An MDP-Based Time Domain Impulse Jamming Mitigation Scheme

Yutao Wang[1], Yingtao Niu[2], Jianzhong Chen[2] and Fang Fang[1]
[1] *Graduate School, Army Engineering University, Houbiaoying Road, Nanjing, P.R.China*
[2]*The Sixty-third Research Institute, National University of Defence Technology, Nanjing, P.R.China*

Keywords: Markov decision process, Anti-jamming, Impulse jamming.

Abstract: This paper proposed a method through a time domain Markov decision process as a countermeasure of random periodic impulse jamming for a user in a time slotted environment. First, the random periodic impulse jamming is modelled. Then the time domain MDP-based anti-jamming communication model is proposed and the optimal transition probability on each state is calculated. Finally, we proposed an online learning algorithm to approach the optimal transition probabilities. Simulation results show that our method is better than other countermeasures of impulse jamming.

1 INTRODUCTION

Impulse jamming can corrupt the data transmission of communication system(Poisel, 2011) in various applications like IoT systems(Landa et al., 2017), OFDM systems(Epple and Schnell, 2017) et al. A short form periodic jamming (SFPJ) attack can cause huge reduction of packet delivery ratio (PDR) with little cost and traditional anti-jamming schemes such as spread spectrum techniques in frequency domain is not appropriate to the situation due to the impulse signal has a wide spectrum density(Debruhl and Tague, 2013). One usual pulse jamming pattern is called periodic impulse jamming, which generates impulse jamming periodically. The jamming source is widely distributed in practice, such as high-voltage equipment(Lin et al., 2015). Despite the interferences generated by nature, impulse jamming is also commonly used by malicious users to corrupt communication links. Jie et al. derived a closed form of BER (Bit Error Rate) of optimal periodic impulse jamming for QPSK system(Jie et al., 2017). As a countermeasure, the detection of periodic impulse jamming is studied. Yuan Yuan He, et al.(He et al., 2008) used wavelet transforming method to estimate impulse jamming.

Instead of periodic impulse jamming, malicious user can use variants of periodic impulse jammings to improve jamming effect and to avoid being detected. Random periodic impulse jamming is a kind of impulse jamming whose occurrence time obeys some distribution. We proposed a Markov decision process (MDP) based countermeasures to mitigate the jam effects.

This paper is organized as follow. The system model is introduced in Section 2. In Section 3, we calculated the optimal transmission probability under certain jamming probability. An online learning algorithm is provided in section 4 to obtain the optimal transition probability vector. Section 5 presents computer simulation results. Finally, in Section 6, some concluding remarks are provided.

2 SYSTEM MODEL

Consider the situation where a synchronized time-slotted communication system consists 2 licensed users, one of which sender, the other receiver. In each time slot, the sender sends a frame with length t_L. A malicious jamming generates impulse jamming sequentially with some distributions based on certain period $T, T >> t_L$. We consider each pulse duration is too short than a frame length, but can corrupt the frame. In this paper, We assume that the launched time of k 's jamming impulse is independent normally distributed with a mean of $kT, k = 1,2,3,...$, variance σ^2.

Wang, Y., Niu, Y., Chen, J. and Fang, F.
An MDP-Based Time Domain Impulse Jamming Mitigation Scheme.
DOI: 10.5220/0008099500490054
In *Proceedings of the International Conference on Advances in Computer Technology, Information Science and Communications (CTISC 2019)*, pages 49-54
ISBN: 978-989-758-357-5

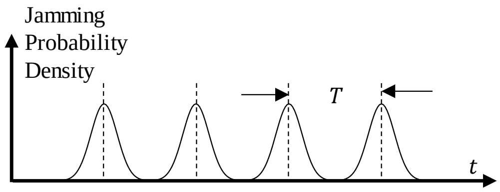

Figure 1: Jamming probability density.

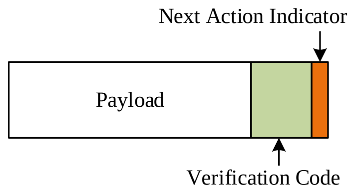

Figure 2: Component of a frame.

