electronics

Article

Design of Multi-User Collaborative Anti-Jamming System Under Sensing Heterogeneity

Shiqi Gao [1,2], Yingxin Huang [1], Nan Qi [1,2], Xiaonan Mu [1], Luliang Jia [3] and Daolong Wu [2,*]

[1] College of Electronic Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China; luguoshanye@nuaa.edu.cn (S.G.); huangyingxin@nuaa.edu.cn (Y.H.); qinan@nuaa.edu.cn (N.Q.); muxiaonan@nuaa.edu.cn (X.M.)
[2] National Key Laboratory of Multi-domain Data Collaborative Processing and Control, The 20th Research Institute of China Electronics Technology Group Corporation, Xi'an 710068, China
[3] School of Space Information, Space Engineering University, Beijing 101416, China; jiall@hgd.edu.cn
* Correspondence: rain1985_wu@163.com

Abstract

Dynamic spectrum access enables efficient anti-jamming in cognitive radio systems. However, in a multi-user distributed decision scenario, differences in spectrum states make collaboration among users a major challenge, especially when the sensing devices are heterogeneous. In order to solve this issue, we propose a collaborative anti-jamming cognitive radio system architecture based on historical jamming knowledge. Devices exhibiting high sensing performance support those exhibiting low sensing performance. An online reinforcement learning algorithm is used to learn the jamming patterns in real time. Finally, a multi-user collaborative anti-jamming system is developed using a software-defined radio platform. The anti-jamming performance of the system is verified experimentally under both internal communication jamming and external malicious jamming scenarios, achieving a jamming probability below 0.1.

Keywords: collaborative anti-jamming; device heterogeneity; cognitive radio

check for updates

Academic Editor: Hector E. Nistazakis

Received: 19 September 2025
Revised: 28 October 2025
Accepted: 29 October 2025
Published: 30 October 2025

Citation: Gao, S.; Huang, Y.; Qi, N.; Mu, X.; Jia, L.; Wu, D. Design of Multi-User Collaborative Anti-Jamming System Under Sensing Heterogeneity. *Electronics* **2025**, *14*, 4264. https://doi.org/10.3390/electronics14214264

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

With the rapid development of the Internet of Things (IoT), the proliferation of wireless communication devices has significantly intensified the competition for limited spectrum resources, making efficient and intelligent spectrum management increasingly critical. However, as communication systems become more intelligent, they also face more serious threats, among which malicious jamming has emerged as a serious concern. Such attacks can severely degrade system performance by disrupting decision mechanisms in cognitive radio (CR) systems [1]. This has led to an urgent need for robust anti-jamming solutions.

Spectrum sensing and dynamic spectrum access (DSA) are effective methods in CR systems for solving spectrum jamming problems. Spectrum sensing can be used to identify jamming bands and idle bands [2]. With DSA, spectrum holes that occur during communication can be opportunistically utilized, thereby reducing the impact of jamming [3]. The emergence of software-defined radio (SDR) platforms has enabled real-time adjustment of communication frequency bands. A representative example is GNU Radio, which can directly control parameters of the universal software radio peripheral (USRP) via the USRP hardware driver (UHD) [4]. Consequently, the SDR platform enables the demonstration and verification of DSA for anti-jamming. Reinforcement learning (RL) has better adaptability and enables more efficient decision making in dynamic environments. Numerous

RL algorithms have been applied to anti-jamming in dynamic spectrum access (DSA) [5]. Since the jamming environment faced by each user is similar, sharing learning experiences can improve the overall anti-jamming capability of the system [6]. In [7], a collaborative dynamic spectrum access approach based on the Q-learning algorithm was proposed, emphasizing the role of collaboration. The method demonstrated a superior performance over conventional overlay and underlay strategies in terms of throughput, jamming resistance, and reliability. A CR anti-jamming platform integrating RIS-based signal modulation and the Dyna-Q algorithm for DSA was proposed in [8]. It achieved more than a 50% improvement in throughput under dynamic jamming scenarios. In [9], an anti-jamming algorithm based on the Double Deep Q-Network (DDQN) was proposed. The algorithm does not require prior information about jamming but uses wideband spectrum sensing to obtain real-time channel states, and the superiority of the algorithm is verified under various jamming modes. Another study [10] integrated deep reinforcement learning with transfer learning, aiming to effectively resist jamming attacks in wireless networks. However, in solving the complex multi-user anti-jamming problem, few papers have considered cases such as an inconsistent spectrum state and the heterogeneity of sensing devices.

