Path Selection using Available Bandwidth Estimation in Overlay-based Video Streaming

Manish Jain and Constantine Dovrolis

College of Computing, Georgia Institute of Technology

Abstract. IP networks present a challenging environment for video streaming because they do not provide throughput, jitter, or loss rate guarantees. In this work, we focus on improving the perceived quality of video streaming through dynamic path selection. Selecting one of several Internet paths is possible using multihoming and/or an overlay routing infrastructure. We conduct an experimental comparison of various measurement-based path selection techniques for video streaming. The path selection is based on the measurement of network-layer metrics, such as loss rate, jitter or available bandwidth, while the video quality is evaluated based on the VQM tool. Our experiments show that the most effective technique for adaptive path selection relies on an estimate of the lower bound of the available bandwidth variation range. We show how to perform such measurements using the video packets, eliminating the measurement overhead in the selected path. Finally, we show that adaptive path selection is more effective than a simple, but commonly used, form of FEC.

1 Introduction

As the "last mile" access capacity continues to grow, IP video streaming becomes more popular among users and content providers. Many experts believe that IPTV is the next "killer-application" in the Internet [7]. However, supporting video streaming and IPTV presents significant challenges. IP networks often suffer from several network impairments, including packet losses, significant jitter and one-way delays, as well as outages of unpredictable duration. Additionally, most IP networks today do not offer deterministic or statistical QoS guarantees.

Since the early nineties, several approaches for adaptive video streaming applications have been proposed. One approach is to adjust the encoding scheme and/or video frame rate in response to changes in the network state [11, 23, 4]. The main drawback of such schemes, however, is that the perceived video quality varies with time, causing user dissatisfaction. Another class of approaches is to use proactive error correction techniques, such as Reed-Solomon FEC codes, or to retransmit lost packets through standard ARQ schemes [18, 13, 12]. The major drawback of FEC schemes is that they introduce bandwidth overhead even when the network does not drop packets. The drawback of retransmissions is that they require a playback delay of a few round-trip times, additional

state/buffering at the sender, and reverse-path traffic (e.g., negative ACKs). Another approach is to mask the effect of lost or discarded (i.e., late) packets through the use of codec-specific error concealment techniques [19, 14]. The effectiveness of such techniques is limited however.

Even though IP networks typically use a single path from one host to another, the recent popularity of multihoming and overlay networks allows content providers to choose between several network paths towards a given receiver [26]. Such path diversity gives video streaming one more adaptation option: to dynamically switch from one path to another depending on the observed (or predicted) performance in the candidate paths. This technology uses network-level measurements, such as loss or jitter, and it has been shown that it can quickly react to congested paths or outages [22, 2]. A variation of this approach is to combine path diversity with Multi-Description Coding (MDC) techniques [5, 3], and use multiple paths simultaneously. The studies in this area rely on loss rate, delay or TCP throughput measurements, and they typically perform these measurements using "dummy" probing packets.

In this work, we consider an overlay-based video streaming architecture in which the objective is to maximize the perceived video quality through dynamic overlay path selection. A novel aspect of our study is that the network measurements that drive the path selection process rely on available bandwidth (avail-bw) estimation. The avail-bw of a network path is defined as the residual capacity at the path's bottleneck, and so it represents the maximum additional load that the path can carry before it becomes saturated [9]. The reason we focus on avail-bw is because this metric can determine whether a path has enough capacity to carry a video stream *before* we switch the stream to that path. Other network-layer metrics, such as jitter or packet loss rate, can only determine whether a path is *already* congested, causing degradation in the video quality at the receiver [20, 21, 6]. We also show how to modify an existing avail-bw estimation technique, described in [9], so that the measurements are performed using application packets, rather than "dummy" probing packets, eliminating the measurement overhead in the currently selected path. We evaluate the video quality based on the VQM technique described in the ITU-T recommendation J.144 [1]. It has been shown that VQM is superior to other video quality metrics, such as PSNR, because the VQM score is more representative of the user-perceived video quality [15]. With a series of repeatable experiments in a controlled environment we compare the VQM score of path selection schemes based on jitter, loss rate, and various percentiles of the avail-bw distribution. The main result of this experimental study is that performing path selection based on *the estimated lower bound of the avail-bw variation range* performs significantly better in terms of VQM score, path switching frequency, and probability of aborting an ongoing video stream.

