IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 18 June 2024, accepted 2 July 2024, date of publication 5 July 2024, date of current version 15 July 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3423840

RESEARCH ARTICLE

Machine Learning Assisted Cross-Layer Joint Optimal Subcarrier and Power Allocation for Device-to-Device Video Transmission

SHU-MING TSENG[1], (Member, IEEE), JUN-JIE WU[2], AND CHAO FANG[3], (Senior Member, IEEE)

[1]Department of Electronic Engineering, National Taipei University of Technology, Taipei City 106, Taiwan
[2]Lenovo Global Technology (Taiwan) Ltd., Taipei City 115, Taiwan
[3]Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China

Corresponding author: Shu-Ming Tseng (shuming@ntut.edu.tw)

This work was supported in part by the National Science and Technology Council, Taiwan, under Grant NSTC 112-2221-E-027-079-MY2; and in part by the National Taipei University of Technology-Beijing University of Technology (NTUT-BJUT) Joint Program under Grant NTUT-BJUT-112-02.

ABSTRACT The previous scheme used imitation learning by classification of branching or pruning, combined with Branch and Bound (B&B) algorithm to solve physical-layer-only joint optimal subcarrier and power allocation problem in the device-to-device communications. In this paper, we propose joint source encoding rate control and machine learning assisted cross-layer joint optimal subcarrier/ power allocation. The proposed scheme has source encoder rate control and can adaptively adjust video rate to increase the video quality, peak signal to noise ratio (PSNR). The previous physical-layer-only scheme did not use the content-based video rate adaption. Furthermore, the proposed scheme uses the objective function of PSNR directly and allocates the subcarrier and power considering the different rate-distortion function of users' videos. The previous physical-layer-only scheme could only treat the users' video equally. Under the new minimum PSNR constraint of the cellular user (CU), we derive a new objective function that is independent of the transmission power of the CU to simplify the optimization problem formulation. The previous scheme considered the physical-layer-only objective function and constraints. Finally, in addition to imitation learning, the proposed scheme adopts ensemble learning with downsampling the majority set {prune} to alleviate the class imbalance problem and improves performance. The simulation results show that in the scenario where the number of CUs is 5, the number of subcarriers is equal to the number of CUs, the bandwidth is 15k Hz, and the number of D2D pairs is 2, the PSNR of the previous physical-layer-only scheme is 31.03 dB, while our proposed cross-layer allocation scheme is 35.67 dB, a 4.64 dB gain. The trained model trained at 5 CUs can generalize without re-training to 10 CUs with only 5.91% gap to the optimal PSNR and 20.43 times speed (95% execution time reduction) when compared to the globally joint optimal subcarrier/power allocation B&B algorithm.

INDEX TERMS Video communications, machine learning, resource management, Internet of Things, data imbalance, cross layer design.

I. INTRODUCTION

With the advancement of internet technology and the popularization of communication devices, a large number of devices are interconnected. In order to meet the needs of a large number of growing mobile users and improve the efficiency of spectrum use, device-to-device (D2D) communication in the future Internet of Things (IoT) network will be one of the indispensable and important technologies [1]. D2D technology communicates directly through neighboring user equipment without relying on the base station (BS), and improves the network spectrum efficiency by reusing

The associate editor coordinating the review of this manuscript and approving it for publication was Renato Ferrero.

the spectrum resources of the cellular user (CU) [2]. Thus, D2D communication has been incorporated into the Long Term Evolution- Advanced (LTE-Advanced). In recent years, D2D communication based on cellular networks has aroused extensive discussions in academia due to its potential to improve spectral efficiency and cell throughput. However, D2D communication may cause serious interference to existing CUs due to spectrum reuse. Therefore, the performance of the system is improved by power control and subcarrier allocation, thereby reducing the channel interference between the D2D pair and the CU [3], [4], [5].

In wireless communication, resource allocation such as channel and power is critical for performance improvement. The resource allocation problem is viewed as a mixed integer nonlinear programming (MINLP) problem. Most of these MINLP problems are NP-hard problems. Therefore, some studies had turned to suboptimal scheme to reduce complexity, such as [6] based on game theory to solve power allocation in cellular network-assisted D2D communication systems, [7] solved the channel and power assignment problem. in D2D -based IoT networks. Reference [8] proposed separate channel allocation, power allocation, and matching in D2D-based IoT systems. Reference [9] proposed a resource allocation scheme which is robust with channel uncertainties in underlaid D2D communication network. However, the above papers [6], [7], [8], [9] dealt with the physical layer only and did not consider joint source rate encoder rate and resource management optimization in D2D video communications.

Cross layer resource allocation in cellular [10], [11], [12] and D2D [5], [13], [14] video communications considers more than one layer in TCP/IP layers and gets better performance and user experience in video communications, the majority of today's data traffic [15]. Reference [10] considered adaptive modulation and coding and cross PHY/APP resource allocation with fixed source encoding rate in cellular systems. The fixed data size cannot adapt to time varying channels. Reference [11] proposed a content-based adaptive source encoding rate and cross PHY/APP layer resource allocation scheme in cellular systems. Reference [12] updated rate distortion parameters in higher rate and compensated nonperfect source encoder rate control in cellular systems. Reference [5] proposed a separately suboptimal approach-subcarrier allocation in outer loop and power allocation in the inner loop with objective of minimizing sum video MSE distortion in D2D video communications. It considered the channel status information (CSI) in the physical (PHY) layer and the video rate distortion function in the application (APP) layer. Reference [13] proposed resource block (RB) allocation and D2D pairing in D2D video communications. Reference [14] proposed power allocation with the objective to maximized the data rate weighted by video popularity in video caching of D2D video delivery. However, the adaptive video source encoder rate and joint optimal subcarrier (or RB) and power allocation for D2D video communications has not been addressed in previous D2D cross layer resource allocation studies [5], [13], [14].

In recent years, machine learning has become a breakthrough technology and has been widely used in resource allocation problems in wireless communications, because the resource allocation scheme based on machine learning can effectively reduce the computational complexity and program execution time [16], [17], [18], [19], [20], [21], [22], thus causing huge research attention. In [16], the SCMA codebook assignment problem in wireless communication was solved through deep neural networks. In [17], a deep reinforcement learning based approach was used to solve the sum rate maximization problem of D2D communication. The machine learning methods mentioned above follow an end-to-end model, and learn the relationship between input and output through various machine learning methods. This method was more suitable for a single kind of optimization problem such as power allocation without channel allocation in [18] because the end-to-end learning model was difficult to effectively solve the complex MINLP problem [19]. In [18], a DNN approximate weighted minimum mean squared error (WMMSE) algorithm was proposed to lower the complexity. Reference [20] proposed the power allocation based on homogeneous graph neural network (GNN) for D2D communications. Reference [21] proposed the power allocation based on heterogeneous GNN) for D2D communications where the number of antennas is different in D2D device groups. however, there is significant performance gap compared to the optimal solution because the training samples are obtained through a sub-optimal algorithm, plus the error of prediction using deep learning [22]. Furthermore, the disadvantage of using the deep learning method is the requirement of big training data. As a result, the cost of collecting training data will be very expensive, which is not feasible in reality. However, more complex MINLP problems are difficult to solve through this end-to-end approach, so it is better to combine mathematical optimization techniques and with machine learning to solve the MINLP problem [19], [22].

