Ph.D. DISSERTATION

Resolving Performance Anomaly and Hidden Node Problem in IEEE 802.11n Wireless LANs

IEEE 802.11n 무선 랜에서의 비정상적 성능 저하 현상과 히든 노드 문제의 해결

BY

MINHO KIM

FEBRUARY 2013

DEPARTMENT OF ELECTRICAL ENGINEERING AND
COMPUTER SCIENCE
COLLEGE OF ENGINEERING
SEOUL NATIONAL UNIVERSITY

Ph.D. DISSERTATION

Resolving Performance Anomaly and Hidden Node Problem in IEEE 802.11n Wireless LANs

IEEE 802.11n 무선 랜에서의 비정상적 성능 저하 현상과 히든 노드 문제의 해결

BY

MINHO KIM

FEBRUARY 2013

DEPARTMENT OF ELECTRICAL ENGINEERING AND
COMPUTER SCIENCE
COLLEGE OF ENGINEERING
SEOUL NATIONAL UNIVERSITY

Resolving Performance Anomaly and Hidden Node Problem in IEEE 802.11n Wireless LANs

IEEE 802.11n 무선 랜에서의 비정상적 성능 저하 현상과 히든 노드 문제의 해결

지도교수 최 종 호

이 논문을 공학박사 학위논문으로 제출함

2012 년 11 월

서울대학교 대학원

전기컴퓨터 공학부

김 민 호

김민호의 공학박사 학위논문을 인준함

2012 년 12 월

위 원 장	:	박	세 웅
부위원장	:	최	종 호
위 원	:	최	성 현
위 원	:	오	성 회
위 원	:	박	은 찬

Abstract

The demand for wireless local area network (WLAN) has drastically increased due to the prevalence of the mobile devices such as smart phones and tablet PCs. In this dissertation, we study the two defective phenomenons in IEEE 802.11 WLANs, which are *performance anomaly* and *hidden node problem*. Each node may have a different amount of airtime because the basic channel access mechanism, distributed coordination function (DCF), in IEEE 802.11 was originally designed to provide fair chance to access the channel, regardless of packet size and data rate. This can lead to the degradation of overall network throughput and airtime fairness among nodes, which is known as performance anomaly. The hidden node problem occurs due to the failure of the carrier sense multiple access with collision avoidance (CSMA/CA) mechanism. When a transmitter cannot carrier sense the signal of another node even though its signal can interfere with the frame reception at the receiver, this node is called as a hidden node which leads to degradation of network performance. Recently, as service providers and personal users install access points (APs) without considering nearby basic service sets (BSSs), a large number of BSSs can overlap with each other, causing the hidden node problem to occur more frequently.

To resolve the performance anomaly, we first study the problem of optimal frame aggregation in IEEE 802.11n, when a wireless channel is shared by heterogeneous nodes of different data rates and packet sizes. We employ a generalized two-level frame aggregation scheme that combines the conven-

tional aggregated medium access control (MAC) service data unit (A-MSDU) and aggregated MAC protocol data unit (A-MPDU) schemes, and formulate an optimization problem to maximize the achievable throughput while assuring airtime fairness. On the basis of the solution of the optimization problem, we propose a frame size adaptation (FA) scheme that adjusts the number of packets in a frame according to the data rate and packet size. The FA scheme is fully compatible with the IEEE 802.11n standard and works in a distributed manner, which neither modifies the channel access mechanism nor resorts to a centralized scheduling algorithm. The simulation results confirm that the FA scheme tightly regulates the airtime usage of each node to be almost the same and significantly improves the overall network throughput compared to other existing schemes.

The solutions to alleviate the hidden node problem, such as the request-to-send/clear-to-send (RTS/CTS) exchange, often contain protocol overhead that may lead to throughput degradation when there is no hidden node. Therefore, it is necessary to develop a hidden node detection mechanism that can be used to trigger a resolution mechanism for hidden nodes. By utilizing frame aggregation, block acknowledgement (ACK), and fast link adaptation (FLA) in IEEE 802.11n, we propose a novel hidden node detection (HD) mechanism that takes three main causes of frame losses, which are collisions, hidden nodes, and channel impairments, into consideration. The proposed HD mechanism detects hidden nodes based on measurable MAC layer statistics and the received block ACK frame, and determines whether or not to use the RTS/CTS exchange. We show in a simulation study that the HD mechanism can detect hidden nodes well under various circumstances and that the network throughput can

be improved by using the RTS/CTS exchange adaptively in conjunction with the HD mechanism.

Finally, we propose an enhanced mechanism for resolving the hidden node problem in conjunction with the FA scheme and the HD mechanism. We first classify all possible scenarios of the hidden node problem according to the relative location of nodes in two flow interaction. Through a simulation study, we investigate the effectiveness of the RTS/CTS mechanism for each scenario and show that the malfunction of the RTS/CTS mechanism is due to the limitation of the effective CTS range and the carrier sensing mechanism at the transmitter-side. The proposed mechanism enhance the capability of the RTS/CTS mechanism in two ways: by extending the effective CTS range, and by adopting the receiver-oriented contention (ROC) mechanism. We provide extensive simulation results to show that the proposed mechanism can effectively resolve the hidden node problem, and the network throughput and fairness can be significantly improved.

Keywords: Wireless LAN, IEEE 802.11n, RTS/CTS, Frame aggregation, Performance anomaly, Airtime fairness, Hidden node problem, Hidden node detection

Student Number: 2007-20936

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The explosive growth of mobile devices such as smart phones and tablet PCs accelerates the demand for wireless Internet access. The wireless local area network (WLAN) is one of the most popular wireless communication technology thanks to its ease of deployment and low installation cost [1], [2]. To satisfy the increasing demand for higher throughput of WLANs, the IEEE 802.11n standard [3] introduces new physical (PHY) layer and medium access control (MAC) layer specifications. By using advanced PHY layer technologies such as multiple-input multiple-output (MIMO) antenna, orthogonal frequency division multiplexing (OFDM), adaptive channel coding, and channel bonding, the data rate in the PHY layer reaches up to 600 Mb/s with a 40 MHz channel bandwidth and 4×4 MIMO configuration. However, MAC-layer overheads such as the MAC header, contention time, and acknowledgement (ACK) limit the actual throughput. To reduce these overheads, the MAC layer in IEEE 802.11n introduces several enhancement mechanisms including frame aggregation, block

ACK, and reverse direction.

In a typical and real WLAN environment, both the channel state and user application are quite different among nodes. When a wireless channel is shared among many heterogeneous nodes of different data rates and packet sizes, sharing the channel in a fair and efficient manner is imperative. The basic channel access mechanism adopted in IEEE 802.11, i.e., the distributed coordination function (DCF), was designed to provide long-term equal channel access opportunity for all contending nodes. This design philosophy was intended to provide *throughput-fairness* among nodes that have similar-sized packets. In a multi-rate environment, however, the DCF can significantly degrade the overall network throughput. Actually, the throughput of a node with high data rate is limited by a node with low data rate, which is well known as *performance anomaly* [4]. For example, let us consider an example in Fig. 1.1. For simplicity, we assume that the transmitters of both links have the same size of packets, and the transmitter of `Link1` uses a lower data rate than that of `Link 2`, which results in the transmission time for the data and ACK frame to be twice larger than that of `Link2`. The two links fairly share the channel resulting from the DCF, and thus, they may have the same number of data transmissions for a given duration of T as shown in Fig. 1.1. In Fig. 1.1, both links equally have four chances to transmit data and ACK frame transmissions within T. Consequently, the throughput of both links are the same regardless of the different data rate. Even though the transmitter of `Link2` has a higher data rate, it only achieves the similar throughput as `Link1`. This problem neutralizes many possible advantages of employing higher data rates and it is exacerbated when the difference among the data rates increases Similarly, a performance anomaly

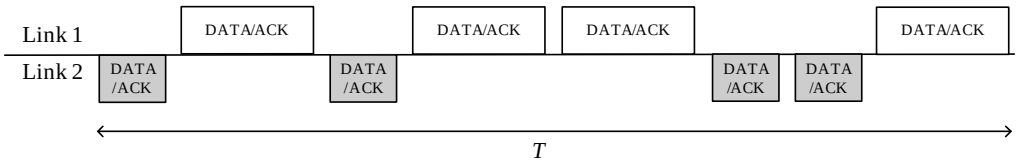


Figure 1.1 Example of the performance anomaly of two links with different data rate.

problem also arises among nodes that transmit packets of different sizes.

On the other hand, in the IEEE 802.11 [5] based wireless LANs (WLANs), the hidden node problem [6] can arise in various forms due to the nature of carrier sense multiple access with collision avoidance (CSMA/CA) mechanism. When a transmitter cannot carrier sense the signal of another node even though its signal can interfere with the frame reception at the receiver, this node is called as a hidden node which leads to degradation of network performance. For example, let us consider a network topology depicted in Fig. 1.2. There are two overlapped BSSs, and let us assume that mobile node 1 (`MN1`) transmits data frames to access point 1 (`AP1`) in the uplink and `AP2` transmits data frames to `MN2` in the downlink. The transmission of `AP2` cannot be sensed by `MN1`, and thus, `MN1` may simultaneously transmit data frames to `AP1` even though the transmission of AP2 interferes with the transmitted signal. This problem can severely degrade the throughput of `MN1`. Recently, as service providers and personal users install APs without considering nearby basic service sets (BSSs), a large number of BSSs can overlap with each other, causing the channel interference between BSSs to become a serious problem. In such an environment, the hidden node problem [6] can occur more frequently and the CSMA/CA mechanism in WLAN is likely to fail, leading to a significant performance degradation.

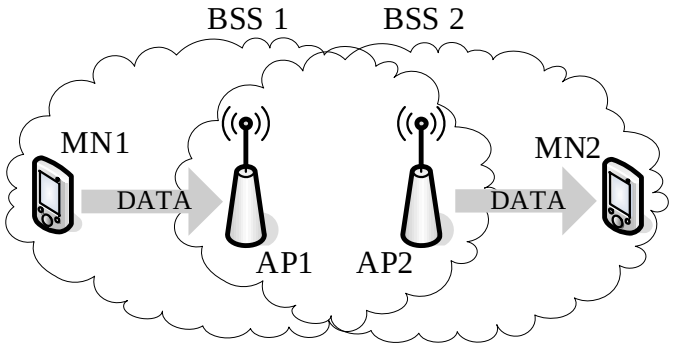


Figure 1.2 Example of the hidden node problem in the overlapped BSS network.

We first study the problem of sharing a wireless channel among heterogeneous nodes that have different data rates and packet sizes to resolve the performance anomaly. The objective of this work is two-fold: (*i*) to strictly enforce airtime fairness regardless of the data rate and packet size and (*ii*) to maximize the network throughput by considering the channel error rate, data rate, and MAC overhead. For this, we propose a frame size adaptation (FA) scheme by employing a generalized two-level frame aggregation scheme that combines the aggregated MAC service data unit (A-MSDU) and aggregated MAC protocol data unit (A-MPDU) schemes. Extensive simulation results confirm that the proposed scheme significantly outperforms other existing schemes in terms of the aggregate throughput and channel access delay while maintaining a high level of airtime fairness. The FA scheme also plays one of the key roles in resolving the hidden node problem in the next part of the dissertation.

The solutions to resolve the hidden node problem such as the RTS/CTS mechanism often contains an unnecessary protocol overhead when there is no hidden node, leading to throughput degradation. Therefore, it is important to develop a detection mechanism for hidden nodes, which can be used to trigger a

resolution mechanism for such nodes. For this, we propose a novel hidden node detection (HD) mechanism while considering the three main causes of frame losses with the help of the new features in IEEE 802.11n [3]. Frame losses can be classified into two types in IEEE 802.11n: entire and partial frame losses. We first classify a frame loss according to its type and cause. Interference from hidden nodes can cause both entire and partial frame losses. In order to detect both types of frame losses due to hidden nodes, we propose two hidden node detection methods. One is based on the statistics of measurable parameters in the MAC layer, and the other is based on the received block ACK frame. When no hidden nodes are detected, the transmitter only adjusts its modulation and coding scheme (MCS) and contention window (CW) values to cope with channel errors and collisions. When hidden nodes are detected, the transmitter employs the RTS/CTS exchange if it is expected to give better throughput performance than the basic access mode. We evaluate the performance of the proposed HD mechanism through simulation. It detects the hidden nodes well under various circumstances and the network throughput is improved through the use of the HD mechanism with RTS/CTS.

Finally, we propose an enhanced resolution mechanism for the hidden node problem with the help of the FA scheme and HD mechanism. We first investigate the effectiveness of the RTS/CTS mechanism via simulation study for all possible scenarios of the hidden node problem. To improve the capability of the RTS/CTS mechanism, we propose a enhanced mechanism for resolving the hidden node problem in two ways: *(i)* by extending the effective CTS range and *(ii)* by adopting the receiver-oriented contention (ROC) mechanism. The FA scheme and the HD mechanism help to extend the effective CTS range to the

carrier sensing range and to adaptively use the ROC mechanism, respectively. Through extensive simulation under various network topologies, we show that the proposed mechanism effectively resolve the hidden node problem, and the network throughput and fairness can be significantly improved.

The dissertation is organized as follows. Chapter 2 demonstrates the operation of the FA scheme for resolving performance anomaly. In Chapter 3, the HD mechanism is proposed to detect the presence of hidden nodes, by which the RTS/CTS exchange can be used depending on the network environment. In Chapter 4, the resolution mechanism for hidden node problem is proposed by applying the FA scheme and HD mechanism. The organizational details of Chapters 2, 3, and 4 can be found in the end of Sections 2.1, 3.1, and 4.1, respectively. Finally, we conclude the dissertation in Chapter 5.

Chapter 2

Frame Size Adaptation for Resolving Performance Anomaly

2.1 Background and Related Work

In a typical and real WLAN environment, both the channel state and user application are quite different among nodes. For example, a certain node may transmit small-size VoIP packets[1] at a high data rate (e.g., 20-byte packets at 600 Mb/s), while another node may transmit large-size data packets at a low data rate (e.g., 1500-byte packets at 6.5 Mb/s). In this extreme case, the channel occupation times to transmit a single packet differ by an order of magnitude even when considering the MAC/PHY headers. Therefore, when a wireless channel is shared among many heterogeneous nodes of different data

[1]Unless otherwise stated, we refer to a *packet* as what MAC receives from the upper layer and a *frame* as what MAC transfers to the lower layer. We will specifically define these terms in the next section.

rates and packet sizes, sharing the channel in a fair and efficient manner is imperative. The basic channel access mechanism adopted in IEEE 802.11, i.e., the distributed coordination function (DCF), was designed to provide long-term equal channel access opportunity for all contending nodes. In a multi-rate environment, however, the DCF can significantly degrade the overall network throughput, which is well known as *performance anomaly* [4]. This problem neutralizes many possible advantages of employing higher data rates and it is exacerbated when the difference among the data rates increases (the data rate supported in IEEE 802.11n ranges from 6.5 Mb/s to 600 Mb/s). Similarly, a performance anomaly problem also arises among nodes that transmit packets of different sizes.

The performance anomaly problem can be resolved by enforcing fairness in terms of the channel access time, referred to as *airtime fairness*, or *temporal fairness* [7, 8]. There are several proposals for airtime fairness in the literature, and most of them adjust the channel access probability according to the data rate and packet size [9–13]. In [9], a transmitter adjusts its contention window size to be inversely proportional to the data rate under the assumption that the packet size is similar for all nodes. The studies in [10–13] extend the basic idea in [9]. The authors in [13] provide the optimal channel access probability for each node to achieve airtime fairness and to maximize the aggregate throughput. The transmission opportunity (TXOP) mechanism in IEEE 802.11e [14] can be used to achieve airtime fairness. The amount of airtime consumed by each node can be brought to a comparable level by transmitting multiple back-to-back data packets within the same TXOP duration. However, the granularity of control in TXOP is coarse.

On the other hand, the IEEE 802.11n [3] standard introduces two types of frame aggregation schemes, an aggregated MAC-level service data unit (A-MSDU) and an aggregated MAC-level protocol data unit (A-MPDU), to increase the throughput by decreasing the MAC-layer overheads. The standard only specifies the frame format and maximum size of the aggregated frame; however, the optimal frame size and operation mode are not standardized and are left undetermined. Recently, several analytical studies have investigated the effect of A-MPDU and/or A-MSDU frame sizes on the achievable throughput and their optimal sizes to improve throughput [15–19].

In this chapter, we study the problem of sharing a wireless channel among heterogeneous nodes that have different data rates and packet sizes. The objective of this chapter is two-fold: (*i*) to strictly enforce airtime fairness regardless of the data rate and packet size and (*ii*) to maximize the network throughput by considering the channel error rate, data rate, and MAC overhead. For this purpose, we propose a frame size adaptation (FA) scheme by employing a generalized two-level frame aggregation scheme that combines the A-MSDU and A-MPDU schemes. Compared with other existing studies, this study makes the following contributions.

- In contrast to the approaches for airtime fairness in [9–13], the proposed mechanism is novel and effective; it (*i*) utilizes the frame aggregation scheme without modifying the basic channel access mechanism to fully comply with the IEEE 802.11n standard and (*ii*) introduces the notion of target airtime and controls the actual airtime tightly close to the target value.

- Unlike the frame aggregation schemes to increase throughput in [15–19], this work proposes a generalized solution; it (*i*) aims to improve both efficiency and fairness and (*ii*) employs two-level aggregation to take advantage of both A-MSDU and A-MPDU.

- In determining the optimal size for frame aggregation, most current analytical models do not consider the effect of link adaptation and the optimal size cannot be represented in a closed form. In contrast, the analysis model in this study incorporates the link adaptation algorithm [3] employed in 802.11n. In addition, it formulates the problem as an optimization problem and provides a closed-form solution for optimal frame aggregation.

Extensive simulation results confirm that the proposed scheme significantly outperforms other existing schemes in terms of the aggregate throughput and channel access delay while maintaining a high level of airtime fairness.

The rest of the chapter is organized as follows. In Section 2.2, we describe the preliminaries of this chaper. We formulate the problem as an optimization problem in Section 2.3 and provide an effective method to adjust frame size in Section 2.4. In Section 2.5, we validate the proposed scheme via simulations and compare its performance with other existing schemes. Finally, we conclude the chapter in Section 2.6.

2.2 Preliminaries

In this section, we first present a motivating example that demonstrates the performance anomaly of the IEEE 802.11 DCF. We then introduce frame ag-

Table 2.1 Performance anomaly of IEEE 802.11 DCF due to differences in packet size and data rate.

	N1	N2	N3	N4
Packet size (bytes)	250	1000	250	1000
Data rate (Mb/s)	13	13	65	65
Unit transmission time (ms)	0.312	0.774	0.150	0.242
Attempt probability	0.093	0.093	0.090	0.091
Throughput (Mb/s)	1.061	4.109	1.073	4.197
Airtime ratio	0.166	0.397	0.080	0.127

gregation schemes in IEEE 802.11n, which will be used as basic tools for the proposed solution to resolve the performance anomaly problem.

2.2.1 Performance anomaly in a heterogeneous network environment

The IEEE 802.11 DCF was basically designed to provide a fair channel access opportunity for all contending nodes regardless of the data rate and packet size. Therefore, it results in a performance anomaly in a heterogeneous network environment where the nodes have different packet sizes and data rates [4]. To illustrate this problem, we performed a preliminary simulation under a simple scenario where four nodes (denoted as `N1` $\sim$ `N4`) transmit data packets of different sizes at different data rates as described in Table 2.1. Here, we assume an ideal error-free channel. The *unit transmission time* in Table 2.1 is calculated as the time required to transmit a single data frame according to the IEEE 802.11n specification.[2] Also, throughout the chapter, the *airtime ratio* of

[2]The unit transmission time does not take into account the backoff time but includes the time for the PHY/MAC headers, short inter-frame space (SIFS) and ACK transmission. It is noted that the ACK is transmitted at the lowest data rate (with the most robust modulation

a node is calculated by accumulating all unit transmission times of successful frame transmissions for the entire simulation time and then dividing it by the simulation time. It is noted that the sum of the airtime ratios is less than one because the airtime does not account for the backoff time and the time consumed for unsuccessful transmissions. The *attempt probability* is measured as the probability that a node makes an attempt for frame transmission in an idle slot. From the results listed in Table 2.1, we observe the following.

- DCF assures almost an equal chance of channel access for all contending nodes, regardless of their data rate and packet size, i.e., the difference in the attempt probability among nodes is negligible. This implies that the number of packets transmitted by each node is almost the same, leading to *throughput fairness* among nodes that have the same packet size. For example, the throughput of `N1` is comparable with that of `N3`, even though the data rate of `N3` is five times higher than that of `N1`.

- Performance anomaly also occurs due to a difference in the packet size. When nodes have the same data rate, the throughput of a node is almost proportional to its packet size, e.g., the throughput of `N2` is about four times higher than that of `N1`.

- The airtime is mostly consumed by the node transmitting large packets at the low data rate (`N2`); only a small fraction of airtime is consumed by the node transmitting small packets at the high data rate (`N3`). Despite the high data rate, its advantage cannot be fully utilized under the DCF.

and coding scheme) to enhance the transmission reliability.

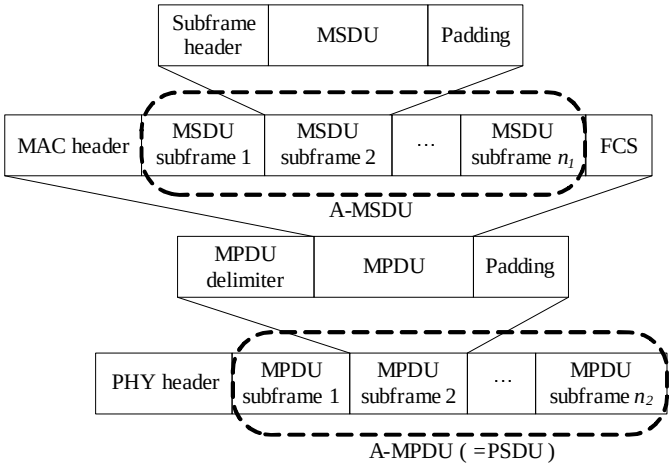


Figure 2.1 Frame structure of two-level frame aggregation scheme that combines A-MSDU with A-MPDU.

Consequently, the performance anomaly not only causes unfairness in airtime but also degrades the overall network throughput. Providing airtime fairness among nodes can be an effective solution to this problem.[3] It can be achieved by adjusting the attempt probability according to the packet size or data rate [9–13], as already addressed. This problem can also be mitigated by making use of the TXOP mechanism. In contrast to other existing approaches, we provide a novel approach to achieve airtime fairness based on the frame aggregation scheme introduced in IEEE 802.11n.

2.2.2 Frame aggregation in IEEE 802.11n

The IEEE 802.11n standard introduces two types of frame aggregation schemes, A-MSDU and A-MPDU, to enhance the MAC efficiency by reducing MAC-layer overheads. The A-MSDU scheme aggregates multiple MSDUs into a single MPDU, whereas the A-MPDU scheme allows multiple MPDUs to be aggregated

[3]The airtime fairness is equivalent to the proportional fairness [11, 20].

into a single PHY-layer service data unit (PSDU). Compared with A-MPDU, A-MSDU has less header overhead because it can transmit multiple packets with a single PHY/MAC header, whereas A-MPDU transmits them with separate MAC headers under a common PHY header. If the channel is error free, A-MSDU outperforms A-MPDU and the throughput in both cases increases as the number of aggregated packets increases. However, A-MPDU is more effective than A-MSDU when the channel is error-prone. If some of the transmitted MPDU sub-frames are corrupted due to channel error, A-MPDU can selectively retransmit only the corrupted sub-frames, which is possible thanks to block ACK and the frame check sequence (FCS) field in the individual MAC header for each sub-frame. In contrast, A-MSDU has to retransmit the entire frame to recover the corrupted sub-frames. Therefore, when the bit error rate (BER) of a channel is not negligible, A-MPDU outperforms A-MSDU. Recently, a hybrid frame aggregation scheme known as *two-level aggregation*, which combines A-MSDU and A-MPDU, has been introduced [21, 22], and its performance has been evaluated via simulations. However, the effect of this scheme on throughput or fairness is not analyzed in depth. Fig. 2.1 shows the frame structure of the two-level aggregation scheme. Let us denote n_1 and n_2 as the numbers of MSDU and MPDU sub-frames, respectively, in the two-level aggregation scheme. We can take advantage of both the A-MSDU and A-MPDU schemes by optimally determining the values of n_1 and n_2, which is the starting point of this work.

Now, we generalize several aggregation schemes in terms of n_1 and n_2 as follows:

Table 2.2 Notations and values of PHY/MAC headers and timing overheads.

