A Scalable Rate Adaptation Mechanism for IEEE 802.11e Wireless LANs

Hakyung Jung, Taekyoung Kwon, Yanghee Choi
Seoul National University, Seoul, Korea
Multimedia and Mobile communications Lab.
{hkjung, tk, yhchoi}@mmlab.snu.ac.kr

Yongho Seok
LG Electronics Institute of Technology, Seoul, Korea
Information and Technology Lab.
yhseok@lge.com

Abstract— **Although many rate adaptation schemes have been introduced to efficiently utilize the multiple transmission rates of the IEEE 802.11 standard, Automatic Rate Fallback (ARF) is the most widely implemented scheme on the market because of its simplicity. Our study reveals that the ARF malfunctions severely when it is used over IEEE 802.11e WLANs. The reason is that the intolerably shortened range of contention window for voice access category induces frequent collisions even when a small number of users contend. This paper demonstrates the inefficiency of the ARF using an analytic model and proposes a new rate adaptation scheme that maintains scalability. Simulation reveals that our scheme performs consistently well compared to the ARF scheme.**

I. INTRODUCTION

The IEEE 802.11 standards provide multiple transmission rates using combinations of different modulation and channel coding scheme. Since the 802.11 standard does not state how to choose the transmission rate depending on communication environments, many rate adaptation mechanisms have been proposed to efficiently utilize the multiple transmission rates. Among them, Automatic Rate Fallback (ARF) has been the most widely deployed rate adaptation scheme on the 802.11 WLAN market because of its simplicity.

On the other hand, for enhancing the original 802.11 MAC layer to support Quality of Service (QoS), IEEE 802.11e was standardized in 2005. In 802.11e, a new MAC layer function called the *hybrid coordination function* (HCF) is introduced. HCF uses both a contention-based channel access method, called *enhanced distributed channel access* (EDCA) mechanism for prioritized QoS and a polling-based method, referred to as the *HCF-controlled channel acess* (HCCA) for parameterized QoS. This paper focuses only on the case of EDCA.

In order to differentiate the channel access among different user priority traffic, 802.11e EDCA varies its own parameters. For example, a voice traffic, which has the highest priority, is given the smallest size of contention window (CW) than the others. As a negative consequence, however, this increases a frame collision probability to a significant degree even when a small number of users generate traffic. In general, rate adaptation schemes that do not differentiate frame collisions from frame transmission failures caused by channel errors show poor performance when collisions occur frequently since they may decrease the transmission rate unnecessarily. Unfortunately, previous studies [1] [2] which are capable to handle collisions are not suitable to be employed on 802.11e since they utilize Request-to-Send/Clear-to-Send(RTS/CTS) exchanges, which induce significant overhead for VoIP applications with small frame size.

Based on the above observations, we propose a new rate adaptation scheme, called Scalable Auto Rate Fallback (S-ARF). The key idea of S-ARF is deferring the decision on the reason of transmission failures until a predetermined number of consecutive transmission failures occur to surmise the reason of failures cautiously. In addition, it tolerates an intermittent failure for responsiveness to the variation of wireless channel.

The two key contributions of this paper are as follows. First, we reveal the malfunction of the ARF over 802.11e, and demonstrate it using an analytical model. Second, we design and evaluate a new rate adaptation scheme, which addresses the identified issue.

The rest of the paper is organized as follows. Related work is presented in Section II. We show the inefficiency of ARF by providing an analytical model in Section III, and evalute the S-ARF by means of simulation in Section V. Finally, this paper concludes in Section VI.

II. BACKGROUND

A. IEEE 802.11e EDCA

Based on the legacy 802.11 DCF, EDCA is designed to provide prioritized services. In the EDCA, traffic of different priorities is assigned to one of four transmit queues, which respectively correspond to four access categories (ACs): background, best effort, video and voice, in ascending order of their priorities. The basic idea of the EDCA priority mechanism is to differentiate the parameter values of each AC. That is, AIFS[AC], CWmin[AC], CWmax[AC] are used instead of DIFS, CWmin, CWmax of DCF respectively for the contention to transmit a frame. Obviously, the smaller AIFS[AC] and CWmin[AC], the higher the probability of winning the contention with the other ACs, and hence the more bandwidth share for a given traffic condition.

As shown in Table I, the range of CW for voice access category is much shorter than those of the others. As a negative consequence, it seems reasonable to conjecture that this shortened range results in a dramatic increase of a frame collision probability even when a small number of users generate voice traffic. Therefore, a rate adaptation scheme which runs over 802.11e should handle this problem properly.

