Access Delay and Throughput Evaluation of Block ACK under 802.11 WLAN

Yean-Fu Wen, Frank Yeong-Sung Lin, Kun-Wei Lai
{d89002; yslin; r91006}@im.ntu.edu.tw
Tel)+886-2-2363-0231#2812, FAX)+886-2-236-1327
Department of Information Management, National Taiwan University (NTU)

Abstract-**The transmission bit rate on wireless media has increased, and will continue to increase, rapidly. However, the high overhead of 802.11 MAC (Media Access Control) protocol degrades the theoretical bandwidth. We therefore propose a block ACK approach to reduce the overhead and increase system capacity. Using block ACK to increase packet size overcomes the constraint of co-existing with Ethernet's maximum packet size. Our approach includes block ACK for the basic mode, the RTS/CTS mode, and IABA (Initial ACK then Block ACK). To understand delay and throughput effects, we propose an extended analytical model. The numerical results show the total throughput of the basic mode degrades as the number of MHs (Mobile Host) increases, but the IABA approach degrades smoothly. The RTS/CTS mechanism not only avoids hidden terminal problems, but also improves the total throughput so that it is higher than the maximum throughput of the original mechanism. Considering the delay issues, the numerical results show that the number of packets per block can not be increased infinitely.**

Key Words: Wireless LANs, Wireless Networks, DCF, Block ACK, Access delay, and Throughput.

I. INTRODUCTION

The Wireless Local Area Networks (WLANs) have been widely deployed in many public areas such as airports and hotels. It has become the network on demand for people to retrieve information anywhere. More and more applications such as multi-media service and VOIP etc. need more bandwidth. To make the most efficient use of the wireless medium, we need to take a systems' view and design the operation of the MAC so that it effectively uses and manages the services of the physical layer while providing a lower overhead to the higher layers.

The WLAN, which is similar to Ethernet and Token Ring provides sharing of the wireless media, using the same two main approaches as the MAC protocol: (1) PCF (Point Coordination Function) uses a "poll-and-response" protocol to eliminate contention among wireless stations; and (2) DCF (Distributed Coordination Function), executes a "listen before talk", p-persistence and immediate ACK strategy for all MHs [6]. In this paper, we are interested in the widely available access method, DCF which uses the CSMA/CA protocol.

Two medium access techniques are used for packet transmission in DCF: the basic access mechanism and the optional Request-to-Send/Clear-to-Send (RTS/CTS) mechanism [6]. The basic access mechanism is similar to Ethernet, but does not have collision detection. The sender treats the packet as a collision case unless an ACK is received and schedules its retransmission, according to the exponential back-off algorithm. However ACK timeout requires a long time for retransmission and therefore wastes system resources.

For the RTS/CTS access mechanism, the special RTS and CTS frames are utilized to avoid long term collision. The MH senses that the channel is free after a DCF inter-frame space (DIFS). It then sends the RTS frame, but the data frame and destination host responds with a CTS frame after a short inter-frame space (SIFS). If the MH can receive the CTS frame correctly, the data frame will be transmitted. The RTS frame is retransmitted according to the binary exponential back-off procedure if the CTS frame is not received within CTS timeout. All other MHs that receive an RTS or CTS frame stop transmitting, or pause the back-off, and re-schedule it to resume based on the Network Allocation Vector (NAV).

The remainder of this paper is organized as follows. In Section II, we briefly describe the concept of block ACK. In Section III, an analytical framework used for evaluating our proposed approach is discussed. In Section IV, the performance of block ACK is numerically analyzed with MATLAB tools. Finally, in Section V, we present our conclusions.

II. THE BLOCK ACK CONCEPT

The DCF basic mode has a low handshaking generates extensive overhead protocol overhead, but the transmissions are prone to collision. In constrast, the RTS/CTS mechanism reduces the probability of collisions of data packets, but the four way handshake generates extensive overhead. For example, about 18.2% of available data rate is consumed by inter-frame spaces (IFS) and MAC-level ACK. However, If RTS/CTS is enabled, it can approach 38.9%. Moreover, the contention mechanism of DCF exponentially increases the frame delivery time to the peer station. Frequent retransmissions also cause unpredictable access delays in the order of tens to hundreds of milliseconds (ms) and transmission of pending frames can be blocked.