Notation	Description	Value
L_{sub}	Subframe header size (for A-MSDU)	14 bytes
L_{mac}	MAC header size	34 bytes
L_{pad}	Padding size	0∼3 bytes
L_{deli}	MPDU delimiter size (for A-MPDU)	4 bytes
L_{FCS}	Frame check sequence (FCS) size	4 bytes
T_{phy}	PHY header transmission time	32μs
T_{SIFS}	Short inter-frame space (SIFS)	16 μs
T_{DIFS}	Distributed inter-frame space (DIFS)	34 μs
T_{slot}	Slot time	9 μs

1. A-MSDU ($n_1 > 1$, $n_2 = 1$): A PSDU consists of n_1 MSDU sub-frames each of which contains a packet.

2. A-MPDU ($n_1 = 1$, $n_2 > 1$): A PSDU consists of n_2 MPDU sub-frames.

3. Two-level aggregation ($n_1 > 1$, $n_2 > 1$): A PSDU consists of n_2 MPDU sub-frames, and each MPDU sub-frame consists of n_1 MSDU sub-frames.

It is noted that the total numbers of packets transmitted in a single PSDU are n_1, n_2, and $n_1 \cdot n_2$ in the cases of A-MSDU, A-MPDU, and two-level aggregation, respectively. Let us focus on the header overhead incurred during the aggregation process. Table 2.2 lists several notations of the PHY/MAC headers and timing overheads considered in this study. We define $L_{oh}(n_1, n_2)$ as the overhead to construct a single PSDU, which can be represented as

$$L_{oh}(n_1, n_2) = \begin{cases} n_1 a + b, & \text{for A-MSDU,} \\ n_2(b + c), & \text{for A-MPDU,} \\ n_1 n_2 a + n_2(b + c), & \text{for Two-level,} \end{cases} \tag{2.1}$$

where $a = L_{sub} + L_{pad}$, $b = L_{mac} + L_{FCS}$, and $c = L_{deli} + L_{pad}$. It is noted that there exist several limits on n_1 and n_2. According to IEEE 802.11n, the maximum lengths of A-MSDU and A-MPDU are 3839 (or 7935) and 65536 bytes, respectively. In addition, because of the length of the block ACK information field, at most 64 sub-frames can be aggregated in an A-MPDU frame. In the case of two-level aggregation, the length of an MPDU sub-frame should not exceed 4095 bytes for compatibility. It is a challenging problem to determine the optimal values of n_1 and n_2, because they depend on various factors such as the BER, data rate, packet size, and header overhead. In the subsequent sections, we present an objective function for the optimal operation of frame aggregation by considering airtime fairness and achievable throughput. We then provide a simple and effective method to find the optimal solution of the objective function.

2.3 Problem Statement

In this section, we first describe how the data rate is adjusted by the fast link adaptation (FLA) algorithm in IEEE 802.11n, which is closely related to how the optimal size of frame aggregation is determined. Then, we formulate an optimization problem to derive the optimal frame size.

2.3.1 Fast link adaptation in IEEE 802.11n

The FLA algorithm is usually employed in IEEE 802.11n [3]. A transmitter requests for an appropriate modulation and coding scheme (MCS) to a receiver via the MCS request sub-field. At the request, the receiver determines an appropriate MCS by considering the throughput and/or target frame error rate

and notifies it via the MCS feedback sub-field. It is difficult to estimate the block error rate (BLER) of an L_B-byte block in a MIMO-OFDM system because the signal to noise ratio (SNR) levels among the sub-carriers and/or spatial streams may differ due to frequency selectivity [23]. Therefore, a simple but accurate mapping method is needed to estimate the BLER as a function of the SNR for each sub-carrier and spatial stream. The exponential effective SNR mapping (EESM) method [24] was proposed for this purpose and can be used for FLA in IEEE 802.11n [25, 26]. It translates an instantaneous set of post-processing SNRs into a scalar link quality metric known as the effective SNR. This mapping procedure should be calibrated according to each MCS to estimate the BLER of an L_B-byte block by considering the relationship between the SNR and BLER for each MCS in an additive white Gaussian noise (AWGN) channel. Let us define γ_m^{eff} as the effective SNR with an MCS index of m, which can be calculated as in [25]. Then, the effective error rate of a PSDU consisting of several L_p-byte packets, P_e, can be expressed as

$$P_e(n_1, n_2) = 1 - \{1 - BLER(\gamma_m^{eff})\}^{\frac{L(n_1,n_2)}{L_B}}, \tag{2.2}$$

where $L(n_1, n_2)$ is

$$L(n_1, n_2) = \begin{cases} n_1(a + L_p) + b, & \text{for A-MSDU,} \\ L_p + b, & \text{for A-MPDU,} \\ n_1(a + L_p) + b, & \text{for Two-level.} \end{cases} \tag{2.3}$$

Here, $BLER(\gamma_m^{eff})$ in (2.2) is obtained by the SNR-BLER curve of the L_B-byte block in an AWGN channel. Hereafter, we assume that the receiver estimates

$P_e(n_1, n_2)$ for each MCS according to (2.2) and then selects the MCS that gives the highest data rate among the MCSs that have a $P_e(n_1, n_2)$ value below a target error rate η. From (2.2) and (2.3), we note the following points, which agree with intuition and provide insights into determining the optimal values of n_1 and n_2.

- For A-MSDU, $P_e(n_1, n_2)$ increases as n_1 increases, i.e., A-MSDU becomes vulnerable to channel error as its size increases. There is a trade-off between the header overhead and the error rate. Although a large value of n_1 contributes to increased throughput by reducing the header overhead, a more robust MCS is required to maintain an acceptable error rate, which results in the decrease of throughput.

- For A-MPDU, $P_e(n_1, n_2)$ is independent of n_2, i.e., the error rate is independent of the number of aggregated MPDU sub-frames because of the selective repeat mechanism. This comes at the cost of a larger header overhead compared to A-MSDU (see (2.1)).

2.3.2 Optimization problem for frame aggregation

From the viewpoint of airtime fairness, the basic concept of the proposed frame size adaptation scheme is to adjust the values of n_1 and n_2 such that the per-node airtime is regulated around the target reference value, T_{ref}. One of the most important issues is how to set T_{ref} properly. We discuss this later in Section 2.4.3.

Another objective of the proposed scheme is to maximize throughput. Now, let us derive the throughput that can be obtained with the two-level aggrega-

tion scheme under the constraint of airtime fairness. We only focus on the generalized two-level aggregation scheme because the single-level A-MSDU and A-MPDU schemes can be approximately considered as special cases of two-level aggregation. We emphasize that our objective is to find the optimal values of n_1 and n_2, and not to derive an accurate analysis model of aggregate throughput. For this purpose, we only focus on the case where a node gets a chance to access the channel after competing with other nodes and tries to maximize its own throughput by adjusting the values of n_1 and n_2. In the proposed scheme, the optimal values of n_1 and n_2 are determined for each node in a distributed manner, depending on its data rate and packet size. Moreover, the proposed scheme is completely decoupled from the contention window control that deals with collisions among nodes[4] and adopts the standard binary exponential backoff mechanism to resolve collisions without modifying the channel access mechanism. We define $T_{data}(n_1, n_2)$ as the time required to transmit a single aggregate PHY-layer protocol data unit (PPDU) consisting of n_2 MPDU sub-frames and n_1 MSDU sub-frames per MPDU sub-frame. Then, $T_{data}(n_1, n_2)$ is represented as

$$T_{data}(n_1, n_2) = T_{phy} + \frac{8(n_1 n_2 L_p + L_{oh}(n_1, n_2))}{R_m}, \tag{2.4}$$

where T_{phy} is the time required to transmit the PHY header including the preamble, and R_m is the data rate with MCS index m, which is selected by the FLA algorithm at the receiver. Under the condition that there is no trans-

[4]According to the result in [15], the optimal size of an A-MSDU that maximizes the aggregate throughput is mainly affected by the BER but hardly affected by the number of contending nodes.

mission failure due to channel error, the maximum throughput $S(n_1, n_2)$ is expressed as

$$S(n_1, n_2) = \frac{8 n_1 n_2 L_p}{T_{data}(n_1, n_2) + T_{oh}}, \tag{2.5}$$

where $T_{oh} = T_{DIFS} + T_{BO} + T_{SIFS} + T_{back}$. Here, T_{BO} and T_{back} are the average backoff time[5] and block ACK transmission time, respectively.

To find the appropriate values for n_1 and n_2, we formulate the optimization problem as follows.

$$\begin{aligned} (n_1^*, n_2^*) &= \arg \max_{(n_1, n_2) \in \mathbb{N}^2} S(n_1, n_2) \left(1 - P_e(n_1, n_2)\right), \\ \text{subject to } & |T_{ref} - T_{data}(n_1, n_2)| < T_\epsilon, \\ & n_1 L_p + L_{oh}(n_1, 1) \leq L_{max}^{msdu}, \\ & n_1 n_2 L_p + L_{oh}(n_1, n_2) \leq L_{max}^{mpdu}, \\ & n_2 \leq N_{max}^{mpdu}, \end{aligned} \tag{2.6}$$

Here, T_ϵ is introduced to allow some tolerance between $T_{data}(n_1, n_2)$ and T_{ref} because it is difficult to make $T_{data}(n_1, n_2)$ exactly equal to T_{ref} with integer values of n_1 and n_2. Also, in (2.6), L_{max}^{msdu} (= 3839 bytes) and L_{max}^{mpdu} (= 65536 bytes) are the maximum sizes of A-MSDU and A-MPDU frames respectively, and N_{max}^{mpdu} (= 64) is the maximum number of MPDU sub-frames. These three constraints are imposed by the IEEE 802.11n standard, whereas the first one stems from airtime fairness. The problem in (2.6) is an integer programming

[5]The average backoff time is approximated as $T_{BO} = T_{slot}\ (CW_{min} - 1)/2$, where CW_{min} is the minimum contention window size to achieve the maximum throughput. The underlying rationale is that the maximum throughput depends mainly on the collision probability, but little on the values of n_1 and n_2 [15].

problem, which is more difficult to solve than the case when the variables are real numbers. Each transmitter has to solve this optimization problem every time before constructing an aggregated frame, which may be too time-consuming to solve the problem in real-time. In addition, the transmitter cannot accurately estimate the frame error rate $P_e(n_1, n_2)$ for the currently transmitting frame. These difficulties demand a more efficient method of finding the optimal solution, (n_1^*, n_2^*), which will be provided in the subsequent section.

2.4 Adaptive Frame Aggregation for Fairness and Efficiency

In this section, we provide a simple and effective method to determine n_1^* and n_2^*. We first transform the optimization problem into an unconstrained, single-variable optimization problem, and then provide a heuristic solution, which will be used to adjust the frame size.

2.4.1 Unconstrained single-variable optimization problem

To simplify the optimization problem, we replace n_1 and n_2 with x_1 and x_2, respectively, which are positive real values without upper bounds so that $T_{data}(x_1, x_2)$ can be made exactly equal to T_{ref}. Then, we can release all the constraints and replace the objective function in (2.6) with

$$(x_1^*, x_2^*) = \arg \max_{(x_1,x_2)\in\ \mathbb{R}^2} Kx_1x_2\left(1 - P_e(x_1, x_2)\right), \tag{2.7}$$

where constant K is

$$K = \frac{8L_p}{T_{ref} + T_{oh}},$$

because $T_{data}(x_1, x_2) = T_{ref}$. We further simplify (2.7) by expressing x_2 in terms of x_1. By solving the equation $T_{data}(x_1, x_2) = T_{ref}$ with respect to x_1 and x_2, x_2 becomes

$$x_2 = \frac{C}{Ax_1 + B}, \tag{2.8}$$

where $A = a + L_p$, $B = b + c$, and $C = R_m(T_{ref} - T_{phy})/8$. It is noted that x_2 is mapped to a unique x_1, as shown in (2.8).

Next, we consider the frame error rate. As the transmitter is only informed of the appropriate MCS index m for the previously transmitted frame, it cannot readily predict the frame error rate for the current frame. Let us define (x_1', x_2') as the numbers of A-MSDU and A-MPDU sub-frames used in the aggregation for the previously transmitted frame and define $P_e(x_1', x_2')$ as its frame error rate. To simplify the problem, we make the following assumptions: (i) the frame transmission time is sufficiently shorter than the channel coherence time so that the channel condition does not change significantly between two consecutive frames; (ii) the FLA algorithm works properly so that the error rate is slightly lower than the target error rate η ($<<1$)[6]. With these assumptions,

[6]A node may suffer from deep fading, and thus, the frame error rate may exceed the target value even though the node uses the most robust MCS. However, this happens rarely.

we approximate $P_e(x_1', x_2') \approx \eta$ and relate $P_e(x_1, x_2)$ with $P_e(x_1', x_2')$ as

$$\begin{aligned} P_e(x_1, x_2) &= 1 - \{1 - P_e(x_1', x_2')\}^{\frac{L(x_1, x_2)}{L(x_1', x_2')}} \\ &\approx \frac{L(x_1, x_2)}{L(x_1', x_2')}\eta, \end{aligned} \tag{2.9}$$

because $\eta \ll 1$. Then, $(1 - P_e(x_1, x_2))$ is represented as

$$\begin{aligned} 1 - P_e(x_1, x_2) &= 1 - \frac{x_1(a + L_p) + b}{L(x_1', x_2')}\eta \\ &= D - A\eta' x_1, \end{aligned} \tag{2.10}$$

where $\eta' = \eta / L(x_1', x_2')$ and $D = 1 - b\eta'$.

Through the above procedure, we can convert the original problem in (2.6) to the following unconstrained single-variable optimization problem:

$$x_1^* = \arg\max_{x_1 \in \mathbb{R}} \; KCf(x_1), \tag{2.11}$$

where

$$f(x_1) = \frac{Dx_1 - A\eta' x_1^2}{Ax_1 + B}. \tag{2.12}$$

$f'(x_1)$ is readily shown to be a strictly decreasing function of x_1, i.e., $f(x_1)$ is concave. This confirms that a unique x_1^* exists. We can find x_1^* from the derivative of $f(x_1)$ and x_2^* from (2.8), as

$$x_1^* = \frac{B}{A}\left(\sqrt{1 + \frac{D}{B\eta'}} - 1\right), \quad x_2^* = \frac{C}{Ax_1^* + B}. \tag{2.13}$$

2.4.2 Heuristic solution for frame size adaptation

Based on the results of x_1^* and x_2^* given in (2.13), we find the feasible set of (n_1, n_2) that are positive integer values. We assume that a node has a sufficient number of packets to aggregate. We first determine n_1 as the integer closest to n_1^* and then determine n_2 such that $T_{data}(n_1, n_2)$ becomes as close as possible to T_{ref}, i.e.,

$$n_1 = [x_1^*], \quad n_2 = \left[\frac{C}{An_1 + B}\right], \tag{2.14}$$

where $[x]$ is the integer closest to x. In (2.14), n_2 is determined by using (2.8) instead of $[x_2^*]$ to minimize the difference between $T_{data}(n_1, n_2)$ and T_{ref}, which is due to the rounding error. However, the set of (n_1, n_2) obtained in this way cannot completely be free of rounding errors.

To compensate for the deviation of $T_{data}(n_1, n_2)$ from T_{ref}, we alternatingly use two sets of (n_1, n_2), denoted as (n_1^u, n_2^u) and (n_1^l, n_2^l), each of which makes the actual value of T_{data} comparable to T_{ref} but slightly larger and smaller than T_{ref}, respectively. We obtain these two sets as follows. If the set of (n_1, n_2) obtained from (2.14) results in $T_{data}(n_1, n_2) > T_{ref}$, then we set $(n_1^u, n_2^u) = (n_1, n_2)$. To find (n_1^l, n_2^l), we increase n_1^l by one from n_1^u while recalculating n_2^l as

$$n_1^l = n_1^u + 1, \quad n_2^l = \left[\frac{C}{An_1^l + B}\right]. \tag{2.15}$$

Recall that the header overhead is smaller for the MSDU sub-frame than for the MPDU sub-frame. These changes in n_1 and n_2 (i.e., the increase in n_1 and

Table 2.3 Comparison between proposed method and total enumeration method ($L_p = 500$ bytes and $T_{ref} = 3$ ms).

	Proposed				Total enumeration				Proposed				Total enumeration			
Data rate (Mb/s)	n_1^u	n_2^u	T_{data} (ms)	thput (Mb/s)	n_1^u	n_2^u	T_{data} (ms)	thput (Mb/s)	n_1^l	n_2^l	T_{data} (ms)	thput (Mb/s)	n_1^l	n_2^l	T_{data} (ms)	thput (Mb/s)
6.5	3	2	3.97	5.83	5	1	3.28	5.83	4	1	2.64	5.78	1	4	2.82	5.59
13	2	5	3.36	11.57	1	9	3.16	11.20	3	3	2.99	11.61	3	3	2.99	11.68
26	2	9	3.03	23.10	2	9	3.03	23.20	3	6	2.99	23.21	3	6	2.99	23.35
39	2	14	3.14	34.68	4	7	3.07	35.04	3	9	2.99	34.82	3	9	2.99	35.03
52	2	18	3.03	46.21	2	18	3.03	46.39	3	12	2.99	46.43	3	12	2.99	46.71
65	2	23	3.10	57.78	2	23	3.10	58.01	3	15	2.99	58.03	3	15	2.99	58.39
78	2	27	3.03	69.31	4	14	3.07	70.07	3	18	2.99	69.64	3	18	2.99	70.06
104	2	36	3.03	92.41	2	36	3.03	93.11	3	24	2.99	92.85	3	24	2.99	93.42
117	2	41	3.07	103.99	3	28	3.10	105.16	3	27	2.99	104.46	3	27	2.99	105.10
130	2	45	3.03	115.52	3	31	3.09	116.84	3	30	2.99	116.07	3	30	2.99	116.77

the decrease in n_2) reduce the total header overhead and decrease $T_{data}(n_1^l, n_2^l)$. If $T_{data}(n_1^l, n_2^l)$ remain larger than T_{ref} even with the set in (2.15), we keep increasing n_1^l by one at a time until $T_{data}(n_1^l, n_2^l)$ becomes less than T_{ref}. Conversely, if the set of (n_1, n_2) obtained from (2.14) makes $T_{data}(n_1, n_2)$ smaller than T_{ref}, we set $(n_1^l, n_2^l) = (n_1, n_2)$. In a similar way, we obtain (n_1^u, n_2^u) by decreasing n_1^u by one from n_1^l and recalculating n_2^u as

$$n_1^u = n_1^l - 1, \quad n_2^u = \left\lfloor \frac{C}{An_1^u + B} \right\rfloor, \tag{2.16}$$

Once a transmitter selects (n_1^u, n_2^u) and (n_1^l, n_2^l), it performs frame aggregation with either (n_1^u, n_2^u) or (n_1^l, n_2^l). Here, we introduce a weighting factor w $(0 < w < 1)$ to alternate between these two sets. When a transmitter gets a chance to access a channel (i.e., the backoff counter reaches zero), it generates a random number X uniformly distributed between zero and one. If X is less than w, the transmitter performs frame aggregation with $(n_1, n_2) = (n_1^u, n_2^u)$. Otherwise, it performs frame aggregation with $(n_1, n_2) = (n_1^l, n_2^l)$. To make the average value of T_{data} as close as possible to T_{ref}, we set the value of w such that

$$T_{ref} = wT_{data}(n_1^u, n_2^u) + (1 - w)T_{data}(n_1^l, n_2^l), \tag{2.17}$$

i.e.,

$$w = \frac{T_{ref} - T_{data}(n_1^l, n_2^l)}{T_{data}(n_1^u, n_2^u) - T_{data}(n_1^l, n_2^l)}. \tag{2.18}$$

To validate the effectiveness of this method, we compared the results ob-

tained by the proposed and total enumeration methods via simulations. In the total enumeration method, (n_1^u, n_2^u) and (n_1^l, n_2^l) were obtained by solving the original optimization problem in (2.6) under the assumption that the transmitter knows the actual value of $P_e(n_1, n_2)$.[7] We first obtained the candidate sets of (n_1, n_2) that satisfy the four constraints in (2.6). Here, T_ϵ was initially set to 0.1 ms. If there was no candidate set, we relaxed the constraint on T_ϵ by doubling its value until at least one set of (n_1, n_2) satisfied the constraints. We then calculated the objective function for all candidate sets of (n_1, n_2) and determined the optimal set that maximizes the objective function. This method gives an ideal solution that can be used to evaluate the effectiveness of the proposed method.

Table 2.3 lists the sets of (n_1^u, n_2^u) and (n_1^l, n_2^l) obtained by these two methods, along with T_{data} and throughput (indicated as *thput*). We set the packet size L_p as 500 bytes and T_{ref} as 3 ms. The results in Table 2.3 show that particularly when the data rate is low, (n_1^u, n_2^u) and (n_1^l, n_2^l) obtained by the proposed method somewhat differ from those obtained by the total enumeration method, and the difference between T_{data} and T_{ref} is greater with the proposed method than with the total enumeration method. The reason for this difference is that, when the data rate is low, the number of candidate sets of (n_1, n_2) is small, and thus, the rounding error and approximation error become relatively large. In some low data rate cases, however, the throughput is larger for the proposed method than for the total enumeration method, at the expense of a larger deviation of T_{data} from T_{ref}. Recall that such a deviation can be com-

[7] (n_1^u, n_2^u) and (n_1^l, n_2^l) are the same as the optimal values of (n_1^*, n_2^*) obtained by solving (2.6) with the first constraint $(T_{data}(n_1^u, n_2^u) - T_{ref}) < T_\epsilon$ and $(T_{ref} - T_{data}(n_1^l, n_2^l)) < T_\epsilon$, respectively.

pensated for by alternatingly adopting (n_1^u, n_2^u) and (n_1^l, n_2^l) when constructing aggregated frames; this will be confirmed via simulations in the next section. Despite these differences between the proposed and total enumeration methods, the throughput difference between the two methods is negligible for most data rates. Therefore, we conclude that the proposed method is suitable for adjusting frame size, as it yields performance comparable to that of the total enumeration method with reduced computational time.

2.4.3 Comments on several issues

We end this section by discussing the following issues.

Guidelines to set T_{ref}

When assuring airtime fairness, T_{ref} is used as the target reference value. If this value is too small, only a few packets can be aggregated. This decreases the possibility of reducing the header and timing overheads, which results in the loss of opportunities to enhance the throughput. However, a large value of T_{ref} may have several drawbacks. We need to carefully consider the following points in setting T_{ref}.

- **Link adaptation and channel coherence time**: If T_{ref} exceeds the channel coherence time, the link adaptation algorithm may work poorly. In this case, the estimated channel quality may not hold during the entire transmission time of an aggregated frame, and the transmission of subframes may fail frequently. Consequently, T_{ref} should be set sufficiently shorter than the channel coherence time,[8] so that the channel condition

[8] Although a receiver may estimate the channel coherence time and report it to a transmitter

is kept nearly constant during T_{ref}.

- **Delay and short-term fairness**: The value of T_{ref} affects short-term fairness, as well as delays in channel access, queuing, and retransmission. As T_{ref} increases, a node occupies the shared channel for a longer time, during which other nodes have to defer channel access. Thus, the increase of T_{ref} increases the channel access delay between two consecutive aggregated frames as well as the queuing delay and jitter, which are undesirable for real-time service or streaming service. Moreover, the increase of T_{ref} also increases the retransmission delay for corrupted sub-frames because they can not be retransmitted until the node occupies the channel again. All of these effects contribute to the degradation of short-term fairness among nodes.

- **Feasibility of frame aggregation**: The value of T_{ref} is related with the feasibility of n_1 and n_2. If T_{ref} is large, many packets needs to be aggregated to assure airtime fairness. However, because of limits on n_1 and n_2, excessively large values of T_{ref} make n_1 and n_2 noncompliant with this constraint.

Taking all of these effects into consideration, we suggest that the proper range of T_{ref} is on the order of several milliseconds, which agrees well with the default value of TXOP in IEEE 802.11e.

periodically [27] this possibility is beyond the scope of this study. We consider that the coherence time of a typical WLAN channel is on the order of a few tens of milliseconds [28].

Compliance with the standard

The values of n_1 and n_2 obtained from (2.14) should satisfy the constraints of L_{max}^{msdu}, L_{max}^{mpdu}, and N_{max}^{mpdu} imposed by the IEEE 802.11n standard. Let us define the maximum numbers of n_1 and n_2 that comply with these limits as N_1^{max} and N_2^{max}, respectively. We need to deal with the case where $x_1^* > N_1^{max}$ and/or $x_2^* > N_2^{max}$.

