Energy and Delay Analysis of Wireless Networks with ARQ\textsuperscript{∗}

Shih Yu Chang, Achilleas Anastasopoulos and Wayne E. Stark
Electrical Engineering and Computer Science Department
University of Michigan
Ann Arbor, MI 48109-2122.
email: \{shihyuc,anastas,stark\}@umich.edu

Abstract—In this paper, we consider a variant of the 802.11 protocol augmented with an automatic repeat request (ARQ) mechanism for data transmission. An innovative analysis method that provides the joint distribution of energy and delay via their joint generating function is proposed. The generating function is used to evaluate the average energy and delay of a successful data packet transmission. Our analysis takes into account the effect of channel noise on the transmission of the RTS, CTS, Data, and ACK packets. Our analysis indicates that the proposed 802.11 variant performs almost identically to the original protocol and that both are sensitive to the knowledge of channel conditions.

I. INTRODUCTION

Recently there has been considerable interest in the design and performance evaluation of wireless local area networks (WLANs). According to OSI protocol layers specification, the physical layer (lowest layer) handles the transmission of raw bits through a communication link. The forward error control (FEC) coding technique is used to mitigate the effect of channel noise at the receiver. The error control coding technique reduces the required energy needed for transmission at the expense of increased delay and reduced data rate. The second layer is the data link layer which is responsible for establishing a reliable and secure logical link over the unreliable wireless link. The lower sublayer of data link layer is the medium access control (MAC) protocol layer which resolves conflicts between two users attempting to access the channel. Thus there is some amount of energy and delay spent in reserving the channel. The upper sublayer of data link layer is logical link control (LLC) sublayer that implements error control function. The most common technique to combat errors used in this sublayer is automatic repeat request (ARQ). A tradeoff needs to be considered in this sublayer is energy and delay. In this paper we provide a method to analyze and optimize the energy and delay of WLANs. We incorporate the data link layer including medium access control (MAC) and logical link sublayer with the physical layer in the analysis.

In order to combat the well known hidden terminals problem [1], a request-to-send/clear-to-send (RTS/CTS) mechanism described in IEEE Standard [2] is used for packet transmission. Before transmitting a data packet, a transmitting station reserves the channel by sending a particular short packet: request-to-send (RTS). If the RTS packet is received correctly at the corresponding destination station, the destination station acknowledges the transmitting station by sending back a clear-to-send (CTS) packet. If the above channel reservation is successful, the transmitting station will begin to send a data packet (DT) and wait for the acknowledge (ACK) packet from the destination station. The transmitting station detects the unsuccessful channel reservation by the lack of a CTS packet. The RTS/CTS scheme increases system performance by reducing the probability of a collision when long data packets are transmitted.

Previous performance evaluation of 802.11 has been carried out either by means of simulation ([3] and [4]) or by means of analytical models ([5], [6] and [7]). Most of the above consider the case of an error free channel. In [7], the effects of imperfect channel condition were investigated. However, the relation of error probability to delay and energy has not been considered. In this paper, we investigate the energy and delay relation of networks using a model that incorporates the different packet error probabilities as a function of energy and delay.

The first contribution of this paper is that we give a set of nonlinear system equations that characterize the best packets length which maximize the average system throughput. The second contribution of this paper is that we propose a state diagram to represent the operation of the system. We determine the generating function of the state diagram and use it to characterize the joint energy and delay distribution of the system incorporating physical layer and data link layer characteristics. By using the random coding bound we can optimize the system performance over the code rate and energy per coded bit. Although we use random coding under AWGN channel to represent packet error probabilities, the framework of our analysis can be used for other coding and modulation schemes under various wireless channel. The third contribution is that the IEEE 802.11 original protocol and the proposed one have almost identical performances and are equally sensitive to the knowledge of the channel quality at the transmitter.

The remainder of this paper is organized as follows. In section II, we describe the system we are considering. In section III, we introduce the generating function and determine the joint generating function for energy and delay. In section IV, we present numerical results to demonstrate the energy-delay tradeoff with random coding and the improvement of average throughput. The conclusions are stated in section V.

II. SYSTEM DESCRIPTION

A. Stop-and-Wait ARQ with Finite Transmission (SWARQ-FT)

Suppose each transmitter has \( K \) (fixed) information bits to be transmitted. These \( K \) bits will be encoded to form a packet with length \( N_1 \) and transmit. If the receiver decodes this packet correctly, it sends an ACK to the transmitter. Otherwise, if decoding is not correct, after the expiration of a timeout the transmitter encodes the same number of information bits \( K \) to

\textsuperscript{*} This work was supported by the Office of Naval Research under Grant N00014-03-1-0232

U.S. Government work not protected by U.S. copyright
a packet of length $N_2$ and transmits it again. This procedure will continue (using length $N_l$ packets for the $l$th transmission) and will terminate when either the transmitter receives the ACK correctly or the total number of attempts achieves a maximum allowed number, $d$.

