# UNEQUAL ERROR PROTECTION FOR REAL-TIME VIDEO STREAMING USING EXPANDING WINDOW REED-SOLOMON CODE

*Yufeng Geng*[1], *Xinggong Zhang*[1,3,*], *Chao Zhou*[2], *Zongming Guo*[1,3]

[1]Institute of Computer Science & Technology, Peking University, Beijing, 100871, P.R. China
[2]Media Technology Lab, CRI, Huawei Technologies CO., LTD, Beijing, China
[3]Cooperative Medianet Innovation Center, China
Email: {gengyufeng, zhangxg, guozongming}@pku.edu.cn, zhouchaoyf@gmail.com

## ABSTRACT

Expanding Window FEC is an emerging scheme for robust real-time video streaming over wireless networks, with low latency and reduction of error propagation. In this work, we focus on the problem of Expanding Window FEC redundancy allocation which has not been adequately addressed in current works. We first analyse the error probability of the adopted Expanding Window Reed-Solomon code (EW-RS), and introduce an equivalent error probability to simplify them. Then we are able to formulate the optimal redundancy allocation into a constrained nonlinear optimization problem, where by allocating the redundancy unequally considering the unequal importance of different frames and their dependency based on the expanding window, unequal error protection (UEP) is achieved and the overall distortion is minimized. Moreover, to reduce the computation complexity, a high-efficiency hill-climbing algorithm is developed to obtain the suboptimal allocation. At last, the experimental results demonstrate the effectiveness of both the proposed allocation scheme and solution algorithm.

***Index Terms***— UEP, Distortion-Optimized, Redundancy Allocation

## 1. INTRODUCTION

Recent years there have seen a rapid growth of real-time video communications over wireless networks. However, the potentially low bandwidth and high packet error rates become significant obstacles for high-quality multimedia communications. To overcome such obstacles, ways of robust video transmission have been explored and Forward Error Correction (FEC) is one of the most widely used schemes. On the other hand, since different parts of video content have unequal importance, allocating FEC unequally can achieve unequal error protection so as to improve video transmission quality[1],[2],[3],[4]. However, the performance of most existing UEP schemes relies on sufficiently large FEC coding block size, which introduce considerable latency as the need to wait for the whole block to be decoded, thus fail to serve real-time video streaming[5].

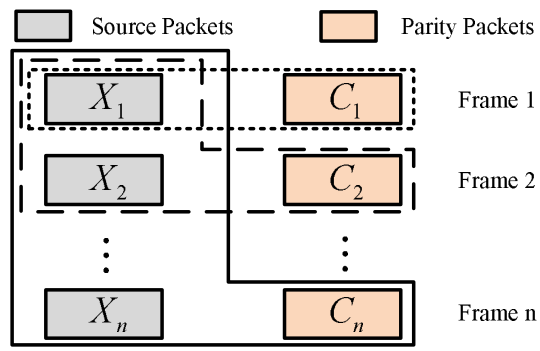


**Fig. 1**: Expanding Window FEC

Recently, an Expanding Window FEC coding scheme has been proposed[6]-[9]. As shown in Fig.1, the FEC parity packets of the $n$th frame, $C_n$, are generated with all source packets in frame $i$ with $1 \leq i \leq n$ of the current GOP. By this windowing approach, the block size is guaranteed, and UEP is also achieved as parity packets can help to recover all its former frames. Moreover, no delay is introduced as no packets of the following frames are used. In [6] and [7], a representative Expanding Window Fountain code using a windowing technique is proposed. While Expanding Window Fountain codes based UEP for scalable video coding (SVC) is introduced in [8] and [9]. A recent work in [10] proposed an expanding window code for real-time video streaming based on randomized RS code. However, all these expanding window codes allocates the parity packets greedily and can not achieve the maximal possible video quality. Besides, to the best of our knowledge, none of them have paid effort to the theoretical analysis of the FEC allocation problem for expanding window codes.

