Cross-Layer FEC-Based Mechanism for Packet Loss Resilient Video Transmission

Roger Immich[1,⋆], Eduardo Cerqueira[2,⋆], and Marilia Curado[1]

[1] Department of Informatics Engineering, University of Coimbra
Pinhal de Marrocos, 3030-290, Coimbra, Portugal
{immich,marilia}@dei.uc.pt

[2] Faculty of Computer Engineering, Federal University of Para
Av. Augusto Correa, 01, 66.075-110, Para, Brazil
cerqueira@ufpa.br

Abstract. Real-time video transmission over wireless networks is now a part of the daily life of users, since it is the vehicle that delivers a wide range of information. The challenge of dealing with the fluctuating bandwidth, scarce resources and time-varying error levels of these networks, reveals the need for packet-loss resilient video transport. Given these conditions, Forward Error Correction (FEC) approaches are desired to ensure the delivery of video services for wireless users with Quality of Experience (QoE) assurance. This work proposes a Cross-layer Video-Aware FEC-based mechanism with Unequal Error Protection (UEP) scheme for packet loss resilient video transmission in wireless networks, which can increase user satisfaction and improve the use of resources. The advantages and disadvantages of the developed mechanism are highlighted through simulations and assessed by means of both subjective and objective QoE metrics.

Keywords: Forward Error Correction (FEC), Video-aware FEC, QoE, Cross-layer, Unequal Error Protection (UEP).

1 Introduction

At present, it is becoming increasingly common to adopt multi-hop wireless networks for different purposes, including a wide range of real-time video services, including streaming. Video streaming has been used by many companies as a part of a business drive to increase productivity, improve collaboration, reduce costs, and streamline and optimize business operations. Following the same trend, non-professional users are producing, sharing and accessing thousands of videos by using both wired and wireless systems. As an example, in October 2011 more than 200 billion videos were viewed online in only one month [1]. This figure means that, on average, each person on the planet has to watch around one online video per day.

⋆ CNPq Fellow - Brazil.

E. Biersack et al. (Eds.): Data Traffic Monitoring and Analysis, LNCS 7754, pp. 320–336, 2013.

However, new mechanisms for increasing the transmission quality are required to support the growth of video traffic. Additionally, video quality may be affected by several factors, some of which are owing to the video characteristics, such as codec type, bitrate, format and the length of the Group of Pictures (GoP) and, even, the content/genre of the video [2]. A further factor is that the perceptual quality is not affected by all the packets in the same way, because there is a link between the packet content and the impact it has on the user's perception of video quality. When this is taken into account, the most important information should be best protected, and thus encourage the conception and use of the Unequal Error Protection scheme (UEP). The video content also plays an important role during the transmission. Videos with a small degree of movement and few details tend to be more resilient to packet loss. In contrast, videos with high levels of detail and movement are more susceptible to losses and the flaws will be more noticeable [3].

Quality of Experience (QoE) metrics are used to assess the level of the video impairments and can be defined in terms of how users perceive the quality of an application or service [4]. Provided that most of the video services are real-time applications, they need a steady and continuous flow of packets, which can be affected by a number of factors specially in wireless environments. The channel conditions in these networks can suddenly change over time due to noise, co-channel interference, multipath fading, and also, the mobile host movement [5]. Despite the problems outlined above, Wireless Mesh Network (WMN) provides a cost-efficient way of distributing broadband Internet access. Another advantage is its flexibility and reliability for a large set of applications in a wide coverage area [6][7]. Nevertheless, one of the major challenges in WMN is how to distribute the available bandwidth fairly among the requesting nodes to support real-time video traffic [8]. For this reason, it is important to optimize the resource usage and thus avoid congestion periods and a high packet loss rate, particularly in resource-consuming applications, such as video streaming.

The adoption of adjustable data protection approaches is of crucial importance to enhance video transmission, and provide both good perceived video quality and low network overhead. Forward Error Correction (FEC) schemes have been used successfully in real-time systems [9]. FEC allows robust video transmission through redundant data sent along with the original set. As a result, if some information is lost, the original data can be reconstructed through the redundant information [10]. However, as mentioned earlier, the resources might be limited and unfairly distributed as well. An adjustable FEC-based mechanism must use UEP schemes to reduce the amount of redundant information. In this approach, the amount of redundancy is chosen in accordance with the relevance of the protected data, and thus give better protection to the most important video details.

This chapter describes a cross-layer and **VI**d**E**o-a**W**are FEC-based mechanism (ViewFEC) for packet loss resilient video transmission with UEP. The objective is to strengthen video transmissions, while increasing user satisfaction and improving the usage of wireless resources. Owing to these factors, the use of video-aware FEC-based mechanisms is suitable to transmit videos with better quality, although it needs additional bandwidth to send the redundant information data. ViewFEC is a novel adaptive mechanism that overcomes this problem by dynamically configuring itself according to the video characteristics and user perception of quality. Using this process only the more sensitive data sets will carry an unequal amount of redundant information, thus maintaining a good video quality and saving resources. The mechanism described here was assessed through simulations with real video sequences, using subjective and objective QoE metrics.

The remainder of this chapter is structured as follows. The related work is shown in Section 2. Section 3 describes ViewFEC and its evaluation is demonstrated in Section 4. Conclusions and future works are summarized in Section 5.

