On Improving Voice Quality Under Dynamic Encoding Algorithms in ATM Networks

Izhak Rubin

Electrical Engineering Department
University of California, Los Angeles
Los Angeles, CA. 90024

Kirk K. Chang

101 Crawfords Corner Road
AT&T Bell Laboratories
Holmdel, NJ. 07737

Abstract

We consider a voice multiplexer which supports ATM packet voice traffic. This system model also applies to the multiplexing operation carried out at the output queue of each output queueing ATM switch. The voice multiplexer implements an encoding algorithm under which the voice quantization levels are dynamically adjusted according to the congestion status of the multiplexer. Using an embedded voice encoding scheme, a voice packet is encoded into lower and higher priority ATM cells. When high congestion conditions are observed, lower priority voice cells (which contain the least significant bits) are discarded. In previous studies, the voice quality under a dynamic encoding algorithm has been described by the long-term average voice encoding rate ($AER_{\infty}$). Due to the strong correlation properties exhibited by the network, we introduce in this paper a "probabilistic" component in the cell-discarding process. A Voice Degradation Indicator (VDI) is introduced to measure the short-term voice quality degradation experienced by a voice stream. Through numerical examples, we identify and compare the system performance under the deterministic and probabilistic algorithms. Under the same $AER_{\infty}$ levels, a probabilistic algorithm is shown to yield better short-term voice quality behavior as well as to improve the mean cell waiting time performance.

1 Introduction

Voice signals are tolerant of temporary quality degradations. Dynamic voice encoding algorithms (formerly identified as bit-dropping algorithms and cell-dropping algorithms in [1,2] and [3], respectively) are developed to take advantage of this property. Under a dynamic voice encoding algorithm, the voice quantization levels are dynamically adjusted to the network congestion levels. For example under an ATM environment, a voice packet is encoded into higher and lower priority ATM cells, containing the most significant bits (MSB's) and least significant bits (LSB's) of the voice samples, respectively. When high congestion conditions are observed, the lower priority cells (also called LSB cells) are discarded.

A significant amount of research work has been devoted to studying such algorithms in various network systems, including packet-switching networks ([1,2]), in which packets are of variable size, and ATM cell-switching networks ([3]), in which voice samples are carried by fixed length ATM cells. It has been shown in [1,2,3] that by using such dynamic voice encoding algorithms, the delay-throughput performance features exhibited by the network are significantly improved.

Under such a dynamic voice encoding algorithm, we improve the network's delay-throughput performance by temporarily degrading the voice stream's quality. In previous works ([1,3,5,6,7,8]), the voice quality behavior under dynamic encoding algorithms have been studied through mathematical analyses and subjective listening tests. In these studies, the steady-state average encoding rate ($AER_{\infty}$), measured in bits per sample, has been used as the key voice quality measure. It has been shown that, in a packet-switching network environment, in order to maintain satisfactory voice quality, we need to have $AER_{\infty} > 3.7$ (bits/sample)[5,6]. In an ATM environment ([3,7,8]), it was observed that up to 10% LSB cell loss rate (among LSB cells), which corresponds to $AER_{\infty} = 3.8$ (bits/sample), leads to imperceptible voice quality degradations.

We consider a voice multiplexer which supports ATM packet voice traffic. Note that this system model applies to the multiplexing operation carried out at the output queue of each output queueing ATM switch. The voice multiplexer implements a dynamic voice encoding algorithm under which the voice quantization levels are dynamically adjusted according to the congestion level of the multiplexer. A voice packet is encoded into lower and higher priority ATM cells. When high congestion conditions are observed, lower priority voice cells (which contain the least significant bits) are discarded.

Noting that the voice packet arrival process is "pseudo-periodic" ([9]), the degradations incurred by the cell discarding process at the multiplexer are not...
distributed uniformly among all active multiplexed voice sources and, over a shorter period of time, some sources will experience higher degradations than others. Therefore, even when a prescribed AER level is satisfied, the dynamic encoding algorithm can cause a voice source to experience significant short-term voice quality degradations.

In this paper, we investigate such a dynamic encoding mechanism for ATM voice cell streams served by a multiplexing system, examining the behavior of the voice quality index over a shorter time scale. To improve performance, rather than using the conventional “deterministic” voice encoding algorithm, we introduce a “probabilistic” voice encoding scheme. In the latter, we incorporate a probabilistic component in determining the quantization level of voice signals. A Voice Degradation Indicator (denoted as VDI) is introduced as a short-term voice quality measure. It represents the probability that the voice quality of a stream is degraded consecutively over a specified period of time. We develop a combined simulation-analytical method to compute this VDI for a multiplexer system. Under this combined method, we first measure the system's performance behavior. We demonstrate that significant improvements in the short-term voice quality and in the mean cell waiting time are attained when a probabilistic algorithm is used.