In this paper, an Expanding Window FEC allocation scheme for real-time video streaming services is formulated

∗corresponding author
This work was supported by National High-tech Technology R&D Program (863 Program) of China under Grant 2013AA013504 and National Natural Science Foundation of China under contract No. 61471009.

and the problem of FEC allocation is studied. We first analyse the error probability of Expanding Window Reed-Solomon code (EW-RS) and propose two propositions to illustrate its codec characteristic. Then we derive the video distortion model by introducing an equivalent error probability to simplify the computation complexity. Unequal amount of redundancy is then allocated to different frames according to the final distortion impact. The optimal redundancy allocation is formulated as a constrained nonlinear optimization problem. To reduce the computation complexity and make it suitable for practical implementation, an efficient hill-climbing algorithm is adopted. Experimental results show that significant PSNR improvement can be obtained by our method under various network conditions.

The rest of the paper is organized as follows. Section 2 illustrates the codec feature of EW-RS and formulates the distortion impact by introducing an equivalent error probability. Section 3 proposes the constrained redundancy allocation optimization problem, and a hill-climbing algorithm is adopted to solve it. The performance of the proposed scheme is evaluated in Section 4, and Section 5 concludes the paper.

## 2. EXPANDING WINDOW FEC ALLOCATION

FEC has been proved to be an effective method for video streaming under error-prone network conditions. However, there is a trade-off between the decoding efficiency and latency that the bigger of the coding block size, the higher of the decoding efficiency and latency, and vice versa. Especially in real-time video streaming, this becomes more challenging.

To balance the requirement of the decoding efficiency and latency, the EW-RS scheme, as in Fig.1, is employed. In this method, at the sender side, the later frame is encoded (channel coding rather than source coding) together with all the former frames of the current GOP. While at the receiver side, the parity check equations involve packets of all the previous frames, and the lost packets may be jointly solved. It is worth noting that UEP is automatically achieved as the parity packets of the later frames can be used to correct the former frames. And real-time decoding is achieved as no packets of the following frames is used in the current FEC coding block.

In the next subsections, we will give the formulation of this EW-RS scheme. Specifically, we have derived the error probability model and the video distortion model under the above-mentioned scheme.

### 2.1. Error Probability of EW-RS

In this section, we derive the error probability model under the EW-RS scheme as shown in Fig.1. For each GOP, suppose there are $N$ frames, and $k_n$ and $r_n$ denote the number of source and parity packets respectively for frame $n$.

As we know, for $RS(\hat{N}, \hat{K})$[11], which is one of the most used FEC method, where $\hat{N}$ is the block size and $\hat{K}$ is the number of source packets in each coding block, the source packets can be recovered if and only if the number of lost packets is no higher than the parity packets, i.e. $(\hat{N} - \hat{K})$. More specifically, for our adopted EW-RS scheme, the parity check equations for decoding of frame $n$ can be written as:

$$\begin{cases} \mathbf{A_1 X_1} & = \mathbf{C_1} \\ \mathbf{A_2}(\mathbf{X_1}^T, \mathbf{X_2}^T)^T & = \mathbf{C_2} \\ \cdots & \\ \mathbf{A_n}(\mathbf{X_1}^T, \mathbf{X_2}^T, \cdots, \mathbf{X_n}^T)^T & = \mathbf{C_n} \end{cases} \tag{1}$$

where $\mathbf{X_i} = (x_{i,1}, x_{i,2}, \cdots, x_{i,k_i})^T$ is the $k_i \times 1$ source packet vector of frame $i$ and $\mathbf{C_i} = (c_{i,1}, c_{i,2}, \cdots, c_{i,r_i})^T$ is the $r_i \times 1$ parity packet vector. Every lost packet corresponds to a variable in $X_i$ (source packet) or $C_i$ (parity packet). $\mathbf{A_i}$ is a $r_i \times \sum_{j=1}^{i} k_j$ matrix, denoting the parity check matrix of frame $i$. Note that $\mathbf{X_i}$ is involved in all parity check equations whose right hand is $\mathbf{C_j}$ with $j \geq i$. Moreover, by using randomization for $\mathbf{A_i}$, the rank of coefficients matrix of (1) is full[10].

