CROSS-ENTROPY BASED ANTENNA SELECTION FOR SCALABLE VIDEO STREAMING OVER MIMO WIRELESS NETWORKS

Chao Zhou, Xinggong Zhang , Zongming Guo*

Institute of Computer Science & Technology, Peking University, Beijing, 100871, P.R. China
Email: {zhouchaoyf, zhangxg, guozongming}@pku.edu.cn

ABSTRACT

In this paper, we investigate the antenna selection (AS) problem for scalable video streaming over MIMO wireless networks. By scheduling scalable video layers over MIMO antennas with different signal strength, the video layers are transmitted with un-equal error protections. Considering layer dependencies and various antenna conditions, it is a non-linear combinatorial problem for AS to minimize the overall end-to-end distortion. To find the optimal solution with low complexity, a cross-entropy based solution, named CEBAS, is proposed. All solutions are indexed by unique binary strings, and the primal problem is reformulated to a binary combination problem. Then, random strings are generated using the probability distribution of solutions, which is updated by the cross-entropy optimization method. The feasibility of solution is guaranteed by our proposed projection strategy. CEBAS is iterative in nature and converges to the global optimum in probability. Simulation results reveal both the effectiveness and efficiency of our proposed algorithm. When comparing CEBAS against other existing algorithms, consistent superior performance has been observed.

Index Terms— AS, MIMO, UEP, Video streaming

1. INTRODUCTION

Recent years have seen a rapid growth of video communications over wireless networks. However, the potentially low bandwidth and high bit error rates become significant obstacles for high quality multimedia communications. To overcome such obstacles, multi-input multi-output (MIMO) system has recently emerged as one of the most prominent techniques. Video delivery over PT-MIMO [1] to provide high bit-rate video content could be found in [2]. However, since video signal contains bits with different importances, it is able to provide more reliable transmission in MIMO. The use of scalable video coding (SVC) [3] is a basis for a number of schemes, where different treatment of layers that corresponds to their importance results in improvements of the transmitted video quality. Thus, providing different levels of error protection (UEP) [4–6] to the video layers is very crucial for scalable video streaming.

One of the most classical research work is ACS-MIMO [6], in which the video layers are periodically switched among multiple antennas according to antenna's signal to noise ratio (SNR) strength. In that fashion, implicit UEP is automatically achieved by antenna selection (AS) for video streaming over MIMO systems. The reconstructed video quality significantly outperforms PT-MIMO [1]. Based on this work, many other UEP schemes [7, 8] have been proposed. Nevertheless, all of these works assume that the channel bandwidth and video bit-rates are consistent, and also do not consider channel coding in transmission. In practical implementations, they may suffer heavy quality deterioration since *i*) channel bandwidth is generally time-varying and unpredictable in MIMO, and *ii*) the bit-rate of video layers are diverse because video content has different motion and spatial detail.

In this paper, we reinvestigate the AS problem for scalable video streaming over MIMO systems. Then an AS based UEP scheme is proposed with the objective to minimize the end-to-end video transmission distortion. In this scheme, antennas' bandwidth, SNR strength, and video layers' bit-rate are jointly considered, and UEP is achieved by mapping the video layers to the appropriate MIMO antennas. We formulate this scheme into a non-linear combinatorial optimization problem. It can be solved by an exhaustive search method, but the computational complexity is prohibitive.

In order to reduce the complexity, we investigate the cross-entropy optimization method, which was first proposed by Rubinstein [9]. Based on the cross-entropy method, a low-complexity near-optimal AS algorithm, named CEBAS, is devised. In CEBAS, each feasible solution is denoted by an unique binary string, and the primal problem is reformulated to a binary combination problem, whose objective is to find the optimal binary string. Then, random strings are generated according to the stochastic properties of optimal solutions, and the probability distribution of the solutions is iteratively updated by the cross-entropy optimization method. The string generated by the global optimal probability distribution corresponds to the optimal solution for the primal problem. Furthermore, a projection strategy is proposed to guarantee the binary string's feasibility. The algorithm is iterative in nature and converges to the global optimum in probability. Simulation results reveal both the effectiveness and efficiency of CEBAS. The quality of reconstructed video demonstrates significant improvement in CEBAS as compared with existing schemes.