Through the use of numerical examples, we study extensively the system's performance behavior. We also compare the voice degradation (VDI) levels achieved under the probabilistic and deterministic voice encoding algorithms, for given long-term AER levels. We demonstrate that significant improvements in the short-term voice quality and in the mean cell waiting time are attained when a probabilistic algorithm is used.

The remainder of this paper is organized as follows. We present the system's model and the descriptions of the voice encoding algorithms in Section 2. The short-term voice quality measure (VDI) is defined in Section 3. In Section 4, we qualitatively study the statistical behavior of the voice traffic streams and explain why the probabilistic encoding algorithms lead to short-term voice quality improvements. Numerical examples are presented in Section 5. Final conclusions are drawn in Section 6.

2 System Model and Descriptions of Voice Encoding Algorithms

We consider an ATM voice multiplexer station consisting of a station buffer loaded by V identical ATM voice sources. The system configuration is described in Fig. 1. The multiplexing scheme of the station is assumed to be first-come-first-served. A dynamic voice encoding algorithm is implemented by the server. The voice samples are encoded by using an Embedded-ADPCM scheme ([1,2,4]). The station buffer's capacity is set to be K (cells). Let R (bits/sec) denote the transmission rate of the output link.

The ATM voice cell arrival pattern is presented in Fig. 2. For each voice stream, we observe talkspurts and silence periods occurring alternately. The length of a silence period is assumed to be governed by an exponential distribution. For the cell arrival pattern within a talkspurt, we consider the encoding scheme described in [3]. According to this scheme, a pair of ATM cells is generated every 22 ms. Each cell-pair contains 176 Embedded-ADPCM coded voice samples. The first cell of each cell-pair contains the first two most significant bits (MSB's) of each sample. Therefore, this cell is identified as the MSB cell. The remaining two least significant bits (LSB's) of each sample are carried by the second cell, which is identified as the LSB cell. The number of ATM cell-pairs generated in a talkspurt is assumed to be governed by a geometric distribution.

The speech activity factor, denoted by μ, is directly computed from the mean lengths of a talkspurt and silence period. For example, if the mean lengths of a talkspurt and silence period are equal to 420ms and 580ms respectively, the speech activity factor is equal to μ = 42%.

The conventional “deterministic” voice encoding algorithm (identified as cell-dropping algorithm in [3]) is described in Fig. 3. Note that the variable Q denotes the queue size (in number of cells) at the start time of a cell transmission, and Th denotes the cell-dropping threshold, respectively. If the cell to be transmitted is a LSB cell and if Q > Th, this cell is discarded.

We generalize the deterministic voice encoding algorithm into a probabilistic voice encoding algorithm by incorporating a probabilistic component in the process of discarding LSB cells. The probabilistic voice encoding algorithm is described in Fig. 4. Note that the constant p (0 ≤ p ≤ 1) is a selectable system parameter and that the variable Rq denotes the outcome of a random number generator which generates random numbers uniformly distributed between 0 and 1. For example, assume the cell to be transmitted to be a LSB cell; if Q > Th, this cell will be discarded with probability p. It is noted that a deterministic encoding algorithm is a special case of a probabilistic algorithm in which p is set to be equal to 1.

Note that one can introduce an additional cell-dropping threshold Th2, Th2 > Th, such that if Th2 ≥ Q > Th, a LSB cell will be discarded with probability p; while if Q > Th, this cell is discarded with probability 1. However, it can be shown that the introduction of this Th2 threshold does not impact system performance.

3 Descriptions of the Voice Degradation Indicator (VDI)

In this Section, we introduce a voice quality measure which characterizes the short-term voice quality degradation. It characterizes the quality of the encoded voice samples over consecutive time periods, with individual encoding rate averages computed over a short time scale.

Consider a single voice source (the tagged source). Assume that the voice quality level of this source is
 siguientes. The length of a talkspurt is in the range of 300 ms to 500 ms, which is much longer than the generated deterministically. By investigating the superpositioned traffic stream feeding the multiplexer, we observe the incoming traffic pattern to be "pseudo-periodic" with a "pseudo-period" of 22 ms ([9]).

For a selected ATM voice source, we observe the waiting time realized by a probabilistic algorithm. Assume that the LSB cell of the j-th cell-pair generated by a source is discarded. Due to the "pseudo-periodicity" of the arrival traffic, it is likely that the LSB cell of the (j+1)-th cell-pair generated by the source will also be dropped. As a result, under a deterministic encoding algorithm, the server tends to consecutively discard LSB cells generated by the same source. Hence, the random variables \{Y_j, j \geq 1\} (defined in Section 3) are not statistically independent, as indicated in [5]. In fact, by examining numerically various systems, we have observed the Y process (under a deterministic encoding algorithm) to be modeled by a k-th order Markov chain, with \(k \geq 3\). The Markovian nature of the Y process can lead to severe degradations in the short-term voice quality of an individual stream. Under a probabilistic encoding algorithm, we incorporate a probabilistic component in discarding cells and thus avoid the "per-stream-consecutive-cell-discarding" phenomena. As a result, the distortion level exhibited by the VDI function (under a probabilistic algorithm) is significantly reduced.

