Delay Analysis of Different Backoff Algorithms in IEEE 802.11

Ivan N. Vukovic
Motorola Inc., 1441 W. Shure Dr., Arlington Heights, IL, 60004 USA, Ivan.Vukovic@motorola.com, (847) 632-5963, fax (847) 435-9970

Natt Smavatkul
Motorola Inc., 1301 E. Algonquin Rd., Schaumburg, IL, 60196 USA, Natt@motorola.com

Abstract - **In this paper, we use a newly developed analytical model to compute average delay of the existing binary exponential backoff (BEB) algorithm in IEEE 802.11 wireless LAN and two different proposals, exponential increase exponential decrease (EIED) and exponential increase linear decrease (EILD) backoff algorithms. A one-dimensional Markov chain model is constructed for each algorithm and used to compare delay under saturation conditions. Our approach simplifies previous analyses of the BEB algorithm which used a 2-dimensional Markov chain model, while the other backoff algorithm delays have not been analyzed before. Additionally, improvements were being made to the previous analysis of BEB to achieve more accurate results. Analytical results are compared to those obtained from the simulation.**

I. INTRODUCTION

One of the critical components in 802.11 WLAN is its binary exponential backoff algorithm (BEB), which aims to efficiently utilize the medium while still reacts quickly to signs of network congestion. This is accomplished by doubling the channel arbitration time (backoff period) every time a station experiences collision or corruption. The channel arbitration time decreases rapidly to its minimum value after a successful transmission. At first glance, this aggressive reduction in the backoff period can lead to performance degradation if there are a large number of stations competing for the resource, since it encourages more collisions after any successful transmission. As a result, there are two recent proposals that suggest a slower reduction in a backoff period after a successful transmission. The first algorithm is called exponential increase exponential decrease (EIED) [2], which doubles the backoff period after a collision and halves it after a successful transmission. The second algorithm is called exponential increase linear decrease (EILD), which is based on MILD originally proposed in [3]. EILD still doubles the backoff period after collision, but linearly decreases the backoff period after a successful transmission.

A substantial work on IEEE 802.11 MAC performance exists to date. Here we focus only on the closely related work. Saturation throughput was first analyzed in [5] and subsequently refined in [1]. In this paper we simplify the approach developed in [1] [5] without relaxing any assumptions. We also extend the analysis to EIED and EILD. EIED was proposed as an enhancement to BEB and analyzed using a 2-dimensional Markov chain model in [7], but their Markov chain does not accurately model proposed changes (e.g., after a successful transmission while in state $W_{k,i}$ with $k>m'$, their Markov chain can still reside in a state with the maximum backoff window CW_{max}). Our analytical model of EIED shows better agreement with the simulations. Average delay of the BEB algorithm in saturation has been analyzed in [8]. We use a similar approach but modify the delay portion due to station's transmission to obtain exact numbers. We also calculate the average of the EIED and EILD that was not previously available. Other relevant work includes [6], which presents an alternative analysis of BEB based on mean value approach.

II. MARKOV MODELS OF BACKOFF ALGORITHMS

The IEEE 802.11 MAC protocol is based on a carrier sensing multiple access with collision avoidance (CSMA/CA) technique. At the end of its own transmission the station picks a random deference time (uniformly distributed between [0, W_i-1]) and begins counting idle slots. Index i defines the retransmission stage of the current packet. If the whole slot was observed to be idle the counter is decremented. When the medium is sensed busy the countdown is suspended until the end of the DIFS period when the counter is decremented by one. When the counter reaches zero the station begins transmission. All three backoff schemes are based on the CSMA/CA, but differ in the way they calculate the counter value W_i. These differences will be explained along with their Markov models.

