Roger Kreutz Immich

# Mechanisms for Resilient
# Video Transmission

·U C·

UNIVERSIDADE DE COIMBRA

DEPARTMENT OF INFORMATICS ENGINEERING
FACULTY OF SCIENCES AND TECHNOLOGY
UNIVERSITY OF COIMBRA

# MECHANISMS FOR RESILIENT VIDEO TRANSMISSION

Roger Kreutz Immich

Doctoral Program in Information Science and Technology
PhD Thesis submitted to the University of Coimbra

Advised by Prof. Dr. Marilia Pascoal Curado
and by Prof. Dr. Eduardo Cerqueira

February, 2017

Departamento de Engenharia Informática
Faculdade de Ciências e Tecnologia
Universidade de Coimbra

# Mecanismos para transmissão resiliente de vídeos

Roger Kreutz Immich

Programa de Doutoramento em Ciências e Tecnologias da Informação
Tese de Doutoramento apresentada à Universidade de Coimbra

Orientado pela Prof. Dr. Marilia Pascoal Curado
e pelo Prof. Dr. Eduardo Cerqueira

February, 2017

1,3,7-Trimethylxanthine, *thank you*

# Acknowledgements

First of all, I would like to say thank you to everyone that was involved in one way or another in this journey.

I would like to give a special mention to both of my advisors, Marilia Curado and Eduardo Cerqueira, for the trust put in me in the first place and also for all the insights, support, and patience. Thanks for making me feel like home since the beginning.

To all my friends in Coimbra, that I will not name here because it would require so many additional pages to do it right. Having said that, one friend I have to name here. Bruno, not only this but thanks for holding the fort while I was away and for facing all the bureaucratic and lagging processes in my name.

Thanks, mom, dad, and sis for always, always, always believing in me unconditionally.

Eve, thanks for giving me the space that I need to finish the Thesis and also for doing everything you could to help me.

# Abstract

Wireless networks are envisaged to be one of the most important technologies to provide cost-efficient content delivery, including for video applications. They will allow thousands of thousands of fixed and mobile users to access, produce, share, and consume video content in an ubiquitous way. Real-time video services over these networks are becoming a part of everyday life and have been used to spread information ranging from education to entertainment content. However, the challenge of dealing with the fluctuating bandwidth, scarce resources, and time-varying error rate of these networks, highlights the need for error-resilient video transmission.

In this context, the combination of Forward Error Correction (FEC) and Unequal Error Protection (UEP) approaches is known to provide the distribution of video applications for wireless users with Quality of Experience (QoE) assurance. In order to correctly perform the UEP it is necessary to identify the most important parts in the video sequences.

To tackle this issue, this thesis proposed a procedure to assess the video characteristics, such as the codec type, the frame type and size, the length and format of the group of pictures as well as the motion vectors, and their related impact on the perceived quality to end-users. Furthermore, as the video content plays an important role on the perceived quality, this thesis also proposes a method to characterise the video's motion intensity. This involves conducting an exploratory data analysis in bootstrap time and then the use of several techniques in real-time to use the found results. The purpose of the above-mentioned proposals is to give support for the main goal of this thesis, which is to propose mechanisms for resilient video transmission.

Taking everything into consideration, this thesis proposes a series of cross-layer video-aware and FEC-based mechanisms with UEP to enhance video transmission in several types of wireless networks. A number of methods to set an adaptive amount of redundancy were used in these mechanisms, such as heuristic techniques, random neural networks, ant colony optimisation, and fuzzy logic.

In the first one, heuristic techniques, the mechanisms rely on human experience to define the best strategy. In addition, the aim is not to reach a perfect solution, but a practical one with satisfactory results. In the random neural networks methods, the neurones are trained and validated before run-time until they are able to perform an adequate numeric categorisation. The ant colony optimisation techniques use a defined metaheuristic based on the ant's pheromones and pre-set rules to compute a precise amount of redundancy. The last one, fuzzy logic techniques, the mechanisms depend on fuzzy rules and sets to find an adequate redundancy ratio.

The advantages and drawbacks of the proposed mechanisms were demonstrated in realistic simulations using real video sequences and actual network traces. The assessments were conducted with well-known QoE metrics. The results show that all the proposed mechanisms were able to outperform the competitors on both perceived video quality and network footprint.

# Resumo

As redes sem fios estão entre as tecnologias mais importantes para prover a entrega de conteúdos a um custo acessível, inclusive no caso de aplicações de vídeo. Esta tecnologia vai permitir que milhares de utilizadores móveis tenham acesso, produzam, partilhem e consumam conteúdo de vídeo de uma maneira ubíqua. Nestas redes, os serviços de distribuição de vídeos em tempo real já se tornaram parte do dia-a-dia e são utilizados para difundir desde conteúdo educacional como também de entretenimento. Entretanto, devido às dificuldades de gerir a flutuação da largura de banda disponível, os recursos escassos, e os erros que variam de tempo-em-tempo, revela-se a necessidade de uma transmissão de vídeo resistente a perdas.

Neste contexto, a utilização conjunta de correção antecipada de erros (do Inglês *Forward Error Correction* - FEC) e também da proteção desigual contra erros (do Ingês *Unequal Error Protection* - UEP) podem auxiliar na distribuição de serviços de vídeo para utilizadores de rede sem fios com garantia de qualidade de experiência (do Inglês *Quality of Experience* - QoE). Procurando aplicar de forma adequada a proteção desigual de erros é necessário identificar corretamente as partes mais importantes das sequências de vídeo.

Para resolver esta situação, esta tese propôs um procedimento para avaliar as características dos vídeos, tais como o tipo de codec, o tipo e tamanho dos quadros, o comprimento do grupo de imagens e também os vetores de movimentações, bem como o impacto destas características na qualidade percebida pelos utilizadores finais. O conteúdo do vídeo também é importante na definição da qualidade. Esta tese também propôs um método para melhor caracterizar a intensidade de movimentação. Isto envolve realizar uma analise exploratória de dados antes de iniciar o sistema, assim como a utilização de diversas técnicas para aceder aos resultados obtidos em tempo real. As técnicas acima foram propostas para dar suporte ao principal tema desta tese, que é o de projetar e construir mecanismos para a transmissão resiliente de vídeos.

Levando em consideração as informações acima, esta tese propôs diversos mecanismos utilizando técnicas de "cross-layer" e "video-aware" baseadas na correção antecipada dos erros e com proteção desigual da informação. O objetivo principal é melhorar a qualidade dos vídeos transmitidos em diversos tipos de redes sem fio. Para chegar a este resultado, diversos mecanismos adaptativos foram utilizados, tais como técnicas heurísticas, redes neuronais aleatórias, otimização por colónias de formigas e lógica difusa.

As técnicas heurísticas referem-se a métodos que utilizam a experiência humana sobre o assunto para definir a melhor estratégia. É importante frisar que o principal objetivo não é encontrar a solução perfeita, mas sim uma solução prática que tenha resultados satisfatórios. Outra técnica utilizada são as redes neuronais aleatórias, onde os neurónios são treinados e validados ante da execução

do mecanismo até que estes atinjam a capacidade de realizar adequadamente a categorização numérica das informações que recebem. No método de otimização por colónia de formigas, a quantidade de redundância é calculada utilizando uma meta-heurística que é baseada nos feromônios das formigas e em regras prédefinidas. O último método é a lógica difusa que utiliza regras e um conjunto de dados para encontrar a quantidade mais adequada de redundância.

As vantagens e desvantagens dos mecanismos propostos foram demonstradas através de simulações realísticas com a utilização de sequências de vídeo reais e arquivos de registos de redes. A avaliação dos mecanismos foram realizadas através de métricas conhecidas de qualidade de experiência. Os resultados obtidos demonstram que os mecanismos propostos foram capazes de obter uma melhor performance que os competidores tanto na questão da qualidade dos vídeos como na questão da sobrecarga da rede.

**Palavras-chave:** Correção antecipada de erros (FEC); FEC orientado ao vídeo; Qualidade de experiência (QoE); Concepção cross-layer; Proteção desigual dos dados

# Foreword

The work detailed in this thesis was accomplished at the Laboratory of Communication and Telematics (LCT) of the Centre for Informatics and Systems of the University of Coimbra (CISUC), within the context of the following projects:

The outcome of the design, experiments, and assessments of several mechanisms on the course of this work resulted in the following publications:

**Journal papers:**

- Immich, R. and Cerqueira, E. and Curado, M., "**Shielding video streaming against packet losses over VANETs**", The Journal of Mobile Communication, Computation and Information, Wireless Networks, Volume 22, Issue 8, pp 2563-2577, Springer, 2015

- Immich, R. and Borges, P. and Cerqueira, E. and Curado, M., "**QoE-driven video delivery improvement using packet loss prediction**", International Journal of Parallel, Emergent and Distributed Systems, Volume 30, Issue 6, pp 478-493, Taylor & Francis, 2015

- Rosário, D. and Cerqueira, E. and Neto, A. and Riker, A. and Immich, R. and Curado, M., "**A QoE handover architecture for converged heterogeneous wireless networks**", The Journal of Mobile Communication, Computation and Information, Wireless Networks, Volume 19, Issue 8, pp 2005–2020, Springer, 2013

**Book chapters:**

- Immich, R. and Cerqueira, E. and Curado, M., "**Improving video QoE in Unmanned Aerial Vehicles using an adaptive FEC mechanism**", in Wireless Networking for Moving Objects: Models, Approaches, Techniques, Protocols, Architectures, Tools, Applications and Services, Volume 8611, pp 198-216, Springer LNCS, 2014

- Immich, R. and Cerqueira, E. and Curado, M., "**Cross-layer FEC-based Mechanism for Packet Loss Resilient Video Transmission**", in Data Traffic Monitoring and Analysis: From measurement, classification and anomaly detection to Quality of experience, Volume 7754, pp 320-336, Springer LNCS, 2013

**Conference papers:**

- Immich, R. and Cerqueira, E. and Curado, M., "**Towards a QoE-driven Mechanism for Improved H.265 Video Delivery**", in the 15th IFIP Annual Mediterranean Ad Hoc Networking Workshop (MED-HOC-NET), 2016

- Immich, R. and Cerqueira, E. and Curado, M., "**Adaptive QoE-driven video transmission over Vehicular Ad-hoc Networks**", in the IEEE Conference on Computer Communications Workshops (INFOCOM), 2015

- Immich, R. and Cerqueira, E. and Curado, M., "**Towards the Enhancement of UAV Video Transmission with Motion Intensity Awareness**", in the IEEE IFIP Wireless Days, 2014

- Immich, R. and Borges, P. and Cerqueira, E. and Curado, M., "**Adaptive Motion-aware FEC-based Mechanism to Ensure Video Transmission**", in the 19th IEEE Symposium on Computers and Communications (ISCC), 2014

- Immich, R. and Borges, P. and Cerqueira, E. and Curado,M., "**Ensuring QoE in Wireless Networks with Adaptive FEC and Fuzzy Logic-based Mechanisms**", in the IEEE International Conference on Communications (ICC), 2014

- Immich, R. and Borges, P. and Cerqueira, E. and Curado, M., "**AntMind: Enhancing Error Protection for Video Streaming in Wireless Networks**", in the 5th IEEE International Conference on Smart Communications in Network Technologies (SaCoNET), 2014

- Cerqueira, E. and Quadros, C. and Neto, A. and Riker, A. and Immich, R. and Curado, M. and Pescapée, A., "**A Quality of Experience Handover System for Heterogeneous Multimedia Wireless Networks**", in the IEEE International Conference on Computing, Networking and Communications (ICNC), 2013

- Immich, R. and Cerqueira, E. and Curado, M., "**Adaptive Video-Aware FEC-based Mechanism with Unequal Error Protection Scheme**", in the 28th Annual ACM Symposium on Applied Computing (SAC), 2013

- Zhao, Z. and Braun, T. and Rosário, D. and Cerqueira, E. and Immich, R. and Curado, M., "**QoE-aware FEC Mechanism for Intrusion Detection in Multi-tier Wireless Multimedia Sensor Networks**", in the International Workshop on Wireless Multimedia Sensor Networks (WMSN), 2012

- Cerqueira, E. and Curado, M. and Neto, A. and Riker, A. and Immich, R. and Quadros, C., "**A Mobile QoE Architecture for Heterogeneous Multimedia Wireless Networks**", in the GC'12 Workshop: The 4th IEEE International Workshop on Mobility Management in the Networks of the Future World (MobiWorld), 2012

- Cerqueira, E. and Neto, A. and Immich, R. and Curado, M. and Riker, A. Barros, H., "**A Parametric QoE Video Quality Estimator for Wireless Networks**", in the GC'12 Workshop: IEEE Workshop on Quality of Experience for Multimedia Communications (QoEMC), 2012

**Co-advisor of MSc thesis:**

- Borges, P. and Immich, R. and Curado, M., "**Mechanisms for resilient video transmission in wireless networks**", University of Coimbra, 2014

**Special session presentations:**

- Borges, P. and Immich, R. and Cerqueira, E. and Curado, M., "**Mechanisms for resilient video transmission in wireless networks: Adaptive FEC mechanism with random neural network classification and ant colony optimization**". 18 Seminário Rede Temática de Comunicações Móveis (RTCM), 2014

- Immich, R., Cerqueira, E. and Curado, M., "**Cross-layer FEC-based Mechanism to ensure Quality of Experience in Video Transmission**". 15 Seminário Rede Temática de Comunicações Móveis (RTCM), 2012

# Contents

# List of Figures

# List of Algorithms

# List of Tables

# Abbreviations and Acronyms

**4G**       4th Generation Networks

**5G**       5th Generation Networks

**A-STAR**   Anchor-based Street and Traffic Aware Routing

**ACK**      Acknowledgement

**ACR**      Absolute Category Rating

**AC**       Access Category

**AODV**     Ad-hoc On-demand Distance Vector

**AP**       Access Points

**ARQ**      Automatic Repeat reQuest

**AU**       Application unit

**AWARE**    Autonomous self-deploying and operation of Wireless sensor-actuator networks cooperating with AeRial objEcts

**B-Frames** Bi-directionally predictive coded frames

**BFP**      Bentley Faust Preparata

**CATT**     Contention-Aware Transmission Time

**CBR**      Constant Bit Rate

**CIF**      Common Intermediate Format

**CSMA**     Carrier Sense Multiple Access

**DCR**      Degradation Category Rating

**DOLSR**    Directional Optimized Link State Routing Protocol

**DSCQS**    Double-Stimulus Continuous Quality-Scale

**DSR**      Dynamic Source Routing

**ECC**      Error-Correcting Code

**EC**       Error Correction

**EDCA**     Enhanced Distributed Channel Access

**ETT**      Expected Transmission Time

**ETX**      Expected Transmission Count

**FANETs**   Flying Ad-Hoc Networks

**FEC**   Forward Error Correction

**FLC**   Fuzzy Logic Controller

**GGR**   Guaranteed Geocast Routing

**GPMOR**   Geographic Position Mobility Oriented Routing

**GPSR**   Greedy Perimeter Stateless Routing

**GSR**   Geographic Source Routing

**GoP**   Group of Picture

**HEVC**   High Efficiency Video Coding

**HFS**   Hierarchical Fuzzy System

**HVS**   Human Visual System

**HWMP**   Hybrid Wireless Mesh Protocol

**I-Frames**   Intra-coded frames

**ITS**   Intelligent Transportation Systems

**ITU-R**   ITU Radiocommunications Sector

**ITU-T**   ITU Telecommunication Standardization Sector

**ITU**   International Telecommunication Union

**LAETT**   Load Aware Expected Transmission Time

**LAN**   Local Area Network

**LDPC**   Low-Density Parity-Check

**LD**   Linkage Distance

**LTE**   Long-Term Evolution

**LTE-Advanced**   Long-Term Evolution Advanced

**MANET**   Mobile Ad-Hoc Network

**MIC**   Metric of Interference and Channel-switching

**MIND**   Metric for INterference and channel Diversity

**MIROSE**   Minimum Interference Route Selection

**MMSE**   Minimum Mean Squared Error

**MOS**   Mean Opinion Score

**MPCA**   Mobility Prediction Clustering Algorithm

**MPEG**   Moving Picture Experts Group

**MSE**   Mean Squared Error

**NACK**    Negative Acknowledgement

**NSDVCR**  Network State Dependent Video Compression Rate

**OBU**     On Board Unit

**OLSR**    Optimized Link State Routing Protocol

**OSI**     Open Systems Interconnection

**P-Frames** Predictive coded frames

**PSNR**    Peak Signal to Noise Ratio

**QoE**     Quality of Experience

**QoS**     Quality of Service

**RF**      Radio Frequency

**RNN**     Random Neural Networks

**RSSI**    Received Signal Strength Indication

**RSU**     RoadSide Unit

**RS**      Reed-Solomon

**RTP**     Real-time Transport Protocol

**RTSP**    Real Time Streaming Protocol

**RTT**     Round-Trip Time

**SDN**     Software-Defined Networks

**SI13**    13-pixel Spatial Information filter

**SNR**     Signal-to-Noise Ratio

**SSIM**    Structural Similarity

**TACR**    Trust dependent Ant Colony Routing

**TCP**     Transmission Control Protocol

**TS**      Transport Stream

**UAV**     Unmanned Aerial Vehicle

**UDP**     User Datagram Protocol

**UEP**     Unequal Error Protection

**UGC**     User-Generated Content

**V2D**     Vehicle-to-Device

**V2G**     Vehicle-to-Grid

**V2I**     Vehicle-to-Infrastructure

**V2P**     Vehicle-to-Pedestrian

**V2V**      Vehicle-to-Vehicle

**V2X**      Vehicle-to-Everything

**VANETs** Vehicular Ad-Hoc Networks

**VQEG**    Video Quality Experts Group

**VQMT**    MSU Video Quality Measurement Tool

**VQM**     Video Quality Metric

**VoD**      Video-on-Demand

**WANETs** Wireless Ad-Hoc Networks

**WAVE**    Wireless Access for Vehicular Environments

**WCETT-LB** Weighted Cumulative Expected Transmission Time with Load
          Balancing

**WLAN**    Wireless Local Area Network

**WMNs**    Wireless Mesh Networks

# Chapter 1

# Introduction

> The saddest aspect of life right now is that science gathers knowledge faster than society gathers wisdom.
>
> *(Isaac Asimov, Book of Science and Nature Quotations)*

## Contents

This thesis contemplates the necessity of providing high-quality video delivery for end-users. In order to do that, several different approaches were adopted solving a considerable number of current issues. In this process, the contributions reached include an approach to quantify the impact of the video characteristics on its quality, a method to characterise the motion intensity of video sequences, as well as several mechanisms to improve the video quality over diverse network scenarios. The research background and motivation of this thesis are presented and investigated followed by the discussion of the proposed objectives together with the respective contributions, as well as the thesis outline.

## 1.1 Background and Motivation

In the last few years, there has been a rapid proliferation of a wide range of real-time video services and applications. This growth can be partly attributed to the recent development and improvement of mobile devices, such as notebooks, tablets, and smartphones, as well as advances in wireless networks [comScore, 2013, Adobe Digital Index, 2014]. This improvement allows an increased number of services and applications to be available to the end-users. The adoption of

video services can be used as means to provide users with both information and entertainment content. Furthermore, the whole Internet, and not just the mobile end-users, is experiencing a considerable growth in traffic that is in part led by these novel real-time video services. According to Cisco, by 2020 the global Internet traffic will be 95 times higher than it was in 2005. Additionally, the video IP traffic will represent over 82% of the global IP traffic [Cisco, 2016].

This situation is easily explained by the large amount of new forms of information and entertainment that are being released every day by thousands of users. This includes, but is not limited to, User-Generated Content (UGC), news websites, social networking, and e-learning. Moreover, a considerable number of television studios, film production companies, and recording studios are also increasing the variety of video content that is available over the Internet. Just to give one example, by 2020 it is predicted that every second almost a million minutes of video content will be sent over the global Internet [Cisco, 2016].

It is clear that the distribution of video content over the Internet is rapidly expanding. End-users are increasingly consuming this type of content that might be anything from homemade videos to professionally developed programming. However, real-time video transmission imposes more challenges than non-real-time data transfer. Keeping a satisfactory perception of video quality requires the provision of sufficient bandwidth, low interference, as well as low delay and jitter throughout the video transmission.

In the light of this, several multi-hop wireless networking technologies are being increasingly employed [Meng et al., 2016], such as Wireless Mesh Networks (WMNs), Flying Ad-Hoc Networks (FANETs), and Vehicular Ad-Hoc Networks (VANETs). The use of these wireless networks introduces additional challenges that have to be addressed to improve the video quality from the end-users perspective. Several of these challenges can be attributed to a more dynamic and error-prone environment. First of all, the wireless channels are shared among the nodes, and because of this, they must collaborate to allow simultaneous video transmissions, as well as other network/Internet traffic.

A further aggravating factor is that this method of transmission is susceptible to severe physical conditions, such as multipath fading, Radio Frequency (RF) interference, shadowing, and background noise. In contrast with the wired links, packet loss does not necessarily mean network congestion, and can often be related to random physical causes, leading to time-varying communications impairments and recurring link interruption. Since most of the video services are real-time applications, they need a steady and continuous flow of packets, which can be affected by several of the above-mentioned factors in wireless environments. These challenges become even more serious in multi-hop wireless networks.

In addition, all the video transmissions in a wireless network are subjected to a shared channel medium and also exposed to a high packet drop rate, delay, and jitter during the transmission. Additionally, the underlying routing protocol might impose some message exchange overhead due to route updates. Burst losses are also fairly common since packets may be sent on non-working routes

before the routing protocol detects the disruption.

To make the matter worst, the growth of the network leads to higher node density which can increase packet collisions, due to the concurrent transmission. Nevertheless, one of the main challenges in this type of network is how to evenly distribute the available bandwidth among the requesting nodes for real-time traffic [Liu and Liao, 2009]. The network related issues and features are further discussed in Chapter 2.

In addition to the network parameters, there are other issues that might influence the video quality. Some arise from the video characteristics, such as the frame and codec type, bitrate, video format and size of the Group of Picture (GoP), and even the content of the video [Yuan et al., 2006, Khan et al., 2010]. Apart from that, not all the packets have an equal impact on perceptual quality. There is a correspondence between the type of information that the packet carries and the impact it has on the user perception of video quality [Greengrass et al., 2009b].

Taking everything into consideration, is necessary a flexible and adjustable mechanism to improve the video quality. Hence, these requirements are best met by using a Unequal Error Protection (UEP) scheme. Through a UEP-based scheme, it is possible to assign different amounts of redundancy to the original data. This helps to better protect the most important information, allowing the transmission of videos with high quality, while introducing less redundant information, therefore decreasing the network overhead. An elaborate review about the video characteristics and the impact on its quality can be found in Chapter 3.

Quality of Experience (QoE) metrics are commonly used to assess the video quality level, allowing also to identify in what situations the flaws are more noticeable [Serral-Gracià et al., 2010]. QoE metrics measure the video degradation and can be defined as how users subjectively perceive the quality of an application or service [Piamrat et al., 2009]. This means that the performance should be measured end-to-end and must reflect the user standpoint. QoE is related to Quality of Service (QoS) metrics, but differs from it by providing a more comprehensive assessment and was designed to provide an appraisal of human expectations in regards to the video service.

QoE metrics can be classified into two general approaches, namely subjective and objective. Subjective evaluations are based on human individuals rating the video quality. Because of this, it is possible to capture a great variety of details that might impact on the perceived quality. To apply these metrics, however, tend to be time-consuming, expensive, and hard to employ in real-time, as well as the results are not reproducible. On the other hand, objective metrics try to mimic the human vision system by using mathematical models to identify and measure the video impairments. Therefore, they do not rely on human interventions and are both verifiable and reproducible. More details about QoE metrics can be found in Chapter 3.

In the light of the aforementioned issues and features, the employment of an adaptive data protection is essential to overcome the transmission challenges in

wireless networks, providing both high-perceived video quality and low network overhead. One way to achieve this level of protection is by resending packets if lost or dropped, i.e., Automatic Repeat reQuest (ARQ) method. Another possibility is to add redundant information to the original data set, in this case, if some information is lost it can be rebuilt, i.e. Forward Error Correction (FEC) schemes. Both techniques have advantages and disadvantages which are explained and discussed in Chapter 4.

A number of different solutions have been proposed to address the aforementioned challenges. Several of these proposals are reviewed and summarised in Chapter 4. However, as far as the literature goes, none of them was able to present a holistic mechanism which takes into consideration QoE-sensitive information, video details, as well as the network conditions. To this end, this thesis proposes, discusses, and assesses several mechanisms, tailored to diverse wireless network characteristics, which improve the video quality perceived by the end-users in wireless networks while reducing the network overhead.

## 1.2 Objectives and Contributions

The main goals of this thesis are to define a method which enables the characterization of the motion intensity in arbitrary video sequences and also to design, implement, and assess several FEC-based mechanisms with content-awareness to enhance the quality of video delivery from the point-of-view of end-users over diverse types of wireless networks. This is accompanied by a comprehensive analysis of the state of the art of video transmission optimisation mechanisms, with a particular focus on multi-hop wireless networks.

In addition, the proposed mechanisms were evaluated by using the Network Simulator 3 (NS-3) due to the necessity of large-scale networks with a large (and expensive) variety of hardware equipment. The simulator offers a practical feedback allowing the designer to investigate the correctness and efficiency of the proposed mechanisms in a controlled and reproducible environment. Therefore, it is easier to explore the unforeseen interaction between the multiple elements involved in the experiments [Breslau et al., 2000]. In addition, it also allows a straight and fair comparison of results in a transverse manner among several research efforts. Despite the use of a simulator, real video sequences and traces of real network operations and characteristics were used. This provides closer results to real world implementations.

The specific goals of this thesis are as follows:

**Goal 1** - Define a method to provide video content characterization according to its motion intensity, as well as the video parameters;

**Goal 2** - Propose a set of adaptive FEC-based content- and video-aware mechanisms, which aim to support video distribution to wireless users over error-prone networks. These mechanisms should assure the QoE and optimise the usage of the wireless channels resources;

**Goal 3** - Study the performance of the proposed mechanisms in various network
environments. The assessment should always include how the video quality
is perceived by the end-users and also the network overhead impact.

Taking into consideration the specific goals, this thesis has succeeded in produc-
ing the follow main contributions:

**Contribution 1, Assessing the impact of the video characteristics on the
video quality level**

This contribution is related to the study of video sequences details, such
as the frame type and size, the codec type, the length and format of the
group of pictures, as well as the motion vectors, and their impact on the
perceived video quality. All methods and mechanisms proposed in this
thesis, from Chapter 5 to 7, benefit from the findings generated in this
contribution;

**Contribution 2, A method to characterise the motion intensity of videos**

The motion intensity plays an important role on the amount of redund-
ancy needed to improve or even maintaining a high video quality over
networks with high error rates. The proposed method encompasses sev-
eral procedures, such as an exploratory data analysis using a hierarchical
clustering approach to separate and group video sequences with similar
motion intensities. After the clustering process, several techniques can
be applied to provide an easy way to access the cluster's characteristics,
such as a heuristic routine (as seen in Chapter 5), a random neural net-
work (as seen in Chapter 5), and fuzzy logic sets and rules (as seen in
Chapters 5, 6, and 7). The characterisation method is always performed
offline, where a database, sets, and rules are generated to be used in the
real-time mechanisms;

**Contribution 3, Heuristic mechanism**

The proposal, design, and evaluation of an adaptive heuristic-based mech-
anism to enhance video transmissions over wireless networks is one of the
contributions of this thesis. The proposed mechanism uses the human
experience and knowledge of the tackled issues to define a strategy to
improve the video quality without adding unnecessary redundancy, thus
saving the wireless resources. The main objective of heuristic methods is
not to provide a perfect solution, but rather a practical process of dealing
with a problem with satisfactory results. This contribution is studied in
Section 5.2.

**Contribution 4, Random neural networks mechanism**

A different mechanism was proposed using the random neural networks
as means to both define the video content category related to the motion
intensity and to define in real-time an adequate amount of redundancy.
In order to do that, first, the mechanisms have to be trained prior to the
execution. This involves feeding the neural network with a large enough
number of video sequences to perform a numeric categorisation of the
frames in terms of motion intensity. After trained and validated it is
ready to perform the video content characterisation and to attribute an

adaptive redundancy. Section 5.3 presents this contribution.

**Contribution 5, Ant colony optimisation mechanism**

The conceptual definition, implementation, and assessment of mechanisms that uses ant colony optimisation to overcome the network-related problems regarding the video transmission is another contribution of this thesis. These mechanisms rely on ant colony metaheuristic to define a precise amount of redundancy according to the degree of motion intensity, as well as other video characteristics. This contribution is evidenced in Section 5.4.

**Contribution 6, Fuzzy logic mechanism**

Another contribution of this thesis is the design and evaluation of several mechanisms that employ fuzzy logic to decide how much redundancy is needed for each type of video sequence. The definition of the fuzzy logic rules and sets is an offline process that depends on the feedback obtained from Contribution 2. However, once defined, they are ready to be used in real-time. In addition, hierarchical fuzzy logic structures are also designed and assessed. This allows building even-more-complex systems at lower computation costs. This contribution can be found in several mechanisms in Chapters 6 and 7.

## 1.3 Thesis Outline

This thesis is organised as follows.

**Chapter 2 - Network Conceptualization and Feature Highlights**

Introduces the related network concepts and main definitions about several types of networks that are used in the experiments, namely wireless mesh networks, flying ad-hoc networks, and vehicular ad-hoc networks. The usage of the wireless channels and the routing protocol characteristics are also investigated in this chapter, as well as the ways to implement a cross-layer design. In addition, the dynamic nature and the time-varying channels conditions of these networks are explored because they can have a significant impact on delay-sensitive applications such as the video services.

**Chapter 3 - Video Coding Design and Quality of Experience**

Addresses the video-related components and details, such as compression techniques, different types of frames, the group of pictures format, as well as the concept of macroblocks and blocks. In addition, it is also examined how the video sequences are distressed by impairments, the importance of the different types of frames, and the impact of the motion intensity on the quality. At the end of the chapter, quality of experience assessment methods are discussed.

**Chapter 4 - Advances on Improved Video Transmissions**

Discusses the concepts surrounding error-correcting codes. Two types are

investigated, namely the automatic repeat request and the forward error correction, along with the advantages and disadvantages of both. This chapter also shows the literature review on mechanisms to enhance the video transmission quality. At the end, the summary and open issues are presented.

**Chapter 5 - Mechanisms for Resilient Video Transmission over WMNs**

Proposes and assesses three mechanisms to shield video transmissions over wireless mesh networks. All the proposed mechanisms adjust the redundancy amount in real-time according to several video characteristics. The first mechanism (ViewFEC) uses a heuristic method to adjust the redundancy. The second mechanism (neuralFEC) adopts random neural networks to both classify the motion intensity and in the decision-making process. The last mechanism (PredictiveAnts) is based on a random neural network, to categorise motion intensity, and uses ant colony optimisation for the dynamic redundancy allocation.

**Chapter 6 - Mechanisms for Resilient Video Transmission over FANETs**

Proposes and assesses two adaptive video-aware mechanisms to safeguard unmanned aerial vehicle real-time video transmissions against packet loss. The first mechanism (uavFEC) uses motion vectors to define the spatial video complexity and fuzzy logic to adjust the redundancy amount. The second mechanism (MINT-FEC) follows the same procedure, but differs by using additional parameters, such as the temporal video intensity and the ability to cope with videos of arbitrary size.

**Chapter 7 - Mechanisms for Resilient Video Transmission over VANETs**

Proposes and assesses two adjustable mechanisms to increase the video transmission resiliency over vehicular ad-hoc networks. The former mechanism (CORVETTE) uses a hierarchical fuzzy logic model to define a specific amount of redundancy according to the video characteristics and the network conditions. The latter mechanism (SHIELD) adopts a similar procedure for the video details, however, a much more comprehensive method is used to diagnose and evaluate the network conditions.

**Chapter 8 - Conclusions and Future Work**

Presents the final remarks and conclusions, as well as the outline of future research to further advance this work.

# Chapter 2

# Network Conceptualization and Feature Highlights

Big Brother is Watching You !

*(George Orwell, Nineteen eighty-four)*

## Contents

The optimisation of video transmission over wireless networks is a complex task which involves a number of concepts and definitions. In this chapter, several of these components will be discussed, providing a general understanding on this subject. First, the cross-layer design will be described and after that, the characteristics of wireless networks will be discussed.

## 2.1 Introduction

Wireless networks have become a cost-effective way to distribute variate contents. Another advantage is the untethered convenience allowing a good range of mobility to the end-users, as well as the self-management and multi-hopping features. Because of that, the widespread deployment of wireless networks, as well as the usage of mobile devices, has increased significantly in recent years. These

networks are well known for their dynamic nature along with their time-varying channel conditions.

Apart from that, several new services have stringent requirements and work in real-time. This situation is a cause of great concern when it comes to video transmission, which can be considered to be one of the most demanding services. Environments that do not provide the necessary conditions will have a poor performance in this type of service and the video impairments will be very noticeable to the end-users. In this case, the intrinsic network details must be exploited in order to understand how to provide the necessary conditions to better serve these resource-demanding services.

The remainder of this chapter presents the concepts of cross-layer design in Section 2.2 and the main characteristics of the wireless channels in Section 2.3. This is followed by Section 2.4 which describes wireless ad-hoc networks. The chapter's summary is given in Section 2.5.

## 2.2 Cross-layer Design

The purpose of a layering principle is to reduce the constraints and the complexity of the design. The Open Systems Interconnection (OSI) model [Day and Zimmermann, 1983] proposes seven conceptual layers, where each one is only able to communicate and provide information to the adjacent layers. This means that non-adjacent layers are not allowed to exchange information. The ever-increasing network traffic, along with the emergence of novel services, e.g., video transmission which has stringent requirements, means that it is a very challenging task to meet the end-to-end requirements without interaction between non-adjacent protocol layers [Setton et al., 2005].

The cross-layer design was proposed to overcome this limitation. It can be defined as the violation of the referenced layered communication architecture through direct communication or the sharing of information between non-adjacent layers. There are several ways that cross-layer design can be carried out, for example, by creating new layer interfaces and allowing direct communication between non-adjacent layers. An alternative approach could be to perform a vertical calibration across layers, where the values to set parameters could span across non-adjacent layers. Merging adjacent layers using a joint design can also be considered to be cross-layer [Srivastava and Motani, 2005].

Figure 2.1 (adapted from [Lindeberg et al., 2011]) depicts several examples of information exchange using a cross-layer design. For instance, from the application layer, it is possible to set packet priorities directly to the link layer or to define the delay and bandwidth requirements in the routing layer. Furthermore, as well as defining information, it is also possible to have access to other's layers data. For example, the application layer can have information about the offered load from the transport layer or the number of available routers, bandwidth, and packet losses from the routing layer.
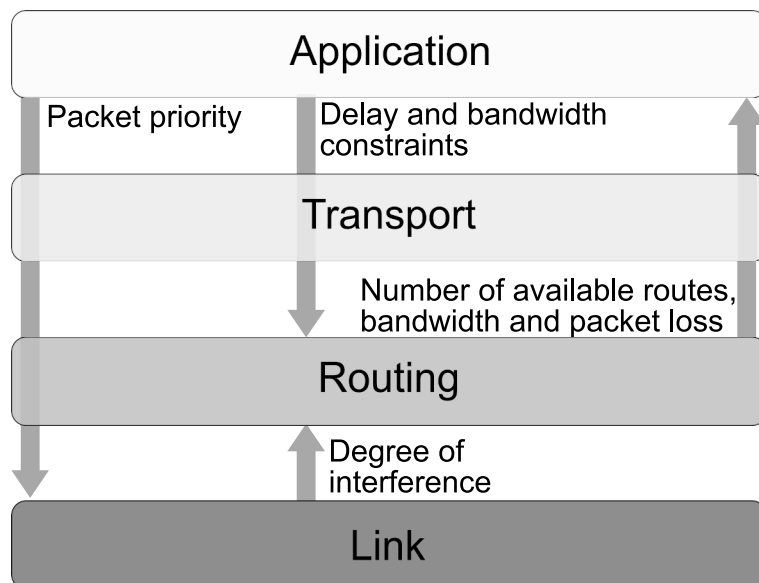
Figure 2.1: Cross-layer information exchanges

Since the use of this technique is widespread, several studies have been proposed for addressing the question of cross-layer design, and some of them have achieved promising results in improving the system performance of wireless networks [Wang et al., 2010, Oh and Chen, 2010, Liu et al., 2008, Andreopoulos et al., 2006]. However, a number of problems may arise, e.g. incompatibility with existing protocols, unclear protocol-layer abstractions, increased design complexity, difficulty in maintenance and management, among others.

Overall, the cross-layer techniques are a flexible means of both accessing and defining information in non-adjacent protocol layers. This is especially important when assessing the end-user perception of video quality because the use of this information provides a better chance of interpreting the general user perception of the service correctly. Nevertheless, it is important to take into account the relevance of compatibility and standardisation, as otherwise the interoperability of the solution will be impaired.

## 2.3 Wireless Channels

A channel can be considered a medium to convey a data signal from the sender to the receiver. While the nodes in a wired network transmit information using links that have static and well-known characteristics, wireless nodes communicate via a much more dynamic and unpredictable environment. A radio transmits electromagnetic waves through the wireless channels over a non-uniform and unknown medium, where the properties of each scenario might be radically different and oscillate over time [Adlakha et al., 2007]. These properties can be affected by surrounding objects, communication between other nodes and even themselves, as well as radio frequency (RF) interference from other equipment. This means that, as the signal travels, it is subjected to a wide range of physical

effects that can cause a deterioration of the quality.

These abnormalities sometimes occur as a result of events like multipath fading and shadowing which lead to a decrease of signal quality and thus tend to increase packet error rates. Multipath fading is the consequence of more than one interfering reflections of the same signal. Shadowing occurs when large objects block (or partially block) the propagation path, causing severe signal attenuation [Lindeberg et al., 2011].

In addition to the fact that a shared medium is used, the radio antennas are unable to sense and transmit simultaneously, which makes matters worse. Another known issue is the hidden terminal problem, which happens when a node is visible from one Access Point (AP), although not from the other nodes connected to the same AP. It is also possible to name the exposed node issue that occurs whenever a node is kept from transmitting packets due to transmissions from other neighbouring nodes. These constraints increase the packet loss rates even further, as a result of a higher number of collisions. This situation becomes even worse in high node density networks. Moreover, as the distance between the nodes increases, the signal strength gradually decreases.

Thus, owing to the time-varying link characteristics and the unpredictable nature of wireless channels, mechanisms are needed to enhance the transmission. This is especially true in view of the stringent delay, jitter, and packet loss requirements imposed by video transmission services.

## 2.4 Wireless Ad-Hoc Networks (WANETs)

A WANET is composed of a collection of wireless nodes which do not rely on a pre-existing infrastructure. This is a self-organized and decentralised type of network where the communications are performed by multi-hop wireless paths. Additionally, these networks can be spontaneously created as the nodes dynamically choose where to forward the data according to the network connectivity. On top of that, due to its quick deployment and a reduced configuration, they are very convenient for several types of applications.

The ascendance of video services, however, has created the need for a stable and trustworthy communication link between the nodes in order to guarantee a high level of QoS and QoE support. Considering that WANETs are highly dynamic, have near-unpredictable node mobility, and suffer from time-varying network conditions, the QoE and QoS provisioning is very complex as well as challenging. Because of that, new methods and mechanisms have to be proposed to address these issues, being this one of the main goals of this thesis.

According to its application, WANETs can be further divided into more specialised categories. The ones related to this work are presented below.

## 2.4.1 Wireless Mesh Networks (WMNs)

WMNs have been used as a possible means of extending the reach of wireless APs, proving to be a viable solution for spreading reliable wireless Internet across wide areas [Pal and Nasipuri, 2011]. This type of network is formed by a set of mesh routers, generally static, which use multi-hop transmissions to create a backbone network, allowing, in this way, communication between mesh clients [Akyildiz and Wang, 2008]. A WMN does not have a fixed structure or a defined and immutable path to convey the information because these networks rely on dynamic paths which can be changed and discovered through the routing protocol in real-time. In doing that, it is possible to assume that every router is responsible for maintaining the information flow between other routers in the neighbourhood.

The mesh client, in these networks, can be any type of device that has some kind of wireless connectivity, namely a notebook, a smartphone, or a tablet. Usually, one or more mesh routers are connected to the Internet acting as gateways to the clients. In general, WMNs incur low installation and maintenance costs and offer a reliable service due to their dynamic self-configuration and self-organization capabilities [Akyildiz and Wang, 2005]. This type of network has been adopted for many industrial and academic deployments, in small and large offices, as well as in home applications. This means that this technology can be used in any place that does not have a wired local area network (LAN) or has high-deployment costs, e.g., instant surveillance systems, and back-haul service for large-scale wireless sensor networks [Kim and Shin, 2006].

Recently, these multi-hop wireless networks have come to be regarded as the most important wireless technology to provide last mile access in future wireless multimedia networks. WMNs will allow thousands of fixed and mobile users to access, produce and share multimedia content in an ubiquitous way. These multi-hop communications are expected to meet quality requirements, namely QoS and QoE, to achieve this goal. They are also expected to ensure easy deployment, flexibility, and high reliability [Bruno et al., 2005, Zwinkels, 2004].

**Routing Protocols and Metrics**

Due to the dynamic nature of the WMNs, a routing protocol is needed to find and maintain connectivity between the nodes, allowing it to achieve the best possible overall performance. In other words, the routing protocol can be described as the process to determine the end-to-end path that connects the source node to the destination node [Waharte et al., 2006]. Another role of this type of protocol is to share the routing information with the network nodes as a means of providing the knowledge required to build the network topology.

There is a wide range of routing protocols, but generally, the approaches can be divided into proactive, reactive, or hybrid categories. In proactive routing protocols, there is a periodic exchange of routing messages to keep the routing tables updated, even when there is no network traffic, and hence a waste of

resources. Reactive or on-demand routing protocols were developed to serve as an alternative to the constant message exchanges. These protocols discover routes upon request avoiding the waste of resources. The problem arising from this approach is that there is a higher initial delay when retrieving a routing path, and depending on the service type, this may not be acceptable. Lastly, there are hybrid routing protocols which exploit the advantages of both proactive and reactive protocols [Ancillotti et al., 2011].

The routing protocols can use several metrics to find the best end-to-end path. Hop count is the simplest metric, where only the number of hops between the source and destination are considered. A different performance metric is known as Expected Transmission Count (ETX) [De Couto et al., 2005], which takes into account the data loss and the number of retransmissions. Another metric is the Expected Transmission Time (ETT) [Draves et al., 2004], which is an improvement of the ETX metric that includes the link bandwidth to calculate the route path [Parissidis et al., 2011]. A few more examples of metrics used in other studies include the location-dependent contention, such as the Contention-Aware Transmission Time (CATT) [Genetzakis and Siris, 2008], Metric of Interference and Channel-switching (MIC) [Yang et al., 2006b], the number of per-link admitted flows e.g., Load Aware ETT (LAETT) [Aiache et al., 2008], interference and load through passive monitoring in Metric for INterference and channel Diversity (MIND) [Borges et al., 2009], as well as the load-dependent cost, such as Weighted Cumulative Expected Transmission Time with Load Balancing (WCETT-LB) [Ma and Denko, 2007].

Furthermore, QoS-based routing is also desirable due to the recent improvements in mobile devices as well as the emergence of real-time video services. In this approach, the end-to-end paths are determined by using information about the resource availability in the network in conjunction with the QoS requirements of the flows. Additionally, the routing protocol is able to find a path that satisfies a specific flow requirement, increasing the network resource utilisation. An important feature of this approach is that it has a collision detection mechanism which enables it to avoid disruptive interferences [Luo and Ephremides, 2012, Bouhouch and El Fatmi, 2007].

Another important feature of some WMN routing protocols is the support for multipath routing. These routing algorithms take advantage of the multiple connections between the mesh nodes to provide more than one path between source-destination pairs [Lindeberg et al., 2011]. One advantage of multipath routing approaches is that they can provide alternative paths according to the required QoS, without a frequent route discovery [Yang et al., 2006a]. A further benefit is the ability to transmit information, from the source node, through a number of disjoint paths, to the destination node, by enabling bandwidth aggregation, load balancing, and reduced delay as well as improving fault tolerance. The fault tolerance feature is especially important for real-time transmission because when a high degree of multipath diversity is achieved, it is less feasible that a link disruption will occur in more than one path at the same time, thus making the chance of a packet loss less likely [Tsirigos and Haas, 2004]. Therefore, if this happens, it will reduce the impact of a single link failure on the video

quality.

## 2.4.2 Vehicular Ad-Hoc Networks (VANETs)

VANETs enable vehicles to create and maintain wireless communication links between them in the absence of a central base station. One of the main goals of this technology is to provide ubiquitous connectivity as well as to allow efficient vehicle-to-vehicle communications, which is necessary to implement an Intelligent Transportation Systems (ITS) [Al-Sultan et al., 2014]. This network uses the IEEE 802.11p Wireless Access for Vehicular Environments (WAVE) [Jiang and Delgrossi, 2008] standard and is envisioned to be used in a multitude of environments, such as in highways as well as in rural and urban scenarios, including support for many applications.

A typical VANET is composed of several components, as described below:

**On board unit (OBU)** This is an onboard device inside the vehicles to provide connectivity with other OBUs or Roadside units (RSUs). Usually, it is equipped with a specialised interface for ad-hoc communication and short-range wireless capabilities using IEEE WAVE. The OBU is also responsible for providing data forward to and from both others OBUs or RSUs.

**Roadside unit (RSU)** This component is a WAVE-enabled device typically fixed that is placed along the roadside or in predetermined locations. The main goal of this device is to extend de transmission range of other RSUs or OBUs. In order to do that, this device is usually equipped with a WAVE radio and more often than not with another communication link directly to the infrastructural network. Because of that, it is also possible to provide Internet connectivity to passing by OBUs.

**Application unit (AU)** This element stores and acts as interface for the applications and/or services supported by the provider. This element can be a dedicated hardware or be implemented using the same physical components of the OBU. This means that the difference between the OBU and AU could be only software-based, i.e., applications and interface. Nevertheless, the network functions are still the OBU responsibility.

The above-described components allow distinct types of communication setup as briefly described below:

**Vehicle-to-Vehicle (V2V)** This is an ad-hoc network allowing the direct communication between the vehicles, e.g., OBU to OBU. In this format, there is no need for a fixed or pre-existing infrastructure. The main applications of V2V are safety information, warning functions, security, co-operative assistance, as well as the dissemination of application-related data.

**Vehicle-to-Infrastructure (V2I)** In this type of communication the vehicles exchange information directly with the roadside infrastructure, e.g., OBU to RSU. This model can offer a one-hop connection to the Internet as well as a high bandwidth link.

**Hybrid architecture** This design allows both V2V and V2I communication. In doing this, it enables multi-hop transmissions between vehicles and/or the roadside infrastructure. For example, vehicles that are far away from each other can use one or multiples RSUs to maintain contact. Another possible scenario to be considered is vehicles distant from an RSU that can use another vehicle's OBU as a hop to extend the delivery range, thus reaching the closest RSU.

**Vehicle-to-Everything (V2X)** It adopts a broader approach than Hybrid architectures by allowing the communication between the vehicles and any other entity, such as V2V, V2I, Vehicle-to-Grid (V2G), Vehicle-to-Pedestrian (V2P), and Vehicle-to-Device (V2D). In other words, it comprises each and every message exchange in the network. It works in ad-hoc mode and therefore it does not require a pre-existing infrastructure, which is an advantage in little developed or remote areas.

There are some key characteristics of this model that are worth mentioning. For example, due to the high mobility of the nodes, this network environment has a very dynamic topology. Another characteristic related to the high node mobility is the frequent disconnections. These characteristics can be explained by the speed of the vehicles, which can be high, especially on highways. In addition, the vehicles may also be moving in opposite directions. This means that as quickly as they can join a network they can also leave, leading to the mentioned dynamic topology and the frequent disconnections.

Other aspects of VANETs are the predictable mobility patterns and the absence of power constraints. Despite the very dynamic topology, due to the mobility patterns, it is feasible to predict the future position of the vehicles as they move on pre-defined roadways. Furthermore, they have to respect the speed limit, traffic lights and signals, as well as being restricted to traffic conditions. An additional feature that can bring advantages is the power supply. Vehicles are equipped with long-life batteries which can be recharged as they move. This allows processing more complex and computationally intensive applications.

Together with useful services and applications, VANETs also present new challenges that have to be addressed to ensure a good QoE. For example, in the ad-hoc mode, the lack of an infrastructure transfers the responsibilities to the vehicles. In doing that, they are responsible not only for their communication but also for the message forwarding of other vehicles or entities associated with that network. It is also possible to highlight several other challenges such as signal fading and the small effective diameter of the network, incurring in a weak connectivity between the nodes, as well as bandwidth limitations.

**Routing Protocols and Metrics**

Due to the dynamic nature of VANETs, which have highly mobile nodes and suffer from rapid topology changes, the design of an efficient routing protocol is very challenging. This protocol should allow delivering the packets in the minimum period of time with the smallest packet loss rate (PLR) possible. They

are also responsible for avoiding transmission conflicts and reducing the degree of interference.  The routing protocols can be classified into several categories, such as topology-driven routing, position-based routing, cluster-based routing, and geocast routing, just to name a few.

The topology-driven or ad-hoc routing uses several known mobile ad-hoc network (MANET) protocols to create and maintain the delivery paths, such as Ad-hoc On-demand Distance Vector (AODV) [Perkins et al., 2003] and Dynamic Source Routing (DSR) [Johnson and Maltz, 1996].  However, these protocols are designed for general purpose ad-hoc networks and tend to have poor performance in VANETs [Liu et al., 2004].

A different approach is taken by the position-based routing, where information about the neighbouring environment and the vehicles well-defined mobility patterns can be employed allowing the selection of better routing paths [Fonseca and Vazão, 2013].  The Geographic Source Routing (GSR) [Lochert et al., 2003] is one example of a position-based protocol.  It uses the city topology, with aid of a pre-loaded map, and geographic routing to determine the best nodes to receive and forward the messages.  Another proposal is the Anchor-based Street and Traffic Aware Routing (A-STAR) [Seet et al., 2004].  This protocol is similar to GSR but differs by adding traffic awareness, through statistically and dynamically rated maps, in the path selection.  In doing this, it tries to solve the issue of unevenly distributed vehicles which tend to concentrate more on some roads than other.

