Cross-Layer Scheduling Algorithm for WLAN Throughput Improvement

Sung Won Kim

School of Electrical Engineering and Computer Science, Yeungnam University,
Gyeongsangbuk-do, 712-749, Korea
ksw@ieee.org

Abstract. Throughput improvement is critical in wireless communication networks, since the wireless channel is often shared by a number of nodes in the same neighborhood. With cross-layer design, bandwidth can be shared more efficiently by competing flows in proportion to their channel conditions. In this paper, we propose a cross-layer design for throughput improvement in IEEE 802.11 wireless local area networks (WLANs). Our protocol is derived from the Distributed Coordination Function (DCF) in the IEEE medium access control (MAC) protocol. Simulation results show that the proposed method achieves the improved throughput compared with IEEE 802.11. An important feature of the proposed method is its backward compatibility, which allows the proposed method can work with legacy IEEE 802.11 nodes.

1 Introduction

IEEE 802.11 wireless local area networks (WLANs) [1] have become increasingly prevalent in recent years. In IEEE 802.11 WLANs, a channel is shared by all nodes in the neighborhood of an access point (AP). Dividing the limited channel bandwidth efficiently among nodes is an important and challenging problem.

Currently, there has been a shift in the design of recent generation wireless networks to support the multimedia services [2]–[4], so-called cross-layer design. To improve the system throughput by using the cross-layer design, bandwidth should be shared by all competing nodes proportional to a channel condition of each link. Links that have a better channel condition must be assigned higher priority, so that they can obtain higher bandwidth. The key challenge in WLAN is that there is no centralized scheduling server, as in the case of a router output port in a wireline environment. Instead, the scheduling operation is distributed among wireless nodes with packets to transmit.

An opportunistic scheduling algorithm that exploits the inherent multi-user diversity has been implemented as the standard algorithm in the third-generation cellular system IS-856 [5] (also known as high data rate, HDR). To enable the opportunistic multi-user communications, timely channel information of each link is required for an effective scheduling. Just as all the previous schemes have assumed, the exploitation of timely channel information is possible in cellular networks where the base station acts as a central controller and control channels are available for channel state feedback.

G. Min et al. (Eds.): ISPA 2006 Ws, LNCS 4331, pp. 322–331, 2006.

When it comes down to WLANs, it is difficult to utilize the multi-user diversity. The AP cannot track the channel fluctuations of each link because of the single shared medium and the distributed Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) Medium Access Control (MAC) protocol. Wang *et al*. [6] presented the opportunistic packet scheduling method for WLANs. The key mechanisms of the method are the use of multicast RTS (Request-To-Send) and priority-based CTS (Clear-To-Send) to probe the channel status information. Since their method requires the modification of RTS and CTS in the standard, the scheme cannot be directly applied into widely deployed IEEE 802.11 typed WLANs.

On the other hand, this form of the multi-user wireless system produces asymmetric traffic loads where most of the traffic loads converge into APs. For example, Internet access or mobile computing uses transmission control protocol (TCP) or user datagram protocol (UDP) in which the offered traffic load is strongly biased toward the downlink (from AP to nodes) against the uplink (from nodes to AP) or the direct link (from nodes to nodes). Thus, these traffic flows for the downlink are completely blocked due to the CSMA/CA MAC protocol in distributed environments.

To alleviate the bottleneck problem in the downlink and exploit the multi-user diversity in WLANs, we propose a cross-layer design combining the opportunistic downlink packet scheduling and MAC protocol. The remainder of this paper is organized as follows. The next section presents related works. Section 3 describes the proposed method. In Section 4, we investigate the enhancement of the proposed method with some numerical results. Finally, the paper is concluded in Section 5.

2 Related Work

2.1 IEEE 802.11 DCF

MAC protocol in the IEEE 802.11 standard consists of two coordination functions: mandatory Distributed Coordination Function (DCF) and optional Point Coordination Function (PCF). In the DCF, a set of wireless nodes communicates with each other using a contention-based channel access method, CSMA/CA. CSMA/CA is known for its inherent fairness between nodes and robustness. It is quite effective in supporting symmetric traffic loads in ad hoc networks where the traffic loads between nodes are similar.

