Defence Science Journal, Vol. 70, No. 2, March 2020, pp. 183-189, DOI : 10.14429/dsj.70.14232

Modelling and Simulation of Cognitive Electronic Attack under the Condition of System-of-systems Combat

Zhang Yang[@,*], Si Guangya[@], and Wang Yanzheng[@]

[*]*Luoyang Electronic Equipment Test Center, Luoyang - 471 003, China*
[@]*National Defence University, Beijing - 100 091, China*
[*]*E-mail: zhangyang1986128@sina.com*

ABSTRACT

From the height of system-of-systems combat and operational perspective, the operations of cognitive electronic warfare (CEW) was analysed, and its main process and links were described. Secondly, the jamming effectiveness evaluation (JEE) model of cognitive electronic attack (CEA) operations was established based on the interference side, in which the change of threat degree was used as the measure index of jamming effectiveness. Then, based on the Q-learning model, an intelligent countermeasure strategy generation (ICSG) model was established, and the main steps in the model were given. Finally, on the basis the JEE model and the ICSG model, the simulation experiment was carried out for CEA operations. The result showed that combining the JEE model with the ICSG model can express the main process of the operations of CEW, as well as proved the validity of these models.

Keywords: Cognitive electronic warfare; Reinforcement learning; Modelling and simulation

1. INTRODUCTION

The design of the new operational concept of cyber war operations is a powerful driving force to promote the development of the equipment system, as well as the innovation of the tactic and operational method of cyber war. Cognitive electronic warfare (CEW) operational concept, technology and equipment system are important support for strengthening the cyberspace capacity building in the US "Third Offset Strategy"[1]. The progress of the research and the application of combat are being widely concerned by the domestic and international professionals in the cyberspace. By carrying out the modelling and simulation of CEW to the high-level operational concept under the background of joint operations, it will be able to promote the research of combat theory, equipment systems and key technologies of CEW.

At present, there are few research results on CEW at home and abroad, focusing on two aspects: First, from the qualitative analysis, the professionals analyse the basic concepts, functional components and operational applications of CEW[2,3]. Zhang[2] analyses the characteristics of CEW and the possible technical composition from the perspective of single platform, which lays a foundation for the research of the CEW architecture. Zhou[3] analyses the intelligent principles of the cognitive radio, cognitive radar and typical CEW projects, as well as the relationship between CEW and cyber warfare, from the perspective of the observation, orientation, decision, action combat loop (OODA loop). These articles establish the basis of the quantitative analysis of the CEW and let the high-level decision makers know the importance of CEW. The second is to start from the signal level, the professionals study the jamming effectiveness evaluation, intelligent countermeasure strategy generation and waveform optimisation[4-7]. Investigator Guang proposes a possible organisation and operation model for the cognitive reconnaissance module in CEW, and gives two processing procedures of the module, which are proposed for the parameter measurement in the frequency domain and the spatial domain of the CEW equipment[4]. Amuru[7] has developed a new framework to enable the jammer continuously changing the jamming strategy in some typical scenarios based on the reinforcement learning (RL) algorithms, where the reward is associated with the state transitions rather than itself. These achievements drive the development of CEW from different specific research fields by a long way.

The type of the systems-of-systems (SoS) confrontation is the basic operation model in the future, as well as the joint operational SoS is the background and target of CEW. The research of modelling and simulation of CEW as a complete operational concept from the perspective of SoS confrontation is very important, that can facilitate high-level professionals understanding of the operational mechanism and technical mechanism of CEW operations, recognise the enormous combat capability of CEW operations, and highlight the operational effects, while the research achievements are rare currently. This paper starts with the analysis of CEW from the perspective of the OODA loop, and studies the modelling and simulation of CEA operations as an important component of CEW operations, combining the reinforcement learning model with cognitive electronic attack (CEA) operations.

Received : 06 March 2019, Revised : 22 December 2019
Accepted : 06 February 2020, Online published : 09 March 2020

2. OPERATIONAL ANALYSIS OF CEW BASED ON THE OODA LOOP

Cognitive electronic warfare operations dynamically adjust the attack and protection strategies based on real-time environmental situation awareness, operational effectiveness assessment, and knowledge learning and accumulation results. Through the closed loop of the operational process, intelligent and efficient EW will be implemented. CEW operations will effectively respond to the complex electromagnetic situation in future battlefields. From the operational view analysis, CEW operations are subversive and autonomous cyber war operations, and the combat process CEW operations conforms to the OODA loop model. The analysis of CEW from the perspective of OODA helps to understand the whole process and key links of CEW, and then guides the modelling of CEW operations. From the level of SoS confrontation, CEW operations are mainly composed of four processes such as "cognitive interception—cognitive processing—cognitive decision-cognitive jamming", and constitute a closed loop from reconnaissance to interference and evaluation.

