CHAPTER 6

Advanced Infrared Collision Avoidance Procedures

This chapter presents an analytical model for the Collision Avoidance (CA) procedures of the Reserved transfer modes of the AIr MAC protocol. It also employs this model to study the effectiveness of the CA procedures on channel utilization assuming an error free channel and no Repetition Rate implementation (RR=1). Chapter 7 employs the model developed in this chapter and derives new analytical models for the AIr retransmission schemes to study utilization performance when transmission errors occur and RR coding is implemented.

The CA procedures are summarized as follows. A station wishing to transmit first contends for medium access. The contention period is slotted and a contending station can only transmit at the beginning of a slot. The slot period is referred to as Collision Avoidance Slot (CAS). To minimize collision probability with other contenting stations, a station first selects a number of CAS to wait and assigns the selected value to the CAS timer. This number is randomly selected from the range (0,*CW*-1), where *CW* is the current Contention Window (CW) size. If during the deferral period a reservation from another station is observed, the station stops the CAS timer and decreases it again when the current transmission is over and the next contention period is started. The station transmits when the CAS timer reaches zero. As the hidden station problem [31][61] is likely to appear in infrared LANs, the AIr protocol defines a long CAS duration. The CAS duration σ is greater than the RTS transmission time plus TAT plus the time required to hear the beginning portion (PA and SYNC fields) of the responding CTS frame ($\sigma > T_{RTS} + \mathrm{TAT} + TT_{PA} + TT_{SYNC}$). Such a long CAS duration ensures that contending stations hidden from the RTS transmitter (and thus not being able to hear the RTS frame) but not from the receiver will receive the beginning of the CTS frame within a single CAS duration. As a result, contending stations with even a single CAS difference at the beginning of the contention period will not collide even if they are hidden from each other.

This approach provides an effective solution to utilization degradation caused by collisions from hidden stations. However the long CAS duration may still degrade utilization if the number of empty and colliding CAS during the contention periods is high. This number depends on the number of the competing stations and on the *CW* values used by these stations. As stations can only adjust their *CW* values after successful reservations and collisions, the implemented *CW* adjustment algorithm becomes of great importance if maximum utilization is to be achieved.

This chapter is outlined as follows. Section 6.1 defines saturation utilization and the parameters used in the analysis. Section 6.2 presents a simple and accurate analytical model to calculate the performance of the CA procedures of the AIr protocol assuming a finite number of stations. The model is based on the CW adjustment procedures proposed in AIr LM specification [42]. The model is validated by comparing analytical with simulation results in section 6.3. Section 6.4 employs the analytical model to evaluate AIr utilization performance for SDATA frame employment under the assumption that no transmission errors occur. In particular, the effect of CW size to utilization is explored and the optimum CW size is derived by differentiating the utilization equation. The effectiveness to utilization of the proposed CW size adjustment scheme and of MAC and PHY layer parameters, such as the Turn Around Time (TAT), minimum and maximum CW size values, is finally explored.

6.1 Saturation utilization and parameter definitions

This work considers channel utilization which is defined in section 2.7 as the time portion during which the infrared medium successfully transfers payload data at the base rate (4 Mbit/s). Transmission of error frames does not contribute to channel utilization. According to the definition, if the transmitted information is repeated RR times, only the first transmission contributes to utilization. Thus, if the transmitter implements RR=2 to cope with transmission errors, utilization is halved.

It is known that several random access schemes show an unstable behavior when the offered load varies. A wide discussion of the instability problem can be found in [10]. The CA scheme of the 802.11 protocol is similar to that of the AIr protocol and exhibits some form of instability according to [21][38]. Simulation results for the 802.11 protocol when the offered load linearly increases with the simulation time are presented in [12]. When the offered load is low compared to channel capacity, 802.11 utilization closely follows the offered load. As load increases, utilization also increases until a maximum value is reached. Further increasing the offered load results in utilization degradation until a steady value is reached [12]. This steady value is referred to as saturation utilization. Saturation utilization is lower than the maximum utilization the LAN can reach, when the offered load reaches a specific value. However, the maximum utilization figure is practically meaningless as the offered load is not a controllable parameter and the network can reach the maximum utilization figure only for short periods of time under varying load conditions. Saturation utilization is defined as the network utilization in steady state conditions. For simplicity, saturation utilization is referred to as 'utilization' in this work.

An infrared LAN consisting of n transmitting stations under saturation conditions is considered. In saturation status, it is assumed that a station always has a window of frames available for transmission. All stations employ the Reserved transfer mode with sequenced data, as described in section 5.3.2.3. When a station successfully captures the channel, it transmits a window of w frames with fixed payload size of l bits at C bit/s. Current analysis also assumes that reservation control frames (RTS, CTS, EOB and EOBC) are always received error free. This is a realistic assumption because, as described in section 5.2, control frames CTS, EOB and EOBC contain only an RH portion which is transmitted using the maximum repetition rate RR=16 to minimize transmission errors. The RTS control frame has also an MBR field consisting of only 48 bits, which is transmitted using variable RR. This MBR length is extremely small for the expected link quality and the assumption that the RTS frames are always received error free also holds true. As the considered indoor links operate at very short distances, current model also assumes that the one way propagation delay is very small and can be safely neglected.

The analytical model also assumes that there are no hidden stations. Thus, all stations will always receive the RTS and the CTS frames of a successful reservation. Therefore, there is no fairness problem as all stations have an equal chance to reserve the infrared medium. Based on AIr CA procedures, the current model considers a CW increase by 4 after a collision and a CW decrease by 4 after a successful reservation [42][96]. The specified CW lower limit of 8 and upper limit of 256 are also implemented. The time required for the transmission of reservation control frames and of frame elements is presented in Table 6.1 [94]. The implemented timers and time delay values are presented in Table 6.2. All suggested values by the standard are applied in the analysis and calculations with the exception of the Wait for CTS (WFCTS) time

frame / frame element		duration	time (μsec)
TT_{PA}	(frame element)	-	64
TT_{SYNC}	(frame element)	-	40
TT_{RH}	(frame element)	-	128
T_{RH}	(frame)	$TT_{PA}+TT_{SYNC}+TT_{RH}$	232
T_{RTS}	(frame)	T_{RH}+48/(C * RR)	244
T_{CTS}	(frame)	T_{RH}	232
T_{EOB}	(frame)	T_{RH}	232
T_{EOBC}	(frame)	T_{RH}	232

Table 6.1 AIr frame and frame element transmission times for C=4Mbit/s and RR=1.

timer/delay	restriction	suggested (μsec)	implemented (μsec)
Turn Around Time (TAT)	-	200	200
CAS duration (σ)	$>T_{RTS}+TAT+TT_{PA}+TT_{SYNC}$	800	800
WFCTS Timer	$>=TAT+T_{RH}$	632	556
EXIT1 Timer	$=TAT+T_{EOBC}+TAT$	632	632
EXIT2 Timer	$=TAT$	200	200

Table 6.2 AIr timers and time delays

out value. A contending station starts the WFCTS timer when it transmits an RTS frame and stops the timer when the corresponding CTS frame is received. If two or more stations are transmitting at the same CAS, the transmitted RTS frames collide, no CTS frames are generated and the WFCTS timer expiration that follows signals the collision to these stations. A WFCTS timer value that obeys the protocol restriction (T_{WFCTS}>=TAT+T_{RH}) and synchronizes the colliding with the remaining stations in contending for medium access at exactly the same time is implemented. The current model also adjusts CAS timers to the beginning of the next CAS when an RTS or a CTS frame is received.

6.2 Analytical model

The key assumption of our analytical model is that an RTS frame transmission collides with the same probability p regardless of the station's current CW value. It is assumed that, in saturation conditions, the collision probability is constant and independent of the experienced collisions and successful reservation attempts. The analytical model first examines the behavior of a single station in order to compute the collision probability p, as well as the probability τ that a station transmits in a randomly selected CAS for a network with n transmitting stations. Subsequently, by considering that only three events can occur in a CAS, a) a successful reservation, b) a collision and c) the CAS is empty, utilization and the average delay for a successful reservation are expressed as a function of τ .

