Cross-layer Feedback Control for Video Communications via Mobile Ad-hoc Networks

H. Gharavi and K. Ban

National Institute of Standards and Technology
100 Bureau Drive, Gaithersburg, MD 20899-8920

***Abstract*-This paper presents a cross-layer feedback control scheme for video communications over unstructured mobile networks for tactical operations. A peer-to-peer mobile ad-hoc network has been considered for the experimental testbed. Since ad-hoc networks can suffer greatly from heavy packet loss, in a multihop transmission we have considered a number of protection schemes to maximize the perceived quality of the video signal. In particular, we developed a rate control as well as a packet recovery scheme based on the network characteristics derived from the underlying ad-hoc routing protocol. In addition, to further enhance the quality of service, a redundant packet transmission scheme is presented for lossy recovery of the missing packets.**

Keywords: RTP video, H.264/AVC, mobile ad-hoc networks (MANET), IEEE 802.11, WLAN, testbed.

I. INTRODUCTION

For many tactical operations, communication between mobile nodes (such as robotic vehicles) requires high performance sensing channels capable of transmitting high quality video signals over unstructured mobile networks. This necessitates a peer-to-peer network assembly of wireless multihop ad-hoc routing. In addition, recent advances in Wireless LAN (WLAN) technologies such as the IEEE 802.11 standard [1], have provided new opportunities to develop an experimental platform to design and assess ad-hoc networks for transmission of multimedia information in realistic environments [2]-[4]. The IEEE 802.11 standard defines the carrier sense multiple access protocol combined with collision avoidance (CSMA/CA) [1]. This protocol can support WLAN in two different modes: infrastructure and ad-hoc. In the infrastructure mode, mobile nodes should communicate with each other via an access point (AP) that is based on a centralized protocol known as Point Coordination Function (PCF) [1]. For ad-hoc operation however, the standard specifies another access protocol called Distributed Coordination Function (DCF). DCF, unlike PCF, is not suitable for real-time applications due to unbounded delays at higher traffic loads. Furthermore, in multihop transmission the network throughput performance is shown to suffer dearly as the signal traverses over more hops [2]-[3]. This is mainly the result of co-channel interference, which tends to increase as the signal hops through more intermediate nodes. Another characteristic of ad-hoc networks is the large delay caused by the effect of route changes, which tend to occur more frequently under high mobility conditions [4]. This would result in a loss of significant numbers of video packets that can seriously reduce any chance of recovery at the destination node. Therefore, the most important challenge in providing video communication in ad-hoc network environments would be to maximize the perceived quality-of-service (QoS). Furthermore, for a network where resources such as the bandwidth and delays are expected to vary considerably, one effective approach would be to adopt a control mechanism via a feedback channel mechanism. Under the best effort transmission conditions, such a mechanism can be more conveniently utilized at the application layer. For instance, when Real-time Transport Protocol (RTP) is utilized for video streaming, the reception report via its control protocol (RTCP) can provide a packet loss assessment of the channel [5]. However, under high mobility conditions where the network topology is expected to change frequently, a more periodic feedback channel assessment may become necessary. Under current RTCP specifications, this could result in a significant increase in the network traffic. In addition, despite the fact that the RTCP reception report can provide useful information (e.g., packet-loss rate, sequence numbers of missing packets) due to the round trip delay it may arrive a little too late for any possible preventive actions.

Alternatively, we present a zero-delay, traffic-free channel assessment technique that can allow the transmitting node to adapt itself to the expected channel conditions (e.g. network resources). The proposed approach is simply based on acquiring the critical information from the underlying routing protocol. Note that information such as hop-count and route change indicator can play a crucial role in evaluating the channel condition. This information can then be utilized to appropriately control the coding strategy at the transmitting node. In particular, based on the extracted routing information, we present a packet recovery protection scheme where the number of parity packets is adaptively controlled in accordance with the multihop characteristics of the transmission path. Furthermore, against a burst of packet drops we present a redundant packet transmission scheme

This project is supported by DARPA NMS Grant: 8922385.

that can recover the missing packets at the expense of a graceful reduction in the received video quality. For video compression we have considered H.264/AVC video coding standard [6] for transmission in the form of RTP/UDP/IP packets. Finally, the paper describes our experimental system for evaluating the overall network performance in terms of the quality of the received video under various test conditions.