A counter slot is defined as a time period at the end of which the backoff counter changes value. While the medium is idle the backoff counter is decremented by 1 every T_{slot}. When the medium becomes busy the subsequent transmission period is considered a counter slot at the end of which the backoff counter is decremented by 1 or, if the station transmitted, it is incremented by a random amount based on the backoff state and the outcome of the transmission. With this definition we can embed a Markov chain (MC) representing the state of the counter at the end of each transmission period. The state of the MC is defined as the maximum value of the deference interval for the current transmission period. Note that the number of counter slots as we defined them above per transmission cycle is Y+1 (counter counts down to zero), where Y is the random window size. We will use this fact to calculate the probability of a station starting transmission in a particular slot. The present scheme defined in the standard is the Binary Exponential Backoff (BEB) shown in Fig. 1, where the state W_i implies the maximum deference time of 2^iW-1 slots, and where W is the minimum window size set at $W = W_0+1 = 32$.

The probability of collision is denoted by p following the notation in [1]. We assume that p is constant, regardless of the current backoff state W_i, which allows us to collapse a 2-dimensional Markov chain [1] into a 1-dimensional model.

The implicit assumption in this section (as in [1]) is that there is no limit on the number of retransmissions and thus the MC will stay in state W_m until eventual success. By assuming the backoff counter stays in state W_m after a packet is dropped due to finite retransmission limit would make the model in Fig. 1 valid for the finite retransmission case as well. MC with finite retransmissions will also be analyzed in Section III.

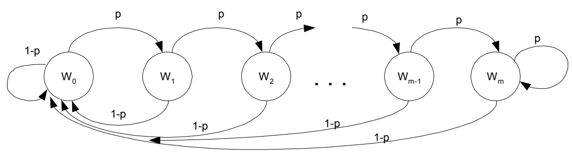

Figure 1. MC model of BEB.

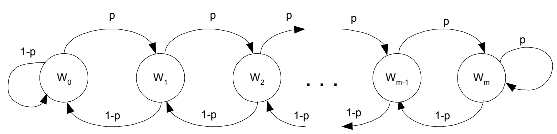

Figure 2. MC model of EIED with (r_i=2, r_d=2).

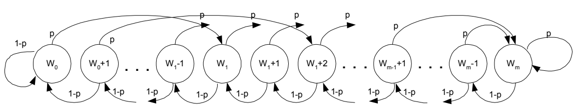

Figure 3. MC model of EILD with 2^mW-W+1 states.

Several modifications have been proposed for the baseline BEB scheme in 802.11. Recognizing that returning immediately from state W_m to state W_0 is too aggressive and can lead to increased collision probability the approach was to reduce the window more slowly after successful transmissions. One such approach is to both reduce and increase the window size exponentially, i.e., exponential increase exponential decrease (EIED) scheme. This approach is shown in Fig. 2 for equal increase and decrease rates (r_i=2, r_d=2). Generalization of this approach was proposed in [2] where the exponential increase and decrease are not at the same rate. We presented a 1-dimentionsal MC analysis of EIED where the rate of exponential increase is twice larger than the rate of exponential decrease (r_i=4, r_d=2) in [9]. Other approaches such as exponential increase linear decrease (EILD or MILD – Multiplicative ILD) were also proposed [3]. The MC model for EILD or MILD is shown in Fig. 3. Due to linear decrease after success the number of states is large (2^mW-W+1).

In Figs. 2-3 we assumed that a new packet transmission after a success can start in any backoff state excluding state W_m. This is desired so that memory of congestion is not erased immediately after a successful transmission. Limited retransmission attempts can be modeled using Figs. 2-3 if a new packet transmission, after the previous packet reached retransmission limit and was dropped, starts in backoff state W_m. The difference between the model in Fig. 2 and the one proposed in [7] for the same scheme is that in [7] it is possible to stay in state W_m and use maximum backoff window even after a successful transmission. We argue that, contrary to the results in [7], the operation of a backoff scheme that maintains maximum window size after a packet drop should not depend on the retransmission limit, as long as the retransmission limit is greater than the index of the maximum backoff stage (m).

III. Markov Chain Analysis

In this section we solve for the steady-state probabilities of the MC's illustrated in Figs. 1-3. Note that BEB was solved in [1] using a 2-dimension MC approach. Our approach simplifies the calculation for this case. Other schemes have not been investigated analytically.

