Time Split-Token Based Contention Resolution (TS-TBCR) Protocol for Integrated Voice and Data Service over Wireless LAN's

Dhadesugoor R. Vaman & Sharad Kumar
Dept. of EE/CS
Stevens Institute of Technology
Hoboken, NJ 07030
Tel: (201) 216 5057
Fax: (201) 216 8246
Email: bis_dvaman@sitvxc.stevens-tech.edu

Abstract

This paper proposes a time split channel access with token based contention resolution (TS-TBCR) protocol for integrated voice and data applications on wirelesss LAN's. This protocol provides perfect scheduling of information by reserving a small bandwidth for contention resolution. This guarantees contention resolution before the start of an information slot thus preventing collisions and conserving bandwidth. For contention resolution a token based scheme requiring little centralized control from the base station is suggested. This protocol provides statistical multiplexing of voice and data traffic. Performance of the protocol for voice traffic was evaluated using a computer simulation. The results show that a large number of information sources could be multiplexed on the wireless channel keeping the voice packet loss probability within acceptable limits. The protocol provides bounded packet delay for voice sources and thus can be utilized for provision of integrated services.

1. Introduction

Most present day wireless networks are confined to providing voice and limited data communication facility. Wireless LAN is a relatively new concept that bridges the gap between mobile and cordless telephony and realizes the tremendous need for short range wireless communications in private and business environments. In contrast to wire-based LAN's which are mainly used for data transmission, wireless LAN's must support both voice and data traffic. Besides allowing easy reconfiguration without need for any recabling, the biggest plus points of a wireless LAN is its ready ability to connect to a high speed backbone network. This paper proposes a time-split channel access with token based contention resolution (TS-TBCR) protocol for voice/data integration applications on wireless LAN's.

Packet contention methods allow for multiplexing of a large number of terminals, each with a low average data rate and a high peak rate. Most of the contention resolution protocols used in LAN environments suffer from throughput degradations and long packet delays at heavy loads due to excessive collisions and thus can't be used for voice/data integration. Carrier Sense Multiple Access with Collision Detection (CSMA/CD) is widely used in wire-based LAN's. The CSMA/CD protocol requires collision detection and information transfer to take place on the same channel. This leads to collision of information packets which requires the transmitting terminal to transmit a bit sequence called 'jam' upon detecting the collision. This reduces the effective throughput on the channel.[1]

The 1-Persistent Dual Channel LAN (1P-DC-LAN)[2] overcomes this drawback to provide near-perfect scheduling throughputs for data transmission and bounded voice packet delays. This protocol employs two separate channels: one for scheduling and one for information transfer. It has been shown that using a separate channel for scheduling information transfer leads to better performance since collisions on the information channel are avoided. In this paper we present a variation of the 1P-DC-LAN protocol which is more suitable for voice/data integration on wireless LAN's.

Traditionally Time Division Multiple Access (TDMA) protocol has been used to provide packet radio networks with collision free communication and spatial reuse of the radio channel. Using TDMA for serving terminals with a low average data rate leads to idling of slots in the frame when they can be used by other terminals. Packet Reservation Multiple Access (PRMA)[3] allows multiplexing of a large number of terminals onto the same wireless access channel. PRMA is a combination of reservation ALOHA, R-ALOHA and TDMA. Terminals in active state contend for the channel using the ALOHA protocol and on successfully transmitting the first packet in a slot reserve that slot in subsequent frames. PRMA suffers from the drawback that in the event of a collision an entire slot is wasted.

This paper explores a multiple access protocol, TS-TBCR, for transmission of a mixture of packetized voice and data in a local wireless environment. This protocol provides perfect scheduling in the information slots using a dual channel concept and a token based distributed contention resolution scheme. More specifically the protocol defines two types of slots on the radio channel: scheduling slots and information slots. Scheduling slots are used for contention resolution to schedule transmission in the information slot. This scheme guarantees contention resolution before the start of an information slot thus preventing collisions and conserving bandwidth. This allows for multiplexing of a larger number of information sources on the wireless channel.

2. System Architecture

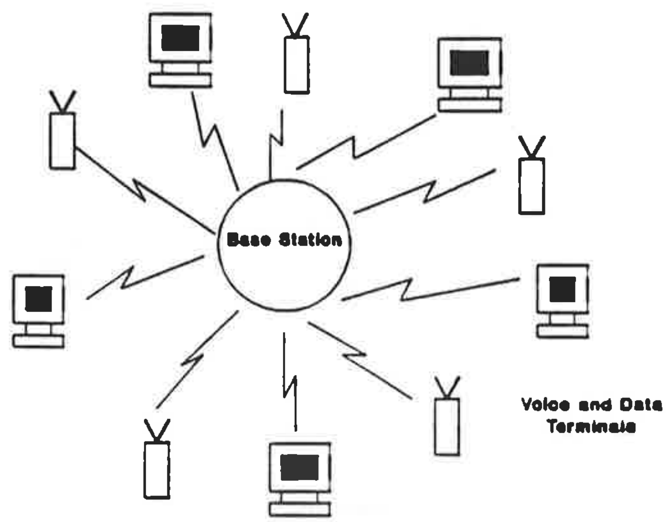


Fig.1 Configuration of a Wireless LAN

We consider a wireless LAN with a star topology as shown in Fig.1. All terminals transmit to a central base station (radio port) using a single frequency band. For simplicity of protocol description, the channel is assumed to be slotted in time. The system is assumed to be Frequency Division Duplex (FDD), i.e., downstream traffic (base station to portable) is assumed to be transmitted on a separate channel. The protocol calls for time frames with durations matched to periodic rate of voice traffic. Each frame is divided into slots that are dynamically used by active information sources.

