Robust Delay Estimation for Internet Multimedia Applications

Kyung-Joon Park[1], Won Jong Jeon[2], Klara Nahrstedt[2], Tamer Başar[3], and Chong-Ho Choi[1]

[1] School of Electrical Engineering and Computer Science and ASRI
Seoul National University, Seoul 151-742 KOREA
{kjpark, chchoi}@csl.snu.ac.kr
[2] Department of Computer Science
University of Illinois at Urbana-Champaign, IL 61801 USA
{wonjeon, klara}@cs.uiuc.edu
[3] Department of Electrical and Computer Engineering and Coordinated Science Laboratory
University of Illinois at Urbana-Champaign, IL 61801 USA
tbasar@decision.csl.uiuc.edu

Abstract. Traffic estimation and modeling has been a crucial issue in many research areas of communication networks. For example, intermediate routers in networks estimate the rate of packet flows via queue size information in order to maximize throughput and provide fairness between flows. Also, many senders in the end system such as multimedia applications estimate end-to-end delay between a sender and a receiver to reduce dropping probability of packet. In this position paper, we propose a robust adaptive estimation scheme for Internet multimedia applications. The proposed scheme adopts an autoregressive (AR) model for the process and identifies the parameters of the AR process via a robust identification algorithm. This robust identification algorithm usually leads to better performance when the noise is correlated and/or nonstationary, and it is also more robust to modeling uncertainties. Here, we consider the problem of end-to-end delay estimation for audio playout mechanisms. We rigorously formulate the proposed scheme in the realm of audio playout mechanisms and give some preliminary simulation result which shows effectiveness of the proposed scheme. Even though we apply the proposed scheme to the estimation of end-to-end delay in audio applications, the scheme itself is very general and we expect that it can be applied to many other estimation problems in Internet multimedia applications.

1 Introduction

Recently, there has been a significant increase of interest in Internet multimedia applications, i.e. video conferencing, Internet telephony, video-on-demand etc. This kind of applications require quality of service (QoS), such as throughput, packet loss, delay, and jitter. However, the current Internet multimedia applications commonly employ the UDP transport mechanism, which is not capable of

F. Boavida et al. (Eds.): IDMS/PROMS 2002, LNCS 2515, pp. 255–262, 2002.

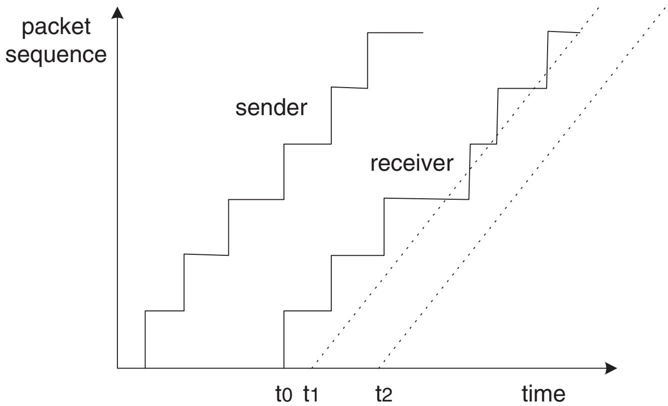


Fig. 1. Playout Jitter Problem.

congestion control. Consequently, QoS is usually provided by application-level end-user adaptation [1], [2], [3].

In many applications, the problem of traffic modeling and estimation at the end users is a crucial issue for providing QoS [2], [3], [4]. One of the applications is the estimation of packet delay for audio playout mechanisms [5]. In Internet audio applications such as real-time voice communication and packet radio service, delay jitter are the most crucial factor for QoS. In order to alleviate the problem caused by unpredictability of delay and delay jitter in the Internet, the receiver usually needs to buffer some amount of packets before it actually plays them. Small buffering delay cannot tolerate severe delay jitter and also lead to significant packet loss, whereas large buffering delay causes large startup delay. Therefore, the amount of buffered packets and timing of playout are very important for the performance of audio applications. Figure 1 shows the sample sequence of operations on a sender and receiver of an audio application. Two solid lines represent the sequence number of audio packets sent and received respectively, and two dotted lines represent the sequence number of packets played by the receiver. Delay jitter during transmission causes the uneven arrivals of the packets at the receiver. In order to reduce such delay jitter, the receiver delays the initiation of playout of received packets. t_1 and t_2 represent two different playout times. In case of t_1, some of packets are delivered to the receiver after their playout time, therefore they are dropped without playing. In case of t_2, the receiver has larger delayed startup time, i.e. $t_2 - t_0$ than that in case of t_1, which is $t_1 - t_0$. Consequently, the main objective of playout algorithms is to choose a small startup time which can also keep acceptable packet loss rate.