It is also possible to use cluster-based routing in VANETs.  The same method used in MANETs is adopted here, where the cluster members can perform intra-cluster communication and the cluster head is responsible for both intra- and inter-cluster message exchange.  However, due to the fast changing topology and high-speed vehicles, the clustering techniques used in MANETs are not appropriated for VANETs.  To address this issue, the Trust dependent Ant Colony Routing (TACR) [Sahoo et al., 2012] uses the real-time position and an ant colony routing technique based on a trust value for each vehicle.  In addition, it also considers the direction and the relative speed in this process.  This allows it to better define the cluster members, providing higher scalability and fewer transmissions failures.

A further example of VANETs routing methods is the geocast protocols.  The aim of these protocols is to deliver the packets in a form of a location-based multicast [Singh and Agrawal, 2014, Kaiwartya and Kumar, 2014].  This means that all nodes within a specific geographical region are able to receive these messages.  One example is the Guaranteed Geocast Routing (GGR) protocol [Kaiwartya and Kumar, 2015].  It uses acknowledgements to ensure the one-hop deliveries as well as a heuristic procedure to select the next hop vehicle.

## 2.4.3 Flying Ad-Hoc Networks (FANETs)

This network model can be defined as a form of MANET, and a subset of VANET, where the nodes are UAVs [Bekmezci et al., 2013]. In the same ways as in the VANETs, the communications links can be delineated as UAV-to-UAV, UAV-to-Ground, and hybrid. However, there are several differences between this network and other ad-hoc networks. Firstly, the node mobility has very specific characteristics, for example, it is much faster than in VANETs. In addition, it also has a higher degree of movement, which means that the node positions are far less predictable. Even in autonomous mode, the UAVs are constantly correcting and adjusting the predefined flight plan [Sahingoz, 2014]. Another difference is that these nodes are capable of making fast and sharp turns, changing direction very easily, which has to be considered in the mobility models. Apart from that, the node density is lower than in other ad-hoc networks. This happens because the UAVs are scattered in the sky and can be positioned far away from each other.

Besides the node mobility and density, the radio propagation model is also peculiar to these networks. In VANETs the nodes are just above the ground and in several cases, there is no line-of-sight communication between them. Because of that, the radio signals are disturbed by the terrain features and geographical structures. On the other hand, in FANETs the nodes tend to be far away from the ground allowing line-of-sight communications, not being directly impacted by the ground structures and the terrain. However, this creates a different source of interference, namely the ground reflection effect. Another relevant issue in UAVs is the environmental weather conditions.

Furthermore, the UAV platform offers small energy constraints and the possibility of high computation power. Modern UAVs are equipped with high-capacity batteries that can provide energy for an extended period of time. However, this can be an issue with small form factor UAVs. In addition, UAVs are outfitted with state-of-the-art processors which offer a considerable high computation power. Both of these features are only constrained by the size and weight of the desired hardware. Lighter UAVs have extended fly range and additional payload capacity, which can be used to deploy supplementary sensors or any other necessary peripherals.

### Routing Protocols and Metrics

Several routing protocols have been proposed for MANETs and VANETs networks, however, due to the specific issues that have to be addressed these routing strategies are not ideal. Because of the fast pace and high mobility nodes in FANETs, a resourceful routing protocol is needed to provide the best possible communication channels between the UAVs.

One example of a position-based routing protocol for FANETs is the Geographic Position Mobility Oriented Routing (GPMOR) [Lin et al., 2012b]. This protocol uses the movement prediction of UAVs to choose the next hop and adjust the

routing tables. Another example is the Directional Optimized Link State Routing Protocol (DOLSR) [Alshabtat et al., 2010]. It uses OLSR as base, modifying, however, the procedure of choosing the multipoint relay nodes through the use of directional antennas, reducing the end-to-end latency.

Another example of FANET routing protocols is the Greedy Perimeter Stateless Routing (GPSR) [Karp and Kung, 2000]. This protocol uses only the position of routers and the packet destination to construct the routing tables. In doing that, it provides very good scalability in densely deployed networks. Furthermore, hierarchical protocols also tackle scalability issues. The UAVs are divided into groups and each cluster head communicates directly with an upper layer UAV or a satellite. One of the main issues with this approach is how to predict the mobility patterns of the UAVs and the clusters. The Mobility Prediction Clustering Algorithm (MPCA) [Zang and Zang, 2011] uses a dictionary based on a digital tree algorithm and the link expiration timeout to minimise this issue.

A different approach is adopted by Data-centric routing algorithms. These models can be adapted to FANETs and usually have a publish-subscribe model type as communication architecture [Erman et al., 2008]. The UAVs that are producing data (publishers) are connected directly with other UAVs or the ground station that wants to consume the data (subscribers). Because of that, the routing algorithm creates the tables based on the data content. The use of data aggregation algorithms is desirable to provide energy-efficient content dissemination. The Autonomous self-deploying and operation of Wireless sensor-actuator networks cooperating with AeRial objEcts (AWARE) [Gil et al., 2007] is one example of this approach.

## 2.5 Summary

In this chapter important components of wireless networks are discussed, including cross-layer design, wireless channels, wireless ad-hoc networks, and routing protocol characteristics, First of all, the concept of cross-layer implementation was presented in Section 2.2. This technique can be employed to overcome the restricted form of communication of the network layers. Because of that, it can be considered a violation of the referenced layered communication to provide more flexible mechanisms to improve the network transmissions.

Section 2.3 outlined the wireless channels characteristics, such as the time-varying link properties and the unpredictable communication quality. In addition, the common issues found in these networks, such as shadowing, multipath fading, and signal attenuation were also pointed out.

Furthermore, Section 2.4 detailed the WANETs characteristics and presented the specialised network categories related to this work, namely WMNs, VANETs, and FANETs. The WMNs are commonly used to extend the reach of wireless access points by creating a multi-hop network. There is no fixed structure in these networks, therefore, they are easy to expand and reliable, as they do not

have a single point of failure. VANETs enable ubiquitous connectivity between the network members. The main applications are related to transport efficiency and information/entertainment services. The last network category, FANETs, are responsible for providing reliable communication channels for flying nodes or UAVs. They are a subset of VANETs having several similar characteristics but differ in regards to the node mobility and density, as well as in the radio propagation model.

# Chapter 3

# Video Coding Design and Quality of Experience

His thin dark face had become animated, his eyes had lost their mocking expression and grown almost dreamy. - It's a beautiful thing, the destruction of words.

*(George Orwell, Nineteen eighty-four)*

## Contents

R eal-time video services are becoming a large part of the daily routines of people all over the world. They have been used to spread information ranging from education to entertainment content. In addition, this technology has been used by many companies as a part of a business drive as well as by non-professional users, resulting in a considerable growth of online videos consumption. In this chapter, several details about the video format and the quality of experience will be explored to provide an overall awareness on this subject.

## 3.1 Introduction

The rapid growth of real-time video services is evident in recent years, particularly from wireless mobile devices [Adobe Digital Index, 2016]. This leap is related to the technological advances in broader connectivity and the widespread adoption of smart devices, as well as the rise in popularity of this type of content. Many companies have used video streaming as part of a business drive to increase productivity, improve collaboration, reduce costs, and streamline and optimise business operations. Following the same trend, non-professional users are producing, sharing and accessing thousands of videos by using both wired and wireless systems [Cisco, 2016].

The video quality of these services can be impacted by several factors, including the video characteristics. The type of codec, the streaming bitrate, the frame type, as well as the format and the length of the Group of Pictures (GoP), can alter how the video quality is perceived. The video content also plays an important role on the perceived quality. Videos with a small degree of movement and few details tend to be more resilient to packet loss, keeping the quality high. In contrast, videos with high levels of details and movement are more susceptible to losses and the flaws will be more noticeable [Khan et al., 2010].

In order to assess the video integrity, Quality of Experience (QoE) methods are desirable. These metrics evaluate the level of the video impairments and can be defined in terms of how users perceive the quality of an application or service [Piamrat et al., 2009]. In other words, the QoE metrics assess the video quality considering the end-users point-of-view.

The remainder of this chapter features the main concepts and definitions needed to understand the characteristics of video formats (Section 3.2). This is followed by the examination of how video is distressed by impairments in Section 3.3. Section 3.4 presents an analysis of the metrics and standards used for video quality assessment. At the end, Section 3.5 gives the summary.

## 3.2 The Video Format

Video technology is used to record, display, and broadcast multimedia content through electronic media. In a simple way, a video flow is composed of a sequence of temporal-related images arranged to create one fluid moving picture. Video coding formats, or standards, aim to create efficient digital video archives. These documents describe how the pictures have to be stored and transmitted. Examples of video standards include Theora [Xiph.Org, 2011], MPEG-4 Part 10 (H.264) [ITU-T H. 264/AVC, 2005], and High Efficiency Video Coding (HEVC) [Sullivan et al., 2012]. On the other hand, the implementation of these formats is called codec, such as x264, OpenH264, and Xvid.

### 3.2.1 Video Compression

In the video technology, the compression or coding is one of the main processes [Richardson, 2004]. It is responsible for converting the video signal so it takes less storage space and consequently uses less bandwidth during the transmission. The compression process can be in both lossy and lossless format. In the lossless compression, the original data can be perfectly reconstructed, which does not allow high compression rates. On the other hand, in the lossy compression, several video details can be suppressed, achieving higher compression rates. However, with the increase in the compression ratio, there is a decrease in the video quality.

The video compression is performed by excluding some types of redundant information. Usually, compression algorithms exploit three main scenarios, namely spatial (intra-frame), temporal (inter-frame), and perceptual redundancy. The first two are also called statistical redundancy and can be used in lossless compression because all the information is only reorganised and not dropped. The spatial redundancy compression takes advantage of the significant correlation between neighbouring pixels in a given area of the frame. All the pixels with the same characteristics are represented by a pointer, reducing the amount of information that needs to be stored. The temporal redundancy compression capitalises on the statistical correlation among pixels in subsequent video frames. In the same way as in the spatial compression, a pointer can be used to represent the characteristics of a block of pixels that has motion attributes, however, with static stored values.

The last one (perceptual redundancy) is only used in lossy compression because it requires dropping non-essential details. This compression is based on the fact that the Human Visual System (HVS) does not perceive all the optical information equally. Just to give one example, the human eye has a better perception of luminance variation than to colours deviations. These small details, which cannot be clearly recognised, can be neglected without incurring in a perceptual alteration of the video quality.

### 3.2.2 Coding Elements

The coding format is another component of the video technology. The H.264 is, to date, the most commonly used format of video content [Ma et al., 2016]. Figure 3.1 depicts an overview of the H.264 structure and components. The video sequence is composed of frames with different types (1), which are, in turn, divided in the GoP (2). Each frame (3) has its own type and can be divided into slices (4). The slices can hold one or more macroblocks (5) that can be broken down in blocks (6). These components are explained below.

A video frame can be defined as an amount of data after compression. There are three major types of frames: Intra-coded frames (I-Frames), Predictive-coded frames (P-Frames), and Bi-directionally predictive-coded frames (B-Frames). The compression in I-Frames is carried out by reducing the spatial redundancy
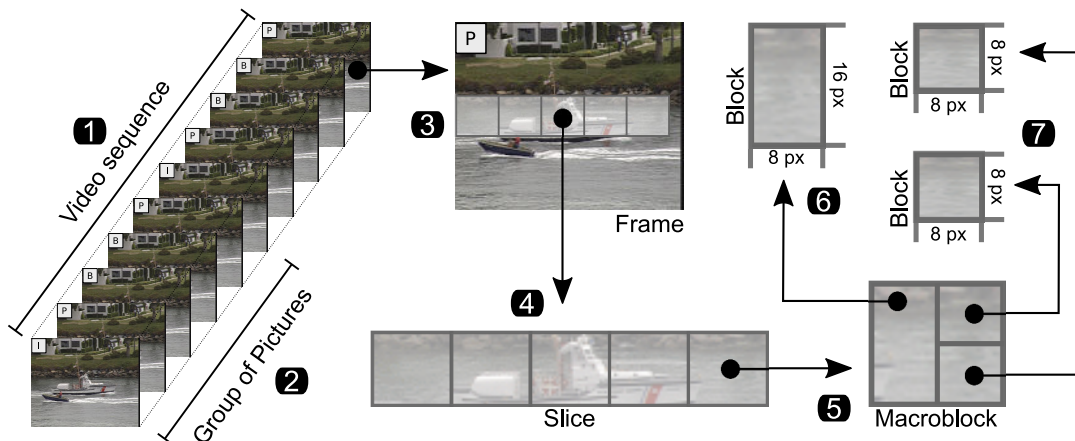
Figure 3.1: H.264 standard video structures

only in the current raw frame. Such compression takes advantage of the human eye inability to detect certain minor changes in the image. This frame is like a conventional static image and because of this, these frames are reconstructed independently. Because of that, as soon as the information that this frame holds is needed, it can be displayed without having to consult other frames. On the other hand, P- and B-Frames only enclose segments of the image information, reduce the coding data, and thus improve the video compression rates. On the negative side, these frames need other I- or P-Frames to be decoded, respectively.

A P-Frame only contains the changes of the actual image when compared with the previous I- or P-Frame. With this type of prediction, less coding data is required ($\approx 50\%$ less in contrast to I-Frame size). B-Frames also use prediction to reduce the amount of coding data, but unlike the P-Frame, they use both the previous and following frames to determine the content. The size of a B-Frame is usually $\approx 25\%$ smaller than a P-Frame [Wiegand et al., 2003]. In both types of frame the data is expressed as motion vectors and transforms coefficients which are used in prediction correction and motion compensation.

GoP are sequences of frames grouped together. The MPEG standard uses a hierarchically-structured GoP. This structure is a full sequence, which means that it contains all the information that is needed to decode the video images, within that period of time. Therefore, it enables random access into that portion of the video. Figure 3.2 shows an example of a GoP structure, where $M$ represents the distance between successive P-Frames and $N$ defines the separation between the adjacent I-Frames. The GoP frame ratio is $N : M$, in the example, the ratio is 9:3. This means that it contains one I-Frame at every 9 frames and one P-Frame after each three B-Frames. This structure is flexible, and both the frame types and their location within a GoP are adjusted in accordance with the encoding format desired.

A slice is another MPEG structure, which are special regions defined during the coding process. Each slice represents a spatially distinct region inside a frame. In view of this, everything within the slice's area will be encoded inde-
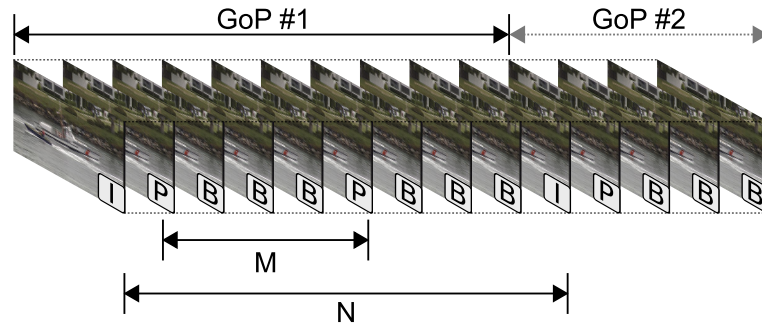
Figure 3.2: GoP structure

pendently from the remaining of the frame. This gives the opportunity to the codec to decide to use a distinct amount of compression inside the same frame. In other words, a slice is composed of an arbitrary and successive number of macroblocks.

The macroblocks are fundamental processing units for motion compensation. Each one can be associated with a unique motion vector. Due to the hierarchical MPEG structure, each frame type has a predefined way to handle motion compensation. For example, I-Frames can only have intra-coded macroblocks. This means that it has to be self-contained. On the other hand, P-frames can have both intra-coded or predicted motion compensation. Finally, B-Frames can have intra-coded, predicted, and/or bi-predicted motion compensation macroblocks.

Blocks are the smallest coding unit defined by the MPEG format. This is the most basic element in frame coding. Usually the frames are segmented into blocks of 8x8 pixels, however, the H.264 standard allows blocks with 4x4, 16x8, and 16x16 pixels, which can be defined per-macroblock basis [ITU-T H. 264/AVC, 2005].

## 3.3 Video Impairments

In general, video impairments tend to be transient, thus it is difficult to detect and act upon them. Consequently, this is a hard problem to solve in a satisfactory way, especially in wireless networks. The situation is aggravated by the fact that video transmission is more sensitive to impairments than data distribution. This can lead to a serious perceived quality degradation even with a low rate of packet loss. Apart from this, not all packets have an equal impact on the perceptual video quality. Each packet, or group of packets, can carry different types of information about the frames that are being transmitted, and some are more important than others. This originates from the video frame arrangement carried out by the encoding/compressing process. Besides that, subjective factors also play an important role; for example, the motion and complexity levels, as well as the amount of detail in the video sequence can determine how the packet loss is perceived.

Additionally, the encoding and compression techniques performed by the codec will have an influence on the video quality as well. The basic principle of the compression techniques is to reduce the amount of information that is stored and/or transmitted. There are several ways to achieve this reduction, such as by limiting the redundant data, discarding less important information, and using prediction schemes [Seeling et al., 2004]. With higher compression, less data needs to be sent through the transmission channel, however, any packet loss that may happen will have a greater impact on the video quality.

Another important trait to be considered is the video content. Studies have shown that videos with slight movement have better packet loss resilience, whereas videos with rapid movement are less resilient and the flaws generated are more noticeable. Generally, the video content can be classified into three categories according to its motion intensity [Lotfallah et al., 2006]. The first one is low intensity, which includes sequences with a small moving region of interest, generally someone's face and traditionally on a static background. This type of video sequence can handle up to 20% of packet loss and still have an acceptable video quality. The second classification, medium intensity, contains sequences where adjacent scenes are modified. This category usually has frames with higher image details. The packet loss natively supported is up to 10%. Finally the last classification, high motion intensity, includes wide-angle image sequence and a lot of movement, as usually the entire picture is moving. This classification is more sensitive to packet losses, thus it can handle up to 6%. Table 3.1 summarises the relationship between the video content, motion intensity, and natural resilience to packet loss [Khan et al., 2009].

Table 3.1: Impact of packet loss on different motion intensities

| Motion intensity | Video characteristic | Packet loss |
|---|---|---|
| Low | small moving region | up to 20% |
| Medium | higher movements and image details | up to 10% |
| High | wide-angles and non-static backgrounds | up to 6% |

Furthermore, the MPEG's hierarchically-structured GoP also has an influence on how the packet loss will be perceived. Because of its ranked order used to organise the frames, which allows reducing the file size, some pieces of information are more important than others. As a result, the packet losses may affect the video quality in different ways, depending on the information that was lost. Figure 3.3 depicts three hypothetical scenarios. In Scenario (A), the network drops packets within a B-Frame. In this case, just this particular frame will be damaged, because it needs information from other frames, but none of the others needs its information. Scenario (B) describes dropped packets in a P-Frame. This will produce impairments that are extended through the rest of the GoP. This means that the error will only be corrected with the arrival of an I-Frame. Finally, Scenario (C) shows one I-Frame being damaged by packet losses. In this case, the impairment will spread throughout the remainder of the GoP. This happens because all other frames need the information carried by the I-Frame, and thus the video quality will only improve when the decoder receives

the next I-Frame.
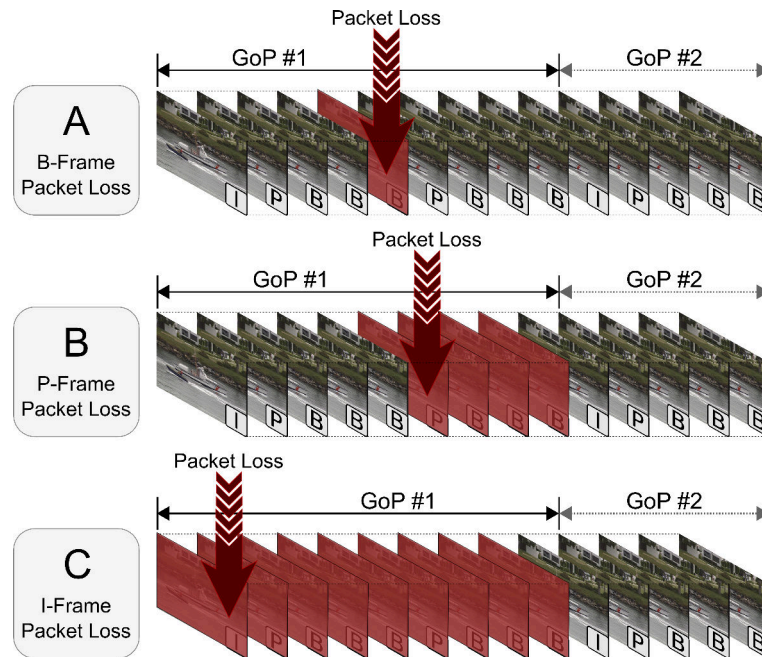


Figure 3.3: Packet loss repercussion on the GoP

## 3.4 Video Quality Assessment

The video quality measurement can be defined as an estimation of the video degradation. Several types of artifacts and/or distortions can be introduced during the encoding, decoding, and transmission process of video content, which can negatively impact the end-user perception of quality. It is not difficult for human subjects to point out video defects, especially if they are happening over an extended period of time. On the other hand, it is not a straightforward task to quantify these impairments. Therefore, QoE metrics are used to provide support for this task.

Quality of Experience (QoE) can be defined as a measure of customer's personal observation of a service, product, or application. QoE is related to Quality of Service (QoS) metrics, which measure the network operation conditions, such as crosstalk, noise, lost or dropped packets, however differs from it because adopts a holistic approach assessing the end-to-end connection as well as the final user requirements or expectations. It is possible to say that QoE extends the concept of QoS by encompassing additional factors, aiming to provide an appraisal of human expectations in respect to a particular activity.

In regards to the QoE, there are two types of metrics that can be used to quantify video impairments: subjective and objective. Objective metrics use mathematical models with a set of indicators that are correlated with the user's perception of quality. These metrics obviate the need for human intervention and can have instant results. Subjective evaluations use human individuals to

assess the video quality. These metrics are able to capture all the details that might affect the user experience, however, they need long rating process, are expensive, and the results are not reproducible.

### 3.4.1 Subjective metrics

One of the most widely used approaches for subjective video evaluation is the Mean Opinion Score (MOS) [ITU-R P.910, 1999, ITU-R BT.500-13, 2012]. Using this technique, it is possible to capture the full degree of subjectivity that end-users adopt to perceive the impairments in the video sequence. MOS is recommended by the ITU-R and uses a group of people voting in video sequences, according to a predefined quality scale, to rate the quality. The MOS scale goes from 1 to 5, where 5 is the best possible score, as shown in Table 3.2.

Table 3.2: Mean Opinion Score scale

| MOS | ACR | Impairment |
|-----|-----|------------|
| 5 | Excellent | Imperceptible |
| 4 | Good | Perceptible but not annoying |
| 3 | Fair | Slightly annoying |
| 2 | Poor | Annoying |
| 1 | Bad | Very annoying |

The recommendation ITU-R BT.500-13 [ITU-R BT.500-13, 2012] describes several subjective quality test methods, such as Absolute Category Rating (ACR), Degradation Category Rating (DCR), and Double-Stimulus Continuous Quality-Scale (DSCQS). In the ACR method, the video sequences to be ranked are presented one by one. The viewers watch the sequences only once and then rate the quality using the scale shown in Table 3.2. The ACR scale is related to the MOS scale. This is considered one of the most suitable methods for conducting the quality assessment of video applications [Seshadrinathan et al., 2010].

In the DCR method, the video sequences are displayed in pairs. The first one is the original video, and the second is the impaired sequence which is a result of the experiments. In this procedure, voters are asked to rank the impairment of the second video using as reference the first one. Finally, in the DSCQS method, the video sequences are showed in pairs as well, similarly to DCR, one is the impaired and another is the reference. However, the subjects do not know which one is the reference and are asked to score the quality of both videos. The resulting scale is present in pairs containing both the scores for the reference and the impaired videos.

The subjective methods provide good results, however, they are fairly expensive to set up, need the collaboration of a considerable number of humans voters, and it takes a long time to get the results [Winkler and Mohandas, 2008]. Taking this into consideration, objective metrics are desirable. They are fast and intend to be unbiased. In addition, they are computed through mathematical calculations, and thus, measurable and verifiable.

## 3.4.2 Objective metrics

One of the most common objective metrics is the Peak Signal to Noise Ratio (PSNR) [Ma et al., 2011]. This is a non-perception-based metric that is widely used to evaluate the performance of several optimisation techniques. This metric gives the distortion measurements between two video images, which are averaged over the video sequence. Through this approach, a single well-defined metric is provided and it has proven useful to compare different optimisation mechanisms using a single scale. However, the absolute PSNR values of different video content may not be directly related [Huynh-Thu and Ghanbari, 2008a]. Additionally, due to the use of averages, this metric is not able to differentiate between individual loss occurrences over a period of time.

Equation 3.2 defines the PSNR calculation, where the output result is a logarithmic decibel value. This scale is used because of the wide dynamic range of several signals. The PSNR calculation is based on the Mean Squared Error (MSE), defined in Equation 3.1. Table 3.3 summarises the notations adopted.

Table 3.3: PSNR Notation

| Parameter | Meaning |
|---|---|
| $MAX$ | Upper pixel value limit of the input image |
| $A$ and $B$ | Assessed images |
| $m$ | Image width in pixels |
| $n$ | Image height in pixels |
| $m \times n$ | Number of pixels |

$$MSE = \frac{1}{m \times n} \sum_{i=1}^{m} \sum_{j=1}^{n} (|A_{(i,j)} - B_{(i,j)}|)^2 \qquad (3.1)$$

$$PSNR(dB) = 10 \times \log_{10} \left( \frac{MAX^2}{MSE} \right) \qquad (3.2)$$

At the end, the PSNR is a common objective metric to assess data fidelity. However, it is based on a byte-by-byte comparison disregarding what the information actually represents. Additionally, PSNR does not recognise the pixel structure in the image nor the spatial relationship between the pixels, thus, it does not consider the visual importance of each pixel [Huynh-Thu and Ghanbari, 2008b]. Since the results obtained from PSNR do not correlate with subjective human perception, other metrics have been proposed. Two of the most widely adopted objective QoE metrics that correlate with subjective perception are Structural Similarity Metric (SSIM) and Video Quality Metric (VQM) [Chikkerur et al., 2011].

The SSIM [Wang et al., 2004] assesses video quality through frame-to-frame measurements. This metric employs the video structural distortion as an estimate of the visual quality. This provides a good correlation with the perceptual

quality because the human visual system is very effective in the extraction of structural information, thus, any impairment in this information will be proportional to the perceived errors.

SSIM indexes are composed of three basic components, namely, luminance, contrast and structural similarity. The index values are calculated only for carefully selected blocks inside each frame, and not the whole frame. This allows saving computation resources while providing precise video quality assessments. At the end, an average of all SSIM indexes will give the frame quality. This information is used together with the local luminance to define a final perceptual quality.

Finally, the general video quality is found by the weighted sum of all frames indexes. The weighting value is based on each frame's motion intensity. The overall index value is a decimal number between 0 and 1, where 1 stands for precisely the same video. A simplified SSIM calculation is defined through Equation 3.3 and Table 3.4 specifies the parameters used.

Table 3.4: SSIM Notation

| Parameter | Meaning |
|:---:|:---|
| $c$ | Contrast correlation |
| $l$ | Luminance correlation |
| $s$ | Structure correlation |
| $x$ and $y$ | Measurement between two windows of equal size |
| $\alpha, \beta, \gamma$ | Significance of each component |

$$SSIM(x,y) = [c(x,y)]^{\alpha} \times [l(x,y)]^{\beta} \times [s(x,y)]^{\gamma} \qquad (3.3)$$

VQM [Pinson and Wolf, 2004] is a modified discrete cosine transform-based metric to assess video quality. Values closer to 0 correspond to the best video quality. In order to calculate these values, this metric uses the same features as the human eye to perceive the video quality, including colour and block distortion, blurring and global noise. More specifically, this model employs a linear combination with seven independent parameters. Four parameters are extracted from spatial gradients ($SI_{loss}$, $SI_{gain}$, $HV_{loss}$, $HV_{gain}$), two are obtained from a chrominance vector ($CHROMA_{spread}$, $CHROMA_{extreme}$), and the last one is derived from absolute temporal and contrast details ($CT\_ATI_{gain}$). These parameters are defined as flows:

- PARAMETER $SI_{loss}$ - It is responsible for detecting differences in the spatial information. It uses a 13-pixel spatial information filter (SI13) that evaluates any meaningful visual edge impairments (e.g., blurring);

- PARAMETER $SI_{gain}$ - This parameter measures the opposite situation of $SI_{loss}$. This means that any enhancement in the video quality provided by any edge sharpening technique will be reflected here;

- PARAMETER $HV_{loss}$ - It is in charge of identifying and measuring any edge

fluctuation from horizontal/vertical position to a diagonal orientation. In doing this, it is able to catch horizontal/vertical edges that suffer more blurring than diagonal edges;

- PARAMETER $HV_{gain}$ - This parameter computes the reversed edge movement of $hv_loss$. In other words, it detects diagonal-oriented edges that had tilted to horizontal/vertical-oriented edges (e.g., blocking artifacts);

- PARAMETER $CHROMA_{spread}$ - It identifies the variation of two-dimensional colour samples. In the same way as in the other parameters, the colour variation is only reported if it is over a large area which can impact significantly on the perceived quality;

- PARAMETER $CHROMA_{extreme}$ - As the parameter's name suggest it is also responsible for detecting colour variations. Its operation is similar to the $CHROMA_{spread}$, but differs by exploring acute colour impairments in a restricted area (e.g., common impairment produced by digital transmission errors);

- PARAMETER $CT\_ATI_{gain}$ - This parameter infers the motion intensity of a defined region by computing the product of temporal and spatial intensity. This information is important to grade how much the impairments will impact on the perceived video quality. A high motion intensity region will have more perceptual artifacts than a low motion intensity region (e.g., noise and error blocks).

After the definition of all parameters, the VQM applies a linear combination to score the video quality. This metric is optimised to obtain a good correlation between objective and subjective assessments.

## 3.5 Summary

This chapter described several components and details related to video technology. In Section 3.2 video compression techniques were discussed, along with the description of several methods that allow reducing the redundant video information. These algorithms include statistical compression (spatial and temporal), as well as perceptual redundancy compression. Additionally, the H.264 structure and components were described, including the GoP format, the different types of frames, macroblocks, and blocks.

Section 3.3 presented how the video sequences are impacted by packet loss. It was evidenced that not all packets have equal importance on how the quality is perceived. This means that some packets (or group of packets) are more important than others. Following the same idea, I-Frames are more important than, P-Frames, which in turn, are more important than B-Frames. Furthermore, the position inside the hierarchically-structured GoP is equally important. Frames closer to the beginning of the GoP are more important than frames closer to the end. Additionally, the motion intensity of the videos also plays an important

role. Videos with low motion intensity are more resilient to packet loss than a video with high motion intensity.

The video quality assessment was examined in Section 3.4. The concept of QoE was defined as a measure of customer's personal perception of a service or application. This metric is related to QoS but differs by encompassing additional factors providing an end-to-end assessment. The methods used to score the QoE can be objective or subjective. The former uses mathematical models and the latter requires human voters to quantify the video quality.

# Chapter 4

# Advances on Improved Video Transmission

And even technological progress only happens when its products can in some way be used for the diminution of human liberty.

*(George Orwell, Nineteen eighty-four)*

## Contents

T he design of techniques to improve the video transmission quality is a complex task which involves a number of concepts and definitions, especially error-correcting codes. This chapter discusses several of these details to provide a general understanding on this subject. It also presents a literature review of mechanisms that aim to enhance the video transmissions with and without error-correcting codes.

## 4.1 Introduction

Video services are very demanding applications, as they require a steady and continuous flow of packets. Taking this into account, both network and application-level technologies must be adopted to enable the delivery of a video stream with

satisfying quality. At the network level, QoS techniques can be used to safeguard video transmissions. However, this does not guarantee that the perceived quality by the end-user is at the best possible. Thus, other mechanisms or techniques are required to provide the end-users with high perception video.

Video transmissions over wireless channels introduce a set of challenges. The omnidirectional signal has the tendency to endure time-varying physical effects which can result in a degraded video quality. The outcome of shadowing, multipath fading, hidden terminal, and the antenna range may originate errors in packets leading them to be dropped. Additionally, the strict bandwidth, delay requirements, and node overpopulation can also aggravate the situation creating channels with high error rates.

To mitigate these issues, error correction techniques have been successfully used to protect the transmission of real-time video services [Nafaa et al., 2008]. These techniques provide robust video transmission through redundant information that is sent along with the original data set (FEC-based), or by resending the lost packets (ARQ-based). However, despite the latest developments, there is still a shortage of adaptive QoE-driven mechanisms to improve real-time video transmissions [Jiang et al., 2012, Bellalta et al., 2014]. These mechanisms need to be able to perform on unforeseen situations to protect the most QoE-sensitive data, while not adding unnecessary network overhead. In order to do that, several aspects of the video details, along with the network characteristics and condition need to be considered. As a result, it is possible to assign an optimal amount of redundant data only to QoE-sensitive data, or carefully choose which information should be sent again in case of losses. These are important features in any mechanism, especially if it will be used in highly dynamic networks.

To address the aforementioned issues, this chapter introduces and discusses error correction techniques in Section 4.2. After that, several important studies about the optimisation of video transmission are presented and reviewed in Section 4.3, followed by the summary and open issues in Section 4.4.

## 4.2 Error-Correcting Codes (ECC)

ECC or error correction (EC) techniques are used to enhance the communication over unreliable channels. This type of approach is especially needed in wireless channels which are subject to a series of issues. Using EC techniques it is possible to reconstruct the original error-free data if losses occur during the transmission. Generally speaking, the error correction can be carried out with two distinct methods, namely Forward Error Correction (FEC) and Automatic Repeat reQuest (ARQ). FEC-based schemes send redundant information (parity bits) along with the original data set, which can be used to recover the original data in case of loss. On the other hand, ARQ-based schemes use error-detection codes along with positive and negative acknowledgements (ACK). If a negative ACK is received (or a time-out happens), the sender performs a retransmission to correct erroneous or lost data. Both FEC and ARQ methods are discussed

below.

## 4.2.1 Automatic Repeat reQuest (ARQ)

ARQ is an error-control (or error-correction) method based on the retransmission of the packets that fail to reach their destination and/or arrive in a damaged state. This retransmission process is generally triggered by two conditions: (1) packets that are not acknowledged on time, in the case of packet loss (delay-constrained retransmission), and (2) through data request, when erroneous packets arrive (parity retransmission). Taking this into consideration, the use of a reliable feedback channel is of critical importance to achieve a good performance. This is a closed-loop mechanism also known as backwards error correction, and is commonly adopted in unicast protocols [Rizzo, 1997]. These techniques are particularly helpful if the end-to-end network statistics are unknown [Setton et al., 2008].

There are basically three types of ARQ, namely Stop-and-wait, Go-Back-N, and Selective Repeat ARQ (ARQ-SR). The first, Stop-and-wait ARQ, is the most naive method because it sends one frame at a time, and then, waits for a positive ACK, a negative acknowledgement (NACK), or a time-out incident. In the case of an ACK, the next frame is transmitted; however, in the case of a NACK or time-out, the same frame is transmitted again, until an ACK is received. This method ensures that the data is not lost, as well as correct packet order. It is clear that this method raises several real-life implementation issues. For instance, if an ACK is lost or damaged, the same frame will be re-sent; the latency can significantly increase; the channel throughput is just a fraction of what is expected. This issue is addressed in the other ARQ methods through the addition of a larger sequence number [Zhang et al., 2012]. In this way, it is possible to send several packets at the same time, and then wait for an ACK to acknowledge the whole set.

The second method, Go-Back-N ARQ, sends the number of frames defined by a sliding window, before receiving any ACK from the receiver. To keep track of the sequence number, the receiver process always acknowledges what has been sent with the number of the last frame received correctly, and in the right sequence. If this number matches some point in the sliding window, it slides to that point, and the transmission or retransmission of the missing frames is started. On the other hand, if the number does not match any point, all the frames inside must be retransmitted. This process continues until there are no more packets. It is a more efficient method because there is no need to wait for an ACK for each packet, and hence, time is not spent waiting for another packet to be sent [Chen and Wu, 2012]. The drawback is that this method can involve having to send duplicate frames because if one frame in the window is lost or damaged, every subsequent frame has to be retransmitted.

The last method, ARQ-SR, also continues to send the frames inside the sliding window even after losses. However, it differs from the Go-Back-N as it only resends the lost frames. On the one hand, this makes it more bandwidth efficient,

on the other hand, it needs a more complex process to receive and acknowledge the frames. This means that it has to keep track of the earliest frame that has not been received, and reports to the sender the ACKs and NACKS. As a result, the sender will use the ACKs to move the sliding window and the NACKs will be the only ones that are retransmitted [Cai and Yang, 2008].

These error correction methods are appropriate as a means of recovering from a few errors when the network is temporarily congested or undergoing some type of interference. If the conditions of the network have deteriorated, even the ARQ-SR scheme will produce a large number of packets that need to be retransmitted, generating a feedback implosion problem, which can degrade the network even further [Li et al., 2007]. Another drawback is that, although ARQ does not consume unnecessary bandwidth, it can increase the round-trip time (RTT) and the delay [Cai and Yang, 2008]. In this case, it is only helpful to retransmit packets if they have a chance to arrive on time, or in other works, before its playout time. Additionally, several enhancement techniques have been proposed for the ARQ methods [Tsai et al., 2009, Han et al., 2010, Hassan and Landolsi, 2010], including some in combination with FEC-based schemes. Further details can be found in Section 4.3.2.

## 4.2.2 Forward Error Correction (FEC)

FEC are methods that, if correctly employed, can enhance the video quality in noisy or lossy networks. They are open-loop mechanisms that send redundant data along with the original set. This means that if some original information is lost, the error-free data can be reconstructed using the redundant information, without the need for retransmissions. One advantage of this technique lies on the fact that no further interaction between the sender and receiver is required. This happens because the receiver is able to reconstruct the lost information solely through the correctly received packets.

One advantage of this technique is to ensure a good performance when applied to multiple receivers in an error-prone environment. In this type of scenario, the video impairments tend to occur due to uncorrelated and time-varying losses. A good application of this technique, for example, is to provide reliable multicast on the Internet [Rizzo, 1997].

This section presents three categories of FEC codes as shown below. Although there are several other implementations of EC techniques, our objective here is not to make a survey of these mechanisms but just to name a few examples.

- **Block codes**: These ECCs work on a fixed-size block of $k$ input bits (or symbols) generating $n$ bits in the output $(n, k)$. This means that the sender builds blocks of predetermined size $(k)$, adds a predefined amount of redundancy (parity bits), and transmits $(n)$ to the receiver. The receiver, in turn, uses a decoding mechanism to rebuild the error-free data. The overall performance and success of the transmissions rely upon the parameters chosen for the block code configuration. Examples of block codes are the

Hamming(7,4) [Hamming, 1950], the Reed-Solomon (RS) code [Reed and Solomon, 1960], and the Low-density parity-check (LDPC) codes [Gallager, 1962].

- **Convolutional codes**: It can be characterised as continuous or codes with arbitrary block length. These codes differ from the previous, as their size is not settled by any algebraic properties. The generation of the parity bits (or symbols) is generally applied using a sliding operation of a function with boolean polynomial proprieties. Examples of convolutional codes are the Turbo code [Berrou and Glavieux, 1996] and the Serial concatenated convolutional codes (SCCC) [Benedetto et al., 1998].

- **Fountain codes**: This technique produces an endless supply of encoded symbols. The original error-free data can be recovered from any subset of encoded symbols that is larger or at least equal in size to the source set. These codes are considered rateless because they do not have a fixed code rate. Examples of fountain codes are the LT codes [Luby, 2002] and the Raptor codes [Shokrollahi, 2006].

All FEC techniques add some type of redundant information which is sent along with the original data set. In order to provide a good result at the receiver side, a suitable amount of redundancy has to be added. One disadvantage of this scheme is the increased computational effort required to compose the redundancy. There are however FEC schemes which provide good performance and can be used in real-time video transmission, RS codes, which will be detailed below. Another drawback is the network overhead, as it needs a larger bandwidth to add a larger amount of redundant information for all the data [Neckebroek et al., 2010, Abboud et al., 2011].

This problem can be overcome by employing coding optimisations that include content-aware models with a UEP scheme [Nafaa et al., 2008]. The UEP scheme can take advantage of the knowledge about video details obtained through the cross-layer models. In this way, these methods can provide different amounts of redundancy to the video packets that are more important, and will thus cause a bigger impact if they are lost. In doing this, packets with more QoE-sensitive data are effectively protected, and this reduces the impairments in the video sequence, while saving network resources.

The RS code [Reed and Solomon, 1960] is one of the most widely used FEC-based methods. RS code belongs to the linear block codes category and it is a non-binary cyclic ECC that uses univariate polynomials over blocks with a predetermined size. This code can be use to provide error control in a great variety of digital systems, such as CDs, DVDs, Blu-ray discs, barcodes, digital television, ADSL, xDSL, satellite data links, as well as in wireless and mobile communications.

This ECC method can detect and correct multiple errors in each block. In order to do that, the encoder receives the source blocks and adds a determined amount of redundancy. This information is sent to the receiver and the decoder will attempt to detect any possible error and use the redundant data if needed

to rebuild the original error-free data, as depicted in Figure 4.1.
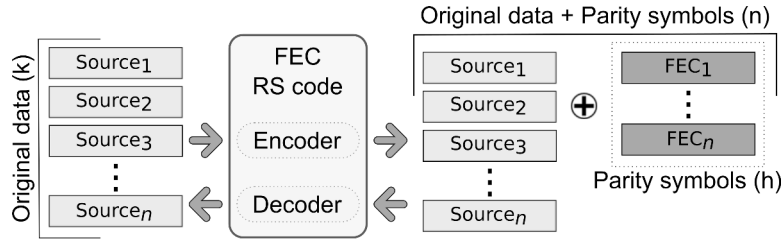


Figure 4.1: Reed-Solomon code

Table 4.1 shows the RS notation. The code can be specified as $RS(n, k)$. This means that the total block size, including the redundancy data symbols, is represented by $n$. The parameter $k$ indicates the original data set size, resulting in the $(n, k)$ parity code. The last parameter is $h$, and it defines the amount of redundancy (or parity symbols), which is the same as $h = (n - k)$. In order to recover the entire original data set $k$, at least $(n - h)$ packets have to arrive successfully. The robustness to losses is determined by the size of $h$, and the error recovery against an average packet loss rate can be expressed as $h/n$ or $(n - k)/n$.

Table 4.1: Reed-Solomon Notation

| Notation | Meaning |
|:---:|:---|
| $k$ | Original data set |
| $n$ | Original data set + parity symbols |
| $h$ | Redundancy amount (or parity symbols) |
| $h/n$ | Error recovery rate |

The RS erasure code has low complexity and therefore offers a better performance for real-time services [Neckebroek et al., 2010]. This is important when considering its adoption in services that include video delivery as they demand a continuous flow of packets as well as low delay.

## 4.3 Video Transmission Optimisation

Owing to the wide range of solutions, the mechanisms presented in this section were divided into Non-EC and EC-based methods. It is possible to enhance video transmission with several techniques such as packet prioritisation, link adaptation, and video bitrate adjustment. However, without using an EC-based scheme this improvement will only be attainable to a limited degree. At the end, an error correction mechanism is needed to ensure a high video quality as perceived by end-users, whatever network adversity may occur.

## 4.3.1 Non-EC based Optimisation Mechanisms

Several techniques have been proposed to enhance the quality of video transmissions without using EC-based methods. One approach is the combination of an adaptive algorithm in the link layer and a scalable video codec in the application layer [Haratcherev et al., 2006] to improve the quality of real-time video streams over wireless channels. Its architecture is based on a cross-layer technique that continuously adapts the video encoding rate (through the use of scalable video codec) at the application layer, in accordance with its link layer status. Another advantage of this approach is that it employs a link adaptation algorithm. This algorithm is able to efficiently manage the modifications in the wireless channel conditions by using both SNR and the wireless channel statistics to automatically adjust the radio and link layer parameters, with the aim of obtaining the best possible packet transmission quality.

One of the weaknesses of this architecture is that it does not take into account the video content that is being delivered. It is known, for instance, that videos with lower spatial or temporal resolution are more loss resilient and in this case, the encoding rate should not be changed, since it leads to unnecessary modifications, and consumes time and resources. Another problem is that the encoding characteristics of the scalable video have to be chosen prior to the running time. If these parameters are not carefully chosen, there may be some network conditions that do not fit any of the existing encoding systems, and this will lead to the selection of non-optimal coding.

Another cross-layer proposal allows the video packets to be mapped, from the application layer to the appropriate Access Categories (ACs), at the link layer [Chilamkurti et al., 2010, Lin et al., 2009]. This classification is performed depending on the importance of the video information that the packets are carrying. In doing this, it enables a content-aware categorization to prioritise the video packets in terms of their frame type. This means that I-frames will have higher priority, followed by P-frames, and B-frames will have lower priority. In addition, downward probabilities are employed. This means that, if the high priority AC queue exceeds a pre-determined threshold, the less important packets are mapped to lower priority ACs.

This solution provides good results when the network is not heavily loaded. However, when packet losses are caused by network congestion (e.g., concurrent transmissions from other nodes), it is not sufficient to prioritise the packets. In this case, some type of error recovery technique must be applied to improve (or even maintain) the video quality. A further weakness is the lack of assessment of the motion activity, as frame types have different degrees of importance depending on the motion activity of each frame. Finally, the position of each frame inside the GoP also plays an important role. It should be taken into account that frames close to the beginning of the GoP produce larger video impairments if lost than frames close to the end of the GoP [Greengrass et al., 2009a].

Another proposed solution is a technique to improve the quality of the video transmission through optimising the network resource utilisation in accordance

with the video content classification [Khan et al., 2010]. The optimisation is achieved by determining the initial quality and then adapting the Video Sender Bitrate and the Frame Rate to the users QoE requirements. This allows the impact of the QoS parameters to be mapped to the end-to-end perception of video quality. The need of preprocessed videos reduce the applicability of this solution and also require a lot of processing power to carry out the content classification since they are not suitable for real-time services. Additionally, the video content which could improve the video performance, even more, is not taken into account.

Another approach is employ a border proxy (located at the edge of the WMNs) to improve the video traffic delivery and provide two services [Qiu et al., 2011]. The first service is a route selection algorithm, called Minimum Interference Route Selection (MIROSE), which is able to choose routes based on the amount of interference. In doing this, the route with minimal interference can be chosen and thus allow a better usage of network resources. The second service is responsible for establishing the optimal video streaming rate according to the dynamic and unpredictable nature of the wireless channel conditions and is called Network State Dependent Video Compression Rate (NSDVCR) [Qiu et al., 2009]. In carrying this out, the NSDVCR algorithm can temporarily buffer the video stream to optimise the compression rates so that it suits the network conditions of the end-to-end path. This re-compression is based on a heuristic that determines the optimal compression rate that corresponds to the current network status.

This technique has several weaknesses; for example, re-coding is a processor-intensive activity which raises serious scalability issues, and the same applies to the usage of a buffer in the gateway. As a result, it would have to include a comprehensive storage space to accommodate a higher number of concurrent streams. This architecture has a single point of failure, which means that if the gateway is down, the video enhancement will not work. Some of these issues can be remedied with a high availability and load-balancing system, although this increases deployment and maintenance costs.

A different method was proposed for VANETs. It adopts an adaptive multi-objective Medium Access Control (MAC) retransmission limit strategy [Asefi et al., 2012]. At the Road Side Units (RSUs), channel statistics and packet transmission rate are used as input to the optimisation framework in order to tune the MAC retransmission limit. Although this optimisation improves the performance of video transmission, it only aims to minimise the playback freezes and reduce the start-up delay. These are important characteristics, however, QoE metrics should be used to assess the image quality. This evaluation would provide a more comprehensive assessment of the proposed mechanism. Additionally, the authors only took into account the use of RSUs and two-hop communications. It is known that the major advantages of VANETs come from the communication directly between the vehicles, without the need for a fixed infrastructure. This severely restricts the application of the mechanism.

Another VANET proposal relies on routing protocol adaptations, such as the QoE-based routing protocol for video streaming over VANETs (QOV) [Pham

et al., 2014]. In QOV, the perceptual quality of the videos is assessed in real-time, at the receivers, using the Pseudo-Subjective Quality Assessment (PSQA) [Rubino, 2005] metric. After that, the results are announced to the neighbours throughout `Hello` packets. This allows the routing protocol to choose the best paths to deliver the video sequences. Nevertheless, VANETs are very dynamic networks and because of that, the proposed mechanism would have to update very quickly the PSQA result announcement, overloading the network with `Hello` packets. Another weakness of this proposal is that it does not include any type of EC. As aforementioned, the video quality can be maintained only up to a certain level without using EC, however, if the network has a high packet loss rate the quality will decrease.

The mechanisms presented in this section are able to improve the performance of the video transmissions, however, there is a lack of any type of EC on them. As previously mentioned, without using an EC technique the video quality can only be sustained up to some degree. After that, an impact on the visual video quality will be noticed if the number of errors overcomes the natural video resilience to packet loss.

## 4.3.2 EC-based Optimisation Mechanisms

Several EC-based techniques have been proposed in the past years to improve the perceived video quality. One example is the Adaptive Multi-Hop FEC (AM-FEC) protection scheme [Tsai et al., 2009] to improve the quality of video streaming data. The proposal seeks to reduce the end-to-end delay and FEC computational costs at the same time. Through a heuristic algorithm, the AM-FEC protection minimises the amount of redundant information sent with the original data set. The video frame rate is also used to dynamically adjust the FEC scheme on each link.

One disadvantage of this approach is that it does not use QoE metrics to improve the performance. The QoE methods are only used in the assessment. Other important parameters that are missing in the AM-FEC protections are codec type and motion complexity, which have proved to be efficient in this kind of scheme.

The Adaptive Cross-Layer FEC (ACFEC) mechanism uses packet-level error correction [Han et al., 2010]. Through a cross-layer design, these packets are monitored by the mechanism at MAC layer, and when a loss occurs, a failure counter is increased. The information held by the failure counter determines whether the number of FEC recovery packets is increased or decreased. In this way, when the counter is zero, it means that there is no packet loss and the wireless connection is good, and thus recovery packets are not generated which results in less redundant traffic. However, no assessment of the network overhead is conducted, which is very important, especially in a wireless environment.