Collaboration can be categorized into two types: collaborative sensing and collaborative decision making. However, collaborative decision making is often ineffective in scenarios where users experience inconsistent spectrum states. Even when adopting collaborative sensing, precise time synchronization is critical, as timing discrepancies during data acquisition can lead to misaligned observations across nodes. As highlighted in [11], such synchronization errors—though small—can undermine the consistency of shared sensing data, especially in systems with heterogeneous devices and variable sensing delays. This can result in temporal misalignment of spectrum sensing results among users, even when operating within the same decision interval. This work adopts a collaborative sensing approach, where users perform distributed and autonomous decision making. Compared to other works on multi-user collaborative anti-jamming, the AND-rule, the OR-rule, or the K/M-rule is not adopted in this system. This is because users with limited sensing capabilities generate wideband spectrum data with a certain degree of delay. Considering the issue of the inconsistent spectrum state faced by these heterogeneous sensing devices, a collaborative spectrum estimation method based on historical jamming knowledge is proposed. The corresponding multi-user collaborative anti-jamming strategy is studied. Online Q-learning is embedded into the anti-jamming algorithm of each subsystem. To mitigate the adverse impact of sensing delays on learning accuracy, a two-slot reward assignment mechanism is introduced which extends the reward feedback window, thereby reducing the negative effects caused by temporal misalignment. Finally, by collaboratively sharing learning experiences and sensing data, a multi-user collaborative anti-jamming platform under heterogeneous sensing devices is developed. The effectiveness of the system's anti-jamming capabilities is validated in a scenario involving both internal jamming and external malicious jamming.

2. System Architecture for Multi-User Collaborative Anti-Jamming

A multi-user communication scenario is illustrated in Figure 1a, which includes a primary user (PU) communication pair, a secondary user (SU) communication pair, and multiple malicious jamming sources. SUs need to connect to the system without blocking communication between PUs. Therefore, PUs and SUs must solve external malicious jamming and internal communication jamming between users.

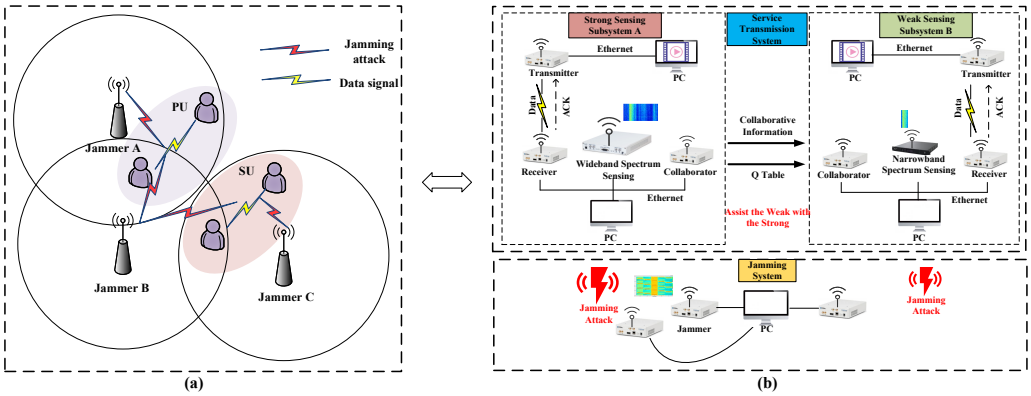


Figure 1. Multi-user collaborative anti-jamming system. (**a**) Communication scenario; (**b**) system architecture.

Unlike existing research on multi-user collaborative anti-jamming, we introduce device heterogeneity in the context of a heterogeneous spectrum state. Specifically, there are differences in the performance of sensing devices between PUs and SUs. PUs perform wideband spectrum sensing across the entire bandwidth B_w to detect the spectrum state of all channels. In contrast, SUs rely on narrowband sensing and are limited to observing the spectrum state only within their own communication channel with bandwidth B_{sub}. The conflict between the jammers and users is modeled as a dynamic game process, and the system architecture of Figure 1b is built for this scenario. It is mainly composed of a service transmission system and a jamming system. The service transmission system is divided into a strong sensing subsystem A and a weak sensing subsystem B.

