

Resource Allocation for Uncoded Multi-user Video Transmission over Wireless Networks

Dian Liu¹ · Hao Cui¹ · Jun Wu¹ · Chong Luo²

Published online: 28 March 2016
© Springer Science+Business Media New York 2016

Abstract In wireless networks, multi-user video streaming under limited resource is a challenging problem. The main challenge is how to meet the transmission requirements under the different channel condition and video content complexity. In this paper, we propose an uncoded video transmission framework to deliver the multi-user video over wireless networks. In order to evaluate the overall performance of multi-user network more practically, three optimization strategies are proposed in this paper: 1) minimizing the total distortion; 2) minimizing the maximal distortion; 3) minimizing the summation of square root distortion. Furthermore, the corresponding joint resource allocation algorithms are developed to solve the optimization problems. The simulation results demonstrate that different optimization strategies have different resource allocation and performance for each user. The optimization strategy 1) performs the best in terms of average PSNR of all users, the optimization strategy 2) achieves more fair result, and the optimization strategy 3) achieves a good balance. The

strategy 3) can improve the performance for users with bad channel condition, while little loss is caused for users with good channel condition.

Keywords Multi-user · Uncoded video transmission · Optimal resource allocation

1 Introduction

The mobile traffic surges with prevalence of smart phone and development of wireless access technology. Among the mobile traffic, the video traffic dominates. According to Cisco Visual Networking Index (VNI), the video traffic will increase to 67 % percentage of mobile traffic by 2017 [1]. The increase of wireless infrastructure and spectrum resource lags behind traffic, thus we need to utilize wireless resource in more efficient way. How to guarantee the multi-user service quality over resource-limited wireless network is one of the most challenging problems.

In a multi-user video streaming service, the video contents required by the users are different. The base station (BS) needs to transmit all these video streams at the same time, thus facing some challenges. First, the rate-distortion (R-D) curve of a video is content depended. The resource scheduler at the base station needs to estimate the heterogeneous R-D curves and update them periodically. Second, the resource at the base station is limited. Third, the channel condition for each user is heterogeneous. In order to efficiently allocate the resource, the base station needs each user to feedback his channel condition and keep track of it. To achieve a good overall performance, BS is necessary to perform a joint optimization of resource allocation based on the video content complexity and channel condition subject to some limits.

✉ Hao Cui
hao.cui@live.com

Dian Liu
1310518@tongji.edu.cn

Jun Wu
wujun@tongji.edu.cn

Chong Luo
chong.luo@microsoft.com

¹ College of Electronics and Information Engineering, Tongji University, Shanghai 201804, People's Republic of China

² Microsoft Research Asia, Beijing 100080, People's Republic of China

In traditional digital video communication system, a video source is compressed into bit streams according source rate. The channel bit rate is adaptively determined by the channel condition. However, in the multi-user scenario, different video content have different R-D characteristics. It is difficult to derive the source rate and channel bit rate. Uncoded video transmission has attracted a lot of attention since Softcast [2] emerged. However, most researches focus on broadcasting and point to point unicast till now. In this paper we investigate uncoded video transmission for multi-user streaming. In uncoded video transmission, the video source is first decorrelated by an orthogonal transform, e.g. DCT. Without quantization and entropy coding, the DCT coefficients are scaled by a linear factor and then are directly transmitted by the amplitude modulation. Uncoded transmission can adapt to channel condition automatically. The resource scheduler only needs to know the coefficient distribution instead of complicated R-D curve.

The analysis of Softcast [2] has shown that the uncoded scheme could achieve nearly the same end-to-end distortion as the more complicated digital system for a single video and unknown channel condition case. However when it is applied in multi-user video streaming service, the video content diversity and channel condition diversity should be considered in resource allocation. Intuitively, high complexity video should to be transmitted with more resources. However, a multi-user video transmission system should consider not only the reconstructed video quality of each individual user but also network resource and status. So under the bandwidth and power resources constraints, it is a challenging task to allocate the resources to different users to achieve the optimal network performance.

In this paper, we propose an uncoded video transmission framework in multi-user scenario. Firstly, we formulate the multi-user resource allocation as a three-dimension optimization problem, which is constrained by power, channel signal to noise power ratio (SNR) and bandwidth. Secondly, we apply three optimization strategies: 1) minimizing the total distortion, 2) minimizing the maximal distortion, 3) minimizing the summation of square root distortion. These optimization problem can be solved by the joint bandwidth and power allocation. Lastly, the proposed algorithm is validated through simulation.

This paper is organized as follows. Section 2 describes the related work in recent years. The system architecture for multi-user uncoded video transmission is described in Section 3. Section 4 focuses on different optimization strategies and how to solve these optimization problems. Simulation results are presented in Section 5 and conclusions are drawn in Section 6.

2 Related work

2.1 Uncoded video transmission

K. H. Lee [3] proposed and proved the optimal linear coding for vector channels in his pioneering theoretic work. After more than 30 years, Softcast applied this theory in wireless video, with considering the scalability and robustness of video. Softcast performs a 3D-DCT to a group of pictures and a linear coding proposed in [3] to the DCT coefficients with the total power constraint. The emergence of Softcast brings a surge of interest in uncoded video transmission systems. Follow-up researches includes WaveCast [4], Cactus [5], and Robust [6]. Among these schemes, WaveCast based on lossy transmission and 3D wavelet transform, utilizes motion compensated temporal filter (MCTF) to exploit inter frame redundancy. The motion vectors are transmitted in BPSK with 1/2 coding rate, and the coefficients are linear coded, quantized and transmitted in wireless channel. Cactus uses open-loop video compressing method to remove inter and intra redundancy, and transmits the residual in an uncoded way. Cactus transmits the pixel values for denoising purpose while Softcast transmits the DCT coefficients. These two system just like Softcast, only take AWGN channel into consideration. With consideration of fast fading channel, Robust [6] allocates the bandwidth and power according to priority of data, which obtains performance gain from the channel diversity in fading channel. But all these systems consider single video transmission only, not multi videos.

