MAML-Based D2D Power Control Scheme in User-Variable Scenario

Zhenyu Fan[1(✉)] and Xinyu Gu[1,2]

[1] School of Artificial Intelligence, Beijing University of Posts and Telecommunications, Beijing 100876, China
{salam,guxinyu}@bupt.edu.cn
[2] Purple Mountain Laboratories, Nanjing 211111, China

Abstract. Meta-Learning has been extensively studied since it has the ability of quickly learning new skills by leveraging prior few-shot tasks, which is capable of relieving the problem of relying on large amount of data sample existed in deep learning. In this paper, we apply an algorithm of model-agnostic meta-learning (MAML) to cope with Device-to-Device (D2D) transmit power control issue in user-variable scenario. Specifically, MAML first learns good weight initializations of D2D power control neural network in initial D2D scenario, contributing to a meta-learner. When the number of D2D user changes, the network loads the meta learner and quickly adapts to a new scenario on a few shots of samples. Simulation results demonstrate that MAML shows good performance in generalization and MAML better conducts D2D user-variety power control issues than regular deep neural network power control methods.

Keywords: Model-agnostic meta-learning · deep neural network · user-variable · weight initialization · optimization

1 Introduction

Device-to-Device (D2D) communication has been considered as a promising technology, which enables nearby D2D user equipment (DUE) to communicate directly without the transition of the base station (BS) and undertakes an ability of communicating in the same channel to increase spectrum utilization efficiency. Nevertheless, channel reuse may prompt severe mutual interference among DUEs and it impairs D2D system throughput. Transmit power control becomes a critical technique to alleviate DUE mutual interference. For the purpose of adjusting transmit power of different DUEs in the same channel to maximize total throughput, most researches formulate it as non-convex problems, which entail large quantities of iterations and has high computational complexity. Machine learning (ML), e.g., deep learning has been exploited in D2D power control since ML based methods can approximate traditional optimization algorithms with much lower computational complexity. However, most neural networks in ML possess fixed-dimensional input and output which is only applied to D2D scenarios with specific number of D2D users [1], when facing a scenario where the number of D2D

F. Gao et al. (Eds.): ChinaCom 2022, LNICST 500, pp. 362–374, 2023.
https://doi.org/10.1007/978-3-031-34790-0_28

users is changing, the behavior of former network may be degraded, a new power control network has to be established and the training process would be conducted again.

Due to the correlation between the network structure and the number of D2D pairs in the scene, applying deep neural network (DNN) power control network to user-variable scenarios may lead to a result that the size of power control DNN for previous moment is not available for this moment, which is inappropriate for real-time operation and leads to the redundancy of rebuilding and retraining new power control network. Meta-learning is a category of machine learning that learns the representation on various learning tasks, forming prior experience and yielding a meta-learner, such that meta-learner teaches to initialize a base-learner and base-learner adapts new tasks quickly using only a small number of training samples. Moreover, model-agnostic meta-learning (MAML) proposed in [2] is a popular meta-learning approach which can be directly applied to any learning problem and model trained with a gradient descent procedure without the constraints on the number of learning parameters or model architecture. Compared with other machine learning methods, MAML can bring better generalization performance to the network and has attracted wide attention of many researchers. The author in [3] applied a modified model-agnostic algorithm capable of performing tasks just trained on a few shots of samples. In [4], the author involved L_1 regularization in standard MAML and proposed a sparse model-agnostic meta-learning (SMAML) to further enhance the efficiency in MAML. [5] applied a practical implementation of MAML to conduct image classification tasks and the result showed that MAML can transfer the prior experience extracted from pre-training on to new tasks and led to good generalization performance. Considering the transfer characteristic of MAML, in this paper, we exploit MAML method to D2D power control in user-variable scenario, i.e., first, we build a DNN power control network and acquire initial parameters in an initial scenario, second, when D2D user number changes, the network reloads initial parameters such that get trained for power control in a new scenario quickly, thereby enhancing network adaptability and real-time operation.

The rest of this paper is organized as follows. Section 2 describes the considered D2D communication scenario and formulates the D2D power control problem. In Sect. 3, we introduce MAML algorithm and propose MAML-based power control policy. Numerical results are shown and discussed in Sect. 4 and we finally conclude the paper in Sect. 5.