Thus the TS-TBCR protocol allows for statistical multiplexing. As shown in Fig.2 the TS-TBCR protocol defines two type of slots within a frame: scheduling slots and information slots.

An information slot is always preceded by a fixed number of scheduling slots which are used to schedule the transmission in this information slot. The TS-TBCR protocol guarantees contention resolution before the start of an information slot and thus maximizes bandwidth by allowing only one terminal to transmit in the information slot. However the penalty is the extra bandwidth required for scheduling. The size of the scheduling and information slots is fixed. The information slot is used by portables to transmit speech, data, control, and signalling information. A contention cycle is defined as a set of scheduling slots and the following information slot. Each contention cycle starts with a 8-bit preamble being transmitted by the base station which is used for synchronization of the portables. There is no contention on the downstream channel. In the TS-TBCR protocol, each portable possesses two tokens: a dynamic token and a static token. While the dynamic token of a portable is a function of the delay experienced by a packet waiting to be transmitted, the static token is assigned at the time of registration with the base station. This paper doesn't highlight the static token assignment procedure, but only concentrates on the channel access scheme assuming that each portable has already been assigned a static token.

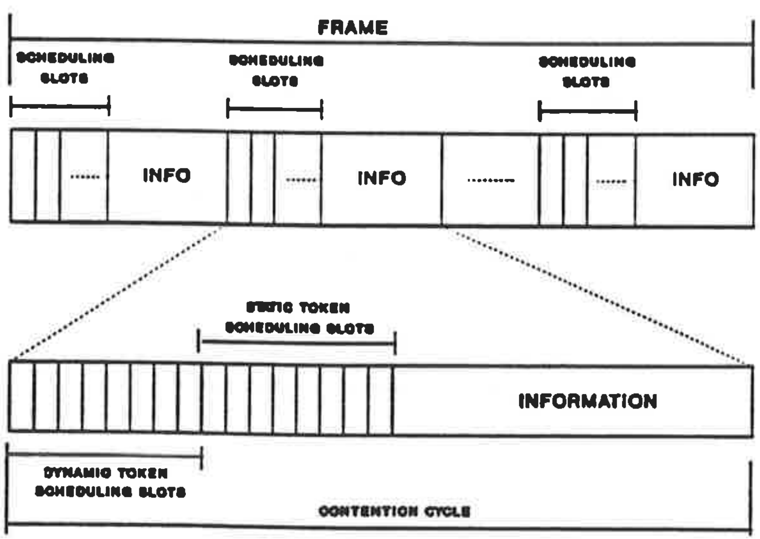


Fig.2 Type of Slots on the Channel

The short roundtrip propogation delay, around ten microseconds, associated with a local wireless environment makes it possible for terminals to obtain instantaneous response from the base station [4]. Each terminal is equipped with a voice activity detector, a 32 kbps voice coder, and a packetizer. Each packet consists of speech and other non-speech information. Based on the maximum delay allowed by the network, only D_{max} ms of speech can be buffered at each terminal. Thus, any voice packet that can not be transmitted within D_{max} ms is discarded. In contrast the buffer size for data packets can be quite large which causes large delays rather than packet loss.

Fig.3 shows in detail the architecture of the scheduling slots. The scheduling slots are also divided into two parts. The first part is used for contention resolution based on the dynamic token and second based on the static token. The architecture defines 16 scheduling slots in each contention cycle. First 8 of these slots are used for the dynamic token and the next 8 for the static. Each of these scheduling slot is equivalent of an 8-bit period. Each of these 8-bit periods are further divided into two parts, T and R. First half of each of these slot (4-bits) is used for transmission and the second for obtaining a response from the base station. The portable only transmits in the first 4-bit period and receives in the following 4-bit period. This architecture assumes that the small hand-held portables are capable of transmitting and receiving at the same time.

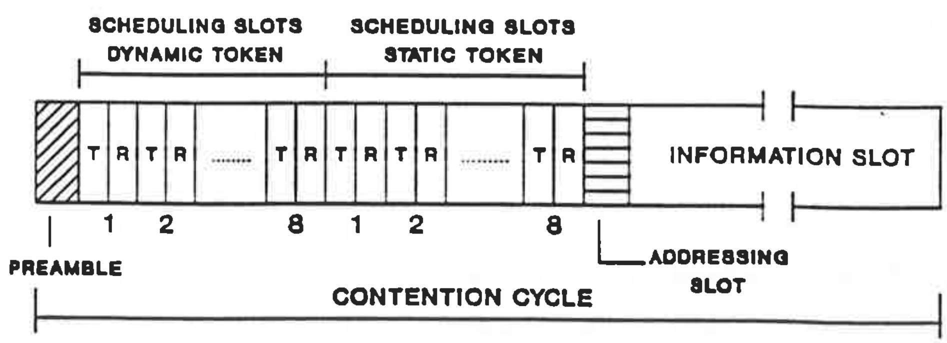


Fig.3 Architecture of Scheduling Slots

3. Token Based Contention Resolution

The contention slots are used to schedule the transmission in the information slot. Since both the dynamic and static tokens are 8-bit period long, there are 8 scheduling slots for dynamic token and 8 for the static token in one contention cycle.