2.2 Multiuser video transmission systems

In digital video communication, optimal video transmissions in multiuser systems have been well studied. Joint source/channel rate adaptation and power allocation are proposed to achieve the optimal objectives subject to constraints on limited network resources, e.g. [7–11]. In [7], G. M. Su formulated two optimal problems, one considering the fairness aiming to minimizing the maximal expected distortion, and the other considering efficiency aiming to minimizing the total expected distortion. By constructing a continuous R-D function through linear interpolation of the discrete R-D function, these two optimal problems were solved. In [8], a fairness scheme was proposed to minimize the maximum the expected distortion among all users. By utilizing a piecewise linear R-D model, the optimal encoding rate can be determined and the end-to-end BER can also be satisfied by a water-filling power allocation. In [9], the optimal objectives aiming to minimize the sum distortion was solved in a distributed way, the local users reported

their distortion price to the base station and the base station updated resources allocation according to the price. In [12], a dynamic programming framework was proposed to maximize the expected discounted accumulated video quality. However the framework requires detailed knowledge of the statistics of the experienced environment dynamics, which is often unavailable before transmission. So they also proposed an online learning algorithm to overcome this obstacle. In [10], a more accurate end-to-end distortion estimation is adopted. Combining with the packet loss problem, an optimal resources allocation scheme is proposed aiming to minimizing the total expected distortion. In [11], V. Joseph et al. considered the impact of Perceived Video Quality (PVQ) across the sequence of scenes, and proposed an optimal online algorithm to maximize user QoE by considering tradeoffs among mean, variance and fairness.

The main difference among [7–10] is the R-D model. Su et al. [7, 8] adopt the linear model, and [9] improves the R-D accuracy in a distributed way but at a cost of communication overhead. Huang et al. [9] assumes that the distortion of all sequence segments are the same. Maani et al. [10] adopts a more accurate nonlinear R-D model. Representing distortion precisely is the key point to allocate resources in digital multi-user video communications. In the uncoded schemes, the distortion can be written as a close-form expression of bandwidth and power, and it is convex, optimization solution can be derived for optimal resource allocation.

3 System model

In this section, we present an overview of proposed framework. Figure 1 illustrates the system model of uncoded multi-user video transmission in the single base station (BS) scenario. The BS delivers the video sources to several users, who may be in different place and have different channel conditions. The multiple users need share the limited bandwidth and power resources to transmit their own video streaming. For wireless multi-user networks, the main concern is not only the individual user quality, but also the overall performance. Equal allocation of bandwidth and power is inefficient and unfair, so we proposed an optimal bandwidth and power allocation in this paper.

We first introduce the analog video transmission scheme and discuss the relationship between distortion and resource allocation. Then we formulate the joint bandwidth and power allocation problem as a distortion-cost function.

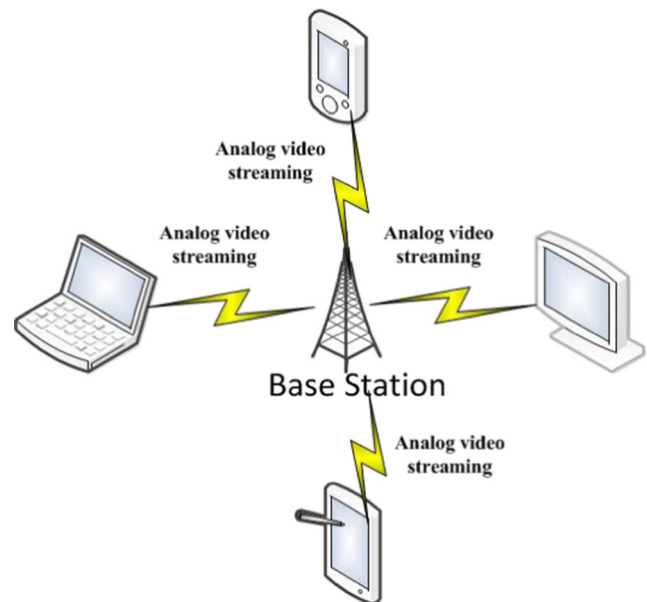


Fig. 1 System model of uncoded multi-user video transmission

3.1 Overview of uncoded video transmission

Considering a multi-user wireless network as shown in Fig. 2, where the base station transmits the video sources to each user in an uncoded way. With the limited bandwidth and power resources, the base station allocates the resources to each user aiming to achieve the best overall performance.

Before transmission, each video divides frames into group of pictures (GOP), and the size of GOP can vary from 4, 8, 16 or 32. We adopt GOP size 8 meaning that there are 8 pictures in one GOP. Then, we apply a 3D-DCT to each GOP for the purpose of de-correlation in both space and time domain, and the DCT coefficients are divide into equal-sized chunks. In our scheme, each frame is divided into multiple 8x8 chunks for producing few side information.

Usually the content complexity is shown by how heavily unequal energy distribution of DCT coefficients in a GOP are. In relative static videos, its energy concentrates in a few low frequency coefficients. Thus discarding these low energy coefficients may have negligible effects on video recovery quality. So these high energy coefficients need to be well protected. This gives us hint that relative static videos can be transmitted with less bandwidth and the saved power can use to protect other high-priority coefficients. On the other hand, the coefficients of high-motion videos distribute smoothly, and most of them have a certain effect on video recovery quality. So it must be transmitted with more bandwidth.

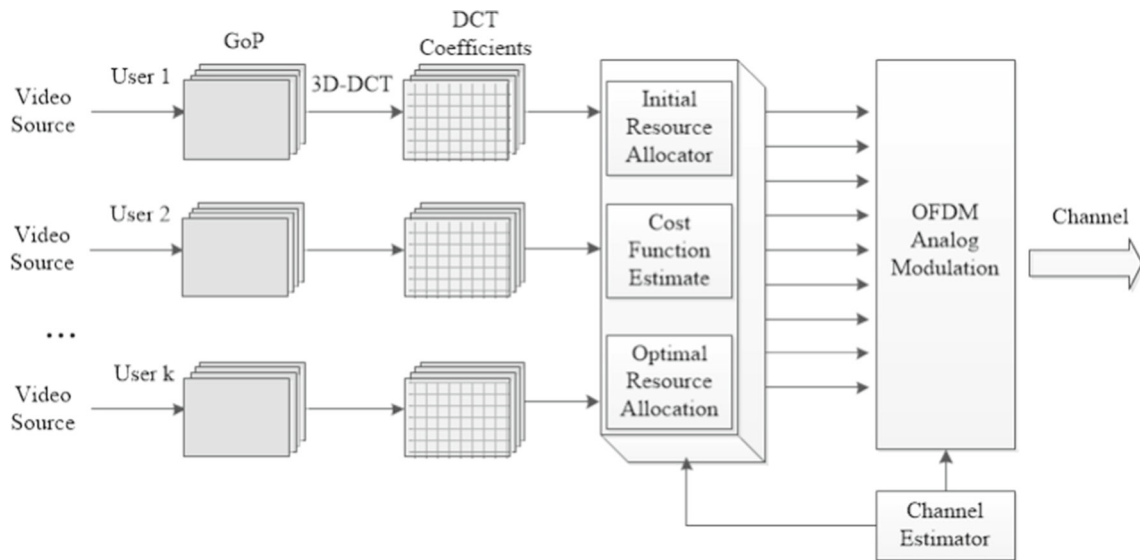


Fig. 2 Overview of uncoded schemes

According to the coefficients and the channel condition, the scheduler will allocate the bandwidth and power to improve the overall performance among all users. If the bandwidth and power are not abundant, some coefficients will be discarded. The coefficients to be transmitted will be multiplied by a scaling factor as a linear coding for error protection. The scaled coefficients are directly mapped to the I and Q plane. In addition, the meta data, including the variance of the chunks and the variance of the GOP, should be transmitted through conventional digital modulation at a reliable transmission rate (e.g. BPSK with 1/2 coding). Based on the estimated channel condition, the DCT coefficients can be recovered by MMSE (minimum mean square error) detection. Finally, the entire GOP can be reconstructed through de-scaling and inverse 3D-DCT.

