electronics

MDPI

Article

An Electronic Jamming Method Based on a Distributed Information Sharing Mechanism

Pan Zhang [1,2], Yi Huang [1,2] and Zhonghe Jin [1,2,*]

[1] Micro-Satellite Research Center, Zhejiang University, Hangzhou 310027, China; zhangpan_zju@zju.edu.cn (P.Z.)
[2] Zhejiang Micro-Satellite Research Key Laboratory, Zhejiang University, Hangzhou 310027, China
* Correspondence: jinzh@zju.edu.cn

Abstract: In an electronic jamming system, the ability to adequately perceive information determines the effectiveness of an electronic countermeasures strategy. This paper proposes a new method based on the combination of a multi-agent electronic jammer and an information sharing mechanism. With the development of intelligent technology and deep learning, these technologies have been applied in electronic countermeasure game systems. Introducing intelligent technology into the electronic confrontation system can greatly improve decision-making efficiency. At the same time, a multi-agent electronic countermeasure cooperative system based on the information sharing method can break through the limited information perception capabilities of a single agent, thereby greatly improving the survivability of jamming systems in electronic warfare. Experimental results show that our method requires a lower jamming-to-signal ratio than the single jammer method to achieve effective electronic jamming. In addition, the electronic jamming parameters can be updated automatically as the external electromagnetic environment changes quickly, realizing a more intelligent electronic jamming system.

Keywords: information sharing; multi-agent electronic jammer; Q-learning; situational awareness

Citation: Zhang, P.; Huang, Y.; Jin, Z. An Electronic Jamming Method Based on a Distributed Information Sharing Mechanism. *Electronics* **2023**, *12*, 2130. https://doi.org/10.3390/electronics12092130

Academic Editors: Hirokazu Kobayashi and Toshifumi Moriyama

Received: 20 March 2023
Revised: 25 April 2023
Accepted: 28 April 2023
Published: 6 May 2023

1. Introduction

An electronic jammer can jam targets by controlling the jamming signal parameters (e.g., signal center carrier frequency (Fc), signal bandwidth (BW), pulse repetition frequency (PRF), and pulse width (PW)) [1,2]. However, there is room for improvement in terms of perception ability and adaptive jamming signal parameters for electronic jammer systems.

Certain tasks in complicated environments are difficult to complete due to the limitations of a single jammer. Distributed jamming systems have been constantly evolving in recent years due to their vast application potential [3]. Electronic jammer swarms demonstrate group behavior by allowing individuals to interact locally in order to solve a global collaborative task [4]. Meanwhile, the present trend in electronic countermeasures (ECMs) is moving in the direction of intelligence and distribution. Distributed ECMs, as opposed to conventional centralized ECMs, take advantage of their number and area to achieve better jamming of targets [5]. The effect of distributed suppression and deception jamming on radar detection effectiveness is discussed in [6].

A number of experts have introduced artificial intelligence technology to electronic warfare. It was not until 2010 that DARPA [7,8] reported on adaptive electronic warfare learning, and the value of adaptive radar confrontation [9,10] was developed gradually. Wang Shafei [11] of the PLA Military Academy proposed a cognitive electronic warfare system architecture combining artificial intelligence with electronic warfare, greatly improving the ability of the electronic warfare system to perceive threat signals and make jamming decisions. Xing Qiang and Zhu Weigang [12] proposed an intelligent radar confrontation method based on Q-learning. By analyzing the convergence time and cycle times of the Q

matrix, the jamming effect was realized along with independent perception and intelligent decision-making, and the adaptability of the radar confrontation system was improved as well.

This paper proposes an information sharing method based on a multi-agent electronic jamming system that realizes the real-time perception and efficient processing of three-dimensional electromagnetic situations. Simultaneously, multi-agent jamming parameters can be autonomously adjusted based on situational information sharing mechanisms between different jammers, ensuring that the electromagnetic situation of agents in the system is balanced. Meanwhile, this method can avoid jamming decision errors due to poor electromagnetic information.

2. Multi-Agent Cooperative Jamming Method

2.1. Range Gate Pull-Off Electronic Jamming Model

Figure 1 shows a schematic diagram of traditional static electronic jamming. In Figure 1, the active radar performs dynamic electromagnetic sensing of the external environment by transmitting signals with a specific modulation mode.

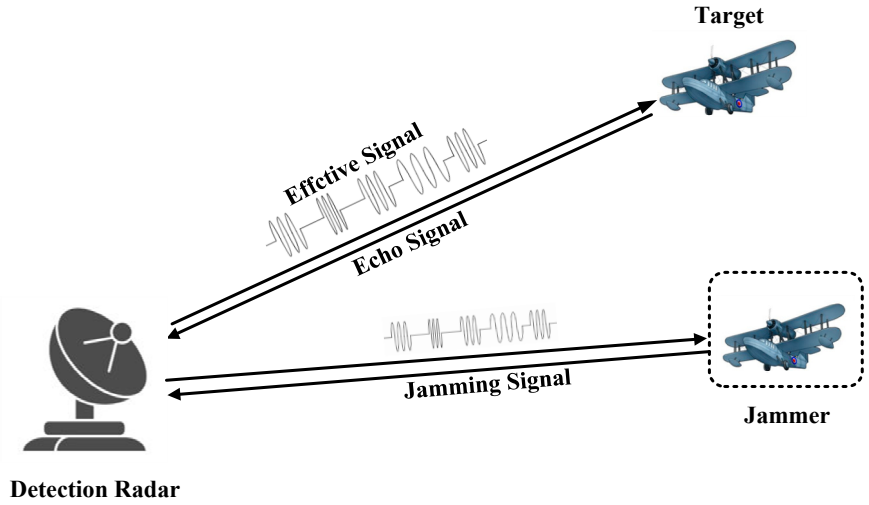


Figure 1. Electronic jamming system.

The electronic jammer intercepts the effective reconnaissance signal emitted by the radiation source to achieve rapid parameter estimation of the effective signal, then makes a specific electronic jamming decision for the radar source.

The signal $S(t)$ received by the radar can be expressed as

$$S(t) = A \cdot \exp\left[j \cdot 2\pi \cdot \left(f_0 t + \frac{1}{2}\mu t^2\right)\right] \tag{1}$$

where A is the amplitude of the signal, f_0 is the carrier frequency, μ is the frequency modulation slope of the signal, and t is the sampling time.

If the distance between the radar and the target is R, the receiver receives the radar echo signal after a delay of $\frac{2R}{c}$. Therefore, the echo signal $S_{rec}(t)$ received by the radar receiver can be expressed as

$$S_{rec}(t) = S(t - \frac{2R}{c}) + n(t) \tag{2}$$

where R is the distance between the radar and the target, c is $3e^8 m/s$, and $n(t)$ is modeled by white Gaussian noise with a distribution obeying the following formula:

$$n(t) \sim N(0, \sigma^2) \tag{3}$$

The receiver detects the effective linear frequency modulation (LFM) signal according to the energy accumulation method. The effective signal detection process is shown in Figure 2.