6.2.1 Transmission probability

As stations operate in saturation conditions, a station has immediately a window of frames ready for transmission after it successfully captures the channel and transmits a window of frames. Let *b(t)* be the stochastic process that represents the backoff time counter for a specific station. Process *b(t)* takes only integer values and represents the remaining CAS the station will defer before transmitting. Time scale *t* is also slotted and *t* takes only integer values; *t* represents the beginning of a CAS and *t*+1 represents the beginning of the next CAS. Stations increment *t* at the beginning of a CAS. We will use the term 'slot' to denote an increase in this discrete time scale. If the CAS is empty or contains a collision, an increase in the discrete time scale *t* (slot) corresponds to a CAS duration (σ) in the system time t_s. However, if the CAS contains a successful reservation, a slot corresponds to the total time required for a transmission of *w* frames during the successful reservation (see Fig. 6.1).

The stochastic backoff process *b(t)* decrements by one for every *t* increase. When *b(t)* reaches zero, the station transmits an RTS frame and the new *b(t)* value is randomly selected by a flat distribution in the range (0,*CW*-1). The current *CW* value depends on the number of successful reservations and collisions the station has experienced in the past. As a result, process *b(t)* is non-markovian. We use for convenience the notation W=CW_{min}. We define 'maximum backoff stage' *m* as CW_{max}=W+4*m* and

$$W_i = W + 4i, \quad i \in (0, m) \tag{6.1}$$

where *i* is defined as the 'backoff stage'. According to the definition, W=W_0=CW_{min} and W_m=CW_{max}. According to AIr LM specification [42], CW_{min}=8, CW_{max}=256 and *m*=62. We use *s*(*t*) to denote the stochastic process representing the backoff stage $i \in (0, m)$ of the station at time *t*,

$$s(t) = i, \quad i \in (0, m) \tag{6.2}$$

Considering that W_i is the current *CW* value of a station and that process *b(t)* represents the remaining CAS a station will defer transmission, Figure 6.1 shows *s(t)* and *b(t)* values for the contention process presented in Figure 5.5. Figure also shows how discrete time scale *t* relates to system time t_S and presents the slot duration. Discrete time scale *t* increases only at the beginning of a CAS and takes only integer

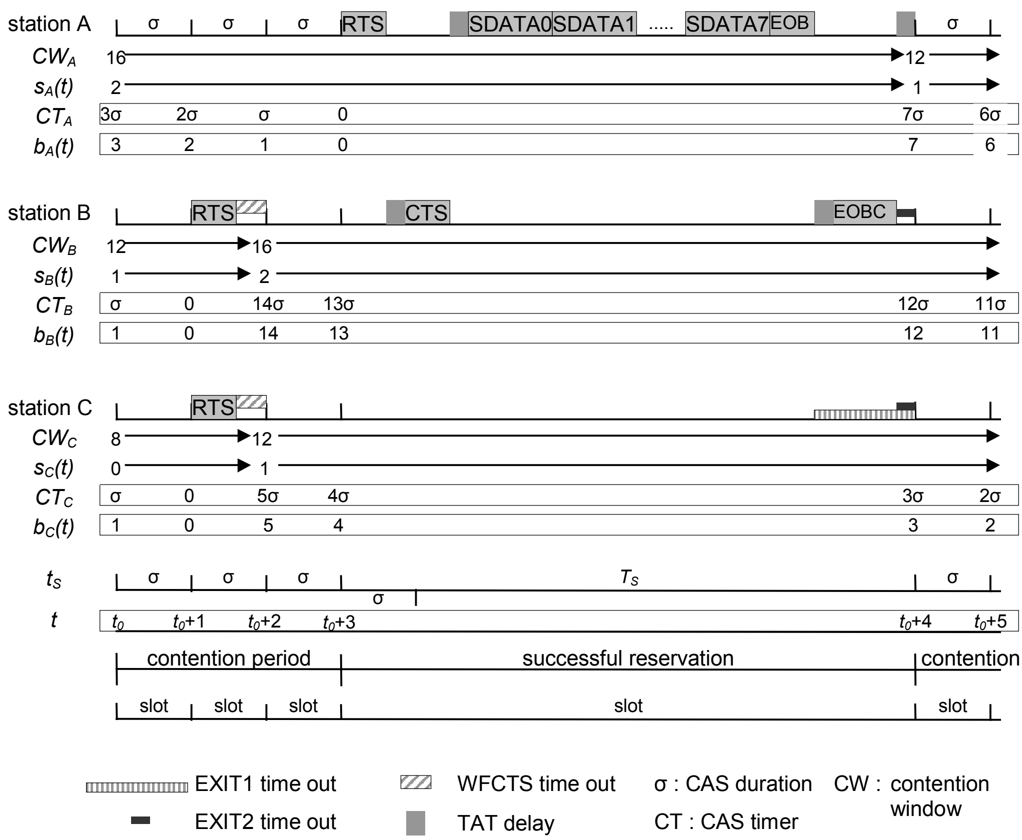


Figure 6.1. Stochastic processes s(t) and b(t)

values. At the beginning of the contention (t=t_0), station A is at backoff stage 2 ($s_A(t_0)$=2) because its current *CW* value is 16 (CW_A=16=8+4x2). Station A also has $b_A(t_0)$=3 because it will defer for three CAS before transmitting an RTS. Similarly, station B has $s_B(t_0)$=1 and $b_B(t_0)$=1 and station C has $s_C(t_0)$=0 and $b_C(t_0)$=1. As the first CAS is empty, stations do not change backoff stage ($s_X(t_0+1)$=$s_X(t_0)$, X={A,B,C}), decrement their $b(t)$ value ($b_X(t_0+1)$=$b_X(t_0)$-1, X={A,B,C}) and the slot duration is σ. As the second CAS contains a collision of stations B and C, these stations increase their backoff stages ($s_B(t_0+2)$=$s_B(t_0+1)$+1=2 and $s_C(t_0+2)$=$s_C(t_0+1)$+1=1) and randomly select new $b(t)$ values ($b_B(t_0+2)$=14 and $b_C(t_0+2)$=5). Station A decreases its $b(t)$ value because it does not receive a valid frame and the slot duration is σ. The third CAS is empty and at (t_0+3) stations have $s_A(t_0+3)$=2, $b_A(t_0+3)$=0, $s_B(t_0+3)$=2, $b_B(t_0+3)$=13, $s_C(t_0+3)$=1 and $b_C(t_0+3)$=4. The fourth CAS contains a successful RTS transmission and, as discrete time scale t increases only at the beginning of a CAS, the slot duration is T_S and contains the successful RTS/CTS exchange, the transmission time of the SDATA