We use the ns2 network simulation program to obtain the initial total throughput versus the packet size for a variable number, *n*, of MHs with 802.11b standard parameters, which lists as Table 1. The environment we consider is a single cell with an AP. Figure 1 shows each MH that intends to transmit a packet has to forward its packet to the access point (AP) first, even if it is destined for a MH located in the same cell. The communication channel is error-free and of no obstacle. Figures 2 and 3 show the maximum throughput as about 6.57 Mbps and 4.75 Mbps with saturation bandwidth which set the upper layer packet size to 1,500 bytes for basic mode and RTS/CTS mode, respectively. The results show that as the packet size increases, the total throughput slow up increases. The basic throughput of the basic mode is higher than RTS/CTS mode, but the increase rate of the RTS/CTS mode is larger than the basic mode.

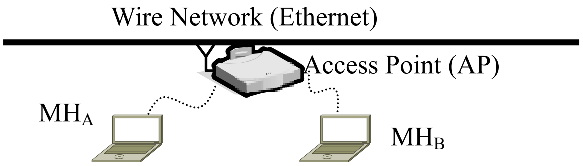


Fig. 1 Single hop wireless where each MH intends to transmit a packet has to forward its packet to the AP first.

Table 1 Parameter values used for simulation

Parameters	Values	Parameters	Values
MSDU size	1500bytes	ACK length	14bytes
MAC header	34bytes	PHY header	16bytes
RTS payload	20bytes	CTS payload	14bytes
Slot time	20µs	DIFS	50µs
SIFS	10µs	Propagation time	1µs
CW_{min}	32	CW_{max}	1023
L_{max}	2304	B_{max}	11

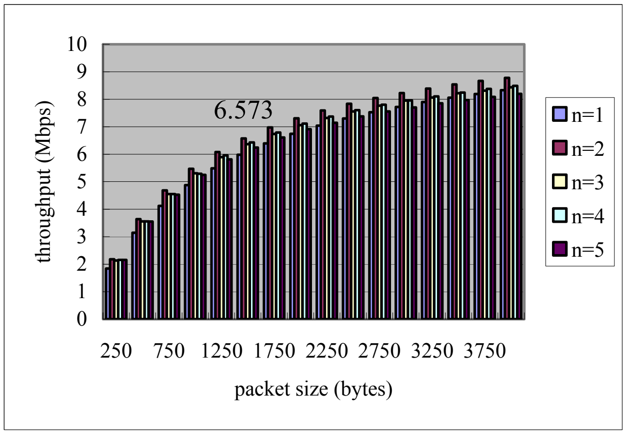


Fig. 2 Maximum throughput of basic mode using NS2

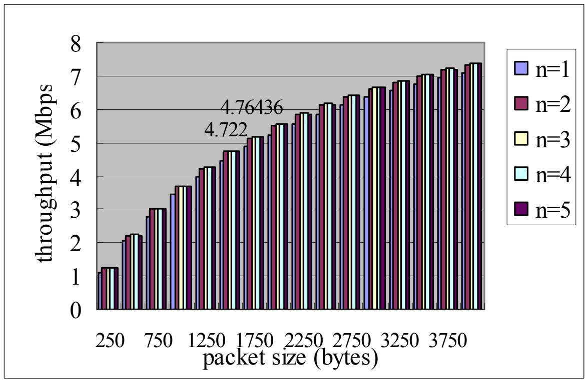


Fig. 3 Maximum throughput of RTS/CTS mode using NS2

In order to accommodate the packet size with a general limitation of 1,500bytes, we must cut the long frame into small sizes. Figures 4(a) and 4(d) show the original mode, transmitting one small size frame must be followed an ACK frame that is transmitted with the lowest bit rate. The block ACK approach shown in Fig 4(b) and 4(e) eliminates intermediate ACK with multiple back-to-back packets, but only responds with an ACK after the final frame to identify successful transmission. To avoid a long collision avoidance (CA) period under the basic mode, we also propose the IABA approach to shorten the collision period of the basic mode. Figure 4(c) shows the initial ACK sent to avoid the collision which is similar to the CTS function, and then follow-up multiple back-to-back packets are sent with only one ACK frame at the end of transmission cycle.