The rest of the paper is organized as follows. §2 presents an overlay-based video streaming architecture. §3 describes the path selection techniques that we evaluate. The experimental methodology is described in §4, and the results are

presented in §5. §6 gives a brief comparison between adaptive path selection and a simple (but commonly used) FEC scheme. We conclude in §7.

2 VDN architecture and in-band measurements

In this section, we present the high-level architecture of an overlay-based video streaming architecture, referred to as the *Video Distribution Network (VDN)*. A content provider constructs a VDN by deploying several overlay nodes that will act as either overlay ingress/egress nodes or as intermediate nodes (see Figure 1). Each VDN node runs a *Measurement Module (M-module)* to measure the performance (jitter, loss rate, avail-bw) of the overlay links from that VDN node to its neighbors. The VDN also runs a link-state protocol so that each node is aware of the latest state in all VDN links. We only consider VDN paths with at most one intermediate overlay node, based on the results of [26]; additional intermediate nodes are rarely needed in practice. The path of a video stream is determined at the ingress VDN node, i.e., we use source routing. The egress VDN node removes any VDN headers and delivers the stream to the receiver.
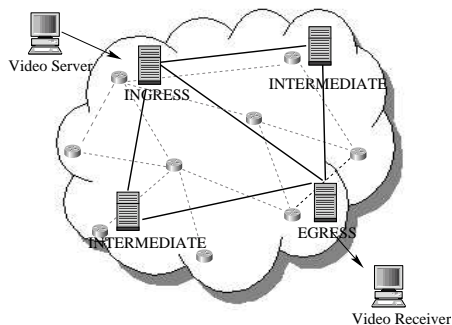

Fig. 1. VDN architecture.

For the purposes of this paper, the most important aspect of the VDN architecture is the M-module. In our implementation, the M-Module relies on active measurement to estimate the packet loss rate, jitter, and the avail-bw variation range at a given overlay link. The loss rate is measured as the fraction of lost packets in a 1-Mbps stream of 1500-byte packets. The jitter is estimated as the maximum absolute difference between the spacing of consecutive packets at the receiver relative to their spacing at the sender, using the same probing stream. The avail-bw variation range is measured as described in our earlier work [9, 10].

One feature of the M-Module is that it uses the application's video packets to perform *in-band* network measurement. The in-band approach eliminates the measurement overhead, at least in the path that is currently selected for video

streaming (we still use out-band measurements with empty probing packets in paths that do not transfer any video streams). Using video packets for the estimation of loss rate or jitter is relatively straightforward. The estimation of avail-bw, on the other hand, requires shaping the video stream at a different rate than the transmission rate at the sender. In the following paragraph we describe how the M-module shapes the video stream to a particular rate at the input of each overlay link.
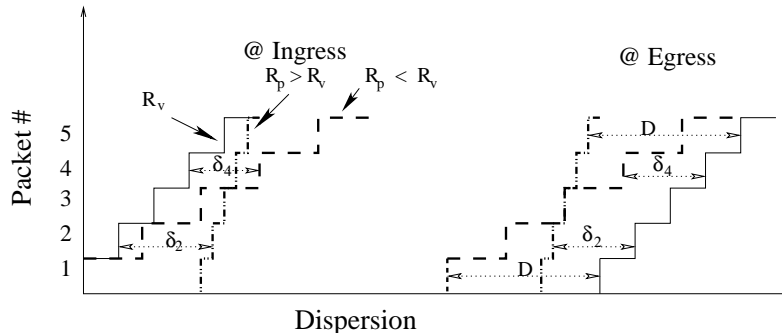
Fig. 2. Shaping of video packets to a probing rate R_p for avail-bw estimation.