As mentioned earlier, this type of network generally does not have a fair bandwidth allocation which, means that increasing the overhead of one node may af-

fect the communication between the others. In addition, the approach discussed above does not consider the video content, and, as is well-known, this information has a direct influence on how the video is resilient to packet loss [Aguiar et al., 2012b]. Although the ACFEC mechanism seems to be a good solution when the network is healthy and there is sporadic packet loss, when network congestion occurs, this mechanism will start to generate more and more FEC redundancy packets, which will increase the congestion.

Another technique to enhance the quality of the video transmission employs a forward error correction and retransmission-based adaptive source-channel rate control [Hassan and Landolsi, 2010]. This scheme uses real-time monitoring of the decoder buffer occupancy and the channel state, to calculate the optimal parameters for FEC redundancy. This information is regularly returned, through a feedback channel or out-of-band using Real Time Streaming Protocol (RTSP) [Schulzrinne et al., 1998], to the video encoder at the server site, which proceeds to adapt it to its own transmission parameters.

The authors claim that there has been an improvement in the QoE for end-users, however, the main objective of this scheme is to ensure the continuity of video playback with unpredictable channel variations and avoid unnecessary FEC redundancy. Moreover, information such as video content and frame type were not considered in the definition of the proposal [Aguiar et al., 2012b]. This approach does not assess QoE metrics, as it relies on packet loss values to predict QoE levels, and it does not measure the overhead that has been introduced.

A distinct proposal to enhance video transmission over wireless local area networks are based on a method which adapts in real-time the amount of FEC redundancy and the transmission rate [Alay et al., 2010]. In order to adjust the FEC redundancy and the transmission rate, the receivers periodically send the packet error rate information to the Access Point (AP). Using this information, the AP can identify the worst channel's condition and then adjust the transmission rate and FEC. The application level FEC redundancy adaptation is done by multiple pre-encoded videos with different bit rates and FEC rate, so, in order to adapt theses parameters, the system has to switch to a different bit stream. The bit-stream switching is always initiated when the next frame is an I-frame because it is independently decodable.

One of the downsides of this proposal is the need of a pre-processed video which reduces the applicability of this solution. It also demands high processing power and storage space, since there is the need to encode multiple times the same video with different bit rates and FEC redundancy. Moreover, the FEC overhead amount introduced by this mechanism was 48% (without taking into account the feedback messages overhead), which is higher than the proposed mechanisms as shown in Chapters 5, 6, and 7.

The Adaptive Hybrid Error Correction Model (AHECM) solution adopts a dynamic FEC block length [Tsai et al., 2011a]. This FEC block can be adjusted in real-time depending on Markov models to estimate the PLR and the number of continuous losses, to boost video transmissions. This mechanism is heavily based on network parameters leaving out important QoE-sensitive information,

such as frame size and type, as well as the motion intensity. Furthermore, the mechanism uses a buffer to cope with the impact of packet disordering. This should increase delay and lead to the discarding of the packets by the encoder due to playback time out. Following the same pattern as the studies outlined above, no attempt is made to measure the network overhead.

An alternative mechanism is the Adaptive Packet and Block length FEC (APB-FEC) [Tsai et al., 2011b]. This mechanism uses smaller packet lengths in order to increase the size of the FEC block. A feedback channel is used to receive packet loss information in order to adapt the video sequences to the network characteristics. Additionally, the mechanism also adapts the buffer size according to the network conditions. The use of buffers is not optimal and can increase the delay. Also, relying on information from the receiver can be problematic due to the fact that if the communication is hindered the feedback information may not reach the sender. In addition, the use of past PLR can lead to an inaccurate characterization of the network due to outdated information. The performance assessment is based on the effective packet loss rate, the network overhead, and PSNR metric. However, the PSNR metric is known to not correlate well with how the QoE is perceived by end-users.

A FEC-based mechanism was proposed to enhance the quality of video streaming using video-aware techniques [Diaz et al., 2011]. In this approach, some parameters are used to define the importance of a data set, namely the frame relevance, channel state, and bitrate constraints. With this information, the proposed algorithm is able to select the most suitable frames to give protection in real-time. This mechanism also employs an unequal loss protection technique, providing the possibility to select the most suitable frames to add redundancy, and consequently, securing the delivery of the most important information.

One of the weaknesses of this solution is that it only uses the frame type in order to define the relevance of the data and does not take into consideration the motion activity levels, which can have a considerable impact in the perceived impairments [Greengrass et al., 2009a]. It also requires the addition of extra tags in RTP encapsulation, which means that the videos have to be reprocessed.

The Cross-Layer Mapping Unequal Error Protection (CLM-UEP) assigns a different level of redundancy according to the frame type of the video sequences and the packet loss rate [Lin et al., 2012a]. Moreover, this mechanism has an adaptive mapping algorithm to direct the video data and redundant packets to the suitable Access Category (AC) queues. This operation also takes into consideration the frame type and the packet loss rate, as well as the AC queue occupancy to avoid congestion-induced packet losses. The amount of redundancy is defined through the analysis of the frame type and the past PLR.

The mechanism was assessed using the Playable Frame Ratio (PFR) and the PSNR score. Nonetheless, the past PLR may not repeat in the near future leading to a mischaracterization of the network. Moreover, the average PLR will not capture fast time-varying changes in the wireless network channels. Another major drawback of this mechanism is the lack of use of important video characteristics, such as levels of motion activity and position of the frames within

the GoP. As evidenced before, this information plays a substantial role to define the best amount of redundancy, allowing the system to save important network resources.

An additional work proposes the use of a 2-state Markov hierarchical model to predict the short-term losses and hidden Markov models (HMM) to forecast the longer-term network losses [Silveira and Silva, 2012]. In doing this, both the PLR and burstiness are categorised and used as input to configure the amount of redundancy added by the FEC scheme. The assessment is performed with the Perceptual Evaluation of Speech Quality (PESQ) and MOS metrics. This proposal does not consider the video characteristics in the decision-making process. These characteristics are known to have a direct impact on the video resiliency to packet loss and consequently, on the QoE for the end-users.

The "Transport Audiovisuel avec Protection Inégale des Objets et Contrôle d'Admission" (TAPIOCA) mechanism divides each GoP by layers assigning different priority values to each one [Lecuire, 2012]. This allows the protection of the most important layers. The assessments of the mechanism were performed by the Decodable Frame Rate (DFR) and the Protection System Efficiency (PSE) metrics. These assessment techniques are so unique that they do not provide much information about the QoE performance of the scheme. The process of dividing each GoP into layers is both computationally heavy and time-dependent, making this scheme unsuitable for real-time use. Additionally, the mechanism does not take into consideration the motion intensity of the video sequences, which can have a considerable influence in the perceived impairments.

The Optimised Cross-Layer FEC (OCLFEC) [Talari et al., 2013] computes priority values based on the mean squared error of each frame. Two error correction codes are used, namely Luby Transform (LT) and Rate-compatible Parity Check (RCPC). The former is used to encode the data and the latter to add check bits. Performance optimisations are made on both for specific situations. The GoP information is the obtained using a cross-layer technique and then assigned different priorities according to the commutative mean squared error of the entire GoP. The GoPs are encoded and cyclic redundancy check bits are added to detect coding errors. Afterwards, the FEC codes are optimised with different parameters for different situations. This mechanism has several optimisation phases, for each frame, that are very time-consuming, increasing the delay and degrading the QoE.

The mechanism is assessed in terms of QoS performance, which does not guarantee a good QoE for end-users. The only metric used to assess the performance of the mechanism is the PSNR, which by itself does not tell much information about QoE. Besides that, the OCLFEC does not take into account the motion intensity and the network state, leaving out important characteristics that should be considered to protect video sequences.

A different mechanism to improve the video quality over wireless networks compares the efficiency of Random Linear Coding (RLC) and XOR-based coding [Rezende et al., 2013]. The benchmark results show that both erasure codes are able to improve the video quality by increasing the number of successfully re-

ceived packets over error-prone networks. The results also show that XOR-based coding outperforms the RLC scheme.

In addition, the proposed mechanism finds the optimal packet block size, which allows adding a more precise amount of redundancy. While this is true, important network and video characteristics are not taken into consideration. Some of these features, such as packet loss, video content, and codec, are important in the optimisation process to provide a way to compute a precise amount of redundancy leading to both high video quality and low network overhead.

The Hybrid Video Dissemination Protocol (HIVE) [Naeimipoor and Boukerche, 2014] uses a multi-layer strategy to improve the video quality. The HIVE multi-layer strategy is based on the joint use of traffic congestion control scheme, node selection method, and application layer erasure coding technique. This allows higher packet delivery ratio while keeping latency and packet collisions low.

The results show improvement in the PSNR assessment, leading the authors to claim that they improved the QoE for end-users. However, relying on only one metric is not enough to prove that, especially considering that the PSNR results do not correlate well with the human vision system [Huynh-Thu and Ghanbari, 2008b]. Another issue is the lack of video characteristics assessment. It is known that these video details have a considerable impact on how resilient a video sequence is when experiencing packet loss.

Table 4.2 compiles the main characteristics of all above-named works. The parameters are defined as follows:

- **FEC-based**: accounts for mechanisms that employ FEC;

- **ARQ-based**: mark mechanisms that use ARQ;

- **QoE-sensitive data**: this parameter demonstrates mechanisms that identity and/or considerate the video content to define the EC policy;

- **Video-aware**: check mark is given to mechanisms that use any video characteristics to define the amount of redundancy and/or retransmission;

- **High-quality video**: it is marked if the mechanisms are using videos equal or higher than 720p (HD ready);

- **Network status**: this parameter defines if the mechanisms use the information about the network healthy to define the redundant data;

- **UEP-enabled**: means that different amounts of redundancy are being added to distinct portions of the video.

- **QoE assessment**: used if the mechanism's assessment is performed by subjective and/or objective QoE metric;

- **Network assessment**: this parameter reflects the network-related evaluation, especially the network overhead caused by the redundancy;

Table 4.2: State of the Art of EC-based optimisation mechanism

| Proposals / Parameters | FEC-based | ARQ-based | QoE-sensitive data | Video-aware | High-quality video | Network status | UEP-enabled | QoE assessment | Network assessment |
|---|---|---|---|---|---|---|---|---|---|
| [Tsai et al., 2009] | ✔ | ✔ | ✗ | ✗ | ✗ | ✗ | ✗ | ✔/✗[1] | ✗ |
| [Alay et al., 2010] | ✔ | ✗ | ✗ | ✔ | ✗ | ✔ | ✗ | ✔/✗[1] | ✔ |
| [Han et al., 2010] | ✔ | ✔ | ✗ | ✗ | ✗ | ✔/✗[2] | ✗ | ✔/✗[1] | ✗ |
| [Hassan and Landolsi, 2010] | ✔ | ✔ | ✔ | ✗ | ✔ | ✔/✗[3] | ✔ | ✗ | ✗ |
| [Tsai et al., 2011a] | ✔ | ✗ | ✗ | ✗ | ✗ | ✔ | ✗ | ✗ | ✔ |
| [Tsai et al., 2011b] | ✔ | ✗ | ✗ | ✗ | ✗ | ✔ | ✗ | ✗ | ✔ |
| [Diaz et al., 2011] | ✔ | ✗ | ✗ | ✔ | ✗ | ✔ | ✔ | ✔/✗[1] | ✗ |
| [Lin et al., 2012a] | ✔ | ✗ | ✗ | ✔ | ✗ | ✔ | ✔ | ✔/✗[1] | ✔ |
| [Silveira and Silva, 2012] | ✔ | ✗ | ✗ | ✗ | ✗ | ✔ | ✗ | ✔ | ✔ |
| [Lecuire, 2012] | ✔ | ✗ | ✗ | ✔ | ✗ | ✗ | ✔ | ✗ | ✔ |
| [Talari et al., 2013] | ✔ | ✗ | ✔ | ✔ | ✗ | ✗ | ✔ | ✔/✗[1] | ✗ |
| [Rezende et al., 2013] | ✔ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✔ |
| [Naeimipoor and Boukerche, 2014] | ✔ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✔/✗[1] | ✔ |

[1] Uses only PSNR, which does not correlate well with how humans see video impairments
[2] Based only in the last PLR
[3] Uses the probability of errors, no real measure nor actual feedback

As evidenced, these works fell short to produce an holistic mechanism to provide resilient video transmission without imposing unnecessary network overhead. Some of the proposed works do not provide a QoE-driven procedure to compute the necessary amount of redundancy. This is an important step towards an efficient mechanism, as the most QoE-sensitive information has to be better protected than other less important data. Without this, an unnecessary network overhead can be generated. This leads to another problem; several studies do not provide a clear account of this matter, and this hampers any comparisons from being made. Other proposals fail to take into consideration the video-aware data. It is known that this content plays an important role in the video transmission process [Khan et al., 2010], and without this information, the proposed mechanism cannot achieve the most advantageous performance.

## 4.4 Summary and Open Issues

The widespread deployment of wireless networks, as well as the usage of mobile devices, has increased significantly in recent years. These networks are well known for their dynamic nature along with their time-varying channel conditions. In addition, several new services have stringent requirements and need to be operated in real-time. This is especially true when dealing with video transmissions which can be considered one of the most demanding services in terms of delay, packet loss rate, and bandwidth. Environments that do not provide the necessary conditions will have poor performance in this type of service and

the video impairments will be very noticeable to the end-users.

Section 4.2 presented several concepts about error-correcting codes. The first method discussed was the ARQ, where packets are retransmitted if lost or after a determined time-out time. The advantage of this technique is that it adds very little network overhead, i.e., only small messages on the feedback channel. The disadvantage is that it can add a delay in the packets delivery. If the delay is high, the packet may be received after its play-out time, meaning that it can no longer be used. This trait is not appropriate for delay-sensitive applications with real-time transmissions, such as the video services.

The FEC method, instead of retransmitting the lost information, adds an amount of redundant data (or parity bits), allowing the receiver rebuild the original error-free data if something is lost or damaged. These types of error correction are widely diffused and have been used to improve the reliability of storage data, as well as to shield the communication in error-prone channels. There are several FEC codes with different characteristics and applications. One of the most used is the Reed-Solomon code because of its simplicity and real-time features.

This chapter also presented a state of the art revision about the mechanisms that improve the video transmission quality (Section 4.3). Two types of mechanisms were discussed, non-EC based and EC-based. The former does not use any type of error correction, being limited the amount of improvement that it can provide, especially in high-error rate networks. The latter adopts error correction, which can be ARQ or FEC and a combination between both as well.

As evidenced in the review there is a lack of QoE-driven, Video-aware, and high-resolution ready mechanisms. This can be explained by a large number of interfering components (e.g., video characteristics, network status, how the human vision perceives video quality) and the complex relationships between them. Taking everything into consideration, these EC methods have to be well configured to improve the video quality without adding unnecessary redundancy or delay.

In the case of FEC-based mechanisms, a number of open issues arose. To start off, one challenge is how to provide a reliable way to calculate a proper FEC block size. Smaller blocks are more efficient in slower upload links and are more resilient to burst packet losses. The downside is that they lead to a lower FEC encoding efficiency. On the other hand, bigger blocks, have better encoding efficiency but they are more susceptible to burst losses.

Once decided the FEC block sizes, it is necessary to configure the encoding FEC algorithm which will generate the redundant data. This is a challenging process because it involves the selection of the appropriate encoding rates to ensure that the missing data can be reconstructed. If a low level of redundancy is added, a small amount of network overhead is produced, however, it may not be possible to reconstruct the original error-free block due to the limited information. Conversely, if a high level of redundancy is chosen, these FEC blocks will be more resilient to packet loss, however, a larger network overhead

occurs.

On top of that, different encoding rates can be used to protect distinct categories of data in determined network conditions. For example, high rates can be used to protect the most QoE-sensitive data when the network presents high PLR. On the other hand, low rates can be used if the network is healthy. In addition, the mid-level data on the scale of importance can be protected with low encoding rates even if the network has a high-errors incidence as they will not have great impact if lost.

# Chapter $5$

# Mechanisms for Resilient Video Transmission over WMNs

> You can't consume much if you sit still and read books.
>
> *(Aldous Huxley, Brave New World)*

## Contents

The video delivery over wireless networks is now a part of the daily life of users since it is a solution that delivers a wide range of information. Despite the issues previously outlined, 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 [Zhu, 2011, Akyildiz and Wang, 2005]. Meanwhile, several difficulties can impair the

success of the transmission, such as limited network resources and high error rates, as well a fluctuating signal strength that may lead to variable bandwidth. The use of these error-prone networks unveils the need for an adaptive mechanism to improve the video transmission. Adaptive FEC-based techniques that are able to assure a high QoE for end-users are a convenient means of delivering video data to wireless users in this case. This chapter describes and assesses three proposed FEC-based adaptive mechanisms to shield video transmissions over WMN.

## 5.1 Introduction

The usage of online video services has been increasing rapidly in recent years, particularly from wireless mobile devices [Adobe Digital Index, 2016]. This upswing is related to several technological improvements in mobile devices, as well as the rise in popularity of this type of service. For example, several companies are using live video services to reduce costs and increase both collaboration and productivity.

Following the same trend, the number of non-professional users creating, sharing, and consuming online videos is growing apace. Cisco predicted that over 82% of all Internet traffic by 2020 will be some sort of video content [Cisco, 2016]. As an example, roughly a million minutes of IP video will be crossing the network every second. This figure means that a single individual would have to spend over 5 million years to watch the entire video content transmitted in only one month. Because of the video traffic ascendance, the probability of errors arising from interference and network congestion increases. This unveils the need for an adaptive mechanism to shield the video delivery, otherwise, the above factors will impact on the video quality, degrading the QoE for the end-users.

Taking this into consideration, new mechanisms for increasing the transmission quality are required to support the growth of video traffic, as the video quality may be affected by several factors. Some of them are owing to the video characteristics, such as codec type, bitrate, format, and the length of the GoP, as well as, the content/genre of the video [Yuan et al., 2006].

In addition, when packet losses occur, the perceptual quality is not harmed in the same way by all the packets. This happens 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 encouraging the use of UEP-based mechanisms. In addition, the video content also plays an important role during the transmission. Videos with a small degree of movement and fewer details tend to be more resilient to packet loss. In contrast, videos with higher levels of detail and movement are more susceptible to losses and the impairments will be more noticeable [Khan et al., 2010].

Provided that most of the video services are real-time applications, they need a steady and continuous flow of packets. This constraint can be affected by a

number of factors, especially 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 [Lindeberg et al., 2011]. Nevertheless, as afore-stated 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 [Liu and Liao, 2009]. For this reason, it is important to optimise 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 an adjustable data protection is of crucial importance to enhance video transmission, providing both well-perceived video quality and low network overhead. FEC-based schemes have been successfully used in real-time systems [Nafaa et al., 2008]. FEC allows robust video transmission through redundant data, which is sent along with the original set. As a result, if some information is lost, the original data can be reconstructed through the redundant information [Lee et al., 2011]. However, the resources might be limited and unfairly distributed as well. An adjustable FEC-based mechanism must use UEP schemes to reduce the volume of redundant information. In the UEP approach, the amount of redundancy is chosen in accordance with the relevance of the protected data and thus giving better protection to the most important video details.

Several techniques can be adopted to enable the use of UEP. These techniques can vary from heuristic-based mechanisms to sophisticated schemes using Random Neural Networks (RNN) [Abraham, 2005] and Ant Colony Optimization (ACO) [Dorigo et al., 1996]. In heuristic-based mechanisms, the steps to solve a problem usually come from the knowledge of past problems with similar characteristics. The main objective is not to guarantee an optimal or perfect solution, but a practical and/or quick method that offers a satisfactory outcome to the proposed problem [Pearl, 1984].

RNNs can also be used to the same end. They are a type of Neural Network (NN) that provides an information-processing paradigm based on the central nervous system. Through a learning process, based on pattern reading and connection weight adjustment, it is possible to configure these techniques for a specific application, being widely used in pattern recognition and data classification [Mohamed and Rubino, 2002, Aguiar et al., 2012a].

Another option is the use of ACO, which is a probabilistic algorithm, based on the behaviour of ants. This strategy is used to dynamically solve computational problems by finding the best path in a graph. In this method, ants span through the paths between the nodes to find a solution. In every path followed, a pheromone marker is deposited. At the end, the paths with the greater amount of pheromone represent the best-fitted solutions [Dorigo et al., 2006].

Considering the aforementioned issues, this chapter describes the design and evaluates three adaptive FEC-based mechanisms. The first one is the adaptive cross-layer VIdEo-aWare FEC-based Mechanism with Unequal Error Protection scheme (ViewFEC) in Section 5.2. The second proposed mechanism is the

adaptive Video-aware Random Neural Networks (RNN) based mechanism (neuralFEC), in Section 5.3. The last one is the QoE-driven motion- and video-aware mechanism (PredictiveAnts), presented in Section 5.4.

## 5.2 Adaptive Video-aware FEC-based Mechanism (ViewFEC)

Motivated by the open issues afore-identified, this section proposes and validates the adaptive cross-layer VIdEo-aWare FEC-based Mechanism with Unequal Error Protection scheme (ViewFEC).

### 5.2.1 ViewFEC Overview

The aim of ViewFEC 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 an 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.

In the ViewFEC mechanism, decisions are made at the network layer resorting to two modules, the Cluster Analysis Knowledge basE (CAKE) and the Cross-LAyer inforMation (CLAM). The decision-making process at the network layer provides deployment flexibility and allows the ViewFEC mechanism to be implemented in access points, routers, or video servers. The analysis of the information obtained from these two modules enables the proposed mechanism to estimate the most advantageous redundancy ratio necessary to sustain a good video quality, without adding unnecessary network overhead.

Figure 5.1 depicts the ViewFEC mechanism. There are three distinct stages. In the first stage, the proposed mechanism identifies several key video characteristics, such as motion and complexity levels, as well as the 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. By the aid of these details, the ViewFEC mechanism will be able to correctly identify the video characteristics needed to configure the amount of redundancy in the next stage. The offline process is important because it allows a fast and more accurate real-time execution since few variables need to be handled. The construction of the FEC blocks and the UEP redundancy assignment takes place in the third stage. Further details of each module are described later.
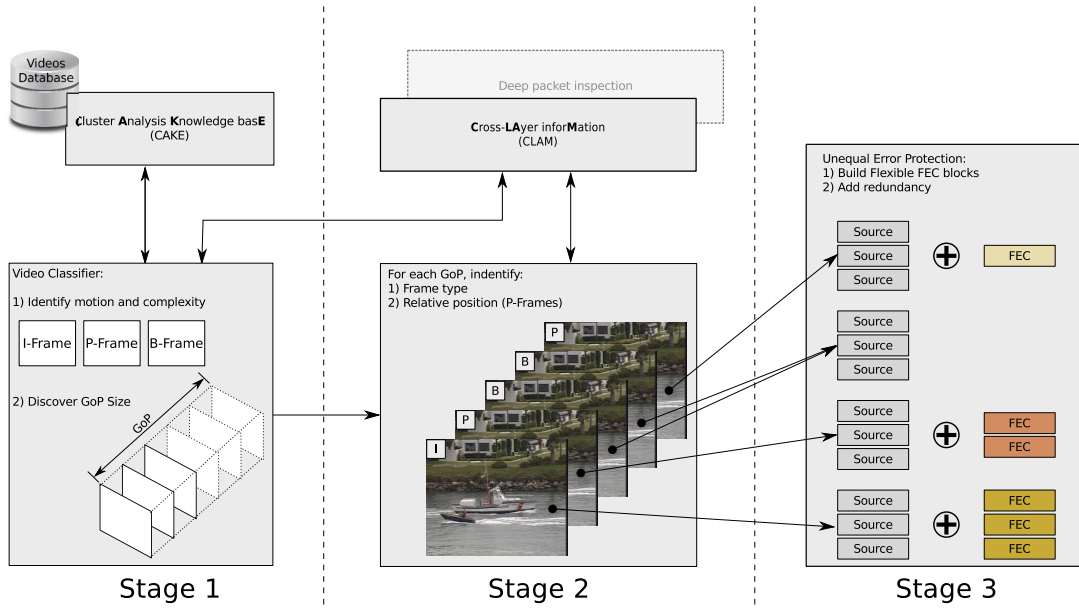
Figure 5.1: ViewFEC stages

## 5.2.2 Towards the design of ViewFEC

First of all, an exploratory analysis using hierarchical clustering is performed to conceive a knowledge database in the CAKE module. The hierarchical clustering is a statistical method of partitioning data into groups that are as homogeneous as possible [Revelle, 1979]. 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. This directory also stores information about the relation between video characteristics and their impact on video quality.

The video sequences of the experiments were chosen in compliance with the recommendations of the VQEG [Staelens et al., 2011] and ITU [ITU-T J.247, 2008]. 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. The video sequences also contain distinct temporal and spatial details, luminance stress, and still and cut scenes.

Video motion and complexity are commonly classified into three categories, namely low, medium, and high [Khan et al., 2010, Aguiar et al., 2012b] (see Figure 5.2 at linkage distance (ld) 1). Nevertheless, throughout the experiments, videos with both medium and high complexities behaved roughly the same. Therefore, the linkage distance of the cluster analysis algorithm was chosen to only produce two clusters (Figure 5.2 at ld 2). The linkage criterion is

responsible for setting the observation distance between the nodes of the cluster. In other words, this is how the elements will be grouped together. The proposed mechanism also employs the Ward method which seeks to reduce the sum of squares between the samples inside the cluster, this better reflects the experiment's findings.
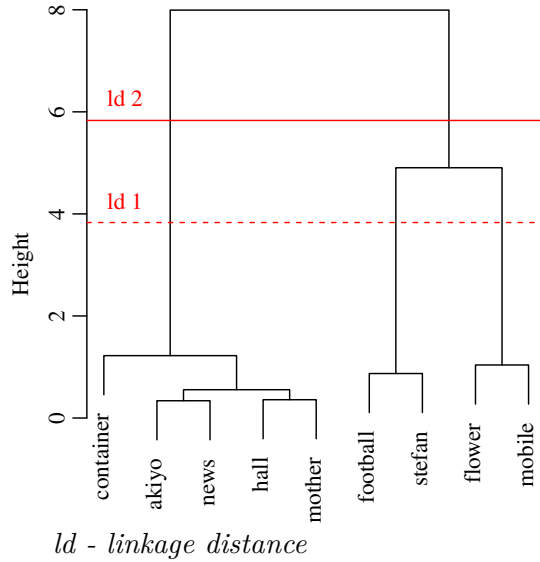


Figure 5.2: Cluster Dendrogram

The relationship between motion and complexity levels, as well as frame size (in bytes) and frame position, is shown in Figures 5.3 and 5.4. The former depicts the Mobile video sequence and the latter the Akiyo video sequence, one from each motion cluster. Only the first GoP of each video was considered, which made it easier to visualise the results. The Mobile sequence has uninterrupted scene modification and a wide-angle camera, and thus, high motion and complexity levels. For this reason, the video has larger frames and also a greater difference, in terms of size, between P- and B-Frames, as shown in Figure 5.3. In contrast, the Akiyo video sequence has only a small region of interest that is moving, which is concentrated around the face and shoulders, and also a static background. As a result, it has a low motion and complexity levels, leading to a smaller difference in size between P- and B-Frames, as depicted in Figure 5.4.

Additionally, both figures show the SSIM scores with frames been deliberately discarded. The measurement of this metric is fairly simple, even though it is consistent with the human visual system, and yields good scores [Wang et al., 2004]. The SSIM results were acquired by removing the frame which occupied that position, i.e. the first SSIM value was calculated without the first frame, and the same is applied to all other frames. In the Mobile video results (Figure 5.3), it is clear that I- and P-frames have greater significance, and also that the frames closest to the beginning of the GoP have more impact on the video quality when discarded. As expected, the Akiyo sequence behaves differently (Figure 5.4). It has a lower motion and complexity levels being more resilient to packet loss and achieving higher SSIM values [Khan et al., 2010]. The CAKE module is
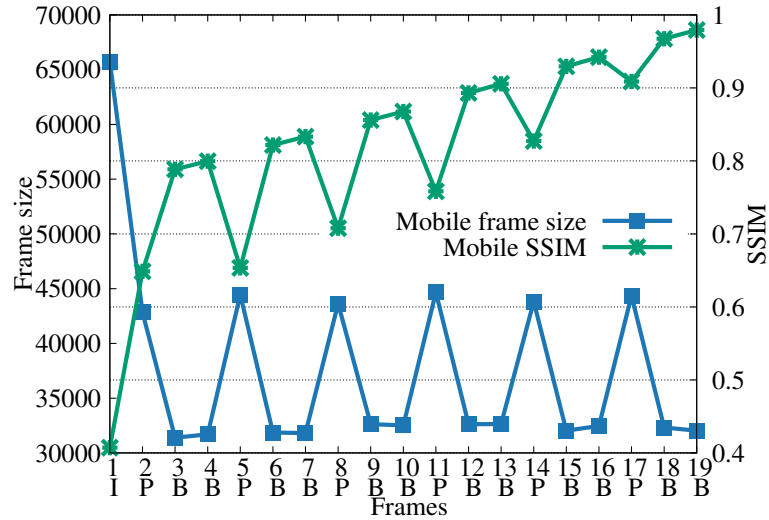
Figure 5.3: Frame size x QoE (SSIM) - Mobile video sequence
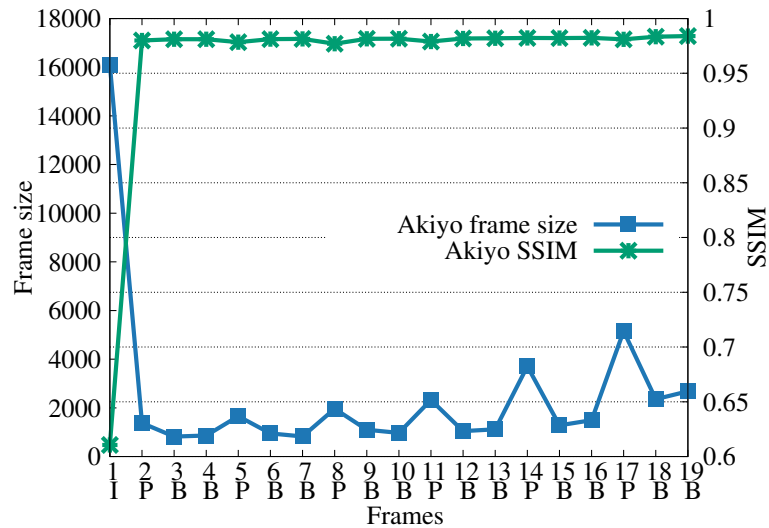
Figure 5.4: Frame size x QoE (SSIM) - Akiyo video sequence

aware of these video characteristics and can determine the motion activity and complexity levels of each GoP that is being transmitted. This is done because it is possible to have a different motion and complexity levels inside the same video sequence, as expected for Internet videos.

### 5.2.2.1 Video-aware Information Module

The next step (Stage 2) is the CLAM module. The functions of this module are implemented using cross-layer techniques. This allows accessing information from the application layer, such as the video characteristics, to the network layer, where the module is deployed. It has three basic functions. The first one 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.

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. Another important remark is that video sequences with a high level of spatial complexity tend to have larger I-Frames in 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 mean 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 increases, the size of both B-, and especially P-Frames, also increases.

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.

In addition, the ViewFEC mechanism has a flexible structure, making possible to swap 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 analysing the packet header of some of the protocols, such as UDP, RTP and TS, it is possible to discover information about codec type and coding parameters, among others [Schulzrinne et al., 2003]. On the other hand, the video content information can only be accessed through deep packet inspection scheme.

### 5.2.2.2 Defining an Adaptive Redundancy Amount

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. The RS code was adopted because it offers less complexity, and consequently achieves a better performance for real-time services [Neckebroek et al., 2010]. Nevertheless, any other alternative scheme could be used if needed. 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 mechanism settles the parity code in real-time. In other words, both $n$ and $h$ parameters are adjusted at Stage 3. This is done 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 unique 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. Consequently, 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 most beneficial 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.

The reduction of the network overhead is important, because as the network grows larger, the number of concurrent transmissions increases, and this may cause serious interference problems. The situation gets worse if more overhead is added 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.

Algorithm 5.1 shows a pseudo-code of the ViewFEC operation. It illustrates how the GoP length and motion detection are performed, and also, the steps taken to assign a tailored amount of redundancy. The algorithm has two nested loops which are responsible for going through the GoP and also the frames within each GoP. For each frame, several conditions are assessed. 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. At the end, is only added an adjusted amount of redundancy according to each frame's characteristic.

Equation 5.1 shows 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 5.2. If $\gamma > 0$, some level of redundant information will be added to the frame under discussion. 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. The notation used in the equations is shown in Table 5.1.

---

**Algorithm 5.1:** ViewFEC pseudo-code

---

**1** *for each GoP do*
**2** $\quad$ CAKE.getGopMotion(GoP);
**3** $\quad$ CLAM.getGopLength(GoP);
**4** $\quad$ *for each frame do*
**5** $\quad\quad$ *case (CLAM.getFrameType(frame)) do*
**6** $\quad\quad\quad$ *I-Frame:*
**7** $\quad\quad\quad\quad$ buildFFBlock(frame);
**8** $\quad\quad\quad\quad$ addRedundancy(frame);
**9** $\quad\quad\quad\quad$ sendFrame(frame);
**10** $\quad\quad\quad$ *P-Frame:*
**11** $\quad\quad\quad\quad$ CLAM.getRelativePosition(frame);
**12** $\quad\quad\quad\quad$ buildFFBlock(frame);
**13** $\quad\quad\quad\quad$ addRedundancy(frame);
**14** $\quad\quad\quad\quad$ sendFrame(frame);
**15** $\quad\quad\quad$ *B-Frame:*
**16** $\quad\quad\quad\quad$ sendFrame(frame);
**17** $\quad\quad$ *end case*
**18** $\quad$ *end for*
**19** *end for*

---

Table 5.1: 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 |

$$R_{GoP} = \sum_{i=0}^{GoPLength} \left[ FS_i \times FT_i \times C_{GoP} \times \frac{1}{RP_i} \right] \tag{5.1}$$

$$FT_i = \begin{cases} \gamma > 0 & \text{, send frame with redundancy} \\ 0 & \text{, frame without redundancy} \end{cases} \tag{5.2}$$

The parameter $C_{GoP}$ in Equation 5.3 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 $(1, 0.5)$, for example. This means that 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 vectors could be defined as ($\alpha High, \alpha Medium, \alpha Low$), with elements $(1, 0.5, 0.25)$, for example. In this configuration, three levels of motion intensity will be addressed, high, medium and low, respectively.

$$C_{GoP} = \begin{cases} 1 & , \text{ high motion and complexity} \\ 0 \leq \alpha < 1 & , \text{ otherwise} \end{cases} \tag{5.3}$$

$RP_i$ is the last parameter in Equation 5.1, 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, especially in video sequences with larger GoP length.

The total amount of redundant information within a video sequence can be computed by 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 using Equation 5.4.

$$\bar{R} = \frac{1}{N_{GoP}} \sum_{i=0}^{N_{GoP}} R_{GoP(i)} \tag{5.4}$$

## 5.2.3 ViewFEC Performance Evaluation and Results

The main objective of the ViewFEC mechanism is to reduce the network overhead introduced by FEC-based schemes while maintaining videos with an acceptable level of quality.

### 5.2.3.1 Experiment settings

The assessment of the benefits and impact of ViewFEC on WMNs were carried out by using NS-3. The evaluation scenario comprises six nodes distributed in a grid form (3x2); each node is 90 meters away from its closest neighbour [Oh and Chen, 2010]. Optimized Link State Routing Protocol (OLSR) [Clausen and Jacquet, 2003] was used as the routing protocol, although any other protocol can be used, such as the Hybrid Wireless Mesh Protocol (HWMP) [Bahr, 2007]. A Constant Bit Rate (CBR) was set as background traffic at 800 kbps and ten video sequences were used in the evaluation scenario [Fitzek and Reisslein, 2012], with Common Intermediate Format (CIF) size (352x288), H.264 codec. The GoP size was set at a 19:2 ratio, which means that it contains one I-Frame at every 19 frames and one P-Frame after each two B-Frames. A Frame-copy technique is used as error concealment method, which means that the decoder will replace each lost frame with the last good one received. Table 5.2 shows the simulation parameters.

Apart from the background traffic, a two-state Markov chain model was implemented to better reflect the network environments in practice. This model is also known as the Gilbert-Elliot loss model. This is considered a realistic model to simulate network losses because of its burst loss pattern representation, which

Table 5.2: ViewFEC Simulation parameters

| Parameters | Value |
|---|---|
| Display size | CIF - (352 x 288) |
| Frame rate mode | Constant |
| Frame rate | 29.970 fps |
| GoP | 19:2 |
| Video format | H.264 |
| Codec | x264 |
| Container | MP4 |
| Error concealment method | Frame-copy |
| Wireless standard | IEEE 802.11g |
| Propagation model | FriisPropagationLossModel |
| Background traffic | 800 kbps CBR |
| Routing Protocol | OLSR |
| Number of nodes | 6 nodes (grid of 3x2) |
| Error model | Gilbert-Elliot |

is commonly found in wireless channels [Wilhelmsson and Milstein, 1999]. Figure 5.5 shows this model which is composed of two states ($G$, $B$) and has four parameters ($PG$, $PB$, $r$, $k$). In state $G$, the highest number of packets must be received correctly. The $PG$ parameter holds a probability that indicates when a packet has successfully arrived without any kind of errors. The opposite is true for state $B$, and the parameter $PB$ indicates the probability of a packet being lost or damaged. The state transition is given by $k$ and $r$, where $k$ represents the probability of transition from state $G$ to $B$, and $r$ represents the probability of transition from state $B$ to $G$.

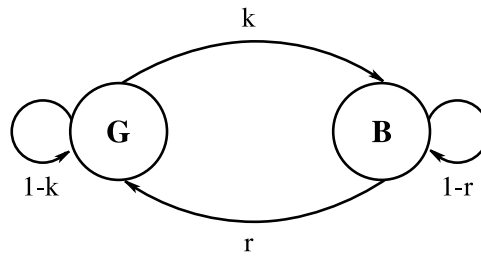

Figure 5.5: Gilbert-Elliot loss model

It is possible to compute the steady-state probability for the $G$ and $B$ states, $\varphi G$ and $\varphi B$ respectively, for $0 < k, r < 1$ with Equations 5.5 and 5.6.

$$\varphi G = \frac{r}{(r + k)} \tag{5.5}$$

$$\varphi B = \frac{k}{(r + k)} \tag{5.6}$$

It is also possible to calculate the average packet loss probability, $\bar{P}_{loss}$, using

Equation 5.7.

$$\bar{P}_{loss} = PG \times \varphi G + PB \times \varphi B \qquad (5.7)$$

Three scenarios with different mechanisms were assessed. The first experiment, which served as a baseline, was carried out without any enhancement (Without FEC). The second was implemented with a non-adaptive Video-aware FEC approach (Video-aware FEC), where a fixed amount of data redundancy (80%) was statically added to both I- and P-frames. This amount of redundancy was selected according to an extensive set of experiments, which showed the best video quality situation taking into consideration the characteristics of the scenario defined for the experiment. Finally, the last scenario adopted the proposed adaptive approach with unequal error protection (ViewFEC). Each one of these three experiments was simulated 20 times with different packet loss patterns due to distinct initial seeds for random number generation [Salyers et al., 2008] used by the Gilbert-Elliot model. The average loss was approximately 20%.

The video quality obtained in the different evaluation scenarios was assessed through objective and subjective measurements. The objective metrics used to assess the video quality were SSIM and VQM since there was a lack of correlation in PSNR values according to the subjective human perception. Both SSIM and VQM are among some of the most widely used to this end [Chikkerur et al., 2011]. The SSIM analyses the structural similarity, contrast and luminance of the transmitted images to rank it according to the likeness from the original data. Values closer to one represent better video quality. VQM uses a discrete cosine transform to assess the spatial-temporal property of the human visual system, allowing it to evaluate the image distortion. Values closer to zero represent better video quality. The objective quality assessment of the video sequences was carried out with EvalVid [Klaue et al., 2003] and MSU Video Quality Measurement Tool (VQMT) [Vatolin et al., 2011].

Additionally, a subjective evaluation was performed based on the MOS with single-stimulus and ACR. The subjective experiments were conducted using a Desktop PC with Intel Core i5, 4GB RAM and a 21" LCD monitor, with an application that displays the video sequences and collects the user scores. All the sequences are played in a random order in the middle of the screen. A neutral grey background is displayed to avoid distracting the attention of the observer. 25 observers participated in the experiments; they all had normal vision and their ages ranged from 18 to 45. The observers included undergraduates, postgraduate students, and university staff.

### 5.2.3.2 QoE assessments

Figure 5.6 shows the average SSIM scores of all the video sequences. In the SSIM metric, values closer to one indicate better video quality. It is possible to notice that when there is an increase in packet loss rate, there is a sharp decrease in the video quality of sequences that are being transmitted without

any type of protection mechanism. At the same time, video sequences that are using either type of FEC-based mechanisms, are able to maintain a good quality. Another important aspect that is worth highlighting, 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 resilience to a certain amount of packet loss. Generally speaking, video sequences with low spatial and temporal complexities are more resilient to losses, achieving better results in the QoE assessment. Other sequences, with high spatial and temporal complexity, had poorer results and despite the similar average, the standard deviation was higher with a packet loss rate of 10%. In other words, the QoE assessment values that were obtained are more distant from each other.
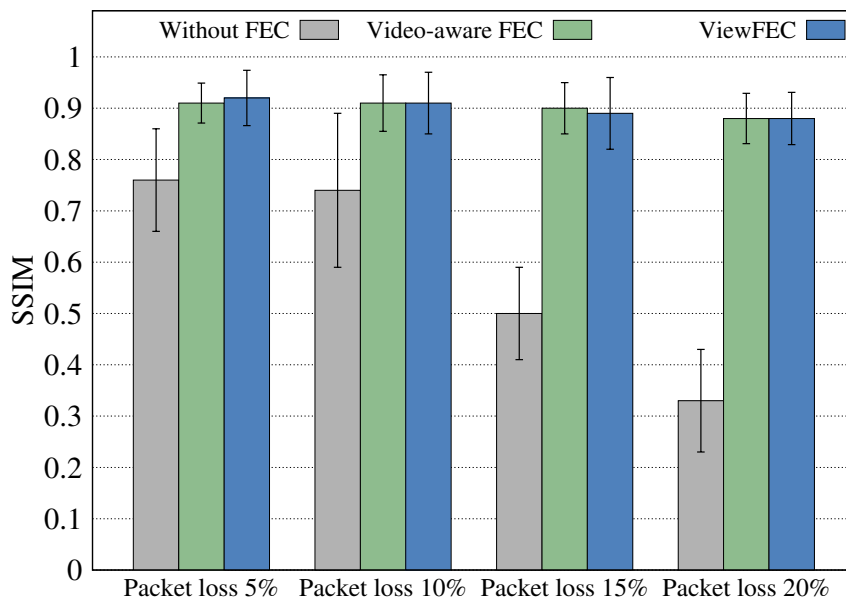


Figure 5.6: Average SSIM values for all the video sequences

Almost the same pattern is discernible in the VQM values showed by Figure 5.7. In this metric, videos with a better quality score close to zero. With a packet loss rate of 5% or 10%, the VQM values are also very close to each other. This is not so evident as in the SSIM metric because VQM tends to be more rigid with regard to video impairments, because of that, videos with fewer flaws have poorer results. For the same reason, the standard deviation of this metric tends to be higher than the SSIM metric.

Figure 5.7 also shows a large standard deviation for the scenario without FEC. The reason for this is that the VQM metric tends to rapidly increase the negative scores of the impaired videos. If no error correction technique is employed, some videos will have more defects than others. Therefore, because this metric has the tendency to assign really low scores to videos with high degrees of impairments, the sequences transmitted without error correction will receive fairly bad scores, resulting in a larger standard deviation.

