

Techniques for Packet Voice Synchronization

WARREN A. MONTGOMERY, MEMBER, IEEE

Abstract—Packet switching has been proposed as an effective technology for integrating voice and data in a single network. An important aspect of packet-switched voice is the reconstruction of a continuous stream of speech from the set of packets that arrive at the destination terminal, each of which may encounter a different amount of buffering delay in the packet network. The magnitude of the variation in delay may range from a few milliseconds in a local area network to hundreds of milliseconds in a long-haul packet voice and data network.

This paper discusses several aspects of the packet voice synchronization problem, and techniques that can be used to address it. These techniques estimate in some way the delay encountered by each packet and use the delay estimate to determine how speech is reconstructed. The delay estimates produced by these techniques can be used in managing the flow of information in the packet network to improve overall performance. Interactions of packet voice synchronization techniques with other network design issues are also discussed.

I. INTRODUCTION

ARGUMENTS have been advanced to show that packet switching can be an effective technology to integrate voice and data in a single network [1], [2], [3]. A packet network can exploit the bursty nature of data and voice to achieve a reduction in the transmission bandwidth needed to carry a particular mix of traffic over a circuit network. Packet switching offers many other advantages, including a more flexible allocation of bandwidth to individual calls, and interfaces that are well matched to modern terminals.

A number of significant technical problems exist in sending voice through a packet network. One of the most significant involves the reconstruction of a continuous stream of voice from a set of packets sent through the network [4], [5], [6]. With packetized voice, each packet contains a segment of speech plus some header information. Packets may either be fixed or varying in length. Because of the statistical nature of packet switching, each packet can encounter a different amount of delay in traversing the packet network. If the packets of a given call are routed independently, where each may follow a different path through the network, packets may even arrive at the destination out of order. The significant issue involved is, thus, how to design an appropriate voice reconstruction strategy that reproduces acceptable quality speech from packets that arrive with varying transit delay and may, in some cases, arrive out of order. The remainder of this paper discusses the packet voice synchronization problem and some potential solutions.

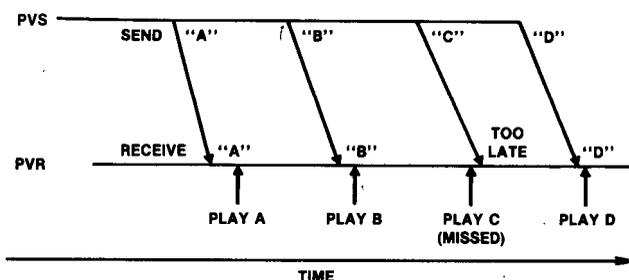


Fig. 1. Packet voice synchronization.

II. THE PACKET VOICE SYNCHRONIZATION PROBLEM

As discussed, the fundamental problem that we are addressing is to reconstruct a continuous stream of speech from packets that arrive with varying transit delay. This is illustrated by Fig. 1, which shows the timing relationships in a typical packet voice network. Packets are produced at the packet voice sender (PVS) and sent through the network. Typically, packet voice systems produce packets at a regular time interval. The selection of the packet production interval is governed by the desire to keep buffering delay and overhead small. Low delay requires small packets, while low overhead requires large packets. Packet intervals of 10–50 ms are typical [1], [7]. We assume that packets are not produced during silent intervals in the call.

As the packets pass through the packet network, each can encounter a varying amount of queuing delay in the statistically multiplexed links. The variation in delay depends on the nature of the packet network (i.e., local area network (LAN) or long-haul transport network), the traffic on the network, and the speed of the network facilities. For a local area network, variable delay is typically small, less than 10 ms [8], [9]. For a long-haul network, variable delay can be significantly larger, typically up to 100 μ s. [10]. The packet voice synchronization problem is, thus, generally more significant in a long-haul network than in a local network application of packet voice.

