IEEE ICME, July 2003.

LOW-LATENCY WIRELESS VIDEO OVER 802.11 NETWORKS USING PATH DIVERSITY

Allen Miu◇, John G. Apostolopoulos†, Wai-tian Tan† and Mitchell Trott†*

†Streaming Media Systems Group
Hewlett-Packard Labs, Palo Alto, CA 94304

◇Laboratory for Computer Science
Massachusetts Institute of Technology, Cambridge, MA

ABSTRACT

Wireless local area networks, such as 802.11b, are becoming widespread as they provide simple wireless connectivity and *data* delivery. This paper examines low-latency (conversational) video communication over 802.11b networks. The challenges to enable low-latency video include overcoming the highly variable delays, losses, and bandwidth of 802.11b wireless networks. To overcome these challenges we (1) employ the H.264/MPEG-4 Advanced Video Coding (AVC) standard for high video compression efficiency and good resilience to losses, (2) use low-latency best-effort transport mechanisms, and (3) exploit the potential path diversity between each mobile client and multiple access points in the infrastructure, where we use multiple paths simultaneously or switch between multiple paths (site selection) as a function of channel characteristics. Our results indicate that the proposed system can provide significant benefits over conventional single access point (single path) systems.

1. INTRODUCTION

Wireless LANs provide simple and low cost methods of connectivity. In particular, 802.11b wireless networks are being quickly adopted to provide wireless infrastructure in businesses and educational institutions, and 802.11 radios are being integrated as standard components within mobile computing devices such as laptops and PDAs. Most current uses involve data delivery, however the desire for low-delay applications such as voice over IP (VoIP) and video over IP (video phone) is apparent. In contrast to data communications, which are very sensitive to packet losses but typically not sensitive to delays, voice and video communications can tolerate some losses, but have strict delay requirements. Specifically, if a voice or video packet arrives late, it is useless and is equivalent to a lost packet.

The goal of this work is to achieve conversational (low-delay) video communication over 802.11b wireless networks. For example, a mobile device may connect via an 802.11b wireless link to an access point (AP) and then over the wired infrastructure to another wired or wireless node. The wireless channels may be shared by other users, and may be afflicted by interference from microwave ovens, Bluetooth devices, or other 802.11b devices. Generally the wireless links, rather than the wired infrastructure, will be the bottlenecks. The challenges that must be overcome to enable low-delay video include overcoming the highly variable delays, losses, and bandwidth provided by the 802.11b wireless network.

*This work was performed during a summer internship at HP Labs. The authors would also like to thank John Ankcorn and Susie Wee of HP Labs for their contributions to this work.

There has been considerable prior work on 802.11 networking for *data* delivery, and also some prior work on low latency communication over 802.11b, including examination of the link-layer behavior for UDP traffic as a function of packet size [1] and the effect of Bluetooth or microwave oven interferers on 802.11 bandwidth and delay performance. Prior work on video over 802.11 includes [2, 3]. In [3] FEC is used to overcome time-varying wireless losses, for example by adding 50 parity packets to every 100 data packets, leading to a delay of 100–150 packets depending on the loss pattern. This approach is probably not appropriate for low delay applications. Error-resilient source coding coupled with path diversity has been examined for multiple description video coding and path diversity over packet networks [4], low-latency voice over IP [5], using the distributed infrastructure of a content delivery network (CDN) to achieve path diversity between multiple senders in a CDN and each client [6], and using path diversity over ad-hoc wireless networks for MD image communication [7]. However, we are not aware of prior work that examine the simultaneous transmission of video from multiple wireless 802.11 access points in order to overcome the wireless challenges.

2. PROPOSED SOLUTION

To enable low-delay video communication over 802.11b wireless networks we investigate a number of techniques across the entire system, including source coding, end-to-end streaming, 802.11b link-layer adjustments, and the simultaneous use of multiple access points in the distributed infrastructure of access points that are often available. Specifically, to overcome the aforementioned challenges we (1) base our video compression on H.264/MPEG-4 AVC (previously known as H.26L) which provides high compression efficiency and good resilience to losses, (2) use low-latency best-effort transport mechanisms, and (3) use potential path diversity from multiple 802.11b access points, where we can use multiple paths simultaneously or switch between multiple paths (site selection) as a function of channel characteristics.

High Quality, Low-Latency, and Error-Resilient Video Compression The MPEG-4 and H.263 Version 2 video compression standards, as well as the emerging ITU H.264/ISO MPEG-4 Part 10 AVC standard (previously known as H.26L), are reasonable compression standards for this application. In this paper, we use H.264 [8] since it provides the best video compression performance, is specifically designed to be resilient to packet losses, and supports low-latency applications. Further details on how we perform compression are given in Section 3.

