Journal of Zhejiang University SCIENCE A
ISSN 1009-3095 (Print); ISSN 1862-1775 (Online)
www.zju.edu.cn/jzus; www.springerlink.com
E-mail: jzus@zju.edu.cn

Performance analysis of IEEE 802.11e EDCA in wireless LANs*

ZHANG Wei†1, SUN Jun[1], LIU Jing[1,2], ZHANG Hai-bin[1]
([1]*Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai 200240, China*)
([2]*State Key Laboratory of Integrated Service Networks, Xidian University, Xi'an 710071, China*)
†E-mail: zh.wei@sjtu.edu.cn

Received July 4, 2006; revision accepted Oct. 4, 2006

Abstract: In this paper, we propose an analytical model for the performance evaluation of IEEE 802.11e enhanced distributed channel access (EDCA). Different from most previous analytical studies based on the saturation assumption, we extend the analytical model to non-saturation conditions. An empty state is introduced into the Markov chain to represent the status of transmission queue being empty. This model can be used to calculate the traffic priority, throughput, and MAC layer delay with various configurations of contention parameters. A detailed simulation is provided to validate the proposed model. With the help of this model, the contention parameters can be configured appropriately to achieve specific quality-of-service (QoS) requirements.

Key words: IEEE 802.11e, Saturation assumption, Non-saturation assumption, Enhanced distributed channel access (EDCA)
doi:10.1631/jzus.2007.A0018 **Document code:** A **CLC number:** TN92

INTRODUCTION

IEEE 802.11e is designed to support multimedia applications. The main and mandatory scheme of IEEE 802.11e standard is enhanced distributed channel access (EDCA), which adopts service differentiation in configuration. The performance analysis of EDCA has been extensively studied by analytical or numerical means in recent years (Kong *et al.*, 2004; Xiao, 2005; Zhu and Chlamtac, 2005; Hui and Devetsikiotis, 2005; Zhang *et al.*, 2006). With the help of these works, performance metrics such as throughput and delay can be accurately obtained. However, these models are all based on the saturation assumption. Although performance under saturation condition provides the fundamental bounds on system throughput and delay, it cannot reveal the best working scenarios. It is proved that the maximum protocol capacity of IEEE 802.11 can only be achieved in the non-saturated case and is almost independent of the number of active stations (Zhai *et al.*, 2005). So a framework capable of analyzing the performance under both saturation and non-saturation can be very helpful in achieving deeper understanding of the EDCA mechanism. In this paper, we present a new analytical model for the EDCA mechanism under both non-saturation and saturation.

Engelstad and Østerbø (2005a; 2005b; 2005c; 2006a; 2006b) presented a queuing analysis model for IEEE 802.11e EDCA with virtual collision handler mechanism. Zhai *et al.*(2004) presented an analytical model for IEEE 802.11 distributed coordination function (DCF) and derived an approximate probability distribution of the MAC layer service time and performed M/G/1/K queuing analysis to obtain a few performance metrics of wireless LANs.

PERFORMANCE MODEL

Overview of the EDCA

Compared with DCF, which uses the same DCF interframe space (DIFS), $CW_{\min}$, and $CW_{\max}$, EDCA offers differentiated access through EDCA parameter set $AIFS[AC]$, $CW_{\min}[AC]$, $CW_{\max}[AC]$, and $TXOPlimit[AC]$ for a corresponding AC (AC=0, 1, 2, 3). The

* Project (No. 60332030) supported by the National Natural Science Foundation of China

arbitration interframe space *AIFS*[*AC*] is determined by *AIFS*[*AC*]=*SIFS*+*AIFSN*[*AC*]**aSlotTime*, where *AIFSN*[*AC*] is an integer indicating the number of slots after a short interframe space (SIFS) duration that a station should defer to before either invoking a backoff or starting a transmission. With the transmission opportunity (TXOP) scheme, a backoff entity can transmit multiple packets within one TXOP with a maximum length up to *TXOPlimit*[*AC*].

According to 802.11e specification, the contention window (CW) size is reset to $CW_{\min}$ after each successful transmission. The backoff mechanism is also used after a successful transmission before sending the next frame, even if there is no other pending MAC service data unit (MSDU) to be delivered, this is often referred to as post-backoff, as it is done after the transmission, not before. The post-backoff ensures that there is at least one backoff interval between two consecutive transmissions. There is one exception to the rule that a backoff has to be performed before any MAC protocol data unit (MPDU) transmission. If an MSDU from the higher layer arrives at the station when: (1) the transmission queue is empty, (2) the latest post-backoff has finished already, (3) the medium has been idle for at least one DIFS, it may be delivered immediately without performing the backoff procedure.