- **Case 1**: $x_1^* > N_1^{max}$ and $x_2^* < N_2^{max}$

  In this case, we set

  $$n_1 = N_1^{max}, \quad n_2 = \left[\frac{C}{AN_1^{max} + B} \right].$$

  This configuration may incur more header overhead due to the decrease in n_1 and increase in n_2. On the other hand, the decrease in n_1 contributes to reduce the frame error rate.

- **Case 2**: $x_2^* > N_2^{max}$ and $x_1^* < N_1^{max}$

  As in the first case, we set

  $$n_2 = N_2^{max}, \quad n_1 = \left[\frac{C}{AN_2^{max}} - \frac{B}{A} \right].$$

  In this case, n_2 is decreased from n_2^* while n_1 is increased from n_1^*. Contrary to the first case, the header overhead decreases whereas the error rate increases.

- **Case 3**: $x_1^* > N_2^{max}$ and $x_2^* > N_2^{max}$

  In this case, there is no choice but to set $n_1 = N_1^{max}$ and $n_2 = N_2^{max}$.

These three cases mainly occur when either T_{ref} is very large or the packet size is small but the data rate is high. As long as we carefully determine T_{ref} according to the guidelines above, we can avoid such cases.

Comparison with conventional A-MSDU and A-MPDU

The *two-level* aggregation scheme is a generalized scheme combining A-MSDU and A-MPDU. Therefore, it can replace the *single-level* A-MSDU and A-MPDU by setting $n_2 = 1$ and $n_1 = 1$, respectively. However, it is not desirable to apply the proposed scheme to the conventional single-level A-MSDU scheme because the gain will be marginal due to the strict limit on L_{max}^{msdu} (3896 bytes) and the absence of a selective repeat mechanism. In addition, the gain of the two-level aggregation scheme will diminish if applied to the conventional A-MPDU scheme, because (*i*) an MPDU sub-frame has a larger header overhead than an MSDU sub-frame and (*ii*) the A-MPDU scheme supports aggregation of up to 64 packets, which is the maximum limit in a block ACK, whereas the two-level aggregation scheme can aggregate several packets per MPDU sub-frame. The simulation results in the next section compare the two-level aggregation scheme with the conventional single-level A-MPDU scheme.

Dealing with unsaturated cases

Thus far, we have assumed that each node has a sufficient number of packets in the transmission queue so that it can arbitrarily adjust the number of packets when constructing an aggregated frame. However, we need to consider the case of unsaturated traffic condition where the number of backlogged packets for

frame aggregation (denoted as N_{pkt}) is insufficient. The proposed scheme can easily cope with this case. If $N_{pkt} < x_1^* \cdot x_2^*$, the node performs frame aggregation with n_1 and n_2 such that

$$n_1 = \min(N_{pkt}, \lceil x_1^* \rceil), \quad n_2 = \lfloor N_{pkt} / n_1 \rfloor, \tag{2.19}$$

where $\lfloor x \rfloor$ is the largest integer that does not exceed x. This can avoid unnecessary delay resulting from the waiting time of packet arrivals for optimal frame aggregation, but it makes T_{data} less than T_{ref}. The unused portion of airtime by unsaturated nodes can be well used by other saturated nodes. Therefore, this policy does not debase the overall network efficiency and maintains airtime fairness among the saturated nodes.

2.5 Performance Evaluation

2.5.1 Simulation configuration

In this section, we evaluate the performance of the proposed frame size adaptation scheme via simulations by using MATLAB and compare its performance with several existing schemes. We implemented MAC/PHY layers according to the system model of IEEE 802.11n operating at the 5 GHz frequency band. The system parameters related to IEEE 802.11n can be found in Table 2.2 and [3]. TGn channel models implemented by MATLAB [29, 30] were used to generate MIMO channel matrices for the models defined in [28]. Simulations were conducted for channel model B, which models residential homes or small offices. The velocity of the moving environment was assumed to be 1.2 km/h, which corresponds to a coherence time of approximately 15 ms. The transmit and the

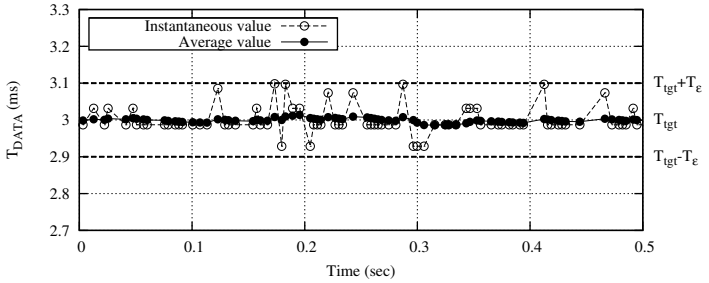
background noise powers were set to 15 dBm and -87 dBm, respectively. FLA was employed as a link adaptation algorithm, and the target error rate η was set to 0.01. The data rate was in the range of 6.5 to 130 Mb/s with a 2×2 MIMO configuration for a channel bandwidth of 20 MHz and a guard interval of 800 ns. User datagram protocol (UDP) packets were generated by constant bit rate (CBR) sources so that the transmitter always had a sufficient number of packets for frame aggregation. The performances were evaluated and compared for the following four schemes.

- `DCF`: This is a baseline scheme that only employs DCF. It suffers from the performance anomaly problem.

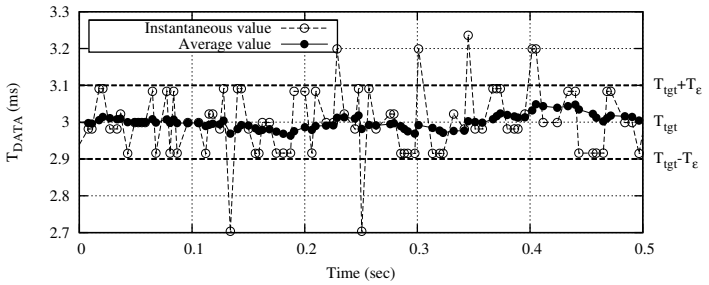
- `STFMAC` [13]: This scheme adjusts the channel access probability of the DCF to achieve airtime fairness and to maximize the aggregate throughput. It has the same objective as the scheme proposed in this chapter.

- `TXOP`: This scheme uses the TXOP mechanism, with which airtime fairness can be roughly achieved. A node can transmit multiple back-to-back packets within the TXOP limit. The TXOP duration is set equal to the value of T_{ref} (= 3 ms) for a fair comparison with the proposed scheme.

- `FA`: This is the proposed frame size aggregation scheme. We consider two versions of this scheme based on the proposed and total enumeration methods (Section 2.4.2), which are denoted as `FA-H` and `FA-T`, respectively.



(a) `N1` (L_p =500 bytes, 30 m away from AP).



(b) `N2` (L_p =1000 bytes, 60 m away from AP).

Figure 2.2 Variation of T_{data} in the proposed scheme.

2.5.2 Validation of airtime fairness

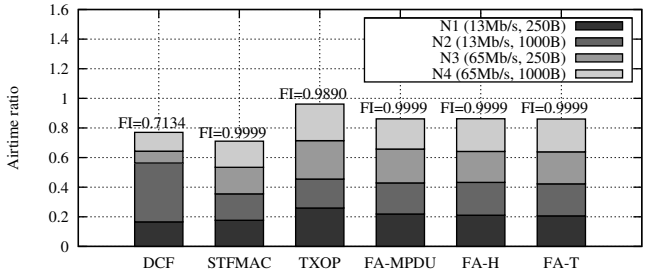
We first verify that the proposed scheme adaptively controls the size of the aggregated frame to assure airtime fairness. The simulation was performed under the network configuration where two nodes (`N1` and `N2`) were respectively located 30 and 60 m away from an AP and transmitted data frames to the AP. `N1` had 500 byte data packets, whereas `N2` had 1000 byte data packets. Each node suffered different fading environment and adjusted its data rate by using the FLA algorithm. It is noted that this configuration allows `N1` to have a higher data rate than `N2`.

Fig. 2.2 shows how `FA-H` adjusts T_{data} for these two nodes. The instantaneous values indicate the actual T_{data} at each transmission instant, whereas

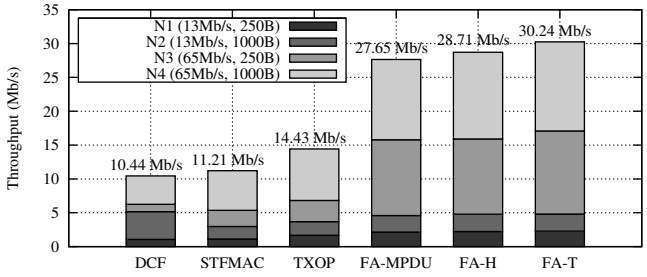
the average values are obtained by a weighted moving average. In the case of N1, the instantaneous value of T_{data} is more tightly regulated around T_{ref} than in the case of N2, because FA-H provides better granularity of control for N1 because of the smaller packet size and higher data rate. However, the average value of T_{data} can be made very close to T_{ref} for both nodes because of the compensation mechanism with two sets of (n_1^u, n_2^u) and (n_1^l, n_2^l) (see (2.17)). These simulation results confirm that the proposed scheme can maintain the average airtime of each node close to T_{ref}, regardless of the data rate and packet size.

2.5.3 Solution to the performance anomaly

This simulation is designed to see how various schemes resolve the performance anomaly of DCF. We performed simulations under the same condition described in Section 2.2.1. We also considered the case where the proposed scheme is implemented with the conventional A-MPDU scheme, to compare the performance of two-level aggregation with that of single-level A-MPDU (denoted as FA-MPDU). In the case of FA-MPDU, a transmitter finds n_2^u and n_2^l by the total enumeration method while fixing n_1 to 1, and the limitation on n_2 is removed for the purpose of comparison. Figs. 2.3(a) and 2.3(b) show the per-node airtime ratio and throughput for various schemes, respectively. All schemes except for the DCF provide good airtime fairness, as shown Fig. 2.3(a); hence, they are free of the performance anomaly problem. Jain′s fairness index (FI) is calculated with the airtime ratio and indicated in Fig. 2.3(a). The FI values in the cases of STFMAC, FA-MPDU, FA-H, and FA-T are very close to the ideal value; however, the FI value of TXOP is slightly smaller than that of the others, except for DCF. However, in the case of TXOP, the sum of the airtime ratios is slightly greater



(a) Airtime ratio



(b) Throughput

Figure 2.3 Performance comparison of several mechanisms to resolve performance anomaly of DCF.

than that of the other schemes. This is attributed to the property of the TXOP mechanism in response to collision; if the transmission of the first frame fails due to collision (recall that the channel is error-free in this simulation), transmission of the pending packets within the remaining TXOP time is prohibited. Therefore, only the first frame will suffer from transmission failure, and the remaining TXOP time can be used by other nodes. However, in our schemes (`FA-MPDU`, `FA-H`, `FA-T`), the collision may last during the entire transmission of an aggregated frame, which is inevitable because the proposed scheme does not modify DCF at all. In addition, there are no notable differences in the FI and the per-node airtime ratio among `FA-MPDU`, `FA-H`, and `FA-T`.

Next, we focus on the throughput shown in Fig. 2.3(b), where the aggregate throughput is indicated along with per-node throughput. Three schemes, FA-MPDU, FA-H, and FA-T, remarkably outperform the others. These superior performances mainly stem from the frame aggregation that reduces several MAC overheads, e.g., contention time, header overhead, inter-frame space time, and ACK transmission time. Although FA-MPDU gives a throughput performance comparable to that of FA-H and FA-T, its throughput is smaller than that of FA-H and FA-T by approximately 4% and 9%, respectively. This confirms the advantage of two-level aggregation over a single-level aggregation. The DCF shows the worst throughput performance due to the performance anomaly problem. In the case of STFMAC, its throughput gain over DCF is small even though it greatly improves airtime fairness. This is because STFMAC controls the channel access probability in an optimal way, but cannot take advantage of frame aggregation. It is interesting to note that the aggregate throughput of TXOP is quite smaller than that of the other frame aggregation schemes, although its total airtime ratio is slightly higher. This can be explained as follows. The airtime is defined to include the inter-frame space time, ACK transmission time, and MAC header transmission time, which do not make any contribution to improve throughput. Although TXOP transmits several frames within the TXOP limit, each frame has its own PHY and MAC headers, and a corresponding ACK follows instead of a block ACK. As a result, TXOP has an aggregate throughput of at most half that of FA-H or FA-T.

2.5.4 Performance comparison

Here, we compare the performance of several schemes in terms of throughput, fairness, and delay. The simulations were carried out under a realistic heterogeneous network configuration where N_n (ranging from 1 to 30) nodes were randomly located within a circle area of 60 m radius and transmitted data frames to an AP located in the center of the circle. The packet size of each node was randomly selected between 200 and 1000 bytes. Figs. 2.4(a), 2.4(b), and 2.4(c) represent the aggregate throughput, Jain′s FI based on the per-node airtime ratio, and the average MAC-to-MAC delay[9], respectively. As shown in these figures, all of the performance indices of `FA-H` are nearly equal to those of `FA-T` for the entire range of N_n, implying that the proposed method with reduced computational complexity is highly effective. Hereafter, we only refer to the performance of `FA-H`.

First, we can see from Fig. 2.4(a) that the throughput gain of `STFMAC` over `DCF` is only 5∼10% and that of `TXOP` over `STFMAC` is approximately 50∼80%. The throughput enhancements by the `FA-H` over `STFMAC` and `TXOP` reach up to 190% and 70%, respectively. As N_n increases, the aggregate throughput decreases for all schemes, due to the increase in collision probability. However, it is important to note that, in the case of `FA-H`, although the absolute value of throughput decreases, its relative gain over that of the other schemes is maintained nearly constant.

In terms of airtime fairness, with the exception of `DCF`, the FI values are

[9]We define the *MAC-to-MAC delay* as the time interval between the instant when a packet is delivered to the MAC interface queue in a transmitter and the instant when it is successfully received by a receiver and delivered to the upper layer from the MAC layer.

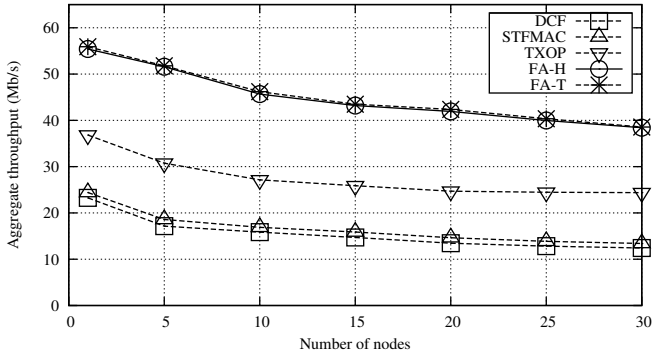
nearly close to one regardless of N_n. However, the fairness of DCF deteriorates as N_n increases, i.e., it decreases from 0.90 to 0.82 as N_n increases from 5 to 30.

Next, we compare the delay performance in Fig. 2.4(c). The average delay of FA-H also outperforms that of the other schemes; it is smaller than that of STFMAC and TXOP by approximately 62∼68% and 35∼60%, respectively. FA-H can transmit multiple packets in an aggregated frame, and thus, the overhead time per packet can be effectively reduced. Consequently, FA-H can shorten the delay considerably. As N_n increases, the delay increases for all schemes due to the contention among nodes; however, the relative gain of FA-H over the others does not decrease. This is another important advantage of the proposed aggregation scheme.

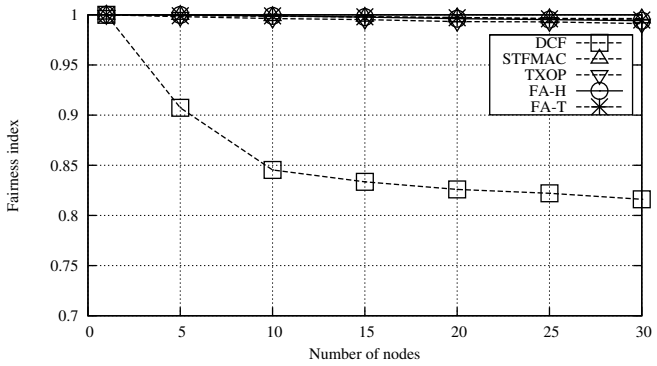
2.6 Chapter Summary

In this chapter, we proposed the FA scheme by employing the generalized two-level frame aggregation scheme in IEEE 802.11n WLANs, to achieve airtime fairness and enhance overall network throughput. The FA scheme tightly regulates the airtime of each node around a target time and thus effectively resolves the performance anomaly problem that arises due to different data rates and packet sizes among nodes. At the same time, the FA scheme adjusts the frame size so that the throughput can be maximized by considering the channel error rate, data rate, and MAC overhead. We formulated the problem of frame aggregation as an optimization problem and proposed a simple and effective method to reduce the computational complexity. The simulation results confirm that

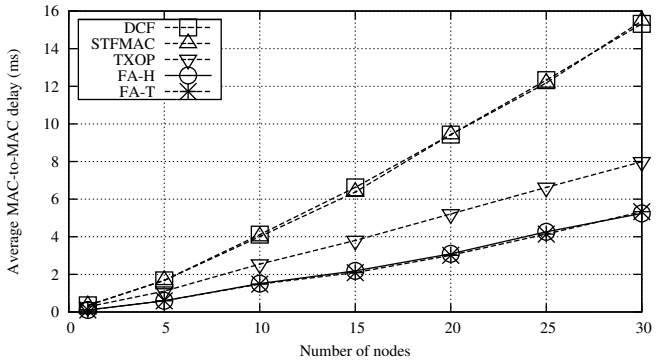
the FA scheme significantly outperforms other existing schemes in terms of the aggregate throughput and channel access delay while maintaining a high level of airtime fairness.



(a) Aggregate throughput



(b) Airtime fairness (Jain′s fairness index)



(c) Average MAC-to-MAC delay

Figure 2.4 Performance comparison of several schemes in the heterogeneous network configuration.

Chapter 3

Hidden Node Detection Mechanism

3.1 Background and Related Work

The optional RTS/CTS mechanism in the IEEE 802.11 standard [5] and its extensions [31–33] are possible solutions to alleviate the hidden node problem. However, these solutions incur protocol overhead that may lead to throughput degradation when there is no hidden node. Therefore, it is necessary to develop a detection mechanism for hidden nodes that can be used to trigger a resolution mechanism for such nodes. Detecting hidden nodes depends on the capability of a transmitter to differentiate its frame losses according to the cause. If the transmitter identifies the cause of the frame losses, it can then select an appropriate countermeasure for the next transmission. There are three main causes of frame losses as described below.

1) **Collisions (Synchronous interference)** : interference signals due to frame transmissions by other nodes within the carrier sensing range of a transmitter.

2) **Hidden nodes (Asynchronous interference)** : interference signals due to frame transmissions by other nodes out of the carrier sensing range of a transmitter.

3) **Channel impairments** : noise, fading, and distortion, etc., which result in channel errors.

One way to overcome a collision is to double the CW size so that the probability of a collision is reduced. The hidden node problem can be alleviated by exchanging RTS/CTS frames before the transmission of a data frame. When channel impairments affect frame transmission, an appropriate MCS can be selected according to the channel condition, which is known as link adaptation [34–40]. However, a legacy 802.11 transmitter cannot exactly determine the causes of frame losses because the success of a frame transmission is determined solely by the successful reception of the corresponding ACK frame.

Differentiating a collision from a channel error is one of the crucial issues to consider when developing a link adaptation scheme. A straightforward solution is to have a receiver to report the channel information to the transmitter so as to adjust the MCS. In RBAR [34], the authors proposed a feedback mechanism at a receiver to report an appropriate MCS to the transmitter using the RTS/CTS exchange. There is no need for the RBAR transmitter to differentiate a collision from a channel error given that it can adjust the MCS according to the report by the receiver. RBAR, which always uses the RTS/CTS mechanism,

can achieve good performance even in the presence of hidden nodes. However, the RTS/CTS exchange may incur large overhead when there is no hidden node, and the IEEE 802.11a/b/g standard does not support a feedback mechanism via the RTS/CTS exchange. Therefore, several link adaptation schemes have been developed in an effort to differentiate a collision from a channel error by the transmitter itself. The most popular and widely deployed link adaptation scheme, ARF [35], is used as the baseline scheme in several studies [36–38] due to the simplicity of its implementation. ARF, however, will experience significant performance degradation when the number of contending nodes increases because it ignores frame losses due to collisions [41]. The RTS/CTS option can be used with ARF to mitigate the effect of a collision, but it imposes a considerable overhead. To reduce the overhead incurred by RTS/CTS, CARA [36] and RRAA [37] adaptively use the RTS/CTS exchange. RTS probing, in which a transmitter transmits an RTS frame only after a frame loss, was originally proposed in CARA. RRAA improves the RTS probing of CARA by transmitting RTS frames more aggressively based on a parameter which is adjusted adaptively to the network condition. CARA and RRAA can achieve good performance, even in the presence of hidden nodes, because they turn on the RTS/CTS option when frame transmissions fail due to hidden nodes. However, there is still room for improvement because they use RTS/CTS whenever there is a frame loss. There are some link adaptation schemes that differentiate a collision from a channel error without the help of the RTS/CTS mechanism. The scheme known as Adaptive Thresholds [38] employs dynamic up/down thresholds for ARF by offsetting the collision probability. In [39], the negative ACK (NAK) frame is introduced to differentiate a collision from a channel

error, but it requires modification of the standard. RA-MCE [40] estimates collision and channel error probabilities using MAC layer statistics. However, the performance of these scheme can slip in the presence of hidden nodes because they do not take the impact of hidden nodes into consideration. They need to use the RTS/CTS exchange adaptively to prevent performance degradation due to hidden nodes.

On the other hand, hidden node detection mechanisms have been proposed by several authors [42–44]. In [42], a transmitter detects hidden nodes when an ACK frame is observed without observing the corresponding data transmission. It can be only applied in a single BSS network and does not work properly in general network topologies. Nishide et al. proposed a hidden node detection mechanism which implements an AP cooperation system in a densely deployed network [43]. However, it is difficult to implement, because most APs are installed without planning in advance. In other research [44], hidden nodes are detected by observing the relative timing between a transmitted signal and an interference signal. Because this method does not consider the effect of channel impairments, it may not operate properly when frame losses occur due to any mixture of the three causes presented above.

In this chapter, we propose a novel hidden node detection (HD) mechanism while considering the three main causes of frame losses with the help of the new features in IEEE 802.11n [3]. The new enhanced features in the MAC layer are frame aggregation and block ACK, which are designed to reduce the overhead such as PHY headers and the contention time. In contrast to the IEEE 802.11a/b/g standards, a receiver transmits a block ACK frame even when some parts of the data frame are corrupted. Fast link adaptation (FLA)

can also be employed in IEEE 802.11n. The basic idea of FLA is identical to that used in RBAR except for the usage of the RTS/CTS option. A FLA receiver estimates the channel condition during the reception of the PHY header and selects an appropriate MCS by considering the throughput and target frame error rate (FER). The selected MCS is reported back to the transmitter using the reserved field in a block ACK frame, and the transmitter adjusts its MCS for the next transmission. Link adaptation schemes based on ARF are simple to implement, but they should be modified if they are to be compatible to the IEEE 802.11n system due to the block ACK. Therefore, we choose FLA as the baseline link adaptation scheme for the proposed HD mechanism.

Frame losses can be classified into two types in IEEE 802.11n: entire and partial frame losses. We first classify a frame loss according to its type and cause. Interference from hidden nodes can cause both entire and partial frame losses. In order to detect both types of frame losses due to hidden nodes, we propose two hidden node detection methods. One is based on the statistics of measurable parameters in the MAC layer, and the other is based on the received block ACK frame. When no hidden nodes are detected, the transmitter only adjusts its MCS and CW values to cope with channel errors and collisions. When hidden nodes are detected, the transmitter employs the RTS/CTS exchange if it is expected to give better throughput performance than the basic access mode. We evaluate the performance of the proposed HD mechanism through simulation. It detects the hidden nodes well under various circumstances and the network throughput is improved through the use of the HD mechanism with RTS/CTS.

The contributions of this chapter are as the follows:

- Under a realistic environment in which frame losses can be caused by any combination of the three main causes described above, we show how to classify frame losses according to their type and cause in IEEE 802.11n WLAN.

- We propose a hidden node detection mechanism that helps transmitter to take the appropriate action against frame losses according to their type and cause. The proposed mechanism does not need to modify the IEEE 802.11n standard.

- We show that the network throughput can be improved by using an appropriate hidden node resolution mechanism such as RTS/CTS in conjunction with the proposed HD mechanism.