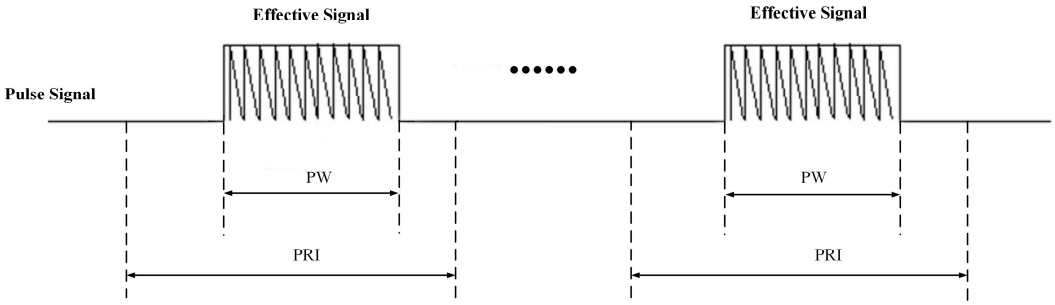


Figure 2. Pulse energy detection and pulse width measurement.

In the time-domain signal flow, the energy of the noise is much smaller than that of the effective signal. After the ADC sampling, the energy accumulation value of the 16-point discrete time-domain echo signal received by the jammer is compared with the threshold value. Here, we set the threshold amplitude to four times the amplitude energy of the echo signal. To ensure the detection of effective signals while reducing the probability of false alarms caused by noise, we adopt a 16-point energy accumulation method to detect LFM signals. When the cumulative value is continuously greater than the threshold, it represents the starting point of the effective signal, i.e., the time of arrival (TOA). Sliding detection is continuously carried out on the time-domain signal flow. When the cumulative value is continuously lower than the threshold value, it represents the time of end (TOE) of the effective signal. The time interval from the start point TOA to the end point TOE is the pulse width, Pw. The time interval between the first TOA and the second TOA is the pulse repetition period, PRI. The pulse repetition rate PRF is the reciprocal of the pulse repetition period PRI [13–15], i.e., $PRF = \frac{1}{PRI}$.

The carrier frequency Fc of the LFM signal can be obtained by multiplying the signal delay conjugate $S_{\text{rec}}(T_1 - t)^*$ by the original signal $S_{\text{rec}}(t)$, where T_1 is the time delay. The effective bandwidth BW of a pulse signal can be calculated using the pulse width Pw and pulse signal modulation slope μ, that is, $BW = \mu \cdot Pw$. At this point, the pulse descriptor word (PDW) parameter estimation process for the target radar's signal is completed.

The electronic reconnaissance process is shown in Figure 3, including the estimated parameters.

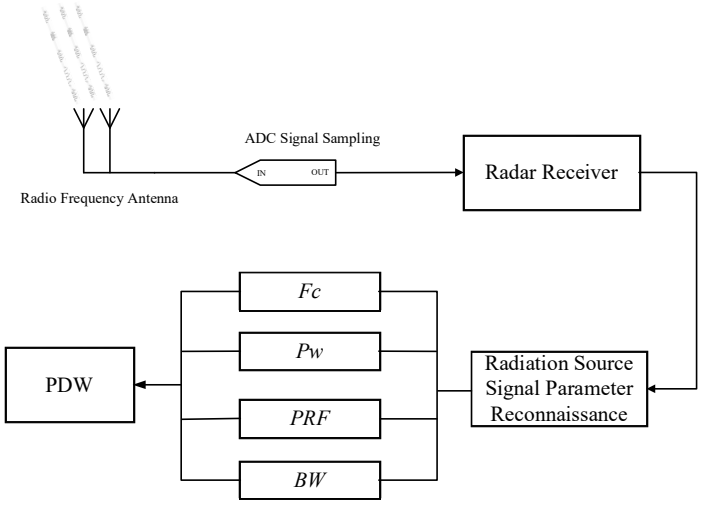


Figure 3. The estimation process of electromagnetic parameters of radiation sources.

RGPO jamming adds a time delay to the received signal $S(t)$, resulting in a difference in distance between the jamming signal $S_j(t)$ and the real echo signal $S_{rec}(t)$ that interferes with the target's recognition of the real echo signal. The jamming range gate can be either positive or negative. A positive distance means that the jamming signal is far from the radar receiver, while a negative distance means that the interference signal is close to the radar receiver.

Here, we apply a jamming range gate at a distance R_j to the radar signal. Therefore, the electronic jamming signal after range gate pull-off (RGPO) jamming is as follows:

$$S_j(t) = S\left[t - \frac{2\left(R + R_j\right)}{c}\right] + n(t) \tag{4}$$

where R_j is the range gate. The jamming range gate can be either positive or negative.

2.2. Balancing Parameters of the Multi-Agent Situation Information

Assuming the electronic jamming system contains N electronic jammer nodes, the situational information of a single electronic jammer agent can be expressed as follows:

$$\begin{aligned} X_i &= C_i\left(\{\text{Information of Agent } i\}_{i\in N}\right) \\ &= C_i : (Fc, Pw, PRF, BW) \end{aligned} \tag{5}$$

where Fc is the center carrier frequency, Pw is the signal pulse width, PRF is the pulse repetition rate, and BW is the effective bandwidth of the signal. Moreover, $C_i : (Fc, Pw, PRF, BW)$ represents the situation information of the parameter set (Fc, Pw, PRF, BW) at the ith electronic jammer agent.

Then, the situation information that the multi-agent system can perceive is expressed as

$$C = \sum_{i=1}^{N} (X_i | i \in N_i) \tag{6}$$

where $\sum_{i=1}^{N}$ is the sum situation information of N jammers.

Here, we assume that the weight of the parameter is γ. The parameter weight is represented as follows:

$$\begin{cases} 4 \cdot \gamma = 1 \\ \alpha_{Fc} = \alpha_{Pw} = \alpha_{PRF} = \alpha_{BW} = \gamma \end{cases} \tag{7}$$

where $\{\alpha_{Fc}, \alpha_{Pw}, \alpha_{PRF}, \alpha_{BW}\}$ corresponds to the weight of Fc, Pw, PRF, BW.

Because the situation parameter information between different jammer agents is not exactly the same at the same time, it is necessary to observe and balance the parameters between the different agents. A partial derivation of the reconnaissance situation parameters in the adjacent electronic jammer is carried out; the expression is as follows:

$$\begin{aligned} \frac{\partial(X_i - X_{i-1})}{\partial Fc} &= \frac{\partial(C_i : Pa(Fc, Pw, PRF, BW))}{\partial Fc} \\ &\quad - \frac{\partial(C_{i-1} : Pa(Fc, Pw, PRF, BW))}{\partial Fc} \\ &= \frac{\partial(C_i | Fc - C_{i-1} | Fc)}{\partial Fc} \end{aligned} \tag{8}$$

where $\frac{\partial F}{\partial Fc}$ is the derivative of Fc with respect to the function F.

The expression of the situational information transmission of the jamming parameters (Fc, Pw, PRF, BW) between different jammer agents can be expressed as

$$\begin{cases} \frac{\partial(X_i - X_{i-1})}{\partial Pw} = \frac{\partial(C_i|Pw - C_{i-1}|Pw)}{\partial Pw} \\ \frac{\partial(X_i - X_{i-1})}{\partial PRF} = \frac{\partial(C_i|PRF - C_{i-1}|PRF)}{\partial PRF} \\ \frac{\partial(X_i - X_{i-1})}{\partial BW} = \frac{\partial(C_i|BW - C_{i-1}|BW)}{\partial BW} \end{cases} \tag{9}$$

where $\frac{\partial F}{\partial(Pw, PRF, BW)}$ is the derivative of (Pw, PRF, BW) with respect to the function F.