Firstly, we discuss the optimal uncoded video transmission along with the corresponding distortion characteristics. Given a video source, the optimal transmission is minimizing the expected distortion. For user k , the DCT coefficients of a GOP s_k are divided into N chunks. Each chunk is modeled as an independent zero-mean Gaussian variable with the variance λ_{ki} . A linear coding is applied to the coefficients at the transmitter end:

$$X_k = g_k s_k \tag{1}$$

where g_k is the linear coding diagonal matrix. Its diagonal elements are power scaling factors. If the channel bandwidth is less than the source bandwidth, some chunks with the smallest variance will be discarded and their power scaling factors equal 0. Then the analog data is transmitted over an additive white Gaussian noise (AWGN) channel with variance σ_k^2 .

Given the known noise variance and the received signal Y_k , the receivers perform linear least squares estimation (LLSE). The λ_k is delivered to the receiver via a reliable way. So the estimated signal can be derived as follows

$$\hat{X}_k = \frac{g_k \lambda_k}{g_k^2 \lambda_k + \sigma_k^2} Y_k \tag{2}$$

Hence, we can calculate the expected distortion of user k

$$\mathbb{E}_k[m_k, g_k, \sigma_k] = \sum_{i=1}^{m_k} \frac{\lambda_{ki} \sigma_k^2}{g_{ki}^2 \lambda_{ki} + \sigma_k^2} + \sum_{i=m_k+1}^N \lambda_{ki} \tag{3}$$

or an uncoded water-filling result

$$\mathbb{E}_k[m_k, p_k, \sigma_k] = \frac{\left(\sum_{i=1}^{m_k} \sqrt{\lambda_{ki} \sigma_k^2}\right)^2}{p_k + \sum_{i=1}^{m_k} \sigma_k^2} + \sum_{i=m_k+1}^N \lambda_{ki} \tag{4}$$

where p_k denotes the allocated send power and m_k denotes the allocated bandwidth, so the distortion can be written as $\mathbb{E}(p, m, \sigma)$.

In multi-user scenario, the optimal transmission goal can be expressed as

$$\mathbb{E}[\mathbf{m}, \mathbf{p}, \boldsymbol{\sigma}] = f[\mathbb{E}_k[m_k, p_k, \sigma_k]] \tag{5}$$

where $\mathbf{m} = (m_1, \dots, m_k)$ is the bandwidth allocation vector of all users, the $\mathbf{p} = (p_1, \dots, p_k)$ is the power allocation vector of all users, and f is an utility function which generally has a convexity and its argument is the expected distortion of each user.

3.2 Problem statement

In single user scenario, the optimization objective is to minimize the mean distortion under the constraints on total power. Given the total transmitting power P , the power allocation problem is a convex problem, which can be solved by the Lagrange method. The optimal result can be written as

$$p_{ki} = g_{ki}^2 \lambda_{ki} = \frac{\sqrt{\lambda_{ki} \sigma_k^2} (P + \sum_{i=1}^{m_k} \sigma_k^2)}{\sum_{i=1}^{m_k} \sqrt{\lambda_{ki} \sigma_k^2}} - \sigma_k^2 \quad (6)$$

Since p_{ki} is greater than 0, the variance λ_{ki} should be greater than a threshold ϕ which depends on the noise variance

$$\lambda_{ki} > \phi = \left(\frac{\sum_{i=1}^{m_k} \sqrt{\lambda_{ki} \sigma_k^2}}{p_k + \sum_{i=1}^{m_k} \sigma_k^2} \right) \sigma_k^2 \quad (7)$$

However, in multi-user networks, the individual optimal result may not be fair to all users. The network utility optimization helps to improve the overall performance. Supposing that there are K users in networks with different channel conditions, we can know the rough channel condition from the channel estimator, and each channel noise's variance is σ_k^2 ($k = \{1, \dots, K\}$). The total available power is P , and the total available bandwidth is M . The problem can be formalized

$$\begin{aligned} \min & \sum_{k=1}^K f[\mathbb{E}_k[m_k, p_k, \sigma_k]] \\ \text{s.t.} & \sum_{k=1}^K m_k \leq M \\ & \sum_{k=1}^K p_k \leq P \\ & m_k \in \mathbb{Z} \end{aligned} \quad (8)$$

Note that \mathbb{E}_k is a multi-variable function, the main variable is the bandwidth and power, and f is the utility function to be discussed. The above optimization problem is a kind of Knapsack problem. It is a mixed integer nonlinear programming (MINLP) problem, which has been proven to be NP-hard. To reduce the complexity, we divide the optimization problem into two sub-problems, namely power allocation and bandwidth allocation.

The optimization objective Eq. 8 can be rewritten as

$$\min_{\mathbf{m}} \left\{ \min_p \{E[\mathbf{M}, \mathbf{p}, \boldsymbol{\sigma}] \mid \sum_{k=1}^K m_k = M\} \right\} \quad (9)$$

With this expression, it is clear that one sub-problem is to find the optimal power allocation under a given bandwidth allocation, and the other sub-problem is to find the optimal bandwidth allocation given that the optimal power allocation has been solved in the first sub-problem.

4 Joint resource allocation

Different optimization objectives should be considered in different scenarios while jointly allocating bandwidth and power in wireless multi-user networks. In the following subsections, we will discuss three optimization objectives: 1) Minimizing the total distortion; 2) Minimizing maximal distortion; 3) Minimizing the summation of square root distortion. The performances are briefly analysed for each objective.

