A Multiple Relay-Based Medium Access Control Protocol in Multirate Wireless Ad Hoc Networks with Multiple Beam Antennas

Yuh-Shyan Chen¹, Chih-Shun Hsu², and Po-Ta Chen³

¹Department of Computer Science and Information Engineering
National Taipei University, Taiwan, R.O.C.
E-mail: yschen@csie.ntpu.edu.tw

²Department of Information Management
Shih Hsin University, Taiwan, R.O.C.

³Department of Computer Science and Information Engineering
National Chung Cheng University, Taiwan, R.O.C.

Abstract

The advanced technique of multiple beam antennas is recently considered in wireless networks to improve the system throughput by increasing spatial reuse, reducing collisions, and avoiding co-channel interference. The usage of multiple beam antennas is similar to the concept of Space Division Multiple Access (SDMA), while each beam can be treated as a data channel. Wireless networks can increase the total throughput and decrease the transmission latency if the physical layer of a mobile node can support multirate capability. Multirate wireless networks incurs the anomaly problem, because low data rate hosts may influence the original performance of high data rate hosts. In this work, each node fits out multiple beam antennas with multirate capability, and a node can either simultaneously transmit or receive multiple data on multiple beams. Observe that the transmitting or receiving operation does not happen at the same time. In this paper, we propose a multiple relay-based medium access control (MAC) protocol to improve the throughput for low data rate hosts. Our MAC protocol exploits multiple relay nodes and helps the source and the destination to create more than one data channel to significantly reduce the transmission latency. Observe that low data rate links with long-distance transmission latencies are distributed by multiple relay nodes, hence the anomaly problem can be significantly alleviated. In addition, the ACK synchronization problem is solved to avoid the condition that source nodes do not receive ACKs from destination nodes. An adjustment operation is presented to reduce unnecessary relay nodes during the fragment burst period. Finally, simulation results illustrate that our multiple relay-based MAC protocol can achieve high throughput and low transmission latency.

Index Terms: multiple beam antennas, multirate, relay, medium access control, wireless ad hoc networks.

I. INTRODUCTION

A wireless ad hoc network (MANET) [3] is made up of identical mobile nodes, each node with a limited wireless transmission range to communicate with neighboring nodes. The MANET can be applied to military, disaster, industry, and mobile learning system [2] due to its self-organizing and adaptive features. Wireless communication is unreliable due to its low channel quality, high noise interference, high
packet collision, and high packet loss rate. The novel design of Medium Access Control (MAC) protocol can significantly improves the reliability and throughput of wireless transmissions. The IEEE 802.11 standard [5] is the most well-known MAC protocol for designing a MANET.

In the IEEE 802.11 standard, Distributed Coordination Function (DCF) and Point Coordination Function (PCF) are defined in the MAC sub-layer [5]. DCF supports the contention-based media access for infrastructure/ad hoc networks and PCF provides contention-free media access for infrastructure networks. The Carrier Sense Multiple Access/Collision Avoidance (CSMA/CA) mechanism is used in DCF to avoid collisions and hidden terminal problem [20]. Clear Channel Assessment (CCA) function at the physical layer is used to detect the busy or idle state of the wireless medium. To solve the hidden terminal problem and avoid collision, RTS (Request to Send)/CTS (Clean to Send) message exchange is developed in IEEE 802.11 to set the network allocation vector (NAV). Consequently, the CCA function and RTS/CTS handshake mechanism are the key operations of the CSMA/CA mechanism. In addition, IEEE 802.11 supports the fragment burst to increase the throughput since control messages, such as RTS and CTS messages, are overheads for each data transmission. Existing IEEE 802.11 networks assume that each mobile node is equipped with an omnidirectional antenna [4]. Many research results [4][9] point out that the property of omnidirectional antennas leads to high collisions and low throughput in MANETs. A mobile node equipped with an omnidirectional antenna always listens to signals from all directions and sends signals to all directions. Because of that, the beam width of an omnidirectional antenna is 360°. The drawback of the omnidirectional antenna is that a mobile node in the transmission mode may collides and/or interferes with other unconcerned mobile nodes in the reception mode. The deserved throughput is insidiously degraded. The drawback of the omnidirectional antenna [8][21] can be significantly alleviated by using the directional antenna or multiple beam antennas. It is surely that existing IEEE 802.11 MAC protocol can not be directly applied to MANET with directional antenna or multiple beam antennas. To increase the spatial reuse, avoid co-channel interference and reduce packet collision, many existing protocols were recently designed to consider the models of directional antennas [10][14][15] and multiple beam antennas [11][12][16][17]. Especially, the use of multiple beam antennas can improve the throughput if multiple beam antennas simultaneously communicate with multiple neighboring nodes [11]. This work investigates a new MAC protocol under the same model of multiple beam antennas in [11].

The rate-adaptive MAC protocol is recently considered in [6] to dynamically adjust the power level and/or transmission rate to improve the network throughput and the transmission latency. IEEE 802.11 a/b/g standard [5] also define the physical layer to support multirate capability to improve the throughput of wireless networks. Unfortunately, the standard does not explicitly specify algorithms to utilize multirate capability. Multirate wireless networks incurs the anomaly problem [6], because low data rate hosts may influence the original performance of high data rate hosts. The channel occupying time of low data rate hosts is usually longer than that of high data rate hosts. Low data rate hosts seize the available channel time of high data rate hosts. More recently, relay-based MAC protocols [19][22][23] is designed to significantly reduce the anomaly problem. With the rate-adaptive/multirate capability, relay-based MAC protocol is developed with the assistance of a relay node to speed up the transmission time.
In this paper, we propose a multiple relay-based medium access control (MAC) protocol to improve the throughput for low data rate hosts. In this investigation, each node is equipped with multiple beam antennas with multirate capability, and a node can either simultaneously transmit or receive multiple data on multiple beams. Observe that the transmitting or receiving operation does not happen at the same time. Our MAC protocol exploits multiple relay nodes and helps the source and the destination to create more than one data channels to significantly reduce the transmission latency. Observe that low data rate links with long-distance transmission latencies are distributed by multiple relay nodes, hence the anomaly problem can be significantly alleviated. In addition, the ACK synchronization problem is solved to avoid the condition that source nodes do not receive ACKs from destination nodes. An adjustment operation is presented to reduce unnecessary relay nodes during the fragment burst period. Finally, simulation results illustrate that the proposed multiple relay-based MAC protocol achieves the result of high throughput and low transmission latency.

The rest of this paper is organized as follows. Section II discusses related works. The system model, basic idea, and challenges are described in section III. Section IV presents the details of the multiple relay-based MAC protocol. Performance evaluation is examined in section V. Finally, section VI concludes this paper.

II. RELATED WORK

This section introduces some related works [7][11][13][18][19][22][23] about multiple beam antennas, multirate, and relay-based MAC protocols as follows.

A. Protocols of Multiple Beam Antennas

Fig. 1 gives examples of the omnidirectional antenna, the directional antenna, and multiple beam antennas, respectively. The omnidirectional antenna transmits to or receives from all directions. Mobile hosts with an omnidirectional antenna easily suffer from interferences or collisions. To avoid the drawbacks of the omnidirectional antenna, the models of the directional antenna and multiple beam antennas are developed, because they have novel properties of increasing spatial reuse, avoiding co-channel interference and reducing collisions. Multiple beam antennas can simultaneously communicate with multiple nodes by exploiting Space Division Multiple Access (SDMA) concept.

The model of multiple beam antennas has the novel property of concurrent communications with multiple nodes. The IEEE 802.11 DCF-based MAC protocol is only suitable for the omnidirectional antenna model. Jain et al [11] proposed a cross-layer MAC protocol, called Explicit Synchronization via Intelligent Feedback (ESIF), to utilize the advantages of multiple beam antennas. Fig. 2 illustrates the basic ESIF operation. Figs. 2 (a) and (b) gives the transmission details of many-to-one communication and one-to-many communication, respectively. Observe that ESIF removes the contention window because that contention window is the main obstruction of utilizing the advantages of multiple beam antennas. To solve collisions by removing the contention window, each node must hear special RTS/CTS, RTS with Intelligent Feedback (RIF), CTS with Intelligent Feedback (CIF), and schedule (SCH) which are
Fig. 1. Antenna types: (a) the omni-directional antenna, (b) the directional antenna, and (c) multiple beam antennas.

Fig. 2. Basic operations of ESIF protocol of (a) many-to-one communication and (b) one-to-many communication.
specially designed by ESIF. This result can conjecture transmission schedule of all neighbor nodes. In the cross-layer design of ESIF, the MAC layer of each mobile node additionally acquires the routing information from the network layer to avoid collisions and obtain better judgements. The model of ESIF is the multiple beam antennas and single data rate capability.

B. Multirate Protocols

It is known that carrier sensing range and transmission range are mainly determined by the factor of transmission power and data rate. Example is given in Fig. 3(a). When a mobile node is in the carrier sensing range of a source node $S$, it means that the mobile node can correctly hear signals but cannot correctly decode signals sent from $S$. If a mobile node is in the transmission range, the mobile node can correctly hear and decode signals from $S$. Fig. 3(a) shows when node $N_1$ transmits a frame to nodes $N_2$, $N_3$ and $N_4$. Nodes $N_2$ and $N_4$ can correctly receive and decode the frame, but node $N_3$ cannot correctly decode the frame since node $N_3$ is in carrier sensing range of node $N_1$. Fig. 3 (b) illustrates the transmission range and carrier sensing range in multirate wireless networks. The usage of higher transmission power can increase the carrier sensing range and the transmission range. Given a fixed transmission power and a carrier sensing range, it is observed that the transmission range decreases as the data rate increases. The transmission range is usually the half of the carrier sensing range.