B. Related work

Automatic Rate Fallback (ARF) is the first published and the most widely implemented rate adaptation algorithm [3]. However, it may work inefficiently when it is used over 802.11e WLANs. In ARF, the transmission rate is determined based on keeping track of received and missed acknowledgment (ACK) frames. When two consecutive ACKs are missed at current rate, the second retry and subsequent transmissions are performed at the next lower (*fallback*) rate. When the number of successively received good ACKs reaches 10, the transmission rate is increased to a next higher rate. However, if the first transmission (*probation*) after the rate increase fails, the rate is immediately lowered rather than retrying once again.

The rationale is that a single frame transmission failure is very likely to be due to collision, while multiple consecutive failures are to be due to the deteriorated channel condition. In case of voice category in 802.11e, however, this assumption may not be true. Because of its shortened CW size, a collision probability of voice traffic becomes much larger than that of legacy DCF even when a small number of users generate traffic, thus ARF decreases its transmission rate unnecessarily. Moreover, the increased collisions make 10 consecutive successful transmissions, which are required for the attempt to raise the transmission rate, a rare event.

On the other hand, to distinguish collisions from frame errors, the scheme called CARA uses RTS/CTS exchange when data frame transmission fails [1]. The underlying assumption is that exchanging RTS/CTS frames before transmitting data frames reduces collisions because small RTS/CTS frames can reserve the wireless channel in advance with a relatively smaller collision probability than that of data frames. However, exchanging RTS/CTS frames before transmitting voice frame may just waste the precious wireless bandwidth. The reason is that the packet size of typical VoIP application is quite smaller than that of typical data frames, thus being comparable to that of RTS/CTS frames. Therefore, we suggest that rate adaptation schemes in 802.11e WLANs not employ RTS/CTS exchanges to identify the reason of transmission failures.

III. AUTO RATE FALLBACK OVER IEEE 802.11E

In this section, we present a Markov chain model to prove that the ARF scheme shows unscalable performance over 802.11e WLANs, where the shortened range of CW in voice category suffers from the frame collision probability significantly even when a small number of contenders exist in the network.

TABLE I

EDCA DEFAULT SETTING OF CW$_{min}$ AND CW_{max}.

Access category	CWmin	CWmax
AC_BK	aCWmin	aCWmax
AC_BE	aCWmin	aCWmax
AC_VI	aCWmin/2	aCWmin
AC_VO	aCWmin/4	aCWmin/2

We assume that the channel conditions are ideal (i.e., there is no hidden terminal) and that all the stations operate in saturation: stations always have a packet available for transmission. We also assume that all the transmission failures result only from frame collisions. We use a two-dimensional discrete-time Markov chain to model the behavior of ARF as shown in Fig 1. The states are defined as couples of two integers $\{r, s\}$, with r representing the transmission rate of 802.11 and s representing the consecutive successful transmission count (0, 1, 2, ..., 10). Let p denote the probability of successful transmission, then $1-p$ is the probability of transmission failure, namely the collision probability. Let $S_{r,s}$ be the stationary distrubution of the Markov chain, then we have

$$S_{1,j} = pS_{1,j-1} \Rightarrow S_{1,j} = p^j S_{1,0} \quad \text{for } j = 1 \sim 9 \tag{1}$$

$$S_{1,0} = (1-p)S_{2,10} + (1-p)\sum_{k=0}^{9} S_{1,k} \tag{2}$$

$$S_{i,j} = pS_{i,j-1} \Rightarrow S_{i,j} = p^{j-1} S_{i,1} \quad \text{for } j = 2 \sim 9 \tag{3}$$

$$S_{i,1} = p(S_{i,0} + S_{i,10}), \quad S_{i,0} = (1-p)S_{(i+1),10} \tag{4}$$

$$S_{i,10} = (1-p)\sum_{k=0}^{9} S_{i,k} + pS_{(i-1),9} \tag{5}$$

$$\text{for } i = 2, 3, \cdots, n-1 \tag{6}$$

$$S_{n,0} = (1-p)S_{n,1} + pS_{(n-1),9} \tag{7}$$

$$S_{n,1} = p(S_{n,0} + S_{n,1}) \tag{8}$$

$$\sum_{\forall i,j} S_{i,j} = 1 \tag{9}$$

Using equations (1)∼(9), every state probabilities can be obtained for given p, and hence the average transmission rate can be expressed as:

$$E[r] = \sum_{\forall r,s} S_{r,s} * r$$

Since the collision probability $1-p$ is determined by the number of stations, we adopt the equations and the parameters from [4] [5] [6]. The number of stations is varied from 2 to 10.

