PERFORMANCE ANALYSIS OF PRACTICAL COMBINATORIAL MECHANISM WITH EXPONENTIAL BACKOFF AND RETRANSMISSION CUTOFF

Ying Weng, Chaohuan Hou

Department of Integrated Digital System, Institute of Acoustics, Chinese Academy of Sciences
yingweng@sohu.com, hch@mail.ioa.ac.cn

ABSTRACT

This paper investigates the combinatorial mechanism of a general case of exponential backoff (EB) with factor *r* and retransmission cutoff with traffic sources consisting of infinite number of stations in ideal channel conditions. A new, simple but more exact infinite-station bi-dimensional discrete-time analytical model based on Markov chain, to reflect the actual behavior of this combinatorial mechanism, is presented. By means of the proposed model, we provide an extensive performance evaluation, and new analytical results are given. We also obtain the analytical expressions for the maximum saturation throughput, which converges to a non-zero constant as the station number goes to infinity. The accuracy of the analysis is verified by elaborate simulation results. In addition, the packet rejection rate is discussed for various values of system offered load and initial minimum contention window.

1. INTRODUCTION

Medium access control (MAC) protocols are the key part for wireless local area networks (WLANs). Among them, exponential backoff (EB) scheme and retransmission cutoff scheme are of practical interest from the viewpoint of feasibility to implement. Since they require neither to track feedback information nor to estimate the number of backlogged packets.

In existing networks, retransmissions of a packet will be cut off and also a packet will be discarded after a certain number of unsuccessful transmissions in order to avoid excessive collisions among retransmitted packets and to satisfy delay constraints associated with a packet. The value of retransmission cutoff should be determined according to quality of service to be guaranteed. In [1, 5], it has been assumed that packets which fail in correct reception have to be retransmitted until correctly received, whereas the packet retry limit is not be considered. In this paper, we restrict the number of retransmission trials to a finite value and permit packet droppings.

Most of papers, such as [4, 6], study EB and/or retransmission cutoff in the context of a specific network MAC protocol; however, the characteristics of those protocols seem to have as much or more effect on network performance than the intrinsic behavior of EB and retransmission cutoff. Here we focus on combinatorial mechanism of EB and retransmission cutoff itself.

Performance evaluation of combinatorial mechanism with all its details and under realistic traffic conditions has been considered difficult. Therefore, many papers have assumed simpler traffic conditions and/or operations; furthermore, simplified and/or modified models are often used to make analyses more tractable. As a result, the scope of such performance analyses is somewhat limited. Because in a WLAN, the number of active stations is generally quite large, we assume that traffic sources consist of an infinite number of stations. This hypothesis approximates a large population in which each station generates and transmits infrequently. In this paper, we succeed in proposing a new, simple but more exact model to reflect the actual behavior of this combinatorial mechanism, moreover show that the Markov analysis works well. New analytical results and expressions are given after we provide an extensive performance evaluation. As proven by comparison with simulations, the results are extremely accurate and practically exact.

This paper is organized as follows. Section 2 describes the practical combinatorial mechanism of EB and retransmission cutoff. A Markov model is presented in Section 3. We obtain analytical results for performance evaluation in Section 4, and verify them by simulations in Section 5. Conclusions are drawn in Section 6.

2. COMBINATORIAL MECHANISM

We consider a transmission medium shared by infinite *n* stations, which are fed by a certain arrival process and try to access the channel according to combinatorial mechanism: general case of EB with factor *r* and retransmission cutoff. Time is subdivided into slots of equal length, and all packets are assumed to be of the same duration, equal to the slot time. Furthermore, all stations are synchronized so that every transmission starts at the beginning of a slot and ends before the next slot on an ideal channel with no transmission errors.