*Corresponding author. E-mail:zhangxg@pku.edu.cn

This work was supported by National Natural Science Foundation of China under contract No.60902004, National Basic Research Program of China under contract No.2009CB320907 and Doctoral Fund of Ministry of Education of China under contract No. 20090001120027.

The paper is organized as follows: Section 2 formulates the proposed scheme and a low-complexity near-optimal antenna-selection algorithm is presented in Section 3. Finally, we present illustrative simulation results in Section 4 and conclude the paper in Section 5.

2. PROBLEM FORMULATION

Consider a video streaming link from source terminal equipped with N_t transmit antennas to the destination terminal equipped with N_r receive antennas, and $N_t \leq N_r$ for guaranteeing the Degrees-of-Freedom (DOF) constraint. Assume that channel matrix is known perfectly at the receiver[10], the approximated BER with M-QAM using gray coding of i-th($\forall 1 \leq i \leq N_t$) transmit antenna is derived as[11]:

$$\rho_i \approx \frac{\sqrt{M-1}}{\sqrt{M}\log_2\sqrt{M}}\mathrm{erfc}\left(\sqrt{\frac{3\gamma_i}{2(M-1)}}\right) \quad (1)$$

where $\mathrm{erfc}(\cdot)$ is the complementary error function, γ_i is the post-processing SNR of i-th transmit antenna. We further assume the bandwidth vector $\mathbf{R} = [R_1, R_2, ..., R_{N_t}]$ is known perfectly at the transmitter, where R_i is i-th transmit antenna's bandwidth.

The input video sequence is also encoded into N_t layers (scalability in SNR quality dimension) [3] with bit-rate vector $\mathbf{r} = [r_1, r_2, ..., r_{N_t}]$, where r_i denotes i-th video layer's bit-rate, and assume that $R_i \geq r_j, \forall 1 \leq i, j \leq N_t$. All these assumptions can be easily met by adjusting the encoder parameters (*layer-number, QP, etc.*) in the config file. We employ the distortion model proposed in [5], and write the end-to-end video transmission distortion as

$$\mathcal{D} = \sum_{i=1}^{N_t} w_i(1 - \prod_{k=1}^{i}(1 - P_k)) \quad (2)$$

where w_i is the weight of i-th video layer, denoting its priority, P_k denotes k-th video layer's PER.

We define $\mathbf{A}_{N_t \times N_t}$ as AS matrix to denote the mapping relationship between video layers and transmit antenna elements, and its entries $a_{ij} \in \{0,1\}, \forall 1 \leq i, j \leq N_t$. Let $a_{ij} = 1$ if i-th video layer is transmitted over j-th transmit antenna, and $a_{ij} = 0$ otherwise. Then we have:

$$P_i = [\mathbf{A}_{N_t \times N_t}]_i \times \left[\mathbf{P}^c_{N_t \times N_t}\right]_i^{\mathrm{T}}, \ \forall 1 \leq i \leq N_t \quad (3)$$

where $[\cdot]_i$ denotes i-th row of the matrix, $[\cdot]^{\mathrm{T}}$ denotes the transpose operation. $\mathbf{P}^c_{N_t \times N_t}$ is the FEC coding error matrix, and its entries P^c_{ij} denotes the error probability of FEC when i-th video layer transmitted over j-th transmit antenna. In order to protect the video layers against channel error, linear FEC coding is adopted and the error probability is

$$P^c_{ij} = 1 - \sum_{k=0}^{m_{ij}} \binom{N}{k} \rho_j^k (1 - \rho_j)^{(N-k)} \quad (4)$$

where m_{ij} is the number of partial bits in each packet, and N is the packet size. Noting that P^c_{ij} is a monotonic decreasing function of m_{ij}, we set $m_{ij} = N \times (1 - \frac{r_i}{R_j})$, which is the maximal number of partial bits under bandwidth constraints.