The purpose of having multiple sensing nodes in the electronic jamming system is to improve the accuracy of the system regarding the situation parameter information of the radiation source and to reduce the misjudgment of the situation information caused by environmental factors. If the reconnaissance parameter information of different agents is unbalanced, the parameters of each agent need to be adjusted accordingly. For example, if a node does not work or is interfered with, resulting in a large difference in the derivative of the parameter set (Fc, Pw, PRF, BW) between adjacent agents, then the problematic situation information should be ignored or discarded.

Therefore, the effective inter-agent situation parameter information $C_i|(Fc, Pw, PRF, BW)$ should be the situation information after the problematic parameters are removed. The modified situation information expression is as follows:

$$\begin{cases} C_i|Fc = Mean\left(\prod\limits_{k \in N_i}(C_k|Fc)\right) \\ C_i|Pw = Mean\left(\prod\limits_{k \in N_i}(C_k|Pw)\right) \\ C_i|PRF = Mean\left(\prod\limits_{k \in N_i}(C_k|PRF)\right) \\ C_i|BW = Mean\left(\prod\limits_{k \in N_i}(C_k|BW)\right) \end{cases} \tag{10}$$

where the function $\prod\limits_{k \in N_i}(\Xi)$ represents the average value of parameter set Ξ in a certain range $\|Mean(\Xi)\cdot(1-\ell), Mean(\Xi)\cdot(1+\ell)\|$. Here, ℓ is 0.1.

Then, the multi-agent situation parameters after information sharing can be expressed as

$$\underset{i \in N_i}{X'} = C_{Balance} : \left(Fc', Pw', PRF', BW'\right) \tag{11}$$

where $C_{Balance}$ represents the balanced agent situation information and Fc', Pw', PRF', BW' represent the balanced jamming parameters' values.

Based on the mechanism of multi-agent parameter information sharing, dynamic situation information perception of the external electromagnetic environment is efficiently achieved.

2.3. Autonomous Decision of Jammer Agent Parameters

Through information sharing mechanisms, multi-agent jammers realize the collaborative perception of electromagnetic situation information in the whole region through cooperative means. At the same time, in order to make the jamming signal of the electronic jammer adapt to the new situation, it is necessary to control the jamming parameters dynamically and independently in real time.

Suppose that the parameter of electronic jamming agent i at time t is

$$\underset{i}{Par}^t{}_{\text{jam}} = {K_i}^t : \{\gamma\cdot(\mu_{Fc}, \mu_{Pw}, \mu_{PRF}, \mu_{BW})\} \tag{12}$$

where ${K_i}^t$ represents the total situation of the jamming parameters of agent i at moment t, $(\mu_{Fc}, \mu_{Pw}, \mu_{PRF}, \mu_{BW})$ represents the corresponding parameter values of (Fc, Pw, PRF, BW), and γ indicates the weight of each jamming parameter.

It is important to measure the rate of change between the current jamming parameters and the radar radiation source parameters. Understanding the current parameter variation trend is a key index for the subsequent jamming parameter set (Fc, Pw, PRF, BW) decision at time $t+1$. The transformation rate of the parameter set (Fc, Pw, PRF, BW) at a time between t and $t+1$ can be obtained by taking the partial derivative of parameter set (Fc, Pw, PRF, BW) of the function $\left(\underset{i}{Par}^{t}_{\text{jam}} - \underset{i}{X}^{t+1}\right)$. Therefore, the transformation rate $\nabla^{t,t+1}{}_{(Fc,Pw,PRF,BW)}$ can be expressed as

$$\begin{cases} \nabla^{t,t+1}{}_{Fc} = \dfrac{\partial\left(\underset{i}{Par}^{t}_{\text{jam}} - \underset{i}{X}^{t+1}\right)}{\partial Fc}, \\ \nabla^{t,t+1}{}_{Pw} = \dfrac{\partial\left(\underset{i}{Par}^{t}_{\text{jam}} - \underset{i}{X}^{t+1}\right)}{\partial Pw}, \\ \nabla^{t,t+1}{}_{PRF} = \dfrac{\partial\left(\underset{i}{Par}^{t}_{\text{jam}} - \underset{i}{X}^{t+1}\right)}{\partial PRF}, \\ \nabla^{t,t+1}{}_{BW} = \dfrac{\partial\left(\underset{i}{Par}^{t}_{\text{jam}} - \underset{i}{X}^{t+1}\right)}{\partial BW}. \end{cases} \tag{13}$$

where $\frac{\partial\left(\underset{i}{Par}^{t}_{\text{jam}} - \underset{i}{X}^{t+1}\right)}{\partial(Fc,Pw,PRF,BW)}$ is the derivative of (Fc, Pw, PRF, BW) with respect to the function $\left(\underset{i}{Par}^{t}_{\text{jam}} - \underset{i}{X}^{t+1}\right)$.

The trend of the situation parameters for the next moment is updated for the same agent at different times. Using the situational information sharing method, the next changing process of jamming parameter ${K_i}^{t+1}|(Fc, Pw, PRF, BW)$ is as follows:

$$\begin{cases} {K_i}^{t+1}|Fc = \mu_{Fc} + \varepsilon \cdot \nabla^{t,t+1}{}_{Fc}, \\ {K_i}^{t+1}|Pw = \mu_{Pw} + \varepsilon \cdot \nabla^{t,t+1}{}_{Pw}, \\ {K_i}^{t+1}|PRF = \mu_{PRF} + \varepsilon \cdot \nabla^{t,t+1}{}_{PRF}, \\ {K_i}^{t+1}|BW = \mu_{BW} + \varepsilon \cdot \nabla^{t,t+1}{}_{BW}. \end{cases} \tag{14}$$

where $(\mu_{Fc}, \mu_{Pw}, \mu_{PRF}, \mu_{BW})$ represents the corresponding parameter values of (Fc, Pw, PRF, BW) and ε is a constant with a value of 0.1.

The information about the electronic jamming parameters $\underset{i}{Par}^{t+1}_{\text{jam}}$ updated at moment $t+1$ can be expressed as follows:

$$\begin{aligned} \underset{i}{Par}^{t+1}_{\text{jam}} &= {K_i}^{t+1}\left(\left\{\begin{array}{c}\text{Jamming Paraments} \\ \text{of Agent } i\end{array}\right\}_{i \in Ni}\right) \\ &= {K_i}^{t+1} : (\mu'_{Fc}, \mu'_{Pw}, \mu'_{PRF}, \mu'_{BW}). \end{aligned} \tag{15}$$

where $(\mu'_{Fc}, \mu'_{Pw}, \mu'_{PRF}, \mu'_{BW})$ represents the weight of (Fc, Pw, PRF, BW) at moment $t+1$ for the ith electronic jammer agent.

3. Evaluation of Jamming Effectiveness

3.1. Jamming-to-Signal Ratio Definition

To effectively jam a radar signal, it is necessary to create a jamming signal at the target receiver for the radar signal. The jamming-to-signal ratio (*JSR*) is a good indicator for quantifying the jamming effect. The *JSR* is similar to the signal-to-noise ratio (*SNR*), where *J* represents the received jamming signal energy and *S* represents the received radar signal energy.