Consider a station which implements a deterministic voice encoding algorithm. Assume that the LSB cell of the j-th cell-pair generated by a source is discarded. Due to the "pseudo-periodicity" of the arrival traffic, it is likely that the LSB cell of the (j+1)-th cell-pair generated by the source will also be dropped. As a result, under a deterministic encoding algorithm, the server tends to consecutively discard LSB cells generated by the same source. Hence, the random variables \{Y_j, j \geq 1\} (defined in Section 3) are not statistically independent, as indicated in [5]. In fact, by examining numerically various systems, we have observed the Y process (under a deterministic encoding algorithm) to be modeled by a k-th order Markov chain, with \(k \geq 3\). The Markovian nature of the Y process can lead to severe degradations in the short-term voice quality of an individual stream. Under a probabilistic encoding algorithm, we incorporate a probabilistic component in discarding cells and thus avoid the "per-stream-consecutive-cell-discarding" phenomena. As a result, the distortion level exhibited by the VDI function (under a probabilistic algorithm) is significantly reduced.

Note that one can implement an "individually-recorded" voice encoding algorithm to avoid the consecutive-discarding phenomena described above. For example, an "individually-recorded" algorithm can be designed to record the short-term average encoding rate of each individual voice source and discard LSB accordingly to reduce the measured VDI levels. In this manner, one is able to uniformly distribute degradations among active voice sources. However, such an "individually-recorded" algorithm requires a large amount of memory space and computation time, and is not practically implementable.

5 Numerical Examples

In this Section, numerical examples are presented. We first identify the optimal performance for deterministic and probabilistic algorithms under all 7th and p values. Then, we demonstrate the performance improvement in the short-term voice quality and in cell waiting time realized by a probabilistic algorithm.

Assume that the system is symmetrically loaded. The mean lengths of talkspurts and silence periods of each voice source are equal to 420ms and 580ms, respectively, which correspond to a 42% speech activity factor (\(\mu\)). Assume that the voice encoding scheme described in [3] is used. Under this encoding scheme, the offered load of a voice source in a talkspurt is equal to 38.5 kbps and the average offered load of a voice source is computed as 38.5 x 0.42 = 16.2kbps. The multiplexer's buffer size (K) is assumed to be equal to 200 cells. The VDI values shown in this Section have been obtained through the use of the combined simulation-analytical method described in [10]. All the remaining performance measures have been obtained from simulations.

In Fig. 5 and Fig. 6, the precision of the combined simulation-analytical computational algorithm is verified. We consider two loading scenarios, under which the transmission rate of the output link is equal to 1.536 Mbps and 39.97 Mbps, respectively. As-

4 Short-term Voice Quality Improvements under the Probabilistic Encoding Algorithm

In this section, we qualitatively study the statistical behavior of the ATM voice cell arrival process and explain why the probabilistic encoding algorithm leads to reduced values of VDI(\(t, q\)) and, hence, improvements in short-term voice quality.

For a selected ATM voice source, we observe the following. The length of a talkspurt is in the range of 300 ms to 500 ms, which is much longer than the in-talkspurt cell-pair inter-arrival time (22 ms). Further note that during a talkspurt cell-pairs are generated deterministically. By investigating the superpositioned traffic stream feeding the multiplexer, we observe the incoming traffic pattern to be "pseudo-periodic" with a "pseudo-period" of 22 ms ([9]).
sume that a deterministic algorithm is used and that the Y (or X) process is a 3-rd order Markov chain (i.e. k = 3). We compare the VDI values obtained from the combined method with those obtained from simulations. In Fig. 5, we consider a multiplexing system for which the transmission rate of the output link is equal to 1.536 Mbps and the number of loading voice sources is 95. The cell-dropping threshold (Th) is equal to 50. In Fig. 6, the transmission rate of the output link is equal to 39.97 Mbps and the number of loading voice sources is 95. The cell-dropping threshold (Th) is equal to 100. We show that the VDI values obtained by the combined method are very close to those obtained from simulations. Note that we also have shown (not in figures) the combined method to behave in a similar precise manner when the probabilistic cell discard algorithm is employed.