Analytical model

In this subsection, we give an analytical model of EDCA irregardless of saturation. To analyze the EDCA protocol, we make the following assumptions: (1) ideal channel conditions without hidden terminals; (2) the EDCA works in ideal synchronized slot time, a slot being equal to the duration of DIFS; (3) a finite and fixed number N of contending stations; (4) each station has the same N_t ACs, and without loss of generality, the AC with smaller number has higher priority.

For convenience, in the sequel all parameters associated with the nth AC have a subscript n. Fig.1 describes the state of the nth AC by $\{n, i, k\}$. Here, i (i=0, 1, …, l_n) represents the backoff stage, where l_n denotes the retry limit. k (k=0, 1, 2, …, $W_{n,i}-1$) represents the value of backoff counter, where $W_{n,i}$ is the contention window size in the backoff stage i for the nth AC.

In all backoff states the backoff instance

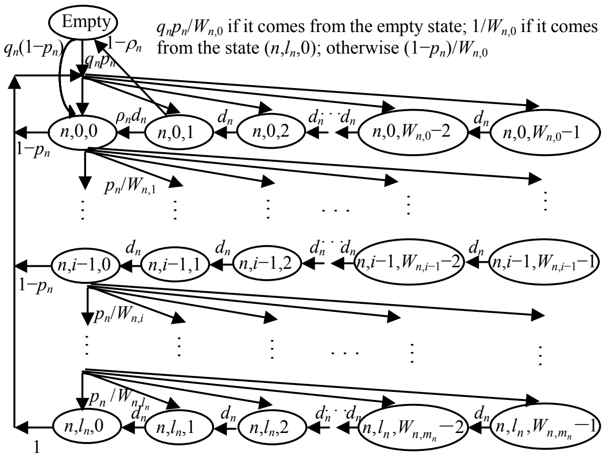

Fig.1 Proposed Markov chain model of EDCA

decreases its backoff counter at a probability d_n, which is also called as the backoff state transition rate. A transmission attempt succeeds at a probability $1-p_n$, where p_n is the conditional collision probability that a station in the backoff stage for the nth AC senses that the channel is busy. The contention window starts at the minimum value $CW_{n,\min}$ and is doubled after each collision until it reaches its maximum value. Let m_n be the maximum backoff stage. We have

$$W_{n,i}=\begin{cases}2^i(CW_{n,\min}+1), & i\le m_n,\\ W_{n,m_n}, & m_n<i\le l_n.\end{cases} \quad (1)$$

The backoff instance makes transmission attempt in any state of (n, i, 0). Let $b_{n,i,k}$ and b_{empty} denote the stationary distributions of the chain in (n, i, k) and empty, respectively. Using the relationship of $b_{n,i,0}=p_n^i b_{n,0,0}$, the transmission probability τ_n can be written as

$$\tau_n=\sum_{i=0}^{l_n} b_{n,i,0}=b_{n,i,0}\frac{1-p_n^{l_n+1}}{1-p_n},\quad 1\le n\le N_t. \quad (2)$$

To describe the network behavior in non-saturation status, the traffic characters should be taken into account. The traffic flow is characterized by its packet arrival pattern and payload statistics. In Fig.1, the probabilities ρ_n and q_n are used to represent the status of the nth AC's transmission queue. As we have shown, the backoff mechanism undergoes post-backoff each time it successfully transmitted a packet.

ρ_n represents the probability that there is a packet waiting for transmission in the queue after the post-backoff phase. If there are no packets in the transmission queue in state (n, 0, 1), the post-backoff enters the empty state and waits for a packet to arrive. If there are packets in the queue in state (n, 0, 1), the post-backoff enters the state (n, 0, 0) and starts transmission attempt.

At the empty state, packets arrive at a probability q_n. After a packet arrives, the backoff instance does a "listen-before-talk". If the backoff instance senses the channel is busy, at a probability p_n, it enters one of the states (n, 0, i). Otherwise, it moves to state (n, 0, 0) and starts transmission attempt.

