Performance of H.264 compressed video streams over 802.11b based MANETs

Carlos T. Calafate, M. P. Malumbres, P. Manzoni
Department of Computer Engineering
Polytechnic University of Valencia (UPV), Valencia, SPAIN
{*calafate, mperez, pmanzoni*} @*disca.upv.es*

Abstract

This paper addresses the problem of video streams transmission in 802.11b based Mobile Ad-hoc Networks (MANETs). Through the study of the delivery of a live encoded video stream, we expose the characteristics of different routing protocols and the infeasibility to provide QoS. The analysis shows the impact of a MANET on H.264 real-time video flows in terms of packet loss, end-to-end delay, jitter and distortion, and the behavior of H.264 error resilience tools in order to determine their effectiveness on such network scenarios. The results show that video traffic has demands that are hard to be met by a standard MANET, and that improvements are required in terms of routing protocols and QoS provisioning either on the MAC layer or at IP level using traffic shaping tools. Also, most of H.264 error resilience tools are not so effective as expected with this kind of networks, being the random macroblock updating the most effective one we have tested.[1]

1. Introduction

The increasing use of mobile devices and the demand for video oriented applications is leading companies and researchers to look for solutions in the field of mobile multimedia. Several improvements related to video compression technology were made in recent years resulting in the ISO MPEG-4 Part 2 [1] standard and ITU-T Recommendation H.263 [2]. The JVT H.264/MPEG-4 part 10 is a new standard that offers an enhanced video technology which provides superior compression performance and better error-resilience, as well as many other features as will be exposed in section 3. Such improvements pave the way for ubiquitous human-to-human video communication, even when using low-bandwidth and error-prone network environments.

Until now, most of the studies done about MANETs and related performance issues have relied on overall statistic results regarding packet losses and other parameters of significance. In this paper we follow a different strategy in order to provide an accurate study of real-time video on 802.11b based MANETs. Our analysis focuses on a single H.264 video stream, so that the effects of different routing protocols and CSMA/CA radio technology are put into evidence in terms of packet losses, packet loss patterns, end-to-end delay and jitter. At the same time, we will be able to analyze the behavior of the H.264 error-resilience tools, to evaluate their effectiveness in terms of perceived video quality distortion.

We shall measure the impact of several 802.11b MANET aspects over the final video quality perceived by the end user, like ad-hoc routing algorithms, mobility and traffic patterns, etc. We shall also evaluate current video error-resilience techniques.

Concerning the structure of this paper, in the next section we introduce some important aspects related to 802.11b based MANETs. Section 3 presents the H.264 video codec and the available error-resilience mechanisms, and in section 4 we describe the simulation framework. Simulation results are presented in section 5, and concluding remarks are made in section 6, along with some guidelines about future work.

2. Issues concerning 802.11b based MANETs

IEEE's 802.11b standard [3] is being increasingly used throughout corporations worldwide due to its good balance of cost, range, bandwidth and flexibility. The bandwidths set by the standard range from 1 to 11 Mbps, but other standards in the same family aim at higher bandwidths. The 802.11 standard offers operation modes named Point Coordination Function (PCF) and Distributed Coordination Function (DCF). PCF is used in infrastructure mode, where Access Points are responsible for coordinating the transmissions from nodes. DCF, on the other hand, is a distributed mechanism through which each node has the responsibility of sensing the medium, to avoid and react to collisions.

The medium access technique (CSMA/CA) is currently being enhanced by the IEEE P802.11 task group E in order to provide a framework for QoS. Our analysis is focused on 802.11b networks with Distributed Coordination Function.

Due to their nature, MANETs are very unstable due to frequent route changes caused by node movement, node on/off activity or even noise. In an ideal situation, messages associated with routing on MANETs should be given a high priority since that, due to mobility, route changes are very frequent. Also, mobile nodes are usually battery bounded, which means that sending data though invalid paths should be avoided whenever possible.