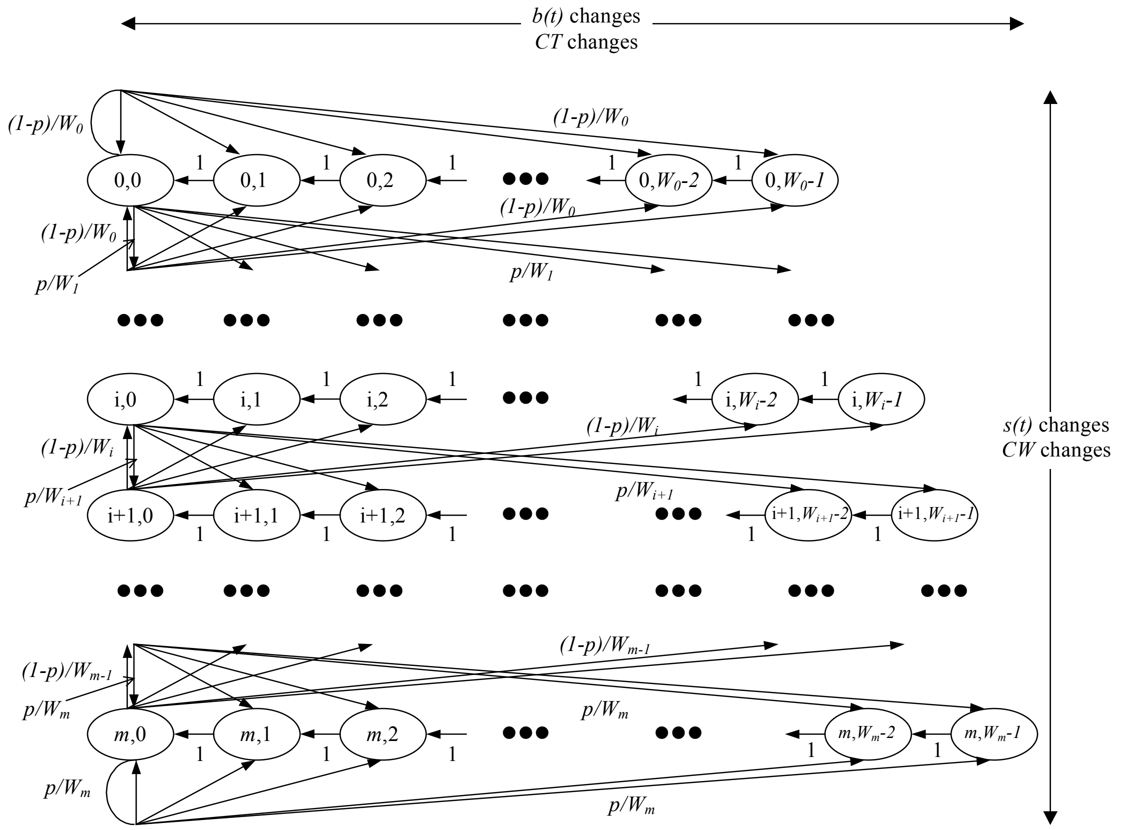


Figure 6.2 Markov Chain model for back off CW

frames and the EOB/EOBC frame exchange. Stations B and C decrease their $b(t)$ values ($b_B(t_0+4)=12$ and $b_C(t_0+4)=3$) and station A decreases its back off stage ($s_A(t_0+4)=1$) and randomly selects a new $b(t)$ value ($b_A(t_0+4)=7$) because it accomplished a successful reservation.

Assuming that the collision probability p is constant and independent of the *CW* used to select the current deferral period, the bidimentional process $\{s(t),b(t)\}$ can be represented by the discrete-time Markov chain drawn in Fig. 6.2. Lets assume that a station's bidimentional process $\{s(t),b(t)\}$ is currently at state (i,k), $i=0,1,...,m$, $k=0,1,...,W_i-1$. As $s(t)=i$, the station's current *CW* value is W_i; as $b(t)=k$, the station will defer k slots before transmitting an RTS. As $b(t)$ decrements at the discrete time scale, a decrement may correspond to an empty CAS or to an RTS collision by other stations or to a successful reservation of another station. After k steps, the station reaches state $(i,0)$ and transmits an RTS. As this transmission collides with probability p, the station will transition to state $(i+1,k)$ with probability p/W_{i+1}, where k is randomly selected in the range $(0,W_{i+1}-1)$; as this RTS transmission is successful with probability $(1-p)$, the station will transition (at the end of the reservation, see Fig. 6.1) to state (*i-1*,*k*) with probability (*1-p*)/W_{i-1}, where *k* is randomly selected in the range (0,W_{i-1}-1) in this case.

The only non-null one-step transition probabilities of the chain are (using the short notation $P\{i_1,k_1 \mid i_0,k_0\} = P\{s(t+1)=i_1, b(t+1)=k_1 \mid s(t)=i_0, b(t)=k_0\}$):

$$\begin{cases} P\{i,k \mid i,k+1\} = 1 & k\in(0,W_i-2) \quad i\in(0,m) \\ P\{i,k \mid i+1,0\} = (1-p)/W_i & k\in(0,W_i-1) \quad i\in(0,m-1) \\ P\{i+1,k \mid i,0\} = p/W_{i+1} & k\in(0,W_{i+1}-1) \quad i\in(0,m-1) \\ P\{0,k \mid 0,0\} = (1-p)/W_0 & k\in(0,W_0-1) \\ P\{m,k \mid m,0\} = p/W_m & k\in(0,W_m-1) \end{cases} \tag{6.3}$$

The first equation in (6.3) accounts for the fact that the backoff time counter is decreased until it reaches zero. The second and third equations represent the *CW* decrease and the *CW* increase after a successful reservation and a collision respectively. The fourth equation considers that if the current backoff stage is 0, the *CW* value is not further decreased even after a successful reservation. Finally, the fifth equation considers that the *CW* is not increased after a collision if the maximum backoff stage *m* is reached.

We define that $b_{i,k} = \lim_{t\to\infty} P\{s(t)=i, b(t)=k\},\ i\in(0,m),\ k\in(0,W_i-1)$ represents the stationary distribution of the chain. Considering that $b_{00} = (1-p)b_{00} + (1-p)b_{10}$, b_{10} is given by:

$$b_{10} = \frac{p}{1-p} b_{00} \tag{6.4}$$

Since $b_{10} = pb_{00} + (1-p)b_{20}$, b_{20} is given by

$$b_{20} = \left(\frac{p}{1-p}\right)^2 b_{00} \tag{6.5}$$

Consequently, it can be easily shown that

$$b_{i0} = \left(\frac{p}{1-p}\right)^i b_{00},\ i\in(0,m) \tag{6.6}$$

Owing to chain regularities, for each $k\in(0,W_i-1)$,

$$b_{0k} = \frac{W_0-k}{W_0}\left(b_{10}(1-p) + b_{00}(1-p)\right)$$

$$b_{ik} = \frac{W_i - k}{W_i}\left(b_{i+1,0}(1-p) + b_{i-1,0}p\right),\ i \in (1, m-1) \tag{6.7}$$

$$b_{mk} = \frac{W_m - k}{W_m}\left(b_{m-1,0}p + b_{m0}p\right)$$

After some algebra

$$b_{0k} = \frac{W_0 - k}{W_0} b_{00}$$

$$b_{ik} = \frac{W_i - k}{W_i} b_{i0},\ i \in (1, m-1) \tag{6.8}$$

$$b_{mk} = \frac{W_m - k}{W_m} b_{m0}$$

Equations (6.8) can be rewritten as

$$b_{ik} = \frac{W_i - k}{W_i} b_{i0},\ i \in (0, m), k \in (0, W_i - 1) \tag{6.9}$$

As a result, equations (6.6) and (6.9) express all $b_{i,k}$ values as a function of b_{00} and probability p. By applying the normalization factor

$$1 = \sum_{i=0}^{m}\sum_{k=0}^{W_i-1} b_{ik} = \sum_{i=0}^{m} b_{i0} \sum_{k=0}^{W_i-1}\frac{W_i - k}{W_i} = \sum_{i=0}^{m} b_{i0}\frac{W_i + 1}{2} = \frac{b_{00}}{2}\sum_{i=0}^{m}\left(\frac{p}{1-p}\right)^i (W_i + 1) \tag{6.10}$$

and by substituting W_i form equation (6.1),

$$1 = \frac{b_{00}}{2}\left[(W+1)\sum_{i=0}^{m}\left(\frac{p}{1-p}\right)^i + 4\sum_{i=0}^{m} i\left(\frac{p}{1-p}\right)^i\right] \tag{6.11}$$

After some algebra we derive

$$1 = \frac{b_{00}}{2}\left[(W+1)\frac{(1-p)^{m+1} - p^{m+1}}{(1-2p)(1-p)^m} + \frac{4p(1-p)}{(1-2p)^2}\left(1 - \frac{p^m(1+m) - p^{m+1}(2m+1)}{(1-p)^{m+1}}\right)\right] \tag{6.12}$$

Equation (6.12) expresses b_{00} as a function of the collision probability p, the smallest contention window size W and the number of backoff stages m. This analysis allows us to evaluate the station's transmission probability τ. Considering that a station transmits when the backoff timer reaches zero,

$$\tau = \sum_{i=0}^{m} b_{i0} = b_{00}\sum_{i=0}^{m}\frac{p^i}{(1-p)^i} = b_{00}\frac{(1-p)^{m+1} - p^{m+1}}{(1-2p)(1-p)^m} \tag{6.13}$$

Substituting the value of b_{00} from eq. (6.12) to eq. (6.13), τ is given by

$$\tau(p):\tau=\frac{2}{(W+1)+\frac{4p\left((1-p)^{m+1}+(2m+1)p^{m+1}-(m+1)p^{m}\right)}{\left((1-p)^{m+1}-p^{m+1}\right)(1-2p)}} \tag{6.14}$$