A. BEB with unlimited retransmissions

Let the steady-state probability of MC being in state Wi be $P[W_i]=q_i$. We can write the following set of equation

$$q_i = \begin{cases} p \cdot q_{i-1}, & 1 \le i < m, \\ p \cdot \left(q_{i-1} + q_i\right) & i = m. \end{cases} \tag{1}$$

Solving for q_0 using normalization condition we obtain the steady state probabilities for all states as follows

$$q_i = \begin{cases} p^i \cdot (1-p), & 0 \le i < m, \\ p^m & i = m. \end{cases} \tag{2}$$

The average number of slots E[Z] spent in each state between transitions, averaged over all states, is given by 1 + E[backoff interval], i.e.,

$$\begin{aligned} E[Z] &= 1 + \sum_{i=0}^{m} \frac{2^i W - 1}{2} q_i \\ &= \frac{1 + W - p(W+2) - 2^m p^{m+1} W}{2(1-2p)} \end{aligned} \tag{3}$$

Since only one slot is used for transmission between state transitions the probability that a backlogged station transmits in a random slot is given by

$$\tau = \frac{1}{E[Z]} = \frac{2}{(W+1) + \dfrac{pW(1-(2p)^m)}{1-2p}} \tag{4}$$

This expression is the same as the expression derived in [1]. The rest of the analysis follows the same approach as in [1]. Before we present delay expressions we continue with the other cases.

B. BEB with limited retransmissions

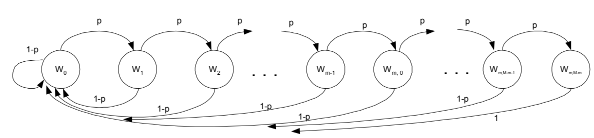

Figure 4. MC model of BEB with finite retransmission limit.

$$q_i = \frac{p^i \cdot (1-p)}{1 - p^{M+1}}, \qquad 0 \le i \le M. \tag{5}$$

$$\tau = \frac{2(1-p^{M+1})}{(W+1) + \frac{pW(1-(2p)^m)}{1-2p} - p^{M+1}(1+2^m W)} \tag{6}$$

C. EIED with (r_i=2, r_d=2)

$$q_i = \frac{r^i(1-r)}{1-r^{m+1}}, \qquad r = \frac{p}{1-p},\ 0 \le i \le m. \tag{7}$$

$$\tau = \frac{2\left((1-p)^{m+1} - p^{m+1}\right)}{\left((1-p)^{m+1} - p^{m+1}\right) + \frac{(1-2p)}{(1-3p)}\left((1-p)^{m+1} - (2p)^{m+1}\right)W} \tag{8}$$

D. EILD

$$q_i = \begin{cases} \dfrac{p}{(1-p)^i} q_0, & 1 \le i \le W+1, \\ \dfrac{1}{1-p}(q_{i-1} - (i \bmod_2) \cdot p \cdot q_{\left\lfloor \frac{i-1}{2} \right\rfloor}), & W+1 < i. \end{cases} \tag{9}$$

$$\tau = \frac{1}{E[Z]} = \frac{1}{1 + \sum_{i=0}^{2^m W - W} \frac{i+W-1}{2} q_i} \tag{10}$$

E. Saturation Throughput Analysis

The subsequent analysis follows closely [1]. We have so far expressed the probability that an individual station transmits in a random slot as a function of the overall collision probability p. The two quantities are related through the following equation [1]

$$p = 1 - \left(1 - \tau(p)\right)^{n-1} \tag{11}$$

The above non-linear equation can be solved for p, obtaining p^* and consequently $\tau^* = \tau(p^*)$. The throughput analysis of each backoff algorithm is presented in [9].

IV. Delay Analysis

In this section we show how the 1-dimensional model that was presented in this paper can be used to calculate average delay for BEB, EIED and EILD scheme. We investigate 3 cases: (1) Infinite retransmission limit, (2) Finite retransmission limit with backoff state in W_m after packet drop, (3) Finite retransmission limit with transition to W_0 after packet drop. The first 2 versions are modeled by the MC shown in Figs. 1-3. Version 3 is modeled by the MC in Fig. 4. To calculate average delay we use the approach proposed in [8] for the 2-dimensional model and apply it to our 1-dimensional model. Our analysis provides a more accurate result than [8] since we include an exact expression for the average waiting time each station spends transmitting its own packets. Previous work [8] assumed this delay component was part of the backoff delay. Finally, we offer a simple expression for calculating average delay with infinite retransmissions.

