Voice and Data transmission over an 802.11 Wireless network

Matthijs A. Visser[1]
visser@dutetvd.et.tudelft.nl

Magda El Zarki[1,2]
magda@ee.upenn.edu

[1]Department of Electrical Engineering, Delft University of Technology, Mekelweg 4, 2628 CD Delft, The Netherlands
[2]Department of Electrical Engineering, University of Pennsylvania, 200 South 33rd Street, Philadelphia, PA,

Abstract — This paper analyzes the transmission of voice and data over an 802.11 Wireless Local Area Network (WLAN). The data is transmitted in a contention based access period, while the voice samples are transmitted during a contention free period, based on a polling scheme. Because statistical multiplexing can be utilized, speech may be outdated when a poll arrives. The portion of outdated speech is then clipped to decrease the load on the channel. We analyze the quality of the voice conversations in terms of the percentage of bits clipped as well as the throughput of the data for various parameters. We show the boundary conditions involved in the transmission of voice over the WLAN and demonstrate the impact of a time-bounded service on the throughput during the contention period. The results show that the high overhead introduced by the 802.11 WLAN standard results in a low number of possible voice conversations. It can also be concluded that the cooperation of the contention based and contention free periods results in a poor performance. Further, variation of the maximum payload size reveals that the largest possible maximum payload size must be selected to minimize the percentage of clipped bits and maximize the throughput. Finally, we show that a larger Superframe length provides the opportunity for more voice conversations or a higher data throughput, but requires increasing the Time To Live for the speech bits to retain an acceptable quality.

1. Introduction

Currently the IEEE 802.11 committee is developing a standard for a Wireless Local Area Network (WLAN) [1, 2]. This standard uses Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) as the basic channel access protocol; called the Distributed Coordination Function (DCF). In addition to the DCF, the standard specifies a Point Coordination Function (PCF) which utilizes a polling scheme to grant stations access to the channel. These two channel access schemes are discussed in more detail in section 2.

As a result of the polling scheme the PCF will be able to offer a time-bounded service for real-time sources, as long as the maximum length of the polling list can be controlled. Possible real-time applications are, for instance, video conferencing and voice.

The 802.11 standard currently specifies bandwidths of 1 Mbps and 2 Mbps. Due to these low bit rates, the time-bounded source has to be low bit rate as well. A speech source with Speech Activity Detection (SAD), resulting in an on-off pattern, with a constant bit rate during the on period, is a good candidate for an application utilizing the PCF. The speech generation model will be further discussed in section 3.

Due to the on-off nature of the speech source we employ statistical multiplexing to increase the number of speech conversations that can be accommodated. As a result of the statistical multiplexing and the variance in the speech patterns of the different stations, we encounter situations in which a station has to wait longer for a poll than the average inter-poll time. If the waiting time exceeds a threshold, called the Time To Live (TTL), we use a method known as speech clipping to discard speech when transmission capacity is not available.

2. 802.11 Wireless LAN MAC Protocols

In this section we describe the two medium access schemes that are used by an 802.11 WLAN. The information given in this section is completely based on the P802.11/D1 Draft standard [1]. Because it is a draft standard, certain parts may still be changed, although we expect that the basic access protocols described here will not undergo any further modification.

2.a Distributed Coordination Function

The fundamental access method of the 802.11 standard, the DCF, is present in all stations within an 802.11 network. In addition to the CSMA/CA protocol there is a random backoff algorithm. This algorithm

This paper previously appeared in the Proceedings of the IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC '95), Toronto, Canada, September 1995.

The research for this paper has been performed at the University of Pennsylvania as part of Matthijs Visser's Masters thesis.

reduces the collision probability between multiple stations accessing the medium at the point where a collision would most likely occur, just after the medium becomes free following a busy medium.

The basic access method employed by the DCF is shown in figure 1. The various Inter Frame Spaces (IFS) are used to differentiate between different types of transmissions. Before each transmission a station has to wait DIFS in order not to disturb other ongoing transmissions, which might only be separated by a Short Inter Frame Space (SIFS) or a Priority Inter Frame Space (PIFS). A complete Medium Access Control (MAC) Protocol Data Unit (MPDU), for instance, existing of a Data Frame and an Acknowledgment (ACK) frame, is only separated by a Short Inter Frame Space.

