Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 154 (2019) 199–205

Procedia Computer Science

www.elsevier.com/locate/procedia

8th International Congress of Information and Communication Technology, ICICT 2019

Q-Learning-Based Spinal Code Transmission Strategy in Long Distance FSO Communication

Jun Ao[a] , Na Li[a]*, Chun Bo Ma[a]

[a]*School of Information and Communication,Guilin University of Electronic Technology,Guilin,Guangxi,541004,China*

Abstract

In this paper, a spinal code transmission strategy based on Q-learning algorithm is proposed for long distance free space optical communication. The strategy adopts the Q-learning algorithm with the main objective of reducing the retransmission times to optimize the decoding waiting time and improve the transmission efficiency. The simulation results illustrate that, under the same turbulent conditions, the average retransmission times of the proposed transmission strategy is maintained at about one time, which is 94-99% less than the traditional spinal code transmission strategy and 78-92% less than the linear transmission strategy. Thus effectively improves the transmission efficiency of long-distance optical communication.

Keywords: FSO comunication , Spinal code, Q-learning, Interleaver, De-interleaver

1. Introduction

Rate-adaptive FSO systems can provide reliable communication in different weather conditions while efficiently exploiting the channel capacity.In long-distance FSO communication, in order to improve the availability of FSO channels degraded by atmospheric turbulence, the effects of channel gain variations must be compensated [1] [2].FEC technologies such as LDPC codes with excellent performance[3] are limited by the characteristics of fixed encoding rate.

* Corresponding author.Tel:13517530729.
E-mail address:1140042924@qq.com

10.1016/j.procs.2019.06.030

Rateless code is a kind of code without coding rate constraint. Due to its forward increasing redundancy, it can automatically adapt to the dynamic change of link without feedback and become an ideal means to make full use of channel transmission capacity under different weather conditions [4].Spinal code[5]is a novel rate-free encoding method capable of near-capacity limited transmission in both BSC channel and AWGN channel [6]. Although spinal code has good channel adaptability, in order to improve channel utilization and reduce unnecessary decoding waiting time, it is necessary to avoid transmission of unnecessary coding passes. Therefore, spinal code requires a bit rate control strategy to coordinate the transceiver and receiver to solve this problem [7] [8] [9] .This paper proposes a new spinal code transmission strategy based on Q-learning algorithm for long distance FSO communication applications, especially for space-earth communication with high transmission delay and low feedback retransmission efficiency.

2. Atmospheric Optical Channel

Defining a channel gain h, which takes into account both the effects of turbulence and attenuation including absorption and scattering, geometric loss, and insertion loss, the FSO channel can be described as

$$y = hx + n \tag{1}$$

where x is the OOK signal modulated at signalling rate R and n is additive white Gaussian noise (AWGN), with zero mean and variance σ^2 .

In this work, a packetized Spinal code is considered, where messages are grouped into packets for transmission over the channel. This practical technique avoids huge bitlevel processing which can limit the performance of highspeed links such as the underlying rate-adaptive FSO system.

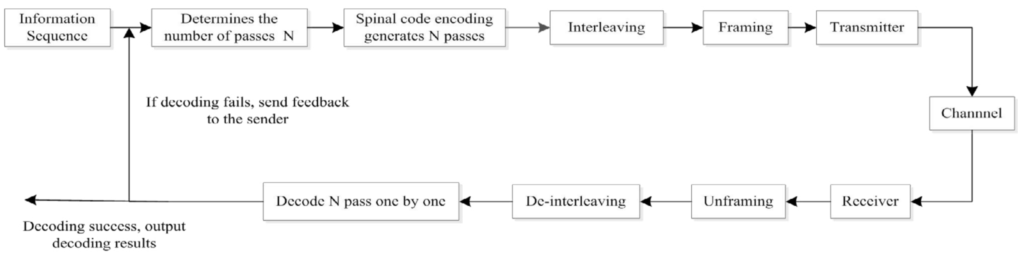

Figure 1: Block diagram of Spinal coded FSO communication system

3. Spinal Code Transmission Strategy Based on Q-Learning Algorithm

In this section, a spinal code transmission strategy based on Q-learning algorithm was proposed to dynamically adjust the number of passes. A block diagram of Spinal coded FSO communication system is illustrated in Fig.1.At the transmitter, spinal code coding is performed for an information sequence with a length of n bits. N passes are obtained according to Q-learning strategy. N encoding passes are modulated and sent after interleaving. According to the feedback, the sender continuously updates the Q-learning strategy and determines the number of coding passes to be sent next time.