3.1 Dynamic Token Based Contention Resolution

The dynamic token possesed by each portable is a function of the waiting time of the buffered packet. Therefore as packet delay increases so does the token value. The token can be thought of as priority, the higher the token value the higher the priority. In this scheme the system gives priority to packets with the longest waiting time. Maximum allowable delay for a speech packet is denoted by D_{max}, i.e., any voice packet that cannot be transmitted within D_{max} ms is discarded. Each portable has a timer associated with it which keeps track of the of the delay of the pending voice packet. Besides a timer each portable also has a temporary n-bit counter corresponding to the n-bit dynamic token. On the arrival of a voice packet, both the counter and the token are initialized to zero. An n-bit counter(token) can define 2^n levels, therefore an 8-bit counter can represent 256 levels (0-255). The n-bit counter is incremented by one every $T_n = D_{max}/2^n$ ms. A speech packet is discarded T_n ms after the counter has reached its maximum value, i.e., all 1's. At the beginning of each contention cycle the counter value is downloaded into the dynamic token which is then used by the portable for contention. Thus, the dynamic token possesed by each portable changes only at the beginning of the contention cycle.

The contention resolution proceeds as follows:

1) A portable having an information packet to transmit, i.e., data, speech or signalling information, waits for the beginning of a contention cycle and then synchronizes on the preamble. As mentioned earlier, each half contention slot is equivalent of a 4-bit period. In the first half of the scheduling slot-1 each of these portables map the most significant bit of their 8-bit binary token, transmitting a busy tone for "1" and not transmitting otherwise.
2) After receiving the transmission from these portables the port takes the following action: The port detects the signals, interpreting a bit as a "0" if it receives no signal and "1" otherwise. The resulting binary value is transmitted in the next 4-bit period, i.e., second half of the scheduling slot-1.
3) On receiving this bit pattern, all portables compare the most significant bit of their token with that of the received pattern. If the received signal is interpreted as a "1", all portables which didn't transmit a busy tone in the most significant bit period backoff based on some predetermined policy.

All the remaining portables contend in the next contention slot repeating the above procedure except they compare the second most significant bit of their token with that of the received pattern. The portables keep on contending till the last slot where they compare the least significant bits. If a match occurs, the portable has won the contention based on the dynamic token.

There is a finite probability that two or more portables have the same token while contending in the last scheduling slot reserved for the dynamic token. This is only possible if two or more portables receive a voice packet within a T_n ms interval. The probability associated with this event is shown in the Appendix. Only those portables who win contention based on the dynamic token are allowed to contend based on the static token in the following scheduling slots of the contention cycle. Shown below is a representative example of contention resolution between three portables based on the dynamic token.

	Dynamic Token
Portable A:	1 0 1 1 1 0 0 1
Portable B:	0 1 1 0 0 0 1 1
Portable C:	1 1 0 0 0 1 1 0

In the 'T' part of scheduling slot 1 each portable maps it most significant bit onto the slot as follows:

Portable A transmits: 1 1 1 1
Portable B transmits: 0 0 0 0
Portable C transmits: 1 1 1 1

Base station interprets: 1 1 1 1
Returns this value in the corresponding 'R' part of the slot

After receiving the response from the base station each portable compares it to the value it transmitted: Portable C transmitted a '0' but receives a '1', and thus it backsoff. In scheduling slot 2 only portables A and C contend, mapping their second most significant bit onto the slot.

Portable A transmits: 0 0 0 0
Portable B transmits: 1 1 1 1

Base station interprets: 1 1 1 1
Returns this value in the corresponding 'R' part of the slot

Since the bit value received by portable A doesn't match the one it transmitted, it also backsoff. Thus portable C has won the contention and would transmit in the corresponding information slot.

3.2 Static Token Based Contention Resolution

Only those terminals who have won contention based on the dynamic token are allowed to contend with their static token. Thus number of stations contending at this point are significantly lesser than those at the beginning of the contention cycle. Each portable already has a unique 8-bit static token assigned to it. The static token contention proceeds exactly as the dynamic token contention. There are a total of 8 scheduling slots for the static token contention. Since static tokens are unique it guarantees contention resolution. That is, at the end of these scheduling slots there is only one station left to transmit. This is the station that possesses the largest dynamic and static token. This leads to perfect scheduling in the information slot and prevents any collision thus maximizing bandwidth.

Terminals who backoff during this contention cycle contend in the next contention cycle in a similiar manner.

4. Simulation Methodology

A computer simulation was performed to evaluate the performance of the TS-TBCR protocol in a local wireless environment with a channel rate of 720 kb/s. At this point only voice traffic was simulated on the network.

