Transmitting Scalable Video with Unequal Error Protection over 802.11b/g

Robin D'haenens, Jeroen Doggen
University College of Antwerp, Artesis
Paardenmarkt 92, 2000 Antwerp, Belgium
Telephone: 0032 (0)3 213 79 41
Email: j.doggen@ha.be

Dirk Bakker, Tim Dams
Vrije Universiteit Brussel, ETRO
Pleinlaan 2, 1050 Brussel, Belgium
Telephone: 0032 (0)2 629 16 75
Email: tdams@etro.vub.ac.be

***Abstract*—We developed a simulation set-up that can test the behaviour of streaming applications over an error-prone wireless networks: 802.11b/g specifically. The application we tested is a video coder that adds unequal error protection to the scalable extension of H.264/SVC. This protection mechanism enables the recovery of lost packets. We adapted a radio wave propagation model to define Rayleigh fading which is able to model packet reception correctly for indoor environments. We investigated how the video encoder creates RTP packets and we noticed a very small payload, resulting in a high protocol overhead of 54%. We recommended the encoded designers to use UDP lite because it does not discard corrupt packets, as UDP does, but allows them to pass to the video decoder. This ensures more incoming information that can be evaluated at the application layer. The simulation of the 802.11b/g networks with bad wireless network conditions proves the correcting capabilities of the video decoder. The decoder could repair lost packets, or ensure graceful degradation. However this graceful degradation could not be maintained when the packet losses become too high, as for example when moving out of reach of an access point.**

***Index Terms*—802.11b/g, RF propagation, H.264/SVC, mobility, SWiNWiMob.**

I. Introduction

Wireless networks are becoming more and more integrated into everyday life. Mobile devices support wireless access to give the user the ability to access wireless networks. For example surfing the internet on your cell phone, accessing your company network through your laptop over Wi-Fi or downloading your agenda to your smartphone over Bluetooth.

The market is emerging by providing mobile users multimedia and giving them the freedom to choose the content and the moment to consume them. Services like real-time streaming, video-on-demand and VoIP are getting integrated as we write. One of the challenges for researchers in the field of multimedia and telecommunication are to meet the requirements (e.g. delay, quality, etc.) of the end-users by taking into account their limitations like processing power, screen resolution and battery capacity. Another challenge is to make sure the multimedia is transported over this wireless network, with carrying quality, without losing too much data. A possible solution for these problems lies in Unequal Error Protection in Scalable Video Streams [1] (UEP in SVS), which is a protection mode for the scalable extension of H.264/AVC [2].

The objective of this research is to test the robustness of this UEP algorithm in SVC. We do this by simulating this stream over a wireless network, which is more vulnerable to transmission errors than wired network channels.

This paper starts with a brief discussion of H.264/AVC. This is necessary to understand its extension: Scalable Video Coding, which can be found in the next section. We then explore the priority-based error protection for the scalable extension of H.264/AVC. Next, we explore the simulation environment and we discuss how we tried to make the wireless environment as realistic as possible. We tackle some problems such as the implementation of radiowave propagation models and error behaviour of wireless transmissions. Finally we present the results and our conclusion.

II. Unequal Error Protection in SVC

H.264/AVC [2] is a product from two institutes who creates video compression standards: ITU-T Video Coding Experts Group (VCEG), and the Moving Pictures Experts group (MPEG). H.264/AVC improves previous standards by improving network support. This is accomplished by introducing NAL units. These units are an extra logical layer which provides enhanced support for network transmissions by dividing the encoded bitstream into coherent datapackets.

The Scalable Video Coding extension is a technique where a video is encoded only once, and can be decomposed out of multiple layers which affect a videostream in the temporal or spatial domain. This feature gives the receivers the possibility to request only a base layer (with the lowest quality), for low-end devices or a base layer with all the enhancement layers (which add quality) for HDTV's.

Here we provide a brief introduction of this process, which you can see in fig. 1. A more thorough study of priority-based error protection for the scalable extension of H.264/SVC can be found in [1].