Fig. 2 shows the average transmission PHY rate chosen by the ARF scheme as the number of contending stations increases. The result shows that the stations under voice category choose lower transmission PHY rates than those under the other categories due to increased collisions. One should note that the flows for the voice traffic get to use quite low transmission PHY rate even when only two stations exist. Furthermore, the ARF scheme erroneously makes voice flows choose the lowest transmission rate when three or more

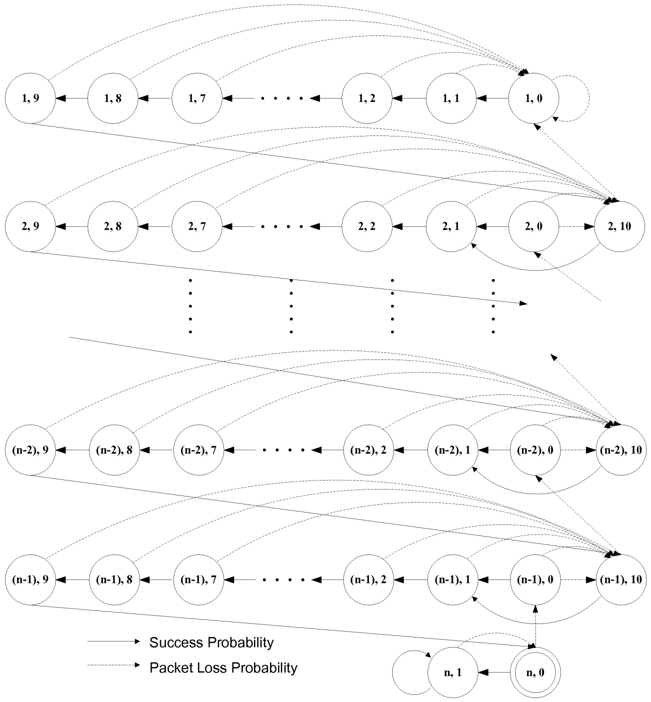

Fig. 1. Markov chain model for ARF.

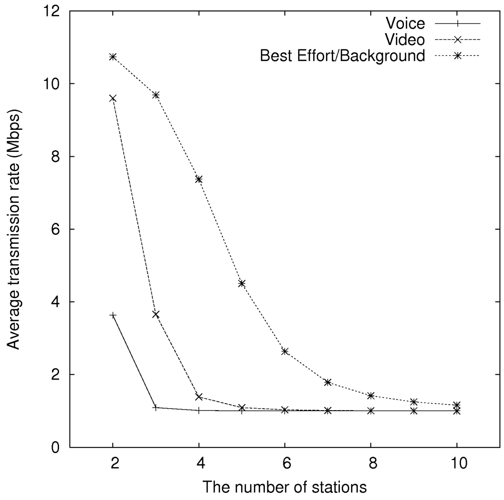

Fig. 2. The average transmission PHY rate chosen by the ARF scheme according to the varying number of contending stations

stations contend in the network. Our analytic model convinces that our conjecture of the inefficient performance of the ARF was true and a new scalable rate adaptation mechanism is needed for 802.11e WLANs.

IV. SCALABLE AUTO RATE FALLBACK (S-ARF)

This section describes the proposed Scalable Auto Rate Fallback scheme, called S-ARF. The main objective of S-ARF is attaining scalable performance of the voice flows in 802.11e WLANs.

As analyzed in Section III, the ARF scheme which has been widely deployed in the market cannot determine the proper transmission rate for voice flows when it is used over 802.11e WLANs. The shortened range of CW for voice traffic induces frequent frame collisions; thus the ARF scheme which is likely to misidentify the reason of transmission failures makes the overall system performance worse. Moreover, the high collision probability makes 10 consecutive successful transmissions a rare occasion, which triggers the attempt to increase the transmission rate to the next higher layer.

Unfortunately, previous research that is capable of distinguishing collisions from frame errors seems not feasible to be employed when voice applications are used as discussed in Section II. Our proposed scheme, S-ARF, remedies this problem without exchanging RTS/CTS frames by modifying the operation of ARF as follows.