The DCF achieves automatic medium sharing between compatible nodes through the use of CSMA/CA. Before initiating a transmission, a node senses the channel to determine whether or not another node is transmitting. If the medium is sensed idle for a specified time interval, called the distributed interframe space (DIFS), the node is allowed to transmit. If the medium is sensed busy, the transmission is deferred until the ongoing transmission terminates.

If two or more nodes find that the channel is idle at the same time, a collision occurs. In order to reduce the probability of such collisions, a node has to perform a backoff procedure before starting a transmission. The duration of this backoff is determined by the Contention Window (CW) size which is initially set to CW_{min}. The CW value is used to randomly choose the number of slot times in the range of [0, $CW - 1$], which is used for backoff duration. In case of an unsuccessful transmission, the CW value is updated to $CW \times 2$ while it does not exceed CW_{max}. This will guarantee that in case of a collision, the probability of another collision at the time of next transmission attempt is further decreased.

A transmitter and receiver pair exchanges short RTS and CTS control packets prior to the actual data transmission to avoid the collision of data packets. An acknowledgement (ACK) packet will be sent by the receiver upon successful reception of a data packet. It is only after receiving an ACK packet correctly that the transmitter assumes successful delivery of the corresponding data packet. Short InterFrame Space (SIFS), which is smaller than DIFS, is a time interval between RTS, CTS, data packet, and ACK packet. Using this small gap between transmissions within the packet exchange sequence prevents other nodes from attempting to use the medium. As a consequence, it gives priority to completion of the ongoing packet exchange sequence.

2.2 Rate Adaptation

In [11], the auto-rate fallback (ARF) protocol for IEEE 802.11 has been presented. If the ACKs for two consecutive data packets are not received by the sender, the sender reduces the transmission rate to the next lower data rate and starts a timer. When, the timer expires or ten consecutive ACKs are received, the transmission rate is raised to the next higher data rate and the timer is canceled. However, if an ACK is not received for the immediately next data packet, the rate is lowered again and the timer is restarted. The ARF protocol is simple and easy to incorporate into the IEEE 802.11. However, as pointed out in [12], it is purely heuristic and cannot react quickly when the wireless channel conditions (e.g. signal to noise ratio, SNR) fluctuate.

In the above algorithms, the rate adaptation is performed at the sender. However, it is the receiver that can perceive the channel quality, and thus determine the transmission rate more precisely. Observing this, the authors in [13] have presented a receiver-based auto-rate (RBAR) protocol assuming that the RTS/CTS mechanism is there. The basic idea of RBAR is as follows. First, the receiver estimates the wireless channel quality using a sample of the SNR of the received RTS, then selects an appropriate transmission rate for the data packet, and piggybacks the chosen rate in the responding CTS packet. Then, the sender transmits the data packet at the rate advertised by the CTS. The simulation results in [13] show that the RBAR protocol can adapt to the channel conditions more quickly and in a more precise manner than does the ARF protocol, and thus it improves the performance greatly. Heusse *et al.* [14] have observed that in multi-rate WLANs, when certain mobile nodes use a lower bit rate than others, the performance of all nodes is considerably degraded. Specifically, the throughput of a high-bit nodes is down-equalized to that of the lowest bit-rate peer.

2.3 Throughput Fairness

Recently, throughput unfairness between the uplink and the downlink in IEEE 802.11 WLANs has received attention. In [7], the authors observe a significant unfairness between the uplink and the downlink flows when the DCF is employed in a WLAN. The TCP fairness issues between the uplink and the downlink in WLANs has been studied in [8]. Uplink flows receive significantly higher throughput than downlink flows. They find that the buffer size at the AP plays a key role in the observed unfairness, and propose a solution based on TCP receiver window manipulation. The fairness problem between uplink and downlink traffic flows in IEEE 802.11 DCF is also identified in [9]. Since in DCF, the AP and the nodes have equal access to the channel, when the downlink has a higher traffic load than the uplink, the downlink becomes a bottleneck. To solve this problem, the paper proposed a controllable resource allocation scheme between uplink and downlink flows, which adapts the parameters according to the dynamic traffic load. The scheme also improves the system utilization by reducing the collision probability. Wu and Fahmy [10] proposed a bandwidth sharing algorithm to achieve long-term throughput fairness in IEEE 802.11 WLANs. The algorithm does not require any change of the MAC frame format, which allows legacy IEEE 802.11 nodes to seamlessly coexit with the proposed method.