II. CROSS-LAYER RATE CONTROL

As real-time multimedia communications find their way into military applications, efforts to support QoS are becoming increasingly important. In particular, under best effort ad-hoc network environments where the routing and channel characteristics are expected to vary frequently (e.g., under high mobility conditions), achieving an acceptable video quality is becoming a challenging task. For instance, assuming a unicast transmission, two major factors affecting the ad-hoc channel performance are: i) large delays due to the route discovery process in the event of a route change [4], and ii) deterioration of the network throughput performance as the number of transmission hops increase. For instance, Fig. 1 shows the maximum throughput performance as a function of hop-count (number of hops from the source to the destination). These results were obtained using our experimental testbed where the maximum bitrate of all 802.11 WLAN cards (ad-hoc nodes) was set to 2 Mbit/s with retry limit 2 (up to 1 retransmission).

The results shown in Fig. 1 verify the degree in which the network performance deteriorates with increased numbers of hops. This behavior suggests that to improve the perceived quality of video it would be necessary to develop a rate control mechanism. Such a mechanism can be developed based on prior knowledge of the network characteristics. For instance, information about the transmission path or warning of a route change can provide a valuable cross-layer feedback to the source node in order to set its coding strategy accordingly. For instance, in ad-hoc routing protocols, such as AODV (Ad-hoc On Demand Distance Vector) [8] and OLSR (Optimized Link State Routing) [9], each node maintains the routing table for an entry (destination) with the hop-count (number of hops from source to destination).

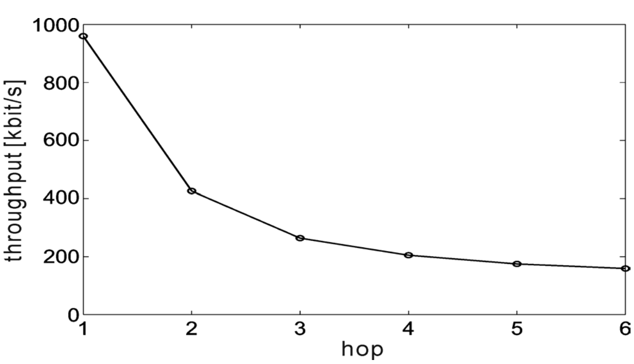

Fig. 1. Hop count versus maximum throughput performance (UDP packet size is 612-byte).

Although it is difficult to predict the channel condition (e.g. packet loss rate, delay, fading, etc.) between the transmitter and receiver, it is mostly true that the hop-count can provide valuable information about the expected channel conditions as far as the ad-hoc routing is concerned. For example, we can use this information to adaptively control the transmission rate as well as the amount of error protection.

In the case of AODV, if a route from source to destination has already been established, each time the source node tries to send packets, the source first checks its routing table entries to obtain the hop-count from the source to the destination. If the destination node is no longer listed in the table, the source initiates a route request (RREQ) to find a new route to the destination. As soon as the new route has been established, the source can then obtain the hop-count before passing it to the application layer.

When a route change occurs, the route error (RERR) message caused by the link breakage will be sent to the source node. The source node can use the reception of RERR or the initiation of RREQ as an indication of the route change so that it can change its transmission rate and coding strategy accordingly. Since a route change usually coincides with a change of hop-count, the source node may simply monitor a hop-count to detect the route change.

III. CODING STRATEGY

Although forward error correction (FEC) schemes can help to restore the missing packets, they have the unfortunate effect of increasing the transmitting bitrate. Therefore, without reducing the transmission rate it is not always a very efficient way to use a larger amount of FEC or redundant packets when the destination node is too many hops away. For a given number of hops, any protection strategy should be based on a two-way compromise between the video coding bitrate and the amount of FEC overhead. Although to achieve the best result may require finding an optimized solution, in this paper our rate/distortion control mechanism was based on a predefined set of parameters to change the video bitrate as well as the protection overhead on a hop-count basis. For instance, the video transmission bitrate can be controlled via its quantizer parameter (QP). In addition, a packet recovery control parameter has been defined in such a way that it can assign a higher degree of packet protection or redundant packet transmission to a route undergoing a large number of hops. As mentioned earlier, another important factor that can affect the performance of real-time transmission of video signals is the large delay caused by the process of route discovery. This would consequently result in the loss of a significant number of packets, which could make the process of resynchronizing the video stream at the decoder almost impossible. Under these conditions and based on the feedback that the video encoder receives from the ad-hoc routing table, the encoder would then stop transmitting packets. As soon as a new path has been established, the packet transmission is resumed by encoding the incoming frame in an intraframe mode (i.e., I-frame). We should point out that at the expense of additional coding delay, the encoder (prior to the I-frame) can send packets for the missing frames in the form of redundant packets (see next section). This arrangement can help continuity in displaying the video signal at the receiver.