Orthogonal frequency-division multiplexing (OFDM) effectively mitigates the inter-symbol interference (ISI) caused by the delayed spread of wireless channels [12]. Therefore, it has been widely used in many wireless systems. Specifically, we first convert the original file to be transmitted into a bit stream. For services with large data volumes (such as pictures, videos, etc.), it is necessary to split the data bit stream and complete the preliminary data encapsulation by adding frame headers and CRC checksums. Next, we perform symbol mapping (such as QPSK) on each encapsulated data stream and convert it into multiple low-speed data streams through serial-to-parallel conversion. Then, the frequency domain signal is converted into the time domain through IFFT, and a cyclic prefix is added to each OFDM symbol to form an OFDM-modulated signal. Finally, the signal is up-converted into the channel f_m and sent with a USRP [13]. The communication link employs OFDM, with data being transmitted over the available channel set $\mathcal{M} = \{f_1, \ldots, f_M\}$. An acknowledgment (ACK) control link is established to ensure consistency between the receiver's and the transmitter's decisions. Specifically, in each time slot k, the receiver uses the spectrum state to decide the communication channel $f_m(k)$ of the next time slot $k+1$. Then, the decision information is transmitted back to the transmitter through the ACK control link to switch to the same frequency. The ACK control link is set to Gaussian Minimum Shift Keying (GMSK) modulation.

To enhance the anti-jamming capability of SUs and facilitate collaboration between PUs and SUs, a collaborative link is established between them, which is configured with GMSK modulation. The collaborative information transmitted may include the PUs' action strategies, learning experiences, and sensing data when necessary. Traditional jammers follow preset jamming rules, such as frequency sweeping and a fixed frequency. Intelligent jammers possess spectrum sensing capabilities and certain jamming strategies. To simulate various types of malicious jammers, we set up a traditional jammer and an intelligent jammer. The jammers are distributed across different locations to create inconsistencies in

the spectrum state. The traditional jammer generates jamming based on a probability matrix. The intelligent jammer is equipped with spectrum sensing performance and generates jamming of the channel after detecting continuous data signals. When the energy emitted by a jammer is concentrated on a single channel for an extended period, it increases the likelihood of being detected. Therefore, to ensure the security of the intelligent jammer, it initially remains silent and performs only spectrum sensing. The jammer initiates jamming on that channel when it detects that the channel utilization exceeds a certain threshold τ. The jammer is implemented using a USRP connected to a PC. All digital signal processing in this paper is implemented using the GNU Radio.

3. Algorithm Design for Anti-Jamming

In the communication scenario considered in this paper, SUs lack wideband spectrum sensing capabilities, and the spectrum states of the PUs and SUs may differ. Under such conditions, collaborative sensing between PUs and SUs is required to achieve collaborative anti-jamming. To address this, a collaborative spectrum estimation method is proposed in Section 3.1 to enable the PUs to assist SUs in spectrum sensing. Subsequently, in Section 3.2, an anti-jamming algorithm is designed by leveraging the data obtained from wideband spectrum sensing in conjunction with an online Q-learning algorithm. Finally, in Section 3.3, a multi-user collaborative anti-jamming algorithm is developed, which integrates intra-subsystem anti-jamming with inter-subsystem collaboration.

3.1. Collaborative Spectrum Estimation Method

Since PUs have a wideband spectrum sensing performance, they can obtain the spectrum states more accurately. Therefore, the purpose of this subsection is to study how SUs can accurately and promptly restore the spectrum states under the condition that the spectrum states are different and the sensing ability is limited. It is widely accepted that anti-jamming algorithms lacking wideband spectrum sensing capabilities significantly underperform compared to those with such capabilities. This performance gap is particularly evident in terms of the convergence time and effectiveness when the system operates under relatively complex jamming scenarios [14]. When the sensing ability is limited, one approach to obtaining a complete spectrum state is to use sequential sensing, which involves step-by-step scanning and restoration. However, this method can introduce significant sensing delays. Another approach is to use compressive sensing technology, which can reconstruct signals at lower sampling rates. However, its performance is constrained by the assumption of signal sparsity and faces challenges such as high complexity in reconstruction algorithms and difficulty in hardware implementation. To fully leverage the advantage of accelerated convergence provided by wideband spectrum sensing, this subsection proposes a specialized collaborative sensing method tailored to the considered scenario. The goal of this method is to enable rapid and accurate restoration of the spectrum states of the SUs.

According to the SUs' possible locations, there are various scenarios regarding the spectrum states. In Figure 1a, assuming that the SUs appear in the non-jammer C's range, the worst spectrum state under malicious jamming is the same as that for the PUs. In this scenario, the PUs can learn two sets of orthogonal action strategies based on their spectrum states. The PUs can then pass another action to the SUs. When the SUs are within the jammer C's range, the spectrum states faced by the PUs and the SUs become inconsistent. Therefore, the SUs cannot successfully avoid jamming by directly using the PUs' actions or sensing results. The SUs must possess a learning performance which aligns with the current mainstream approach of distributed multi-user collaborative anti-jamming [15]. However, the key distinction lies in our consideration and analysis of device heterogeneity.