Kamerman et al. [13] proposed the auto rate fallback (ARF) protocol. If two consecutive ACKs are not received by a source node, the source node decreases the data rate for the next transmission. If consecutive ten ACKs are successfully received by the source node, the source node increases the data rate for the next transmission. Holland et al. [7] proposed a receiver-based auto rate (RBAR) protocol to accurately estimate the data rate. The receiver measures the signal strength of the RTS frame, and then inserts proper transmission rate into CTS frame according to the estimated SNR, so as to notify the source node. Finally, the source node transmits a frame at the proper transmission rate to the receiver node. In addition, RBAR modifies the original 802.11 RTS/CTS frames and MAC header of the DATA frame to set the proper NAV.
Sadeghi et al. [18] proposed an opportunistic auto rate (OAR) protocol to improve RBAR. OAR exploits the concept of the fragment burst to improve the throughput of high data rate nodes. Unfortunately, these schemes can not solve the anomaly problem [6] incurred by the multirate wireless network.

C. Relay-based MAC Protocols

Zhu et al. [22] recently proposed a relay-enabled medium access control protocol, namely rDCF. rDCF performs the triangular handshake to improve the throughput of lower data rate nodes in multirate MANETs. According to the triangular handshake, the destination chooses direct transmission or relay and notifies the source. Zou et al. [23] proposed a relay-aided media access (RAMA) protocol to improve DCF-based MAC protocol, RBAR. RAMA is similar to rDCF. The source chooses direct transmission or relay according to RTS/CTS handshake information and a relay list is obtained from a previous invitation frame of the relay node. More recently, K. Tan et al. [19] proposed CODE protocol to further improve the single relay-base MAC protocol, rDCF and RAMA, by using multiple relay nodes. The concept of multiple relay nodes is originated by lantern which is proposed by Chen et al. in [3]. The lantern is constructed by multiple relay nodes between a pair of source and destination. Chen et al. utilize the lantern to develop a QoS on-demand routing protocol in MANETs [3]. However, all of the mentioned relay-based MAC and routing protocols assume that a node is equipped with an omnidirectional antenna. No multiple relay-based MAC protocol is designed for the multirate MANETs with multiple beam antennas. The contribution of this work is to develop a new multiple relay-based MAC protocol in multirate MANETs with multiple beam antennas.

III. PRELIMINARY

In this section, the system model of multiple relay-based MAC protocol is described. The basic idea of the proposed scheme is then presented. The challenge of our work is finally explained.

A. System Model

The system model mainly follows the model in [11]. The model in [11] is multiple beam antennas with fixed rate capability. The model of this work is multiple beam antennas with multirate capability. Each mobile node forms $M$ fixed-directional and non-overlapping beams, each beam can simultaneously transmits its data by a different data rate. That is, a mobile node can at most simultaneously communicate with $M$ neighboring nodes by $M$ beams. The angle of each beam is $360°/M$ and the beam shape is assumed to be conical. A node can either simultaneously transmit or receive multiple data on multiple beams. Observe that the transmitting and receiving operations do not happen at the same time. The interferences of side and back lobes are ignored in this work. Observe that the carrier sensing range and transmission range of multiple beam antennas are the same as that of an omnidirectional antenna. The wireless channel is assumed to be ideal and symmetric, and received power gains are equal in both directions. In the multiple beam antennas, the data rate information is obtained by precisely calculating the angle of arrival of the received signal and the estimated SNR of the received power gain. Our work is
based on ESIF protocol [11] to support the parallel multiple-communications. The physical layer transmits data by an assigned data rate requested by the MAC sublayer. The physical layer provides the channel status, received power gain, and SNR to the MAC layer. To avoid the hidden terminal problem, control frames are transmitted by the lowest data rate. Data frames can be transmitted by any data rates supported from the physical layer. To simplify our discussion throughout this work, multiple beam antennas are supposed to have eight beams, and each beam supports the data rates of 1, 2, 5.5, and 11 Mbps.

**B. Basic Idea**

Given a pair of source and destination nodes, $S$ and $D$, let a node $R$ be a neighboring node of $S$ and $D$, $R$ is regarded as a relay node if

$$\frac{1}{\text{Rate}_{SR}} + \frac{1}{\text{Rate}_{RD}} < \frac{1}{\text{Rate}_{SD}}$$
where \( Rate_{SD}, Rate_{SR}, Rate_{RD} \) are data rates between \( S \) and \( D \), \( S \) and \( R \), and \( R \) and \( D \), respectively. The basic idea of our approach is described as follows.

As mentioned before, the technique of multiple beam antennas is similar to the technique of Space Division Multiple Access (SDMA). \( M \) beams in multiple beam antennas can be treated as \( M \) individual data channels. Only one direct data channel exists between \( S \) and \( D \). With the assistance of relay nodes, more extra data channels can be utilized except for the direct data channel if relay nodes exist between other \( M - 1 \) beams of \( S \) and \( D \). These relay nodes establish extra data channels between \( S \) and \( D \), \( S \) sends frames to \( D \) via relay nodes. With multiple relay nodes, a data frame of \( S \) is divided into many sub-frames, and each sub-frame is distributed to disjoint data channels established by different relay nodes. The transmission latency is significantly reduced if \( S \) has enough extra data channels for the data transmission. The transmission latency is about \( 1/N \) of the original transmission latency if \( S \) has one direct data channel and \( N - 1 \) extra data channels with relay nodes to \( D \).

Fig. 4 illustrates the basic idea of the multiple relay-based MAC protocol. Assume that \( S \) sends a frame with \( L \) bits to \( D \), let nodes \( R, R_1, R_2 \) and \( R_3 \) be relay nodes between nodes \( S \) and \( D \). \( T_a \) is the transmission latency of direct transmission between \( S \) and \( D \). \( T_b \) is the transmission latency with one relay node \( R \). \( T_c \) is the transmission latency by two disjoint paths, one is the direct transmission between \( S \) and \( D \) and another is the path passed through a relay node \( R \). \( T_d \) is the transmission latency by three disjoint paths, these paths are passed through relay nodes \( R_1, R_2 \) and \( R_3 \). Examples are given in Figs. 4 (a), (b), (c), and (d).

\[
T_a = \frac{L}{Rate_{SD}} \quad (1)
\]

\[
T_b = \frac{L}{Rate_{SR}} + \frac{L}{Rate_{RD}} \quad (2)
\]

\[
T_c = \max \left( \frac{L}{2 \times Rate_{SD}} + \frac{L}{2 \times Rate_{SR}} + \frac{L}{2 \times Rate_{RD}} \right) \quad (3)
\]

\[
T_d = \max \left( \frac{L}{3 \times Rate_{SR_1}} + \frac{L}{3 \times Rate_{SR_2}} + \frac{L}{3 \times Rate_{SR_3}} + \frac{L}{3 \times Rate_{R_1D}} + \frac{L}{3 \times Rate_{R_2D}} + \frac{L}{3 \times Rate_{R_3D}} \right) \quad (4)
\]

We can observe that \( T_d < T_c < T_b < T_a \). The motivation of this work is that the transmission latency of low data rate link can be significantly reduced by distributing a frame into multiple disjoint paths through different relay nodes. Consequently, the anomaly problem can be alleviated by exploiting multiple relay nodes in multirate MANET with multiple beam antennas.

C. Challenge

Two challenges are considered in a multirate MANET with multiple beam antennas; one is the ACK synchronization problem and another one is the unnecessary relay problem. Fig. 5 indicates the ACK synchronization problem. In this paper, each node uses multiple beam antennas with multirate capability,
Fig. 5. The ACK synchronization problem: (a) many-to-one multirate communication and (b) one-to-many multirate communication.

Fig. 6. The unnecessary relay problem: (a) original direct transmission, (b) helpless relay, and (c) helpful relay.
the node can simultaneously communicate with multiple neighboring nodes. But all beams must have an uniform communication mode at the same time. Different beam has different communication latency by different data rate and thus \( S \) may not be able to correctly receive an ACK from \( D \) before occurring the ACK timeout even if the data frame is correctly received by \( D \).

Fig. 5 (a) illustrates the ACK synchronization problem for many-to-one multirate communications. Nodes \( S_1 \) and \( S_2 \) simultaneously send equal-sized frame to \( D \), where \( \text{Rate}_{S_1D} < \text{Rate}_{S_2D} \). Node \( S_2 \) cannot receive ACK\(_{S_2} \) message from \( D \) until the time \( t \) when ACK\(_{S_1} \) message sending from \( D \) to \( S_1 \). If ACK\(_{S_2} \) timeout occurs before time \( t \), \( S_2 \) supposes that \( D \) does not correctly receive the frame, node \( D \) will re-submit the current data frame at the next transmission. This is because that \( D \) is still receiving the frame from node \( S_1 \); \( D \) cannot change from the reception mode into transmission mode unless \( D \) finishes receiving frames from \( S_1 \). However, \( S_2 \) can fully understand that \( D \) correctly received the frame if node \( S_2 \) waits for a long ACK timeout time to receive ACK\(_{S_2} \). But it may increase the transmission latency.

Fig. 5 (b) gives a similar ACK synchronization problem for one-to-many multirate communications.

