

1

A multiple relay-based medium access control protocol in multirate wireless *ad hoc* networks with multiple beam antennas

3

Yuh-Shyan Chen^{1,*}, Chih-Shun Hsu² and Po-Ta Chen³

5

¹Department of Computer Science and Information Engineering, National Taipei University, Taiwan²Department of Information Management, Shih Hsin University, Taiwan

7

³Department of Computer Science and Information Engineering, National Chung Cheng University, Taiwan

SUMMARY

9 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. 11 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 13 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 15 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 17 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 19 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 21 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 23 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 25 fragment burst period. Finally, simulation results illustrate that our multiple relay-based MAC protocol can achieve high throughput and low transmission latency. Copyright © 2009 John Wiley & Sons, Ltd.

Received 18 November 2008; Revised 17 October 2009; Accepted 20 October 2009

27 KEY WORDS: multiple beam antennas; multirate; relay; medium access control; wireless *ad hoc* networks

*Correspondence to: Yuh-Shyan Chen, Department of Computer Science and Information Engineering, National Taipei University, Taiwan.

†E-mail: yschen@csie.ntpu.edu.tw, yschen@mail.ntpu.edu.tw

Contract/grant sponsor: National Science Council of the Republic of China; contract/grant numbers: #NSC-95-2221-E-305-008, #NSC-96-2213-E-194-007

1

1. INTRODUCTION

2 A wireless ad hoc network (MANET) [1] is made up of identical mobile nodes, each node with
3 a limited wireless transmission range to communicate with neighboring nodes. The MANET can
4 be applied to military, disaster, industry, and mobile learning system [2] due to its self-organizing
5 and adaptive features. Wireless communication is unreliable due to its low channel quality, high
6 noise interference, high packet collision, and high packet loss rate. The novel design of Medium
7 Access Control (MAC) protocol can significantly improve the reliability and throughput of wireless
8 transmissions. The IEEE 802.11 standard [3] is the most well-known MAC protocol for designing
9 a MANET.

10 In the IEEE 802.11 standard, Distributed Coordination Function (DCF) and Point Coordination
11 Function (PCF) are defined in the MAC sub-layer [3]. DCF supports the contention-based
12 media access for infrastructure/*ad hoc* networks and PCF provides contention-free media access
13 for infrastructure networks. The Carrier Sense Multiple Access/Collision Avoidance (CSMA/CA)
14 mechanism is used in DCF to avoid collisions and hidden terminal problem [4]. Clear Channel
15 Assessment (CCA) function at the physical layer is used to detect the busy or idle state of the
16 wireless medium. To solve the hidden terminal problem and avoid collision, RTS (Request to
17 Send)/CTS (Clean to Send) message exchange is developed in IEEE 802.11 to set the network
18 allocation vector (NAV). Consequently, the CCA function and the RTS/CTS handshake mecha-
19 nism are the key operations of the CSMA/CA mechanism. In addition, IEEE 802.11 supports the
20 fragment burst to increase the throughput since control messages, such as RTS and CTS messages,
21 are overheads for each data transmission. Existing IEEE 802.11 networks assume that each mobile
22 node is equipped with an omnidirectional antenna [5]. Many research results [5, 6] point out that
23 the property of omnidirectional antennas leads to high collisions and low throughput in MANETs.
24 A mobile node equipped with an omnidirectional antenna always listens to signals from all direc-
25 tions and sends signals to all directions. Because of that, the beam width of an omnidirectional
26 antenna is 360° . The drawback of the omnidirectional antenna is that a mobile node in the trans-
27 mission mode may collide and/or interfere with other unconcerned mobile nodes in the reception
28 mode. The deserved throughput is insidiously degraded. The drawback of the omnidirectional
29 antenna [7, 8] can be significantly alleviated by using the directional antenna or multiple beam
30 antennas. It is sure that the existing IEEE 802.11 MAC protocol cannot be directly applied to
31 MANET with directional antenna or multiple beam antennas. To increase the spatial reuse, avoid
32 co-channel interference and reduce packet collision, many existing protocols were recently designed
33 to consider the models of directional antennas [9–11] and multiple beam antennas [12–15]. Espe-
34 cially, the use of multiple beam antennas can improve the throughput if multiple beam antennas
35 simultaneously communicate with multiple neighboring nodes [12]. This work investigates a new
36 MAC protocol under the same model of multiple beam antennas in [12].

37 The rate-adaptive MAC protocol is recently considered in [16] to dynamically adjust the power
38 level and/or transmission rate to improve the network throughput and the transmission latency.
39 IEEE 802.11 a/b/g standard [3] also defines the physical layer to support multirate capability
40 to improve the throughput of wireless networks. Unfortunately, the standard does not explicitly
41 specify algorithms to utilize multirate capability. Multirate wireless networks incur the *anomaly*
42 problem [16], because low data rate hosts may influence the original performance of high data
43 rate hosts. The channel occupying time of low data rate hosts is usually longer than that of high
44 data rate hosts. Low data rate hosts seize the available channel time of high data rate hosts. More
45 recently, relay-based MAC protocols [17–19] are designed to significantly reduce the anomaly

1 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.

3 In this paper, we propose a multiple relay-based MAC protocol to improve the throughput for
low data rate hosts. In this investigation, each node is equipped with multiple beam antennas with
5 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
7 time. 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
9 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
11 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
13 burst period. Finally, the simulation results illustrate that the proposed multiple relay-based MAC
protocol achieves the result of high throughput and low transmission latency.

15 The remainder of this paper is organized as follows. Section 2 discusses related works. The
system model, basic idea, and challenges are described in Section 3. Section 4 presents the details
17 of the multiple relay-based MAC protocol. The Performance evaluation is examined in Section 5.
Finally, Section 6 concludes this paper.

19 2. RELATED WORK

21 This section introduces some related works [12, 17–22] about multiple beam antennas, multirate,
and relay-based MAC protocols as follows.

2.1. Protocols of multiple beam antennas

23 Figure 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.
25 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
27 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 simul-
29 taneously communicate with multiple nodes by exploiting Space Division Multiple Access (SDMA)
concept.

31 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
33 antenna model. Jain *et al.* [12] proposed a cross-layer MAC protocol, called Explicit Synchroniza-
tion via Intelligent Feedback (ESIF), to utilize the advantages of multiple beam antennas. Figure 2
35 illustrates the basic ESIF operation. Figures 2(a) and (b) give the transmission details of many-
to-one communication and one-to-many communication, respectively. Observe that ESIF removes
37 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,
39 each node must hear special RTS/CTS, RTS with Intelligent Feedback (RIF), CTS with Intelli-
gent Feedback (CIF), and schedule (SCH) which are specially designed by ESIF. This result can
41 conjecture transmission schedule of all neighbor nodes. In the cross-layer design of ESIF, the

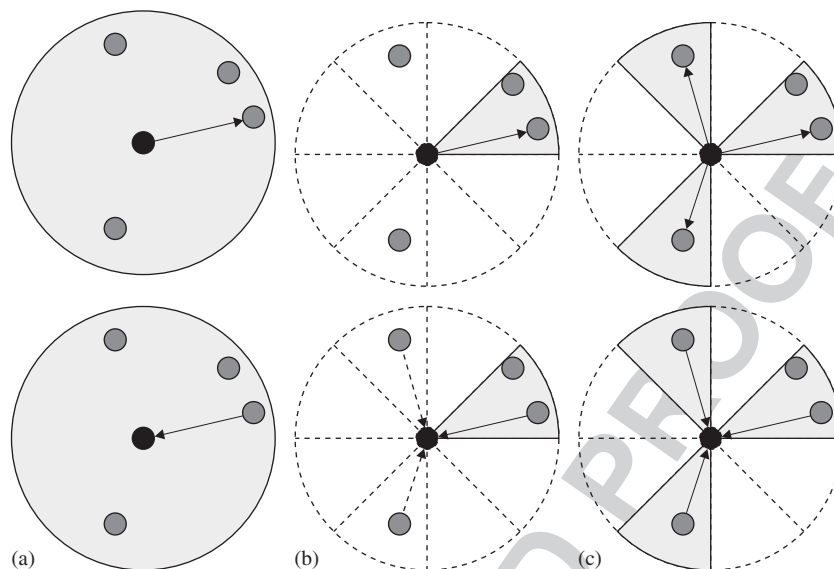


Figure 1. Antenna types: (a) the omni-directional antenna; (b) the directional antenna; and (c) multiple beam antennas.

1 MAC layer of each mobile node additionally acquires the routing information from the network
 2 layer to avoid collisions and obtain better judgements. The model of ESIF is the multiple beam
 3 antennas and single data rate capability.

2.2. Multirate protocols

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

17 Kamerman and Monteban [21] proposed the auto rate fallback (ARF) protocol. If two consecutive
 18 ACKs are not received by a source node, the source node decreases the data rate for the next
 19 transmission. If consecutive 10 ACKs are successfully received by the source node, the source
 node increases the data rate for the next transmission. Holland *et al.* [20] proposed a receiver-based
 21 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

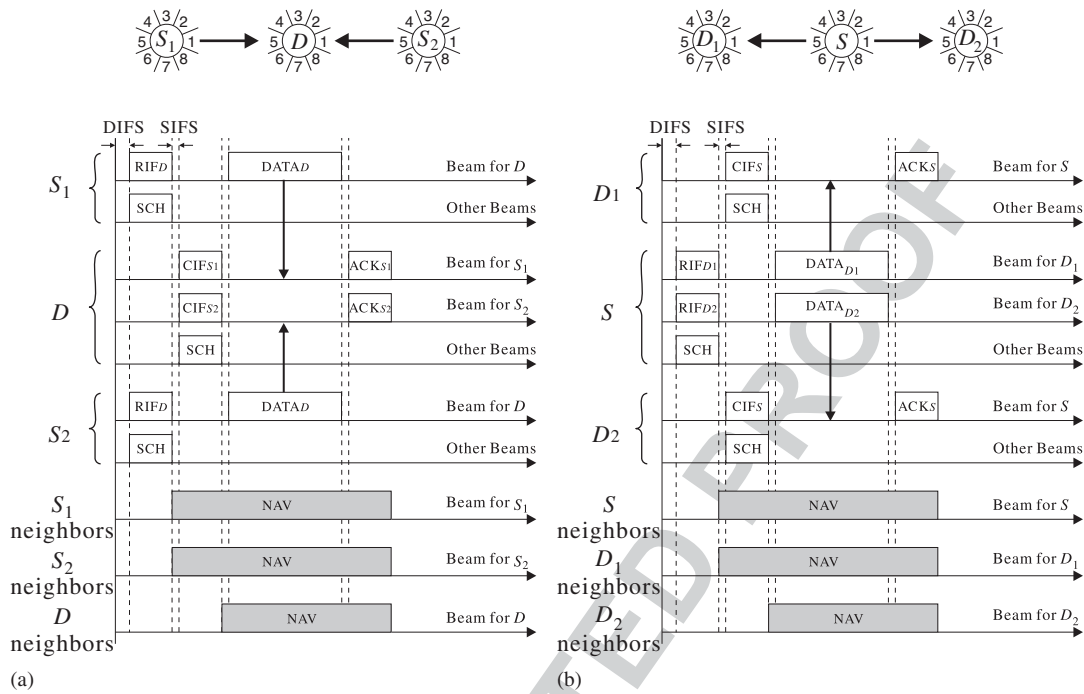


Figure 2. Basic operations of ESIF protocol of: (a) many-to-one communication and (b) one-to-many communication.

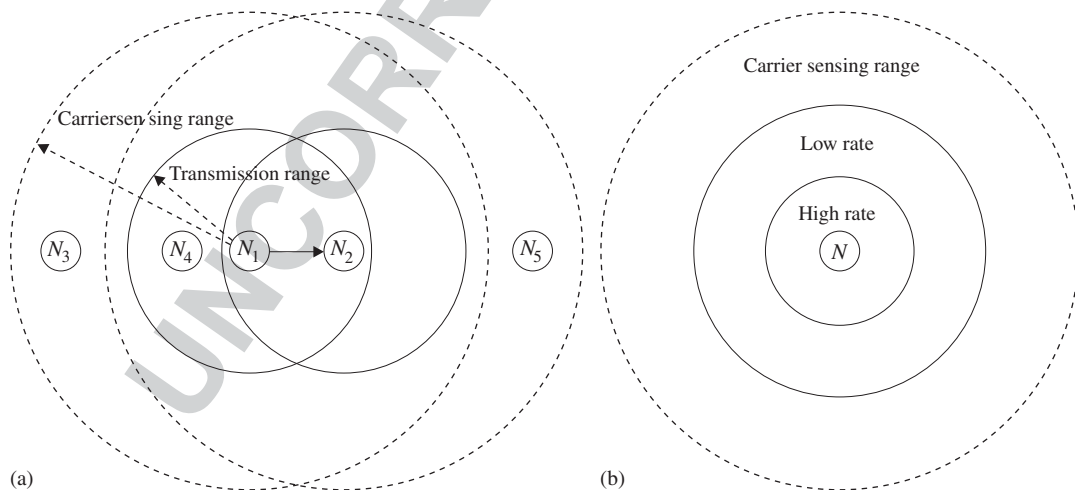


Figure 3. Range relationships: (a) transmission range and carrier sensing range and (b) low data rate range and high data rate range.