2 Related Work

A wide range of techniques have been proposed to improve the quality of video over wireless networks. The Adaptive Cross-Layer FEC (ACFEC) mechanism uses packet-level error correction at the MAC layer [11]. All the video packets are monitored and, depending on the extent of the losses, the number of FEC recovery packets is either increased or decreased. Nevertheless, no assessment of the network overhead has been conducted, which is very important to assure a fair usage of resources. This approach also does not take account of the video content, and it is well-known that this information has a direct influence on the packet loss resilience of the video [12]. In theory, the ACFEC appears to be a viable solution when the network is healthy and there is sporadic packet loss. However, when congestion occurs, this mechanism will lead to an incremental rise in the number of redundancy packets, which will increase the congestion and cause an unfair distribution of resources.

Another technique that is employed to strengthen the video transmission quality is carried out through a FEC- and retransmission-based adaptive source-channel rate control [13] scheme. The main purpose of this scheme is to ensure the video has playback continuity under uncertain channel variations, as a means of avoiding unneeded FEC redundancy. Video characteristics, such as content and frame type, were not considered in this proposal. In addition, this approach does not use QoE metrics to assess the video quality, as it relies on packet loss levels to predict the QoE values, and finally, it does not measure the network overhead that has been introduced.

A mechanism to improve video transmission over local area wireless networks which use real-time FEC redundancy adjustment and transmission rate was proposed in [14]. To perform the FEC adjustment and transmission rate, all the receivers have to send the network state information to the Access Point (AP) periodically. This mechanism uses application level FEC and pre-encoded video sequences, which means that it will need multiple pre-encoded videos with different bit rates and FEC rate. After this, the system has to switch between different bit streams in real-time. The need for a larger number of pre-processed videos reduces the applicability of this solution in real systems. At the same time, it also requires a high processing power and storage space, because it has to encode the same video multiple times using distinct bit rates and FEC redundancy data.

A proposal to enhance video streaming transmission using concurrent multipath and path interleaving was proposed in [15]. This scheme also uses a dynamic FEC block size which is adapted to the average packet loss rate for each path, allowing concurrent interleaved data to be sent, with FEC protection, over multiple paths. This solution only takes into account the network parameters, and leaves out video characteristics, such as codec and frame type, GoP length, motion and complexity. Following the same pattern as the studies outlined above, no attempt is made to measure the network overhead. Due to the factors mentioned earlier, the ViewFEC mechanism aims to enhance the video transmission quality without adding unnecessary network overhead.

3 Video-Aware FEC-Based Mechanism

In the light of the open issues mentioned earlier, this study describes the Video-aware FEC-based Mechanism (ViewFEC) for resilient packet loss video transmission with UEP to enhance video transmission over wireless networks. In ViewFEC, decisions are made at the network layer by resorting to two modules, the **C**luster **A**nalysis **K**nowledge bas**E** (CAKE) and the **C**ross-**LA**yer infor**M**ation (CLAM). The decision-making process at the network layer provides better deployment flexibility, and allows the ViewFEC mechanism to be implemented at access points, routers or in the video server. The analysis of the information obtained from these two modules enables the ViewFEC mechanism to estimate the optimal redundancy ratio necessary to sustain a good video quality, without adding unnecessary network overhead.

The ViewFEC mechanism is depicted in Figure 1. There are 3 distinct stages. In the first stage, through information fetched from CAKE and CLAM, our mechanism identifies several key video characteristics such as motion and complexity levels, as well as GoP length. In the second stage, further details about the video sequence are gathered, namely the type and relative position of the frames within its GoP. The construction of the FEC blocks and the UEP redundancy assignment take place in the third stage. Further details of each module are described later.

CLAM is one of the main modules of ViewFEC and supports the functions that will define the required amount of redundancy needed to maintain a good video quality. These functions are implemented using cross-layer techniques, accessing information of the application layer, such as the video characteristics, from the network layer, where the module was deployed. The objective of the first function is to identify the GoP length. As previously discussed, when the GoP length is larger, the packet loss has a greater influence on the video impairments. This happens partially because, a new I-Frame that is needed to fix the error, will take longer to arrive. Another important remark is that video sequences with high level of spatial complexity tend to have larger I-Frames in

