RATE-DISTORTION OPTIMIZED STREAMING OF AUTHENTICATED VIDEO

Zhishou Zhang[1,2], *Qibin Sun*[1], *Wai-Choong Wong*[1,2], *John Apostolopoulos*[3] *and Susie Wee*[3]

[1]Institute for Infocomm Research, Singapore
[2]Department of ECE, National University of Singapore, Singapore
[3]Hewlett-Packard Labs, USA

ABSTRACT

Stream authentication methods usually impose overhead and dependency among packets. The straightforward application of state-of-the-art rate-distortion (R-D) optimized streaming techniques produce highly sub-optimal R-D performance for authenticated video, since they do not account for the additional dependencies. This paper proposes an R-D optimized streaming technique for authenticated video, by accounting for authentication dependencies and overhead. It schedules packet transmission based on packets' importance in terms of both video quality and authentication dependencies. The proposed technique works with any stream authentication method as long as the verification probability can be quantitatively computed from packet loss probability. Simulation results based on H.264 JM 10.1 and NS-2 demonstrate that the proposed authentication-aware R-D optimized streaming technique substantially outperforms authentication-unaware R-D optimized streaming techniques. In particular, when the channel capacity is below the source rate, the PSNR of authenticated video quickly drops to unacceptable levels using conventional R-D optimized streaming techniques, while the proposed technique still maintains R-D optimized video quality.

Index Terms— ***Video streaming, Authentication, R-D Optimization***

1. INTRODUCTION

Video streaming applications are becoming increasingly popular and important, which is evident by the emerging commercial services like movie-on-demand, video conference, video surveillance, and so on. However, the security issues, like integrity, source authentication, confidentiality, and secure adaptation are serious concerns [1]. This paper deals with integrity and source authentication issues.

A recent advance in media streaming is the Rate-Distortion Optimized (*RaDiO*) [2] streaming technique, which takes into account packet importance and knowledge about channel using Lagrangian cost function. It computes a packet transmission schedule that minimizes the expected end-to-end distortion subject to a constraint on the average transmission rate. The performance improvement of *RaDiO* over heuristic streaming techniques is significant, and low-complexity versions of *RaDiO* are being developed for video streaming, e.g. [3].

Authenticated video is the decoded video that results from packets which are *both received and verified*. A packet which is received but not verified is discarded. For stream authentication, common approaches [4]-[8] are to amortize a signature among a group of packets in order to reduce overhead and complexity. The packets are connected as a directed acyclic graph, where nodes correspond to packets and edges correspond to hash links. An edge from packet *A* to *B* is implemented by appending *A*'s hash to *B*. The graph has only one packet carrying signature, and each node has at least one directed path to the signature packet. At the receiver, lost packets are removed from the graph, and a packet is verifiable if it has at least one path to the signature packet. Fig. 1 gives an example to illustrate the basic idea. Note that the authentication graph imposes inter-dependency among the packets, e.g. in Fig. 1, packet P_3 depends on P_1 and P_{sig}, and packet P_2 depends on P_{sig}, P_0 and P_1 for verification. For instance, if P_{sig} is lost, all other packets will not be verifiable.

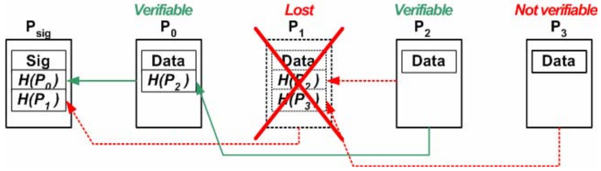

Fig. 1 – An example of authentication graph

Stream authentication changes both the relative packet importance (due to authentication dependencies) and packet sizes (due to authentication overhead). Existing R-D optimized streaming techniques, like *RaDiO* [2] [3], do not consider these issues. These streaming techniques can be referred to as ***authentication-unaware***, and they cannot achieve optimal performance for authenticated video where only the received packets which are verifiable are decoded. To solve this problem, we need ***authentication-aware*** streaming techniques, which take into account the authentication dependencies and overheads.

This paper proposes an *authentication-aware* technique to enable R-D optimized streaming for authenticated video. It works with any authentication method as long as the verification probability can be quantitatively computed from loss probability. We use butterfly authentication method [4] to illustrate the idea in this paper. Being *authentication-aware*, it accounts for authentication dependencies and overhead, in addition to the original packet importance and packet size. As such, we are able to formulate the R-D optimization problem, of minimizing the expected distortion of the authenticated video at the receiver.