4.1 Frame Format

The bit rate on the channel is assumed to be 720 kb/s. With a speech coding rate of R kb/s, a portable in an active state generates a K bit speech packet every T_g ms, where $K=R.T_g$. Therefore each frame duration on the channel is T_g ms. We assumed a 32 kb/s speech coding rate. Depending upon T_g, a portable transmits a 512 bit packet every 16 ms, or a 1024 bit packet every 32 ms. In addition, there is an N bit header associated with each speech packet. Therefore each information slot contains $K+N$ bits. These N bits correspond to signalling and other non-speech information. In our simulations, we considered two values of K, 512 and 1024 corresponding to generation times of 16 and 32 ms respectively [5], assuming a 64 bit overhead for both cases.

Each contention cycle starts with a 8-bit preamble, followed by 16-8 bit (128 bits) slots for scheduling, and a 8-bit addressing slot. Therefore each slot length is actually K+N+144 bits. A channel rate of 720 kb/s gives 16 and 18 information slots for generation times of 16 and 32 ms respectively.

4.2 Speech Model

Typical behaviour of a voice source is modelled as a two state process. A voice source is "active" when a talker is speaking, and "inactive" when the speaker is silent. In the active state, the voice source generates fixed length packets periodically [6][7]. No packets are generated in the inactive state. A voice source alternates between the two states. It has been shown that the active and inactive durations can be approximated by exponential distributions. Fig.4 shows a two state model of a voice source. Let P_a denote the probability of being in the active state, and P_i the probability of being in the inactive state. The activity factor then defined as P_a/P_i, is the proportion of the time a source is in the active state.

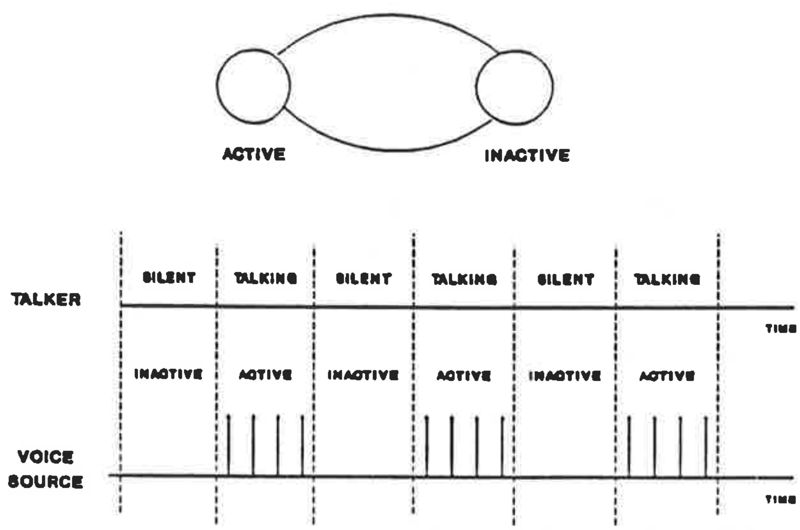


Fig.4 Voice Source Behavior

For our simulations, we used an activity factor of 0.36. Active and inactive periods were assumed to be exponentially distributed with average durations of 0.36 and 0.64 sec. respectively. Voice sources in active state generate a packet every T_g ms. T_g is a design variable in our simulations. At each terminal, information from the voice encoder is combined with the header to form a packet to be transmitted.

4.3 Expected Value and Standard Deviation of Delay

In the TS-TBCR protocol portables in talkspurt don't reserve a slot in every frame. A portable has to contend for each and every packet it needs to transmit. Therefore voice packets don't get transmitted in the same slot every frame and there is a finite variance of delay associated with each speech packet, as opposed to TDMA or PRMA protocols. Simulations were performed to evaluate the expected value and standard deviation of voice packet delay.

5. Simulation Results

The performance of the TS-TBCR protocol was evaluated only for voice traffic. In all the simulations a maximum of 50 simultaneous conversations were realized using the speech model. A speech activity factor of 0.36 was assumed throughout the course of our simulations. All simulations were performed for a coding rate of 32 kb/s. About 12 min. of speech from each voice source was simulated. One of the parameter of interest was the maximum number of simultaneous conversations that could be supported with a packet dropping probability, P_{drop} of less than 1 percent. This quantity is denoted by $S_{0.01}$. The effect of different values of T_g and D_{max} on $S_{0.01}$ was also studied. Voice packet delay as a function of number of simultaneous conversations and T_g was another parameter of interest.

For the same channel rate of 720 kb/s, a TDMA based wireless system with the same packet overhead can support 16 simultaneous converations with no packet dropping. For a generation time of 16 ms, the TS-TBCR protocol can support about 34-35 simultaneous conversations, i.e., $S_{0.01}$=35. Fig. 1 depicts P_{drop} as a function of number of simultaneous conversations on the channel. Fig. 2 shows that for a generation time of 16 ms PRMA protocol gives $S_{0.01}$=37 which is slightly better than that for TS-TBCR protocol [3][4]. Fig.2 shows a comparison of P_{drop} for the two protocols.