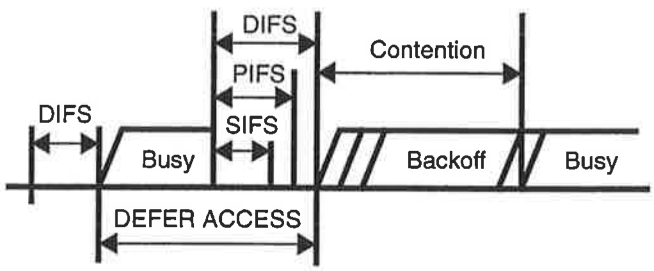

Figure 1 Basic Access mechanism.

2.b Point Coordination Function

The Point Coordinator (PC) invokes the Point Coordination Function (PCF), which generates the Superframe that sets the framework for the access procedures. It is divided into a Contention Free Period, with the PCF Access Procedure, and a Contention Period, with the DCF Access Procedure as shown in figure 2. When the PCF is present all stations within the region where the PCF is active obey the PCF rules.

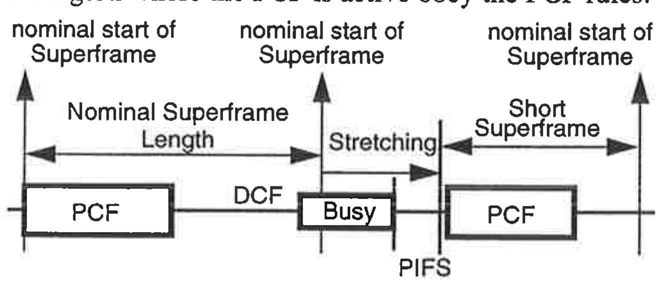

Figure 2 Superframe structure.

As a result of the CSMA/CA protocol, the PC might be unable to gain control of the channel at the nominal beginning of the Superframe. If, for instance, a station starts a transmission during the DCF period which lasts longer than the remaining time between the start of the transmission and the nominal start of the next superframe, the PC has to defer the start of its transmission until the medium has been free for a PCF Inter FRame Space (PIFS). This mechanism is shown in figure 2.

The event in which the Contention Free period starts later than the nominal start of the Superframe is called stretching. This stretching imposes a boundary condition on the length of a Superframe and on the length of the Contention Free period. This relationship is shown in figure 3.

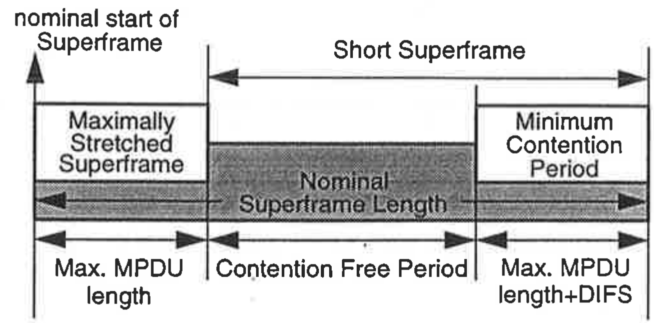

Figure 3 Effect of Superframe stretching.

After the PC has gained control of the channel it starts transmitting CF-Down frames according to its polling list. If a station has a CF-Up frame queued, then it transmits this frame after a SIFS period. If a station does not have a frame queued for transmission the PC will transmit its next CF-Down after waiting for a PIFS. This mechanism is shown in figure 4.

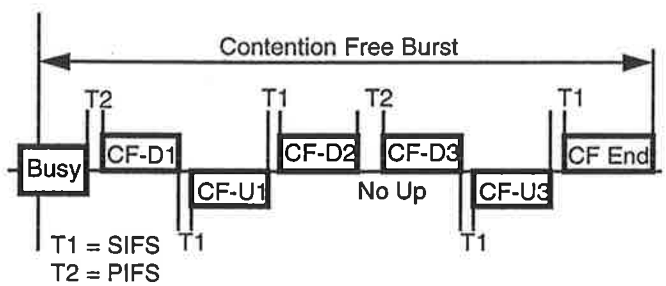

Figure 4 PCF Access Procedure.