Low-latency Packet Delivery over Lossy Channels The end-to-end latency is affected by end-to-end (host-to-host) behavior and 802.11b link behavior. The end-to-end behavior includes flow control and potential retransmits (which delay subsequent packets). To minimize the end-to-end latency, we compress at an approximately constant bit rate (CBR) and packetize the video so that it is resilient to packet loss. The resilience to packet loss, coupled with the fact that most losses occur on the wireless link (where we may use fast link-layer retransmits), obviates the need for end-to-end retransmits. The compressed video is sent as RTP/UDP/IP.

The 802.11b wireless network operates in the unlicensed 2.4 GHz ISM band, and is often afflicted by a variety of impairments. The quality of an 802.11b link generally varies with time, and depends on radio propagation in the local environment, mobility, cross traffic, and interference from microwave ovens, Bluetooth transmitters and other sources. These effects contribute to packet losses in the wireless link. While low signal quality can corrupt data transmission, contention from both exposed and hidden nodes can cause packet collisions.

When link losses occur, 802.11b may invoke link-layer retransmissions to recover lost packets. Each retransmission incurs a delay of 2–22 ms (the delay grows as the randomized backoff increases with each successive failure). It has been shown that with modifications to existing 802.11b drivers, one can intelligently set a maximum retransmission limit for *each* packet to tradeoff latency and loss in a fine-grain manner [2]. We have not experimented with per-packet adaptive retransmission algorithms, as our focus is to understand how the conventional approach behaves under transmission schemes using single and multiple access points.

Low-latency Packet Delivery over Busy Channels Another source of latency arises from the carrier sense mechanism used in 802.11. Before sending a packet, the transmitter detects whether the channel is busy (either from an ongoing transmission or a non-802.11 interferer). If the channel is busy the sender blocks until the channel is free, causing a *non-deterministic* delay. Much prior research has examined how to perform priority scheduling to minimize jitter for real-time traffic in 802.11 networks (e.g., [9]). However, these approaches can not be easily supported on the current implementations of 802.11b.

We can overcome the above problems by using multiple access points. Even though a client conventionally talks with a single AP at any point in time, there are often a number of nearby APs in the infrastructure. Each of these APs is often strategically placed at a different location, and therefore offers a different relationship to the client with respect to distance, obstructions, multipath, signal strength, contention, available bandwidth, neighboring interferers, and potential hidden nodes. This suggests that one can use the *distributed infrastructure of APs* surrounding a client to provide sender or receiver *diversity* and thereby improve the client's communication from the wireless to the wired space. Current commercial 802.11 cards allow simultaneous communication via multiple APs only using ad-hoc mode, however it appears that this capability may be possible in infrastructure mode using only software changes.

3. EXPERIMENTAL SETUP AND RESULTS

To evaluate the performance of using multiple access points for video communication with low-latency constraints, we conducted a number of experiments to collect appropriate packet traces. As in Figure 1, a single source sends a uniformly spaced sequence of 1500 byte packets to each of two 11 Mb/s 802.11b wireless access points, AP1 and AP2, over a 100 Mb/s wired Ethernet network. The packet spacing is 1/30th of a second, to emulate video frame rate of 30 frames/s. Each packet contains a departure timestamp and a sequence number. Whenever a packet arrives, the wireless access points immediately forward it to the same *mobile* receiver over the same channel using 802.11b ad-hoc mode. The access points are physically separated by about 25 meters of lab space area that is occupied by open cubicles and have been tested to be well within radio range of each other. The APs are configured with RTS/CTS and packet fragmentation disabled and a factory default transmission retry limit of 16. To minimize queuing delay, we set the transmission queue length to 6 packets.

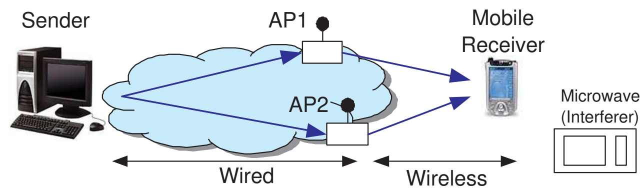

Fig. 1. Wireless video over 802.11 to a mobile receiver using path diversity provided by multiple access points.