Delay in the case of a saturated network is defined as the waiting time of a packet from the time it enters the head-of-line of the queue until an ACK following successful transmission is received. The average delay is conditioned on the event that the packet is not dropped due to excessive number of retransmissions. This is not an issues when $M \to \infty$. The time between two backoff counter decrements, referred to as the counter slot, depends on the channel activity. If the channel was idle, the counter slot is equal to the slot length (T_{slot}). If the channel was busy the counter slot length is either T_s or T_c (time for successful transmission and collision [9]) depending on whether the transmission results in a success or collision, respectively. The counter slot length is also weighted by probability of success P_S [9]. Given the probability of each station transmission, τ, the probability of at least one stations transmitting in a given slot P_{tr} is

$$P_{tr} = 1 - \left(1 - \tau(p)\right)^n \tag{12}$$

The average length of a counter slot is thus given by

$$E[\text{counter slot}] = (1 - P_{tr})T_{slot} + P_{tr}P_S T_S + P_{tr}(1 - P_S)T_C \tag{13}$$

While in backoff state Wi, the average number of slots spent waiting for the next transmission opportunity is $(2^i W-1)/2$.

Case 1: Infinite retransmission limit case

Since every new transmission for BEB starts in backoff state W_0 the probability of a packet succeeding in stage W_j is given by $p^j(1-p)$. Given that the packet is successfully transmitted in stage W_j its average delay should have three components: delay from transmitting j times unsuccessfully, delay from (j+1) backoff periods before all transmission attempts and, finally, the delay for the successful packet transmission. The average delay until the packet is successfully transmitted and an ACK is received, including the times when the station transmitted unsuccessfully, is given by

$$\begin{aligned} E_{BEB}[\text{D}] &= \sum_{j=0}^{\infty} E[D \mid j \text{ retransmission attempts}] \cdot p^j \cdot (1-p) \\ &= \sum_{j=0}^{\infty}\left(\left(j \cdot T_C + T_S + E[\text{counter slot}] \cdot \sum_{i=0}^{j} \frac{2^{\min(i,m)}W - 1}{2}\right) \cdot p^j \cdot (1-p)\right) \\ &= \frac{p}{1-p}T_C + T_S + E[\text{counter slot}] \cdot \sum_{j=0}^{\infty}\sum_{i=0}^{j} \frac{2^{\min(i,m)}W-1}{2} \cdot p^j \cdot (1-p) \end{aligned} \tag{14}$$

For EIED scheme with infinite retransmissions we need to condition on the initial backoff state for a new transmission. Since a successful transmission moves the MC to the lower state, the state W_m cannot be the initial backoff state for a new transmission. Let I_k be an impulse function equal to unity for k=0 and 0 otherwise. The probability of a new transmission starting in state W_k is given by

$$P[\text{new tx starts in } W_k] = \frac{(1-p) \cdot (I_k q_k + q_{k+1})}{(1-p)}, \quad 0 \le k < m. \tag{15}$$

The term in the numerator represents the probability of entering state W_k after success, while the denominator (1-p) is the normalization constant. Steady-state probabilities q_k are derived in (7). Given the initial state for the new transmission W_k and the number of retransmissions before the final success j it is clear how to calculate the average delay. Unconditioning with respect to these two quantities we obtain by the following expression for EIED with infinite retransmissions

$$
\begin{aligned}
&E_{EIED}[\mathrm{D}]= \\
&=\sum_{k=0}^{m-1}\left((I_k \cdot q_k + q_{k+1})\cdot \sum_{j=0}^{\infty} E[D \mid j \text{ retransmission attempts}]\cdot p^j \cdot (1-p)\right) \qquad (16) \\
&=\sum_{k=0}^{m-1}\left((I_k \cdot q_k + q_{k+1})\cdot\right. \\
&\quad\left.\cdot\left(\frac{p}{1-p}\cdot T_C + T_S + E[\text{counter slot}]\cdot \sum_{j=0}^{\infty}\sum_{i=k}^{k+j}\frac{2^{\min(i,m)}W-1}{2}\cdot p^j \cdot (1-p)\right)\right)
\end{aligned}
$$