SVC bitstream: this is the complete scalable bitstream where all the raw video information is encoded. The bitstream gets divided into NAL units.

Priority assigner: this module calculates the priority that each NAL unit deserves. This priority level is assigned by the following formula.

$$pr = TL(QL_{max} + 1) + QL$$

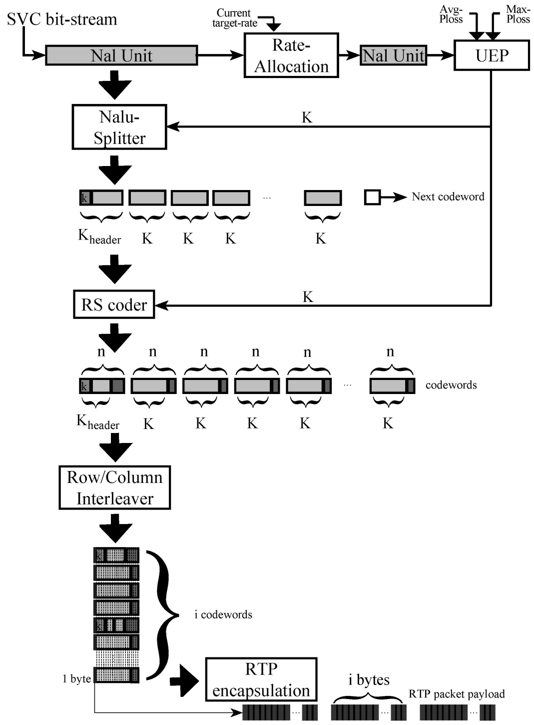

Fig. 1. Encoding Process of Priority-Based Error Protection for the Scalable Extension of H.264/AVC.

Where pr is priority, TL is temporal level, QL quality level of the current NAL unit and QL_{max} the maximal quality level. The current temporal level is the property that enables the scaling in the temporal domain which means scaling in frame-rate. The outcome of this formula indicates the importance of the NAL unit where a low pr means a high priority.

Rate Allocation with UEP uses the calculated priority, network channel conditions (maximum and average loss) and the video encoding bitrate for a Group Of Pictures (GOP) to decide whether the NAL units should be encoded. Network channel conditions are received from RTCP (Real Time Control Protocol[3]) reports. These parameters give important information about the content of the NAL unit. This algorithm ensures the encoding of the base layer while the enhancement layers are only coded if the available encoding rate is not yet achieved.

Reed-Solomon Coder: The NAL unit splitter creates code words from the encoded NAL units. These code words receive an extra protection: Forward Error Correction (FEC), by adding a certain amount of parity bytes. The exact amount of parity bytes depends on the priority of the NAL unit [4].

Row/Column interleaver: the interleaver adds an extra protection mechanism. This interleaver is also called a row-column interleaver because incoming records are imported like rows and read out like columns [4]. This means that a loss of an entire packet results in a loss of 1 byte for several packets, which is easier to correct for the Reed-Solomon coder.

RTP encapsulation: the last step before sending it over the network is to form Real Time Transport Protocol (RTP) packets. RTP contains a timestamp and an ID so the receiver can order the packets before sending them to the decoder. RTP is sent over UDP which offers no retransmission mechanism, as this is unnecessary for streaming applications which are bound to strict timing rules.

At last, the *decoder* inverses the process of the encoder. It converts the received RTP packets back into the video stream. The interleaver detects lost packets, and adds empty packets so the Read-Solomon decoder can recover them.

III. Simulation Model

Simulation is an approach which can be used to predict the behaviour and performance of large, complex stochastic systems. When a model is not very accurate, one can make the wrong conclusions from the simulation results. The basic problem is that every simulation model is inherently wrong, ranging from lightly flawed up to totally wrong [5]. We will dedicate this section to introduce the configuration of our simulation using the open source network simulator ns-2 [6].