5.1 Effect of Speech Packet Generation Time on P_{drop}

The speech packet generation time, T_g was also varied to observe its effect on P_{drop}. Two values of T_g were used, 16 ms and 32 ms, and the corresponding frame sizes were also 16 and 32 ms. Fig.3 shows packet dropping probability as a function of number of simultaneous conversations and T_g. P_{drop} is seen to have dramatically reduced with increased T_g. This is due to the lower percentage of scheduling overhead which leads to 18 slots on the channel instead of 16. $S_{0.01}$=35 for T_g=16 ms and 40 for T_g=32 ms which is a marked improvement. The corresponding value for the PRMA protocols are 37 and 30 respectively.

Therefore the TS-TBCR protocol is seen to favor larger packet generation times over shorter ones.

5.2 Effect of Maximum Allowable Delay on P_{drop}

D_{max} is the maximum allowable delay for a speech packet, i.e., if a packet can't be transmitted within D_{max} ms it is discarded from the buffer. From Fig.4 P_{drop} is seen to improve as D_{max} is increased from 16 to 32 ms. $S_{0.01}$ is approximately 36-37 for D_{max}=32 ms.

5.3 Expected Value and Standard Deviation of Packet Delay

Fig.5 shows expected voice packet delay, E_d as a function of number of simultaneous conversations on the channel and the packet generation time. Again we see a marked reduction in E_d as the packet generation time is increases to 32 ms. For both cases, $S_{0.01}$=35 for T_g=16 ms and $S_{0.01}$=40 for T_g=32 ms, E_d is around 4 ms.

From Fig.6 the standard deviation of voice packet delay is observed to be about 4 ms for both cases, i.e., $S_{0.01}$=35 for T_g=16 ms and $S_{0.01}$=40 for T_g=32 ms.

6. General Discussion

Altough the performance of the protocol has only been evaluated for voice traffic, it is expected to perform well in an integrated voice/data environment. This protocol inherently provides statistical multiplexing and hence treates both voice and data similarily. It can give higher priority to voice traffic by allowing voice terminals to increment their dynamic tokens at a faster rate as compared to terminals serving data traffic. Also voice terminals can be assigned higher static tokens. Contention resolution is solely based on tokens possesed by the terminals at the time of contention irrespective of the type of information to be transmitted. This prevents data traffic from being penalized.

At this point it is important to compare the TS-TBCR protocol with other schemes suggested for local wireless environments. Since TDMA reserves a slot in every frame for a communicating terminal it leads to idling of channel if terminals being served have a low average data rate. Thus it is not suitable for voice/data integration. Another multiple access technique suggested is PRMA. Since PRMA is a reservation based protocol it provides high priority to voice by restricting data traffic to access only those slots that are not being used by voice traffic. Therefore if a large number of voice terminals are using the channel, it could lead to unacceptable delays for data traffic. This problem can be extremely acute if the data messages are signalling messages, e.g., call setup messages. This would result in relatively large call setup times which are generally unacceptable in a local environment. On the other hand the TS-TBCR protocol can provide higher priority to signalling messages by assigning them the highest dynamic token. The performance of this protocol for a mixture of packetized voice and data is presently under study.

7. Conclusions

The TS-TBCR protocol analyzed here allows multiplexing of a large number of voice sources on the channel keeping packet losss probability within acceptable limits. The protocol can support about 35-40 simultaneous conversations for $P_{drop} <= 0.01$. The protocol provides bounded voice packet delays. The simulations also show that the protocol favors larger packet generation times as opposed to smaller ones. The protocol performs as well as, and in some cases even better than PRMA for voice traffic. Overall TS-TBCR protocol seems to have a tremendous potential for voice/data integration applications in local wireless environment.

REFERENCES

[1] W. Stallings, *Data and Computer Communications*. MacMillan, 1988, ch. 11, pp. 349-361.

[2] D. R. Vaman, C. T. Wu, and K. Lu ", 1-Persistent dual channel local area network for integrated voice and data application", *Int. J. Dig. Anlg. Cab. Sys.*, vol. 1, pp. 95-106, 1988.

[3] D. J. Goodman and S. X. Wei, "Efficiency of packet reservation multiple access," *IEEE Trans. Vehic. Technol.*, vol. 40, no. 1, pp. 170-176, Feb 1991.

[4] D. J. Goodman, R. A. Valenzuela, K. T. Gayliard, and B. Ramamurthi, "Packet reservation multiple access for local wireless communications," *IEEE Trans. Commun.*, vol. 37, pp. 885-890, Aug. 1989.

[5] C. J. Weinstein and J. W. Forgie, "Experience with speech communication in packet networks," *IEEE J. Select. Area Commun.*, vol. SAC-1, pp. 963-980, Dec. 1983.

[6] J. N. Daigle and J. D. Langford, "Models for analysis of packet voice communications systems," *IEEE J. Select. Area Commun.*, vol. SAC-4, no. 6, pp. 847-855, Sept. 1986.

[7] P. T. Brady, "A model for generating on-off speech patterns in two way conversations," *Bell Syst. Tech. J.*, vol. 48, pp. 2445-2472, Sept 1969.

[8] S. Ross, *A First Course in Probability*, Third Edition. New York: Macmillan, 1988, ch. 6, pp. 225-226.