Branch and Bound (B&B) algorithm is a mathematical optimization technique which can obtain the optimal solution to the MINLP problem. Its time complexity is high. A number of literatures [19], [22], [23] show that imitation learning can accelerate the branching process of B&B algorithm to reduce the computational complexity. In [23], the concept of applying imitation learning to solve the MINLP problem was first proposed. Reference [22] used the B&B algorithm and deep neural network (DNN) classifier by imitation learning in wireless communication to solve the power allocation problem of Cloud-RANs. Reference [19] used B&B algorithm and support vector (SVM) classifier by imitation learning to solve the problem of jointly subcarrier and power allocation in D2D communications. However, previous machine-learning based D2D resource allocation studies [18], [20], [21] considered only power allocation and not channel allocation; previous jointly optimal power and subcarrier allocation D2D resource allocation study [19] considered the physical layer only but did not take source encoding rate control and application layer metric into consideration.

A. CONTRIBUTION

In this paper, we propose joint source encoding rate and imitation/ensemble learning based cross-layer resource allocation optimization for uplink D2D video transmission systems. The proposed imitation learning of B&B (jointly optimal channel/power allocation) in a cross layer fashion allows the proposed scheme to i) outperforms classical suboptimal and separate channel/power allocation schemes such as [13] ii) scale/generalize well without re-training while typical deep learning based approaches do not scale/generalize to larger number of users Combined with the proposed downsampling training and ensemble learning (Sec. IV-E), the proposed imitation learning of B&B scheme solves the class unbalance problem and is expected to improve the performance further [25].

When compared with aforementioned previous studies [5], [6], [7], [8], [9], [13], [14], [18], [19], [20], [21], this paper has following contributions:

1) We propose joint video source encoder rate adaption and cross layer joint channel/power allocation scheme in D2D video communications by B&B algorithm and imitation learning. For comparison, [6], [7], [8], [9] considered the physical layer only resource allocation in D2D communications and did not take source encoding rate control and application layer metric into consideration. References [5], [13], and [14] considered cross layer resource allocation in D2D video communications, but did not consider both channel and power allocation or allocated the channel and power allocation in a separately suboptimal way. References [18], [20], and [21] considered deep-learning based power allocation and not channel allocation in D2D networks. Reference [19] considered the physical layer only machine learning-based joint channel and power allocation in D2D communications, but did not take source encoding rate control and application layer metric into consideration.
2) We take source encoding rate control and application layer metric into consideration. The video rate control is employed, according to radio resource constraints. The objective function is changed to maximize the minimum PSNR (peak signal-to-noise ratio), the video quality, instead of data rate in the physical layer, of the D2D pair. The minimum CU performance constraint is changed to minimum PSNR instead of data rate, of CUs. When training the SVM classifier, we change the original training feature data rate in [19] to the video mean-square-error (MSE) distortion. We improve the performance of the classifier by changing the problem-depend features. Based on the minimum PSNR constraint, we derive a novel expression of CU power allocation. We then simplify the resource allocation problem by removing the CU power allocation variable, thereby reducing the computational complexity.
3) We propose ensemble learning similar to [24] to improve performance. The difference is that the training datasets are obtained to downsample the majority class {prune} such that {prune} class has the same size as the minority class {branch} in a way similar to [25]. Thus the class-imbalance problem is compensated and the precision is expected to be improved [25]. For comparison, physical layer baseline [19] does not deal with class imbalance in the dataset and thus the precision still has room to improve.
4) The proposed scheme with the above contributions 1)-3) outperforms the state-of-the-arts (SOTAs) physical layer baseline scheme [19], and cross layer baseline scheme [13] by 4.64 dB, and 3.66 dB, respectively, in PSNR in the simulation results. This is because the proposed scheme is imitation learning based approach on the globally joint optimal B&B algorithm; the cross layer baseline scheme [13] is suboptimal and does not jointly allocation channel and power, and physical layer baseline scheme [19] bases on the globally joint optimal B&B algorithm but physical layer only and does not have ensemble learning to deal with class imbalance and improve accuracy.
5) The trained model trained at 5 CUs can generalize without re-training to 10 CUs with only 5.91% gap to the optimal PSNR and 20.43 times speed (95% execution time reduction) when compared to the conventional B&B algorithm. Therefore, the proposed scheme is near globally joint optimal B&B with only 5% of its time complexity.

II. SYSTEM MODEL

In this paper, we propose joint video source encoder rate control and imitation learning based cross-layer resource allocation for uplink D2D video transmission systems. In subsection A, we give the system overview. In subsection B, we describe the PHY model of uplink D2D communications system. In subsection C, we describe the APP layer model of video rate distortion model and video quality PSNR.

A. VIDEO TRANSMISSION SYSTEM OVERALL ARCHITECTURE

Fig. 1 is the uplink D2D video transmission system, the system architecture is based on [5], [11], and [12]. The resource allocation block not only assign the channel and power, but also control the source encoder rate, as indicated by the arrow to H.264 encoder block. H.264 rate control is applied to each group of pictures (GOP) at discrete rates. The difference is that we add D2D communication to the cellular wireless OFDMA video transmission system, allowing the D2D pair to reuse the CU's uplink channel to transmit data.

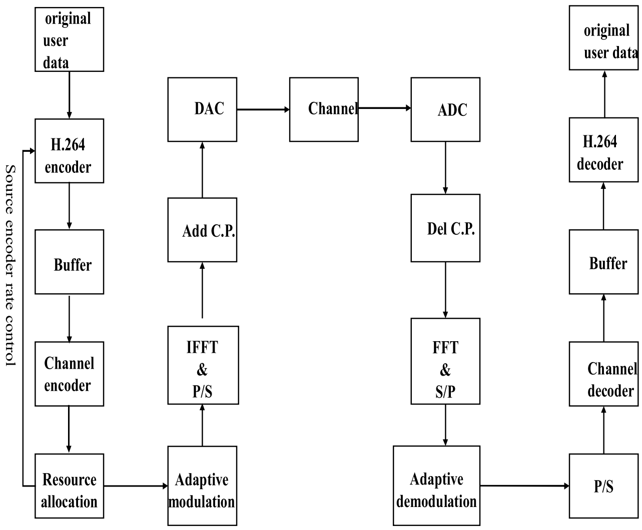

FIGURE 1. The overall architecture of the proposed video source encoder rate control and imitation learning based cross-layer resource allocation for uplink D2D video transmission systems.

B. UNDERLAY D2D COMMUNICATION SYSTEM MODEL IN PHY LAYER

As shown in Figure 2, we consider an uplink single-cell LTE-Advanced (LTE-A) cellular network-assisted D2D communication system, which consists of a base station (BS), K cellular users (CU), and M D2D pairs. Each CU uses an OFDMA subcarrier channel, so the number of subcarriers and CUs is equal. A D2D pair can occupy the CU's uplink channel. CU index k = {1,2,...,K} is equal to the subcarrier index without loss of generality. D2D pair index m = {1,2,...,M}. Subcarrier allocation vector $\alpha_{K\times M} = [\alpha_{km}]$ denotes the subcarrier allocation results. If D2D pair m reuse the channel of CU k, then $\alpha_{km} = 1$, otherwise $\alpha_{km} = 0$.