4.1 Minimizing the total distortion

Each chunk is modeled as i.i.d Gaussian source. According to the uncoded transmission described in Section 2, the joint bandwidth and power allocation problem aiming at minimizing the total distortion for the network can be mathematically formulated as follows

$$\begin{aligned} \min \mathbb{E}_{total} &= \sum_{k=1}^K \sum_{i=1}^{m_k} \frac{\lambda_{ki} \sigma_{ki}^2}{p_{ki} + \sigma_k^2} + \sum_{k=1}^K \sum_{i=m_k+1}^N \lambda_{ki} \\ \text{s.t.} & \sum_{k=1}^K \sum_{i=1}^{m_k} p_{ki} \leq P \\ & \sum_{k=1}^K m_k \leq M \\ & m_k \in \mathbb{Z} \end{aligned} \quad (10)$$

m_k is the allocated bandwidth to user k , which must be integer. This optimal problem can be solved by solving two sub-problems.

Sub-problem 1) To solve the power allocation problem under a given bandwidth allocation, we relax the constraints on bandwidth. Assuming there are abundant bandwidth resources, each user can transmit all the coefficients. The power allocation problem can be written as

$$\begin{aligned} \min \mathbb{E}_{total} &= \sum_{k=1}^K \sum_{i=1}^{m_k} \frac{\lambda_{ki} \sigma_{ki}^2}{p_{ki} + \sigma_k^2} + \sum_{k=1}^K \sum_{i=m_k+1}^N \lambda_{ki} \\ \text{s.t.} & \sum_{k=1}^K \sum_{i=1}^{m_k} p_{ki} \leq P \\ & m_k = N \end{aligned} \quad (11)$$

Note that this problem is convex, which can be solved by the Lagrange method

$$p_{ki} = g_{ki}^2 \lambda_{ki} = \frac{\sqrt{\lambda_{ki} \sigma_k^2} (P + \sum_{k=1}^K \sum_{i=1}^{m_k} \sigma_k^2)}{\sum_{k=1}^K \sum_{i=1}^{m_k} \sqrt{\lambda_{ki} \sigma_k^2}} - \sigma_k^2 \quad (12)$$

The power allocation algorithm can be described as Algorithm 1. Its complexity is $O(N)$, because the main computation is calculating p_{ki} . With the constraint (12) described in Section 2, the bandwidth allocation \mathbf{m} should be updated based on the solved power here. That's sub-problem 2.

Algorithm 1 Power allocation algorithm

Input: λ, σ, N, K, P
Output: \mathbf{m}
 Initialization: $m_k = N, updateFlag = 1$
while $updateFlag = 1$ **do**
 $updateFlag = 0$
 $p_{ki} \leftarrow Poweralloc(\lambda, \sigma, N, P)$
 for $k=1$ **to** K **do**
 for $i=1$ **to** N **do**
 if $p_{ki} \leq 0$ **then**
 $m_k \leftarrow m_k - 1;$
 $updateFlag = 1;$
 end
 end
 end
 $\mathbf{m} = (m_1, m_2, \dots, m_K)$
end

Sub-problem 2) This sub-problem is trying to solve bandwidth allocation according to the power allocation and the initial bandwidth assumption in sub-problem 1. If total available bandwidth is more than the initial bandwidth assumption, the initial bandwidth is an optimal bandwidth allocation result. The sub-problem 2 has been solved. Otherwise, a greedy algorithm is proposed to solve the bandwidth allocation problem. To satisfy the bandwidth constraint, some coefficients have to be discarded, which causes distortion increment. According to the objective function Eq. 9, the distortion increment can be calculated as

$$Inc_{ki} = \lambda_{ki} - \frac{\lambda_{ki} \sigma_k^2}{p_{ki} + \sigma_k^2} \tag{13}$$

Thus we need to discard the chunk with least Inc_{ki} . Theoretically, power should be reallocated after discarding data. But from engineering perspective, it may be not necessary to reallocate power so frequently. Supposing the redundant power is v , according the Taylor series, the degradation without power reallocation can be written as

$$\mathbb{E}_{total}(p-v) - \mathbb{E}_{total}(p) = -\nabla \mathbb{E}_{total}(p)v + \frac{1}{2} \nabla^2 \mathbb{E}_{total}(p)v^2 \tag{14}$$

Making a tradeoff between efficiency and accuracy, we set a degradation threshold δ to determine whether reallocate power or not. According to Eqs. 9 and 14, the reallocation condition is

$$v_{re} \geq \frac{\nabla \mathbb{E}_{total}(p) + \sqrt{\nabla \mathbb{E}_{total}(p)^2 + 2 \nabla^2 \mathbb{E}_{total}(p) \delta}}{\nabla^2 \mathbb{E}_{total}(p)} \tag{15}$$

The bandwidth allocation algorithm can be described as Algorithm 2. Its complexity is $O(N^2)$.

Algorithm 2 Bandwidth allocation algorithm

Input: $\lambda, \sigma, K, P, M, p_{ki}, \mathbf{m}, \delta$
Output: \mathbf{m}, p_{ki}
 Initialization: $v = 0, Inc_{ki} \leftarrow Increment(\lambda, \sigma, p_{ki})$
while $\mathbf{m} \geq M$ **do**
 Discarding the least $Inc_{ki};$
 $m_k \leftarrow m_k - 1;$
 $\mathbf{m} = (m_1, m_2, \dots, m_K)$
 $v \leftarrow v + p_{ki};$
 if $v \geq v_{re}$ **then**
 $p_{ki} \leftarrow Poweralloc(\lambda, \sigma, \mathbf{m}, P);$
 end
end
 $\mathbf{m} = (m_1, m_2, \dots, m_K)$
 $p_{ki} \leftarrow Poweralloc(\lambda, \sigma, \mathbf{m}, P)$

Given the nature of the DCT coefficients distribution, the power is extremely concentrated on low frequency components. The condition can only be met when it need to drop most chunks. We can solve this problem by sorting the Inc_{ki} , and discarding the several least significant chunks to satisfy the bandwidth constraint directly. The simulation result shows that it causes almost no degradation in performance. The complexity of this step is $O(N \log N)$ mainly caused by sorting. So the bandwidth allocation without iteration can be described as Algorithm 3.

