

Throughput enhancement by exploiting spatial reuse opportunities with smart antenna systems in wireless ad hoc networks



Chao-Tsun Chang^{a,*}, Chih-Yung Chang^b, Tzu-Lin Wang^b, Yun-Jung Lu^b

^a Department of Information Management, Hsiuping University of Science & Technology, Taichung, Dali, Taiwan

^b Department of Computer Science and Information Engineering, Tamkang University, Taipei, Tamsui, Taiwan

ARTICLE INFO

Article history:

Received 9 April 2012

Received in revised form 26 January 2013

Accepted 3 March 2013

Available online 6 June 2013

Keywords:

Smart Antennas

Scheduling

Protocol

Ad Hoc Networks

ABSTRACT

Smart antenna (or multi-beam antenna) systems can support simultaneous transmissions (or receptions) of multiple packets in different beams using the same channel. However, network performance is highly dependent on transmission scheduling. This study develops two transmission-scheduling schemes for use with smart antenna systems to improve network throughput and reduce transmission delay. The proposed scheduling mechanisms, which are designed to minimize the average latency and maximize network throughput, exploit the opportunities for parallel transmissions and consider communication restrictions and packet sizes. The performance results indicate that the proposed schemes outperform previous schemes in network throughput and transmission delay.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Wireless ad hoc networks are characterized by low cost and high transmission rates without requiring the use of any infrastructural support. The greatest challenge for improving the network throughput of ad hoc networks involves exploiting spatial reuse opportunities at which multiple packets can be exchanged simultaneously in an area without colliding with each other. Recent advances in antenna technology have increased spatial reuse opportunities.

Omnidirectional antennas transmit or receive signals equally well in all directions. Transmissions using omnidirectional antennas might easily create spatial bottlenecks that severely limit the overall network throughput. Unidirectional antennas [1–3,11,15] focus radio-frequency (RF)

energy in a particular direction and exploit stronger beam-forming, thereby saving energy and resulting in a smaller interference area compared to that of omnidirectional antenna systems. Although unidirectional antennas were originally applied for exploiting spatial reuse, they permit only one transmission or reception by a node at a given time. Recently, multiple-beam smart antennas [6–8,10,12–14] used in ad hoc networks have received attention because of their potential to improve network throughput. Through the application of complex digital signal processing techniques, multiple-beam smart antennas can support the simultaneous transmissions (or receptions) of multiple packets in different beams by using the same channel. This is the case with switched-beam smart antenna comprised of multiple beam antenna array (SB-MBA) [7,13]. Numerous MAC scheduling approaches [6,7,14–16] have been proposed for increasing the opportunities for simultaneous transmissions (or receptions) and improving network throughput through smart antennas. However, enhanced scheduling to increase the parallel degree of transmissions with smart antenna systems is required.

* Corresponding author. Address: Department of Information Management, Hsiuping University of Science & Technology, 11 Gungye Road, Dali City, Taichung, Taiwan. Tel.: +886 4 24961123x3110; fax: +886 4 24961118.

E-mail address: cctas@mail.hust.edu.tw (C.-T. Chang).

Several studies [6,7] have proposed scheduling protocols with smart antennas to improve the spatial reuse opportunities for unidirectional antennas. Although smart antenna systems can support multiple-beam formations in multiple directions, they cannot be used for data transmissions and receptions simultaneously. This constraint is referred to as the *Tx/Rx constraint* in this study. A host with smart antennas has the *data diversity constraint*, which restricts its communication with only one neighboring host in each beam; that is, a host equipped with k -beam smart antennas may communicate with a maximum of k different hosts during any given period. How to protect the packet transmission from interference for a given set of communication requests and minimize the transmission latency by considering the *Tx/Rx* and *data diversity* constraints are the key challenges for maximizing network throughput.

This study develops transmission scheduling schemes with multiple-beam smart antennas to improve network throughput and reduce delay time. The proposed mechanisms use clustering, and they centralize the intra-cluster scheduling but schedule the inter-clustering by using a distributed approach. In intra-cluster scheduling, each cluster header collects all transmission demands from its members, schedules parallel transmission pairs, and broadcasts the scheduling results to the cluster members. The proposed intra-cluster scheduling schemes exploit the opportunities for parallel transmissions and consider communication restrictions and packet sizes, to minimize the average latency and maximize the network throughput. Furthermore, the proposed schemes consider the transmission time of packets and regulate the orders of packet transmissions to minimize their packet delays.

The remainder of this study is organized as follows: Section 2 introduces related studies. Section 3 specifies the network environment and problem formulation. The proposed transmission scheduling algorithms for intra-cluster scheduling are presented in Sections 4 and 5 shows the proposed algorithms for scheduling inter-cluster transmissions. Section 6 provides a performance evaluation of the proposed schemes in contrast to those of existing studies, and finally, Section 7 offers a conclusion.

2. Related studies

Exploring spatial reuse opportunities to increase network capacity has become one of the most critical challenges for wireless ad hoc networks in recent years. Generally, advanced antenna technologies can be classified into unidirectional and multiple-beam smart antennas. Although multiple-beam smart antennas are more expensive than unidirectional antennas, they can sense neighboring hosts in each beam and support simultaneous transmissions (or receptions) of multiple packets in different beams. However, the hidden terminal and deafness [1–5,12,13] problems can significantly degrade network throughput. The following paragraphs present a survey of related studies, which involve unidirectional or multiple-beam smart antennas in MAC protocol designs.

Several MAC protocols that were developed for unidirectional antennas adopted directional control packets to prevent control packet collisions for improving network throughput. The D-MAC [1] scheme uses control packets, such as ORTS/DRTS and OCTS, and the DVCS scheme [2] uses ORTS or DRTS packets based on the receiver's location. Related studies [3,4] have introduced the circular-based DMAC (CRD-MAC) protocol to reduce the effects of the hidden-terminal and deafness problems and exploit space usage by using ORTS and DCTS packets. However, when the receiver loses the ORTS packet, its neighboring nodes suffer from the deafness problem. In [5], the MAC protocol for directional antennas (MDA) was designed to manage the hidden terminal and deafness problems by constructing a Deafness Table (DT) and using DRTS/DCTS messages. It also considers the optimization of the circular RTS transmissions of control frames. However, the circular RTS transmissions create a significant overhead to neighboring nodes of senders and receivers when the intent of these nodes is not to communicate with the sender or receiver. The multiHop RTS MAC (MMAC) protocol [16] adopts the DRTS/DCTS exchange to exploit the benefits of beamforming by identifying shorter routes. However, the performance of the protocol cannot be guaranteed because the benefits are highly dependent on the topology and flow patterns in the networks. The CW-DMAC protocol proposed in [12] uses the ORTS and OCTS packets to reduce the effects of the deafness and directional hidden terminal difficulties for unidirectional antenna-based wireless multi-hop networks. However, using ORTS and OCTS packets reduces spatial reuse. Bianchi et al. [14] proposed a space division – time division MAC protocol to manage unidirectional antennas for increasing the capacity. Although this protocol guarantees interference-free transmission of control messages, the support for spatial and time frequency divisions in the ad hoc network is complex.

Multiple-beam smart antennas have been commonly applied in ad hoc networks [12,15–18] for exploring spatial reuse opportunities to improve network capacity. The multiple-beam smart antenna is referred to as a multiple-beam adaptive array (MBAA) in [6]. These smart antennas can support simultaneous transmissions (or receptions) of multiple packets in different beams by using the same channel through the application of complex digital signal processing techniques.

Asynchronous MAC protocols based on IEEE 802.11 DCF were designed in several studies [6,7] to enable multi-packet transmissions (or receptions) for multiple-beam antennas. The hash-based scheduling mechanism referred to as ROMA [6], extends the range of scheduled hosts to two hops for improving scheduling efficiency. The ROMA mechanism implements hash functions to randomly separate nodes into senders or receivers per time slot and designates the priority of incoming links. The avoidance of hidden terminals is considered in the scheduling process. Although ROMA might arrange one-hop neighbors to increase opportunities for simultaneous transmissions or receptions, it also increases the control overheads by collecting transmission requirements from two-hop neighboring hosts. Furthermore, hash-based scheduling might arrange large and small packets for transmission simulta-

neously; thus, the beams that complete their transmissions might be idle while waiting for other busy beams.

The authors of [7] proposed the ESIF protocol, a MAC scheduling approach based on routing information. The ESIF protocol changes IEEE 802.11 RTS/CTS messages to RIF (RTS with intelligent feedback) and CIF (CTS with intelligent feedback) to piggyback feedback on neighboring nodes. The control messages used in ESIF are comprised of frame control, duration, receiver address, transmitter address, priority, N , and FCS fields. The duration field represents the estimated duration of communication, and the N field denotes the number of potential transmitters in the beam. When the sender and receiver exchange an RIF/CIF message in certain beams, they simultaneously send SCH messages to their neighbors in other beams. After receiving SCH messages, the neighbors of the sender or receiver sense which beams from the sender and receiver still possess potential for transmission, and update their directional network allocation vector (DNAV) accordingly. The SCH message permits a neighbor to determine whether it is located in the active beam of the sender or receiver. Therefore, a node can decide whether to defer the transmissions for one-hop neighbors, thereby reducing the effects of the deafness and hidden terminal problems. Furthermore, this scheme removes the backoff policy to allow simultaneous transmissions or receptions of numerous packets through smart antennas. However, the performance of parallelism highly depends on the grouping scheme which aims to construct a set of parallel transmissions as a group and arrange the transmissions in a group to be transmitted in parallel. In ESIF, each node does not determine the schedules of other neighboring nodes by considering all transmission requirements, including its own and its neighbors' transmission requirements. Although ESIF minimizes energy and latency overheads, the potential for parallel transmissions is not exploited efficiently in transmission scheduling.

This study presents decentralized scheduling mechanisms created by modifying the ESIF MAC for exploiting the opportunities for parallel transmissions in clusters, and by considering communication restrictions and packet sizes to minimize the average latency and maximize network throughput. Table 1 shows the characteristics of the proposed scheme compared to related schemes. The proposed scheme is a novel scheduling method that improves the parallel degree of transmissions by using smart antennas without additional support.

3. Network model and problem formation

3.1. Network model

Let V denote the set of n hosts $h_1, h_2, \dots, h_n$ and E denote the set of neighboring connections between hosts, where $E \in V \times V$. The topology of the ad hoc network can be defined by $G = (V, E)$. Let N_i represent the set of single-hop neighbors of host h_i . Each host has a unique ID and is equipped with a wide azimuth switched-beam smart antenna (SB-MBA), which comprises a k -beam antenna array [7,13] to sense neighboring hosts in each beam of smart antennas. The antenna beam patterns of the SB-MBA are predetermined by shifting the signal phase of each element to cover all directions. Each station is assumed to create non-overlapping multiple beams with gain, and the impact of a side-lobe interference [6,7] or the benefits of nulling are not considered [7]. Furthermore, each node is assumed to be able to determine precisely the angle of arrival (AoA) of a received signal [2] and the channel is symmetric. SB-MBAs support simultaneous transmissions (or receptions) of multiple packets in different beams at any given time through the use of complex digital signal processing techniques. The *neighbor matrix* L , which presents the neighboring relationship among stations, is described as follows.