Cognitive interception intercepts the environment signals about the targets through the dynamic perception of the battlefield electromagnetic environment, and measures and stores the signals. With the support of the signal feature knowledge base (SFKB), cognitive processing uses signal feature mining and other techniques to extract, locate and identify signal features, and to perform signal threat assessment and judgment. Through unified, standardised knowledge engineering, new knowledge of new signal characteristics and threats are generated, then the knowledge base is updated. Cognitive decision-making, with the support of the jamming performance and jamming strategy knowledge base, analyses the current target signal characteristics, evaluates the previous interference performance, dynamically generates a new and more appropriate countermeasure strategy, and updates the jamming performance knowledge base (JPKB) and jamming strategy knowledge base (JSKB). Cognitive jamming, with the support of the power management knowledge base (PMKB), reasonably allocates the energy of each jamming beam, performs power control and management, and releases interference for the target.

In CEW operations, cognitive interception and cognitive processing constitute cognitive electronic reconnaissance (CER) operations together, while cognitive decision-making and cognitive jamming constitute cognitive electronic attacks operations (CEA) together. Because cognitive interception and cognitive jamming focus on describing the flow and utilisation process of signals, they are problems that are concerned with signal level modelling and simulation. The model resolution is high, which is difficult to describe in the SoS confrontation simulation. At the same time, because cognitive processing and cognitive decision-making focus on the process of generating and utilising information, they are problems which are concerned in the SoS-level. The model resolution is relatively lower, which can be simplified and portrayed in SoS confrontation simulation. The result of CER operations is to provide data, information, and intelligence support for CEA operations. In this paper, aiming at the urgent need of SoS confrontation simulation, CEA operation is the main modelling object, and a CEA operation model for high-level operational concept demonstration is constructed.

3. CEA OPERATIONAL MODELLING METHOD BASED ON RL MODEL

The emergence of CEW technology and equipment has changed the form of traditional electronic warfare, so that when the CEW equipment performs CEA operations, it can continuously "trial and error", that is, transform the electronic attack style, and detect the changes of the signal characteristics and working status of the enemy electronic information equipment, so as to evaluate the effect of the CEA operations online, then continuously adjust the EA style to finally learn the best confrontation style under different circumstances. This "trial and error" feature of CEA has strong similarities with the RL model. In fact, RL is a high level of abstraction of CEA operational in the field of machine learning. The head of the BAE Josh Niedzwiecki believes that[1] "the learning process of the CEW system is highly consistent with the RL model, and the CEW system adapts to the learning environment through RL."

In view of the similarity between the RL model and the CEA operations, the RL model can provide guidance and reference for the modelling of CEA. The classical RL model is as shown in Fig. 1. The entities of the model includes two categories: the agent and the environment. The agent can perform actions, while the environment is an interactive object whose behaviour is uncontrollable. The main work-flow of the agent is as follows: first, the agent observes the environment to obtain the state of the environment; secondly, the agent selects the action according to a certain policy and interacts with the environment to get the reward of the action (Reward). Again, the agent repeat the above steps until the end. The work-flow of the environment is as follows: first, receiving an action of the agent and reacting to the action; secondly, passing the state of the environment and the reward of the action to the agent. The ultimate goal of the RL model is to maximise the rewards.

According to the RL model, based on the high abstraction and generalisation of CEA, a CEA operations model can be constructed. The flow chart is as shown in Fig. 2.

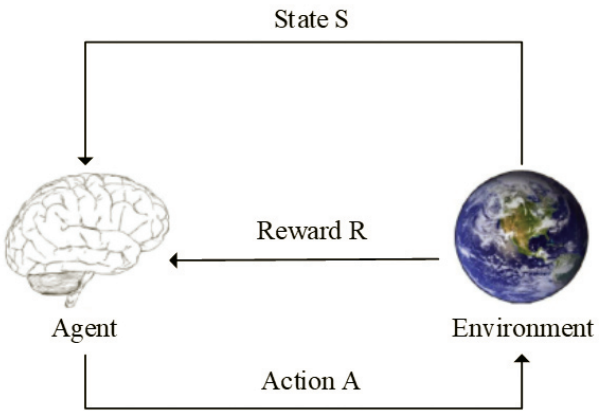


Figure 1. Diagram of RL model.

