Delay Distribution Analysis of IEEE 802.11 with Variable Packet Length

P. Raptis, V. Vitsas, A. Banchs, K. Paparrizos

Abstract—The distribution of the delay is a very important measure to determine the performance of a WLAN. Indeed, real-time applications typically require that their packets reach their destination within a certain delay with a given probability, and to guarantee this it is necessary to perform the computation of the delay distribution. In this paper we develop a novel method to compute the delay distribution of IEEE 802.11 Distributed Coordination Function (DCF) for a WLAN where the stations transmit packets of variable length. We develop an analytical model that applies to basic, RTS/CTS, and hybrid (coexistence of basic and RTS/CTS) access modes. The analytical results are validated by simulation results.

Index Terms—IEEE 802.11; Performance; Delay distribution; Variable packet lengths.

I. INTRODUCTION

The most deployed protocol for Wireless Local Area Networks (WLANs) worldwide is the IEEE 802.11. IEEE 802.11 specifies a Medium Access Control (MAC) protocol in order to share the medium [1]. MAC employs the Distributed Coordination Function (DCF) for contention-based channel access which uses the Binary Exponential Backoff (BEB) algorithm to minimize collision probability. DCF provides two techniques to transmit data packets; the basic access and the Request-To-Send/Clear-To-Send (RTS/CTS) access method.


Utilizing z-transform of the packet delay, [6],[7] and [8] calculated the probability generating function (pgf) which in turn is used to compute the probability distribution function (pdf) of the packet delay. The method employed by [6],[7] and [8] is computationally very expensive. Author in [9] proposed an effective model that computes the probability of a packet to be successful transmitted with delay time lower than a given value. Although less complex than [6],[7] and [8], the method of [9] still involves a number of relatively expensive operations. In [10] we proposed a simple, accurate and effective delay distribution analysis for basic access mode while in [11] for RTS/CTS mechanism. The model in [11] is also accurate, effective but more complex than [10].

The distribution of the delay is a very important measure to determine the performance of a WLAN. Indeed, real-time applications typically require that their packets reach their destination within a certain delay with a given probability, and to guarantee this it is necessary to perform the computation of the delay distribution. Previous analyses of the delay distribution [6-11] are commonly based on the unrealistic assumption of fixed length packets, which prevents their use to analyze realistic WLAN environments. In this paper we extend [11] and develop a novel method to compute the delay distribution of a WLAN under variable packet lengths. We develop an analytical model that applies to basic, RTS/CTS and hybrid access (coexistence of basic and RTS/CTS mechanisms).

The rest of the paper is organized as follows. Section II briefly describes DCF. Section III describes the mathematical model of the DCF. Section IV presents the delay distribution analysis. Section V describes the variable packet length analysis of the proposed model. Section VI presents the delay distribution with variable packet length analysis. Section VII validates the model by comparing analytical to simulative results and, finally, Section VIII presents the conclusions.

II. DISTRIBUTED COORDINATION FUNCTION

A. Basic access

Each station senses the channel before transmitting a packet. If the channel is idle for a period of Distributed Inter Frame Space (DIFS) then the station transmits its packet. If the channel is busy, the station defers until an idle DIFS is detected and then generates a random backoff interval before transmitting. The backoff time is sliced in terms of a slot time...
that a station 

+2σ

+ (10)

+ (5)

+ (1)

+ (8)

+ (6)

collides with constant and independent probability p regardless of the backoff stage. The probability τ that a station transmits a packet in a randomly chosen slot time can be expressed as [3]:

\[ \tau = \frac{2(1-p)(1-p^R)}{W(1-(2p)^{n-1})} \]

The probability \( p \) that a transmitted packet encounters a collision is given by:

\[ p = 1 - (1 - \tau)^{n-1} \]

Equations (1) and (2) represent a non-linear system with two unknown values, \( \tau \) and \( p \), which can be solved using numerical methods and has a unique solution.