3 Cross-Layer Scheduling Algorithm

3.1 System Model

Each node can directly communicate only with the AP (uplink or downlink), since we focus on AP-coordinated wireless network. We propose that the AP determines the downlink channel access method according to the operation mode, DCF and ACF (AP Coordination Function). In DCF, the AP accesses the downlink channel by using the CSMA/CA. Nodes and AP use the DCF mechanism with RTS/CTS handshaking, where the next channel access should wait for DIFS and backoff window time after previous ACK packet. A two-way handshaking technique without RTS/CTS handshaking called basic access mechanism is not considered in this paper although our proposed method can be easily extended to the basic access mechanism. In ACF, the AP waits only for SIFS period instead of DIFS and backoff period. By shorting the interval period, the AP can access the channel without collision because all other nodes should wait at least DIFS period which is longer than SIFS period.

3.2 Multi-user Diversity

To switch between the two channel access methods, we propose that the AP has counters for the uplink and the downlink, denoted by $U(n)$ and $D(n)$, respectively. The counter values increase whenever there is a successful packet transmission in the uplink or the downlink. For example, when a packet is transmitted through the uplink at time n, the counter values are updated as

$$U(n) = U(n-1) + 1, \tag{1}$$

$$D(n) = D(n-1). \tag{2}$$

When $D(n) \geq U(n)$, which means the accumulated number of the downlink successful packet transmission is larger than that of the uplink, the operation mode of the AP is set to the DCF. On the contrary, when $D(n) < U(n)$, the operation mode of the AP is changed to the ACF. The two counters, $U(n)$ and $D(n)$, also update the values in the ACF and the operation mode will be changed to the DCF as soon as $D(n) \geq U(n)$. By using these two operation modes, more throughput is allocated to the downlink.

In DCF, the packet scheduling algorithm adopts the first-in first-out (FIFO) algorithm. In ACF, the AP schedules the packet based on the channel quality. Thus, the AP has to track the channel information. In order to track the latest channel quality, it is necessary to send the control packet to the node. However, this method will increase the overhead and need the modification of the IEEE 802.11 standard. Our design goal is that the proposed method can be implemented without the modification of the nodes already deployed in the system. Thus, we propose that the AP updates the channel quality of each link after every successful packet transmissions. The channel quality is reported from the physical layer by measuring the SNR of the CTS and ACK control packets. This estimation of the channel quality may not be the timely information. However, the estimation error is in the acceptable range as will be shown in the next section. Moreover, the proposed method can be implemented without the modification of the deployed nodes.

The AP lists all the communication links according to the channel quality. When the AP is in the ACF, the link that recorded the best channel quality in the previous successful transmission is given the first chance to transmit the packet in the queue. When there is no packet in the queue for that link, the next best channel quality link is given the second chance.

4 Performance Analysis

4.1 IEEE 802.11 DCF

Let N be the number of active nodes except AP. Then the probability that the successful packet transmission is performed by node n is given as

$$P_n = \frac{1}{N+1}, \quad \text{for } n = 1, 2, ..N. \tag{3}$$

The same probability applies to the AP. Let Γ be the maximum available system throughput. Then, the system throughput allocated to the downlink, Γ_d, and the uplink, Γ_u, are given as

$$\Gamma_d = \Gamma \times P_n = \Gamma \frac{1}{N+1}, \tag{4}$$

$$\Gamma_u = \Gamma \times (1 - P_n) = \Gamma \frac{N}{N+1}, \quad \text{for } n = 1, 2, ..N, \tag{5}$$

where the packet size is assumed to be the same for all the transmission. When the packet sizes for the uplink, S_u, and for the downlink, S_d, are different, (4) and (5) are changed to