As the packets arrive at the packet voice receiver (PVR), they are reconstructed into a continuous stream of voice samples and delivered to the destination customer. Typically, this reconstruction is done by choosing a target playout time for each incoming packet as a fixed interval, D_m , after the packet is produced [5]. Each packet that arrives before its target playout time is placed in the proper sequence in a queue of packets from which the speech is reconstructed. If a packet does not arrive before its playout

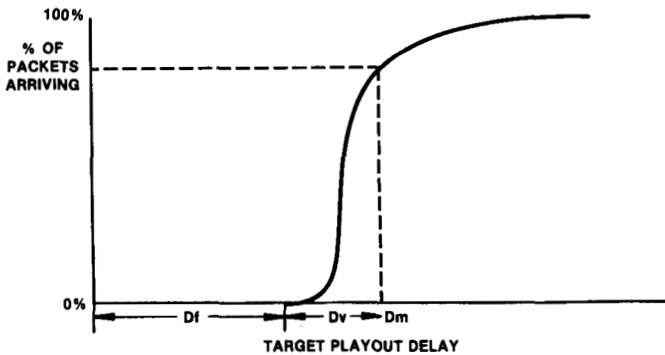


Fig. 2. Delay variation in a packet network.

time, the queue will be empty, and the packet is effectively lost.¹ In Fig. 1, packets *A*, *B*, and *D* arrive in the prescribed interval and are used, while packet *C* is excessively delayed.

We assume that the target playout time for each packet is a fixed time interval following the production of that packet at the sender. This is required to ensure that the entire system introduces a fixed delay into the speech path, and that continuous speech can be reconstructed without varying delay. The choice of the time interval should be made to optimize the delay and packet loss.

Fig. 2, which illustrates this tradeoff in a typical packet network, shows qualitatively the proportion of packets that arrive in less than a given target transit delay for a particular packet connection. Fewer packets are lost if a longer delay is allowed because a larger proportion of packets will experience that delay, or less, in the network. The delay curve is determined from the delay characteristics of the packet network.

The delay experienced by a packet can be divided into a fixed delay (D_f), which is the same for each packet in a call, and a variable delay (D_v). The fixed delay arises from the propagation of signals representing the packets, and from fixed buffering delays in the network, the PVS, and the PVR. The variable delay results from queueing, and from other variable processing delays in the network.²

It should be emphasized that the variable delay is the key factor. Variable delay introduces the tradeoff between the delay experienced in the call and the rate of packet loss illustrated in Fig. 2. The time interval designated D_v in the figure shows the range of delay variation in this case. Nearly all of the distribution's actual transit time (before buffering to compensate for variable delay) is contained in this interval. Wide delay variation in a particular network will result in a wide interval for D_v and more difficulty in packet voice synchronization. The nature of the variable delay is highly dependent on the nature of the network.

- In a local network, the variable delay is the result of contention for the network transport medium. The nature

of the delay is highly dependent on the particular characteristics of the medium and the access protocols. For most proposed packet voice protocols, the variable delay is quite small.

- In a long-haul network where packets traverse multiple links, the variable delay results from queueing at each link. Variable delay depends on the number of links, utilization, packet size, and link speed. Delays may often be over 100 ms [4].

- If a long-haul network allows the packets of one call to travel different routes, the difference in propagation delays between the different routes followed appears as variable delay. This increases delay variation and may not be appropriate for packet voice on certain networks. [10].

The delay target noted earlier is limited at the upper end by the amount of delay that the customer will tolerate in a call. While there is no hard limit to this delay, studies on customer acceptance of delay in calls [11], [12], [13] suggest that 250 μ s may be appropriate as an upper limit on one-way delay for a packet voice call in a public telephone network.

Previous work [5], [10] has studied the tradeoff between delay and packet loss extensively. The work reported in this paper concentrates on the design of a synchronization mechanism, given particular targets for delay and packet loss.

Once a delay target is chosen for a packet voice call, a mechanism is needed to determine the playout time for each incoming packet. To do so, the PVR must determine the production time for each incoming packet, or equivalently the actual delay experienced by each packet.

In general, the delay or production time determination need not be performed on every incoming packet. Once an estimate of the delay or production time has been made for one packet, the relative production time of subsequent packets can be encoded in information sent in each packet, such as a sequence number. These packets are then placed in proper sequence in the playout buffer to await delivery.

This procedure can be used to encode relative timing information between talkspurts as well as within one talkspurt. During a silent interval, the PVS increments its sequence information to reflect the duration of the silent period, so that the PVR can correctly infer the time between the last delay measurement and the production time of packets in the talkspurt following the silent interval. Delay measurements could be made once per packet, once per talkspurt, once per call, or at any other interval during a call. The scheme relies on the clock at the PVR to be close enough in frequency to the clock at the PVS so that clock skew does not distort the timing.