A detailed analysis was conducted with the scenario that the proposed mechanism achieved the best results (20% packet loss rate). Figure 5.8 shows the
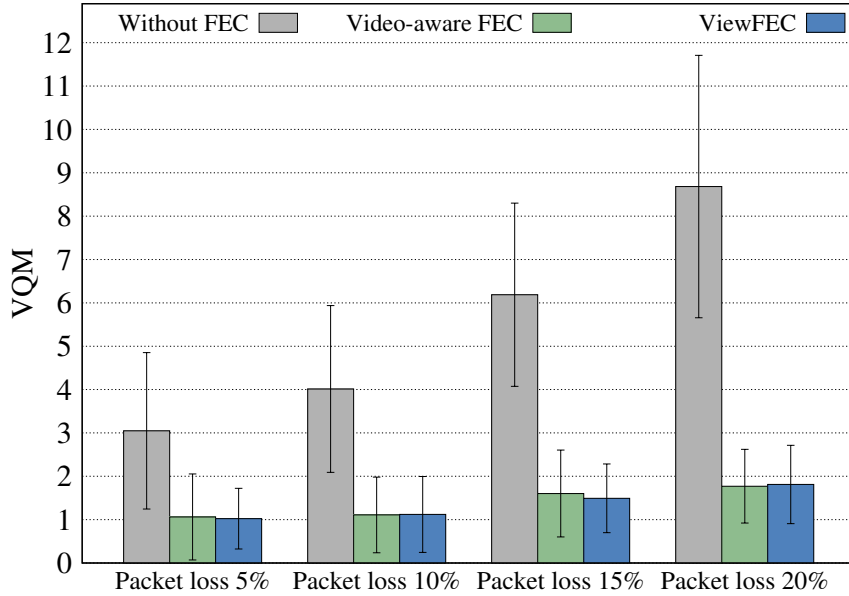
Figure 5.7: Average VQM values for all the video sequences

results of the subjective experiments. Without using a FEC-based scheme to protect the transmission, the average MOS was 2.05, which is considered poor video quality with annoying impairments. When the non-adaptive Video-aware FEC and ViewFEC mechanisms were employed, the MOS average values were 4.39 and 4.37, respectively. These values are between good and excellent quality, with perceptible but not annoying impairments. The results showed that one of the objectives of ViewFEC mechanism had been attained, which was to maintain the video quality.
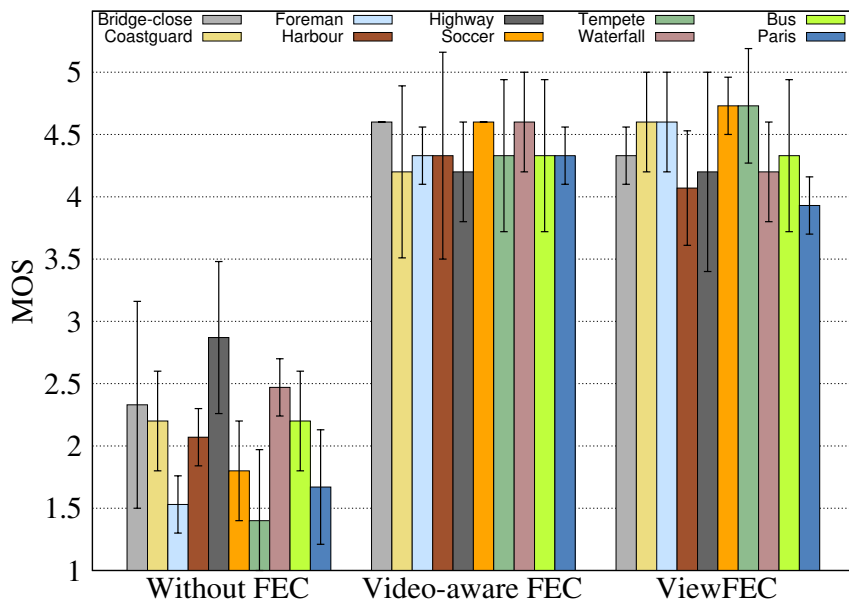


Figure 5.8: Average MOS per video sequence

Figure 5.9 and 5.10 show the SSIM and VQM values of each video sequence

used in the tests. The SSIM average value (when the FEC schemes were not used) was 0.33 and the VQM value was 8.68, representing low-quality levels and confirming what was found in the subjective assessment. On the other hand, the SSIM average of the non-adaptive Video-aware FEC and ViewFEC mechanism was 0.88, and the VQM values were 1.81 and 1.77, respectively. These scores show a good video quality, once again, corroborating the subjective findings. The distinct QoE scores reached by the different video sequences in the experiments are due to the unique characteristics of each video. As mentioned before, small differences in motion and complexity levels can influence the obtained values. In view of this, it is important to make use of several types of videos when conducting the experiments.
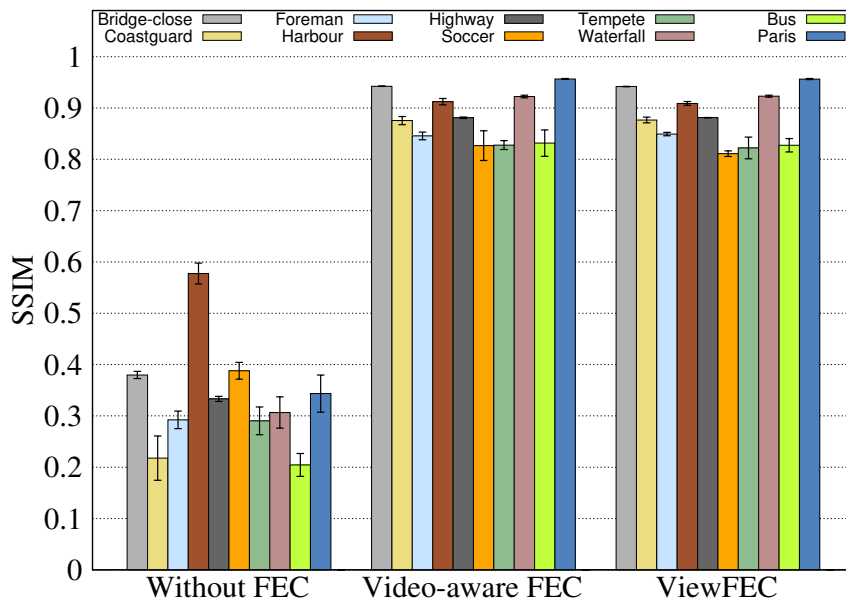


Figure 5.9: Average SSIM per video sequence

### 5.2.3.3 Network footprint analysis

All the QoE assessments demonstrated that the ViewFEC mechanism was able to maintain a good video quality. Nevertheless, the main goal of this mechanism was to reduce the network overhead. This is important in wireless networks, due to the limited channel resources, uneven bandwidth distribution, and interference caused by concurrent transmissions. In the experiments, the network overhead is the sum of all video frames transmitted, which includes the redundant data, minus the original frame size. The non-adaptive Video-aware FEC mechanism was able to deliver videos with a network overhead between 53% and 78%, as shown in Figure 5.11. On the other hand, when the ViewFEC mechanism was used, the network overhead remained between 34% and 47%. The ViewFEC mechanism imposes, on average, 40% less network overhead than the non-adaptive Video-aware FEC, with equal or slightly better video quality, as shown in Figure 5.8, 5.9 and 5.10.
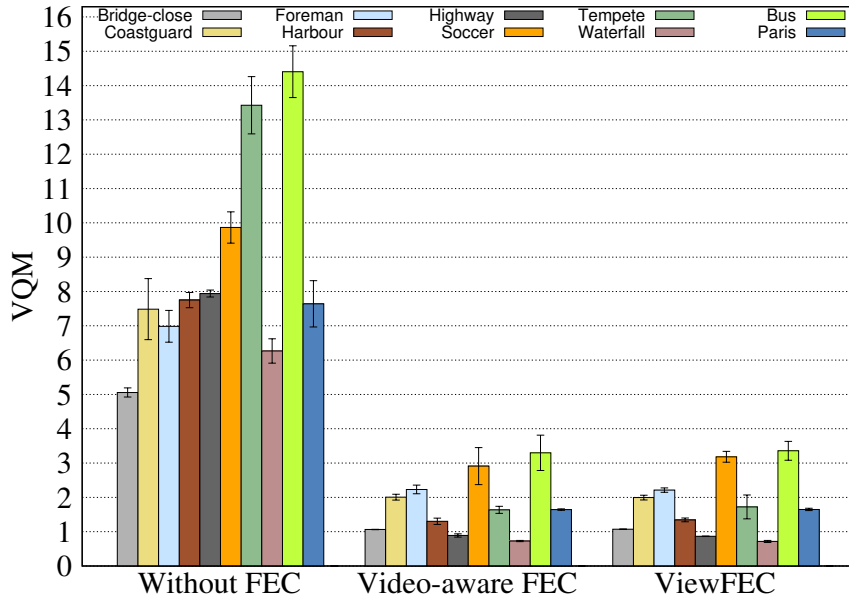
Figure 5.10: Average VQM per video sequence
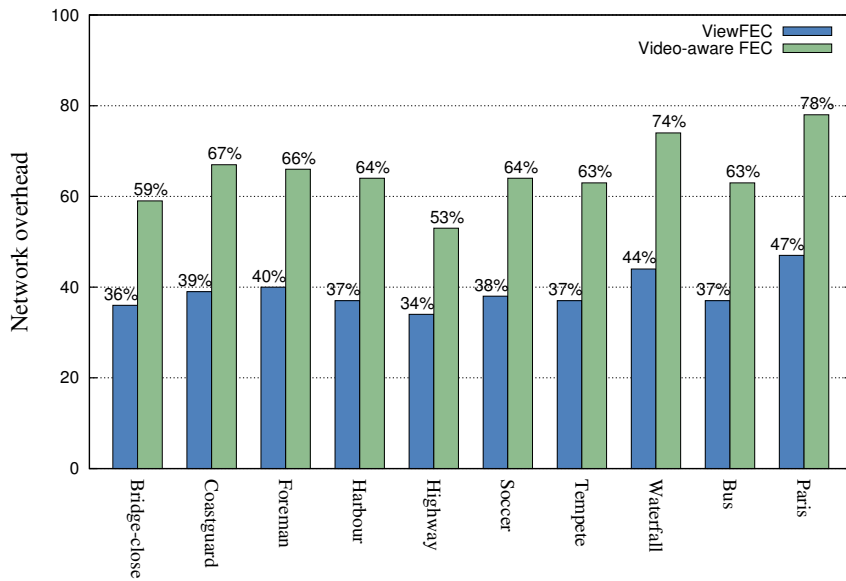


Figure 5.11: Network overhead (%)

The video sequence Highway had the smallest reduction in network overhead (36%) and the Coastguard sequence had the largest (42%). This can partially be explained by the size of the I-, P- and B-Frames. Figure 5.12 shows the size of the frames of all videos. The y-axis shows the number of sent packets, including the redundant information. When the Highway values are analysed, one can notice that over 61% of the packets belong to B-Frames, which are not considered in either the non-adaptive Video-aware FEC or in ViewFEC, because they lead to minor impairments if lost. This means that less than 39% of the packets are optimised by the ViewFEC mechanism, and this result in a smaller

reduction in the overhead. Conversely, the Coastguard sequence has more than 46% of the packets in I- or P-Frames, which can be optimised, resulting in a greater reduction of network overhead.
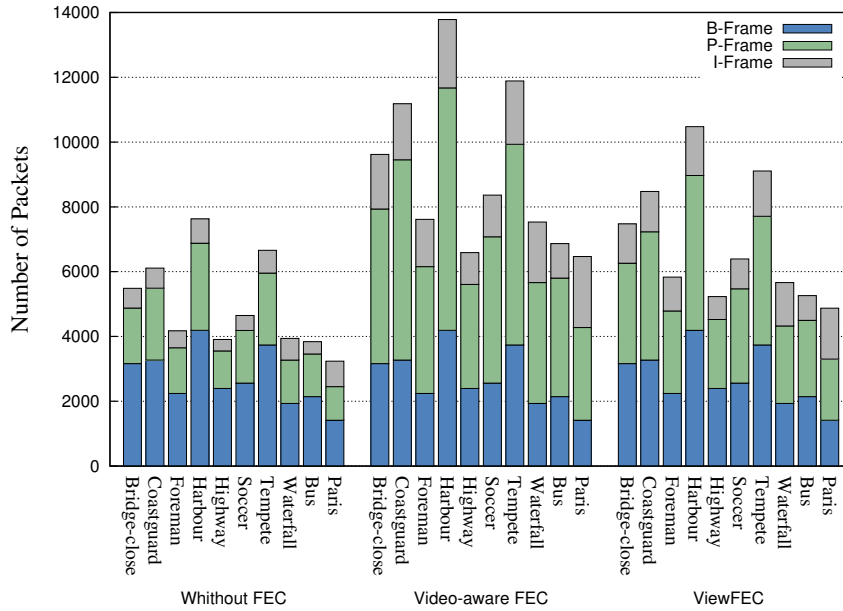


Figure 5.12: Network overhead in packets

### 5.2.3.4 Overall results

Table 5.3 summarises the results and shows the improved video quality (in percentage) in each scenario for both mechanisms. In the VQM metric, the more negative the percentage showed in the table is the better the result (e.g., ↓79.60%). On the other hand, the opposite is true for the SSIM metric. As expected, both FEC-based mechanisms produce more valuable results when the network has a higher PLR. When the average PLR was 20%, for example, it was possible to achieve a reduction of over 79% in VQM values; this means $\approx 4.9$ times smaller scores. With the SSIM metric, there was an increase of over 166% in the results, meaning $\approx 2.66$ times higher values.

Table 5.3: ViewFEC 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% |

Taking into consideration the results of the experiments, it is possible to say that the proposed ViewFEC mechanism showed good performance. Additionally, it

also highlights the fact that it is feasible to improve the QoE of video sequences delivered over wireless networks.

# 5.3 Video-aware and RNN-based mechanism (neuralFEC)

The mechanism described in Section 5.2 provided positive results, however, several issues were identified during the experiments. Most of these are related to the heuristic used in the process to classify the motion intensity. A more accurate technique to perform this task could improve the mechanism.

Using more information about the video sequences on the classification method can also be beneficial. An efficient way to quantify the pace of action is through motion vectors. These vectors play a key part in the video compression process, allowing the system to store changes from adjacent frames, including both previous and future frames. Therefore it is possible to quantify the motion intensity of a given frame using the information inside its vectors. Taking these issues into account, this section proposes the adaptive Video-aware Random Neural Networks (RNN) based mechanism (neuralFEC). This mechanism was proposed in collaboration with MSc student Pedro Borges, which performed his work under my co-supervison.

## 5.3.1 neuralFEC Overview

The neuralFEC mechanism aims at overcoming the limitations of non-adaptive schemes, such as the inability to take into consideration a precise classification of the video's motion intensity, which is crucial to a high QoE. The mechanism proposed in this section mitigates these problems by adaptively selecting the amount of redundancy given to individual frames, showing better results than the previously proposed mechanism. The adaptive redundancy is chosen according to the analysis of the frame type and the motion characteristics using a RNN [Abraham, 2005]. Neural Networks are computational models inspired by biological central nervous systems, which are able to go through the process of machine learning and pattern recognition. The NN can be trained by feeding to it learning patterns and letting it change the weights according to some learning rule. Random Neural Networks are known for they success in pattern recognition and classification problems [Mohamed and Rubino, 2002], making this type of NN suitable for the proposed mechanism.

Figure 5.13 depicts the overall operation of the neuralFEC mechanism. First of all, the same procedure adopted by ViewFEC has been used again in the offline process. This means that an exploratory analysis using hierarchical clustering was carried out, however, here the results are used to train the RNN. The RNN was validated through the human experience about the intrinsic video characteristics and several simulation experiments. After this, the RNN can be used

in real-time. The decision-making process conducted by the RNN determines a specific amount of redundancy needed by each frame. This allows the neural-FEC to shield only the QoE-sensitive data against packet loss, resulting in better video quality as perceived by the end-users, while saving network resources. A detailed explanation of the proposed mechanism is presented afterwards.
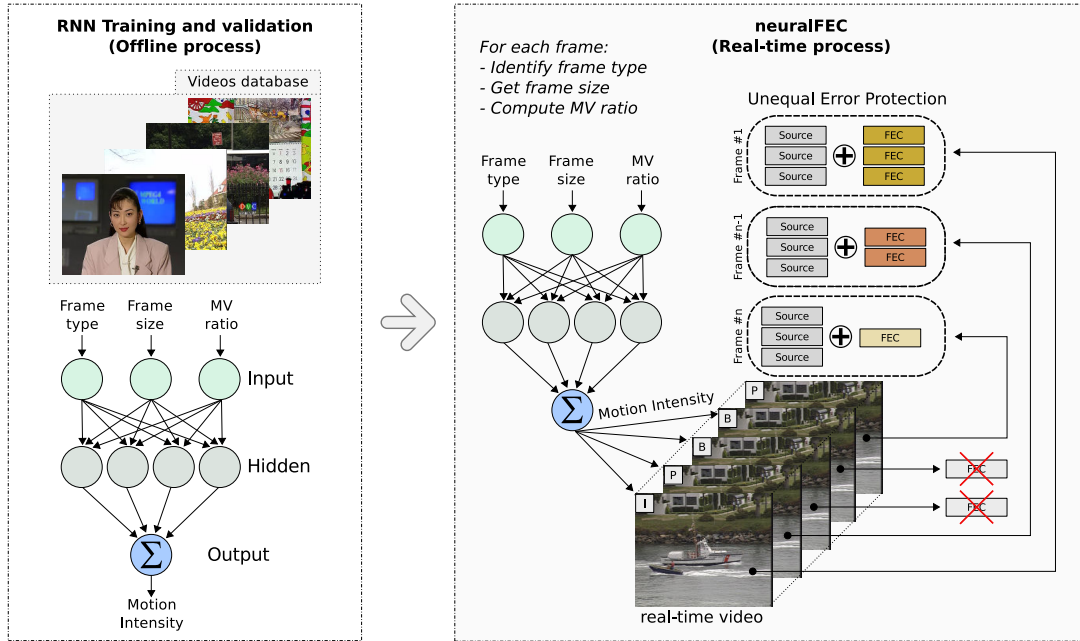


Figure 5.13: neuralFEC mechanism

## 5.3.2 Towards the design of neuralFEC

In order to perform the classification of each frame according to its motion intensity, a RNN was employed. As other NNs, this method has the capability for learning and generalisation, providing particularly good results in pattern recognition and classification problems [Mohamed and Rubino, 2002]. By training the network successfully with an adequate range of video samples, it can be used in real-time to classify a given video sequence according to the intensity of the movement. This is achieved by attributing a specific value to determine frames with different motion intensities. After that, the neuralFEC is able to select, in real-time, the appropriate amount of redundancy to be transmitted so that the network overhead is minimised and QoE is maximised.

The RNN structure consists of three input nodes, seven hidden layer nodes and one output. The three input nodes represent each frame's characteristics specifically frame size, frame type, and motion vectors ratio (the total number of vectors divided by the distance described by them). The ratio was used because a certain frame can have several vectors pointing to a close distance while other frames can have fewer vectors pointing further away, however, and consequently defining a situation of higher motion intensity. Finally, the output node provides the motion intensity classification value, computed by the network

from the given inputs. Through these parameters, it is possible to characterise the video motion intensity and choose the optimal amount of redundancy on a frame-by-frame basis.

To adequately train the RNN, an exploratory hierarchical cluster analysis using Ward's method [Ward Jr, 1963] was performed to categorise selected video sequences which represent different types of movement. As always, the video sequences were selected according to the recommendations of the VQEG [Staelens et al., 2011] and ITU [ITU-T J.247, 2008]. A set of 15 videos was selected to perform the hierarchical cluster analysis. Each video was broken down into three parameters, namely about the frame size and type, and motion vectors ratio.

Using the exploratory analysis results, the video sequences were classified into three categories of motion intensity, namely low, medium and high intensity. Afterwards, two videos of each motion intensity category were randomly selected to train the RNN. The training of the RNN consisted in feeding the information of this set of selected videos to the inputs of the network for about 600 iterations which were the point at which the Minimum Mean Squared Error (MMSE) stabilised. After the training period it was validated with a different set of video sequences, the remaining 9 videos from the exploratory analysis, which also cover all three motion intensity characteristics.

After the training and validation phases, the RNN can be used in the real-time process. Using cross-layer techniques, the neuralFEC mechanism is able to obtain important information about several video characteristics, namely frame type and size, as well as the number of motion vectors and the Euclidean distance pointed by them. All these details are fed to the RNN which in turn provides, in real-time, an accurate motion intensity value for each frame.

Once the video frame is classified, it is encoded with the amount of redundancy selected by the RNN. Through this procedure, a precise UEP amount can be assigned to each frame, where only the QoE-sensitive data will be protected. The result of this is better video quality while reducing the amount of redundancy data needed. Therefore, not adding unnecessary redundancy will allow more users to access services, improving the overall system performance.

Algorithm 5.2 shows the pseudo-code portraying the neuralFEC real-time operation. All procedures are performed inside a for-loop, at line 01, which will go through all the frames in the video sequence. At line 02, the frame type is identified to be used in the selection control mechanism (if statement) at line 03. This allows the change in the control flow according to neuralFEC needs, which is to assign a tailored redundancy amount to I- and P-Frames, and send B-Frames without additional data. At lines 04, 05, 06, and 07 it is possible to observe the identification of the frame size, the computation of the motion vectors ratio, the classification of the video frame motion intensity using the RNN, and the assignment of an unequal amount of redundancy to the most QoE-sensitive data, respectively. Lines 08 and 10 are responsible for sending the frame with or without redundancy.

---

**Algorithm 5.2:** neuralFEC pseudo-code

---

**1** **for** *each Frame* **do**

**2**      FT $\leftarrow$ GETFRAMETYPE(*Frame*);

**3**      **if** *(FT equal (I- or P-Frame))* **then**

**4**          FS $\leftarrow$ GETFRAMESIZE(*Frame*);

**5**          MVratio $\leftarrow$ CALCULATERATIO(getMV(*Frame*));

**6**          MotionIntensity $\leftarrow$ RNN(*FT, FS, MVratio*);

**7**          ADDREDUNDANCY(RS(*MotionIntensity*));

**8**          SENDFRAME(*Frame + Redundancy*)

**9**      **else**

**10**          SENDFRAME(*Frame*)

**11**      **end**

**12** **end**

---

### 5.3.3 neuralFEC Performance Evaluation and Results

The neuralFEC goal is to use an optimised UEP scheme according to the video frame's motion intensity characteristics, making it possible to better protect the most QoE-sensitive frames. Therefore, providing both the reduction in the impact of packet loss on the video quality and the use of only the necessary amount of redundancy. Through this, it is possible to distribute the redundancy in a way that will improve the QoE for the end user, while sparing network resources.

#### 5.3.3.1 Experiment settings

In order to assess the performance of the proposed mechanism in wireless networks, several experiments were performed by using NS-3. The scenario for evaluation is comprised of 25 nodes in a grid disposition (5x5), separated by 50 meters. The OLSR was used as the routing protocol. Ten video sequences were used in this scenario, namely Bowing, Coastguard, Container, Crew, Foreman, Hall, Harbour, Mother and Daughter, News and Soccer. These particular sequences were selected in order to have a great variety of motion intensities. They are in CIF with a resolution of 352x288 and coded with the H.264 codec. The GoP size was set to 19:2 thus after each I- or P-frames, come two B-frames. Table 5.4 shows the simulation parameters.

A two-state discrete-time Markov chain model was implemented following a simplified Gilbert-Elliot packet-loss model [Razavi et al., 2009], which approximates the behaviour of a wireless network. It produces simulation results which are closely related to those of burst loss patterns of wireless channels [Wilhelmsson and Milstein, 1999]. This model differs from the Gilbert-Elliot used in Section 5.2 because the probability of packet loss in the Good state (G) is set at 0, which means no losses. On the other hand, the probability of packet loss in the Bad state (B) is set at 1, where all packets are lost. The PLR can be obtained

Table 5.4: neuralFEC Simulation parameters

| Parameters | Value |
| --- | --- |
| Display size | CIF - (352 x 288) |
| Frame rate mode | Constant |
| Frame rate | 29.970 fps |
| GoP | 19:2 |
| Video format | H.264 |
| Codec | x264 |
| Container | MP4 |
| Error concealment method | Frame-copy |
| Wireless standard | IEEE 802.11g |
| Propagation model | FriisPropagationLossModel |
| Background traffic | 800 kbps CBR |
| Routing Protocol | OLSR |
| Number of nodes | 25 nodes (grid of 5x5) |
| Error model | simplified Gilbert-Elliot |

by Equation 5.8, where $P_{BG}$ represents the probability of transitioning from the Bad state to the Good state and vice-versa with $P_{GB}$.

$$PLR = \frac{P_{BG}}{P_{BG} + P_{GB}} \qquad (5.8)$$

To validate and compare the results, three experiments with different mechanisms were performed. The first experiment serves as a baseline as there was no FEC mechanism in use. The second experiment was performed with a non-adaptive video-aware FEC mechanism (Video-aware FEC), where a fixed amount of 38% of redundancy was added only to I- and P-frames. This amount of redundant data was chosen after several experiments and represents the best video quality in the considered scenarios. The last experiment is the proposed adaptive mechanism with RNN and UEP (neuralFEC). Each of the three experiments was simulated 10 times with an error rate of 20% representing an average loss [Immich et al., 2013b] obtained through a simplified Gilbert-Elliot packet loss model.

The video quality of each evaluation scenario was assessed through an objective measurement, namely the SSIM. The objective quality assessment of the video sequences was performed with EvalVid and the MSU VQMT.

### 5.3.3.2 Network footprint analysis

Figure 5.14 shows the results in terms of network overhead. The non-adaptive Video-aware FEC mechanism had an overhead between 35% and 43%. On the other hand, when the neuralFEC mechanism was employed, the amount of overhead remained between 13% and 24%. This means that the average redundancy added by the non-adaptive mechanism was around 38% in contrast to only 19%

added by neuralFEC. It is also clear that the proposed mechanism can assess the importance of frames according to motion intensity. This assessment is performed by the RNN, which attributes a higher classification for frames with a great amount of movement, and a lower classification for frames with less amount of movement. In doing that, a greater amount of redundancy was attributed to video sequences such as Crew, Soccer, Harbour and Coastguard. On the contrary, video sequences which are classified as being of lesser motion intensity, such as Bowing, Mother and Hall are given less redundancy. These results show that the neuralFEC mechanism performs better than the non-adaptive Video-aware FEC mechanism in terms of overhead, by reducing in average a half of the redundancy needed to protect the data.
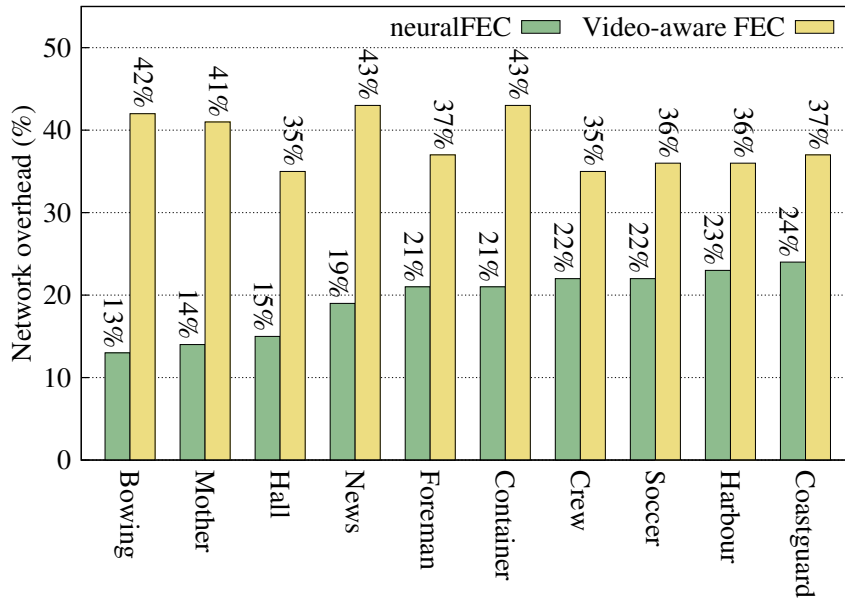


Figure 5.14: Network Overhead

### 5.3.3.3 QoE assessments

Besides saving the already scarce network resources by not adding unnecessary redundancy, it is also important to provide good video quality. In order to verify this situation, a set of assessments was performed using the SSIM metric. Figure 5.15 depicts the SSIM values for each video sequence while using the three aforementioned protection schemes. The results show the neuralFEC mechanism obtained an average SSIM value of 0,831 against a value of 0,819 for the video-aware FEC mechanism and 0,726 for the mechanism that did not use any type of protection. This represents a slight improvement of around 1,5%, on average, in terms of SSIM value for the adaptive neuralFEC mechanism in comparison to the non-adaptive video-aware FEC mechanism. In further detail, the SSIM score achieved by neuralFEC for the Harbour video sequence was of 0,675 against 0,662 for the video-aware mechanism and 0,485 for the mechanism without FEC. Although all videos were transmitted with the same PLR, the

SSIM score obtained by the same three mechanisms for the Bowing sequence was of 0,915, 0,914, and 0,920, respectively. This can be explained by the different characteristics of these two sequences. The Harbour video sequence has a greater amount of motion compared to the Bowing sequence, meaning that packet loss has a greater effect on this type of sequences. This results in lower SSIM scores for sequences with a higher degree of motion intensity and also shows that videos with a lower degree of motion intensity have greater resilience to packet loss. Due to this, it is important to employ adaptive FEC mechanisms, such as neuralFEC to protect the contents of the video taking into account its motion intensity characteristics.
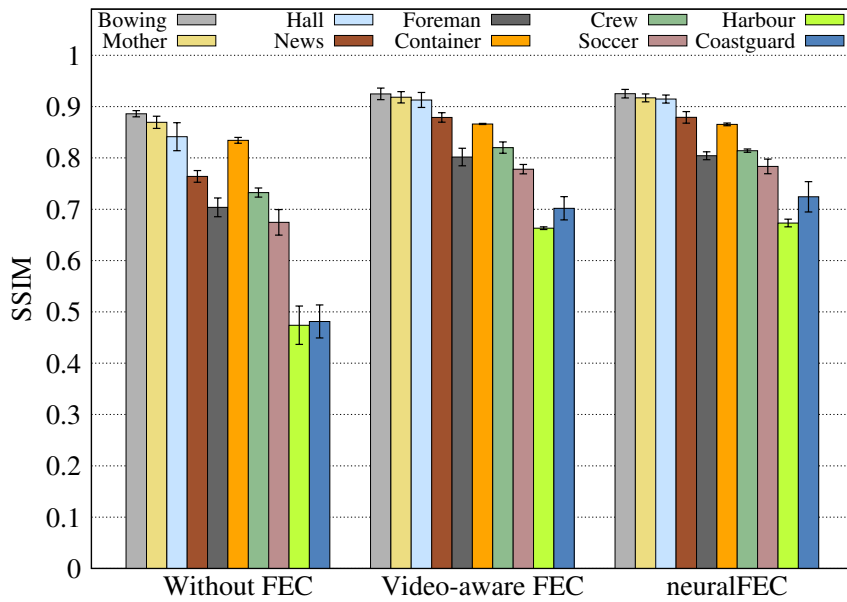


Figure 5.15: neuralFEC Objective QoE assessment (SSIM)

### 5.3.3.4 Overall results

Table 5.5 summarises the results presenting the average SSIM and network overhead for all video sequences. It demonstrates that the proposed neuralFEC mechanism had a slightly improved video quality. Most importantly, it was able to do so while drastically reducing the network overhead by not adding unnecessary redundancy. This is of great importance in wireless networks, due to the limited nature of the channel resources, which can be aggravated by packet loss due to interference from concurrent transmissions and network congestion.

Table 5.5: Average SSIM and network overhead of neuralFEC

|  | neuralFEC | Video-aware FEC | Without FEC |
|---|---|---|---|
| SSIM | 0,831 | 0,819 | 0,726 |
| Overhead | 19,334% | 38,460% | – |

The results showed that the neuralFEC mechanism, through an accurate motion intensity classification of video sequences with distinct characteristics, is able to add a precise amount of protection. In doing that, it can offer less overhead during transmission in a wireless mesh network setting while providing as good video quality as non-adaptive FEC mechanisms.

# 5.4 QoE-driven Motion- and Video-aware Mechanism (PredictiveAnts)

The aforementioned mechanism was able to provide a notable outcome including a slightly better video quality as well as the reduction of the unnecessary network overhead. Nevertheless, there is still room for improvement in the classification method, by using a higher number of video characteristics. Additionally, both ViewFEC and neuralFEC lack the capability to assess the network conditions to adjust the redundancy amount, which can have a significant impact.

Considering the open issues aforementioned, particularly the absence of motion-aware mechanisms that use a broad amount of video characteristics together with a network assessment, this section presents the proposed QoE-driven motion- and video-aware mechanism (PredictiveAnts). This mechanism was proposed in collaboration with MSc student Pedro Borges.

## 5.4.1 PredictiveAnts Overview

The adaptive FEC-based mechanism proposed in this section uses several video characteristics and packet loss rate prediction to shield real-time video transmission over wireless mesh networks, improving both the user experience and the usage of resources. This is possible through a combination of a RNN, to categorise motion intensity of the videos, and an ACO scheme, for dynamic redundancy allocation, allowing the protection of the most important information.

Figure 5.16 depicts the PredictiveAnts mechanism. It is composed of two processes, one is performed offline and the other one in real-time. In the same way as in the other mechanisms, the Offline process is responsible for the RNN training and validation steps. The main objective of the RNN is to characterise the motion intensity of video sequences according to several inputs, such as the frame type and size, the number of motion vectors, and the Euclidean distance pointed by these vectors. Since it is an offline process, it needs to be executed only once. After that, the RNN can be used in real-time. The offline process is important because it allows a fast and more accurate real-time execution since few variables need to be handled.

The real-time process consists of several modules, each one having very peculiar tasks, as follows:
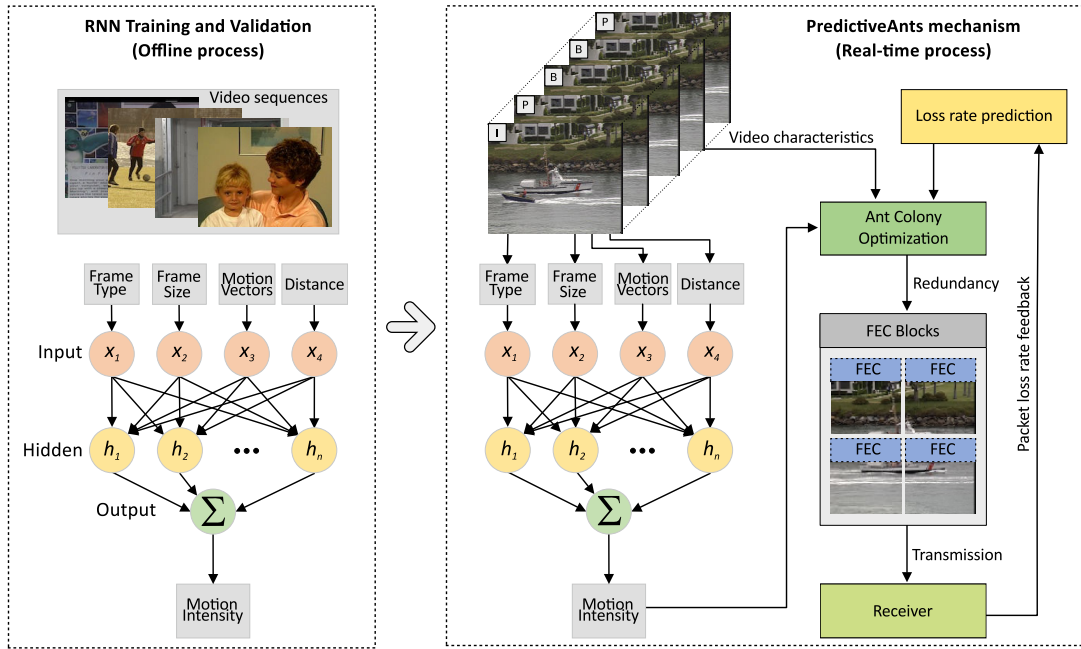
Figure 5.16: PredictiveAnts mechanism

- MOTION INTENSITY - The motion intensity characterization is performed by the RNN in real-time since it was already trained and validated in the offline process;

- FEEDBACK RECEIVER - The feedback mechanism is responsible for the retrieval of loss statistics. The information is collected by the receiver and sent to the transmitter;

- LOSS RATE PREDICTION - Using the feedback statistics, the properties of the error probability are estimated on the server side;

- VIDEO CHARACTERISTICS - This module fetches information from the video sequences that are being transmitted to identify video characteristics such as the frame type and size, as well as the motion vectors;

- ANT COLONY OPTIMIZATION - The ACO is responsible for making a joint analysis of all the information gathered by the other modules, establishing the most suitable amount of redundancy to each FEC block;

- FEC BLOCKS - The FEC blocks are built and a specific amount of redundancy designed by the ACO is assigned to each one.

The ACO is a probabilistic algorithm, based on the behaviour of ants, used to dynamically solve computational problems by finding the best path in a graph. In this solution, ants span through the paths between the nodes to find a solution. In every path followed, a pheromone marker is deposited. At the end, the paths with a greater amount of pheromone represent the best-fitted solutions [Dorigo et al., 1996].

## 5.4.2 Towards the design of PredictiveAnts

As previously mentioned, the PredictiveAnts mechanism comprises several processes and modules that are going to be detailed in this section. First of all, in the same way as in the previous mechanism, it is necessary to train and validate the RNN for the motion intensity categorization. The RNN is composed of four input nodes, seven hidden nodes and one output node. The input nodes are the frame size, the frame type, the number of motion vectors and the Euclidean distance described by the vectors. The hidden nodes are generated through a stochastic process whereas the output node gives the motion intensity value. The RNN was trained using a set that comprises distinct motion scenarios and validated with a different set. The selection of the sets was performed through an exploratory hierarchical cluster analysis to group the video sequences according to the motion intensity. Several video characteristics are used in this analysis, such as frame type, frame size and motion vectors. The results are well-defined clusters that can be used in the RNN. As aforementioned, the offline process needs to be executed only once and after that, the RNN can be used in real-time.

After the offline process, all further computations are done in real-time. One of the main improvements of this proposed solution is the design and use of a simple error prediction scheme. This is performed, instead of just using the instantaneous network loss rate, to attribute a customised amount of redundancy to the video sequence being transmitted. The use of a loss prediction scheme enables a further reduction of added overhead, by balancing the allocation of redundancy data between the network's good and bad states.

Forecasting future events is a very important task, as the predicted data can be used as input for the decision-making process. There are several proposals to forecast the PLR using, for example, time-series, sparse basis models, and hidden Markov models. However, as it is needed a fast mechanism, which should be able to run in real-time, a simpler model was designed. Thereby, the error prediction scheme was developed based on the concept of good and bad gaps [Karner et al., 2007]. In order to do that, a feedback mechanism was implemented to enable the retrieval of loss statistics. This feedback information comprises the distribution of good and bad gaps during transmission. As shown in Figure 5.17, a good gap is defined as the interval of packets that were successfully received between two bad gaps (white squares). A bad gap is the interval of packets during which a burst of errors is occurring (red squares). The feedback information is collected by the receiver and sent to the transmitter in the form of a vector containing the size of every gap of each type. From the measured statistics, the characteristics of the error probability can be computed on the server side. This information is used as a predicting value of a higher probability of the occurrence of an error in the next block of packets to be transmitted. Therefore, the error prediction scheme has an influence on ACO, leading to an adjustment in redundancy based on the prediction of the occurrence/non-occurrence of an error.

Another enhancement in PredictiveAnts is the use of FEC blocks. This means that, instead of performing redundancy allocation on a frame-by-frame basis,
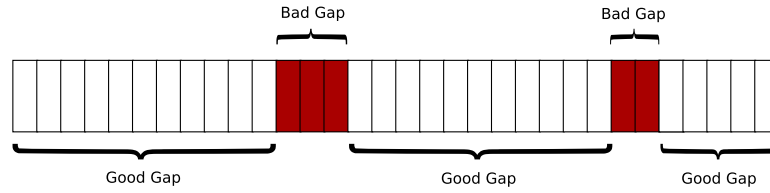
Figure 5.17: Packet gaps during transmission

the mechanism was improved to support the transmission of each frame on blocks of several packets. The size of the blocks can be adjusted according to video and/or network characteristics. Particularly in this case, the size of the blocks was set to 10 network packets. This value was selected through extensive experimentation to provide more flexibility of the mechanism while dealing with the error correcting code, and better control over the data. Furthermore, this provides enough granularity to comply with the sizes of the gaps. Therefore, the selected packet block size was that which allowed for the block to be isolated (i.e., whole block inside a good gap) from an error gap through error prediction in all packet loss rate conditions.

Once the information about the motion intensity, video characteristics and the packet loss prediction is gathered, the ACO mechanism can be defined. It also has an offline process and after that, through the aforementioned information, the exact amount of redundancy needed for each FEC block is computed in real-time. The ACO metaheuristic that needs to be defined are the construction graph, the creation of the candidate list, the delimitation of the heuristic information, and the pheromone trails definition.

### 5.4.2.1 The "Construction Graph" details

The construction graph is one of the principal elements of the ACO metaheuristic. It is used to map the problem under consideration onto a graph [Dorigo et al., 2006], so the feasible solutions are encoded as walks on the graph. This means that, as the ants traverse the construction graph, they construct a solution to the problem. In other words, the result value of the objective function in each walk corresponds to a viable solution to the original problem [Gutjahr, 2000].

A hierarchical graph was used to meet the needs of the mechanism. This means that, once the ants start to walk they can only go to the next layer and always forward. Since it is only possible to move forward and from the previous to the next layer, the construction graph is not fully connected and the number of vertices is equal to the number of layers. Additionally, the distance between the nodes is directly proportional to the amount of redundancy required to improve, or at least maintain, a superior QoE. By taking this into consideration, the construction graph is built to better reflect this condition, enabling PredictiveAnts to find the best possible solution for each scenario.

The resulting construction graph for this problem is described as a connection

graph $G_c = (C, L)$, where nodes $C$ are the components and $L$ represents the set of partially connected components $C$, also called connections. The problem constraints are given by the function $\Omega$ and follow these conditions:

(1) In $G_c$, there is only one start node and it is located at the first layer;

(2) Let $Q$ be the set of tours (complete walks) $q$ in $G_c$ which satisfy the conditions below:

    (i) $q$ always starts at the start node of $G_c$ in the first layer;

    (ii) $q$ contains exactly one node of each layer of $G_c$;

    (iii) The last node on $q$ belongs to the last layer of $G_c$;

Then $\Omega$ maps the set $Q$ onto the collection of attainable solutions for this specific problem instance. Following this definition, the construction graph $(G_c, \Omega)$ gives the set of all feasible solutions. The construction graph refers to the association of a set of QoE- and network-related parameters (e.g. motion intensity nodes, frame type and size nodes, as well as the packet loss rate nodes) with a set of vertices in the graph, meaning the connections between the nodes. Using the results of the exploratory data analysis it is possible to create an efficient construction graph.

### 5.4.2.2 The "Heuristic Information" details

The heuristic information, also called heuristic value, provides the ability to exploit problem-dependent knowledge obtained prior to the execution or at run-time if retrieved from a different source other than the ants. This information will guide the ants' probabilistic solution, meaning that the ants have to take into consideration fewer options to decide how to move on the graph. In doing that, it will strongly reduce the local search spectrum and consequently improve the solutions. Owing to this fact, the ACO algorithm is able to provide good performance in real-time.

Using the results of the exploratory data analysis, together with the knowledge database and human expertise the heuristic information $I_h$ is defined. This information is composed of the length $d_{ij}$ of the arc connecting the nodes $i$ and $j$. Therefore, it is possible to define the heuristic information as $I_h = 1/d_{ij}$. As mentioned before, the length of the *arc* is directly proportional to the amount of redundancy required to improve or at least maintain a superior QoE. For this reason, the longer the tour the ants are walking the higher the redundancy amount needed. In the PredictiveAnts mechanism, all the $I_h$ is pre-computed once at the bootstrap time, a table with all the possible values is generated, and it remains unchanged during the whole mechanism's run.

### 5.4.2.3 The "Candidate List" details

The candidate list is used to reduce the number of possible choices that have to be considered at every construction step. In order to accomplish such task, this list holds a small number of promising choices of next stop. The static lists are built utilising prior knowledge of the problem, however, they can also be generated dynamically with information gather on-the-fly. Since the proposed mechanism uses a hierarchical graph and it does not change over time, it can use a static candidate list composed of all the nodes of the next layer.

Let $L_c$ be the candidate list of any specific node, an arc $(i, j)$ is included in this list if the following conditions are met:

(1) The arc $(i, j)$ it is not already included in $L_c$;

(2) The arc $(i, j)$ establishes a connection from the origin layer to a higher one, which means that it does not create cycles or backwards links;

(3) The arc $(i, j)$ holds that $I_{h(ij)} > 0$, which implies that this connection needs to add some useful heuristic information;

The adoption of candidate lists is twofold. Fist of all, it restricts the walking path of the ants to certain conditions. This is of primordial importance to PredictiveAnts due to its hierarchical graph design, and thus not allowing the ants to walk horizontally inside the same layer, but just between the layers. Secondly, it strongly reduces the dimension of the search space of each ant, improving the real-time performance and therefore speeding up the solution process.

### 5.4.2.4 The "Pheromone Trails" details

The pheromone trails in the ACO metaheuristic help to guide the ants to make probabilistic decisions, and thus, construct possible solutions for the problem that is being solved. These trails are composed of numerical information distributed in the paths along the graph. During the algorithm's execution, the ants adapt the pheromone value to express their search knowledge.

### 5.4.2.5 PredictiveAnts ACO structure

Figure 5.18 shows PredictiveAnts ACO graph. It has fourteen nodes characterising video and network details. These nodes were chosen because they represent a combination of factors that directly affect the video quality, as follows:

- START - The first node is just the starting point. As the total amount of redundancy is given by the travelled path all ants must start at the same point;

- MOTION INTENSITY - These three nodes feature the RNN classification in terms of motion intensity, which can be low, medium and high motion;
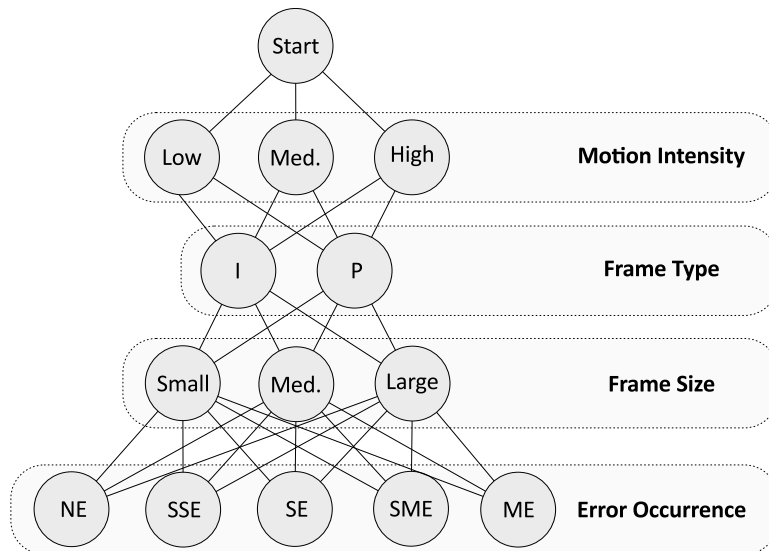
Figure 5.18: Graph used in the ACO mechanism

- FRAME TYPE - The frame type, I- or P-frame, is represented by these two nodes. These are the most important frames in the MPEG standard. The loss of one I- or P-Frame will be more noticeable by the end-user because the error will only be corrected when another I-Frame arrives, in other words, in the beginning of the next GoP. Thus, those frames need to be protected with redundant information;

- FRAME SIZE - These three nodes characterise the frame size, which can be small, medium, and large;

- ERROR OCCURRENCE - The last layer of nodes represents the possibility of occurrence of an error instead of the instantaneous packet loss rate. The five nodes represent five different scenarios which can occur:

  - NE (NO ERROR) - It is a scenario where no error is accounted for the current FEC block;

  - SSE (SHARED SINGLE ERROR) - In this scenario, a single error is predicted, which will be shared by this FEC block and the next;

  - SE (SINGLE ERROR) - A single error is predicted only for the current FEC block;

  - SME (SHARED MULTIPLE ERRORS) - It is a scenario where the occurrence of two or more error gaps is predicted in the current FEC block continuing to the next one;

  - ME (MULTIPLE ERRORS) - In this scenario, two or more blocks of errors occur only on the current FEC block.

A simple ACO model was used with both the number of iterations and ants set to 10. The values were reached through extensive experimentation to obtain a solution which did not worsen the delay of the PredictiveAnts mechanism. In run-time, the ants search the graph while leaving pheromone in the travelled

path, this reinforces the best solutions for the problem which can be re-used for similar conditions. The value computed by the ACO mechanism will be used to configure the amount of redundancy in the RS algorithm [Reed and Solomon, 1960]. This algorithm is of low complexity being suitable for real-time use. By adding a tailored amount of redundancy to each FEC block, it is possible to better protect the most QoE-sensitive data, maximising the video quality and, at the same time, minimising the network overhead.

Algorithm 5.3 shows the PredictiveAnts mechanism pseudo-code. All operations are repeated for each frame (01). The first step is to get the frame type (02). Since only I- and P-Frames are protected it is necessary to check this condition (03). If false, the frame is sent immediately (13). If true, other information about the frame is required, such as frame size (04), the number of motion vectors (05) and the Euclidean distance of the motion vectors (06). All these are input information to the RNN for motion intensity categorization (07). Once this value is found, it can be fed to the ACO mechanism (08), together with the frame type, frame size and the loss rate prediction, to compute the amount of redundancy (09). Afterwards, the FEC blocks are built using the original frame and the redundancy (10) and then the blocks are sent (11). The feedback information about the packet loss is received (15) and the loss rate prediction is calculated (16).

---

**Algorithm 5.3:** PredictiveAnts pseudo-code

---

1 **for** *each Frame* **do**
2    FT ← GETFRAMETYPE($Frame$);
3    **if** *(FT equal (I- or P-Frame))* **then**
4       FS ← GETFRAMESIZE($Frame$);
5       MV ← GETMOTIONVECTORS($Frame$);
6       MVDist ← COMPUTEDISTANCE($MV$);
7       MotionIntensity ← RNN($FT, FS, MV, MVDist$);
8       RSpar ← ACO($MotionIntensity, FT, FS, LP$);
9       Redundancy ← RS($RSpar$);
10      BUILDFECBLOCKS($Frame + Redundancy$);
11      SENDFECBLOCKS($FB1, FB2, ..., FBn$);
12   **else**
13      SENDFRAME($Frame$);
14   **end**
15   LOSS ← RECEIVERFEEDBACK();
16   LP ← CALCULATELOSSRATEPREDICTION($LOSS$);
17 **end**

---

### 5.4.2.6 Computational complexity of PredictiveAnts

The computational complexity of the main PredictiveAnts components is as follows. Equation 5.9 represents the ACO complexity of finding a solution with an expected number of iterations in a graph with $n$ nodes, $m$ edges and where $\rho$

is the evaporation rate of the pheromone used by the ants [Attiratanasunthron and Fakcharoenphol, 2008].

$$O(\ \frac{1}{\rho}\ n^2\ m\ log\ n)\ \tag{5.9}$$

Concerning Reed-Solomon, the encoding computational complexity is comprised of two steps, namely the pre-computing of the Generator Matrix (GM) of the code, followed by the multiplication of the source vector by the GM. Equation 5.10 represents the total computational complexity of the encoding per element [Lacan, 2009], where $k$ represents the rows and $n$ represents the columns of the GM matrix.

$$O(\ \frac{k}{(n-k)}\ *(log\ k)^2 + log\ k)\ \tag{5.10}$$

The decoding steps of the Reed-Solomon code involve the computation of the $k*k$ sub-matrix of the GM. Afterwards, this matrix is inverted and multiplied by the received vector in order to recover the original vector. The computational complexity [Lacan, 2009] per element of these steps is represented by Equation 5.11 where $k$ is the number of received elements.

$$O((log\ k)^2)\ \tag{5.11}$$

Overall, the PredictiveAnts mechanism has the capability to be used in realtime. Moreover, due to the accurate categorization of the motion intensity in the video sequences and the PLR prediction, the adaptive PredictiveAnts mechanism can downsize the network overhead, reducing the video delivery footprint while improving the video quality.

## 5.4.3 PredictiveAnts Performance Evaluation and Results

The PredictiveAnts mechanism aims to improve the usage of wireless network resources by reducing the overhead while assuring a good perceived video quality. The performance evaluation goal is to show that the PredictiveAnts mechanism can effectively decrease the network overhead while still providing high QoE.

### 5.4.3.1 Experiment settings

The evaluation experiments were carried out by using the NS-3. The scenario is composed of a grid of 25 static nodes (5x5), 90 meters apart from each other. The OLSR was used as the routing protocol. A data set of ten video sequences in CIF format, GoP length of 19:2, and H.264 codec was used. The selected video sequences are different from those used to train the RNN. These videos

cover different distortions and subjects, which represent content usually found in on-line video services. The Frame-Copy error concealment method was used. Table 5.6 shows the simulation parameters.

Table 5.6: PredictiveAnts Simulation parameters

| Parameters | Value |
|---|---|
| Display size | CIF - (352 x 288) |
| Frame rate mode | Constant |
| Frame rate | 29.970 fps |
| GoP | 19:2 |
| Video format | H.264 |
| Codec | x264 |
| Container | MP4 |
| Error concealment method | Frame-copy |
| Wireless standard | IEEE 802.11g |
| Propagation model | FriisPropagationLossModel |
| Routing Protocol | OLSR |
| Number of nodes | 25 nodes (grid of 5x5) |
| Error model | Simplified Gilbert-Elliot |

In order to simulate the burst loss patterns found in wireless networks [Wilhelmsson and Milstein, 1999], a simplified two-state discrete-time Markov chain scheme following the Gilbert-Elliot (GE) packet-loss model [Yu et al., 2005, Razavi et al., 2009] was implemented. By adjusting these probabilities, it is possible to generate different error patterns, which can be translated to specific PLR values. In the experiments, the values were set to 5%, 10%, 15%, and 20%, which are commonly present in wireless networks.

Five different cases were simulated as follows: (1) without any type of FEC. This case will serve as baseline to compare with the others; (2) Video-aware Equal Error Protection (VaEEP) (where both I- and P-Frames are equally protected) with a pre-defined amount of redundancy set to 38%; (3) Video-aware UEP (VaUEP), here again both I- and P-Frames are protected, this time, however, with a different amount of redundancy depending on the type. An average of 30% redundancy amount is added. It is important to notice that the protection of only I- and P- frames is a common practice in the video transmission industry. The redundancy amounts used by both VaEEP and VaUEP mechanisms were attained after a thorough set of simulation studies. They showed, in average, a good tradeoff between video quality and network overhead under the different PLR; the next case is the AntMind mechanism which uses a combination of an RNN and ACO for Unequal Error Protection [Immich et al., 2014b]; Finally, the last case adopts the mechanism proposed in this section, the PredictiveAnts mechanism.

### 5.4.3.2 QoE assessments

Two main QoE metrics were employed to carry out the video quality assessment, namely SSIM and VQM. The objective quality assessment was conducted using Evalvid and MSU VQMT.