$$\varGamma_d = \varGamma \frac{S_d}{S_u N + S_d}, \quad (6)$$

$$\varGamma_u = \varGamma \frac{S_u N}{S_u N + S_d}. \quad (7)$$

In this case, the ratio between the uplink throughput $\varGamma_u$ and the downlink throughput $\varGamma_u$ is given as

$$\frac{\varGamma_d}{\varGamma_u} = \left(\varGamma \frac{S_d}{S_u N + S_d}\right) \Big/ \left(\varGamma \frac{S_u N}{S_u N + S_d}\right) = \frac{S_d}{S_u} \times \frac{1}{N}. \quad (8)$$

Thus, in DCF, the allocated downlink throughput decreases as the number of nodes increases because the system throughput is shared equally between nodes.

This method is not efficient when the traffic load is asymmetric between the uplink and the downlink such as TCP and UDP. Even in the case of symmetric traffic load, the downlink traffic in DCF gets less throughput than that of the uplink and this causes the increased delay of the downlink traffic.

4.2 Simulation Parameters

We evaluate the performance of the proposed method by computer simulations. The IEEE 802.11 DCF is compared with the proposed method, named CROSS. The parameter values used to obtain numerical results for the simulation runs are based on the IEEE 802.11b direct sequence spread spectrum (DSSS) standard [1].

To reflect the fact that the surrounding environmental clutter may be significantly different for each pair of communication nodes with the same distance separation, we use the log-normal shadowing channel model [15]. The path loss PL in dB at distance d is given as

$$PL(d) = PL(d_0) + 10n \log(d/d_0) + X_\sigma, \quad (9)$$

where d_0 is the close-in reference distance, n is the path loss exponent, and X_σ is a zero-mean Gaussian distributed random variable with standard deviation σ. We set n to 3.25 and σ to 5.2 according to the result of measurements for an office building model [15]. To estimate $PL(d_0)$, we use the Friis free space equation

$$P_r(d_0) = \frac{P_t G_t G_r \lambda^2}{(4\pi)^2 d_0^2 L}, \quad (10)$$

where P_t and P_r are the transmit and receive power, G_t and G_r are the antenna gains of the transmitter and receiver, λ is the carrier wavelength, and L is the system loss factor which is set to 1 in our simulation. Most of the simulation parameters are drawn from the data sheet of Cisco 350 client adapter. The received power is

$$P_r(d) = P_t - PL(d). \quad (11)$$

The minimum received power level for the carrier sensing is set to -95 dBm, which is the noise power level. The long-term signal-to-noise ratio (SNR) is

$$SNR_L = P_t - PL(d) - \eta + PG, \tag{12}$$

where η is the noise power set to -95 dBm and PG is the spread spectrum processing gain given by

$$PG = 10 \log_{10} \frac{C}{S}, \tag{13}$$

where C is the chip rate and S is the symbol rate. Since each symbol is chipped with an 11-chip pseudonoise code sequence in the IEEE 802.11 standard, PG is 10.4 dB. The received SNR is varied by the Ricean fading gain δ. Under this model, the SNR of the received signal is

$$SNR = 20 \log_{10} \delta + SNR_L. \tag{14}$$

For the data rate in the physical layer for each communication link, we assume that the system adapts the data rate by properly choosing one from a set of modulation scheme according to the channel condition. The set of modulation schemes used in our simulation studies are BPSK, QPSK, 16QAM, 64QAM, and 256QAM. For simplicity, we ignore other common physical layer components such as error correction coding. With 1 MHz symbol rate and the above modulation schemes, the achieved data rates are 1, 2, 4, 6, and 8 Mbps, respectively.

We assume that all nodes except the AP are randomly distributed in the circle area with diameter 150 meters and move randomly at speed 0.1m/sec. The AP is located at the center of the area. To evaluate the maximum performance, traffic load is saturated in each nodes and the destination addresses of the packets are the AP. In the AP, there are N connections, each for one node, and packets are generated for each connections with the same pattern as those in each nodes. To make an asymmetric traffic load condition between uplink and downlink, the size of the downlink and uplink packets are 1024 and 64 bytes, respectively. The number of node N is set to 25. The effects of the uplink packet size and the number of nodes on the performance are also evaluated by the simulation.