Definition (Neighbor Matrix (L)). A neighbor matrix L with size $n \times n \times k$ is used to represent the beams where the neighboring hosts are located. Each entry of the matrix $l_{i,j,q}$ stores a binary value to denote whether host h_j is located in the beam sector q of host h_i , where $1 \leq q \leq k$. If host h_j is located within q , the value of entry $l_{i,j,q}$ is one; otherwise, it is zero.

Fig. 2a shows the neighbor matrix of the network shown in Fig. 1a. The entries $l_{1,2,2}$, $l_{1,3,2}$, $l_{1,4,3}$, and $l_{1,5,1}$ store the value one because hosts h_2 , h_3 , h_4 , and h_5 are located in beam₂, beam₂, beam₃, and beam₁ of host h_1 , respectively, as shown in Fig. 2a.

The total transmission duration is partitioned into several equal-length intervals, with each contains T slots, meaning that all hosts complete their transmission requirements in T slots. Let c_{ij} denote a communication pair where the intent of host h_i is to send data to host h_j . Several terms for presenting the problem formulation are defined as follows.

Table 1
Characteristics of proposed scheme compared to related schemes.

Schemes	Antenna type	Objectives	CSMA/CA characteristics	Extra support
D-MAC [1]	Unidirectional	Collision avoidance	DRTS/ORTS, OCTS	No
DVCS [2]	Unidirectional	Channel capability	DRTS/ORTS, DCTS	Location information
CRD-MAC [3,4]	Unidirectional	Reduce hidden-terminal and deafness	ORTS, DCTS	No
MDA [5]	Unidirectional	Reduce hidden-terminal and deafness	DRTS, DCTS	No
MMAC [16]	Unidirectional	Exploit benefits of beamforming	DRTS, DCTS	Routing information
CW-DMAC [12]	Switch-beam	Reduce hidden-terminal and deafness	ORTS, OCTS	No
[14]	Switch-beam	Interference-free transmission	ORTS, OCTS	Freq./Time divisions
ROMA [6]	MBAA	Multi-packet communication	DRTS, DCTS	Two-hop trans. req.
ESIF [7]	SM-MBA	Parallel transmissions (beams of a node)	RIF/SCH, CIF/SCH	No
Our scheme	SM-MBA	Parallel transmissions (cluster)	RIF/SCH, CIF/SCH, CIR, SIA	No

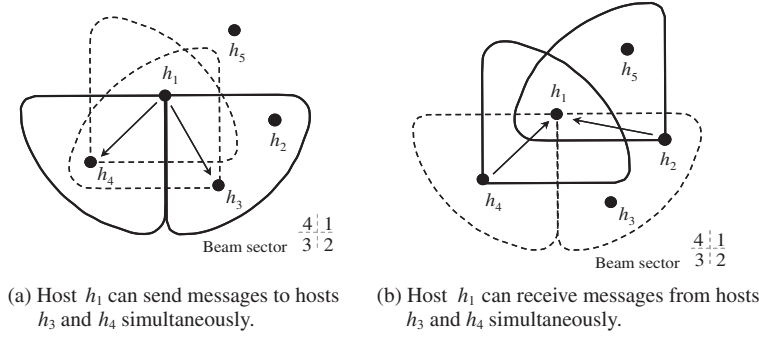


Fig. 1. Example of WLAN equipped with four-beam smart antenna and propagation models.

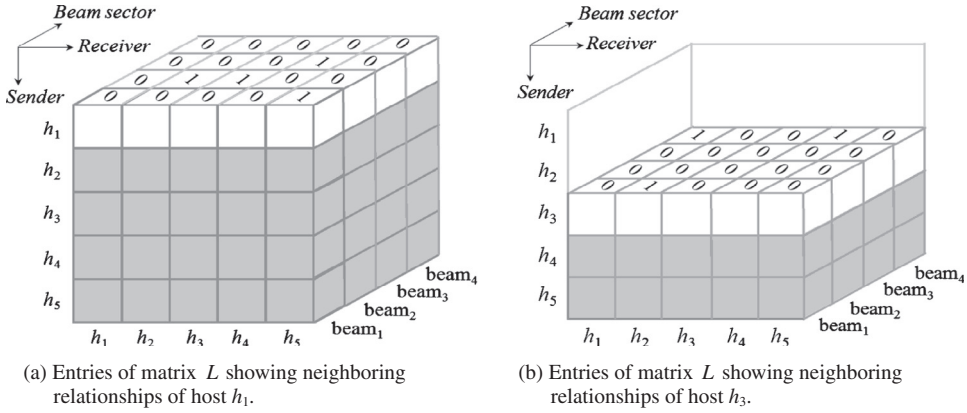


Fig. 2. Examples of neighbor matrix of hosts h_1 and h_3 from Fig. 1a.

Definition (Transmission Demand Matrix (D)). Let Transmission Demand Matrix $D = [d_{ij}]_{n \times n}$ represent the data units of transmission requirements for a given network. The entry d_{ij} records the data units of transmission demand associated with c_{ij} .

Definition (Scheduling Matrix (S)). A Scheduling Matrix $S = [s_{ij,t}]_{n \times n \times T}$ defines a transmission schedule, where

$$s_{ij,t} = \begin{cases} 1, & \text{reserved for host } h_i \text{ sends data to } h_j \\ & \text{during slot } t, \\ 0, & \text{otherwise.} \end{cases}$$

Fig. 3 shows an example of the scheduling matrix where time slot t_2 is allocated to $c_{5,1}$, $c_{5,2}$, and $c_{4,3}$. Let r and $[t_1, t_2]_{ij}$ denote the transmission rate and the transmission interval from t_1 to t_2 , allocated to c_{ij} , respectively. We have

$$|[t_1, t_2]_{ij}| = \left\lceil \frac{d_{ij}}{r} \right\rceil = |\{t | \forall t \in [t_1, t_2]_{ij} \text{ and } s_{ij,t} = 1\}|.$$

Let δ_{ij} represent the set of transmission slots from 0 to T allocated to c_{ij} , where

$$\delta_{ij} = \{t \in [0, T] | s_{ij,t} = 1\}.$$

Let $\text{delay}(i, j)$ denote the waiting time of c_{ij} . If the transmission interval $[t_1, t_2]_{ij}$ is allocated to c_{ij} , then the following equation holds.

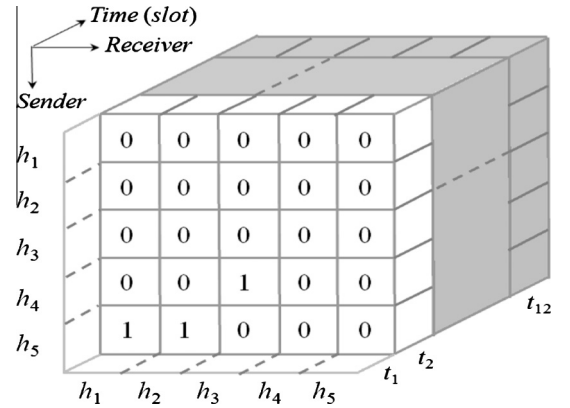


Fig. 3. Example of scheduling matrix; pairs $c_{5,1}$, $c_{5,2}$, and $c_{4,3}$ will use slot t_2 .

$$|\delta_{ij}| = |t_2 - t_1 + 1| \text{ and } \text{delay}(i, j) = t_1$$

3.2. Problem formulation

Given a network $G = (V, E)$, assume that the transmission requirement for each host is stored in matrix $D = [d_{ij}]_{n \times n}$. Two operators are defined before introducing scheduling.

Definition (Positive AND operator ($\odot$)). Given two positive integers x and y , the operator $\odot$ defines the Positive AND operation as follows:

$$x \odot y = \begin{cases} 1, & \text{if } (x > 0) \text{ and } (y > 0), \\ 0, & \text{otherwise.} \end{cases}$$

Definition (Positive XOR operator ($\otimes$)). Given two positive integers x and y , the operator $\otimes$ defines the Positive XOR operation as follows:

$$x \otimes y = \begin{cases} 1, & \text{if } (x \neq y), \\ 0, & \text{otherwise.} \end{cases}$$

Although the k -beam smart antennas system permits a host to communicate with k different hosts at any given time, the exploitation of parallel transmissions should satisfy several constraints, thus avoiding the interference and deafness problems. These constraints are introduced in the following subsections.

3.2.1. Beam mode constraint

All beams of an antenna system should be set to the same mode (either sending or receiving mode) at any given time. This hardware constraint is referred to as a *beam mode constraint*. The scheduling schemes cannot assign a host to send a message using one beam and receive a message using another beam at any given slot t . Therefore, any host h_i cannot be allowed to have both values of $s_{i,x,t}$ and $s_{y,i,t}$ to be 1s at slot t , given that h_x and h_y are neighbors of h_i . Expression (1) shows the *beam mode constraint*.

$$s_{i,x,t} \odot s_{y,i,t} = 0, \quad \forall h_x, h_y \in N_i, \quad \forall t \in T, \quad \forall h_i \in V. \quad (1)$$

3.2.2. Data diversity constraint

Each beam of the smart antenna system can send messages to a maximum of one neighboring host at any given time slot. This hardware constraint is referred to as the *data diversity constraint*. Let neighboring hosts h_x and h_y be located in beam _{q} of host h_i . Expressions (2) and (3) can be applied to ascertain whether host h_i is arranged to send/receive messages to/from its neighboring hosts h_x and h_y in beam _{q} at any given slot t . The following constraint should be satisfied.

$$(s_{i,x,t} \otimes l_{i,x,q}) \odot (s_{i,y,t} \otimes l_{i,y,q}) = 0, \quad \forall h_x, h_y \in N_i, \quad \forall t \in T, \quad 1 \leq q \leq k. \quad (2)$$

$$(s_{x,i,t} \otimes l_{i,x,q}) \odot (s_{y,i,t} \otimes l_{i,y,q}) = 0, \quad \forall h_x, h_y \in N_i, \quad \forall t \in T, \quad 1 \leq q \leq k. \quad (3)$$

3.2.3. Interference constraint

Each beam of the smart antenna system can receive messages from at most one neighboring host at any given slot. This hardware constraint is referred to as the *interference constraint*. Let neighboring hosts h_i and h_x be located in beam _{q} of host h_j . Expression (4) verifies whether $c_{x,y}$ interferes with host h_j at any given slot t when it is arranged to receive messages from host h_i .