Fig. 6 presents the unnecessary relay problem of many-to-one communication. Suppose that nodes \( S_1 \) and \( S_2 \) simultaneously transmit the equal-sized frames (\( L \) bits) to node \( D \), where \( \text{Rate}_{S_1D} < \text{Rate}_{S_2D} \). Figs. 6 (b) and (c) illustrate that node \( R \) is a relay node of both high data rate link and low data rate link. Based on Figs. 6 (a), (b), and (c), three reception latencies are denoted as \( T_{a'} \), \( T_{b'} \), and \( T_{c'} \) as follows.

\[
T_{a'} = \max\left( \frac{L}{\text{Rate}_{S_1D}}, \frac{L}{\text{Rate}_{S_2D}} \right)
\]

\[
T_{b'} = \max\left( \frac{L}{\text{Rate}_{S_1D}}, \frac{L}{\text{Rate}_{S_2R}} + \frac{L}{\text{Rate}_{RD}} \right)
\]

\[
T_{c'} = \max\left( \frac{L}{\text{Rate}_{S_1R}} + \frac{L}{\text{Rate}_{RD}}, \frac{L}{\text{Rate}_{S_2D}} \right)
\]

, where \( T_{a'} \) is the reception latency of node \( D \) without the assistance of relay node. It is obvious that both \( T_{b'} \) and \( T_{c'} \) are the reception latency of node \( D \) with the assistance of relay nodes. Observe that \( T_{c'} < T_{a'} = T_{b'} \). The total reception latency of node \( D \) is constrained by node \( S_1 \), it is observed that reducing the reception latency of node \( S_2 \) is useless. It is better to reduce the reception latency of node \( S_1 \) with the assistance of relay node \( R \) as shown in Fig 6 (c). In addition, there is a similar problem for one-to-many communication. Unfortunately, nodes \( S_1 \) and \( S_2 \) cannot know this situation in advance. It may waste the communication resources. This condition can be improved under the situation that nodes \( D, S_1, \) and \( S_2 \) have the burst data transmission. To overcome the unnecessary relay problem, our new MAC protocol is developed by notifying node \( S_2 \) that relay node \( R \) is useless as shown in Fig 6.

**IV. MULTIPLE RELAY-BASED MAC PROTOCOL**

Our new MAC protocol is based on ESIF protocol [11] to achieve the parallel communication capability. This paper extends the ESIF Network Allocation Vector (ENAV) in [11] by adding a relay node table to keep the information of multiple relay nodes. In addition, control frames, such as RIF, CIF, SCH, and MAC header frame in [11] are modified to support multiple relay nodes. Fig. 7 shows frame formats of our approach. The ”Priority/N” field, which exists in ESIF protocol [11], is used to inform neighboring
node its individual schedule since ESIF protocol removes the contention window. "Type" and "Subtype" fields of Frame Control field can be used to distinguish these control frames. The first three bits of the Information field records the beam number of a mobile node, where each mobile node assumes to have eight beams in this work. Other fields are introduced later.

Our multiple relay-based MAC protocol is divided into three phases as follows.

1. **Relay nodes discovery**: The possible relay nodes is discovered to provide the information of multiple relay nodes in the multiple relay-based MAC protocol.

2. **Multiple relay-based MAC operation**: The kernel operation of the multiple relay-based MAC protocol.

3. **Eliminating unnecessary relay nodes**: The unnecessary relay nodes are eliminated in the fragment bursting.

The detail of the multiple relay-based MAC protocol is presented as follows.

**A. Relay Nodes Discovery**

Given a pair of nodes S and D, node S should keep the information of multiple relay nodes between S and D. This information is obtained by performing the relay nodes discovery. According to Eq. (1), a relay region exists between S and D, where relay nodes exist in the relay region. Assume that each mobile node has \( M \) beams. Given that beam "i" of S and beam "j" of D can aim at each other's directions. The relay region is contained between beams "\( i \)",”(\( i - 2 + M \))%\( M \)+1”, ”(\( i + M \))%\( M \)+1”, of S and beams ”j”, ”(\( j - 2 + M \))%\( M \)+1”, ”(\( j + M \))%\( M \)+1” of D. Without loss of generality, example is given as shown in Fig. 4 (a), if \( i = 1 \) and \( j = 5 \), then relay regions are contained between beams 1, 2, 8 of S and beams 4, 5, 6 of D through this investigation. Node S and D can possibly speculate the positions of relay nodes. Suppose that transmission ranges of 1, 2, 5.5, and 11 Mbps are 250, 186.75, 167.75, and...
Fig. 8. Examples of the relay region: (a) 1, 2, and 5.5 Mbps and (b) 2, 5.5, and 11 Mbps.

120.5 meters, respectively. Using Eq. (1), Fig. 8 (a) shows a relay region whose $\text{Rate}_{SD} = 1$ Mbps. Fig. 8 (b) illustrates a relay region whose $\text{Rate}_{SD} = 2$ Mbps.

1) Relay Nodes Discovery Operation: Assume that $S_i$ and $S_j$ communicate with $D$ by $\text{Rate}_{S_iD}$ and $\text{Rate}_{S_jD}$. The relay nodes discovery is performed to search for the possible relay nodes between $S_i$ and $D$, and $S_j$ and $D$. Example is given in Fig. 9 for $S_1$ and $S_2$ communicate with $D$ by $\text{Rate}_{S_1D}$ and $\text{Rate}_{S_2D}$. Let $S_i(b_1, ..., b_m)$ denotes the node $S_i$ which simultaneously sends specific message through beams $b_1, ..., b_m$, where $1 \leq m < \text{Max beam number}$. Let $S_i(b_1, ..., b_m)$ denotes the node $S_i$ which simultaneously sends specific message through all other beams except for beams $b_1, ..., b_m$, where $1 \leq m < \text{Max beam number}$. The relay nodes discovery operation is stated as follows.

A1. Nodes $S_i(b_1)$ and $S_j(b_1)$ simultaneously send an RIF message to node $D$.

A2. Nodes $S_i(b_2, ..., b_m)$ and $S_j(b_2, ..., b_m)$ simultaneously send a new control frame, called the relay discovery (RD), to attempt to find relay nodes, where the RD message of source node $S$ contains MAC address of the desired destination node $D$ and transmission beam number. Nodes $S_i(b_1, ..., b_m)$ and $S_j(b_1, ..., b_m)$ send the SCH message for the transmission schedule. Example is given that $S_1(1, 2, 8)$ and $S_2(4, 5, 6)$ simultaneously send the RD messages, and $S_1(3, 4, 5, 6, 7)$ and $S_2(1, 2, 3, 7, 8)$ sends the SCH message.

A3. Node $D$ simultaneously receives two RIF messages from $S_i$ and $S_j$ and then estimates $\text{Rate}_{S_iD}$ and $\text{Rate}_{S_jD}$ according to the measured SNR.
A4. Node $D$ fills the information of reception beam number, $Rate_{S_i,D}$, and $Rate_{S_j,D}$ into the CIF message, and then responds the CIF message back to nodes $S_i$ and $S_j$ after waiting for an SIFS period.

A5. Node $D$ sends the RD message to help nodes $S_i$ and $S_j$ to find relay nodes. For example, the RD messages of beams 4 and 6 contain the reception beam number, MAC addresses of $S_1$, $D$, and $Rate_{S_1,D}$, and the RD messages of beams 2 and 8 contain the MAC addresses of $S_2$, $D$, and $Rate_{S_2,D}$. $D(3,7)$ sends an SCH message for the transmission schedule.

A6. Assumed that there is a node $R_X$ between nodes $S_i$ and $D$. Node $R_X$ estimates $Rate_{S_i,R_X}$ and $Rate_{R_X,D}$ and obtains $Rate_{S_1,D}$ from the CIF or RD message from $D$ after hearing the RIF or RD message from $S_i$. Node $R_X$ applies Eq. (1) to check whether $R_X$ is a relay node between $S_i$ and $D$.

A7. After $S_i$, $S_j$, and $D$ finishing the current transmission, relay node $R_X$ unicasts a service advertisement.
14

(SD) message to \(S_i\) and \(S_j\), where the SD message contains \(Rate_{S_i,R_X}\), \(Rate_{RX,D}\), receiving beam of \(R_X\), receiving beam of \(D\), and MAC addresses of \(S_i\), \(D\), and \(R_X\). Node \(S_i\) records the relay node information into the relay table after receiving the SD message. In addition, the relay table of \(S_i\) has to record \(Rate_{S_i,D}\) from the CIF message of \(D\).

2) Solving the ACK Synchronization Problem: Fig. 9 (b) gives an example of the relay node discovery. Node \(S_1\) discovers relay nodes \(R_1\) and \(R_2\), and \(S_2\) discovers relay nodes \(R_3\), \(R_4\), and \(R_5\). Let’s recall the ACK synchronization problem. Suppose \(S_1\) and \(S_2\) send out the same size frames to \(D\) by data rates \(Rate_{S_1,D}\) and \(Rate_{S_2,D}\), respectively. Transmission latencies of \(S_1\) and \(S_2\) is not equal since \(Rate_{S_1,D}\) (1 Mbps) is smaller than \(Rate_{S_2,D}\) (2 Mbps). The ACK synchronization problem occurs if the transmission delay of \(S_1\) is greater than that of \(S_2\). It is possible that \(S_2\) cannot receive the ACK message in-time after sending a data frame and \(S_1\) will still transmit data frame to \(D\).

To solve the ACK synchronization problem, each mobile node should try to let each transmission delay be the maximal transmission delay. The maximal transmission delay is the transmission time of the maximum data frame with the maximal payload. Fig. 9 (c) demonstrates an example. The method of solving the ACK synchronization problem is given as follows.

B1. Assumed that \(S_i\) and \(S_j\) do not know each other’s transmission latencies. Nevertheless, \(S_j\) can receive the ACK message before occurring the ACK timeout if the ACK timeout time of \(S_j\) is the same as that of \(S_i\).