Equ.(4) shows that matrix $\mathbf{P}^c$ is affected by not only SNR strength, but also antenna's bandwidth and video layer's bit rates. Therefore, when we implement AS in Equ.(3), all these factors should be considered, rather than simply according to SNR strength [6]. At last, the proposed antenna-video-layer mapping problem can be formulated as

$$\mathbf{A}^* = \arg\min_{\{\mathbf{A}\}} \sum_{i=1}^{N_t} w_i(1 - \prod_{k=1}^{i}(1 - P_k(\mathbf{A}, \mathbf{R}, \mathbf{r})))$$

s.t.

$$\begin{aligned} & a_{ij} \in \{0,1\}, \ \forall 1 \leq i, j \leq N_t \\ & \sum_{i=1}^{N_t} a_{ij} = 1, \ \forall 1 \leq j \leq N_t \\ & \sum_{j=1}^{N_t} a_{ij} = 1, \ \forall 1 \leq i \leq N_t \end{aligned} \quad (5)$$

The constraints denote that this is a one-to-one mapping problem, and the number of video quality layers is equal to the number of antennas. The case that they are not equal is beyond the scope of this paper, and we will investigate it in our future work. Besides, this is a non-linear combinational optimization problem due to video layer's inter-dependence. The most straightforward method to determine the optimal AS matrix $\mathbf{A}^*$ is to perform an exhaustive search over the whole solution space, whose size is $N_t!$. However, the required computational complexity is prohibitive.

3. CROSS ENTROPY BASED ANTENNA SELECTION ALGORITHM

In this section, a low-complexity near-optimal cross-entropy based antenna selection algorithm is devised. First, each solution in the solution space is indexed with an unique identifier, i.e. a binary string. Then we reformulate the original problem to a binary combination problem. Through the cross-entropy optimization method, we continually update the probability distribution of feasible solutions to approach the global optimal probability distribution. The string generated by the global optimal probability distribution corresponds to a optimal solution for the primal problem. Moreover, a projection strategy is proposed to guarantee the feasibility of strings.

3.1. Problem Reformulation

Since the whole solution space is $N_t!$, we can use a certain binary string $\mathbf{u}$ with length $\mathcal{K} = \lceil \log(N_t!) \rceil$ as an unique identifier for each solution $\mathbf{A}$. We denote $\mathcal{U}$ as the set of feasible $\mathbf{u}$, whose feasibility is guaranteed by $\mathrm{DEC}(\mathbf{u}) \leq N_t!$, where $\mathrm{DEC}(\cdot)$ denotes binary-to-decimal conversion. Also, each $\mathbf{u}$ corresponds to a streaming distortion $\mathcal{I}(\mathbf{u})$ as Equ. 2.

Now, the primal problem in (5) is reformulated as

$$\mathbf{u}^* = \arg\max_{\mathbf{u} \in \mathcal{U}} \mathcal{I}(\mathbf{u}) \quad (6)$$

where $\mathcal{I}(\mathbf{u}) = -\sum_{i=1}^{N_t} w_i(1 - \prod_{k=1}^{i}(1 - P_k(\mathbf{u}, \mathbf{R}, \mathbf{r})))$.

3.2. Cross-entropy Iterations

We define the indicator function $h_+(\mathcal{X})$ which returns one if event $\mathcal{X}$ is true, and zero otherwise. Let u_i denote the i-th element of binary string $\mathbf{u}$, and p_i is the probability of $u_i = 1$, where $i = 1, 2, ..., \mathcal{K}$. Then, given $\mathbf{p}$, the probability of the string $\mathbf{u}$ is:

$$f(\mathbf{u}, \mathbf{p}) = \prod_{i=1}^{\mathcal{K}} p_i^{h_+(u_i=1)} (1 - p_i)^{1-h_+(u_i=1)} \tag{7}$$

where $\mathbf{p}$ is the set of p_i.

We further define a collection of outputs $\{h_+(\mathcal{I}(\mathbf{u}) \geq d)\}$ in the space $\mathbf{u} \in \mathcal{U}$ for various d. From [12], the cross entropy of random selected string set $\{\mathbf{u}[n]\}$ is defined as:

$$\hat{\mathcal{Q}}(\mathbf{p}) = -\frac{1}{N_u} \sum_{n=1}^{N_u} h_+(\mathcal{I}(\mathbf{u}[n]) \geq d) \ln f(\mathbf{u}[n], \mathbf{p}) \tag{8}$$

where $\mathbf{u}[n] (1 \leq \forall n \leq N_u)$ is the n-th random string generated by the previous $\mathbf{p}$ (updated in the last iteration), N_u is the number of generated strings.