Let \( q_o \) be the probability with that at least one station (out of \( n \)) transmits in a considered slot time:

\[ q_o = 1 - (1 - \tau)^n \]

Let \( q_i \) be the probability that a transmission occurring on the channel is successful. This can be computed as the probability that only one station transmits and the \( n-1 \) remaining stations do not, given the condition that a transmission occurs on the channel. Probability \( q_i \) is given by:

\[ q_i = \frac{n \cdot \tau \cdot (1 - \tau)^{n-1}}{1 - (1 - \tau)^n} \]

Let \( E[\text{slot}] \) be the average length of a slot time. \( E[\text{slot}] \) is given by:

\[ E[\text{slot}] = (1 - q_o) \sigma + q_o q_s T_s + q_o (1 - q_s) T_c \]

where \( \sigma \) is the duration of an empty slot, and \( T_s \) and \( T_c \) are the time durations the channel is sensed busy during a successful transmission and a collision, respectively.

The time duration of \( T_s \) and \( T_c \) depends on the channel access method employed. For the basic access method, we have:

\[ T_s^{\text{bas}} = T_c^{\text{bas}} = O^{\text{bas}} + \frac{l}{C} \]

where \( l \) is the packet length and \( O^{\text{bas}} \) is the packet overhead

\[ O^{\text{bas}} = DIFS + \frac{H}{C} + 2\delta + SIFS + \frac{ACK}{C} \]

III. DCF Mathematical Model

In this study we assume that the channel conditions are ideal (no transmission errors, no hidden stations and no capture effect), the number of contending stations \( n \) is fixed and each station has always a packet available for transmission (i.e., the station works in saturated conditions).

Let \( b(t) \) and \( s(t) \) be the stochastic processes representing the backoff time counter and the backoff stage \((0,...,R)\) respectively for a given station at time \( t \). The two-dimensional process \( \{ s(t), b(t) \} \) is a discrete-time Markov chain. In this study we utilize the Markov chain model of [3]. The key approximation in this model is that each packet transmission collides with constant and independent probability \( p \) regardless of the backoff stage. The probability \( \tau \) that a station transmits a packet in a randomly chosen slot time can be expressed as [3]:

\[ \tau = \frac{2(1-p)(1-p^R)}{W(1-(2p)^{n-1})} \]

The probability \( p \) that a transmitted packet encounters a collision is given by:

\[ p = 1 - (1 - \tau)^{n-1} \]

Equations (1) and (2) represent a non-linear system with two unknown values, \( \tau \) and \( p \), which can be solved using numerical methods and has a unique solution.

Let \( q_o \) be the probability with that at least one station (out of \( n \)) transmits in a considered slot time:

\[ q_o = 1 - (1 - \tau)^n \]

Let \( q_i \) be the probability that a transmission occurring on the channel is successful. This can be computed as the probability that only one station transmits and the \( n-1 \) remaining stations do not, given the condition that a transmission occurs on the channel. Probability \( q_i \) is given by:

\[ q_i = \frac{n \cdot \tau \cdot (1 - \tau)^{n-1}}{1 - (1 - \tau)^n} \]

Let \( E[\text{slot}] \) be the average length of a slot time. \( E[\text{slot}] \) is given by:

\[ E[\text{slot}] = (1 - q_o) \sigma + q_o q_s T_s + q_o (1 - q_s) T_c \]

where \( \sigma \) is the duration of an empty slot, and \( T_s \) and \( T_c \) are the time durations the channel is sensed busy during a successful transmission and a collision, respectively.

The time duration of \( T_s \) and \( T_c \) depends on the channel access method employed. For the basic access method, we have:

\[ T_s^{\text{bas}} = T_c^{\text{bas}} = O^{\text{bas}} + \frac{l}{C} \]

where \( l \) is the packet length and \( O^{\text{bas}} \) is the packet overhead

\[ O^{\text{bas}} = DIFS + \frac{H}{C} + 2\delta + SIFS + \frac{ACK}{C} \]

\( H \) is the packet header size (equal to the sum of MAC and physical headers), \( \delta \) is the propagation delay and \( C \) is channel bit rate.