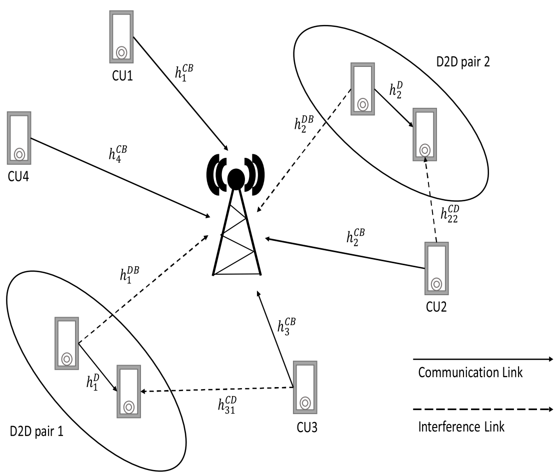

FIGURE 2. D2D communication system scenario underlay LTE-A network.

In the uplink D2D communication system model of Fig. 2, h_k^{CB} is the channel gain from CU k to BS, h_m^D is the channel gain of D2D pair m, and h_{km}^{CD} is the interference channel gain from CU k to the receiver of D2D pair m, h_m^{DB} is the interference channel gain from the transmitter of D2D pair m to the base station. p_k^C is the transmit power of CU k, and p_{km}^D is the transmit power of the D2D pair reusing the CU k channel.

Because each CU channel can be reused by no more than one D2D pair, the signal-to-interference-plus-noise ratio (SINR) of the D2D pair m reused CU k channel is

$$SINR_{km}^D(p^C, p^D, \alpha) = \frac{\alpha_{km} h_m^D p_{km}^D}{h_{km}^{CD} p_k^C + \sigma^2} \tag{1}$$

where σ^2 is the power of additive white Gaussian noise (AWGN).

The SINR of CU k is

$$SINR_k^C(p^C, p^D, \alpha) = \frac{h_k^{CB} p_k^C}{\Sigma_m \alpha_{km} h_m^{DB} p_{km}^D + \sigma^2} \tag{2}$$

According to [11], BW is the bandwidth of each subcarrier, $\eta = 3\left[Q^{-1}\left(\frac{SER_t}{4}\right)\right]^{-2}$ is the difference between theoretical and actual channel capacity.

The data rate of CU k is

$$R_k^C(p^C, p^D, \alpha) = BW log_2(1 + \eta SINR_k^C) \tag{3}$$

The data rate of D2D pair m can be expressed as

$$\begin{aligned} R_m^D(p^C, p^D, \alpha) &= \sum\nolimits_k \alpha_{km} R_{km}^D(p^C, p^D, \alpha) \\ &= \sum_k \alpha_{km} BW log_2(1+\eta SINR_{km}^D) \end{aligned} \tag{4}$$

where R_{km}^D is the data rate for D2D pair m occupying CU k's channel.

C. VIDEO RATE-DISTORTION MODEL FOR RESOURCE ALLOCATION

According to the video rate distortion model of [5], [11], [12], [13], [16], [29], [30], and [31], the video MSE distortion of the m-th D2D pair can be expressed as

$$MSE_m^D(p^C, p^D, \alpha) = a_m + \frac{b_m}{R_m^D + c_m} \tag{5}$$

where a_m, b_m, c_m are constants that depend on the video content. As indicated in subsection A, the source encoder H.264 rate control is applied to the frames inside one group of pictures (GOP) at discrete rates. The operational points are nonlinearly fit the video MSE distortion in (5) to obtain a_m, b_m, c_m.

The PSNR of the m-th D2D pair can be expressed as

$$PSNR_m^D(p^C, p^D, \alpha) = 10 log_{10} \frac{255 * 255}{MSE_m^D} \tag{6}$$

This PSNR in (6) using video rate distortion model in (5), also called error free PSNR [11], is widely used for cross layer resource allocation [5], [11], [12], [13], [16], [29], [30], [31]. There are no comparing two video frames in a pixel-by-pixel way, and no associated processing delay/latency. This PSNR in (6) using video rate distortion model in (5), also called error free PSNR [11] does not include channel error, RD curve fitting error, and imperfect encoder rate control, but the differences from the PSNR comparing two video frames, also called PSNR decoder [11], is small [5], [11], [12], [13], [16].

Because there is no video frame comparison to obtain PSNR and no associated latency, obtaining PSNR is not a bottleneck to conduct resource allocation [5], [11], [12], [13], [16], [29], [30], [31].

III. BASELINE SCHEME—PHYSICAL LAYER RESOURCE ALLOCATION [19]

The baseline scheme is slightly modified version of [19]. The difference is 1) we add a video source encoder to [19] so video communications is considered in the baseline scheme. 2) [19] considered data rate per unit bandwidth but the data rate here includes subcarrier bandwidth, and the gap between theoretical and actual channel capacity η is considered here but [19] did not.

A. VIDEO SOURCE ENCODER

The video-content-based video rate adaption is not applied because the baseline scheme is a physical layer only one.

B. PHYSICAL LAYER PROBLEM FORMULATION

The goal is maximizing the minimum data rate of D2D pair m. The physical layer resource allocation (channel, CU power, D2D pair power) problem of D2D communications is formulated as follows:

$$\max_{\{p^C,p^D,\alpha\}} \max_{m\in M} R_m^D(p^C,p^D,\alpha) \tag{7}$$

$$\text{subject to} \quad \alpha_{km} \in \{0,1\},\ \forall k \in K,\ \forall m \in M \tag{7a}$$

$$\sum_{m\in M} \alpha_{km} \leq 1,\ \forall k \in K \tag{7b}$$

$$\sum_{k\in K} \alpha_{km} p_{km}^D \leq P_{max}^D,\ \forall m \in M \tag{7c}$$

$$R_k^C(p^C,p^D,\alpha) \geq R_{min}^C,\ \forall k \in K \tag{7d}$$

$$p_k^C \leq P_{max}^C,\ \forall k \in K \tag{7e}$$

where, R_{min}^C is the minimum data rate constraint for the CU, P_{max}^D is the maximum transmit power limit of the D2D pair, P_{max}^C is the maximum allowable transmit power of the CU.

C. BRANCH AND BOUND (B&B) ALGORITHM

n the problem formulation in the previous subsection, the integer variables are binary, so B&B could obtain the global optimal solution by iteratively searching the tree [22]. The problem is solved by branching the binary tree, and each node that branches off is a sub-problem. In the calculation, the uncertain branch variables are relaxed into continuous variables in [0,1], and local solutions are obtained by solving the corresponding sub-problems.

The search process of the B&B algorithm consists of the following steps.

- Selecting node: select a node from the tree's list of unvisited nodes.
- Computation: solve the nonlinear sub-problems corresponding to the nodes and obtain their local solutions.
- Fathom decision: use the local solution and the global solution (the current optimal solution to the objective function) to decide if we remove this node from the node list.

An illustrating example of B&B algorithm is shown in Fig. 3. We first set the global solution to $+\infty$ and obtain x = 5/2, y = 3/2 by solving the constraints. Then select the branch variable y to branch to get the sub-problem, and get the local solution $-13/3$ by solving the sub-problem. Then continue to branch downwards. In step 3, we get the integer solution x = 1, y = 1, and the local solution is -3. According to the B&B pruning policy, we delete this node and update the global solution to -3. And so on until all nodes complete the visit.