With (1), before deriving the error probability for each frame, we give two propositions as follows:

**Proposition 2.1.** *In the EW-RS, if frame $n$ can be successfully decoded, all the previous frames must be successfully decoded too.*

*Proof.* As in Fig.1 and (1), parity packets of frame $n$ also encodes all source packets of the previous frames. According to RS codes, if lost packets of frame $n$ is decoded, all other packets encoded by the same parity packets should also be decoded. □

**Proposition 2.2.** *Let $l_i$ be the number of lost packets for frame $i$, then the conditions that frame $n$ can be successfully decoded are:*

$$\begin{cases} l_n \leq r_n \\ l_{n-1} + l_n \leq r_{n-1} + r_n \\ \cdots \\ \sum_{j=1}^{n} l_j \leq \sum_{j=1}^{n} r_j \end{cases} \tag{2}$$

*Proof.* According to Fig.1 and (1), $X_i$ $(\forall 1 \leq i \leq n)$ is involved in all parity check equations of frames $i$ to $n$, denoted by set $\mathcal{E}$. The rank of $\mathcal{E}$ is $\sum_{j=i}^{n} r_j$, which is also the number of all parity packets of frames $i$ to $n$. For constraints in (2), assume there exists a violation denoted as $\sum_{j=i}^{n} l_j > \sum_{j=i}^{n} r_j$, under which frame $n$ can still be successfully decoded. According to proposition 2.1, frames 1 to $n-1$ can also be successfully decoded. Thus, we can further assume that frame $j$ $(\forall 1 \leq j < i)$, can be successfully decoded. Then for $\mathcal{E}$, the number of variables will be more than $\sum_{j=i}^{n} l_j$, which is larger than its rank, and makes it unsolvable. So, the assumption does not hold. □

Proposition 2.1 shows that if a frame can be decoded, all its previous frames can also be decoded, thus the dependency among the frames is eliminated. Proposition 2.2 give the conditions for successfully decoding any frame. Combining propositions 2.1 and 2.1, to evaluate the transmission distortion of the video sequence, we need to obtain the loss probability for each frame. However, it is not an easy job to derive the expression of the frame loss probability since conditions (2) includes $n$ sub-conditions and they are interdependent. Thus, instead of directly deriving the loss probability for each frame which is challenging, a recursive method is designed to derive the frame loss probability approximatively. For frame $n$, we write its loss probability as

$$\begin{aligned} &P(n) \\ =&1-\sum_{i=0}^{r_n}(\mathcal{F}(k_n+r_n,i,p)\sum_{j=0}^{r_n-i}\mathcal{F}(\sum_{z=1}^{n-1}k_z,j,\overline{P}(n-1))) \end{aligned} \tag{3}$$

where $\mathcal{F}(a,b,c) = \binom{a}{b}c^b(1-c)^{a-b}$ denotes the binomial probability density function, and $p$ is the channel transmission error rate. $\overline{P}(n-1)$ here is defined as the Equivalent Error Probability (EEP), which estimates the overall packet loss probability of frames 1 to $n-1$. By definition, the probability of all the frames being decoded (channel decoding) calculated by either $\overline{P}(n-1)$ or $P(n-1)$ should be the same, written as

$$(1-\overline{P}(n-1))^{\sum_{i=1}^{n-1}k_i} = 1-P(n-1) \tag{4}$$

the left of which represents the probability of decoding frames 1 to $n-1$ using EEP, meaning that all the packets are received.