1 the estimated SNR, so as to notify the source node. Finally, the source node transmits a frame at
2 the proper transmission rate to the receiver node. In addition, RBAR modifies the original 802.11
3 RTS/CTS frames and MAC header of the DATA frame to set the proper NAV. Sadeghi *et al.* [22]
4 proposed an opportunistic auto rate (OAR) protocol to improve RBAR. OAR exploits the concept
5 of the fragment burst to improve the throughput of high data rate nodes. Unfortunately, these
6 schemes cannot solve the anomaly problem [16] incurred by the multirate wireless network.

7 2.3. Relay-based MAC protocols

8 Zhu and Cao [18] recently proposed a relay-enabled MAC protocol, namely *rDCF*. *rDCF* performs
9 the triangular handshake to improve the throughput of lower data rate nodes in multirate MANETs.
10 According to the triangular handshake, the destination chooses direct transmission or relay and
11 notifies the source. Zou *et al.* [19] proposed a relay-aided media access (RAMA) protocol to
12 improve DCF-based MAC protocol, RBAR. RAMA is similar to *rDCF*. The source chooses direct
13 transmission or relay according to RTS/CTS handshake information and a relay list is obtained
14 from a previous invitation frame of the relay node. More recently, Tan *et al.* [17] proposed CODE
15 protocol to further improve the single relay-base MAC protocol, *rDCF* and RAMA, by using
16 multiple relay nodes. The concept of multiple relay nodes is originated by *lantern* that is proposed
17 by Chen and Ko [1]. The lantern is constructed by multiple relay nodes between a pair of source
18 and destination. Chen and Ko utilize the *lantern* to develop a quality of service on-demand routing
19 protocol in MANETs [1]. However, all of the mentioned relay-based MAC and routing protocols
20 assume that a node is equipped with an omnidirectional antenna. No multiple relay-based MAC
21 protocol is designed for the multirate MANETs with multiple beam antennas. The contribution
22 of this work is to develop a new multiple relay-based MAC protocol in multirate MANETs with
23 multiple beam antennas.

3. PRELIMINARY

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

27 3.1. System model

28 The system model mainly follows the model in [12]. The model in [12] is multiple beam antennas
29 with fixed rate capability. The model of this work is multiple beam antennas with multirate
30 capability. Each mobile node forms M fixed-directional and non-overlapping beams, each beam
31 can simultaneously transmit its data by a different data rate. That is, a mobile node can at most
32 simultaneously communicate with M neighboring nodes by M beams. The angle of each beam is
33 $360^\circ/M$ and the beam shape is assumed to be conical. A node can either simultaneously transmit
34 or receive multiple data on multiple beams. Observe that the transmitting and receiving operations
35 do not happen at the same time. The interferences of side and back lobes are ignored in this work.
36 Observe that the carrier sensing range and the transmission range of multiple beam antennas are
37 the same as that of an omnidirectional antenna. The wireless channel is assumed to be ideal and
38 symmetric, and received power gains are equal in both directions. In the multiple beam antennas,
39 the data rate information is obtained by precisely calculating the angle of arrival of the received
40 signal and the estimated SNR of the received power gain. Our work is based on ESIF protocol [12]

1 to support the parallel multiple-communications. The physical layer transmits data by an assigned
 2 data rate requested by the MAC sublayer. The physical layer provides the channel status, received
 3 power gain, and SNR to the MAC layer. To avoid the hidden terminal problem, control frames are
 4 transmitted by the lowest data rate. Data frames can be transmitted by any data rates supported
 5 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.

7 *3.2. Basic idea*

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

$$\frac{1}{\text{Rate}_{SR}} + \frac{1}{\text{Rate}_{RD}} < \frac{1}{\text{Rate}_{SD}}$$

11 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.

13 As mentioned before, the technique of multiple beam antennas is similar to the technique of
 SDMA. M beams in multiple beam antennas can be treated as M individual data channels. Only
 15 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 $M - 1$
 17 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
 19 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
 21 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 .

Figure 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 Figures 4(a)–(d).

$$T_a = \frac{L}{\text{Rate}_{SD}} \tag{1}$$

$$T_b = \frac{L}{\text{Rate}_{SR}} + \frac{L}{\text{Rate}_{RD}} \tag{2}$$

$$T_c = \max \left(\frac{L}{2 \times \text{Rate}_{SD}}, \frac{L}{2 \times \text{Rate}_{SR}} + \frac{L}{2 \times \text{Rate}_{RD}} \right) \tag{3}$$

$$T_d = \max \left(\frac{L}{3 \times \text{Rate}_{SR_1}} + \frac{L}{3 \times \text{Rate}_{R_1D}}, \frac{L}{3 \times \text{Rate}_{SR_2}} + \frac{L}{3 \times \text{Rate}_{R_2D}}, \frac{L}{3 \times \text{Rate}_{SR_3}} + \frac{L}{3 \times \text{Rate}_{R_3D}} \right) \tag{4}$$

23 We can observe that $T_d < T_c < T_b < T_a$. The motivation of this work is that the transmission
 25 latency of low data rate link can be significantly reduced by distributing a frame into multiple

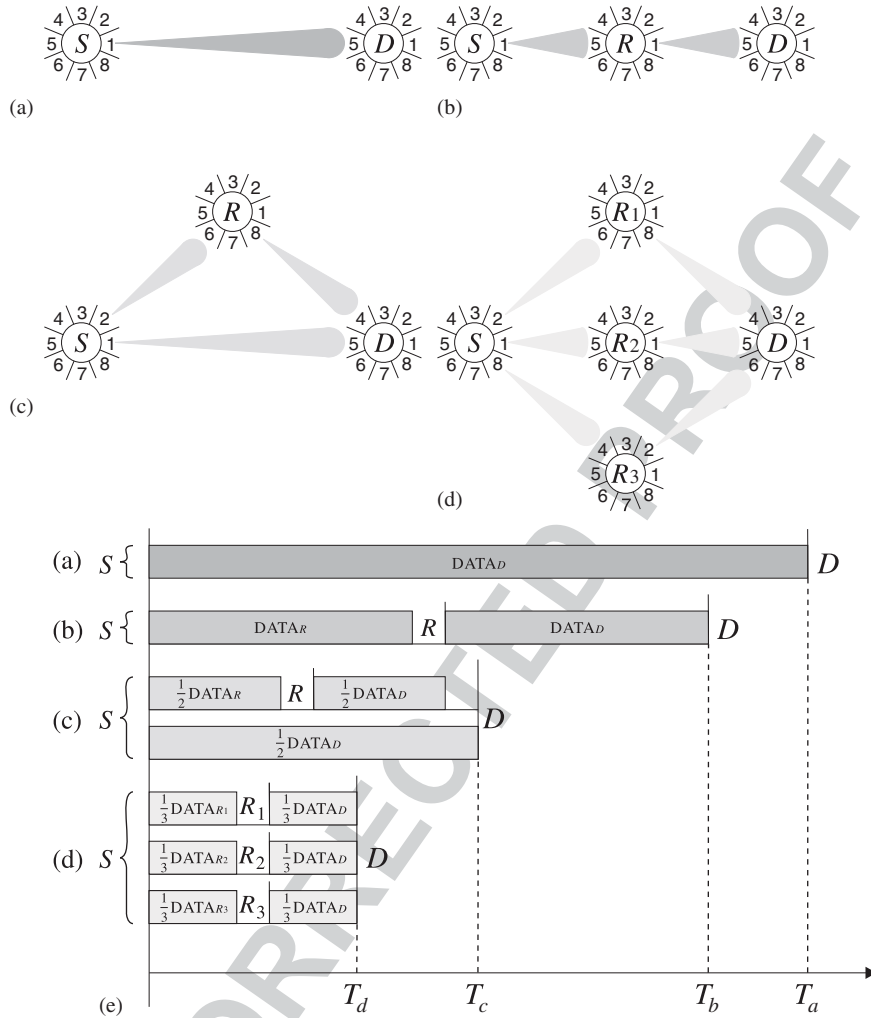


Figure 4. Basic idea of multiple relay-based MAC protocol: (a) original direct transmission; (b) direct relay; (c) combine direct transmission with indirect relay; and (d) multiple relay transmission.

1 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.

3 3.3. 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. Figure 5 indicates the ACK synchronization problem. In this paper, each node uses multiple beam antennas with multirate capability, the node can simultaneously communicate with multiple neighboring nodes. But all beams must have an uniform communication mode at the same time. Different beams have different communication latency by different data rate and thus S may not be able to correctly

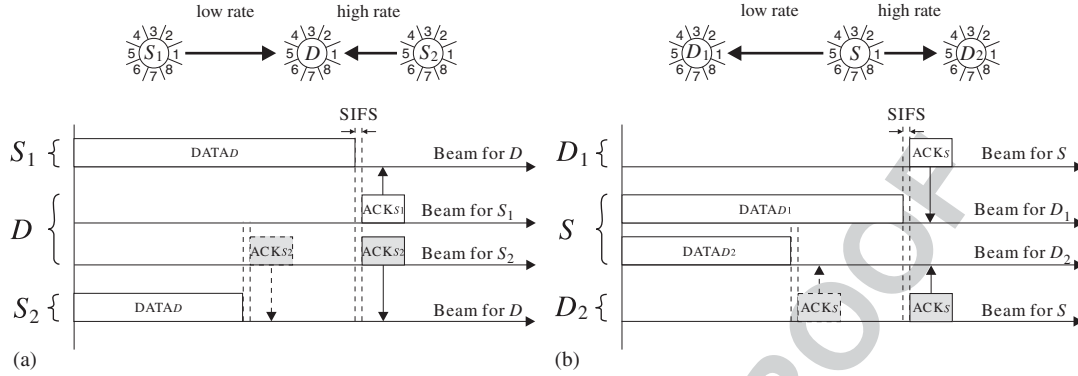


Figure 5. The ACK synchronization problem: (a) many-to-one multirate communication and (b) one-to-many multirate communication.

- 1 receive an ACK from D before occurring the ACK timeout even if the data frame is correctly received by D .
- 3 Figure 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 $Rate_{S_1 D} < Rate_{S_2 D}$.
- 5 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. Figure 5(b) gives a similar ACK synchronization problem for one-to-many multirate communications.
- 11

Figure 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 $Rate_{S_1 D} < Rate_{S_2 D}$. Figures 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 Figures 6(a)–(c), three reception latencies are denoted as $T_{a'}$, $T_{b'}$, and $T_{c'}$ as follows:

$$T_{a'} = \max\left(\frac{L}{Rate_{S_1 D}}, \frac{L}{Rate_{S_2 D}}\right) \quad (5)$$

$$T_{b'} = \max\left(\frac{L}{Rate_{S_1 D}}, \frac{L}{Rate_{S_2 R}} + \frac{L}{Rate_{RD}}\right) \quad (6)$$

$$T_{c'} = \max\left(\frac{L}{Rate_{S_1 R}} + \frac{L}{Rate_{RD}}, \frac{L}{Rate_{S_2 D}}\right) \quad (7)$$

- 13 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
- 15 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

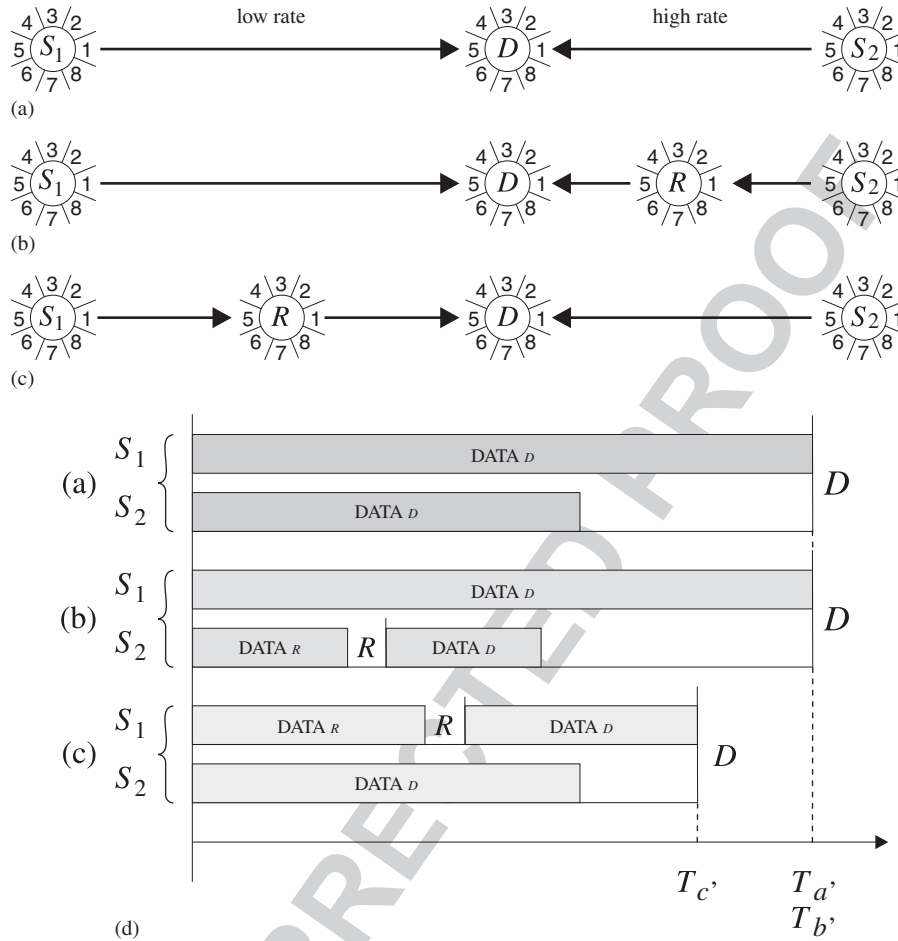


Figure 6. The unnecessary relay problem: (a) original direct transmission; (b) helpless relay; and (c) helpful relay.