2 System Model and Problem Formulation

2.1 System Model

As can be seen in Fig. 1, we consider a Device-to-Device communication scenario with the size of $D \times D$, where a BS is located in the center to manage the power control information and channel state information (CSI) for D2D. Meanwhile, there are N pairs of D2D users, whose maximum distance between transmitter and receiver is 25 m [6], randomly distributed in the area and they simultaneously reuse the same frequency for transmissions. Let $\mathbb{I} = \{1, 2, 3, ..., N\}$ and $\mathbb{J} = \{1, 2, 3, ..., N\}$ respectively represents the set of D2D transmitters and receivers, accordingly, a D2D receiver $k \in \mathbb{J}$ may receive interference from other D2D transmitter $i \in \mathbb{I}\backslash\{k\}$. To address the cochannel interference, transmit power control is necessary such that cochannel interference can be relieved and D2D system throughput is maximized.

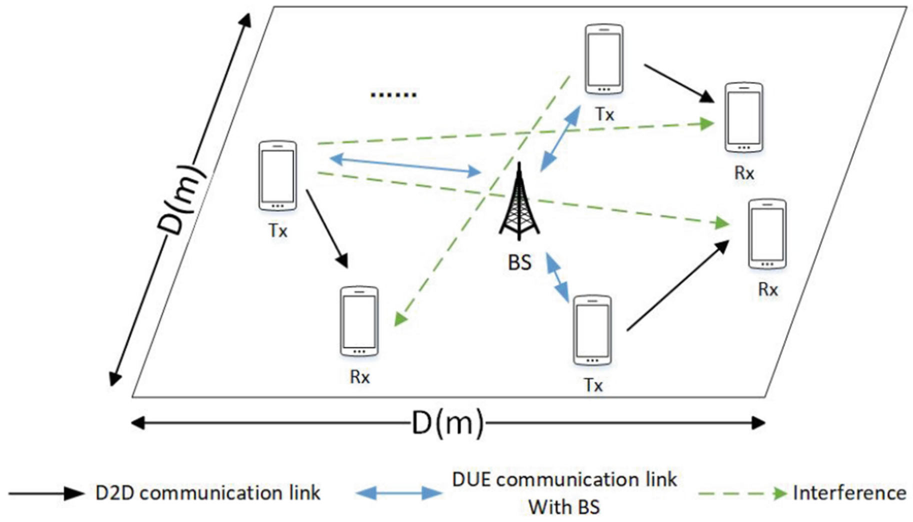


Fig. 1. D2D communication scenario network topology.

2.2 Problem Formulation

In the present work, we evaluate the D2D system weighted sum-rate (WSR) and the mean squared error (MSE) between predicted power and theoretical optimal power to characterize the merit of ML-based power control scheme, while comparing the adaptability of regular neural network and MAML to user-variable areas. Denote the distance from the i-th D2D transmitter to the j-th receiver as $d_{i,j}$ and the multipath fading between the i-th D2D transmitter and the j-th receiver as $G_{i,j}$, specifically, each multipath fading $G_{i,j}$ satisfies $G_{i,j} \sim N(0, 1)$. Thus, the channel gain from the i-th D2D transmitter to the j-th receiver can be written as $h_{i,j} = \beta(d_{i,j})^{-\alpha}|G_{i,j}|$ [7], where β and α represent the path loss coefficient and path loss exponent, and $\mathbf{H}$ is the matrix of $h_{i,j}$.

We analyze the overall system throughput and formulate the transmit power control problem in the work based on the Signal to Interference plus Noise Ratio (SINR) of D2D receivers. Let P_i denote the predicted transmit power of the i-th transmitter, where $0 \leq P_i \leq P_{max}$, and $\mathbf{P}$ is the matrix of P_i. Furthermore, we assume that N_0 is noise spectral density and W is the carrier bandwidth that D2D users access [8], and the WSR model can be expressed as follows:

$$\text{maximize} \sum\nolimits_{k=1}^{N} W log_2(1 + \frac{h_{i,j}P_i}{N_0 W + \sum_{k \in \mathbb{I} \setminus \{i\}} h_{k,i}P_k}) \tag{1}$$

The objective of power control scheme is to maximize the D2D WSR model such that an ideal power control policy for CSI in a certain scenario can be achieved.