B. IEEE 802.11 MAC Protocol with SWARQ-FT

The wireless networks that we analyze here have the following specifications. First, each station with a fixed position can hear (detect) the transmission of $n - 1$ other neighbor stations in the network. Second, stations always have a packet ready to transmit. Third, the receivers have no multiple-access capability (i.e., they can only receive one packet at a time). Fourth, all stations have fixed position (no mobility). Finally, each station uses 802.11 type protocol for medium access control (MAC).

In the following we give a brief description of the most important features of the IEEE 802.11 MAC protocol (more details can be found in [2]). When a station is ready to transmit a packet, it picks random a number $j_i$ uniformly distributed in $\{0, 1, \ldots, W_i - 1\}$, where $W_i = 2^i W$ is the contention window (CW) size, $i$ is the contention stage (initially $i = 0$), and $W$ is the minimum CW size. A backoff time counter begins to count down with an initial value $j_i$: it decreases by one for every idle slot of duration $\sigma$ seconds as long as the channel is sensed idle, stops the count down when the channel is sensed busy, and reactivates when the channel is sensed idle again. The station transmits a request-to-send (RTS) packet when the counter counts down to zero. After transmitting the RTS packet, the station will wait for a clear-to-send (CTS) packet from the receiving station. If there is a collision of the RTS packet with other competing stations or a transmission error occurs in the RTS or CTS packet, the transmitting station doubles the CW size (increases the contention stage $i$ by one) and picks a random number $j_i$ from a set of size $W_{i+1}$. If there are no collisions or errors in the RTS and CTS packets, the station begins to transmit the data packet by using SWARQ-FT that is described in section II-A. However, if the source station cannot receive the ACK packet correctly after $d$ (maximum allowed transmission number) transmissions, then the CW size will also be doubled (the contention stage $i$ will increase by one) and the transmitting station will join the contention period again. The contention stage is reset ($i$ is set to zero) when the transmitting station receives an ACK correctly. A time diagram indicating the sequence of these events is depicted in Fig. 1. It is also noted that there is a maximum CW size (or equivalently, a maximum contention stage, $m$); when the transmitter is in this maximum stage and needs to join the contention period again, it does not increase further the contention window, but picks a random number in $\{1, \ldots, 2^m W\}$.

III. Analysis

In this section, we analyze the energy and delay characteristics of the wireless networks described above. The delay $T_d$ of each data packet is defined as the time duration from the moment that the backoff procedure is initiated until DIFS

seconds after the ACK packet is received correctly by the transmitting station, as shown in Fig. 1. Similarly, the energy $E_i$ is defined as the energy consumed by both transmitting and receiving stations in the duration of $T_d$.

We begin by calculating some important event probabilities that are crucial in later analysis. The performance evaluation is based on the Markov chain model described in [6] and [7]. The state of the Markov chain represents the value of the backoff counter and the contention window stage. Our model differs from the model proposed in [6], [7] in that packet error probabilities are included in the CW stage transition probability. By modifying the Markov chain model described in [6], [7], we can incorporate the packet error probabilities into the analysis. We denote the error probability of an RTS packet as $P_e,_{RTS}$, a CTS packet as $P_e,_{CTS}$ and an ACK packet as $P_e,_{ACK}$. The probability $P_e,_{DT}$ represents the error probability of $i$th transmission of a data packet, where $1 \leq i \leq d$. We assume that the channel is memoryless between packets. These probabilities depend on the particular channel, channel coding, modulation etc. (a specific example will be given in Section IV). Let $p_{ce}$ represent the transition probability (as defined in [6]) from one CW stage to the next. This is also the probability of an unsuccessful transmission attempt seen by the transmitting station when its packet is being transmitted on the channel. However, in this work, an unsuccessful transmission can happen not only due to packet collision, but also due to packet errors. There are four possible cases when the channel is sensed busy. The first case is that the RTS packet suffers a collision or packet error. The second case is that the RTS packet is correctly received and collision free but there is an error in the CTS packet transmission. The third case is that the RTS packet is correctly received and collision free, there is no error in CTS packet, but there is an error in all $d$ transmissions of the DT packet. The last case is that the data packet is correctly received in one of these $d$ transmissions but there is an error in the corresponding ACK packet. Because the event of packet collision and the event of packet error are
independent, \( p_{cc} \) can be expressed as