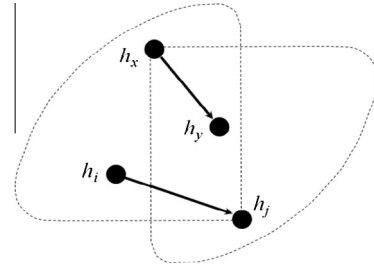


Fig. 4. Example illustrating the interference constraint for c_{ij} and $c_{x,y}$ at time slot t .

$$(s_{i,j,t} \otimes l_{i,j,q}) \odot (s_{x,y,t} \otimes l_{y,j,q}) = 0, \quad \forall h_i, h_x, h_y \in N_j, \quad \forall t \in T, \quad 1 \leq q \leq k. \quad (4)$$

Fig. 4 shows an example of the *interference constraint* when c_{ij} and $c_{x,y}$ are scheduled at any given slot t .

The challenge addressed herein entails developing efficient transmission scheduling schemes for exploiting the opportunities for parallel transmissions and minimizing total latency. The exploitation of parallel transmissions should satisfy several constraints, including the beam mode, data diversity, and interference constraints. The proposed scheduling mechanisms are operated on the cluster-based [13] network topology where a cluster consists of a head, members and gateways. Let m denote the number of constructed hosts in the cluster. Let $\delta_{i,j}$ and $\text{delay}(i,j)$ denote the set of scheduled transmission intervals allocated for a communication pair $c_{i,j}$ and the latency of $c_{i,j}$, respectively. The proposed scheduling mechanism attempts to allocate the transmission interval $\delta_{i,j}$ to achieve the following goals.

$$\min \cdot \frac{1}{m} \sum_{i=1}^m \left| \bigcup_{j=1}^m \delta_{i,j} \right| \text{ and } \min \cdot \sum_{i=1}^m \sum_{j=1}^m \text{delay}(i,j)$$

The first goal is to minimize the average transmission intervals, which can be achieved by exploiting the opportunities for parallel transmissions. The second goal is to minimize the total latency (i.e., the sum of transmission latency for each communication), which can be achieved by minimizing the interference among the transmissions.

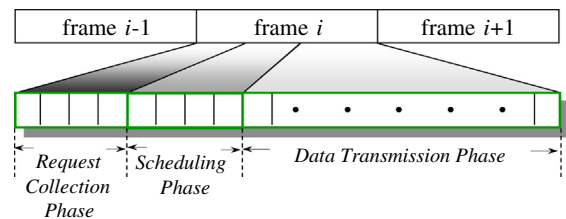


Fig. 5. The duration of each frame in the proposed schematic is comprised of three phases: request collection, scheduling, and data transmission.

4. Scheduling mechanisms for a smart antenna system

Two scheduling mechanisms, referred to as the *Intra-S³* and *S³* schemes, are proposed for the smart antenna system. The *Intra-S³* scheme is proposed to schedule the communication pairs for each cluster, and the *S³* scheme extends the *Intra-S³* mechanism to the inter-cluster environment. In a cluster, the head is responsible for collecting the transmission requirements of the members and for arranging the transmission slots for the requirements. The proposed *Intra-S³* scheme partitions the duration of each frame into three sub-frames, which are used for performing operations such as requesting collection, scheduling, and data transmission as shown in Fig. 5. The *Intra-S³* scheduling algorithm is mainly comprised of these three major tasks and is executed in every frame. The following section details the *Intra-S³* scheduling scheme and illustrates its concepts with examples.

4.1. Request collection phase of *Intra-S³*

In this phase, the head collects the neighboring relationships, i.e., the existing links for the hosts in its cluster and constructs the neighbor matrix L . First, the head polls each neighbor to collect neighbor requests. A polling operation is comprised of at most k messages simultaneously sent from k different beams and prompts hosts located in different beam sectors of the head to transmit their transmission requests at the next time slot. After receiving the polling message, all polled members can simultaneously send their transmission requirements to the head in the next slot. A common transmission rate is used in this phase. The transmission requirement specifies the volume of transmission data and the transmission rate. Upon receiving transmission requests, the head establishes the transmission matrix D , transmission rate matrix R , and null scheduling matrix S ; the scheduling phase can then be initiated.

Figs. 6 and 7 provide examples of the entire scheduling algorithm. In Fig. 6a, host h_1 serves as the head of the cluster.

Each host is assumed to be equipped with a four-beam smart antenna system. In this phase, host h_1 polls each neighbor to collect neighbor requests and creates the neighbor matrix L of the neighboring hosts, as shown in Fig. 6b. The head h_1 arranges the available slots for each member according to the neighbor matrix L , thus allowing them to send transmission requirements to the head without collision. Fig. 7a shows the potential communication pairs, and Fig. 7b provides the data volume for each transmission requirement for the network, as shown in Fig. 6a.

4.2. Parallel transmission construction for scheduling

Parallel transmission construction is designed to exploit the opportunities for parallel transmissions. This phase mainly consists of two tasks: *Host Parallel set Construction (HPC)* and *Cluster Parallel sets Construction (CPC)*. In the HPC, the cluster head considers one member host h_i at a time and constructs several *Host-centric Parallel Sets (HPS)* for h_i ; thus, all parallel communications in the various beams of host h_i can be arranged. The cluster head then applies the CPC to construct the *Cluster-centric Parallel Sets (CPS)*. Each CPS can be regarded as a maximal merge set of HPSs from different hosts and allows several hosts to execute their communications simultaneously through their multi-beam smart antenna systems.

The HPC executed by the cluster head attempts to exploit the opportunities for parallel transmissions from multiple beams under the *beam mode constraint* and *data diversity constraint*. Let $b_{i,k}$ be the transmission requirement set of host h_i in beam k :

$$b_{i,k} = \{c_{ij} | d_{ij} > 0 \text{ and } l_{ij,k} = 1\}.$$

In the HPC, the head h_{Head} selects a communication pair from each $b_{i,k}$ for host h_i to construct an HPS if $b_{i,k}$ is non-empty; thus, the HPC may produce several non-empty HPSs for each host h_i , which are referred to as $g_{i,1}, g_{i,2}, \dots$, and $g_{i,w}$. Fig. 8 presents the process of HPC for the example provided in Fig. 7a. All transmissions associated with the

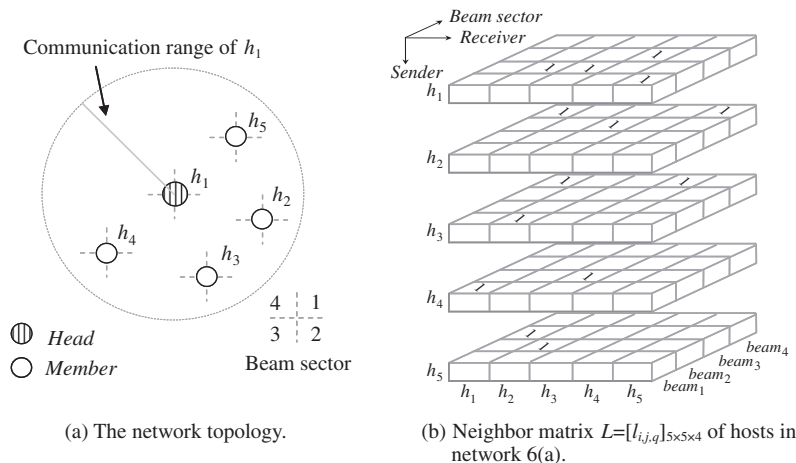


Fig. 6. Example of network topology and neighbor matrix of the network.

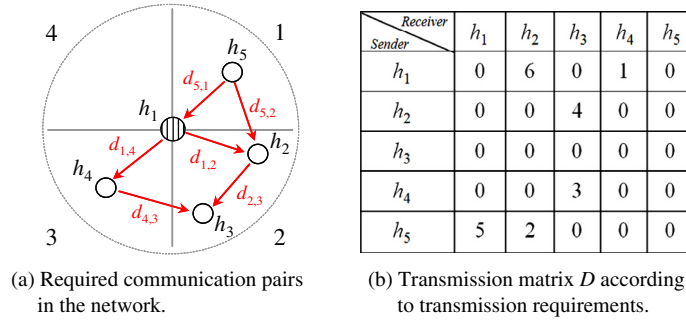


Fig. 7. Transmission requirements collected in the Head are represented by transmission matrix D .

pairs collected in an HPS can be simultaneously transmitted through different beams of h_i ; h_{Head} then ranks the HPSs of all hosts in ascending order regarding the element number of each HPS and renames each HPS by parallel group g'_x , where $1 \leq x \leq m$ and $|g'_1| \geq |g'_2| \geq \dots \geq |g'_m|$. Therefore, all parallel groups are denoted by $P_T(h_{Head})$ and can be obtained through the union of parallel sets of different hosts:

$$P_T(h_{Head}) = \cup_{i=1}^m g'_i.$$

The CPC merges the parallel groups of $P_T(h_{Head})$ to generate CPSs, which exploit more opportunities for simultaneous transmissions in the multiple beams of multiple hosts. The CPC should consider the communication constraints when constructing the CPSs according to the neighbor matrix L and transmission demand matrix D . The constraint checking function $f_{PT}(c_{ij}, c_{xy})$ is defined as follows:

$$f_{PT}(c_{ij}, c_{xy}) = \begin{cases} \text{true,} & \text{if (1)–(3) hold,} \\ \text{false,} & \text{otherwise.} \end{cases}$$

If function $f_{PT}(c_{ij}, c_{xy})$ returns false, the communication pairs c_{ij} and c_{xy} cannot be scheduled in the same time slot, which avoids the violations of communication constraints. The communication pair c_{ij} can be collected into an existing CPS if $f_{PT}(c_{ij}, c_{xy})$ is true for each communication pair c_{xy} in the CPS. When the pair c_{ij} is collected into the CPS, it should be deleted from the original HPS. Let ζ be the new set of $P_T(h_{Head})$ obtained by applying HPC; h_{Head} then ranks the parallel groups in ascending order according to the element number of each group and renames each group CPS_x , where $1 \leq x \leq |\zeta|$ and $|CPS_1| \geq |CPS_2| \geq \dots \geq |CPS_{|\zeta|}|$. The computational complexity of the HP algorithm is $O(N \times u)$, where N and u denote the number of neighboring hosts and beams, respectively. Let ζ be the number of parallel groups of $P_T(h_{Head})$. The computational complexity of the CP algorithm is $O(\zeta^2)$.

Consider the example of Fig. 6a. The transmission requirement sets of all hosts are $b_{1,2} = \{c_{1,2}\}$, $b_{1,3} = \{c_{1,4}\}$, $b_{2,3} = \{c_{2,3}\}$, $b_{4,2} = \{c_{4,3}\}$, $b_{5,2} = \{c_{5,2}\}$, and $b_{5,3} = \{c_{5,2}\}$, and others are empty sets. Fig. 7 shows the HPC and CPC procedures applied to this example. In Fig. 8, the non-empty HPSs are $g_{1,1} = \{c_{1,2}, c_{1,4}\}$, $g_{2,1} = \{c_{2,3}\}$, $g_{4,1} = \{c_{4,3}\}$, and $g_{5,1} = \{c_{5,2}, c_{5,1}\}$, which are constructed by applying the HPC. The $P_T(h_{Head})$ set can then be obtained from the union of g'_1 , g'_2 , g'_3 , and g'_4 , where $g'_1 = \{c_{1,2}, c_{1,4}\}$, $g'_2 = \{c_{5,2}, c_{5,1}\}$, $g'_3 = \{c_{2,3}\}$, and $g'_4 = \{c_{4,3}\}$. Following the application of the CPC, the sets in the $P_T(h_{Head})$ are grouped into several

CPSs including $CPS_1 = \{c_{5,2}, c_{5,1}, c_{4,3}\}$, $CPS_2 = \{c_{1,2}, c_{1,4}\}$, and $CPS_3 = \{c_{2,3}\}$.

