Deep Reinforcement Learning Based Green Resource Allocation Mechanism in Mobile Edge Network for Ubiquitous Power IoT

Mo YANG¹,², Xiu-li HUANG², Wei-wei MIU³, Peng-fei YU², Wei Li³ and Ru-xia YANG²

¹State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing, China
²Global Energy Interconnection Research Institute Co., Ltd, Beijing, China
³Communication branch, State Grid Jiangsu Electric Power Co., Ltd, Jiangsu, China

*Corresponding author

Keywords: Deep Reinforcement Learning, Green Resource Allocation, Mobile Edge Network.

Abstract. Ubiquitous Power IoT is an important part of smart grid or energy Internet. Edge network is a key component of smart grid and future energy Internet. IoT contains various terminals, diverse access networks and equipment. The allocation of energy-efficient resources under QoS guarantee is an effective means to ensure the reliable operation of the network. Due to the heterogeneity of the services carried and in order to reduce business operation risks and impact on the power grid, it is necessary to provide highly reliable resource allocation method to the ubiquitous power IoT business for the energy Internet. Edge network can reduce end-to-end latency and reduce backhaul link traffic. However, mobile edge networks moving the function of storage and computing down to the edge nodes of network, which makes resource management more complex. Moreover, the computing resources in the edge network is limited and the effect to the ecological environment and economic cost should also be considered. So how to allocate resources such as bandwidth and power more effectively while meeting the needs of users becomes an important issue. Even if the Deep Reinforcement Learning (DRL) algorithm has been used to a lot of the work which are correlated to edge networks, there lacks the applications for green resource allocation. In this paper, a mechanism based on Deep Reinforcement Learning (DRL) for the resource allocation problem is proposed oriented to edge networks. The mechanism aims at how to allocate resources more efficiently and energy-saving while satisfying the requirements of each user. The simulation results show that the energy efficiency value could be obtained when the algorithm converges and stabilizes. The Energy Efficiency (EE) of the proposed mechanism and the productiveness in satisfying user requirements and implementing green resource allocation are validated.

Introduction

The ubiquitous technology in the network has enabled the Internet of Things to be used in all aspects of production, dissemination and consumption, and the full penetration of the service industry. The ubiquitous power infrastructure of the Internet of Things is derived from full-time staff and national policies and industries. Management, energy management and business users. [1]. The technology of wireless communication has a rapid development in recent years. For the purpose of meeting the higher demand of smart grid for latency and effectively relieve backhaul link pressure, the academic and industrial research on mobile edge networks. [2].
Ubiquitous Power IoT is an important part of smart grid or energy internet. IoT contains various terminals, diverse access networks and equipment. The allocation of energy-efficient resources under QoS guarantee is an effective means to ensure the reliable operation of the network. Due to the heterogeneity of the services carried and in order to reduce business operation risks and impact on the power grid, it is necessary to provide highly reliable resource allocation method to the ubiquitous power IoT business for the energy Internet.

Paper [3] proposes a fair resource allocation method to maximize overall network throughput under the constraints of the minimum transmission rate of each mobile user. By using time-sharing variables, the paper obtains an approximate optimal bargaining resource allocation strategy for mixed integer nonlinear programming optimization. However, this algorithm does not consider energy consumption and energy efficiency, which may result in large cost loss and energy consumption. Deep learning (DL) has achieved a number of exciting achievements [4]. Notably the combination of deep learning (DL) and reinforcement learning, i.e. deep reinforcement learning (DRL)[5]. The current studies on the DRL algorithm has achieved important success in a lot of aspects, including application to MEC. However, more work is needed to achieve green resource allocation by using DRL.

In the paper, a DRL-based edge network green resource allocation framework is proposed. Our main contribution is summarized in the following:

- Because the computing resources of the edge network is limited, a green resource allocation framework based on DRL is proposed.
- With the goal of minimizing energy efficiency while meeting the needs of each mobile user, the DRL agent defined the space of state, the space of action, and reward function

System Model

The paper considered the downlink transmission situation in a SDN-enabled heterogeneous network (HetNet) which consists of a set of Base Station (BS) \( I := \{1, ..., i, ..., I \} \) and a set of users \( J := \{1, ..., j, ..., J \} \) and a core network \( N \). The set of BSs \( I \) as well as the core router (core network) \( N \) are connected using the wired backhaul links[6]. The transmission between the user's device \( j \) and the BS \( i \) are via wireless links. Fig. 1 shows the scenario.

