LSTM versus simple RNN and FNN in MC-DLMA design for this paper:

MAC Protocol for Multi-channel Heterogeneous Networks Based on Deep Reinforcement Learning

Xiaowen Ye, Yiding Yu, Liqun Fu

I. Introduction

This is a supplementary document to the paper: MAC Protocol for Multi-channel Heterogeneous Networks Based on Deep Reinforcement Learning. In this document, we compare the performance of LSTM with simple RNN and FNN in MC-DLMA design.

A. Deep Neural Network

As shown in Fig. 1, when the feedforward neural network (FNN) [1] is used as DNN, each neuron is independent of each other, and the output Q_t depends only on the input state s_t at each time slot t. However, in DSA, the dynamics of the combination of all channels are non-Markovian and determined jointly by nodes using different protocols, thereby FNN performs not well in addressing the DSA issue. Unlike FNN, recurrent neural network (RNN) [2] is in possession of memory ability—all the neurons of RNN are connected circularly, and the input of each neuron is not only related to the input state s_t, but also to the output h_{t-1} of the neurons from previous time slots. Fig. 1 shows that, the input state s_t will bind h_{t-1} to form a new state, then this new state will be entered into all the neurons to generate Q_t. As a result, RNN is capable of learning the correlation for each channel on different time slots, by fusing channels' observations over time.

A limitation of simple RNN (i.e., the most basic RNN) is that its memory is short-term. Overlong state sequence s_t will make the gradient $\nabla L_{N_{\mathrm{E}}}(\boldsymbol{\theta})$ in (1) too larger or too small, resulting in the problem of vanishing or exploding gradient [3].

$$\nabla L_{N_{\mathrm{E}}}(\boldsymbol{\theta}) = \mathbb{E}_{i\in B}\Big(r_{i+1} + \gamma \max_{a'} Q(s_{i+1},\, a';\, \boldsymbol{\theta}_{-}) - Q(s_i,\, a_i;\, \boldsymbol{\theta})\Big)\nabla Q(s_i,\, a_i;\, \boldsymbol{\theta}). \tag{1}$$

As a solution, long short-term memory (LSTM) [4] is developed on the basis of simple RNN. To preserve memory for long state sequences, LSTM selectively remembers important memory and forgets non-significant memory by employing a state control unit c. As presented in Fig. 1, at each time slot t, LSTM analyzes both the input state s_t from the environment and the output h_{t-1} of neurons in the last time slot, and then remembers only useful knowledge, rather than all knowledge like simple RNN. Selective memory makes LSTM more closer to the human brain in making judgements, thereby this approach performs well in dynamic processing and natural language processing [5]–[7]. In [8], LSTM is used to predicte the photovoltaic power, while the authors of [9] applied it to classify arrhythmias. The results of aforementioned works are promising. In this paper, we adopt LSTM as DNN in DQN, which can solve the DSA problem by reasoning the temporal correlation of multi-channel HetNets.

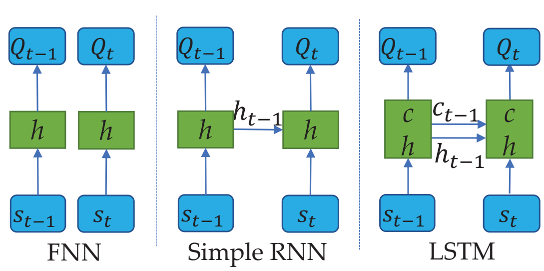
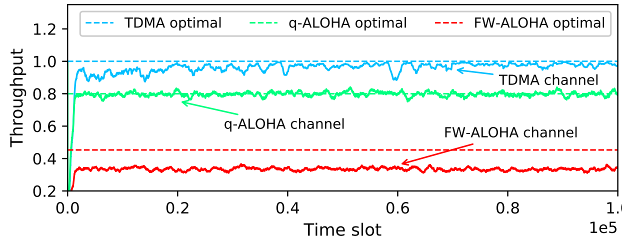
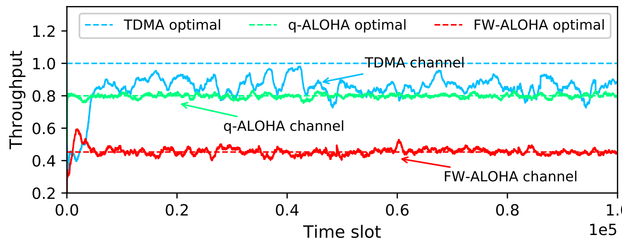


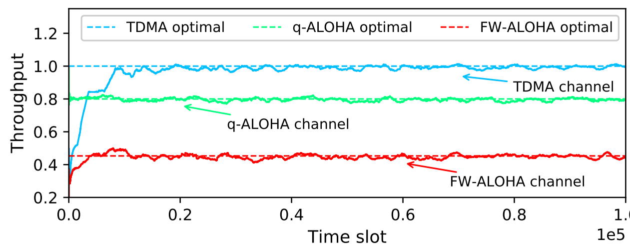
Fig. 1. Feedforward neural network (FNN), simple recurrent neural network (RNN) and long short-term memory (LSTM).