Algorithm 3 Bandwidth allocation without iteration algorithm

Input: $\lambda, \sigma, K, P, M, p_{ki}, \mathbf{m}, \delta$
Output: \mathbf{m}, p_{ki}
 Initialization: $v = 0, Inc \leftarrow Increment(\lambda, \sigma, p_{ki})$
 Sorting $Inc;$
 Discarding the least $\mathbf{m} - M$ $Inc;$
 Updating $m_k \mathbf{m} = (m_1, m_2, \dots, m_K)$
 $p_{ki} \leftarrow Poweralloc(\lambda, \sigma, \mathbf{m}, P)$

4.2 Minimizing the maximal distortion

In this subsection, we study minimizing the overall maximal distortion of all users under the given constraints, which is the min-max problem. Typically the min-max problem embodies fairness better. The optimization problem can be expressed as

$$\begin{aligned} \min \max & \mathbb{E}_k \\ \text{s.t.} & \sum_{k=1}^K \sum_{i=1}^{m_k} p_{ki} \leq P \\ & \sum_{k=1}^K m_k \leq M \\ & m_k \in \mathbb{Z} \end{aligned} \tag{16}$$

This problem can be solved by a greedy algorithm. The constraint in Eq. 6 indicating the least power p'_{ki} to transmit a chunk with variance λ_{ki} should satisfy

$$\begin{cases} \lambda_{ki+1} \leq \phi_{ki+1} \\ \lambda_{ki} > \phi_{ki} \end{cases} \quad (17)$$

Through equation $\lambda_{ki+1} = \phi_{ki+1}$, p_{ki} can be derived as

$$p_{ki} = \left(\sqrt{\lambda_{ki}\sigma_k^2} - \sqrt{\lambda_{ki+1}\sigma_k^2} \right) \frac{p_k + (i+1)\sigma_k^2}{\sum_{j=1}^{i+1} \sqrt{\lambda_{kj}\sigma_k^2}} + p_{ki+1} \quad (18)$$

The min-max power allocation algorithm can be written as Algorithm 4.

Algorithm 4 Power allocation algorithm

Input: $\lambda, \sigma, k, j_k, P$
Output: p
 solve p_{ki}
while $\lambda_{ki} \leq \phi_{ki}$ **do**
 $p_{k(i+1)} = p_{k(i+1)} + \varepsilon$
 solve p_{ki}
end
 $p = (p_1, p_2, \dots, p_K);$

The distortion of user k with the least power allocation is

$$D_{ki} = \left(\frac{\left(\sum_{i=1}^{m_k} \sqrt{\lambda_{ki}\sigma_k^2} \right)^2}{p_k + \sum_{i=1}^{m_k} \sigma_k^2} + \sum_{i=m_k+1}^N \lambda_{ki} \right) \quad (19)$$

Under the least power allocation to each chunk, we can gradually allocate the bandwidth to the users which have the maximal distortion. Once the bandwidth constraint is satisfied, it means that the optimal bandwidth allocation is finished and we continue to allocate the abundant power to achieve the minimal distortion. But if the power constraint is firstly satisfied, it means the abundant bandwidth can not use because of lacking power.

So the min-max bandwidth allocation algorithm can be written as

Algorithm 5 Bandwidth allocation algorithm

Input: $\lambda, \sigma, K, M, P,$
Output: m, p
 Initialization: $m = 1, \dots, 1, j_k = 1$
while $sum(p_k) \leq P$ **do**
 while $sum(m_k) \leq M$ **do**
 $Ref = \min D_{k,j_k}$
 for $k=1$ **to** K **do**
 while $D_{k,j_k} > Ref \ \&\& \ j \leq N$ **do**
 $j \leftarrow j + 1;$
 $p_{k,j_k} = \text{LeastPowerallocation}(\lambda_{k,j_k})$
 calculating $D_{k,j_k};$
 end
 end
 end
 $m = (m_1, m_2, \dots, m_K);$
 $p = (p_1, p_2, \dots, p_K);$
end

4.3 Minimizing the summation of square root distortion

In this subsection, we proposed using square root function as the proportional fairness function. Following the derivation in Eq. 8, the joint bandwidth and power allocation problem aiming at minimizing the square root utility function as follows

$$\begin{aligned} \min \mathbb{E}_{sqrt} &= \sum_{k=1}^K \sqrt{\mathbb{E}_k[m_k, p_k, \sigma_k]} \\ \text{s.t.} \quad &\sum_{k=1}^K \sum_{i=1}^{m_k} p_{ki} \leq P \\ &\sum_{k=1}^K m_k \leq M \\ &m_k \in \mathbb{Z} \end{aligned} \quad (20)$$

where m_k is allocated bandwidth of user k , which must be integer. p_{ki} is allocated power of chunk i in video k . As the analysis described in Section 2, this optimal problem can be solved by solving two sub-problem.

Sub-problem 1) The first sub-problem is to solve the power allocation problem under a given bandwidth allocation. Given bandwidth allocation m , the power allocation optimization objective can be expressed as follows

$$\begin{aligned} \min \mathbb{E}_{sqrt} &= \sum_{k=1}^K \sqrt{\left(\frac{\left(\sum_{i=1}^{m_k} \sqrt{\lambda_{ki}\sigma_k^2} \right)^2}{p_k + \sum_{i=1}^{m_k} \sigma_k^2} + \sum_{i=m_k+1}^N \lambda_{ki} \right)} \\ \text{s.t.} \quad &\sum_{k=1}^K p_k \leq P \\ &(m_1, m_2, \dots, m_K) = m \end{aligned} \quad (21)$$

Note that the objective function is convex and the constraint on power is linear. Therefore, the problem is convex. In general, constraint convex optimization problem can be solved by Lagrangian method

$$\begin{aligned} L(m, \gamma) &= \sum_{k=1}^K \sqrt{\left(\frac{\left(\sum_{i=1}^{m_k} \sqrt{\lambda_{ki}\sigma_k^2} \right)^2}{p_k + \sum_{i=1}^{m_k} \sigma_k^2} + \sum_{i=m_k+1}^N \lambda_{ki} \right)} \\ &\quad + \gamma \left(\sum_{k=1}^K p_k - P \right) \end{aligned} \quad (22)$$

For convenience, we define

$$\begin{aligned} A_k &= \left(\sum_{i=1}^{m_k} \sqrt{\lambda_{ki}\sigma_k^2} \right)^2 \\ B_k &= \sum_{i=1}^{m_k} \sigma_k^2 \\ C_k &= \sum_{i=m_k+1}^N \lambda_{ki} \end{aligned} \quad (23)$$

So the Lagrangian L can be rewritten as:

$$L(\mathbf{m}, \gamma) = \sum_{k=1}^K \sqrt{\frac{A_k}{p_k + B_k}} + C_k + \gamma \left(\sum_{k=1}^K p_k - \mathbf{P} \right) \tag{24}$$

For solving the optimal power allocation, we derive a equation set through the Lagrangian L :

$$\begin{cases} \gamma = -\frac{1}{2} \left(\frac{A_1}{p_1 + B_1} + C_1 \right)^{-\frac{1}{2}} \frac{A_1}{(p_1 + B_1)^2} \\ \dots \\ \gamma = -\frac{1}{2} \left(\frac{A_k}{p_k + B_k} + C_k \right)^{-\frac{1}{2}} \frac{A_k}{(p_k + B_k)^2} \\ \mathbf{P} = \sum_{k=1}^K p_k \end{cases} \tag{25}$$

However, we observe that the Eq. 25 are nonlinear equations with high dimensions and high orders. It is difficult to get the analytic expression, especially with the increase of user number K , which will bring the exponential increase of computation complexity. So we propose to use gradient descent method to solve this problem.