In the considered network, a set \( F := \{1, ..., f, ..., F \} \) of flows are running in. Each of the flows has a required packet size \( m_f \) and required data rate \( r_f \) and the transmission time is

\[
 r_f = \frac{m_f}{t_0} \quad (1)
\]

Assuming one link supports one flow, \( r_{ij}^f \) is the data rate of flow \( f \) between one of the BSs and one of the users.
The connection relationship between the BS $i$ and the user $j$ is expressed as $s_{ij} \in \{0,1\}$, $s_{ij} = 1$ indicating that the user $j$ is served by the BS $i$, $s_{ij} = 0$ indicating that the user $j$ has no connection relationship with the base station $i$ [7].

The transmitted energy depends on the wireless transmission energy and the backhaul transmission energy. $t_0$ is the operation time. The wireless transmission power is expressed as $p_{ij}$, $p_{ij}$ represents the power allocated by the BS $i$ to the user $j$.

Among them, the wireless transmission energy is:

$$E_T(p_{ij}, s_{ij}) = \sum_{i \in I} \sum_{j \in J} t_0 p_{ij} s_{ij}$$

(3)

Thus, the energy efficiency is:

$$E_b = \frac{E_T(p_{ij}, s_{ij})}{\sum_{f \in F} r_f}$$

(4)

The optimization problem is illustrated as follows:

$$\min_{S, P} E_b$$

(5)

Subject to:

$$s_{ij} \in \{0,1\} \quad \forall \ i \in I, j \in J$$

(6)

$$\sum_{i \in I} s_{ij} = 1 \quad \forall j \in J$$

(7)

$$\sum_{j \in J} p_{ij} s_{ij} \leq P_{\text{max}}, \forall i \in I$$

(8)

$$\sum_{j \in J} x_{ij} s_{ij} \leq 1 \quad \forall i \in I$$

(9)
\[ \text{SINR} \geq \phi \quad \forall i \in I, j \in J \quad (10) \]

In the constraint (6-7), \( s_{ij} \in \{0,1\} \) is a binary decision variable that indicates whether the user is connected to a certain BS. One user could only be served by one BS at one time. The constraint (8-9) reflects that for any one BS, its total transmit power and allocated bandwidth cannot exceed the total power and total bandwidth that it can provide. The constraint (10) reflects the user’s SINR constraint.

**DRL-based Green Resource Allocation Framework**

In this part, a framework for green resource allocation for edge networks based on DRL is proposed. The goal is to minimize the energy efficiency of the edge network while meeting the needs of each user and not exceeding the maximum power and bandwidth load of each base station.

The state space, action space as well as the reward function of the DRL-based framework are defined as follows:

- **State Space:** If the number of users is \( K \), then \( \mathcal{M} = \{s_1, s_2, \ldots, s_K\} \) can be used to indicate a certain connection relationship between all users and the base station. So our state space can be expressed as \( S = \{\mathcal{M}_1, \mathcal{M}_2, \ldots, \mathcal{M}_N\} \), where \( N = 3^K \) represents the total number of connection relationships.

- **Action Space:** The action space is defined as \( A = \{1,2, \ldots, N\} \).

- **Reward:** The instant reward is formulated as \( E_{\text{max}} - E \), where \( E_{\text{max}} \) indicates the maximum energy consumption value, \( E \) indicates the energy consumption value after taking this action.