For RTS/CTS method we have:

\[ T_s^{\text{rts}} = O^{\text{rts}} + \frac{l}{C} \]

\[ T_c^{\text{rts}} = DIFS + RTS + SIFS + CTS \]

where \( O^{\text{rts}} \) is the packet overhead

\[ O^{\text{rts}} = DIFS + RTS + 3SIFS + CTS + \frac{H}{C} + \frac{ACK}{C} \]

IV. Delay Distribution

The packet delay is defined as the time interval elapsed between the moment the packet is at the head of the MAC queue and the time when an ACK for this packet is received. The packet delay distribution is a discrete distribution whose the smallest time unit of the backoff counter is one time slot [7].

The first step of our mathematical model is to group packet time delays according to the number of collisions suffered by
a packet. Note that this number corresponds to the stage at which the successful transmission occurs [5].

The delay $D$ of a packet successfully transmitted at backoff stage $j$ takes a value of the form:

$$D = T_j + N_e + N_r + N_s + jT_e + T_i$$

for $0 \leq j \leq R$  \hspace{1cm} (11)

where $N_e$ is the number of empty slots that a packet encounters before its successful transmission, $\sigma$ the duration of an empty slot, $N_r(N_s)$ is the number of successful (collided) transmissions from the rest of the stations that a packet encounters before its successful transmission, $jT_e$ is the time that the packet occupies the channel with collisions until it reaches the $j$ stage and $T_i$ is the time to transmit successfully from the $j$th stage.

The key assumption upon which we base our analysis is the following one: we assume that the average relationship between $N_e$, $N_r$ and $N_s$ is the same one equal to the average, i.e.

$$\frac{N_e}{P_e} = \frac{N_r}{P_r} = \frac{N_s}{P_s}$$

\hspace{1cm} (12)

Note that, in average, for every $P_e$ slot times with a transmission (successful $P_s$ or unsuccessful $P_r$), we have $P_e$ empty slot times, where $P_r$ is the probability that at least one station out of $n-1$ transmits in a considered slot time, and $P_e$ is the probability that the channel is idle and are given by:

$$P_e = 1 - (1 - \tau)^{n-1}$$ and $P_r = 1 - P_e$

\hspace{1cm} (13)

From the above, we can express $D$ as a function of one variable: $N_r$

$$D = N_r(T_s + T_c(P_e/P_r) + \sigma(P_r/P_s)) + T_i$$

\hspace{1cm} (14)

From (14) we can compute $N_r$ for a given delay $D$

$$N_r = \text{round}_{int}\left(\frac{D - T_s}{T_c(P_e/P_r) + \sigma(P_r/P_s)}\right)$$

\hspace{1cm} (15)

The probability $P$ that a successfully transmitted packet encounters $N_r$ transmissions is:

$$P(D) = \sum_{j=0}^{k} P(D_j)$$

\hspace{1cm} (16)

where $P_j$ is the probability that the successful transmission occurs at stage $j$. The probability $P_j$ is computed as following:

$$P_j(N_j) = A_j \cdot x_j[N_j]$$ for $0 \leq j \leq m$

\hspace{1cm} (17)

$$P_m+N_u+1(N_s) = A_m+N_u+1 \cdot x_{m+N_u+1}[N_s]$$ for $m < m + u + 1 \leq R$

\hspace{1cm} (18)

where: $x$ are $R+1$ one-dimensional arrays

$$x_f = f_j[j]$$ for $j = 0$

\hspace{1cm} (19)

$$x_f = x_{j-1} \cdot f_j$$ for $1 \leq j \leq m$

\hspace{1cm} (20)

$$x_{m+N_u+1} = x_{m+N_u+1} \cdot f_m$$ for $m < m + u + 1 \leq R$

\hspace{1cm} (21)

($f$ are $m$ one-dimensional arrays)

$$A_j = \frac{(1-p)^{p_j}}{1-p^{p_j+1}} \frac{1}{2^{j(j-1)/2} W^{j/2}} \left(\frac{P_e}{P_s}\right)^{N_s}$$ for $0 \leq j \leq m$

\hspace{1cm} (22)

$$A_m+N_u+1 = \frac{(1-p)^{p_{m+N_u+1}}}{1-p^{p_{m+N_u+1}+1}} \frac{1}{2^{m(m+3)/2} W^{m+2}} \left(\frac{P_e}{P_s}\right)^{N_s}$$

\hspace{1cm} (23)

and $u=0,1,2…$ is the remaining number of retransmissions trials in order to reach $R$. In (20) and (21) the * denotes convolution operation (convolution is the same operation as multiplying the polynomials whose coefficients are the elements of $x$ and $f$).