Algorithm. Parallel Transmission Constructing

Algorithm : Parallel Transmission Constructing^a

Input : Neighboring Matrix L , Transmission Demand Matrix D ^a

Output : $P_T(h_{Head})$ ^a

HP Let the number of beam sectors be u and the transmission requirement sets be
 $b_{i,k} = \{c_{ij} \mid d_{ij} = 1 \text{ and } l_{ij,k} = 1\}$
 $P_T(h_{Head}) = \zeta = \{\}$;
 For each host h_i {
 $t = 1$;
 Do {
 $g_{i,t} = \{\}$;
 For $j = i+1$ to u
 If $|b_{i,j}| \neq 0$ {
 select c from $b_{i,j}$; $g_{i,t} = g_{i,t} \cup \{c\}$; }
 If $(|g_{i,t}| \neq 0)$ { $\zeta = \zeta \cup \{g_{i,t}\}$; $t++$; }
 } until $g_{i,t} = \{\}$;
 }
 Rename the elements of ζ such that
 $|g'_1| \geq |g'_2| \geq \dots \geq |g'_{|\zeta|}|$;

CP For $i = 1$ to $|\zeta| - 1$
 For $j = i+1$ to $|\zeta|$
 For each $c_{a,b} \in g'_j$
 If $f_{PT}(c_{a,b}, c_{xy}) = \text{true}$, for all $c_{xy} \in g'_i$ {
 $g'_i = g'_i \cup \{c_{a,b}\}$; $g'_j = g'_j - \{c_{a,b}\}$; }
 $\zeta = \zeta - g'_i$, if $|g'_i| = 0$;
 Rename the elements of ζ such that
 $|g'_1| \geq |g'_2| \geq \dots \geq |g'_{|\zeta|}|$;
 Return $P_T(h_{Head})$;

4.3. Scheduling phase of Intra-S³

This phase treats the CPSs which are produced in the PTC sub-phase as its inputs. Two scheduling schemes, Max Parallel Transmission (MaxPTran) and Delay Aware Scheduling (DAS) are proposed to maximize the number of parallel transmissions and minimize the average latency, respectively.

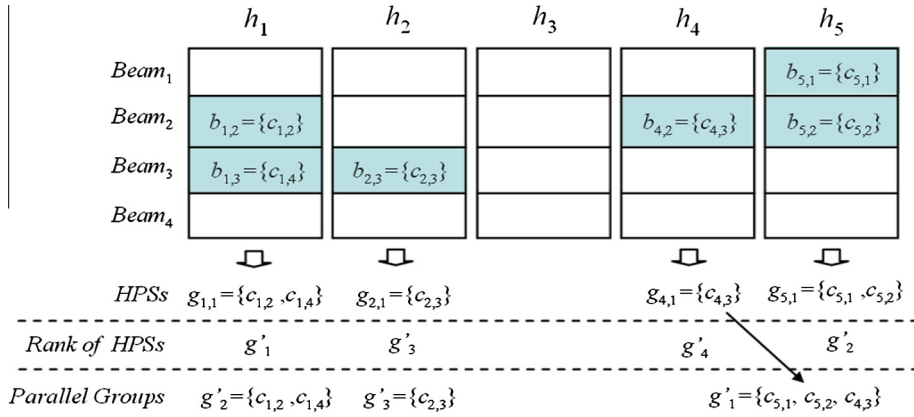


Fig. 8. Process of HPC using the example in Fig. 6a.

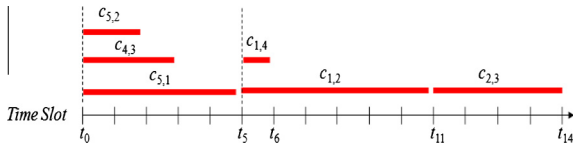


Fig. 9. Scheduling results of the example shown in Fig. 6a through the application of the MaxPTran algorithm.

4.3.1. Max parallel transmission scheduling (MaxPTran)

The MaxPTran algorithm schedules the transmission pairs of the $P_T(h_{Head})$ set in a slot-based manner. The cluster head initially resets the initiation time (i.e., t_s). Let r and $d_{max}(i)$ be the transmission rate and maximal volume of transmission data for the communication pairs in parallel group g'_i , respectively. The number of time slots, referred to as $Slots(i)$, is assigned for each parallel group g'_i , where

$$Slots(i) = \left\lceil \frac{d_{max}(i)}{r} \right\rceil.$$

Therefore, the transmission interval $[t_s, t_e]$ is reserved for the parallel group g'_i , where t_e is equal to $t_s + Slots(i)$, and the interval

$$\delta_{a,b} = \left[t_s, t_s + \left\lceil \frac{d_{a,b}}{r} \right\rceil \right]_{a,b}$$

is set for each communication pair $c_{a,b} \in g'_i$, where entry $d_{a,b}$ denotes the data units of transmission demand associated with $c_{a,b}$. The scheduling matrix S can be filled according to the slots assigned for each communication pair of group g'_i . The start time t_s is then set at t_e for the following parallel groups. When the time slot assignments of all parallel groups are finished, the MaxPTran algorithm completes, and the header broadcasts the scheduling matrix S to the members of the cluster. Let q be the number of the transmissions required in the cluster. The computational complexity of the MaxPTran algorithm is $O(q)$.

Consider Fig. 6a as an example. The set $P_T(h_{Head})$ contains $g'_1 = \{c_{5,2}, c_{5,1}, c_{4,3}\}$, $g'_2 = \{c_{1,2}, c_{1,4}\}$, and $g'_3 = \{c_{2,3}\}$. The time slot assignment is applied to each parallel group

in the following sequence: $g'_1 > g'_2 > g'_3$. Assume that the transmission rate is 1 data unit/slot. The values of $Slots(1)$, $Slots(2)$, and $Slots(3)$ are equal to 5, 6, and 3, respectively. The entries $\delta_{5,1} = [0,4]_{5,1}$, $\delta_{5,2} = [0,1]_{5,2}$, and $\delta_{4,3} = [0,2]_{4,3}$ of scheduling matrix S are then filled by the head for the communication pairs $c_{5,1}$, $c_{5,2}$, and $c_{4,3}$, respectively. Following the same process applied for group g'_1 , the entries $\delta_{1,4} = [5,5]_{1,4}$ and $\delta_{1,2} = [5,10]_{1,2}$ of the scheduling matrix S are filled to represent the schedules of the communication pairs $c_{1,4}$ and $c_{1,2}$, respectively. Finally, the entry $\delta_{2,3} = [11,14]_{2,3}$ of the scheduling matrix S is filled according to the assigned time slots of communication pair $c_{2,3}$ belonging to group g'_3 . Fig. 9 shows the scheduling results of applying the MaxPTran algorithm onto the example shown in Fig. 7a. The total transmission slots and latency are 15 and 21, respectively.

Algorithm. MAXimal Parallel TRANsmissions

Algorithm : MAXimal Parallel TRANsmissions

Input : $P_T(h_{Head})$

Output : Transmission Schedule Matrix S

Step 1.

$t_s = 0; t_e = 0;$

Step 2.

For $i=1$ to $|P_T(h_{Head})|$ {

 Compute $Slots(i)$ for the group $g'_i \in P_T(h_{Head})$;

$t_e = t_s + Slots(i)$;

 For each $c_{a,b} \in g'_i$ {

$\delta_{a,b} = \left[t_s, t_s + \left\lceil \frac{d_{a,b}}{r} \right\rceil \right]_{a,b};$

$t_s = t_e; \}$

Step 3.

Broadcast the scheduling Matrix S ;

4.3.2. Delay aware scheduling (DAS)

DAS is a delay-sensitive scheduling algorithm that assigns a higher transmission priority to the parallel group with lower numbers of time slots. DAS attempts to reduce

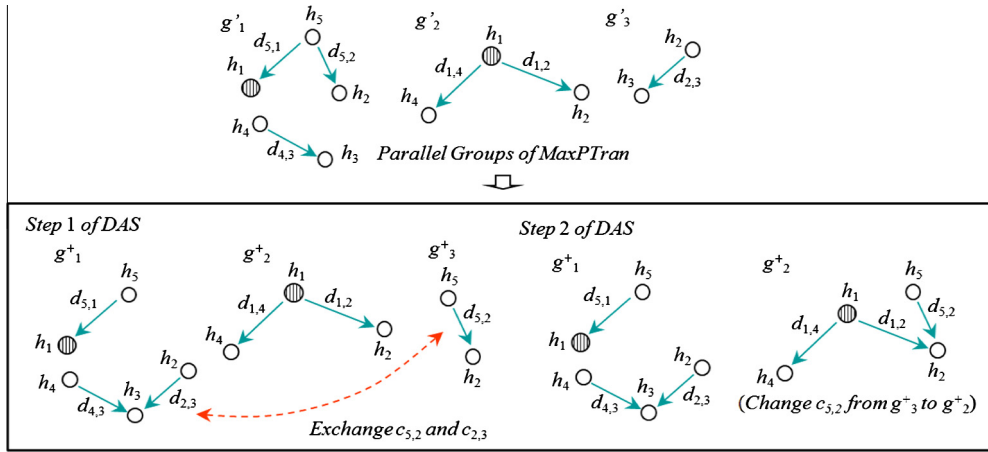


Fig. 10. Steps of DAS for illustrating the changes of parallel groups in Fig. 6a.

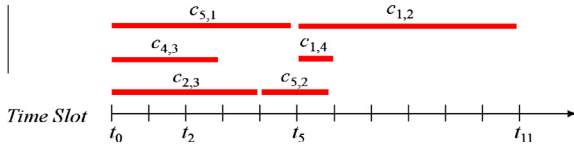


Fig. 11. Scheduling results of the example shown in Fig. 6a through the application of the DAS algorithm.

the total transmission time and average latency. The basic concept of DAS is to adjust the communication pairs among the parallel groups obtained from MaxPTTran to reduce the waiting time of each group. The variance function $\varphi(g'_i)$ is used to evaluate the variance transmission duration of the communication pairs in each parallel group g'_i , where

$$\varphi(g'_i) = \sum_{c_{a,b} \in g'_i} \left(\frac{1}{|g'_i|} \sum \left[\frac{d_{a,b}}{r} \right] - \left[\frac{d_{a,b}}{r} \right] \right)^2.$$

Let g_i^+ be the changed g'_i after executing the DAS algorithm. The DAS algorithm intends to adjust the orders of the communication pairs among the parallel groups to further minimize the value of $\Sigma\varphi(g'_i) - \Sigma\varphi(g_i^+)$. Let q be the number of transmissions required in the cluster. The computational complexity of the DAS algorithm is $O(q^2)$.