Derivation of the loss probability of all frames now becomes an iteration, while the initial value $P(1)$ is

$$P(1) = 1-\sum_{i=0}^{r_1}\mathcal{F}(k_1+r_1,i,p) \tag{5}$$

### 2.2. Transmission Distortion Model

Generally, the loss of previous frames will affect the quality of later frames due to the dependency among frames, this is also known as error propagation. For instance, if the I frame is lost, it may lead to the decoding failure of all frames which refer to it.

In this work, according to proposition 2.1, as long as the current frame can be decoded, all referenced frames that have been lost will get decoded, and the reference buffer will be updated to stop the error propagations. Therefore, we can simply use the PSNR of each frame to evaluate its weight. For the $n$th frame, its PSNR is denoted as $\gamma(n)$ and the expected distortion at the receiver side is denoted as $D(n)$, we have

$$D(n) = \gamma(n)P(n) \tag{6}$$

Based on the loss probabilities and the distortion impacts of video frames, the expected distortion impacts of a GOP can be expressed as

$$\overline{\mathcal{D}}(\mathbf{r}) = \sum_{i=1}^{N}D(i) \tag{7}$$

**Algorithm 1** The hill-climbing algorithm for (LP1)

$\mathbf{r}, \mathbf{r}^* \leftarrow (0,\cdots,0)$
$\overline{\mathcal{D}} \leftarrow +\infty$
**for** $i \leftarrow 1, R$ **do**
  $\mathbf{r} \leftarrow \mathbf{r}^*$
  **for** $j \leftarrow 1, N$ **do**
    $r_j \leftarrow r_j + 1$
    $\overline{\mathcal{D}}_t \leftarrow \overline{\mathcal{D}}(\mathbf{r})$
    **if** $\overline{\mathcal{D}}_t < \overline{\mathcal{D}}$ **then**
      $\overline{\mathcal{D}} \leftarrow \overline{\mathcal{D}}_t$
      $\mathbf{r}^* \leftarrow \mathbf{r}$
    **end if**
    $r_j \leftarrow r_j - 1$
  **end for**
**end for**

where $\mathbf{r}$ denotes the redundancy allocation vector $(r_1,\cdots,r_N)$.

## 3. PROBLEM FORMULATION AND OPTIMIZATION

To improve the performance of FEC, our objective is to minimize the expected distortion impacts. The optimization of redundancy allocation $\mathbf{r}$ can be formulated as

$$\begin{aligned} \mathbf{r}^* =& \arg\min_{\mathbf{r}} \quad \overline{\mathcal{D}}(\mathbf{r}) && \text{(LP1)} \\ \text{s.t.} \quad & \sum_{i=1}^{N}(k_i+r_i) \le C && \text{(i)} \\ & r_i \ge 0 \qquad \forall i \in [1,N] && \text{(ii)} \end{aligned}$$

where $C$ denotes the bandwidth limitation, and constraint (i) restricts that the number of source and parity packets must be no more than $C$.

The objective of this optimization problem is to derive the optimal redundancy allocation vector $\mathbf{r}^*$, so that the UEP is achieved with the best performance, and the total expected distortion is minimized. It is a nonlinear integer optimization problem, and performing an exhaustive search over the whole solution space whose size is $N^R$ (with $R$ being the number of total parity packets) would be computationally very expensive and therefore impractical.

Thus, a suboptimal solution is proposed to simplify this problem. With (LP1), the goal is to find the redundancy allocation $\mathbf{r}^*$ so as to obtain the minimum $\overline{\mathcal{D}}$. To reduce the complexity, we develop a hill-climbing algorithm as shown in Algorithm 1. In this algorithm, all the parity packets are allocated to the frames iteratively that for each parity packet, it will be allocated to the frame that minimize the video distortion. The complexity of our proposed algorithm is reduced to $\mathcal{O}(NR)$ while the results prove almost the same with the algorithm that traverses the entire solution space.