\[
p_{cc} = p_c + (1 - p_c)\left[ P_{e,\text{RTS}} + (1 - P_{e,\text{RTS}})P_{e,\text{CTS}} + (1 - P_{e,\text{RTS}})(1 - P_{e,\text{CTS}}) \prod_{i=1}^{d} P_{e,\text{DT}_i} + (1 - P_{e,\text{RTS}}) \right].
\]

where \( p_c \) is the packet collision probability and we set \( P_{e,\text{DT}_0} = 1 \). The packet transmission probability \( p_{tx} \) which is the probability of a transmitting station sending an RTS packet during each backoff slot, and the collision probability \( p_c \) can be expressed as (see [6])

\[
p_{tx} = \frac{2(1 - 2p_{cc})}{(1 - 2p_{cc})(W + 1) + p_{cc}W(1 - (2p_{cc})^m)}
\]

and

\[
p_c = 1 - (1 - p_{tx})^{n-1}.
\]

\( \rho \)From the above three nonlinear equations, (1)–(3), one can evaluate the probabilities \( p_{cc}, p_{tx} \) and \( p_c \). Another important probability that will be used in our analysis later is the probability of a transmission of an RTS packet from exactly one of the remaining \( n-1 \) stations given that at least one of the remaining stations is transmitting. It is denoted by \( p_{tx_1} \) and can be expressed as

\[
p_{tx_1} = \frac{(n - 1)p_{tx}(1 - p_{tx})^{n-2}}{p_c}.
\]

We assume that stations spend energy only in packets transmission and that the energy consumption for each packet transmission is proportional to the packet length. In particular, let \( \alpha E_e \) be the energy required to transmit one coded bit, where \( \alpha > 1 \) is the inverse of the path loss between transmitter and receiver. Then the energy received for a packet of length \( N \) (normalized to the thermal noise energy level, \( N_0 \)) is \( N E_e/N_0 \). Similarly, the transmission delay of each packet is also proportional to the packet length. Let \( R_t \) be the transmitting rate in bits per second. The delay, in seconds, for a packet transmission with packet length \( N \) is \( N/R_t \). Based on the protocol description in Section II, the state flow diagram shown in Fig. 2 can be built.

Let \( T_{\text{RTS}}, T_{\text{CTS}}, T_{\text{DT}_i} \) and \( T_{\text{ACK}} \) denote the time du-
ration for the transmission of RTS, CTS, \( k \)-th DT and ACK packets, respectively. Similarly, let \( E_{\text{RTS}}, E_{\text{CTS}}, E_{\text{DT}_i} \) and \( E_{\text{ACK}} \) denote the (received) energy for the transmission of RTS, CTS, \( k \)-th DT and ACK packets, respectively. Finally, let \( T_{\text{DIFS}} \) and \( T_{\text{SIFS}} \) be system delay parameters defined by the standard. We will now derive the joint generating function of the delay and energy random variables associated with the transitions in the state diagram in Fig. 2. In all the derivations below, functions are of the form \( G(X,Y) \), where the variables \( X, Y \) are the transform variables of the delay and energy, respectively.

The generating function \( G_{rs} \) corresponding to successful channel reservation is

\[
G_{rs}(X,Y) = (1 - p_c)(1 - P_{e,\text{RTS}})(1 - P_{e,\text{CTS}})X^{T_{\text{RTS}}+T_{\text{CTS}}+2T_{\text{SIFS}}}Y^{E_{\text{RTS}}+E_{\text{CTS}}}.
\]

Similarly, the generating function \( G_{ts} \) corresponding to successful transmission of a data packet can be expressed as

\[
G_{ts}(X,Y) = \sum_{k=0}^{N/R_t} \frac{1}{(1 - P_{e,\text{RTS}})(1 - P_{e,\text{ACK}})} \cdot X^{kT_{\text{RTS}}+T_{\text{DIFS}}+T_{\text{ACK}}}Y^{kT_{\text{RTS}}+T_{\text{DIFS}}+T_{\text{ACK}}+T_{\text{SIFS}}+T_{\text{DIFS}}E_{\text{RTS}}+E_{\text{CTS}}}
\]

