Title	Deep Reinforcement Learning for Wireless Multihop Networks: Factor Graph and Nested Lattice Approaches
Author(s)	崔, 志瀚
Citation	
Issue Date	2023-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/18313
Rights	
Description	Supervisor: リム 勇仁, 先端科学技術研究科, 修士(情報科学)

Deep Reinforcement Learning for Wireless Multihop Networks: Factor Graph and Nested Lattice Approaches

2110078 CUI Zhihan

A Wireless Multihop Network (WMN) is set of wirelessly connected nodes without an aid of centralized infrastructure that all the nodes process cooperatively and forward any packet via relaying nodes by multihop fashion. With the direct communications of the multihop fashion, device-to-device communication can extend the range of the transmissions. Using other nodes as the relay nodes, the packets can be transmitted from a source node to a destination node over longer distance. Thus, WMN not only can extend the network coverage, but also can achieve better network capacity by using the relay nodes strategy. Compared with traditional wireless networks, the WMN can use each user in the network as an access point to transmit data without a base station. This will greatly reduce the cost of building base station for the operating company and is expected to be applied to 6G networks.

Using other nodes as the relay nodes, the messages can be transmitted from a source node to a destination node over longer distance. Thus, WMN only can extend the network coverage, but it can degrade the network capacity drastically due to the uncertainty of source node choosing a path to send the message and also the nature of multihop fashion. In particular, the first problem is happened when the number of nodes in the network is increasing, there are many data transmission paths to be selected from a source node to the corresponding destination node by an efficient path selection algorithm. Since the criteria and conditions of each path are different, the result of the path selection has a great influence on the entire network capacity. Second problem is occurred when a node has to send not only its own message, but also other nodes' messages in the network, in which it can lead to high latency for the sent message to reach its destination node. Besides that, the queuing time and processing time will increase drastically when when the number of nodes is large.

To mitigate these two problems, we consider deep rein- forcement learning (DRL) with both factor graph (FG) and nested lattice code (NLC) approaches in our research study. For first aforementioned problem, we apply FG-based DRL (fDRL) scheme to select a relay node for each hop and find best path to the destination node. With these paths, a resultant of network topology that can achieve a better network capacity can be established. For the second problem, NLC that is used to reduce the link error probability of a channel can correct the errors through the integration of computer-and-forward (CoF) strategy. By the way, the entire network capacity can be improved more by reducing the required number of time slots. Besides that, We also propose two novel learning path selection algorithms, i.e., SNR-based learning path selection (NLPS) algorithm focuses on increasing the end-to-end throughput from source node to destination node considering SNR as reward of DRL; and SINR-based learning path selection (INLPS) algorithm uses the results obtained from NLPS and considers SINR as reward during the training phase of DRL to find the appropriate path for any source node.

With the development of hardware devices, the computing speed of processors has increased. Even Machine Learning (ML) that requires a lot of training and a lot of calculations is widely used in various fields, including wireless networks. Machine learning can solve a wide variety of problems. Among them, Reinforcement Learning (RL) is a kind of ML that changes the action in real time according to the environment, and continuously makes choices in training with the strategy of maximizing rewards, which is the most suitable for solving the path selection problem. In RT, Q-learning is a model-free RL algorithm to learn the value of an action in a particular state. It does not require a model of the environment, and it can handle problems with stochastic transitions and rewards without requiring adaptations, which can be applied in WMN.

However, Q-learning requires a large number of iterations, which costs a long computation time. FG is introduced to solve this problem. A FG is a bipartite graph representing the factorization structure of a global function into a product of smaller local functions, each local function contains the product from other factor. Using the sum-product algorithm, the global function can represent the whole FG, which can determine the best network capacity from a tree-based structure topology. Using FG to preprocess the data for Q-learning and filter out structures with low network capacity can greatly shorten the training time. In addition, Nested Lattice Code (NLC) applied to encode and decode the messages and Compute-and-Forward (CoF) strategy is used to send data in the transmission phase to reduces the number of time slots, thereby further increasing network capacity.

Therefore, the objectives of this thesis are to examine the performances of network capacity and computation time in WMN with FG and NLC; and to propose novel Factor-Graph based Deep Reinforcement Learning (fDRL) scheme with two learning path selection algorithms for the Q-Learning to achieve better network capacity while reducing the computation time with the aid of FG.

First, this thesis proposes system models including network model, channel model, interference model and network capacity model. Based on these models, fDRL scheme with two learning path selection algorithms is proposed: SNR-based Learning Path Selection (NLPS) and SINR-based Learn-

ing Path Selection (INLPS). These two algorithms are improved based on Q-learning, making them more suitable for solving the problems presented in this thesis. The NLPS algorithm takes the SNR of the link as a reward, starting from the source node, considering the state to select the path and update the information until reaches the destination node. The path selected by NLPS has high SNR, which means the end-to-end (E2E) throughput is high. Based on the results of NLPS algorithm, INLPS algorithm calculates the SINR of the link as a reward and select the path. The path selected by INLPS has higher SINR and thus higher E2E throughput. To reduce the number of iterations of two algorithms, FG preprocesses the data to select the root node for the algorithms. In the transmission phase, according to the network topology formed by the algorithm, the NLC with CoF strategy pair the nodes and combine the two signals to encode and send, which increase the average network capacity.

According to the theoretical and numerical studies with the assumption of the system model, simulation results reveal that FG can reduce computation time up to 99% for algorithms. As for two proposed algorithms, compared to not using the path selection algorithm, NLPS and INLPS increase the network capacity. When there are 50 nodes, the average network capacity of INLPS (2.52 Mbps) is about 3.19 times higher than NLPS (0.79 Mbps) and 6.46 times higher than use FG and NLC (0.39 Mbps). When the number of nodes is 100, the values become 1.83 Mbps, 0.43 Mbps and 0.36 Mbps respectively and INLPS algorithm is about 4.26 times higher than NLPS and about 5.08 times higher than use FG and NLC. For number of iterations, NLPS uses 101 iterations to reach 98% of highest E2E throughput and INLPS uses 185 iterations when the number of nodes is 50. These two algorithms uses 156 iterations and 263 iterations respectively when the number of nodes is 100.

Based on the result, the thesis concludes that FG can reduce the computation time for algorithms before training. With the aid of FG, NLPS and INLPS can find best path for each node in a shorter time. Both NLPS and INLPS can increase network capacity. The difference is that INLPS increase more but NLPS cost a small number of iterations. With the aid of NLC, the network capacity is increased further.

Keywords: Deep Reinforcement Learning; Factor Graph; Nested Lattice Code; Learning Path Selection Algorithm; Wireless Multihop Networks