Let $b_j$ and $h_j$ be $m$ arrays that store the following values:

$$b_j[i] = P_e^i$$ for $0 \leq i \leq W_j - 1$, $0 \leq j \leq m$

\hspace{1cm} (24)

$$h_j[i,i] = \left(\frac{P_{r_j}}{P_s}\right)^i$$ for $0 \leq i \leq W_j - 1$, $0 \leq j \leq m$

\hspace{1cm} (25)

Let now store the values of the following array multiplication in $f_j$ arrays (of $W_j$ elements):

$$f_j[k] = z_j[k,0:le] \cdot h_j$$ for $0 \leq k \leq W_j - 1$, $0 \leq j \leq m$

\hspace{1cm} (26)

where $z_j$ be $m$ temporary arrays of $W_j \times W_j$ elements that store the following array values:

$$z_j[k,le] = z_j[k,0:le] = t_a[i,k] : le]$$ for $0 \leq k \leq W_j - 1$

\hspace{1cm} (27)

and $k : le$ means from line (column) $k$ to the last line (column) $le$. The $t_a$ is a temporary array of $1 \times W_j$ elements that store the results of the following array multiplication:

$$t_a = b_j \cdot a_{j,k}$$

\hspace{1cm} (28)

and $a_{j,k}$ are $m$ two-dimensional arrays of $W_j \times W_j$ elements where we store the following numbers:

$$a_{j,k,i} = C_i$$

\hspace{1cm} (29)

for $0 \leq k \leq (W_j - j)/2$,

\hspace{1cm} (30)

$$k_s \leq k \leq W_j - 1$$,

\hspace{1cm} (31)

$$k_s + k_c \leq j \leq W_j - 1$$ and $0 \leq j \leq m$

\hspace{1cm} (32)

Equation (29) is inside a triple loop where (30) expresses the outer, (31) the middle and (32) the inner loop.

Let now compute $C_i$ of (29). If we know the value of $C_i$ for $i = k_s + k_c$ then we can compute $C_{i+1}$ for $i+1$ keeping $k_s$ and $k_c$ stable:

$$C_{i+1} = C_{i+1} \cdot \left(\frac{i+1}{(i+1-k_s-k_c)!}\right)$$ for $k_s + k_c \leq i \leq W_j - 1$

\hspace{1cm} (33)

and let be $C_i = V_{k_c}$. Now, if we know the value of $V_{k_c}$ for $k_c = k_s$ then we can compute $V_{k_c+1}$ for $k_c = k_c$ keeping $k_s$ stable and is given by:

$$V_{k_c+1} = V_{k_c} \cdot \frac{(k_c + k_c + 1)}{k_c + 1}$$ for $k_s \leq k_c \leq W_j - 1$

\hspace{1cm} (34)

and let be $V_{k_c} = y_j[k_c]$. We know the value of $y_j[k_c]$ as we can create it as following:

$$y_j[k_c] = 1$$ for $k_c = 0$

\hspace{1cm} (35)

$$y_j[k_c + 1] = y_j[k_c] \cdot (2k_c - 1)(2k_c) / k_c^2$$ for $m < m + u + 1 \leq R$
for $1 \leq k \leq (W - 1)/2$ (36)

where $y_j$ are $m$ one-dimensional arrays of $W_j/2$ elements.