Priority analysis

Owing to the chain regularities, for each $k\in(1, W_{n,j}-1)$, it is

$$b_{n,j,k}=\begin{cases}\dfrac{W_{n,0}-k}{W_{n,0}d_n}[b_{n,0,0}+b_{\text{empty}}q_n p_n], & j=0,\\ \dfrac{W_{n,j}-k}{W_{n,j}d_n}p_n^j b_{n,0,0}, & j=1,\ \ldots,\ l_n.\end{cases} \quad (3)$$

We also have $(1-\rho_n)b_{n,0,1}=q_n b_{\text{empty}}$. Thus, by using the normalization condition for stationary distribution, we have

$$\frac{1}{b_{n,0,0}}=\sum_{i=0}^{l_n}\left(1+\sum_{k=1}^{W_{n,i}-1}\frac{W_{n,i}-k}{W_{n,i}d_n}\right)p_n^i+(1-\rho_n)\frac{\displaystyle\sum_{k=1}^{W_{n,0}-1}\frac{W_{n,0}-k}{W_{n,0}d_n}q_n p_n+1}{\dfrac{W_{n,0}d_n}{W_{n,0}-1}q_n-q_n(1-\rho_n)p_n}. \quad (4)$$

Eq.(4) represents a set of nonlinear equations, which can be solved numerically. Note all probabilities should be between 0 and 1. Compared with (Xiao, 2005), it is easy to see that the first sum in Eq.(4) represents the saturation part, while the second part denotes the non-saturation. When $\rho_n\to1$, the Markov chain behavior approaches that of the saturation case. There are still some probabilities that need to be determined to solve the nonlinear equation set. Next, we explain how to derive them.

A transmitted packet collides when one more transmission queue also transmits during a slot time. And the medium is busy when at least one transmission queue transmits during a slot time. Taken the virtual collision handler mechanism into account, the conditional collision probability p_n and channel busy probability $p_{\rm b}$ can be expressed as

$$p_{\rm b}=1-\prod_{i=0}^{N_{\rm t}-1}(1-\tau_i)^N, \quad (5)$$

$$p_n=1-\frac{1-p_{\rm b}}{\prod_{i=n}^{N_{\rm t}-1}(1-\tau_i)}. \quad (6)$$

In two cases the nth AC's backoff counter decreases its value by one: (1) the channel is idle for $AIFS_n$ following a busy period; (2) after the $AIFS_n$ duration, the backoff counter decreases its value every idle time slot (Fig.2). In fact, d_n can be represented as the sum of two conditional probabilities: the probability that the channel is idle for $AIFS_n$ given it is busy in the previous time slot, and the probability that the channel is idle in the current slot given it has passed the $AIFS_n$ duration. Therefore, the sufficient condition for the backoff counter to decrease its value is that the channel is idle for at least $AIFS_n$, so we have $d_n=(1-p_{\rm b})^{AIFS_n}$.

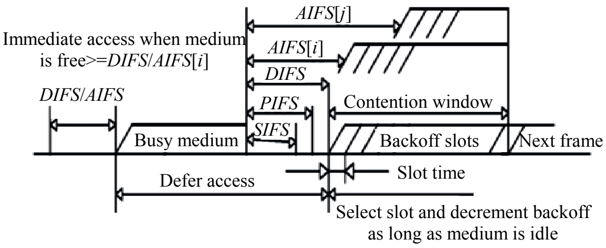

Fig.2 IEEE 802.11e EDCA mechanism parameters (IEEE 802.11 WG, 2005)

The probabilities ρ_n and q_n are used to represent the status of the nth AC's transmission queue, which is correlated with the traffic properties and MAC layer service time distribution. Let λ_n denote the traffic rate in terms of arrival packets per second and D_n denote the average medium access delay of the nth AC. For a G/G/1 queue, the probability that the transmission queue is not empty ρ_n can be expressed as $\rho_n=\lambda_n D_n$. Note that the post-backoff phase is included in the medium access delay regardless of whether the queue is saturation or not.