Fig.5 Packet Dropping Probability as a function of number of simultaneous conversations for TS-TBCR Protocol

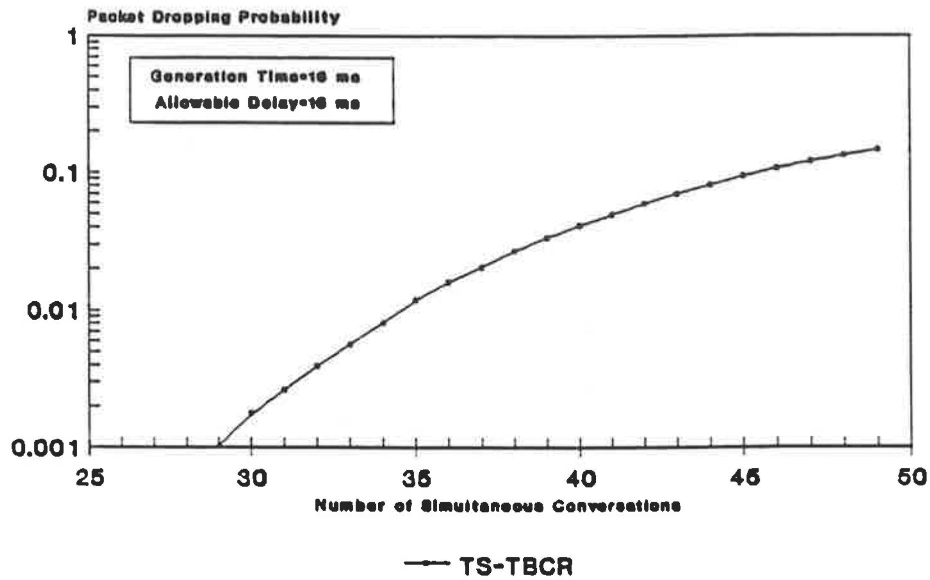


Fig.6 Performance Comparison of TS-TBCR protocol and PRMA protocol

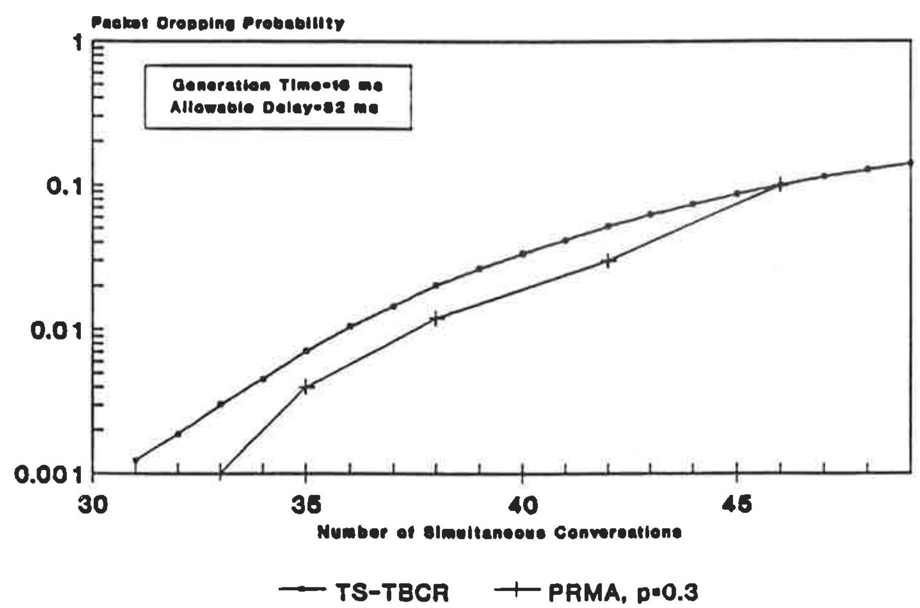


Fig.7 Packet Dropping Probability as a function of number of simultaneous conversations and packet generation time

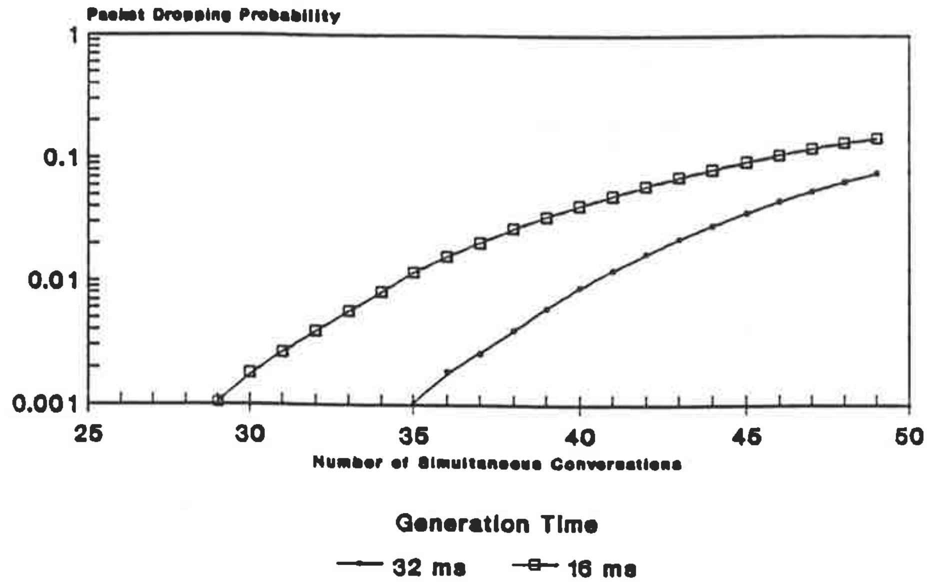


Fig.8 Packet Dropping Probability as a function of number of simultaneous conversations and allowable packet delay