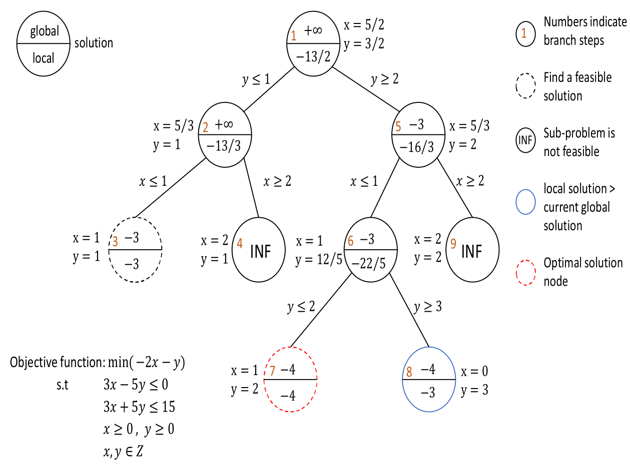

FIGURE 3. An example of B&B algorithm.

The original B&B pruning policy includes the following three cases:

- Infeasible sub-problems: if the sub-problems generated by relaxing the constraints cannot meet other constraints, then the node will be pruned (step 4 in Fig. 3).
- Find a feasible solution: as shown in Figure 3.1. If x and y are integer vectors whose sub-problems satisfy other constraints and are also feasible solutions, the node will also be pruned (step 3 in Fig. 3).
- The local solution is worse than the current global solution: since the objective function of Fig. 3 is to find the minimum value, if the local solution of the node > the current global solution, the node will continue to branch downward and will not find a better solution, so the node is pruned (Step 8 in Fig. 3).

However, the exponential computational complexity of this algorithm is impractical, prompting us to speed it up through imitation learning techniques.

D. IMITATION LEARNING OF PRUNING POLICY BY SVM CLASSIFIER

We can greatly improve the efficiency of B&B algorithms by learning optimal pruning policy to prune more nodes during the search process.

The B&B algorithm in Fig. 3 is a sequential decision process containing a state space S, an action space A, and a policy space Π. S is the set of all visited nodes, with the corresponding global solution and the current optimal solution. A is $\{prune, branch\}$. $\pi \in \Pi$ represents the mapping relationship between states and actions, that is, $\pi(s) = a$. Best pruning policy π^*, supplies the best action a^* for any state $s \in S$. In theory, we can use supervised learning to learn a policy in this sequential decision problem if we have the features and labels of all nodes. The architecture of imitation learning is similar to reinforcement learning [19], and the goal of both is to train the policy. The difference is that reinforcement learning learns through reward without knowing the best policy, while the best policy for imitation learning is known (B&B algorithm result), so it can be trained through supervised learning [19], [22].

The imitation learning problem we discuss here can be transformed into a classifying problem with feature maps through appropriate design, and then we use SVM to train the classifier and think about how to design appropriate feature maps. Designing feature is important for training classifiers, especially features that are closely related to states $s \in S$. In [22], problem-independent features and problem-dependent features are used to further improve the performance of the classifier. The following eight features are used in [19] to train the classifier.

i. Problem-independent Features:

Such features are obtained from the binary tree constructed by B&B and applicable to all MINLP problems [23]. Independent features include the following three categories.

- Node features: the depth of the node, the search depth of the node, and the local solution of the node.
- Branch feature: the branch variable of the binary tree.
- Tree features: the current global solution, the number of solutions currently obtained.

ii. Problem-dependent Features:

To choose appropriate problem-dependent features, we know from the objective function that channel and power constraints are two key factors.

- Data rate
- Power features

IV. PROPOSED RESOURCE ALLOCATION ACROSS PHYSICAL/APPLICATION LAYERS

In this paper, we propose joint source encoding rate and imitation/ensemble learning based cross-PHY-and-APP-layer resource allocation optimization for cellular network-assisted D2D video communications. The differences from the baseline scheme in Sec, III are 1) content-based source encoder rate adaption 2) the objective function in the problem formulation is PSNR of the application layer (eq. (8)) instead of the sum data rate in the physical layer. The constraint of minimum data rate in the physical layers is replaced by minimum PSNR constraint in the application layer (eq. (8d)) 3) one of the problem-dependent feature of SVM classification (branching or pruning in imitation learning of optimal pruning policy) is changed from data rate to video MSE distortion (eq.(19)) 4) we derive novel power allocation of CU based on the minimum PSNR limit of CU (Proposition 1) and then reduce the problem formulation in (8) to simplified problem formulation in (18) 5) we apply ensemble learning-based testing with downsampling based training to deal with class-imbalance and improve performance.

A. VIDEO SOURCE ENCODER

As shown in Fig. 1 (arrow for the resource allocation block to the H.264 encoder block), the proposed scheme adopts standard H.264 rate control [11], [12] to adaptively adjust the video bit rate in order to improve the video quality (PSNR). This can enhance the video transmission performance compared to schemes that only optimize physical layer resources.

B. CROSS LAYER PROBLEM FORMULATION

We design a resource management problem according to [19], changing the objective function to maximizing the minimum PSNR of D2D pairs with the following constraints. Each CU channel can only be used by one D2D pair at most, the PSNR of each CU cannot be less than a given threshold, and the transmission power has an upper limit. Thus, the resource management problem could be expressed as:

$$\underset{\{p^C,p^D,\alpha\}}{max} \max_{m\in M} PSNR_m^D(p^C, p^D, \alpha) \tag{8}$$

$$\text{subject to} \quad \alpha_{km} \in \{0, 1\}, \ \forall k \in K, \ \forall m \in M \tag{8a}$$

$$\sum\nolimits_{m\in M} \alpha_{km} \leq 1, \ \forall k \in K \tag{8b}$$

$$\sum\nolimits_{k\in K} \alpha_{km} p_{km}^D \leq P_{max}^D, \ \forall m \in M \tag{8c}$$

$$PSNR_k^C \geq PSNR_{min}^C, \ \forall k \in K \tag{8d}$$

$$p_k^C \leq P_{max}^C, \ \forall k \in K \tag{8e}$$

where α_{km} in (8a) is the channel allocation indicator for allocating channel k (CU k occupies) to D2D pair m. (8b) states that each channel k can be reused by at most one D2D pair only. (8c) states the maximum transmit power constraint P_{max}^D for each D2D pair which can occupy more than one channel. $PSNR_{min}^C$ in (8d) is the minimum PSNR limit for the CU, P_{max}^C in (8e) is the maximum transmit power limit of the CU.

C. SIMPLIFIED CROSS LAYER PROBLEM FORMULATION

In order to simplify the resource allocation problem and reduce the complexity, we derive a new expression of the CU power p^C in terms of p^D and the minimum PSNR limit parameter V_k (a_k, b_k, c_k) in the following Proposition 1. This leads to the problem of cross-layer resource allocation without CU power p^C, a simpler problem formulation.