Consider a self-defense electronic jamming situation in which each target has a cognitive jammer. In the case of a radar and a jammer, the jammer focuses on optimizing the jamming efficacy by learning the radar's strategy in order to protect the target from detection [16]. To make the study easier, we consider the target to be a point target with

a radar cross-section (RCS) of ρ. We assume that the radar is jammed during each beam dwell time, which is referred to as a jamming round in this article. The number of radar pulses transmitted during a jamming round is determined by the beam dwell duration and the pulse repetition interval (*PRI*). For the nth radar pulse in a jamming round, the carrier frequency Fc is $Fc_r^{(n)}$, the bandwidth BW is $Bw_r^{(n)}$, the PRI is $pri_r^{(n)}$ which represents the time between the rising edge of the $(n-1)$th and nth radar pulses, the pulse width PW is $Pw_r^{(n)}$, and the transmission power is $P_r^{(n)}$. The jammer tries to align the jamming signal with the radar signal in both the time and frequency domains at each pulse. For the nth jammer pulse, the Fc is $f_{jam}^{(n)}$, the BW is $Bw_r^{(n)}$, the pulse delay time is $dt_{jam}^{(n)}$, which indicates the time between receiving the $(n-1)$th radar pulse and sending the next jamming pulse, the PW is $Pw_{jam}^{(n)}$, and the transmission power is $P_r^{(n)}$. Furthermore, the distance between the radar and the target is R and the wavelength of the radar signal is λ. The radar and jammer antenna gains are G_r and G_{jam}, respectively. The radar and jamming signal propagation losses are L_r and L_{jam}, respectively. The polarization matching loss coefficient between the jammer signal and the radar signal is ψ. The power of the echo at the radar receiver can be expressed as

$$P_{rs}^{(n)} = \frac{P_r^{(n)} G_r^2 \rho \lambda^2}{(4\pi)^3 D^4 L_r} \tag{16}$$

The power of the nth jamming pulse at the radar receiver is

$$P_{rj}^{(n)} = \frac{P_{jam}^{(n)} G_{jam} G_r \lambda^2 \psi}{(4\pi)^2 D^2 L_{jam}} \tag{17}$$

Using an effective jamming coefficient to modify the *JSR* calculation formula, the average *JSR* for the nth radar pulse is computed as follows:

$$JSR^{(n)} = \frac{P_{rj}^{(n)} \cdot X_f^{(n)} \cdot X_t^{(n)}}{P_{rs}^{(n)}} = \frac{P_j^{(n)} G_{jam} \mu 4\pi D^2 L_r X_f^{(n)} X_t^{(n)}}{P_r^{(n)} G_r \rho L_{jam}} \tag{18}$$

where $X_f^{(n)}$ and $X_t^{(n)}$ are the effective jamming factors in the frequency and time domains, expressed as

$$X_f^{(n)} = \frac{\Delta f^{(n)}}{B_j^{(n)}} \cdot \mathrm{sgn}\left(\Delta f^{(n)}\right) X_t^{(n)} = \frac{\Delta t^{(n)}}{Pw_j^{(n)}} \cdot \mathrm{sgn}\left(\Delta t^{(n)}\right) \tag{19}$$

where $\Delta f^{(n)}$ and $\Delta t^{(n)}$ are the overlapping rates in the frequency domain and the time domain, provided as follows:

$$\begin{aligned} \Delta f^{(n)} &= \min\left(Fc_j^{(n)} + Bw_j^{(n)}/2, Fc_r^{(n)} + Bw_r^{(n)}/2\right) - \max\left(Fc_j^{(n)} - Bw_j^{(n)}/2, Fc_r^{(n)} - Bw_r^{(n)}/2\right) \\ \Delta t^{(n)} &= \min\left(dt_j^{(n)} + Pw_j^{(n)}, PRI_r^{(n)} + Pw_r^{(n)}\right) - \max\left(dt_j^{(n)}, PRI_r^{(n)}\right) \end{aligned} \tag{20}$$

where $\mathrm{sgn}(x)$ can be written as

$$\mathrm{sgn}(x) = \begin{cases} 1, & x > 0 \\ -1, & \text{otherwise} \end{cases} \tag{21}$$

3.2. Q-Learning Method for Jamming Policy

The jamming process can be represented as a quaternion $\{\mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}\}$ through a finite Markov decision process (MDP) [17,18]. Here, $\mathcal{S}$ is a finite collection of radar states, with state $s \in \mathcal{S}$ determined by the radar pulse parameters (Fc, Pw, PRF, BW), and $\mathcal{A}$ is a finite set of jammer actions, where action $a \in \mathcal{A}$ is defined by the jamming pulse parameters.

When the jammer performs action $a^{(n)}$, the transition probability $\mathcal{P}\left(s^{(n+1)} \mid s^{(n)}, a^{(n)}\right)$ describes how the present state $s^{(n)}$ transitions to the next state $s^{(n+1)}$. In addition, $\mathcal{R}$ is the reward after each action is taken.

Reinforcement learning is an effective method for solving MDP problems, with the key being the determination of the optimal policy $\pi : \mathcal{S} \rightarrow \mathcal{A}$ to determine which action should be performed in each state. The state-value function for policy π is proposed to evaluate the effect of a policy, as follows:

$$v_\pi(s) = E_\pi\left[\sum_{m=0}^{\infty} \xi^m \mathcal{R}_m \mid s_m = s\right] \tag{22}$$

where $E_\pi[\cdot]$ stands for the expected value with the given policy π and $\xi \in [0, 1)$ is the discount rate of the reward R, which means that a long-term reward is considered and its influence decreases with time.

The update strategy of the electronic jammer parameters based on Q-learning [19–22] is as follows:

$$Q^*(s,a) = Q(s,a) + \eta \cdot (r(s,a) + \xi \cdot \max_{a'} Q(s',a') - Q(s,a)) \tag{23}$$

where s is the current jamming parameter status of the electronic jammer, a is the current action taken, $r(s,a)$ indicates the reward after the operation of action a, $\max_{a'} Q(s',a')$ is the best parameter adjustment action taken by the jammer agent at the next moment, η is the learning rate parameter, and ξ is the discount factor parameter. In continuous tasks, the discount factor ξ is usually set in the range [0,1) in order to ensure that the reward $r(s,a)$ does not diverge to infinity.

In our work, the jammer applies jamming to the radar by changing the signal characteristic parameters. The more realistic the electronic jamming signal is compared to the real echo signal, the better the interference effect. Here, a represents the action space of each of the jamming parameters (Fc, Pw, PRF, BW) and s represents the current state of each jamming parameters. At the beginning of moment t, when we perform action a on the jamming parameter (Fc, Pw, PRF, BW) the jammer obtains the current local state information as s and the current Q table as $Q(s,a)$. At the same time, it receives a reward $r(s,a)$ for executing action a. Obviously, the reward function is mainly used to reward expected behaviors and punish undesirable actions. Then, we use the $r(s,a) + \xi \cdot \max_{a'} Q(s',a')$ corresponding to action a' as the estimated value. Finally, $Q^*(s,a)$ is updated using Formula (23). The above operation is repeated until the jamming parameters (Fc, Pw, PRF, BW) reach the optimal state.

The dynamic cognitive electronic jamming strategy based on Q-learning can realize real-time perception based on the environmental electromagnetic situation and update the electronic jamming parameters of the system in real time according to the perception results, ensuring that the jamming strategy of the electronic jamming system always has the best possible status. The system flow is shown in Figure 4.