The pdf of the arrival interval is the convolution of the pdf of 2 adjacent jammings, denoted as $f(x)$. Apparently $f(x)$ is normally distributed with mean T and variance $2\sigma^2$. If current time slot is $t_0 = nT$. For convenience, we denote the conditional probability of jamming next time slot is

$$p_J(n+1) = \frac{\int_{t_0}^{t_0+t_L} f(x)}{\int_{t_0}^{t_{\max}} f(x)} = \frac{Q\left(\frac{t-t_0}{\sigma}\right) - Q\left(\frac{t-(t_0+t_L)}{\sigma}\right)}{Q\left(\frac{T-t_0}{\sigma}\right)} \quad (1)$$

Where

$$Q(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_x^{+\infty} e^{-\frac{t^2}{2}} dt \quad (2)$$

As a countermeasure which depicted in Figure 2, in each time slot the sender sends a frame consists 3 elements: payload, verification code and the Next Action Indicator (NAI). The verification code part is used to check whether this frame is corrupted by the jammer, while the NAI indicates whether to continue send signal or keep silent next frame, i.e. indicates the receiver whether a legitimate frame comes next time slot. We consider the duration of the NAI part is too small that cannot be influenced by the jamming. To make the problem clear to understand, we assume the receiver can immediately sensor the communication status i.e. whether the channel is jammed by the malicious attacker.

We define the current the state of the current time slot S_n, At the end of each time slot, the sender observes the state of the current time slot S_n, and select the corresponding action $a_n \in \{0,1\}$ with a probability p_n, "0" represents "to silent", "1" represents "to continue". When taking action "to silent", the sender would stop sending message until a jamming is detected; when taking action "to continue", the sender would continue sending message next time slot. If this is the *K*th consecutive slot with successful transmission, the state is denoted by $S_n = K$. The transmitter receives an immediate payoff $U(n)$ in the *n*th time slot, decided by

$$U(n) = \text{Send} \cdot \mathbf{1}\left(R \cdot \mathbf{1}(\text{Transmitted}) - L \cdot \mathbf{1}(\text{Jammed})\right) \quad (3)$$

where R represents communication gain, while L represents jammed loss. $\mathbf{1}(\cdot)$ is the indicator function returns 1 when the statement in the parenthesis holds TRUE and 0 otherwise.

The transition of states can be described as a Markov chain, as show in Figure 3. The transition probabilities depend on the action taken by the transmitter. We use $P(S_m | S_n, 0)$ and $P(S_m | S_n, 1)$ to represent the transition probability from the current state S_n to a new state S_m when taking action 0 and 1, respectively. Obviously,

$$\begin{aligned} &P(S_{n+1} | S_n, 1) = 1 - p_J(n) \\ &P(S_0 | S_n, 1) = p_J(n) \\ &P(S_0 | S_n, 0) = 1, n > 0 \end{aligned} \quad (4)$$

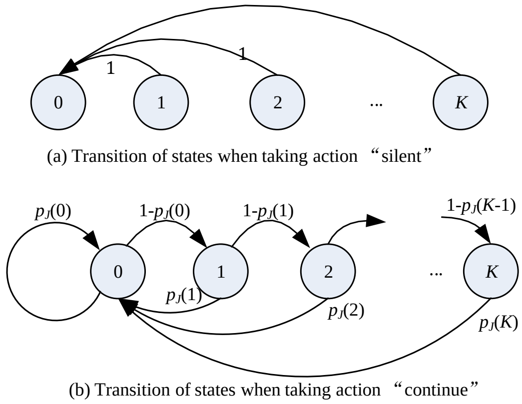

Figure 3: Markov chains of state transitions when different actions are taken.

Note that the jammer may jam the channel with some probability, the transmitter will have possibility being jammed when taking action "continue" at a certain state. The state of the next time slot depends on the action of the current time slot, the jamming condition and the state of current time slot, hence we can model this scenario as a Markov decision process (MDP), from which the defense strategy is obtained.