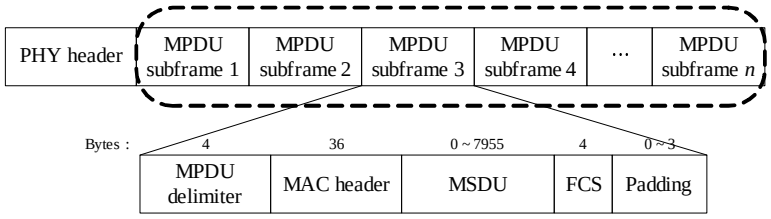
The rest of the chapter is organized as follows. Section 3.2 describes the system model considered throughout this chapter. Section 3.3 presents the hidden node detection mechanism and Section 3.4 evaluates the performance of the proposed mechanism. Finally, Section 3.5 concludes the chapter.

3.2 System Model

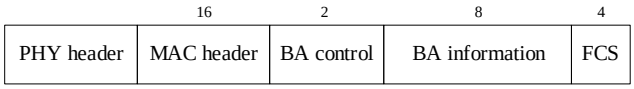
In this section, we briefly introduce the basic features and models in the MAC and PHY layers of IEEE 802.11n for the development of this chapter.

3.2.1 Frame aggregation and block ACK in the MAC Layer

Throughout this chapter, we consider only A-MPDU, shown in Fig. 3.1(a), which can achieve higher throughput performance than A-MSDU in an error-prone channel by selectively retransmitting only corrupted MPDU subframes



(a) A-MPDU data frame.



(b) Block ACK frame.

Figure 3.1 A-MPDU data and block ACK frame formats in IEEE 802.11n.

[15]. The concept of A-MPDU aggregation is to gather multiple MPDU subframes with a single leading PHY header as in Fig. 3.1(a). It begins by appending each MSDU with its own MAC header and frame check sequence (FCS) to form an MPDU subframe. Then, an MPDU delimiter is inserted before each MPDU subframe. Padding bits are also added so that the length of each subframe is a multiple of four bytes, which can facilitate subframe delineation at a receiver. Finally, all of the MPDU subframes are concatenated to form a single large frame.

The block ACK mechanism was originally proposed with the TXOP mechanism in the IEEE 802.11e standard [14]. Unlike the legacy 802.11, where each data frame is acknowledged individually, the block ACK is used to acknowledge several consecutive data frames in one response. This can help to decrease the protocol overhead. The block ACK can be used with the A-MPDU in IEEE 802.11n. Figure 3.1(b) shows the frame format of the block ACK. The receiver conducts a FCS for each MPDU subframe while decoding the transmitted A-MPDU frame. Then it acknowledges each MPDU subframe to the transmitter

using the BA information field in the block ACK frame. We assume that the receiver immediately responds to the transmitted A-MPDU frame with a corresponding block ACK frame.

3.2.2 Channel and interference models in the PHY Layer

OFDM is a modulation scheme used in 802.11n PHY which employs 64 subcarriers, where 52 subcarriers are for data and the remaining subcarriers are for pilot or null signals. Forward error correction (FEC) is performed by convolutional coding that uses the generator polynomials $g_0 = 131_8$ and $g_1 = 171_8$ of the code rate 1/2 [5]. Higher code rates of 2/3, 3/4, and 5/6 can be obtained by puncturing the original 1/2 code. The coded bits are interleaved and then mapped to symbols from one of the allowed constellations (BPSK, QPSK, 16QAM, and 64QAM). Table 3.1 shows the multiple MCSs considered in this chapter.

We assume that the transmission power P_t is identical for all transmitters and the the propagation of the signal follows the log-distance path-loss model with the path-loss exponent of α [45]. The gain G from the transmitter to the receiver can be expressed as

$$G(r) = \frac{1}{r^{\alpha}}, \tag{3.1}$$

where r is the distance between the transmitter and the receiver. The instantaneous received signal-to-interference-plus-noise ratio (SINR) at the receiver

is

$$\gamma = \frac{P_t G(r)}{P_N + \sum P_t G(r')}, \tag{3.2}$$

where P_N and r' are the background noise power and the distance from an interfering transmitter, respectively. We denote γ_m as the SINR threshold for each MCS m, which is given in Table 3.1. We assume that a receiver can decode a frame successfully with probability larger than 95% when the received SINR is consistently above the SINR threshold during transmission of the frame. We define the transmission range as the maximum distance that the signal can be successfully received with probability larger than 95% without interference. Then, the transmission range for each MCS, r_m, can be calculated as follows:

$$r_m = \left(\frac{1}{\gamma_m} \frac{P_t}{P_N} \right)^{1/\alpha}. \tag{3.3}$$

Table 3.1 shows r_m for each MCS where P_t, P_N, and α are 15 dBm, -87 dBm, and 5, respectively. We assume that the transmitter's carrier is sensed by carrier detection [5]. If a transmitter senses power that exceeds the carrier sense threshold P_{th}, it does not transmit and freezes its back-off countdown process. This will prevent frame transmissions from other transmitters when a transmitter is transmitting data. The carrier sensing range, r_{cs}, is mapped from the carrier sense threshold P_{th}, as follows:

$$r_{cs} = \left(\frac{P_t}{P_{th}} \right)^{1/\alpha}. \tag{3.4}$$

P_{th} is assumed to be -85 dBm, which corresponds to r_{cs} of 100 meters.

Table 3.1 IEEE 802.11n PHY (For mandatory 20 MHz channel bandwidth and 800 nanoseconds guard interval)

MCS index (m)	Modulation scheme	Code rate	Data rate	SINR threshold (γ_m)	Transmission range (r_m)
1	BPSK	1/2	6.5 Mb/s	6.8 dB	80 m
2	QPSK	1/2	13 Mb/s	7.9 dB	76 m
3	QPSK	3/4	19.5 Mb/s	10.6 dB	67 m
4	16QAM	1/2	26 Mb/s	13.0 dB	60 m
5	16QAM	3/4	39 Mb/s	17.0 dB	50 m
6	64QAM	2/3	52 Mb/s	21.8 dB	40 m
7	64QAM	3/4	58.5 Mb/s	24.7 dB	35 m
8	64QAM	5/6	65 Mb/s	28.1 dB	30 m

We also assume that a transmitter utilizes the FLA scheme that can be employed in the IEEE 802.11n system [3]. A receiver estimates the channel condition during the reception of the PHY header and selects the highest MCS among the MCSs that satisfy $\gamma > \gamma_m$. The selected MCS is reported back to the transmitter using the reserved field in a block ACK frame, after which the transmitter can adjust its MCS for the next transmission. In this way, a transmitter can adjust its transmission rate according to the channel condition.

3.3 Hidden Node Detection Mechanism

In this section, we first show how frame losses are classified according to their type and cause in IEEE 802.11n that employs the frame aggregation and block ACK schemes. Then, a hidden node detection mechanism is proposed for a transmitter to check whether there are hidden nodes, which enables the transmitter to take appropriate counter measures in such a case.

3.3.1 Classification of frame losses

In the legacy 802.11 WLANs, the receiver sends an ACK frame to the transmitter only when the entire data frame is decoded successfully. This results in only one type of frame loss (entire frame loss), and it is sufficient to classify frame loss according to its cause. However, in IEEE 802.11n that employs the frame aggregation and block ACK schemes, the receiver can send a block ACK frame back to the transmitter once the PHY header of a frame is decoded successfully, even when the subframes are not successfully received. Therefore, there are two types of frame losses: entire frame loss and partial frame loss. We subdivide frame losses according to their type and cause, as shown in Table 4.2. We first categorize frame losses into entire (E) and partial (P) frame losses according to the receiver's response. For the entire frame loss type (E), we denote frame losses caused by a collision, hidden nodes, and channel impairments as events (E1), (E2), and (E3), respectively. A collision always results in entire frame loss not partial frame loss, because the frame transmissions that result in collision always start at the same slot. Therefore, we only consider partial frame losses caused by hidden nodes and channel impairments, denoting them as (P2) and (P3), respectively. The transmitter can get more information on the transmission result due to the block ACK, and this can help to differentiate frame losses according to their causes, as discussed in the next subsection.

3.3.2 Detecting hidden nodes

In this subsection, we propose a hidden node detection (HD) mechanism. We first present how to detect events (E2) and (P2). Then, we present a method

[4]The frame body corresponds to A-MPDU in Fig. 3.1(a).

Table 3.2 Classification of frame loss event

Type of frame losses	Causes of frame losses	Receiver response
Entire frame loss (E)	(E1) Collisions (synchronous interference) (E2) Hidden nodes (asynchronous interference on PHY header) (E3) Channel impairments (during PHY header or block ACK frame transmission)	No response
Partial frame loss (P)	(P2) Hidden nodes (asynchronous interference on frame body[4]) (P3) Channel impairments (during frame body transmission)	Transmit block ACK

for a transmitter to determine whether or not to use the RTS/CTS exchange.

Detecting (E2)

The PHY headers and conventional control frames such as RTS, CTS, and ACK frames are transmitted at the lowest MCS for reliable transmission in WLANs [5]. The block ACK frame in IEEE 802.11n can be transmitted at a higher MCS, but the IEEE 802.11n standard does not specify which is the best MCS for transmitting a block ACK frame. A transmitter can know the best MCS for transmitting a data frame while a receiver cannot, because the link adaptation is employed only for data frames. One possible approach is to transmit a block ACK frame using the same MCS for data frames. This will make sense when the forward and reverse links have the same channel condition. However, wireless channels are asymmetric in general, and thus this approach cannot guarantee the transmission of a block ACK frame to be error-free. This is troublesome for 802.11n nodes employing the frame aggregation because the entire data frame has to be retransmitted at the failure of decoding a block ACK frame. Unless the receiver knows the best MCS for a block ACK frame, it

is better to transmit a block ACK frame at the lowest MCS as in transmitting the conventional control frames for reliable transmission. Hence, we proceed further under the assumption that block ACK frames are always transmitted at the lowest MCS. Although the size of a block ACK frame (56 bytes including PHY header) is larger than those of the conventional control frames, the error probability of a block ACK frame is at most 4.5% in a very noisy channel (e.g., BER of 10^{-4}). Therefore, when a transmitter fails to receive a block ACK frame after transmitting a data frame, it can interpret that the transmission failure is due to a collision (E1) or due to hidden nodes (E2), considering that the channel error probabilities of transmitting a PHY header and a block ACK frame are negligible. [1]

The transmitter can estimate the probability of (E2) using measurable MAC layer statistics, as described below. We consider an arbitrary node i with N_i other transmitters within its carrier sensing range. Also, there may be hidden transmitters out of the carrier sensing range of node i. Let us denote p_i^{ctr} as the probability that there is at least one node among N_i nodes transmitting with node i at the same slot. It is noted that not all of the concurrent transmissions by the nodes in the carrier sensing range result in a collision. The receiver can still decode a frame when the received SINR is larger than the SINR threshold, which is known as the capture effect [46], and we denote this probability as p_i^{cap}. The probability that node i successfully receives a block ACK frame, p_i^{ack}, can

[1] In the simulation study of Section 3.4, we demonstrate that the channel error probability of a block ACK frame has little effect on the performance of the proposed HD mechanism.

be expressed as

$$p_i^{ack} = p_i^{ctr} p_i^{cap} + (1 - p_i^{ctr})(1 - p_{i,E}^{hid}), \tag{3.5}$$

where $p_{i,E}^{hid}$ is the probability of the PHY header becoming corrupted by interference from hidden nodes (E2). If we denote τ_i as the attempt probability of node i, p_i^{ctr} can be derived as

$$p_i^{ctr} = 1 - \prod_{j=1}^{N_i} (1 - q_{i,j}\tau_j), \tag{3.6}$$

where $q_{i,j}$ is the conditional probability that node j senses the channel as idle given that node i senses the channel as idle. If the carrier sensing areas of nodes i and j are identical, the conditional probability $q_{i,j}$ then becomes 1. However, the carrier sensing areas can be different, and $q_{i,j}$ is less than 1 in general. We also define p_i^{idle} as the probability that no node initiates transmission within the carrier sensing range of node i in an idle slot. This can be represented as

$$p_i^{idle} = (1 - \tau_i) \prod_{j=1}^{N_i} (1 - q_{i,j}\tau_j). \tag{3.7}$$

It is difficult for node i to know the values of $q_{i,j}$ and N_i without exchanging information between the nodes in the carrier sensing range. Fortunately, we can eliminate $q_{i,j}$ and N using the relationship between (3.6) and (3.7). In this case, p_i^{ctr} becomes

$$p_i^{ctr} = 1 - \frac{p_i^{idle}}{1 - \tau_i}. \tag{3.8}$$

Because we can express $p_{i,E}^{hid}$ in (3.5) as

$$p_{i,E}^{hid} = 1 - \frac{p_i^{ack} - p_i^{ctr} p_i^{cap}}{1 - p_i^{ctr}}, \tag{3.9}$$

and together with p_i^{ctr} in (3.8), we can estimate $p_{i,E}^{hid}$ by estimating p_i^{ack}, p_i^{cap}, τ_i, and p_i^{idle}. The probabilities p_i^{ack}, τ_i, and p_i^{idle} can be estimated using MAC layer statistics. However, it is difficult to estimate p_i^{cap} because this value depends on the relative location of node i and the nodes that concurrently transmit, which may change with time.[2] Instead, we can easily obtain a lower bound $\underline{p}_{i,E}^{hid}$ for $p_{i,E}^{hid}$ by setting $p_i^{cap} = 0$, which can be expressed as

$$\underline{p}_{i,E}^{hid} = 1 - \frac{p_i^{ack}(1 - \tau_i)}{p_i^{idle}}. \tag{3.10}$$

At this point, we present how to estimate τ_i, p_i^{ack} and p_i^{idle} using measurable MAC layer statistics similar to the estimation in [40]. We define the *transmission cycle* of DCF for node i as the time period consisting of backoff slots (i.e., idle slots) and times for busy medium due to other nodes' transmissions, node i's data transmission and the corresponding SIFS, block ACK transmission, and DIFS, as shown in Fig. 3.2(a). For ease of explanation, these terms, excluding the backoff slots and the busy medium time, are merged into a *transmission instant*. As each transmission starts and ends at the slot boundaries, we can abstractly draw the transmission instants and busy medium times due to other nodes' transmissions as black dots, as shown in Fig. 3.2(b). Here, Δ_k denotes the DATA and block ACK transmission times plus DIFS and SIFS in the k-th

[2]It is possible to numerically calculate the average p_i^{cap} for a given spatial distribution of nodes and channel model, as in [46]. Note that the average capture probability becomes negligible as the number of nodes increases.

transmission cycle. In Fig. 3.2(b), there are two transmission cycles that occur during the time intervals $(t_k + \Delta_k, t_{k+1} + \Delta_{k+1}]$ and $(t_{k+1} + \Delta_{k+1}, t_{k+2} + \Delta_{k+2}]$. We define the *observation period* as the time taken for a node to complete l consecutive transmission cycles. For example, if l is set to 2, the time interval $(t_k + \Delta_k, t_{k+2} + \Delta_{k+2}]$ in Fig. 3.2(b) then corresponds to one observation period. Due to the characteristics of CSMA/CA of the IEEE 802.11 MAC, node i can accurately count the numbers of idle slots I_i, its own missing block ACKs F_i, and the transmissions of other nodes in the carrier sensing range O_i during the observation period. From the definitions of τ_i, p_i^{ack} and p_i^{idle}, we can estimate these values as follows:

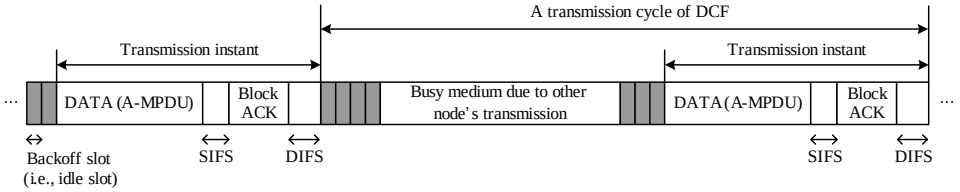
$$\tau_i = \frac{l}{I_i}, \quad p_i^{ack} = 1 - \frac{F_i}{l}, \quad p_i^{idle} = 1 - \frac{l + O_i}{I_i}. \tag{3.11}$$

Finally, $\underline{p}_{i,E}^{hid}$ is low-pass filtered with the coefficient λ ($0 < \lambda < 1$) as follows:

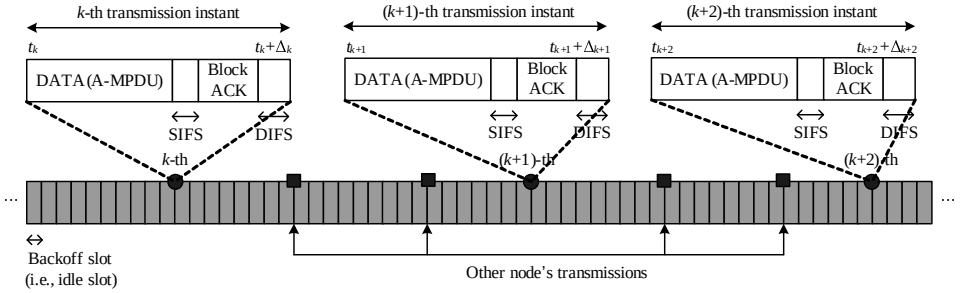
$$\underline{p}_{i,E}^{hid} \leftarrow \lambda \cdot \underline{p}_{i,E}^{hid} + (1 - \lambda) \cdot \left[1 - \frac{p_i^{ack}(1 - \tau_i)}{p_i^{idle}} \right]. \tag{3.12}$$

Once $\underline{p}_{i,E}^{hid}$ is estimated, it can be used to help the transmitter's decision. When $\underline{p}_{i,E}^{hid}$ is less than a certain threshold $\eta^{(E2)}$, [3] the transmitter considers that an entire frame loss is caused mainly due to a collision and doubles its CW value to reduce the collision probability. Otherwise, it considers that there are hidden nodes and moves to the next step to determine whether or not to use the RTS/CTS mechanism, which is discussed later.

[3] If $\eta^{(E2)}$ is set too low, the transmitter may misinterpret that there is a hidden node due to the estimation error in $\underline{p}_{i,E}^{hid}$. We set $\eta^{(E2)}$ as 0.1 to avoid such possible misinterpretation. Actually, it is more important to properly determine whether or not to use RTS/CTS, which will be discussed later.



(a) In the actual time-line.



(b) Abstraction.

Figure 3.2 Transmission cycles of DCF for an arbitrary node i.

Detecting (P2)

Detecting (E2) is based on the probability of MAC layer statistics because a transmitter does not receive a block ACK. In the case of partial frame loss, however, a transmitter can receive the transmission result of each MPDU subframe from the received block ACK frame. When a transmitter receives a block ACK frame, it only needs to differentiate partial frame losses due to hidden nodes (P2) from channel errors (P3), considering that a collision cannot occur in such a case.

A receiver estimates the channel condition during the reception of the PHY header and reports an appropriate MCS to the transmitter using a block ACK frame. The transmitter then updates its MCS based on the received block ACK frame. When the PHY header of the A-MPDU frame is corrupted by

interference from hidden nodes, the block ACK cannot be transmitted to the transmitter and thus the transmitter does not update its MCS. A frame body can be decoded with high probability when there is no interference from hidden nodes, given that a transmitter can adjust its MCS according to the channel state. When the error probability of an MPDU subframe is higher than expected, the hidden nodes are considered to be the main cause of the error. The transmitter can easily calculate the error probability for an MPDU subframe by simply counting the number of corrupted MPDU subframes in an A-MPDU frame. Let us define $p_{i,P}^{hid}$ as the error probability of an MPDU subframe due to interference from hidden nodes (P2). Upon receiving a block ACK, $p_{i,P}^{hid}$ is low-pass filtered with a coefficient λ ($0< \lambda <1$) as

$$p_{i,P}^{hid} \leftarrow \lambda \cdot p_{i,P}^{hid} + (1-\lambda) \cdot \frac{X}{n}, \tag{3.13}$$

where n and X are the numbers of all MPDU subframes and corrupted MPDU subframes in an A-MPDU frame, respectively. The transmitter can detect the presence of hidden nodes when $p_{i,P}^{hid}$ is larger than a certain threshold $\eta^{(P2)}$. Here, $\eta^{(P2)}$ needs to be set appropriately based on the capability of the link adaptation scheme. According to the system model of the FLA scheme described in Section 3.2, the error probability of an MPDU subframe is guaranteed to be less than 5% when there is no hidden node. In such a case, it is reasonable to set $\eta^{(P2)} = 0.05$. Similarly in practical FLA schemes such as [25] and [26], the MCS is adjusted to meet the target FER which result in the average FER performance to be maintained around the target value; thus in this case, it is reasonable to set $\eta^{(P2)}$ around the target FER.

When to employ RTS/CTS

When (E2) or (P2) is detected, the proposed HD mechanism determines whether or not to initiate the RTS/CTS exchange depending on the estimated values of $p_{i,E}^{hid}$ and $p_{i,P}^{hid}$. It is better to initiate the RTS/CTS exchange when the RTS/CTS mode is expected to yield higher throughput than the basic access mode. Therefore, we need to derive an analytic model that can estimate the throughput performance with/without turning on the RTS/CTS exchange. The studies in [47–49] give throughput models of 802.11 WLANs where each node employs the frame aggregation and block ACK schemes by taking both collisions and channel errors into consideration. However, the models are based on the assumption that all the nodes have the same attempt probability, and the effect of hidden nodes is not taken into account. On the other hand, a per-flow throughput model can be derived based on [40] by taking the three main causes of frame losses into consideration. Still there are difficulties in comparing the basic access mode and RTS/CTS mode with respect to throughput, because τ_i and p_i^{ctr} may vary depending on whether the RTS/CTS exchange is employed or not. To make the derivation of the throughput model simple, we ignore the variations of τ_i and p_i^{ctr} by the RTS/CTS exchange, and only focus on the case where a node gets a chance to access channel after competing with other nodes. Also, we assume that the RTS/CTS exchange can effectively resolve the hidden node problem. Let us denote S_i as the throughput for node i after competing with other nodes. Then, S_i can be calculated as

$$S = \frac{8 n_s L_p}{T_{data} + T_{oh}}, \tag{3.14}$$

where n_s and L_p are the number of successfully transmitted MPDU subframes and the packet size, respectively. Recall that T_{oh} is the overhead required for completing the transmission of single data frame, which is calculated as

$$T_{oh} = \begin{cases} T_{back} + T_{SIFS} + T_{DIFS} + T_{BO}, & \text{for basic access mode,} \\ T_{rts} + T_{cts} + T_{back} + 3\ T_{SIFS} + T_{DIFS} + T_{BO}, & \text{for RTS/CTS mode,} \end{cases} \tag{3.15}$$

where T_{rts}, T_{cts}, are the time durations for transmitting RTS and CTS frames, respectively. In our model, each MPDU subframe in a data frame is affected only by the hidden nodes in the basic access mode. On the other hand, all MPDU subframes in a data frame can be successfully transmitted in the RTS/CTS mode. Then, the number of successfully transmitted MPDU subframes can be estimated as

$$n_s = \begin{cases} (1 - \underline{p}_{i,E}^{hid})(1 - p_{i,P}^{hid})n_2, & \text{for basic access mode,} \\ n_2, & \text{for RTS/CTS mode.} \end{cases} \tag{3.16}$$

Let S_i^B and S_i^R be the throughput for the basic access mode and RTS/CTS mode, respectively. Node i calculates S_i^B and S_i^R based on the estimated $\underline{p}_{i,E}^{hid}$ and $p_{i,P}^{hid}$ by using (3.14), (3.15), and (3.16), and employs the RTS/CTS exchange when S_i^R/S_i^B is larger than 1.

We end this section by summarizing the proposed HD mechanism. Figure 3.3 shows a flow chart of the HD mechanism at a transmitter. After transmitting a data frame, the transmitter updates $\underline{p}_{i,E}^{hid}$ using the MAC layer statistics, as in (3.12). When the transmitter does not receive a block ACK until the ACK timeout, it checks whether (E2) has occurred using the updated $\underline{p}_{i,E}^{hid}$.