4.3 Numerical Results

The effect of the uplink packet size on the downlink and uplink throughtput is shown in Fig. 1. The uplink packet size in the figure is normalized to 64 bytes. As explained in (7), the uplink throughput of DCF increases as the uplink packet size increases. This trend also applies to CROSS because the relative overhead size such as RTS, CTS, ACK, backoff, and collision period is reduced as the uplink packet size increases. DCF provides larger uplink throughput because it gives the same channel access chance to all the active nodes. Note that the most of the system throughput of DCF is allocated to the uplink which leads to the downlink bottle neck problems in asymmetric traffic load conditions. As explained in (8),

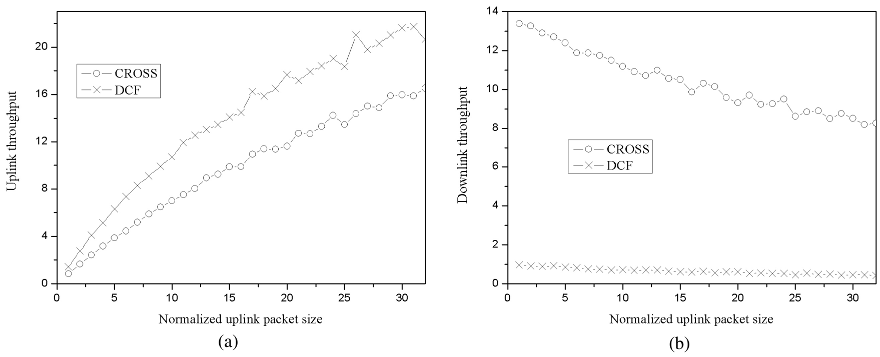

Fig. 1. Downlink and uplink throughput versus normalized uplink packet size

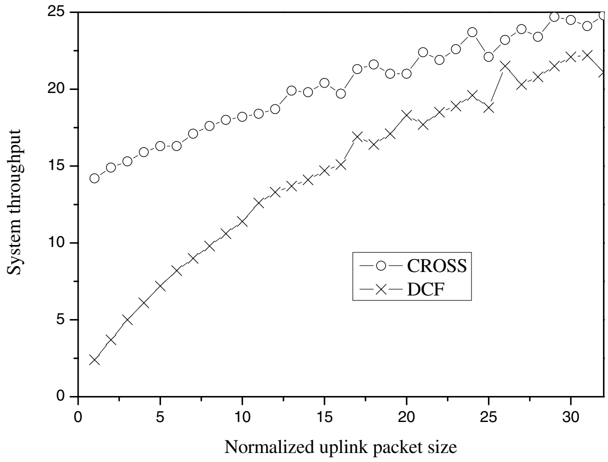

Fig. 2. The effect of the uplink packet size on the system throughput

the downlink throughput of DCF decreases as the uplink packet size increases. This trend also applies to CROSS. Compared with DCF, the proposed method provides larger throughput to the downlink and can mitigate the bottleneck problem of asymmetric traffic load condition.

The effect of the uplink packet size on the system throughput is shown in Fig. 2. As the uplink packet size increases, the system throughput increases. This is because the throughput increase of the uplink is more than that of the downlink. It is also shown that the proposed method outperforms than DCF.

The system throughput of the proposed method is compared with DCF in Fig. 3 by changing the number of nodes. In DCF, the system throughput decreases as the number of nodes increases. This decrease of the system throughput mainly comes from the increased collision between the packet transmissions. The probability of the packet collision increases as the number of nodes increases. On the contrary, the proposed method maintains a constant system throughput