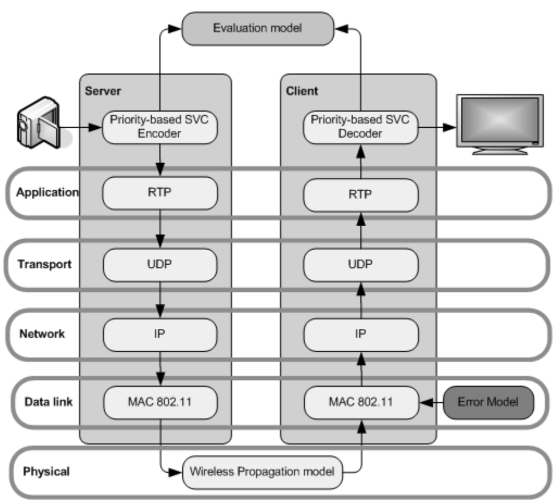

Fig. 2. Schematic overview of the simulation

Fig. 2 shows the architecture of the simulation. We simulate a transmitter (e.g. camera) which sends its raw data to our video encoder. The encoded data gets packetized into RTP packets and then is further encapsulated into UDP, IP and MAC packets, respectively. Next, the packet is sent over the network by a realistic propagation model (see chapter III-A). Unfortunately, this process is not very accurate and wireless links are almost perfect in ns. However, our mission is to test the wireless propagation errors so we introduced an extra error model at the data link layer of the receiver.

Finally, the data follows the protocol stack to the application layer where all the received packets are injected to the decoder.

The results of the decoder are compared with the input file by the evaluation model.

A. Physical Layer - Radiowave Propagation

An important aspect of a reliable simulation of a wireless channel is the modelling of the radiowaves. Wireless channels suffer from reflection, multipath fading, diffraction, scattering, etc. This causes interference which influences the quality of your wireless link. If it is absolute necessary to model *all* of those interferences, you'll need to make a three-dimensional map with properties of the used materials (such as permittivity) [7]. This approach is not desirable for non-site-specific simulations so we require other parameters to implement a propagation model.

One of the main responsibilities of the propagation models is to calculate a nodes received power in a proper way. Ns-2 has 4 models to implement the wireless behaviour: freespace model, two-ray ground reflection model, shadowing model and Nakagami model.

1) Freespace Model: This model uses the formula of Friis [8] which states that the received power is inversely proportional to the distance.

$$P_{rx}(d) = \frac{P_{tx}G_tG_r\lambda^2}{4^2\pi^2d^2L} \text{ [9]}$$

where $P_{rx}(d)$ is the received power at distance d from the transmitter, P_{tx} the transmitted power, G_t transmitting antenna gain, G_r receiving antenna gain and λ carrier wavelength. Ns-2 added L, a loss variable to allow the user to alter the received signal strength so that transmission ranges can be modified. This formula results in the successfully reception of data within a perfect circle around the transmitter. However, this is not realistic.

2) Two-Ray Ground Reflection Model: The Freespace model does not take into account reflections or the height of the antenna and receiver which makes the model unrealistic. A two-ray ground reflection model [10] does consider those parameters and is able to calculate two wave receptions. A direct wave and a reflected wave over a perfect flat plane.

$$P_{rx}(d) = \frac{P_{tx}G_tG_rh_t^2h_r^2}{d^4L} \text{ [9]}$$

where h_t and h_r respectively are the height of the transmitter and receiver.
While this model might be usable for long distance radiowave communication, it is not really usable for short distances with no line-of-sight like Wi-fi. Ns has implemented this model in cooperation with the Freespace model because the received signal strength would vary too much for short distances.