Thus, in each iteration, we generate some new strings from the previous probability distribution of solutions, and obtain a new distribution $f(\cdot, \mathbf{p})$ by the cross-entropy optimization method. The distribution function $f(\cdot, \mathbf{p})$ is iteratively updated to approach the global optimal distribution $f(\cdot, \mathbf{p}^*)$. As discussed in [13], the probability updating rule is obtained by minimizing the Kullback-Leibler divergence. This is equivalent to solving

$$\hat{\mathbf{p}}^* = \arg\min_{\mathbf{p}} \hat{\mathcal{Q}}(\mathbf{p}) \tag{9}$$

and $\hat{\mathcal{Q}}(\mathbf{p})$ is concave. In order to minimize $\hat{\mathcal{Q}}(\mathbf{p})$, we set $\frac{\partial \hat{\mathcal{Q}}(\mathbf{p})}{\partial \mathbf{p}} = 0$, leading to the updating rule as

$$p_i = \frac{\sum_{n=1}^{N_u} h_+(\mathcal{I}(\mathbf{u}[n]) \geq d) h_+(u_i[n] = 1)}{\sum_{n=1}^{N_u} h_+(\mathcal{I}(\mathbf{u}[n]) \geq d)} \tag{10}$$

This updating rule is iteratively executed with an aim to generate a sequence of increasing thresholds $d[0] \leq d[1] \leq$ ... (where $d[t]$ denotes the threshold at the t-th iteration), until the convergence to d^* is reached.

3.3. Projection Strategy

It is important to note that the samples generated based on Equ.(10) cannot guarantee that the samples are feasible, or $\mathrm{DEC}(\mathbf{u}) \leq N_t!$. A projection strategy is proposed to convert the infeasible samples into feasible ones. Given a random string $\mathbf{u}$, which consists of ones and zeros following the Bernoulli p.d.f. by the parameter vector in Equ.(10), the following projection strategy is applied:

Case 1) If $\mathbf{u} \in \mathcal{U}$, then no projection is needed.

Case 2) If $\mathbf{u} \notin \mathcal{U}$, i.e., $\mathrm{DEC}(\mathbf{u}) > N_t!$. We define $\mathcal{Z} = \{i \,|u_i = 1, 1 \leq i \leq \mathcal{K}\}$. Then, the feasibility is ensured by iteratively seting $u_k = 0$ and updating $\mathcal{Z} = \mathcal{Z} - \{k\}$, where $k = \arg\min_{i \in \mathcal{Z}} p_i$. The iteration ends if $\mathrm{DEC}(\mathbf{u}) \leq N_t!$.

3.4. Solution Algorithm

To end this section, we present the pseudo-code of CEBAS algorithm as in Algorithm 1:

Algorithm 1 Cross Entropy Based Antenna Selection Algorithm

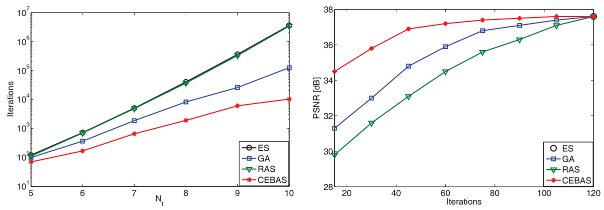
1: initialize iteration counter $t = 1$, probability vector $\mathbf{p}[0] = \{p_i^0\}_{i=1}^{N_u} = \{0.5\}_{i=1}^{N_u}$, $d[0] = -\infty$;
2: generate N_u random binary strings, $\{\mathbf{u}[n, t]\}_{n=1}^{N_u}$, by exploiting the p.d.f. $f(\cdot, \mathbf{p}[t-1])$;
3: **for** each $n \in [1, N_u]$ **do**
4:     **if** $\mathbf{u}[n, t] \notin \mathcal{U}$ **then**
5:         modify $\mathbf{u}[n, t]$ until that $\mathbf{u}[n, t] \in \mathcal{U}$ by the projection strategy in Sec. 3.3;
6:     **end if**
7:     compute the function $\mathcal{I}(\mathbf{u}[n, t])$;
8: **end for**
9: $d[t] = \max\left(\eta \max\left(\{\mathcal{I}(\mathbf{u}[n, t])\}_{n=1}^{N_u}\right), d[t-1]\right)$;
10: calculate the vector $\mathbf{p}[t]$ according to Equ.(10);
11: $\mathbf{p}[t] := \xi \mathbf{p}[t] + (1 - \xi) \mathbf{p}[t-1]$;
12: **if** a convergence criterion is satisfied **then**
13:     stop;
14: **else**
15:     $t = t + 1$, go to 2;
16: **end if**