1 of node S_1 with the assistance of relay node R as shown in Figure 6(c). In addition, there is a
 3 similar problem for one-to-many communication. Unfortunately, nodes S_1 and S_2 cannot know this
 5 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 Figure 6.

7 4. MULTIPLE RELAY-BASED MAC PROTOCOL

9 Our new MAC protocol is based on ESIF protocol [12] to achieve the parallel communication
 capability. This paper extends the ESIF Network Allocation Vector (ENAV) in [12] by adding

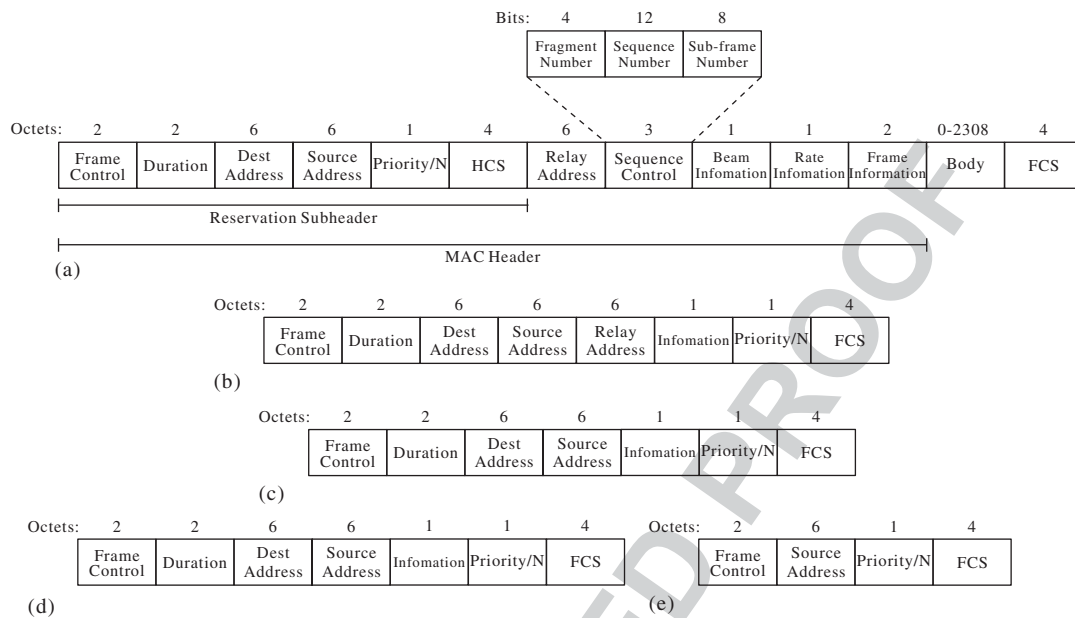


Figure 7. Frame formats of multiple relay-based MAC protocol: (a) MAC frame format; (b) RIF, ATS, SCH, RD; (c) CIF, WTS, SCH, RD, ACK; (d) PTS; and (e) RESET.

1 a relay node table to keep the information of multiple relay nodes. In addition, control frames,
 3 such as RIF, CIF, SCH, and MAC header frame in [12] are modified to support multiple relay
 5 nodes. Figure 7 shows frame formats of our approach. The ‘Priority/N’ field, which exists in
 7 ESIF protocol [12], is used to inform neighboring node its individual schedule since ESIF protocol
 9 removes the contention window. ‘Type’ and ‘Subtype’ fields of Frame Control field can be used
 11 to distinguish these control frames. The first three bits of the Information field record the beam
 13 number of a mobile node, where each mobile node assumes to have eight beams in this work.
 15 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 are discovered to provide the information about 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.

17 4.1. Relay nodes discovery

Given a pair of nodes S and D , node S should keep the information about multiple relay nodes
 19 between S and D . This information is obtained by performing the relay nodes discovery. According
 to Equation (1), a relay region exists between S and D , where relay nodes exist in the relay region.

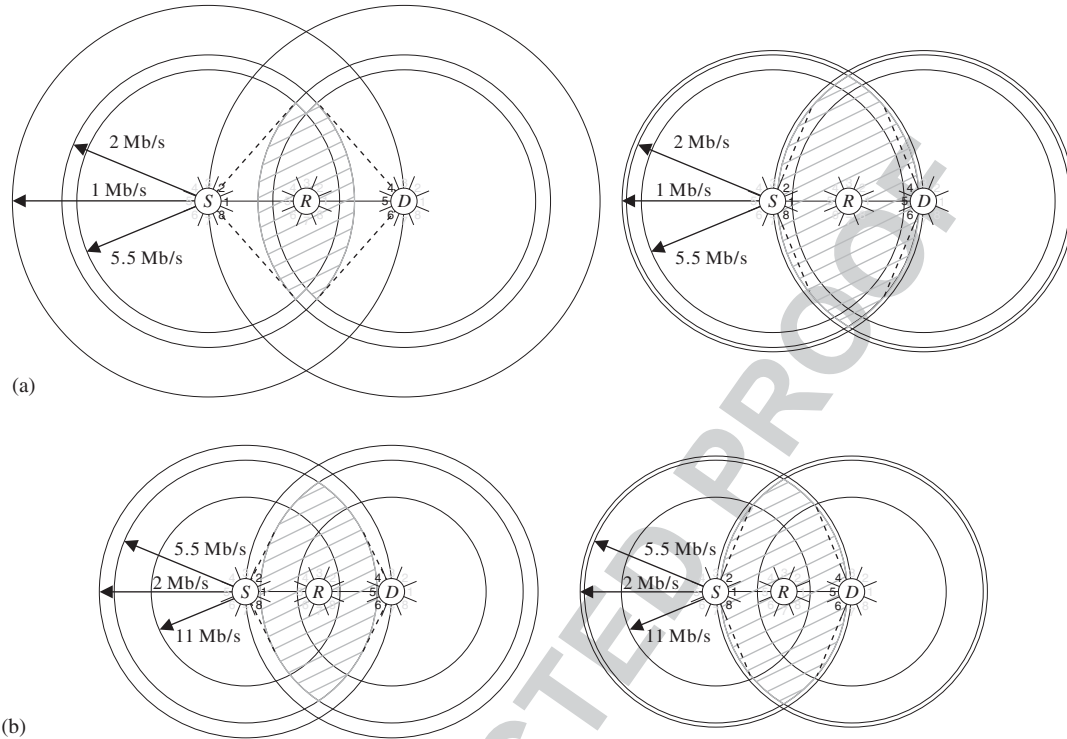


Figure 8. Examples of the relay region: (a) 1, 2, and 5.5 Mbps and (b) 2, 5.5, and 11 Mbps.

1 Assume that each mobile node has M beams. Given that beam ‘ i ’ of S and beam ‘ j ’ of D can aim
 2 at each other’s directions. The relay region is contained between beams ‘ i ’, ‘ $(i - 2 + M) \% M + 1$ ’,
 3 ‘ $(i + M) \% M + 1$ ’, of S and beams ‘ j ’, ‘ $(j - 2 + M) \% M + 1$ ’, ‘ $(j + M) \% M + 1$ ’ of D . Without loss
 4 of generality, example is given as shown in Figure 4(a), if $i = 1$ and $j = 5$, then relay regions are
 5 contained between beams 1, 2, 8 of S and beams 4, 5, 6 of D through this investigation. Node
 6 S and D can possibly speculate the positions of relay nodes. Suppose that transmission ranges of
 7 1, 2, 5.5, and 11 Mbps are 250, 186.75, 167.75, and 120.5 m, respectively. Using Equation (1),
 8 Figure 8(a) shows a relay region whose $Rate_{SD} = 1$ Mbps. Figure 8(b) illustrates a relay region
 9 whose $Rate_{SD} = 2$ Mbps.

10 (1) *Relay nodes discovery operation*: Assume that S_i and S_j communicate with D by $Rate_{S_i D}$
 11 and $Rate_{S_j D}$. The relay nodes discovery is performed to search for the possible relay nodes
 12 between S_i and D , and S_j and D . Example is given in Figure 9 for S_1 and S_2 communicate with
 13 D by $Rate_{S_1 D}$ and $Rate_{S_2 D}$. Let $S_i(b_1, \dots, b_m)$ denotes the node S_i which simultaneously sends
 14 specific message through beams b_1, \dots, b_m , where $1 \leq m < Max_beam_number$. Let $S_i(\overline{b_1, \dots, b_m})$
 15 denotes the node S_i that simultaneously sends specific message through all other beams except
 16 for beams b_1, \dots, b_m , where $1 \leq m < Max_beam_number$. The relay nodes discovery operation
 17 is stated as follows.

- 18 A1. Nodes $S_i(b_1)$ and $S_j(b_1)$ simultaneously send an RIF message to node D .
- 19 A2. Nodes $S_i(b_2, \dots, b_m)$ and $S_j(b_2, \dots, b_m)$ simultaneously send a new control frame, called
 the relay discovery (RD), to attempt to find relay nodes, where the RD message of source

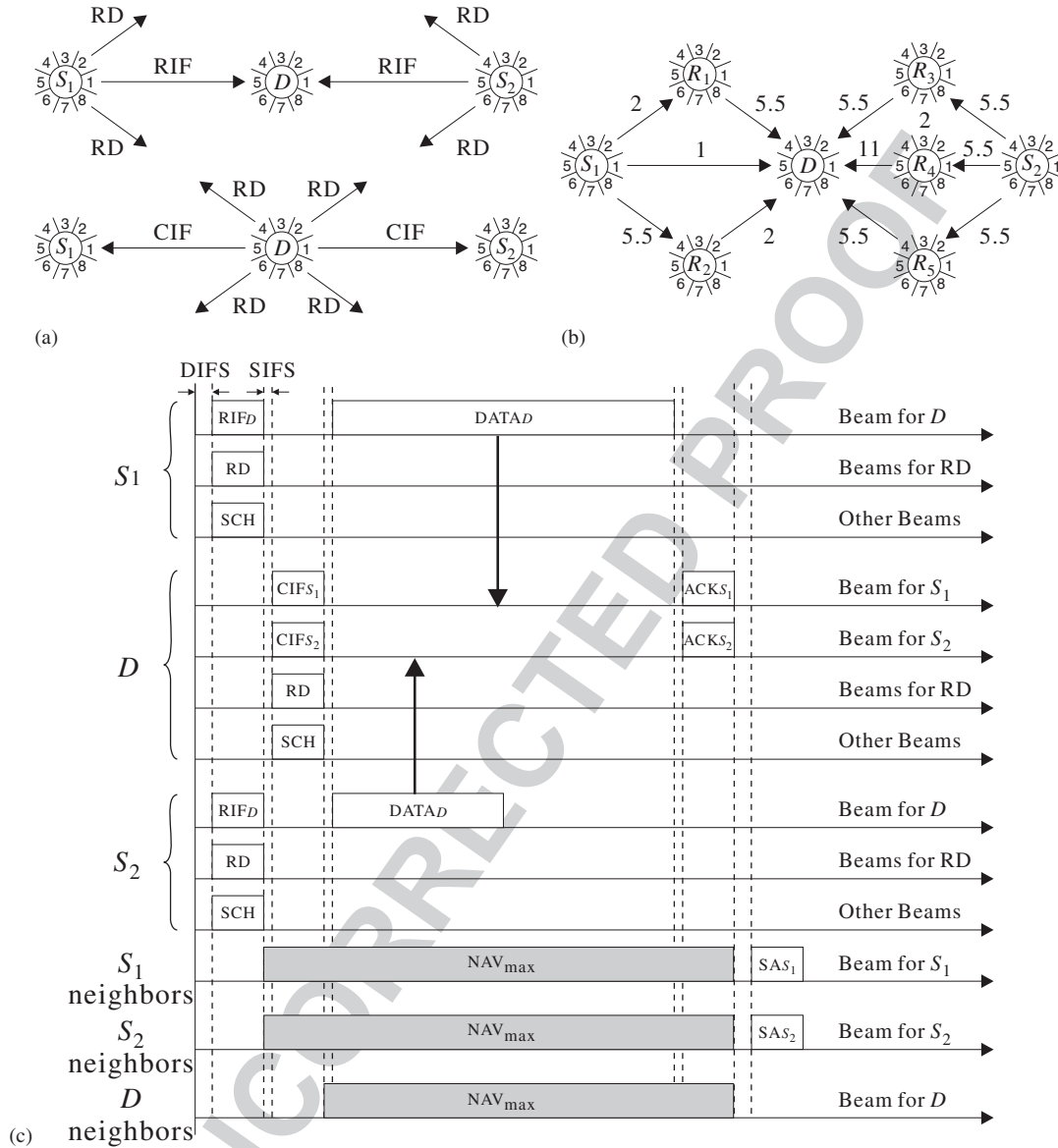


Figure 9. Example of relay discovery: (a) beams for relay discovery; (b) relay discovery result; and (c) the detailed procedure.

1 node S contains MAC address of the desired destination node D and transmission beam
 number. Nodes $S_i(b_1, \dots, b_m)$ and $S_j(b_1, \dots, b_m)$ send the SCH message for the transmission
 3 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 $S_2(1, 2, 3, 7, 8)$ send the SCH message.

