Performance Evaluation of a Packetized Voice System - Simulation Study[1]

Tatsuya Suda
Department of Computer Science
450 Computer Science
Columbia University
New York, N.Y. 10027, U.S.A.

Hideo Miyahara
Department of Information and Computer Science
Faculty of Engineering Science
Osaka University
Toyonaka, 560, Japan

Toshiharu Hasegawa
Department of Applied Mathematics and Physics
Faculty of Engineering
Kyoto University
Kyoto, 606, Japan

[1] to appear in IEEE Trans. on Commun.

Table of Contents

ABSTRACT

Introduction of the packet switching technique into digitized voice communication may afford great advantages in efficient use of the channel, compared to both circuit-switched and DSI systems. Detailed characteristics, however, have not been obtained because of difficulty in the exact analysis. Hence, simulation models are developed in this paper for the packetized voice transmission system, and various characteristics such as transmission delays and loss probability of voice packets are obtained. We further evaluate three types of voice packet reassembly strategy at the receiving terminal, and obtain the optimal packet length, which keeps each of overall packet transmission delay and packet loss probability less than a certain permissible values. Comparison among three strategies is also stated.

1. INTRODUCTION

Introduction of packet switching technique into digitized voice communication [1] affords great advantages in efficient use of the channel, flexibility to the network traffic fluctuations, etc., in comparison with conventional circuit-switched systems [2] and DSI systems [3].

In the circuit-switched system, when a call arrives, a transparent line is established for realizing smooth conversation between the origin and destination terminals. This system, however, cannot make efficient use of the channel, because the channel capacity must be assigned to the call according to its peak data rate (not the average rate), and also because the line must be held even if no talkspurt is transmitted.

In the DSI system, channel capacity is assigned on a demand assignment basis only to talkspurts in each call. Hence, channel efficiency is approximately twice that of the circuit-switched system if the call population is large enough [4]. Some portion of talkspurts, however, may be lost because of contention among talkspurts in different calls. (This loss probability should be kept less than 1% for reasonable voice quality [5].)

In the packetized voice communication system [4]-[11], each talkspurt only is encoded and organized into packets, each packet then being transmitted through the network on a store and forward basis. Since voice packets can wait at intermediate nodes in the network until outgoing channels become free, the system will achieve higher channel utilization than is possible in the DSI system, although some delay will result [6]. (It still remains a research question to determine the extent and significance of this trade-off.) Further, it is possible to change and adjust the voice coding rate according to the network traffic congestion, and also to use existing packet-switched data network for voice communication.

Despite the many advantages associated with the packetized voice system, detailed characteristics have not been obtained because of difficulty in the exact analysis. In this paper, simulation models for packetized voice system are developed under three types of packet reassembly strategy for evaluation of various system characteristics such as voice packet transmission delay, loss probability, statistical fluctuations between original and played-out silence intervals, and so on. We also obtain the optimal voice packet length for each of the reassembly strategies, keeping both overall packet transmission delay and loss probability less than some permissible values. Comparison of these three strategies is also stated.

2. PACKETIZED VOICE COMMUNICATION NETWORK

2.1. Packetized Voice Communication Network

In the packetized voice communication network (Fig.1), speech is digitized at an uniform rate by the A/D encoder in the transmitting terminal, and then, organized into constant length packets by the packetizer. The speech detector judges each packet as to whether it contains active parts of the voice or not, and only non-silent packets are transmitted through the network on a store and forward basis. At the receiving terminal, voice packets are stored in the packet voice receiver, and then, decoded into acoustic sound by the D/A decoder.

Since each voice packet waits at intermediate nodes in the network until the outgoing channels become free, packets will arrive at the receiving terminal with randomized inter-arrival times. Because of fidelity requirement of voice, voice packet reassembly strategy, which will play out voice packets at the same uniform rate as they were generated, is required at the packet voice receiver. In this paper, the following three types of reassembly strategy are assumed.

1. N.T.I. (Null Timing Information) Strategy (Fig.2-a) [8]

   The packet voice receiver delays every first packet of the talkspurt by a given amount T of time (control time) and plays out succeeding packets at the same uniform rate as they were generated. If a packet is not received by its played-out time, that packet is considered to be lost. This strategy, requiring no network synchronization, is easy to realize, however, overall packet transmission delay may be relatively large and fidelity of played-out silence intervals may be low.

2. C.T.I. (Complete Timing Information) Strategy (Fig.2-b) [8]

   If the network delay of a packet is less than a given control time T, that packet is additionally delayed at the receiver by an amount equal to the control time T minus its network delay, and then is played out. A packet with a delay greater than T is considered to be lost, even if it be the first one of the talkspurt. This strategy requires network to be synchronized, and also requires timing information in the packet header. However, unlike the N.T.I. strategy, this will keep overall packet transmission delay less than some constant, and will ensure relatively high fidelity of played-out silence intervals.

3. N.T.I.-C.T.I. Mix Strategy

   If the network delay of the first packet of the talkspurt is less than a given control time T, then that packet and successive ones are played out in the same way as that in the C.T.I. strategy. If the network delay of the first talkspurt packet is greater than T, the packet voice receiver plays out that packet immediately upon receiving it, and continues to play out successive packets at the same uniform rate as they were generated. Packets which are not received by their played-out times are considered to be lost.