If the clock in the PVR runs slower than that in the PVS, packets will ultimately begin to accumulate in the buffer. If the PVR clock is faster, the buffer will empty and the packet loss rate will rise, as the target delay will shrink by the amount by which the clock in the PVR is ahead of that in the PVS. To avoid this, the delay target of the first packet (the one for which a delay or production time is available) must be extended by the anticipated maximum drift between the clocks for the period over which relative timing is to be used. For example, with clocks that differ in

¹With some speech coding strategies, it may be possible to make use of a late packet. We assume as the worst case that all of the information in the delayed packet is lost. Whether or not the late packet is used, the speech produced by the PVR is degraded by the late arrival.

²Temperature variations, buffering in multiplexors, and failure recovery procedures in transmission systems can produce variations in the propagation delay of signals along a single link. These variations are small compared to the variable delays of queueing and processing, and can, for the most part, be ignored in analyzing synchronization of packet voice systems.

frequency by 1 part in 10^6 , the drift amounts to $1 \mu\text{s}$ in somewhat more than 16 min of continuous operation. Thus, the effect will be negligible for many calls if only one estimate is made for the entire call. For calls lasting for a long period of time, correction of the delay target will be required. In practice, the delay target can be corrected by increasing or decreasing the duration of a silent interval in the speech with little impact on the speech quality likely.

III. DELAY ESTIMATION METHODS

Many methods can be used to estimate either the packet production time or transport delay of an incoming packet. In this section, we discuss four general methods:

- **Blind delay:** the receiver makes a worst case assumption about the delay encountered by a packet.
- **Roundtrip estimation:** the roundtrip delay between the PVS and PVR is used to estimate the one-way delay of a particular packet.
- **Absolute timing:** synchronized clocks are maintained at the PVS and PVR and packets carry absolute time-stamps.
- **Accumulated variable delay:** the packet network keeps track of the delay experienced by a packet as it travels through the network.

In addition, some variations on these schemes, such as changing the delay estimate by some adaptive process, are possible and are discussed below.

A. Blind Delay

The simplest strategy for estimating the production time of an arriving packet is to make a worst case assumption. Once the arrival time has been estimated, the PVR uses sequence information in subsequent packets to determine the proper playout time for each. We call this blind delay, because the PVR makes its estimate blindly, with no information on the actual packet production time or transit delay.

The worst case assumption for the PVR assumes that the packet on which the estimate is based arrives with minimum transit delay (D_f only), and that other packets may be delayed by significantly more time. Hence, the PVR must set the target playout time for the first packet to be its arrival time plus a maximum variable delay ($\text{Max } D_v$). $\text{Max } D_v$ is the difference between the delay target (D_m) and the fixed network delays (D_f). $\text{Max } D_v$ is chosen so that packets that experience more than this much delay can be lost in the playout procedure without unacceptable degradation of speech quality. An additional small delay is added to the target playout time to allow for clock drift between the PVS and the PVR since relative timing will be used for subsequent packets.

The delay estimate can be revised at any time since a new estimate can be made on any packet. A new estimate may differ from the current estimate by at most $\text{Max } D_v$. The delay target can be adjusted based on a new estimate by extending or contracting a silent interval.

This scheme is quite simple and requires no additions to the packet network to aid in packet voice reconstruction. The PVS must put relative timing information, such as

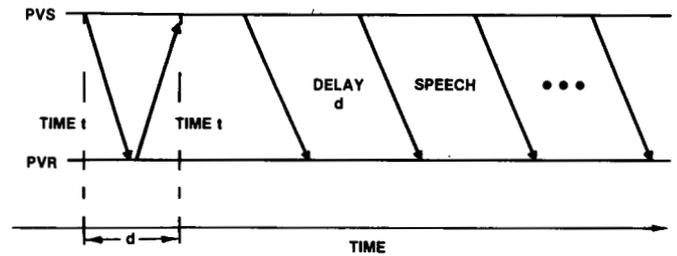


Fig. 3. Roundtrip delay estimation.

sequence numbers, into each packet to be used in reconstruction. The control logic in the PVR is quite simple.