The strategy of exchanging space for time is adopted, and sequential sensing is employed exclusively during the construction of the historical jamming knowledge, not in other operational phases. Based on the historical jamming knowledge, we employ energy detection to estimate the impact of jamming on each channel and convert these estimates into an average power vector $\overline{P} = \{\overline{p}_1, \ldots, \overline{p}_M\}$. This average power is then used to estimate the jamming power when jamming occurs in each respective channel. Even if there is a delay in detecting jamming, as long as the characteristics of the jamming (such as power, bandwidth, and center frequency) are accurate, we can accurately reflect the average jamming situation for that channel. We consider that for a jammer close to the SUs, if its influence range does not include the PUs, the demodulation threshold can still be met under such jamming conditions. Therefore, although the jamming is not recognized as effective by the PUs, we can still determine the presence of jamming by examining the jamming power $P^A = \{p_1^A, \ldots, p_M^A\}$ transmitted from the PUs. Once it is confirmed that there is jamming in the channel f_m, $\overline{p}_m$ can be directly extracted from $\overline{P}$ to estimate the impact of jamming on the SUs.

3.2. Anti-Jamming Algorithm Design Based on Wideband Spectrum Sensing

We create a game process between jamming and communication as a Markov Decision Process (MDP). Specifically, in each time slot k, the receiver needs to decide on an available channel for communication. Therefore, we define the state $s(k) = f_j(k)$ to express the receiver's sensing of the spectrum state, where $f_j(k)$ is the channel j occupied by the jamming signal. The action $a(k) = m$ expresses the channel number corresponding to the channel f_m selected by the receiver from the set of available channels $\mathcal{M}$. RL algorithms have been studied in the field of anti-jamming. Commonly used algorithms include Q-learning, Deep Q-Network (DQN), Deep Deterministic Policy Gradient (DDPG), and Soft Actor–Critic (SAC). Each RL algorithm has its strengths and limitations. Since the actions and states in this paper are discrete, we consider selecting between DQN and Q-learning algorithms. However, DQN involves significantly more parameters than the Q-table used in Q-learning. Meanwhile, to ensure the accuracy of DQN's network parameters, it is necessary to design and implement a specific transmission protocol, which increases the system's complexity and transmission delay. Therefore, the system platform developed in this paper uses the online Q-learning algorithm.

$$Q(s,a) = Q(s,a) + \alpha[r + \gamma \max_{a'} Q(s',a') - Q(s,a)], \tag{1}$$

where α is the learning rate, γ is the discount rate, r is the reward of action a taken by the user in state s, s' is the next state transitioned to from state s, and a' is the action selected in the next state s'.

Under the actual communication–jamming game relationship, the communicating party, as a passive adjuster, often only has access to the time slot situation of its own algorithmic decision. However, the jamming switching is stochastic, which leads to it being difficult to maintain consistency in the jamming switching and anti-jamming algorithm decision on the time slot. At this point, the method proposed in [8] possesses certain limitations; in this paper, when the jamming switching is aligned with the decision time slot of the anti-jamming algorithm, the reward value r can be based on the spectrum semantic value obtained from the sensing of a time slot. At this time, the anti-jamming algorithm proposed in this paper has a good performance. However, as shown in Figure 2, since the reward is obtained based on the sensing at t_2, using this method only guarantees avoidance of the jamming state corresponding f_1 at t_1, while for $t_0 \sim t_1$, there is a certain possibility of being jammed; this is because the jamming starts switching to f_1 only at t_1, while at $t_0 \sim t_1$, the jamming is still in f_0. When the throughput at the end of each time

slot is used as the reward value r for feedback calculation, the reliability of this parameter cannot be guaranteed in the actual platform implementation because the transmitter is not always transmitting data and due to the inconsistency of frequency switching or feedback delay. In this regard, without loss of generality, for this paper based on the time slot situation in Figure 2, the reward r is calculated based on the channel conditions of two time slots.

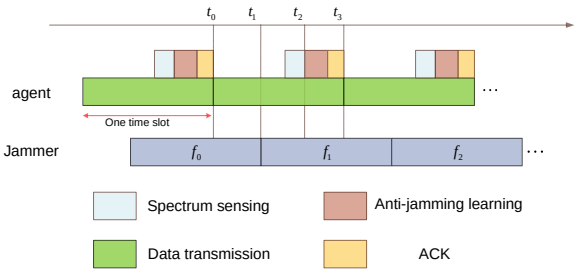


Figure 2. An illustration of the transmission slot structure.

$$r(k) = \begin{cases} 1 & \text{if } a(k) \neq f_j(k) \text{ and } a(k) \neq f_j(k+1) \\ 0 & \text{otherwise} \end{cases}. \tag{2}$$

where $r(k)$ is the reward obtained at time slot k, $a(k)$ is the action taken at time slot k, $f_j(k)$ denotes the channel occupied by the jammer at time slot k, and $f_j(k+1)$ denotes the channel occupied by the jammer at time slot $k+1$.