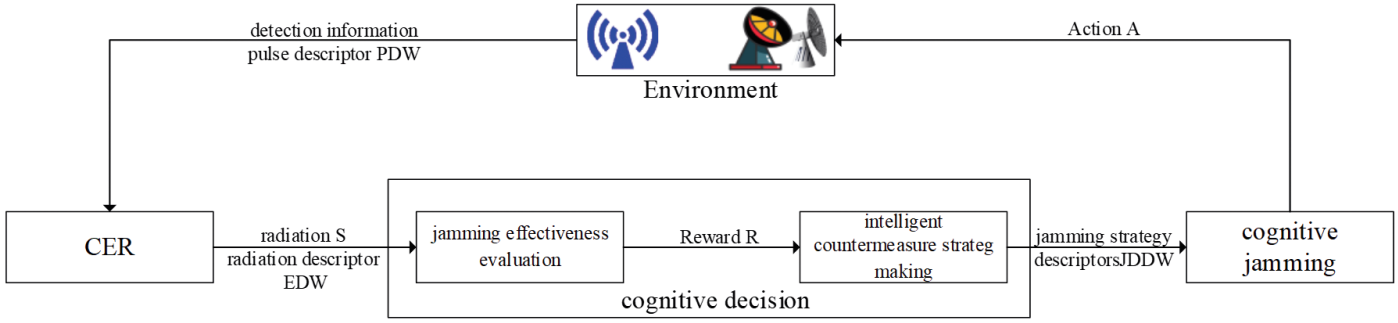


Figure 2. Diagram of the process of cognitive electronic attack action model.

The work-flow of the CEA operations model is as follows:

Step 1: The CEA operations obtains the emitter descriptor word (*EDW*) generated by the CER operations, which is analogous to the state S of the agent acquisition from environment in the RL model, and goes to Step 2.

The *EDW* is used to describe the main parameters describing the characteristics of the radiation obtained by the process of sorting, identifying, locating and threatening the radiation, including radiation position(*POS*), pulse width(*PW*), pulse repetition frequency (*PRI*), radiation system(*SYS*), radiation application(*APP*), radiation type(*TYPE*), radiation working state(*ST*), and radiation threaten level(*TL*). The BNF form of the *EDW* is as Eqn. (1):

$$EDW ::=< POS >< PW >< PRI >< SYS > \\ < APP >< TYPE >< ST >< TL > \quad (1)$$

Step 2: According to the change of the *EDW* before and after the jamming, using the jamming effectiveness evaluation (JEE) model which is based on jamming-side perspective to valuate the effectiveness of the interference, analogous to the agent obtaining the reward R of a certain action, go to Step 3.

The available jamming performance evaluation methods include jamming-based evaluation methods and jammed-based evaluation methods. The evaluation rules include information rules, power rules, efficiency rules, etc. The influencing factors include power, distance, frequency, timing, style, and system, etc.

Step 3: Learning the interference situation, and using Intelligent confrontation strategy generation (ICSG) model to select a better interference pattern based on the jamming strategy, generating an jamming decision description word (*JDDW*), analogous to the policy in the RL model, and going to Step 4.

The *JDDW* is to optimise and control interference patterns, interference waveform, and interference resources. In view of the requirements of the SoS confrontational and simulation, this paper regards the decision of interference pattern (*JTYPE*) as the most important interference strategy decision content, and can be expressed as the *JDDW* , which is formally described as is as Eqn (2):

$$JDDW ::=< JTYPE > \quad (2)$$

The available reinforcement learning methods are mainly various types of model-free reinforcement learning algorithms, including Monte Carlo (MC) algorithm and Timeporal Differences (TD) algorithm.

Step 4: Generate an interference waveform according to the *JDDW* , optimise the jamming resources, and jam the radiation again, analogous to the agent performing action A in the RL model.

4. JEE MODEL BASED ON THE JAMMING-SIDE PRESPECTIVE

Jamming effectiveness evaluation is an important guarantee to realize the closed loop of CEW. It transforms the information acquired by CER operations into the decision-making basis needed for cognitive decision-making, and provides optimised direction driving for CEA operations. Only accurate and correct evaluating interference performance can provide the correct feedback for the learning behavior of CEA. Interference effectiveness assessment based on the interference side perspective is one of the important features of CEW advanced in traditional electronic warfare. It is one of the important capabilities of CEW to achieve auto countermeasure. By detecting of the information and analysing the changes of the signal characteristics, behaviour characteristics and working state of the radiation, due to the interference, CEA operations can evaluate the effectiveness of the jamming and provide a quantitative basis for the adjustment of the interference pattern. Based on the SoS confrontation level, this paper takes radar jamming to one multi-phase array radar (R) as an example. Based on the existing research results, this paper initially explores the JEE model based on the jamming side perspective under the SoS confrontation condition.