- 1 A3. Node D simultaneously receives two RIF messages from S_i and S_j and then estimates
Rate $_{S_i D}$ and Rate $_{S_j D}$ according to the measured SNR.
- 3 A4. Node D fills the information about the 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
5 for an SIFS period.
- 7 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
9 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.
- 11 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 $_{S_i D}$ from the CIF or RD message from D after hearing the
13 RIF or RD message from S_i . Node R_X applies Equation (1) to check whether R_X is a relay
node between S_i and D .
- 15 A7. After S_i , S_j , and D finishing the current transmission, relay node R_X unicasts a service
advertisement (SD) message to S_i and S_j , where the SD message contains Rate $_{S_i R_X}$,
17 Rate $_{R_X D}$, 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
19 message of D .

(2) *Solving the ACK Synchronization Problem:* Figure 9(b) gives an example of the relay node
21 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 us recall the ACK synchronization problem. Suppose that S_1 and S_2 send out the same size
23 frames to D by data rates Rate $_{S_1 D}$ and Rate $_{S_2 D}$, respectively. Transmission latencies of S_1 and S_2
are not equal since Rate $_{S_1 D}$ (1 Mbps) is smaller than Rate $_{S_2 D}$ (2 Mbps). The ACK synchronization
25 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
27 frame to D .

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

- 33 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 .
- 35 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
37 addition, the NAV durations of RIC, CIF, RD, and SCH messages are set to be the maximal
transmission delay.
- 39 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
41 to reset the NAV durations of all neighboring nodes.
- 43 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.

1 4.2. Multiple relay-based MAC operation

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

17 One-to-many communication case may be occurred, and S simultaneously communicates
 19 with multiple destination nodes D_i , where $i > 1$. The source node S attempts to allot possible
 21 relay nodes to all the destination nodes. To consider the unnecessary relay problem, S may
 23 have high priority to allot usable relay nodes from the relay node tables. Note that S does
 25 not allot any D_k , if D_j cannot allot at least a relay node, where Rate_{SD_j} is the lowest
 27 among all Rate_{SD_k} and $k \leq i$. This is because that the total transmission delay is limited by
 29 the lowest data rate node even if transmission delays of other lower data rate nodes can be
 31 improved.

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

37 (1) *Allotting relay-nodes operation*: The algorithm of S allotting proper relay nodes to desired
 39 D is given in Algorithm 1. Before describing the algorithm for one-to-many communication,
 41 some notations are defined. D is a one-dimensional array to record the desired destination nodes
 43 ($D_i, 1 \leq i \leq n$) of S . C is a one-dimensional array to record the next hop nodes ($C_i, 1 \leq i \leq n$)
 45 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
 47 for D_i uses indirect transmission. Note that C is used when there is at most one relay node for
 49 each D_i . If there are more relay nodes, C' is used. Thus, C' is a two-dimensional array to store
 51 more relay nodes of D_i and C'_i is a one-dimensional array to store extra relay nodes allotted
 53 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
 55 are one-dimensional arrays. Elements of B are unassigned free beams for D_i . Element of T is
 57 the transmission latency between S and D_i . T is counted by procedure *CountSource*. Example is
 59 given in Figure 10, since node D is initially not allotted any relay node, initial setting of node
 61 S_2 is $D_1 = \text{'nodeD'}$ and $C_1 = \text{'nodeD'}$. $C'_1 = \emptyset$ and $B = \{1, 2, 3, 4, 5, 6, 7, 8\}$ because S_2 can use
 63 eight beams to communicate. T_1 is the direct transmission latency between nodes S_2 and D_1 .
 65 After the initial setting, S_2 then performs *AllotRelayNodes* procedure. The main operations of
 67 *AllotRelayNodes* procedure are explained as follows.

- 69 C1. All beams for desired destinations D are removed since these beams are used to exchange
 71 control frames. For instance, $B - B_5 = \{1, 2, 3, 4, 6, 7, 8\}$ (line 2 of Procedure *AllotRelayNodes*).
 73 C2. Let T_{S, D_i} be the transmission latency between S and D_i , $1 \leq i \leq n$, let $T_{\max} = \max_{i=1..n} T_{S, D_i}$.
 75 If $T_{S, D_i} = T_{\max}$, S will try to allot relay nodes. Some temporal parameters are used to store
 77 results of allocating relay nodes (lines 5–6 of Procedure *AllotRelayNodes*).
 79 C3. S performs *Allocate* procedure S to allot relay node to D_i . The value of *True* or *False*
 81 returned from *Allocate* procedure represents the success or failure of the allocation operation
 83 (line 7 of Procedure *AllotRelayNodes*).
 85 C4. Some parameters are required to be updated by temporal parameters if *Allocate* procedure
 87 returns *True* value, where *cycle* is the allocation time. Finally, *AllotRelayNodes* procedure

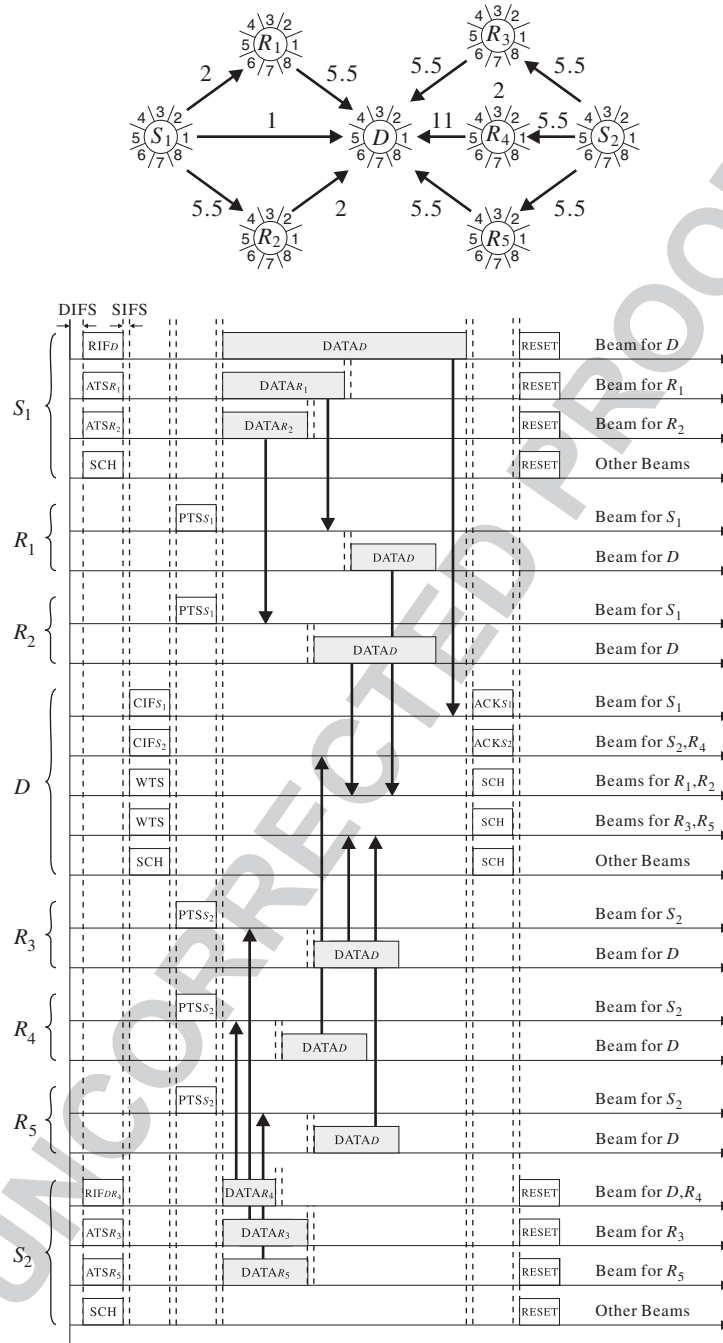


Figure 10. Example of multiple relay-based MAC protocol and the detailed procedure.

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 $\begin{cases} \text{if } S \text{ sends to } D_i \text{ directly} \\ C_i = D_i \\ \text{if } S \text{ sends via } R_x \text{ indirectly} \\ C_i = R_x \end{cases}$
 $C' = \{ C'_i \mid 1 \leq i \leq n \}$, where $\begin{cases} \text{if } D_i \text{ has no extra relay nodes} \\ C'_i = \emptyset \\ \text{if } D_i \text{ has extra relay nodes} \\ C'_i = \{ R_j \mid 1 \leq j \leq \gamma \}, \end{cases}$
 $B = \{ B_i \mid 1 \leq i \leq M, n \leq M \}$, 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 .

```

1 Procedure AllotRelayNodes()
2   B = B-beams for  $D_1 \sim D_n$ ;
3   repeat = True; cycle = 0;
4   while repeat == True do
5      $T_{max} = \max\{T_1, T_2, \dots, T_n\}$ ;
6      $b = B$ ;  $c = C$ ;  $c' = C'$ ;  $t = T$ ;
7     repeat = Allocate( $T_{max}, S, D, b, c, c', t$ );
8     if repeat == True then
9        $B = b$ ;  $C = c$ ;  $C' = c'$ ;  $T = t$ ;
10    cycle = cycle + 1;
11  end if
12  end while
13  return( cycle, C, C' )
14 end Procedure

```

1 returns cycle, C, and C' to S by executing Allocate procedure (lines 8-10 of Procedure AllotRelayNodes).

3 C5. Go to step C2.

The operation of Allocate procedure is stated as follows.

5 D1. Repeatedly performing B2–B3 steps for $|d|$ times.

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

9 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 Allocate).

11 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).

13 For instance, as shown in Figure 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.

Procedure : Allocate()

```

1 Procedure Allocate(  $T_{max}$ ,  $S$ ,  $D$ ,  $b$ ,  $c$ ,  $c'$ ,  $t$  )
2    $allot = False$  ;
3   for  $i = 1$  to  $|d|$ 
4     if  $t_i < T_{max}$  then continue ;
5      $allot = False$  ;
6     if  $c_i = D_i$  && beam for  $D_i$  has a relay node  $R_x$  then
7       if  $t_i > CountSource( S, D_i, c_i \cup c'_i \cup R_x )$  then
8          $c_i = R_x$  ;  $allot = True$  ;
9          $t_i = CountSource( S, D_i, c_i \cup c'_i )$  ;
10        continue ;
11      end if
12    end if
13     $allot = False$  ;
14    for  $j = 1$  to  $|b|$ 
15      if beam  $b_j$  has a relay node  $R_y$  for  $D_i$  then
16        if  $t_i \leq CountSource( S, D_i, c_i \cup c'_i \cup R_y )$  then continue ;
17         $c'_i = c'_i \cup R_y$  ;
18         $b = b - b_j$  ;
19         $t_i = CountSource( S, D_i, c_i \cup c'_i )$  ;
20         $allot = True$  ;
21      end if
22    next  $j$ 
23  next  $i$ 
24  return  $allot$  ;
25 end Procedure

```

Procedure : CountSource()

```

1 Procedure CountSource(  $S$ ,  $D$ ,  $\Gamma$  )
2    $k = |\Gamma|$  ;
3    $t = \max( T_{overhead_1} + \frac{MAC + \frac{Payload_D}{k}}{Rate_1}$ ,  $T_{overhead_2} + \frac{MAC + \frac{Payload_D}{k}}{Rate_2}$ , ...
4     ,  $T_{overhead_k} + \frac{MAC + \frac{Payload_D}{k}}{Rate_k}$  ) ;
5     , where  $\begin{cases} \text{if } \Gamma_p \neq D \\ T_{overhead_p} = 2T_{physical} + SIFS, Rate_p = \frac{Rate_{S\Gamma_p} \times Rate_{\Gamma_p D}}{Rate_{S\Gamma_p} + Rate_{\Gamma_p D}} \\ \text{if } \Gamma_p == D \\ T_{overhead_p} = T_{physical}, Rate_p = Rate_{SD} \end{cases}$ 
6   if  $k == 1$  &&  $\Gamma_1 == D$  then
7      $t = t + RIF + CIF + ACK + 3SIFS$  ;
8   else
9      $t = t + RIF + CIF + PTS + ACK + 4SIFS$  ;
10  end if
11  return  $t$  ;
12 end Procedure

```

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

- 4 E1. Let k represent the number of data channels for D_i . The source divides the current data
 5 frame into k sub-frames, and each one is simultaneously transmitted to all nodes of Γ_p (line
 6 1 of Procedure *CountSource*).
 7 E2. The transmission delay t is calculated (line 3 of Procedure *CountSource*).
 8 E3. The actual transmission delay t is calculated with (or without) relay nodes (lines 4–8 of
 9 Procedure *CountSource*).

10 For example, in Figure 10, S_1 uses Algorithm 1 to search for relay nodes to communicate with
 11 D . The returned values for S_1 are $C_1 = D$ and $C'_1 = \{R_1, R_2\}$. Similarly, S_2 also uses Algorithm
 12 1 to search for relay nodes to communicate with D . The returned values for S_2 are $C_1 = R_4$ and
 13 $C'_1 = \{R_3, R_5\}$, which means that S_2 divides its data frame into three sub-frames. These sub-frames
 14 are transmitted to D through relay nodes R_3, R_4 , and R_5 . Another example is that S_1 divides its data
 15 frame into three sub-frames, one is through direct communication and the other two sub-frames
 16 are through relay nodes R_3 and R_4 .