3. Traffic Models

3.a Speech Model

For the speech source we use a simple on-off speech model, as described in [7, 8]. This speech model ignores events such as mutual silence, double talk, etc., as described in [3, 4]. It can be represented by a two state Markov model with a silent state and a talking state. In our model we use the parameters specified in [8], $T_{TLK} = 1$ s and $T_{SIL} = 1.35$ s. Therefore, the percentage of time spent in the silent state is $1.35 / (1+1.35) * 100\% \approx 57\%$ and the percentage of time in the talking state is $1/(1+1.35) * 100\% \approx 43\%$.

To improve our speech model we made two modifications described in [3, 4, 5]. First, we fill in silence gaps shorter than 200 ms because these are a result of stop consonants instead of the end of a talk spurt which effectively shifts the probability distribution of the silence gaps 200 ms to the left. In order to comply with an average silence period of 1.35 s we use an average of 1.15 s for our negative exponential distribution of the silence periods and add 200 ms afterwards. Second, we consider talk spurts shorter than 15 ms to be impulse noise. From the negative exponential distribution we calculate that with an average of 1 s only 1.48% of the talk spurts will be shorter than 15 ms. We consider this to be very small and do not change the probability distribution function of the talk spurts.

3.b Data Model

To see the effect of variable stretching on the transmission of speech over an 802.11 WLAN we send traffic during the DCF part of the Superframe (see figure 2). To model this data traffic we use a negative exponential distribution for the inter arrival times of the data frames, as well as for the length of the data frames.

The Superframe structure imposes a boundary condition on the length of the data frames, as will be explained in section 4. Therefore, the payload length of the data frames might have an upper limit which is smaller than the maximum payload of 2304 bytes as specified in the 802.11 standard. We enforce this upper limit by reducing the payload size to the maximum payload if our random number generator selects a frame length which is too long. Because we want to retain a negative exponential distribution for the payload length, we choose the average payload length such that no more than 5% of the generated payload lengths will have a length greater than the maximum.

The load of the data traffic on the channel will be 98%, including overhead based on the percentage of bandwidth available for data traffic. If, however, the Contention Free period ends before its maximum length then the effective load might be less than 98%. This ending of the Contention Free period before its nominal ending might be the result of one of the following: 1. a non-stretched preceding superframe; 2. a low number of ongoing speech conversations; 3. a low number of stations in a talk spurt.

4. Boundary conditions

During the Contention Free period of the Superframe the stations that are part of a voice conversation are polled according to the polling list. We implement the polling list as a simple Round Robin system. If during one CF period the polling list cannot be completed, then in the next CF period the Polling Coordinator will start the polling sequence with the next station on the polling list. If, the polling sequence is not started with the first station on the list and the end of the polling list is reached before the end of the CF period, then the PC will continue the polling sequence with the first station on the polling list. The PC will always end a polling sequence if all the stations on the polling list have been polled during one CF period.

As a result of this setup the minimum Inter Poll Time (T_{IP}) will be approximately equal to the Superframe Time (T_{SF}). If N_C represents the maximum number of voice samples that may be transmitted during one Contention Free Period, then the T_{IP} is as given in equation 1, where N is the actual number of ongoing conversations and N_O is the maximum number of possible conversations in the case of optimal statistical multiplexing.

$$T_{IP} \approx T_{SF} \qquad , \; N \le N_O$$
$$T_{IP} = \frac{N}{N_C \times \left[1 + \frac{T_{SIL}}{T_{TLK}}\right]} \times T_{SF} = \frac{N}{N_O} \times T_{SF}, \quad N > N_O \tag{1}$$

The Superframe time T_{SF} consists of a Contention Free period with length T_{CF} and a Contention period with length T_{CP}. To guarantee that N_C voice samples can be transmitted during each CF Period we need to add the maximum MPDU length to the length of the CF period (see also figure 3). This relationship is shown in equation 2. $T_{maxMPDU}$ and T'_{CF} are further defined in equations 3 - 5.