3) Shadowing Model: This model tries to characterise wireless channels in which the line-of-sight is obstructed by large objects. This will result in poor signal quality behind the object, called shadowing. The ns-2 Shadowing model consists of two parts that influence the signal: large scale fading and small scale fading. The first is a slow variation of the received signal strength that models the obstruction of large objects. The second, small scale fading, tries to define the variations that result in rapidly changing signal strengths over a short time over a short distance, such as diffraction. The following formula states how it is mathematically defined in ns.

$$\left[\frac{P_{rx}(d)}{P_{rx}(d_0)}\right]_{dB} = -10\beta log\left(\frac{d}{d_0}\right) + N(0,\sigma)$$

where d_0 is the reference distance (typical 1m for indoor cells), β is the path loss exponent which is the large scale fading. Finally, $N(0,\sigma)$ is a Gaussian random variable with zero mean and a standard deviation (also called shadow deviation) of σ. This deviation defines small scale fading. The manual of ns-2 provides some empirically determined values for β and σ for both in and outside environments, together with different levels of movement and obstacles.

4) Nakagami Model: The newest propagation model in ns-2 is the Nakagami model [11]. This model calculates the received signal strength according to the formula of Friis, which is used in the Freespace model, and adds fading by a random variable that is distributed as the Nakagami-m distribution. The probability density function is stated below.

$$p(x,\mu,\omega) = \frac{2\mu^\mu}{\Gamma(\mu)\omega^\mu}x^{2\mu-1}exp(-\frac{\mu}{\omega}x^2)$$

where μ is the shape (or fading) parameter and ω is the spread. This implementation offers some advantages. First, this distribution can have the same behaviour as a Rayleigh distribution if you select the appropriate parameter (μ = 1). This distribution is very important as it is often used to model indoor radiowave propagation when there is no clear line of sight [12]. Second, ns-2 allows the user to define three distributions where every distribution models the behaviour for a certain distance interval. This gives us the possibility to model different behaviour over distance.

5) Simulation: We simulated a simple scenario where one node transmits data to another. We evaluated the packet reception on several distances. We defined the sensing threshold to a standard Linksys PCI wireless adapter (WMP54G) which is -85dBm for a datarate of 11Mbps and -70dBm for 54Mbps. Ns-2 does not implement the Auto Rate Fallback (ARF) mechanism so we decided to work with -85dBm. The Linksys adapter has a transmission power of 19dBm. We configured the radio propagation models for indoor use where possible. This means μ = 1 for Nakagami and Shadowing model has a β of 4 and a σ of 7 dB [9].

Fig. 3 shows some remarkable results. The range of the Freespace and Two-ray ground propagation models indicate line-of-sight as they reach 400-600m respectively. Nakagami starts losing quality at 250m end has no reception at all at 600m. Remember that Nakagami is based on Freespace model, which explains the long range. Notice the simplistic transition between perfect reception (100%) and no reception (0%) reception for Freespace and Two ray ground reflection model. There is no smooth degradation of probability on reception, unlike Nakagami and Shadowing.

Although Shadowing seems to have the most realistic properties, we decided to use the Nakagami model because

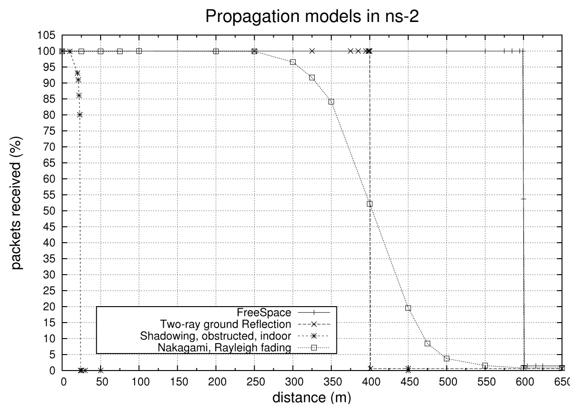

Fig. 3. Behaviour of the Propagation Models in ns-2. Shadowing model reaches the most realistic range for indoor radio propagation, while Nakagami has a richer model for signal degradation.

of its flexibility. The different behaviour over distance and the possibility of implementing Rayleigh fading are two important features that the other propagation models don't have. We implemented Rayleigh fading because most Wi-Fi networks are used in indoor environments.