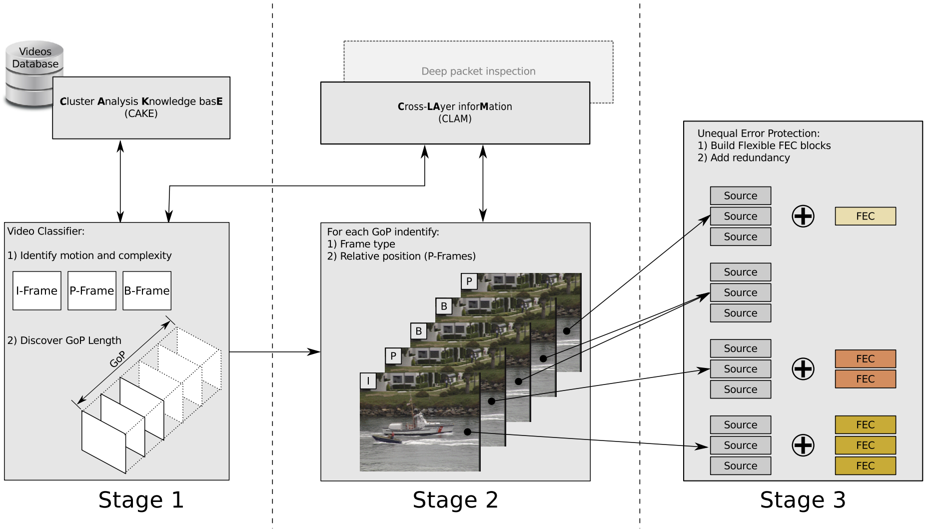


Fig. 1. ViewFEC stages

relation to P- and B-Frames. On the other hand, videos with higher temporal activities tend to have larger P- and B-Frames. Larger frames means that more network packets will be required to carry their data, increasing the chance of these packets being lost. Thus, these packets need more redundancy. Incidentally, as the GoP length is increased, the size of both B-, and especially P-Frames, also increases. In consequence, P-Frames close to the beginning of the GoP take on greater importance regarding the video quality. The role of the second function is to identify the frame type. Different frame types need distinct amounts of redundancy. For example, the loss of an I-Frame will cause more impairments than the loss of a P-Frame, and hence, the loss of a P-Frame will be worse than the loss of a B-Frame. The task of the last function is to identify and compute the relative position of P-Frames inside the GoP. P-Frames that are closer to the beginning of the GoP have more impact, if lost, than those close to the end, and as a result, need more redundancy packets. The combined use of these functions enables ViewFEC to enhance the video quality transmission without adding unnecessary network overhead, and thus, support a higher number of simultaneous users sharing the same wireless resource.

The ViewFEC has a flexible structure, which makes it possible to change the modules to obtain the desired behaviour. When it is not feasible to use a cross-layer design to obtain application layer information, e.g. in a router device, the CLAM module can be exchanged for one that uses another technique to obtain the desired information, for instance, packet and deep packet inspection. By analyzing the packet header of some of the protocols, such as User Datagram Protocol (UDP), Real-time Transport Protocol (RTP) and Transport Stream (TS), it is possible to discover information about codec type and coding parameters, among others. On the other hand, the video content information can only be accessed through deep packet inspection.

The CAKE module optimizes the video transmissions (Figure 1 - Stage 1) through the use of a database with video motion and complexity which is built off-line (before the video distribution). The classification of the video motion and complexity levels needed to create the database is achieved by performing a hierarchical clustering with Euclidean distance. This is a statistical method of partitioning data into groups that are as homogeneous as possible. The video sequences are clustered according to the size of the I-, P- and B-Frames, because they tend to have similar motion activity. This operation only has to be performed once during the setup phase of the mechanism. Afterwards, when the mechanism is running, the relationship between the database information and the videos that are being transmitted in real-time is used to determine a couple of video characteristics, namely motion activity and complexity levels.

The video sequences of the experiments were chosen in compliance with the recommendations of the Video Quality Experts Group (VQEG) [16] and International Telecommunication Union (ITU) [17]. A total of 20 different videos were assessed. Ten of them were used to assemble the database and another set of ten was used to evaluate the ViewFEC mechanism. While remaining in compliance with the recommendations, the videos cover different distortions and content, since they are representative of regular viewing material. As well as this, these sequences contain distinct temporal and spatial details, luminance stress, and still and cut scenes (see Section 4 for more details about the videos).

Video motion and complexity are commonly classified in three categories, namely low, medium and high [3][12] (see Figure 2 at linkage distance (ld) 1). Nevertheless, throughout the experiments, videos with both medium and high complexities behaved roughly the same. Because of this, the linkage distance of our cluster analysis algorithm was chosen to only produce the clusters (Figure 2 at ld 2). This mechanism also employs the Ward method which seeks to reduce the sum of squares between the samples inside the cluster, which better reflects our findings.

The relationship between motion and complexity levels, and frame size (in bytes) and frame position, is shown in Figure 3. This diagram depicts two video sequences - Mobile (A) and Akiyo (B) - each from a different cluster. Only the first GoP of each video was considered, which made it easier to visualize the results. The Mobile sequence has uninterrupted scene modification and a wide-angled camera, and thus, high motion and complexity levels. For this reason, the video has larger frames and greater differences in size between P- and B-Frames, as shown by Figure 3-A. In contrast, the Akiyo video only has a small moving region of interest, just most of the face and shoulders, and also a static background. As a result, it has low motion and complexity levels leading to a smaller difference in size between P- and B-Frames, as depicted by Figure 3-B.

Additionally, the same Figure 3 shows an assessment of the Structural Similarity (SSIM) with frames that have been deliberately discarded. The measurement of this metric is fairly simple, even though it is consistent with the human visual system, and yields good results [18]. The SSIM results were acquired by removing the frame which occupied that position, i.e. the first SSIM value was

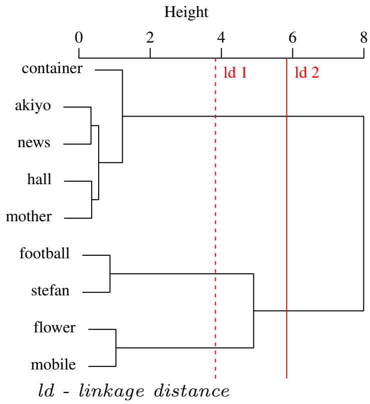
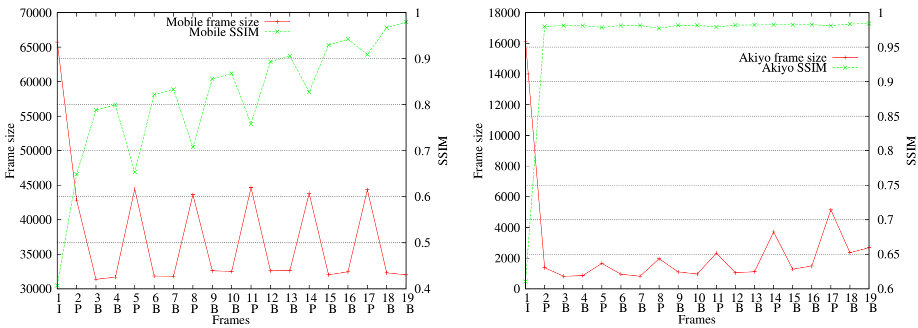


Fig. 2. Cluster Dendrogram

calculated without the first frame, and so forth. It is clear on the basis of these findings that, apart from the fact that in the Mobile video, I- and P- frames have greater significance, the frames closest to the beginning of the GoP also have more impact on QoE video quality when discarded. As expected, the Akiyo sequence behaves differently. It has lower motion and complexity levels being more resilient to packet loss and achieving higher SSIM values [3].