- ARF increases the consecutive failure counter (C_{fail}) whenever it fails to receive an ACK frame, and lowers the transmission rate when the value of C_{fail} reaches the failure threshold (F_{TH}). In S-ARF we surmise that the reason of transmission failures was due to the deteriorated channel condition if the value of C_{fail} reaches a predetermined number (T_{limit}), and increase the consecutive frame error counter (C_{err}). S-ARF lowers the transmission rate if the value of C_{err}, not C_{fail}, becomes F_{TH}.
- ARF resets the value of C_{succ} if an ACK frame is missed. Instead, S-ARF resets it only if the value of C_{err} was added, that is, T_{limit} consecutive failures occurred. Moreover, the value of C_{succ} is increased by $(C_{fail}+1)$ for responsiveness.

The detailed procedure of the S-ARF operation is illustrated in a state transition diagram of Fig. 3, and its notations are summarized in Table II. The state diagram shows how the transition to Fallback state is triggered starting from Initial state, assuming that there exist only two transmission rates for simplicity. As shown in Fig. 3, the value of C_{fail} is increased when an ACK frame is missed, and then C_{fail} is compared with T_{limit}. If C_{fail} reaches T_{limit}, the value of C_{err} is increased by T_{limit} and C_{fail} is reset to zero. On the other hand, the value of C_{succ} is increased by $(C_{fail}+1)$ and C_{fail} is reset to zero when ACK frame is successfully received.

TABLE II

NOTATIONS FOR THE S-ARF STATE DIAGRAM.

Notations	Description
C_{succ}	consecutive success counter
C_{fail}	consecutive failure counter
C_{err}	consecutive frame error counter
F_{th}	failure threshold
S_{th}	success threshold (the default value is 10)
T_{limit}	pre-determined value

V. PERFORMANCE EVALUATION

In this section, we validate the performance of our scheme using the ns-2 simulator. We simulate the IEEE 802.11b PHY with the feature of IEEE 802.11e [7].

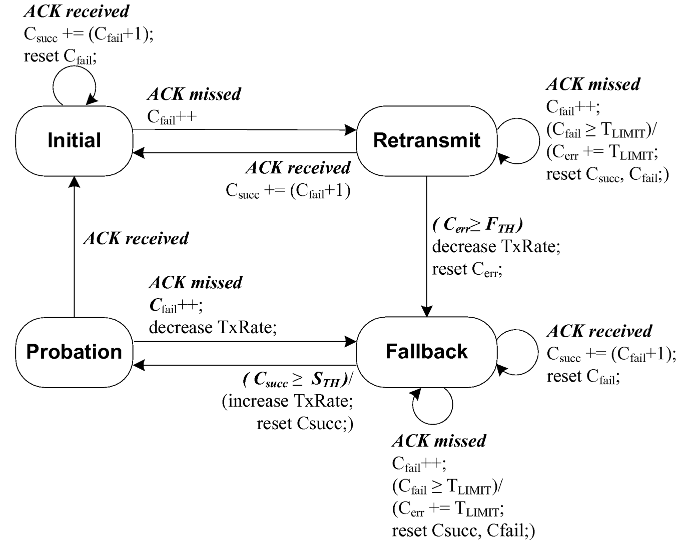

Fig. 3. State diagram of basic S-ARF operation. We assume that there exist only two transmission rates for simplicity. The expression in parentheses is a condition and expressions following a slash will be executed if the condition is satisfied.

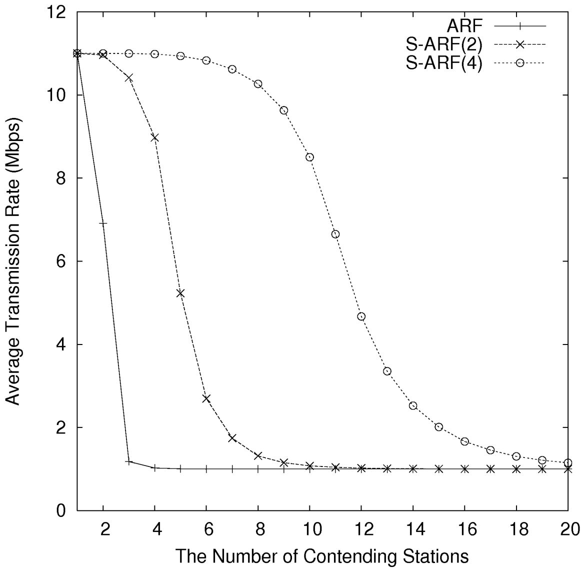

Fig. 4. Average transmission rate comparison of our proposed scheme (S-ARF(2) and S-ARF(4)) against ARF for star-topology networks with various number of contending stations in saturation.