The basic algorithm of delay estimation used in audio conferencing tools such as NeVot 1.4 [6] has been influenced by RFC 793 TCP RTT estimation [7]. The estimate of average packet delay d_i is as follows:

$$d_i = \alpha d_{i-1} + (1-\alpha) n_i \tag{1}$$

where n_i is the delay suffered by the i-th packet in the network, and α is a weighting factor which controls the rate of convergence of the algorithm. The variation in this delay, v_i is estimated by

$$v_i = \alpha v_{i-1} + (1-\alpha)|d_i - n_i|. \tag{2}$$

This is used to fix the end-to-end delay, *ted* for playing out the next packet as follows:

$$ted = d_i + \beta v_i \tag{3}$$

where β is called a safety factor used to guarantee that the estimated delay is larger than the actual delay with a high probability. The most significant shortcoming of this basic algorithm is that it does not adapt to network traffic efficiently. Once the value of α is given, the model is fixed and this can degrade estimation performance. We should not ignore that TCP can somehow adapt itself to networks, but it does not adapt itself efficiently, especially when there is an abrupt change of network condition.

Many researchers have suggested modified algorithms for better performance [1], [5], [8], [9], [10]. However, most of the proposed schemes were not based on estimation theory, but some kind of heuristics. Recently, P. DeLeon and C. Sreenan [11] have proposed a scheme which used a simple normalized least-mean-square (NLMS) algorithm in adaptive filter theory [12].

In this paper, we propose a robust adaptive estimation algorithm based on recent development in robust control [13]. The proposed scheme is not only adaptive in its nature, but also robust to non-stationary noise and modeling uncertainties. We will also explain that most of the existing schemes correspond to special cases of the proposed scheme. We further expect that the proposed scheme can be applied to many estimation problems in Internet multimedia applications. The rest of the paper is organized as follows: In Section II, we will formulate the delay process as an autoregressive (AR) model and give a design criterion which is of H^∞ type [14]. In Section III, we derive the update rules for the estimate of α and the estimate of the variance σ^2. We give some preliminary simulation result in Section IV. Finally, conclusion and future work follows in Section V.

2 Problem Formulation

Here, we will closely follow the methodology introduced in [15]. This methodology is a discrete-time version of the original algorithm in [13]. The AR model we adopt is as follows:

$$d_{n+1} = \sum_{i=1}^{p} \alpha_i d_{n+1-i} + \phi_n \tag{4}$$

where d_i is the i-th packet delay, α_i's are parameters of the AR process to be identified, and ϕ_n is an unknown noise sequence. Here, p is called the order of the AR process. Note that the complexity and the computational burden of the algorithm is determined by p. The AR model (4) is a generalization of (1) since the model (1) corresponds to the special case of $p = 1$ of (4). Furthermore, while the basic algorithm (1) uses a fixed weighting factor α, the proposed algorithm identifies the parameters α_i's via a robust identification method as in the next section.

We can express (4) with the following vector notation:

$$d_{n+1} = \boldsymbol{\alpha}^T \eta_n + \phi_n \tag{5}$$

where $\boldsymbol{\alpha} := (\alpha_p, ..., \alpha_1)^T$ and $\eta_n := (d_{n+1-p}, d_{n+2-p}, ..., d_n)^T$. What we are going to do is to identify $\boldsymbol{\alpha}$ based on the previous data, i.e. η_n.