This paper is organized as follows. Section 2 gives an overview of the conventional *RaDiO* streaming technique for un-authenticated video; Section 3 analyzes the authentication dependencies and overhead, followed by the description of the R-D optimized streaming technique for authenticated video. Section 4 validates the proposed technique with simulation results. The paper is concluded in Section 5.

2. OVERVIEW OF RATE-DISTORTION OPTIMIZED STREAMING

The *RaDiO* streaming techniques (e.g., [2]) assume a compressed media stream that has been assembled into packets. Each packet is associated with the packet size *B*, deadline *T* and distortion increment Δd. For instance, if a packet P_l is received before its deadline T_l, the overall distortion will be reduced by Δd_l. Here the distortion model is assumed to be additive, meaning that the inter-dependence between the effects of lost packets is ignored. Although this is not necessarily true for bursty loss, this model can accurately estimate the distortion when the lost packets are spaced sufficiently far apart from each other with respect to the intra-refresh period.

In streaming scenario with sender-driven re-transmission, packet P_l is assigned with transmission policy π_l, dictating whether or not P_l will be sent at every transmission opportunity before its deadline T_l. Associated with π_l are the cost function $\rho(\pi_l)$ and the error function $\varepsilon(\pi_l)$, where $\rho(\pi_l)$ is the expected number of transmissions and $\varepsilon(\pi_l)$ is the probability that it is received before T_l. Given a group of *N* packets, the goal is to find the optimized policy $\pi = [\pi_0, \pi_1, \ldots, \pi_{N-1}]$ that minimizes the Lagrangian cost function

$$J(\pi) = D(\pi) + \lambda R(\pi) \tag{1}$$

In (1), λ controls the trade-off between the distortion $D(\pi)$ computed by (2) and rate $R(\pi)$ computed by (3).

$$D(\pi) = D_0 - \sum_{l=0}^{N-1} \Delta d_l (1 - \varepsilon(\pi_l)) \tag{2}$$

$$R(\pi) = \sum_{l=0}^{N-1} B_l \rho(\pi_l) \tag{3}$$

The optimization problem can be solved in an iterative manner. Each iteration searches for optimal policy for only one packet, keeping the policy fixed for the rest of the packets. This is repeated until the Lagrangian cost converges. For instance, at certain iteration, the optimal policy for packet P_l is given by:

$$\pi_l^* = \arg\min_{\pi_l} \quad \varepsilon(\pi_l) + \lambda_l \rho(\pi_l) \tag{4}$$

The new multiplier, $\lambda_l = \lambda B_l / \Delta d_l$, is determined by λ, B_l and Δd_l. A packet will have more opportunities to be sent if it has smaller size and larger distortion increments. This technique [3] has much lower complexity than the original *RaDiO* [2], because the distortion increment is defined to be the total distortion caused by loss of a packet, which implicitly accounts for decoding dependency and error concealment.

3. RATE-DISTORTION OPTMIZED STREAMING OF AUTHENTICATED VIDEO

This section describes the proposed R-D optimized streaming technique for authenticated video. First, we analyze the authentication dependency and overhead. Second, we describe how to achieve R-D optimized performance by accounting for authentication dependency and overhead.

3.1 Video Authentication

The authentication dependency can be interpreted as how much the loss of one packet will affect the verification (and therefore the decoding) of the others. For instance, in Fig. 1, given P_1 is received, the verification probability of P_2 is $(1\text{-}\varepsilon_{sig})$; given P_1 is lost, P_2's verification probability is reduced to $(1\text{-}\varepsilon_{sig})(1\text{-}\varepsilon_0)$. Thus, for P_2, the significance of P_1 is the difference, $(1\text{-}\varepsilon_{sig})\varepsilon_{0.}$ Similarly, for P_3, the significance of P_1 is $(1\text{-}\varepsilon_{sig})$. In total, the authentication importance of P_1 is $(1\text{-}\varepsilon_{sig})\varepsilon_0(1\text{-}\varepsilon_2)\Delta d_2+(1\text{-}\varepsilon_{sig})(1\text{-} \varepsilon_3)\Delta d_{3.}$ In this way, we can compute packets' authentication importance, which will be used together with original packet importance and size in the R-D optimization problem. However, some authentication methods, like EMSS [7] and Augmented Chain [8], do not allow this computation due to their complex graph structure. Our proposed technique can use Simple Hash Chain [6], Tree-authentication [5] or butterfly authentication [4]. Nevertheless, other stream authentication methods can also be used as long as packets' authentication importance can be quantitatively computed from loss probability of other packets. As Simple Hash Chain is not robust against packet loss and Tree-authentication has too much authentication overhead, we use butterfly authentication method to illustrate the idea. Readers are referred to [4] for more details and discussion of butterfly stream authentication.