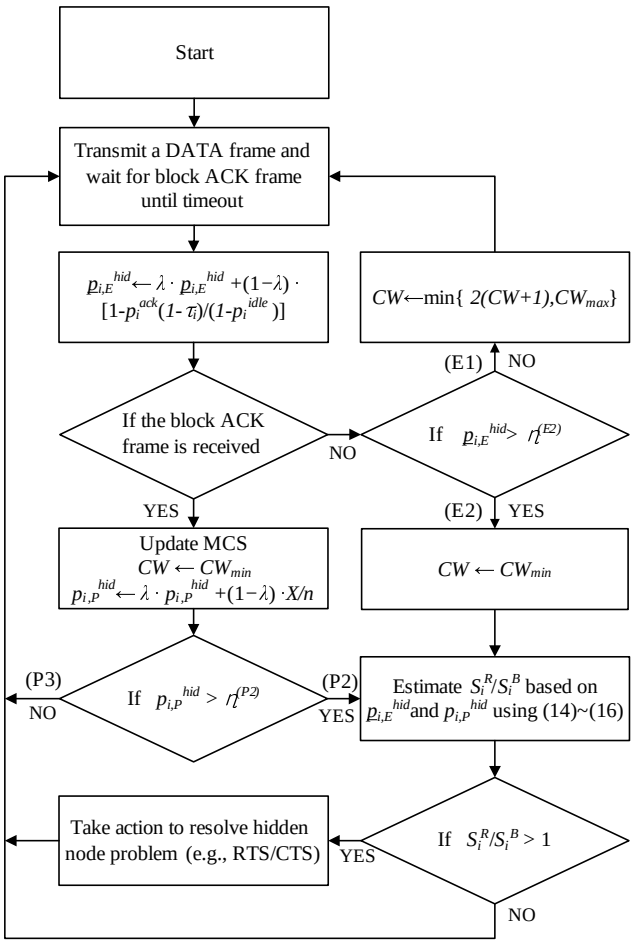


Figure 3.3 Flow chart of the proposed HD mechanism.

The transmitter doubles its CW value to reduce the collision probability when $p_{i,E}^{hid} < \eta^{(E2)}$. Otherwise, (E2) is detected, and the transmitter initialize its CW value and moves to the next step of estimating the ratio S_i^R/S_i^B. When the transmitter successfully receives a block ACK, it first updates its MCS and initializes its CW value because a collision cannot result in partial frame loss. It then updates $p_{i,P}^{hid}$ as in (3.13) by inspecting the BA information field in the received block ACK frame. It takes no action when $p_{i,P}^{hid}$ is less than $\eta^{(P2)}$.

Otherwise, (P2) is detected and the transmitter moves to the next step of estimating the ratio S_i^R/S_i^B. After estimating S_i^R/S_i^B, the transmitter takes a necessary action to resolve the hidden node problem when $S_i^R/S_i^B>1$.

3.4 Performance Evaluation

In this section, we evaluate the performance of the proposed HD mechanism via simulation using the ns-2 simulator [50]. We implemented MAC and PHY layers according to the system model described in Section 3.2. The thresholds for the proposed HD mechanism, $\eta^{(E2)}$ and $\eta^{(P2)}$, were set to 0.1 and 0.05, respectively. The parameters l and λ were set to 5 and 0.75, respectively. UDP packets were generated by CBR sources so that the transmitters always had packets to transmit and the frames were transmitted using the MCSs in Table 3.1. The packet size L_p and the number of MPDU subframes n were assumed to be 512 bytes and 5, respectively. Other system parameters related to IEEE 802.11n can be found in [3]. The performances levels of the following four schemes are evaluated through simulations.

- `FLA` : This is a baseline scheme that only employs FLA. Its throughput performance degrades when there are hidden nodes.

- `FLA-R` : The RTS/CTS exchange is always used prior to the data transmission irrespective of the existence of hidden nodes. Its throughput performance is expected to be less than that of FLA when there is no hidden node.

- `FLA-HD(o)-AR` : The hidden node detection mechanism proposed in [42]

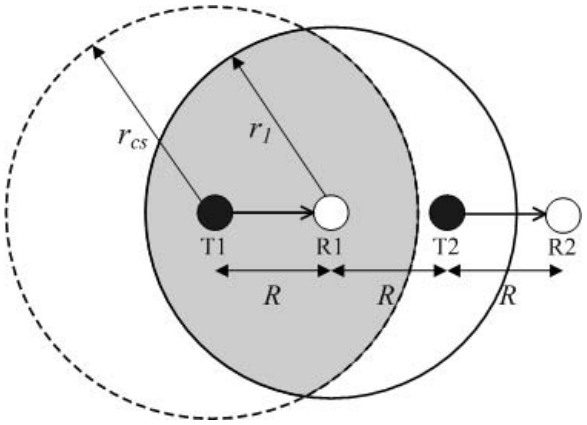


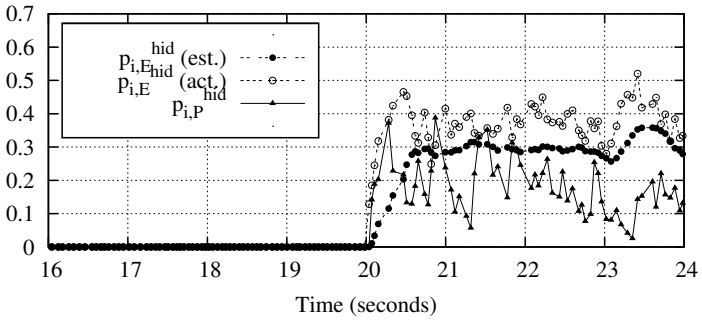
Figure 3.4 The network topology for testing the proposed HD mechanism. (There are additional ten transmitters that transmit data frames to `R1` within the shaded area.)

is employed so that the RTS/CTS exchange can be adaptively used only when there are hidden nodes.

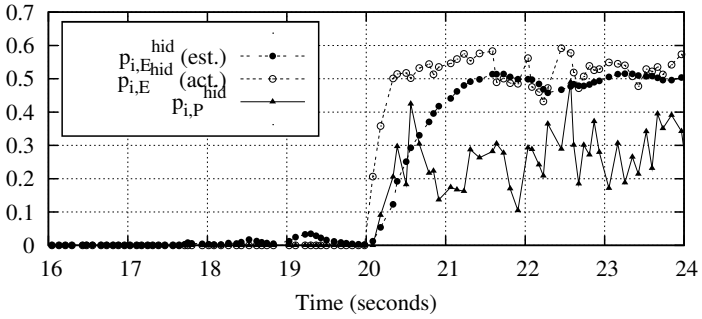
- `FLA-HD(p)-AR` : The proposed HD mechanism is employed so that the RTS/CTS exchange can be adaptively used only when there are hidden nodes.

3.4.1 Validation of the proposed HD mechanism

We first show the effectiveness of the proposed HD mechanism in detecting hidden nodes and initiating RTS/CTS. Let us consider a network topology with two links, `Link1` and `Link2`, which consist of transmitter-receiver pairs (`T1-R1`) and (`T2-R2`), respectively, as depicted in Fig. 3.4. There are additional ten transmitters that transmit data frames to `R1` within the shaded area in Fig. 3.4. We set R so that it is larger than 50 meters, implying that `T1` and `T2` cannot carrier sense each other. The frame transmission of `T2` can interfere with the



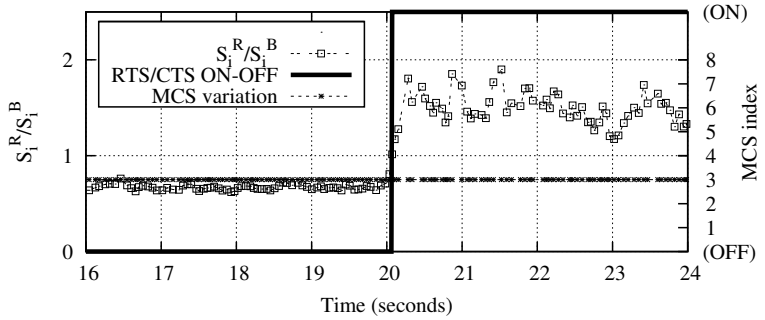
(a) $R = 65$ meters.



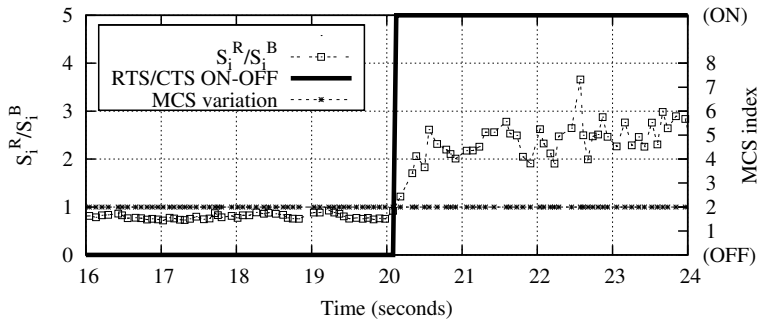
(b) $R = 75$ meters.

Figure 3.5 Variation of $p_{i,E}^{hid}$ and $p_{i,P}^{hid}$ in `T1`.

frame transmission of `T1`, whereas the frame transmission of `T1` does not affect the frame transmission of `T2`. In the simulation, `T2` starts to send data frames at 20 seconds, which implies that `Link1` experiences frame losses due to hidden node after 20 seconds. We observe how the HD mechanism works in `T1`. Figures 3.5(a) and 3.5(b) show the actual $p_{i,E}^{hid}$ and its estimate value, $\underline{p}_{i,E}^{hid}$, with time when R is 65 and 75 meters, respectively. To have a reference in evaluating the proposed HD mechanism, we implemented a genie-aided receiver that can perfectly differentiate the cause of frame losses and give the actual value of $p_{i,E}^{hid}$. It is noted that the transmitter that employs FLA selects MCS 3 (19.5 Mb/s) and 2 (13 Mb/s) for $R = 65$ and $R = 75$, respectively. Therefore,



(a) $R = 65$ meters.



(b) $R = 75$ meters.

Figure 3.6 Variation of estimated S_i^R/S_i^B (for initiating RTS/CTS) and MCS index in `T1`.

the time duration for transmitting the frame body when $R = 75$ is two times longer compared to when $R = 65$. The estimation result of $p_{i,E}^{hid}$ is appropriate to use for both cases. The gap between the actual value and its estimate is due to estimation error and the capture effect that is not considered when calculating $p_{i,E}^{hid}$. Figures 3.5(a) and 3.5(b) also show the variation of $p_{i,P}^{hid}$ with time. These figures show that partial frame losses increase as the transmission time of frame body increases. Figures 3.6(a) and 3.6(a) show the variation of estimated S_i^R/S_i^B with time when R is 65 and 75 meters, respectively. It is clear that the HD mechanism quickly detects the frame loss due to hidden nodes and

appropriately turns on RTS/CTS. These figures also show the variation of the MCS index adjusted by the FLA scheme. Thanks to block ACK, the FLA scheme can appropriately adjust the MCS, even when there are hidden nodes. This is an advantage of FLA, considering that other link adaptation schemes may misinterpret frame losses due to hidden nodes as due to channel error and consequently lowers the MCS level.

We performed simulations to compute the false alarm ratio of the proposed HD mechanism in general network topologies. We determine the existence of hidden nodes based on actual measurements of $p_{i,E}^{hid}$ and $p_{i,P}^{hid}$, which can be obtained by the genie-aided receiver. Let us define the event U where a node detects hidden nodes by the HD mechanism even though the actual measurements of $p_{i,E}^{hid}$ and $p_{i,P}^{hid}$ does not exceed thresholds $\eta^{(E2)}$ and $\eta^{(P2)}$, respectively. We also define the event V where a node does not detect hidden nodes by the HD mechanism, even though the actual measurements of $p_{i,E}^{hid}$ or $p_{i,P}^{hid}$ exceeds the thresholds. The occurrence ratios of the events U and V are the false positive and false negative alarm ratios, respectively. We assume a network topology where N transmitter-receiver pairs (i.e., links) are randomly located in a square area of 200m×200m, and performed simulation for 30 different random topologies. To show that the channel error probability of a block ACK frame is negligible in the HD mechanism, we performed simulation for various values of BER on the block ACK frame. Table 3.3 shows the false positive and negative alarm ratios of the HD mechanism for various values of N and BER. Detecting (E2) in the HD mechanism results in the occurrence of the event U (i.e., false negative alarm) because the HD mechanism uses the estimate of the lower bound, $\underline{p}_{i,E}^{hid}$. As the values of BER and N increase, the false negative

Table 3.3 False alarm ratios in the proposed HD mechanism.

N	10			20			30		
BER	10^{-6}	10^{-5}	10^{-4}	10^{-6}	10^{-5}	10^{-4}	10^{-6}	10^{-5}	10^{-4}
False-positive	0.020	0.017	0.022	0.022	0.023	0.023	0.018	0.023	0.024
False-negative	0.035	0.042	0.061	0.053	0.056	0.067	0.077	0.081	0.088

alarm ratio slightly increases due to the increase in the estimation errors of $p_{i,E}^{hid}$, but they are still very low (3∼9%). On the other hand, detecting (P2) in the HD mechanism results in the occurrence of the event V (i.e., false positive alarm) because the HD mechanism may misinterpret channel error as hidden node when the estimated $p_{i,P}^{hid}$ is slightly larger than $\eta^{(P2)}$. The positive alarm ratio does not depend on the values of BER and N, and is less than the false negative alarm ratio for all values of BER and N. From the above simulation result, we can conclude that the proposed HD mechanism can detect hidden nodes with low false alarm ratios regardless of the number of links and the channel error probability of a block ACK frame.

3.4.2 Performance comparison

At this point, we compare the throughput performance of the four schemes to study how the proposed HD mechanism with adaptive RTS/CTS can improve network throughput performance. The network topologies that we have considered are described below.

- A *symmetric star topology network* where an AP is located at the center of the network and ten transmitters are uniformly located on a circle with a radius of R (meters). The hidden node problem occurs for all nodes when

R is larger than 50 meters.

- An *asymmetric line topology network* where two links `Link1` and `Link2` are located on the same line. All of the distances between `T1-R1`, `R1-T2`, and `T2-R2` are R. This topology is identical to that in Fig. 3.4, except that there are no additional links. The hidden node problem occurs only for `Link1` when R is larger than 50 meters.

- An *overlapped BSS network* where two APs are located in the middle of the network and the distance between the APs is 50 meters. For each AP, twenty transmitters are randomly located around the AP and they transmit data frames to the AP.

- An *ad hoc network* where thirty links are randomly located in a square area of 200m×200m.

It is noted that the interference from hidden nodes can be a serious problem in complex network topologies such as the overlapped BSS and ad hoc networks.

Figure 3.7 shows the aggregate throughput of the four schemes in the symmetric star topology network as R increases. The figure shows that the aggregate throughput of each scheme decreases as R increases because the transmitters select a lower MCS as R increases. `FLA-R` that always uses the RTS/CTS exchange achieves a higher throughput performance than `FLA` when there are hidden nodes (R>50), but the protocol overhead due to RTS/CTS results in throughput degradation compared to `FLA` when there is no hidden node (R<50). `FLA-HD(p)-AR` shows the highest throughput performance among the four schemes throughout the range of R, as it turns on RTS/CTS only when

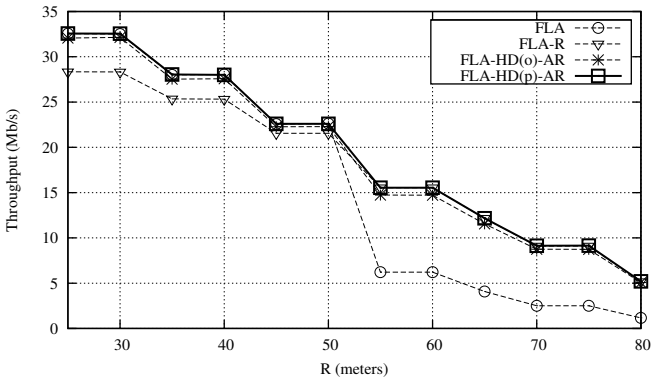
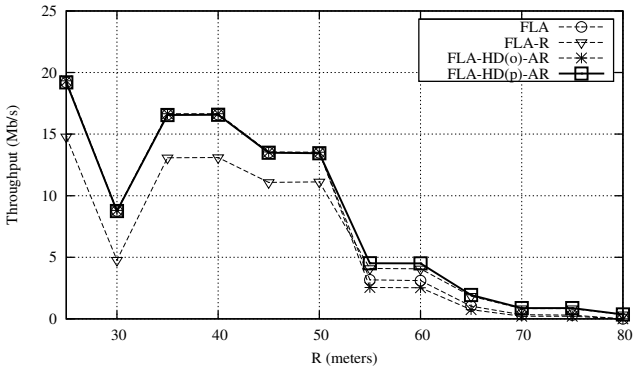


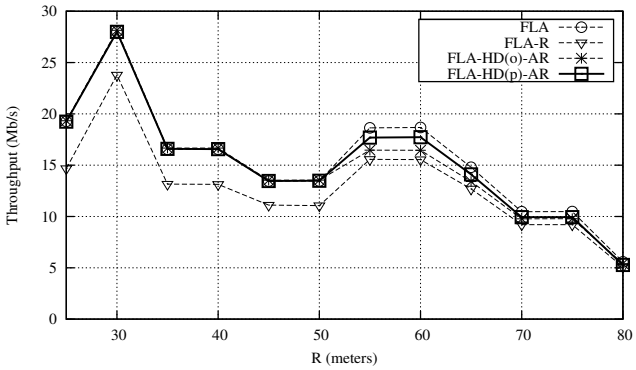
Figure 3.7 Aggregate throughput of the four schemes in the symmetric star topology network.

hidden nodes are detected. `FLA-HD(o)-AR` also gives comparable throughput performance to `FLA-HD(p)-AR`, because it can well detect hidden nodes in such a symmetric network. As already discussed in Section 3.1, however, the performance of `FLA-HD(o)-AR` may be degraded in general network topologies, which will be shown below.

Figures 3.8(a) and 3.8(b) show the link throughput of the four schemes in the asymmetric line topology network. When R is 30 meters, `Link2` achieves more throughput than that of `Link1` for all schemes even though `T1` and `T2` can carrier sense each other. This is because only the transmission of `T1` is successful due to the capture effect when `T1` and `T2` transmit data frame at the same slot. Also, the throughput of `Link2` increases as R increases in the range of $50<R<60$, because `T2` cannot carrier sense the signal of `T1` when R is larger than 50 meters. Thus, it does not defer its transmission during the transmissions of `T1`, which results in an increase of the `Link2` throughput. In the case of `FLA`, `Link1` starves when `T2` becomes a hidden node ($R > 50$), because `T2` greedily transmits data frames as if there are no other nodes around



(a) Link (T1-R1).



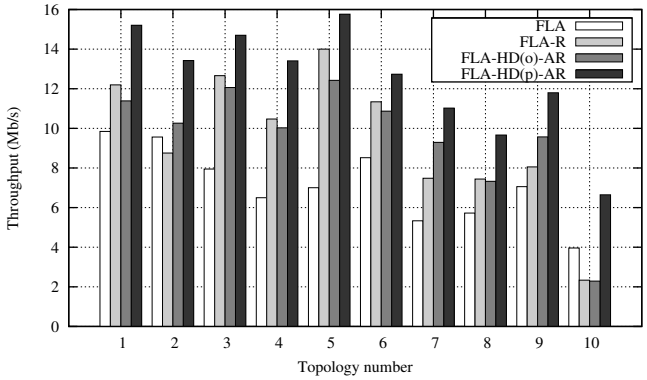
(b) Link (T2-R2).

Figure 3.8 Link throughput of the four schemes in the asymmetric line topology network.

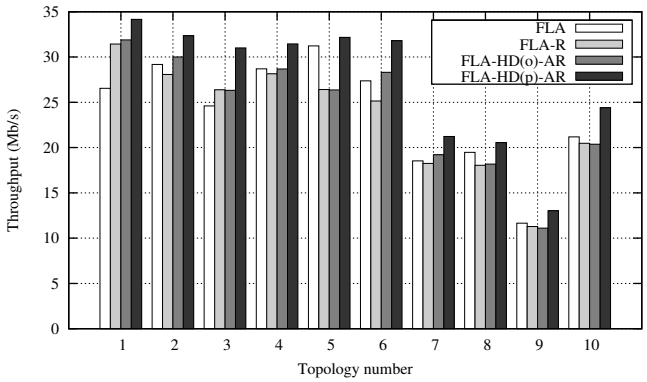
T2. In the case of FLA-R, Link1 achieves more throughput than that of FLA due to the RTS/CTS exchange, which improves the probability of a successful transmission of Link1. However, Link2 achieves lower throughput than that of FLA because Link2 does not need to use the RTS/CTS exchange. In the case of FLA-HD(o)-AR, both links falsely detect hidden nodes when R is larger than 50 meters. T1 can detect a hidden node when the block ACK transmission of R2 is observed even though the data transmission of T2 is not observed. However, R2 is out of the carrier sensing range of T1, and thus T1 fails to detect a hidden node

and does not turn on the RTS/CTS exchange. On the other hands, T2 misinterprets that there is a hidden node and thus turns on the RTS/CTS exchange, because it can carrier sense the block ACK transmission of R2 but not the data transmission of T1. Therefore, the throughput of Link1 in FLA-HD(o)-AR is similar to that in FLA, while the throughput of Link2 in FLA-HD(o)-AR is similar to that in FLA-R. When FLA-HD(p)-AR is used, Link1 uses the RTS/CTS exchange whereas Link2 does not when R is farther than 50 meters. Consequently, FLA-HD(p)-AR mitigates the starvation of Link1, as in FLA-R. On the other hand, the throughput of Link2 is larger than that of FLA-R because Link2 does not use RTS/CTS. The aggregate throughput of FLA-HD(p)-AR is similar to that of FLA considering that the throughput improvement of Link1 comes at the expense of Link2. Therefore, we can conclude that FLA-HD(p)-AR prevents starvation due to hidden nodes and provides throughput comparable to that of FLA in the asymmetric network topology.

Figures 3.9(a) and 3.9(b) show the aggregate throughput of the four schemes in the overlapped BSS and ad hoc networks, respectively. For each network, we tested the four schemes with ten different random topologies, and the figures show the result for each topology. In the case of overlapped BSS network (Fig. 3.9(a)), RTS/CTS can effectively alleviate the effect of hidden nodes; thus, FLA-R outperforms FLA in most cases. FLA-HD(o)-AR frequently fails to detect hidden nodes in such a complex network topology, and thus FLA-HD(o)-AR sometimes has less throughput than FLA-R. FLA-HD(p)-AR, in which only the nodes that are affected by hidden nodes turn on the RTS/CTS exchange appropriately, shows the highest throughput. The throughput improvements over FLA, FLA-R, and FLA-HD(o)-AR are 78%, 46% and 43%, respectively, on aver-



(a) Overlapped BSS network.



(b) Ad hoc network.

Figure 3.9 Aggregate throughput of the four schemes in the overlapped BSS and ad hoc networks.

age. In the case of ad hoc network (Fig. 3.9(b)), however, RTS/CTS alone does not solve the hidden node problem effectively and thus the performance gap between `FLA` and `FLA-R` is smaller than that in the overlapped BSS network. The usage of the RTS/CTS mechanism may rather degrade throughput performance due to the overhead caused by the RTS/CTS mechanism when RTS/CTS cannot resolve the hidden node problem. Although the RTS/CTS exchange could not completely resolve the hidden node, `FLA-HD(p)-AR` shows

higher throughput performance compared to other schemes and the throughput improvements over `FLA`, `FLA-R`, and `FLA-HD(o)` are 14%, 17% and 13%, respectively, on average.

3.5 Chapter Summary

In this chapter, we proposed the HD mechanism that is simple to implement by using new features in IEEE 802.11n system. The HD mechanism, which does not need to modify the IEEE 802.11n standard, detects hidden nodes well in a practical network environment in which frames can be lost due to any combination of collisions, hidden nodes, and channel impairments. It can be used to trigger a resolution mechanism for hidden nodes (e.g., RTS/CTS exchange), and the network throughput can be significantly improved. In the subsequent chapter, we present an enhanced mechanism to resolve the hidden node problem more effectively in conjunction with the HD mechanism.

Chapter 4

Hidden Node Resolution Mechanism

4.1 Background and Related Work

In the IEEE 802.11 [5] based WLANs, the hidden node problem [6] can arise in various forms due to the nature of the CSMA/CA mechanism. When a transmitter cannot carrier sense the signal of another node even though its signal can interfere with the frame reception at the receiver, this node is called as a hidden node which leads to degradation of network performance.