4.1 Qualitative Analysis of Jamming Performance and Radar State

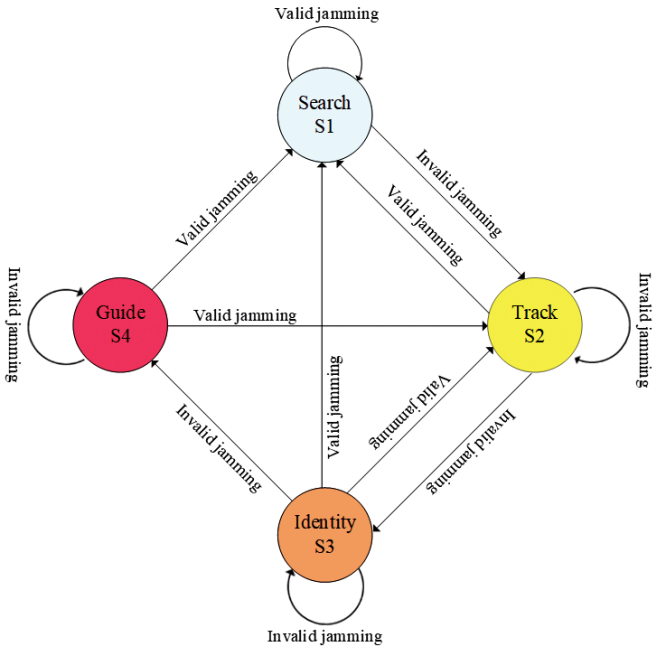
From a macro perspective, there is a certain mapping relationship between interference performance and radar state. This is the basis for the modelling of jamming effectiveness assessment based on the interference side perspective. When the radar R is in different working states, the interference performance of different interference patterns of CEA operations may be different, even the interference performance of the same jam pattern may be different[8,9]. Changes of the radar's working state after being disturbed can be used to assess whether the jamming is valid.

When the radar R is in the search state, the effective interference makes the radar signal processing and the data

processing unable to detect the target signal normally. At this time, the radar R can only keep the search state unchanged. Otherwise, the invalid jamming will not affect the normal processing of the target signal by the radar, and the radar R will go to the tracking state. When the radar R is in the tracking state, the effective interference will make the radar unable to track the target stably, even lose the target, re-enter the search state, or maintain the tracking state unchanged. Otherwise, the invalid interference will not affect the radar R tracking of the target, then the radar will transfer to identify other false targets for imaging, affecting the radar's imaging process for the target, maintaining the recognition status or moving to tracking and search status. When the radar R is in a guided state, effective jamming will cause the radar to lose its target or track to other targets and move to the tracking state or search state.

The relationship between the interference performance and the radar working state can be represented by a state transition diagram, as shown in Fig. 3(a).

Assuming that the working state $S_1 = Search$, $S_2 = Track$, $S_3 = Identity$, $S_4 = Guide$ of the radar R constitutes a working state set $S = \{S_i\}(1 \le i \le 4)$. Interference patterns include $J_1 = Aim$, $J_2 = Block$, $J_3 = Scan$, $J_4 = Distance\ \ Cheat$, $J_5 = Speed\ Cheat$, $J_6 = Angle\ Cheat$, $J_7 = Target\ Cheat$, which constitute jamming pattern set $J = \{J_i\}(1 \le i \le 7)$. For Fig. 3(a), the state transition matrix can be used for abstract drawing, as shown in Fig. 3(b). ST is the state transition matrix, and the row and column are all working states. "1" indicates that the state $S_p(1 \le p \le 4)$ of a row can be directly converted to the state $S_q(1 \le q \le 4)$ of a column, while "-1" indicates that the state S_p can not be directly converted to the state S_q. For example, $ST(2,3) = 1$ indicates that S_2 can be directly converted to the state S_3.



(a)

$$ST = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \end{array} \begin{array}{c} \begin{array}{cccc} S_1 & S_2 & S_3 & S_4 \end{array} \\ \begin{pmatrix} 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & -1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & 1 \end{pmatrix} \end{array}$$

(b)

Figure 3. (a) Diagram of radar working state conversion and (b) Matrix of radar working state conversion.

4.2 Jamming Performance Evaluation Function based on Threat

At the time t, the CEA operations use the interference pattern $J_t \in J$ to jam the radar radar R which the working state is $S_t \in S$. After the interference, the CER operations detect the radar and obtains the observation O_{t+1}, and the working state of the radar R change to $S_{t+1} \in S$. When the interference reduces the threat of the working state (except the search state), the jamming is effective. Otherwise, while the threat is continuously increased or unchanged, the jamming is invalid. Therefore, the change of the radar threat degree due to the interference can be used as a measure of the interference performance.

This paper defines the jamming performance function $Eff(\bullet)$ which used to calculate the jamming performance of J_t. The expression is as shown in Eqn. (3):

$$\begin{aligned} Eff(J_t \mid S_t, S_{t+1}) &= P(S_{t+1} \mid J_t, S_t) \times R(S_t, S_{t+1}) \\ &= P(S_{t+1} \mid J_t, S_t) \times (W(S_{t+1}) - W(S_t)) \end{aligned} \quad (3)$$

In this Eqn. (3), $W(\bullet)$ is the threat degree function, indicating the threat degree of the working state of the radar. $P(S_{t+1} \mid J_t, S_t)$ is the effective jamming probability, which indicates that when the radar working state is S_t, the probability that the working state will change to S_{t+1} by using the interference pattern J_t. $P(S_{t+1} \mid J_t, S_t)$ can be given by the simulation test. $R(S_t, S_{t+1})$ is the reward function for work status changes, and $R(S_t, S_{t+1}) = (W(S_{t+1}) - W(S_t))$. In general, when the interference is effective, the farther away from the state S_t and S_{t+1}, the smaller the effective interference probability. The working state S_t can be obtained by the CER operations processing the observation O_t, and is included in the radiation descriptor EDW_t for the radar R.