While simple, this scheme does not achieve the best possible performance [5]. The actual delay which is inserted into the call will be equal to the actual transit delay of the packet on which the estimate is made, plus $\text{Max } D_v$. In the worst case, the packet on which the estimate is made experiences a delay of $\text{Max } D_v$ to begin with, and a total delay of $\text{Max } D_v$ longer than the maximum network delay.

For a local area network application, D_v , and thus $\text{Max } D_v$ is relatively small, and other fixed delays are small as well. Blind delay may be very appropriate strategy, with the estimate being made either on the first packet of the call only, or on the first packet of each spurt of speech. The adjustment of silent intervals to update the delay target should not pose a problem because the amount of adjustment will be small. For a long-haul network, however, $\text{Max } D_v$ may be too large for the overall delay introduced by this process to be tolerated.

B. Roundtrip Measurement

Blind delay is a simple technique, but it may not be adequate in a long-haul network. A second delay estimation technique commonly used in maintaining synchronized clocks in a distributed network is to actually measure the roundtrip delay in the communication path between the PVS and PVR, and assume that delay is equally distributed between both directions.

This concept is illustrated in Fig. 3. The measurement is made by sending a packet containing a local clock reading from the PVS. When the packet arrives at the PVR, it is immediately sent back to the PVS. When it arrives back at the PVS, the PVS calculates the roundtrip delay by subtracting the clock value in the packet from the current time. The roundtrip delay is then sent to the PVR, and all subsequent packets sent by the PVS to the PVR contain timing information relative to the first packet sent by the PVS.

While this technique gives an accurate measurement of roundtrip delay, estimation of one-way delay may not be accurate, because the delay in both directions may not be equal. In the worst case, delay in one direction is fixed delay only (D_f), while delay in the other direction includes maximum variable delay ($D_f + \text{Max } D_v$). The maximum error is, thus, half of $\text{Max } D_v$. Thus, the PVS must treat this estimate suspiciously, as it does with blind delay estimation, and allow additional delay to ensure that packets arrive on time.

While roundtrip delay estimation cannot provide a completely accurate delay measurement, it does reduce errors

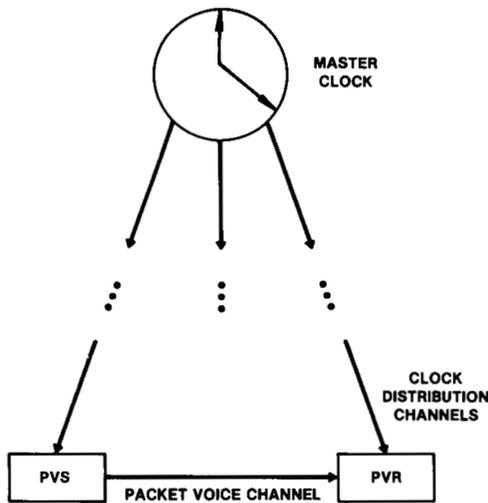


Fig. 4. Clock distribution for absolute timing.

substantially over the blind delay method. In practice, this improvement may be sufficient to permit its use in a long-haul network with low variable delay. In addition, the worst case error described above may not occur in a particular network due to network management practices. In this case, less added delay would be needed to compensate for error in the delay estimate.

C. Absolute Timing

The third technique that can be used is to maintain clocks in the PVS and PVR synchronized to the same absolute time reference. In this case each packet carries an indication of its production time, and the PVR uses that to compute the target playback time. The timestamp need not be capable of encoding the absolute time, but must encode time to a sufficient resolution to allow the PVR to unambiguously determine when the packet was produced. An 8 bit value with the least significant bit representing $1 \mu\text{s}$ should be sufficient for networks where variable delay is less than 250 ms. The PVR can infer higher order bits not sent with the packet by examining the timestamp and the absolute clock. Here, the PVS and PVR are quite simple, and the problem is now one of maintaining synchronized clocks.