A number of studies have been addressed the hidden node problem in the literature. The experimental studies in [51, 52] show that the overall network throughput and fairness are significantly degraded by the hidden nodes. The studies in [53–55] derived analytic models to estimate the impact of the hidden nodes on the network performance for simple network topologies. However, it is

very difficult to extend the analytic models to the the case of a general network topology. To alleviate the hidden node problem, two control frames named RTS and CTS were first introduced in the multiple access with collision avoidance (MACA) scheme [56]. After a transmitter sends the RTS frame, the receiver sends the CTS frame to reply the RTS frame. Whenever the neighboring nodes overhear the RTS and/or CTS frames, they defer channel access for the expected duration of the transmission, which can be known by the reserved field of the RTS and/or CTS frames. It is noted that the disadvantage of MACA is the absence of acknowledgement of a successful transmission. The MACA scheme inspired the four way handshake including ACK frame in the IEEE 802.11 standard [5]. The MACA by invitation (MACA-BI) scheme [31, 57] was designed based on the MACA scheme to suppress the transmission of RTS frame at the transmitter. The receiver sends the ready to receive (RTR) frame, which is a renamed CTS frame and serves as the polling frame. Upon receiving an RTR frame, the transmitter sends a data frame to the receiver. It is noted that the receiver needs to know whether the transmitter has packets to send before transmitting an RTR frame. The MACA-BI scheme outperforms the MACA scheme when the hidden node problem is predominant [31]. The dual busy tone multiple access (DBTMA) mechanism [58] uses busy tone signals to keep hidden nodes silent during transmission. It divides a single common channel into two sub-channels, which are data and control channels. After exchanging the RTS/CTS frames, both transmitter and receiver transmit busy tone signals during data transmission, which helps to avoid any unwanted transmission because every node that senses busy tone signal do not start its own transmission. However, it requires additional transceivers and channels, and is not compatible

with the IEEE 802.11 standard because it does not use the ACK frame. Moreover, it is difficult to appropriately divide the channel into the control and data sub-channels, and sufficient spectral separation between the data and control sub-channels is needed to avoid inter-channel interference.

In this chapter, we intend to resolve the hidden node problem in IEEE 802.11n [3] WLANs with the help of the FA scheme[1] and the HD mechanism presented in the previous chapters. We first investigate the effectiveness of the RTS/CTS mechanism via simulation study for all possible scenarios of hidden node problem. To enhance the capability of the RTS/CTS mechanism, we propose an enhanced mechanism for resolving hidden node problem in two ways: *(i)* by extending the effective CTS range and *(ii)* by adopting the receiver-oriented contention (ROC) mechanism. Specifically, we make the following contributions.

- We classify all the possible scenarios of the hidden node problem according to the relative location of nodes in two flow interaction. Through a simulation study for each scenario, we show that the poor network performance caused by the RTS/CTS mechanism is due to the limitation of the effective CTS range and the carrier sensing mechanism at the transmitter-side.

- To extend the effective CTS range, we propose a method to identify the CTS frame for a node out of the transmission range but within the carrier sensing range of the CTS frame. After identifying the CTS frame, the

[1]Throughout this chapter, for simplicity, we assume that only the A-MPDU scheme is used for the FA scheme and each transmitter adjusts the number of packets in a frame so that T_{data} is less than but close to T_{ref} as much as possible for a give data rate and packet size.

node can appropriately set its network allocation vector (NAV) value with the help of the FA scheme even though it does not know the duration of the upcoming data transmission.

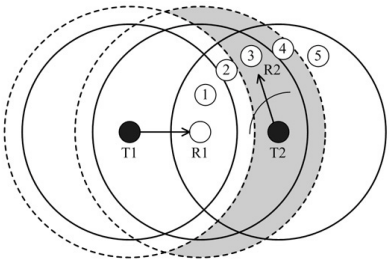
- To compensate the defect of the carrier sensing mechanism at the transmitter-side, we adopt the ROC mechanism, which is effective in presence of the hidden node problem. The ROC mode is initiated only when the interference from hidden nodes is detected via the HD mechanism.

- Through extensive simulations under various network topologies, we show that the proposed mechanism effectively resolves the hidden node problem, and the network throughput and fairness can be significantly improved.

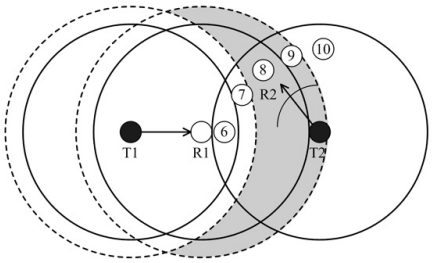
The rest of the chapter is organized as follows. Section 4.2 presents the various forms of hidden node problem and investigate the effectiveness of the RTS/CTS mechanism. Section 4.3 describes the proposed mechanism for resolving the hidden node problem, and Section 4.4 evaluates the performance of the proposed mechanism. Finally, Section 4.5 summarizes the chapter.

4.2 Various Network Configuration Causing the Hidden Node Problem

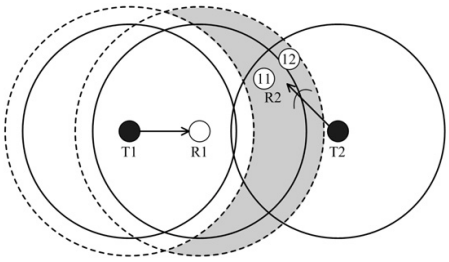
In this section, we classify all possible scenarios of the hidden node problem by considering the relative location of nodes. For each scenario, we investigate how the RTS/CTS mechanism affect on the network performance via simulation study.



(a) `T2` is located within the transmission range of `R1`.



(b) `T2` is located out of the transmission range but within the carrier sensing range of `R1`.



(c) `T2` is located out of the carrier sensing range of `R1`.

Figure 4.1 Classification of the hidden node problems according to the relative location of nodes. `T2` has to be located out of carrier sensing range of `T1` so that `T2` is the hidden node for `T1`.

4.2.1 Geographical network configurations

We classify the hidden node problem according to the relative location of nodes in two flow interaction, which can be used as a basis to understand more complex scenarios. Let us consider two links, `Link1` and `Link2`, which consist of the transmitter-receiver pairs (`T1`-`R1`) and (`T2`-`R2`), respectively. We classify the hidden node problem from the view point of `Link1` with fixed locations of `T1` and `R1`, and variable locations of `T2` and `R2`, as shown in Fig. 4.1. It is noted that the circles with the solid-line and the dotted-line indicate the transmission range and carrier sensing range, respectively. We assume that the interference

range is equal to the carrier sensing range. Then, Link1 is interfered by the hidden node when there are at least one node, whether it is a transmitter or a receiver, in the shaded area of Fig. 4.1. For a fixed Link1, we first locate T2 within the transmission range of R1, out of the transmission range but within the carrier sense range of R1, and out of the carrier sense range of R1, as in Figs. 4.1(a), 4.1(b), and 4.1(c), respectively. For the three different locations of T2, there are twelve scenarios that can cause the hidden node problem depending on the location of R2, which are summarized as the following.

- When T2 is located within the transmission range of R1, there are five possible scenarios depending on the location of R2, as shown in Fig. 4.1(a). We label these from S1 to S5. The transmission of T2 interferes with Link1 for these five scenarios whereas the transmission of R2 interferes with Link1 only in S3 and S4.

- When T2 is located out of the transmission range but within carrier sensing range of R1, there are five possible scenarios depending on the location of R2, as shown in Fig. 4.1(b). We label these from S6 to S10. The transmission of T2 interferes with Link1 for these five scenarios whereas the transmission of R2 interferes with Link1 only in S8 and S9.

- When T2 is located out of the carrier sensing range of R1, there are two possible scenarios depending on the location of R2, as shown in Fig. 4.1(c). We label these as S11 and S12. It is noted that only the transmission of R2 interferes with Link1 in S11 and S12.

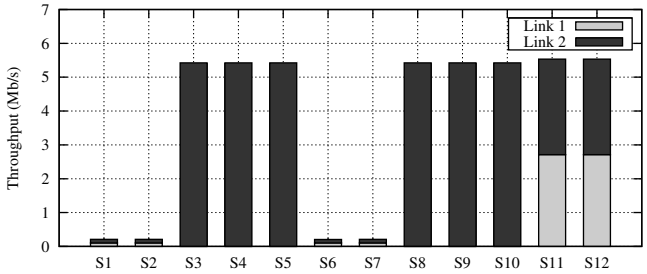
Although there are twelve possible scenarios for the hidden node problem in two flow interaction, some of scenarios may have a similar behavior. The twelve

scenarios can be grouped into three categories as in [54], which are symmetric incomplete state (SIS), asymmetric incomplete state (AIS), and interfering destinations incomplete state(IDIS). In short, `S1`, `S2`, `S6`, and `S7` correspond to the SIS category, `S3`, `S4`, `S5`, `S8`, `S9`, and `S10` correspond to the AIS category, and `S11` and `S12` correspond to the IDIS category. For the basic access mode, each scenario that belongs to a same category gives a similar performance. However, the categorization in [54] do not consider the effect of the RTS/CTS mechanism, even though the use of the RTS/CTS mechanism makes a difference among scenarios of the same category because of the difference between the transmission range and the carrier sensing range of the CTS frame. Therefore, it is necessary to consider all the twelve scenarios to investigate the effectiveness of the RTS/CTS mechanism.

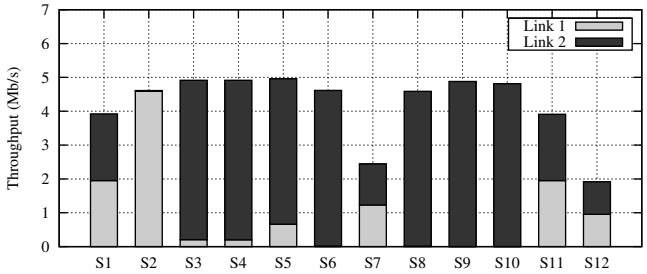
4.2.2 Effectiveness of the RTS/CTS mechanism

We investigate the performance characteristic of each scenario via simulation study. In the simulation, the distance between transmitter and receiver was set 77 meters, for which MCS 1 (6.5 Mb/s) is the best transmission rate. Other simulation settings are the same as in Section 4.4.

Figure 4.2 shows the link throughput of each scenario for the basic access and RTS/CTS modes. In the case of the basic access mode, some of the scenarios show the same behavior, as shown in Fig. 4.2(a). The scenarios `S1`, `S2`, `S6`, `S7`, `S11`, and `S12` can be grouped into a symmetric case because one link plays the same role as the other link in the aspect of transmission and interference. The throughput performance is nearly the same for both links. The link throughput is much higher for `S11` and `S12` compared to the others



(a) Basic access mode (`BASIC`).



(b) RTS/CTS mode (`RTS/CTS`).

Figure 4.2 Link throughput for the twelve hidden node scenarios.

because only ACK transmissions of the receiver interferes with the other link. On the other hands, the scenarios `S3`, `S4`, `S5`, `S8`, `S9`, and `S10` can be grouped into an asymmetric case because `Link1` starves whereas `Link2` greedily occupies the channel. The scenarios `S5` and `S10` are a little bit different from the others because the ACK transmissions of `R2` do not interfere with `Link1`. However, this difference is negligible in our simulation as shown in Fig. 4.2(a), because the ACK transmission time is much shorter than data transmission time.

In the case of the RTS/CTS mode, the performance characteristic becomes more complex compared to the basic access mode, as shown in Fig. 4.2(b). The scenarios `S1`, `S7`, `S11`, and `S12` can be grouped into a symmetric case, because two links are located symmetrically and achieve the similar throughput

performance. The aggregate throughput is larger in S11 and S12 compared to the others because only the ACK transmission of R2 interferes with Link1. The throughput difference between S1 and S7 is due to the NAV setting after the transmission of the CTS frame in the other link. The transmitters in S7 cannot decode the CTS frames transmitted by the other link, and thus, it cannot set the NAV value appropriately, which results in the degradation of throughput. The throughput difference between S11 and S12 can be explained similarly. In S2, the throughput of Link1 is much larger than that of Link2. When the transmission of CTS frame by R1 is successful, T2 can decode the CTS frame and appropriately set the NAV value. However, when the transmission of CTS frame by R2 is successful, T1 cannot decode the CTS frame and cannot set the NAV value appropriately, which degrades the throughput of Link2. It is noted that S6 is the opposite case of S2. The other scenarios S3, S4, S5, S8, S9, and S10 can be grouped into an asymmetric case because Link1 starves whereas Link2 greedily occupies the channel. The throughput of Link1 is larger in S3, S4, and S5 compared to S8, S9, and S10 because the transmission of Link1 can be protected when CTS frame transmissions by R1 are successful. Moreover, the throughput of Link1 is larger in S5 compared to S3 and S4 because it is not interfered with ACK transmissions by R2.

From the simulation results in the above, we can see that the throughput performance in the RTS/CTS mode depends on the transmission and carrier sensing range of the CTS frame. The study in [59] show that the effectiveness of RTS/CTS mechanism depends on the ratio between the radii of carrier sensing range and transmission range. The throughput performance improves as the ratio increases. However, the RTS/CTS mechanism cannot resolve the

hidden node problem effectively, especially in the case of asymmetric cases. The interference from the hidden nodes is inevitable even with the application of the RTS/CTS mechanism as long as a transmitter participates in channel contention via CSMA/CA mechanism.

4.3 Resolving the Hidden Node Problem

In this section, we propose an effective resolution mechanism for the hidden node problem by extending the effective CTS range and introducing the ROC mechanism.

4.3.1 Extending the effective CTS range

The RTS/CTS mechanism was designed to prevent the transmission of the other nodes around a transmitter and a receiver by announcing the information on the upcoming data transmission via the RTS and CTS frames. In particular, the transmission of the CTS frame is effective in preventing the transmission of hidden nodes, because they defer the channel access by setting the NAV value appropriately when overhearing the CTS frame. In the IEEE 802.11 standard [5], the NAV value after receiving the CTS frame is calculated by the duration field in the CTS frame. We define the effective CTS range as the range that the transmission of CTS frame can prevent the transmission of other nodes during the subsequent data and ACK transmissions. Then, the effective CTS range becomes equal to the transmission range of the CTS frame because only a node that can decode the CTS frame appropriately calculates the NAV value. When a neighboring node is located out of the transmission range but within the carrier sensing range of the CTS frame, it cannot decode

the CTS frame but only can sense the transmission of a frame. In this case, the neighboring node defers its channel access by setting extended inter-frame space (EIFS) instead of NAV, and the upcoming data transmission cannot be protected because the duration of EIFS is much shorter than that of NAV. This can be explained by the throughput difference between `S1` and `S7` in Fig. 4.2(b). The transmitters in `S7` only can sense the frame transmissions by the receiver in the other link, and thus, they defer the channel access only for EIFS. Therefore, the subsequent data and ACK frame transmission cannot be protected, which results in a smaller throughput than that of `S1`.

Now, we investigate how to extend the effective CTS range to the carrier sensing range of the CTS frame. For this, two major problems have to be resolved, which are how to identify the CTS frame when a node can only sense the transmission of a frame, and how to set the NAV value to protect the upcoming data transmission. Nandy and Gupta in [60] proposed a method to identify the transmission of a CTS frame by observing the transmission duration of frames. There are three types of control frames, i.e., RTS, CTS, and ACK frames, which are transmitted at the lowest data rate. To identify the control frames, each control frame should have a unique size. However, both the CTS and ACK frames have the same size of 14 bytes in the legacy 802.11a/b/g, and thus, it was proposed to make the CTS frame size of 17 bytes to differentiate from the ACK frame [60]. In the IEEE 802.11n standard, CTS and ACK frames can be differentiated by transmission duration because block ACK frame is much larger than that of CTS frame. After identifying the CTS frame, a node has to defer channel access to protect the upcoming data transmission, but it still does not know the appropriate duration of NAV.

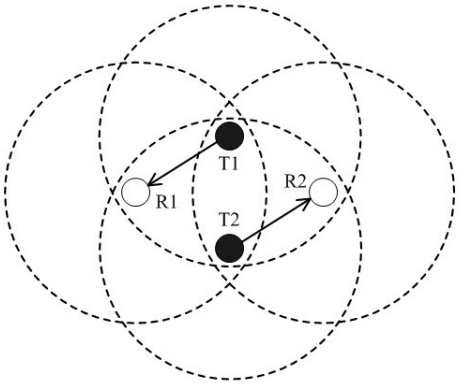


Figure 4.3 The problem of the ROC mechanism. (The transmitters are located within the carrier sensing range each other, but the receivers are not.)

In [60], a node sets the NAV value corresponding to the maximum packet size, which result in a waste of radio resources. We try to find the answer to this problem in the FA scheme. Recall that the FA scheme adjusts the frame size so that the transmission time of data frame, T_{data}, to be close to T_{ref} as much as possible by adjusting the number of packets in a frame. Therefore, a node can defer the channel access appropriately by simply setting the duration of NAV as T_{ref} after identifying the CTS frame, which surely protects the upcoming data transmission.

4.3.2 Receiver-oriented contention (ROC) mechanism

In the previous section, we observed that there is a limitation for the RTS/CTS mechanism in resolving the hidden node problem because the exchange of control packets also suffers from the hidden node problem caused by the carrier sensing mechanism at the transmitter. Let us consider the case where a receiver participates in channel contention with the CSMA/CA mechanism instead of a transmitter. In all the hidden node scenarios except `S5` and `S10`, for example,

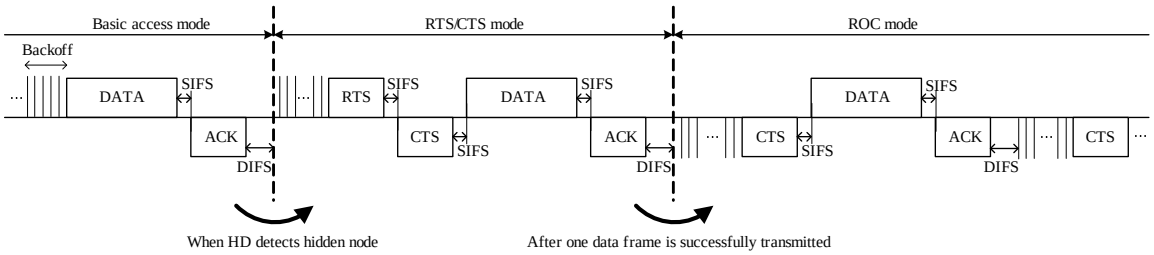


Figure 4.4 Basic operation of the ROC mechanism in conjunction with the HD mechanism.

the CSMA/CA mechanism can eliminate the hidden node problem if `R1` and `R2` participate in channel contention, because they can carrier sense each other. Based on this observation, we propose the ROC mechanism inspired by the basic idea of MACA-BI [31]. Recall that, in MACA-BI, the part corresponding to the RTS frame transmission is suppressed and a receiver polls a transmitter for data transmission via RTR frame, which is the renamed CTS frame. The ROC mechanism can be simply incorporated in IEEE 802.11 system by allowing a receiver to participate in the contention to access the channel with the CSMA/CA mechanism. For this, a receiver has its own back-off counter, decrements the counter by one when the channel is idle, and transmits a CTS frame whenever the counter reaches zero. Upon receiving the CTS frame, the transmitter transmits a data frame to the receiver and subsequently the receiver transmits an ACK frame to the transmitter. The ROC mechanism can save the transmission time of RTS frames and protect the interference from hidden nodes because the receiver senses the channel and contends with its neighboring nodes.

However, there are some disadvantages in this mechanism. The most critical one is that different forms of hidden node problem may occur. Let us consider

the network topology depicted in Fig. 4.3. In the RTS/CTS mode where transmitters contend for channel, T1 and T2 can share the wireless channel fairly because they can carrier sense each other. When the ROC mechanism is applied, however, T1 and T2 may transmit the data frame simultaneously because R1 and R2 are out of carrier sensing range. Moreover, the hidden node problem in some of the scenarios cannot be resolved by the ROC mechanism. In S5 and S10, for example, the CSMA/CA mechanism does not properly operate despite the use of the ROC mechanism because R1 and R2 cannot carrier sense each other. One possible solution to this difficulty is to apply the ROC mechanism adaptively depending on the network condition. We propose to use the ROC mechanism with the HD mechanism, so that only the links that suffer from the hidden node problem use the ROC mechanism. We can explain the basic operation of the ROC mechanism with the HD mechanism by Fig. 4.4. Initially, the link operates in basic access mode and the transmitter detects the presence of a hidden node via the HD mechanism. When the interference from the hidden node is detected, the transmitter initiates the RTS/CTS mechanism to alleviate the interference from the hidden node. It is noted that the RTS/CTS exchange may still not be successful due to the hidden nodes. In such a case, the transmission in the RTS/CTS mode needs to be continued until one data frame transmission is successful. After one successful transmission of a data frame in the RTS/CTS mode, the link switches to transmit in the ROC mode to resolve the hidden node more effectively. The ROC mechanism with HD can resolve the problem that can occur as in Fig. 4.3 because T1 and T2 detect no hidden node, and thus, the transmitter keeps operating in the basic access mode. Also, the hidden node problem in S5 and S10 can be resolved because

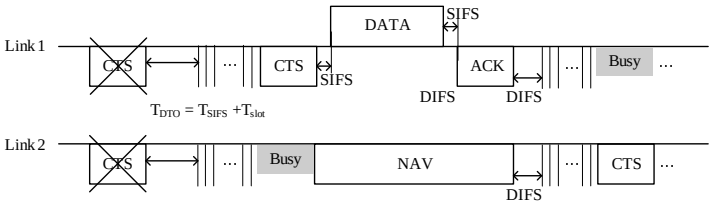
only T1 detects a hidden node, and R1 and T2, which can carrier sense each other, participate in channel contention.

4.3.3 Comments on several issues

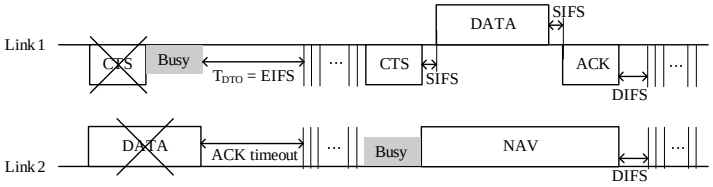
We end this section by discussing the following issues related to the proposed mechanism.

Collisions in the ROC mode

In the ROC mode, when a CTS frame transmission is successful, the subsequent data and ACK frame transmissions are guaranteed to be successful because the neighboring nodes of the receiver defer channel access. However, collisions still occur when the backoff counters of the receiver and some of the neighboring nodes reach zero at the same time. There are two types of collision as in Fig. 4.5; a collision between two receivers operating in the ROC mechanism (CTS-CTS collision) and a collision between a receiver operating in the ROC mode and a transmitter operating in the basic mode (CTS-DATA collision). For each type of collision, it is needed to appropriately set the timeout duration of the data frame, T_{DTO}. When a CTS-CTS collision occurs (Fig. 4.5(a)), both receivers set their T_{DTO} equal to SIFS plus a slot time to reduce unnecessary delay in channel access. When a CTS-DATA collision occurs (Fig. 4.5(b)), the receiver of Link1 sets its T_{DTO} equal to EIFS after the medium become idle, and the transmitter of Link2 sets its T_{DTO} equal to ACK timeout after completing the transmission of data frame. To set the T_{DTO} value properly, a receiver first has to be able to differentiate a CTS-CTS collision from a CTS-DATA collision. If a receiver senses the channel busy after transmitting a CTS



(a) CTS-CTS collision.



(b) CTS-DATA collision.

Figure 4.5 Two types collision in the ROC mode.

frame, it must be a CTS-DATA collision. Otherwise, it is assumed to be a CTS-CTS collision. After a collision, the receiver doubles the CW value to make the collision probability smaller irrespective of the collision type.

Traffic estimation in the ROC mode

In the ROC mode, a receiver serves as a polling node by transmitting the CTS frame to initiate the data transmission at a transmitter. For bursty traffic patterns, however, several CTS frame transmissions can be wasted in an inactive period, which results in performance degradation. To avoid the waste of the CTS frames, a receiver should know the traffic condition at a transmitter. This problem can be easily solved due to the nature of the FA scheme. Recall that the FA scheme performs frame aggregation with packets in the current queue when the number of packets for frame aggregation is insufficient to avoid unnecessary delay resulting from the waiting time of packet arrival. Therefore, T_{data} becomes

much less than T_{ref} at the beginning of the inactive period, which can be used as a signal of finishing the ROC mode. Then, both transmitter and receiver go back to the basic access mode and wait for the next packets.

Effect of the exposed node problem

Another drawback of the CSMA/CA mechanism is the exposed node problem [61,62], which limits spatial reuse by blocking a node to initiate data transmission even though its transmission does not interfere with ongoing transmission. However, it is difficult to resolve the hidden node and exposed node problems simultaneously because the two problems conflict with each other. For example, the RTS/CTS mechanism that helps to alleviate the hidden node problem may make the exposed node problem worse [62]. The proposed mechanism can effectively resolve the hidden node problem while minimizing the effect of the exposed node because it initially operates in the basic access mode and moves to the ROC mode only when a node detects hidden nodes by the HD mechanism. Moreover, the RTS frame transmission is suppressed in the ROC mode which helps to reduce the exposed nodes. We show the performance of the proposed mechanism in presence of the exposed node in Section 4.4.3.