2.2. Performance Criteria for Packetized Voice Network

Voice packet transmission delay may be one of the most important performance criteria for packetized voice network. We define voice packet total delay W as time interval from the beginning of packetization to its played-out time. W becomes

$$W = W_p + W_q + W_t + R + W_r \tag{1}$$

where

- W_p ; packet generation period
- W_q ; sum of queueing delays at intermediate nodes in a network
- W_t ; sum of voice packet transmission times in a network
- R ; propagation delay in a network
- W_r ; packet demodulation delay (time interval from the arrival at the packet voice receiver to its played-out time)

We further define transmission delay W_s of a voice packet as time interval from the beginning of its packetization to its arrival time at the packet voice receiver. W_s becomes (see Fig.2)

$$W_s = W_p + W_q + W_t + R = W - W_r \tag{2}$$

Here, W_p and W_q are

$$W_p = P/V \tag{3}$$

$$W_t = n[(P + H)/C] \tag{4}$$

where

- P ; voice packet length (excluding header)

- H ; voice packet header length

- V ; voice coding rate

- C ; channel speed

- n ; the number of channels where the packet is transmitted

It is clear from eqs.(3) and (4) that values of W_p and W_t are in proportion to the packet length. W_r, under the fixed reassembly strategy, is in proportion to control time T. W_q depends on the degree of network congestion.

Because of real time requirement of voice transmission, overall delay for each packet should be kept less than a certain permissible value. (Ref.[9] shows that overall delay should be less than 200 m sec. for smooth conversation.) Hence, voice packet length may become a very critical factor. If the packet length is too long, packet generation period W_p and packet transmission time W_t will be very large. If too short, nodal queueing delay W_q will become large due to the packet header overhead.

Voice packet loss probability P_r should also be kept under some permissible value to maintain voice quality. With the voice packet length fixed, if control time T increases, packet loss probability will decrease, however, overall transmission delay will increase. Hence, there exists an optimal control time which minimizes overall packet transmission delay, while keeping packet loss probability under some permissible value.

Following the above presented view, we obtain both the optimal packet length and optimal control time through simulations. We also evaluate statistical fluctuations between original and played-out silence intervals.

3. SIMULATION MODELS

3.1. 1-Hop Model

We consider two simulation models, 1-hop and multi-hop (network) models. Fig.3 shows the 1-hop model configuration, where node i supports N number of calls. Each of N calls begins to be packetized at the transmitting terminal from its arrival instant, and then, non-silent packets only are transmitted to node i. Node i has infinite buffer capacity, and transmits incoming packets to node j on a first-come first-served (FCFS) basis.

3.2. Network Model

In the network model (Fig.4), all packets generated from the same call are transmitted through the same unique path. The reason of our assuming fixed routing scheme is that, in voice communication, voice packets must be played out in the order of their generation. In Fig.4, voice packets generated from a certain call, test packets, are transmitted through a fixed route (node 1-2-...- node M). Voice packets coming from all the other routes are assumed to arrive with the rate λ_i at an intermediate node i. Each node has infinite buffer capacity to store packets. Incoming packets make different queues according to their outgoing lines for their transmissions, and queues are processed independently on a FCFS basis. It is assumed at node i that test packets and the ratio ($1-q_i$) of the rest packets will be transmitted through the line (node i - node i+1) on the fixed route (node 1-2-...-node M). In Fig.4, the real and broken lines show streams of test packets and the other packets, respectively.

4. SIMULATION RESULTS

4.1. Parameters

Simulations are carried out under the following parameter settings. In the 1-hop model, N (number of calls), C (capacity of the line between node i and j), V (voice coding rate) and H (packet header length) are N = 70, C = 1.544 M bits/s, V = 16 K bits/s and H = 100 bits, respectively. V, H and P (packet length) are assumed to be same among all calls. Talkspurt and silence intervals in a call are assumed to obey exponential distributions with mean 1.23 sec. and 1.77 sec., respectively [10]. In simulations, termination of calls and generation of new calls are not considered, that is, N is kept fixed. The reason is that the statistical fluctuations in the presence of talkers are much slower than the statistical fluctuations in the generation and transmission of voice packets [4]. Propagation delay R is assumed to be zero.

In the network model, M (number of nodes), C (channel capacity), λ_i (arrival rate of voice packets from the outside of the fixed route) and q_i (probability of transmission toward the outside of the fixed route) are M = 3, C = 56 K bits/s, $\lambda_1 = \lambda_2 = \lambda_3 = 2$ packets/packet generation period and $q_1 = q_2 = q_3 = 0.7$, respectively. (This approximately corresponds to $1 + \lambda_1$(talkspurt + silence interval length)/(talkspurt length) = 6 calls at the first node, 7.5 calls at the node 2 and 8 calls at the node 3.) Packet arrival process at an intermediate node from the outside of the fixed route is assumed to be Poisson process. This is obviously an approximation of a real system, but this seems to be reasonable because of the Palm-Khintchine's theorem [12]. The theorem guarantees that sum of n independent renewal processes obeys Poisson process if n is sufficiently large [12]. Distributions of talkspurt and silence intervals in a call, V (voice coding rate) and H (packet header length) are same respectively as those in 1-hop model, and furthermore, V and P are same in all calls. Propagation delay R is assumed to be zero.

4.2. Simulation Results

Figs. 5 and 6 are simulation results for the 1-hop model. Fig.5 shows mean packet transmission delay $E[W_s]$ as a function of the packet length P. This shows that there exists an optimal packet length which minimizes $E[W_s]$. (The reason will be explained in the network model results.) The optimal packet length and the minimum $E[W_s]$ are around 75 bits and 5 m sec., respectively. Minoli shows approximate analysis for this 1-hop model in ref.[11]. Using his results, the optimal voice packet length and the minimum $E[W_s]$ in this case become 63 bits and 7.6 m sec., respectively. Our simulation results coincide well with his results. Mean packet total delay E[W] in the N.T.I. strategy is shown in Fig.6 as a function of the packet length. In this figure, control time T is taken so as to minimize E[W], satisfying the condition that packet loss probability P_r is less than 1 %. It can be seen that the optimal packet length (without header) which minimizes E[W] is around 80 bits.