The same approach can be used for calculating EILD delay. In that case index k takes values from [0, 2^mW-W+1]. However, knowing the average saturation throughput for all three schemes with infinite retransmissions one can calculate the average delay in a much more elegant way as follows

$$
E_X[D] = \frac{n \cdot L}{R \cdot S_X} \qquad (17)
$$

where n is the number of contending stations, L is the packet length, R is the bit-rate, and S_x is a normalized throughput [9]. Subscript X stands for either BEB, EIED or EILD scheme. It should be pointed out that (17) holds only for infinite retransmission ($M \rightarrow \infty$). We refer to the calculation in (17) as the departure-time based (DTB).

Case 2: Finite retransmission limit with backoff state in W_m after packet drop

In this case we do not allow the number of retransmissions to be larger than M, i.e., the total number of transmissions should not be bigger than M+1, and the backoff state after packet drop due to retransmission limit stays at W_m. As opposed to the previous case, the new transmission can start in state W_m. We assume that $M \geq m$ which is typically the case. For BEB we distinguish two possibilities that the new transmission starts in either state W_0 in the case of success (with probability $1-p^{M+1}$) or in state W_m in the case of collision (with probability p^{M+1}). Consistent with the previous delay definition, delay for dropped packet is not included in this calculation. The probability that the packet succeeds after j retransmissions now needs to be conditioned on the packet not being dropped ($1-p^{M+1}$). Thus, we have

$$
\begin{aligned}
&E_{BEB}[\mathrm{D}] = P[\text{previous tx successful}]\cdot \\
&\quad\cdot\sum_{j=0}^{M}\left(\left(j\cdot T_C + T_S + E[\text{counter slot}]\cdot\sum_{i=0}^{j}\frac{2^{\min(i,m)}W-1}{2}\right)\cdot P[\text{success in } W_j \mid \text{packet not dropped}]\right) \qquad (18) \\
&\quad + P[\text{previous tx unsuccessful}]\cdot \\
&\quad\cdot\sum_{j=0}^{M}\left(\left(j\cdot T_C + T_S + E[\text{counter slot}]\cdot(j+1)\cdot\frac{2^{m}W-1}{2}\right)\cdot P[\text{success in } W_j \mid \text{packet not dropped}]\right) \\
&= (1-p^{M+1})\cdot\sum_{j=0}^{M}\left(\left(j\cdot T_C + T_S + E[\text{counter slot}]\cdot\sum_{i=0}^{j}\frac{2^{\min(i,m)}W-1}{2}\right)\cdot\frac{p^j\cdot(1-p)}{1-p^{M+1}}\right) \\
&\quad + p^{M+1}\cdot\sum_{j=0}^{M}\left(\left(j\cdot T_C + T_S + E[\text{counter slot}]\cdot(j+1)\cdot\frac{2^{m}W-1}{2}\right)\cdot\frac{p^j\cdot(1-p)}{1-p^{M+1}}\right)
\end{aligned}
$$

In the case of EIED with finite retransmissions we also distinguish two cases based on the success or drop of the previous packet. Similar to (16) and (18) we have