In Figure 4, the distributed electronic jammer can obtain spatial electromagnetic situation information among multiple agents through collaborative perception. The agent situation after information sharing is expressed as X'. The reconnaissance situation information at the current moment and the previous moment is input into the intelligent decision system. The optimal jamming strategy for the next epoch implemented by the electronic jamming machine is realized through the Q-learning decision algorithm.

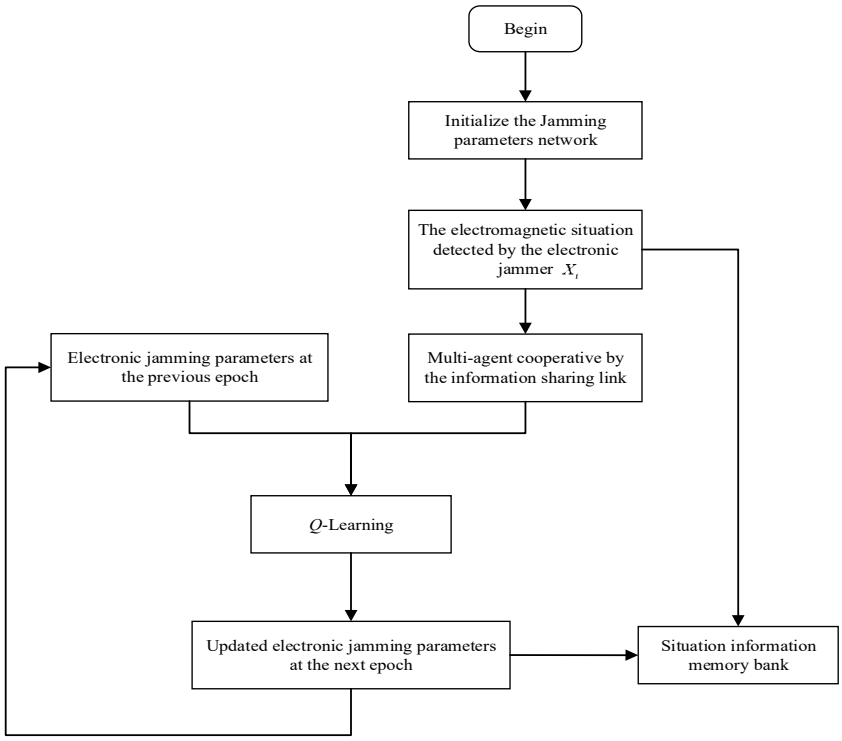


Figure 4. Electronic jamming flow chart based on multi-agent information sharing and *Q*-learning methods.

4. Experiment and Analysis

A distributed multi-agent cooperative electronic jamming system uses a process of continuous correction during the process of interaction with the external electromagnetic environment, in which the jamming parameters change dynamically. In the process of electronic confrontation, electronic jammer agents constantly learn from and evolve with the electromagnetic environment to achieve the optimal jamming effect.

In order to verify the effectiveness of this method, the following experiments were designed.

Table 1 shows the change trend of the electromagnetic situation information of the radiation sources in different stages. The period from T0 to T5 indicates the changes in the radiation source parameters at every 10-min interval. It can be seen from the table that the signal parameters emitted by the electromagnetic radiation source are changed in different stages, which can effectively avoid the problem of tracking brought about by the single parameters of the radiation source emission signal. The radiation source characteristics with parameter variations have a certain anti-jamming function. Electronic jammers need to perceive the change trend of the radiation source situation information quickly in a dynamic and complex electromagnetic environment. Moreover, the information is used as prior information to realize autonomous perception and a dynamic cognitive jamming strategy.

Table 1. Electromagnetic parameter information at different times.

Stage	Fc/GHz	Pw/us	PRF/Hz	BW/MHz
T0	3.25	22.5	5000	10
T1	5	60	10,000	10
T2	4.375	40	5000	20
T3	2.6	50	2500	20
T4	5.45	20	10,000	15
T5	6.25	35	5000	15

4.1. Situation Information Sharing

Based on the parameter situation information of the radiation source emission signal in Table 1, the electronic jammer is able to realize the electromagnetic parameter estimation and modulation type recognition of the radiation source through the passive electronic reconnaissance system.

In this experiment, six electronic jammer agents were distributed isometrically around the radiation source to realize the dynamic perception of the spatial global electromagnetic situation information. The signal-to-noise ratio of the environment was set to 0 dB, and the noise was assumed to be white Gaussian noise with a power density of one half. Because the situation parameter information between each agent is not completely consistent, it is necessary to balance the situation parameters obtained by the distributed electronic agents. The balanced jammer agent parameters can be expressed as $\underset{i \in N_i}{X'} = C_{\text{Balance}} : Pa(Fc', Pw', PRF', BW')$. The results are shown in Figure 5.

The different characteristic parameters (including Fc, Pw, PRF, BW) of the radiation source perceived by the six electronic jammers at different sampling times are shown in Figure 5. In this experiment, the signal-to-noise ratio was 0db. Among them, the transverse coordinate is the agent number, the longitudinal coordinate is the value of the corresponding characteristic parameters, the virtual line represents the theoretical parameter value, and the real line represents the actual measurement value after the balanced treatment.

The radiation source signal parameter estimation results based on information sharing are shown in Table 2.

Table 2. Measured values of electromagnetic parameters at different times.

Stage	Fc/GHz	Pw/us	PRI/us	μ/Hz/s
T0	3.2539	22.367	198.3	4.492×10^{11}
T1	4.9758	59.76	102.76	1.634×10^{11}
T2	4.3863	40.432	200.843	4.9317×10^{11}
T3	2.6036	50.372	399.346	3.9705×10^{11}
T4	5.4620	19.874	100.264	7.5475×10^{11}
T5	6.2510	35.006	201.149	4.285×10^{11}

The relationships between PRI and PRF can be written as

$$PRI = \frac{1}{PRF} \tag{24}$$

The relationships between μ, Pw, and Bw are as follows:

$$\mu = \frac{Bw}{Pw} \tag{25}$$

The multi-agent situation information sharing method can fully characterize the current electromagnetic situation. At the same time, the situation parameter imbalance caused by insufficient single jammer agent perception ability can be avoided.