In non-saturation status, the Markov chain will stop in the empty state and wait for a packet to arrive. q_n is the probability that at least one packet will arrive in the transmission queue during the following generic time slot under the condition that the queue is empty at the beginning of the slot. Assuming the traffic arriving in the transmission queue is Poisson distributed, q_n is given by

$$q_n = 1 - \sum_{i=0}^{N_t-1} (p_{i,s} e^{-\lambda_n T_{i,s}} + p_{i,c} e^{-\lambda_n T_{i,c}}) - (1-p_b) e^{-\lambda_n T_e}, \quad (7)$$

here, T_e, $T_{i,c}$ and $T_{i,s}$ represent the duration of an empty slot time, the collision time and the successful transmission time of the ith AC, respectively. $p_{i,s}$ denotes the probability that a packet of the ith AC is transmitted successfully. $p_{i,c}$ represents the probability that a transmission with collision occurs in a slot time that the traffic flow with the largest payload belongs to the ith AC. Without loss of generality, we assume that from the first to the last AC, the packet length monotonously increases. We have $p_{i,s}=N\tau_i(1-p_i)$ and $p_{i,c} = [1-(1-\tau_i)^N] \prod_{j=i+1}^{N_t-1} (1-\tau_j)^N - p_{i,s}$. The correctness of these two equations can be verified by $(1-p_b) + \sum_{i=0}^{N_t-1} p_{i,s} + \sum_{i=0}^{N_t-1} p_{i,c} = 1$.

If all ACs have the same payload length, Eq.(7) can be rewritten as

$$q_n = 1 - \left[p_s e^{-\lambda_n T_s} + (1-p_b) e^{-\lambda_n T_e} + (p_b - p_s) e^{-\lambda_n T_e} \right], \quad (8)$$

where $p_s=\sum p_{i,s}$, $T_s=T_{n,s}$ and $T_c=T_{n,c}$.

$T_{n,s}$ and $T_{n,c}$ can be written as for the RTS/CTS scheme

$$\left.\begin{aligned} T_{n,s} &= RTS + 3\cdot SIFS + H + E[P_n] + CTS + ACK, \\ T_{n,c} &= RTS + SIFS + CTS\_Timeout. \end{aligned}\right\} \quad (9)$$

and for the basic scheme

$$\left.\begin{aligned} T_{n,s} &= H + E[P_n] + SIFS + ACK, \\ T_{n,c} &= H + E^*[P] + SIFS + ACK\_Timeout, \end{aligned}\right\} \quad (10)$$

where $E^*[P]$ is the average length of the longest packet payload involved in a collision and $E[P_n]$ is the average payload length of the nth AC.

Throughput and delay analysis

The throughput of the nth AC, S_n, is

$$S_n = \frac{p_{n,s} E[P_n]}{(1-p_b)T_e + \sum_{i=0}^{N_t-1} p_{i,s} T_{i,s} + \sum_{i=0}^{N_t-1} p_{i,c} T_{i,c}}. \quad (11)$$

In this paper, we define two kinds of delays, i.e., the MAC layer service time and the medium access delay. The MAC layer service time $E[T_{ser,n}]$ is defined as the duration of time taken for a state transition from a packet beginning to be served to it being dropped or successfully transmitted. The medium access delay D_n, which is used to calculate ρ_n, is defined as the duration of time from a packet beginning to be served to the time when the post-backoff is completed, viz., the Markov transition process reaches (n, 0, 0) or empty state.

In saturation condition, there are always packets waiting to be served in the transmission queue. The backoff instance enters the post-backoff phase with packets in its queue. Therefore, the MAC layer service time equals the medium access delay. However, in non-saturation conditions, the post-backoff may be already completed before a new packet arrives in the transmission queue. Thus, under these conditions the post-backoff will not add to the transmission delay. To get the MAC layer service time, the post-backoff duration needs to be subtracted from the medium access delay.

During the deference, the backoff counter may be frozen due to other node's occupation of the channel. Thus both the medium access delay and the MAC layer service include the duration of deference, freezing and the transmission time whether success or failure.

$$D_n = \sum_{i=0}^{l_n} p_{tx,n,i} D_{n,i} + (1-p_{tx,n}) D_n^{drop}, \quad (12)$$

where $D_{n,i}$ is the average delay in case of i retries, and D_n^{drop} is the average delay in case of being dropped. $p_{tx,n}$ and $p_{tx,n,i}$ are the probabilities that a packet of the

nth AC is not dropped and that it is successfully transmitted with i retries, respectively. We have $p_{\mathrm{tx},n}=\sum_{i=0}^{l_n} p_{\mathrm{tx},n,i}=1-p_n^{l_n+1}$, and $p_{\mathrm{tx},n,i}=(1-p_n)p_n^i$.