Suppose that N video packets of size L arrive at a VDN node at a rate R_v. The arrival time of the ith packet is t_i^a. The objective of the local M-module is to shape those N packets at a probing rate R_p, so that it can measure whether R_p is smaller or larger than the avail-bw in the outgoing overlay link. To do so, the N incoming packets are delayed so that their output spacing is L/R_p. If t_i^d is the departure time of the ith packet, then $t_i^d - t_{i-1}^d = L/R_p$ and $t_i^d = t_i^a + \delta_i$, where δ_i is the delay introduced to the ith packet. For instance, if $R_p > R_v$, the M-module shapes the packet stream by introducing a decreasing amount of delay in successive packets; otherwise, the M-module introduces an increasing amount of delay in successive packets (see Figure 2). The M-module adds the value δ_i into the VDN header of packet i. In some cases, the receiving application may demand that the video stream arrives at the rate R_v in which it was sent to (e.g., for clock synchronization). In that case, the egress VDN node can reshape the video stream rate back to the initial transmission rate at the sender, by delaying each packet by $D - \delta_i$, where D is the maximum cumulative delay that can be introduced to the packet from the M-modules at the ingress and intermediate nodes and δ_i is the corresponding cumulative delay that the ith packet experienced.

3 Path selection schemes

In this section, we describe the four path selection schemes we evaluate. The schemes are distinguished based on the choice of the key measured network performance metric.

Loss based path selection (LPS): In LPS, we monitor the average loss rate in all candidate paths during 3-second periods. The path with the minimum loss rate is selected. If the currently used path has zero loss rate, then we do not switch to another path even if there are other loss-free paths.

Jitter based path selection (JPS): The jitter of successive packet pairs is also measured over 3-second periods. The path with the minimum 90th percentile of jitter measurements is selected. If the minimum jitter is practically the same in more than one paths, then JPS selects the path with the lowest loss rate. If the loss rate is also equal, then JPS stays at the current path if that is one of the best paths, or it randomly picks one of the best paths otherwise.

Avail-bw based path selection (APS): This scheme has two variations. In the first, we use the average avail-bw (A-APS). In the second, we use the lower bound of the avail-bw variation range (L-APS) (see [10] for more details). A new avail-bw estimate results in almost every 3 seconds, similar to LPS and JPS. If the avail-bw estimate (average or lower bound) in the currently selected path is greater than twice the video transmission rate, then we stay in that path. Otherwise, we choose the path with the highest avail-bw estimate; note that this may still be the currently selected path.

4 Experimental setup

We have evaluated the performance of the previous path selection schemes with controlled experiments in the testbed of Figure 3. The video stream is transmitted from node B to E through either the direct path, or through the path that traverses node D. The video stream is a 2-minute clip formed by combining SMPTE test video sequences [8] and encoded in MPEG-2 with average rate 6Mbps. The VLC player [24] transmits the stream to the network. The stream is initially routed through the direct path. The M-modules run at nodes B, D, and E, and they perform both network measurement and the path selection process. Both paths carry cross traffic and they are occasionally congested. The cross traffic in each path is generated by replaying NLANR packet traces, collected from various university access links [16]. The average cross traffic rate is set to the desired value by scaling the packet interarrivals by the appropriate factor.

We compare the performance of various path selection algorithms based on three criteria: *video quality*, *user-abort probability* and *path switching frequency*. The video quality is measured using the VQM tool [25], which implements the