If the performance function $Eff(J_t \mid S_t, S_{t+1}) > 0$, the threat level is reduced and the interference pattern J_t is valid. Otherwise, if $Eff(J_t \mid S_t, S_{t+1}) \le 0 (S_t \ne S_1)$, the threat level is increased or remains unchanged, and J_t is invalid. The value of the jamming performance function $Eff(\bullet)$ is equivalent to the reward in RL model, and can provide a basis for the optimisation of the jamming action (i.e., the interference pattern).

5. ICSG MODEL BASED ON Q-LEARNING ALGORITHM

The description and construction of the intelligent confrontation strategy generation process in the CEA operations can drive the CEA operations models auto and autonomously in the SoS simulation. The ICSG model is the key and difficult point of modelling the CEA operations. Q-learning (QL) algorithm proposed by Watkins is an important improvement

of TD algorithm in RL. Due to its simple algorithm, rapid convergence and convenient use, it is widely used and regarded as an important milestone in RL development[10,11]. Considering the small amount of sample data and the high real-time simulation requirements, based on the table Q-learning algorithm, this paper only uses the intelligent countermeasure strategy for the radar R as an example, and constructs the ICSG model to intelligently generate countermeasures against radiation.

5.1 Description of the Problem

Assume that the working state set S of the radar R is known, and the action a ($a \in J$) is a certain interference pattern. All jamming patterns can be used for any state s, that is $\forall s \in S$, $A(s) = J$. The Definition of the state-action value function $Q(s,a)$ is as given in Definition (1).

Definition (1): state-action value function $Q(s,a)$

$Q(s,a)$ is the Q function, which indicates that when the state is s, the expectation of the maximum discount future reward obtained after the action a is performed, and represents the quality of a particular action in a given state, as shown in Eqn. (4)[5]:

$$Q^{\pi}(s,a) = E_{\pi}\{R_t \mid S_t = s, A_t = a\} = E_{\pi}\{\sum_{k=0}^{\infty} \gamma^k R_{t+k+1} \mid S_t = s, A_t = a\} \tag{4}$$

In Eqn. (4), π is a strategy, which indicates a rule for selecting an action in each state. γ ($0 \le \lambda \le 1$) is a discount factor, which indicates a relative proportion of future rewards. R_t is a reward when the state is $S_t = s$ and the action is $A_t = a$, which can be represented by a rewards function $R(s,a)$.

Knowing the threat degree $W(s)$ of each state s and the interference function $Eff(J_t \mid S_t, S_{t+1})$, the goal of the ICSG model is to find the optimal strategy π^*, so that the radar R will be finally in the target state $s_E = S_1$ with the greatest expected reward regardless of the initial state.

The task of the ICSG model is to train the Q function, get the best action corresponding to each state, at last form the optimal strategy π^*. When the training is over, any initial state $s_s \in S$ is as given, and the optimal action and interference pattern can be generated autonomously.

5.2 The Main Process of the ICSG Model

The progress of the ICSG model based on Q-learning algorithm are as shown in Fig. 4. The main steps of this model are as follows.

Step 1: Initialisation Q function, and going to Step 2.

Due to the number of elements of the radar working state set is $|S| = 4$, and the number of elements of the jamming pattern set is $|J| = 7$, so the Q function is initialised to a 4×7