Maintaining synchronized clocks in a geographically distributed system is, in general, a complex problem. What is usually required is the distribution of a standard reference frequency to each local clock via a reliable channel [14] to obtain frequency synchronization. Time synchronization is obtained by distributing the absolute time reading of a master clock to the local clocks over communication channels with known propagation delay. This process is used to maintain synchronization in the digital telecommunication network. Fig. 4 illustrates the procedure. The provision of an absolute time reference for the sole purpose of packet voice synchronization may prove to be too costly. There are cases, however, where either the problem is simpler or the time reference exists for other purposes.

In a local area network, clock synchronization is a relatively easy problem because propagation delay is short. A single station on the local area network that periodically

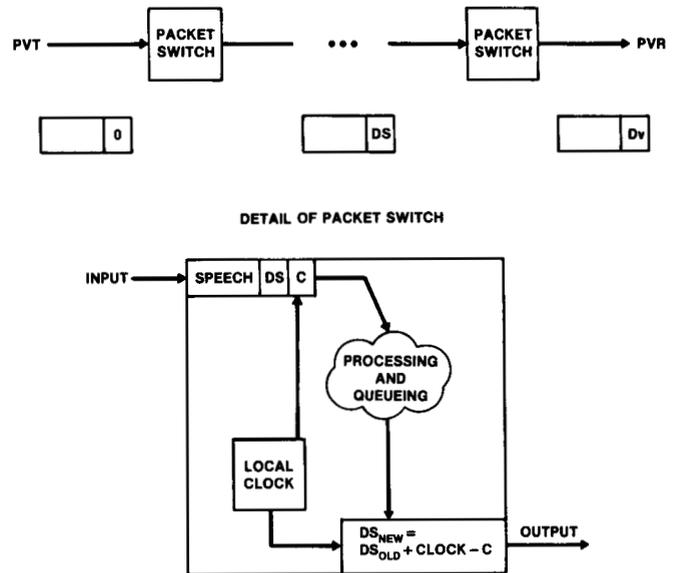


Fig. 5. Added variable delay computation.

sends an absolute time reference to other stations would be sufficient to give adequate synchronization for packet voice reconstruction. Recall, however, that the other synchronization techniques also perform well in this environment, since variable delay is small.

D. Added Variable Delay

The fourth method for estimating the delay experienced by a packet in a packet network is to actually measure the variable delay where it occurs. As indicated before, variable delay results from queueing and processing delays in packet switches and packet terminals. The variable delay measurement can be made by carrying a "delay stamp" indicating the accumulated delay in each packet. Each network element adds its delay to the delay stamp as the packet passes through.

The delay calculation can be made by each network element, using only its local clock. For example, when a packet enters a packet switch, it can be tagged with the reading of the local clock upon entry. The packet will then experience variable delay in processing and queueing. As the packet is about to leave the switch, after any variable delay processing, the tag and the local clock are used to update the delay stamp in the packet. Frequency drift among the local clocks does not affect the delay measurement since the process uses only one clock in computing each component of the delay, and no two clocks need be synchronized. Fig. 5 illustrates the computation of variable delay as a packet traverses a long-haul network.

The delay stamp can be used to encode either total delay, by accounting for the propagation delays on the transmission links as well as the switching delays, or can encode only the variable delay in switches and terminals. It is important only that all variable delay is included in the code, and that the PVR knows what the maximum expected delay encoded in the delay stamp is. This will either be $\text{Max } Dv$ or the maximum total delay.

Fig. 6 illustrates the added variable delay scheme. The

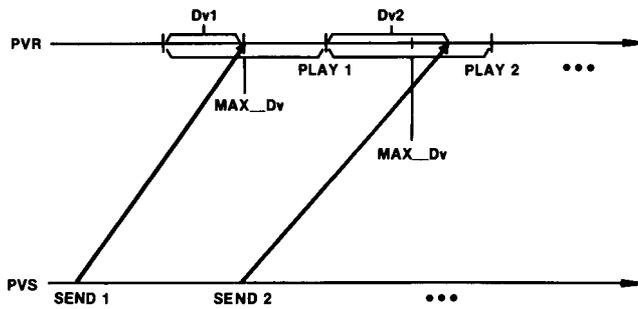


Fig. 6. Packet voice synchronization using delay stamps.

PVR uses the delay stamp in an arriving packet to determine a playout time. This time is simply the current time plus the difference between the maximum expected delay stamp value, and the actual value of the delay stamp in the packet. Intuitively, the process used by the PVR to determine target playout time is as follows.