ITU-T J.144 recommendation [1]. VQM compares the original video stream with the received video stream, and it reports a metric between 0 and 1. Note that *a lower VQM scores correspond to better video quality*. It has been shown that the VQM score correlates very well with the user-perceived video quality (MOS score). The VQM software supports five models to evaluate video quality, described in detail in the NTIA Handbook [17]. In this work, we use the *television model*. In the following graphs we report the minimum, average, and maximum VQM score from the five runs of each experiment. The five runs differ in terms of the initial phase between the video clip and the cross traffic traces.
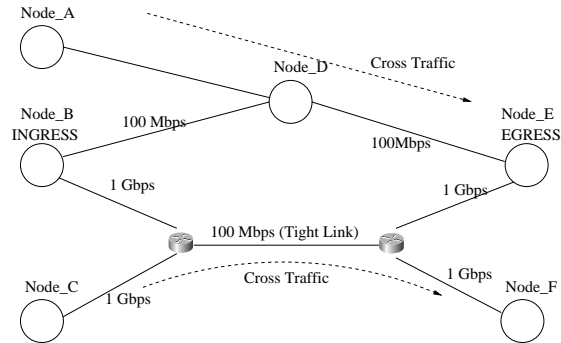

Fig. 3. Testbed.

The *user-abort probability* focuses on the short-term variations of the VQM score. The idea is that if the VQM score is too high (poor quality) during a time window, then the user would either abort the video stream or she would be unsatisfied. We measure the VQM score of consecutive 10-second video segments. A video stream is considered *aborted* if one of the following two conditions is met: either the VQM score in a single segment is larger than 0.55, or two consecutive video segments have VQM scores larger than 0.35. We chose these values based on extensive subjective tests of several video streams under different conditions. To estimate the user-abort probability, we measured the fraction of aborted video stream in 30 experiments.

The last evaluation metric is the total number of path switching events. Even though the path switching frequency does not affect the video quality, it is an important aspect of any dynamic routing mechanism from the network operator's perspective. Frequent path switching of large traffic volumes can affect the network stability and traffic engineering. Consequently, even though our primary interest is to optimize video streaming quality, we would also like to avoid unnecessary path switching.

5 Results

Figures 5, 6 and 7 show the performance of the considered path selection schemes under different load conditions (i.e., utilization of the bottleneck link in each path). The PSC trace shown in Figure 4 is replayed at the direct path, while the FRG trace is replayed at the indirect path. We adjust the average rate of each trace to achieve the desired utilization. Even though we set the long-term average rate of the two traces at the same value, there are time periods where one path is congested while the other is not. Note that there are time periods, mainly in higher load conditions, where both paths are congested. Obviously, path switching techniques cannot avoid congestion in that case, but they can still choose the least congested path.
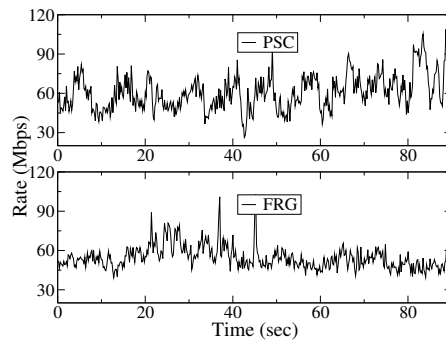

Fig. 4. Exmaples of stationary and non-stationary NLANR cross traffic traces.
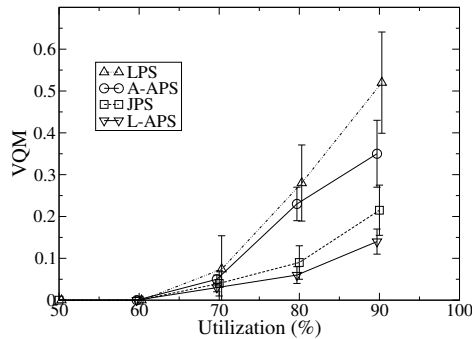
Fig. 5. VQM scores for the four path selection schemes.