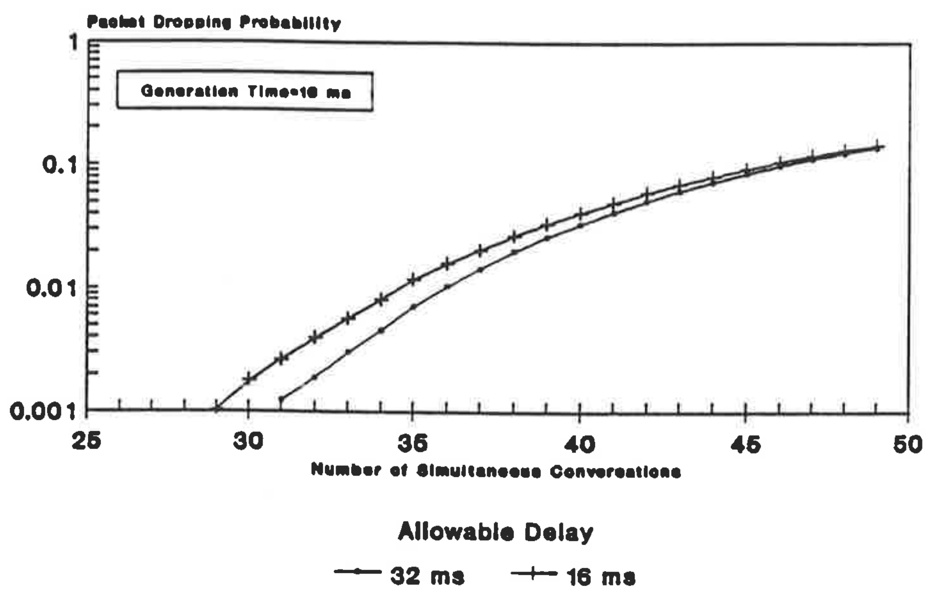


Fig.9 Expected Voice Packet Delay as a function of number of simultaneous conversations and generation time

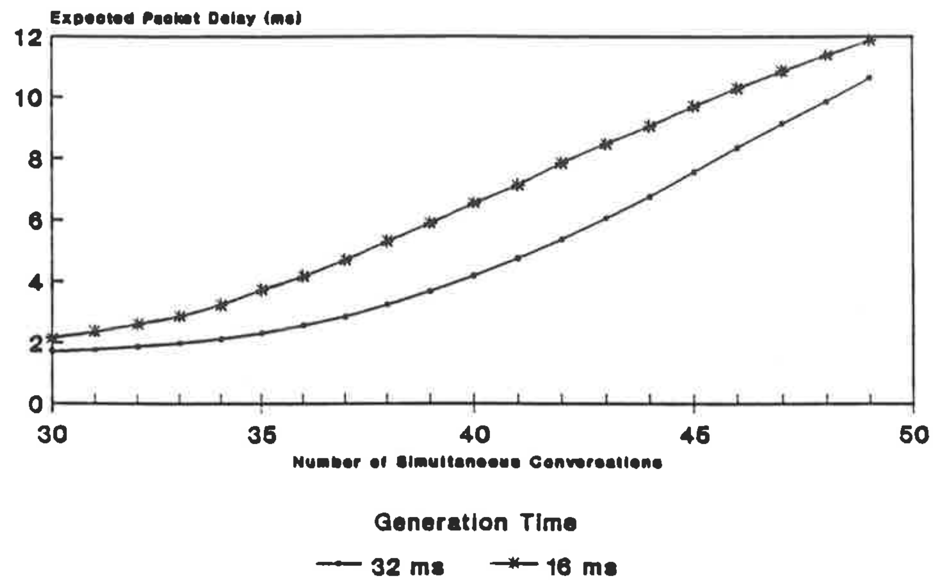
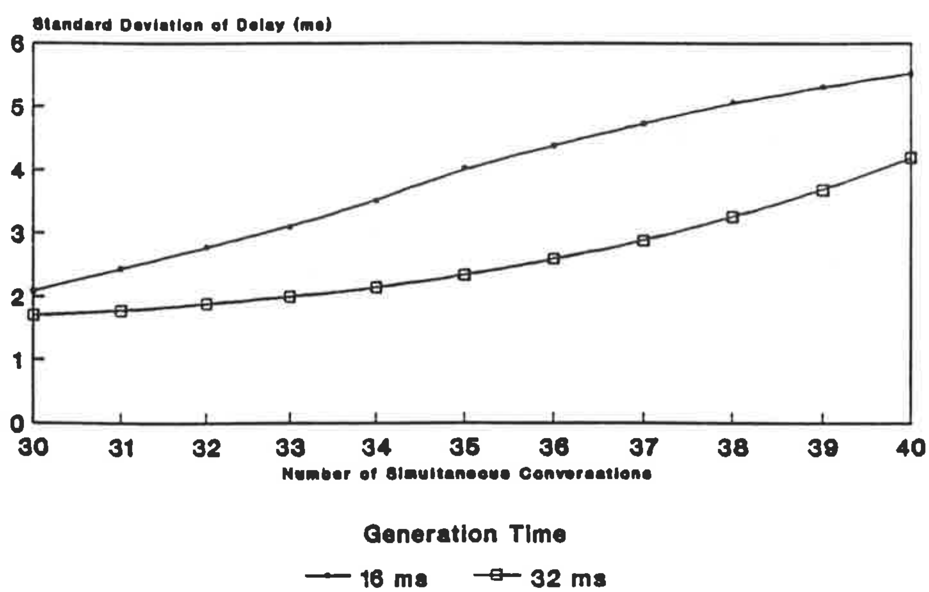