Given that weighted minimum mean square error (WMMSE) is capable of reaching a stationary solution of non-convex problem, e.g., problem (1) [9], we regard WMMSE as the upper limit of D2D network communication performance and denote $\mathbf{P_{WMMSE}}$ as the theoretical optimal power generated by WMMSE. Since $\mathbf{P_{WMMSE}}$ generated according to certain CSI channel gain utilizing WMMSE scheme represents an optimal power allocated policy in a specific scenario, the MSE between $\mathbf{P_{WMMSE}}$ and $\mathbf{P}$ is capable of indicating the adaptability of ML power control method. The adaptability of ML power control method to user-variable areas can be written as:

$$\text{minimize}\,\mathrm{L} = \sqrt{(\mathbf{P_{WMMSE}} - \mathbf{P})^2} \tag{2}$$

The smaller the value of L, the greater approximating performance of ML power control and the better adaptability to certain D2D scenario can be achieved. Thus, the objective of power control scheme includes minimizing the MSE between theoretical optimal power and predicted power.

3 Model-Agnostic Meta-learning Based D2D Power Control Scheme

In this section, the detailed MAML algorithm is presented first. Then, we construct the neural network used for determining D2D transmit power and apply the MAML to ML-based power control procedure for D2D power control in user-variable scenarios.

3.1 Model-Agnostic Meta-learning Algorithm

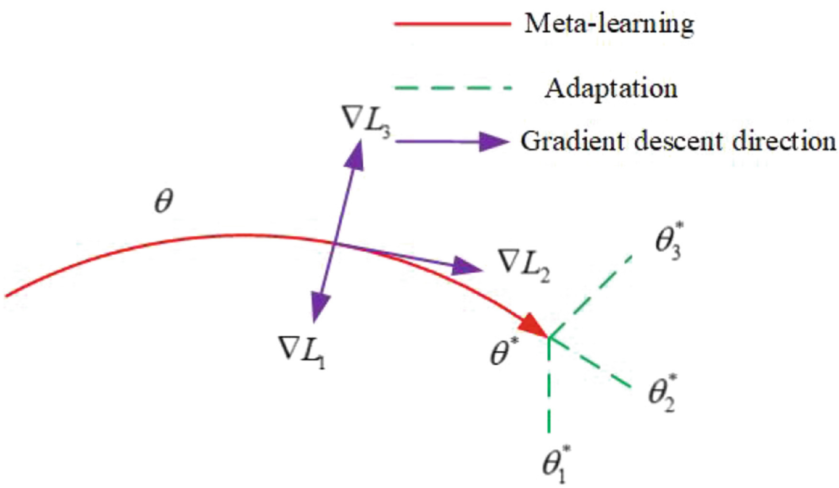


Fig. 2. Diagram of model-agnostic meta-learning algorithm (MAML).

As depicted in Fig. 2, meta-learning possesses an ability of training good model's initial parameters θ, contributing to a generic meta-learner on various similar tasks, such that the learner can quickly create adapted parameters $\theta_1^*, \theta_2^*, \theta_3^*$, thereby achieving rapid adaptation and providing an ideal feature representation towards most deep learning models trained for new tasks [2], and that is defined as model-agnostic. In K-shot MAML scenario, we consider a distribution over tasks $T \sim p(T)$ where H batches of task, $\{T_i\}_{i=1}^H$, are collected from this distribution and K samples are drawn in each task T_i. Accordingly, a generic model f_θ with parameters θ is trained on K samples in each task from $p(T)$, i.e., when training on T_i, the model will adapt to new parameters $\theta_i^{'}$, which is calculated via gradient descent update, as is given as follows:

$$\theta_i^{'} = \theta - \gamma_1 \nabla_\theta L_{T_i}(f_\theta) \tag{3}$$

where $L_{T_i}(f_\theta)$ is the loss computed using K samples for task T_i, γ_1 is the learning rate in the first step of gradient updating. To achieve good adaptability for f_θ, the model parameters are trained by minimizing total loss from tasks in $p(T)$, and the meta-objective can be expressed as follows:

$$\min_\theta \sum_{T_i \sim p(T)} L_{T_i}(f_{\theta_i^{'}}) = \sum_{T_i \sim p(T)} L_{T_i}(f_{\theta - \gamma_1 \nabla_\theta L_{T_i}(f_\theta)}) \tag{4}$$

with the total loss computed across tasks from $p(T)$ minimized, the optimal initialization parameters for meta-learner can be acquired as follows:

$$\theta \leftarrow \theta - \gamma_2 \nabla_\theta \sum_{T_i \sim p(T)} L_{T_i}(f_{\theta - \gamma_1 \nabla_\theta L_{T_i}(f_\theta)}) \tag{5}$$

where γ_2 is the learning rate in the second step of gradient update, and the K-shot MAML algorithm is outlined in Algorithm 1.

Algorithm 1 K-shot Model-Agnostic Meta-Learning

Require: $p(T)$: distribution over tasks
Require: γ_1, γ_2: learning rates of two steps
1: randomly initialize θ
2: **while** not done **do**
3: Sample H batches of tasks $\{T_i\}_{i=1}^H \sim p(T)$
4: **for all** T_i **do**
5: Evaluate $\nabla_\theta L_{T_i}(f_\theta)$ with respect to K samples
6: Compute adapted parameters with gradient descent: $\theta_i^{'} = \theta - \gamma_1 \nabla_\theta L_{T_i}(f_\theta)$
7: **end for**
8: Update $\theta \leftarrow \theta - \gamma_2 \nabla_\theta \sum_{T_i \sim p(T)} L_{T_i}(f_{\theta - \gamma_1 \nabla_\theta L_{T_i}(f_\theta)})$
9: **end while**

3.2 MAML-Based Power Control Scheme

In ML D2D transmit power allocation problem, a DNN is constructed to optimize problems in (1) and (2), thereby determining an optimal transmit power allocation policy of all D2D transmitters according to channel state information $\mathbf{H}$, whose size is $N \times N$ and is related to the number of D2D transceiver pairs N in the scenario. As shown in Fig. 3, a four-layer fully connected neural network is employed for feature extraction, and the channel state information $\mathbf{H}$ is flattened into one-dimensional vector, whose length is N^2, contributing to the input of DNN. In each hidden layer, 40 hidden neurons are included and the Rectified Linear Unit (ReLU) is used as the activation function to relieve gradient vanishing and overfitting.

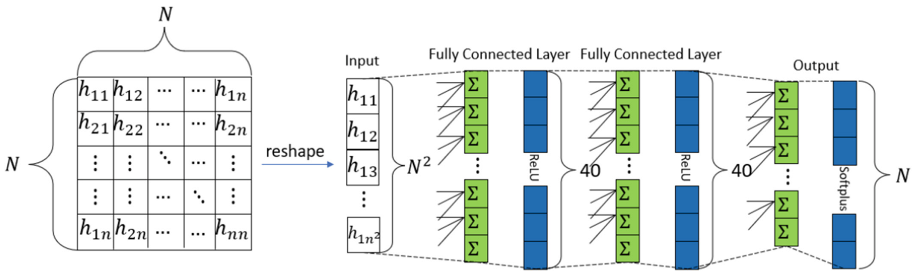


Fig. 3. Deep neural network architecture for power control.

The output of the network keeps the length of N and is fed into the Softplus part, i.e., when the input of Softplus part is x_S, the i-th output of Softplus part comes to $\log(e^{[x_S]_i} + 1)$, which restricts the output range and prevents negative values such that it can be treated as the predicted power allocation result.

In terms of DNN power control in D2D user-variable scenarios, the number of D2D transceiver pairs N is always changing, since the length of input layer and output layer are related to N, a phenomenon may happen that a well-trained DNN used for power control for previous second may not be applicable for this second and we have to establish a new DNN, initializing the whole model parameters again, which leads to unnecessary repetition of steps. Meta-learning has an ability of leveraging previous experience to train a model that can quickly adapt to new tasks using only a few samples and training iterations, furthermore, MAML is a well-studied meta-learning algorithm that has shown impressive results over many problems, e.g., supervised regression, classification, and reinforcement learning, and its feature has potential to handle the phenomenon. Therefore, in this paper, we propose a MAML-based power control scheme to address the changing of D2D transceiver pairs in the scenario.