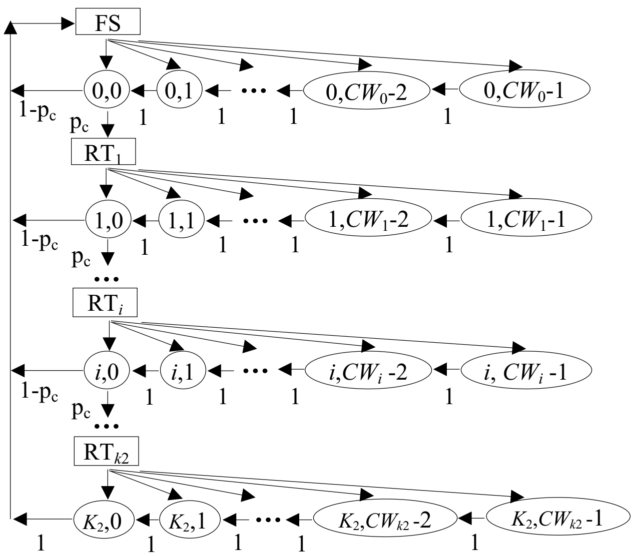

Fig.1 Markov chain model for combinatorial mechanism

At the beginning of a slot, each station is in one of two states: fresh (FS) state (state 0), with packet generated newly; or retransmission i (RT_i) state, with packet experienced i unsuccessful transmissions. Each station in FS state generates and transmits a new packet in a time slot. While each station in RT state is blocked in the sense that it cannot generate a new packet. RT state is further subdivided into k_2 state: 1, 2, … , k_2. Stations that have succeeded in the (re)transmission return to FS state. On the contrary, state transitions of unsuccessful stations depend on the previous state. Unsuccessful stations in FS state enter RT_1 state and those in RT_i state enter RT_{i+1} state (i=1, 2, … , k_2-1). A packet dropping occurs when a station in RT_{k2} state fails its retransmission, so that it moves back to FS state.

At each packet transmission, a random backoff time is uniformly chosen in the interval [0, CW-1] defined as contention window. Its value depends on the number of transmissions failed for the packet. For the packet's first transmission, $CW_{\min}$ is set to be CW. After each collision, CW is multiplied by r until it reaches $CW_{\max}$. Let $T(t)$ be the stochastic process representing the backoff timer for a given station at slot time t. The backoff timer is decremented as long as the channel is sensed idle. It is "frozen" when a transmission is detected on the channel, and reactivated when the channel is sensed idle again. The station transmits when the backoff timer reaches zero.

3. MARKOV MODEL

The behavior of a single station in this combinatorial mechanism is studied with an infinite-station bi-dimensional discrete-time Markov chain model.

Here $i \in [0,\ k_2]$ is called "backoff stage". Let k_2 represent maximum backoff stage, as specified this value could be larger than maximum contention window index k_1, while the CW will hold after that, as shown in (1). Let $S(t)$ be the stochastic process representing the backoff stage [0, … , k_2] of the station at slot time t. Then we have

$$\begin{cases} CW_i = r^i CW, & i \le k_1 \\ CW_i = r^{k_1} CW, & k_1 < i \le k_2 \end{cases} \quad (1)$$

where $CW=CW_{\min}$, and $r^{k_1}CW = CW_{\max}$.

Let p_t be the stationary probability that a station will transmit in an arbitrary slot time. At each transmission attempt, and regardless of the number of retransmissions suffered, each packet collides with constant and independent probability p_c. Therefore, the bi-dimensional process $\{S(t),\ T(t)\}$ is a discrete-time Markov chain, depicted in Fig.1. In this model, we define $P\{i_{n+1}, k_{n+1} | i_n, k_n\} = P\{ S(t+1)= i_{n+1}, T(t+1)= k_{n+1} | S(t)= i_n, T(t)= k_n\}$, then the only non-null one-step transition probabilities are

$$\begin{cases} P\{i, j-1 | i, j\} = 1, & j \in [1, CW_i - 1], i \in [0, k_2] \\ P\{0, j | i, 0\} = (1-p_c)/CW_0, & j \in [0, CW_0 - 1], i \in [0, k_2 - 1] \\ P\{i+1, j | i, 0\} = p_c / CW_{i+1}, & j \in [0, CW_{i+1} - 1], i \in [0, k_2 - 1] \\ P\{0, j | k_2, 0\} = 1/CW_0, & j \in [0, CW_0 - 1] \end{cases} \quad (2)$$

The first transition probability equation in (2) accounts for that, at the beginning of each slot time, the backoff timer is decremented. The second accounts for that a new packet following a successful packet transmission starts with backoff stage 0. The third accounts for that an unsuccessful retransmission makes the backoff stages increase. The fourth accounts for that, at the maximum backoff stage, the CW will be reset if the transmission is unsuccessful or restart the backoff stage for new packet if the transmission is successful.