Proposition 1: If D2D pair m occupies the channels of CU k, the CU k's optimal power of could be expressed as

$$p_k^C = \left(2^{\frac{V_k}{BW}} - 1\right) \frac{h_m^{DB} p_{km}^D + \sigma^2}{\eta h_k^{CB}} \tag{9}$$

where $V_k(a_k, b_k, c_k)$ is the minimum PSNR limit parameter containing the CU, a_k, b_k, c_k are constant that depends on the video content, the value of each CU k is different.

$$V_k(a_k, b_k, c_k) \triangleq \frac{b_k}{\frac{255*255}{10^{PSNR_{min}^C/10}} - a_k} - c_k \tag{10}$$

Proof: Appendix

According to the constraints (c)(e), we can get the new upper bound of p_{km}^D as $P_{km}^D \leq P_{km}^{max}$, where

$$p_{km}^{max} = min\left\{P_{max}^D, \left(\frac{1}{h_m^{DB}}\right)\left(\frac{\eta h_k^{CB} p_{max}^C}{2^{\frac{V_k}{BW}} - 1} - \sigma^2\right)\right\} \tag{11}$$

D2D pair m reusing the CU k channel has the following data rate:

$$\begin{aligned}
&R_{km}^D\left(p^C, p^D, \alpha\right) \\
&= BW log_2\left(1 + \eta \frac{\alpha_{km} h_m^D p_{km}^D}{h_{km}^{CD} p_k^C + \sigma^2}\right) \\
&= BW log_2\left(1 + \eta \frac{\alpha_{km} h_m^D p_{km}^D}{\frac{\left(2^{\frac{V_k}{BW}} - 1\right)(h_m^{DB} p_{km}^D + \sigma^2) h_{km}^{CD}}{\eta h_k^{CB}} + \sigma^2}\right) \\
&\triangleq \widehat{R}_{km}^D\left(p^D, \alpha\right)
\end{aligned} \tag{12}$$

For the convenience of writing, we define two variables y_{km} and z_{km}

$$y_{km} \triangleq \frac{\sigma^2}{h_m^D} + \frac{\left(2^{\frac{V_k}{BW}} - 1\right) h_{km}^{CD} \sigma^2}{\eta h_k^{CB} h_m^D} \tag{13}$$

$$z_{km} \triangleq \frac{\left(2^{\frac{V_k}{BW}} - 1\right) h_{km}^{CD} h_m^{DB}}{\eta h_k^{CB} h_m^D} \tag{14}$$

We rewrite

$$\begin{aligned}
R_m^D(p^C, p^D, \alpha) &= \sum\nolimits_k \alpha_{km} R_{km}^D(p^C, p^D, \alpha) \\
&= \sum_k \alpha_{km} BW log_2(1 + \eta \frac{\alpha_{km} h_m^D p_{km}^D}{h_{km}^{CD} p_k^C + \sigma^2}) \\
&= \sum_k BW log_2(1 + \eta \frac{\alpha_{km} h_m^D p_{km}^D}{h_{km}^{CD} p_k^C + \sigma^2}) \\
&= \sum_k BW log_2(1 + \eta \frac{\alpha_{km} p_{km}^D}{y_{km} + z_{km} p_{km}^D}) \\
&\triangleq \widehat{R}_m^D\left(p^D, \alpha\right)
\end{aligned} \tag{15}$$

Then the MSE of D2D pair m can be rewritten as:

$$\begin{aligned}
MSE_m^D\left(p^C, p^D, \alpha\right) &= a_m + \frac{b_m}{R_m^D + c_m} \\
&= a_m + \frac{b_m}{\widehat{R}_m^D + c_m} \\
&\triangleq \widehat{MSE}_m^D\left(p^D, \alpha\right)
\end{aligned} \tag{16}$$

Then the PSNR of (3.2) D2D pair m can be rewritten as:

$$\begin{aligned}
PSNR_m\left(p^C, p^D, \alpha\right) &= 10 log_{10} \frac{255 * 255}{MSE_m^D} \\
&= 10 log_{10} \frac{255 * 255}{\widehat{MSE}_m^D} \\
&\triangleq \widehat{PSNR}_m^D\left(p^D, \alpha\right)
\end{aligned} \tag{17}$$

Finally, we reduce the original resource allocation problem (8) to the following simplified problem

$$\max_{\{p^D, \alpha\}} \min_{m \in M} \widehat{PSNR}_m^D\left(p^D, \alpha\right) \tag{18}$$

$$\text{subject to} \quad \alpha_{km} \in \{0, 1\}, \ \forall k \in K, \ \forall m \in M \tag{18a}$$

$$\sum\nolimits_{m \in M} \alpha_{km} \leq 1, \ \forall k \in K \tag{18b}$$

$$\sum\nolimits_{k \in K} \alpha_{km} p_{km}^D \leq P_{max}^D, \ \forall m \in M \tag{18c}$$

$$P_{km}^D \leq P_{km}^{max}, \ \forall k \in K, \ \forall m \in M \tag{18d}$$

D. ACCELERATING METHOD THROUGH IMITATION LEARNING

The baseline scheme [19] designed the data rate as a problem-dependent feature according to the channel parameters of D2D communication. This method only considered the physical layer. Here we change the problem-dependent feature to consider the MSE of the application layer.

The problem-dependent features now include

- Video MSE distortion features (proposed) (19)
- Power features

The problem-independent features are:

- the depth of the node
- the search depth of the node
- the local solution of the node.
- the branch variable of the binary tree.
- the current global solution
- the number of solutions currently obtained.

SVM binary classifiers are trained with supervised learning methods. The training dataset contains all searched nodes by B&B. But as the number of nodes grows exponentially, this approach consumes a lot of memory. An iterative training method is commonly used in imitation learning, named "Dataset Aggregation" (DAgger) [26], to solve the problem of massive memory consumption. The main concept of Dagger is to do this by collecting a dataset with the current policy at each iteration, and then using the aggregation of all the collected datasets to train a new policy.

The biggest difference between imitation learning and traditional supervised learning is the way of collecting features and labels of classifiers. For supervised learning, all features and labels must be available before the training process. In contrast, imitation learning uses DAgger to iteratively collect data to update the learned policy. Most papers show that DAgger has been widely used for imitation learning and provides strong performance guarantees [23], [26].

The pseudocode of training processing with Dagger is shown in Algorithm 1.

Algorithm 1 Training process for imitation learning

Initialization:
In the first iteration we take π^* as the initial strategy and denote it as $\pi^{(1)}$ and the initial training dataset $\mathcal{D}$ is empty.

The number of iterations is n. Then we use $\pi^{(1)}$ to search the problem set Q and collect the data into the training data set $\mathcal{D}$.

We stack the six problem independent features and two problem dependent features (one feature in (19) is different from baseline scheme [19]) in an eight-dimensional vector
for $n = 1$ to N do
  for problem Q in D do
    $\mathcal{D}^{(Q)} \leq$ ftarrow DataCollection(Q, π^n)
    $\mathcal{D} \leq$ ftarrow $D \cup \mathcal{D}^{(Q)}$
  end for
  $\pi^{(n+1)} \leq$ ftarrow train SVM using D
end for
After n iterations, we choose the best performing policy as π^*
return π^*

E. DOWNSAMPLING TRAINING AND ENSEMBLE LEARNING TESTING

The class imbalance in the datasets impact negatively the accuracy of the classification task. In the problem of prune policy learning in previous subsection, the {prune} class is significantly larger than {branch} class. We can either upsample the minority class {branch} or downsample the majority class{prune}. However, downsampling performs better in general [25].

We are motivated to propose ensemble learning similar to [24] to improve performance. The difference is that the training datasets are obtained by downsampling the {prune} class such that {prune} class has the same size as the {branch} class (data is balance), in a way similar to [25]. Specifically, the set of {prune} is downsampled such that M disjoint subset $S_{prune,1}$, $S_{prune,M}$, is formed. The set of branch, S_{branch}, is paired with $S_{prune,1}$, $S_{prune,M}$, respectively, to trained M classified. In the testing stage, Ensemble learning is applied to pick the best one to increase classification accuracy.

