Research on Anti-Jamming for UAV Airborne Jamming Method Based on Relay Optimization Strategy

Fang Fang1,2*, Xuewen Liu3, Yonggui Li2 and Qing Shen1

1 Army Engineering University of PLA, Nanjing, Jiangsu, 239000, P. R. China
2 The Sixty-Third Research Institute, National University of Defense Technology, Nanjing, Jiangsu, 239000, P. R. China
3 The Company Commander of 31618 Troops 41 Element, Fuzhou, Fujian, 350000, P. R. China

*Corresponding author’s e-mail: 520093773@qq.com

Abstract. Aiming at the new threat of Unmanned Aerial Vehicle (UAV) airborne jamming faced by wireless mobile communication networks, a communication network anti-jamming technology method based on relay optimization selection strategy is proposed. In order to achieve the maximum transmission rate of the selected relay node, a system model based on multi-arm bandit (MAB) is established. It is proved by simulation that the method of selecting the best relay node by Upper Confidence Bound (UCB) algorithm is feasible. It can avoid the nodes that are most affected by jamming and improve the anti-jamming ability of the communication system.

1. Introduction

The wireless Ad hoc networks are widely used because of their advantages such as temporary opening, mobile access and dynamic structure adjustment. As an emerging convenient carrier, the UAV is not only small in size and light in weight, but also has the advantages of strong mobility and high security. The emergence of UAV airborne jamming has greatly increased the flexibility and concealment of jamming, which has brought challenges to the research of anti-jamming technology. The jammers in this paper no longer implement attack for signal parameters and key information. Instead, the UAV is used to jam key nodes in the communication network, such as relay nodes, thereby effectively reducing the communication performance of the wireless network.

Therefore, this paper proposes an anti-jamming strategy based on the multi-arm bandit (MAB) model and simulates the performance of the algorithm. The simulation results show that the anti-jamming strategy based on UCB algorithm can avoid the nodes most affected by jamming. At the same time, the anti-jamming strategy utilizes the best relay nodes while efficiently exploring other relay nodes. This continuously improves the anti-jamming ability of the communication system.

2. System model and problem description

Due to the openness of wireless communication, wireless Ad hoc networks are vulnerable to a variety of security threats. Malicious jamming by attackers is a major threat to wireless communication systems. As shown in Figure 1, it is assumed that there is a pair of communication parties and m relay nodes in a wireless communication network. Due to factors such as path loss and shadow, transmitter A must select one of the relay nodes to forward to reach receiver B. The jammers observe the location range of the relay
node through the reconnaissance device. Then the UAVs implement wideband suppression jamming with period \( T_1, T_2 (T_1 < T_2) \) on a circular path with radius \( r_1, r_2 (r_1 < r_2) \). And the effective jamming range of the two UAVs \( J_1, J_2 \) can cover \( m \) relay nodes. In the case that the communication network cannot obtain the jamming information, how to ensure the maximum communication capacity is an urgent problem that the communication party needs to solve. The formula for channel capacity is

\[
C = B \log_2 \left(1 + \frac{S}{(J + N)}\right)
\]

Where \( B \) is the channel bandwidth (Hz) and \( S \) is the signal power (W) and \( N \) is the noise power (W) and \( J \) is the jamming signal power (W).

**Figure 1. Wireless Mobile Communication Under UAV Airborne Jamming**

As shown in Figure 1, the two jammers have a track radius of \( r_1, r_2 (r_1 < r_2) \) and a period of \( T_1, T_2 (T_1 < T_2) \). Therefore, it has been in a periodic motion state and the position is uncertain. It causesthat the jamming power received by the relay node in each time slot is dynamically changing. The receiving power of the receiver is

\[
P_R = \alpha P_t \times d^{-1}, 0 < \alpha < 1
\]

Where \( P_t \) is the transmit power of the signal and \( P_R \) is the received power of the signal and \( d \) is the distance between the transmitter and the receiver, and \( \alpha \) is the positive correlation operation. Assuming that the jamming power of the UAV jammer is \( P_j \), then the jamming power \( J(t) \) received by each relay node is

\[
J_1(t) = \alpha P_j \times d_{1,i}^{-1}, 0 < \alpha < 1
\]

\[
J_2(t) = \beta P_j \times d_{2,i}^{-1}, 0 < \beta < 1
\]

\[
J(t) = J_1(t) + J_2(t)
\]

\( d_{1,i}, d_{2,i} \) represent the distance between jammer 1 and jammer 2 to the relay node.