However, in order to sense the unpredictable neighborhood, MANET nodes typically make use of broadcasting. Broadcast packets, unlike unicast packets, are not acknowledged and also do not benefit from the RTS/CTS mechanism. This means that they are transmitted only once, and so there is no assurance that the packet will be correctly received by any surrounding node, suffering therefore of interference, collisions, or time-varying channel effects. Moreover, the fact that they are sent at the lowest possible rate (1 Mbps) increases the transmission time, which also increases the probability that these frames collide. Despite of these drawbacks, most routing protocols broadcast packets in many occasions, such as to advertise themselves through "Hello messages", request a route to the neighbor nodes, or both.

So, we can advance that the nature of wireless channels and the behavior of routing protocols can have a severe impact on video transmission performance.

3. H.264 related issues

The recent video coding standard H.264 [4], part of an activity on-going since 1997 named H.26L, was developed by the Joint Video Team (JVT), an alliance formed by the former ITU-T VCEG and ISO MPEG-4 groups. This new standard is not application-specific, and performs significantly better than the available ISO MPEG-4 Part 2 standard [1] and ITU-T Recommendation H.263 [2] in terms of compression, network adaptation and error robustness.

In the design of the H.264 codec some mechanisms were included on both encoder and decoder envisioning enhanced performance in lossy environments, such as wireless networks or the Internet. By tuning certain parameters, the user can obtain a trade-off between compression rate and error resilience.

The most commonly used methods to stop temporal propagation of errors when no feedback channel is available are the random intra macroblock updating and the insertion of intra-coded pictures (I-frames). While intra frames reset the prediction process, avoiding error propagation, their use has a generally high bandwidth cost, causing also severe bit rate variations. The use of random intra macroblock updating is more effective than I-frames at high loss rates because it not only aids in generating streams with more constant bit-rate, but can also provide better results by statistically resetting the error for each of the macroblocks.

Multi-frame prediction is another tool targeting to increase both compression performance and error-resilience. This is achieved by using more than one reference frame in the prediction process. As exposed in [5], this technique is particularly useful after the loss of a full frame when some of the previous reference frames are available, enabling partial motion compensation.

Concerning the decoder, it plays a fundamental role in error resilience since it is responsible for error concealment tasks.

4. Simulation framework

In a previous work [6] we performed a detailed analysis of different tuning parameters and behaviors integrated in H.264 codec. The performance of the H.264 codec was evaluated using the reference software JM3.9a.

Taking into account the results from that work, the Hadamard transform, CABAC and Rate Distortion Optimization were used since they offered the best results. The use of adaptive block transforms for inter and intra macroblocks was set to the fully flexible mode. Concerning error-resilience issues, the best options were enabling random intra macroblock updates - set to 1/3 of frame size - and applying FMO reordering.

The chosen test sequence is the well-known QCIF Foreman whose size may be adequate for current PDAs and other mobile devices' displays.

The frame rate is set to 10 frames per second, and the average bit rate is 178.64 kbps. In order to perform the desired evaluations we used the Network Simulator (NS-2) [7] version 2.1b9a. NS-2 is a discrete event simulator. The physical layer for the simulation uses two-ray ground reflection as the radio propagation model. The link layer is implemented using IEEE 802.11b Distributed Coordination Function (DCF), and the Media Access Control Protocol (MAC) is CSMA/CA - Carrier Sense Multiple Access with Collision Avoidance. This module was modified in order to correctly update the *contention window size* and the *short retry count*.

The transmission range for each of the mobile nodes is set to 250m and the bandwidth to 11Mbps (full rate).

To evaluate the desired video streams the RTP output from the H.264 encoder was converted to the NS-2's native input format. That way, we are able to stress the network with real-life video traffic instead of relying on CBR flows.

Our evaluation is done over 100 simulated seconds and all results presented are average results from 20 random simulation processes.

After the NS-2's simulation process ends, we process the output results in order to determine the reconstructed video sequence according to the packet loss pattern. This method aims at performing evaluations as real as possible, in order to predict the effects of MANET networks and video codecs on the video quality perceived by the final user.

5. Performance Results

In this section we start with a preliminary evaluation of several ad-hoc routing algorithms in order to determine their average re-routing times. If re-routing times are long MANETs will have problems to deliver compressed video streams, being a very important factor for this kind of application.

Afterwards we measure the impact of node mobility in a typical scenario, followed by an analysis of the final delivered video quality under variable network congestion.

Mobility and congestion are two different aspects that may affect the video quality performance at different degrees. For that reason we also test node mobility and network congestion independently.