- Subtract the delay stamp from the local clock to determine the time at which the packet was produced, plus any fixed delays not encoded in the delay stamp.
- The playout target time is then determined by adding the maximum expected delay stamp value to produce the production time of the packet plus the maximum expected transit delay.

As with other techniques, the delay estimate may be used either on each packet, or only occasionally, with relative timing information used on other packets. With this technique, however, it may be simpler for the network to perform the delay stamp calculation on every packet and for the PVR to perform the target time calculation on each, rather than to include separate relative timing information from the PVS in each packet and separate delay handling procedures in the PVR.

Like absolute timing, the added variable delay method provides a good measurement of packet network delay, and thus can achieve packet voice synchronization with minimal total delay.

E. Adaptive Strategies

None of the above techniques produces a completely accurate delay measurement. Error in the delay measurement procedure requires that extra delay be introduced in the call to avoid packet loss when the delay estimate is in error. Blind delay and roundtrip delay estimation may produce estimates with substantial errors, while absolute timing and added variable delay are subject to small errors due to clock drift and propagation delay variation. These techniques can be improved by making use of additional information to adapt the delay estimate as the call progresses.

The adaptation can be made to change the target delay during silent intervals, so that the silence is compressed or expanded by a small amount. Compression or expansion of silence by small amounts is not noticeable in the reconstructed speech.

The adaptive change in playout can be based either on the amount of speech in the buffer (i.e., the length of time between packet arrival and target playout time), or on the basis of repeated roundtrip delay measurements. In either case, the repeated measurements give an indication of whether the current estimate is too high or too low, and the playout targets are adjusted accordingly. Strategies also exist to change the rate at which speech is played out while preserving the pitch of the speech [15].

Adaptive approaches can also be of benefit with absolute timing and with added variable delay in determining the maximum expected transit delay. Most of the time the network will operate with considerably less variable delay than that experienced in the worst case, and transit delays can be reduced by reducing buffering in the PVR without losing any packets. In the case of both this and the use of adaptive techniques with the other strategies, two points are important to remember.

- Adapting playout delay only on the basis of the number of packets arriving too late for playout is quite difficult because few lost packets can be tolerated to give acceptable performance. Thus, the occurrence of lost packets should be infrequent with respect to changes in delay in the packet network, and it may be quite difficult to adjust playout delay based on lost packet rate and still achieve an acceptable lost packet rate.

- Network conditions can change rapidly during intervals where no packets are sent. Thus, a conservative delay estimate should be used after a prolonged silent period.

The various techniques considered in this section provide the basis for reconstructing speech from a stream of packets. The simpler techniques appear to be quite adequate for a local area network application of packet voice, while more complex techniques may be needed to achieve adequate performance in a long-haul network with larger variable delay. The next section considers how the packet voice synchronization techniques interact with other aspects of network design.

IV. THE IMPACT OF SYNCHRONIZATION TECHNIQUES ON NETWORK DESIGN

In practice, the packet voice synchronization techniques will interact with other aspects of the design of a packet voice network. We have already seen one example of this; absolute timing is simpler if packetizing and packet reassembly are done within the network, where a large number PVS's and PVR's can share a common time reference. Other areas of interaction occur in the protocol structure and network management.

A. Protocol Impact of Packet Voice Synchronization

The blind delay, roundtrip delay estimation, and absolute delay techniques require the transport of timing or sequencing information between the PVS and PVR. The

network does not process the timing information, and thus treats it the same as it treats the coded speech. In the ISO model for open systems interconnection [16], this places packet voice synchronization at the transport layer or above.

With absolute timing, there is a slight additional complication. The network must provide the absolute time reference to the PVS and PVR. The distribution of the time reference represents another interface for consideration. If the time reference is distributed through the same physical path used for packet communication, the physical and link layers of the packet protocol may be affected in deriving the channel for clock distribution.

With the added variable delay scheme, timing information must be maintained in each packet and updated in the network. This places the timing function in the network layer (level 3) of the OSI model. This may place an extra burden on nonvoice applications of the packet network to support the timing measurements for packet voice.