In the pseudo-code, line 9 increases the threshold $d[t]$ iteratively until the convergence to d^* is reached, and $\eta = 0.9 \sim 1$ typically. In line 11, a smoothing process is adopted to avoid the local optimum, and $0 < \xi < 1$ is the smoothing factor. The convergence criterion in line 12 may be a maximum number of iterations, after which $d[t]$ has not made noticeable improvement. The convergence proof of this algorithm can refer to the literature [12].

4. SIMULATION RESULTS

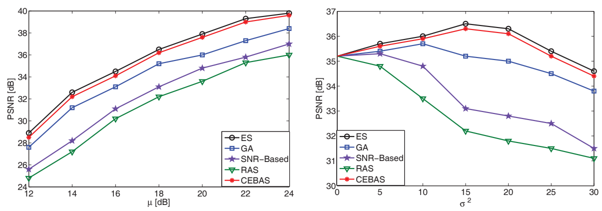
In order to evaluate the performance of CEBAS, we conduct extensive experiments by simulations. Some existing AS methods, including exhaustive search (ES), conventional genetic algorithm (GA), SNR-Based AS [6], and random antenna-selection (RAS), are implemented for comparisons. We consider a MIMO channel which is characterized with the average SNR of antennas is proportional to a truncated normal distribution, i.e., $\gamma \sim N_t(\mu, \sigma^2)$ and bounded in $[\gamma_{\min}, \gamma_{\max}]$. The test video sequence, *akiyo*, is encoded into N_t layers using the JSVM reference software Ver9.12 [3], and transmitted over a MIMO system.

We first compare the complexity of the four schemes under different antenna number (video layers) N_t. The convergence criterion for all algorithms is when the quality loss of reconstructed video is less than 0.1dB compared with ES. The results is illustrated in Fig.1(a). It shows that for various number of antennas, CEBAS always needs the least number of iterations. We also plot the curves of quality increased with the number of iterations. For the scenario with 5 antennas, Fig.1(b) shows that CEBAS always provides a better performance than others under the same complexity (iterations). In other words, CEBAS is able to achieved the same video quality with lower complexity. The figure also shows that when



(a) Various antenna numbers. (b) Various iterations.

Fig. 1. Complexity comparisons under the fixed channel conditions $\mu = 18$dB, $\sigma^2 = 15, \gamma_{\min} = 0$dB, $\gamma_{\max} = 35$dB.



(a) Varying average SNR (b) Varying channel variance

Fig. 2. PSNR under different channel conditions with $N_t = 5, \gamma_{\min} = 0$dB, $\gamma_{\max} = 35$dB.

the iterations is greater than 40, 80, and 110 for CEBAS, GA and RAS respectively, the performance improvements are negligible. Thus, we choose the iterations of 30 for CEBAS, and 60 for both GA and RAS in the following simulations. This also demonstrates the efficiency of CEBAS.

We also validate the algorithms under various network conditions. By varying the average SNR μ and variance σ, the quality of reconstructed video is shown in Fig.2. ES method always yields the optimum performance with the highest complexity (120 iterations) as expected. CEBAS has a very close performance with ES but in a much lower complexity. Moreover, it outperforms GA, RAS, and SNR-Based methods for all channel conditions. We also find that PSNR is increasing as μ increases as Fig.2(a) shows, while the monotonicity is lost in Fig.2(b) as σ^2 increases. This is mainly because that when the variance of antennas SNR is very small, all the antennas nearly have the same quality and UEP is hard to be achieved by AS. On the other hand, when the variance is large, SNR-Based AS and RAS method do not fully exploit the diverse of antennas, and their performances are deteriorated monotonously.