V. VARIABLE PACKET LENGTH

We consider that a packet length takes a value $l$ of the set $L$ with probability $P_l$, where $L$ is the set of all possible packet lengths. For simplicity, we assume that all stations pick the packet length from the same distribution ($l, P_l$) (the analysis would be very similar in the case when the stations pick the packet length from different distribution sets).

Let $E[l]$ be the average packet length of all possible $l$ in the set $L$. $E[l]$ is computed as following:

$$E[l] = \sum_{l \in L} l \cdot P_l$$ (37)

Let $P_k$ be the probability that exactly $k$ stations are involved in one collision:

$$P_k = \binom{n}{k} \frac{\tau^k (1 - \tau)^{n-k}}{q_{y_0} (1 - q_{y_0})} \text{ for } k \geq 2$$ (38)

Let $P_y$ be the sum of probabilities that corresponds to packet lengths that are shorter or equal to the longest data packet $l$ in a collision:

$$P_y = \sum_{l < L} P_l$$ (39)

Let $P_{l,k}$ be the probability that a packet with length $l$ is the longest packet in a collision when $k$ stations are involved in a collision:

$$P_{l,k} = \left( \frac{k}{1} \right) P_l P_{l+1} - \left( \frac{k}{2} \right) P_l^2 P_{l+2} + \left( \frac{k}{3} \right) P_l^3 P_{l+3} - \cdots \left( \frac{k}{k} \right) P_l^k P_0$$

for $2 \leq k \leq n$ (40)

So, for $k=2$ and $k=3$ the (40) becomes:

$$P_{l,2} = 2P_l P_y - P_l^2$$ (41)

$$P_{l,3} = 3P_l^2 P_y - 3P_l P_{l+1} + P_{l+1}$$ (42)

From eq. 40 we can compute the probability that any number of stations can be involved in one collision.

Let us now compute the time duration of $T_s$ and $T_c$ for the different access schemes.

A. Basic Access

The time $T_s$ for a successful transmission is the sum of packet’s overhead $O^{bas}$ plus the average payload:

$$T_s = O^{bas} + \frac{1}{C} \cdot E[l]$$ (43)

In the basic access collisions occur between data packets. The time duration of a collision is the time of the longest data packet involved in a collision:

$$T_c = O^{bas} + \frac{1}{C} \sum_{l \in L} \sum_{k=2}^{n} P_{l,k} P_{l,k}$$ (44)

B. RTS/CTS Access Mode

In the RTS/CTS scheme the computation of $T_s$ is similar to that of basic access, while a collision lasts $T_c^{rsa}$ (9):

$$T_s = O^{rsa} + \frac{1}{C} \cdot E[l]$$ (45)

$$T_c = T_c^{rsa}$$ (46)

C. Hybrid Access

The times $T_s$ and $T_c$ in hybrid access depend on the RTS-threshold value $l_{TH}$. If the data packet is less than the threshold then basic access is used otherwise RTS/CTS scheme is used.

The $T_s$ time is computed as following:

$$T_s = O^{bas} \sum_{l \in L} P_l + O^{rsa} \sum_{l \in L} P_l + \frac{1}{C} \cdot E[l]$$ (47)

In the hybrid access mode may occur three possible collision scenarios: 1) collision between data packets, 2) collision between RTS frames and data packets, and 3) collision between RTS frames. In the first and the second scenario the collision depends on the longest data packet involved in the collision as the packet header $H$ is always higher than the length of a RTS frame [1]. In the third scenario the duration of a collision is the time of a RTS frame. For the first and the second scenario the collision time is computed as:

$$T_{c,1,2} = O^{bas} \sum_{l \in L} P_l + \frac{1}{C} \sum_{l \in L} \sum_{k=2}^{n} P_{l,k}$$ (48)

and for the third scenario as:

$$T_{c,3} = O^{rsa} \sum_{l \in L} \sum_{k=2}^{n} P_{l,k}$$ (49)

Finally, from (48) and (49) we compute the collision time $T_c$ for the hybrid access:

$$T_c = T_{c,1,2} + T_{c,3}$$ (50)

VI. DELAY DISTRIBUTION WITH VARIABLE PACKET LENGTH

Let $P_{p,l}$ be the partial probability that a packet of length $l$ (of the set $L$) is successfully transmitted after experiencing delay $D$ is given by:

$$P_{p,l} = P_l \cdot P(D)$$ (51)

In order to compute $P_{p,l}$ we follow the next steps, 1) we decide the access mode for a WLAN, 2) we compute $T_s$ and $T_c$ for the chosen access mode (as described in Section V), 3) for a given delay $D$ we compute $N_e$ from eq. 15, 4) we compute total probability $P$ from eq. 16, and finally 5) we substitute $P$ to eq. 51.