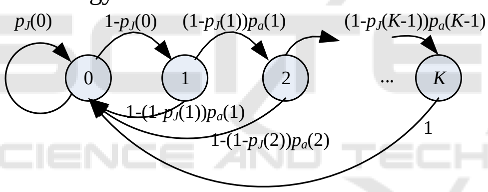

Figure 4: Total transition probability.

An MDP consists of four important components, namely, a finite set of states, a finite set of actions, transition probabilities, and immediate payoffs(Wu et al., 2011). As the attacker jams the channel with some distribution, if the transmitter continues sending message, it will be eventually blocked by the jammer. Thus, the state S will be finite. Denote the maximum possible According to (3), the immediate payoff depends on both the state and the action of previous time slot, i.e.

$$U(S_n, a_n) = \begin{cases} R, & \text{if } S_n = \{0,1,2,...\}, a_n = 1, \textit{Unjammed} \\ -L, & \text{if } S_n = \{1,2,3,...\}, a_n = 1, \textit{Jammed} \\ 0, & \text{if } S_n = \{0,1,2,...\}, a_n = 0 \end{cases} \tag{5}$$

3 CALCULATION

Our goal is to find the appropriate $\boldsymbol{p_a}$ that can maximize the sum of immediate payoff and expected payoff conditioned on the current action probability.

$$U(\boldsymbol{S}) = \sum_{n=1}^{N} U(S_n, a_n) \tag{6}$$

In all the scenarios above, the interval between 2 jamming signals is not infinite in practice, so we can be informed that the states are also finite. Denote $N = \left\lfloor \frac{t_{\max}}{t_L} \right\rfloor$ as the maximum states count, where $f(t_{\max}) < p_{thresh}$.

$$p_J(n) = \frac{f(t_L \cdot n < t < t_L \cdot n + t_L \mid t = t_L \cdot n)}{f(t > t_L \cdot n \mid t = t_L \cdot n)} \tag{7}$$

Denote $\boldsymbol{S} = (S_0, S_1, ...S_N)$ which represents state vector, $\boldsymbol{a} = (a_0, a_1, ...a_N)$ represents action vector; $\boldsymbol{p_S} = (p_S(0), p_S(1), ...p_S(N))$ represents vector of probabilities of each state; $\boldsymbol{p_a} = (p_a(0), p_a(1), ...p_a(N))$ represents vector of probabilities of action on each state; $\boldsymbol{p_J} = (p_J(0), p_J(1), ...p_J(N))$ represents vector of jamming probability on each state.

From the definition, $p_a(0) = 1$, $p_J(N) = 1$, $p_a(N) = 0$. With the transition probability $\boldsymbol{p_a}$ we can derive the total state transition probability matrix $\mathbf{P}$.

$$\mathbf{P} = \begin{pmatrix} 1-(1-p_J(0))\,p_a(0) & (1-p_J(0))\,p_a(0) & ... & 0 \\ 1-(1-p_J(1))\,p_a(1) & 0 & ... & 0 \\ ... & ... & ... & ... \\ 1 & 0 & ... & 0 \end{pmatrix} \tag{8}$$

***Proposition* 1**: The states probability $\boldsymbol{p_S}$ is determined by $\boldsymbol{p_J}$ and $\boldsymbol{p_a}$.

Proof: from (8), the following stationary distribution of the Markov chain can be formulated as below:

$$\boldsymbol{p_S} = \boldsymbol{p_S}\mathbf{P} \tag{9}$$

$$\begin{aligned} p_S(0) &= \sum_{n=0}^{N} p_{n0} p_S(n) \\ p_S(1) &= p_{01} p_S(0) \\ p_S(2) &= p_{12} p_S(1) \\ &\dots \\ p_S(N) &= p_{N-1,N} p_S(N-1) \end{aligned} \tag{10}$$

Subject to

$$\sum_{n=0}^{N} p_S(n) = 1 \tag{11}$$

p_{mn} denotes the transition probability of state m to n. Thus, the states' probability can be figured out as

$$p_S(0) + \sum_{n=1}^{N} \left(p_S(n) \cdot \prod_{k=0}^{n} p_{k,k+1} \right) = 1 \tag{12}$$