B. Network Management Impact of Packet Voice Synchronization

The packet voice synchronization technique can produce significant information that is useful in managing the resources of the network.

Roundtrip delay estimation gives an accurate view of the roundtrip transport delay of a particular connection. This can be used to monitor load in the network, or to refuse service if the transit delay is excessive when the call is set up.

Absolute timing and added variable delay both encode information in each packet that can be used to determine the total packet delay for that packet. This can be used in a variety of ways inside of the network.

- Packets that have been excessively delayed can be discarded in the network, thus reducing the load on the network and the delay for the remaining successful packets.

- Accumulated delay or production time can be used in queueing packets for delivery. This may allow a reduction in maximum delay under heavy load, by queueing packets that have already encountered a large amount of delay in front of packets that have accumulated little delay.

- The delay measurement in each packet can be used to detect congestion in the network and to invoke overload management strategies to divert traffic away from areas of the network with excessive packet load.

In addition to these uses, with the added variable delay method, the calculation done in each switch gives an indication of the congestion in that switch and in the links traversed by the packet. This local estimate can be used to trigger overload management strategies in a switch experiencing long packet delays.

As can be seen from the above discussion, absolute timing and added variable delay both provide information that can be very useful in network management. This

information may offset the added complexity of implementing these techniques.

V. SUMMARY AND CONCLUSIONS

This paper has presented a discussion of the issues involved in reconstructing speech from voice packets. Several techniques were discussed to provide the necessary synchronization. Blind delay and roundtrip delay estimation are relatively simple techniques that should provide adequate performance in a local network environment. More complex techniques may be needed to achieve adequate performance in a long-haul network. For this purpose, absolute timing and added variable delay serve adequately. Both of these techniques involve simple implementations at the packet voice sender and packet voice receiver, but may complicate the long-haul network to provide the necessary information. The added variable delay technique appears to be simpler in the case where absolute time references are not maintained for some purpose other than packet voice synchronization.

The absolute timing and added variable delay techniques provide a number of benefits for network management. Delay information produced by these techniques can be used to aid in the detection and control of network congestion. This makes these techniques very attractive for use in a long-haul network for packet voice.

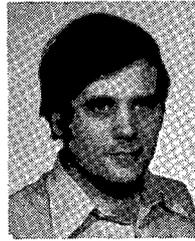
ACKNOWLEDGMENT

The author would like to acknowledge contributions of others at Bell Laboratories to the formation of these ideas. The author is particularly indebted to M. Decina, formerly Professor of Telecommunications at the University of Rome, Rome Italy, for stimulating conversations on packet voice synchronization while on sabbatical leave to work at Bell Laboratories.

REFERENCES

- [1] J. W. Forgie and A. G. Nemeth, "An efficient packetized voice/data network using statistical flow control," in *Proc. IEEE Commun. Conf.*, vol. 3, 1977.
- [2] R. Cole, "Packet voice: When it makes sense," *Speech Technol.*, Sept. 1982.
- [3] I. Gittman and H. Frank, "Economic analysis of integrated voice and data networks: A case study," *Proc. IEEE*, vol. 66, pp. 1549-1570, Nov. 1978.
- [4] J. G. Gruber, "Delay related issues in integrated voice and data networks," *IEEE Trans. Commun.*, vol. 29, pp. 786-800, June 1981.
- [5] G. Barberis and D. Pazzaglia, "Analysis, design and buffer sizing of a packet voice receiver," *IEEE Trans. Commun.*, vol. COM-28, pp. 217-227, Feb. 1980.
- [6] J. W. Forgie, "Speech transmission in packet-switched store and forward networks," in *Proc. AFIPS*, vol. 44, part 1, May 1975, pp. 137-142.
- [7] C. J. Weinstein and H. M. Heggstad, "Multiplexing of packet speech on an experimental wideband satellite network," in *Proc. 7th AIAA Commun. Satellite Syst. Conf.*, San Diego, CA, Mar. 1982.
- [8] N. F. Maxemchuk, "A variation on CSMA/CD that yields movable TDM slots in integrated voice/data local networks," *Bell Syst. Tech. J.*, vol. 61, p. 1527, Oct. 1982.

