autumn-Sim Berlin Springer
[9] Mingan Tong and Linsen Li 1999 Air Combat Decision of Cooperative Multi-target Attack and Its Neural Net Realization Acta Aeronautica ET Astronautica Sinica 20 309-312
[10] Yong Chang, Changsheng Jiang and Zhiwei Chen 2011 Decision-Making Based on Fuzzy Neural Network for Air Combat of Multi-Aircraft against Multi-Target Electronics Optics & Control 18 13 –17
[11] An Xu, Yingxin Kou and Lei Yu 2012 Steady Engagement Maneuvering Strategy with Q-learning Based on RBFNN for Air Vehicles,Systems Engineering and Electronics 34 98 –102
[12] Yaofei Ma, Guanghong Gong and Xiaoyuan Peng 2010 Cognition Behavior Model for Air Combat Based on Reinforcement Learning Journal of Beijing University of Aeronautics and Astronautics 36 379 - 383
[13] Changqiang Huang and Shangqin Tang 2016 From Alphago to Alphaeagle: on the Intelligent Autonomous Air Combat Technology for UCAV Journal of Command and Control 2 261-264
[14] Rui Zhou and Zongji Chen 2010 Learning Evasive Maneuvers Using Genetic Algorithms Control and Decision 16 465-472
[15] Shaowu Dai, Bin Zuo and Yunan Hu 2006 Learning Evasive Maneuvers Using Genetic-Annealing Algorithms Flight Dynamics 24 77-80
[16] Deyun Zhou, Feng Li and Xiaobo Pu 2002 On Improving Tactical Planning in Air Combat with Genetic Algorithm Journal of Northwestern Polytechnical University 20 109-112

[17] James S, Jonathan P, Brian W, et al 2010 Air-combat strategy using approximate dynamic programming Journal of Guidance, Control and Dynamic 33 1641-1654

[18] Chang H S, Fu M C, Hu J Q 2007 An asymptotically efficient simulation-based algorithm for finite horizon stochastic dynamic programming IEEE Trans on Automatic Control 52 89-94.

[19] Iwata K, Ikeda K, Sakai H. 2006 A statistical property of multi-agent learning based on Markov decision process IEEE Trans Neural Networks 17 829 - 842.

[20] Hong X, Chen 2009 A new RBF neural network with boundary value constraints IEEE Trans. on Systems Man and Cybernetic-Part B Cyberneticw 39 298 -303