A. Simulation Setup

We simulate the WLAN where each station uses Orinoco 802.11b PC card in indoor environment and transmits with 15 dBm power. We use the empirical BER (Bit Error Rate) to SNR (Signal-to-Noise ratio) curves for the various codings, provided by Intersil [8] to estimate a frame error rate (FER). To simulate the indoor channel condition, a shadowing propagation model is used with path loss exponent 4 [9]. Each station transmits frames for 200 seconds and results of the last 100 seconds is measured to avoid initial dynamics. We repeated this run 20 times and averaged them.

We implemented and evaluated the ARF scheme (referred to as ARF), and our proposed rate adaptation schemes: S-ARF(2) (the values of T_{limit} and F_{th} are two and four, respectively) and S-ARF(4) (the values of T_{limit} and F_{th} are four and eight, respectively). These schemes were compared with each other in terms of average transmission rate (in Mbps) and aggregate system throughput (in Mbps).

B. Star Topology with Saturated Traffic

To study how effectively S-ARF deals with collisions, we first compared the performance of S-ARF with that of ARF in a star topology. With this scenario, various number of contending stations are equidistantly placed on a circle around the AP with the radius of 10 meters, which is close enough to transmit data frames at the highest transmission rate (11 Mbps) all the time. All stations are static and always have a packet available to transmit to the AP.

Fig. 4 shows the average transmission rate when stations transmit data frames, as the number of stations increases. It can be found that ARF cannot choose high transmission rate even when the number of contenders is small. This result is analogous to that of ARF analysis in Section III. S-ARF, however, is able to select a relatively high transmission rate up to a moderate number of stations.

C. Star Topology with Voice and Video Traffic

We now compare the aggregate throughput of the voice flows in star topology where voice and video traffic flows coexist. Four stations transmit video flows to the AP and various number of stations transmit voice flows to the AP. We use CBR traffic with 64 Kbps and 160 bytes of a packet to emulate VoIP service, and use CBR traffic with 500Kbps and 1000 bytes of a packet to emulate video streaming.

Fig. 5 shows that S-ARF outperforms ARF in terms of aggregate throughput in every case. Since ARF has no mechanism that considers the effect of collisions, its performance is severely degraded as the number of voice flows increases. Our scheme is less affected by frame collisions than the ARF scheme because it defers the decision of the transmission failure until T_{limit} consecutive failures occur. Although the aggregate throughput of S-ARF is also degraded regardless of its parameter when the number of voice flows in the network exceeds nine, we presume that such a case will not be common in practice.

D. Random Mobility

Since it is probable that VoIP application is used while users are roaming around, we evaluated the performace of S-ARF using the random waypoint mobility model with reflection. The mobility domain is a two-dimensional square with each side being 60 meters long, and the AP is placed in the middle. We set the average speed to 5 m/s, average pause time to 1 sec, and average travel time to 10 sec, respectively. The simulation setup for video and voice traffic is the same as Section V-C. We vary the number of voice flows from one to fourteen and

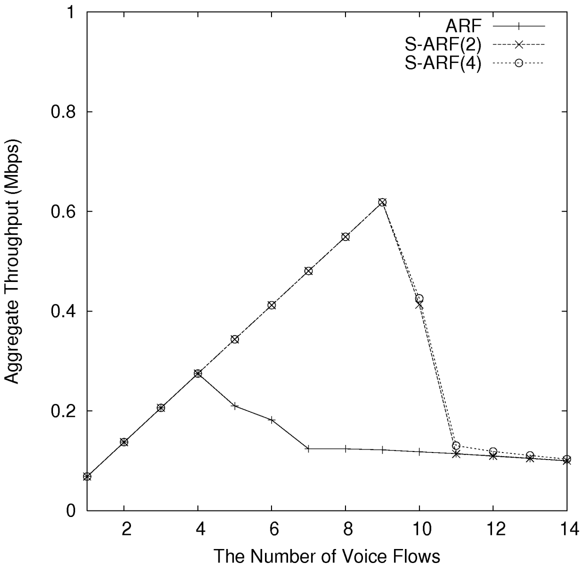

Fig. 5. Aggregate throughput comparison of our proposed scheme (S-ARF(2) and S-ARF(4)) against ARF for star-topology networks according to various number of voice flows with four background video flows.