B2. The ACK timeout time of each node should be based on the maximal transmission delay to avoid ACK message not to be received in-time before occurring the ACK timeout. In addition, the NAV durations of RIC, CIF, RD, and SCH messages are set to be the maximal transmission delay.

B3. If data rates of \(S_i\) and \(S_j\) are not the lowest data rate, the NAV duration would be set too large. To solve this problem, other beams of the destination should send the SCH message to reset the NAV durations of all neighboring nodes.

B4. After waiting for an SIFS time period, \(S_i\) and \(S_j\) send a new control frame, called RESET, at all beams after receiving all ACK messages if the NAV duration of neighboring nodes is set too large. Note that, RESET has no duration field because the NAV duration of RESET is always zero.

B. Multiple Relay-Based MAC Operation

After performing the relay nodes discovery, relay node tables of \(S_i\) and \(S_j\) keep all the information of possible relay nodes. The detail of multiple relay-based MAC protocol is given in Fig. 10. This case is many-to-one communication. To reduce the total transmission latency, \(S_i\) and \(S_j\) transmit frames to \(D\) by exploiting relay nodes for each data transmission. Nodes \(S_i\) and \(S_j\) endeavor to use all usable relay nodes from the relay node table. The total communication delay is limited by the maximum of transmission delay of \(S_i\) and \(S_j\), where \(S_i\) and \(S_j\) simultaneously transmit frames to \(D\).

One-to-many communication case may be occurred, and \(S\) simultaneously communicates with multiple destination nodes \(D_i\), where \(i > 1\). The source node \(S\) attempts to allot possible relay nodes to all the destination nodes. To consider the unnecessary relay problem, \(S\) may have high priority to allot usable relay nodes from the relay node tables. Note that \(S\) does not allot any \(D_k\), if \(D_j\) cannot allot at least a
Fig. 10. Example of multiple relay-based MAC protocol and the detailed procedure.
relay node, where \( \text{Rate}_{SD_i} \) is the lowest among all \( \text{Rate}_{SD_k} \) and \( k \leq i \). This is because that the total transmission delay is limited by the lowest data rate node even if transmission delays of other lower data rate nodes can be improved.

Two multiple relay-based MAC operations, namely allotting relay-nodes operation and multiple relay-based frame exchange operation, are shown as follows.

1) Allotting Relay-Nodes Operation: The algorithm of \( S \) allotting proper relay nodes to desired \( D \) is given in Algorithm 1. Before describing the algorithm for one-to-many communication, some notations are defined. \( D \) is a one-dimensional array to record the desired destination nodes \( (D_i, 1 \leq i \leq n) \) of \( S \). \( C \) is a one-dimensional array to record the next hop nodes \( (C_i, 1 \leq i \leq n) \) of \( S \) from \( S \) to \( D_i \). \( C_i \) is \( D_i \) if the beam for \( D_i \) uses direct transmission. \( C_i \) is \( R_x \) if the beam for \( D_i \) uses indirect transmission. Note that, \( C \) is used when there is at most one relay node for each \( D_i \). If there are more relay nodes, \( C' \) is used. Thus, \( C' \) is a two-dimensional array to store more relay nodes of \( D_i \) and \( C'_i \) is a one-dimensional array to store extra relay nodes allotted to \( D_i \), where \( 1 \leq i \leq n \). Note that, \( C'_i = \emptyset \) if \( S \) cannot allot any relay node to \( D_i \). Both \( B \) and \( T \) are one-dimensional arrays. Elements of \( B \) are unassigned free beams for \( D_i \). Element of \( T \) is the transmission latency between \( S \) and \( D_i \). \( T \) is counted by procedure \( \text{CountSource} \). Example is given in Fig. 10, since node \( D \) is initially not allotted any relay node, initial setting of node \( S_2 \) is \( D_1 = " \text{node D} \)” and \( C_1 = " \text{node D} \). \( C'_1 = \emptyset \) and \( B = \{1, 2, 3, 4, 5, 6, 7, 8\} \) because \( S_2 \) can use eight beams to communicate. \( T_1 \) is the direct transmission latency between nodes \( S_2 \) and \( D_1 \). After the initial setting, \( S_2 \) then performs \( \text{AllotRelayNodes} \) procedure. The main operations of \( \text{AllotRelayNodes} \) procedure are explained as follows.

C1. All beams for desired destinations \( D \) are removed since these beams are used to exchange control frames. For instance, \( B - B_5 = \{1, 2, 3, 4, 6, 7, 8\} \). (line 2 of Procedure \( \text{AllotRelayNodes} \))

C2. Let \( T_{SD_i} \) be the transmission latency between \( S \) and \( D_i \), \( 1 \leq i \leq n \), let \( T_{\max} = \max_{i=1..n} T_{SD_i} \). If \( T_{SD_i} = T_{\max} \), \( S \) will try to allot relay nodes . Some temporal parameters are used to store results of allocating relay nodes. (lines 5-6 of Procedure \( \text{AllotRelayNodes} \))

C3. \( S \) performs \( \text{Allocate} \) procedure to allot relay node to \( D_i \). The value of \( \text{True} \) or \( \text{False} \) returned from \( \text{Allocate} \) procedure represents the success or failure of the allocation operation. (line 7 of Procedure \( \text{AllotRelayNodes} \))

C4. Some parameters are needed to be updated by temporal parameters if \( \text{Allocate} \) procedure returns \( \text{True} \) value, where \( \text{cycle} \) is the allocation time. Finally, \( \text{AllotRelayNodes} \) procedure returns \( \text{cycle} \), \( C \), and \( C' \) to \( S \) by executing \( \text{Allocate} \) procedure. (lines 8-10 of Procedure \( \text{AllotRelayNodes} \))

C5. go to step C2.

The operation of \( \text{Allocate} \) procedure is stated as follows.

D1. Repeatedly performing B2-B3 steps for \( |d| \) times.

D2. If \( T_{SD_i} = T_{\max} \) then \( S \) tries to communicate with \( D_i \) through relay nodes by performing B3. (line 4 in Procedure \( \text{Allocate} \))

D3. The source node \( S \) checks its relay node table to allot possible relay node to \( D_i \) by a fixed beam, where the fixed beam is toward \( D_i \). (lines 5-12 in Procedure \( \text{Allocate} \))
D4. The source node $S$ checks its relay node table to allot possible relay node to $D_i$ by the remaining beams. (lines 13-22 in Procedure Allocate)

For instance as shown in Fig. 10, the returned values for $S_1$ are $C_1 = R_4$ and $C'_1 = \{R_3, R_5\}$. It means that $S_2$ divides a data frame into three sub-frames, and use relay nodes $R_3, R_4$, and $R_5$ to transmit these three sub-frames concurrently. Procedure CountSource is shown as follows.

Let $\Gamma$ consist of $c_i$ and $c'_i$, where $c_i$ is the direct beam (data channel) for $D_i$, and $c'_i$ are extra beams (data channels) for $D_i$. The main operations of CountSource procedure are expressed as follows.

E1. Let $k$ represent the number of data channels for $D_i$. The source divides the current data frame into $k$ sub-frames, and each one is simultaneously transmitted to all nodes of $\Gamma_p$. (line 1 of Procedure CountSource)

E2. The transmission delay $t$ is calculated. (line 3 of Procedure CountSource)

E3. The actual transmission delay $t$ is calculated with (or without) relay nodes. (lines 4-8 of Procedure CountSource)

For example in Fig. 10, $S_1$ uses Algorithm 1 to search for relay nodes to communicate with $D$. The returned values for $S_1$ are $C_1 = D$ and $C'_1 = \{R_1, R_2\}$. Similarly, $S_2$ also uses Algorithm 1 to search for relay nodes to communicate with $D$. The returned values for $S_2$ are $C_1 = R_4$ and $C'_1 = \{R_3, R_5\}$, which means that $S_2$ divides its data frame into three sub-frames. These sub-frames are transmitted to $D$ through relay nodes $R_3, R_4$, and $R_5$. Another example is that $S_1$ divides its data frame into three sub-frames, one is through direct communication, and the other two sub-frames are through relay nodes $R_3$ and $R_4$. 

---

**Algorithm 1: Allotting Multiple Relay Nodes for Destinations**

**Input:**

$D = \{D_i \mid 1 \leq i \leq n\}$, where $D_i$ is a desired destination.

$C = \{C_i \mid 1 \leq i \leq n\}$, where

- if $S$ sends to $D_i$ directly: $C_i = D_i$
- if $S$ sends via $R_k$ indirectly: $C_i = R_k$
- if $D_i$ has no extra relay nodes: $C'_i = \emptyset$
- if $D_i$ has extra relay nodes: $C'_i = \{R_{ij} \mid 1 \leq j \leq \gamma\}$

$B = \{B_i \mid 1 \leq i \leq M, n \leq M\}$, where $B_i$ is the beam $i$ of $S$.

$T = \{T_i \mid 1 \leq i \leq n\}$, where $T_i$ is the transmission latency to $D_i$.