Recall that c_{ij} and $c_{x,y}$ can be scheduled in the same time slot if function $f_{PT}(c_{ij}, c_{x,y})$ returns true. Step 1 in the DAS algorithm aims to exchange the communication pairs $c_{a,b} \in g'_i$ and $c_{x,y} \in g'_j$ if the following two conditions are satisfied: 1) the value of function $f_{PT}(c_{p,q}, c_{x,y})$ returns true for each $c_{p,q} \in g'_i - \{c_{a,b}\}$, and the value of function $f_{PT}(c_{p,q}, c_{a,b})$ returns true for each $c_{p,q} \in g'_j - \{c_{x,y}\}$, and 2) the value of $\Sigma\varphi(g'_i - \{c_{a,b}\} + \{c_{x,y}\}) + \Sigma\varphi(g'_j - \{c_{x,y}\} + \{c_{a,b}\})$ should be lower than $\Sigma\varphi(g'_i) + \Sigma\varphi(g'_j)$. Step 2 in the DAS algorithm aims to change the communication pair $c_{x,y}$ into the group g'_i if $c_{x,y}$ satisfies the condition that the value of

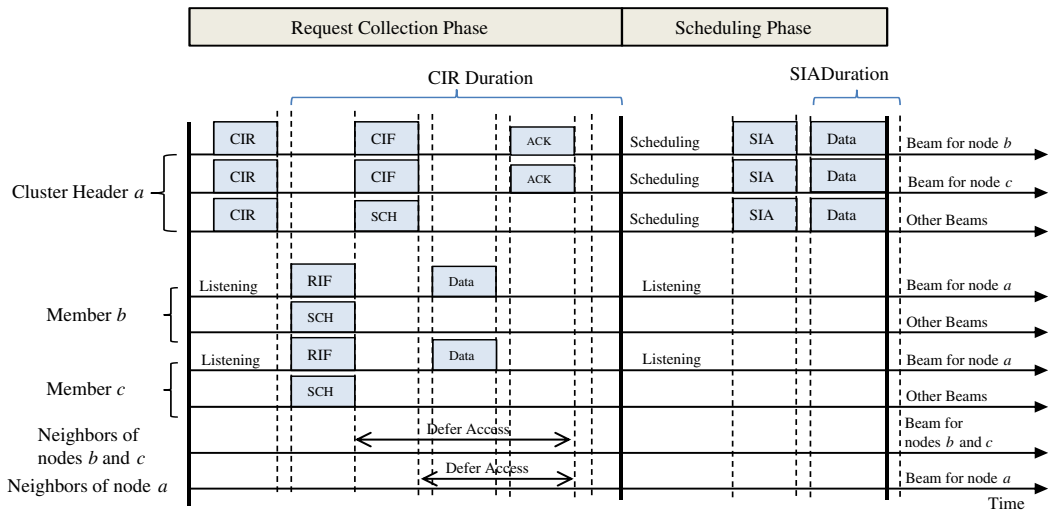


Fig. 12. Example illustrating the basic concepts of the extended ESIF MAC in the proposed Intra-S³ scheme.

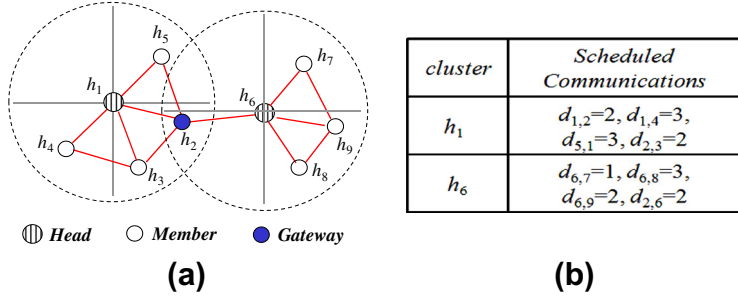


Fig. 13. (a) Example of network topology for inter-cluster scheduling with smart antenna system. (b) Scheduled communication pairs of the example shown in (a).

function $f_{PT}(c_{a,b}, c_{x,y})$ is true for each $c_{a,b} \in g'_i$. The waiting cost function $W(g'_i)$ is defined as $W(g'_i) = \text{Slots}(i)$.

Algorithm. Delay Aware Scheduling (DAS)

Algorithm : Delay Aware Scheduling (DAS)

Input : $P_T(h_{Head})$

Output : Transmission Schedule Matrix S

Step 1.

For any two g'_i and g'_j {
 $c_{a,b} \in g'_i, c_{x,y} \in g'_j$;
 If $f_{PT}(c_{x,y}, c_{p,q})$, where $c_{p,q} \in g'_i - \{c_{a,b}\}$
 If $f_{PT}(c_{a,b}, c_{p',q'})$, where $c_{p',q'} \in g'_j - \{c_{x,y}\}$ {
 $U = g'_i - \{c_{a,b}\} + \{c_{x,y}\}$;
 $V = g'_j - \{c_{x,y}\} + \{c_{a,b}\}$;
 If $\phi(U) + \phi(V) < \phi(g'_i) + \phi(g'_j)$ {
 $g'_i = U; g'_j = V;$ } } }

Step 2.

For any two g'_i and g'_j {
 If $f_{PT}(c_{x,y}, c_{a,b})$, where $c_{a,b} \in g'_j$, for all $c_{x,y} \in g'_i$ {
 $g'_i = g'_i + \{c_{a,b}\}$;
 $g'_j = g'_j - \{c_{a,b}\}$ } }

Step 3.

$t_s = 0; t_e = 0$;
 For each g'_i , following the value of $W(g'_i)$ increasing order {
 Compute $\text{Slots}(i)$ for the group g'_i ;
 $t_e = t_s + \text{Slots}(i)$;
 For each $c_{a,b} \in g'_i$ {
 $\delta_{a,b} = [t_s, t_s + \left\lceil \frac{d_{a,b}}{r} \right\rceil]_{a,b}; t_s = t_e; \}$ }

Step 4.

For each g'_i , following the value of $W(g'_i)$ increasing order {
 For each $c_{a,b} \in g'_i$ {
 $\delta_{a,b} = [t_s, t_e]_{a,b}$;
 while $[t_s - 1, t_e - 1]_{a,b}$ can not interfere with $c \in \cup g'_j$
 $\delta_{a,b} = [t_s --, t_e --]_{a,b}; \}$ }

Step 5.

Broadcast the scheduling Matrix S

Consider Fig. 7a for example. The set $P_T(h_{Head})$ returned from the PTC procedure contains $g'_1 = \{c_{5,2}, c_{5,1}, c_{4,3}\}$, $g'_2 = \{c_{1,2}, c_{1,4}\}$, and $g'_3 = \{c_{2,3}\}$. Assume that the transmission rate is 1 data unit/slot. First, the value of $\Sigma\phi(g'_i)$ is determined by using

$$\begin{aligned} \Sigma\phi(g'_1) &= (1.67^2 + 1.33^2 + 0.33^2) + 2 \times (2.5^2) + 0 \\ &= 17.17. \end{aligned}$$

The communication pair $c_{5,2}$ belonging to g'_1 can be exchanged with $c_{2,3}$, which belongs to g'_3 because $\Sigma\phi(g'_1) < \Sigma\phi(g'_i)$, where

$$\Sigma\phi(g'_1) = (1^2 + 0^2 + 1^2) + 2 \times (2.5^2) + 0 = 14.5.$$

Thereafter, $c_{5,2}$ can be changed from g'_1 to g'_2 . The changed set $P_T(h_{Head})$ contains $g'_1 = \{c_{2,3}, c_{5,1}, c_{4,3}\}$ and $g'_2 = \{c_{1,2}, c_{1,4}, c_{5,2}\}$, and the group g'_3 is discarded because it is empty. The parallel groups of the DAS are determined by modifying the results obtained through the application of MaxP-Tran, as shown in Fig. 10. In Fig. 10, the parallel groups obtained in MaxP-Tran are improved by exchanging $c_{5,2} \in g'_1$ and $c_{2,3} \in g'_3$ for minimizing the value of $\Sigma\phi(g'_i) - \Sigma\phi(g'_i)$ in Step 1 of DAS. Thereafter, the results can be improved further by merging $c_{5,2} \in g'_3$ into g'_2 to increase the sizes of the parallel pairs of g'_2 .

The values of $\text{Slots}(1)$ and $\text{Slots}(2)$ are 6 and 5, respectively. Because the parallel group with a lower number of slots is scheduled with a higher transmission priority, the time slot assignment is applied to the parallel groups in the following sequence: $g'_1 > g'_2$, where $W(g'_1)$ and $W(g'_2)$ are 5 and 6, respectively. Therefore, the entries $\delta_{5,1} = [0, 4]_{5,1}$, $\delta_{2,3} = [0, 3]_{2,3}$, and $\delta_{4,3} = [0, 2]_{4,3}$ of scheduling matrix S are filled by the head according to the slot assignments of $c_{5,1}$, $c_{2,3}$ and $c_{4,3}$, respectively. Thereafter, the entries $\delta_{1,2} = [5, 10]_{1,2}$, $\delta_{1,4} = [5, 5]_{1,4}$, and $\delta_{5,2} = [5, 6]_{5,2}$ of scheduling matrix S are filled for pairs $c_{1,2}$, $c_{1,4}$, and $c_{5,2}$, respectively. In particular, pairs $c_{5,1}$ and $c_{4,3}$ belong to group g'_1 and interfere with $c_{2,3}$. However, they can work in parallel with $c_{5,2}$; thus, the entries $\delta_{5,2}$ can be changed to $\delta_{5,2} = [4, 5]_{5,2}$ for reducing the waiting duration. The scheduling results of the DAS algorithm are shown in Fig. 11. The total transmission slots and latency are 11 and 14, respectively.

4.4. Synchronization of Intra- S^3

ESIF MAC [7] handles synchronization by modifying the IEEE 802.11 RTS/CTR messages to RIF/CIF messages with intelligent feedback. Based on the synchronization approach proposed in [7], this study further defines several control packets such as the Collecting Information Request

(CIR) and Scheduling Information Announce (SIA) to satisfy the synchronization requirements of the *Intra-S³*. The CIR and SIA packets further assist the cluster header to gather scheduling information from its neighbors and announce the scheduling results to them, respectively. Each node is aware of the number of the neighboring nodes by receiving RIF and CIF messages, and listening to the SCH messages from the neighbors. An introduction of the synchronization of *Intra-S³* is shown in Fig. 12.