A. RTP Packetization

The first step in RTP video streaming is to encapsulate a compressed video bitstream into RTP packets. The encapsulation can be accomplished with greater flexibility using the emerging H.264/AVC coding standard [6]. H.264 is designed to provide a "network-friendly" packet-based video representation. It is based on the conceptual separation between a video coding layer (VCL) and a so-called network abstraction layer (NAL) [6]. The NAL header consists of one-byte defining the payload type (6-bit packet-type, 1-bit error indication bit, 1-bit reserved bit). Following the NAL header, the RTP payload is comprised of the slice output of the VCL and its header includes parameter set, picture structure (progressive frame picture, top field picture, bottom field picture, etc), slice type (Intra, Inter, B, etc.), address of the first macroblock (MB) in the slice, and so on. The first macroblock address provides useful information in finding the number of macroblocks (MBs) that may have been lost in the preceding packets.

The coded elements within a packet consist of the slice-layer syntax elements such as; macroblock type (or MB-mode) defining a type of prediction used to encode the macroblock, coded block pattern (CBP) identifying blocks with non-zero coefficients in a macroblock, quantizer parameters which are transmitted as a delta value from the previous value of the quantizer parameter (delta QP), motion vectors, and the coded coefficient data. Based on the QP value, we can increase or reduce the average transmission rate dynamically upon a change of hop-count. The selection of the QP value also depends on the amount of parity packets generated for FEC protection. Indeed, based on our own experiments, a higher degree of protection (which would be at the expense of lower video bitrate e.g., high QP value) can have a better impact on the overall quality of the received video.

Therefore, based on the hop-count, a different degree of error protection has also been considered in our approach. A video frame synchronized RTP-based FEC protection strategy has been applied to protect H.264/AVC coded video packets [7]. In addition, to combat the burst of packet loss typical of mobile ad-hoc environments, we have also deployed a lossy packet recovery scheme, which is designed to handle bursts of packet drops [7]. This is accomplished by generating a redundant packet aimed at reducing the overhead by transporting only the most sensitive information in the video frame. In our current implementation, redundant packets have been applied only to P-frames (prediction frames) in the hope of minimizing the propagation of distortion. For I-frames (intra frames), where the effect of packet drops can have a serious impact on the subjective quality of the received video, an FEC protection scheme has been considered. The redundant packet strategy considered in this paper is based on the fact that recovering only the motion vectors can have a profound effect on concealing the propagation distortion, despite a loss of data coefficients. In this strategy, a redundant packet is generated for every frame using the same header as the first slice. The payload of the redundant packet consists of runlength codes for skip macroblocks, the macroblock mode, components of the motion vector, and finally the coded block pattern (CBP).

At the decoder a successfully received redundant packet is then utilized to regenerate the missing data packet(s). This initially requires identifying the first and last macroblock addresses in the missing packet (i.e., slice), which can be easily obtained from the successfully received neighboring packets. Based on this information, the relevant data is then extracted from the redundant packet to regenerate only the missing packet. Note that, depending on the degree of a packet-drop's burstiness, the redundant packet, which is generated for every coded video frame, can be transmitted a number of times per frame (interleaved within the data packets) by using the same RTP time-stamp. In this case also, the number of redundant packet transmissions can vary depending on the value of the hop-count. Finally we should point out that due to the video-frame synchronized FEC and redundant packet transmission the video rate control via QP parameter is set on a frame-by-frame basis.

IV. NETWORK SET-UP

An experimental testbed for a peer-to-peer mobile ad-hoc network has been successfully constructed to demonstrate the feasibility of video communication for mobile sensor networks. All the mobile nodes were implemented using IEEE 802.11. For ad-hoc routing we deployed the AODV routing protocol [8]. In our experiments, the bit rate for all the IEEE 802.11 devices has been set to 2 Mb/s. In addition, the maximum number of retransmssions on all WLAN deveices was limitted to one.

For RTP/UDP/IP streaming, H.264/AVC encoder software [10] has been imported to the UCL-VIC software [11]. In addition, we have modified the VIC architecture to enhance its packet delivery operation. To evaluate the performance of the network, various tests were carried out to measure distortion and received video quality in a multihop chain transmission. To carry out these measurements under the same conditions, a pre-captured video sequence was used instead of live video. The sequence was captured in QCIF format and at 8 frames/s. This rate is selected in accordance with the average frame rates that our network is capable of handling live video with respect to the limited processing power of the mobile nodes (i.e., iPAQs).