Finally, we will test the behavior of the H.264 video codec, analyzing the effectiveness of its error-resilience tools at different network congestion levels.

5.1. Preliminary evaluation

Protocols used for routing in MANETs are usually divided into two main categories: reactive and proactive. Moreover, another division can be made according to the way in which they detect link failures. While the method of sending "Hello" messages is more universal, 802.11b enables the use of a more effective and efficient method to detect link breaks by using the information it provides. Awareness of the link layer allows nodes to react to broken links more quickly, avoiding sending packets to nowhere.

Broken links are the main cause of long packets-loss bursts in MANETs. In fact, long packets-loss bursts can be a major source of problems for video streams. This problem is more evident when Hello packets are used to detect broken links. Typical "Hello" intervals [8, 9], range from 1 to 2 seconds, and so re-routing times can be as high as 6 seconds or more - a connection is considered lost usually after 3 missing "Hellos". Since such failures are too long to be handled even by the most versatile video codec, we recommend enabling protocols with link-level awareness in order to perform re-routing tasks as soon as possible.

In this preliminary evaluation we use a single UDP flow to determine the re-routing times of different protocols. The evaluation is made in worse scenario situations and the number of hops to destination is varied between 2 and 5 hops. Table 1 presents average results for some well-known routing protocols.

Routing protocol	Re-routing time (s)
AODV	0.248023
DSR	0.275674
TORA	0.497529
AODV-H	2.7729
OLSR	12.1324

Table 1. Average re-routing times in worse case scenarios

AODV-H and OLSR use "Hello" messages for link sensing, and so perform significantly worse than link aware protocols as expected. Moreover, re-routing time for "Hello" based protocols depends essentially on the "Hello" period and on the number of missed "Hellos" until the link is considered lost. In OLSR the "Hello" period is 2 seconds, twice that in AODV; both consider the link is lost after 3 failed "Hellos". This implementation of OLSR also requires that a node receives 3 "Hellos" from a neighbor before the link between both can be used, which explains why this value (worse case) is twice the one in a normal case.

5.2. Mobility evaluation in a typical scenario

After this initial evaluation, we devised a scenario with 30 nodes in a 670×670 area. Mobility was generated through the random waypoint model available in the NS tool with maximum speed set to 3 m/s for low mobility and 8 m/s for high mobility. In addition to the video flow, 5 background FTP flows are also set (1 every 6 nodes).

Routing protocol	Low mobility		High mobility	
	PSNR	% loss	PSNR	% loss
AODV	26.78	28.4	22,84	51.8
AODV-H	25.04	37.9	22.09	55,6
TORA	26.45	31.9	27.02	31.4
DSR	24.11	43.0	25.02	44.1

Table 2. Performance of different routing protocols relative to the video test sequence

Table 2 shows the results achieved by using different routing protocols, with this scenario, in terms of distortion and packet loss rate. "Hello" based AODV performs relatively well in situations of low mobility because route changes do not occur so often. Also, there are less chances that background congestion causes one link to be considered lost (3 consecutive "Hellos" have to be lost). TORA shows the best overall behavior under this scenario, showing good distortion levels at all speeds and good ability to maintain the packet loss rate at high mobility. DSR is also

able to maintain steady levels of distortion and packet loss rate, although not so efficiently as TORA.

This analysis does not pretend to evaluate the goodness of different routing protocols, but rather to evaluate the video performance achieved on a congested network using different routing methods. Please refer to works such as [10] for a more general study on the performance of different routing protocols.