Equation (6.14) expresses probability τ as a function of the unknown probability *p*. Assuming that the number of transmitting stations *n* is constant and that all contending stations 'see' the discrete-time Markov chain drawn in Fig. 6.2 at the steady state and transmit with probability τ, the transmission collision probability *p* can be expressed as the probability that at least an extra one of the remaining (*n*-1) stations transmit at the chosen CAS

$$p=1-(1-\tau)^{n-1} \tag{6.15}$$

Equations (6.14) and (6.15) form a non-linear system with the unknowns τ and *p* which can be solved by employing numerical methods. Finally, τ and *p* are evaluated for a certain *W*, *m* and *n* combination. It is easy to demonstrate that there is only one solution to the non-linear system. Equation (6.15) can be rewritten as:

$$\tau^{*}(p):\tau=1-(1-p)^{1/(n-1)} \tag{6.16}$$

The function $\tau^{*}(p)$ is a continuous and monotonically increasing function in the range $p\in(0,1)$ and increases from $\tau^{*}(0)=0$ to $\tau^{*}(1)=1$. Function $\tau(p)$ is a continuous and monotonically decreasing function is the same range. It decreases from $\tau(0)=\frac{2}{W+1}$ to $\tau(1)=\frac{2}{W+1+4m}$. There is only one solution to the non-linear system because $\tau(0)>\tau^{*}(0)$ and $\tau(1)<\tau^{*}(1)$. Continuity of function $\tau(p)$ with respect to critical value *p*=1/2 is easily shown by taking into account that in this case equation (6.6) reduces to $b_{i0}=b_{00}$, $i\in(0,m)$ and equation (6.10) becomes

$$1=\frac{b_{00}}{2}\sum_{i=0}^{m}(W_i+1)=\frac{b_{00}}{2}\sum_{i=0}^{m}(W+4i+1)=\frac{b_{00}}{2}\left((m+1)(W+1)+4\sum_{i=0}^{m}i\right) \tag{6.17}$$

and $b_{00}=\frac{2}{(m+1)(W+2m+1)}$. In this case $\tau(1/2)=(m+1)b_{00}=\frac{2}{W+2m+1}$.

6.2.2 Utilization analysis

Based on the calculated collision probability *p* and transmission probability τ, we can now analyse all possible events that can occur in a randomly chosen slot. Let P_{tr} be the probability that at least one station transmits an RTS frame in the considered slot. For a LAN of n contending stations and a station transmission probability τ,

$$P_{tr} = 1-(1-\tau)^n \tag{6.18}$$

Let P_s be the conditional probability that an occurring RTS transmission is successful (P(success/trans)). P_s can be evaluated as the probability that, in the considered slot, one station transmits an RTS and the remaining (n-1) stations defer transmission provided that at least one station transmits an RTS,

$$P_s = \frac{P_{success}}{P_{tr}} = \frac{n\tau(1-\tau)^{n-1}}{1-(1-\tau)^n} \tag{6.19}$$

A successful reservation occurs with probability P_sP_{tr} in a randomly selected slot and the time utilized for transmitting payload information is wt, where w is the window size and t is defined as the time required for transmitting payload information data in an SDATA frame. The value of t is given by

$$t = \frac{RR\ l}{C} \tag{6.20}$$

where RR is the Repetition Rate, l is the payload data length and C is the base data rate. Considering that a randomly selected slot is empty with probability 1-P_{tr} and contains a collision with probability P_{tr}(1-P_s), utilization can be evaluated by dividing the time utilized for transmitting payload data by the average slot duration

$$U = \frac{1}{RR} \cdot \frac{P_{tr}P_s wt}{(1-P_{tr})\sigma + P_{tr}P_sT_s + P_{tr}(1-P_s)\sigma} \tag{6.21}$$

where T_s is the slot duration when a successful reservation is accomplished and σ is the CAS duration. Equation (6.21) considers that a collision lasts exactly one CAS duration. Utilization equation (6.21) can be easily reduced to

$$U = \frac{1}{RR} \cdot \frac{P_{tr}P_s wt}{P_{tr}P_sT_s + \sigma - P_{tr}P_s\sigma} \tag{6.22}$$

When the Reserved transfer mode with sequenced data is utilized

$$T_S = D + w(t + F_S + p_1) \tag{6.23}$$

where D is the reservation overhead that includes the transmission time of the RTS, CTS, EOB and EOBC frames and the TAT delays that follow these frames, F_S is the transmission time of the SDATA frame overhead (preamble, robust header, CRC etc.) and p_1 is the preparation time of an SDATA frame. Assuming that the RTS MBR field is always transmitted using RR=1, D=1.74 msec. The value of F_S is given by

$$F_S = T_{RH} + \frac{RR\, l_S^{'}}{C} \tag{6.24}$$

where $l_S^{'}$ is the length of the MBR overhead of an SDATA frame and T_{RH} is as defined in section 5.2. T_{RH} is the transmission time of a frame with no MBR field. According to Fig. 5.3, $l_S^{'}$ =80 bits and section 5.2 evaluates that T_{RH}=232 μsec.

Our analytical model allows measurement of the time portion utilized on all tasks affecting AIr utilization. Such an evaluation reveals the impact of physical and link layer parameters to utilization. It is valuable for link designers in achieving high utilization at a reasonable cost and for link implementers in selecting suitable parameter values in order to maximize utilization. Considering that a randomly selected slot is empty with probability 1-P_{tr} and that the CAS duration is σ, the time portion utilized in empty CAS is given by

$$U_{empty} = \frac{(1-P_{tr})\sigma}{P_{tr}P_sT_s + \sigma - P_{tr}P_s\sigma} \tag{6.25}$$

A randomly selected slot contains a collision with probability $P_{tr}(1-P_s)$ and the time portion utilized on collisions when two or more stations are simultaneously trying to reserve the infrared channel is

$$U_{coll} = \frac{P_{tr}(1-P_s)\sigma}{P_{tr}P_sT_s + \sigma - P_{tr}P_s\sigma} \tag{6.26}$$

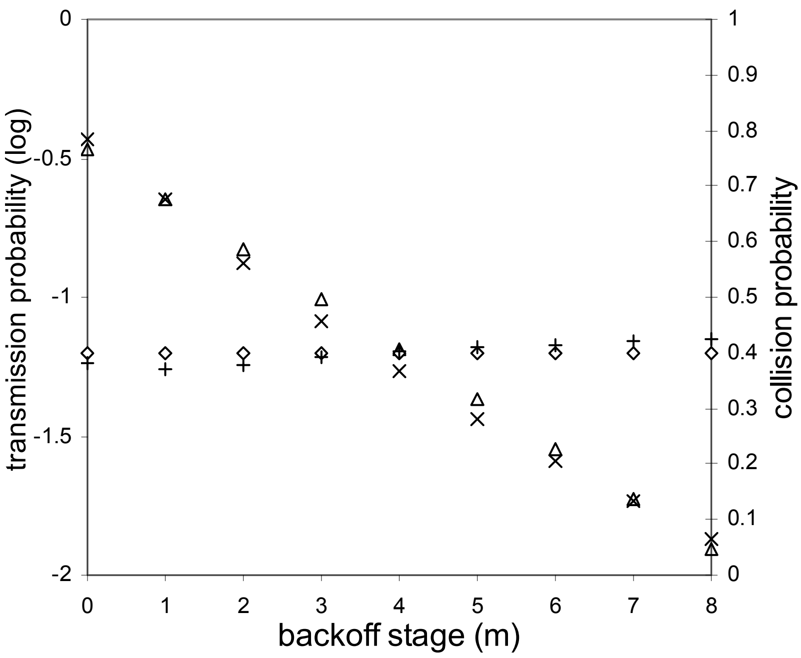
The time portion utilized on transmitting data frame overheads, reservation control frames (RTS, CTS, EOB and EOBC) and the associated TAT delays is

$$U_{over} = \frac{P_{tr}P_s(T_s - wt)}{P_{tr}P_sT_s + \sigma - P_{tr}P_s\sigma} \tag{6.27}$$

The average contention period for a successful reservation, C_p, is given by

$$C_p = \frac{1-P_{tr}P_s}{P_{tr}P_s}\sigma \tag{6.28}$$

The calculated C_p value represents the overhead introduced by the CA procedures for a successful reservation and is particularly useful if the LC layer implements error control. In this case, a MAC successful reservation may contain only an LC acknowledgment frame that informs the transmitter of the correctly received frames in the previous window transmission. It will be used in chapter 7 which considers channel



+ *collision probability p_i, simulation*
◇ *collision probability p, analysis*
x *transmission probability q_i, simulation*
△ *transmission probability q_i, analysis*

Figure 6.3 Transmission (q_i) and collision (p_i) probabilities versus back off stage, W=8, m=62, n=5

errors, LC acknowledgement overhead and retransmission delays.