$$
\begin{aligned}
&E_{EIED}[\mathrm{D}] = (1-p^{M+1})\cdot\sum_{k=0}^{m-1}\left((I_k\cdot q_k + q_{k+1})\cdot\sum_{j=0}^{M}E[D\mid j \text{ retransmission attempts}]\cdot\frac{p^j\cdot(1-p)}{1-p^{M+1}}\right) + \\
&\qquad + p^{M+1}\cdot\sum_{j=0}^{M}\left(\left(j\cdot T_C + T_S + E[\text{counter slot}]\cdot(j+1)\cdot\frac{2^{m}W-1}{2}\right)\cdot\frac{p^j\cdot(1-p)}{1-p^{M+1}}\right) \\
&\quad = (1-p^{M+1})\cdot\sum_{k=0}^{m-1}\left((I_k\cdot q_k + q_{k+1})\cdot\right. \\
&\qquad\left.\cdot\sum_{j=0}^{M}\left(\left(j\cdot T_C + T_S + E[\text{counter slot}]\cdot\sum_{i=k}^{k+j}\frac{2^{\min(i,m)}W-1}{2}\right)\cdot\frac{p^j\cdot(1-p)}{1-p^{M+1}}\right)\right) + \\
&\qquad + p^{M+1}\cdot\sum_{j=0}^{M}\left(\left(j\cdot T_C + T_S + E[\text{counter slot}]\cdot(j+1)\cdot\frac{2^{m}W-1}{2}\right)\cdot\frac{p^j\cdot(1-p)}{1-p^{M+1}}\right)
\end{aligned}
\qquad (19)
$$

As in the previous case EILD analysis follows the approach as in (19).

Case 3: Finite retransmission limit with transition to W_0 after packet drop

Here we present results for BEB only and use MC in Fig. 6. In this case each new packet transmission starts in the state W_0. Given that a transmission succeeds in stage j and that the packet is not dropped, the time the packet waits until the final ACK is received is the sum of a three components: sum of backoff times in each stage, sum of the transmission time when the packet was unsuccessful and finally the transmission time including ACK. Thus we have

$$
E_{BEB}[\mathrm{D}] = \sum_{j=0}^{M}\left(\left(j\cdot T_C + T_S + E[\text{counter slot}]\cdot\sum_{i=0}^{j}\frac{2^{\min(i,m)}W-1}{2}\right)\cdot\frac{p^j\cdot(1-p)}{1-p^{M+1}}\right) \qquad (20)
$$

V. NUMERICAL RESULTS

In Fig. 5 we show average delay as a function of the number of backlogged stations for the BEB algorithm. We illustrate all 3 cases presented in Section IV. With retransmission limit at M=6 [4] there is almost no difference between staying in state W_m after packet is dropped or transitioning to state W_0, because the packet dropping probability is small. Additionally we also validate that both approaches for calculating delay for case 1 yield the same result, namely direct calculating and Departure time based (DTB, equation (17)).

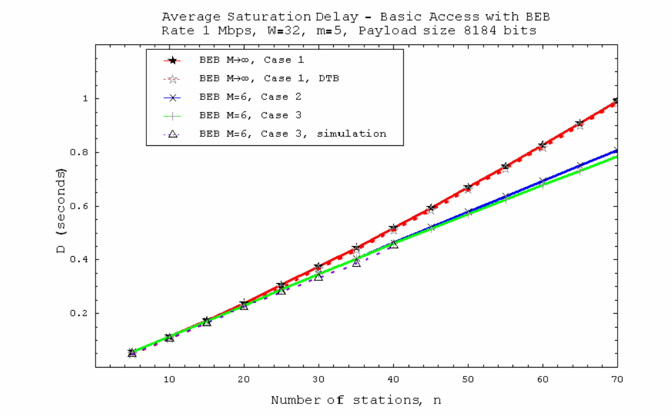

Figure 5. Average delay for various versions of BEB.

In Fig. 6 we compare delay performance for the three backoff schemes. First thing to notice is that for EIED and EILD the difference between $M \rightarrow \infty$ and M=6 is negligible compared to BEB. EILD scheme exhibits the best delay performance. It should also be noted that EIED performs slightly worse than BEB when M=6.

The simulation results show that all three backoff schemes take an approximately 1 second to reach steady state, regardless of loading. After reaching steady state, the delay variance is significantly smaller for EILD than the other two schemes, because EILD is the least likely to be impacted by collision. In term of fariness, Fig. 7 shows W_i of each EILD user when n=10. Although most users reach steady state in a few seconds, two users take over 15 seconds to reach the steady state. During that time, those two users have an unfair advantage over other users because a smaller W_i allows them to transmit a packet with lower delay. This unfair advantage is reduced when the number of station increases, because it is more likely to collide. For BEB and EIED, there is no per-user steady state, because W_i reverts back to W_0 at some point for each user. As a result, existing user will occasionally use W_0, just like a newly activated user.