We wish to obtain a sequence of estimates for $\boldsymbol{\alpha}$, denoted $\hat{\boldsymbol{\alpha}}_n$ at step n, so that $\hat{\boldsymbol{\alpha}}_n$ would depends on all the past and the present value of $d_{(\cdot)}$, i.e. $\hat{\boldsymbol{\alpha}}_n = \hat{\boldsymbol{\alpha}}(d_n, d_{n-1}, ..., d_0)$. The criterion to be minimized is of the H^∞ type [14], which is the gain from the energy of the unknowns to a weighted quadratic identification error:

$$J(\{\hat{\boldsymbol{\alpha}}_n\}_{n=0}^{\infty}) := \sup_{\{\phi_n\}_{n=0}^{\infty}, \boldsymbol{\alpha}} \frac{\sum_{n=0}^{\infty} (\boldsymbol{\alpha} - \hat{\boldsymbol{\alpha}}_n)^T Q_n (\boldsymbol{\alpha} - \hat{\boldsymbol{\alpha}}_n)}{\sum_{n=0}^{\infty} |\phi_n|^2 + (\boldsymbol{\alpha} - \overline{\boldsymbol{\alpha}}_0)^T \overline{Q}_0 (\boldsymbol{\alpha} - \overline{\boldsymbol{\alpha}}_0)} \tag{6}$$

where $\overline{\boldsymbol{\alpha}}_0$ is some initial estimate for $\boldsymbol{\alpha}$, $\overline{Q}_0 > 0$ is a fixed weighting matrix, and $Q_n \geq 0, n = 0, 1, ...$ is a sequence of weighting matrices which will be specified later. Intuitively, the criterion J is an index that measures worst-case attenuation from additive disturbance and error in the initial estimate to the estimation error over an interval of interest. Consequently, our objective is to find a sequence $\hat{\boldsymbol{\alpha}}^* := \{\hat{\boldsymbol{\alpha}}_n^*\}_{n=0}^{\infty}$, with the properties that

$$J(\hat{\boldsymbol{\alpha}}^*) = \inf_{\hat{\boldsymbol{\alpha}} = \{\boldsymbol{\alpha}_n\}_{n=0}^{\infty}} J(\hat{\boldsymbol{\alpha}}) =: (\gamma_{id}^*)^2 \tag{7}$$

and $\lim_{n\to\infty} \hat{\boldsymbol{\alpha}}_n^* = \boldsymbol{\alpha}$.

3 Robust Identification Algorithm

3.1 Update Rule for Estimation of α

Now, from [15], we derive update rules for estimate of $\boldsymbol{\alpha}$. First we introduce Σ_n, which is a sequence of $p \times p$-dimensional positive-definite matrices. Note that Σ_n is necessary in the way of estimation of $\boldsymbol{\alpha}$ and does not have significant physical meaning. The following is the update rule for Σ_n:

$$\Sigma_{n+1} = \Sigma_n + \eta_{n-1}\eta_{n-1}^T - \frac{1}{\gamma^2} Q_n, \tag{8}$$

$$\Sigma_1 = \overline{Q}_0 - \frac{1}{\gamma^2} Q_0. \tag{9}$$

Here, if we let $\overline{Q}_0 = Q_0 = I_p$, where I_p is the $p \times p$ identity matrix, and $Q_n = \eta_{n-1}\eta_{n-1}^T$ in which case $\gamma_{id}^* = 1$ in (7), then we get the following update rule for Σ_n:

$$\Sigma_{n+1} = \Sigma_n + (1-\gamma^{-2})\eta_{n-1}\eta_{n-1}^T, \tag{10}$$

$$\Sigma_1 = (1-\gamma^{-2})I_p. \tag{11}$$

Here γ is a parameter of the algorithm and should be larger than 1.

Now, we have the following update rule for $\hat{\boldsymbol{\alpha}}_n$:

$$\hat{\boldsymbol{\alpha}}_{n+1} = \hat{\boldsymbol{\alpha}}_n + (d_{n+1} - \hat{\boldsymbol{\alpha}}_n^T\eta_n)(\Sigma_{n+1} + \eta_n\eta_n^T)^{-1}\eta_n, \tag{12}$$

$$\hat{\boldsymbol{\alpha}}_0 = \boldsymbol{\alpha}_0. \tag{13}$$

Here $\hat{\boldsymbol{\alpha}}_n$ denote the estimate of $\boldsymbol{\alpha}$ for n-th iteration and $\boldsymbol{\alpha}_0$ an initial value. Note that $\hat{\boldsymbol{\alpha}}_n$ is a p-dimensional vector.