6.3 Model validation

The model presented in the previous section is validated by comparing analysis with simulation results obtained from the OPNET simulator presented in section 5.6. We first focus on the behavior of a single station. The probability that a station's transmission is at stage i is denoted by q_i. The analytical model calculates probabilities q_i by

$$q_i = \frac{b_{i0}}{\tau}, \ i=0,1, \ldots ,m \tag{6.29}$$

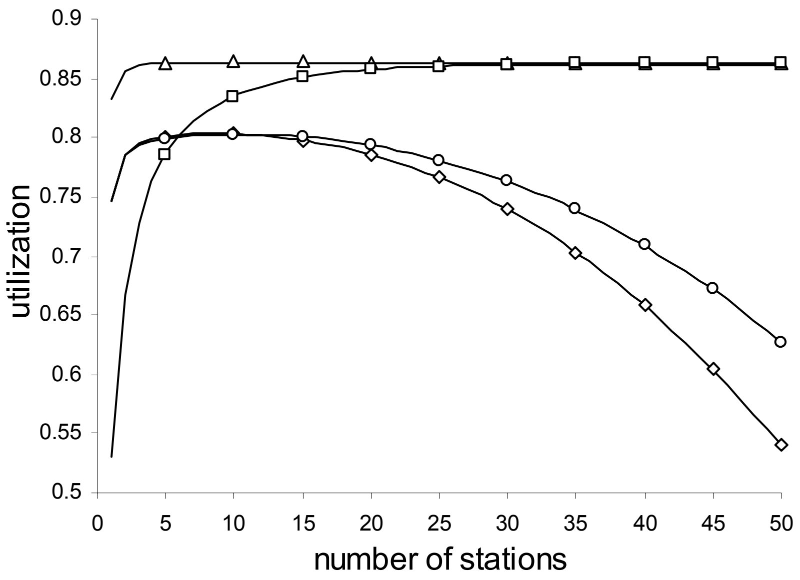
Probability p_i is defined as the collision probability of a station's transmission at stage i. The average collision probability, p_{ave}, of a station's transmission is given by

$$p_{ave} = \sum_{i=0}^{m} q_i p_i \tag{6.30}$$

The key assumption of the analytical model can now be expressed as

$$p = p_{ave} = p_i, \ i=0,1, \ldots ,m \tag{6.31}$$

Fig. 6.3 compares simulation and analysis results for the q_i and p_i probabilities for a

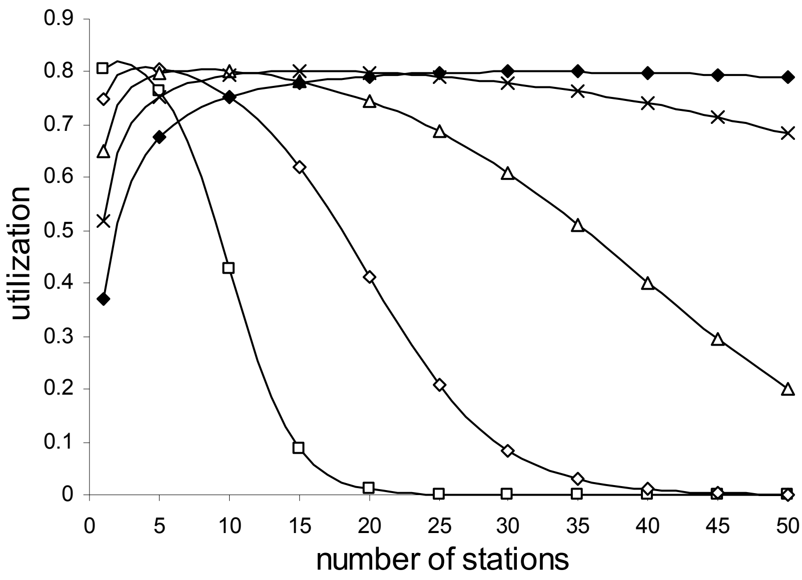


△ *W=8, m=62, l=16Kbits, w=8, C=4Mbit/s, RR=1*
□ *W=64, m=62, l=16Kbits,w=8, C=4Mbit/s, RR=1*
◇ *W=8, m=4, l=16Kbits, w=4, C=4Mbit/s, RR=1*
○ *W=8, m=5, l=16Kbits, w=4, C=4Mbit/s, RR=1*

Figure 6.4 Utilization: analysis versus simulation

network with five transmitting stations using the proposed values for W and m parameters. The figure includes only stages with transmission probability higher than 1% and shows that high back off stages are rarely used for n=5. Simulation results are acquired with a 95% confidence interval lower than 0.002. Slight differences are noticed between the p_i values and the p value that approximates them. The simulation calculated that the average collision probability p_{ave} is 0.385 and it is effectively approximated by the analytical model's constant collision probability p of 0.397. In addition, simulation results indicate that the collision probability p_i is not significantly different for higher stages, as it might have been expected. This result validates the key assumption of the proposed analytical model.

Considering the behavior of a network with n stations, analysis and simulation utilization versus network size is plotted in Fig. 6.4 for different W and m values. In all cases, the analytical model is proven extremely accurate as analytical results (lines) coincide with simulation results (symbols). Figure shows that the slight differences between p_i and p probabilities observed in Fig. 6.3 have negligible impact on utilization. Simulation results are again calculated with a 95% confidence interval lower than 0.002.



□ *CW (slots)=4* ◇ *CW (slots)=8*
△ *CW (slots)=16* x *CW (slots)=32*
◆ *CW (slots)=64*

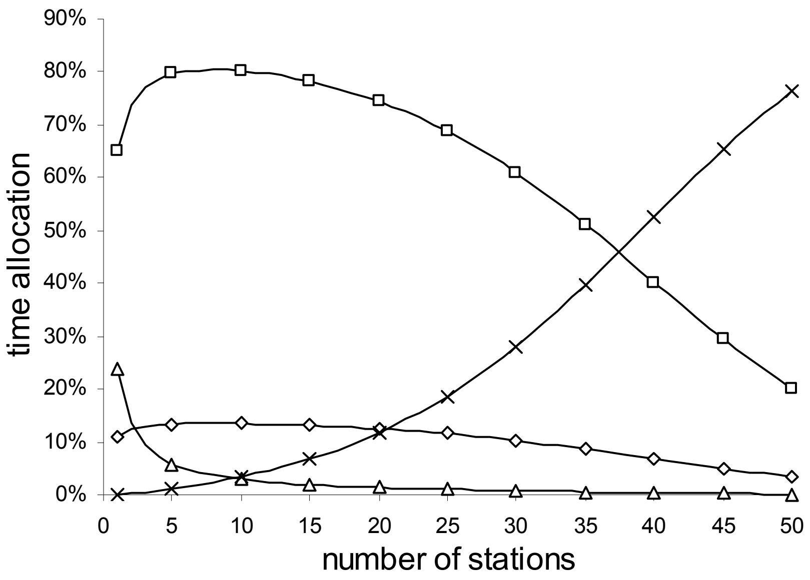
Figure 6.5 Utilization versus n for fixed CW size, l=16Kbits, w=4, C=4Mbit/s, RR=1

6.4 Performance evaluation

This section presents a performance evaluation of the AIr protocol by employing the analytical model developed in section 6.2. The considered SDATA frame employment is referred to as Sequential No Frame Level Acknowledgment (SEQ-NoFLACK) protocol in the next chapter. This evaluation assumes that p_1 is 40 μsec, the infrared channel is error free and that no Repetition Rate coding (RR=1) is implemented.