**Procedure AllotRelayNodes()**

1. **repeat** = True; cycle = 0;
2. **while** **repeat** == True **do**
3. $T_{max} = \max\{T_1, T_2, ..., T_n\}$
4. $b = B \setminus c = \{C, C'\}; t = T$;
5. **repeat** = Allocate($T_{max}$, $S$, $D$, $b$, $c$, $c'$, $t$);
6. **if** **repeat** == True **then**
7. $B = b; C = c; C' = c'; T = t$;
8. cycle = cycle + 1;
9. **end if**
10. **end while**
11. return (cycle, $C$, $C'$)
12. **end Procedure**
2) **Multiple relay-based frame exchange operation:** Without loss of generality, we consider the many-to-one communication scenario of $S_1$ and $S_j$, who simultaneously communicate with $D$ to illustrate the multiple relay-based frame exchange operation. Example is given in Fig. 10, $S_1$ and $S_2$ have a common destination $D$. After performing Algorithm 1, the return values for $S_1$ are $C_1 = D$ and $C_1' = \{R_1, R_2\}$ and the return values for $S_2$ are $C_1 = R_4$ and $C_1' = \{R_3, R_5\}$. Therefore, we have the scenario of $S_1(1)$ communicates with $D(5)$ by a direct communication, $S_1(2, 8)$ simultaneously communicates with $D(4, 6)$ through relay nodes $R_1$ and $R_2$; $S_2(4, 5, 6)$ simultaneously communicates with $D(2, 1, 8)$ through relay nodes $R_3$, $R_4$ and $R_5$. The multiple relay-based frame exchange operation is explained as follows.

**F1.** Initially, $S_1(b_1)$ and $S_j(b'_1)$ directly sends out the RIF message to $D$. $S_1(b_2, \ldots, b_m)$ send out new control
frames, called *appoint to send* (ATS) message, to relay nodes $R_1, R_2, \cdots, R_{m-1}$, respectively. $S_j(b'_2, b'_3, \cdots, b'_n)$ send out the ATS message to relay nodes $R'_1, R'_2, \cdots, R'_{n-1}$, respectively. The ATS message is used to appoint one neighbor node as a relay node.

F2. For $S_i$ after sending the ATS message for all relay nodes $R_1, R_2, \cdots, R_{m-1}$, relay nodes $R_1, R_2, \cdots, R_{m-1}$ should recognize that $S_i$ asks all relay nodes to relay sub-frames to $D$. For $S_j$ after sending the ATS message for all relay nodes $R'_1, R'_2, \cdots, R'_{n-1}$, each relay node, $R'_q$, $1 \leq q \leq n - 1$, recognizes that its MAC address is in the ATS message.

F3. Node $D$ responds the CIF message to $S_i$ and $S_j$ after $D$ successfully received the RIF message and waiting for an SIFS period. Observe that, the information fields of the CIF message for $S_i$ and $S_j$ have the latest $Rate_{SD}$ and $Rate_{SJ,D}$, respectively. According to the information fields of RTS, corresponding beams of node $D$ send new control frames, called the *wish to send* (WTS) message through all relay nodes to $S_i$ and $S_j$.

F4. Assume that a candidate relay node $R_c$ acquires $Rate_{SD}$ by acquiring this information from the WTS message and calculates the $Rate_{SR_c}$, $Rate_{Ra,D}$, and $Rate_{SD}$ according to the measured SNR of the ATS and WTS messages. If $R_c$ satisfies Eq. (1), it will become an appointed relay node $R_a$. When node $R_a$ is identified, $R_a$ sends a new control frames, called the *promise to send* (PTS) message, to $S_i$ (or $S_j$). Nodes $S_i$ and $S_j$ then send multiple sub-frames after receiving the PTS message from all appointed relay nodes.

F5. The PTS message keeps the latest $Rate_{SR_a}$ and $Rate_{Ra,D}$. The source node $S$ recomputes the number of relay nodes, since some appointed relay nodes do not actually respond the PTS message or the information of the PTS message is not the same as the record in the relay node table. The relay node may receive RIF (ATS) message of multiple sources at the same time; however, the relay node responds the PTS message to only one source node $S$. After receiving RIF (ATS) and CIF (WTS), the relay node can send to one source node according to its information fields. Observe that $S$ chooses the lowest rate to send the PTS message. The relay node with the same data rate randomly chooses one to send the PTS message.

F6. Finally, all source nodes $S_i$ and $D$ send the RESET or SCH messages to reset the redundant NAV duration for all neighboring nodes because that NAV durations setting by RIF, CIF, SCH, ATS, and WTS messages are set to the maximal transmission time. This may reduce the NAV length.

Example is given in Fig. 10, $S_1$ divides a data frame into three sub-frames. One sub-frame is directly transmitted to $D$ and two sub-frames are transmitted to $D$ through relay nodes $R_1$ and $R_2$. Node $S_2$ additionally divides a data frame into three sub-frames, which are sent to $D$ through $R_3$, $R_4$, and $R_5$. Node $D$ finally receives the six sub-frames and combines them as the original data frame, and then responds ACK messages to $S_1$ and $S_2$. To implement the combination, the *sub-frame number* field in IEEE 802.11 MAC header frame is used to solve this combination.

C. Eliminating Unnecessary Relay Nodes

Our multiple relay-based MAC protocol supports a dynamic relay-node management by dynamically eliminating unnecessary relay nodes under the fragment burst transmission. To take the unnecessary relay
Fig. 11. Example of fragment burst with adjusting relay nodes: (a) Phase I: before adjusting, (b) Phase II: after adjusting, and (c) the detailed procedure.
Algorithm 2: Adjusting Unnecessary Relay Nodes for Sources

Input:

\[ S = \{ S_i \mid 1 \leq i \leq n \} \], where \( S_i \) is a desired source.

\[ C = \{ C_i \mid 1 \leq i \leq n \} \], where

\[
C_i = \begin{cases} 
S_i & \text{if } D \text{ receives from } S_i \text{ directly} \\
R_x & \text{if } D \text{ receives from } R_x \text{ indirectly} \\
\emptyset & \text{if } S_i \text{ has no extra relay nodes}
\end{cases}
\]

\[ C' = \{ C'_i \mid 1 \leq i \leq n \} \], where

\[
C'_i = \begin{cases} 
\emptyset & \text{if } S_i \text{ has no extra relay nodes} \\
\{ R_j \mid 1 \leq j \leq \gamma \} & \text{if } S_i \text{ has extra relay nodes}
\end{cases}
\]

\[ T = \{ T_i \mid 1 \leq i \leq n \} \], where \( T_i \) is the reception latency to \( S_i \).

1 Procedure AdjustRelayNodes()
2 \( T_{\text{max}} = \max\{T_1, T_2, ..., T_n\} \);
3 for \( i = 1 \) to \( |S| \)
4 \( \text{if } T_i \geq T_{\text{max}} \text{ then continue } ; \)
5 \( \text{adjust} = \text{True} ; \)
6 \( c = C_i ; c' = C'_i ; \)
7 \( \text{while } \text{adjust} == \text{True} \&\& c' \neq \emptyset \text{ do } \)
8 \( \text{adjust} = \text{Remove}(T_{\text{max}}, S_i, D, c, c', \text{True}) ; \)
9 \( \text{if } \text{adjust} == \text{True} \text{ then } C'_i = c' ; \)
10 end while
11 \( \text{if } c \neq S_i \text{ then } \)
12 \( \text{adjust} = \text{Remove}(T_{\text{max}}, S_i, D, c, c', \text{False}) ; \)
13 \( \text{if } \text{adjust} == \text{True} \text{ then } C_i = S_i ; \)
14 end if
15 next \( i \)
16 return \( (C, C') \) ;
17 end Procedure

Procedure : Remove()

1 Procedure Remove\( (T_{\text{max}}, S, D, c, c', \text{extra}) \)
2 \( \text{if } \text{extra} == \text{True} \&\& c' \neq \emptyset \text{ then } \)
3 \( \text{Remove one element of } c' ; \)
4 else if \( \text{extra} == \text{False} \text{ then } \)
5 \( c = S ; \)
6 end if
7 \( t = \text{CountDestination}(S, D, c \cup c') ; \)
8 \( \text{if } t \leq T_{\text{max}} \text{ then } \)
9 \( \text{return } \text{True} ; \)
10 \( \text{return } \text{False} ; \)
11 end Procedure

Procedure : CountDestination()

1 Procedure CountDestination\( (S, D, \Gamma) \)
2 \( k = \mid \Gamma \mid ; \)
3 \( t = \max( T_{\text{overhead}} + \frac{\text{Payload}_D}{\text{Rate}_1}, T_{\text{overhead}} + \frac{\text{Payload}_D}{\text{Rate}_2}, ... , T_{\text{overhead}} + \frac{\text{Payload}_D}{\text{Rate}_g} ) ; \)
4 \( \text{if } \Gamma_p \neq D \)
5 \( T_{\text{overhead}} = 2T_{\text{physical}} + \text{SIFS} , \text{Rate}_p = \frac{\text{Rate}_{\text{ST}} \times \text{Rate}_{\text{TD}}}{\text{Rate}_{\text{SD}}} \)
6 \( T_{\text{overhead}} = T_{\text{physical}} , \text{Rate}_p = \text{Rate}_{\text{SD}} \)
7 \( t = t + \text{ACK} + \text{SIFS} ; \)
8 \( \text{return } t ; \)
9 end Procedure
node problem into account, this work aims to provide a dynamic relay-node management to release some unnecessary relay nodes to improve the system throughput.

A destination node $D$ knows the whole relay-node topology and transmission latency of all the source node $S_x$ after receiving data frames from all $S_x$, where $1 \leq x \leq n$. Algorithm 2 is performed by the destination node $D$ if $D$ has already received the first data frame (phase I) from all $S_x$ during the fragment burst period. A destination node $D$ uses the ACK message to notify all $S_x$ the adjusted information.

Before describing procedure AdjustRelayNodes, two procedures are defined. Procedure Remove ($T_{\text{max}}, S, D, c, c', \text{extra}$) is used to remove a relay node $c'$ from $c$ if extra is equal to True. Procedure CountDestination ($S, D, \Gamma$) is used in Procedure Remove to calculate the total transmission delay for a giving pair of the source node $S$ and the destination node $D$. The main operation of AdjustRelayNodes procedure is explained as follows.