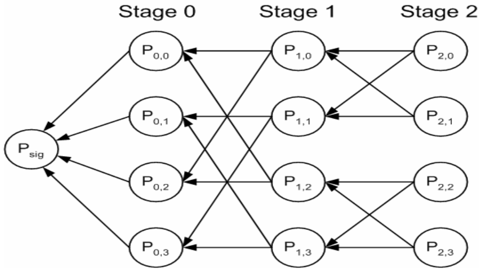

Fig. 2 – An example butterfly authentication graph

Fig. 2 gives an example butterfly authentication graph with 13 packets. Given such a graph with $N = M(\log_2 M+1)+1$ packets, the signature packet is denoted by P_{sig}, and other packets are denoted by $P_{s,j}$, where *s* indicates the stage, and *j* indicates the packet within stage. P_{sig} carries the signature and hashes of all packets in stage 0, and packet $P_{s,j}$ has its hash appended to $P_{s-1,j}$ and $P_{s-1,k}$, where *k* and *j* are $\log_2 M$-bit numbers differing at only $(s\text{-}1)^{th}$ most significant bit. As such, $P_{s,j}$ is verifiable if either $P_{s-1,j}$ or $P_{s-1,k}$ is received and authenticated. Thus, its verification probability $V_{s,j}$ can be expressed in (5), assuming P_{sig} is always received, and $\varepsilon_{s,j}$ is loss probability of $P_{s,j}$.

$$V_{s,j} = \begin{cases} \begin{pmatrix} (1-\varepsilon_{s-1,j})V_{s-1,j} + (1-\varepsilon_{s-1,k})V_{s-1,k} - \\ (1-\varepsilon_{s-1,j})V_{s-1,j}(1-\varepsilon_{s-1,k})V_{s-1,k} \end{pmatrix} & 0 < s \le \log_2 M \\ 1 & s = 0 \end{cases} \tag{5}$$

By repeatedly applying (5), $V_{s,j}$ can be expressed with loss probabilities of all packets in {$P_{s',j'} \mid 0 \le s' < s$, *Path exists from* $P_{s,j}$ *to* $P_{s',j'}$}, which is referred to as the *determining set* $A(P_{s,j})$.

On the other hand, the loss probability of $P_{s,j}$ affects the verification probability of all packets in $\{P_{\hat{s},\hat{j}} \mid s < \hat{s} \le \log_2 M$, *Path exists from* $P_{\hat{s},\hat{j}}$ *to* $P_{s,j}\}$, which is referred to as the *determined set* $B(P_{s,j})$. For example, in **Fig. 2**, $A(P_{1,1})=\{P_{0,1}, P_{0,3}\}$ and $B(P_{1,1})=\{P_{2,0}, P_{2,1}\}$.

In butterfly authentication graph, for any packet $P_{\hat{s},\hat{j}}$ in $B(P_{s,j})$, its verification probability $V_{\hat{s},\hat{j}}$ is a linear function of $\varepsilon_{s,j}$, where a and b are positive numbers. Its proof is given in [11].

$$V_{\hat{s},\hat{j}} = -a\varepsilon_{s,j} + b \qquad (6)$$

Using the butterfly method, the authentication overhead of P_{sig} is $O_{sig}=g+Mh$, where g and h are signature size and hash size, respectively. The overhead for the other packets is:

$$O_{s,j} = \begin{cases} 2h & 0 \le s < \log_2 M \\ 0 & s = \log_2 M \end{cases} \qquad (7)$$

So, the total overhead is $O=g+h(M+2M\log_2 M)$ when using butterfly authentication method.

3.2 Rate-Distortion Optimized Streaming

At the receiver, the authenticated video must be exclusively decoded from packets that are received and successfully authenticated before their deadlines. This definition allows packet loss and delay, but prevents packet alteration. Our goal is to compute the transmission policy that minimizes the distortion of the authenticated video by taking into account of authentication overhead and dependency.