(A) Mobile video sequence *(B) Akiyo video sequence*

Fig. 3. Frame size x QoE (SSIM)

The CAKE module is aware of these video characteristics and can determine the motion activity and complexity levels of each GoP that are being transmitted. The GoP length is another important factor, which was fetched from the CLAM module. Despite the fact that the GoP length remains the same, all these parameters are assigned GoP by GoP (Figure 1 - Stage 1). This is done because it is possible to have different motion and complexity levels inside the same video sequence, as expected for Internet videos. The next step (Stage 2) is responsible for retrieving details of the frame type and relative frame position inside the GoP (for P-Frames) from the CLAM module. By the aid of these details, our mechanism will be able to correctly identify the video characteristics needed to configure the amount of redundancy in the next stage.

In the last step (Stage 3), the amount of redundancy needed is calculated in accordance with the details obtained from the previous stages. This tailored amount of redundancy is used to optimally adjust the FEC scheme. In these experiments, a Reed-Solomon (RS) code was used as an erasure code, because it offers less complexity, and consequently achieves a better performance for real-time services [19], but any other alternative scheme could be used. A RS code consists of n, s, and h elements. The total block size, including the redundancy data, is represented by n, and s indicates the original data set size, therefore the parity code is (n, s). Finally, the parameter h defines the amount of redundancy, which could also be represented as $h = n - s$. Before the original data set s can be restored, at least $(n - h)$ packets have to arrive successfully. The recovery rate can be expressed as h/n or $(n - s)/n$, which means that the robustness to losses is given by the size of h.

The ViewFEC settles the parity code in real-time. In other words, both n and h parameters, are adjusted at Stage 3, based on video characteristics found at Stages 1 and 2, obtained from the CAKE and CLAM modules, respectively. The first parameter of the parity code, n, is used to build the Flexible FEC Block (FFBlock) scheme. This scheme involves dividing the I- and P-Frames into groups of packets, allowing each group to have an individual redundancy data size. This individual size is defined by the second parameter, h, and provides a tailored amount of redundancy for each FFBlock. Hence, rather than using a single redundancy amount to all the frames and video sequences, the ViewFEC mechanism uses an adjustable amount, since it is capable of yielding good results in different network conditions and also supporting a wide range of video characteristics.

The adjustable amount of redundancy data assigned by ViewFEC is the outcome of the joint evaluation of the frame type and position inside the GoP as well as the video motion and complexity levels. By adopting this procedure, we are able to infer the spatio-temporal video characteristics, and as a result, to choose the optimal redundancy amount, h, for each FFBlock. Owing to this, the ViewFEC mechanism is able to achieve better video quality, and has the further advantage of reducing the amount of data that needs to be sent through the network, decreasing the overhead and providing a reasonable usage of wireless resources. This is a very important achievement, because as the network grows larger, the number of concurrent transmissions is increasing with it, and this may cause serious interference problems. The situation gets even worse if we add more overhead due to redundant information. This means that, if the overhead is reduced, a larger number of users will be able to receive more videos with better quality, thus boosting the overall capabilities of the system.

A pseudo-code of the ViewFEC operation is shown in Figure 4. This illustrates how the GoP length and motion detection are performed, and also, the steps taken to assign a tailored amount of redundancy. The algorithm starts with a loop, at line 1, that passes through all the GoPs in a video sequence. At line 4, there is a second loop, which is inside the first one, and will walk through all the frames within a GoP, and only apply the redundancy that is needed. The information retrieval from CAKE and CLAM modules occurs through lines 2, 3, 5, and 11. Since the redundancy amount of P-Frames also depends on their relative position inside the GoP, it has to be treated differently from the I-Frames; this difference is noticeable at line 11.

```
01 for each GoP
02   CAKE.getGopMotion(GoP)
03   CLAM.getGopLength(GoP)
04   for each frame
05       case (CLAM.getFrameType(frame))
06           I-Frame:
07               buildFFBlock(frame)
08               addRedundancy(frame)
09               sendFrame(frame)
10           P-Frame:
11               CLAM.getRelativePosition(frame)
12               buildFFBlock(frame)
13               addRedundancy(frame)
14               sendFrame(frame)
15           B-Frame:
16               sendFrame(frame)
17       end case
18   end for
19 end for
```