1. $T_{\text{max}}$ represents the maximum transmission delay of all source nodes $S_x$, where $1 \leq x \leq n$, $T_{\text{max}} = \max\{T_1, T_2, \ldots, T_n\}$, $T_x$ is the transmission delay between $S_x$ to $D$. (line 2 of Procedure AdjustRelayNodes)

2. Any of the source nodes $S_x$, $1 \leq x \leq n$, whose transmission delay is smaller than $T_{\text{max}}$, has opportunity to remove unnecessary relay nodes.

3. For each $S_x$, $1 \leq x \leq n$, execute G3 and G4.

4. call Procedure Remove ($T_{\text{max}}, S_x, D, c, c', \text{extra}$) if $c' \neq \emptyset$.

5. call Procedure Remove ($T_{\text{max}}, S_x, D, c, c', \text{extra}$) if $c' \neq S_x$.

6. goto G3.

Example is given in Fig. 11 to illustrate the dynamic relay-node management under the fragment bursting, while the initial topology is given in Fig. 10. To illustrate the effects of fragment bursting, two consecutive data fragmentation transmissions are illustrated in Fig. 11(a) and (b) (phases I and II). Observe that Fig. 11(a) (phase I) is the same as Fig. 10, and the result of eliminating unnecessary relay nodes is given in Fig. 11(b) (in phase II). The key difference is that three relay nodes $R_3, R_4, \text{and } R_5$ exist between $S_2$ and $D$ in phase I, but only relay node $R_4$ exists between $S_2$ and $D$ in phase II. This is because it is useless to keep relay nodes $R_3$ and $R_5$ in phase II. Observe that the total transmission delay is bounded by $S_1$ because $S_1$ has the maximal transmission delay. It is meaningless to utilize more relay nodes for $S_2$ since the total transmission delay for $S_2$ to $D$ cannot be reduced even if more number of relay nodes is utilized. Therefore, all of the unnecessary relay nodes are released. Before phase I, nodes $S_1$ and $S_2$ cannot know each other’s transmission delays. Therefore, the operation of eliminating unnecessary relay nodes is occurs after phase I. The ACK message in phase I from $S_1$ and $S_2$ is used to provide the information of transmission delay.

V. PERFORMANCE EVALUATION

This section presents the analysis of transmission delays and simulation results. The analysis of the transmission delays of the cases of no relay, direct relay, direct relay with one extra relay, and direct relay with two extra relays are first investigated. Performance comparisons between ESIF and the multiple relay-based MAC protocol are then illustrated.
A. Analysis of Transmission Delays

The analysis of the transmission delays of the cases of no relay, direct relay, direct relay with one extra relay, and direct relay with two extra relays are investigated as follows.

1) NAV Duration: Tables 1, 2, and 3 show NAV durations of various control frames. Table 1 represents that the source and destination nodes do not meet the fragment burst. The first row (RIF, ATS, SCH) contains the control frames sent by the source nodes. The second row (CIF, WTS, SCH) contains the control frames sent by the destination nodes. The NAV durations of the above control frames are assumed to be the longest transmission latency because of the ACK synchronization problem. In the third row, ACK and SCH are sent by the destination nodes and RESET is sent by the source nodes. Tables 2 and 3 are the NAV durations of the destination and source nodes during the fragment burst period, respectively. Successively multiple data frames and ACKs may play the role of RIF and CIF, respectively, to set the suitable NAV when the fragment burst occurs. The destination knows the transmission conditions of each source after receiving all data frames. Hence, the next fragment burst duration is the maximum transmission latency of all the source nodes. The destination responds ACKs with the maximum NAV duration to the sources. Other beams of the destination also send SCH with the maximum NAV duration. The sources understand the duration of the next fragment burst after receiving ACKs. The sources send the next fragment an SIFS later. Other beams of the sources also send SCH to set NAV.

Table 1: NAV Duration of Control Frames

<table>
<thead>
<tr>
<th>Frame</th>
<th>NAV duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>RIF, ATS, SCH</td>
<td>$CIF + T_{\text{physical}} + \frac{MAC + \text{Payload}_{\text{max}}}{\text{the lowest data rate}} + ACK + 3SIFS$</td>
</tr>
<tr>
<td>CIF, WTS, SCH</td>
<td>$T_{\text{physical}} + \frac{MAC + \text{Payload}_{\text{max}}}{\text{the lowest data rate}} + ACK + 2SIFS$</td>
</tr>
<tr>
<td>ACK, SCH, RESET</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 2: NAV Duration During Fragment Burst Period of Destination Node

<table>
<thead>
<tr>
<th>Frame</th>
<th>NAV duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACK_1, ACK_n-1</td>
<td>$\max{DATA_{S_{1}}^{\max}, \ldots, DATA_{S_{i}}^{\max}} + ACK + 2SIFS$</td>
</tr>
<tr>
<td>SCH_1, SCH_n-1</td>
<td>$\max{DATA_{S_{1}}^{\max}, \ldots, DATA_{S_{i}}^{\max}} + ACK + 2SIFS$</td>
</tr>
<tr>
<td>ACK_n</td>
<td>0</td>
</tr>
</tbody>
</table>

$(n = \text{number of fragment})$

Table 3: NAV Duration during Fragment Burst Period of Source Node

<table>
<thead>
<tr>
<th>Frame</th>
<th>NAV duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>DATA_i at beam_j</td>
<td>$\text{NAV Duration of } ACK_{i-1} - SIFS - (\text{sub)frame transmission latency at beam}_j$</td>
</tr>
<tr>
<td>SCH_{2n}</td>
<td>$\text{NAV Duration of } ACK_{i-1} - SIFS - SCH$</td>
</tr>
</tbody>
</table>

$(n = \text{number of fragment}, M = \text{number of beams})$

(2in, 1jM))

2) Network parameters: Network parameters of the mathematic analysis are given in Table 4. A transmission delay includes the Inter-frame Space (IFS), several control frames and a data frame. A
data frame consists of a physical preamble, a physical header, a MAC header and a payload. All the control frames are transmitted at the lowest data rate to avoid hidden terminal problem. The MAC header and payloads are transmitted at the data rate which is recorded in the physical header.

\[ T_{\text{physical}} = T_{\text{preamble}} + T_{\text{PLCP header}} \] (8)

where \( T_{\text{preamble}} \) and \( T_{\text{PLCP header}} \) are transmitted at the lowest data rate. \( T_{\text{PLCP header}} \) has the information about the data rate of the MAC header and payload.

### Table 4: Network Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( T_{\text{physical}} )</td>
<td>Transmission time of physical layer.</td>
</tr>
<tr>
<td>( L )</td>
<td>The size of payload. (bytes)</td>
</tr>
<tr>
<td>( \text{Rate}_{\alpha \beta} )</td>
<td>The data rate between nodes ( \alpha ) and ( \beta ). (Mbps)</td>
</tr>
<tr>
<td>( T_{n}^{\text{DATA}} )</td>
<td>Transmission delay of a data frame using ( n ) extra relay nodes.</td>
</tr>
<tr>
<td>( T_{\hat{n}}^{\text{DATA}} )</td>
<td>Transmission delay of a data frame using ( n - 1 ) extra relay nodes and 1 direct relay node.</td>
</tr>
<tr>
<td>( T_{\hat{n}} )</td>
<td>Transmission time using ( n - 1 ) extra relay nodes.</td>
</tr>
<tr>
<td>( T_{\hat{\hat{n}}} )</td>
<td>Transmission time using ( n - 1 ) extra relay nodes and 1 direct relay node.</td>
</tr>
</tbody>
</table>

3) Transmission Delays: Suppose that the payload transmission delay is long-winded. The transmission delays of the four relay conditions (no relay, direct relay, direct relay + one extra relay, and direct relay + two extra relays) are listed as follows.

As shown in Fig. 4 (a), node \( S \) has an impending frame which transmits to node \( D \) directly. Suppose that the data rate is the lowest data rate and the payload length \( L \) is the maximum length.

\[ T_{0}^{\text{DATA}} = T_{\text{physical}} + \frac{MAC + 8L}{\text{Rate}_{SD}} \] (9)

\[ T_{0} = DIFS + 3SIFS + RIF + CIF + T_{0}^{\text{DATA}} + \text{ACK} \] (10)

Given a relay node \( R \) between the direct transmission beam of node \( S \) and the direct reception beam of \( D \). Equation (11) shows the actual data rate if node \( S \) transmits a frame to node \( D \) via node \( R \).

\[ \frac{1}{\text{Rate}_{SR}} + \frac{1}{\text{Rate}_{RD}} = \frac{\text{Rate}_{SR} \times \text{Rate}_{RD}}{\text{Rate}_{SR} + \text{Rate}_{RD}} \] (11)

Using Equation (11), the transmission delay of the direct relay is

\[ T_{1}^{\text{DATA}} = T_{\text{physical}} + \frac{8L}{\text{Rate}_{SR}} + SIFS + T_{\text{physical}} + \frac{8L}{\text{Rate}_{RD}} \]

\[ = 2T_{\text{physical}} + SIFS + 8L \times \frac{\text{Rate}_{SR} \times \text{Rate}_{RD}}{\text{Rate}_{SR} + \text{Rate}_{RD}} \] (12)
Equation (13) is the transmission delay of the direct relay. The transmission delay of the direct relay includes the time of control frames and $T_{DATA}^{DATA}$.

\[
T_1 = DIFS + 5SIFS + RIF + CIF + PTS + T_{DATA}^{DATA} + ACK + RESET
\]

(13)