In our first set of experiments, the source node transmits UDP packets to the destination node over the mobile ad-hoc network. For packet recovery, we used the FEC or redundant packet scheme for P-fames together with the FEC scheme for I-frames. The percentage of FEC overhead for I-frames is set to almost the same as that for P-frames. We also used three quantizer parameters (QPs), which were selected dynamically and on a frame-by-frame basis to control the video bitrate in accordance with the hop-count value. As discussed earlier, the hop-count value was extracted from the AODV routing table. Fig. 2 and Fig. 3 show the average distortion of the received video frames when the destination is 2-hop and 6-hop away from the source, respectively.

In the case of 2-hop, the FEC scheme performs better than the redundant packet scheme. Since the packet loss rate is relatively small, the FEC scheme can almost restore the video frame completely, while redundant packet scheme only recovers the motion vector information. However, the difference in subjective visual quality is not very noticeable between FEC and redundant packet schemes.

In the case of 6-hop, the redundant packet scheme is more suitable than the FEC scheme. As the numbers of hops increase, we encounter burst packet loss situations more often. Since only one redundant packet is sufficient to restore a whole frame regardless of the number of lost packets (despite a loss of quality), the error protection by the redundant packet is more robust than FEC when large and/or burst packet losses are expected.

We should also point out that lowering the bitrate (i.e., via QP) has a very positive effect on the video quality when the source and destination nodes are located far away from each other.

Our final experiment was based on a scenario where the destination node moves away from the source node in such a way that the received video goes through steps of one to six hops. Each time a route change occurs the source node receives feedback indicating a new path is underway. The video source encoder then stops transmitting packets until the new route is established. Subsequently, the encoder, which has already stored data for the missing frames in the form of redundant packets, resumes transmission by sending these packets before transmitting the next incoming frame as an I- frame. Following the I-frame, the encoder switches to P-frame mode using the new QP value (e.g., to reduce the bitrate). The main motivation for sending a coded P-frame in the form of a redundant packet (instead of a sequence of regular packets for P-frame) is to avoid a substantial increase in the transmission rate.

With respect to packet recovery, the overhead associated with FEC and/or redundant packet transmissions increases as the destination node moves further away.

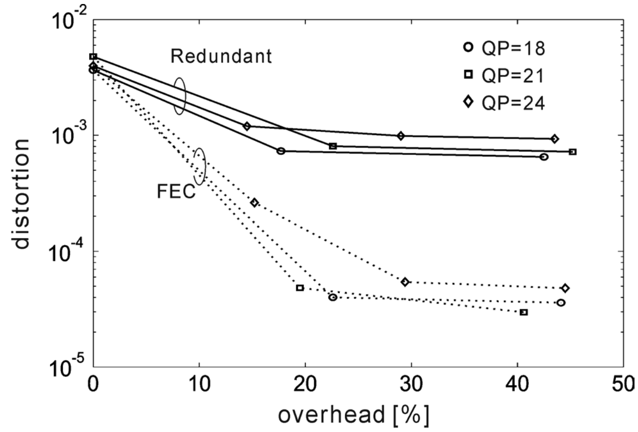

Fig. 2. Average distortion (2-hop)

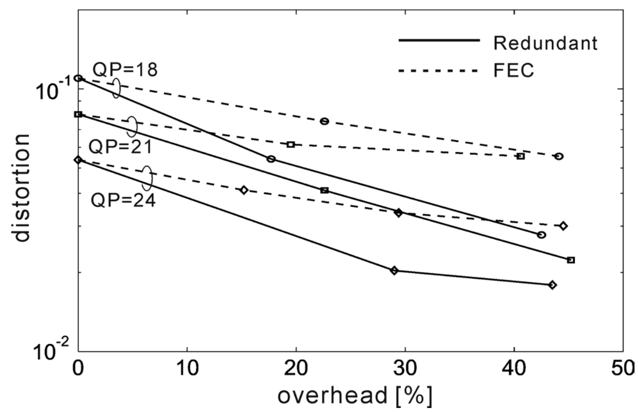

Fig. 3. Average distortion (6-hop)

In order to investigate impact of the feedback control scheme with rate control and dynamic packet recovery overheads, Fig. 4 shows the performance difference in trems of avarage distortion under the above test scenario. In this figure, we also include the results in the absence of feedback control (i.e., using a fixed QP value) mechanism with or without packet recovery for the purpose of comparison.

In the feedback control scheme we used QP=19 with 20 % FEC overhead for 1 – 2 hops, QP=21 with three redundant packets (33 % overhead) per frame for 3 – 4 hops, and QP=24 with four redundant packets (45 % overhead) per frame for more than 4 hops. For non-feedback transmissions, we used QP=19 without error protection, and QP=21 with 41 % FEC overhead so that all these schemes have almost the same average transmitting bitrate. Note that in these experiments the QP values and the amount of overhead associated with FEC and/or the number of redundant packets were selected heuristically and in accordance with the hop-count.