- [9] J. F. Shoch, "Carrying voice traffic through an Ethernet local network, a general overview," in *Proc. IFIP WG 6.F Int. Workshop Local Area Comput. Networks*, Zurich, Aug. 1980, pp. 429-434.
- [10] G. J. Coviello, "Comparative discussion of circuit versus packet switched voice," *IEEE Trans. Commun.*, vol. COM-27, pp. 1153-1160, Aug. 1979.
- [11] E. T. Klemmer, "Subjective evaluation of delay and echo suppressors in telephone communications," *Bell Syst. Tech. J.*, vol. 42, Nov. 1963.
- [12] —, "Subjective evaluation of transmission delay in telephone conversations," *Bell Syst. Tech. J.*, vol. 56, July-Aug. 1967.
- [13] J. W. Forgie, "Network speech system implications of packetized speech". Massachusetts Inst. Technol., Lincoln Lab., Tech. Annu. Rep. Sept. 30, 1976.
- [14] J. E. Abate *et al.*, "The switched digital network plan," *Bell Syst. Tech. J.*, vol. 56, pp. 1297-1320, Sept. 1977.
- [15] S. E. Youngberg, "Rate/pitch modification of speech using the constant Q transform," in *Rec. ICASSP*, Washington, DC, Apr. 1979, pp. 748-751.
- [16] H. Zimmermann, "OSI reference model—The ISO model of architecture for open systems interconnection," *IEEE Trans. Commun. (Special Issue on Computer Network Architectures and Protocols)*, vol. COM-28, Apr. 1980.



Warren A. Montgomery (M'83) received the A.B. degree from Dartmouth College in 1973, in mathematics and engineering. While at Dartmouth he worked on the operating system for the Dartmouth Time-Sharing System (DTSS). He received the M.S., Engineers, and Ph.D. degrees in computer science from the Massachusetts Institute of Technology, Cambridge, in 1976, 1976, and 1980, respectively. His thesis research for the M.S. was in the area of secure operating systems, and for the Ph.D. in control of updates for distributed databases.

He held an NSF graduate fellowship from 1973 to 1976 and subsequently worked as a Research Assistant in the Laboratory for Computer Science, M.I.T. Since joining Bell Labs he has worked in the Exploratory Telecommunication Services Department on software structures to support new telecommunication services, and on enhanced signaling and other communication protocols. He is currently Supervisor of the Switching Architecture Group, studying architectures for integrated services switching equipment.

Dr. Montgomery is a member of the Association for Computing Machinery and Sigma Xi.

Coded Speech in Packet-Switched Networks: Models and Experiments

GIULIO BARBERIS, MARIO CALABRESE, LIVIO LAMBARELLI, AND DANIELE ROFFINELLA

Abstract—This paper considers the possibility of introducing packetized voice traffic into a packet-switched network.

It is well known that the network must assure voice packets sufficient delay characteristics for conversational speech, i.e., low delay between speaker and listener and low delay jitter or variance. To reach these goals, simplified protocols and priority rules for voice handling are proposed and evaluated.

A model of a packet switching node structure capable of handling both data and voice is derived for both analytical and simulation approaches.

The use of low bit rate voice encoders is considered. The necessity of avoiding the transmission of silent intervals is discussed in relation to the behavior of packet voice receivers.

Proposed strategies are compared by means of analytical tools and simulation experiments considering the presence of voice, interactive, and batch data packets.

Manuscript received April 11, 1983; revised September 20, 1983. This paper was presented at GLOBECOM, San Diego, CA, November 28-December 1, 1983.

The authors are with CSELT, Centro Studi E Laboratori Telecomunicazioni, 10148 Torino, Italy.

I. INTRODUCTION

PACKET switching is a technique widely used in specialized data communication networks with a growing number of actual networks working in the world [1].

The integration of different types of information over such media becomes attractive for large scale economy and for better utilization of these powerful resources [2]. In comparison to other information sources, the human voice seems to be quite different; nevertheless, if it is digitized and suitably compressed, it may resemble the behavior of data from the point of view of its "bandwidth" and "burstiness" characteristics [3]. Moreover, interactive voice inherently demands from the network a better quality of service in terms of average delay and overall delay variance [4].

The aim of this work is to assess the feasibility of introducing packetized voice traffic as a new service in a