Next we show results for the network model. Mean nodal queueing delay $E[W_q]$, packet generation period W_p, packet transmission time W_t and mean packet transmission delay $E[W_s]$ are shown in Fig.7 as a function of the packet length P. $E[W_s]$ and $E[W_q]$ are obtained through simulations, while W_p and W_t being obtained by eqs. (3) and (4). (Note that these values don't depend on packet reassembly strategy.) There exists an optimal packet length which minimizes $E[W_s]$. This is due to the following trade-off relation. If the packet length decreases, W_p and W_t will decrease. However, W_q will increase due to packet header overhead. This figure shows that, when the packet is long, W_p and W_t greatly contribute to W. The voice packet length should then be relatively smaller than that of the usual data packet (1000 - 2000 bits). Fig.8 shows density function of nodal queueing delay W_q for various values of the packet length. The value of ρ shows traffic intensity at the final node (node 3).

Packet loss probability P_r and mean packet total delay E[W] in the N.T.I. strategy are shown in Fig.9 as a function of control time T. The packet length P is 150 bits in this figure. There exists an optimal control time T^* which minimizes E[W], while keeping P_r under some permissible value. The optimal control time T^*, for example, under the condition of $P_r \leq 1$ % is 39.6 m sec., and the corresponding value of E[W] is 65.5 m sec..

Fig.10 shows mean packet total delay E[W] in the N.T.I. strategy. For each value of the packet length, control time is taken optimal so as to minimize E[W] under the condition that $P_r \leq 1$ %. This figure shows that there exists an optimal packet length P^* which minimizes E[W]. (Mean packet total delays in C.T.I. and N.T.I.-C.T.I. mix strategies have also been obtained, and there are not significant difference among characteristics of these three.) Tab.1 shows the optimal packet length P^*, optimal control time T^* and the corresponding value of E[W] ($E[W]^*$) for each of the reassembly strategies.

Next we consider statistical fluctuation of silence intervals. Here we define fluctuation of silence intervals S as

$$S = \begin{matrix}\text{played-out silence} \\ \text{interval length}\end{matrix} - \begin{matrix}\text{original silence} \\ \text{interval length}\end{matrix} \tag{5}$$

Fig.11 and Tab.2 show density function and the first and second moments of S, respectively. The packet length and control time are taken as optimal values as shown in Tab.1. There is not significant difference between the C.T.I. and the N.T.I.-C.T.I.

mix strategies as for the silence interval fluctuation S. The reason is as follows. The difference between the C.T.I. and the N.T.I.-C.T.I. mix strategies lies in whether loss of the first packets of talkspurts will occur (the C.T.I strategy) or not (the N.T.I.-C.T.I. mix strategy). However, when packet loss probability P_r is kept less than 1 %, the first talkspurt packets are rarely discarded. Hence, there becomes no significant difference in the silence interval fluctuation S. Fig.11 and Tab.2 show that the C.T.I and the N.T.I.-C.T.I. mix strategies are favored over the N.T.I. strategy with respect to the silence interval fluctuation S.

4.3. Considerations

The above simulation results show that there exists an optimal packet length which minimizes overall packet transmission delay, while keeping packet loss probability under a permissible value (1 %). In the above examples, the optimal packet length (including the header) is around 180 bits in the 1-hop model and 250 - 300 bits in the network model. Considering the usual data packet length (1000 - 2000 bits), these optimal values are relatively short. This is because, when the voice packet length is long, the packet generation period greatly contributes to overall packet transmission delay.

Simulations have been executed for three types of the packet reassembly strategy. There is no significant difference among these strategies with respect to overall packet transmission delay, while the C.T.I. and the N.T.I.-C.T.I. mix strategies are superior to the N.T.I strategy as for silence interval fluctuation. However, the N.T.I. strategy does not require network synchronization, resulting in easy implementation and also in reduction of packet header. Ref.[6], for instance, shows that the packet header length can be reduced to 32 bits. Hence, the N.T.I. strategy might have better characteristics than those shown in this paper. Considering the above facts, the N.T.I. strategy is most favored with respect to overall packet transmission delay. However, if the fluctuation of silence intervals in the played-out speech is critical for speech quality, the C.T.I. and the N.T.I.-C.T.I. mix strategies become more favored than the N.T.I. strategy. The study on effects of silence interval fluctuation to played-out speech awaits future investigations.

5. CONCLUSIONS

Various characteristics of the packetized voice communication network such as overall packet transmission delay and packet loss probability are obtained through simulations in this paper. Three types of packet reassembly strategy are also evaluated. We show that there exist both an optimal packet length and an optimal control time which minimize overall packet transmission delay while keeping packet loss probability less than a certain permissible value. The packetized voice communication system is still at its beginning, and many problems are remain unsolved. These problems await future analysis.

References

1. B. Gold, "Digital Speech Network", Proc. of the IEEE, Vol.65, No.12, 1977, pp.1636 - 1656.

2. E. A. Harrington, "Voice/Data Integration Using Circuit Switched Network", IEEE Trans. on Commun., Vol.COM-28, No.6, 1980, pp.781 - 793.

3. M. Schwartz, K. Mase and D. R. Smith, "Priority Channel Assignment in Tandem DSI", IEEE Trans. on Commun., Vol.COM-28, No.10, 1980, pp.1802 - 1809.

4. C. J. Weinstein and E. M. Hofstetter, "The Tradeoff Between Delay and TASI Advantage in Packetized Speech Multiplexer", IEEE Trans. on Commun., Vol.COM-27, No.10, 1979, pp.1716 - 1720.