6.4.1 The effect of Contention Window (CW) size

The impact of the CW size to utilization is considered first. Fig. 6.5 plots utilization versus number of stations for fixed *CW* values, i.e. no *CW* adjustment algorithm is enforced. Utilization results are calculated by employing the analytical model proposed in section 6.2 for no backoff stages (*m*=0). Results for *n*=1 are also included because *n* represents the number of the transmitting stations and not the number of the participating stations in the LAN. Fig. 6.5 shows that utilization is low for networks with a small number of transmitting stations when a large *CW* value is used. When the number of stations increases, maximum utilization performance is achieved. Further



□ *payload data transmission (utilization)*
◇ *overheads*
△ *empty CAS*
x *collisions*

Figure 6.6 Time allocation of various AIr tasks versus n for fixed CW size=16 slots, l=16Kbits, w=4, C=4Mbit/s, RR=1

network size increase results in utilization degradation, especially for small *CW* values. Fig. 6.5 reveals that the *CW* value that results in maximum utilization increases if the number of transmitting stations increases. Utilization performance is significantly decreased if unsuitable *CW* values are implemented. Fig. 6.5 is identical to Fig. 5.11, which plots simulation results for the same network scenarios. This is an additional validation of the analytical model and of the results it produces. Fig. 6.6 plots all factors affecting utilization versus n for a fixed *CW* value of 16. Time portion utilized on empty CAS, collisions and overheads are plotted using eq. (6.25), (6.26) and (6.27) respectively. For small networks, utilization degrades mainly due to the increased number of empty CAS. For large networks, the increased number of collisions results in significant utilization degradation. Maximum utilization is observed when a suitable *CW* for the network size is used that results in a very small number of collisions and empty CAS simultaneously.

6.4.2 Optimum CW size

Optimum *CW* (W_{opt}) value for a LAN with n transmitting stations can be evaluated by employing the proposed analytical model for m=0 and by setting the first derivative of the utilization equation versus τ equal to zero. When m=0, the model's key assumption that a transmission collides with the same probability regardless of the *CW* value used to select the deferral period is always true because *CW* is fixed in this case. By rearranging eq. (6.22)

$$U = \frac{wt/RR}{T_s - \sigma + \dfrac{\sigma}{P_{tr} P_s}} \tag{6.32}$$

Considering that w, t, T_s, RR and σ are constants, utilization is maximized when the expression

$$u = P_{tr} P_s \tag{6.33}$$

is maximized. By substituting P_{tr} and P_s values from equations (6.18) and (6.19), equation (6.33) becomes

$$u = n\tau(1-\tau)^{n-1} \tag{6.34}$$

By taking the first derivative of equation (6.34) versus τ and setting it equal to zero, after some algebra we obtain

$$\tau_{opt} = \frac{1}{n} \tag{6.35}$$

To reach optimum transmission probability τ_{opt} stations should employ the optimum *CW* size W_{opt}. When m=0, equation (6.14) reduces to

$$\tau = \frac{2}{W+1} \tag{6.36}$$

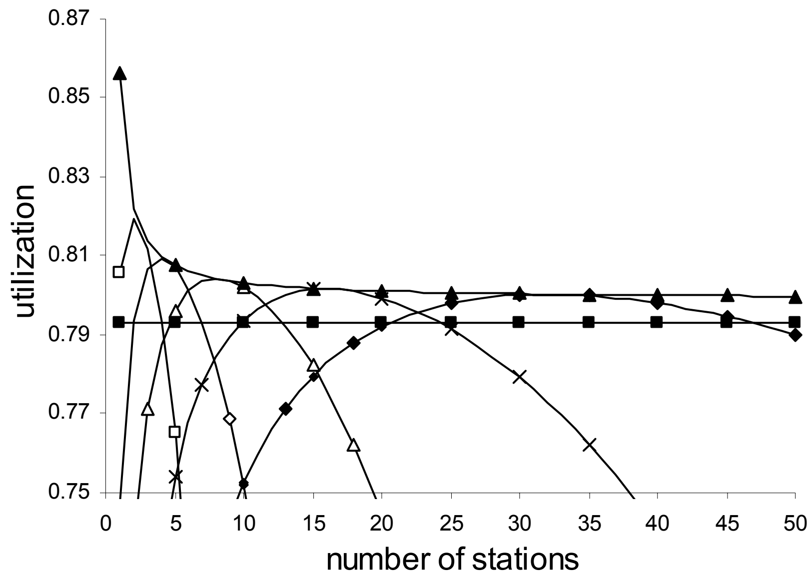
Combining equations (6.35) and (6.36)

$$W_{opt} = 2n - 1 \tag{6.37}$$

Maximum utilization can be evaluated from equation (6.32) if we substitute P_{tr} and P_s from equations (6.18) and (6.19) for τ_{opt} given from equation (6.35)

$$U_{max} = \frac{wt/RR}{T_s - \sigma + \sigma\left(\dfrac{n}{n-1}\right)^{n-1}} \tag{6.38}$$

As a result, maximum utilization depends on the number of transmitting stations in the LAN. However, for large n maximum utilization reaches a steady value. This



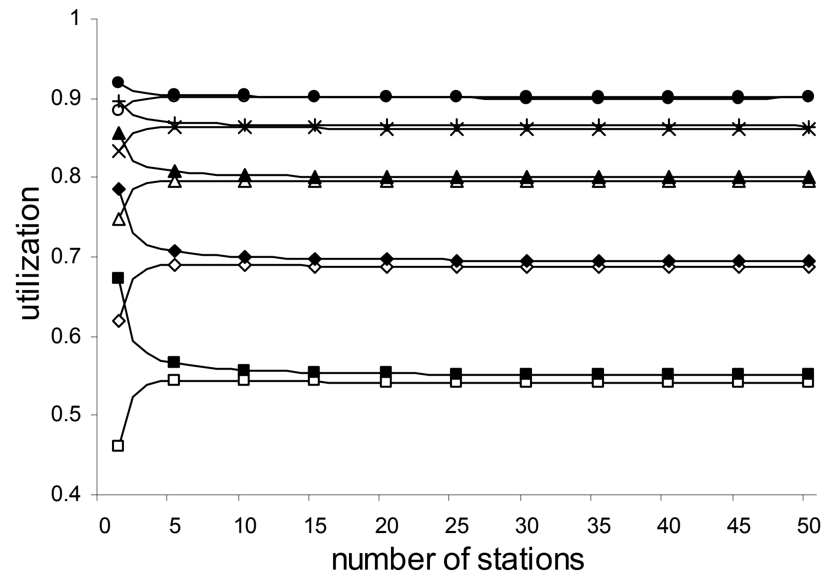
□ *CW (slots)=4*
◇ *CW (slots)=8*
△ *CW (slots)=16*
x *CW (slots)=32*
◆ *CW (slots)=64*
▲ *maximum* U_{max}
■ *maximum* U_{appr}

Figure 6.7 Utilization versus n for fixed CW size, l=16Kbits, w=4, C=4Mbit/s, RR=1

conclusion is slightly different with the expressed conclusion in [12][13] that maximum utilization is independent of *n* for the exponential backoff adjustment scheme of the IEEE 802.11 protocol although linear and exponential backoff schemes coincide when no adjustment (m=0) is allowed. Different conclusions arise from the approximations necessary for calculating maximum utilization in [12][13] because a collision lasts several CAS time periods in the 802.11 protocol. Fig. 6.7 plots utilization versus number of stations for fixed *CW* values and focuses on the maximum achievable utilization (note that the y-axis ranges from 0.75 to 0.87). It also plots U_{max} given from eq. (6.38) and the approximated maximum utilization U_{appr} calculated by performing the approximations presented in [12][13] for AIr's physical and link layer parameter values. The figure shows that when collisions last exactly one CAS duration, as in the AIr protocol, the approximations result in a lower calculated utilization, especially for very small LANs.