Remark 1: The estimator (12) is a generalized LMS filter [13]. Hence, the NLMS estimator in [11] corresponds to a special case of the proposed algorithm. Furthermore, if we let $\gamma \uparrow \infty$, the proposed algorithm will be precisely the least-squares (LS) estimator [13], [15]. In general, the proposed estimator ranges from the LS estimator to the generalized LMS estimator for certain choices of γ.

Remark 2: The proposed algorithm has been derived based on H^∞ optimal control theory [14]. Hence, it usually shows better performance when the noise is correlated and/or non-stationary, and it is also more robust to modeling uncertainties. Consequently, unlike the NLMS estimator, the proposed algorithm does not require pre-filtering methods such as Discrete Wavelet Transform (DWT) in [9] for decorrelation.

3.2 Update Rule for Estimation of the Variance

Here we obtain an estimate of the variance σ^2 of ϕ_n. We need the variance σ^2 when we calculate the playout time as in (3). First we define the autocorrelation function estimator at step n as follows:

$$\hat{R}^n[k] = \frac{1}{n+1}\sum_{i=k}^{n} d_i d_{i-k} \tag{14}$$

where $k = 0, 1, ..., p$. Then we have the following recursive relation:

$$\hat{R}^{n+1}[k] = \frac{n+1}{n+2}\hat{R}^n[k] + \frac{1}{n+2}d_{n+1}d_{n+1-k}. \tag{15}$$

Now we have the following equation for the estimate of variance σ_n^2:

$$\hat{\sigma}_n^2 = \hat{R}^n[0] - \sum_{i=1}^{p} \hat{l}_n^i \hat{R}^n[i]. \tag{16}$$

Here $\hat{l}_n^i$ is the i-th element of $\hat{l}_n$, where $\hat{l}_n$ is the LS estimate of α at time n. Note that $\hat{l}_n$ is a p-dimensional vector. $\hat{l}_n$ can be obtained as follows:

$$\hat{l}_{n+1} = \hat{l}_n + (d_{n+1} - \hat{l}_n^T \eta_n)(\sum_{i=0}^{n} \eta_i \eta_i^T + I_p)^{-1} \eta_n. \tag{17}$$

3.3 Robust Playout Algorithm

Now, the overall playout algorithm for Internet multimedia applications can be summarized as follows:

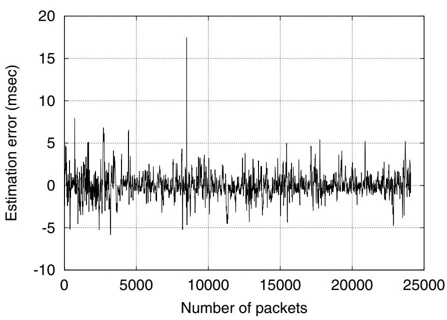
For each talkspurt,
  For every i-th packet,
    Calculate $\hat{\alpha}_i$ and $\hat{\sigma}_i^2$ by (12) and (16).
  End
  Next playout time $t_p = \hat{d}_{n+1} + \beta\hat{\sigma}_n$, where $\hat{d}_{n+1} = \hat{\alpha}_n^T \eta_n$.
End

Here, instead of calculating $\hat{\sigma_n}$ by (16), we can simply use (2) to lessen the computational burden caused by (16).

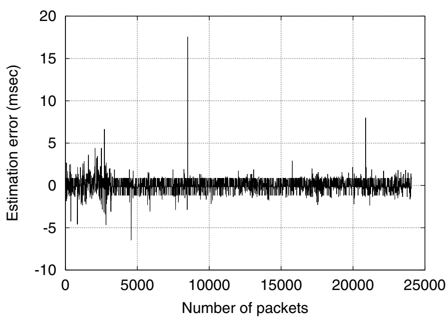
4 Simulation

In this section, we give some preliminary simulation result. We compare the proposed algorithm with the basic algorithm (1). In both cases, we used (2) for simple calculation of the variation. Here, we adopted the traces used in [5]. We only show simulation result for trace 2 in [5] for space limitations.

We used $\alpha = 0.99802$ for the basic algorithm and $\beta = 4.0$ for the variation as in [5] and we set $p = 2$ for the AR model (4) . Here, we compare the estimation error, i.e. the error between the estimated delay and the actual delay for all



(a) Basic algorithm.



(b) Proposed algorithm.