5. J. W. Forgie, "Voice Conferencing in Packet Network", Proc. of ICC-80, 1980, pp.21.3.1 - 21.3.4.

6. J. W. Forgie and A. G. Nemeth, "An Efficient Packetized Voice/Data Network Using Statistical Flow Control", Proc. of ICC-77, 1977, pp.38.2-44 - 38.2-48.

7. R. V. Cox and R. E. Crochiere, "Multiple User Variable Rate Coding for TASI and Packet Transmission Systems", IEEE Trans. on Commun., Vol.COM-28, No.3, 1980, pp.334 -344.

8. C. Barberis and D. Pazzaglia, "Analysis and Optimal Design of a Packet-Voice Receiver", IEEE Trans. on Commun., Vol.COM-28, No.2, 1980, pp.217 - 227.

9. M. J. Ross, A. C. Tabbot and J. A. Waite, "Design Approaches and Performance Criteria for Integrated Voice/Data Switching", Proc. of the IEEE, Vol.65, No.9, 1977, pp.1283 - 1295.

10. M. J. Fisher, "Delay Analysis of TASI with Random Fluctuations in the Number of Voice Calls", IEEE Trans. on Commun., Vol.COM-28, No.11, 1980, pp.1883 - 1889.

11. D. Minoli, "Optimal Packet Length for Packet Voice Communication", IEEE Trans. on Commun., Vol.COM-27, No.3, 1979, pp.607 -611.

12. J. Cohen, "The Single Server Queue", North Holland, 1969.

Figure and Table Captions

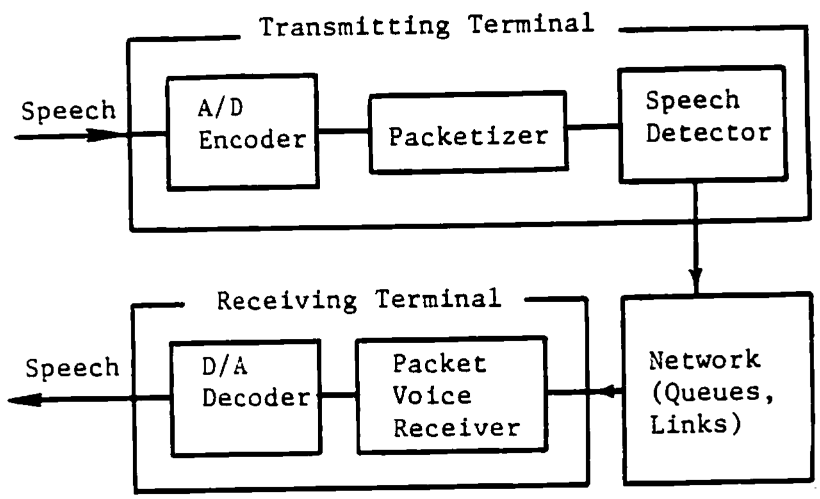
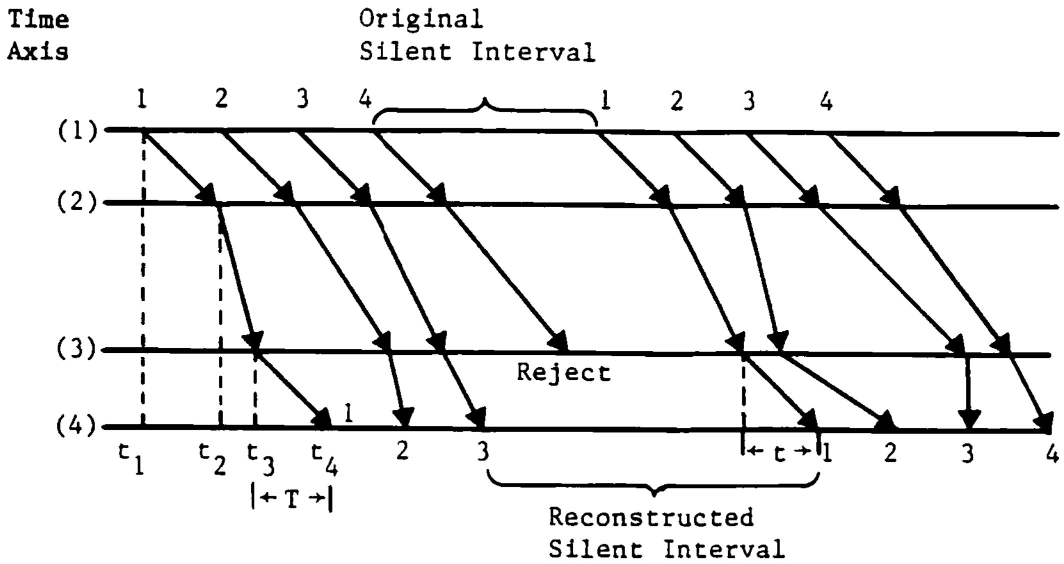
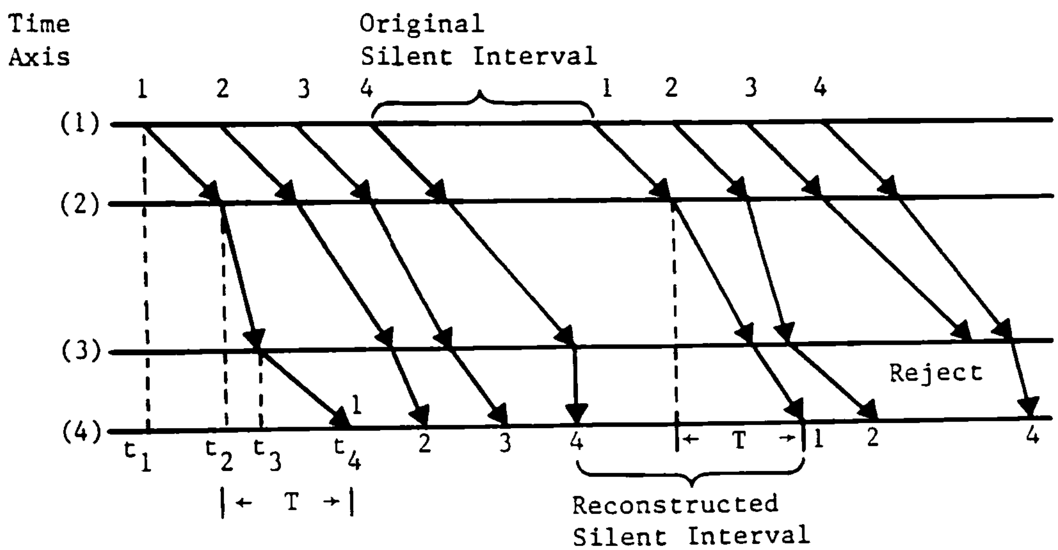