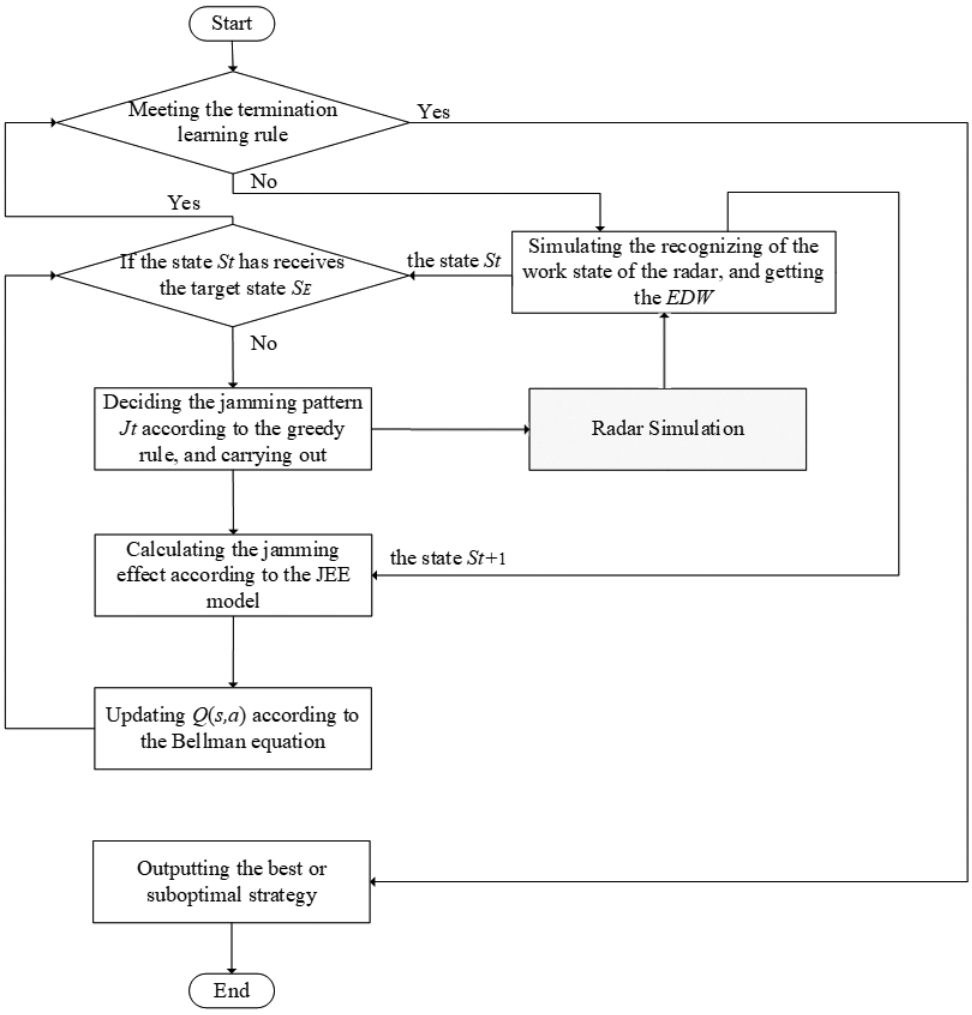


Figure 4. The flow chart of the ICSG model.

zero matrix, in which the row of the matrix represents the state of the radar, and the column represents the action that may be taken.

Step 2: When the termination learning rule is met, go to Step 8. Otherwise, going to STEP 3.

The termination learning rule includes two rules, one is the Q value convergence rule $\Re(QStop)$, the other is the iteration time rule $\Re(QNum)$. $\Re(QStop)$ means that the learning process is terminated when the difference between the value $Q(s,a)$ of the last two experiments is less than or equal to the convergence threshold ρ^{Q}. $\Re(QNum)$ means that when the times of experiments is greater than the iterations times threshold N^{Q}, the learning process is terminated.

Step 3: According to the results of the CER operations, identify the current status of the radar R, go to STEP 4.

Step 4: When the target state $s_E = S_1$ is not reached, go to Step 5. Otherwise, go to STEP 2.

Step 5: According to the ε greedy rule $\Re(\varepsilon Max)$, select the interference pattern J_t as the action a_t in the jamming pattern set J, and go to Step 6.

$\Re(\varepsilon Max)$ refers to the relationship between the exploration of unknown knowledge (randomness) and the use of known knowledge (greediness) in the process of reinforcement learning, so that the RL process uses both the most rewarding actions which is known, and the unknown actions.

Step 6: After performing the action a_t (that is after the interference pattern J_t is used to interfere with the radar R, according to the radar state transition diagram and the result of the CER operations on the radar R, the radar working state s_t becomes $s_{t+1} \in S$ (s_{t+1} obtained by the SoS confront simulation environment). Iteratively update the value function $Q(s_t, a_t)$ according to the Bellman equation of Eqn (5) below, and go to Step 7.

$$Q(s_t,a_t) \leftarrow Q(s_t,a_t) + \alpha \times [R(s_t,a_t) + \gamma \max_a Q(s_{t+1},a) - Q(s_t,a_t)] \tag{5}$$

In this equation, α is the learning efficiency factor, and the reward function is $R(s_t,a_t) = Eff(J_t \mid s_t, s_{t+1})$ based on the JEE model.

Step 7: Let the status s_t be the current status s_{t+1}, go to Step 4.

Step 8: Output the optimal strategy π^* according to Eqn (6) and end the process of the ICSG model.

$$\pi^* = \arg\max_a (Q^*(s,a)) \tag{6}$$

6. SIMULATION EXPERIMENT

In the Matlab environment, this paper conducts simulation experiments on CEA operations, especially the core model-the ICSG model. The result expresses the autonomic decision-making process of CEA operations under SoS confrontation conditions.