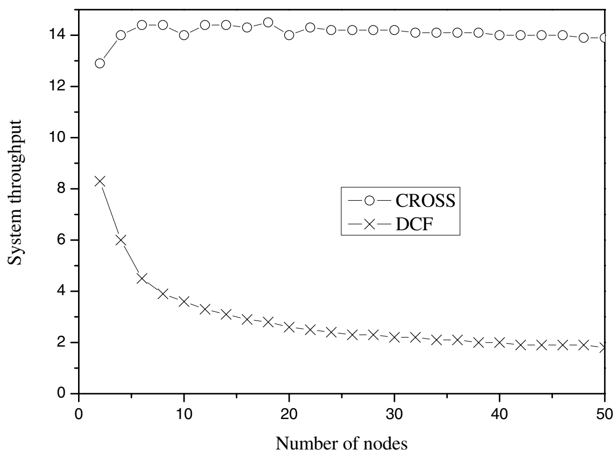

Fig. 3. The effect of the number of nodes on the system throughput

because it provides contention-free access method for the AP. Also note that the proposed method provides more system throughput than DCF because faster data rate is provided for the packet transmission during the ACF.

5 Conclusion

In order to increase the system throughput of WLAN, efficient cross-layer methods are actively worked. In this paper, we proposed the cross-layer method that combines the scheduling method, MAC layer protocol, and physical layer information. Depending on the channel conditions, channel access method and the scheduling method are dynamically changed. In the performance analysis and the simulation results, we showed that IEEE 802.11 DCF has the problem of throughput unfairness between the uplink and the downlink. It is also shown that the proposed method provides more system throughput than IEEE 802.11 DCF and alleviates the problem of throughput unfairness.

References

1. IEEE Std 802.11b-1999: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Higher-Speed Physical Layer Extension in the 2.4 GHz Band. (1999)
2. Zhang, Y.J., Letaief, K.B.: Adaptive resource allocation and scheduling for multiuser packet-based OFDM networks. In: Proc. IEEE ICC 2004. Volume 5., Paris, France (2004) 2949–2953
3. Johnsson, K.B., Cox, D.C.: An adaptive cross-layer scheduler for improved QoS support of multiclass data services on wireless systems. IEEE J. Select. Areas Commun. **23** (2005) 334–343

4. Shakkottai, S., Rappaport, T.S., Karlsson, P.C.: Cross-layer design for wireless networks. IEEE Commun. Mag. (2003) 74–80
5. IS-856: CDMA 2000 standard: High rate packet data air interface specification. (2000)
6. Wang, J., Zhai, H., Fang, Y.: Opportunistic packet scheduling and media access control for wireless LANs and multi-hop ad hoc networks. In: Proc. IEEE WCNC 2004, Atlanta, Georgia (2004) 1234–1239
7. Grilo, A., Nunes, M.: Performance evaluation of IEEE 802.11e. In: Proc. IEEE PIMRC 2002, Lisboa, Portugal (2002)
8. Pilosof, S., Ramjee, R., Raz, D., Shavitt, Y., Sinha, P.: Understanding TCP fairness over wireless LAN. In: Proc. IEEE Infocom 2003, San Francisco, CA, USA (2003)
9. Kim, S.W., Kim, B., Fang, Y.: Downlink and uplink resource allocation in IEEE 802.11 wireless LANs. IEEE Trans. Veh. Technol. **54** (2005) 320–327
10. Wu, Y., Fahmy, S.: A credit-based distributed protocol for long-term fairness in IEEE 802.11 single-hop networks. In: Proc. IEEE WiMob 2005. (2005) 98–105
11. Kamerman, A., Monteban, L.: WaveLAN-II: A high-performance wireless LAN for the unlicensed band. Bell Labs Tech. J. **2** (1997) 118–133
12. Qiao, D., Choi, S., Shin, K.G.: Goodput analysis and link adaptation for IEEE 802.11a wireless LANs. IEEE Trans. Mob. Comput. **1** (2002) 278–292
13. Holland, G., Vaidya, N., Bahl, P.: A rate-adaptive MAC protocol for multi-hop wireless networks. In: Proc. IEEE/ACM MOBICOM 2001, Boston, MA, USA (2001) 236–251
14. Heusse, M., Rousseau, F., Berger-Sabbatel, G., Duda, A.: Performance anomaly of 802.11b. In: Proc. IEEE Infocom 2003, San Francisco, CA, USA (2003)
15. T. S. Rappaport: Wireless communications: principles and practices, 2nd Ed. Prentice Hall (2002)