Fig. 2. End-to-end delay estimation for audio application.

packets at the receiver. As we can see from Figure 3.3, the estimation error of the proposed scheme is kept smaller than that of the basic algorithm and also the variation of the proposed scheme is much smaller than that of the basic algorithm. Hence, we can verify the potential effectiveness of the proposed algorithm.

5 Conclusion and Future Work

We have proposed a robust adaptive estimation algorithm which adopt an AR model. The parameters of an AR model are identified by a robust identification algorithm which is based on recent development in robust control [13]. This identification algorithm is not only more effective when the noise is correlated and/or non-stationary, but is also more robust to model uncertainties than the usual estimation schemes. We have applied the proposed scheme to the problem of Internet audio playout mechanisms. Our preliminary simulation result shows the potential effectiveness of the method. We are currently working on performance comparison between the proposed algorithm and other adaptive schemes such as those in [5], [8], [9], [11]. Since the most salient feature of the proposed scheme is its robustness, we are also working on extensive simulation to show the robustness of the proposed scheme over the existing algorithms. We expect that the proposed algorithm can be applied to many estimation problems in Internet multimedia applications.

Acknowledgements. The work of K.-J. Park and C.-H. Choi has been partly supported by the BK21 SNU-UIUC Collaborative Program. The work of W. J. Jeon and K. Nahrstedt has been partly supported by the Kodak fellowship.

References

1. X. Wang and H. Schulzrinne, “Comparison of adaptive Internet multimedia applications,” *IEICE Transactions on Communications*, vol. E82-B, no. 6, June 1999.
2. B. Li and K. Nahrstedt, “A control theoretical model for quality of service adaptations,” in *Proceedings of the IEEE International Workshop on Quality of Service (IWQoS)*, Napa, CA 1998.
3. B. Li, “Agilos: a middleware control architecture for application-aware quality of service adaptations,” *Ph.D Dissertation*, Department of Computer Science, University of Illinois at Urbana-Champaign, May 2000.
4. V. Paxon, “Measurements and analysis of end-to-end Internet dynamics,” *Ph.D Dissertation*, Department of Computer Science, University of California, Berkeley, 1997.
5. R. Ramjee, J. Kurose, D. Towsley, and H. Schulzerine, “Adaptive playout mechanisms for packetized audio applications in wide area networks,” in *Proceedings of IEEE INFOCOM*, 1994, pp. 680–686.
6. H. Schulzrinne, “Voice communication across the Internet: a network voice terminal,” *Technical Report*, Department of Computer Science, University of Massachusetts, Amherst, MA, July 1992.

7. V. Jacobson, "Congestion avoidance and control," in *Proceedings of ACM SIGCOMM*, 1998.
8. C. J. Sreenan, J. C. Chen, P. Agrawal, and B. Narendran, "Delay reduction techniques for playout buffering," *IEEE Transactions on Multimedia*, vol. 2, no. 2, pp. 88–100, June 2000.
9. A. Kansal and A. Karandikar, "Adaptive delay estimation for low jitter audio over Internet," in *Proceedings of IEEE GLOBECOM*, 2001.
10. S. B. Moon, J. Kurose, and D. Towsley, "Packet audio playout delay adjustment: performance bounds and algorithms," *Multimedia Systems*, vol. 6, pp 17–28, Jan. 1998.
11. P. DeLeon and C. J. Sreenan, "An adaptive predictor for media playout buffering," in *Proceedings of IEEE ICASSP*, Phoenix, AZ, March 1999, pp. 3097–3100.
12. S. Haykin, *Adaptive Filter Theory*, Prentice Hall, 3rd edition, Englewood-Cliffs, NJ, 1998.
13. G. Didinsky, Z. Pan, and T. Başar, "Parameter identification for uncertain plants using H^∞ methods," *Automatica*, vol. 31, no. 9, pp. 1227–1250, September 1995.
14. T. Başar and P. Bernhard, *H^∞-optimal Control and Related Minimax Design Problems: A Dynamic Game Approach*, Birkhäuser, Boston, MA, August 1995.
15. Omar Ait-Hellal, Eitan Altman, and Tamer Başar, "A robust identification algorithm for traffic models in telecommunications," in *Proceedings of the IEEE International Conference on Decision & Control*, 1999, pp. 3071–3076.