Thanks to wideband spectrum sensing, the system can obtain jamming conditions on different channels. In each time slot, it will evaluate whether each action is affected by jamming according to the spectrum sensing results to obtain the instantaneous return reward vector $\mathbf{r}$ of all actions. Since this process does not involve actual communication, it is referred to as virtual learning. This not only accelerates the algorithm's convergence but also forms an available set of actions $A(k) \subseteq \mathcal{M}$. We continuously record the action selected by the receiver, where $\mathbf{x} = (x_1, x_2, \ldots, x_M)$ denotes the vector that records the number of times each action $m \in \mathcal{M}$ is selected over a time horizon $k = 1, 2, \ldots, N$, where $x_m = \sum_{k=1}^{N} \mathbb{I}(a(k) = m)$ and $\mathcal{M} = \{1, 2, \ldots, M\}$. $\mathbb{I}(\cdot)$ is the indicator function, which returns 1 when the condition inside the parentheses is true, and 0 otherwise. Then, we calculate the cumulative probability $\theta_k(a)$ of each action and assign weights $\mu_k(a)$ to the elements in $A(k)$. Finally, the action interval is assigned according to the weights $\mu_k(a)$, and then, the specific action $a(k) = \text{SampleAction}(A(k), \mu_k)$ is sampled from this interval by using a random function.

$$\theta_k(a) = \frac{x_a}{\sum_{a \in \mathcal{M}} x_a}, \quad \forall a \in \mathcal{M}. \tag{3}$$

$$\mu_k(a) = \frac{1 - \theta_k(a)}{\sum_{a \in A(k)} (1 - \theta_k(a))}, \quad \forall a \in A(k). \tag{4}$$

3.3. Algorithm Design for Multi-User Collaborative Anti-Jamming

As shown in Section 3.1, the spectrum states exhibit common characteristics, which makes the anti-jamming experience learned by the PUs beneficial for the SUs. Therefore, before the SUs access the system, they first use the collaborative spectrum estimation method proposed in Section 3.1 to build a historical jamming knowledge base. After the SUs access the system, they first acquire learning experiences from the PUs through the collaborative link. In each time slot, the PUs transmit the spectrum state and action to the SUs to assist in achieving wideband spectrum estimation. The SUs then use the anti-jamming algorithm described in Section 3.2 for autonomous learning. We combine

the previous components and present a complete multi-user collaborative anti-jamming algorithm, which is shown in Algorithm 1.

Algorithm 1: Multi-user collaborative anti-jamming.

Initialize parameters α, β, γ, $Q^A(s^A, a^A)$
for $t = 0, 1, 2, \ldots, K-1$ **do**
  $s'^A \leftarrow$ wideband spectrum sensing
  Observe resultant reward **r**
  $a^A \leftarrow \text{SampleAction}(A(k), \mu_k)$
  Update Q^A
  Send an ACK message to the transmitter
  if $t < l$ **then**
    $\overline{P} \leftarrow$ historical jamming knowledge
  end if
  if $t \geq l$ **then**
    if $t = l$ **then**
      Send Q^A, P^A, a^A to B
      Initialize Q^B based on Q^A
    else
      Send P^A, a^A to B
    end if
    $s'^B \leftarrow$ collaborative spectrum estimation
    Observe resultant reward **r**
    $a^B \leftarrow \text{SampleAction}(A(k), \mu_k)$
    Update Q^B
    Send an ACK message to the transmitter
  end if
end for

4. Platform Demonstration and Result Analysis

4.1. Platform Demonstration

The multi-user collaborative anti-jamming platform developed by us is shown in Figure 3. In this system, the transmitter and the receiver of the strong sensing subsystem A represent the communication pair between PU and PU, while the transmitter and the receiver of the weak sensing subsystem B represent the communication pair between SU and SU. In order to reflect the heterogeneity of the sensing devices, the sampling rates of the sensing devices in subsystems A and B are set to $B_1 = B_w$ and $B_2 = B_{sub}$. The collaborative distance d_1 is defined as the distance between the receivers of subsystems A and B. The distance d_2 of transmission refers to the distance between the transmitter and the receiver within the subsystem. We set up three different sources of jamming, including an intelligent jammer, a traditional jammer distributed between subsystems A and B, and a traditional jammer near subsystem B. The jamming probability matrices of the two traditional jammers are illustrated in Figure 4. The probability of the sensing signal of the intelligent jammer is set to $\tau = 0.4$, which means that the jammer will jam a channel when the utilization rate of the channel is greater than 0.4. Relevant experiments are conducted to verify the multi-user anti-jamming performance, and the parameters are shown in Table 1.