As can be seen in Fig. 4, the scheme is divided into two phases, meta-training phase and meta-test phase [10], specifically, data samples used for both meta-training phase and meta-test phase are generated through system-level simulation of respective D2D scenarios. In meta-training phase, the number of D2D transceiver pairs in scenario is N_0, where a DNN is initialized and a K-shot MAML is conducted. The data sample used for meta-training phase includes M_0 batches and the value of M_0 includes 10000 and 20000.

Each batch contains K sample shots, consequently, the size of input feature fed into DNN becomes $[M_0, K, N_0, N_0]$ and the size of output label becomes $[M_0, K, N_0]$. After meta-training phase, a well-trained DNN emerges and we save the model parameters, which is defined as meta-learner. In meta-test phase, we aim to apply the meta-learner to the power control network in scenarios with different number of D2D pairs. Considering the number of D2D pairs is variable in meta-test phase and the corresponding power control network has different size of input and output layer, we denote the number of D2D transceiver pairs in user-variable scenario as the variable parameter N and reinitialize the input and output weight parameters of power control network with N dimensions while in hidden layers the network directly loads weight parameters of the meta-learner, and then the network transferred with meta-learner contributes to a base-learner. After loading meta-learner, the base-learner will be finetuned under only a few data samples with M batches to adapt to a new specific D2D scenario, accordingly, the size of input feature becomes $[M, K, N, N]$ and that of output label becomes $[M, K, N]$. Notably, during finetuning, the loss function is considered the same as optimization problem (2), and the optimizer comes to Adam with the learning rate of 0.001.

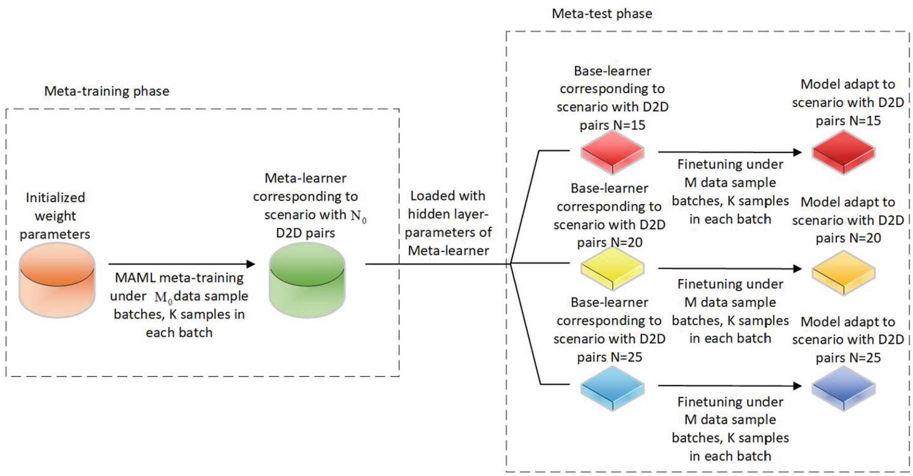


Fig. 4. Schematic of MAML-based power control scheme.

Through MAML, the base-learner can adapt to user-variable scenarios quickly and behave better generalization performance under only a few training samples without rebuilding a new power control network or conducting large number of training iterations.

4 Simulation Results and Analysis

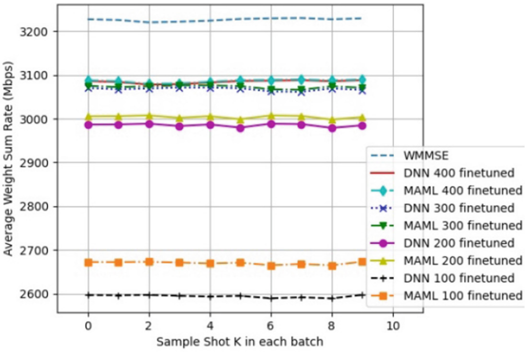
In this section, we conduct extensive numerical experiments to verify the functionality and performance of the proposed MAML-based power control scheme and compare the performance of the proposed scheme with that of the following schemes: 1) WMMSE-based power control policy [9]; 2) regular DNN power control policy mentioned in [6].