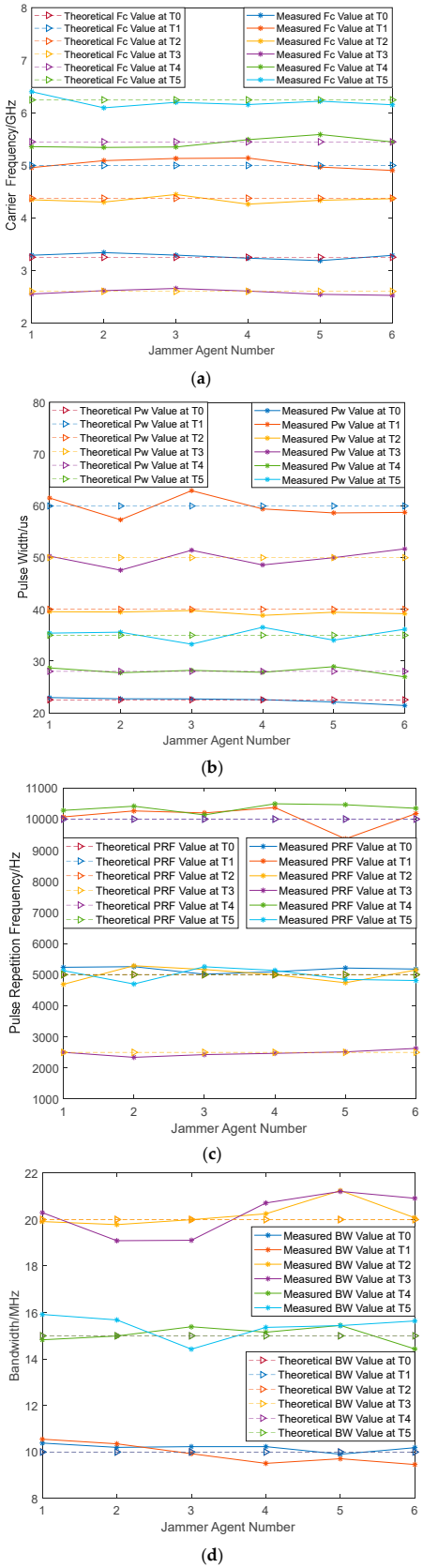


Figure 5. The multi-agent electronic jammer measures the electromagnetic situation information at different times. (**a**) The measurements and theoretical values of *Fc*; (**b**) the measurements and theoretical values of *PW*; (**c**) the measurements and theoretical values of *PRF*; (**d**) the measurements and theoretical values of *BW*.

4.2. Optimization of the Jamming Parameters in Real Time

The characteristic parameters X_i' of the balanced radiation source signal are input into the electronic jammer intelligence system as environmental perception information, and the real-time intelligent jamming strategy for the radiation source target is realized using the *Q*-learning method.

Hyperparameters are elements that must be set in *Q*-learning in order to adjust learning strategies. In this experiment, the learning rate η was set to 0.01 and the discount factor ξ was set to 0.9. The *Q* table of all electronic agents was initialized to 0 and the state s of each agent was randomly initialized.

Based on the shared electromagnetic situation information obtained in Section 4.1, the jamming parameters of the electronic jammers can change independently with the change in the radiation source parameters and achieve accurate and effective jamming of the radiation source. The reward function and the loss function based on the Q-learning method in a certain time period are shown in Figures 6 and 7.

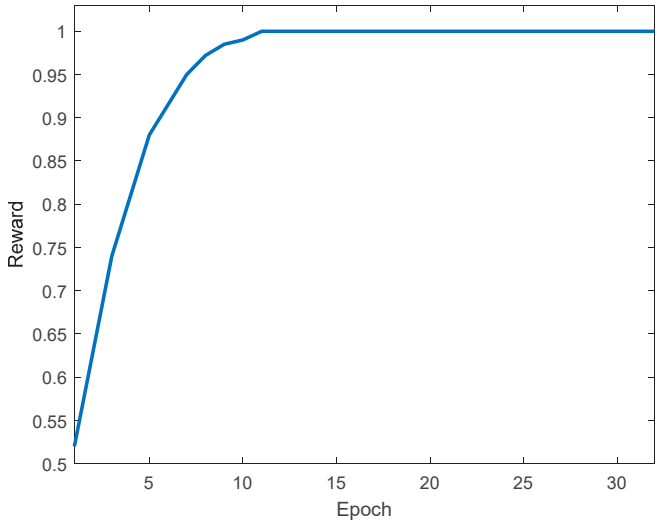


Figure 6. The reward–return curve within one period.

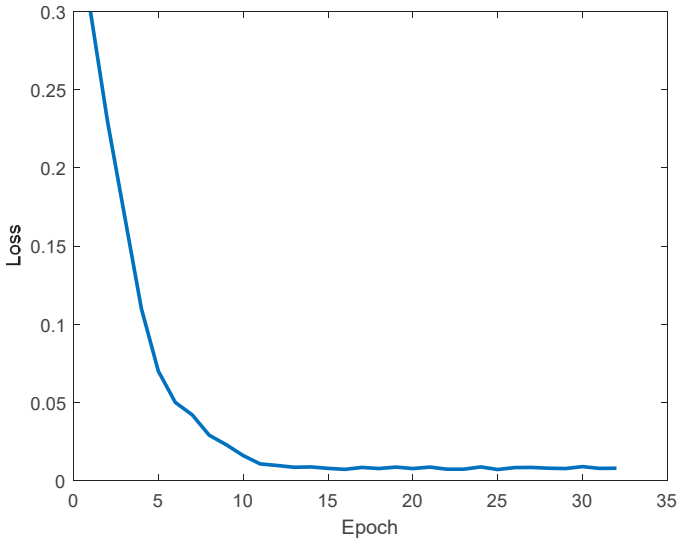


Figure 7. The loss function curve within one period.

Figures 6 and 7 show the reward–return function curve and loss function curve for one period. They demonstrate that the electronic jamming parameters achieve convergence in a

very short period of ten epochs. Moreover, the loss value reaches 0.0016 and the reward value reaches up to 0.9924, thereby realizing the rapid updating of the electronic jamming parameters.

Figure 8 shows a real-time jamming parameter response based on *Q*-learning with the radiation source situation parameters in Table 1 changed.

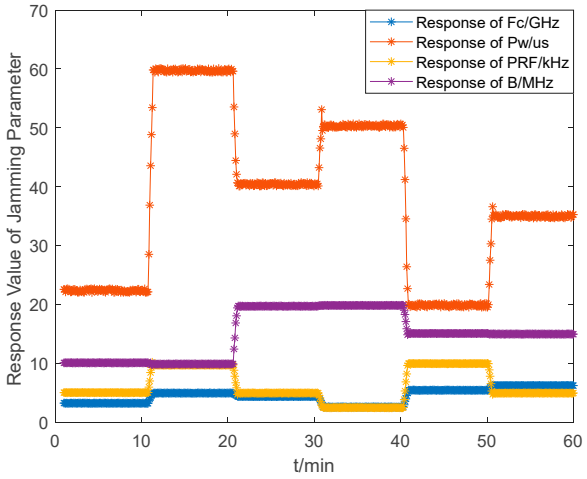


Figure 8. Dynamic response curve of the jamming parameter matrix over time.

As can be seen in Figure 8, the electronic jammers can quickly estimate the current electromagnetic situation and adjust the jamming parameters in the next epoch to adapt to the current electromagnetic environment when the radiation source changes.

4.3. Effect of the Number of Jammers on Perception Accuracy

Considering the influence of environmental factors, we set the reconnaissance error probability of the jamming parameters to 5%. The ability of different numbers of electronic jammers to perceive the electromagnetic situation information was analyzed after 1000 independent Monte Carlo experiments.

The accuracy result for different jammers is shown in Figure 9.

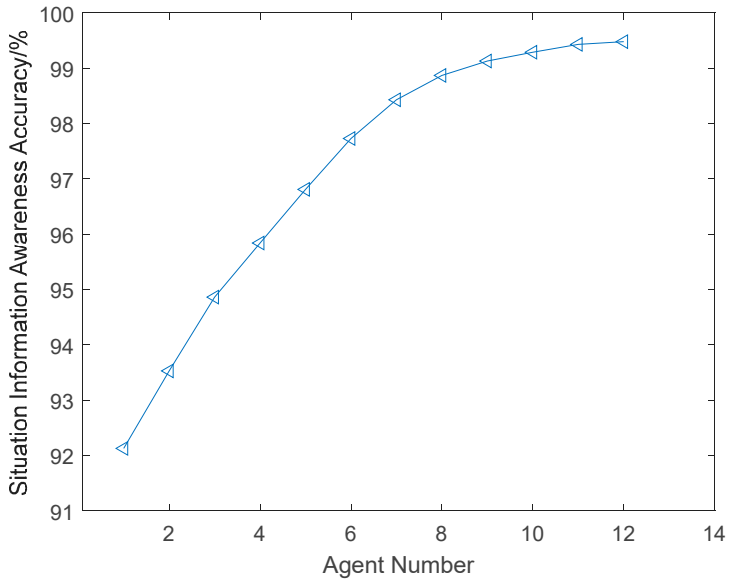


Figure 9. The accuracy of situational information awareness trends with the number of distributed jammers.