As can be observed from this figure, the feedback control scheme with rate-control and dynamic error protection successfully demonstrates its effectiveness in maximizing the perceived QoS by utilizing the routing information from the underlying layer.

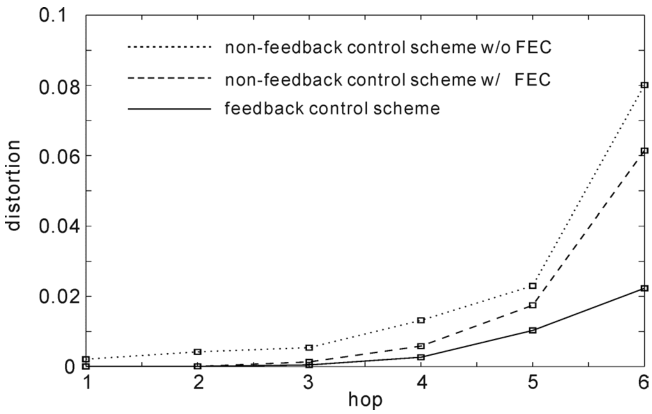

Fig. 4. Average distortion with feedback control

With regards to subjective evaluations, we compared the proposed feedback control approach with non-feedback video transmission by viewing the received video sequences on the screen in real-time. Indeed, the dynamic control strategy shown to have a profound effect on improving the quality of the received video. More importantly, the cross-layer control feedback not only offers much better quality but also provides continuity of video display, which is an essential criterion for reliability of video communication in sensitive tactical operations.

V. CONCLUSION

Our main objective has been to evaluate the feasibility and reliability of video communication over an unstructured mobile network. Evaluations were performed experimentally using a peer-to-peer ad-hoc network. For the ad-hoc routing aspect of the network, the AODV protocol was considered.

This network has been used to transport the encapsulated H.264/AVC coded video in the form of RTP/UDP/IP packets. For packet recovery, a combination of FEC and redundant packet transmission was considered. The main problem however, has been the heavy loss of packets when a change of routing path occurs or signal traverse over more hops. Therefore under the best effort scenario of RTP/UDP/IP transmission, to improve the perceived QoS we propose an overhead-free, feedback control approach that can estimate the channel condition as far as the ad-hoc routing is concerned. Based on our experimental testbed, we observed such an approach could improve the reliability of video communication over ad-hoc networks.

ACKNOWLEDGMNET

The authors wish to thank John Cambiotis of NIST for his help in developing the video streaming aspects of these investigations.

REFFERENCES

[1] ANSI/IEEE Std 802.11 1999 Edition, Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, Institute of Electrical and Electronic Engineers, Aug. 1999.

[2] A. Ganz, A. Phonphoem, N. Llopis, I. Kim, K. Wongthavarawat, and Z. Ganz, “Converged voice, video and data wired-wireless LANs testbed,” Conference Proceedings, IEEE MILCOM, Vol. 2, pp. 1297 –1301, Oct/Nov 1999.

[3] H. Gharavi and K. Ban, “ Vision-based Ad-hoc Sensor Networks for Tactical Operations,” World Wireless Congress, 3Gwireless’2002, San Francisco, May 28-31, 2002.

[4] H. Gharavi, K. Ban, “Multihop Sensor Network Design for Wideband Communications,” PROCEEDINGS OF THE IEEE, vol. 91, NO. 8, August, 03.

[5] H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson, “RTP: A Transport Protocol for Real-Time Applications, ” RFC1889, Jan. 1996.

[6] T. Weigand, “Document: JVT-G050, “Joint Video Specification (ITU-T Rec. H.264 | ISO/IEC 14496-10 AVC), March 2003.

[7] H. Gharavi and K. Ban, “Multihop Ad-hoc Network Design For Multimedia Communications,” unpublished.

[8] C.E. Perkins, E.M. Royer, and S.R.DAS, “Ad Hoc On Demand Distance Vector (AODV) Routing, ” IETF Internet Draft, draft-ietf-manet-aodv-12.txt, Nov.2002.

[9] T. Clausen, P. Jacquet, A. Laouiti, P. Minet, P. Muhlethaler, A. Qayyum, and L.Viennot, “Optimized Link State Routing Protocol,” IETF Internet Draft, draft-ietf-manet-olsr-06.txt, Sept.2001.

[10] H.264/AVC Software Coordination: http://bs.hhi.de/~suehring/

[11] http://www-mice.cs.ucl.ac.uk/multimedia/software/vic.