After receiving N encoding passes, the receiver de-interleaves firstly, and then decode N passes one by one. If the decoding of one pass is successful, the decoding of this information segment is completed, and feedback is sent to the sending end to enter the encoding transmission of the next information segment. If the receiver fails to decode N passes, it also sends feedback to the sender and decodes them after receiving the encoding passes resent by the sender. Until the decoding is complete or the maximum numbers of retransmission is reached, the encoded to the next information segment. The spinal code transmission strategy implementation process based on Q-learning algorithm is show in figure 2.

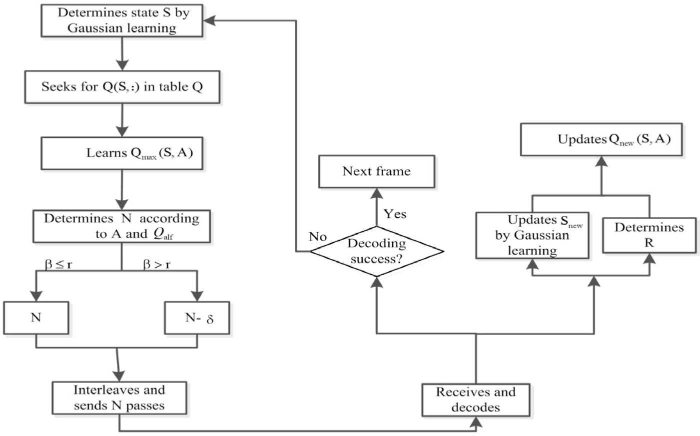

Figure 2: Spinal code transmission strategy implementation process based on Q-learning algorithm

3.1. Encoding spinal code

Spinal code encoding process is as follow. Firstly, information sequence M with length n bit is divided into n/k groups with length k bits, i.e. $M_0, M_2, ..., M_{n/k-1}$ .We denote it as M_i ,then, the sequence of states is simply $s_i = \text{Hash}(s_{i-1}, M_i)$, $s_0 = 0^v$ .Each of the n/k states ,or spine values with length v bits is used to seed a RNG(random number generator). Each RNG generates a sequence of pseudo-random number with length of c bits: RNG:$\{0,1\}^v \times N \rightarrow \{0,1\}^c$ .Denotes each sequence of pseudo-random as x_i ,then each batch of output symbols $\{x_1, ..., x_i\}$ constitutes an encoding pass, i.e. $X_j = \{x_{1,j}, ..., x_{i,j}\}$,where X_j represents the encoding pass of j-th .Encoding passes can be generated continuously multiple times by RNG and encoded symbols can be sent. When the encoding end receives a successful decoding feedback signal from the decoding end, or the encoding end decides to abandon the current information segment, the encoding finish.

The number of passes coding can be modeling by defining the decoding cumulative distribution function (CDF) at the sending end[4]. Studies have shown that Gaussian distribution should be a reasonable approximation $f(n)$ of decoding CDF. According to the number of passes required for successful decoding before, the mean μ and variance σ^2 of $f(n)$ can be learned through gaussian learning algorithm, so as to determine the current state of the channel. Finally, Q-learning strategy determines the number of coding passes that should be continuously sent by the current sending end according to the learned state. Q_a is gaussian learning factor with parameters ranging from 0 (no memory) to 1 (unlimited memory). The larger the value is, the greater the impact of previous moment decoding CDF on the estimation of current moment decoding CDF will be. Considering that the spinal code decoding requires a large number of coding passes in a large SNR range, the skewness Q_{alf} is introduced.