17 (2) *Multiple relay-based frame exchange operation*: Without loss of generality, we consider
 18 the many-to-one communication scenario of S_i and S_j , who simultaneously communicate with
 19 D to illustrate the multiple relay-based frame exchange operation. Example is given in Figure
 20 10, S_1 and S_2 have a common destination D . After performing Algorithm 1, the return values
 21 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 $C'_1 = \{R_3, R_5\}$.
 22 Therefore, we have the scenario of $S_1(1)$ communicates with $D(5)$ by a direct communication,
 23 $S_1(2, 8)$ simultaneously communicates with $D(4, 6)$ through relay nodes R_1 and R_2 ; $S_2(4, 5, 6)$
 24 simultaneously communicates with $D(2, 1, 8)$ through relay nodes R_3, R_4 and R_5 . The multiple
 25 relay-based frame exchange operation is explained as follows.

- 26 F1. Initially, $S_i(b_1)$ and $S_j(b'_1)$ directly send out the RIF message to D . $S_i(b_2, \dots, b_m)$ send out
 27 new control frames, called *appoint to send* (ATS) message, to relay nodes R_1, R_2, \dots , and
 28 R_{m-1} , respectively. $S_j(b'_2, \dots, b'_n)$ send out the ATS message to relay nodes R'_1, R'_2, \dots , and
 29 R'_{n-1} , respectively. The ATS message is used to appoint one neighbor node as a relay node.
 30 F2. For S_i after sending the ATS message for all relay nodes R_1, R_2, \dots , and R_{m-1} , relay nodes
 31 R_1, R_2, \dots , and R_{m-1} should recognize that S_i asks all relay nodes to relay sub-frames to
 32 D . For S_j after sending the ATS message for all relay nodes R'_1, R'_2, \dots , and R'_{n-1} , each
 33 relay node, $R'_q, 1 \leq q \leq n-1$, recognizes that its MAC address is in the ATS message.
 34 F3. Node D responds the CIF message to S_i and S_j after D successfully received the RIF
 35 message and waiting for an SIFS period. Observe that, the information fields about the CIF
 36 message for S_i and S_j have the latest $\text{Rate}_{S_i D}$ and $\text{Rate}_{S_j D}$, respectively. According to the
 37 information fields of RTS, corresponding beams of node D send new control frames, called
 38 the *wish to send* (WTS) message through all relay nodes to S_i and S_j .
 39 F4. Assume that a candidate relay node R_c acquires $\text{Rate}_{S D}$ by acquiring this information
 40 from the WTS message and calculates the $\text{Rate}_{S R_x}$, $\text{Rate}_{R_x D}$, and $\text{Rate}_{S D}$ according to the
 41 measured SNR of the ATS and WTS messages. If R_c satisfies Equation (1), it will become
 42 an appointed relay node R_a . When node R_a is identified, R_a sends a new control frames,
 43 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.

- 1 F5. The PTS message keeps the latest $\text{Rate}_{S R_a}$ and $\text{Rate}_{R_a D}$. The source node S recomputes the
 3 number of relay nodes, since some appointed relay nodes do not actually respond the PTS
 5 message or the information about the PTS message is not the same as the record in the relay
 7 node table. The relay node may receive RIF (ATS) message of multiple sources at the same
 9 time; however, the relay node responds the PTS message to only one source node S . After
 11 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.
- 9 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,
 11 SCH, ATS, and WTS messages are set to the maximal transmission time. This may reduce
 the NAV length.

13 Example is given in Figure 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
 15 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
 17 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
 19 combination.

4.3. Eliminating unnecessary relay nodes

21 Our multiple relay-based MAC protocol supports a dynamic relay-node management by dynam-
 ically eliminating unnecessary relay nodes under the fragment burst transmission. To take the
 23 unnecessary relay node problem into account, this work aims at providing a dynamic relay-node
 management to release some unnecessary relay nodes to improve the system throughput.

25 A destination node D knows the whole relay-node topology and transmission latency of all the
 source nodes S_x after receiving data frames from all S_x , where $1 \leq x \leq n$. Algorithm 2 is performed
 27 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
 29 the adjusted information.

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

- 35 G1. T_{\max} represents the maximum transmission delay of all source nodes S_x , where $1 \leq x \leq n$,
 $T_{\max} = \max\{T_1, T_2, \dots, T_n\}$, T_x is the transmission delay between S_x to D (line 2 of Procedure
 37 *AdjustRelayNodes*).
- G2. Any of the source nodes S_x , $1 \leq x \leq n$, whose transmission delay is smaller than T_{\max} , has
 39 opportunity to remove unnecessary relay nodes.
- G3. For each S_x , $1 \leq x \leq n$, execute G3 and G4.
- 41 G4. Call Procedure *Remove* ($T_{\max}, S_x, D, c, c', extra$) if $c' \neq \emptyset$.
- G5. Call Procedure *Remove* ($T_{\max}, S_x, D, c, c', extra$) if $c' \neq S_x$.
- G6. goto G3.

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

$C_i = S_i$	$\left. \begin{array}{l} \text{if } D \text{ receives from } S_i \text{ directly} \\ \text{if } D \text{ receives from } R_x \text{ indirectly} \end{array} \right\}$
$C_i = R_x$	

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

$C'_i = \emptyset$	$\left. \begin{array}{l} \text{if } S_i \text{ has no extra relay nodes} \\ \text{if } S_i \text{ has extra relay nodes} \end{array} \right\}$
$C'_i = \{ R_j \mid 1 \leq j \leq \gamma \}$	

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

```

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

```

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

Procedure : Remove()

```

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

```

Procedure : CountDestination()

```

1 Procedure CountDestination(  $S, D, \Gamma$  )
2    $k = | \Gamma |$  ;
3    $t = \max( T_{overhead_1} + \frac{MAC + \frac{Payload_D}{k}}{Rate_1}, T_{overhead_2} + \frac{MAC + \frac{Payload_D}{k}}{Rate_2}, \dots$ 
      ,  $T_{overhead_k} + \frac{MAC + \frac{Payload_D}{k}}{Rate_k} )$  ;
      , where  $\left\{ \begin{array}{l} \text{if } \Gamma_p \neq D \\ \quad T_{overhead_p} = 2T_{physical} + SIFS, \ Rate_p = \frac{Rate_{S\Gamma_p} \times Rate_{\Gamma_p D}}{Rate_{S\Gamma_p} + Rate_{\Gamma_p D}} \\ \text{if } \Gamma_p == D \\ \quad T_{overhead_p} = T_{physical}, \ Rate_p = Rate_{SD} \end{array} \right.$ 
4    $t = t + ACK + SIFS$  ;
5   return  $t$  ;
6 end Procedure

```

1 5. PERFORMANCE EVALUATION

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

5.1. Analysis of transmission delays

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

9 (1) *NAV Duration*: Tables I–III show NAV durations of various control frames. Table I represents
10 that the source and destination nodes do not meet the fragment burst. The first row (RIF, ATS,
11 SCH) contains the control frames sent by the source nodes. The second row (CIF, WTS, SCH)
12 contains the control frames sent by the destination nodes. The NAV durations of the above control
13 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

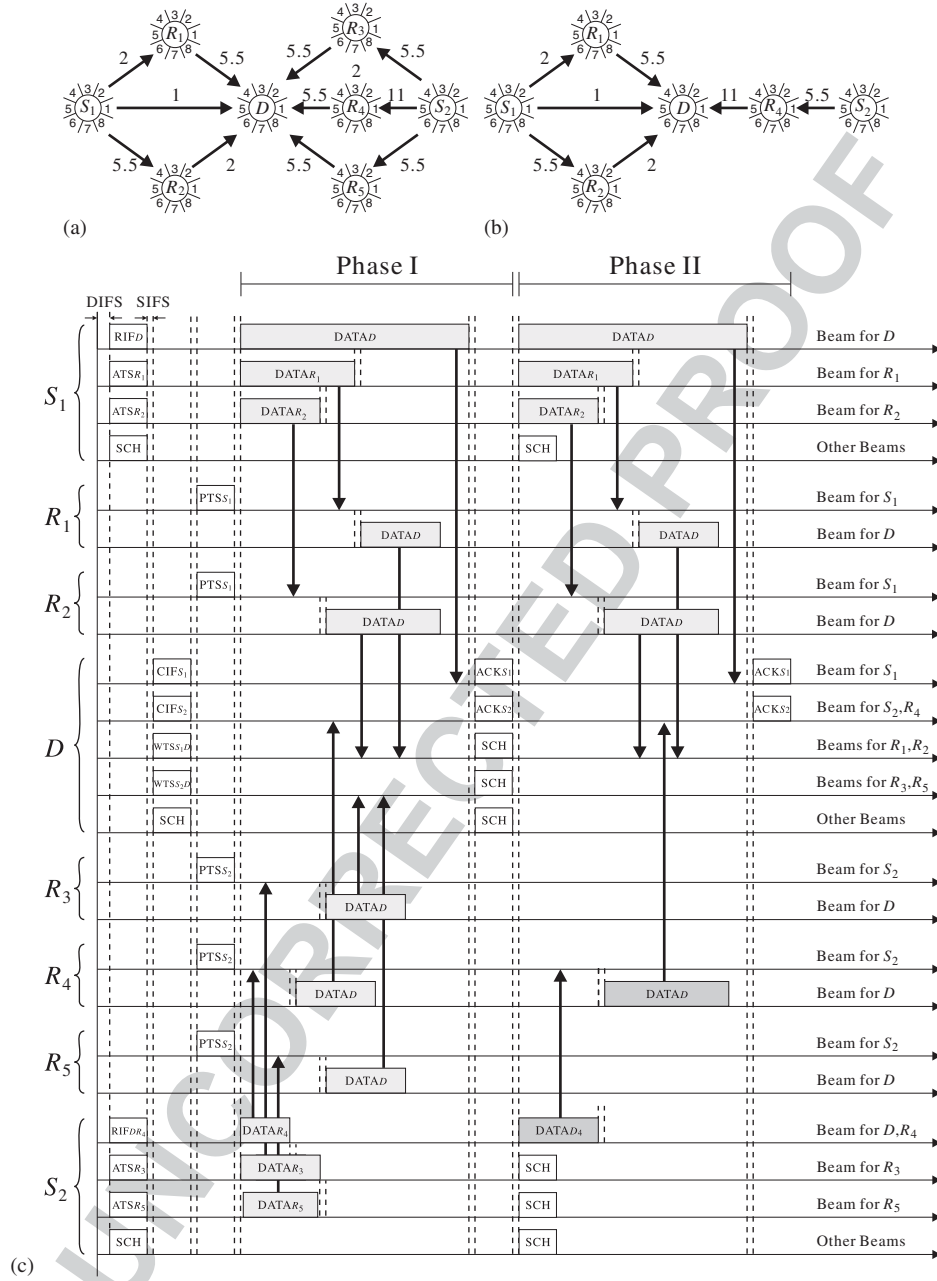


Figure 11. Example of fragment burst with adjusting relay nodes: (a) Phase I: before adjusting; (b) Phase II: after adjusting; and (c) the detailed procedure.

Table I. NAV duration of control frames.

Frame	NAV duration
RIF, ATS, SCH	$CIF + T_{\text{physical}} + \frac{\text{MAC} + \text{Payload}_{\text{max}}}{\text{the lowest data rate}} + \text{ACK} + 3\text{SIFS}$
CIF, WTS, SCH ACK, SCH, RESET	$T_{\text{physical}} + \frac{\text{MAC} + \text{Payload}_{\text{max}}}{\text{the lowest data rate}} + \text{ACK} + 2\text{SIFS}$ 0

Table II. NAV duration during fragment burst period of the destination node.

Frame	NAV duration
ACK ₁ ACK _{n-1}	$\max\{\text{DATA}_{\text{max}}^{S_1}, \dots, \text{DATA}_{\text{max}}^{S_i}\} + \text{ACK} + 2\text{SIFS}$
SCH ₁ SCH _{n-1} ACK _n	$\max\{\text{DATA}_{\text{max}}^{S_1}, \dots, \text{DATA}_{\text{max}}^{S_i}\} + \text{ACK} + 2\text{SIFS}$ 0

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

Table III. NAV duration during fragment burst period of the source node.

Frame	NAV duration
DATA _i at beam _j	NAV Duration of ACK _{i-1} – SIFS – (sub)frame transmission latency at beam _j
SCH _{2n}	NAV Duration of ACK _{i-1} – SIFS – SCH

$(n = \text{number of fragment}, M = \text{number of beams}).$
 $((2in, 1jM))$.

- 1 by the source nodes. Tables II and III are the NAV durations of the destination and source nodes
 2 during the fragment burst period, respectively. Successively multiple data frames and ACKs may
 3 play the role of RIF and CIF, respectively, to set the suitable NAV when the fragment burst occurs.
 4 The destination knows the transmission conditions of each source after receiving all data frames.
 5 Hence, the next fragment burst duration is the maximum transmission latency of all the source
 6 nodes. The destination responds ACKs with the maximum NAV duration to the sources. Other
 7 beams of the destination also send SCH with the maximum NAV duration. The sources understand
 8 the duration of the next fragment burst after receiving ACKs. The sources send the next fragment
 9 an SIFS later. Other beams of the sources also send SCH to set NAV.

(2) *Network parameters*: Network parameters of the mathematic analysis are given in Table IV.

- 11 A transmission delay includes the Inter-frame Space (IFS), several control frames and a data frame.
 12 A data frame consists of a physical preamble, a physical header, a MAC header and a payload.
 13 All the control frames are transmitted at the lowest data rate to avoid hidden terminal problem.
 14 The MAC header and payloads are transmitted at the data rate which is recorded in the physical
 15 header.

$$T_{\text{physical}} = T_{\text{preamble}} + T_{\text{PLCPheader}}$$

- 17 where T_{preamble} and $T_{\text{PLCPheader}}$ are transmitted at the lowest data rate. $T_{\text{PLCPheader}}$ has the infor-
 mation about the data rate of the MAC header and payload.