The power allocation optimization objective can be simplified as a constrained optimization problem with equality

$$\begin{aligned} \min \mathbb{E}_{sqr t} \\ \text{s.t. } \sum_{k=1}^K p_k = \mathbf{P} \end{aligned} \tag{26}$$

We can use Newton’s method to solve this problem. Similar to the power allocation problem in subsection A, we need to update bandwidth allocation and reallocate the power. Given a initial power allocation \mathbf{p}_0 , the Newton step $\Delta \mathbf{p}$ and decrement δ are

$$\begin{aligned} \Delta \mathbf{p} &= -\frac{\nabla \mathbb{E}_{sqr t}(\mathbf{p}_0)}{\nabla^2 \mathbb{E}_{sqr t}(\mathbf{p}_0)} \\ \delta &= (\Delta \mathbf{p}^T \nabla^2 \mathbb{E}_{sqr t}(\mathbf{p}_0) \Delta \mathbf{p})^{1/2} \end{aligned} \tag{27}$$

Given the nature of Newton method, convergence can be extremely rapid when the starting point is near the optimal solution. We can relax the bandwidth constraint, set $C_k = 0$, the Lagrangian equation set can be solved. So the starting point is near the optimal solution before bandwidth allocation. The Newton iteration becomes much more efficient. So the power allocation algorithm can be described as Algorithm 6.

Algorithm 6 Power allocation algorithm

```

Input:  $\lambda, \sigma, N, K, \mathbf{p}_0, \varepsilon, \delta$ 
Output:  $\mathbf{m}$ 
Initialization:
 $\mathbf{m} = N, \dots, N, \text{updateFlag} = 1, p_k = \mathbf{p}/N, t = 1$ 
while  $\text{updateFlag}$  do
     $\text{updateFlag} = 0$ 
    repeat
         $\Delta \mathbf{p} = -\frac{\nabla \mathbb{E}_{sqr t}(\mathbf{p}_0)}{\nabla^2 \mathbb{E}_{sqr t}(\mathbf{p}_0)}$ ;
         $\delta = (\Delta \mathbf{p}^T \nabla^2 \mathbb{E}_{sqr t}(\mathbf{p}_0) \Delta \mathbf{p})^{1/2}$ ;
         $\mathbf{p}_0 = \mathbf{p}_0 + t \Delta \mathbf{p}$ ;
    until  $\delta^2 \leq 2\varepsilon$ ;
     $(p_1, \dots, p_k) \leftarrow \mathbf{p}_0$ 
    for  $k=1$  to  $K$  do
        for  $i=1$  to  $N$  do
            if  $p_{ki} \leq 0$  then
                 $m_k \leftarrow m_k - 1$ ;
                 $\text{updateFlag} = 1$ ;
            end
        end
    end
     $\mathbf{m} = (m_1, m_2, \dots, m_K)$ 
end
    
```

Sub-problem 2) The second sub-problem is to solve the bandwidth allocation problem under a given power allocation. Similar to the bandwidth allocation in subsection A, the objective function increment can be calculated as

$$Inc_{ki} = \sqrt{\mathbb{E}_k[m_{k(i-1)}, p_k, \sigma_k]} - \sqrt{\mathbb{E}_k[m_{k(i)}, p_k, \sigma_k]} \tag{28}$$

where $\mathbb{E}_k[m_{k(i-1)}, p_k, \sigma_k]$ denotes the utility function result with discarding the (i) th chunk of user k , and $\mathbb{E}_k[m_{k(i)}, p_k, \sigma_k]$ denotes the utility function result with retaining the (i) th chunk in video k . In one video, Inc_{ki} is positive correlation with variance λ . In order to save the computation, we can just compare the variance of last chunk and discard the smallest. The bandwidth allocation algorithm can be described as Algorithm 7.

Algorithm 7 Bandwidth allocation algorithm

```

Input:  $\lambda, \sigma, K, \mathbf{P}, M, p_{ki}, \mathbf{m}, \delta, v_{re}$ 
Output:  $\mathbf{m}, p_{ki}$ 
Initialization:  $v = 0$ 
 $Inc_{k_{i,ast}} \leftarrow \text{Increment}(\lambda, \sigma, p_{ki})$ 
while  $\mathbf{m} \geq M$  do
    Discarding the least  $Inc_{k_{i,ast}}$ ;
     $m_k \leftarrow m_k - 1$ ;
     $\mathbf{m} = (m_1, m_2, \dots, m_K)$ 
     $v \leftarrow v + p_{ki}$ ;
    if  $v \geq v_{re}$  then
        repeat
             $\Delta \mathbf{p} = -\frac{\nabla \mathbb{E}_{sqr t}(\mathbf{p}_0)}{\nabla^2 \mathbb{E}_{sqr t}(\mathbf{p}_0)}$ ;
             $\delta = (\Delta \mathbf{p}^T \nabla^2 \mathbb{E}_{sqr t}(\mathbf{p}_0) \Delta \mathbf{p})^{1/2}$ ;
             $\mathbf{p}_0 = \mathbf{p}_0 + t \Delta \mathbf{p}$ ;
        until  $\delta^2 \leq 2\varepsilon$ ;
         $(p_1, \dots, p_k) \leftarrow \mathbf{p}_0$ 
    end
end
     $\mathbf{m} = (m_1, m_2, \dots, m_K)$ 
     $p_{ki} \leftarrow \text{Poweralloc}(\lambda, \sigma, \mathbf{m}, \mathbf{P})$ 
    
```


As the analysis in subsection A, and the power allocation need an iterative process. For efficiency purpose, we just need to reallocate the power one time after the bandwidth allocation finished. The simulation result shows that it cause little degradation in performance.

5 Simulation results

We carry out the evaluation based on multi-user videos, which are 720p with resolution 1280×720 . The frame rate of all video is 30 FPS (frame per second). The user number is assumed to 3 or 6. Similar to the test sequence used in Softcast, our test sequence is composed of multiple standard video test sequences including *City*, *ShuttleStart*, *Shields*, *Parkrun*, *Stockholm*, *Jets*, *Intotree*.

Since this paper focuses on uncoded video transmission, Softcast is considered as a reference system. In Softcast, we just assign all the available resources to all users in average. With limited bandwidth and power resources, every user discards the same number of chunks and allocated the same power. For a fair comparison, all systems use GOP size 8 and equal chunk division with 8×8 chunks per frame. The transmitting video content include 32 frames(4 GOPs). The available channel resources is assumed to be limited. The ratio of channel bandwidth to the source bandwidth is set to 1/2. The mean power also set as 1/2. For brief expression, we quantize the mean power and bandwidth to 1 per chunk, so the total available power resource is quantize by multiplying ‘chunk number’, ‘frame number of one GOP’, ‘power ratio’ and ‘user number’, as well as bandwidth. For instance, in Softcast average scheme, each user will allocate 256 units power resources and 256 units bandwidth resources.