4.4 Performance Evaluation

In this section, we evaluate the performance of the proposed mechanism for resolving the hidden node problem via simulation using the ns-2 simulator [50]. We implemented MAC and PHY layers according to the system model described in Section 3.2. UDP packets were generated by CBR sources so that

the transmitters always had packets to transmit and the frames were transmitted using the MCSs in Table 3.1. Other system parameters related to IEEE 802.11n can be found in [3]. The performances were evaluated and compared for the following five schemes.

- `BASIC` : This is the baseline scheme which operates in the basic access mode.
- `RTS/CTS` : This scheme always employs the RTS/CTS mechanism.
- `RTS/CTS-E` : This scheme always employs the RTS/CTS mechanism with the extended CTS range.
- `ROC-E` : This scheme always employs the ROC mechanism with the extended CTS range.
- `ROC(HD)-E` : This is the proposed scheme in which the ROC mechanism is adaptively employed by the HD mechanism with the extended CTS range.

It is noted that the FA scheme is employed for all scheme for fair evaluation. The packet size and T_{ref} were set to 512 bytes and 1 milliseconds, respectively.

.

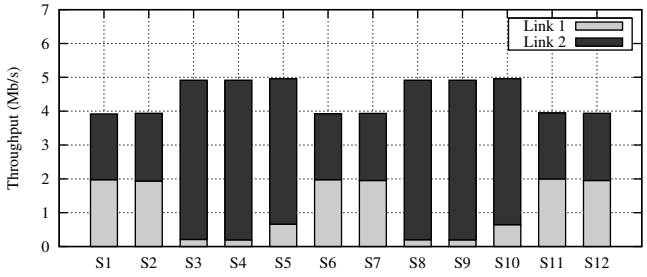
4.4.1 The twelve hidden node scenarios

We first evaluate the throughput performance of the five schemes for the twelve hidden node scenarios in Fig. 4.1. Figure 4.6 shows the link throughput of `RTS/CTS-E`, `ROC-E`, and `ROC(HD)-E`, and the performance of `BASIC` and `RTS/CTS` are already presented in Fig. 4.2. We can see that `RTS/CTS-E` can suppress

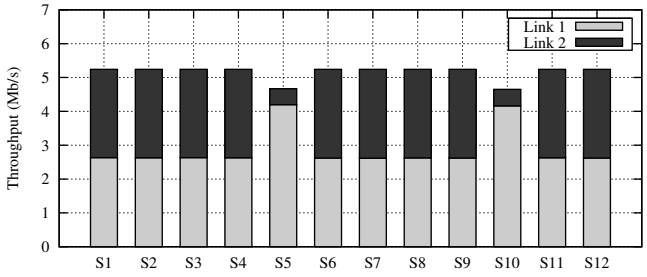
the performance differences in some of the scenarios by extending the effective CTS range. For example, two links come to have nearly the same throughput in the case of S2 and S7 and also in the case of S11 and S12. However, in the asymmetric scenarios such as S3, S4, S5, S8, S9, and S10, the extension of the effective CTS range is insufficient to resolve the hidden node problem. When ROC-E is applied, two receivers participate in channel contention instead of the transmitters, and thus, the hidden node problem for all the scenarios except for S5 and S10 can be resolved because two receivers can carrier sense each other. Therefore, the throughputs of two links are almost the same, which implies that the CSMA/CA mechanism operates wall and two links fairly share the channel. In S5 and S10, however, the receivers cannot carrier sense each other, and thus, the hidden node problem is still left unsolved. When ROC(HD)-E is employed, only the link that suffers from the hidden node uses the ROC mechanism thanks to the HD mechanism. In the symmetric scenarios such as S1, S2, S6, S7, S11, and S12, both receivers participate in channel contention and the throughput of each link is similar to the corresponding throughput of ROC-E. In the asymmetric scenarios such as S3, S4, S5, S8, S9, and S10, R1 and T2 participate in channel contention and thus all of the hidden node problem can be resolved because R1 and T2 can carrier sense each other. From the simulation result, we can see that ROC(HD)-E can effectively resolve all of the possible hidden node scenarios that can occur by the interaction of two flows.

4.4.2 Line topologies of three flows and double ring topologies

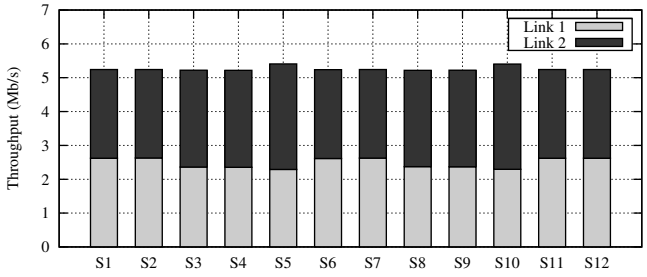
To show the effectiveness of the proposed mechanism for resolving the hidden node problem in general network topologies, we performed simulation for vari-



(a) `RTS/CTS-E`



(b) `ROC-E`



(c) `ROC(HD)-E`

Figure 4.6 Link throughput of `RTS/CTS-E`, `ROC-E`, and `ROC(HD)-E` in the twelve hidden node scenarios.

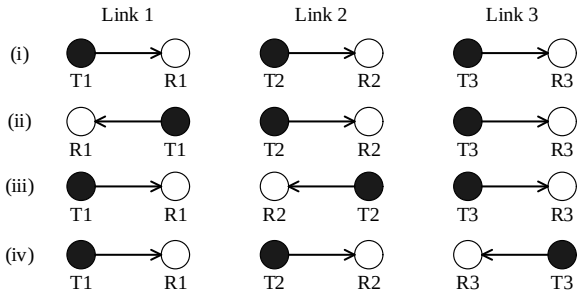


Figure 4.7 Four kinds of line topologies of three flows.

ous network topologies. We first tested the five schemes for the line topologies of three flows as shown in Fig. 4.7. It is noted that there are four different flow patterns depending on the direction of the data transmission and they are all asymmetric. Table 4.1 shows the link throughput and aggregate throughput of each scheme for the four topologies in Fig. 4.7, and we summarize the result as the following.

- **Line topology in Fig. 4.7(i)**: The relations between `Link1` and `Link2`, and `Link2` and `Link3` are the same as `S3` (or `S4`) in Fig. 4.1(a). Therefore, `Link1` and `Link2` suffer from the hidden node problem, and they have very low throughput in `BASIC`. In particular, due to the nature of the CSMA/CA mechanism, the link located in the middle (i.e., `Link2`) necessarily suffers severe starvation. There is no performance difference between `RTS/CTS` and `RTS/CTS-E` because the extension of the effective CTS range is insufficient to resolve the hidden node problem in such an asymmetric network. In the case of `ROC-E`, the receivers participate in channel contention, and the adjacent receiver can carrier sense each other. Therefore, the hidden node problem can be resolved and each link

Table 4.1 Link throughput of the five schemes in the line topologies of three flows (in Mb/s).

Topology	Scheme	Link1	Link2	Link3	Aggregate
Line (i)	BASIC	1.225	0.089	15.069	16.383
	RTS/CTS	6.664	0.319	13.227	20.210
	RTS/CTS-E	6.578	0.320	13.221	20.119
	ROC-E	11.071	0.374	10.456	21.901
	ROC(HD)-E	11.459	0.422	11.856	23.737
Line (ii)	BASIC	14.841	0.001	14.233	29.075
	RTS/CTS	12.249	0.060	13.460	25.769
	RTS/CTS-E	8.937	0.219	13.315	22.471
	ROC-E	11.528	0.123	11.868	23.519
	ROC(HD)-E	14.661	0.459	14.646	29.766
Line (iii)	BASIC	14.320	0.002	14.201	28.523
	RTS/CTS	12.828	0.077	12.684	25.589
	RTS/CTS-E	12.553	0.206	12.011	24.770
	ROC-E	10.743	0.310	10.385	21.438
	ROC(HD)-E	13.619	0.554	14.572	28.745
Line (iv)	BASIC	0.908	1.129	0.845	2.882
	RTS/CTS	7.299	1.010	7.790	16.099
	RTS/CTS-E	11.789	0.246	11.813	23.848
	ROC-E	11.721	1.546	10.945	24.212
	ROC(HD)-E	11.724	1.544	10.956	24.224

throughput shows symmetric behavior. When ROC(HD)-E is applied, R1, R2, and T3 participate in channel contention, and thus, the performance behavior is similar with ROC-E. The aggregate throughput is somewhat larger than that of ROC-E because Link3 operates in the basic access mode.

- **Line topology in Fig. 4.7(ii)**: The relation between Link1 and Link2 is the same as the two links in Fig. 4.3, whereas the relation between Link2 and Link3 is the same as S3 (or S4) in Fig. 4.1(a). Therefore, only Link2 suffers from the hidden node problem and it have much lower throughput compared to other links in BASIC. We can see that the extension of CTS

range contributes to increase the throughput for Link2 when comparing RTS/CTS and RTS/CTS-E. When ROC(HD)-E is applied, T1, R2 and T3 participate in channel contention, which enables R2 and T3 to carrier sense each other and resolves the hidden node problem.

- **Line topology in Fig. 4.7(iii)**: The relation between Link1 and Link2 is the same as S7 in Fig. 4.1(b), whereas the relation between Link2 and Link3 is the same as the two links in Fig. 4.3. Therefore, Link1 and Link2 suffers from the hidden node problem. As in the case of (ii), we notice that the extension of the effective CTS range contributes to increase the throughput for Link2 by comparing RTS/CTS and RTS/CTS-E. When ROC(HD)-E is applied, R1, R2 and T3 participate in channel contention, which enables R1 and R2 to carrier sense each other and resolves the hidden node problem.

- **Line topology in Fig. 4.7(iv)**: The relation between Link1 and Link2 is the same as S3 (or S4) in Fig. 4.1(a), whereas the relation between Link2 and Link3 is the same as S7 in Fig. 4.1(b). Therefore, all links suffer from the hidden node problem and BASIC gives poor throughput performance. We can see that the extension of CTS range contributes to increase the throughput of Link1 at the expense of the throughput of Link2 by comparing RTS/CTS and RTS/CTS-E. For ROC-E and ROC(HD)-E, all receivers participate in channel contention, and thus, they show similar performance. The hidden node problem can be resolved because the adjacent receiver can carrier sense each other.

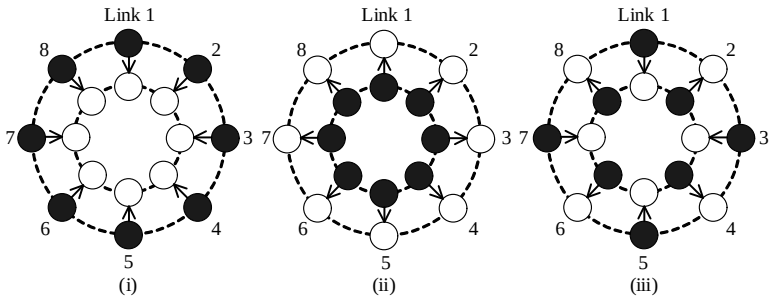


Figure 4.8 Three kinds of double ring topologies.

For each case in Fig. 4.7, `ROC(HD)-E` makes the links in the asymmetric topology have the symmetric behavior, and the hidden node problem can be resolved. `ROC(HD)-E` has the highest aggregate throughput for each line topology of three flows; however, the links in the middle starves, which is inevitable as long as we use DCF as the basic channel access mechanism.

We then tested the five schemes in the double ring topologies as shown in Fig. 4.8. There are three kinds of flow patterns depending on the direction of the data transmission. The radius of inner and outer circle are 15 and 80 meters, respectively. Then, the link distance is 40 meters where the MCS 3 (19.5 Mb/s) is the best transmission rate. Table 4.1 shows the link throughput and aggregate throughput of each scheme for the three topologies in Fig. 4.8, and we summarize the result as the following.

- **Double ring topology in Fig. 4.8(i)** : The transmitters and receivers are uniformly located in the outer and inner circles, respectively. Transmission of each link is interfered by hidden nodes, and thus, `BASIC` shows poor performance. Although `RTS/CTS` somewhat alleviates the effect of the hidden node as shown in Table 4.1, the improvement of throughput is

not much. For each link, the data and ACK frame transmissions cannot be protected even when the RTS/CTS exchange is successful because the link on the opposite side can only sense but cannot decode the CTS frame. `RTS/CTS-E` gives three times higher throughput than that of `RTS/CTS` because the extension of the effective CTS range helps to appropriately set the NAV value even though a node can only sense the CTS frame. When `ROC-E` or `ROC(HD)-E` applied, the receivers participate in channel contention instead of the transmitters. This effectively resolve the hidden node problem and all the links fairly share the channel, because all the receivers can carrier sense each other.

- **Double ring topology in Fig. 4.8(ii)**: In contrast to the topology (i), the transmitters and receivers are uniformly located in the inner and outer circle, respectively. There are no hidden nodes for all the links because all the transmitters can carrier sense each other, and thus, all links fairly share the channel when `BASIC`, `RTS/CTS`, or `RTS/CTS-E` is applied as shown in Table 4.1. In the case of `ROC-E`, however, the receivers unnecessarily participate in channel contention instead of the transmitters, which results in another form of the hidden node problem. This causes a significant throughput degradation similar to the case of `RTS/CTS` in topology (i). When `ROC(HD)-E` is applied, the transmitters keep on participating in channel contention and operate in basic access mode because the HD mechanism does not detect the hidden node. `ROC(HD)-E` is similar to `BASIC` and is somewhat higher compared to `RTS/CTS` and `RTS/CTS-E` in throughput, because the transmitters does not use the RTS/CTS ex-

change.

- **Double ring topology in Fig. 4.8(iii)**: The topology (iii) is a mixture of (i) and (ii), where half of the transmitters transmit the data frames inwards and the other half of the transmitters transmit the data frames outwards. `Link1`, `Link3`, `Link5`, and `Link7` suffer from the hidden node problem and the others do not. `BASIC` shows poor performance due to the hidden node problem. In the case of `RTS/CTS`, the throughput of `Link1`, `Link3`, `Link5`, and `Link7` are improved compared to `BASIC`, but are still much smaller than that of the other links as shown in Table 4.1 because the RTS/CTS mechanism cannot effectively resolve the hidden node problem. When `RTS/CTS-E` is applied, the extension of the effective CTS range contributes to increase the throughputs of `Link1`, `Link3`, `Link5`, and `Link7`. In the case of `ROC-E`, the receivers participate in channel contention, and thus, `Link2`, `Link4`, `Link6`, and `Link8` suffer from the hidden node problem. When `ROC(HD)-E` is applied, only the receivers of `Link1`, `Link3`, `Link5`, `Link7` participate in channel contention. Therefore, the hidden node problem can be effectively resolved and all the links can fairly share the channel, because the receivers of `Link1`, `Link3`, `Link5`, and `Link7` and the transmitters of `Link2`, `Link4`, `Link6`, and `Link7` can carrier sense each other.

In the case of double ring topologies, `ROC(HD)-E` makes the node in the inner circle participate in channel contention, whether it is a transmitter or receiver, with the help of the HD mechanism. Therefore, the hidden node problem can be effectively resolved and all the links in the network can fairly share the channel

Table 4.2 Link throughput of the five schemes in the double ring topologies (in Mb/s).

Topology	Scheme	Link1	Link2	Link3	Link4	Link5	Link6	Link7	Link8	Aggregate
Double ring (i)	BASIC	0.013	0.016	0.014	0.015	0.014	0.013	0.015	0.015	0.115
	RTS/CTS	0.263	0.258	0.256	0.263	0.259	0.259	0.263	0.258	2.079
	RTS/CTS-E	0.742	0.796	0.771	0.791	0.767	0.767	0.780	0.766	6.180
	ROC-E	1.751	1.782	1.748	1.756	1.749	1.733	1.768	1.744	14.031
	ROC(HD)-E	1.741	1.754	1.762	1.728	1.771	1.797	1.744	1.743	14.040
Double ring (ii)	BASIC	1.790	1.788	1.823	1.843	1.861	1.872	1.816	1.834	14.627
	RTS/CTS	1.663	1.692	1.732	1.662	1.685	1.682	1.703	1.767	13.586
	RTS/CTS-E	1.721	1.703	1.719	1.741	1.687	1.695	1.675	1.644	13.585
	ROC-E	0.343	0.335	0.333	0.338	0.334	0.323	0.336	0.326	2.668
	ROC(HD)-E	1.789	1.857	1.892	1.883	1.867	1.878	1.788	1.819	14.773
Double ring (iii)	BASIC	0.045	0.409	0.042	0.456	0.050	0.386	0.046	0.436	1.870
	RTS/CTS	0.177	2.054	0.172	1.910	0.198	1.995	0.190	2.007	8.703
	RTS/CTS-E	0.509	2.163	0.512	2.104	0.502	2.227	0.488	2.128	10.633
	ROC-E	1.999	0.254	1.990	0.240	2.002	0.229	2.024	0.244	8.982
	ROC(HD)-E	1.743	1.753	1.717	1.782	1.735	1.735	1.724	1.740	13.929

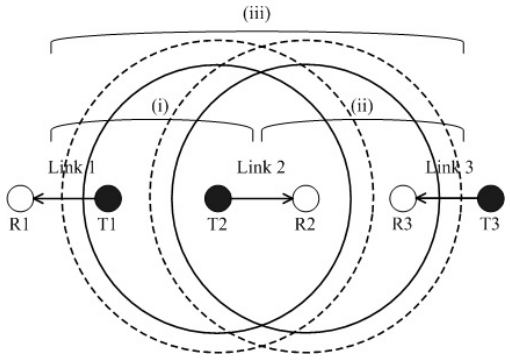


Figure 4.9 Three kinds of scenarios of the exposed node problem.

because all the nodes in the inner circle can carrier sense each other.

4.4.3 Performance of ROC(HD)-E in presence of exposed nodes

In this section, we show the performance of `ROC(HD)-E` in presence of exposed nodes to verify that `ROC(HD)-E` can effectively resolve the hidden node problem while minimizing the effect of the exposed node problem. Table 4.3 shows the link throughput and aggregate throughput of each scheme for the topologies in Fig. 4.9, which shows the three kinds of scenarios of the exposed node problem. We summarize the result as the following.

- **Exposed node problem in Fig. 4.9 (i)** : In this scenario, `Link1` and `Link2` are assumed to be active. Two links can simultaneously transmit data frames because data transmission of each transmitter does not interfere the frame reception of the link on the opposite side. In the basic access mode, `RTS/CTS`, and `RTS/CTS-E`, however, each link occupies only half of the channel because `T1` and `T2` can carrier sense each other and transmit data when the channel is idle. When `ROC-E` is applied, `T1` and `T2` can simultaneously transmit data frame because R1 and R2 cannot

carrier sense each other. However, CTS or ACK frame transmission by the receiver interferes the data transmission of the link on the opposite side, which results in lower throughput than that of the basic access mode. `ROC(HD)-E` is similar to the basic access mode in throughput because there are no hidden nodes and thus it maintains the basic access mode.

- **Exposed node problem in Fig. 4.9 (ii)** : In this scenario, `Link2` and `Link3` are assumed to be active. The hidden node and exposed node problems coexist in this case. The ACK transmission of a receiver interferes the data frame reception of the link on the opposite side, which results in the hidden node problem. When `R2` overhears the CTS transmission of `R1` on the opposite side, it cannot reply the RTS transmission by `T2`, which results in the exposed node problem. As shown in the basic access mode, the link throughput becomes less compared to the result in (i) because of the hidden node problem. When `RTS/CTS-E` is applied, the exposed node problem becomes dominant due to the use of the CTS frame, which results in less throughput compared to the basic access mode. When `ROC-E` or `ROC(HD)-E` is applied, `R1` and `R2`, which can carrier sense each other, participate in channel contention, and thus, fairly share the channel.

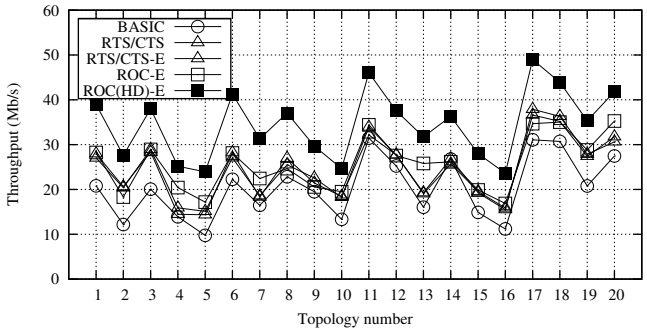
- **Exposed node problem in Fig. 4.9 (iii)** : In this scenario, all links are assumed to be active. We can see that the use of RTS/CTS makes the throughput of `Link2` lower because of the exposed node problem. When `ROC-E` is applied, `Link1` and `Link2` can transmit data frame simultaneously, and thus, the throughput of `Link2` increases compared to `RTS/CTS` and `RTS/CTS-E`. When `ROC(HD)-E` is applied, `T1`, `R2`, and `R3` participate

Table 4.3 Link throughput of the five schemes in presence of exposed nodes (in Mb/s).

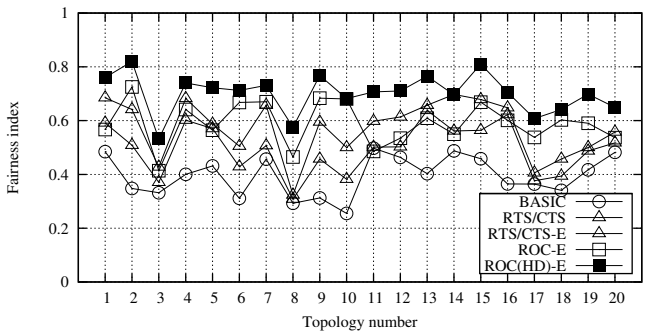
Topology	Scheme	Link1	Link2	Link3	Aggregate
Exposed node (i)	BASIC	7.831	7.727	-	15.558
	RTS/CTS	6.974	6.934	-	13.908
	RTS/CTS-E	6.812	6.994	-	13.806
	ROC-E	6.967	7.063	-	14.030
	ROC(HD)-E	7.795	7.818	-	15.613
Exposed node (ii)	BASIC	-	6.311	6.409	12.720
	RTS/CTS	-	3.240	3.174	6.414
	RTS/CTS-E	-	3.172	3.190	6.362
	ROC-E	-	7.328	7.357	14.685
	ROC(HD)-E	-	7.314	7.362	14.676
Exposed node (iii)	BASIC	8.590	3.037	11.391	23.018
	RTS/CTS	9.566	0.658	12.083	22.307
	RTS/CTS-E	9.967	0.625	11.978	22.570
	ROC-E	10.921	2.285	12.584	25.790
	ROC(HD)-E	12.144	3.844	11.768	27.756

in contention. Unlike ROC-E, Link1 operates in the basic access mode in ROC(HD)-E, which results in the increase of throughput.

The use of the RTS/CTS mechanism significantly degrades throughput performance in presence of exposed node whereas it contributes to throughput enhancement in presence of hidden nodes. Therefore, appropriate use of the RTS/CTS mechanism is required and the HD mechanism helps to this. Recall that ROC(HD)-E initially operates in the basic access mode and moves to the ROC mode only when a node detects hidden nodes. Consequently, ROC(HD)-E gives at least similar throughput performance with either BASIC or ROC-E in presence of the exposed node as shown in Table 4.3. Moreover, it can achieve much higher throughput performance compared to the other schemes in general network environment where a large number of hidden and exposed nodes



(a) Aggregate throughput.



(b) Airtime fairness.

Figure 4.10 Aggregate throughput and airtime fairness for the five schemes in twenty random topologies.

coexist, which is shown in the subsequent subsection.

4.4.4 Random topologies

Finally, we tested the five schemes in random topologies. We consider an ad hoc network where thirty links are randomly located in a square area of 200m×200m. In such a complex network topology, a number of hidden and exposed nodes may coexist and they result in significant degradation of network performance.

Figure 4.10(a) shows the aggregate throughput of the five schemes in the

random ad hoc networks. We tested the five schemes with twenty different random topologies, and the figure shows the result for each topology. Although the RTS/CTS mechanism may make the exposed node problem worse, `RTS/CTS` gives larger throughput compared to `BASIC`. This is because the RTS/CTS mechanism not only alleviates the hidden node problem but also reduces the waste of channel resource when collisions occur in such a dense network topology. `RTS/CTS-E` gives a higher throughput than `RTS/CTS` because it can properly protect an ongoing transmission once the transmission of CTS frame is sensed. We can see that `ROC-E` shows higher throughput than `RTS/CTS` and `RTS/CTS-E` in several cases because there may be quite a few hidden nodes and the ROC mechanism is more effective to resolve the hidden node problem. However, it sometimes gives similar or less throughput compared to `RTS/CTS` and `RTS/CTS-E`, which means that the adaptive use of the ROC mechanism is required. `ROC(HD)-E`, in which only the nodes that are affected by hidden nodes operate in the ROC mode, shows the highest throughput, and the throughput improvements over `BASIC`, `RTS/CTS`, `RTS/CTS-E`, and `ROC-E` are 78%, 45%, 43% and 36%, respectively, on average.