Table IV. Network parameters.

Parameter	Description
T_{physical}	Transmission time of physical layer
L	The size of payload (bytes)
$\text{Rate}_{\alpha\beta}$	The data rate between nodes α and β (Mbps)
T_n^{DATA}	Transmission delay of a data frame using n extra relay nodes
$T_{\hat{n}}^{\text{DATA}}$	Transmission delay of a data frame using $n-1$ extra relay nodes and 1 direct relay node
T_n	Transmission time using n extra relay nodes
$T_{\hat{n}}$	Transmission time using $n-1$ extra relay nodes and 1 direct relay node

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

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

$$T_0^{\text{DATA}} = T_{\text{physical}} + \frac{MAC + 8L}{\text{Rate}_{SD}} \quad (8)$$

$$T_0 = DIFS + 3SIFS + RIF + CIF + T_0^{\text{DATA}} + ACK \quad (9)$$

- 5 Given a relay node R between the direct transmission beam of node S and the direct reception
 7 beam of D . Equation (10) shows the actual data rate if node S transmits a frame to node D via
 7 node R .

$$\frac{1}{\text{Rate}_{SR}} + \frac{1}{\text{Rate}_{RD}} = \frac{\text{Rate}_{SR} \times \text{Rate}_{RD}}{\text{Rate}_{SR} + \text{Rate}_{RD}} \quad (10)$$

Using Equation (10), the transmission delay of the direct relay is

$$\begin{aligned} T_{\hat{1}}^{\text{DATA}} &= T_{\text{physical}} + \frac{8L}{\text{Rate}_{SR}} + SIFS + T_{\text{physical}} + \frac{8L}{\text{Rate}_{RD}} \\ &= 2T_{\text{physical}} + SIFS + 8L \times \frac{\text{Rate}_{SR} \times \text{Rate}_{RD}}{\text{Rate}_{SR} + \text{Rate}_{RD}} \end{aligned} \quad (11)$$

- 9 Equation (12) is the transmission delay of the direct relay. The transmission delay of the direct
 11 relay includes the time of control frames and $T_{\hat{1}}^{\text{DATA}}$.

$$T_{\hat{1}} = DIFS + 5SIFS + RIF + CIF + PTS + T_{\hat{1}}^{\text{DATA}} + ACK + RESET \quad (12)$$

Given a relay node R_1 that 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 (18), 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 (19) represents the total transmission latency of direct relay + one extra relay. By using Equation (10), actual data rates of the direct relay (node R_2) and the extra relay (node R_1) are shown in Equations (13) and (14), respectively.

$$\frac{1}{\text{Rate}_{SR_2}} + \frac{1}{\text{Rate}_{R_2D}} = \frac{\text{Rate}_{SR_2} \times \text{Rate}_{R_2D}}{\text{Rate}_{SR_2} + \text{Rate}_{R_2D}} = x \quad (13)$$

$$\frac{1}{\text{Rate}_{SR_1}} + \frac{1}{\text{Rate}_{R_1D}} = \frac{\text{Rate}_{SR_1} \times \text{Rate}_{R_1D}}{\text{Rate}_{SR_1} + \text{Rate}_{R_1D}} = y \quad (14)$$

1

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 equation shows how to divide a frame according to different data rates of the two data channels;

$$\frac{L_1}{x} = \frac{L_2}{y} \quad (15)$$

$$L_1 = \frac{x}{x+y} \quad (16)$$

$$L_2 = \frac{y}{x+y} \quad (17)$$

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

$$T_2^{\text{DATA}} = \max\left\{2T_{\text{physical}} + \text{SIFS} + 8L \times L_1 \times \frac{\text{Rate}_{SR_1} \times \text{Rate}_{R_1D}}{\text{Rate}_{SR_1} + \text{Rate}_{R_1D}}, \right. \\ \left. 2T_{\text{physical}} + \text{SIFS} + (8L - 8L \times L_1) \times \frac{\text{Rate}_{SR_2} \times \text{Rate}_{R_2D}}{\text{Rate}_{SR_2} + \text{Rate}_{R_2D}}\right\} \quad (18)$$

$$T_2 = \text{DIFS} + 5\text{SIFS} + \text{RIF} + \text{CIF} + \text{PTS} + T_2^{\text{DATA}} + \text{ACK} + \text{RESET} \quad (19)$$

3

Figure 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 (10), the actual data rate of the extra relay (node R_3) is shown as follows:

5

$$\frac{1}{\text{Rate}_{SR_2}} + \frac{1}{\text{Rate}_{R_3D}} = \frac{\text{Rate}_{SR_3} \times \text{Rate}_{R_3D}}{\text{Rate}_{SR_3} + \text{Rate}_{R_3D}} = z \quad (20)$$

7

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

$$\frac{L_1}{x} = \frac{L_2}{y} = \frac{L_3}{z} \quad (21)$$

$$L_1 = \frac{x}{x+y+z} \quad (22)$$

$$L_2 = \frac{y}{x+y+z} \quad (23)$$

$$L_3 = \frac{z}{x+y+z} \quad (24)$$

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.

$$T_3^{\text{DATA}} = \max \left\{ \begin{aligned} &2T_{\text{physical}} + \text{SIFS} + 8L \times L_1 \times \frac{\text{Rate}_{SR_1} \times \text{Rate}_{R_1D}}{\text{Rate}_{SR_1} + \text{Rate}_{R_1D}}, \\ &2T_{\text{physical}} + \text{SIFS} + 8L \times L_2 \times \frac{\text{Rate}_{SR_2} \times \text{Rate}_{R_2D}}{\text{Rate}_{SR_2} + \text{Rate}_{R_2D}}, \\ &2T_{\text{physical}} + \text{SIFS} + (8L - 8L \times L_1 - 8L \times L_2) \times \frac{\text{Rate}_{SR_3} \times \text{Rate}_{R_3D}}{\text{Rate}_{SR_3} + \text{Rate}_{R_3D}} \end{aligned} \right\} \quad (25)$$

$$T_3 = \text{DIFS} + 5\text{SIFS} + \text{RIF} + \text{CIF} + \text{PTS} + T_3^{\text{DATA}} + \text{ACK} + \text{RESET} \quad (26)$$

1

Equations (9), (12), (19), and (26) 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 .

3

5

5.2. Simulation results

7

The simulation is based on NCTUns 3.0 [23] 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 V and the performance metrics are shown as follows.

9

11

- *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.

13

15

17

19

(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

Table V. Simulation parameters.

Parameter	Value
1 Mbps range	250 m
2 Mbps range	186.75 m
5.5 Mbps range	167.75 m
11 Mbps range	120.5 m
Carrier sensing range	550 m
DIFS	50 μ s
SIFS	10 μ s
Total beams	8
Data size	1400 bytes
Packet buffer size	50
Packet queuing lifetime	50 packet duration

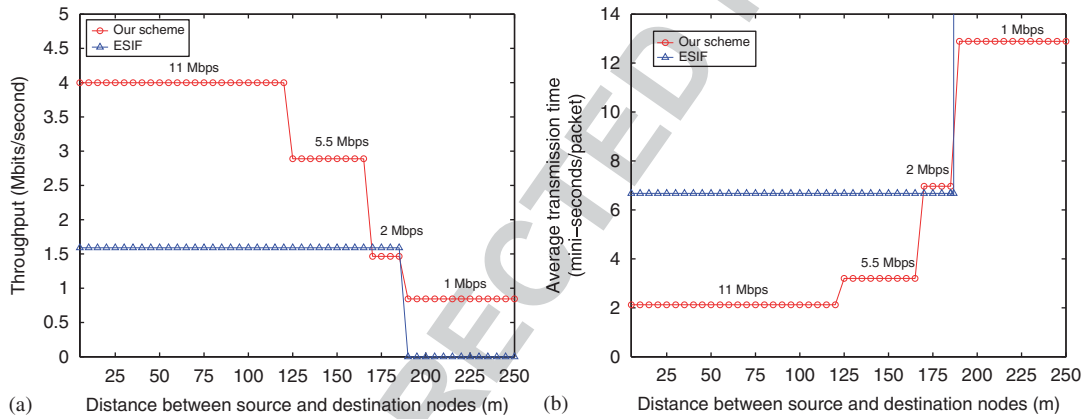


Figure 12. Performances vs distances between source and destination nodes: (a) throughput and (b) average transmission time.

1 assumes that each node uses multiple beam antennas with multirate capability (1, 2, 5.5, and
 2 11 Mbps). Figure 12 shows the difference between the multiple relay-based MAC protocol and
 3 ESIF. The scenario only has a source node and a destination node. Distances between the source
 4 and destination nodes are changed from 5 to 250 m. In Figure 12(a), when the distances between
 5 the source and destination nodes are 250–186.65, 186.75–167.76, 167.75–102.5, and 120.5–5 m,
 6 the throughput of our scheme are 0.84, 1.46, 2.88, and 3.99 Mbps, respectively. In Figure 12(b),
 7 when the distances between the source and destination nodes are 250–186.65, 186.75–167.76,
 8 167.75–102.5, and 120.5–5 m, the packet transmission latencies of our scheme are 12.88, 6.96,
 9 3.19, and 2.12 ms, respectively. When the distances between the source and destination nodes
 10 are 186.75–5 m, the throughput and average packet transmission time of ESIF are 1.59 Mbps and
 11 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.

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

9 (2) *Various relay conditions*: Figures 13 and 14 show the throughput, average transmission
10 latency, and packet dropped rate by simulating four relay conditions. The four relay conditions
11 are 'no relay', 'direct relay', 'direct relay + one extra relay', and 'direct relay + two extra relay'.
12 In Figure 13, the distance between the source and destination nodes is 250 m and the data rate is
13 1 Mbps. In Figure 14, the distance between the source and destination nodes is 186 m and the data
14 rate is 2 Mbps. Three relay nodes are placed in the relay region between the source and destination
15 nodes. And the three relay nodes are located at different beams of the source and destination nodes
16 to establish three data channels for relay. These conditions are run for 30 s to obtain the results.

17 Figure 13(a) shows the throughput of the four relay conditions of 1 Mbps under various packet
18 arrival rates. Figures 13(b) and (c) illustrate the average transmission latency of the four relay
19 conditions of 1 Mbps under various packet arrival rates. Figure 13(d) displays the packet dropped
20 rate of the four relay conditions of 1 Mbps under various packet arrival rates. The maximum
21 throughput of 'no relay' still does not exceed 0.84 Mbps even if the packet arrival rate is greater than
22 140 packets/s. This is because the transmission latency of 'no relay' is 12.89 ms. The transmission
23 time is greater than the packet arrival interval. The packet buffer accumulates more and more
24 packets when the packet arrival rate becomes large. Finally, a packet is dropped if a packet stays in
25 the buffer too long. From Figure 13(c), we can observe that the average packet transmission time
26 of 'direct relay', 'direct relay + one extra relay', and 'direct relay + two extra relay' is 6.04, 3.9,
27 and 3.2 ms, respectively. Hence, using more relay nodes to transmit can enhance the throughput
28 evidently since the transmission time is shortened.

29 Figure 14(a) demonstrates the throughput of the four relay conditions of 2 Mbps under various
30 packet arrival rates. Figure 14(a) shows that the maximum throughput of 'no relay', 'direct
31 relay', 'direct relay + one extra relay', and 'direct relay + two extra relay' are 1.46, 2.5,
32 3.27, and 3.65 Mbps, respectively. Figures 14(b) and (c) show the average packet transmis-
33 sion latency of the four relay conditions of 2 Mbps under various packet arrival rates. From
34 Figure 14(c), the average packet transmission latencies of 'no relay', 'direct relay', 'direct relay
35 + one extra relay', and 'direct relay + two extra relay' are 6.96, 3.86, 2.8, and 2.43 ms,
36 respectively. Figure 14(d) illustrates the packet dropped rate of the four relay conditions of
37 2 Mbps under various packet arrival rates. The shorter transmission time makes smaller packet
38 dropped rate.