V. SIMULATION RESULTS

As shown in Fig 2, we consider a single-cell cellular network-assisted D2D communication system with a radius of 500 meters. The BS is located in the center of the cell, and the CUs are uniformly distributed in the cell, and the D2D pairs are also uniformly distributed in the cell. Table 1 summarizes our simulation parameters. According to [27] and [28], we set the bandwidth of each subcarrier to 15 kHz.

TABLE 1. Simulation parameters.

Parameters	Value
Cellular users (K)	5/7
D2D pairs (M)	2/3
The radius of the cell	500 m
The distance of D2D, r_{min}, r_{max}	15 m, 50 m
Pass loss CU	128.1+37.6log(distance)
Pass loss D2D	148+40log(distance)
Shadowing	10 dB
Max transmit power CU/D2D	20 dBm
Subcarriers bandwidth	15 kHz
Training samples	100
Testing samples	20

We compare the PSNR performance, the video quality, of the proposed cross layer scheme in Sec. IV, the physical-layer-only baseline scheme in Sec. III [19], and the cross layer baseline scheme in [13].

As shown in Fig. 4, our proposed joint source encoding rate control and imitation/ensemble learning based cross-layer resource allocation scheme is 4.03∼5.37 dB higher than the physical layer baseline scheme that only considered the physical layer in different scenarios. The reasons are three-folded: 1) The proposed scheme has source encoder rate control and can adaptively adjust video rate to increase PSNR. The baseline scheme does not. 2) The proposed scheme uses the objective function of PSNR directly and allocates the subcarrier and power considering the different rate-distortion function of users' videos. The baseline scheme disregards the different rate-distortion function of users' videos. 3) The proposed scheme uses imitation learning with downsampling the majority set to compensate the class imbalance problem and improve classification accuracy. The proposed cross layer scheme also outperforms the cross layer baseline scheme [13] because the proposed scheme is imitation learning based on

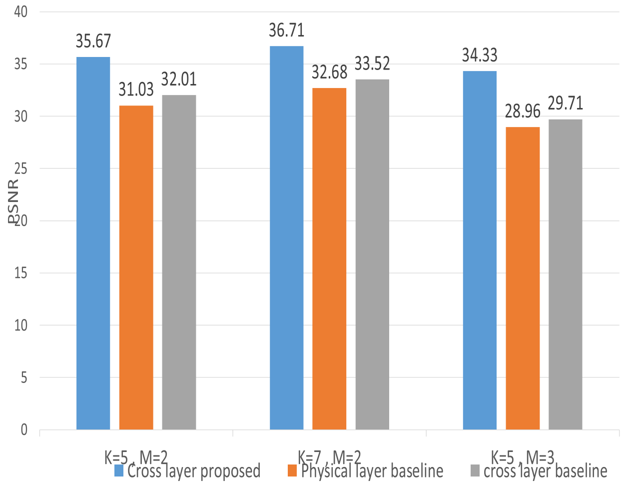

FIGURE 4. Average PSNR of the proposed cross layer scheme, the physical layer baseline scheme [19], and cross layer baseline scheme [13] in different scenarios.

the globally joint optimal B&B algorithm, and the cross layer baseline scheme [13] is suboptimal. The cross layer baseline scheme [13] is slightly better than the physical layer baseline scheme [19].

The main purpose of using imitation learning is learning the pruning policy to speed up the branching process of B&B. In addition, it is very difficult and expensive to obtain a large number of training samples in wireless networks. The algorithm of imitation learning combined with B&B can complete the training process with only a small amount of training data [22]. We set the number of training samples and test data samples to 100 and 20, respectively. We can take advantage of the exponentially increasing number of nodes in a B&B tree, only 100 training samples are sufficient [19]. Because the 100 training samples already contain tens of thousands of nodes [19], [22].

To compare the cross-layer resource allocation performance of accelerated B&B using imitation learning, three parameters are used here to measure the simulation results. Gap is the objective function gap between the traditional B&B algorithm and accelerating method through imitation learning. Speed is the ratio of the number of search nodes between the traditional B&B algorithm (as 1x) and accelerating method through imitation/ensemble learning. The higher the ratio, the less computing time and the better the acceleration result of imitation learning. Pruning accuracy is the accuracy of pruning non-optimal nodes using the pruning policy via imitation/ensemble learning.

The three results for different number of CUs (K) and D2D pairs (M) are compared in Table 2. In the scenario of $K = 5$, $M = 2$, the imitation/ensemble learning-based pruning policy makes the B&B search process 2.25 times faster, while the gap to the optimal B&B bound is only 2.05%. As the number of users and subcarriers increases, in a more complex scenario $K = 5$, $M = 3$, this method is 13.99 times speed as

TABLE 2. The effectiveness of imitation/ensemble learning in different user scenarios.

Parameter	K=5, M=2
Gap	2.05%
Speed	2.25x
Pruning accuracy	95.57%

Parameter	K=7, M=2
Gap	3.80%
Speed	4.98x
Pruning accuracy	93.91%

Parameter	K=5, M=3
Gap	5.13%
Speed	13.99x
Pruning accuracy	90.05%

B&B, while the gap with respect to B&B bound is 5.13%. The method we use can achieve a small objective function gap with a very small amount of training data (about 100), while also reducing the computational complexity of B&B. Compared with the one million training samples in [18] and tens of thousands of training samples in [29], imitation learning combined with B&B is more likely to be implemented in real life [22].

Good generalization ability allows the model training in a small network to be used in a larger network without retraining. In Table 3, the trained $K = 5$, $M = 2$ model is used to test the performance in different scenarios. The simulation results show that using the smaller scenario ($K = 5$, $M = 2$) to predict the larger scenario ($K = 7$, $M = 2$) has only 0.45% more in gap, and at comparable speed. It can be seen from this that if

TABLE 3. The generalization ability of imitation/ensemble learning in different user scenarios.

Parameter	Policy (K=7, M=2) in Scenario (K=7, M=2)	Policy (K=5, M=2) in Scenario (K=7, M=2)
Gap	3.80%	4.25%
Speed	4.98x	5.03x

the training data is insufficient in the larger scenario (training sample is more difficult to obtain), the model already trained in smaller scenario can be used to assist.

The generalization capability is further studied in Table 4. The policy (K = 5, M = 2) perform acceptably well even for Scenario (K = 10, M = 2), whose complexity is $2^{10*2}/2^{5*2} = 2^{10}$ times Scenario (K = 5, M = 2) and its training samples are extremely difficult to obtain. It shows the proposed scheme has good generalization capabilities and can learn the policy from the small scale (small K*M) scenario and apply the policy to large scale (larger K*M) scenario even for 2^{10} times bigger scenario.

TABLE 4. The generalization ability of imitation learning of policy (K = 5, M = 2) in different user scenarios.

Parameter	Scenario (K=7, M=2)	Scenario (K=8, M=2)	Scenario (K=10, M=2)
Gap	4.25%	5.01%	5.91%
Speed	5.03x	9.18x	20.43x

A. TIME COMPLEXITY COMPARISON OF PROPOSED SCHEME AND PHYSICAL LAYER BASELINE SCHEME

We explore the time complexity of the proposed cross layer scheme and baseline physical-layer-only scheme. The computing platform we use is a desktop PC with CPU i7-12700, without GPU, and with python version 3.6.

We follow the physical layer baseline scheme [19] and use execution time as the time complexity. At each testing sample, the network realization (channel state information etc.) is randomly generated. The B&B tree and execution time of different testing samples would be different. The execution time averaged over 20 testing samples is shown in Table 5 for scenario K = 7, M = 2. It indicates the computational complexity/ latency of the proposed scheme is only slightly larger than that of physical layer baseline scheme.