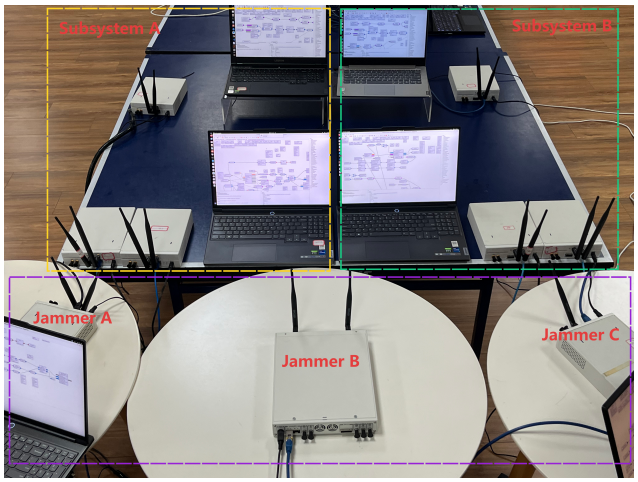


Figure 3. Multi-user collaborative anti-jamming platform.

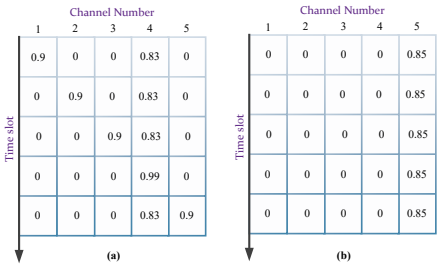


Figure 4. The jamming probability matrices of jammers (**a**) between A and B and (**b**) near B.

Table 1. Parameter settings.

Parameter Name	Value
Center frequency (MHz)	2129.6, 2129.8, 2130, 2130.2, 2130.4
Number of channels	5
Jamming signal power (dBm)	5
Antenna gain (dB)	15
Constellation mapping	QPSK
Jamming signal bandwidth (KHz)	200
Data signal bandwidth (KHz)	200
Wideband spectrum sensing sampling rate (MHz)	1
Narrowband spectrum sensing sampling rate (KHz)	200
Collaborative distance (m)	2
Transmission distance (m)	1

4.2. Result Analysis

Figure 5 presents a spectrum waterfall plot illustrating the impact of the intelligent jammer. In this plot, the vertical axis represents time, and the horizontal axis denotes the communication channels. When the data signal remains concentrated on a particular channel for an extended duration, the intelligent jammer detects its presence and launches a strong suppressive jamming signal on the same channel. As observed in the figure, once the jammer detects a continuous transmission lasting approximately 3 s, it identifies the active channel and effectively suppresses the communication by overwhelming it with interference. Moreover, when the legitimate user switches to a new channel, the jammer promptly recognizes the change and shifts its suppression accordingly. The plot demonstrates three successive channel switches, all of which are accurately detected and immediately jammed.

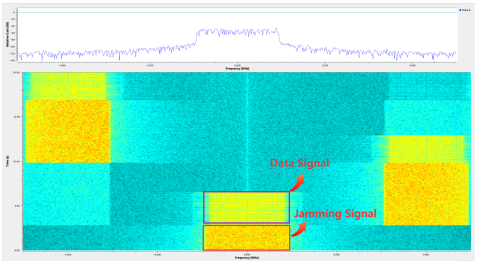


Figure 5. The intelligent jammer senses and releases the jamming signal.

Figure 6 shows the spectrum state obtained through wideband spectrum sensing at the start of subsystem A. To clearly show the results of the sensing and algorithmic decisions, a spectrum semantic map based on the spectrum waterfall has been constructed. In this map, green bands represent the communication channels, red bands are the channels affected by jamming, and yellow bands show the communication channels that have been blocked due to jamming.

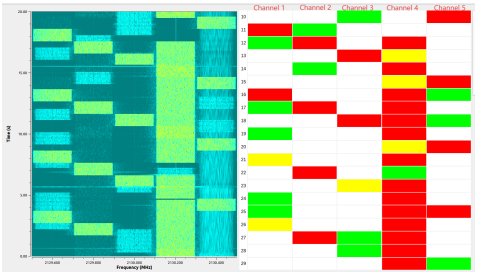


Figure 6. The result of subsystem A's wideband spectrum sensing.