Figure 5 shows the overall VQM score for each path selection scheme. LPS clearly performs poorly since it only reacts after congestion has affected the currently selected path. The A-APS scheme does not perform much better than

LPS. The reason is that the average avail-bw does not capture the variability of the avail-bw distribution, and so the two paths appear as almost equally good in most of the time. The JPS and L-APS schemes have comparable performance and they are clearly better than A-APS and LPS. This is because both JPS and L-APS are able to detect the onset of queuing delays in the currently selected path, before that path becomes congested. Note that L-APS is slightly better than JPS, especially in the case of 90% utilization.
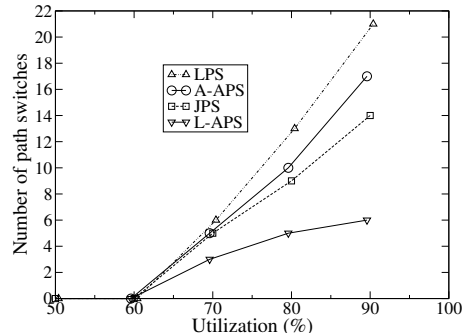
Fig. 6. Path switching frequency.

Figure 6 shows the number of path switching events in the same set of experiments. The main observation is that L-APS has the lowest path switching frequency. JPS causes significantly more path changes, and is comparable to A-APS. This is because JPS relies on a comparison of the maximum jitter in the two paths, and so a minor variation in the jitter, which can result from an short-lived cross-traffic burst, may trigger JPS to switch paths. Instead, L-APS does not switch paths if the currently selected path provides a large safety margin, in terms of avail-bw, for the given video stream rate.

Figure 7 shows the user-abort probability, i.e., the fraction of aborted video streams. The ranking of the four path selection schemes is as in the case of the long-term VQM score in Figure 5.

We next show some results for the traces shown in Figure 8. These traces include instances of traffic non-stationarity. In the PSC trace, the traffic rate varies abruptly between 20Mbps and 45Mbps during the second half (level shifts). On the other hand, the AMP trace exhibits periods of slowly increasing traffic load (notice the traffic "ramp"). We are interested to examine the effectiveness of the considered path selection schemes under such traffic conditions. In these experiments, we use the same trace in both paths. In one path, the trace playback starts from the beginning, while in the other path the trace playback starts from the middle.

Figures 9 and 10 show the VQM scores for the PSC and AMP traces, respectively. Note that, overall, the level shifts of the PSC trace cause higher VQM scores compared to the smoother AMP trace. L-APS performs clearly
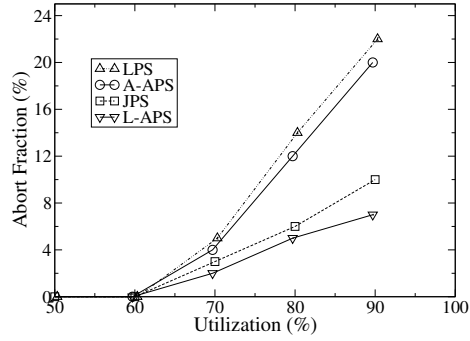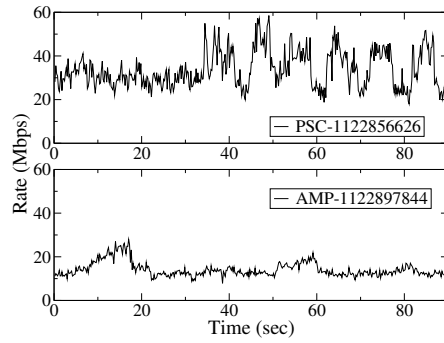
Fig. 7. User-abort probability.

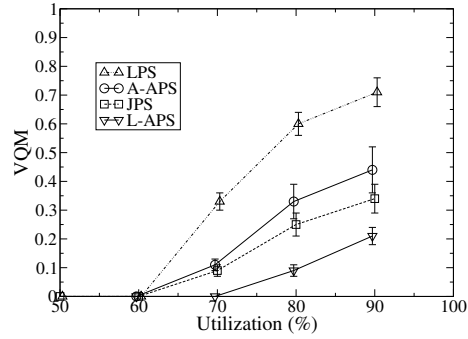**Fig. 8.** Cross traffic traces with instances of non-stationarity.