For a comprehensive understanding of our scheme, we apply control variate method to discuss the results of proposed optimization strategies in three cases: 1) All users(3 users) require the same video content but with different channel condition. 2) All users(3 users) require different video content but with the same channel condition. 3) All users(6 users) with different video content and channel condition. In each case, we compare the bandwidth and power allocation, the performance in terms of PSNR(Peak Signal to Noise Ratio). In result figures, the *Aver* method means average all the available resources to all users; the *MinMSE* method represents minimizing the total distortion; the *MinSqrt* method represents minimizing the summation of square root distortion; the *Minmax* method represents minimizing the maximal distortion. In **Case 1** and **Case 2**, we concentrate on the analysis of the difference among proposed optimization objectives. In **Case 3**, we concentrate on the performance gain of the proposed optimization objectives.

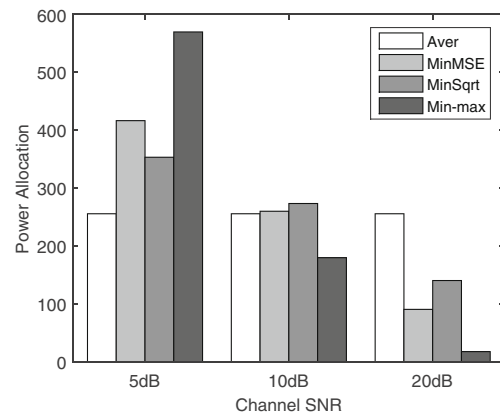


Fig. 3 Power allocation in Case 1

Case 1 All users(3 users) require the same video content but with different channel condition. This case shows the difference of resource allocation and how performance is affected by the channel noise.

1) Power allocation

As shown in Fig. 3, this experiment demonstrates the difference of power allocation among four methods. It shows that the users with low channel SNR have been allocated more power resources to some extent. In multi-user networks, to improve the overall performance, the users in bad channel condition need more protection. However, the protection of users in bad channel is at the cost of the loss of users in good channel condition. For instance, the *Minmax* method is the extremely fair protection, it has improved performance of users under bad channel condition but it also causes big loss to the users under good channel condition. The *MinMSE* method and *MinSqrt* method seem to balance better.

2) Bandwidth allocation

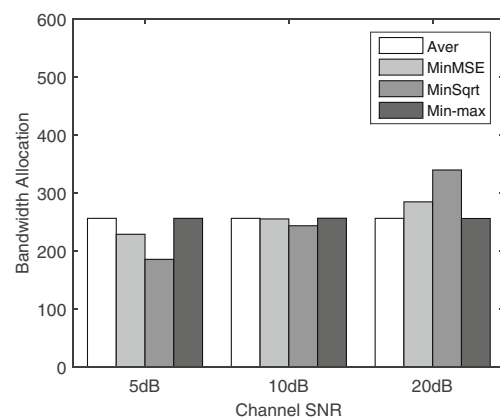


Fig. 4 Bandwidth allocation in Case 1

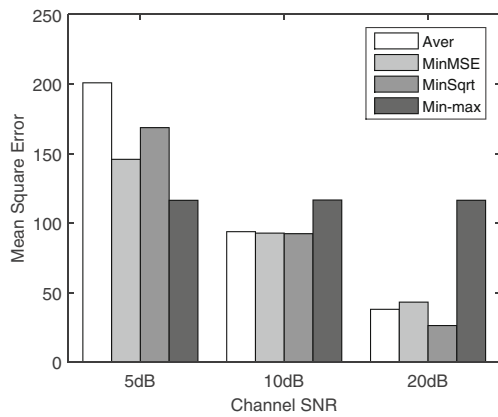


Fig. 5 Mean Square Error in Case 1

As shown in Fig. 4, this experiment demonstrates the difference of bandwidth allocation among four methods. Unlike the power allocation, the bandwidth allocation does not change too much compared to the *Aver* method, because the bandwidth allocation is mainly determined by the coefficient distribution. In case 1, the videos are the same and the bandwidth allocation does not vary like power allocation.

3) Mean Square Error

As shown in Fig. 5, this experiment demonstrates optimal results of all methods. The MSE has been quantized for a brief and clear compare. It shows that *Aver* method brings the MSE gap naturally, but *Minmax* method makes all users the same through allocating the resources. The *MinMSE* method just protect the high MSE users without much loss to other users. The result of *MinSqrt* method is confused, which the lowest MSE users still decreased, it proves that the optimization strategies suit different scenarios.

Case 2 All users(3 users) require different video content but with the same channel condition. This case shows the

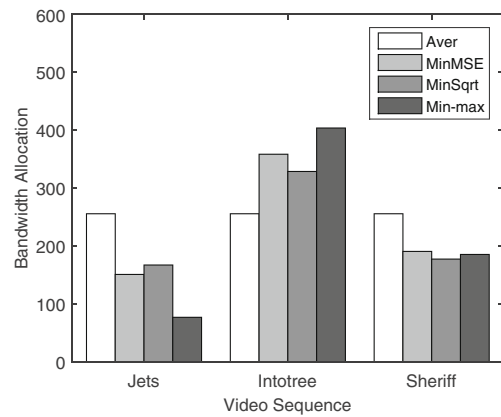


Fig. 7 Bandwidth allocation in Case 2

difference of resource allocation and performance effected by the video content.

1) Power allocation

As shown in Fig. 6, this experiment demonstrates the difference of power allocation caused by the video content. The DCT coefficients of three sequences have different variances. The bigger variance is, the more power is allocated. In frequency domain, the low frequency coefficients have higher energy, so it needs more power to transmit them. Thus the video with high energy in frequency domain tend to be allocated more power resources.

2) Bandwidth allocation

As shown in Fig. 7, the bandwidth allocation is almost the same with power allocation. Under the same channel condition, the distribution of DCT coefficients determines the resource allocation.

3) Mean Square Error

As shown in Fig. 8, just like case 1, the *Aver* method brings the gap and the optimal methods decrease the gap.