Fig.1 Packetized Voice Communication Network



(2-a) N.T.I. (Null Timing Information) Strategy



(2-b) C.T.I. (Complete Timing Information) Strategy

Fig.2 Packetized Voice Reassembly Strategies

(continue to the next page)

Time Axis

1 --- Beginning Time of Packetization

2 --- Beginning Time of Voice Packet Transmission

3 --- Packet Arrival Time at the Packet Voice Receiver

4 --- Beginning Time of Packet Demodulation

T --- Control Time

$t_4 - t_1$ --- Packet Total Delay W

$t_3 - t_1$ --- Packet Transmission Delay W_s

$t_2 - t_1$ --- Packet Generation Period W_p

$t_3 - t_2$ --- Nodal Queueing Delay W_q + Packet Transmission Time W_t + Propagation Delay R

$t_4 - t_3$ --- Packet Demodulation Delay W_r

Fig.2 Packetized Voice Reassembly Strategies

(Continued)

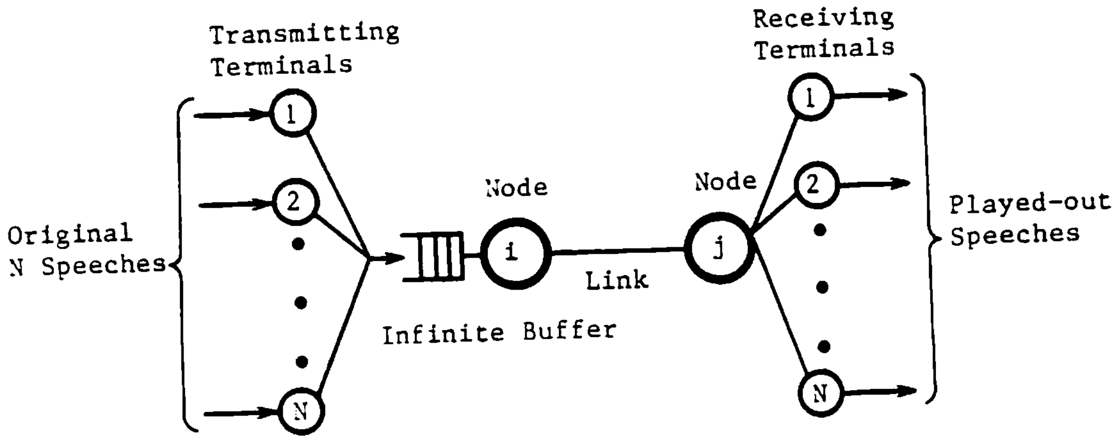


Fig.3 1 - hop Model Configuration

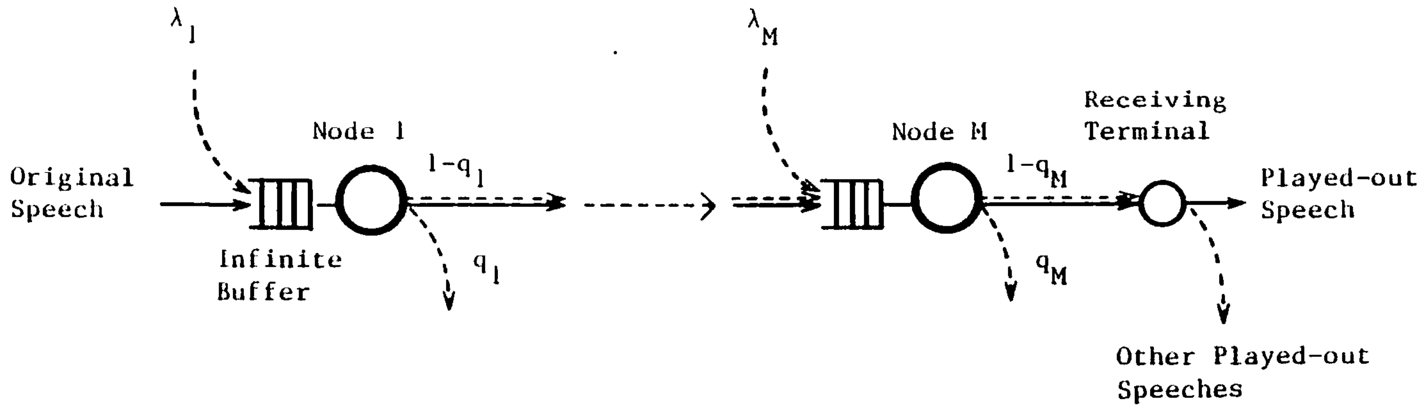


Fig.4 Network Model

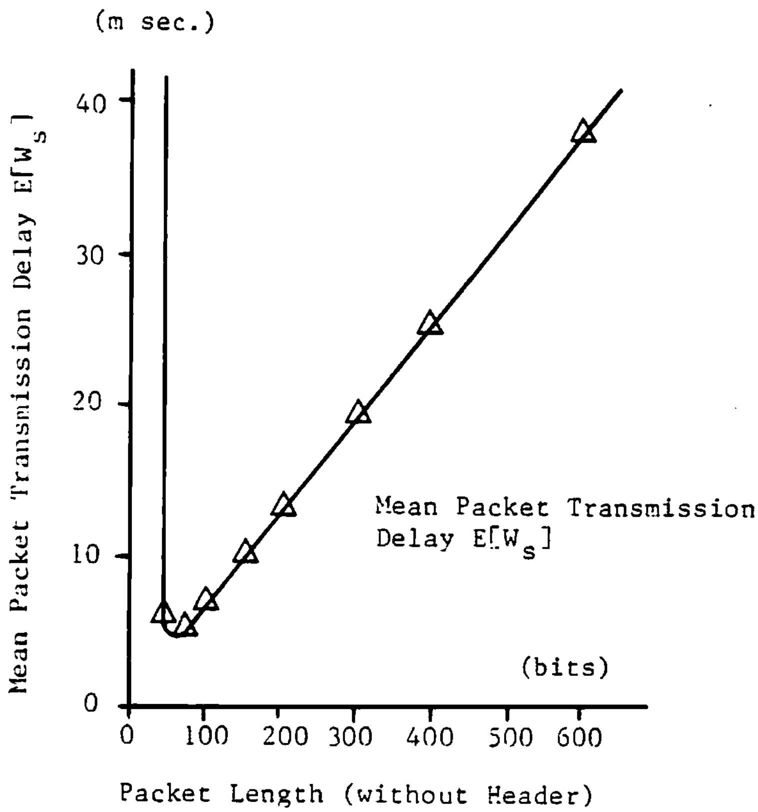


Fig.5 Mean Packet Transmission Delay $E[W_s]$ (1-Hop Model)

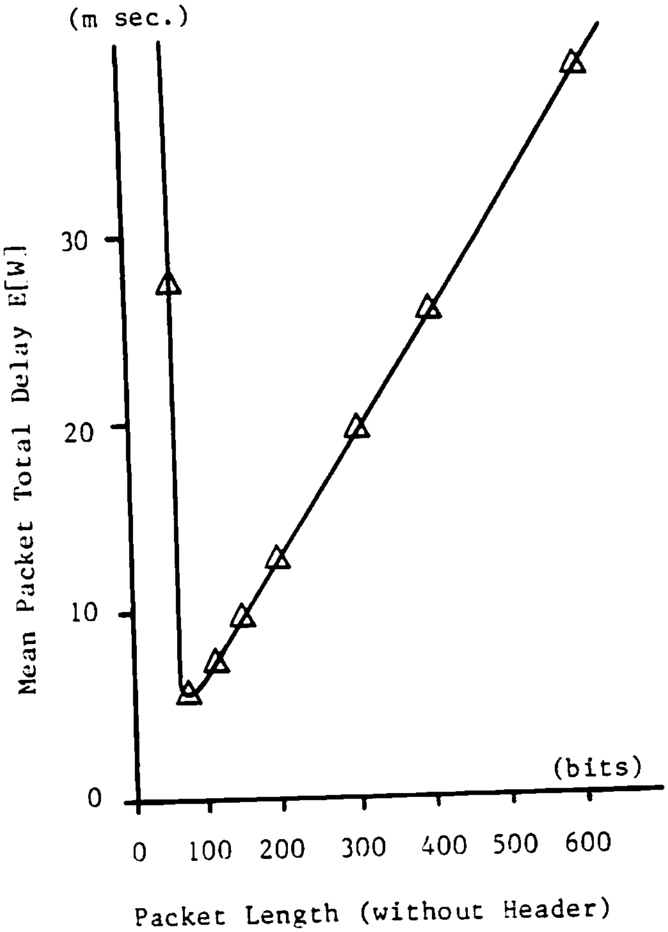


Fig.6 Mean Packet Total Delay in the N.T.I. Strategy (1-Hop Model)