5. CONCLUSION

In this paper, we have investigated the AS problem for scalable video transmission over MIMO systems. Unlike previous works, antenna's bandwidth, SNR strength, and video bit rates are jointly considered in AS, and UEP is achieved. We formulate the AS problem into a non-linear combinatorial optimization problem with the object to minimize the end-to-end video transmission distortion. Furthermore, a low-complexity cross-entropy based AS algorithm is devised. The simulation results reveal both the effectiveness and efficiency of CEBAS.

References

[1] H. Zheng and K.J.R. Liu, "Space-Time Diversity for Multimedia Delivery over Wireless Channels," *IEEE Int. Symp. Circuits and Syst.*, vol. 4, pp. 285–288, May. 2000.

[2] Chih-Hung Kuo, Chang-Su Kim, and C.-C.J. Kuo, "Robust Video Transmission over Wideband Wireless Channel Using Space-Time Coded OFDM Systems," *IEEE Wirel. Commun. and Networking Conf.*, vol. 2, pp. 931–936, Mar. 2002.

[3] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the Scalable Video Coding Extension of the H.264/AVC Standard," *IEEE Trans. Circuits. Syst. Video Technol.*, vol. 17, no. 9, pp. 1103 –1120, Sep. 2007.

[4] M. K. Jubran, M. Bansal, and L. P. Kondi, "Low-Delay Low-Complexity Bandwidth-Constrained Wireless Video Transmission Using SVC Over MIMO Systems," *IEEE Trans. Multimedia*, vol. 10, no. 8, pp. 1698 –1707, Dec. 2008.

[5] Y. P. Fallah, H. Mansour, S. Khan, P. Nasiopoulos, and H. M. Alnuweiri, "A Link Adaptation Scheme for Efficient Transmission of H.264 Scalable Video Over Multirate WLANs," *IEEE Trans. Circuits. Syst. Video Technol.*, vol. 18, no. 7, pp. 875 –887, Jul. 2008.

[6] D. Song and Chang Wen Chen, "Scalable H.264/AVC Video Transmission Over MIMO Wireless Systems With Adaptive Channel Selection Based on Partial Channel Information," *IEEE Trans. Circuits. Syst. Video Technol.*, vol. 17, no. 9, pp. 1218 –1226, Sep. 2007.

[7] Shengjie Zhao, Yi Zhang, and Luoning Gui, "Optimal Resource Allocation for Video Delivery over MIMO OFDM Wireless Systems," *IEEE Global Telecommun. Conf.*, pp. 1–5, Dec. 2010.

[8] J. Abot, M. Nauge, C. Perrine, C. Larabi, C. Bergeron, Y. Pousset, and C. Olivier, "A Roubst Content-Based JPWL Transmission over A Realistic MIMO Channel under Perceptual Constraints," *IEEE Int. Conf. Image Process.*, pp. 3302–3305, Sep. 2011.

[9] R. Y. Rubinstein, "Optimization of Computer Simulation Models with Rare Events," *Eur. J. Oper. Res.*, vol. 99, no. 1, pp. 89–112, May 1997.

[10] B. Hassibi and B.M. Hochwald, "How Much Training is Needed in Multiple-Antenna Wireless Links?," *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 951–963, Apr. 2003.

[11] Bernard Sklar, *Digital communications - fundamentals and applications*, Prentice Hall, NJ, 2001.

[12] Reuven Y. Rubinstein and Dirk P. Kroese, "The Cross-Entropy Method: A Unified Approach to Combinatorial Optimization, Monte-Carlo Simulation, and Machine Learning," *Springer Verlag*, 2004.

[13] S. Kullback and R. A. Leibler, "On Information and Sufficiency," *Annals Mathematical Statistics*, vol. 22, pp. 79–86, 1951.