6.1 Simulation Experiment Environment Setting

It is assumed that after CER operations, the radiation description of the radar R is EDW formed by the CEW equipment, and the state of the radar R is s_t after the state recognition process. Assume that the time t is close to the time $t+1$, the CEW equipment has not widely moved. The value of the threat degree $W(\bullet)$ is quantified by the expert analysis method, such as $W(S_1) = 0$, $W(S_2) = 30$, $W(S_3) = 70$, $W(S_4) = 100$. Therefore, a threat degree transfer matrix WR can be obtained as shown in Eqn. (7). Each element $WR(x,y)(1 \le x, y \le 4)$ in this equation represents the threat change value when transitioning from state S_x to state S_y.

$$WR = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \end{array} \begin{array}{c} \begin{array}{cccc} S_1 & S_2 & S_3 & S_4 \end{array} \\ \begin{pmatrix} 100 & -30 & -\text{inf} & -\text{inf} \\ 30 & 0 & -40 & -\text{inf} \\ 70 & 40 & 0 & -30 \\ 100 & 70 & -\text{inf} & 0 \end{pmatrix} \end{array} \tag{7}$$

Simply, it is assumed that the effects of the seven jamming patterns on the radar working state transition are independent of the initial state, and the jamming factors of these jamming patterns on the state transition are $JamFac(i) = (|J| - i + 1)/|J|$.

Based on the threat degree transfer matrix WR nd the jamming pattern influence factor $JamFac$, the $|S| \times |S| \times |J|$ third-order jamming performance matrix $R = WR \otimes JamFac$ is constructed as the reward matrix, as shown in Eqn. (8):

$$R(x,y,i) = WR(x,y) \times JamFac(i) \tag{8}$$

In the SoS confrontation simulation environment, it is necessary to simulate the change of the radar working state after the interference. If the jamming is not affected, the radar state transition probability matrix PT can be fixed as shown in Eqn. (9) assumingly, which is based on the prior knowledge according to some experts in the electronic warfare field, and its value can effect the simulation result.

$$PT = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \end{array} \begin{array}{c} \begin{array}{cccc} S_1 & S_2 & S_3 & S_4 \end{array} \\ \begin{pmatrix} 0.3 & 0.7 & 0 & 0 \\ 0.2 & 0.3 & 0.5 & 0.0 \\ 0.1 & 0.2 & 0.2 & 0.5 \\ 0.05 & 0.25 & 0 & 0.7 \end{pmatrix} \end{array} \tag{9}$$

6.2 Simulation Experiment Conclusion and Analysis

In the experiment, the greedy factor $\varepsilon = 0.1$, the learning efficiency factor $\alpha = 0.3$, the discount factor $\gamma = 0.8$, the iteration times threshold $N^{Q} = 10000$, and the convergence threshold $\rho^{Q} = 1$, the resulting Q function matrix is as shown in Table 1.

According to the Q function, the final interference strategy corresponding to the radar state can be obtained. For example, when the radar state is S_1 which represents search, the interference strategy J_1 should be selected so as

Table 1. Q function array

Q Table	J_1	J_2	J_3	J_4	J_5	J_6	J_7
S_1	482.67	466.90	447.86	431.98	414.59	404.00	391.18
S_2	411.51	408.21	404.23	402.47	397.10	393.62	386.60
S_3	366.42	359.25	349.20	343.10	339.91	331.84	324.77
S_4	395.71	383.11	376.11	364.33	358.67	344.08	331.98

$\pi^*(S_1) = J_1$ at this time. It can be seen that the interference strategy corresponding to each state at this time is the jamming pattern J_1, which is consistent with the good jamming effect of the jamming pattern J_1 in any of the preset states.

To study the influence of the learning factor on the ICSG model, the greedy factor $\varepsilon = 0.1$, the discount factor $\gamma = 0.8$, the iteration number threshold $N^Q = 10000$, and the convergence threshold $\rho^Q = 1$. The learning efficiency factor α are taken as [0.2, 0.4, 0.6, 0.8, 1.0] respectively, then the relationship between the sum of the maximum Q value of the states and the learning efficiency factor-the number of iterations in the Q value function matrix is obtained, as shown in Fig. 5.

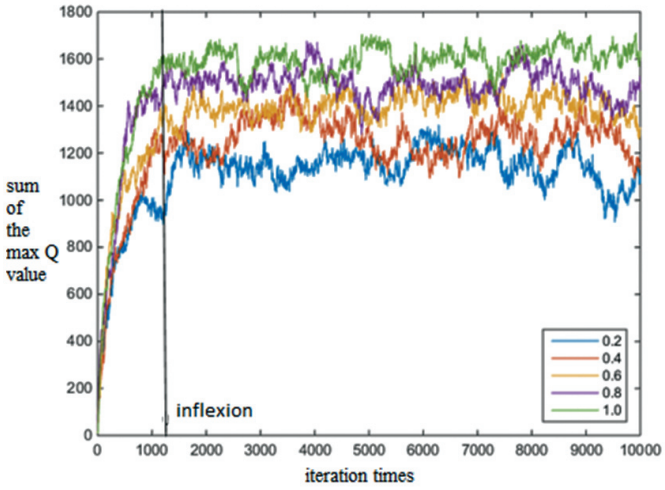


Figure 5. Diagram of the relation between the iteration times and the sum of the maximum Q value.