Figure 4.10(b) shows the Jain's fairness index of the five schemes in the random ad hoc networks to show the airtime fairness among nodes. It is noted that the FA scheme gives ideal fairness performance when all the nodes in the network can carrier sense each other. However, the fairness index may not a good performance index in such an ad hoc network because the carrier sensing area of each node may differ from each other, and thus, an appropriate airtime usage for each node also should be differ from each other. Nevertheless, we compared the fairness index of the five schemes to show the degree of star-

vation due to hidden nodes, which can be an indication whether the proposed mechanism can effectively resolve the hidden node problem. From Fig. 4.10(b), we can see that `ROC(HD)-E` shows the highest fairness index compared to the other schemes because it can effectively prevent starvation due to hidden node by appropriately initiating the ROC mode.

4.5 Chapter Summary

In this chapter, we proposed a resolution mechanism for the hidden node problem by enhancing the capability of the RTS/CTS mechanism. Through a simulation study for the twelve hidden node scenarios, we have demonstrated that the poor network performance caused by the RTS/CTS mechanism is due to the limitation of the effective CTS range and the carrier sensing mechanism at the transmitter-side. We first extended the effective CTS range to the carrier sensing range of the CTS frame with the help of the FA scheme that makes the transmission time each node almost the same. We then proposed the ROC mechanism to compensate the defect of the carrier sensing mechanism at the transmitter-side. The HD mechanism was employed to adaptively operate in the ROC mode. The proposed mechanism effectively resolve the hidden node problem, which leads to improvement of the network throughput and fairness.

Chapter 5

Conclusions

The DCF and the CSMA/CA mechanism in IEEE 802.11 WLANs cause the defective phenomenons known as performance anomaly and hidden node problem, respectively. The performance anomaly restricts any possible advantage of using higher data rates and degrades overall network throughput even though all nodes can carrier sense each other and DCF provides fair chances to access the channel. The interference from hidden nodes becomes a serious problem because of the increase of the number of overlapped BSS and ad hoc networks. In such a complex network environment, the CSMA/CA mechanism is likely to fail, leading to significant performance degradation.

The performance anomaly can be resolved by enforcing fairness in terms of channel access time referred to as airtime fairness. In this context, we proposed the FA scheme by employing the generalized two-level frame aggregation scheme in IEEE 802.11n WLANs, to achieve airtime fairness and enhance overall network throughput without modifying DCF. The FA scheme tightly regulates the

airtime of each node around a target time and thus effectively resolves the performance anomaly problem that arises due to different data rates and packet sizes among nodes. At the same time, the FA scheme adjusts the frame size so that the throughput can be maximized by considering the channel error rate, data rate, and MAC overhead. We formulated the problem of frame aggregation as an optimization problem and proposed a simple and effective method to reduce the computational complexity.

As the first step to resolve the hidden node problem, we proposed the HD mechanism to adaptively use the RTS/CTS mechanism that contains the protocol overhead. The HD mechanism, which does not need to modify the IEEE 802.11n standard, detects hidden nodes well in a practical network environment where frames can be lost due to any combination of collisions, hidden nodes, and channel impairments. To enhance the capability of the RTS/CTS mechanism, we proposed an enhanced mechanism for resolving hidden node problem by adaptively using the ROC mechanism in conjunction with the HD mechanism. Also, we extend the effective CTS range by employing the FA scheme that makes the transmission time of all nodes almost the same.

The simulation results confirm that the `ROC(HD)-E` scheme, which employs all the proposed mechanism in this dissertation, can effectively resolve the performance anomaly and hidden node problem, and thus, the network throughput and fairness can be significantly improved. The advantage of `ROC(HD)-E` is that it can be used without modifying the DCF and does not depend on the channel access probability. This opens a possibility of adopting more advanced MAC protocol that adjusts the channel access probability depending on the network environment to improve the network throughput and fairness.

Bibliography

[1] R. Berezdivin, R. Breinig, and R. Topp, "Next-generation wireless communications concepts and technologies," *IEEE Communications Magazine*, vol. 40, no. 3, pp. 108–116, March 2002.

[2] L. Munoz, R. Agero, J. Choque, J. Lrastorza, L. Sánchez, M. Petrova, and P. Mahonen, "Empowering next-generation wireless personal communication networks," *IEEE Communications Magazine*, vol. 42, no. 5, pp. 64–70, May 2004.

[3] "IEEE standard for information technology - Telecommunications and information exchange between systems - Local and metropolitan area networks - Specific requirements part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications amendment 5: Enhancements for higher throughput," *IEEE Std. 802.11n*, Oct. 2009.

[4] M. Heusse, F. Rousseau, G. Berger-Sabbatel, and A. Duda, "Performance anomaly of 802.11b," in *Proceedings of IEEE International Conference on Computer Communication (INFOCOM)*, April 2003.

[5] "Part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications," *IEEE Std. 802.11-2007*, 2007.

[6] S. Khurana, A. Kahol, and A. Jayasumana, "Effect of hidden terminals on the performance of IEEE 802.11 MAC protocol," in *Proceedings of IEEE Local Computer Networks (LCN)*, Oct. 1998.

[7] G. Tan and J. Guttag, "Time-based fairness improves performance in multi-rate WLANs," in *Proceedings of USENIX*, June 2004.

[8] I. Tinnirello and S. Choi, "Temporal fairness provisioning in multi-rate contention-based 802.11e WLANs," in *Proceedings of IEEE International Symposium on World of Wireless Mobile and Multimedia Networks (WOWMOM)*, June 2005.

[9] H. Kim, S. Yun, I. Kang, and S. Bahk, "Resolving 802.11 performance anomalies through QoS differentiation," *IEEE Communications Letters*, vol. 9, no. 7, pp. 655–657, July 2005.

[10] C. Chou, K. Shin, N. Shankar *et al.*, "Contention-based airtime usage control in multirate IEEE 802.11 wireless LANs," *IEEE/ACM Transactions on Networking*, vol. 14, no. 6, pp. 1179–1192, Dec. 2006.

[11] A. Banchs, P. Serrano, and H. Oliver, "Proportional fair throughput allocation in multirate IEEE 802.11e wireless LANs," *Wireless Networks*, vol. 13, no. 5, pp. 649–662, Oct. 2007.

[12] T. Joshi, A. Mukherjee, Y. Yoo, and D. Agrawal, "Airtime fairness for IEEE 802.11 multirate metworks," *IEEE Transactions on Mobile Computing*, vol. 7, no. 4, pp. 513–527, April 2008.

[13] H. Lee and C.-H. Choi, "Achieving airtime fairness and maximum throughput in IEEE 802.11 under various transmission durations," *IEICE Transactions on Communications*, vol. E94-B, no. 11, pp. 3098–3106, Nov. 2011.

[14] "Part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications. Amendment 8: Medium access control (MAC) quality of service enhancements," *IEEE Std. 802.11e-2005*, 2005.

[15] Y. Lin and V. Wong, "Frame aggregation and optimal frame size adaptation for IEEE 802.11n WLANs," in *Proceedings of IEEE Global Telecommunications Conference (GLOBECOM)*, Dec. 2006.

[16] B. Ginzburg and A. Kesselman, "Performance analysis of A-MPDU and A-MSDU aggregation in IEEE 802.11n," in *Proceedings of IEEE Sarnoff Symposium*, May 2007.

[17] B. S. Kim, H. Y. Hwang, and D. K. Sung, "Effect of frame aggregation on the throughput performance of IEEE 802.11n," in *Proceedings of IEEE Wireless Communications and Networking Conference (WCNC)*, March 2008.

[18] X. He, F. Y. Li, and J. Lin, "Link adaptation with combined optimal frame size and rate selection in error-prone 802.11n networks," in *Proceedings of IEEE International Symposium on Wireless Communication Systems (ISWCS)*, Oct. 2008.

[19] T. Li, Q. Ni, D. Malone, D. Leith, Y. Xiao, and T. Turletti, "Aggregation with fragment retransmission for very high-speed WLANs," *IEEE/ACM Transactions on Networking*, vol. 17, no. 2, pp. 591–604, April 2009.

[20] B. Sadeghi, V. Kanodia, A. Sabharwal, and E. Knightly, "Opportunistic media access for multirate ad hoc networks," in *Proceedings of ACM International Conference on Mobile Computing and Networking (MobiCom)*, Sept. 2002.

[21] D. Skordoulis, Q. Ni, H. Chen, A. Stephens, C. Liu, and A. Jamalipour, "IEEE 802.11n MAC frame aggregation mechanisms for next-generation high-throughput WLANs," *IEEE Wireless Communications*, vol. 15, no. 1, pp. 40–47, Feb. 2008.

[22] C.-Y. Wang and H.-Y. Wei, "IEEE 802.11n MAC enhancement and performance evaluation," *Mobile Networks and Applications*, vol. 14, no. 6, pp. 760–771, Dec. 2009.

[23] M. Lampe, H. Rohling, and W. Zirwas, "Misunderstandings about link adaptation for frequency selective fading channels," in *Proceedings of IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Sept. 2002.

[24] S. Tsai and A. Soong, "Effective-SNR mapping for modeling frame error rates in multiple-state channels," *3GPP2-C30-20030429-010*, April 2003.

[25] G. Martorell, F. Riera-Palou, and G. Femenias, "Cross-layer fast link adaptation for MIMO-OFDM based WLANs," *SPRINGER Wireless Personal Communications*, pp. 599–609, Feb. 2011.

[26] T. Jensen, S. Kant, J. Wehinger, and B. Fleury, "Fast link adaptation for MIMO OFDM," *IEEE Transactions on Vehhicular Technology*, vol. 59, pp. 3766–3778, Oct. 2010.

[27] C. Yeo, "Improving IEEE 802.11n link adaptation using coherence time estimation," *Master Thesis*, Seoul National University, Seoul, Korea, Feb. 2010.

[28] V. Erceg *et al.*, "TGn channel models," *IEEE 802.11 document 03/940r4*, May 2004.

[29] L. Schumacher and B. Dijkstra, "Description of a MATLAB implementation of the indoor MIMO WLAN channel model proposed by the IEEE 802.11 TGn channel model special committee," *Implementation note version*, vol. 5, 2004.

[30] L. Schumacher, "WLAN MIMO channel MATLAB program," *URL: http://www.info.fundp.ac.be/∼lsc/Research/IEEE_80211 _HTSG_CMSC/distribution_terms.html*, 2003.

[31] F. Talucci, M. Gerla, and L. Fratta, "MACA-BI (MACA by invitation)-A receiver oriented access protocol for wireless multihop networks," in *Proceedings of IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Sept. 1997.

[32] Y. Wang and J. J. Garcia-Luna-Aceves, "A new hybrid channel access scheme for ad hoc networks," *ACM Wireless Networks Journal, Special Inssue on Ad Hoc Networking*, vol. 10, no. 4, April 2004.

[33] L. Du and L. Chen, "Receiver initiated network allocation vector clearing method in WLANs," in *Proceedings of IEEE Asia-Pacific Conference on Communications (APCC)*, June 2005.

[34] G. Holland, N. Vaidya, and P. Bahl, "A rate-adaptive MAC protocol for multi-hop wireless networks," in *Proceedings of ACM International Conference on Mobile Computing and Networking (MobiCom)*, July 2001.

[35] A. Kameraman and L. Monteban, "WaveLAN 2: A high-performance wireless LAN for the unlicensed band," *Bell Labs Technical Journal*, vol. 2, no. 3, pp. 118–133, Aug. 2002.

[36] J. Kim, S. Kim, S. Choi, and D. Qiao, "CARA: Collision-aware rate adaptation for IEEE 802.11 WLANs," in *Proceedings of IEEE International Conference on Computer Communication (INFOCOM)*, April 2006.

[37] S. H. Y. Wong, H. Yang, S. Lu, and V. Bharghavan, "Robust rate adaptation for 802.11 wireless networks," in *Proceedings of ACM International Conference on Mobile Computing and Networking (MobiCom)*, Sept. 2006.

[38] J. Choi, J. Na, K. Park, and C.Kim, "Collision-aware design of rate adaptation for multi-rate 802.11 WLANs," *IEEE Journal of Selected Areas in Communications*, vol. 26, no. 8, pp. 1366–1375, Oct. 2008.

[39] Q. Pang, S. Liew, and V. Leung, "Design of an effective loss-distinguishable MAC protocol for 802.11 WLAN," *IEEE Communications Letters*, vol. 9, no. 9, pp. 781–783, Sept. 2005.

[40] M. Kim, Y. Kim, and C. Choi, "Rate adaptation by estimating channel quality in IEEE 802.11 wireless LAN," *IEICE Transactions on Communications*, vol. E95-B, no. 1, pp. 243–253, Jan. 2012.

[41] S. Choi, K. Park, and C.Kim, "On the performance characteristics of WLANs: revisited," in *Proceedings of ACM SIGMETRICS*, June 2005.

[42] Y. Kim, J. Yu, S. Choi, and K. Jang, "A novel hidden station detection mechanism in IEEE 802.11 WLAN," *IEEE Communications Letters*, vol. 10, no. 8, pp. 608–610, Aug. 2006.

[43] K. Nishide, H. Kubo, R. Shinkuma, and T. Takahashi, "Detecting hidden terminal problems in densely deployed wireless networks," in *Proceedings of IEEE Global Communications Conference (GLOBECOM)*, Dec. 2010.

[44] H. Ma, J. Zhu, and S. Roy, "On loss differentiation for CSMA-based dense wireless network," *IEEE Communications Letters*, vol. 11, no. 11, pp. 877–879, Nov. 2007.

[45] T. Rappaport, *Wireless communications: principles and practice*. Upper Saddle River, NJ:Prentice-Hall, 1996.

[46] X. Li and Q. A. Zeng, "Capture effect in the IEEE 802.11 WLANs with Rayleigh fading, shadowing, and path loss," in *Proceedings of IEEE International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, June 2006.

[47] T. Li, Q. Ni, D. Malone, D. Leith, Y. Xiao, and T. Turletti, "Aggregation with fragment retransmission for very high-speed WLANs," *IEEE/ACM Transactions on Networking*, vol. 17, no. 2, pp. 591–604, April 2009.

[48] T. Li, Q. Ni, and Y. Xiao, "Investigation of the block ACK scheme in wireless ad-hoc networks," *WILEY Journal of Wireless Communications and Mobile Computing*, vol. 6, no. 6, pp. 877–888, Sept. 2006.

[49] Q. Ni, T. Li, T. Turletti, and Y. Xiao, "Saturation throughput analysis of error-prone 802.11 wireless networks," *WILEY Journal of Wireless Communications and Mobile Computing*, vol. 5, no. 8, pp. 945–956, Dec. 2005.

[50] NS2 Network Simulator, http://www.isi.edu/nsnam/ns.

[51] P. Ng, S. Liew, K. Sha, and W. To, "Experimental study of hidden node problem in IEEE 802.11 wireless networks," *Sigcomm Poster*, 2005.

[52] K. Hung and B. Bensaou, "Distributed rate control and contention resolution in multi-cell IEEE 802.11 WLANs with hidden terminals," in *Proceedings of ACM International Symposium on Mobile Ad hoc Networking and Computing (MobiHoc)*, Sept. 2010.

[53] A. Tsertou and D. Laurenson, "Revisiting the hidden terminal problem in a CSMA/CA wireless network," *IEEE Transactions on Mobile Computing*, vol. 7, no. 7, pp. 817–831, July 2008.

[54] S. Razak, V. Kolar, and N. Abu-Ghazaleh, "Modeling and analysis of two-flow interactions in wireless networks," *ELSEVIER Ad Hoc Networks*, vol. 8, no. 6, pp. 564–581, Aug. 2010.

[55] Y. Kim and C.-H. Choi, "Analysis of the IEEE 802.11 back-off mechanism in presence of hidden nodes," *IEICE transactions on communications*, vol. 92, no. 4, pp. 1291–1299, April 2009.

[56] P. Karn, "MACA - A new channel access method for racket radio," in *Proceedings of ARRL/CRRL Amateur radio computer networking conference*, Sept. 1990.

[57] R. Singh and D. Lobiyal, "Performance modeling of slotted MACA-BI MAC protocol for mobile ad hoc networks," in *Proceedings of ACM International Conference on Interaction Sciences: Information Technology, Culture and Human*, Dec. 2009.

[58] Z. Haas and J. Deng, "Dual busy tone multiple access (DBTMA)- A multiple access control scheme for ad hoc etworks," *IEEE Transactions on Communications*, vol. 50, no. 6, pp. 975–985, June 2002.

[59] K. Xu, M. Gerla, and S. Bae, "Effectiveness of RTS/CTS handshake in IEEE 802.11 based ad hoc networks," *ELSEVIER Ad Hoc Networks*, vol. 1, no. 1, pp. 107–123, 2003.

[60] Z. Li, S. Nandi, and A. Gupta, "ECS: An enhanced carrier sensing mechanism for wireless ad hoc networks," *ELSEVIER Computer Communications*, vol. 28, no. 17, pp. 1970–1984, Oct. 2005.

[61] D. Shukla, L. Chandran-Wadia, and S. Iyer, "Mitigating the exposed node problem in IEEE 802.11 ad hoc networks," in *Proceedings of IEEE International Conference on Computer Communications and Networks (ICCCN)*, Oct. 2003.

[62] J. Sobrinho, R. de Haan, and J. Brázio, "Why RTS-CTS is not your ideal wireless LAN multiple access protocol," in *Proceedings of IEEE Wireless Communications and Networking Conference (WCNC)*, March 2005.

초록

최근 스마트폰, 태블릿 PC 등의 모바일 기기 사용이 급증함에 따라 무선 랜 (wireless local area network (WLAN)) 을 통한 인터넷 사용이 증가하고 있다. 본 논문에서는 IEEE 802.11n 표준 기반 무선 랜의 두 가지 문제점인 비정상적 성능 저하 현상 (performance anomaly) 과 히든 노드 문제 (hidden node problem) 에 대해 다루었다. 무선 랜 노드들은 일반적으로 서로 다른 전송 속도와 패킷 (packet) 크기를 가지지만, IEEE 802.11 표준의 MAC (medium access control) 프로토콜인 DCF (distributed coordination function) 는 이에 대한 고려 없이, 각 노드들에게 채널 접근 기회를 공평하게 제공하도록 설계되었다. 따라서, 노드 간 채널 점유 시간의 공평성 (airtime fairness) 이 저하되고, 높은 전송 속도 또는 작은 패킷 크기를 가지는 노드들의 처리량 (throughput) 성능이 제한되어 네트워크의 전체 처리량 성능이 크게 감소하는 현상이 생기는데, 이를 비정상적 성능 저하 현상이라 한다. 히든 노드 문제는 송신 노드에서 수행하는 CSMA/CA (carrier sensing multiple access with collision avoidance) 메커니즘이 근본적인 원인이다. 송신 노드는 주변의 간섭 (interference) 신호를 미리 감지하고 간섭 신호가 없을 때 데이터 프레임 (frame) 을 전송한다. 그러나 송신 노드가 감지할 수 없는 수신 노드 주변의 간섭 신호가 데이터 프레임의 전송을 방해 할 수 있으며, 이를 히든 노드 문제라 한다. 최근 무선 랜 서비스 제공자와 개인 사용자들이 무선 랜 접속 장치 (access point(AP)) 를 사전계획 없이 자유롭게 설치함으로써 기본 서비스 영역 (basic service set (BSS)) 간의 채널 간섭이 심해짐에 따라 히든 노드 문제가 더욱 빈번하게 발생하고 있다.

비정상적 성능 저하 현상을 해결하기 위해, 본 논문에서는 IEEE 802.11n

표준의 프레임 결합 (frame aggregation)을 활용하여 각 노드들의 전송 속도와 패킷 크기에 따른 최적의 프레임 결합 방법에 대해 연구하였다. 두 종류의 프레임 결합 방법인 A-MSDU (aggregated MAC service data unit) 와 A-MPDU (aggregated MAC protocol data unit) 의 장점을 모두 활용할 수 있는 2단계 프레임 결합을 사용하여, 채널 점유 시간 공평성을 보장하면서 처리량 성능을 최대화하기 위한 목적 함수를 세우고, 이를 기반으로 한 최적화 문제를 세웠다. 세워진 최적화 문제의 최적의 해를 수학적으로 유도하고, 이를 바탕으로 목적 함수를 달성하기 위한 최적의 프레임 크기와 이를 실시간으로 조절하는 방법 (frame size adaptation (FA) scheme) 을 제안하였다. 제안된 프레임 사이즈 조절 방법은 IEEE 802.11 표준과 쉽게 호환이 가능하며 DCF를 수정할 필요가 없는 장점이 있다. 시뮬레이션을 통해, 제안된 프레임 크기 조절 방법이 기존의 해결 방법과 비교하여 네트워크의 처리량 성능을 크게 향상 시키면서 노드 간 채널 점유 시간을 공평하게 제공함을 보였다.

히든 노드 문제는 RTS/CTS (request to send/clear to send)와 같은 컨트롤 (control) 프레임을 교환함으로써 완화시킬 수 있지만 히든 노드가 없는 상황에서는 오버헤드 (overhead) 에 의한 성능 저하를 겪게 된다. 즉, 히든 노드가 있는 상황에서만 RTS/CTS 메커니즘을 사용하는 것이 효과적이므로, 이를 위해서는 히든 노드를 감지할 수 있는 메커니즘이 필요하다. 본 논문에서는 IEEE 802.11n 표준의 새로운 기능인 프레임 결합, block ACK 등을 활용하여 프레임 전송 실패의 세 가지 원인 (전송 충돌, 채널 에러, 히든 노드) 중 히든 노드에 의한 전송 실패를 감지하는 방법 (hidden node detection (HD) mechanism) 을 제안하였다. 프레임 결합과 block ACK의 사용으로 인해 프레임 전송 실패의 유형을 전체 프레임 손실 (entire frame loss) 과 부분 프레임 손실 (partial frame loss) 로 분류할 수 있다. 전체 프레임 손실에 히든 노드가 미치는 확률을 MAC 계층에서 측정할 수 있는 값들을 통해 추정하고, 부분 프레임 손실에 히든 노

드가 미치는 확률은 block ACK의 비트맵 필드 (bitmap field) 를 통해 추정하였다. 추정된 확률값을 바탕으로, 히든 노드에 의해 처리량 성능이 저하될 가능성이 높은 노드들만 RTS/CTS를 사용하도록 함으로써 불필요한 오버헤드를 줄이는 동시에 히든 노드 문제를 완화시킬 수 있다. 시뮬레이션을 통해, 제안된 히든 노드 감지 메커니즘이 다양한 환경에서 히든 노드의 존재를 효과적으로 감지하며, 이에 따라 RTS/CTS를 선별적으로 사용함으로써 네트워크의 처리량 성능이 증가됨을 보였다.

마지막으로, 히든 노드 문제를 좀 더 효과적으로 해결하기 위해 RTS/CTS 메커니즘의 성능을 개선 하였다. 우선, 다양한 히든 노드 문제 상황에 대한 시뮬레이션을 통해 RTS/CTS 메커니즘의 문제점이 CTS 프레임의 제한된 유효 거리와 송신단에서 행해지는 CSMA/CA 메커니즘 때문이라는 것을 보였다. 제안된 히든 노드 문제 해결 메커니즘은 프레임 크기 조절 방법을 활용하여 CTS 프레임의 유효 거리를 확장하고, 히든 노드 감지 메커니즘을 활용하여 히든 노드가 감지 되었을 때 수신 노드에서 대신 CSMA/CA를 수행하는 ROC (receiver oriented contention) 메커니즘 사용함으로써 송신단에서 행해지는 CSMA/CA 메커니즘의 문제점을 보완하였다. 시뮬레이션을 통해, 제안된 메커니즘이 히든 노드 문제를 효과적으로 해결해 주고, 이에따라 네트워크 전체 처리량 성능과 채널 점유 시간 공평성이 향상됨을 보였다.

주요어 : 무선 랜, IEEE 802.11n 표준, RTS/CTS, 프레임 결합, 비정상적 성능 저하 현상, 채널 점유 시간 공평성, 히든 노드 문제, 히든 노드 감지
학번 : 2007-20936