39 (3) *Concurrently multiple communications*: Our MAC protocol is based on ESIF. The feature of
40 ESIF is concurrent multiple communications. Figures 15 and 16 illustrate the performances of the
41 many-to-one communication and cross concurrent communications, respectively. The throughput
42 and transmission time are derived every 0.5 s so as to calculate the averages. The two scenarios
43 are run for 15 s. Node R_1 is awakened in the fifth second and Node R_2 is awakened in the tenth
44 second. In Figure 15(a), Both nodes S_1 and S_2 communicate with node D . In ESIF, the data rate
45 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

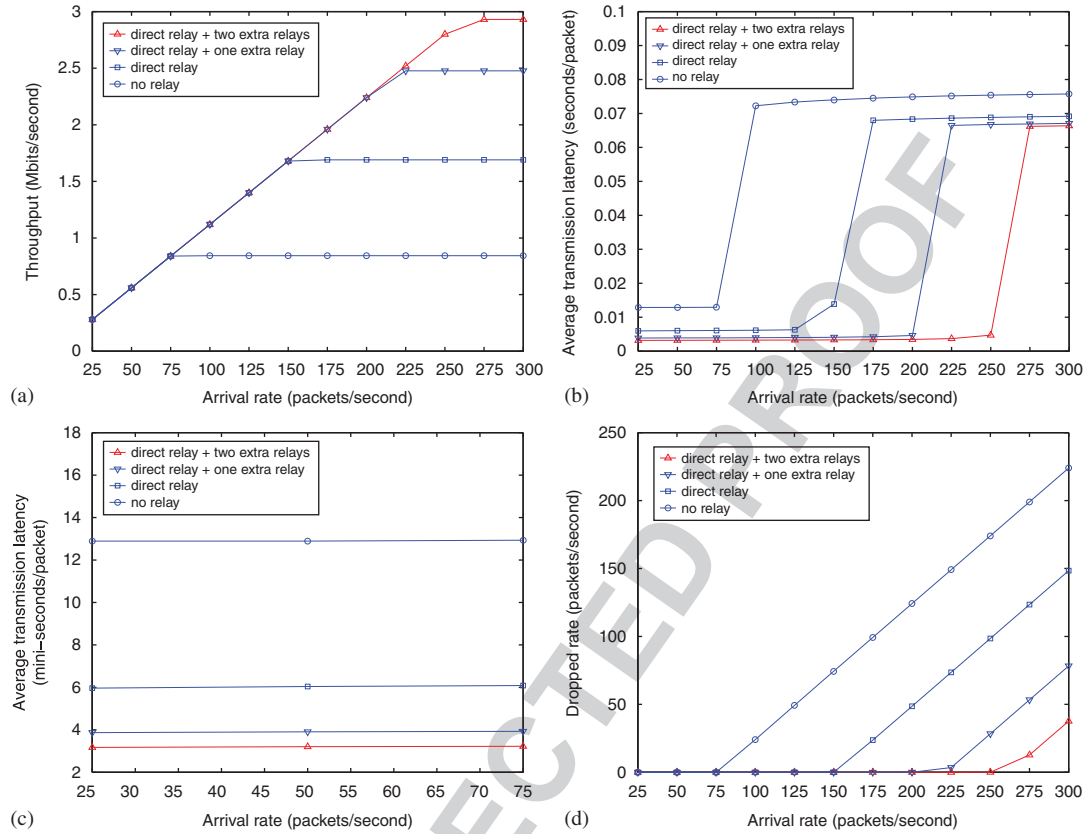


Figure 13. Four relay conditions of 1 Mbps: (a) throughput; (b) average transmission latency; (c) average transmission latency before buffer overflow; and (d) packet dropped rate.

1 node S_1 when nodes S_1 and S_2 communicate with node D concurrently. Hence, our scheme is
 2 rarely worse than ESIF after the fifth second. In the fifth seconds, node R_1 awakens and helps node
 3 S_1 . The transmission time between nodes S_1 and D is shortened. In the tenth seconds, node R_2
 4 awakens and helps node S_1 . The transmission time between nodes S_1 and D is further shortened.
 5 Hence, the throughput and the transmission time are improved by nodes R_1 and R_2 as shown in
 6 Figures 15(b) and (c).

7 In Figure 16(a), nodes S_1 and S_2 send frames to nodes D_1 and D_2 via node C , respectively.
 8 Node C can relay the frames of nodes S_1 and S_2 to nodes D_1 and D_2 simultaneously when nodes
 9 S_1 and S_2 transmit frames via node C simultaneously. The transmission time is constrained by
 10 node S_1 when nodes S_1 and S_2 transmit simultaneously. The transmission time is constrained by
 11 node D_2 when node C transmits to nodes D_1 and D_2 simultaneously. Figures 16(b) and (c) show
 12 the throughput and the transmission time are improved after node R_1 and R_2 are awakened. The
 13 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 .

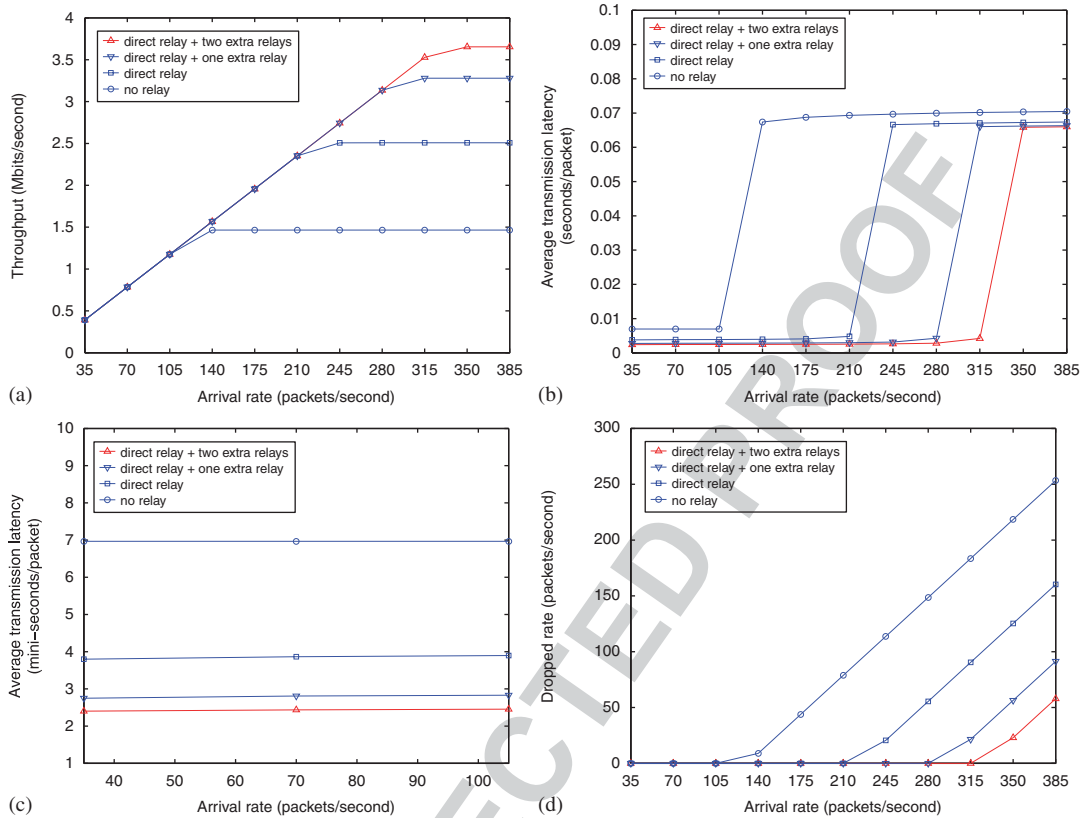


Figure 14. Four relay conditions of 2 Mbps: (a) throughput; (b) average transmission latency; (c) average transmission latency before buffer overflow; and (d) packet dropped rate.

1 (4) *Random topology*: Figure 17 shows the one-to-many communication under a fully connected
 2 topology. Ten nodes are put randomly in a 180m × 180m area. One of the ten nodes acts as the sender and two of the remaining nine nodes act as receivers. The sender communicates
 3 with the two receivers concurrently. With the raised packet arrive rate, the throughput of ESIF
 4 only stays at 3Mbps since the data rate is always 2Mbps. However, the throughput of our
 5 scheme can continue to be enhanced when the packet arrival rate becomes more and more
 6 higher because the transmission latencies of the two links are shortened by relay nodes. In
 7 Figure 18, the multihop topology is adopted. Thirty nodes are put randomly in a 1000m ×
 8 500m area. The shortest path routing is adopted. We simulate one traffic flow in the area.
 9 Compared with ESIF, our scheme can still achieve higher throughput and lower transmission
 10 latency.
 11

12 (5) *Fragment Burst with Adjusting Relay Nodes*: Figure 19 shows the wireless medium utilization
 13 of the adjusting relay nodes algorithm. The throughput and transmission time are derived from the
 14 average of the results of every 0.25 s. In Figure 19(a), S_1-D and S_2-D have a heavy traffic and
 15 S_3-R_3 has a light traffic. In order to alleviate the heavy traffic and transmit faster, nodes S_1 and S_2

Color Online, B&W in Print

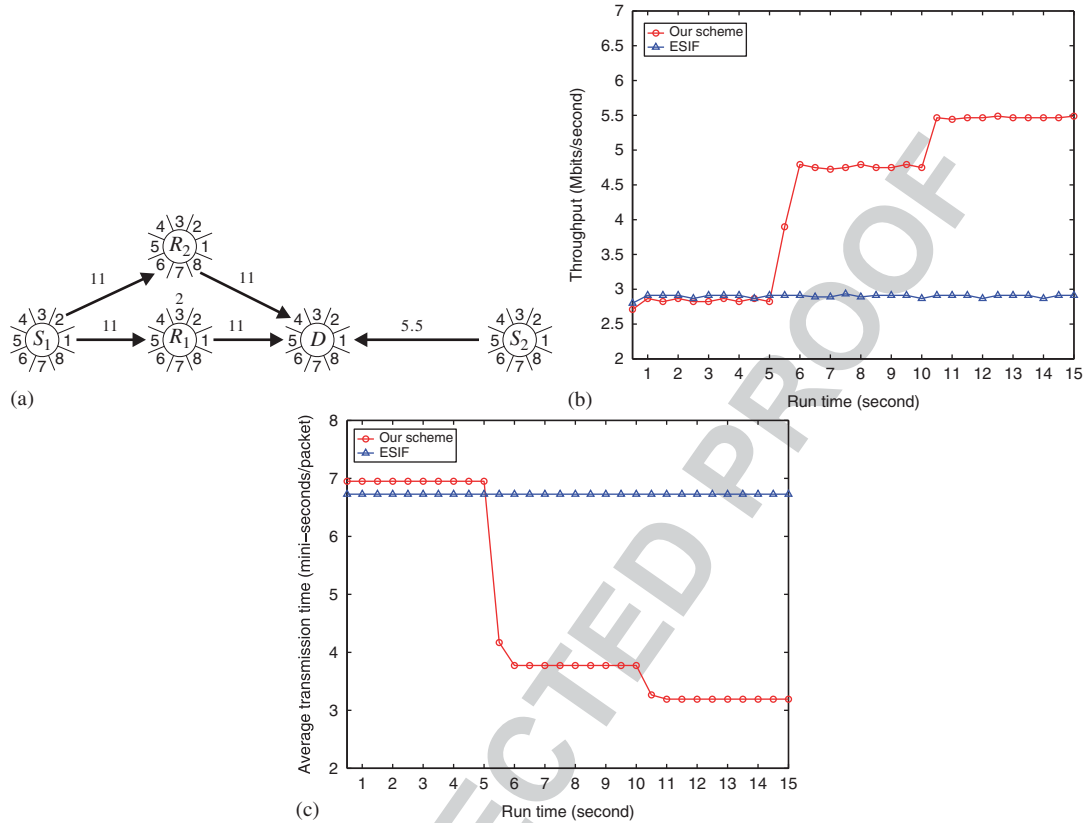


Figure 15. Many-to-one communication with multiple relay: (a) topology; (b) throughput; and (c) average transmission time.

1 often perform the fragment burst and multiple relay mechanism. Nodes R_1 and R_2 are relay nodes
 2 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
 3 destination of node S_3 . For this reason, node S_3 cannot communicate with node R_3 when node R_3
 4 acts as the relay node of S_2 . In this scenario, node R_3 is often requested to be the relay node of
 5 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
 6 for a long time if S_2 performs the fragment burst and multiple relay mechanism. However, nodes
 7 R_3 and R_5 are unnecessary relay nodes when nodes S_1 and S_2 exploit multiple relay mechanism
 8 to transmit simultaneously.
 9 Nodes S_1 and S_2 do not know each other's transmission delay. Accordingly, nodes R_3 and R_5 are
 10 hard to shorten the total transmission delay during the concurrent fragment burst period of nodes
 11 S_1 and S_2 . The wireless medium utilization is low if node R_3 acts as a unnecessary relay node
 12 for a long time. This is because that node S_3 has no opportunity to communicate with node R_3 .
 13 However, node D can response ACKs with relay information to the source nodes after receiving
 14 all (sub)frames. Node R_3 can be released the role of the unnecessary relay node by using the
 15 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

Color Online, B&W in Print

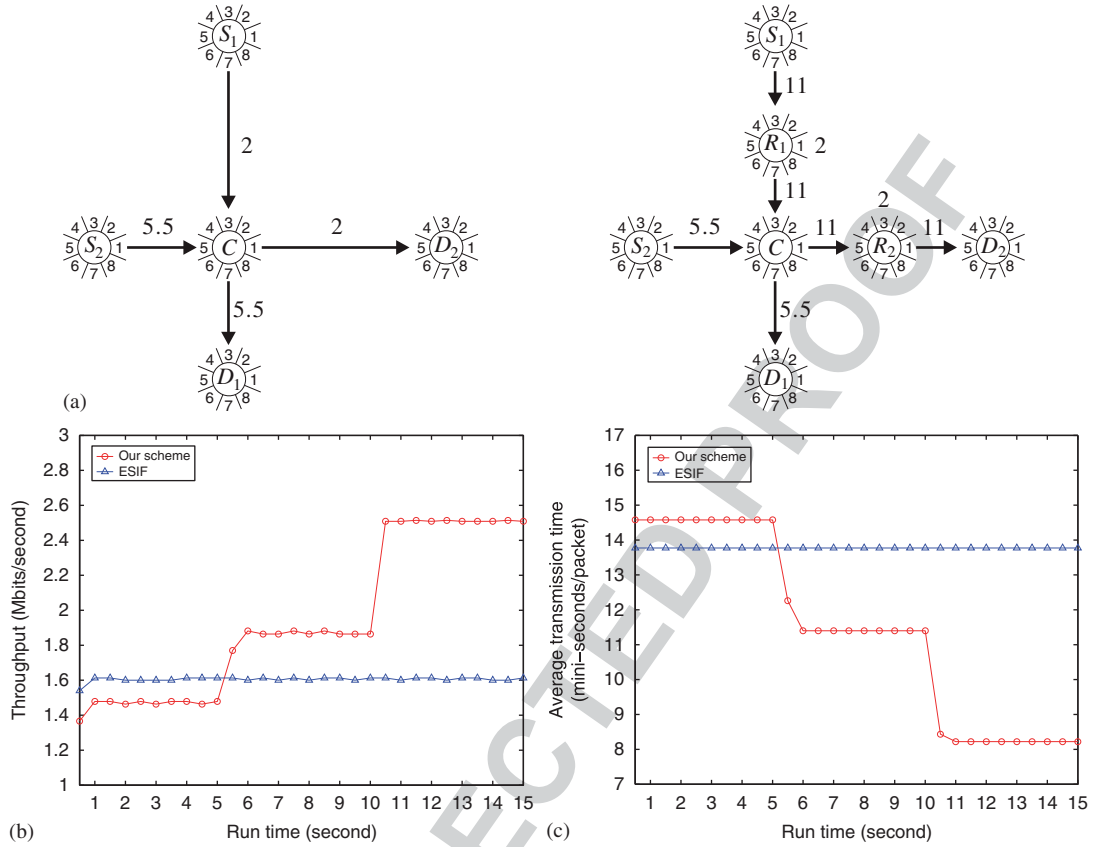


Figure 16. Cross concurrent communications with multiple relay: (a) topology; (b) throughput; and (c) average transmission time.