The average number of backoff slots that a station needs to transmit a packet successfully at the ith retry stage is $(W_{n,i}-1)/2$. Therefore, for a packet that is successfully transmitted after the ith retry, the corresponding average number of backoff slots is $EB_{n,i}=\sum_{j=0}^{i}(W_{n,j}-1)/2$. The average freezing slot time can be expressed as $T_{\mathrm{slot}}=\sum_{i=0}^{N_t-1}(p_{i,s}T_{i,s}+p_{i,c}T_{i,c})/p_b$. Finally, the following expression for $D_{n,i}$ is obtained

$$D_{n,i}=EB_{n,i}\left(T_e+\frac{1-d_n}{d_n}T_{\mathrm{slot}}\right)+T_{n,s}+iE[T_c], \quad (13)$$

where $E[T_c]$ is the average collision duration an AC suffers. If all the ACs have the same packet length, we have $E[T_c]=T_c=T_{n,c}$. However, if each AC has different packet length, $E[T_c]$ can be expressed as $E[T_c]=\sum_{n=0}^{N_t-1}p_{n,c}T_{n,c}\Big/\sum_{n=0}^{N_t-1}p_{n,c}$.

If a packet is dropped, the average delay is

$$D_n^{\mathrm{drop}}=EB_{n,l_n}\left(T_e+\frac{1-d_n}{d_n}T_{\mathrm{slot}}\right)+(l_n+1)E[T_c]. \quad (14)$$

The average post-backoff duration is

$$D_{n,\mathrm{post}}=\frac{(W_{n,0}-1)}{2}\left(T_e+\frac{1-d_n}{d_n}T_{\mathrm{slot}}\right). \quad (15)$$

The MAC layer service time in non-saturation status can be calculated as $E[T_{\mathrm{ser},n}]=D_n-D_{n,\mathrm{post}}$, and in saturation status $E[T_{\mathrm{ser},n}]=D_n$.

SIMULATION RESULTS

In this section, we conduct simulations using ns-2 (http://www.isi.edu/nsnam/ns/) to validate the proposed analytical model. The values of the parameters used to obtain numerical results for both the analytical model and the simulation runs are summarized in Table 1. For simplicity, here we only consider the basic mechanism of EDCA. The transmission rate is 2 Mbps.

Table 1 Calculation and simulation parameters

Parameter	Value
PHY header (including the preamble)	192 bits
MAC header (including the CRC bits)	272 bits
ACK frame	PHY header+112 bits
ACK timeout	*DIFS+ACK*
Data rate	2 Mbps
Time slot	20 μs
SIFS	1 time slot
DIFS	*SIFS*+2·*aSlotTime*
AIFSN	{2, 2, 3, 7}
Propagation delay	1 μs
CW_{min}	{7, 15, 31, 31}
CW_{max}	{15, 63, 1023, 1023}

In simulation, we consider a heterogeneous traffic scenario. The traffic generator generates packets according to the distribution of packet interarrival time and packet size. Here, all data packets arrive from the upper layer as Poisson sequence, with fixed packet length for each AC.

Fig.3 and Fig.4 show the analytical and simulation results in downlink scenario, where a QoS-enabled Access Point (QAP) transmits packets to multiple QoS-enabled Stations (QSTAs). The payload length of each AC is fixed at 1024 bytes. Simulations agree well with the analytical results. However, the MAC layer service time is not as well simulated as the throughput. As we only consider perfect saturation and non-saturation cases in delay analysis, the analytical results of MAC layer service time during the transition from perfect non-saturation to perfect saturation cannot match the simulation results exactly. This influences the accuracy of the analysis results on MAC layer service time.

It can be noted that the EDCA mechanism provides an effective way of differentiation. As the traffic load in the network increases, the throughput of higher priority traffic AC always achieves equal or much higher bandwidth than the lower priority traffic AC. This phenomenon is also revealed by the results of the MAC layer service time. The increasing traffic load leads to more collisions and a longer backoff time for each AC, which increases the MAC layer service time for each AC with the lower priority traffic ACs increase much faster than higher priority traffic AC.