5.3. Performance under congestion

After the mobility evaluation we chose both TORA and AODV protocols to proceed with our analysis. We evaluate their performance when submitted to different levels of congestion at user mobility levels (low mobility). These results were achieved using the same 30 node square scenario described in the previous subsection.
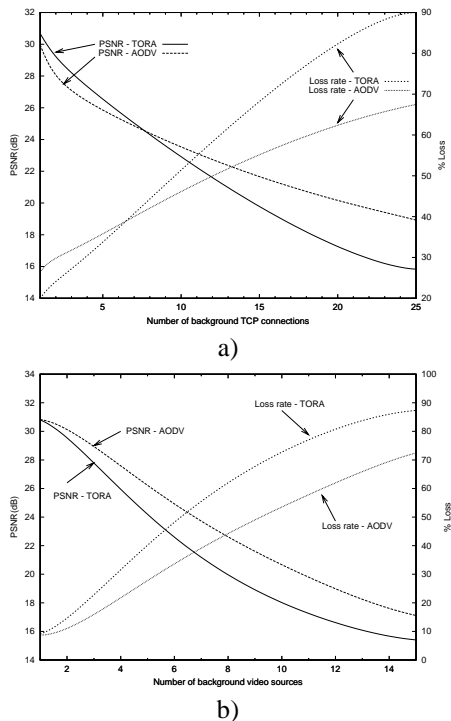

a)

b)

Figure 1. PSNR and packet loss rate performance for a variable number of background a) TCP connections b) video connections

Figure 1.a) allows us to compare the performance of TORA and AODV with a variable number of TCP connections in the background (TCP traffic is currently the most common - FTP, Web, Telnet, Database Access, etc. - but perhaps not in the future). We can see from that figure that acceptable distortion levels cannot be reached with more than 10 background connections using either TORA or AODV. TORA is, therefore, the best choice for this range and, even though AODV performs significantly better un-

der critical levels of congestion, the results in terms of distortion are almost at noise levels.

Figure 1.b) shows a similar analysis, but now all the background traffic is composed of video flows identical to the one under evaluation. In this scenario AODV always performs better than TORA and, in overall, we consider AODV to be an adequate choice to support video flows as reliably and uninterruptedly as possible.

5.4. Results on the effects of re-routing and background traffic

To complete our analysis, we change the scenario shape keeping the same number of nodes and area size. Now the scenario is rectangular (1500×300 meters) to increase the average number of hops. Envisaging a differentiated analysis of mobility and congestion, we started with a situation having neither background traffic nor mobility. We then analyzed separately the effect of allowing high mobility to all nodes (maximum speed of 10 m/s and no background traffic) and the effect of congesting the network by setting all the nodes to transmit a moderated amount of CBR traffic (no movement). In all situations, the average (or exact) number of hops was three; the routing protocol used was AODV.
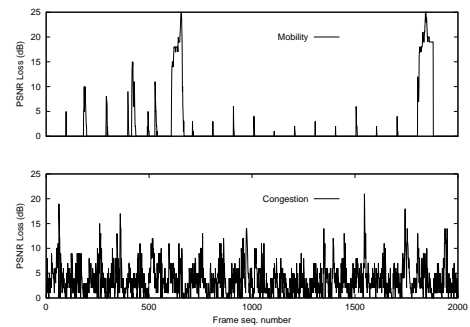

Figure 2. Effect of congestion and mobility on user perceived PSNR

Figure 2 shows the effects of mobility and congestion on user perceived video distortion. As it can be seen, mobility affects distortion in a bursty fashion, typically causing the loss of multiple frames and consequently freezing the image. On the other hand, traffic congestion causes packets to be lost in a more random fashion, so that distortion variation is smoother though more frequent.

The delay analysis also evidences the nature of both kinds of losses, as presented in figure 3.a).

In the reference situation (still), more than 99,9% percent of the packets arrive before 7 ms; with high mobility, 92% of the packets arrive in less than 10 ms. Point X is the frontier of two distinct regions: the one on the right where a very small number of packets have very high delays (as much as 6 seconds), and the one on the left where packet forwarding is uninterrupted.

In the "mobility" scenario, although the average number of hops is 3, this value varies throughout the simulation, explaining why some of the packets arrive earlier than those in the reference scenario and others arrive later (before X). The phenomena whereby some packets arrive with very high delays (after X) is expected since AODV causes packets to wait in a queue when re-routing tasks are being performed.
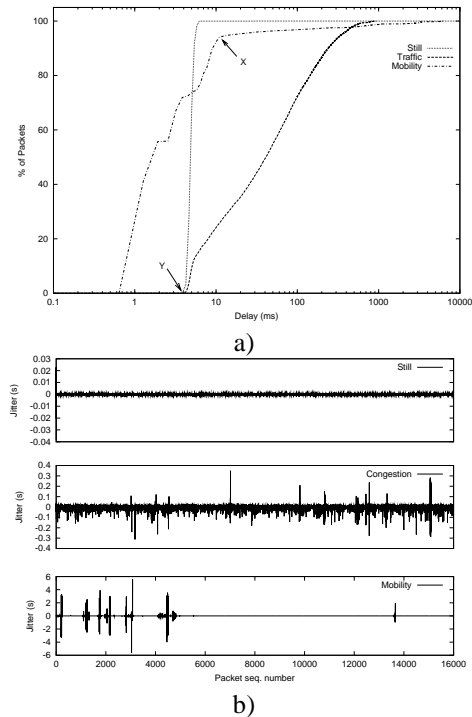