Fig. 9. VQM scores for the PSC trace.

better in this case than the JPS scheme. Note that there is a difference in the relative performance of JPS and A-APS in the two traces. The reason is that, even though JPS is more proactive than A-APS in switching paths, in the AMP trace the A-APS scheme performs slightly better because it can detect the slowly decreasing level of avail-bw during the traffic ramp.
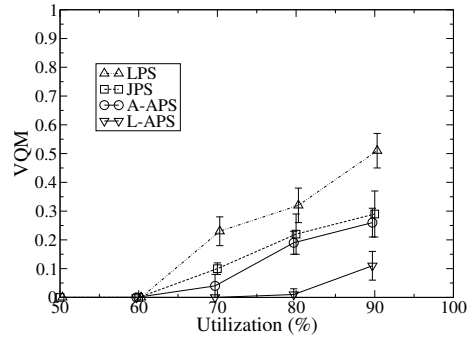
Fig. 10. VQM scores for the AMP trace.

6 Path switching versus FEC

In this section, we conduct a preliminary comparison between path switching techniques and FEC-based loss recovery. A commonly used FEC scheme is the Reed-Solomon (RS) code [18, 13, 12]. In an (n, k)-RS code, $n-k$ out of n packets carry FEC packets. An (n, k) RS-code can recover from all losses in a block of n packets if at least k of those n packets are received. The main drawback of FEC-based schemes is their transmission overhead $n - k/k$.
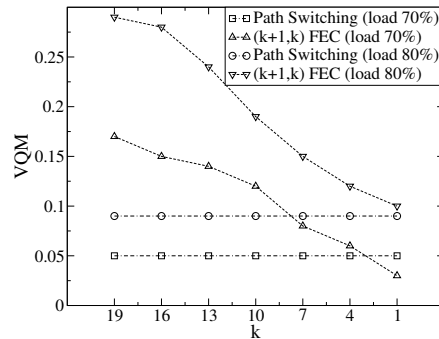

Fig. 11. Path switching versus FEC.

Here, we evaluate the simplest form of FEC in which $n = k+1$. This instance of RS-coding is equivalent to sending a single parity packet after every k data packets. We compare the VQM score of this technique with L-APS for two load conditions: 70% and 80% bottleneck utilization using the NLANR traffic trace (BWY-1063315231). Figure 11 shows the results for different values of k. At the far left, k=19 corresponds to a transmission overhead of about 5%, while at the far right, k=1 corresponds to 100% overhead. Note that path switching with L-APS performs consistently better than FEC, except the case of k=1. The main

reason is that path switching can often avoid congestion altogether, while FEC attempts to recover from the effects of congestion, which is not always possible. Additionally, the FEC scheme we evaluate here is not effective in dealing with the bursty nature of congestion-induced packet losses.

7 Summary

This work focused on the use of measurement-driven path selection techniques for video streaming. We showed that if the path selection is driven by a conservative estimate of the available bandwidth, then the resulting video streaming performance is significantly improved compared to other commonly used network metrics. An interesting open problem is to develop path switching mechanisms that are driven by direct video quality measurements at the receiver. Another open problem is to evaluate the performance of a hybrid approach using both FEC and path selection mechanisms.

Acknowledgement

References

1. Objective Perceptual Video Quality Measurement Techniques for Digital Cable Television in the presence of Full Reference. ITU-T Recommendation J.144 rev. 1, 2003.
2. Y. Amir, C. Danilov, S. Goose, D. Hedqvist, and A. Terzis. 1-800-OVERLAYS: Using Overlay Networks to Improve VoIP Quality. In *Proceedings of NOSSDAV*, 2005.
3. J. Apostolopoulos, T. Wong, W. Tan, and S. Wee. On Multiple Description Streaming with Content Delivery Networks. In *Proceedings of INFOCOM*, 2002.
4. A. Balk, D. Maggiorini, M. Gerla, and M. Y. Sanadidi. Adaptive MPEG-4 Video Streaming with Bandwidth Estimation. In *Proceedings of QoS-IP*, 2003.
5. A. Begen, Y. Altunbasak, O. Ergun, and M. Ammar. Multi-path Selection for Multiple Description Video Streaming over Overlay Networks. *Signal Processing: Image Communication*, 20:39–60, 2005.
6. J. Boyce and R. Gaglianello. Packet Loss Effects on MPEG Video Sent Over the Public Internet. In *Proceedings of Multimedia*, 1998.
7. S. Cherry. The Battle for Broadband. *IEEE Spectrum*, 42(1):24–29, Jan. 2005.
8. C. Fenimore. Mastering and Archiving Uncompressed Digital Video Test Materials. In *Proceedings of 142nd SMPTE Technical Conference*, 2000.