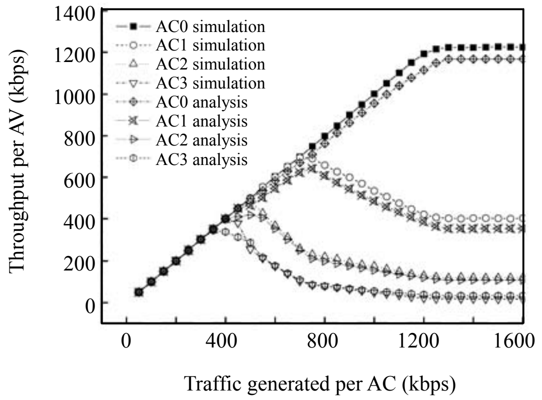

Fig.3 Numerical and simulation results for throughput

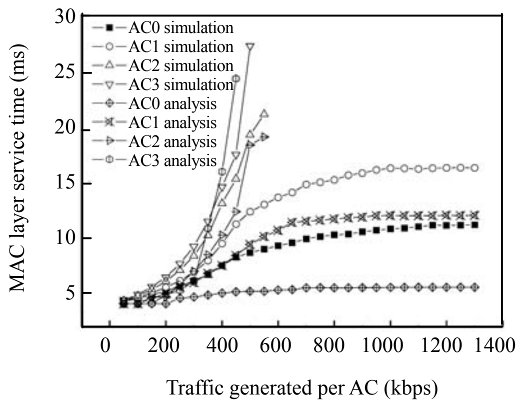

Fig.4 Numerical and simulation results for the MAC layer service time

CONCLUSION

In this paper, by taking the traffic characters into account, we present an analytical model to analyze the performance of EDCA in both saturation and non-saturation conditions. With the help of the proposed analytical model, both throughput and delay performance can be obtained.

The analytical model can provide help in studying the ability of EDCA in supporting QoS, the effect on service differentiation for each contention parameter, and provide help for parameterization for some types of traffic and development of access admission control schemes for WLANs.

References

Engelstad, P.E., Østerbø, O.N., 2005a. Differentiation of Downlink 802.11e Traffic in the Virtual Collision Handler. Proceedings of the IEEE Conference on Local Computer Networks, p.639-646.

Engelstad, P.E., Østerbø, O.N., 2005b. Delay and Throughput Analysis of IEEE 802.11e EDCA with Starvation Prediction. Proceedings of the IEEE Conference on Local Computer Networks, p.647-655.

Engelstad, P.E., Østerbø, O.N., 2005c. Non-saturation and Saturation Analysis of IEEE 802.11e EDCA with Starvation Prediction. Proceedings of the 8th ACM International Symposium on Modeling, Analysis and Simulation of Wireless and Mobile Systems. Montreal, Canada, p.224-233.

Engelstad, P.E., Østerbø, O.N., 2006a. Queueing Delay Analysis of 802.11e EDCA. Proceedings of the 3rd Annual Conference on Wireless on Demand Network Systems and Services. Les Menuires, France.

Engelstad, P.E., Østerbø, O.N., 2006b. Analysis of the Total Delay of IEEE 802.11e EDCA. Proceedings of the IEEE International Conference on Communications. Istanbul, Turkey.

Hui, J., Devetsikiotis, M., 2005. A unified model for the performance analysis of IEEE 802.11e EDCA. *IEEE Trans. Commun.*, **53**(9):1498-1510. [doi:10.1109/TCOMM.2005.855013]

IEEE 802.11 WG, 2005. Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification, Amendment 8: Medium Access Control (MAC) Quality of Service Enhancements.

Kong, Z., Tsang, D.H.K., Bensaou, B., Gao, D.Y., 2004. Performance analysis of IEEE 802.11e contention-based channel access. *IEEE J. Select. Areas Commun.*, **22**(10):2095-2106. [doi:10.1109/JSAC.2004.836019]

Xiao, Y., 2005. Performance analysis of priority schemes for IEEE 802.11 and IEEE 802.11e wireless LANs. *IEEE Trans. Wirel. Commun.*, **4**(4):1506-1515. [doi:10.1109/TWC.2005.850328]

Zhai, H.Q., Kwon, Y., Fang, Y.G., 2004. Performance analysis of IEEE 802.11 MAC protocols in wireless LANs. *Wireless Commun. Mob. Comput.* **4**(8):917-931. [doi:10.1002/wcm.263]

Zhai, H.Q., Chen, X., Fang, Y.G., 2005. How well can the IEEE 802.11 wireless LAN support quality of service? *IEEE Trans. Wirel. Commun.*, **4**(6):3084-3094. [doi:10.1109/TWC.2005.857994]

Zhang, W., Sun, J., Liu, J., Zhang, H.B., 2006. Performance analysis for IEEE 802.11e EDCA. *IEICE Trans. Commun.*, in press.

Zhu, H., Chlamtac, I., 2005. Performance analysis for IEEE 802.11e EDCF service differentiation. *IEEE Trans. Wirel. Commun.*, **4**(4):1779-1788. [doi:10.1109/TWC.2005.847113]