According to (8)(10)(12),

$$\begin{aligned} p_S(0) &= \frac{1}{1 + \sum_{n=1}^{N} \left(\prod_{k=0}^{n-1} p_{k,k+1} \right)} \\ p_S(1) &= \left(1 - p_J(0)\right) p_S(0) \\ &\dots \\ p_S(N) &= \left(1 - p_J(0)\right) \cdot \dots \cdot \left(1 - p_J(N-1)\right) p_S(0) \end{aligned} \tag{13}$$

Thus, the states' probability $\boldsymbol{p_S}$ can be determined by $\boldsymbol{p_J}$ and $\boldsymbol{p_a}$.

The returning to state S_0 from state S_0 through several states is defined as a cycle. It is obvious that the state will eventually return to S_0. The total probability from each path in a cycle is 1, i.e.

$$\sum_{k=0}^{N} \Pr\{S_0 \mapsto S_1 \mapsto \dots \mapsto S_k \mapsto S_0\} = 1 \tag{14}$$

The expectation of total payoff in one cycle is denoted as $U^*(\boldsymbol{S})$, which can be obtained by the equation below.

$$U^*(\boldsymbol{S}) = \sum_{n=0}^{N} \left(\begin{array}{l} \left(\frac{p_S(n)}{p_S(0)} p_a(n) p_J(n)(nR - L) \right) \\ + \left(\frac{p_S(n)}{p_S(0)} \left(1 - p_a(n)\right) \cdot nR \right) \end{array} \right) \tag{15}$$

From the equation, it can be inferred that $U^*(\boldsymbol{S})$ is determined by $\boldsymbol{p_a}$ and $\boldsymbol{p_J}$. Our algorithm aims to achieve the goal of maximizing $U^*(\boldsymbol{S})$, the following equation is needed.

$$\begin{aligned} \boldsymbol{p_a} &= \arg\max\left(U^*(\boldsymbol{S})\right) \\ &= \arg\max \left(\sum_{n=0}^{N} \left(\begin{array}{l} \frac{p_S(n)}{p_S(0)} p_a(n) p_J(n)(nR - L) \\ + \frac{p_S(n)}{p_S(0)} \left(1 - p_a(n)\right) \cdot nR \end{array} \right) \right) \end{aligned} \tag{16}$$

It is theoretically possible to calculate the optimal $\boldsymbol{p_a}$. However, the calculation is too complex to solve due to $\boldsymbol{p_a}$ is a $\left\lfloor \frac{t_{max}}{t_L} \right\rfloor$ dimension vector. If the probability resolution is dp, the solution space will be $\left(\frac{1}{dp}\right)^{\left\lfloor \frac{t_{max}}{t_L} \right\rfloor}$. To find the max value and its index in acceptable time, the Simulated Annealing Algorithm(Ogbu and Smith, 1990) is introduced into this model. We set the vector $\boldsymbol{p_a}$ as argument. At each epoch, we make a random change of $\boldsymbol{p_a}$ and compare the expectation of corresponding $U^*(\boldsymbol{S})$. The change is accepted according to metropolis criterion. After sufficiently large time, the probability of $\boldsymbol{p_a} = \arg\max\left(U^*(\boldsymbol{S})\right)$ will be 1.

4 ONLINE LEARNING APPROAH

In practice, the jamming probability vector $\boldsymbol{p_J}$ cannot be obtained from the environment immediately. To approach the optimal transition probability vector $\boldsymbol{p_a}$, We propose an online learning algorithm. According to the simulation results, the optimal $\boldsymbol{p_a} = \{0,1\}$, and all "1" occur before a certain time slot and all "0" after that certain time slot. So we translate the problem into find this certain time slot as boundary before which continuously transmitting signal. We formulate the problem as a multi-armed bandit problem. The user selects Kth arm represents continuously transmit K time slots. Facing this problem, the trade-off between exploration and exploitation is the focus of the problem. If the user chooses the exclusively on the arm that he thinks is best(exploit), he may miss the actual best arm. If the user keeps trying out all the arms and gathering statistics ("exploration"), he may fail to play the best arm often enough to get a high return(Auer et al., 2011).