Therefore, with the multiple packets approach, an MH is allowed to send multiple frames consecutively by setting *more_frag* = 1 in the MAC control frame after gaining access to the medium [6]. Thus, the block ACK for the basic mode and RTS/CTS mechanisms have only one DIFS and one back-off time for multiple packets. The performance of a similar approach for 802.11b is studied in [10]. Sadeghi et al. also introduce the Opportunistic Auto Rate (OAR), an enhanced protocol for multi-rate IEEE 802.11 in wireless ad hoc networks [1]. The data flushing data transfer (DFDT) approach [10] not only concentrates on the stability of the network by using extra overhead but also considers other factors such as: contention, packet length, and packet arrival rate. These factors also affect the stability and throughput of the network. In 802.11e MAC-level ACK has become optional [5]. This means that when the "no ACK" policy is used, the MAC does not send an ACK when it has received a frame. It also means that reliability of "no ACK" traffic is reduced, but it improves the overall MAC efficiency for time-sensitive traffic, such as VoIP, where the data has a very strict lifetime.

Our approach focuses on generating data service to improve the system capacity. The RTS/CTS block ACK approach apportions the RTS/CTS frame overhead, which is added to avoid hidden terminal problems in the original mode. The IABA approach combines the basic mode and the CTS function to avoid a long collision avoidance period.

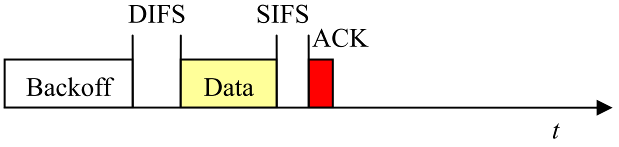
In order to understand block ACK effects, we not only evaluate throughput but also block access delay time to choose the appropriate number of packets per block.

III. ANALYTICAL MODEL

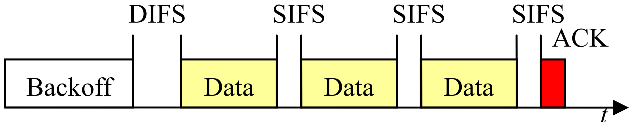
We extend Bianchi's model in [3] to limit the back-off to a finite state for block ACK with an AP. A two or four way handshaking mechanism is adopted in the MAC protocol. Table 2 lists the main notations and given modified parameters in our extended analytical model.

A. Markov Analysis

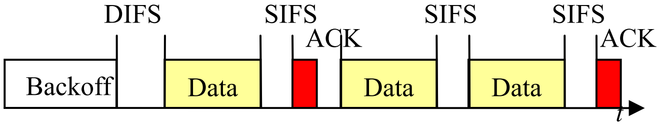
We begin by estimating the probability of a collision. Let *[t, t+1)* denote a discrete and integral time scale representing a logical time unit. Each MH decreases its back-off counter, or transmits a packet, at the beginning of each logical time unit. Let $p(t)$ denote the collision probability to transmit a packet at time t. Assume that $p(t)$ is constant and independent of time, i.e. $p(t)= p$ for all integers $t \geq 0$. Let $S(t)$ denote the back-off stage at time t, where $0 \leq S(t) \leq m+u$. Figure 5 shows the finite state of the back-off Markov chain. Its probability distribution is (1).