TABLE 5. Execution time for physical-layer baseline scheme and proposed scheme.

	Physical Layer Baseline Scheme	Proposed Scheme
Execution time	1.44sec	1.54sec

The explanation is as follows. By replacing PHY data rate feature by APP layer video MSE distortion feature in (19) in the SVM binary classifier, the number of features is kept the same as that of physical layer baseline scheme, and thus the complexity of SVM binary classifier is about the same. Furthermore, B&B algorithm objective is changed to PSNR from data rate, so there is additional mapping from data rate to PSNR by (5) and (6). Therefore, the execution time/ computational complexity of the proposed scheme is just a little more than that of the physical layer baseline scheme.

VI. CONCLUSION

In this paper, we proposed joint content-based video encoder rate adaption and optimal cross layer subcarrier/power allocation based on imitation /ensemble learning for video transmission in cellular networks assisted uplink D2D communications. The proposed scheme has source encoder rate control and can adaptively adjust video rate to increase PSNR. For comparison, the baseline scheme does not use the content-based video rate adaption. Furthermore, the proposed scheme uses the objective function of PSNR directly and allocate the subcarrier and power considering the different rate-distortion function of users' videos. Using the minimum PSNR limit of the cellular user, we derive an objective function that is independent of the transmission power of the cellular user to simplify the optimization problem formulation. For comparison, the baseline scheme uses physical layer only objective function and constraints. Finally, the proposed scheme adopt ensemble learning with downsampling the majority set {prune} to alleviate the class imbalance problem and improve performance. The simulation results show the proposed scheme outperforms the baseline scheme by 4∼5 dB in PSNR. Compared with the traditional B&B, the speed the proposed scheme is increased by 2.05 to 20.43 times (execution time is 50%-5%), depending on the scale of the scenario. In addition, the practice of accelerating the B&B branch through imitation/ensemble learning has a certain degree of generalization ability, and only needs one hundred training data. The proposed scheme has good generalization capabilities and can learn the policy from the small scale scenario and apply the policy to even 210 times larger scenario.

APPENDIX
PROOF OF PROPOSITION 1

If D2D pair m reuses the channels of CU k, the following equations must be satisfied to guarantee the minimum PSNR limit of CU k.

$$R_{min}^{C} = BWlog_2\left(1+\eta\frac{h_k^{CB}p_k^C}{h_m^{DB}p_{km}^D+\sigma^2}\right) \tag{A-1}$$

Then, we can calculate the maximum MSE and minimum PSNR of CU K as:

$$MSE_{max}^{C} = a_k + \frac{b_k}{BWlog_2\left(1+\eta\frac{h_k^{CB}p_k^C}{h_m^{DB}p_{km}^D+\sigma^2}\right)+c_k} \tag{A-2}$$

$$PSNR_{min}^{C}=10log_{10}\left(\frac{255*255}{a_k+\frac{b_k}{BWlog_2\left(1+\eta\frac{h_k^{CB}p_k^C}{h_m^{DB}p_{km}^D+\sigma^2}\right)+c_k}}\right) \tag{A-3}$$

where a_k, b_k, c_k are constant that depends on the video content and is different for each CU.

After that, we can get

$$p_k^C=\left(2^{\frac{\frac{b_k}{\frac{255*255}{10^{PSNR_{min}^C/10}}-a_k}-c_k}{BW}}-1\right)\frac{h_m^{DB}p_{km}^D+\sigma^2}{\eta h_k^{CB}} \tag{A-4}$$

For the convenience of writing, we define a variable here V_k

$$V_k(a_k,b_k,c_k)\frac{b_k}{\frac{255*255}{10^{PSNR_{min}^C/10}}-a_k}-c_k \tag{A-5}$$

Finally, get

$$p_k^C=\left(2^{\frac{V_k}{BW}}-1\right)\frac{h_m^{DB}p_{km}^D+\sigma^2}{\eta h_k^{CB}}. \tag{A-6}$$

ACKNOWLEDGMENT

(Shu-Ming Tseng, Jun-Jie Wu, and Chao Fang contributed equally to this work.)

REFERENCES

[1] P. Khuntia and R. Hazra, "An efficient channel and power allocation scheme for D2D enabled cellular communication system: An IoT application," *IEEE Sensors J.*, vol. 21, no. 22, pp. 25340–25351, Nov. 2021.

[2] Y.-J. Chen, K.-M. Liao, M.-L. Ku, and F. P. Tso, "Mobility-aware probabilistic caching in UAV-assisted wireless D2D networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Waikoloa, HI, USA, Dec. 2019, pp. 1–6.

[3] M.-L. Ku and J.-W. Lai, "Joint beamforming and resource allocation for wireless-powered device-to-device communications in cellular networks," *IEEE Trans. Wireless Commun.*, vol. 16, no. 11, pp. 7290–7304, Nov. 2017.

[4] M. Jo, T. Maksymyuk, B. Strykhalyuk, and C.-H. Cho, "Device-to-device-based heterogeneous radio access network architecture for mobile cloud computing," *IEEE Wireless Commun.*, vol. 22, no. 3, pp. 50–58, Jun. 2015.

[5] P. Wu, P. C. Cosman, and L. B. Milstein, "Subcarrier assignment and power allocation for device-to-device video transmission in Rayleigh fading channels," *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 5967–5981, Sep. 2017.

[6] H.-H. Nguyen, M. Hasegawa, and W.-J. Hwang, "Distributed resource allocation for D2D communications underlay cellular networks," *IEEE Commun. Lett.*, vol. 20, no. 5, pp. 942–945, May 2016.

[7] T. Chen, X. Zhang, M. You, G. Zheng, and S. Lambotharan, "A GNN-based supervised learning framework for resource allocation in wireless IoT networks," *IEEE Internet Things J.*, vol. 9, no. 3, pp. 1712–1724, Feb. 2022.

[8] S. A. Gbadamosi, G. P. Hancke, and A. M. Abu-Mahfouz, "Interference avoidance resource allocation for D2D-enabled 5G narrowband Internet of Things," *IEEE Internet Things J.*, vol. 9, no. 22, pp. 22752–22764, Nov. 2022.

[9] Y. Xu, Z. Liu, C. Huang, and C. Yuen, "Robust resource allocation algorithm for energy-harvesting-based D2D communication underlaying UAV-assisted networks," *IEEE Internet Things J.*, vol. 8, no. 23, pp. 17161–17171, Dec. 2021.

[10] Y.-H. Jung, Q. Song, K.-H. Kim, P. Cosman, and L. B. Milstein, "Cross-layer resource allocation using video slice header information for wireless transmission over LTE," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 28, no. 8, pp. 2024–2037, Aug. 2018.

[11] D. Wang, L. Toni, P. C. Cosman, and L. B. Milstein, "Uplink resource management for multiuser OFDM video transmission systems: Analysis and algorithm design," *IEEE Trans. Commun.*, vol. 61, no. 5, pp. 2060–2073, May 2013.

[12] F. Li, T. Wang, and P. C. Cosman, "Joint rate adaptation and resource allocation for real-time H.265/HEVC video transmission over uplink OFDMA systems," *Multimedia Tools Appl.*, vol. 78, no. 18, pp. 26807–26831, Sep. 2019.