Fig. 4. ViewFEC pseudo-code

In Equation 1, it is possible to calculate the amount of redundancy added by the ViewFEC mechanism to each GoP (R_{GoP}). FS_i describes the number of packets of the frame that are being transmitted and FT_i holds the frame type, as shown in Equation 2. If $\gamma > 0$, some level of redundant information will be added to that frame. If we have the vector $(\gamma I, \gamma P, \gamma B)$ with elements $(1, 1, 0)$, for example, only I- and P-Frames will receive redundant information. Additionally, if there is a need to further improve the video quality even if this leads to an increased overhead in the network, the elements of the vector could be $(2, 1, 0)$, meaning that the I-Frames would receive twice the amount of redundant information than the P-Frames (assuming that the other parameters were equal). The parameter C_{GoP} in Equation 3 describes the motion and complexity levels. If the mechanism is using two distinct video clusters, it is possible to define the vectors $(\alpha High/Medium, \alpha Low)$, with elements $(1, 0.5)$, and $(\alpha High, \alpha Medium, \alpha Low)$, with elements $(1, 0.5, 0.25)$. In the former, the cluster with high motion and complexity levels would receive twice the amount of redundancy than the cluster with low levels. If more redundancy levels are needed, the latter could be used, which means that three levels of motion and complexity will be addressed, high, medium and low, respectively. RP_i is the last parameter in Equation 1, which defines the relative distance of the P-Frames inside the GoP. As previously mentioned, frames closer to the end of the GoP are likely to receive less redundant information because the impact of packet loss will be smaller than a loss near the beginning of the GoP, specially in video sequences with larger GoP length. The notation used in the equations is shown in Table 1.

Table 1. Adopted Notation

Notation	Meaning
R_{GoP}	ViewFEC redundancy amount per GoP
FS_i	Frame size in packets of number i_{th} frame
FT_i	Frame type of number i_{th} frame
C_{GoP}	GoP motion and complexity level
RP_i	Relative position of number i_{th} P-Frame
N_{GoP}	Number of GoPs in the video sequence

$$R_{GoP} = \sum_{i=0}^{GoPLength} \left[FS_i \times FT_i \times C_{GoP} \times \frac{1}{RP_i} \right] \tag{1}$$

$$FT_i = \begin{cases} \gamma > 0 \text{ , } \textit{send frame with redundancy} \\ 0 \text{ , } \textit{frame without redundancy} \end{cases} \tag{2}$$

$$C_{GoP} = \begin{cases} 1 \text{ , } \textit{high motion and complexity} \\ 0 \leq \alpha < 1 \text{ , } \textit{otherwise} \end{cases} \tag{3}$$

To compute the total amount of redundant information within a video sequence, just perform the sum of all the redundant information of each GoP, which is given by R_{GoP}. On the other hand, the average amount of redundant data, $\bar{R}$, can be found in Equation 4.

$$\bar{R} = \frac{1}{N_{GoP}} \sum_{i=0}^{N_{GoP}} R_{GoP(i)} \tag{4}$$

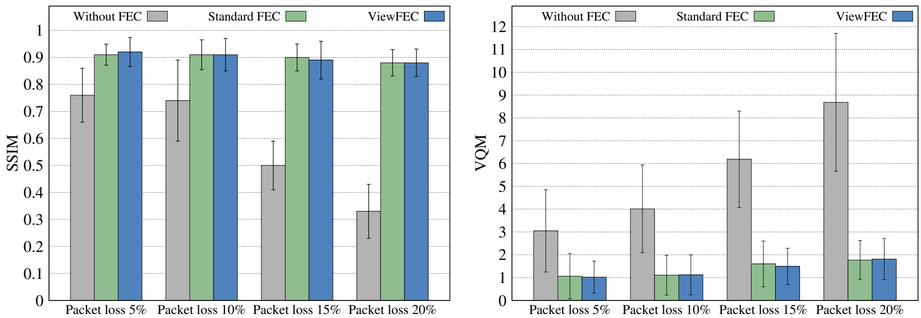
4 Performance Evaluation and Results

The primary goal of the ViewFEC mechanism is to reduce the unneeded overhead, while maintaining videos with an acceptable level of quality. The evaluation experiments were carried out by using Network Simulation 3 (NS-3) [20]. The evaluation setting comprises nine nodes placed in a grid form (3x3), 90 meters apart from the closest neighbour. The routing protocol used was Optimized Link State Routing Protocol (OLSR). An 800 kbps Constant Bit Rate (CBR) background traffic was set. Ten different video sequences were employed in the evaluation scenario [21], with Common Intermediate Format (CIF) size (352x288), H.264 codec and 300 Kbps. All the videos have a GoP length of 19:2, meaning that every 19 frames another I-Frame will be placed and after each two B-Frames, there will be one P-Frame. The error concealment method used by the decoder was Frame-Copy, that is, the lost frames will be replaced by the last good one received. The Gilbert-Elliot loss model is used to produce realistic wireless loss patterns and four different packet loss rates were used: 5%, 10%, 15% and 20%.

Three different cases were used for the video transmission protection. The simplest case (1) is without any type of enhancement (Without FEC). The second case (2) adopts a video-aware FEC-based approach (where both I- and P-Frames are protected in an equal way) with a static amount of redundancy set to 80% (Standard FEC). This amount of redundancy showed the best video quality under the highest packet loss rate defined and was achieved on the basis of a set of detailed experiments. Finally, the last case (3) adopts our proposed approach of an adaptive unequal error protection (ViewFEC). Each of these three cases was simulated 20 times (for packet loss rates of 5%, 10%, 15% and 20%), which means that 80 simulations were carried out for each case, resulting in 240 simulations per video. A total of 10 videos were used, resulting in 2400 simulations. Owing to the distinct initial seeds used to generate the random number, each simulation has a different packet loss pattern.

Subjective and objective QoE metrics were used to assess the video quality obtained from the different cases, namely Structural Similarity Metric (SSIM) [18], Video Quality Metric (VQM) [22] (which were adopted because both are the most widely used objective metrics [23]) and Mean Opinion Score (MOS) [24][25]. With the aid of a set of indicators correlated to the user's perception of quality, the objective metrics perform the assessment without human intervention. The quality assessment was conducted by Evalvid [26] and MSU Tool [27].