(a) The MC-DLMA protocol using FNN.



(b) The MC-DLMA protocol using simple RNN.



(c) The MC-DLMA protocol using LSTM.

Fig. 2. Individual throughputs of each channel when the MC-DLMA node using different DNNs coexists with one TDMA node, one q-ALOHA node and one FW-ALOHA node.

B. Performance Evaluation

This section demonstrates that the advantages of using LSTM as DNN in multi-channel HetNets over FNN or simple RNN as DNN. In particular, we consider the coexistence of one MC-DLMA node and a mix of one TDMA node, one q-ALOHA node and one FW-ALOHA node, wherein the TDMA node transmits 3 time slots out of 10 time slots within a frame, the transmission probability of the q-ALOHA node is 0.8, and the contention window size of the FW-ALOHA node is 5. We compare the individual throughputs of each channel when the MC-DLMA node uses one of these three DNNs (i.e., FNN, RNN, and LSTM), respectively. For FNN, we use the ResNet [10] with 64 neurons per hidden layer, which has been proved to perform well in single channel scenarios [11]. Meanwhile, when we delve into the performance of the MC-DLMA node using RNN, we directly replace LSTM with RNN. For the performance upper bound, we replace the MC-DLMA protocol with the model-aware policy that has the global network information to coexist with other nodes.

As we can see from Fig. 2 (a), FNN performs poorly due to the lack of the ability to aggregate the observation information of multi-channel over time slots. In particular, when the MC-DLMA node uses FNN as DNN, the throughput of both the TDMA channel and the q-ALOHA channel approximates the optimal benchmark, whereas the throughput of the FW-ALOHA channel is much lower than the optimal benchmark. On the other hand, although simple RNN can learn the temporal correlation of multi-channel HetNets in different time slots, the learning process of the MC-DLMA node is not stable enough because it is not good at processing long input sequences. Particularly, the throughput of both the q-ALOHA channel and the FW-ALOHA channel is close to the optimal value, but the throughput on the TDMA channel cannot converge: after the throughput reaches the upper bound in the 42000-th time slot, it quickly drops and then fluctuates continuously, as shown in Fig. 2 (b). Unlike FNN and simple RNN, LSTM enables the MC-DLMA node to maximize the throughput of each channel in the network, as presented in Fig. 2 (c). This is because LSTM can selectively aggregate the channel observations over time slots, and then effectively employ the temporal correlation of the observation information to make a decision in each time slot. This experimental results clearly demonstrate that compared with FNN and simple RNN, LSTM with selective memory is more power powerful in solving DSA problems in multi-channel HetNets.

Xiaowen Ye and Liqun Fu are with the School of Information, and the Key laboratory of Underwater Acoustic Communication and Marine Information Technology Ministry of Education, Xiamen University, Xiamen, China. Email:{xiaowen@stu.xmu.edu.cn, liqun@xmu.edu.cn}

Yiding Yu is with the Department of Information Engineering, The Chinese University of Hong Kong, Shatin, Hong Kong Special Administrative Region, China. Email:yy016@ie.cuhk.edu.hk

References

[1] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press, 2016.

[2] T. Mikolov, M. Karafiát, L. Burget, J. Černockỳ, and S. Khudanpur, “Recurrent neural network based language model,” in *Eleventh annual conference of the international speech communication association*, 2010.

[3] Y. Bengio, P. Simard, P. Frasconi *et al.*, “Learning long-term dependencies with gradient descent is difficult,” *IEEE transactions on neural networks*, vol. 5, no. 2, pp. 157–166, 1994.

[4] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[5] D. Soutner and L. Müller, “Application of lstm neural networks in language modelling,” in *International Conference on Text, Speech and Dialogue*. Springer, 2013, pp. 105–112.

[6] T.-H. Wen, M. Gasic, N. Mrksic, P.-H. Su, D. Vandyke, and S. Young, “Semantically conditioned lstm-based natural language generation for spoken dialogue systems,” *arXiv preprint arXiv:1508.01745*, 2015.

[7] S. Merity, N. S. Keskar, and R. Socher, “Regularizing and optimizing lstm language models,” *arXiv preprint arXiv:1708.02182*, 2017.

[8] M. Abdel-Nasser and K. Mahmoud, “Accurate photovoltaic power forecasting models using deep lstm-rnn,” *Neural Computing and Applications*, vol. 31, no. 7, pp. 2727–2740, 2019.

[9] O. Yildirim, U. B. Baloglu, R.-S. Tan, E. J. Ciaccio, and U. R. Acharya, “A new approach for arrhythmia classification using deep coded features and lstm networks,” *Computer methods and programs in biomedicine*, vol. 176, pp. 121–133, 2019.

[10] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[11] Y. Yu, T. Wang, and S. C. Liew, “Deep-reinforcement learning multiple access for heterogeneous wireless networks,” *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 6, pp. 1277–1290, 2019.