In Figure 9, the accuracy increases as the number of electronic jamming agents grows. The accuracy of the situational information $P(X)$ perceived by the distributed jamming agents can be expressed as follows:

$$P(X) = \sum_{l=1}^{N} \frac{P(X_l|_{Fc}) + P(X_l|_{Pw}) + P(X_l|_{PRF}) + P(X_l|_{BW})}{4} \tag{26}$$

where $P\left(X_l\Big|_{(Fc,Pw,PRF,BW)}\right)$ represents the accuracy of the parameters (Fc, Pw, PRF, BW) in terms of the situation information X_l.

The estimation error $\Re\left(Par_i\right)$ of the parameter set (Fc, Pw, PRF, BW) can be expressed as follows:

$$\Re\left(Par_i\right) = \frac{\sum_{n=1}^{1000} \left\| Par_{\text{i\_real}} - Par_{\text{i\_estimate}} \right\|_2}{\frac{1000}{Par_{\text{i\_real}}}} \times 100\% \tag{27}$$

where $Par_{\text{i\_real}}$ is the theoretical value and $Par_{\text{i\_estimate}}$ is the actual measurement value of the parameter set (Fc, Pw, PRF, BW).

Compared with the perception of radiation source information by a single agent, multiple agents are able to obtain a more accurate situational awareness of the radiation sources. When the number of electronic jamming agents reaches seven, the recognition rate can reach 98.43%.

4.4. Analysis of Jamming Effectiveness Evaluation

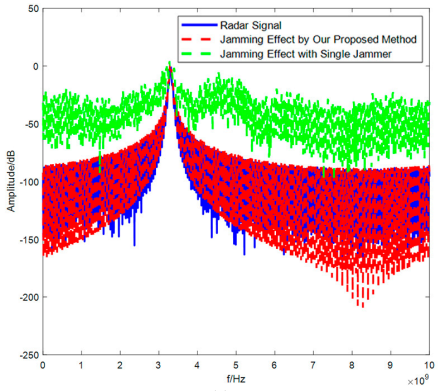
In order to verify the effectiveness of our proposed jamming method, range gate pull-off (RGPO) jamming was carried out on the target. The distance between the radar and the target was 100 km, and the range gate for the jamming target was 10 km. Other parameters in our simulation were as follows: G_r = 20 dB, G_{jam} = 5 dB, L_r = 10 dB, L_{jam} = 7 dB, ρ = 1 m^2. Considering that radar antennas are typically linearly polarized, whereas jammer antennas are circularly or tangentially polarized, γ was set to 0.25, κ was set to 0.8, and ε was set to 0.1. The other radar parameters were set as shown in Table 1 in different stages. In particular, six electronic jamming agents were distributed in the system.

We compared the performance of our algorithm with six jammers based on a distributed information sharing method and jamming with a single jammer. Figure 10 shows the jamming effect of the two methods from T0 to T5.

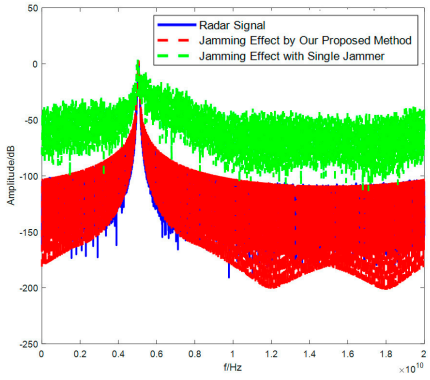
In Figure 10, compared to the jamming method with a single jammer, the jamming signal spectrum of our proposed method is closer to the real radar signal. Additionally, the sidelobe power of the jamming signal taken by our proposed method is lower than that taken by the method with a single jammer, which reduces the threat of the jamming signal, thereby improving the survivability of the jammers.

According to the definition of the *JSR* in Section 3.1, we calculated the *JSR* value of the jamming signal of the two methods in different stages.

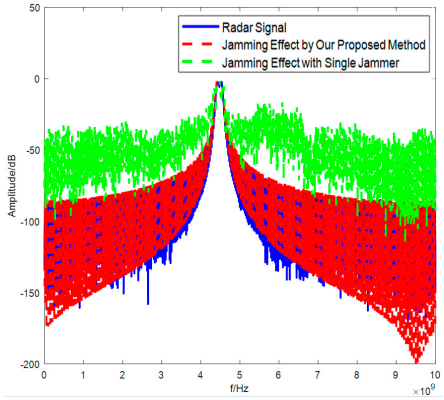
In Table 3, it can be seen that the *JSR* value of the jamming signal generated by our method is smaller than that generated by the jamming method with a single jammer in each stage. This result further validates the effectiveness of our method. Table 4 shows the feasibility verification of our method for the jamming parameter set (Fc, Pw, PRF, BW).



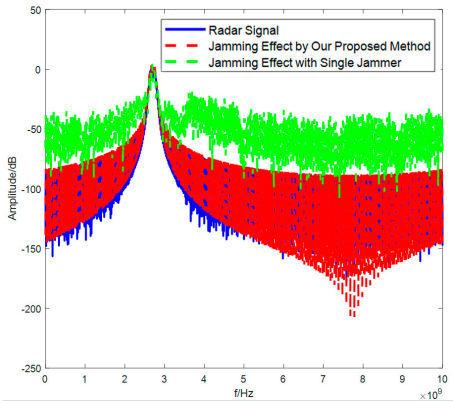
(**a**)



(**b**)



(**c**)



(**d**)

Figure 10. *Cont.*

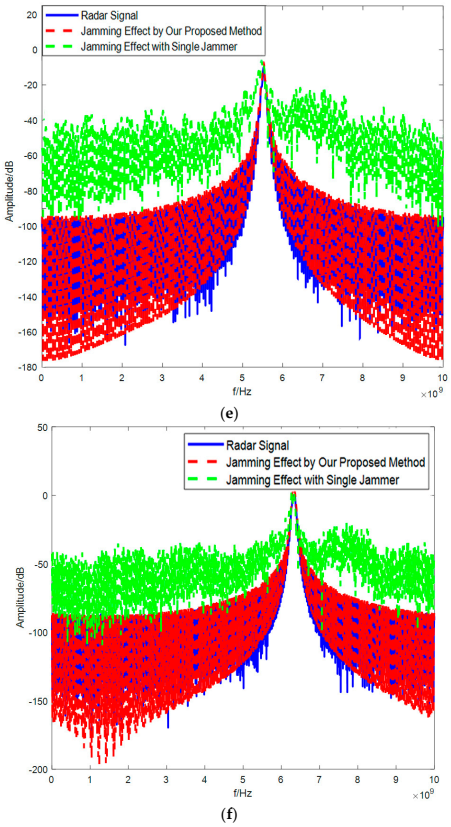


Figure 10. The jamming effect performance of our algorithm with six jammers based on the distributed information sharing method compared to jamming with a single jammer. (**a**) Comparison of the jamming effects of the two methods in stage T0. (**b**) Comparison of the jamming effects of the two methods in stage T1. (**c**) Comparison of the jamming effects of the two methods in stage T2. (**d**) Comparison of the jamming effects of the two methods in stage T3. (**e**) Comparison of the jamming effects of the two methods in stage T4. (**f**) Comparison of the jamming effects of the two methods in stage T5.