Figure 5 shows the average number of the SSIM and VQM values for all the video sequences. In the SSIM metric, the values closer to one indicate a better video quality. As Figure 5-A illustrates, with an increase in the packet loss rate, there is a sharp decline in the video quality of sequences that are transmitted without any type of protection mechanism. At the same time, the video sequences that use either type of FEC-based mechanisms, were able to maintain a good quality. Another important factor that should be noted, is that with 5% and 10% of packet loss rates, the video quality of sequences without FEC are, on average, virtually the same. This can be explained by the natural video resiliency to a certain amount of packet loss. Generally speaking, video sequences with low spatial and temporal complexities are more resilient to loss, and achieve better results in the QoE assessment. Other sequences, with high spatial and temporal complexities, had poor results, and despite the similar average, the standard deviation is higher with a packet loss rate of 10%. This means that the obtained QoE assessment values are more distant from each other. Almost the same pattern is found in the VQM values in Figure 5-B. In this metric, videos with better quality score close to zero. With a packet loss rate of 5% and 10%, the VQM values are also fairly close to each other. This is not as evident as in the SSIM metric because VQM tends to be more rigid when there are video impairments, and yields poor results to videos with fewer flaws. For the same reason, the standard deviation of this metric has a tendency to be higher than the SSIM metric.



(A) SSIM average values *(B) VQM average values*

Fig. 5. Average QoE values for all video sequences

Table 2 summarizes the results and shows the improvement of each scenario in percentage terms. The best values for the VQM metric are the smallest, while for SSIM, they are the highest. As expected, both FEC-based mechanisms produce more valuable results when the network has a higher packet loss rate, in our scenario an average of 20%. For example, it was possible to achieve a reduction of over 79% in VQM values, this means ≈ 4.9 times smaller scores. With the SSIM metric, there was an increase of over 166% in the results, meaning ≈ 2.66 times higher values.

Although objective tests can easily assess video quality, they fail to capture aspects of human vision, and thus, subjective evaluations are also required. However this type of assessment tends to be expensive and time-consuming, and in view of this, we decided to select our best-case scenario to perform these experiments. With

Table 2. QoE values and improvement

Packet loss rate	QoE Metric	Without FEC	Video-aware FEC	Video-aware FEC Improvement	ViewFEC	ViewFEC Improvement
Packet loss 5%	VQM	3.05	1.06	↓65.14%	1.02	↓66.48%
	SSIM	0.76	0.91	↑19.74%	0.92	↑21.05%
Packet loss 10%	VQM	4.01	1.11	↓72.36%	1.12	↓72.09%
	SSIM	0.74	0.91	↑22.97%	0.91	↑22.97%
Packet loss 15%	VQM	6.19	1.60	↓74.09%	1.49	↓75.87%
	SSIM	0.50	0.90	↑80.00%	0.89	↑78.00%
Packet loss 20%	VQM	8.68	1.77	↓79.60%	1.81	↓79.14%
	SSIM	0.33	0.88	↑166.67%	0.88	↑166.67%

a packet loss rate of 20%, both Standard and ViewFEC mechanisms achieved better results, and the most significant differences appeared in the videos transmitted without protection mechanism. MOS is one of the most widely used approaches for subjective video evaluation. This follows one of the ITU-T recommendations and uses a predefined quality scale for a group of people scoring video sequences. The MOS scale ranges from 1 to 5, where 5 is the best possible score. A Single Stimulus (SS) method (standard ITU-R BT.500-11 [24]) was chosen because it was a suitable means of carrying out the quality assessment of emerging video applications [28].

The results of the subjective experiments are depicted in Figure 6. The case without FEC has, on average, 2.05 of MOS, which is considered as a poor video quality with annoying impairments. On the other hand, when the FEC-based mechanisms (Standard FEC and ViewFEC) were used, the MOS average values were 4.39 and 4.37, respectively. This indicates that the video quality is between good and excellent, with minor but not annoying impairments, once again, corroborating the objective findings. The different video assessment values, which can be visualized in the results, are due to the unique characteristics of the video sequences. Small differences in motion and complexity levels can influence the obtained values. As a result, it is important to use various types of video when conducting the experiments. This means that a part of the main objective of ViewFEC was achieved, which was to maintain the video quality.

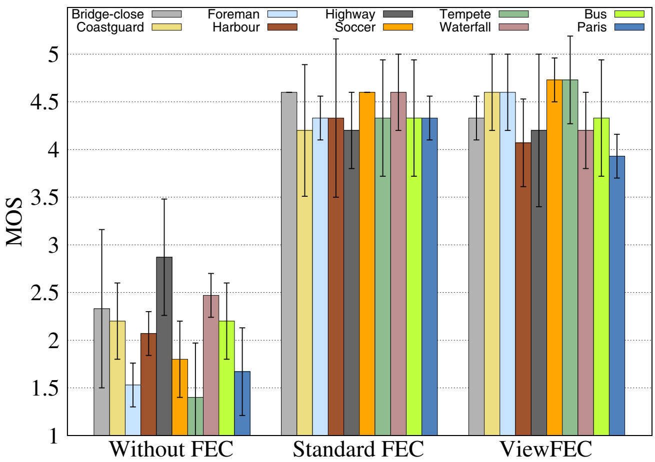


Fig. 6. Average MOS per video sequence with 20% packet loss

Both of the objective and subjective QoE assessments that were employed demonstrated that the ViewFEC mechanism was able to maintain a good video quality in different scenarios. However, another goal of our mechanism is to reduce the network overhead. Due to the limited wireless channel resources, the uneven bandwidth distribution and the interference caused by concurrent transmissions present in WMN, this is a very important issue. In our set up, the network overhead is given by the summation of the size of all video frames that are transmitted. This allows to measure the specific overhead added by the mechanism. Up to now, neither ViewFEC mechanism nor Standard FEC have been able to adjust the FEC parameters to the state of the network; hence, all packet loss rates have the same FEC overhead. As shown in Figure 7, the network overhead added by the Standard FEC was between 53% and 78%. Conversely, the ViewFEC had considerably less overhead and remains between 34% and 47%. This means that the ViewFEC mechanism imposes, on average, 40% less network overhead, with equal or slightly better video quality, as illustrated by Figures 5 (A and B), and 6.