The goal of our experiment is to compare the performance of low-latency video over single versus multiple access points. Ideally, we would like our results to be representative of the average performance experienced by a mobile wireless receiver such as a PDA, laptop, etc. However, the results are highly location-dependent; moving the receiver by only a few inches can cause a 10–20 dBm change in signal strength and over 50% change in packet loss rate. We address this problem by collecting a 15-minute packet trace from a mobile receiver that is pushed on a cart at walking speed. The mobile receiver remains within 15 meters of at least one AP (and at most 40 meters from the other) at all times. Thus, the packet trace represents a continuous sampling of streaming performance at different locations in range of the APs.

Another challenge is the uncontrolled interference in the environment. Ideally, we would like to collect performance statistics for streaming over a single access point in one trial, repeat the experiment using two access points in a separate trial, then compare the two. However, spurious interference—generated by the daily activities of the building occupants—changes with every trial. Our solution is to use the same 15-minute packet trace to generate five sub-sampled traces representing different transmission schemes. Because the sub-sampled traces are generated from the same packet trace, they all experience the same interference pattern, which then allows us to make fair comparisons.

We generate the five sub-sampled traces as follows. Recall that the original 15-min packet trace contains constant packet rate streams from both access points. First, we generate a sub-sampled trace, *AP1*, by selecting only the packets streamed from access point AP1 in the 15-min packet trace. Thus, *AP1* represents a scenario with 30 packets/s sent from AP1 and 30 packet/s cross traffic sent from AP2. Likewise, *AP2* represents a scenario where the video is streamed from AP2 while the cross traffic is sent from AP1. The third sub-sampled trace, *Balanced*, represents a scenario where two access points stream a 30 packet/s video, each access point simultaneously transmitting exactly one half the stream at 15 packets/s. This is a "dumb" path diversity approach since it uses AP1 and AP2 equally, regardless of their respective link conditions. We generate this trace by selecting every other packet from each of the two streams from our original 15-min packet trace.

The fourth trace, *Site Selection*, represents a simple site selection algorithm that adaptively selects the access point with the lowest error rate. In detail, the preferred AP supplies 95% of the packets, while the other AP supplies the remaining 5% (for probing). The selection is recomputed at each time using the reception statistics from the 300 previous packets. The final trace, *Oracle*, is included to provide a bound on the performance of any site selection algorithm. We generate this trace by declaring a packet to be received if it is received from either access point. Note that the performance of *Oracle* is equal to that of repetition coding, in which the same packets are sent from both APs (but at the cost of twice the channel occupancy).

Figure 2 shows the performance of access points AP1 and AP2. The top plot shows the evolution of the received packet signal strength (RSS). There are times where the signal quality of AP1 is much better than AP2 and vice versa. The next two plots graph the average packet loss rate (PLR_i) and the number of loss events (L_i^{Burst}) of burst length $\geq$ 2 for traces *AP1* and *AP2*. Each statistic is computed over an interval i of 150 packets (5 secs). When comparing the plots we see a high correlation among RSS, PLR_i, and L_i^{Burst}. The high PLR_i values are surprising, as the maximum retry limit is 16 and the streams are sent at a rate of 360 kb/s, which is well below the saturation rate of approximately 6 Mb/s in 802.11b [1]. We have verified that packets are rarely dropped at the access points' queues; thus, almost all the losses are caused by wireless transmission. The bottom plot shows the normalized one-way delay for each packet in the trace (normalized by subtracting the minimum delay over all packets). Due to spurious interference and the large transmission retry limit, there often exist delays much greater than 150 ms, an unacceptable value in many low-latency communication applications.

Results for the traces *Balanced*, *Site Selection*, and *Oracle* are displayed in Table 1, Table 2, and Figure 3. The single access point results *AP1* and *AP2* are included for comparison. To model the sensitivity of applications to delay, we introduce a delay threshold D_{thr}. If a packet experiences a one-way delay above D_{thr} it is assumed lost.

Referring to Table 1 and the upper plots in Figure 3, we see that *AP1* and *AP2* have similar performance. As expected, the packet loss rate PLR of *Balanced* is the approximate average of *AP1* and *AP2*, but the number of burst events is much lower. In *Site Selection*, we see significant reduction in PLR compared to *Balanced*, while the number of burst events falls between *Balanced* and *AP1*/*AP2*. Evidently our simple site selection algorithm does not adapt quickly enough to avoid bursts. The *Oracle* trace performs best of all. This highlights the maximum gain that may be achieved by distributing packets adaptively across access points.