Let $\pi_{i,j} = \lim_{t\to\infty} P\{S(t)=i, T(t)=j\}$, $i \in [0,\ k_2]$, $j \in [0,\ CW_i - 1]$ be the stationary distribution of the chain. It is easy to obtain a closed-form solution for this Markov chain. Firstly

$$p_c \pi_{i,0} = \pi_{i+1,0},\ 0 \le i \le k_2 - 1 \ \rightarrow \ \pi_{i,0} = p_c^i \pi_{0,0},\ 0 \le i \le k_2 \quad (3)$$

Owing to the chain regularities, for each $j \in [0,\ CW_i - 1]$, we get

$$\pi_{i,j} = \frac{CW_i - j}{CW_i} \begin{cases} (1-p_c)\sum_{m=0}^{k_2-1} \pi_{m,0} + \pi_{k_2,0}, & i = 0, \\ p_c \pi_{i-1,0}, & 0 < i \le k_2 \end{cases} \quad (4)$$

Making use of the fact that $(1-p_c)\sum_{i=0}^{k_2-1} \pi_{i,0} + \pi_{k_2,0} = \pi_{0,0}$, and by the relation between (3) and (4), $\pi_{0,0}$ is finally determined by imposing the normalization condition, that simplifies as follows

$$1 = \sum_{i=0}^{k_2} \sum_{j=0}^{CW_i - 1} \pi_{i,j} = \sum_{i=0}^{k_2} \pi_{i,0} \frac{CW_i + 1}{2} \rightarrow \pi_{0,0} = \left(\sum_{i=0}^{k_2} p_c^i \frac{CW_i + 1}{2}\right)^{-1} \quad (5)$$

As any transmission occurs when the backoff timer is equal to zero, regardless of the backoff stage, we have

$$p_t = \sum_{i=0}^{k_2} \pi_{i,0} = \frac{1 - p_c^{k_2+1}}{1 - p_c} \pi_{0,0} \tag{6}$$

From (1), (5) and (6), we obtain

$$p_t = \begin{cases} \dfrac{2H}{CW[1-(rp_c)^{k_2+1}](1-p_c)+H}, & k_2 \le k_1 \\ \dfrac{2H}{CW[1-(rp_c)^{k_1+1}](1-p_c)+H+CWr^{k_1}p_c^{k_1+1}(1-rp_c)(1-p_c)^{k_2-k_1}}, & k_2 > k_1 \end{cases} \tag{7}$$

where $H = (1-rp_c)(1-p_c^{k_2+1})$.

The numerical value of p_t is also constrained by the fact that p_c can be expressed in terms of p_t, that is

$$p_c = 1 - (1 - p_t)^{n-1} \tag{8}$$

where $(1 - p_t)^{n-1}$ is the probability that none of the other n -1 stations transmits. Solving (8) for p_t, we get

$$p_t = 1 - (1 - p_c)^{1/(n-1)} \tag{9}$$

Since (7) and (9) are two constraining equations for p_t as a function of p_c, the unique intersection of these two equations gives the values of p_c and p_t for given n, CW, r, k_1 and k_2. Therefore, (7) and (9) represent a nonlinear system in the two unknowns p_t and p_c, which can be solved by numerical results.

Note that we must have $p_t \in (0, 1)$ and $p_c \in (0, 1)$, then $\lim_{n\to\infty} (1 - p_c)^{1/(n-1)} = 1$, thus from (7) and (9) we have

$$\lim_{n\to\infty} (1 - rp_c)(1 - p_c^{k_2+1}) = 0, \quad \rightarrow \lim_{n\to\infty} p_c = 1/r \tag{10}$$

In fact, p_c is always less than $1/r$, which is a necessary condition for the system to reach steady state.

Since the values of denominator and numerator in (7) both equal to zero when p_c =1/r, the probability p_t converges as follows by using L-Hospital

$$\lim_{\substack{n\to\infty \\ p_c \to 1/r}} p_t = \begin{cases} \dfrac{2(1-r^{-(k_2+1)})}{1-r^{-(k_2+1)}+CW(k_2+1)(1-r^{-1})}, & k_2 \le k_1 \\ \dfrac{2(1-r^{-(k_2+1)})}{1+CW(k_1+1)-r^{-(k_2+1)}+CW(k_2-2k_1)r^{-1}-CWr^{k_1-k_2-1}}, & k_2 > k_1 \end{cases} \tag{11}$$

Note that p_t converges to a non-zero constant as n goes to infinity and p_c goes to $1/r$.