B. Data Link Layer - MAC 802.11

We already discussed the problems that may arise at the physical layer and how ns-2 implements those issues. The next layer, the data link layer, also has to deal with some problems.

1) Overhead: 802.11 MAC layer uses CSMA/CA which is responsible for the reservation of the shared wireless medium. To avoid collisions, nodes send a Request To Send (RTS) message to announce the need of the medium. This request is answered by a Clear To Send (CTS) message that offers the authorisation to start sending the data. When the data transmission is over, the receiver replies with an acknowledge (ACK) to indicate a correct reception of the data. Although RTS/CTS has a very good efficiency for avoiding collisions, it is responsible for a non-neglible overhead. In fact, the frame size is an important parameter whether to decide if this mechanism is useful. Our video encoder creates small RTP (12Byte overhead) packets with a payload of 64 bytes. After encapsulation of UDP (8B), IPv4(20B) or IPv6(40B) and MAC(36B) is this a framesize of 140B (IPv4). This is an efficiency of only 46%. But there is more overhead: RTS (frame size of 20B), CTS (14B) and ACK (14B) needs to be send to finalise the transmission of 1 packet. If we add this to our data framesize, we have a total of 188B for 64B of video data. These numbers show that disabling RTS/CTS saves 22% of data over the network. However, turning this off enlarges the chance of collisions so more retransmissions will be necessary. Turning RTS/CTS off is only desirable when the amount of nodes are low and the network is not heavily used. Finally we add a remark that these calculations don't include retransmissions and don't take in account backoff timings and waiting periods between the transmission of subsequent frames.

2) ARQ: 802.11 MAC provides reliability on top of the physical layer with the Stop-and-Wait Automated Repeat reQuest (ARQ-SW). When frames are not received or if the Frame Check Sequence (FCS) indicate an error, the frame is discarded and no ACK is sent. The sender transmits the frame again, until an ACK is received, or a maximum amount of retransmissions is reached. This amount may vary from 4 retransmissions for short messages and 7 retransmissions for big messages [13].

We tested the performance of ARQ by simulating a constant bit stream of 10,000 packets of 64 bytes over 802.11g. We have intentionally chosen a slow throughput rate of 60kbps over a 54Mbps line. This makes sure we test the retransmission mechanism without saturating the link. We configured a bursty channel which is modelled as a two-state Markov chain (see chapter III-C). This bursty character has an average frame loss of 5-6 % while the amount of errors that propagate to the packet level, strongly depends on the amount of retransmissions of the ARQ.

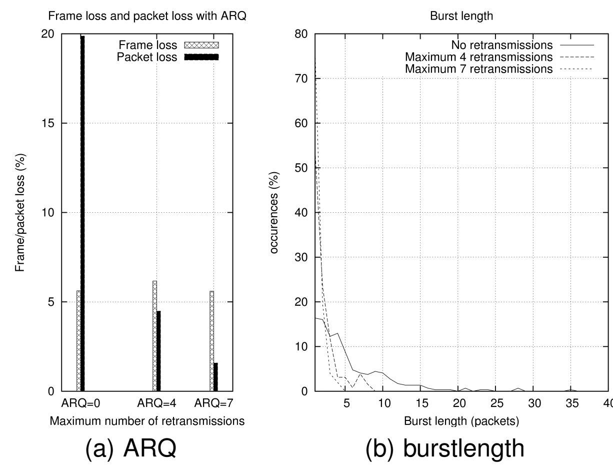

Fig. 4. ARQ has an important role to ensure correct data delivery. An average of 5% *frame* loss can result in a *packet* loss of 19.5%.

The results in fig. 4a shows that a frame loss of 5.5% leads to a packet loss of 19.5% when we disable ARQ. An ARQ of maximum 4 reduces 6.17% frame loss to 4.5% packet loss while an ARQ of 7 can repair 5.6% frame loss to 1.5% packet loss. We want to emphasise the difference between frames and packets as frames are at the data link layer and packets are at the transport layer. This means that 1 packet consists at least out of 4 frames (RTS, CTS, data, ACK).