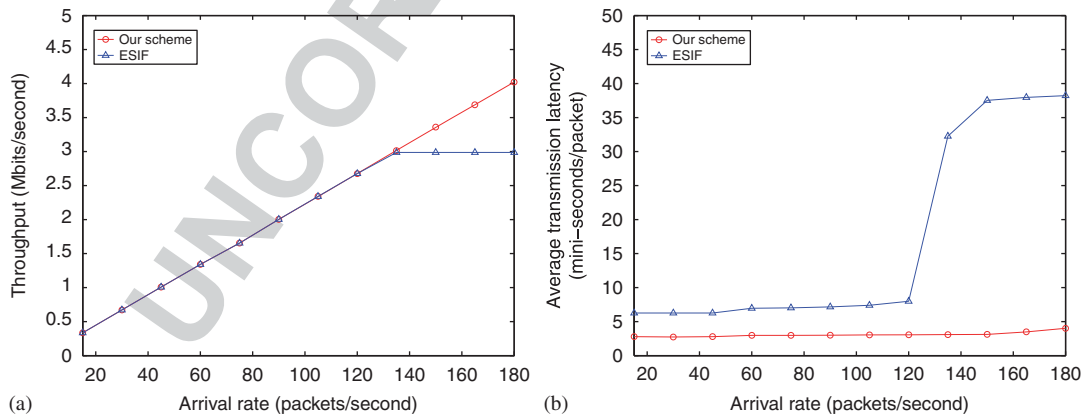


Figure 17. Performances comparison of one-to-many communication under one hop: (a) throughput and (b) average transmission latency.

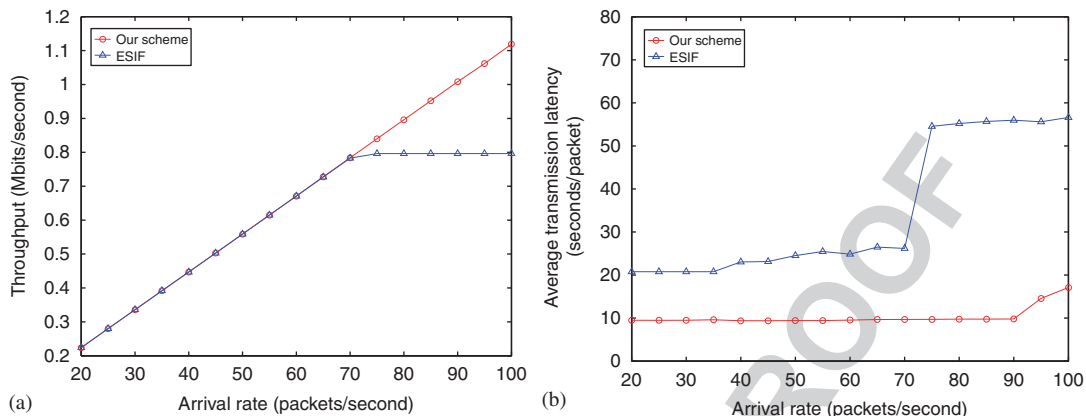


Figure 18. Performances comparison of multihop: (a) throughput and (b) average transmission latency.

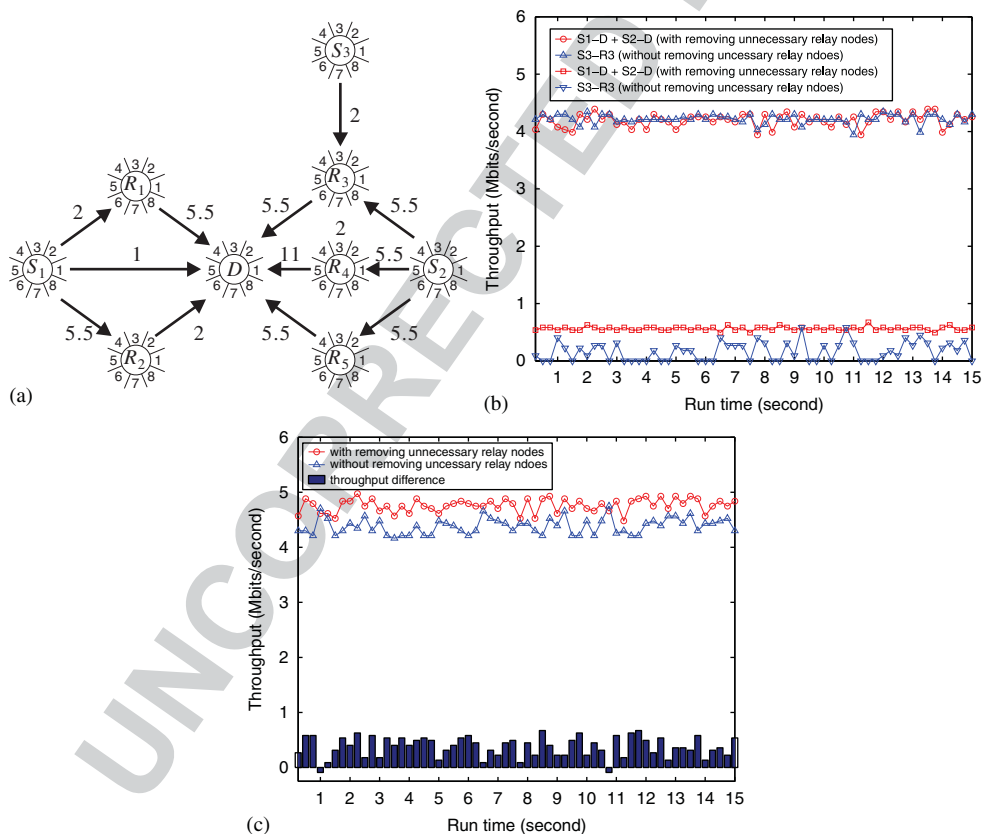


Figure 19. The wireless medium utilization of adjusting relay nodes algorithm: (a) throughput of $S_1 - D + S_2 - D$ and $S_3 - R_3$ and (b) total throughput of $S_1 - D + S_2 - D$ and $S_3 - R_3$.

1 DATA/ACK) every three transmissions. The packet arrival rate of S_3-R_3 is 50 packets/s. Figure
 2 19(b) shows the performance with and without the removing algorithm. Without the removing
 3 algorithm, node R_3 is often to be a unnecessary relay nodes when nodes S_1 and S_2 perform the
 4 multiple relay mechanism and fragment burst simultaneously. Node S_3 is difficult to communicate
 5 with node R_3 , and hence the throughput of S_3-R_3 is very low. With the removing algorithm,
 6 node R_3 can be released the role of the unnecessary relay node. Node S_3 has the opportunity to
 7 communicate with node R_3 even when nodes S_1 and S_2 perform the multiple relay mechanism and
 8 fragment burst simultaneously. Therefore, the throughput of S_3-R_3 becomes higher. Figure 19(c)
 9 shows that the total throughput and wireless medium utilization are better by using the removing
 unnecessary relay nodes algorithm.

11

6. CONCLUSION

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

25

ACKNOWLEDGEMENTS

This work is supported by the National Science Council of the Republic of China under Grants #NSC-95-2221-E-305-008 and #NSC-96-2213-E-194-007.

REFERENCES

- 27 1. Chen Y-S, Ko Y-W. A lantern-tree-based QoS on-demand multicast protocol for a wireless mobile ad hoc network. *IEICE Transactions on Communications* 2004; 717–726.
- 29 2. Chen Y-S, Kao T-C, Sheu J-P. A mobile learning system for scaffolding bird watching learning. *Journal of Computer Assisted Learning* 2003; 19(3):347–359.
- 31 3. Group IW. Part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications, *ANSI/IEEE Std. 802.11*, 1999.
- 33 4. Yang D, Shao W, Shen D, Li VO. Hybrid busy tone and carrier sense-based MAC protocol in multi-hop ad hoc networks. *International Journal of Ad Hoc and Ubiquitous Computing* 2007; 2(4):208–214.
- 35 5. Choudhury RR, Yang X, Ramanathan R, Vaidya NH. On designing MAC protocols for wireless networks using directional antennas. *IEEE Transaction on Mobile Computing* 2006; 477–491.
- 37 6. Huang Z, Shen C-C. Multibeam antenna-based topology control with directional power intensity for ad hoc networks. *IEEE Transaction on Mobile Computing* 2006; 508–517.

Q2

- 1 7. Huang Z, Shen C-C. A comparison study of omnidirectional and directional MAC protocols for ad hoc networks. *IEEE Global Telecommunications Conference*, 2002; 57–61.
- 3 8. Yi S, Pei Y, Kalyanaraman S. On the capacity improvement of ad hoc wireless networks using directional antennas. *ACM International Symposium on Mobile Ad Hoc Networking and Computing*, 2003; 108–116.
- 5 9. Huang Z, Shen C-C, Srisathapornphat C, Jaikaeo C. A busy-tone based directional MAC protocol for ad hoc networks. *IEEE Military Communications Conference*, 2002; 1233–1238.
- 7 10. Ko Y-B, Shankarkumar V, Vaidya NH. Medium access control protocols using directional antennas in ad hoc networks. *IEEE International Conference on Computer Communication*, 2000; 13–21.
- 9 11. Korakis T, Jakllari G, Tassioulas L. A MAC protocol for full exploitation of directional antennas in ad hoc wireless networks. *ACM International Symposium on Mobile Ad Hoc Networking and Computing*, 2003; 98–107.
- 11 12. Jain V, Gupta A, Lal D, Agrawal DP. A cross layer MAC with explicit synchronization through intelligent feedback for multiple beam antennas. *IEEE Global Telecommunications Conference* 2005; 3196–3200.
- 13 13. Jain V, Gupta A, Lal D, Agrawal DP. IEEE 802.11 DCF based MAC protocols for multiple beam antennas and their limitations. *IEEE Mobile Adhoc and Sensor Systems Conference (MASS)*, 2005.
- 15 14. Lal D, Agrawal DP. A multiple-beam antenna protocol at a wireless access point for exploiting spatial parallelism. *IEEE/Sarnoff Symposium on Advances in Wired and Wireless Communication*, 2004; 23–26.
- 17 15. Lal D, Toshniwal R, Radhakrishnan R, Agrawal DP, Caffrey J. A novel MAC layer protocol for space division multiple access in wireless ad hoc networks. *IEEE International Conference on Computer Communications and Networks*, 2002; 614–619.
- 19 16. Heusse M, Rousseau F, Berger-Sabbatel G, Duda A. Performance anomaly of 802.11b. *IEEE International Conference on Computer Communication*, 2003; 836–843.
- 21 17. Tan K, Wan Z, Zhu H, Andrian J. CODE: cooperative medium access for multirate wireless ad hoc network. *IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks*, 2007.
- 23 18. Zhu H, Cao G. rDCF: a relay-enabled medium access control protocol for wireless ad hoc networks. *IEEE Transaction on Mobile Computing* 2006; 1201–1214.
- 25 19. Zou S, Li B, Wu H, Zhang Q, Zhu W, Cheng S. A relay-aided media access (RAMA) protocol in multirate wireless networks. *IEEE Transactions on Vehicular Technology* 2006; 1657–1667.
- 27 20. Holland G, Vaidya N, Bahl P. A rate-adaptive MAC protocol for multi-hop wireless networks. *ACM International Conference on Mobile Computing and Networking*, 2001; 236–251.
- 29 21. Kamerman A, Monteban L. WLAN-II: A high-performance wireless LAN for the unlicensed band. *Bell Labs Technical Journals* 1997; 253–270.
- 31 22. Sadeghi B, Kanodia V, Sabharwal A, Knightly E. Opportunistic media access for multirate ad hoc networks. *ACM International Conference on Mobile Computing and Networking*, 2002; 24–35.
- 33 23. NCTUns Network Simulator and Emulator 3.0. Available from: <http://nsl.csie.nctu.edu.tw/nctuns.html>.

Q3

Q4

Q5