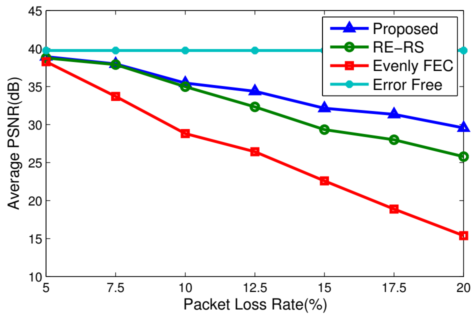


**Fig. 2**: Average PSNR under different packet loss rate with bandwidth=1Mbps, *Foreman* coded 871 Kbps is used

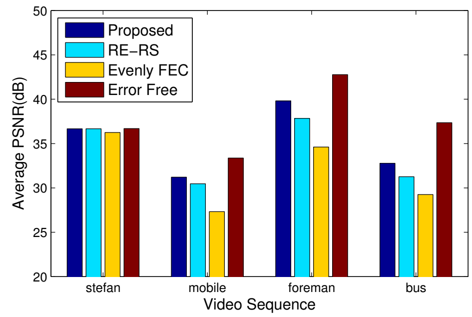


**Fig. 3**: Average PSNR under different sequences with packet loss rate=10%, bandwidth=2Mbps. Sequence bitrates set as, *Stefan*:1441Kbps, *Mobile*:1786Kbps, *Foreman*:1812Kbps and *Bus*:1890Kbps

## 4. EXPERIMENTS

In this section, we evaluate the performance of our proposed FEC allocation approach. Four sequences, Foreman, Bus, Mobile, and Stefan, are tested. All sequence are encoded by JM18[12] with the length of 90 frames. The GOP structure is IPPP with 30 frames and the reference frame number is one. For performance comparison, RE-ES scheme [10], which has been proved to outperform many state-of-the-art approaches, and Evenly FEC, which also meet the real-time constraint but offers no UEP, are also implemented.

Firstly, we compare the PSNR performance under different packet error rate of our all the three approaches. In this experiment, the PSNR performance of all approaches is compared under various packet loss rate. The sequence *Foreman* is tested whose rate is 871kbps, and the bandwidth is set to 1Mbps. The results are shown in Fig.2 and the PSNR of the original sequence is plotted as a benchmark. From the figure, we can clearly find that our proposed approach outperforms all the other approaches under all packet loss rates. It is also worth to notice that as the packet loss rate increases, the gap between our approach and the others becomes bigger. This is mainly because our approach changes the allocation of redundancy to adapt different network conditions while the RE-RS approach simply allocate the parity packet equally. And in the Evenly FEC method, the parity packet are allocated without considering the dependency and unequal importance of the frames. Specifically, when packet loss rate is 20%, the proposed scheme provides nearly 5.0 dB average gain over the RE-RS approach.

Then, the performance is evaluated on all sequences with different video rate, and the bandwidth is set to 2Mbps. The result shown in Fig.3 has demonstrated that our proposed approach always has higher PSNR than the other approaches. Moreover, as the video rate increases, the advantage of UEP is more notable and this also demonstrates the effectiveness of our proposed FEC allocation approach.

At last, for sequence *Foreman*, we show the PSNR curves of each frame, the packet loss rate is set to 10% and the bandwidth is set to 1Mbps. It is shown that for almost all the frames, the average PSNR of our proposed scheme is higher than that of the RE-RS and Evenly FEC scheme. In both our proposed approach and RE-RS, the Expanding Window FEC allocation scheme is adopted and the distortion propagation among frames is eliminated, therefore, much better performance is obtained compared with Evenly FEC. However, compared with RE-RS, the FEC is allocated more effectively and higher average PSNR is obtained.