We next examine the application-layer performance in terms of reconstructed video quality. Video sequences are compressed using JM 2.0 of the emerging H.264/MPEG-4 AVC video compression standard and are appropriately framed into packets which are sent as RTP/UDP/IP. We use four standard video test sequences in QCIF format: *Foreman*, *Claire*, *Mother-Daughter*, and *Salesman*. Each has 300 frames at 30 frames/s, and is coded with a constant quantization level. The average PSNR and the bit rate for each sequence are given in the column headings of Table 2. The first frame of each sequence is coded as an I-frame, and all subsequent frames are coded as P-frames. To improve error-resilience, a slice in every 4^{th} frame is intra updated, corresponding to an intra update every $N = 4 \times 9 = 36$ frames. Both the packet framing and the intra update are as recommended in JM 2.0. Each

Fig. 2. Performance statistics for 15-minute packet trace, where $+$ and $\circ$ indicate *AP1* and *AP2*. Received packet signal strength (top plot), average packet loss rate (second), number of loss events of burst length $\geq$ 2 (third), and one-way delay (bottom), all as a function of packet sequence number. In the second and third plots, each point represents an interval of 150 packets (5 secs). In the third plot $\bigtriangledown$ indicates *Balanced*.

P-frame fits within a 1500-byte payload, i.e., within a single transmitted packet, while the first I-frame requires a number of packets. The transmission of each video sequence is simulated by assuming the packet loss patterns of the five traces *AP1*, *AP2*, *Balanced*, *Site Selection*, and *Oracle*. Each 10 sec video sequence is replicated 90 times to span the 15 minute packet trace. This also corresponds to an I-frame every 10 secs.

The measured distortion for various video sequences, transmission schemes, and delay thresholds are given in Table 2 and the bottom two plots of Figure 3. The average PSNR over the entire 15 min test is displayed. In addition, Table 2 gives the number of times the average PSNR over a 10-second window (period between I-frames) drops below 30.0 dB (28.0 dB for Foreman). This latter metric provides an indication of the frequency of objectionable events during the trace.

Comparing Tables 1 and 2, the primary cause of improvement in average PSNR is reduced PLR. The *Site Selection* trace improves the average PSNR relative to the conventional approach of using only a single access point (either AP1 or AP2) by 1.6–3.0 dB. The *Balanced* trace in addition indicates that reducing the occurrence of burst loss events—even at the same PLR—has a secondary but still significant beneficial effect, ranging from 0.1 to 1.7 dB (consistent with prior work that showed that burst losses produce greater total distortion in the reconstructed video than an equal number of isolated losses [10, 4].)

Based on the results in Tables 1 and 2, at the same D_{thr}, the *Site Selection* trace provides a reduction in PLR by 2–4.5% and an improvement in PSNR_{Ave} of 1.6–3.0 dB as compared to the conventional case of using a single AP. Alternatively, as shown by the arrows in Figure 3, for the same PLR and PSNR_{Ave}, *Site Selection* allows a reduction in the required delay threshold by about one-third, e.g. from 100ms to 60ms, thereby improving interactivity.

Scheme	Delay Threshold (D_{thr})								
	40ms			80ms			∞		
	PLR	PLR$_B$	Bursts	PLR	PLR$_B$	Bursts	PLR	PLR$_B$	Bursts
AP1	16.41	11.97	956	8.32	4.98	386	6.56	2.86	288
AP2	18.20	13.29	1074	9.00	5.32	415	7.00	3.01	323
Balanced	17.19	5.48	545	8.58	1.45	131	6.70	0.63	67
Site Selection	13.89	9.01	818	6.00	2.74	243	4.58	1.35	144
Oracle	3.73	2.06	184	0.92	0.59	38	0.26	0.13	9

Table 1. Average packet loss rate PLR, average burst loss rate PLR$_B$, and number of burst loss events Bursts. The PLR$_B$ is the ratio of packets lost in bursts (loss episodes of length ≥ 2) to the total number of packets transmitted.

Scheme	Sequence							
	Foreman (35.8 dB at 156.2 kb/s)		Claire (39.6 dB at 39.2 kb/s)		Mother & Daughter (36.2 dB at 68.6 kb/s)		Salesman (34.9 dB at 67.6 kb/s)	
	PSNR$_{\text{Ave}}$	N_{thresh}	PSNR$_{\text{Ave}}$	N_{thresh}	PSNR$_{\text{Ave}}$	N_{thresh}	PSNR$_{\text{Ave}}$	N_{thresh}
AP1	24.34	71	31.61	8	31.23	15	30.19	16
AP2	24.01	69	31.05	9	31.12	17	29.85	10
Balanced	24.40	82	32.78	6	31.81	9	30.76	7
Site Selection	25.77	58	34.11	2	32.79	4	32.03	3
Oracle	31.58	7	37.11	1	35.12	2	33.92	2

Table 2. Reconstructed video quality for four video sequences and five transmission traces, with D_{thr} = 80 ms: average PSNR (PSNR$_{\text{Ave}}$) and number of times (N_{thresh}) the PSNR falls below a threshold. The bit rate and (error-free) PSNR$_{\text{Ave}}$ are given in the column headings.