Given a relay node $R_1$ which can establish an extra data channel to distribute the transmission latency, and a relay node $R_2$ is located between the direct beams of nodes $S$ and $D$. The payload is divided into two equal-sized sub-frames. One sub-frame is relayed to node $D$ via node $R_2$, The other sub-frame is relayed to node $D$ via node $R_1$. Hence, in Equation (19), two sub-frames’ transmission time is shown and max means the maximum of the two transmission delays, which is actually the total transmission delay. Equation (20) represents the total transmission latency of direct relay + one extra relay. By using Equation (11), actual data rates of the direct relay (node $R_2$) and the extra relay (node $R_1$) are shown in Equations (14) and (15), respectively.

\[
\frac{1}{Rate_{SR_2}} + \frac{1}{Rate_{R_2D}} = \frac{Rate_{SR_2} \times Rate_{R_2D}}{Rate_{SR_2} + Rate_{R_2D}} = x
\]

(14)

\[
\frac{1}{Rate_{SR_1}} + \frac{1}{Rate_{R_1D}} = \frac{Rate_{SR_1} \times Rate_{R_1D}}{Rate_{SR_1} + Rate_{R_1D}} = y
\]

(15)

The transmission time equals to the frame size divided by the data rate. A data frame is transmitted simultaneously via a direct relay data channel and an extra relay data channel. Hence, the payload may be divided into proper proportions according to the actual data rate of each data channel. Data channels of higher transmission rate transmit longer sub-frames. Data channels of lower transmission rate transmit shorter sub-frames. The following equations shows how to divide a frame according to different data rates of the two data channels.

\[
\frac{L_1}{x} = \frac{L_2}{y}
\]

(16)

\[
L_1 = \frac{x}{x + y}
\]

(17)

\[
L_2 = \frac{y}{x + y}
\]

(18)

where $L_1$ is the sub-frame of the direct relay and $L_2$ is the sub-frame of the extra relay.

\[
T_2^{DATA} = \max \{ 2T_{physical} + SIFS + 8L \times L_1 \times \frac{Rate_{SR_1} \times Rate_{R_1D}}{Rate_{SR_1} + Rate_{R_1D}}, 2T_{physical} + SIFS + (8L - 8L \times L_1) \times \frac{Rate_{SR_2} \times Rate_{R_2D}}{Rate_{SR_2} + Rate_{R_2D}} \}
\]

(19)

\[
T_2 = DIFS + 5SIFS + RIF + CIF + PTS + T_2^{DATA} + ACK + RESET
\]

(20)
Fig. 4 (d) demonstrates that nodes $R_1$, $R_2$, and $R_3$ help node $S$ to relay three sub-frames, respectively. The payload of a frame is divided into three sub-payloads. The total transmission latency $T_3$ is shown as follows. By using Equation (11), the actual data rate of the extra relay (node $R_3$) is shown in Equation (21).

$$\frac{1}{\text{Rate}_{SR_2}} + \frac{1}{\text{Rate}_{R_3,D}} = \frac{\text{Rate}_{SR_3} \times \text{Rate}_{R_3,D}}{\text{Rate}_{SR_3} + \text{Rate}_{R_3,D}} = z$$ \hspace{1cm} (21)

The following equations show how to divide a frame according to different data rates of the three data channels.

$$\frac{L_1}{x} = \frac{L_2}{y} = \frac{L_3}{z}$$ \hspace{1cm} (22)

$$L_1 = \frac{x}{x+y+z}$$ \hspace{1cm} (23)

$$L_2 = \frac{y}{x+y+z}$$ \hspace{1cm} (24)

$$L_3 = \frac{z}{x+y+z}$$ \hspace{1cm} (25)

where $L_1$ is the sub-frame of the direct relay. $L_2$ and $L_3$ are sub-frames of the two extra relays respectively.

$$T_3^{DATA} = \max\{2T_{physical} + SIFS + 8L \times L_1 \times \frac{\text{Rate}_{SR_3} \times \text{Rate}_{R_3,D}}{\text{Rate}_{SR_3} + \text{Rate}_{R_3,D}}, 2T_{physical} + SIFS + 8L \times L_2 \times \frac{\text{Rate}_{SR_3} \times \text{Rate}_{R_3,D}}{\text{Rate}_{SR_3} + \text{Rate}_{R_3,D}}, 2T_{physical} + SIFS + (8L - 8L \times L_1 - 8L \times L_2) \times \frac{\text{Rate}_{SR_3} \times \text{Rate}_{R_3,D}}{\text{Rate}_{SR_3} + \text{Rate}_{R_3,D}}\}$$ \hspace{1cm} (26)

$$T_3 = DIFS + 5SIFS + RIF + CIF + PTS + T_3^{DATA} + ACK + RESET$$ \hspace{1cm} (27)

Equations (10), (13), (20), and (27) are compared as follows. $T_1$ is smaller than $T_0$ since the long-winded transmission time of a data frame is improved by node $R$ although control overheads, such as $PTS$, and $T_{physical}$ are increased. $T_2$ is shorter than $T_1$ because $T_2$ uses an extra relay node. Similarly, $T_3^{DATA}$ is close to one-third of $T_0^{DATA}$, and hence $T_3$ is smaller than $T_0$, $T_1$, and $T_2$.

B. Simulation results

The simulation is based on NCTUns 3.0 [1] simulator and our extensions. We modify the original MAC80211 and Wphy modules of NCTUns 3.0 simulator to support concurrently multiple communications and multiple beams with multirate capability. Simulation parameters are shown in Table 5 and the performance metrics are shown as follows.

- Throughput: total number of received data bits for all the destination hosts.
- Average transmission time: includes the transmission time of a data frame, total transmission time of the control frames (RIF, CIF, ACK), and total time of Inter-Frame Space (DIFS, SIFS).
- Average transmission latency: the time from a packet entering the queue of the source to the time it is delivered to the destination. Hence, the average transmission latency consists of the propagation delay, transmission time, and queuing delay.

### Table 5: Simulation Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Mbps Range</td>
<td>250 m</td>
</tr>
<tr>
<td>2 Mbps Range</td>
<td>186.75 m</td>
</tr>
<tr>
<td>5.5 Mbps Range</td>
<td>167.75 m</td>
</tr>
<tr>
<td>11 Mbps Range</td>
<td>120.5 m</td>
</tr>
<tr>
<td>Carrier Sensing Range</td>
<td>550 m</td>
</tr>
<tr>
<td>DIFS</td>
<td>50 $\mu$s</td>
</tr>
<tr>
<td>SIFS</td>
<td>10 $\mu$s</td>
</tr>
<tr>
<td>Total Beams</td>
<td>8</td>
</tr>
<tr>
<td>Data Size</td>
<td>1400 bytes</td>
</tr>
<tr>
<td>Packet Buffer Size</td>
<td>50</td>
</tr>
<tr>
<td>Packet Queuing Lifetime</td>
<td>50 packet duration</td>
</tr>
</tbody>
</table>

1) Differences between Multiple Relay-Based MAC Protocol and ESIF: ESIF supposes that each node is equipped with multiple beam antennas and the data rate is 2 Mbps. However, this paper assumes that each node uses multiple beam antennas with multirate capability (1, 2, 5.5, and 11 Mbps). Fig. 12 shows the difference between the multiple relay-based MAC protocol and ESIF. The scenario only has a source node and a destination node. Distances between the source and destination nodes are changed from 5 to 250m. In Fig. 12 (a), when the distances between the source and destination nodes are 250 m $\sim$ 186.65 m, 186.75 m $\sim$ 167.76 m, 167.75 m $\sim$ 102.5 m, and 120.5 m $\sim$ 5 m, the throughput of our scheme are 0.84, 1.46, 2.88, and 3.99 Mbps, respectively. In Fig. 12 (b), when the distances between the source and destination nodes are 250 m $\sim$ 186.65 m, 186.75 m $\sim$ 167.76 m, 167.75 m $\sim$ 102.5 m, and 120.5 m $\sim$ 5 m, the packet transmission latencies of our scheme are 12.88, 6.96, 3.19, and 2.12 ms, respectively. When the distances between the source and destination nodes is 186.75 m $\sim$ 5 m, the throughput and average packet transmission time of ESIF are 1.59 Mbps and 6.67 ms, respectively. But the throughput of ESIF is zero when the distance is greater than 186.75 m because the distance exceeds the transmission range of 2 Mbps.

The performance of our scheme is slightly worse than ESIF when the data rate between the source and destination nodes is 2 Mbps. This is because the control frames and data frames of ESIF are transmitted at 2 Mbps. However, in order to avoid hidden terminal problem, the control frames of our scheme are transmitted at the lowest data rate (1 Mbps). For performing multiple relay scheme, this paper adds relay MAC address and information fields to the original frame formats of ESIF, respectively. Furthermore, in
order to reset the redundant NAV duration of neighboring nodes, the source may send a RESET after receiving an ACK. These reasons incur slightly worse performance.

2) Various Relay Conditions: Figs. 13 and 14 show the throughput, average transmission latency, and packet dropped rate by simulating four relay conditions. The four relay conditions are "no relay", "direct relay", "direct relay + one extra relay", and "direct relay + two extra relay". In Fig. 13, the distance between the source and destination nodes is 250m and the data rate is 1 Mbps. In Fig. 14, the distance between the source and destination nodes is 186m and the data rate is 2 Mbps. Three relay nodes are placed in the relay region between the source and destination nodes. And the three relay nodes are located at different beams of the source and destination nodes to establish three data channels for relay. These conditions are run for 30 seconds to obtain the results.