As shown in Figure 7, subsystem A can reduce the probability of being jammed to 0.1 after about 25 iterations. After performing wideband spectrum sensing, subsystem A does not classify channel 5 as jamming because this jammer is deployed closer to subsystem B than to subsystem A. The demodulation threshold is still satisfied when the channel is selected for communication. So, this jamming signal is not reflected in the spectrum semantic map. In this situation, subsystem B faces a similar but not identical jamming environment to subsystem A. In other words, the jamming situation faced by subsystem B is more complex than subsystem A. This results in a high probability of jamming at the beginning of subsystem B's iteration. However, with the collaboration of subsystem A, subsystem B can still achieve the same effect as having wideband spectrum sensing.

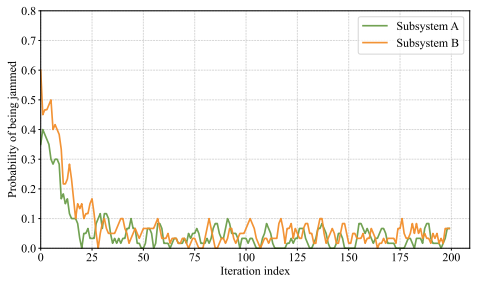


Figure 7. The jamming probability of the system.

Figure 8 shows the selection action probabilities of the system. The probability of selecting all actions is lower than τ, effectively avoiding the intelligent jammer. Subsystem A has a selection rate of only 0.019 on channel 4, indicating that the jammer continuously releases a jamming signal on this channel. Subsystem B has selection rates of 0.037 and 0.029 on channels 4 and 5, which is consistent with the jammer's impact on subsystem

B. In Figure 9, the blue bands are the action of subsystem A, and the green bands are the action of subsystem B. It can be observed that both subsystems A and B successfully avoid jamming during iterations 61 to 80, demonstrating the system's anti-jamming performance.

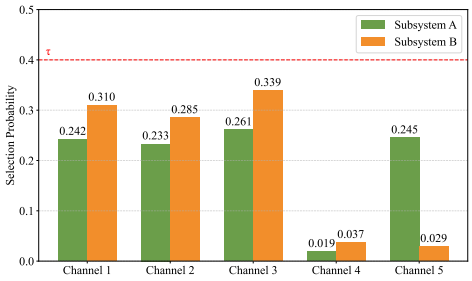


Figure 8. Cumulative action selection probabilities of the system after 200 iterations.

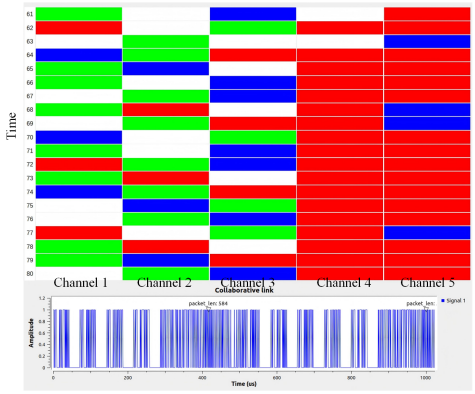


Figure 9. The spectrum semantic map during iterations 61 to 80.

5. Conclusions

A collaborative anti-jamming system framework for multi-user distributed decision making under the heterogeneity of sensing devices has been proposed, utilizing an online Q-learning algorithm to achieve anti-jamming in each subsystem. In addition, the data link, the ACK control link, and the collaborative link have been designed to enable collaborative anti-jamming. A collaborative spectrum estimation method has been proposed. The weak sensing subsystem builds a historical jamming knowledge base and restores the spectrum state in collaboration with the strong sensing subsystem. Finally, a multi-user collaborative anti-jamming system suitable for heterogeneous sensing devices has been developed. The anti-jamming performance of each subsystem has been analyzed based on the probability of action selection and the probability of being jammed. As the algorithm has converged, the actions of all subsystems have become uniformly distributed across the available channels. Moreover, the jamming probability has remained below 0.1, further confirming the system's anti-jamming performance.

Author Contributions: Conceptualization, S.G. and Y.H.; methodology, S.G. and N.Q.; software, S.G. and Y.H.; validation, S.G., X.M. and N.Q.; formal analysis, L.J.; investigation, S.G.; resources, D.W. and N.Q.; data curation, S.G.; writing—original draft preparation, S.G.; writing—review and editing, D.W. and N.Q.; visualization, S.G.; supervision, L.J.; project administration, N.Q. and D.W. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the National Aerospace Science Foundation of China under Grant 2023Z021052002, the National Key Laboratory of Multi-domain Data Collaborative Processing and Control under Grant MDPC20240402, the National Natural Science Foundation of China under Grant 62271253 and Grant 62471491, and the Postgraduate Research and Practice Innovation Program of NUAA under Grant xcxjh20240406.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