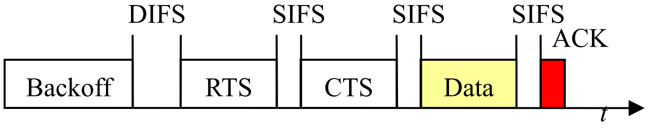
(a) Basic mode original mechanism



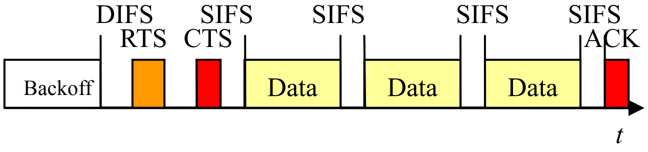
(b) Block ACK with basic mode mechanism



(c) Block ACK with IABA mechanism



(d) RTS/CTS original mechanism



(e) Block ACK with RTS/CTS mechanism

Fig. 4 The concept of block ACK

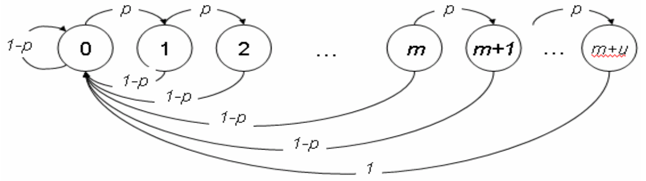


Fig 5. The finite state diagram of the back-off Markov chain

Table 2 Main notation lists

Notations	Description
p	The collision probability.
q	Packet transmission probability.
ρ	The saturation total throughput.
$E[T_I]$	The average time lengths of all idle periods.
$E[T_C]$	The average time lengths of all colliding periods.
$E[T_S]$	The average successful transmission time per cycle.
$E[T]$	The average time length of a transmission cycle.
N_c	The number of colliding transmission periods.
p_c	The collision probability.
N_s	The number of time slots contained in an idle period
D_s	The average delay of successful transmission by others.
D_c	The average collision time between two successful packets.
D_I	The total idle slot time.
n	The number of active MHs.
m	Maximum back-off stage.
u	The number of Finite back-off stages after stage *m*
R	The transmission bit rate.
W	The value of mimimum contention windows (CW_{min}.)
L	The packet size (MSDU).
T_{pro}	Propagation delay for all packets.
T_{slot}	The slot time.
T_{RTS},	The time required to transmit a RTS (include physical header and propagation delay).
T_{CTS}	The time required to transmit a CTS (include physical header and propagation delay).
T_{ACK}	The time required to transmit an ACK (include physical header and propagation delay).
T_{PHY}	The time lengths required to transmit a physical header
T_{MAC}	The time lengths required to transmit a MAC header
Given Modified Parameters	
b	The amount of multiple back-to-back packets of one block

Let B_s denote the back-off counter that the station will be chosen in back-off stage s, where $0<B_s<W-1$. Then the distribution of B_s is given in (2).

$$\Pr\{S=s\}=\begin{cases}\dfrac{1-p}{1-p^{m+u+1}} & ,s=0;\\ \dfrac{1-p}{1-p^{m+u+1}}p^{s} & ,1\leq s\leq m+u;\\ 0 & ,s>m+u.\end{cases} \tag{1}$$

$$\Pr\{B_S=i\}=\begin{cases}\dfrac{1}{2^{S}W}, \text{for } i=0,\ 1,\cdots,2^{S}W-1; 0\leq s\leq m-1\\ \dfrac{1}{2^{m}W}, \text{for } i=0,\ 1,\cdots,2^{S}W-1; m\leq s\leq m+u\end{cases} \tag{2}$$

Since the back-off counter follows a uniform distribution, the mean value of B with condition probability at state s is:

$$E[B \mid S = s] = \begin{cases} \dfrac{2^s W - 1}{2}, 0 \le s \le m-1 \\ \dfrac{2^m W - 1}{2}, m \le s \le m+u \end{cases} \tag{3}$$

Then summate all the probabilities of equation (2) and multiply them by equation (3), we get the average number of B of all back-off states by (4)

$$E[B] = \sum_{s=0}^{m+u} E[B \mid S = s] \Pr\{S = s\} \tag{4}$$

$$= \frac{(1-2p)(W-1-2^m p^{m+u+1} W + p^{m+u+1}) + pW(1-(2p)^m)}{2(1-2p)(1-p^{m+u+1})}$$

Note that Bianchi evaluated $p< 0.5$ to avoid *(1-2p)* equal to zero error.