Specially, for comparing the adaptability of DNN policy and MAML-based policy, we train the raw DNN network with M batches of data samples, the same as the finetuning process of base-learner in meta-test phase, and we subsequently compare the trained DNN with the finetuned base-learner in MAML-based power control. Moreover, the structure of raw DNN is the same as that of the base-learner such that other irrelevant variables can be controlled. The simulation program is operated in Python 3.6.2 with Tensorflow 2.6.2 on the computer equipped with 16GB of RAM and one 8-core Intel CPU, besides, Keras library is used for the construction and training of a neural network. The specific parameter settings are given in Table 1.

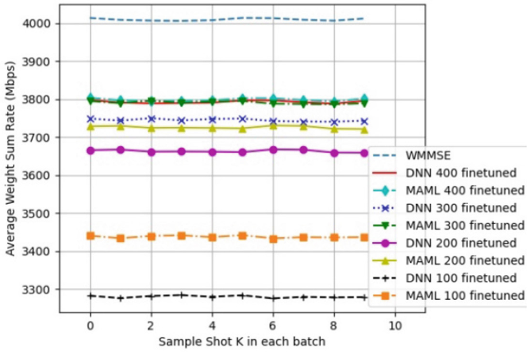
Table 1. System Parameters.

Parameter	Value
Number of D2D pairs in meta-training phase N_0	10
Number of D2D pairs in meta-training phase N	{15, 20, 25}
Number of samples in each batch K	10
Number of batches in meta-training phase M_0	{10000, 20000}
Number of batches in meta-test phase M	{100, 200, 300, 400}
D2D Max/Min transmit power P_{max}/P_{min}	24/0 dBm
Size of D2D area D	700 m
Channel Bandwidth W	20 MHz
Noise spectral density N_0	−174 dBm/Hz
Learning rate in the first step γ_1	0.01
Learning rate in the second step γ_2	0.001
Path loss coefficient β	$10^{-3.453}$
Path loss exponent α	3.8

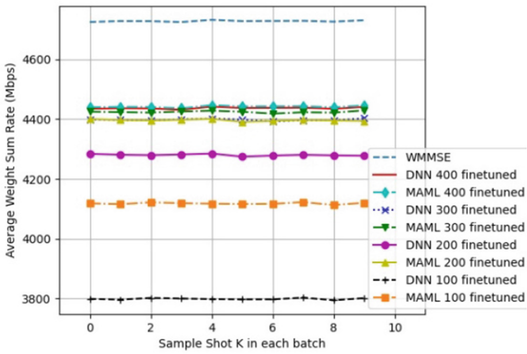
Firstly, the simulation evaluates the WSR performance of D2D scenarios with different number of D2D pairs when exploiting WMMSE, MAML-based and regular DNN power control policy. In Fig. 5(a), 5(b) and 5(c), we show the average WSR result finetuned by different number of training batches in scenarios with different number of D2D pairs. It can be observed that for arbitrary sample shot, no matter how many batches conducted in the finetuning process, the WSR of MAML-based power control policy always performs better than regular DNN power control policy, and both of them performs slightly inferior to WMMSE-based scheme. Besides, as the number of batches M increases in finetuning process, both the WSR of MAML-based power control and that of DNN power control achieve closer to optimal WSR value while the gap between them decreases or even decreases to 0 under certain conditions. This is because with the increase of batches, both raw DNN and base-learner have been finetuned more thoroughly such that both of them have reached saturation and their performance gap may not be obvious, meanwhile, we may infer from the result that the superiority of MAML-based policy is more obviously highlighted with the descent of training batches or with the growth of D2D pairs.



(a) WSR performance in scenario where N=15.



(b) WSR performance in scenario where N=20.

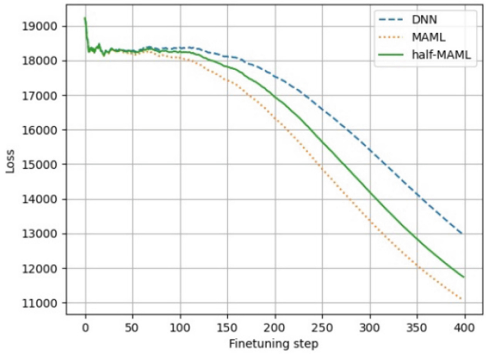


(c) WSR performance in scenario where N=25.