a)

b)

Figure 3. Effects of congestion and mobility in terms of a) delay b) jitter

Congestion causes a very different behavior, so that all packets that arrive to the destination do so in less than 1 second, though the delay between consecutive packets can vary greatly. The start point (Y) for both reference and congestion scenarios is common because the destination is 3 hops away on both.

The jitter analysis of figure 3.b) also aids at visualizing the behavioral difference between both. Even though the jitter peaks occur rather infrequently, they are an order of magnitude superior than those caused by congestion. We conclude that jitter peaks usually translate into a change of route when using reactive protocols.

As it could be inferred from previous results, tightening the limits on packet delay causes more negative effects in high-congestion scenarios than in high-mobility ones. However, these effects can be countered by QoS policies at either the MAC or higher levels. Transmission breaks due to mobility are much more difficult to counter and are more crit-

ical. Solutions to this problem could be introduced at the MAC level itself by giving routing traffic a higher priority using the developing standard 802.11e. Due to the nature of the wireless channel, though, we are not able to provide a 100% delivery guarantee even to a single surrounding node.

5.5. Evaluation of video codec choices

Our evaluation concerning the video codec parameters focuses on two topics: the number of reference frames for motion estimation and the best method for intra-macroblock updating. The evaluation relative to the number of reference frames was done using the heavy congestion scenario presented in the previous subsection.
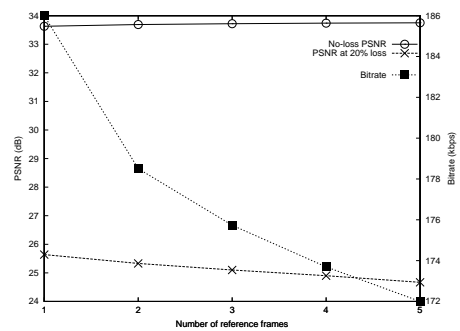

Figure 4. Performance under high congestion of a variable number of reference frames

In figure 4 we present the distortion achieved in this scenario. As it can be seen in that figure, the use of multiple reference frames improves the reference distortion and reduces the bit-rate slightly, being therefore the expected result. In terms of error-resilience there is a monotonous distortion decrease and a degradation of 1 dB is achieved by using 5 reference frames instead of just 1.

Since this result was unexpected according to [5], we completed our analysis by evaluating the performance of this parameter in the situation it was originally proposed for: entire frame losses. Instead of running a high mobility scenario (known to cause that kind of losses), we have directly tested the effects of loosing 1 to 5 consecutive frames, so that the error propagation effect was presented as clearly as possible.

The results found again point to using one reference frame in order to obtain the best performance, recovering from error up to 10 times faster comparing with the use of 5 reference frames. A proper tuning choice for H.264 in MANETs would be, therefore, to use only one reference frame; demands in terms of memory on both encoder and decoder are also reduced by this setup.

Concerning intra-updating of macroblocks, H.264 provides several choices to the user. We have evaluated the main available choices in the reference software, which are:

use of I frames, intra update a pre-defined number of macroblocks randomly and intra update a whole macroblock line randomly chosen for each frame.

In this process, all test files are encoded at the same bitrate and the scenario is the same one used in previous subsections. It is considered as an example of high congestion, with a packet loss ratio of 20%. Besides this scenario, we also created one with low congestion (4% loss) to provide a more consistent and general evaluation.