ARQ also has a strong influence on the length of burst errors. While the same bursty error model was assigned to the link, the burst length diminished seriously as the amount of retransmissions increase.

Fig. 4 shows that burst length can be very long when no retransmissions are sent. Only 20% of all the packet drops, have a burst length of 1 packet when no retransmissions are enabled, while 4 and 7 retransmissions have respectively 50% and even 76% of only 1 packet burst length. The longest burst error for this scenario reaches 35 packets when retransmissions are disabled. 4 and 7 retransmissions bring this number back to 8 and 4 packets, respectively.

The retransmission mechanism of 802.11 has some disadvantages. A delay is introduced for every retransmission and the RTS/CTS mechanism uses valuable time at the shared medium that might endanger the following frames of the stream.

C. Error Model

We discussed the problems that may arise at both the physical and the data link layer. The implementation of the physical layer in ns-2 is not very accurate, especially errors on a wireless link are not realistic so they must be defined by the end-user. We used a two-state Markov chain to simulate the bursty character of errors. A Markov chain has two states: a good state (correct reception of packets) and bad state (when an error occurs). The state of the next packet can be calculated, only depending on the current state. We used a previous study which set-up an experiment and created the Markov chain for 802.11b [14] to describe bursty errors.

D. Video Traffic

The video encoder uses rate allocation which tries to keep the bitrate constant. This constant bitrate is fed into the row/column interleaver which also acts as a buffer. That interleaver is a circular buffer that contains 64 records of 255 bytes, or 16.320KB. This means when the actual throughput of the network channel is smaller than 16.320KB/sec, a constant bitrate is maintained in the network, otherwise we get a bursty character. However, we assumed a filled buffer at all times which causes a constant bitrate.

IV. Results

The simulation setup has 2 nodes where a sender transmits a protected stream to a receiver over an access point. This simulation introduces packet loss at the receiver while the transmitting node, and the transport to the access point is assumed ideal.

A. Scenario 1

This scenario simulates the transmission of video sequence 'football' over a 802.11g link at 54Mbps. The first 180 frames of the sequence are encoded at 1Mbps with a frame rate of 30 fps. This stream needs an actual throughput of 1.04Mbps and is protected from channel loss of 10 to 20 %.

Fig. 5 shows the peak-signal-to-noise ratio (PSNR) for the sequence on the left y-axis and the error model on the right y-axis while it is send over the network in function of time (x-axis). We ran this simulation twice, each time with another error model. Fig. 5a shows 2 quality drops at the beginning of the video stream and fig. 5b, shows more quality drops. Although they don't match on the time-axis, the causes of the drops can be assigned to the peaks in packet loss who crosses the 20% threshold. The unfinished part of this research, the implementation of delays from encoder and decoder, can be seen in both figures. We did not implement delays which means that the video starts to play, when the first bit is received. However, the delay can easily be derived from the figure because the packet loss peaks result in PSNR drops at a constant time shift of 0.5 seconds.

If we compare both error models of this scenario, we can conclude that an higher average loss, like in fig. 5b, does not affect the PSNR of the video streams, as long as the maximum protection threshold is not exceeded. Successive simulation runs showed similar results.