[13] S. Ullah, K. Kim, A. Manzoor, L. U. Khan, S. M. A. Kazmi, and C. S. Hong, "Quality adaptation and resource allocation for scalable video in D2D communication networks," *IEEE Access*, vol. 8, pp. 48060–48073, 2020.

[14] Q.-N. Tran, N.-S. Vo, M.-P. Bui, T.-M. Phan, Q.-A. Nguyen, and T. Q. Duong, "Spectrum sharing and power allocation optimised multihop multipath D2D video delivery in beyond 5G networks," *IEEE Trans. Cognit. Commun. Netw.*, vol. 8, no. 2, pp. 919–930, Jun. 2022.

[15] *Cisco Visual Networking Index: Forecast and Trends*, Cisco Syst., Inc., San Jose, CA, USA, 2022.

[16] S.-M. Tseng and W.-Y. Chen, "Cross-layer codebook allocation for uplink SCMA and PDNOMA-SCMA video transmission systems and a deep learning-based approach," *IEEE Syst. J.*, vol. 17, no. 1, pp. 294–305, Mar. 2023.

[17] D. Ron and J.-R. Lee, "DRL-based sum-rate maximization in D2D communication underlaid uplink cellular networks," *IEEE Trans. Veh. Technol.*, vol. 70, no. 10, pp. 11121–11126, Oct. 2021.

[18] H. Sun, X. Chen, Q. Shi, M. Hong, X. Fu, and N. D. Sidiropoulos, "Learning to optimize: Training deep neural networks for interference management," *IEEE Trans. Signal Process.*, vol. 66, no. 20, pp. 5438–5453, Oct. 2018.

[19] M. Lee, G. Yu, and G. Y. Li, "Learning to branch: Accelerating resource allocation in wireless networks," *IEEE Trans. Veh. Technol.*, vol. 69, no. 1, pp. 958–970, Jan. 2020.

[20] Y. Shen, Y. Shi, J. Zhang, and K. B. Letaief, "Graph neural networks for scalable radio resource management: Architecture design and theoretical analysis," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 1, pp. 101–115, Jan. 2021.

[21] X. Zhang, H. Zhao, J. Xiong, X. Liu, L. Zhou, and J. Wei, "Scalable power control/beamforming in heterogeneous wireless networks with graph neural networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2021, pp. 01–06.

[22] Y. Shen, Y. Shi, J. Zhang, and K. B. Letaief, "LORM: Learning to optimize for resource management in wireless networks with few training samples," *IEEE Trans. Wireless Commun.*, vol. 19, no. 1, pp. 665–679, Jan. 2020.

[23] H. He, H. Daume III., and J. M. Eisner, "Learning to search in branch and bound algorithms," in *Proc. Adv. Neural Inf. Process. Syst.*, Dec. 2014, pp. 3293–3301.

[24] T. G. Dietterich, "Ensemble learning," in *The Handbook of Brain Theory and Neural Networks*, vol. 2. Cambridge, MA, USA: MIT Press, 2002, pp. 110–125.

[25] S. Tyagi and S. Mittal, "Sampling approaches for imbalanced data classification problem in machine learning," in *Proc. ICRIC*. Cham, Switzerland: Springer, 2019, pp. 209–221.

[26] S. Ross, G. Gordon, and D. Bagnell, "A reduction of imitation learning and structured prediction to no-regret online learning," in *Proc. 14th Int. Conf. Artif. Intell. Statist.*, 2011, pp. 627–635.

[27] G.-M. Kang, J. Lee, and O.-S. Shin, "Power allocation scheme for D2D communications in an OFDM-based cellular system," in *Proc. Int. Conf. Inf. Netw. (ICOIN)*, Jan. 2015, pp. 388–389.

[28] R. Yin, C. Zhong, G. Yu, Z. Zhang, K. K. Wong, and X. Chen, "Joint spectrum and power allocation for D2D communications underlaying cellular networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 4, pp. 2182–2195, Apr. 2016.

[29] S.-M. Tseng, G.-Y. Chen, and H.-C. Chan, "Cross-layer resource management for downlink BF-NOMA-OFDMA video transmission systems and supervised/unsupervised learning based approach," *IEEE Trans. Veh. Technol.*, vol. 71, no. 10, pp. 10744–10753, Oct. 2022.

[30] F. Li, F. Shuang, Z. Liu, and X. Qian, "A cost-constrained video quality satisfaction study on mobile devices," *IEEE Trans. Multimedia*, vol. 20, no. 5, pp. 1154–1168, May 2018.

[31] W.-Y. Lin, T.-H. Chang, and S.-M. Tseng, "Deep learning-based cross-layer power allocation for downlink cell-free massive multiple-input–multiple-output video communication systems," *Symmetry*, vol. 15, no. 11, p. 1968, Oct. 2023.

SHU-MING TSENG (Member, IEEE) received the B.S. degree in electrical engineering from National Tsing Hua University, Taiwan, in 1994, and the M.S. and Ph.D. degrees in electrical engineering from Purdue University, West Lafayette, IN, USA, in 1995 and 1999, respectively. From 1999 to 2001, he was with the Department of Electrical Engineering, Chang Gung University, Taiwan. Since 2001, he has been with the Department of Electronic Engineering, National Taipei University of Technology, Taipei City, Taiwan, where he has been a Professor, since 2007. He was a Visiting Scholar with the University of California at San Diego, USA, from 2014 to 2015. He is the author of 66 SCI journal articles, including 21 in IEEE. His research interests include AI, B5G, wireless resource allocation, cross-layer design, and computer vision. He was a TPC Member of the IEEE IWCMC 2022 LowPowerIoT Symposium. He was an External Member of Taiwan Stock Exchange Listing Review Committee. He was the TPC Co-Chair of the IEEE 2021 30th Wireless and Optical Communications Conference (WOCC 2021). He is an Editor of *KSII Transactions on Internet and Information Systems*.

JUN-JIE WU received the M.S. degree in electronic engineering from the Department of Electronic Engineering, National Taipei University of Technology, Taiwan, in 2022. His research interests include B5G transmission technology, cross-layer resource allocation, and deep learning.

CHAO FANG (Senior Member, IEEE) received the B.S. degree in information engineering from Wuhan University of Technology, Wuhan, China, in 2009, and the Ph.D. degree from the State Key Laboratory of Networking and Switching Technology in Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing, China, in 2015. From 2013 to 2014, he was funded by China Scholarship Council to visit Carleton University, Ottawa, ON, Canada, as a Joint Doctorate. In 2016, he joined Beijing University of Technology, Beijing. He is currently an Associate Professor. Moreover, he is a Visiting Scholar with the University of Technology Sydney, Sydney, NSW, Australia; the Commonwealth Scientific and Industrial Research Organization, Canberra, ACT, Australia; The Hong Kong Polytechnic University, Hong Kong; Kyoto University, Kyoto, Japan; Muroran Institute of Technology, Muroran, Japan; and the Queen Mary University of London, London, U.K. His research interests include information-centric networking, big data for networking, cloud-edge-terminal computing, resource management, and content delivery. He was a recipient of the Best Paper Award of IEEE ICFEICT'2022. From 2022 to 2023, he was the Vice Chair of the Technical Affairs Committee in IEEE ComSoc Asia/Pacific Region. He was also the Session Chair of ICC NGN'2015 and ICCC NMNRM'2021 and the Poster Co-Chair of HotICN'2018. Moreover, he is the Leading Editor of *Electronics* and *Symmetry* Special Issues.