6.4.3 CW adjustment algorithm

The effectiveness of the proposed CW adjustment algorithm in estimating suitable

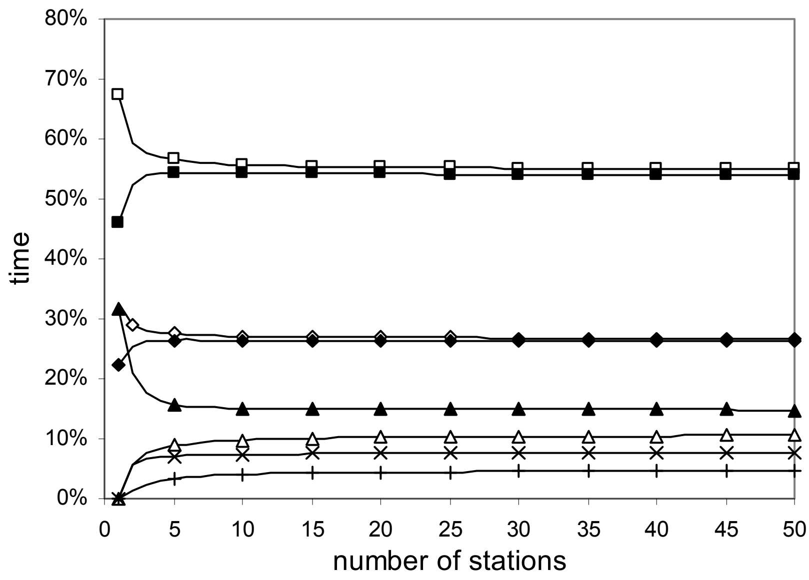


□ *w=1* ■ *w=1*, U_{max}
◇ *w=2* ◆ *w=2*, U_{max}
△ *w=4* ▲ *w=4*, U_{max}
x *w=8* + *w=8*, U_{max}
○ *w=16* ● *w=16*, U_{max}

Figure 6.8 Utilization versus n, l=16Kbits, W=8, m=62, C=4Mbit/s, RR=1

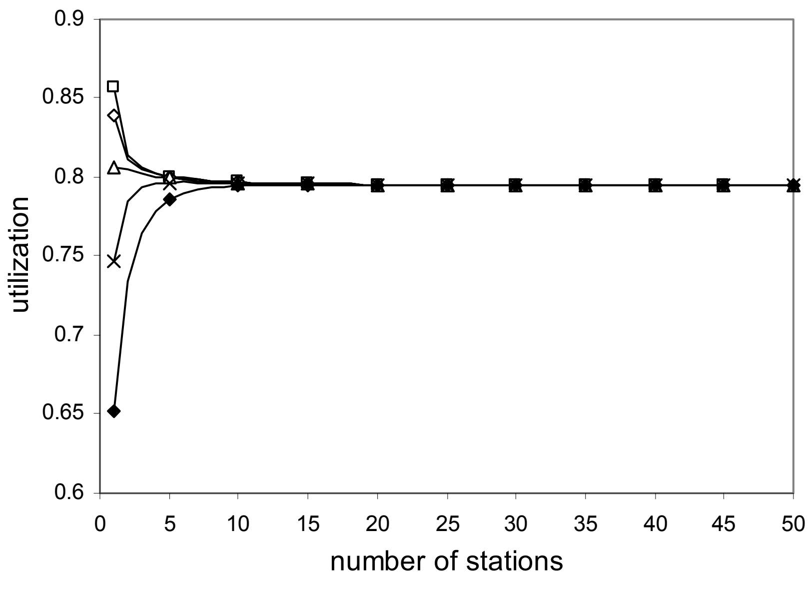
CW values is explored in Fig. 6.8, which compares maximum with achieved utilization for the protocol proposed CW_{min}=8 and *m*=62 values. The proposed algorithm performs effectively but for small networks the achieved utilization is significantly lower than the maximum possible. Fig. 6.8 also shows that utilization is very low when *w* is small and that applications should employ large window sizes in order to achieve high utilization. AIr is an inefficient choice for applications wishing to transmit small amounts of information data in every successful reservation.

Fig. 6.9 plots the time portion utilized by all time consuming tasks for the proposed parameter values. It also plots for comparison the same results when the optimum *CW* value (W_{opt}) is implemented. W_{opt} is calculated by eq. (6.37). For large networks, the proposed adjustment algorithm and parameter values result in more empty CAS and less colliding CAS than optimal. As utilization is slightly lower than optimal for large networks, the proposed adjustment algorithm is considered to perform satisfactorily. However, for small networks, the increased number of empty CAS results in a significantly lower utilization than optimal. Low AIr utilization for small networks was first noticed by evaluating simulation results presented in Fig. 5.12.



□ *payload data transmission (utilization) (optimal)* ■ *payload data transmission (utilization)*
◇ *overheads (optimal)* ◆ *overheads*
△ *empty CAS (optimal)* ▲ *empty CAS*
x *collisions (optimal)* + *collisions*

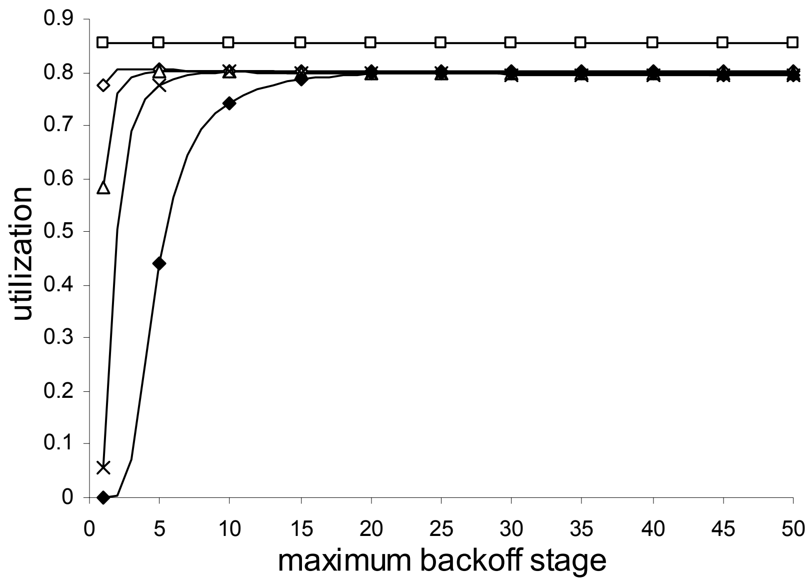
Figure 6.9 Time allocation for various AIr tasks versus n, l=16Kbits, w=1, W=8, m=62, C=4Mbit/s, RR=1



□ *W=1* ◇ *W=2*
△ *W=4* x *W=8*
◆ *W=16*

Figure 6.10 Utilization versus n, m=62, l=16Kbits, w=4, C=4Mbit/s, RR=1

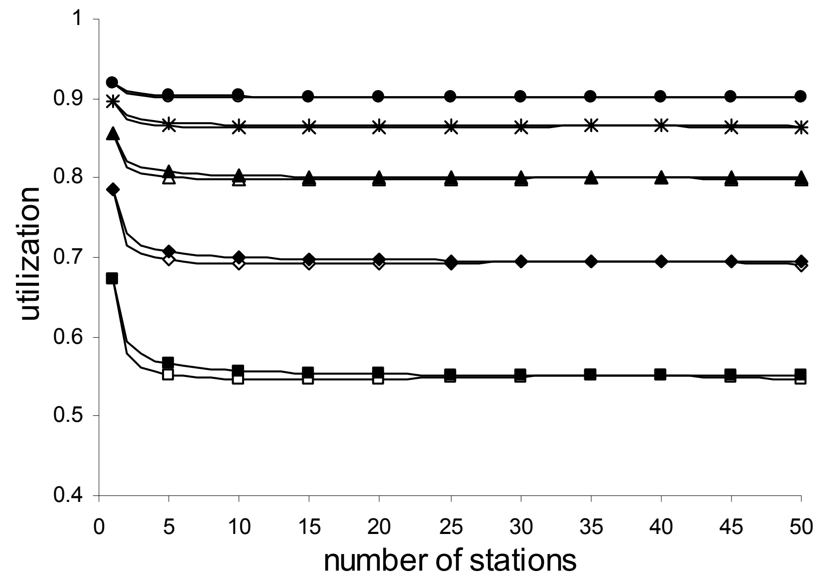
Fig. 6.10 considers the efficiency of the proposed parameter values for small networks by plotting utilization versus n for different CW_{min} (W) values. The figure