Fig. 13 (a) shows the throughput of the four relay conditions of 1 Mbps under various packet arrival rates. Figs. 13 (b) and (c) illustrate the average transmission latency of the four relay conditions of 1 Mbps under various packet arrival rates. Fig. 13 (d) displays the packet dropped rate of the four relay conditions of 1 Mbps under various packet arrival rates. The maximum throughput of "no relay" still does not exceed 0.84 Mbps even if the packet arrival rate is greater than 140 packets/second. This is because the transmission latency of "no relay" is 12.89 ms. The transmission time is greater than the packet arrival interval. The packet buffer accumulates more and more packets when the packet arrival rate becomes large. Finally, a packet is dropped if a packet stays in the buffer too long. From Fig. 13 (c), we can observe that the average packet transmission time of "direct relay", "direct relay + one extra relay", and "direct relay + two extra relay" are 6.04, 3.9, and 3.2 ms, respectively. Hence, using more relay nodes to transmit can enhance the throughput evidently since the transmission time is shortened.

Fig. 14 (a) demonstrates the throughput of the four relay conditions of 2 Mbps under various packet arrival rates. Fig. 14 (a) shows that the maximum throughput of "no relay", "direct relay", "direct relay + one extra relay", and "direct relay + two extra relay" are 1.46, 2.5, 3.27, and 3.65 Mbps, respectively. Figs. 14 (b) and (c) show the average packet transmission latency of the four relay conditions of 2 Mbps under various packet arrival rates. From Fig. 14 (c), the average packet transmission latencies of "no
Fig. 13. Four relay conditions of 1 Mbps: (a) throughput, (b) average transmission latency, (c) average transmission latency before buffer overflow, and (d) packet dropped rate.

relay”, ”direct relay”, ”direct relay + one extra relay”, and ”direct relay + two extra relay” are 6.96, 3.86, 2.8, and 2.43 ms, respectively. Fig. 14 (d) illustrates the packet dropped rate of the four relay conditions of 2 Mbps under various packet arrival rates. The shorter transmission time makes smaller packet dropped rate.

3) Concurrently Multiple Communications: Our MAC protocol is based on ESIF. The feature of ESIF is concurrent multiple communications. Figs. 15 and 16 illustrate the performances of the many-to-one communication and cross concurrent communications, respectively. The throughput and transmission time are derived every 0.5 second so as to calculate the averages. The two scenarios are run for 15 seconds. Node $R_1$ is awakened in the fifth second and Node $R_2$ is awakened in the tenth second. In Fig. 15 (a), Both nodes $S_1$ and $S_2$ communicate with node $D$. In ESIF, the data rate is 2 Mbps and all control and data frames are transmitted at 2 Mbps. In our scheme, Node $S_1$ transmits at 2 Mbps and node $S_2$ transmits at 5.5 Mbps. The transmission time is constrained by node $S_1$ when nodes $S_1$ and $S_2$ communicate with node $D$ concurrently. Hence, our scheme is rarely worse than ESIF after the fifth second. In the fifth seconds, node $R_1$ awakens and helps node $S_1$. The transmission time between nodes $S_1$ and $D$ is shortened. In the tenth seconds, node $R_2$ awakens and helps node $S_1$. The transmission time between nodes $S_1$ and $D$ is further shortened. Hence, the throughput and the transmission time are improved by nodes $R_1$ and $R_2$.
Fig. 14. Four relay conditions of 2 Mbps: (a) throughput, (b) average transmission latency, (c) average transmission latency before buffer overflow, and (d) packet dropped rate.

as shown in Figs. 15 (b) and (c).

In Fig. 16 (a), nodes $S_1$ and $S_2$ send frames to nodes $D_1$ and $D_2$ via node $C$, respectively. Node $C$ can relay the frames of nodes $S_1$ and $S_2$ to nodes $D_1$ and $D_2$ simultaneously when nodes $S_1$ and $S_2$ transmit frames via node $C$ simultaneously. The transmission time is constrained by node $S_1$ when nodes $S_1$ and $S_2$ transmit simultaneously. The transmission time is constrained by node $D_2$ when node $C$ transmits to nodes $D_1$ and $D_2$ simultaneously. Figs. 16 (b) and (c) show the throughput and the transmission time are improved after node $R_1$ and $R_2$ are awakened. The transmission time between nodes $S_1$ and $C$ is shortened by $R_1$. The transmission time between nodes $c$ and $D_2$ is shortened by $R_2$.

4) Random Topology: Fig. 17 shows the one-to-many communication under a fully connected topology. Ten nodes are put randomly in a $180m \times 180m$ area. One of the ten nodes acts as the sender and two of the remaining nine nodes act as receivers. The sender communicates with the two receivers concurrently. With the raised packet arrive rate, the throughput of ESIF only stays at 3 Mbps since the data rate is always 2 Mbps. However, the throughput of our scheme can continue to be enhanced when the packet arrival rate becomes more and more higher because the transmission latencies of the two links are shortened by relay nodes. In Fig. 18, the multihop topology is adopted. Thirty nodes are put randomly in a $1000m \times 500m$ area. The shortest path routing is adopted. We simulate one traffic flow in the area. Compared to
ESIF, our scheme can still achieve higher throughput and lower transmission latency.

5) Fragment Burst with Adjusting Relay Nodes: Fig. 19 shows the wireless medium utilization of the adjusting relay nodes algorithm. The throughput and transmission time are derived from the average of the results of every 0.25 second. In Fig. 19 (a), $S_1-D$ and $S_2-D$ have a heavy traffic and $S_3-R_3$ has a light traffic. In order to alleviate the heavy traffic and transmit faster, nodes $S_1$ and $S_2$ often perform the fragment burst and multiple relay mechanism. Nodes $R_1$ and $R_2$ are relay nodes of node $S_1$; nodes $R_3$, $R_4$, and $R_5$ are relay nodes of node $S_2$. Besides, node $R_3$ is also the desired destination of node $S_3$. For this reason, node $S_3$ can not communicate with node $R_3$ when node $R_3$ acts as the relay node of $S_2$. In this scenario, node $R_3$ is often requested to be the relay node of node $S_2$ because node $S_2$ has a heavy traffic to node $D$. Nodes $R_3$ may be the relay node of $S_2$ for a long time if $S_2$ performs the fragment burst and multiple relay mechanism. However, nodes $R_3$ and $R_5$ are unnecessary relay nodes when nodes $S_1$ and $S_2$ exploit multiple relay mechanism to transmit simultaneously.

Nodes $S_1$ and $S_2$ do not know each other’s transmission delay. Accordingly, nodes $R_3$ and $R_5$ are hard to shorten the total transmission delay during the concurrent fragment burst period of nodes $S_1$ and $S_2$. The wireless medium utilization is low if node $R_3$ acts as a unnecessary relay node for a long time. This is because that node $S_3$ has no opportunity to communicate with node $R_3$. However, node $D$ can response
Fig. 16. Cross concurrent communications with multiple relay: (a) topology, (b) throughput, and (c) average transmission time.

Fig. 17. Performances comparison of one-to-many communication under one hop: (a) throughput and (b) average transmission latency.
Fig. 18. Performances comparison of multihop: (a) throughput and (b) average transmission latency.

Fig. 19. The wireless medium utilization of adjusting relay nodes algorithm: (a) throughput of $S_1-D + S_2-D$ and $S_3-R_3$ and (b) total throughput of $S_1-D + S_2-D$ and $S_3-R_3$. 
ACKs with relay information to the source nodes after receiving all (sub)frames. Node $R_3$ can be released the role of the unnecessary relay node by using the removing unnecessary relay nodes algorithm of node $D$ when nodes $S_1$ and $S_2$ exploit multiple relay mechanism to transmit simultaneously. Nodes $S_1$ and $S_2$ perform the fragment burst (five DATA/ACK) every three transmissions. The packet arrival rate of $S_3$-$R_3$ is 50 packets/sec. Fig. 19 (b) shows the performance with and without the removing algorithm. Without the removing algorithm, node $R_3$ is often to be a unnecessary relay nodes when nodes $S_1$ and $S_2$ perform the multiple relay mechanism and fragment burst simultaneously. Node $S_3$ is difficult to communicate with node $R_3$, and hence the throughput of $S_3$-$R_3$ is very low. With the removing algorithm, node $R_3$ can be released the role of the unnecessary relay node. Node $S_3$ has the opportunity to communicate with node $R_3$ even when nodes $S_1$ and $S_2$ perform the multiple relay mechanism and fragment burst simultaneously. Therefore, the throughput of $S_3$-$R_3$ becomes higher. Fig. 19 (c) shows that the total throughput and wireless medium utilization are better by using the removing unnecessary relay nodes algorithm.

VI. CONCLUSION

In this paper, we have proposed a multiple relay-based MAC protocol based on ESIF. Each node in the wireless ad hoc network is supposed to equipped with multiple beam antennas with multirate capability. We attempt to exploit the multiple relay concept to shorten the transmission latency and alleviate the anomaly problem. However, the ACK synchronization problem are raised since multirate causes the difference of the transmission latency of each beam. Our scheme can solve the ACK synchronization problem. The proposed algorithm let the source nodes allocate proper relay nodes to low rate beams. The source nodes cannot know each other’s transmission latencies during the many-to-one communication. Therefore, the unnecessary relay problem occurs. Unnecessary relay nodes continue to be used during the fragment burst period. By exploiting the adjusting algorithm, unnecessary relay nodes can be removed during the fragment burst period. Hence these unnecessary relay nodes can communicate with other nodes during the remaining fragment burst period. Finally, simulation results show that the proposed multiple relay-based MAC protocol can achieve higher throughput and lower transmission latency.

REFERENCES