Given a group of $N=M(\log_2 M+1)+1$ packets that are connected into a butterfly authentication graph, packet $P_{s,j}$ (P_{sig}) is associated with 4 quantities: packet size $B_{s,j}$ (B_{sig}), authentication overhead $O_{s,j}$ (O_{sig}), deadline $T_{s,j}$ (T_{sig}), and distortion increment $\Delta d_{s,j}$ (Δd_{sig}), and its transmission policy is $\pi_{s,j}$ (π_{sig}). To transmit the group of packets with policy $\pi = [\pi_{sig}, \pi_{0,0}, \pi_{0,1}, ..., \pi_{\log_2 M, M-1}]$, the expected transmission cost is computed by summing up the cost of individual packets, as shown in (8).

$$R(\pi) = (B_{sig} + O_{sig})\rho(\pi_{sig}) + \sum_{P_{s,j}} (B_{s,j} + O_{s,j})\rho(\pi_{s,j}) \qquad (8)$$

The expected distortion of authenticated video is computed by subtracting the total distortion from D_0, which is the distortion when no packet is decoded, as shown below.

$$D(\pi) = D_0 - (1-\varepsilon(\pi_{sig}))\left(\Delta d_{sig} + \sum_{P_{s,j}} \Delta d_{s,j}(1-\varepsilon(\pi_{s,j}))V_{s,j}\right) \qquad (9)$$

Substituting (8) and (9) into (1), we get the Lagrangian cost function.

$$J(\pi) = D_0 + \left(-(1-\varepsilon(\pi_{sig})\Delta d_{sig} + \lambda(B_{sig} + O_{sig})\rho(\pi_{sig})\right) + \sum_{P_{s,j}} \begin{pmatrix} -\Delta d_{s,j}(1-\varepsilon(\pi_{sig})(1-\varepsilon(\pi_{s,j})V_{s,j} \\ +\lambda(B_{s,j} + O_{s,j})\rho(\pi_{s,j}) \end{pmatrix} \qquad (10)$$

This optimization problem can be solved in an iterative manner, i.e. optimizing the policy for one packet at time, until $J(\pi)$ converges. For instance, the policy can be decided by (11) for P_{sig} and by (12) for $P_{s,j}$.

$$\pi^*_{sig} = \arg\min_{\pi_{sig}} \varepsilon(\pi_{sig}) + \lambda'_{sig}\rho(\pi_{sig}) \qquad (11)$$

$$\pi^*_{s,j} = \arg\min_{\pi_{s,j}} \varepsilon(\pi_{s,j}) + \lambda'_{s,j}\rho(\pi_{s,j}) \qquad (12)$$

where $\lambda'_{sig} = \lambda(B_{sig}+O_{sig})/S_{sig}$ and $\lambda'_{s,j} = \lambda(B_{s,j}+O_{s,j})/S_{s,j}$. The sensitivity factors, S_{sig} and $S_{s,j}$, can be obtained by taking partial derivative of $D(\pi)$ with respect to ε_{sig} and $\varepsilon_{s,j}$, respectively. Recall that for any packet $P_{\hat{s},\hat{j}} \in B(P_{s,j})$, $V_{\hat{s},\hat{j}}$ is linear function of $\varepsilon_{s,j}$. Therefore, the sensitivity factors can be computed as

$$S_{sig} = \Delta d_{sig} + \sum_{P_{s,j}} \left(\Delta d_{s,j}(1-\varepsilon(\pi_{s,j})V_{s,j}\right) \qquad (13)$$

$$S_{s,j} = (1-\varepsilon(\pi_{sig}))\left(\Delta d_{s,j}V_{s,j} + \sum_{P_{\hat{s},\hat{j}} \in B(P_{s,j})} \Delta d_{s,j}(1-\varepsilon(\pi_{s,j}))a^{\hat{s},\hat{j}}_{s,j}\right) \qquad (14)$$

where $V_{\hat{s},\hat{j}} = -a^{\hat{s},\hat{j}}_{s,j}\ \varepsilon_{s,j} + b^{\hat{s},\hat{j}}_{s,j}$, both $a^{\hat{s},\hat{j}}_{s,j}$ and $b^{\hat{s},\hat{j}}_{s,j}$ are positive numbers.

The sensitivity factor is the amount by which the distortion will increase if the packet is lost. From (13) and (14), we can see that packet $P_{s,j}$ has greater sensitivity factor $S_{s,j}$ if one or more of the following criteria are met: 1.) $\Delta d_{s,j}$ is greater; 2.) $V_{s,j}$ is higher; 3.) there are more packets in $B(P_{s,j})$; 4.) Packets in $B(P_{s,j})$ have greater distortion increments; 5.) Packets in $B(P_{s,j})$ have lower loss probability; 6.) $P_{s,j}$ has higher impact to the verification probability of packets in $B(P_{s,j})$. In particular, the signature packet P_{sig} has the highest sensitivity factor, because all packets will not be verifiable if P_{sig} is lost. Note that the authentication dependency is implicitly accounted for when the sensitivity factors are computed using (13) and (14).