From the overall trend of each curve, the larger the learning efficiency factor, the larger the sum of the maximum Q value, and the better the overall interference effect of the "trial and error" process. At the same time, the timing of the inflection points of each curve is not much different. The simulation shows that the learning efficiency factor can influence the impression of the ICSG model, deciding the best learning effect, while has little effect on the learning speed.

To sum up, this simulation experiment shows that the ICSG model based on the Q-learning algorithm can express the main process of CEA operations, as well as the influence of the learning efficiency factor. The experiment verifies the effectiveness of the ICSG model, the JEE model and some algorithms proposed in this paper.

7. CONCLUSIONS

CEW closely combines artificial intelligence technology with electronic warfare technology, greatly improving the operational effectiveness and changing the "rules of the game" of electronic warfare. In this paper, based on the requirements of SoS confrontation simulation and combat concept demonstration, the main processes of CEA operations are analysed from the perspective of OODA loop. The JEE model based on interference side view and the ICSG model based on Q-Learning are established, and initially realised the modelling and simulation of CEA operations. The whole simulation experiment time is about 5.2s and is a little slow to generate the jamming strategy in real-time. Because the state-action matrix is big, and the radar working state is continuously transferred according to the state transition probability. In the future, deep Q network technology can be used to cope with more complex state-action spaces and improve iteration speed.

REFERENCES

1. Knowles, J. Regaining the advantage cognitive electronic warfare. *J. Elec. Def.*, 2016, **12**, 8-9. doi: 10.1109/CIP.2010.5604178
2. Zhang, W. & Yang, D. Some thoughts on cognitive electronic warfare. *Elec. War. Tech.*, 2017, **32**(4), 6-9. doi: 10.3969/j.issn.1674-2230.2017.04.002
3. Zhou, B.; Dai, H. & Qiao, H. Research on recognition EW and cyberspace operation based on "OODA loop" *CAEIT. Thea. J.*, 2014, **9**(6), 556-562. doi: 10.3969/j.issn.1673-5692.2014.06.002
4. Guang, X. Cognitive reconnaissance radar signal processing technology. *Elec. Sci. Tech.*, 2016, **29**(7), 143-146. doi: 10.16180/j.cnki.issn1007-7820.2016.07.041
5. Li, Y.; Zhu, X. & Gao, M. Design of cognitive radar jamming based on Q-Learning algorithm. *Trans. Bei. Ins. Tech.*, 2015, **35**(11), 1194-1199. doi: 10.15918/j.tbit1001-0645. 2015.11.017
6. Amuru, S. & Buehrer, R. Optimal jamming using delayed learning. *In* Proceeding of IEEE Military Communication Conference, 2014, Oct 2014, pp. 252-260. doi: 10.1109/MILCOM.2014.252
7. Hanawal, M.; Abdel-Rahman, M. & Krunz, M. Joint adaptation of frequency hopping and transmission rate for anti-jamming wireless systems. *IEEE Trans. Mob. Comp.* 2016, **15**(9), 2247-2259. doi: 10.1109/TMC.2015.2492556
8. Zhao, Y. & Xu, W. A method of real-time electronic attack effectiveness evaluation based on state transition of radar. *Elec. Inf. War. Tech.*, 2016, **31**(3), 42-46. doi: 10.3969/j.issn. 1674-230.2016.03.009
9. Nicholas, O. & Warren, P. Threat evaluation and jamming allocation. *IET Radar. Son. Navi.*, 2017, **11**(3), 459-465. doi: 10.1049 /iet-rsn.2016.0277
10. Manz, B. Cognition: EW gets brainy. *Elec. Def. J.,* 2012, **10**(35), 8-10. doi: 10.2139/ssrn.1261980
11. Ameer, S. Cognitive electronic warfare system. *Proc. IEEE Cog. Radio. Net. 2016*, Oct 2016, pp. 150-155. doi: 10.13140/RG.2.2.10939.62240

CONTRIBUTORS

Dr Zhang Yang received the PhD in National Defence University, Beijing, China, in 2019.
He is the correspondent author of this paper and is responsible for writing and the simulation. He establishes the JEE model and CSG model.

Mr Si Guangya is a professor in National Defence University, Beijing, China.
He is responsible for writing and guides the research as a whole.

Dr Wang Yanzheng received the PhD in National Defence University, Beijing, China, in 2013.
He is responsible for writing as well as proofreading the manuscript.