Figure 5.19 shows the SSIM assessment for all of the video sequences in each of the five schemes. The values are an average of all PLRs for each video. The scheme without FEC averaged a value of 0,806. The VaEEP mechanism averaged a value of 0,880 and the VaUEP obtained 0,881. The AntMind mechanism had an average of 0,876 and the PredictiveAnts score 0,884, which was the highest average value. The distinct values for the different video sequences are due to the unique characteristics of each video, this highlights the need for a motion- and video-aware mechanism. These results show therefore that the PredictiveAnts mechanism offers a better video quality than its competitors. Taking this into account and the reduction of the added overhead, it is possible to say that it provides a more precise protection scheme.
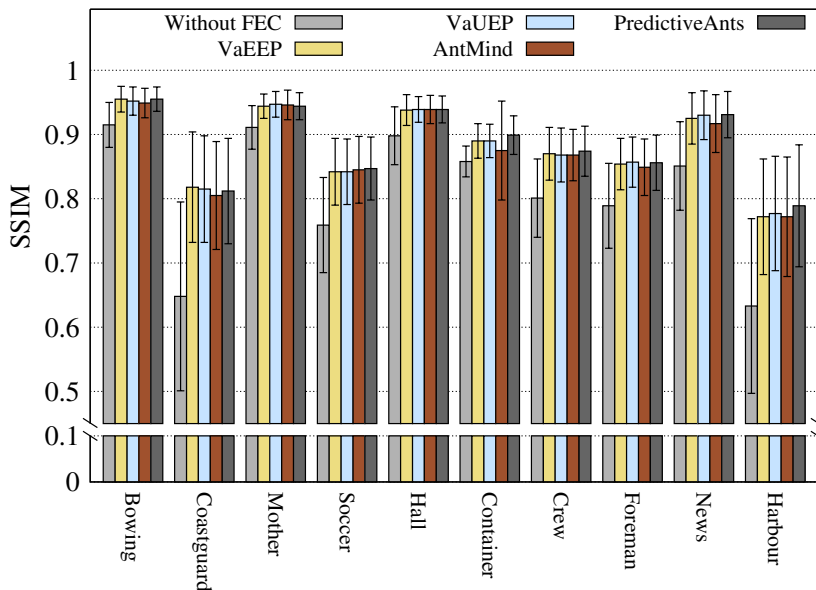


Figure 5.19: Objective QoE assessment (SSIM)

Figure 5.20 presents the VQM scores. The scheme without FEC averaged a value of 5,277. The VaEEP and VaUEP mechanisms had an average of 3,895 and 3,860, respectively. The AntMind mechanism achieved an average of 3,940 and the PredictiveAnts scores 3,664. The same way as in the SSIM assessment, the PredictiveAnts had the better video quality. This proves once again that the improvements made to the PredictiveAnts were able to reduce the network overhead, while improving the video quality.
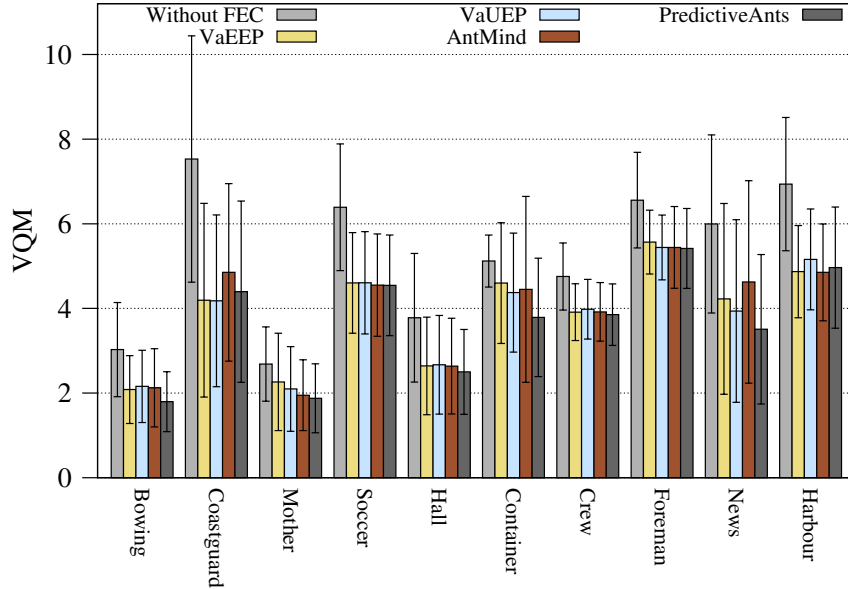
Figure 5.20: Objective QoE assessment (VQM)

### 5.4.3.3 Network footprint analysis

Figure 5.21 shows the network overhead results of all PLRs using the four FEC schemes. The first scheme, without FEC, is not shown because it does not produce overhead. VaEEP's average overhead was 38% with values ranging from 35% to 43%, and VaUEP's average overhead was 30% with values ranging from 25% to 36%. The AntMind mechanism had an average overhead of 15%, with values between 9% and 19%. This is a notable result, with an overall reduction of more than 50% in the redundancy amount (60% over VaEEP and 50% over VaUEP). The proposed PredictiveAnts was able to produce even better results, providing an average overhead of 11%, with values ranging from 7% to 13%. This represents a further improvement of on average over 27% less redundancy. This means that far less redundancy data is used by the PredictiveAnts opposed to VaEEP, VaUEP, and AntMind.

Additionally, it is worth pointing out that PredictiveAnts correctly characterises the importance of the frames according to their motion intensity details. In all video sequences, PredictiveAnts outperforms the AntMind mechanism. However, the biggest reductions in the network overhead were achieved on the video sequences which have greater amounts of motion intensity, such as Harbour (36%) and Coastguard (33%). The lowest reductions in the network overhead were found on the videos that are opposite to these two in terms of motion intensity, specifically Bowing (11%) and Mother (17%). Since the PredictiveAnts mechanism is an enhancement of AntMind, the greater gain in the overhead reduction has already been achieved. Therefore, this explains why the video sequences with lower intensities of motion had a slight reduction in the network overhead.
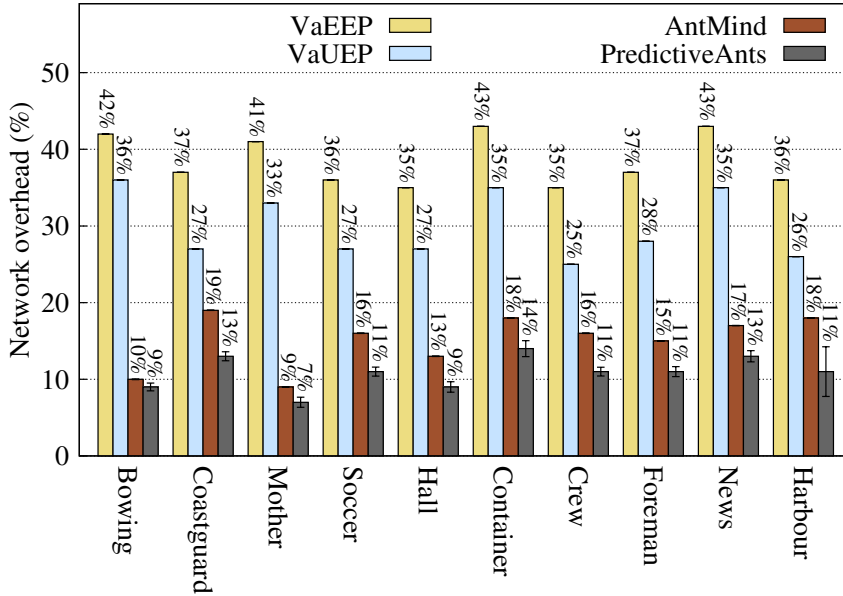
Figure 5.21: PredictiveAnts Network Overhead

### 5.4.3.4 Overall results

Table 5.7 summarises the SSIM, VQM, and network overhead results. It demonstrates that the proposed mechanism was able to considerably cut down the network overhead by not adding unnecessary redundancy. The PredictiveAnts achieved an average of 67% of redundancy reduction over the non-adaptive schemes (71% less redundancy than VaEEP and 63% less than VaUEP). It also provides a very good result over the AntMind mechanism, saving more than 27% in the network overhead. Additionally, the PredictiveAnts mechanism managed to achieve the best results in terms of video quality. These outcomes are very important in wireless environments due to the already scarce network resources.

Table 5.7: PredictiveAnts Average SSIM, VQM and network overhead

|  | PredictiveAnts | AntMind | VaEEP | VaUEP | Without FEC |
|---|---|---|---|---|---|
| SSIM | 0,884 | 0,876 | 0,881 | 0,880 | 0,806 |
| VQM | 3,664 | 3,940 | 3,895 | 3,860 | 5,277 |
| Overhead | 10,970% | 14,898% | 38,460% | 29.827% | – |

On the basis of the results referred to above, the PredictiveAnts mechanism showed that it is able to considerably reduce the network overhead. This is only possible due to an accurate categorization of the motion intensity details. Additionally, the packet loss prediction scheme allows anticipating the amount of redundancy that will be needed before the transmission. In doing that, it shields the video delivery against losses by ensuring an adequate protection to any kind of video sequence. This leads to an improved QoE for end-users.

## 5.5 Summary

This chapter described and assessed three proposed mechanisms to safeguard video transmission resiliency over WMN. The ViewFEC mechanism proposed in Section 5.2 adjusts the redundancy amount according to several video characteristics. This proactive adjustment occurs in real-time and all the information needed is gathered through cross-layer techniques. It also benefits from a motion activity database which is built before the execution of the mechanism. The main goal of this mechanism was to achieve the best video quality possible without adding unnecessary network overhead, due to the redundancy packets. This ability was confirmed through the simulation results where the ViewFEC mechanism outperformed non-adaptive FEC-based schemes in both the video quality and the network overhead. Generally speaking, the ViewFEC mechanism presented good results, it had, however, several issues in regards to the heuristic used in video classification procedure. In this case, a more precise technique could be used to better classify the motion intensity of the video sequences.

The neuralFEC mechanism (detailed in Section 5.3) improves on the above issues by using RNN in both the classification process and the decision-making steps. In doing that, it provides the possibility to shield the video transmission in wireless networks, protecting only the most QoE-sensitive data, maximising the video quality, while saving network resources by not sending unnecessary redundancy. The experimental simulation results showed that neuralFEC was able to reduce the amount of network overhead by 50% while maintaining or even improving the QoE for the end user. This is a considerable enhancement over non-adaptive FEC mechanisms and also reinforces the importance of using adaptive FEC-based scheme which takes into account motion intensity when protecting a video stream with varying characteristics. Altogether, both the ViewFEC and the neuralFEC mechanisms provided very good results, especially in reducing unnecessary network overhead. However, considering the video quality, the mechanisms still presented some deficiency that can be improved upon. Additionally, both mechanisms do not assess the network state to adjust the redundancy, which has an impact on the final result.

In the light of the aforementioned issues, the PredictiveAnts mechanism, described in Section 5.4, uses several video characteristics and packet loss rate prediction to shield real-time video transmission over static wireless mesh networks. This allows the improvement of both the user experience and the usage of resources. The proposed mechanism is based on a combination of a RNN, to categorise motion intensity of the videos, and an ACO scheme, for dynamic redundancy allocation. The experiment results evidenced that PredictiveAnts was able to enhance the video quality without adding an unnecessary amount of redundancy. In comparison to the non-adaptive mechanism, it has reduced the network overhead by 67% on average. When compared to the adaptive mechanism, it provides a further 27% savings in the network overhead. This is a great enhancement over both non-adaptive and adaptive FEC mechanisms.

It only reinforces the relevance of using adaptive FEC mechanisms, which take into consideration the motion intensity and packet loss prediction to protect a video streaming with fluctuating characteristics.

The literature review and the proposed mechanisms described in this chapter culminated in the following publications:

**Journal papers:**

- Immich, R. and Borges, P. and Cerqueira, E. and Curado, M., "**QoE-driven video delivery improvement using packet loss prediction**", International Journal of Parallel, Emergent and Distributed Systems, Volume 30, Issue 6, pp 478-493, Taylor & Francis, 2015

- Rosário, D. and Cerqueira, E. and Neto, A. and Riker, A. and Immich, R. and Curado, M., "**A QoE handover architecture for converged heterogeneous wireless networks**", The Journal of Mobile Communication, Computation and Information, Wireless Networks, Volume 19, Issue 8, pp 2005–2020, Springer, 2013

**Book chapter:**

- Immich, R. and Cerqueira, E. and Curado, M., "**Cross-layer FEC-based Mechanism for Packet Loss Resilient Video Transmission**", in Data Traffic Monitoring and Analysis: From measurement, classification and anomaly detection to Quality of experience, Volume 7754, pp 320-336, Springer LNCS, 2013

**Conference papers:**

- Immich, R. and Borges, P. and Cerqueira, E. and Curado, M., "**Adaptive Motion-aware FEC-based Mechanism to Ensure Video Transmission**", in the 19th IEEE Symposium on Computers and Communications (ISCC), 2014

- Immich, R. and Borges, P. and Cerqueira, E. and Curado,M., "**Ensuring QoE in Wireless Networks with Adaptive FEC and Fuzzy Logic-based Mechanisms**", in the IEEE International Conference on Communications (ICC), 2014

- Immich, R. and Borges, P. and Cerqueira, E. and Curado, M., "**AntMind: Enhancing Error Protection for Video Streaming in Wireless Networks**", in the 5th IEEE International Conference on Smart Communications in Network Technologies (SaCoNET), 2014

- Immich, R. and Cerqueira, E. and Curado, M., "**Adaptive Video-Aware FEC-based Mechanism with Unequal Error Protection Scheme**", in the 28th Annual ACM Symposium on Applied Computing (SAC), 2013

- Zhao, Z. and Braun, T. and Rosário, D. and Cerqueira, E. and Immich, R. and Curado, M., "**QoE-aware FEC Mechanism for Intrusion**

**Detection in Multi-tier Wireless Multimedia Sensor Networks**”, in the International Workshop on Wireless Multimedia Sensor Networks (WMSN), 2012

**Special session presentations:**

- Borges, P. and Immich, R. and Cerqueira, E. and Curado, M., “**Mechanisms for resilient video transmission in wireless networks: Adaptive FEC mechanism with random neural network classification and ant colony optimization**”. 18 Seminário Rede Temática de Comunicações Móveis (RTCM), 2014

- Immich, R., Cerqueira, E. and Curado, M., “**Cross-layer FEC-based Mechanism to ensure Quality of Experience in Video Transmission**”. 15 Seminário Rede Temática de Comunicações Móveis (RTCM), 2012

# Chapter 6

# Mechanisms for Resilient Video Transmission over FANETs

Who controls the past controls the future; who controls the present controls the past

*(George Orwell, Nineteen eighty-four)*

## Contents

U AV are rising in popularity together with video applications for both military and civilian use. This unveils the need for an adaptive video-aware mechanism capable of overcoming a number of challenges related to the scarce network resources, device movement, as well as high error rates, to ensure a good video quality delivery. Adaptive FEC-based techniques are known to be suitable to enhance the QoE of video transmitted over error-prone wireless networks with high mobility, which is an intrinsic characteristic of FANETs using UAV-to-Ground connection model. Additionally, as mentioned before, the unique characteristics of each video sequence, such as the spatial complexity and the temporal intensity, strongly affect how the QoE will be impacted by

the packet loss. This chapter describes two adaptive video-aware mechanisms to safeguard UAV real-time video transmissions against packet loss, providing a better user experience, while saving resources.

## 6.1 Introduction

The rapid growth of both, autonomous and nonautonomous UAV [Kumar et al., 2001], with the objective of video surveillance, exploitation, and reconnaissance is evident in the last years. The deployment of these vehicles is no longer exclusive of military and special operation applications, as the civilian use of small UAVs has also increased due to ease operation, robust, and cost-effective wireless networking technologies, such as 4th Generation Networks (4G) Long-Term Evolution (LTE).

The adoption of UAVs can be helpful in a broad range of situations, more often than not replacing fixed video cameras due to their mobility and low-cost operation in contrast to manned systems. Some examples of UAV applications are in traffic surveillance, sports events, festivals, public parades, or at any place that has the potential of gathering plenty of people [Puri, 2005, Bekmezci et al., 2013]. It is also worth highlighting the use for monitoring and inspection of critical infrastructures, such as harbours, large industrial areas, railways, long pipelines, power plants, as well as to cover large areas with lack of infrastructure, such as interior border control, countryside properties or even in natural disaster sites and rescue missions [Bernard et al., 2011].

The benefits of UAV with video capability are clear, however, even with proper equipment, robust data integration and visualisation tools, poor-quality video streaming can compromise the usability of the system. These video streams are watched by humans, and a good quality is essential to, for example, identify faces, damaged power lines or pipes, as well as track conditions. However, as mentioned before, real-time video transmissions with ensured QoE are resource-demanding services, especially over wireless networks.

Incidentally, FANETs tend to have poor connectivity quality [Frew and Brown, 2009]. In addition, channel conditions can quickly fluctuate over time owing to the high mobility of the nodes and terrain structures, as well as other wireless communication issues like noise, multipath fading, and channel interference [Lindeberg et al., 2011].

Another challenge is to fairly use the available bandwidth [Liu and Liao, 2009]. It is critical to make an efficient use of resources preventing the induction of network congestion and a high packet loss rate. This is especially important in resource-consuming services like video transmission. Therefore, an optimised distribution of live video streams with QoE support is one of the main challenges in highly dynamic wireless environments, such as FANETs with UAV-to-Ground model. Choosing the proper adaptive redundancy control mechanism with QoE and network-awareness is decisive for an efficient use of resources while increasing

the video quality as perceived by end-users.

This chapter describes two video-aware mechanisms that use motion vectors details, FEC, and Fuzzy logic to improve the resilience of UAV video transmission with both UEP and QoE-awareness. The first one is the cross-layer adaptive video-aware FEC mechanism (uavFEC), in Section 6.2, and the second is the Motion INTensity and video-aware mechanism (MINT-FEC), in Section 6.3.

# 6.2 Adaptive Video-aware Fuzzy Logic Mechanism (uavFEC)

Considering the open issues aforementioned this section describes and evaluates the proposed cross-layer adaptive video-aware FEC mechanism (uavFEC). The proposed mechanism aims to supply an alternative considering the lack of QoE- and video-aware proposals which include clear indicators of motion intensity, as well as, the network conditions. The main goal of the proposed mechanism is to enhance the video transmission of small UAV.

## 6.2.1 uavFEC Overview

The uavFEC mechanism uses motion vectors details, FEC, and Fuzzy logic to improve the resilience of UAV video transmission with both UEP and QoE-awareness. The uavFEC mechanism dynamically configures itself, using fuzzy logic, to send redundant information of only the most important data. This improves the human experience when watching live video flows while providing users and authorities (e.g., firefighters and paramedics) with a high perception of videos. This is important because it allows reducing the human reaction times in case of an emergency.

Figure 6.1 depicts the overall operation of the proposed mechanism. First of all, the video is captured, packetized, and delivered to uavFEC. After that, the mechanism will gather information about the video characteristics, such as the distance pointed by the motion vectors, frame type, GoP length, and relative position of P-Frames. This information is obtained through cross-layer techniques and loaded on the fuzzy interface engine to compute a suitable redundancy amount. Another important feature of uavFEC is the use of the network status to improve even further the amount of redundancy, which allows enhancing the video quality without adding unnecessary network overhead.

## 6.2.2 Towards the design of uavFEC

In order to conceive uavFEC, a knowledge database needs to be created. In order to do that, an exploratory analysis using hierarchical clustering was adop-
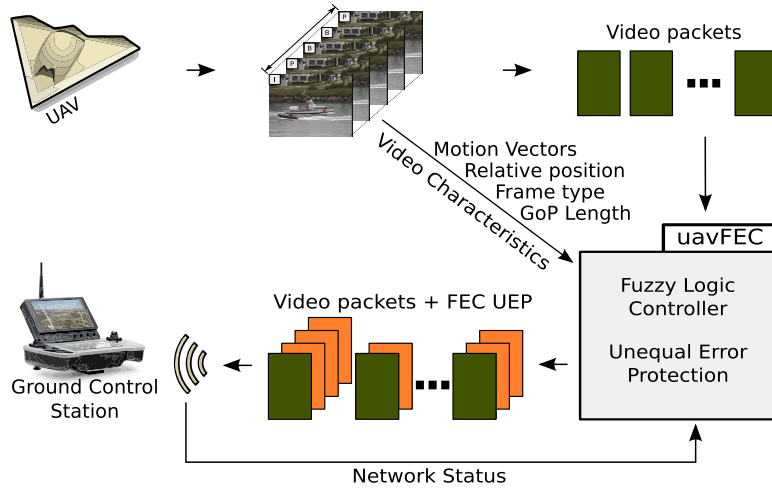
Figure 6.1: General view of the uavFEC mechanism

ted. This database stores information about the relation between several video characteristics and their impact on the quality of the videos. The combined use of this knowledge database and human expertise allows the definition of several fuzzy rules and sets. The offline process needs to be executed only once. Following this analysis, the information is loaded into the fuzzy interface engine and can be used in the real-time decision-making process. This is an important step since the real-time mechanism can be faster and more accurate, as fewer variables need to be handled.

uavFEC also uses the network state as one of the inputs to the adaptive mechanism. This information is jointly employed with the GoP length, motion vectors distance, frame type, and the relative position, to determine a suitable amount of redundancy. After that, using an improved UEP technique, a proper amount of redundancy will be added, sparing resources while increasing the video quality. A detailed description of the adaptive mechanism is presented below.

The use of fuzzy logic in the proposed mechanism allows it to be more comprehensive and dynamic because it can take into consideration a larger number of video and network details and still be fast enough to operate in real-time schemes as expected in a highly dynamic UAV network. Additionally, fuzzy logic can be considered a problem-solving methodology that aims to define what the system should do rather than attempting to fully understand its operation. It adopts a simple approach to provide definitive conclusions relying on imprecise, ambiguous, or vague information.

In order to use fuzzy logic, it is necessary to define several components, such as rules, sets, and membership functions. The rules define how the system behaves. The fuzzy sets, in contrast to classical sets that an element either belongs or does not belong to, are capable of having a degree of membership. Finally, the membership functions are designed to represent the significance of each element in the fuzzy set.

The process of designing the fuzzy logic components that will be used in the uav-

FEC mechanism enfolds a series of exploratory analysis to define the behaviour and value of each one of them. The first step is to quantify the motion intensity. In order to do that, an exploratory analysis using hierarchical clustering with Euclidean distance was conducted. This is a statistical method of partitioning data into groups that are as homogeneous as possible [Revelle, 1979]. Motion vectors data is used to create these clusters. As mentioned before, the idea of motion vectors was obtained from classical mechanics and their vector-oriented model of motion. This model describes the movement of objects as simple as the sequence of small translations on a plane [Le Gall, 1991].

To produce a comprehensive database, the motion vectors of several UAVs video sequences were extracted. Then, through Euclidean distance, it was computed how far each vector is pointing and summed together with all others in the same frame. This was used instead of just counting them because one frame can have several vectors pointing to a close distance where another frame can have fewer vectors, but pointing much farther away, thus presenting higher motion intensity.

Figures 6.2 and 6.3 depict an example of the aforementioned situation. At frame #21 (Figure 6.2) the UAV is turning right, thus, it is possible to see that the motion vectors are longer than in frame #34 (Figure 6.3). In the frame #34 the UAV finished the turn and starts hovering, therefore, the motion vectors are smaller. Incidentally, several of them are so small that they are represented by dots instead of arrows. The Euclidean distance sums of all motion vectors in these frames are 109300 and 14117, for frame #21 and #34 respectively. However, the total number of vectors in each frame is 4959 (frame #21) and 4963 (frame #34). This means that even tough frame #34 has more vectors, they are describing less motion than those stored at frame #21, more precisely, they are 7.74 times smaller.

Initially, the distance described by all the motion vectors in all frames is computed. After that, the frames are clustered together according to the motion intensity. Based on the linkage distance between the clusters, the motion intensity was divided into three groups, namely "small", "medium", and "high". Using this information, the motion intensity set can be defined as presented in Algorithm 6.1.

---

**Algorithm 6.1:** Motion Intensity input set

---

*InputLVar\* **Motion** = new InputLVar("**MotionIntensity**");*
  Motion → addTerm( ShoulderTerm("*LOW*", 10000, 30000, true));
  Motion → addTerm( TriangularTerm("*MEDIUM*", 21000, 80000));
  Motion → addTerm( ShoulderTerm("*HIGH*", 60000, 130000, true));
engine.addInputLVar(**Motion**);

---

After defining the sets, it is necessary to set up the membership functions. This definition is a complex and problem-dependent task. Taking this into account, it is preferable to use piecewise linear functions (formed by straight-line sections), because they are simple and more efficient with respect to computability and
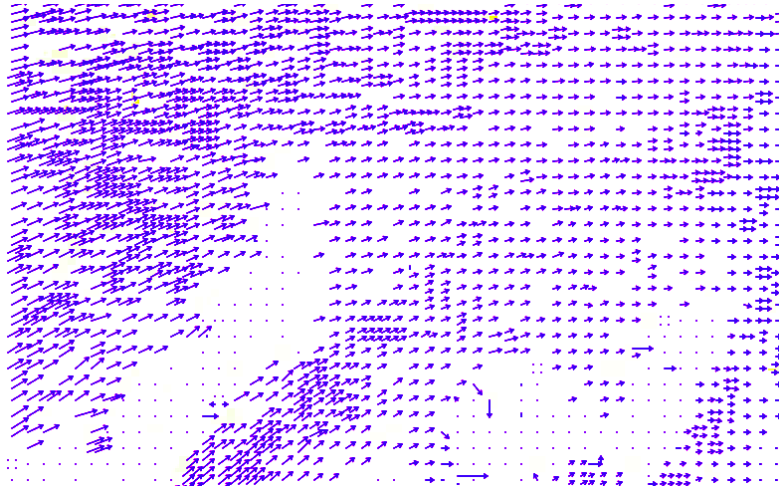
Figure 6.2: Motion vectors of frame #21
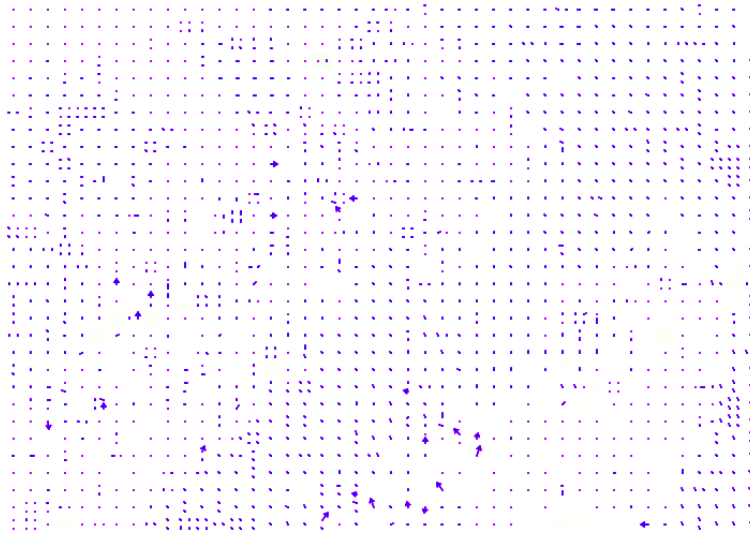


Figure 6.3: Motion vectors of frame #34

resource requirements. Figure 6.4 shows the graphical representation of the membership functions.

After delineating the motion intensity, the packet loss rate set must be defined. The aim of this activity is to quantify the packet loss rate against the video quality in terms of QoE. In other words, a loss rate of 10% can be considered low in the proposed approach however, it might be unacceptable in other applications, such as a voice over IP call. To define this set, a number of network simulations with several packet loss rates as well as a broad collection of UAV video sequences were carried out. On average, the video quality was considered good when the network losses were between 0% and 10%. Between 5% and 20%, a tolerable video quality was perceived, but over 15% the quality quickly decreased, soon becoming unacceptable. In view of this, three categories were defined, namely "low", "medium", and "high", as showed in Algorithm 6.2.

Another stage is to delineate the redundancy set. The main goal of this set is
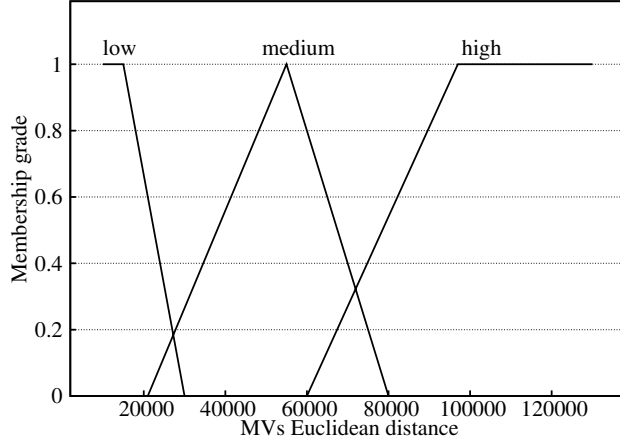
Figure 6.4: Motion intensity membership function

---

**Algorithm 6.2:** Packet loss rate input set for strict video resolutions

---

*InputLVar* ***PLR*** *= new InputLVar("**PacketLossRate**");*

> PLR → addTerm( TriangularTerm("*LOW*", 0, 15));
> PLR → addTerm( TriangularTerm("*MEDIUM*", 5, 30));
> PLR → addTerm( TriangularTerm("*HIGH*", 20, 100));

engine.addInputLVar(**PacketLossRate**);

---

to establish the output value which will be used to add the redundancy. Here again, a combination of experiments and human knowledge in the field was used to specify what could be considered a "small", "medium", and "large" amount of redundancy. The values obtained and the graphical representations of the membership functions are displayed in Algorithm 6.3 and Figure 6.5, respectively.

---

**Algorithm 6.3:** Motion activity output set

---

*OutputLVar* ***Redundancy*** *= new OutputLVar("**RedundancyAmount**");*

> Redundancy → addTerm( ShoulderTerm("*SMALL*", 0.55, 0.70, true));
> Redundancy → addTerm( TriangularTerm("*MEDIUM*", 0.60, 0.80));
> Redundancy → addTerm( TriangularTerm("*LARGE*", 0.75, 1));

engine.addOutputLVar(**RedundancyAmount**);

---

After defining all the fuzzy sets, the IF-THEN structure must be created. This is a straightforward procedure, because if the transmitted video has low levels of motion activity (according to the motion vectors) and the packet loss rate is low as well, then the uavFEC will attribute also a low redundancy. The same procedure is valid for "medium" and "high" motion activities and packet loss rate as partially depicted in Algorithm 6.4.

After defining the rules and sets, they need to be loaded in the fuzzy logic controller. This activity has to be performed just once, during the system setup period (bootstrap). After the definition, the controller will calculate the de-
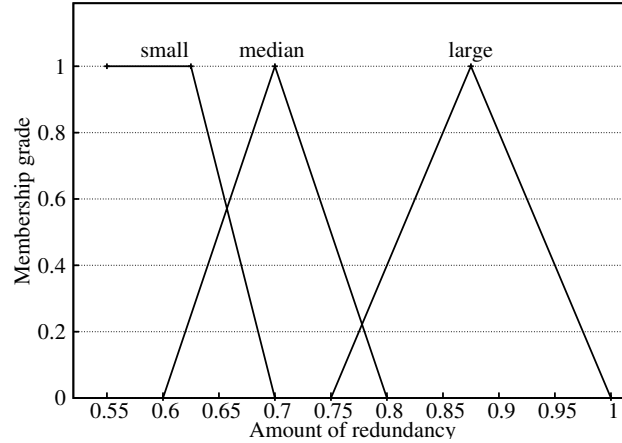
Figure 6.5: Redundancy amount membership function

---

**Algorithm 6.4:** Packet loss x redundancy amount rules

*RuleBlock\* **block** = new RuleBlock();*

*block → addRule( new MamdaniRule("*
    ***if** (Motion is LOW and PacketLossRate is LOW)* **then**
        RedundancyAmount is SMALL;
  *", engine));*

*block → addRule( new MamdaniRule("*
    ***if** (Motion is MEDIUM and PacketLossRate is MEDIUM)* **then**
        RedundancyAmount is MEDIUM;
  *", engine));*

*block → addRule( new MamdaniRule("*
    ***if** (Motion is HIGH and PacketLossRate is HIGH)* **then**
        RedundancyAmount is LARGE;
  *", engine));*

---

gree of membership of each input information, resulting in a precise amount of redundancy on-the-fly.

This is important because video transmission is delay-sensitive, meaning that if a frame is received after its decode deadline it cannot be displayed. Moreover, unlike neural networks or genetic algorithms, the fuzzy logic controller does not need a period of online training or convergence, making it a proper tool for real-time control. Additionally, the calculations can be very simple, especially when triangular or trapezoidal membership functions are adopted [Pedrycz, 1994], and even further reduce to a simple operation through fuzzy control surface.

## 6.2.3 uavFEC Performance Evaluation and Results

The main goal of the uavFEC mechanism is to improve the perceived video quality without adding unnecessary network overhead, thus saving wireless network resources.

### 6.2.3.1 Experiment settings

The evaluation scenario is composed of up to four UAVs, equipped with a 4G LTE radio at 800MHz. These UAVs can be operated in autonomous or nonautonomous mode. In a surveillance scenario, for example, it is possible to have a human operating the UAV. This allows having an instant change of direction and speed during the pursuit of a suspect. Therefore, the mobility model was defined as random waypoint [Bouachir et al., 2013].

All UAVs are in line-of-sight and communicate directly with an ad-hoc connection to the ground control station which was equipped with a portable base station and antenna. To simulate the video transmission, a set of twenty real UAV video sequences in high definition (720p), GoP length of 19:2, and H.264 codec was used. Due to the ad-hoc communication and the high definition videos, the flying range is limited to a radius of 900 meters from the base station.

A Frame-Copy error concealment method is active, this means that lost frames are replaced by the last good one received. The PLR varies according to the movement of the UAVs, namely distance from the portable base station and velocity, and also due to concurrent transmissions of others UAVs. Owing to the aforementioned details, the PLR can range from 0% to 45%. Table 6.1 shows the simulation parameters.

Table 6.1: uavFEC Simulation parameters

| Parameters | Value |
|---|---|
| Display size | 1280 x 720 |
| Display aspect ratio | 16:9 |
| Frame rate mode | Constant |
| Frame rate | 29.970 fps |
| GoP | 19:2 |
| Video format | H.264 |
| Codec | x264 |
| Container | MP4 |
| Propagation model | FriisPropagationLossModel |
| UAV velocity | 45-65 km/h (28-40 mph) |
| LTE Frequency band | 800MHz |
| LTE Mode | FDD |
| LTE Bandwidth | 5 MHz |
| eNodeB Operating Power | 22 dBm |
| Antenna Gain | 16 dBi |

In order to compare the results, five different cases were simulated. The first is without FEC, serving as a baseline to compare with the others. The second case is a non-adaptive video-aware FEC-based approach. In this case, only I- and P-Frames are protected with an equal amount of redundancy, which was set at 65%. This amount was chosen because it provides a good tradeoff between video quality and network overhead under several PLRs. The next case is the previously proposed mechanism in Section 5.2, which uses a simple adaptive unequal error protection (ViewFEC). Another case is an implementation of the Cross-Layer Mapping Unequal Error Protection (CLM-UEP) [Lin et al., 2012a]. The last case is the proposed uavFEC mechanism. The simulation setup is composed of 20 real UAVs video sequences and 5 cases, each one was simulated 30 times with each video, resulting in 3.000 simulations in total.

### 6.2.3.2 QoE assessments

Figure 6.6 depicts the average SSIM for all video sequences when only one UAV is transmitting. The measurement of this metric is fairly simple however, it is consistent with the human visual system, given good results [Wang et al., 2004]. In SSIM, values closer to one indicate a better video quality. As expected, when the UAV is far away from the ground control station there is a decline in the video quality. In the baseline case, without FEC, a sharp decline in the video quality after 400m is perceived. Conversely, the UAVs using a FEC-based mechanism are able to sustain a better video quality longer, and it is only noticeable after 500m for case 2, and after 700m for cases 3 to 5.
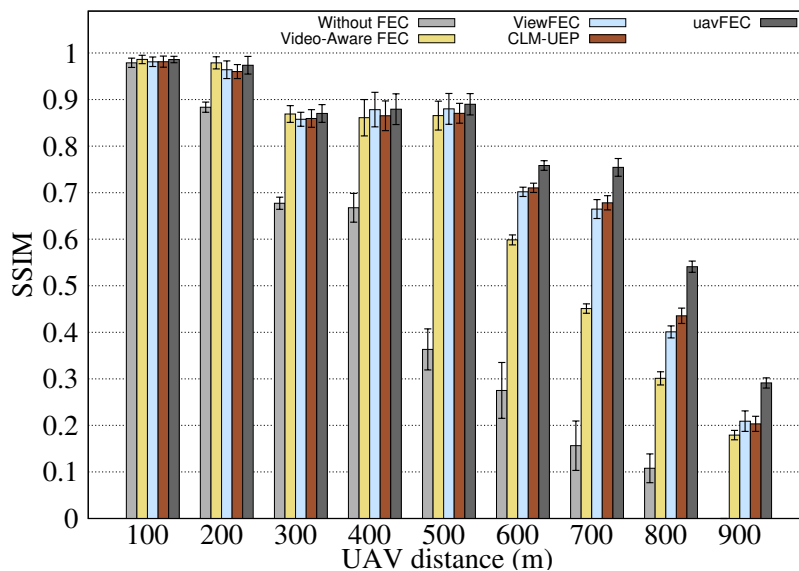


Figure 6.6: SSIM for all scenarios with one UAV

Almost the same behaviour is shown in Figure 6.7 which demonstrates the results for 2 UAVs transmitting simultaneously. One clear difference between these two scenarios is the increase in the standard deviation of the baseline case. This can be explained by the natural resiliency of some videos to packet loss due

to different video characteristics. Video sequences with low motion intensity are more resilient to losses, and generally, have better results in the QoE-aware assessment. On the other hand, videos with high motion intensity tend to have poor results.
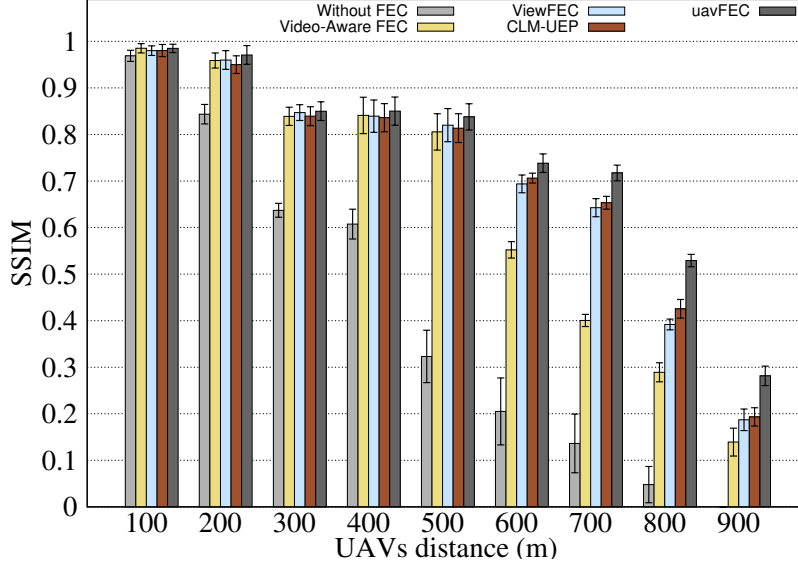


Figure 6.7: SSIM for all scenarios with two UAVs

As the number of video sequence flows begins to increase, the quality of the transmitted video starts to decrease sooner than before. Figures 6.8 and 6.9 depict this tendency. In the first two scenarios (with one and two UAVs), the uavFEC managed to keep the SSIM above 0.7 up to 700m (other approaches only up to 600m). However, with three and four UAVs, the uavFEC was able to maintain the SSIM over 0.7 only up to 600m, after that, there is a sharp decline in the video quality in all of the assessed mechanisms. This can be attributed to a more congested network due to several transmissions together with the distance from the ground control station.

### 6.2.3.3 Network footprint analysis

Throughout the QoE assessment was demonstrated that the uavFEC mechanism enhances the video quality over several scenarios, having particularly good results over higher distances and with increased network traffic. Besides the video quality, the uavFEC was also designed to add as little as possible redundancy, to maintain a low overhead and thus saving resources.

The network overhead was computed by summing the size of all video frames transmitted by each mechanism. This means that, if the original frame size is subtracted, it is possible to find the specific amount of redundancy added only by the approaches. Two mechanisms assessed are non-adaptive, video-aware FEC and ViewFEC, and because of that, they have the same network overhead in all distances, which was 65.10% and 38.90%, respectively, as showed in Figure 6.10.
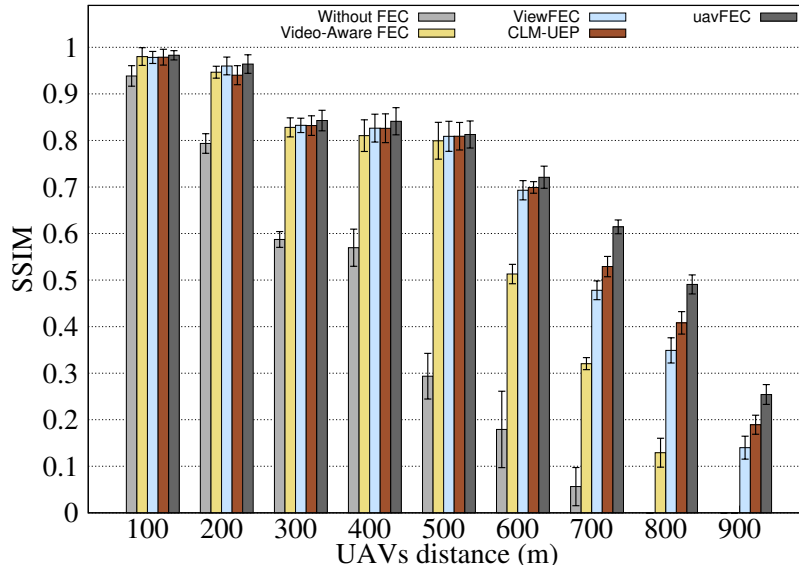
Figure 6.8: SSIM for all scenarios with three UAVs
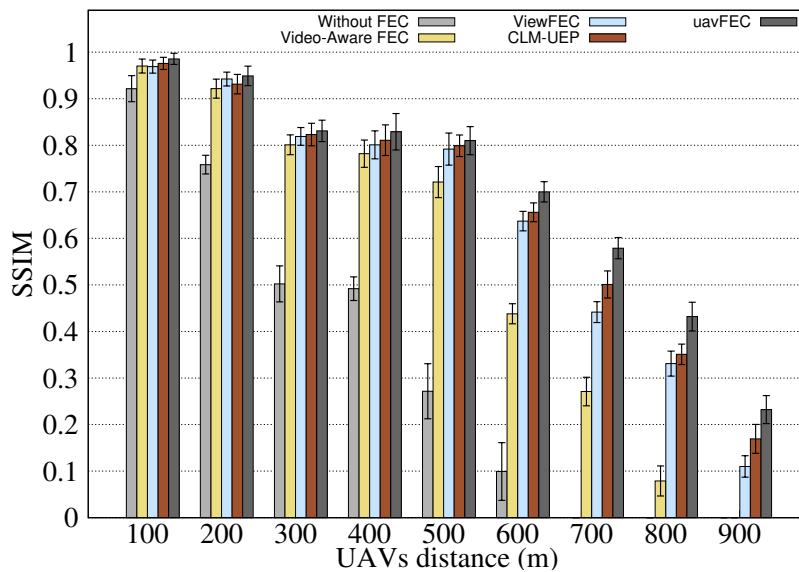


Figure 6.9: SSIM for all scenarios with four UAVs

These mechanisms are not appropriate because even when the UAVs are close to the ground control station they add a considerable amount of redundancy, wasting resources. The same figure depicts the results for uavFEC and CLM-UEP. Both mechanisms perform close to each other up to 600m, but in average the uavFEC has lower network overhead. Over 600m, the uavFEC starts to add more redundancy, increasing the network overhead, however, providing better video quality.
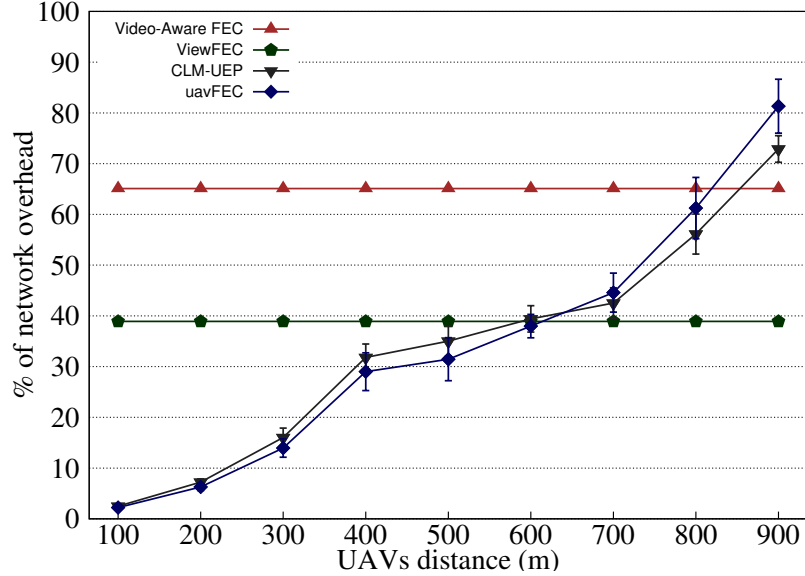
Figure 6.10: Network overhead for all scenarios

### 6.2.3.4 Overall results

Figure 6.11 depicts the comparison of uavFEC and the related work (CLM-UEP) [Lin et al., 2012a]. The graph shows the average percentage of QoE improvement against the amount of redundancy added by the mechanisms in all scenarios (from 1 to 4 UAVs). A positive percentage means that uavFEC had better QoE results than CLM-UEP. In all four scenarios, uavFEC presented a slightly better video quality until 600m, on average between 0.59% and 5.00% better. The real advantage of uavFEC is noticeable after the 700m when it enhances even further the video quality. The uavFEC was able to achieve improvements, on average, between 11.59% and 28.52%, better than CLM-UEP. Taking this into consideration, it is clear that the proposed mechanism performs better in a higher distance, where the PLR is also higher. This gives uavFEC the capability to operate in wide coverage areas.

It is also shown in Figure 6.11 the comparison of the amount of redundancy added by both, CLM-UEP and uavFEC. A negative percentage means that the proposed mechanism adds less redundancy than CLM-UEP. In all four scenarios, uavFEC added less redundancy until 600m, which was around 3.75% and 15.12% less on average and still managed to transmit the videos with higher QoE. This means that uavFEC was able to improve the video quality and at the same time save resources. After 700m, the uavFEC mechanism begins to increase the redundancy. This happens because the uavFEC was developed to enhance the video quality over higher distances, which make the networks more susceptible to errors. Considering this, the mechanism will have to increase the protection of the most important video data, adding more overhead. For example, at 700m the uavFEC mechanism added on average 4.49% more redundancy, and at 900m added 10.19%. Increasing the redundancy is an expected response of the proposed mechanism to further improve the video quality, which can be
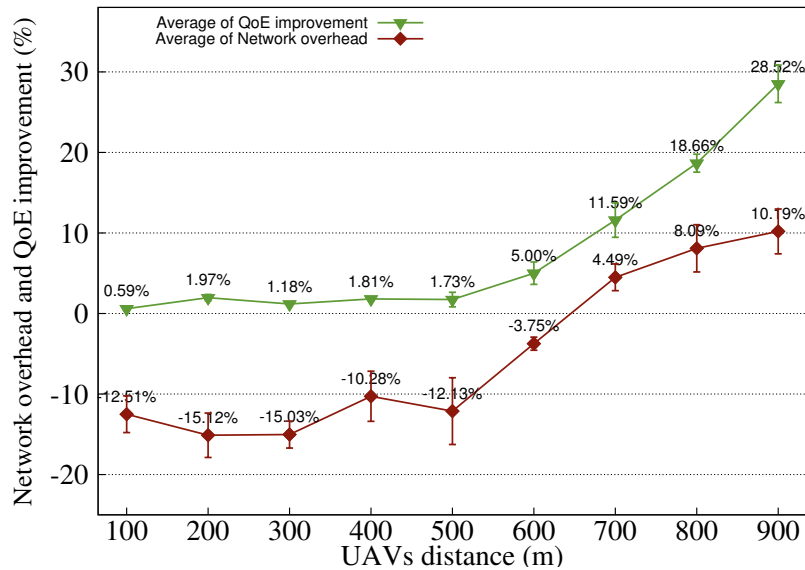
Figure 6.11: QoE and Redundancy against UAV distance

confirmed through the QoE assessment in the same figure. In summary, uavFEC provides a good tradeoff between video quality and network overhead.

A further analysis of Figure 6.11 shows that up to 500m both mechanisms had similar QoE results, with uavFEC having a modest higher video quality. The major difference was the considerably smaller network overhead, this means that uavFEC, through its QoE- and Video-aware techniques, was able to add redundancy to the most important video data only. At 600m, the uavFEC mechanism still adds less redundancy than the related work, but it is already showing better results, with an improvement of 5.00% on QoE. After this threshold, considering the increasing distance and in order to improve the video quality, the proposed mechanism starts to add a larger amount of redundancy. The result of this approach are videos transmitted on average with more than 28% of better quality than CLM-UEP while adding less than 11% of redundancy.

The uavFEC mechanism achieved good results making the video transmission more resilient to packet loss and thus, enabling a longer video transmission range for the UAVs. The results are particularly beneficial in a higher distance with several UAVs, providing a better video quality of live video flows, allowing end-users, such as civilians and/or authorities, to have a high-quality perception of videos and thus reducing reaction times.

## 6.3 Adaptive Motion Intensity Awareness mechanism (MINT-FEC)

The mechanism presented in Section 6.2 obtained favourable results however, it presented some drawbacks in regard to the motion intensity classification and the strict video resolution dependence. To improve on these issues, the

adaptive FEC-based mechanism with motion intensity awareness (MINT-FEC) was proposed.

## 6.3.1 MINT-FEC Overview

The MINT-FEC mechanism aims to enhance the resilience of UAV real-time video transmissions. One of the major weaknesses in the mechanisms found in the literature, including the uavFEC, is the use of unnecessary redundancy. To tackle this issue, the MINT-FEC dynamically adapts itself by using fuzzy logic, to add a precise amount of redundancy to only the most QoE-sensitive data, while ensuring high-quality video and downsizing the usage of scarce wireless resources.

In the MINT-FEC mechanism, the motion intensity is now given by combining the spatial complexity and temporal intensity. Spatial complexity is how distinct one frame is from another, as well as the colour and luminance saturation. The temporal intensity can be defined as how fast and how much the image is changing frame-by-frame. Their joint use provides a more accurate motion classification to be used for adding a adequate amount of redundancy. Another improvement from the work described in Section 6.2 is the video resolution independence. By normalising the values of all the video characteristics, as well as using the motion vector distance and macroblock size, it is possible to add, on-the-fly, an adaptive amount of redundancy to videos with arbitrary resolution. A detailed description of all the novel components is given next.

Following the same core structure as uavFEC, the MINT-FEC mechanism also depends on Fuzzy logic and thus several fuzzy components need to be defined, such as the sets, membership functions, and rules. The offline process needs to be executed only once. After that, all the generated information can be loaded into the fuzzy interface engine to be used in real-time.