9. M. Jain and C. Dovrolis. End-to-End Available Bandwidth: Measurement Methodology, Dynamics, and Relation with TCP Throughput. *IEEE/ACM Transactions on Networking*, 11(4):537–549, Aug. 2003.
10. M. Jain and C. Dovrolis. End-to-end Estimation of Available Bandwidth Variation Range. In *Proceedings of SIGMETRICS*, June 2005.
11. K. Jeffay, D. L. Stone, T. Talley, and F. D. Smith. Adaptive, Best-Effort Delivery of Digital Audio and Video Across Packet-Switched Networks. In *Network and Operating System Support for Digital Audio and Video*, 1993.
12. W. Jiang and H. Schulzrinne. Comparison and Optimization of Packet Loss Repair Methods on VOIP Perceived Quality under Bursty Loss. In *Proceedings of NOSSDAV*, 2002.
13. L. Kontothanassis, R. Sitaraman, J. Wein, D. Hong, R. Kleinberg, B. Mancuso, D. Shaw, and D. Stodolsky. A Transport Layer for Live Streaming in a Content Delivery Network. *Proceedings of IEEE*, 92, 2004.
14. W. M. Lam and A. Reibman. An Error Concealment Algorithm for Images Subject to Channel Errors. *IEEE Transactions on Image Processing*, 4(5):533–542, May 1995.
15. X. Lu, S. Tao, M. E. Zarki, and R. Guerin. Quality-Based Adaptive Video Over the Internet. In *Proceedings of CNDS*, 2003.
16. NLANR MOAT. Passive Measurement and Analysis. http://pma.nlanr.net/PMA/, 2006.
17. M. Pinson and S. Wolf. NTIA HB 06434: In-Service Video Quality Metric Users Manual. http://www.its.bldrdoc.gov/pub/ntia-rpt/06-434/, 2005.
18. I. Reed and G. Solomon. Polynomial Codes over Certain Finite Fields. *Journal of the Society for Industrial and Applied Mathematics*, 8(2), 1960.
19. P. Salama, N. Shroff, E. Coyle, and E. Delp. Error Concealment Techniques for Encoded Video Streams. In *Proceedings IEEE International Conference on Image Processing*, pages 9–12, 1995.
20. W. Tan and A. Zakhor. Real-Time Internet Video Using Error Resilient Scalable Compression and TCP-Friendly Transport Protocol. *IEEE Transactions on Multimedia*, 1(2), 1999.
21. S. Tao, J. Apostopoulos, and R. Guerin. Real-Time Monitoring of Video Quality in IP Networks. In *Proceedings of NOSSDAV*, 2005.
22. S. Tao and R. Guerin. Application-specific Path Switching: A Case Study for Streaming Video. In *Proceedings of ACM International Conference on Multimedia*, 2004.
23. B. Vandalore, W. chi Feng, R. Jain, and S. Fahmy. A Survey of Application Layer Techniques for Adaptive Streaming of Multimedia. *Real-Time Imaging*, (3), 2001.
24. VLC Media Player. http://www.videolan.org/vlc, 2006.
25. S. Wolf. VQM Software. http://www.its.bldrdoc.gov/n3/video/vqmsoftware.htm, 2006.
26. Y. Zhu, C. Dovrolis, and M. Ammar. Dynamic Overlay Routing Based on Available Bandwidth Estimation: A Simulation Study. *Computer Networks Journal*, 2006.