The sensitivity factor, together with the size (packet size plus authentication overhead), determines how the bandwidth is allocated among the packets. In the resulting optimized policy, a packet will have more transmission opportunities if its Lagrangian multiplier is smaller, i.e. smaller size and greater sensitivity factor.

4. EXPERIMENTAL RESULTS

To evaluate the proposed technique, we implemented five systems using NS-2 [9] and H.264/AVC JM10.1 [10]. The first system, *RaDiO*, implements the *RaDiO* video streaming technique [3] for un-authenticated video, whose performance is used for reference. The second system, *RaDiO_butterfly_Aware*, implements our proposed technique with butterfly authentication [4]. The third system, *RaDiO_Butterfly_Unaware*, implements *RaDiO* streaming [3] and Butterfly authentication [4]. The fourth system, *RaDiO_EMSS*, implements *RaDiO* streaming [3] and EMSS authentication [7]. The fifth system, *RaDiO_AC*, implements *RaDiO* streaming [3] and Augmented Chain authentication [8]. However, in the last three systems, the *RaDiO* streaming is *authentication-unaware*, i.e. it does not recognize authentication dependency. For all systems with authentication, we use SHA-1 for hashing (16 bytes) and RSA for signature (128 bytes), a signature is amortized among 33 packets (around 1-second video data), which corresponds to around 8Kbps overhead. We compare the R-D performance at various loss rates and transmission rates. All authentication methods are configured with their respective optimal parameters.

In our experiment, we consider sender-driven retransmission streaming scenario, where the receiver acknowledges every received packet. The channels are packet-erasure. Packet loss and delay are random and independent in both forward and backward channels. Packet delay follows a shifted Gamma distribution with parameter k, n and α, where n is the number of routers in the path, k/n is the constant delay per router, $1/\alpha$ and $1/\alpha^2$ are the mean and variance of the queuing delay per router. Both forward and backward channel follows the same delay distribution: $k_F = k_B = 50ms$, $n_F = n_B = 2$, $1/\alpha_F = 1/\alpha_B = 25ms$. The time interval between consecutive transmission opportunities is $T=100ms$ and playout delay is $\delta=600ms$. At time t, packets whose deadline falls in $[t+k, t+k+\delta]$ are eligible for transmission. A fixed λ is used for each streaming session. The video sequence, *Foreman*, has 400 QCIF frames at 30 frames/s, and is encoded at about 150Kbps. Each GOP comprises of one I-frame followed by 14 P-frames.

The simulation results show that *RaDiO_Butterfly_Aware* consistently outperforms *RaDiO_Butterfly_Unaware*, *RaDiO_EMSS* and *RaDiO_AC* at various loss rates (0.03, 0.05, 0.1 and 0.2) and transmission rates. Due to space limitation, we only give R-D curves at loss rate 0.1 (e_F=e_B=0.1) in Fig. **3** and 0.05 (e_F=e_B=0.05) in Fig. **4**, where e_B and e_F are loss rates of the backward and forward channels. For comparison purpose, we also measure the R-D performance of *RaDiO* for un-authenticated video with (1) no loss and no delay, and (2) loss but no delay. At low bandwidth, *RaDiO_Butterfly_Unaware*, *RaDiO_EMSS* and *RaDiO_AC* produce unacceptable performance, because the Y-PSNR drops quickly due to their lack of awareness of authentication dependencies. When bandwidth is scarce, packets with smaller Δd will have less transmission opportunities, leading to high chance of loss. However, these packets can be very important for verifying other packets, and their loss therefore greatly degrades video quality. The steep slope and quick dropoff in performance for the *authentication-unaware* techniques may be reduced by increasing the packets' verification probability, but this would require significant additional authentication overhead which would negatively impact the overall R-D performance. *RaDiO_Butterfly_Aware*, being *authentication-aware*, has much better performance. Its R-D curve closely follows *RaDiO without authentication.* The performance gap is around 2dB, because *RaDiO_Butterfly_Aware* has 8Kbps overhead for authentication and also the original packets' importance is not fully aligned with packets' authentication importance.