## 6.3.2 Towards the design of MINT-FEC

The design process of MINT-FEC starts with the definition of the fuzzy components. In a similar fashion to the uavFEC mechanism, the first step is to quantify the spatial complexity. This component represents how much static information a frame is carrying compared to the previous one. The most common way to compute this difference is using the Sum of Absolute Differences (SAD) [Vanne et al., 2006]. It is not a complex operation however, it is very time-consuming because it compares each pixel from both frames, making this impractical in real-time.

Another way to locate this information is through the frame sizes. The problem of using the frame size is that several video characteristics can impact on it, such as different resolutions (picture size), content, as well as temporal intensity. To be able to compare the frame sizes among different videos, it is necessary to normalise all the information. Using Equation (6.1) the average frame size is

calculated, and the same operation is also executed for P- and B-frames. After that, through Equation (6.2), all frame sizes are normalised, as before, this is also done for P- and B-frames. This process is performed for each video sequence separately. Table 6.2 shows the adopted notation.

$$\mu I_s = \frac{1}{nF} \sum_{i=0}^{nF-1} I_{s(i)} \tag{6.1}$$

$$\hat{\mu} I_s = \frac{\mu I_s}{\mu I_s + \mu P_s + \mu B_s} \tag{6.2}$$

Table 6.2: MINT-FEC Adopted Notation

| Notation | Meaning |
|---|---|
| $\mu I_s, \mu P_s, \mu B_s$ | Frame size average |
| $\hat{\mu} I_s, \hat{\mu} P_s, \hat{\mu} B_s$ | Normalised frame size average |
| $I_{s(i)}, P_{s(i)}, B_{s(i)}$ | Frame size of the $i_{th}$ frame |
| $nF$ | Number of frames in the video sequence |
| $|MV|$ | Euclidean distance of a motion vector |
| $\left|MV_{(i)}\right|$ | Euclidean distance of the $i_{th}$ motion vector |
| $MB_h$ | Macroblock height |
| $MB_w$ | Macroblock width |
| $aMB$ | Macroblock area |
| $aMB_{(i)}$ | Area of the $i_{th}$ macroblock |
| $nMB$ | Number of macroblock in the frame |
| $TI_{\Delta t}$ | Temporal intensity |

Once all the frame sizes are normalised, it is possible to perform an exploratory analysis to cluster all frames of all video sequences together according to their sizes. Based on the linkage distance of the clusters it was possible to divide them into three distinct groups, namely "small", "medium", and "large". After defining the clusters, a boxplot was used to summarise and display the distribution of the data. This is an important tool in the exploratory analysis because it displays the shape of the distribution of each cluster along with the central value and the variability. Figure 6.12 shows the boxplot for the spatial complexity. The fuzzy sets for spatial complexity were defined using the information displayed by the boxplot, as showed by Algorithm 6.5.

After defining the set, the membership functions need to be outlined. This is problem-dependent, as well as a complex task being difficult to find an optimal solution [Wong et al., 2005]. Considering that, it is better to use piecewise linear functions (formed of straight-line sections). These functions are both simpler and more efficient regarding computability, leading to lesser resource requirements. Figure 6.13 shows the graphical representation of the chosen membership functions for the frame sizes.

Apart from the spatial complexity, the fuzzy components for the motion intensity also need to be created. The analysis of this criterion is performed through
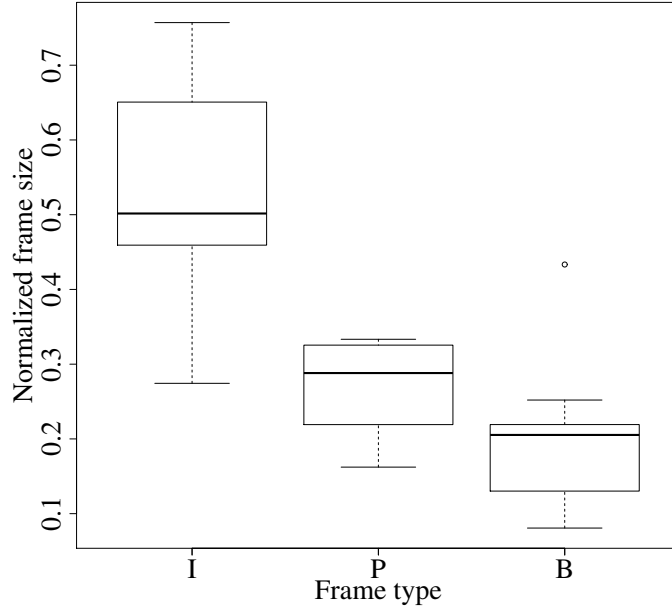
Figure 6.12: Spatial Complexity

---

**Algorithm 6.5:** Spatial complexity (Frame size sets)

---

FuzzyOperator& **op** = FuzzyOperator::DefaultFuzzyOperator();
FuzzyEngine engine("complex-mamdani", op);

*InputLVar\* **Isz** = new InputLVar("**I-size**");*
　Isz → addTerm( ShoulderTerm("*SMALL*", 0.274, 0.459, true));
　Isz → addTerm( TriangularTerm("*MEDIUM*", 0.274, 0.651));
　Isz → addTerm( ShoulderTerm("*LARGE*", 0.502, 0.757, true));
engine.addInputLVar(**Isz**);

*InputLVar\* **Psz** = new InputLVar("**P-size**");*
　Psz → addTerm( ShoulderTerm("*SMALL*", 0.162, 0.219, true));
　Psz → addTerm( TriangularTerm("*MEDIUM*", 0.162, 0.325));
　Psz → addTerm( ShoulderTerm("*LARGE*", 0.288, 0.333, true));
engine.addInputLVar(**Psz**);

*InputLVar\* **Bsz** = new InputLVar("**B-size**");*
　Bsz → addTerm( ShoulderTerm("*SMALL*", 0.081, 0.13, true));
　Bsz → addTerm( TriangularTerm("*MEDIUM*", 0.081, 0.219));
　Bsz → addTerm( ShoulderTerm("*LARGE*", 0.205, 0.252, true));
engine.addInputLVar(**Bsz**);

---

motion vectors details. In order to better infer the amount of movement described by the vectors, instead of counting them, it is computed how far each one is pointing using the Euclidean distance. This represents the motion intensity better because it is possible to have one frame with several vectors pointing to a close distance meanwhile, another frame with fewer vectors, pointing farther away although, and thereby having higher motion intensity.

As defined in the MPEG standard, the motion vectors describe the movement
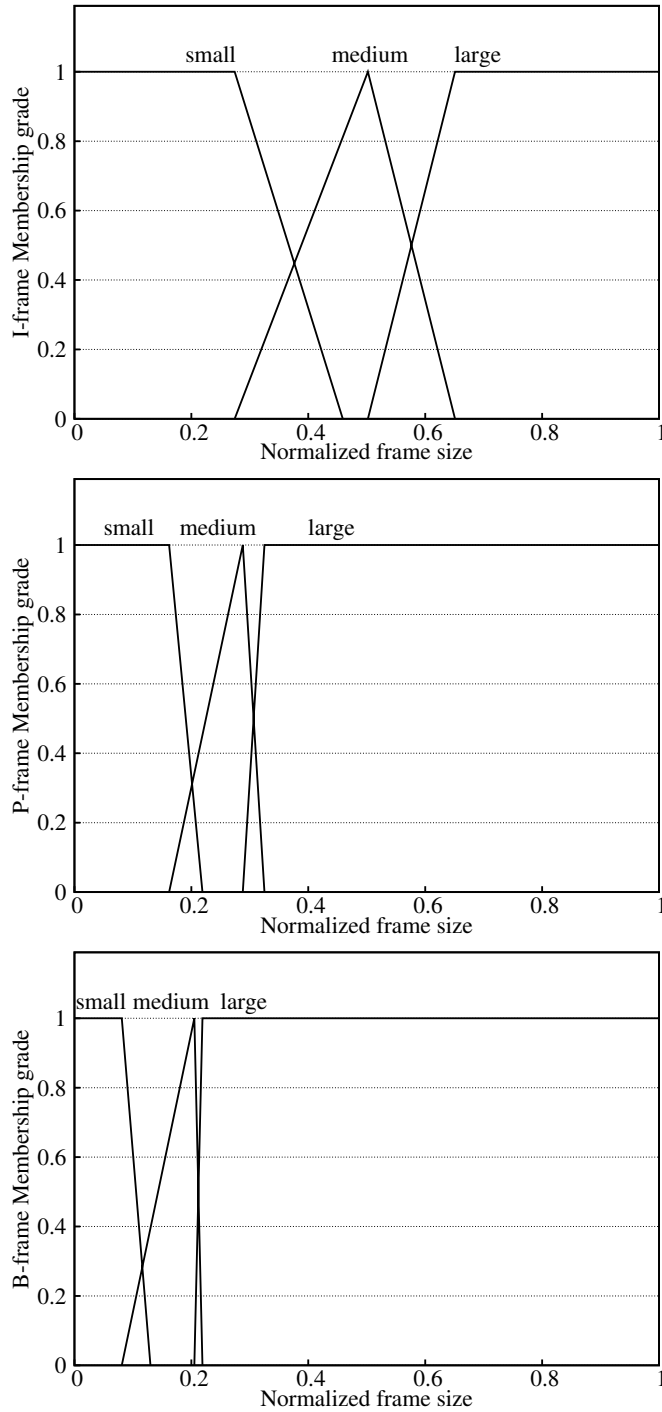
Figure 6.13: Frame size membership function

of macroblocks from some position in one frame to another position in another frame. It is important to note that not all MB have the same size, as well as videos with higher resolution will have more macroblocks than videos with lower resolution. To be able to compare video sequences with different macroblocks sizes and resolutions, the MB area is used, given by Equation (6.3), together with the motion vectors. Additionally, using Equation (6.4) it is possible to calculate for each macroblock how many pixels have been moved and how far away, which can be translated as temporal intensity.

$$aMB = MB_h \times MB_w \tag{6.3}$$

$$TI_{\Delta t} = \frac{1}{nMB} \sum_{i=0}^{nMB-1} aMB_{(i)} \times \left| MV_{(i)} \right| \tag{6.4}$$

Using the aforementioned details, another exploratory analysis was performed to classify the video sequences in terms of temporal intensity. This time, instead of breaking the video sequences in frames, the whole video was analysed. The values found through Equation (6.4) were used to cluster the videos into three distinct groups, namely "low", "medium", and "high" temporal intensity. Additionally, in the same way as before, a boxplot was used to summarise and display the data distribution, as well as to create the sets, as presented in Figure 6.14.
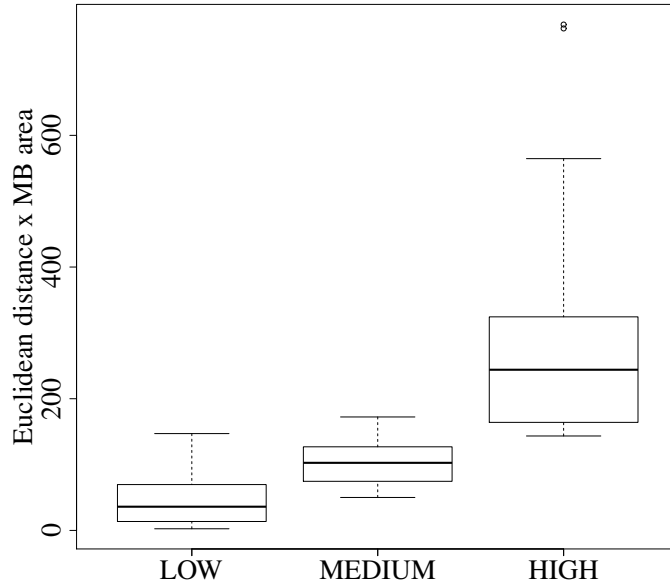


Figure 6.14: Temporal Intensity

Another important step is to define the PLR. The primary objective is to find out the influence of different PLR in the QoE for a set of videos. The same procedure adopted by uavFEC was applied to MINT-FEC however, different values were found. To find the PLR that best represents the video quality, a number of network simulations using a broad collection of UAV video sequences were carried out. As the MINT-FEC can handle arbitrary video resolution, several of them were used in the experiments. For PLR between 0% and 10%, the video quality was good, in comparison to the uavFEC, where the good range was between 0% and 15%. A tolerable video quality was observed for most of the videos between 5% and 20% of PLR (uavFEC had values between 5% and 30%). Over 15% of PLR the quality quickly decreased in videos with higher motion intensity, and over 25% it became unacceptable. In general, the PLR input set is lower in MINF-FEC than in uavFEC because videos with higher resolution tend to consume more network resources. Based on the results, the PLR set was defined, as shown in Algorithm 6.6.

---

**Algorithm 6.6:** Packet loss rate input set for arbitrary video resolutions

---

*OutputLVar\* **PLR** = new OutputLVar("**PacketLossRate**");*
  PLR → addTerm( TriangularTerm("*LOW*", 0, 10));
  PLR → addTerm( TriangularTerm("*MEDIUM*", 5, 20));
  PLR → addTerm( TriangularTerm("*HIGH*", 15, 100));
engine.addOutputLVar(**PLR**);

---

The last fuzzy set is the redundancy set. The MINT-FEC mechanism uses the same set as uavFEC, which was explained in Section 6.2.2 and Algorithm 6.3 defines it. With all sets delineated, it is necessary to create the rules. This activity also involves human knowledge about the video characteristics, namely spatial complexity and temporal intensity, as well as the frame type, and the PLR. As mentioned before, in videos with high spatial complexity the I-Frame needs a greater amount of protection, because it holds a large amount of information. On the other hand, in videos with high temporal intensity, the I-Frame also needs to be protected. The P-Frame also plays an important role because it holds the temporal information about that sequence, and needs to have almost the same protection as the I-Frame. Algorithm 6.7 shows two rules that represent this case.

---

**Algorithm 6.7:** Packet loss x video characteristics rules

---

*RuleBlock\* **block** = new RuleBlock();*

*block → addRule( new MamdaniRule("*
  ***if** (SpatialComplexity is HIGH and*
      *PacketLossRate is HIGH and*
      *FrameType is I)*
  ***then***
      RedundancyAmount is HIGH;
  *", engine));*

*block → addRule( new MamdaniRule("*
  ***if** (TemporalIntensity is HIGH and*
      *PacketLossRate is HIGH and*
      *FrameType is I or P)*
  ***then***
      RedundancyAmount is HIGH;
  *", engine));*

---

MINT-FEC utilises the same core structure of uavFEC, so once all the fuzzy rules and sets are defined, they are employed in real-time in the fuzzy logic controller. In the same way as before, the offline process needs to be performed just once, after that the controller will be able to compute a suitable amount of QoE-aware redundancy on-the-fly.

## 6.3.3 MINT-FEC Performance Evaluation and Results

The MINT-FEC goal is to improve on uavFEC (Section 6.2) to ensure an even higher perceived QoE for end-users, while avoiding unnecessary network overhead.

### 6.3.3.1 Experiment settings

The assessment scenario consists of up to four UAVs operating in autonomous mode, with 4G LTE radio at 800MHz. To better reflect a UAV scenario, the Gauss-Markov distribution mobility model was used. This model provides a uniform spatial distribution of the nodes and also simulates inertia in the movements, which are a characteristic of UAVs in autonomous mode. The ground control station is equipped with a portable base station and antenna. All UAVs are in line-of-sight and communicating in ad-hoc mode. Only real UAV video sequences were used in the experiments. To be more precise, twenty of each video resolution (1080p, 720p, and SVGA), giving a total of sixty video sequences. All of them were encoded with both same GoP length of 19:2 and same H.264 codec. Considering the portable base station power, the ad-hoc communication, and the very demanding high definition videos, the flying range was limited to a radius of 2000 meters from the base station. Due to the harsh environment and the low-gain antenna, the PLR can range from 0% to 35%. Figure 6.15 shows the packet loss distribution in the experiments. At the receiver side, a Frame-Copy error concealment was used. Table 6.3 shows the simulation parameters.
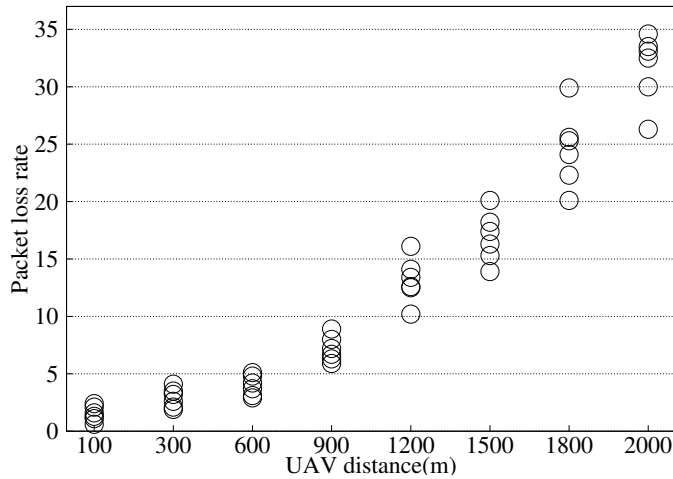


Figure 6.15: MINT-FEC's experiment PLR distribution

Five different schemes were simulated as follows: (1) without any FEC mechanism. This is only to serve as a baseline for comparison with the others; (2) a non-adaptive video-aware FEC (I- and P-Frames are equally protected) using a pre-set value of 75% of redundancy (Video-aware FEC). This value was chosen because it showed a good tradeoff between QoE and network overhead in several PLR; (3) the adaptive FEC-based mechanism (uavFEC), presented

Table 6.3: MINT-FEC Simulation parameters

| Parameters | Value |
|---|---|
| Display sizes | 1920x1080, 1280x720, and 800x600 |
| Frame rate mode | Constant |
| Frame rate | 29.970 fps |
| GoP | 19:2 |
| Video format | H.264 |
| Codec | x264 |
| Container | MP4 |
| Propagation model | FriisPropagationLossModel |
| Mobility model | Gauss-Markov |
| UAV velocity | 45-65 km/h (28-40 mph) |
| LTE Frequency band | 800MHz |
| LTE Mode | FDD |
| LTE Bandwidth | 5 MHz |
| eNodeB Operating Power | 22 dBm |
| Antenna Gain | 16 dBi |

in Section 6.2; (4) a related work implementation of the Cross-Layer Mapping Unequal Error Protection (CLM-UEP) [Lin et al., 2012a]. At last, (5) adopts the proposed MINT-FEC mechanism.

### 6.3.3.2 QoE assessments

Figure 6.16 shows the average of the SSIM results. A foreseen situation can be clearly noticed, the farther away the UAVs are from the ground control station the worst is the video quality. In the first case (without FEC), a good video quality is noticed up to 600m. This is expected on the grounds that some video sequences tend to have a natural resiliency to packet loss. In particular, this situation is true in videos with low motion intensity, which usually scores higher results in QoE assessment. Between 600m and 900m the video quality is already affected and a sharp decline is perceived after that. At the same time, in the FEC-based schemes, such as Video-aware FEC, uavFEC, CLM-UEP and, MINT-FEC, the video quality was good for a long distance, until 1200m. Notwithstanding, the proposed mechanism outperforms all its competitors in terms of video quality, providing even better results over higher distances. A comprehensive comparison analysis is given further.

### 6.3.3.3 Network footprint analysis

MINT-FEC can provide enhanced video quality, especially over higher distances, however, it is equally important to do so with lower network overhead. The Video-aware FEC scheme is non-adaptive and due to that, it has a constant network overhead, as showed in Figure 6.17. This is not suitable for UAVs because even when they are close to the base station, with a low PLR, a large
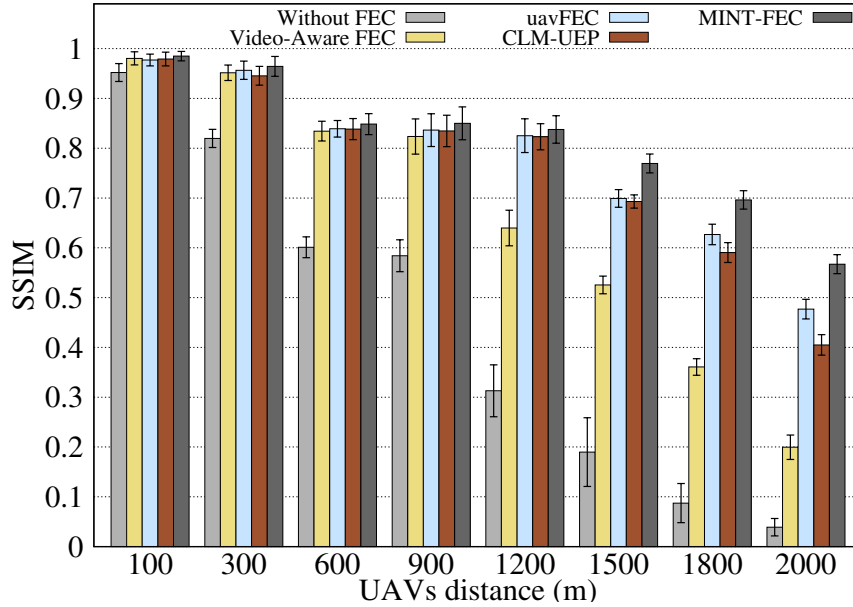
Figure 6.16: Average SSIM QoE for all scenarios

amount of redundancy is added, wasting resources.  On the other hand, the
adaptive mechanisms (uavFEC, CLM-UEP, and MINT-FEC) allow a better use
of the network resources, as also shown in Figure 6.17.
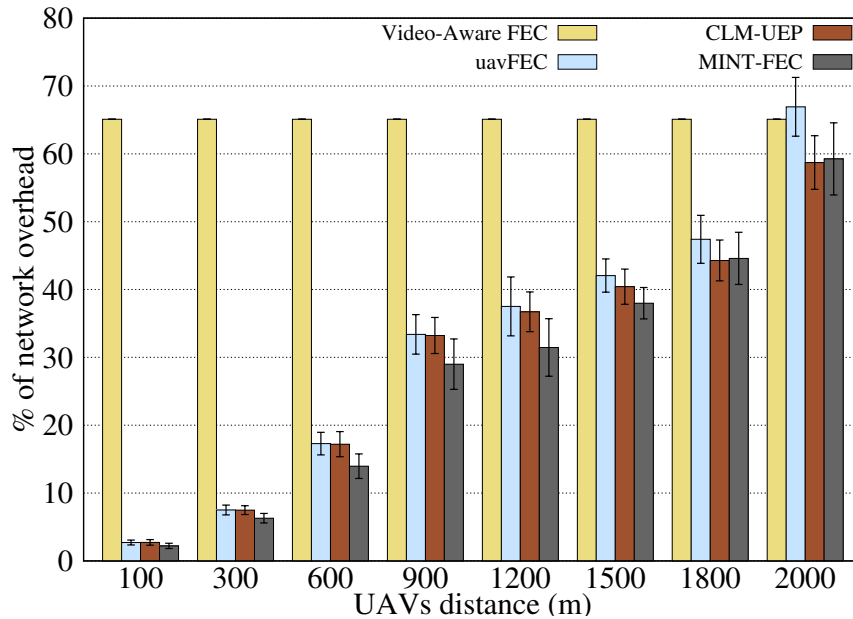


Figure 6.17: Average Network overhead

In all three mechanisms, the initial amount of redundancy is small and starts
to become larger as the UAVs move away from the ground station.  Both the
uavFEC mechanism and CLM-UEP perform close to each other up to 1200m.
After that, uavFEC starts to add more redundancy to provide better video
quality.  Here again, the MINT-FEC performs better than the others.  Up to

1500m, it induces less network overhead, while providing higher video quality. Subsequently, it adds a slightly higher redundancy and still smaller than the other schemes, in favour of a considerably better video quality. This proves that it was possible to identify the most important video portions and protect them accordingly.

### 6.3.3.4 Overall results

To further understand the MINT-FEC achievements, a comparison against CLM-UEP and uavFEC is given in Figure 6.18. The first case is the comparison between CLM-UEP and MINT-FEC, and the second one, uavFEC against MINT-FEC. The graph shows the average percentage of QoE and network overhead improvement. In the QoE assessment, a positive percentage means that the proposed mechanism achieved higher video quality, which is desirable. On the other hand, in the network evaluation, a negative percentage means that the MINT-FEC generated less overhead, which is also advantageous.
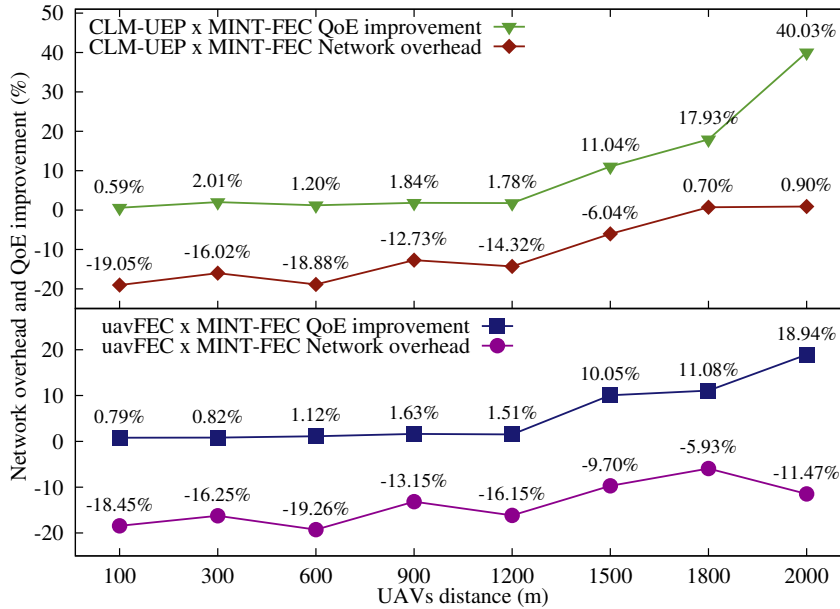


Figure 6.18: QoE and Redundancy comparison

In both cases, MINT-FEC presented a slightly better video quality until 1200m, which was between 0.59% and 2.01%, and between 0.69% and 1.63%, respectively. While outperforming the other mechanisms in terms of quality, the proposed mechanism was also able to considerably reduce the network overhead. It added on average 16.20% less redundancy than CLM-UEP and 16.65% less redundancy than uavFEC, up to 1200m. This proves that MINT-FEC is capable of better identifying the most QoE-sensitive data and adds a precise amount of QoE-aware redundancy to it, resulting in higher video quality and less network overhead. After this threshold, the MINT-FEC starts to increase the amount of redundancy to improve the video quality. This happens for the reason that the proposed mechanism was designed to sustain a higher video quality over

long distances when the connection is more susceptible to errors. In doing that, the videos are received with up to 40% better quality in comparison to the CLM-UEP mechanism and up to 18% higher quality than the uavFEC mechanism. Another important advance of the MINT-FEC over the uavFEC was the network overhead reduction beyond 1200m. In this case, the MINT-FEC mechanism managed to reduce the overhead by up to 11.74%. This is an additional proof that the proposed mechanism is doing a better work to infer the motion intensity using spatial complexity and temporal intensity, allowing the definition of a precise amount of redundant information to the most sensitive data. Thanks to that, it was possible to deliver higher video quality leading to a better user perception.

## 6.4 Summary

This chapter investigated the performance of the two proposed mechanisms to improve the resiliency to packet loss over FANETs using UAV-to-Ground model, namely uavFEC and MINT-FEC. The results shown in Section 6.2.3 demonstrated that the uavFEC mechanism improves video transmissions over highly dynamic networks, maximising the QoE without adding unnecessary network overhead. The uavFEC mechanisms was able to provide videos in the best-case scenario with, on average, between 11.59% and 28.52% better QoE than the main competitor (CLM-UEP). Overall, the uavFEC mechanism attained good results. It had, nevertheless, some drawbacks especially due to its strict video resolution dependence and basic motion intensity classification.

To improve on these issues, the MINT-FEC mechanism was proposed. The achieved results, shown in Section 6.3.3, attested that the proposed mechanism outperforms the other mechanisms in the experiments. In scenarios with up to 1200m, it was able to deliver videos with slightly higher quality and at the same time generating substantially less network overhead. Therefore, an enhanced video quality was perceived without wasting wireless resources. On the other hand, over 1200m, due to harsher conditions, the MINT-FEC mechanism starts to increase the redundancy providing a considerably higher QoE than the other mechanisms. This is a desired tradeoff between network overhead and video quality. The results support the claim that MINT-FEC was able to identify the motion intensity video sequences as well as to handle arbitrary video resolutions. At the end, MINT-FEC was able to better protect the most QoE-sensitive data, which is translated in higher video quality.

The studies and proposed mechanisms in this chapter resulted in the following publications:

**Book chapter:**

- Immich, R. and Cerqueira, E. and Curado, M., "**Improving video QoE in Unmanned Aerial Vehicles using an adaptive FEC mechanism**", in Wireless Networking for Moving Objects: Models, Approaches,

Techniques, Protocols, Architectures, Tools, Applications and Services, Volume 8611, pp 198-216, Springer LNCS, 2014

**Conference papers:**

- Immich, R. and Cerqueira, E. and Curado, M., "**Towards the Enhancement of UAV Video Transmission with Motion Intensity Awareness**", in the IEEE IFIP Wireless Days, 2014

- Cerqueira, E. and Quadros, C. and Neto, A. and Riker, A. and Immich, R. and Curado, M. and Pescapée, A., "**A Quality of Experience Handover System for Heterogeneous Multimedia Wireless Networks**", in the IEEE International Conference on Computing, Networking and Communications (ICNC), 2013

- Cerqueira, E. and Neto, A. and Immich, R. and Curado, M. and Riker, A. Barros, H., "**A Parametric QoE Video Quality Estimator for Wireless Networks**", in the GC'12 Workshop: IEEE Workshop on Quality of Experience for Multimedia Communications (QoEMC), 2012

# Chapter 7

# Mechanisms for Resilient Video Transmission over VANETs

All animals are equal, but some animals are more equal than others.

*(George Orwell, Animal Farm)*

## Contents

V ANETs are envisioned to offer support for a large variety of distributed applications such as road and traffic alerts, autonomous driving capabilities and video distribution. The use of video-equipped vehicles, with support for live transmission, unveils a set of challenges that can range from the scarce network resources and vehicles movement to the time-varying channel conditions and high error rates. Here again, to overcome these challenges an adaptive FEC-based scheme can be tailored to shield the video transmission with QoE assurance. This chapter details two proposed adaptive mechanisms to improve the video transmissions over VANETs.

# 7.1 Introduction

VANET is considered the core component of intelligent transportation systems, providing support to many applications. The endorsement of video-based services can be beneficial to a broad range of situations, such as road safety, driver awareness, traffic status, and infotainment applications. Besides the users' experience, the video quality is also important to allow a better assessment of each situation. For example, it can give police officers, paramedics, and firefighters an accurate representation of the scene they will attend, thereby reducing the response time. Beyond the traffic related services, a sports venue or a festival could broadcast a live feed to incoming fans caught in a traffic jam. These are only simple examples of a limitless number of alternatives to make available rich-format services. These services, however, have to deal with the unreliable wireless connection of the VANETs, which are highly dynamic in nature and strongly prone to packet loss [Zhou et al., 2011b, Gerla et al., 2014]. Therefore, it is imperative to strengthen the video transmissions against losses [Immich et al., 2013b, Immich et al., 2014c]. This calls for an adaptive mechanism to enhance the video delivery to provide higher QoE.

In VANETs, there is still a lack of adaptive QoE-driven mechanisms to better support live video transmissions [Soldo et al., 2011, Shen et al., 2011, Jiang et al., 2012, Bellalta et al., 2014]. This can be attributed to the challenging combination of the VANETs' dynamic topology and the stringent video requirements. In order to surpass these adversities, a good mechanism has to take into consideration several aspects of the intrinsic network characteristics and video details, being able to correctly identify and protect the most QoE-sensitive data.

As mentioned before, several techniques have been proposed to tackle the VANETs challenges in the last few years. Some of them are trying to solve these issues throughout adaptive routing protocols [Marwaha et al., 2004, Zeng et al., 2013, Pham et al., 2014, Wu and Ma, 2014, Zhang et al., 2015]. The results show that a reliable routing protocol has a major influence on improving the video quality. This improvement, however, is restricted to a specific level. After this level, to increase or even sustain the video quality it is crucial to resort to some amount of redundant data, which allow reconstructing the original data set in case of packet losses. A known approach to supply this redundancy is using FEC techniques. However, due to the video requirement of a timely delivery of a considerable amount of data [Zhou et al., 2011a], along with the shared wireless channel resources, a self-adaptive FEC-based mechanism is advisable. This mechanism needs to have the capability to operate under unforeseen conditions in order to increase the human perception while reducing the network overhead.

This chapter proposes and evaluates two QoE-aware mechanisms tailor-made to overcome any network mishaps, therefore providing the capability to deliver with high QoE for end-users. The first mechanism described is the adaptive QoE-driven COntent-awaRe VidEo Transmission opTimisation mEchanism (CORVETTE) in Section 7.2 and the latter is the self-adaptive FEC-based proactive

error recovery mechanism to shield video transmissions over VANETs (SHIELD), detailed in Section 7.3.

# 7.2 QoE-driven and Content-aware Video Delivery (CORVETTE)

Owing to the open issues aforementioned, particularly the shortage of adaptive QoE-driven mechanisms that efficiently consider the motion activity of the videos together with VANETs features, this section outlines and evaluates the adaptive QoE-driven COntent-awaRe VidEo Transmission opTimisation mEchanism (CORVETTE). The proposed mechanism enhances the work described in Sections 6.2 and 6.3. The main improvements and new experiment setups are described next.

## 7.2.1 CORVETTE Overview

The main goal of the CORVETTE mechanism is to improve the resilience of video transmissions over VANETs. One common issue found in the existing mechanisms is the lack of QoE-awareness. This means that important video characteristics, from the human perspective, are neglected, resulting in unnecessary redundancy. To tackle this issue, the CORVETTE uses a Hierarchical Fuzzy System (HFS) together with intrinsic VANET characteristics, to add a precise amount of redundancy exclusively on QoE-sensitive data. This ensures high-quality video transmission, while downsizing the network overhead.

Figure 7.1 depicts an overview of the proposed mechanism. The first step is to assess the network conditions (1). In order to do that, different parameters are evaluated in a combined way, namely the network density, PLR, and the node's position. To calculate the density, first, the network area is found using a hull algorithm. After that, the total number of 1-hop nodes is divided by the area, which gives the network density. All these parameters are necessary because none of them by itself is accurate enough to characterise the quality of the network links [Vlavianos et al., 2008, Wan et al., 2015]. The combination of them, however, can provide a very good estimation of the network conditions. Thereafter, using cross-layer techniques, important details about the video characteristics are collected (2). In the video-aware procedure of the mechanism, several details are analysed, such as the image resolution, frame type and size, motion vectors, and macroblock configuration. At the end, all the gathered data are fed to the fuzzy inference engine, which will compute a specific amount of redundancy (3).

Provided that the network conditions are not the same at all intermediate nodes, this parameter has to be reassessed at each hop (4). On the other hand, the video characteristics do not change during the transmission. In view of this, they are embedded in each packet header by the server node. This eliminates
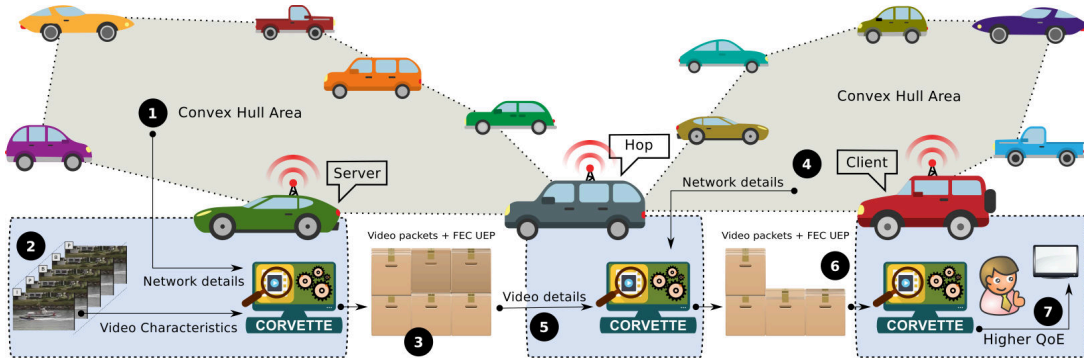
Figure 7.1: General view of the CORVETTE mechanism

the need for processor intensive tasks (e.g. deep packet inspection) on each and every packet. The IPv6 optional hop-by-hop header was chosen to store this information [Martini et al., 2007]. This means that it is always ready to use whenever needed (5,6). Owing to this, the task of adjusting the redundancy amount on each hop is facilitated. The result is a higher video quality, and consequently, superior QoE is perceived by the end-users (7).

## 7.2.2 Towards the design of CORVETTE

This section describes the manifold procedure and modules that the COR-VETTE mechanism is composed of. Primarily, to enable the CORVETTE real-time capabilities a knowledge database is needed. The procedure to build this database uses the same core structure as in the mechanisms described before with the additional improvements to accommodate the new parameters. A comprehensive description of this process can be found in Chapters 5 and 6.

An important feature to enable the CORVETTE real-time capabilities is the use of fuzzy logic. This allows building a dynamic and comprehensive scheme, which takes into consideration several network and video characteristics while still manages to perform in real-time. Nevertheless, in conventional fuzzy logic systems, the rules grow exponentially according to the number of variables. Therefore, it is common to have a rule-explosion situation when handling a lot of variables, making the fuzzy logic controller very hard to implement. To address this issue, the CORVETTE mechanism was designed to use a HFS, where low-dimensional fuzzy systems can be arranged in a hierarchical form, reducing the global number of rules because the system grows linearly.

The combined use of the knowledge database and human expertise enables setting up the fuzzy sets, rules, and hierarchical levels. Figure 7.2 shows the several hierarchical levels used by CORVETTE. The output of each low-level component in the previous layer is used as input to the next layer components. As mentioned before, on each network hop the amount of redundancy is adjusted according to several factors, such as the network density, the PLR, and the distance to the next hop (or final destination). This operation is represented by the (A) portion of the HFS structure. On the other hand, at the server node,

the full HFS structure is performed, which is composed of the (A) and (B) parts.
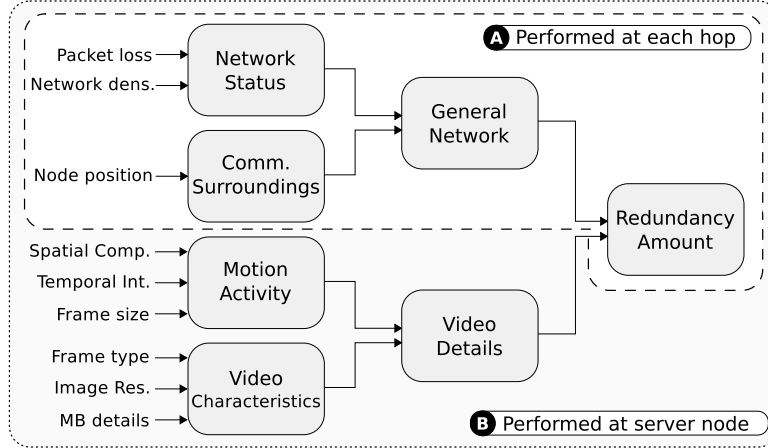


Figure 7.2: CORVETTE Hierarchical Fuzzy Logic structure

The design of HFS follows the same method as in standard fuzzy logic schemes. This means that several fuzzy components have to be defined, such as sets, rules, membership functions and the inference engine. Since this is a complex process, it has to be executed offline and only once. After loading all the generated information into the fuzzy engine, it can be used on-the-fly and with a high performance. A detailed explanation of this process is given below.

### 7.2.2.1 The "General Network" layer conception

The "General Network" component accounts for the definition of the network health. The characterization of a good or bad channel is not an easy task and it cannot rely upon a single metric, especially in wireless networks [Vlavianos et al., 2008]. With this in mind, the CORVETTE mechanism uses three metrics to better establish a network quality indicator. These metrics are divided into two specific components, namely "Network status" and "Communication surroundings". The former is defined by the combined assessment of the PLR and network density. The latter is given by the position of the vehicles. Each one of these metrics is described next.

To compute the network density, the number of nodes is divided by the network area. Because there is no fixed structure on VANETs and they can quickly change over time, it is challenging to estimate the network surface area. To solve this problem, the proposed mechanism uses a convex hull algorithm. In this type of algorithm, a convex polygon is drawn including all the nodes on the network. A polygon is convex when it is non-intersecting and a segment between any two points on the boundary lies entirely inside of it. Successively, a convex hull is the smallest polygon containing all the points.

There are several algorithms to find the convex hull, however, due to performance issues, the QuickHull method was chosen [Barber et al., 1996]. It uses a divide-and-conquer algorithm, and because of that, it is easier to find the convex hull

of small sets, and at the end, the discovered hulls are merged. Once the convex polygon is found, it is possible to calculate the surface area and use it to find the network density.

Another component of the "Network status" is the PLR. The main goal in defining this parameter is to identify the impact of different PLRs in the QoE. As mentioned before, it is well-known that video sequences have a natural resiliency to packet loss [Immich et al., 2013a]. To better define the PLR category, a broad number of network simulations with a large set of video sequences were carried out. The results show that the QoE is good from the PLR 0% up to 11%. Additionally, in most of the cases, from 5% as far as 22% a tolerable video quality for end-users was noticed. When exceeding 17% however, a quick decrease in the video quality was observed, especially in high motion intensity videos. The majority of the cases, over 34% of PLR, the QoE became unbearable. Using fuzzy logic it is possible to create classes of PLR exactly as found in the experiments, even with overlapping values. Algorithm 7.1 shows the PLR set definition.

---

**Algorithm 7.1:** Packet loss rate input set

---

**1** *InputLVar\* **PLR** = new InputLVar("**PacketLossRate**");*

**2**     PLR → addTerm( TriangularTerm("*LOW*", 0, 11));

**3**     PLR → addTerm( TriangularTerm("*MEDIUM*", 5, 22));

**4**     PLR → addTerm( TriangularTerm("*HIGH*", 17, 100));

**5** engine.addInputLVar(**PLR**);

---

The last component of the "General network" layer is the "Communication surroundings". In CORVETTE, this component is established by the node position. This is a straightforward, but very important information. Because of signal attenuation and radio-frequency interference, nodes further away from each other tend to require a higher amount of redundancy to preserve a good video quality. This information becomes even more valuable used in conjunction with the other input parameters. For example, a much higher amount of redundancy will be required if the network is very dense and the nodes are far apart than if the network was not so heavily populated.

### 7.2.2.2 The "Video Details" layer conception

Besides the network conditions, the video characteristics also play an important role when defining a precise amount of redundancy. In CORVETTE's hierarchical fuzzy system, the "Video details" criteria layer is divided into two specific components, namely "Motion activity" and "Video characteristics".

The motion activity parameter is defined by the combined use of temporal intensity and spatial complexity. The temporal intensity, in the CORVETTE mechanism, is given by the motion vector details. This structure is responsible for tracking the movement of macroblocks from one position, in any given frame, to another position, in the next one. Since the MPEG standard allows

the use of distinct macroblock sizes, the CORVETTE mechanism adopts the same procedure of the uavFEC and MINT-FEC mechanisms by computing the area of each macroblock and using the number of pixels that are being moved. This enables a better representation of the intensity of the motion in arbitrary resolutions.

In addition, the Euclidean distance of each motion vector is also computed, resulting on how far each and every macroblock is being moved. This information gives more precise results than just counting the number of motion vectors. All the input parameters are normalised, allowing the protection of videos with arbitrary resolutions on the fly. Following the same idea as presented before, an exploratory analysis using hierarchical clustering is performed to find the best classes that represent the temporal intensity. After finding the classes, the fuzzy set and the membership function can be defined. Finding the best-fitted membership function is a complex and problem-dependent task [Wong et al., 2005], being difficult to attain the optimal solution. Consequently, piecewise linear functions are desired.

As previously mentioned, the spatial complexity is also used to quantify the amount of the motion activity. This parameter represents the difference between the static information that the actual frame has when compared to the frame before. One way to compute this value is using the Sum of Absolute Differences (SAD) [Vanne et al., 2006]. This process, however, compares each and every pixel of both frames resulting in a very complex and time-consuming operation. Taking this into consideration, the CORVETTE mechanism uses the normalised frame size to the same end. This enables a much faster operation and, on top of that, it also allows the use of arbitrary video resolutions.

The same process used to find the different classes in the temporal intensity is also used to define the clusters here. This means that, once all the frame sizes are normalised, an exploratory analysis is performed to divide the data into the most homogeneous groups. After that, using the linkage distance between the clusters was possible to separate them into five distinct groups, namely "very small", "small", "medium", "large", and "very large", as showed in Figure 7.3.

With the definition of the fuzzy sets completed, the fuzzy rules must be designed. As mentioned before, this is an intricate task, which requires a joint knowledge of the network details, VANETs properties, and video characteristics. Aiming to reduce this complexity during the design phase of the rules, as well as to have a better performance in real-time, the CORVETTE mechanism uses HFS. This layered system allows handling fewer input parameters at the same time. At the end, the result is a system with a small number of simple rules, which lead to better performance.
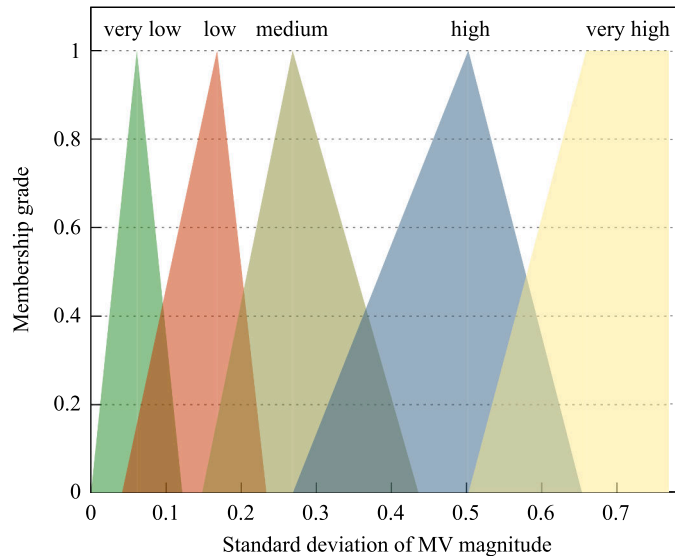
Figure 7.3: Temporal intensity membership function

## 7.2.3 CORVETTE Performance Evaluation and Results

The main goal of the CORVETTE mechanism is to improve the QoE for end-users. At the same time, it avoids unnecessary network overhead, preserving wireless resources.

### 7.2.3.1 Experiment settings

The simulated scenario is composed of up to 360 vehicles, using IEEE 802.11p WAVE. The speed range varies from 5 to 17 m/s. Several real video sequences with three distinct resolutions, namely 1080p, 720p, and SVGA, were used in the experiments. The videos are examples of regular viewing material, covering different content and distortions. These sequences also include still and cut scenes, colour and luminance stress, as well as several motion intensities. They were all encoded with H.264 codec and GoP length of 19:2.

Using the OpenStreetMAP [Haklay and Weber, 2008], a clipping of 2 by 2 km of the Manhattan borough (New York City) was obtained. This clipping was used as input for the Simulation of Urban MObility (SUMO) [Behrisch et al., 2011], which considers the map structure, driving patterns, routes, crossings, roundabouts, traffic lights, to generate the mobility traces, which were used in the NS-3 simulations.

Additionally, in this work, the V2V communication characteristics are explored to better adjust the proposed mechanism to the actual network conditions. Even though a VANET environment enables roadside infrastructure, the V2V communication was chosen because it is unlikely that such infrastructure will cover all the highways and cities in the near future. Consequently, if the infrastructure is available it can be used however, the optimisations will only be performed on the communication between the vehicles.

The adopted routing protocol is the Cross-Layer, Weighted, Position-based Routing (CLWPR) [Katsaros et al., 2011]. As the name implies, this is a position-based routing protocol that uses mobility information from the nodes to tailor itself for VANETs environments. Receiver nodes are using Frame-Copy error concealment; this means that if a frame is lost, the last good one will take its place. Table 7.1 shows the simulation parameters.

Table 7.1: CORVETTE Simulation parameters

| Parameters | Value |
|---|---|
| Display sizes | 1920x1080, 1280x720, and 800x600 |
| Frame rate mode | Constant |
| Frame rate | 29.970 fps |
| GoP | 19:2 |
| Video format | H.264 |
| Codec | x264 |
| Container | MP4 |
| Propagation model | FriisPropagationLossModel |
| Mobility | SUMO mobility traces |
| Routing protocol | CLWPR |
| Wireless | IEEE 802.11p (WAVE) |
| Radio range | 250-300m |
| Internet layer | IPv6 |
| Transport layer | UDP |
| Location | Manhattan borough (New York City) |
| Map size | 2.000 m x 2.000 m |
| Vehicles speed | 18-61 km/h (11-38 mph) |

Five distinct mechanisms were simulated as stated next: (1) Without FEC, to be used as a baseline; (2) Video-aware Equal Error Protection FEC (VaEEP), where I- and P-frames are equally protected; (3) Video-aware Unequal Error Protection FEC (VaUEP). This mechanism protects I- and P-frames with a specific amount of redundancy according to the importance of each one; (4) the adaptive FEC-based mechanism (AdaptFEC), which takes into consideration several video characteristics and the network state [Immich et al., 2014d]. At last, (5) the proposed CORVETTE mechanism.