The total probability $P$ is the sum of the all partial probabilities for a given delay $D$:
\[ P(D) = \sum_{i\in L} P_{i,j} \] (52)

VII. VALIDATION AND RESULTS

In order to validate our model we compare simulative to analytical results. The parameter values used for both simulation and analytical results follow the values specified for the Direct Spread Sequence Spectrum (DSSS) employed in the IEEE 802.11b standard and are shown in Table I. The packet length \( l \) is taken from the set \( L = \{1, 5, 10, 12\} \) Kbits with corresponding probabilities \([0.5, 0.05, 0.15, 0.3]\).

### Table I. System Parameter Values

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Channel bit rate</td>
<td>1 Mbit/s</td>
</tr>
<tr>
<td>Packet Payload</td>
<td>Variable length</td>
</tr>
<tr>
<td>MAC header</td>
<td>224 bits</td>
</tr>
<tr>
<td>PHY header</td>
<td>192 bits</td>
</tr>
<tr>
<td>ACK</td>
<td>112 bits + PHY header</td>
</tr>
<tr>
<td>RTS</td>
<td>160 bits + PHY header</td>
</tr>
<tr>
<td>CTS</td>
<td>112 bits + PHY header</td>
</tr>
<tr>
<td>Propagation delay, ( \delta )</td>
<td>1 ( \mu )s</td>
</tr>
<tr>
<td>Slot time, ( \sigma )</td>
<td>20 ( \mu )s</td>
</tr>
<tr>
<td>SIFS</td>
<td>10 ( \mu )s</td>
</tr>
<tr>
<td>DIFS</td>
<td>50 ( \mu )s</td>
</tr>
<tr>
<td>Minimum ( W, W_o )</td>
<td>32</td>
</tr>
<tr>
<td>Number of ( W ) sizes, ( m )</td>
<td>5</td>
</tr>
<tr>
<td>Short retry limit, ( R )</td>
<td>6</td>
</tr>
</tbody>
</table>

Fig. 1 plots the delay distributions for packet lengths of the set \( L \) (partial probabilities vs. delay) for a network of \( n=25 \) using hybrid access. Fig. 2 and Fig. 3 plot the same as Fig. 1 but for basic access and RTS/CTS access respectively.

For the hybrid access we used threshold value \( l_{TH} = 1860 \) bits. The model is accurate as the analytical match the simulation results for all access modes. All simulation results are taken with a 95% confidence interval lower than 0.0001. The results of the hybrid access in figure 1 and the RTS/CTS access in figure 3 are very similar as the threshold value we used in the hybrid access had little effect on the delay distributions for the chosen packet lengths (set \( L \)) and \( P_j \)'s. If we zoom Fig. 1 in x-axis from 0.0-0.5 secs we get Fig. 4 (note that y-axis is in linear scale). The figure confirms that the model is accurate as the simulative lines closely follow analytical lines.
The assumption that the stations in a WLAN transmit packets of variable sizes (such as voice and data packets) is more realistic than the assumption of fixed packet sizes. In this paper we develop a novel method to compute the delay distribution of IEEE 802.11 DCF for a WLAN where the stations transmit packets of variable length. We develop an analytical model that applies to basic, RTS/CTS, and hybrid access modes. Our model can compute the probability that a packet (of any length) will be successfully transmitted for a given delay value (the packet lengths are taken from a given distribution). The analytical results are validated by simulation results.

VIII. CONCLUSIONS

REFERENCES