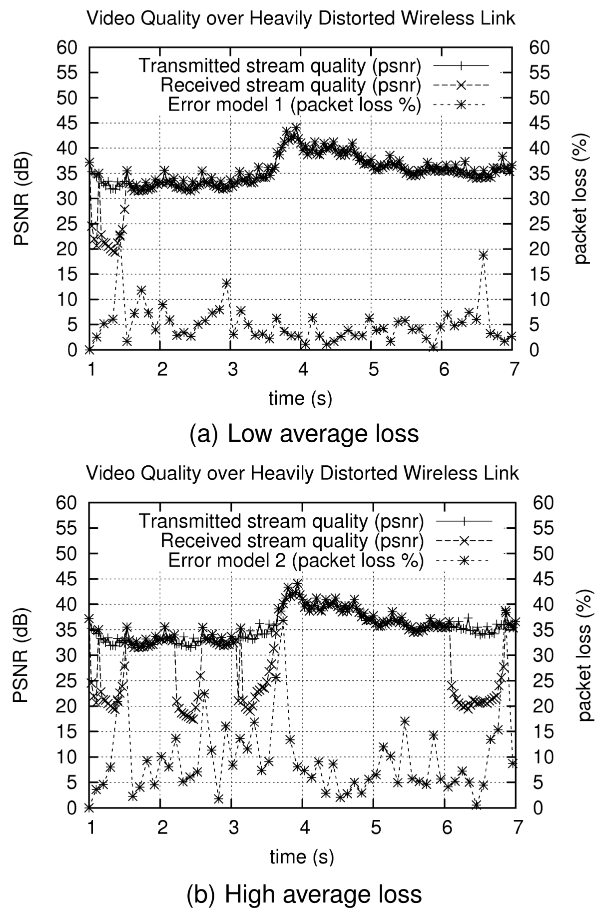

Fig. 5. Distorted Links Causes over 20% Losses. The decoder can repair the packet loss as long as the actual loss is lower than the protected loss. Otherwise the packet loss results in graceful degradation.

B. Scenario 2: Mobility

We implemented Rayleigh fading for this scenario as we discussed in chapter III-A5. Although, we added a slight modification. We assigned a loss variable (L=25) to model the transmission range to a more common range with Wi-Fi. We used that propagation model to find out how the video decoder reacts when a receiver gets out of range. Fig. 6 shows the distance between the access point and the receiver. This receiver starts at a distance of 50m, moves further away to 90m and returns. The figure also shows the amount of packets that are dropped in percent, because the signal could not be received. We must emphasise that we did not include an additional error model.

We see an almost perfect packet reception when the receiver is less than 60m away from its access point. The further we move away, the larger the fluctuations are of dropped packets. You can see that the amount of dropped packets correlate to the distance.

We simulated a transmission of a videostream over the network and let the receiver slowly move away from the access point in order to achieve a graceful degradation until the decoder no longer receives sufficient information to reconstruct the video sequence. We saw in scenario 1 that the decoder can protect the stream for a certain channel loss but when the actual loss is higher than the loss against which the stream is protected to, the decoder suffers from gracefull degradation. These simulations showed that no reconstruction was possible because the fluctuations of packet drops were too great. Losses that are much higher than the protected loss can't be reconstructed. Unfortunately, these fluctuations are inherent to indoor wave propagation so we can conclude that no visual degradation is possible at the edges of the wireless network.

Successive simulation runs showed similar results. In this particular case, averaging is not useful, as we are interested in the behaviour of the video system in a realistic environment.

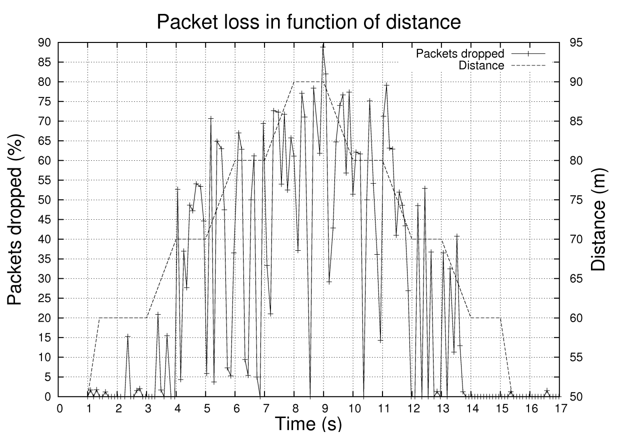

Fig. 6. Scenario 3: Mobility. The dropped packets highly fluctuate when the distance between the access point and the receiver is higher than 60m. Notice the correlation between the distance and the dropped packets.