$$T_{SF} = T_{CF} + T_{CP} = T_{maxMPDU} + T'_{CF} + T_{CP} \tag{2}$$

where $T_{CP} \ge T_{maxMPDU} + DIFS$

$$T_{maxMPDU} = T_{DATA} + SIFS + T_{ACK} \tag{3}$$

$$T_{DATA} = \frac{PA + H + Payload}{R_C} \tag{4}$$

$$T_{tot} = T_{CF-Down} + SIFS + T_{CF-Up} + SIFS$$
$$= 2 \cdot \left[\left(\frac{PA + H + T_{SF} \cdot R_S}{R_C}\right) + SIFS\right] \tag{5}$$

$$T'_{CF} = N_C \times T_{tot} \tag{6}$$

In equation 3, T_{DATA} and T_{ACK} are the times needed to transmit a data frame and its corresponding acknowledgment respectively. In equation 4, PA represents the Preamble length in bits, H the overhead per frame in bits, Payload the Payload size of the data frame in bits, and R_C the transmission speed in bits per second (bps). In equation 5, R_S is the source coding rate or sample rate of the speech encoder in bps.

In equation 5 we have to note that T_{tot} is not under all circumstances the time required for a station to be polled and transmit its frame, but an upper bound of this time. Using this upper bound slightly increases the overhead in the case of an underloaded system, where N is smaller than N_O.

From equations 3- 5 we can deduce a relationship between the maximum possible payload size and N_C with T_{SF} as a parameter. This is shown in figure 5. From these equations we can also find a relationship between the percentage of the available bandwidth used for the transmission of speech and N_C with T_{SF} as a parameter. In equation 7 we define this relationship, this is shown graphically in figure 6.

$$\%B_{Voice} = \frac{T'_{CF}}{T_{SF}} \cdot 100\% \quad (7)$$

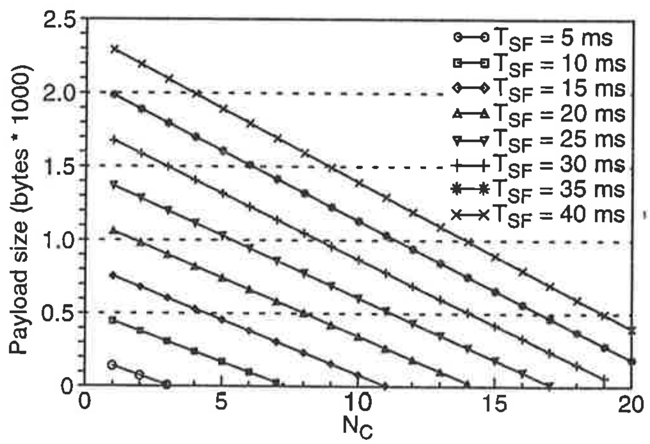

Figure 5 Maximum possible payload size as a function of N_C and T_{SF}

5. Simulation Study

For our simulations we built an extensive model of the 802.11 standard in BONeS® (Block Oriented Network Simulator) Designer™. The model consists of 15 workstations that can transmit data as well as voice. All 15 workstations are involved in the transmission of data, while the number of workstations involved in the transmission of voice is variable.

The simulations use the parameter values in table 1, for the parameters that are yet unspecified by the standard. These values are chosen because they are large enough to be detected in a real system, i.e. greater than the slot-time which is approximately 11 µs, but small enough to have a negligible effect on our simulation results. The parameters that are specified by the 802.11 draft standard are presented in table 2. Furthermore, we introduce several parameters that have the values given in table 3, unless a specific plot shows a different value for a parameter.