When spinal code transmission strategy based on Q-learning is implemented, first the state S of the channel is determined by gaussian learning, then the corresponding row Q(S,:) in Q table is found according to the state S, and finally the action A corresponding to the largest element $Q_{max}(S, A)$ in the row is found to obtain the number of coding passes that should be sent at present. The Q table stores the action a in state S and the cumulative reward Q(S,a) corresponding to a. The reward value is used to measure the cumulative reward that can be obtained by taking an action under a specific state S. It is composed of the expectation of immediate reward and long-term reward, and the weight between them is measured by Q learning factor η . η =0,ignoring long-term reward and thinking only about immediate reward while η =1 long-term reward are as important as immediate ones. In the establishment of Q table, according to the learned mean μ and variance σ^2 , the states were divided into n types, and the relationship between the number of coding channels corresponding to the actions was determined. The action performed according to the Q table is exactly to select the optimal number of coding channels N under the current state S, so as to maximize the cumulative reward.

The reward function R is defined as:

$$R: \; r(s,a) = -\gamma \frac{M_{orepass}}{M_{orepassMax}} - (1-\gamma)\frac{R_{etransTime}}{R_{etransTimeMax}} \tag{2}$$

where γ is used to measure the proportional relationship between the number of passes and the number of retransmission. The larger the value is, the more inclined it is to reduce the number of channels sent, and vice versa. $M_{orepass}$ is the number of passes currently sent, and $M_{orepassMax}$ is the maximum number of passes allowed to be sent. $R_{etransTime}$ is the current number of retransmissions, and $R_{etransTimeMax}$ is the maximum number of retransmissions allowed. In this paper, the main application scenario is satellite-to-ground optical communication.

In Q-learning, there is a certain probability to explore unknown learning, so an exploration factor β is introduced. The smaller the value of β, the greater the probability of executing action A and vice versa. After determining the optimal action A, it is necessary to determine the final number of sending as N or $N-\delta$ based on the value of β. After that, it is necessary to re-learn the mean μ and variance σ^2 of the current channel according to the number of selected coding channels N or $N-\delta$ to determine the new state S_{new} and new action A_{new}, and then update $Q_{max}(S,A)$ of the S row in Q table according to the row $Q_{max}(S_{new},A_{new})$ where the new state S_{new} is located and the reward function R under the current decoding state. The updating formula of the Q table is: $Q(s,a) \leftarrow r(s,a) + \eta \cdot \max_{a_{new}}(S_{new},A_{new})$, Where η is Q-learning factor.

3.2. Decoding spinal code

The receiver decodes N passes received after de-interleaving.Spinal code decoding is a maximum likelihood (ML) decoding algorithm. The basic decoding process is as follows: take s_0 as the root node, consider the possible value of $M_0, M_2, ..., M_{n/k-1}$ in order to obtain 2^n species possibilities, then traverse from the root node to the leaf node, and the leaf node with the minimum total decoding path overhead is the decoding result with the ML.

ML decoding requires traversing 2^n possibilities, which is very complex and greatly reduces the probability of correct decoding. The literature[6] has proved that alternate codewords with comparable path costs to the ML decoding differ only in the last $O(\log n)$ bits. According to this theoretical, a ML decoding algorithm based on pruning the tree is obtained and only B nodes with the lowest decoding cost are retained in each level .

3.3. Interleaving technique

Interleaving technique is considered to solve the sudden channel errors in transmission and improve the probability of successful decoding. The international telecommunication union in 3GPP standards recommended by 3GPP standard interleaver[10], it has important effect to improve the reliability of data transmission. The basic process is: First, the n-bit information segment is written to a matrix of size $R \times C$ on a row basis, then interleaving between rows and interleaving within rows to obtain a new matrix, finally The new matrix is pruned and the output bitstream is read in columns.

4. Performances Analysis

The spinal code transmission strategy based on Q-learning, the linear transmission strategy, and the traditional spinal code transmission strategy were compared with the non-turbulent and with turbulent (weak turbulence,

σ_I^2 =0.1[11]) channel under the premise of interleaving. The simulation parameters are set as follows : $M=64\text{bits}, K = 8\text{bits}, B = 3, C = 16$; $Q_a = 0.8, \eta = 0.4, \gamma = 0.1, \beta = 0.11$.