We now develop expressions for the generating functions associated with failure to either reserve a channel or to transmit data. As can be seen from the state diagram, these functions are products of two generating functions. Each product contains a factor that is independent of the particular contention stage, and a factor that depends on the contention stage \( i \). We first describe the constant factors. The generating function \( G_{fr} \), corresponding to failure to reserve the channel is given by

\[
G_{fr}(X,Y) = [p_c + (1 - p_c)P_{e,\text{RTS}}]X^{T_{\text{RTS}}+T_{\text{DIFS}}Y^{E_{\text{RTS}}} + (1 - p_c)(1 - P_{e,\text{RTS}})P_{e,\text{CTS}}}X^{T_{\text{RTS}}+T_{\text{CTS}}+2T_{\text{SIFS}}+T_{\text{DIFS}}Y^{E_{\text{RTS}}+E_{\text{CTS}}}}.
\]
captured by the generating function $G_{bf}$ as follows

$$G_{bf}(X,Y) = \prod_{i=1}^{d} P_{c,DT_i} X_{i=1}^{d} T_{DT_k} + (d-1) T_{SIFS} + T_{DIFS},$$

$$Y_{i=1}^{d} E_{DT_k} + \sum_{i=1}^{d-1} \prod_{j=0}^{i-1} P_{c,DT_j}(1-P_{c,DT_i}) P_{c,ACK}.$$  

$$X_{i=1}^{d} T_{DT_k} + T_{ACK} + i T_{SIFS} + T_{DIFS} Y_{i=1}^{d} E_{DT_k} + E_{ACK}.$$  

(8)

We now evaluate the generating functions denoted by $G_{bs,i}$ of the state diagram. This generating function characterizes the delay for the transmitting station from the instant of starting the backoff procedure to the instant that the backoff counter reaches to zero at stage $i$. We do not need to consider energy consumption here since the transmitting station stops any transmission during this period. The event of each slot being sensed busy due to the transmission of other stations is independent for each slot and has the same probability, $p_c$, for each slot (from our previous assumptions). At backoff stage $i$, the range of possible backoff slots is from 1 to $2^i W$. Let $j$ be the backoff slots randomly chosen uniformly from the above range, then the number of the occupied slots in these $j$ slots is binomially distributed with parameters $(j, p_c)$. Hence, the generating function, $G_{bs,i}$, of backoff procedure at stage $i$ is

$$G_{bs,i}(X) = \sum_{j=1}^{2^i W} \frac{1}{2^i W} \sum_{k=0}^{j} \binom{j}{k} (1-p_c) X^j - (p_c G_{oc}(X))^k,$$

(9)

where $G_{oc}(X)$ is the generating function for an occupied slot. We define an occupied slot as a slot when the transmitting station senses the channel is busy due to the transmission of one of the remaining $n-1$ stations. Using arguments similar to the ones used in the derivation of (1), we can express $G_{oc}(X)$ as follows

$$G_{oc}(X) = \left[ (1-p_{tx1}) + p_{tx1} P_{c,RTS} \right] X^{T_{RTS} + T_{DIFS} + T_{RTS} + T_{CTS} + T_{SIFS} + T_{DIFS} + T_{RTS} + T_{CTS}} +$$

$$p_{tx1}(1-P_{c,RTS})(1-P_{c,CTS}) \prod_{i=1}^{d} P_{c,DT_i},$$

$$X^{T_{RTS} + T_{CTS} + T_{DIFS} + T_{RTS} + T_{CTS} + T_{SIFS} + T_{DIFS}} +$$

$$p_{tx1} (1-P_{c,RTS})(1-P_{c,CTS}) \sum_{i=1}^{d} P_{c,DT_i},$$

$$X_{i=1}^{d} T_{DT_k} + T_{ACK} + (i+2) T_{SIFS} + T_{DIFS}.$$  

(10)

From the state diagram and Mason’s gain formula, we obtain the following backward recursive generating function to characterize the energy and delay of the system. Let $G_{i}$ be the system generating function. Then the recursion is

$$f_{m}(X,Y) = \frac{G_{r} G_{ts}}{1 - G_{f} G_{bs,m} - G_{r} G_{bf} G_{bs,m}}$$

$$f_{i-1}(X,Y) = G_{r} G_{ts} + (G_{f} + G_{r} G_{bf}) G_{bs,i} f_{i}$$

$$G_{s}(X,Y) = G_{bs,0} f_{0},$$  

(11)

where $i$ is the index of the CW stage from 1 to $m$. Hence, we can evaluate any joint statistics of the energy and delay of a successful packet transmission. In particular, the mean values are given by

$$T_d = \frac{\partial G_s}{\partial X} \bigg|_{X=Y=1}, \quad \mathbb{E} = \frac{\partial G_s}{\partial Y} \bigg|_{X=Y=1}.$$  

(12)

IV. NUMERICAL RESULTS

We will use channel reliability based bounds for an AWGN channel to estimate the packet error probability. Let $K$ be the number of information bits in a packet (specified by the 802.11 standard) and $N$ be the number of coded bits in a packet. Then there exist an encoder and decoder with packet error probability $P_{c,K,N}$ bounded as

$$P_{c,K,N} \leq 2^{K-N} R_0,$$  

(13)

where $R_0 = 1 - \log_2(1+e^{-E_c/N_0})$ is the cutoff rate determined by signal-to-noise ratio.