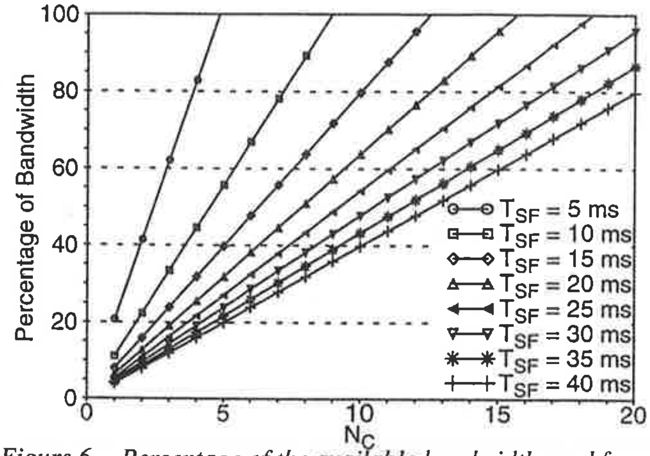

Figure 6 Percentage of the available bandwidth used for speech transmission as a function of N_C and T_{SF}

Description	Value
SIFS	30 µs
DIFS	60 µs
PIFS	90 µs

Table 1 Unspecified values.

Description	Value
ACK	18 bytes
H	32 bytes
PA	24 bytes

Table 2 Specified values.

Description	Value
R_S	8 kbps
T_{SF}	20 ms
N_C	5

Description	Value
Max. Payl.	743 bytes
TTL	25 ms
R_C	1 Mbps

Table 3 Parameter values.

In our simulations we model the transmission channel as an error free channel, and do not take into account the probability of collision. The resulting clipping and throughput statistics are therefore uniquely the result of the protocol and not of lost frames.

6. Results

As a performance measure for the transmission of coded voice we use the percentage of generated bits that is clipped. In our simulations we can have multiple clips per speech burst, thus we use Midspeech Burst Clipping (MSC) [6]. The percentage of clipped bits is calculated over all voice connections. Because R_S is 8 kbps each coded speech bit has a duration of 125 µs and the total buffer length needs to be R_S*TTL.

As a performance measure for the data frames we use the throughput in bits per second (bps). The throughput is the total average user throughput, i.e. of the payload. When $N < N_C$, the system is loaded less than 98% because the Polling Coordinator sends a CF-End frame before T'_{CF} has elapsed. For $N > N_C$ it is possible that the system will be overloaded, because the Polling Coordinator may poll during the complete Contention Free period and thus also during the period that is added to absorb stretching.

From figure 6 we conclude that due to the relatively large overhead each voice conversation requires much more bandwidth than the source coding rate, resulting in a small number of possible voice conversations. Figure 5 demonstrates that the combination of the Superframe structure and the large percentage of the available bandwidth needed for voice transmission results in limitation of the maximum payload size and thus of the maximum throughput.

In figures 7 and 8 we plot our performance measures versus the number of ongoing conversations with N_C as a parameter. Figure 7 shows that due to the low number of voice conversations that can be accommodated, the statistical multiplexing is not as effective. However, if we allow 2 percent clipping we can accommodate approximately $2 \cdot N_C$ voice conversations.

Figures 9 and 10 reveal that the maximum Payload should be as large as possible. A large maximum Payload obviously yields a larger data throughput than a small maximum Payload. However, from figure 9 we conclude that the largest possible maximum payload also yields a minimum percentage of clipped bits. This is due to an increased T_{CF} for a higher maximum payload while the T'_{CF} remains the same. The Polling Coordinator can therefore poll more stations if no full stretching occurs.

From figures 11 and 12 we determine that it is more advantageous to select a larger Superframe length. However, this would also involve a necessary increase of the TTL parameter, otherwise the percentage of clipping will drastically increase (see figure 11 for TTL = T_{SF} = 25 ms). Note that we set the maximum payload to the maximum possible value for different values of T_{SF}.

7. Conclusions

This paper has presented results for the transmission of a combination of time-bounded speech traffic and data traffic over an 802.11 Wireless Local Area Network. The results are generated by performing simulations of a 15 workstation model, which was built using the simulation tool BONeS.

We show that the high overhead for the voice frames has a significant effect on the number of possible voice conversations. Also, the cooperation of the DCF and PCF within the Superframe structure limits the number of possible voice conversations and also limits the maximum payload size that can be transmitted. This limit on its turn limits the maximum achievable throughput. We, therefore, conclude that under the current conditions the Superframe structure operates poorly for both the voice and the data transmission.