The DRL based green resource allocation framework is described in Algorithm 1.
Simulation Results

Table 1. Simulation Parameters.

| Parameter          | Value  |
|--------------------|--------|
| Deployment         | 3 BSs  |
| Max BS Power       | 3W     |
| Packet of flows    | 10M    |
| Processing time    | 1s     |
| Max bandwidth      | 20MHz  |
| Bandwidth ratio    | 0.2    |
| \( \phi \)         | -3dB   |

The DRL based green resource allocation framework in edge network in this paper includes three base stations and several users. The parameters are as shown in Table 1. The approximate values of steps and corresponding results when DRL converging in the four cases are shown in Table 2.
Table 2. Approximate value for convergence.

| Users  | Converge Step | Reward | Loss  | Energy Efficiency |
|--------|---------------|--------|-------|-------------------|
| 6 users| 500           | 21.76  | 0.163 | 3.772             |
| 5 users| 350           | 21.09  | 0.114 | 3.456             |
| 4 users| 250           | 20.38  | 0.082 | 2.965             |
| 3 users| 220           | 20.25  | 0.035 | 2.207             |

Next, six users are used to compare the energy efficiency values and power consumption values under three different strategies. The DA indicates that the user selects the nearest base station to receive the service, and the DRL indicates the value of the DRL-based edge network green resource allocation framework mentioned in this paper when converging. The UC indicates that any two users are clustered and assigned to one of the base stations for service. It can be seen from the figure that the energy efficiency value and power consumption value of DRL are the lowest, followed by the strategy of selecting the nearest distance (DA), and the energy efficiency value and power consumption value of the method of any two user clustering (UC) are the highest. The results are shown in Fig. 2.

Figure 2. Comparison of energy efficiency and power consumption under three strategies.

Summary

The DRL-based framework proposed in the paper solves the problem of green resource allocation in edge networks. Minimizing energy efficiency while ensuring the needs of each user. A well-trained network can achieve the goal of green energy-saving, also solve the effective resource allocation problem, and the results achieved the lowest energy efficiency compared with the other two strategies. However, in the article, the simulation scenario is kind of simple. The future work would take more complex scenarios into consideration, for example including more users and more base stations.

Acknowledgement

This work is supported by Science and Technology Project from of State Grid Corporation of China: “Key technology Research and Application for Ubiquitous Inter-networking Security for Power wireless private network services (No. SGJSXT00GCJS1900120)”

Reference

[1] Y. Zhang, K. Liang, S. Zhang and Y. He, "Applications of edge computing in PIoT," 2017 IEEE Conference on Energy Internet and Energy System Integration (EI2), Beijing, 2017, pp. 1-4.

[2] B. Dab, N. Aitsaadi and R. Langar, "Q-Learning Algorithm for Joint Computation Offloading and Resource Allocation in Edge Cloud," 2019 IFIP/IEEE Symposium on Integrated Network and Service Management (IM), Arlington, VA, USA, 2019, pp. 45-52.
[3] T. Yang, Y. Hu, M. C. Gursoy, A. Schmeink and R. Mathar, "Deep Reinforcement Learning based Resource Allocation in Low Latency Edge Computing Networks," 2018 15th International Symposium on Wireless Communication Systems (ISWCS), Lisbon, 2018, pp. 1-5.

[4] T. Yang, Y. Hu, M. C. Gursoy, A. Schmeink and R. Mathar, "Deep Reinforcement Learning based Resource Allocation in Low Latency Edge Computing Networks," 2018 15th International Symposium on Wireless Communication Systems (ISWCS), Lisbon, 2018, pp. 1-5.

[5] I. John, A. Sreekantan and S. Bhatnagar, "Efficient Adaptive Resource Provisioning for Cloud Applications using Reinforcement Learning," 2019 IEEE 4th International Workshops on Foundations and Applications of Self* Systems (FAS*W), Umea, Sweden, 2019, pp. 271-272.

[6] W. Liao, M. Hong, H. Farmanbar, X. Li, Z. Luo and H. Zhang, "Min Flow Rate Maximization for Software Defined Radio Access Networks," in IEEE Journal on Selected Areas in Communications, vol. 32, no. 6, pp. 1282-1294, June 2014.

[7] H. Zhu, S. Wang and D. Chen, "Energy-efficient user association for heterogenous cloud cellular networks," 2012 IEEE Globecom Workshops, Anaheim, CA, 2012, pp. 273-278.