### 7.2.3.2 QoE assessments

Figure 7.4 shows the SSIM results. The first thing to be noticed is that when the simulation is performed with a few vehicles, e.g., 40, the network is sparse and suffers from connectivity issues. The CORVETTE mechanism proves that it can handle those situations by adjusting the amount of redundancy correctly, thus, outperforming competitors. It is also important to perceive that the standard deviation is high in all mechanisms. This can be expected since some video sequences tend to be naturally more resilient to packet loss than others, resulting in distinctive quality scores. This is especially true in low motion intensity videos. On the other hand, when the network is very dense, e.g., above 300 cars,

the mechanisms have to deal with interference and a degraded network connection. Here again, CORVETTE outperforms the other mechanisms, providing better video quality.
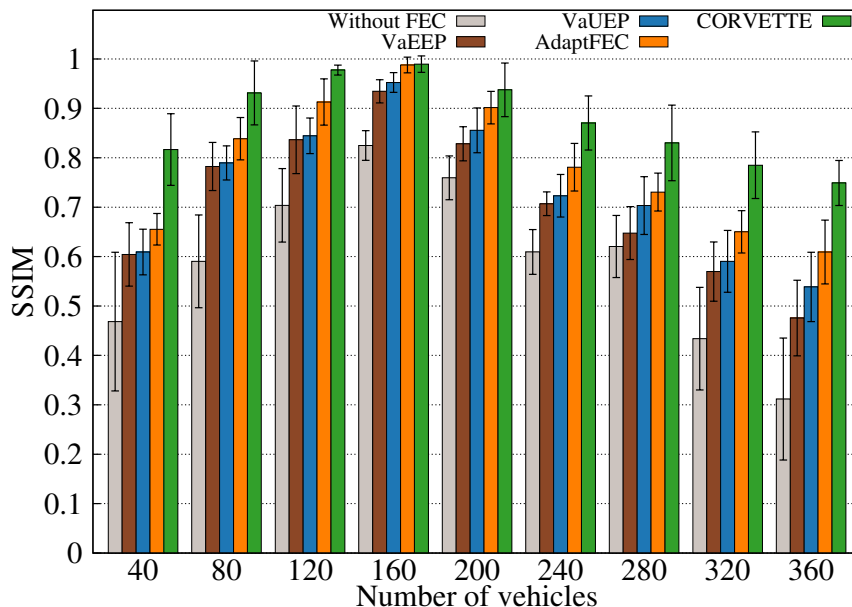


Figure 7.4: Average SSIM for all mechanisms

Overall, the proposed mechanism exceeds by 48% the without FEC scheme when it comes to video quality. Against VaEEP and VaUEP schemes, the CORVETTE mechanism granted over 23% and 19% better scores, respectively. In comparison to AdaptFEC, the proposed mechanism assured more than 11% higher video quality. At first, this value does not seem to be very high however, it was able to achieve it while adding 41% less network overhead.

Figure 7.5 presents the VQM scores. Almost the same pattern found in the SSIM metric is also present here. When the network is sparse, the videos tend to have lower quality, especially without any FEC-based mechanism. The best-case scenario in the experiments is between 160 and 240 vehicles. In this range, the videos were transmitted with better quality. Similarly to the SSIM assessment, using VQM the CORVETTE also outperformed all other mechanisms. Once again, this proves that the transmission enhancements performed by the proposed mechanism lead to a better video quality for the end-users. On average, the CORVETTE mechanism provided 65% better video quality than the scheme without FEC, 42% and 40% higher scores than VaEEP and VaUEP, respectively. Against the AdaptFEC scheme, the proposed mechanism presented 30% better results.

### 7.2.3.3 Network footprint analysis

Apart from higher video quality, a lower network overhead is equally desirable, which is shown in Figure 7.6. The VaEEP and VaUEP schemes are non-adaptive,
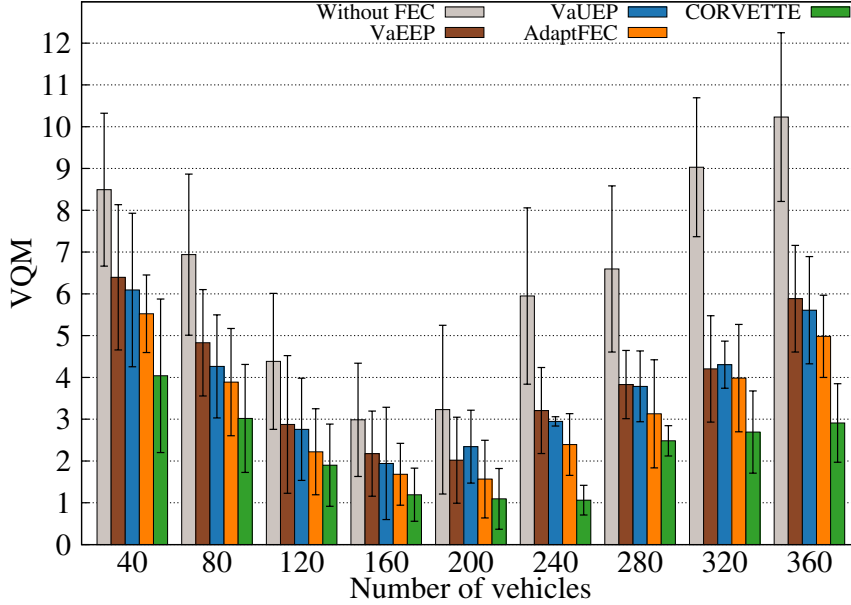
Figure 7.5: Average VQM for all mechanisms

thus producing a constant network overhead through the experiments. As can be seen in the graph, the non-adaptive protection schemes produce a large amount of network overhead, especially in the VaEEP case. The VaEEP protection is also not very efficient because it shields certain parts that do not require protection. Therefore, the VaUEP scheme takes into consideration the video characteristics, namely the frame type, and adds a specific amount of redundancy to each one. This grants the reduction of the network footprint while improving the video quality. Unlike the VaEEP, the graph shows a standard deviation for the VaUEP values because different video sequences require distinctive amounts of redundancy.

On the other hand, the adaptive mechanisms (AdaptFEC and CORVETTE) yielded better results in terms of network overhead, allowing a proper usage of the wireless resources. Another important difference is the standard deviation, where the CORVETTE mechanism has higher values. In this case, this is very desirable, because it means that the proposed mechanism was able to better identify the most important portions of the video sequences and protect them accordingly. This demonstrated a tailored protection, providing both superior video quality and low network overhead. On average, the CORVETTE mechanism added 41% less overhead than AdaptFEC. When compared to the VaEEP and VaUEP schemes, the proposed mechanism was able to generate 70% and 58% less overhead, respectively.

### 7.2.3.4 Overall results

Table 7.2 summarises the average SSIM, VQM, and network overhead scores. It shows that the CORVETTE mechanism was able to provide a substantial reduction in network overhead. This was possible by adding a specific amount
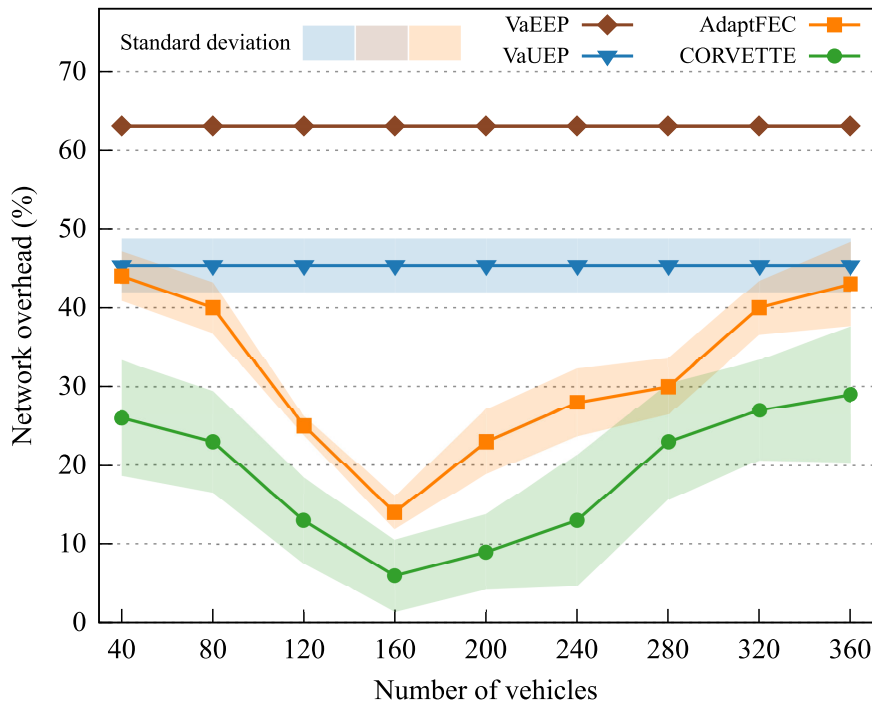
Figure 7.6: Average network overhead

of redundancy to each video sequences, and thus, avoiding any unnecessary redundancy. Additionally, it did so while increasing the video quality, leading to higher QoE for end-users.

Table 7.2: Average SSIM, VQM, and network overhead

|  | **CORVETTE** | **AdaptFEC** | **VaUEP** | **VaEEP** | **Without FEC** |
|---|---|---|---|---|---|
| SSIM | 0,876 | 0,785 | 0,734 | 0,709 | 0,591 |
| VQM | 2,266 | 3,264 | 3,783 | 3,935 | 6,425 |
| Overhead | 18,777% | 31,888% | 45,342% | 63,102% | — |

Taking everything into consideration, the CORVETTE mechanism showed results with better video quality while reducing considerably the network overhead. This is only possible because it performs an accurate motion intensity classification, which is used along with other video and network characteristics. The combined use of all this information enables the mechanism to add an adequate protection to each video sequence improving the quality for end-users.

## 7.3 Self-adaptive Mechanism to Shield Video Transmissions (SHIELD)

The CORVETTE mechanism was able to attain encouraging results, notwithstanding, there was still room for further improvement. For example, more

parameters could be considered to enhance the reliability of the network conditions assessment, as well as, a reduced number of operations could be possible by changing the convex hull approach. To address these open issues, the self-adaptive FEC-based proactive error recovery mechanism to shield video transmissions over VANETs (SHIELD) was proposed. The main enhancements are described next.

## 7.3.1 SHIELD Overview

The aim of the mechanism proposed in this section is to offer videos with higher QoE to video-equipped vehicles while, at the same time, downsizing the network overhead footprint. To this end, SHIELD extends the CORVETTE mechanism by using several video characteristics and specific VANETs details to safeguard real-time video streams against packet losses. This means that meaningful video aspects related to the human point-of-view, are not neglected, which leads to the addition of a very specific amount of redundancy. One of the main contributions of this work over the CORVETTE mechanism is the combined used of network density, SNR, PLR, and the vehicle's position. This allows SHIELD to better protect the video sequences and enhance the QoE.

## 7.3.2 Towards the design of SHIELD

The SHIELD mechanism is an improvement over the CORVETTE work presented in Section 7.2. In the same way as presented before, it needs a knowledge database which is populated before the mechanism execution. Following the same basic core structure as before, SHIELD also employs a HFS. This allows it to reduce the global number of rules by arranging them in a hierarchical form, thus, making it possible to be performed in real-time. Despite the fact that the same HFS structure was used, all rules and sets regarding the network parameters were overhauled. On the other hand, the video related details were kept the same. Figure 7.7 depicts the hierarchical levels adopted by SHIELD, as well as the components that were upgraded. The main improvements are detailed below.

There are three layers in SHIELD HFS, namely (A) Objective function, (B) General criteria, and (C) Specific criteria. The output of each low-level layer is used as input to the next layer. The first layer (A) represents the amount of redundancy that SHIELD will add to a specific portion of video data. The main goal is to find the amount of redundancy for a system that, given its constraints, results in less network overhead and better QoE. The second layer (B) encompasses the overall details that the proposed mechanism uses to determine the redundancy amount, namely the network details and the video characteristics. The bottom layer (C) is responsible for handling the input parameters of each feature used by the fuzzy logic inference system. This layer has a subdivision (C)(2), which is performed at each network hop. All the input parameters (C)(1) are only taken into consideration at the server node.
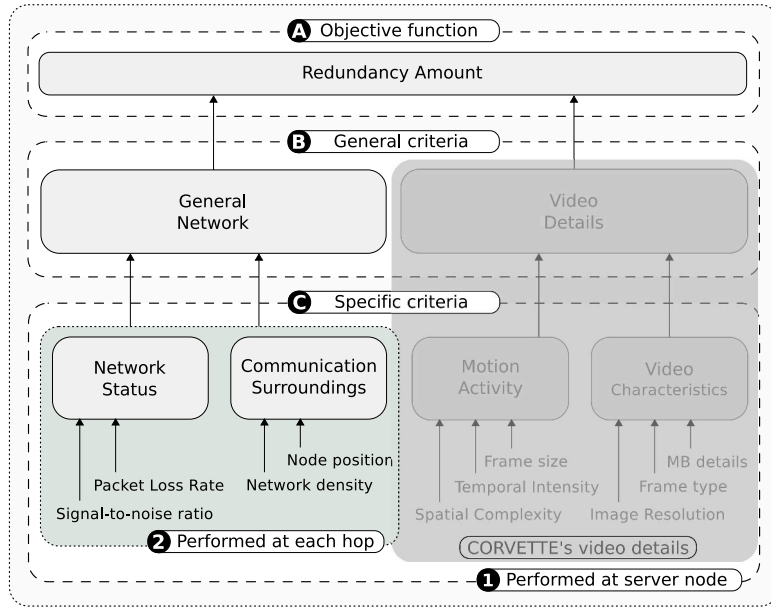
Figure 7.7: SHIELD HFS structure

The "General Network" criteria account for the definition of the network conditions. This component was completely upgraded in the new mechanism. All the hierarchical rules were rewritten, the input parameters were modified, as well as a new one was added, namely the SNR. This enables the mechanism to use more metrics to better establish a network quality indicator. In total, the SHIELD mechanism uses 4 metrics. These metrics are divided into two specific criteria, namely "Network status" and "Communication surroundings". The former is defined by the combined assessment of the SNR and the PLR. The latter is given by the network density and the position of the vehicles. Each one of these metrics is described next.

The SNR is the level of the desired signal against the level of background noise. This is a good indicator for the physical medium, especially for spectrum sensing. While this is true, it cannot be considered a reliable general network quality indicator by itself. This steams from the fact that a strong channel signal will not always produce a good network connection [Vlavianos et al., 2008]. On the other hand, a very weak signal will yield a low-quality network connection. As a result, to create a more holistic indicator more than one metric has to be used. Another obvious candidate to define the network quality is the PLR. In general, the SNR and PLR have a negative correlation, meaning that when one increases, the other decreases and vice versa. However, they complement each other because the SNR takes into consideration the physical spectrum part of the transmission and the PLR provides a point of view closer to the application layer.

The PLR input set was also redesigned. In the same way as before several experiments were carried out to find the relationship between the losses and QoE, using a more comprehensive number of video sequences. The results made evident that it is possible to have a good QoE with packet loss between 0%

and 12%, a small difference from CORVETTE's PLR input set (between 0% and 11%), yet a step forward to better characterise the network conditions. In most of the cases, an acceptable video quality for end-users was perceived with losses from 5% up to 23%. Here again, another small difference in the found values, which in CORVETTE case were between 5% and 22%. Additionally, after a threshold of 19% (was 17% in the CORVETTE mechanism) the video quality starts to decrease apace, particularly in videos with high resolution and motion intensity. In the experiments with more than 36% (34% in CORVETTE implementation) of PLR, the QoE reached intolerable levels. Algorithm 7.2 shows only one of the many fuzzy sets defined in the SHIELD mechanism.

---

**Algorithm 7.2:** SHIELD PLR input set

---

**1** *InputLVar\* **PLR** = new InputLVar("**PacketLossRate**");*
**2**   PLR → addTerm( TriangularTerm("*LOW*", 0, 12));
**3**   PLR → addTerm( TriangularTerm("*MEDIUM*", 5, 23));
**4**   PLR → addTerm( TriangularTerm("*HIGH*", 19, 100));
**5** engine.addInputLVar(**PLR**);

---

Another component of the "General network" criteria is the "Communication surroundings". This component also was enhanced during SHIELD design. It uses the network density and the position of the vehicles to provide more information about the network in which the video sequences are being transmitted. These parameters are updated at each beacon exchange in the routing protocol. The network density is given by the number of nodes, in this case vehicles, divided by the network area. It is important to notice that VANETs are very dynamic networks with a decentralised structure, proving to be a challenge for the estimation of the network surface area. To address this issue, the proposed mechanism uses an approximate convex hull algorithm.

A convex hull algorithm is able to find the smallest boundary polygon containing all the points inside of it, using only non-intersecting segments, as showed in Figure 7.8(a). There are several algorithms to find the convex hull of a given set of points. In the CORVETTE mechanism, the QuickHull method was used. It uses a divide-and-conquer algorithm with average complexity of $\mathcal{O}(n \log n)$ and at the worst case, it could take $\mathcal{O}(n^2)$. However, it is not imperative for the mechanisms to use high-precision value for the area size, instead, a good approximation is sufficient to provide very good results. Owing to this, and to improve the general performance, SHIELD uses the Bentley Faust Preparata (BFP) [Bentley et al., 1982] approximation convex hull algorithm as showed in Figure 7.8(b).

The BFP algorithm, which runs in $\mathcal{O}(n)$ time, replaces the sort operation by dividing the plane into vertical strips. In each strip, the minimum and maximum points are found and added to the boundary. This algorithm is an approximation because a non-extreme point, in a given strip, can be discarded even if it is on the convex hull boundary. Nevertheless, the point will not be far from the convex hull, resulting in a good approximation of the actual convex hull.

Figure 7.9 shows the comparison between the number of nodes and the resulting amount of operations in both QuickHull and BFP algorithm. On average, the

(a) Convex Hull   (b) Approximate Convex Hull
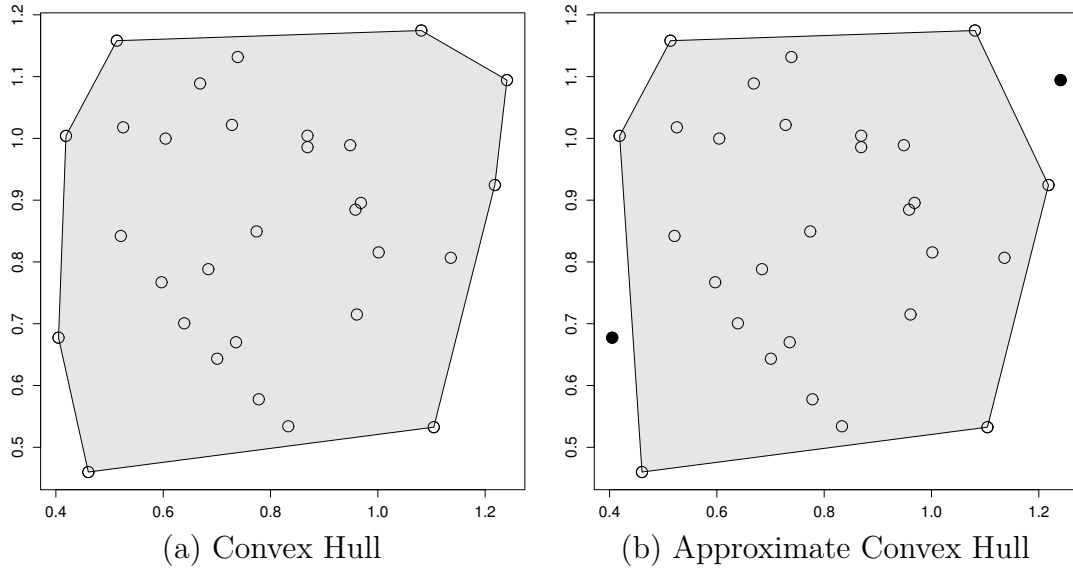
Figure 7.8: Convex Hull and Approximate Convex Hull

QuickHull algorithm has fairly good performance, it can degrade, however, up to exponential in the worst case. On the contrary, the BFP algorithm has a steady linear performance, providing results more quickly even with a small number of nodes.
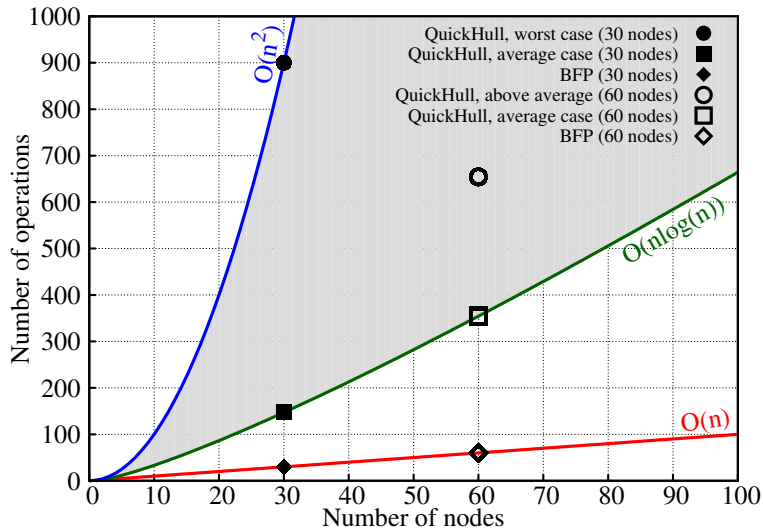


Figure 7.9: Complexity of QuickHull and BFP

There is a multitude of advantages of performing fewer operations. First of all, due to the time-sensitive video data, it is important that the client node receives the information as soon as possible, thus performing fewer operations allows dispatching the video quickly. In addition, because of the fast time-varying network conditions, the faster this information is made available the more accurate it is. At last, performing a minimum number of operations means less energy consumption as well as more available processor power to perform other tasks.

Another noteworthy betterment in SHIELD is the approximation of the node position and the network density inputs. The latter parameter was moved from the "Network status" component to the "Communication surroundings". In doing that, a different set of rules can now be created to better correlate the movement of the nodes and the number of nodes with 1-hop distance in a given area. These parameters provide more accurate results when analysed together than independently, enabling SHIELD to yield a precise network assessment.

In short, the adoption of another input (SNR) and the adjustment of the "Communication surroundings" component led to a complete rearrangement of the "General network" module. Which in terms, led better-consolidated rules that express with higher precision the needs of the proposed mechanism. At the end, this new "General network" layer presented is responsible for the better integration of SHIELD in VANETs. In other words, it allows the proposed mechanism to ascertain the best-fitted QoE-aware amount of redundancy according to each video sequence.

## 7.3.3 SHIELD Performance Evaluation and Results

The main goal of the SHIELD mechanism is to enhance the QoE while avoiding any unnecessary network overhead. In doing that, it improves end-users satisfaction and preserves the already scarce wireless resources at the same time.

### 7.3.3.1 Experiment settings

In order to better characterise the performance of the proposed mechanism two distinct environments were assessed: urban and highway. Each of these surroundings features a variety of unique challenges. In the urban environment, there are buildings and many other structures that will affect the signal propagation. On the other hand, in the highway environment, there is much more free space, which facilitates the signal propagation. Besides that, the mobility patterns are also very distinctive. The urban scenario presents a lot of driving options, such as avenues and streets close to each other. On the highway is quite different, as there are no crossroads and just a few exits and entrances. In addition, the speed of the vehicles has very particular properties in each one of these environments. In the urban case, the velocity usually is between 20 km/h and 60 km/h, and it changes frequently due to traffic lights, speed bumps, and crosswalks. Meanwhile, on the highway, the speed variance is very low, generally staying from 80 km/h to 120 km/h.

Several configurations are shared between the two environments, such as the wireless and network technology (IEEE 802.11p WAVE with V2V and CLWPR routing protocol), as well as the video content and codification parameters. All videos were sent using Evalvid Tool and encoded with H.264, GoP length of 19:2. Additionally, three different resolutions were used, namely 1080p, 720p, and SVGA. For each resolution, 10 videos were chosen to be transmitted [Xiph.Org, 2016]. A multi-flow scenario was adopted. This means that up to 10 videos are

transmitted simultaneously[1].  All the receiver nodes are enabled with Frame-Copy error concealment.

The mobility traces for both environments were generated using SUMO. For the urban environment, a clipping of 2 by 2 km of the Manhattan borough (New York City) was used. This environment was simulated with up to 360 vehicles at speeds ranging from 20 km/h and 60 km/h. To simulate this environment a clipping of 10 km of US Interstate Highway 78 (I-78) was used. The number of vehicles is the same, up to 360, with a velocity between 80 km/h and 120 km/h.

Two different propagation models were used to better represent each environment. In the highway scenario, the logDistance propagation model was used [Mittag et al., 2011]. This is because of the open spaces and the reduced number of sources of interference existent in this environment. This leads to easier communication between the nodes. On the other hand, in the urban environment, there are plenty of sources generating interference. In view of this, the Nakagami-m propagation model was added on the top of the logDistance model. This allows simulating the fast fading characteristics commonly found in this environment [Taliwal et al., 2004]. Table 7.3 summarises the simulation parameters.

Figure 7.10 shows the steps involved in the experiment. First of all, the mobile traces are required from the SUMO application (1). After that, SUMO will use real map clippings from the OpenStreetMap (2) to generate the traces. The traces enable a realistic simulation, providing more accurate results. Following this, all the information is loaded in the SHIELD mechanism (3). Next, the proposed mechanism will assess the network conditions (4) and request the video to be transmitted (5). Real video sequences are used in the experiment (6). Afterwards, the SHIELD mechanism optimises and secures the video transmission against packet loss (7). The next step is to deliver the video sequences to the receiver (8). At the end, the original (9) and the transmitted (10) videos are assessed using objective QoE metrics (11).

Five different scenarios were assessed in both urban and highway environments. The first one is without any type of FEC. The results of this experiment will be used as a baseline for the others. The second scenario is the Video-aware Equal Error Protection FEC (VaEEP) mechanism. The Video-aware Unequal Error Protection FEC (VaUEP) mechanism is the third scenario. The fourth scenario is using the adaptive QoE-driven COntent-awaRe VidEo Transmission opTimisation mEchanism (CORVETTE), presented in Section 7.2. The fifth and last scenario is the proposed SHIELD mechanism.

---

[1]Samples of the transmitted videos are available in `http://www.youtube.com/channel/UCsBOSdKpCKD2GS6aXzB-FUQ/videos`

Table 7.3: Simulation parameters

| Parameters | Value |
|---|---|
| Display sizes | 1920x1080, 1280x720, and 800x600 |
| Frame rate mode | Constant |
| Frame rate | 29.970 fps |
| GoP | 19:2 |
| Video format | H.264 |
| Codec | x264 |
| Container | MP4 |
| Wireless technology | IEEE 802.11p (WAVE) |
| Communication | Vehicle To Vehicle (V2V) |
| Routing protocol | CLWPR |
| Mobility | SUMO mobility traces |
| Radio range | 250m |
| Internet layer | IPv6 |
| Transport layer | UDP |
| **Highway environment** | |
| Propagation model | logDistance |
| Location | I-78 |
| Map size | 10.000 m |
| Vehicles speed | 80-120 km/h (50-75 mph) |
| **Urban environment** | |
| Propagation model | logDistance + Nakagami-m |
| Location | Manhattan borough(New York City) |
| Map size | 2.000 m x 2.000 m |
| Vehicles speed | 20-60 km/h (12-37 mph) |
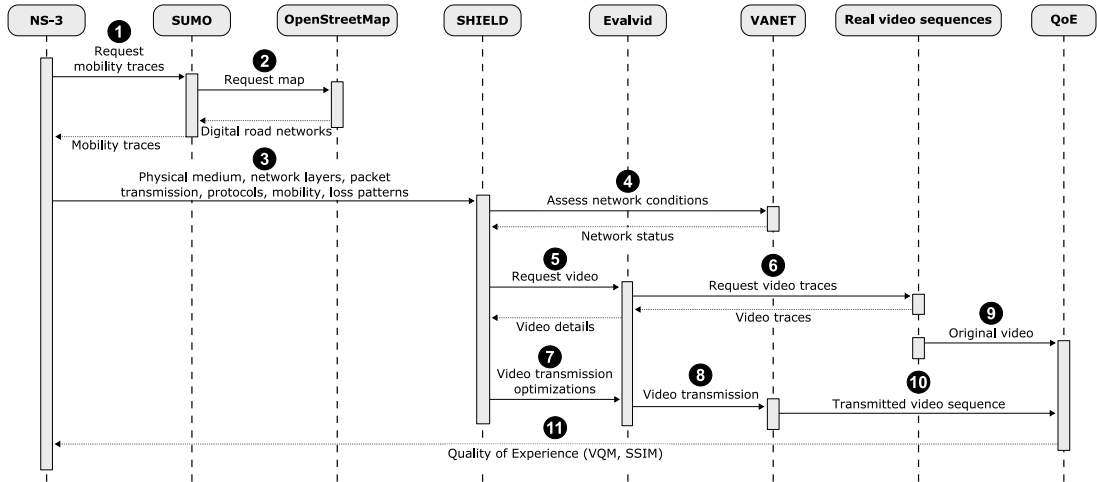


Figure 7.10: Steps involved in the experiment

## 7.3.3.2 QoE assessments

Figure 7.11 shows QoE results of the urban scenario. (a) depicts the SSIM average and (b) shows the QoE improvement in comparison to the base line. In (a),

it is possible to notice that the simulation starts with a few vehicles and the QoE results, for all mechanisms, can be considered low. This can be credited to the fact that the network is suffering from connectivity issues because it is relying on very few and scattered nodes to transmit all the video data. Even in this scenario, the SHIELD mechanism was able to protect the most important parts of the video sequences, producing better results. As showed in (b), this led to an improvement of more than 90% on the video quality when compared to the baseline (without FEC). The second best result was the CORVETTE mechanism with 60% of SSIM improvement.



(a) Average SSIM

(b) Percentage of SSIM improvement
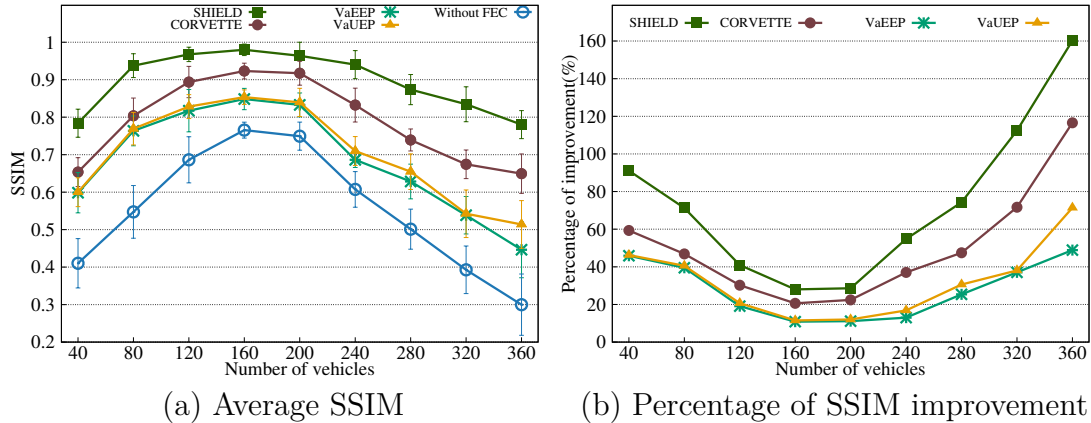
Figure 7.11: SSIM assessment of the urban environment

Figure 7.11(a) also shows that the best QoE results for all mechanisms are obtained when the network has 160 and 200 vehicles. This number of nodes provides the best coverage of the whole area, while it does not cause excessive interference. Because of the improved network conditions, the baseline also has better results, thus reducing the SSIM improvement perceived by the other mechanisms. This situation is clearly evidenced in Figure 7.11(b) for 160 and 200 vehicles. On the other hand, when the network becomes very dense, e.g., above 280 vehicles, the mechanisms have to face increasingly degraded network connections. Once again, the SHIELD surpassed the other mechanisms, providing up to 160% higher SSIM scores in comparison to the baseline.

Furthermore, Figure 7.12(a) presents the VQM average and (b) depicts the percentage of QoE improvement of the mechanisms in comparison to the baseline. Although this metric differs from SSIM, almost the same pattern can be found in (a). At the beginning of the experiment, the network is sparse and the videos have low quality. VQM gives them high scores, which in this case are not good. This is especially true for the baseline because it does not use any type of FEC-based mechanism to secure the transmissions. The best-case scenario in the VQM scores is the same as in the SSIM results, for 160 and 200 vehicles. This confirms the notion that the videos are transmitted with better quality with this configuration. In the same way as in the SSIM assessment, the VQM scores demonstrate that the SHIELD mechanism outperforms all other mechanisms.

Additionally, Figure 7.12(b) shows a pattern similar to the SSIM results. The

(a) Average VQM

(b) Percentage of VQM improvement

Figure 7.12: VQM assessment of the urban environment

highest improvements are accomplished when the network is sparse, between 40 and 120 vehicles, or in a very dense network, above 240 vehicles. On average, the proposed mechanism provided scor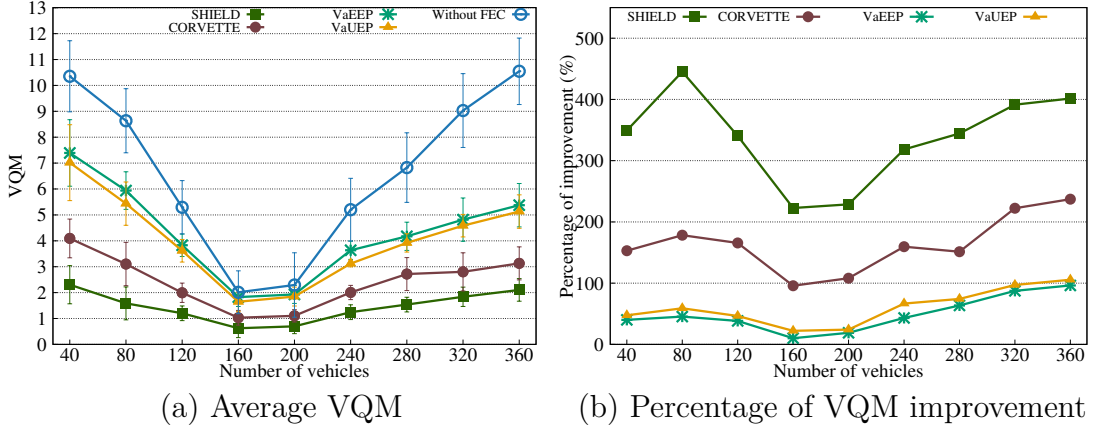es 78% better than the baseline. Additionally, it achieved 66% and 63% higher marks than VaEEP and VaUEP, respectively, and over 40% better scores in comparison to the CORVETTE mechanism.

In addition to the urban scenario, a highway environment was also assessed with both SSIM and VQM metrics. Figure 7.13 shows the QoE assessment, whereas (a) depicts the average SSIM and (b) shows the percentage of improvement achieved by the mechanism against the baseline. In (a), the first thing to be noticed is that the QoE results are closer to one another in this environment. This happens because the network conditions are not as harsh as in the urban scenario. At first, there are some connectivity issues when the network is sparse, e.g., 40 vehicles. After this threshold, a better video quality is being provided. The best results are evidenced for 120 and 240 vehicles. In (b), it is possible to notice that the highest improvements are reached when connectivity issues affect the network, for example, when the number of deployed vehicles is 40 and 80. In addition, major improvements are also perceived when there is a higher level of interference, such as above 280 vehicles. Here again, the SHIELD mechanism outperforms all its competitors.

The average VQM is shown in Figure 7.14(a) and the percentage of VQM improvement by each mechanism is shown in (b). In (a), the results follow the same tendency as the SSIM scores. This means that the VQM results are also closer to one another, especially above 120 vehicles. This is evidenced because the highway environment is not as rough as the urban setting. In (b), it is clear that the highest percentage of improvement is achieved when the nodes are sparse. This means that there are connectivity issues in the network, e.g., for 40 and 80 vehicles. After this threshold, the network conditions improve and the enhancements provided by the mechanisms decrease. Nevertheless, the SHIELD mechanism is able to surpass the competitors.

(a) Average SSIM        (b) Percentage of SSIM improvement
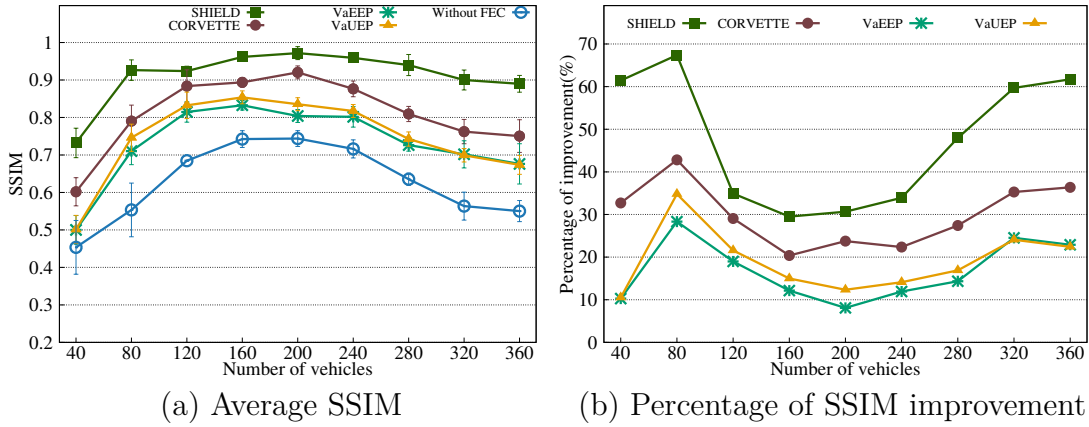
Figure 7.13: SSIM assessment of the highway scenario



(a) Average VQM        (b) Percentage of VQM improvement
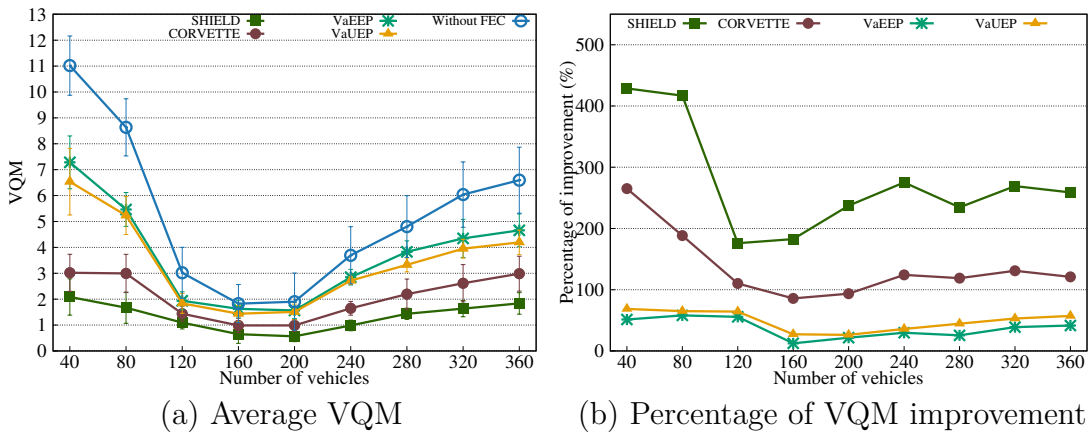
Figure 7.14: VQM assessment of the highway scenario

### 7.3.3.3 Network footprint analysis

Figure 7.15 shows the network overhead of all mechanisms in both (a) urban and (b) highway environments. The non-adaptive VaEEP and VaUEP schemes yield a constant network footprint in both scenarios because they do not adapt the amount of redundancy according to the network conditions. As depicted in the graph, these non-adaptive schemes add a considerably larger amount of redundancy. On top of that, the protection is not very efficient because, in the VaEEP case, the protection is added equally to all video data. As highlighted before, not all video packets need the same degree of protection. To tackle this issue VaUEP considers the frame type to add a specific amount of redundancy. This results in less network overhead and, at the same time, improves the video quality.

The VaEEP mechanism does not have a standard deviation because it uses a unique and pre-defined amount of redundancy, which is applied equally to all videos. The VaUEP mechanism also has a pre-defined amount of redundancy, but it is not unique. This means that each type of video frame will have a specific amount of redundancy. Additionally, each video frame has a different

(a) Urban scenario

(b) Highway scenario
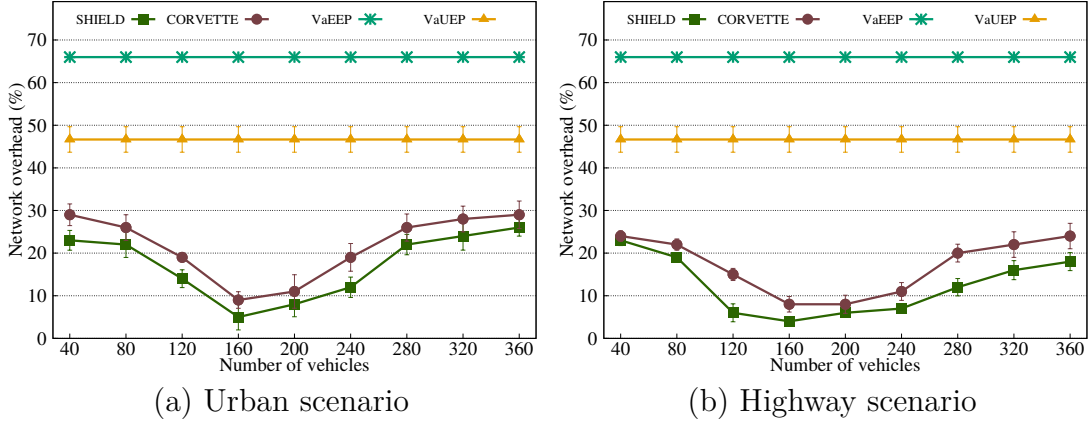
Figure 7.15: SHIELD Network overhead

size, leading to a variation in the amount of redundancy, and thus, a standard deviation is displayed.

The adaptive mechanisms, CORVETTE and SHIELD, were able to produce a lower network overhead, improving the wireless resources usage. In both mechanisms, when the network condition is better the footprint decreases. In the urban environment, this is evidenced when the simulation has 160 vehicles. The SHIELD mechanism produces a network overhead of only 5%, while CORVETTE is producing 9%. In the highway environment, the best conditions are experienced between 120 and 240 vehicles. The overhead produced by the SHIELD mechanism was between 4% and 7%, while CORVETTE is producing between 8% and 15%.

On average, the SHIELD mechanism added 20% less overhead in the urban environment and 28% less in the highway scenario, in comparison to the CORVETTE mechanism. When compared to the VaEEP and VaUEP mechanisms, the SHIELD mechanism produced 73% and 63% less overhead, respectively in the urban scenario. In the highway scenario, the network overhead downsize was 81% and 73%, respectively. In the end, the proposed mechanism was able to produce a tailored protection, enabling a higher video quality and lower network overhead.

### 7.3.3.4 Overall results

The overall results demonstrate that the SHIELD mechanism outperforms all its competitors as showed in Table 7.4, which summarises the average SSIM, VQM, and the network footprint. The SHIELD mechanism enables downsizing the network footprint in both urban and highway environment. This stems from the fact that a tailored amount of redundancy, based upon video characteristics and the network conditions, is added to each video sequence, preventing any unnecessary redundancy. Furthermore, the proposed mechanism also enhanced the quality of the video delivered, thus providing higher QoE for the end-users.

Table 7.4: Average SSIM, VQM, and network footprint

| | **SHIELD** | **CORVETTE** | **VaUEP** | **VaEEP** | **Without FEC** |
|---|---|---|---|---|---|
| **Urban environment** | | | | | |
| SSIM | 0,895 | 0,787 | 0,701 | 0,684 | 0,551 |
| VQM | 1,459 | 2,441 | 4,034 | 4,323 | 6,688 |
| Overhead | 17,333% | 21,778% | 46,660% | 65,984% | – |
| **Highway environment** | | | | | |
| SSIM | 0,911 | 0,809 | 0,744 | 0,729 | 0,627 |
| VQM | 1,328 | 2,095 | 3,414 | 3,728 | 5,281 |
| Overhead | 12,333% | 17,112% | 46,660% | 65,984% | – |

# 7.4 Summary

This chapter described and assessed two proposed mechanisms to increase the video transmission resiliency over VANETs. The CORVETTE mechanism detailed in Section 7.2 enabled safeguarding video delivery over highly dynamic networks, as well as improving the QoE perceived by end-users. In the VQM assessments, it was able to surpass the AdaptFEC mechanism presenting results with 30% higher QoE. When it comes to the SSIM assessment, it generated videos with 11% higher quality. This value was reached in conjunction with a 41% decrease on the network overhead, which proves to be a good overall result. Regardless of the CORVETTE mechanism satisfactory results, there was room for improvement in the way it handled the network-related parameters.

In order to improve on the aforementioned issues, the SHIELD mechanism, described in Section 7.3, was proposed. In the urban environment, SHIELD achieved between 27% and 57% better QoE in comparison to the CORVETTE mechanism. On the same scenario, it produced only 17% of network overhead, at the same time, CORVETTE was producing 22%. In the highway environment, the SHIELD QoE improvement was between 20% and 87% better than CORVETTE, and also producing less network overhead, only 12% in comparison to the 17% of CORVETTE. The assessment results clearly show that the latter mechanism was able to outperform the other competitors, including the CORVETTE mechanism. This corroborates that the improvements made to the network module result in both, better QoE and lower network footprint.

The following publications resulted from the work carried out in this chapter:

**Journal paper:**

- Immich, R. and Cerqueira, E. and Curado, M., "**Shielding video streaming against packet losses over VANETs**", The Journal of Mobile Communication, Computation and Information, Wireless Networks, Volume 22, Issue 8, pp 2563-2577, Springer, 2015

**Conference papers:**

- Immich, R. and Cerqueira, E. and Curado, M., "**Towards a QoE-driven Mechanism for Improved H.265 Video Delivery**", in the 15th IFIP Annual Mediterranean Ad Hoc Networking Workshop (MED-HOC-NET), 2016

- Immich, R. and Cerqueira, E. and Curado, M., "**Adaptive QoE-driven video transmission over Vehicular Ad-hoc Networks**", in the IEEE Conference on Computer Communications Workshops (INFOCOM), 2015

- Cerqueira, E. and Curado, M. and Neto, A. and Riker, A. and Immich, R. and Quadros, C., "**A Mobile QoE Architecture for Heterogeneous Multimedia Wireless Networks**", in the GC'12 Workshop: The 4th IEEE International Workshop on Mobility Management in the Networks of the Future World (MobiWorld), 2012

# Chapter 8

# Conclusions and Future Work

As if one believed anything by instinct!
One believes things because one has
been conditioned to believe them.

*(Aldous Huxley, Brave New World)*

## Contents

The video delivery over wireless networks is growing apace, every day hundreds of thousands of users and companies produce, share, and consume a vast amount of this type of content. This is an important way to convey educational information and news reports, as well as the most varied forms of entertainment. Despite numerous efforts on this field, several open issues still remain to be solved. This thesis proposed several mechanisms to improve the quality of the video transmissions in diverse wireless networks contributing to the advancement of the state-of-the-art in this topic. This chapter revises the addressed problems emphasising the major contributions accomplished in this thesis. At the end, future research opportunities and directions are outlined.

## 8.1 Synthesis of the Thesis

An effective method for making video transmission resilient to losses is of critical importance for the success of video transmission over wireless networks, this is even more evident in multi-hops networks. Due to the recent technological advances, the end-users of today desire high-quality video stream, services that do not follow this tendency will have poor acceptance. However, to guarantee a video transmission with high perceived quality is a complex task that depends on a multitude of elements.

One of the most evident parameter that has an impact on the video quality are the network-related details. Taking this into consideration, Chapter 2 presented an overview of the network technologies involved in this thesis. It started with a review of cross-layer techniques, which are necessary to provide the flexibility needed to implement any network transmission optimisation. After that, particular details of each multi-hop wireless networks are explored, such as the routing protocols and channel characteristics, as well as the mobility and traffic patterns.

In addition to the network-related features, the video characteristics also play an important role on how the perceived quality is impaired by losses. In the light of this, Chapter 3 discussed several video-related peculiarities that are known to have an impact on the quality, such as the video format, the type of the frames and the hierarchical structure of the video format, as well as the motion intensity classification. It is also reviewed how the human vision system perceives the impairments and the assessment approaches to quantify the video quality.

Under the given conditions, several mechanisms have been proposed to tackle the open issues related to the video transmission. Chapter 4 presented the most relevant ones in the literature. The main characteristics, advantages, and weakness of each one were discussed. In addition, this chapter also described the off-the-shelf techniques to improve the video transmission quality and why they are not enough to ensure best possible perceived quality to end-users. Owing to the open issues in the literature review the main contributions of this thesis were proposed.

Chapter 5 introduced three proposed mechanisms to shield video transmissions. These mechanisms adopted different approaches, such as how to classify the motion intensity and how to compute the most suited amount of redundancy, to enable high-quality video delivery over wireless networks. Several techniques are used to this end, e.g., a heuristic method, random neural networks, and ant colony optimisation. The results were assessed using quality of experience metrics, as well as the network footprint impact.

Chapter 6 corresponds to the advancements made on improving video transmissions using unmanned aerial vehicles. The proposed mechanisms take into consideration several video characteristics, such as the frame type, the motion vectors, the spatial video complexity, and the temporal video intensity, as well as been capable of handling videos with arbitrary size. In addition, fuzzy logic methods are used to both classify the video details and to set the redundancy amount. In the same way as before, the experiments were assessed using quality of experience metrics.

Chapter 7 specified the achievements on enhancing video delivery over vehicular ad-hoc networks. The mechanisms proposed in this chapter contemplated both the video characteristics and the intrinsic details of this type of network, including even the prediction of future packet loss rates. This was only possible by the combined use of several methods, such as fuzzy logic and hierarchical fuzzy logic, as well as convex hull and approximated convex hull algorithms.

# 8.2 Contributions

The main goals of this thesis were delineated in Chapter 1 and are briefly revisited here. One objective was to define a method to provide video content characterization. This should be done according to the motion intensity and details of each video sequence. Another objective was to propose a series of adaptive FEC-based content- and video-aware mechanisms. The objective of these mechanisms is to shield video transmission over wireless networks providing both high QoE and low network overhead. The third general goal was to study the performance of the proposed mechanisms in different network environments, such as WMNs, FANETs, and VANETs.

These goals have tailored the work presented in this thesis and have led to the following contributions:

**Contribution 1, Assessing the impact of the video characteristics on the video quality level**
> This is a transversal contribution, which is present in all mechanisms described in Chapters 5, 6, and 7. The proposed method uses several video characteristics to assess the impact of them in the perceived video quality. The video details used were the codec type, the frame type and size, the length and format of the group of pictures as well as the motion vectors. According to this information, the proposed mechanisms were able to correctly identify the most important parts of the video sequences.

**Contribution 2, A method to characterise the motion intensity of videos**
> This is the second transversal contribution in this thesis that is also used by all mechanisms presented in Chapters 5, 6, and 7. In the proposed method, the video content is put through a series of procedures to characterise the intensity pace. This is important because different intensities have distinct impacts on the perceived quality if some information is lost. The proposed mechanisms use this knowledge to better protect the most QoE-sensitive data, which allows providing high perceived quality without unnecessary network overhead.

**Contribution 3, Heuristic mechanism**
> This contribution was achieved in Chapter 5, Section 5.2. The main goal of the proposed mechanism was to provide a practical solution with satisfactory results. In order to do that, the mechanism used human experience and knowledge to delineate several strategies to shield the video transmission against losses. The results show that it was possible to maintain or even increase the video quality without adding unnecessary network redundancy.