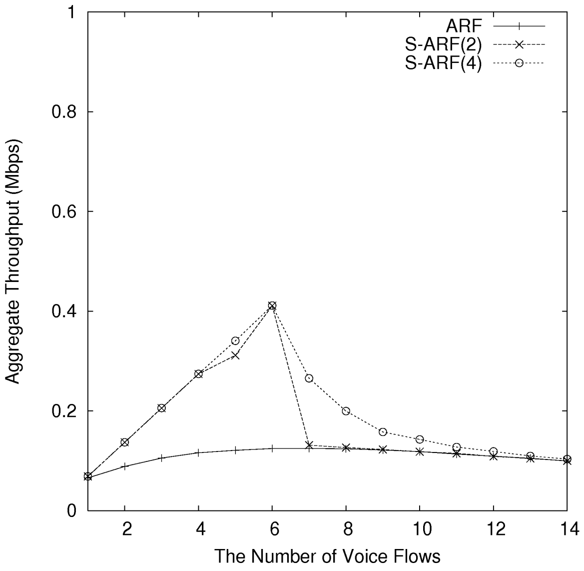

Fig. 6. Aggregate throughput comparison of our proposed scheme (S-ARF(2) and S-ARF(4)) against ARF for random mobility model according to various number of voice flows with four background video flows.

fix the number of background video flows to four. Note that, in order to obtain high throughput in this scenario, the rate adaptation scheme should be able to adapt its transmission rate quickly at a given distance and have a mechanism for handling transmission failures due to collisions.

Fig. 6 compare the performance of S-ARF with that of ARF in terms of aggregate throughput. Results show that S-ARF achieves better performance than ARF regardless of the number of contending stations. In case of ARF, ten consecutive successful transmissions are very rare events if two or more stations are contending. Thus, once it lowers its transmission rate, it is hard to restore its transmission rate again. However, ten consecutive successful transmissions are not much rare events for S-ARF since it regards a transmission failure followed by a success as a successful transmission when it updates a successful transmission counter. As a result, S-ARF may increase its transmission rate better than other schemes.

VI. CONCLUSION

In this paper, we propose a new scalable rate adaptation scheme, called S-ARF, for IEEE 802.11e WLANs by deferring the decision on the reason of transmission failures until a predetermined number of consecutive transmission failures occur to surmise the reason of failures cautiously. The two key contributions of this paper are as follows. First, we reveal the malfunction of the ARF over 802.11e, and demonstrate it using an analytical model. Second, we design the new rate adaptation scheme which addresses the identified issue, and evaluate the performance of S-ARF using simulations in various scenarios. Results reveal that the S-ARF performs consistently well compared to the ARF scheme. In the future, we plan to work on the optimization/adaptation of operational parameters including the pre-determined value (T_{limit}) and the failure threshold (F_{TH}).

REFERENCES

[1] J. Kim, S. Kim, S. Choi, and D. Qiao, “CARA: Collision-aware Rate Adaptation for IEEE 802.11. WLANs,” in Proceedings of the IEEE INFOCOM, Barcelona, Spain, 2006.

[2] S. H. Y. Wong, H. Yang, S. Lu, V. Bharghavan, “Robust Rate Adaptation for 802.11 Wireless Networks,” in Proceedings of the 12th annual international conference on Mobile computing and networking Los Angeles, CA, USA: ACM Press, 2006.

[3] A. Kamerman and L. Monteban, “WaveLAN/sup R/-II: a high-performance wireless LAN for the unlicensed band,” Bell Labs Technical Journal, vol. 2, pp. 118-133, 1997.

[4] G. Bianchi, “Performance analysis of the IEEE 802.11 distributed coordination function,” IEEE Journal on Selected Areas in Communications, vol. 18, pp. 535-547, 2000.

[5] IEEE Computer Society. 802.11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, June 1997.

[6] IEEE Computer Society. 802.11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications Amendment 8: Medium Access Control (MAC) Quality of Service Enhancements (802.11e), Nov. 2005.

[7] Technical Report, S. Wietholter, M. Emmelmann, C. Hoene, A. Wolisz, “TKN EDCA Model for ns-2,” Telecommunication Networks Group, Technische Universitat Berlin, June 2006.

[8] Intersil, “HFA3861B; Direct Sequence Spread Spectrum Baseband Processor”, January 2000.

[9] T. S. Rappaport, *Wireless Communications: Principle and Practice*, Englewood Cliffs, NJ:Prentice-Hall, 1996.