Although the number of possible voice conversations is low, the polling scheme is able to deliver the required quality in terms of the percentage of bits clipped under the condition that the length of the polling list can be controlled.

This analysis shows that from both the perspective of voice quality and data throughput the largest maximum payload size should be selected.

Finally, we conclude that a larger Superframe length yields an increased throughput or provides the opportunity to accommodate more voice conversations. However, we have to ensure that the Time To Live is higher than the T_{SF}, otherwise an unacceptable quality results.

8. Acknowledgment

This study has been made possible by the Universiteitsfonds of Delft University of Technology.

9. References

[1] P802.11/D1, Draft Standard IEEE 802.11, "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications," November 1994.

[2] M. A. Visser, "Performance Study of the IEEE 802.11 Wireless LAN Standard," Masters Thesis, Delft University of Technology, June 1995.

[3] Paul T. Brady, "A Statistical Analysis of On-Off Patterns in 16 Conversations," Bell System Technical Journal, January 1968.

[4] Paul T. Brady, "A Model for Generating On-Off Speech patterns in Two-Way Conversation," Bell System Technical Journal, September 1969.

[5] John G. Gruber, "A Comparison of Measured and Calculated Speech Temporal Parameters Relevant to Speech Activity Detection," IEEE Transactions on Communications, Vol. COM-30, No. 4, April 1982.

[6] John G. Gruber and Leo Strawczynski, "Subjective Effects of Variable Delay and Speech Clipping in Dynamically Managed Voice Systems," IEEE

Transactions on Communications, Vol. COM-33, No. 8, August 1985.

[7] Sanjiv Nanda, David J. Goodman and Uzi Timor, "Performance of PRMA: A Packet Voice Protocol for Cellular Systems," IEEE Transactions on Vehicular Technology, Vol. 40, No. 3, August 1991.

[8] Hai Xie and David J. Goodman, "PRMA System Performance With Variable Number of Calls," Technical Report WINLAB-TR-36, June 1992.

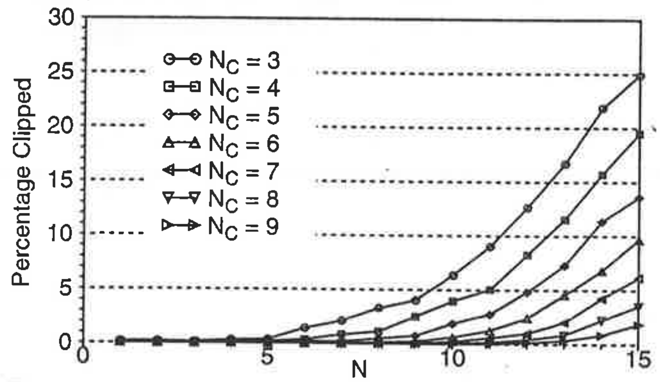

Figure 7 *Percentage of clipped bits versus the number of ongoing conversations (N), with N_C as a parameter.*

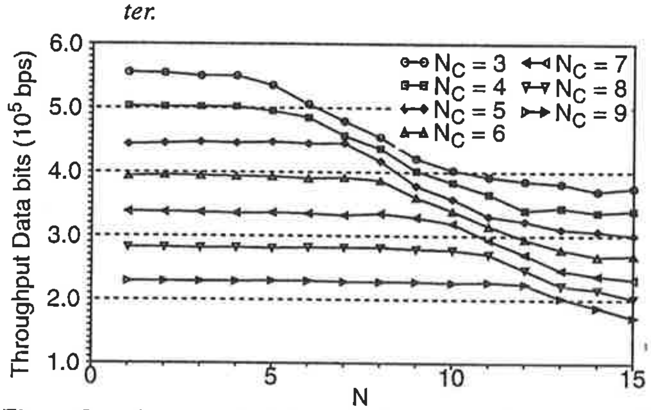

Figure 8 *Throughput of the data bits versus the number of ongoing conversations (N), with N_C as a parameter.*