From sender's point of view, the channel capacity is $(1-e_B)(1-e_F)R$, where R is the bandwidth. To transmit all data packets (150+8 Kbps), the required bandwidth is 158/(1-0.1)(1-0.1)=195Kbps when e_F=e_B=0.1, and 158/(1-0.05)(1-0.05)=175Kbps when e_F=e_B=0.05. When the bandwidth is smaller than this, the sender will transmit only the more important packets. This is the reason why the *RaDiO_Butterfly_Aware* curve starts dropping at 175Kbps in Fig. 4 and 195Kbps in Fig. 3.

5. CONCLUSION

This paper proposes an *authentication-aware* R-D optimized streaming technique. The sender is able to dedicate the limited bandwidth to those packets that are more important for video quality and packet authentication. Experimental results show substantial performance gains over both *authentication-unaware* R-D optimized streaming techniques and authentication-aware non-R-D-optimized streaming techniques, especially when the channel capacity is below the source rate.

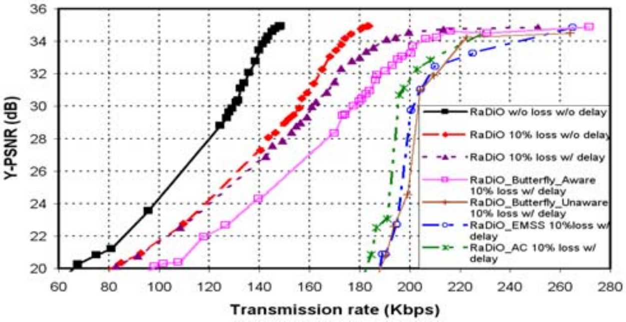

Fig. 3 – R-D performances when loss rates e_F=e_B=0.1

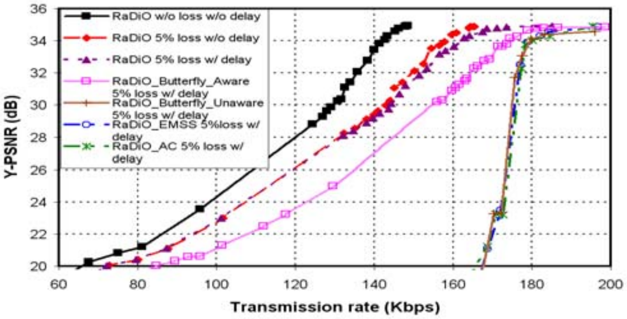

Fig. 4 – R-D performances when loss rates e_F=e_B=0.05

ACKNOWLEDGEMENT: The authors would like to thank Yan Li (Stanford University) and Karsten Suehring (Fraunhofer Institute) for their help on H.264 JM software.

7. REFERENCES

[1] J. Apostolopoulos, "Secure media streaming & secure adaptation for non-scalable video," IEEE International Conference on Image Processing, October 2004

[2] P.A. Chou and Z. Miao, "Rate-distortion optimized streaming of packetized media," Microsoft Research Technical Report MSR-TR-2001035, February 2001

[3] J. Chakareski, J. Apostolopoulos and B. Girod ``Low-Complexity Rate-Distortion Optimized Video Streaming," IEEE International Conference on Image Processing, October 2004

[4] Z. Zhang, Q. Sun and W.C. Wong, "A proposal of butterfly-graph based stream authentication over lossy networks," in Proc. IEEE International Conference on Multimedia & Expo, July 2005

[5] C.K. Wong and S. Lam, "Digital Signature for Flows and Multicasts," The University of Texas at Austin, Department of Computer Sciences, Technical Report TR-98-15, July 1998

[6] R. Gennaro and P. Rohatgi, "How to sign digital streams,", in Advances in Cryptology – CRYPTO'97, pp. 180-197

[7] A. Perrig, R. Canetti, J. Tygar and D. Song, "Efficient authentication and signing of multicast streams over lossy channels," in Proc. of IEEE Symposium on Security and Privacy, 2000, pp. 56-73

[8] P. Golle and N. Modadugu, "Authentication streamed data in the presence of random packet loss," ISOC Network and Distributed System Security Symposium, 2001, p.13-22

[9] The Network Simulator (NS-2), http://www.isi.edu/nsnam/ns/

[10] H.264/AVC Reference Software, JM Version 10.1, http://iphome.hhi.de/suehring/tml

[11] Z. Zhang, Q. Sun, W.C. Wong, J. Apostolopoulos S. Wee, "Rate-distortion optimized streaming of authenticated video," *Technical Report*: I2R-2005-031