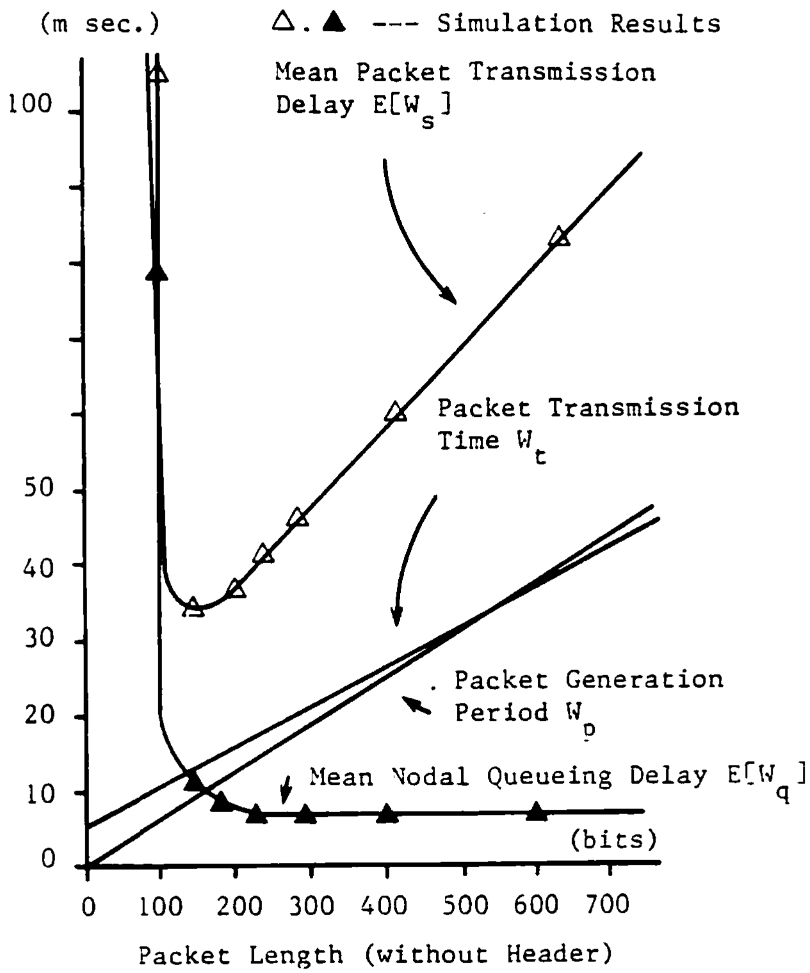


Fig.7 Voice Packet Delays (Network Model)

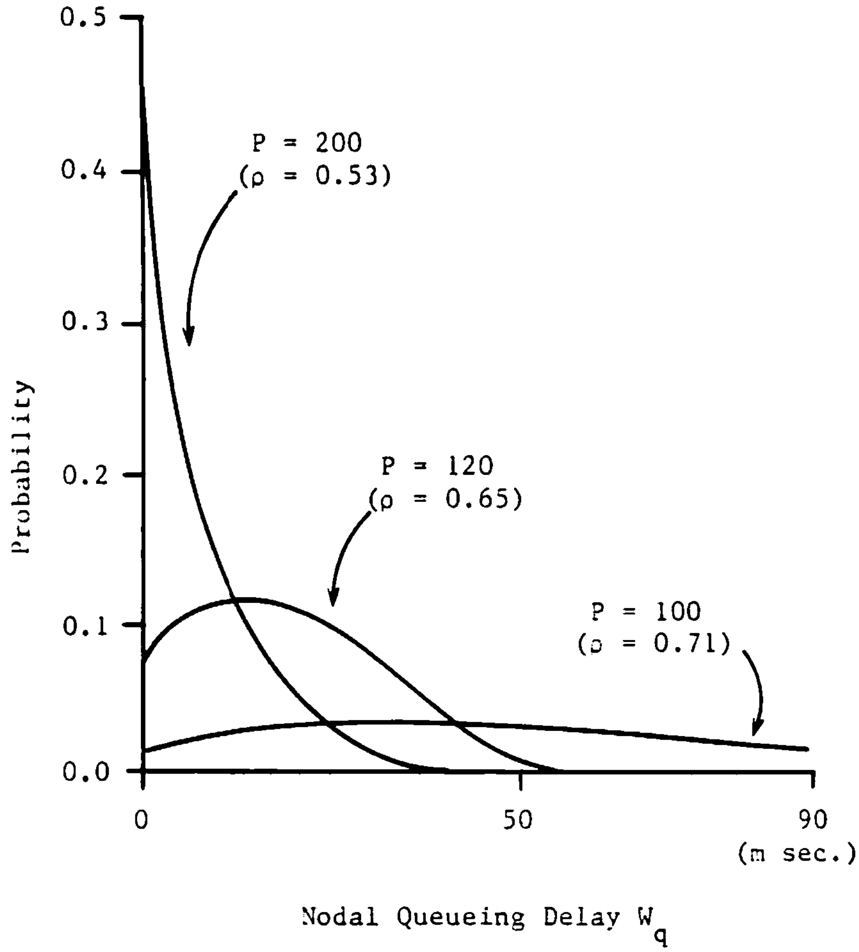


Fig.8 Density Function of Nodal Queueing Delay W_q (Network Model)

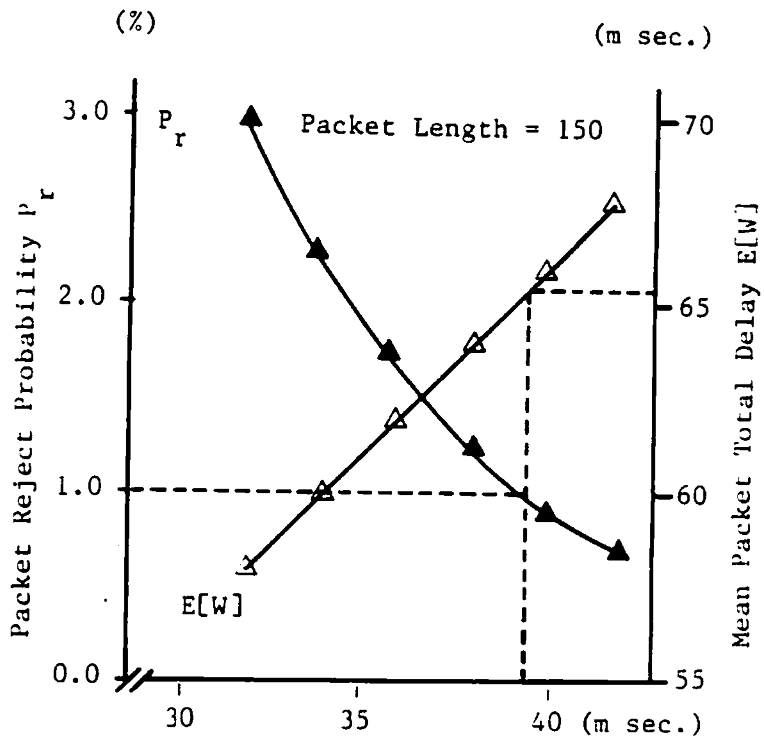


Fig.9 Packet Reject Probability and Mean Packet Total Delay in the N.T.I. Strategy (Network Model)