At a steady state, the station has to wait $E[B]$ logical time units before it can transmit a packet. In other words, the probability q of the station to transmit a packet at any logical time unit is listed as (5).

$$q = \frac{1}{E[B]+1} \tag{5}$$

$$= \frac{2(1-2p)(1-p^{m+u+1})}{(1-2p)(W+1-2^m p^{m+u+1} W - p^{m+u+1}) + pW(1-(2p)^m)}$$

The probability of one or more other stations transmitting packets at the same logical time unit as the station follows the geometric distribution, so we get (6).

$$p = \left(1-(1-q)^{n-1}\right) \tag{6}$$

B. Throughput Analysis

The saturation throughput of the DCF access method has been extensively studied in recent literature [2] [3] [4] [8] [11]. Figure 6 shows the renewal and reward transmission cycle, which includes the idle period, collision period and success period. An idle period is a time interval in which the channel remains idle due to the back-off procedure. Success period T_s denotes sender has successfully received ACK. According to the IEEE 802.11 specifications [6], T_s for the basic mode, IABA approach and RTS/CTS mechanisms can be calculated as (7) to (9), respectively.

$$E[T_s^{bas}] = T_{DIFS} + b*(T_{PHY} + T_{MAC} + E[L]/R + T_{pro} + T_{SIFS}) + T_{ACK} \tag{7}$$

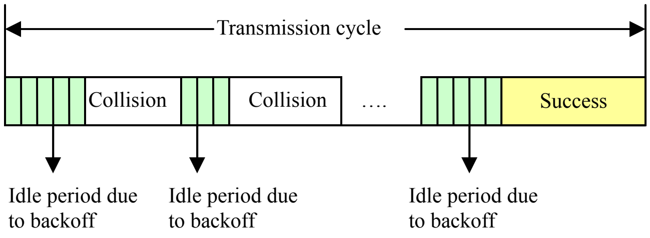


Fig. 6 Renewal and reward transmission cycle

$$E[T_s^{IABA}] = T_{DIFS} + b*(T_{PHY} + T_{MAC} + E[L]/R + T_{pro} + T_{SIFS}) + 2*T_{ACK} \tag{8}$$

$$E[T_s^{rts/cts}] = T_{DIFS} + T_{RTS} + T_{SIFS} + T_{CTS} + b*(T_{SIFS} + T_{pro} + T_{PHY} + T_{MAC} + E[L]/R + T_{pro} + T_{SIFS}) + T_{ACK} \tag{9}$$

Then the saturation bandwidth is:

$$\rho = \frac{bL}{E[T_I + T_c + T_s]} = \frac{bL}{E[T_I] + E[T_c] + E[T_s]} \tag{10}$$

Since the MH detects the collision by ACK frame, the T_c of each approach can be computed as (11) to (13) for basic mode, IABA approach and RTS/CTS mode, respectively.

$$T_c^{bas} = E[N_c](T_{DIFS} + b*(T_{PHY} + T_{MAC} + E[L]/R + T_{pro} + T_{SIFS}) + T_{ACK}) \tag{11}$$

$$T_c^{IABA} = E[N_c](T_{DIFS} + T_{PHY} + T_{MAC} + E[L]/R + T_{pro} + T_{SIFS} + T_{ACK}) \tag{12}$$

$$T_c^{rts/cts} = E[N_c](T_{DIFS} + T_{RTS} + T_{SIFS} + T_{CTS}) \tag{13}$$

The p_c can be computed as at least one transmitting station, namely p_c = Pr{the number of transmitting stations >=2 | number of transmitting stations = 1}. It is also equal to (1-Pr{number of transmitting station = 0}-Pr{number of transmitting stations=1}) / Pr{number of transmitting stations >=1}. Then the p_c is listed as (14).

$$p_c = \frac{1-(1-q)^n - nq(1-q)^{n-1}}{1-(1-q)^n} \tag{14}$$

The distribution of N_c follows the geometric distribution given its mean value as (15).

$$\Pr\{N_c = i\} = (1-p_c)p_c^i, \quad \text{for } i = 0,\ 1,\ 2, \cdots, \tag{15}$$

$$E(N_c) = \frac{p_c}{1-p_c} \tag{16}$$

We assume that the time lengths of idle periods are independently and identically distributed. T_I can be computed as (17).