The energy-delay curves of the IEEE 802.11 protocol with SWARQ-FT $(d = 2)$ is evaluated for different number of users and data packet lengths. For comparison, we also present similar curves for the original 802.11 protocol, which is essentially an SWARQ-FT protocol with $d = 1$. For both protocols there is an optimal $N$ for each kind of packets. For small $N$ the packet error probability is large, which increases the system delay and energy due to the high chance of packet retransmission. On the other hand, if $N$ is large, the packet error probability is small but the transmission time is large. Our goal is to find the best $N$ for each kind of packet to minimize system delay. We first fix $E_c/N_0$ and minimize $\bar{T}_d$ from (12) to get the corresponding optimal packet lengths $N_{RTS}^*, N_{CTS}^*, N_{ACK}^*$ and $N_{DTI}^*, N_{DT2}^*$ (note that in the case of $d = 2$ we need to optimize over two data packet lengths, $N_{DTI}$ and $N_{DT2}$). Using these values, we can evaluate both average delay and average energy consumption. Repeating the above procedure for different $E_c/N_0$, we get the energy-delay curves. Table I summarizes the system parameters used in our numerical evaluations.

We first plot in Fig. 3 the energy-delay curves for the original protocol and the IEEE 802.11 MAC Protocol with SWARQ-FT $(d = 2)$. We observe that the energy-delay curves for two protocols are almost identical. This shows that if we can select packet lengths properly, the SWARQ-FT does not give significant improvement. To compare these two protocols with respect to their robustness to signal-to-noise ratio (SNR), we fix the packet lengths, vary $E_c/N_0$ and plot energy and delay for both protocols. We observe that both systems are equally and very sensitive to SNR. Essentially, the reduction in

| TABLE I
<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>System Parameters for Numerical Results</strong></td>
</tr>
<tr>
<td>$K_{RTS}$</td>
</tr>
<tr>
<td>$K_{CTS}$</td>
</tr>
<tr>
<td>$K_{ACK}$</td>
</tr>
<tr>
<td>Channel Bit Rate, $R_c$</td>
</tr>
<tr>
<td>Slot Time, $\sigma$</td>
</tr>
<tr>
<td>$T_{SIFS}$</td>
</tr>
<tr>
<td>$T_{DIFS}$</td>
</tr>
</tbody>
</table>
delay the SW ARQ achieves by immediately retransmitting an erroneously received packet, is compensated by the increased delay needed to access the channel, since other users use the same protocol. While ARQ provides some robustness to SNR uncertainty in conventional single user system, this is not the case in this multiuser scenario.

In Fig. 4, the energy-delay tradeoff is plotted for the SW ARQ protocol for different values of quantity $N_{DT} - K_{DT}/R_0(E_c/N_0)$. This figure shows that the optimal packet lengths are on the order of 60 bits larger than the minimum length implied by the cutoff rate. For high SNR values as small as 10 will provide almost optimal performance, while for low SNR, the number of redundant bits has to increase.

![Fig. 3. Energy-delay curves for $n = 10$ and $K_{DT} = 1400$. Solid lines represent the curves after packets optimization. Dashed lines represent the energy-delay curves for $E_c/N_0$ of 0 and 3 dB using the optimal packet lengths.](image)

![Fig. 4. The dashed lines represent SWARQ after optimization. The number beside the curve is the redundant bits.](image)

V. CONCLUSION

The primary contribution of this work is to analyze the IEEE 802.11 protocol with SW ARQ-FT using a state diagram by obtaining the joint generating function of the energy and delay. After optimizing the system delay over the code rate for each packet, we obtained the energy-delay tradeoff curves. Our results show that both protocols are sensitive to the knowledge at the transmitter of the received SNR. This is a consequence of the steep falloff of packet error probability with respect to SNR for random coding. We expect similar results when LDPC or turbo codes are used.

REFERENCES