□ *n=1* ◇ *n=5*
△ *n=10* x *n=20*
◆ *n=50*

Figure 6.11 Utilization versus maximum backoff stage m, W=1, l=16Kbits, w=4, C=4Mbit/s, RR=1

shows that utilization is independent of *W* for large networks but highly increases for low *W* values for small networks. Significant utilization increase in observed by reducing *W* to one if only one or two stations contend for medium access, a usual case in real life wireless LANs. The situation is explained by considering that, for such small LANs, the proposed *W* value of 8 is significantly greater than the optimum W_{opt} value calculated by equation (6.37). Thus, utilization decreases due to the increased number of empty CAS. As a conclusion, for the considered network scenarios, *W* can be reduced to one in order to significantly increase utilization for small networks. Fig. 6.11 explores the suitability of the proposed CW_{max} value by plotting utilization versus maximum backoff stage *m* for different network sizes. Maximum backoff stage has minimal effect on utilization for small networks but a higher value is required for large networks. Utilization reaches its maximum value if *m* is greater than 20 even for LANs consisting of 50 transmitting stations. This value is significantly lower than the value of 62 proposed by the standard [42]. As a conclusion, new values for the *W* and *m* parameters are proposed; *W*=1 and *m*=20. Fig. 6.12 compares maximum with achievable utilization when the proposed *W* and *m* values are implemented. It shows that the achieved utilization is very close to the maximum. Direct comparison with Fig. 6.8 shows that



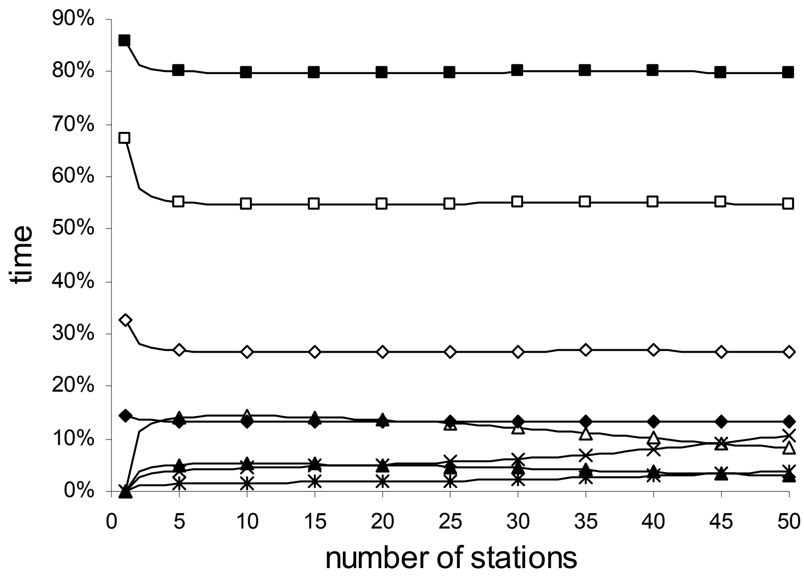
□ *w=1*	■ *w=1*, U_{max}
◇ *w=2*	◆ *w=2*, U_{max}
△ *w=4*	▲ *w=4*, U_{max}
x *w=8*	+ *w=8*, U_{max}
○ *w=16*	● *w=16*, U_{max}

Figure 6.12 Utilization versus n, l=16Kbits, W=1, m=20, C=4Mbit/s, RR=1

utilization is greatly improved for small networks due to the implementation of the lower *W* value.

6.4.4 The effect of high RH RR value and TAT delay

Fig. 6.13 plots the time portion utilized by all time consuming tasks versus network size for two window size values (*w*=1 and *w*=4) to demonstrate the factors resulting in utilization impairment for small payload window sizes. It shows that a higher window size value (*w*=4) results in significant utilization improvement. Time portion utilized on empty CAS and collisions is always minimal but for the small *w* value of one, a significant time portion is utilized in overheads. This time portion includes a) SDATA frame header transmission, b) control (RTS/CTS/EOB/EOBC) frame transmission and c) the TAT delays that follow control frames. The situation is explained by considering that if *w*=1, four control frames are exchanged for every successful reservation that includes only one data frame transmission. The RH field of these control frames is always transmitted using RR=16 to ensure maximum coverage. As a result, to cope with potential hidden nodes, applications must transmit large amounts of information data



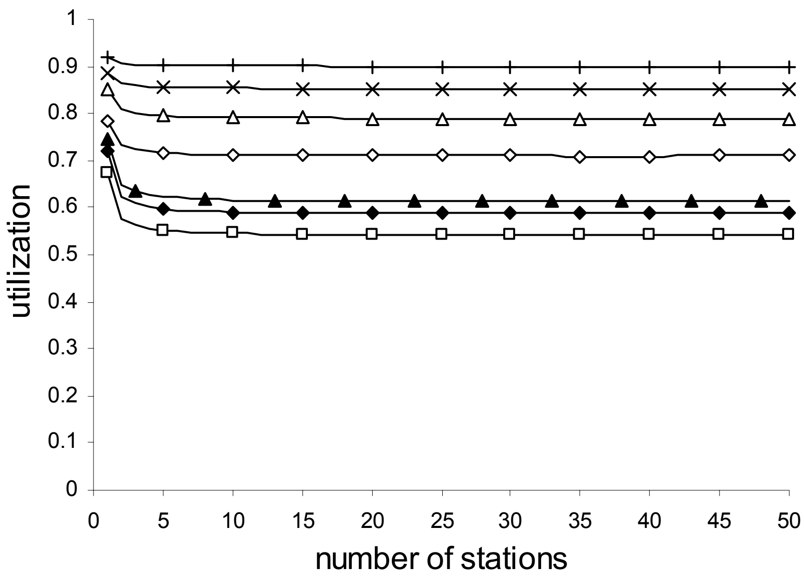
□ *payload data transmission (utilization), w=1* ■ *payload data transmission (utilization), w=4*
◇ *overheads, w=1* ◆ *overheads, w=4*
△ *empty CAS, w=1* ▲ *empty CAS, w=4*
x *collisions, w=1* ⁎ *collisions, w=4*

Figure 6.13 Time allocation of various AIr tasks versus n, W=1, m=20, l=16Kbits, C=4Mbit/s, RR=1

during a successful reservation. As a conclusion, the long CAS duration and the proposed *CW* adjustment algorithm do not result in significant utilization degradation.

Fig. 6.14 presents the utilization improvement if the RH field is transmitted using *RR*=1 and/or if the TAT duration (a physical layer parameter) is reduced for applications transmitting only 16Kbits (*w*=1) of information data at *RR*=1 during a successful reservation. Small improvement is observed if only the TAT delay is reduced. However, transmitting the RH field using RR=1 results in significant utilization improvement. Further reducing the TAT delay is beneficial. For all cases, a smaller CAS duration is considered, always obeying the restriction that $\sigma >= T_{RTS}+TAT+TT_{PA}+TT_{SYNC}$ [45] to ensure that a station not capable of hearing the RTS frame will hear the beginning of the corresponding CTS frame in the CAS duration.

As a conclusion, the AIr protocol requires high window size implementation for achieving high utilization due to the high RR used in transmitting the RH field; the long CAS duration and the proposed CW adjustment algorithm do not result in significant utilization degradation. A high utilization can be achieved by transmitting the RH field



□ *RH field in RR=16, TAT=200μsec, σ=800μsec*
◆ *RH field in RR=16, TAT=100μsecs, σ=700 μsec*
▲ *RH field in RR=16, TAT=50μsec, σ=650μsec*
◇ *RH field in RR=1,TAT=200μsec, σ=300μsec*
△ *RH field in RR=1, TAT=100μsec, σ=200μsec*
× *RH field in RR=1, TAT=50μsec, σ=100μsec*
+ *RH field in RR=1, TAT=10μsec, σ=50μsec*

Figure 6.14 Utilization versus n, W=1, m=62, l=16Kbits, w=1, C=4Mbit/s, RR=1

using RR=1, reducing the TAT delay and implementing the proposed CW adjustment algorithm.