During the request collection phase, the cluster header broadcasts the CIR messages in all beams to initiate and synchronize the operations in the phase. The appropriate duration included in the CIR messages is used to terminate the phase. On receiving the CIR message, each neighboring node sends the information including the neighboring relationships in the beams and transmission requests to the cluster header using the RIF/CIF/SCH messages before the duration ends. The cluster header can determine the neighbor matrix and transmission demand matrix for the next phase according to the information received from neighboring nodes. The Scheduling Phase initiates after the Request Collection phase is complete. The cluster header executes the *Intra-S³* algorithm to schedule the transmission interval allocated for the communication pairs, minimizing the average latency and maximizing the network throughput. Thereafter, the neighboring nodes wait for the scheduling results announced by the cluster head.

The cluster header announces the scheduling results by broadcasting SIA messages in all beams. On receiving the scheduling results, the neighboring nodes follow the scheduled timing to initiate the transmissions. The N field of the SIA messages shows the repeat times to apply the scheduling, meaning that each scheduling result of the *Intra-S³* can be applied for a synchronizing transmission period; thus, the control overhead of *Intra-S³* is cost efficient. Fig. 12 shows an example of the basic concepts of the extended ESIF MAC for the proposed *Intra-S³* scheme. In Fig. 12, the cluster header a broadcasts CIR messages to collect the transmission requests of its member nodes b and c , and SIA messages to announce the scheduling results.

5. The S³ scheme

The proposed S³ scheme extends *Intra-S³* to permit the consideration of inter-cluster scheduling. The applied inter-cluster scheduling operates on a cluster-based [9] network topology where several stations referred to as gateways participate in more than one cluster.

There are two challenges when extending the *Intra-S³* algorithm from an intra-cluster to inter-cluster scenario. First, according to the *beam mode constraint*, a gateway cannot be scheduled to send a message from one cluster and receive a message from another cluster at any given slot t . Second, determining how to minimize the average latency is critical, especially for cross-cluster routing.

The proposed inter-cluster scheduling applies an adaptive round-robin method to satisfy the *beam mode constraint* and minimize latency. In the proposed scheduling

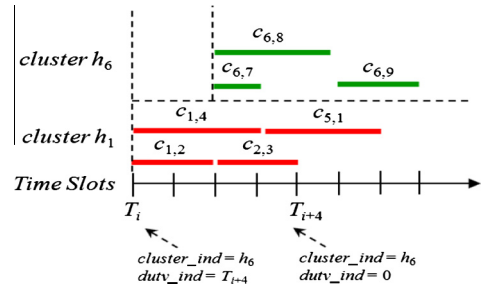


Fig. 14. Time slot arrangement that considers the inter-cluster scenario shown in Fig. 13.

Table 2

Abbreviations of Compared Mechanisms.

ROMA	Receiver-oriented multiple access of [6]
ESIF	Explicit synchronization via intelligent feedback of [7]
<i>Intra-S³</i>	Proposed scheduling algorithm
S3	Proposed scheduling algorithm

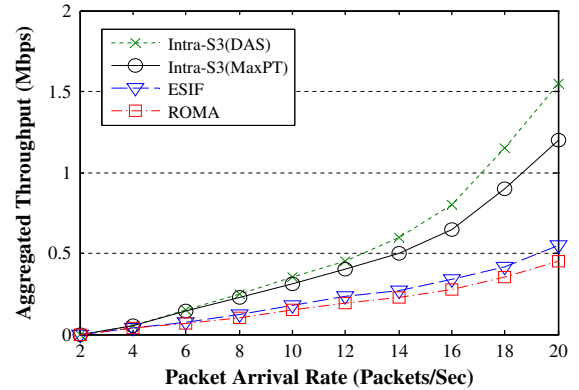


Fig. 15. Comparison of network throughput under various packet arrival rates.

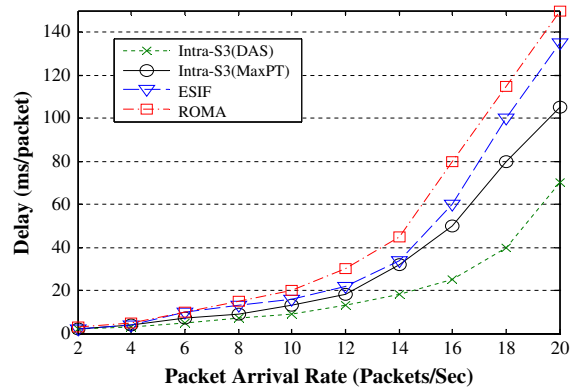


Fig. 16. Impact of packet arrival rate on transmission delay.

process, each gateway maintains a cluster indicator and duty indicator to perform decisions regarding which cluster it will participate in and what its participation

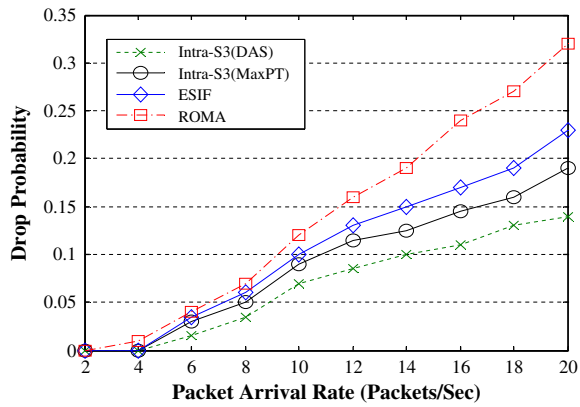


Fig. 17. Comparison of drop probability under various packet arrival rates.

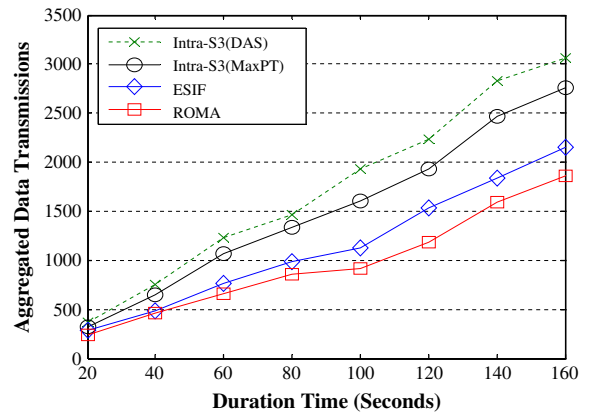


Fig. 20. Comparison of aggregated data transmissions of proposed schemes and ROMA and ESIF in the duration time.

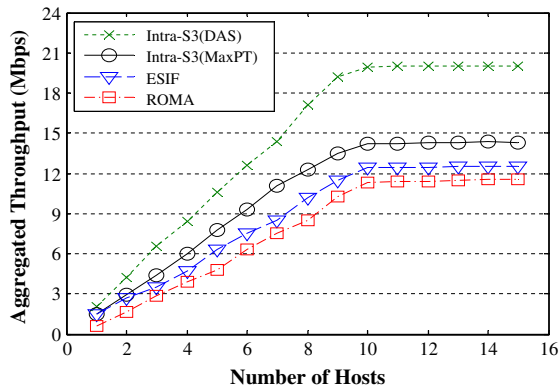


Fig. 18. Impact of number of hosts on the aggregated network throughput.

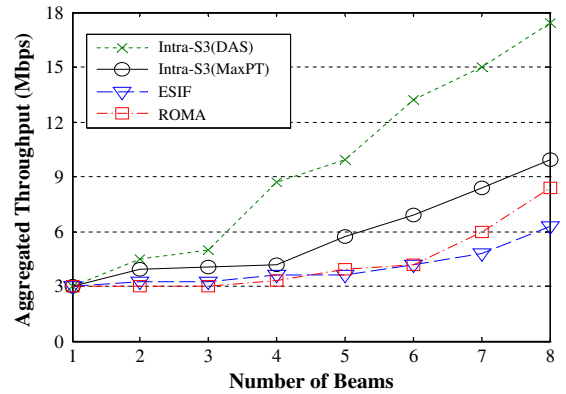


Fig. 21. Comparison of network throughput by varying the number of beams.

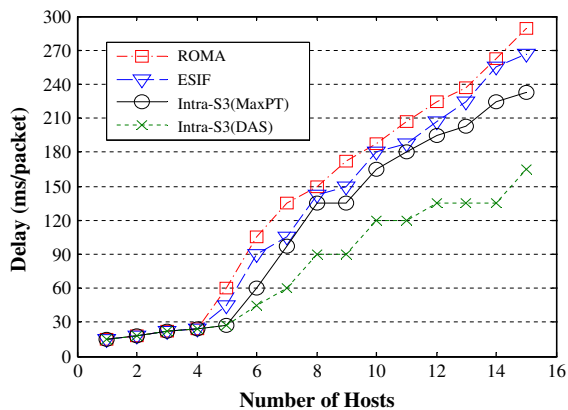


Fig. 19. Comparison of transmission delay of proposed schemes and ROMA and ESIF under various network densities.

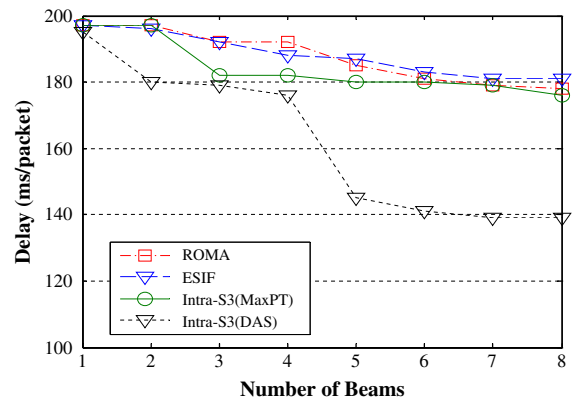


Fig. 22. Comparison of transmission delay by varying the number of beams.

duration is. The gateway participates in exactly one cluster at a time. Fig. 13 gives an example to illustrate the inter-cluster scheduling algorithm. In Fig. 13, two clusters share the common gateway h_2 . Let $cluster_ind$ and $duty_ind$ be

the cluster indicator and duty indicator maintained by station h_2 , respectively. Assume that gateway h_2 currently participates in cluster h_1 and that $cluster_ind$ is set at h_6 . At T_i , the cluster head h_1 announces the time slot arrange-

ment to its members; thus, gateway h_2 knows that the end of the participation time of this cluster is T_{i+4} . Furthermore, the gateway h_2 sets the *duty_ind* to T_{i+4} . If the cluster head h_6 initiates its intra-clustering between T_i and T_{i+4} , then its scheduling does not consider h_2 because h_2 does not participate in the cluster. At T_{i+4} , the gateway h_2 switches from cluster h_1 to cluster h_6 and waits for the next scheduling cycle of cluster h_6 . Fig. 14 shows the resultant scheduling for the two clusters.