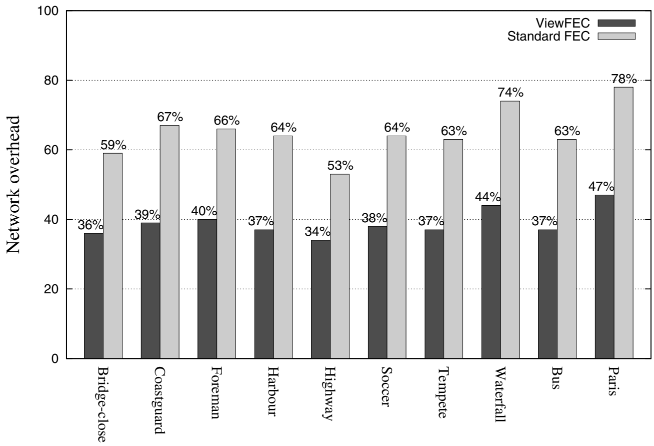


Fig. 7. Network overhead (%)

Owing to a lack of space, the Coastguard video, which is one of the best-case scenarios, was chosen to visualize the results from the viewpoint of the user. Some frames of a video sequence used in the tests, were selected at random and are displayed in Figure 8. In this case, the ViewFEC achieved 4.6 in MOS scale as well as reducing the network overhead, as shown in Figure 6. An improvement of more than 109% was achieved when compared with the video sent without mechanisms to improve the quality (without the FEC scheme, it only reached 2.2 in MOS scale). If compared with the Standard FEC mechanism, which reached 4.2, the ViewFEC still achieves better results with more than 21% improvement in video quality. Furthermore, the standard deviation is considerably smaller, meaning that ViewFEC gets results that are more closely bunched, which indicates a more stable and reliable mechanism.

ViewFEC yielded good results and enabled packet loss resilient video transmission, thus, improving the video quality in the WMN. At this time, the mesh network is only used as a test scenario, and we expect that our mechanism will show its real benefits once it is tailored to these kinds of networks.

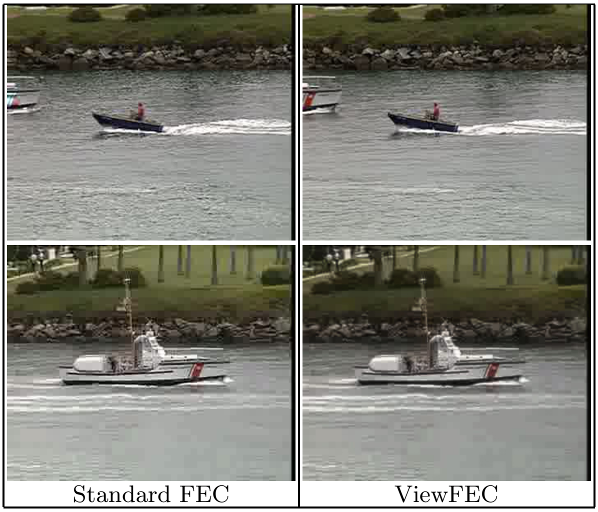


Fig. 8. Frames 10 and 206 of Coastguard video for Standard FEC and ViewFEC

5 Conclusions and Future Works

An effective approach to increase packet loss resilience in video transmission is essential for the growth of video streaming over wireless networks. The Video-Aware FEC-based mechanism for packet loss resilient video transmission provides it with the capacity to enhance video transmission without adding unnecessary network overhead, leading to a better usage of the already scarce wireless resources. A set of controlled experiments was carried out that took into account the different types, complexities, and motions of the video sequences. Network configurations were adopted with different packet loss rates. The simulation results show that the ViewFEC outperforms non-adaptive FEC-based schemes in terms of video quality and in particular, network overhead. The use of Flexible FEC blocks increases the resilience of video transmission to packet losses (by allowing a higher recovery rate), and as a result, achieves better video quality. Video codecs tend to be resilient to a certain amount of packet loss, and because of this, we realized that there was no need to protect all the packets to enhance video quality, especially the information closer to the end of the GoP. The network overhead perceived in the simulations using ViewFEC was between 34% and 47% (on average, 40% less than Standard FEC). As a future work, a systematic variation of the GOP length and structure, as well as the burst length, given by the GE model, will be performed providing a broadest way to validate the results. Also in the next stages, the mechanism will be adjusted to obtain the network state, as a means of enabling it to better adapt to these scenarios as well. Another important point, as previously mentioned, is that through the use of Flexible FEC block, it will be easier to seek out the unique optimization opportunities that WMN offer, i.e. concurrent multipath transmission, network coding and, opportunistic routing.

Acknowledgment. This work was partially funded by the Portuguese Ministry of Science (scholarship contract SFRH/BD/79094/2011) and by FCT UBIQUIMESH project (PTDC/EEATEL/105472/2008).