We define 2 variables, exploration rate α and temperature t to solve the problem. Both the two variables decrease over iteration. If the exploration rate is higher, the user is more likely to choose the new state to calculate the new payoff and vise versa. If the temperature is higher, the user is more likely to accept the new arm and vice versa.

The user selects the first arm at first transmitting cycle and calculate the sum payoff according to actual jammings. Our algorithm is depicted as Figure 5.

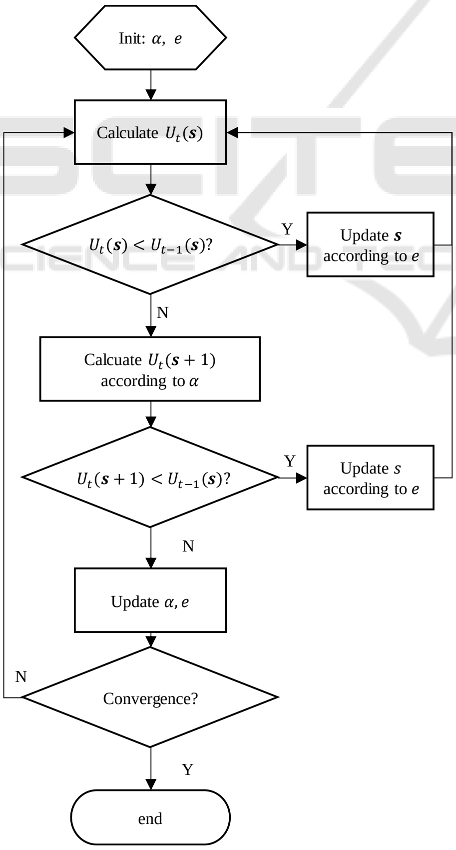

Figure 5: Online learning algorithm to approach optimal boundary.

5 SIMULATION RESULTS

First, we calculated the effect of different parameters to transmitting probability $\boldsymbol{p_a}$. We set transmission gain $R = 10$, time slot length $t_L = 0.01s$. For a normal distributed jamming, the jamming period $T = 0.5$ and variance $\sigma^2 = 0.1^2$. With different $L = 10, 20, 30, 40, 50$. The result is shown in Figure 6. The transmitting probability goes from 1 to 0 with as sharp gradient. Different jamming loss result in different transmitting probability vector. With greater jamming loss, the transmitting time slot trends to be shorter.

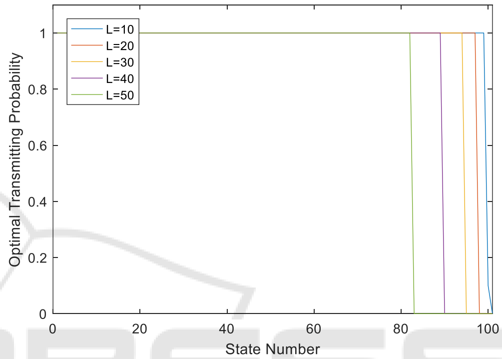

Figure 6: Optimal transmitting probability with different jamming loss with $\sigma^2 = 0.1^2$.

In Figure 7, we compared our scheme with different combinations of the attack and the defense strategy. The transmission gain and jamming loss are set to 10 and 30 respectively. We plotted the average sum payoff of a cycle in all the 5 situations. First, the communication is under a non-jamming environment. The equivalent average sum payoff of a cycle is the highest. After that, a periodic impulse jamming occurs and makes the average sum payoff a great loss. As a countermeasure, the authorized user takes periodic transmission to withstand the impulse jamming. The malicious attacker then chooses random periodic impulse jamming strategy, which drops the average sum payoff most. To mitigate the jamming effect, the authorized user then chooses MDP-Based impulse jamming mitigation scheme that rises the average sum payoff curve.