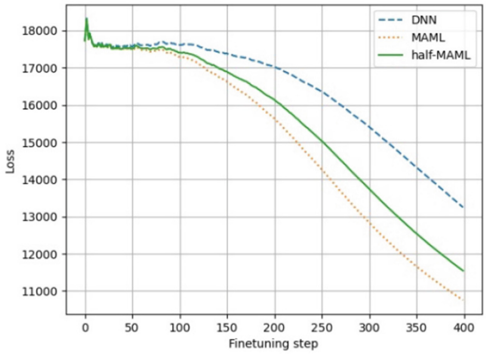
Fig. 5. The WSR performance of D2D scenarios with different number of D2D pairs.

Secondly, the simulation evaluates the MAML model transfer ability and adaptability towards D2D user-variable scenarios under 400 sample batches of finetuning process, mainly by comparing the MSE value respectively derived from raw DNN, base-learner loaded with half-trained meta-learner and base-learner loaded with totally trained meta-learner. Particularly, half-trained meta-learner is trained under 10000 batches in meta-training phase where M_0 is 10000, while meta-learner is trained under 20000 batches in meta-training phase where M_0 is 20000. As can be seen from results in Fig. 6(a), 6(b) and 6(c), regardless of the number of D2D pairs in the scenario, base-learner always experiences lower MSE than DNN with direct weight initializations when trained with any number of batches, moreover, the base-learner equipped with totally trained meta-learner experiences lower MSE value than base-learner equipped with half-trained meta-learner. Simulation results demonstrate that when the number of D2D pairs changes, the base-learner can better approximate the WMMSE policy than raw network with direct weight initializations. Besides, the more batches applied to the training of meta-learner, the more prior knowledge base-learner would accumulate, thereby accomplish better generalization performance. MAML-based power control behaves better adaptability in D2D user-variable scenario.

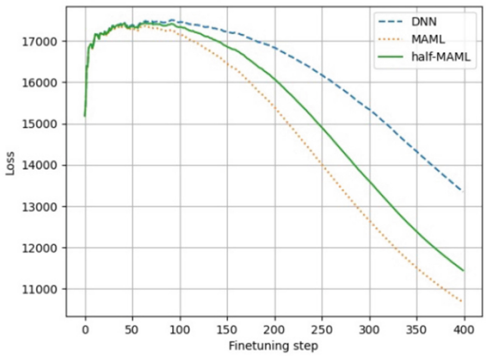
Thirdly, to analyze the real-time operation of MAML-based power control scheme, the simulation evaluates the MAML model time complexity towards D2D user-variable scenario under 400 sample batches of finetuning process, mainly by comparing the probability histogram of time consumption derived from raw DNN, base-learner loaded with half-trained meta-learner and base-learner loaded with totally trained meta-learner. We can infer from results in Fig. 7(a), 7(b) and 7(c) that regardless of the number of D2D pairs in scenario, the time consumption distribution of MAML-based power control is more concentrated than that of raw DNN, i.e., MAML-based power control time consumption distribution is mainly concentrated in the low time consumption interval while the time consumption of raw DNN still exist individual distributions in the range of high time consumption interval. Simulation results indicate that with respect of D2D user-variable scenario, power control network loaded with prior knowledge could adapt to changes of the number of D2D pairs quickly, MAML-based power control achieves better real-time operation.



(a) MSE value in scenario where N=15.

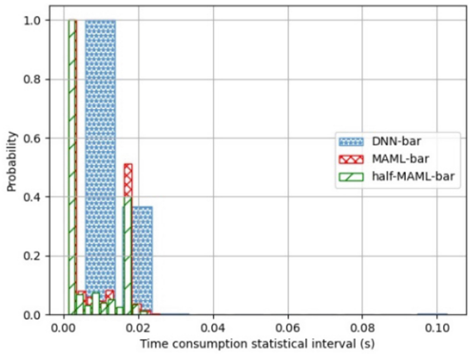


(b) MSE value in scenario where N=20.