In the simulation, according to the above analysis, reducing the number of retransmission is the main objective, and $\gamma = 0.1$. Figure 3 respectively show the number of retransmission times and the number of corresponding passes required for successful decoding of the three transmission strategies under the difference turbulence condition

As can be seen from figure 3 ， the average number of retransmission for successful decoding of spinal code transmission strategy based on Q-learning is about 1 times, which is far less than that required by the other two strategies and relatively stable as a whole. So it is concluded that Q-learning strategy requires a large average number of passes, but it significantly reduces the number of retransmission, which is 94-99% lower than the traditional spinal code transmission strategy. There is 78%-92% retransmission numbers less than the linear spinal code transmission strategy. For the long-distance optical communication system with strict requirements on time delay, it is of great significance to reduce the decoding delay and interrupt probability of the system by sacrificing some transmission resources within the acceptable range to improve the reliability of communication.

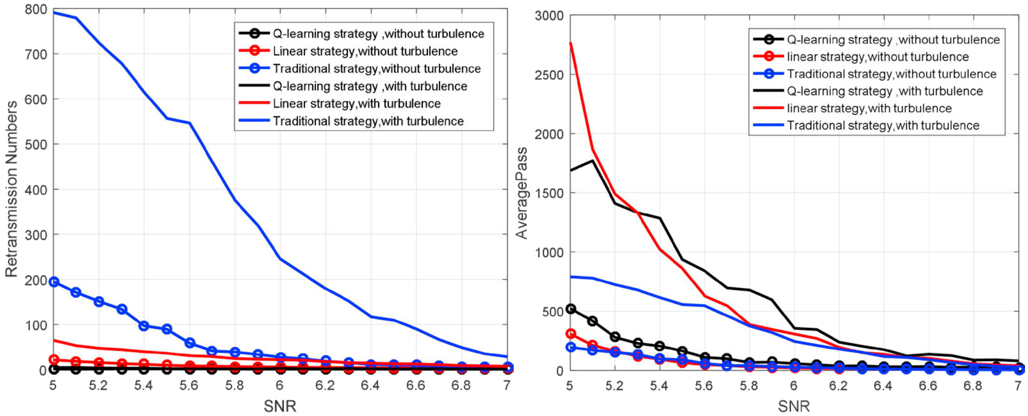

Figure3. The number of retransmissions and passes in different turbulent conditions

Spinal code transmission strategy based on Q-learning algorithm is described in figure 4 to figure 5 with respect to two parameters: β (indicated by beta in figure note) and Q_a .

β represents the probability of leaving the learned optimal value and exploring other values in Q-learning. The smaller the value is, the more inclined it is to choose N. It can be seen in figure 4, the smaller the β is, the more inclined it is to reduce the number of retransmission, and the overall change trend is more stable. So in the Q-learning strategy with reducing the number of retransmission as the main goal, the smaller value should be selected.

As can be seen from figure 5, different Q_a under different channel conditions have different influences on the number of retransmission times and coding channels required for the final successful decoding. In the case of no turbulence channel, the smaller Q_a is, the more likely it is to reduce the number of retransmission times and the corresponding number of coding channels is increased. In turbulent channel, the higher the value of Q_a , the less retransmission times are needed. So the value of Q_a should be determined according to the specific application and requirements.