V. Conclusions and Recommendations

We designed a wireless simulation environment with an extra error model that implements 802.11b/g. We used this environment to send a video stream from a sender to a receiver and examined the decoded stream. When we discussed the 802.11 MAC layer in depth, we noted a very low payload efficiency of 46%, used by the video encoder. This inefficiency leads to a very low actual throughput of the wireless channel. These small packets also have benefits, especially from point of view of the video encoder designers. The impact of packet loss is smaller because it contains less data and buffer size can be minimised which keeps delays low. However, we advise the designers of the video encoder to find a way to add more payload in the RTP packets, otherwise it might be difficult to deliver high quality video over 802.11b/g.

Another recommendation from our study of 802.11b/g, is the implementation of UDP lite[15]. This relative new protocol is designed especially for multimedia transmission and is based on the fact that corrupted received packets contain more information than discarded packets. This protocol can be used with 802.11 at the datalink layer with the same mechanism. Modifications to the video encoder's unequal error protection mechanism can be used to repair the corrupted packets.

We simulated the transmission of the encoded video stream over severely disturbed wireless networks. The simulations demonstrate the ability of the decoder to reconstruct the lost packets when the encoder receives correct parameters of the state of the network channel. Sudden increases of lost packets can result in a temporary graceful degradation of the video quality.

The last scenario showed us that packet reception at the edges of an indoor wireless network is very unreliable. The large amount of packet drops disables the decoder from reconstructing the protected video stream, so that graceful degradation can not be achieved here.

References

[1] D. Bakker, D. Cromboom, T. Dams, A. Munteanu, and J. Barbarien, "Priority-based error protection for the scalable extension of h.264/svc," Strasbourg, France, 7-10 april 2008, photonics Europe 2008 International Symposium.

[2] G. B. A. L. Thomas Wiegand, Gary J. Sullivan, "Overview of the h.264/avc video coding standard," *IEEE Transactions on circuits and systems for video technology*, vol. 13, p. 17, 2003.

[3] C. Huitema, "Real time control protocol (RTCP) attribute in session description protocol (SDP)," Internet Engineering Task Force, RFC 3605, Oct. 2003. [Online]. Available: http://www.rfc-editor.org/rfc/rfc3605.txt

[4] S. Lin and D. J. Costello, *Error Control Coding: Fundamentals and Applications*. Pearson Prentice Hall, 2004.

[5] J. Flood, *Telecommunication Switching, Traffic and Networks*. New York: Prentice-Hall, 1998.

[6] "The network simulator - ns-2," http://www.isi.edu/nsnam/ns/, version 2.33.

[7] V. Sridhara and S. Bohacek, "Realistic propagation simulation of urban mesh networks," *Computer Networks,*, vol. 51, p. 21, 2007.

[8] T.Pratt and C. W. Bostian, *Satellite Communications*. John Wiely and sons, 1986.

[9] *The ns manual*, June 2008.

[10] T. S. Rappaport, *Wireless Communications: Principles and Practice (2nd Edition)*. Prentice Hall, 1996, ch. Chapter 3: Mobile Radio Propagation: Large-Scale Path Loss, pp. 69-138.

[11] M. Nakagami, *The m-distribution: A general formula of intensity distribution of rapid fading*. New York: Pergamon, 1960, ch. Satistical Methods in Radio Wave Propagation, pp. 3-36.

[12] D. Laurenson, "Indoor radio channel propagation modelling by ray tracing techniques," Ph.D. dissertation, University of Edinburg, 1994.

[13] "IEEE Standard for Information technology Telecommunications and information exchange between systems Local and metropolitan area networks Specific requirements," 11 2007.

[14] S. A. Khayam, S. Karande, H. Radha, and D. Loguinov, "Performance analysis and modeling of errors and losses over 802.11b lans for high-bit-rate real-time multimedia," *Signal processing: Image Communication*, vol. 18, pp. 575–595, 2003.

[15] *UDP Lite for Real Time Multimedia Applications*. Vancouver, Britisch Columbia, Canada: IEEE International Conference of Communications (ICC), june 1999.