Each time the data is sent, the transmitter needs to find a node \( i \) from the \( m \) relay nodes for transmission. As shown in Figure 2, the maximum transmission rate from transmitter A to the relay node is

\[
r_{i,A} = C_{i,A}, i \in m
\]

The maximum transmission rate of relay node \( i \) to receiver B is

\[
r_{i,B} = C_{i,B}, i \in m
\]

So the above problem is equivalent to how to choose the best relay node to maximize the transmission rate

\[
r = \text{arg max}(r_{i,A} + r_{i,B}), i \in m
\]
3. Anti-jamming strategy based on UCB algorithm

Therefore, we model the communication system as a MAB model in the absence of jamming information from the communicating party. The MAB model was first proposed by Robbins, and the classic case is a gambler game. The MAB has a number of arms, each of which produces a corresponding benefit. When the gambler pulls an arm, it gets a benefit. Obviously, the gain is related to the relevant distribution of the push arm. Before the game starts, the gambler is unable to know the distribution of the gain. In order to gain maximum benefits, we need to take some method. The gambler chooses one of the arms to push many times until the number of times reaches a certain value. The statistical distribution of the gain corresponding to the arm can be obtained. In this chapter, the arm is defined as a relay node, so there are \( m \) arms. Then, the instantaneous gain of the \( K \)th selection relay node is \( \mu_i(k) \). It is

\[
\mu_i(k) = r_{i,a}(k) + r_{i,b}(k)
\]

After the selection of \( t \) times, the average value of the gain of node \( i \) is

\[
\tilde{\mu}_i(t) = \frac{1}{T_i(t)} \sum_{k=1}^{t} \mu_i(k)
\]

Where \( T_i(t) \) represents the number of times that node \( i \) is selected after \( t \) times. As a basis for judging the merits and demerits of the algorithm, the regret value is usually adopted. In this chapter, the regret value \( R(t) \) is the difference between the total gain obtained by selecting the best node and the total gain obtained after adopting an algorithm after \( t \) time slots. Its mathematical expression is

\[
R(t) = t\mu^* - \sum_{i=1}^{m} \mu_i(k)
\]

Where \( \mu^* \) represents the average gain of the best node

\[
\mu^* = \max\{ \mu_1, \mu_2, \ldots, \mu_m \}
\]

According to the MAB model established above, in order to minimize the regret value or maximize the gain, we need to know the gain, that is, the transmission rate, of each relay node selected in the action set. In order to solve this problem well, a relay optimization selection strategy based on UCB algorithm is proposed.

First, \( m \) relay nodes are randomly distributed in a range of \( L \times L \) squares, where the relay node \( i \) coordinates are \((x_i, y_i)\). Two UAV’s motion trajectory is centered on the center of the \( L \times L \) square. Their flight radiuses are \( r_1, r_2 \) \((r_1 < r_2)\), and their jamming period are \( T_1, T_2 \) \((T_1 < T_2)\), and their power of both jammers are \( p \). Assume that the coordinates of the two UAVs at time \( t \) are \((x_{J1,t}, y_{J1,t}), (x_{J2,t}, y_{J2,t})\). Then, the transmission rate of the relay node \( i \) is

\[
r_i = B \log_e \left( 1 + \frac{S}{\left( (\alpha + \beta) p + N \right)} \right)
\]

\[
\alpha \propto \sqrt{(x_{J1,t} - x_i)^2 + (y_{J1,t} - y_i)^2}
\]

\[
\beta \propto \sqrt{(x_{J2,t} - x_i)^2 + (y_{J2,t} - y_i)^2}
\]

Therefore, the gain of the relay node \( i \) is \( \mu_i = r_i \). Since this paper assumes that the jammers are moving in the scenario, the relay node with the highest transmission rate at each moment is changing. Therefore, you cannot always select the node with the highest transmission rate at the previous time. So how to choose the best node? Is it to use historical data to make judgments or to find new nodes to explore? If you always select the node with the highest known transmission rate, this method cannot estimate the transmission rate of each node very well, but you may not be able to select the optimal node. If each node is selected in turn to obtain the transmission rate of each node, this method can well estimate
the transmission rate of each node, but you may miss many opportunities to select the optimal node.

Score each arm according to the UCB algorithm’s scoring formula (14)

\[
\text{arm}(t) = \arg \max_{i=1,2,\ldots,m} (\hat{\mu}_i(t) + \frac{2\ln(t)}{T_i(t)})
\] (14)

Where \( \hat{\mu}_i \) is the historical average gain and its iteration formula is

\[
\hat{\mu}_i(t) = \begin{cases} 
\frac{\hat{\mu}_i(t-1) \times T_i(t-1) + \mu_i(t)}{T_i(t-1) + 1}, & \text{choose relay node } i \\
\hat{\mu}_i(t-1), & \text{else}
\end{cases}
\] (15)

\( T_i(t) \) is the number of times that the node \( i \) is selected, and its iteration formula is

\[
T_i(t) = \begin{cases} 
T_i(t-1) + 1, & \text{choose relay node } i \\
T_i(t-1), & \text{else}
\end{cases}
\] (16)

Each time the highest-scored relay node is selected, then the transmission rate is observed, and the parameters of the arm are updated according to the feedback, and the iteration is repeated.

According to formula (14), it can be seen that the first part refers to the expected average gain of the current node, which reflects the utilization of the data, and the second part represents a confidence factor, which reflects the exploration of the data. Next time which node will be selected is determined by two parts. As the number of selecting nodes \( i \) increases, the confidence factor decreases, but the average gain is expected to increase. When the number of times of selecting node \( i \) is sufficiently large, the value of \( \text{arm}(t) \) approaches the average estimated gain of the node, which is the upper bound of the confidence interval.