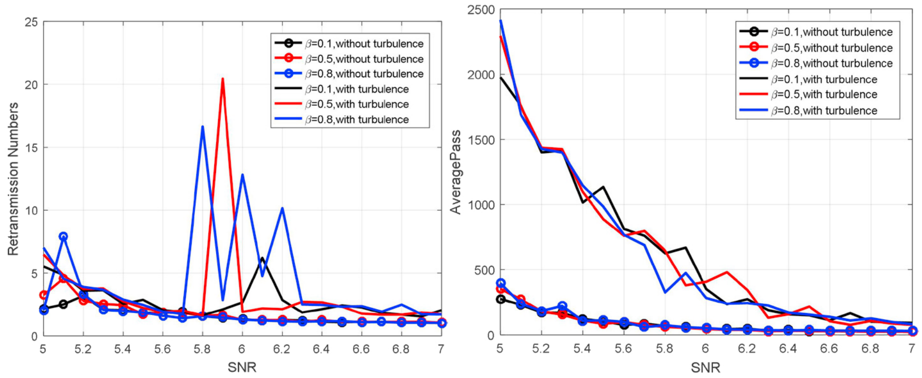

Figure 4: Influence of different β in different turbulent conditions

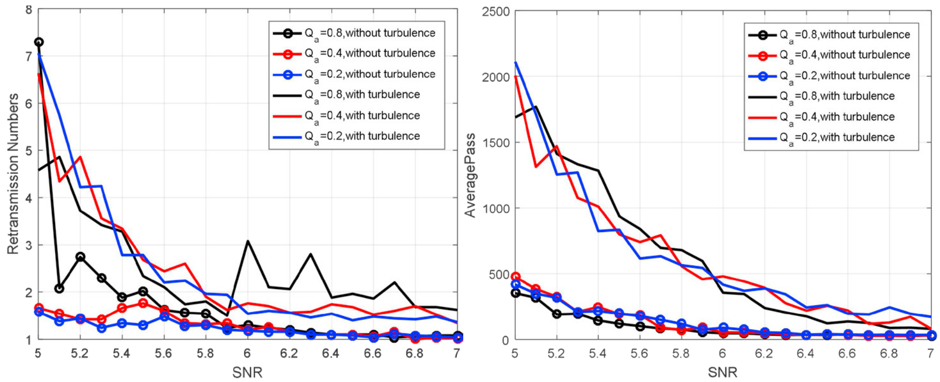

Figure 5: Influence of different Q_a in different turbulent conditions

5. Conclusion

In this paper, the performance of a Q-learning strategy -based spinal-coded FSO communication systems was investigated over different turbulent conditions. The results show that using different interleaves for different retransmissions our Q-learning strategy-based spinal coded technique can provide significantly higher throughputs particularly in turbulence conditions.

6. References

1. Akyildiz I F , Akan O B , et al. The state of the art in interplanetary Internet[J]. IEEE Communications Magazine, 2004, 42(7):108-118.
2. Hemmati H . Deep Space Optical Communications [J]. 2006,231-232.
3. Ao J , Liang J , Ma C , et al. Optimization of LDPC Codes for PIN-Based OOK FSO Communication Systems[J]. IEEE Photonics Technology Letters, 2017, 29(9):727-730.
4. Bursalioglu O Y , Caire G , Divsalar D . Joint Source-Channel Coding for Deep-Space Image Transmission using Rateless Codes[C]// Information Theory and Applications Workshop (ITA), 2011. IEEE, 2011.
5. Perry J, Iannucci P A, Fleming K E, et al. Spinal Codes[C]//Proceedings of the ACM SIGCOMM 2012 Conference on Applications, Technologies, Architectures and Protocols for Computer Communication, Helsinki, 2012: 49-60.
6. Perry J, Balakrishnan H, Shah D. Rateless Spinal Codes[J]. Association for Computing Machinery, 2013: 1-5
7. Iannucci P A, Perry J, Balakrishnan H, et al. No symbol left behind: a link-layer protocol for rateless codes[C]. Proceedings of the 18th annual international conference on Mobile computing and networking. ACM, 2012: 17-28.
8. Ao Jun, Liang Jiwei, Ma Chunbo, et al. Q-Learning-Based Raptor10 Code Decoding Strategy in Free-Space Optical Communication[J]. Chinese Journal of Lasers, 2017, 44(9):0906007.

9. Sheng Haiyang. Research on the spinal code based transmission mechanism for deep space communication[D]. Harbin Institute of Technology,2014.
10. 45.003, 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; GSM/EDGE Channel coding[S].211-214.
11. Xarcha V, Stassinakis A N, Nistazakis H E, et al. Wavelength diversity/or Free Space Optical Systems: Performance evaluation over log normal turbulence channels[C]. Microwave Radar and Wireless Communications (MIKON),2012 19th International Conference on. IEEE, 2012, 2: 678-683.