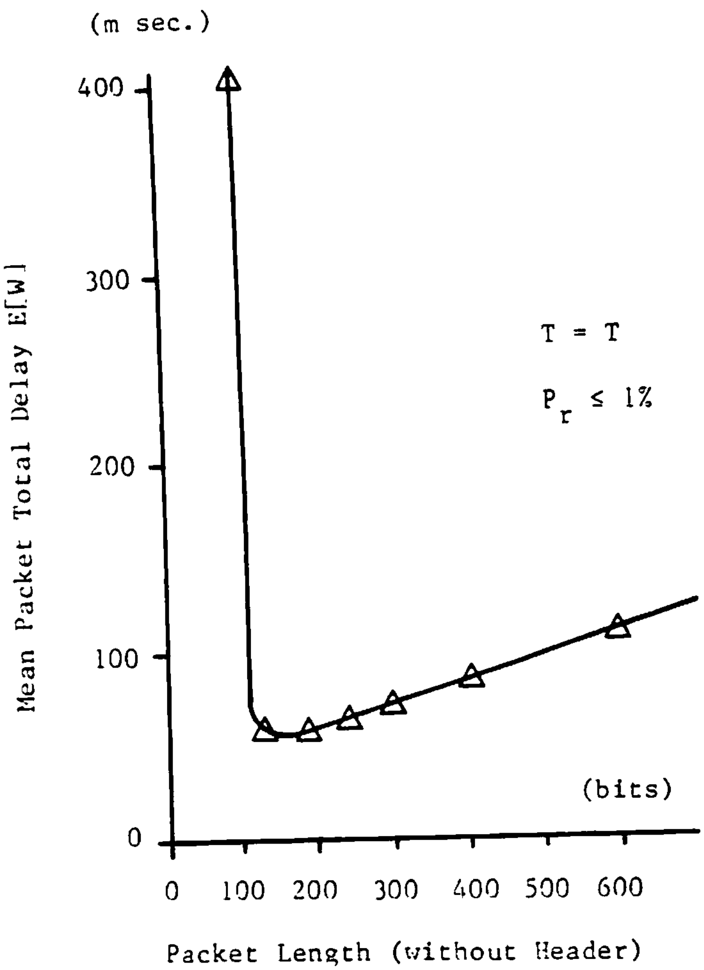


Fig.10 Mean Packet Total Delay in the N.T.I. Strategy (Network Model)

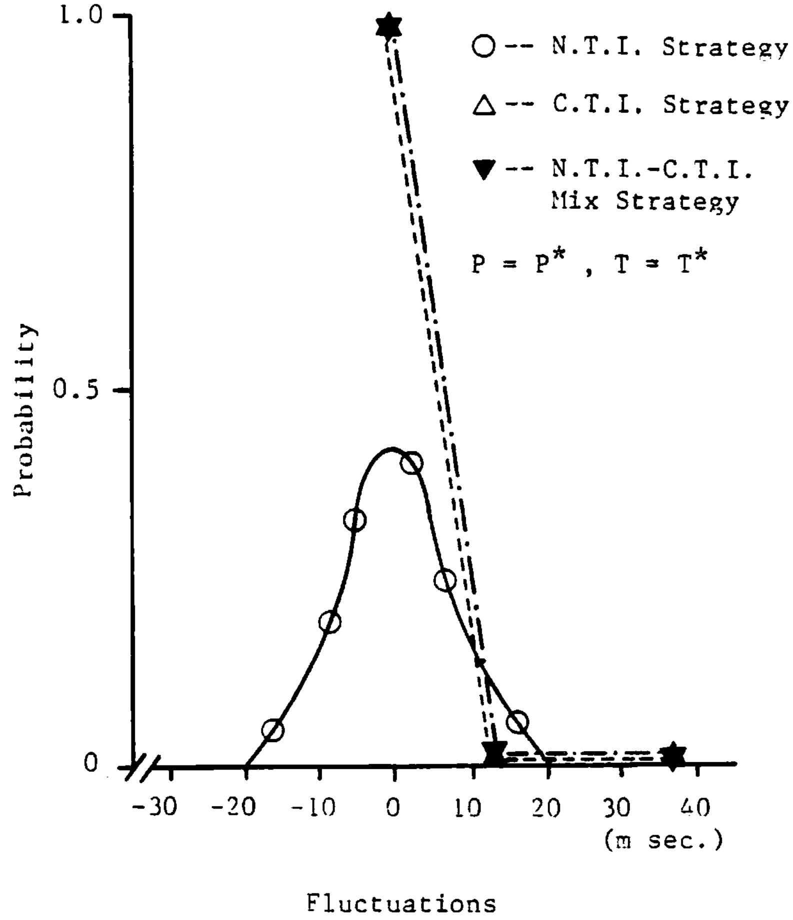


Fig.11 Density Function of Fluctuations Between Original and Played-out Silent Intervals (Network Model)

	N.T.I. Strategy	C.T.I. Strategy	N.T.I.-C.T.I. Mix Strategy
P* (bits)	200	200	200
T* (m sec.)	31	50	50
$E[W]^*$ (m sec.)	63.0	62.2	62.2

Tab.1 Optimal Values (Network Model)

	N.T.I. Strategy	C.T.I. Strategy	N.T.I.-C.T.I. Mix Strategy
Mean (m sec.)	0.228	0.195	0.195
Second Moment (m sec.^{2})	37.204	4.870	4.870

Tab.2 Fluctuations Between Original and Played-out Silent Intervals (Network Model)