強化学習を用いた動的な通信環境での符号化制御

路　　　姍*

Reinforcement Learning Based Code Rate Control in Dynamic Communication Environment

Shan LU*

The error correction codes add some extra data to a message used to recover the corrupted data in the communication system. The code rate is the ratio of the original data length and the transmitted data length and varies to grantee the transmission performance in the communication environment. The switches of the code rates in the dynamic channel environment cause a considerable transmission delay. Therefore, this research project adopts a code rate control algorithm based on Reinforcement Learning (RL) approach to optimize the average transmission rate to balance the time delay and transmission performance.

1. Introduction

In the communication system, communication channels are subject to channel noise during the transmission from the source to a receiver, and thus errors are introduced to the transmission data. The error correction techniques add redundancy data to the message to detect such errors, while error correction enables the reconstruction of the original corrupted data. The code rate is the ratio of the original data length and the transmitted data length, which usually switches based on the channel environment. However, in the dynamic channel environment, the frequent switches of the code rates cause a considerable transmission delay. Therefore, the objective of this research is to balance the transmission time and performance.

Therefore, we propose a code rate control algorithm based on RL in this research. The main idea is to optimize the average transmission rate by decreasing the number of switching times of code rate at a required performance level, that is, to maintain a constant code rate at a specific transmission time. The problem is how to find the optimal constant code rate in the dynamic environment.

First, we define a Markov decision process (MDP) model for the transmission model. Then, to get the optimal code rate, we use supervised learning to train a specific code rate selection Q-table, representing the expectation of the future reward for performing a code rate under a particular state. Then, a reinforcement learning algorithm is developed to find the optimal selection of coding rate. Eventually, the simulation result shows that the proposed code rate selection algorithm significantly improved the average transmission rate in a dynamic environment.

2. Reinforcement Learning based Code Rate Control

Figure 1 shows a framework that incorporates a communication system and a RL algorithm to select the optimal code rates for a dynamic environment. The environment part is the communication system with a channel encoder and decoder and a dynamic channel. We categorize the dynamic channel based on characteristics and parameters into three classes: high, medium, or low dynamics channels and define the specific transition probabilities. Then, we obtain an MDP model of the transmission model based on the dynamic channel model.

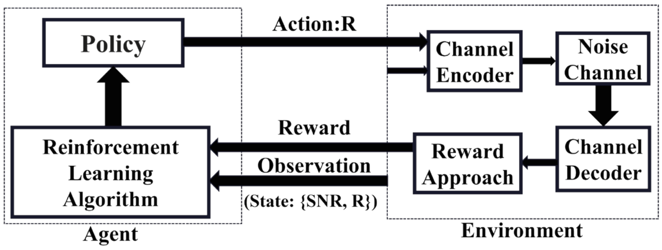

Fig. 1 Framework of code rate selection based on RL in dynamic environment

2022年3月4日 受理
* 豊田理研スカラー
岐阜大学工学部電気電子・情報工学科情報コース

The agent aims to provide an optimal code rate to the environment. Supervised learning is applied to learn a Q-table on the defined MDP model. In training, the environment part first feedbacks a reward to the agent for each transmission based on the block error and code rates. Then, the agent applies the proposed reinforcement learning algorithm to efficiently update the policy to provide optimum code rate selection under the current environment. The learned Q-table represents the future reward's expectation for performing a code rate under a particular state, which shows the optimal code rate selection to cope with the dynamic environment.

After training, based on the Q-table, the agent controls the code rate of the transmission part under the current transmission environment. The encoder switches the code rate and sends the codeword to the channel and decoder after receiving the new code rate from the agent.

3. Simulation results

The LDPC code [1] with a codelength of bits is used for error correction in the communication system. In the training proceeding, 12000 blocks LDPC codewords are transmitted for training. After training, 10000 blocks LDPC codewords are transmitted for evaluating the performance. Let the transmission time per block be 1 *ms*. The code rate switch costs 5 or 10 *ms*.

Table 1 The comparison between code rate control with RL and without RL.

	With RL	Without RL	With RL	Without RL
Code rate switch cost (*ms*)	5		10	
Number of switch time	5020	22762	5020	22762
Switch time delay (*ms*)	25100	113810	50200	227620
total transmission time (*ms*)	125100	213810	150200	327620
average transmission rate (*ms*)	609	377	504	246

Table 1 shows the number of block switches, total transmission time, and average transmission rates using the leaned Q-table. For comparison, the scheme results without using a Q-table are also shown. The table shows that the time delay obtained by using the Q-table is much smaller than without using Q-table. Therefore, the code rate switching algorithm can avoid frequent and quick switching of the code rate for the wireless communication system in the same period. Furthermore, the average transmission rate obtained using the Q-table is much larger than without Q-table. Therefore, we have that the proposed code rate switching algorithm can significantly improve the average transmission rate.

4. Conclusions

We proposed a code rate control algorithm based on RL in the dynamic channel environment to balance the transmission time and performance. The proposed code rate switching algorithm based on the training Q-table can significantly improve the average transmission rate.

謝辞

本研究は公益財団法人豊田理化学研究所の豊田理研スカラー助成によって行われたものであり，ご支援いただきました公益財団法人豊田理化学研究所にこの場を借りて深く感謝申し上げます．

REFERENCES

1) R. Gallager, “Low-density parity-check codes,” *IRE Transactions on Information Theory*, **8** (1962) 21-28, doi: 10.1109/TIT.1962.1057683.