4. MAXIMUM THROUGHPUT ANALYSIS

We investigate the maximum saturation throughput by calculating the probability that there is a successful transmission in a time slot, when the number of stations n goes to infinity. Maximum saturation throughput $S_{\max}$ is a fundamental performance figure defined as the limit reached by the system throughput as the offered load increases, and represents the maximum load that the system can carry in stable conditions.

A successful transmission occurs when there is only one transmitting station. Thus the probability that there will be a successful transmission in a time slot is as

$$p_s = C_n^1 p_t (1 - p_t)^{n-1} = np_t (1 - p_t)^{n-1} \tag{12}$$

where C_n^1 is the number of ways of choosing 1 out of n stations.

We normalize the slot time as the unit time; then in any given unit time duration the average number of packets successfully transmitted is p_s. If we ignore the packet overhead, the normalized throughput is simply p_s.

As we see, since from (1), where $CW = CW_{\min}$, $r^{k_1} CW = CW_{\max}$, and the backoff time is uniformly chosen, we get that p_s grows up to the maximum saturation throughput for given r, k_1 and k_2, when CW is set as the optimal contention window size, that is

$$CW = \begin{cases} r^{k_2/2}, & k_2 \le k_1 \\ r^{k_1/2}, & k_2 > k_1 \end{cases} \tag{13}$$

Therefore the probability p_s converges as follows

$$\lim_{n\to\infty} p_s = \frac{r-1}{r}[\ln\frac{r}{r-1} + \begin{cases} \dfrac{2(1-r^{-(k_2+1)})}{1+(k_2+1)(1-r^{-1})r^{k_2/2}-r^{-(k_2+1)}}], & k_2 \le k_1 \\ \dfrac{2(1-r^{-(k_2+1)})}{1+(k_1+1)r^{k_1/2}+(k_2-2k_1)r^{k_1/2-1}-r^{-(k_2+1)}-r^{3k_1/2-k_2-1}}], & k_2 > k_1 \end{cases} \tag{14}$$

5. SIMULATION

To validate the model and support our analysis, the event-driven customer simulation is written in the C++ programming language, which attempts to emulate as closely as possible the real operation of each station. It is assumed that traffic sources with infinite stations collectively form a Poisson process with offered load on system G. In order to make it easy to compare with the results obtained in other papers, we select factor r=2, which is called Binary Exponential Backoff (BEB), the special case of EB. The parameters used here are: maximum contention window index k_1 is equal to 10, then $CW_{\max} = 2^{k_1} = 1024$; packet retry limit k_2 is equal to 16 for various values of $CW_{\min}$. Simulation results are obtained by running 1,000,000 time slots.

The simulation results for the normalized system saturation throughput S_{sat} in Table 1 and Fig. 2 agree with those obtained from our analysis, and show that with an appropriate setting of the initial minimum contention window size as optimal window size, the theoretical maximum throughput bound can be reached. Simulation result for the extreme case is $CW_{\min}$ =32, G=0.356, $S_{\max}$ =35.019%. The analytical result is $S_{\max}$ =35.004%, which practically coincides with the simulation result, in a

99.957% confidence interval lower than 0.00015, and shows that the analytical model is extremely accurate.

TABLE 1
SIMULATION RESULTS FOR SATURATION THROUGHPUT

CW_{min}	G	S_{sat}
2	0.296	29.054%
4	0.298	29.455%
8	0.340	32.801%
16	0.355	34.417%
32	0.356	35.019%
64	0.348	34.016%
128	0.346	33.753%
256	0.359	34.812%
512	0.350	34.117%
1024	0.338	33.452%