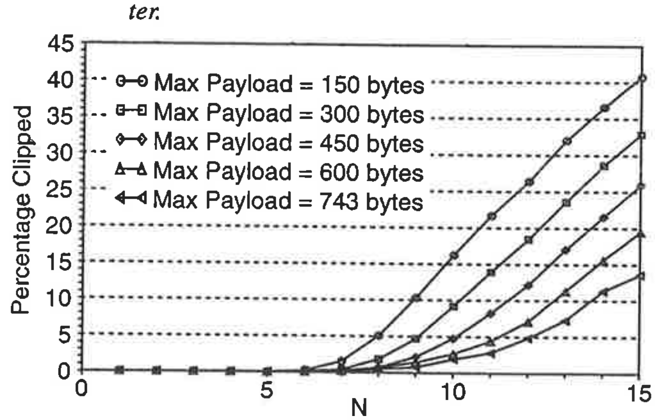

Figure 9 *Percentage of clipped bits versus the number of ongoing conversations (N), with the Maximum Payload size as a parameter.*

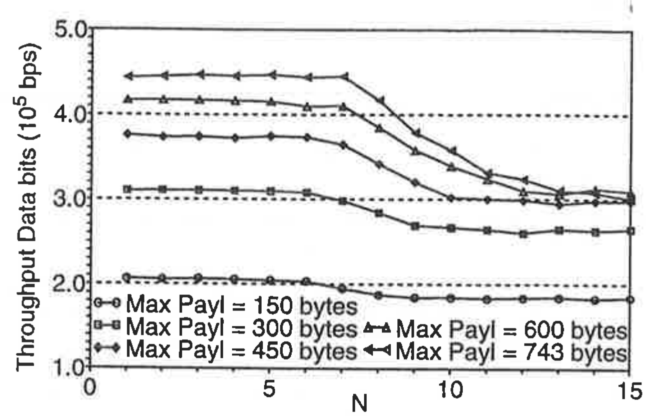

Figure 10 *Throughput vs. the number of ongoing conversations, with the Max. Payload size as a parameter.*

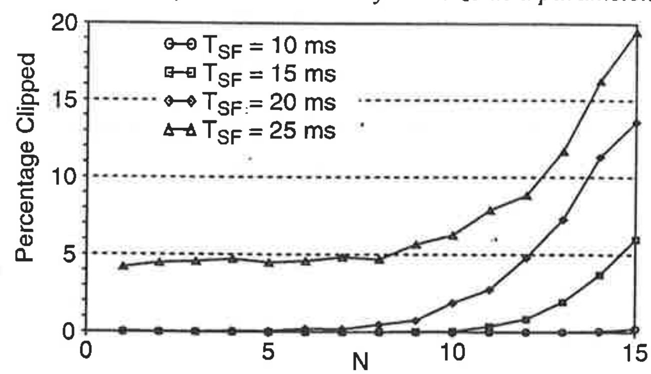

Figure 11 *Percentage of clipped bits versus the number of ongoing conversations (N), with the Superframe length (T_{SF}) as a parameter.*

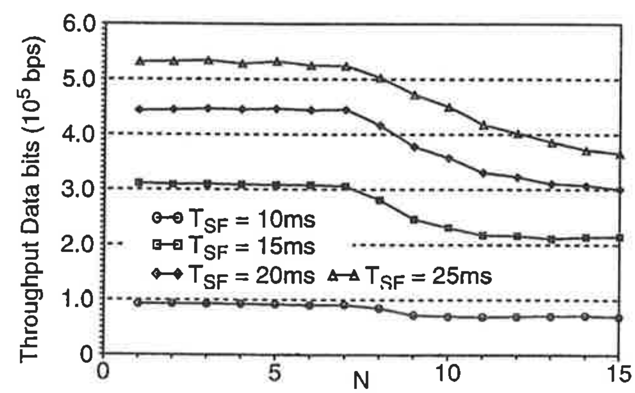

Figure 12 *Throughput versus the number of ongoing conversations (N), with T_{SF} as a parameter.*