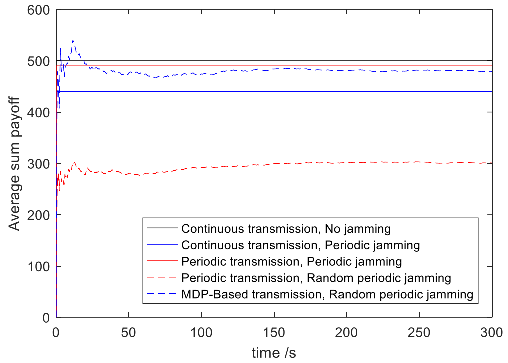

Figure 7: Comparison of different anti-jamming schemes.

In Figure 8, the proposed algorithm in section 4 is verified. We generated normally distributed jamming with a mean of $kT, k=1,2,3,\ldots$, $T=0.5$, variance $\sigma^2=0.1$. Transmitting gain $R=10$, jamming loss $L=50$. The theory value optimal state is calculated based on optimal $\boldsymbol{p}_a$ obtained at the beginning of this section.

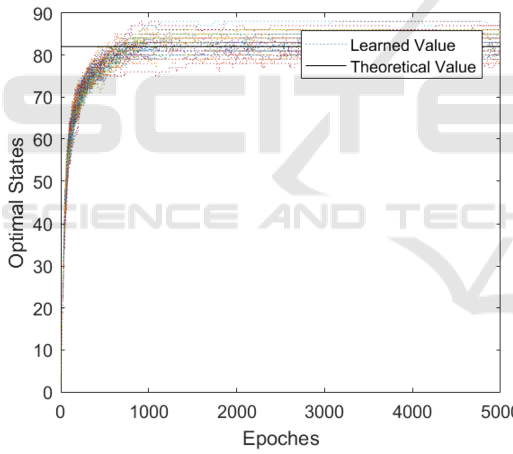

Figure 8: Comparison between learned value and theoretical value.

6 CONCLUSIONS

In this paper, we proposed a Markov decision process based impulse jamming mitigation scheme and used simulated annealing algorithm to obtain the numerical result. The optimal transmitting probability is either 0 or 1, and the continuously transmission slot is shorter when jamming loss L grows. We have shown that our scheme is better than periodic transmission scheme under same random periodic impulse jamming environment.

REFERENCES

Auer, P., N. Cesa-Bianchi, Y. Freund and R. E. Schapire, 2011. The Non-Stochastic Multi-Armed Bandit Problem. In *Siam Journal on Computing* 32(1):48-77.

Debruhl, B. and P. Tague, 2013. How to jam without getting caught: Analysis and empirical study of stealthy periodic jamming. In *IEEE International Conference on Sensing*.

Epple, U. and M. Schnell, 2017. Advanced Blanking Nonlinearity for Mitigating Impulsive Interference in OFDM Systems. In *IEEE Transactions on Vehicular Technology* 66(1):146-158.

He, Y. Y., C. J. Yu, T. F. Quan and X. Jin, 2008. Impulsive interference detection method based on Morlet wavelet and maximum likelihood estimation. In *IEEE International Conference on Industrial Informatics*.

Jie, Y., H. Bingyang and X. Jingying, 2017. Optimal Periodic Pulse Jamming Signal Design for QPSK Systems. In *Journal of Beijing Institute of Technology* 26(3):381-387.

Landa, I., A. Blazquez, M. Velez and A. Arrinda, 2017. Indoor measurements of IoT wireless systems interfered by impulsive noise from fluorescent lamps. In *European Conference on Antennas & Propagation*.

Lin, J., T. Pande, I. H. Kim, A. Batra and B. L. Evans, 2015. Time-Frequency Modulation Diversity to Combat Periodic Impulsive Noise in Narrowband Powerline Communications. In *IEEE Transactions on Communications* 63(5):1837-1849.

Ogbu, F. A. and D. K. Smith, 1990. The application of the simulated annealing algorithm to the solution of the n / m / C max flowshop problem. In *Computers & Operations Research* 17(3):243-253.

Poisel, R. A., 2011. *Modern Communications Jamming Principles and Techniques*, Artech House, Inc.

Wu, Y., B. Wang, K. J. R. Liu and T. C. Clancy, 2011. Anti-Jamming Games in Multi-Channel Cognitive Radio Networks. In *IEEE Journal on Selected Areas in Communications* 30(1):4-15.