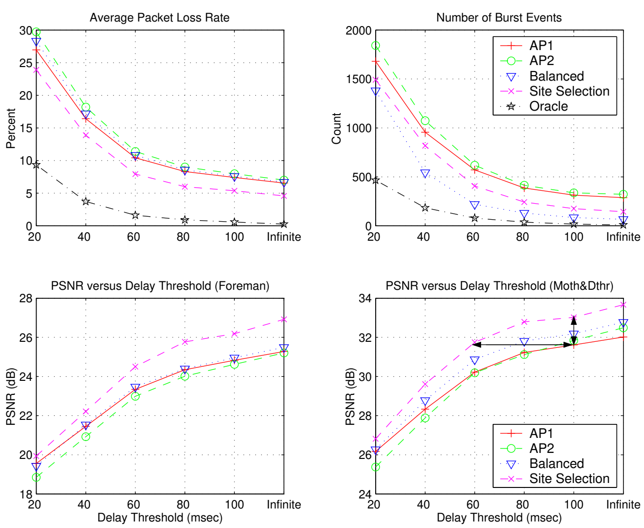

Fig. 3. Packet loss rate, number of burst loss events, and average PSNRs, as a function of delay threshold.

4. SUMMARY

Low-latency video communication over 802.11b wireless networks is challenging because of the highly variable nature of the wireless link. This paper proposed a system to tackle this problem which was composed of H.264/MPEG-4 AVC, low-latency best-effort transport mechanisms, and the use of path diversity between multiple access points and the mobile client. In particular, the use of path diversity from multiple access points can provide significant benefits compared to the conventional case when only a single access point is used. This was shown for both a non-adaptive scheme that alternates access points for each packet, and for a simple adaptive scheme that uses past error statistics to select an access point. The simple adaptive scheme provides significant improvements in PLR and PSNR$_{\text{Ave}}$ for the same delay threshold, or a sizable reduction in the required delay threshold to achieve the same PLR and PSNR$_{\text{Ave}}$. The upper limit on performance has been established using a diversity scheme based on an all-knowing Oracle; the extent to which this performance may be approached is a subject for future research.

5. REFERENCES

[1] L. Munoz M. G. Arranz, R. Aguero and P. Mahonen, "Behavior of UDP-based applications over IEEE 802.11 wireless networks," in *Proc. of PIMRC 2001*, San Diego, CA, Sept. 2001.

[2] D. Hertrich, "MPEG4 video transmission in wireless LANs: Basic QoS support on the data link layer of 802.11b," Minor Thesis, Technical University Berlin, 2002, http://www.daniel-hertrich.de/studienarbeit/.

[3] A. Majumdar, D. Sachs, I. Kozintsev, K. Ramchandran, and M. Yeung, "Multicast and unicast real-time video streaming over wireless LANs," *IEEE Trans. on Circuits and Systems for Video Technology*, June 2002.

[4] J.G. Apostolopoulos, "Reliable video communication over lossy packet networks using multiple state encoding and path diversity," *Visual Communications and Image Processing (VCIP)*, January 2001.

[5] Y.J. Liang, E.G. Steinbach, and B. Girod, "Real-time voice communication over the Internet using packet path diversity," *Proc. ACM Multimedia*, Sept/Oct 2001.

[6] J.G. Apostolopoulos, T. Wong, W. Tan, and S.J. Wee, "On multiple description streaming with content delivery networks," *IEEE INFOCOM*, June 2002.

[7] N. Gogate, D. Chung, S.S. Panwar, and Y. Wang, "Supporting image/video applications in a mobile multihop radio environment using route diversity and multiple description coding," *IEEE Trans. on Circuits and Systems for Video Technology*, September 2002.

[8] ITU-T and ISO/IEC, *Draft Text for Joint Video Specification (ITU-T Rec. H.264 and ISO/IEC 14496-10 AVC*, October 2002.

[9] Nitin H. Vaidya, Paramvir Bahl, and Seema Gupta, "Distributed fair scheduling in a wireless LAN," in *Proc. of 6th ACM MOBICOM*, Boston, MA, Aug. 2000.

[10] Y.J. Liang, J.G. Apostolopoulos, and B. Girod, "Analysis of packet loss for compressed video: Does burst-length matter?," *ICASSP*, April 2003.