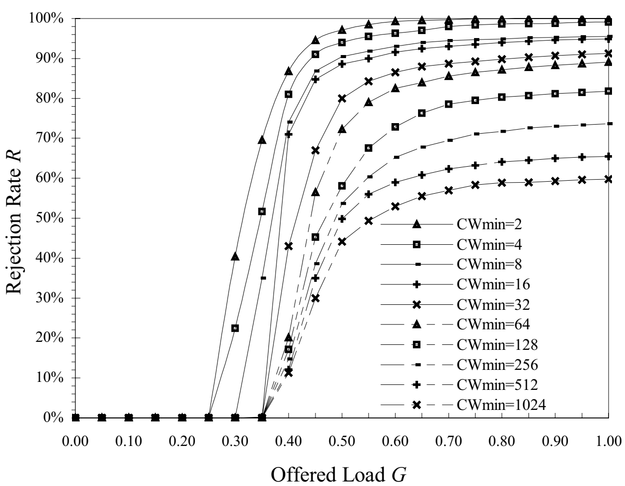

Fig.3 Rejection rate versus offered load

Moreover, the packet rejection rate R is also investigated here, which is defined as the percentage of packets that are discarded. In most realistic situations, a small rejection rate is acceptable, especially if the rejected packets have been delayed too long to be important. Fig. 3 shows that rejection rate R mounts up nonlinearly with system offered load G. By comparing the rejection rate performance for various values of CW_{min}, we get that rejection rate R consistently drops down with continuous increase of CW_{min}.

6. CONCLUSIONS

The new and simple bi-dimensional discrete-time Markov chain model developed in this paper is suited for analyzing the combinatorial characteristics of EB and retransmission cutoff. It is assumed an infinite number of stations and ideal channel conditions. Using the proposed model, we obtain the analytical performance expressions. Comparison with simulation results shows that the model is extremely accurate in predicting the system maximum saturation throughput, and in close compliance with

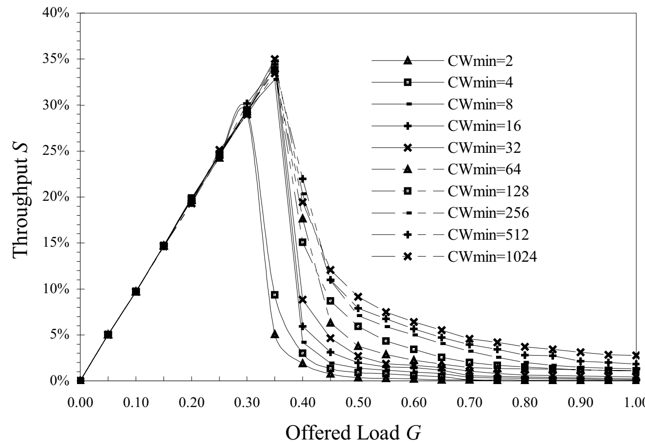

Fig.2 Throughput versus offered load

practical cases under the same network configuration. From the analysis results, we have that this combinatorial mechanism guarantees a certain amount of throughput no matter how many stations are present in the network. The packet rejection rate R is also considered, and shown that R rises nonlinearly with system offered load G, but decreases all along with continuous increase in CW_{min}. We conclude that this combinatorial mechanism efficiently enhances the system capability for resolving collisions, and improves the fairness of network contention. Moreover, this analytical method can also be applied to analyze network protocols using EB and retransmission cutoff.

7. REFERENCES

[1] G. Bianchi, "Performance Analysis of the IEEE 802.11 Distributed Coordination Function," *IEEE JSAC*, vol. 18, no. 3, pp. 535-547, March 2000.

[2] F. Cali, M. Conti, and E. Gregori, "Dynamic Tuning of the IEEE 802.11 Protocol to Achieve a Theoretical Throughput Limit," *IEEE/ACM Trans. Networking*, vol. 8, no. 6, pp. 785-799, December 2000.

[3] H. Wu, Y. Peng, K. Long, S. Cheng, and J. Ma, "Performance of Reliable Transport Protocol over IEEE 802.11 Wirless LAN: Analysis and Enhancement," *in Proc. IEEE INFOCOM 2002*, vol. 2, pp. 599-607, 23-27 June 2002.

[4] Y. Liu, "Performance Analysis of Frequency-Hop Packet Radio Networks With Generalized Retransmission Backoff," *IEEE Trans. Commun.*, vol.1, no. 4, pp. 703-711, October 2002.

[5] B.J. Kwak, N.O. Song, and L.E. Miller, "Analysis of the Stability and Performance of Exponential Backoff," *in Proc. IEEE WCNC 2003*, vol. 3, pp. 1754-1759, 16-20 March 2003.

[6] K. Sakakibara, T. Seto, and D. Yoshimura, "Effect of Exponential Backoff Scheme and Retransmission Cutoff on the Stability of Frequency-Hopping Slotted ALOHA Systems," *IEEE Trans. Commun.*, vol. 2, no. 4, pp. 714-722, July 2003.