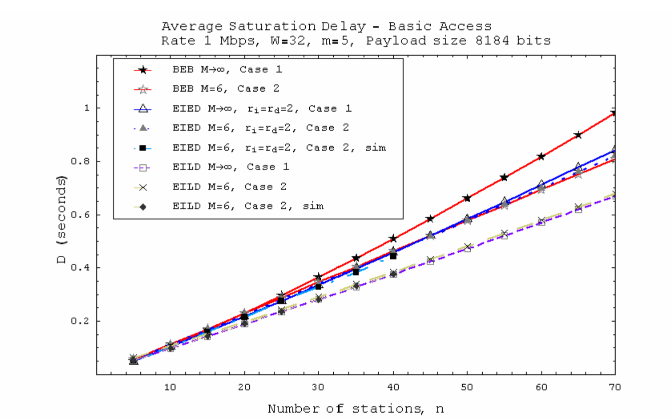

Figure 6. Comparison between BEB, EIED and EILD.

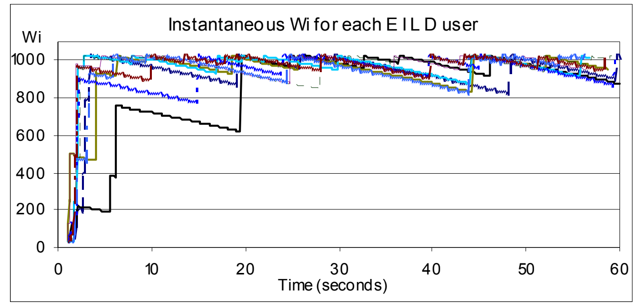

Figure 7. Instantenous Wi for each EILD user (n=10).

VI. Conclusion

We presented a new analytical model to compute average delay of the BEB, EIED and EILD backoff algorithms for IEEE 802.11 wireless LAN. A one-dimensional Markov chain model is constructed for each algorithm and used to compare delay under overload conditions. Our analysis improves the accuracy of the previous work and extends it to the proposed backoff algorithms EIED and EILD that were not analyzed previously. We show that EILD has the best delay performance. Additionally, the difference in average delay between infinite and finite retransmissions is minimal for EIED and EILD.

VII. References

[1] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function", *IEEE Journal on Selected Areas in Communications*, pp. 535–547, Vol. 18, no. 3, March 2000.

[2] N. Song, B. Kwak, J. Song, and L. E. Miller, "Enhancement of IEEE 802.11 distributed coordination function with exponential increase exponential decrease backoff algorithm," *The 57th IEEE Semiannual Spring Vehicular Technology Conference*, Vol. 4 , pp. 2775 –2778, April 2003.

[3] V. Bharghavan, A. Demers, S. Shenker, and L. Zhang, "MACAW: A Media Access Protocol for Wireless LAN's," *Proceeding ACK SIGCOMM'94*, London, England, 1994, pp. 212-225.

[4] ANSI/IEEE Std 802.11: "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications", 1999.

[5] G. Bianchi, "IEEE 802.11 – Saturation Throughput Analysis", *IEEE Communications Letters*, pp. 318-320, Vol. 2, no. 12, December 1998.

[6] Y.C. Tay and K.C. Chua, "A Capacity Analysis for the IEEE 802.11 MAC Protocol", *ACM Wireless Networks*, Kluwer Academic Publishers, pp. 159-171, no. 7, 2001.

[7] H. Wu, S. Cheng, Y. Peng, K. Long, J. Ma," IEEE 802.11 Distributed Coordination Function (DCF): Analysis and Enhancement" *IEEE ICC 2002*, May 2002.

[8] P. Chatzimisios, A.C. Boucouvalas, and V. Vitsas, "IEEE 802.11 Packet Delay – A Finite Retry Limit Analysis," IEEE Globecom 2003, Dec 2003.

[9] I. Vukovic and N. Smavatkul, "Saturation throughput analysis of different backoff algorithms in IEEE 802.11," Accepted for publication in IEEE PIMRC 2004, Barcelona, Spain, September 2004.