Algorithm anti-jamming algorithm based on relay optimization

1. Data initialization: UAV’s flight radius \( r_1, r_2 \), jamming power \( P \) and jamming period \( T_1, T_2 \), number of relay nodes \( m \) and location \( (x, y) \), time slot \( N \) and UCB algorithm regret value \( R \).
2. Calculate the real-time revenue \( \mu \) of each node and calculate the score value \( \text{arm}(t) \), and select the relay node with the highest score to transmit.
3. Update the main parameters \( \hat{\mu} \) and \( T(t) \) according to the selected relay node feedback.
4. According to the updated parameters \( \hat{\mu} \) and calculate the new score value \( \text{arm}(t) \), select the relay node corresponding to the maximum value for transmission from the new score value \( \text{arm}(t) \).
5. If \( T(t) \leq N \), repeat 3~5.
6. If \( T(t) > N \), it ends.

Figure 3. Optimal Relay Node Selection Flow Chart

In practice, it is difficult for the communicating party to accurately obtain the jamming information. Especially in the dynamic environment where two UAVs launch jamming in different periods, it is difficult to extract effective jamming information from the complex jamming electromagnetic environment. Therefore, the anti-jamming problem under unknown jamming information is built into the MAB model. Then using the online learning UCB algorithm to find the best relay node is one of the effective methods. The process is shown in Figure 3.
4. Simulation Results and Analysis

In order to verify the anti-jamming performance of the proposed algorithm, this paper uses the MATLAB simulation tool to simulate the system model. It is assumed that the wireless communication network has $m$ relay nodes in the range of $1000 \times 1000m^2$. The signal transmission power of the communication party is $P_t = 100mW$. The flying radiuses of the UAVs are $r_1 = 1200m, r_2 = 1500m$. And the jamming power of the UAVs are $P_j = 1000mW$.

Firstly, $m$ nodes are randomly generated within the range of $1000 \times 1000m^2$. Secondly, the distances from all nodes to the two jammers are calculated. Then the jamming powers that the nodes receive are calculated according to formulas (6)-(8). Finally, according to the gain of all nodes, select the maximum gain node as the action of the next moment.

When $m = 10$, and $T_1 = \pi / 5(rad/s), T_2 = \pi / 10(rad/s)$, we get Fig. 4 and Fig. 5. According to the simulation diagram, as the number of selections increases, the regret value gradually decreases. When the number of selections reaches 50, the regret value is basically 0, and the accumulated regret value remains basically $3.5 \times 10^8 bit/s$.

![Figure 4](image1.png)  
**Figure 4.** The regret value of the node when the number of nodes is 10

![Figure 5](image2.png)  
**Figure 5.** The accumulated regret value of the node when the number of nodes is 10

When $m = 5$, and $T_1 = \pi / 5(rad/s), T_2 = \pi / 10(rad/s)$, we get Fig. 6 and Fig. 7. According to the simulation diagram, as the number of selections increases, the regret value gradually decreases. Especially the convergence speed is getting faster. When the number of selections reaches 30, the regret value is basically 0, and the accumulated regret value remains basically $5.8 \times 10^8 bit/s$. We fix other parameters and study the effect of changing the number of relay nodes on the UCB algorithm. It can be seen from Fig. 6 and Fig. 7 that when the number of nodes is reduced from 10 to 5, the algorithm converges faster. It shows that the anti-jamming ability of the system is obviously improved.
5. Summary
In this paper, the MAB model is applied to the dynamic anti-jamming technology of wireless communication. The UCB method is used to optimally solve the relay node optimization selection problem. The simulation analysis shows that under the condition where the UAV dynamic jamming and the communication party are not fully aware of the wireless communication environment, the UCB algorithm can avoid the nodes with the greatest impact and improve the anti-jamming ability of the communication system. The system model in this paper is relatively simple, and it is necessary to continue to explore the dynamic anti-jamming problem in complex communication environments in the future work.

References
[1] Lu J, Lun L, Dan S, et al. Dynamic Multi-Arm Bandit Game Based Multi-Agents Spectrum Sharing Strategy Design [J]. 2017.
[2] Amuru S D, Tekin C, Schaar M V D, et al. Jamming Bandits - A Novel Learning Method for Optimal Jamming [J]. IEEE Transactions on Wireless Communications, 2016, 15(4): 2792-808
[3] Bonifonte A, Chen Q. Restless Multi-Arm Bandits Problem: An Empirical Study [J].
[4] Besbes O, Gur Y, Zeevi A. Stochastic multi-armed-bandit problem with non-stationary rewards[C]; proceedings of the International Conference on Neural Information Processing Systems, 2014:199-207
[5] Melián-Gutiérrez L, Modi N, Moy C, et al. Hybrid UCB-HMM: A Machine Learning Strategy for Cognitive Radio in HF Band [J]. IEEE Transactions on Cognitive Communications & Networking, 2016, 1(3): 347-58.
[6] Zhang J, Jiang H S. Spectrum sharing based on multi-armed bandit online learning [J]. Computer Engineering & Design, 2014.
[7] Zhu J, Chen H, Xiong J. Channel Selection Based on Multi-armed Bandit [J]. Telecommunication Engineering, 2015.