$$E[T_I] = (E[N_c]+1)(\delta E[N_s]) \tag{17}$$

The distribution of N_s is:

$$\Pr\{N_s = i\} = \left(1-(1-q)^n\right)\left((1-q)^n\right)^i, \ for\ i = 0,1,2.. \tag{18}$$

Then the mean value of N_s is as follows:

$$E[N_s] = \sum_{i=0}^{\infty} i\left(1-(1-q)^n\right)\left((1-q)^n\right)^i = \left(\frac{(1-q)^n}{1-(1-q)^n}\right) \tag{19}$$

C. Average packet delay

For the delay model, we refer to Wang's model [4] to support multiple back-to-back packets environments. Let D be the access delay described above paragraph. To compute D, we divide it into four parts. (1) T_s, successful transmission time which defined as above paragraph; (2) D_s, the average time of successful transmission by other stations; (3) D_c, the average collision time between two successful blocks; and (4) D_I, the total idle slot time, which includes the total back-off time and backlog time of each MH. So we get:

$$D = T_S + D_S + D_C + D_I \tag{20}$$

Equations (7) to (9) have already defined the T_s for the basic mode, IABA approach and RTS/CTS mode, respectively. We also define the T_c by equations (11) to (13). During the interval between two continuous successful transmissions in a station, the time for a successful transmission in each other station is $T_s *N$, where N is the number of successful transmissions by other stations. Heusse et al. have shown that the long term channel access probability is equal for common situations in a wireless environment [7]. Therefore during the interval of two continuous successful transmissions in an MH, every other MH must also have a successful transmission respectively. Thus, we obtain

$$D_S = \begin{cases} T_S^{bas}*(n-1) & \text{,for basic mode;} \\ T_S^{IABA}*(n-1) & \text{,for IABA appraoch;} \\ T_S^{rts}*(n-1) & \text{,for RTS/CTS mode;} \end{cases} \tag{21}$$

Considering the whole network, there are $E[N_c]$ continuous collisions in the time between two random continuous, successful transmissions. According to the analysis above, there are N successful transmissions in the period of time of D. Therefore, we have:

$$D_c = \begin{cases} n*E[N_c]*T_c^{bas} & \text{,for basic mode;} \\ n*E[N_c]*T_c^{IABA} & \text{,for IABA appraoch;} \\ n*E[N_c]*T_c^{rts} & \text{,for RTS/CTS mode;} \end{cases} \tag{22}$$

According to the above analysis, there is a back-off interval before each successful transmission or each collision, and there are N successful transmissions and N_c collisions in the period of time of D according to the analysis above, so the total time between two successful transmissions of idle slots is:

$$D_I = (E[N_c] + n) * E[N_s] * T_{slot} \tag{23}$$

IV. ACCESS DELAY AND THROUGHPUT EVALUATION

The system total throughput is evaluated by saturation bandwidth which means MH always have packet to send. To find the maximum throughput, we assume the bandwidth of all MHs is the highest transmission rate, i.e. 11Mbps for 802.11b. The main parameter values are listed as Table 1.

Figure 7 compares the total throughput of both basic DCF and variable amount frames versus the number of MHs n, with and without the RTS/CTS mechanism using NS2 simulation. It can be seen that, for both protocols, if the RTS/CTS mechanism is not used, the throughput degradation is dramatic, even without considering the hidden terminal problems. It becomes more critical as the number of frames required to form one block increases. The adaptive block ACK only outperforms the basic protocol when there are a few MHs within a cell. On the other hand, the total throughput remains high with the RTS/CTS mode; the total throughput is more efficient than the basic mode as the number of packets is equal to 2.