Updating method	Avg. PSNR (20% loss)	Avg. PSNR (4% loss)
1/3 random MB updates	25,58	30,63
IPP GOP sequence	24,01	30,19
IPPPPP GOP sequence	23,35	29,32
Random line intra update	22,79	28,93
No intra MB updates	20,62	25,84

Table 3. Average PSNR results using different strategies for intra MB updating

Table 3 presents the average distortion values for this scenario for low and high congestion levels. The use of random macroblock updates proves to be the best option in terms of error-resilience, showing its effectiveness with respect to no updating (around 5 dB of difference).

6. Conclusions and future work

We presented the main issues related to 802.11b based MANETs, taking into account the requirements of real-time video. The results from previous works related to H.264 were used to tune the video flows to achieve good error-resilience under severe losses. A preliminary analysis focused on typical re-routing times associated with common MANET routing protocols. That analysis evidenced the effectiveness of link-level aware routing protocols in re-routing tasks.

We proceeded with a mobility evaluation under average congestion, where TORA offers the best distortion results to the video stream. Variable congestion tests followed using TORA and AODV. Using TCP as background traffic, TORA has only provided slightly better results with less than 10 connections, with AODV offering a better overall performance. In fact, up to four extra video connections can be achieved with AODV relative to TORA maintaining the same level of distortion.

The obtained results evidenced that even though routing protocols detect broken links in milliseconds, they are not able to perform re-routing tasks as quickly as it would be desired. This phenomena occurs because, due to collisions, they are not always able to successfully broadcast routing packets, causing long transmission breaks. In fact,

increasing background traffic intensifies this problem, causing routing tasks to become more and more unfeasible.

An analysis of delay and jitter followed, showing the effects of congestion and mobility on video streams separately. Here, the ON/OFF behavior with high mobility can cause the loss of communication during long time periods (i.e., 10 seconds or more), being therefore prone to cause annoyance to the receptor. This point will require special consideration in further enhancements.

Concerning the H.264 video codec, we have also showed that the tuning performed was effectively resilient in terms of macroblock updating. The use of more than one reference frame, though effective in reducing bit-rate, increases the temporal error propagation and it should be avoided, expect for situations where the media is reliable (CD, DVD, or hard-disk).

Future work will focus on finding techniques suitable for offering good QoS to video streams by differentiating traffic flows, as well as by making routing related communication more reliable.

References

[1] ISO/IEC IS,Coding of Audio-Visual Objects, part 2: Visual (MPEG-4). Information Technology, November 2001.

[2] Video coding for low bitrate communication. ITU-T Recommendation H.263, 1995.

[3] IEEE/IEC Std 802.11. *Wireless LAN Medium Access Control(MAC) and Physical Layer (PHY) specifications.* The Institute of Electrical and Electronics Engineers, Inc., August 1999.

[4] ITU-T Rec. H.264 — ISO/IEC 14496-10 AVC. http://bs.hhi.de/ wiegand/JVT.html, March 2003.

[5] M. Budagavi and J. D. Gibson. Error Propagation in Motion Compensated Video over Wireless Channels. *Proceedings of the IEEE International Conference on Image Processing, Santa Barbara, USA*, pages 89–92, October 1997.

[6] Carlos T. Calafate and M. P. Malumbres. Testing the H.264 Error-Resilience on Wireless Ad-hoc Networks. *4th EURASIP Conference focused on Video / Image Processing and Multimedia Communications, Zagreb, Croatia*, July 2003.

[7] K. Fall and K. Varadhan. NS notes and documentation. The VINT project, UC Berkeley, LBL USC/ISI, and Xerox PARC, November 1997.

[8] Charles E. Perkins and Elizabeth M. Royer. Ad hoc On-Demand Distance Vector Routing. In *Proceedings of the 2nd IEEE Workshop on Mobile Computing Systems and Applications, New Orleans, LA*, pages 90–100, February 1999.

[9] T. Clausen, P. Jacquet, A. Laouiti, P. Muhlethaler, A. Qayyum, and L. Viennot. Optimized link state routing protocol. *International Multi Topic Conference, Pakistan*, 2001.

[10] S.-J. Lee, C.-K. Toh, and M. Gerla. Performance Evaluation of Table-Driven and On-Demand Ad Hoc Routing Protocols. In *Proceedings of IEEE PIMRC'99, Osaka, Japan*, pages 297–301, September 1999.