**Contribution 4, Random neural networks mechanism**
> Chapter 5, more specifically Section 5.3, describe how this contribution was achieved. In the proposed mechanism, random neural networks were used to both categorise the video content according to its motion intensity and to set, in real-time, an adaptive amount of redundancy. The mechanism

has a training and validation phase, which is offline and after that, it is able to perform at run-time. The QoE assessment shows that it outperform others mechanisms, generating also, less network overhead.

**Contribution 5, Ant colony optimisation mechanism**

Section 5.4 describes the mechanism built to support this contribution. It uses an ant colony metaheuristic to assess and classify several video characteristics as well as network conditions, which allows choosing a precise amount of redundancy. In doing this, the perceived video quality is kept at the maximum level possible while the network footprint stays low.

**Contribution 6, Fuzzy logic mechanism**

This contribution is spread in Chapters 6 and 7. Several of the proposed mechanisms use Fuzzy logic as well as Hierarchical fuzzy logic. These techniques were broadly adopted because of their flexibility, allowing the mechanisms to better adapt to a variety of situations. They are also easy to reconfigure and appropriate to manage abstract concepts. Additionally, it allows building complex systems, which are able to handle a large number of inputs at low computation costs. The proposed mechanism's results show that they outperform other competitors, producing less network overhead while increasing the QoE for end-users.

The following subsection outlines further research directions in the fields addressed in this thesis.

## 8.3 Future Work

The relevance and the applicability of mechanisms to improve the quality of video transmissions have been discussed in this thesis, highlighting the need for reliable solutions. The methods and procedures contemplated in this work have provided motivating results, as well as showing the ability to outperform existing solutions in a varied of scenarios. However, there are still several upgrades that could be addressed in the future including, but not limited to, concurrent multipath transmission, opportunistic routing, network coding, and path interleaving.

A distinct aspect that could be addressed is the impact of the proposed mechanisms in an environment where they need to compete for resources against other applications.

The adoption of a smart retransmission scheme could also be considered. Such technique could use a number of information in the decision-making processes, assessing whether it is worth to retransmit a specific packet. The information can range from video characteristics, such as frame type, motion activity and play-out time, and network information, such as delay and link quality, to FEC-related details, for example, what is the minimum number of packets needed to reconstruct the original data. In doing that, it would enable the retransmission

of only what is really needed from the user-point-of-view, respecting the strict constraints commonly found in these services.

Additionally, owing to the ever growing number of wireless-capable devices and the broader amount of communication technologies that they can use, future mechanisms will have to accommodate these heterogeneous networks and devices allowing them to coexist while providing satisfactory levels of user satisfaction. In multihoming environments, where a single device with several interfaces can be connected to multiple communication networks, the necessity of mechanisms to assess the multiple link qualities and adjust the transmission parameters is even more evident. In this context, wireless mobile telecommunications technologies such as Long-Term Evolution Advanced (LTE-Advanced) and 5th Generation Networks (5G) can be used together with Wireless Local Area Network (WLAN). The mechanisms should consider each communication channel's characteristic to adapt the video transmissions. Furthermore, load-balancing schemes can also be considered to further improve the reliability of the transmissions.

Along with the wireless communication, the underlying network infrastructure is also changing. The adoption of Software-defined network (SDN) and Cloud based systems provides a way to build dynamically adaptable and cost-effective networks by decoupling the network control from the forwarding functions. However, one of the main challenges is how to autonomic configure these networks and protect the transmissions in order to provide the best possible service while not using unnecessary network resources. In order to achieve that, several link quality indicators, such as the bit-error rate and signal-to-noise and distortion ratio, as well as mobility patterns, signal strength, and the importance of the data that is being transmitted have to be taken into consideration to programmatically configure the network control layer. This allows maximising the resources usage in a quick and dynamic fashion leading to a better transmission quality.

Finally, sustainable computing with energy-aware devices has gained importance in recent years. The improvement of the overall energy-efficient video transmission over wireless networks can be accomplished, for example, by minimising the amount of redundant information, as proposed in this thesis and thus reducing the network activity. However, there are several other ways to improve on this issue. In a multihoming environment, for instance, it is possible to take into consideration the specific wireless technology's energy consumption when choosing the more appropriated communication method. Furthermore, it is also possible to incorporate the end-user preferences and behaviours to adapt the video transmission mechanisms in order to find the most desirable energy/quality ratio.

# Bibliography

[Abboud et al., 2011] Abboud, O., Pussep, K., Kovacevic, A., Mohr, K., Kaune, S., and Steinmetz, R. (2011). Enabling resilient p2p video streaming: survey and analysis. *Multimedia Systems*, 17:177–197.

[Abraham, 2005] Abraham, A. (2005). Artificial neural networks. *handbook of measuring system design.*

[Adlakha et al., 2007] Adlakha, S., Zhu, X., Girod, B., and Goldsmith, A. (2007). Joint capacity, flow and rate allocation for multiuser video streaming over wireless ad-hoc networks. In *Communications, 2007. ICC'07. IEEE International Conference on*, pages 1747–1753. IEEE.

[Adobe Digital Index, 2016] Adobe Digital Index (Q1 2016). U.S. digital video benchmark report. Technical report, Adobe.

[Adobe Digital Index, 2014] Adobe Digital Index (Q2 2014). U.S. digital video benchmark report. Technical report, Adobe.

[Aguiar et al., 2012a] Aguiar, E., Riker, A., Abelém, A., Cerqueira, E., and Mu, M. (2012a). Video quality estimator for wireless mesh networks. In *Quality of Service (IWQoS), 2012 IEEE 20th International Workshop on*, pages 1–9. IEEE.

[Aguiar et al., 2012b] Aguiar, E., Riker, A., Cerqueira, E., Jorge Gomes Abelem, A., Mu, M., and Zeadally, S. (2012b). Real-time qoe prediction for multimedia applications in wireless mesh networks. In *IEEE 4th Future Multimedia Networking (IEEE FMN 2012).*

[Aiache et al., 2008] Aiache, H., Conan, V., Lebrun, L., and Rousseau, S. (2008). A load dependent metric for balancing internet traffic in wireless mesh networks. In *Mobile Ad Hoc and Sensor Systems, 2008. MASS 2008. 5th IEEE International Conference on*, pages 629 –634.

[Akyildiz and Wang, 2005] Akyildiz, I. and Wang, X. (2005). A survey on wireless mesh networks. *Communications Magazine, IEEE*, 43(9):S23–S30.

[Akyildiz and Wang, 2008] Akyildiz, I. and Wang, X. (2008). Cross-layer design in wireless mesh networks. *Vehicular Technology, IEEE Transactions on*, 57(2):1061–1076.

[Al-Sultan et al., 2014] Al-Sultan, S., Al-Doori, M. M., Al-Bayatti, A. H., and Zedan, H. (2014). A comprehensive survey on vehicular ad hoc network. *Journal of network and computer applications*, 37:380–392.

# Bibliography

[Alay et al., 2010] Alay, O., Korakis, T., Wang, Y., and Panwar, S. S. (2010). Dynamic rate and fec adaptation for video multicast in multi-rate wireless networks. *Mobile Networks and Applications*, 15:425–434.

[Alshabtat et al., 2010] Alshabtat, A. I., Dong, L., Li, J., and Yang, F. (2010). Low latency routing algorithm for unmanned aerial vehicles ad-hoc networks. *International Journal of Electrical and Computer Engineering*, 6(1):48–54.

[Ancillotti et al., 2011] Ancillotti, E., Bruno, R., Conti, M., and Pinizzotto, A. (2011). Load-aware routing in mesh networks: Models, algorithms and experimentation. *Computer Communications*, 34(8):948 – 961.

[Andreopoulos et al., 2006] Andreopoulos, Y., Mastronarde, N., and van der Schaar, M. (2006). Cross-layer optimized video streaming over wireless multihop mesh networks. *Selected Areas in Communications, IEEE Journal on*, 24(11):2104–2115.

[Asefi et al., 2012] Asefi, M., Mark, J. W., and Shen, X. (2012). A mobility-aware and quality-driven retransmission limit adaptation scheme for video streaming over vanets. *Wireless Communications, IEEE Transactions on*, 11(5):1817–1827.

[Attiratanasunthron and Fakcharoenphol, 2008] Attiratanasunthron, N. and Fakcharoenphol, J. (2008). A running time analysis of an ant colony optimization algorithm for shortest paths in directed acyclic graphs. *Information Processing Letters*, 105(3):88 – 92.

[Bahr, 2007] Bahr, M. (2007). Update on the hybrid wireless mesh protocol of ieee 802.11s. In *IEEE Internatonal Conference on Mobile Adhoc and Sensor Systems*, pages 1–6. IEEE.

[Barber et al., 1996] Barber, C. B., Dobkin, D. P., and Huhdanpaa, H. (1996). The quickhull algorithm for convex hulls. *ACM Trans. Math. Softw.*, 22(4):469–483.

[Behrisch et al., 2011] Behrisch, M., Bieker, L., Erdmann, J., and Krajzewicz, D. (2011). Sumo–simulation of urban mobility. In *The Third International Conference on Advances in System Simulation (SIMUL 2011), Barcelona, Spain*.

[Bekmezci et al., 2013] Bekmezci, İ., Sahingoz, O. K., and Temel, Ş. (2013). Flying ad-hoc networks (fanets): A survey. *Ad Hoc Networks*, 11(3):1254 – 1270.

[Bellalta et al., 2014] Bellalta, B., Belyaev, E., Jonsson, M., and Vinel, A. (2014). Performance evaluation of ieee 802.11p-enabled vehicular video surveillance system. *Communications Letters, IEEE*, 18(4):708–711.

[Benedetto et al., 1998] Benedetto, S., Divsalar, D., Montorsi, G., and Pollara, F. (1998). Serial concatenation of interleaved codes: Performance analysis, design, and iterative decoding. *IEEE Transactions on information Theory*, 44(3):909–926.

[Bentley et al., 1982] Bentley, J. L., Preparata, F. P., and Faust, M. G. (1982). Approximation algorithms for convex hulls. *Commun. ACM*, 25(1):64–68.

[Bernard et al., 2011] Bernard, M., Kondak, K., Maza, I., and Ollero, A. (2011). Autonomous transportation and deployment with aerial robots for search and rescue missions. *Journal of Field Robotics*, 28(6):914–931.

[Berrou and Glavieux, 1996] Berrou, C. and Glavieux, A. (1996). Near optimum error correcting coding and decoding: Turbo-codes. *IEEE Transactions on communications*, 44(10):1261–1271.

[Borges et al., 2014] Borges, P., Immich, R., Cerqueira, E., and Curado, M. (2014). Mechanisms for resilient video transmission in wireless networks. adaptive fec mechanism with random neural network classification and ant colony optimization. In *18 Seminario Rede Tematica de Comunicacoes Moveis (RTCM)*.

[Borges et al., 2009] Borges, V. C., Pereira, D., Curado, M., and Monteiro, E. (2009). Routing metric for interference and channel diversity in multi-radio wireless mesh networks. In *Proceedings of the 8th International Conference on Ad-Hoc, Mobile and Wireless Networks*, ADHOC-NOW '09, pages 55–68, Berlin, Heidelberg. Springer-Verlag.

[Bouachir et al., 2013] Bouachir, O., Garcia, F., Larrieu, N., and Gayraud, T. (2013). Ad hoc network qos architecture for cooperative unmanned aerial vehicles (uavs). In *Wireless Days*, pages 1–4. IEEE.

[Bouhouch and El Fatmi, 2007] Bouhouch, H. and El Fatmi, S. G. (2007). Qos routing in ad hoc networks by integreting activity in the olsr protocol. In *Proceedings of the Second International Conference on Systems and Networks Communications*, ICSNC '07, pages 1–, Washington, DC, USA. IEEE Computer Society.

[Breslau et al., 2000] Breslau, L., Estrin, D., Fall, K., Floyd, S., Heidemann, J., Helmy, A., Huang, P., McCanne, S., Varadhan, K., Xu, Y., et al. (2000). Advances in network simulation. *IEEE computer*, 33(5):59–67.

[Bruno et al., 2005] Bruno, R., Conti, M., and Gregori, E. (2005). Mesh networks: commodity multihop ad hoc networks. *IEEE Communications Magazine*, 43(3):123–131.

[Cai and Yang, 2008] Cai, W.-Y. and Yang, H.-B. (2008). A hierarchical qos framework for wireless multimedia network. In *2008 International Conference on Communications, Circuits and Systems*, pages 732–736, Xiamen.

[Cerqueira et al., 2012] Cerqueira, E., Neto, A., Curado, M., Riker, A., Immich, R., Barros, H., Aguiar, E., and Abelem, A. (2012). A parametric qoe video quality estimator for wireless networks. In *Globecom Workshops*, pages 1282–1286.

[Chen and Wu, 2012] Chen, Z. and Wu, W. (2012). Research on error control of multicast in ad hoc. *Computer Science and Electronics Engineering, International Conference on*, 2:513–516.

[Chikkerur et al., 2011] Chikkerur, S., Sundaram, V., Reisslein, M., and Karam, L. (2011). Objective video quality assessment methods: A classification, review, and performance comparison. *Broadcasting, IEEE Transactions on*, 57(2):165 –182.

[Chilamkurti et al., 2010] Chilamkurti, N., Zeadally, S., Soni, R., and Giambene, G. (2010). Wireless multimedia delivery over 802.11e with cross-layer optimization techniques. *Multimedia Tools Appl.*, 47(1):189–205.

[Cisco, 2016] Cisco (2016). White paper: Cisco VNI forecast and methodology, 2015-2020. Technical report, Cisco.

[Clausen and Jacquet, 2003] Clausen, T. and Jacquet, P. (2003). Optimized Link State Routing Protocol (OLSR). RFC 3626.

[comScore, 2013] comScore (2013). Brazilian online video audience reaches 43 million unique viewers in december 2012. Technical report, comScore inc.

[Day and Zimmermann, 1983] Day, J. and Zimmermann, H. (1983). The osi reference model. *Proceedings of the IEEE*, 71(12):1334–1340.

[De Couto et al., 2005] De Couto, D. S. J., Aguayo, D., Bicket, J., and Morris, R. (2005). A high-throughput path metric for multi-hop wireless routing. *Wirel. Netw.*, 11(4):419–434.

[Diaz et al., 2011] Diaz, C., Cabrera, J., Jaureguizar, F., and Garcia, N. (2011). A video-aware fec-based unequal loss protection scheme for rtp video streaming. In *Consumer Electronics (ICCE), 2011 IEEE International Conference on*, pages 3 –4.

[Dorigo et al., 2006] Dorigo, M., Birattari, M., and Stutzle, T. (2006). Ant colony optimization. *IEEE computational intelligence magazine*, 1(4):28–39.

[Dorigo et al., 1996] Dorigo, M., Maniezzo, V., and Colorni, A. (1996). Ant system: optimization by a colony of cooperating agents. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 26(1):29–41.

[Draves et al., 2004] Draves, R., Padhye, J., and Zill, B. (2004). Routing in multi-radio, multi-hop wireless mesh networks. In *Proceedings of the 10th annual international conference on Mobile computing and networking*, MobiCom '04, pages 114–128, New York, NY, USA. ACM.

[Erman et al., 2008] Erman, A. T., van Hoesel, L., Havinga, P., and Wu, J. (2008). Enabling mobility in heterogeneous wireless sensor networks cooperating with uavs for mission-critical management. *IEEE Wireless Communications*, 15(6):38–46.

[Fitzek and Reisslein, 2012] Fitzek, F. H. and Reisslein, M. (2012). Mpeg-4 and h.263 video traces for network performance evaluation library, http://www-tkn.ee.tu-berlin.de/research/trace/trace.html.

[Fonseca and Vazão, 2013] Fonseca, A. and Vazão, T. (2013). Applicability of position-based routing for vanet in highways and urban environment. *Journal of Network and Computer Applications*, 36(3):961–973.

[Frew and Brown, 2009] Frew, E. and Brown, T. (2009). Networking issues for small unmanned aircraft systems. *Journal of Intelligent and Robotic Systems*, 54(1-3):21–37.

[Gallager, 1962] Gallager, R. (1962). Low-density parity-check codes. *Information Theory, IRE Transactions on*, 8(1):21 –28.

[Genetzakis and Siris, 2008] Genetzakis, M. and Siris, V. (2008). A contention-aware routing metric for multi-rate multi-radio mesh networks. In *Sensor, Mesh and Ad Hoc Communications and Networks, 2008. SECON '08. 5th Annual IEEE Communications Society Conference on*, pages 242 –250.

[Gerla et al., 2014] Gerla, M., Wu, C., Pau, G., and Zhu, X. (2014). Content distribution in VANETs. *Vehicular Communications*, 1(1):3 – 12.

[Gil et al., 2007] Gil, P., Maza, I., Ollero, A., and Marrón, P. (2007). Data centric middleware for the integration of wireless sensor networks and mobile robots. In *proc. 7th Conference on Mobile Robots and Competitions, ROBOTICA*.

[Greengrass et al., 2009a] Greengrass, J., Evans, J., and Begen, A. C. (2009a). Not all packets are equal, part 2: The impact of network packet loss on video quality. *IEEE Internet Computing*, 13:74–82.

[Greengrass et al., 2009b] Greengrass, J., Evans, J., and Begen, A. C. (2009b). Not all packets are equal, part i: Streaming video coding and sla requirements. *IEEE Internet Computing*, 13:70–75.

[Gutjahr, 2000] Gutjahr, W. J. (2000). A graph-based ant system and its convergence. *Future Generation Computer Systems*, 16(8):873 – 888.

[Haklay and Weber, 2008] Haklay, M. and Weber, P. (2008). Openstreetmap: User-generated street maps. *IEEE Pervasive Computing*, 7(4):12–18.

[Hamming, 1950] Hamming, R. W. (1950). Error detecting and error correcting codes. *Bell System technical journal*, 29(2):147–160.

[Han et al., 2010] Han, L., Park, S., Kang, S.-S., and In, H. P. (2010). An adaptive fec mechanism using cross-layer approach to enhance quality of video transmission over 802.11 wlans. *TIIS*, pages 341–357.

[Haratcherev et al., 2006] Haratcherev, L., Taal, J., Langendoen, K., Lagendijk, R., and Sips, H. (2006). Optimized video streaming over 802.11 by cross-layer signaling. *Communications Magazine, IEEE*, 44(1):115 – 121.

[Hassan and Landolsi, 2010] Hassan, M. and Landolsi, T. (2010). A retransmission-based scheme for video streaming over wireless channels. *Wirel. Commun. Mob. Comput.*, 10:511–521.

[Huynh-Thu and Ghanbari, 2008a] Huynh-Thu, Q. and Ghanbari, M. (2008a). Scope of validity of psnr in image/video quality assessment. *Electronics Letters*, 44(13):800 –801.

[Huynh-Thu and Ghanbari, 2008b] Huynh-Thu, Q. and Ghanbari, M. (2008b). Scope of validity of PSNR in image/video quality assessment. *Electronics Letters*, 44:800–801(1).

[Immich et al., 2014a] Immich, R., Borges, P., Cerqueira, E., and Curado, M. (2014a). Adaptive motion-aware fec-based mechanism to ensure video transmission. In *Computers and Communication (ISCC), 2014 IEEE Symposium on*, pages 1–6.

[Immich et al., 2014b] Immich, R., Borges, P., Cerqueira, E., and Curado, M. (2014b). Antmind: Enhancing error protection for video streaming in wireless networks. In *Smart Communications in Network Technologies (SaCoNeT), 2014 International Conference on*, pages 1–6.

[Immich et al., 2015a] Immich, R., Borges, P., Cerqueira, E., and Curado, M. (2015a). Qoe-driven video delivery improvement using packet loss prediction. *Int. J. Parallel Emerg. Distrib. Syst.*

[Immich et al., 2012] Immich, R., Cerqueira, E., and Curado, M. (2012). Cross-layer fec-based mechanism to ensure quality of experience in video transmission. In *15 Seminario Rede Tematica de Comunicacoes Moveis (RTCM)*.

[Immich et al., 2013a] Immich, R., Cerqueira, E., and Curado, M. (2013a). Adaptive video-aware fec-based mechanism with unequal error protection scheme. In *Proceedings of the 28th Annual ACM Symposium on Applied Computing*, pages 981–988. ACM.

[Immich et al., 2013b] Immich, R., Cerqueira, E., and Curado, M. (2013b). Cross-layer fec-based mechanism for packet loss resilient video transmission. In Biersack, E., Callegari, C., and Matijasevic, M., editors, *Data Traffic Monitoring and Analysis*, volume 7754 of *Lecture Notes in Computer Science*, pages 320–336. Springer–Verlag, Berlin, Heidelberg.

[Immich et al., 2014c] Immich, R., Cerqueira, E., and Curado, M. (2014c). Ensuring qoe in wireless networks with adaptive fec and fuzzy logic-based mechanisms. In *Communications (ICC), 2014 IEEE International Conference on*, pages 1687–1692.

[Immich et al., 2014d] Immich, R., Cerqueira, E., and Curado, M. (2014d). Improving video qoe in unmanned aerial vehicles using an adaptive fec mechanism. In *Wireless Networking for Moving Objects*, Lecture Notes in Computer Science. Springer Berlin Heidelberg.

[Immich et al., 2014e] Immich, R., Cerqueira, E., and Curado, M. (2014e). Towards the enhancement of uav video transmission with motion intensity awareness. In *Wireless Days (WD), 2014 IFIP*.

[Immich et al., 2015b] Immich, R., Cerqueira, E., and Curado, M. (2015b). Adaptive qoe-driven video transmission over vehicular ad-hoc networks. In *Computer Communications Workshops (INFOCOM WKSHPS), 2015 IEEE Conference on*.

[Immich et al., 2015c] Immich, R., Cerqueira, E., and Curado, M. (2015c). Shielding video streaming against packet losses over vanets. *Wireless Networks*, pages 1–15.

[Immich et al., 2016] Immich, R., Cerqueira, E., and Curado, M. (2016). Towards a qoe-driven mechanism for improved h.265 video delivery. In *2016 Mediterranean Ad Hoc Networking Workshop (Med-Hoc-Net)*, pages 1–8.

[ITU-R BT.500-13, 2012] ITU-R BT.500-13 (2012). Methodology for the subjective assessment of the quality of television pictures. Recommendation BT.500-13.

[ITU-R P.910, 1999] ITU-R P.910 (1999). Subjective video quality assessment methods for multimedia applications. Recommendation P.910.

[ITU-T H. 264/AVC, 2005] ITU-T H. 264/AVC (2005). H.264: Advanced video coding for generic audiovisual services. *ITU-T Rec. H. 264-ISO/IEC 14496-10 AVC.*

[ITU-T J.247, 2008] ITU-T J.247 (2008). Objective perceptual multimedia video quality measurement in the presence of a full reference. Recommendation J.247.

[Jiang and Delgrossi, 2008] Jiang, D. and Delgrossi, L. (2008). Ieee 802.11p: Towards an international standard for wireless access in vehicular environments. In *Vehicular Technology Conference, 2008. VTC Spring 2008. IEEE*, pages 2036–2040.

[Jiang et al., 2012] Jiang, T., Wang, H., and Vasilakos, A. V. (2012). Qoe-driven channel allocation schemes for multimedia transmission of priority-based secondary users over cognitive radio networks. *Selected Areas in Communications, IEEE Journal on*, 30(7):1215–1224.

[Johnson and Maltz, 1996] Johnson, D. B. and Maltz, D. A. (1996). Dynamic source routing in ad hoc wireless networks. In *Mobile computing*, pages 153–181. Springer.

[Kaiwartya and Kumar, 2014] Kaiwartya, O. and Kumar, S. (2014). Geocast routing: Recent advances and future challenges in vehicular adhoc networks. In *Signal Processing and Integrated Networks (SPIN), 2014 International Conference on*, pages 291–296. IEEE.

[Kaiwartya and Kumar, 2015] Kaiwartya, O. and Kumar, S. (2015). Guaranteed geocast routing protocol for vehicular adhoc networks in highway traffic environment. *Wireless Personal Communications*, 83(4):2657–2682.

[Karner et al., 2007] Karner, W., Nemethova, O., and Rupp, M. (2007). Link error prediction in wireless communication systems with quality based power control. In *Communications, 2007. ICC '07. IEEE International Conference on*, pages 5076–5081.

[Karp and Kung, 2000] Karp, B. and Kung, H.-T. (2000). Gpsr: Greedy perimeter stateless routing for wireless networks. In *Proceedings of the 6th annual*

*international conference on Mobile computing and networking*, pages 243–254. ACM.

[Katsaros et al., 2011] Katsaros, K., Dianati, M., Tafazolli, R., and Kernchen, R. (2011). Clwpr - a novel cross-layer optimized position based routing protocol for vanets. In *Vehicular Networking Conference (VNC), 2011 IEEE*, pages 139–146.

[Khan et al., 2009] Khan, A., Sun, L., and Ifeachor, E. (2009). Impact of video content on video quality for video over wireless networks. In *Proceedings of the 2009 Fifth International Conference on Autonomic and Autonomous Systems*, pages 277–282. IEEE Computer Society.

[Khan et al., 2010] Khan, A., Sun, L., Jammeh, E., and Ifeachor, E. (2010). Quality of experience-driven adaptation scheme for video applications over wireless networks. *IET Commun.*, 4(11):1337–1337.

[Kim and Shin, 2006] Kim, K.-H. and Shin, K. G. (2006). On accurate measurement of link quality in multi-hop wireless mesh networks. In *Proceedings of the 12th annual international conference on Mobile computing and networking*, MobiCom '06, pages 38–49, New York, NY, USA. ACM.

[Klaue et al., 2003] Klaue, J., Rathke, B., and Wolisz, A. (2003). Evalvid–a framework for video transmission and quality evaluation. *International Conference on Modelling Techniques and Tools for Computer Performance Evaluation*, pages 255–272.

[Kumar et al., 2001] Kumar, R., Sawhney, H., Samarasekera, S., Hsu, S., Tao, H., Guo, Y., Hanna, K., Pope, A., Wildes, R., Hirvonen, D., et al. (2001). Aerial video surveillance and exploitation. *Proceedings of the IEEE*, 89(10):1518–1539.

[Lacan, 2009] Lacan, J. (2009). Reed-solomon forward error correction (fec) schemes.

[Le Gall, 1991] Le Gall, D. (1991). Mpeg: A video compression standard for multimedia applications. *Communications of the ACM*, 34(4):46–58.

[Lecuire, 2012] Lecuire, V. (2012). Unequal error protection under bitrate constraint for video streaming over internet. *Computer Communications*, 35(3):287 – 297.

[Lee et al., 2011] Lee, J.-W., Chen, C.-L., Horng, M.-F., and Kuo, Y.-H. (2011). An efficient adaptive fec algorithm for short-term quality control in wireless networks. In *Advanced Communication Technology (ICACT)*, pages 1124 – 1129.

[Li et al., 2007] Li, J., Cui, Y., and Chang, B. (2007). Peerstreaming: design and implementation of an on-demand distributed streaming system with digital rights management capabilities. *Multimedia Systems*, 13:173–190. 10.1007/s00530-006-0073-6.

[Lin et al., 2009] Lin, C.-H., Shieh, C.-K., Ke, C.-H., Chilamkurti, N., and

Zeadally, S. (2009). An adaptive cross-layer mapping algorithm for mpeg-4 video transmission over ieee 802.11e wlan. *Telecommunication Systems*, 42:223–234. 10.1007/s11235-009-9182-9.

[Lin et al., 2012a] Lin, C.-H., Wang, Y.-C., Shieh, C.-K., and Hwang, W.-S. (2012a). An unequal error protection mechanism for video streaming over ieee 802.11e wlans. *Computer Networks*, 56(11):2590 – 2599.

[Lin et al., 2012b] Lin, L., Sun, Q., Li, J., and Yang, F. (2012b). A novel geographic position mobility oriented routing strategy for uavs. *Journal of Computational Information Systems*, 8(2):709–716.

[Lindeberg et al., 2011] Lindeberg, M., Kristiansen, S., Plagemann, T., and Goebel, V. (2011). Challenges and techniques for video streaming over mobile ad hoc networks. *Multimedia Systems*, 17:51–82.

[Liu et al., 2008] Liu, C., Gkelias, A., Hou, Y., and Leung, K. (2008). Cross-layer design for qos in wireless mesh networks. *Wireless Personal Communications*, 96:64–76.

[Liu et al., 2004] Liu, G., Lee, B.-S., Seet, B.-C., Foh, C.-H., Wong, K.-J., and Lee, K.-K. (2004). A routing strategy for metropolis vehicular communications. In *International Conference on Information Networking*, pages 134–143. Springer.

[Liu and Liao, 2009] Liu, T. and Liao, W. (2009). Interference-aware qos routing for multi-rate multi-radio multi-channel ieee 802.11 wireless mesh networks. *Wireless Communications, IEEE Transactions on*, 8(1):166 –175.

[Lochert et al., 2003] Lochert, C., Hartenstein, H., Tian, J., Fussler, H., Hermann, D., and Mauve, M. (2003). A routing strategy for vehicular ad hoc networks in city environments. In *Intelligent Vehicles Symposium, 2003. Proceedings. IEEE*, pages 156–161. IEEE.

[Lotfallah et al., 2006] Lotfallah, O., Reisslein, M., and Panchanathan, S. (2006). Adaptive video transmission schemes using mpeg-7 motion intensity descriptor. *Circuits and Systems for Video Technology, IEEE Transactions on*, 16(8):929 –946.

[Luby, 2002] Luby, M. (2002). Lt codes. In *Foundations of Computer Science, 2002. Proceedings. The 43rd Annual IEEE Symposium on*, pages 271 – 280.

[Luo and Ephremides, 2012] Luo, J. and Ephremides, A. (2012). A new approach to random access: Reliable communication and reliable collision detection. *Information Theory, IEEE Transactions on*, 58(2):989–1002.

[Ma et al., 2011] Ma, K. J., Bartoš, R., and Bhatia, S. (2011). Review: A survey of schemes for internet-based video delivery. *J. Netw. Comput. Appl.*, 34(5):1572–1586.

[Ma and Denko, 2007] Ma, L. and Denko, M. K. (2007). A routing metric for load-balancing in wireless mesh networks. In *Proceedings of the 21st International Conference on Advanced Information Networking and Applications*

*Workshops - Volume 02*, AINAW '07, pages 409–414, Washington, DC, USA. IEEE Computer Society.

[Ma et al., 2016] Ma, X., Zeng, W. K., Yang, L. T., Zou, D., and Jin, H. (2016). Lossless roi privacy protection of h.264/avc compressed surveillance videos. *IEEE Transactions on Emerging Topics in Computing*, 4(3):349–362.

[Martini et al., 2007] Martini, M., Mazzotti, M., Lamy-Bergot, C., Huusko, J., and Amon, P. (2007). Content adaptive network aware joint optimization of wireless video transmission. *Communications Magazine, IEEE*, 45(1):84–90.

[Marwaha et al., 2004] Marwaha, S., Srinivasan, D., Tham, C. K., and Vasilakos, A. (2004). Evolutionary fuzzy multi-objective routing for wireless mobile ad hoc networks. In *Evolutionary Computation, 2004. CEC2004. Congress on*, volume 2, pages 1964–1971. IEEE.

[Meng et al., 2016] Meng, T., Wu, F., Yang, Z., Chen, G., and Vasilakos, A. V. (2016). Spatial reusability-aware routing in multi-hop wireless networks. *IEEE Transactions on Computers*, 65(1):244–255.

[Mittag et al., 2011] Mittag, J., Papanastasiou, S., Hartenstein, H., and Strom, E. (2011). Enabling accurate cross-layer PHY/MAC/NET simulation studies of vehicular communication networks. *Proceedings of the IEEE*, 99(7):1311–1326.

[Mohamed and Rubino, 2002] Mohamed, S. and Rubino, G. (2002). A study of real-time packet video quality using random neural networks. 12(12):1071–1083.

[Naeimipoor and Boukerche, 2014] Naeimipoor, F. and Boukerche, A. (2014). A hybrid video dissemination protocol for vanets. In *Communications (ICC), 2014 IEEE International Conference on*, pages 112–117.

[Nafaa et al., 2008] Nafaa, A., Taleb, T., and Murphy, L. (2008). Forward error correction strategies for media streaming over wireless networks. *IEEE Communications Magazine*, 46(1):72–79.

[Neckebroek et al., 2010] Neckebroek, J., Moeneclaey, M., and Magli, E. (2010). Comparison of reed-solomon and raptor codes for the protection of video on-demand on the erasure channels. In *Information Theory and its Applications, International conference, Proceedings*, pages 856–860. IEEE.

[Oh and Chen, 2010] Oh, B. J. and Chen, C. W. (2010). A cross-layer adaptation hcca mac for qos-aware h.264 video communications over wireless mesh networks. In *IEEE International Symposium on Circuits and Systems (ISCAS)*, pages 2259 –2262.

[Pal and Nasipuri, 2011] Pal, A. and Nasipuri, A. (2011). A quality based routing protocol for wireless mesh networks. *Pervasive and Mobile Computing*, 7(5):611 – 626. <ce:title>QoS and Service Provisioning for Integrated Wireless Networks</ce:title>.

[Parissidis et al., 2011] Parissidis, G., Karaliopoulos, M., Spyropoulos, T., and

Plattner, B. (2011). Interference-aware routing in wireless multihop networks. *IEEE Transactions on Mobile Computing*, 10(5):716–733.

[Pearl, 1984] Pearl, J. (1984). Heuristics: intelligent search strategies for computer problem solving.

[Pedrycz, 1994] Pedrycz, W. (1994). Why triangular membership functions? *Fuzzy Sets and Systems*, 64(1):21 – 30.

[Perkins et al., 2003] Perkins, C., Belding-Royer, E., and Das, S. (2003). Ad hoc on-demand distance vector (aodv) routing. Technical report.

[Pham et al., 2014] Pham, T. A. Q., Piamrat, K., and Viho, C. (2014). Qoe-aware routing for video streaming over vanets. In *Vehicular Technology Conference (VTC Fall), 2014 IEEE 80th*, pages 1–5.

[Piamrat et al., 2009] Piamrat, K., Viho, C., Bonnin, J.-M., and Ksentini, A. (2009). Quality of experience measurements for video streaming over wireless networks. In *Information Technology: New Generations, Third International Conference on*, pages 1184–1189.

[Pinson and Wolf, 2004] Pinson, M. H. and Wolf, S. (2004). A new standardized method for objectively measuring video quality. *IEEE Transactions on broadcasting*, 50(3):312–322.

[Puri, 2005] Puri, A. (2005). A survey of unmanned aerial vehicles (uav) for traffic surveillance. *Department of computer science and engineering, University of South Florida*.

[Qiu et al., 2009] Qiu, X., Liu, H., Ghosal, D., Mukherjee, B., Benko, J., Li, W., and Bajaj, R. (2009). Adaptive video compression rate optimization in wireless access networks. In *Local Computer Networks, 2009. LCN 2009. IEEE 34th Conference on*, pages 420 –427.

[Qiu et al., 2011] Qiu, X., Liu, H., Ghosal, D., Mukherjee, B., Benko, J., Li, W., and Bajaj, R. (2011). Enhancing the performance of video streaming in wireless mesh networks. *Wireless Personal Communications*, 56:535–557. 10.1007/s11277-010-9988-5.

[Quadros et al., 2013] Quadros, C., Cerqueira, E., Neto, A., Pescape, A., Riker, A., Immich, R., and Curado, M. (2013). A quality of experience handover system for heterogeneous multimedia wireless networks. In *Computing, Networking and Communications (ICNC), 2013 International Conference on*, pages 1064–1068.

[Quadros et al., 2012] Quadros, C., Cerqueira, E., Neto, A., Riker, A., Immich, R., and Curado, M. (2012). A mobile qoe architecture for heterogeneous multimedia wireless networks. In *Globecom Workshops*, pages 1057–1061.

[Razavi et al., 2009] Razavi, R., Fleury, M., and Ghanbari, M. (2009). Adaptive packet-level interleaved fec for wireless priority-encoded video streaming. *Adv. MultiMedia*, 2009:3:1–3:14.

[Reed and Solomon, 1960] Reed, I. S. and Solomon, G. (1960). Polynomial Codes Over Certain Finite Fields. *Journal of the Society for Industrial and Applied Mathematics*, 8(2):300–304.

[Revelle, 1979] Revelle, W. (1979). Hierarchical cluster analysis and the internal structure of tests. *Multivariate Behavioral Research*, 14(1):57–74.

[Rezende et al., 2013] Rezende, C., Almulla, M., and Boukerche, A. (2013). The use of erasure coding for video streaming unicast over vehicular ad hoc networks. In *Local Computer Networks (LCN), 2013 IEEE 38th Conference on*, pages 715–718.

[Richardson, 2004] Richardson, I. E. (2004). *H. 264 and MPEG-4 video compression: video coding for next-generation multimedia.* John Wiley & Sons.

[Rizzo, 1997] Rizzo, L. (1997). Effective erasure codes for reliable computer communication protocols. *SIGCOMM Comput. Commun. Rev.*, 27:24–36.

[Rosario et al., 2013] Rosario, D., Cerqueira, E., Neto, A., Riker, A., Immich, R., and Curado, M. (2013). A qoe handover architecture for converged heterogeneous wireless networks. *Wireless Networks*, 19(8):2005–2020.

[Rubino, 2005] Rubino, G. (2005). Quantifying the quality of audio and video transmissions over the internet: the psqa approach. *Design and operations of communication networks: a review of wired and wireless modeling and management challenges. Imperial College Press, London.*

[Sahingoz, 2014] Sahingoz, O. K. (2014). Networking models in flying ad-hoc networks (fanets): Concepts and challenges. *Journal of Intelligent & Robotic Systems*, 74(1-2):513.

[Sahoo et al., 2012] Sahoo, R. R., Panda, R., Behera, D. K., and Naskar, M. K. (2012). A trust based clustering with ant colony routing in vanet. In *Computing Communication Networking Technologies (ICCCNT), 2012 Third International Conference on*, pages 1–8.

[Salyers et al., 2008] Salyers, D., Striegel, A., and Poellabauer, C. (2008). Wireless reliability: Rethinking 802.11 packet loss. In *World of Wireless, Mobile and Multimedia Networks, 2008. WoWMoM 2008. 2008 International Symposium on a*, pages 1 –4.

[Schulzrinne et al., 2003] Schulzrinne, H., Casner, S., Frederick, R., and Jacobson, V. (2003). Rtp: A transport protocol for real-time applications. Technical report.

[Schulzrinne et al., 1998] Schulzrinne, H., Rao, A., and Lanphier, R. (1998). Real time streaming protocol (rtsp).

[Seeling et al., 2004] Seeling, P., Reisslein, M., and Kulapala, B. (2004). Network performance evaluation using frame size and quality traces of single-layer and two-layer video: A tutorial. *IEEE Communications Surveys & Tutorials*, 6(3):58–78.

[Seet et al., 2004] Seet, B.-C., Liu, G., Lee, B.-S., Foh, C.-H., Wong, K.-J., and Lee, K.-K. (2004). A-star: A mobile ad hoc routing strategy for metropolis vehicular communications. In *International Conference on Research in Networking*, pages 989–999. Springer.

[Serral-Gracià et al., 2010] Serral-Gracià, R., Cerqueira, E., Curado, M., Yannuzzi, M., Monteiro, E., and Masip-Bruin, X. (2010). An overview of quality of experience measurement challenges for video applications in ip networks. In *International Conference on Wired/Wireless Internet Communications*, pages 252–263. Springer.

[Seshadrinathan et al., 2010] Seshadrinathan, K., Soundararajan, R., Bovik, A., and Cormack, L. (2010). Study of subjective and objective quality assessment of video. *IEEE Transactions on Image Processing*, 19(6):1427 –1441.

[Setton et al., 2008] Setton, E., Baccichet, P., and Girod, B. (2008). Peer-to-peer live multicast: A video perspective. *Proceedings of the IEEE*, 96(1):25 –38.

[Setton et al., 2005] Setton, E., Yoo, T., Zhu, X., Goldsmith, A., and Girod, B. (2005). Cross-layer design of ad hoc networks for real-time video streaming. *Wireless Communications, IEEE*, 12(4):59–65.

[Shen et al., 2011] Shen, Z., Luo, J., Zimmermann, R., and Vasilakos, A. V. (2011). Peer-to-peer media streaming: Insights and new developments. *Proceedings of the IEEE*, 99(12):2089–2109.

[Shokrollahi, 2006] Shokrollahi, A. (2006). Raptor codes. *IEEE/ACM Trans. Netw.*, 14(SI):2551–2567.

[Silveira and Silva, 2012] Silveira, F. and Silva, E. d. S. (2012). Predicting packet loss statistics with hidden markov models for {FEC} control. *Computer Networks*, 56(2):628–641.

[Singh and Agrawal, 2014] Singh, S. and Agrawal, S. (2014). Vanet routing protocols: issues and challenges. In *Engineering and Computational Sciences (RAECS), 2014 Recent Advances in*, pages 1–5. IEEE.

[Soldo et al., 2011] Soldo, F., Casetti, C., Chiasserini, C., and Chaparro, P. (2011). Video streaming distribution in vanets. *Parallel and Distributed Systems, IEEE Transactions on*, 22(7):1085–1091.

[Srivastava and Motani, 2005] Srivastava, V. and Motani, M. (2005). Cross-layer design: a survey and the road ahead. *Communications Magazine, IEEE*, 43(12):112–119.

[Staelens et al., 2011] Staelens, N., Sedano, I., Barkowsky, M., Janowski, L., Brunnstrom, K., and Le Callet, P. (2011). Standardized toolchain and model development for video quality assessment - the mission of the joint effort group in vqeg. In *Quality of Multimedia Experience (QoMEX)*, pages 61 –66.

[Sullivan et al., 2012] Sullivan, G. J., Ohm, J.-R., Han, W.-J., and Wiegand, T.

(2012). Overview of the high efficiency video coding (hevc) standard. *Circuits and Systems for Video Technology, IEEE Transactions on*, 22(12):1649–1668.

[Talari et al., 2013] Talari, A., Kumar, S., Rahnavard, N., Paluri, S., and Matyjas, J. (2013). Optimized cross-layer forward error correction coding for h.264 avc video transmission over wireless channels. *EURASIP Journal on Wireless Communications and Networking*, 2013(1):1–13.

[Taliwal et al., 2004] Taliwal, V., Jiang, D., Mangold, H., Chen, C., and Sengupta, R. (2004). Empirical determination of channel characteristics for dsrc vehicle-to-vehicle communication. In *Proceedings of the 1st ACM International Workshop on Vehicular Ad Hoc Networks*, VANET '04, pages 88–88, New York, NY, USA. ACM.

[Tsai et al., 2011a] Tsai, M.-F., Chilamkurti, N. K., Zeadally, S., and Vinel, A. (2011a). Concurrent multipath transmission combining forward error correction and path interleaving for video streaming. *Comput. Commun.*, 34:1125–1136.

[Tsai et al., 2011b] Tsai, M.-F., Ke, C.-H., Liang, H.-M., and Huang, H.-Y. (2011b). Forward error correction with interleaving mechanism combining cognitive technology for video streaming over wireless networks. In *Wireless and Pervasive Computing (ISWPC), 2011 6th International Symposium on*, pages 1 –6.

[Tsai et al., 2009] Tsai, M.-F., Shieh, C.-K., Hwang, W.-S., and Deng, D.-J. (2009). An adaptive multi-hop forward error correction protection scheme for video streaming over wireless mesh networks. *Int. J. Commun. Syst.*, 22(10):1297–1318.

[Tsirigos and Haas, 2004] Tsirigos, A. and Haas, Z. J. (2004). Analysis of multipath routing-part i: the effect on the packet delivery ratio. *Trans. Wireless. Comm.*, 3(1):138–146.

[Vanne et al., 2006] Vanne, J., Aho, E., Hamalainen, T., and Kuusilinna, K. (2006). A high-performance sum of absolute difference implementation for motion estimation. *Circuits and Systems for Video Technology, IEEE Transactions on*, 16(7):876–883.

[Vatolin et al., 2011] Vatolin, D., Moskvin, A., Petrov, O., Putilin, S., Grishin, S., and Marat, A. (2011). Video quality measurement tool.

[Vlavianos et al., 2008] Vlavianos, A., Law, L., Broustis, I., Krishnamurthy, S., and Faloutsos, M. (2008). Assessing link quality in ieee 802.11 wireless networks: Which is the right metric? In *Personal, Indoor and Mobile Radio Communications, 2008. PIMRC 2008. IEEE 19th International Symposium on*, pages 1–6.

[Waharte et al., 2006] Waharte, S., Boutaba, R., Iraqi, Y., and Ishibashi, B. (2006). Routing protocols in wireless mesh networks: challenges and design considerations. *Multimedia Tools Appl.*, 29(3):285–303.

[Wan et al., 2015] Wan, Z., Xiong, N., and Yang, L. (2015). Cross-layer video transmission over ieee 802.11e multihop networks. *Multimedia Tools and Applications*, 74(1):5–23.

[Wang et al., 2010] Wang, Y.-C., Lin, C.-H., Tsai, M.-F., and Shieh, C.-K. (2010). Cross-layer unequal error protection mechanism with an interleaved strategy for video streaming over wireless networks. In *Wireless Communications and Networking Conference (WCNC), 2010 IEEE*, pages 1 –6.

[Wang et al., 2004] Wang, Z., Bovik, A., Sheikh, H., and Simoncelli, E. (2004). Image quality assessment: from error visibility to structural similarity. *Image Processing, IEEE Transactions on*, 13(4):600–612.

[Ward Jr, 1963] Ward Jr, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American statistical association*, 58(301):236–244.

[Wiegand et al., 2003] Wiegand, T., Sullivan, G., Bjontegaard, G., and Luthra, A. (2003). Overview of the h.264/avc video coding standard. *Circuits and Systems for Video Technology, IEEE Transactions on*, 13(7):560 –576.

[Wilhelmsson and Milstein, 1999] Wilhelmsson, L. and Milstein, L. (1999). On the effect of imperfect interleaving for the gilbert-elliott channel. *Communications, IEEE Transactions on*, 47(5):681 –688.

[Winkler and Mohandas, 2008] Winkler, S. and Mohandas, P. (2008). The evolution of video quality measurement: From PSNR to hybrid metrics. *IEEE Transactions on Broadcasting*, 54(3):660–668.

[Wong et al., 2005] Wong, K.-W., Tikk, D., Gedeon, T., and Koczy, L. (2005). Fuzzy rule interpolation for multidimensional input spaces with applications: a case study. *Fuzzy Systems, IEEE Transactions on*, 13(6):809–819.

[Wu and Ma, 2014] Wu, H. and Ma, H. (2014). Opportunistic routing for live video streaming in vehicular ad hoc networks. In *World of Wireless, Mobile and Multimedia Networks (WoWMoM), 2014 IEEE 15th International Symposium on a*, pages 1–3.

[Xiph.Org, 2011] Xiph.Org (2011). Theora format specification, http://www.theora.org/doc/theora.pdf.

[Xiph.Org, 2016] Xiph.Org (2016). Xiph.org video test media [derf's collection], http://media.xiph.org/video/derf/.

[Yang et al., 2006a] Yang, C.-Y., Ting, Y.-W., and Yang, C.-C. (2006a). A loop-free multi-path routing with qos for wireless ad hoc network. In *Proceedings of the 20th International Conference on Advanced Information Networking and Applications - Volume 02*, AINA '06, pages 179–185, Washington, DC, USA. IEEE Computer Society.

[Yang et al., 2006b] Yang, Y., Wang, J., and Kravets, R. (2006b). Load-balanced routing for mesh networks. *SIGMOBILE Mob. Comput. Commun. Rev.*, 10(4):3–5.

[Yu et al., 2005] Yu, X., Modestino, J., and Tian, X. (2005). The accuracy of gilbert models in predicting packet-loss statistics for a single-multiplexer network model. In *INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE*, volume 4, pages 2602–2612 vol. 4.

[Yuan et al., 2006] Yuan, Y., Cockburn, B., Sikora, T., and Mandal, M. (2006). A gop based fec technique for packet based video streaming. In *Proceedings of the 10th WSEAS international conference on Communications*, ICCOM'06, pages 187–192.

[Zang and Zang, 2011] Zang, C. and Zang, S. (2011). Mobility prediction clustering algorithm for uav networking. In *GLOBECOM Workshops (GC Wkshps), 2011 IEEE*, pages 1158–1161. IEEE.

[Zeng et al., 2013] Zeng, Y., Xiang, K., Li, D., and Vasilakos, A. V. (2013). Directional routing and scheduling for green vehicular delay tolerant networks. *Wireless networks*, 19(2):161–173.

[Zhang et al., 2012] Zhang, C., Zhang, J., Wei, G., and Ren, P. (2012). An efficient cooperative arq protocol for wireless relay networks. *Computer Communications*.

[Zhang et al., 2015] Zhang, X. M., Zhang, Y., Yan, F., and Vasilakos, A. V. (2015). Interference-based topology control algorithm for delay-constrained mobile ad hoc networks. *Mobile Computing, IEEE Transactions on*, 14(4):742–754.

[Zhao et al., 2012] Zhao, Z., Braun, T., do Rosario, D., Cerqueira, E., Immich, R., and Curado, M. (2012). Qoe-aware fec mechanism for intrusion detection in multi-tier wireless multimedia sensor networks. In *Wireless and Mobile Computing, Networking and Communications (WiMob), 2012 IEEE 8th International Conference on*, pages 689–696.

[Zhou et al., 2011a] Zhou, L., Chao, H.-C., and Vasilakos, A. V. (2011a). Joint forensics-scheduling strategy for delay-sensitive multimedia applications over heterogeneous networks. *Selected Areas in Communications, IEEE Journal on*, 29(7):1358–1367.

[Zhou et al., 2011b] Zhou, L., Zhang, Y., Song, K., Jing, W., and Vasilakos, A. V. (2011b). Distributed media services in p2p-based vehicular networks. *Vehicular Technology, IEEE Transactions on*, 60(2):692–703.

[Zhu, 2011] Zhu, R. (2011). Intelligent rate control for supporting real-time traffic in wlan mesh networks. *Journal of Network and Computer Applications*, 34(5):1449 – 1458.

[Zwinkels, 2004] Zwinkels, P. (2004). Wireless mesh network. *Journal Of The Communications Network*, 3:47–51.