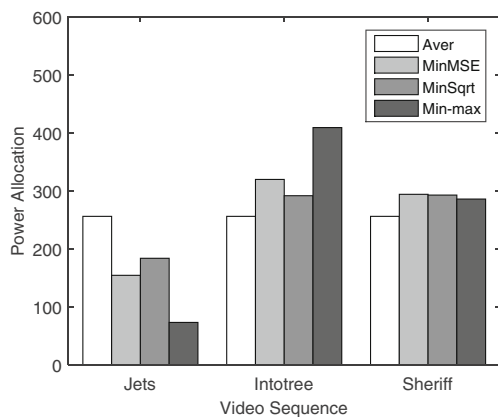


Fig. 6 Power allocation in case 2

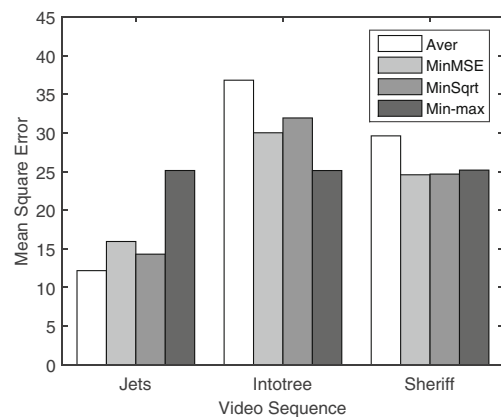


Fig. 8 Mean Square Error in Case 2

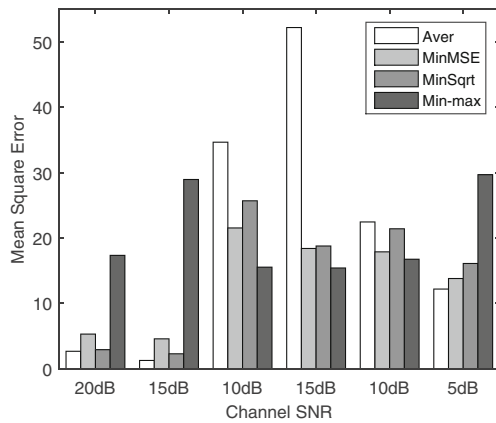


Fig. 9 Mean square error in Case 3

However, the *MinSqrt* method performs better in this case, meanwhile it causes fewer loss to sequence *Jets* but improve other users performance.

Case 3 All users(6 users) with different video content and channel condition. The six sequence are *City*, *ShuttleStart*, *Shields*, *Parkrun*, *Stockholm*, *Jets*. In this case, we concern about the MSE and the video quality in terms of PSNR.

1) Mean Square Error

As shown in Fig. 9, the *MinMSE* method decrease 16.43dB in terms of total MSE. The *MinSqrt* method decreases 15.83dB in terms of total MSE. But *MinMax* method emphasizes the fairness of all users and it only decreases 2.42dB in terms of total MSE.

2) Video Quality

As shown in Fig. 10, the *MinMSE* method and *MinSqrt* method decreases the gap among all users, and *Min-*

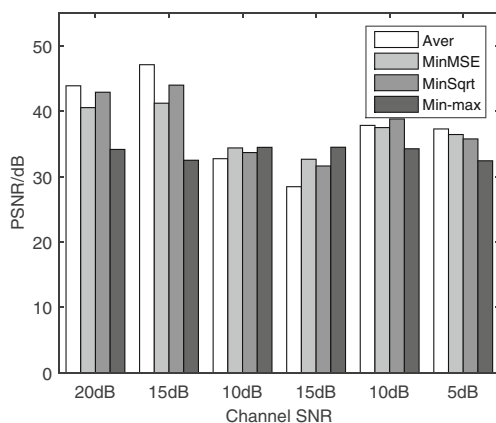


Fig. 10 Video quality in Case 3

Sqrt method seems more reasonable. The *MinMax* method emphasizes the fairness of all users but it causes a performance loss in terms of total PSNR.

6 Conclusion

In this paper, we present an analog multi-user video transmission framework over wireless networks and apply three different optimization strategies to solve the resource allocation problems. The optimization strategies evaluate the overall network performance from different perspective. The optimization to minimize the total distortion performs the best in terms of average PSNR of all users; the optimization to minimize the maximal distortion achieves more fair result, but sometimes it may be unreasonable because it has reduced the performance of users with the best channel condition; The optimization to minimize the summation of square root distortion seems to be a better choice, which can improve the performance of users with bad channel conditions, meanwhile causing little loss of good condition users.

Acknowledgments This work was supported in part by National Natural Science Foundation of China (Nos. 61390513, 61571329 and 61502341), National Science and Technology Major Project of China (Nos. 2014ZX03003003).

References

1. Cisco (2013) Cisco visual networking index: global mobile data traffic forecast update, 2012–2017. Tech. rep. http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-520862.html
2. Jakubczak S, Katabi D (2011) A cross-layer design for scalable mobile video. In: Proceedings of ACM Mobicom, pp 289–300
3. Lee KH, Petersen DP (1976) Optimal linear coding for vector channels. IEEE Trans Commun 24(12):1283–1290
4. Fan X, Xiong R, Wu F, Zhao D (2012) Wavecast: wavelet based wireless video broadcast using lossy transmission. In: Proceedings of IEEE visual communications and image processing (VCIP), pp 1–6
5. Cui H, Song Z, Yang Z, Luo C, Xiong R, Wu F (2013) Cactus: a hybrid digital-analog wireless video communication system. In: Proceedings of ACM MSWiM, pp 273–278
6. Cui H, Luo C, Chen CW, Wu F (2014) Robust uncoded video transmission over wireless fast fading channel. In: Proceedings of IEEE INFOCOM, pp 73–81
7. Su GM, Han Z, Wu M, Liu KR (2006) A scalable multiuser framework for video over OFDM networks: fairness and efficiency. IEEE Trans Circuits Syst Video Technol 16(10):1217–1231
8. Su GM, Han Z, Wu M, Liu KR (2007) Joint uplink and downlink optimization for real-time multiuser video streaming over WLANs. IEEE J Sel Top Sign Proces 1(2):280–294

9. Huang J, Li Z, Chiang M, Katsaggelos AK (2008) Joint source adaptation and resource allocation for multi-user wireless video streaming. *IEEE Trans Circuits Syst Video Technol* 18(5):582–595
10. Maani E, Pahalawatta PV, Berry R, Pappas TN, Katsaggelos AK (2008) Resource allocation for downlink multiuser video transmission over wireless lossy networks. *IEEE Trans Image Process* 17(9):1663–1671
11. Joseph V, de Veciana G (2012) Jointly optimizing multi-user rate adaptation for video transport over wireless systems: mean-fairness-variability tradeoffs. In: *Proceedings of IEEE INFOCOM*, pp 567–575
12. Fu F, van der Schaar M (2010) A systematic framework for dynamically optimizing multi-user wireless video transmission. *IEEE J Sel Areas Commun* 28(3):308–320