Table 3. Jamming signal *JSR* value of the two methods in different stages.

Method	*JSR*					
	T0	**T1**	**T2**	**T3**	**T4**	**T5**
Our algorithm	1.0215	0.9987	1.0032	1.0386	1.0108	1.1250
Jamming with a single jammer	1.6732	1.7851	1.5738	1.6210	1.8327	1.5248

Table 4. The feasibility verification of our method for the jamming parameter set (Fc, Pw, PRF, BW).

NO.	Evaluation Indicator	Correlation with Evaluation Result
1	Fc	positive
2	Pw	positive
3	PRF	positive
4	BW	positive

5. Conclusions

In this article, we have proposed a new electronic jamming method based on a distributed information sharing mechanism. This method, which is based on Q-learning,

can realize the real-time perception of the external electromagnetic environment and the autonomous updating of the jamming parameters. At the same time, it can realize a fast and accurate response from the jamming system and improve the concealment of jamming signals. Compared with the jamming method with a single jammer, the proposed distributed multi-agent electronic jamming method based on information sharing can achieve effective jamming with a lower *JSR*. Therefore, the concealment of the jammer system is improved, and the survival probability of the jammer is enhanced. Moreover, the proposed method provides practical advantages for intelligent electronic jamming systems and can serves as inspiration for future cognitive electronic jammers.

Author Contributions: Conceptualization, Z.J. and P.Z.; methodology, P.Z.; software simulation, P.Z. and Y.H.; data processing and analysis, P.Z. and Y.H.; conclusion analysis, Y.H.; writing-original draft preparation, P.Z.; writing-review and editing, Z.J.; funding acquisition, Y.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the NSFC under Grant No. U19A2054 and the China National Funds for Distinguished Young Scientists under Grant No. 61525403. This research was also supported by the Zhejiang International Youth Talent Fund of Zhejiang Lab.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Cywiński, A.; Ostrowski, R.; Strzelec, M. Electronic warfare in the optical band: Main features, examples and selected measurement data. *Def. Technol.* **2020**, *17*, 1636–1649. [CrossRef]
2. Smith, C.R.; Grasso, R.; Pledger, J.; Murarka, N. Trends in electro-optical electronic warfare. *Proc. SPIE* **2012**, *8543*, 854392. [CrossRef]
3. Zhou, Y.; Rao, B.; Wang, W. UAV swarm intelligence: Recent advances and future trends. *IEEE Access* **2020**, *8*, 183856–183878. [CrossRef]
4. Zhou, L.; Leng, S.; Liu, Q.; Wang, Q. Intelligent UAV swarm cooperation for multiple targets tracking. *IEEE Internet Things J.* **2021**, *9*, 743–754. [CrossRef]
5. Wang, X.; Huang, T.; Liu, Y. Resource allocation for random selection of distributed jammer towards multistatic radar system. *IEEE Access* **2021**, *9*, 29048–29055. [CrossRef]
6. Zhou, Z.; Rao, B.; Xie, X. The influence mechanism of UAV group on the detection performance of air defense radar. In Proceedings of the 2018 3rd International Conference on Automation, Mechanical Control and Computational Engineering (AMCCE 2018), Dalian, China, 12–13 May 2018; Atlantis Press: Beijing, China, 2018; pp. 338–343.
7. DARPA. Behavior Learning for Adaptive Electronic Warfare. Available online: https://www.fbo.gov (accessed on 6 October 2010).
8. Kingsley, N.; Guerci, J.R. Adaptive amplifier module technique to support cognitive RF architectures. In Proceedings of the IEEE Radar Conference, Cincinnati, OH, USA, 19–23 May 2014; pp. 1329–1332.
9. Fu, J.; Wan, Y.; Wen, G.; Huang, T. Distributed robust global containment control of second-order multiagent systems with input saturation. *IEEE Trans. Control. Netw. Syst.* **2019**, *6*, 1426–1437. [CrossRef]
10. Yan, Y.; Huang, J. Cooperative output regulation of discretetime linear time-delay multi-agent systems. *IET Control. Theory Appl.* **2016**, *10*, 2019–2026. [CrossRef]
11. Wang, S.; Bao, Y.; Li, Y. The architecture and technology of cognitive electronic warfare. *Sci. Sin. Inf.* **2018**, *48*, 1603–1613. [CrossRef]
12. Xing, Q.; Zhu, W.G.; Jia, X. Intelligent radar countermeasure based on Q-learning. *Syst. Eng. Electron.* **2018**, *40*, 1031–1035.
13. He, J.-H. *FPGA Software Design for Reconnaissance and Jamming Integration Processor*; Harbin Engineering University: Harbin, China, 2017.
14. Zuo, S.; Song, Y.; Lewis, F.L.; Davoudi, A. Output containment control of linear heterogeneous multi-agent systems using internal model principle. *IEEE Trans. Cybern.* **2017**, *47*, 2099–2109. [CrossRef] [PubMed]
15. Li, Z.; Ren, W.; Liu, X.; Fu, M. Distributed containment control of multi-agent systems with general linear dynamics in the presence of multiple leaders. *Int. J. Robust Nonlinear Control.* **2013**, *23*, 534–547. [CrossRef]
16. Osner, N.R.; du Plessis, W.P. Threat evaluation and jamming allocation. *IET Radar Sonar Navig.* **2017**, *11*, 459–465. [CrossRef]
17. Rabiner, L.R. A tutorial on hidden markov models and selected applications in speech recognition. *Proc. IEEE* **1989**, *77*, 257–286. [CrossRef]
18. Han, L.; Ning, Q.; Chen, B.; Lei, Y.; Zhou, X. Ground threat evaluation and jamming allocation model with markov chain for aircraft. *IET Radar Sonar Navig.* **2020**, *14*, 1039–1045. [CrossRef]

19. Feng, H.Z.; Liu, H.W.; Yan, J.K.; Dai, F.Z.; Fang, M. A fast efficient power allocation algorithm for target localization in cognitive distributed multiple radar systems. *Signal Process.* **2016**, *127*, 100–116. [CrossRef]
20. Bui, V.H.; Nguyen, T.T.; Kim, H.M. Distributed operation of wind farm for maximizing output power: A multi-agent deep reinforcement learning approach. *Access IEEE* **2020**, *8*, 173136–173146. [CrossRef]
21. Zhang, Y.; Peng, L.; Xu, R.; Li, J. A distributed low-redundancy information sharing algorithm in ad hoc networks with directional antennas. *Procedia Comput. Ence* **2018**, *131*, 1142–1149. [CrossRef]
22. Peng, T.; Leckie, C.; Ramamohanarao, K. Information sharing for distributed intrusion detection systems. *J. Netw. Comput. Appl.* **2007**, *30*, 877–899. [CrossRef]