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

References

1. Pirayesh, H.; Zeng, H. Jamming Attacks and Anti-Jamming Strategies in Wireless Networks: A Comprehensive Survey. *IEEE Commun. Surv. Tutorials* **2022**, *24*, 767–809. [CrossRef]
2. Liu, X.; Guan, M.; Zhang, X.; Ding, H. Spectrum Sensing Optimization in an UAV-Based Cognitive Radio. *IEEE Access* **2018**, *6*, 44002–44009. [CrossRef]
3. Si, P.; Yu, F.R.; Yang, R.; Zhang, Y. Dynamic spectrum management for heterogeneous UAV networks with navigation data assistance. In Proceedings of the 2015 IEEE Wireless Communications and Networking Conference (WCNC), New Orleans, LA, USA, 9–12 March 2015; pp. 1078–1083.
4. Supriyatno, B.I.; Hidayat, T.; Susksmono, A.B.; Munir, A. Development of Radio Telescope Receiver Based on GNU Radio and USRP. In Proceedings of the 2015 1st International Conference on Wireless and Telematics (ICWT), Manado, Indonesia, 17–18 November 2015 ; pp. 1–4.
5. Li, X.; Chen, J.; Ling, X.; Wu, T. Deep Reinforcement Learning-Based Anti-Jamming Algorithm Using Dual Action Network. *IEEE Trans. Wirel. Commun.* **2023**, *22*, 4625–4637. [CrossRef]
6. Yao, F.; Jia, L. A Collaborative Multi-Agent Reinforcement Learning Anti-Jamming Algorithm in Wireless Networks. *IEEE Wirel. Commun. Lett.* **2019**, *8*, 1024–1027. [CrossRef]
7. Liu, X.; Sun, C.; Yau, K.L.A.; Wu, C. Joint Collaborative Big Spectrum Data Sensing and Reinforcement Learning Based Dynamic Spectrum Access for Cognitive Internet of Vehicles. *IEEE Trans. Intell. Transp. Syst.* **2024**, *25*, 805–815. [CrossRef]
8. Liu, Y.; Qi, N.; Pang, Z.; Zhang, X.; Wu, Q.; Jin, S.; Wong, K.K. Metasurface-Based Modulation with Enhanced Interference Resilience. *IEEE Commun. Lett.* **2023**, *27*, 1447–1451. [CrossRef]
9. Ali, A.S.; Lunardi, W.T.; Bariah, L.; Baddeley, M.; Lopez, M.A.; Giacalone, J.P.; Muhaidat, S. Deep Reinforcement Learning Based Anti-Jamming Using Clear Channel Assessment Information in a Cognitive Radio Environment. In Proceedings of the 2022 5th International Conference on Advanced Communication Technologies and Networking (CommNet), Marrakech, Morocco, 12–14 December 2022; pp. 1–6.
10. Janiar, S.B.; Wang, P. Intelligent Anti-Jamming Based on Deep Reinforcement Learning and Transfer Learning. *IEEE Trans. Veh. Technol.* **2024**, *73*, 8825–8834. [CrossRef]
11. Bono, F.M.; Polinelli, A.; Radicioni, L.; Benedetti, L.; Castelli-Dezza, F.; Cinquemani, S.; Belloli, M. Wireless Accelerometer Architecture for Bridge SHM: From Sensor Design to System Deployment. *Future Internet* **2025**, *17*, 29. [CrossRef]
12. Hwang, T.; Yang, C.; Wu, G.; Li, S.; Li, G. OFDM and Its Wireless Applications: A Survey. *IEEE Trans. Veh. Technol.* **2009**, *58*, 1673–1694. [CrossRef]
13. Liu, Y.; Du, Z.; Zhang, F.; Zhang, Z.; Yu, W. Implementation of Radar-Communication System based on GNU-Radio and USRP. In Proceedings of the 2019 Computing, Communications and IoT Applications (ComComAp), Shenzhen, China, 26–28 October 2019; pp. 417–421.
14. Slimeni, F.; Chtourou, Z.; Scheers, B.; Nir, V.L.; Attia, R. Cooperative Q-learning based channel selection for cognitive radio networks. *Wirel. Netw.* **2018**, *25*, 4161–4171. [CrossRef]
15. Jing, X.; Wang, R.; Lei, H.; Liu, H.; Chen, Q. Multi-Agent Discrete Soft Actor-Critic Algorithm-Based Multi-User Collaborative Anti-Jamming Strategy. *IEEE Trans. Inf. Forensics Secur.* **2025**, *20*, 5025–5038.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.