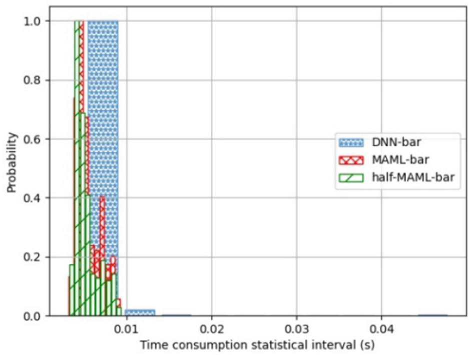


(c) MSE value in scenario where N=25.

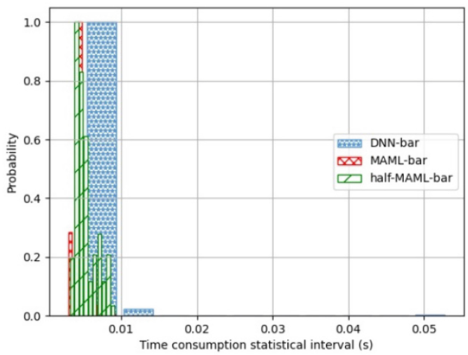
Fig. 6. The MSE value between optimal power and predicted power in D2D scenarios with different number of D2D pairs.



(a) CDF curve in scenario where N=15.



(b) CDF curve in scenario where N=20.



(c) CDF curve in scenario where N=25.

Fig. 7. Time consumption in D2D scenarios with different number of D2D pairs.

5 Conclusion

The Machine learning based power control issue considering dynamic user changing scenarios is more complicated than regular neural network-based power control due to the correlation between the network structure and the number of D2D pairs, and directly applying raw DNN to power control in D2D user-variable scenarios may leads to repetition of rebuilding and retraining new power control network, which increase time complexity. In this work, we analyze the characteristic of meta-learning and apply MAML to pre-train a meta-learner for DNN power control methods. Simulation results show that in D2D user-variable areas, the MAML-based power control method achieves better weighted sum rate and lower mean squared error than DNN power control without meta-learners under the same training conditions. Furthermore, the power control network loaded with meta-learner parameters can converge to optimal results faster and regress to similar performance with less training samples compared with regular power control network. Intuitively, a better adaptability can be accomplished for a user-variable power control network in view of MAML algorithm.

Acknowledgement. This work was supported in part by the State Major Science and Technology Special Projects (Grant No.2018ZX03001024) and in part by the National Key Research and Development Program (Grant No. 2022YFF0610303).

References

1. Pei, E., Yang, G.: A deep learning based resource allocation algorithm for variable dimensions in d2d-enabled cellular networks. In: 2020 IEEE/CIC International Conference on Communications in China (ICCC), pp. 277–281. IEEE (2020)
2. Finn, C., Abbeel, P., Levine, S.: Model-agnostic meta-learning for fast adaptation of deep networks. In: Proceedings of the 34th International Conference on Machine Learning, in PMLR, vol. 70, pp. 1126–1135 (2017)
3. Pawar, A.: Modified model-agnostic meta-learning. In: 2020 IEEE International Conference on Machine Learning and Applied Network Technologies (ICMLANT). IEEE (2020)
4. Gai, S., Wang, D.: Sparse model-agnostic meta-learning algorithm for few-shot learning. In: 2019 2nd China Symposium on Cognitive Computing and Hybrid Intelligence (CCHI), pp. 127–130. IEEE (2019)
5. So, C.: Exploring meta learning: parameterizing the learning-to-learn process for image classification. In: 2021 International Conference on Artificial Intelligence in Information and Communication (ICAIIC), pp. 199–202 (2021)
6. Lee, W., Kim, M., Cho, D.H.: Transmit power control using deep neural network for underlay device-to-device communication. IEEE Wirel. Commun. Lett. **8**(1), 141–144 (2019)
7. IEEE 802.20 WG: IEEE 802.20 channel models document. IEEE 802.20 PD-08rl (2007)
8. 3GPP TR 38.889: 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Study on NR-based access to unlicensed spectrum. Rel-16 V16.0.0, pp. 64–65 (2018)
9. Shi, Q., Razaviyayn, M., Luo, Z.Q., He, C.: An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel. IEEE Trans. Signal Process. **59**(9), 4331–4340 (2011)
10. Sun, Q., Liu, Y., Chua, T.S., et al.: Meta-transfer learning for few-shot learning. In: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE (2019)