6. Simulation study

This section examines the performance improvement of the proposed *Intra-S³* and *S³* schemes against the ROMA [6] and ESIF [7] approaches in terms of aggregated throughput and average packet delay. Each node is equipped with an SB-MBA [13], and the transmitting range of the beam is set at 100 units at a data rate of 11 Mbps. The MAC protocol depicted in Subsection 4.4 is applied to the proposed schemes. Each node generates packets with sizes randomly chosen from 500, 1000, 1500, and 2000 bytes. Similar to IEEE 802.11, the SIFS, AIFS, and DIFS durations are set at 10, 20, and 50 μ s, respectively. Similarly, the packet sizes for the RTS, CTS, and ACK are 44, 38, and 38 bytes, respectively. The packet sizes of the RIF/CIF/SCH/CIR/SIA are set at 21 bytes [7]. The generation of packets at the source node is modeled as a Poisson process with Rayleigh fading channels. Furthermore, each node has a buffer that can hold a maximum of 30 packets. Each performance metric is obtained by averaging the respective metrics of a 1000 experiments. The 95% confidence interval is always less than 5% of the reported values. Table 2 shows the abbreviations of each compared mechanism. The network simulator ns-2 is used to perform all simulations.

6.1. Simple topology

To facilitate the understanding of the benefits in the proposed schemes compared to those of related studies, the simple network topology shown in Fig. 13a and b is first considered. Each node is equipped with the four-beam smart antennas. Figs. 15 and 16 examine the performance of each scheme in terms of network throughput and end-to-end delay when the packet arrival rate ranges from 2 to 20 packets/s.

Intra-S³ and ESIF generally outperform ROMA in terms of network throughput and delay. Although ROMA improves the scheduling efficiency by considering the two-hop neighboring hosts, the overhead for collecting the communication requests leads to lower throughput compared to the other two mechanisms. Moreover, the hash-based policy for deciding the link priority can be further improved by applying a more efficient policy. In ESIF, each node determines its own transmission schedule based on the transmission demands of neighboring nodes. However, the performance of parallelism highly depends on the grouping scheme which aims to construct a set of parallel transmissions into a group. By comparison, the proposed *Intra-S³* determines the schedule by considering all requested data transmissions in a cluster. Therefore,

Intra-S³ exploits more opportunities for parallel transmissions than ESIF does, thus resulting in a higher network throughput, as shown in Fig. 15. Regarding Fig. 15, *Intra-S³* with DAS has a higher throughput than *Intra-S³* with *MaxPTran* because it adjusts the transmission schedule among the parallel groups to further reduce the total number of time slots required for all transmissions. As shown in Fig. 16, *Intra-S³* with DAS achieves a shorter delay than the *Intra-S³* with *MaxPTran* because it further adjusts the transmission schedule among the parallel groups. When the packet arrival rate is larger than six packets/sec, the probability of packet drop is increased with the degree of data congestion. As shown in Fig. 17, the proposed *Intra-S³* demonstrates higher performance than the other two mechanisms do in terms of the drop probability, because the *Intra-S³* achieves superior parallelism, thereby reducing the phenomenon of data congestion.

6.2. Ad hoc topologies

The following simulation environment considers a network containing 200 nodes that is distributed in an area of 500 * 500 square units. The number of beams equipped in each node ranges from 1 to 8.

6.2.1. Performance study of the *Intra-S³* scheme

The performance of an eight-beam smart antenna system is shown in Figs. 18, 19, 20, 21, 23 and 24. Fig. 18 shows a comparison of the aggregated network throughputs of the compared approaches that was executed by varying the number of neighboring hosts ranging from 1 to 16. In general, the network throughputs of the compared approaches increase with the number of neighboring hosts. However, the network throughput remains the same when the number of neighboring hosts is higher than 10. This is because the number of parallel transmissions is bound by the number of beams. By comparison, the proposed *Intra-S³* with DAS and *MaxPTran*, outperforms ROMA and ESIF, because it exploits more spatial reuse opportunities by arranging more simultaneous transmissions (or receptions) of multiple packets in different beams. In particular, *Intra-S³* with DAS has a more favorable throughput compared to that of *MaxPTran*, because it further adjusts the transmission pairs among parallel groups to reduce the total number of required time slots for all transmissions. ROMA has a poor network throughput compared to other schemes because its control overhead increases with the number of neighboring nodes.

Fig. 19 shows the impact of network density on transmission delay. In general, both the neighboring interference and control overhead increase with the number of neighboring hosts. ROMA attempts to collect communication requests from two-hop neighbors; thus, ROMA has a larger delay compared to other schemes. Because *Intra-S³* with either DAS or *MaxPTran* approaches attempt to arrange a maximal number of parallel transmissions in different beams, the waiting time of most transmissions is significantly reduced; thus, the two proposed approaches outperform the existing ROMA and ESIF approaches in transmission delay. *Intra-S³* with DAS further adjusts several transmissions by applying the *MaxPTran* schedule to

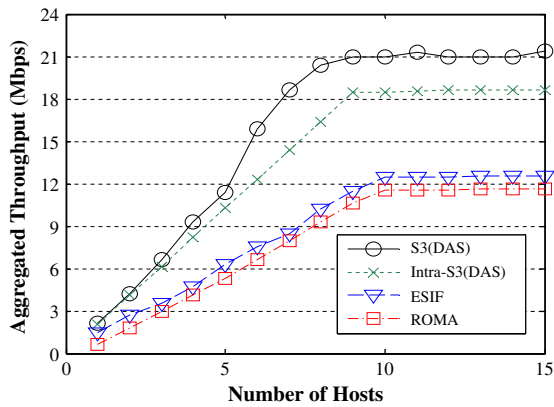


Fig. 23. Comparison of network throughput for proposed $Intra-S^3$, S^3 schematics, ROMA, and ESIF under various network densities.

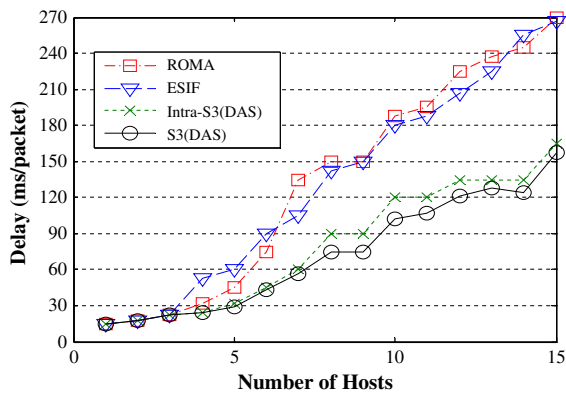


Fig. 24. Comparison of delay time for proposed $Intra-S^3$, S^3 schemes, ROMA, and ESIF under various network densities.

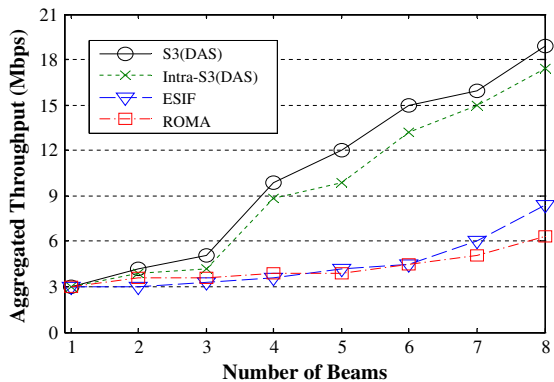


Fig. 25. Comparison of network throughput for $Intra-S^3$, S^3 , ROMA, and ESIF by varying the number of beams.

complete all transmissions of each parallel group earlier. Therefore, $Intra-S^3$ with DAS reduces the delay considerably compared to $Intra-S^3$ with MaxPTran.

Fig. 20 shows the aggregated data transmissions of the proposed schemes as compared with ROMA and ESIF. A

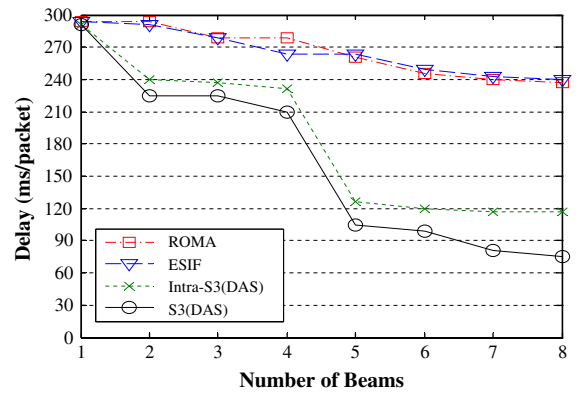


Fig. 26. Comparison of delay time for $Intra-S^3$, S^3 , ROMA, and ESIF by varying the number of beams.

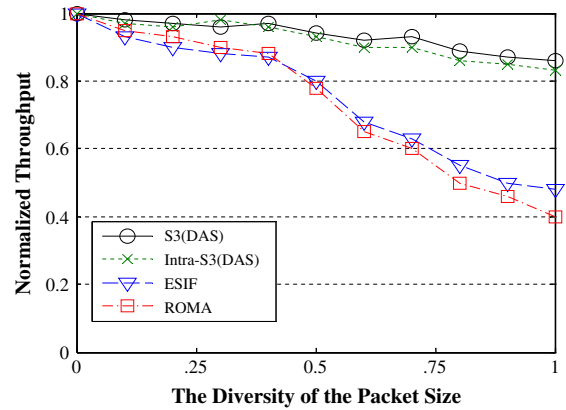


Fig. 27. Comparison of normalized throughput for $Intra-S^3$, S^3 , ROMA, and ESIF by varying the diversity of the packet size.

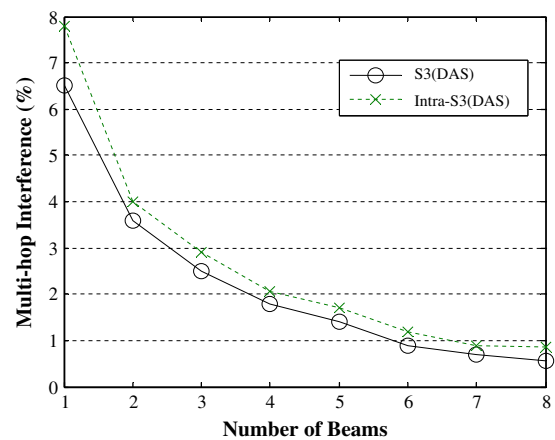


Fig. 28. Comparison of multi-hop interference for $Intra-S^3$ and S^3 by varying the number of beams.

node applying the ESIF MAC can arrange its parallel transmissions in beams and avoid the deafness and hidden terminal problems. Therefore, ESIF has lower control

overhead and shorter transmission delay than ROMA. *Intra-S³* determines the schedule by considering all requested data transmissions in a cluster. Therefore, *Intra-S³* exploits more opportunities for parallel transmissions than ROMA and ESIF do, and as a result. The proposed *Intra-S³* has shorter delay and more aggregated data transmissions, as shown in Fig. 20.

Fig. 21 shows the impact of the number of beams on the network throughput. The four compared approaches have superior network throughputs when the number of beams is increased, because the number of opportunities for spatial reuse increases with the number of beams. *Intra-S³* with *MaxPTran* and *DAS* outperforms ROMA and ESIF because it uses the multiple beams more efficiently. However, *Intra-S³* with *DAS* outperforms *Intra-S³* with *MaxPTran* because it reduces the number of required time slots to complete all transmissions.