In Fig. 8, we present the performance of the system under the probabilistic encoding algorithm for different Th and p values, with the output transmission rate set equal to 1.536Mbps. The number of the loading voice sources is equal to 95, which corresponds to an offered load of 1.0 (i.e. 1.536Mbps).

In Fig. 7, we present the VDI(2, lsec, q) curves, for q = 3.5, 3.4, 3.3 and 3.2, as a function of Th = 10, 30 and 60. A noticeable short-term voice quality degradation is observed as p increases. In Fig. 8, we present the performance of the mean cell waiting time, the average voice encoding rate (AER∞) and the MSB cell blocking probability (induced by buffer overflow). We note from Fig. 8(c) that severe cell blocking occurs when p ≤ 0.3, under all Th values. In Fig. 8(b), we observe a noticeable increase in the mean cell waiting time as p decreases. In Fig. 8(b), we also observe an asymptotic behavior in mean cell waiting time, i.e., for p > 0.4, no noticeable improvement in the cell-waiting level is achieved. In conjunction with the observation (using Fig. 7) that noticeable voice quality degradation is observed as p increases, we conclude that the optimal value for p to be equal to 0.4 (for all Th levels).

The optimal value of Th is obtained through Figs. 9-10. In Fig. 9, we show the values exhibited by VDI(2, lsec, q), q = 3.5, 3.4, 3.3 and 3.2, as a function of Th, for p = 1.0 (a deterministic algorithm), and for p = 0.4 (the optimal probabilistic algorithm). It is observed that the VDI values decrease as Th increases. In turn, in Fig. 9 we note the VDI function to exhibit an asymptotic behavior. For example, with p = 1.0, only a limited level of short-term voice quality improvement is achieved by increasing Th beyond 30.

Performance curves of the mean cell waiting time and the average voice encoding rate (AER∞) are shown in Fig. 10. The MSB cell blocking probability (due to buffer overflow) is negligibly small and is therefore not shown. In Fig. 10(a), we note that the values of AER∞ under the two p values under consideration are almost identical (i.e. the numbers of discarded LSB cells under probabilistic and deterministic algorithms are the same). This phenomena can be explained as follows. During a congestion period, under the probabilistic algorithm, the probability of discard-

<table>
<thead>
<tr>
<th>Probabilistic Algorithm (Th=15, p=0.4)</th>
<th>Deterministic Algorithm (Th=30)</th>
</tr>
</thead>
<tbody>
<tr>
<td>VDI(2, lsec, 3.5X(9))</td>
<td>1.43</td>
</tr>
<tr>
<td>VDI(2, lsec, 3.4X(9))</td>
<td>0.58</td>
</tr>
<tr>
<td>VDI(2, lsec, 3.3X(9))</td>
<td>0.25</td>
</tr>
<tr>
<td>VDI(2, lsec, 3.2X(9))</td>
<td>0.16</td>
</tr>
<tr>
<td>Mean Cell Waiting Time (ms)</td>
<td>3.03</td>
</tr>
<tr>
<td>Average Encoding Rate (bps)</td>
<td>3.80</td>
</tr>
</tbody>
</table>

Table 1: Performance Comparison Between Probabilistic and Deterministic Algorithms

We have also studied the loading scenario under which the output transmission rate is set equal to 39.97Mbps. We have shown that the same level of improvement as that shown in Table 1 can also be realized for multiplexing system operating at higher data rates.

6 Conclusions
We consider a voice multiplexer which supports ATM packet voice traffic. This system model also ap-

Authorized licensed use limited to: Univ of Calif Los Angeles. Downloaded on March 4, 2009 at 03:26 from IEEE Xplore. Restrictions apply.
plies to the multiplexing operation carried out at the output queue of each output queueing ATM switch. The voice multiplexer implements a dynamic voice encoding algorithm under which the voice quantization levels are dynamically adjusted according to the congestion status of the multiplexer. A voice packet is encoded into lower and higher priority ATM cells. When high congestion conditions are observed, lower priority voice cells (which contain the least significant bits) are discarded.

We have extended the currently used deterministic cell-discarding algorithm into a probabilistic cell-discarding algorithm. A Voice Degradation Indicator (VDI) has been introduced to describe the short-term voice degradation levels.

Through the use of numerical examples, we have investigated extensively the system's behavior. We have compared the voice degradation (VDI) levels achieved under the probabilistic and deterministic algorithms, for given long-term AER∞ levels. We have demonstrated that significant improvements in the short-term voice quality and in the mean cell waiting time are attained when a probabilistic algorithm is used.

References


Fig. 3 Deterministic Cell-Dropping Algorithm

Fig. 4 Probabilistic Cell-Dropping Algorithm

Fig. 5 Precision of the Combined Algorithm when Output Rate is 1.536Mbps

Fig. 6 Precision of the Combined Algorithm when Output Rate is 39.97Mbps