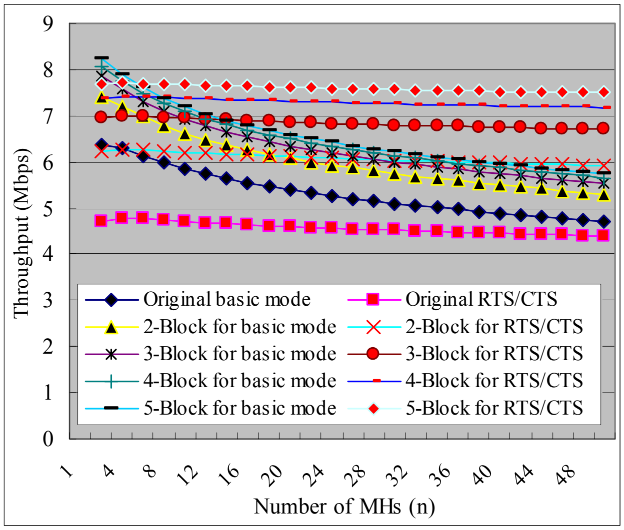
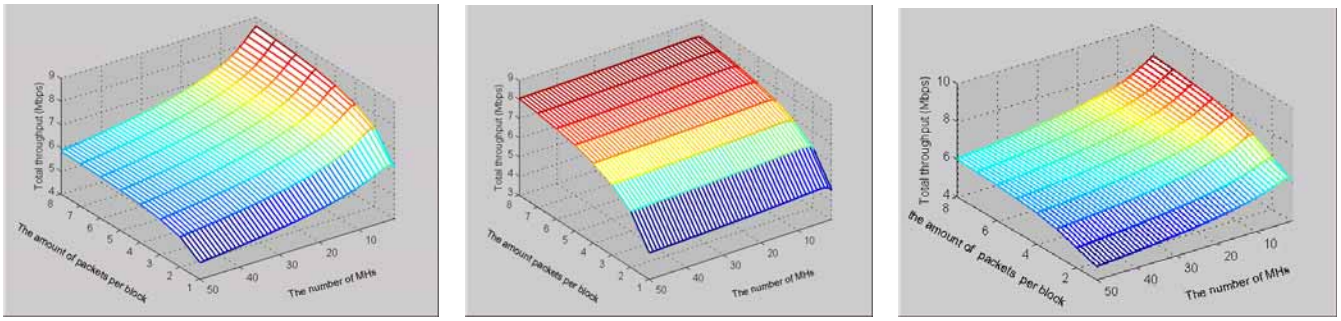


Fig 7 basic mode block ACK compare with RTS/CTS mode

When the RTS/CTS mechanism is considered, performances are good even with many MHs. Moreover, block ACK not only compensates for the RTS/CTS frames overhead, but also has the advantage of reducing collision time. In order to avoid the hidden-terminal problem, the RTS/CTS is able to transmit a large packet. To replace a data frame with an RTS/CTS frame degrades collision time T_c.

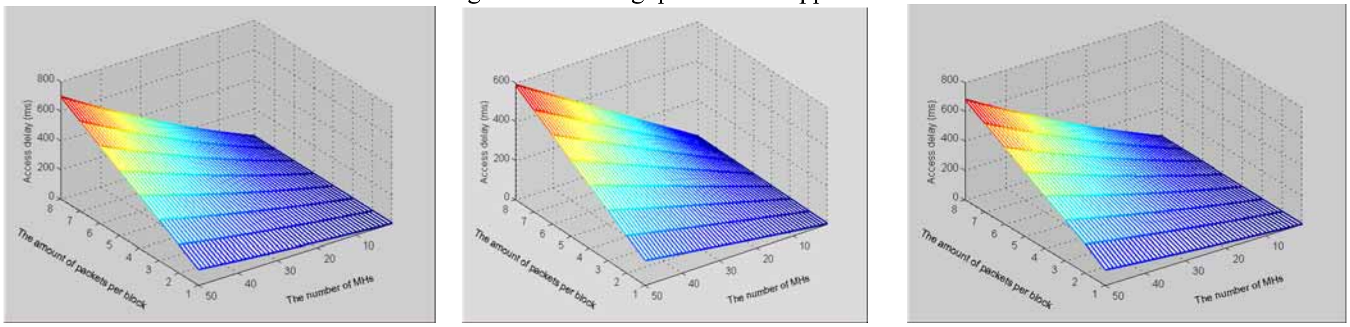
Figures 8(a), 8(b), and 8(c) show the respective results of the basic mode, IABA approach and RTS/CTS mode for the number of MHs versus the number of packets in a block using MATLAB tools. The total throughput of all of three approaches increases as the number of packets in a block increases. Due to the low collision period of the RTS/CTS mode, its total throughput not only maintains the maximum theoretical throughput, but also slows up the rate of increase as the number of packets in a block increases. Although the total throughput of the basic mode degrades rapidly, the IABA approach shortens the collision detection period so that its total throughput does not degrade as fast as the basic mode.