References

1. comScore, "More than 200 billion online videos viewed globally in october," comScore inc., Tech. Rep. (2011), `http://www.comscore.com/Press_Events/Press_Releases/2011/12/More_than_200_Billion_Online_Videos_Viewed_Globally_in_October`
2. Yuan, Y., Cockburn, B., Sikora, T., Mandal, M.: A gop based fec technique for packet based video streaming. In: Conference on Commun (WSEAS). ICCOM 2006, pp. 187–192 (2006)
3. Khan, A., Sun, L., Jammeh, E., Ifeachor, E.: Quality of experience-driven adaptation scheme for video applications over wireless networks. IET Commun. 4(11), 1337–1337 (2010)
4. Piamrat, K., Viho, C., Bonnin, J.-M., Ksentini, A.: Quality of experience measurements for video streaming over wireless networks. In: Third International Conference on Information Technology: New Generations, pp. 1184–1189 (2009)
5. Lindeberg, M., Kristiansen, S., Plagemann, T., Goebel, V.: Challenges and techniques for video streaming over mobile ad hoc networks. Multimedia Systems 17, 51–82 (2011)
6. Zhu, R.: Intelligent rate control for supporting real-time traffic in wlan mesh networks. Journal of Network and Computer Applications 34(5), 1449–1458 (2011)
7. Akyildiz, I., Wang, X.: A survey on wireless mesh networks. IEEE Communications Magazine 43(9), S23–S30 (2005)
8. Liu, T., Liao, W.: Interference-aware qos routing for multi-rate multi-radio multi-channel ieee 802.11 wireless mesh networks. IEEE Transactions on Wireless Communications 8(1), 166–175 (2009)
9. Nafaa, A., Taleb, T., Murphy, L.: Forward error correction strategies for media streaming over wireless networks. IEEE Communications Magazine 46(1), 72–79 (2008)
10. Lee, J.-W., Chen, C.-L., Horng, M.-F., Kuo, Y.-H.: An efficient adaptive fec algorithm for short-term quality control in wireless networks. In: Advanced Communication Technology (ICACT), pp. 1124–1129 (February 2011)
11. Han, L., Park, S., Kang, S.-S., P., H.: An adaptive fec mechanism using cross-layer approach to enhance quality of video transmission over 802.11 wlans. In: TIIS, pp. 341–357 (2010)
12. Aguiar, E., Riker, A., Cerqueira, E., Jorge Gomes Abelem, A., Mu, M., Zeadally, S.: Real-time qoe prediction for multimedia applications in wireless mesh networks. In: IEEE 4th Future Multimedia Networking, IEEE FMN 2012 (2012)
13. Hassan, M., Landolsi, T.: A retransmission-based scheme for video streaming over wireless channels. Wirel. Commun. Mob. Comput. 10, 511–521 (2010)
14. Alay, O., Korakis, T., Wang, Y., Panwar, S.S.: Dynamic rate and fec adaptation for video multicast in multi-rate wireless networks. Mobile Networks and Applications 15, 425–434 (2010)
15. Tsai, M.-F., Chilamkurti, N.K., Zeadally, S., Vinel, A.: Concurrent multipath transmission combining forward error correction and path interleaving for video streaming. Comput. Commun. 34, 1125–1136 (2011)

16. Staelens, N., Sedano, I., Barkowsky, M., Janowski, L., Brunnstrom, K., Le Callet, P.: Standardized toolchain and model development for video quality assessment - the mission of the joint effort group in vqeg. In: Quality of Multimedia Experience (QoMEX), pp. 61–66 (September 2011)
17. ITU-T Recommendation J.247, Objective perceptual multimedia video quality measurement in the presence of a full reference, International Telecommunications Union Std. (2008)
18. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Process. 13(4), 600–612 (2004)
19. Neckebroek, J., Moeneclaey, M., Magli, E.: Comparison of reed-solomon and raptor codes for the protection of video on-demand on the erasure channels. In: Proceedings of Information Theory and its Applications, International Conference, pp. 856–860. IEEE (2010)
20. Weingartner, E., vom Lehn, H., Wehrle, K.: A performance comparison of recent network simulators. In: IEEE Int. Conf. Commun (ICC), pp. 1–5 (2009)
21. Mpeg-4 and h.263 video traces for network performance evaluation, `http://www.tkn.tu-berlin.de/research/trace/trace.html`
22. Pinson, M.H., Wolf, S.: A new standardized method for objectively measuring video quality. IEEE Trans. on Broadcasting 50(3), 312–322 (2004)
23. Chikkerur, S., Sundaram, V., Reisslein, M., Karam, L.: Objective video quality assessment methods: A classification, review, and performance comparison. IEEE Transactions on Broadcasting 57(2), 165–182 (2011)
24. ITU-R Recommendation BT.500-11, Methodology for the subjective assessment of the quality of television pictures, International Telecommunications Union Std. (2002)
25. ITU-R Recommendation P.910, Subjective video quality assessment methods for multimedia applications, International Telecommunications Union Std. (1999)
26. Klaue, J., Rathke, B., Wolisz, A.: Evalvid - a framework for video transmission and quality evaluation. In: 13th Internation Conference on Modeling Techniques and Tools for Computer Performance Evaluation, pp. 255–272 (2003)
27. Vatolin, D., Moskin, A., Pretov, O., Trunichkin, N.: Msu video quality measurement tool, `http://compression.ru/video/quality_measure/video_measurement_tool_en.html`
28. Seshadrinathan, K., Soundararajan, R., Bovik, A., Cormack, L.: Study of subjective and objective quality assessment of video. IEEE Transactions on Image Processing 19(6), 1427–1441 (2010)