Fig.10 Std. Deviation of Voice Packet Delay as a function of number of simultaneous conversations and Generation Time



APPENDIX

The probability that no two portables possess the same dynamic token at the time of contention is equivalent to the probability that no two portables receive voice packets within a T_n ms interval. Since in the "active" state each source generates fixed length packets periodically, this probability is the same as the probability that no two sources begin their talkspurt within T_n ms of each other.

Let S denote the total number of conversing voice sources and N the number of sources active in one frame duration, T_g ms, where N <= M. Therefore N voice packets are observed in one frame duration. Let X_1, X_2, ..., X_N denote the ordered arrival times of these N voice packets. Then,

$$0 \le X_1 \le X_2 \le X_3 \le \ldots\ldots\ldots\ldots \le X_N \le T_g$$

Since each source spends an exponentially distributed amount of time in both "active" and "inactive" states, the arrival time of the first packet of the talkspurt from each source will be uniformly distributed in the interval $(0,T_g)$.

From [8],

$$P_r\left(\frac{different\ dynamic\ tokens}{N\ active\ sources}\right) = [1-(N-1)\frac{T_n}{T_g}]^N$$

The desired probability is obtained by unconditioning over N

$$P_r(different\ dynamic\ tokens) = \sum_{N=1}^{S}[1-(N-1)\frac{T_n}{T_g}]^N \cdot P_r(N\ active)$$

The probability of N out of M speech sources being in "active" state is given by

$$P_r(N\ active) = \binom{S}{N} P_a^{\,N}(1-P_a)^{S-N}$$

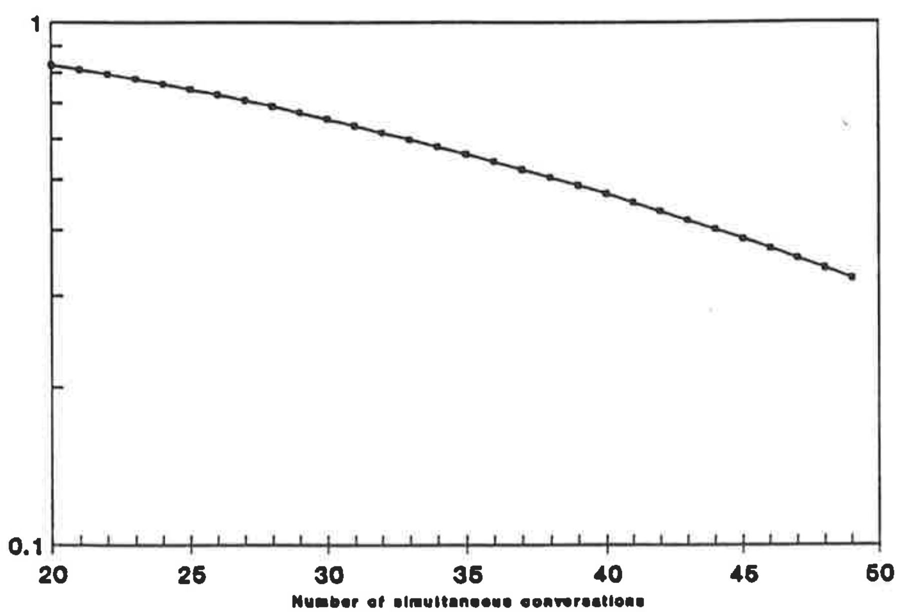
15

where P_a is the probability of a source being in the active state as defined in section 4.2

$$P_r(\textit{same dynamic token}) = \sum_{N=1}^{S} [1-(N-1)\frac{T_n}{T_g}]^N \cdot \binom{S}{N} P_a^N (1-P_a)^{S-N}$$

For a 8 bit dynamic token with a frame generation time of T_g=16 ms, T_n=0.0625 ms, the activity factor P_a=0.36, and the total number of conversing sources S=50, the desired probability is shown in Fig.11.

Fig.11 Probability of no two portables having the same dynamic token



From the figure we observe that for S=35, there's a 55% probability of contention being resolved solely on the basis of the dynamic token.