Because multiple back-to-back packets hold the channel time much than the original mode, we evaluate average block delay versus the number of packets and number of MHs to determine the delay bound. Figures 9(a), 9(b), and 9(c) show the average block delay increases as the number of packet per block increases using MATLAB. The access delay in block ACK for the basic mode increases rapidly, which also causes its throughput to degrade rapidly. Due to the short collision period of the RTS/CTS mode, its access delay does not increase rapidly, but the delay monotonically increases as the number of packets increases. This means that the number of packets per block needs to be limited.



(a) basic mode with block ACK (b) RTS/CTS mode with block ACK (c) the IABA approach

Fig. 8 Total throughput of three appraoches



(a) basic mode with block ACK (b) RTS/CTS mode with block ACK (c) the IABA approach

Fig. 9 Access delay of three approaches

V. CONCLUSIONS

In this paper, we address block ACK with multiple back-to-back packets without immediate ACK to improve the throughput. As expected, the performance of the basic mode only improves when the traffic is light. When the number of MHs increases, its performance degrades rapidly. Conversely, the IABA approach can improve the total throughput by avoiding long block ACK. The best performance of block ACK is the RTS/CTS mode which compensates the RTS/CTS overhead to avoid hidden problems. Its total throughput increases as the number of packets increases, but the access delay time increases slowly. Ultimately, the numerical results show the delay monotonically increases as the number of packets increases; the number of packets in block needs to be limited. The hidden terminal problem will be the subject of our future works.

REFERENCES

[1] B. Sadeghi, V. Kanodia, A. Sabharwal, and E. Knighlty, Opportunistic media access for multirate ad hoc networks, in Proceedings of ACM MobiCom'02, Atlanta, Giorgia, September 2002.

[2] G. Bianchi, IEEE 802.11: saturation throughput analysis, *IEEE Communications Letters 2(12)*, 1998, 318-320.

[3] G. Bianchi, Performance analysis of the IEEE 802.11 distributed coordination function, *Selected Areas in Communications, IEEE Journal on, 18(3),* March 2000, 535-547.

[4] G.Wang, Y.Shu, L.Zhang, O.W.W.Yang, Delay analysis of the IEEE 802.11 DCF, Proc. IEEE PIMRC, Beijing, China, 2003.

[5] IEEE 802.11e/D3.0, Draft Supplement to Part 11: Wireless Medium Access Control (MAC) and physical layer (PHY) specifications: Medium Access Control (MAC) Enhancements for Quality of Service (QoS), May 2002.

[6]"IEEE 802.11b Standard"; standards.ieee.org/getieee802 /download /802.11b-1999.pdf.

[7] M. Heusse; F. Rousseau, G. Berger-Sabbatel, A.Duda, Performance anomaly of 802.11b, INFOCOM 2003. Twenty-Second Annual Joint Conferences of the IEEE Computer and Communications Societies. IEEE, 2(30), April 2003, 836-843.

[8] N. Gupta and SR Das. A capacity and utilization study of mobile ad hoc networks. In 26th Annual IEEE Conference on Local Computer Networks, LCN 2001, November 2001, 576-583.

[9] P.Chatzimisios, A.C.Boucouvalas, V.Vitsas, IEEE 802.11 packet delay-a finite retry limit analysis, Proc. IEEE GLOBECOM, San Francisco, CA, December 2003.

[10]ST.Sheu, T.Chen, J.Chen, F.Yu, An improved data flushing MAC protocol for IEEE 802.11 wireless ad hoc network, Vehicular Technology Conference, 2002. Proceedings, IEEE 56th, Volume: 4, 2002, 2435-2439, IEEE Semiannual Vehicular Technology Conference, 24-29.

[11]X. Dong, M. Ergen, P. Varaiya, A. Puri, Improving the Aggregate Throughput of Access Points in IEEE 802.11 Wireless LANs", IEEE WLN, Bonn, Germany, October 2003.