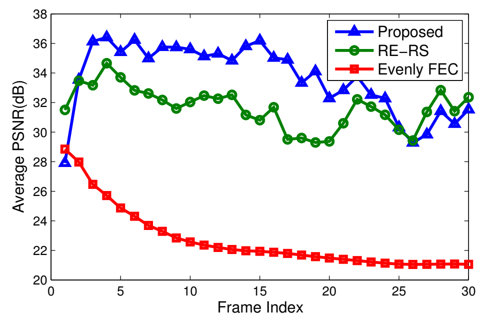


**Fig. 4**: Frame-by-frame video quality in one GOP with packet loss rate=10%, bandwidth=1Mbps, *Foreman* coded 871 Kbps is used

## 5. CONCLUSION

In this paper, an Expanding Window FEC allocation scheme for real-time video streaming services is formulated and the problem of FEC allocation is studied. We first analyze the error probability of EW-RS, and propose two propositions to illustrate its codec characteristic. Then we derive the video distortion model by introducing an equivalent error probability to simplify the computation complexity. Unequal amount of redundancy is then allocated to different frames according to the final distortion impact. The optimal redundancy allocation is formulated as a constrained nonlinear optimization problem. Moreover, to reduce the computation complexity, a high-efficiency hill-climbing algorithm is developed. Experimental results show that significant PSNR improvement can be obtained by our method under various network conditions.

## 6. REFERENCES

[1] Xiaokang Yang, Ce Zhu, Zheng Guo Li, Xiao Lin, and Nam Ling, "An unequal packet loss resilience scheme for video over the internet," *Multimedia, IEEE Transactions on*, vol. 7, no. 4, pp. 753–765, 2005.

[2] Yongfei Zhang, Shiyin Qin, and Zhihai He, "Transmission distortion-optimized unequal loss protection for video transmission over packet erasure channels," in *Multimedia and Expo (ICME), 2011 IEEE International Conference on*. IEEE, 2011, pp. 1–6.

[3] Xinggong Zhang, Chao Zhou, and Zongming Guo, "A novel jscc scheme for scalable video transmission over mimo systems," in *Image Processing (ICIP), 2012 19th IEEE International Conference on*. IEEE, 2012, pp. 2277–2280.

[4] Chao Zhou, C-W Lin, Xinggong Zhang, and Zongming Guo, "A novel jscc scheme for uep-based scalable video transmission over mimo systems," *Circuits and Systems for Video Technology, IEEE Transactions on*, 2014.

[5] Yao Wang, Stephan Wenger, Jiantao Wen, and Aggelos K Katsaggelos, "Error resilient video coding techniques," *Signal Processing Magazine, IEEE*, vol. 17, no. 4, pp. 61–82, 2000.

[6] Dino Sejdinovic, Dejan Vukobratovic, Angela Doufexi, Vojin Senk, and Robert J Piechocki, "Expanding window fountain codes for unequal error protection," *Communications, IEEE Transactions on*, vol. 57, no. 9, pp. 2510–2516, 2009.

[7] Sajid Nazir, Vladimir Stankovic, and Dejan Vukobratovic, "Expanding window random linear codes for data partitioned h. 264 video transmission over dvb-h network," in *Image Processing (ICIP), 2011 18th IEEE International Conference on*. IEEE, 2011, pp. 2205–2208.

[8] Dejan Vukobratovic, Vladimir Stankovic, Dino Sejdinovic, Lina Stankovic, and Zixiang Xiong, "Scalable video multicast using expanding window fountain codes," *Multimedia, IEEE Transactions on*, vol. 11, no. 6, pp. 1094–1104, 2009.

[9] Cornelius Hellge, David Gomez-Barquero, Thomas Schierl, and Thomas Wiegand, "Layer-aware forward error correction for mobile broadcast of layered media," *Multimedia, IEEE Transactions on*, vol. 13, no. 3, pp. 551–562, 2011.

[10] Jimin Xiao, Tammam Tillo, and Yao Zhao, "Real-time video streaming using randomized expanding reed–solomon code," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 23, no. 11, pp. 1825–1836, 2013.

[11] Stephen B Wicker and Vijay K Bhargava, *Reed-Solomon codes and their applications*, John Wiley & Sons, 1999.

[12] [Online], "http://iphome.hhi.de/suehring/tml/," 2014.