Fig. 22 shows the results of an additional examination of the impact of the beam number on transmission delay. *Intra-S³* with *DAS* has the shortest delays compared to other approaches because it reduces the number of required time slots to complete all transmissions. *Intra-S³* with *DAS* has the smallest delay compared to other approaches because it attempts to reduce the transmission delay of each parallel group. *Intra-S³* with *MaxPTran* significantly arranges a maximal number of simultaneous transmissions in different beams, resulting in higher transmission delay performance compared to ROMA and ESIF.

6.2.2. Performance Study of the *S³* Scheme

The *S³* scheme is designed to extend *Intra-S³* to consider inter-cluster scheduling for avoiding a scheduling collision between clusters. The performances of *Intra-S³* and *S³* with the *DAS* are compared with those of ROMA and ESIF. An eight-beam smart antenna system was used in the experiments. Figs. 23–26 show the simulation results, where *S³* (*DAS*) denotes *Intra-S³* with *DAS*.

Fig. 23 shows a comparison of the aggregated network throughputs of the compared schemes that was conducted by varying the number of neighboring hosts ranging from 1 to 16. The *S³* scheme outperforms *Intra-S³* because it avoids the violation of the *beam mode constraint* of gateways. However, the *Intra-S³* approach still reduces the transmission delay considerably, as compared to the ROMA and ESIF approaches, as shown in Fig. 24.

Fig. 25 shows the impact of the number of beams on the network throughput for the proposed *Intra-S³* and *S³* schemes compared to ROMA and ESIF approaches. The proposed *Intra-S³* and *S³* schemes have superior network throughputs when the number of beams is increased because the opportunities for spatial reuse increase with the number of beams. The proposed *S³* approach effectively avoids the violation of the *beam mode constraint* for gateways; thus, it outperforms *Intra-S³* in terms of network throughput. Furthermore, *S³* attempts to arrange the maximal number of parallel transmissions in different beams without collisions in the schedule; thus, the delay time of most transmissions can be significantly reduced. Therefore, *S³* has a shorter transmission delay than the other three approaches, as shown in Fig. 26.

Fig. 27 shows the effect of packet size diversity on performance. In this study, the diversity of packet size is defined as the probability that the size of new generated packet is different to the size of a previously generated packet. The size of a newly generated packet is always different to the size of a previously generated packet in case the diversity of the packet size is set at one. The evaluation results of ESIF and ROMA drop substantially when the diversity approaches one, as shown in Fig. 27. This is because both ESIF and ROMA do not schedule packet transmissions according to packet sizes and available time in each channel, thus leading to significant bandwidth holes. The proposed *S³* and *Intra-S³* apply the proposed scheduling mechanisms and arrange the orders of packet transmissions to reduce the number of bandwidth holes. Therefore, the results of *S³* and *Intra-S³* drop only slightly when the diversity approaches one.

The proposed *S³* and *Intra-S³* apply the neighboring graph to represent the interference relationship for exploiting the opportunities for parallel transmissions. Fig. 28 shows the interference impact of multi-hop neighbors on the network throughput, which is caused by varying the number of beams. The percentages of interference are measured according to the ratio of interference created by multi-hop neighbors to the total interference on the receiver side. In Fig. 28, the multi-hop interference decreases with the number of beams because it can be effectively reduced when increasing spatial reuse opportunities. Conversely, *S³* has less multi-hop interference than *Intra-S³* because it further applies the inter-cluster scheduling mechanism. The difference between the multi-hop interference of *S³* and *Intra-S³* is reduced with the number of beams, as shown in Fig. 28. Thus, the proposed schemes have superior network throughputs and delays, even if they apply the neighboring graph to represent the interference relationship for exploiting the opportunities for parallel transmissions.

7. Conclusion

This paper proposes *Intra-S³* and *S³* approaches for scheduling transmissions using smart antenna systems. The *Intra-S³* scheme with the *MaxPTran* scheduling approach considers the *beam mode*, *data diversity*, and *interference* constraints, and is designed to arrange a maximal number of simultaneous transmissions in different beams for exploiting spatial reuse opportunities. Furthermore, *Intra-S³* with *DAS* further adjusts the transmissions among parallel groups to initiate each parallel group earlier, thereby reducing the number of required time slots for completing all transmissions. The *S³* approach is proposed for use in the inter-cluster environment. The *S³* approach successfully reduces the latency of inter-cluster scheduling. The performance results reveal that the proposed *S³* and *Intra-S³* with *MaxPTran* and *DAS* policies outperform ESIF and ROMA in network throughput and latency.

References

- [1] Y.B. Ko, V. Shankarkumar, N.H. Vaidya, Medium access control protocols using directional antennas in ad hoc networks, IEEE INFOCOM (2000) 13–21.

- [2] M. Takai, J. Martin, A. Ren, R. Bagrodia, Directional virtual carrier sensing for directional antennas in mobile ad hoc networks, *ACM MobiHoc* (2002) 183–193.
- [3] T. Korakis, G. Jakllari, L. Tassiulas, CDR-MAC: a protocol for full exploitation of directional antennas in ad hoc wireless networks, *IEEE Transactions on Mobile Computing* 7 (2) (2008) 145–155.
- [4] G. Jakllari, I. Broustis, T. Korakis, S.V. Krishnamurthy, L. Tassiulas, Handling asymmetry in gain in directional antenna equipped ad hoc networks, *IEEE PIMRC* (2005) 1282–1288.
- [5] H. Gossain, C. Cordeiro, D.P. Agrawal, MDA: an efficient directional MAC scheme for wireless ad hoc networks, *IEEE Globecom* (2005) 3633–3637.
- [6] L. Bao, J.J. Garcia-Luna-Aceves, Receiver-oriented multiple access in ad hoc networks with directional antennas, *Wireless Networks* 11 (2005) 67–79.
- [7] V. Jain, A. Gupta, D. Lal, D.P. Agrawal, A cross layer MAC with explicit synchronization through intelligent feedback for multiple beam antennas, *IEEE Globecom* (2005) 3196–3200.
- [8] R. Ramanathan, On the performance of ad hoc networks with beamforming antennas, *ACM MobiHoc* (2001) 95–105.
- [9] C.C. Shen, C. Srisathapornphat, R. Liu, Z. Huang, C. Jaikao, E.L. Lloyd, CLTC: a cluster-based topology control for ad hoc networks, *IEEE Trans. on Mobile Computing* 3 (2004) 18–32.
- [10] Y. Jing, H. Jafarkhani, Network beamforming using relays with perfect channel information, *IEEE Transactions on Information Theory* 55 (2009) 2499–2517.
- [11] R. Mudumbai, D.R. Brown, U. Madhow, H.V. Poor, Distributed transmit beamforming: challenges and recent progress, *IEEE Communications Magazine* 47 (2009) 102–110.
- [12] A.P. Subramanian, S.R. Das, Addressing deafness and hidden terminal problem in directional antenna based wireless multi-hop networks, *Wireless Networks* 16 (2010) 1557–1567.
- [13] D. Lai, V. Jain, Q.A. Zeng, D.P. Agrawal, Performance evaluation of medium access control for multiple-beam antenna nodes in a wireless LAN, *IEEE Transactions on Parallel and Distributed Systems* 15 (2004) 1117–1129.
- [14] G. Bianchi, D. Messina, L. Scalia, I. Tinnirello, A space-division time-division multiple access scheme for high throughput provisioning in WLANs, *IEEE ICC* 4 (2005) 2728–2733.
- [15] F. Babich, M. Comisso, Theoretical analysis of asynchronous multi-packet reception in 802.11 networks, *IEEE Transactions on Communications* 58 (6) (2010) 1782–1794.
- [16] R.R. Choudhury, X. Yang, R. Ramanathan, N.H. Vaidya, On designing MAC protocols for wireless networks using directional antennas, *IEEE Transactions on Mobile Computing* 5 (5) (2006) 477–491.
- [17] P. Nicopolitidis, On performance improvement of wireless push systems via smart antennas, *IEEE Transactions on Communications* 60 (2012) 312–316.
- [18] H. Liu, S. Gao, T. Loh, Electrically small and low cost smart antenna for wireless communication, *IEEE Transactions on Antennas and Propagation* 60 (3) (2012) 1540–1549.



Chao-Tsun Chang received the Ph.D. degree in Computer Science and Information Engineering from National Central University, Taiwan, in 2006. He is with the Department of Department of Information Management, Hsiuping University of Science & Technology, Taiwan, as an Associate Professor in 2012. In the recent ten years, he has directed eleven researching projects including seven national NSC projects and seven information system developments. He published thirteen SCI indexed journal papers, including *IEEE*

TVT, *IEEE Sensors*, *Computer Networks*, *ACM/Baltzer Journal of Wireless Networks*, and *Journal of Parallel and Distributed Computing*. He is a member of

the IEEE Computer Society and Communication Society. His current research interests include Ad Hoc wireless networks, wireless sensor networks, cooperative radio networks, and mobile computing.



Chih-Yung Chang received the Ph.D. degree in Computer Science and Information Engineering from National Central University, Taiwan, in 1995. He is currently a Full Professor with the Department of CSIE at Tamkang University, Taiwan. He served as an Associate Guest Editor of many SCI indexed Journals, including *International Journal of Ad Hoc and Ubiquitous Computing* (IJAHUC 2011 and 2012), *International Journal of Distributed Sensor Networks* (2012), *IET Communications* (2011), *Telecommunication Systems* (TS, 2010), *Journal of Information Science and Engineering* (JISE, 2008), and *Journal of Internet Technology* (JIT, 2004 and 2008). He was an Area Chair of *IEEE AINA*'2005, *TANET*'2000, *TANET*'2010, Vice Chair of *IEEE WisCom*'2005, *EUC*'2005 and *IEEE ITRE*'2005, *IEEE AINA* 2008, Program Co-Chair of *IEEE MNSA*'2005, *UbiLearn*'2006, *WASN*'2007, *ACM SAMnet*'2008, *IEEE AHUC*'2008, *iCube*'2010, *iCube*'2011, Workshop Co-Chair of *MSEAT*'2003, *MSEAT*'2004, *IEEE INA*'2005, *ICS*'2008, *NCS*'2009, *IEEE VCNA*'2009 and Publication Chair of *MSEAT*'2005 and *SCORM*'2006. He is a member of the IEEE Computer Society and Communication Society. His current research interests include Internet of Things, Wireless Sensor Networks, Ad Hoc Wireless Networks, and WiMAX broadband technologies.



Tzu-Lin Wang received the B.S. and M.S. degrees in computer science and information engineering in 2007 and 2009, respectively, from Aletheia University, Taiwan. She is currently working toward the Ph.D. degree in the Department of Computer Science and Information Engineering at Tamkang University, Taiwan. She won numerous scholarships in Taiwan and has participated in many wireless sensor networking projects. Her current research interests include wireless sensor networks, Ad Hoc wireless networks, mobile/

wireless computing, and WiMAX.



Yun-Jung Lu received the Ph.D. degree from Computer Science and Information Engineering of Tamkang University, Taiwan, in 2012. Currently, he is a software engineer in NVIDIA Corporation.